Urban Planning
2025 • Volume 10 • Article 8562
https://doi.org/10.17645/up.8562

ARTICLE

Open Access Journal

A Machine Learning Approach to Adapt Local Land Use Planning to Climate Change

Julia Forster [1], Stefan Bindreiter [1], Birthe Uhlhorn [2], Verena Radinger-Peer [2], and Alexandra Jiricka-Pürrer [2]

[1] Institute of Spatial Planning, TU Wien, Austria
[2] Institute of Landscape Development, Recreation and Conservation Planning, University of Natural Resources and Life Sciences, Austria

Correspondence: Julia Forster (julia.forster@tuwien.ac.at)

Submitted: 30 April 2024 **Accepted:** 6 August 2024 **Published:** in press

Issue: This article is part of the issue "AI for and in Urban Planning" edited by Tong Wang (TU Delft) and Neil Yorke-Smith (TU Delft), fully open access at https://doi.org/10.17645/up.i388

Abstract

The impacts on living conditions and natural habitats deriving from planning decisions require complex analysis of cross-acting factors, which in turn require interdisciplinary data. At the municipal level, both data collection and the knowledge needed to interpret it are often lacking. Additionally, climate change and species extinction demand rapid and effective policies in order to preserve soil resources for future generations. Ex-ante evaluation of planning measures is insufficient owing to a lack of data and linear models capable of simulating the impacts of complex systemic relationships. Integrating machine learning (ML) into systemic planning increases awareness of impacts by providing decision-makers with predictive analysis and risk mitigation tools. ML can predict future scenarios beyond rigid linear models, identifying patterns, trends, and correlations within complex systems and depicting hidden relationships. This article focuses on a case study of single-family houses in Upper Austria, chosen for its transferability to other regions. It critically reflects on an ML approach, linking data on past and current planning regulations and decisions to the physical environment. We create an inventory of categories of areas with different features to inform nature-based solutions and backcasting planning decisions and build a training dataset for ML models. Our model predicts the effects of planning decisions on soil sealing. We discuss how ML can support local planning by providing area assessments in soil sealing within the case study. The article presents a working approach to planning and demonstrates that more data is needed to achieve well-founded planning statements.

Keywords

GIS analysis; machine learning; nature-based solutions; spatial analysis; spatial planning

1. Introduction

Municipal planning is crucial for climate adaptation and mitigation (e.g., Measham et al., 2011; Storbjörk, 2010). Local governments in Austria, through mayors and councils, use various informal and formal instruments for spatial planning. Spatial planning involves balancing diverse interests and strategic coordination with planning goals and policies at the regional, national, and international levels. One area receiving increased political attention is the preservation of soil as a resource in order to arrive at nature-based solutions (NBS) and mitigate climate change as well as adapt to it (Seddon et al., 2020). NBS are broadly defined as solutions to societal challenges that are inspired and supported by nature (European Commission, 2015). In contrast to many engineered solutions, NBS have the potential to address both climate mitigation and adaptation challenges at relatively low cost while delivering multiple additional benefits for people and nature (Seddon et al., 2020). While urban areas are often beneficiaries of NBS (e.g., urban greening), in smaller regions and municipalities the relevance of unsealed soil as a resource and the application of NBS are often overlooked. At the municipal level, there is often a lack of data and expertise for NBS in planning, resulting in shortcomings in addressing climate change and biodiversity loss locally, as pointed out by recent studies (Boehnke et al., 2023). There is therefore an urgent need to plan strategically for more efficient NBS use in settlement development and create transparency based on sufficient and good-quality data. Machine learning (ML) and automated analysis offer a potential solution. Spatial planning research is exploring artificial intelligence, including ML, to aid data collection, processing, and interpretation. A survey of the literature, however, reveals a gap: While ML is used for predicting land use changes based on time series or extracting imaging landscape elements and features, possible correlations and patterns with underlying planning regulations have not so far been analyzed.

A main aim of this study is to support the development of simulations for the review of planning regulations relevant to the preservation of soil as well as the implementation of NBS and the actual planning decisions taken based on a case study in Austria. The study targets smaller municipalities with limited capacities that often overlook broader implications. In order to develop a robust concept and database for impact analysis as a basis for NBS assessment within an existing regulatory framework, we pursue a methodological approach to building ML models towards the goal of making unsealed soil more visible as a resource and analyzing areas with NBS potential. The prediction of planning decisions on soil sealing is tested via three ML models (k-nearest neighbors [KNN], random forest, and support vector machines). Backcasting planning information and developing qualitative area categories for NBS provides the training dataset for the ML model. Our study provides insights into linking data on planning regulations and decisions to data on the physical environment in Upper Austria.

Following a literature review outlining the research gap (Section 2), the article details the five-step method design (Section 3) and implementation of Steps 1–3 with results (Section 4). The concluding discussion (Section 5) and outlook (Section 6) provide an interpretation of the results and offer directions for further research.

2. State of the Art and Conceptual Framework

2.1. The Role of Municipalities in the Preservation of Soil/Land for Climate Change Adaptation and NBS

At the local level, immediate and directly tangible climate change impacts converge with concrete opportunities for action and local knowledge about vulnerabilities (Radinger-Peer et al., 2015). Local governments are, on the one hand, confronted with multiple global crises (e.g., climate crisis, financial crisis, biodiversity crisis) plus the overarching pressure on land as a valuable resource for multiple purposes, and, on the other hand, with the day-to-day challenges in their place-specific local context (Haase et al., 2018). They often lack the ability to track the consequences of sealing resulting from their planning regulations and decisions and thus overlook the opportunities for climate change adaptation and mitigation as well as preservation of biodiversity (Raymond et al., 2017). Unsealed soil is a key resource for climate change adaptation. In order to strategically minimize climate change impacts, unsealed land is needed to strategically position and connect NBS (Seddon et al., 2020). Particularly in small- to medium-sized municipalities, soil sealing is happening with increasing velocity, counteracting climate change adaptation and mitigation. At the same time, small- and medium-sized municipalities and their role in climate change adaptation, mitigation, and biodiversity conservation remain surprisingly under-researched (Fünfgeld et al., 2023). Many planning processes lack systematic integration of NBS based on strategies for preservation of soil and strategic planning of adaptation measures.

One central aspect that hinders municipalities is limited (institutional) adaptive capacity, owing to limited resources, knowledge, and political will (Buschmann et al., 2022; Fünfgeld et al., 2023). Further barriers on the municipal level include the difficulty in understanding climate science (Fünfgeld, 2010), lack of staffing capacity (Bierbaum et al., 2013), and limited financial resources (Vringer et al., 2021). As a consequence, small- and medium-sized municipalities are characterized by much greater pragmatism in their planning decisions as opposed to strategic planning approaches (Bardt, 2018). Oijstaeijen et al. (2022) point out that smaller municipalities are more affected by knowledge gaps and, therefore, planning decisions that do take NBS and green infrastructure into account are almost solely cost-driven and fail to recognize the full range of benefits (Brokking et al., 2021). Unrecognized benefits of NBS include increased climate resilience, quality of life for inhabitants, as well as biodiversity (Pan et al., 2021). In summary, there is a significant gap in both practice and research on how small- and medium-sized municipalities can be supported in making informed decisions based on transparent data and how planning regulations and subsequent decisions affect opportunities to implement NBS for more resilient land use planning.

2.2. Artificial Intelligence and ML in Spatial Planning

Artificial intelligence and ML offer many possibilities for planning and decision-making, from test automation and simulation to uncovering patterns in data. ML algorithms, including random forest, support vector machines, and neural networks, are used in land use planning and have proven suitable for developing predictive models based on soil data and environmental variables (Chaturvedi & de Vries, 2021). These models are often trained on time series and time-lapse analyses of remote sensing data, segmented to identify/classify objects like roads, buildings, and vegetation (Dornaika et al., 2016; Shorter & Kasparis, 2009; Zhou & Chang, 2021), or classify urban functional areas (Chen et al., 2021). Automatic (Karila et al.,

2023) and manual labeling (González-Collazo et al., 2023) of remote sensing data provides the ground truth for quantitative evaluation of the trained model. If available, it would still be preferable to use already vectorized data with attribute information. ML methods can serve as a fallback option to fill gaps in the vectorized data needed for our project. In Austria and EU countries, land use plans are already available in digital formats, e.g., via the INSPIRE interface (Directive of 14 March 2007, 2007). However, there is no generalized, comprehensive information available for zoning plans and building regulations. For landscape features, vectorized data is partially available through OpenStreetMap, and open government data (OGD) and at small scales via INSPIRE, such as CORINE Land Cover (Büttner et al., 2004). Higher resolution data in the form of point clouds (e.g., LiDAR or terrestrial laser scans) is available but complex to analyze due to the need for attribution and feature extraction.

Nagappan and Daud (2021) show that while ML models are used to identify land-use or land-cover patterns from various data/time series, possible correlations with planning regulations are not analyzed. In their 2022 study, Takouabou et al. (2022) argue that static models are inadequate for understanding the complex urban dynamics of contemporary cities. They proposed that ML algorithms could offer a more effective approach to data processing, facilitating the reorganization of urban planning processes. This is also true for rural areas. The pressing need for planning support for accurate estimation of decision outcomes and interactions is essential for robust and sustainable development in rural areas. A review of the existing research reveals a lack of case studies and practical ML applications in spatial planning, soil quality and land use analysis, and natural structures serving as NBS. Nor do they establish a link to the underlying planning regulations.

Analyzing open spaces in connection with underlying planning regulations is the innovative part of creating the datasets. This article critically discusses the opportunities such a model offers and the need to embed it in additional methodological approaches. As an initial conceptual study, we discuss the feasibility of ML application in spatial planning at the community level, using open data to train the model. We also identify the conditions necessary for NBS to foster climate change adaptation and biodiversity preservation and examine the likely impact of an ML-supported model in planning practice.

3. Case Study and Methodology

The article focuses on a case study of single-family houses in Hörsching, Upper Austria, chosen for its transferability to other regions. It covers 10 ha, chosen because single-family homes constitute 64% of residential buildings in Austria (Statistik Austria, 2023). This dominance is notable in the outskirts of large cities like Vienna, Graz, and Linz.

Figure 1 shows the settlement structure and key spatial regulations of the zoning plans for the use case area, provided by representatives of Hörsching. The settlement database represents planning regulations for different development phases (from west to east and from the 1980s to the present). Additionally, municipal strategic planning documents were examined, including the local development concept and sectoral development concepts for traffic, building land, and green space.

This section describes the classification process and the development of the ML training data and model. The MLbase4NBS method (see Figure 2) presented in this article is part of our five-step concept (ML4Nature) for the automated assessment of NBS potentials. To quantitatively assess NBS potentials on

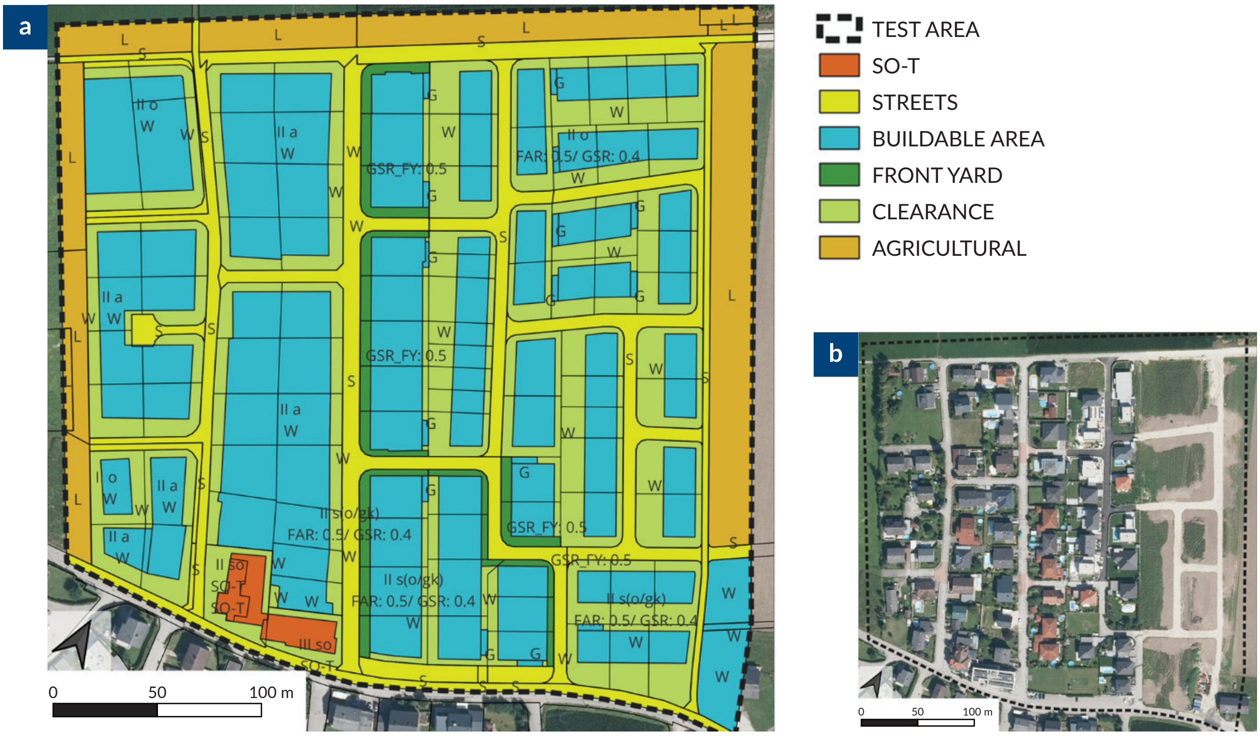

Figure 1. Use case area for the presented method including (a) zoning plan and (b) aerial view. Source: Adapted from the digital cadastral map (BEV, 2023) and map data (Geoland, 2024).

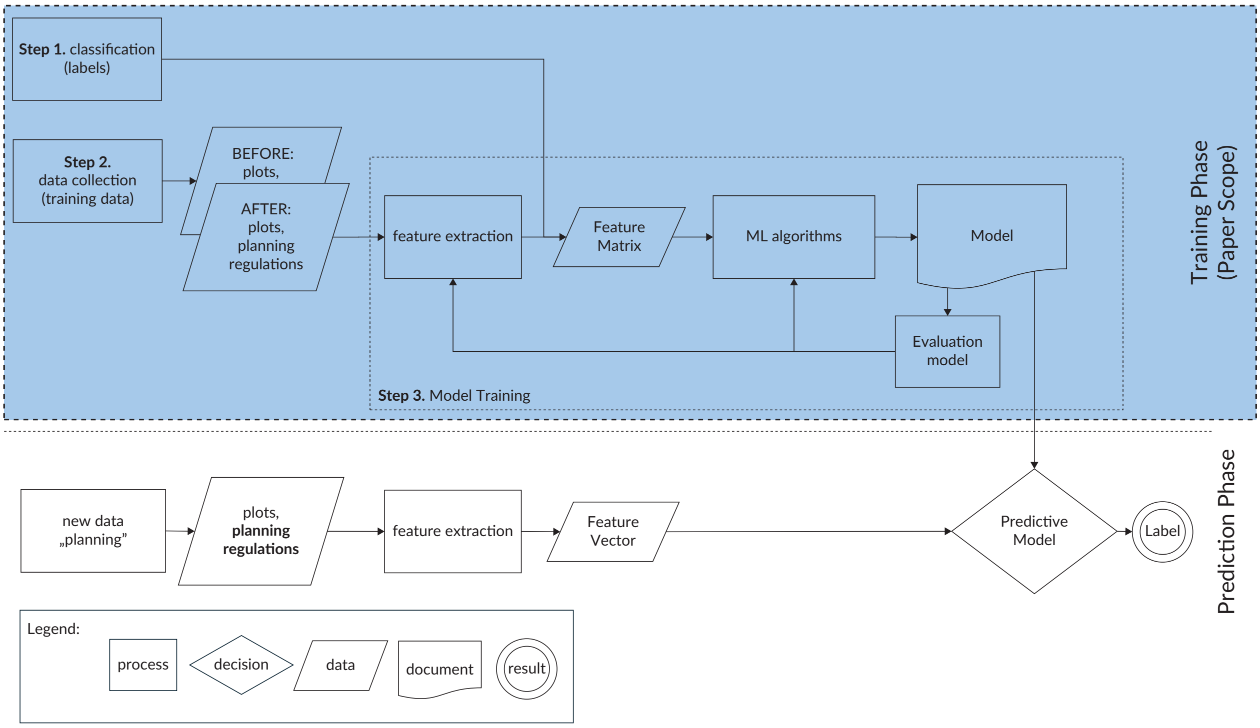

Figure 2. MLbase4NBS method: Training Phase (Steps 1–3, paper scope) and Prediction Phase.

properties, the analysis divided and classified properties into qualitative subareas (Steps 1 and 2). The ML model was then trained on these subareas along with building and planning regulations (Step 3). This model predicts qualities of subareas on new plots based on proposed planning regulations (Step 4), enabling quantification of NBS potentials (Step 5). The cumulative potentials are derived from the expected effects of the subareas, considering their features, size, and location (see data structure in Figure 3).

3.1. Step 1: Elaborating the Classification Model

By analyzing past and current planning regulations of the research area, past planning decisions can be processed for detailed backcasting and measure evaluation. The backcasting information is processed as a learning/training dataset for future decisions, measures, and actions. This process builds the inventory of "spatial" NBS, based on orthophoto mapping and site visits to derive different area category attributions for computer-aided processing.

Initially, all categories of different areas on a parcel must be defined and elaborated, considering their impacts on environment, habitat, and climate. These areas are categorized according to criteria such as biodiversity indices, evaporation and infiltration potential, or sealing. The classification model is intended for use in an automated evaluation process for NBS.

Figure 3 outlines the process: Specific quantitative and qualitative properties were collected for different areas. These include general information on location, geometry, and formal use, such as the natural green space characteristics, which were distinguished into ML4Nature (surface) categories and ML4Nature (areas) equipment (see Table 1).

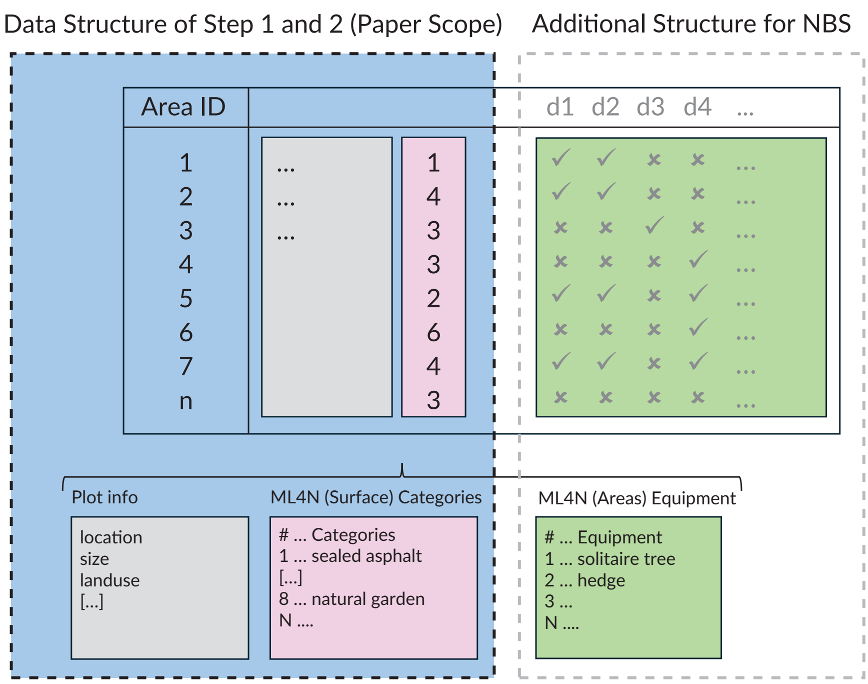

Figure 3. Data structure concept: surface-related properties (paper scope) and NBS equipment of areas.

Table 1. ML4Nature surface categories within the settlement structure.

ML4Nature_CAT_ID	ML4Nature Surface Category	Drachenfels Category
1	Sealed asphalt, dark concrete	
2	Sealed asphalt, light-coloured concrete	
3	Sealed water pool	
5	Fruit and vegetable garden	12.6.2
6	Home garden with large trees	12.6.3
7	Modern ornamental garden	12.6.4
8	Natural garden	12.6.5
9	Heterogeneous home garden area	12.6.6
[...]	[...]	[...]
60	Agricultural area, field	
99	Buildings	

The ML4Nature categories were based on Drachenfels (2021) after screening several categorization schemes. They refer to structures common in small- to medium-scale municipal settlement areas such as the sub-group "vegetation-determined biotope complexes and types of use in green spaces." They were supplemented by additional areas (sealed surfaces, buildings) not defined as "green spaces," which were preeminent in the first analysis presented in this article as they are a pre-requisite of many NBS to climate change adaptation next to structures attached to buildings themselves. ML4Nature equipment holds key features of green spaces (based on Drachenfels, 2021) with a minimum set of facilities by area (see Figure 2).

General plot information supplemented by Categories and Equipment form the basis for the automated and rule-based evaluation of open spaces. They constitute the starting point for NBS analysis: The NBS potentials (potentials for cooling, evaporation, infiltration) are defined for the different area categories. These depend on the vegetation and soil characteristics of the area.

3.2. Step 2: Development of Training Dataset

To predict the distribution of areas on parcels based on planning regulations with ML, a geospatial training dataset was developed, containing digital planning regulations categorized for impact analysis. Pre-trained models for roads, buildings, and swimming pools can assist in the preparation and pre-classification of spatial features within the dataset. Data can be sourced from established databases, GIS, and OGD platforms such as:

- National cadastre-info (https://data.bev.gv.at) on land;
- Federal GIS DORIS (https://www.doris.at);
- National open data (https://data.gv.at);
- European initiatives like Copernicus (https://www.copernicus.eu/en);
- Various GitHub repositories (https://github.com/zhouenbo/awesome-satellite-imagery-datasets).

ML algorithms can be created and tested using Python-based open-source tools like Tensorflow (https://www.tensorflow.org) and Pytorch (https://pytorch.org), which offer pre-trained models for image segmentation and object detection, applicable to aerial image analysis and land cover classification.

The preparation of our datasets included the preparation of GIS data at parcel level, as well as the extraction and digitization of building regulations. The resulting dataset contained information on past and existing land use and land cover distributions, and the planning regulations that influence them. Additional planning information and parcel-specific details (terrain, height, orientation, geometry/form) were then added.

Establishing the ground truth for the training data involves various methods of collection, calculation, and classification of areal features per building plot. Geoprocessing tools based on OGD, pre-trained models for certain features (e.g., buildings, water areas, streets), and manual identification and labeling of open and green spaces according to defined quality criteria and categories are used for pre-processing and classification.

3.3. Step 3: Model Training

Finding an optimal model that neither overfits nor underfits requires an iterative approach. This is why a supervised learning model is preferable to an unsupervised model, as it can represent individual decisions more clearly. The quality of an ML project depends on three important factors: data collection, data pre-processing, and data labeling. Accurate labeling is essential for robust results.

The dataset (split into training and validation data) is used to train an ML model. The dataset allows qualifying which regression procedure or decision tree/forest fits best. The ground truth enables quantitative evaluation of the tested models and selection of the best fitting model according to evaluation criteria. The trained model can forecast future scenarios based on planning regulation settings, estimating the distribution of areas on parcels and climate impact based on the areal share.

When selecting a model for a relatively small dataset with less than 1,000 samples and no linear relationship between features, models that handle non-linear relationships were considered. KNN is intuitive and effective for simple, small datasets. Random forest, based on multiple decision trees, prevents overfitting and handles non-linear relationships well. Support vector regression (SVR) uses kernel tricks to model complex relationships but requires careful parameter tuning. Within this methodological approach, the ML model was selected based on cross-validation to determine the best fit for the data's characteristics.

The selection of ML models for testing this use case was based on the following considerations:

- KNN: The idea behind KNN is that similar data points in the data space will have similar target values, assuming that soils and areas/points in similar regions have similar characteristics. The basic premise of the model is that where most features are similar, a similar result is likely to occur.
- Random forest: This model is robust to overfitting and can handle complex non-linear relationships between input features and target values. In our application, the relationships between input values are non-linear, making random forest an appropriate choice. In general, both KNN and random forest are good models for prediction scenarios where the relationships between values are unclear or how parameters affect outcomes is not well defined.
- SVR: Particularly suitable when the data are not linearly separable or the relationships are non-linear. SVR is robust to outliers, making it a strong candidate for dealing with diverse and complex datasets where standard linear models fail.

Our study empirically tested the prediction of area categories (classification problem) and the rate of sealed area on plots (regression problem). Utilizing planning data, cadastral data, aerial photographs, and on-site inspections, an evaluation with these two target variables appears feasible and realistic.

3.4. Step 4: Prediction Phase, and Step 5: Integration and Consideration of Structures Serving as NBS

As shown in the lower part of Figure 2, the prediction (Step 4) of area properties and the area category forms the basis of the automated investigation of NBS potentials for future building developments based on existing zoning regulations. The specific implementation of this link (Step 5) was not part of this study. For example, the degree of sealing of properties does not yet allow specific conclusions to be drawn about the infiltration capacity of areas. The challenge of assessing the climate impact of open spaces thus remains unresolved in this article. By collecting the spatial and green space defining characteristics of the properties, a future assessment of the NBS potential can be implemented based on quantitative data and on prediction of the area category through a set of statistical rules (depending on size, location, and presence) together with a larger dataset. The accuracy of the data does not allow for property-specific assessments, but it does allow for the estimation of potential within ranges for different planning alternatives in advance.

4. ML Implementation and Results (Steps 1–3)

This section describes the execution and implementation of the MLbase4NBS method illustrated in Section 3 (Steps 1–3) and the results of the iterative process of model training (see Section 4.2).

4.1. Set-Up of a Training Dataset and Related Opportunities and Limitations (Steps 1 and 2)

4.1.1. Data Sources and Quality

The base data for analysis are development and zoning plans, which were georeferenced manually. As GIS base, orthophotos of varying recency and sources (Geoland, 2024; Google Maps) are used. Further, cadastral shape data from the Federal Office of Metrology and Surveying (BEV, 2023) are added. Among key spatial information in the plans, the following points were homogeneously regulated for the entire study area:

- Green space percentage on the properties (min. 40%);
- Floor area ratio (0.5);
- Max. 2 floors (in the special designation for tourism max. 3 floors);
- Open or semi-detached construction within the polygons declared as building areas;
- Building setback of 3–5 meters from the street;
- Areas specifically designated as front gardens with separate minimum greening requirements.

Furthermore, regulations in textual parts of the zoning plans necessary for NBS potentials of the areas are included:

- The mandatory greening of flat roofs > 50 m² (although there are numerous exceptions);
- The requirement for car parking in designated areas to be grassed rather than paved;
- Special regulations for enclosures;

- The requirement to infiltrate stormwater on the properties themselves.

For the case study, the different areas were manually assigned to the ML4Nature categories by subdividing the cadastral base areas (see Figure 4a). Next, eight zoning plan documents were digitally uniformized (see Figure 4b). Via intersection, different subareas were created; they form the basis for the training dataset (Figure 4c).

4.1.2. Processing of Planning Regulations and Their Implications on the Existence of NBS

The challenges of combining textual and geometric planning regulations (Sections 3 and 4.1) were due to the difficulty of extracting common regulations from regulation plans and translating textual planning regulations to parcels. Therefore, the textual regulations and the regulatory plans were processed in parallel. By geoprocessing and assigning certain areas from the plan to the parcels, the identified findings were attached to the individual objects via attributes. A triangulation of the following methods was used to prepare the base data for supervised ML and a practicable categorization of open spaces: GIS analysis, systematic literature review, historical document analysis, orthophoto mapping.

The 1,404 different subareas were assigned to 10 different ML4Nature categories (see Table 1) and ML4Nature equipment features were examined for the areas. The equipment has not yet been included in the ML questions tested here. Overall, more than 47% of the total case study area is sealed, either by buildings, roads, or other sealed surfaces.

4.1.3. Training Data (Limitations and Opportunities)

National and federal OGD data sources (attributed vector data, orthophotos) were used to carry out the method on this use case and were updated and supplemented by visiting the site. Here it also became apparent that an attempt is being made to answer two specific questions based on the small sample: a classification problem and a regression problem.

4.1.3.1. Classification Problem

Figure 4 shows that the combined information on open space and development regulations form the attributes for the classification problem. Table 2 shows the 24+2 attributes identified in the model.

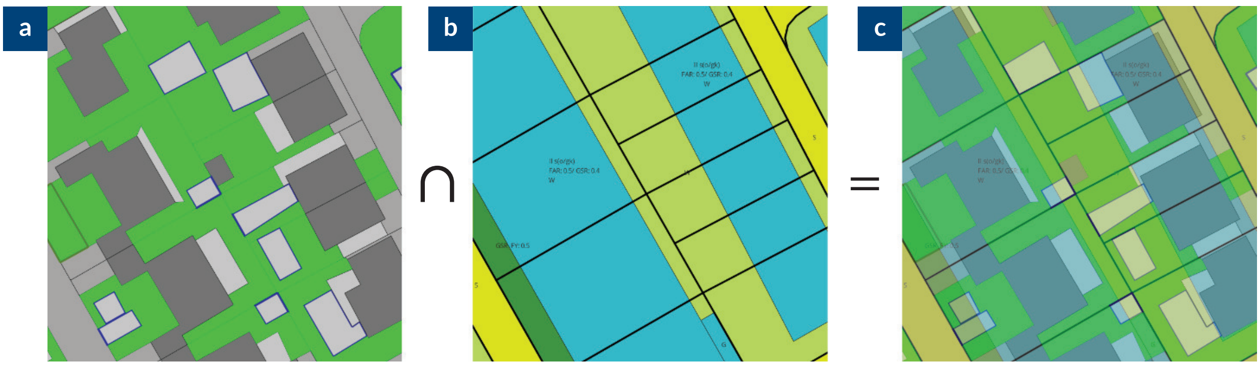

Figure 4. Intersection of ML4Nature categories with development plan as base for the training dataset.

Table 2. Overview of tested ML problems.

ML Problem	# Features	Attributes	Target Value
Classification	$n = 1{,}404$	26 attributes: size of plot, perimeter area ratio ($Z = A \times 100 / P^2$), [Previous State:] (overall) use category from Austrian land register, use categories as list from Austrian land register, share of building area on plot, share of sealed area on plot, land use/zoning category, front yard (t/f), sealed surface (t/f), [Development Plan:] number of floors on plot, floor area ratio, building eaves height on parcel, building ridge height on parcel, special regulations for front yard (t/f), special regulation for building clearance/side setback (t/f), dedicated land use/zoning category for plot parts (e.g., outbuildings), number of floors on plot part, special building arrangement for plot part, min. green space ratio on plot, floor area ratio for plot part, min. green space ratio for front yard plot parts, [Location:] point x, point y	ML4Nature-CATEGORY OF PLOT PART
Regression	$n = 132$	15 attributes: size of plot, perimeter area ratio ($Z = A \times 100 / P^2$), [Previous State:] (overall) use category from Austrian land register, use categories as list from Austrian land register, share of building area on plot, share of sealed area on plot, land use/zoning category on plot, overall dedicated land use/zoning category for plot, [Development Plan:] building arrangement (e.g., open, closed, or semi-detached construction) for whole plot, number of floors on plot, floor area ratio, building eaves height on plot, building ridge height on plot, share of dedicated building clearance areas on plot, share of dedicated front yard areas on plot	RATE OF SEALED AREA ON PLOT

In addition to the general property attributes and the attributes describing the previous state of the areas, there are attributes containing development regulations. Attributes 25 and 26 (as X and Y coordinates for a point grid) are included because the dataset is subsequently artificially enlarged (see Figure 5b).

Table 2 (row 1) attributes show which open space ML4Nature (surface) category (Table 1) is predicted for an area by the classification model.

4.1.3.2. Regression Problem

In order to predict the proportion of sealed area on a plot, the plot level attributes were collected. Table 2 (row 2) shows the 15 attributes used for the prediction.

Figure 5 shows the current rate of building on plots as an analysis step during data collection (Figure 5a) and the rate of sealed areas on plots (Figure 5b). The rate of sealed area on plot (Table 2, row 2) is the target for the regression model. The model predicts the rate of sealed areas per plot based on the plot data collected. Table 3 shows the attributes selected for the regression model. The training datasets are small. A total of 1,404 different subareas were identified on 132 plots.

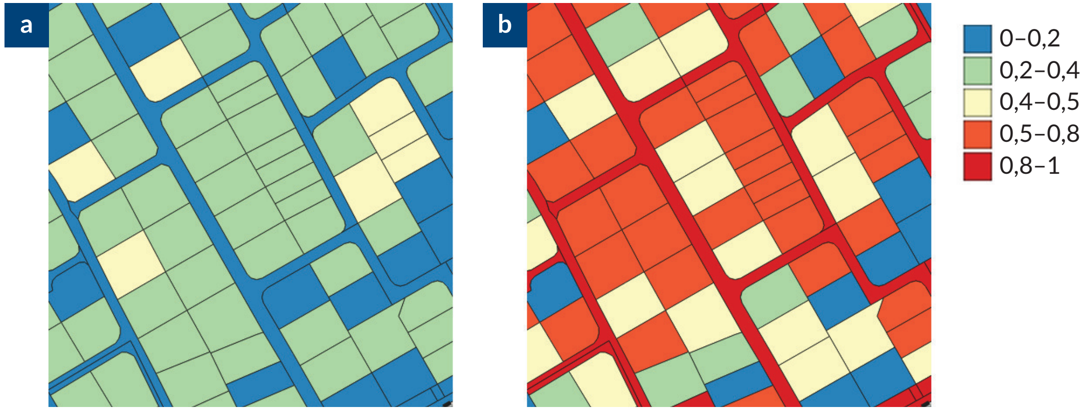

Figure 5. Rates of (a) building areas on plots and (b) sealed area on plot.

Table 3. Accuracy classification of ML4Nature category.

	Polygons (*N* = 1,404, test-size 30%, 24 features)		Point Grid (*N* = 24,703, test-size 30%, 24 features)		Point Grid Without Streets (*N* = 20,682, test-size 30%, 24 features)	
Classifier Model	Accuracy	Parameter	Accuracy	Parameter	Accuracy	Parameter
KNN	0.38388625592417064	5 neighbors	0.9666711644852247	3 neighbors	0.9618049959709911	3 neighbors
Random Forest	0.3818683026279804	n_estimators: 50 max_depth: None CV = 3	0.9830557483229239	n_estimators: 200 max_depth: None CV = 3	0.9825242372382754	n_estimators: 200 max_depth: 100 CV = 3

4.2. ML Model Training (Step 3) and Results

4.2.1. Prediction of Category (Classification Problem)

In the following, the selected models for classification (prediction of area category on subareas of features) are presented and the results are compared. Table 3 shows the accuracy classification of the ML4Nature category of polygons, point grid, and an alternative of point grid analysis with removed road surfaces.

The polygon-based analysis shows that the accuracy of KNN and random forest are approximately equally poor. The accuracy ranges from 0 to 1, where 1 would indicate overfitting of the model. With 10 possible categories in the training dataset, a rate of 38% is better than chance (10%), but still far from an accurate prediction. This is most likely due to the small training dataset.

To check if the prediction works better with a larger dataset, a point grid (2m) was placed over the polygons, with each individual point receiving the information of the area on which it lies. This immediately gave us a test dataset of over 24,000 records (see Figure 6b). Using the point-grid dataset, the accuracy of both models jumped to almost 1, indicating strong overfitting. This is because the algorithm is too well trained on the training set, which is certainly also due to the small number of different initial values, resulting in too little noise in the model.

In particular, road surfaces (which make up about 20% of the area of the test site) can probably be predicted very well on the basis of the input variables. When road surfaces are completely removed from the training dataset, as expected, the accuracy drops slightly, but this suggests that the input variables were not badly chosen and that the model could work quite well with a larger sample.

Overall, the small datasets are challenging and need to be extended to larger samples with higher variance. Strategies for automatically increasing the variance of the data or artificially enriching and enlarging the test data are discussed in Section 5.1.

4.2.2. Prediction of Sealed Area (Regression Problem)

For the prediction of sealed areas, KNN and random forest were used. Additionally, the results of the SVR were examined. The results are compared in Table 4. Streets with high accuracy were removed from the training dataset.

Test A, a test run with a very small plot dataset ($N = 132$), shows that the accuracy value of KNN cannot be trusted. It jumps significantly with each new calculation. SVR and random forest have a similar range of values, which could indicate a lack of quality in the dataset, most likely only in its size.

Therefore, the extension via a point grid of 2m, as in the classification problem, was chosen to increase the dataset. The enlarged training dataset ($N = 24{,}709$) leads to an accuracy of 1 in all three models, due to massive overfitting of the larger but still simple dataset of low complexity.

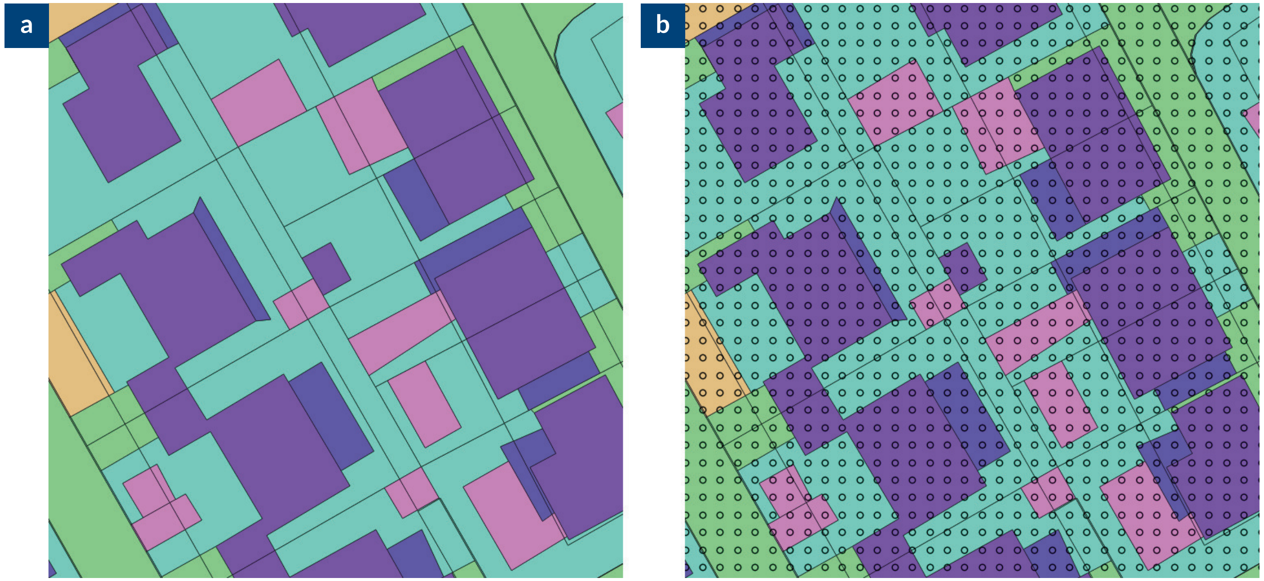

Figure 6. Extending the training dataset by using a point grid (b) instead of the polygon layer (a).

Table 4. Accuracy of the regression models: Test comparison.

	Test A (N = 132, test-size 30%, 15 features)		Test B (N = 495, test-size 30%, 15 features)		Test C (N = 495, test-size 30%, 15 features)	
Regression Model	Accuracy	Parameter	Accuracy	Parameter	Accuracy	Parameter
KNN	0.541316	5 neighbors	0.869972	5 neighbors	0.810307	5 neighbors
Random Forest	0.3423376622194893	n_estimators: 50 max_depth: 5 CV = 5	0.9072332695334303	n_estimators: 100 max_depth: 15 CV = 3	0.8612529237675896	n_estimators: 150 max_depth: None CV = 3
SVR	0.35775595006039373	C = 1.0, epsilon = 0.2	0.7140296158996142	C = 1.0, epsilon = 0.2	0.744826052468326	C = 1.0, epsilon = 0.2

Test B in Table 4 shows that reducing the number of points ($N = 495$) by considering only the 50th largest point leads to more variation in the dataset and more reasonable results.

To further improve the training dataset, the variance of the examples was increased by adding noise to certain features by randomly increasing and decreasing the feature size by +/– 5%. The ratio describing the relationship between area and feature perimeter was also randomized by +/– 5%. Noise was also added to the target variable (SEAL_PERC). Test C shows the expected changes in the accuracy of the models.

5. Discussion

This article has presented a conceptual approach to classifying subareas and predicting soil sealing on areal plots. The methodological approach provides a planning tool for estimating categorization probabilities based on planning rules. With more features in the training dataset and additional variables (e.g., NBS equipment of the areas), target variables can be predicted in detail, and linking regulations with implications on soil sealing as well as NBS is possible.

5.1. Feasibility and Limitations in the Set-Up and Application of the ML4Nature Concept

There is potential bias in data and parameter settings during training: Selecting hyperparameters based on best-fit criteria is tempting but can lead to overfitting. Consistency in predictions must be ensured for small data samples. The validity of cross-validation depends on the quality of the subsets, which small samples cannot guarantee. Extending test datasets reduces overfitting tendencies, indicating that the proposed method can provide conclusive results. The ML model predicts area distributions of previously classified areas.

There are challenges owing to incomplete data: While all current planning data was available for our small use case, evolutionary steps and initial data were not fully available. On a larger scale, data is likely to be incomplete. Missing data can be substituted with assumptions for initial testing, but this can lead to distortions. Techniques like iterative imputation estimate missing values based on existing data, reducing the impact of gaps. Leveraging local expertise and historical land maps can provide preliminary insights and help make educated guesses about missing data. This strategy addresses immediate challenges and sets a pathway for continuous improvement, enhancing model resilience and effectiveness in regional planning and development assessments. Our test results show that prediction accuracy improves by inflating and varying datasets.

The most effort-intensive part of the process lies in data collection. Algorithms capable of classifying orthophotos can automate some analytical processes. Attention must be paid to uniform spatial/temporal resolution of data sources (e.g., orthophotos from similar flight times should be analyzed together to minimize bias). While missing data could be added manually in this article, further research is needed on the applicability of automated algorithms to enrich training data and expand samples.

There are issues related to implementation and scalability: For practical implementation, training data should be collected and compiled at the federal level. Training of the models should be organized jointly for regions with homogeneous spatial and settlement structures. Trained models will then be used by municipalities for prediction, since single municipalities cannot manage this by themselves. Transferring models across Austria will require generalization and homogenization of data, as planning regulations differ between federal states (land use categories, building regulations, etc.). Natural and green areas may also differ between eastern and Alpine regions in Austria but can be generalized based on European standards (e.g., implemented in INSPIRE and SENTINEL data). For urban areas, models are likely more transferable but depend on regional building regulations.

5.2. Outreach and Expected Impact of the ML4Nature Model for Planning

Adaptation to climate change and the development of NBS are strongly linked to the planning decisions and actions of local policymakers because of their responsibilities in land use regulation and land use planning, as is the case in Austria. Transparency on the impacts of soil sealing and the opportunities for the application of NBS could lead to more comprehensive planning and decision-making processes. A variety of studies confirm that both the diversity and structural variety of green spaces significantly influence biodiversity, emphasizing the importance of conservation in both large areas and small green elements along streets or in backyards (Fuller & Gaston, 2009; Matthies et al., 2017)—the availability of space and unsealed soil is a prerequisite for their implementation. The ML approach outlined in this article shows the need for larger training dataset developments to create transparency not only on sealing and availability of open space but also on individual structures such as old trees and hedges, which provide crucial habitats for various species and contribute to biodiversity based on their physical characteristics and the surrounding environment (Gosling et al., 2016). This will also allow reflection on vegetation along traffic structures, which plays a key role for both conservation in settlement areas (Helldin et al., 2015; Thomas et al., 2003) and climate change adaptation (Morakinyo et al., 2020). Only then is the impact of planning regulations more specifically traceable. The model has potential to reflect specific capacities for both biodiversity conservation and climate change adaptation via more detailed recognition of the natural structures. Reducing heat through shading and transpiration, for instance, varies by tree species and interrelationships among environmental conditions (Zölch et al., 2019). Factors such as the width of hedges, an important factor in ecological value, would require additional development of the model. In addition to biological factors, such as tree size or the leaf area index, microclimatic aspects, such as radiation, wind direction, or speed and soil conditions, e.g., soil moisture and temperature, influence the transpiration performance of trees and vary depending on the spatial structure (Offerle et al., 2007). Integrating this more specific information to connect to climate change adaptation more precisely will be part of future studies in cooperation with climatologists and biologists. An actor-centered approach, including methods such as interviews and stakeholder workshops, could further support reflection on the role of data and transparency in planning decisions and the potentials and implications of an ML application.

The proposed approach (including data acquisition, training data, model training, and NBS potentials derived from prediction and status quo), however, serves as a base data layer to support policymakers in their actions. This applies particularly with regard to creating transparency on the most crucial factors for adaptation to climate change, in keeping open spaces (natural hazards/ventilation), avoiding sealing, increasing infiltration capacity, greening and networking green corridors and green structures, promoting blue infrastructure, and a combination of these.

6. Conclusion and Directions for Further Research

Green and open spaces are often neglected in spatial planning and treated as residual areas. The Sustainable Development Goals and their targets are difficult for municipalities to implement in planning practice. This is mainly due to a lack of knowledge about the effects of planning decisions and a lack of action support for planning actors within a planning process. In this study, a regression model was developed that can serve as a supporting tool for strategic analysis of past and future planning impacts. It provides the conceptual fieldwork foundation for the development of learning datasets. These are designed to allow the strategic use of analysis algorithms and can be used in other models and methods where transparent land use classification and nature and climate impact analysis are required.

When trained for different regions, spatial situations, and planning regimes, these models can provide valuable insights and a basis for iterative planning and decision-making processes in local spatial planning. In addition, the models should aid and encourage authorities to monitor and enforce the implementation of planning guidelines. In order to minimize data gaps and increase the power of interpretation, participatory approaches also offer potential for future studies.

To address the concept of virtual NBS elements and their implementation, average tables for different setups based on collected and analyzed data are needed. These could serve as benchmarks or reference points for configuring training datasets to ensure real-world reflections. The average setups allow input variables to be standardized across models for more consistent and comparable results.

Further research should identify current knowledge and technology gaps to better preserve soil and reduce sealing, while also evaluating the effectiveness of specific NBS implementations. This could involve developing advanced data analysis algorithms, exploring innovative materials and designs for NBS, and gaining a deeper understanding of their socio-economic impacts. Implementing pilot projects as part of this research can provide real-world testing and refinement of theories and models, yielding practical insights.

Acknowledgments

The authors acknowledge TU Wien Bibliothek for financial support through its Open Access Funding Program and for financial support for proofreading and thank the municipality of Hörsching for providing the development plans for the study area.

Conflict of Interests

The authors declare no conflict of interests.

Data Availability

The research data can be provided on request.

References

Bardt, H. (2018). Kommunale Umweltpolitik zwischen strategischer Planung und Pragmatismus—Das Beispiel der Anpassung an den Klimawandel. In A. Gadatsch, H. Ihne, J. Monhemius, & D. Schreiber (Eds.), *Nachhaltiges Wirtschaften im digitalen Zeitalter* (pp. 323–332). Springer Gabler.

BEV. (2023). *Kataster Stichtagsdaten*. https://www.bev.gv.at/Services/Produkte/Kataster-und-Verzeichnisse/Kataster-Stichtagsdaten.html#download-05-1

Bierbaum, R., Smith, J. B., Lee, A., Blair, M., Carter, L., Stuart Chapin, F., III, Fleming, P., Ruffo, S., Stults, M., McNeeley, S., Wasley, E., & Verduzco, L. (2013). A comprehensive review of climate adaptation in the United States: More than before, but less than needed. *Mitigation and Adaptation Strategies for Global Change*, *18*, 361–406. https://doi.org/10.1007/s11027-012-9423-1

Boehnke, D., Jehling, M., & Vogt, J. (2023). What hinders climate adaptation? Approaching barriers in municipal land use planning through participant observation. *Land Use Policy*, *132*, Article 106786. https://doi.org/10.1016/j.landusepol.2023.106786

Brokking, P., Mörtberg, U., & Balfors, B. (2021). Municipal practices for integrated planning of nature-based solutions in urban development in the Stockholm region. *Sustainability*, *13*(18), Article 10389. https://doi.org/10.3390/su131810389

Buschmann, D., Koziol, K., Bausch, T., & Reinhard, S. (2022). Adaptation to climate change in small German municipalities: Sparse knowledge and weak adaptive capacities. *Natural Resources Forum*, *46*(4), 377–392.

Büttner, G., Feranec, J., Jaffrain, G., Mari, L., Maucha, G., & Soukup, T. (2004). The CORINE Land Cover 2000 project. *EARSeL eProceedings*, *3*(3), 331–346

Chaturvedi, V., & de Vries, W. T. (2021). Machine learning algorithms for urban land use planning: A review. *Urban Science*, *5*(3), Article 68. https://doi.org/10.3390/urbansci5030068

Chen, C., Yan, J., Wang, L., Liang, D., & Zhang, W. (2021). Classification of urban functional areas from remote sensing images and time-series user behavior data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 1207–1221. https://doi.org/10.1109/JSTARS.2020.3044250

Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an infrastructure for spatial information in the European community (INSPIRE). (2007). *Official Journal of the European Union*, L 108.

Dornaika, F., Moujahid, A., El Merabet, Y., & Ruichek, Y. (2016). Building detection from orthophotos using a machine learning approach: An empirical study on image segmentation and descriptors. *Expert Systems with Applications*, *58*, 130–142. https://doi.org/10.1016/j.eswa.2016.03.024

Drachenfels, O. v. (2021). *Kartierschlüssel für Biotoptypen in Niedersachsen—Unter besonderer Berücksichtigung der gesetzlich geschützten Biotope sowie der Lebensraumtypen von Anhang I der FFH-Richtlinie* (Stand März 2021). Niedersächsischer Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz.

European Commission. (2015). *Towards an EU research and innovation policy agenda for nature-based solutions & re-naturing cities*. https://doi.org/10.2777/765301

Fuller, R. A., & Gaston, K. J. (2009). The scaling of green space coverage in European cities. *Biology Letters*, *5*(3), 352–355. https://doi.org/10.1098/rsbl.2009.0010

Fünfgeld, H. (2010). Institutional challenges to climate risk management in cities. *Current Opinion in Environmental Sustainability*, *2*(3), 156–160.

Fünfgeld, H., Fila, D., & Dahlmann, H. (2023). Upscaling climate change adaptation in small- and medium-sized municipalities: Current barriers and future potentials. *Current Opinion in Environmental Sustainability*, *61*, Article 101263.

Geoland. (2024). *Basemap.at Verwaltungsgrundkarte Raster Österreich*. Data.gv.at. https://www.data.gv.at/katalog/dataset/2c54f4d6-5712-4e5a-a025-b7f7a396c39b

González-Collazo, S. M., Balado, J., González, E., & Nurunnabi, A. (2023). A discordance analysis in manual labelling of urban mobile laser scanning data used for deep learning based semantic segmentation. *Expert Systems with Applications*, *230*, Article 120672. https://doi.org/10.1016/j.eswa.2023.120672

Gosling, L., Sparks, T. H., Araya, Y., Harvey, M., & Ansine, J. (2016). Differences between urban and rural hedges in England revealed by a citizen science project. *BMC Ecology*, *16*(Suppl. 1), Article 15. https://doi.org/10.1186/s12898-016-0064-1

Haase, A., Bedtke, N., Begg, C., Gawel, E., Rink, D., & Wolff, M. (2018). On the connection between urban sustainability transformations and multiple societal crises. In S. Kabisch, F. Koch, E. Gawel, A. Haase, S. Knapp, K. Krellenberg, J. Nivala, & A. Zehnsdorf (Eds.), *Urban transformations* (pp. 61–76). Springer.

Helldin, J.-O., Wissman, J., & Lennartsson, T. (2015). Abundance of red-listed species in infrastructure habitats—"Responsibility species" as a priority-setting tool for transportation agencies' conservation action. *Nature Conservation*, *11*, 143–158. https://doi.org/10.3897/natureconservation.11.4433

Karila, K., Matikainen, L., Karjalainen, M., Puttonen, E., Chen, Y., & Hyyppä, J. (2023). Automatic labelling for semantic segmentation of VHR satellite images: Application of airborne laser scanner data and object-based image analysis. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, *9*, Article 100046. https://doi.org/10.1016/j.ophoto.2023.100046

Matthies, S. A., Rüter, S., Schaarschmidt, F., & Prasse, R. (2017). Determinants of species richness within and across taxonomic groups in urban green spaces. *Urban Ecosystems*, *20*(4), 897–909. https://doi.org/10.1007/s11252-017-0642-9

Measham, T. G., Preston, B. L., Smith, T. F., Brooke, C., Gorddard, R., Withycombe, G., & Morrison, C. (2011). Adapting to climate change through local municipal planning: Barriers and challenges. *Mitigation and Adaptation Strategies for Global Change*, *16*(8), 889–909. https://doi.org/10.1007/s11027-011-9301-2

Morakinyo, T. E., Ouyang, W., Lau, K. K.-L., Ren, C., & Ng, E. (2020). Right tree, right place (urban canyon): Tree species selection approach for optimum urban heat mitigation—Development and evaluation. *Science of the Total Environment*, *719*, Article 137461. https://doi.org/10.1016/j.scitotenv.2020.137461

Nagappan, S. D., & Daud, S. M. (2021). Machine learning predictors for sustainable urban planning. *International Journal of Advanced Computer Science and Applications*, *12*(7), 772–780.

Offerle, I., Eliasson, C. S. B., Grimmond, B., & Holmer, B. (2007). Surface heating in relation to air temperature, wind and turbulence in an urban street canyon. *Boundary-Layer Meteorology*, *122*(2), 273–292.

Oijstaeijen, W. V., Passel, S. V., Back, P., & Cools, J. (2022). The politics of green infrastructure: A discrete choice experiment with Flemish local decision-makers. *Ecological Economics*, *199*, Article 107493. https://doi.org/10.1016/j.ecolecon.2022.107493

Pan, Q., Amy, J. S., Yongfei, B., Jianhui, H., Jianguo, W., Shahid, N., Dima, C., Dashuan, T., Qibing, W., & Xingguo, H. (2021). Biodiversity–productivity relationships in a natural grassland community vary under diversity loss scenarios. *Journal of Ecology*, *110*(1), 210–220. https://doi.org/10.1111/1365-2745.13797

Radinger-Peer, V., Penker, M., Chiari, S., Danzinger, G., Enengel, B., Kuhnel, F., & Sammer, K. (2015). Regional vulnerability to the challenges of climate change and energy provision: Lessons learned from transdisciplinary assessments in Austria and Germany. *GAIA – Ecological Perspectives for Science and Society*, *24*(4), 261–270.

Raymond, C., Frantzeskaki, N., Kabisch, N., Berry, P., Breil, M., Razvan Nita, M., Geneletti, D., & Calfapietra, C. (2017). A framework for assessing and implementing the co-benefits of nature-based solutions in urban areas. *Environmental Science & Policy*, *77*, 15–24. https://doi.org/10.1016/j.envsci.2017.07.008

Seddon, N., Chausso, A., Berr, P., Girardin, C., Smith, A., & Turner, B. (2020). Understanding the value and limits of nature-based solutions to climate change and other global challenges. *Philosophical Transactions of the Royal Society B*, *375*(1794), Article 20190120. http://doi.org/10.1098/rstb.2019.0120

Shorter, N., & Kasparis, T. (2009). Automatic vegetation identification and building detection from a single nadir aerial image. *Remote Sensing*, *1*(4), 731–757. https://doi.org/10.3390/rs1040731

Statistik Austria. (2023). *Zensus Gebäude- und Wohnungszählung 2021: Ergebnisse zu Gebäuden und Wohnungen aus der Registerzählung*. https://www.statistik.at/services/tools/services/publikationen/detail/1650,20240701

Storbjörk, S. (2010). 'It takes more to get a ship to change course': Barriers for organizational learning and local climate adaptation in Sweden. *Journal of Environmental Policy & Planning*, *12*(3), 235–254. https://doi.org/10.1080/1523908X.2010.505414

Takouabou, S. C. K., Diop, B., Azmi, R., Jaligot, R., & Chenal, J. (2022). Reviewing the application of machine learning methods to model urban form indicators in planning decision support systems: Potential, issues and challenges. *Journal of King Saud University – Computer and Information Sciences*, *34*(8B), 5943–5967. https://doi.org/10.1016/j.jksuci.2021.08.007

Thomas, J. A., Snazell, R. G., & Ward, L. K. (2003). Are roads harmful or potentially beneficial to butterflies and other insects? In B. Sherwood, D. Cutler, & J. A. Burton (Eds.), *Wildlife and roads: The ecological impact* (pp. 203–222). Imperial College Press. https://doi.org/10.1142/9781860949586_0019

Vringer, K., de Vries, R., & Visser, H. (2021). Measuring governing capacity for the energy transition of Dutch municipalities. *Energy Policy*, *149*, Article 112002. https://doi.org/10.1016/j.enpol.2020.112002

Zhou, P., & Chang, Y. (2021). Automated classification of building structures for urban built environment identification using machine learning. *Journal of Building Engineering*, *43*, Article 103008. https://doi.org/10.1016/j.jobe.2021.103008

Zölch, T., Rahman, M. A., Pfleiderer, E., Wagner, G., & Pauleit, S. (2019). Designing public squares with green infrastructure to optimize human thermal comfort. *Building and Environment*, *149*, 640–654. https://doi.org/10.1016/j.buildenv.2018.12.051

About the Authors

Julia Forster is an architect and head of the Spatial Simulation Lab (Simlab) at TU Wien, where her research work focuses on innovative data analysis and digital spatial visualization solutions to provide holistic views of systems and facilitate agile collaboration among experts from different fields. She completed her PhD in spatial planning under the auspices of the URBEM doctoral college and in 2017 was awarded the TU Wien Ressel Prize for her doctoral dissertation.

Stefan Bindreiter is a spatial planner and software developer and holds an MSc in spatial planning from TU Wien and an MSc in digital media from the University of Applied Sciences in Hagenberg. He is a PhD student and project researcher at TU Wien's Spatial Simulation Lab with a research focus on algorithm-based analytical tools for spatial planning and the concrete application of these methodologies in the creation and implementation of planning strategies.

Birthe Uhlhorn is a research associate at the Institute of Landscape Development, Recreation and Conservation Planning at the University of Natural Resources and Life Sciences Vienna (BOKU University). In her doctoral dissertation, she examines the impact of advancing digitalization and the continuously improving data situation in the field of monitoring and follow-up in environmental assessments, and generates hypotheses on the future role of advanced digitalization in environmental planning.

Verena Radinger-Peer trained as a landscape planner at the University of Natural Resources and Life Sciences, Vienna (BOKU). As a post-doc researcher, she worked at the Fraunhofer Institute in Karlsruhe (Germany) as well as the University of Twente (The Netherlands). Since 2023 she has held a tenure-track professorship in sustainable landscape development at the University of Natural Resources and Life Sciences, Vienna (BOKU). Her research focuses on climate adaptation, sustainable spatial and landscape development, as well as transdisciplinary research.

Alexandra Jiricka-Pürrer is a senior scientist at the University of Natural Resources and Life Sciences Vienna with a habilitation in environmental planning. She is an expert in climate change adaptation and mitigation as well as in biodiversity preservation in the field of environmental and spatial planning. Her research activities have focused on various decision-making contexts, from strategic policies, plans, and programs, and from national and federal-state level to local level.