Diplomarbeit
(Diploma Thesis)

Ecological Urbanistic Analysis and its possible Implementation through Machine Learning

A Conceptual Framework for the Integration of Eco-Complexity into Analytical Tasks in Multi-Species Urban Design with Potentials and Challenges for its Implementation through Machine Learning Classification

ausgeführt zum Zwecke der Erlangung des akademischen Grades Diplom-Ingenieur / Diplom-Ingenieurin an der TU-Wien, Fakultät für Architektur und Raumplanung

submitted in satisfaction of the requirements for the degree of Diplom-Ingenieur / Diplom-Ingenieurin at the TU Wien, Faculty of Architecture and Planning

von

Alexander Authried, BSc
00626359

Betreuer: Univ.Prof. Arch. Dipl.-Ing. Michael Ulrich **Hensel**, PhD
Mitbetreuung: Assistant Prof. Milica **Vujovic**, PhD
E259 Institut für Architekturwissenschaften
E259-01 Digitale Architektur und Raumplanung
Technische Universität Wien,
Karlsplatz 13, 1040 Wien, Österreich

Wien, am **eigenhändige Unterschrift**

Abstract

Urbanization exerts pressure on natural environments within and outside urban areas. This affects human life as well as other species and threatens the goals of sustainable development. While most attempts in architecture and urban design to cope with the challenges imposed on cities are rather following a normative ideal of the 'green city', recent approaches are aiming at incorporating ecological knowledge through performance-oriented and data-driven design methods. The 'ECOLOPES' project is developing a recommendation system for multi-species building design in urban contexts. Within this project, urban classification is determined to identify potential project sites with similar urban and ecological conditions. Such classifications are a crucial analytical foundation to purposefully develop ecologically sophisticated design proposals. Although urban classification has been a research topic with many applications, a conceptual approach to develop classifications as analytical method, addressing the complex behaviour of ecological systems, has not yet been undertaken. Machine learning provides a set of tools which are promising to engage in new ways with big data and its underlying patterns, but research in the context of ecological urban design is sparse. As machine learning studies are characteristically experimental and case-based, and theoretical foundations for interdisciplinary application are still lacking, this thesis sets out to conceptualize *'ecological urbanistic analysis' (EUA)* as a framework for analysing complex ecological urban systems at multiple spatio-temporal scales and further assess the potentials and challenges to facilitate the implementation of EUA in urban design through machine learning methods. This thesis is methodically based on grounded theory and uses elements of literature reviews to support the comprehensibility of the results. The main part is a two-step synthesis: 1) a transect of theories of biodiversity, ecological systems and spatial analysis as a conceptual framework for EUA in the context of multi-species urbanism, which incorporates different dimensions of eco-complexity, and 2) an assessment of current machine learning methods and algorithms in the context of urban ecology with a short review of case studies. This thesis concludes that machine learning offers a novel and interesting approach to analysing complex ecological systems within urban areas, where different algorithms are suited for different aspects of eco-complexity. However, there are also important trade-offs and challenges that come with machine learning as computational method. Although up to date only few studies at small scales, addressing issues of urban ecology and biodiversity, have been conducted, there is a lot of potential for future research, and ultimately a chance to mainstream ecological knowledge into urban design and architecture, to support long-term sustainable urban development.

Keywords: urban analysis; urban ecology; biodiversity; eco-complexity; machine learning; urban design; classification

Kurzfassung

Urbanisierung bringt natürliche Umwelten, sowohl innerhalb als auch außerhalb von Städten unter Druck. Betroffen sind sowohl menschliches Leben als auch andere Spezies und das Ziel einer nachhaltigen Entwicklung insgesamt. Während Versuche, durch Architektur und Design mit diesen Herausforderungen einen adäquaten Umgang zu finden, mehrheitlich durch ein normatives Ideal der 'grünen Stadt' geprägt sind, zielen neuere Herangehensweisen auf die Integration ökologischen Wissens durch performance-orientierte und datengesteuerte Methoden ab. Das 'ECOLOPES' Projekt entwickelt einen Empfehlungsdienst zur Inklusion von Artenvielfalt im Gebäudedesign urbaner Regionen. Im Rahmen dieses Projekts sollen urbane Klassifikationen potenzielle Bauplätze mit ähnlichen urbanen und ökologischen Bedingungen identifizieren. Solche Klassifikationen sind eine essenzielle analytische Grundlage, um zielgerichtet ökologisch bedeutsame Designvorschläge zu entwickeln. Obwohl urbane Klassifikation bereits breite Anwendungen gefunden hat, fehlt ein konzeptueller Ansatz um Klassifikationen als analytische Methode, die das Verhalten komplexer ökologischer Systeme integriert, weiterzuentwickeln. Machine Learning bietet eine Auswahl an Werkzeugen, die einen neuen Zugang zu Big Data und den Daten zugrundeliegender Muster versprechen, aber Forschung im Bereich ökologisch-urbanen Designs ist rar. Da Studien zu Machine Learning meistens experimentell und fallorientiert gestaltet sind und theoretische Grundlagen zur interdisziplinären Anwendbarkeit noch fehlen, hat diese Diplomarbeit zum Ziel *'ökologisch-urbanistische Analyse' (ÖUA)* als spezifischen Ansatz für die Analyse komplexer ökologischer, urbaner Systeme auf multiplen räumlichen und zeitlichen Ebenen zu konzeptualisieren und weiters die Potenziale und Herausforderungen zu evaluieren, um die Implementierung von ÖUA im Städtebau durch Methoden des Machine Learning voranzutreiben. Diese Diplomarbeit setzt als Methodik 'Grounded Theory' mit Elementen von Literaturreviews ein, um die Nachvollziehbarkeit der Ergebnisse zu unterstützen. Daraus geht eine zweistufige Synthese hervor: 1) ein Transekt verschiedener theoretischer Ansätze zur Biodiversität, ökologischen Systemen und räumlicher Analyse, als konzeptuelles Framework für ÖUA und 2) eine Evaluierung aktueller Methoden und Algorithmen im Machine Learning im Kontext urbaner Ökologie, inklusive eines kurzen Reviews durchgeführter Studien. Diese Diplomarbeit kommt zum Schluss, dass Machine Learning Methoden einen neuartigen und vielversprechenden Ansatz zur Analyse ökologischer Systeme in urbanen Umgebungen bietet, wobei unterschiedliche Algorithmen verschiedene Aspekte von Öko-Komplexität erfassen können. Allerdings existieren bedeutende Abwägungen und Herausforderungen, die mit Machine Learning als computergestützter Methode einhergehen. Obwohl bis dato nur einige wenige Analysen, die Themen von urbaner Ökologie und Biodiversität beinhalten, in kleinem Maßstabe durchgeführt wurden, bietet sich großes Potenzial für weitere Forschung, und in letzter Konsequenz die Gelegenheit ökologisches Wissen in Städtebau und Architektur zu implementieren, um langfristig nachhaltige urbane Entwicklung zu unterstützen.

Schlagwörter: urbane Analyse; urbane Ökologie; Biodiversität; Ökokomplexität; Machine Learning; Städtebau; Klassifikation

Acknowledgements

My gratitude goes to my supervisor Univ.Prof. Arch. Dipl.-Ing. Michael Hensel, PhD for giving me the opportunity to write this thesis and his encouragement to engage with this demanding topic. Further, I want to thank him and my co-supervisor Assistant Prof. Milica Vujovic, PhD for their support in writing this thesis.

From the bottom of my heart I want to thank my parents, Eva and Kurt, for the many years of love, support, and open doors. Special thanks go to my mother, who lived with me through the highs and lows and further helped proofreading this thesis.

Special thanks also to Neda, who pushed me to finally write this thesis. She was there as emotional support and to renew my self-belief when necessary. She also would not let me forget that academic achievements are only one aspect of a fulfilled life.

Finally, thanks to all my 'silent' companions who stayed at my side over the years.

Danksagung

Meinem Betreuer Univ.Prof. Arch. Dipl.-Ing. Michael Hensel, PhD danke ich für die Möglichkeit diese Arbeit zu schreiben und seine Ermutigungen, ein solch anspruchsvolles Thema zu behandeln. Des Weiteren möchte ich ihm und meiner Co-Betreuerin Assistant Prof. Milica Vujovic, PhD für ihre Unterstützung beim Verfassen der Arbeit danken.

Aus tiefstem Herzen möchte ich mich bei meinen Eltern, Eva und Kurt, für die vielen Jahre der Liebe, Unterstützung und offenen Türen bedanken. Insbesondere gilt dies für meine Mutter, die jedes Hoch und Tief mit mir durchlebt und bei dieser Arbeit durch Korrekturlesen mitgeholfen hat.

Besonderer Dank gilt auch Neda, die mich schlussendlich dazu gebracht hat diese Diplomarbeit zu schreiben. Sie war mir immer wieder eine seelische Stütze und hat meinen Glauben an mich selbst gestärkt wenn dies nötig war. Außerdem ließ sie mich nicht vergessen dass ein erfülltes Leben aus mehr als akademischer Leistung besteht.

Zuletzt danke ich auch noch allen 'stillen' Begleitern die über die Jahre an meiner Seite geblieben sind.

List of Abbreviations

AHC Analytical hierarchical clustering.
AI Artificial intelligence.
ANN Artificial neural network.
ART Adaptive resonance network.

BPNN Backpropagation neural network.

CLARA Clustering large applications.
CNN Convolutional neural network.

DNLP Dynamic network link prediction.
DNN Deep neural network.
DT Decision tree.

EIM ECOLOPES information model.
EUA Ecological urbanistic analysis.

GAN Generative adversarial network.
GAT Graph attention neural network.
GCNN Graph convolution neural network.
GIS Geographic information system.
GNN Graph neural network.
GW-RF Geographically weighted random forest.

KPI Key performance indicator.

LISA Local indicators of spatial association.
LSTM Long short-term memory.
LULC Land use and land cover.

MAUP Modifiable areal unit problem.
ML Machine learning.
MLPNN Multi-layer perceptron neural network.
MVP Minimum viable population.

NDVI Normalized difference vegetation index.

OCNN Object-based convolutional neural network.

PAM Partitioning around medoids.
PCA Principal component analysis.
PPCA Probabilistic principal component analysis.

REDCAP Regionalisation with dynamically constrained agglomerative clustering and partitioning.
RF Random forest.
RNN Recurrent neural network.

SKATER Spatial 'K'luster analysis by tree edge removal.
SVM Support vector machine.

UGCoP Uncertain geographic context problem.

VAE Variational autoencoder.

xAI Explainable artificial intelligence.

Contents

1. Introduction

Rapid urbanization is posing unforeseen challenges for architecture and urban design. The unprecedented scale and speed of urban growth has been documented extensively, concluding that the integrity of earth's planetary ecosystem will be greatly dependent on the conditions of future urban development (Alberti, 2005; McKinney, 2006; Ahern, 2013; Weisser et al., 2023). Only about two decades ago, with the beginning of the next millennium, an interdisciplinary surge in interest about sustainability and resilience of cities appeared (Ahern, 2013; Elmqvist et al., 2019). Although today, green cities have become an almost standard formula within the architectural and urban design textbooks, design practitioners tend to oversimplify dynamic ecological processes and assume design outcome based on linear and static thinking.

This introductory section will draw a short recapitulation of different concepts of urbanism with regards to their understanding and integration of nature and ecological processes. Of course, such an outline cannot address developments in depth and may be seen as overly simplifying, but it merely intends to demonstrate how the conceptualization of the natural and the cultural sphere have changed, reflected by the organization of cities in Western European and North American societies. First, a historic outline will highlight the recent development in the understanding of the occurrence of natural phenomena in cities from the turn of the 20th century on. Second, the 'ECOLOPES' project will be introduced as an example of current development in data-driven, computational, multi-species architectural and urban design, by its approach to define urban environments by their social-ecological condition and design capacity. Third, the possibility to deal with complex systems through urban design projects will raise important considerations, and fourth, the possibility to use machine learning methods to address those considerations will underline and exemplify the research concept and questions accordingly.

1.1 The Human-Nature Dichotomy and the City

For most of history, amongst other important reasons, settlements and cities had been fortifications against nature and its perils, a 'wilderness' threatening human survival. From medieval to baroque cities, efforts to keep building structures as dense as possible led to several other hazards, such as fires, pests and diseases. A famous example of how analytical inquiry into spatial patterns could help such problems is the investigation of a cholera outbreak in London in 1854 by John Snow. A mapping of cholera deaths in combination with the location of communal water pumps revealed the connection of contaminated water sources and infections (Tulchinsky, 2018). In the mid-nineteenth century, European cities started to tackle the problems of overly dense cities by radically demolishing and restructuring parts of the old urban fabric, amongst others, the Haussmannian renovation of Paris or the construction of the Viennese 'Ringstraβe'. Natural elements, such as street trees, parks and gardens played an important role as aesthetic and representative elements, as well as for recreational and leisure activities.

With the advent of the modernist era at the onset of the 20th century, the idea of industrialization took over in architecture and urban design, viewing cities as mechanistic apparatuses with processes which were assumed to be predictable and linear. This glorification of human technology and machinery culminated in plans and depictions of 'ideal' cities such as the 'Hilberseimer Hochhausstadt'. The 'Congrès Internationaux ď Architecture Moderne' (CIAM) developed the Charter of Athens, which proclaimed urban design and planning as a task of functional zoning, which recognised the problematics of urbanisation and the loss of natural environments for human wellbeing. Although 'open spaces' within cities should integrate existing natural features, they were mainly addressed as buffer zones between housing and industry or other detrimental functions to human well-being, as well as providing recreational and sporting amenities (n.n., 1963). Natural elements within cities were seen as a planned constant, which would have to be managed to remain in a specific condition.

Over the past century, cities have been set up by these modernistic ideals, leading to urban sprawl, developing into patchworked landscapes with markedly decreased ecological function. Urban design and planning were oriented towards energy- and infrastructure-intensive forms of living, coupled with the maxim to create architecture as an engineering task to minimize impact of natural phenomena and improve hygienic conditions. In contrast to what has been intended, this development has led to severe degradation of life-supporting conditions for humans, as well as other organisms.

1.1.1 The Ideal of the 'Green City'

From the late 1970s on, a series of investigations into ecological degradation of human-dominated landscapes and limited planetary capacities to support life, together with the mainstreaming of environmental protection, the world has turned its attention to 'sustainable development'. This term was famously coined by the 'World Commission on Environment and Development' in 1987, which defined sustainable development of a society as having the capacity "to ensure that it meets the needs of the present without compromising the ability of future generations to meet their own needs (United Nations, 1987)".

With the new millennium the political agenda was set to the Millenium Development Goals, which in 2015 were superseded by the 'Sustainable Development Goals' (United Nations, s.a.). To this point, architecture and urban design efforts towards sustainability were mainly focused on experimenting with urban form and engineering energy and material efficiency (Ahern, 2013). The turn to sustainability and resilience was met with a renewed interest in the rediscovery of nature as design element. In face of a changing climate, which exerts additional pressure on cities as densely populated human habitats, lacking resilience to adapt to increasing heat waves or other disturbances while pertaining function as human habitats, the most promising and relatable measures to promote human well-being are reintroducing vegetation and green spaces into the built environment (Lee and Maheswaran, 2010; Perini et al., 2021), as well as strategic design of urban built-up areas (e.g. wind channels for air circulation). Although such measures are backed by scientific research, a systematic approach from an ecosystem perspective has been missing until now. This is due to conventional planning regimes which rarely consider urban ecosystems, the services and disservices they produce, and the way they contribute to urban biodiversity, or the loss of it (Canepa et al., 2022; McPhearson et al., 2015; Fineschi and Loreto, 2020).

Such planning is usually limited to few key parameters, such as "planting a small selection of plants" (Canepa et al., 2022), isolated as individual projects, facultatively connected by guiding principles at larger scales of urban development. Even within such individual projects, the implementation of vegetation and other natural elements is usually strictly separated from building construction and designed by different professions - namely architects and landscape designers, prohibiting integration of natural and man-made design elements. As urban design and planning are deeply rooted in normative practices, still dominated by a modernistic design philosophy, green infrastructure in cities is usually conceived of as an 'ideal' condition (Ahern, 2013). Once reached, plants and animals are heavily controlled and managed (pest control, gardening, etc.) (Canepa et al., 2022). Such controlled biotopes, with a limited function to enhance human comfort and well-being, are marked by reduced biodiversity and ecological function compared to natural biotopes, which can develop and maintain high levels of biodiversity through natural processes and cycles (Canepa et al., 2022). This suggests that creating green infrastructures should only be viewed as a first step (Hostetler et al., 2011).

1.1.2 Cities are Ecosystems

The presented historic practices of urban planning led to urbanized areas developing into severely, and often abruptly altered and degraded environments (Alberti, 2005) compared to the pristine ecosystems they replaced. But 'green' developments to combat the loss of ecosystem function are usually imposing 'nature' rather as a dogmatic idea, merely accounting for resource efficiency and design optimization (Ahern, 2013). Although scientific research about biological organisms within cities was initiated by biologists already in the 1930s (Wu, 2014), only with the involvement of ecologists and the unfolding discourse about complex systems, approaches were slowly shifting from an 'ecology in cities' towards an 'ecology of cities' which also viewed human artifacts in urban agglomerations as parts of a bigger ecosystem (Wu, 2014). However, it is assumed that cities as ecosystems function in completely different ways than rural or native ones, and traditional ecological models are not sufficient to describe their patterns and processes (Alberti, 2005). Such ecosystems differ in several ways from other natural or cultivated landscapes (Alberti, 2005):

1. Urban systems are structurally unique, due to their highly altered topography through built-up masses.
2. They are heterotrophic in that sense, that they need resources, especially nutrients, which must be imported from outside the physical urban boundaries.
3. Since there is no possibility for natural nutrient cycling inside cities, there is also a concentrated outflow of waste, that must be treated at the urban fringes or outside.
4. Due to the spatial modification cities experience adverse effects of immissions (e.g. radiation as heat) and emissions (particulate matter) respectively.
5. Due to an overall lowered resilience (high population density, modified environment with little biomass and altered hydrological flows), cities are prone to disturbances (heat waves, pests, etc.).

Further, natural processes are highly limited due to (Alberti, 2005):

1. Lack of habitat patches, offering space for resources to animals, plants, fungi and microbiota;
2. Invasion of non-native species through human induced import;
3. Heavily managed green spaces suppressing natural community assemblies and plant successions.

1.1.3 A Shift towards Ecosystem Services and Biodiversity

Several measurements and standards have been developed to address and value the function of ecosystems. Ecosystem services (ES) are a key concept to value ecosystem functions which are beneficial to human life, whereas many ES are directly accountable for human health and wellbeing (e.g. climate regulation in dense built-up urban areas) (Costanza et al., 1997). The assessment of ES targets the quantitative and qualitative services that humans receive through processes in ecosystems (e.g., O_2 production from photosynthesis) as services or benefits, and adverse effects (e.g. human-wildlife conflicts) or costs (Döhren and Haase, 2015).

Although approaches such as ES focus on ecological processes, they are still conceptually human-centred (Weisser et al., 2023). This includes the reproduction of cultural practices through design and construction of the built environment, covering a range from creating highly specialized infrastructures (e.g., streets are covering vast areas in pavement, only to be passable for vehicles) which in turn must be maintained in this very specific state to function properly (Weisser et al., 2023). This in turn inhibits ecological processes and cycles, and consequently, such areas are marked by a low to vanishing population of organisms. Müller (2005) pointed out the importance of differing between ES as an anthropocentric approach, contrasting the notion of self-organizing and regulating capacities of ecosystems, which provide a basis for a multi-species perspective.

Recent development is shifting away from a notion of ES as a purely anthropocentric perspective, to a biophilic or 'bio-inclusive' approach. 'Regenerative design' is alluding to the idea of the possibility to restore natural conditions or at least regain a considerable amount of natural functionality and ES in the process (Perini et al., 2021). Trying to move beyond an anthropocentric approach to a holistic and inclusive perspective for multiple lifeforms, 'biodiversity' has been postulated as a measure for the richness and abundance of biotic factors inside ecosystems as an indicator for its condition. *Biodiversity* is perceived as being facilitated by ecosystem functions, and has more indirect, but nonetheless crucial effects for creating a favourable human habitat, because in turn, biodiversity is a requirement for the provision of several ES (e.g. pollination) (Pedersen Zari, 2015). Until recently, researchers unanimously conceived of urban areas as inherently adverse to biodiversity. This notion has since been contested by research, which showed a more nuanced picture (see the differentiation between an 'ecology *in* cities' and an 'ecology *of* cities'). Depending on many factors, cities can offer a suitable habitat to several species, in some cases even produce a niche for endangered ones (Gentili et al., 2024).

1.2 Cities from a Non-Anthropocentric Perspective

Conservation efforts for individual species are a highly specialized topic, but the achievement and support for a high general level of biodiversity within urban areas is shifting towards being a structural necessity for a sustainable developing urbanism (Pedersen Zari, 2015; Weisser et al., 2023). This differentiation between human centred and holistic views of natural processes has also found resonance in architectural research, as Hensel (2013) argued for a shift of architecture towards non-anthropocentric design. As a specialization of sustainable or regenerative design, such approaches to non-anthropocentric design are referred to as 'multi-species' or 'wildlife-inclusive' design. Aside the relatively young scientific discourse about architecture and urban design, the idea of humans living in mutually beneficial relationships with other species is of course not new at all. For thousands of years, vernacular architecture has shown ways of symbiotic existence between humans and animals (Rudofsky, 1964). Such an approach has of course severe implications not only for the potential beneficial effects (Döhren and Haase, 2015). Conflicts between humans and other organisms have impacted the development of cities drastically as sketched out in Section 1.1.

The principle of conviviality between humans and other species has been taken up by a group of researchers, driving the development of urban and building design towards the inclusion of wildlife (Apfelbeck et al., 2020; Weisser et al., 2023). Under the concept of 'animal-aided design', the design and development of buildings is concerned with the integration of habitat space (Hauck and Weisser, 2015). A fundamental new aspect of bio-inclusive design approaches compared to the vernacular architectures is that the benefit of providing living space for other species is also seen as a 'cultural' one. Additionally, design is oriented towards full life cycles of targeted species, instead of just supporting certain activities (e.g. nesting sites) (Hauck and Weisser, 2015). This concept is of major important to ensure long-term success (Hauck and Weisser, 2015).

The idea of 'multi-species' design is trying to implement biodiversity as a guiding principle into urban and building design (Weisser et al., 2023). At larger scales, focus on a single or few species for conservation purposes needs to be abstracted to the management of potentially habitable spaces (Perini et al., 2021). Depending on the objectives of urban design cases, numerous different aspects might be of importance to achieve a successful implementation wildlife-inclusive urban design (Perini et al., 2021). Examples for research questions or design problems could be:

- Where are ecological boundaries within the city?
- Which are the most important corridors for flying pollinators to access a site?
- How did a species adapt to urban conditions based on the most important habitat niches?

New non-anthropocentric design paradigms are essential in establishing practical guidelines and provide a firm basis for research by design (Hensel, 2013). However, design studies are representing a case-to-case approach, where every species needs profiling for its specific design needs, which in the context of urban design is limiting the approach to developing rather isolated prototypes. This leaves big potential for a systematized way towards an integrated design approach. Up until now, only few projects are concerned with the embedding of ecology through scientific and performance-based as well as data-driven approaches, amongst which the 'ECOLOPES' project is an outstanding undertaking to offer a design recommendation systems to practitioners (Perini et al., 2021).

Ecological considerations are commonly applied in urban and architectural design practice as an 'afterthought', when the most crucial design parameters have already been determined (Weisser et al., 2023). But once the primary geometry of a project is defined, options for finding ecologically sustainable solutions get narrowed down significantly (Yoffe et al., 2023). As early design stages and decisions are known to have the greatest impact on the overall design outcome, such a *modus operandi* is leaving ecological design efforts as merely a kind of 'virtue signalling', without concise and cohesive implementation of ecological aims and goals. Temporal scales, i.e. the dynamics of ecological systems, often remain completely unconsidered (e.g. plant successions or day-night changes) (Weisser et al., 2023).

Design space is a concept regarding the entirety of all possible solutions to a design task within defined constraints and spatial boundaries, i.e. the boundary conditions to a design problem (Weisser et al., 2023). Within this design space permutations of any design parameter against all others are thinkable. This concept is important for the integration of ecological aspects into architecture and urban design, because - as outlined in Seciton 1.1.1 - traditional design approaches have been developed by coarse guidelines, abstract building codes or functional and aesthetic requirements for human use. But to include biodiversity in architectural and urban design at an ecosystem level, decisions need to be made about the kind of intervention, the biotic communities involved, the existing versus the necessary habitat conditions, plant successions, and beneficial or adverse human-nature interactions (Weisser et al., 2023). For this task it is utmost important to set an optimal 'design space' which is considering the existing conditions, the built environment and human activity to make assumptions or forecast an optimized set of ecological parameters on spatial and temporal scales, including site conditions, neighbourhood environment, habitat connectivity, rural-urban gradients, indigenous species, climatic development (Weisser et al., 2023).

If ecological principles are to be explicitly considered from the design brief onwards, usually architects and planners are cooperating with domain experts as consultants (Weisser et al., 2023). As a drawback, such interdisciplinary cooperation often takes a lot of time and work in terms of finding a common language or common methods and tools to work with, which is an important hurdle lowering the implementation of ecological sustainability and biodiversity enhancement (Moscovitz and Barath, 2022; Yoffe et al., 2020). As early design stages are often marked by high pressure on time and budget, possible exchange of ideas in interdisciplinary teams is regularly inhibited (Yoffe et al., 2023). Another aspect is that the implementation of ecology and sustainability in CAD has until now been dominated by qualitative methods. To achieve a shift in urban design practice from merely normative and qualitative evaluation towards quantifiable methods, metrics have been developed by different scientific domains. Landscape ecologists and landscape planners have come up with a variety of such measurements, often referred to as 'Key Performance Indicators' (KPI) (Selvan et al., 2023a; Yoffe et al., 2023).

1.2.1 The 'ECOLOPES' Project

'ECOLOPES' is an ongoing project, running from 2021 until 2025, funded by the European Commission. The goal is to introduce "a radically new integrated ecosystem approach to architecture, equally focused on humans, plants, animals and associated organisms (European Commission)." The ECOLOPES project considers the information of design with ecological knowledge at early design stages as crucial success factor for integrating a multi-species perspective into architecture and urban design (Perini et al., 2021). The trajectory of this project is directed at the improvement of living conditions in urban areas through the enhancement of biodiversity and the integration of wildlife into the built urban environment (Perini et al., 2021). The ECOLOPES project develops a computational design recommendation system facilitating the design of multi-species building envelopes, guided by ecological expert knowledge, relying on a data-driven design recommendation system (Canepa et al., 2022).

The computational architecture consists of a frontend plugin for CAD modelling software, and a backend featuring an *'ECOLOPES Information Model (EIM)'* where an ontology is describing the relationship and interdependencies of all involved biotic and abiotic elements, including human artifacts (i.e. architecture) (Perini et al., 2021; Weisser et al., 2023). This model is supposed to inform the designer about the specific environmental requirements for targeted biological communities. An evolutionary algorithm is generating design variants, which will be evaluated by human users, facilitated through the EIM (Perini et al., 2021).

Selvan et al. (2023a) developed a multi-objective, multi-level optimization approach, correlating abstract design objectives (e.g., "enhance biomass growth") to *key performance indicators (KPI).* KPI are quantifiable variables, that can be manipulated through architectural and landscape design. In ecology such variables are also referred to as 'proxies', which enable to evaluate effects that cannot be directly measured, allowing the assessment of the design variants' ecological performance. KPI in the ECOLOPES design approach are to be defined by an interdisciplinary design team starting from the formulation of a design brief (Weisser et al., 2023).

As a foundation for the development of new design projects, the ECOLOPES design approach intends the development of a method for urban classification to „select design cases in a systematic way, and to facilitate the comparison of the performances achieved by an 'ecolope' under different sets of environmental and architectural conditions (Perini et al., 2021)". The characterization of similar urban and ecological conditions is articulated as a set of variables and sub-variables on an urban, as well as an architectural scale. Urban classification is seen as a mean for strategic site selection, and should indicate potential sites, suitable to achieve goals in multi-species urban design on different scales, e.g. improvement of biomass production (Perini et al., 2021; Selvan et al., 2023a). Such an implementation should also enable designers to apply the knowledge from simulated ECOLOPES, generated in a mini modelling experiment (MIMO) onto different urban contexts, and support analogies (Vogler et al., 2022). Such classifications will form a dataset within an expert database for ecological modelling by other components of the ECOLOPES architecture (Vogler et al., 2022).

Perini et al. (2021) identified machine learning (ML) approaches, such as *Hierarchical Clustering Analysis (HCA)*, as a suitable method to establish the intended urban classification for the ECOLOPES project. The clustering should facilitate an "integrated classification workflow for urban to architectural scale (Perini et al., 2021)". The referenced implementation is a hierarchical clustering method with Bayesian inference (Araldi et al., 2021), extracting building types by several geometric variables. This algorithm makes pairwise comparisons of the probability that one datapoint is dependent on another (i.e. shares similarities). Datapoints are further grouped pairwise by their similarity, producing k clusters with two datapoints on every level of the hierarchy (Araldi et al., 2021). The big advantage of hierarchical clustering methods is thus, that the resulting clusters can be modified afterwards, by adjusting the desired granularity of differentiation for every branch of clusters on different levels of hierarchy. This allows to estimate the best fitting number of clusters by the intended differentiation of datapoints (Araldi et al., 2021).

The ECOLOPES approach considers two scales of investigation for urban classification (Perini et al., 2021), as shown in Figure 1.1: 1) 100m landscape/urban resolution, 2) 10m architectural resolution.

The definition of cell sizes with corresponding radii suggests, that adjacent cells ought to be considered for the classification. Although there is no explanation about the intent of such radii available in the published papers while writing this thesis, it implies assumed interdependencies between the cells. Perini et al. (2021) further identify variables for urban classification at the two different scales. The multiplicity of variables considered in combination with the clustering approach allows for creating a classification informed by complex and interdependent socio-economic and environmental factors, while balancing the number of total clusters. This information will serve as basis for the development of design projects through further computational processing (Perini et al., 2021).

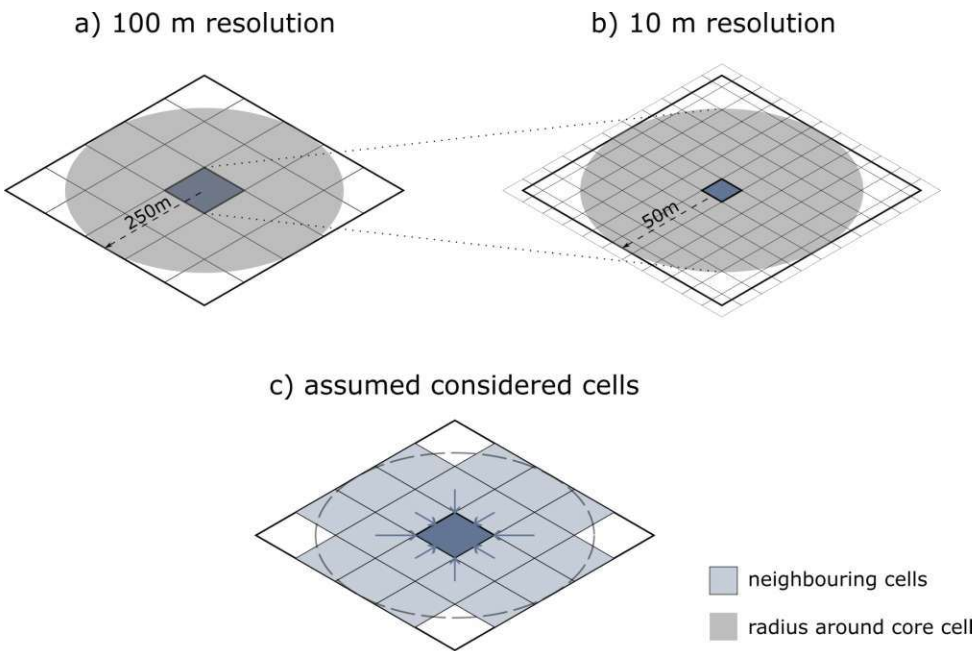

Figure 1.1: a) and b): Planned scales and radii for urban classification in the ECOLOPES approach, after Perini et al. (2021); c): assumed cells to be considered through clustering analysis.

The ECOLOPES project is a spearheading example of a stand-alone, holistic design recommendation system, facilitating multi-species architectural design by coupling knowledge databases, ecological models and ontologies, and CAD software (Perini et al., 2021). However, the shift towards biodiversity support and wildlife-inclusive design poses new challenges for fundamentally rethinking our engagement with architecture and urban design, which will need to be implemented as a new analytical perspective (Weisser et al., 2023).

Designing within urban ecosystems of course cannot be done in isolation. While the ECOLOPES project promises to revolutionize the way how architectural design is incorporating ecological knowledge, ecological inquiry from the neighbourhood to the urban scale remains conceptually vague. Urban design and planning have become standard practice for structuring the available space within and around urban areas, to facilitate an efficient allocation of infrastructures, but architectural projects are commonly treated as ‘stand-alone’ solutions. Clearly such an approach is prone to shortcomings at higher system levels for designing new and developing existing biotopes through design and planning.

1.2.2 Classification as analytical tool for multi-species design

The comparison of the complex relationship of ecological factors can show unexpected results and reveal underlying patterns, that a seemingly unrelated factor is contributing strongly to the optimization of another (e.g. the open space distribution to achieve affordable housing demands in an investor project (Moscovitz and Barath, 2022)). It has become clear, that urban ecosystems need to be evaluated as socio-ecological dynamic systems, to attribute the complexity of interactions between humans and ecological processes, and diverse socioeconomic and biophysical factors need to be taken into account simultaneously (Alberti, 2005). Complex behaviour can be categorized into *disorganized* and *organized*, where the first comprises of independent actions of all agents, and the latter of interactions among them (Jacobs, 1961), and Hanna (2022) argued that cities comprise the latter case. He further states that an essential problem of analysing phenomena of increasing complexity is their representation (Hanna, 2022).

Deterministic modelling has long been the status quo for any investigation in ecological phenomena beyond the mere mapping or regression of variables. Although these models have proved to be of tremendous value, they show limitations for analysing complex systems, since there have to be assumptions about the variables involved, their correlation and distribution, and non-linearity is usually not considered.

Another important aspect about modelling is the fact, that although the prediction of individual behaviour is limited with the currently available tools and data, at larger scales of investigation patterns emerge, which correlate to real-world observations (Hanna, 2022). Such behaviour, which cannot be explained by isolated observations, is also referred to as *emergent* (Alberti, 2005). Still, it is unclear to what extent the patterns and their emergence are dependent on the selection of specific features, at specific scales.

Urban development has several negative impacts on the natural remnant habitats within and around cities, e.g. fragmentation and isolation of habitats, modification of regional energy and material metabolism (Alberti, 2005). It has been acknowledged by many scholars, that a comparison of urban ecosystems across cities and regions will help to identify and establish universal patterns and processes as well as their causalities and interdependencies (Aronson et al., 2016). But heterogeneity and fragmentation add a complexity to urban habitats (Groffman et al., 2017), in which conservative ecological models have limited applicability due to the dominant and constant presence of human agency. Alberti (2005) stated that until then most studies relating to the impact of urbanization on ecosystems had taken coarse aggregated measures of urbanization to correlate to changes in environmental systems.

Although the correlation between ecological characteristics and built urban environment will be a crucial task for multi-species urban design and planning (Weisser et al., 2023), until recently a relatively small amount of research was concerned with the classification of urban ecosystems, at spatial scales where the changing patterns of human infrastructure stocks (i.e. the built environment) are considered. Existing classification schemes in the context of urban ecology typically only account for heterogeneity of physical land cover properties or static land use parameters. This might be due to the long-standing approach of coarse-scale land use and land cover (LULC) classifications as a methodically well-developed research method (Cadenasso et al., 2007).

This entails several problems as classification also means some kind of discretization : 1) In conventional classification approaches high complexity is usually reduced to one single, or few key factors or indicators, or their representation and expression through aggregated classes, 2) if data of different features is far apart, conclusions combining these inputs can be biased, based on the assumption that a variable can change quickly over time or in space, as natural systems are in continuous change, exchange, and development (e.g. the status of vegetation based on the *normalized difference vegetation index (NDVI)* will vary greatly based on the time of year, as might soil humidity), and 3) due to the complex relationships of system parts, continuous updating and strong feedback loops are necessary.

Recently, some scholars have tried to find new, context-specific classifications, such as Stewart and Oke (2012), who combined basic forms of land cover classification with urban morphology to form "local climate zones". Other researchers are developing more precise classifications, tailored to a specific landscape type (Wu et al., 2021). Although this approach is classifying landscapes under the specific lens of climate, classifications remain dependent only on land cover information. 'HERCULES' distinguishes elements of buildings, surfaces and vegetation by features of typology, cover, bare soil, paved, and coarse or fine texture (Cadenasso et al., 2007). Coming from the perspective of landscape ecology, established ecological classification schemes at finer scales still are oriented towards description of vegetational structures (Bunce et al., 2008; Farinha-Marques et al., 2017). But important factors to describe environmental characteristics are dynamics, stochasticity, heterogeneity, pattern-process relationships and associated ecological processes (Dramstad et al., 1996; Ahern, 2013). Cadenasso et al. (2007) attest that conventional classification approaches, either at coarse or fine scales, merge information of land structure and land use in a way, which impedes their application for testing links between structural and functional aspects (Cadenasso et al., 2007). Therefore, over the last decades, reductionist approaches to classifications have been criticised due to their inability to capture emergent behaviour of complex systems (Lu and Yang, 2022).

But classification tasks are not merely about categorizing data into abstract classes of representation, they also imply the need to evaluate the underlying data and apply some form of analytic evaluation or modelling to achieve consistent categorization (i.e. thresholds, anomaly detection, etc.). Urban classification in this sense would feature the analysis of a complex dynamic system based on spatially explicit data of one or multiple points in time, to delineate structure-function relationships for characteristics of interest. To exemplify this assumption, García-Pardo et al. (2023) used a clustering approach to relate building structure to seasonal vegetation patterns. Although of limited complexity, it shows how classifications can be used to relate ecological phenomena to the built urban environment.

1.2.3 Knowledge Discovery through ML methods

If animals, plants, and other species are to be included into architecture and urban design, practitioners are confronted with non-anthropocentric functional and aesthetic requirements (Weisser et al., 2023). Other species have different needs, especially when they are supposed to live in close vicinity or direct contact with a dense human population (i.e. nesting sites, shelter, hunting grounds, etc.), as well as a different perception of their environment (Grobman et al., 2023). Designing and planning cities for biodiversity and as multi-species biotopes has thus the intention to create a “constructed ecosystem”, aiming at providing certain functions or ecosystem services, which are greatly dependent and influenced by the species involved (Lundholm, 2015).

Scholars are not unanimous about the applicability of rules and models from ‘natural’ ecosystems onto urban ones. Many such observations have been published (Alberti, 2005). Although urban ecology has become an established scientific field, the interactions and resulting effects of humans and other species are still not well understood, especially at smaller scales (Aronson et al., 2016). Knowledge derived from other, non-urban, nature-like ecosystems must be reviewed under urban conditions, with qualitative and quantitative approaches, since these systems might differ substantially (Aronson et al., 2016). For such an endeavour it will be necessary to build theoretical foundations linking ecological and architectural knowledge, as Weisser et al. (2023) identify the missing relationships between ecology and architecture as one major challenge for multi-species design to implement.

With the rise of *big data*, the potentials and requirements for analysis of ecological phenomena have changed drastically. Kitchin (2014) is arguing for a regime shift towards a ‘fourth’ paradigm of data-driven knowledge generation, formulating the following characteristics of big data:

> - Huge in volume, consisting of terabytes or petabytes of data;
> - High in velocity, being created in or near real-time;
> - Diverse in variety, being structured and unstructured in nature;
> - Exhaustive in scope, striving to capture entire populations or systems ($n = all$);

As the ECOLOPES approach shows, classifications in urban ecological applications can today potentially address the complexity of urban ecosystems by a plethora of biotic and abiotic properties, facilitated by the massive availability of data. However, our ability to process and analyse such huge amounts of heterogenic data, which is prone to error and generally still sparse in nature has not evolved as quickly as the development of data collection and generation methods (Reichstein et al., 2019).

Looking at recent developments in *artificial intelligence (AI)*, *machine learning (ML)* has gained popularity through the dissemination of ready-to-use algorithms and pretrained models for a variety of computational tasks. These tools have successfully been used for discovering and predicting complex pattern in high dimensional datasets within many scientific disciplines (Scowen et al., 2021).

Machine learning has been conceptually formulated and defined in 1959 by Samuel (1959) as the ability of computer programs to learn from data without explicit programming. Such 'learning' can be measured as improvement in accomplishing given tasks (e.g. to distinguish images of different animal species) (Samuel, 1959). It is today a wide-spread field of research with a plethora of methods, algorithms, and possible applications. The rapidly growing body of knowledge tries to make distinctions in form of algorithmic architecture, overall complexity, and methodical applications (Joshi, 2023).

Machine learning has been applied to 'surrogate modelling' approaches and is being developed to lower the threshold of accessing data-driven analytical design methods (Chaillou, 2022). ML works with mathematical and statistical methods to estimate or iteratively determine patterns in data. In contrast to deterministic models, most ML machine learning methods do not necessitate assumptions about the relationships and distributions of data (Jung, 2022). ML methods are further characterised through (Carta, 2022) as:

- Machine learning differs from other analytic or generative digital tools by the fact, that through 'learning' the patterns in data, algorithms do not only reliably and consistently classify, or cluster given datapoints into meaningful sets, but can also infer this process to unseen or future data, i.e. making estimations (sometimes also called 'predictions').
- Machine learning is facilitated by adapting non-task-specific algorithms to fit onto a problem, either with human support or all by itself. This offers the potential for practitioners to use machine learning on problems, which could not be solved manually. It is not necessary for users to understand the underlying patterns in data to receive meaningful results.

However, there are several limitations to the use of ML algorithms which may limit the applicability for laymen in each domain substantially (Carta, 2022):

- Knowledge about the individual impact of chosen variables, as well as their relationships and interdependencies are crucial to develop reliable models with high performance scores.

- The interpretation of the ML output is not necessarily straightforward, and the interpretability of ML solutions is, depending on the algorithm and the data used, limited.

This opens the question how to classify urban ecological systems, subject to the task of integrating biodiversity into architecture and urban design. While there is a need to capture the complexity of such a system (Aronson et al., 2016), not all factors might have significance to a specific question or aspect of a system (e.g. while the *amount* of fresh water may influence a land-living species, the *type* of water sources may not have a significant impact). Although a classification, supposed to answer specific design questions, will be context dependent, other classifications might be more generalizable (e.g. which potential projects sites share similar conditions for biomass growth). To address this problem, knowledge about the behaviour of cities as social-ecological systems will be necessary.

1.3 Aims and Scope

Ecology and its scientific approaches might be bewildering for many architects and urban designers. But indeed, urbanism could be viewed as an inquiry into an archetype of human habitats. Cities have been described as organism by different thought schools, such as Metabolists, and analogies have established as common language, such as the urban 'tissue'. One of the most intriguing arguments for the investigation of cities as social-ecological systems has been laid out by Alexander et al. (1977) with their seminal works addressing the systems of built environment and human activity as 'patterns'. This system has been analysed and classified according to scale, layout, predominant forms and morphology, function and use, and social aspects. In the sense of a highly differentiated analysis of a specific human habitat form, "A pattern language" was a survey and classification of socio-ecological ecosystems from a sociological and urbanistic or architectural perspective. As much as this analysis still holds true, leaving the anthropocentric perspective leaves us with little evidence about other organisms living amongst humans. Gaining an understanding of how patterns in cities impact other species, calls for new approaches to analytical inquiry.

In contrast to engaging in multi-species architectural design, which needs to take biodiversity as an objective in the design brief (Grobman et al., 2023), where ecologists formulate ecological targets, architects and urban designers usually do not have the knowledge nor literacy to account for ecological potentials when analysing urban space. In addition to sophisticated, data-driven design-solutions like ECOLOPES, methods and tools for understanding urban space as dynamic social-ecological system, with interdependencies at multiple scales, are lacking. Classifications of similar ecological conditions within urban areas, as intended for the ECOLOPES workflow, are a major analytical concept which is yet lacking proper theoretical background for application, because it remains unclear how such complex relationships could be addressed methodically.

Some scholars in the field of applied ecology argued for an integrated approach to ecological modelling, requiring a consistent translational relationship between theory building and empirical research (Peters and Okin, 2017). Other authors emphasized the importance of theoretical knowledge building as foundation for research studies (Pickett et al., 2017). As ML is almost exclusively driven by experimental research, the lack of theoretical background of many studies may lead to biased assumptions about the applicability of algorithms and impede the exploitation and development of ML concepts for the integration of biodiversity in urban design. Following this premisis for the need of knowledge building, I propose that the analysis of design problems in the context of urban eco-systems profits from an integration of field-specific knowledge of biodiversity and urban ecology into a more general theoretical framework of ecological system knowledge, applicable for the analysis of design problems in multi-species urban design.

Ecological Urbanistic Analysis (EUA) thus aims at being an integral part of analytical processes for urban design projects, by providing a conceptual framework, and evaluating its possible implementation through ML methods, by its capabilities to address biodiversity issues by analysing urban ecological conditions as a classification problem with multi-variate, multi-scalar, and spatio-temporal data.

1.3.1 Preliminary literature search

A first search of the phrases "urban ecological classification" and "ecological urban classification" in the databases of *Scopus* (Title, Abstract, Keywords), *Web of Science* (Topic) and *Core UK* yielded zero results. An additional search on *Google Scholar* revealed 8 hits for "urban ecological classification". Three of those records were dismissed due to Chinese language or duplicate data. This initial reporting is to establish the scope and context of the thesis and to see if a conceptualization or a framework has already been established in scientific literature. In summary, the articles found reveal a stark contrast in topics, methods of classification, and spatial explicitness. This leads to the assumption that the topic of urban classification integrating urban ecology has not yet been evolved as a distinct field of research.

In a next step, database searches were executed to identify all relevant research concerning urban classification and analysis under an ecological lens. A literature search on Scopus (in English language) to clarify prior or alternative use of the terms (title, abstract, keywords) *"urban ecolog* classif"*, *"urban ecolog* analy"* (Oct. 3rd, 2023), *"urban classification" AND (ecolog* OR ecosystem OR biodiversity)*, and *"urban analysis" AND (ecolog* OR ecosystem OR biodiversity)* (Sept. 22nd, 2023) yielded few results. The records consisted of different formats (data sheets, articles, conference papers). A screening of the documents showed a wide range of topics and an inconsistent use of the terms 'classification' and 'analysis'. This initial review suggested that the term combinations have neither been used frequently nor consistently. Additionally, 'urban classification' is lacking small-scale and ecological considerations, while 'ecological classification' was lacking consideration of urban systems.

1.3.2 Research questions and structuring

The lack of a clear conception of analytical classifications as tool for research prior to architectural and urban design, as envisioned by EUA, and how to implement them through machine learning methods is the motivation for writing this thesis. To fill these knowledge gaps, the following research questions will be addressed:

- How can ecological urbanistic analysis (EUA), with special regard to biodiversity and multi-species design, be conceptualized, and what are the most important factors? How can an analytical classification of ecological conditions within urban areas be achieved?
- Which types of ecological, analytical classifications of urban environments have been conducted through ML, and where do they fall short in addressing complex ecological systems?
- Which potentials and challenges does state-of-the-art ML entail to deal with highly complex ecological systems in urban analysis and classification for urban and architectural design approaches such as 'ecolopes' and which ML algorithms are suitable for different analytical tasks?

The nature of such an investigation is clearly a highly interdisciplinary one. McPhearson et al. (2016) describe a list of 23 scientific disciplines involved into urban ecological research, with architecture, urban design, and urban planning among them. To address the research topic without a biased perspective from disciplinary customs of architecture and urban design disciplines, literature will be reviewed based upon a thematic search within multiple domains of scientific literature. This puts emphasis on an ecological understanding of patterns and processes within the urban context and relating, translating and integrating such knowledge into a *conceptual framework*. To satisfy this goal, the following sections of the thesis will include:

1. A methodological foundation for the selection of suitable methods to encompass the broad scope of investigation;
2. Establishing space as a common denominator and main object of investigation of landscape ecology and urban design;
3. A review of theoretical paradigms and concepts of biodiversity and ecosystems within the urban context to find relationships and guiding principles, which can be addressed through analytical processes;
4. A conceptual framework for Ecological Urbanistic Analysis (EUA) to connect ecological and urban design theory;
5. An overview of ML in the context of ecology and urban design with a short review of case studies;
6. A discussion of potentials and challenges to promote the integration of biodiversity and multi-species-oriented urban design and architecture through ML.

2. Methods and Materials

Commonly in scientific research, especially with narrow research questions, methodical emphasis is given to the procedural part of methods (Elsbach and van Knippenberg, 2020; Post et al., 2020). As an example, in most review papers, the type of review is selected and the processes and steps within the according method are described. But interdisciplinary research topics should not only be addressed as a procedural but also as a structural problem (Defila and Di Giulio, 2015). Before selecting and planning the method of research or literature search, the topic, research question, expected results and the methods of synthesis should be established in some form and critically appraised, to ensure accurate procedural methods and valid results through synthesis (Defila and Di Giulio, 2015). To maximise transparency and reproducibility of this thesis, the research structure and process will be documented, substantiated by methodological theory in this section.

2.1 Research Structure and Process

After an initial reading of the published ECOLOPES research articles (Weisser et al., 2023; Perini et al., 2021; Canepa et al., 2022; Selvan et al., 2023b,a), a coherent, concise and transparent method had to be developed to ensure arriving at a synthesis, which is not biased by domain-specific perspectives. The difficulty lay in covering knowledge from multiple fields in an exploratory way, where judgement about thematic inclusion must be developed iteratively. To ensure a holistic selection of literature to build this thesis on, the methodical approach needed to support a relatively quick evaluation of topic-specific literature, as well as being able to expand on field-specific topics for later synthesis. To reduce the risk of bias or incompleteness, inter- and transdisciplinary methods were used to structure the research process.

2.1.1 Inter- and Transdisciplinary Research

Interdisciplinary research on social-ecological dynamics within urban ecosystems can be traced back to the beginning of the 1990s (Andrade et al., 2021). Inter- and transdisciplinary research is commonly considered as a collaboration of either scholars of different academic disciplines or of academic scholars with researchers outside the academic realm (i.e. institutions, practitioners or citizens). Although this thesis cannot be considered as inter- or transdisciplinary in the understanding of a group of people with different backgrounds working together, the research questions address knowledge from different disciplines, with field specific methods, vocabulary and perspectives (Defila and Di Giulio, 2015). Knowledge integration is considered the core issue of inter- and transdisciplinary research, therefore addressing epistemic and methodical validity is a main concern (Defila and Di Giulio, 2015).

To this end, findings and approaches must be selected in terms of their contribution to the common answers, they must be reprocessed, related and brought together. The common result is the integrated knowledge produced in this process, the so-called *knowledge synthesis* (Defila and Di Giulio, 2015), where the outcome of an interdisciplinary research process is achieved by *integrative actions.* To this end, Defila and Di Giulio (2015) have developed an *inventory of synthesis* as a methodical analysis to identify what is needed, for what purpose, at which stage in the research progress, as well as to address the epistemic structure of the desired results of integration.

Benda et al. (2002) define five categories of knowledge structure in context of interdisciplinary research: 1) history and form of disciplinary knowledge, 2) spatial and temporal scales, 3) precision, 4) accuracy of predictions, 5) availability of data for construction, calibration and testing. This can also serve as categorization for possible problems, mismatches and biases when constructing interdisciplinary knowledge. The aim of defining knowledge structures at the onset of an interdisciplinary research process is to generate “solvable problems”, that is a research question that can be answered “within defined limits of precision and certainty” (Benda et al., 2002). Therefor it is vital to define the boundaries of a research project (Elsbach and van Knippenberg, 2020).

2.1.1.1 Boundary Judgement and Boundary Object

A first step in inter- and transdisciplinary research is the identification of involved disciplines. Although the thesis will put an emphasis on architectural and urban design, other disciplines are necessary to adopt a holistic perspective. As already mentioned, McPhearson et al. (2016) identified involved disciplines as seen in Section 1.3.2. Additionally, the assessment of ML methods in the context of EUA will involve at least computer science, data science, and to some degree statistics. With such a great number of disciplines to be considered, the research topic needs to be defined by conceptual or semantic focusing, which can be achieved by a boundary critique.

The declared aim of this paper is to build a conceptual framework for EUA in the context of multi-species urban design through a type of analytical classification, which is not relying on priori established reductionist measurements, but which is able to infer from a highly complex array of features of different patterns and processes in urban environments, regarding the design intent or inquiry about ecosystems. Frameworks are frequently used for fields of research, which are not well established. They are supposed to clarify the most important concepts and condense and organize them in a consistent way (Cadenasso et al., 2003). As language might be ambiguous, especially in interdisciplinary research fields, exclusion of terms might bias the result. From the preliminary readings I concluded to focus the research on biodiversity, opposed to the other discourses of sustainability and resilience, and ecosystem services, which are inherently anthropocentric. However, although this helps to establish a focus area, it is to be expected that the other topics regarding sustainability cannot be separated entirely.

Since a distinct discourse about ecological analysis within the discipline of architecture and urban design could not be identified in the preliminary literature search other than what has been outlined through the ECOLOPES project publications, the present thesis needs to establish a basic relationship of ecological and urbanistic or architectural knowledge. Since the two disciplines are from different knowledge domains, ecology within the applied natural sciences and architectural and urban design as applied technological and design disciplines, a common subject should be identified to facilitate knowledge integration (Defila and Di Giulio, 2015). Since architecture is strictly spatially related, a common notion of space will help to translate ecological concepts into architectural considerations. Although there are far more aspects to designing and planning within socio-ecological systems, such as human stakeholders, governance and policy, and design project considerations, to include such would go beyond the already broad scope of this thesis.

As the concept of urban ecological analysis or classification is just emerging in the literature, especially in combination with the rapidly evolving genre of artificial intelligence, it is difficult to set a proper scope for identifying the relevant literature a priori. Boundaries can also be expressed as a 'boundary object', a concept first introduced by Star and Griesemer (1989). A boundary object needs to be at the same time abstract enough to be shared among involved disciplines, yet at the same time specific enough for every discipline to address it in their specific terms (Defila and Di Giulio, 2015). Such an object can either be an established concept or term, or it can be a 'latent' object, which has not explicitly been expressed, although several disciplines have undertaken research addressing the concept (Defila and Di Giulio, 2015). Hence "ecological urbanistic analysis" seems to be a suitable boundary object to capture multi-species and biodiversity-inclusive design with special regard to complex and multi-variate dynamics in urban systems, using ML methods of classification and clustering, as it incorporates three general topics, yet through the combination they point into a specific direction, which can be addressed from different fields of research.

After the establishment of a boundary object, the research structure can be addressed in more detail. The "inventory of synthesis" framework developed by Defila and Di Giulio (2015) spans the interdisciplinary research process around the boundary object, as will be shown in the next part.

2.1.1.2 Inventory of Synthesis

Within the "inventory of synthesis" as presented by Defila and Di Giulio (2015), the synthesis consists of "building blocks" related to each other. These building blocks consist of "elements" which are empirical evidence, theories, concepts, etc. These elements can either be direct contributions from scientific fields or be the result of integration processes of other elements, but elements do not need to relate to each other (Defila and Di Giulio, 2015). This structure is the basis for identifying necessary steps and defining methods of integration to be applied. Methodological aspects as establishing common terminology, joint categorial system, or developing common approaches to describe the object of research are integrated results themselves, although not being part of the building blocks of synthesis (Defila and Di Giulio, 2015). The inventory is depicted as hierarchical flowchart, containing disciplines, elements, and building blocks as components and indication of integrative actions (Defila and Di Giulio, 2015). Figure 2.1 shows the established inventory of synthesis for this thesis, adopted from Defila and Di Giulio (2015) for structuring of the different, field specific research tasks.

Although Defila and Di Giulio (2015) note that the inventory of synthesis is not a temporal or procedural diagram, from this hierarchy of synthesis a fitting procedure needs to be developed, to uphold the established structure. This process will be described in the following, as a multi-staged synthesis procedure which adapts grounded theory and literature reviews as the central research methods.

2.1.2 Grounded Theory

Grounded Theory offers a conceptual approach, pursuing methodological development of 'theory building', 'systematic collections and analysis of data', and 'theories emerging from data'. For example, Charmaz (2006) defines grounded theory as "a method of conducting qualitative research that focuses on creating conceptual frameworks or theories through building inductive analysis form the data". Coding of data tries to capture emergent patterns and concepts. It is necessary to set these newly found codes into relation to concepts in existing research. *Open coding* describes the initial stage of data analysis, where concepts are extracted from literature as they emerge. *Axial coding* involves bringing the extracted concepts in relationship with each other, identifying more important concepts, and bringing them into order. *Selective coding* is the final step of analysis and aims at establishing an overarching category, that binds together all relevant concepts and expresses the argument of the thesis (Wolfswinkel et al., 2013).

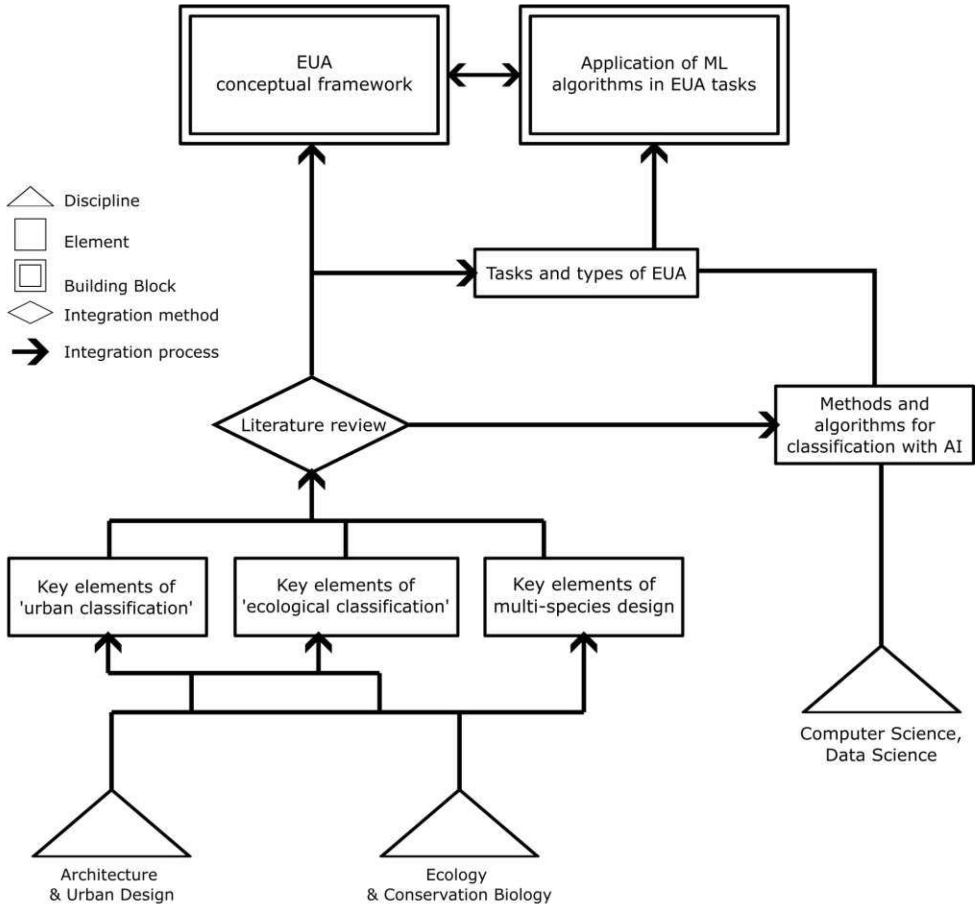

Figure 2.1: Inventory of Synthesis for Ecological Urbanistic Analysis through Machine Learning; adopted from Defila and Di Giulio (2015).

As work in grounded theory relies on data, a sophisticated selection of theoretical literature will be necessary to formulate this thesis. High quality theoretical sampling is considered the core requirement of grounded theory research (Timonen et al., 2018). Its purpose is the identification of emerging concepts within data, and to assess their underlying properties, dimensions and variations (Timonen et al., 2018). Through refinement of the theoretical focus and guiding research questions, the aim is to reach a "saturation" in data, i.e. no new meaningful information or insights can be gained (Timonen et al., 2018). Although the background literature of the ECOLOPES project is likely to contain many highly relevant documents, it cannot be assumed that the specific concepts of this thesis are fully encompassed within this set of literature or its references. For this reason, an auxiliary method for initial literature retrieval as 'data' to work on, is adopted to cope with the unknown theoretical extent.

The combination of literature reviews and grounded theory has been discussed as a disputed, but fruitful combination to engage scientific rigour in grounded theory (Wolfswinkel et al., 2013; Dunne, 2011). Hence, a literature review will be a central part of this thesis, serving three purposes with respect to an approach based on grounded theory: 1) identifying important theoretical literature to base an argument about the analysis of urban environments under a perspective of biodiversity and multi-species design, 2) supporting the first stage of coding, informing about underlying important concepts, 3) selecting a set of articles which inform about the possible applications of the stipulated framework. Relating to this logic as a core principle of grounded theory, the literature review supports multiple tasks in constructing knowledge.

2.1.3 Literature Review

Several techniques for exploration of existing literature have been well established: *term-related searches*, *snowballing*, *co-citation networks*, *citation-cited-by inferencing*, and others. Although the methods relating to author citation networks are good fits for evaluating research topics which are narrow and highly specialized, a weakly defined field of research might suffer from missing related strains of research which could contribute to the results. Through the general organization of research work in field and domain specific categories, a two-dimensional representation of interdisciplinary concepts would result in separated clusters. Therefore, literature search needs to address the shortcomings of field- and domain-specific organization, jargon, and concept creation. A possible logic to address this issue is shown in Figure 2.2.

Literature reviews have gained momentum over the last years in general scientific research, as well as in the design and planning field (e.g. Tyc et al. (2023); Ullah (2021); Vujovic et al. (2023)). This might be due to the rapidly growing body of literature, which demands increasing efforts of collection and synthetization. Accordingly, methodology on literature reviews and other structured methods of knowledge synthesis has grown substantially over the last decades. An overview and general distinction of review types can be drawn from many sources (Snyder, 2019). Post et al. (2020) define a review as "a study that analyses and synthesizes an existing body of literature by identifying, challenging, and advancing the building blocks of a theory through an examination of a body (or several bodies) of prior work". Further they attest that "review articles can connect research findings from various disparate sources in original ways so that a new perspective or phenomenon emerges (Post et al., 2020)." The objective of the literature review should determine the format of the review.

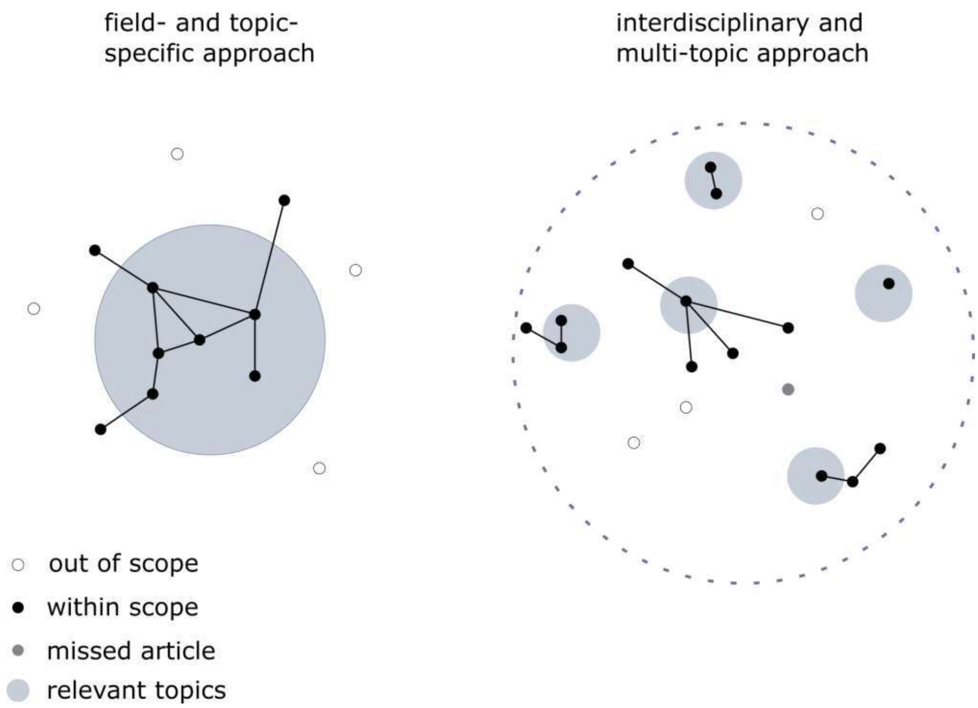

Figure 2.2: The possibly related literature concerning a research topic depends on the scope. Under the assumption that all possible literature is represented in two-dimensional space, narrowly defined research questions can be assessed by addressing a certain point. But broader, interdisciplinary, or not-well established topics might find meaningful information at many different points. However, with growing complexity of the research topic, a search approach which aims at multiple thematic domains may also miss relevant literature, if it is not well embedded in citation networks, or the jargon used highly differs from the dominant domains.

Owing to the intense application of this method, a plethora of review types has formed over the years. The systematic literature review is often referred to as the gold standard of literature reviews, because it defines clear guidelines to ensure validity for the knowledge synthesis. Although systematic reviews exist as a separate category, based on primary empirical studies (Whittemore and Knafl, 2005), for reasons of transparency, reproducibility, reducing the likelihood of bias and ensuring a comprehensive corpus of literature, adhering to a systematic approach should be used in any format of review (Booth et al., 2012). The *PRISMA statement* is a guideline for conducting systematic literature reviews, which offers a robust structure and checklist for systematic literature reviews (Liberati et al., 2009). To give this thesis a reasonable basis in terms of retrieval and assessment of existing literature from various domains of scientific research, literature will be searched adhering to systematic literature review guidelines. Although these guidelines will help to improve transparency and reproducibility of this thesis, the reporting format and synthesis of knowledge of systematic literature reviews are empirically oriented and not appropriate for this thesis. Such an approach will miss important input from other disciplines, describing similar epistemological processes, scientific methods or empirical observations with a different vocabular from different perspectives (Casali et al., 2022).

Integrative reviews are indicated for either mature topics which need a reconceptualization for fragmented literature, or emerging topics in need of an initial synthesis, even of disparate sources, to identify important concepts, topics and other discursive elements, as well as contradictions (Post et al., 2020). In the latter case, integrative reviews support in assessing a state-of-the-art, but also gaps in literature. Boundary conditions are typically broad for an emerging topic and the theoretical perspectives might be weakly connected to this point (Elsbach and van Knippenberg, 2020). Integrative reviews differ from systematic ones in some respects. Integrative reviews incorporate both empirical and theoretical literature (Hopia et al., 2016). For an integrative review it is not necessary to cover all relevant literature, although the search strategy should aim to do so (Elsbach and van Knippenberg, 2020). The lack of methodology in assessing literature is also a distinctive trait in comparison to conceptual or theory-building manuscripts, which also differ in their foundational narrowness and the specific concept or theory to be explained. To estimate the relevance of an integrative review its justification as well as its boundary conditions need to be established (Elsbach and van Knippenberg, 2020).

Although still a relatively vague concept, integrative reviews found ample application. Early articles about writing integrative reviews originated in health care (Whittemore and Knafl, 2005) and human resource planning (Torraco, 2005). Elsbach and van Knippenberg (2020) distinguish integrative reviews from other types of research (conceptual and theoretical manuscripts), which may also include a review of literature, as not having an a priori set argument. Insights and perspectives are produced by the review and its synthesis, rather than just reporting bibliometrics, or testing an a priori postulated hypothesis (Post et al., 2020). As Elsbach and van Knippenberg (2020) write: "In addition to this general definition, we also argue that the insights or perspectives offered arise from the review, rather than guide the review." This conception shows strong relatedness to the concept of grounded theory, which in turn supports the coupling of both methods for the synthesis of novel research topics.

Since it is not possible, nor intended to fully assess the complexity of all topics included in this thesis, the knowledge synthesis will be partly guided by grounded theory and partly by integrative literature review standards. Therefore, a term-related search, guided by literature review methodology was chosen, to define a set of vocabulary, which is describing a body of literature by conceptual characterizations, rather than specific keywords. The following section will supply a scaffolding for the conduction of the literature search, coding and knowledge synthesis.

2.2 Operationalization of the Research Process

To review the latent topic of "ecological urbanistic analysis" on a broader scale, a critical mass of literature, to establish a first scaffolding for a conceptual framework, needs to be reviewed. As mentioned before, to approach the concept from an interdisciplinary perspective, several scientific domains need to be considered for delivering important contributions. For this purpose, an expansive literature review will be undertaken to collect necessary evidence to answer the research questions. To approach this strongly interdisciplinary topic with sufficient openness, the literature review will consist of the following steps:

1. An initial scoping search of "urban classification" and "ecological classification" together with an evaluation of the ECOLOPES literature to extract the most important concepts and terms for search query construction. "Classification" is chosen over "analysis" for this first search iteration, as it represents a narrower understanding of the analytical aspects of interest in combination with ML methods;
2. Building a more informed, broad search term, to cover interdependencies with other concepts and disciplines;
3. An integrative Literature review of theoretical background papers;
4. Finding gaps in the body of literature and pick conceptual supplements from citation tracking or topic specific auxiliary papers;
5. Conceptualization of EUA as framework with related tasks for machine learning methods;
6. A literature search for ML methods in an urban and ecological context;
7. A systematic review of ML-related papers within the retrieved urban ecological corpus;
8. Synthesizing the EUA framework, reviewed applications and ML theory into potentials and challenges for promoting biodiversity and multi-species design through the integration of eco-complexity.

2.2.1 Definition of the Problem Scope

Due to the broad and interdisciplinary scope of the research questions, an extended form of literature search has been developed, as sketched out above. Because the term of "ecological urban analysis" has not been established by the scientific literature under research in this thesis (see Section 1.3.1), to establish a relevant search basis for the main review part, the terms 'urban classification' and 'ecological classification' will be queried in scientific databases to extract the most important terms related to these topics. 'Classification' is chosen over 'analysis' as in the initial searches for 'classification' showed a narrower thematic boundary which would better fit the scope of this thesis. Although techniques exist to generate search queries from minimal information, such as partial research questions, tests have shown, that a larger set of information, describing the scope of a research project have improved the performance of such query enrichment techniques (Badami et al., 2022).

To build a rich search query from the relevant ECOLOPES papers, citation tracked references and the search results for "urban classification" and "ecological classification", the most important terms will be extracted through keyword extraction and text mining methods. All the extracted terms will be used to construct a search string encompassing all identified relevant topics and concepts. In a next step, the resulting query will be taken as input for an integrative review, where the most relevant sources will be identified, screened, and analysed for synthesis.

2.2.2 Use of Literature Databases

For the individual search queries, scientific literature databases will be used as follows:

- Initial searches for "urban ecological classification" and analogous terms (see Section 1.3 was conducted on *Scopus* (Elsevier, s.a.b) and *WOS* (Clarivate, s.a.a) to establish the maturity of the research topic.
- A term extraction search for "urban classification" and "ecological classification" will be conducted using Scopus, WOS, and CORE UK (CORE, s.a.).
- The main literature search to establish a body for synthesis of EUA will use Scopus and WOS to ensure a minimum quality standard, while additionally limiting results to journal articles, conference papers, and books. Books, although usually not considered in reviews can help to gain overviews of different domain-specific topics, while conference papers support the state-of-the art evaluation.
- Additional literature searches for ML topics will make use of Scopus or WOS databases.

2.2.3 Literature Search

As it cannot be expected for the literature search to cover all relevant information within one search term, additional literature retrieval techniques will be needed, based on the main corpus of literature. Citation tracking was found to be increasing studies relevancy as a standalone or supplementary search method (Hirt et al., 2023). The idea behind this method is, that literature relevant for the research topic is likely to be citing or be cited by other literature relevant to the topic. Co-citing papers are papers also citing references cited by the *key paper* and co-cited papers are sharing citing papers with the key paper (Belter, 2016). This search strategy is either limited by a defined number of tracking levels, or by not finding any more relevant papers. A major advantage to improve the effectiveness and efficiency of literature research through citation tracking is the selection of papers, which are identified to be most accurately addressing the research topic, and/or have a high citation index, thus increasing the possibility to be influential and important sources in a field.

Zwakman et al. (2018) have categorized *golden bullets* as literature that aligns with the scope of the research and should be an essential part of the search result. They are used for *feature extraction* but also to validate the search result, as an inclusion of golden bullets is an indicator for a suitable search strategy or query. As the core topics of this thesis must be developed iteratively, a possible reference body of literature is provided by the 'ECOLOPES' references which will be taken as a measure of validity of the search results.

2.2.3.1 Boolean Search Query Construction

Search queries are one crucial element for finding an accurate set of literature to perform a valid review, covering all relevant results with the least amount of screening and processing (Badami et al., 2022). This is especially true if the authors knowledge on the review topic, or at least parts of it, are limited at the beginning of the review. While there has been research about query building and refinement, this research is domain-specific, and methods of automation rely on discipline-specific training data or thesauri and most require an initial query formulation to work with (Badami et al., 2022).

With the emergence of new concepts in science, terms are often borrowed from other disciplines, new terms are coined simultaneously, and several terms are used interchangeably (Badami et al., 2022; Garousi and Felderer, 2017). Language for this emerging and multidisciplinary topic is still forming and consolidating, i.e. a phenomenon can be addressed by multiple terms, also the vocabulary may not always be precise. During the preliminary research for this thesis, it has shown that a term can occur in completely different contexts, and that ambiguous terms exist, like "architecture”, used in computer science to describe the concept and structuring of software, or “environment” used in several disciplines, which can negatively impact the literature search.

An initial query, which can be partially defined by research questions or a collection of relevant abstracts, can be used to build a seed for an informed search query (Badami et al., 2022). This thesis will take selected literature published or referenced by the ECOLOPES project team, as well as a preparatory search of the topics of “ecological classification” and “urban classification” in scientific databases, to describe the fields in question and extract important search terms. Similar approaches to keyword building have been undertaken (Vujovic et al., 2023). Grames et al. (2019) developed the tool *litsearchr* (Grames, s.a.) implemented in *R* (R Foundation, s.a.), to semi-automatically generate search queries from an initial *naive search*, generating a seed corpus. This tool will be used to extract possible search terms from the three initial corpora as described above.

2.2.3.2 Retrieval of Relevant Literature

Once the literature is retrieved, several steps of processing will be necessary. Different sets in database-related formats will need to be standardized for deduplication, which will in turn be done in *R* with the tool *ASySD* (Hair and Wallrich, s.a.) developed by Hair et al. (2023). Evaluating the retrieved literature will reveal important concepts, which are not satisfyingly addressed, which can in turn be retrieved through citation tracking, and specified literature searches. Common methods to incrementally add relevant literature are *cherry-picking* and *pearl growing* (Zwakman et al., 2018), which will be used to supplement the main literature search.

2.2.4 Knowledge Synthesis

Many scholars refer to knowledge synthesis methods as a holistic methodical approach to knowledge generation. In this sense, knowledge synthesis methods commonly described are all forms of literature reviews, meta-synthesis, mapping approaches, etc. (Whittemore et al., 2014). There is an ambiguity in the conception of synthesis, as many authors of literature review guidelines are referring to 'synthesis' as the last step of a literature review (Hopia et al., 2016; Whittemore and Knafl, 2005). Garritty et al. (2019) define knowledge synthesis as scientific studies derived from primary sources. Using reproducible and reliable methods, bias is minimized, and the validity of the summarization of the data and following conclusions is improved.

Elsbach and van Knippenberg (2020) name *critical analysis* and *creative synthesis* as important factors of impactful literature reviews. Critical analysis is set to identify "themes, patterns, relationships and gaps", whereas creative synthesis aims at integrating insights from analysis into existing frameworks and formulating new perspectives on a topic (Elsbach and van Knippenberg, 2020). They further distinguish the development of an integrative conceptual framework in contrast to a theoretical framework. Conceptual frameworks aim to identify processes of interaction among components of a concept, which are in dynamic exchange at multiple levels of analysis (Elsbach and van Knippenberg, 2020).

Rossini and Porter (1979) defined four types of knowledge integration. For this thesis the only reasonable approach is the 'modelling' approach, since it can be conducted with one researcher, and it reduces bias by referring to an established framework of knowledge or is building the synthesis on a systematic way of synthesis. There exists a plethora of terms regarding the types of representation of knowledge, which are similar. Conceptual models have been used as a method to consolidate fragmented research into holistic representations of elements and relationships in complex systems (Battesini et al., 2021).

Conceptual frameworks are apt to depict the current state of knowledge, identify research gaps and outline a methodical foundation for a research project (Varpio et al., 2020). Ullah (2021) proposed a method for developing conceptual frameworks through literature reviews, based on the PRISMA statement. Once the relevant literature has been processed by *basic analysis* the reviewer must manually assess the *critical success factors (CSF)* to formulate the conceptual framework. Then a clustering of the proposed CSF can be undertaken to form layers of the conceptual framework. This clustering can either be "natural", expert-based or literature-based (Ullah, 2021). For this thesis, the foundational theory will be used to construct such a clustering. Once the clusters are established, the conceptual framework can be formulated by subsuming the CSF to the clusters of the framework (Ullah, 2021).

The knowledge synthesis will be two-fold: 1) by integratively reviewing theory found within the corpus of literature, and semi-systematically reviewing the applications of ML methods addressing the topic of EUA. This is supposed to complement a rather qualitative approach from grounded theory with case studies as 'proof of concept', and underline the importance of future research.

2.3 Conduction of Literature Search and Review

The literature search and review has been conducted in several, iterative steps from November 2023 until April 2024. After an initial reading of the research articles, published by the ECOLOPES project team, four articles addressing the thematic complex were chosen for citation tracking, by the description of multi-species buildings design within the urban context at a holistic level (Perini et al., 2021; Canepa et al., 2022; Selvan et al., 2023a; Weisser et al., 2023). All references were screened by title for inclusion into the next step. Inclusion criteria were: 1) the article addresses aspects of urban systems, urban ecology, biodiversity or a related topic, or 2) addresses urban modelling, analysis, classification or clustering. Pooled with the ECOLOPES articles, the search resulted in a set of 124 articles with title and, where eligible, abstract.

The first stage of the database search, the so-called *naive search* (Grames et al., 2019), was conducted on December 17th, 2023. The search resulted for “urban classification”/“ecological classification” as follows: 1) 318/573 (Scopus; Title, Abstract, Keywords), 2) 197/422 (WOS; Topic), and 3) 2,470/3,294 (Core UK; API query) records. Which sum up per search term to: 1) “urban classification” 2,985, and 2) “ecological classification” 4,289 records.

The records were processed in following steps (record counts refer to the search terms “urban classification”/“ecological classification” in total):

1. Retrieval of topic information for the Scopus dataset;
2. Unification of record format with application of processing Ids to facilitate processing;
3. Merging of datasets from all three databases: 2,988/4,289 records remaining;
4. Exclusion of records not containing author, title, or abstract information: 2,500/3,584 records remaining;
5. Language detection with *Spacy* language detector module in Python. The manual control of the result revealed some missed records which were removed manually. Exclusion of articles other than English language in abstracts (also bilingual abstracts) based on detected language and database language information: 2,339/2,898 remaining records;
6. Preprocessing of the fields author, title, and abstract to remove special characters;
7. Conversion of column names for deduplication via *ASysD* in R;
8. Deduplication with manual inspection of possible duplicates: 2,021/2,448 records remaining;
9. Separation of records into datasets with and without topics;
10. Classification of records without topics for containment of anthropocentric topics in Python using *pytorch* (Linux Foundation, s.a.) and *huggingface* (Hugging Face Inc., s.a.) libraries;

11. Exclusion of records containing anthropocentric topics: 637/2,195 records remaining;
12. Extraction of possible search terms from the remaining records via *litsearchr* tool in *R*.

An initial screening of the retrieved records confirmed the assumptions about gaps of "urban classification" being discussed in a very narrow sense of almost exclusively land use and land cover studies, and "ecological classification" focusing on native and rural ecosystems, and organisms within them. While the WOS dataset already contained topic information, similar information for the Scopus dataset had to be retrieved via API separately. This topic information revealed that the dataset for "urban classification" showed many records related to medical topics. A keyword co-occurrence analysis conducted with *VOS Viewer* (Leiden University, s.a.) revealed a big cluster of keywords revolving around topics like "medicine", "health", and "social sciences". As the approach to the research topic should be non-anthropocentric, this could potentially skew the results of keyword analysis and compositions towards such topics.

As the Core UK records, as well as some of the other records did not contain any topic information a method had to be implemented to keep a rich set of records for search term development while sorting out unwanted records to tighten the thematic boundary. To use ML for this task, the records containing topic information were normalized and put into a consistent format. This was done by setting up a matrix containing all occurring topics in the original format and summarizing them under a new and simplified categorization. As there is no common understanding on how to categorize scientific fields, the categorization was informed by the ASJC (Elsevier, s.a.a), nature (Springer Nature Limited, s.a.), JACS (Higher Education Statistics Agency, s.a.), and WOS topic systematics (Clarivate, s.a.b).

The implementation of the ML method for topic classification took considerably more time than expected. All modelling was done with Python, and the libraries *sci-kit multilearn* (Szymański, s.a.), *pytorch* and *huggingface*. A first approach towards multi-label classification of the abstract texts using *sci-kit multilearn* and a pre-trained *sciBERT* (Beltagy et al., 2019) model did not show satisfactory results. Since the dataset needed balancing, under-sampling led to some topics being represented by very few datapoints, which could account for the weak classifier performance.

In a second attempt *sciBERT* models were fine-tuned with *pyTorch* and *huggingface* libraries, to identify four topic groups, i.e. 'medi - medicine', 'psyc - psychology', 'heal - health' and 'soci - social sciences', from the records' abstracts, through binary classification. This approach carried the drawback that every record could potentially contain multiple topics, and a binary classification was only able to determine the inclusion of a specific topic. But since the aim was only to minimize misleading terms for the final search term construction, this ambiguity was deemed acceptable. After multiple iterations of training with different hyperparameters and training data splits, the best performing models were evaluated by accuracy and f1-scores.

The models did not reach an ideal classification score for all topics, and thus the *ROC curve* of every model was evaluated and the decision threshold for every topic was adapted accordingly to optimize the performance. The high variance of the same base model on different topics is likely due to the varying scope and broadness, as well as the specificity of language and terms used in the respective research field.

After classification, all records containing the anthropocentric topics were dismissed, which reduced the record set to: 1) “urban classification” 637, and 2) “ecological classification” 2,195 records. A screening of the remaining records made apparent, that classification in the sense of EUA stipulated in Section 1.3 is not always referred to explicitly as classification, but many analytical processes which imply some form of classification, such as ‘analysis’ or ‘mapping’ can refer to some process of classification. This implied that the next search step should address a broader notion of analytical tasks, in order not to limit the result to narrow mechanistic and deterministic classification approaches such as land use and land cover classifications.

2.3.1 Creating the Main Search String

At this point, the *litsearchr* tool in *R*, developed by Grames et al. (2019), was used for search term suggestion. Three sets of title and abstracts of “urban classification”, “ecological classification” and the ECOLOPES references were separately evaluated. Parameters were set to a minimum gram size of 1, and inclusion in at least 5 percent of all records to be significant. This approach was chosen after an initial attempt to create a specific set with relatively few search terms, within a high threshold of inclusion, to identify common vocabulary, did not provide meaningful results. Instead, from a very generic set of multiple search terms, the scope of the research topic should be formulated. After all suggested search terms were pooled, they were analysed for query construction. Since the litsearchr tool is not performing word stemming, many related word forms occurred. Therefore, all interesting and possibly relevant 1-grams were extracted as noun form. Some of the suggested terms were deliberately omitted because they seemed to address a scale (“world”, “geo*”) or context (“benthic”, “aquatic”, “marine”) exceeding the scales to be considered in research questions.

The challenge in establishing a search term with a big set of individual non-specific terms was to group the terms by concept, to reduce the conditions to a suitable minimum, and prevent exclusion from potentially interesting records. This approach leaves a lot of ambiguity towards the meaning of words. This was compensated by the abundance of ambiguous words, grouped into large conceptual groups. The strategy chosen was to put more specific terms into smaller groups, representing higher levels of search terms, with a high confidence of topic relevance, and less specific terms in larger groups to grant more leeway for record matching. The concepts were logically grouped to address the three core components of the initial search: 1) the urban sphere 2) ecology and biodiversity, and 3) classification or any related form of analysis. The resulting structure was a complex of 5 hierarchical levels with a total of 11 conceptual groups.

The resulting levels were then iteratively aggregated in querying the *Scopus* and *WOS* databases. The results showed that the first two levels, with a higher certainty of topic relevance reduced the results significantly, whereas at the 3rd level the decrease in resulting records flattened dramatically. This suggested that the more obscure and uncertain levels of search term concepts did not alter the results substantially, but as the resulting record number was still high, all five levels of search term concepts were chosen for the final query.

The final search query can be seen in Appendix A. Figure 2.3 shows the conceptual grouping of the individual search terms.

2.3.2 Literature retrieval for theory on urban and ecological theory on analysis

In the second review step only the databases of *Scopus* and *Web of Science* were queried, to ensure literature quality and to facilitate the data processing. The search yielded 25,422 records in *Scopus* and 20,691 records in *WOS* respectively. From this result only articles, reviews, conference papers, and books or book sections in English language were further processed.

For this review step the following record processing was applied (record counts refer to the databases Scopus/WOS):

1. Search query of databases: 25,422/20,691 records;
2. Retrieval of topic information for Scopus dataset;
3. Exclusion of records other than English language, from dataset language info: 23,269/20,389 records remaining;
4. Unification of record format;
5. Pooling of records: 43,658 records remaining;
6. Deduplication via the AsySD tool in R: 29,339 records remaining;
7. Exclusion of records not containing information on author, title, or abstract, plus exclusion of records other than journal articles, books, book chapters, conference papers, or conference proceedings, plus exclusion of anthropocentric topics: 24,627 records remaining.

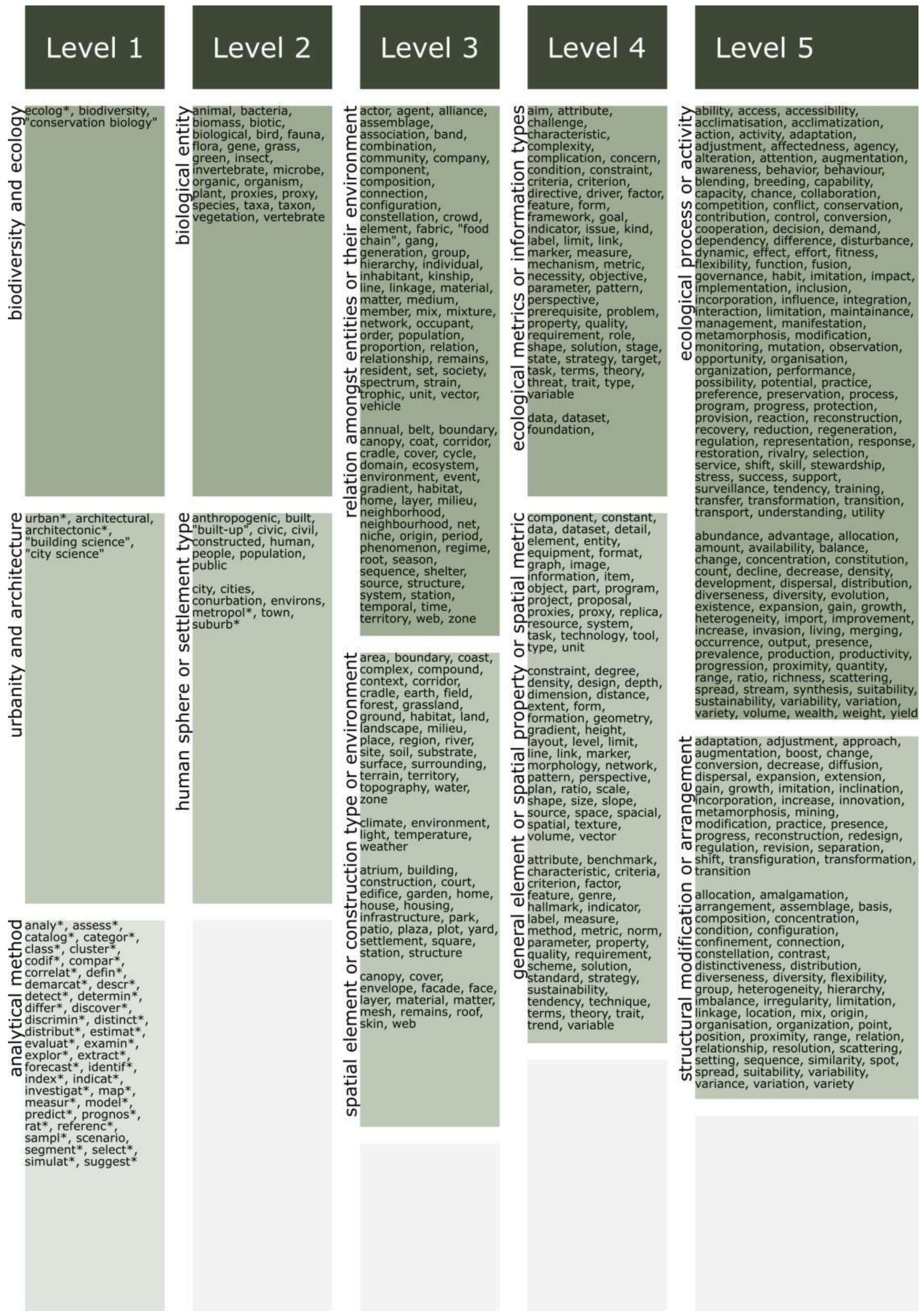

Figure 2.3: Final search terms grouped by concept and term level.

First, relevant articles were tried to be identified through a topic modelling approach in *Python* and the *GenSim* library (Řehůřek, s.a.). Topic clusters, depending on the cluster number, either shared a lot of terms such as 'green infrastructure' or 'urban green space', making a meaningful distinction impossible, or were defined by very specific terms, which could not be interpreted into more general topics. In a second attempt for the identification of highly relevant papers, a document relevance score, i.e. a document ranking approach. was implemented to evaluate the corpus by average co-authors citations, adjusted document citations, and inclusion of 4 term categories in record titles. The score was calculated from three indicators: 1) weighted average co-authors citations, 2) weighted document citations, and 3) topic specificity.

Weighted average co-authors citations were calculated first by evaluating for every individual author, for every document he or she authored or co-authored, the citations were evaluated and weighted by their actuality as

$$wght\_ac = \frac{\sum doc\_c + doc\_c * (1 - \frac{e^{(doc\_age/10)}}{e^{(pub\_span/10)}})}{\sum auth\_d} \tag{2.1}$$

$$wght\_dc = doc\_c + doc\_c * (1 - \frac{e^{(doc\_age/10)}}{e^{(pub\_span/10)}}) \tag{2.2}$$

wght_ac = weighted author citations wght_dc = weighted document citations doc_c = document citations doc_age = document age in years pub_span = time span from oldest to newest publication year in the dataset in years auth_d = number of documents in the record set from same author

and topic specificity was calculated by record title inclusion of terms of four categories, where for any partial occurrence within a category a value of 1 was added to the topic specificity score, which was then averaged by all categories:

1. architecture or urban science = 'architect', 'urban design', 'urban plan', 'building science', 'landscape design';
2. biodiversity or spatial ecology = 'biodiversity', 'biological diversity', 'conservation biology', 'conservation ecology', 'landscape ecology', 'urban ecology';
3. document type describing concepts = 'review', 'approach', 'guide', 'overview', 'meta-analysis', 'framework', 'synthesis';
4. document addressing spatial or geometric aspects = 'space', 'spatial', 'morphology', 'form', 'geometry', 'typology'.

$$norm\_ts = \frac{ts}{4} \tag{2.3}$$

ts = topic specificity score norm_ts = normalized topic specificity of a document by title

All parameters were normalized and summed. The result was then again normalized to achieve a score between 0 and 1. From the resulting relevance ranking, the top 500 documents (ca. 2 percent of the corpus) were manually screened by titles and abstracts for inclusion in read through evaluation.

As the variety of topics was still substantial, inclusion and exclusion criteria were formulated as:

- Included were documents which addressed several aspects of urban scale, biodiversity or ecosystems, machine learning theory; use of machine learning for result generation, marine/river related if spatial relation expressed;
- Excluded were documents with topics of soundscapes, air pollution, human perception, wildfire, papers whole city or regional level, agriculture if not related to urban areas/buildings, building typologies or indoor evaluation, habitat related to humans, genetics or diseases, reviews of case studies; case study, mapping or feature extraction, economic, use of machine learning for literature retrieval, heat islands, air pollution, cultural landscapes, green perception, land use land cover change, smart city, 'green urban spaces' or 'green infrastructure', urban-rural comparisons, human health and wellbeing related, focused on a certain species or taxa, sustainability, energy.

89 articles were retrieved for read through evaluation, as far as accessible through the institutional account of the university or open access resources, full texts were screened, and potentially contributing articles were read through. After read through evaluation 35 documents were considered as a thematic seed covering multiple contributing topics. Further literature was then identified by citation tracking as mentioned in Section 2.3.2.

2.3.3 Machine Learning Theory in the Context of EUA

To substantiate the review about ML applications, ML specific articles were queried in a separate, simplified approach. The two databases 'Scopus' and 'WOS' were queried for: "("artificial intelligence" OR "deep learning" OR "learning algorithm" OR "learning system" OR "machine learning" OR "machine-learning" OR "neural net*" OR "neural-net*" OR "reinforcement learning" OR "*supervised algorithm" OR "*supervised learning") AND (overview OR guidelines OR guide OR review OR introduction OR "textbook" OR "text book" OR reader OR summary OR synthesis) AND (biolog* OR ecolog* OR architect* OR cities OR city OR urban OR geograph*)". From the resulting set, suitable articles were chosen for screening by title. After the screening, articles for establishing a theoretical background on ML in the context of urban ecology were chosen and evaluated. This set was later enhanced by review articles addressing ML in spatial and temporal contexts.

2.3.4 Machine Learning Applications for EUA

For the scoping review part concerning the application of ML methods in the context of EUA, a subset of the main literature corpus (24,627 records) was created, by filtering documents containing ML related key terms. Those key terms were collected from another 'Scopus' search with the query string "machine learning" and from the related keywords provided by the database. Although some articles do not explicitly mention "machine learning" or related high-level concepts in their titles or abstracts, but do make use of ML methods, those records were dismissed because it increased the chance to obtain meaningful records which consciously used machine learning approaches and describe them methodically. This was viewed as being beneficial, as the review is not intended as being comprehensive, but rather explorative. Also, manual screening effort would be reduced. To generate a set of articles for review the dataset was queried for the containment of one of the following terms in title or abstract: 'artificial intelligence, deep learning, learning algorithm, learning system, machine learning, machine-learning, neural net, neural-net, reinforcement learning, supervised algorithm, supervised learning'. The resulting ML-related subset was then screened for inclusion, if documents addressed one or several of the identified tasks in EUA from the literature for the EUA literature. After screening and read-through evaluation, 34 articles were reviewed.

For the reporting format of the ecological applications of ML in urban areas, the papers are analysed and organized by key characteristics to receive an overview of what has so far been tried to accomplish. During the assessment of the ML theory reviews, several studies offered a set of variables, from which the best fitting ones were assembled for this review (Rubbens et al., 2023; Stupariu et al., 2022). The chosen characteristics are 1) authors 2) publishing year 3) title 4) location 5) data used 6) data processing, 7) features used, 8) study purpose, 9) ML tasks, 10) ML algorithms, 11) evaluation of ML performance, 12) contribution to EUA, and 13) spatio-temporal awareness.

2.3.5 Critical Appraisal of the Reviewed Documents and the Research Concept

From the preliminary results of the ecological theoretical literature, the two sets of records regarding "urban classification" and "ecological classification" from the previous screening step (Scopus and WOS databases only) were screened for the term "ecolog* OR ecosystem OR biodivers* OR conservat*" for the "urban classification" and "urban OR cities OR city OR architect* OR building" for the "ecological classification" set respectively. Only few papers could be identified fitting the scope of the thesis, updated by the screening of the search results from Section 2.3. This strengthened the prior result that the individual search terms did not deliver meaningful results for the aims of this thesis.

One possible evaluation of the literature body received through Section 2.3 was to check the containment of the ECOLOPES papers and their references. In sum 35 articles of a total of 95 articles from this reference body listed on Scopus were retrieved through the search query. This corresponds to approximately one third of all selected references, which indicates a good query formulation, since not all reference articles were central to the topic of the thesis.

One factor that was not considered in the search query formulation, and a drawback of the conceptual groups concept, was that 'ecosystem' was not grouped in the top-level search. This was done to keep the focus on biodiversity and multi-species results. However, it later turned out that the two concepts are interconnected, and no sharp boundary is drawn within the ecological literature. This was compensated by citation tracking within the theoretical literature, resulting in additional research effort. As the review of ML applications aims at identifying case studies addressing biodiversity and multi-species research, the omittance of ecosystems in the top-level concept group keeps validity.

The systematic evaluation of the main dataset retrieved in Section 2.3 for this thesis did not yield many results to underpin the applications of EUA through ML methods. Therefore, a small exemplary selection of articles, which came up during the writing of this thesis, were chosen to supplement the results, and act as a 'proof of concept' for the framework.

2.3.6 Structuring of the Knowledge Integration

The following section presents the results of the reviewed literature. As described in 2.2, this process was done iteratively by extracting key concepts of ecology and biodiversity in the context of urban design and planning and is integratively presented to promote readability and to arrive at a conceptual framework for EUA.

A conceptual framework will require the identification of critical success factors or central theoretical concepts (Ullah, 2021). These will be abstracted in the following by:

1. Presenting bridging concepts between urban design and landscape ecology, to find a common spatial representation;
2. Then important principles of biodiversity applicable in multi-species urban design will be analysed;
3. The identified high-level principles of ecosystems will be addressed by analytical considerations;
4. From the theoretical background and the analytical methodologies, a unifying conceptual framework for EUA will be developed.

3.

Common Ground for Ecological Urbanism and Urban Ecology

To develop an understanding how ecological and spatial concepts relate in urban areas, it is crucial to understand the concepts of cities as artificially altered ecosystems (McPhearson et al., 2016). As mentioned in Section 1.1.3, this includes an understanding going beyond green infrastructures and natural remnants within cities, towards the inclusion of artifacts and human activities as essential constituents of ecological systems (Wu, 2014).

Recent research in urban biodiversity has changed the perspective on cities and their potential contribution to biological conservation. Not only that cities can in fact offer valuable habitats to a number of species, counterintuitively they can even support threatened ones (Ives et al., 2016). Urban areas are usually a highly heterogeneous patchwork of buildings, impervious surfaces, and 'natural' remnants. While historically the 'natural' or 'green' urban spaces have received most attention towards ecological research, it has been acknowledged in recent years, that urban environments may resemble natural ones in key aspects (Young et al., 2009). Thus, cities generally increase the availability of heterogeneous habitats (Spotswood et al., 2021), where even artificial habitat patches are able to replicate structural and functional aspects, allowing for the utilization by various species (Farinha-Marques et al., 2017). It has also been found that cities might even affect evolutionary processes (Diamond and Martin, 2021).

It is of interest to define cities as ecosystems to understand the assumptions which must be made to translate this concept from natural systems to social-ecological ones. After Tansley (1935), ecosystems can be generally defined as an "organism-complex and all the physical factors forming the environment of the biome, which have inherent structure, processes and ways of functioning." As any system is also defined through its boundaries, it is important to notice that urban ecosystems are open systems with dynamic boundaries and are in constant exchange with their fringe environments (Chen et al., 2014).

There is potential to design analogues to natural ecosystems intentionally for conservation purposes, but to do so, knowledge about the characteristics of urban environments and their impact on biodiversity processes is of key importance. Additionally, to the design and planning, the need for an ongoing conservation management of newly created biotopes within the existing urban fabric, must be of primary concern (Farinha-Marques et al., 2017). There is already ample evidence, that cities: 1) may provide highly specialized, complex and even unique habitats, which do not or only rarely occur in natural landscapes with small patch sizes, and 2) may, although pollution is a major topic, protect certain species from harmful chemicals, which are commonly found in semi-natural (e.g. agricultural) landscapes and might even transgress into natural ones (Casiker et al., 2021; Gentili et al., 2024). But while these findings imply that cities can be an important factor for biological conservation in the future, it must be considered that many species are adversely affected by urban environments.

Heymans et al. (2019) conducted a literature review about ecological urban planning and design. The authors provide the concept of *'urban consonance'*, which is integrating the identified streams of research from either the lens of ecological sustainability or spatiality into a holistic framework. The six main themes are 1) ecosystem services, 2) socio-ecological systems, 3) resilience, 4) biodiversity, 5) landscape, 6) green infrastructure. Additionally, 'integration and holism' is a superset characteristic to reach urban consonance. This conceptualization clarifies the different streams of research recently conducted, and it hints to which ones are potentially supporting a spatially explicit analysis, suited for application in multi-species architectural and urban design.

Although the importance of holistic concepts for urban ecological research must be acknowledged, and the broader topic of EUA needs to account for that, this thesis is limited to aspects concerning multi-species urban design. Hence, only the three meta-concepts of biodiversity, landscape and social-ecological systems will be addressed explicitly, as ecosystem services and green infrastructure are generally anthropocentric perspectives, and resilience is considered an important meta-aspect, which goes beyond the aims of this thesis.

3.1 Urban Ecology and Ecological Urbanism - A Spatial Divide

Only about 20 years ago researchers started investigating the relationship between patterns of urban development and their influence on ecological processes (Alberti, 2005). A fundamental problem, why ecological studies rarely find application in urban design and planning, are the different notions of spatiality within and between the disciplines (Alberti, 2005; Cadenasso et al., 2006b). Ecologists commonly regard space as one variable of many, influencing dynamic processes, cycles and behaviour of systems or organisms. Customary spatial representation within ecological studies, especially the urban-rural gradient and coarse landscape classifications, have been criticised as bearing little to no information about heterogeneity in spatial configuration and physical features of urban habitats, making interpretations about their contributions to biodiversity difficult, next to impossible (Beninde et al., 2015).

Owing to the different schools of ecology, spatial characteristics have played a minor role in urban ecological studies, with reduced and simplified consideration of urban structures (Wu, 2014; Kattel et al., 2013). Spatial features have long been represented by aggregate measures, such as demographics or built-up surface area (Alberti, 2005), which gave little input for architects and urban designers to work on. Therefore, spatiality is mostly expressed in an abstract or aggregate manner, either in form of two-dimensional maps with discrete grid units, representing measured or interpolated datapoints, or gradients as continuous graphs from the urban centre to the rural outskirts (Marcus et al., 2019a). Although adequate for describing certain ecological phenomena in a general and abstracted manner, gradients do not take into account spatial distributions (Alberti, 2005).

On the other hand, architects and designers tend to lack ecological knowledge, leading to biased decisions and choices, favouring single species for aesthetic reasons, leading to hyper-specific solutions, which neglect the complexity of ecosystem processes (Grobman et al., 2023). Designers and planners work mainly through qualitative assessment of space as main determinant of socio-ecological processes. Hence, architecture and urban design are deeply rooted in Euclidean three-dimensional space. Thus, it is necessary to establish a framework which is incorporating crucial ecological principles and describing their manifestation in spatially explicit urban terms. To facilitate an integration, a common spatial representation allowing for the translation of concepts is needed.

3.1.1 A Common Language for Urban Ecology and Ecological Urbanism

This unveils a core problem of the disciplinary evaluation methods. While ecology normally is focusing on interaction of system components, design and planning disciplines are focusing on the spatial configuration of real-world entities and their functional implications. Although there is high similarity in means of description, design and planning disciplines and ecology are constituting two distinct ontologies of geomorphological representation (Marcus et al., 2019a).

To arrive at a crossroads of urban and landscape ecology with architecture and urban design, a characterization of physical entities can be categorized into means and objects of description (Marcus et al., 2019a). Means of description have usually at least some level of abstraction, allowing to represent different entities with a finite set of elements. In architecture and urban design, the most common means of representation are geometric primitives such as points, lines or polygons in two-dimensional space. While these tools grant the possibility to refer to spatial entities of various domains, it also gives ambiguity to their meaning (Marcus et al., 2019a).

Proposed fine-scale classifications, which consider individual elements of landscapes, still do not reflect their explicit configuration in space (Cadenasso et al., 2007). Further, traditional ecology regards man-made elements as irrelevant or a type of 'negative' to vegetational structures, leading to their exclusion from consideration (Wu, 2014). But this is not only neglecting all humans as biotic entities and animal species, it also denies the interfaces of human habitats with broader biotic communities. Especially in urban areas, these interfaces need intense research and consideration (Apfelbeck et al., 2020). While in natural conditions most biotopes are connecting smoothly or are gradually shifting, the threshold of human and mixed species biotopes is spatially condensed to liminal spaces and hard borders with high potential for conflict (Dramstad et al., 1996).

Landscape ecology is providing a vital link for human-nature interactions, where the concept of ‘landscape’ offers a common reference system, able to address ecological processes in a spatially explicit manner (Heymans et al., 2019; Kattel et al., 2013). While landscape ecology has been found to contribute to ecological pattern-process understanding through describing landscape patterns, it has also been criticised as having little impact on planning and decision making (Nassauer and Opdam, 2008). While traditional ecology regarded ‘landscape’ as something rather constant with slow rates of change, cities and their surrounding agricultural landscapes are changing at much higher intervals.

Urban areas are changing at an unprecedented speed in their morphology, surface cover, flow of energy and nutrients, physical phenomena (such as wind etc.). Still, we can identify cities as structurally and spatially different form other types of landscape and other types of settlements, with relatively sharp borders, boundaries, and fringes along a rural-urban gradient (McKinney, 2006). But the traditional notion of an opposite of urban areas and land considered as ‘rural’ is not to be set equal with ‘natural’ or not human dominated land. The function of cities is highly dependent on intensive agricultural use of surrounding or even remote areas, thus being highly modified landscapes for urban use (Wu and David, 2002).

Although the notion of a city as ecosystem demands attention for spatial qualities of patch structures and their importance for species survival, the effects of and within built-up space, the background of most human activity and its supportive and adverse effects for living organisms remain a ‘background noise’ (Alberti, 2005; Goddard et al., 2010). Many ecological studies within urban areas still constrain their spatial assessment to urban green spaces, factoring the built-up areas only as ‘distance’ to the next green space (Matthies et al., 2017). While this might have been sufficient for most ecological research, urban design and architecture are also concerned with objects on smaller scales, such as individual buildings and their parts. Hence it seems necessary to reassess and add such concepts into ecological spatial frameworks with increasing resolution of analysis (Farinha-Marques et al., 2017).

Notably there is a fundamental difference in the description of urban space solely by land use and land cover and the actual patterns and processes, due to immaterial constraints such as legal limits and barriers (Marcus, 2005). Immaterial constraints may have severe consequences on the ecological composition and development in urban areas. While the concept of a ‘plot’ encompasses only an immaterial good, based on property rights, almost all decisions on land development will be made upon the immaterial rights and duties attached to the plot (Marcus, 2005). As the immaterial borders commonly do not coincide with landscape features (such as creeks or escarpments), habitats tend to be interrupted by artificial borders (Dramstad et al., 1996). This may have consequences for the dynamics which are constrained as opposed to an undulating natural landscape, i.e. species which might profit in their fitness from changing habitat patches due to different reasons, might be inhibited and not eligible for urban areas (Aronson et al., 2016). This shows that ecological processes are not necessarily limited to visible properties of landscape patterns and puts the validity of analyses constrained to material boundaries and visible entities into question.

A first vital step to relate ecological research to urban design and architecture is to delineate commonalities and to clarify how the two disciplines can communicate through space as common denominator. A short summary of spatial conceptions in ecology will be presented and set in relation to urban typo-morphology as dominating concept in urban design.

3.1.2 Ecological Concepts of Space

Ecological concepts of space are oriented towards changes in landscape or habitat patterns by key structural properties in a one- or two-dimensional spatial representation (Cadenasso et al., 2007). Habitats have been defined as "element of the land surface that can be consistently defined spatially in the field in order to define the principal environments in which organisms live (Bunce et al., 2005, 2008)". But as ecological research initially was based on natural or semi-natural landscapes, such definitions where calibrated on regional aspects of climate, topography, and vegetational patterns (Marcus et al., 2019a). Urban areas are characterized by a high variability of environmental and structural characteristics at small scales (Cadenasso et al., 2007). This limits the applicability of coarse scale approaches to architectural and urban design.

3.1.2.1 Rural-Urban Transects or Gradients

One of the most common methods for assessing urban ecosystems are rural-urban gradients. Ecological pattern and processes, as well as effects of the urban environment are analysed along an imaginary section from the urban core to the fringes or the hinterland (Turner, 1989). Although spatial factors are limited to one dimension (such as population density or mean building height), some general observations can be drawn, e.g. the variation of patch sizes (Goddard et al., 2010). Several important effects have been reported in urban-rural gradients related to biodiversity. McKinney (2002) described a general decline in species richness from the fringe towards the urban centre. Also, different species can be observed along a urban-rural gradient depending on their ability to capitalize on the altered environment: 1) *urban exploiters* in the centre, 2) *urban adapters* towards suburban areas and 3) urban avoiders at the urban fringe and beyond (McKinney, 2002). But there has been criticism about gradients in terms of not being spatially explicit, hence not being able to give information about variations in density and multi-centralities within cities (Goddard et al., 2010).

3.1.2.2 Land use and Land Cover (LULC) Mapping

Until the early 2000s, coarse-scale land classifications were merely used to capture urban growth. Since then, ecologists have engaged with the ecological functions in urban areas, for which such classifications proved to be of little use (Cadenasso et al., 2007). Land use and land cover (LULC) are common and often combined categorizations for geospatial classifications, derived from geography and remote sensing images. They have been developed to extend existing demographically oriented classifications, and as standardized schemes at national levels (Cadenasso et al., 2007). The primary differentiation of those traditional schemes is a division into urban and vegetated land cover. Such a division has been efficient for coarse, low-resolution data and big scales of investigation, but it does not account for the high heterogeneity within urban areas (Cadenasso et al., 2006b). For ecological considerations at a larger scale, distinctions of urban morphologies are usually considered negligible (Marcus et al., 2019a), and space is commonly reduced to a generic 'metric', not paying attention to sub-scale differences and creating the problem of 'outliers' either being incorporated into main categories or becoming a convolution without attributing to specificity. For use in urban ecology, the concept has been refined to different spatial scales and degrees of synthesis (Cadenasso et al., 2007; Farinha-Marques et al., 2017). But there has been criticism that such coarse information might be neglecting key factors in defining biodiversity resources in urban areas (Mathieu et al., 2007; Farinha-Marques et al., 2017).

3.1.2.3 The Patch-Corridor-Mosaic Model

In contrast to other fields in ecology restricting spatial characteristics to gradients or LULC classifications, landscape ecology has developed the concept of *patches* and their *edges* as fundamental spatial unit, which has been adopted by urban ecology as well (Alberti, 2005). The concept is differing between patches and *corridors* forming a so-called *mosaic*, representing the entire landscape (Dramstad et al., 1996; Marcus et al., 2019a). Patches refer to similar ecological conditions of habitat quality, landscape structure, or similar characteristics. Corridors on the other hand are patches with a distinct linear shape, usually connecting other patches. They facilitate dispersal among patches. The mosaic is the entire landscape consisting of patches and corridors with different properties (Dramstad et al., 1996; Marcus et al., 2019a).

Patches can vary in size and either one patch can constitute a habitat by itself, or a habitat can spread over several patches, a so-called *meta-patch* (Marcus et al., 2019b). A patch is characterized by its general shape, size and resulting edge zones influencing habitat quality by defining resource availability, competition and predator-prey relations (Alberti, 2005). Through patches, urban landscapes can be described as an assemblage of multiple patch types with varying ecological traits (Scolozzi and Geneletti, 2011). The concept of a 'patch' further enables to map and model landscape structures and their changes, but adding multi-species buildings further complicates the ecological mosaic and demands even finer resolutions in the ecological assessment of urban space.

Patch limits also mark ecological boundaries (Dramstad et al., 1996). They establish zones between neighbouring patches with altered characteristics compared to the core zones of the individual patches (Cadenasso et al., 2003). This is referred to as the *edge effect*, which describes changes in biophysical processes and species dynamics towards patch boundaries (Dramstad et al., 1996). Urban landscapes are characterized by sharp boundaries, resulting from human land use preferences (Alberti, 2005), narrowing such edge effects. Most studies do not consider urban patterns as spatially explicit features, but rather as areas fragmenting natural habitat patches, creating edges and affecting resource availability and species dynamics (Alberti, 2005).

Patches can occur naturally, but in urban contexts they are more often induced by human intervention (Marcus et al., 2019a). Patches and corridors are defined as distinctive elements within a larger *matrix* of contrasting, either homogenous or at least comparably similar ecological conditions (Marcus et al., 2019a). It is often used in population ecology (Beninde et al., 2015), but especially in urban ecology it describes for animals and vegetation seemingly uninhabitable areas of urban space. To this end, the matrix is still in use as a concept for the space in between semi-natural or natural urban habitat patches.

Some researchers claim that the matrix has been neglected due to reasons of private ownership and difficulties of conducting research within urban areas (Rigó and Barina, 2020). Other researchers claim that for ecologists, built-up land has for the longest time been considered as not valuable for conservation purposes, due to the assumption, that certain landscape patterns do not affect ecological processes other than building barriers and fragmenting habitat patches (Wu, 2014). But this conception has been criticised due to the mere fact that in some cities, the amount of green space within and in between buildings makes up a considerable proportion compared to public green spaces underlines the potential importance to analyse all parts of cities for biodiversity conservation (Werner, 2011). This critique is based on the investigation of small scales habitat patches within such matrices, finding that they do indeed provide habitat to numerous organisms and further can also promote the movement or dispersal between habitat patches (Werner, 2011). Adding multi-species buildings further complicates the ecological mosaic and demands even finer resolutions in urban ecological assessment.

3.1.3 Urban Morphology

The aim of the patch-corridor-mosaic model in landscape ecology is to describe biotope or habitat types by their predominant two-dimensional geometry (Marcus et al., 2019a). Like urban elements, the spatial geometry is also producing certain functional aspects (e.g. a linear, narrow biotope is mainly considered as a corridor to reach non-adjacent habitat patches). Further, patches and corridors do not primarily give information about the topographical and morphological characteristics, as patches can be flat, with small vegetation such as meadows, while corridors, such as shrubs or trees promoting the movement of some species, can be perceived as barriers for humans (Scolozzi and Geneletti, 2011; Marcus et al., 2019a).

While there has been comparison between urban morphology and landscape ecology by Marcus et al. (2019a), the problem has not been resolved, that urban morphology is spatially explicit, while landscape ecology is in plan representation implicitly representing biotic and abiotic factors. While elements of urban morphology such as streets, buildings, and street-blocks are to a certain extent scale-dependent, this is not evident for ecological pattern. These are mainly informed by their function and general geometry in contrast to surrounding elements. A matrix can refer to a natural or agricultural landscape around an urban agglomeration, as well as to the urban fabric of buildings and streets, surrounding patches of green like parks (Marcus et al., 2019a). This leaves the question of how to identify the smallest unit with homogeneous characteristics, which could also be seen as structural unit, from which an urban mosaic is built (Palazzo, 2022).

The approach of Marcus et al. (2019a), seen in Figure 3.1 unifies ecological and morphological configurations, at the expense of geometric and material information. It combines the information of human-made and natural inhabited parcels of land. Two-dimensional research has been the status quo in landscape and urban ecology for decades, due to various reasons as described in Section 3.1.2. In 2017 Alavipanah et al. (2017) conducted a review about the integration of the third dimension into urban ecosystem service research, and concluded, that although the potential impact of a three-dimensional integration of urban space can hold valuable insights into the understanding of the human environment, such studies had been exceptions. However, the approach by Marcus et al. (2019a) offers a valuable translation of spatial conceptions from one discipline to the other and should be assumed as best practice for interdisciplinary research at the moment.

Despite the efforts to integrate patch ecology with urban morphology by Marcus et al. (2019a), patches are primarily defined by ecological characteristics and do not necessarily have the same boundaries as urban and architectural objects. Depending on the scale of the patch definition, only parts of architectural objects, or multiple entities at once may be included in the patch definition. The most common approach to define the urban structure within a patch are metrics and indicators. This is facilitated by typomorphology (Biljecki and Chow, 2022) and other descriptors (Weinstock, 2011).

The patch-corridor-mosaic model, although lacking three-dimensional information, can be used as a primary spatial, and hierarchically organisation of research. But while spatial representations are focused on structural aspects and their changes in a spatial reference system, it is crucial to understand how different species access and populate urban areas and the available habitat patches within. These aspects are described in the following.

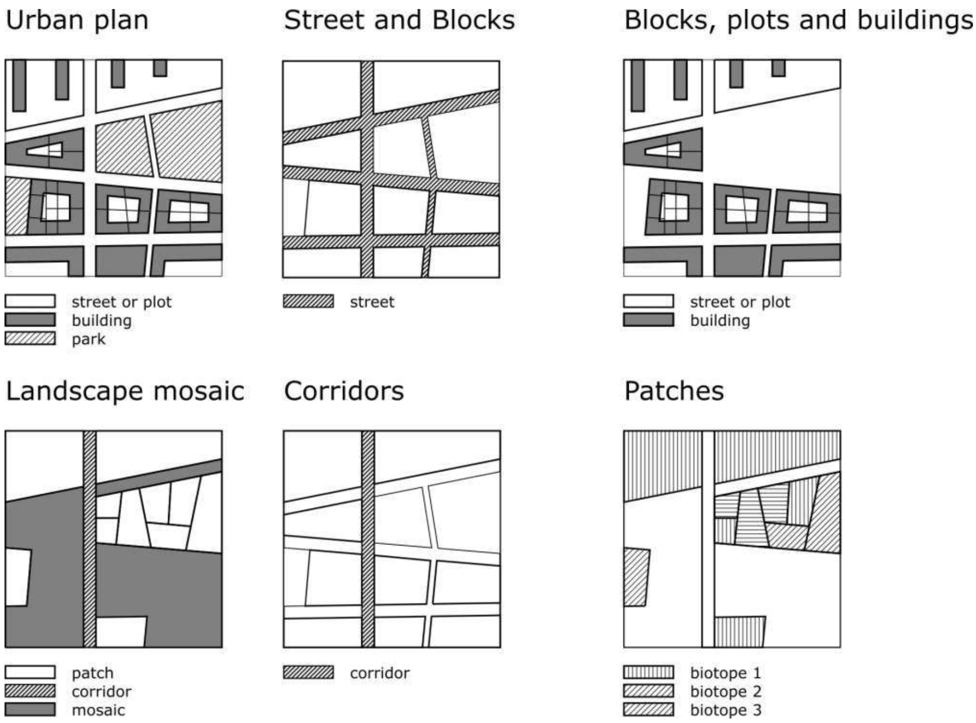

Figure 3.1: A correlation of descriptive elements in landscape ecology and urban morphology, after Marcus et al. (2019a).

4.
Biodiversity in Urban Ecological Systems

Biodiversity must not be understood as one single construct, but rather as a concept with several facets, which make application in urban design rather difficult (Desjardins-Proulx et al., 2019). Ecological studies generally focus on specific species or taxa, and their responses to urbanization can vary a lot, especially when looking at an abstracted spatial conception like the urban-rural gradient (Beninde et al., 2015). Biodiversity can refer to the richness or relative abundance of either 1) genes 2) species 3) ecosystem functions, or 4) ecosystems (Rebele, 1994). *Species richness* can be defined as the mean number of species per unit area. *Diversity* addresses not only to the absolute number of species, but also their relative evenness (Rebele, 1994). Further biological diversity is differentiated with regard to the spatial extent and scale in various biodiversity measurements, as seen in Table 4.1 (Rebele, 1994).

Table 4.1: Different measurements for biodiversity related to the spatial scale, as in Rebele (1994).

Name	**Spatial extent**
α-diversity	Local scale, most often used for species occurrence studies, depending on the home range of a species.
β-diversity	Connecting α- and γ-diversity; can also be described as the diversity between communities.
γ-diversity	Regional scale or overall diversity of a landscape.

4.1 Occurrence of Species in Urban Environments

Two important distinctions must be made according to the type of species occurring in urban areas: First, species might either be *'native'* to a regional ecosystem or *'non-native'*, *'invasive'* or *'alien'*. In urban areas, due to the heavily modified environmental conditions, alien species often have advantages in acclimatizing to the modified environment and making use of the artificial resources (Colléony and Shwartz, 2020). Native species generally decline in population in urbanized areas due to the reduction of natural, favourable habitats and their intolerance to human presence (Alberti, 2005). Both, native and alien species are affected by urban habitat fragmentation (Wu, 2014). Second, species might either be *urban followers*, *urban adapters*, or *urban avoiders*. Urban followers profit in general from the proximity to urban environments, adapters are able to compensate parts of their natural requirements, and urban avoiders cannot survive in cities (Colding, 2007). *Urban avoiders* are restricted to habitats outside cities or at their fringes, other species mostly inhabit very particular habitat patches within urban areas, whereas others may be able to make use of many different environments within cities by exerting various activities throughout the surrounding landscape (Hong et al., 2004).

4.1.1 Hierarchical Filters

There are many ways in which organisms belonging to distinct species are accessing urban habitats and their resources. Especially within urban areas, behavioural patterns are highly differentiated, due to the heterogeneous, fragmented, and abruptly changing spatial structures and environments (Spotswood et al., 2021). An important species trait in this context is the distinction of sessile and mobile species. There are many variations according to anatomical characteristics and the physical environment, such as terrestrial, aquatic, and airborne movement, as well as differences in distances that can be travelled or bridged (Spotswood et al., 2021). These specific traits, together with the distribution of resources, and interspecific interactions form the individual *home range* (Börger et al., 2008; Buchmann et al., 2011). Species with larger home ranges are enabled to exploit the full rural-urban gradient or use urban habitat patches as stopovers during migratory activity (Spotswood et al., 2021). This shows that the spatial scale, important for biotic processes of individual species varies from micro-habitats to the size of whole regions and continents. In addition to the home range Apfelbeck et al. (2019) note *dispersal abilities* and *barriers* as limiting factors.

Urbanization influences environmental and landscape characteristics not only by small-scale development of heterogeneous habitat patches, but also at bigger spatial scales through modified soils and altered hydrological networks and watershed patterns (Pickett et al., 2011; Wu, 2014). Human agency is altering components of the urban environment to improve human habitat conditions, but through such modifications of the environment, diverse filters are applied, selecting and determining available species for the altered and fragmented biotopes (Andrade et al., 2021; Aronson et al., 2016). Such biotopes are usually at a disequilibrium, which means that the observed state of the biotope is not stable over time (Holling, 1992). In this context *extinction debt* signifies an overabundance of species in a landscape, compared to the same landscape at an equilibrium state (Hahs et al., 2009).

The capability of species to colonize urban environments is dependent on structural patterns at multiple scales, and the factors limiting the availability of species are thus called *hierarchical filters* (Aronson et al., 2016). At coarse scales several factors of hierarchical filtering in urban areas have been identified, as seen in Table 4.2. It is important to understand that these filters act simultaneously (Williams et al., 2009). At a landscape ecological scale such filters may impose barriers which limit, or corridors which facilitate dispersal of organisms (Apfelbeck et al., 2019).

The *regional pool* consists of species, which are indigenously inhabiting a region or have become part of the biocenosis over time. Thus, they are naturally available as candidates for populating urban areas (Apfelbeck et al., 2019). The *urban species pool* narrows the regional pool down to represent those species, which are actually capable of populating and surviving under urban conditions. The *local pool* signifies those species which can reach and populate a specific habitat patch (i.e. a park, garden, etc.) (Aronson et al., 2016).

Table 4.2: Hierarchical filters from Aronson et al. (2016) and Williams et al. (2009)

Aronson et al. (2016)	**Williams et al. (2009)**
Regional climate	Urban environment
Biogeography and land use	Habitat transformation
Urban form and development history	Fragmentation
Human mediated biotic interchange	Human preferences
Socioeconomic cultural influences	-
Local human facilitation	-
Species interactions	-

The concept of hierarchical filters explains the reasons why any species occurs in an urban environment at all. This is important, because not all species living in a certain region will be able to access and populate a given plot within the urban mosaic, even if conditions at the specific habitat patch are favourable.

4.1.2 Species Composition and Abundance

Species *composition* in urban areas is determined by many natural and anthropogenic factors, such as hierarchical filters. Literature is suggesting that, although individually highly heterogeneous, at a global scale, cities have many uniform characteristics, such as impervious surfaces, fragmentation of green spaces, high-rate disturbances, which facilitate a relatively narrow and homogeneous species pool, based on traits (Aronson et al., 2016). Although most species interactions occur within local and neighbouring patches, spatial relations at larger scales play a significant role (Aronson et al., 2016). It is argued that geographic location is even the main factor determining species richness (Rigó and Barina, 2020). Other factors such as altered patch composition and configuration through human agency needs to be addressed in terms of spatial properties, which have at least a certain degree of explicitness.

Many approaches in biological conservation focus on a single species, which is problematic due to interspecific processes necessary to uphold an ecosystem. Multiple species will form a *community of species* with certain patterns of interaction. The assembly of communities is determined by functional traits and life history of species (Aronson et al., 2016). The concepts of *metapopulation* and *metacommunity* describe the abundance and occurrence of groups of species and communities respectively at regional scales (Chase et al., 2020). The metacommunity concept is offering four core processes determining species composition: 1) environmental filtering due to abiotic conditions (see Section 4.1.1), 2) biotic interactions, 3) dispersal among habitat patches, and 4) ecological drift (Chase et al., 2020).

But even if accessibility through urban filters for a species is granted, the patch itself needs to be suited for the species to fulfil its different needs to complete its *life cycle* (Hauck and Weisser, 2015). Life cycles refer to the concept of all actions and interactions an individuum needs to undertake from birth to death and constitute a plethora of activities, depending on the species, such as photosynthesis, foraging, predation, sleep, procreating, and nesting (Hauck and Weisser, 2015). These activities might require different environments, which means that many species are forced to move between different biotopes.

Three important concepts regarding how species establish populations in ecosystems are (Rebele, 1994):

- *Colonization* describes the process of accessing, populating and initially altering an existing ecosystem by a species.
- *Succession* refers to the development of species communities over time. Typically, depending on the environment only specialised species are capable of colonizing an ecosystem. After several generations, the environment might be significantly changed (such as soil building, shading through tree canopies), and become available for other species, which might coexist or supersede other species.
- *Interaction* are all processes of simultaneous exchange and influence that different species may exert on each other (predation, symbiosis, competition). Biological interaction among individual organisms and species happens after Rebele (1994) through: 1) competition, 2) predation, 3) mutualism.

After Grilo et al. (2022) *functional traits* are differentiated in *response traits* and *effect traits*. While the prior is addressing the performance or fitness of individuals, and can influence “environmental tolerances, habitat requirements, and responses to pressures”, the latter addresses influence on ecosystem structure and function.

The *functional group* concept collects all species which hold similar traits and can therefor exert similar roles in a community or biocenosis (Lundholm et al., 2010). This has also led the scientific discourse of landscape ecology to focus on *functional diversity* rather than on species diversity within the urban ecological paradigm (Alberti, 2005). Similar functional traits are further granting a certain substitutability of species to maintain key functions within altered ecosystems (Grilo et al., 2022).

Umbrella species are indicative of other species, which are assumed to cohabit biotopes due to interspecific activities (Ardiantiono et al., 2024), whereas *indicator species* are supposed to be indicative of characteristics of entire ecosystems (Löfvenhaft et al., 2002). Another important distinction of biodiversity research is between biotic factors and abiotic ones, since research has shown that biodiversity within cities depends on geophysical factors, as well as biophysical ones (Beninde et al., 2015; Aronson et al., 2014). Individual species differ in their dependence on ecosystem characteristics and properties, hence some species might indicate larger scale conditions, while other species may indicate smaller scales (Alberti, 2005).

Several processes and interactions can be distinguished, which are directly or indirectly affecting species *abundances* in urban areas compared to natural landscapes, as seen in Table 4.3 (Alberti, 2005):

Table 4.3: Direct and indirect factors, affecting urban species abundance, from Alberti (2005).

Direct	**Indirect**
Ecosystem processes	Changed predation
Habitat alteration	Interspecific competition
Food supply	Diseases

Although all of the mentioned factors are to some degree themselves influenced by the spatial configuration of the abiotic environment (Andrade et al., 2021), habitat alteration is the component which is directly impacted by urban design and planning. But the effect of environmental conditions found in intra-urban habitat patches on biodiversity, is species dependent (Beninde et al., 2015). There are trade-offs even for individual species, and the impact of urban environments might vary from negative to positive (Gentili et al., 2024; Spotswood et al., 2021).

4.1.3 Habitat Suitability and Potential

Habitat patches can be assessed by habitat type and habitat quality (Riitters et al., 1997). However, using landscape indices for inference of other ecological features, such as biodiversity or habitat modelling has been criticised for not producing valid results and therefore little usefulness for designers and planners (Corry and Nassauer, 2005). Habitat suitability describes the aptness of a habitat patch for the dispersal, population, and maintenance of a viable population size.

Habitat requirements are species-specific and have to be assessed individually or by (meta-)community models (Chase et al., 2020). Planning and design considerations need to account for the achievability of habitat requirements, as some might need to be present already, due to slow development processes, while other can be realized through construction or management (Apfelbeck et al., 2019).

There are different possibilities to assess the suitability of habitat patches for biodiversity purposes. Löfvenhaft et al. (2004) define a classification based on the population success of organisms on an ordinal scale: 1) hostile, 2) unsuitable, 3) dispersal, 4) survival, and 5) breeding. *Metapopulation ecology* describes the processes of species populating fragmented patches, where populations in some patches might become extinct, but there is a possibility of repopulation through other patches in the vicinity (Chase et al., 2020). Scolozzi and Geneletti (2011) proposed *habitat potential (HP)* as rule-based assessment method for the impacts of land use changes on biodiversity at local scales.

Table 4.4: Different ecological models as presented by Andrade et al. (2021)

Model name	Characteristics
Patch dynamics	Homogenous environment, limited dispersal of species, colonization and competition shape community composition.
Species sorting	Heterogenous environment, niche partitioning and resource gradients drive biodiversity patterns.
Mass effects	Species are associated with certain habitats; patch structure creates source-sink dynamics.
Neutral theory	Functional equivalency of species, random demographic effects and dispersal limitation.

In ecological networks, risks are minimized through redundancies and spread over the network (Opdam et al., 2006). How urban development is affecting biodiversity through causing varying concentrations, connections, and increasing heterogeneity is not well understood (Alberti, 2005). Nevertheless, has it been acknowledged that relationships between patches play a major part in urban ecosystems. Ecologists tend to distinguish between features of local habitat suitability and the surrounding landscape's connectivity or permeability for biodiversity measurements (Beninde et al., 2015; Apfelbeck et al., 2019).

4.2 Abiotic Factors of Multi-Species Urban Ecosystems

After having established basic considerations of how species react and interact with urban environments, it seems necessary to identify spatio-temporal properties, through which biodiversity can be addressed in analysis and classifications, to abstract the complex actions of individual organisms, populations, and communities to the patch as a spatial reference unit. As already shown in Section 3.1.2.3, patches in urban systems can be defined in several ways, and characterized by biophysical structures, built structures, social structures, and immaterial boundaries (Grove et al., 2005; Cadenasso et al., 2006b).

Models aiming to predict urban biodiversity, make different assumptions about the respective weight of social, environmental, and spatial factors. Explanation of the influence of factors can be abstracted into models of 'species sorting' assuming no influence of distance between spatial factors, 'mass effects' assuming a degressive influence, 'patch dynamics' putting emphasis on spatial distance and 'neutral theory' which tries to explain urban biodiversity as a null hypothesis solely by spatial distances (Leibold and Chase, 2017; Andrade et al., 2021). The models are summarized in Table 4.4.

A possibility to determine intra-urban variation of biodiversity is to isolate and quantify factors of influence (Beninde et al., 2015). Factors of influence are important to establish systemic dependencies. Andrade et al. (2021) describe three main factors influencing urban biodiversity as 1) environmental, 2) social, 3) spatial. With respect to cities and the dominant human influence, such factors could be further categorized into static topographic features (i.e. the layout or configuration) and dynamic anthropogenic processes (management) (Beninde et al., 2015). Spatial heterogeneity is an important variable, but it has to be considered that the correlation between biodiversity and heterogeneity varies depending on the specific metric (e.g. slope is negatively correlated with species richness) (Wu and Qi, 2000). Pickett et al. (2011) formulated factors of influence on urban biodiversity as: 1) prevailing climate, 2) substrate, 3) resident organisms and their residual effects, 4) landscape relief, and 5) time. Another possible differentiation comes from landscape characteristics as *structure*, *function*, and *change* (McGarigal and Marks, 1995; Botequilha Leitão and Ahern, 2002).

These factors of influence can be used to gain information about different aspects of biodiversity by either mapping them to gain information about the underlying landscape structure, constructing ecological networks, or analysing interactions between elements of ecosystems.

4.2.1 Habitat Structure

The overview of biodiversity in urban areas, important concepts, drivers and indicators shows the extent and complexity of studying the presence of other species under urban conditions. On the other hand, such high specificity for individual species or communities suggests a necessary flexibility in the concept of habitats, while generating a consistent, scalable system, which is spatially explicit, i.e. can be universally translated into various spatial configurations (Farinha-Marques et al., 2017).

Byrne (2007) presented a framework centring habitat structure as link between coarse scale ecosystem and landscape assessments, and sociocultural systems and the resulting human activities. Habitat structure is “the amount, composition and three-dimensional arrangement of physical matter (both abiotic and biotic) at a location (Byrne, 2007)”, with special regard to human altered structures and artifacts. Byrne (2007) conceived of habitat structure as the local-scale environments, influencing biotic and abiotic conditions and constituting landscape patterns at larger scales. Although originally developed for soil research, it was thought of as incorporating enough flexibility to be applicable in a broad context.

Landscape archetypes can either be real world examples, or abstracted idealized landscape units (Cullum et al., 2017). *Ecological archetypes* are not a model per se, but the analysis of recurrent patterns in socio-ecological systems at an intermediate level of abstraction (Oberlack et al., 2019). These approaches enable the identification of models which are able to explain a phenomenon under a set of specific conditions.

The concept of *ecotopes* describes a classification approach of habitat patches based on their predominant vegetation. They constitute the smallest possible distinct landscape features in habitat classification (Ellis et al., 2000). Ecotopes are established at species-specific scales, as every species responds with varying intensity to different aspects of landscape complexity (Hong et al., 2004). In the context of urban environments, such units have to be adapted to the built environmental structure. Hong et al. (2004) emphasized the importance of accounting for human activities in establishing ecotopes. Within the concept of ecotopes, boundaries are also referred to as 'ecotones' (Hong et al., 2004).

The 'keystone structure concept' states that, together with other habitat-specific spatial processes, the mentioned operational factors cause that different species show different reactions to habitat heterogeneity at specific scales, causing peaks in habitat heterogeneity effects to species groups at specific scales (Tews et al., 2004). The *extended keystone hypothesis* states that all terrestrial ecosystems are dependent on a small number of biotic and abiotic entities and their interactions, which affect temporal and spatial structures on a wide range of scales (Holling, 1992).

Around the new millennium the *'ecosystem functional type' (EFT)* concept, an approach to classifying and modelling ecosystem processes has been developed. The hypothesis is that within a given land cover type, ecosystems correspond (Wu and David, 2002). For this time, such an approach was feasible with the remote imagery available from Landsat and other satellites. While still relevant for higher-level evaluation, such an approach is disregarding the actual structuring of landscapes at smaller scales. Functional types, as well as species composition and richness affect the mode and efficiency of resource cycling within ecosystems, and with it the pattern and processes (Alberti, 2005). Therefore, biodiversity can be a good indicator of the condition of an entire ecosystem.

Rigó and Barina (2020) conducted a classification of intra-urban habitats within Budapest with territorial units at the architectural and neighbourhood scale. They analysed public and semi-public urban spaces, dividing them into street sections (linear), squares and yards (extensive), and respective overlaps of these categories. Within these broader habitat groups, they distinguished many different green space categories, even detailing 'walls' and 'cracks' as possible habitats. Their research resulted in a selective study of plant species habitat preferences. This approach is enhancing classifications by structural elements of built urban environments, partly at building or sub-building scale.

These considerations show the importance of detailed evaluation of habitat structures, to include all relevant properties and find a patch definition that is suited to the scales at which individual species or communities act.

4.2.2 Urban Ecological Networks

Heterogeneity is a key characteristic of urban environments, and although this offers suitable habitats for a number of different species, individual habitats tend to be heavily fragmented throughout the urban landscape (Alberti, 2005). Three levels of habitat fragmentation can be distinguished according to Verboom and Pouwels (2004):

1. low fragmentation: at least one patch in a network is large enough to support a minimum viable population (MVP);
2. medium fragmentation: at least one patch can support a key population, which is depending on other patches providing opportunity for immigration;
3. high fragmentation: no key patches occur in a network, but the total carrying capacity of the network can compensate fragmentation.

Fragmentation of habitat patches is not only directly influencing the connectivity of the landscape mosaic but also influencing diversity, structure and distribution of vegetation (Alberti, 2005). There has been further evidence that certain factors are more influential than others, e.g. habitat area is more important than connectivity (Beninde et al., 2015). Ample research exists on the occurrence of species within urban areas, trying to measure effect sizes of various ecosystem factors. Although such studies give informative insights for specific case studies, they generally provide little input to design requirements (Beninde et al., 2015).

While some ecological processes are influenced only from within a patch, adjacencies to other patches can play an important role, as well as large scale connections through material and energy transport, e.g. watersheds (Alberti, 2005). The *landscape complementation* hypothesis by Dunning et al. (1992) proposes that in heterogeneous landscapes, species need to move in between patches to fulfil their critical biotic requirements (Colding, 2007). This implies that species, from individuals to communities, are affected by the composition and configuration of the urban landscape, and ecological functional units might be dispersed over several patches (Colding, 2007).

Connectivity is the inverse to the fragmentation of habitat patches and informs about the influence of landscape patterns on the ability of species to move and disperse amongst habitat patches (Liu et al., 2024). Connectivity is altered by modification of the terrain and resulting changes in the biophysical structure (Alberti, 2005). From a metapopulation perspective, wider temporal and spatial scales play a role in defining the possibility for exchange amongst subpopulations via *stepping stones* or corridor patches [Colding (2007); SimberloffCox1987]. *Stepping stones* are habitat patches which are not suitable for permanent occupation, but offer crucial resources for migration, so that they can be used to overcome habitat fragmentation (Colding, 2007). There is a distinction between *structural connectivity* and *functional connectivity* (Liu et al., 2024). The basic components of urban ecological networks are shown in Figure 4.1. While structural connectivity is limited to the spatial arrangement of habitat patches, functional connectivity considers the actual biological processes which are coupled with movement and dispersal between patches. It is argued that landscape characteristics can be substituted by other characteristics providing similar features, such as movability, e.g. corridors substituted by stepping-stone habitats which reduce the distance to other favourable patches (Beninde et al., 2015).

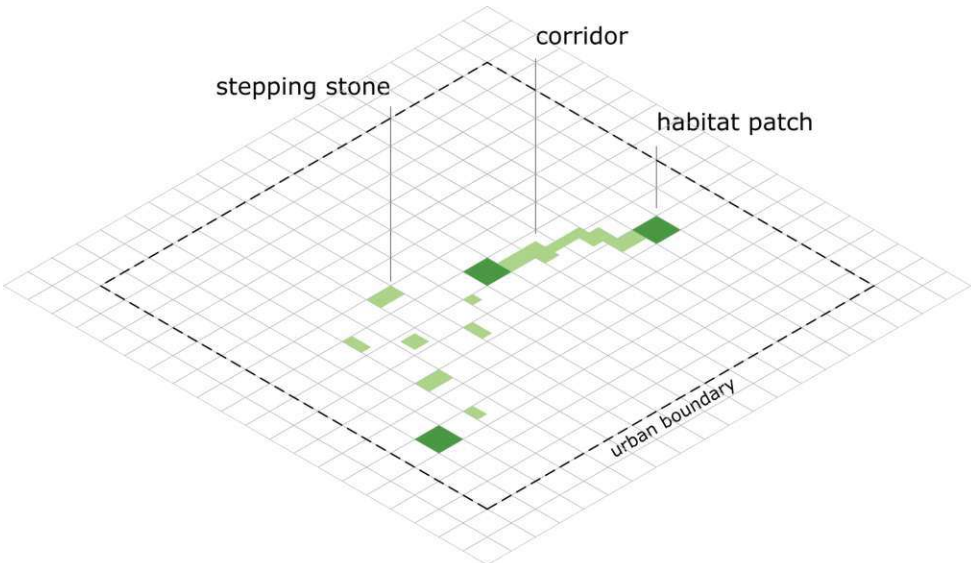

Figure 4.1: Types of ecological connectivity between habitat patches.

Although species occurrence and environmental variables can inform about important habitat patches, anthropogenic threat sources and specific environmental variables can help to identify *ecological resistance* (Luo et al., 2022). Information about suitable habitat patches and occurring resistance can be used to construct an ecological network, allowing for clustering habitat patches and identifying inter-cluster corridors (Luo et al., 2022).

Resistance surfaces, describe the difficulty, also referred to as *impedance* (Kong et al., 2010), for individual species to cross certain patches within the urban mosaic, according to their structural characteristics in form of *barriers*. *Ecological corridors* are narrow patches surrounded by differing landscape patterns, which are permitting or even promoting the dispersal of organisms (Peng et al., 2018). Species are not necessarily found only in habitats which are beneficial to them and support viable populations. Animals can be attracted to inferior habitat patches, producing *ecological traps* (Battin, 2004). Spatial scales of landscape structure are impacting the perception and response of taxa and species in very specific and differing ways (Goddard et al., 2010).

Ecological networks can be more generally characterized as sets of patch systems, which are spatially or functionally linked by flows of materials, organisms, and other ecological processes and which are in exchange with the surrounding mosaic of patches (Oh et al., 2011; Opdam et al., 2006). Networks are complementary to the patch-corridor-mosaic model. Ecological networks are divided into *core area*, *corridor*, and *buffer area*, details given in Table 4.5 from Oh et al. (2011). Ecological networks can vary considerably in size and might transgress urban boundaries, as seen in Figure 4.2.

Table 4.5: Ecological network components from Oh et al. (2011).

Classification	**Explanation**
Core Area	Existing important ecosystem areas, satisfying habitat requirements for species.
Corridor	Functionally connected areas which facilitate the dispersal and migration of species in ecosystems.
Buffer Area	Areas surrounding core areas with differing landscape characteristics, potentially absorbing negative impacts and protecting core areas.

The aspect of functional connectivity is important in ecological network construction, since not all relationships are due to spatial adjacency. Several ecological processes might not be easily identifiable by spatial characteristics, such as watersheds, other connections and networks might be of cyclical or periodical nature, such as creeks which might dry up seasonally.

Concepts of landscape complementation and landscape supplementation (Colding, 2007) underline the necessity to evaluate not only a specific site for its suitability as a habitat, but to enclose the latent functionality of adjacent or nearby patches. It also shows that the alteration of a site might change its functionality as habitable patch for species already occupying nearby patches of land (Goddard et al., 2010). This broadens the discourse of ecologically informed design and shows the high sensitivity in land development decisions, according to landscape connectivity. Furthermore, it shows that even when there is no specific target species or community to be established on a site, ecologically informed design will benefit biodiversity on larger scales (Colding, 2007; Goddard et al., 2010).

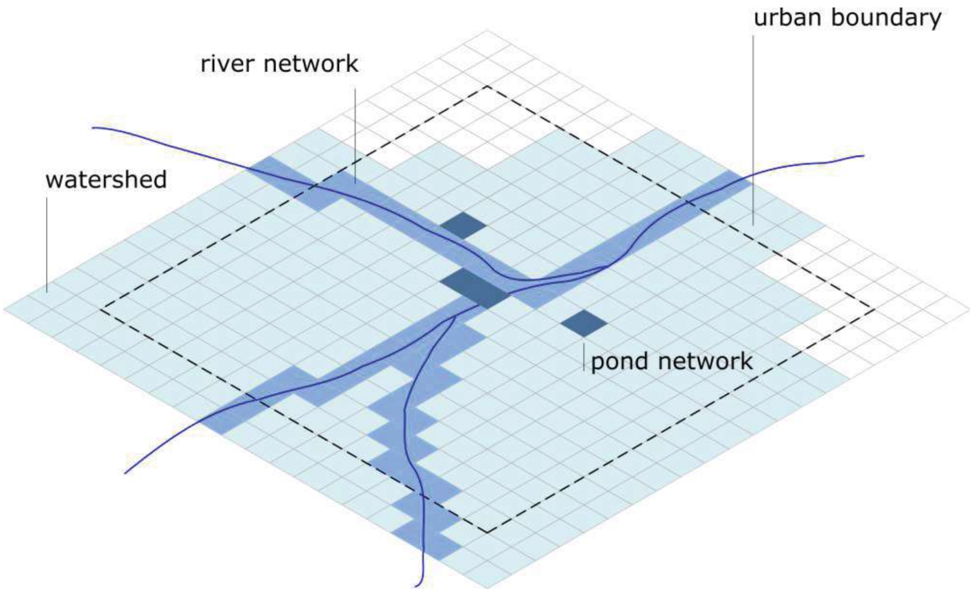

Figure 4.2: An example of functional and structural connectivity where a stream is passing a habitat patch, impacting the patch through material transport and building a corridor to remote patches.

4.2.3 Patch Dynamics and Change

Patches change in time (Cadenasso et al., 2006b), as their structure is continually developing, due to the biophysical processes and interactions of all biotic and abiotic *agents* within an ecological system (Pickett et al., 2011). Some development might create recurring patterns, such as seasonal vegetation cycles, others change the patch structure permanently, such as plant successions. Patch dynamics establish the need to consider that while design might be targeting specific species, the patch might become available for unintended species after development. On the other hand, isolated habitat patches, which have higher suitability than others, might not be reachable at a certain point in time. Species will occupy less qualitative, but easier accessible habitat patches, until the network configuration changes, and the high qualitative habitat patch becomes accessible (Cadenasso et al., 2013).

4.3 Principles of Complex Ecological Systems

Biodiversity in multi-species urban ecosystems can be summarized by three key factors as presented above. *Habitat structure* explains the within-patch composition (Byrne, 2007), i.e. the biotic and abiotic properties, whereas *ecological networks* describe the connection and exchange in between patches (Verboom and Pouwels, 2004). They could also be considered as the underlying infrastructure, with distinct spatial properties, for dynamic processes and species interactions. But the real-world observations, necessary to detail the above-mentioned properties and characteristics are limited to reductionist methods, which are not suitable to gain a general understanding of which mechanisms are central to answering urban ecological questions.

Urban ecosystems are evolutionary systems, where pattern at higher levels emerge from local interactions (Alberti and Marzluff, 2004). Complexity theory is an important keystone to understand how ecosystems differ from deterministic ways of thinking, often applied in design and planning tasks. However, such abstract understandings need refinement and recontextualization if they are supposed to be able to be described in qualitative or quantitative terms, which is the prerequisite for analytical research (Wu and Qi, 2000).

4.3.1 Hierarchy and Scale Coupling

Establishing an explicit representation of space requires the identification of structural and functional units at a scale where processes of interest occur (Wu and David, 2002). After Wu and David (2002), *patch dynamics* promote the conception of ecological systems showing *emergent properties*, which can only be described by the interactions at a horizontal patch level, but verticality within ecosystems can be addressed by *hierarchy theory*. *Hierarchical structure* is a conceptual discretization enabled through *loose coupling*, which allows for the separation of system levels on a vertical axis, and distinct subsystems at a horizontal one (Wu, 1999; Wu and David, 2002). This decomposability is the foundation of systems to be analysed and modelled at scales of interest (Wu and David, 2002). This is facilitated by the assumption that interactions between subsystems operate at longer intervals, and long-termed dynamics of the entire system are determined by slow processes, thus enabling the isolation of short-termed dynamics of individual subsystems (Wu and David, 2002). However, perturbations are capable of creating non-linear effects at multiple system levels (Wu and David, 2002).

Nested hierarchies describe the full enclosure of components at a given level by components at a higher systems level (Wu and David, 2002). For practical reasons, scientific domains usually define lower and upper boundaries of systems. For example, in landscape ecology, individual organisms are the basic units, forming, together with other biotic units and abiotic factors, local patch ecosystems, which generate larger scale flow systems, forming the landscape as upper boundary (Wu and David, 2002). This logic of decomposable, distinct hierarchies within ecosystems allows to make their complexity graspable, and further define top-down constraints as well as bottom-up pattern and processes (Wu and David, 2002).

For any *focal level*, there exists a higher level forming constraints, setting the context and exerting control, while a lower level defines the components and mechanisms of the focal level (Wu and David, 2002). Lower system's levels are described by smaller patterns and higher process frequencies, while higher system's levels describe larger patterns and slower process frequencies (Wu and David, 2002). For most ecological phenomena or processes, there are many factors influencing them, but there are usually only few dominant factors at any spatio-temporal domain of scale (Wu and David, 2002). Considering the choice of scales, it might become necessary to transgress the boundary of 'urban areas' to validly address influencing factors, since different species reflect environmental conditions (and vice versa) over larger scales than other (Alberti, 2005).

To understand the interdependencies between pattern, process, and system function, it is necessary to address the relationships at several temporal, spatial, and organizational scales [Alberti (2005); Cottingham2002]. This regards the *extent* as well as the *granularity* of data, as only if pattern and processes act on the same hierarchical level, interaction is observable (Wu and David, 2002).

The *hierarchical patch dynamics paradigm* unifies vertical and horizontal coupling of ecological systems, holding several tenets stated by Wu and David (2002), as seen in Table 4.6.

Table 4.6: Tenets of the hierarchical patch dynamics paradigm from Wu and David (2002)

Tenets of hierarchical patch dynamics
Ecological systems are spatially nested patch hierarchies, in which larger patches are made of smaller patches.
Dynamics of an ecological system can be studied as the composite dynamics of individual patches and their interactions at adjacent hierarchical levels.
Pattern and process are scale dependent, and they are interactive when operating in the same domain of scale in space and time.
Non-equilibrium and stochastic processes are not only common, but also essential for the structure and functioning of ecological systems.
Ecological stability frequently takes the form of meta-stability that is achieved through structural and functional redundancy and incorporation in space and time.

4.3.2 The Development of Complex Adaptive Systems

In the context of biodiversity in urban ecosystems, *resilience*, which is the ability to compensate disturbances, and, crossing certain thresholds, shifting to a new regime (Elmqvist et al., 2019), is dependent on species abundance, distribution, diversity and interaction on different spatial and temporal scales (Alberti, 2005). Elmqvist et al. (2019) addressed the implementation and management of resilience of cities as complex systems, to gain capacity to secure and stay on a developmental trajectory towards sustainability. Along with a changing understanding of sustainability and resilience as two different but related concepts, the notion of cities as self-organizing systems demands a turn in urban planning away from an idealistic anticipation of spatial form, towards and adaptive and maximizing capacity building for resilience (Ahern, 2013).

Core characteristics of urban ecosystems as self-organizing or *complex adaptive systems (CAS)* are high diversity of components and their spatial heterogeneity, non-linear behaviour (Levin, 1998), multi-scale interactions, and self-regulation (Wu, 2014). Such systems are characterized through dynamics of unpredictable interplay between development, disturbances (internal or external) and adaptation (Holling, 1992; Wu, 2014). Although the notion of ecosystems as CAS is neutral towards shifts into new regimes, it is clear, that the support and the quality of human and non-human life is dependent on certain margins and conditions of ecosystems, which are comparably narrow.

The perspective of urban systems as complex adaptive systems is suggesting some key properties, which have to be taken into account for analytical purposes. These properties are described by Preiser et al. (2018), as seen in Table 4.7. Hierarchy is a core property of CAS and occurs in nature often as modularity (Wu and David, 2002), as mentioned above.

Human control over such complex behaviour is clearly limited and without intervention, regimes will always shift to ecologically more favourable or efficient ones, which in densely populated urban areas will probably come with negative effects for traditional forms of human settlement and technical infrastructure. While a basic assumption of CAS is the impossibility to predict the development of CAS, there have been efforts to describe urban ecological systems by structure and function.

Table 4.7: CAS properties from Preiser et al. (2018).

Types of organizing principles that bring about CAS features	**Related concepts and attributes that characterize CAS**
Constituted relationally	Netlike structure, hierarchies, holarchic, diverse components, built-in redundancy, heterogeneity.
Adaptive capacities	Self-generation, self-organization, decentralized control, memory, evolutionary and concurrent persistence and change (resilience), anticipatory capacities.
Dynamic processes	Far-from-equilibrium, multiple-trajectories possible, periods of fast and slow change (punctuated equilibria), nonlinear interactions, attractors, thresholds, tipping points, regime shifts, feedback loops (enabling and constraining), cross-scale interactions.
Radically open	Porous boundaries, embeddedness, nestedness, exchange of matter, information, energy, teleconnections.
Contextually determined	Function changes as system changes, components with multiple context dependent identities.
Novel qualities emerge through complex causality	Circular/recursive causality, large webs of causality, multiple pathways of causality, high levels of stochasticity, same starting conditions that produce different outcomes, emergent properties.

4.3.3 Patterns and Processes of Urban Ecosystems

A basic understanding of ecosystem theory are the objects of description. Ecologists, who have adapted complexity as important concept already in the late 1990s, commonly distinguish between *pattern* and *process* of ecological systems (Cadenasso et al., 2006a). Patterns represent distinct spatio-temporal occurrences and distributions of ecologically relevant factors, which set the boundary conditions for processes to happen, while processes describe the agency of biophysical factors coupled to identifiable mechanisms, either directly affecting patterns or intermediate processes (Alberti et al., 2003). Alternatively, ecosystems can be described through three main factors: 1) structure, 2) function, and 3) spatial heterogeneity (Alberti, 2005). Chenetal2014 on the other hand differentiate between structure, process and function.

A possible synthesis of the different descriptions could be as follows [Li and Wu (2007); Li and Reynolds (1995); Chenetal2014]:

- *Structure* addresses the complexity and variability of system properties, i.e. their biophysical components, in space and time, as well as their *patterns* of dispersal.
- *Process* refers to the change within or the exchange in between biophysical components.
- *Function* is giving purpose to structures and processes, as maintaining and developing an ecosystem.

The influence of patterns and processes is interchangeable (Botequilha Leitão and Ahern, 2002). In general, dynamics of urban systems, can be viewed from an ecological perspective in a cyclical manner (Alberti, 2005):

1. Changes in land cover affect biophysical factors (diversity, net primary production, soil quality) as well as geophysical factors (runoff, watersheds, sedimentation, etc.);
2. This changes biotic processes by altering the availability of nutrients and water, affecting size and dynamics of populations, communities, and overall ecosystems;
3. Through physical factors and biotic processes the ability to regenerate habitats (regulate microclimate, air quality, etc.) is affected;
4. Anthropogenic activity is influenced by the habitat quality and in response further changes land cover, biotic processes, which in turn again influences habitat quality.

Table 4.8: Main components of social-ecological systems from Alberti et al. (2003).

Component	**Examples**
Pattern	Land use, Land cover, Heat Islands.
Process	Nutrient cycle, predation, community development.
Effect	Natural productivity, salmon runs, community dynamics.
Driver	Population growth, climate, economic growth.

Alberti et al. (2003) extended the pattern-process model to four main components (see Table 4.8) which are in a cyclical relationship: 1) Patterns, 2) processes, 3) effects/change, and 4) drivers. Effects can be characterized as processes at an intermediate level from a functional perspective, while drivers are activities, either anthropogenic or geophysical, able to induce change in ecological patterns (Alberti et al., 2003). Patterns and processes, together with disturbances, form multi-dimensional pathways and permutations of interactions within ecosystems (Geary et al., 2020).

4.3.4 Disturbances and Transient Dynamics

Alberti (2005) proposed that ecosystem functions are affected by urban development patterns, whereas the mechanisms are represented by land cover change and the modification of natural disturbance. *Disturbances* are understood as discrete events, of either natural or anthropogenic origin, that abruptly change ecosystems and their components (Alberti, 2005; Kearney et al., 2019). Processes, natural events and human intervention in urban ecosystems can be categorized at different scales with different frequencies and magnitudes. Intentional human agency, such as the foundation of new settlements are rare events, compared to the everyday traffic within cities (Batty, 2005). But while disturbances are often only conceived of as the act of a phenomenon, disturbances can also show a persistent character, if the structure of the socio-ecological system is modified (Alberti et al., 2020). *Transitions* mark an ongoing regime shift that will result in persisting new traits (Bestelmeyer et al., 2017). Slower, or gradually occurring changes (e.g. microclimate, morphology, hydrology) modify natural disturbance regimes (Alberti, 2005), and might be considered as a 'constant' ecosystem characteristic, depending on the time scale of investigation. Disturbances as well as other activities, such as agriculture, do not only impact ecosystems while they occur, but they inscribe into the landscape (Ramalho and Hobbs, 2012), and alter ecosystems permanently.

Several characteristics of urban landscapes, produced by human intervention, create potential disturbances for ecological processes, which are not always obvious. Urban features tend to create sharp boundaries, with hard edges, while high variability and heterogeneity create small patches, unnatural shape complexity, fragmentation, homogenization of natural patterns, chronic stresses, etc. (Alberti, 2005). Additionally, natural disturbances in urban landscapes are usually altered in their quality or quantity (introduction of invasive species, altered drainage and runoff patterns, etc.), due to modified landscape characteristics (Alberti et al., 2003; Alberti, 2005). Land management, which could be defined as a repetitive or sustained intervention of varying magnitude, is having a substantial influence on existing vegetation and its ecological potentials (Cadenasso et al., 2006b). Such management is usually informed by social and legal boundaries and rules, thus producing abrupt differences in landscape, disturbance and community patterns (Cadenasso et al., 2006b). Figure 4.3 is giving an overview of different disturbance regimes, with their respective magnitude in temporal and spatial dimensions.

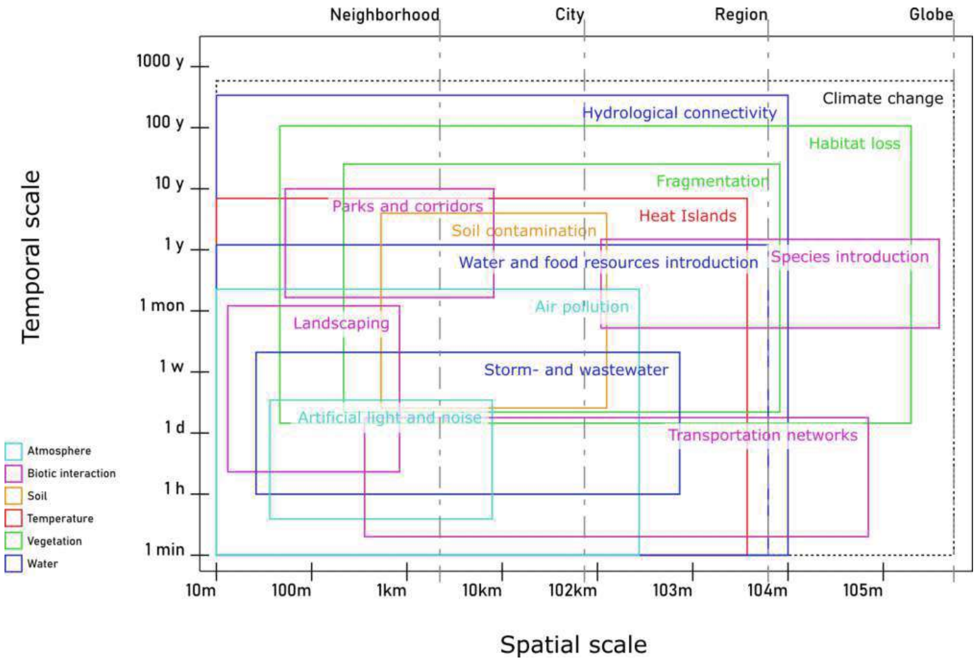

Figure 4.3: Disturbance regimes and their spatial and temporal extent, after Alberti et al. (2020) and McDonnell and Hahs (2015).

Hastings (2010) has defined *transient dynamics* as a contrasting concept of temporarily changed system behaviour over short and long time scales. He finds that this is usually caused by some kind of disturbance, which only occurs for a limited period, after which the systems will return to their 'default' setting. This does imply, that an observed systems state, influenced by disturbances with remote sources, might change over time without changes in the system setting itself. This poses two fundamental problems to identifying and classifying systems and their parts by observation: 1) if a system's behaviour has changed over time, will it be persistent or not, and 2) can the change be explained by the system's setting itself?

4.3.5 Dimensions of Eco-Complexity

The two main properties of engagement with eco-complexity in biodiversity research are changes and shifts in *environmental composition* and *spatial configuration* of biotopes or habitat patches (Andrade et al., 2021; Aronson et al., 2016). 'Composition' and 'configuration' are also two distinct concepts in urban ecology. While composition is regarding properties without spatial reference, such as number and proportion of elements, configuration is the spatial setting - arrangement, shape, contrast to neighbours, connectivity and anisotropy - of these elements (Wu and David, 2002; Li and Wu, 2007). Cadenasso et al. (2007) further distinguishes a triadic set of dimensions of eco-complexity as *heterogeneity*, *connectivity*, and *contingency* with further concretization at increasing levels of complexity.

Cadenasso et al. (2006a) describe *spatial heterogeneity* as the composition and configuration of patches. *Patch richness* refers to the abundance of different habitat patches (e.g. forest, grasslands, etc.), whereas *patch frequency* describes the ratio of habitat specific patch abundances. *Patch configuration* further distinguishes actual layout of patches, and their mutual adjacencies. *Patch change* describes the transition of individual habitat patches from one to another condition. *Shifting mosaics* finally describe the changing constellation of different patch types as discrete events or states.

Organizational connectivity describes intra- and inter-patch processes and interactions. *Within- unit processes* are confined within the boundaries of a single patch or unit. *Unit interactions* are happening directly between adjacent patches. *Boundary regulation* and *cross-unit interaction* describe processes with intermediate patches involved. *Functional patch dynamics* describe the overall interaction patterns with feedbacks at a systems level (Cadenasso et al., 2006a).

Temporal contingency addresses the level of simultaneity and immediacy of interactions that might occur. *Contemporary direct links* and *contemporary indirect links* both work at the same time interval, whereas *lagged links* have to be accounted to prior states of other patches. *Legacies* work directly but are caused through prior patch states, and *slowly emerging indirect links* are taking effect indirectly in space and time (Cadenasso et al., 2006a). In information theory such lagged, or memory effects are also referred to as 'non-Markovian' behaviour.

Table 4.9: Principles of ecological complexity from Zipperer et al. (2000).

Principle	**Definition**
Content	The structural and functional attribute of a patch where "structure" is the physical arrangement of ecological, physical, and social components, and "function" refers to the way the components interact
Heterogeneity	The spatial and temporal distribution of patches across a landscape. Heterogeneity creates the barrier or pathways to the flow of energy, matter, species, and information
Context	The patch's location relative to the rest of the landscape as well as the adjacent and nearby land units that are in direct contact or linked to a patch by active interaction
Connectivity	How spatially or functionally continuous a patch, corridor, network or matrix of concern is
Dynamics	How a patch or patch mosaic changes structurally and functionally through time
Hierarchy	A system of discrete functional units that are linked but operate at two or more scales, Proper coupling of spatial and temporal hierarchies provides a key to simplifying and understanding the complexity of urban landscapes

Although the definitions by Cadenasso et al. (2006a) are very detailed, there seems much ambiguity about describing connectivity through interactions and contingency through links. Another similar formulation of ecological complexity is found in Zipperer et al. (2000), which shows similar attributes, but expands 'heterogeneity' into three distinct categories of 'content', 'context', and 'heterogeneity', as seen in Table 4.9.

The reviewed principles of complex ecological systems have delivered four main components which need to be considered in EUA: 1) hierarchical coupling, 2) urban ecosystems as CAS, 3) patterns and processes, 4) dimensions of eco-complexity. These components need to be addressed by analytical methods. Therefore, the next section will summarize approaches to ecosystem modelling and important parameters.

5. Analysis and Modelling of Ecological Urban Systems and their Components

“The complex relationships between pattern and processes in all fields of ecology, their sensitivity to space, time and scale with a multiplicity of interrelated impacting factors require analytical methods and tools, able to cope with such high complexity in a nonlinear logic, approximating complex systems under uncertainty and dynamic behaviour (Stupariu et al., 2022).” The presented theoretical concepts explain basics of urban biodiversity and how important spatial factors are linked to more abstract concepts of ecosystems as complex system. This shows how detailed questions in multi-species urban design, such as the ecological equivalence of urban habitat patches, could be analysed by few general principles of eco-complexity. This also creates a set of characteristics and requirements for the applicability in the analysis and modelling of urban social-ecological systems.

5.1 Deterministic Ecological Modelling

“Ecosystem models attempt to incorporate ecosystem components [...] and processes [...] into one modelling framework (Geary et al., 2020).” These models aim at mimicking ecological processes of at least two ecosystem components (Geary et al., 2020). Ecosystems are represented as interactions between different components and processes. Geary et al. (2020) distinguish several objectives of ecosystem models as 1) description and understanding of current ecosystems, 2) forecast or hindcast scenarios, and 3) decision on management actions. Furthermore, it is of key importance to select a modelling approach according to the objective of research. However, only few of those ecosystem models are accounting for spatial heterogeneity and connectivity (Geary et al., 2020). Topology and adjacency of neighbouring habitat patches do have a significant impact on ecological phenomena such as biodiversity (Colding, 2007). Spatial explicitness further implies expressing the relational position of two neighbouring patches logically.

Different approaches to logic demonstrate trade-offs in addressing real-world complexity. While mathematical logic is establishing precise relationships, with general applicability and transferability from one discipline to another (Desjardins-Proulx et al., 2019), it is incapable to account for stochasticity. Logics of uncertainty can account for randomness but still only offer a binary way of description. Logic of vagueness is introducing a probability of an observation being true or false (Desjardins-Proulx et al., 2019; Geary et al., 2020). While from a standpoint of logics, fuzzy approaches might be more correct, they are of little use to describe ecosystems for design and planning tasks.

The presented aspects of modelling ecosystems are general guidelines for research in urban ecological systems. Ecosystem modelling generally bears trade-offs between specificity, being able to describe highly complex processes and general applicability, allowing transfer and extrapolation of knowledge at the cost of complexity (Geary et al., 2020). Uncertainty in ecosystem modelling typically rises with structural complexity, i.e. with a growing number of variables to account for, but also models of low complexity are prone to biases in assumption of influential parameters, component relationships, and scale selection (Geary et al., 2020). That means, that it is to be expected that stipulated principles underlying the formulas, are only true in most cases (Desjardins-Proulx et al., 2019).

Geary et al. (2020) state varying levels of abstraction, at which ecological models work:

> 1. Establishing interaction networks as an unparametrized graph representation;
> 2. Probabilistic interaction events (e.g. Bayesian belief network);
> 3. Simple parametrization of interactions through weights and regression calculations;
> 4. Deterministic dynamic system representation.

There are several major components or *state factors* relevant for the analysis of ecological systems. A comparison of three authors is given in Table 5.1 (Pickett et al., 2011; Wentz et al., 2018).

Table 5.1: Comparison of ecological components.

Pickett et al. (2011)	**Wentz et al. (2018)**
-	Human constructed elements
Soil characteristics	Soil-plant continuum
Resident organisms	-
-	Water elements
-	Two- and three-dimensional space
Terrain characteristics (relief, elevation, slope, aspect)	Spatial pattern
History of the system	Time
Contemporary climate	-

5.2 Hierarchy

As results have shown, hierarchical scales play an imperative role in biocomplexity theory as well as ecosystem assessments (Wu, 1999). Structural, multi-scalar constraints are an important consideration when planning for biodiversity in urban ecosystems (Andrade et al., 2021). It helps to move away from a normative ‘optimum’ to an operative set of relative conditions, opportunities and limitations facilitating the choice of target communities and possible or needed improvements in species composition and biotope characteristics.

However, when addressing a specific ecological aspect, such as species richness in urban areas, there should be a specified *focal* scale from which relationships across scale can be examined (Ascher, 2001; Wu and David, 2002). The focal scale is dependent on the expected outcome, i.e. the ecosystem process under investigation (Wu and David, 2002). It has been suggested that the meso-scale, which lies between the city level and individual plots (e.g. neighbourhood or district level), is suitable for spatial analysis in ecological contexts (Pauleit and Breuste, 2011). With the focal scale spatial analysis also has to define the *extent* (i.e. the chosen system boundary), as well as *spatial units* (discrete, hierarchically structured parts of the system) (Wu and Qi, 2000). Spatial units can also show overlaps. When referred to a specific species, home ranges can be used to select core areas for analysis (Oh et al., 2011).

But if interactions of patterns and processes are only observable at same or similar spatio-temporal scales (Wu and David, 2002), measurements must properly represent pattern and process at the desired scale of interaction. A conceptual framework was proposed by Wu (1999), summarized by Wu and David (2002), as the *scaling ladder approach*:

1. Identifying appropriate patch hierarchies: this step involves the decomposition of the complex system either by a top-down partitioning or bottom-up aggregation.
2. Making observations and developing models at focal levels: ecological processes can be studied by choosing a proper grain size, i.e. the spatio-temporal resolution, as well as the extent of observation
3. Extrapolating information across domains of scale hierarchically

The problems behind the scalarity of ecological processes has been addressed as 'modifiable areal unit problem' (MAUP) and 'scale effects' (Wu and Qi, 2000). This problem refers to the sensitivity of results according to the definition of geographic units (Nikparvar and Thill, 2021; Liu and Biljecki, 2022). *Uncertain geographic context problem (UGCoP)* refers to the sensitivity of results according to the changing delineations of contextual units (Nikparvar and Thill, 2021). Contextual influences change across data aggregations (Nikparvar and Thill, 2021).

Frequencies also play an important role in finding the appropriate scale and proper metrics to describe pattern and processes. Wu and David (2002) state, that for any hierarchical level with a total time span of investigation of T and the time span or frequency of observations of τ, any process significantly slower than $1/T$ can be assumed to be constant at the focal level, whereas any process with a frequency much faster than $1/\tau$ can be regarded as 'noise' and it is sufficient to assume averaged behaviour as influence at the focal level.

Scaling information received at a focal level must consider non-linear behaviour in space and time (Wu and David, 2002). A general approach to scaling, i.e. changing grain and extent of investigation, is to link models hierarchically. This is done by using output from one model as input at another scale (Wu and David, 2002). Besides choosing specific scales of investigation it is also imperative to be able to relate different scales of research with each other (Wu and Qi, 2000). This has profound consequences on the measurements of space, and requires a spatial system which is either scalable, or at least organized by hierarchical, aggregate measurements.

Farinha-Marques et al. (2017) have developed a method for habitat classification to contribute to the issues of spatial diversity and complexity within urban areas. Their method is applicable at a microlevel, attributing to the vegetational structure within a habitat patch. This approach shows how fixed spatial scales can be too rigid to give a valid classification for different species. Instead, scales need to be chosen in accordance with the home range of the targeted species or communities, as for smaller species, microhabitat structuring and patch composition are relevant factors (Farinha-Marques et al., 2017).

5.3 Spatial Analysis

Traditional approaches to modelling ecosystems, such as land use change or urban growth modelling, usually separated spatial and non-spatial data. The improvement of remote sensing in combination with the development of geographic information systems (GIS) facilitated the integration of spatial and other factors (Chaturvedi and de Vries, 2021). Hence, spatially explicit models have gained importance for making management decisions (Geary et al., 2020). Analytical methods for addressing spatial characteristics of heterogeneity are *ecological patch assessment*, *eco-profiling* or *ecotope classification*.

Spatial dependence, or *spatial autocorrelation*, also referred to as the first law of geography as “near things are more related than distant things” (Tobler, 1970), poses a basic assumption in spatial data analysis. *Stationary* spatiality refers to a relationship of data, which is explained only by their relative configuration (e.g. distance as scalar unit). *Non-stationary* spatiality on the other hand considers that relationships might change due to the absolute position of data in a reference system (Goodchild and Janelle, 2004). Spatial non-stationarity describes the variation of residuals due to geographic (place-related) factors, seen in Figure 5.1. Spatial dependence has also implications for the sampling of data (cluster proximity) (Jemeļjanova et al., 2024).

Spatial analysis requires the determination of a spatial organisation. The basic unit can either be an *object*, a *border*, or a *place* (Wang et al., 2010). Additionally, there is the possibility of producing spatially discrete units derived from a reference system by either using point patterns, geostatistical data or lattice data (Cressie, 1993). Geostatistical data and lattice data bring the advantage of having identical units, with constant spatial relations with each other. These considerations also emphasize an underlying problem of classification in spatial context.

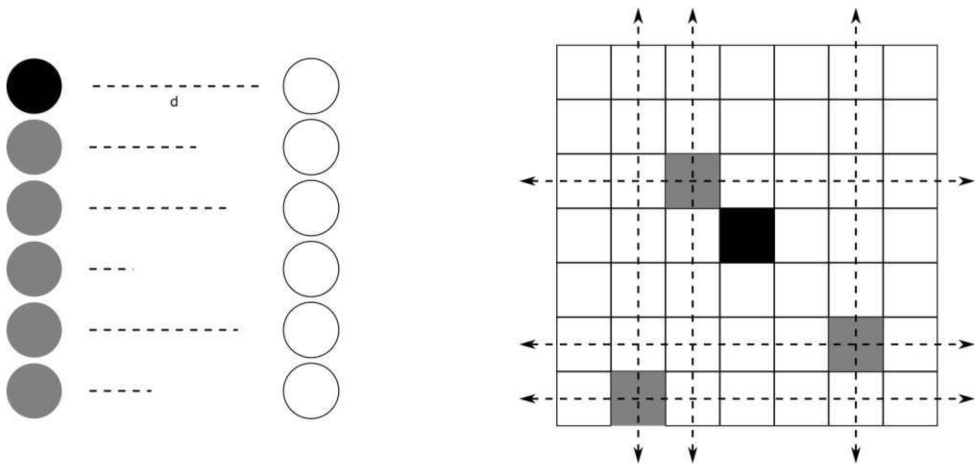

Figure 5.1: Left: spatial stationarity assumes no influence from the location of observation, and is often expressed as adjacency or distance between units. Right: Spatial non-stationarity assumes variance depending on the location of observations, and needs to be addressed by spatial co-variates.

The most coherent way to build basic units is to tile space into a gridded structure, where every grid tile has a uniform shape. Such information can be represented as pixel in two-dimensional or voxel in three-dimensional space (Cressie, 1993). As pixel or voxel information is the main data source for many kinds of data (e.g. solar irradiance, elevation models, etc.). The ECOLOPES approach uses voxels to translate and evaluate building designs in their ecological models (Tyc et al., 2023). Such spatial information is comparably easy to process, while the entire plane space is accounted for, but on the other hand there exist trade-offs.

Grid-based approaches have several advantages compared to entity-based ones. Grids are abstractions of space which are independent from real-world complexity. They are equally sized, evenly distributed and adjacent, without overlaps, can be scaled and interpolated according to evaluation purpose, and are therefore easier to compare and evaluate (Tyc et al., 2023). On the other hand, they do not represent real-world borders and linkages, resulting in the loss information (average values, details lost etc.) or the introduction of bias (Cressie, 1993; Nikparvar and Thill, 2021). They also give little information about the scale dependency of evaluated ecosystems.

The main problem are mismatches of the extent of the grid tile and the extent of information within. For example, ground cover might vary a lot inside a single grid tile, if tiles exceed certain sizes. As the grid tile represents the smallest unit of analysis, i.e. one grid tile is one data point, information needs to be aggregated by means of calculating averages or finding dominant features. The main problems with mismatches in extent of grid tiles and ground truth are depicted in Figure 5.2.

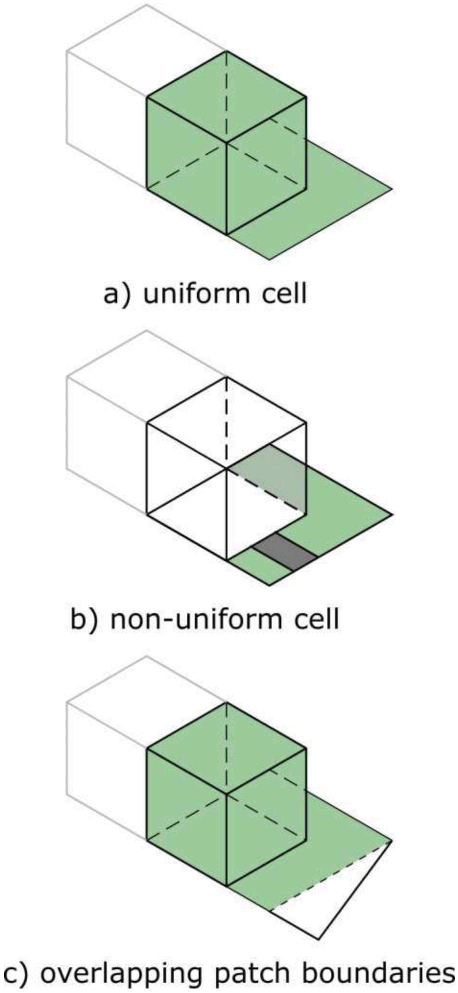

Figure 5.2: Possible mismatches between gridded data and ground truth can lead to biased results: a) uniform patch properties in a cell are favourable for analysis, but at larger scales unlikely as landscape fragmentation in urban areas is high, b) non-uniform patches need to represent different properties either as average or some kind of metric, c) overlapping patch boundaries are mismatches between biophysical or immaterial boundaries and cell sizes.

On the other hand, using real-world entities as basic units, it is easier to form concise units. But spatial relationships are much more difficult to describe, because adjacencies with other units cannot be simply expressed by a determined number of uniform shared edges but need to be evaluated separately or solved topologically. A third approach is the use of geographic regions, formed by immaterial boundaries. Demographic divisions should be considered, as they often represent similar socio-economic conditions, but also landscaping and management types (Geronimus and Bound, 1998). These trade-offs need to be considered when defining a basic unit for analysis. As it is more common to use a grid structure, uniform cells need a set of metrics to describe characteristics of ecological and spatial nature.

The specific representation of spatial data as points, lines, polygons, grids, etc. and the measurement carry along potential difficulties through information loss or generalization (Nikparvar and Thill, 2021). Three-dimensional point data is heavily influenced by the measurement technique (distance of recorded point from sensor) in its densities (Nikparvar and Thill, 2021). The mapping of three-dimensional point data into two-dimensional space entails problems if there are multiple z-values (third dimension) for a point, and irregularly spaced points. Methods to address the mapping into two-dimensional space include *voxel feature encoding (VFE)*, or the projection of point clouds on multiple synthetic two-dimensional images and the labelling of pixels (Nikparvar and Thill, 2021). It is important to notice that urban metrics are not independent of domain specific perspectives. Hence, the development of appropriate metrics for interdisciplinary research is still an unsolved issue (Ramalho and Hobbs, 2012).

Spatial heterogeneity refers to the variability of observed patterns due to interactions, dispersion, diffusion and exchange. Heterogeneity is also influenced on multiple spatial and temporal scales, where intensity and duration play significant roles (Nikparvar and Thill, 2021). Spatial references, entities and phenomena carry additional information (within-object and between-object information) which can be used as features for an observation matrix (Nikparvar and Thill, 2021). Between-object information includes connectivity, contiguity, distance, association, and direction (Nikparvar and Thill, 2021).

5.4 Network Analysis

The concept of landscape connectivity is based on *graph theory*. Graph structures are built from *nodes* or *vertices*, representing an entity or basic unit, e.g. an individual patch, and *edges*, which signify the connection and its quality between the nodes (Liu et al., 2024). There are different methods for evaluating graph structures. *Least-cost path* analysis can identify potential corridors with least resistance values from a source patch to all others (Kong et al., 2010). *Gravity modelling* assesses the interactions between nodes, with the level of interaction representing efficiency of corridors and significance of linked nodes (Kong et al., 2010). Figure 5.3 shows a translation of urban morphological elements into different graph representations, analogous to Figure 3.1 (Marcus et al., 2019b). Graphs can either represent *meta patches*, i.e. all habitat patches which are accessible by adjacency or proximity, or *corridor networks* representing possible vectors of movement (Marcus et al., 2019b).

Coming from an ecologist's understanding of landscape as an arrangement of patches, with emerging patterns, may seem somewhat familiar to architects and urban designers. In fact, Hillier (2007) draws this metaphor of spaces and forms emerging as social-cultural patterns. In fact, the study of *space syntax* offers some valuable tools for analysing urban networks in both qualitative and quantitative ways. Berghauser Pont et al. (2017) have found that visual distances between habitat patches have greater importance for the ability of species to move around than Euclidean distances. With respect to this, we can further distinguish patch connectivity and 'visual connectedness' of patches (Kaczorowska and Pont, 2019).

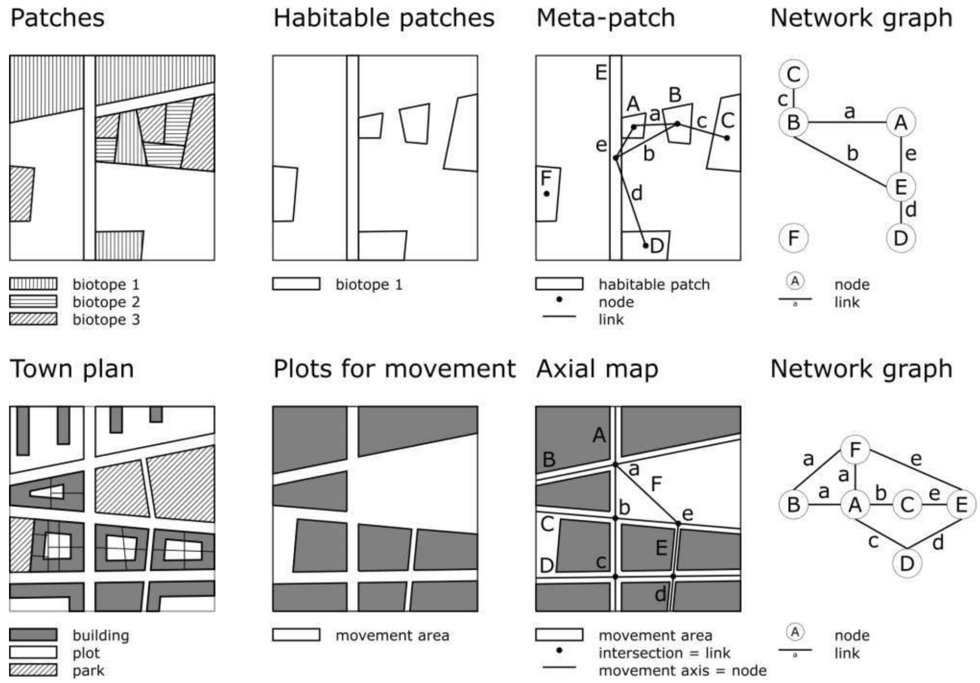

Figure 5.3: To represent ecological networks, either patch-based or axial graphs can be used, after Marcus et al. (2019b).

Urban ecological network planning consists after Oh et al. (2011) in three main steps: 1) green space assessment, 2) ecological network development, 3) examination of land use plan. As it has been argued before, urban habitats are heterogeneous and more recent studies support a differentiated perspective on which urban patches can be considered as potential habitats. The aim of the green space assessment is to classify core and buffer areas, as well as developmental risk areas. Common forms of analysis of networks are *connectivity network analysis* (Kaczorowska and Pont, 2019), graph theory and gravity modelling (Kong et al., 2010), or *network cohesion* (Opdam et al., 2003).

5.5 Dynamic and Agent-Based Analysis

Bascompte (2009) explains how wildlife interactions can be represented as ecological networks. At any scale, connectedness between agents is the precondition for interaction, and hence, processes in complex systems (Kaczorowska and Pont, 2019). Within urban socio-ecological systems patterns emerge from either geophysical, biophysical or socio-economic activity. The smallest unit of such activity is commonly addressed as 'agent' (Batty, 2005). Agent-based modelling is a type of modelling approach in ecosystem research, which is investigating individual agents and their behaviour, trying to abstract emerging patterns at bigger scales. These patterns are determining the emergent ecosystem functions, which are only observable at aggregate levels of analysis (Wu and David, 2002; Hanna, 2022). Reciprocally, big-scale emergent functions do inform the behaviour of individual agents within the ecosystem, closing a circular loop representing the whole ecosystem (Wu and David, 2002).

The Entrainment Hypothesis states that "within any one ecosystem, time-series data for biotic variables should have periodicities that cluster into a small number of sets, reflecting the generation time of one of the critical structuring variables, i.e., periodicities (or frequencies) should be discontinuously distributed in a predictable way (Holling, 1992)." Common methods for analysing dynamics of ecosystems are *community detection* (Lu and Yang, 2022), *species presence prediction*, *agent-based modelling*.

A key mechanism described in landscape ecology is the exchange or interaction between habitat patches. Such habitat patches can be, analogous to organisms, viewed as 'agents' with internal development impacting other patches. Batty (2005) presented a concept of cities as system of cells, or agents, which is vital for an understanding of abiotic entities as agents in a spatial context. Such agents and their connectedness can be represented by graph theory (Marcus et al., 2019b). This puts emphasis on networks as being able to project the outcome of dynamics and interactions into a static representation.

A useful aspect comes from *space syntax*, (Hillier, 2007) as it tries to abstract human activity to more general functional notions, such as 'movement' without the need to define the specific mode. Similarly, it is argued, that by applying more abstract notions of habitat functionality and traits, ecologically informed design will be facilitated through higher principles. This grants the opportunity to move away from the requirement to define target species and optimize only for those, towards a more general notion of supporting biodiversity (Marcus et al., 2019b).

Hence, dynamics can either be expressed as a 'change' of elements or networks, or through agent-based observations. As described above, one aspect of complex ecological systems is, that they are influenced by present, as well as past conditions and events (Aronson et al., 2016). To address dynamics (i.e. change and interaction) in space and time, Ramalho and Hobbs (2012) proposed a 'dynamic urban framework' based on the assumption that a static assessment of urban-rural gradients is insufficient to account for temporal complexity of urban remnant configurations. They defined *remnant age*, *past remnant and landscape attributes*, and *landscape fragmentation drivers* as three core factors which can be developed from discrete time series of data.

6.

A Conceptual Framework for Ecological Urbanistic Analysis

This chapter presents an interim discussion of the results of the first research component “ecological urbanistic analysis - conceptual framework”. The review of theoretical backgrounds on biodiversity, urban ecology and possible interfaces with urbanistic analysis has brought up a plethora of disciplinary concepts, which could only be presented in short.

6.1 Strengthening EUA through an Integrative Framework

After establishing the theoretical implications of eco-complexity for modelling and analysis, foundational considerations of biodiversity in cities, and a translational descriptor between ecology and urban design have been established, this section aims, as an intermediary step, at conceptualizing a framework for Ecological Urbanistic Analysis, integrating the presented theoretical background by establishing practical guidelines. From the results of the ecological literature, it is stipulated that due to the multiple complexities of biodiversity in urban ecosystems, a differentiated analytical approach is needed. The wide variety of potential species populating urban habitats, with their individual traits and inter-specific dynamics within communities, necessitate to define what constitutes similar conditions according to the goal of intervention.

The results from the reviewed literature about biodiversity in urban contexts suggests that there are several factors influencing the occurrence, composition and abundance, as well as long-term viability of specific populations and communities within cities. While there are many detailed factors concerning inter-species interactions, life history of species and others, there are several important spatial factors, which can be addressed directly through urbanistic analysis. But the common approaches in urban and landscape ecology, to analyse aspects of biodiversity are mainly oriented towards vegetation pattern and habitat sizes, which give little input for urban designers and architects to work with. It is therefore of interest to comprehend the relationship of organisms within and as part of ecosystems on a more abstract level.

6.2 Boundary Conditions and Requirements of EUA

From the reviewed literature several reoccurring patterns across the scales, from biodiversity as real-world occurrence of species within cities, over important abiotic factors of landscapes, to the general principles of ecosystems as complex systems, can be identified. The principles of eco-complexity as presented in Section 4.3, especially hierarchy theory (Wu and David, 2002) and the frameworks of biocomplexity (Cadenasso et al., 2007; Zipperer et al., 2000) offer vital concepts to address complex ecological systems at a theoretical level. Several boundary conditions and requirements can be derived as follows.

Cities must be thought of as social-ecological systems, where there is a constant exchange between humans, non-human organisms and the abiotic environment (Alberti et al., 2003). This still leaves us with a plethora of relationships and aspects to account for. But on an even more abstract level, pattern, processes and functions of ecosystems can be described by a set of concepts with growing level of analytical complexity. The four main dimensions that need to be addressed are: 1) hierarchy 2) heterogeneity, 3) connectivity, and 4) contingency (Cadenasso et al., 2006a).

The three dimensions of eco-complexity are deeply interwoven, and pattern and processes are always a result of their interplay. As urban designers and architects are used to work with spatially explicit information, analysis of ecological patterns and processes should be expressed through spatial representation. But spatial concepts are also domain specific and there seems to be a gap between traditional urbanists and ecologists approaches to space.

As described in Section 4.3.1, the focal scale mainly depends on the pattern, process, or species and their behavioural traits under investigation. But further, available data and the desired basic unit of analysis need to be considered (Li et al., 2019), as properties of ecological entities need some kind of organisation to be described in space. This definition of basic units bear important trade-offs.

While spatial properties are characterised by basic units, the relationships between these basic units, need to be addressed as ecological networks. An ecosystem is constantly evolving over space and time (Bestelmeyer et al., 2017). This means that spatial heterogeneity and connectivity are changing with the location and the state of a system. Future system states are informed by current and past interactions and processes (Ramalho and Hobbs, 2012). This has to be accounted for by non-stationary and dynamic modelling approaches (Goodchild and Janelle, 2004).

To substantiate the framework, an extended definition of EUA could be as follows: EUA is the investigation of ecological complexity within urban socio-ecological systems with spatially and temporally explicit methods. Urban social-ecological systems and their coupled patterns and processes are described by addressing spatio-temporal heterogeneity, contingency and resulting connectivity, by determination of constraining factors at higher scales, and the emergent, non-linear behaviour of agents at lower scales.

6.3 Observables of EUA

The presented frameworks by Cadenasso et al. (2006a) and Zipperer et al. (2000) provide a platform on which particular models, with regards to their specific research question, can select variables for investigation. While such a framework enables an overview of available perspectives and factors, it does not imply connectedness amongst all of its constituents (Cadenasso et al., 2006a). Hence the selection process has to be articulated under consideration of the objectives of analysis. The applicability of the presented dimensions of ecological complexity is depending on a mid-level definition which is commensurable by standards of applied sciences. Such a bridging concept is important to relate low-level concepts in multi-species design to higher principles of eco-complexity.

Regarding the specific requirements and concepts for biodiversity in urban areas as a basis for multi-species urban design, and their relationship with higher-level ecological principles, scale coupling and hierarchy theory play a key role. The concept of hierarchical coupling allows for observing processes and patterns of ecosystems at different level of abstraction (Wu and Loucks, 1995; Wu and David, 2002). Following this, the principles of eco-complexity adhering to spatial heterogeneity, connectivity, and contingency (Cadenasso et al., 2006a) have validity on any scale of observation. Where detailed aspects of interest are exceedingly complex and interdependencies to other aspects are unknown, abstractions of such problems to more general rules of ecosystems can provide an analytic interface.

Observables are a concept in physics describing properties or characteristics which can be measured, it is reasonable to abstract from applied ecosystem analysis such observables to be addressed through EUA. Hence, I will borrow this concept as an analytical tool, to describe the representation of real-world phenomena at different spatio-temporal scales through underlying, scale-invariant characteristics of complex systems.

Scale is an essential aspect of hierarchically nested complex systems. Hierarchical patch dynamics suggest that interactions of processes and patterns can only be observed, if both operate on the same or similar spatio-temporal scales (Wu and David, 2002). Scale specific effects need definition of a focal scale depending on the object of investigation (Wu and David, 2002). Small-scale and big-scale phenomena are coupled hierarchically, where smaller-scale patterns and processes are only observable as ‘noise’ on a focal scale, and higher-scale patterns and processes form constraints (Wu and Loucks, 1995). See Figure 6.1.

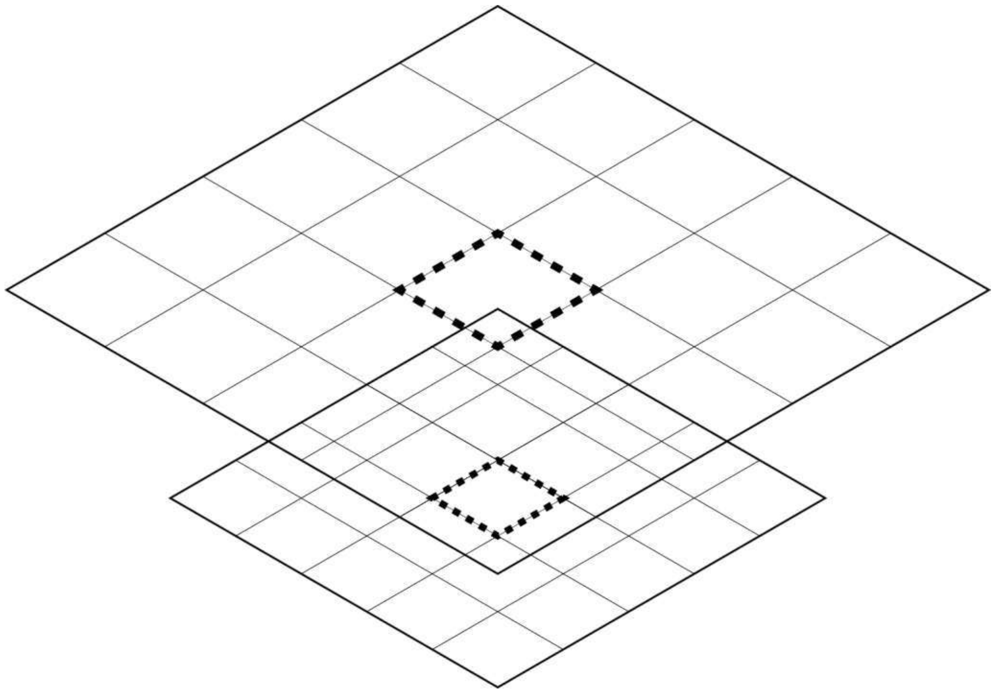

Figure 6.1: The observable *scale* describes a focal level of spatio-temporal resolution.

Extending the concept of Cadenasso et al. (2006a), *composition* refers to all within unit properties, as well as their richness with regard to the overall variation of the units of analysis. See Figure 6.2.

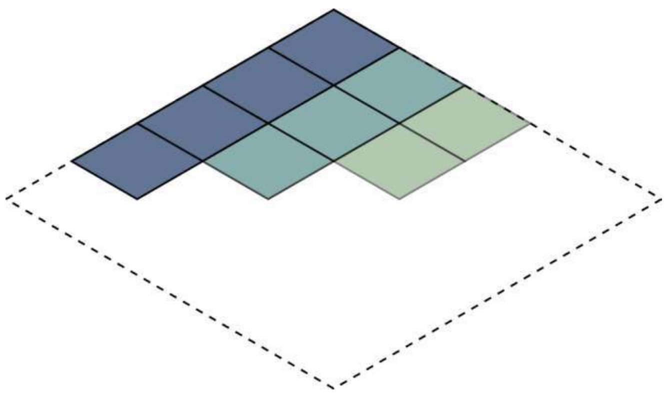

Figure 6.2: The observable *composition* describes the ecological variation within an urban ecosystem.

Configuration refers to the spatially and temporally explicit setting of all units (Cadenasso et al., 2006a), whereas configuration implies non-stationarity. See Figure 6.3.

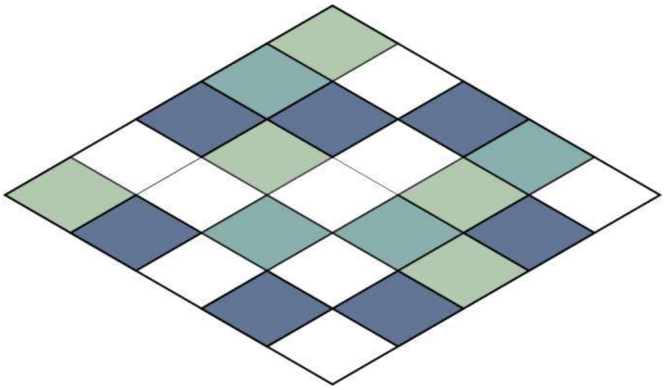

Figure 6.3: The observable *configuration* describes constellations of basic units within a referential system.

Connectivity must be accounted for in spatial and temporal dimensions, as it is the precondition for interaction. Hence, *connectedness* refers to the spatially and temporally explicit links among all units (similar to Kaczorowska and Pont (2019)). Connectedness can be influenced by configuration and contingency (freezing of lakes as corridors).

Connectedness of geophysical flows and species dispersal influences the ecological conditions of habitat patches via indirect and lagged effects (Cadenasso et al., 2006a; Chase et al., 2020). Connectedness also allows for the spread of disturbances, such as human traffic or diseases. This leads further to the assumption that negative impacts, adverse processes or barriers on adjacent or 'bridging' patches may also affect neighbouring habitat patches and must therefore be attributed in the evaluation (Colding, 2007). See Figure 6.4.

Interaction refers to processes between units of analysis, which can in this context also be seen as agents. However, within-unit processes will not be observable at the focal scale, other than as emergent properties. See Figure 6.5.

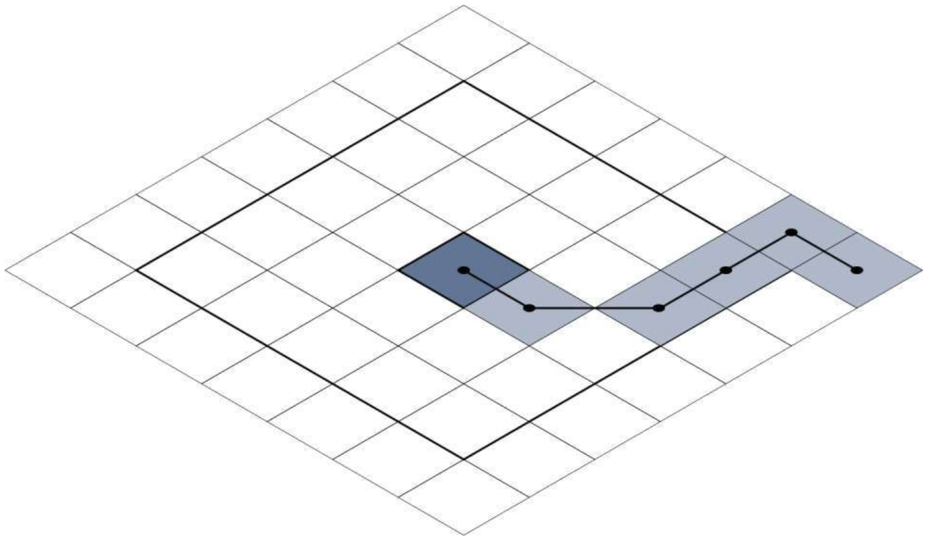

Figure 6.4: The observable *connectedness* accounts for all existing links between basic units of analysis. It emerges from configuration and is a precondition of interaction.

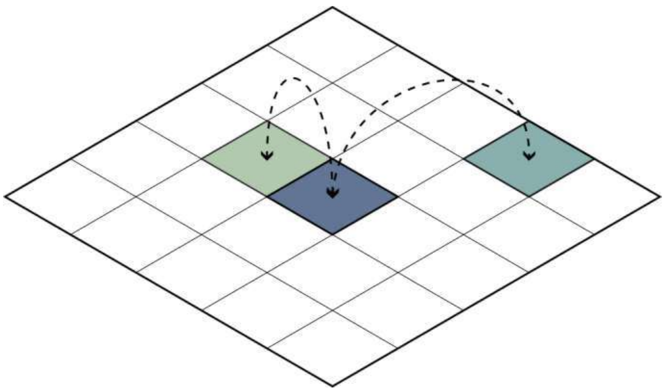

Figure 6.5: The observable *interaction* describes the influence of basic units amongst each other. Interaction might lead to change of a basic unit.

Change refers to the difference in composition, configuration, connectedness, and interaction over time (Ramalho and Hobbs, 2012). Change can only be observed as a differential between two discrete time steps. See Figure 6.6.

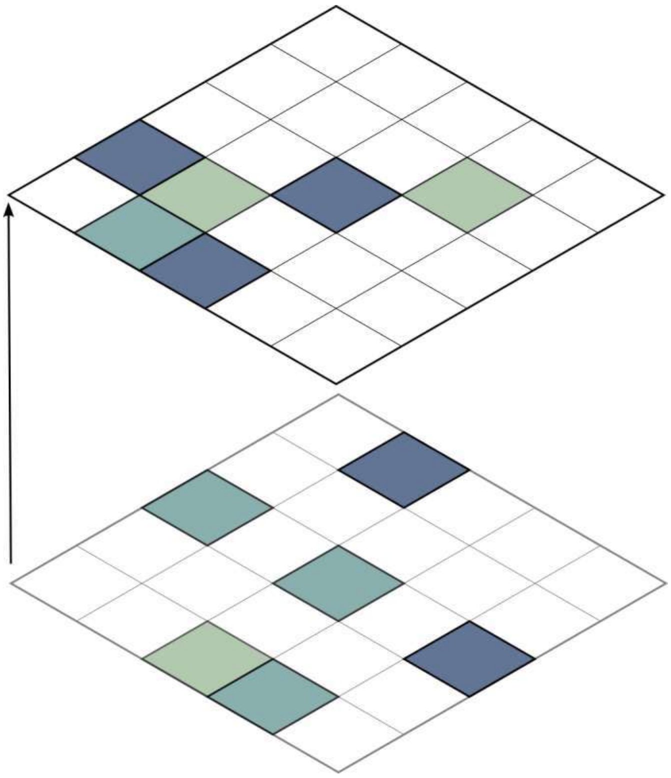

Figure 6.6: The observable *change* addresses the differences which ecological systems and their agents produce in time.

As system is defined by its current characteristics, but also by previous system states, legacy effects and transient dynamics need to be factored in, due to differing development trajectories (Ramalho and Hobbs, 2012). *Memory* refers to the persistence of past interactions and change in composition, configuration and connectedness of units (Brierley, 2010). See Figure 6.7.

Although composition, configuration, connectedness, and memory could be described as state variables, and interaction and change as temporal variables, there should be no strict distinction, because all aspects are connected in space and time. From composition to memory, there is increasing complexity in the spatio-temporal analysis, as composition is comparably easily identifiable through measurement, while memory entails the concept of *entropy* and *legacy effects* over long time spans. The integration of the observables into the frameworks presented in Section 4.3 is shown in Figure 6.8.

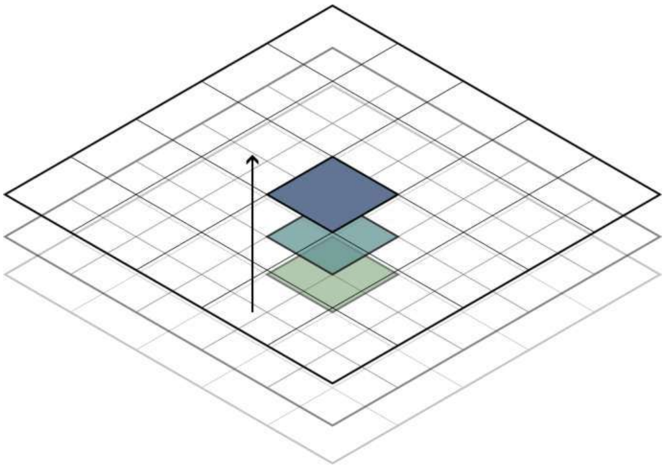

Figure 6.7: The observable *memory* refers to past pattern-process-dynamics, which are not active anymore, but have transformed the ecological system profoundly.

6.4 Application Protocols

As EUA is not an end in itself, there needs to be a specific research question to apply the proposed framework to. Such a question might be concerned with the implementation of design or planning projects. Adopting the three protocols from Pickett et al. (2016), we can limit the application of Ecological Urbanistic Analysis to:

- Well understood and defined processes that can be tested against changes in urban variables and environmental influences. This could be a biodiversity indicator in comparison to sealed surfaces. A typical application is regression analysis.
- The importance of individual variables for the outcome of processes can be tested through statistical models. If it is unknown if, and to what degree a variable is impacting a phenomenon (i.e. an observed variable), several methods exist to test and quantify such relationships. *Principal component analysis* (PCA) is amongst the most common methods.
- Ecological processes and variables can be tested against human decision and management outcomes. This application directly tries to establish a relationship between observed phenomena after human intervention (without investigating individual variables changed).

For the integration of the postulated observables into an empirical analytical task, a general workflow should include the following steps (Liu et al., 2024; Löfvenhaft et al., 2004; García-Pardo et al., 2023):

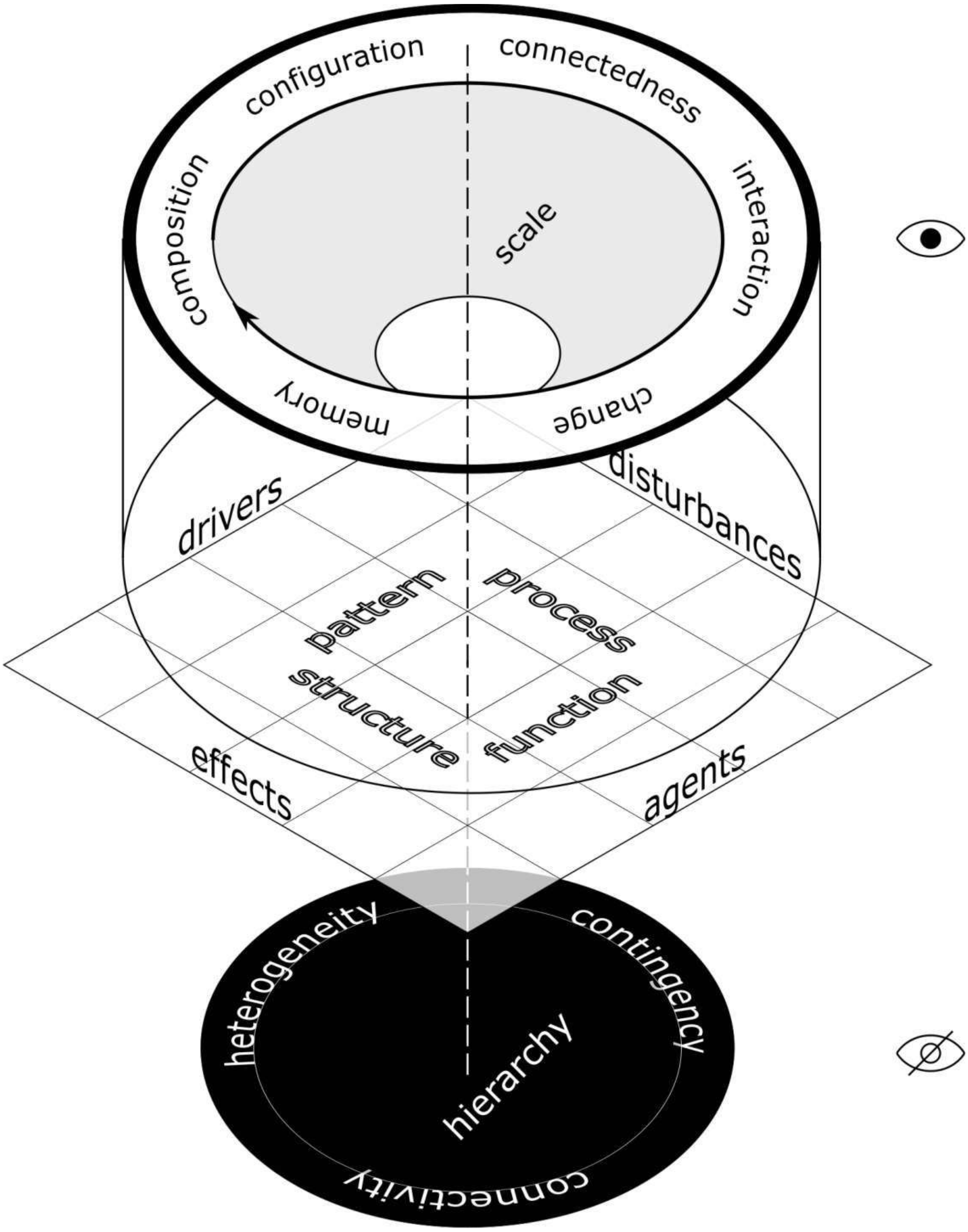

Figure 6.8: The integration of the postulated *observables* of EUA with the reviewed concepts describing complex ecological systems. While the principles of eco-complexity, as described by Cadenasso et al. (2006a) are not directly observable, and the individual patterns, processes, effects, drivers, agents, and disturbances can be conceived of as 'noise', the observables work as an abstraction lens to focus EUA by describing complex processes at a meta-level which is still addressable by common analytical methods.

1. Defining objectives and targets of biodiversity or multi-species design;
2. Defining analytical methods: 1) habitat suitability 2) ecotopes, 3) network analysis, 4) agent-based modelling, 5) patch dynamics;
3. Choosing an analytical approach: geographic units, graph structures, sequences;
4. Defining the study area and its extent in spatio-temporally explicit terms;
5. Defining a focal scale in accordance with a meaningful basic unit;
6. Defining variables of concern and data sources;
7. Execution of analysis;
8. Interpretation and Synthesis.

Tasks in ML are the different use cases for ML models. Which task can be done and how well the model can learn from the training data depends on many factors. Machine Learning offers support for a range of core tasks in urban design and planning, as well as ecology (Casali et al., 2022; Stupariu et al., 2022). Depending on the application and the purpose of research, several applications of ML can be differentiated:

1. Dimensionality Reduction of high-dimensional data;
2. Extracting features from background data, i.e. the detection of features and/or the segmentation of data into feature and background;
3. Mapping entities (artifacts, organisms, etc.) with respect to a spatially explicit system;
4. Regression analysis and causal inference to explore dependencies and influences amongst variables;
5. Classification and Clustering, i.e. the organization of datapoints into distinct sets;
6. Monitoring changes amongst datapoints within sequential datasets;
7. Prediction of unseen data or future states of datapoints;
8. Generating new datapoints from learned patterns for data augmentation.

Nikparvar and Thill (2021) list several applications of ML considering spatial and temporal factors as: 1) land use and land cover classification, 2) cross-sectional characterization, 3) longitudinal change, 4) urban growth, 5) gentrification; 6) disaster management, 7) agriculture and crop yield, 8) prediction, 9) infectious disease emergence and spread, 10) transportation and crash analysis, 11) map visualization and cartography, 12) delineation of geographic regions, 13) habitat mapping, 14) geographic information retrieval and text matching, 15) POI and region recommendation, 16) trajectory and movement pattern prediction, 17) point cloud classification, 18) spatial interaction, 19) spatial interpolation, and 20) spatio-temporal prediction.

7. Machine Learning in the Context of Ecological Urbanistic Analysis

To engage with ecological knowledge through machine learning methods, complexity must be addressed at a conceptual, abstractable level which needs to be represented by the logics of machine learning algorithms. To this ends the presented framework for EUA has deducted seven spatio-temporal dimensions, which are of importance to approximate complex ecological systems through analysis.

The field of machine learning has developed exponentially over the last few years. This section is supposed to be an exploratory introduction into state-of-the-art Machine Learning in the context of EUA. Emphasis will be given to properties and theoretical capabilities of ML algorithms. It is therefore not comprehensive and can only highlight aspects which are important for the understanding of potentials and challenges for applications in EUA for multi-species design. This chapter is structured as follows: 1) Establishing conceptual foundations of ML, 2) a typical ML workflow, 3) an overview of spatio-temporal ML, and 4) a summary of ML algorithms which have found broad application in ecology or similar disciplines and have interesting properties with regard to the requirements of EUA.

In contrast to classic ecological modelling techniques, which tried to model suggested causal links of phenomena into mechanistic and deterministic mathematical functions, Machine Learning is *data-driven*, the causality and dependency within data is derived by inference and approximation (Casali et al., 2022). In contrast to deterministic modelling, ML approaches center around generic algorithmic structures which use probabilistic methods or can be trained on approximating patterns within input data by minimizing a general *loss function* [Pichler and Hartig (2023); Rubbensetal2023]. While probabilistic modelling assumes a certain distribution of data (e.g. gaussian or binomial), and therefore is able to address uncertainty by metrics (confidence interval or ρ-values), the lack of this assumption behind other ML algorithms, such as artificial neural networks (ANN), required the development of non-parametric approaches for estimating uncertainty (Pichler and Hartig, 2023). An important distinction of ML methods to conventional classification or analysis is that data (besides preprocessing, normalization and vectorization) does not need to be processed and converted into predetermined metrics or indicators, but can be, or even should be fed into the models directly (Hanna, 2022).

Machine learning is a subset of artificial intelligence, and deep learning is a subset of ML itself. Machine learning is an artificial intelligence technique using different algorithms to train abstract computational models in recognizing patterns inherent to data describing real-world patterns and processes, enabling inference and prediction (Casali et al., 2022). Mitchell (1999) gave a definition for machine learning as: "a computer program is said to learning from experience E with respect to some class of task T and performance measure P, if its performance at tasks in T, as measured by P, improves with the experience E (Mitchell, 1999)." Although this definition does not mention data as a variable, there are differences in which type, and which amount of data can be handled by different ML algorithms. Hence, there is an interdependency between the task T which corresponds to the desired output O, the algorithm A, and the input data D defining which performance measure P is best suited for evaluation and what improvement can be expected through experience.

Traditional modelling is concerned with combining explanatory variables as functional algorithms to make truthful representations or predictions. Although there is a tendency in research towards predictive ML modelling, another fundamental strain of research is committed towards the explanatory capabilities of ML via causal inference (Pichler and Hartig, 2023). In Machine Learning, the approach to being able to explain an outcome or to predict it, are generally separate approaches. *Explanatory modelling* tries to identify variables that have significant relationships with an outcome, which requires to infer causal dependencies. The performance of algorithms in *predictive modelling* on the other hand, is not necessarily determined by the ability of the algorithm to infer causality (Pichler and Hartig, 2023).

Another important distinction of ML architectures is between *discriminative* and *generative* ML (Jung, 2022). Prior is learning to "divide" data into sets according to similarity or other patterns, latter is trying to learn how to create new instances of data which could pass as real.

The recent surge in Machine Learning was facilitated by the exponential increase in big data availability (Kitchin, 2014; Hanna, 2022). Still, there are restrictions to the availability of uniformly distributed and standardized data in quantities required by most ML algorithms (Reichstein et al., 2019). Hence, new techniques are being developed, aiming at drastically reducing the amount of training data needed, such as *transfer learning*, *few-shot learning*, or *single-shot learning*. Such approaches are of particular interest for ecological tasks (Pichler and Hartig, 2023).

A major critique of machine learning in general, and deep learning in particular, is that such methods pose a 'black-box' problem. Especially deep learning algorithms have been criticised for producing their output intransparently, prohibiting explanation of why an algorithm made decisions (Liu and Biljecki, 2022). Even more so, sometimes it is not even possible to examine the models properly (Liu and Biljecki, 2022). This significantly lowers the potential to use ML algorithms as a decision support for design and planning tasks. *Shapley values* and *partial dependence plots* are methods to partly overcome the problem (Jemeļjanova et al., 2024). The concerns raised about the opaqueness of ML models is fuelled by the insight that the predictive ability is not necessarily correlated with correct causal relationships within the model (Pichler and Hartig, 2023).

In the past years there have been approaches to new ML methods which enable at least a certain amount of verifiability, so called *white boxes* (Pichler and Hartig, 2023). This is also known under the paradigm of explainable artificial intelligence (xAI). Explainable Artificial Intelligence (xAI) is a recent development due to the drawbacks of many machine learning techniques described above (Pichler and Hartig, 2023).

7.1 Hybrid Methods

Machine Learning is methodically not confined within itself, and can be combined in various ways either with other techniques of Artificial intelligence as well as other computation techniques in general (such as agent-based modelling, knowledge-graphs, ontologies, etc.) (Guo and Liu, 2024). Especially traditional statistics are often combined with ML methods in sequential steps of analysis, as methods in statistics can offer advantages over ML with respect to *interpretability* and *computational cost* (Pichler and Hartig, 2023).

Analogous to Guo and Liu (2024) such hybrid methods can be distinguished into: 1) methods which are also types of AI, 2) methods which are supported by ML (ML is used to preprocess data), and 3) methods which support ML (ML is used on pre-processed data).

Within this distinction there are many possible constellations of AI and other computation techniques, including (Guo and Liu, 2024):

1. Provision of datasets through AI;
2. Integration into ensemble methods;
3. AI as optimization of computation;
4. AI sets parameters of other computational methods;
5. AI as an evaluation of other computational methods.

7.2 Learning Concepts

ML methods can be described by several learning concepts, which are not mutually exclusive. ML methods can be distinguished most basically into *supervised*, which means that the ground truth, often referred to as target or response variable, for at least a part of the data is known or elaborated beforehand (even if the relationship among the features provided are unknown) (Pichler and Hartig, 2023). This involves different mechanisms of control and guidance how the model can 'learn' from those examples to map the input variables to the output. *Unsupervised*, where datapoints with a given set of features, but no results are provided to infer output from, hence the model has to learn and decide by itself how to relate and sort the datapoints (Jung, 2022; Nikparvar and Thill, 2021). Table 7.1 gives an overview of the most important learning concepts and their descriptive characteristics.

Table 7.1: Different learning concepts for ML.

Learning Concept	**Description**
Supervised learning	In supervised learning, models are trained for classification or regression of datapoints, by providing data with values or labels for the target variables (i.e. the correct or expected result) (Tufail et al., 2023). After training is completed, unseen, unlabelled data can be presented for inference to the model.
Unsupervised learning	Unsupervised learning is applied for datasets which do not contain any target variables. Relationships amongst datapoints are to be found by the algorithm by finding similarities amongst parameters (e.g. distances) or calculating some kind of loss function
Semi-supervised learning	Semi-supervised learning can be distinguished from supervised learning by the training data, which contains only a relatively small set of labelled data, but also unlabelled data is used in the training process (Tufail et al., 2023)
Transfer learning	Is a learning concept where models are trained with a basic or multiple basic tasks in a layered structure. After training, top-level layer can be exchanged or new layers are added to modify the basic model to fulfil a similar but specific task. This approach saves time and resources for training complex models on very large datasets, as only part of them is adjusted to the new task.
Reinforcement learning	The basic concept is that the algorithm is an agent performing actions and making decision, which are in turn rewarded (or penalized). This concept relies on the evaluation of discrete steps, which results in a trial-and-error strategy, building sequences of actions to optimize output (i.e. overall reward) (Nikparvar and Thill, 2021).
Ensemble learning	A concept in ML where multiple models are tied together, and the final result is estimated through the different individual results of the individual models. The best-known method in ensemble learning are *random forests (RF)*.

For supervised methods, the output can either be continuous or at least ordinal, which is referred to as *regression*, or discrete which is the narrow definition of *classification* in ML (Rubbens et al., 2023). Regression, despite not being within the scope of this thesis, has many applications regarding ecological analysis. It can either be used as stand-alone solution or can be used in combination with other ML methods for feature correlation analysis (Guo and Liu, 2024).

Classification in its narrow sense means the discretization of datapoints into two or more predefined groups. This might either be simple binary classification of TRUE or FALSE, 0 or 1, or more complex classifications like multi-class, multi-label or multi-output classifications (Jung, 2022). The definition of desirable output categories has restrictions due to:

- The training data, where the input data is accurately representing the output data;
- Expert knowledge needed to define meaningful classes;
- Restricted applicability in different contexts (e.g. land cover classifications based on vegetation structures might not be applicable in desert regions).

Although these restrictions show the potential of classifying urban ecological units by pattern within the input data itself, i.e. unsupervised classification or clustering, there are some considerations for classifications if the desired output is already known (e.g. corridor networks).

Semi-supervised classification works with only partially labelled datasets, which usually means that the proportion of labelled datapoints (i.e. where the target variable is known) is small, compared to unlabelled ones (Casali et al., 2022; Rubbens et al., 2023). *Active learning* is a subset of semi-supervised classification approaches, where the ML model is getting feedback from human supervisors to enhance learning performance (Nikparvar and Thill, 2021).

Unsupervised methods, which are also called 'unsupervised classification', are producing a specified number of categories or *clusters* to divide the datapoints into (Hanna, 2022). Clustering can be distinguished by the axis of data discrimination into *partitioning* and *hierarchical* clustering. While the first concept treats similarities of datapoints as equal within the defined number of clusters, the latter establishes nested sets, representing a hierarchy of similarity (Wang et al., 2010). Hierarchical clustering can be further categorized in *agglomerative* and *divisive* clustering methods. Agglomerative clusters are being built from each datapoint as cluster of size 1 successively until a defined halting criterion is met. Divisive methods start from a cluster containing all datapoints and discriminate until a halting criterion is met (Wang et al., 2010). The proposed clustering approach by Perini et al. (2021) for the ECOLOPES project is a subvariant of hierarchical clustering called *agglomerative hierarchical clustering (AHC)* (Araldi et al., 2021).

Clustering methods can further be distinguished depending on their definition of how clusters are formed (Wang et al., 2010):

1. *Distance-based Clustering* is performed through a distance or dissimilarity metric;
2. *Density-based Clustering* are based on density of datapoints, forming regions which are clustered together;
3. *Model-based Clustering* assumes a specific statistical distribution of each cluster, the whole dataset forming a mixture of distribution models.

Constraint-based clustering is a special form of clustering, considered as semi-supervised learning method, where clustering algorithms are informed by some kind of supervised restriction or boundary condition (Nikparvar and Thill, 2021). *Contiguity-constrained agglomerative clustering* is a type of spatially explicit clustering, as it prohibits the merging of two clusters if they are not spatially contiguous (Barbierato et al., 2020). Edges in graphs are *first order* if they connect two neighbouring nodes, whereas a graph is spatially contiguous if all edges are of first order (Guo, 2008). Two graphs (or trees) are spatially contiguous if they can be connected by a first order edge. Contiguity-constrained clusters are formed either by only considering first order edges, or by calculating the distance between all available edges between clusters, which is called *full order constraining* (Guo, 2008).

Although classification and clustering can be seen as distinct tasks, which refers to the intended output format, classification is an underlying principle of ML, defining how an algorithm makes decisions. Even in tasks like image segmentation, the decisions if a condition is met or not, in this case if a pixel represents the boundary of an object, is made by some algorithmic 'classifier' (Jung, 2022). In this respect, classification refers to a logic of decision making, more than to a specific task. Also, as ML tasks are not mutually exclusive, several tasks can be necessary to conduct EUA.

7.3 Spatio-Temporal Machine Learning

The previous overview of machine learning concepts provided a basic understanding how machine learning works methodically. To answer, with respect to the presented framework for EUA and the importance to include eco-complexity into urbanistic analysis, how machine learning can specifically address spatial and temporal aspects of eco-complexity, theory from different disciplines concerning spatial ML will be presented. *Spatial machine learning* is a category of machine learning, specialized on the processing of geospatial data. Such data includes a geo-referential component (location) which is connected to other data, e.g. population statistics, which can either be tangible or intangible (Casali et al., 2022).

When spatial dependencies must be assumed, spatial data can improve the performance of ML models, or even provide new insights into relationships amongst data (Nikparvar and Thill, 2021). However, spatial data shows critical properties, which must be considered, including spatial dependency (i.e. spatial autocorrelation), heterogeneity and scale (Nikparvar and Thill, 2021). Not accounting for SA violates the underlying dependencies and may lead to biased evaluation metrics (Jemeļjanova et al., 2024).

Table 7.2: Variants of spatial covariates from (Jemeļjanova et al., 2024).

Covariate method	**Description**
Coordinates	Usually added as covariate in the form of longitude and latitude within a geographic reference system, or as X,Y values in Cartesian coordinate systems.
Neighbourhood covariates	Are usually given in the form of a weighted sum of the n nearest neighbours, whereas the weights are calculated by a distance metric.
Spatial weight matrices	Are an nxn representation of the spatial dependencies between all datapoints. The advantage of describing the (directional) relationships between each pair of datapoints can potentially lead to the need of exhaustive calculation resources.
Distance measures	Refer the distance of the datapoint to some other specified point, such as the corners, edges, or centre of the study boundary.
Various *resolutions*	Covariates can be added at various *resolutions* to improve the model performance.
Spatial (pre-)clustering	Sorts datapoints depending on their spatial association. This can be achieved by *local indicators of spatial association (LISA)* as proposed by Anselin (1995), for example by *Moran's I.*

There are several ways in which spatial data can be addressed through ML models. Jemeļjanova et al. (2024) distinguish between 1) using explicit spatial covariates, 2) dataset formation, 3) algorithmic calculation, and 4) independent exploratory analysis. As the last option is a hybrid approach going beyond the topic of my thesis, I refer the reader to Jemeļjanova et al. (2024).

7.3.1 Explicit Spatial Covariates

Explicit spatial covariates are provided as an input parameter together with other non-spatial parameters to introduce spatial dependency in the data pattern. Table 7.2 gives an overview about different possibilities of explicit spatial covariates.

7.3.2 Data Splitting

If spatial dependencies are not considered in the process of data splitting, i.e. some form of random sampling, model training is prone to biases due to clustering and spatial proximity (Jemeļjanova et al., 2024). Table 7.3 gives an overview of methods for spatially sensitive data splitting.

Table 7.3: Data splitting methods for spatial ML (Jemeļjanova et al., 2024).

Method	**Description**
Spatial block CV	Splits datapoints spatially into folds via clustering, creating folds of different or equal sizes and shapes to stratify the data sample.
Distance-based sampling	Omits the selection of training or respective testing points, if they are within certain distances of each other.
Replicating density distribution	Creates datasets for testing by approximating the distribution of the training dataset.
Weighted random CV	Uses residuals from a model where a random CV has been applied, multiplied by a weight to avoid spatial clustering.

Table 7.4: Spatial ML through model calculations (Jemeļjanova et al., 2024).

Method	**Description**
Local or weighted models	Account for spatial dependence either by fitting data locally or globally.
Kriging residuals	Uses the residuals of a fitted model, applies a semivariogram and kriging to predict values.
Representing spatial dependency as a surface function	Includes an additional function to spatially represent the predicted values as a surface function.
Building decision trees accounting for spatial dependence	Account for spatial dependency via bootstrapping sampling or accounting for the variance reduction through Moran's I or Geary's C.
Restricting coefficient values	Sets a regression coefficient to account for value similarity.

7.3.3 Model Calculations

Spatial dependency can also be addressed by the model directly, either by integrating additional methods into the model or using an algorithm which is aware of SA (Jemeļjanova et al., 2024). An overview of spatial model calculations is given in Table 7.4.

7.3.4 Spatial Clustering Methods

Referring to unsupervised ML methods, Kopczewska (2022) identifies several clustering methods which can be applied to map data point in space by:

1. Ignoring location attributes;
2. Clustering of points in space;
3. Clustering locations with values;
4. Clustering regression coefficients;
5. Clustering based on density.

7.3.5 Networks as Spatial Representation

After Lu and Yang (2022) there are three main ways to translate network information into data structures suitable for ML application: 1) Only a sample of few datapoints containing local geometric information is considered at a time, 2) long distance information is incorporated as path trajectories, and 3) evaluating aggregate information at a global scale. *Graph construction* translates spatial information from a Euclidean into a Non-Euclidean domain (Wang et al., 2024). Analog to image representation as graph structure, spatial entities within cities can be abstracted from applying a grid or lattice structure, or directly using real-world entities as graph nodes. To characterize each node as part of an n-dimensional network, nodes can be assigned to n-dimensional vectors (Lu and Yang, 2022).

7.3.6 Trade-Offs in Spatio-Temporal ML

The main advantage of using explicit spatial covariates is, that basically a form of ML can be applied (e.g. RF or SVM) regardless of the algorithms incorporating some form of spatial awareness (Nikparvar and Thill, 2021). This is also in line with recent developments towards data-centric forms of ML (Nikparvar and Thill, 2021). On the other hand, data-centricity entails potential problems if the dataset is not well prepared. If environmental covariates are correlated to spatial covariates, such might be superfluous (Jemeļjanova et al., 2024). If a large number of input variables are involved, variables might either be redundant if they are strongly correlated or might be correlated in contradicting ways (Pichler and Hartig, 2023; Nikparvar and Thill, 2021). Adding spatial covariates generally improve model performance, while minimally impacting modelling time. They can be a source of overfitting, and potentially spatial covariates mask the relationships of environmental covariates with the target variable (Jemeļjanova et al., 2024).

Recent research suggests that spatially explicit models considerably increase performance compared to generic models for the application to spatial data (Janowicz et al., 2020). It is important to acknowledge that non-linear spatio-temporal dependence of data lacks research (Janowicz et al., 2020). But fully accounting for spatio-temporal dependence might result in data volume exceeding computational capacities of current technological resources.

7.4 ML Workflow

Application of ML methods consist of several steps. Any of these steps will have profound impact on the quality of the model output, and a certain degree of computational knowledge will be necessary to draw meaningful conclusions. However, over the last few years, many ML libraries have been developed, offering pipelines for ML methods, which facilitate data handling and model training. *End-to-End approaches* describe the necessity to chain together multiple components to solve problems of higher complexity (Joshi, 2023). The most important steps in applying ML methods are described in the following.

7.4.1 Define Biodiversity/Multi-Species Goals or Targets

The identification of research questions or use cases needs to be informed by background knowledge about meaningful relationships in data. The presented framework for EUA helps to relate principles of eco-complexity to concrete problems in analysis and design. Additionally, the choice of ML methods and algorithms should be developed in accordance with the use case in EUA and the most important characteristics of eco-complexity.

7.4.2 Data Acquisition and Processing

The selection of data does not necessarily follow the definition of biodiversity goals but should rather be considered an integral part of objective formulation. Data can be selected either by expert databases providing KPI or causal inference with statistical or ML methods. Data should be chosen either by causality or correlation, i.e. they have meaningful relationships with the observed phenomenon (Jung, 2022).

The selection of variables to use as input is a crucial task for valid results. For some data the necessity might be obvious, but too many variables might even cause problems with interpretation. This might happen if variables have contradicting correlations amongst each other. To obtain a set of variables and input parameters, which represent the phenomenon under investigation well, and is restricted to a computationally efficient minimum, several techniques can be applied. *Causal inference* tries to estimate correct effects of variables (Pichler and Hartig, 2023). Variables, which do not seem of interest, can be *confounding* to other variables and must be included in causal inference. On the other hand, *colliding* variables need to be excluded (Pichler and Hartig, 2023). Correlation analysis (Pearson's correlation coefficient) helps to identify and reduce correlated variables (Li et al., 2019). To promote interpretability and reduce workload, several techniques can be used to examine variables and their relationship. There are several methods to extract weights of individual variables (Liu et al., 2024). One of the most widely used techniques for dimensionality reduction is *principal component analysis (PCA)*.

The choice of a fitting set of data to support valid and high performing ML output, is crucial. Three approaches for the selection of parameters can be considered:

- Existing research and modelling of the subject of inquiry that can be translated into an ML problem.
- Interdisciplinary work with ecological experts, formulating hypotheses and choosing variables.
- Prior evaluation of parameters and their sensitivity to change with statistical or ML methods.

Feature selection can be achieved by genetic algorithm or dimensionality reduction methods, where principal components analysis (PCA), factor analysis, independent components analysis and self-organizing maps (SOM) are most prominent (Nikparvar and Thill, 2021). PCA transforms large sets of interrelated variables into smaller sets of uncorrelated variables. As PCA assumes spatial stationarity, enhanced methods such as locally weighted PCA or geographically weighted PCA can be applied (Nikparvar and Thill, 2021). See Table 7.5 for an overview of effect size estimation methods.

Input data for ML purposes can come from a variety of sources. While remote sensing has been the main source for decades with ever increasing temporal and spatial resolution, lately social media data has gained the attention of many researchers (Martí et al., 2019). Although broadly available for many parts of the world, high quality data is still distributed very unevenly (Casali et al., 2022). Lacking standards of procedural methodology for collecting and formatting data is adding risk towards producing and using biased datasets (Casali et al., 2022). Often, due to technical problems, or especially in ecology, the effort and cost of data acquisition prevent the completeness of data needed to gain meaningful insights (e.g. species occurrence data) (Pichler and Hartig, 2023). *Resolution* of data has long been, and remains an important topic in many disciplines. Although highly resolved satellite imagery is becoming a standard product, even at the scale of 10 m and below, there is no guaranteed availability for all locations within a time span. Even more important, scale mismatches of data might be an issue in producing meaningful ML output. Therefore, interpolation and aggregation of data might be needed.

Table 7.5: Methods for effect size estimation of variables

Technique	**Description**
Dimensionality Reduction	Describes a set of techniques to aggregate multivariate data into fewer groups by finding commonalities in the data. The resulting 'clusters' might either be an end in itself as to receive an explanatory set or to reduce the dimensionality for further computation.
Principal Component Analysis (PCA)	Is a process of reducing multivariate data by finding correlated variables and combining them into new ones, while pertaining most of the information.
Hierarchical Clustering Analysis	Is a process of pairwise finding the most alike datapoints and clustering them hierarchically until only one cluster remains. This structure allows for 'choosing' at which level of aggregation the clusters represent the sought description best.
Pearson's Correlation Coefficient (r)	Measures the linear correlation between two datasets.

Data sparsity refers to a typical characteristic of high-dimensional data, where not all values are present. This is common with big datasets, where some or many dimensions have '0'-values or no values at all. *Missing data* can be independent on other data points, dependent on neighbourhoods or show specific patterns. Methods to handle missing data is aggregating data, removing observations, or imputing values (Nikparvar and Thill, 2021). Spatial prediction or interpolation used to impute missing data by spatial statistics are *GWR*, geostatistical models, *kriging*, and *probabilistic principal component analysis (PPCA)* (Nikparvar and Thill, 2021).

To improve data quality for model training and evaluation, several methods of data processing can be applied to raw data. An overview is given in Table 7.6.

Table 7.6: Methods and concepts for data processing.

Data Processing Method	**Description**
Data augmentation	If critical data is to be found below pixel level, analysis and predictions will be biased, if sub-pixel heterogeneity is high, and expressed as average or dominant type values. One solution is to infer data from available sub-pixel fractions to the rest of the data set. If the spatial coverage of data is generally incomplete other methods, such as Monte Carlo technique may be used to estimate the uncertainty of variability and affectedness of model output (Li and Wu, 2007). If to little data is available to achieve sufficient model training, generative ML methods can produce additional datapoints based on the available dataset (Nikparvar and Thill, 2021).
Data interpolation	If not all necessary datapoints within an area can be retrieved, the missing datapoints can be interpolated either by statistics or ML methods.
Data standardization	Data often comes from different sources, e.g. building polygons from different municipal geographic services, and can be of various formats and different interpretations or conventions of how to represent data (Casali et al., 2022). Therefor *standardization* is a crucial step, in which all datapoints are brought into a common format. This might be due to buildings being represented as ground perimeters in one source and differentiated in main and sub-building polygons in another set.
Data normalization	Another major issue in data preparation is the scalarity of input data. Depending on the issue, features that vary in scale, might generate undesirable results. If one feature is measured as fraction between 0 and 1, while another feature is represented on an open scale (e.g. street segment in m), algorithms might overestimate the importance of the second feature due to its numerical magnitude compared to the first one. Another example from natural language processing would be texts of varying length, where algorithms look at the positioning of words, will tend to look in the 'wrong' place, which is prevented by normalizing dimensions through masking operations.

7.4.3 Model Training, Evaluation and Application

As ML algorithms and their efficiency depend on the number of features and dimensions they are fed with, trade-offs should considered. Generally, with more features/dimensions effectiveness of models is increased, but on the downside calculation time and resource use rise considerably, lowering the overall efficiency of the model (Pichler and Hartig, 2023). Computational complexity is a measure of resources needed to execute an algorithm. The most common measure is time complexity, which is expressed as a maximum of elementary operations, as a function of the input size n. *Parameters* are the variables used as input into a model, while *hyperparameters* refer to variables influencing or constraining several factors of model training, such as *learning rate* (Joshi, 2023). Due to the complex tasks in ML it is not easily predictable which algorithm will perform best. It is common practice to compare two or more different algorithms and evaluate their respective performance on a given problem (Pichler and Hartig, 2023).

As the architecture of ML algorithms is very heterogeneous, some based on traditional statistics, others on neural networks, there are few commonalities. Although some algorithms do not need any training at all (e.g. k-nearest neighbours), those who are trained, need to calculate some kind of parameter to estimate the success of their learning. This is done via minimization of a *loss function*, to optimize the model's performance (Joshi, 2023). This is commonly accomplished by the concept of *gradient descent*, where the function is iteratively being minimized (Pichler and Hartig, 2023). There are two main problems with this sort of optimization (Chaillou, 2022; Jung, 2022). First, it is unknown if the function will have only one minimum. That means that the algorithm will first possibly find a *local minimum* instead of the *global minimum*. If iteration steps are set too closely, the algorithm will always move back to the local minimum and miss the global one. Second, if iteration steps are chosen to big, the algorithm might move out of bounds and find no minimum at all.

Generalization describes the ability of models to produce valid results if provided with formerly unseen data (Pichler and Hartig, 2023). During training, datasets are commonly split into *training*, *validation*, and *test* datasets. The first two are used in model training, while the last is used to check generalization capabilities before applying the model onto unlabelled data. Because models learn to fit their algorithmic pattern to the training data, they might adjust so well, that they perform almost perfectly on the specific training data but will produce larger errors, when new data is provided. This phenomenon is called *overfitting* (Pichler and Hartig, 2023). Due to the basic concept of machine learning, the performance of algorithms varies highly depending on the given input and the expected output. To be able to ascertain a satisfying validity of a model, certain procedures and metrics have been developed (Balogun et al., 2021). The most common are based on a *confusion matrix*. This matrix contains for categories of output: 1) true positive, 2) false positive, 3) true negative, and 4) false negative.

7.5 An Overview of Machine Learning Algorithms

There are several different approaches to categorize ML algorithms into methods, depending on the underlying logical assumptions (probabilistic, non-linear, non-parametric, etc.), the intended mode of learning (supervised vs. unsupervised), and several others (Hsieh, 2009; Joshi, 2023). Through the fast development of the research field, it is hardly possible, nor is it strictly necessary to fully encompass the field of ML to arrive at potentials and challenges for EU.

Although Cellular Automata are not considered ML in a narrow sense, these algorithms are often used for similar purposes. Cellular Automata represent a fixed layout of cells around a focal cell, which influence each other's state. The main characteristic of cellular automata is that if one cell changes it state, neighbouring cells are also affected (Le and Huang, 2023). They do this based on the *Moore's* or *von Neuman* neighbourhood logic. Moore's neighbourhood counts every cell adjacent to the core (n=8), while von Neumann neighbourhood only counts neigbhours which share edges with the core (n=4). These algorithms are not to be considered as ML since the interactions follow an a priori fixed set of rules, which is neither dependent on the specific input data, nor is there any adaptation or learning process involved. On the other hand, CA have proven as robust, reliable and relatively well interpretable. Hence, they still find application within further developed ML frameworks (see Ha and Jeong (2021) for example).

In the following an overview of algorithms sorted by their learning concept, and kinships is given. Recent developments and broadly applied algorithms (such as CNNs) are explained in more depth than others.

7.5.1 Supervised Learning Algorithms

Supervised algorithms can distinguished in comparably old, but broadly used algorithms, and more recent deep neural net algorithms which interesting learning capabilities. The first ones include amongst others support vector machines and random forests, while the latter ones could be exemplified by convolutional neural networks and recurrent neural networks.

7.5.1.1 Naive Bayes Classifier

A *Bayes classifier* is a statistical algorithm, based on *Bayesian probability* and *maximum likelihood*. The basic mechanism is to check the probability of a datapoint being in either of binary classes as an expression of a loss probability of $P_{err} = p(y \neq h(x))$, while a *naive* classifier is assuming independence between the individual features of each datapoint (Jung, 2022; Kopczewska, 2022).

7.5.1.2 Support Vector Machine (SVM)

SVM is a type of binary classifier, which chooses datapoints from both class clusters to create a *hyperplane* in between, i.e. the assumption that data is linearly separable in a continuous hypothesis space, defining the maximum distance, called 'margin', between clusters of information (Crisci et al., 2012; Chaturvedi and de Vries, 2021; Pichler and Hartig, 2023). SVM work well with small training samples, dependent on choice of kernel and regularization parameters (Chaturvedi and de Vries, 2021). If datapoints are not linearly separable, a non-linear feature space transformation can be applied [Pichler and Hartig (2023); Kopczewska2022]. SVM are commonly used for image classification in ecological applications (Pichler and Hartig, 2023). Through the maximization of the margin width (ϵ) and minimization of misclassifications (ξ) (Nikparvar and Thill, 2021).

Support vector random field is an extension that explicitly models spatial dependencies using *conditional random fields (CRF)*. An observation-matching potential function models the relationship between observations and classes, while a local-consistency potential function penalizes spatial discontinuity (Nikparvar and Thill, 2021).

7.5.1.3 Decision Tree (DT)

Decision trees are a very popular group of algorithms. The basic algorithmic structure works on a set of fixed hypotheses represented by one or more decision trees, with a defined feature (X) and label space (Υ) (Jung, 2022). The algorithm starts at the *root node* of a tree and works through the individual hypotheses towards *leaf nodes* which represent a decision. In contrast to algorithms with a linear hypothesis space (i.e. linear and logistic regression, and SVM) DTs are capable of approximating non-linearity (Jung, 2022).

Spatial entropy-based decision trees use spatial autocorrelation and incremental learning to select tree nodes in a spatial raster framework (Nikparvar and Thill, 2021). 'Salt and pepper' noise is a common problem in image classification with DT (Nikparvar and Thill, 2021).

7.5.1.4 Ensemble Methods

Ensemble Methods are based on the assumption that more complex algorithms in general have lower prediction errors (Pichler and Hartig, 2023). So called 'weak learners', which are commonly simple methods such as decision trees or Naive Bayes classifiers, are bundled together producing average estimations, which enable low error estimations, even done the individual error is high (Pichler and Hartig, 2023). The best-known ensemble method is arguably Random Forest (RF), which bundles a number of individual decision trees and randomly combines their output to improve the overall estimate.

Bagging is a method based on bootstrapping from statistical analysis, which aims at augmenting datasets. From a dataset of size n, through resampling with replacement, m datasets are generated. Further, m models are being trained with one dataset each. The bagging predictor is then calculating the mean model output (Crisci et al., 2012).

A random forest is a variant of a bagging ensemble and is based on multiple, randomized decision trees, where a majority vote is predicting the final output (Crisci et al., 2012), while additionally subsampling features in each node (Breiman, 2001; Pichler and Hartig, 2023). The algorithm works well with many different input features, i.e. multi-variate datasets. Also, the tendency to overfitting is low (Chaturvedi and de Vries, 2021). Because it is generally one of the most efficient learning algorithms and has a relatively simple setup, it has been one of the most widely use ML algorithms. *Geographically weighted random forests (GW-RF)* are similar to *geographically weighted regressions*, with a spatial weight matrix integrated into a local regression analysis framework, which avoids the problem of GWR in handling collinearity among predictors (Nikparvar and Thill, 2021).

Boosting refers to the sequential setting of several weak learners, where either a general differentiable 'cost' function or iteratively the residual errors of the predecessor are being minimized (Joshi, 2023). The most popular algorithm of this set is AdaBoost (Crisci et al., 2012; Pichler and Hartig, 2023). A common application of boosting is also in *boosted regression trees* where an overall loss function is minimized (Pichler and Hartig, 2023). *Gradient boosting* is creating additive decision trees, which are being developed iteratively. Residual errors from the prior tree are treated through prioritizing them (Kopczewska, 2022). *Extreme gradient boosting (XGBoost)* is an advanced version of gradient boosting (Wieland et al., 2019).

7.5.1.5 Artificial Neural Networks (ANN)

Artificial neural networks are currently the fastest developing class of ML algorithms. The basic unit of Artificial Neural Networks (ANN) is the so called 'perceptron'. *Perceptrons* are connected in layers, whereas every perceptron in every layer is connected to every perceptron in another layer, but not within the same layer (Joshi, 2023). Input signals are passed together with an adaptable weight, and a bias term. The idea behind this solution is that a static network of generic nodes learns to adapt to a problem by adjusting the weights (Joshi, 2023). After processing the individual signals into an aggregated unit, an *activation function* determines whether the signal is passed on (i.e. significant) or not (Joshi, 2023). This activation function induces non-linear behaviour. At the output layer signals are collected and aggregated (Joshi, 2023). ANN are non-parametric, i.e. they do not need any assumption about distributions underlying the input data (Chaturvedi and de Vries, 2021).

Classic Artificial Neural Networks consist of input, output, and ‘hidden layer’ (Pichler and Hartig, 2023). The simplest version is referred to as ‘feedforward’ network, which shows some significant weaknesses of weight adjustment. To improve learning abilities, the ‘backpropagation’ mechanism in so called *back propagation neural networks (BPNN)* is signalling the computed gradients back through the network (backpropagation) to the input layer to adjust weights efficiently (LeCun et al., 2015).

Radial basis function networks (RBF) are simple networks with one hidden layer. In contrast to conventional MLP neural networks, the weighted norm (distance) of input vectors and the neurons is calculated with a radially symmetric (gaussian) activation function (Nikparvar and Thill, 2021). RBF show some advantages in spatial modelling compared to conventional MLP, but can also be combined (Nikparvar and Thill, 2021).

7.5.1.6 Deep Neural Networks (DNN)

Deep learning is a subset of ML (Balogun et al., 2021). It distinguishes the *multi-layer perceptron (MLP)* from common ANN by a more complex architecture, containing three or more ‘hidden’ layers (Pichler and Hartig, 2023). Currently, the most popular deep neural networks for spatio-temporal data analysis are *Convolutional neural networks (CNNs)*, *Graph Neural Networks (GNNs)*, *Recurrent neural networks (RNNs)* in combination with CNNs and GNNs, and *Generative adversarial neural networks (GANs)* (Nikparvar and Thill, 2021).

7.5.1.7 Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNNs) are a variant of DNNs which have been very popular recently for image-based classification tasks, such as feature extraction, or image segmentation (Pichler and Hartig, 2023). One main reason for their supremacy in image-related tasks is the so called ‘convolution’ of raster data before evaluating vectors with an MLPNN. Main components of the CNN are: 1) convolution layer, 2) pooling layer, 3) flatten layer, and 4) activation function (Wang et al., 2024).

The convolution filter is a raster of usually 3 x 3 to 5 x 5 Pixels, sliding along the entire image, multiplicating the pixel values with the filter values, thus creating a probability for the centred pixel to contain a certain feature (Joshi, 2023). This value is combined with an activation function to account for non-linearity. After a full convolution, a pooling layer reduces every n x n Pixels through a pooling function (usually Max pool selects the maximum value) (Joshi, 2023). The process of convolution, activation and pooling is repeated several times until the features are flattened and fed into a fully-connected MLPNN for further evaluation (Wang et al., 2024).

CNN algorithms have proven successful in clustering morphological urban features (Cai et al., 2021). As objects and regions in morphological and similar analyses are usually irregularly shaped, the use of CNN is restricted, due to the need for regularly shaped images (Nikparvar and Thill, 2021). To overcome this problem, the *Object-based convolutional neural network (OCNN)* has been developed. In two steps, images are segmented into linearly and polygonally shaped features with homogeneous spectral and spatial properties, which are then evaluated by two separate CNNs, using different resolution sizes (Nikparvar and Thill, 2021).

7.5.1.8 Recurrent Neural Networks

Recurrent Neural Networks use backpropagation for processing sequenced data (LeCun et al., 2015). Within the 'hidden layers' neurons hold a 'state vector' holding information about past data sequences. A major drawback of training RNNs is that over time, gradients tend to move out of bounds (LeCun et al., 2015). *Long Short-term Memory (LSTM)* RNN are a special variant compensating for the quickly decaying memory or conventional RNNs, featuring an *input gate*, *output gate* and *forget gate*. LSTM have shown outstanding abilities to predict sequences (LeCun et al., 2015).

7.5.1.9 Graph Neural Networks

As the name suggests, *graph neural networks (GNN)* work on the idea that every datapoint represents a node in a graph structure. Nodes can be connected by edges, facilitating information exchange. The main advantage of graphs in comparison to grid structures in Euclidean domains is the possibility to define irregularly structured relationships between data (Nikparvar and Thill, 2021). Because irregularly structured data cannot be processed by most other DL algorithms, GNNs are specifically constructed for graph representations. The main principle of GNNs is the evaluation of nodes by aggregating the variables of the node and its neighbourhood, similar to conventional convolutions on grid data (Nikparvar and Thill, 2021). GNNs show high effectiveness in representing geometric structure in the original data (Nikparvar and Thill, 2021), and the literature on GNN is growing rapidly.

Spectral approaches to GNN create spectral representations of graphs by applying the graph Fourier transformation. This is facilitated through the eigenvalues and eigenvectors in the graph's Laplacian matrix (Nikparvar and Thill, 2021). Spectral graphs are sensitive to changes in the graph structure (Nikparvar and Thill, 2021). Spatial approaches directly process convolutions on neighbouring nodes in the graph simultaneously (Nikparvar and Thill, 2021).

Further two kinds of GNN can be distinguished in *Graph convolution neural networks (GCNN)* and *Graph attention neural networks (GAT)*. They differ in the way how information from adjoining nodes is being processed. *Graph Convolution Neural Networks* (GNN) are similar to CNN where the datapoints are not raster data but a graph. The convolution filter is aggregating information by sliding through the nodes and their respective neighbours (Liu and Biljecki, 2022). *Spatial regression graph convolutional neural networks* the mapping between the graph structure and the spatial weights matrix enables a specific capability to capture lagged spatial effects in the observed variables (Zhu et al., 2022). non-stationarity, less sensitive to standard deviation, good fit for insufficient observed variables (Zhu et al., 2022). *Graph attention networks* Graph attention networks are comparable to spatially oriented GNNs, but they are further able to impose weights on the edges connecting the vertices, and hence put attention to important connections in the graph by iteratively estimating true weights (Nikparvar and Thill, 2021).

7.5.1.10 Generative Adversarial Networks (GAN)

A *generative Adversarial Network (GAN)* is using two networks which are working against each other: a *generative network* tries to imitate an input set of data with randomized features, while a *discriminative network* tries to classify the received input as real or generated data (Nikparvar and Thill, 2021). Both networks learn to optimize through adversarial tasks (Nikparvar and Thill, 2021). In most use cases, training is complete when the generator outperforms the discriminator, and the generated features are not longer recognized.

7.5.1.11 Variational Autoencoder (VAE)

This form of network is based on the idea of extracting features in datasets. An encoder network transforms input data into ‘latent’ variables in low-dimensional space (similar to principal components), from which the decoder network tries to reconstruct the original input data. Therefore, the input nodes of the encoder must be symmetric to the decoder’s output nodes (Kingma and Welling, 2019; Nikparvar and Thill, 2021). Errors are minimized over a residual network (Kopczewska, 2022).

7.5.1.12 Adaptive Resonance Theory Network (ART)

A supervised, self-organizing, and self-stabilizing neural network for non-stationary environments. Can be coupled with fuzzy logic (ARTMAP) with two modules where one module adaptively changes topologies of networks and the other one maintains class labels (Nikparvar and Thill, 2021).

7.5.2 Unsupervised Learning algorithms

Self-organizing maps and k-nearest neighbours might be the most prominent unsupervised learning algorithms, but others also have interesting properties for spatio-temporal applications.

7.5.2.1 K-means

For a defined number of clusters k with centroids the distances to data points are calculated, with iteratively optimizing the location of centroids by minimizing the distance (Kopczewska, 2022; Nikparvar and Thill, 2021).

7.5.2.2 K-Nearest Neighbour

As the name suggests, for the *K-nearest Neighbour* algorithm, datapoints are evaluated in multidimensional feature space by distance to their k nearest neighbours. Every datapoint is evaluated and, most commonly by majority voting, is assigned to the most frequent class (Pichler and Hartig, 2023; Kopczewska, 2022). K-Nearest Neighbour is one of the simplest ML algorithms and does not inflict any form of optimization. Despite having high computational cost for finding the nearest neighbours for every point, and working on a simplistic logic, the algorithm delivers good performance if data is densely populated (Joshi, 2023).

7.5.2.3 PAM and CLARA

Partitioning around medoids (PAM) is assuming k core points (medoids), which must belong to the sample, where the total distance of points is minimized iteratively, to find the ideal set of clusters (Kopczewska, 2022). *Clustering large applications (CLARA)* is a derivation of PAM for large datasets, where only a subsample is clustered, and the rest of the points is assigned as k-nearest neighbours (Kopczewska, 2022).

7.5.2.4 SKATER and REDCAP

Spatial 'K'luster analysis by tree edge removal (SKATER) was originally proposed by AssunÇão et al. (2006). Data points are clustered depending on their location. The algorithm calculates the total similarity distance between all variables of individual contiguous areas (i.e. neighbours) as graphs. Through pruning, the graph complexity is limited to by removing edges with high dissimilarity, resulting in a graph with k nodes where every node is connected to a maximum of two other nodes with *k - 1* edges, also called a *spanning tree*, leaving fewer possible partitions (AssunÇão et al., 2006). Cluster building is achieved by removing edges, producing subsets of spanning trees. Edges to remove are calculated by minimizing intra-cluster deviation (AssunÇão et al., 2006).

Regionalisation with dynamically constrained agglomerative clustering and partitioning (REDCAP) is an extension of SKATER using hierarchical agglomeration constraining, where distances are defined through linkages and spatial constraining strategies (first-order or full-order) of neighbourhoods (Kopczewska, 2022). The spatial contiguous tree is the partitioned through the minimum spanning tree method, by calculationg suqared deviations (Guo, 2008).

7.5.2.5 DBSCAN

Density-based spatial clustering of applications with noise (DBSCAN) does not use distance metrics but is analysing density and sparsity of datapoints in multi-dimensional space (Kopczewska, 2022). Clusters are generated sequentially from random points, where neighbourhoods of radius ϵ are evaluated and points are classified as belonging to the cluster or as 'noise'. Subsequentially clusters are formed until all points are within a cluster (Kopczewska, 2022).

7.5.2.6 Markov Random Field

Markov random field (MRF) is a probabilistic method which is able to represent the dependency between pixel and regional information (Chaturvedi and de Vries, 2021). MRF have been used for image segmentation, texture analysis, as well as land use change classifications (Chaturvedi and de Vries, 2021).

7.5.2.7 Self Organising Maps (SOM)

Self-Organizing Maps (SOM) could also be characterised as neural network without hidden layers (Nikparvar and Thill, 2021). But while most neural networks type algorithms are oriented towards supervised learning techniques, SOM work with a different kind of cost function, which accounts for the similarity of datapoints with their neighbourhood. This means that SOM are a type of clustering algorithm that is capable of projecting multi-variate data into a low-dimensional (most commonly two-dimensional) presentation, while preserving topological relationships and relative distances among the original datapoints (Joshi, 2023; Chen and Ma, 2023). The input data points are projected onto an array of weights (Lek and Guégan, 1999). Neurons are organized in a one-, two-, or three-dimensional grid, where dissimilar units are placed farther away than similar ones (Nikparvar and Thill, 2021). SOM can be applied to spatial and non-spatial data. In contrast to *k-means*, SOM compares all neurons pairwise and visualizes the relation by the distance in topological space (Nikparvar and Thill, 2021). While SOM offer a more robust representation of features as components than PCA at the first level, PCA is able to create subcomponents to fully describe data variation (Joshi, 2023).

GeoSOM addresses the ignorance of geographic references in SOM, forcing the algorithm to give emphasis to geographically close neurons (Nikparvar and Thill, 2021). One-dimensional SOMs can be used as a mode of dimensionality reduction to produce spatial clusters capturing the variation in geographic locations (Nikparvar and Thill, 2021).

7.5.3 Combined and Stacked Algorithms

The algorithms presented above show a high level of abstraction to elucidate the algorithmic principles. However, in recent literature, algorithms are increasingly modified by combining one algorithmic architecture with another, to suppress shortcomings or enhance abilities (e.g. Convolutional Neural Networks combined with LSTM implementations). In the following section, practical applications in the context of EUA will be analysed to showcase the potentials and challenges of using ML algorithms to address aspects of ecological complexity.

8. Review of Machine Learning Applications for EUA

As described in Section 2.3.5, the review of is two-fold: first, a comparison of research found within the search results of the theoretical foundation for EUA. Second, individually selected articles addressing the *observables* presented in Section 6.3.

8.1 Articles systematically selected from the EUA Theory Dataset

The whole set is marked by a high diversity in terms of objects of research and methods. The most prominent topic is the mapping of LULC change. In terms of spatial scales, most articles are either referring to regional aspects or mappings at city level. Only few articles are differentiating elements of urban space, such as street networks, buildings, or parks and gardens. In general, scales of investigations tended to exceed the local or neighbourhood scale.

Many articles do not provide detailed information about the methodological background. This has been also reported from literature reviews in similar research fields (Grekousis, 2019). These results are supported by other recent reviews concerning the relationships of urban areas and ecological or environmental aspects (Li et al., 2023; Alavipanah et al., 2017; Middel et al., 2014; Morin et al., 2022).

8.1.1 Data Sources, Processing and Input Features

Many articles are using raster image data from remote sensing sources (n=15). Google Street View had been used by Zhang et al. (2023) and Ringland et al. (2021). Other data was based on surveys, GIS data, or calculated metrics. Some articles used time-series as in put data (Li and Fan, 2022; Lin et al., 2018; Morshed et al., 2022; Zubair et al., 2021). None of the articles used spatially explicit data in form of global co-variates.

Only few articles described methodically how data was interpolated and normalized regarding different scales of input data. Geographic information was used as elevation or slope (Bai et al., 2022; Bonilla-Bedoya et al., 2021). Artificial structures were mainly considered in covariates as distance metric (Bonilla-Bedoya et al., 2021; Dimopoulos et al., 1999; Karapinar Senturk, 2022). Other parameters were either raster images socio-demographic data, or calculated metrics such as NDVI.

8.1.2 Machine Learning Tasks addressing EUA

Within the reviewed articles, a number of tasks could be identified. The presented tasks are referring to the partial task accomplished by ML algorithms or models.

8.1.2.1 Sensitivity and Key Performance Factor Analysis

Mostly in a bigger frame of analysis, some articles were concerned with the extraction of crucial information about the input data and their relevance for results. Bai et al. (2022), investigating the construction of urban green space ecological networks used a back propagation neural network (BPNN), another name for multi-layer perception neural networks with back propagation to extract indicator weights to omit human bias in doing so. Similarly, Jiao and Han (2022) also used BPNN to identify weights and thresholds for planning UGS networks in sponge cities.

8.1.2.2 Enhance Performance and reduce Computational Cost

Ban et al. (2022) used BPNN coupled with a generative algorithm to evaluate the GA's individual evaluation function, reducing execution cost.

8.1.2.3 Image Segmentation and Feature Extraction

Dong et al. (2022) used DNN for extraction of roof surfaces in urban settings and later classification of their potential as green roofs. Similarly, Luhua et al. (2022) used a deep learning approach (D-LinkNet) for the detection of building shapes and green roof potential analysis. Ringland et al. (2021) used a CNN to detect plant species based on Google Street Maps images without spatial relation. Zhang et al. (2023)

8.1.2.4 Spatial Clustering and Classification

Barbierato et al. (2020) was one of the few examples who, after segmenting images of vegetation from street view imagery, used the spatially explicit REDCAP algorithm, to receive concise and cohesive clusters of information about the occurrence of vegetational patterns at a district level. Bonilla-Bedoya et al. (2021) used RF to predict (i.e. classification of unseen examples) the spatial distribution of soil characteristics, concluding that the method proved effective, but other classifier could improve performance in urban application. Chapman et al. (2020) tried to predict human activity and accompanied species invasion in freshwater networks with Generalized Boosting Model, distributed RF, DLNN and super learning, the last one being a stacked ensemble of algorithms. They reported mixed algorithmic performance dependent on specific activities to predict. Chen et al. (2023b) used SOM in combination with PCA to segment and classify land cover based on a set of reduced ecosystem services. Nölke et al. (2023) used a CNN in combination with hierarchical clustering to determine the composition and configuration of urban infrastructure based on three hierarchical scales.

8.1.2.5 Correlation of Ecological Factors within Urban Environments

Dimopoulos et al. (1999). He et al. (2022) tried to estimate plant parameters in relation to ... through CNN. Jutras et al. (2002) and Jutras et al. (2009) used ANN and BPNN to predict urban tree growth regarding several plant characteristics and built environment parameters, with good performance on predictions in both experiments. Li and Fan (2022) tried to predict ecological variables from UGS parameters using a PSO-BPNN. Sun et al. (2021) used ELM and XGBoost to correlate geographic factors with the cooling efficiency of urban parks.

8.1.2.6 Species Occurrence or Abundance Mapping

Some articles are concerned with the prediction of species occurrence or abundance in relation to landscape characteristics (Bergerot et al., 2011). Algorithms used where SOM, ANN, ... which proved promising and could even account for the adjacency of spatial units. Karapinar Senturk (2022) tried to predict the presence of amphibian species by evaluating satellite images of potential habitats with ANN. Łopucki and Kiersztyn (2020) evaluated species presence supported by Kernel Density Estimation, PSO, DT, and NN. Steenberg et al. (2019) used MLPNN to predict mortality in trees through neighbourhood-related indicators. Wang et al. (2021) used KNN, RF, SVM and BPNN for the classification of tree species. Wellmann et al. (2020) predicted the presence of avian species in relation to spatial heterogeneity. Wiese et al. (2019) predicted species distribution and causal inference through *MaxEnt*.

8.1.2.7 Change of Ecological Systems

Labib (2019) used ANN to predict land-use changes with various environmental indicators. They concluded that although the explanatory factors were weak, the predictive performance was satisfying. Le and Huang (2023) used an experimental approach, using CA in combination with a GAN (CycleGAN) to predict mycorrhizal networks of urban trees and estimate suitable tree planting locations. Lin et al. (2018) changes in wetland with Kernel Extreme Learning Machine, ELM, SVM and MLC. Morshed et al. (2022) used CA for the classification and prediction or LULC changes. dos Santos et al. (2021) used CA-Markov and ANN to model land use change at an urban scale. Zaleckis et al. (2022). Zubair et al. (2021).

8.1.3 Spatio-Temporality and Scale

The review articles showed in general only a weak reference to explicit spatio-temporal considerations. Explicit spatiality in form of pixel-based image, which can be seen as a local (image-wise) explicit spatiality was evaluated by 20 articles. Only Zaleckis et al. (2022) used a graph-based approach. Most other articles used multivariate data with different levels of dimensionality. No article specifically referred to hierarchical effects of models.

8.1.4 Use of Algorithms

In general, most studies used common algorithmic approaches, which are usually found in broader ecological research, such as RF, nearest neighbour, artificial neural networks. Especially neural networks are widely used for different tasks. Convolutional neural networks are almost always used for feature extraction tasks. With respect to spatially explicit tasks, the REDCAP algorithm used by Barbierato et al. (2020), and SOM, e.g. Bergerot et al. (2011) or Hassell et al. (2021), were the only algorithms used. The prevalent learning architecture are neural networks. Other studies also used SOM, RF, and regression models. Depending on the publishing year, there is a clear tendency towards DL methods, such as multi-layer perceptron neural networks (MLPNN) or CNN methods.

8.1.5 Summary

In summary, the reviewed articles only showed weak attribution of the principles of eco-complexity expressed by the framework presented in Chapter 6. Although some topics were of interest, the methodical approaches largely disregarded spatio-temporal aspects. The two most interesting articles in the context of EUA were Barbierato et al. (2020), because they used a contiguity-constraint clustering method to generate adjacency-aware clusters of urban forest patches, as well as Le and Huang (2023), which used a version of a GAN (CycleGAN) in combination with a CA to predict optimized future planting location of street trees, to enhance the urban ecological network.

8.2 Selected Machine Learning Case Studies

The literature found within the search results to construct the conceptual framework for EUA did not provide satisfactory evidence for the applicability and relevance of eco-complexity within the analysis of urban complex system. For this reason, literature found through several additional database searches with more precise search terms, as well as literature discovered non-systematically during the process of writing this thesis, will help to provide a ‘proof of concept’, demonstrating how in recent years the focus in research is shifting towards including aspects of complexity and complex systems in ML methods and algorithm creation. The description of the research is entirely citing the work of the respective authors, and for readability reasons only cited at the end.

8.2.1 Scale - "‘Reading’ cities with computer vision: a new multi-spatial scale urban fabric dataset and a novel convolutional neural network solution for urban fabric classification (Fang et al., 2020) "

Fang et al. (2020) developed their framework ‘UrbanClassifier’ to classify urban land parcels from multi-channel imagery by aspects of 1) city of origin, 2) morphological type, and 3) historical period. Image evaluation is facilitated by a classic CNN architecture. The idea for this learning algorithm is to take the representation of multiple scales as input to improve prediction performance. In a first step, a dataset was created as a multi-channel city map, representing buildings, the street network, and land parcels. Then all land parcels were annotated with additional information about urban morphological types and development periods. To establish a multi-scaled dataset, the prepared multi-channel images were cropped at random locations at different, fixed resolutions values. The performance was compared to three baseline models, where one only accounted for local geometric features, one added surrounding context, and the third one used a multi-scale approach with three different scales.

The multi-scale baseline model performs convolutions on each of the three image scales and concatenates the resulting flattened vectors to feed them into a fully connected classifier network for each of the three classification tasks. The main model ‘MMSSModel’ produces *path weights* (i.e. different combinations of the multiple scales available) specifically for every task and the individual path convolutions are multiplied with these weights and then summed to arrive at a classification result. The multi-scale convolutional learning approach is shown in Figure 8.1. The results of this experiment were not straightforward. Noise was introduced through the combination of several scales, which forced the baseline model to equally consider all scales, which decreased performance compared to the best-performing single-scale model. However, the MMSSModel which could compensate for the varying explanatory qualities of scales achieved better results at two of three tasks (Fang et al., 2020).

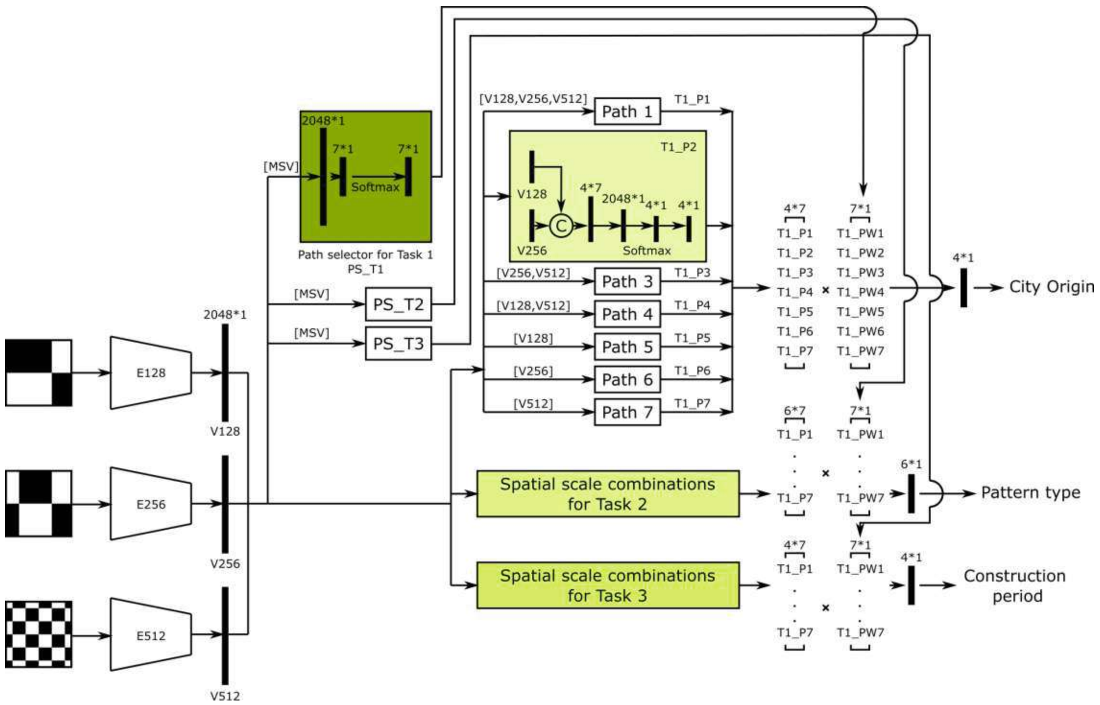

Figure 8.1: UrbanClassifier uses multi-scale image datasets and evaluates the performance of scale combinations to improve predictions; after Fang et al. (2020).

8.2.2 Composition - "Mapping Vegetation types by different Fully Convolutional Neural Network Structures with Inadequate Training Labels in Complex Landscape Urban Areas (Chen et al., 2023)"

Chen et al. (2023) proposed to overcome the issue of CNN needing excessive amounts of data to be trained properly by combining the advantages of DL with ensemble and transfer learning. This was achieved by pre-training three implementations of UNet++ networks using LULC data to extract *deep features*. As data input they used Sentinel-2 satellite images with 13 bands and 10, 20, and 60 m resolution. From these images sample sets were prepared by adding vegetation indices and label points for training. The main algorithmic module to extract 'deep features' from satellite images was a combination of three different implementations of the UNet framework (Chen et al., 2023).

The main idea behind the *UNet* architecture is to process high-resolution images in a series of convolutions reducing the resolution, similar to conventional CNN architectures which extract deep features, such as localization. But UNet differs from CNN as the down-sampled images are then up-sampled again through convolutions, to apply localization and feature information to a high-resolution image (Ronneberger et al., 2015).

The *Relief-F* algorithm was then applied to filter the deep features extracted from the image data. The selected features and labels were first used as input for a stacked classifier, combining SVM, gradient boosting decision tree, and k-nearest neighbour algorithms for their specific advantages as base classifiers to analyse the probability distribution. The output of the base classifiers was then concatenated and used as input into an RF meta-classifier. The results showed that the classification of the extracted deep features in images enabled significantly higher classification accuracy compared to common classification algorithms such as RF, U-Net, DeepLab V3+ (Chen et al., 2023). The algorithmic logic of VGG16-UNet++ is shown in Figure 8.2.

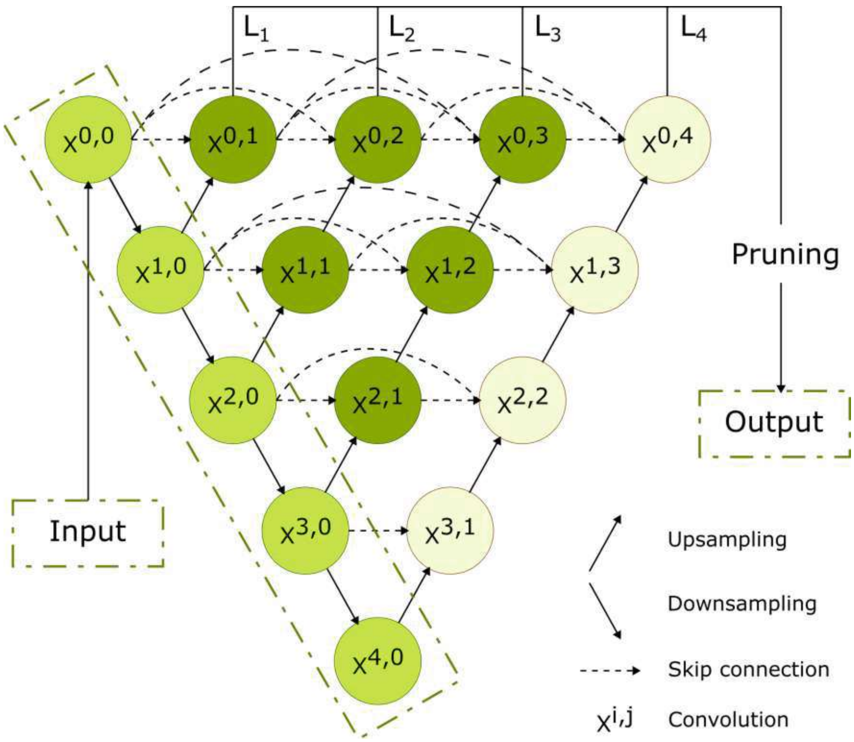

Figure 8.2: The functional principle of UNet++ to extract deep features after Chen et al. (2023).

8.2.3 Configuration - "Assessment of the urban habitat quality service functions and their drivers based on the fusion module of graph attention network and residual network (Wang et al., 2024)"

Wang et al. (2024) proposed a new method to assess habitat quality through image classification. The idea is to avoid the ignorance of adjacent image elements as in conventional convolution kernels, while also avoiding the comparably high computational cost of full-resolution adjacency matrices for graph-based computation. Their solution is combination of a GAT and CNN. The method is based on remote sensed imagery and land use classifications. First, a 1 x 1 convolution is performed to reduce noise and redundant information within the remote sensing data.

The graph-based branch of the framework is then performing *simple linear iterative clustering (SLIC)* through K-means to divide the image into 'superpixels' which are connected and hold similar spectral characteristics (Wang et al., 2024). Full-resolution pixels information is then transformed into one-dimensional vectors for each superpixel centroid. These new superpixel features are then fed into the GAT network to obtain network representation and calculate new weighted feature vectors, as seen in Figure 8.3. The GNN is able to enhance interaction between neighbouring image elements and to remove redundant information, through reweighting the edges in between nodes.

The other branch consists of a *spatial attention module* and a *channel attention module*. Both modules are using CNN performing different kinds of convolutions to obtain attention weights for spatial and channel information. Features are finally fused through weighted fusion and habitat quality is classified with a classifier function. The framework performs well and makes significantly fewer misclassifications than other ML algorithms (SVM, ResNet, and others). The authors conclude that natural factors are more important for habitat quality than socioeconomic factors (Wang et al., 2024).

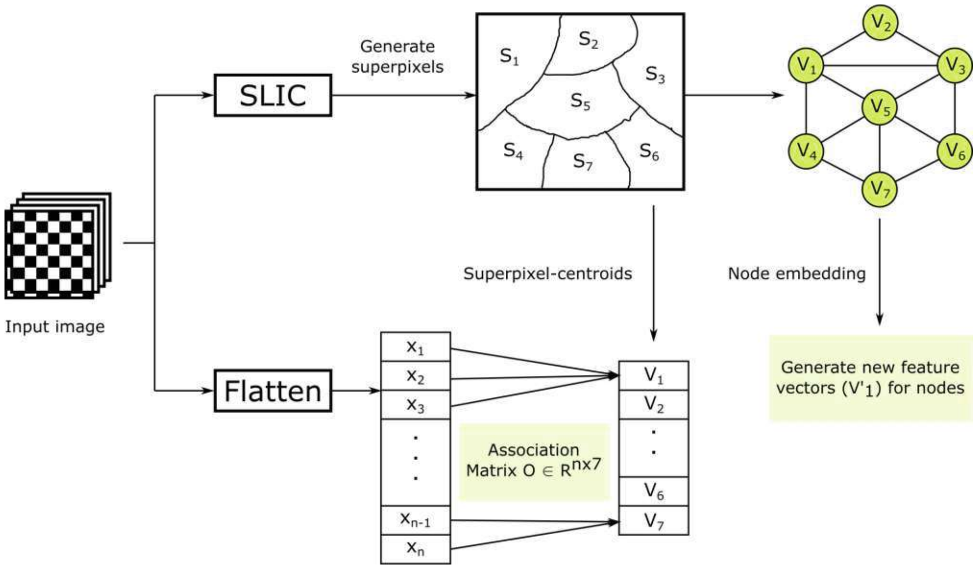

Figure 8.3: The GAT module features the creation of 'superpixels' to facilitate regionalization and graph attention. After Wang et al. (2024).

8.2.4 Connectedness - "Finding key players in complex networks through deep reinforcement learning (Fan et al., 2020)"

Fan et al. (2020) proposed 'FINDER' as a framework for the identification of key nodes within networks, which are important for network function. The implementation is a deep reinforcement learning framework. The purpose is to find nodes in a network whose activation or removal would maximally impact the network functionality, finding an optimal set of key nodes in graphs. FINDER is a data-driven deep learning framework, where key players of a network are found in a trial-and-error approach via a Markov decision process, sequentially removing nodes from the network, which is rewarded by decreasing *accumulated normalized connectivity (ANC)* (Fan et al., 2020). FINDER can handle two different connectivity measures, $\sigma_p air(\cdot)$ and $\sigma_g cc(\cdot)$.

In a first step, FINDER is trained on synthetically generated, small graphs until reaching sufficient performance. Then it can be applied to either larger synthetic networks or real-world networks. The algorithm performed significantly better than other available methods on both kinds of network. FINDER is working especially efficient, significantly reducing computation cost (Fan et al., 2020). The main procedure during real-world applications is shown in Figure 8.4.

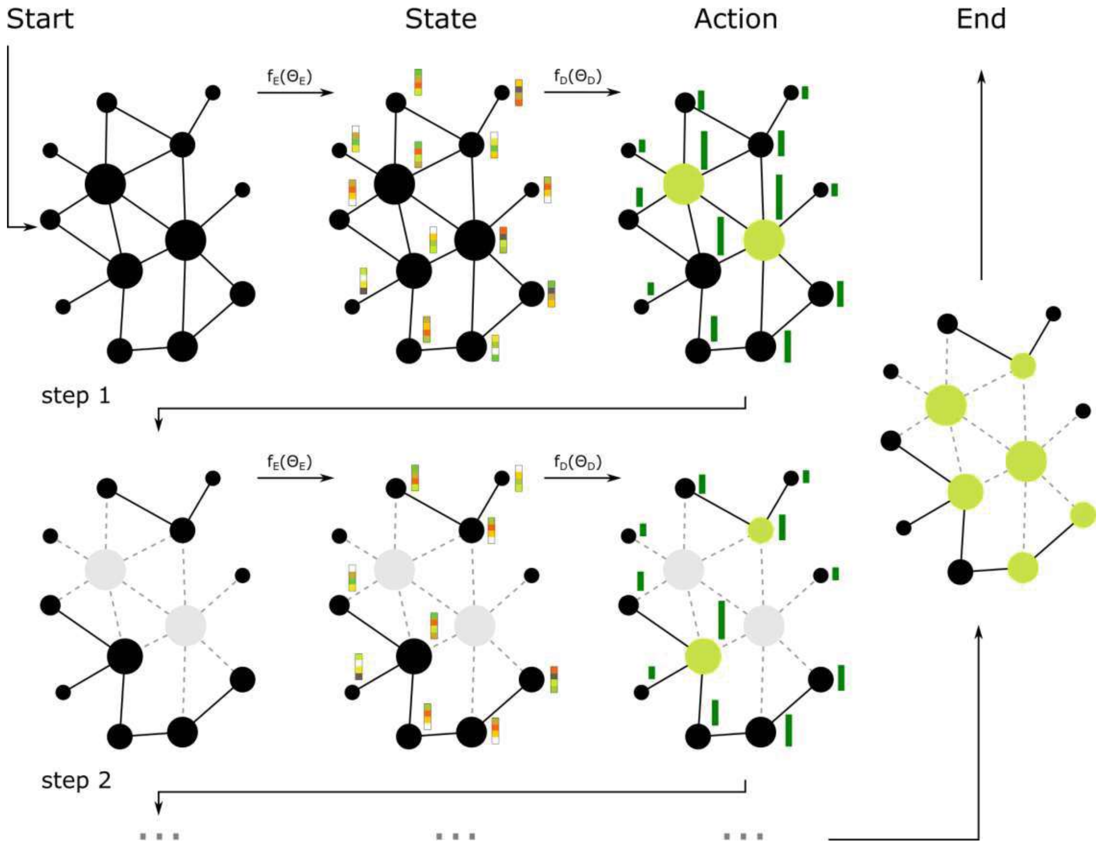

Figure 8.4: The conceptual workflow of FINDER to detect key nodes within networs, after Fan et al. (2020).

8.2.5 Interaction - "Uraveling hidden interactions in complex systems with deep learning (Ha and Jeong, 2021)"

Ha and Jeong (2021) published their framework for *AgentNet* in 2021. This framework is based on the idea of agent-based modelling to predict complex system behaviour. The architecture is based on a GAT module. The model initially assumes a fully connected graph structure, implying possible connection amongst all agents, while true connections are evaluated over training iterations. The logic of changing states of agents within the system can be described in four steps: 1) neighbourhood awareness, 2) apply strength, 3) interact, 4) transition (Ha and Jeong, 2021).

The agent-based system is implemented under several assumptions, 'state' and 'global external variables', and a time-series based vector (Ha and Jeong, 2021). Agent-based models calculate a *transition function* of its constituting parts through multiple time steps. Interactions within the systems are expressed pairwise. In contrast to conventional GAT architectures, interaction functions account for variable-wise interaction strengths. State variables are assumed to be impacted by continuity, stochasticity and memory. The model uses LSTM to produce sequential predictions, i.e. to capture memory effects of the system. The implementation with different complex system models (Cellular automata, Vicsek model, Active Ornstein-Uhlenbeck particle, Chimney Swift flock) allowed for the evaluation of different system characteristics to be addressed by *AgentNet*. A drawback of AgentNet is that it is currently limited to pairwise interaction evaluation. Three or higher order interactions could potentially better approximate complex system behaviour. On the other hand, *AgentNet* is a universally applicable framework, scalable for arbitrary numbers of agents (due to the GNN approach) (Ha and Jeong, 2021).

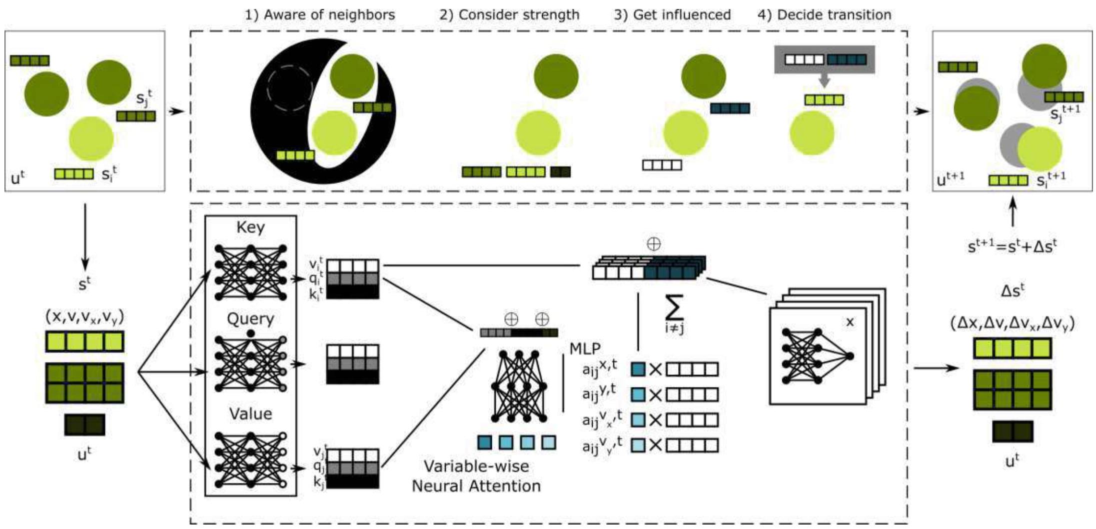

Figure 8.5: The framework functionality of AgentNet is based on pairwise attention. After (Ha and Jeong, 2021).

8.2.6 Change - "E-LSTM-D: A Deep Learning Framework for Dynamic Network Link Prediction (Chen et al., 2021)"

Chen et al. (2021) addressed the evolution of real-world networks, where nodes occur and vanish over time with *dynamic network link prediction (DNLP).* They proposed E-LSTM-D as an extended version of an LSTM network which is coupled with an encoder-decoder architecture. The framework of the E-LSTM-D model is depicted in Figure 8.6. The encoder-decoder architecture learns network representations automatically, while the stacked LSTM module learns to predict temporal features of changing links in networks. First a sequency of graphs with length N is mapped through an encoder into lower-dimensional latent space, transforming the graph matrix into a matrix representing structural features. A series of LSTM cells is then learning the evolutionary patterns within these feature matrices, returning predicted feature maps back into the encoder-decoder architecture, where the decoder projects the feature matrices back to full graphs. In contrast to the VAE, this framework is a supervised learning method. *E-LSTM-D* is capable of making long-term predictions with comparably low performance decline. The framework shows promising capabilities to lower the threshold for graph-based studies and can profit from further reduction of computational complexity to increase performance on large-scale networks (Chen et al., 2021).

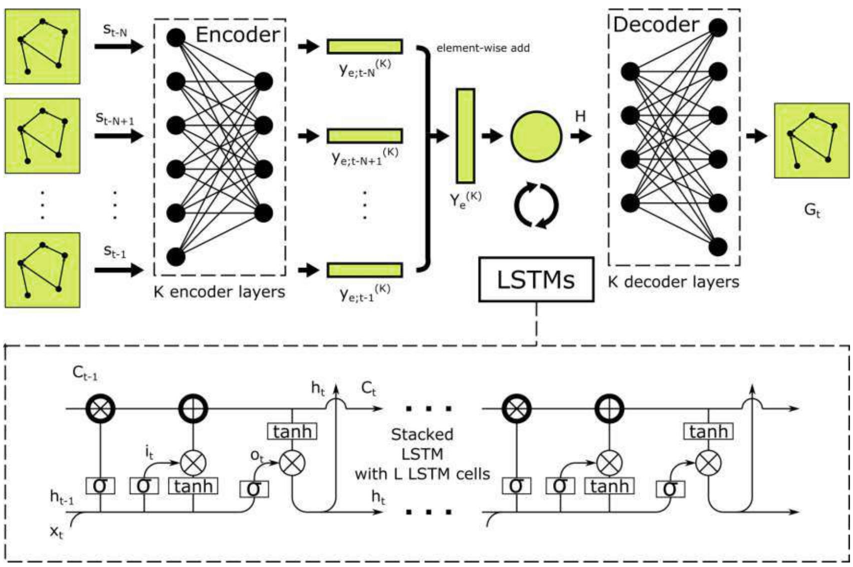

Figure 8.6: The E-LSTM-D framework by Chen et al. (2021) features an end-to-end graph encoder-decoder architecture coupled with an LSTM module to learn evolutionary patterns, after Chen et al. (2021).

8.2.7 Memory - "What makes the difference between memory and face of a landscape? A machine learning approach applied to the federal state Brandenburg, Germany (Wieland et al., 2019)"

The approach of Wieland et al. (2019) utilises two different models, one to evaluate the 'memory' and one for the 'face' of the landscape. Both models are used to estimate historical forests around 1880. The 'memory' model tries to identify the influence of basic landscape factors on the long-term forest distribution. The 'face' model uses big data to identify variables influencing the short-term changes of landscapes. The models do not address spatio-temporality explicitness and use different factors of influence related to topography, soil, land management, and biotopes. The idea is to use data from today, to restore a forest landscape from 1880. The models apply eXtreme Gradient Boosting (XGBoost), which facilitates gradient boosting on decision trees. A main criterion for choosing this algorithm was its ability to estimate importance of features . Both models were able to reflect the distribution of forests in 1880 with a comparably high degree of accuracy (Wieland et al., 2019).

9. Discussion

To arrive at an assessment of the challenges and potentials which ML poses for EUA, we will first recapitulate the main points of concern. Machine learning and deterministic modelling share similarities in aiming at sufficiently describing real-world phenomena, while abstracting their complexity (Geary et al., 2020; Hanna, 2022; Pichler and Hartig, 2023). In contrast to deterministic modelling approaches, machine learning does not a priori assume a certain relationship amongst the data, but it adjusts weights and functions from a generic probabilistic or non-linear model to approximate the complex system and discern underlying reproducible patterns (Hanna, 2022).

9.1 A Framework for EUA - Recapitulation

After reviewing and presenting a large volume of literature, it becomes evident, that although some efforts have been made to unify methodological concepts from ecology, and urban design and planning (e.g. Marcus et al. (2019a)), those frameworks remain at an abstract theoretical level. Further, theoretical concepts of biodiversity, landscape, and urban ecosystems are pertaining to different spatio-temporal scales and levels of abstraction. Although they share conceptual commonalities, biodiversity research is guided by reductionist approaches (Beninde et al., 2015). However, hierarchical coupling provides a vital theoretical link to overcome the gap between specific phenomena of biodiversity and theoretical knowledge concerning eco-complexity and complex adaptive systems (Wu and David, 2002).

Furthermore, ecological conceptions of space are in some key respects differing from those in architectural and urban design. Most prominently, they rarely concern with explicit spatial representations in two- or three-dimensional domains (Cadenasso et al., 2007; Alavipanah et al., 2017). However, the patch-corridor-mosaic model can serve as a suitable interface for interdisciplinary analysis at a meso-scale (Pauleit and Breuste, 2011; Marcus et al., 2019a). This enables the formulation of *observables*, a concept borrowed from physics, as measurements which are mutually valid in both knowledge domains and are scale-invariant. However, although the observables can be used as an approach to establish spatio-temporal topographies and topologies, EUA remains at a conceptual level, without definition of methods or metrics for application. On the other hand, the observables address distinct characteristics of complex ecological systems and provide a basis for the evaluation of ML methods as applied form of EUA.

9.2 Applicability of ML for EUA

With the current surge in ML research, the question has arisen, to which degree analysis of complex systems, such as social-ecological systems, are dependent on specific data or features and the method of inquiry (Hanna, 2022). It has been claimed that with the use of ML and AI in general, designers could be enabled to find structures and patterns in data, which would either be too difficult or complex to comprehend otherwise (Carta, 2022), or which go beyond a researcher's or practitioner's field of expertise. However, such expectations should be seen cautiously as there are several important requirements and trade-offs which have to be taken into account.

While the framework presented in this paper should serve as a foundation for the integration of aspects of eco-complexity in social-ecological urban systems, several factors must be given attention to:

1. The setting of ecological objectives is crucial for finding the right parameters and approaches for analysis.
2. Parameter selection necessitates an intricate knowledge of which ecosystem components might be important, but there are also several possibilites to estimate parameter correlation and impact through statistical and ML methods.
3. Data acquisition based on selected parameters is dependent on the availability, the format and resolution to be used as features.
4. Interpretation of ML outputs requires some sort of metrics, but even more so, only the use of multiple analytical approaches, whether ML based or not, will give a fuller picture, necessary to assure the validity of the model output.
5. ML algorithms are a conglomerate of very different forms of logic, mathematical or computational methods, and levels of complexity. Choosing the right algorithm will depend on the task at hand.

9.3 Potentials and Challenges for Multi-Species Design

Data has become abundant in recent years. However, high quality data has been a prerequisite for any empirical evaluation, and this still holds true for machine learning. But high-quality data from survey and aggregate statistics are comparably expensive and are usually updated only periodically. Although the availability of data has significantly improved for certain types, such as remote sensed imagery, other types of relevant data such as species abundance and distribution, or soil characteristics remain relatively scarce (Pichler and Hartig, 2023). More recent approaches to machine learning try to leverage big data from social networks, but as this data is provided voluntarily with limited metadata or methodical consistency, quality is a major issue, introducing uncontrollable bias, which can be propagated and covered-up by black-box models (Xing and Sieber, 2018; Liu and Biljecki, 2022).

Potentially one of the biggest issues for the mainstreaming of ML into analytical tasks in architecture and urban design for ecological applications is the selection of data. Not only must the data be available, the choice of 'meaningful' data, together with the analysis of possible interferences within variables will for many use cases only be possible by developing expertise or consulting experts (Pichler and Hartig, 2023). However, there are several methods at hand, supporting the selection of data and minimizing biased results. Also, pre-trained models and modelling pipelines facilitate an expert-informed application of ML methods.

Even if data is already massively available in high resolution for certain factors such as building height or land cover and is updated frequently, AI demands the conversion of all input data into machine-readable form, mostly into vector form, to be processed by algorithms. This increases the amount of preprocessing, as well as the risk of errors if not done with expert knowledge. Further, the ad hoc nature of many machine learning studies leads to lacking reproducibility or comparability of results due to issues of coarse scaling measures, such as 'neighbourhood level', which is impacting the function of certain algorithms (such as CA) (Kopczewska, 2022; Jemeļjanova et al., 2024). The *modifiable areal unit problem (MAUP)* addresses the bias introduced by using point data which, when aggregated to larger scales, is distorted by the aggregation unit's shape or scale. This problem especially affects algorithms which are based on the assumption of spatially evenly distributed data in grid form (such as CNN) (Nikparvar and Thill, 2021).

A major drawback to the applicability of ML algorithms is their relatability to practical concerns. Especially the design of DL methods, posing 'black boxes' complicates analytical application (Pichler and Hartig, 2023). Although we can evaluate the performance of different models either by comparison or by specific metrics, the question whether a certain percentage of a score is representative for its performance, and if the output is satisfyingly accurate to base major decisions for long-term development with possible lock-in effects upon, remains unanswered. This depends on numerous factors and must be decided most often by efficiency reasons. A major inhibitor of comparatively evaluating ML approaches from different disciplines, or for different applications is the lack of interdisciplinary guidelines and standards on reporting ML results, as they have been established in specific fields, e.g. biology (Kopczewska, 2022).

9.4 Strengths and Weaknesses of Algorithms

Classifications support a uniform evaluation of ecological landscapes. They are indicated if a problem is well understood, and all variables and their correlation are known. They produce a consistent output, that is directly comparable to different situations (Kopczewska, 2022). Clustering on the other hand will produce different outputs in different situations or will fail to address newly introduced relationships. Clusters from individual cities will either need new clusters for new cities, or risk that some clusters are not representative. Further, variables are for many analytical tasks not easily identifiable and can differ regionally, as Araldi et al. (2021) note about the definition of building types via clustering methods, that there is no fixed set of formal characteristics to differentiate building types, since important properties differ from one type to another. However, as clustering is achieved without predefined output classes, it offers the opportunity to 'explore' data from an experimental point of view, or compare it to classification results for evaluation purposes (Pichler and Hartig, 2023).

From the algorithms presented in the reviewed literature, there are some indications that machine learning might be able to address dimensions of eco-complexity effectively. The wide variety of algorithms to choose from leaves the question which algorithm is the best fit for a given task. Although this is a difficult decision and an understanding of functionality of an algorithm as well as the specific task at hand is imperative, some generalisations can be made. On the other hand, the nature of complex systems leads to the assumption, that there cannot be a "one fits all" solution (Jemeļjanova et al., 2024). Because there is no definitive way to say which algorithm will yield the best predictive result for a given task, most studies engage at least two different algorithmic approaches, and compare the outcome.

Traditional supervised ML algorithms have been optimized to work well with comparably small datasets, as well as structured data (Pichler and Hartig, 2023). More recent developed DL approaches show advantages in tasks with high-dimensional, large datasets. Convolutional Neural Networks are specifically good with image data and have a high capacity in learning the representation of features which are sensitive to neighbouring patterns. This enables classification approaches to account for adjacency between datapoints in high-resolution datasets. Similar capabilities are present in spatial ML approaches to unsupervised learning such as SKATER. However, the approach of rasterized convolutions limits the possibilities to address spatial dependency on a global scale (e.g. the geographic position to distinguish data from different cities).

Recurrent Neural Networks are designed to process sequential data (Pichler and Hartig, 2023). Through the implementation of LSTM, RNNs are equipped with a long-term 'memory' and are superior in processing and predicting time-series. Such models could account for multi-temporal processes and could be used for the analysis of individual plots (e.g. prediction of plant successions), or coupled with spatially explicit methods at neighbourhood or district scales.

GNN are potentially capable of maintaining relationships by topological connections for datapoints which are not adjacent but connected by higher scale patterns. GAT are further able to steer the attention of the model towards more important links through weighting the edges of the graph. At the time of writing this thesis, GNNs incorporating temporal awareness (such as LSTM) appear to be the most promising strain of ML methods to account for spatio-temporal complexity in data patterns (Nikparvar and Thill, 2021).

For unsupervised algorithms, possible applications for EUA remain unclear. There are several clustering methods considering spatial dependency, but applications addressing complex systems in spatio-temporal domains are scarce. But as the presented articles in Section 8.2 were chosen unsystematically, by preference to showcase the potentials to address observables of EUA, unsupervised methods might not be represented equally to supervised ones.

Hierarchical clustering approaches (e.g. agglomerative hierarchical clustering), as well as constraint-based clustering (e.g. SKATER) have found ample use in land cover mapping studies. However, drawbacks such as the need for expert rules or supervised thresholds, and the number of variables as limiting factor for the interpretability of clusters at individual levels can hamper the applicability of such algorithms (Václavík et al., 2013).

Self-organizing maps are able to work with minimal input to produce distinct clusters. They are especially relevant for uncovering pattern in multi-dimensional datasets (Lek and Guégan, 1999). They preserve similarity distances when reducing dimensionality into the two-dimensional output matrix (Václavík et al., 2013). Modified versions of SOM are also capable of considering geographic closeness of datapoints (Nikparvar and Thill, 2021).

This summary is restricted to algorithms which are considered to be important in addressing spatio-temporal dependencies. As the reviewed literature has shown, there is a tendency to combine different ML methods to compensate for shortcomings of individual algorithms and enhance model performance. Within such combinations, specialised algorithms such as CNN, RNN or GAT can be used to address spatio-temporal dependencies, while the results of these operations can be used as input into simpler, but more robust classifiers, such as RF (e.g. Chen et al. (2023)).

However, some studies also concluded that algorithmic optimization or individualization, while probably increasing the result marginally might result in substantially raised resource demands, which cannot be ignored by sustainability concerns (Jemeļjanova et al., 2024).

9.5 Application of EUA through Machine Learning Methods

The case studies in Section 8.2 exemplified the observables of EUA. The studies were mainly chosen due to their innovative approach to address complexity in data. However, these examples are all forms of classification approaches. Although it may be more difficult to find suitable analogues of spatially explicit clustering of urban data, the work of Barbierato et al. (2019) provides an interesting example of contiguity-constrained based clustering of cities.

9.5.1 Scale

This first study from Fang et al. (2020), although addressing urban typo-morphology, is of interest for EUA because of the deliberate incorporation of multiple spatial *scales* for analysis. This approach could easily be transferred conceptually to include other aspects of ecological interest, to classify habitat structures or for similar tasks in EUA. If image data was enriched by additional information, such classifications could hold important information about emergent pattern in urban ecosystems.

9.5.2 Composition

The identification and extraction of vegetation or other biophysical entities is an important first step, if detailed inventories at urban levels are not available. Therefore, the approach by Chen et al. (2023) to inventorize vegetation types is a good first fit to represent *composition* analysis through ML. Composition is relatively easy to analyse through conventional remote sensing methods. With improvement of resolution, such information can be retrieved at ever finer scales. Although vegetation information is important for biodiversity and network analysis, most approaches are still restricted to vegetation extraction, which is excluding information about the built environment.

9.5.3 Configuration

The habitat quality assessment through LULC data from Wang et al. (2024) shows potential to use high-resolution images for habitat classification. Although it is based on image classification through CNN, within images spatial relationships of features are expressed through a GAT. If such an approach was enhanced with other important variables of urban eco-systems, this approach could address complex configurations of urban ecological systems.

9.5.4 Connectedness

The framework from Fan et al. (2020) was successfully applied to a set of different complex real-world networks and showed very good generalization capabilities. Such an approach could be formulated as a problem for optimizing ecological networks within cities and could help in identifying the most important sites or patches for network conservation or enhancement through urban design measures and nature-based solutions.

9.5.5 Interaction

The approach of *AgentNet* from Ha and Jeong (2021) shows several interesting capabilities for working with complex systems. First, the GAT module is initially assuming fully connected graphs with equal weights. True connections and weights are estimated iteratively through graph convolutions. It would be possible to abstract grid cells or objects in urban areas as agents (Batty, 2005). Under such an assumption, an agent-based approach via graph-based networks would enable the description of changes from complex interactions, which could be mapped back into a spatially explicit form for evaluation. Such an approach could be transferred to EUA via the definition of grid cells or entities as graph nodes. Depending on the input variables, true connections in complex urban networks could be revealed iteratively. The additional implementation of an LSTM module further enables sensitivity for temporal development. This poses an interesting application for ecological network construction, which is informed by development and change, as well as source-sink dynamics (e.g. soil contamination), which require long-term observations. Additionally, this method can be used two observe pair-wise interactions within larger networks.

9.5.6 Change

Similar to Ha and Jeong (2021), but comparably less complex, the potential of the approach of Chen et al. (2021) lies in the ability to predict future change in urban networks, which will enable more efficient planning scenario creation for many use cases, such as to see if a species will be able to access certain habitat patches with the future development of the urban ecological network.

9.5.7 Memory

The study from Wieland et al. (2019) was chosen because it directly addresses the concept of 'landscape memory'. Although at regional scales, it perfectly showcases how ML is capable of extracting fast and slow variables influencing the change of landscapes over long time spans. Such approaches could provide valuable information how variables influenced the long-term development of urban ecosystems by identifying important factors and drivers to determine development trajectories.

9.6 Research Gaps

After the discussion of the most important aspects about applying ML methods to tasks in EUA, to address eco-complexity as spatio-temporal problem, it is vital to delineate future research directions and questions.

9.6.1 Representation and Analytic Workflows for Architecture and Urban Design

The proposal of EUA with this thesis as a conceptual framework was a consequence of the lacking literature about theory and methods combining ecology and urban design into an analytical approach. This framework aims to connect field-specific knowledge in conservation biology to high-level ecological systems research. However, emphasis was put on ecological aspects of urban biodiversity. As cities are considered social-ecological systems, the impact of human agency and the effects of ecological processes on human life and well-being need to be considered equally.

The growing body of theoretical knowledge in landscape and urban ecology has led to a plethora of frameworks, concepts, and theories with a lot of overlap. This is a potential hurdle for interdisciplinary research. Although this thesis tries to overcome some of these inconsistencies, there remains much work. Continuous integration of theoretical and empirical knowledge will be key to providing meaningful results, applicable in urban design and planning (Pickett et al., 2017).

9.6.2 From Theory to Application of EUA

Although there is some evidence, a coherent overview for concrete application scenarios in urban design and planning could not be synthesized in this thesis. Since urban designers are usually lacking expert knowledge in ecology, such an 'inventory' of EUA applications in biodiversity and sustainability planning could lower the threshold for applied research studies.

9.6.3 White Boxes and Ready-To-Use Models

A major drawback of most ML methods is that they do not offer a 'reasoning' for their output. This complicates the interpretation of results, especially if the expertise in the addressed field of research is limited. This problem could either be addressed by developing so-called 'white boxes' which are enabling a certain amount of comprehensibility of the ML output, or by the use of pre-trained models. Such models would be developed by domain experts and their application could be facilitated through transfer learning, helping to involve designers with ML methods.

9.6.4 Algorithm Modification vs. Hybrid Methods

Although algorithms are currently being developed to better fit to specific tasks and data types, many authors in the reviewed literature about ML expressed caution about developing current strains of ML methods to adapt better to specific problems. The main reason is that in many experiments, although specialised algorithms performed better than generic ones, it is not entirely clear, if a comparably small increase in performance is justified looking at the immense resources of time and energy needed to produce and train such models. Especially since design and planning domains are inherently heuristic in their approaches to problem solving, even generic algorithms might produce a 'good enough' solution for an inquiry into urban phenomena. Furthermore, the application of ensemble methods, transfer learning and hybrid methods shows the potential to adapt existing ML methods to specific tasks.

9.6.5 Integration of Spatio-Temporal Sensitivity for the Analysis of Complex Systems

The full potential of using ML methods to uncover unrecognized relationships and patterns within cities from a spatio-temporal perspective, has clearly not been researched sufficiently. The review of potential applications EUA through ML methods has shown that complexity is mainly addressed in spatial or temporal terms, only rarely explicitly as spatio-temporal problem, whereas hierarchy is only considered implicitly through selection of variables. *Complex network theory* in combination with deep learning offers a flexible possibility to analyse complex systems (Lu and Yang, 2022). This puts emphasis on urban ecological networks and GAT as promising coupling. However, there are other computational principles on the rise, such as reservoir computing (Yan et al., 2024), distributed intelligent systems (Guleva et al., 2020), or physics-oriented ML (Karniadakis et al., 2021), which are supporting even more complex aspects of spatio-temporal phenomena.

9.6.6 Bridging Urban and Architectural Scales

As the current standard approach for analysing ecosystems within urban areas is based on two-dimensional representations, patterns and processes at finer resolutions, especially three-dimensional properties, can only be considered as coarse metrics. While this might be sufficient for analytical tasks in urban design, architectural design, such as envisioned by the ECOLOPES project, needs more detailed information. Geometric ML (Bronstein et al. (2017); Wang et al. (2019)) is researching ML methods for Non-Euclidean spatial domains, i.e. three-dimensional complex shapes. This approach to machine learning could hold interesting possibilities to surpass the two-dimensional representation of urban environments, to include three-dimensional details at architectural scales, and promote multi-scales approaches for EUA.

9.6.7 Empirical Evidence for Integrative Research

The reviewed literature in this thesis clearly shows that there is still little methodical knowledge combining natural sciences, and design and planning disciplines. If the already vast theoretical and empirical evidence from urban biodiversity and urban ecology research is to be successfully mainstreamed in urban design and planning tasks, integrative knowledge production could be key.

Although theoretical and conceptual syntheses are important to clarify ambiguities or set foundations for new perspectives and understanding of cities as complex social-ecological systems (Peters and Okin, 2017), empirical research and research by design will be the driving force to promote the potentials of such an emerging design paradigm. Hence, the presented framework will profit from empirical evidence, and future development to establish methodical distinctness.

Furthermore, practical application of EUA through ML methods, implemented into a computational design workflow, together with real-world case studies and long-term evaluation will be necessary to estimate the true potential to improve biodiversity and living conditions in urban areas.

10. Conclusion

This thesis intended to investigate the potentials and challenges for using ML methods in EUA, to facilitate the integration of biodiversity and eco-complexity in urban design. To make a first step towards formulating Ecological Urbanistic Analysis EUA as distinct analytical approach to complex urban social-ecological systems, this thesis integrated literature from multiple scientific fields to propose a conceptual framework. However, the framework remains on an abstract level, and needs to implement reductionist approaches to gain applicability. Also, the framework is limited to ecological aspects, and needs to be developed to incorporate cultural aspects to gain full potential for the application in urban social-ecological systems. Hence, the established theoretical knowledge and framework should be considered as a modular approach, and needs further elaboration and vertical as well as horizontal integration, guided by empirical evidence and research by design. Nevertheless, the synthesized framework should help to further engage with eco-complexity within urban design and architecture from a systemic and holistic perspective, facilitating better understanding for complex pattern-process relationships occurring in urban ecological systems. Also, the established framework needs proof-of-concept through implementation in software solutions and case studies for verification of potentials and shortcomings, as well as development of theory through research by design.

It appears to be an immense undertaking to synchronize conceptual abstractions of complex systems with the ubiquitous and fragmented generation of practical implementations for design-specific affordances. But I fully endorse the call from other scholars for the need to integrate knowledge to a point of universality, where researchers and practitioners from different fields are able to communicate their field-specific results. Machine learning alludes to the idea of dealing with problems where the user lacks intricate knowledge of the studied phenomena. I would cautiously advise not to be overoptimistic about the potentials of this technology in its current state. Experts are needed to formulate meaningful questions and pick best fit data to represent complex phenomena accurately. However, the development of tools in cooperation with domain experts may yield applications for promoting ecological knowledge in urban design and architecture. The potentials of ML to integrate ecology into urban planning and design remain largely untested and offer an interesting new field of research to be developed, as Machine learning (ML) carries the potential to overcome disciplinary limitations in terms of methods and language, as it represents an approach to formal modelling, which is interpretable universally.

Machine learning provides the possibility to address complex problems, without the necessity of deterministic mathematical approaches. However, this comes at a price: analysis of ML outputs is currently lacking interpretability. Explainable AI (xAI) is promising to improve on that trade-off in near future. However, for more complex projects, domain experts will be indispensable for the analytical process to support meaningful results. The reviewed literature suggests that machine learning is still in its infancy when it comes to analysing hierarchically organised, spatio-temporal complex ecological processes at the urban scale. This might be due to the conventions in ecology to assess larger regions, where landscape mosaics are coarser, and patch variation and sizes are better suited for analysis via LULC and similar metrics. This implies, especially for design and planning, to develop a new digital literacy and an openness to deal with ambiguity and degrees of uncertainty, as scientific disciplines have done. As spatial machine learning in the presented context is restricted to two-dimensional spatial representations, where space is limited to 'place-based' measures, future advancements such as geometrical machine learning provide an interesting moonshot to further integrate EUA for interactive responsive integration into the architectural design process. With evolving models, striving for universal AI, such as LLMs (e.g. Grok, ChatGPT, etc.), the accessibility to make use of specialised ML methods might also be lowered.

An intricate understanding of cities as social-ecological complex systems will be key to safeguard future habitat development. If urban design aims to promote biodiversity and multi-species biotopes, high levels of analytic competence to predict future development will be necessary. This means to be able to develop highly individualised, but networked cities, guided by shared knowledge of eco-complexity. The mainstreaming of ecological knowledge into analysis and early-stage urban design will probably not be as visible as many had envisioned before, but a performance-oriented integration might help to alleviate many of today's problems and perhaps even evolve the concept of sustainability to a next stage, where we do not merely try to not limit the possibilities of future generations, but start building a stable foundation on which future generations can thrive.

Bibliography

Ahern, J. Urban landscape sustainability and resilience: the promise and challenges of integrating ecology with urban planning and design. *Landscape Ecology*, 28(6):1203–1212, 2013. ISSN 1572-9761. doi: 10.1007/s10980-012-9799-z. URL `https://doi.org/10.1007/s10980-012-9799-z`.

Alavipanah, S., Haase, D., Lakes, T., and Qureshi, S. Integrating the third dimension into the concept of urban ecosystem services: A review. *Ecological Indicators*, 72:374–398, 2017. ISSN 1470-160X. doi: https://doi.org/10.1016/j.ecolind.2016.08.010. URL `https://www.sciencedirect.com/science/article/pii/S1470160X1630471X`.

Alberti, M. The effects of urban patterns on ecosystem function. *International Regional Science Review*, 28(2):168–192, 2005. doi: 10.1177/0160017605275160. URL `https://journals.sagepub.com/doi/abs/10.1177/0160017605275160`.

Alberti, M. and Marzluff, J. M. Ecological resilience in urban ecosystems: Linking urban patterns to human and ecological functions. *Urban Ecosystems*, 7(3):241–265, 2004. ISSN 1573-1642. doi: 10.1023/B:UECO.0000044038.90173.c6. URL `https://doi.org/10.1023/B:UECO.0000044038.90173.c6`.

Alberti, M., Marzluff, J. M., Shulenberger, E., Bradley, G., Ryan, C., and Zumbrunnen, C. Integrating humans into ecology: Opportunities and challenges for studying urban ecosystems. *BioScience*, 53(12):1169–1179, 2003. ISSN 0006-3568. doi: 10.1641/0006-3568(2003)053[1169:IHIEOA]2.0.CO;2. URL `https://doi.org/10.1641/0006-3568(2003)053[1169:IHIEOA]2.0.CO;2`.

Alberti, M., Palkovacs, E., Roches, S., Meester, L., Brans, K., Govaert, L., Grimm, N. B., Harris, N. C., Hendry, A. P., Schell, C. J., Szulkin, M., Munshi-South, J., Urban, M. C., and Verrelli, B. C. The complexity of urban eco-evolutionary dynamics. *BioScience*, 70(9):772–793, 2020. ISSN 0006-3568. doi: 10.1093/biosci/biaa079. URL `https://doi.org/10.1093/biosci/biaa079`.

Alexander, C., Ishikawa, S., and Silverstein, M. *A Pattern Language: Towns Buildings Construction*. Oxford University Press, New York, 1977.

Andrade, R., Franklin, J., Larson, K. L., Swan, C. M., Lerman, S. B., Bateman, H. L., Warren, P. S., and York, A. Predicting the assembly of novel communities in urban ecosystems. *Landscape Ecology*, 36(1):1–15, 2021. ISSN 1572-9761. doi: 10.1007/s10980-020-01142-1. URL `https://doi.org/10.1007/s10980-020-01142-1`.

Anselin, L. Local indicators of spatial association—lisa. *Geographical Analysis*, 27(2):93–115, 1995. ISSN 0016-7363. doi: https://doi.org/10.1111/j.1538-4632.1995.tb00338.x. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1538-4632.1995.tb00338.x`.

Apfelbeck, B., Jakoby, C., Hanusch, M., Steffani, E. B., Hauck, T. E., and Weisser, W. W. A conceptual framework for choosing target species for wildlife-inclusive urban design. *Sustainability*, 11(24):6972, 2019. ISSN 2071-1050. URL `https://www.mdpi.com/2071-1050/11/24/6972`.

Apfelbeck, B., Snep, R. P. H., Hauck, T. E., Ferguson, J., Holy, M., Jakoby, C., Scott MacIvor, J., Schär, L., Taylor, M., and Weisser, W. W. Designing wildlife-inclusive cities that support human-animal co-existence. *Landscape and Urban Planning*, 200:103817, 2020. ISSN 0169-2046. doi: https://doi.org/10.1016/j.landurbplan.2020.103817. URL `https://www.sciencedirect.com/science/article/pii/S0169204619308035`.

Araldi, A., Emsellem, D., Fusco, G., Tettamanzi, A., and Overal, D. Exploring building typologies through fast iterative bayesian clustering. SAGEO 2021 Proceedings, Avignon, UMR ESPACE, pages 113–124, La Rochelle, France, 2021. ISBN 978-2-910545-12-1. URL `https://hal.science/hal-03228379`.

Ardiantiono, Deere, N. J., Ramadiyanta, E., Sibarani, M. C., Hadi, A. N., Andayani, N., Ginting, Y., Bull, J. W., and Struebig, M. J. Selecting umbrella species as mammal biodiversity indicators in tropical forest. *Biological Conservation*, 292:110511, 2024. ISSN 0006-3207. doi: https://doi.org/10.1016/j.biocon.2024.110511. URL https://www.sciencedirect.com/science/article/pii/S0006320724000727.

Aronson, M. F. J., La Sorte, F. A., Nilon, C. H., Katti, M., Goddard, M. A., Lepczyk, C. A., Warren, P. S., Williams, N. S. G., Cilliers, S., Clarkson, B., Dobbs, C., Dolan, R., Hedblom, M., Klotz, S., Kooijmans, J. L., Kühn, I., MacGregor-Fors, I., McDonnell, M., Mörtberg, U., Pyšek, P., Siebert, S., Sushinsky, J., Werner, P., and Winter, M. A global analysis of the impacts of urbanization on bird and plant diversity reveals key anthropogenic drivers. *Proceedings of the Royal Society B: Biological Sciences*, 281(1780), 2014. doi: doi:10.1098/rspb.2013.3330. URL https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2013.3330.

Aronson, M. F. J., Nilon, C. H., Lepczyk, C. A., Parker, T. S., Warren, P. S., Cilliers, S. S., Goddard, M. A., Hahs, A. K., Herzog, C., Katti, M., La Sorte, F. A., Williams, N. S. G., and Zipperer, W. Hierarchical filters determine community assembly of urban species pools. *Ecology*, 97(11):2952–2963, 2016. ISSN 0012-9658. doi: https://doi.org/10.1002/ecy.1535. URL https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1002/ecy.1535.

Ascher, W. Coping with complexity and organizational interests in natural resource management. *Ecosystems*, 4(8):742–757, 2001. ISSN 1435-0629. doi: 10.1007/s10021-001-0043-y. URL https://doi.org/10.1007/s10021-001-0043-y.

AssunÇão, R. M., Neves, M. C., Câmara, G., and Da Costa Freitas, C. Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees. *International Journal of Geographical Information Science*, 20(7): 797–811, 2006. ISSN 1365-8816. doi: 10.1080/13658810600665111. URL https://doi.org/10.1080/13658810600665111. doi: 10.1080/13658810600665111.

Badami, M., Benatallah, B., and Baez, M. *Advanced Information Systems Engineering*, book section Systematic Literature Review Search Query Refinement Pipeline: Incremental Enrichment and Adaptation, pages 129–146. Springer International Publishing, 2022. ISBN 978-3-031-07472-1.

Balogun, A.-L., Tella, A., Baloo, L., and Adebisi, N. A review of the inter-correlation of climate change, air pollution and urban sustainability using novel machine learning algorithms and spatial information science. *Urban Climate*, 40: 100989, 2021. ISSN 2212-0955. doi: https://doi.org/10.1016/j.uclim.2021.100989. URL https://www.sciencedirect.com/science/article/pii/S2212095521002194.

Barbierato, E., Bernetti, I., Capecchi, I., and Saragosa, C. Remote sensing and urban metrics: An automatic classification of spatial configurations to support urban policies. *Trends in earth observation*, 1:187–190, 2019.

Barbierato, E., Bernetti, I., Capecchi, I., and Saragosa, C. Integrating remote sensing and street view images to quantify urban forest ecosystem services. *Remote Sensing*, 12(2), 2020. ISSN 2072-4292. URL https://www.mdpi.com/2072-4292/12/2/329.

Bascompte, J. Disentangling the web of life. *Science*, 325(5939):416–419, 2009. doi: doi:10.1126/science.1170749. URL https://www.science.org/doi/abs/10.1126/science.1170749.

Battesini, M., ten Caten, C. S., and Pacheco, D. A. d. J. Key factors for operational performance in manufacturing systems: Conceptual model, systematic literature review and implications. *Journal of Manufacturing Systems*, 60:265–282, 2021. ISSN 0278-6125. doi: https://doi.org/10.1016/j.jmsy.2021.06.005. URL https://www.sciencedirect.com/science/article/pii/S0278612521001291.

Battin, J. When good animals love bad habitats: Ecological traps and the conservation of animal populations. *Conservation Biology*, 18(6), 2004.

Batty, M. Agents, cells, and cities: New representational models for simulating multiscale urban dynamics. *Environment and Planning A: Economy and Space*, 37(8):1373–1394, 2005. doi: 10.1068/a3784. URL https://journals.sagepub.com/doi/abs/10.1068/a3784.

Beltagy, I., Lo, K., and Cohan, A. Scibert: A pretrained language model for scientific text. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pages 3615–3620. Association for Computational Linguistics, 2019. doi: 10.18653/v1/D19-1371. URL https://aclanthology.org/D19-1371https://doi.org/10.18653/v1/D19-1371.

Belter, C. W. Citation analysis as a literature search method for systematic reviews. *Journal of the Association for Information Science and Technology*, 67(11):2766–2777, 2016. ISSN 2330-1635. doi: https://doi.org/10.1002/asi.23605. URL https://doi.org/10.1002/asi.23605.

Benda, L. E., Poff, L. N., Tague, C., Palmer, M. A., Pizzuto, J., Cooper, S., Stanley, E., and Moglen, G. How to avoid train wrecks when using science in environmental problem solving. *BioScience*, 52(12):1127–1136, 2002. ISSN 0006-3568. doi: 10.1641/0006-3568(2002)052[1127:HTATWW]2.0.CO;2. URL https://doi.org/10.1641/0006-3568(2002)052[1127:HTATWW]2.0.CO;2.

Beninde, J., Veith, M., and Hochkirch, A. Biodiversity in cities needs space: a meta-analysis of factors determining intra-urban biodiversity variation. *Ecology Letters*, 18(6):581–592, 2015. ISSN 1461-023X. doi: https://doi.org/10.1111/ele.12427. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ele.12427.

Berghauser Pont, M., Ahrné, K., Gren, , Kaczorowska, A., and Marcus, L. Integrating visibility graph analysis (vga) with connectivity analysis in landscape ecology. Proceedings of the 11th Space Syntax Symposium, 2017.

Bestelmeyer, B. T., Ash, A., Brown, J. R., Densambuu, B., Fernández-Giménez, M., Johanson, J., Levi, M., Lopez, D., Peinetti, R., Rumpff, L., and Shaver, P. *Rangeland Systems: Processes, Management and Challenges*, book section State and Transition Models: Theory, Applications, and Challenges, pages 303–345. Springer International Publishing, Cham, 2017. ISBN 978-3-319-46709-2. doi: 10.1007/978-3-319-46709-2_9. URL https://doi.org/10.1007/978-3-319-46709-2_9.

Biljecki, F. and Chow, Y. S. Global building morphology indicators. *Computers, Environment and Urban Systems*, 95:101809, 2022. ISSN 0198-9715. doi: https://doi.org/10.1016/j.compenvurbsys.2022.101809. URL https://www.sciencedirect.com/science/article/pii/S0198971522000539.

Booth, A., Papaioannou, D., and Sutton, A. *Systematic Approaches to a Successful Literature Review*. SAGE Publications Ltd, London, United Kingdom, 2012.

Botequilha Leitão, A. and Ahern, J. Applying landscape ecological concepts and metrics in sustainable landscape planning. *Landscape and Urban Planning*, 59(2):65–93, 2002. ISSN 0169-2046. doi: https://doi.org/10.1016/S0169-2046(02)00005-1. URL https://www.sciencedirect.com/science/article/pii/S0169204602000051.

Breiman, L. Random forests. *Machine Learning*, 45(1):5–32, 2001. ISSN 1573-0565. doi: 10.1023/A:1010933404324. URL https://doi.org/10.1023/A:1010933404324.

Brierley, G. J. Landscape memory: the imprint of the past on contemporary landscape forms and processes. *Area*, 42(1):76–85, 2010. ISSN 0004-0894. doi: https://doi.org/10.1111/j.1475-4762.2009.00900.x. URL https://rgs-ibg.onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-4762.2009.00900.x.

Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A., and Vandergheynst, P. Geometric deep learning: Going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017. ISSN 1558-0792. doi: 10.1109/MSP.2017.2693418.

Buchmann, C. M., Schurr, F. M., Nathan, R., and Jeltsch, F. An allometric model of home range formation explains the structuring of animal communities exploiting heterogeneous resources. *Oikos*, 120(1):106–118, 2011. ISSN 0030-1299. doi: https://doi.org/10.1111/j.1600-0706.2010.18556.x. URL https://doi.org/10.1111/j.1600-0706.2010.18556.x.

Bunce, R. G. H., Metzger, M. J., Jongman, R. H. G., Brandt, J., de Blust, G., Elena-Rossello, R., Groom, G. B., Halada, L., Hofer, G., Howard, D. C., Kovář, P., Mücher, C. A., Padoa-Schioppa, E., Paelinx, D., Palo, A., Perez-Soba, M., Ramos, I. L., Roche, P., Skånes, H., and Wrbka, T. A standardized procedure for surveillance and monitoring european habitats and provision of spatial data. *Landscape Ecology*, 23(1):11–25, 2008. ISSN 1572-9761. doi: 10.1007/s10980-007-9173-8. URL https://doi.org/10.1007/s10980-007-9173-8.

Bunce, R., Groom, G., Jongman, R., Padoa-Schippa, E., and Metzger, M. *Handbook for Surveillance and Monitoring of European Habitats. First Edition*. Alterra-rapport. Alterra, Netherlands „ 2005.

Byrne, L. B. Habitat structure: A fundamental concept and framework for urban soil ecology. *Urban Ecosystems*, 10(3): 255–274, 2007. ISSN 1573-1642. doi: 10.1007/s11252-007-0027-6. URL `https://doi.org/10.1007/s11252-007-0027-6`.

Börger, L., Dalziel, B. D., and Fryxell, J. M. Are there general mechanisms of animal home range behaviour? a review and prospects for future research. *Ecology Letters*, 11(6):637–650, 2008. ISSN 1461-023X. doi: https://doi.org/10.1111/j.1461-0248.2008.01182.x. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1461-0248.2008.01182.x`.

Cadenasso, M. L., Pickett, S. T. A., and Grove, J. M. Dimensions of ecosystem complexity: Heterogeneity, connectivity, and history. *Ecological Complexity*, 3(1):1–12, 2006a. ISSN 1476-945X. doi: https://doi.org/10.1016/j.ecocom.2005.07.002. URL `https://www.sciencedirect.com/science/article/pii/S1476945X05000838`.

Cadenasso, M. L., Pickett, S. T. A., Weathers, K. C., and Jones, C. G. A framework for a theory of ecological boundaries. *BioScience*, 53(8):750–758, 2003. ISSN 0006-3568. doi: 10.1641/0006-3568(2003)053[0750:Affato]2.0.Co;2. URL `https://doi.org/10.1641/0006-3568(2003)053[0750:AFFATO]2.0.CO;2`.

Cadenasso, M. L., Pickett, S. T., and Grove, M. J. Integrative approaches to investigating human-natural systems: the baltimore ecosystem study. *Natures Sciences Sociétés*, 14(1):4–14, 2006b. doi: https://doi.org/10.1051/nss:2006002. URL `http://www.sciengine.com/publisher/EDPSciences/journal/NaturesSciencesSociÃĪtÃĪs/14/1/10.1051/nss:2006002`.

Cadenasso, M. L., Pickett, S. T. A., and Schwarz, K. Spatial heterogeneity in urban ecosystems: reconceptualizing land cover and a framework for classification. *Frontiers in Ecology and the Environment*, 5(2):80–88, 2007. ISSN 1540-9295. doi: https://doi.org/10.1890/1540-9295(2007)5[80:SHIUER]2.0.CO;2. URL `https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/1540-9295%282007%295%5B80%3ASHIUER%5D2.0.CO%3B2`.

Cadenasso, M. L., Pickett, S. T. A., McGrath, B., and Marshall, V. J. *Resilience in Ecology and Urban Design: Linking Theory and Practice for Sustainable Cities*, book section Ecological Heterogeneity in Urban Ecosystems: Reconceptualized Land Cover Models as a Bridge to Urban Design. Springer Science+Business Media, Dordrecht, Netherlands, 2013.

Cai, C., Guo, Z., Zhang, B., Wang, X., Li, B., and Tang, P. Urban morphological feature extraction and multi-dimensional similarity analysis based on deep learning approaches. *Sustainability*, 13(12):6859, 2021. ISSN 2071-1050. URL `https://www.mdpi.com/2071-1050/13/12/6859`.

Canepa, M., Mosca, F., Barath, S., Changenet, A., Hauck, T. E., F., L., Pianta, M., Roccotiello, E., Selvan, S. U., Vogler, V., and Perini, K. Ecolopes, beyond greening - a multi-species approach for urban design. *AGATHÓN – International Journal of Architecture*, (11):8, 2022. ISSN 2464-9309. doi: doi.org/10.19229/2464-9309/11212022.

Carta, S. *Machine Learning and the City*, book section Introduction. John Wiley & Sons Ltd, 2022. ISBN 9781119815075. doi: https://doi.org/10.1002/9781119815075.ch1.

Casali, Y., Aydin, N. Y., and Comes, T. Machine learning for spatial analyses in urban areas: a scoping review. *Sustainable Cities and Society*, 85:104050, 2022. ISSN 2210-6707. doi: https://doi.org/10.1016/j.scs.2022.104050. URL `https://www.sciencedirect.com/science/article/pii/S2210670722003687`.

Casiker, C. V., Jagadishakumara, B., Sunil, G. M., Chaithra, K., and Devy, M. S. *The Rural-Urban Interface: An Interdisciplinary Research Approach to Urbanisation Processes Around the Indian Megacity Bengaluru*, book section Bee Diversity in the Rural–Urban Interface of Bengaluru and Scope for Pollinator-Integrated Urban Agriculture, pages 171–182. Springer International Publishing, Cham, 2021. ISBN 978-3-030-79972-4. doi: 10.1007/978-3-030-79972-4_18. URL `https://doi.org/10.1007/978-3-030-79972-4_18`.

Chaillou, S. *Artificial Intelligence and Architecture - From Research to Practice*. Birkhäuser, Berlin, Boston, 2022. ISBN 9783035624045. doi: doi:10.1515/9783035624045. URL `https://doi.org/10.1515/9783035624045`.

Charmaz, K. *Constructing grounded theory: A practical guide through qualitative analysis*. Sage, London, 2006.

Chase, J. M., Jeliazkov, A., Ladouceur, E., and Viana, D. S. Biodiversity conservation through the lens of metacommunity ecology. *Annals of the New York Academy of Sciences*, 1469(1):86–104, 2020. ISSN 0077-8923. doi: https://doi.org/10.1111/nyas.14378. URL https://nyaspubs.onlinelibrary.wiley.com/doi/abs/10.1111/nyas.14378.

Chaturvedi, V. and de Vries, W. T. Machine learning algorithms for urban land use planning: A review. *Urban Science*, 5 (3):68, 2021. ISSN 2413-8851. URL https://www.mdpi.com/2413-8851/5/3/68.

Chen, J., Zhang, J., Xu, X., Fu, C., Zhang, D., Zhang, Q., and Xuan, Q. E-lstm-d: A deep learning framework for dynamic network link prediction. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51(6):3699–3712, 2021. ISSN 2168-2232. doi: 10.1109/TSMC.2019.2932913.

Chen, L. and Ma, Y. A framework for assessing trade-offs and synergies in green space system services based on ecosystem services bundles. *Forests*, 14(8):1614, 2023. ISSN 1999-4907. URL https://www.mdpi.com/1999-4907/14/8/1614.

Chen, S., Chen, B., and Fath, B. D. Urban ecosystem modeling and global change: Potential for rational urban management and emissions mitigation. *Environmental Pollution*, 190:139–149, 2014. ISSN 0269-7491. doi: https://doi.org/10.1016/j.envpol.2014.03.032. URL https://www.sciencedirect.com/science/article/pii/S0269749114001237.

Chen, S., Zhang, M., and Lei, F. Mapping vegetation types by different fully convolutional neural network structures with inadequate training labels in complex landscape urban areas. *Forests*, 14(9):1788, 2023. ISSN 1999-4907. URL https://www.mdpi.com/1999-4907/14/9/1788.

Colding, J. 'ecological land-use complementation' for building resilience in urban ecosystems. *Landscape and Urban Planning*, 81(1):46–55, 2007. ISSN 0169-2046. doi: https://doi.org/10.1016/j.landurbplan.2006.10.016. URL https://www.sciencedirect.com/science/article/pii/S0169204606002179.

Colléony, A. and Shwartz, A. When the winners are the losers: Invasive alien bird species outcompete the native winners in the biotic homogenization process. *Biological Conservation*, 241:108314, 2020. ISSN 0006-3207. doi: https://doi.org/10.1016/j.biocon.2019.108314. URL https://www.sciencedirect.com/science/article/pii/S000632071931434X.

Corry, R. C. and Nassauer, J. I. Limitations of using landscape pattern indices to evaluate the ecological consequences of alternative plans and designs. *Landscape and Urban Planning*, 72(4):265–280, 2005. ISSN 0169-2046. doi: https://doi.org/10.1016/j.landurbplan.2004.04.003. URL https://www.sciencedirect.com/science/article/pii/S0169204604000854.

Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R. V., Paruelo, J., Raskin, R. G., Sutton, P., and van den Belt, M. The value of the world's ecosystem services and natural capital. *Nature*, 387(6630):253–260, 1997. ISSN 1476-4687. doi: 10.1038/387253a0. URL https://doi.org/10.1038/387253a0.

Cressie, N. A. C. *Statistics for Spatial Data*. John Wiley & Sons, Inc., New York, 1993. ISBN 0-471-00255-0.

Crisci, C., Ghattas, B., and Perera, G. A review of supervised machine learning algorithms and their applications to ecological data. *Ecological Modelling*, 240:113–122, 2012. ISSN 0304-3800. doi: https://doi.org/10.1016/j.ecolmodel.2012.03.001. URL https://www.sciencedirect.com/science/article/pii/S0304380012001081.

Cullum, C., Brierley, G., Perry, G. L., and Witkowski, E. T. Landscape archetypes for ecological classification and mapping:the virtue of vagueness. *Progress in Physical Geography: Earth and Environment*, 41(1):95–123, 2017. doi: 10.1177/0309133316671103. URL https://journals.sagepub.com/doi/abs/10.1177/0309133316671103.

Defila, R. and Di Giulio, A. Integrating knowledge: Challenges raised by the "inventory of synthesis". *Futures*, 65:123–135, 2015. ISSN 0016-3287. doi: https://doi.org/10.1016/j.futures.2014.10.013. URL https://www.sciencedirect.com/science/article/pii/S0016328714001748.

Desjardins-Proulx, P., Poisot, T., and Gravel, D. Artificial intelligence for ecological and evolutionary synthesis. *Frontiers in Ecology and Evolution*, 7, 2019. ISSN 2296-701X. doi: 10.3389/fevo.2019.00402. URL https://www.frontiersin.org/articles/10.3389/fevo.2019.00402.

Diamond, S. E. and Martin, R. A. Evolution in cities. *Annual Review of Ecology, Evolution and Systematics*, 52(Volume 52, 2021):519–540, 2021. ISSN 1545-2069. doi: https://doi.org/10.1146/annurev-ecolsys-012021-021402. URL https://www.annualreviews.org/content/journals/10.1146/annurev-ecolsys-012021-021402.

Dramstad, W. E., Olson, J. D., and Forman, R. T. T. *Landscape Ecology Principles in Landscape Architecture and Land-Use Planning*. Island Press, Washington D.C., USA, 1996.

Dunne, C. The place of the literature review in grounded theory research. *International Journal of Social Research Methodology*, 14(2):111–124, 2011. ISSN 1364-5579. doi: 10.1080/13645579.2010.494930. URL `https://doi.org/10.1080/13645579.2010.494930`. doi: 10.1080/13645579.2010.494930.

Dunning, J. B., Danielson, B. J., and Pulliam, H. R. Ecological processes that affect populations in complex landscapes. *Oikos*, 65(1):169–175, 1992. ISSN 00301299, 16000706. doi: 10.2307/3544901. URL `http://www.jstor.org/stable/3544901`.

Döhren, P. and Haase, D. Ecosystem disservices research: A review of the state of the art with a focus on cities. *Ecological Indicators*, 52, 2015. URL `https://doi.org/10.1016/j.ecolind.2014.12.027`.

Ellis, E. C., Li, R. G., Yang, L. Z., and Cheng, X. Long-term change in village-scale ecosystems in china using landscape and statistical methods. *Ecological Applications*, 10(4):1057–1073, 2000. ISSN 1051-0761. doi: https://doi.org/10.1890/1051-0761(2000)010[1057:LTCIVS]2.0.CO;2. URL `https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/1051-0761%282000%29010%5B1057%3ALTCIVS%5D2.0.CO%3B2`.

Elmqvist, T., Andersson, E., Frantzeskaki, N., McPhearson, T., Olsson, P., Gaffney, O., Takeuchi, K., and Folke, C. Sustainability and resilience for transformation in the urban century. *Nature Sustainability*, 2(4):267–273, 2019. ISSN 2398-9629. doi: 10.1038/s41893-019-0250-1. URL `https://doi.org/10.1038/s41893-019-0250-1`.

Elsbach, K. D. and van Knippenberg, D. Creating high-impact literature reviews: An argument for 'integrative reviews'. *Journal of Management Studies*, 57(6):1277–1289, 2020. ISSN 0022-2380. doi: https://doi.org/10.1111/joms.12581. URL `https://doi.org/10.1111/joms.12581`.

European Commission. Ecological building envelopes: a game-changing design approach for regenerative urban ecosystems, 2023-12-28 . URL `https://cordis.europa.eu/project/id/964414`. last accessed: 2024-04-14.

Fan, C., Zeng, L., Sun, Y., and Liu, Y.-Y. Finding key players in complex networks through deep reinforcement learning. *Nature Machine Intelligence*, 2(6):317–324, 2020. ISSN 2522-5839. doi: 10.1038/s42256-020-0177-2. URL `https://doi.org/10.1038/s42256-020-0177-2`.

Fang, Z., Qi, J., Yang, T., Wan, L., and Jin, Y. "reading" cities with computer vision: a new multi-spatial scale urban fabric dataset and a novel convolutional neural network solution for urban fabric classification tasks. Proceedings of the 28th International Conference on Advances in Geographic Information Systems, page 507–517. Association for Computing Machinery, 2020. doi: 10.1145/3397536.3422240. URL `https://doi.org/10.1145/3397536.3422240`.

Farinha-Marques, P., Fernandes, C., Guilherme, F., Lameiras, J. M., Alves, P., and Bunce, R. G. H. Urban habitats biodiversity assessment (urhba): a standardized procedure for recording biodiversity and its spatial distribution in urban environments. *Landscape Ecology*, 32(9):1753–1770, 2017. ISSN 1572-9761. doi: 10.1007/s10980-017-0554-3. URL `https://doi.org/10.1007/s10980-017-0554-3`.

Fineschi, S. and Loreto, F. A survey of multiple interactions between plants and the urban environment. *Frontiers in Forests and Global Change*, 3(30), 2020. ISSN 2624-893X. URL `https://www.frontiersin.org/articles/10.3389/ffgc.2020.00030`.

García-Pardo, K. A., Moreno-Rangel, D., Domínguez-Amarillo, S., and García-Chávez, J. R. Urban classification of the built-up and seasonal variations in vegetation: A framework integrating multisource datasets. *Urban Forestry & Urban Greening*, 89:128114, 2023. ISSN 1618-8667. doi: https://doi.org/10.1016/j.ufug.2023.128114. URL `https://www.sciencedirect.com/science/article/pii/S1618866723002856`.

Garousi, V. and Felderer, M. *Experience-based guidelines for effective and efficient data extraction in systematic reviews in software engineering*. 2017. doi: 10.1145/3084226.3084238.

Garritty, C., Stevens, A., Hamel, C., Golfam, M., Hutton, B., and Wolfe, D. Knowledge synthesis in evidence-based medicine. *Semin Nucl Med*, 49(2):136–144, 2019. ISSN 1558-4623 (Electronic) 0001-2998 (Linking). doi: 10.1053/j.semnuclmed.2018.11.006. URL https://www.ncbi.nlm.nih.gov/pubmed/30819393. Garritty, C Stevens, A Hamel, C Golfam, M Hutton, B Wolfe, D eng Research Support, Non-U.S. Gov't Review 2019/03/02 Semin Nucl Med. 2019 Mar;49(2):136-144. doi: 10.1053/j.semnuclmed.2018.11.006. Epub 2019 Feb 4.

Geary, W. L., Bode, M., Doherty, T. S., Fulton, E. A., Nimmo, D. G., Tulloch Ayesha, I. T., Tulloch Vivitskaia, J. D., and Ritchie, E. G. A guide to ecosystem models and their environmental applications. *Nature Ecology & Evolution*, 4 (11):1459–1471, 2020. doi: https://doi.org/10.1038/s41559-020-01298-8. Copyright - © The Author(s), under exclusive licence to Springer Nature Limited 2020.

Gentili, R., Quaglini, L. A., Galasso, G., Montagnani, C., Caronni, S., Cardarelli, E., and Citterio, S. Urban refugia sheltering biodiversity across world cities. *Urban Ecosystems*, 27(1):219–230, 2024. ISSN 1573-1642. doi: 10.1007/s11252-023-01432-x. URL https://doi.org/10.1007/s11252-023-01432-x.

Geronimus, A. T. and Bound, J. Use of census-based aggregate variables to proxy for socioeconomic group: Evidence from national samples. *American Journal of Epidemiology*, 148(5):475–486, 1998. ISSN 0002-9262. doi: 10.1093/oxfordjournals.aje.a009673. URL https://doi.org/10.1093/oxfordjournals.aje.a009673.

Goddard, M. A., Dougill, A. J., and Benton, T. G. Scaling up from gardens: biodiversity conservation in urban environments. *Trends in Ecology & Evolution*, 25(2):90–98, 2010. ISSN 0169-5347. doi: https://doi.org/10.1016/j.tree.2009.07.016. URL https://www.sciencedirect.com/science/article/pii/S0169534709002468.

Goodchild, M. and Janelle, D. *Spatially Integrated Social Science*, book section Thinking Spatially in the Social Sciences, pages 3–17. Oxford Academic, New York, USA, 2004. ISBN 0-19-515270-0. doi: 10.1093/oso/9780195152708.003.0001.

Grames, E. M., Stillman, A. N., Tingley, M. W., and Elphick, C. S. An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. *Methods in Ecology and Evolution*, 10(10):1645–1654, 2019. ISSN 2041-210X. doi: https://doi.org/10.1111/2041-210X.13268. URL https://doi.org/10.1111/2041-210X.13268.

Grekousis, G. Artificial neural networks and deep learning in urban geography: A systematic review and meta-analysis. *Computers, Environment and Urban Systems*, 74:244–256, 2019. ISSN 0198-9715. doi: https://doi.org/10.1016/j.compenvurbsys.2018.10.008. URL https://www.sciencedirect.com/science/article/pii/S0198971518302928.

Grilo, F., McPhearson, T., Santos-Reis, M., and Branquinho, C. A trait-based conceptual framework to examine urban biodiversity, socio-ecological filters, and ecosystem services linkages. *npj Urban Sustainability*, 2(1):32, 2022. ISSN 2661-8001. doi: 10.1038/s42949-022-00077-7. URL https://doi.org/10.1038/s42949-022-00077-7.

Grobman, Y. J., Weisser, W., Shwartz, A., Ludwig, F., Kozlovsky, R., Ferdman, A., Perini, K., Hauck, T. E., Selvan, S. U., Saroglou, S., Barath, S., Schloter, M., and Windorfer, L. Architectural multispecies building design: Concepts, challenges, and design process. *Sustainability*, 15(21):15480, 2023. ISSN 2071-1050. URL https://www.mdpi.com/2071-1050/15/21/15480.

Groffman, P. M., Cadenasso, M. L., Cavender-Bares, J., Childers, D. L., Grimm, N. B., Grove, J. M., Hobbie, S. E., Hutyra, L. R., Darrel Jenerette, G., McPhearson, T., Pataki, D. E., Pickett, S. T. A., Pouyat, R. V., Rosi-Marshall, E., and Ruddell, B. L. Moving towards a new urban systems science. *Ecosystems*, 20(1):38–43, 2017. ISSN 1435-0629. doi: 10.1007/s10021-016-0053-4. URL https://doi.org/10.1007/s10021-016-0053-4.

Grove, J. M., Burch, W. R. J., and Pickett, S. T. A. *Community and Forestry: Continuities in the Sociology of Natural Resources*, book section Social Mosaics and Urban Forestry in Baltimore, Maryland. University Washington Press, Seattle, 2005.

Guleva, V., Shikov, E., Bochenina, K., Kovalchuk, S., Alodjants, A., and Boukhanovsky, A. Emerging complexity in distributed intelligent systems. *Entropy*, 22(12):1437, 2020. ISSN 1099-4300. URL https://www.mdpi.com/1099-4300/22/12/1437.

Guo, D. Regionalization with dynamically constrained agglomerative clustering and partitioning (redcap). *International Journal of Geographical Information Science*, 22(7):801–823, 2008. ISSN 1365-8816. doi: 10.1080/13658810701674970. URL https://doi.org/10.1080/13658810701674970. doi: 10.1080/13658810701674970.

Guo, Z. and Liu, X. How artificial intelligence cooperating with agent-based modeling for urban studies: A systematic review. *Transactions in GIS*, n/a(n/a), 2024. ISSN 1361-1682. doi: https://doi.org/10.1111/tgis.13152. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/tgis.13152.

Ha, S. and Jeong, H. Unraveling hidden interactions in complex systems with deep learning. *Scientific Reports*, 11(1): 12804, 2021. ISSN 2045-2322. doi: 10.1038/s41598-021-91878-w. URL https://doi.org/10.1038/s41598-021-91878-w.

Hahs, A. K., McDonnell, M. J., McCarthy, M. A., Vesk, P. A., Corlett, R. T., Norton, B. A., Clemants, S. E., Duncan, R. P., Thompson, K., Schwartz, M. W., and Williams, N. S. G. A global synthesis of plant extinction rates in urban areas. *Ecology Letters*, 12(11):1165–1173, 2009. ISSN 1461-023X. doi: https://doi.org/10.1111/j.1461-0248.2009.01372.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1461-0248.2009.01372.x.

Hair, K., Bahor, Z., Macleod, M., Liao, J., and Sena, E. S. The automated systematic search deduplicator (asysd): a rapid, open-source, interoperable tool to remove duplicate citations in biomedical systematic reviews. *BMC Biology*, 21(1):189, 2023. ISSN 1741-7007. doi: 10.1186/s12915-023-01686-z. URL https://doi.org/10.1186/s12915-023-01686-z.

Hanna, S. *Machine Learning and the City*, book section Urban Complexity, pages 1–13. Wiley & Sons Ltd, 2022. ISBN 9781119815075. doi: https://doi.org/10.1002/9781119815075.ch1. URL https://doi.org/10.1002/9781119815075.ch1.

Hastings, A. Timescales, dynamics, and ecological understanding. *Ecology*, 91(12):3471–3480, 2010. ISSN 0012-9658. doi: https://doi.org/10.1890/10-0776.1. URL https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/10-0776.1.

Hauck, T. E. and Weisser, W. W. Aad - animal aided design, 2015. ISBN 978-3-00-047519-1.

Hensel, M. *Performance-Oriented Architecture: Rethinking Architectural Design and the Built Environment*. John Wiley & Sons Ltd, 2013.

Heymans, A., Breadsell, J., Morrison, G. M., Byrne, J. J., and Eon, C. Ecological urban planning and design: A systematic literature review. *Sustainability*, 11(13):3723, 2019. ISSN 2071-1050. URL https://www.mdpi.com/2071-1050/11/13/3723.

Hillier, B. *Space is the machine: A configurational theory of architecture*. Space Syntax, London, UK, 2007.

Hirt, J., Nordhausen, T., Appenzeller-Herzog, C., and Ewald, H. Citation tracking for systematic literature searching: A scoping review. *Research Synthesis Methods*, 14(3):563–579, 2023. ISSN 1759-2879. doi: https://doi.org/10.1002/jrsm.1635. URL https://doi.org/10.1002/jrsm.1635.

Holling, C. S. Cross-scale morphology, geometry, and dynamics of ecosystems. *Ecological Monographs*, 62(4):447–502, 1992. ISSN 0012-9615. doi: https://doi.org/10.2307/2937313. URL https://doi.org/10.2307/2937313.

Hong, S.-K., Kim, S., Cho, K.-H., Kim, J.-E., Kang, S., and Lee, D. Ecotope mapping for landscape ecological assessment of habitat and ecosystem. *Ecological Research*, 19(1):131–139, 2004. ISSN 0912-3814. doi: https://doi.org/10.1111/j.1440-1703.2003.00603.x. URL https://esj-journals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1440-1703.2003.00603.x.

Hopia, H., Latvala, E., and Liimatainen, L. Reviewing the methodology of an integrative review. *Scandinavian Journal of Caring Sciences*, 30(4):662–669, 2016. ISSN 0283-9318. doi: https://doi.org/10.1111/scs.12327. URL https://doi.org/10.1111/scs.12327.

Hostetler, M., Allen, W., and Meurk, C. Conserving urban biodiversity? creating green infrastructure is only the first step. *Landscape and Urban Planning*, 100(4):369–371, 2011. ISSN 0169-2046. doi: https://doi.org/10.1016/j.landurbplan.2011.01.011. URL https://www.sciencedirect.com/science/article/pii/S016920461100048X.

Hsieh, W. W. *Machine Learning Methods in the Environmental Sciences: Neural Networks and Kernels*. Cambridge University Press, Cambridge, 2009. ISBN 9780521791922. doi: DOI:10.1017/CBO9780511627217. URL https://www.cambridge.org/core/books/machine-learning-methods-in-the-environmental-sciences/494FDC330DA43C8AE8B649D897B441C6.

Ives, C. D., Lentini, P. E., Threlfall, C. G., Ikin, K., Shanahan, D. F., Garrard, G. E., Bekessy, S. A., Fuller, R. A., Mumaw, L., Rayner, L., Rowe, R., Valentine, L. E., and Kendal, D. Cities are hotspots for threatened species. *Global Ecology and Biogeography*, 25(1):117–126, 2016. ISSN 1466-822X. doi: https://doi.org/10.1111/geb.12404. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/geb.12404.

Jacobs, J. *The Death and Life of Great American Cities*. Vintage Books, New York, 1961.

Janowicz, K., Gao, S., McKenzie, G., Hu, Y., and Bhaduri, B. Geoai: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *International Journal of Geographical Information Science*, 34(4):625–636, 2020. ISSN 1365-8816. doi: 10.1080/13658816.2019.1684500. URL https://doi.org/10.1080/13658816.2019.1684500. doi: 10.1080/13658816.2019.1684500.

Jemeļjanova, M., Kmoch, A., and Uuemaa, E. Adapting machine learning for environmental spatial data - a review. *Ecological Informatics*, 81:102634, 2024. ISSN 1574-9541. doi: https://doi.org/10.1016/j.ecoinf.2024.102634. URL https://www.sciencedirect.com/science/article/pii/S1574954124001766.

Joshi, A. V. *Machine Learning and Artificial Intelligence*. Springer Cham, 2nd edition, 2023.

Jung, A. *Machine Learning - The Basics*. Machine Learning: Foundations, Methodologies, and Applications. Springer Singapore, 2022. ISBN 978-981-16-8195-0.

Kaczorowska, A. and Pont, M. B. Modelling urban environments to promote ecosystem services and biodiversity: Case of stockholm. *International Journal of E-Planning Research (IJEPR)*, 8(3):1–12, 2019. ISSN 2160-9918. doi: 10.4018/IJEPR.2019070101. URL https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/IJEPR.2019070101.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., and Yang, L. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021. ISSN 2522-5820. doi: 10.1038/s42254-021-00314-5. URL https://doi.org/10.1038/s42254-021-00314-5.

Kattel, G. R., Elkadi, H., and Meikle, H. Developing a complementary framework for urban ecology. *Urban Forestry & Urban Greening*, 12(4):498–508, 2013. ISSN 1618-8667. doi: https://doi.org/10.1016/j.ufug.2013.07.005. URL https://www.sciencedirect.com/science/article/pii/S1618866713000824.

Kearney, S. P., Coops, N. C., Stenhouse, G. B., and Nelson, T. A. Ecoanthromes of alberta: An example of disturbance-informed ecological regionalization using remote sensing. *Journal of Environmental Management*, 234:297–310, 2019. ISSN 0301-4797. doi: https://doi.org/10.1016/j.jenvman.2018.12.076. URL https://www.sciencedirect.com/science/article/pii/S0301479718314981.

Kingma, D. P. and Welling, M. An introduction to variational autoencoders. *Found. Trends Mach. Learn.*, 12(4):307–392, 2019. ISSN 1935-8237. doi: 10.1561/2200000056. URL https://doi.org/10.1561/2200000056.

Kitchin, R. Big data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 2014. doi: 10.1177/2053951714528481. URL https://journals.sagepub.com/doi/abs/10.1177/2053951714528481.

Kong, F., Yin, H., Nakagoshi, N., and Zong, Y. Urban green space network development for biodiversity conservation: Identification based on graph theory and gravity modeling. *Landscape and Urban Planning*, 95(1):16–27, 2010. ISSN 0169-2046. doi: https://doi.org/10.1016/j.landurbplan.2009.11.001. URL https://www.sciencedirect.com/science/article/pii/S0169204609002333.

Kopczewska, K. Spatial machine learning: new opportunities for regional science. *The Annals of Regional Science*, 68(3):713–755, 2022. ISSN 1432-0592. doi: 10.1007/s00168-021-01101-x. URL https://doi.org/10.1007/s00168-021-01101-x.

Le, Y. and Huang, S.-Y. Prediction of urban trees planting base on guided cellular automata to enhance the connection of green infrastructure. *Land*, 12(8):1479, 2023. ISSN 2073-445X. URL `https://www.mdpi.com/2073-445X/12/8/1479`.

LeCun, Y., Bengio, Y., and Hinton, G. Deep learning. *Nature*, 521(7553):436–444, 2015. ISSN 1476-4687. doi: 10.1038/nature14539. URL `https://doi.org/10.1038/nature14539`.

Lee, A. and Maheswaran, R. The health benefits of urban green spaces: a review of the evidence. *Journal of Public Health*, 33(2):212–222, 2010. ISSN 1741-3842. doi: 10.1093/pubmed/fdq068. URL `https://doi.org/10.1093/pubmed/fdq068`.

Leibold, M. A. and Chase, J. M. *Metacommunity Ecology, Volume 59*, book section The Theories of Metacommunities, pages 23–48. Princeton University Press, Princeton, 2017. ISBN 9781400889068. doi: doi:10.1515/9781400889068-003. URL `https://doi.org/10.1515/9781400889068-003`.

Lek, S. and Guégan, J. F. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecological Modelling*, 120(2):65–73, 1999. ISSN 0304-3800. doi: https://doi.org/10.1016/S0304-3800(99)00092-7. URL `https://www.sciencedirect.com/science/article/pii/S0304380099000927`.

Levin, S. A. Ecosystems and the biosphere as complex adaptive systems. *Ecosystems*, 1(5):431–436, 1998. ISSN 1432-9840. doi: 10.1007/s100219900037. URL `https://doi.org/10.1007/s100219900037`.

Li, E., Parker, S. S., Pauly, G. B., Randall, J. M., Brown, B. V., and Cohen, B. S. An urban biodiversity assessment framework that combines an urban habitat classification scheme and citizen science data. *Frontiers in Ecology and Evolution*, 7, 2019. ISSN 2296-701X. doi: 10.3389/fevo.2019.00277. URL `https://www.frontiersin.org/articles/10.3389/fevo.2019.00277`.

Li, H. and Reynolds, J. F. On definition and quantification of heterogeneity. *Oikos*, 73(2):280–284, 1995. ISSN 00301299, 16000706. doi: 10.2307/3545921. URL `http://www.jstor.org/stable/3545921`.

Li, H. and Wu, J. *Key Topics in Landscape Ecology*, book section Landscape pattern analysis: key issues and challenges, pages 39–61. Cambridge Studies in Landscape Ecology. Cambridge University Press, Cambridge, 2007. ISBN 9780521616447. doi: DOI:10.1017/CBO9780511618581.004. URL `https://www.cambridge.org/core/product/31CE02E5706749615EC5874898926558`.

Li, Y., Ren, C., Ho, J. Y.-e., and Shi, Y. Landscape metrics in assessing how the configuration of urban green spaces affects their cooling effect: A systematic review of empirical studies. *Landscape and Urban Planning*, 239:104842, 2023. ISSN 0169-2046. doi: https://doi.org/10.1016/j.landurbplan.2023.104842. URL `https://www.sciencedirect.com/science/article/pii/S0169204623001615`.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., and Moher, D. The prisma statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLOS Medicine*, 6(7):e1000100, 2009. doi: 10.1371/journal.pmed.1000100. URL `https://doi.org/10.1371/journal.pmed.1000100`.

Liu, P. and Biljecki, F. A review of spatially-explicit geoai applications in urban geography. *International Journal of Applied Earth Observation and Geoinformation*, 112:102936, 2022. ISSN 1569-8432. doi: https://doi.org/10.1016/j.jag.2022.102936. URL `https://www.sciencedirect.com/science/article/pii/S1569843222001339`.

Liu, Z., Yin, H., Wang, Y., Cheng, Q., and Wang, Z. Research progress on animal habitat constructions from the perspective of urban biodiversity improvement. *Frontiers in Environmental Science*, 11, 2024. ISSN 2296-665X. doi: 10.3389/fenvs.2023.1133879. URL `https://www.frontiersin.org/articles/10.3389/fenvs.2023.1133879`.

Lu, D. and Yang, S. A survey of the analysis of complex systems based on complex network theory and deep learning. *Int J Performability Eng*, 18(4):241–250, 2022. doi: 10.23940/ijpe.22.04.p2.241250. URL `https://www.ijpe-online.com/CN/abstract/article_4673.shtml`.

Lundholm, J. The ecology and evolution of constructed ecosystems as green infrastructure. *Frontiers in Ecology and Evolution*, 3, 2015. ISSN 2296-701X. doi: 10.3389/fevo.2015.00106. URL `https://www.frontiersin.org/articles/10.3389/fevo.2015.00106`.

Lundholm, J., MacIvor, J. S., MacDougall, Z., and Ranalli, M. Plant species and functional group combinations affect green roof ecosystem functions. *PLOS one*, 5(3), 2010.

Luo, Y., Zhu, Z., Wu, J., Zhang, Y., Li, X., Zhao, W., Yuan, Y., Duanmu, Z., and Li, M. Exploring habitat patch clusters based on network community detection to identify restored priority areas of ecological networks in urban areas. *Urban Forestry & Urban Greening*, 78:127771, 2022. ISSN 1618-8667. doi: https://doi.org/10.1016/j.ufug.2022.127771. URL https://www.sciencedirect.com/science/article/pii/S1618866722003144.

Löfvenhaft, K., Björn, C., and Ihse, M. Biotope patterns in urban areas: a conceptual model integrating biodiversity issues in spatial planning. *Landscape and Urban Planning*, 58(2):223–240, 2002. ISSN 0169-2046. doi: https://doi.org/10.1016/S0169-2046(01)00223-7. URL https://www.sciencedirect.com/science/article/pii/S0169204601002237.

Löfvenhaft, K., Runborg, S., and Sjögren-Gulve, P. Biotope patterns and amphibian distribution as assessment tools in urban landscape planning. *Landscape and Urban Planning*, 68(4):403–427, 2004. ISSN 0169-2046. doi: https://doi.org/10.1016/S0169-2046(03)00154-3. URL https://www.sciencedirect.com/science/article/pii/S0169204603001543.

Marcus, L. Plot syntax - a configurational approach to urban diversity. Fifth International Space Syntax Symposium, 2005.

Marcus, L., Berghauser Pont, M., and Barthel, S. Towards a socio-ecological spatial morphology: integrating elements of urban morphology and landscape ecology. *Urban Morphology*, 23(2):115–124, 2019a. doi: 10.51347/jum.v23i2.4084. URL https://journal.urbanform.org/index.php/jum/article/view/4084.

Marcus, L., Berghauser Pont, M., and Barthel, S. Towards a socio-ecological spatial morphology: a joint network approach to urban form and landscape ecology. *Urban Morphology*, 24(1):21–34, 2019b. doi: 10.51347/jum.v24i1.4088. URL https://journal.urbanform.org/index.php/jum/article/view/4088.

Martí, P., Serrano-Estrada, L., and Nolasco-Cirugeda, A. Social media data: Challenges, opportunities and limitations in urban studies. *Computers, Environment and Urban Systems*, 74:161–174, 2019. ISSN 0198-9715. doi: https://doi.org/10.1016/j.compenvurbsys.2018.11.001. URL https://www.sciencedirect.com/science/article/pii/S0198971518302333.

Mathieu, R., Aryal, J., and Chong, A. K. Object-based classification of ikonos imagery for mapping large-scale vegetation communities in urban areas. *Sensors*, 7(11):2860–2880, 2007. ISSN 1424-8220. URL https://www.mdpi.com/1424-8220/7/11/2860.

Matthies, S. A., Rüter, S., Schaarschmidt, F., and Prasse, R. Determinants of species richness within and across taxonomic groups in urban green spaces. *Urban Ecosystems*, 20(4):897–909, 2017. ISSN 1573-1642. doi: 10.1007/s11252-017-0642-9. URL https://doi.org/10.1007/s11252-017-0642-9.

McDonnell, M. J. and Hahs, A. K. Adaptation and adaptedness of organisms to urban environments. *Annual Review of Ecology, Evolution, and Systematics*, 46(1):261–280, 2015. doi: 10.1146/annurev-ecolsys-112414-054258. URL https://www.annualreviews.org/doi/abs/10.1146/annurev-ecolsys-112414-054258.

McGarigal, K. and Marks, B. J. Fragstats: spatial pattern analysis program for quantifying landscape structure. Report, U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, 1995. URL http://dx.doi.org/10.2737/PNW-GTR-351.

McKinney, M. L. Urbanization, biodiversity, and conservation: The impacts of urbanization on native species are poorly studied, but educating a highly urbanized human population about these impacts can greatly improve species conservation in all ecosystems. *BioScience*, 52(10):883–890, 2002. ISSN 0006-3568. doi: 10.1641/0006-3568(2002)052[0883:Ubac]2.0.Co;2. URL https://doi.org/10.1641/0006-3568(2002)052[0883:UBAC]2.0.CO;2.

McKinney, M. L. Urbanization as a major cause of biotic homogenization. *Biological Conservation*, 127(3):247–260, 2006. ISSN 0006-3207. doi: https://doi.org/10.1016/j.biocon.2005.09.005. URL https://www.sciencedirect.com/science/article/pii/S0006320705003563.

McPhearson, T., Andersson, E., Elmqvist, T., and Frantzeskaki, N. Resilience of and through urban ecosystem services. *Ecosystem Services*, 12:152–156, 2015. ISSN 2212-0416. doi: https://doi.org/10.1016/j.ecoser.2014.07.012. URL https://www.sciencedirect.com/science/article/pii/S2212041614000837.

McPhearson, T., Pickett, S. T. A., Grimm, N. B., Niemelä, J., Alberti, M., Elmqvist, T., Weber, C., Haase, D., Breuste, J., and Qureshi, S. Advancing urban ecology toward a science of cities. *BioScience*, 66(3):198–212, 2016. ISSN 0006-3568. doi: 10.1093/biosci/biw002. URL `https://doi.org/10.1093/biosci/biw002`.

Middel, A., Häb, K., Brazel, A. J., Martin, C. A., and Guhathakurta, S. Impact of urban form and design on mid-afternoon microclimate in phoenix local climate zones. *Landscape and Urban Planning*, 122:16–28, 2014. ISSN 0169-2046. doi: https://doi.org/10.1016/j.landurbplan.2013.11.004. URL `https://www.sciencedirect.com/science/article/pii/S0169204613002120`.

Mitchell, T. M. Machine learning and data mining. *Commun. ACM*, 42(11):30–36, 1999. ISSN 0001-0782. doi: 10.1145/319382.319388. URL `https://doi.org/10.1145/319382.319388`.

Morin, E., Herrault, P.-A., Guinard, Y., Grandjean, F., and Bech, N. The promising combination of a remote sensing approach and landscape connectivity modelling at a fine scale in urban planning. *Ecological Indicators*, 139:108930, 2022. ISSN 1470-160X. doi: https://doi.org/10.1016/j.ecolind.2022.108930. URL `https://www.sciencedirect.com/science/article/pii/S1470160X22004010`.

Moscovitz, O. and Barath, S. *A generative design approach to urban sustainability rating systems during early-stage planning*. 2022. doi: 10.52842/conf.caadria.2022.1.171.

Müller, F. Indicating ecosystem and landscape organisation. *Ecological Indicators*, 5(4):280–294, 2005. ISSN 1470-160X. doi: https://doi.org/10.1016/j.ecolind.2005.03.017. URL `https://www.sciencedirect.com/science/article/pii/S1470160X05000257`.

Nassauer, J. I. and Opdam, P. Design in science: extending the landscape ecology paradigm. *Landscape Ecology*, 23(6):633–644, 2008. ISSN 1572-9761. doi: 10.1007/s10980-008-9226-7. URL `https://doi.org/10.1007/s10980-008-9226-7`.

Nikparvar, B. and Thill, J.-C. Machine learning of spatial data. *ISPRS International Journal of Geo-Information*, 10(9):600, 2021. ISSN 2220-9964. URL `https://www.mdpi.com/2220-9964/10/9/600`.

n.n. The ciam charter of athens, 1933: Outcome of a similar effort. *Ekistics*, 16(95):263–267, 1963. ISSN 00132942. URL `http://www.jstor.org/stable/43622757`.

Oberlack, C., Sietz, D., Bürgi Bonanomi, E., de Bremond, A., Dell'Angelo, J., Eisenack, K., Ellis, E. C., Epstein, G., Giger, M., Heinimann, A., Kimmich, C., Kok, M. T. J., Manuel-Navarrete, D., Messerli, P., Meyfroidt, P., Václavík, T., and Villamayor-Tomas, S. Archetype analysis in sustainability research: meanings, motivations, and evidence-based policy making. *Ecology and Society*, 24(2), 2019. doi: 10.5751/ES-10747-240226. URL `https://www.ecologyandsociety.org/vol24/iss2/art26/`.

Oh, K., Lee, D., and Park, C. Urban ecological network planning for sustainable landscape management. *Journal of Urban Technology*, 18(4):39–59, 2011. ISSN 1063-0732. doi: 10.1080/10630732.2011.648433. URL `https://doi.org/10.1080/10630732.2011.648433`. doi: 10.1080/10630732.2011.648433.

Opdam, P. F. M., Verboom, J., and Pouwels, R. Landscape cohesion: An index for the conservation potential of landscapes for biodiversity. *Landscape Ecology 18 (2003) 2*, 18, 2003. doi: 10.1023/A:1024429715253.

Opdam, P., Steingröver, E., and Rooij, S. v. Ecological networks: A spatial concept for multi-actor planning of sustainable landscapes. *Landscape and Urban Planning*, 75(3):322–332, 2006. ISSN 0169-2046. doi: https://doi.org/10.1016/j.landurbplan.2005.02.015. URL `https://www.sciencedirect.com/science/article/pii/S016920460500054X`.

Palazzo, E. Bridging urban morphology and urban ecology: a framework to identify morpho-ecological periods and patterns in the urban ecosystem. *Journal of Urban Ecology*, 8(1):juac007, 2022. ISSN 2058-5543. doi: 10.1093/jue/juac007. URL `https://doi.org/10.1093/jue/juac007`.

Pauleit, S. and Breuste, J. H. *Urban Ecology: Patterns, Processes, and Applications*, book section Land-Use and Surface-Cover as Urban Ecological Indicators. Oxford University Press, 2011. ISBN 9780199563562. doi: 10.1093/acprof:oso/9780199563562.003.0004. URL `https://doi.org/10.1093/acprof:oso/9780199563562.003.0004`.

Pedersen Zari, M. Ecosystem services analysis in response to biodiversity loss caused by the built environment. *S.A.P.I.E.N.S*, 7:1–14, 2015.

Peng, J., Yang, Y., Liu, Y., Hu, Y., Du, Y., Meersmans, J., and Qiu, S. Linking ecosystem services and circuit theory to identify ecological security patterns. *Science of The Total Environment*, 644:781–790, 2018. ISSN 0048-9697. doi: https://doi.org/10.1016/j.scitotenv.2018.06.292. URL `https://www.sciencedirect.com/science/article/pii/S0048969718323684`.

Perini, K., Canepa, M., Barath, S., Hensel, M., Mimet, A., Uthaya Selvan, S., Roccotiello, E., Selami, T., Sunguroglu Hensel, D., Tyc, J., Vogler, V., Weisser, W., and Mosca, F. *ECOLOPES: A multi-species design approach to building envelope design for regenerative urban ecosystems.* 2021.

Peters, D. P. C. and Okin, G. S. A toolkit for ecosystem ecologists in the time of big science. *Ecosystems*, 20(2):259–266, 2017. ISSN 1435-0629. doi: 10.1007/s10021-016-0072-1. URL `https://doi.org/10.1007/s10021-016-0072-1`.

Pichler, M. and Hartig, F. Machine learning and deep learning—a review for ecologists. *Methods in Ecology and Evolution*, 14(4):994–1016, 2023. ISSN 2041-210X. doi: https://doi.org/10.1111/2041-210X.14061. URL `https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.14061`.

Pickett, S. T. A., Cadenasso, M. L., Grove, J. M., Boone, C. G., Groffman, P. M., Irwin, E., Kaushal, S. S., Marshall, V., McGrath, B. P., Nilon, C. H., Pouyat, R. V., Szlavecz, K., Troy, A., and Warren, P. Urban ecological systems: Scientific foundations and a decade of progress. *Journal of Environmental Management*, 92(3):331–362, 2011. ISSN 0301-4797. doi: https://doi.org/10.1016/j.jenvman.2010.08.022. URL `https://www.sciencedirect.com/science/article/pii/S0301479710002707`.

Pickett, S. T. A., Cadenasso, M. L., Rosi-Marshall, E. J., Belt, K. T., Groffman, P. M., Grove, J. M., Irwin, E. G., Kaushal, S. S., LaDeau, S. L., Nilon, C. H., Swan, C. M., and Warren, P. S. Dynamic heterogeneity: a framework to promote ecological integration and hypothesis generation in urban systems. *Urban Ecosystems*, 20(1):1–14, 2017. ISSN 1573-1642. doi: 10.1007/s11252-016-0574-9. URL `https://doi.org/10.1007/s11252-016-0574-9`.

Pickett, S. T. A., Cadenasso, M. L., Childers, D. L., McDonnell, M. J., and Zhou, W. Evolution and future of urban ecological science: ecology in, of, and for the city. *Ecosystem Health and Sustainability*, 2(7):e01229, 2016. doi: 10.1002/ehs2.1229. URL `https://doi.org/10.1002/ehs2.1229`. doi: 10.1002/ehs2.1229.

Post, C., Sarala, R., Gatrell, C., and Prescott, J. E. Advancing theory with review articles. *Journal of Management Studies*, 57(2):26, 2020. ISSN 0022-2380. doi: https://doi.org/10.1111/joms.12549. URL `https://doi.org/10.1111/joms.12549`.

Preiser, R., Biggs, R., De Vos, A., and Folke, C. Social-ecological systems as complex adaptive systems: organizing principles for advancing research methods and approaches. *Ecology and Society*, 23(4), 2018. doi: 10.5751/ES-10558-230446. URL `https://www.ecologyandsociety.org/vol23/iss4/art46/`.

Ramalho, C. E. and Hobbs, R. J. Time for a change: dynamic urban ecology. *Trends in Ecology & Evolution*, 27(3):179–188, 2012. ISSN 0169-5347. doi: https://doi.org/10.1016/j.tree.2011.10.008. URL `https://www.sciencedirect.com/science/article/pii/S0169534711003028`.

Rebele, F. Urban ecology and special features of urban ecosystems. *Global Ecology and Biogeography Letters*, 4(6):173–187, 1994. ISSN 09607447. doi: 10.2307/2997649. URL `http://www.jstor.org/stable/2997649`.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat. Deep learning and process understanding for data-driven earth system science. *Nature*, 566(7743):195–204, 2019. ISSN 1476-4687. doi: 10.1038/s41586-019-0912-1. URL `https://doi.org/10.1038/s41586-019-0912-1`.

Rigó, A. and Barina, Z. Methodology of the habitat classification of anthropogenic urban areas in budapest (hungary). *Biologia Futura*, 71(1):53–68, 2020. ISSN 2676-8607. doi: 10.1007/s42977-020-00011-x. URL `https://doi.org/10.1007/s42977-020-00011-x`.

Riitters, K. H., O'Neill, R. V., and Jones, K. B. Assessing habitat suitability at multiple scales: A landscape-level approach. *Biological Conservation*, 81(1):191–202, 1997. ISSN 0006-3207. doi: https://doi.org/10.1016/S0006-3207(96)00145-0. URL `https://www.sciencedirect.com/science/article/pii/S0006320796001450`.

Ronneberger, O., Fischer, P., and Brox, T. U-net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, pages 234–241. Springer International Publishing, 2015. ISBN 978-3-319-24574-4. doi: 10.1007/978-3-319-24574-4_28.

Rossini, F. and Porter, A. Frameworks for integrating interdisciplinary research. *Research Policy*, 8:70–79, 1979. doi: 10.1016/0048-7333(79)90030-1.

Rubbens, P., Brodie, S., Cordier, T., Destro Barcellos, D., Devos, P., Fernandes-Salvador, J. A., Fincham, J. I., Gomes, A., Handegard, N. O., Howell, K., Jamet, C., Kartveit, K. H., Moustahfid, H., Parcerisas, C., Politikos, D., Sauzède, R., Sokolova, M., Uusitalo, L., Van den Bulcke, L., van Helmond, A. T. M., Watson, J. T., Welch, H., Beltran-Perez, O., Chaffron, S., Greenberg, D. S., Kühn, B., Kiko, R., Lo, M., Lopes, R. M., Möller, K. O., Michaels, W., Pala, A., Romagnan, J.-B., Schuchert, P., Seydi, V., Villasante, S., Malde, K., and Irisson, J.-O. Machine learning in marine ecology: an overview of techniques and applications. *ICES Journal of Marine Science*, 80(7):1829–1853, 2023. ISSN 1054-3139. doi: 10.1093/icesjms/fsad100. URL `https://doi.org/10.1093/icesjms/fsad100`.

Rudofsky, B. *Architecture without architects, an introduction to nonpedigreed architecture*. The Museum of Modern Art, New York, 1964.

Samuel, A. L. Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3):210–229, 1959. ISSN 0018-8646. doi: 10.1147/rd.33.0210.

Scolozzi, R. and Geneletti, D. Spatial rule-based assessment of habitat potential to predict impact of land use changes on biodiversity at municipal scale. *Environmental Management*, 47(3):368–383, 2011. ISSN 1432-1009. doi: 10.1007/s00267-011-9613-8. URL `https://doi.org/10.1007/s00267-011-9613-8`.

Scowen, M., Athanasiadis, I. N., Bullock, J. M., Eigenbrod, F., and Willcock, S. The current and future uses of machine learning in ecosystem service research. *Science of The Total Environment*, 799:149263, 2021. ISSN 0048-9697. doi: https://doi.org/10.1016/j.scitotenv.2021.149263. URL `https://www.sciencedirect.com/science/article/pii/S0048969721043369`.

Selvan, S. U., Saroglou, S. T., Mosca, F., Tyc, J., Joschinski, J., Calbi, M., Vogler, V., Weisser, W. W., Grobman, Y. J., and Barath, S. Multi-species building envelopes - developing a multi-criteria design decision-making methodology for cohabitation. volume 2 of *Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia - CAADRIA*, pages 643–652, Hong Kong, 2023a. Association for Comptuer-Aided Architectural Design Research in Asia.

Selvan, S. U., Saroglou, S. T., Joschinski, J., Calbi, M., Vogler, V., Barath, S., and Grobman, Y. J. Toward multi-species building envelopes: A critical literature review of multi-criteria decision-making for design support. *Building and Environment*, 231:110006, 2023b. ISSN 0360-1323. doi: https://doi.org/10.1016/j.buildenv.2023.110006. URL `https://www.sciencedirect.com/science/article/pii/S0360132323000331`.

Snyder, H. Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104: 333–339, 2019. ISSN 0148-2963. doi: https://doi.org/10.1016/j.jbusres.2019.07.039. URL `https://www.sciencedirect.com/science/article/pii/S0148296319304564`.

Spotswood, E. N., Beller, E. E., Grossinger, R., Grenier, J. L., Heller, N. E., and Aronson, M. F. J. The biological deserts fallacy: Cities in their landscapes contribute more than we think to regional biodiversity. *BioScience*, 71(2):148–160, 2021. ISSN 0006-3568. doi: 10.1093/biosci/biaa155. URL `https://doi.org/10.1093/biosci/biaa155`.

Star, S. L. and Griesemer, J. R. Institutional ecology, ‘translations’ and boundary objects: Amateurs and professionals in berkeley’s museum of vertebrate zoology, 1907-39. *Social Studies of Science*, 19(3):387–420, 1989. doi: 10.1177/030631289019003001. URL `https://journals.sagepub.com/doi/abs/10.1177/030631289019003001`.

Stewart, I. D. and Oke, T. R. Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society*, 93(12):1879–1900, 2012. ISSN 00030007, 15200477. URL `http://www.jstor.org/stable/26219368`.

Stupariu, M.-S., Cushman, S. A., Pleşoianu, A.-I., Pătru-Stupariu, I., and Fürst, C. Machine learning in landscape ecological analysis: a review of recent approaches. *Landscape Ecology*, 37(5):1227–1250, 2022. ISSN 1572-9761. doi: 10.1007/s10980-021-01366-9. URL `https://doi.org/10.1007/s10980-021-01366-9`.

Tansley, A. G. The use and abuse of vegetational concepts and terms. *Ecology*, 16(3):284–307, 1935. ISSN 0012-9658. doi: https://doi.org/10.2307/1930070. URL https://esajournals.onlinelibrary.wiley.com/doi/abs/10.2307/1930070.

Tews, J., Brose, U., Grimm, V., Tielbörger, K., Wichmann, M. C., Schwager, M., and Jeltsch, F. Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. *Journal of Biogeography*, 31(1):79–92, 2004. ISSN 0305-0270. doi: https://doi.org/10.1046/j.0305-0270.2003.00994.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1046/j.0305-0270.2003.00994.x.

Timonen, V., Foley, G., and Conlon, C. Challenges when using grounded theory:a pragmatic introduction to doing gt research. *International Journal of Qualitative Methods*, 17(1):1609406918758086, 2018. doi: 10.1177/1609406918758086. URL https://journals.sagepub.com/doi/abs/10.1177/1609406918758086.

Tobler, W. R. A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46:234–240, 1970. ISSN 00130095, 19448287. doi: 10.2307/143141. URL http://www.jstor.org/stable/143141.

Torraco, R. J. Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, 4 (3):356–367, 2005. ISSN 15344843. Copyright - Copyright SAGE PUBLICATIONS, INC. Sep 2005.

Tufail, S., Riggs, H., Tariq, M., and Sarwat, A. I. Advancements and challenges in machine learning: A comprehensive review of models, libraries, applications, and algorithms. *Electronics*, 12(8), 2023. doi: 10.3390/electronics12081789.

Tulchinsky, T. H. *Case Studies in Public Health*, book section John Snow, Cholera, the Broad Street Pump; Waterborne Diseases Then and Now, pages 77–99. Academic Press, 2018. ISBN 978-0-12-804571-8. doi: https://doi.org/10.1016/B978-0-12-804571-8.00017-2. URL https://www.sciencedirect.com/science/article/pii/B9780128045718000172.

Turner, M. G. Landscape ecology: The effect of pattern on process. *Annual Review of Ecology, Evolution, and Systematics*, 20(Volume 20, 1989):171–197, 1989. ISSN 1545-2069. doi: https://doi.org/10.1146/annurev.es.20.110189.001131. URL https://www.annualreviews.org/content/journals/10.1146/annurev.es.20.110189.001131.

Tyc, J., Selami, T., Hensel, D. S., and Hensel, M. A scoping review of voxel-model applications to enable multi-domain data integration in architectural design and urban planning. *Architecture*, 3(2):137–174, 2023. doi: 10.3390/architecture3020010.

Ullah, F. A beginner's guide to developing review-based conceptual frameworks in the built environment. *Architecture*, 1 (1):5–24, 2021. doi: 10.3390/architecture1010003.

United Nations. Report of the world commission on environment and development "our common future". Report, 1987.

Varpio, L., Paradis, E., Uijtdehaage, S., and Young, M. The distinctions between theory, theoretical framework, and conceptual framework. *Academic Medicine*, 95(7), 2020. ISSN 1040-2446. URL https://journals.lww.com/academicmedicine/fulltext/2020/07000/the_distinctions_between_theory,_theoretical.21.aspx.

Verboom, J. and Pouwels, R. *Ecological networks and greenways; concept, design, implementation*, book section Ecological functioning of ecological networks: a species perspective. Cambridge University Press, Cambridge, UK, 2004.

Vogler, V., Mimet, A., Mognieh, A., Selvan, S. U., Mosca, F., Selami, T., Barath, S., Weisser, W., and Fraguada, L. Ecolopes d3.1: Prototype technical requirements. Report, 2022.

Vujovic, M., Stojanovic, D., Selami, T., and Hensel, M. Design and science: Content analysis of published peer-reviewed research over the last four decades. *Frontiers of Architectural Research*, 12(4):613–629, 2023. ISSN 2095-2635. doi: https://doi.org/10.1016/j.foar.2023.04.001. URL https://www.sciencedirect.com/science/article/pii/S2095263523000249.

Václavík, T., Lautenbach, S., Kuemmerle, T., and Seppelt, R. Mapping global land system archetypes. *Global Environmental Change*, 23(6):1637–1647, 2013. ISSN 0959-3780. doi: https://doi.org/10.1016/j.gloenvcha.2013.09.004. URL https://www.sciencedirect.com/science/article/pii/S0959378013001532.

Wang, C., Guo, J., and Liu, J. Green roofs and their effect on architectural design and urban ecology using deep learning approaches. *Soft Computing*, 28(4):3667–3682, 2024. ISSN 1433-7479. doi: 10.1007/s00500-024-09637-8. URL https://doi.org/10.1007/s00500-024-09637-8.

Wang, X., Gu, W., Ziebelin, D., and Hamilton, H. An ontology-based framework for geospatial clustering. *International Journal of Geographical Information Science*, 24(11):1601–1630, 2010. ISSN 1365-8816. doi: 10.1080/13658811003702147. URL https://doi.org/10.1080/13658811003702147. doi: 10.1080/13658811003702147.

Wang, Y., Sun, Y., Liu, Z., Sarma, S. E., Bronstein, M. M., and Solomon, J. M. Dynamic graph cnn for learning on point clouds. *ACM Trans. Graph.*, 38(5):Article 146, 2019. ISSN 0730-0301. doi: 10.1145/3326362. URL https://doi.org/10.1145/3326362.

Weinstock, M. The metabolism of the city: The mathematics of networks and urban surfaces. *Architectural Design*, 81(4): 102–107, 2011. ISSN 0003-8504. doi: https://doi.org/10.1002/ad.1275. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/ad.1275.

Weisser, W. W., Hensel, M., Barath, S., Culshaw, V., Grobman, Y. J., Hauck, T. E., Joschinski, J., Ludwig, F., Mimet, A., Perini, K., Roccotiello, E., Schloter, M., Shwartz, A., Hensel, D. S., and Vogler, V. Creating ecologically sound buildings by integrating ecology, architecture and computational design. *People and Nature*, 5(1):4–20, 2023. doi: https://doi.org/10.1002/pan3.10411. URL https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1002/pan3.10411.

Wentz, E. A., York, A. M., Alberti, M., Conrow, L., Fischer, H., Inostroza, L., Jantz, C., Pickett, S. T. A., Seto, K. C., and Taubenböck, H. Six fundamental aspects for conceptualizing multidimensional urban form: A spatial mapping perspective. *Landscape and Urban Planning*, 179:55–62, 2018. ISSN 0169-2046. doi: https://doi.org/10.1016/j.landurbplan.2018.07.007. URL https://www.sciencedirect.com/science/article/pii/S0169204618306455.

Werner, P. The ecology of urban areas and their functions for species diversity. *Landscape and Ecological Engineering*, 7(2): 231–240, 2011. ISSN 1860-188X. doi: 10.1007/s11355-011-0153-4. URL https://doi.org/10.1007/s11355-011-0153-4.

Whittemore, R. and Knafl, K. The integrative review: updated methodology. *Journal of Advanced Nursing*, 52(5):546–553, 2005. ISSN 0309-2402. doi: https://doi.org/10.1111/j.1365-2648.2005.03621.x. URL https://doi.org/10.1111/j.1365-2648.2005.03621.x.

Whittemore, R., Chao, A., Jang, M., Minges, K. E., and Park, C. Methods for knowledge synthesis: An overview. *Heart & Lung*, 43(5):453–461, 2014. ISSN 0147-9563. doi: https://doi.org/10.1016/j.hrtlng.2014.05.014. URL https://www.sciencedirect.com/science/article/pii/S0147956314001897.

Wieland, R., Wulf, M., and Meier, K. What makes the difference between memory and face of a landscape? a machine learning approach applied to the federal state brandenburg, germany. *Spatial Information Research*, 27(2):237–246, 2019. ISSN 2366-3294. doi: 10.1007/s41324-018-0228-5. URL https://doi.org/10.1007/s41324-018-0228-5.

Williams, N. S., Schwartz, M. W., Vesk, P. A., McCarthy, M. A., Hahs, A. K., Clemants, S. E., Corlett, R. T., Duncan, R. P., Norton, B. A., Thompson, K., and McDonnell, M. J. A conceptual framework for predicting the effects of urban environments on floras. *Journal of Ecology*, 97(1):4–9, 2009. ISSN 0022-0477. doi: https://doi.org/10.1111/j.1365-2745.2008.01460.x. URL https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2745.2008.01460.x.

Wolfswinkel, J. F., Furtmueller, E., and Wilderom, C. P. M. Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, 22(1):45–55, 2013. ISSN 0960-085X. doi: 10.1057/ejis.2011.51. URL https://doi.org/10.1057/ejis.2011.51. doi: 10.1057/ejis.2011.51.

Wu, J. Hierarchy and scaling: Extrapolating information along a scaling ladder. *Canadian Journal of Remote Sensing*, 25 (4):367–380, 1999. ISSN 0703-8992. doi: 10.1080/07038992.1999.10874736. URL https://doi.org/10.1080/07038992.1999.10874736. doi: 10.1080/07038992.1999.10874736.

Wu, J. Urban ecology and sustainability: The state-of-the-science and future directions. *Landscape and Urban Planning*, 125:209–221, 2014. ISSN 0169-2046. doi: https://doi.org/10.1016/j.landurbplan.2014.01.018. URL https://www.sciencedirect.com/science/article/pii/S0169204614000322.

Wu, J. and David, J. L. A spatially explicit hierarchical approach to modeling complex ecological systems: theory and applications. *Ecological Modelling*, 153(1):7–26, 2002. ISSN 0304-3800. doi: https://doi.org/10.1016/S0304-3800(01)00499-9. URL https://www.sciencedirect.com/science/article/pii/S0304380001004999.

Wu, J. and Loucks, O. L. From balance of nature to hierarchical patch dynamics: A paradigm shift in ecology. *The Quarterly Review of Biology*, 70(4):439–466, 1995. ISSN 00335770, 15397718. URL http://www.jstor.org/stable/3035824.

Wu, J. and Qi, Y. Dealing with scale in landscape analysis: An overview. *Geographic Information Sciences*, 6(1):1–5, 2000. ISSN 1082-4006. doi: 10.1080/10824000009480528. URL https://doi.org/10.1080/10824000009480528. doi: 10.1080/10824000009480528.

Wu, Z., Lu, Q., Lei, S., and Yan, Q. Study on landscape ecological classification and landscape types evolution: A case study of a mining city in semi-arid steppe. *Sustainability*, 13(17):9541, 2021. ISSN 2071-1050. URL https://www.mdpi.com/2071-1050/13/17/9541.

Xing, J. and Sieber, R. Propagation of uncertainty for volunteered geographic information in machine learning. 10th International Conference on Geographic Information Science, 2018. doi: 10.4230/LIPIcs.GIScience.2018.93.

Yan, M., Huang, C., Bienstman, P., Tino, P., Lin, W., and Sun, J. Emerging opportunities and challenges for the future of reservoir computing. *Nature Communications*, 15(1):2056, 2024. ISSN 2041-1723. doi: 10.1038/s41467-024-45187-1. URL https://doi.org/10.1038/s41467-024-45187-1.

Yoffe, H., Plaut, P., Fried, S., and Grobman, Y. J. Enriching the parametric vocabulary of urban landscapes. Anthropologic: Architecture and Fabrication in the cognitive age - Proceedings of the 38th eCAADe Conference, pages 47–56, 2020.

Yoffe, H., Raanan, N., Fried, S., Plaut, P., and Grobman, Y. J. Sustainable urban landscapes: a computation framework for enhancing sustainability in early-stage design. *Archnet-IJAR: International Journal of Architectural Research*, ahead-of-print(ahead-of-print), 2023. ISSN 2631-6862, 1938-7806. doi: 10.1108/ARCH-06-2023-0152. URL https://doi.org/10.1108/ARCH-06-2023-0152.

Young, C., Jarvis, P., Hooper, I., and Trueman, I. C. *Landscape Ecology Research Trends*, book section Urban landscape ecology and its evaluation : a review, pages 45–69. New York : Nova Science, 2009 2009. ISBN 9781604566727. URL http://hdl.handle.net/2436/92006.

Zhu, D., Liu, Y., Yao, X., and Fischer, M. M. Spatial regression graph convolutional neural networks: A deep learning paradigm for spatial multivariate distributions. *GeoInformatica*, 26(4):645–676, 2022. ISSN 1573-7624. doi: 10.1007/s10707-021-00454-x. URL https://doi.org/10.1007/s10707-021-00454-x.

Zipperer, W. C., Wu, J., Pouyat, R. V., and Pickett, S. T. A. The application of ecological principles to urban and urbanizing landscapes. *Ecological Applications*, 10(3):685–688, 2000. ISSN 1051-0761. doi: https://doi.org/10.1890/1051-0761(2000)010[0685:TAOEPT]2.0.CO;2. URL https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/1051-0761%282000%29010%5B0685%3ATAOEPT%5D2.0.CO%3B2.

Zwakman, M., Verberne, L. M., Kars, M. C., Hooft, L., van Delden, J. J. M., and Spijker, R. Introducing palette: an iterative method for conducting a literature search for a review in palliative care. *BMC Palliative Care*, 17(1):82, 2018. ISSN 1472-684X. doi: 10.1186/s12904-018-0335-z. URL https://doi.org/10.1186/s12904-018-0335-z.

Review Articles

Bai, H., Li, Z., Guo, H., Chen, H., and Luo, P. Urban green space planning based on remote sensing and geographic information systems. *Remote Sensing*, 14(17):4213, 2022. ISSN 2072-4292. URL `https://www.mdpi.com/2072-4292/14/17/4213`.

Ban, N., Liu, X., and Zhu, C. Landscape ecological construction and spatial pattern optimization design based on genetic algorithm optimization neural network. *Wireless Communications and Mobile Computing*, 2022:4976361, 2022. ISSN 1530-8669. doi: 10.1155/2022/4976361. URL `https://doi.org/10.1155/2022/4976361`.

Barbierato, E., Bernetti, I., Capecchi, I., and Saragosa, C. Integrating remote sensing and street view images to quantify urban forest ecosystem services. *Remote Sensing*, 12(2), 2020. ISSN 2072-4292. URL `https://www.mdpi.com/2072-4292/12/2/329`.

Bergerot, B., Fontaine, B., Julliard, R., and Baguette, M. Landscape variables impact the structure and composition of butterfly assemblages along an urbanization gradient. *Landscape Ecology*, 26(1):83–94, 2011. ISSN 1572-9761. doi: 10.1007/s10980-010-9537-3. URL `https://doi.org/10.1007/s10980-010-9537-3`.

Bonilla-Bedoya, S., López-Ulloa, M., Mora-Garcés, A., Macedo-Pezzopane, J. E., Salazar, L., and Herrera, M. Urban soils as a spatial indicator of quality for urban socio-ecological systems. *Journal of Environmental Management*, 300:113556, 2021. ISSN 0301-4797. doi: https://doi.org/10.1016/j.jenvman.2021.113556. URL `https://www.sciencedirect.com/science/article/pii/S0301479721016182`.

Chapman, D. S., Gunn, I. D. M., Pringle, H. E. K., Siriwardena, G. M., Taylor, P., Thackeray, S. J., Willby, N. J., and Carvalho, L. Invasion of freshwater ecosystems is promoted by network connectivity to hotspots of human activity. *Global Ecology and Biogeography*, 29(4):645–655, 2020. ISSN 1466-822X. doi: https://doi.org/10.1111/geb.13051. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/geb.13051`.

Chen, J., Zhang, J., Xu, X., Fu, C., Zhang, D., Zhang, Q., and Xuan, Q. E-lstm-d: A deep learning framework for dynamic network link prediction. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51(6):3699–3712, 2021. ISSN 2168-2232. doi: 10.1109/TSMC.2019.2932913.

Chen, S., Zhang, M., and Lei, F. Mapping vegetation types by different fully convolutional neural network structures with inadequate training labels in complex landscape urban areas. *Forests*, 14(9):1788, 2023a. ISSN 1999-4907. URL `https://www.mdpi.com/1999-4907/14/9/1788`.

Chen, Y., Wang, W., and Chen, X. M. Bibliometric methods in traffic flow prediction based on artificial intelligence. *Expert Systems with Applications*, 228:120421, 2023b. ISSN 0957-4174. doi: https://doi.org/10.1016/j.eswa.2023.120421. URL `https://www.sciencedirect.com/science/article/pii/S0957417423009235`.

Dimopoulos, I., Chronopoulos, J., Chronopoulou-Sereli, A., and Lek, S. Neural network models to study relationships between lead concentration in grasses and permanent urban descriptors in athens city (greece). *Ecological Modelling*, 120(2):157–165, 1999. ISSN 0304-3800. doi: https://doi.org/10.1016/S0304-3800(99)00099-X. URL `https://www.sciencedirect.com/science/article/pii/S030438009900099X`.

Dong, J., Guo, F., Lin, M., Zhang, H., and Zhu, P. Optimization of green infrastructure networks based on potential green roof integration in a high-density urban area—a case study of beijing, china. *Science of The Total Environment*, 834:155307, 2022. ISSN 0048-9697. doi: https://doi.org/10.1016/j.scitotenv.2022.155307. URL `https://www.sciencedirect.com/science/article/pii/S0048969722024007`.

dos Santos, A. R., Anjinho, P. d. S., Neves, G. L., Barbosa, M. A. G. A., de Assis, L. C., and Mauad, F. F. Dynamics of environmental conservation: Evaluating the past for a sustainable future. *International Journal of Applied Earth Observation and Geoinformation*, 102:102452, 2021. ISSN 1569-8432. doi: https://doi.org/10.1016/j.jag.2021.102452. URL https://www.sciencedirect.com/science/article/pii/S0303243421001598.

Fan, C., Zeng, L., Sun, Y., and Liu, Y.-Y. Finding key players in complex networks through deep reinforcement learning. *Nature Machine Intelligence*, 2(6):317–324, 2020. ISSN 2522-5839. doi: 10.1038/s42256-020-0177-2. URL https://doi.org/10.1038/s42256-020-0177-2.

Fang, Z., Qi, J., Yang, T., Wan, L., and Jin, Y. "reading" cities with computer vision: a new multi-spatial scale urban fabric dataset and a novel convolutional neural network solution for urban fabric classification tasks. Proceedings of the 28th International Conference on Advances in Geographic Information Systems, page 507–517. Association for Computing Machinery, 2020. doi: 10.1145/3397536.3422240. URL https://doi.org/10.1145/3397536.3422240.

Ha, S. and Jeong, H. Unraveling hidden interactions in complex systems with deep learning. *Scientific Reports*, 11(1): 12804, 2021. ISSN 2045-2322. doi: 10.1038/s41598-021-91878-w. URL https://doi.org/10.1038/s41598-021-91878-w.

Hassell, J. M., Bettridge, J. M., Ward, M. J., Ogendo, A., Imboma, T., Muloi, D., Fava, F., Robinson, T. P., Begon, M., and Fèvre, E. M. Socio-ecological drivers of vertebrate biodiversity and human-animal interfaces across an urban landscape. *Global Change Biology*, 27(4):781–792, 2021. ISSN 1354-1013. doi: https://doi.org/10.1111/gcb.15412. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.15412.

He, Q., Ng, J. L., and Noh, N. I. F. M. Plant landscape configuration method of regional characteristic rainwater garden based on deep learning. Multimedia Technology and Enhanced Learning, pages 354–367. Springer Nature Switzerland, 2022. ISBN 978-3-031-18123-8. doi: 10.1007/978-3-031-18123-8_27.

Jiao, H. and Han, J. Urban green space planning and design for sponge city. *Scientific Programming*, 2022, 2022. ISSN 1058-9244. doi: 10.1155/2022/5333231. URL https://doi.org/10.1155/2022/5333231.

Jutras, P., Prasher, S. O., Yang, C.-C., and Hamel, C. Urban tree growth modelling with artificial neural network. volume 2 of *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No.02CH37290)*, pages 1385–1389 vol.2, 2002. ISBN 1098-7576. doi: 10.1109/IJCNN.2002.1007718.

Jutras, P., Prasher, S. O., and Mehuys, G. R. Prediction of street tree morphological parameters using artificial neural networks. *Computers and Electronics in Agriculture*, 67(1):9–17, 2009. ISSN 0168-1699. doi: https://doi.org/10.1016/j.compag.2009.02.008. URL https://www.sciencedirect.com/science/article/pii/S0168169909000441.

Karapinar Senturk, Z. Amphibian species detection in water reservoirs using artificial neural networks for ecology-friendly city planning. *Ecological Informatics*, 69:101640, 2022. ISSN 1574-9541. doi: https://doi.org/10.1016/j.ecoinf.2022.101640. URL https://www.sciencedirect.com/science/article/pii/S1574954122000899.

Labib, S. M. Investigation of the likelihood of green infrastructure (gi) enhancement along linear waterways or on derelict sites (ds) using machine learning. *Environmental Modelling & Software*, 118:146–165, 2019. ISSN 1364-8152. doi: https://doi.org/10.1016/j.envsoft.2019.05.006. URL https://www.sciencedirect.com/science/article/pii/S1364815218305541.

Le, Y. and Huang, S.-Y. Prediction of urban trees planting base on guided cellular automata to enhance the connection of green infrastructure. *Land*, 12(8):1479, 2023. ISSN 2073-445X. URL https://www.mdpi.com/2073-445X/12/8/1479.

Li, S. and Fan, Z. Evaluation of urban green space landscape planning scheme based on pso-bp neural network model. *Alexandria Engineering Journal*, 61(9):7141–7153, 2022. ISSN 1110-0168. doi: https://doi.org/10.1016/j.aej.2021.12.057. URL https://www.sciencedirect.com/science/article/pii/S1110016821008693.

Lin, Y., Yu, J., Cai, J., Sneeuw, N., and Li, F. Spatio-temporal analysis of wetland changes using a kernel extreme learning machine approach. *Remote Sensing*, 10(7):1129, 2018. ISSN 2072-4292. URL https://www.mdpi.com/2072-4292/10/7/1129.

Luhua, L. U. O., Mingjie, C., Lulu, D., Wei, S. U., Xin, L. I., Xiaodong, H. U., Xin, Z., Chen, L. I., Weiming, C., Hanning, S. H. I., and Jiancheng, L. U. O. Intelligent identification of building patches and assessment of roof greening suitability in high-density urban areas: A case study of chengdu. *Journal of Resources and Ecology*, 13(2):247–256, 2022. ISSN 1674-764X. doi: 10.5814/j.issn.1674-764x.2022.02.008. URL https://www.jorae.cn/EN/10.5814/j.issn.1674-764x.2022.02.008.

Morshed, S. R., Fattah, M. A., Haque, M. N., and Morshed, S. Y. Future ecosystem service value modeling with land cover dynamics by using machine learning based artificial neural network model for jashore city, bangladesh. *Physics and Chemistry of the Earth, Parts A/B/C*, 126:103021, 2022. ISSN 1474-7065. doi: https://doi.org/10.1016/j.pce.2021.103021. URL https://www.sciencedirect.com/science/article/pii/S1474706521000541.

Nölke, N., Fehrmann, L., Plieninger, T., and Kleinn, C. Categorization of green and grey infrastructure complexity in the rural–urban interface of bengaluru, india: an unsupervised volumetric approach with relevance for urban quality. *Urban Ecosystems*, 26(1):161–172, 2023. ISSN 1573-1642. doi: 10.1007/s11252-022-01300-0. URL https://doi.org/10.1007/s11252-022-01300-0.

Ringland, J., Bohm, M., Baek, S.-R., and Eichhorn, M. Automated survey of selected common plant species in thai homegardens using google street view imagery and a deep neural network. *Earth Science Informatics*, 14(1):179–191, 2021. ISSN 1865-0481. doi: 10.1007/s12145-020-00557-3. URL https://doi.org/10.1007/s12145-020-00557-3.

Steenberg, J. W., Millward, A. A., Nowak, D. J., Robinson, P. J., and Smith, S. M. A social-ecological analysis of urban tree vulnerability for publicly-owned trees in a residential neighborhood. *Arboriculture and Urban Forestry*, 45:15, 2019.

Sun, Y., Gao, C., Li, J., Gao, M., and Ma, R. Assessing the cooling efficiency of urban parks using data envelopment analysis and remote sensing data. *Theoretical and Applied Climatology*, 145(3):903–916, 2021. ISSN 1434-4483. doi: 10.1007/s00704-021-03665-2. URL https://doi.org/10.1007/s00704-021-03665-2.

Wang, C., Guo, J., and Liu, J. Green roofs and their effect on architectural design and urban ecology using deep learning approaches. *Soft Computing*, 28(4):3667–3682, 2024. ISSN 1433-7479. doi: 10.1007/s00500-024-09637-8. URL https://doi.org/10.1007/s00500-024-09637-8.

Wang, Y., Wang, J., Chang, S., Sun, L., An, L., Chen, Y., and Xu, J. Classification of street tree species using uav tilt photogrammetry. *Remote Sensing*, 13(2):216, 2021. ISSN 2072-4292. URL https://www.mdpi.com/2072-4292/13/2/216.

Wellmann, T., Lausch, A., Scheuer, S., and Haase, D. Earth observation based indication for avian species distribution models using the spectral trait concept and machine learning in an urban setting. *Ecological Indicators*, 111:106029, 2020. ISSN 1470-160X. doi: https://doi.org/10.1016/j.ecolind.2019.106029. URL https://www.sciencedirect.com/science/article/pii/S1470160X19310258.

Wieland, R., Wulf, M., and Meier, K. What makes the difference between memory and face of a landscape? a machine learning approach applied to the federal state brandenburg, germany. *Spatial Information Research*, 27(2):237–246, 2019. ISSN 2366-3294. doi: 10.1007/s41324-018-0228-5. URL https://doi.org/10.1007/s41324-018-0228-5.

Wiese, D., Escalante, A. A., Murphy, H., Henry, K. A., and Gutierrez-Velez, V. H. Integrating environmental and neighborhood factors in maxent modeling to predict species distributions: A case study of aedes albopictus in southeastern pennsylvania. *PLOS ONE*, 14(10):e0223821, 2019. doi: 10.1371/journal.pone.0223821. URL https://doi.org/10.1371/journal.pone.0223821.

Zaleckis, K., Kamičaitytė, J., Mlinkauskienė, A., and Jankauskaitė-Jurevičienė, L. Modeling of changes in four urban capitals using up-to-date information systems and mathematical graph-based simulative models for urban regeneration (kaunas case). *Sustainability*, 14(24):17014, 2022. ISSN 2071-1050. URL https://www.mdpi.com/2071-1050/14/24/17014.

Zhang, L., Wang, L., Wu, J., Li, P., Dong, J., and Wang, T. Decoding urban green spaces: Deep learning and google street view measure greening structures. *Urban Forestry & Urban Greening*, 87:128028, 2023. ISSN 1618-8667. doi: https://doi.org/10.1016/j.ufug.2023.128028. URL https://www.sciencedirect.com/science/article/pii/S1618866723001991.

Zubair, S., Jain, S. K., and Somvanshi, S. Comparative analysis and prediction of ecological quality of delhi. Advances in Energy and Environment, pages 163–177. Springer Singapore, 2021. ISBN 978-981-33-6695-4. doi: 10.1007/978-981-33-6695-4_16.

Łopucki, R. and Kiersztyn, A. The city changes the daily activity of urban adapters: Camera-traps study of apodemus agrarius behaviour and new approaches to data analysis. *Ecological Indicators*, 110:105957, 2020. ISSN 1470-160X. doi: https://doi.org/10.1016/j.ecolind.2019.105957. URL https://www.sciencedirect.com/science/article/pii/S1470160X19309525.

Web Resources

Clarivate. Web of science, s.a.a. URL https://www.clarivate.com. last accessed: 2024-09-03.

Clarivate. Web of science research areas, s.a.b. URL https://incites.help.clarivate.com/Content/Research-Areas/wos-research-areas.htm. last accessed: 2024-09-03.

CORE. Core uk, s.a. URL https://core.ac.uk/. last accessed: 2024-09-03.

Elsevier. All science journal classification codes (asjc), s.a.a. URL https://service.elsevier.com/app/answers/detail/a_id/15181/supporthub/scopus/. last accessed: 2024-09-03.

Elsevier. Scopus, s.a.b. URL https://www.scopus.com. last accessed: 2024-09-03.

Grames, E. M. litsearchr, s.a. URL https://elizagrames.github.io/litsearchr/index.html.

Hair, K. and Wallrich, L. Asysd, s.a. URL https://camaradesuk.github.io/ASySD/.

Higher Education Statistics Agency. Jacs 3.0: Principal subject codes, s.a. URL https://www.hesa.ac.uk/support/documentation/jacs/jacs3-principal. last accessed: 2024-09-03.

Hugging Face Inc. Hugging face, s.a. URL https://huggingface.co/. last accessed: 2024-09-03.

Leiden University. Vosviewer, s.a. URL https://www.vosviewer.com/. last accessed: 2024-09-03.

Linux Foundation. Pytorch, s.a. URL https://pytorch.org/. last accessed: 2024-09-03.

R Foundation. R, s.a. URL https://www.r-project.org/. last accessed: 2024-09-03.

Springer Nature Limited. Subjects in scientific reports, s.a. URL https://www.nature.com/srep/browse-subjects. last accessed: 2024-09-03.

Szymański, P. scikit-multilearn, s.a. URL http://scikit.ml/. last accessed: 2024-09-03.

United Nations. Sustainable development goals, s.a. URL https://sdgs.un.org/. last accessed: 2024-07-31.

Řehůřek, R. Gensim: topic modelling for humans, s.a. URL https://radimrehurek.com/gensim/.

List of Figures

List of Tables

Appendices

Appendix A - Final Search Query

(ecolog* OR biodiversity OR “conservation biology”) AND (animal OR bacteria OR biomass OR biotic OR biological OR bird OR fauna OR flora OR gene OR grass OR green OR insect OR invertebrate OR microbe OR organic OR organism OR plant OR proxies OR proxy OR species OR taxa OR taxon OR vegetation OR vertebrate) AND ((actor OR agent OR alliance OR assemblage OR association OR band OR combination OR community OR company OR component OR composition OR connection OR configuration OR constellation OR crowd OR element OR fabric OR “food chain” OR gang OR generation OR group OR hierarchy OR individual OR inhabitant OR kinship OR line OR linkage OR material OR matter OR medium OR member OR mix OR mixture OR network OR occupant OR order OR population OR proportion OR relation OR relationship OR remains OR resident OR set OR society OR spectrum OR strain OR trophic OR unit OR vector OR vehicle) OR (annual OR belt OR boundary OR canopy OR coat OR corridor OR cradle OR cover OR cycle OR domain OR ecosystem OR environment OR event OR gradient OR habitat OR home OR layer OR milieu OR neighborhood OR neighbourhood OR net OR niche OR origin OR period OR phenomenon OR regime OR root OR season OR sequence OR shelter OR source OR structure OR system OR station OR temporal OR time OR territory OR web OR zone)) AND ((aim OR attribute OR challenge OR characteristic OR complexity OR complication OR concern OR condition OR constraint OR criteria OR criterion OR directive OR driver OR factor OR feature OR form OR framework OR goal OR indicator OR issue OR kind OR label OR limit OR link OR marker OR measure OR mechanism OR metric OR necessity OR objective OR parameter OR pattern OR perspective OR prerequisite OR problem OR property OR quality OR requirement OR role OR shape OR solution OR stage OR state OR strategy OR target OR task OR terms OR theory OR threat OR trait OR type OR variable) OR (data OR dataset OR foundation OR information OR knowledge OR resource OR supplement)) AND ((ability OR access OR accessibility OR acclimatisation OR acclimatization OR action OR activity OR adaptation OR adjustment OR affectedness OR agency OR alteration OR attention OR augmentation OR awareness OR behavior OR behaviour OR blending OR breeding OR capability OR capacity OR chance OR collaboration OR competition OR conflict OR conservation OR contribution OR control OR conversion OR cooperation OR decision OR demand OR dependency OR difference OR disturbance OR dynamic OR effect OR effort OR fitness OR flexibility OR function OR fusion OR governance OR habit OR imitation OR impact OR implementation OR inclusion OR incorporation OR influence OR integration OR interaction OR limitation OR maintainance OR management OR manifestation OR metamorphosis OR modification OR monitoring OR mutation OR observation OR opportunity OR organisation OR organization OR performance OR possibility OR potential OR practice OR preference OR preservation OR process OR program OR progress OR protection OR provision OR reaction OR reconstruction OR recovery OR reduction OR regeneration OR regulation OR representation OR response OR restoration OR rivalry OR selection OR service OR shift OR skill OR stewardship OR stress OR success OR support OR surveillance OR tendency OR training OR transfer OR transformation OR transition OR transport OR understanding OR utility) OR (abundance OR advantage OR allocation OR amount OR availability OR balance OR change OR concentration OR constitution OR count OR decline OR decrease OR density OR development OR dispersal OR distribution OR diverseness OR diversity OR evolution OR existence OR expansion OR gain OR growth OR heterogeneity OR import OR improvement OR increase OR invasion OR living OR merging OR occurrence OR output OR presence OR prevalence OR production OR productivity OR progression OR proximity OR quantity OR range OR ratio OR richness OR scattering OR spread OR stream OR synthesis OR suitability OR sustainability OR variability OR variation OR variety OR volume OR wealth OR weight OR yield)) AND (urban* OR architectural OR architectonic* OR “building science” OR “city science”) AND ((anthropogenic OR built OR “built-up” OR civic OR civil OR constructed OR human OR people OR population OR public) OR (city OR cities OR conurbation OR environs OR metropol* OR town OR suburb*)) AND ((area OR boundary OR coast OR complex OR compound OR context OR corridor OR cradle OR earth OR field OR forest OR grassland OR ground OR habitat OR land OR landscape OR milieu OR place OR region OR river OR site OR soil OR substrate OR surface OR surrounding OR terrain OR territory OR topography OR water OR zone) OR (climate OR environment OR light OR temperature OR weather) OR (atrium OR building OR construction OR court OR edifice OR garden OR home OR house OR housing OR infrastructure OR park OR patio OR plaza OR plot OR yard OR settlement OR square OR station OR structure) OR (canopy OR cover OR envelope OR facade OR face OR layer OR material OR matter OR mesh OR remains OR roof OR skin OR web)) AND ((component OR constant OR data OR dataset OR detail OR element OR entity OR equipment OR format OR graph OR image OR information OR item OR object OR part OR program OR project OR proposal OR proxies OR proxy OR replica OR resource OR system OR task OR technology OR tool OR type OR unit) OR (constraint OR degree OR density OR design OR depth OR dimension OR distance OR extent OR form OR formation OR geometry OR gradient OR height OR layout OR level OR limit OR line OR link OR marker OR morphology OR network OR pattern OR perspective OR plan OR ratio OR scale OR shape OR size OR slope OR source OR space OR spacial OR spatial OR texture OR volume OR vector) OR (attribute OR benchmark OR characteristic OR criteria OR criterion OR factor OR feature OR genre OR hallmark OR indicator OR label OR measure OR method OR metric OR norm OR parameter OR property OR quality OR requirement OR scheme OR solution OR standard OR strategy OR sustainability OR tendency OR technique OR terms OR theory OR trait OR trend OR variable)) AND ((adaptation OR adjustment OR approach OR augmentation OR boost OR change OR conversion OR decrease OR diffusion OR dispersal OR expansion OR extension OR gain OR growth OR imitation OR inclination OR incorporation OR increase OR innovation OR metamorphosis OR mining OR modification OR practice OR presence OR progress OR reconstruction OR redesign OR regulation OR revision OR separation OR shift OR transfiguration OR transformation OR transition) OR (allocation OR amalgamation OR arrangement OR assemblage OR basis OR composition OR concentration OR condition OR configuration OR confinement OR connection OR constellation OR contrast OR distinctiveness OR distribution OR diverseness OR diversity OR flexibility OR group OR heterogeneity OR hierarchy OR imbalance OR irregularity OR limitation OR linkage OR location OR mix OR origin OR organisation OR organization OR point OR position OR proximity OR range OR relation OR relationship OR resolution OR scattering OR setting OR sequence OR similarity OR spot OR spread OR suitability OR variability OR variance OR variation OR variety)) AND (analy* OR assess* OR catalog* OR categor* OR class* OR cluster* OR codif* OR compar* OR correlat* OR defin* OR demarcat* OR descr* OR detect* OR determin* OR differ* OR discover* OR discrimin* OR distinct* OR distribut* OR estimat* OR evaluat* OR examin* OR explor* OR extract* OR forecast* OR identif* OR index* OR indicat* OR investigat* OR map* OR measur* OR model* OR predict* OR prognos* OR rat* OR referenc* OR sampl* OR scenario OR segment* OR select* OR simulat* OR suggest*)

Appendix B - Review Details of ML Applications from the Main Literature Search

Authors	Publication year	Title	Publication
Bai et al. (2022)	2022	Urban Green Space Planning Based on Remote Sensing and Geographic Information Systems	Remote Sensing
Ban et al. (2022)	2022	Landscape Ecological Construction and Spatial Pattern Optimization Design Based on Genetic Algorithm Optimization Neural Network	Wireless Communications and Mobile Computing
Barbierato et al. (2020)	2020	Integrating Remote Sensing and Street View Images to Quantify Urban Forest Ecosystem Services	Remote Sensing
Bergerot et al. (2011)	2011	Landscape variables impact the structure and composition of butterfly assemblages along an urbanization gradient	Landscape Ecology
Bonilla-Bedoya et al. (2021)	2021	Urban Soils as a spatial indicator of quality for urban socio-ecological systems	Journal of Environmental Management
Chapman et al. (2020)	2020	Invasion of freshwater ecosystems promoted by network connectivity to hotspots of human activity	Global Ecology and Biogeography
Chen et al. (2023b)	2023	A Framework for Assessing Trade-Offs and Synergies in Green Space System Services Based on Ecosystem Services Bundles	Forests
Dimopoulos et al. (1999)	1999	Neural network models to study relationships between lead concentration in grasses and permanent urban descriptors in Athens city (Greece)	Ecological Modelling
Dong et al. (2022)	2022	Optimization of green infrastructure networks based on potential green roof integration in a high-density urban area - A case study of Beijing; China	Science of the Total Environment
Hassell et al. (2021)	2020	Socio-ecological drivers of vertebrate biodiversity and human-animal interfaces across an urban landscape	Global Change Biology
He et al. (2022)	2022	Plant Landscape Configuration Method of Regional Characteristic Rainwater Garden Based on Deep Learning	Proceedings of the 4th EAI International Conference; ICMTEL 2022
Jiao and Han (2022)	2022	Urban Green Space Planning and Design for Sponge City	Scientific Programming
Jutras et al. (2002)	2002	Urban tree growth moedlling with artificial neural network	Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02
Jutras et al. (2009)	2009	Prediction of street tree morphological parameters using artifical neural networks	Computers and Electronics in Agriculture
Karapinar Senturk (2022)	2022	Amphibian species detection in water reservoirs using artificial neural networks for ecology-friendly city planning	Ecological Informatics
Labib (2019)	2019	Investigation of the likelihood of green infastructure (GI) enhancement along linear waterways or on derelict sites (DS) using machine learning	Environmental Modelling and Software
Le and Huang (2023)	2023	Prediction of Urban Trees Planting Based on Guided Cellular Automata to Enhance the Connection of Green Infrastructure	Land
Li and Fan (2022)	2022	Evaluation of urban green space landscape planning scheme based on PSO-BP neural network model	Alexandria Engineering Journal
Lin et al. (2018)	2018	Spatio-Temporal Analysis of Wetland Changes Using a Kernel Extreme Learning Machine Approach	Remote Sensing
Łopucki and Kiersztyn (2020)	2020	The city changes the daily activity of urban adapters: Camera-traps study of Apodemus agrarius behaviour and new approaches to data analysis	Ecological Indicators
Luhua et al. (2022)	2022	Intelligent Identification of Building Patches and Assessment of Roof Greening Suitability in High-density Urban Areas: A Case Study of Chengdu	Journal of Resources and Ecology
Morshed et al. (2022)	2022	Future ecosystem service value modeling with land cover dynamics by using machine learning based Artificial Neural Network model for Jashore city; Bangladesh	Physics and Chemistry of the Earth
Nölke et al. (2023)	2023	Categorization of green and grey infrastructure complexity in the rural–urban interface of Bengaluru; India: an unsupervised volumetric approach with relevance for urban quality	Urban Ecosystems
Ringland et al. (2021)	2021	Automated survey of selected common plant species in Thai homegardens using Google Street View imagery and a deep neural network	Earth Science Informatics
dos Santos et al. (2021)	2021	Dynamics of environmental conservation: Evaluating the past for a sustainable future	International Journal of Applied Earth Observation and Geoinformation
Steenberg et al. (2019)	2019	A Social-Ecological Analysis of Urban Tree Vulnerability for Publicly Owned Trees in a Residential Neighborhood	Arborculture & Urban Forestry
Sun et al. (2021)	2021	Assessing the cooling efficiency of urban parks using data envelopment analysis and remote sensing data	Theoretical and Applied Climatology
Wang et al. (2024)	2024	Green roofs and their effect on architectural design and urban ecology using deep learning approaches	Soft Computing
Wang et al. (2021)	2021	Classification of Street Tree Species Using UAV Tilt Photogrammetry	Remote sensing
Wellmann et al. (2020)	2020	Earth observation based indication for avian species distribution models using the spectral trait concept and machine learning in an urban setting	Ecological Indicators
Wiese et al. (2019)	2019	Integrating environmental and neighborhood factors in MaxEnt modeling to predict species distributions: A case study of Aedes albopictus in southeastern Pennsylvania	PLoS ONE
Zaleckis et al. (2022)	2022	Modeling of Changes in Four Urban Capitals Using Up-to-Date Information Systems and Mathematical Graph-Based Simulative Models for Urban Regeneration (Kaunas Case)	Sustainability
Zhang et al. (2023)	2023	Decoding urban green spaces: Deep learning and google street view measure greening structures	Urban Forestry & Urban Greening
Zubair et al. (2021)	2021	Comparative Analysis and Prediction of Ecological Quality of Delhi	Advances in Energy and Environment - Selected Proceedings of TRACE 2020

Authors	Study area	Data	Data pre-processing
Bai et al. (2022)	Fendgdong New City; Xi'ian New Area; China	Landsat 8 OLI_TIRS (30m); ALOS terrain elevation (12m); daily rainfall; organic carbon content in soil; soil texture spatial distribution; nightlight image data; annual surface climate; annual climate;	Calculation of various metrics
Ban et al. (2022)	Not specified	Not specified	Not specified
Barbierato et al. (2020)	Viareggio; Tuscany; Italy	UltraCam Xp multispectral frames (RGB+NIR; 0.2m); LiDAR image: topographic regional database (1:2000); OpenStreetMaps: street blocks; GoogleStreetView images;	Alignment of data at 1x1m resolution
Bergerot et al. (2011)	Paris; France	26 butterfly species occurrence data; habitat type; temperature; cloud cover; ground cover	Not specified
Bonilla-Bedoya et al. (2021)	Quito; Ecuador	Field survey: soil samples; DEM; SPOT7 imagery; demographic and geographic data	Spectroscopy/not specified
Chapman et al. (2020)	England; UK	Car park distance; fishery distance; boat access distance; human population; boat navigation distance; river length; elevation; lake area; northing; easting; protected area	Calculation of indices
Chen et al. (2023b)	Foshan City; China	Land cover from satellite images (Google Earth; 0.5m)	Not specified
Dimopoulos et al. (1999)	Athens; Greece	Species occurrence and contamination; spatial metrics	Calculation of descriptor variables
Dong et al. (2022)	Beijing; China	RS imagery Google Maps (0.46m)	Data augmentation (5000 out of 30 samples)
Hassell et al. (2021)	Metropolitan Nairobi (stratified samples)	ESRI World Imagery (1:500); demographic data; survey data	Classify land use; calculate Simpson's index; elevation
He et al. (2022)	Not specified	Rgb images	Not specified
Jiao and Han (2022)	Not specified	Not specified	Not specified
Jutras et al. (2002)	Montreal; Canada	Soil compaction; mechanical damage of trees; available light; aerial growing space; underground growing space; de-icing salt exposure; street type; street orientation; width of the street; disttance tree to nearest building; volume tree pit; tree: genus; species; cultivar; age; DBH; crown width; tree height; phytopathological condition; crown development; general condition; diameter growth index; tree hieght growth index; total crown growth index	Not specified
Jutras et al. (2009)	Montreal; Canada	DBH; Annual DBH increment; height; annual height increment; crown diameter/DBH; crown volume/DBH; height/DBH; crown volume; annual crown volume increment; irradiation; street width; distance tree to building; distance tree to curb; tree pit volume; urban zone; surficial deposit	Data normalization
Karapinar Senturk (2022)	Poland	GIS; satellite images; EIA reports	Not specified
Labib (2019)	Greater Manchester City Region; UK	GIS data from various sources	Not specified
Le and Huang (2023)	Camden; London; UK	Green Space; Urban trees; NDVI; Land cover (woodland); carbon emissions	Not specified
Li and Fan (2022)	Wanning district	Patch density; shannon diversity; avg. Perimeter area ration; spread degree	Not specified
Lin et al. (2018)	Dongtan wetland; Chongming Island; Shanghai; China	Temporal Landsat images (1986-2013)	Construction of reference dataset based on ground survey
Łopucki and Kiersztyn (2020)	Camera traps (18 rural; 15 urban sites)	Camera trap images	Not specified
Luhua et al. (2022)	Chengdu; China	GF-2 images (0.8m); Landsat 8 (30m) satellite images; ancilliary population; area data (Chengdu Statistical Yearbook); building height; road network	Radiometric calibration; atmospheric correction; vectorization of road network; calculation of density metrics; resampling of rasta data
Morshed et al. (2022)	Jashore City; Bangladesh	Landsat images (2000;2010;2020)	Radiometric and atmospheric correction; generation of composite band combinations
Nölke et al. (2023)	Bengaluru; India	WorldView-3 (1.2m/0.3m)	Data fusion; interpolation of multispectral bands to 0.3m
Ringland et al. (2021)	Chiang Mai; Lamphun; Lampang; Phayao; and Phrae provinces; Thailand	Google Street View raster images	Image tagging with 11 species
dos Santos et al. (2021)	Sao Paulo; Brazil	Land use data; geographic and map data	Standardization
Steenberg et al. (2019)	Harbord Village/ Toronto; Ontario; Canada	Field data and public tree inventory (2007/08 and 2014); species; DBH; location; indicators of urban forest vulnerability	Aggregated indexing of tree condition
Sun et al. (2021)	Shanghai; China	Landsat 8 TIRS images (August 2023 and February 13; 2017)	estimation of LST; calculated NDVI
Wang et al. (2024)	Shanghai; China	Delphi method (expert evaluation); remote sensing images	Expert survey for evaluation indicators of CASBEE method; Tefel method
Wang et al. (2021)	Beijing and Zhangjiakou City; China	UAV imagery	Pix4Dmapper; 3d point cloud generation; tree segmentation
Wellmann et al. (2020)	Leipzig; Germany	RapidEye EO data	NDVI; PCA; indicator calculation
Wiese et al. (2019)	Pennsylvania; USA	Mosquito traps; neighborhood factors: census data; land cover raster	Raster resampling (232m)
Zaleckis et al. (2022)	Kaunas; Lithuania	Open Street Map	Calculation of graph centralities: closeness centrality; betweenness centrality; simple density; accessible length at 4 different radii
Zhang et al. (2023)	Livingston; New Jersey; USA	Google Street View raster images; road network data	Calculation of PVGVI
Zubair et al. (2021)	Delhi; India	Sentinel 2A (10m)	Image band processing

Authors	Features	Research Task	ML task	ML algorithm used
Bai et al. (2022)	Biodiversity; water resource conservation capacity slope; relative location; vegetation cover; soil conservation capacity; outdoor recreation intensity; ecological unit scarcity	Construction of urban green space ecological networks	Analysis of urban green space ecological sources; extraction of indicator weights	Back Propagation Neural Network
Ban et al. (2022)	Not specified	Optimization of spatial pattern	Simulate individual evaluation function in GA to reduce execution cost	BPNN
Barbierato et al. (2020)	Rgb image (400x400)/ PCA calculated parameters	Quantification of urban forest ecosystem services	Semantic segmentation; spatial clustering	Nearest Neighbour; CNN; REDCAP
Bergerot et al. (2011)	Not specified	Clustering of butterfly assemblages	Species abundance correlation to environmental metrics	SOM
Bonilla-Bedoya et al. (2021)	Elevation; slope; temperature; precipitation; PM2.5; spectral bands; population density; land cover; green infrastructure; distance to soil sample; distance to roads; vehicular traffic	Predict ecological quality from soil samples at geographic location	Predict spatial distribution	RF
Chapman et al. (2020)	See data	Connections of freshwater networks; human activity and species invasion	Predict various human activities; model fishing	Generalized Linear Model; Generalized Boosting Model; Distributed Random Forests; DLNN; Super Learning (Stacked ensemble of algorithms)
Chen et al. (2023b)	Not specified	Trade-offs and synergies in UGS for ES bundles	Segmentation and classification of land cover; dimensionality reduction to ES bundles	PCA; SOM
Dimopoulos et al. (1999)	Mean density; mean vegetation height; wind velocity; mean height of adjacent buildings; distance sample point to nearest adjacent street; traffic volume	Relationship between plant lead concentration and urban descriptors	Inference/prediction plant characteristics and urban descriptors; study of influencing factors through linear regression	MLP (SPSS)
Dong et al. (2022)	Rgb image(1000x1000)	GI infrastructure improvement by integration of green roof potential	Classification of potential green roofs/ feature extraction	DLNN
Hassell et al. (2021)	Habitat diversity; elevation; % grassland; % shrubs; % trees; % artificial; % bareground; total pigs; weath index; household area; etc.	Drivers of vertebrate biodiversity in an urban landscape	Co-occurrence of wildlife and livestock species; composition of urban vertebrate	SOM
He et al. (2022)	Not specified	Predicting plant configuration	Estimation of plant parameters	CNN
Jiao and Han (2022)	Patch density; shannon diversity; avg. Perimeter area ration; spread degree	Planning of UGS networks in sponge cities	Network training to infer useful weights and thresholds	BPNN
Jutras et al. (2002)	See data	Urban tree growth modelling	Prediction of tree growth	ANN
Jutras et al. (2009)	See data	Prediction of street tree morphological parameters	Prediction of DBH; annual DBH increment; crown volume; height; annual height increment; crown volume; annual crown increment	BPNN
Karapinar Senturk (2022)	Surface of water reservoir; water reservoirs in habitat; type of water reservoires; dominant types of land cover; type of shore; vegetation intensity;maintenance status; planned use of water reservoirs; fishing activity; distance to buildings; distance to roads; percentage access from edges to open areas	Consideration of amphibian occurrence according to environmental variables	Prediction of species occurrence	ANN
Labib (2019)	Site size; surrounding tree coverage; air pollution; NO{sub}2 concentration; population density; accessibility; surrounding built up area	Green infrastructure enhancement along waterways	Explain/predict land-use changes	ANN; ANFIS
Le and Huang (2023)	Water conservation; plant health; NDVI; sunlight; tree canopy; tree species;	Estimate tree planting location based on health prediction	Image transformation urban pattern and mycorrhizahl networks	CycleGAN (image transformation); CA
Li and Fan (2022)	Landscape fragmentation; spatial heterogeneity; landscape capital structure; optimal landscape pattern	Analytical support for urban green space planning	Predict ecological variables from urban green space layouts	PSO-BPNN
Lin et al. (2018)	NDVI; wetness (Kanth-Thomas); optimal bandwith combination	Spatio-temporal analysis of wetland changes	Change detection	Kernel Extreme Learning Machine (K-ELM); ELM; SVM; MLC
Łopucki and Kiersztyn (2020)	Species presence	Animal behaviour in urban environments	Species detection	Kernel Density Estimation; PSO; decision tree; NN
Luhua et al. (2022)	Building sample points and surfaces	Identification of building patches for green roof suitability	Detection of building shapes	D-LinkNet (DL)
Morshed et al. (2022)	Various band composite images	Future ecosystem value modeling	LULC classification and change prediction	Maximum Likelihood estimation; Cellular Automata-Markov Chain (CA-MC); MLP-MC network
Nölke et al. (2023)	Images (112x112); probability map (72x72); DSM (from satellite stereo pairs)	Classification of green and grey infrastructure at an rural-urban interface	Determine composition and configuration of urban green and grey infrastructure at 3 distinct scales: imperviousness; neighborhood imperviousness; infrastructure complexity	CNN (U-Net); hierarchical clustering (factoextra R)
Ringland et al. (2021)	Raster image (3840x2160)	Survey of plant species in homegardens	Supervised learning detection of plant species	CNN (RetinaNet)
dos Santos et al. (2021)	Raster image of land use; soil type; slope; drainage system; urban area; roads	Predict levels of environmental conservation	Model land use change	CA-Markov; ANN
Steenberg et al. (2019)	Tree mortality; diameter growth rate; neighbourhoods-derived tree condition indices	Tree vulnerability analysis at neighborhood level	Prediction of mortality outcomes based on vulnerability indicators	IBM SPSS Statistics 24 (MLPNN)
Sun et al. (2021)	geogrpahic variables; cooling indices	cooling effieciency of urban parks	causal inference of geographic factors on cooling efficiency	ELM; XGBoost
Wang et al. (2024)	Level 1 indicators: services provided; outdoor physical environment; outdoor resource protection; indoor environmental quality	Effect of green roofs on design	Evaluation based on grades; i.e. classification	CNN-LSTM
Wang et al. (2021)	Tree height; crown width; crown height; canopy volume; 4 ratio metrics	Classification of tree species	See overall task	KNN; RF; SVM; BPNN
Wellmann et al. (2020)	Spatial heterogeneity; vegetation density	Estimation of avian species distributions in urban area	Predict presence/absence of single species; species clusters; and species richness	RF
Wiese et al. (2019)	Mosquito Presence and Absence; Below Poverty; Best Housing Conditions; Education; Median Household Income; Housing Density; Population Density; Urban Population; Vacant Housing Units; Worst Housing Conditions; Imperviousness of the surfaces; Land Cover; Tree Canopy; Average Precipitation; 3-month avg. Precipitation; avg. Temperature; 3-month avg. Temperature; avg. EVI; avg. NDWI; Elevation; Slope; Flow Accumulation	Predict species distribution	Predict species distributions/ infer causal variables	MaxEnt
Zaleckis et al. (2022)	Street network; inhabitants density; mean land prices	Model urban capitals	Data modeling (prediction) indicator development of urban capitals	NN (Matlab)
Zhang et al. (2023)	Raster image	Identification and measurement of green structures	Image segmentation; and classification of vegetation structure	DeepLabv3+ (CNN); SegNet; PSPNet
Zubair et al. (2021)	NDMI; NDBI; NDVI; geospatial ecological impact index	Prediction of ecological quality	Mapping forecast of environmental indicator	ANN (no details available)

Authors	**Evaluation of ML**	**Description of eco-complexity**	Observables addressed	**Spatio-temporal complexity**
Bai et al. (2022)	ML to reduce human subjectivity; no critical appraisal or evaluation	Static distribution of ecological sources and inferred indicator weights	Connectedness	Spatial (pixel; local)
Ban et al. (2022)	Combination of NN and GA reduced time cost without affecting performance	Inconclusive	Inconclusive	None
Barbierato et al. (2020)	Good performance on segmentation; sensitivity to correlated attributes of clustering regionalization; solved by PCA	Spatially bound geographic clustering; considering contiguity	Configuration	Spatial (pixel; local)
Bergerot et al. (2011)	SOM and ANN proved promising tools; clustering mechanism sensitive to contiguity/neighborhoods;	Static distribution sensitive to contiguity/neighborhoods	Configuration	None
Bonilla-Bedoya et al. (2021)	Effective; but other predictors might improve performance - for urban application	Static spatial distribution with multivariate parameters	Configuration	None
Chapman et al. (2020)	Different algorithms predicted different activities best	Static prediction of (human) behaviour based on indices	Inconclusive	None
Chen et al. (2023b)	Limited representation of UGS due to dimensionality reduction	Synergies and trade-offs	Inconclusive	Spatial (pixel; local)
Dimopoulos et al. (1999)	Good performance; possible connection to buildings	Explanatory modeling of static factors	Inconclusive	None
Dong et al. (2022)	Acceptable performance on feature extraction	Inconclusive	Configuration	Spatial (pixel; local)
Hassell et al. (2021)	Not specified	Static relationships of ecological factors and species diversity	Configuration	None
He et al. (2022)	Not specified	Functional capacity according to	Inconclusive	Spatial (pixel; local)
Jiao and Han (2022)	Minimized errors through according sample sizes	Estimation of influencing factors	Inconclusive	None
Jutras et al. (2002)	Good performance for prediction	Modelling future discrete state of organisms/population	Change	None
Jutras et al. (2009)	Average prediction of 91%	Modelling future discrete state of organisms/population	Change	None
Karapinar Senturk (2022)	Accuracy of 80-90%	Predict impacted species	Configuration	None
Labib (2019)	Weak capability of explanatory factors; good prediction of changes	Predict neighborhood changes	Change	None
Le and Huang (2023)	Limited by few features included in tree species evaluation	Evolutionary population dynamics; spatially determined	Memory	Spatial (pixel; local)
Li and Fan (2022)	Implementation of PSO with BPNN is improving learning ability	Design feedback	Configuration	Spatial (pixel; local); sequential
Lin et al. (2018)	K-ELM outperforms other algorithms	Spatio-temporal prediction	Change	Spatial (pixel; local); sequential
Łopucki and Kiersztyn (2020)	PSO and NN good results	Species occurrence; activity	Interaction	None
Luhua et al. (2022)	Building shapes were quickly and accurately obtained	Supporting process; static evaluation	Configuration	Spatial (pixel; local)
Morshed et al. (2022)	High accuracy for change prediction	Change prediction based on sequential data	Change	Spatial (pixel; local); sequential
Nölke et al. (2023)	High accurarcy in classification task; differentiated clusters attained	3-dimensional clustering of morphological complexity	Configuration	Spatial (pixel; local)
Ringland et al. (2021)	Satisfactory for most species	Mapping of species distribution; missing complexity	Configuration	Spatial (pixel; local)
dos Santos et al. (2021)	Relatively low prediction values were accepted	Dynamic 2D distribution of classes	Change	Spatial (pixel; local)
Steenberg et al. (2019)	Not specified	Inference of ecological condition/ change	Change	None
Sun et al. (2021)	acceptable performance	ecological performance; explanatory factors 2-dimensional	Inconclusive	Spatial (pixel; local)
Wang et al. (2024)	LSTM superior to CNN; mixed results in total	Inconclusive	Inconclusive	Spatial (pixel; local)
Wang et al. (2021)	BP NN best results; then KNN; RF; SVM; results depending on species	Biotic inventory; static;	Configuration	Spatial (pixel; local)
Wellmann et al. (2020)	Methods supports assumption that vegetation pattern represent species through dietary traits	Biodiversity factors with spatial relevance	Configuration	Spatial (pixel; local)
Wiese et al. (2019)	Accuracies of three models around 75 %	Inference of neighborhood factors (non-spatial)	Configuration	Spatial (pixel; local)
Zaleckis et al. (2022)	Positive validation	Identification of ecological allocation at different scales	Configuration	Spatial (graph)
Zhang et al. (2023)	CNN better performance	Identifiaction of organisms; structuring of landscapes	Configuration	Spatial (pixel; local)
Zubair et al. (2021)	Not specified	Spatial (2D); temporal (time-step)	Configuration	Sequential