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

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.

Table 1
Numeric codes for observed phenological phases of agricultural crops (Deutscher Wetterdienst, 2015a).

Phenological phase	Numeric code (phase ID)
Beginning of flowering	5
Full flowering	6
Beginning of sowing	10
Emergence	12
Closed stand	13
4th leaf unfolded	14
Beginning of shooting/stem elongation	15/67
Beginning of bud formation	17
Beginning of heading	18/66
Beginning of milk ripening	19
Early dough ripening	20
Beginning of yellow ripening	21
Beginning of full ripening	22
Harvest	24
Beginning of tassel emergence	65

A certain degree of quality control is already applied to these data, which is based on a multi-step cross validation procedure using ±25 days observation error as first threshold and ±15 days error in the second step (Hense and Müller, 2007). The most frequently cultivated crops in Germany (Statistisches Bundesamt, 2011) and the related, continuously reported phases are listed in Table 2. The annually collected phenological observations of these predominantly grown crops can be obtained via FTP server (Deutscher Wetterdienst, 2015a).

The distribution of the phenological stations is irregular both between regions of different environmental conditions and within regions of similar suitability for cultivation of a specific crop. For instance, in regions where winter wheat is cultivated frequently (Central Germany, and southern Germany apart from the mountainous regions) the point density varies between 1.5 and 3 stations per 1000 km^2, but no regional trend is visible. In mountainous and coastal regions, demanding crops cannot or only rarely be grown so that observation density is much lower. Additionally, the number and location of observations vary between different years due to crop rotation practices. Since the observers define the extent of their observation area according to the current cultivation status, positional accuracy of the phenological stations is approximately 2–5 km in position and 50 m in elevation (Deutscher Wetterdienst, 2015b). This limits the applicability for modeling phenology on medium spatial resolution, because it is impossible to consider temperature differences on mesoclimatic scale. Despite quality control is operationally applied to the data by the DWD, there are still outliers in the final product. In this study, a statistical filter is therefore applied which only considers observations within an interval of 1.5 standard deviations around the mean of the total data set. Exemplary, all stations that reported the phenological phase *beginning of yellow ripening* in 2011 are shown in Fig. 1b.

2.2. Temperature data

Daily mean temperatures were obtained for 503 stations per day on which the DWD is running daily automatic quality control procedures to detect and remove measurement errors. This error removal results in slight daily differences of the actual number of temperature observations per day. The positional uncertainty of the stationary temperature stations is assumed to be not significant. The temperature stations are not regularly distributed (Fig. 1b). Population density determines the network density as well as meteorological criteria. For instance, in mountainous areas where more short-term weather changes occur, the density is higher than in lowland areas of more constant weather conditions.

2.3. Elevation data

Since temperature is strongly dependent on elevation and topography, a digital elevation model (DEM) is used as explanatory variable for temperature interpolation. For this study, the *SRTM DEM* (USGS, 2004, see Fig. 1a) was selected. The data set has been produced by the space-borne STS-99 Shuttle Radar Topography Mission (SRTM) sensor and is freely available outside the United States at 90 m resolution with a vertical accuracy of 20 m and a horizontal accuracy of 16 m (Rabus et al., 2003). The DEM has been filtered following Lee (1980) to reduce signal noise. Elevations below sea level were set to 0 m.

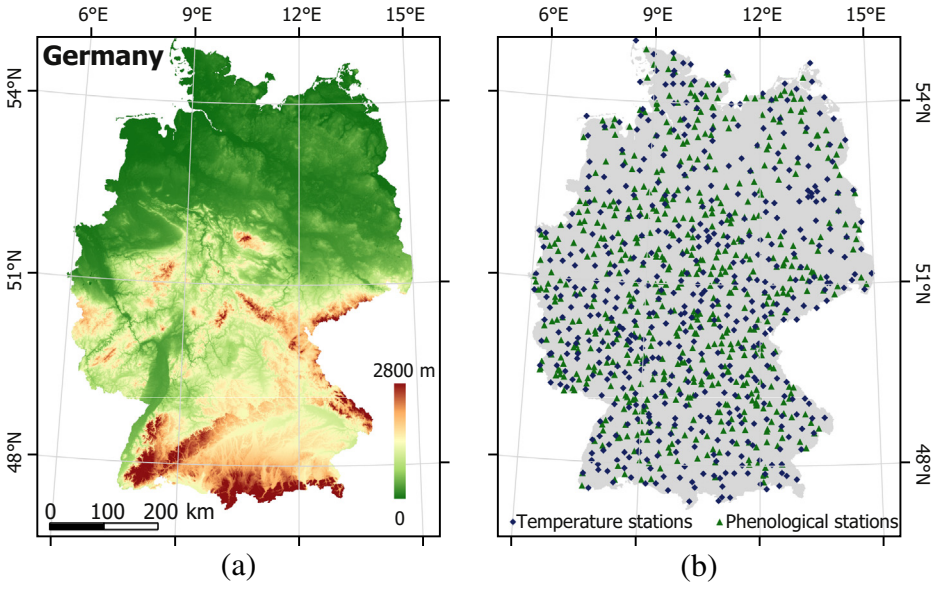

Fig. 1. Germany-wide input data sets. 1 km digital elevation model (a); all phenological and meteorological DWD stations that observed the phase *beginning of yellow ripening* for winter wheat in 2011 and all stations that provided daily temperature measurements on 1st January 2011 (b). Projection: EPSG code 25632 (Spatialreference, 2015). Data sources: USGS (2004) and Deutscher Wetterdienst (2015a).

The DEM raster cells were aggregated to 1×1 km pixel size with a total number of 358,320 pixels. The 1 km^2 resolution was chosen to consider the uncertainty in position and altitude of the phenological observations.

2.4. Soil moisture data

The DWD also provides daily raster data which represent the soil moisture in per cent of accessible field capacity for grass covering entire Germany (Deutscher Wetterdienst, 2015a). Spatial resolution of the raster data sets is 1×1 km. According to the data set publisher (DWD Climate Data Center (CDC), 2016), the data set is produced using observational data of approximately 280 synoptic climate stations. The soil moisture is calculated using a modified Penman–Monteith rationale for determination of evapotranspiration (Löpmeier and Deutscher Wetterdienst, 1983). The station-specific soil moisture values are interpolated using multiple linear regression and triangulation. No quality assessment routines are applied by the DWD, hence the data uncertainty cannot be quantified. Apart of the algorithm-dependent uncertainty, two further restrictions decrease the usability for the phenological model. First, sandy loam is used as soil type for entire Germany without taking regional differences into account. Second, the interpolated values for field capacity are only valid for unspecified grass vegetation.

3. Methods

3.1. Methodological background

3.1.1. Phenological modeling

Numerous studies have shown that in temperate climates phenological development correlates with accumulated heat sums (Sitch et al., 2003; Miller et al., 2001; McMaster and Wilhelm, 1997; Russelle et al., 1984). They are calculated from the measured daily mean temperatures above a base temperature T_B, which is the lower threshold for photosynthetic activity. The underlying concept has been described by Réaumur (1735) and applied for prediction of crop phenology for centuries. This relation is often used for modeling plant phenology and is mostly referred as *growing degree day* (*GDD*):

$$GDD = 0.5 \times (T_{max} - T_{min}) - T_B \quad (1)$$

According to Eq. (1), a *GDD* is the daily contribution to the accumulated heat sum, T_{max} and T_{min} are the daily observed maximum and minimum temperatures, and T_B is the base temperature. T_B can be either species-specific or individual for each phase of a crop species (Slafer and Savin, 1991). The daily *GDD* – beginning at a starting point in time (DOY_{start}) – is then accumulated over a vegetation cycle. Because winter crops are sown during autumn of the previous year, the often-used 1st January is not necessarily the optimal starting day. This day does not take environmental differences into account which are expected to lead to inaccurate extraction of the critical temperature sum. Station-specific starting days like the day of sowing or the start of snow melt can be chosen instead to respect topographic or climatic differences. Since other factors than climate are also influencing the day of sowing, using snow melt as starting point is also inappropriate. Hence we expect the reported day of sowing (phase ID: 10, see Table 1) as most reliable starting day.

On the day on which the accumulated temperature sum exceeds a phase-specific threshold, the plant reaches the next stage of its phenological cycle. However, this relation is only applicable for phenological phases with a physiological background and not for the phases *sowing and tilling* or *harvest*.

Photosynthetic activity can start even on days with a negative contribution to *GDD*, if T_{max} is higher than T_B and T_{min} is lower than T_B resulting in a mean temperature below 0 °C (McMaster and Wilhelm, 1997). This can be especially the case during the early months of a growing season. To determine the portion of the day with temperatures higher than T_B, hourly temperature measurements are required. Since the use of hourly observations would increase the amount of required data rapidly and we assume the effect of on the modeling results to be negligible, we used observed daily mean temperatures.

3.1.2. Kriging

Kriging is a widely-applied geostatistical method for interpolation of spatial data that was first presented by Krige (1951) for improving ore reserve estimations. It is an algorithm that uses the decreasing autocorrelation between two sample points with increasing spatial distance to predict intermediate values. The algorithm is based on an empirical variogram and a fitted variogram model to predict a variable at a location where no sample exist. Several theoretical variogram models exist, but the Matérn (Matérn, 1960) model is strongly suggested for the interpolation of spatial data (Stein, 1999) and can be used universally both for short and long distance variation models (Hengl, 2009). This variogram model includes a smoothing parameter *kappa* and a Gaussian model as limiting case as well as an exponential model as special case and thus is more flexible to local behavior of the observations than other models.

Kriging allows the consideration of possible correlations of the predicted variable with various explanatory variables (Hengl et al., 2007). For instance, this is the case for air temperature, which shows a dependency to elevation and surface topography. Kriging using explanatory variables is referred as Universal Kriging, Kriging with external drift or Regression Kriging. Other methods like Inverse Distance Weighting interpolation, averaging values per polygon and Ordinary Kriging can be understood as special cases of the Regression Kriging method (Hengl, 2009), which hence is understood as best linear unbiased predictor (Stein, 1999; Hengl, 2009).

3.1.3. Cross validation

Cross validation is an often applied method for accuracy assessment for interpolations of environmental variables (Kuhn and Johnson, 2013). It compares predicted values on a location with the observed values on the same location. Leave-one-out cross validation excludes one sample from the input data set and predicts this sample using the surrounding observation points. Afterwards, the difference between the predicted value and the observed value is used as estimate of the prediction performance. 10-fold cross validation is an alternative technique, which is a variant of bootstrapping and is insensitive against outliers (Hengl, 2009). Here, the data set is split into 10 parts (folds) of equal size and each fold is used for cross validation and calculation of the parameters *Root Mean Squared Error* (*RMSE*) and R^2 as metrics for the model's *goodness of fit*. *RMSE* is a measure to describe the difference between observed and modeled values and is calculated according to Eq. (2).

$$RMSE = \sqrt{\frac{(y_{obs} - y_{pred})^2}{N_{obs}}} \quad (2)$$

Here, y_{obs} and y_{pred} represent observed and modeled values and N_{obs} is the number of observations. Cross validation provides global accuracy measures as well as the residuals for each observation which can be used to detect observations of high uncertainty that in turn can affect the entire prediction result negatively.

3.2. Basic model conception

The core PHASE model follows the growing degree days approach (see Eq. (1)) but includes day length as proxy for latitude influences. It determines the number of required heat units to reach a phenological phase, which is representative for the complete study area. First, a station- and phase-specific threshold of required heat units is determined (Eq. (3)).

$$T_{sum}^{eff}[j] = \sum_{i=DOY_{start}}^{DOY_{obs}} \left((\overline{T}_{ij} - T_B) \times \frac{DL_i}{24} \right) \quad (3)$$

$T_{sum}^{eff}[j]$ in heat units (HU) represents the accumulated effective temperatures on a specific phenological station j between a fixed starting day DOY_{start} and the observed phenological event DOY_{obs}. $\overline{T}$ is the daily mean temperature and DL the day length at the phenological stations, T_B is the base temperature.

Next, a critical HU value T_P is determined as a quantile of the distribution of all values of T_{sum}^{eff} (Eq. (4)).

$$T_P = Q_{opt}(T_{sum}^{eff}) \quad (4)$$

T_P is the critical temperature sum required for a specific phase and Q_{opt} is the optimal quantile of T_{sum}^{eff} of all phenological stations. Finally, the first day DOY_P on which the cummulated heat units are equal or exceed the critical temperature sum T_P is calculated for each station.

The work flow in Fig. 2 can be distinguished in three parts:

1. Data pre-processing is performed which includes the interpolation of daily temperature observations to retrieve a daily mean temperature raster data set covering entire Germany ($\overline{T_{ij}}$ in Eq. (3); see Section 3.2.1) and the outlier detection of the phenological observations,
2. Fitted values of the parameters of Eqs. (3) and (4) are estimated by finding the configuration that leads to the lowest possible *RMSE*, resulting in the calculation of accumulated daily effective temperatures (T_{sum}^{eff}).
3. DOY_P is determined and interpolated to produce a raster data set representing the day of entry of the target phenological phase. Additionally, cross-validation is performed to assess the prediction accuracy (Section 3.3).

Due to the automatic fitting procedures, the model is completely data-driven and no prior assumptions on the data distribution are required for modeling.

The model has been implemented within the statistical computing environment *R* (R Core Team, 2015). All geodata are imported using the *R*-packages *rdgdal* (Bivand et al., 2015) and *raster* (Hijmans, 2015).

3.2.1. Temperature interpolation

The daily mean temperatures are interpolated using local Universal Kriging (see Section 3.1.2) which is implemented within the *gstat* package (Pebesma, 2004). Elevation is used as independent variable to respect topography influences in the prediction. A Matérn variogram model is automatically fitted to the empirical semivariogram using the function *fit.variogram()* which is also included in the *gstat* package. A crucial parameter is the number of stations to include in the temperature prediction at a certain location (*nmax*). In principle, the higher *nmax* the more accurate predictions can be assumed. On the other hand, Kriging is based on the assumption that more distant observations are almost ineffective for the prediction result (Hengl, 2009). Since computation time increases strongly with increasing number of considered stations, the number of the included stations should be reduced to limit processing time.

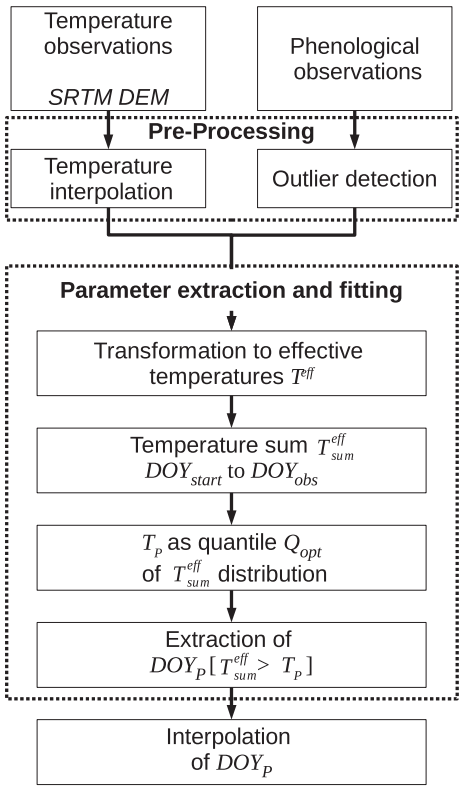

Fig. 2. Structure of the basic model including preprocessing, fitting and spatial interpolation. *SRTM DEM* - SRTM digital elevation model; T_{sum}^{eff} - accumulated temperatures from DOY_{start} to observed day DOY_{obs}; T_P - required heat units; Q_{opt} - Quantile of T_{sum}^{eff} distribution that leads to lowest RMSE values; DOY_P - day on which accumulated temperature sum exceeds T_P.

3.2.2. Calculation of effective temperatures

The interpolated mean temperatures $\overline{T}$ are converted to station-specific daily effective temperatures T^{eff} (Eq. (3), comparable with *GDD* in Eq. (1)). First, the base temperature is subtracted from the mean temperature of each day. Next, the resulting temperatures are adjusted to take the dependence of photosynthesis on available sunlight (photoperiod) into account. Here, the observed mean temperature are multiplied by the ratio of day length (DL [h]) to the whole day.

DL is calculated using the *R* function *daylength()* that is included in the package *geosphere* (Hijmans et al., 2014). T_B is the base temperature below which no growth occurs. Although different values for T_B exist, the most often applied default values which are used independently of crop type and region are 0 °C, 5 °C or 10 °C. The actual base temperature for a certain plant or cultivar and phase can be also calculated by statistical formulas (Yang et al., 1995) or selected using standard look-up-tables. In this study, a sequence of T_B values is tested ranging from 0 °C to 10 °C.

3.2.3. Extraction of DOY_P

The plant- and station-specific critical temperature sum (T_{sum}^{eff}) for a phenological phase on a specific station is calculated by summing up all daily temperatures from a starting point DOY_{start} to the day on which the phenological phase has been observed (DOY_{obs}). In this study, DOY_{start} is defined by the reported sowing day. All negative effective temperatures are considered as not significant for photosynthetic activity.

The day DOY_P, on which the cumulative sum exceeds the indicator temperature sum for the entry of a target phaseT_P, that is used as indicator for the entry of a target phase, is calculated for each station. T_P is automatically detected by testing 19 quantiles ($Q \in [0.05, 0.95]$) in steps of 5% of the T_{sum}^{eff} distribution. The quantile Q_{opt} which leads to the lowest *RMSE* between observed and estimated *DOY* is used for determination of T_P.

3.2.4. Spatial interpolation of DOY_P

The modeled days are passed to a Kriging interpolation with elevation as independent variable using an automatically fitted Matérn variogram model (see Section 3.2.1). The initial values for fitting were set according to Hiemstra et al. (2009), the smoothing parameter *kappa* was set to 1. Here, the optimal value for *nmax* is determined using a 10-fold cross validation technique, and the *nmax* value with the lowest *RMSE* is used for the final prediction.

3.3. Accuracy assessment

An estimate of uncertainty is provided as part of the Kriging procedure. This estimate is referred as Kriging variance and represents the variance of the predicted result on each location in comparison to the optimum. The Kriging variance can be used to assess spatial patterns in uncertainty of the prediction.

RMSE and R^2 were computed using 10-fold cross validation (see Section 3.1.3). For comparison, *RMSE* and R^2 achieved by Inverse Distance Weighting interpolation (IDW) and Ordinary Kriging (OK) applied on the raw phenological observation data using a Matérn variogram model and the *fit.variogram()* function were computed.

3.4. Model extension using soil moisture

The model can be extended by further parameters that influence plant phenological development, e.g., soil moisture. Although soil moisture is less effective for modeling plant development in temperate climates than temperature (Fu et al., 2014), it is of special importance for the model's transferability to other, especially arid and semi-arid climates where soil water availability more influences plant development. Independently of the data limitations mentioned in Section 2.4 and the, for the test site with its temperate climate, expected low significance for the modeling results, a second normalization of the daily effective temperatures (T^{eff}) was tested (Eq. (5)).

$$T^{eff} = (\overline{T} - T_B) \times \frac{DL}{24} \times \frac{1 - \cos(\pi \times \theta_{fc})}{2} \quad (5)$$

Here, T^{eff} is the daily contribution of a day to T_{sum}^{eff}. $\overline{T}$ and T_B are the mean and base temperature, *DL* is the length of daylight and θ_{fc} is the field capacity representing the soil moisture accessible by plants. The additional normalization term is formulated to equal 1 when the plant available field capacity is 100% and declines in conditions of stress due to soil wetness and drought, so that water deficits and surplus reduce the daily contribution of T^{eff} to T_{sum}^{eff}. In doing so, a 10% deficit as well as 10% surplus, for instance, result in the weighting factor of 0.976. Consequently, warm days with many day light hours and 100% plant available field capacity are understood as the most effective days for plant development.

4. Results and discussion

The model has been applied for all crop types and corresponding phenological phases with a plant physiological background (see Table 2) that were observed in 2011. This includes the spring and summer phases of crops that were harvested in 2011, and the autumn phases of the next vegetation cycle that was terminated in 2012. The results are discussed with special focus on the phases *beginning of shooting* and *beginning of yellow ripening* of winter wheat, which is the most frequently grown crop type in Germany, with the exception of perennial grasslands. In 2011, winter wheat was grown on 19% of the total area under cultivation (Statistisches Bundesamt, 2011), followed by maize (13%), winter barley (9%) and winter oilseed rape (8%). Fig. 3 shows the corresponding *DOY* distributions for the total area of Germany whose outliers were removed by statistical filtering (see Section 2.1).

Table 2
Reported phenological phases for the eight typical field crops in Germany (DWD, 2015). See Table 1 for full phase names.

Crop	Crop (scientific name)	Reported phases
Winter wheat	*Triticum aestivum L.*	10, 12, 15, 18,19, 21, 24
Winter rye	*Secale cereale L.*	10, 12, 15, 5, 6, 18, 21, 24
Winter barley	*Hordeum vulgare L.*	10, 12, 15, 18, 21, 24
Oilseed rape	*Brassica napus L.*	10, 12, 14, 67, 17, 5, 22, 24
Oat	*Avena sativa L.*	10, 12, 15, 66, 19, 21, 24
Maize	*Zea mays L.*	10, 12, 67, 5, 65, 19, 20, 21, 24
Potato	*Solanum tuberosum L.*	10, 12, 5, 24
Sugar beet	*Beta vulgaris* subsp. *vulgaris*	10, 12, 13, 24

4.1. Temperature interpolation

Daily temperatures between 2010 and 2011 were interpolated covering all days for which observations exist. According to Hengl (2009), 30–60 included observations result in sufficient accuracy of the final prediction, while Webster and Oliver (2007) stated that 100–150 stations are required at least. A 10-fold cross validation has been applied on all days to determine the number of stations to include in the temperature prediction at a certain location (*nmax*) by subsequent testing of values from 10 to 20, 25, 30, 40, 50, 60, 100 and 200 (Section 3.2.1). The processing times for the predictions of one single day are displayed in Fig. 4. In order to balance prediction accuracy and computation time, the *nmax*

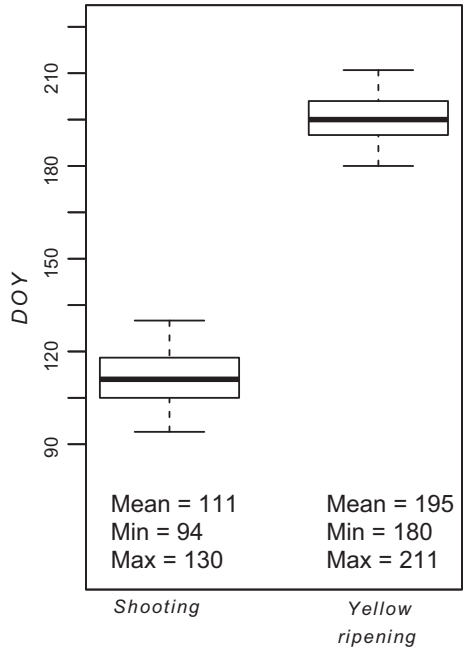

Fig. 3. Boxplots of the observation days for winter wheat phases *beginning of shooting* and *beginning of yellow ripening* in 2011 for the total area of Germany. *DOY* = 111 represents 21st April and *DOY* = 195 corresponds to 14th July 2011.

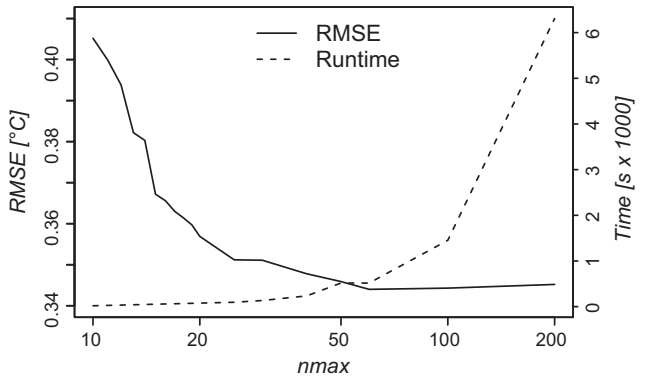

Fig. 4. Mean *RMSE* values of all temperature interpolations derived by 10-fold cross validation using different values for *nmax* (solid line) and processing time of corresponding Kriging interpolation (dashed line).

value was set to 30 for all daily interpolations which provides only slightly higher *RMSE* values than the optimum in most of the test cases while it needs significantly less processing time of about 134 s on a standard PC (4 GB RAM, 2.7 GHz, Microsoft Windows 7OS 64bit). The calculated *RMSE* values for different values for *nmax* vary about 0.06 °C with a decreasing trend to higher *nmax* values. The *RMSE* values for all temperature interpolations are below 1 °C on 89% of the days.

4.2. Extraction of DOY_P and model calibration

4.2.1. Base temperature

Daily temperatures for 2011 were transformed to T^{eff} by normalization regarding photoperiod and subtraction of the base temperatures that have been determined iteratively by testing all values between 0 °C and 10 °C. As Slafer and Rawson (1995) stated for winter wheat, both base temperature as well as sensitivity to photoperiod can vary between different cultivars of a species. The base temperature that leads to the lowest errors in the final prediction was used. Despite $T_B = 0$ °C was used in most of the cases (see Table 3), significant differences exist between both crops and also phases of an individual crop. According to e.g. Salazar-Gutierrez et al. (2013), T_B can vary between 1.6 and 8.4 °C for different winter wheat cultivars. Nevertheless, T_B can be considered as a statistical variable in this study which enables the optimization of model results.

In Fig. 5, the mean errors of the predictions are plotted in relation to the tested base temperatures. The errors for *shooting* clearly exceed the errors for *yellow ripening*. For *yellow ripening*, the base temperature is less decisive for values between 0 °C and 6 °C than for *shooting*, for which the errors increase almost linearly with increasing base temperatures. The stable errors for *yellow ripening* for T_B between 0 °C and 6 °C correspond to the most often reported T_B values for winter wheat, which are between 0 °C and 5.5 °C (McMaster and Wilhelm, 2003; Barrett, 2013). The almost linear

Table 3
Optimal model parameters and accuracy metrics for crop phenological phases observed in 2011.

Plant	Phase	T_B [°C]	Q_{opt}	T_P [HU]	*nmax*	*RMSE*			R^2		
						Mod	Raw	IDW	Mod	Raw	IDW
Winter wheat	Emergence	0	0.35	58	60	11.45	9.33	9.60	0.89	0.15	0.11
	Shooting	0	0.50	423	40	5.75	8.87	9.07	0.65	0.01	0.02
	Heading	10	0.50	118	40	2.14	4.96	5.05	0.72	0.21	0.18
	Milk ripening	5	0.45	588	50	3.57	8.80	8.92	0.65	0.03	0.04
	Yellow ripening	0	0.45	1265	14	3.83	7.11	7.04	0.70	0.08	0.10
Winter rye	Emergence	0	0.40	55	60	8.58	8.24	8.28	0.82	0.03	0.04
	Shooting	0	0.55	413	18	5.92	6.98	7.14	0.79	0.01	0.02
	Beg. of flowering	10	0.50	116	30	2.75	5.53	5.61	0.69	0.18	0.16
	Full flowering	10	0.50	135	60	3.09	6.63	6.88	0.62	0.09	0.05
	Heading	8	0.55	141	50	3.66	5.05	5.10	0.70	0.14	0.13
	Yellow ripening	0	0.45	1283	50	4.51	8.47	8.64	0.61	0.11	0.08
Winter barley	Emergence	0	0.45	73	40	5.04	5.64	5.87	0.81	0.09	0.05
	Shooting	0	0.55	451	50	5.19	7.36	7.45	0.70	0.05	0.05
	Heading	0	0.50	632	60	3.59	4.45	4.66	0.68	0.25	0.18
	Yellow ripening	6	0.50	1064	50	3.88	7.79	7.96	0.70	0.03	0.02
Oilseed rape	Emergence	0	0.45	107	60	7.16	7.10	7.30	0.81	0.04	0.02
	4th leaf unfolded	0	0.40	260	60	7.98	8.88	9.15	0.67	0.01	0.01
	Shooting	0	0.75	603	25	4.90	5.16	5.46	0.72	0.22	0.10
	Bud formation	0	0.70	633	60	5.90	5.45	5.56	0.71	0.19	0.16
	Full flowering	0	0.55	694	50	4.06	4.23	4.30	0.67	0.37	0.35
	Full ripening	0	0.45	1484	20	4.90	8.01	8.33	0.64	0.12	0.07
Oats	Emergence	0	0.40	66	60	5.04	6.00	6.14	0.81	0.15	0.13
	Shooting	1	0.50	284	60	4.12	8.66	9.11	0.63	0.02	0.00
	Heading	0	0.45	541	50	3.23	6.24	6.57	0.66	0.03	0.01
	Milk ripening	0	0.45	855	18	3.29	8.78	9.26	0.57	0.02	0.01
	Yellow ripening	0	0.45	1059	25	3.43	8.31	8.62	0.65	0.08	0.04
Maize	Emergence	3	0.45	79	30	5.06	5.14	5.25	0.82	0.18	0.14
	Shooting	8	0.45	160	40	3.24	9.19	9.54	0.40	0.01	0.00
	Tassel emergence	3	0.45	625	60	3.20	5.72	5.86	0.65	0.16	0.12
	Beg. of flowering	0	0.45	902	60	3.68	7.10	7.18	0.63	0.11	0.10
	Milk ripening	0	0.45	1211	60	3.96	7.64	7.81	0.67	0.09	0.06
	Early dough ripening	0	0.45	1364	30	4.30	8.24	8.59	0.68	0.13	0.07
	Yellow ripening	0	0.40	1468	19	5.32	8.33	8.55	0.70	0.05	0.04
Potato	Emergence	0	0.45	172	30	5.75	7.35	7.23	0.73	0.07	0.09
	Closed stand	4	0.45	295	20	3.88	7.05	7.15	0.64	0.10	0.08
Sugar beet	Emergence	3	0.45	73	30	5.35	10.63	7.61	0.71	0.00	0.02
	Closed stand	3	0.50	421	40	3.21	7.43	7.94	0.47	0.06	0.01

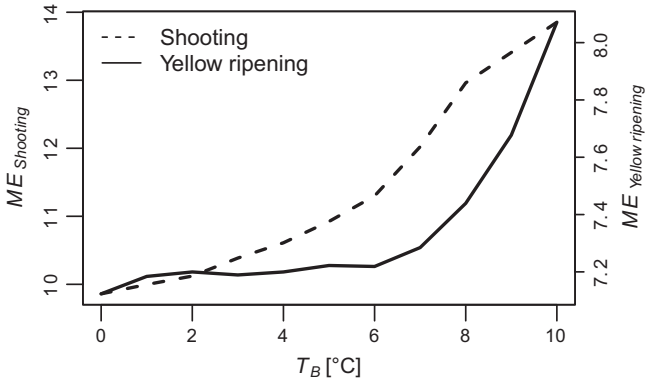

Fig. 5. Mean errors between DOY_P and DOY_{obs} for different base temperatures.

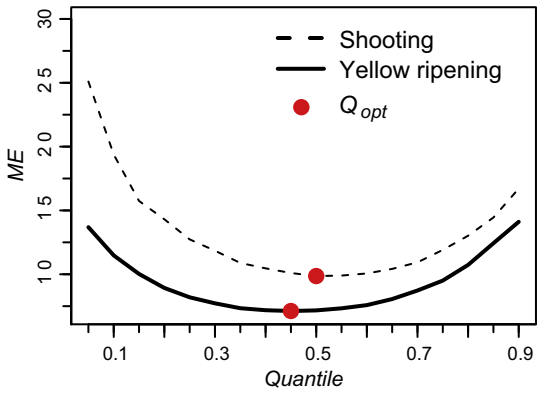

Fig. 7. Mean errors between DOY_{obs} and DOY_P and the optimal quantiles of the T_{sum}^{eff} distributions.

increase of the mean error for *shooting* is probably caused by the exclusion of days with a mean temperature below T_B but photosynthetically effective hours (McMaster and Wilhelm, 1997). This effect decreases for summer phases like *yellow ripening*, but is visible in all further modeling steps (Table 3).

4.2.2. Temperature sums and optimal quantiles

For each phenological station and day, T^{eff} for *shooting* and *yellow ripening* of winter wheat were calculated to retrieve the accumulated temperature sum T_{sum}^{eff} between DOY_{start} and DOY_{obs} by using $T_B = 0\,°C$. According to Fig. 6, the median of T_{sum}^{eff} for *shooting* and *yellow ripening* are 423 HU and 1273 HU with standard deviations of 84.12 and 102.07, respectively.

The DOY_P value on which the temperature sum exceeds each of the tested quantiles was determined. The mean errors of the observed and modeled values and corresponding Q values are shown in Fig. 7. Accordingly, Q_{opt} is equal to the expected median for *shooting* and slightly below for *yellow ripening* where Q_{opt} equals the 45%-quantile.

All data points where the residual of DOY_P and DOY_{obs} is larger than the standard deviation σ of DOY_{obs} were excluded from the data set. This filtering process reduces the considered stations by about 49% and 41% respectively (see Fig. 8). As a result, the model fitting increases significantly. R^2 and *RMSE* for the filtered point data set are 0.65 and 5.75 for *shooting* and 0.70 and 3.83 for *yellow ripening*. Point scattering is higher for *shooting* than for *yellow ripening*, which indicates a better model performance for the prediction of summer phases than on spring phases due to less uncertainty in the calculation of T^{eff} mentioned in Section 4.2.

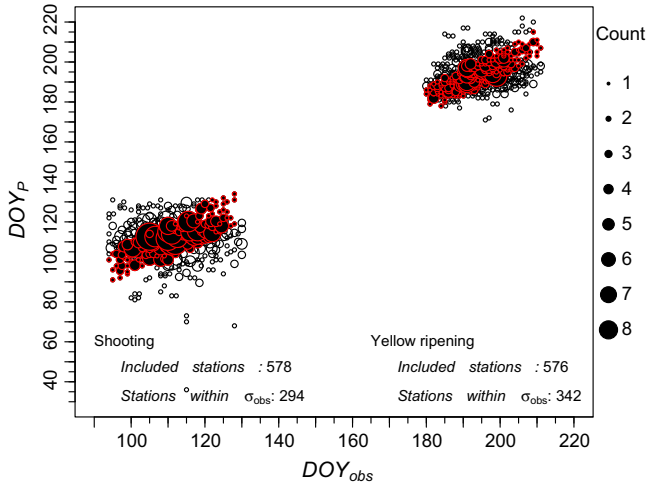

Fig. 8. Density plot of DOY_P and DOY_{obs} for *shooting* (left) and *yellow ripening* (right) of winter wheat in 2011. The diameter of the points indicates the frequency of DOY_{obs}-DOY_P-combinations. All data points within the standard deviation σ are highlighted in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

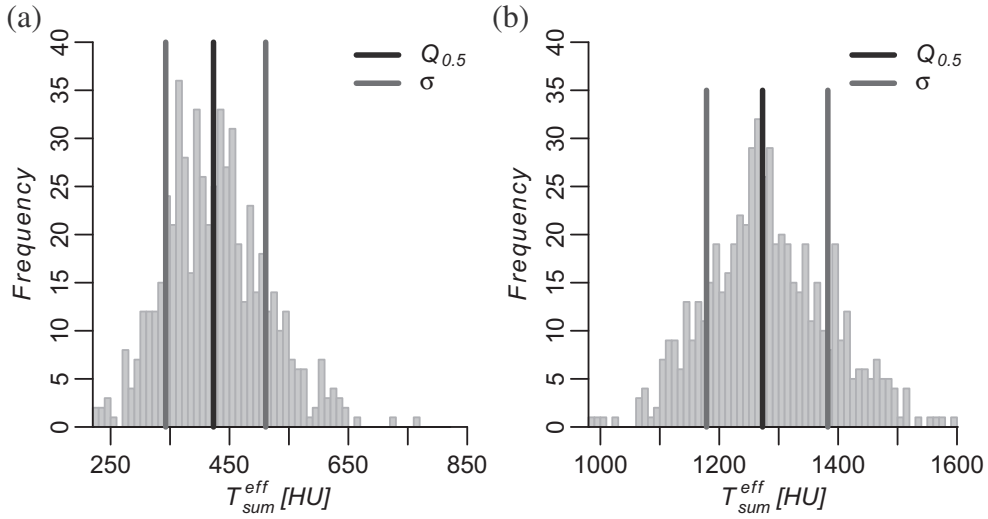

Fig. 6. Distribution of T_{sum}^{eff} calculated for all phenological stations from sowing to the observation date of *shooting* (a) and *yellow ripening* (b) for winter wheat in 2011.

For all other phases and plants, the optimal quantiles as well as T_B and *RMSE* values are listed in Table 3.

4.3. Interpolation of DOY_P

The model results and Kriging uncertainties for the phenological phases *shooting* and *yellow ripening* of winter wheat are shown in Fig. 9. To retrieve an uncertainty measure in days, the square root of the Kriging variance (*Kriging standard deviation*) is plotted. Relatively early dates were modeled for the Rhine valley and the favored regions in eastern Germany, while in mountainous and coastal regions the plant development is delayed. In some regions, the distribution of the observations is relatively coarse (see Fig. 1 and Section 2.1), which causes higher uncertainties (>6 days) compared to regions of dense observations (≈4 days). The prediction in mountainous regions is of high uncertainty (>8 days) for both phases.

The prediction is performed even in areas where a particular crop cannot be grown due to unsuitable environmental conditions. For instance, winter wheat can only be grown in regions with sufficient precipitation, mild winters and warm summers as well as on fertile soils which are not present in entire Germany. However, the model produces information independently of these requirements. This problem is particularly obvious in north-eastern Germany. There, the soils are very sandy and thus not very fertile, and winter wheat is consequently only rarely cultivated and almost no observations exist from this area. However, the Kriging algorithm predicts the phase using the closest locations, in this case from the very fertile regions in Central Germany. Consequently, regions of actually poor environmental conditions appear to be favored for wheat cultivation. The integration of a mask derived from soil types or detailed land use information that margins the prediction on potential cultivation areas could solve this problem in the future.

4.4. Optimal parametrization and model accuracy

The results of the applied 10-fold cross validation (see Section 3.3) are listed in Table 3, where the determined optimal model parameters *nmax*, Q_{opt} and T_P and the accuracy metrics $RMSE_{mod}$ and R^2_{mod} of the interpolated model output for all tested crops and phases are combined. The metrics $RMSE_{raw}$ and R^2_{raw},

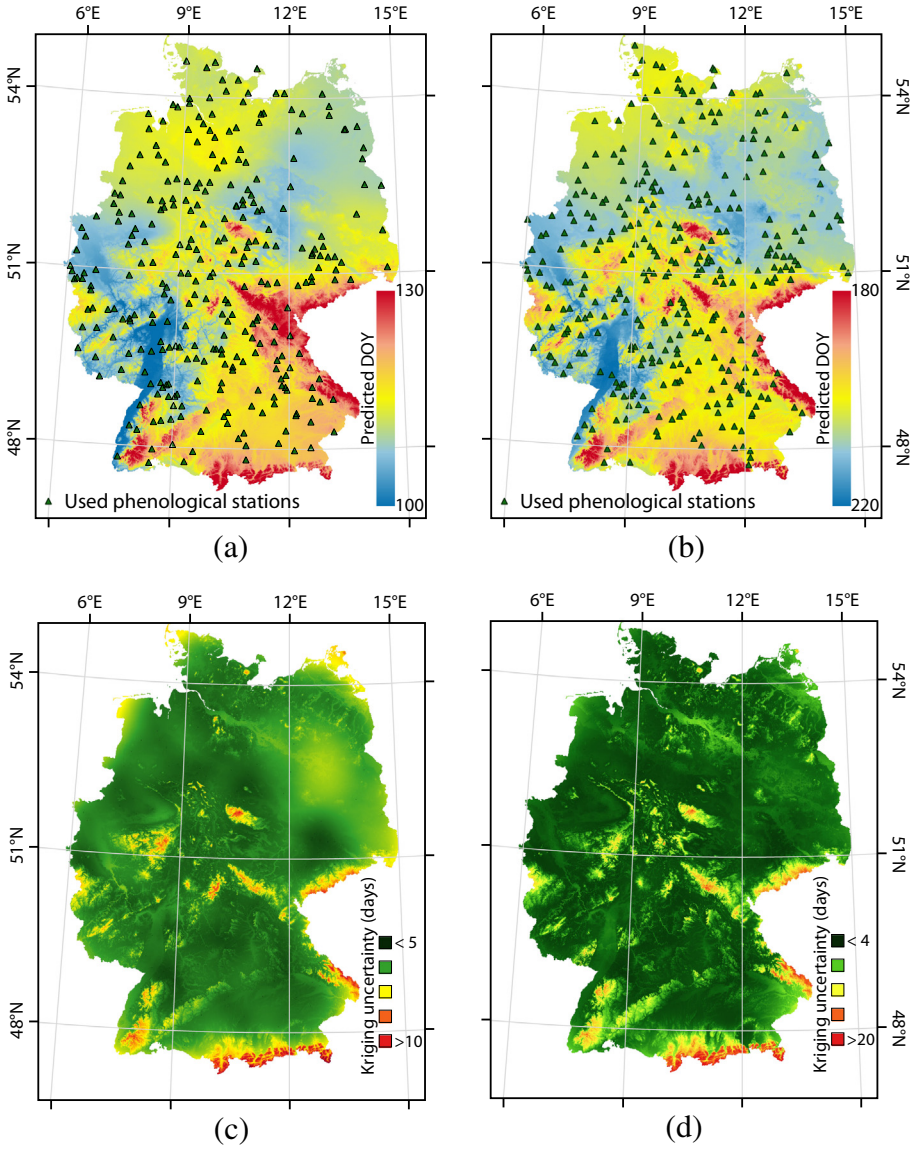

Fig. 9. Interpolated modeling results for the phenological phases *beginning of shooting* (a) and *yellow ripening* (b) of winter wheat in 2011 and corresponding Kriging standard deviations (c and d). Projection: EPSG code 25632 (Spatialreference, 2015).

derived by Ordinary Kriging of the raw phenological observations and $RMSE_{IDW}$ and R^2_{IDW}, calculated using the widely applied Inverse Distance Weighting interpolation for the original and unfiltered observations, are added for comparison.

Q_{opt} is mostly ±5% around the median, with the exceptions of *emergence* of winter wheat, winter rye and oats. In case that T_B is differing between two subsequent phases, T_P can be lower for the later than for earlier phase. This is the case for the phases *shooting* and *emergence* of winter wheat, for which T_B differs by 10 °C. However, the determined T_P in general correlates to *GDD* values published in the literature (e.g., Miller et al., 2001). The modeling of the early phases shows higher $RMSE_{mod}$ values and smaller differences to $RMSE_{raw}$ values compared to later phases, while no declining trend is visible for $RMSE_{raw}$ values. Since *emergence* occurs only a few days after sowing, this phase depends highly on management practices and less on temperature, and the model accuracy is consequently reduced for this phase. $RMSE_{mod}$ is higher than $RMSE_{raw}$ for some phases e.g., for *emergence* of winter wheat, winter rye, oilseed rape and potatoes, which indicates deficits of the model for early phases.

Oilseed rape phases are modeled with lower accuracy than other crops. For two phases, *shooting* and *bud formation*, the optimal quantile is significantly above the median (75% and 70%, respectively). Additionally, $RMSE_{mod}$ for *bud formation* is higher than $RMSE_{raw}$ and $RMSE_{IDW}$. This higher uncertainty indicates a stronger dependency of their phenological development on factors which are not considered within the model like comparatively high requirements to nitrogen, potassium and sulfur nutrition (Grant and Bailey, 1993) or fertilization.

4.5. Integration of soil moisture

For each of the 1 × 1 km DEM and temperature raster cell, the corresponding soil moisture value was extracted for every day and used for calculation of the modified effective temperature term (see Eq. (5)). The consideration of the soil moisture normalization term for the calculation of T^{eff} resulted in reduced *RMSE* for 18 of the modeled phases, while the remaining 19 showed no effect or slightly increased *RMSE*. The improvement is highest for *emergence* of winter rye, with *RMSE* decreasing by more than 4 days. The improvement for *early dough ripening* of maize is close to one day, while the other performance gains are less than 0.5 days. In general, the effect of incorporation of soil moisture on the model accuracy is low, only for two phases it is close to one day or higher (see Table 4).

4.6. Correlation between date of sowing and phenological events

The date on which a plant was sown is an additional, anthropogenic influencing factor on plant phenological development and modeling. To quantify this influence, Pearson's correlation coefficients *r* were calculated between the date of sowing and the DOY_P variants, with and without consideration of soil moisture, as well as between the date of sowing and DOY_{obs} (Table 5).

As expected, DOY_{obs} and the date of sowing are almost linear correlated for *emergence* and show high significance. For *shooting*, *r* is moderate with very high significance, while for all later phases *r* is low. For *heading*, all *r* values are close to zero and not significant, while for the two ripening phases they are slightly higher. In general, correlation between the raw observations is lower than the correlation of the modeled phases and the observed day of sowing. Incorporation of soil moisture slightly increases the correlation coefficient.

4.7. Error sources

Volunteered Geographic Information in general are of limited reliability (Flanagin and Metzger, 2008). Here, for the phenological observations data set, this concerns especially locational uncertainties, cultivar differences and the lack of information regarding fertilization and other management practices (see Sections 2.1 and 4.4) which is one source for model uncertainties.

Due to the spatial interpolation of the observations, uncertainties of the used temperatures are unavoidable. The average uncertainty and an approach to balance accuracy and processing time were discussed in Section 4.1. Other sources for mean temperatures are satellite data, which provide spatially explicit values. However, these values are only valid under cloud-free conditions.

The resolution of the underlying DEM suppresses small-scale variations in elevation and thus reduces interpolation accuracies. The usage of a higher resolution DEM from satellite data e.g. native the *SRTM DEM*, *WorldDEM* (90 m and 12 m resolution, respectively)

Table 4
Difference of *RMSE* between modeling results with and without consideration of soil moisture. Negative values indicate accuracy improvement when soil moisture is considered.

	W. wheat	W. rye	W. barley	Oil. rape	Oats	Maize	Potato	Sug. beet
Emergence	0.55	−4.38	0.49	0.22	0.02	−0.01	−0.18	−0.12
Shooting	−0.03	−0.05	0.03	0.12	0.00	−0.05	–	–
Heading	0.27	0.23	0.02	–	−0.11	–	–	–
Milk ripening	0.09	–	–	–	−0.01	−0.05	–	–
Yellow ripening	0.07	0.1	0.27	–	0.08	−0.35	–	–
Beg. of flowering	–	0.03	–	0.09	–	−0.05	–	–
Full flowering	–	−0.09	–	–	–	–	–	–
4th leaf unfolded	–	–	–	0.88	–	–	–	–
Bud formation	–	–	–	−0.04	–	–	–	–
Full ripening	–	–	–	−0.08	–	–	–	–
Early dough ripening	–	–	–	–	–	−0.94	–	–
Tassel emergence	–	–	–	–	–	0.04	–	–
Closed stand	–	–	–	–	–	–	−0.45	0.14

Table 5
Correlation coefficients *r* and *p*-values for DOY^{0}_{P}, DOY_P and DOY_{obs} against the observed date of sowing for winter wheat.

	Emergence		Shooting		Heading		Milk ripening		Yellow ripening	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
DOY_P	0.98	<2.2e−16	0.49	<2.2e−16	−0.04	0.38	0.28	7.1e−7	0.25	2.8e−6
DOY^{0}_{P}	0.98	<2.2e−16	0.50	<2.2e−16	−0.03	0.56	0.29	1.9e−7	0.26	1.4e−6
DOY_{obs}	0.94	<2.2e−16	0.31	8.6e−8	−0.05	0.30	0.21	1.3e−4	0.15	4.1e−3

or DEM derived from airborne laser scanning with resolutions down to a few centimeters could increase interpolation accuracy, but due to computational limitations study site size and DEM resolution must be balanced.

Further model uncertainties are consequences of the model's intended simplicity. A more sophisticated model that includes parameters like other cardinal temperatures (e.g., upper thresholds, optimum temperatures), which can be freely obtained for many WMO-listed weather stations worldwide, could increase the model accuracy. However, the higher the input data requirements are, the smaller are the application scenarios for the model due to different data availability between regions, countries, phenological observations, etc.

4.8. Comparison with other models

The PHASE model is based on only free and operationally available Germany-wide information on temperature, phenological events and elevation in order to create on-demand maps of up-to-date or past phenological conditions at any location in Germany. On the one hand, the resulting phenological maps represent spatially explicit interpolation results of a specific and adaptable geometric resolution (here: $1 \times 1\ km^2$). This is in contrast to comparable *GDD* approaches applied in Germany which aggregate phenological observation values to reference units like eco-regions (Siebert and Ewert, 2012) or large-scaled grid cells of 50×50 or $100 \times 100\ km^2$ (van Bussel et al., 2011). On the other hand, the thematic depth of the PHASE interpolation results is superior and more crop-specific compared to spatially explicit and satellite-based parameters (e.g. Hird and McDermid, 2009).

Unlike other statistical phenological models, PHASE does not include features and capabilities which might lead to more accurate predictions. Some statistical phenological models include species-specific chilling requirements during the period of dormancy (Luedeling et al., 2009). Consideration of chilling requirements could further improve modeling accuracies for some species, since chilling effects are mostly effective for winter crops but less for summer cultivars.

Furthermore, other models allow the prediction of phenological responses to climate projections into the future (Schröder et al., 2014). This is a very useful capability for a large variety of climate-related sciences, but it is beyond the scope of the PHASE model. One example is the *Promotor Inhibitor Model* (PIM; Schaber and Badeck, 2003) which has recently been applied to investigate the impact of future climate(s) on tree phenology (Lange et al., 2016). PIM is a generalized physiology-based model and fitted on long-term phenological observations, e.g. DWD observational network. Consequently, PIM can assess and explain the importance of underlying drivers like temperature in the sense of chilling requirements or day length which may potentially vary between phenological phases and species. Once robustly fitted, it can also be used for comparable climates or species even when no phenological observations are available. If the underlying mechanism – as exhibited in the long-term observational set – changes, PIM's transferability can be substantially restricted. PHASE, in contrast, uses (geo-) statistics of a single phenological cycle. This ensures for a single year and phase a better predictive performance (2–3 days less MAE compared to PIM).

5. Conclusions and outlook

In this article, we presented the geo-statistical model PHASE which makes the automatic, Germany-wide and spatially explicit prediction of phenological phases possible. The model uses publicly available input data and is characterized by the following features:

- PHASE enables an automatic optimization of all critical parameters which makes the model independent of user-specific parametrization. Thus, the model allows the spatial interpolation of any phase for which phenological observations exist. Each modeling result is characterized by accuracy metrics.
- Only three types of input data are needed including daily provided point data on temperature and phenological events and digital elevation model.
- As demonstrated for the example of soil moisture, additional explaining variables can be integrated in the model to possibly increase the results' accuracy.
- The predictions' geometric resolution is free adaptable. In the presented model setting, the spatial resolution of the resulting phenological raster data set is 1×1 km. Currently, no publicly available phenological model of equivalent performance regarding spatial and thematic resolution as well as processing time is known.

Although the introduced workflow refers to German data situation, the approach is, however, transferable to other temperate regions where a sufficient data base of phenological observations and daily temperature measurements exist or are continuously reported. These requirements are fulfilled in some other European countries, e.g. the Netherlands, Belgium and Great Britain (Rodriguez-Galiano et al., 2015) and the US (Rosemartin et al., 2014). Apart from networks driven by state authorities, the amount of volunteered geographic information including phenological observations has increased in recent years due to the improvement in online communication and positional tracking techniques (Mehdipoor et al., 2015). In addition to human-based observations, stationary near-surface cameras (phenocams) will increasingly provide a permanent visual record of phenological developments (Richardson et al., 2013).

A model transferability is also to be expected to other crop types like soybeans whose growth and development is influenced by heat units and by photoperiod (Setiyono et al., 2007), but which are rarely or almost not cultivated in Central Europe. Depending on the region-specific availability of relevant open-source input data sets, the model can be extended or modified accordingly.

The modeling results are of interest for remote sensing applications, especially for crop classification which has been one of the key applications for remote sensing data over decades. However, the number of distinct cultivated crops that can be discriminated accurately is mostly limited on a few and partly aggregated classes. Class separation on species level is often hampered by high spectral similarity of closely related species e.g., cereal species like wheat, rye and barley. Phenological information can work as indicators for the selection of optimal satellite images to achieve more accurate classification results on species level (Gerstmann et al., 2016).

A further information gain by the usage of PHASE outputs could be achieved in mapping of invasive species. Within the study by Bradley and Mustard (2006), invasive cheatgrass was mapped in northern Nevada, USA, using phenology as indicator for classification and for derivation of invasion risk. For this purpose, Landsat images were selected arbitrary under the knowledge, that cheatgrass has its highest greenness in mid-May, slightly earlier than the grass species which are native in northern Nevada steppes. This arbitrary selection can be assisted and potentially improved by modeled phenology as provided by PHASE.

Turner et al. (2003) stated, that remote sensing based information on plant phenological development are important variables for mapping of biodiversity and determination of vegetation patterns down to species levels. However, remote sensing methods provide a less detailed thematic resolution than phenological information modeled using ground observations. Adapted to the

specific application, PHASE thus has the potential to improve results of biodiversity studies using remote sensing data.

In addition, the coupling of up-to-date phenological predictions with corresponding remotely sensed imagery or simulated times series opens up opportunities for the monitoring of ecosystem services like soil erosion control (Guerra et al., 2014). In doing so, temporal windows and specific reference units like parcels can be detected where soils are potentially covered by sparse or dense vegetation, crop residues or are free of coverage, which enables dynamic soil erosion modeling (Möller et al., 2015).

It is planned to enable public access of operationally produced modeling results within a WebGIS environment. Therefore, for time-consuming parts of the PHASE model, like the prediction of daily temperatures or the optimization of base temperatures (Section 3.2.1) and accumulated effective temperatures (Section 3.2.3), appropriate parallelization techniques (Schiele et al., 2012) are currently tested to enable computational efficient processing.

Acknowledgement

This study was conducted within the research projects *PhenoS* (www.paradigmaps.geo.uni-halle.de/phenos) and *DynaC* (www.paradigmaps.geo.uni-halle.de/dynac), which are funded by the German Ministry for Economics and Technology (BMWi) and managed by the German Aerospace Centre (DLR), contract numbers 50EE1262 and 50EE1230.

References

Barrett, E.C., 2013. Introduction to Environmental Remote Sensing, fourth ed. Routledge.
Berger, M., Moreno, J., Johannessen, J.A., Levelt, P.F., Hanssen, R.F., 2012. ESA's Sentinel missions in support of Earth system science. Rem. Sens. Environ. 120, 84–90.
Bivand, R., Keitt, T., Rowlingson, B., 2015. rgdal: Bindings for the Geospatial Data Abstraction Library. R Package Version 1.0-7.
Bradley, B.A., Mustard, J.F., 2006. Characterizing the landscape dynamics of an invasive plant and risk of invasion using remote sensing. Ecol. Appl. 16 (3), 1132–1147.
Chmielewski, F.-M., Müller, A., Bruns, E., 2004. Climate changes and trends in phenology of fruit trees and field crops in Germany, 1961–2000. Agric. For. Meteorol. 121 (1-2), 69–78.
Chuine, I., Kramer, K., Hänninen, H., 2003. Plant development models. In: Schwartz, M.D. (Ed.), Phenology: An Integrative Environmental Science, Tasks for Vegetation Science, vol. 39. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 217–235.
Deutscher Wetterdienst, 2015b. Vorschriften und Betriebsunterlagen für die phänologischen Beobachter des Deutschen Wetterdienstes. Offenbach, Germany.
Drusch, M., Bello, U.D., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Rem. Sens. Environ. 120, 25–36.
Deutscher Wetterdienst, 2015a. Climate Data Center. URL <ftp://ftp-cdc.dwd.de/pub/CDC/> (last access: 15-03-2015).
DWD Climate Data Center (CDC), 2016. Daily Grids of Soil Moisture under Grass and Sandy Loam, Version 0.x. Tech. Rep.
Estrella, N., Sparks, T.H., Menzel, A., 2007. Trends and temperature response in the phenology of crops in Germany. Glob. Change Biol. 13 (8), 1737–1747.
Ewert, F., Rodriguez, D., Jamieson, P., Semenov, M., Mitchell, R., Goudriaan, J., Porter, J., Kimball, B., Pinter Jr., P., Manderscheid, R., et al., 2002. Effects of elevated CO_2 and drought on wheat: testing crop simulation models for different experimental and climatic conditions. Agric. Ecosyst. Environ. 93 (1), 249–266.
Flanagin, A.J., Metzger, M.J., 2008. The credibility of volunteered geographic information. GeoJournal 72 (3–4), 137–148.
Foerster, S., Kaden, K., Foerster, M., Itzerott, S., 2012. Crop type mapping using spectral-temporal profiles and phenological information. Comput. Electron. Agric. 89 (0), 30–40.
Fu, Y., Zhang, H., Dong, W., Yuan, W., 2014. Comparison of phenology models for predicting the onset of growing season over the northern hemisphere. PLoS ONE 9 (10), e109544.
Gerstmann, H., Möller, M., Gläßer, C., 2016. Optimization of spectral indices and long-term separability analysis for classification of cereal crops using multi-spectral RapidEye imagery. Int. J. Appl. Earth Obs. Geoinf. 52, 115–125.
Grant, C.A., Bailey, L.D., 1993. Fertility management in canola production. Can. J. Plant Sci. 73 (3), 651–670.
Guerra, C.A., Pinto-Correia, T., Metzger, M.J., 2014. Mapping soil erosion prevention using an ecosystem service modeling framework for integrated land management and policy. Ecosystems 17 (5), 878–889.
Hänninen, H., Kramer, K., 2007. A framework for modelling the annual cycle of trees in boreal and temperate regions. Silva Fenn. 41 (1), 167–205.
Hengl, T. 2009. A Practical Guide to Geostatistical Mapping of Environmental Variables. Tech. Rep. European Commission, Joint Research Centre, Institute for Environment and Sustainability.
Hengl, T., Heuvelink, G.B., Rossiter, D.G., 2007. About regression-kriging: from equations to case studies. Comput. Geosci. 33 (10), 1301–1315.
Hense, A., Müller, M., 2007. Geostatistische Modellierung und Qualitätskontrolle von phänologischen Beobachtungen. Promet 33 (1/2), 7–13.
Hiemstra, P., Pebesma, E., Twenhöfel, C., Heuvelink, G., 2009. Real-time automatic interpolation of ambient gamma dose rates from the dutch radioactivity monitoring network. Comput. Geosci. 35 (8), 1711–1721.
Hijmans, R.J., 2015. raster: Geographic Data Analysis and Modeling. R Package Version 2.4-20. URL <http://CRAN.R-project.org/package=raster>.
Hijmans, R.J., Williams, E., Vennes, C., 2014. Geosphere: Spherical Trigonometry for Geographic Applications.
Hird, J., McDermid, G., 2009. Noise reduction of NDVI time series: an empirical comparison of selected techniques. Rem. Sens. Environ. 113 (1), 248–258.
Holen, C., Dexter, A., 1996. A growing degree day equation for early sugarbeet leaf stages. Sugarbeet Res. Ext. Rep. 27, 152–157.
Huang, C.-y., Asner, G.P., 2009. Applications of remote sensing to alien invasive plant studies. Sensors 9 (6), 4869–4889.
Idso, S.B., Jackson, R.D., Reginato, R.J., 1978. Extending the degree day concept of plant phenological development to include water stress effects. Ecology 59 (3), 431–433.
Jamieson, P., Porter, J., Goudriaan, J., Ritchie, J., Van Keulen, H., Stol, W., 1998. A comparison of the models AFRCWHEAT2, CERES-wheat, sirius, SUCROS2 and SWHEAT with measurements from wheat grown under drought. Field Crops Res. 55 (1), 23–44.
Jönsson, A., Eklundh, L., Hellström, M., Bärring, L., Jönsson, P., 2010. Annual changes in MODIS vegetation indices of swedish coniferous forests in relation to snow dynamics and tree phenology. Rem. Sens. Environ. 114 (11), 2719–2730.
Kaduk, J., Heimann, M., 1996. Assessing the climate sensitivity of the global terrestrial carbon cycle model SILVAN. Phys. Chem. Earth 21 (5–6), 529–535.
Kirby, E., Porter, J., Day, W., Adam, J.S., Appleyard, M., Ayling, S., Baker, C., Belford, R., Biscoe, P., Chapman, A., et al., 1987. An analysis of primordium initiation in avalon winter wheat crops with different sowing dates and at nine sites in England and Scotland. J. Agric. Sci. 109 (01), 123–134.
Kramer, K., Leinonen, I., Loustau, D., 2000. The importance of phenology for the evaluation of impact of climate change on growth of boreal, temperate and mediterranean forests ecosystems: an overview. Int. J. Biometeorol. 44 (2), 67–75.
Krige, D.G., 1951. A statistical approach to some basic mine valuation problems on the Witwatersrand. J. Chem. Metall. Min. Soc. S. Afr. 52 (6), 119–139.
Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer.
Lange, M., Schaber, J., Marx, A., Jäckel, G., Badeck, F.-W., Seppelt, R., Doktor, D., 2016. Simulation of forest tree species bud burst dates for different climate scenarios: chilling requirements and photo-period may limit bud burst advancement. Int. J. Biometeorol.
Lee, J.S., 1980. Digital image enhancement and noise filtering by use of local statistics. IEEE Trans. Pattern Anal. Mach. Intell. 2, 165–168.
Löpmeier, F.-J.Deutscher Wetterdienst, 1983. Agrarmeteorologisches Modell zur Berechnung der aktuellen Verdunstung (AMBAV). Dt. Wetterdienst (DWD), Zentrale Agrarmeteorologische Forschungsstelle Braunschweig.
Luedeling, E., Zhang, M., McGranahan, G., Leslie, C., 2009. Validation of winter chill models using historic records of walnut phenology. Agric. For. Meteorol. 149 (11), 1854–1864.
Lunetta, R.S., Knight, J.F., Ediriwickrema, J., Lyon, J.G., Worthy, L.D., 2006. Land-cover change detection using multi-temporal MODIS NDVI data. Rem. Sens. Environ. 105 (2), 142–154.
Masle, J., Doussinault, G., Farquhar, G.D., Sun, B., 1989. Foliar stage in wheat correlates better to photothermal time than to thermal time. Plant Cell Environ. 12 (3), 235–247.
Matérn, B., 1960. Spatial variation, . Lecture Notes in Statistics, second ed., vol. 36. Springer, Berlin.
McMaster, G.S., White, J.W., Weiss, A., Baenziger, P.S., Wilhelm, W.W., Porter, J.R., Jamieson, P.D., 2009. Simulating crop phenological responses to water deficits. In: Ahuja, L.R., Reddy, V.R., Anapalli, S.A., Yu, Q. (Eds.), Modeling the Response of Crops to Limited Water: Recent Advances in Understanding and Modeling Water Stress Effects on Plant Growth Processes. ASA-SSACSSA, pp. 277–300.
McMaster, G.S., Wilhelm, W.W., 1997. Growing degree-days: one equation, two interpretations. Agric. For. Meteorol. 87 (4), 291–300.
McMaster, G.S., Wilhelm, W.W., 2003. Phenological responses of wheat and barley to water and temperature: improving simulation models. J. Agric. Sci. 141 (2), 129–147.
Mehdipoor, H., Zurita-Milla, R., Rosemartin, A., Gerst, K., Weltzin, J., 2015. Developing a workflow to identify inconsistencies in volunteered geographic information: a phenological case study. PLoS ONE 10 (10), 1–14.
Menzel, A., 2007. Phänologische Modelle – phenological models. Promet 33 (1/2), 20–27.

Miller, P., Lanier, W., Brandt, S., 2001. Using Growing Degree Days to Predict Plant Stages. Montana State University, USA (Extension Service).
Möller, M., Gerstmann, H., Thürkow, D., Gao, F., Förster, M., 2015. Coupling of phenological information and synthetically generated time-series for crop types as indicator for vegetation coverage information. In: 8th International Workshop on the Analysis of Multitemporal Remote Sensing Images (Multi-Temp).
Möller, M., Müller, S., Doktor, D., Gläßer, C., 2012. Phenological structuring of multi-temporal RapidEye imagery. In: Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International, pp. 4934–4937.
Nellis, M., Price, K.P., Rundquist, D., 2009. Remote sensing of cropland agriculture. In: Warner, T.A., Nellis, M.D., Foody, G.M. (Eds.), The SAGE Handbook of Remote Sensing, vol. 1. SAGE Publications, London, UK, pp. 368–380.
Pebesma, E.J., 2004. Multivariable geostatistics in s: the gstat package. Comput. Geosci. 30, 683–691.
Peng, C., Guiot, J., Wu, H., Jiang, H., Luo, Y., 2011. Integrating models with data in ecology and palaeoecology: advances towards a model–data fusion approach. Ecol. Lett. 14 (5), 522–536.
Picard, G., Quegan, S., Delbart, N., Lomas, M.R., Toan, T., Woodward, F., 2005. Bud-burst modelling in Siberia and its impact on quantifying the carbon budget. Glob. Change Biol. 11 (12), 2164–2176.
Prishchepov, A.V., Radeloff, V.C., Dubinin, M., Alcantara, C., 2012. The effect of Landsat ETM/ETM+ image acquisition dates on the detection of agricultural land abandonment in Eastern Europe. Rem. Sens. Environ. 126, 195–209.
R Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL <http://www.R-project.org/>.
Rabus, B., Eineder, M., Roth, A., Bamler, R., 2003. The shuttle radar topography mission–a new class of digital elevation models acquired by spaceborne radar. ISPRS J. Photogramm. Remote Sens. 57 (4), 241–262.
Réaumur, M., 1735. Observations du thermométre faites á Paris pendant l'année 1735 comparées avec celles qui ont été faites sous la ligne á l' Ile de France, á Alger et en quelques-unes de nos íles de l'Amérique. Académie Royale Sci., 545–576
Richardson, A., Keenan, T., Migliavacca, M., Ryu, Y., Sonnentag, O., Toomey, M., 2013. Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. Agric. For. Meteorol. 169, 156–173.
Rodriguez-Galiano, V., Dash, J., Atkinson, P., 2015. Characterising the land surface phenology of Europe using decadal MERIS data. Remote Sens. 7 (7), 9390–9409.
Rosemartin, A., Crimmins, T., Enquist, C., Gerst, K., Kellermann, J., Posthumus, E., Denny, E., Guertin, P., Marsh, L., Weltzin, J., 2014. Organizing phenological data resources to inform natural resource conservation. Biol. Conserv. 173, 90–97.
Russelle, M., Wilhelm, W., Olson, R., Power, J.F., 1984. Growth analysis based on degree days. Crop Sci. 24 (1), 28–32.
Salazar-Gutierrez, M., Johnson, J., Chaves-Cordoba, B., Hoogenboom, G., 2013. Relationship of base temperature to development of winter wheat. Int. J. Plant Prod. 7 (4).
Schaber, J., Badeck, F.-W., 2003. Physiology-based phenology models for forest tree species in Germany. Int. J. Biometeorol. 47 (4), 193–201.
Schiele, S., Möller, M., Blaar, H., Thürkow, D., Müller-Hannemann, M., 2012. Parallelization strategies to deal with non-localities in the calculation of regional land-surface parameters. Comput. Geosci. 44, 1–9.
Schröder, W., Schmidt, G., Schönrock, S., 2014. Modelling and mapping of plant phenological stages as bio-meteorological indicators for climate change. Environ. Sci. Eur. 26 (5), 2–13.
Schwartz, M.D. (Ed.), 2006. Phenology: an integrative environmental science. Tasks for Vegetation Science, vol. 39. Kluwer Academic Publishers, Dordrecht, The Netherlands.
Setiyono, T., Weiss, A., Specht, J., Bastidas, A., Cassman, K., Dobermann, A., 2007. Understanding and modeling the effect of temperature and daylength on soybean phenology under high-yield conditions. Field Crops Res. 100, 257–271.
Siebert, S., Ewert, F., 2012. Spatio-temporal patterns of phenological development in Germany in relation to temperature and day length. Agric. For. Meteorol. 152, 44–57.
Sitch, S., Smith, B., Prentice, I.C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J.O., Levis, S., Lucht, W., Sykes, M.T., et al., 2003. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. Glob. Change Biol. 9 (2), 161–185.
Slafer, G., Rawson, H., 1995. Base and optimum temperatures vary with genotype and stage of development in wheat. Plant Cell Environ. 18 (6), 671–679.
Slafer, G., Savin, R., 1991. Developmental base temperature in different phenological phases of wheat (Triticum aestivum). J. Exp. Bot. 42 (8), 1077–1082.
Sobrino, J.A., Julien, Y., Soria, G., 2013. Phenology estimation from Meteosat second generation data. IEEE J. Sel. Topics Appl. Earth Observations Remote Sens. 6 (3), 1653–1659.
Solantie, R., 2004. Daytime temperature sum – a new thermal variable describing growing season characteristics and explaining evapotranspiration. Boreal Environ. Res. 9, 319–333.
Spatialreference, 2015. Catalogs of Spatial Reference Systems. URL <http://spatialreference.org> (accessed on 22nd February 2015)
Statistisches Bundesamt, 2011. Land- und Forstwirtschaft, Fischerei. Bodennutzung der Betriebe (landwirtschaftlich genutzte Flächen). Agrarstrukturerhebung. online. URL <https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/DEHeft_derivate_00004301/2030312117004.pdf> (2016-04-22).
Stein, M., 1999. Interpolation of Spatial Data: Some Theory for Kriging. Springer Series in Statistics. Springer, New York.
Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., Steininger, M., 2003. Remote sensing for biodiversity science and conservation. Trends Ecol. Evol. 18 (6), 306–314.
USGS, 2004. Shuttle Radar Topography Mission.
van Bussel, L.G.J., Ewert, F., Leffelaar, P.A., 2011. Effects of data aggregation on simulations of crop phenology. Agric. Ecosyst. Environ. 142 (1-2), 75–84.
Van Niel, T.G., McVicar, T.R., 2004. Determining temporal windows for crop discrimination with remote sensing: a case study in south-eastern australia. Comput. Electr. Agric. 45 (1-3), 91–108.
Vrieling, A., de Beurs, K.M., Brown, M.E., 2011. Variability of African farming systems from phenological analysis of NDVI time series. Climatic Change 109 (3–4), 455–477.
Wang, E., Engel, T., 1998. Simulation of phenological development of wheat crops. Agric. Syst. 58 (1), 1–24.
Webster, R., Oliver, M., 2007. Geostatistics for Environmental Scientists. Statistics in Practice. Wiley.
Wielgolaski, F.-E., 1999. Starting dates and basic temperatures in phenological observations of plants. Int. J. Biometeorol. 42 (3), 158–168.
Xiao, W., Sun, Z., Wang, Q., Yang, Y., 2013. Evaluating MODIS phenology product for rotating croplands through ground observations. J. Appl. Remote Sens 7 (1), 073562.
Yang, S., Logan, J., Coffey, D.L., 1995. Mathematical formulae for calculating the base temperature for growing degree days. Agric. For. Meteorol. 74 (1-2), 61–74.
Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C., Gao, F., Reed, B.C., Huete, A., 2003. Monitoring vegetation phenology using MODIS. Remote Sens. Environ. 84 (3), 471–475.
Zhao, M., Peng, C., Xiang, W., Deng, X., Tian, D., Zhou, X., Yu, G., He, H., Zhao, Z., 2013. Plant phenological modeling and its application in global climate change research: overview and future challenges. Environ. Rev. 21 (1), 1–14.