Simulating the effect of El Niño Southern Oscillation on the worldwide wheat prices

Edmondo Di Giuseppe*

Institute of BioEconomy,
National Research Council,
Rome, 00185, Italy
Email: edmondo.digiuseppe@cnr.it
*Corresponding author

Gianfranco Giulioni

Department of Pedagogical, Philosophical, and
Economic-Quantitative Sciences,
University 'G.d.' Annunzio of Chieti-Pescara,
Pescara, 65127, Italy
Email: gianfranco.giulioni@unich.it

Massimiliano Pasqui

Institute of BioEconomy,
National Research Council,
Rome, 00185, Italy
Email: massimiliano.pasqui@cnr.it

Abstract: This work analyses the impact of the large-scale atmospheric-oceanic phenomenon known as El Niño Southern Oscillation on wheat production, and the consequent changes in prices at the global scale by using computer simulations. Several intermediate results are obtained on the way to the final goal. The identification of geographic areas relevant to the international wheat market and the integration of heterogeneous datasets are two of them. Building on these two results, the local effects of the El Niño Southern Oscillation phases on the wheat yield are quantified using robust ANOVA regression, and their potential impacts on the aggregate production of each area are estimated. Finally, these estimates are provided as inputs to the computational model, which outputs, among others, the wheat prices of 12 internationally relevant production areas. Simulation results show how the cross-section distributions of prices, conditional on the occurring of El Niño and La Niña, spread to the right compared to that observed for the neutral phase. Therefore, both non-neutral phases imply an increase of average and dispersion of prices, although the effect of La Niña is weaker than that of El Niño.

Keywords: computational model; wheat international markets; climate variability; robust ANOVA regression; price cross-section distributions.

Copyright 20XX Inderscience Enterprises Ltd.

Reference to this paper should be made as follows: Di Giuseppe, E., Giulioni, G. and Pasqui, M. (xxxx) 'Simulating the effect of El Niño Southern Oscillation on the worldwide wheat prices', *Int. J. Computational Economics and Econometrics*, Vol. x, No. x, pp.xxx–xxx.

Biographical notes: Edmondo Di Giuseppe received his PhD in Methodological Statistics at the La Sapienza University of Rome, Italy. He is a researcher at the Institute of BioEconomy, National Research Council of Italy. His research interests include issues related to statistical methodologies and software tools to analyse climate phenomena and their impacts. He is the author of several research studies published in national and international journals, conference proceedings as well as book chapters.

Gianfranco Giulioni received his PhD in Economics from the Università Politecnica delle Marche, Italy. He is a Full Professor in Economics at the Department of Pedagogical, Philosophical, and Economic-Quantitative Sciences, University 'G.d.' Annunzio of Chieti-Pescara. His research interests include the investigation of agents' behaviours and interactions to understand their role in generating the aggregate socio-economic systems' outcomes. Business fluctuations, markets' functioning are among the several phenomena of interest. In recent years he focused his research on the application of computational and ITC techniques to understand the international food markets and to enhance food security. He is an author and co-author of articles that appeared in international refereed journals and books.

Massimiliano Pasqui is a climate scientist and graduated in Physics at the University of Tor Vergata of Rome and obtained a PhD in Physical Modeling for Environmental Protection at the University of Bologna, Italy. He is a researcher at the Institute of Bioeconomy of the National Research Council of Italy. His main interests are numerical modelling in support of weather forecasts, climate data analysis, seasonal forecast predictability, and climate change impacts. He participates in a number of national and international projects in these fields. He is the author of numerous scientific papers in international journals/books on atmospheric science and climate change mechanisms and impacts in the Mediterranean basin and Africa.

1 Introduction

Food price dispersion is a key economic issue since the negative effects are both for consumers and producers at the local and global scale. Traditionally, the analysis has been focused on the level of average food staples prices. In fact, rapidly rising food prices are harmful to poor countries who spend as much as 50% to 70% of income on food (Headey and Fan, 2008). A recent interest is to look also at price dispersion because its deeper understanding would give an opportunity to improve local food security policies (Distefano et al., 2019). It could also provide a guide in setting up international commercial relationships when prices dispersion is associated with a high volatility of prices in international food staple markets that would trigger frequent revisions of commercial partnership to limit domestic food prices.

Many factors affect food prices. Wright (2011) highlights how the balance between

consumption, available supply, and stocks seems to be as relevant for our understanding of these markets as it was decades ago. Ott (2014) analyses the role of the stock in supply-side shocks finding that low cereal stocks mostly influence inter-annual volatility while large cereal stocks reduce the effects of shocks on both intra- and inter-annual volatilities. Moreover, this author finds that the derivatives market does not have a significant impact differently from the crude oil price and exchange rate, especially for intra-annual volatility. Climate variability is another prominent factor in influencing food prices' dispersion (Ubilava, 2018). This is mainly due to the impact of large-scale climate oscillation on crop productivity that Heino et al. (2018) demonstrate to involve the two-thirds of global cropland area.

The objective of this study is to investigate what is the role of a large scale climate driver in modulating the global price of wheat, which together with rice, maize, and soybean are employed as a fundamental element of food and have become a staple in over 97% of countries (Khoury et al., 2014). To this end, we develop a computational model that incorporates the varying of wheat quantity derived from climate variability. In particular, we focus on El Niño Southern Oscillation (ENSO)[1] to reproduce such climate variability at a large scale.

Among large scale forcing mechanisms on Earth, ENSO represents the most prominent climate oscillation. It is a coupled oceanic-atmospheric phenomenon that develops in late boreal summer in the Tropical Pacific Ocean. ENSO reaches an intensity peak in the following winter and dissipates a few months later, during Spring. Such a large scale oscillation modulates weather patterns and ecosystem fluctuations worldwide. During its three phases, warm, cold and neutral (commonly known as El Niño, La Niña, and neutral, respectively), it directly forces weather regimes in the Pacific region and provides indirect influences to others through teleconnection mechanisms (Hu et al., 2020). Such modifications on weather regimes modify the distribution of precipitation and temperature patterns in the Tropics and Sub-Tropics areas mainly. Often negative impacts, such as drought periods or floods are produced (Rojas et al., 2014) and their effects are transferred on crop production (see for example Iizumi et al., 2014a) affecting global food.

Several studies examine the role of climate variability in determining global wheat price variation. Most of them are econometrics models which estimate the response of global wheat price dynamic to the modulation of large scale climate drivers. In particular, since climate affects the supply side of the market, they use some climate variability index as shocks in the equilibrium framework. For instance, Ubilava (2018) proposes an econometric framework to account for possible structural change and regime-dependent nonlinearities in the commodity price dynamics. In particular, the author adopts a time-varying smooth transition autoregressive model to account for the nonlinearity of the ENSO-prices relationship due to the nature of climate anomalies. Algieri (2014) investigates the drivers of wheat prices and quantifies their impact through a vector error correction model. Gutierrez (2017) uses a global vector autoregression model to account for the transmission of shocks from one country to others markets and to analyse ENSO impacts on wheat yield anomalies, export prices, exports, and stock-to-use ratios. The present work aims at contributing to this literature through a generative approach (Epstein, 2006) rather than an econometric approach. In the generative approach, the feature under investigation emerges from the interaction of the agents in the model. The system evolves freely without any restriction to achieve a particular state, such as equilibrium. Given the presence of a considerable number

of agents, computational modelling techniques, also used in this study, are at the heart of the generative approach. To our knowledge, any of the existing studies examine the patterns of price variability using this approach at the global scale, though a few introduce the agent-based modelling to determine the wheat price after farm behaviour in the South of Italy (Siad et al., 2017) or the computational approach to predict the economic impact of climate change-induced loss of agricultural productivity in Pakistan (Khan et al., 2020). Thus, the present paper is a novelty in the adoption of economic computational models to analyse price dispersion induced by climate variability.

This study is composed of two models:

1 the regression model to estimate the impact of ENSO on wheat production
2 the computational model to simulate the international wheat market in presence of climate-derived variation in wheat supply.

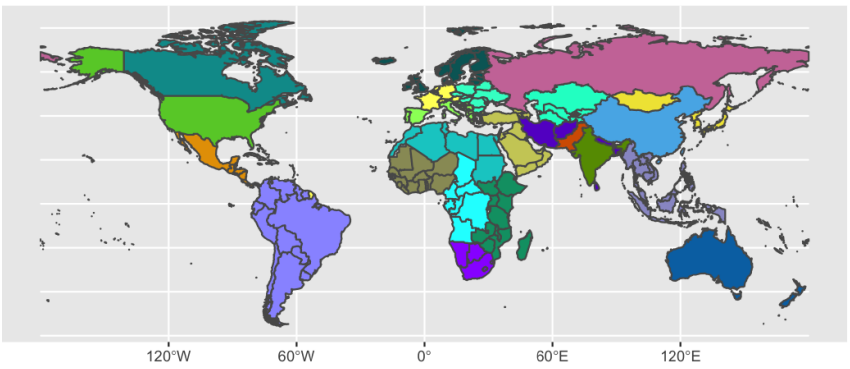
The sequential application of these two models allows identifying how the ENSO phenomenon affects international wheat prices. Important preparatory work was necessary before building and applying the two models. This work is described in Appendices A and B1. The article is organised as follows. Section 2 presents the aggregation/disaggregation of geographic areas made to reach a spatial configuration which can better represent geographic areas most relevant in the international wheat market. Several sources of data were used and combined to estimate the macroscopic effects of the ENSO phenomenon. Section 3 presents some descriptive statistics aiming at highlighting what real-world data tell us on the relationship between ENSO phases and wheat prices. The following two sections describe the regression model and the virtual international wheat market model, respectively. Section 4 illustrates the statistical techniques used to regress yield on a one-dimensional index representing ENSO phases dynamic and the findings of the correlations worldwide. Section 5 presents the computational model used to simulate how the climate effects found in Section 4 affect wheat international markets. The section also discusses the results obtained from simulations. Finally, Section 6 provides conclusions and discusses possible future research.

2 Spatial aspects

Being interested in an international analysis, the geographic setting of the model has an important role. As will be discussed below, several data sources are used in our investigation. Among them, the FAOSTAT datasets from Food and Agriculture Organization (FAO) of United Nations have a prominent role. The spatial domain of the model is therefore basically in line with FAOSTAT classification (which in turn is derived from the United Nations M49 standard) at the 'intermediate regions' level, i.e., sub-continental geographic areas normally composed of several countries. The sub-continental areas considered in this study are reported in the lower part of Figure 1 (Table 7 in Appendix C reports all the countries belonging to each region). The intermediate region world partitioning was amended to explicitly account for countries playing a relevant role in the world wheat production/consumption system. These countries are listed in the last line of the table included in Figure 1 where the superscripts connect each country to the region from which it was parceled out. The

world map highlighting regions and countries is displayed in Figure 1. This geographic setting is also used in a previous work on which we will base the computational part of this work (Giulioni et al., 2019).

Figure 1 Regions and relevant single countries considered in the model (see online version for colours)



Continents	Regions
Africa	Eastern, Middle, Northern, Southern, Western
America	Carribean, Central, Northern[a], South
Asia	Central[c], Eastern[d], South-eastern, Southern[b], Western
Europe	Eastern, Northern, Southern, Western
Oceania	Oceania
Countries	US[a], India[b], Pakistan[b], Russian Federation[c], China[d]

Notes: The superscripts in the table connect each of the five relevant single countries with the region it was a parceled of.

To obtain data for the geographic entities identified here, we use several datasets. Some of them provide data at the country level, while others refer to rectangular portions of the globe which dimension depends on dataset resolution. We refer to the latter as gridded datasets. The datasets we use are the following:

1 the FAOSTAT time series of yearly wheat production, yield, and harvested area (http://www.fao.org/faostat/)

2 the GDHY wheat yield gridded dataset described in Iizumi and Sakai (2020)

3 the Global Crop Planting Dates (GCPD) dataset described in Sacks et al. (2010) that includes spatial patterns of harvested area for winter and spring crop sowing and their corresponding crop calendars

4 the ENSO quarterly time series classified into neutral, El Niño, and La Niña phases

5 the FAOSTAT 'wheat and products' database for the components of wheat demand.

More details on these datasets are given in Appendix A. Since these datasets have a different spatial and temporal resolution, we have to aggregate them to common scales. The spatial aggregation of 1, 2, 3 and 5 is done according to the geographic units shown in Figure 1 while 4 has only a temporal resolution. In particular, for the gridded datasets 2 and 3 that have an equivalent resolution of 0.5° by 0.5°, we aggregate the cell grid values by overlaying 24 raster polygons that represent the shape of the geographical units listed in Table 1. On the other hand, data of 1 and 5 are released both at regional and national levels where the regional aggregations correspond to those chosen for this study. Nevertheless, we have to break up national time series of the US, India, Pakistan, the Russian Federation, and China from their original regions.

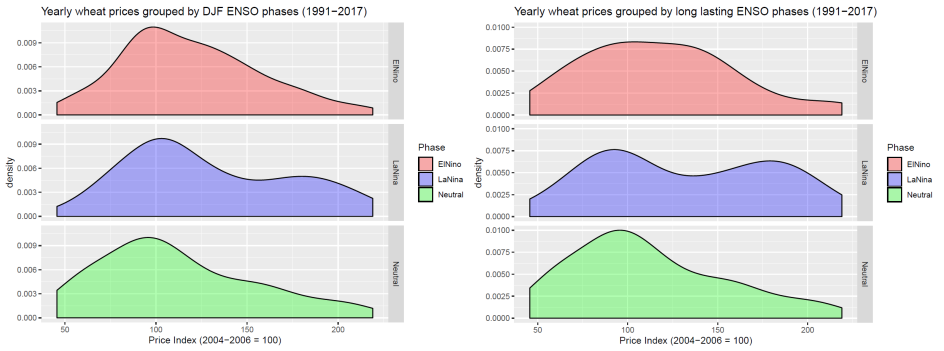
The temporal aggregation used in this study is the year. Data from 1, 2, and 5 are already aligned to it, although with different time windows: 1 – 1961–2018, 2 – 1981–2016, and 5 – 1961–2017. The time series of dataset 4 is composed of 12 quarterly values for each year and is currently updated from 1959 on, while 3 is an instantaneous picture of spatial patterns.

3 Empirical regularities of climate effect on wheat prices

By investigating the distributions of wheat price levels conditional on ENSO phases, we find that there is a strong correlation with wheat prices when El Niño or La Niña are active in the December-January-February (DJF) quarter, especially with La Niña. The result is shown in Figure 2(a). As explained in Appendix A, the ENSO phases we consider are computed on the sea surface temperature anomaly averaged on a three months moving window. The averaging quarters are identified with a three-letter acronym made of the months' initials, e.g., DJF stands for December, January, February; JFM for January, February, and March, and so on. This is the notation employed in Figure 2 and will be used in the paragraphs below. It is also interesting to analyse the same conditional distributions when the ENSO phenomenon begins in DJF and last up to MAM [Figure 2(b)]. The time series of yearly export prices index used here are retrieved from FAOSTAT (http://www.fao.org/faostat/en/#data/PI) and refer to the markets of wheat top exporters. The export price index is calculated by taking as base the period 2004–2006.

In particular, the plots show that both El Niño and La Niña increases the probability of observing high prices. It is also worth noticing that the spread of prices is larger with long-lasting phenomena and, particularly, La Niña causes a bimodal shape of the price distribution. This asymmetric response of prices to the different ENSO phases has been highlighted in other studies (Algieri, 2014; Ubilava, 2017). It is also related to the different local effects manifestation such as the heterogeneity of drought risk associated with ENSO (Rojas et al., 2014; Heino et al., 2018).[2] Such a macro-scale asymmetry modulates wheat and grains cereal production in the Mid-Latitude regions. Local supply shortages are therefore more likely during La Niña.

Figure 2 Distribution of wheat export prices conditioned on the three ENSO phases of El Niño, La Niña, and neutral observed in, (a) DJF (b) long-lasting from DJF to MAM (see online version for colours)



(a) (b)

Note: Prices are from the worldwide top exporters.

4 Modelling climate effects on wheat yields

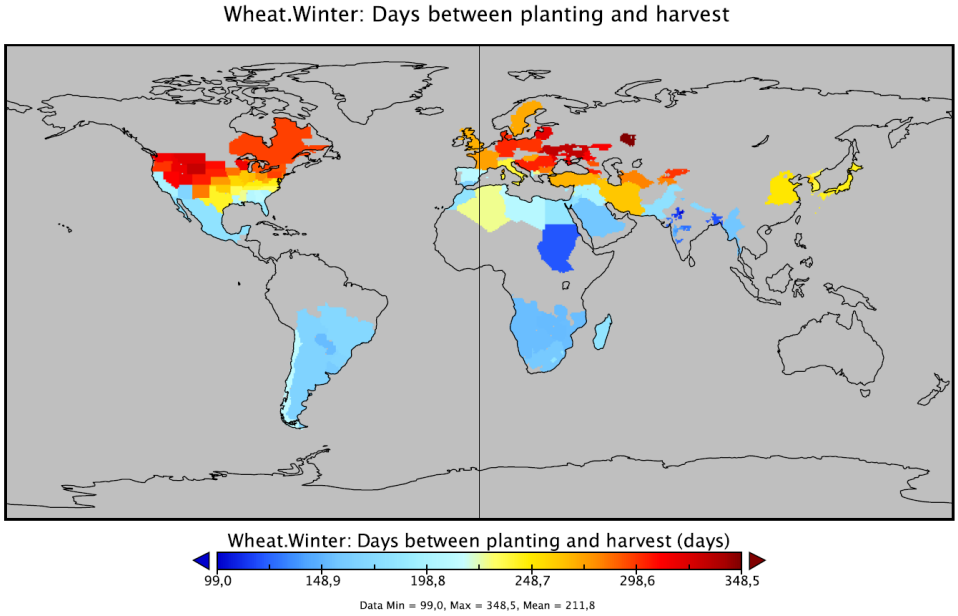
The influence of ENSO on rainfall and temperature patterns has been largely studied in the last decades. An overview of them is shown at https://www.pmel.noaa.gov/elnino/impacts-of-el-nino. In this study, we consider the ENSO impact on the wheat production of 18 groups of countries in addition to India, Pakistan, China, the Russian Federation, and the USA (see Figure 1). We also account for different time-lags of ENSO impact on the wheat-growing season according to the spring and winter crop sowing. In fact, in many places in the world, the sowing of wheat can be done either during winter or spring. Hereinafter, we use the term 'crop sowing' to distinguish between winter and spring sowing. Furthermore, since "yield is more important in determining production because of the large year-to-year variability of yield associated with climatic factors" (Iizumi et al., 2014a), in the analysis of ENSO correlation with wheat production, we analyse yield instead of production.

The components of the approach adopted to this end are explained either in the following paragraphs or in the Appendix. Data preparation, analysis, and modelling, as well as most of the plots, are done with the R software (R Core Team, 2020) and, particularly, by using the set of packages included in *tidyverse* (Wickham et al., 2017). For the raster cropping and variables aggregation we use the *raster* package (Hijmans, 2019). To apply robust ANOVA regression, we use the WRS2 package (Mair and Wilcox, 2019).

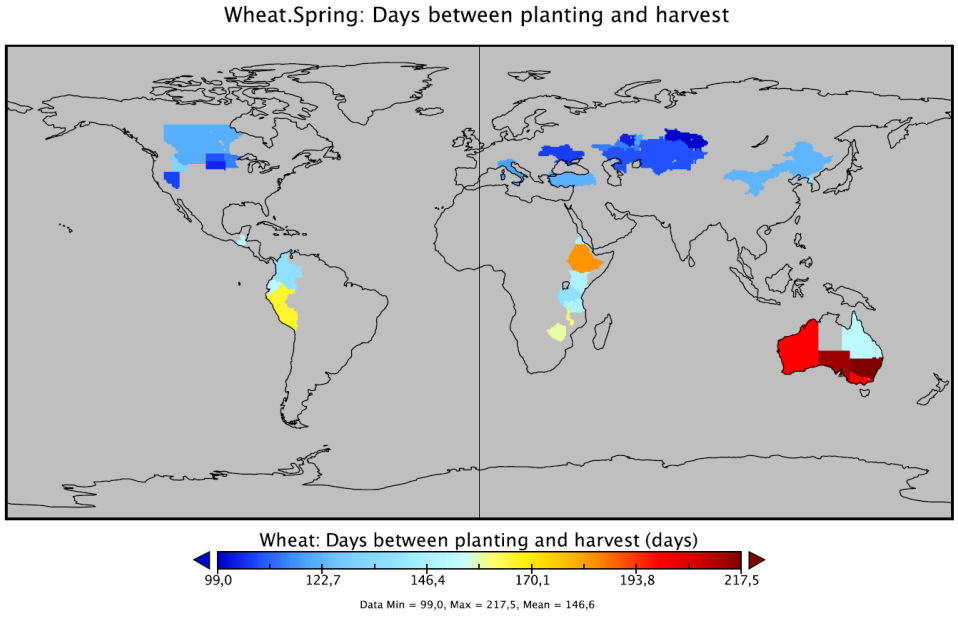
4.1 Selection of potentially impacting ENSO quarters on the growing season

To find a correlation between the time series of ENSO and wheat yield, it is crucial to identify the growing season of this crop for each geographical unit and to find out the period of the ENSO activities more strongly connected to such a growing season. The length of the growing period for winter and spring wheat extracted from the GCPD dataset is displayed in Figure 3.

Figure 3 Global distribution of wheat growing season obtained from Sacks et al. (2010), (a) winter (b) spring (see online version for colours)



(a)



(b)

These periods are important because the longer the growing period, the longer the list of ENSO quarters that might have an impact on the harvest. Using this Figure, we can also identify areas with potential double rotation, i.e., areas where both the winter and spring sowing are managed along the same year.

Table 1 Growing period and ENSO impact time window for winter and spring varieties

Country-group	*Growing winter*	*Growing spring*	*Impact winter*	*Impact spring*
Eastern Africa	(May–October)	(June–December)	DJF -> MAM	ASO -> OND
Middle Africa	(May–October)		DJF -> MAM	
Northern Africa	(November–July)		OND -> MAM	
Southern Africa	(June–November)		ASO -> OND	
Western Africa	(December–July)		NDJ -> MAM	
Northern America	(October–August)	(May–September)	OND -> MAM	DJF -> MAM
Central America	(November–May)		OND -> MAM	
South America	(June–December)	(January–August)	ASO -> OND	DJF -> MAM
Central Asia	(September–July)	(May–September)	OND -> MAM	DJF -> MAM
Eastern Asia	(October–June)	(February–July)	OND -> MAM	DJF -> MAM
Southern Asia	(October–June)		OND -> MAM	
South-Eastern Asia	(October–April)		OND -> MAM	
Western Asia	(December–May)	(April–August)	NDJ -> MAM	DJF -> MAM
Eastern Europe	(October–July)	(May–September)	OND -> MAM	DJF -> MAM
Northern Europe	(October–July)		OND -> MAM	
Southern Europe	(November–June)	(April–August)	OND -> MAM	DJF -> MAM
Western Europe	(November–August)		OND -> MAM	
Oceania		(June–December)		ASO -> OND
Russian Federation	(August–August)	(May–August)	ASO -> OND	DJF -> MAM
India	(November-March)		OND -> FMA	
Pakistan	(August–May)		OND -> MAM	
USA	(September–August)	(May–September)	OND -> MAM	DJF -> MAM
China	(September–June)	(March–August)	OND -> MAM	DJF -> MAM

The columns *growing winter* and *growing spring* in Table 1 report the growing period for the 23 geographical units[3] that results from the aggregation of the gridded dataset. To obtain this aggregation, we firstly crop the GCPD visualised in Figure 3 over the area covered by the 23 geographical units. Secondly, we count the cells with the same planting and harvesting day of the year. Finally, we identify the month to which the most frequent day belongs. For example, the period reported for the winter sowing of Eastern Africa in Table 1 is May–October, this means that May (October) is the month during which the planting (harvesting) activity has been more frequent in that area. Table 1 shows how the winter sowing is largely widespread and that the double rotation is adopted in 12 out of 23 geographic units. It is worth noticing that the unique crop sowing adopted in Oceania is commonly attributed to Spring although the growing season Jun-Dec covers usually the winter months.

In Table 1, the columns *impact winter* and *impact spring* report the period of ENSO activity that can potentially have effects on the wheat growing season. In fact, the most intense activity of this large scale phenomenon is during the autumn-winter time when it reaches a peak. Nevertheless, the effects on weather and climate regimes can persist for a longer period until the thermal inertia of its oceanic component is reached. Thus, its effects are pushed further to the subsequent months such that they extend to the wheat growing season.

4.2 One-way robust ANOVA regression model for wheat yield-ENSO correlation

Regression analysis is used to create models that measure the effect on the dependent variable due to a unit of variation in predictor variables. In our model, the dependent variable is the wheat de-trended yield percentage while the unique predictor is the ONI. It is worth pointing out that the de-trended yield percentage is calculated by dividing the residuals of local polynomial regression by the corresponding fitted values (see Appendix B2), and that ONI is computed by averaging sea surface temperature anomalies in the Niño-3.4 region of the Pacific Ocean and that it also provides a classification of such anomalies corresponding to the three ENSO phases of El Niño (warm), La Niña (cold), and neutral (see Appendix A). Then, ONI has either a continuous or a categorical configuration while the dependent variable is continuous. This opens the possibility to apply a simple linear regression model by taking into account the value of ONI anomalies, but also an analysis of variance (ANOVA) regression model. Nevertheless, a combination of the two is also an option, i.e., the analysis of the covariance (ANCOVA) regression mode. However, as discussed in Section 1, the most important effects pushed by ENSO on worldwide ecosystems come from the modulation of the phases rather than the level of deviation from the neutral situation. As a consequence, we choose the one-way ANOVA (one-way because there is only one predictor) regression to model the effects of the ENSO warm and cold phases on percentage variation of wheat yield in contrast to the level of yield normally expected during the neutral phase. In practice, with the one-way ANOVA, we study the effect of the predictor by pairwise comparison of the variances for each level of the categorical variable, i.e., El Niño, La Niña, and neutral. When the gap is statistically significant, we estimate the value of the percentage difference in yield between El Niño and neutral, La Niña and neutral. However, we find in our data many situations of unequal sample sizes and differences in skewness among the three groups of observations as well as the presence of outliers. These deviations from normality assumptions make the statistical test unuseful. Therefore, also because of the nonhomogeneity of the three variances that is a pre-condition for the application of ANOVA, we adopt its robust version, the so-called robust ANOVA one-way.

The robust estimation was introduced by Tukey (1959), Huber (1964), and Hampel (1968) since the ‘60s. The robust test and estimation used here refer to the work of Wilcox (2012) and, consequently, in the following we call it *Wilcox test*. In particular, we perform the comparison of multiple trimmed group means, with a trimming threshold equals to 0.20. In practice, we exclude the lowest 10% and the largest 10% of values for each group and compute the arithmetic mean on the remaining values. For example, the trimmed mean of yield during El Niño is $\bar{X}^E_{yield} = \sum_{i=1}^{n_E} x_i^E / n_E$ with n_E being the effective sample size after the trimming. Therefore, the null hypothesis of the test is H_0 : $\mu^E_{yield} = \mu^L_{yield} = \mu^N_{yield}$ where μ indicates the arithmetic mean of trimmed observations and E, L, and N stand for El Niño, La Niña, and neutral, respectively. Similarly to classical ANOVA, a significant test is followed up by a post-hoc test to perform multiple pairwise comparisons between groups. Nevertheless, when the objective of the study is to perform a comparison between groups, there is no need to first obtain a significant omnibus ANOVA. Therefore, we directly apply the Wilcox post-hoc test on the trimmed means. This test belongs to the family of *F-distributed Welch-type* test statistics that allow for unequal variances in the groups (Welch, 1951) and its test statistic for two independent groups is of the form:

$$T = \frac{\bar{X}_{g1} - \bar{X}_{g2}}{\sqrt{d_{g1} + d_{g2}}} \tag{1}$$

where $\hat{X}$ is the trimmed mean of the group g and d is a measure of its squared standard errors.

Since we have three independent groups, the test statistic of the multi comparison between two groups can be easily built from equation (1) by applying a linear constraint

$$\hat{\Psi} = \sum_{j=1}^{3} c_j \bar{X}_{g_j} \tag{2}$$

and setting two of the constraints c_j to 1 and –1 and the third to 0. The specificity of the test is determined by the form of the denominator in equation (1), which results in an F-distributed test statistic. Details on its definition can be found in Wilcox (2012). Finally, being respectively $\sigma^2(Y)$ and $\sigma^2(\hat{Y})$ some robust measure of the variation associated with observations and fitted values under the null hypothesis, the estimate of the effect size is $\epsilon = \sqrt{\epsilon^2}$ where

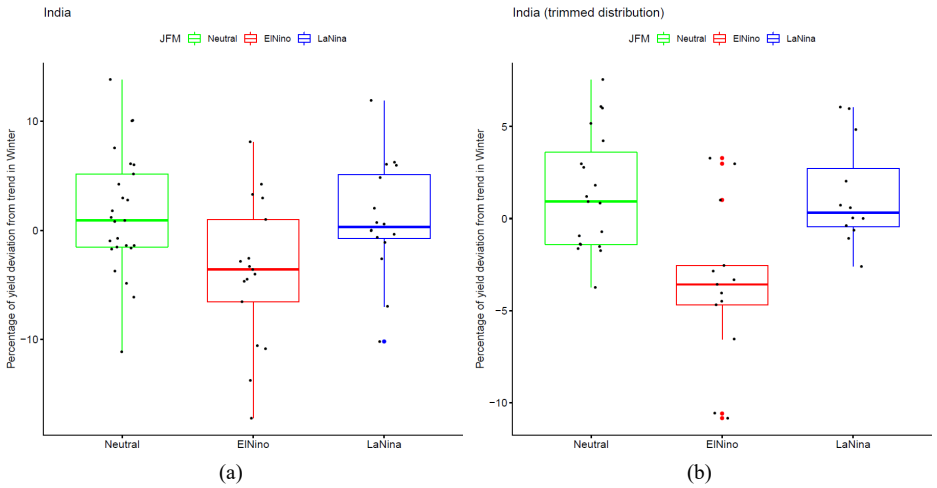
$$\epsilon^2 = \frac{\sigma^2(\hat{Y})}{\sigma^2(Y)}$$

In particular, being $\bar{X}_{yield}$ the grand mean, $\sigma^2(\hat{Y})$ is estimated with

$$(\bar{X}^E_{yield} - \bar{X}_{yield})^2 + (\bar{X}^L_{yield} - \bar{X}_{yield})^2 + (\bar{X}^N_{yield} - \bar{X}_{yield})^2$$

where $\bar{X}_{yield}$,$\bar{X}^E_{yield}$, $\bar{X}^L_{yield}$, and $\bar{X}^N_{yield}$ are the trimmed means.

Figure 4 Distribution of India's yield observations of winter sowing in the three phases of ENSO, (a) complete distribution (b) trimmed distribution (see online version for colours)



(a) (b)

We apply the bootstrap version of this heteroscedastic one-way ANOVA post-hoc test for trimmed means with 999 repetitions over the entire set of quarters reported in the columns *impact winter* and *impact spring* of Table 1 and select the most strong among those that result to be significant. For the sake of clarity, we describe the entire procedure for India while the complete list of significant effects found is reported in Table 3. India adopts a mono-cropping winter sowing and the growing season goes from November to March (see Table 1). The table also shows the list of potential ENSO quarters that are tested: from October-November-December (OND) to February-March-April (FMA). The distribution of yield observations in each phase is displayed in the box-plot of Figure 4.

Firstly, a trimmed mean is applied over these three distributions leading to those shown in Figure 4(b), secondly the Wilcox post-hoc test is performed by applying the function *mcppb20* included in the *WRS2* R package (Mair and Wilcox, 2019). Finally, the significant effects are selected by means of the P-value. The result of the test for India is reported in Table 2.

Table 2 Robust ANOVA-one way test results for India

Group 1	*Group 2*	$\hat{\Psi}$	*ci.lower*	*ci.upper*	*p.value*
Neutral	El Niño	–4.73	–9.61	0.11	0.02
Neutral	La Niña	–4.74	–9.30	–0.21	0.02
El Niño	La Niña	–0.01	–4.14	3.65	0.96

Table 3 Robust ANOVA-one way percentage effects of ENSO on winter and spring yields

Country-group	*CropSowing*	*Quarter*	*Phase*	*Effect (%)*	*Neutral (%)*
Northern Africa	Winter	OND*	El Niño	–7.75	2.90
Northern Africa	Winter	FMA	La Niña	4.12	0.84
Central America	Winter	FMA	El Niño	–2.94	–0.25
South America	Winter	OND	El Niño	1.42	–4.18
Central Asia	Winter	MAM	El Niño	12.01	–1.51
Central Asia	Spring	MAM	La Niña	28.49	–5.94
Southern Asia	Winter	MAM	El Niño	2.41	1.04
Northern Europe	Winter	DJF	La Niña	–2.44	1.60
Oceania	Spring	SON	El Niño	–3.49	0.74
India	Winter	DJF	La Niña	0.08	1.35
India	Winter	JFM	El Niño	–4.74	1.22
USA	Spring	MAM	La Niña	–1.10	–1.47
China	Spring	DJF	La Niña	–1.92	1.61

Notes: The last column reports the trimmed mean of yield during the neutral phase.
*indicates that the identified quarter belongs to the year before the one when the effect is quantified.

There, the effect under linear constraint $\hat{\Psi}$ is shown in the corresponding column while the associated lower and upper confidence intervals are reported in the column *ci.lower* and *ci.upper*, respectively. Note that in order to obtain the explanatory effect size of El Niño or La Niña with respect to neutral, we have to take into account the transformation adopted with the linear constraint [see equation (2)]. That is why the

La Niña value of $\hat{\Psi}$, very similar to that of El Niño, corresponds to an effect size close to 0 (not shown). On the contrary, the El Niño back-transformed value of $\hat{\Psi} = 4.74$ just corresponds to the estimated effect size, i.e., a decrease of yield with respect to the neutral phase equals to 4.74%. This value is also reported in Table 3.

The results show that the wheat yield of 10 out of 23 geographic units is highly correlated with the climate oscillation generated by ENSO. Moreover, the yield of Northern Africa, Central Asia, and India is subject to the variation determined by either El Niño or La Niña. Depending on the type of crop sowing and Hemisphere, the number of significant effects associated with El Niño (7) is almost the same as that associated with La Niña (6). Also, the sign of the effects is more or less equally distributed among the two: 4 (3) negative (positive) with El Niño, 3 (3) negative (positive) with La Niña. On average, the level of the explanatory effect caused by La Niña seems to be more strong than that caused by El Niño, especially because of the high effect found in Central Asia. Although statistically significant, this finding is supported by very few cases happened out of the neutral phase. In the future development of the model, this will require a further investigation of the climate variability at the local level. Nevertheless, in general, these results are consistent with the findings of other studies, for example Iizumi et al. (2014a) and Iizumi et al. (2018) where it clearly appears that the impact of such effects on production strictly depends on local factors such as climate mechanism and the harvested area devoted to wheat cultivation in each geographic unit.

4.3 Estimates of regional wheat production under the ENSO effect

To identify the ENSO impacts on wheat production, we use the estimates of yield variation listed in Table 3 to compute the corresponding variation of production given the harvested area. More specifically, the 13 significant effects found through the statistical model are applied over the years classified by the two phases of El Niño or La Niña, correspondingly, then multiplied by the harvested area devoted to wheat cultivation in those years. Therefore, we obtain a time series of estimated variation of production due to the found correlation with ENSO for each of the interested geographic unit. Finally, the complete time series of production is obtained by summing the estimated variation to the quantity estimated through the *loess* procedure described in Appendix B2. Note that the production of years characterised by a neutral phase as well as that of the geographic units where ENSO has no effect is also estimated through the *loess* procedure. These time series are an input of the computational model presented in the next section..

5 Modelling climate effects on wheat markets

Our previous analysis allows to identify the ENSO effects on wheat yield and to compute the corresponding impact on the production quantity in the considered geographic units. In this section, we will use those results to evaluate the response of the international markets to such effects. The analysis is performed by integrating wheat production and wheat demand.

As one can easily imagine, this market is a complex object whose analysis is not easy. We, therefore, take advantage of computer simulations.

5.1 The computational model

We build our investigation on an existing computational model able to analyse international wheat markets. The starting point is the commodity markets simulator for wheat (CMS-wheat), which is a computational model addressed to the analysis of wheat spot price formation, its dynamics, and the dynamics of exchanged quantities. In this section, we briefly summarise the functioning of the model details of which can be found either in Giulioni et al. (2019) or in the supporting material available at the software Github repository https://github.com/gfgprojects/cms_wheat.

The model handles the dynamics of several producers/suppliers, several consumers/users, and their interactions on markets. The interplay between demand and supply on markets determines prices and exchanged quantities.

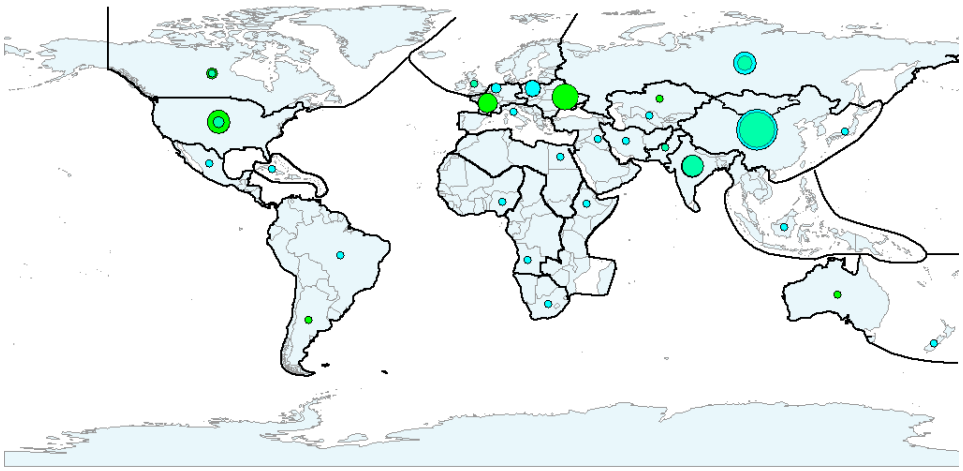
Table 4 Geographic regions and market sessions

Continent	*Region*	*Has an internat. market session?*	*Outgoing hub location*	*Incoming hub location*
Africa	Eastern			Ethiopia
	Middle			Angola
	Northern			Egypt
	Southern			South Africa
	Western			Nigeria
USA	Northern – USA	✓	USA	USA
	Northern except USA	✓	Canada	Canada
	South	✓	Argentina	Brazil
	Central			Mexico
	Caribbean			Cuba
Asia	Southern – India	✓	India	India
	Southern – Pakistan	✓	Pakistan	Pakistan
	Southern except India and Pakistan			Iran
	Central – Russian Federation	✓	Russian Fed.	Russian Fed.
	Central except Russian Federation	✓	Kazakhstan	Uzbekistan
	Eastern – China	✓	China	China
	Eastern except China			Japan
	South-Eastern			Indonesia
	Western			Iraq
Europe	Eastern	✓	Ukraine	Poland
	Northern	✓	UK	UK
	Western	✓	France	Netherland
	Southern			Italy
Oceania		✓	Australia	New Zealand

Producers/suppliers and consumers/users in this model are the wide regions and countries listed in Table 1. On the supply side, each producing area has a wheat outlet. The model distinguishes outlets that are relevant at the international level from those which are relevant at the local level only. The latter are those whose production is always lower than demand. Because the focus of the model is studying the international wheat market, only internationally relevant outlets are kept in the analysis, i.e., demand

and supply of these regions are both considered in the model. Only the net demand, i.e., demand minus production, is considered at the international level for the other regions, and these areas' productions are assumed to be addressed to local users. Internationally relevant outlets are gathered in the international market and will be referred to, more formally, as market sessions.

Figure 5 Regions considered in the model (thick line contour) and commercial hubs (circles) (see online version for colours)



Notes: Green circles denotes outgoing hubs, light blue circles denotes incoming hubs. The size of the circles inform about produced and used quantities.

12 of the 24 areas and countries listed in Table 1 has a production permanently higher than their domestic demand. Therefore they are allowed to sell in the international wheat market. See the 'has an international market session?' column in Table 4. All the 24 areas direct their demand to the international market. In particular, each of them prepares a demand schedule for each of the 12 considered producers and sends it to the relative market session before the market opens. After opening, each producer aggregates the 24 demand schedules received and sets the wheat price to the level that allows it to sell a quantity as close as possible to the wished one. Once the price was established, the quantity assigned to each buyer is computed by using the individual demand curve. Following this assignment, the wheat is moved from the producer outgoing hub to the buyer's incoming hub identified in Table 4 and displayed in Figure 5. The distance between these hubs is used to compute transport costs that are paid by the buyer, and therefore it increases the cost of wheat. Buyers use the information on the costs of wheat obtained from the different producers to modify demand schedules that will be sent to the market before the next opening. Demand schedules are also moved in case the gathered quantity differs from the buyer's target.

The international wheat market opens at regular time intervals, i.e., monthly. Producers and buyers gradually adjust their behaviour over time to achieve their goals. In particular, producers aim at selling the quantity harvested in a year throughout the various market sessions performed in a year. They aim at selling at the highest possible price, provided that the selling price is higher than production costs. Buyers aim at

obtaining a target quantity at the lowest possible cost. Therefore, they manage to move their demand towards those producers who were recently experienced as cheaper.

5.1.1 Changes to the model

The change to the computational model was performed to include the progress made in the above analysis which distinguishes the winter and spring crop sowing. The model from which we start (Giulioni et al., 2019) is a first attempt to investigate the international wheat markets and, being based mostly on the FAOSTAT databases, considers the total yearly production. Therefore, among its inputs, there is a time series of yearly production and an integer number informing on the mean harvesting month (HM) for each producer. In the new version of the model, these inputs are composed of two time series and two integer numbers. They give the winter and spring production and the months in which these productions are harvested, respectively.

The primary consequence of this development is that quantities supplied in the various market sessions change more often, mimicking reality more closely than the original model.

5.2 Simulation procedure

Wheat markets are affected by a large number of factors and historical events. Extrapolating the climate effects is a hard task in a dynamic context where a changing number of these effects merge in time. To identify the climate effects on wheat prices, we favor the *ceteris paribus* analysis made possible by a simulation approach rather than its most frequently used full-blown dynamic ability.

Our approach consists of selecting a reference year in which the production was not significantly affected by a relevant event such as war, import and export ban, and similar. From the climate point of view, this means that the selected year was preceded by a persistent level of ONI close to zero. We then run a simulation giving as inputs the quantities obtained from the databases in the chosen year. Keeping these values unchanged over time, we obtain the output time series (basically prices and quantities exchanged in the various market sessions) to be used as benchmark. In the following runs, the model has the same initialisation, but one of the inputs is changed at a given point in time.

Comparing the set of time series obtained from each simulation run with those obtained from the benchmark we can evaluate the effect of the changes in the inputs on outputs.

5.2.1 Reference year selection

In the last decades, a few periods were characterised by a persistent level of neutral ENSO phase and thus a weak ONI index. According to the ONI index time series[4], the years characterised by a prevalent neutral ENSO phase are: 1980–1981, 1989-1990-1991, 2001–2002, 2003–2004, 2012-2013-2014. Among them, we select

2014 which is the second in a row of three consecutive neutral years, since the identification of the phase depends on the ONI value registered during the winter 2013–2014 and up to that of spring 2014.

Our choice is to use the most recent candidate in the conviction that available data are more reliable. Moreover, we point out that FAOSTAT changed its methodology in the computation of demand components from 2013.[5]

5.2.2 Simulations

As hinted at above, simulations are performed providing to the model data we compute for 2014. In particular, the quantities that fill the supply side of the benchmark simulation are reported in the central column of Table 5.

The other simulations are obtained by changing the harvested quantities at a certain point in time with those that would be observed, according to our previous analysis, if El Niño or La Niña had replaced the ENSO neutral phase happened in reality. More precisely, simulations run on a monthly base, i.e., the international wheat market operates once a month. Each simulation lasts 72 time ticks.

The model has several parameters that are listed and described in the documentation available at https://github.com/gfgprojects/cms_wheat. The most relevant ones, regulate the shape of the demand curve, the level of transport costs and the buyers behaviour on moving demand from most to less expensive producers. Because our analysis focus on one year, the parametrisation of the model is that used in Giulioni et al. (2019) who implemented a computationally intensive procedure for parameter setting using data from 1992 to 2013. There, the calibration process for those parameter that cannot be established using empirical observations was made by an heuristic global optimisation process. The differential evolution algorithm (Storn and Price, 1997) was used for this purpose. Briefly, the process goes through the following steps:

- a population of parameters vectors is taken from the parameter space
- the model is run for each element of the population
- the goodness of each element is evaluated by computing the distance of the yearly prices obtained from the simulation with those observed in reality
- the element of the population producing the minimum distance is selected
- starting from this element, a new population is generated aiming at obtaining lower values for the distance
- if an improvement is realised, the element of the population that provided for this improvement is selected to generate a new population
- the process is iterated until a stopping condition is met.

The outcome of this process provided very good results which are better described in Giulioni et al. (2019).

Table 5 Production in 2014 and estimated change in produced quantities if the ENSO phase would be El Niño or La Niña instead of neutral, CS = cropping systems (w = winter, s = spring) and HM = harvesting month

Area	*Share*	*CS*	*Production (p)*	*HM*	*Δp El Niño*	*Δp La Niña*
Northern America	4.84%		29,758,963			
		w	6,058,129	9		
		s	23,700,834	10		
South America	3.6%		22,122,784			
		w	16,736,353	1	234,329	
		s	5,386,431	9		
Central Asia	3.71%		22,778,168			
		w	9,559,974	8	1,148,153	
		s	13,218,194	10		3,765,863
Eastern Europe	9.74%		59,883,770			
		w	36,948,893	8		
		s	22,934,877	10		
Northern Europe	4.92%		30,263,058			
		w	30,263,058	8		–738,419
		s	0			
Western Europe	10.82%		66,504,216			
		w	66,504,216	9		
		s	0			
Oceania	4.01%		24,623,834			
		w	0			
		s	24,623,834	1	–859,372	
Russian Federation	9.38%		57,661,512			
		w	4,566,060	9		
		s	53,095,452	9		
India	15.14%		93,090,460			
		w	93,090,460	4	–4,412,488	74,472
		s	0			
Pakistan	4.17%		25,652,858			
		w	25,652,858	6		
		s	0			
USA	9.44%		58,043,684			
		w	46,796,563	9		
		s	11,247,121	10		–123,718
China	20.23%		124,355,184			
		w	49,502,258	7		
		s	74,852,926	9		–1,437,176
Total	100%		614,738,491		–3,889,378	1,541,022

5.2.3 Simulating the effects of ENSO phases

In the benchmark simulation, no changes to the inputs are performed during the whole simulation.

In the second run of the model, we start loading the changes to the production that, according to our model, would be observed in 2014 if the ENSO phase were El Niño instead of neutral in period 37. Note that time tick 37 of the simulation corresponds to January of the fourth year. We let the model progress for 36 time steps to avoid transient effects due to initialisation. As seen in the HM column of Table 5, in January, winter wheat is harvested in South America and spring wheat is harvested in Oceania. However, in time tick 37, South America's production is 234,329 tonnes higher than the previous January and in Oceania 859,372 tonnes lower. These changes start affecting the supplies on these two markets possibly causing a change in buyers' wheat cost ranking and, therefore, movements in their demand schedules in the following months. The next change will take place in April (time tick 40 in the simulation) when El Niño would have caused a significant reduction in India's winter wheat production. This event triggers further adjustments in the international wheat market. In June, July, and August, 'normal' productions are realised by Pakistan, China Eastern, and Northern Europe, while Central Asia would have an increase in its winter production. The simulation then progresses by inputting the 'normal' produced quantities for each geographic unit at the due month. Quantities which progressively enter into the model are those emended with the estimated El Niño effects up to time tick 72.

In the third run, the process repeats, but taking as input the La Niña quantities starting from time tick 37.

5.2.4 Simulating the effects of different levels of demand

A second factor that strongly affects the result is the level of demand compared to the available quantity. It is well known that recording data is a hard task, and it is subject to significant distortion even if it is done very carefully. Difficulties concerning the management of demand components are for example discussed in the FAOSTAT documentation cited in footnote 5.

Faced with these difficulties, we choose to perform the previous analysis concerning the climate effect for different levels of demand. In particular, we generate for each buyer four different levels of demand, scaling the sum of the three principal components supplied by the food balance FAOSTAT database (food, feed, and seed) by factors of 1.13, 1.14, 1.15 and 1.16. Note that FAOSTAT identifies four other minor wheat uses: 'processing', 'waste', 'other uses', and 'stock variation'. This justifies the adoption of a multiplier greater than one in our computation of the total demand.

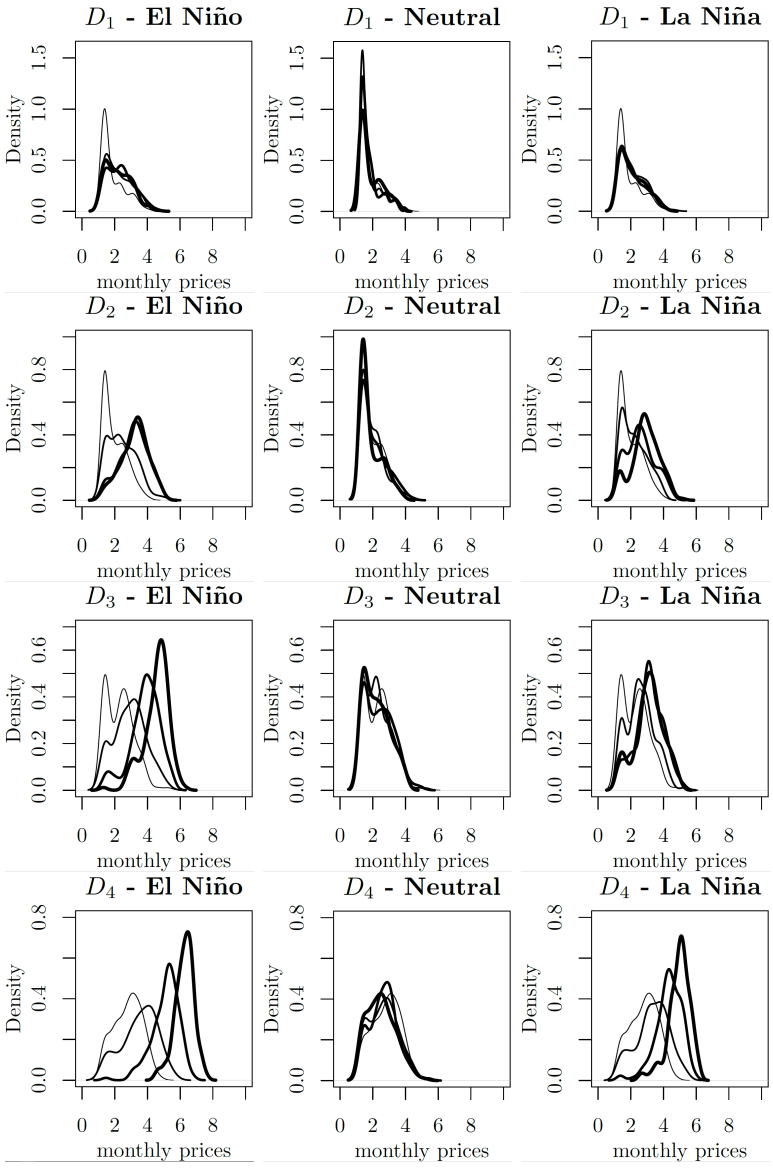
This *what if* approach allows us to compare situations where the total supply of wheat, that in the neutral phase is 614,738,491 tonnes as displayed in Table 5 with the following levels of aggregate demands: D_1 = 604,797,562, D_2 = 610,149,753, D_3 = 615,501,942 and D_4 = 620,854,136.

The comparison of these simulations allows the analysis of the effects of El Niño and La Niña on wheat prices at a different level of market demand.

5.3 Results

In this section, we deal with the effects the ENSO phase would have had on the price distribution of wheat.

Figure 6 Price distribution dynamics



It is useful to recall that the model outputs a price for each production area at each time the market is performed. Therefore, at each time tick, we gather the 12 prices established in the corresponding market sessions. As highlighted above, we run the model on a monthly scale. In a year (12 simulation time ticks) 144 observations are

available. Our analysis consists of evaluating the dynamics of the statistical distribution of these monthly prices. To this aim, we build their kernel density estimation that is displayed in Figure 6.

In this Figure, each chart displays four lines with a different degree of thickness. We will refer to thickness 1 to 4 hereafter. The line with thickness level equal to 1 is the kernel density estimation of the 144 observations obtained from the simulation in time ticks 25–36. That with thickness 2 gives the same for time ticks 37–48 and so on. Comparing these distributions we can monitor the dynamics of the monthly price distribution starting from the thinnest and ending in the thickest.

Figure 6 reports by column the conditional distributions on the three ENSO phases while, by row, distributions are conditionated on the four levels of demand D_1, D_2, D_3 and D_4.

Because for each level of demand, harvested quantities inputted in the simulation change with ENSO phases starting from time tick 37, the three thinnest distributions are the same by row (they relate to time ticks 25–36 where quantities are the same). They, therefore, provide a term of comparison.

A second useful observation concerns the price range. In simulations, prices have a lower boundary represented by production costs. Therefore, a decreasing level of demand would not cause a reduction in minimum observed prices. This explains why density accumulates at a low, but positive value in some chart of Figure 6, especially those with a low level of demand.

It is convenient to start the analysis of the charts in Figure 6 from the results obtained for the neutral phase. In this case, harvested quantities enter the model periodically at the HM identified in Table 5, but, for each producer, they are the same during the whole simulation. Looking at the four charts in the central column of the Figure, we notice that the distribution moves in time, but its shape does not change much. This behaviour signals that the thickness 1 distribution can be taken as a term of comparison for the following distributions when the ENSO phase is switched to El Niño or La Niña, that is from time tick 37 on. Secondly, going down from the top to the bottom chart in the central column of Figure 6 we can see a feature that should not come as a surprise: demand increases create upward pressure on prices. Prices distributions spread to the right as the demand gradually increases from D_1 to D_4.

Some other observations can be done by comparing Figure 6 columns side by side. But two considerations are in order before starting the analysis. First, it is useful to make a consideration of how the sequence of thickness 1–4 distributions can be interpreted in the charts showing El Niño and La Niña cases. As it was highlighted above, the thickness 1 distribution represents the state in the neutral phase. Thickness 2 distribution is the price distribution of the period in which the climate effects are first introduced into the model. As explained above, these changes can be seen as shocks that sequentially perturb the price distribution. From this point of view, this distribution can be seen as representing a transition period. Thickness 3 distribution is that in which all the harvested quantities are constant since at least one production cycle. Therefore, if we want to get rid of transition effects, the comparison has to be done between thicknesses 1 and 3 distributions. Finally, thickness 4 distribution can be used to check if the change process is still ongoing or it has been completed.

Second, to fully appreciate the results, it is useful to highlight ENSO effects on global production. They can be evaluated by looking at the last line in Table 5. According to the estimation based on the analysis in the previous sections, El Niño,

other things being equal, decreases global production by 3,889,378 tonnes, while La Niña increases it by 1,541,022 tonnes. Although these figures should be better qualified in the future (especially for La Niña) because of the 'fragile' estimation of Central Asia effects, we will use them because, as we will show below, they will bring to price behaviour consistent with those observed in reality and discussed in Section 3.

The comparison between the neutral and the El Niño charts (Figure 6 columns 1 and 2) clearly reveals the upward pressure on prices triggered by the reduction of production due to the change in ENSO phase. At the D_1 level of demand, densities move to the right implying a significant increase in dispersion among the monthly prices of the various producers. The distribution is basically unchanged for thickness levels greater than one.

The transitional nature of thickness 2 distribution is evident in the D_2 – El Niño chart. There, thicknesses 1 and 3 distributions are very different, with the second one roughly symmetric and significantly moved to the right. This is also an example in which the adjustment process comes to an end with thickness 3 distribution. In fact, thickness 4 distribution essentially overlaps it. Differently from D_2, the appearance of El Niño when the global demand increases to D_3 triggers a continuous growth of prices. This is represented by the continuous movements of price distribution to the right. Prices growth is even faster the higher the aggregate demand is. The D_4- El Niño chart confirms the result.

Looking at the La Niña charts, i.e., comparing Figure 6 columns 2 and 3, we can observe an unexpected phenomenon: there is a tendency of prices to rise even though the global supply of wheat increases. In other words, despite a downward pressure of prices is expected because of the increase of supply, the changes in prices distributions going through the D_1, D_2, D_3 and D_4 levels of demand are qualitatively the same than those observed after the supply reduction bring about by El Niño. However, the phenomenon is quantitatively less intense. As an example, the D_3 – La Niña distributions behave very similarly to the D_2 – El Niño ones. Similarly, the D_4 – La Niña distributions move like the D_3 – El Niño ones.

This unexpected behaviour points us to take a different perspective which opens a possible way for future research. Looking at La Niña effects in Table 5 we see that the production of some areas is reduced by La Niña. The price dynamics behaviour just commented could point to the fact that their effects on markets prevail on those activated by opposite and quantitatively more relevant effects. In this context, the research can be directed to investigate which market(s) most affect all the others in a particular historical time and global system configuration. As the theory of complex systems teaches us, even small shocks can have macroscopic effects if they hit a component of the system which, in that particular time, is in a critical situation. Distefano et al. (2019) also hint at the international food trade network as a complex system.

6 Conclusions

This paper presents a someway singular work where several heterogeneous elements are carefully orchestrated to travel the broken ground trail going from oceanic surface temperature to wheat prices.

At the beginning of this journey, there is the identification of the geographic units which are relevant to the international wheat markets. The choice is to use

sub-continental areas with some exception in which particular relevant countries are considered individually. The construction of a dataset by aggregating or disaggregating existing data sources with heterogeneous resolution complying with the spatial structure tailored in the previous steps paves the way to the next section of the trail.

In the middle of this traveling, there is the regression model that transforms a so large natural forcing as ENSO into a measurable variable that is directly connected to the ancestral and widespread cultivar of wheat. There are a number of factors affecting wheat production. The climate forcing is perhaps the first to pay attention to. Notwithstanding, its identification is hard. To overcome these difficulties, this work starts from the observation that, other things equal, a 'normal' production follows 'normal' climate conditions. Therefore, the production realised under a neutral phase of ENSO is taken as reference for the analysis. Deviations of production from that observed in the neutral state are measured by means of the one-way robust ANOVA regression coefficients. This change is strictly local-depending and, in fact, the results show that only 10 out of 23 regions/countries that produce wheat is highly correlated with the climate oscillation generated by ENSO. The number of significant effects associated to El Niño, the warm deviation from normality, is almost the same of that associated to La Niña, the cold one, and a similar situation holds for the sign of the effects: 4 (3) negative (positive) with El Niño, 3 (3) negative (positive) with La Niña.

Instead, the last part of the journey is in charge of a computational model that leads straight to another fundamental instrument of the human being's development: the market. The virtual reproduction of the wheat international market realised in the final part of this work by using computer simulations shows how ENSO Phases affect differently the wheat markets and prices according to the level of global demand relative to global supply. In some circumstances, this causes a one-shot movement of the distribution, while in other cases it triggers potentially long-lasting and cumulative effects on price dynamics.

Being based on a long route, this traveling has also many limitations that it is worth discussing here. Moreover, they suggest directions for further research that will hopefully smooth the trail followed in this work.

First comes the need for a more disaggregated dataset. In particular, the availability of data concerning the different crop sowing seasons would have significantly improved the quality of the results obtained in this paper. This option is available for some countries around the world, however, a global dataset that gathers data from country institutional sources, i.e., FAOSTAT is only composed of merged (from winter and spring sowing) time series. Several research groups have devoted their efforts to this end, the GDHY dataset used in this work is an example. However, they often result in some lacunas in the aggregation of quantities due to objective difficulties in collecting such a huge amount of information. It is why we choose to split FAOSTAT production time series into winter and spring ones by using a simple 'weighting'. Additional work in improving this splitting procedure is planned for the future. Furthermore, by aggregating harvested area and production per region to compute regional yield, we *de facto* assume that the technique adopted in each region is the same for all the countries in that region. This assumption could conceivably be true if we accept that the spatial proximity generates a similar 'modernity' evolution.

Beyond these difficulties, this work has several potential applications both in the short and long run. A possible short-term application of this work is the possibility of implementing the computational model in order to predict the wheat prices for the next

season at the regional level, because of the good predictability of ENSO mechanism. On the long run side, what discussed above allows us to glimpse the potential that entices the scholar to complete the model in various directions. On the supply side, decisions on land use, the effects of climate on plant diseases, or the management of the stocks could be modelled. On the demand side, attention could be focused on the determinants of its components. If successful, this will provide in the future a reliable tool to find out policies to enhance food security via better control of main staple average and volatility prices.

Acknowledgements

This research was partially supported by H2020 Project MED-GOLD 'Turning climate-related information into added value for traditional MEDiterranean Grape, OLive and Durum wheat food systems' (https://www.med-gold.eu/) grant number EU-776467. The authors would like to thank the organisers of IWCEE 2019 for the given opportunity of managing a special session and the two anonymous reviewers for their useful suggestions. One of the author would also like to thank the Administration of Fardella, Potenza, Italy for having arranged a smart working place from where this paper has been partially written.

References

Algieri, B. (2014) 'A roller coaster ride: an empirical investigation of the main drivers of the international wheat price', *Agricultural Economics*, July, Vol. 45, No. 4, pp.459–475.

Berk, R.A. (2016) 'Splines, smoothers, and kernels', in in Berk, R.A. (Ed.): *Statistical Learning from a Regression Perspective, Springer Texts in Statistics*, pp.55–127, Springer International Publishing, Cham.

Distefano, T., Chiarotti, G., Laio, F. and Ridolfi, L. (2019) 'Spatial distribution of the international food prices: unexpected heterogeneity and randomness', *Ecological Economics*, Vol. 159, pp.122–132.

Epstein, J.M. (2006) *Generative Social Science: Studies in Agent-Based Computational Modeling*, Princeton University Press, Princeton, Oxford.

Giulioni, G., Di Giuseppe, E., Toscano, P., Miglietta, F. and Pasqui, M. (2019) 'A novel computational model of the wheat global market with an application to the 2010 Russian Federation Case', *Journal of Artificial Societies and Social Simulation*, Vol. 22, No. 3, p.4.

Gutierrez, L. (2017) 'Impacts of El Niño-Southern Oscillation on the wheat market: a global dynamic analysis', *PLOS ONE*, June, Vol. 12, No. 6, p.e0179086.

Hampel, R.F. (1968) *Contributions to the Theory of Robust Estimation*, PhD thesis, University of California, Berkeley.

Headey, D. and Fan, S. (2008) 'Anatomy of a crisis: the causes and consequences of surging food prices', *Agricultural Economics*, November, Vol. 39, No. S1, pp.375–391.

Heino, M., Puma, M.J., Ward, P.J., Gerten, D., Heck, V., Siebert, S. and Kummu, M. (2018) 'Two-thirds of global cropland area impacted by climate oscillations', *Nature Communications*, December, Vol. 9, No. 1, p.1257.

Hijmans, R.J. (2019) *Raster: Geographic Data Analysis and Modeling*, R package version 3.4-5 [online] https://CRAN.R-project.org/package=raster.

Hu, Z-Z., Kumar, A., Jha, B. and Huang, B. (2020) 'How much of monthly mean precipitation variability over global land is associated with SST anomalies?', *Climate Dynamics*, January, Vol. 54, Nos. 1–2, pp.701–712.

Huber, P.J. (1964) 'Robust estimation of a location parameter', *Annals of Mathematical Statistics*, March, Vol. 35, No. 1, pp.73–101.

Iizumi, T. and Sakai, T. (2020) 'The global dataset of historical yields for major crops 1981–2016', *Scientific Data*, March, Vol. 7, No. 1, pp.1–7.

Iizumi, T., Luo, J-J., Challinor, A.J., Sakurai, G., Yokozawa, M., Sakuma, H., Brown, M.E. and Yamagata, T. (2014a) 'Impacts of El Niño Southern Oscillation on the global yields of major crops', *Nat. Commun.*, Vol. 5, Article No. 3712, pp.1–7.

Iizumi, T., Yokozawa, M., Sakurai, G., Travasso, M.I., Romanenkov, V., Oettli, P., Newby, T., Ishigooka, Y. and Furuya, J. (2014b) 'Historical changes in global yields: major cereal and legume crops from 1982 to 2006', *Global Ecology and Biogeography*, March, Vol. 23, No. 3, pp.346–357.

Iizumi, T., Shin, Y., Kim, W., Kim, M. and Choi, J. (2018) 'Global crop yield forecasting using seasonal climate information from a multi-model ensemble', *Climate Services*, Vol. 11, pp.13–23.

Khan, M.A., Tahir, A., Khurshid, N., Ul Husnain, M.I., Ahmed, M. and Boughanmi, H. (2020) 'Economic effects of climate change-induced loss of agricultural production by 2050: a case study of Pakistan', *Sustainability*, February, Vol. 12, No. 3, p.1216.

Khoury, C.K., Bjorkman, A.D., Dempewolf, H., Ramirez-Villegas, J., Guarino, L., Jarvis, A., Rieseberg, L.H. and Struik, P.C. (2014) 'Increasing homogeneity in global food supplies and the implications for food security', *Proceedings of the National Academy of Sciences*, Vol. 111, No. 11, pp.4001–4006.

Mair, P. and Wilcox, R. (2019) 'Robust statistical methods in R using the WRS2 package', *Behavior Research Methods*, April, Vol. 52, pp.464–488 [online] https://doi.org/10.3758/s13428-019-01246-w.

Monfreda, C., Ramankutty, N. and Foley, J.A. (2008) 'Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000', *Global Biogeochemical Cycles*, Vol. 22, No. 1, pp.1–19.

Ott, H. (2014) 'Volatility in cereal prices: intra-versus inter-annual volatility', *Journal of Agricultural Economics*, Vol. 65, No. 3, pp.557–578.

R Core Team (2020) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria [online] https://www.R-project.org/.

Rojas, O.E., Li, Y. and Cumani, R. (2014) *Understanding the Drought Impact of El Niño on the Global Agricultural Areas: An Assessment Using FAO's Agricultural Stress Index (ASI)*, Food and Agriculture Organization of the United Nations.

Sacks, W.J., Deryng, D., Foley, J.A. and Ramankutty, N. (2010) 'Crop planting dates: an analysis of global patterns: global crop planting dates', *Global Ecology and Biogeography*, June, Vol. 19, No. 5, pp.607–620.

Siad, S.M., Gioia, A., Hoogenboom, G., Iacobellis, V., Novelli, A., Tarantino, E. and Zdruli, P. (2017) 'Durum wheat cover analysis in the scope of policy and market price changes: a case study in Southern Italy', *Agriculture*, February, Vol. 7, No. 2, p.12.

Storn, R. and Price, K. (1997) 'Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces', *Journal of Global Optimization*, December, Vol. 11, No. 4, pp.341–359.

Tukey, J.W. (1959) *A Survey of Sampling from Contaminated Distributions*, STRG Technical Report No. 33, Princeton University, Princeton, New Jersey.

Ubilava, D. (2017) 'The ENSO effect and asymmetries in wheat price dynamics', *World Development*, August, Vol. 96, No. 1, pp.490–502.

Ubilava, D. (2018) 'The role of El Niño Southern Oscillation in commodity price movement and predictability', *American Journal of Agricultural Economics*, January, Vol. 100, pp.239–263.

Welch, B.L. (1951) 'On the comparison of several mean values: an alternative approach', *Biometrika*, Vol. 38, Nos. 3/4, pp.330–336.

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Milton Bache, S., Müller, K., Ooms, J., Robinson, D., Paige Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K. and Yutani, H. (2019) 'Welcome to the tidyverse', *Journal of Open Source Software*, Vol. 4, No. 43, p.1686 [online] https://doi.org/10.21105/joss.01686.

Wilcox, R. (2012) 'Chapter 7 – one-way and higher designs for independent groups', in Wilcox, R. (Ed.): *Introduction to Robust Estimation and Hypothesis Testing, Statistical Modeling and Decision Science*, 3rd ed., pp.291–377, Academic Press, Boston.

Wright, B.D. (2011) 'The economics of grain price volatility', *Applied Economic Perspectives and Policy*, Vol. 33, No. 1, pp.32–58.

Appenix A

Description of datasets

Wheat supply database

The time series of production, yield, and harvested area are taken from the 'production-crops' section of the FAOSTAT database. Because of their importance as wheat producers, it is worth detailing the time series of China and the Russian Federation. For the former, Taiwan, Macao and Hong Kong are part of China since 2014, however, we only refer to 'China, Mainland' because of historical homogeneity. For the latter, it is necessary to take into account the significant modification that interested the *Russian territory* since 1991. In fact, from the pre-existent USSR, they have been created several nations: '*Russian Federation*', '*Ukraine*', '*Kazakhstan*', '*Kyrgyzstan*', '*Uzbekistan*', '*Turkmenistan*', '*Tajikistan*', '*Azerbaijan*', '*Armenia*', and '*Belarus*'. Because of this historic happening, the wheat time series of the Russian Federation, as well as those of Central Asia and Eastern Europe, start only from 1992 instead of 1961.

Winter and spring yield dataset

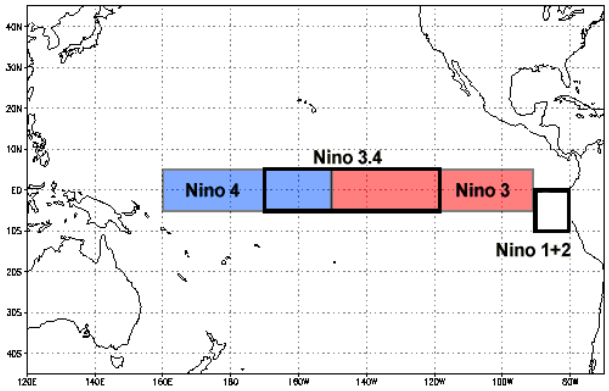
It is important for the goal of this study to know the gathering month since the model in Section 5 takes into account the moment at which the product is ready to be sold. Therefore, the distinction between winter and spring sowing becomes essential. It is also important to know those areas where the double wheat rotation is practiced. Those regions/countries that practice the double rotation have two harvesting periods during the same year, and, consequently, they supply the market two times a year. Furthermore, the climate effects might be significant for one cropping season but not for the other.

For all these reasons, we need a dataset of wheat yield that is able to distinguish between the winter and spring crop sowing. Among others, our choice has fallen on the GDHY dataset from Iizumi and Sakai (2020). This is the updated version of the same dataset described in Iizumi et al. (2014b). As hinted above, GDHY is a yield

gridded dataset of several crops, and wheat is one of them. Its grid size is 0.5° by 0.5° in longitude and latitude while the period is of 36 years (from 1981 to 2016). In GDHY, the winter and spring yield values were obtained from the FAOSTAT yearly time series by integrating them with eco-physiological information from several sources such as remotely sensed leaf area index, the fraction of photosynthetically active radiation, reanalysis of solar radiation, harvested area map, and crop calendars.

We extract from this dataset the fraction of winter and spring production and winter and spring harvested area then use them to split the correspondent FAOSTAT time series. The procedure is explained in Appendix B.

Figure 7 (a) Niño regions: the portion of the Pacific Ocean where the various ENSO indexes are calculated (b) Time series of ENSO phases calculated by means of ONI from 2010 to 2017 (see online version for colours)



(a)

Year	DJF	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ
2010	1.5	1.3	0.9	0.4	-0.1	-0.6	-1.0	-1.4	-1.6	-1.7	-1.7	-1.6
2011	-1.4	-1.1	-0.8	-0.6	-0.5	-0.4	-0.5	-0.7	-0.9	-1.1	-1.1	-1.0
2012	-0.8	-0.6	-0.5	-0.4	-0.2	0.1	0.3	0.3	0.3	0.2	0.0	-0.2
2013	-0.4	-0.3	-0.2	-0.2	-0.3	-0.3	-0.4	-0.4	-0.3	-0.2	-0.2	-0.3
2014	-0.4	-0.4	-0.2	0.1	0.3	0.2	0.1	0.0	0.2	0.4	0.6	0.7
2015	0.6	0.6	0.6	0.8	1.0	1.2	1.5	1.8	2.1	2.4	2.5	2.6
2016	2.5	2.2	1.7	1.0	0.5	0.0	-0.3	-0.6	-0.7	-0.7	-0.7	-0.6
2017	-0.3	-0.1	0.1	0.3	0.4	0.4	0.2	-0.1	-0.4	-0.7	-0.9	-1.0

(b)

Notes: Red for El Niño, blue for La Niña and white for neutral.

Source: Credits to https://www.nws.noaa.gov/credits.php

Harvested area devoted to winter and spring wheat cultivation and crop calendars

To set the wheat growing season in each geographic unit, we use the GCPD dataset released by Sacks et al. (2010) that is composed of data on crop planting and harvesting dates elaborated from multiple institutional sources. In particular, the authors have obtained the crop calendars in each cell grid by examining the relationship between planting dates and temperature, precipitation, and potential evapotranspiration using

30-year average climatologies from the Climatic Research Unit, University of East Anglia. The GCPD dataset is based in turn on the spatial patterns built by Monfreda et al. (2008). We use this dataset also to compensate for some lacks in GDHY dataset.

ENSO index

To represent ENSO variability and to account for its effects, we use an oceanic index, the Oceanic Ninõ Index (ONI), which is a standard metric to represent ENSO phases. Following the adopted definition at the National Oceanic and Atmospheric Administration – Climatic Prediction Center[6], the ONI index is computed by averaging sea surface temperature anomalies in an area of the east-central equatorial Pacific Ocean (5° S to 5° N; 170° W to 120° W), which is called the Niño-3.4 region [see Figure 7(a)]. Finally, the ONI provides a classification of the year according to three phases of ENSO, El Niño, La Niña, and neutral whenever it results in 5 consecutive quarters belonging to the same phase. An example of such a classification is shown in Figure 7(b). Recall that the averaging quarters are identified with a three-letter acronym made of the months' initials. In this study, we use ONI to model the impact of ENSO variability on wheat global production.

Wheat demand database

For the analysis of worldwide wheat consumption, we use 'wheat and products' definition from the FAOSTAT database. The default composition of wheat and products is wheat, flour, bran, macaroni, germ, bread, bulgur, pastry, starch, gluten, cereals breakfast, wafers, mixes and doughs, food preparations, flour, malt extract. These data are available in the two sections of FAOSTAT database:

1 commodity balances

2 food balance sheets.

These two databases have some differences that are described in the FAOSTAT document retrievable at http://fenixservices.fao.org/faostat/static/documents/FBS/KeydifferencesbetweennewandoldFoodBalanceSheet.pdf. However, the various components of the demand maintain their specificity in both of them such that their merging is not inappropriate provided that one check which of the definitions are changed. For example, the US time series of waste in 2 is composed of missing values. The components that result from the merging of the two databases are food, feed, seed, processing, waste, other uses, and stock variation.

Appendix B

B1 Building on FAOSTAT time series of production, yield, and harvested area: from merged to winter and spring ones.

While wheat yield data at the national level are reported by the FAOSTAT, similar data for winter and spring crop sowing are rarely available at that spatial scale. To satisfy

the necessity to have a better correlation between wheat yield and ENSO, we obtain the winter and spring yield time series by splitting the FAOSTAT averaged ones through the information gathered from other gridded datasets, in particular the GDHY and GCPD produced by Iizumi and Sakai (2020) and Sacks et al. (2010), respectively. However, the total yield of a crop, when considering all production systems, is not the sum of the individual yields, but rather the weighted average of their yields. On the other contrary, total harvested area and production are exactly the sums of individual harvested area and production. Thus, we start by splitting total harvested area and production into spring and winter, then we compute the corresponding spring and winter yields. More in details, let us denote the aggregated time series available from the FAOSTAT as p_{FAO} = total wheat production; h_{FAO} = harvested area, and y_{FAO} = yield, then

$$p_{FAO} = h^w y^w + h^s y^s \tag{3}$$

where the superscripts w and s indicate the winter and spring sowing, respectively. Here and in the following paragraphs, the unit of measure of these quantities are: *tonnes* for production, *hectares (ha)* for harvested area, and *hectogram/hectare (hg/ha)* for yield. To compute the unknown quantities h^w, y^w, h^s, and y^s of equation (3) at the geographic unit level, we aggregate over each of them the GDHY values of production and harvested area calculated at gridded cells. To exemplify, we report the maps of China winter and spring yield obtained from GDHY. The map in Figure 8(a) displays the areas of China in which the winter sowing is managed. Colours inform on the yields obtained. Figure 8(b) maps the same for the spring sowing, and the map in Figure 8(c) is the union of the two previously described maps. It is important to notice that in some circumstances the same land can be used twice implementing both the winter and the spring sowing. This clarification will be useful in the following more formal treatment.

The aggregated values of each geographic unit are:

1 $y^w_{GDHY} = \sum_{c=1}^{n^w} y^w_c / n^w$ winter yield averaged over the harvested n^w cells *c* in the geographic unit

2 y^s_{GDHY} the same but spring yield

3 $h^w_{GDHY} = \sum_{c=1}^{n^w} c^w \, Area(c^w)$ harvested area of winter production

4 h^s_{GDHY} = the same but spring harvested area

5 $h^{w+s}_{GDHY} = h^w_{GDHY} + h^s_{GDHY}$ total harvested area.

Note that we could have computed the aggregated quantity of products directly from the yield of each cell grid, however, because of the large variability of harvest area across years and lacunas of yield data in some regions and by following the GDHY authors' suggestion [see usage notes in Iizumi and Sakai (2020)], we prefer to reduce the spatial variability by computing the geographic unit averaged yield. From these quantities, we compute the share of winter and spring production:

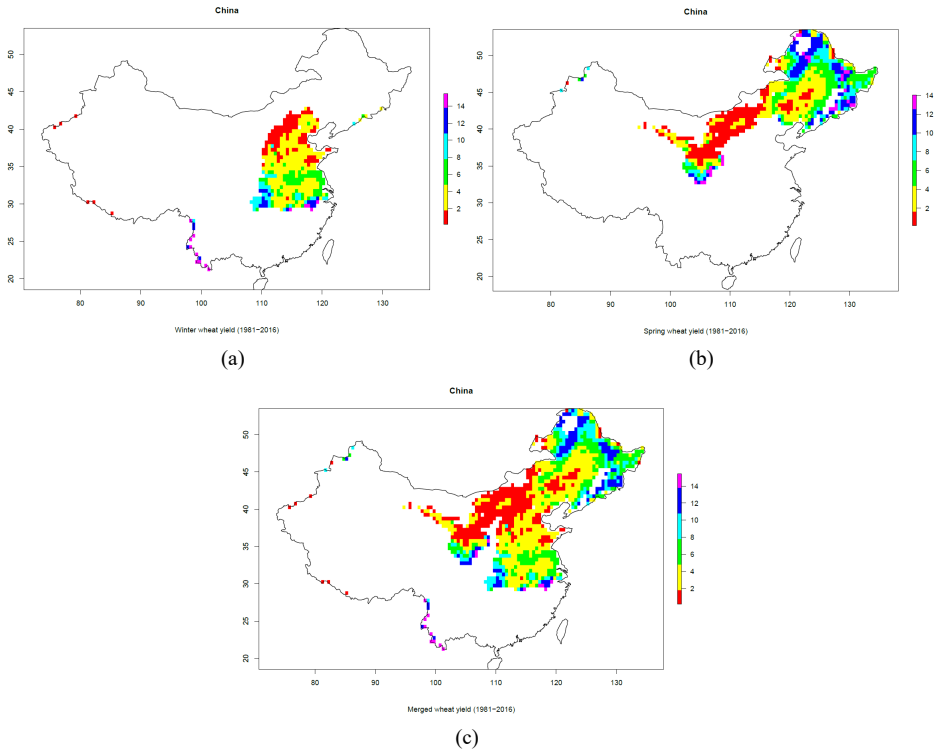
$$\beta^w = \frac{h^w_{GDHY} y^w_{GDHY}}{h^w_{GDHY} y^w_{GDHY} + h^s_{GDHY} y^s_{GDHY}} \quad \beta^s = 1 - \beta^w \tag{4}$$

and the correspondent share of harvested area devoted to winter and spring sowing:

$$\gamma^w = \frac{h^w_{GDHY}}{h^{w+s}_{GDHY}} \quad \gamma^s = \frac{h^s_{GDHY}}{h^{w+s}_{GDHY}} \tag{5}$$

where $1 \leq \gamma^w + \gamma^s \leq 2$ because of the presence of areas that are sowed both in winter and spring.

Figure 8 Spatial distribution of China wheat yield according to Iizumi and Sakai (2020), (a) winter (b) spring (c) merged (see online version for colours)



(a) (b)

(c)

By applying the coefficients β's and γ's of equations (4) and (5) to the FAOSTAT time series, we obtain the split time series of production and harvested area for winter and spring: $p^w_{FAO} = \beta^w p_{FAO}$, $p^s_{FAO} = \beta^s p_{FAO}$, $h^w_{FAO} = \gamma^w h_{FAO}$, and $h^s_{FAO} = \gamma^s h_{FAO}$. Finally, we calculate the FAO yield time series split into winter and spring as follows:

$$y^w_{FAO} = \frac{p^w_{FAO}}{h^w_{FAO}} \quad y^s_{FAO} = \frac{p^s_{FAO}}{h^s_{FAO}} \tag{6}$$

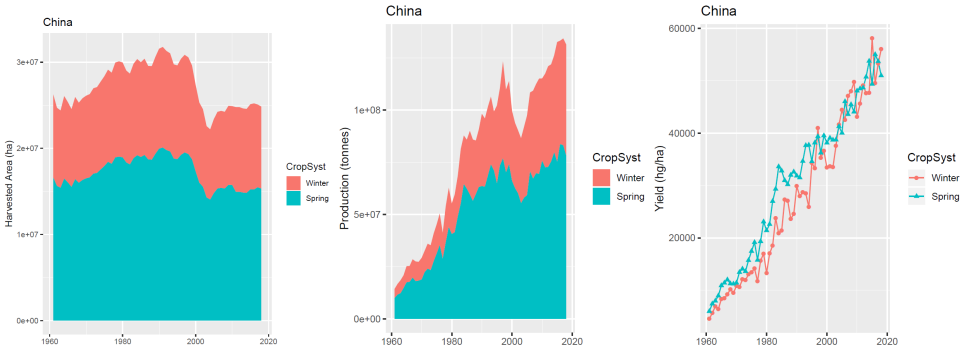
To avoid misspecification by taking into account very small quantities of production for one of the two cropping systems, we make a further check on the $\beta-$weights. In fact, we attribute to a geographic unit a single type of crop sowing if the other has a very small production (lesser or equal to 2%) with respect to the other. Furthermore, when the aggregation that is done by using GDHY dataset gives strange results for any of the geographic unit[7], we directly use the share of the harvested area found in GCPD, that, although referred to a unique year, it let us avoid mistakes in the splitting procedure. The list of deleted systems resulting from such a double-check procedure is given in Table 6.

Finally, the procedure described in this paragraph let us obtain the winter and spring time series of production, harvested area, and yield from 1982 to 2016, that is the time window of GDHY. However, the FAOSTAT dataset covers the period 1961–2018, then we extend the values of β's and γ's to the remaining years by assuming that the mean value of the β's and γ's calculated in the time windows 1982–1991 and 1997–2016 remain fixed for the periods 1961–1981 and 2017–2018, respectively. Using again the example for China, the procedure results in the estimates of the winter and spring time series of harvested area, production, and yield shown in Figure 9.

Table 6 Crop system deleted because of very small production (<2% of the other system)

Country-group	*Notes*
Middle Africa	Production in spring-deleted
Western Africa	Production in spring-deleted
Southern Africa	Production in spring-deleted
Central America	Production in spring-deleted
Southern Asia	Production in spring-deleted
South-Eastern Asia	Production in spring-deleted
Northern Europe	Production in spring-deleted
Western Europe	Production in spring-deleted
Oceania	Production in winter-deleted
India	Production in spring-deleted
Pakistan	Production in spring-deleted

Figure 9 China estimated time series of winter and spring harvested, (a) area (b) production (c) yield from 1961 to 2018 (see online version for colours)



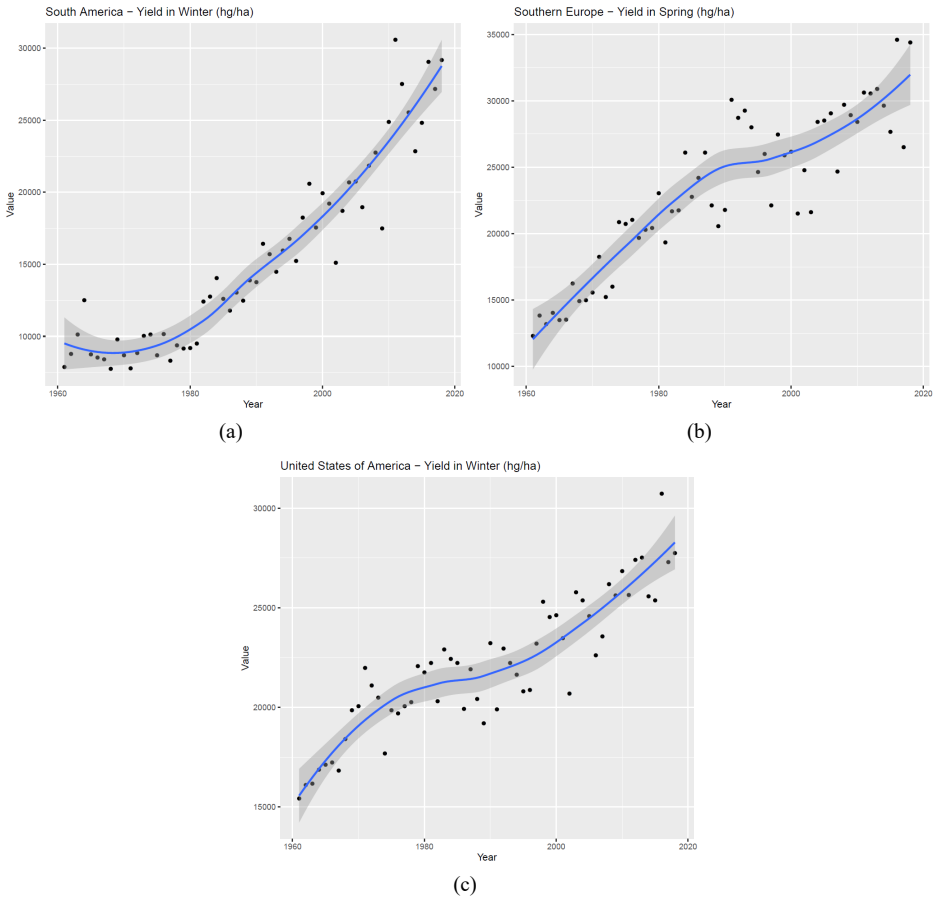
(a) (b) (c)

B2 Yield de-trending

Yield is the quantity of wheat produced per each harvested hectare. When considered in a specific country, it is somehow a measure of the technical capability of that country in producing wheat. Like production, the yield is also affected by an increasing trend. This appears evident for the case of China represented in Figure 9(c), nevertheless, this is true also for most of the other geographical units (not shown). For this reason, it is

straightforward to apply a de-trending procedure before setting any regression model. To this end, we estimate the yield trend by using a nonparametric regression (loess). The loess procedure is widely used. Among many, the reader can refer to the book of Berk (2016) for details. It is based on the idea that for any given value of the predictor (time in our case) a polynomial regression is constructed only from its nearest neighbours' observations. In practice, we estimate a polynomial surface determined by the local fitting of yield on time. To minimise the trade-off between the residual sum of squares and the number of parameters, we setup a time-span equal to 0.75 that is the default option of the *loess* function in the basic *stats* package of R software (R Core Team, 2020). As an example, a visual representation of the loess trend estimate for South America, Southern Europe, and the US is displayed in Figure 10.

Figure 10 Wheat yield time series of, (a) South America (winter) (b) Southern Europe (spring) (c) USA (winter) from 1961 to 2018 and the corresponding estimated trend (see online version for colours)



From this estimate, we compute the de-trended time series of yield, i.e., the yearly series of residuals from local polynomial regression. Finally, we calculate the de-trended yield

percentage by dividing the residuals of local polynomial regression by the corresponding fitted values. The time series of de-trended yield percentage is basically the time series of standardised deviations from its trend and constitutes the dependent variable of the regression analysis on ENSO.

Appendix C

Detailed composition of geographic areas

Table 7 List of countries belonging to each region considered in the model and listed in Figure 1

Africa	Eastern	British Indian Ocean Territory, Burundi, Comoros, Djibouti, Eritrea, Ethiopia, French Southern Territories, Kenya, Madagascar, Malawi, Mauritius, Mayotte, Mozambique, Réunion, Rwanda, Seychelles, Somalia, South Sudan, Uganda, United Republic of Tanzania, Zambia, Zimbabwe
	Middle	Angola, Cameroon, Central African Republic, Chad, Congo, Democratic Republic of the Congo, Equatorial Guinea, Gabon, Sao Tome and Principe
	Northern	Algeria, Egypt, Libya, Morocco, Sudan, Tunisia, Western Sahara
	Southern	Botswana, Eswatini, Lesotho, Namibia, South Africa
	Western	Benin, Burkina Faso, Cabo Verde, Côte d'Ivoire, Gambia, Ghana, Guinea, Guinea-Bissau, Liberia, Mali, Mauritania, Niger, Nigeria, Saint Helena, Senegal, Sierra Leone, Togo
USA	Caribbean	Bahamas, Barbados, Cuba, Others 25 (Carribean)
	Central	Belize, Costa Rica, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama
	Northern	Bermuda, Canada, Greenland, Saint Pierre and Miquelon, USA
	South	Argentina, Bolivia (Plurinational State of), Bouvet Island, Brazil, Chile, Colombia, Ecuador, Falkland Islands (Malvinas), French Guiana, Guyana, Paraguay, Peru, South Georgia and the South Sandwich Islands, Suriname, Uruguay, Venezuela (Bolivarian Republic of)
Asia	Central	Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, Uzbekistan
	Eastern	China, China, Hong Kong Special Administrative Region, China, Macao Special Administrative Region, Democratic People's Republic of Korea, Japan, Mongolia, Republic of Korea
	South-Eastern	Brunei Darussalam, Cambodia, Indonesia, Lao People's Democratic Republic, Malaysia, Myanmar, Philippines, Singapore, Thailand, Timor-Leste, Vietnam

Table 7 List of countries belonging to each region considered in the model and listed in Figure 1

Asia	Southern	Afghanistan, Bangladesh, Bhutan, India, Iran (Islamic Republic of), Maldives, Nepal, Pakistan, Sri Lanka
	Western	Armenia, Azerbaijan, Bahrain, Cyprus, Georgia, Iraq, Israel, Jordan, Kuwait, Lebanon, Oman, Qatar, Saudi Arabia, State of Palestine, Syrian Arab Republic, Turkey, United Arab Emirates, Yemen
Europe	Eastern	Belarus, Bulgaria, Czechia, Hungary, Poland, Republic of Moldova, Romania, Russian Federation, Slovakia, Ukraine
	Northern	Åland Islands, Guernsey, Jersey, Sark, Denmark, Estonia, Faroe Islands, Finland, Iceland, Ireland, Isle of Man, Latvia, Lithuania, Norway, Svalbard and Jan Mayen Islands, Sweden, UK of Great Britain and Northern Ireland
	Southern	Albania, Andorra, Bosnia and Herzegovina, Croatia, Gibraltar, Greece, Holy See, Italy, Malta, Montenegro, North Macedonia, Portugal, San Marino, Serbia, Slovenia, Spain
	Western	Austria, Belgium, France, Germany, Liechtenstein, Luxembourg, Monaco, Netherlands, Switzerland
Oceania		Australia, New Zealand, Papua New Guinea, others 26 (Oceania)

Notes

1 For the interested reader, more details are available at https://iridl.ldeo.columbia.edu/maproom/ENSO/ENSO_Info.html.

2 During El Niño there is a likely increase of droughts in the Tropics, while during La Niña such increasing is more pronounced in the Mid-Latitudes. See the note by the International Research Institute for Climate and Society: https://iri.columbia.edu/news/eight-misconceptions-about-el-nino/.

3 Since it has no production, the analysis by the supply side excludes the Caribbean area resulting in 23 geographical units instead of 24.

4 Cold & Warm Episodes by Season: https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php.

5 http://fenixservices.fao.org/faostat/static/documents/FBS/KeydifferencesbetweennewandoldFoodBalanceSheet.pdf.

6 ONI (version 5) definition and historical data are available from https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php.

7 These limitations are reported by the GDHY authors since they warn to consider that ‘yields in some locations are lacking’ [see usage notes in Iizumi and Sakai (2020)].