Free and Open Source Software for Geospatial (FOSS4G) Conference Proceedings

Volume 18 *Guimarães, Portugal* Article 3

2018

Grassland Recognition with the Usage of Thermal Weights

Alen Mangafić
Geodetic Institute of Slovenia

Nika Mesner
Geodetic Institute of Slovenia

Mihaela Triglav Čekada
Geodetic Institute of Slovenia and Faculty of Civil and Geodetic Engineering, University of Ljubljana

Grassland Recognition with the Usage of Thermal Weights

Optional Cover Page Acknowledgements
This work was financially supported by Slovenian Research Agency by the research project J2-8176.

Grassland Recognition with the Usage of Thermal Weights

Alen Mangafić[a]*, Nika Mesner[a], Mihaela Triglav Čekada[a,b]

[a] *Geodetic Institute of Slovenia, Ljubljana, Slovenia.*

[b] *Faculty of Civil and Geodetic Engineering, University of Ljubljana, Ljubljana, Slovenia*

Abstract: In this paper we apply the usage of thermal weights, a new variable for geostatistical analysis and we present the method for their determination. In the case study we tested a data fusion between Sentinel-2 and Landsat 7/8 data, to incorporate also the thermal factor in the detection of land cover changes. The process distinguishes grasslands from other crops with similar vegetative appearance and offers us the possibility to create a new statistical sample with just grasslands. The data fusion is incorporated in the calculation of Land Surface Temperature (LST^{FU}) by combining the Sentinel-2 derived Normalized Difference Vegetation Index (NDVI), and from it derived land surface emissivity, with the Landsat 7/8 derived Top of Atmosphere Brightness Temperature (TOABT). The experimental LST^{FU} is modified into a normalized assessment variable by a time-series analysis. The result is a thermal weight layer which can help us in further object-based image analyses and classification. The thermal weight is calculated from Sentinel-2 and Landsat 7/8 datasets that has small acquisition time gaps between them. The accuracy assessment due to time gaps and sensor differences was evaluated with Cohens's kappa (κ) and correlation matrix validation. The data fusion is made to test if a Sentinel-2 fusion approach could improve the Thermal Weight created just by Landsat imagery. The purpose was to evaluate the importance of thermal bands for LU/LC cover.

*Corresponding author
Email address: alen.mangafic@gis.si (Alen Mangafić)
URL: https://www.researchgate.net/profile/Alen_Mangafic (Alen Mangafić)

1. Introduction

Multi-spectral data in combination with the high revisit frequency makes Sentinel-2 a great platform for LU/LC change estimation. Unfortunately, some classifications cannot be made precisely with just the input of Sentinel-2 data. Some types of cropland are more demanding for classification because of their biomass and biodiversity characteristics. Some characteristics can be hardly distinguished with an approach that is based just on vegetation index values (NDVI, FAPAR, LAI), their clustering, single-level object-based image analysis and even photointerpretation. Such characteristics are typical also for grasslands. Their changes in time have different dynamics than other croplands. In general, grasslands have lower vegetation indexes than agricultural crops, but their most problematic variety, the overgrown ones, are more demanding for monitoring. The normal growth of agricultural land can be easily mistaken for the overgrowth of grasslands. Their differentiation demands high resolution data for photointerpretation and a robust model for geostatistical analysis. Even with very high-resolution data (aerial orthophotos), photointerpretation can be unsuccessful in some areas. When we deal with automatic grassland recognition techniques for monitoring the minimisation of the margin of error becomes a big challenge.

2. Medium resolution imagery for LU/LC: from SAR to TM variables

To perform an effective geostatistical analysis on medium resolution imagery like Sentinel-2, we must apply adjustment techniques which include auxiliary variables and multilevel analysis. Reducing dimensionality and weighting adjustments can be a very good approach for setting up classification models with means of distinguishing similar land covers. Fitting of data ranges through principal components analysis (PCA) derived from very high-resolution data is a great approach [Deng et al. 2008]. The problem is that high-resolution airborne data, in general, have no adequate acquisition frequencies that monitoring tasks demand. This is different for earth observation data. One of the main advantages of satellite derived imagery is their frequency. Unfortunately, open raster data such as Sentinel-2 data have a medium resolution that cannot give us precise textural information [Yu et al. 2016]. This means that the largest possible variance is not adequate for performing PCA suitable for distinguishing grasslands from growing croplands or low vegetation areas. We must move our deductions to a different empirical level of indexing. Weight variables that are derived also from other sensors rather than just from visible light and near-infrared optical ones can help us with modelling a machine learning process. There are various object-based approaches that integrate C-band Synthetic Aperture Radar data (Sentinel-1) and optical data for land use classification [Tamm et al. 2016]. The processing of C-band SAR data is a much more complicated and time demanding.

There are many case studies oriented towards the use of Landsat TM thermal bands for LU/LC classification which proved that the involvement of thermal imagery with VIS/NIR bands improves the accuracy of classification [e.g. Alavipanah et al. 2001; Sun and Schulz 2015].

The confusion matrix of Landsat 8 imagery (including Thermal Infrared) classified by Sun and Schulz (2015) with the Radom forests algorithm brings out that the crops, which are most commonly misclassified as dense and sparse grasslands are barley, corn and wheat.

Such misclassifications can be reduced with a classification based on imagery with better resolution. Sentinel-2 derived NDVI can supply us with more accurate indexes which improve also the textural information than Landsat 7/8. Our hypothesis is that the growing cycle and the watering of croplands effect their thermal proprieties and differentiate them from other fields which have different growing dynamics. Grassland can be harvested two to four times per season while crops fields are harvested once per season. More studies found that there is a strong positive relationship between soil moisture and NDVI, although the correlation is lower in areas with high NDVI [Chen et al. 2014; Al-Shehhi et al. 2011].

This leads us to suppose, that a combination of Sentinel-2 data and Landsat 7/8 Thermal Infrared data can lead us to a better differentiation of grasslands from the misclassified crops, which have similar vegetative indices and are classified without Thermal Infrared data or are classified in combinations of bands derived just from Landsat 7/8.

3. Methodology

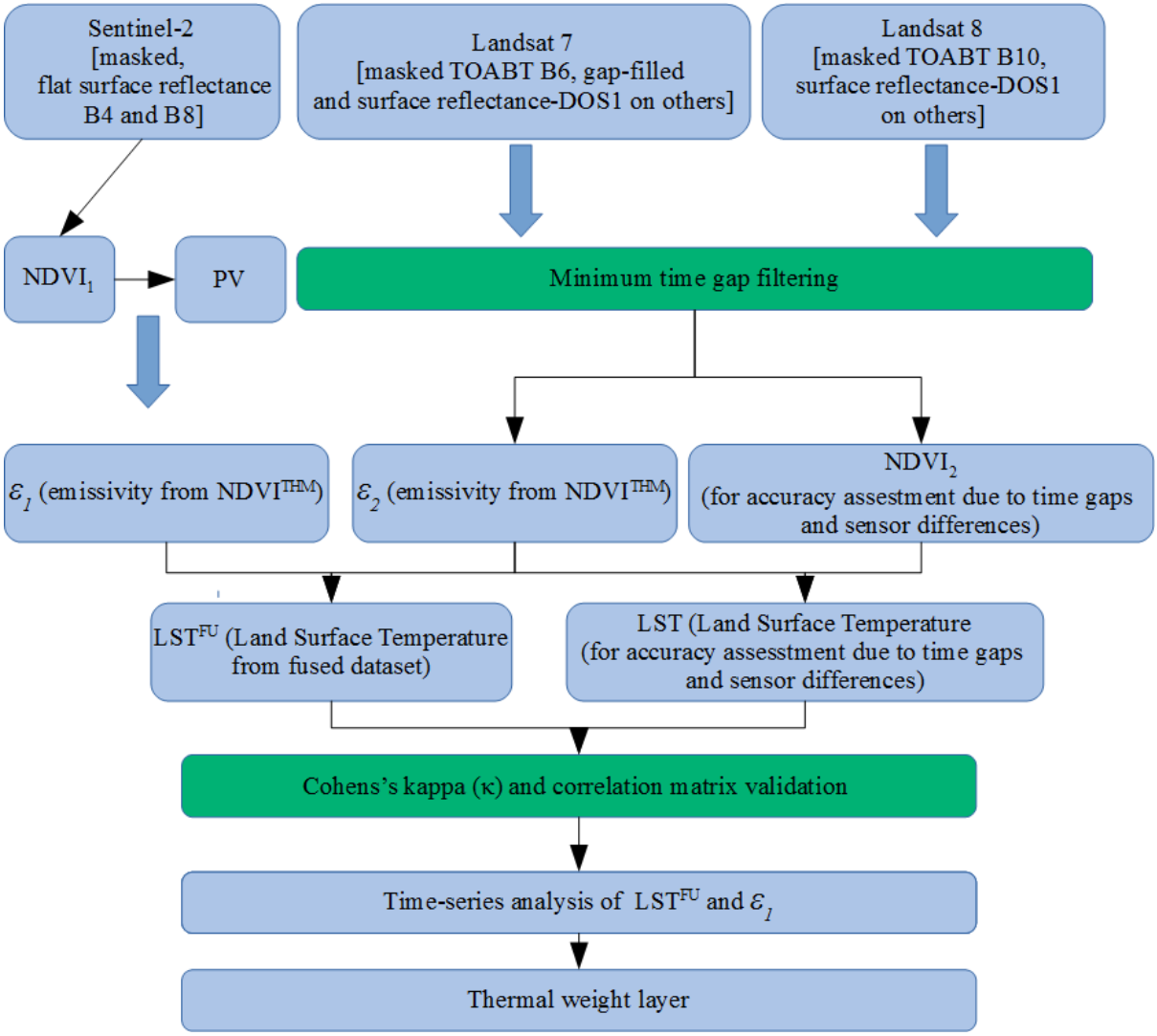

Figure 1: Overview of the processes and the methodology

Figure 1 provides an overview of the processes on which our methodology is constructed. Our methodology is based on data fusion between Sentinel-2 based NDVI and Landsat 7/8 Thermal Infrared data:

1. We applied a NDVI-threshold method ($NDVI^{THM}$) for estimating land surface emissivity (ε) from Sentinel-2 (ε_1) and Landsat 7/8 (ε_2) [Sobrino et al. 2008];
2. The data fusion was incorporated in the calculation of Land Surface Temperature (LST^{FU}) by combining the Sentinel-2 derived land surface emissivity with the Landsat 7/8 derived Top of Atmosphere Brightness Temperature (TOABT). The Thermal Infrared source data is chosen by criteria based on cloud absence with a minimum time gap with prioritizing Landsat 8 Thermal Infrared imagery due to Landsat 7's SLC-off data;
3. The accuracy assessment due to time gaps and sensor differences was evaluated with Cohens's kappa (κ) and correlation matrix validation of $NDVI^{THM}$ data;
4. The experimental LST^{FU} and the ε_1 are modified into a normalized assessment variable by a time-series analysis. The result is a thermal weight layer which can help us in further image analyses and classification;
5. The results are validated and correlated to in-situ data.

3.1. Surface reflectance, gap filing and cloud masking

The preprocessing of satellite imagery was divided into the preprocessing of Sentinel-2 and Landsat 7/8 data. Sentinel-2 data was corrected for atmospheric effects (including the adjacency effects) and slope effect with the multi-temporal MAJA level 2A processor [CESBIO 2015; Lonjou et al. 2016]. The level 2A data was masked from all the clouds (except the thinnest) and all the shadows by MAJA generated cloud mask. We prioritized the use of Landsat 8 Thermal Infrared imagery due to Landsat 7 SLC-off data. Landsat 7 was involved as a data source because of its importance in case of big time gaps between Sentinel-2 and Landsat 8 imagery due to cloudiness suitability. Landsat 7 Thermal Infrared bands has also a better resolution (60 m acquired, 30 m resampled) than Landsat 8 (100 m acquired, 30 m resampled) [USGS 2018a]. We chose to fill the nodata gaps by using an Inverse Distance Weighting Interpolation (IDW) implemented by GDAL (gdal_fillnodata) available in QGIS. There are other geostatistical techniques superior to IDW for SLC-off data interpolation, like the modified AWLHEM, which we did not consider because of its multi-source nature and greater computing time needs [Chen et al. 2012] compared to the efficiency of IDW [Sulong et al. 2015]. Landsat imagery (VIS/NIR) was then corrected by the Dark Object Subtraction 1 (DOS1) image-based atmospheric correction [Moran et al. 1992] implemented in the Semi-Automatic Classification Plugin (SCP) developed for QGIS [Congedo 2016]. The TOABT conversion of the Thermal Infrared bands was made with the SCP, according to the equation 1 [USGS 2018b]:

$$T_{TOAB} = \frac{K_2}{\ln(\frac{K_1}{L_\lambda}+1)} \quad (1)$$

The thermal constants K1 and K2 are provided in the Landsat 8 metadata, but not in the Landsat 7. Therefore, they are calculated from the wavelengths of emitted radiance provided in the Landsat 7 documentation [USGS, 2018a] and the c1 and c2 radiation constants [Congedo

2016]. All the bands were masked with the Fmask method [Zhu et al. 2012] applied with the Cloud Masking QGIS plugin.

3.2. Estimating land surface emissivity (ε) with a NDVI-threshold method

Land Surface Temperature can be calculated from TOABT, according to the equation 2 [Weng et al. 2014; Congedo 2016]:

$$\mathrm{LST} = \frac{\mathrm{T_{TOAB}}}{[1+(\frac{\lambda\, \mathrm{T_{TOAB}}}{c_2})*\ln \varepsilon]} \tag{2}$$

The land surface emissivity ε for its calculation can be determined in various deterministic ways and one of them is the presented NDVITHM proposed for the Landsat 7 TM, according to the equation 3 [Sobrino et al. 2008]:

$$\mathrm{NDVI_{TM}^{THM}} = \begin{cases} \varepsilon = 0.979 - 0.035 * \mathrm{R}, & \mathrm{NDVI} < 0.2 \\ \varepsilon = 0.986 + 0.04 * \mathrm{PV}, & 0.2 \leq \mathrm{NDVI} \leq 0.5 \\ \varepsilon = 0.99, & \mathrm{NDVI} > 0.5 \end{cases} \tag{3}$$

In equation 3, R is used as the Sentinel-2 red band image and PV is the Proportion of Vegetation derived from the Sentinel-2's NDVI, calculated as the Vegetation Condition Index, according to the equation 4 [Orhan et al. 2014]:

$$\mathrm{PV}=(\frac{\mathrm{NDVI}-\mathrm{NDVI_{min}}}{\mathrm{NDVI_{max}}-\mathrm{NDVI_{min}}})^2 \tag{4}$$

The reclassification and calculations were done with GRASS GIS r.mapcalcc.

3.2. Accuracy assessment and validation due to time gaps and sensor differences

The accuracy assessment due to time gaps and sensor differences was evaluated with Cohens's kappa (κ) agreement coefficient and by a further correlation matrix validation, due to kappa's inconsistency cause of the high correlation to overall accuracy [Olofsson et al. 2014; Strahler et al. 2006].

Kappa was calculated by GRASS GIS r.kappa function and the correlation matrix was calculated with GRASS GIS r.covar. The results which had an agreement coefficient higher than κ=70 and strong correlation between NDVI values were used for further time-series analysis which resulted in a Thermal Weight raster layer.

3.3. Thermal Weight calculation

The results which had an acceptable agreement coefficient and correlation were used for further time-series analysis with GRASS GIS r.series. We made two different time-series analysis:

- a raster representing the range of the values of the LST and
- a raster representing the maximum land surface emissivity ε_1 values.

Both rasters were normalized on a scale from 1 to 100. We generated a Thermal Weight layer by summing the normalized rasters.

4. Study Area and Dataset Used

Our dataset consists of an area in the Slovenian Styria region which extends on 362 km2. We obtained a LU/LC shapefile and also in-situ data of validated LU/LC by the Ministry of Agriculture, Forestry and Food (areas I2726 and J2824). We tested the proposed methods on a sample of 4 different time intervals: April, June, July and August. We excluded May's data cause of cloudiness. The data was preprocessed and the Thermal Weight was calculated as described in chapter 3.1. The time gaps between Sentinel-2 and Landsat 7/8 imagery varies from 0 to 2 days, as shown in table 1. Our calculations are based on the hypothesis, that the NDVI and therefore land surface emissivity ε1 values don't have a meaningful change in a time-gap of two days.

Figure 2: Study area (basemap source: GURS, 2018)

Sentinel-2 [10:00h]	Landsat 8 [9:40h]	Time-gap [days]
1.4.2017	3.4.2017	2
20.6.2017	22.6.2017	2
30.7.2017	31.7.2017	1
24.8.2017	25.8.2017	1

Table 1: Time of data acquision and time-gaps with Sentinel-2 data

5. Results and Discussion

Our first analysis was the evaluation of the differences between land surface emissivity due to time gaps and sensor differences. Land surface emissivity it is the effectiveness in emitting energy as thermal radiation (Štefan-Boltzmann law). It's calculation by the proposed $NDVI^{THM}$ brings up some errors, as its theoretically value is up to 1 (perfect black body) and the $NDVI^{THM}$ calculated one has a maximum value of 1.0064. As we use the results for comparison of data derived from the same method and furthermore normalize it as a weight, we can exclude this type of error.

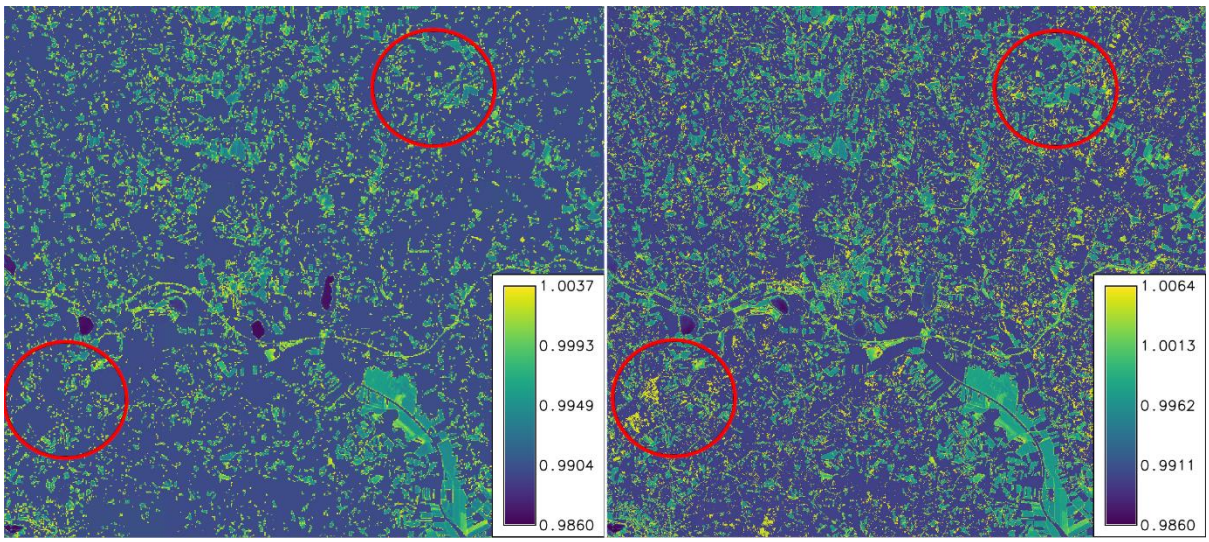

Figure 3: land surface emissivity in April's dataset: Landsat 8 (left) and Sentinel-2 (right)

Nonetheless, already by photointerpretating the emissivity values we can clearly distinguish some arable lands from other land use, as shown in figure 3. The land surface emissivities were both derived by $NDVI^{THM}$ and the Sentinel-2 values have more accurate information due to their better resolution, as in figure 3 we can see some agricultural fields' emissivites which were not recognized by Landsat 8. After calculating the time series, we found that the values of the range of the LST are bigger on agricultural crop fields, which also have the highest maximum land surface emissivity values. As described in the methodology, we generated a Thermal Weight layer by summing the normalized rasters, as shown in figure 4.

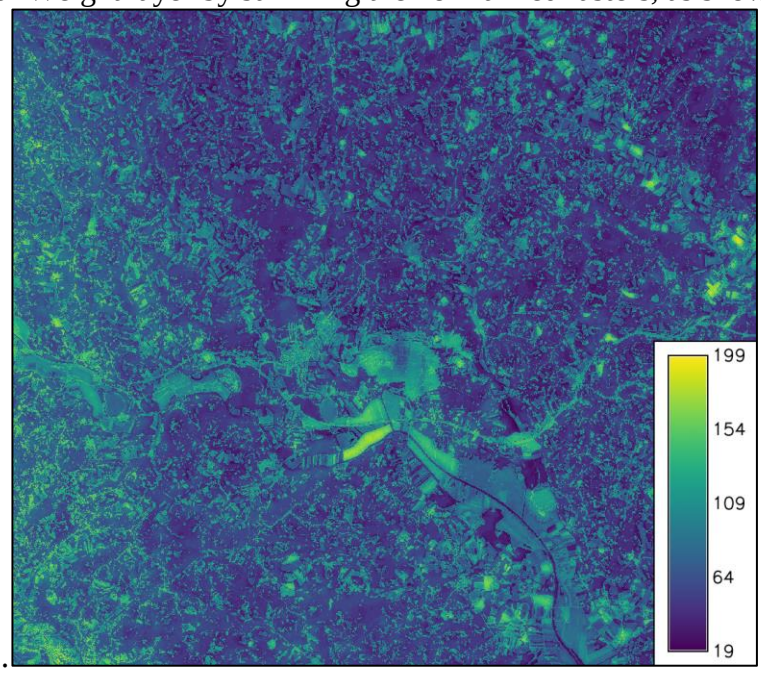

Figure 4: Thermal Weights

We added the average raster values to the LU/LC data, which was validated and categorized into grasslands and agricultural crops polygons bigger than 5000m2. We evaluated the strength and direction of association between the two ranked variables and the Thermal Weights by calculating the Spearman's correlation coefficient for 7662 polygons with R. The calculated coefficient has a negative value of -0.57. This means that the strength of direction of association

between the variables is significant and that the increase of the Thermal Weight means the decrease of the possibility of our polygon to be a grassland one.

6. Conclusion

Thermally weighted raster data can be an important variable for multilevel analysis and including it into classification processes could improve the results. Our purpose was to integrate Thermal Infrared imagery into Sentinel-2 data calculations as a normalized variable, which has a decisive importance in the interpretation process. Our case study chose the diversification of grasslands and agricultural crops cause of their vegetation indices similarity in some seasons. The differences due to thermal and emissivity values were show as important for distinguishing grasslands from agricultural crops. The methodology could be improved by involving a better, Sentinel-2 customized $NDVI^{THM}$ for the calculation of the land surface emissivity. A more robust statistical model could improve the presented methodology. Our methodology is an example, of a possible application of (Landsat) thermal bands in combination with better resolution imagery. Hopefully, soon there will be open Earth observation data products with better or same resolution as Sentinel-2 which will also include a Thermal Infrared sensor capable of medium or high-resolution sensing.

Acknowledgement

This work was financially supported by Slovenian Research Agency by the research project J2-8176.

References

Alavipanah, S. K., M. De Dapper, R. Goossens, and Massoudi, M., January 2001. The Use of TM Thermal Band for Land Cover/Land UseMapping in Two Different Environmental Conditionsof Iran. Journal of Agricultural Science and Technology 3, no. 1: 27–36. Tarbiat Modares University Press. URL http://journals.modares.ac.ir/article-23-2246-en.html

Al-Shehhi, M. R., Saffarini, R., Farhat, Al., Al-Meqbali, N.K. and Hosni Ghedira, 2011. Evaluating the Effect of Soil Moisture, Surface Temperature, and Humidity Variations on MODIS-Derived NDVI Values. IEEE International Geoscience and Remote Sensing Symposium: 3160–63. DOI: 10.1109/IGARSS.2011.6049889.

Chen, F., Zhao, X. and Ye, H., 2012. Making Use of the Landsat 7 SLC-off ETM+ Image Through Different Recovering Approaches. Data Acquisition Applications, edited by Zdravko Karakehayov. InTech, 2012. https://doi.org/10.5772/48535.

Chen, T., de Jeu, R.A.M., Liu, Y.Y., van der Werf, G.R and Dolman, A.J., January 2014. Using Satellite Based Soil Moisture to Quantify the Water Driven Variability in NDVI: A Case Study over Mainland Australia.Remote Sensing of Environment 140: 330–38. https://doi.org/10.1016/j.rse.2013.08.022.

CESBIO, 2015. MACCS/MAJA, how it works. Online, accessed 13-June-2018.
URL http://www.cesbio.ups-tlse.fr/multitemp/?p=6203

Congedo, L., 2016 Semi-Automatic Classification Plugin Documentation. Release 6.0.1.1.
https://doi.org/10.13140/rg.2.2.29474.02242/1.

Deng, J. S., Wang, K., Deng, Y. H., Qi, G. J., 2008. PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data. International Journal of Remote Sensing, 29:16, 4823-4838. DOI: 10.1080/01431160801950162.

Lonjou, V., Desjardins, C., Hagolle, O., Petrucci, B., Tremas, T., Dejus, M., Makarau, A. and Auer, S., 2016. MACCS-ATCOR Joint Algorithm (MAJA). Remote Sensing of Clouds and the Atmosphere XXI, edited by Adolfo Comerón, Evgueni I. Kassianov, Klaus Schäfer, James W. Jack, Richard H. Picard, Konradin Weber, Proc. of SPIE Vol. 10001, 1000107. https://doi.org/10.1117/12.2240935.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E. in Wulder, M. A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote sensing of Environment, 148: str. 42–57. doi:10.1016/j.rse.2014.02.015

Orhan, O., Ekercin, S., & Dadaser-Celik, F., 2014. Use of Landsat Land Surface Temperature and Vegetation Indices for Monitoring Drought in the Salt Lake Basin Area, Turkey. The Scientific World Journal, 2014, 142939. http://doi.org/10.1155/2014/142939

Sobrino, J.A., Jimenez-Muoz J.C., Soria, G., Romaguera, M., Guanter, L., Moreno, J., Plaza, A. and Martinez, P., February 2008. Land Surface Emissivity Retrieval From Different VNIR and TIR Sensors." IEEE Transactions on Geoscience and Remote Sensing 46, no. 2: 316–27. https://doi.org/10.1109/TGRS.2007.904834.

Strahler, A. H., L. Boschetti, G. M. Foody, M. A. Friedl, M. C. Hansen, M. Herold, P. Mayaux, J. T. Morisette, S. V. Stehman, in C. E. Woodcock. 2006. Global Land Cover Validation: Recommendations: for Evaluation and Accuracy Assessment of Global Land Cover Maps, Report of Institute of Environmental Sustainability, Joint Research Centre, European Commission, Ispra, Italy.

Sulong, G., Sadiq, A. and Loay E.G., October 2015. Single and Multi-Source Methods for Reconstruction the Gaps in Landsat 7 ETM+ SLC-off Images. Research Journal of Applied Sciences, Engineering and Technology 11, no. 4: 423–28. https://doi.org/10.19026/rjaset.11.1797.

Sun, L. and Schulz K., June 2015. The Improvement of Land Cover Classification by Thermal Remote Sensing. Remote Sensing 7, no. 7: 8368–90. https://doi.org/10.3390/rs70708368

Tamm T., Zalite K., Voormansik K. and Talgre L., September 2016. Relating Sentinel-1 Interferometric Coherence to Mowing Events on Grasslands. Remote Sensing 8, no. 10: 801. https://doi.org/10.3390/rs8100802

USGS, 2018a. What are the band designations for the Landsat satellites? Online, accessed 11-June-2018. URL https://landsat.usgs.gov/what-are-band-designations-landsat-satellites

USGS, 2018b. Using the USGS Landsat 8 Product. Online, accessed 11-June-2018. URL https://landsat.usgs.gov/using-usgs-landsat-8-product

Zhe, Z. and Woodcock, C.E., March 2012. Object-Based Cloud and Cloud Shadow Detection in Landsat Imagery. Remote Sensing of Environment 118: 83–94. https://doi.org/10.1016/j.rse.2011.10.028.