AUTOMATIC INTERPOLATION OF PHENOLOGICAL PHASES IN GERMANY

M. Möller, C. Gläßer
University Halle-Wittenberg
Department of Remote Sensing and Cartography
Von-Seckendorff-Platz 4
06120 Halle (Saale), Germany

J. Birger
Geoinformation service Birger
Hoppbergsblick 12
06118 Halle (Saale), Germany

ABSTRACT

The German joint project DeCover 2 is developing a methodological framework to cope with the increasing demand for up-to-date land cover information using remote sensing techniques. New satellite systems like RapidEye provide both data of high geometric resolution and high repetition rates. Because of the Germany-wide diversity of natural conditions, same acquisition dates don’t correspond to same phenological phases. Thus, a phenological structuring of the available imagery over the year is needed for the assessment of RapidEye imagery regarding their suitability for the classification and distinction of vegetation classes.

On the example of the phenological phase ‘Yellow Ripeness’ of Winter Wheat in 2010, the presented algorithm demonstrates for the total area of Germany how daily phenological phases can be automatically interpolated on demand, in real-time and considering interpolation accuracies. As input, daily provided point data on temperature and phenological phases from the extensive network of the German Weather Service as well as a SRTM digital elevation model are used. The modeling results enable the identification of temporal phenological windows for specific test sites.

Index Terms— Phenoogy, temporal window, segmentation, SRTM, DeCover

1. INTRODUCTION

Optical remote sensing data are widely used for the classification of vegetation land use/cover classes. The classification results depends on the vegetation type. The type-specific spectral feature is related to phenological phases and corresponding seasonal meteorological conditions like temperature or day length. A variety of satellite systems make available multi-temporal images enabling the classification of vegetation types. A precondition is the identification of phenological phases during the acquisition dates of satellite imagery [1]. This is particularly true since new satellite systems like RapidEye provide data of high repetition rates allowing a more efficient investigation of environmental monitoring issues.

In the article, a simplified temperature-vegetation model (TVM) is presented which enables an automatic, real-time and large-scale interpolation of phenological phases. The selection of phenological phases results from the availability of observations supplied by the German Weather Service (DWD1). The TVM couples observed and vegetation-type specific ‘Days of the Year’ (DOY) with daily temperatures which are also provided by the DWD.

The investigations are embedded in the national joint project DeCover2 which is the German contribution to European and supranational programmes and systems for regular or on-demand surveys of the earth’s surface (e.g. GMES, GEOSS) [2].

2. METHODS AND DATA

The TVM assumes relations between starting days of phenological phases (DOYp) and phase-specific accumulated temperatures of heat units \( \sum_{i=1}^{n} T_{DOY_{i},p} \) (Eq. (1)).

\[
DOY_p \sim \sum_{i=1}^{n} T_{DOY_{i},p} \tag{1}
\]

Figure 1a illustrates the principle TVM workflow. Accordingly, point data sets of daily temperature (DWD_T) and phenological events (DWD_P) are coupled with segmented and overlaid SRTM elevation data (E) as well as natural units (SRTM_NTR). MODEL1 interpolates Germany-wide daily temperatures (T_{DOY_{1...n}}) and calculates for each segment and each DOY unspecific daily temperature sums (cT_{DOY_{1...n}}). Then, the resulting segment data set (SRTM_{T,CTS}) is transferred to the phenological DWD_P point data set. For each phase and each phenological station, MODEL2 calculates specific sums of daily temperatures

---

1http://www.dwd.de
2http://www.decover.info
Using $cTS_{DOY,P1...n}$ and $cTS_{DOY1...n}$ data, Germany-wide and phase-specific temperature sums ($SUM_{P1...n}$) and DOYs ($DOYP1...n$) are interpolated in MODEL 3. Finally, the interpolated DOYs are assessed by usage of three quality measures.

2.1. Input data

The DWD runs a phenological network of observers who report days of phenological events. The basic network consists of about 1500 observers, covers 167 phenological phases, delivers the observation data at the end of each phase and served as independent reference data set. The training data set results from about 500 observers which provide almost daily information about 101 phases immediately after each observed phenological event. Similar, 500 stations of the DWD meteorological observation network measure and deliver daily temperatures (Fig. 1b).

Natural units are areas which have similar landscape-related features regarding soil, terrain and climate. In the study, the natural unit system by Meynen et al. [3] was used as additional information for the interpolation process. Here, the natural unit similarity is determined by the elevation, annual temperature and precipitation means as well as by similar phenological events (Fig. 1b).

The Shuttle Radar Topography Mission (SRTM) resulted in a world-wide and free available digital topographic model (DTM) with a geometric resolution of $90 \times 90$ m. Because of its roughness, the DTM was filtered [4]. Furthermore, a region-growing segmentation algorithm was applied on terrain attributes and natural units. In this way, groups of pixels are aggregated to segments according to a scale-specific homogeneity of terrain attributes [5]. A positive side-effect of segmentation is the reduction of the data volume which becomes important for large data sets. Here, 68,000,000 raster cells were aggregated to 279,857 segments which can be considered as terrain natural units (TNU). The resulting object data set contains information about TNU-related means of elevation as well as the containment to natural units.

2.2. Modeling

The algorithm was developed within the statistical software environment R [4]. Core components are functions executing import, export, interpolation and row-wise summation which were coupled with the FOR loop function. In addition, statistical quality measures were derived.

The direct import of phenological and meteorological data sets was realized with the functions read.fwf [6] and read.xls [7]. While read.fwf enables the import of data sets with fixed width, read.xls reads Excel files.

The function randomForest [8] represents a regression- and ensemble-based decision tree algorithm (DTA). DTA split the feature space of explaining variables (here: elevation, temperature, temperature sums) or thematic attributes...
(here: natural units) until the resulting partitions show the best statistical correlation. The randomForest-algorithm generates a large number of bootstrapped trees (ensembles) whereas the best tree with the best overall regression is used for the final interpolation [6].

The function rowSums allows a row-wise summation whereas the starting and ending point of summation can be defined. By the coupling of rowSums with the FOR loop function, the begin and end of summation can be determined dynamically for each station and phase (see MODEL 2).

The quality measures 'correlation coefficient' r, 'Nash Sutcliff coefficient' NSK and 'Kolmogorov-Smirnov distance between two distributions' D were calculated using the functions cor.test, ks.test and nashS.

3. RESULTS

The results are presented on the example of the phenological phase 'Yellow Ripeness' of 'Winter Wheat' in 2010. Figure 2a visualizes the interpolated DOYs. Accordingly, the spatial DOY distribution represents typical meteorological conditions in Germany. The phase begins in the central and south western part of Germany and continues in eastern direction. The latest DOY can be found in the north, eastern and south part.

The comparison of observed and interpolated DOYs was made by using both the training (N = 169) and the reference data set (N = 773, Fig. 2a). Despite the small number of training (TR) and reference data (RE) in comparison to the total data set (N = 279857), the 'Kolmogorov Smirnov' D-value indicate with 0.24 a relative small difference between observed and simulated empirical cumulative distributions (Fig. 2b). The quality measures show high statistical correlations for the training data set (r = 0.94, NSK = 0.65). The correlation between the reference and interpolation result is also significant but weaker (r = 0.42, NSK = 0.16). A key reason for this is related to the interpolation of Germany-wide and phase-specific temperature sums (SUM_{P1,...,Pn}, MODEL 3) where the explaining variance is only 15%. All other interpolations in MODEL 1 and 3 show significant higher explaining variances with over 90%.

Figure 3 illustrates on the example of different phenological phases of Winter Wheat how temporal windows can be specified for a test site (see Fig. 2a). The y-axis represents DOYs, the hatched areas phenological windows. For the test site, five RapidEye images and some Winter Wheat reference information were available. The black boxplots indicate the temporal and phenological position of the RapidEye images. The boxplot medians represent the corresponding NDVI value distributions and medians. The temporal comparison of NDVI medians reveals a typical course of the year: The values are high during the phases Shooting and Heading. They become quite smaller during the phases Yellow Ripeness and Harvest.

Fig. 2. Interpolation result for of the phenological phase 'Yellow Ripeness' of 'Winter Wheat' for 2010 (a) and corresponding quality measures (b) (DOY = Day of the year).

4. CONCLUSION

The national joint project DeCover 2 is developing a methodological framework for the spatial and thematic updating of existing land use/cover data sets. The updating is based on remote sensing data (here: RapidEye) representing a high spatial and temporal availability. Against this background, the knowledge of up-to-date phenological phases is of importance for the identification of appropriate temporal windows for vegetation discrimination with remote sensing.

The presented study demonstrates how daily phenological phases can be interpolated from meteorological and phenological data supplied by the extensive network of the German Weather Service. DOYs representing 101 specific phenological phases were modeled using R functions for the total area of Germany in an automatic manner. The modeling results enable the identification of temporal windows for specific test sites and are the basis for a large-scale assessment of RapidEye imagery regarding their suitability for the classification and distinction of vegetation classes.

5. REFERENCES


