Analyzing point data with TIMESAT

Summary

Following a successful TIMESAT installation in the previous practical, you will now analyse time series data of 10-day PROBA-V 1-km NDVI composites for individual pixels in northern Iraq. Based on a prepared text-file containing the time series for six points, you will explore different phenology retrieval options with the TIMESAT Graphical User Interface (GUI). There is not necessarily one best option, but this practical should give you some idea of the differences between the options.

Assignment

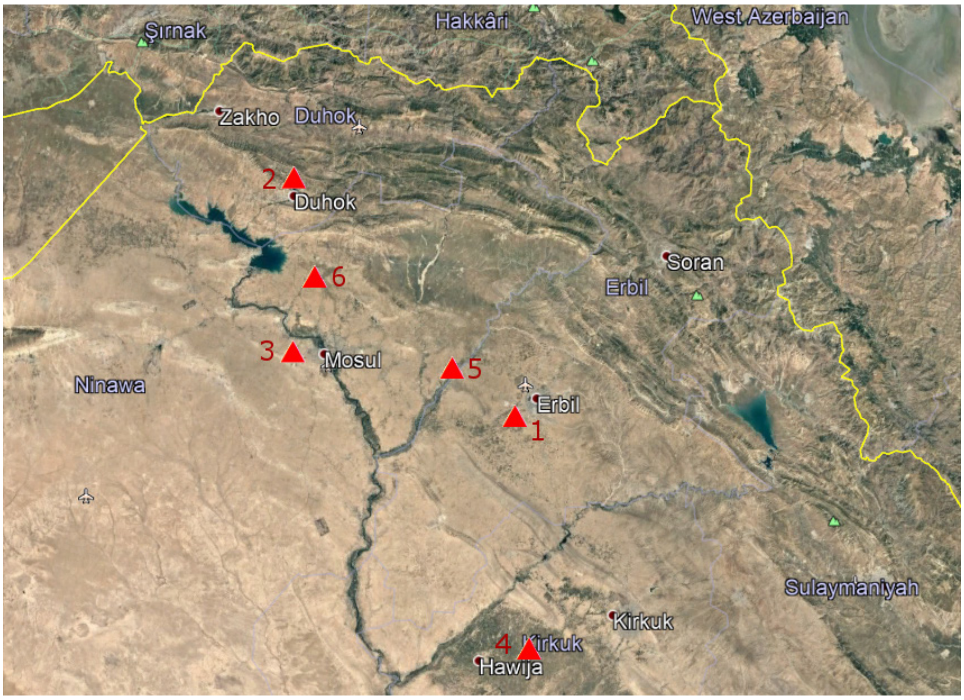
While following the steps in this document, write down your answers to the questions *in green* in a Word Document. In addition, you will be asked to submit an output file, as indicated *in red* in this document.

1. Data and sample locations

The 10-day PROBA-V composites are of Version 2.2 and obtained through the Copernicus Global Land Service product portal. They are constructed by the Flemish Institute for Technical Research (VITO) in Belgium. A detailed product user manual can be found here.

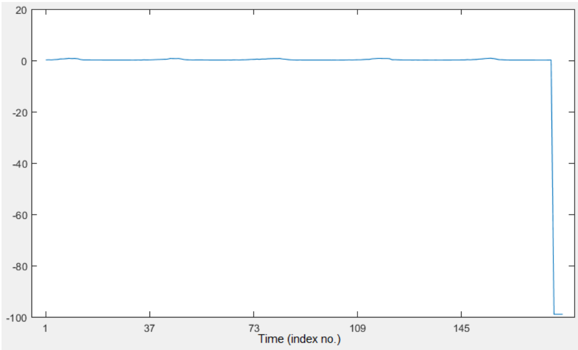
The time series for six sample pixels is stored in *northIraq_ProbaVpoints_scaled.txt*, in which the NDVI is scaled between 0 and 1, with the value -99 indicating missing data (due to e.g. clouds). Each series contains five years of data, from 2013 to 2017. The 2013 data are in fact not derived from the PROBA-V satellite, but from its predecessor SPOT-VEGETATION on the SPOT-5 satellite. At the time of writing, we missed 2017 from the end of November onwards, and those are filled with -99 values.

The figure below shows the locations of each point (with the number indicating the identifiers that are placed in this sequence in the *northIraq_ProbaVpoints_scaled.txt* file. You can further view the precise locations by opening *pixelsProbaV_nIraq.kmz* in Google Earth, which also show the approximate outline of each pixel.

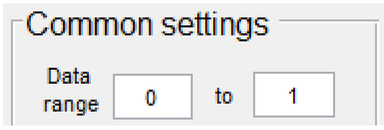


2. Loading the data in the TIMESAT GUI

From the TIMESAT menu system, start the TSM_GUI. In the TSM_GUI window click on **File → Open ASCII data file**, browse to the location where you stored *northIraq_ProbaVpoints_scaled.txt*, select it, click **Open**, and in the next window click **Load data**. You will probably see a figure like:



This is because TIMESAT first needs to know what a valid data range is; for now it is also plotting the -99 values as being valid data values. To change this, adapt in the **Common settings** part of the window, the first Data Range value to 0 (as below). Now you should be able to see the temporal pattern with five distinct years (i.e. from 2013 to 2017).



If you click on the left part of the GUI on **Plot next series**, you can move forward to other rows of data, i.e. other pixels. The row numbers correspond with the identifiers shown on the previous page (and in *pixelsProbaV_nIraq.kmz*). After the last row (here row 6) you will return to the first row when clicking again **Plot next series**.

Note that the x-axis (time) is defined in the same units as the time step in the data: in this case one time step represents 10 days, often referred to also as a dekad. By the way, 'decade' generally refers to 10 years, 'dekad' to 10 days. The dekads are defined to fit precisely to the month definitions, meaning that often the last dekad of each months is not precisely 10 days. For example, dekad 1 refers to 1-10 January, dekad 2 to 11-20 January, and dekad 3 to 21-31 January (so dekad 3 covers 11 days). Similarly, dekad 6 refers to 21-28 February covering 8 instead of 10 days (or 21-29 February for leap years). Although for the PROBA-V composites the precise date selected for each pixel is in principle known (provided with the data), the current version of TIMESAT does not allow to use this information. Therefore, we can assume here that the date corresponding to each 10-day composite is the mid-date of those 10 days; so for 21-31 January, this would be 26 January.

3. Filtering options

Generally the first step for extracting phenology is to fit a model to the data. In the lecture 4 we briefly explained the concept behind the Savitzky-Golay filter. For TIMESAT this is the first option in the list of fitting methods (under **Data Plotting**). Chapter 3 of the user manual (TIMESAT 3.3 Software Manual) explains the various options in more detail.

- Experiment with the Savitzky-Golay, the Assymetric Gaussian, and the Double Logistic functions (without changing other options) by selecting each for the six points. If you hold down the CONTROL key on your keyboard while selecting, you can simultaneously plot the different fittings.

Q1: By only examining the temporal plots, do you find large differences between the different fittings? For which series (i.e. row numbers) is the difference largest? What in the NDVI series may cause the difference to be large?

- Let us continue with the Double Logistic model only. As discussed in Lecture 4 and explained in Section 3.4 of the user manual, noise in NDVI data is negatively biased, meaning that atmospheric effects can lower the NDVI. As a consequence we have more confidence in the larger NDVI values. We can therefore adapt the fitting to the so-called "upper envelope" of the data. In TIMESAT, explore what happens if you can set the number of iterations to 1 (no iteration), 2, and 3 under the **Class-specific settings** with the box behind **No. of envelope iterations**.

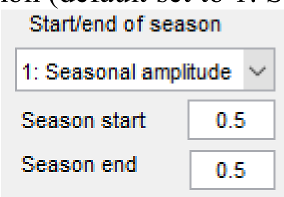
Q2: Describe and explain what happens when we increase the number of iterations. Does it affect the results greatly? For which data row are the differences most pronounced? How can you explain this?

- Other settings exist that we do not treat extensively in this practical. One of them is the **Spike method**, found under **Common settings**. This is explained in Section 3.3 of the user manual. Feel free to explore the effect of changing such settings further. For the remainder of this practical we will keep the value at "0: none".

4. Extracting phenology parameters

After fitting a model to the data, we can estimate phenological parameters. Let us continue to focus on the double logistic model, setting the no. of envelope iterations to 3. Ticking the box **Season start/stop** under **Data plotting** adds the start- and end-of-season (SOS/EOS) estimates to the graph as points on the fitted curve. Note that only for the years 2014-2017 estimates are made of SOS and EOS, because TIMESAT only extracts for the *n-1* center-most full seasons as explained in Section 4.1 of the user manual.

- Under Class-specific settings keep the values at 0.5 and only experiment changing the method of threshold calculation (default set to 1: Seasonal amplitude), as below:



Section 4.3 of the user manual explains the different options.

Q3: When using option 1 (seasonal amplitude), why is the EOS detection for a particular season not on the same NDVI-level as the SOS for the same season? (for example in 2014 for row 2)

Q4: If we change to option 2 (absolute value), we do not get estimates of phenology for each season (for example in row 2): why is that?

Q5: Explain the difference between option 2 (absolute value) and option 3 (relative amplitude).

- Another important setting to vary is the threshold level. Let's focus only on option 3 (relative amplitude) and experiment with changing the thresholds for season start and season end to values between 0.10 and 0.50.

Q6: What effect does a lowering of the threshold have on the dates estimated for SOS and EOS?

Q7: What happens to the SOS/EOS retrieval for the year 2014 for row 2 if we set the threshold for season start at 0.1?

- While one may be tempted to put the thresholds at a low level to make the SOS-date represent the early green-up of the season, we have to consider that the fitted curve for that part of the season may be more affected by noise. Partly for this reason, many studies use a more conservative threshold, such as for example 50% of the seasonal amplitude. Of course, if one is interested in linking the phenology analysis with field data (e.g., on crop emergence dates, or crop productivity), this can be a reason to adapt the method and its options to attain a better fit between estimated parameters and field data. In any case, visual estimation of the effect of the variable settings (as can be done easily with the TIMESAT Graphical User Interface) can be very important before applying settings spatially to multi-temporal image data. This requires a good selection of different profiles of areas of interest within the image.
- We shall use in the remainder of this practical the following settings:
  - Double Logistic model
  - Spike method none (0: none)
  - No. of envelope iterations: 3
  - Start/end of season determined by 3: Relative amplitude
  - Thresholds for season start and end: both 0.25
- The results of the phenology retrieval for a single row of data are found in the bottom right corner of the TIMESAT Graphical User Interface. This information can be saved by clicking: **Output → Write seasonality to file**. The output will be stored in the TIMESAT working directory (*timesat33\run*).

To hand in: Create the seasonality.txt file for the data in Row 6. Rename it to logistic_row6.txt .

- Refer to Figure 16 on p. 26 of the TIMESAT user manual to understand what seasonality parameters are written to the file.

5. Interpretation

Based on the settings defined above, respond to a few interpretation questions:

Q8: For row 6, during which year was the maximum NDVI value the lowest. What could explain this?

Q9: Still for row 6, to what calendar date (approximately) does the start of the season for the first retrieved season correspond? (hint: also read again the last paragraph of Section 2 of this practical)

Q10: Similar as the previous question, to what date does the start of the season for the second retrieved season correspond?

Q11: the large integrated value is often used as a measure of seasonal vegetation productivity, and as such sometimes used for crop yield estimates as well. List the years (2014-2017) in increasing order of the large integrated value.

Q12: How many months/days is the length of season for row 6, for each of the years (2014-2017)?

Q13: Do the other rows of data also show a relatively long season for 2015?

Hand in

- a Word-document in which you write your answers to each of the questions *in green* in this practical.
- The *logistic_row6.txt* file created in part 4 of this practical.