Agricultural and Forest Meteorology 253–254 (2018) 15–30

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Assessment of the agro-climatic indices to improve crop yield forecasting

Jordane A. Mathieu[a,*], Filipe Aires[a,b]

[a] Sorbonne Université, Observatory of Paris, PSL Research University, CNRS, LERMA, 75014 Paris, France
[b] Estellus, Paris, France

ARTICLE INFO

Keywords:
Weather impact model
Statistical model
Corn yield
Agro-climate indices
Agriculture

ABSTRACT

Weather has a major impact on agriculture. Statistical models have been used to estimate or forecast crop yield from weather information. In this paper, a general statistical framework is developed in order to rank and quantify the information content of weather information. The methodology is tested over the US, for corn yield. The weather sensitivity of different corn production areas is first analyzed. More than fifty agro-climatic indices have been compared. The study shows that variability in yield is more likely to be due to weather variability in medium and low production areas. The two best yield predictors are the temperature and then the Standardized Precipitation–Evapotranspiration Index (SPEI) in July. While corn in the north Eastern regions is not significantly affected by weather variation, on the East coast, the weather-based mixed model is able to explain about 32% (corr = 0.57) of the yield variability, but in some particularly weather-sensitive states, such as Virginia, this number can reach 64% (corr = 0.80). Two applications are tested: yield estimation at the end of the year for monitoring purpose, and seasonal prediction. It is shown that some agro-climatic indices can significantly improve crop yield modeling when compared to simpler direct weather information (average correlation increase of 0.12 at the US scale).

1. Introduction

How historical, current and future climate affects socio-economic activities has received tremendous attention in areas such as agriculture (Thompson, 1988; Kotlowski, 2007; Kandiannan et al., 2002), energy (Adams et al., 1998; Kaylen and Koroma, 1991), logistics in sales, or tourism (Jewson and Brix, 2005). In agriculture, crop yield is strongly influenced by several factors (e.g., genetics, soil properties, irrigation) but weather is the major uncontrollable factor influencing the development of crops (Taylor and Carlson, 1997). The accurate prediction of crop yields over large areas is critical for the national food supply, prices, farmer plans, or irrigation (Liang, 2004), so quantifying the impact of weather on crop yield is essential.

Many studies have shown the importance of weather information to explain crop yield. Three major modeling approaches have been used to study this relationship: biophysical or crop simulation models (Hoogenboom, 2000; Jones et al., 2003; Kotlowski, 2007); empirical regression models (Kandiannan et al., 2002; Thompson, 1988; Tannura et al., 2008; Lobell et al., 2007), and functional models (Basso et al., 2013; Chipanshi et al., 2015) which are a simplification and/or a combination of the two. In this paper, we focus on empirical regression models: they are less process-oriented than crop models, but they are more data-driven and they require less auxiliary information such as soil properties. They are calibrated on historical data with as little a priori information as possible.

Some agriculture characteristics are sometimes used (e.g., soil properties, grain size) (NASS, 2016; Asseng et al., 2013) as inputs to the crop yield statistical models. In order to provide the forecasting model information more directly related to crop yield, some agro-climatic indices (i.e., weather-based indices) have been considered as inputs. Agro-climatic indices, especially growing degree-days or some type of water stress indices (e.g., Precipitation-Evaporation), have long been used for crop yield predictions (Lass et al., 1993; Holzkamper et al., 2011). The use of direct weather data became common only in recent decades, when gridded weather data became a major input for regional yield forecasting (Zhang and Huang, 2011; Moreto and Souza, 2015). Traditionally, agro-climatic indices are obtained from direct weather data in order to better represent the link between weather and crop growth and to facilitate decision making in agriculture (Lepage et al., 2012; Caubel et al., 2015). For instance, Lass et al. (1993), Robertson et al. (2013) or Bornn and Zidek (2012) use Growing Degree Days (GDD); Butler and Huybers (2013) use Growing Degree Days (GDD) and Killing Degree Days; Moreto and Souza (2015), Zhang and Huang (2011) use evapotranspiration; and Moral García et al. (2014) uses various agro-climatic indices to assess the wine suitability in a region.

Some studies relate the changes of agro-climatic indices trend to

* Corresponding author.
E-mail address: jordane.mathieu@obspm.fr (J.A. Mathieu).

https://doi.org/10.1016/j.agrformet.2018.01.031
Received 4 July 2017; Received in revised form 19 January 2018; Accepted 22 January 2018
0168-1923/

changes in crop phenology in Eurasia regions (Menzel et al., 2003; Moonen et al., 2002; Genovese et al., 2005; Semenov et al., 2006), or in Sahelian countries (Ben Mohamed et al., 2002), or in the major part of North America (Robeson, 2002; Feng and Hu, 2004; Terando et al., 2011) (the last, with frost days, thermal time, and heat stress index). Torvanger et al. (2004) analyzed the important relationship between yields of potatoes/barley/oats/wheat, and temperature (growing degree days)/precipitation in Norway.

Others studies examined the dependence of agriculture to the weather variability analyzing the time series of several weather indices (Bazgeer et al., 2014). Qian et al. (2010) analyzed a set of agro-climatic indices representing Canadian climatic conditions for field crop production. Qian et al. (2001) used precipitation, maximum temperature, minimum temperature and solar radiation. Bélanger et al. (2000) used water deficit (precipitation–evapotranspiration), effective growing degree-days, and Corn Heat Units (CHU) in order to determined the weather impact on agriculture in Quebec. These indices are considered as the most used weather variables to determine how well suited is a land for cereals growing, according to the Agronomic Interpretation Working Group (1995). Graczyk and Kundzewicz (2016) examined the lengths of the growing season and the frost-free season, the days of occurrence of the last spring frost and first autumn frost, and the annual sums of growing degree-days for three values of temperature threshold in Poland.

Only a few studies (Trnka et al., 2011; Lalic et al., 2013; Peltonen-Sainio et al., 2010) tested more than 4 or 5 agro-climatic indices (such as the Free Frost Period, the Corn Heat Units, the Standardized Precipitation–Evapotranspiration Index, or the soil moisture), in combination to simple weather variables (temperature and precipitation), in a crop yield forecasting application. Here, we propose to analyze more than 50 different agro-climatic indices.

Our application focuses on corn yield over the eastern US. The goal of this paper is: **(1)** to propose an approach to identify the most important agro-climatic indices for crop yield forecasting, **(2)** to identify which US regions are the most weather-sensitive for crop yield, and most importantly, **(3)** to measure the impact of using agro-climatic indices instead of direct weather information for crop yield forecasting. Section 2 presents the materials and methods. The sensitivity analysis is presented in Section 3 and the yield forecast improvements in Section 4.

2. Materials and methods

2.1. Data

2.1.1. Agricultural data

The corn yield data were collected by the US Department of Agriculture NASS (National Agricultural Statistics Service) at the county level. A long historical record is available from 1910 to 2013 but the yield time-series are not complete for all counties. The eastern US produces more corn than the western part because of a more favorable climate and topography. The Corn-Belt region (mainly Iowa, eastern Nebraska, Minnesota, Illinois, and Indiana) is known to be the biggest corn-producing region in the US. In 2016, these five states produced more than 60% of the US corn production (18% for Iowa, 11% for Nebraska, 10% for Minnesota, 15% for Illinois, and 6% for Indiana).

2.1.2. The main corn production basins in the US

Corn ranks first in the US grain production. With more than 38 Mha sown in 2016, it is well ahead of wheat (20 Mha) and soybeans (33 Mha). Maize is mainly grown in 5 major production areas and these different zones produce maize under very variable conditions (Fig. 1):

- Zone 1 (Corn Belt): with a humid continental climate, is located in the center of the United States, on rich and deep soils. The area is particularly well suited to maize, which is mostly grown in rotation with soybeans. Since rainfall is abundant, irrigation is not necessary, and dry yields are the highest in the world: 12 t/ha in 2016. Annual precipitation decreases from over 100 cm in the East to less than 60 cm in the West. The amount of rainfall during the five-month growing season is relatively constant (between 43 and 48 cm) (Neild and Newman, 1990). Thus, only 12% of the maize area is irrigated. In 2016, the Corn Belt produced alone 60% of the US maize production.
- Zone 2: in the Northern Great Plains (North and South Dakota), non-irrigated corn, in rotation with soybean, competes with HRS (Hard Red Spring) wheat. The yield was 9.3 t/ha in 2016, and this area produced 9% of the US maize.
- Zone 3 (the central plains): Nebraska, Kansas, Colorado are the second largest producing area, with much of the maize irrigated from a 450,000 km^2 groundwater (Ogallala aquifer) that extends on 8 states. It is located east of the Rocky Mountain Range, with a dry continental climate and produces a quarter of the US maize production. Corn competes with HRW wheat (Hard Red winter) and soybeans, which are also irrigated. The level of the groundwater table is nevertheless decreasing (-77 cm on average between 1980 and 1995), which could eventually pose irrigation problems. In 2016, this area produced 17% of the maize in the US and had a yield of 9.8 t/ha.
- Zone 4: maize is also grown along the Mississippi River (Lower Mississippi River Basin), where crops are irrigated with moderation. They include SRW (Soft Red winter), soybean, sorghum, and cotton. In 2016, this area yielded 3% of the maize of the US and had a yield of 9.6 t/ha.
- Zone 5: finally, the East Coast produces about 3.4% of maize with an average yield of 6.2 t/ha (all numbers come from the US Department of Agriculture, National Agricultural Statistics Service). Irrigation is limited.

Irrigation is very important in some of the western regions. According to the US Geological Survey in 2005, “the majority of withdrawals (83%) and irrigated acres (74%) were in the 17 western states”. These regions should be discarded in order to obtain a clearer link between precipitation and corn yield. Therefore, all federal states located west of 103° West and some other isolated states (for instance Nebraska and Texas) have not been considered in the study. Finally, the 28 States considered in the study are: Illinois, Indiana, Iowa, Kentucky, Michigan, Minnesota, Missouri, Ohio, and Wisconsin (Zone 1); North Dakota, and South Dakota (Zone 2); Kansas, Nebraska, and Oklahoma (Zone 3); Arkansas, Louisiana, Mississippi, and Tennessee (Zone 4); New Jersey, Delaware, Maryland, Virginia, West Virginia, Pennsylvania, North Carolina, South Carolina, Georgia, Alabama (Zone 5).

The analysis has been realized at the level of each production zone, and also at the US level (considering the 28 states listed above).

Fig. 2 illustrates the natural dispersion and variability of the yield anomalies for the five production zones. In zone 1, there is a lower dispersion (suggesting a lower weather sensitivity and low inter-annual variability), but many negative extreme values (suggesting a higher sensitivity to adverse weather conditions). Areas 2, 3 and 5 have very similar natural variability. Zone 4 has a low dispersion and few extreme values.

2.1.3. Direct meteorological data

Temperature (monthly mean, daily min, daily max and daily mean) and cumulative precipitation data were collected for the period 1979–2013, from the ERA-Interim re-analysis of the European Center for Medium-Range Weather Forecasts (Uppala et al., 2005). We will focus on these 35 years. The US territorial organization and its subdivision into counties are used to project data, from its original 75 km × 75 km regular grid into county-level data. To perform the spatial interpolation, we used the closest pixel technique. It is crude and could be improved (bi-linear interpolation for instance) but in our case,