Computers and Electronics in Agriculture 127 (2016) 726–738

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

PHASE: A geostatistical model for the Kriging-based spatial prediction of crop phenology using public phenological and climatological observations

H. Gerstmann [a,*], D. Doktor [b], C. Gläßer [a], M. Möller [c,a]

[a] Martin Luther University Halle-Wittenberg, Department for Remote Sensing and Cartography, Von-Seckendorff-Platz 4, 06120 Halle (Saale), Germany
[b] Helmholtz Centre for Environmental Research, Department for Computational Landscape Ecology, Permoserstraße 15, 04318 Leipzig, Germany
[c] Martin Luther University Halle-Wittenberg, Institute of Agriculture and Nutrition Science, Department of Farm Management, Karl-Freiherr-von-Fritsch-Straße 4, 06120 Halle (Saale), Germany

ARTICLE INFO

Article history:
Received 8 May 2015
Received in revised form 26 July 2016
Accepted 29 July 2016
Available online 10 August 2016

Keywords:
Phenology
Kriging
Modeling
Growing degree days
Cross validation
Germany

ABSTRACT

Detailed information on plant developmental stages, referred as phenological phases, can assist research, applications and synergies e.g., in land use, climate science and remote sensing. Usually, detailed ground information about phenological phases is only available as point observations. However, in most application scenarios of spatially interpolated phenological information is required. In this article, we present an approach for modeling and interpolation of crop phenological phases in temperate climates on the example of the total area of Germany using statistical analysis and a Kriging prediction process. The presented model consists of two major parts. First, daily temperature observations are spatially interpolated to retrieve a countrywide temperature data set. Second, this temperature information is linked to the day of year on which a phenological event was observed by a governmental observation network. The accumulated temperature sum between sowing and observed phenological events is calculated. The day on which the temperature sum on any location exceeds a phase-specific critical temperature sum, which indicates the day of entry of the modeled phase, is finally interpolated to retrieve a countrywide data set of a specific phenological phase. The model was applied on the example of eight agricultural species including cereals, maize and root crops and 37 corresponding phases in 2011. The results for most of the tested crops and phases show significantly lower *root mean squared errors* (*RMSE*) values and higher *goodness of fit* (R^2) values compared to results computed using Ordinary Kriging (OK) and Inverse Distance Weighting (IDW). The modeling accuracy varies between 2.14 days and 11.45 days for heading and emergence of winter wheat, respectively. The uncertainty of the majority of the modeled phases is less than a week. The model is universally applicable due to automatic parametrization, but model accuracies depend on the crop type and increase during a growing season. The possibility to enhance the model by additional explaining variables is demonstrated by consideration of soil moisture within an extended model setting.

1. Introduction

Phenology studies periodic events in plant development and their dependence on shifting environmental factors such as temperature, day length and precipitation (Kirby et al., 1987; McMaster et al., 2009). Such events and phases are clearly visible developmental stages like blossoming or ripening (Schwartz, 2006).

The main climatic drivers of plant phenology vary in different ecoregions. Temperature is the main driving factor for intra-seasonal timing of phenological events in temperate regions like Central Europe (Chmielewski et al., 2004; Menzel, 2007). Many studies observed that in temperate climates the timing of phenological events is relatively stable and independent of other environmental factors than temperature (e.g., McMaster et al., 2009). Other factors influencing plant phenological development are photoperiod (Masle et al., 1989), daily temperature amplitude (Solantie, 2004), water availability and soil moisture especially in arid and semi-arid climates (McMaster and Wilhelm, 2003; Idso et al., 1978), solar radiation, distance to coasts and settlements, soil

* Corresponding author.
E-mail address: henning.gerstmann@geo.uni-halle.de (H. Gerstmann).

http://dx.doi.org/10.1016/j.compag.2016.07.032
0168-1699/

properties (Zhao et al., 2013) and management factors like date of planting or fertilization practices (Nellis et al., 2009).

Crop phenology in Germany follows several spatial trends. Due to Germany's temperate climatic conditions, phenology is predominantly determined by temperature. Other factors influencing phenology in temperate regions to a lesser extent are precipitation and soil moisture, especially for autumn phases (Menzel, 2007), elevation, sea proximity and population density (Hense and Müller, 2007). Thus plant development in Germany is delayed in coastal and mountainous regions compared to the favored regions in south-western and central lowland regions (Siebert and Ewert, 2012).

Knowledge about plant phenological phases and their timing is of interest for wide application scenarios. Since plants react to changing temperatures and carbon dioxide content, long-term phenological time series can be used to monitor responses of plant phenology to global and regional warming (Estrella et al., 2007). Prevailing phenological information is also required for the assessment of famine risks and food production problems (Vrieling et al., 2011).

Several studies have also shown the potential of phenological information to support land cover classification of remote sensing images, models for crop yield estimation and precision farming on regional and continental scales (Van Niel and McVicar, 2004; van Bussel et al., 2011; Möller et al., 2012; Foerster et al., 2012; Prishchepov et al., 2012). Furthermore, phenology information has the potential to provide valuable input to soil erosion monitoring (Möller et al., 2015), mapping of biodiversity (Turner et al., 2003) or monitoring of invasive plant species (Bradley and Mustard, 2006; Huang and Asner, 2009).

Such support can be expected to continue to gain importance since the temporal availability of medium or high spatial resolution satellite sensors will considerably increase once the Sentinel-2 satellite constellation is working operational (Berger et al., 2012; Drusch et al., 2012). This requires reliable algorithms for data set selection in which phenology can play a major role to detect the most significant data sets for an image classification problem (Möller et al., 2012). In doing so, the required data amount is reduced with minimal loss in accuracy and thus enables an operational use of these data amounts both in environmental and agricultural sciences as well as in policy and decision making.

Detailed phenological data are mostly available as point observations of irregular spatial distribution which represent phenological phases in standardized numeric codes. Spatial information about phenological phases of crops can be also extracted from satellite images of high temporal resolution and corresponding vegetation indices provided for instance by Meteosat (Sobrino et al., 2013) and MODIS (e.g. Zhang et al., 2003; Lunetta et al., 2006; Jönsson et al., 2010; Xiao et al., 2013). However, these methods are mostly applied on only a few clearly visible phases like green-up or onset (Hird and McDermid, 2009).

The mentioned application scenarios require operationally effective and detailed phenological information. To produce such data, point observations have to be spatially interpolated using phenological models. Menzel (2007) and Zhao et al. (2013) distinguish three main types of phenological models:

1. Statistical fitting models which relate climatic variables to phenological development phases (e.g. McMaster and Wilhelm, 1997; Picard et al., 2005).
2. Mechanistic models that are based on cause-effect-relationships (Jamieson et al., 1998; Kramer et al., 2000; Ewert et al., 2002; Hänninen and Kramer, 2007).
3. Theoretical models which focus on plant physiological processes (Kaduk and Heimann, 1996; Schaber and Badeck, 2003; Peng et al., 2011).

Mechanistic and theoretical approaches require a large number of parameters and experimental effort. Statistical fitting methods only require a few input data sets, are of lower complexity and thus more frequently applied. One of the most often applied statistical fitting approach is based on the relation between the observation day of year of a phenological event (DOY_{obs}) and the corresponding accumulated effective temperature (Chuine et al., 2003; Hänninen and Kramer, 2007). This phase- and plant-specific temperature sum is usually referred as *growing degree days* (*GDD*), *heat units*, or *thermal time* (Zhao et al., 2013).

The majority of studies focused on either a region of limited extent or differences in plant parameters, mainly base temperature (Holen and Dexter, 1996; McMaster and Wilhelm, 2003), for different cultivars or cultivation sites of one crop type and between phenological phases (e.g. Wang and Engel, 1998; Ewert et al., 2002; Salazar-Gutierrez et al., 2013). A common problem is that the optimal starting day for *GDD* summation is difficult to determine (Wielgolaski, 1999). Furthermore, most of these studies do not combine phenological models and spatial interpolation since they often refer to pre-defined reference units (e.g. van Bussel et al., 2011; Siebert and Ewert, 2012).

To address these disadvantages, we present a framework which combines a geostatistical method and the *GDD* concept. In doing so, all critical parameters are extracted automatically and dynamically from the input data. After the geostatistical interpolation of daily mean temperatures, temperature sums and observed phenological phases are empirically related in order to extract the entry date of a specific phenological phase. These entry dates are again geostatistically interpolated to obtain area-wide predictions. The model has been designed to be easy-to-use, independent of expert knowledge, extendable, and transferable to any region of temperate climate where phenological observations and temperature measurements are available. The framework consisting of the combined model and the geostatistical interpolation was named PHASE (**PH**enological model for **A**pplication in **S**patial and **E**nvironmental sciences).

In this article, we describe the model structure, its underlying algorithms and methodological background (Section 3.2). We demonstrate its application on a selection of frequently grown crop types with special focus on winter wheat (*Triticum aestivum L.*) for the entire area of Germany (Section 3.2.3) using temperature data and phenological information provided on-demand for free by the German Weather Service.[1] The possibility to enhance the model by additional explaining variables is demonstrated by consideration of soil moisture within an extended model setting.

2. Materials and data

2.1. Phenological data

In Germany, the data base for phenological and meteorological observations is of unique density and quality and thus well-suited for model development. The German Weather Service (*German: Deutscher Wetterdienst* – DWD) operates a phenological monitoring network consisting of about 1200 active stations spread over Germany which report the Julian day of entry (day of year – *DOY*) for numerous phenological phases of agricultural crops, wine and natural plants at the end of each year (Hense and Müller, 2007). Each plant is observed on a different number of stations, depending on the abundance and agrometeorological relevance of the respective crop type. The observations are recorded by volunteers following standardized criteria, and a numeric code is assigned for each phase (Table 1).

[1] http://www.dwd.de.