AGILE: GIScience Series, 5, 8, 2024. https://doi.org/10.5194/agile-giss-5-8-2024
Proceedings of the 27th AGILE Conference on Geographic Information Science, 4–7 Sept. 2024.
Editors: Alison Heppenstall, Mingshu Wang, Urska Demsar, Rob Lemmens, and Jing Yao.
This contribution underwent peer review based on a full paper submission.

Knowledge-Based Identification of Urban Green Spaces: Allotments

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. Urban Green Spaces (UGS) play a crucial role in enhancing the quality of life in cities by providing numerous environmental, social, and health benefits. Among these green spaces, allotment gardens stand out as a unique type that contributes to ecological services, preservation of biodiversity, and the overall well-being of urban dwellers. Unfortunately, the significance of allotment gardens as a specific type of UGS is still disregarded and they are not recognized as a separate category in land use / land cover maps or city maps of green spaces. This is mainly due to the mixed use of allotment areas, their small size and absence of tailored identification or mapping workflows. In this research, we address the latter one by proposing an approach that utilizes various semantic characteristics of allotment gardens to create distinctive spatial representations. The semantic characteristics we consider include the presence, density, and height of garden huts, proximity to water bodies and railroads, as well as the presence of pathways within the allotment gardens. Allotments are delineated using a three-step procedure. This involves utilizing a Random Forest machine learning classifier to create maps of the distribution of green spaces, extracting garden huts employing a threshold, and demarcating the area using a density based clustering technique. Furthermore, we repeat the same workflow in a new study area to assess the applicability of the proposed workflow. With the established workflow, we are able to accurately identify 78% of allotments in Augsburg and 88% in Wuerzburg respectively. Our results demonstrate that the proposed workflow can be a useful approach to validate and extend existing land use and land cover data sets while remaining time and cost effective.

Keywords. urban green spaces, allotments, semantic classification, mapping green spaces

1 Motivation

With a growing urban population worldwide, it is essential that cities are planned and administered in an intelligent and environmentally responsible manner. Urban green spaces (UGS) play a fundamental role in enhancing the overall quality of urban environments and the well-being of their inhabitants. The presence of green spaces within cities has been associated with numerous human well-being benefits. For instance, it has been shown that cognitive recuperation, i.e., when the brain has a chance to rest and heal from cognitive fatigue, positively reflects on improving attention, focus, and overall mental functioning. According to the Attention Restoration Theory of Kaplan and Kaplan (1989), this restoration can be achieved by spending time in inherently fascinating environments. And typically, it is natural environments that provide these restorative opportunities (Kaplan, 1995).

Among the diverse range of UGSs in Europe, allotment gardens have emerged as a response to the rise of unemployment in post-industrialisation (Bell et al., 2016), turning them into primary food production areas in big cities. After 1980, the perception of these places starts to shift from places for solely personal food growth to an alternative use case of UGSs. However, it is possible to observe some differences in specific management practices in allotment gardens at all levels: national, city and municipal. Alongside small-scale agricultural production, parts of allotments can be covered by lawns, orchards and ornamental plants. Therefore, even if various national allotment acts define them as UGS (Dymek et al., 2021), in practice all the existing Land Use and Land Cover (LULC) maps categorize them as leisure activity spaces.

In the existing literature there are many examples of trying to understand effects of allotment gardens on human well-being, biodiversity or others. As such, findings of Wood et al. (2016) reveal that allotment gardeners have a significantly higher self-esteem, and experience less depression and fatigue. Furthermore, an increased frequency of visits to gardening areas is positively related with greater subjective happiness (Mourão et al., 2019). Beyond the benefits for well-being and food production, allotments heavily support biodiversity. Borysiak even suggested to consider allotments as biodiversity hot spots (Borysiak et al. (2017)), after recording 358 species of spontaneous flora in only 11 allotments in Poznań, Poland.

Accurate maps of gardens (private or communal) are crucial for carrying out a reliable analysis. However, in many European cities such maps either do not exist or are incomplete. Furthermore, different gardening types are usually managed by different authorities. For instance, allotment gardens are managed mostly by an "Association of Allotment Gardeners". Herb gardens mostly by the city, while rooftop gardens or balcony gardens are managed either by a "community of owners" or by the companies that own or rent the buildings. Due to the varied management authorities and a lack of consensus among them, there is currently no single differentiated map of urban gardens, neither are they explicitly delineated in typical land use maps. By contrast, the garden areas are commonly found as a part of the "sport and leisure activity" class in authoritative land use data. It is essential to note that this class also includes sport fields and other leisure places such as green corridors. Therefore, extracting only garden relevant information from the given class is nearly impossible. Moreover, there are not many studies that focus on workflows to identify gardening areas in cities, even though necessity of focusing to develop methods for mapping such small-scale UGSs is emphasized (Shahtahmassebi et al., 2021).

An attempt to map gardening areas is made e.g. by Mathieu et al. (2007). In their work, the authors implement an object oriented classification approach using Ikonos imagery to map private gardens in New Zealand. The imagery with a spatial resolution of four meters contains four spectral bands, advantageous to identifying greenery. The authors describe their method as a time saving approach to generate a data set relevant to urban gardens . Degerickx et al. (2020) also emphasizes the need of focusing on various types of UGSs and not only on e.g. large city parks. The authors utilize airborne APEX hyperspectral data, Worldview-2, and airborne LiDAR data in an object-oriented approach to map types of green spaces. Their findings show that the LiDAR data is a most promising data set while suggesting to use multi-temporal image analysis for a better understanding of green areas.

While at a very high resolution, multi or hyper-spectral remote sensing (RS) data combined with novel object-oriented, semantic segmentation methods might provide the most accurate results, the availability of computational power and freely available data sets remains a limiting factor. In this paper, we argue that each type of UGS possesses unique semantic characteristics that define its distinct 'face' or identity. It might be challenging to distinguish between various UGS types using only vegetation information derived from RS imagery, but it this approach is promising if we also utilize their semantic characteristics. A mapping approach using semantics was shown to work, for instance, in case of urban forest mapping (Ismayilova and Timpf, 2023). In this work we are not only interested in accurately mapping allotment gardens, but also in understanding to what extent semantic characteristics of allotments foster the identification process. In order to prove the robustness of the proposed method, we test it in two southern German cities with (slightly) differing topographic and landscape compositions.

In Section 2 we describe in detail our case study and the proposed workflow, while we present the results in Section 3. Section 4 provides discussions, conclusions and suggestions for future work.

2 Case Study

In this paper our main objective is to identify allotment gardens in our study areas by using their distinct characteristics as well as spatial features derived from these characteristics. Afterwards, test the proposed framework in a new study area in order to prove the transferability of the proposed workflow. Therefore, in the following, we describe common gardening types, selected study areas, present utilized data sources as well as the methodological framework we implemented.

2.1 Study Area

Augsburg, a city in southern Germany is the first study area where we build our methodological pipeline. Spanning an area of approximately 146 km2, Augsburg hosts a notable number of different urban gardening areas including, allotment gardens, community gardens, herb gardens and rooftop gardens. An example of an allotment garden can be seen in Fig. 1.

Figure 1. A typical allotment garden in Augsburg during summer.

Each of them exhibits distinct visual characteristics. These characteristics are apparent to the naked eye. Allotments, for instance, are typically composed of near-rectangular plots that are individually managed within a larger allotment agglomeration, allowing city dwellers to cultivate

their own selection of plants, vegetables, or flowers. Community gardens, on the other hand, present a more collective approach. They often appear as areas with a patchwork of vegetable beds, shared compost containers, and sometimes a few communal garden huts where equipment may be stored. These gardens not only provide space for growing produce, but also serve as centers for community interaction. Herb gardens, in contrast, are usually found within public areas like parks and are designed primarily for growing a variety of herbs. These gardens are not only functional, providing fresh herbs for culinary use, they are also educational, showcasing different herb species and their uses. Balcony or rooftop gardens commonly consist of single pots or vegetable beds on balconies or roof terraces. Despite these visual cues, translating such diverse semantic information into a basis for spatial analysis is challenging and might require a very detailed understanding of plant composition, location information and others.
The second study area we choose as a test site is the city of Wuerzburg. Located in the same German state as Augsburg, Wuerzburg covers an area of nearly 88 km2. From a landscape composition point of view, both cities share many similarities, such as the presence of large rivers, a similar vegetation cover and similar types of gardening areas. Yet, the hilly relief and the presence of permanent viticulture crops makes the appearance of both cities differentiable. Furthermore, Wuerzburg accommodates a number of pre-war established allotment areas that contain gardening huts that are larger and higher than commonly accepted ones. Additionally, some of the allotment plots are located on hilly slopes. Therefore, the existing differences could complicate the transferability of the proposed workflow.

2.2 Visual and Spatial Garden Semantics

We previously noted that different types of gardening areas display some similarities as well as some variations in their appearance and purpose. In this paper we focus solely on allotment gardens and establish their applicable semantic features.
When viewed from above, certain dominant features become prominent in both study areas: the presence of small, similarly sized huts in every private parcel within the whole agglomeration; the presence of a larger management building at the entry point to the main area; unpaved paths running out of the main entrance point and making every single plot accessible. The spatial location of the allotment gardens is also catching the eye: many of them are located alongside railroads or close to rivers. However, visually, these plots are comparably smaller or have a more elongated form in contrast to the allotment agglomerations that appear further away from rivers and railroads. The latter ones are larger and more rectangular or square in shape. Some of the described semantic features can be explained with the historical development of these areas as well as by the existing national allotment laws. As such, in the past, allotment plots were given to railway workers to grow produce. Therefore, they are frequently located close to the railroads. Currently, allotment plots can be rented by anyone and therefore, thus additional allotment agglomerations appear in the cities. For cities, allotments are a good way to make use of otherwise "unused" space. Furthermore, under the "small garden" law, allotment gardens should not exceed 400 square meters in size, with garden huts limited to a maximum area of 24 square meters. While there are no strict regulations on the height of garden huts, those built without official building permit (as is typical for garden huts) should not surpass 3.5 meters in height.
After carefully observing allotments in both cities, we establish the following semantic features that will be the building stones of the identification workflow for both cities:

- Hut presence: Every garden plot contains at least a hut or a hut with an extension.
- Hut height: Garden huts are not higher than 3.5 meters and thus can be distinguished from other built-up areas such as houses or buildings, and industrial buildings.
- Paths: The allotment agglomeration is crossed by a network of intersecting paths.
- Proximity: The majority of allotments are in a close proximity to railroads and water bodies, and are not crossed by major roads.

2.3 Data and Software Availability

Some of the allotment agglomerations consist of as few as 8 plots. Therefore, the resolution of the utilised imagery could be a limiting factor. In this paper, we utilize high resolution aerial imagery available to download from the State Office for Digitization, Broadband and Surveying web-page available here. The orthophoto of Augsburg was acquired on the 18th of June 2023, while of Wuerzburg on the 28th of May 2023. The orthophotos are three band RGB-imagery with 40cm spatial resolution. We pre-process this data by setting the right coordinate reference system as well as adjusting the correct extent.
In order to extract the height information of the huts, we calculate the height of objects above earth surface using the freely available digital elevation model (DEM) as well as a digital surface model (DSM). Commonly, above ground height of objects can be represented with the normalised digital surface model (nDSM). nDSM is a derivative elevation product, which we obtain by subtracting a DEM from a DSM.

$$nDSM = DSM - DEM \tag{1}$$

Both DEM and DSM data is acquired from the State Office for Digitization, Broadband and Surveying. DEM is a freely available raster data set with 1 meter resolution. Whereas, DSM is currently not free of charge but can be

bought at a resolution of 40 cm. In order to match the spatial resolution of the data sets, we first resample the DEM to 40cm and then calculate nDSM using the raster calculator expression in ArcGIS Pro version 3.1.1. Furthermore, in order to be able to establish height thresholds, we collect 200 hut points in both cities and calculate the height distribution.
Currently, some allotment data sets can be found in the freely available Open Street Map (OSM). However, these data sets neither accurate nor complete. As such, there are only 23 allotment plots in the city of Wuerzburg according to the official sources, while 88 appear in the OSM data set. Nevertheless, we use other layers of the OSM data sets to refine our results. In order to eliminate areas that we identified but are not allotment-specific features, we utilise land use, roads, railroads and water bodies polygons of OSM data sets available for both study areas. We create a buffer of 2 meters around linear objects to ensure a proper spatial selection afterwards. For further refinements we use the freely available building footprints shapefile from the State Office for Digitization, Broadband and Surveying.
We argue that allotment gardens show a considerable amount of green, and this information is lost when the existing LULC maps are taken into the analysis. Consequently, to confirm our argument, we perform an image classification and extract green space information. For this purpose, we collect 600 pure training points per city for binary image classification. Moreover, to validate the final allotment polygons created using the proposed workflow, we create an allotment validation data set. For this purpose, we use the allotment class within the OSM land use layer and remove or add allotment polygons that are included in the city allotment lists. Table 1 shows in detail the utilised data sets.

Table 1. Data sets utilised for the identification of allotment gardens.

Data set	Access	Type	Resolution
DOP	Free	Raster	40 cm
DEM	Free	Raster	1 m
DSM	Paid	Raster	40 cm
Building footprint	Free	Polygon	-
OSM	Free	Polygon	-
Training data	Free	Points	-
Validation data	Free	Polygon	-

Selected data sets are pre-processed and analysed in ArcGIS Pro version 3.1.1. Furthermore, we utilise RStudio version 4.3.1 to perform a Random Forest image classification. For reproducibility purposes we provide sample data sets, Model Builder file as well as a R code under the following DOI: 10.6084/m9.figshare.25683828.

2.4 Model Development

To achieve our objective, we employ a three-step allotment identification approach that combines classical geoinformatics techniques as well as state-of-the-art machine learning methods. In a first step, we perform a Random Forest (RF) binary classification to identify the extent of "greenery" in both study areas. We then perform height thresholding in order to identify garden huts, as these are key allotment indicators. At this stage, we utilise complementary data sets to refine the outcome of the thresholding. Finally, we perform area delineation techniques in order to form allotment objects. The detailed workflow of the analysis is shown in Fig. 2.

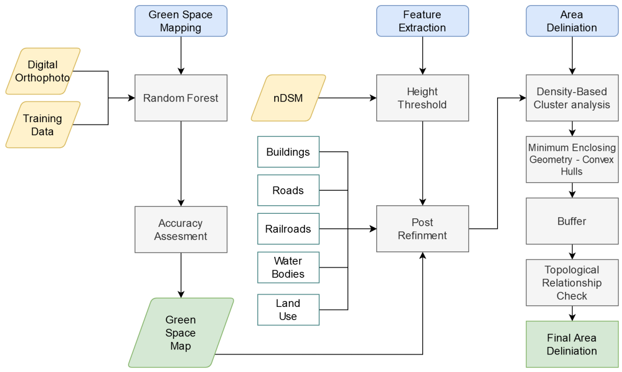

Figure 2. Flowchart of the proposed three-step allotment identification framework.

2.4.1 Green Space Mapping

We map the distribution of green spaces across two study areas with the help of a Random Forest (RF) classifier. Random Forest (RF), proposed by Breiman (Breiman, 2001), is an ensemble learning technique suitable for both classification and regression tasks. RF is chosen for its excellent performance with noisy data and its robustness to parameter initialization. During the classification process, RF employs bootstrap sampling to construct trees within the forest, with the final classification result derived from the majority votes of these trees. Parameter tuning is an integral part of optimizing results with RF, particularly through adjusting the "mtry" parameter, which dictates the number of variables considered for node splitting, and "ntree," which sets the number of trees to grow in the forest.
We construct the RF model using approximately 600 carefully selected training points, which are distributed as green (1) or non-green (0) areas. We ensure a balanced collection of training points for each class, with 300 points representing green and another 300 representing non-green samples. To train the model and assess its accuracy, we implement 10-fold cross-validation. During the process the training data set is split into 10 equal subsamples. Each sub-sample undergoes a sequence of ten iterations where it is used nine times for training and once

for testing. After the satisfactory classification results, we use the trained model to map green spaces within both study areas.

2.4.2 Feature extraction

The main features we intend to extract are garden huts. In order to understand the height distribution of huts, we manually collect around 200 hut centroids. We then extract height data from the nDSM data set for each collected point and compute descriptive statistics, including minimum and maximum values, mean, median, standard deviation, and the 5th and 95th percentiles. Percentiles are commonly used in statistics, because they give a sense of the spread of the data, and they are particularly helpful in identifying the range within which the majority of the data points lie, as well as spotting outliers. Consequently, we take the value of the 5th percentile as the lower threshold and the value of the 95th percentile as the highest threshold to extract the huts. Using the raster calculator, we select nDSM pixels that fall within the defined threshold
The potential drawback of the threshold approach is that it may include numerous objects or pixels within the study area having similar heights. To refine the results and filter out irrelevant objects, we utilize existing and freely available data sets. First, we transform the extracted height pixels into polygons. Subsequently, we mask out those portions of the "height" objects that overlap with green areas. Additionally, we incorporate land use data from the OSM dataset and exclude all objects overlapping with this dataset, ensuring beforehand that the land use dataset does not include the "allotment" class. We also eliminate objects intersecting with railroads, rivers, and roads, while specifically removing the path, footway, and pedestrian classes from the roads dataset. This strategy enables us to preserve objects potentially intersecting with paths commonly found within allotment gardens. Finally, we check whether any of the height objects coincide with building footprints and remove them if necessary.
The clipping and elimination phase significantly reduces the number of remaining objects. However, these manipulations also lead to the creation of some very small parts of larger objects that are unlikely to be garden huts. Therefore, as the final step, we select and remove all objects with an area smaller than 5 square meters. This size criterion is chosen after identifying the minimum size of huts that overlapped with the hut centroid dataset, which was previously used for height threshold extraction.

2.4.3 Area delineation

For area delineation, we concentrate on three main techniques: buffering, clustering, and creating minimum bounding geometries. We define the area of allotment gardens based solely on the presence of garden huts. Although front and backyard gardens or other leisure activity areas might also feature small huts similar to those in allotments, a key difference is that garden huts in allotment areas typically appear at a higher density and are clustered in one specific area. To understand the proximity of the huts to each other, we use the hut examples collected for height threshold estimation. We find that there is usually at least one neighboring hut within a distance of approximately 10 meters. Consequently, we create a 10-meter buffer around every extracted centroid. Following this, we eliminate buffer polygons with an area smaller than 315 square meters, which corresponds to the area of a single 10-meter buffer that does not intersect with any other buffer
To determine which of the remaining centroids are spatially clustered, we implement the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm (Ester et al., 1996). In UGS related studies, DBSAN is mainly used in combination with point cloud data sets. For instance, Li et al. (2022) use DBSCAN to distinguish non-ground points in LiDAR data set, which is later used to identify plant point clouds. Xu et al. (2022) on the other hand implement DBSCAN with point cloud data for tree skeleton extraction in a forest. This approach is particularly successful in detecting noise in the data set. Two parameters for clustering are pre-defined: search distance and minimum number of points in each cluster. Both in Augsburg and Wuerzburg we set the search distance to 10 meters, same as used in buffering step. Minimum number of points in each cluster is defined based on the available information of how many single plots are in each allotment agglomeration. The lowest number of plots in Wuerzburg is 11, whereas in Augsburg it is 8. Therefore, we adjust this parameter to 8 in Augsburg and 10 in Wuerzburg, accordingly.
To construct the outlines of the allotment gardens, we create minimum bounding geometries encompassing each garden hut cluster. Given that garden boundaries are typically irregular, we choose to create minimum enclosing convex hulls. The huts are often situated towards the inner sides of the garden area rather than the outer side. Although the convex hulls are formed by connecting the outermost points with straight line segments, we generate additional buffers around them to more accurately cover the garden areas.
After we identify garden areas we further check for topological relationships. Using spatial location based selection, we identify whether the garden areas are crossed by paths, footway and pedestrian paths from the OSM data set.

3 Results

In both study areas we employ an RF classification to accurately determine the quantity of vegetation present in each allotment agglomeration. Therefore we choose to perform 10-fold cross-validation to train and test the classifier. In Augsburg, the classification reaches 96 % accuracy. Furthermore, both true positives and true

negatives are identified equally well: 48.2 % and 47.8 % out of 50 % respectively. In total, RF identifies nearly 60 km2 of green space in the study area.
In Wuerzburg, the RF model yields an overall classification accuracy of 97 %. Further analysis shows that the identification of the true positive classes reaches 49.8 %, whereas true negatives is 47 %. The results indicate that the "green" class is marginally better identified than the "non-green" class. With the implemented classification approach we identify a total of 52 km2 of vegetation in Wuerzburg.

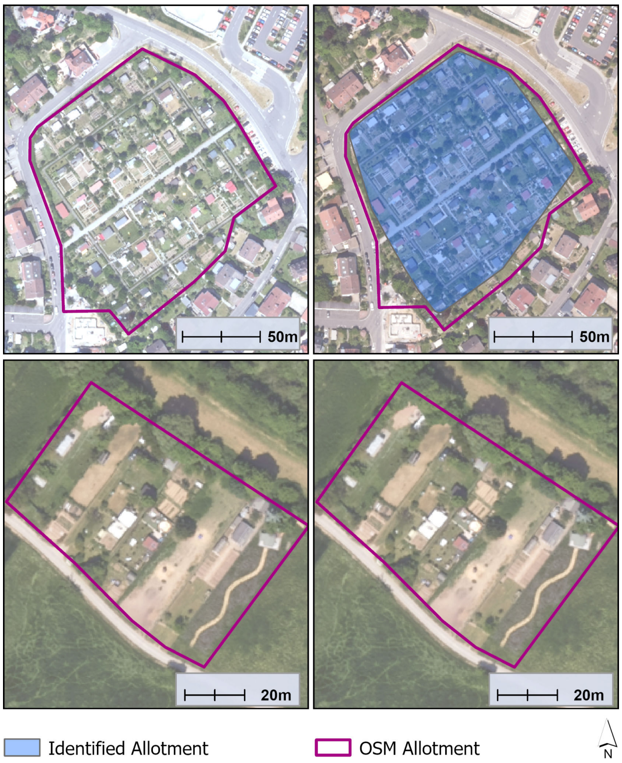

Figure 3. Figure showcasing a successfully identified (above) as well as an unsuccessfully identified (below) allotment agglomerations in the city of Wuerzburg.

In order to establish a height threshold, manually collected hut centroids are used. In Augsburg, the minimum and maximum height values correspond to 1.3 and 3.4 meters respectively. Moreover, the mean value corresponds to 2.39, the median value to 2.35 and the standard deviation to 0.33 meters. In order to include the majority of data points as well as to eliminate outliers, we calculate the 5th and 95th percentiles. Consequently, the minimum threshold equals to 1.9 and the maximum threshold equals to 3 meters.

The minimum and maximum height values extracted from 200 garden huts in Wuerzburg are 1.01 meters and 3.36 meters, respectively. Furthermore, the mean value is equal to 2.37, the median to 2.39, and the standard deviation to 0.41 meters. The 5th and the 95th percentiles

Figure 4. Figure showcasing successfully identified (above) as well as unsuccessfully identified (below) allotment agglomerations in the city of Augsburg

in Wuerzburg equal to 1.72 and 3 meters accordingly. Therefore we set the lowest height threshold to 1.72 and the highest height threshold to 3 meters.
When using the selected height threshold in Augsburg, we acquire around 682'000 objects. Consequently, we utilise additional data sets to limit the number of objects resembling actual hut numbers. With the help of spatial selection using green space, building, land use, roads, railroads, and river layers, we narrow down object numbers to nearly 5000 objects.
In Wuerzburg, the utilised threshold values result in around 650'000 objects. Here again, with the help of green space, building, land use, roads, railroads, and river layers we narrow down the object numbers to almost 1500 objects. Using the remaining extracted objects in both cities, we perform a DBSCAN clustering.
To delineate gardening areas, we enclose the clusters with convex hulls and then expand these hulls by applying a five-meter buffer. Utilizing this methodology, we successfully identify 180 allotments in Augsburg and 30 allotment clusters within the city of Wuerzburg. A comparison of the existing and identified allotment agglomerations in Wuerzburg is presented in Fig, 3, while those in Augsburg are shown in Fig. 4.
We validate our results using the enhanced OSM allotment dataset. Official records indicate that there are 52 allotment agglomerations in Augsburg and 23 in Wuerzburg. However, the OSM allotment dataset lists a total of 120 allotments for Augsburg and 34 for Wuerzburg. We

manually refine these datasets and use them to assess the accuracy of our identification process. Through the applied workflow, we successfully identify 78% of allotments in Augsburg. In Wuerzburg, our identification accuracy reaches 88%.
Furthermore, we calculate the amount of greenery contained within the allotment agglomerations, based on the RF classification dataset. In Augsburg, allotment gardens encompass 0.36 km2 of vegetation, which accounts for 0.6% of the city's total green coverage. In Wuerzburg, we determine that allotment agglomerations contain 0.19 km2 of greenery, representing 0.3% of the city's overall vegetation coverage.
Initially we presupposed that allotment agglomerations are crossed by a network of paths. Therefore, we perform a location based selection between identified areas and paths. Our findings confirm that the identified allotments are indeed crisscrossed by paths. However, we also note that the accuracy of the OSM path data is quite low, capturing only a limited number of paths within these privately-owned allotment agglomerations.

4 Discussion and Conclusions

In this study, we focus on an often overlooked category of UGSs, specifically allotment gardens. These areas are crucial in busy urban settings for both recreational purposes and food production. Contemporary mapping of UGSs predominantly employs Remote Sensing (RS) data, utilizing machine learning or deep learning techniques. These advanced methods, especially through semantic segmentation, achieve remarkable accuracy in identification, yet making it almost impossible to understand why and how each feature was selected. Therefore, our objective was to develop a framework that emphasizes a clear and straightforward decision-making process, ensuring transparency in the identification and selection of areas.
As we explore allotment agglomerations, we establish that garden huts are mostly clustered in one area, making them the primary identifier for allotments. Therefore, we collect sample training data sets and explore height statistics to define extraction thresholds. In order to include the majority of data variance and exclude the outliers, we select the 5th and the 95th percentiles as the minimum and maximum threshold values. Our results indicate that the highest threshold in both cities is nearly identical, approximately 3 meters, while the lowest threshold slightly varies, with values of 1.9 and 1.7 in Augsburg and Wuerzburg, respectively. This variation may result from the training data used to determine the height. But it is evident that the typical hut height ranges between 1.7 to 3 meters, with the maximum height being in line with existing regulations.
Urban environments are complex structures. While we do achieve impressive extraction of garden huts, with the set threshold, additional, non-hut structures are also extracted. Consequently, we use further data sets to refine our results. As such, we use roads, railroads, rivers, and buildings to eliminate unwanted objects. The main drawback of the chosen approach is that the amount of eliminated objects is directly linked to the accuracy of the selection data set. While OSM data sets include all the necessary information (rivers, roads, buildings etc.), the spatial overlap with actual locations sometime could be extremely poor. Furthermore, roads and rivers are provided as linear features, making it difficult to check for spatial relationships with the "hut" objects. Therefore, an extra step to convert them to polygons is necessary.
To identify hut clusters, we utilize a density-based clustering approach, specifically DBSCAN, chosen for its ability to handle noisy data. While this clustering effectively separates hut clusters within the study area, defining DBSCAN parameters is a manual and challenging process. Familiarity with the study area and manually collected training data help determine the appropriate search radius, while the minimum number of points in clusters relies on official allotment information. Furthermore, our results depict some over and under-identification rate in both cities. While looking at a greater detail, we notice that the allotment agglomeration exhibiting common allotment features are identified better than those with varying allotment composition. Varying composition includes allotment agglomerations that are very small in size, that are linear in form, with garden huts distributed linearly, or with only few allotment huts in the center of the larger area. Furthermore, some of the over-identified areas include camping areas near the lakes, where fixed camper houses are located. In a completely new test site, finding suitable model parameters will be challenging and the results may not be as favorable as in this study.
Additionally, we establish that using minimum enclosing convex hulls yields better results, closely resembling reality, than minimum enclosing bounding box or rectangle. However, the effectiveness of the computation of the convex hulls relies on the precision of the hut clusters, which, in turn, depends on the accuracy of the extracted garden huts. Therefore, we observe some convex hulls being too large or small due to the distribution of hut points.
The use of Random Forest (RF) for mapping green spaces has proven powerful in various studies. Thus, in this study, we employ a RF classification approach in combination with high-resolution aerial imagery. Our RF model achieves over 90 % classification accuracy in both study areas. We conduct a detailed visual examination of the classification results in allotment gardens and observe that the RF model struggles to distinguish between natural greenery, such as grass, trees, and crops, and artificial green areas, such as sports fields. Yet, vegetation throughout allotment gardens is adequately classified. Therefore, despite this limitation, we utilize our RF data sets for the final "greenness" calculations.
Another distinctive characteristic of allotment gardens is their intersection with crossing path network. While we examine this topological relationship, we once again encounter accuracy limitations of the OSM data set. Since

allotment agglomerations are not openly accessible areas, the digitization rate of paths in this areas is very low. However, we observe that path digitization rates within allotments is higher in Wuerzburg than in Augsburg.

The implemented workflow manages to better identify larger allotment agglomerations than smaller or linear ones. We observe that this results from the intermediate steps, where we remove height objects overlapping non-allotment land use. As such, we utilize the polygons with green spaces to erase parts of the height objects because huts cannot overlap green areas. However, the aerial imagery used to identify vegetated areas has been taken in summer. Therefore, the study area is well covered with vegetation and large tree canopies sometimes cover garden huts. This, in turn, results in the elimination of height objects during the process. If we were in a position to choose, we would prefer aerial imagery from the beginning of the year.

Furthermore, the demarcation of allotment boundaries show mismatches with original boundaries. This is because of the distribution of garden huts used to create convex hulls. Since we might eliminate more hut objects during the process than aimed at, convex hulls can appear larger or smaller than expected.

A significant part of our work involved establishing threshold values for the extraction of huts, an aspect central to our study. Our objective was not just to determine these thresholds but also to explore the feasibility of transferring the established values to a new study area. This transferability is essential for broadening the applicability of our findings beyond the initial study context. However, our experiments revealed a notable challenge: when applying the same workflow to different study areas, we observed slight variations in the resulting threshold values. This variance highlights the complexity of spatial analysis and the influence of unique geographic, environmental, and historic factors in each study area. It suggests that while our methodology is robust, the specific outcomes it yields are somewhat sensitive to the characteristics of the study area. Given this observation, a potential future work would be a 'blind test' – applying the threshold values derived from our current study to a completely new study area. Such a test would provide valuable insights into the transferability of our findings and the adaptability of our methodology. It would also help in understanding the extent to which our threshold values are influenced by the specific conditions of our initial study areas.

In conclusion, this paper explores the accuracy of identifying allotment gardens based on their distinct semantic characteristics. Two main findings of the analysis include: 1) Almost every large allotment agglomeration with a higher number of single gardening plots is identified better than the smaller, irregular allotments; 2) The outcome of the analysis might not be sufficient to create precise allotment maps in the study areas but might be a reliable source of information to validate existing data sets. We successfully identify 0.36 km2 of vegetation within allotments in Augsburg and 0.19 km2 in Wuerzburg. These findings support our hypothesis, that a considerable amount of greenery is lost if allotments do not appear as a type of UGS in land use and land cover maps. In future work we will concentrate on transferring the learnt knowledge to different study areas, possibly outside of Germany or even Europe.

References

Bell, S., Fox-Kämper, R., Keshavarz, N., Benson, M., Caputo, S., Noori, S., and Voigt, A.: Urban allotment gardens in Europe, Routledge, 2016.

Borysiak, J., Mizgajski, A., and Speak, A.: Floral biodiversity of allotment gardens and its contribution to urban green infrastructure, Urban ecosystems, 20, 323–335, 2017.

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

Degerickx, J., Hermy, M., and Somers, B.: Mapping functional urban green types using high resolution remote sensing data, Sustainability, 12, 2144, 2020.

Dymek, D., Wilkaniec, A., Bednorz, L., and Szczepańska, M.: Significance of allotment gardens in urban green space systems and their classification for spatial planning purposes: A case study of Poznań, Poland, Sustainability, 13, 11 044, 2021.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al.: A density-based algorithm for discovering clusters in large spatial databases with noise., in: kdd, vol. 96, pp. 226–231, 1996.

Ismayilova, I. and Timpf, S.: Semantic Identification of Urban Green Spaces: Forest, AGILE: GIScience Series, 4, 28, 2023.

Kaplan, R. and Kaplan, S.: The experience of nature: A psychological perspective, Cambridge university press, 1989.

Kaplan, S.: The restorative benefits of nature: Toward an integrative framework, Journal of environmental psychology, 15, 169–182, 1995.

Li, X.-X., Tang, L.-Y., Peng, W., Chen, J.-X., and Ma, X.: Estimation method of urban green space living vegetation volume based on backpack light detection and ranging, Ying Yong Sheng tai xue bao= The Journal of Applied Ecology, 33, 2777–2784, 2022.

Mathieu, R., Freeman, C., and Aryal, J.: Mapping private gardens in urban areas using object-oriented techniques and very high-resolution satellite imagery, Landscape and urban planning, 81, 179–192, 2007.

Mourão, I., Moreira, M., Almeida, T., and Brito, L. M.: Perceived changes in well-being and happiness with gardening in urban organic allotments in Portugal, International Journal of Sustainable Development & World Ecology, 26, 79–89, 2019.

Shahtahmassebi, A. R., Li, C., Fan, Y., Wu, Y., Gan, M., Wang, K., Malik, A., Blackburn, G. A., et al.: Remote sensing of urban green spaces: A review, Urban Forestry & Urban Greening, 57, 126 946, 2021.

Wood, C. J., Pretty, J., and Griffin, M.: A case–control study of the health and well-being benefits of allotment gardening, Journal of Public Health, 38, e336–e344, 2016.

Xu, Y., Hu, C., and Xie, Y.: An improved space colonization algorithm with DBSCAN clustering for a single tree skeleton extraction, International Journal of Remote Sensing, 43, 3692–3713, 2022.