Random forest classification analysis of Sentinel-2 and Landsat-8 images over semi-arid environment in the Eastern Mediterranean

Stavros Patsalidis
Department of Civil Engineering and Geomatics,
Cyprus University of Technology,
2-6, Saripolou str., 3603,
Limassol, Cyprus
si.patsalidis@edu.cut.ac.cy

Athos Agapiou
Department of Civil Engineering and Geomatics,
Cyprus University of Technology,
2-6, Saripolou str., 3603,
Limassol, Cyprus
athos.agapiou@cut.ac.cy

Diofantos G. Hadjimitsis
Department of Civil Engineering and Geomatics,
Cyprus University of Technology,
2-6, Saripolou str., 3603,
Limassol, Cyprus
d.hadjimitsis@cut.ac.cy

Abstract

Sentinel-2 land monitoring constellation mission aims to generate products similar with the Landsat-8 images, the world's longest continuously acquired collection of space-based land earth observation data. Though both sensors share similar spectral characteristics, their Relative Spectral Response Filters (RSRFs) are not identical. It is consequently important to assess whether and to what extent end-products, such as land use maps, may vary between these two sensors. For this purpose, the random forest classifier was applied over a semi-arid environment in the Eastern Mediterranean (Cyprus). Initially the Sentinel-2 image was sampled to the Landsat-8 spatial resolution. Then, two different classification strategies have been followed: the first one using an equal (balance) training sample between the 11 land use classes, while the second classification was based on a random training sample. In addition, land use maps were also generated based on maximum likelihood, mahalanobis distance and minimum distance pixel-based supervised classification algorithms. The overall results were evaluated based on kappa, overall, producer's and user's accuracies. Random forest classification has provided the best results with a kappa accuracy of 90% for both datasets while maximum likelihood algorithm has provided a kappa coefficient between 79.06% and 81.27% for Sentinel-2 and Landsat-8 sensors respectively. These results were much more improved compared to the mahalanobis and minimum distance classifiers, with an approximately kappa coefficient 69% and 66% respectively. In addition, the results obtained from the random forest have demonstrated that only a very small variance between the two datasets (Sentinel-2 and Landsat-8) exists (<3 % kappa coefficient), which can be due to the non-identical RSR filters of the sensor.

Keywords: land use classification; random forest; Sentinel-2; Landsat-8; semi-arid environment.

1 Introduction

Land use research is an important variable for many studies involving the Earth surface (Ban et al., 2015). Earth observation is a useful tool for mapping land use for various applications like urban expansion and urban heat phaenomenon (Zhao et al., 2017; Agapiou et al., 2015), hydrological applications (Gashaw et al., 2018; Paule-Mercado et al., 2017), desertification and land degradation (Padonou et al., 2017), land use changes (Akinyemi, 2017), climate change impacts (López et al., 2017) etc. Earth observation programs as the ones of Landsat and Sentinel can support these studies by providing freely distributed and systematic medium resolution optical data around the world.

The primary objective of the optical Sentinel-2 satellite mission is to generate data products comparable with the US Landsat-8 satellite as a result of the close cooperation between ESA and National Aeronautics and Space Administration (NASA) (ESA, 2018) and therefore to assurance data continuity and enhancement of the Landsat missions (Wang et al., 2016).

Indeed, the new Sentinel-2 Multi Spectral Imager instrument (MSI) has a set of bands with very similar spectral wavelengths to the Landsat-8 Operational Land Imager (OLI) (Flood, 2017), but with higher spatial and temporal resolutions (Quintano et al., 2018). Consequently, the Relative Spectral Response (RSR) filter of the Sentinel-2 MSI and Landsat-8OLI sensors have almost equivalent spectral bands. This is in correspondence to the earth observation data service continuity that Sentinel-2A mission should complements the SPOT and Landsat missions. The four bands (blue (490nm), green (560nm), red (665nm) and near infra-red (842nm)) at 10 m resolution of the Sentinel-2, ensure continuity with missions such as SPOT-5 or Landsat-8 and address user requirements, in particular, for basic land-cover classification. The six bands at 20 m resolution (4 narrow bands in the vegetation red edge spectral domain (705nm, 740nm, 775nm and 865nm) and 2 SWIR large bands (1610nm and 2190nm) dedicated to snow/ice/cloud detection, and to vegetation moisture stress assessment, satisfy requirements for enhanced land-cover classification (Baillarin et al., 2012; eoPortal Directory, 2018). In addition, the orbit of Sentinel-2 is fully consistent with SPOT and very close to the Landsat local time, allowing continuous blending of Sentinel-2 data with historical data from legacy missions to build long-term temporal series such as the Landsat images.

Nevertheless, as Mandanici and Bitelli (2016) state in their recent report the Relative Spectral Response (RSR) Filters of the Landsat-8 and Sentinel-8 are not identical, and therefore some differences and variations are expected in the recorded radiometric values. In the same study, the authors have concluded that it is possible to combine Landsat and Sentinel products, however evaluation assessments should be carried out prior to any specific application. It is therefore important to evaluate whether the results obtained from both sensors can provide or not comparable products (such as land use maps).

This study aims to examine whether Landsat-8 and Sentinel-2 sensors can provide comparable land use classification results, in semi-arid environments, within the accuracy limit of their spectral relative uncertainty (± 3%-5%). For this purpose the analysis is carried out using several existing supervised pixel-based classification techniques have been evaluated.

2 Methodology

For the aims of the study, one Landsat-8 and one Sentinel-2 images were used over the same area. The Landsat-8 (path: 176 / row: 36) and Sentinel-2 (granule: L1C_T36SVD_A011601_20170911T083628) images had a relatively small time-window difference, minimizing in this way any seasonal variations between the two datasets (11th of September 2017 and 15th of September 2017 for Landsat and Sentinel images respectively).

Landsat-8 and Sentinel-2 images were downloaded via the Earth Explorer platform at level Tier 1 (T1) and 1C respectively. Both these levels ensure that radiometric and geometric corrections have been applied. Tier 1 (T1) contains the highest quality Level-1 Precision Terrain (L1TP) data considered suitable for time-series analysis. The geo-registration is consistent and within prescribed tolerances (i.e. <12m root mean square error (RMSE)) (Landsat collections, 2018). Level-1C for Sentinel data refers to at-top of atmosphere reflectance with geometric corrections including orthorectification and spatial registration (ESA, 2018). Both datasets are set in the cartographic project of the Universal Traverse Mercator (UTM) and WGS84 ellipsoid. Then the Sentinel's bands were resampled to 30 meters pixel resolution (similar to Landsat-8 image) using a cubic convolution resampling method.

Equivalent bands of Sentinel-2 and Landsat-8 were then extracted for further processing from both datasets. These spectral bands include the visible part of the spectrum, near infra-red and the two short waves infra-red (SWIR) bands.

After the extraction of the specific bands from both sensors, a detail image sampling of training areas was carried out. In total more than 1700 samples were digitized, to be used for training purposes (70% of the total samples) while the rest of the samples (30%) were used to evaluate the classification results. The following land use classes were identified as shown in Table 1:

Table 1: Land use classes and samples used in the study

Class	Name	Samples
Class 1	Natural Waterbodies (deep)	79
Class 2	Natural Waterbodies (shallow)	87
Class 3	Artificial Waterbodies	111
Class 4	Urban Area and Artificial surfaces	172
Class 5	Grey Soil	91
Class 6	High Intensity vegetation	163
Class 7	Medium Intensity vegetation	219
Class 8	Low Intensity vegetation	134
Class 9	Stony Soil	148
Class 10	Agricultural Soil	210
Class 11	Bare Soil	288

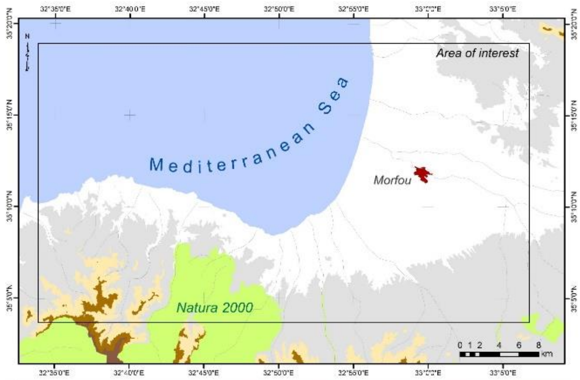
The random forest classification analysis was carried out in the R Project for Statistical Computing environment.

3 Case study area

The case study is located in the western part of Cyprus island covering an area of approximately 35 km length and 20 km width (700 km²) between 32o 30´ 00´´ - 33o 00´ 00´´E and 35o 00´ 00´´ - 35o 20´ 00´´N (WGS 84, Zone 36 North). The topography of the island is dominated by two mountain ranges, namely the Troodos Mountains in the central part of the island the Pentadaktylos Mountains in the northern part. Mesaoria central plain, lies between these two mountains and it is considered the primary agricultural region of the island. Despite its small size, Cyprus has a variety of natural vegetation including forests of conifers and broadleaved trees.

The area of interest is approximately 40-60 kilometres, north of Lemesos and Paphos towns and about 50 kilometres eastwards Lefkosia, the capital of Cyprus. Several artificial areas (small towns and villages) can be found in the area, while most of the area is covered by the Paphos and Troodos forest and other semi-natural areas. The western part is primarily used for agriculture. In addition, another part of the area in the south, is part of the EU-wide network of nature protection areas, Natura 2000 (highlighted with green colour in Figure 1 top).

Figure 1: Detail of the case study area (top), located in the western part of Cyprus (bottom left).



4 Results

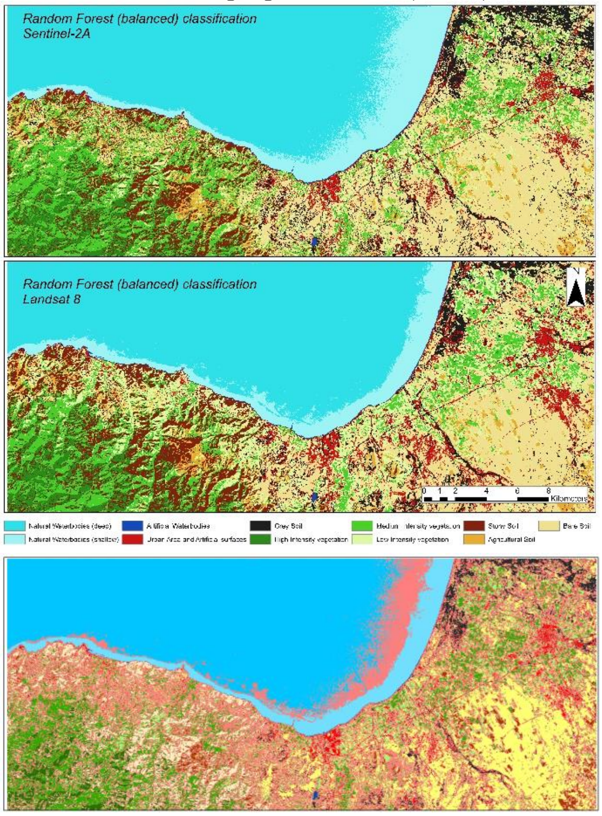
Based on the training areas for all 11 land use classes the random forest classification was applied in the images in the R environment. The balanced and imbalanced random forest classification results for Sentinel-2 and Landsat-8 are presented in Figures 2 and 3 respectively. The land use maps were able to map the Troodos and Paphos forests in the western part of the area as well as the agricultural fields in the eastern part. Urban and other artificial areas have been also spotted in the central and eastern part of the case study. Grey colour soil is noticeable in the northern part of the area near Morfou. This class is linked with the systematic depositing of the rivers.

Overall, both images have provided very similar results: 89.11% and 92.03% of kappa coefficient (for Sentinel-2 and

Landsat-8 respectively), which is within the spectral uncertainty range of the sensors using a balanced training dataset (scenario 1). Similarly, kappa coefficient was estimated to 87.26% and 90.65% based on a random subset of the training dataset (scenario 2). Concerning kappa, a value of 0 corresponds to a total random classification, while a kappa value of 1 represents a perfect agreement between the classification and reference data (Van Vliet, 2009; Yang et al., 2011). The overall accuracy for scenario 1 was found 90.27% for Sentinel-2 while for Landsat-8 was estimated to 92.88%. For the second scenario, these accuracies were calculated to 88.35% and 90.65% respectively (see Table 1).

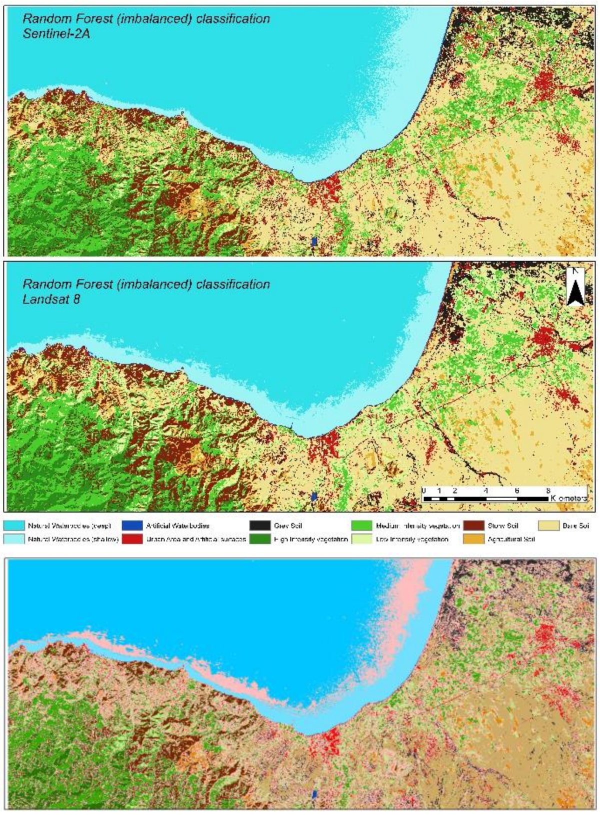
Most of the land use classes have also demonstrated high accuracies (> 74%) while only a limited number of classes such as Class 4 (Urban Area and artificial surfaces), Class 5 (Grey soil) and Class 8 (Low Intensity vegetation) have shown reduced producer's accuracies. Water bodies (Classes 1 -3: Natural Waterbodies (deep); Natural Waterbodies (Shallow); Artificial Water bodies) have been successfully classified with extremely high accuracies. User’s accuracies for balanced training sample strategy, has also provided very high accuracies, since most of the classes have been mapped with more than 85% correctness.

Figure 2: Random forest classification results following a balanced training sampling for Sentinel-2A (top) and Landsat-8 (middle) images. Differences between Sentinel-2A and Landsat 8 are highlighted with red (bottom)



Comparable results have been obtained between balanced and imbalanced training areas for each sensor -individually- based on both producer’s and user’s accuracies. Though, some variances exist for Class 4 (Urban Area and artificial surfaces). The specific class has been already mentioned by Alexakis et al., (2012) as a problematic to be classified (for the area of Cyprus) due to its high correlation of the spectral profile with other soil classes. This strongly affects the performance and the accuracy of the satellite classification. Indeed, the specific class (Urban Area and artificial surfaces) has the lowest ranking -with Class 8 (Low Intensity vegetation) in terms of accuracy between all classes for the random forest classifier.

Figure 3: Random forest classification results following an imbalanced training sampling for Sentinel-2A (top) and Landsat-8 (middle) images. Differences between Sentinel-2A and Landsat 8 are highlighted with red (bottom)



In addition, some significant statistical differences exist between the two sensors for Class 4 (Urban Area and artificial surfaces): producer’s accuracy for Class 4 was around 67% for Sentinel-2 while for Landsat-8 the accuracy was 86%. However, user’s accuracy was high for both images providing a high degree of reliability of the product (i.e. land use map).

To evaluate the potentials of the random forest classifier with other known classifiers, we have compared the results with the maximum likelihood, mahalanobis distance and minimum distance classification products. As mentioned earlier in Section 2 (methodology), the training areas of these classifications have remained the same as the one used before in the random forest. The spectral similarity of the Urban Area and artificial surfaces class (Class 4) with the soil class (Class 11) has resulted an over-estimation of the specific class.

Significant differences exist between the various classifications. For example, the producer's accuracy for Class 4 (Urban Area and artificial surfaces) is less than 32% for Sentinel-2 image after the mahalanobis and minimum distance classification while the score for the maximum likelihood in 70%. Similar difference is found for other classes such Class 5 (Grey soil) and Class 9 (Stony soil), while for Landsat-8 the accuracy varied from 58% up to 76% for all classifications results. Classes such as grey soil (Class 5), high intensity vegetation (Class 6), medium intensity vegetation (Class 7) and low intensity vegetation (Class 8) have provided low producer's and user's accuracies.

Overall the maximum likelihood classification has generated the highest kappa coefficients (79.06% and 81.27% for Sentinel-2 and Landsat-8 respectively) followed by the mahalanobis distance (69.30% and 69.20%) and the minimum distance classifications (66% and 67%). These results are lower than the random classifier (either using a balanced or imbalanced training areas). The overall accuracy of the three classifications range from 69.00% (minimum distance for Sentinel-2) to 83% (maximum likelihood for Landsat-8).

Table 2: Random forest classification accuracies for Sentinel-2 and Landsat-8 images using balanced and imbalanced training sampling among the classes

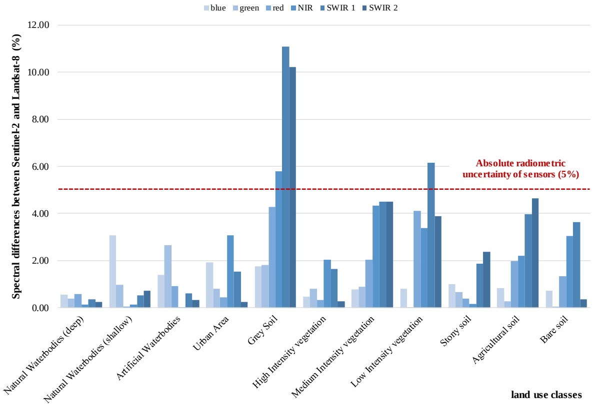
Random Forest classification results		Sentinel-2 Random Forest (Balanced)	Sentinel-2 Random Forest (Imbalanced)	Landsat-8 Random Forest (Balanced)	Landsat-8 Random Forest (Imbalanced)
Kappa statistics		89.11%	87.26%	92.03%	89.69%
Overall Acc.		90.27%	88.35%	92.88%	90.65%
Natural Waterbodies (deep)	**Producer's Acc.**	100.00%	100.00%	100.00%	100.00%
Natural Waterbodies (Shallow)		100.00%	100.00%	100.00%	100.00%
Artificial Water bodies		100.00%	100.00%	96.67%	96.67%
Urban Area & artificial surfaces		65.96%	68.09%	86.79%	84.91%
Grey soil		74.00%	92.59%	87.50%	87.50%
High Intensity vegetation		93.88%	95.92%	98.00%	96.00%
Medium Intensity vegetation		94.12%	86.76%	96.88%	89.01%
Low Intensity vegetation		65.51%	79.31%	73.91%	71.74%
Stony soil		92.31%	80.77%	100.00%	100.00%
Agricultural soil		89.51%	95.52%	98.41%	98.40%
Bare soil		81.61%	72.41%	88.51%	82.76%
Natural Waterbodies (deep)	**User's Acc.**	97.70%	98.30%	98.70%	99.80%
Natural Waterbodies (Shallow)		99.50%	99.00%	97.80%	98.40%
Artificial Waterbodies		99.70%	99.80%	100.00%	99.60%
Urban Area & artificial surfaces		84.40%	86.60%	90.10%	83.20%
Grey soil		81.70%	78.20%	92.10%	76.20%
High Intensity vegetation		95.90%	94.10%	97.10%	95.50%
Medium Intensity vegetation		89.50%	87.90%	88.80%	89.30%
Low Intensity vegetation		73.90%	74.60%	82.50%	76.40%
Stony soil		96.40%	95.60%	96.10%	92.30%
Agricultural soil		95.60%	96.50%	95.50%	94.80%
Bare soil		85.80%	76.50%	89.20%	87.80%

5 Discussion

The overall statistics of the random forest classification, confirm that both Landsat-8 and Sentinel-2 sensors can provide similar classification results (with a range less than 3% for kappa coefficient) in semi-arid environments such as the one of Cyprus. Especially, using a balanced training sample strategy the user's and producer' accuracies are almost identical for most of the classes, providing a unique tool for the systematic land use mapping in the future.

Figure 4 presents the spectral differences observed between the Sentinel-2 and Landsat-8 sensors for the 11 land use classes used in this study. As it is shown, the small spectral differences between the two sensors do not affect significantly the land use classification for most of the classes (9 out of the 11 total classes used here). However, two specific classes (grey soil and low intensity vegetation) seem to exceed the absolute radiometric uncertainty of the sensors (5%), confirming the general conclusions made by Mandanici and Bitelli (2016), that for specific applications (as those of the classification in semi-arid environments) should be firstly evaluated, before any combination of Landsat and Sentinel products. Indeed, these two classes have given different producer's and user's accuracies. The producer's accuracy was estimated to 74.00% and 87.50% for grey soil and 65.51% and 73.91% for low intensity vegetation (for Sentinel-2 and Landsat-8 respectively), while the user's accuracy was found 81.70% and 92.10% for grey soil and 73.90% and 82.50% for low intensity vegetation.

Figure 4: Spectral differences between Sentinel-2 and Landsat-8 sensors for the various land use classes examined in the study.



6 Conclusions

Land use maps are essential tools for various applications. The freely distributed Landsat and Sentinel data can provide systematic cloud-free optical data around the globe. This new generation of space-borne sensors is generating nearly continuous streams of massive remote sensing imageries sending several Terabytes of information every day to the satellite data centres (Ma et al., 2015)].

Though that Sentinel mission aims to provide "identical" spectral data as the one provided by the Landsat series, their RSR filters are not identical. Therefore, it was important to examine whether this small spectral mismatch is affecting classification results and land use maps. Eastern Mediterranean basin -a region characterized by the high frequency of cloud-free images- was selected as a case study. For this purpose, two images over Cyprus with only 4 days difference have been acquired and processed. Training areas have been used in both datasets and then the classification analysis was carried out using the random forest classifier, maximum likelihood, mahalanobis distance and minimum distance. At the end of this analysis, the classification accuracy has been reported.

The results have shown that both Sentinel-2 and Landsat-8 can provide almost identical land use maps especially when sophisticated classifiers as the one of random forest is applied. The results were in general agreement between most of the land use classes used in this supervised classification process. The difference for the kappa coefficient was limited to less than 3%. However, for specific land use classes a variance was observed between the two sensors indicating that further experiments should be made prior to the use of combined Sentinel-2 and Landsat-8 products. Overall, the Sentinel-2 sensor has provided slightly better results compared to the Landsat-8.

Both Landsat and Sentinel optical data can be used as an integrated tool towards the systematic monitoring and mapping of semi-arid environments, providing medium resolution land use maps, taking into considerations the specific problems raised for each case study. The new revised CORINE land use land cover products covering the period 2012-2018, which is still under evaluations, can be also used so as to evaluate further these (and similar) results and therefore to evaluate whether the optical Sentinel 2A and 2B sensors can be integrated as a systematic tool by the Department of Environment of Cyprus for delivering CORINE maps. In the near future, it is expected that the authors would experiment towards the harmonization of the two datasets taking into consideration the different spatial resolution of the satellites.

Acknowledgments

The authors would like to acknowledge Sentinel-2 imagery from European Space Agency (ESA) and Landsat-8 image courtesy of the U.S. Geological Survey. The results are part of the diploma thesis of the S.P.

References

Agapiou A., Alexakis D.D., Lysandrou V., Sarris A., Cuca B., Themistocleous K. and Hadjimitsis D.G. (2015) Impact of urban sprawl to cultural heritage monuments: The case study of Paphos area in Cyprus, *Journal of Cultural Heritage*, 16, 5, 671-680

Akinyemi O.F. (2017) Land change in the central Albertine rift: Insights from analysis and mapping of land use-land cover change in north-western Rwanda, *Applied Geography*, 87, 127-138

Alexakis D., Agapiou A., Hadjimitsis D. and Retalis A. (2012) Optimizing statistical classification accuracy of satellite remotely sensed imagery for supporting fast flood hydrological analysis, *Acta Geophysica*, 60, 3, 959-984

Baillarin S. J., Meygret A., Dechoz C., Petrucci B., Lacherade S., Tremas T., Isola C., Martimort P. and Spoto F. (2012) Sentinel-2 level 1 products and image processing performances, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXIX-B1, 2012, XXII ISPRS Congress, 25 August – 01 September 2012, Melbourne, Australia

Ban Y., Peng G. and Giri C. (2015) Global land cover mapping using Earth observation satellite data: Recent progresses and challenges. *ISPRS Journal of Photogrammetry and Remote Sensing*. 10.1016/j.isprsjprs.2015.01.001.

Corine Land Cover 2012, https://www.eea.europa.eu/data-and-maps/data/external/corine-land-cover-2012 [last accessed: 31st March 2018].

eoPortal Directory, Copernicus: Sentinel-2, https://directory.eoportal.org/web/eoportal/satellite-missions/c-missions/copernicus-sentinel-2 [last accessed: 17th February 2018].

European Space Agency, Observing the Earth, Copernicus, Sentinel-2, https://www.esa.int/Our_Activities/Observing_the_Earth/Cop ernicus/Sentinel-2/Satellite_constellation [last accessed: 17th February 2018].

European Space Agency, Sentinel-2 MSI, 2018 https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/product-types/level-1c [last accessed: 17th February 2018].

Flood, N. (2017) Comparing Sentinel-2A and Landsat 7 and 8 Using Surface Reflectance over Australia, *Remote Sens*. 9, 659.

Gashaw T., Tulu T., Argaw M. and Worqlul W.A. (2018) Modeling the hydrological impacts of land use/land cover changes in the Andassa watershed, Blue Nile Basin, Ethiopia, *Science of The Total Environment*, 619–620, 1394-1408

Landsat collections, 2018, https://landsat.usgs.gov/landsat-collections#C1_Tiers [last accessed: 17th February 2018].

López S., Wright C. and Costanza P. (2017) Environmental change in the equatorial Andes: Linking climate, land use, and land cover transformations, *Remote Sensing Applications: Society and Environment*, 8, 291-303

Ma Y., Wang L., Liu P. and Ranjan R. (2015) Towards Building a Data-intensive Index for Big Data Computing – A Case Study of Remote Sensing Data Processing, *Information Sciences* 319: 171–188. doi:10.1016/j.ins.2014.10.006

Mandanici E. and Bitelli G. (2016) Preliminary Comparison of Sentinel-2 and Landsat-8 Imagery for a Combined Use. *Remote Sens*., 8, 1014.

Padonou A.E., Lykke M.A., Bachmann Y., Idohou R. and Sinsin B. (2017) Mapping changes in land use/land cover and prediction of future extension of bowé in Benin, West Africa, *Land Use Policy*, 69, 85-92

Paule-Mercado M.A., Lee B.Y., Memon S.A., Umer S.R., Salim I. and Lee C.-H. (2017) Influence of land development on stormwater runoff from a mixed land use and land cover catchment, *Science of The Total Environment*, 599–600, 2142-2155

Quintano C., Fernández-Manso A. and Fernández-Manso O. (2018) Combination of Landsat and Sentinel-2 MSI data for initial assessing of burn severity, *International Journal of Applied Earth Observation and Geoinformation*, 64, 221-225.

Van Vliet J. (2009) Assessing the accuracy of changes in spatial explicit land use change models. In: *12th Annual AGILE Int. Conf. on Geographic Information Science 2009,* Leibniz Universitat Hannover, Germany

Wang L., Zhou X., Zhu X., Dong Z. and Guo W. (2016) Estimation of biomass in wheat using random forest regression algorithm and remote sensing data, *The Crop Journal*, 4, 3, 212-219

Yang C., Everitt J.H. and Murden D. (2011) Evaluating high resolution SPOT 5 satellite imagery for crop identification, *Comput. Electron. Agr.* 75, 2, 347-354, DOI: 10.1016/j.compag.2010.12.012.

Zhao Z.-Q., He B.-J., Li L.-G., Wang H.-B. and Darko A. (2017) Profile and concentric zonal analysis of relationships between land use/land cover and land surface temperature: Case study of Shenyang, China, *Energy and Buildings*, 155, 282-295