Early land cover classification with Sentinel 2 satellite images and temperature data

Matej Čerin
Jožef Stefan Institute
Jamova 39, 1000 Ljubljana, Slovenia
matej.cerin@ijs.si

Filip Koprivec
Jožef Stefan Institute
Jamova 39, 1000 Ljubljana, Slovenia
filip.koprivec@ijs.si

Klemen Kenda
Jožef Stefan Institute
Jožef Stefan International Postgraduate School
Jamova 39, 1000 Ljubljana, Slovenia
klemen.kenda@ijs.si

ABSTRACT

The weather is one of the main factors for events that happen on the surface of the earth. Surprisingly, no effort was made to use weather features together with satellite images for land cover classification. In the paper, we use temperature data along with satellite images to improve the accuracy of the classification and to get classification as early in the year as possible. Every year has different conditions, so the temperature can be used as an objective criterion at what time to do classification.

Keywords

remote sensing, weather, earth observation, machine learning, feature selection, classification, temperature

1. INTRODUCTION

Precise classification of land cover is important for agriculture and food security. It is especially important to make the classification as early as possible so the farmers can know the situation of their land better and can take appropriate action.

Since ESA launched the Sentinel 2 mission that provides big amount of data, a lot of research is focusing on satellite data and land cover classification [9, 5, 15, 14]. We found no paper where the weather would be used to improve the land cover classification. Most of the research using weather is focused on predicting the yield [11, 6, 1].

Literature shows that weather plays one of the most important roles in the growth of vegetation and development of crops [8, 7, 16]. It is expected that it will play an even more important role in the future, with climate change trends [7]. According to literature the most important weather variables for the development of vegetation are temperature, precipitation and the duration of sunlight. The most important of them is temperature. The research shows that time for growing is better correlated with temperature than the number of days from planting [1]. The plant growing models talk about the fact that the plants have some optimal temperature above which the plants grow best. It is also important that the temperature does not go above some threshold [8]. Weather variables are also important for farmers decisions. Based on weather condition the farmer can choose when what and how to seed [7].

In the paper, we try to improve our classification prediction with the help of temperature. We are trying to find the moment when the temperature is above some threshold for long enough.

2. DATA

2.1 Data Acquisition

In the article, we use satellite data from the ESA Sentinel-2 mission [3]. The Sentinel-2 mission has two satellites that circle the earth with 180° phase. The same point is visited at least once every five days. Satellites collect data in 13 different spectral bands. Spatial resolution is 10m, 20m or 60m, depending on the band.

We downloaded satellite data with the sentinel-hub library [12] integrated in the eo-learn library [13]. Eo-learn is the library that makes access to and processing of earth observation data easier. We used eo-learn also to preprocess the data. Data were downloaded for times between July 1, 2015 and June 30, 2018.

Temperature data are from the ECMWF (European Centre for Medium-Range Weather Forecasts) archive [2]. ECMWF is the archive that stores accurate historical data, both observed and forecasts weather data for the world. The data used in the article has approximately 15×15 km resolution.

Data for land cover are from the website of the Slovenian Ministry for farming, forests, and food [10]. Data are publicly available and contain 25 classes. The land use data are mostly created with aerial photography called orthophoto. In case that the area is impossible to categorize from the image, the terrain inspection is made.

2.2 Data Preprocessing

Most preprocessing of satellite data are already made by ESA, like atmospheric reluctance or projection [4]. Thus, our data is already clean and ready for use.

The biggest challenge in our satellite data set was missing data when the clouds cower some images or some parts of it. To eliminate that problem we took images provided by ESA and filtered out the pixels that were covered by clouds using the cloud mask. The cloud mask was provided by eo-learn's AddCloudMaskTask() task. That way we got images only with cloudless pixels.

The other concern was that all images were not taken on the same date and now we have some missing data from cloud removal. Therefore we took a time series of each pixel and linearly resampled all bands over time. We resampled it on every 16th day starting on 1.1.2016 and up to 31.12.2017.

That way we produced a data set that had all images at the same timestamp. Linear respelling also filled the gaps from filtering the images with clouds.

The land cover data had 26 different classes, but some classes were too small. We joined the related classes under five more general classes (grass, forest, crop land, urban area and other).

3. METHODOLOGY

The idea of the experiment was that when the temperature is above a certain threshold the plant stars to grow. Because they grow differently, it is easier to classify the areas with different vegetation. Therefore we looked for the time when the temperature is high enough for land covers to be easier to classify.

3.1 Feature Vectors

Experiments were conducted in the area of Slovenia. We used data from years 2016 and 2017. Data from 2016 were used to train the model and data from 2017 for testing. We randomly chose 150 patches in the size of 50×50 pixels ($500\ m \times 500\ m$). Then we sampled from those patches approximately 50 000 pixels with the class that we are interested in and 50 000 pixels that are not from that class. That way we get balanced data sets, that we can use to train our learning algorithm. Thus, we created two vectors for each class, One for learning and one for testing.

For each date, we counted the number of days that average temperature exceeded some maximum temperature (T_{max}). We calculated that for T_{max} from $-10°$C to $26°$C for every $2°$C. Those features we added to the time series of pixels. Because the temperature data has smaller spatial resolution than satellite data, we appended to each pixel the temperature data from the weather data point that is the closest to the coordinates of that pixel.

For all pixels' time series, we found the first timestamp for which the number of days with temperature above T_{max} was higher than the chosen number of days. We took the values of bands at that timestamp for each pixel. That was done for both years, resulting in two feature vectors, one to train the model and the other to test it.

Because some higher temperatures were not reached often enough, some did not include all pixels. If less than 70% of all pixels passed the criteria, the experiment under those criteria was not made.

3.2 Experiment

On data sets from 2016, we trained the decision tree classifier. The decision tree function is from the sci-kit learn python library and was used with default settings. To evaluate models we calculated predictions for the year 2017, and calculated $F1$ score. In all experiments we made two class classification.

We did experiments systematically for all calculated temperatures T_{max} and for all possible numbers of days from 1 to 30.

To compare results we made another experiment where we trained the model on the data from one date and tested it on the closest date next year. We compared $F1$ scores from both experiments, to see if the model, trained with data set chosen with help of temperature, perform better.

4. RESULTS

The maximum $F1$ score from the experiment with the data set determined by temperature is better for all classes than the maximum $F1$ score from the second experiment (table 1). That means that the temperature helped us improve the classification.

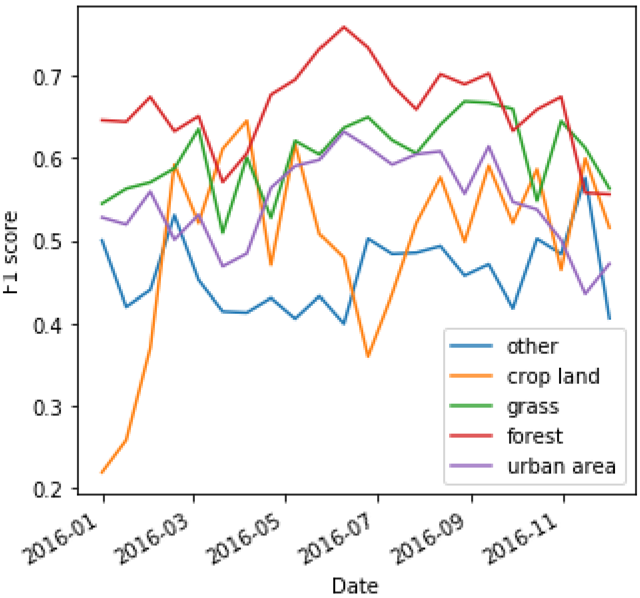

Figure 1: F1 scores from model trained and tested at the same date. Maximum values are used to compare results from first experiment.

	Forest	Grass	Crop	Urban area	Other
Max $F1$ score-temp	0.76	0.74	0.70	0.73	0.59
Max $F1$ score-time	0.74	0.67	0.62	0.64	0.53
Diff	0.02	0.07	0.08	0.09	0.06

Table 1: Table shows maximum F1 scores for all five classes, from both experiments and the difference between them.

Figure 2 shows the difference between F1 scores from the experiment with temperature and the maximum F1 score from

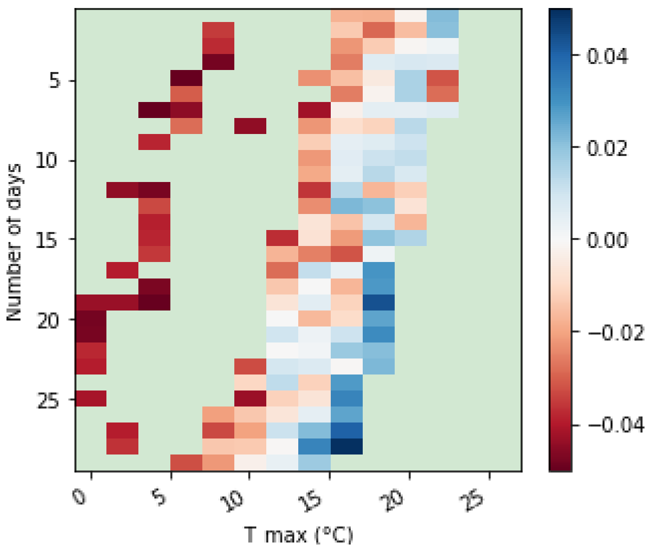

Figure 2: The figure shows the matrix of difference between $F1$ scores from the experiment with temperature and the maximum $F1$ score from the second experiment. Blue color shows times when the $F1$ score is better than in first experiment and red color when the $F1$ score is worse by less than 0.05. Green area is where the classification was worse by more than 0.05 or the times when less than 70% of pixels were available for training. This figure is from the classification of grass.

other experiment. The experiment from figure 2 was made for grass classification. We get similar images also for other classes. On the figure, we notice that we get two islands, one at the temperature around $4°C$ and the other around $16°C$. Each island corresponds to data at different dates. The data from the same island are from similar dates. The distribution over dates for both islands is shown in figure 3.

The classification of grass, forest, and urban area produces the same kind of islands, while the crop and other produces only one big island. We assume that this is due to non-homogeneous vegetation in those two classes.

From figure 1 we see that the classification of the grass in the second experiment is the best at the end of August ($F1 = 0.67$). An even better classification score (up to 0.07, $F1 = 0.74$) can be achieved with most of the classification made before august (blue, orange and green bars from figure 3).

Another useful result from approach with temperature is that the classification can be made earlier in the development of plants. Relatively good classification ($F1 = 0.63$) can be achieved by the end of April (red, purple, brown and pink bars from figure 3).

For forest and urban area, the first island gives also slightly worse classification but we can classify earlier. While the second island gives us better results at the approximately same time. The classification of both classes with one island is approximately at the same time, but it achieves better results.

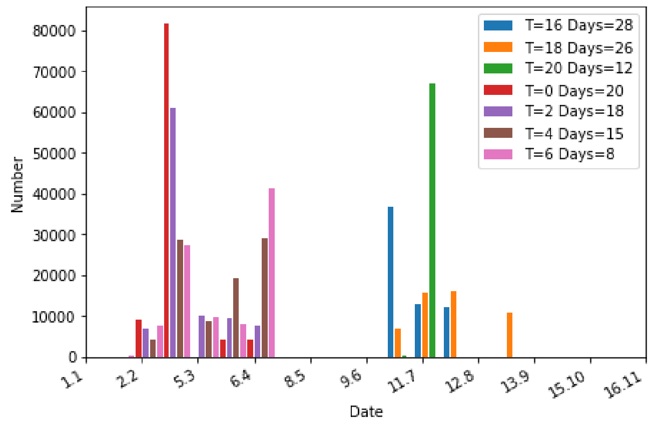

Figure 3: Histogram shows the number of pixels with certain time stamp, at chosen T_{max} and number of days. It shows only some representatives of both islands.

5. CONCLUSIONS

In the article, we showed that temperature can help us determine when is the most appropriate time to do classification. Because the conditions are not the same everywhere, temperature gives us a good and objective tool to determine when is the best time to do classification.

We also showed that we can do classification much earlier than we thought. Plants do not need to fully grow. We can identify its development as early as March or April.

The problem with that method is that we can not know in advance when the optimal time for classification will come. And when that time comes it is not the same for all areas but is determined locally, by local weather condition. Therefore usually, one model can do all classification in one month and a half, for the whole area of Slovenia. But for a farmer who is interested in the growth and development of his plants, that is usually not a problem, because his farm is usually smaller than the resolution of weather data. That means that he can get all classification data for his farm in a day. But if he is from the colder regions of Slovenia he might still wait for some time before getting predictions.

In the future, the goal would be to add other weather features like precipitation or sun duration. Another important use case would be to focus on agriculturally more interesting plants like corn, wheat, and others. That would be important to ensure food security in years with the bad weather condition.

6. ACKNOWLEDGMENTS

This work was supported by the Slovenian Research Agency and the ICT program of the EC under project PerceptiveSentinel (H2020-EO-776115). The authors would like to thank Sinergise for their contribution to EO-learn library along with all the help with data analysis.

References

[1] Jan Dempewolf et al. "Wheat Yield Forecasting for Punjab Province from Vegetation Index Time Series and Historic Crop Statistics". In: *Remote Sensing* 6 (Oct. 2014), pp. 9653–9675.

[2] ECMWF. `https://www.ecmwf.int/`. Accessed 25 August 2019.

[3] ESA. `https://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-2/Satellite_constellation`. Accessed 13 August 2018.

[4] ESA. `https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/processing-levels/level-2`. Accessed 13 August 2018.

[5] Cristina Gómez, Joanne C. White, and Michael A. Wulder. "Optical remotely sensed time series data for land cover classification: A review". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 116 (2016), pp. 55–72. ISSN: 0924-2716.

[6] Wuttichai Gunnula et al. "Normalized difference vegetation index relationships with rainfall patterns and yield in small plantings of rain-fed sugarcane". In: *Australian Journal of Crop Science* 5 (Dec. 2011), pp. 1845–1851.

[7] Toshichika Iizumi. "How do weather and climate influence cropping area and intensity?" In: *Global Food Security* (Nov. 2014).

[8] University of Illinois at Urbana-Champaign. *Illinois agronomy handbook*. 24th ed. Urbana, Ill. : University of Illinois at Urbana-Champaign, College of Agricultural, Consumer and Environmental Sciences, Dept. of Crop Sciences, University of Illinois Extension, 1999. ISBN: 1883097223.

[9] Filip Koprivec, Matej Čerin, and Klemen Kenda. "Crop classification using PerceptiveSentinel". In: (Oct. 2018).

[10] Ministrstvo za kmetijstvo gozdarstvo in prehrano. `http://rkg.gov.si/GERK/`. Accessed 25 August 2019.

[11] Umer Saeed et al. "Forecasting wheat yield from weather data and MODIS NDVI using Random Forests for Punjab province, Pakistan". In: *International Journal of Remote Sensing* 38 (Sept. 2017), pp. 4831–4854. DOI: `10.1080/01431161.2017.1323282`.

[12] Sinergise. `https://github.com/sentinel-hub/sentinelhub-py`. Accessed 14 August 2018.

[13] Sinergise. `https://github.com/sentinel-hub/eo-learn`. Accessed 23 August 2019.

[14] Silvia Valero et al. "Production of a Dynamic Cropland Mask by Processing Remote Sensing Image Series at High Temporal and Spatial Resolutions". In: *Remote Sensing* 8(1) (2016), p. 55.

[15] François Waldner, Guadalupe Sepulcre Canto, and Pierre Defourny. "Automated annual cropland mapping using knowledge-based temporal features". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 110 (2015), pp. 1–13. ISSN: 0924-2716.

[16] Matteo Zampieri et al. "Wheat yield loss attributable to heat waves, drought and water excess at the global, national and subnational scales". In: *Environmental Research Letters* 12 (June 2017), p. 064008.