AGILE: GIScience Series, 3, 38, 2022. https://doi.org/10.5194/agile-giss-3-38-2022
Proceedings of the 25th AGILE Conference on Geographic Information Science, 2022.
Editors: E. Parseliunas, A. Mansourian, P. Partsinevelos, and J. Suziedelyte-Visockiene.
This contribution underwent peer review based on a full paper submission.

Classifying Urban Green Spaces using a combined Sentinel-2 and Random Forest approach

Irada Ismayilova[1] and Sabine Timpf[1]

[1]Geoinformatics Group, Institute of Geography, University of Augsburg, Augsburg, Germany

Correspondence: Irada Ismayilova (irada.ismayilova@geo.uni-augsburg.de)

Abstract. Environmental and human benefits of Urban Green Spaces (UGSs) have been known for a long time. However, the definition of a reasonable greening strategy still remains challenging due to the lack of sufficient baseline information as well as a lack of consensus what constitutes a UGS. Therefore, accurate identification of the existing green spaces in cities is crucial for developing UGS inventories for urban planning and resource management activities. In this paper we explore the potential of freely available highest resolution multi-spectral remote sensing imagery to identify large homogeneous as well small heterogeneous UGSs. The approach of using a Random Forest classification on Sentinel-2 imagery is shown to be useful to identify various UGSs with a 97 % accuracy. Freely available data and a relatively straightforward implementation of the proposed approach makes it a valuable tool for decision and policy makers.

Keywords. urban green spaces, Sentinel-2, random forest

1 Motivation

The list of services provided by Urban Green spaces (UGSs) is very extensive: e.g., reduction of urban heat islands (Sun and Chen, 2017), improvement of micro climate in densely populated urban centers (Buyadi et al., 2013), and positive effects on human well-being (Reyes-Riveros et al., 2021) are only few of these services. Precise maps of UGSs should provide a basis for urban planners and resource managers (Chen et al., 2021). Yet, two major questions have not been adequately addressed: which types of green spaces do exist (ontology) and how to precisely and cost-effectively map all the different UGS types (identification).

A typical source of detailed UGS information are Land Use and Land Cover (LULC) maps. At the European level, the Urban Atlas (UA) (Montero et al., 2014) is one of the most frequently used ones. Additional information can be found for instance in Open Street Map (OSM) or in maps and geo-databases provided by local authorities. In our search, we stumble upon the fact that on the one side detailed local LULC maps are not free of charge or simply do not exist; on the other side, UA is not detailed enough to represent UGSs at different scales, sizes and types. Due to the minimum mapping unit (MMU), only green spaces larger than 0.25 ha are included into the UA (Ludwig et al., 2021). Further, the role of small green spaces is not always well understood (Semeraro et al., 2021), and this could be another reason why e.g., urban gardens that store considerable amount of greenery, are part of the class "sports and leisure activity" instead of the UGS class. These observations show that there is a need to establish which types of green spaces might exist in an urban area, their sizes and whether they are spatially heterogeneous or homogeneous. It is then a policy decision, which of the green spaces should be included in the UGS class. This information should serve as a basis for the creation of detailed UGS maps.

Most novel approaches to develop LULC maps include the usage of remote sensing (RS) data and an implementation (or application) of machine learning (ML) or deep learning (DL) techniques (Abdi, 2020; Singh et al., 2021). The mapping of UGSs is approached in a similar manner (Huerta et al., 2021; Chen et al., 2021). The success of an accurate mapping of UGSs depends on many factors, of which the accuracy of the input data, its spatial resolution and the selected classification methods are the most common ones. Depending on the final aim and level of detail, authors utilize medium resolution Landsat imagery (Huang et al., 2018) as well as high resolution Sentinel-2 imagery (Chen et al., 2021) and very high resolution imagery (Haase et al., 2019; Huerta et al., 2021). The later approach appears to be particularly helpful to identify small-scale UGss, such as greenery in front and backyard gardens (Haase et al., 2019). The drawback of very-high resolution imagery is in its acquisition costs and the increased processing costs. Therefore, many authors utilize freely available geodata and heavily rely on the performance of the selected classification method.

Random Forest (RF) as an ensemble machine learning

(ML) classification method is the most frequently used LULC classification technique. Many studies including, but not limited to Talukdar et al. (2020) and Abdi (2020) confirm that a combination of Sentinel-2 data with especially RF has a great potential to accurately classify LULC. The usage of RS data for UGSs is challenging in terms of the mixed pixels and spectral variability of vegetation (Chen et al., 2021). Yet many studies show that RF can successfully deal with these challenges and produce accurate LULC maps (Phan et al., 2020; Feng et al., 2015). Consequently, in this paper, we investigate the potential of the combination of Sentinel-2 data together with the RF classification method to identify those UGSs in the study area that are both spatially homogeneous (e.g. forest) and heterogeneous (e.g. greenery in gardening areas) and have different sizes (large, small). We also evaluate the quality of the utilized approach in terms of its accuracy and implementation ease for decision and policy makers.

The paper is structured as follows: the following section details the workflow of the automated classification process, including a description of preparing the geodata (deriving diverse indices), training data acquisition as well as the classification modelling using Random Forest. Section 3 presents the results for our case study area. In section 4, we discuss the results and conclude with insights and future work.

2 Workflow of the classification process

We identify UGSs in Augsburg, Germany with a combination of RS data and the RF ML method. The study area is comprised of different UGS types such as forest, parks, green corridors alongside roads and rivers, urban gardening areas and others. The total size of the study area is approximately 147 km2. According to the UA, this includes 38km2 of forest and 6 km2 of green urban areas. These figures do not include greenery in gardening areas that are part of the sports and leisure facility class in the UA as well as greenery between and around residential and industrial areas.
The workflow of the UGS classification approach implemented in this study is given in Fig. 1 and described in the the following sections.

2.1 Preparation of RS data

In this analysis we utilize single date Sentinel-2A imagery from 14th of August 2021. Sentinel-2 is a multi-spectral imagery, collected in 13 spectral bands with a spatial resolution varying between 10-60 meters. For the analysis we download the freely available data from the Copernicus Open Access Hub[1]. Sentinel-2A data has already been radiometrically, geometrically and atmospherically corrected. Therefore,in the pre-processing step we only define the correct coordinate reference system, resample all the bands to a 10m spatial resolution and clip them to the extent of the study area. We further use the bands 2-8 and 8A in the analysis. These pre-processed spectral bands serve as basis for the calculation of vegetation indices.

[1]https://scihub.copernicus.eu

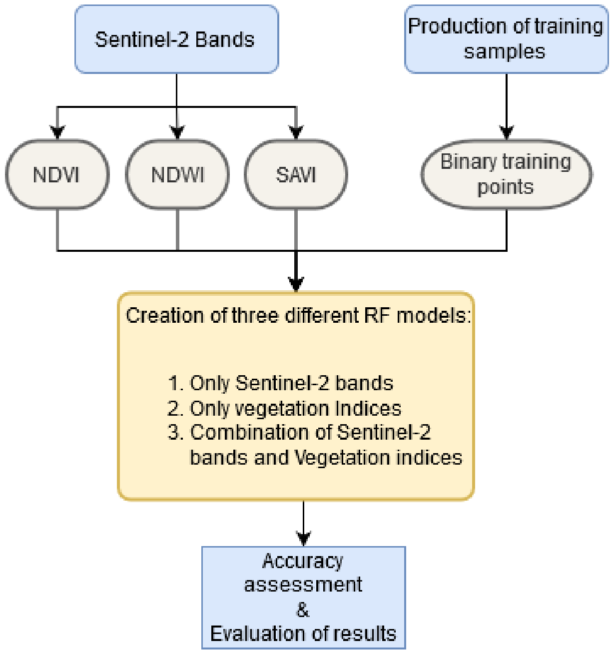

Figure 1. The Flowchart of UGSs classification using RS and ML method.

2.1.1 Vegetation indices

Vegetation Indices (VI) are mathematical combinations of different spectral bands of remotely sensed data. They prove themselves useful as they reduce soil, and atmospheric effects as well as enhance the information contained in single spectral bands. Further, these indices help to bring out the variability in the vegetation characteristics (Viña et al., 2011). As we intend to identify UGSs, which consist of different types of greenery, we choose the following vegetation indices from the Sentinel-2 index database (Henrich et al., 2009): Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI) and Soil-Adjusted Vegetation Index (SAVI). In the following, we provide short descriptions of the selected indices as well as the equations used to calculate them.

NDVI

NDVI is used to differentiate healthy vegetation and based on the knowledge that chlorophyll absorbs the red light while the mesophyll leaf structure scatters near-infrared light (NIR). NDVI values range from -1 to +1, where positive values represent healthy vegetation and negative values indicate an absence of or sparse vegetation (Myneni et al., 1995).

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (1)$$

NDWI

NDWI is a VI that helps to delineate healthy vegetation versus vegetation affected by drought or water shortage as it is sensitive to changes of the liquid water content of vegetation. It is calculated based on the ratio of NIR and green bands (Gao, 1996).

$$NDWI = \frac{Green - NIR}{Green + NIR} \quad (2)$$

SAVI

VIs can be affected by background noises such as soil brightness, soil moisture and saturation effects from high density vegetation. To reduce this noise, transformation techniques are developed and applied to create new indices. SAVI is one of these indices and it helps to minimize the soil brightness noise that might be present in e.g., NDVI. It is calculated based on red and NIR bands while applying an additional canopy background adjustment factor (Huete, 1988).

$$SAVI = \frac{(NIR - Red) * (1 + L)}{NIR + Red + L} \quad (3)$$

where L is a soil brightness correction factor. In this study L is equal to 0.5 as this value is suitable for most land cover types.

2.2 Training data acquisition and Software availability

The amount and the quality of the training data, apart from being a key part of the classification procedure, can have a large impact on the actual classification accuracy (Millard and Richardson, 2015). We collect training data for model development using very-high resolution Google Earth imagery as well an orthophoto. Since we perform two class classification, we sample data that is either UGSs or any other LULC type (e.g. roads, buildings, water). We explicitly pay attention to include samples from all possible greenery types such as deciduous and coniferous trees, herbaceous parks etc. Yet, we keep out areas that do include "greenery" which is not an UGS, e.g. pitches. We randomly distribute the training points over the study area as suggested in Millard and Richardson (2015).
We further confirm the accuracy of the training points samples using a very-high resolution orthophoto. Only correctly sampled points are later included into the analysis. In total, 1500 pure training points, in a 50:50 ratio for both classes (green vs anything else), are utilized in the final models. There is no accepted training data size in the existing literature that can perfectly suit every spatial classification task. Therefore, we proceed with 1500 points as they adequately represent our both classes.
The pre-processing of the data and the classification analysis have been performed using R software (R Core Team, 2020). For reproducibility purposes, we provide bands of the Sentinel-2 image, a portion of the training data as well as the complete R code here: https://doi.org/10.6084/m9.figshare.19551022.v1. We complement the R code with necessary comments for ease of understanding.

2.3 Random Forest classification modelling

In order to identify all the existing green spaces in the study area we use a Random Forest (RF) classification method. RF as proposed by Breiman is an ensemble learning method that can be utilized for both classification and regression tasks (Breiman, 2001). We choose RF due to its proven high performance with noisy data and its insensitivity to the initialization of parameters (Dong et al., 2019). In the process, RF implements bootstrap sampling (random sub sampling) to grow the forest and the majority of votes by each tree determines the final classification result. RF allows the optimization of the the results by a parameter tuning process. Two parameters that can be manipulated are the number of variables used to split a node (*mtry*) and the number of trees grown in the forest (*ntree*). Unlike many other classifiers, an increase in the *ntree* leads to an improved classification performance (Breiman, 2001). Thus, this parameter must be wisely adjusted.
For identification of UGSs, apart from the RF model, we also explore the potential of single spectral bands of Sentinel-2 data (Model 1), VIs derived from the spectral bands (Model 2) and a combination of spectral bands with VIs (Model 3). Thus we build three different RF models including the mentioned variations. Here, we use the 10 fold cross validation approach for training the model. It is a data partitioning strategy, where RF utilizes 9 parts of the data set to train the model and one part to test the model. The process is repeated until all the parts have been trained and tested. This approach is less biased in comparison to the traditional train/test split. The performance of the RF classifier is assessed with the overall classification accuracy and the out of bag error rate (OOB), sensitivity and specificity values. After we find the model with a satisfactory accuracy, we use this model to make predictions on the whole study area.
Based on the final prediction we calculate the total amount of greenery identified with the RF classification. We further inspect the UGS types, which the identified areas belong to. For comparison purposes we refer to the latest Urban Atlas.

3 Results

In this work we classify Sentinel-2 imagery using three RF models in order to identify all UGSs in the study area. We run each RF model three times, each with various combinations of explanatory variables in order to evaluate the efficacy of these variables. (Table 1).

Table 1. Table of implemented RF models and explanatory information.

	Model 1	*Model 2*	*Model 3*
Total number of observations	1453	1453	1453
Total number of explanatory variables	8	3	11
Train/test split	10 fold cross validation	10 fold cross validation	10 fold cross validation
Best mtry	6	2	4
Overall accuracy	~ 90%	~ 98%	~ 97%
OOB error	9.43%	1.4	2.1%
Sensitivity	0.9564	0.8790	0.9444
Specificity	0.8928	0.8578	0.9057

Model 1 contains bands 2-8 and 8A of the Sentinel-2 imagery. The accuracy of the model reaches about 90 %. We then check which predictor variables have the highest influence on the prediction accuracy. The variable importance calculation shows that the red and NIR bands bring the highest contribution to the model accuracy. We repeat the same procedure with Model 2, which contains the VIs, namely NDVI, NDWI and SAVI. The accuracy of the classification hits nearly 98 % with the NDWI and SAVI indices being the slightly more influential explanatory variables. The performance of the final Model 3, which contains raw spectral bands and VIs as predictive variables, is at around 97 %. SAVI, NDVI and Red band appear to be the most contributing predictive variables.
Comparing the three models, Model 3 has a slightly better performance in terms of all accuracy estimates, especially the sensitivity (ability to predict true positives) and the specificity (ability to predict true negatives) of the model. Therefore, we use Model 3 as the final model to make predictions to the whole study area.

In total we classify approximately 56 km2 of UGSs seen on Fig. 2. We further perform detailed examinations of the results and compare them with the Urban Atlas (UA). We calculate the area of the same UGSs class in UA and in the RF classification. The post-classification examination reveals that RF identifies 36 km2 of forest UGSs in the study area that constitutes of 38 km2 in the UA. The class Green urban spaces of the UA covers 6 km2 of the study area while RF classify 4.3 km2 into the same category. The class Sports and Leisure activity of the UA includes urban gardening areas among other classes. The area of this class in the UA is around 6 km2 and within this class RF identifies about 3 km2 of green space. Another 13 km2 accounted for green space by RF are covered by different land uses in UA.

4 Discussion and Conclusions

In this paper, we aim at an automated identification of all possible green spaces in a study area, especially the ones

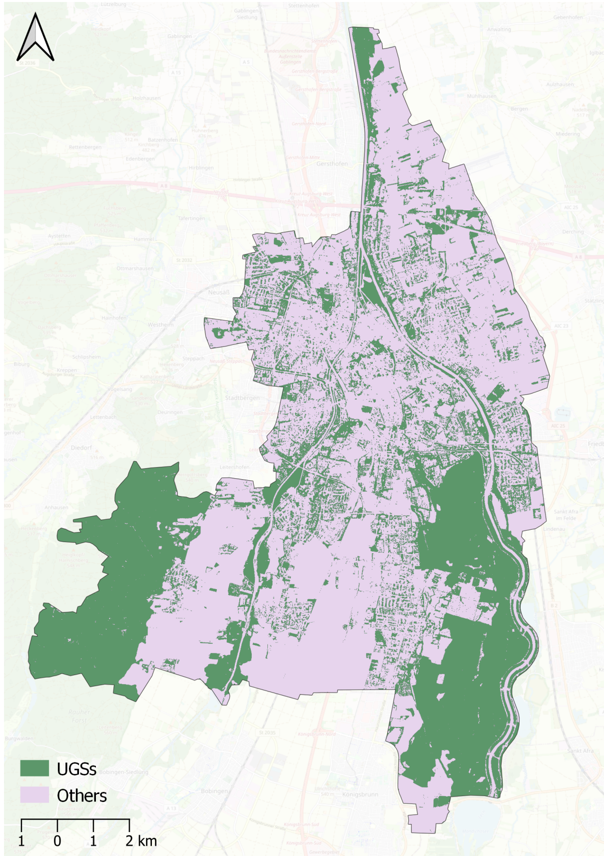

Figure 2. Results of the UGS classification using Sentinel-2 and RF method.

that are not a part of LULC maps or are considered under different LULC classes. For this purpose, we first evaluate whether single spectral bands or Vegetation Indices (VIs) or the combination of both can help to precisely identify green spaces. The results reveal that the selected approach has great potential in green space classification. However, there are no dramatic differences between the implemented models and both single bands as well as VIs are performing well. We use VIs to emphasize the spectral behavior of vegetation in contrast to other targets on the surface and because according to Viña et al. (2011), VIs are shown to be helpful to distinguish vegetation. Yet, similarly to da Silva et al. (2020), we do not observe significant performance differences between the selected VIs. We chose Random Forest (RF) as a classifier based on its high performance in the existing literature, such as Abdi (2020) and Talukdar et al. (2020). RF is also straightforward in its implementation and the results are relatively easy to interpret. Our expectation is mostly fulfilled as in our setup RF reaches a very high classification accuracy in every model.
Many model validation approaches do exist. On one hand, we can calculate statistical measures such as overall classification accuracy, OOB error rate and others (Dong et al., 2019). On the other hand, ground truth data can be used to validate the classified areas (Carranza-García et al., 2019). In this study, we rely on the first approach. The overall

accuracy of the final model is around 97 % with especially high rate of correctly classified true positive and negative classes. Since we use the model trained on nearly 1500 training points and make final predictions on the total study area, there is a need to validate predictions outside of those 1500 training points. However, ground truth data collection is time and labor expensive. Therefore, we conduct a visual confirmation study using very high resolution Orthophotos.

Our visual examination confirms that although overall classification is very good, there are still areas that are wrongly classified. This can be observed in areas such as narrow river banks or in areas with very heterogeneous green-space vs build-up-area composition. We think this might be due to the spatial resolution of the utilized data. Sentinel-2 is considered high resolution remotw sensing data (Chen et al., 2021), yet we find a 10m resolution to be still too coarse for reliable identification of greenery in such complex areas as urban gardens or back and front yard gardens as well as very patchy green spaces e.g. between buildings.

The outcome of this study shows that greenery in large and homogeneous areas such as forests and parks can be very accurately identified using Sentinel-2 data with the RF classification. For instance, RF identifies about 2 km2 less vegetation in the forest than the urban atlas (UA). UA generalizes and smooths the polygons and thus eliminates very small patches that belong to another class. This combination of data and method can also very well distinguish between UGSs and e.g. ball pitches that are also "green". Further, the green corridors alongside water bodies and roads are also well detected with the used methodology unless the corridors are too narrow and the 10m resolution is not sufficient. In this case the usage of even higher resolution data is required. Finally, we can also detect the green areas that exist in urban gardens or back and front yard gardens. Urban gardens are included in the sport and leisure facility class in UA, while our results show that in the study area this class accommodates nearly 3 km2 of greenery. Back and front yard gardens are typically privately owned and do not appear in LULC maps. Further, we acknowledge the need of usage of very high resolution imagery also in this case, as gardening areas a very small and the precise detection of vegetation is not very accurate with Sentinel-2 data. The successful examples of usage e.g. RapidEye for garden vegetation identification, is showcased by Haase et al. (2019). Finally, we identify a substantial amount of vegetation between and around residential and industrial areas. This information, to our knowledge, is not included in any LULC map. However, these areas also contribute to the overall greenness of cities. Similarly to gardening areas, this type of UGSs might be better to identify with very high resolution imagery.

Concluding, the approach proposed in this study is well suited for the studies or decision making processes that require knowledge of the overall amount of greenery in urban areas. However, if the task is to identify how many different green space types exist such as parks, gardens etc., there is a need of a more object oriented classification approach. In our future studies we will concentrate on the latter one by using a combination of RS imagery and machine learning/deep learning methods.

5 Acknowledgment

The author is grateful for the financial support of the Young Researchers Travel Scholarship Program of the University of Augsburg

References

Abdi, A. M.: Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data, GIScience & Remote Sensing, 57, 1–20, 2020.

Breiman, L.: Random forests, Machine learning, 45, 5–32, 2001.

Buyadi, S. N. A., Mohd, W. M. N. W., and Misni, A.: Green spaces growth impact on the urban microclimate, Procedia-Social and Behavioral Sciences, 105, 547–557, 2013.

Carranza-García, M., García-Gutiérrez, J., and Riquelme, J. C.: A framework for evaluating land use and land cover classification using convolutional neural networks, Remote Sensing, 11, 274, 2019.

Chen, Y., Weng, Q., Tang, L., Liu, Q., Zhang, X., and Bilal, M.: Automatic mapping of urban green spaces using a geospatial neural network, GIScience & Remote Sensing, 58, 624–642, 2021.

da Silva, V. S., Salami, G., da Silva, M. I. O., Silva, E. A., Monteiro Junior, J. J., and Alba, E.: Methodological evaluation of vegetation indexes in land use and land cover (LULC) classification, Geology, Ecology, and Landscapes, 4, 159–169, 2020.

Dong, L., Du, H., Mao, F., Han, N., Li, X., Zhou, G., Zheng, J., Zhang, M., Xing, L., Liu, T., et al.: Very high resolution remote sensing imagery classification using a fusion of random forest and deep learning technique—Subtropical area for example, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 113–128, 2019.

Feng, Q., Liu, J., and Gong, J.: UAV remote sensing for urban vegetation mapping using random forest and texture analysis, Remote sensing, 7, 1074–1094, 2015.

Gao, B.-C.: NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space, Remote sensing of environment, 58, 257–266, 1996.

Haase, D., Jänicke, C., and Wellmann, T.: Front and back yard green analysis with subpixel vegetation fractions from earth observation data in a city, Landscape and urban planning, 182, 44–54, 2019.

Henrich, V., Götze, E., Jung, A., Sandow, C., Thürkow, D., and Gläßer, C.: Development of an online indices database: Motivation, concept and implementation, in: Proceedings of the 6th EARSeL imaging spectroscopy sig workshop innovative tool for scientific and commercial environment applications, Tel Aviv, Israel, pp. 16–18, 2009.

Huang, C., Yang, J., and Jiang, P.: Assessing impacts of urban form on landscape structure of urban green spaces in China using Landsat images based on Google Earth Engine, Remote Sensing, 10, 1569, 2018.

Huerta, R. E., Yépez, F. D., Lozano-García, D. F., Guerra Cobián, V. H., Ferriño Fierro, A. L., de León Gómez, H., Cavazos González, R. A., and Vargas-Martínez, A.: Mapping Urban Green Spaces at the Metropolitan Level Using Very High Resolution Satellite Imagery and Deep Learning Techniques for Semantic Segmentation, Remote Sensing, 13, 2031, 2021.

Huete, A. R.: A soil-adjusted vegetation index (SAVI), Remote sensing of environment, 25, 295–309, 1988.

Ludwig, C., Hecht, R., Lautenbach, S., Schorcht, M., and Zipf, A.: Mapping Public Urban Green Spaces Based on OpenStreetMap and Sentinel-2 Imagery Using Belief Functions, ISPRS International Journal of Geo-Information, 10, 251, 2021.

Millard, K. and Richardson, M.: On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping, Remote sensing, 7, 8489–8515, 2015.

Montero, E., Wolvelaer, J. V., and Garzón, A.: The European urban atlas, in: Land use and land cover mapping in Europe, pp. 115–124, Springer, 2014.

Myneni, R. B., Hall, F. G., Sellers, P. J., and Marshak, A. L.: The interpretation of spectral vegetation indexes, IEEE Transactions on Geoscience and Remote Sensing, 33, 481–486, 1995.

Phan, T. N., Kuch, V., and Lehnert, L. W.: Land cover classification using Google Earth Engine and random forest classifier—The role of image composition, Remote Sensing, 12, 2411, 2020.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, 2020.

Reyes-Riveros, R., Altamirano, A., De la Barrera, F., Rozas-Vasquez, D., Vieli, L., and Meli, P.: Linking public urban green spaces and human well-being: A systematic review, Urban Forestry & Urban Greening, 61, 127 105, 2021.

Semeraro, T., Scarano, A., Buccolieri, R., Santino, A., and Aarrevaara, E.: Planning of urban green spaces: An ecological perspective on human benefits, Land, 10, 105, 2021.

Singh, R. K., Singh, P., Drews, M., Kumar, P., Singh, H., Gupta, A. K., Govil, H., Kaur, A., and Kumar, M.: A machine learning-based classification of LANDSAT images to map land use and land cover of India, Remote Sensing Applications: Society and Environment, 24, 100 624, 2021.

Sun, R. and Chen, L.: Effects of green space dynamics on urban heat islands: Mitigation and diversification, Ecosystem Services, 23, 38–46, 2017.

Talukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y.-A., Rahman, A., et al.: Land-use land-cover classification by machine learning classifiers for satellite observations—A review, Remote Sensing, 12, 1135, 2020.

Viña, A., Gitelson, A. A., Nguy-Robertson, A. L., and Peng, Y.: Comparison of different vegetation indices for the remote assessment of green leaf area index of crops, Remote Sensing of Environment, 115, 3468–3478, 2011.