Measuring drought sensitivity of Spanish cereal yields using thermal remote-sensing indicators

David García-León  ·  Sergio Contreras  ·  Johannes Hunink

February 26, 2018

Abstract This study benefits from ESYRCE, a Spanish field-level dataset on crop yields and surfaces spanning from 2002 to 2015 to estimate the impact of agricultural drought on non-irrigated cereals. We use remote sensing indicators to characterise plant’s drought stress. We also explore the evolution of this relationship over time by estimating the reactivity of cereal crops to different drought stress levels, as measured by an environment index. We find that both the VCI and TCI correlate positively with annual yields, that is, agricultural droughts, usually characterised by high land temperature and vegetation health stress and represented by values of these variables below 0.5, will reduce annual yields in an amount of 2% to 20%, depending on the intensity of that drought spell. We also find that this effect is exacerbated the closer we get to the harvesting phase. By looking at the environment index for each crop, we find that the general stagnation of cereal yields in Spain during the last fifteen years is being accompanied by an increase in the resistance of wheat and barley fields to very adverse drought conditions determined by very high temperatures and extreme canopy stress.

Keywords  droughts  ·  drought sensitivity  ·  crop yields  ·  remote sensing

JEL classification  Q1  ·  Q51  ·  Q54  ·  Q59  ·  R11

David García-León acknowledges financial support from the European Commission (H2020-MSCA-IF-2015) under REA grant agreement no. 705408.

David García-León
Fondazione Eni Enrico Mattei, Climate Change Economic Impacts and Adaptation, Isola di San Giorgio Maggiore n.8, 30124 (Venice, ITALY). Tel.: +39 0412700474. Fax: +39 0412700413. E-mail: david.garcia@feem.it.

Sergio Contreras and Johannes Hunink
Future Water, Calle Azucena, 23. 30205 (Cartagena, SPAIN). Tel.: +34 868071286. E-mail: s.contreras@futurewater.es - j.hunink@futurewater.es.
Contents

1 Introduction .................................................. 2
2 Data .......................................................... 3
  2.1 ESYRCE: Spanish survey on acreage and crop yields ......................... 3
  2.1.1 An overview of the cereal sector in Spain .................................. 3
  2.2 Remote sensing thermal measurement of plant’s drought stress ................. 4
3 Methods ....................................................... 7
  3.1 An empirical model of annual yields and drought indicators .................. 7
  3.2 Measuring drought sensitivity: the environment index ......................... 7
4 Results ....................................................... 9
  4.1 Sign and magnitude of remote sensing measures as yield determinants ........ 9
  4.2 Assessing yield trends and their geography .................................... 11
  4.3 Changes in plant’s drought tolerance through the environment index ........ 12
    4.3.1 Wheat .................................................. 12
    4.3.2 Barley ................................................ 13
5 Discussion ................................................... 14
6 Conclusions ................................................ 15
A Integrating field-level yields with satellite drought data ....................... 18
B Additional tables and figures ................................... 18

1 Introduction

Europe in general and Spain in particular, have experienced drought episodes increasingly over the past decades. These ever more frequent events surely can pose threats to food supply security and compromise the stability of the domestic agri-food market. In order to insure agricultural production against those extreme episodes, accurate and efficient agricultural policies should be proposed and measures to adapt production to changing meteorological conditions must be taken. But prior to that, it is required to carefully assess and quantify the effect of drought scenarios on the productivity of our crops. This remains an open empirical question.

Since time immemorial, cereals have been the most important agricultural output in Spain. Apart from their traditional use in the food industry and being the main source for feeding cattle, a number of alternative uses have been created in the last decades, such as the production of biofuels. Particularly relevant in the case of Spain, for example, barley is very much used as input for the beer brewing industry. Overall, their adaptability to different growing environments and climatic conditions make them very suitable to poor soils, as those found in Spain. But, at the same time, their long vegetative cycle makes them be very exposed to environmental and climatic oscillations. We expect, then, that their final productivity will be partly determined by the effect of the environment. In Spain, a net importer of cereals, there are merely a dozen varieties of cereals being grown nowadays. Of them, this paper will consider the most relevant ones, namely, wheat, barley, maize, oat, and rye. Rice, grown in water-lodge regimes, will be left aside this study. These 5 varieties will account for more than 90% of planted areas and total production.

The use of statistical methods to approach the relationship between yields and meteorological variables has proliferated in recent years. The work by Schlenker and Roberts (2009) is a prominent example of this stream of literature. This is corroborated later by other authors, such as Lobell et al. (2011). Tack et al. (2015) document in a subsequent paper that wheat yields have become less resistant to high spring temperatures in the US mid-west. Geographically closer to our case study, Ceglar et al. (2016) identify weather sensitivity pattern in French wheat yields. Another few studies address explicitly the role of droughts on agricultural productivity. For instance, Nelson et al. (2014), for instance, study a battery of biophysical shocks applied to certain crop models to study climate change impacts on yields. Meanwhile, Mysiak et al. (2013) study the impact of droughts in the Po river basin’ agricultural system derived from aggregate

1 Rice, grown in water-lodge regimes, will be left aside this study.
Measuring drought sensitivity of cereals

precipitation figures. In Spain, Iglesias et al. (2000) develop a crop model for wheat in seven Spanish areas analysing the impacts of climate change in terms of temperature, precipitations and CO$_2$ concentrations while Jenkins (2013) assesses direct and indirect losses derived from drought episodes using input-output tables. It is very hard to find, though, empirical studies addressing specifically the role of droughts in the evolution of crop yields. Among them, we would highlight those developed by Lobell et al. (2014) and Quiring and Papakryiaokou (2003) for the US corn belt and Canadian wheat prairies. We are not aware of any rigorous study of the effects of agricultural drought on crop productivity for the case of Europe.

There are a couple of challenges a researcher normally faces when a comprehensive analysis of the role of agricultural droughts on yields. First, there is generally a lack of availability of accurate field-level data on yield performance that span a range of drought conditions and time. We overcome this problem by the use of the Encuesta sobre Superficies y Rendimientos de los Cultivos (ESYRCE), an annual survey conducted by the Spanish Ministry of Agriculture in which field-level data on surface and yields is collected every year.

Second, there is no consensus on a definition of drought. Droughts can be classified into three categories: meteorological, agricultural, and hydrological droughts. Agricultural droughts are those affecting primarily agricultural fields and there does not exist a synthetic measure that measure its level. The channels through which agricultural droughts manifest are various and we should account for them if we want to quantify their impacts. Finding different ground-based measures of agricultural droughts spatially and temporally relevant is almost impossible. We will resort to the use of remote-sensing measures. In particular, we will use thermal imagery to characterise both the state of vegetation and their temperature, variables directly linked to drought stress. In particular, we will characterise drought intensity levels by the combined action of Vegetation Condition Index (VCI) and the Temperature Condition Index (TCI). Both indicators are considered in Svoboda and Fuchs (2016) as relevant drought indicators.

The remainder of this paper is organised as follows. In Section 2, we describe all the data employed in this study and explain how we overcome their differences in spatial resolution. Section 3 explores the statistical and econometric techniques used in this paper. Section 4 describes our main results. In particular, Section 4.1 unveils the role of our set of drought indicators on cereal yields and Section 4.3 analyses how this relation has evolved over the first years of this century. We discuss our findings and suggest future avenues for research in Section 5. Finally, Section 6 concludes.

2 Data

2.1 ESYRCE: Spanish survey on acreage and crop yields

We use ESYRCE (Encuesta sobre Superficies y Rendimientos de los Cultivos), a detailed field-level survey on agricultural surface and crop yields, spanning from years 2002 to 2015. This survey is conducted annually by the Spanish Ministry of Agriculture and serves as the Spanish input for the European Union’s Farm Accountancy Data Network (FADN)$^2$, which describes in fine detail the economic situation of farmers by different groups throughout the European Union. The dataset goes far beyond the needs of this research, as it covers the entire agricultural production carried out in the Spanish territory.

ESYRCE gathers data systematically every year using a conglomerate stratified sampling method. The Universal Transversal Mercator (UTM10) coordinate system is used to conform blocks of 10 km of side covering the full extension of the country, as it is shown in the left panel of Fig. 1. Each of these blocks is then subdivided in 100 equivalent cells of 1km $\times$ 1km (or 100 hectares). Squared segments of 700 meters side anchored in the southwest extreme of each cell are the basic unit that is sampled (see centre and right panel of Fig. 1).

---

$^2$ http://ec.europa.eu/agriculture/rica/
The sampling strategy of ESYRCE is based on the survey of geographical conglomerates (blocks) that are sampled in a stratified manner: blocks are classified into extensive, intensive or very intensive production areas. Accordingly, cells situated in intensive production areas are visited more often. In particular, only three segments from extensive blocks will be sampled every year whereas nine to fifteen segments can be visited in more intensive areas.

In an attempt to isolate our findings from external, human-induced irrigation sources, of which we do not have data, we study in this paper those varieties traditionally regarded as non-irrigated, namely, cereals. As we will discuss in more detail in the next paragraphs, cereal production in Spain is not entirely based on precipitations as source of water. However, we will use those types, observable in our dataset as counterfactual group to test the accuracy of our findings.

2.1.1 An overview of the cereal sector in Spain

Cereal production in Spain, excluding rice (whose growing method, waterlogged soils, is totally unrelated to standard growing techniques) shows a moderate importance within the whole agricultural sector, with an estimated total production value of 3.3 billion euros in 2015, representing 12.7% of total crop production, the second most important agricultural output after fruit and vegetables (MAGRAMA, 2016). Final products, offered largely by a handful of agricultural cooperatives, are headed mainly to human consumption and animal breeding but also to a powerful beer brewing industry or the production of biofuels. Spain, however, is overall a net importer of cereals, featuring a systematic negative trade balance. In the 2015-16 campaign, 14 million tonnes were imported, 73% of total production, most of which (72%) were dedicated to breed cattle.

Around 6 million hectares of cereals are planted on average in Spain every year. Production is widespread around the whole territory but intensively located at the central plateaus and the banks of the main rivers (Ebro, Duero, and Guadalquivir), as evidenced in Fig. 2. There are about a dozen varieties of cereals grown in Spain every year. In this study we focus our attention on to the five most representative, namely, wheat, barley, maize, oat, and rye. These five classes cover 96% total area and 93% total production of cereals. All of them are basically non-irrigated crops, with the exception of maize. In contrast with some other regions in Europe and the US, maize grown varieties are mostly irrigated in Spain (more than 90% of total fields). This poses us with some difficulties to isolate the effect of droughts on production, as weather and drought variations are accommodated with the use of external irrigation.
Traditionally, wheat has been the predominant cereal variety produced in Spain. However, in the late part of the last century, barley has gained importance, as the demand for malted barley of the beer brewing industry raised. As a result, Spain is today the fifth world largest producer of barley and the country in the EU where a largest area is devoted to barley. The surface dedicated to barley has doubled ever since and represents today 43% of total cereal production area. Meanwhile, wheat stands for 36% of total area, oat accounts for 8%, maize for 6.5%, and rye for 3%.

Cereal fields in Spain are characterised by their high degree of atomisation and small dimension, which impacts productivity negatively (see Table S.1). As a consequence, cereal yields in Spain are relatively low, if compared to other European countries and the US. For example, despite being Spain the country in the EU that dedicates a larger area to barley fields, total output is similar in its order of magnitude to that obtained in France or Germany (López-Querol et al., 2016). As shown in Fig. 3, yields of wheat and barley (tons per hectare) have stayed relatively stable during the first 2000’s. Meanwhile, a moderate positive trend can be observed for maize yields. This (almost) flat trend of yields has been already identified for European cereal production by Lobell and Moore (2015). Maize yields behave slightly different, with a positive significant trend slope. Average maize yields have passed from 10 tons per hectare in the early part of the century to almost 12 tons per hectare.
Fig. 3: Agricultural yields over the study period. Horizontal bars inside boxes represent median yield values, boxes represent 1.5*IQR (interquartile range), whiskers the range between 10th and 90th percentile. A solid line connects yield median values.

2.2 Remote sensing thermal measurement of plant’s drought stress

The use of remote sensing measures in the agronomic literature is not new. For instance, Lobell (2013) applies satellite-based data to estimate and analyse the yield gap. Some examples in which these variables are specifically used to measure drought intensity are the works by Govender et al. (2009) or Maes and Steep (2012). The crucial difference between studies carried out during the 1980s and the current is that the level of spatial resolution measures today is far greater and thus, remote sensing variable are much more precise.

Sepulcre-Cantó et al. (2006) as a reference for encouraging the use of thermal imagery as a way to detect water stress of canopy. Siebert et al. (2014) on the importance of measuring the temperature of canopy. Heft-Neal et al. (2017) support the use of satellite temperature instead of air temperature.

The Normalised Difference Vegetation Index (NDVI), a variable first introduced by Gao (1995). The NDVI measures ground and vegetation reflectance through satellite-based optical spectrometers along the visible and red channels. Healthy and vigorous plants absorb more visible light, and thus look dark in the visible channel while they will appear lighter in the infra-red channel. The NDVI formula looks as follows: VCI measures the Normalised Difference Vegetation Index (NDVI) anomaly as an indicator of drought
stress. NDVI, introduced by Kogan (1995), measures greenness and vigour of vegetation and can identify drought-related stress to vegetation

\[
\text{NDVI} = \frac{\text{NIR} - \text{VIS}}{\text{NIR} + \text{VIS}},
\]

where VIS and NIR stand for the spectral reflectance measurements acquired in the visible (red) and near-infrared regions, respectively. The NDVI takes values between 0 and 1, where values close to 1 would denote absence of vegetation, values in the range 0.3-0.8 would state increasing levels of health and values close to 1 would denote a extremely dense canopy. Different values of the NDVI can be observed in Fig. ??, where red areas indicate NDVI values closer to 0 and thus, denote higher drought stress. We exploit the variation in this variable to account for different drought stress levels during the plant’s growing period.

\[
\text{VCI} = \frac{\text{NDVI} - \text{NDVI}_{\text{min}}}{\text{NDVI}_{\text{max}} - \text{NDVI}_{\text{min}}},
\]

(1)

In particular, we make use again of a satellite provided variable called Land Surface Temperature (LST). Satellite-based temperature measures have been proved to be good substitutes of more traditional ground-based air temperature variables to study climate impacts, as evidenced by Heft-Neal et al. (2017).

\[
\text{TCI} = \frac{\text{LST}_{\text{max}} - \text{LST}}{\text{LST}_{\text{max}} - \text{LST}_{\text{min}}},
\]

(2)

3 Methods

3.1 An empirical model of annual yields and drought indicators

For each crop in our sample, we calibrate a multiple regression model relating annual yields and drought indicators at the cell level. To account for potentially positive trends derived from the better use of agronomic practices and/or pesticides and control for location-specific fixed effects, we will include in our regressions a year time trend and allow for provincial (NUTS3) fixed effects, respectively. We will explore different combinations of VCI-TCI indices to understand their significance over the plant’s growing period. Our empirical model will read

\[
y_{it} = \alpha_i + \beta_j + \delta t + X_{it} \gamma + u_{it},
\]

(3)

where \(y_{it}\) stand for annual yields, \(\alpha\) is a cell-level intercept, each \(\beta\) represent a fixed-effect at the province level (\(j \in J\), provinces), the term \(\delta t\) tries to capture time trends attributed to technology improvements or improvements in farming practices, \(X\) are our set of remote-sensing indicators, and \(u\) is the idiosyncratic error term. Parameters are estimated with OLS.

The above estimated empirical crop model enables us to analyse, in a first stage, the trend of yields per crop, breaking them down by productivity potential, as derived from the estimated yield values (\(\hat{y}_{it}\)).

3.2 Measuring drought sensitivity: the environment index

In order to track the evolution of yields sensitivity to drought stress, we need to construct a proper measure of stress. Used to track yield performance by Tollenaar and Lee (2002) among others, the environment index (EI) describes for a particular field in a particular year the level of environmental stress the crop is experiencing. Comparing the evolution of actual harvested yields with those predicted, for particular environmental conditions, gives us a measure of the degree of sensitivity of plants exposed to different
Our regression models derived in the first stage are used to calculate an EI for each cell and year. In our setup, obtaining the EI boils down to calculate for each cell the predicted yields using our regression model, given outstanding weather and drought conditions and having a fixed level of technology. The last aspect is crucial if the pure effect of environmental stress wants to be isolated. We will use in our estimations a constant level of technology fixed at year 2005. The equation would read as

$$EI_{it} = \hat{y}_{it} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\delta}_t,$$  \hspace{1cm} (4)

Fig. 4: Monthly average VCI in different Spanish geographical areas
where the EI for a particular cell in a certain season is the predicted yield, given the environment generated by the optimal number of predictors in the previous stage and their associated coefficients.

These generated data are clustered into groups of cells featuring similar level of stress, from which yield progress is then analysed. Accordingly, different level of stress will be quantified by splitting the distribution of the EI in quintiles, each of them describing different levels of growing conditions or drought severity states. A subsequent analysis of the time trends of each quintiles will provide us with evidence on the relative evolution of yields under different drought scenarios.

Note, however, that our variable of interest is not the level of actual yields any more, but this variable once technology effects and provincial fixed effects have been removed. We will refer to this artificial variable as *yield departure* and will study its trend and components for each EI quintile.

### 4 Results

#### 4.1 Sign and magnitude of remote sensing measures as yield determinants

As indicated in Section 3.1, we calibrate an empirical specification to each cereal variety, accounting for yearly technology developments and provincial (NUTS3) fixed effects. We study the role of average VCI and TCI values and subsequently discriminate their quarterly effects over the growing period. Results derived from the fit of these empirical model of crop yields applied to our five varieties of cereals can be observed in Table 1.

As for wheat, we find that both, the VCI and TCI have positive, significant effects on cereal yield annual levels. In particular, an additional one-tenth unit of VCI would be associated with 65.4 kilograms of cereal per hectare, which represents 2% more output above the average. In the case of TCI, an additional one-tenth, indicative of milder temperatures, would imply an increase in output of 70 kilograms per hectare above the average. The total variance of yields explained by our set of drought indicators is around 27%, a moderate values, consistent with the literature (Quiring and Papakryiaokou, 2003). Given the extension of the territory covered in this study, there may exist a great amount of local characteristics, such as soil quality, soil moisture, use of fertilisers or additional climatic patterns key to explain yield heterogeneity that are not accounted in this specification.

To put these results into the context of an episode of droughts, let us assume that a plant features a normal vegetation health status when VCI is over 0.5. Now, let us imagine that a particular year is characterised by moderate drought conditions (VCI between 0.3 and 0.5). This decay in the average level of VCI would imply a decrease in output from 2% to 4% below the average values solely attributed to the effect of VCI on yields. In a severe drought environment (VCI from 0.1 to 0.3), the shrinkage in output would fluctuate between 4% to 8%. Meanwhile, output losses will be 10% or more under strict drought conditions (VCI below 0.1). Typically, drought scenarios are characterised also by an increase in the average temperature of soils, that is, a decrease in the value of LST. If this happens, we can witness a joint decrease of VCI and LST, which results in a multiplicative adverse effect on yields. As a matter of fact, we could plausibly attribute losses of nearly 20% to annual wheat yields if drought conditions are extreme during a particular season.

Winter wheat in Spain is typically sown during October and is grown on average during the following 9 months until is harvested in mid-June. Subsequently, in order to learn more about the dynamic effect of droughts on plant’s performance, we split the growing period of the plan into three phases: early leaf, flowering, and maturing phases and, accordingly, we break down the average values of VCI and TCI and provide averages in these three phases. As evidenced in Table 1, we identify that the values of VCI and TCI are increasingly important along the growing cycle, being those belonging to the maturing phase vital in determining the final productivity of plants.
The results for barley parallel in many senses those of wheat. Both point estimates are significant and positive, but slightly larger in absolute and relative terms. One-tenth unit deviations of VCI and TCI are associated with increases of 80 and 84 kilograms per hectare in barley yields, respectively. These figures represent each variations of 3% around the average yield. Again, a moderate drought scenario represented by a combined decrease of one-tenth in VCI and TCI would diminish average yields around 6%. This amount would reach 10% to 12% under severe drought scenarios and could go up to 30% falls in output under extreme drought conditions. The role of these indicators during the last phase of the growing period is particularly relevant in the case of TCI whereas it is more evenly distributed in the case of VCI.

As we mentioned in Section 2, most maize production in Spain is supported with external sources of water. Since it is very difficult to find quality data on crop irrigation of maize crops, we cannot control for this variable in our specification. Still, we can regress maize yields on our set of drought indicators to test whether our results from previous crops are meaningful: since the farmer would be able to accommodate weather variations via external water supply to fields, we would expect almost flat profiles in the response of yields to drought and weather indicators. This is corroborated by looking at the third column of Table 1, where we can spot an erratic behaviour of drought indicators when applied to maize data, a clear sign of their lack of confidence in explaining yields of irrigated varieties.

### Table 1: Regression Results

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Barley</th>
<th>Maize</th>
<th>Oat</th>
<th>Rye</th>
</tr>
</thead>
<tbody>
<tr>
<td>VCI$_{\text{avg}}$</td>
<td>0.654***</td>
<td>0.803***</td>
<td>−0.722</td>
<td>(0.052)</td>
<td>(0.042)</td>
</tr>
<tr>
<td>VCI$_{\text{Q1}}$</td>
<td>0.036</td>
<td>0.083*</td>
<td>−0.803**</td>
<td>(0.042)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>VCI$_{\text{Q2}}$</td>
<td>0.224***</td>
<td>0.319***</td>
<td>−0.947***</td>
<td>(0.047)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>VCI$_{\text{Q3}}$</td>
<td>0.334***</td>
<td>0.293***</td>
<td>0.708**</td>
<td>(0.047)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>TCI$_{\text{avg}}$</td>
<td>0.704***</td>
<td>0.836***</td>
<td>1.155**</td>
<td>(0.078)</td>
<td>(0.065)</td>
</tr>
<tr>
<td>TCI$_{\text{Q1}}$</td>
<td>0.159***</td>
<td>0.118**</td>
<td>1.305***</td>
<td>(0.046)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>TCI$_{\text{Q2}}$</td>
<td>0.026</td>
<td>0.120***</td>
<td>0.533*</td>
<td>(0.043)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>TCI$_{\text{Q3}}$</td>
<td>0.739***</td>
<td>0.843***</td>
<td>−1.247***</td>
<td>(0.054)</td>
<td>(0.044)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Barley</th>
<th>Maize</th>
<th>Oat</th>
<th>Rye</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>22151</td>
<td>26336</td>
<td>1868</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R$^2$</td>
<td>0.269</td>
<td>0.254</td>
<td>0.199</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Notes:* In all regressions, the dependent variable is the average annual yield at the cell level (tonnes/hectare). All specifications include time and provincial fixed effects. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.
4.2 Assessing yield trends and their geography

We can use our stylised crop model to develop a primer analysis of the behaviour of yields along time paying attention to the level of drought stress they present. In particular, we distinguish between five levels of drought stress derived from calculating the quintiles of the estimated yields ($\hat{y}_{it}$) and, subsequently, analyse the patterns of actual yields attached to each quintile. With this exercise, we can diagnose the differences in growth of yields in response to drought conditions and assign geographically this behaviour.

A quick look at Fig. 5 enables us to corroborate the general pattern of yields extracted from Fig. 3, that is, yields are basically flat over time regardless drought conditions. This conclusion can be misleading if we do not take into account the effects of local characteristics. As we demonstrate in Fig. 6, there is evidence to believe that certain areas are performing systematically better than others and this has to be accounted for if we want to assess the role of drought stress indicators. Similarly, we find for barley a general flat pattern of yields (Fig. 5) but identify better performances in the banks of river Ebro and Duero (Fig. S.3).

![Fig. 5: Yields trends by quintiles of drought stress. Solid lines show best-fit linear regression for each quintile.](image)

(a) Wheat  
(b) Barley  
(c) Oat  
(d) Rye
4.3 Changes in plant’s drought tolerance through the environment index

In order to isolate the effects of drought indicators on annual yields, we resort to the environment index, as described in Section 3.2. The EI states the expected yield in a certain field, in a certain year, given the outstanding weather conditions. We would test for changes in sensitivity to those conditions by analysing the evolution of yield anomalies with respect to different growing conditions. By splitting the EI in quintiles, we are able to identify five different growing standards with respect to which, we will be able to identify the evolution in performance of each family of crops.

In practical terms, we will use the regression models estimated in Section 4.1 to calculate an EI for each field in each year. In order to study the effect of weather and drought variation, we will set the level of technology fixed at a certain year (2005) for all predictions and subtract province-level fixed effects. We will analyse different levels of drought stress by studying the dynamics of the resulting de-trended variable, that we will know as yield departure. Since the role of drought and weather indicators is only meaningful for non-irrigated varieties, we will not analyse maize in this section.

4.3.1 Wheat

Yield departures of wheat according to different growing conditions (environment index quintiles) and their respective time trends are described in the left panel of Fig. 7.

In general, we observe a flat pattern of average yields for all quintiles except for Q1 (worst growing conditions). Looking at their p-values, all slopes are indistinguishable from 0 except from that relative to Q1. This means that yields under less favourable growing condition (higher drought stress) are experimenting
are relative increase in their performance as they are becoming more resistant or less sensitivity to extreme drought conditions, as measured by soil temperature and vegetation health.

![Graphs showing yield trends for each environment index quintile by year for wheat, barley, oat, and rye.](a) Wheat  (b) Barley  (c) Oat  (d) Rye

Fig. 7: Yield trends for each environment index quintile by year. Solid lines show best-fit linear regressions for each quintile.

We analyse the insights of this response by studying the contributions of our two drought indicators. A detailed analysis of the regression slopes and the contribution of indicators is shown in Fig. 8. The first group of bars illustrates what was previously shown in the left panel of this figure: yields under favourable conditions are stagnated during this period, fields experiencing modest drought stress behave relatively stable and those suffering greater stress clearly converge to total average returns. This is attributable to a general increase of the response of fields to poor VCI and TCI value. Hence, we can state that wheat seeds are more resistant to higher canopy stress and abnormally high temperatures and this can be partly attributed to the development of more resistant seeds (reference?).

4.3.2 Barley

If we replicate the sensitivity analysis to barley fields, we obtain a roughly increasing profile of yields across quintiles, as shown in Fig. 7. This denotes an improvement of barley yields for each variety, regardless the degree of environmental growing conditions. This response derives from a generalised increase in tolerance
with respect to temperatures, as evidenced in Fig. 8. Again, this general improvement is intensified in the case of worst environmental conditions (Q1), where convergence to average yields is observed. This shows the adaptation potential of varieties used in areas more exposed to drought risks. This varieties are characterised by lower average yields and high resistance to extreme water scarcity conditions.

Fig. 8: Contributions of drought indicators to yield trends for each environment index quintile by year.

5 Discussion

The main source of agricultural drought in non-irrigated fields is very clear: the absence of precipitations. The channels through which this water shortage is manifested, though, are various and inter-connected: increase of land temperature, deterioration of canopy health, green-level intensity of plants, soil moisture level, and many others. In order to faithfully test for the impacts on this kind of drought, we should be able to control for most of the mentioned channels. In this study, we control for the first two but we acknowledge that accounting for soil moisture deficit represents a third crucial channel for characterising the level of plant’s drought stress. Hence, soil moisture provides direct evidence of areas suffering rain shortage and this may affect yields.
The use of soil moisture remote sensing measures is gaining importance in the literature in recent years. There are currently two missions devoted to global surface soil moisture monitoring: the Soil Moisture and Ocean Salinity (SMOS) and the Soil Moisture Active and Passive (SMAP). These measures are recently being integrated in drought monitoring programs and drought indices, such as Sánchez et al. (2016). We look forward to integrating a measure of soil moisture in InfoDROUGHT and thus, be able, characterise further dimensions that can have and effect on plant’s drought stress and, thus, on agricultural yields.

Many heterogeneities not accounted in our methodology may impact our final results. In particular, the fact that high yielding varieties are assigned to areas showing consistently better growing conditions over time whereas low-yielding, more resistant varieties are grown in locations exposed to worse growing conditions (as evidenced in Section 4.2) could result in spatially different responses of yields to drought stress. In the event that we suspect that regression coefficients do not remain fixed over space (spatial non-stationarity), we should accommodate this in our methodology. We plan to do so in the future by integrating and applying Geographically Weighted Regressions (GWR) techniques (Brunsdon et al., 1998) to our dataset.

We contribute in this study to quantify the impact of droughts on cereal production in Spain. In particular, we find that drought stress becomes more relevant, the closer we get to the harvesting phase. The results of this study can contribute to the improvement of cereal yield forecasting systems and could help in assessing effective the effective risks a certain field is subject to. Additionally, we find that the resistance of varieties grown in traditionally more drought-exposed areas has raised over the years, which encourages the process of agronomic research on the development of more resistant varieties and represents a reinforcement of national food security under adverse climatic scenarios. We have no reasons to believe that this phenomenon of convergence may be taking place in other countries in Europe. We can easily apply this methodology to different datasets to test locally for this hypothesis.

6 Conclusions

In this paper, we benefit from ESYRCE, a Spanish field-level data set on crop yields and surfaces spanning from the beginning of the 2000’s to 2015 to estimate the impact and reactivity of cereal crops to different drought stress levels. We opt for the use of remote sensing variables to measure plant’s drought stress levels. As highlighted in Lobell (2010), the application of remote sensing data to agronomic studies has proliferated in recent years as a result of the availability of new accurate measures. In particular, we rely on the combination of the Vegetation Condition Index (VCI) and the Vegetation Condition Index (TCI) to characterise the stress of the plant stemming from deficiencies of the canopy health and excess temperatures. The combo VCI-TCI was first used by Kogan (1997) to fully characterise agricultural droughts.

In the first part of this paper, we assess the role of the previous indicators as explanatory factors of yield variability. Despite local yields are explained by an amount of locally specific factors, we find that both the VCI and TCI correlate positively with annual yields, that is, agricultural droughts (characterised by values of these variables below 0.5) can cause yield damages may oscillate from 2% up to more than 20% depending on the crop and the duration and intensity of the drought spell. Moreover, we find that their impact on final output intensifies as we approach the end of the growing period. We also observe that the response of irrigated cereals to these variables is erratic, as evidencing the use of external water resources to accommodate different drought levels.

In a second stage, we use the above estimated model to construct an environment index and analyse the trend of yields under different growing conditions. Exploring actual yields derived from different quantiles of the environment index, we are able to assess how the previous relation has evolved over the past years. We identify an increase in the resistance of wheat and barley fields to extreme drought conditions and certain improvement of performance of barley fields under bad conditions, possibly describing an improvement in agronomic practices and seed resilience.
References


A Integrating field-level yields with satellite drought data

In order to integrate our yield and drought datasets, we face a major conflict, as their spatial resolution are different. On the one hand, the InfoDROUGHT service provides us with satellite pictures taken with a resolution of 1km $\times$ 1km. Meanwhile, ESYRCE yield data has finer detail, as every segment is surveyed extensively so that data for individual fields is reported. The sampling mechanism is described in finer detail in Section 2.1. Since every parcel within sampled segments is geo-referenced, we can assign them to their cell of reference and match the precision level of both yield and drought data. Hence, without a meaningful loss of generality or explanatory power and maintaining we choose 1km $\times$ 1km (100 hectares) as our reference resolution.

However, pixels in the InfoDROUGHT images and cells in ESYRCE (derived from the projection UTM10) do not necessarily coincide. To couple them, we choose to match each south-west coordinate that defines a cell with the respective InfoDROUGHT closest pixel, as obtained by its centroid. After merging both data sets at this (more restrictive) resolution, the overall size of the available data is shrunk up to approximately 25000 yield-year observations of wheat, 30000 of barley and 2200 of maize. Since yields in ESYRCE are expressed on average terms (tons per hectare), we weight each field observation by its total area to aggregate yields at the cell level.

We perform different robustness test of our results using yield anomalies and yield departure over the study period (instead of yield levels) as dependent variable. Due to the sampling strategy of ESYRCE (see Section 2.1), not all cell are visited every year. In order to be able to construct, these alternative measures, we resort further to higher resolution levels and define variables at the block level (equivalent to UTM10 blocks). In this sense, we assign average yield anomalies/departures at the block level with the InfoDROUGHT average value of the pixels contained in that block.

B Additional tables and figures

Table S.1: ESYRCE survey: Field-level descriptive statistics

<table>
<thead>
<tr>
<th>Crop</th>
<th>N</th>
<th>min</th>
<th>max</th>
<th>mean</th>
<th>sd</th>
<th>N</th>
<th>min</th>
<th>max</th>
<th>mean</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barley</td>
<td>64056</td>
<td>1.00</td>
<td>49.00</td>
<td>4.10</td>
<td>4.48</td>
<td>64056</td>
<td>0.20</td>
<td>8.50</td>
<td>2.75</td>
<td>1.37</td>
</tr>
<tr>
<td>Wheat</td>
<td>41683</td>
<td>1.00</td>
<td>49.00</td>
<td>4.10</td>
<td>4.61</td>
<td>41683</td>
<td>0.50</td>
<td>8.88</td>
<td>3.25</td>
<td>1.60</td>
</tr>
<tr>
<td>Oat</td>
<td>7639</td>
<td>1.00</td>
<td>49.00</td>
<td>3.95</td>
<td>4.49</td>
<td>7639</td>
<td>0.10</td>
<td>9.92</td>
<td>1.78</td>
<td>1.12</td>
</tr>
<tr>
<td>Maize</td>
<td>2890</td>
<td>1.00</td>
<td>38.38</td>
<td>4.03</td>
<td>4.27</td>
<td>2890</td>
<td>0.70</td>
<td>14.80</td>
<td>10.63</td>
<td>1.90</td>
</tr>
<tr>
<td>Rye</td>
<td>2481</td>
<td>1.00</td>
<td>39.25</td>
<td>3.21</td>
<td>2.99</td>
<td>2481</td>
<td>0.10</td>
<td>7.00</td>
<td>1.68</td>
<td>0.93</td>
</tr>
</tbody>
</table>
Fig. S.1: Average yields per quintile of drought indicator.
Fig. S.2: Monthly average TCI in different Spanish geographical areas
Fig. S.3: Location of yield quintiles of barley for the period 2003-2015. Darker dots indicate higher stress. Yellowish values indicate best growing conditions.