Supplementary Information

Mapping cropland-use intensity across Europe using MODIS NDVI time series

Stephan Estel ¹*, Tobias Kuemmerle ¹,², Christian Levers ¹, Matthias Baumann ¹, Patrick Hostert ¹,²

¹ Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
² Integrative Research Institute on Transformations in Human-Environment Systems (IRI THESys), Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

* Corresponding author.

Tel: +49 (30) 2093 9341; Fax: +49 (30) 2093 6848

E-mail address: stephan.estel@geo.hu-berlin.de
S1. Study area

Figure S1: Study area, consisting of 19 MODIS tiles, and the cropland extent derived from the GlobCORINE 2005 map. We focused on all classes containing cropland, specifically rainfed and irrigated cropland, complex cropland, and mosaic cropland / natural vegetation, and we masked all other classes.

S2. Calculation and sensitivity analyses of cropping indicators

Cropping frequency

The cropping frequency was calculated using 12 annual, binary maps of fallow active farmland from 2001 to 2012 derived in a previous study (Estel et al. 2015) (Fig. S2). These maps were classified using a geographically well-distributed training dataset for each year and a Random Forests classifier. The validation for each of the 12 active/fallow maps based on independent observations from the field (i.e. Land Use/Land Cover Area Frame Survey; LUCAS) and from satellite images (i.e. Landsat). Overall we collected for each year on average 438 validation points for the fallow and 1870 points for the active class. The overall accuracy of the fallow/active farmland maps were on average 90%. The active farmland class had a higher accuracy, with a producer’s accuracy of 92% on average and a user’s accuracy
96% on average. The fallow class had a producer’s accuracy of 83% on average and a user’s accuracy of 70% on average (see for details Estel et al. 2015).

We calculated the cropping frequency by summarizing the 12 annual maps and counting the number of active years during that time period for each pixel.

\[
\text{Cropping frequency} = \sum_{i=1}^{12} y(i)
\]

Figure S2: Annual maps of fallow and active farmland across Europe from 2001 to 2012, derived from MODIS NDVI time series at a spatial resolution of 231.6 m.
**Multi-cropping**

The multi-cropping indicator was derived by counting the number of annual seasons over the entire time period using the software TIMESAT (Jönsson and Eklundh 2004). The determination of the number of seasons per year based on the base level and the amplitude ratio between the primary and the secondary peak of the two seasons. We mapped for each year in our time series (2001-2012) the number of seasons (single or double-cropping) maps and summarized the number of years with double cropping.

\[
Multi\ cropping = \frac{1}{2} \times \sum_{i=2}^{24} s(i)
\]

The data base for the determination of the annual season was the smoothed and normalized NDVI time series (see for details Estel et al. 2015). The normalization procedure allows the comparison of vegetation phenology all over Europe despite the strong climate gradient and the different timing of green-up and peak vegetation (e.g., shifted vegetation peak in the Mediterranean region, higher seasonality in the North).

**Fallow cycles and fallow cycle index**

The derivation of the fallow cycles based on the 12 annual, binary maps of fallow active farmland from 2001 to 2012 derived by Estel et al.2015, and described above (Fig. S2). We screened the twelve year time series for three different fallow cycle types (FC1, FC2, and FC3). Each cycle type was built by chain segments; a certain number of consecutive fallow years enclosed from two active years. Thus a chain segment of FC 1 consists one fallow year, a chain segment of FC2 two consecutive and a chain segment of FC3 three consecutive fallow years enclosed by active years. We then summarized the total number of chain segments within the entire time series for each fallow cycle type (Fig S3).
Figure S3: Show the total number of chain segments for the three fallow cycle types (FC1, FC2, and FC3). The maximal number of chain segments for FC1 in a twelve year time series is five, for FC2 three and FC3 can consist only two chain segments.

To provide a single indicator for fallow systems and a measurement for the cropland intensity within these fallow systems we attached different weights to the three fallow cycle type. The weights were derived by calculating the ratio of the total number of years in the time series (in our case twelve years) and the maximal possible number of chain segments within the time series. In a twelve year time series FC1 can occurred five times (5/12), FC three times (3/12) and FC3 only two times.
\[
Fallow \ cycle \ index = \sum_{i=1}^{5} \left( FC1 \times \frac{5}{12} \right) + \left( FC2 \times \frac{1}{4} \right) + \left( FC3 \times \frac{1}{6} \right)
\]

Since a single chain segment does not indicate cyclicity or cropland intensity within fallow systems, we considered only time series with at least two chain segments from the same fallow cycle type or combinations of different cycle types.

**Crop Duration Ratio**

The crop duration ratio refers to the relationship between the total length of the growing season (\(L_0\)) and the length of the cropping season at half of the highest peak (i.e., half-maximum, \(L_{50}\)) during the cropping time (Fig. S4). The total length of the growing season was determined by counting the number of days with a land surface temperature above 5°C, i.e., the time between the earliest and latest MODIS acquisition date in a given year when plants are assumed to actively grow (Hickler et al. 2012; Zhang et al. 2004). We derived the crop duration ratio from \(L_0\) and \(L_{50}\) for each non-fallow year between 2001 and 2012 (Fig. S5) and calculated the mean across all years.
Figure. S4: Calculation of Crop duration ratio (CDR) using the total length of the growing season ($L_0$) and the 50% of the peak vegetation ($L_{50}$).

\[
 CDR = \frac{L_{50}}{L_0}
\]

To assess the sensitivity of our results towards the threshold chosen to map the crop duration ratio (the half maximum, 50% peak vegetation threshold, in the default calculation), we derived crop duration ratio maps for 40%, 45%, 50%, 55%, and 60% thresholds for each year (Fig. S4), calculated the average crop duration ratio per pixel, and then derived the standard deviation crop duration ratio per pixel (Fig. S5). Our results showed that the crop duration ratio was relatively robust to the choice of threshold for most cropland areas in Europe (i.e., variation of less than 20%). For some areas, especially in the Mediterranean, where crop duration ratio is lower than in other areas in Europe, we found a higher, but still fairly moderate sensitivity (e.g., in Turkey, on Crimea Fig. S5).
S3. Self-Organizing Maps parameterization

Determination of the optimal cluster number

To determine the optimal number of cluster for the SOM analysis, we carried out a sensitivity analyses with varying clusters numbers and dimensionality (i.e., columns and rows) ranging from 2x2 to 4x4. To identify the optimal cluster number we observed the natural breakpoint in the mean distance of the samples to their cluster centroid (Maulik and Bandyopadhyay 2002) and the Davies-Bouldin cluster validity index which explore the intra- and inter-cluster variability (Davies and Bouldin 1979). A low Davies-Bouldin index indicates a mathematically more satisfactory clustering result. In our case the minimum Davies-Bouldin value was at six. Since the mean distance to the cluster centroid was levelling off around a cluster size of six as well we set the optimal number of cluster to six (Fig S6).
Figure S6: Determination of the optimal cluster number using the mean distance of samples to their cluster centroid and the Davies-Bouldin cluster validity index.

Table S1: Shows the z-score values for each indicator.

<table>
<thead>
<tr>
<th>Z-score</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cropping frequency</td>
<td>0.82</td>
<td>0.72</td>
<td>-1.78</td>
<td>0.85</td>
<td>0.09</td>
<td>0.00</td>
</tr>
<tr>
<td>Multi-cropping</td>
<td>0.12</td>
<td>2.76</td>
<td>-0.80</td>
<td>0.00</td>
<td>-0.53</td>
<td>0.44</td>
</tr>
<tr>
<td>Fallow cycle</td>
<td>-1.25</td>
<td>-0.67</td>
<td>0.73</td>
<td>-1.23</td>
<td>0.73</td>
<td>0.74</td>
</tr>
<tr>
<td>Crop duration ratio</td>
<td>0.22</td>
<td>0.80</td>
<td>0.36</td>
<td>-1.62</td>
<td>-0.35</td>
<td>1.00</td>
</tr>
</tbody>
</table>

References


potential natural vegetation zones with a generalized, tree species-based dynamic vegetation model. 

*Global Ecology and Biogeography, 21, 50-63*

