

# Time Series Analysis

Philipp Satlawa

December 16, 2021

## Abstract

Currently, satellites with mapping sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS) provide an excellent source for monitoring global vegetation activity. With time-series data of the normalized vegetation index (NDVI) we investigate phenology indicators and anomalies for the last decade in the country of Georgia. Furthermore, we examine the differences in phenology of distinct land cover types. Finally, we detected anomalies in particular years at the start of the season (SOS).

## 1 Introduction

Phenology is the examination of the annual biological processes for example the leaf-out and flowering timings. To study the phenology of large areas such as countries, continents or the entire globe we need to exploit remote sensing data from space. Satellites such as Moderate Resolution Imaging Spectroradiometer (MODIS) have provided continuous data over the last two decades and thus given the opportunity to examine vegetation for the whole globe over a long period.

Especially the time series of normalized difference vegetation index (NDVI) has proven to be very convenient for monitoring climate change [[Shabanov et al., 2002](#)], change detection [[Fraser et al., 2000](#)], drought monitoring [[Klisch and Atzberger, 2016](#)], and classification of crops [[Vuolo and Atzberger, 2012](#)].

In this report, we are investigating the penology of Georgia over the last decade. In particular, we are deriving the phenological indicators start of season (SOS), maximum of season (MOS) and end of season (EOS) as well as the variable length of season (LOS) from the NDVI time series generated by MODIS. Moreover, we analyse the data for anomalies and compare the phenology of various land cover types with each other.

## 2 Materials

### 2.1 Study area

Located on the natural boundary between Eastern Europe and Western Asia, Georgia is mainly enclosed by the Greater Caucasus in the north and by the Lesser Caucasus mountain range in the south. Georgia has an area of 69.875 km<sup>2</sup> and its elevation stretches from sea level in the lowlands of eastern Georgia to over 5000 m above sea level in the Greater Caucasus mountains (Figure 1c).

Georgia has despite its relatively small area, a very versatile topography and climate. The eight distinct climatic zones according to the Köppen-Geiger classification (Figure 1e) are mostly determined by the altitude as well as by distance to the sea ([[Curtis, 1995](#)]).

The Likhi Range connects the Greater Caucasus and the Lesser Caucasus mountain ranges thus functions as a natural border between the lowlands and valleys of western and eastern Georgia. While the western part of Georgia is the drains into the Black sea while the eastern part functions as the watershed for the Caspian sea.

The plain in the west ranging from the Black Sea coastline to the inland region called Colchis Lowland (Figure 1e western Cfa region), features a subtropical climate with high humidity and hefty precipitation 1000 to 2000 mm (Figure 1d) due to the unrestricted influence of the Black sea.

In contrast, the eastern plain of Georgia is protected by mountains from the immediate effects of the Caspian Sea and hence has a more continental climate with lower humidity and precipitation

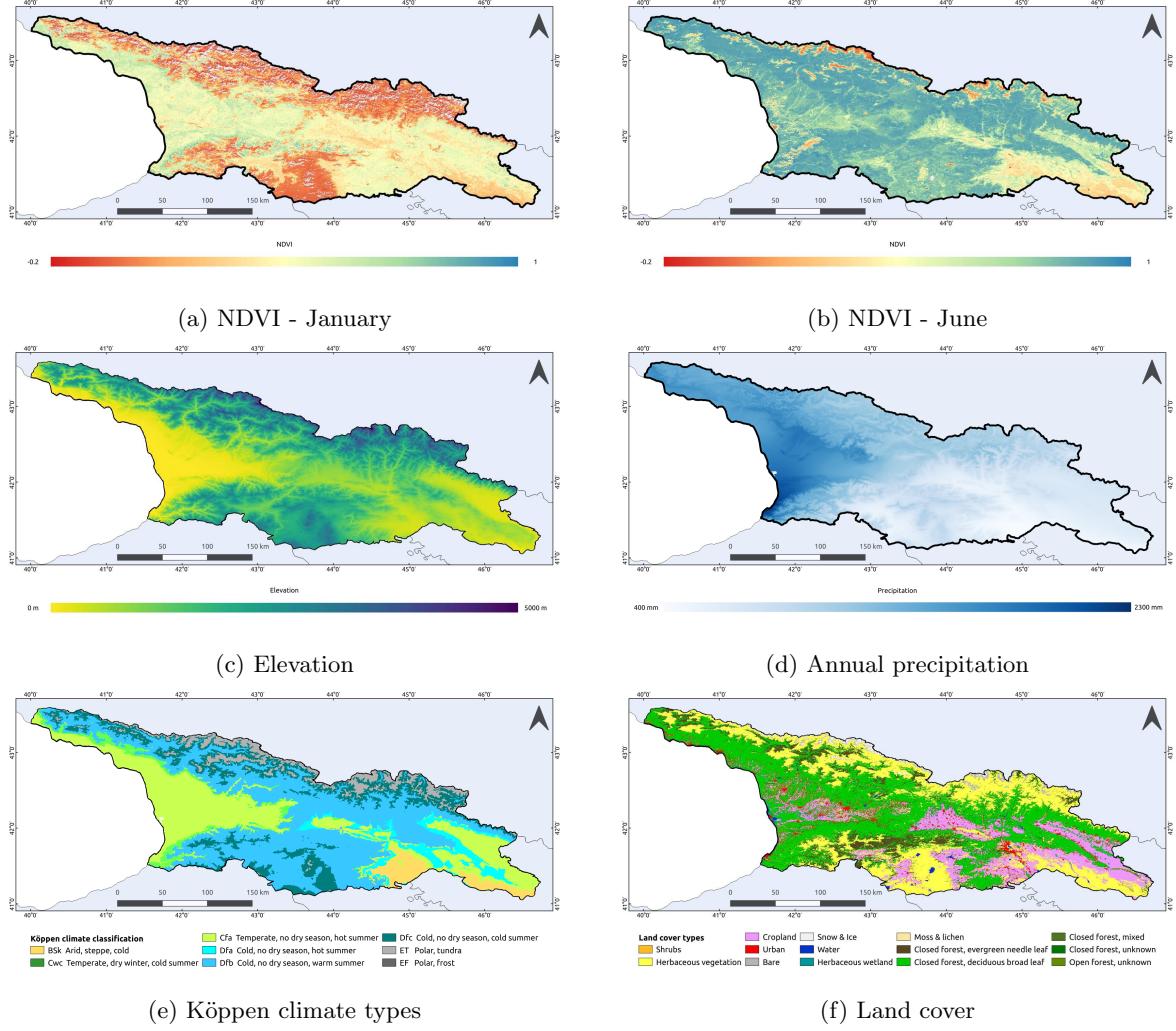


Figure 1: Environmental conditions of Georgia. a) NDVI in January [Vuolo et al., 2012], b) NDVI in June [Takaku et al., 2020], c) Elevation [Harris et al., 2014] and [Fick and Hijmans, 2017], e) Köppen-Geiger climate classification [Beck et al., 2018], f) Land cover [Buchhorn et al., 2020].

compared with the western part (Figure 1d). In the south-eastern region of Georgia the precipitation is at its lowest and the climate is turning into a semi-arid (Figure 1e BSk region).

With the increase in altitude in both north and south of Georgia, the climate changes gradually into a colder and wetter climate. At elevations above about 2100 m, alpine climate conditions are present.

The land cover is comprised mainly of *forest*, *cropland* and *herbaceous vegetation* (Table 1) with agriculture mainly being present in the lowlands and forests being present in the steeper mountainous areas (Figure 1f).

## 2.2 Data

The Moderate Resolution Imaging Spectroradiometer (MODIS) mounted on the two satellites Terra and Aqua provides remote sensing data of the entire world every 1 to 2 days. This temporal resolution allows the creation of MOD13Q1 and MYD13Q1 collection 5 products which are 16-day NDVI composites of the two satellites. The shifted acquisition time allows deriving a global NDVI with an improved temporal frequency of eight days in a spatial resolution of 250 m. The MODIS 16-day NDVI composites for the years 2009 until 2021 was downloaded from the website of the Institute of Surveying, Remote Sensing and Land Information (IVFL) of the University of Natural Resources and

Table 1: Land cover types prevalent in Georgia in percent of total area.

<b>Code</b>	<b>Description</b>	<b>Area [%]</b>
20	Shrubs	0.5
30	Herbaceous vegetation	28.0
40	Cropland	15.8
50	Urban	1.8
60	Bare	1.2
70	Snow and Ice	0.8
80	Water	0.4
90	Herbaceous wetland	0.3
111	Closed forest, evergreen needle leaf	4.6
114	Closed forest, deciduous broad leaf	32.9
115	Closed forest, mixed	1.6
116	Closed forest, unknown	5.0
126	Open forest, unknown	7.1

Applied Life Sciences (BOKU) ([Vuolo et al., 2012]) in the raw data format.

The land cover data was obtained from the Copernicus Global Land Service in a 100m spatial resolution (Moderate Dynamic Land Cover v3) for the year 2019 ([Buchhorn et al., 2020]). Comprised of 23 classes, the land cover map is defined by the Land Cover Classification System (LCCS). The classes relevant for Georgia are listed in Table 1.

## 3 Methodology

### 3.1 data pre-processing

Since the MOD13Q1 and MYD13Q1 collection, the DSM and land cover data had various coordinate reference systems (CRS) we had to align the data before any further processing. Thus all data was aligned to the MOD13Q1 and MYD13Q1 collection with the CRS EPSG:4326 and a spatial resolution of 250m using the nearest neighbour resampling.

Before smoothing, an array of the size 1252x1826x4745 was created. The array shape was determined by the image resolution (1252x1826) and the temporal resolution consisting of the number of days in the 2010 to 2020 period (4745). The individual NDVI images were then assigned to the corresponding days they represent in the 4745 days in the previously created array. This is a necessary step before processing the data with the Whittaker smoother.

### 3.2 Whittaker smoothing

For smoothing the time series in this project we employed the Whittaker smoother due to its fast computation and a good performance against other curve fitting techniques ([Atkinson et al., 2012]). The Whittaker smoother smooths the provided temporal data by applying a penalty on the roughness of the smooth curve as proposed by [Eilers, 2003], which is an adaptation of the original Whittaker's smoother [Whittaker, 1922]. The adaptation of the smother comprises the addition of weights  $w$  and the change of the penalty from the sum of the squares of the third difference to the  $\lambda$  parameter.

In particular the Whittaker smother minimises the loss function:

$$S = \sum_i w_i (y_i - z_i)^2 + \lambda \sum_i (\Delta^2 z_i)^2$$

where  $y$  is a vector with the time series data,  $w$  is a vector with the same length as  $y$  with 0 indicating no data and ones indicating valid data,  $z$  is a vector with the same length as  $y$  with the smoothed data the aim is to obtain vector  $z$  that is similar to  $y$  and additionally interpolates the missing values.

As used in [Atzberger and Eilers, 2011] the Whittaker smoother uses the  $\lambda$  parameter to steer the magnitude of the smoothing. Having a method for computationally obtaining a suitable  $\lambda$  is

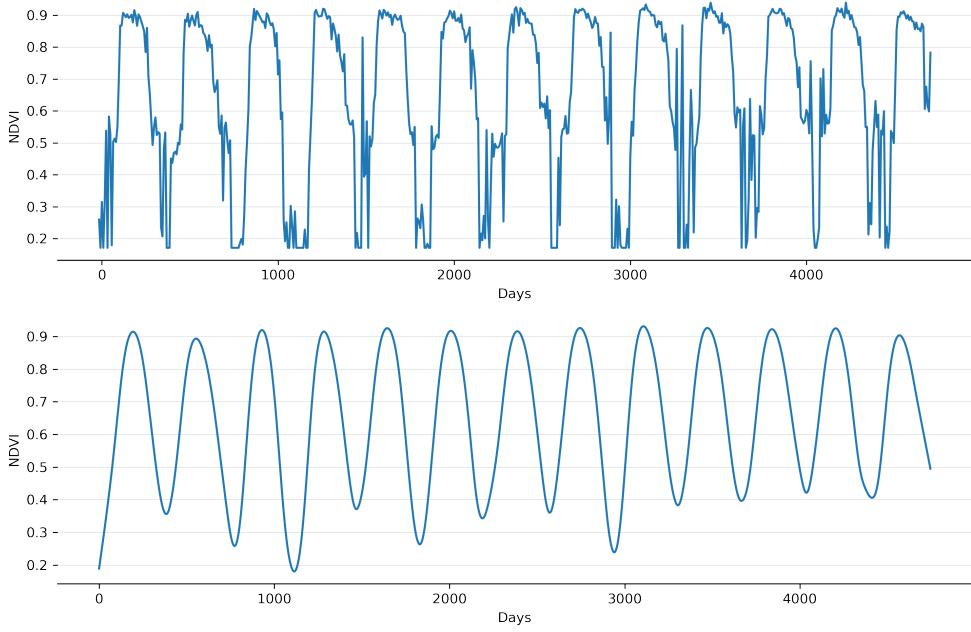


Figure 2: Example of a NDVI time series pixel before (top plot) and after (bottom plot) filtering. The pixel was smoothed by applying the Whittaker smoother with an computationally determined *lambda*.

very desired since it facilitates the smoothing process. [Eilers et al., 2017] showed that an automatic estimation of *lambda* by exploiting the V-curve.

Accordingly, we applied the Whittaker smoother with the automatic *lambda* estimation implemented in Python by [Pesendorfer, 2021] and obtained smoothed daily NDVI values from 2010 to 2020. An example is provided in Figure 2, where the top plot shows the original MODIS data for one pixel and the bottom plot displays the smoothed NDVI data after being processed by the Whittaker smoother.

### 3.3 Phenology indicators and length of season

In this project, we obtain the phenological indicators, start of season (SOS), maximum of season (MOS) and end of season (EOS) as well as the parameter length of season (LOS). Since there is no uniform definition of SOS and EOS we define SOS as the day in which the NDVI-value exceeds 20% of the amplitude (maximum NDVI - minimum NDVI) in the ascending NDVI part.

In order to access the uni- and bi-modal NDVI profiles correctly an approach consisting of mainly two algorithms were implemented. The first algorithm (Algorithm 1) extracts the first relevant maximum of the growing season. The result is fed into the second algorithm (Algorithm 2) that searches for the relevant minimum, computes the amplitude and extracts the SOS. So that only the metrics of the first season are extracted for bi-modal NDVI profiles.

Similarly to the SOS the EOS was derived as the day in which the NDVI-value falls the 20% of the amplitude (maximum NDVI - minimum NDVI) in the descending NDVI part. The algorithms for extracting the EOS is analogous to the algorithms extracting the SOS and can be found in the GitHub repository [Satlawa, 2021].

The computation of MOS was already determined by Algorithm 1, while LOS was calculated by taking the difference between EOS and SOS.

The anomalies for SOS and LOS were extracted by calculating the standard score or z-score. The standard score is defined as follows:

---

**Algorithm 1** Find the first relevant season maximum

---

**Require:** *data* as vector

```
maxima ← find_maxima(data)                                ▷ In-build function to find all local maxima
if maxima is empty then
    maximum ← -1
else if maxima has 1 item then
    maximum ← maxima[0]                                         ▷ [0] means get the first element
else
    global_maximum ← max(data)                                     ▷ In-build function to find global maximum
    for local_maximum in maxima do
        if local_maximum ≥ global_maximum * 0.4 then ▷ The threshold 0.4 was derived through
        trial and error
            maximum ← local_maximum
        end if
    end for
end if
```

---

**Algorithm 2** Find start of season (SOS)

---

**Require:** *data* as vector

```
maximum ← get_maximum(data_all)                               ▷ Calling Algorithm 1
minimum ← min(data<max)                                    ▷ In-build function to find global minimum
amplitude20 ← (maximum - minimum) * 0.2 + minimum      ▷ calculate 20% of NDVI amplitude
sos ← find_index(datamin<max, amplitude20)           ▷ find index of specific value
```

---

$$ZVI = \frac{NDVI - NDVI_{mean}}{NDVI_{std}}$$

### 3.4 Computation

All processes including pre-processing, smoothing extraction of phenology indicators as well as graphs were written in Python utilizing the libraries numpy, gdal, matplotlib, h5py and scipy. The code can be found in the GitHub repository [[Satlawa, 2021](#)]. Maps of Georgia were created with QGIS v3.16. All data was processed on a PC with Ubuntu 20.04 and a Ryzen 5600X with 32GB of RAM.

## 4 Results

### 4.1 Start of season in Georgia

Examining the mean start of season (SOS) in Figure 5, we can deduce that the earliest SOS can be detected in the southeastern part of Georgia that possesses a semi-arid climate (Figure 1e). Analysing further Figure 1d we find that this region some of the lowest annual precipitation of Georgia. Adding the continental climate with higher summer temperatures the growing season in this region is most likely limited by the water availability. Vegetation in such climates has an earlier and lower maximum of season (MOS) as well as a shorter length of season (LOS) (Figure 6), thus moving the defined SOS to an earlier date.

Another striking phenomenon is the later SOS in higher elevations. To investigate this relationship the data was restricted to the land cover type *forest* since it contains vegetation and occurs in a variety of elevations in contrast to the other land cover types that are very elevation specific. Figure 4a shows the correlation between the mean SOS of the land cover type *forest* and the elevation. In general, the relationship between SOS and elevation can be considered moderate due to the coefficient of determination ( $R^2$ ) of 0.58. The moderate correlation can be most probably attributed to the different expositions between the South Caucasus and the North Caucasus as well as noise.

Investigating the progression of the SOS from 2010 to 2020 (Figure 5) we find extreme anomalies in the year 2011 where the SOS of entire Georgia was postponed on average by 18(±14) days. In

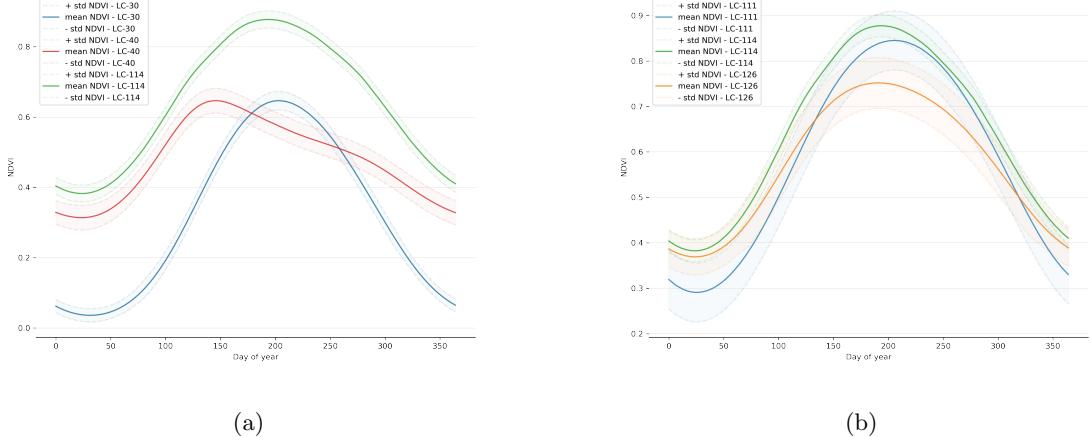


Figure 3: Temporal NDVI profile of land cover types. The graphs present the mean NDVI (opaque line) and the standard variation (translucent area around the line) for several land cover types. a) Temporal NDVI profile of *Closed forest, deciduous broad leaf* (green), *Cropland* (red) and *Herbaceous vegetation* (blue). b) Temporal NDVI profile of *Closed forest, deciduous broad leaf* (green), *Closed forest, evergreen needle leaf* (blue), *Open forest, unknown* (orange).

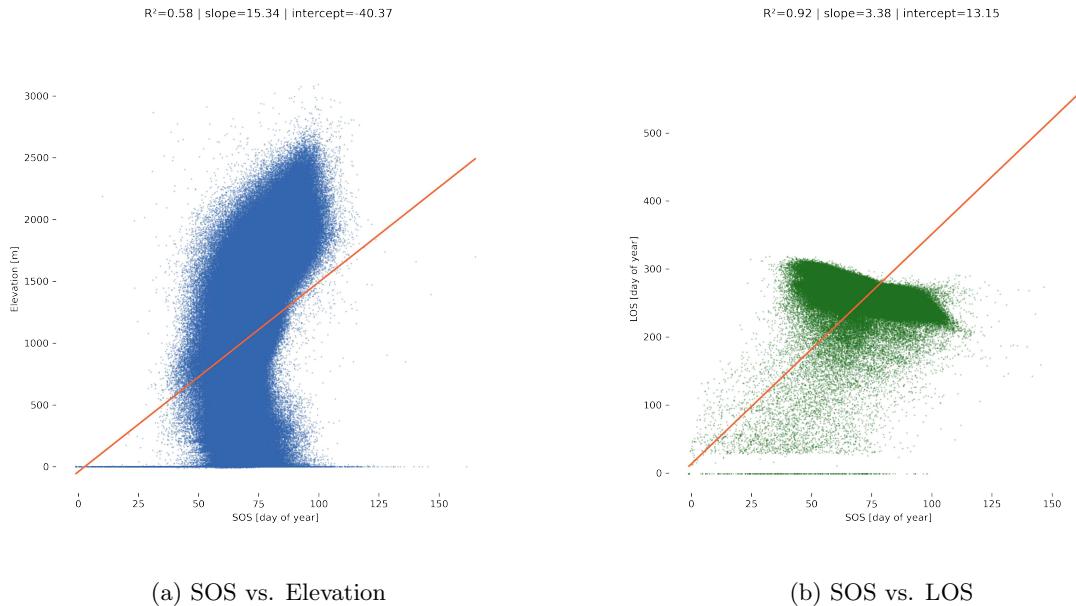


Figure 4: Correlations of start of season (SOS) with Elevation and length of season (LOS). a) Correlation between mean SOS for land cover type *forest* (including all forest types 111 to 126, see Table 1) and the elevation. b) Correlation between mean SOS and the mean LOS, both for land cover type *forest* (including all forest types 111 to 126, see Table 1).

contrast to that, we see an early SOS in 2018 where the SOS started on average  $12(\pm 12)$  days earlier than the mean SOS. Summarizing the findings, we can disclose that the SOS differed in the extreme by an entire month in the period 2010 to 2020.

## 4.2 Length of season in Georgia

Displayed in the upper left in Figure [Figure 6](#) we see the mean LOS for Georgia. Analogous to the mean SOS we can observe shorter LOS in the southeastern region with the semi-arid climate. A second area in the middle of Georgia has a similarly short LOS while being classified as *cropland*. Hence it is very likely that in that region crop is cultivated just once per season and the crop has a short growing period. In contrast to the postponed SOS at higher altitudes, the shortening of LOS by higher altitude is way less pronounced.

Examining the standardized LOS difference for the period 2010 to 2020 we detect 2010 as being a year where most parts of Georgia experienced a much longer ( $27 \pm 27$  days) LOS compared to other years in that decade. This finding is somehow surprising since the standardized SOS difference in [Figure 5](#) showed no pronounced early SOS and the SOS is highly correlated with the LOS ( $R^2=0.92$ , [Figure 4b](#)). Thus if the SOS was not early, the EOS was late that year. Looking in the rightmost column of [Figure 7](#) we can see the late EOS in 2010.

The second year with a pronounced anomaly is 2011 where the LOS was way shorter ( $26 \pm 23$  days) compared to the mean LOS. This anomaly in the LOS coincided with the SOS anomaly in 2011 where the SOS was later than the mean SOS. Finally, in 2018 we have a longer LOS ( $20 \pm 28$  days) that matches with the earlier SOS of the same year.

## 4.3 Phenology of land cover types

[Figure 7](#) show the phenology indicators start of season (SOS), maximum of season (MOS) and end of season (EOS) for the four most prevalent land cover types *Deciduous broad leaf* (33%), *Herbaceous vegetation* (28%), *Cropland* (16%) and *Evergreen needle leaf* (5%) ([Table 1](#)). Generally, we see the same anomalies of SOS over the years in the box plots as in [Figure 5](#).

In particular the *hebaceous vegetation* has predominantly the latest SOS of all the shown land cover types. It is consistent with the fact that this land cover type is prevalent in the higher altitudes where the plants have a shorter photosynthetically active season. The mean temporal NDVI profile in [Figure 3a](#) shows the overall lower NDVI curve compared to the *Evergreen, needle leaf forest* and *Cropland*.

In opposition, the *Cropland* type has the earliest SOS which in turn can be explained by the earlier MOS as compared to the other land cover types. A conspicuousness of *Cropland* is the high variance of the MOS and EOS. This is certainly related to the variation of crop types where diverse crop types have distinct growing patterns. Observing the NDVI-profile in [Figure 3a](#) we can see the *Cropland* profile having an early MOS while afterwards descending slowly with two bends. This bends are mos certainly the reminder of the to double cropping profile that was smoothed out. Due to the various crop types grown in Georgia, the variation of the EOS varies greatly.

Finally, the two forest types *Broad, leave forest* and *Evergreen, needle leaf forest* vary just slightly in SOS, MOS and EOS. In general, the *Evergreen, needle leaf forest* type has a slightly later SOS, MOS and EOS as can be also seen in the NDVI profiles of [Figure 3b](#). This can be explained by the on average higher altitude.

## 5 Conclusion

In this report we were able to compute and analyse the phenology indicators start of season (SOS), maximum of season (MOS) and end of season (EOS) as well as the length of season (LOS) from the Moderate Resolution Imaging Spectroradiometer (MODIS) derived NDVI time series. We showed that the computed phenology indicators give valid results. Furthermore, we analysed anomalies in the period 2010 to 2020 and found:

- 2010 having a very long LOS (+27 days) as well as a late EOS
- 2011 having a shorter LOS (-26 days) and a late SOS (+18 days)

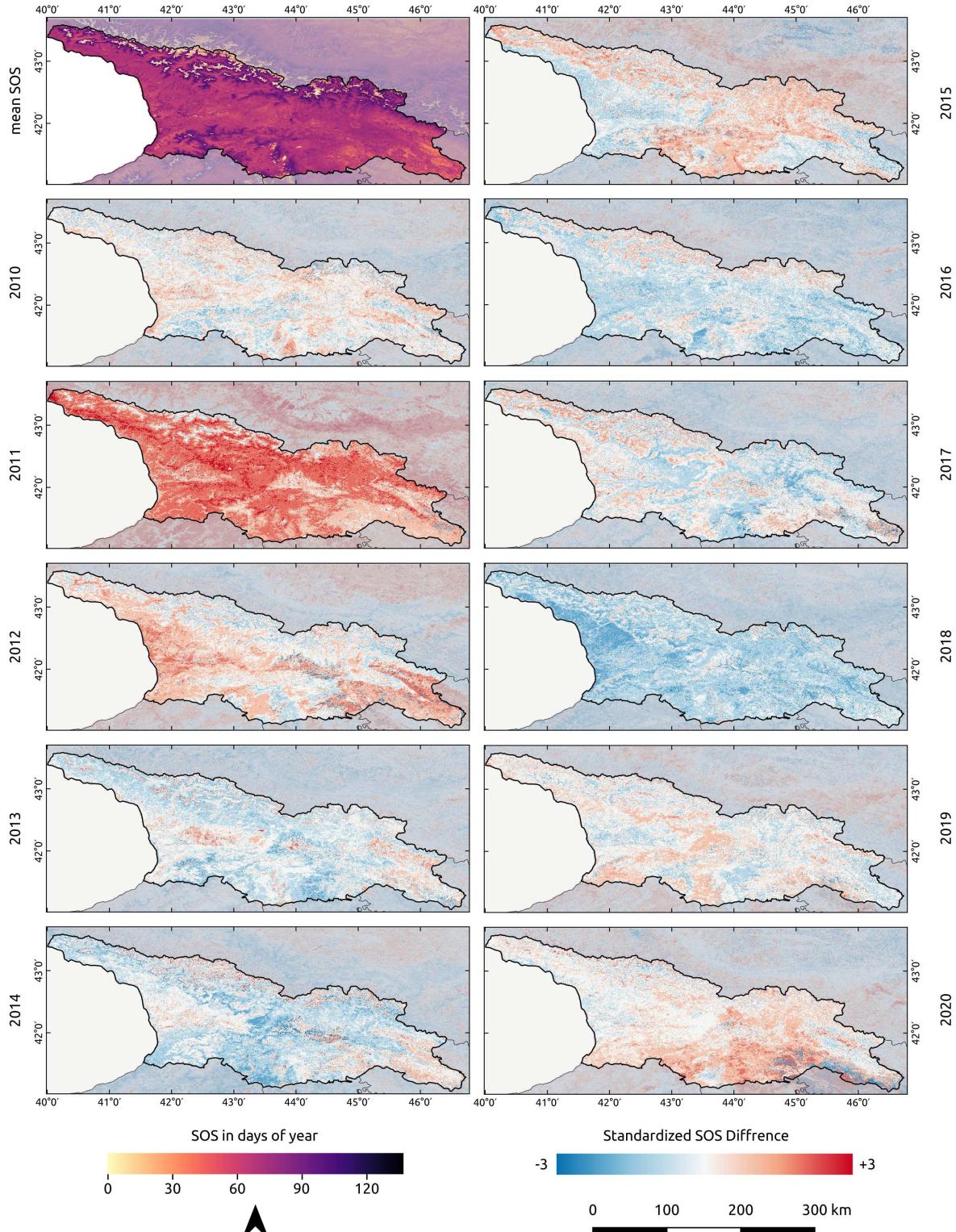


Figure 5: Mean start of season (SOS) and standardized anomalies for Georgia and the years 2010 to 2020. The left uppermost image shows the mean SOS of the period for the period 2010 to 2020. All other images represent the standardized difference (z-score) in standard variations compared to the mean SOS.

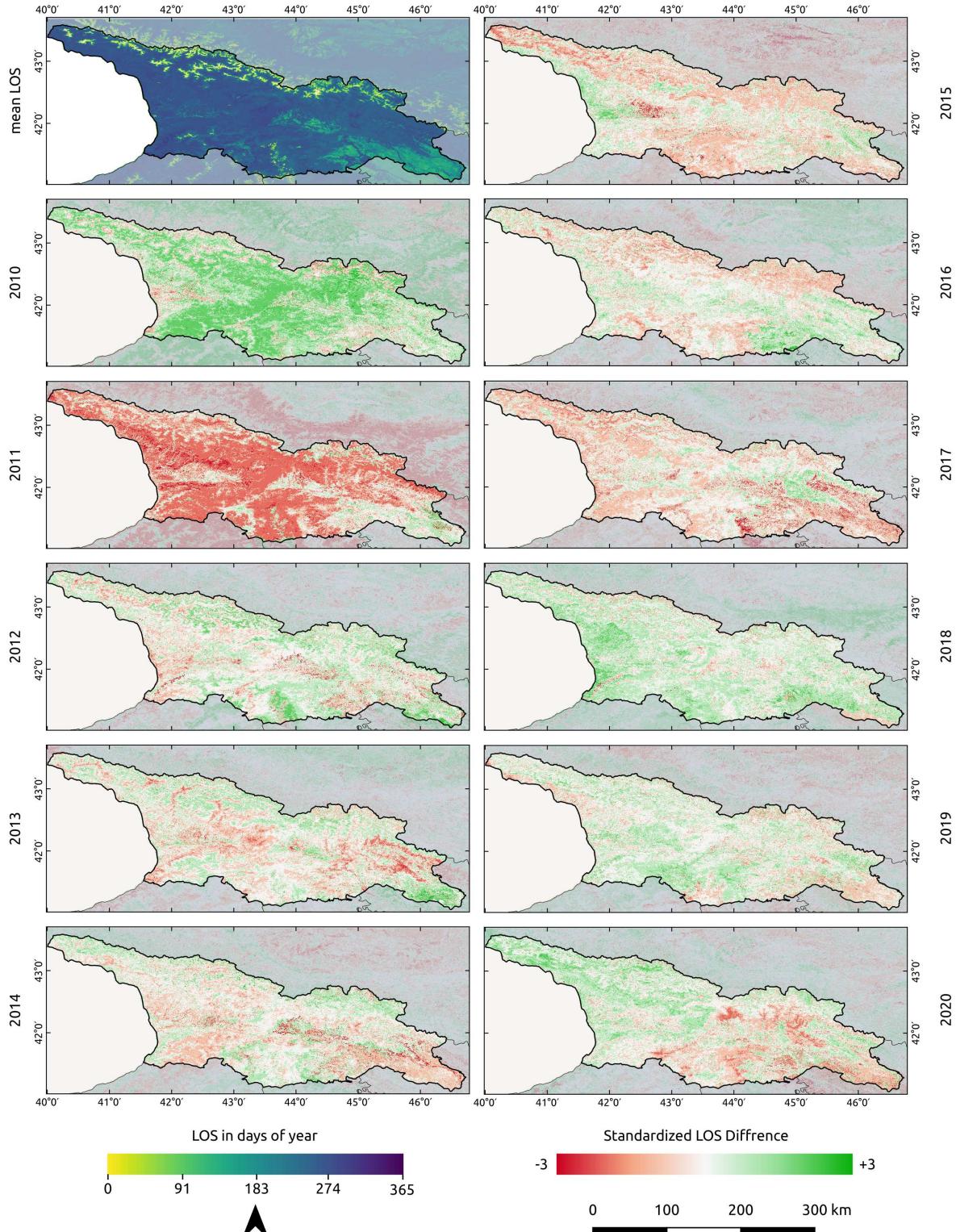


Figure 6: Mean length of season (LOS) and standardized anomalies for Georgia and the years 2010 to 2020. The left uppermost image shows the mean LOS of the period for the period 2010 to 2020. All other images represent the standardized difference (z-score) in standard variations compared to the mean LOS.

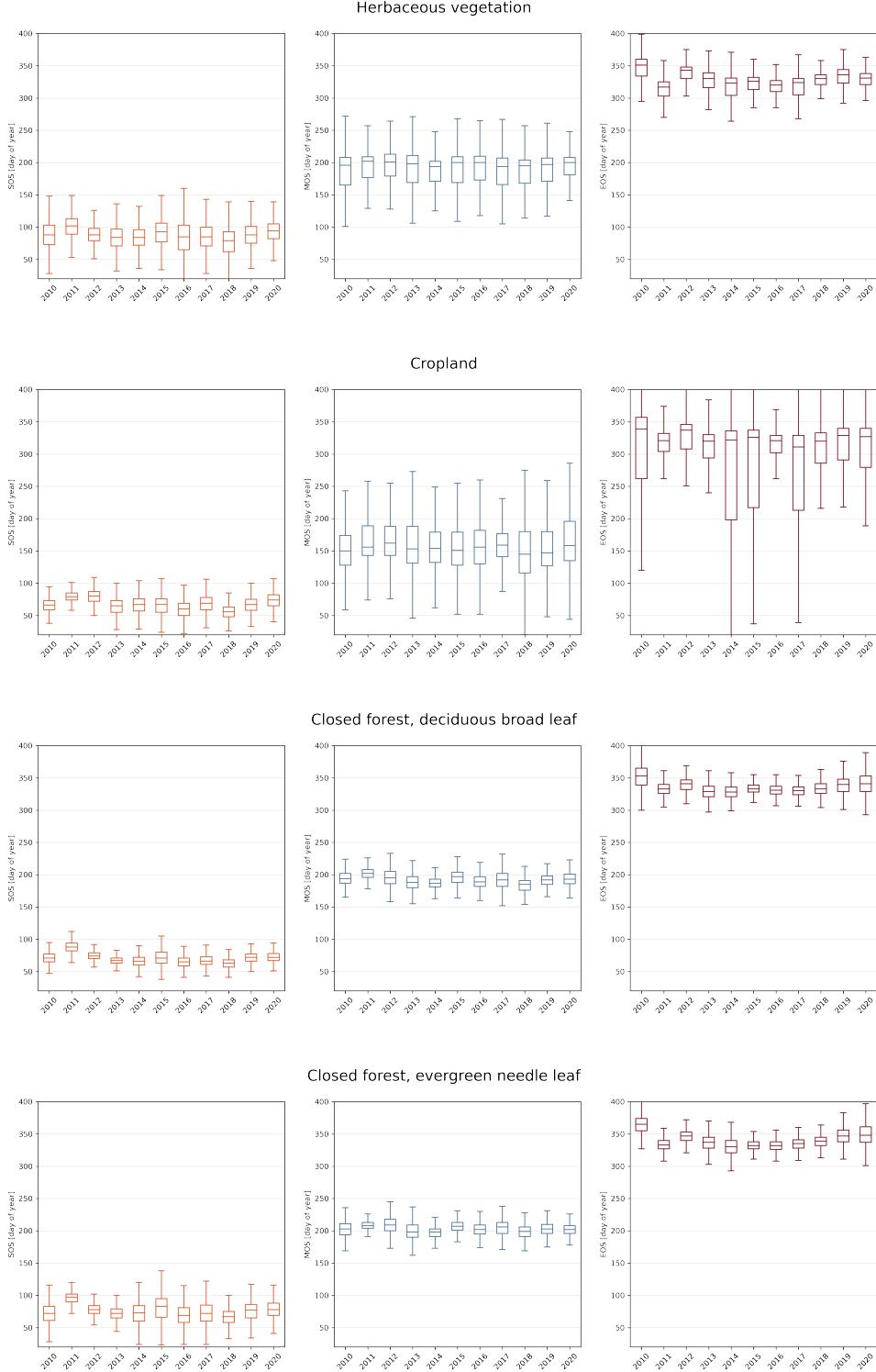


Figure 7: Phenology of the most present land cover types in Georgia. The columns represent the phenology indicators, start of season (SOS), maximum of season (MOS) and end of season (EOS) successively. The rows show the land cover types *herbaceous vegetation*, *cropland*, *deciduous, broad leave forest* and *evergreen, needle leaf forest* from top to bottom. All graphs show the variation of the phenology in a given year (boxes with whiskers) and between between the years 2010 to 2020.

- 2018 having a longer LOS (+20 days) and an early SOS (-12 days)

Moreover, we showed that the phenology indicators of the various land cover types performed according to common sense. This is a first glimpse of the analysis that can be performed with this data. We encourage fellow researchers to investigate further anomalies in phenology in an even longer period to examine the impact of climate change on the vegetation of Georgia and beyond.

## References

- [Atkinson et al., 2012] Atkinson, P. M., Jeganathan, C., Dash, J., and Atzberger, C. (2012). Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. *Remote Sensing of Environment*, 123:400–417.
- [Atzberger and Eilers, 2011] Atzberger, C. and Eilers, P. H. (2011). Evaluating the effectiveness of smoothing algorithms in the absence of ground reference measurements. *International Journal of Remote Sensing*, 32(13):3689–3709.
- [Beck et al., 2018] Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., and Wood, E. F. (2018). Present and future köppen-geiger climate classification maps at 1-km resolution. *Scientific Data*, 5:1–12.
- [Buchhorn et al., 2020] Buchhorn, M., Smets, B., Bertels, L., Roo, B. D., Lesiv, M., Tsednbazar, N.-E., Li, L., and Tarko, A. (2020). Copernicus Global Land Service: Land Cover 100m: Version 3 Globe 2015-2019: Product User Manual. *Copernicus Global Land Operations*, pages 1–93.
- [Curtis, 1995] Curtis, G. E. (1995). *Armenia, Azerbaijan, and Georgia : country studies*. Addison Wesley, Washington, D.C.: Federal Research Division, 1 edition.
- [Eilers, 2003] Eilers, P. H. (2003). A perfect smoother. *Analytical Chemistry*, 75(14):3631–3636.
- [Eilers et al., 2017] Eilers, P. H., Pesendorfer, V., and Bonifacio, R. (2017). Automatic Smoothing of Remote Sensing Data. *2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images, MultiTemp 2017*, (D).
- [Fick and Hijmans, 2017] Fick, S. E. and Hijmans, R. J. (2017). Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12):4302–4315.
- [Fraser et al., 2000] Fraser, R., Li, Z., and Cihlar, J. (2000). Hotspot and ndvi differencing synergy (hands): A new technique for burned area mapping over boreal forest. *Remote Sensing of Environment*, 74(3):362–376.
- [Harris et al., 2014] Harris, I., Jones, P., Osborn, T., and Lister, D. (2014). Updated high-resolution grids of monthly climatic observations – the cru ts3.10 dataset. *International Journal of Climatology*, 34(3):623–642.
- [Klisch and Atzberger, 2016] Klisch, A. and Atzberger, C. (2016). Operational drought monitoring in Kenya using MODIS NDVI time series. *Remote Sensing*, 8(4).
- [Pesendorfer, 2021] Pesendorfer, V. (2021). Modape. <https://github.com/WFP-VAM/modape>.
- [Satlawa, 2021] Satlawa, P. (2021). Phenology georgia. [https://github.com/satlawa/boku\\_phenology](https://github.com/satlawa/boku_phenology).
- [Shabanov et al., 2002] Shabanov, N., Zhou, L., Knyazikhin, Y., Myneni, R., and Tucker, C. (2002). Analysis of interannual changes in northern vegetation activity observed in avhrr data from 1981 to 1994. *IEEE Transactions on Geoscience and Remote Sensing*, 40(1):115–130.
- [Takaku et al., 2020] Takaku, J., Tadono, T., Doutsu, M., Ohgushi, F., and Kai, H. (2020). Updates of aw3d30' alos global digital surface model with other open access datasets. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 43(B4):183–190.

- [Vuolo and Atzberger, 2012] Vuolo, F. and Atzberger, C. (2012). Exploiting the classification performance of support vector machines with multi-temporal moderate-resolution imaging spectroradiometer (MODIS) data in areas of agreement and disagreement of existing land cover products. *Remote Sensing*, 4(10):3143–3167.
- [Vuolo et al., 2012] Vuolo, F., Mattiuzzi, M., Klisch, A., and Atzberger, C. (2012). Data service platform for MODIS Vegetation Indices time series processing at BOKU Vienna: current status and future perspectives. *Earth Resources and Environmental Remote Sensing/GIS Applications III*, 8538(October 2012):85380A.
- [Whittaker, 1922] Whittaker, E. T. (1922). On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society*, 41:63–75.