‘INITIAL STEPS TOWARDS AN OPEN CADASTRE IN THE UK: A CASE STUDY’

Luke Taylor

Thesis submitted in partial fulfilment of the requirements for the degree of

Master of Research in Geospatial Data Science

Nottingham Geospatial Institute

University of Nottingham

30 Triumph Rd, Lenton,

Nottingham, NG7 2TU

United Kingdom

14th September 2023

Abstract

Purpose: To develop a proof-of-concept for a UK community cadastre using open data and open-source software.

Design/Methodology/Approach: The study used a community-led approach, working with a climate change charity in The Meadows, Nottingham, as a case study. A methodology proposed for developing an open-source data processing workflow to provide communities with detailed information on housing archetypes and land ownership. The research developed a proof-of-concept for a UK community cadastre to support local retrofitting measures such as loft insulation to reduce carbon emissions. The result uses open data sources such as Local Authority Data, OpenStreetMap, and INSPIRE Index Land Polygons. It uses open-source software, including QGIS, MMQGIS plugin, Jupyter Notebooks, Python, and Felt.com.

Findings: The study successfully developed a proof-of-concept open-source web application, leveraging exclusively open data and open-source software.

Practical implications: The findings have the following practical implications: 1) The study provides a scalable model that can be adopted by other communities aiming to improve their housing stock's energy efficiency and reduce carbon emissions; 2) The use of open data and open-source software makes this approach financially accessible, even for communities with limited resources. 3) By offering a transparent view of land ownership, the application could potentially democratise land use management, offering an alternative to the existing secretive system in the UK.

Originality/value: This study is original in applying INSPIRE Index Polygons, a data set with limited exploration in geospatial research, and in creating a workflow that brings together relevant open-source software with open data. The developed resource can accelerate community-driven efforts to combat climate change while also beginning to address the complexities and opacity of land ownership in the UK.

Keywords: climate change mitigation, community development, land use management, open source, cadastre

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Professor Doreen Boyd, for her unwavering support and guidance throughout this research. Her insights and expertise have been invaluable, and I am grateful for her willingness to challenge me and help me grow as a researcher.

I would also like to thank Cristina Vrinceanu from the Nottingham University Rights Lab for her technical challenge and expertise. Her work has significantly enriched this research, and I am grateful for her willingness to share her knowledge.

I am indebted to Ambreen, my research "big sister," aka "mi hermana", for her mentorship and crucial role in supporting me - with pep talks and navigational tips.

I deeply appreciate my friends. To AJ and Jarede, my brothers from another mother, thank you for continually challenging, pushing, and sharpening me both personally and academically. To Anne, your wisdom, counsel, and spirit of adventure keep me focused, fighting, and hopeful. Michaela, your support and belief have inspired and strengthened me. Norris, your distractions were always welcome and often needed, providing fun breaks along the way. To Carlo, for your jumps, sticking by me during the early morning gym routine and your company. Love, as always, to my family.

I would also like to thank the staff at Green Meadows for sharing their time, expertise, and data with me. Without their support, this project would not have been possible.

Finally, I would like to thank everyone working to support open data and open-source software. You are the enablers, the catalysts for innovation, and the democratisers of knowledge in the digital age.

I am also grateful for the financial support provided by the EPSRC Centre for Doctoral Training (CDT) in Geospatial Systems (REF EP/S023577/1). This funding allowed me to leave the 9-5 grind and focus on a project that was more aligned with my goals.

Table of Contents

1. Introduction, Aims and Objectives

Background

Environmental management and land ownership are inextricably linked, affecting many factors ranging from sustainable development to carbon emissions(Augustinus, 2023; Jones, 2023; Stanley, 2023). For example, in England, the common practice of burning heather on peatlands – lands which are significant carbon sinks – demonstrates how land use decisions have direct repercussions on carbon emissions ('England's Peatlands: Greenhouse Gases and Carbon Storage. Natural England report NE257', 2010). Indeed, the UK's management of peatland has been the subject of international scientific debate (Davies *et al.*, 2016; Ashby and Heinemeyer, 2021).

The complexity of land ownership has implications for policy and planning initiatives, such as the proposed Manchester congestion charge, which faced substantial lobbying from various landowners and stakeholders (Shrubsole, 2019b). Such examples underscore the necessity for transparent land ownership, as highlighted by Guy Shrubsole's evidence presented to the House of Lords, which attests to the significant influence of undisclosed land ownership on environmental policy ('Written Evidence (LUE0027) to the House of Lords', 2022).

In addition, given that the UK's housing stock is among the oldest and least energy-efficient in Europe, retrofitting these homes becomes imperative for meeting the country's carbon reduction targets (Ürge-Vorsatz and Metz, 2009; Nejat *et al.*, 2015; De Urquia, 2022). This research introduces the concept of a 'community cadastre', an open-source platform that amalgamates land ownership and housing archetype data, as a means to empower local communities. By demystifying the maze of land ownership, this tool aims to facilitate community-driven retrofitting initiatives, ultimately contributing to national carbon reduction and environmental sustainability goals.

Research Question:

- How practical are open geospatial datasets and geospatial open-source software in enabling communities to create their own 'community cadastre'?

Aim:

- To evaluate the potential of open geospatial data and geospatial open-source software in empowering communities to create their community cadastre.

Research Objectives:

- Partner with 'Green Meadows' to **define** a 'community cadastre' and its key components.
- **Identify** open geospatial data and software for cadastre creation.
- Identify gaps in geospatial data and compare machine learning classifiers to **predict** building age.
- **Develop** and **evaluate** a 'community cadastre' proof-of-concept.

Figure 1: Structure of this dissertation.

2. Literature Review

2.1 ‘Cadastre’, Defined

The professional body of surveyors, FIG (International Federation of Surveyors), define a cadastre as:

> *“A Cadastre is normally a parcel based, and up-to-date land information system containing a record of interests in land (e.g. rights, restrictions and responsibilities). It usually includes a geometric description of land parcels linked to other records describing the nature of the interests, the ownership or control of those interests, and often the value of the parcel and its improvements. It may be established for fiscal purposes (e.g. valuation and equitable taxation), legal purposes (conveyancing), to assist in the management of land and land use (e.g. for planning and other administrative purposes), and enables sustainable development and environmental protection.”*

FIG Statement on the Cadastre (FIG, 1995)

Content of a Cadastre

I will now briefly discuss each element of the FIG's definition of a cadastre and highlight the three different elements of the cadastre.

1. A Geometric Description of Land Parcels

The geometric description is a foundational element of cadastres. Without precise geometry, the specific area under consideration may become ambiguous and difficult to identify accurately, echoed by Henssen (Henssen, 1995), who asserts that the primary task of a cadastre is "geometrically oriented, i.e. fixing and representing the parcel". In other words, the absence of rigorous geometric data could render any parcel of land indeterminable.

The question of geometric accuracy often finds debate within the land surveying community. Some professionals prioritise broad coverage with lower accuracy, while others focus on higher precision in surveying techniques. The former approach is favoured by the "Fit for Purpose Land Administration" work (Enemark, 2014; Bennett and Alemie, 2016; Bennett, Koeva and Asiama, 2021), while others focus on high-precision methodologies (Craig and Wahl, 2003; Hanus, Pęska-Siwik and Szewczyk, 2018; Pullar and Donaldson, 2022). The geometric component is essential to the utility and precision of cadastral systems, but the debate over its optimal accuracy level continues to divide the professional community.

2. A Record of Interests in the Land

The records of interests in land have changed over time. In the past, these records were paper-based and held in a centrally located office building by the central government. However, with the increasing use of technology, these records are now most often stored in digital database systems. Depending on the country's preferred infrastructure - systems can be decentralised, open, and highly accessible. Whether paper or digital, these systems all record the same information: interests in land.

The term "interest" can have a wide and varied definition. Dr. Clarissa Augustinus, Chief of the Global Land Tools Network at UN-Habitat, has dedicated her career to advocating for a continuum of land rights, recognising the diversity of land rights (Global Land Tools Network, 2023). In a post-colonial world with a universal declaration of human rights, it is now widely accepted that these differing rights reflect an individual or community's rights to land. Such examples are successfully implemented through political means and put into action, often through the work of UN-Habitat.

With its partner countries, UN-Habitat has worked to create comprehensive national land cadastres in the developing world, where such cadastres were once lacking. As a testament to the GLTN's work, many governments have successfully achieved cadastral systems that reflect the diverse nature of interests in land (Griffith-Charles *et al.*, 2015; Ding *et al.*, 2019). Whether monetary or non-monetary, all of these interests have what might be considered value - the third element of the FIG definition of a cadastre (Zevenbergen *et al.*, 2013).

The evolution of record-keeping systems for cadastres is evident, transitioning from paper-based repositories maintained by central governments to sophisticated digital database systems (Zevenbergen *et al.*, 2013). Technological advances enable these databases to be incredibly decentralised, offering high accessibility through innovations like blockchain and open data (Lakomaa and Kallberg, 2013; Krigsholm *et al.*, 2017; Martyn, 2018). Regardless of the medium—whether paper or digital—the central focus remains the recording of interests in the land.

The notion of "interest" has been substantially influenced by the lifetime contributions of Dr. Clarissa Augustinus, who argued for a continuum of land rights to acknowledge the diversity of interests in land. This diversity of interests is now largely recognised as extending to individual and community land rights in the contemporary, post-colonial context and alignment with universal human rights. UN-Habitat, partnering with nations—particularly in the developing world—works to operationalise this perspective into functional cadastral systems that effectively capture the heterogeneity of land interests (Griffith-Charles, 2011; Zevenbergen *et al.*, 2013; Griffith-Charles *et al.*, 2015). Whether monetary or non-monetary, these interests are encapsulated as a value component, making them a fundamental part of the FIG's definition of a cadastre. Advancements in technology and sociopolitical understanding have changed cadastral records and what those records entail, making them more inclusive and diverse.

3. The Value of the Parcel and its Improvements

The definition of a cadastre *often* includes the value of a land parcel and its improvements. However, this may vary depending on the country's policies and objectives. The primary purpose of this information is to facilitate property taxation by governments (Bahl and Vazquez, 2008). Property taxes can capture the value of land that has increased due to government investments, whether in mature or emerging economies (Dale, 1977; Slack, 2018). Effective cadastral systems should record land value to ensure proper tax collection, which benefits both the government and the nation as a whole.

The term "improvements" in a cadastre can refer to anything from structural modifications, such as adding rooms, to broader regional developments, such as proximity to new transportation infrastructure. These "land value capture" policies are progressive taxes because they primarily burden landowners, potentially reducing inequality (Chi‑Man Hui, Sze‑Mun Ho and Kim‑Hin Ho, 2004; Medda, 2012; Vejchodská *et al.*, 2022). In countries like Denmark, Estonia, Lithuania, Russia, Singapore, and Taiwan, cadastres record these "improvements" for taxation purposes. It could be that the primary utility of a cadastre is for tax or fiscal objectives.

In summary, cadastral systems have three key elements: 1: Geometric descriptions for precise land identification, 2: Records of varied land interests, 3:Valuation for fiscal purposes such as taxation. While these core components remain the same, technological advances and evolving sociopolitical contexts have made modern cadastres increasingly complex and versatile tools for land management.

The Role of Cadastres

The FIG statement provides four reasons for a cadastre:

1. Fiscal purposes;
2. Legal purposes;
3. Assist in managing land and land use, enabling sustainable development and environmental protection;
4. Land Management for Sustainable Development and Environmental Conservation.

Cadastres serve various purposes, but these are the four most common reasons -

1 - Fiscal Purposes

Fiscal objectives often serve as the primary motivation for governments to establish cadastre systems, aligning with broader economic goals related to land management—one of a country's most valuable assets (Amandala, 2007). For example, the World Bank, notably in collaboration with the UN's Global Land Tools Network, has extensively endorsed the concept that secure land tenure is intrinsically linked to agricultural productivity and land value (Feder, 1986; Van Bronkhorst, 2023). These organisations have been pivotal in advocating for cadastre systems as essential frameworks for economic development, particularly in developing nations (Österberg, 2001). The implementation of such policies has successfully facilitated the recording of land rights, yielding advantages at both the individual and governmental levels (Feder, 1986; Österberg, 2001; Van Bronkhorst, 2023).

2 - Legal Purposes

Cadastral systems reinforce the rule of law by maintaining an exhaustive and up-to-date registry of land ownership, boundaries, and valuations(Alemie, Bennett and Zevenbergen, 2015). This robust information infrastructure not only clarifies property rights, which are essential to both legal and economic frameworks, but enhances transparency and minimises land-related disputes. As a result, cadastral systems contribute to stability by providing clear guidelines for land transactions, strengthening property rights, encouraging investment, and promoting socio-economic balance within a given jurisdiction (Alemie, Bennett and Zevenbergen, 2015; Ercan, 2021; Golob and Lisec, 2022).

3 - Assist in managing land and land use, enabling sustainable development and environmental protection.

A cadastre provides a comprehensive and accurate record of land ownership, boundaries, and usage record. This information helps make informed decisions about land use, such as zoning and planning. It also monitors land use change and identifies areas at risk of environmental degradation. By helping to manage land more effectively, cadastres can contribute to sustainable development and environmental protection.

4 - Land Management for Sustainable Development and Environmental Conservation

Cadastral systems play a pivotal role in land management and land use, promoting sustainable development and environmental conservation. These systems provide comprehensive data on land ownership, boundaries, and functions, enabling governments and stakeholders to make informed decisions about land policy, taxation, and zoning (Enemark and Rajabifard, 2009). As a result, an effective cadastre promotes economic growth, equitable land access, and natural preservation. This all-encompassing land information framework is critical for safeguarding land tenure, streamlining land markets, and backing development initiatives in urban and rural settings (Williamson, 2010).

Recent discourse within the FIG Professional Surveyors network has underscored the linkage between land management and climate change, emphasising the critical role of land as both a potential contributor to and mitigator of global warming (Augustinus, 2023; Jones, 2023; Stanley, 2023).

In summary, cadastral systems serve multifaceted roles, encompassing fiscal, legal, and environmental objectives. They provide the foundational structure for governments and stakeholders to build effective land management strategies and enforce the rule of more sustainable development. Through their comprehensive data on land ownership, usage, and value, cadastral systems are indispensable tools for economic growth, social equity, and environmental conservation, underlining their critical importance in contemporary policy-making and governance.

Cadastral Priorities - Accuracy or Ownership

In Less Economically Developed Countries (LEDCs), communities often have limited resources and need formal recognition of their land rights to access broader socio-economic opportunities, such as credit and state services. These communities often collaborate with international organisations like the World Bank to develop cadastres, either at the local or national level (Alemie, Bennett and

Zevenbergen, 2015). This section examines the principles that have emerged from these collaborations to evaluate their applicability in low-resource settings, particularly within local communities in the United Kingdom.

One of the critical frameworks to emerge in this context is fit-for-purpose land administration (FLPA). According to Enemark (Enemark, 2014), this framework posits that an effective land administration system should be flexible, participatory, affordable, reliable, inclusive, upgradeable, and attainable. Case studies, such as Rwanda's Land Registration programme, have demonstrated the efficacy of these principles. This programme utilised aerial photos for mapping, contributing to enhanced women's land rights and more effective property tax collection (Deininger, 2010).

Parallel to this, innovative land tools like the Social Tenure Domain Model (STDM) (Global Land Tools Network, 2023) have come into play. These tools extend upon existing frameworks like the Land Administration Domain Model (LADM) (ISO, 2012) and incorporate methods to represent various land tenure relationships (Griffith-Charles, 2011; Zevenbergen *et al.*, 2013; Rahmatizadeh *et al.*, 2018).

Such innovations occur in numerous, wide-ranging projects. For example, these tools have been employed in Kenya to secure informal settlements and monitor public lands (Koeva *et al.*, 2020). However, gaps and challenges exist in implementing these tools, including professional attitudes toward imprecision and concerns about data access among community members (Griffith-Charles, 2011; Griffith-Charles *et al.*, 2015).

The FLPA approach counters these challenges by emphasising a spatial, legal, and institutional framework designed to be affordable and reliable while recognising multiple forms of legitimate land tenure.

In contrast, conventional cadastral systems often prioritise high levels of accuracy, rigid boundaries, and stringent standards (Enemark, 2014, 2015). The FLPA framework proposes a more dynamic and adaptive approach. It aims to evolve, accommodating varying levels of accuracy and diverse tenure relationships rather than rigidly adhering to technical standards or colonial legacies that may not correspond to the lived realities and needs of the community.

In summary, the FLPA framework and its associated tools present a viable and adaptive approach for land administration, particularly in resource-limited settings. It contrasts conventional cadastral systems by prioritising flexibility, inclusivity, and adaptability, thereby making it a potential candidate for implementation in local communities within the UK that face similar resource constraints.

2.2 The UK Cadastre

This section will provide a brief overview of the UK cadastral system, explaining its history, what system it has, why it remains, and why there has yet to be much push for one. Ultimately, the section aims to explain why there is no desire for a cadastre at the higher levels of power—the government or "The Establishment" (Jones, 2015)—and why it is necessary for a grassroots, community-level approach instead.

UK Geography

The UK is an archipelago off the northwest coast of Europe. It consists of Great Britain (England, Scotland, and Wales), six counties in northeast Ireland, and numerous smaller islands. Agriculture accounts for 63.1% of the land area, with 20% for arable crops and the remainder for grassland. Urban areas comprise 8.7%, and forestry, open land, and water are 20% (Department for Levelling Up, Housing & Communities, 2022). The population is 67 million, 84% in England, 8% in Scotland, 5% in Wales, and 3% in Northern Ireland (Office for National Statistics, 2022).

Current UK system:

The UK lacks a cadastre but has a system of compulsory land registration without a central record of precise boundary locations. The government does not guarantee private boundaries, and Ordnance Survey (OS) is the official mapping agency. Originating from military needs in 1745, the OS evolved to include civilian mapping and became a self-funding government agency by 1990 (Ordnance Survey, 2023).

There is no cadastre in the UK, so no agencies are responsible for cadastral surveying. The official mapping agency is the Ordnance Survey, an executive agency responsible to various government bodies. While there has been a gradual move toward compulsory land registration in the UK, some significant landowners have not yet registered.

Registration is Mandatory

Compulsory land registration in England started in 1926, following the 1925 legal reforms that converted copyholds into freeholds, necessitating a national register (Land Registry, 2019). Land registration is among three regional bodies: the Land Registry for England and Wales, Land & Property Services in Northern Ireland, and Registers of Scotland.

In the beginning, compulsory registration only applied to certain areas, but since 1990, it has covered the entire country. About 83% of titles are registered (Shrubsole, 2019a). Some large landlords, including government bodies, parts of the Church of England, and aristocratic private owners, have not needed to register because they have not encountered any triggering events for registration. Being "immortal" (such as trusts and institutions), these entities are not subject to transfer upon death, which has not been a trigger for registration either. While there has been a gradual move toward compulsory land registration in the UK, some significant landowners still need to register. The efforts to register unrecorded land continue, and the overall system provides advantages to landowners and promotes transparency and accountability.

Land registration documents in the UK show boundary features, but boundaries have no particular legal definition, being lines of minimal width. Under the general boundaries system, a property's boundary is shown about physical features on the ground but is not defined by law (Williamson, 1985). The Land Registry can identify features but can only precisely determine where a boundary lies, as it is typically related to physical boundary features with a substantive legal width (Grover, 2008).

How the System Operates

The original purpose of compulsory land registration in the UK was to prove land title after the feudal system ended. Its primary function is to support the efficient land market by facilitating low-cost, reliable transfers. The central land register has reduced costs and improved reliability in property transfers, with a government guarantee of accuracy (Toms and Lewis, 1974).

Compulsory registration shifted the law of adverse possession, providing something closer to absolute ownership and extinguishing or losing specific remote claims. It also modified the principle of adverse possession, allowing owners to contest a claim, gradually causing this concept to disappear from law (Bullard, 2003).

The Rural Land Register enables subsidy payments to farmers under the EU's Common Agricultural Policy (Department for Environment, Food and Rural Affairs, 2014). The Valuation Office Agency maintains two fiscal cadastres for residential and non-residential properties for tax purposes. Exact boundaries are optional, as taxes relate to market value, with slight variations in land area having little impact on property value.

An ‘Estate’, Not a ‘Parcel’

In the UK, land ownership is recorded as a set of rights, known as an estate, rather than as a specific parcel of land. This system differs from many other countries, which focus on the physical boundaries of land (Bullard, 2003; Grover, 2008). An estate can include a variety of rights, such as the right to use the land, the right to sell it, or the right to pass it on to heirs. Different people can hold these rights at the same time, or they can change over time. For example, a landowner might grant the right to farm their land to a tenant while retaining the right to sell the land.

In summary, the UK's approach to land ownership is unique and complex, focusing on people's various rights over the land rather than the physical land itself. It is a system that has evolved and is a distinctive feature of British land law.

3. Research Gap / The Idea

3.1 Is a Cadastre Needed in the UK?

This section will look at previous attempts at creating a Cadastre for the UK and whether or not one is required, concluding with a full explanation of why this topic is worthy of investigation.

Failed Attempts

There have been proposals for a general cadastre in the past, but they have met with strong political resistance (Toms and Lewis, 1974; Bullard, 2003).

One of the earliest proposals for a general cadastre was made in 1836 by R. K. Dawson, a Royal Engineers officer. Dawson was working on tithe surveys for the Tithe Commission at the time, and he saw the tithe surveys as the potential foundation for a cadastre (Kain, 1975; Grover, 2008). He believed a cadastre would reduce boundary disputes and streamline real estate transfers. However, his proposal met with strong political opposition, and the legislation was subsequently amended to prevent the tithe surveys from being used this way (Kain, 1975).

Another attempt to create a general cadastre was made in 1909. A radical Liberal government introduced a 20% tax on the increased value of land, which required the creation of a fiscal cadastre to identify land ownership (Douglas, 2011). However, the initiative faced such significant opposition that the tax was never collected, and the assembled cadastral data was never used. These two failed attempts illustrate the historical resistance to implementing a general cadastre in the UK.

The Case Against a UK Cadastre

The absence of a cadastre in the UK may be linked to a particular philosophy concerning how land rights are derived and legitimised in British society and the state's role in these processes. The current challenges of introducing a cadastre in the UK present a complex and costly landscape (Grover, 2008). Given the efficiency of the existing property market and the country's robust protection of property rights, the value that a cadastre might add seems limited and perhaps insufficient to justify the substantial expense (Gray and Gray, 2005). The UK's contemporary system already includes elements such as efficient land registration, transparent market practices, professional valuations, and

A well-regulated financial and legal sector. While a cadastre could complement these aspects, it is not considered necessary.

In summary, throughout history, there have been attempts and proposals to create a cadastre system in Britain, but they have met resistance and various challenges. The current infrastructure in the UK already offers strong support for the property market, and the perceived benefits of a cadastre may not outweigh its costs and complexities.

The Case in Support of a UK Cadastre

Guy Shrubsole is a writer and campaigner advocating for introducing a cadastre in the UK. A cadastre is a systematic record of land parcels and their boundaries. It is used in many countries to provide a clear and transparent record of land ownership (Shrubsole, 2019a). Shrubsole argues that the current UK land ownership system is secretive and opaque. This gives wealthy landowners too much power and influence. He cites the example of Peel Holdings, a company that owns 15,000 hectares of land and a £2.3bn property portfolio. Peel's extensive interests include the Manchester Ship Canal, Trafford Centre shopping complex, MediaCityUK site in Salford, airports, fracking, and retail. The company's complex corporate structure, involving over 300 separately registered UK entities, further complicates attempts to trace its holdings and operations (Shrubsole, 2019a).

Shrubsole believes a cadastre would help break wealthy landowners' power and give ordinary people more control over their land. He argues that a cadastre would also improve decision-making about land use. For example, a cadastre could be used to identify areas of land that are currently underused or underutilised. This information could then promote more sustainable land use practices. The UK government has not yet committed to introducing a cadastre. However, Shrubsole and other campaigners are continuing to press for change. They believe a cadastre is essential for a more democratic and sustainable land use system in the UK.

Shrubsole's joint attempts to map the land with data scientist Anna Powell-Smith have had limited results because whilst the addresses of company ownership are available, it does not include the extent to the land. Nonetheless, their attempts, accessible at https://whoownsengland.org, illustrate that it is an essential matter because "who owns land gets to choose how it is used, which has big implications for almost everything. Where we build our homes, how we grow our food, how we protect ourselves from flooding, how much space we set aside for wildlife – all this is hugely affected by who owns land" (Powell-Smith, 2016).

3.2 The Research Gap

This section has outlined the key features of the UK's unique land registration approach, drawing on critical scholars' work. Grover has described the historical development of the UK's system, arguing that it is sufficient for most conventional real estate transactions. Shrubsole, on the other hand, has argued that the lack of transparency in the UK's system perpetuates inequality, particularly by giving wealthy landowners undue influence over environmental policy.

The absence of an open land ownership registry in the UK has been highlighted, which has motivated the present research. However, given the lack of political will and the vested interests of influential stakeholders, the prospects for establishing an open cadastre are limited. This suggests that any effort to create such a system must be at the grassroots or community level.

The following chapter of this dissertation will define the concept of a "community cadastre" and explore identified gaps in the open geospatial data that is available.

4. Research Objective 1: Define

Research Objective 1:
Partner with 'Green Meadows' to define a 'community cadastre' and its key components.

This chapter will discuss the first research objective, looking at how to define a ‘community cadastre’ and what it should contain.

4.1 Methodology

To meet the first objective of defining the concept of a community cadastre, the research methodology involves reviewing the FIG definition and discussing the project's requirements with the specific community. This will ensure the community cadastre is relevant for their situation and needs.

The International Federation of Surveyors (FIG) definition of a cadastre provides a foundational understanding (FIG, 1995), emphasising three main elements:

- Parcel-based information
- Record of interests in land
- Value of the land along with its improvements.

However, this definition also allows for substantial flexibility for community-specific adaptations. The specifics of what constitutes a "community cadastre" should ideally be defined by the community based on its specific needs. For example, data protection requirements may vary depending on the jurisdiction's laws. Similarly, the community's cultural norms and expectations will play a pivotal role in determining whether the cadastre should be openly accessible and to what extent.

Therefore, by working closely with a local community, we aim to outline a framework for a community cadastre that is both compliant with legal requirements and sensitive to local conditions and preferences.

The following sections comprehensively examine the selected community, The Meadows, and delineate its specific requirements and expectations from a "community cadastre." Additionally, the importance of building retrofitting as a crucial measure for climate change mitigation is discussed in depth.

4.2 Findings

A Case Study: The Meadows, Nottingham

The Meadows is a historically significant area in Nottingham that was initially a marshland (See Figure 2). In 1845, the Nottingham Inclosure Act was passed (UK Parliament, 1845), transforming open fields into privately owned land. This helped contribute to the expansion and industrialisation of Nottingham, which was overcrowded at the time (Zadik, 2013). The Act also uniquely allocated 130 acres of land for public recreation, creating numerous parks and recreational spaces within the city.

In the 1870s, The Meadows underwent a significant transformation. A building boom occurred, and vital infrastructure, including flood defences, was established. Trams were introduced, and public transport expanded, making the area more accessible (Lomax, 2013). Leisure activities such as rowing, tennis, cricket, and angling became associated with The Meadows, enhancing its reputation as a hub for recreation. Additionally, commemorative projects, such as the opening of the Victoria Embankment in 1901 and the creation of the Memorial Gardens for the First World War, contributed to The Meadows area's unique character and historical richness.

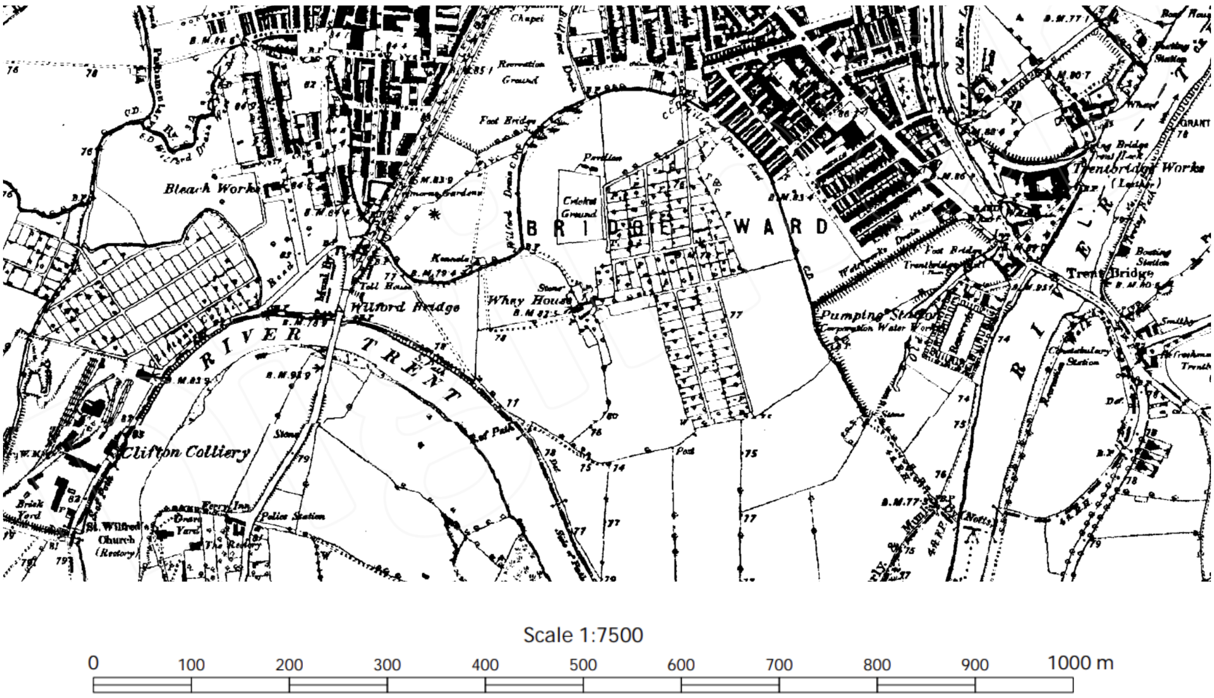


Figure 2: Historical Map of The Meadows, Nottingham. Source: EDINA Digimaps

Nottingham City Council classifies the Old Meadows as a Conservation Area (Nottingham City Council, 2020). The conservation document notes some critical characteristics of the Old Meadows area:

- A variety of red brick terraced houses forming a dense urban grid.
- Substantial green space, including the riverside promenade of Victoria Embankment and The Memorial Gardens, covering about half the designated area.
- Historic and varied architectural styles, including large suburban villas and historic industrial structures.

The conservation area designation intends to protect the character and heritage of The Meadows.

Working with 'Green Meadows'

In 2020, the Nottingham Energy Partnership and Mozes Community Energy Group secured a Lottery Grant for the Green Meadows project to help the Meadows community in Nottingham address climate change. The five-year initiative (Nottingham Climate Assembly, 2021), which began in March 2021, consists of 16 initiatives across three themes:

- **Homes and Buildings**: This theme focuses on mapping the area's carbon footprint and offering energy advice to residents.
- **Living for the Future**: This theme includes workshops and resources to help people combat climate change and prepare for weather changes.
- **Measure for Success**: This theme aims to gauge the project's effectiveness and facilitate learning and collaboration with similar nationwide projects.

In discussions with Green Meadows (Green Meadows, 2023), the staff said they had had trouble understanding what land is in public or private ownership. They also said they would like to create a map of house types to identify general energy-saving retrofitting measures that could apply to each type of house - one of Green Meadows' current priorities.

Green Meadows collaborates with residents in the Meadows area of Nottingham to collectively address climate change through practical learning, training, and advice. One initiative focuses on developing retrofitting guides for approximately 20 housing styles in the area. These guides will identify specific opportunities for each house design to increase energy efficiency.

Initially, a manual survey was conducted, and a map was created to identify the locations of different house styles (See Figure 3). This study explores the application of machine learning techniques to remote sensing imagery, replacing manual survey methods. The methodology aims to accurately classify house styles based on visible features from a bird's eye perspective, including size, shape, materials, and distinguishing characteristics. The original map will be used solely for personal reference to verify the accuracy of the machine-learning classification output.

Determining the ownership of a home or land is also crucial for Green Meadows' work. Allowing for better communication among stakeholders and facilitating discussions on retrofitting measures. Lack of transparency regarding local land ownership hinders the organisation's progress and planning efforts. Therefore, a "community cadastre" concept would be a tool to provide the necessary information that Green Meadows requires.

The Meadows in Nottingham is an appropriate locale for this study due to the diversity of its housing stock. The charity Green Meadows, which the National Lottery funds, is notably active in the area. They have been instrumental in examining the buildings and homes of The Meadows through an archetype study and conducting outreach activities in the form of training and capacity-building to upskill residents. These activities serve to inform residents on how to maintain their homes adequately.

The archetype study by Green Meadows entails surveying around 100 homes within The Meadows. From this, individual "Future-Fit Home Plans" are formulated, offering homeowners tailored advice on enhancing the energy efficiency of their residences. Green Meadows plans to utilise this data and a mapped record of diverse housing archetypes to cultivate a comprehensive understanding of energy efficiency at the community level.

Following discussions with Green Meadows, it was indicated that mapping the geographical distribution of different house archetypes would be beneficial. Consequently, this element has been incorporated into the formulation of the community cadastre.

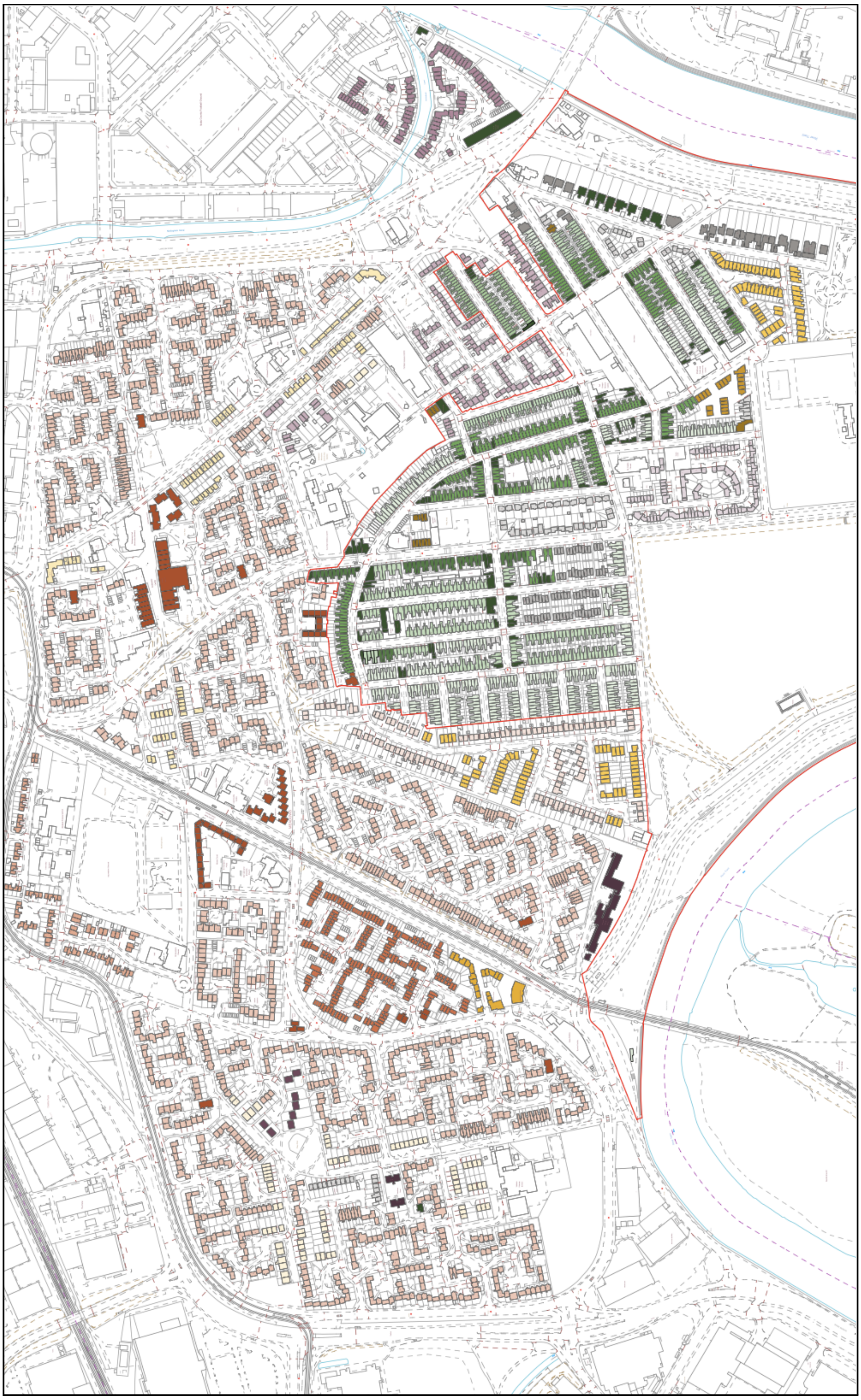


Figure 3: Green Meadow's Archetype Map. Victorian (green), Inter-war (grey), Post-war (pink), Modern (purple), 21st century (yellow). Different hues of the same colour represent different archetypes. The red line is the conservation area boundary. Source: Green Meadows.

The Archetype Map is accompanied by an Archetype Key, which elaborates on the five construction periods by further segmenting them into 63 distinct 'archetypes' in Table 1. These archetypes are more specific and are categorised based on the unique features of the houses.

Building Age Classification	Green Meadows Archetypes
Victorian	VT1a, VT1b, VT2a, VT2b, VT2c, VT2d, VT2e, VT3a, VT5a, VT5b, VT6a, VT6b, VT_NS
Inter-war	IW1a, IW1b, IW2, IW3
Post-war	PW1, PW2a, PW2b, PW2b, PW2c, PW2d, PW2d, PW2e, PW2f, PW3a, PW3b, PW4a, PW4b, PW4c, PW4d, PW5
Modern	M1a, M1b, M2a, M2b, M3a, M3b, M4, M5
21st century	TF1a, TF1b, TF2a, TF2b, TF3a, TF3b, TF3d, TF3e, TF4a, TF4b, TF4c, TF,5a, TF5b, TF5c, TF6a, TF6b, TF6c, TF6d, TF7a, TF7b

Table 1*: Green Meadows' Building Age Classification broken down into further house archetypes. Source: Green Meadows.*

A comprehensive list of these archetypes is provided in Table 1. From here on, the terms building age and construction period will be used interchangeably.

Category	Building Feature
Construction	Solid brick, System built, Cavity
Frontage	One storey, Two storey, Three storey, Purpose built flats
Roof	RIR (Room in Roof), As-built dormer, Half roof dormer, Monopitch, Hipped, Gable, Catslide, Flat
Rear addition	Single Storey, Two Storey, Flat roof cavity wall
Bay	Single storey, Two storey
Other	Alley/underpass, Sloped roof above stairs, GF garage

Table 2: Green Meadows' Building Features to distinguish different house archetypes. Source: Green Meadows.

Based on the discussion. The Green Meadows' community cadastre has three criteria:

1. Identifying the housing archetype of each house, and
2. Determining land ownership (public or private),
3. Viewable on a map.

Figure 4, shows a quick mockup of how such a system might look.

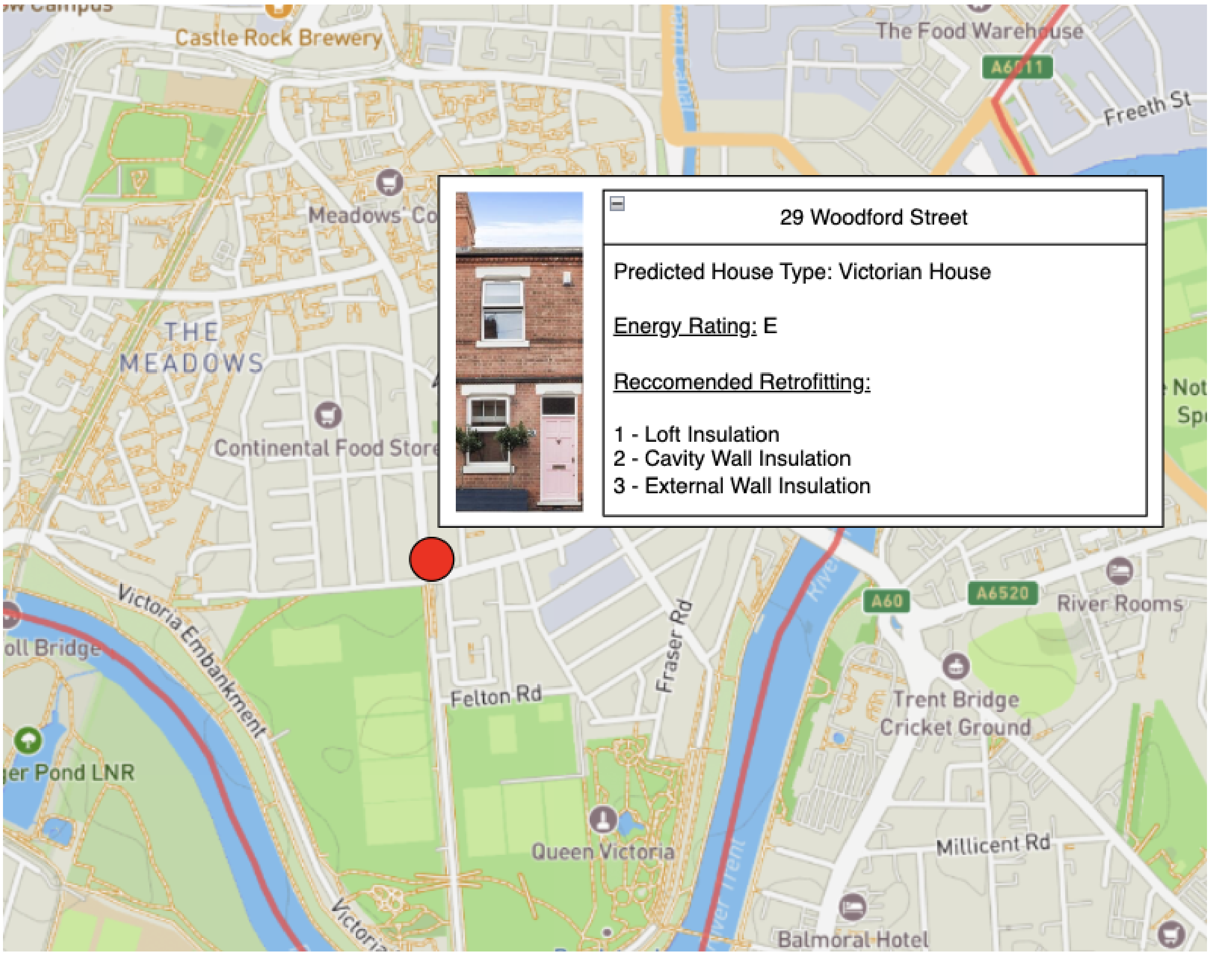


Figure 4: Mock-up of potential mapping product. Source: Author's own.

Retrofitting Buildings to Reduce Climate Change

The United Kingdom's housing stock is among the oldest and least energy-efficient in Europe, posing a significant challenge to reducing carbon emissions. Over half (57%) of the 23.3 million homes in England were built before 1965, when only basic thermal insulation was required by Building Regulations (De Urquia, 2022). As a result, domestic buildings in the UK account for 29% of energy consumption and a significant portion of global carbon emissions. Many older properties were constructed before Part L of the Building Regulations, leading to much lower or non-existent energy efficiency standards (Ministry of Public Building and Works, 1965; Ji, Lee and Swan, 2019). This older housing stock has led to domestic buildings in the UK accounting for 29% of energy usage as estimated in 2015 (BEIS, 2017), contributing a significant percentage of the UK's carbon emissions (Technology Strategy Board, 2013).

Retrofitting is Essential.

Retrofitting existing housing stock to improve thermal and energy performance could significantly impact even though modern building techniques can create energy-efficient new homes. The importance of schemes and incentives aimed at improving energy performance and lowering national carbon emissions is highlighted by this, as noted by the Climate Change Committee. (Ürge-Vorsatz and Metz, 2009; Nejat *et al.*, 2015; Regnier *et al.*, 2018). This emphasises the importance of schemes and incentives aimed at improving energy performance and lowering national carbon emissions, aligning with observations made by The Commission for Climate Change (Commission for Climate Change, 2019).

According to the UK's Department of Energy and Climate Change (DECC), in 2009, the residential sector accounted for 27% of final-user emissions, and the UK's housing stock is particularly old, with a replacement rate of less than 1% per year (Technology Strategy Board, 2013). This means that over 70% of the dwellings in use in 2050 have already been built, highlighting a significant need for eco-refurbishment to improve energy efficiency and provide comfortable living environments (Mohammadpourkarbasi and Sharples, 2013). Customised retrofitting is essential and provides considerable potential for energy conservation and sustainability benefits (Galán, 2015). The Technology Strategy Board estimates that the retrofit market could be worth approximately £200 billion over the next 40 years in the UK (Calcutt, 2007).

The Climate Change Act (UK Parliament, 2008), introduced in 2008, set a legally binding target to reduce greenhouse gas emissions to at least 80% below 1990 levels by 2050. As UK domestic homes account for almost 30% of total national emissions, of which about two-thirds is for space heating (BEIS, 2017), the retrofit of the existing housing stock is unavoidable in the strategy to meet these national targets. Retrofitting represents a vital component in both achieving energy efficiency and fulfilling the country's commitments to sustainability and emission reduction (Department for Business Enterprise & Regulatory Reform, 2007; HM Government, 2013; Technology Strategy Board, 2013). The retrofit of the existing housing stock is essential to meet the UK's legally binding target to reduce greenhouse gas emissions.

Many Homes Have a Similar Design

Figure 5: UK Home Designs Over Time. Source: Build Review.

Many homes in the UK share similar designs due to how they were constructed (See Figure 5), with large plots of land developed by a single developer often resulting in a limited range of standard designs. This uniformity in housing designs can be advantageous in retrofitting for energy efficiency, as the same measures can be applied to multiple houses within an area. A notable example of this can be seen in the UK's terraced houses built between 1837 and 1901 (Historic England, 2019). These homes often feature a standard 'two up, two down' layout with subsequent kitchen and toilet additions, and this commonality in design allows for consistent retrofitting approaches to be taken across numerous properties of the same type. Therefore, the shared design characteristics among many homes present a valuable opportunity for implementing widespread energy-saving measures through a common retrofitting strategy.

Importance of a Cadastre for Climate Change

Despite the limited amount of explicit discussion in the literature on cadastres and their links to climate change, this was highlighted at the FIG Working Week Conference Orlando 2023. Keynote speakers emphasised the importance of land surveyors in recognising that climate change is explicitly linked to climate change (Augustinus, 2023; Jones, 2023; Stanley, 2023). The conference, which is the annual conference for the surveying profession, had numerous talks directly addressing the issue of land and climate change.

A search through the history of papers from the conference returns only a handful of published papers on how cadastres can be used in tackling climate change.

In "Cadastre 2014: A Beacon in Turbulent Times," Pal van der Molen discusses how sustainable monitoring systems, land management systems, and land administration systems can serve as a basis for climate change mitigation and adaptation as well as prevention and management of natural disasters (van der Molen, 2014). However, this paper does not explicitly explain the advantages, disadvantages, or opportunities of how a cadastre full of land knowledge and information could be applied to the climate change crisis we find ourselves in.

The second FIG conference paper to explicitly relate cadastres and climate change is "Climate Change and the Cadastral Surveyor" by Make Stapleton (Stapleton, 2010). This paper summarises the key issues that surveyors are likely to come up against, such as natural disasters such as floods and droughts, which can present significant challenges to property rights and boundaries, affecting the responsibilities of cadastral surveyors who measure these extents. The paper also investigates current land boundary definitions, land tenure, and easement options and contemplates necessary changes in legal definitions and policies. It further explores the existing roles of surveyors in sustainability and climate change and suggests potential expansions and new opportunities in these roles. Stapleton underlines the intricate relationship between climate change and land management, pointing to the need for adaptive strategies in both legal frameworks and professional practices. However, the paper does not mention local efforts or community-driven work to address climate change.

Guy Shrubsole's blog post on his recent "Call for Evidence" to the House of Lords Select Committee on Land Use in England provides an up-to-date and well-informed view on the potential use of an open UK cadastre or open land ownership in the UK (‘Written Evidence (LUE0027) to the House of Lords’, 2022). The submission identifies three critical areas:

- The need for significant changes to land use to meet climate and environmental commitments.
- The challenges associated with the concentrated ownership of land.
- The need for a strategic approach to land use planning.

The document advocates ambitious goals, including doubling England's woodland cover and restoring 100% upland peat by 2045. It criticises the current government targets as inadequate and emphasises the concentration of land ownership, where 1% of the population owns half of the land. The submission provides a valuable overview of the challenges and opportunities facing land use in the UK. It is a call to action for the government to take bold steps to address these challenges and build a more sustainable future for the country.

To Address Climate Change, Action is Required at All Levels

Shrubsole's paper was being considered at the UK national level. However, the imperative for climate change action transcends various levels of society, from national and local governments to individual communities. Existing literature emphasises the importance of multi-level engagement in climate action, recognising that cohesive efforts across these different tiers are essential for effectively combating the challenges of climate change. National and local interventions are crucial for establishing regulatory frameworks and targeted initiatives (Sugiyama and Takeuchi, 2008; Damsø, Kjær and Christensen, 2016; Moloney, Fünfgeld and Granberg, 2018; De Urquia, 2022). Simultaneously, action at the community level is vital for grassroots innovation, fostering local resilience, and engendering public participation in sustainability efforts (Peters, Fudge and Sinclair, 2010; Seyfang, 2010; Bardsley *et al.*, 2019). These collective actions affirm the need for a holistic approach where all levels of society work collaboratively, reflecting a shared responsibility in addressing the global issue of climate change.

Conclusion

Green Meadows is a community-based organisation that operates in the Meadows area. The Meadows area is a good example of the evolution of UK housing over the past 200 years, from constructing terraced houses to demolition in the 1970s to modern infill development. Part of the Meadows area is a conservation area, which is home to Victorian terraced housing. There is also a large council housing estate developed in the 1970s, as well as pockets of modern development. Green Meadows is actively engaged in identifying retrofitting measures for specific housing designs within the community to improve overall energy efficiency. This includes measures such as insulation, double glazing, and solar panels. By retrofitting homes, Green Meadows is helping to reduce energy consumption and emissions, improve air quality, and create a more sustainable community.

Next, Chapter 5 will discuss open data and open-source software.

5. Research Objective 2: Identify

Research Objective 2:
Identify open geospatial data and software for cadastre creation.

This chapter will examine the impact of the government's open data initiative, which has led to increased availability of public data in the UK. It will also explain how the UK's membership of the European Union has facilitated the creation of a new geospatial dataset, the INSPIRE Index Land Polygons, which has made this data more accessible. It will also discuss the open-source software chosen for the study.

5.1 Methodology

To identify the key sources of geospatial data for predicting building attributes in the Meadows area, a comprehensive literature review and discussions with colleagues at the University of Nottingham were conducted. The following are some of the key geospatial data sources that could be used for this project:

- **Government agencies:** The UK Government provides various geospatial data, including land use maps, aerial imagery, and topography data. This data is often of high quality and accuracy, and it is often available at no cost. However, it may not be available for all areas, and it may not be updated as frequently as data from other sources.
- **Private companies:** Several private companies collect and sell geospatial data. These data can be of high quality and accuracy, and they are available for a fee. Given the context of this work focusing on open-source data and software, this was not deemed an option.
- **Academic institutions:** Academic institutions, such as the University of Nottingham, often have the possibility of accessing data on a research licence. However, given that this academic partnership may not be possible for all communities, this was not considered an option to further explore at this stage.
- **OpenStreetMap (OSM)**: OSM is a collaborative project to create a free and open world map. Volunteers contribute the data on OSM, and it is constantly being updated. The data quality can vary, but it can be a valuable resource for projects that require up-to-date information.

Data from government agencies (Land Registry and Local Authorities) and OpenStreetMap were chosen for this project, given their availability for the area of interest, The Meadows, and the ability for other community groups also to be able to access this type of data for their own locality.

Open Data & Open Source Approach

The Open Data (OD) movement is changing how we use and access geospatial information. OD allows anyone to use important data that is easy to read by both humans and machines, with very few restrictions on how it can be used again in the future (Johnson *et al.*, 2017; Coetzee *et al.*, 2020). For example, the Copernicus programme run by the European Union freely provides a large amount of satellite data.

Many government portals, such as data.gov.uk and europeandataportal.eu, have been set up to share this kind of open data. They aim to provide reliable information that can be used to create both economic and social benefits (Charalabidis, Loukis and Alexopoulos, 2014; Viscusi, Castelli and Batini, 2014). A European report even suggests that using OD can lead to more users, more types of business applications, and positive effects on economic growth and social challenges (European Commission. Directorate General for the Information Society and Media. *et al.*, 2015).

In my research, I plan to use OD to predict building ages. Using data from various sources, including the UK government, Nottingham City OpenStreetMap. I believe that this research will help Green Meadows in getting a better understanding of what housing stock exists in the local community and, therefore, which retrofitting measures will be needed.

Datasets

The goal is to merge data from existing open datasets to create a complete vector for each land parcel. A list of the desired polygon attributes is shown in Figure, along with Table 3, which shows details of the data sources used.

House Polygon
inspire_id
national_cadastral_reference
number_of_building_polygons
area_of_building _polygons
perimeter_of_building _polygons
number_of_ukpn_points
shared_perimeter
non-shared_perimeter

Figure 6: Potential attributes for each polygon.

These attributes will be used to create a comprehensive dataset of land parcels in the study area. This dataset will be used to investigate whether missing building attributes can be found through classification.

	Type of Data	Coverage and Accessibility	Purpose of Data	Source of Data	License	File type/size/extent downloaded
a	Nottingham Wards (polygon),	Nottingham, free data	Identification of Meadows area boundary	Nottingham City Council	Open Government Licence	GeoJSON / 528 KB / Nottingham City Council area
b	HM Land Registry's INSPIRE Index Polygons	UK, free data	Position of registered freehold properties in England and Wales	GOV.UK	Open Government Licence	GML / 100,417 KB / Nottingham City Council area
c	Nottingham City Council Owned Land	Nottingham, free data	Identification of council-owned land	Nottingham City Council	Open Government Licence	GeoJSON / 37,322 KB / Nottingham City Council area
d	Residential Building Polygons	Global, free data	Identification of building by house number	Open Street Map	Open Data Commons Open Database License	.pbf / 873 KB / Nottingham /City Council Area

Table 3: Data types, sources, and licensing.

Source 1: OpenStreetMap - Building Footprint Polygons

OpenStreetMap (OSM) (OpenStreetMap contributors, 2017) was founded in the UK in 2004 by Steve Coast, who was dissatisfied with the restrictive copyright of the UK's National Mapping Agency, Ordnance Survey. Initially started as a small project using consumer-grade GPS to map central

London, it quickly gained momentum and had 1,000 registered users within 16 months. By March 2009, it had 100,000 participants. The geospatial data is entirely free and released under a permissive license, which allows anyone to use, modify, and distribute it. However, geospatial data coverage varies by region, being particularly robust in developed areas. Despite initial concerns about the quality of geospatial data contributed by non-professionals, studies such as one by gi have confirmed its reliability, even finding it superior to commercial geospatial data providers in some aspects (e.g. currency and accuracy of geospatial data). OSM's success has led some UK local authorities to reconsider their reliance on National Mapping Agency geospatial data, especially with the advent of alternative large-scale geospatial data from the UK Map.

Source 2: Nottingham City Council - Nottingham City Council Owned Land

UK Government aims to maximise the utility of public sector geospatial data, seeing it as a tool for enhancing public services and benefiting the economy and society (Geospatial Commission, 2023). In 2020, the National Data Strategy was launched to "unlock the power of data", including that generated by the public sector. This links to the Local Government Transparency Code (Department for Levelling Up, Housing and Communities, 2015), which sets out the minimum requirements for local authorities to publish open data for re-use in a timely manner. Nottingham City Council is the relevant local authority, and it has a data portal, opendatanottingham.org.uk, which includes downloadable data on CCTV, tram, pollution, businesses, congestion, buses, the historical environment and many more topics. Moat relevant here is the polygons for the land and property which the council owns. This is useful because if we can distinguish what is owned by the council, then other pieces of land will be owned as freehold land, which is owned privately. Green Meadows wanted to know which land is public (council-owned) and which is private because this allows them to contact the appropriate land owner to discuss potential retrofitting activities or to organise broader community initiatives.

Source 3: INSPIRE Index Land Polygons - Land Boundaries

The INSPIRE Directive (Directive 2007/2/EC) was formulated by the European Parliament and the Council to establish an Infrastructure for Spatial Information in the European Community (European Parliament and Council, 2007). It requires each member state to incorporate this directive into their own national legislation. In the United Kingdom, this was enacted through The INSPIRE Regulations 2009 (UK Parliament, 2018b). The primary aim is to facilitate environmental spatial information sharing among public sector organisations and enhance public access to such information across Europe (Ordnance Survey, 2023).

INSPIRE imposes four principal obligations on public authorities in member states:

1. Standardise metadata to ease the discovery of spatial data.
2. Facilitate straightforward online access to spatial data.
3. Harmonise spatial data utilising common data models and specifications.
4. Develop data-sharing protocols among public bodies.

The HM Land Registry in the UK released its Inspire Polygons, which became accessible under the Open Government Licence as of July 2020. These polygons delineate the location and approximate extent of registered freehold properties and are derived from the Ordnance Survey's MasterMap. As discussed above, when combined with data on publicly-owned land, they can provide useful insights into land ownership within local communities. The data covers 23 million titles across England, Wales, and Scotland (UK Government, 2020).

Initially, there were issues with this dataset. Cho et al. (Cho and Crompvoets, 2019) identified that legal and policy issues have acted as barriers to the prompt execution of the directive. Before this, at the drafting stage, the directive was contentious with the UK due to the Ordnance Survey's status as a government-owned entity. British policymakers negotiated amendments permitting data charging, considered a UK victory. This was a focal point in a broader debate over the commercialisation of public-sector data (The Guardian, 2006).

When the UK departed from the EU in 2020, The INSPIRE (Amendment) (EU Exit) Regulations 2018 (UK Parliament, 2018b) were enacted, and these amendments ensure the continued operability of the INSPIRE Regulations as retained EU law under the European Union (Withdrawal) Act 2018 (UK Parliament, 2018a). This legislative instrument modifies existing statutes to make them applicable post-Brexit. However, given that the UK is not a member of the EU, the UK Government has more ability to repeal a law if it is considered too costly or burdensome for local/national government.

Also, these polygons are limited, as they only provide the address and polygonal outline of the land, not its full extent - this is defined in the land title, and these full details are available from the Land Registry for £3 per title (HM Land Registry, 2023). Despite the challenges and uncertainties introduced by Brexit, the INSPIRE framework remains an essential tool for spatial information management in the UK. This project will make the most of the INSPIRE data that has been made available.

Open Source Software

Open-source geospatial software continues to become more of a viable alternative to closed-source solutions (Coetzee *et al.*, 2020). This development makes it increasingly difficult to justify the initial investment in expensive closed-source licences. The shift towards cloud computing further supports this trend; companies are transitioning from product-centric to service-based models, leading to a culture of openly sharing code. The main benefit is there is limited, or usually no cost to use this software. Because this project aims to create a product that communities can replicate at no cost, open source is the only software to be considered. Table 4 shows a summary of the software to be used.

Name	Usage
QGIS & MMQGIS Plugin (Minn, 2021; QGIS Development Team, 2023)	For initial mapping and handling of geospatial data. Merging geospatial data through spatial joins.
Jupyter Notebooks / Python (Chollet, 2018)	For machine learning (supervised and unsupervised) to classify buildings by building age.
Felt.com (Felt Team, 2023)	To render the final web map.

Table 4: Summary of Software and its usage.

QGIS / MMQGIS Plugin

QGIS will be used in this project because it's the most widely used Geographical Information System. It has a wide user support network and has tools rivalling the propriety software. MMQGIS (Minn, 2021) is a collection of Python plugins designed to enhance the functionality of QGIS. It offers a range of features for vector map layer manipulation, including CSV handling, geocoding, geometry conversion, and buffering. In this project, QGIS and MMQGIS will be used to:

- Collect and process geospatial data
- Perform spatial analysis
- Visualise the results of the analysis.

Jupyter Labs / Python

Python (Van Rossum and Drake, 2009) will be used for machine learning because of its simple syntax and a broad range of specialised libraries like TensorFlow and sci-kit-learn (Pedregosa *et al.*, 2012; Chollet, 2018). These languages will allow for machine learning of some of the most useful rapid development and have strong backing from the community. Its versatility is also evident in its compatibility with geospatial platforms, such as QGIS, which adds another utility layer to geospatial analytics (Sherman, 2008).

Felt.com

Felt (Felt Team, 2023) is an easy-to-use digital mapping tool introduced for versatile planning and navigation tasks, distinguishing itself through its collaborative features. It acts as a digital whiteboard that allows for a wide range of annotations, such as hand-drawn lines, text, and emojis, facilitating collective route planning and marking points of interest. While its user-friendly interface and regular updates make it appealing to various applications, it lacks offline functionality. It has a minimalistic base map, limiting its use to detailed, off-the-grid navigation (BikePacking.com, 2022).

Step-by-Step Process

Software Setup

The first step of the methodology involved setting up the requisite software environment. QGIS was installed along with its MMQGIS plugin, a necessary addition that equips QGIS with advanced geospatial analysis capabilities.

Data Preparation and Preprocessing

Data preparation began with downloading the required files and naming them in a structured format for easier identification later (e.g., 'a_boundary,' 'b_land'). Subsequently, preprocessing was carried out to ensure all datasets adhered to the British National Grid (BNG) projection, identified by the code 27700. Multiple polygon data layers — including wards and land polygons — were cleaned. Specifically, irrelevant wards such as "Meadows" were excluded, and extraneous columns in the attribute tables were removed. In the case of land polygons, spatial selection was used to focus only on polygons intersecting with the pre-selected ward. A size-based filtration was conducted to omit huge land polygons, using the largest single dwelling piece of land as a reference point. A threshold area of 1283 square meters was established for this purpose.

Layer	Attributes to keep	Attributes to create
a_meadows_boundary		Latitude Longitude
b_inspire_land	INSPIREID	land_polygon_area land_polygon_perimeter Land_polygon_vertices Latitude Longitude
c_nottingham_city_council_owned		
d_osm_houses	osm_way_id building other_tags	building_area building_perimeter Building_vertices Latitude Longitude

Table 5: Cleaning process for the different GIS layers.

Table 5 documents the process of cleaning the layers to ensure the layers had the correct attribute information. Table 6 shows the new columns to be created and the attribute calculator code to be entered into QGIS to calculate the correct information.

Attribute	Column Name	Calculation	Formatting notes for field calculator options
>>> Layer: b_land			
Land polygon area	land_area	$area	
Land polygon perimeter	land_perimeter	$perimeter	
Land polygon vertices	land_vertices	(num_points($geometry)-1)	
Land polygon longitude	land_longitude	x(centroid($geometry))	Output field type: decimal number (real) Output field length: 9 Precision : 4
Land polygon latitude	land_latitude	y(centroid($geometry))	Output field type: decimal number (real) Output field length: 9 Precision : 4
>>> Layer: d_buildings			
House footprint polygon area	house_area	$area	
House footprint polygon perimeter	house_perimeter	$perimeter	
House footprint polygon vertices -	house_vertices	(num_points($geometry)-1)	
House footprint polygon longitude	house_long	x(centroid($geometry))	Output field type: decimal number (real) Output field length: 9 Precision : 4
House footprint polygon latitude	house_latitude	y(centroid($geometry))	Output field type: decimal number (real) Output field length: 9 Precision : 4
Proportion of the built up property (%)	b2lpercent	((“house_area”/”land_area”)*100)	
Publicly or privately owned	private	IF(“ncc_owned” = ‘true’, ‘yes’, ‘no’)	
Building age	building_age	To be determined by ML	

Table 6: Calculations entered for the QGIS field calculator.

Data Segmentation

Data segmentation was executed primarily using OpenStreetMap-sourced building polygons. A selective extraction of only residential buildings was performed using field expression calculators. The building types considered were 'detached,' 'house,' 'residential,' 'semidetached_house,' and 'terrace.' Any features falling outside these categories were eliminated. Furthermore, abnormally large residences—defined as having an area greater than 244 square meters—were also omitted. Additionally, any garages identified within the layer were expunged.

Data Enrichment and Calculations

Data enrichment involved appending a 'Private' attribute to the building polygon layer. A spatial join between the land data and building polygons was performed using MMQGIS menu commands. The resulting layer was saved as 'cadastre.' Following this, the building-to-land ratio was calculated as a percentage and integrated into the attribute table. All changes were saved, and the enriched data layer was exported as a .csv file, designated 'cadastre_output.'

Manual Labeling

For additional granularity, manual labelling was employed on a subset of 280 house types. The archetypes were derived from a map provided by Green Meadows but could also be determined through in-person surveys or Google Streetview.

Significance of Enriched Dataset

The enriched dataset, featuring attributes like the building-to-land ratio, was considered integral for subsequent analyses and for addressing the research questions posed in this dissertation. Specifically, the building-to-land ratio offers a valuable metric for classification and achieving a nuanced understanding of land use patterns within the community.

5.2 Findings

Throughout the data processing workflow in QGIS, a meticulous manual review of polygons was conducted. Utilising the MMQGIS plugin, a spatial join operation was executed to merge the 'B land polygons' with 'house footprints'. The field operation selected for this task was the "largest proportion." This approach encountered complications due to layer alignment discrepancies. Specifically, the spatial join resulted in an output of 4,828 features from an initial 8,063 features.

Issue 1: Incomplete Polygon Inclusion

One significant issue arose concerning the inclusion of all relevant polygons (Figure 7). This problem was attributed to the varying degrees of accuracy across datasets. While the OpenStreetMap (OSM) polygons were aligned with the OSM map, they did not align with other data sources such as Google Satellite imagery, INSPIRE Land Index Polygons, or Nottingham City Council Owned Land polygons. Despite multiple attempts to rectify these alignment issues by 'snapping' the polygons to the land polygons, the time constraints of this master's project prevented a full resolution.

Issue 2: Challenges with Large Council-Owned Land Blocks

Another complexity arose due to large tracts of land owned by councils Figure 8. These expansive land areas precluded the possibility of associating a single piece of land with a single OSM house polygon, as depicted in the accompanying figure. Given that these large blocks are council-owned, they had not been individually input into OpenStreetMaps. This is a known limitation of Volunteered Geographic Information (VGI) data and will be explored further in the discussion section of this study.

Figure 7: New blocks of homes (in grey) were not included. Citation: Author's Own.

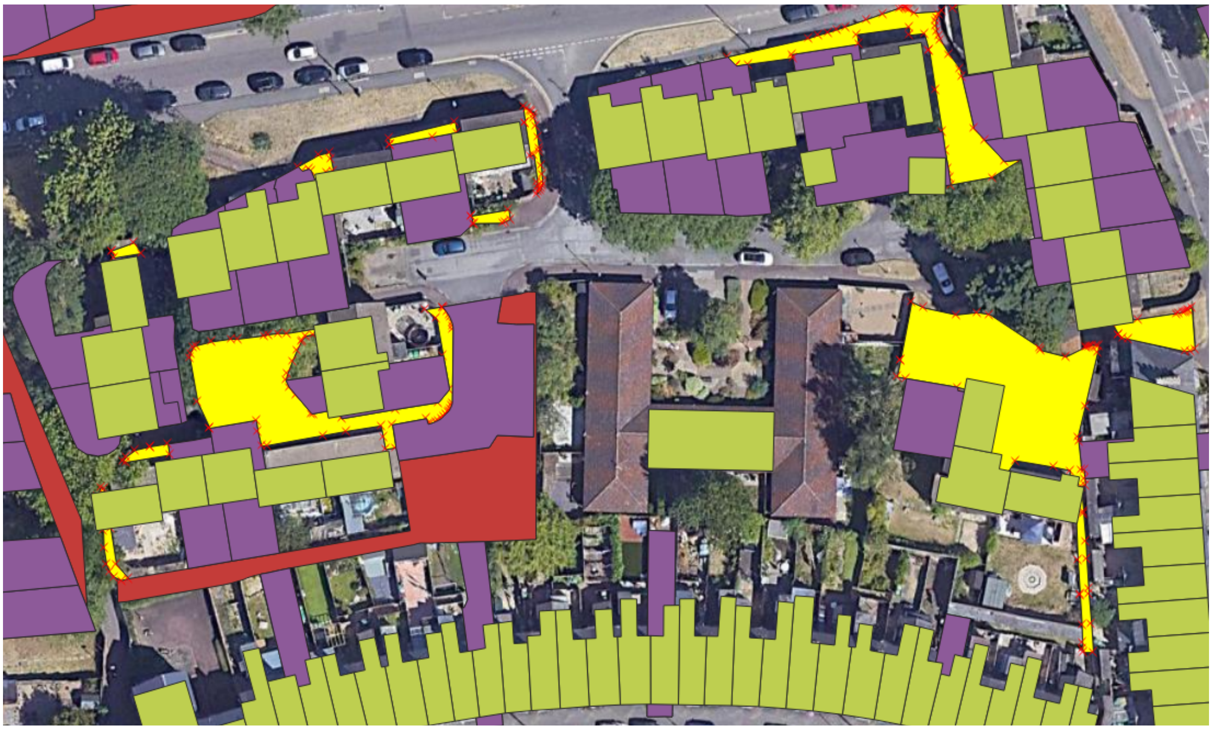


Figure 8: Large blocks of council land (in yellow). Citation: Author's Own.

In summary, this chapter discussed the second research objective of the study, which was to identify open geospatial data and software for cadastre creation. The chapter began by discussing the importance of open data and open-source software and how these can be used to create a more sustainable and equitable society. The chapter then discussed the different sources of open geospatial data, including government agencies, private companies, and academic institutions. The chapter also discussed the different types of open geospatial data, such as vector and raster data. It concluded by discussing the challenges and limitations of using open geospatial data and software.

6. Research Objective 3: Predict

Research Objective 3:
Identify gaps in geospatial data and compare machine learning classifiers to predict building age.

This chapter looks at Research Objective 3, which uses the collected data to predict the missing building age value. The study will focus on the supervised machine learning algorithms - Zero Rate Classifier, Decision Tree, Random Forest, Stacked Generalisation, and eXtreme Gradient Boosting (XGBoost). There will be a comparison of their performance before deciding which one to apply to the unseen dataset and then creating a confusion matrix for further discussion and insight.

6.1 Methodology

A brief literature review reveals various methods for predicting building attributes using geospatial data, as summarised in Table 7.

Citation	Study Location	ML Model	Open Source Data Used?
Nachtigall	Europe	Regression Models	Yes, Open Urban Morphology Data
Atwal et al., Scientific Reports, Vol 12, 2022	Fairfax County, VA; Mecklenburg County, NC; City of Boulder, CO	Supervised Learning	Yes, OpenStreetMap Data
Sturrock et al.	Botswana and Swaziland	Ensemble Machine Learning	Yes, OpenStreetMap Data

Table 7: Summary of building attribute prediction literature.

Beck et al. (Beck *et al.*, 2020) provided a comprehensive methodology for energy modelling of residential buildings in the UK using a blend of quantitative and qualitative metrics. However, their reliance on proprietary datasets such as OSGB MasterMap and AddressBase Plus made their approach less applicable to this study, which prioritises the use of open-source data for creating a community cadastre.

On the other hand, Sturrock et al. (Sturrock *et al.*, 2018) demonstrated the effective use of vectorial data sourced from OpenStreetMaps to classify 'sprayable' and 'not-sprayable' residential buildings in

Botswana and Swaziland for malaria prevention. The study is compelling for several reasons that make it a valuable model to consider for creating a community cadastre to build retrofitting measures. They utilised ensemble machine learning, applied a comprehensive set of building attributes, and subjected their model to rigorous validation.

In contrast, Atwal et al. (Atwal *et al.*, 2022) used the XGBoost algorithm to predict building construction years across the Netherlands, France, and Spain. While their methodology was robust, its larger geographical scale makes it less directly applicable to this research, which focuses on a small community level. It raises questions about the suitability of the XGBoost algorithm when applied to a more limited dataset.

In light of the above literature, this study will employ a range of supervised machine-learning methods to predict building age attributes. The algorithms to be used will be detailed in the following sections.

In summary, based on the above literature, I shall employ ensemble methods to predict the building age attribute. I shall use the following algorithms, Zero Rate Classifier, Decision Tree, Random Forest, Stacked Generalisation, and eXtreme Gradient Boosting (XGBoost), as summarised in Table 8.

	Algorithm	**Application in the Literature**	**Source**
1	Zero Rate Classifier		Scikit-learn (Pedregosa *et al.*, 2012)
2	Decision Tree	(Atwal *et al.*, 2022)	Scikit-learn (Pedregosa *et al.*, 2012)
3	Random Forest	(Nachtigall *et al.*, 2023)	Scikit-learn (Pedregosa *et al.*, 2012)
4	Stacked Generalisation	Sturrock *et al.*, 2018	Scikit-learn (Pedregosa *et al.*, 2012)
5	eXtreme Gradient Boosting (XGBoost)	(Nachtigall *et al.*, 2023)	Scikit-learn (Pedregosa *et al.*, 2012)

Table 8: Summary of algorithms to predict building age

Selection of Machine Learning Algorithms for the Study

1. Zero Rate Classifier

The Zero Rate Classifier, also called a Majority Vote Classifier, serves as a baseline model against which more complex models can be compared. While essential for providing a reference point, especially in the case of imbalanced datasets, this classifier is not designed to offer high predictive accuracy. Its primary function is as a control to assess the efficacy of more sophisticated algorithms. However, relying solely on it would not yield useful or actionable insights for predicting building age.

The Zero Rate Classifier is straightforward. If C represents the classes in the dataset, and c is the most frequent class, the Zero Rate Classifier sets all predictions to c:

$$\text{Prediction} = c, \quad \forall c \in C$$

Figure 9: Equation for Zero Rate Classifier. Source: WolframAlpha.

The Zero Rate Classifier assigns all samples to the majority class, disregarding the feature values.

2. Decision Tree

Decision Trees offer an easily understandable yet powerful model for both numerical and categorical data. Their applicability is evident in the work by Atwal (Atwal *et al.*, 2022), who utilised a mix of quantitative and qualitative metrics. However, Decision Trees are prone to overfitting, especially when the tree is deep, and can be sensitive to noisy data. This could potentially lead to inaccuracies in predicting building ages, especially if the dataset contains outliers or anomalies.

In a Decision Tree, the Gini index or Entropy is used to make splits. The Gini index G for a given node t is:

$$G(t) = 1 - \sum_{i=1}^{C} p(i|t)^2$$

Figure 10: Equation for Decision Tree Classifier. Source: WolframAlpha.

The Gini index quantifies the impurity of the data at node t, where $p(i|t)$ is the proportion of samples that belong to class i at node t.

3. Random Forest

The Random Forest algorithm is an ensemble method consisting of multiple decision trees, offering high accuracy and scalability, as indicated by Sturrock et al.(Sturrock *et al.*, 2018). However, one drawback is that Random Forest models can be computationally intensive, which could be a limitation for community groups with basic computational resources. Also, while they provide higher accuracy, they are not as easily interpretable as single decision trees, which could be a constraint here.

Random Forest doesn't introduce new equations but makes use of those from the decision tree. The final prediction of the classification is a mode of the predictions of all trees.

$$\text{Prediction} = \frac{1}{N}\sum_{n=1}^{N}\text{Prediction}_{\text{Tree}_n}$$

Figure 11: Equation for Random Forest Classifier. Source: WolframAlpha.

4. Stacked Generalisation

Stacked Generalisation, or 'stacking', combines the predictions of multiple algorithms to improve predictive power, as employed effectively by Sturrock et al. (Sturrock *et al.*, 2018). Although stacking has the advantage of hedging the weaknesses of individual models, it comes with the cost of increased complexity. This can make the model difficult to interpret and could also introduce the risk of overfitting if not correctly validated.

Stacking involves training a meta-model M to combine predictions p1, p2 ,...,pN from N base models. Mathematically, this could be represented as

$$\text{Prediction}_{\text{Meta}} = M(p_1, p_2, ..., p_N)$$

Figure 12: Equation for Stacked Generalisation Classifer. Source: WolframAlpha.

The meta-model M takes the outputs of the base models as inputs and produces a final prediction.

5. eXtreme Gradient Boosting (XGBoost)

XGBoost is highly efficient and flexible, as demonstrated by its successful application in the large-scale study by Atwal et al. (Atwal *et al.*, 2022). However, the algorithm's performance at a country level may not directly translate to a smaller geographical scope like a community. Given that machine learning algorithms often perform differently based on the granularity and volume of the data, there's a potential for reduced predictive accuracy at the community level.

XGBoost minimises a loss function L by iteratively adding weak learners f, usually decision trees. The objective function is to be minimised.

$$\mathrm{Obj}(\Theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t)}) + \sum_{k=1}^{K} \Omega(f_k)$$

Figure 13: Equation for XGBoost. Source: WolframAlpha.

Where l is the loss function, and Ω is the regularisation term.

In summary, while these algorithms offer robust and diverse methodological approaches for enhancing predictive accuracy and model interpretability, they are not without limitations. Factors such as computational intensity, risk of overfitting, and loss of interpretability could pose challenges in achieving the most accurate and actionable results. This can be seen in the Data Processing Pipeline in Figure 14.

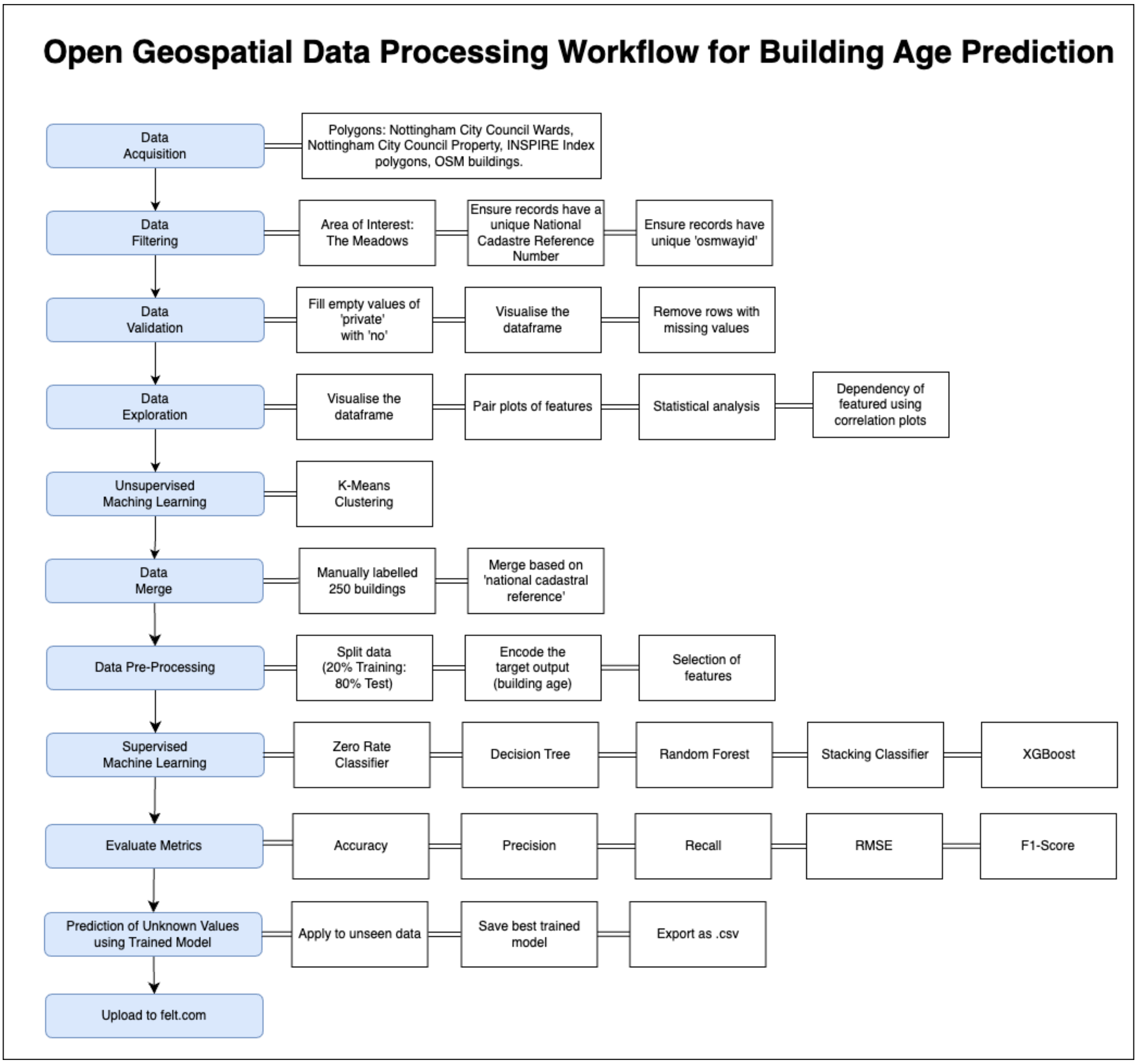


Figure 14: Data Processing Pipeline. Citation: Author's Own.

6.2 Findings

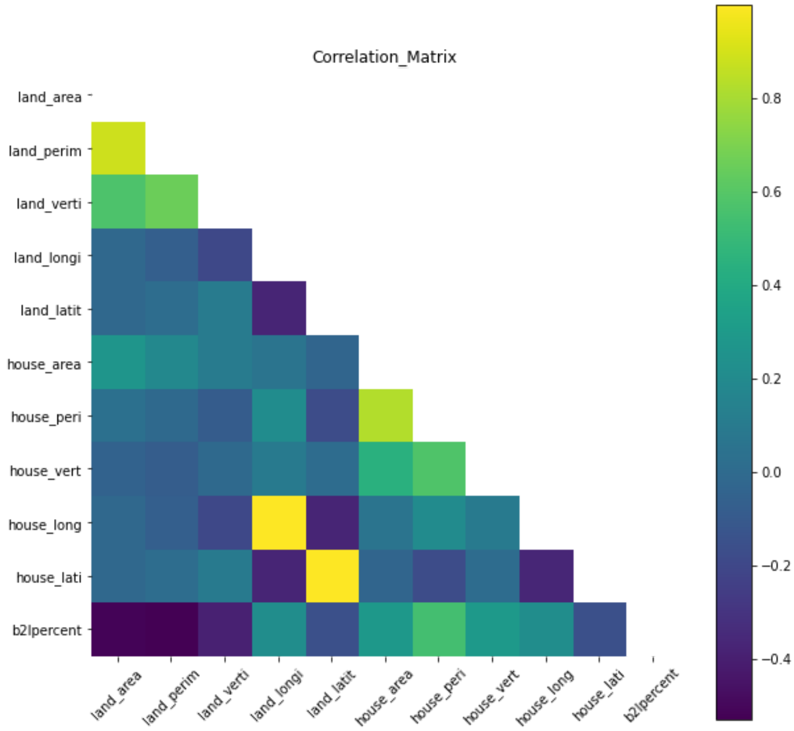


Figure 15: Correlation Heatmap.

This correlation heatmap is a visual guide that shows us how different features of a building, like its location or size, are linked to each other. In this chart, yellow suggests a strong connection between two characteristics, while purple means little to no connection.

Interestingly, the latitude and longitude—basically the geographical coordinates—of houses are very closely connected. This makes sense because homes are often built in clusters for efficiency and cost savings. Previous research has shown that understanding the characteristics of one property can give us valuable information about its nearby neighbours (Sturrock *et al.*, 2018). The neighbouring buildings helped to classify whether buildings were homes and, therefore, whether they should be sprayed with insecticides for malaria prevention.

Usually, in studies like this, we remove features that are too closely linked to avoid duplication. But in this case, geography is a crucial part of what we're studying. So, we've decided to keep this information in, as it helps to make the model more accurate.

On the other end of the correlation spectrum, the size of a building compared to its garden or outdoor space doesn't correlate strongly with other features. This is intriguing, suggesting that you can't make assumptions about garden size based on how big a house is. Whereas it could usually be expected that the bigger the house, the bigger the garden size. However, possibly, because the Meadows is relatively close to the city centre, this may contribute to why the expected relationship is not found.

Unsupervised Clustering

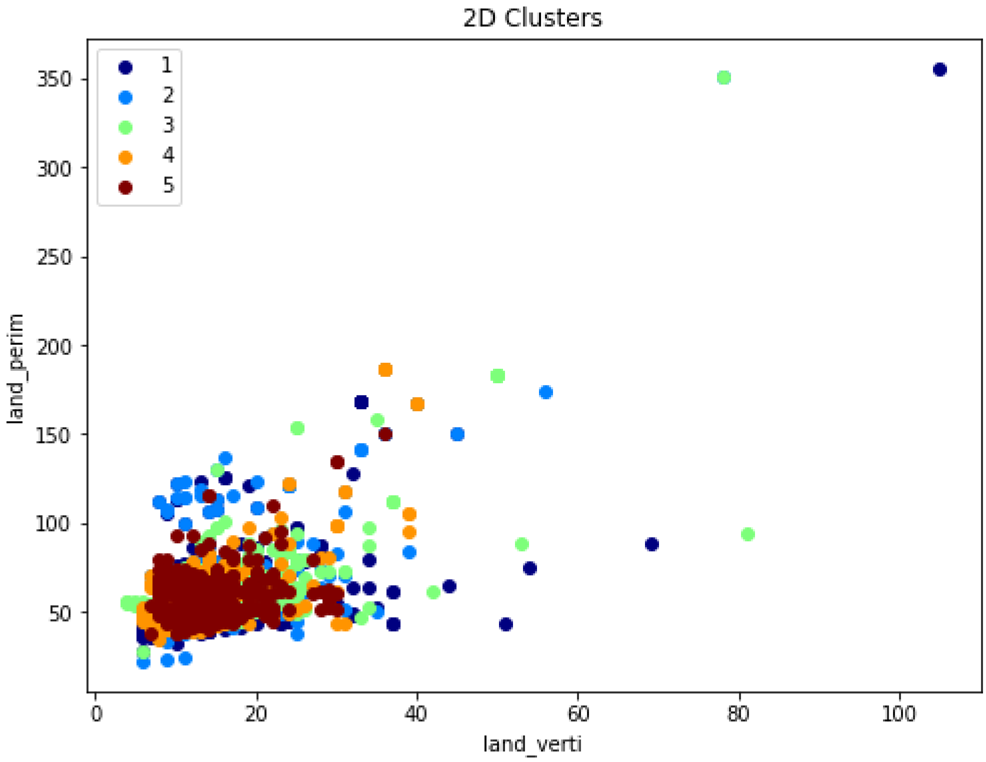


Figure 16: Plot of unsupervised clustering.

In this visualisation, the inability to form distinct clusters is evident, as the data points representing the five intended groups overlap or scatter in a way that doesn't clearly differentiate them. This suggests that the features 'land_verti', 'land_perim', and 'house_area' may not be effective for clustering. This was the same with other clustering options, and so, given this lack of clarity, the current clustering approach was not explored any further.

Model	RMSE	R-Squared	Precision	Recall	Accuracy	F1-score	Avg F1-score
Zero Rate Classifer	1.84	-1.22	0.05	0.22	0.22	0.08	0.07
Decision Tree	1.03	0.3	0.82	0.78	0.78	0.79	0.75
Random Forest	0.43	0.88	0.96	0.95	0.95	0.95	0.96
Stacking Classifier	0.7	0.68	0.93	0.92	0.92	0.92	0.9
XGBoost	0.13	0.99	0.98	0.98	0.98	0.98	0.99

Table 10: Results of the algorithms, including their RMSE, R-Squared, Precision, Recall, Accuracy, F1-Score and Average F1-Score.

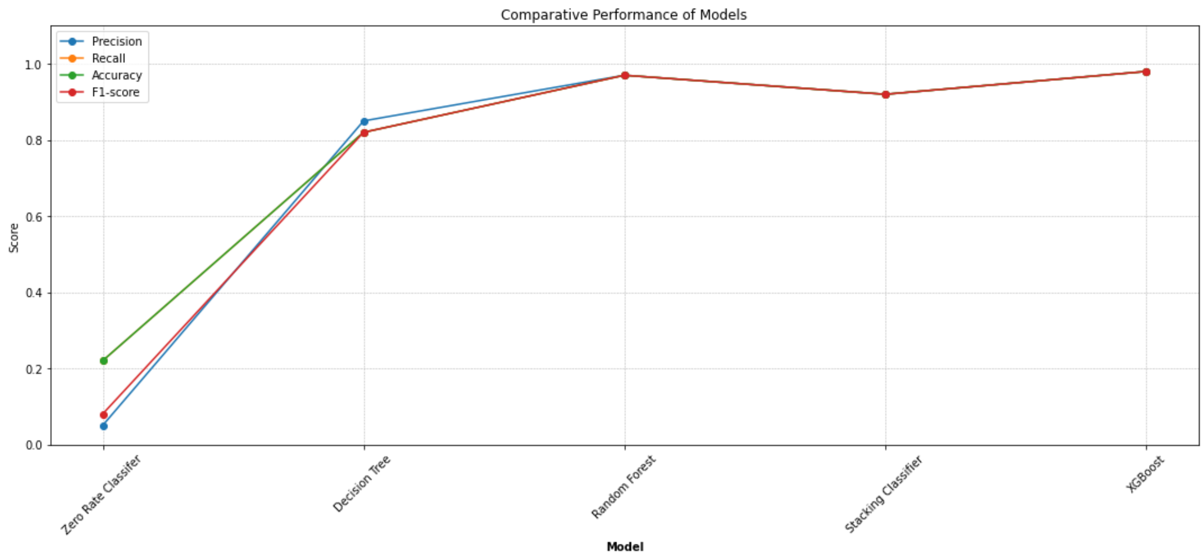


Figure 17: RMSE, R-Squared, Precision, Recall, Accuracy, F1-Score and Average F1-Score plotted and compared on a graph.

Supervised Machine Learning

1. Zero Rate Classifier

The Zero Rate Classifier, which predicts the most frequent class for all observations, achieved an accuracy of 0.2167 and an error rate of 0.7833. This subpar performance is to be expected, as the Zero Rate Classifier is a naive model that does not take into account any of the features of the data (Ren and Bloemraad, 2022). It serves as a baseline for comparison, but a more sophisticated model is required to make accurate predictions.

2. Decision Tree

The decision tree model achieved an accuracy of 0.8333 and an error rate of 0.1667. This suggests the model can effectively capture the salient features relevant to class differentiation. We can see a further break visual breakdown of the roots and nodes in Figure 18 where it is possible to see the number of samples about each node and the Gini coefficients. The accuracy and precision of the model are also at an objectively reasonable level of 0.96 precision and 0.78 recall. In Figure 19, the plot illustrates the Accuracy Over Tree Depth suggesting that the model's performance becomes inconsistent after a tree depth of 5, pointing to a potential overfitting issue.

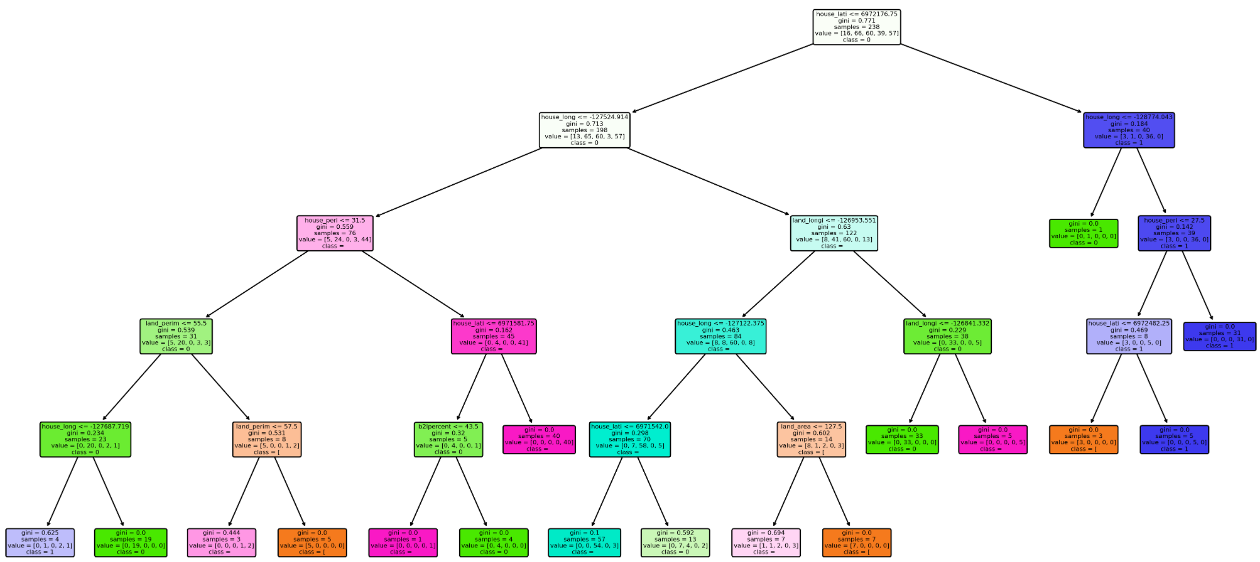


Figure 18: Decision Tree Visualisation.

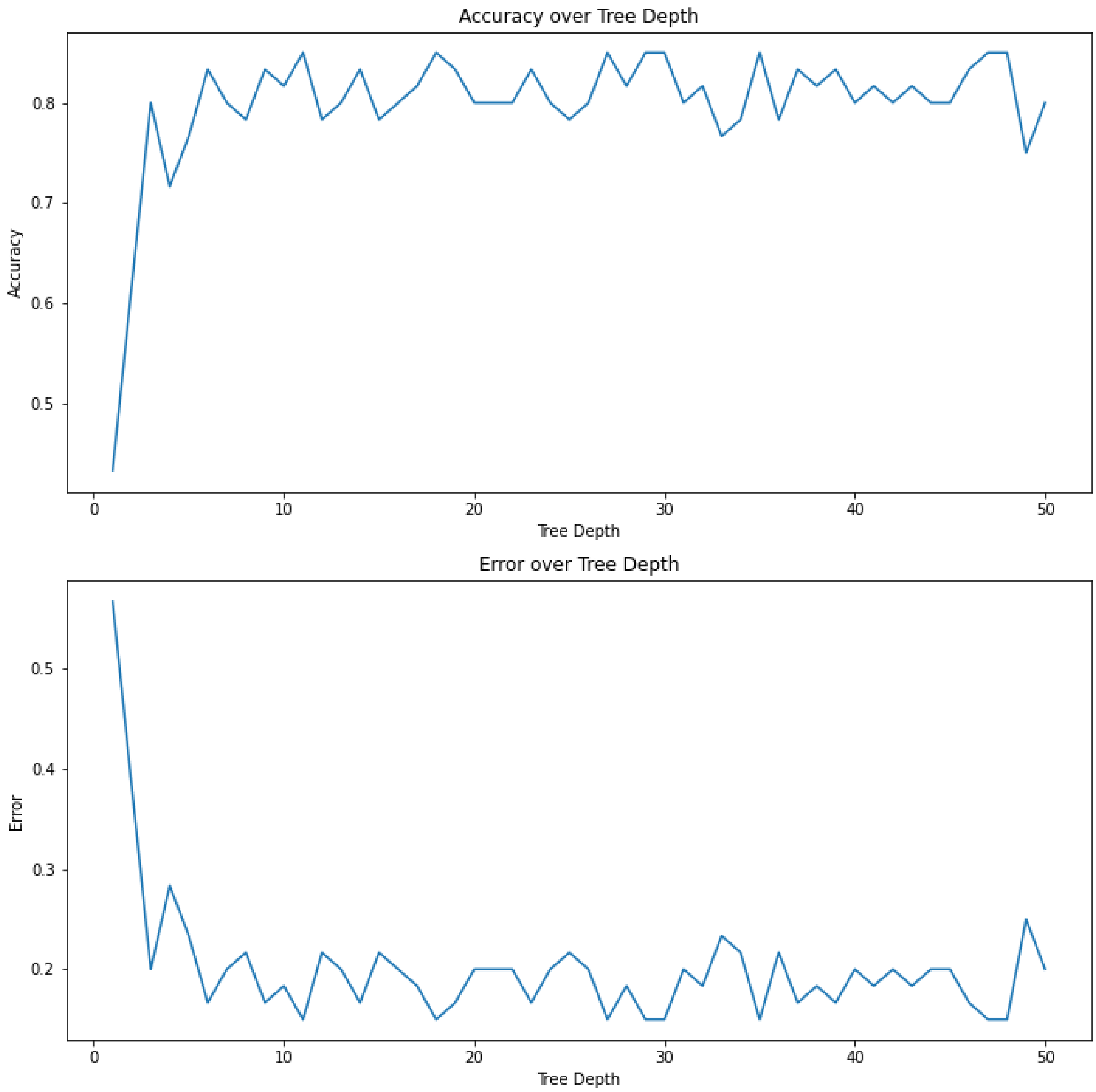


Figure 19: Decision Tree Accuracy and Error graphs.

3. Random Forest

The Random Forest model shows a high accuracy of 0.9667 and a low error rate of 0.0333. The ensemble method, which combines multiple decision trees, appears to substantially improve the model's performance (Segal and Xiao, 2011). The Out of Bag Error over the Trees Plot in Figure 20 indicates that optimal performance is reached at around 70 trees. Beyond this point, the error fluctuates and rises, indicating no further benefit from adding more trees.

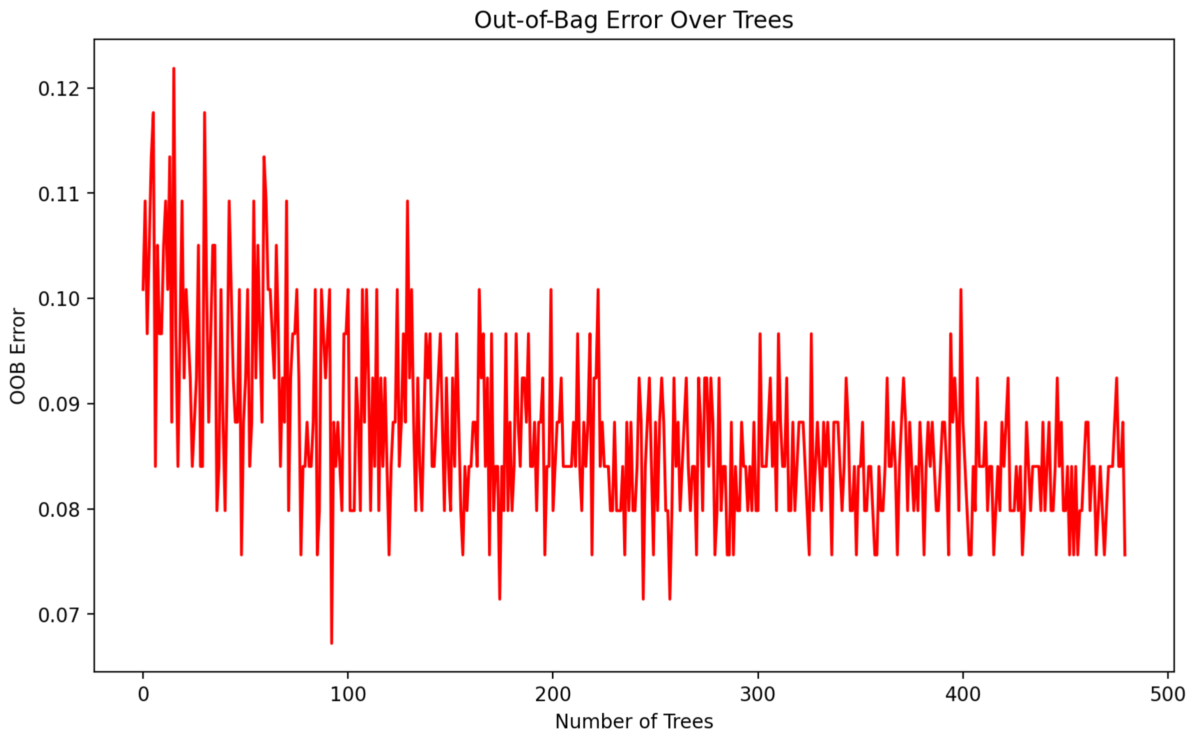


Figure 20: Random Forest Out of Bag Error against Number of Trees.

4. Stacked Generalisation

Stacked Generalisation performed poorly, with an accuracy of 0.2167 and a high error rate of 0.7833. This is similar to the Zero Rate Classifier and could be due to poor model tuning or an ineffective meta-learner (Anifowose, Labadin and Abdulraheem, 2015).

5. eXtreme Gradient Boosting (XGBoost)

The XGBoost model with gradient boosting offered the best performance, with an accuracy of 0.9833 and an error rate of 0.0167. This suggests the model can effectively capture the salient features relevant to class differentiation. The gradient-boosting algorithm optimises the model by iteratively adding new models to the ensemble.

The multi-log loss plot (see Figure 21) shows that the training and testing log loss converge after around ten iterations, suggesting that the model is neither overfitting nor underfitting. Overfitting occurs when the model learns the training data too well and cannot generalise to new data (Ogunleye and Wang, 2020). Underfitting occurs when the model does not learn the training data well enough and cannot make accurate predictions.

The number of features used in the model was 10. The parameters of the gradient boosting algorithm were the learning rate of 0.01 and the number of trees of 100. The evaluation metrics used to measure the performance of the model were accuracy, error rate, and log loss. The steps taken to prevent overfitting were early stopping and regularisation.

The results of this study suggest that the XGBoost model with gradient boosting is a good choice for the classification of building age. The model can achieve high accuracy, and it does not overfit or underfit the training data.

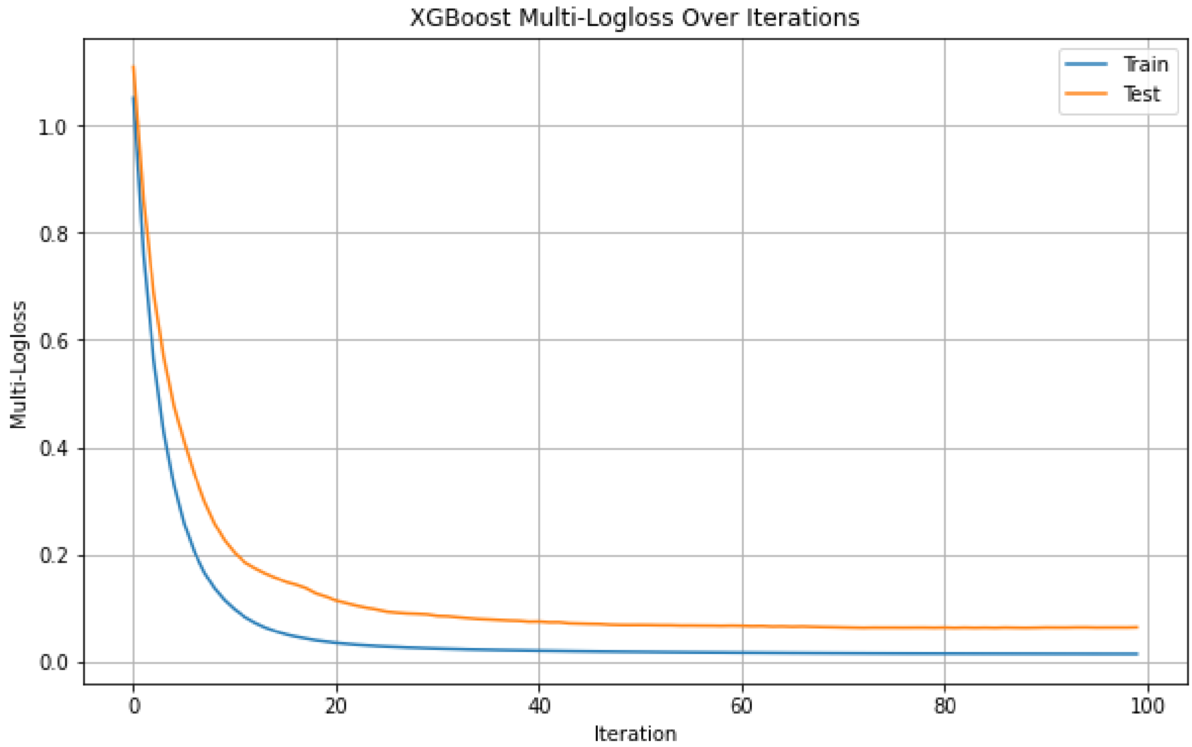
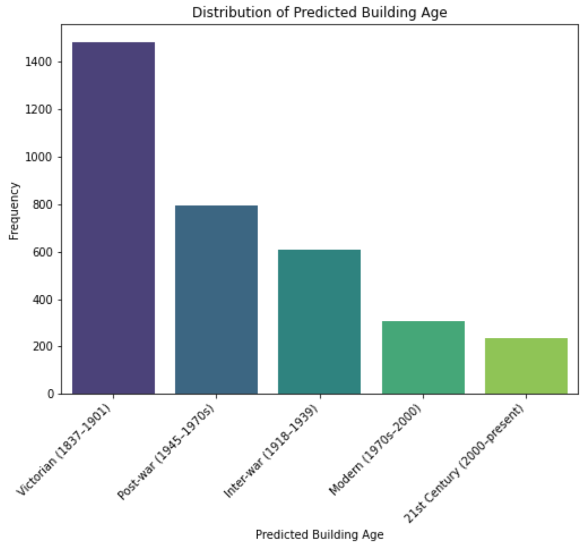


Figure 21: XGBoost Multi-Logloss against the number of iterations.

Summary

Amongst the algorithms tested, XGBoost and Random Forest stands out as the most effective, implying that ensemble methods are particularly well-suited for this building vectorial dataset. The Decision Tree also shows promise but may require tuning to mitigate overfitting; due to limited time at this stage, this was not explored further. Both Stacked Generalisation and Zero Rate Classifier proved inadequate for predicting building age.



Predicted Building Age	Frequency
Victorian (1837-1901)	1484
Post-war (1945-1970s)	792
Inter-war (1918-1939)	608
Modern (1970s-2000)	306
21st Century (2000-present)	236

Figure 22: Bar Plot and table to summarise the number of predicted building ages for The Meadows area.

The results generated by the XGBoost algorithm shed light on the composition of the building stock in the area under study. The most abundant category is the "Victorian (1837–1901), with 1,484 instances indicating that Victorian architecture significantly shapes the local built environment. On the other end of the spectrum, the "21st Century (2000–present)" category appears least frequently, with only 236 instances.

These figures are a historical record of UK government housing policies over the last century. The prevalence of Victorian-era buildings may signify a period of robust construction, possibly driven by the Industrial Revolution and urban expansion. The lesser frequency of 21st-century buildings could indicate a slowdown in new housing developments, perhaps due to more recent governmental policies and broader economic conditions.

In summary, these findings validate the XGBoost model's effectiveness and provide valuable insights into the historical and policy-driven landscape of the Meadow's housing stock.

Confusion Matrix

In our study, creating a confusion matrix by hand is crucial to get a clearer picture of how well our model is doing. Standard measuring accuracy doesn't always give us the whole story (Haghighi *et al.*, 2018). A confusion matrix (Figure 23)lets us dig deeper, showing how often the model got it right (or wrong) for each specific category we are interested in - particularly important when we have uneven numbers in our dataset or more than two categories to predict (Heydarian, Doyle and Samavi, 2022).

		Actual						
		Victorian	Inter-war	Post-war	Modern	21st Century	Classification Overall	User's Accuracy (Precision)
Predicted	Victorian	5	1	0	0	0	6	83%
	Inter-war	0	2	1	2	0	5	40%
	Post-war	0	0	9	0	0	9	100%
	Modern	0	0	0	3	0	3	100%
	21st Century	0	0	0	1	2	3	66%
	Truth Overall	5	3	10	6	2	26	
	Producer's Accuracy (Recall)	100%	66%	90%	50%	100%		

Overall Accuracy: **81%**

Figure 23: Confusion Matrix showing the accuracy of 30 randomly chosen houses.

Individual and Overall Accuracy Metrics

The confusion matrix reveals an overall accuracy of 81%, indicating a generally reliable model. However, this high level of accuracy is not uniformly distributed across all building styles. Specifically, the model performs exceptionally well in identifying Victorian and Post-war buildings, with User Accuracies (also known as Precision) of 83% and 100%, respectively. For Victorian buildings, 5 out of 5 were correctly identified, while for Post-war buildings, 9 out of 9 for Post-war buildings were correctly classified. Conversely, the model's performance dips with Inter-war buildings, achieving only a 40% accuracy rate; only 2 out of 5 predictions were correct for this category.

Producer's Accuracy: A Measure of Recall

In terms of Producer's Accuracy (commonly known as Recall), the model demonstrates strong performance for Victorian and 21st-century buildings, correctly identifying all of the instances, thereby achieving a 100% recall rate. Modern buildings, however, present a challenge with a 50% recall rate; out of 6 actual instances, only three were correctly predicted by the model.

Data Limitations

It's essential to point out that the presented confusion matrix does not account for buildings that were omitted during the polygon pre-processing stage. As such, the current model's actual performance could potentially differ from these results if those omitted instances were included. This is an important consideration for future improvements to the data pre-processing pipeline and model fine-tuning.

To summarise, while the model does well in certain areas, clearly demonstrating its strengths, it also highlights specific styles of buildings where improvement is needed. Such nuanced insights, made possible by the confusion matrix, will serve as crucial reference points for future model refinement.

7. Research Objective 4: Develop & Evaluate

Research Objective 4:
Develop and evaluate a 'community cadastre' proof-of-concept.

This section considers the final research objective, the final stages of creating the proof of concept, and then a brief evaluation of its potential usefulness.

7.1 Methodology

This section will discuss the methodology used to assess the fulfilment of criteria, Green Meadows feedback, and reproducibility of the project.

A) Fulfillment of Criteria

The project's criteria were to identify each house's housing archetype, determine land ownership (public or private), and make the product viewable on a map. To assess the success of meeting these criteria, discuss with Green Meadows again and have a semi-structured conversation about the project, process, outcome, and their thoughts.

B) Green Meadows Feedback

The key questions I asked were:

- What does Green Meadows like about the product?
- What does Green Meadows dislike about the product?
- What other data would be helpful, and who might have that data?
- Is the product accessible to staff and community members?

C) Reproducibility

This project also aimed to look at the reproducibility of the project for other communities in the UK. It will discuss how a community with or without expertise in geospatial data and coding expertise could reproduce this study for their area. Based on experience, it will discuss the extent to which a community would or would not be able to replicate this.

It will integrate the insights gained from the discussion with Green Meadows with analyses and reflections on the outcomes - helping to deepen the overall understanding of the project's impact and limitations and enrich the subsequent phases of this research.

7.2 Findings

At the end of the project, a meeting with the Green Meadows was convened to show them the product, share the created resource and get feedback on whether this would be useful for their work.

The final product was a Felt.com link accessible here: https://tinyurl.com/yck7xsej (see Figure 24)

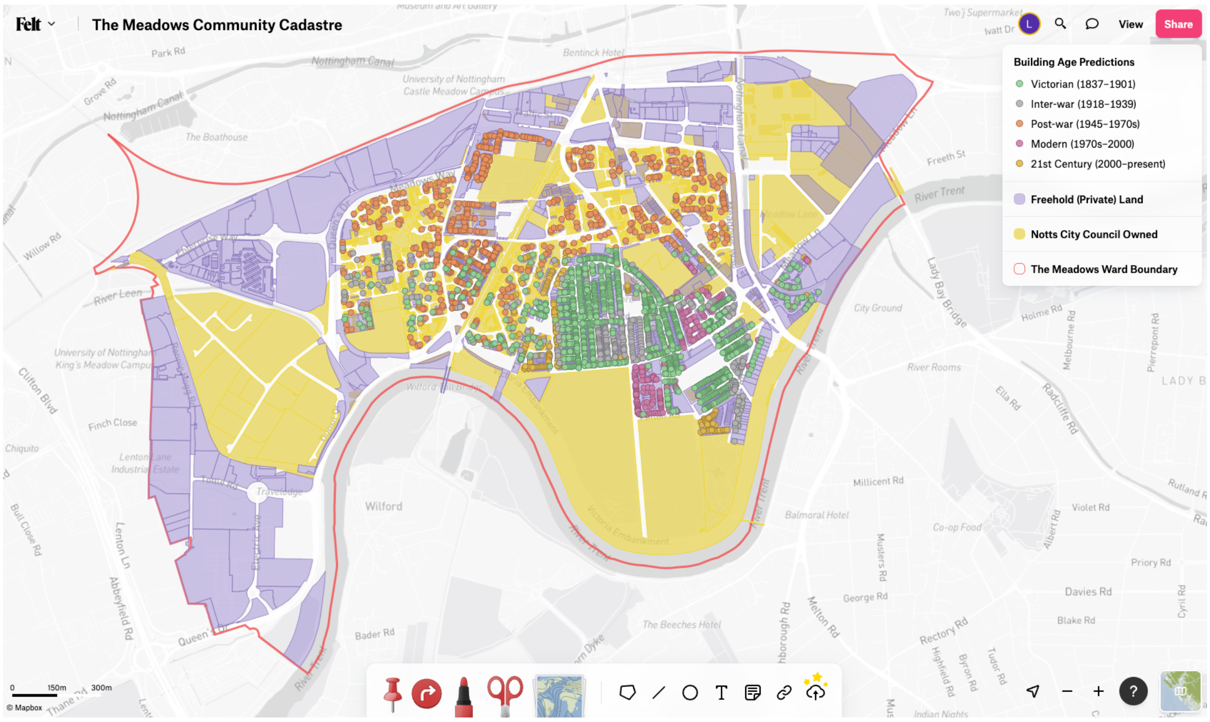


Figure 24: Screenshot of final results, as viewed online at felt.com

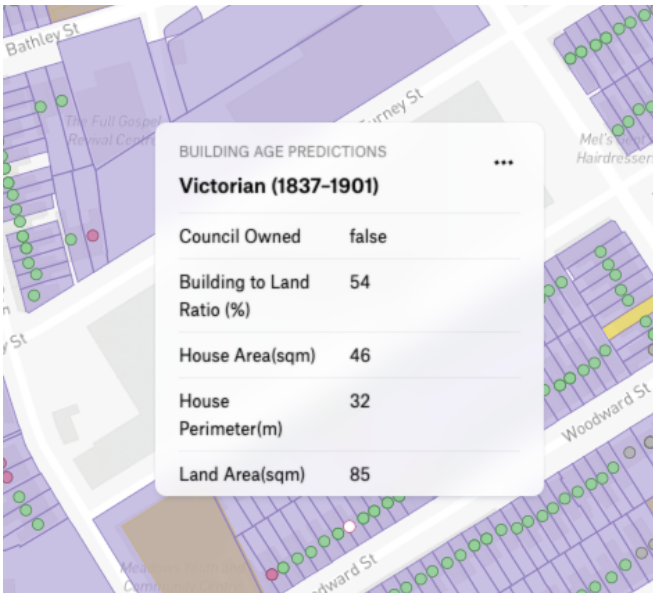


Figure 25: Screenshot of individual house data.

A) Criteria

A screenshot of the finished map can be seen in Figure 24. This shows the user interface of Felt. This mapping software allows users to drag and drop their data onto the map. This was done with four layers in total:

1. Building Age Prediction
2. Freehold (Private Land)
3. Nottingham City Council Owned Land
4. The Meadows Ward Boundary.

Without technical training, users can interact with the same map simultaneously- scribble on top of maps, label, and amend colouring. This is a new mapping tool, which is much more user-friendly than existing desktop-based GIS tools. This would be helpful in community collaboration sessions - discussing and visualising areas that people are talking about and adding interaction capability.

Major Issue Encountered: Alignment of datasets

It should be noted that the projection and alignment of the two datasets is an issue. The OpenStreetMap data is VGI, resulting in some alignment issues; this has previously been an issue with the dataset (Hacar, 2022; Biljecki, Chow and Lee, 2023).

This became an issue when compiling the data with the Nottingham City Council / INSPIRE Index Land polygons - The Nottingham City Council owned land data had been mapped for large areas - for example, where the council owns the land of 10 houses next to each other - this meant that the land and the OSM building polygons could not be individually related to one another on a one-to-one basis. Because of this, these buildings were excluded at the spatial join stage. The resulting large yellow patches of land and the related housing are excluded from the final dataset - as seen in Figure 26. This happened for around 120 properties.

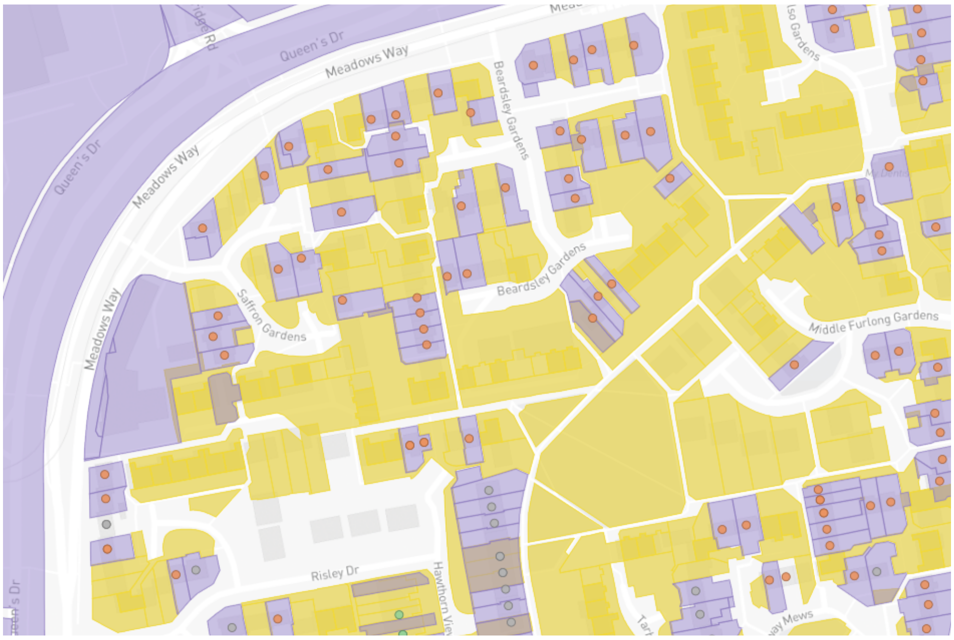


Figure 26: Screenshot of large council-owned patches of land, and houses not included in the study.

B) Green Meadows Feedback

Green Meadows appreciated the map's ability to have an over of the building age predictions at the community level - getting a better understanding of which properties would be grouped together as requiring similar retrofitting interventions. They liked that were able to interact with the map and get property specific details such as size of the plot, building size and building to land ratio.

The charity had not known about the council owned land polygon files that were available to download, and they said this would be useful to understand who owns what land - ie specific small patches of green spaces - for other initiatives that the group were working on - such as a bees biodiversity project and community picnic and teaching events. Knowing htat the council owned a certain piece of land would allow them to liaise with council directly.

To further the land ownership aspect of the work, Green Meadows had said it would be to know which properties were owner-occupied, and which were rented by a private landlord. Similar to the green spaces above, this would ensure that when they were communicating with the property owner, they could prioritise their communications with the owner of the property, who had the ultimate say on whether the retrofitting measures would place or not.

On this point, further data might be available - Nottingham City Council have brought in a ‘Selective Licensing’, where all private landlords have to register their properties in specific areas of the city. The Meadows is an area that is included in the scheme. There is a list of these properties available online - Public Register of Selective Licenses (Notitngham City Council, 2023). Landlords in these areas have to sign up to the scheme - which is in place to improve conditions of these homes for renters - and could be fined upto £30,000 if they do not register. All letting agents dealing with properties also require that landlord provides proof of their registering with the Licensing Scheme. Therefore, this dataset could be very useful for Green Meadows in identifying which properties are owned by landlords, and this had not been included in the initial identification of geospatial data - a limitation of the research.

C) Reproducibility:

A dedicated project website, www.communitycadastre.uk, serves as a repository for this project. The website provides access to the following:

- Screenshots of the Meadows Community Cadastre
- A direct link to the Meadows Community Cadastre
- The underlying codebase
- Step-by-step instructions for communities interested in replicating this project

Technical Expertise Required

It is important to note the level of technical expertise required to replicate this project. The Green Meadows community had limited coding skills, so it is unlikely that they could reproduce this project independently without the assistance of a knowledgeable individual, such as a student with coding experience.

To replicate the project, a foundational understanding of computer science and coding is necessary, particularly for tasks such as downloading code from a GitHub repository. Additionally, primary

Geographic Information System (GIS) skills are essential for downloading geospatial files from the internet and handling them within a QGIS environment.

Therefore, the project is most feasibly replicable by communities with access to someone with at least rudimentary coding and GIS skills. I will elaborate on the implications of this in the discussion chapter.

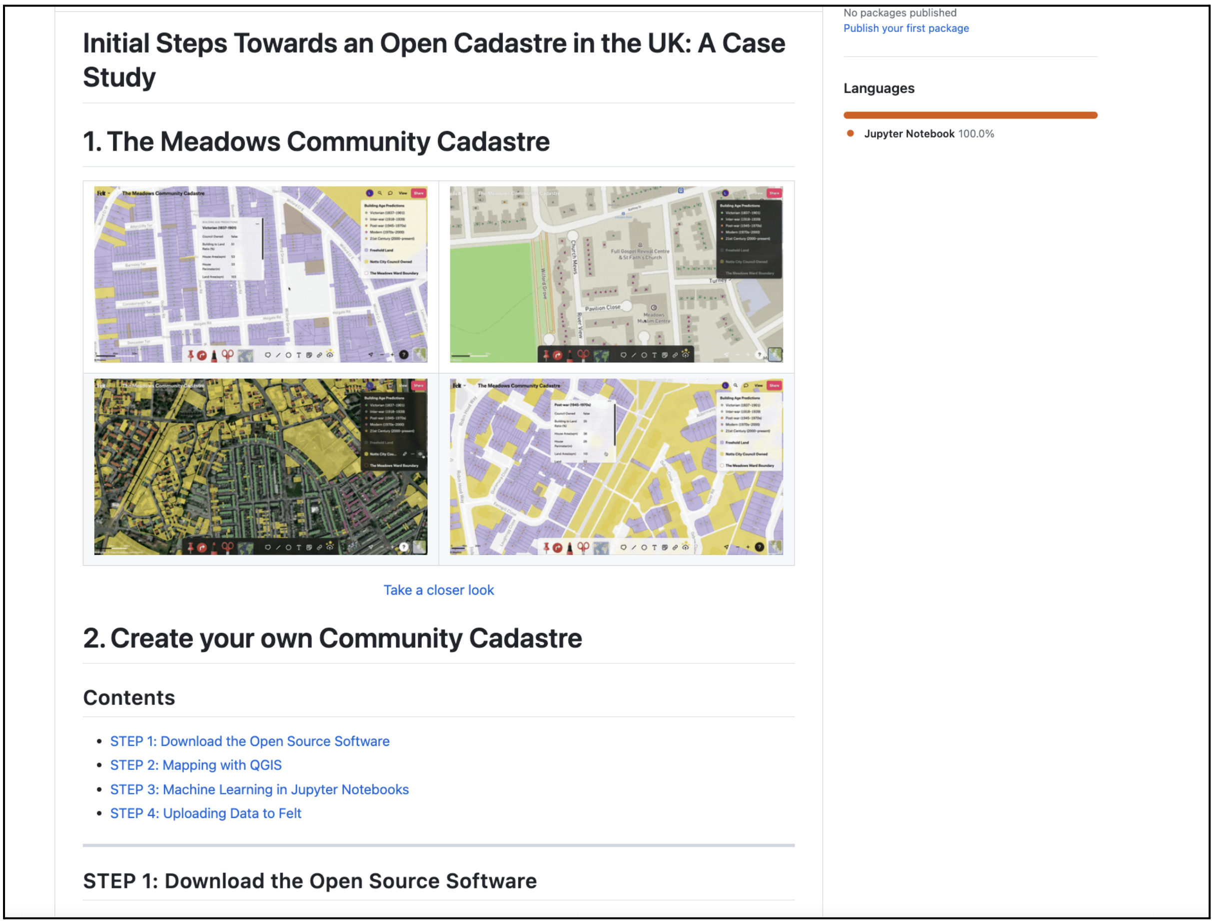


Figure 27: Screenshot of the website with step-by-step instructions and code on GitHub.

8. Discussion

This chapter discusses the findings of the research, as well as the limitations and implications of the study.

A Community's Cadastral Needs

Green Meadows is a charity that works to improve the sustainability of The Meadows area. By forming a partnership with the charity, they shared some of their data, and I carried out machine learning to further validate their dataset. This information will help them better understand the retrofitting interventions for different houses in their area.

Green Meadows specified what they wanted the cadastre to be able to do. To an extent, the project addressed the needs - showing which land was publicly or privately owned, the building age, which related to the house archetype, and the information was accessed via a map. However, the specifics of the archetype were not possible - given the limits of open data quality and time to complete this project.

Data Quality & Availability

The quality of the data was mixed. The INSPIRE Index Polygons data was of high quality. However, the land use data from the OSM had some accuracy and currency issues - for example, there are instances of buildings built outside the dataset. Whilst this can be considered a limitation of the project, it is also only to be used as a starting point - the ability of this project to give communities initial metrics to understand their area better and investigate further could be a positive. Because the project is operating at a local level, the community could then go out and further investigate any anomalies, taking ownership in amending and updating any errors found.

Using open data sources and volunteered geographic information (VGI) to create the map. The open data sources include Nottingham City Council land owned, ward boundaries, and INSPIRE Index Land Polygons. The VGI data included building age data from OpenStreetMap. These datasets are available openly and freely and provide useful results. However, the higher quality datasets that could be obtained either through a community partnering with a private organisation or through the use of an academic license would probably allow a community to get more accurate and reliable results.

An academic partnership might also bring with it motivation and creativity from students. This is a style of project that is used frequently in LEDC countries when addressing planning and cadastral matters and could be a good opportunity in the UK should there be interest in cadastres from other students and/or community groups (Anne Harwood and Zapata, 2006; Winkler, 2013).

Effectiveness of Processing Data

I used a machine learning algorithm to classify the buildings into building ages. The algorithm achieved a high accuracy. However, the granularity of the building age was only one step in the archetype prediction process. To further refine and make this more relevant to Green Meadows, there would need to be more data added to better define the Building beyond its building age and into the specific features of the building age - for example, identifying whether the building has a dormer-style roof or not. Whether there are 2 or 3 floor levels in the property. It is possible to achieve this - by including different data sources, i.e., the Digital Elevation Model available through an academic licence with Ordnance Survey or high spatial resolution aerial imagery to classify the rooftypes.

Using council-owned data raised a brief ethical consideration from the author's view. Given that some people may not want their property information to be private - for example, if they live in social housing, and the potential stigma related to this (Kearns, Kearns and Lawson, 2013). Although all of this data is freely available, to begin to share further and make available, the datasets could result in some community members being unhappy that this data is being more widely made available and shared.

Limitations of the Study

The study had several limitations. First, the data was limited to a small area of Nottingham. At the same time, in an area with a diverse range of houses built over the years, there are still many different variations of housing built, and this may not be widely applicable to other parts of Nottingham, the wider region, or other parts of the UK. Second, the machine learning algorithm was only trained on a minimal dataset - trained with 280 label pieces of data - so it may need to be more generalisable to other areas. Third, the study did not go in-depth into the archetypes of the building ages - this is because, as already stated, much more data would be needed, including building height (Milojevic-Dupont *et al.*, 2020), potentially streetview imagery (Pietro, Marco and Francesca, 2023), and more.

9. Conclusion

The research journey has been a multifaceted exploration of open geospatial datasets and open-source software for community-based cadastral mapping. Partnering with Green Meadows provided invaluable insights into the community's needs, challenges, and aspirations, significantly influencing the project's direction.

The initial question of, ‘How practical are open geospatial datasets and geospatial open-source software in enabling communities to create their own ‘community cadastre’?’, has been explored through the development of a definition, identification of useful open datasets, prediction of the missing building age classifications and creation and evaluation of an open-source community cadastre proof of concept.

The project has been insightful and fun, given that it is an issue close to the author’s heart - poverty, inequality, and land matters. This dissertation, although geospatial in nature, has discussed historical and social issues of power and influence, which continue to be played out in modern British society.

Open data and open-source software are allowing academia and communities to partner and make changes that address some of the biggest and most important issues of today.

The Machine Learning utilised here was limited, but with better datasets and more time to enhance the prediction algorithms, likely, the housing archetypes that Green Meadows wished to identify can indeed be found to a high degree of accuracy. The next steps for this research would be to collect more data and to improve the machine learning algorithm.

The Green Meadows is open to continuing this partnership, and the author hopes to support them in their endeavours to empower The Meadows community and successfully lower its carbon emissions.

Bibliography

Alemie, B.K., Bennett, R.M. and Zevenbergen, J. (2015) 'Evolving urban cadastres in Ethiopia: The impacts on urban land governance', *Land Use Policy*, 42, pp. 695–705. Available at: https://doi.org/10.1016/j.landusepol.2014.10.001.

Amandala (2007) 'Land and people – a country's most valuable resource'. Available at: https://amandala.com.bz/news/land-and-people-a-countrys-most-valuable-resource/.

Anifowose, F., Labadin, J. and Abdulraheem, A. (2015) 'Improving the prediction of petroleum reservoir characterization with a stacked generalization ensemble model of support vector machines', *Applied Soft Computing*, 26, pp. 483–496. Available at: https://doi.org/10.1016/j.asoc.2014.10.017.

Anne Harwood, S. and Zapata, M. (2006) 'Preparing to Plan: Collaborative Planning in Monteverde, Costa Rica', *International Planning Studies*, 11(3–4), pp. 187–207. Available at: https://doi.org/10.1080/13563470601097311.

Ashby, M.A. and Heinemeyer, A. (2021) 'A Critical Review of the IUCN UK Peatland Programme's "Burning and Peatlands" Position Statement', *Wetlands*, 41(5), p. 56. Available at: https://doi.org/10.1007/s13157-021-01400-1.

Atwal, K.S. *et al.* (2022) 'Predicting building types using OpenStreetMap', *Scientific Reports*, 12(1), p. 19976. Available at: https://doi.org/10.1038/s41598-022-24263-w.

Augustinus, C. (2023) 'The Global Land Outlook and Protecting Our Planet (12300) FIG Working Week 2023'. *FIG Working Week 2023*, Orlando, Florida, USA, 28 June. Available at: https://fig.net/resources/proceedings/fig_proceedings/fig2023/papers/ps02/PS02_augustinus_12300_abs.pdf.

Bahl, R. and Vazquez, J. (2008) 'The Property Tax in De ax in Developing Countries: Curr eloping Countries: Current Practice and actice and Prospects', in *In Toward a Vision of Land in 2015: International Perspectives*. (Lincoln Institute of Land Policy), pp. 23–46.

Bardsley, N. *et al.* (2019) 'Domestic thermal upgrades, community action and energy saving: A three-year experimental study of prosperous households', *Energy Policy*, 127, pp. 475–485. Available at: https://doi.org/10.1016/j.enpol.2018.11.036.

Beck, A. *et al.* (2020) 'Automated classification metrics for energy modelling of residential buildings in the UK with *open* algorithms', *Environment and Planning B: Urban Analytics and City Science*, 47(1), pp. 45–64. Available at: https://doi.org/10.1177/2399808318762436.

BEIS (2017) 'Department for Business, Energy & Industrial Strategy. Energy consumption in the UK (ECUK) 2017'. Available at: www.gov.uk/government/uploads/system/uploads/attachment_data/file/633503/ECUK_2017.pdf.

Bennett, R.M. and Alemie, B.K. (2016) 'Fit-for-purpose land administration: lessons from urban and rural Ethiopia', *Survey Review*, 48(346), pp. 11–20. Available at: https://doi.org/10.1080/00396265.2015.1097584.

Bennett, R.M., Koeva, M. and Asiama, K. (2021) 'Review of Remote Sensing for Land Administration: Origins, Debates, and Selected Cases', *Remote Sensing*, 13(21), p. 4198. Available at: https://doi.org/10.3390/rs13214198.

BikePacking.com (2022) 'Introducing Felt, A Digital Collaborative Map'. Available at: https://bikepacking.com/news/felt-digital-collaborative-map/#:~:text=Felt%20also%20doesn't%20replace,that's%20not%20the%20GPX%20trace.

Biljecki, F., Chow, Y.S. and Lee, K. (2023) 'Quality of crowdsourced geospatial building information: A global assessment of OpenStreetMap attributes', *Building and Environment*, 237, p. 110295. Available at: https://doi.org/10.1016/j.buildenv.2023.110295.

Bullard, R. (2003) 'The English and Hong Kong Cadastres', 5(1), pp. 55–60.

Calcutt, J. (2007) 'The Calcutt Review of Housebuilding Delivery'.

Charalabidis, Y., Loukis, E. and Alexopoulos, C. (2014) 'Evaluating Second Generation Open Government Data Infrastructures Using Value Models', in *2014 47th Hawaii International Conference on System Sciences*. *2014 47th Hawaii International Conference on System Sciences (HICSS)*, Waikoloa, HI: IEEE, pp. 2114–2126. Available at: https://doi.org/10.1109/HICSS.2014.267.

Chi‐Man Hui, E., Sze‐Mun Ho, V. and Kim‐Hin Ho, D. (2004) 'Land value capture mechanisms in Hong Kong and Singapore: A comparative analysis', *Journal of Property Investment & Finance*, 22(1), pp. 76–100. Available at: https://doi.org/10.1108/14635780410525153.

Cho, G. and Crompvoets, J. (2019) 'The INSPIRE directive: some observations on the legal framework and implementation', *Survey Review*, 51(367), pp. 310–317. Available at: https://doi.org/10.1080/00396265.2018.1454686.

Chollet, F. (2018) *Deep Learning with Python*. Manning Publications.

Coetzee, S. *et al.* (2020) 'Open Geospatial Software and Data: A Review of the Current State and A Perspective into the Future', *ISPRS International Journal of Geo-Information*, 9(2), p. 90. Available at: https://doi.org/10.3390/ijgi9020090.

Commission for Climate Change (2019) 'UK Housing: Fit for the Future?' Available at: https://www.theccc.org.uk/wp-content/uploads/2019/02/UK-housing-Fit-for-the-future-CCC-2019.pdf.

Craig, B. and Wahl, J. (2003) 'Cadastral survey accuracy standards.', *Surveying and Land Information Science*, 63(2), pp. 87–106.

Dale, P.F. (1977) 'Cadastres and Cadastral Maps', *The Cartographic Journal*, 14(1), pp. 44–48. Available at: https://doi.org/10.1179/caj.1977.14.1.44.

Damsø, T., Kjær, T. and Christensen, T.B. (2016) 'Local climate action plans in climate change mitigation – examining the case of Denmark', *Energy Policy*, 89, pp. 74–83. Available at: https://doi.org/10.1016/j.enpol.2015.11.013.

Davies, G.M. *et al.* (2016) 'The role of fire in UK peatland and moorland management: the need for informed, unbiased debate', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1696), p. 20150342. Available at: https://doi.org/10.1098/rstb.2015.0342.

De Urquia, B.V. (2022) 'Insulating UK homes could save the country "billions" assessments show [news - briefing]', *Engineering & Technology*, 17(10), pp. 10–10. Available at: https://doi.org/10.1049/et.2022.1003.

Deininger, K.W. (2010) *Innovations in land rights recognition, administration, and governance*. Washington, D.C: World Bank.

Department for Business Enterprise & Regulatory Reform (2007) ‘UK Energy in Brief’.

Department for Environment, Food and Rural Affairs (2014) ‘Rural Land Register’. Available at: https://www.gov.uk/government/collections/rural-land-register.

Department for Levelling Up, Housing & Communities (2022) ‘Official Statistics Land use statistics: England 2022’. Available at: https://www.gov.uk/government/statistics/land-use-in-england-2022/land-use-statistics-england-2022.

Department for Levelling Up, Housing and Communities (2015) *Local government transparency code 2015*. Available at: https://www.gov.uk/government/publications/local-government-transparency-code-2015.

Ding, J. *et al.* (2019) ‘A Case Study of the Augmentation and Evaluation of Training Data for Deep Learning’, *Journal of Data and Information Quality*, 11(4), pp. 1–22. Available at: https://doi.org/10.1145/3317573.

Douglas, R. (2011) ‘The Lloyd George Land Taxes’, *Journal of Liberal History*, 73(Winter). Available at: https://liberalhistory.org.uk/wp-content/uploads/2014/10/73_Douglas_Lloyd_George_land_taxes.pdf.

Enemark (2014) ‘Fit-for-Purpose Land Administration’. FIG Publications. Available at: https://vbn.aau.dk/ws/files/196369601/Figpub60_screen.pdf.

Enemark, S. (2015) ‘Understanding the Concept of Fit-For-Purpose Land Administration in Support of the Post 2015 Global Agenda’. FIG Working Week. Available at: http://www.fig.net/resources/proceedings/fig_proceedings/fig2015/papers/ts01b/TS01B_enemark_7614.pdf.

Enemark, S. and Rajabifard, A. (2009) ‘Land Administration for Sustainable Development.’, in. Available at: https://www.researchgate.net/publication/242494736_Land_Administration_for_Sustainable_Development.
‘England’s Peatlands: Greenhouse Gases and Carbon Storage. Natural England report NE257’ (2010). Natural England. Available at: https://publications.naturalengland.org.uk/publication/30021.

Ercan, O. (2021) ‘A closer look at Turkish cadastre and its successful completion’, *Land Use Policy*, 110, p. 104951. Available at: https://doi.org/10.1016/j.landusepol.2020.104951.

European Commission. Directorate General for the Information Society and Media. *et al.* (2015) *Creating value through open data: study on the impact of re use of public data resources*. LU: Publications Office. Available at: https://data.europa.eu/doi/10.2759/328101 (Accessed: 10 September 2023).

Feder, G. (1986) ‘Land Ownership Security and Farm Productivity in Rural Thailand’, in. World Bank. Available at: https://documents1.worldbank.org/curated/en/841451468915314093/pdf/Land-ownership-security-and-farm-productivity-in-rural-Thailand.pdf.

Felt Team (2023) ‘feltmaps.com’. Available at: https://feltmaps.notion.site/Help-Center-9ea41c5449554b75858a76b2d4d21248.

FIG (1995) ‘FIG Statement on the Cadastre’. FIG Publications.

Galán, A. (2015) 'TCS Matrix: Evaluation of Optimal Energy Retrofitting Strategies'.

Geospatial Commission (2023) 'UK Geospatial Strategy 2030'. Available at: https://www.gov.uk/government/publications/uk-geospatial-strategy-2030#:~:text=The%20UK%20Geospatial%20Strategy%202030,technologies%20across%20the%20UK%20economy.

Global Land Tools Network (2023) 'STDM'. Available at: https://github.com/gltn/stdm.

Golob, P. and Lisec, A. (2022) 'Success factors in cadastral boundary settlements based on land surveyor's opinions', *Land Use Policy*, 114, p. 105990. Available at: https://doi.org/10.1016/j.landusepol.2022.105990.

Gray, K. and Gray, S.F. (2005) *Elements of Land Law*. 4th edn. Oxford, UK: Oxford University Press.

Green Meadows (2023) 'Conversation with Green Meadows. 12/06/2023'.

Griffith-Charles, C. (2011) 'The application of the social tenure domain model (STDM) to family land in Trinidad and Tobago', *Land Use Policy*, 28(3), pp. 514–522. Available at: https://doi.org/10.1016/j.landusepol.2010.10.004.

Griffith-Charles, C. *et al.* (2015) 'Key challenges and outcomes of piloting the STDM in the Caribbean', *Land Use Policy*, 49, pp. 577–586. Available at: https://doi.org/10.1016/j.landusepol.2014.12.018.

Grover, R. (2008) 'Why the United Kingdom does not have a cadastre – and does it matter?', in. *Annual Meeting 2008 FIG Commission 7*, Verona, Italy: FIG Publications. Available at: https://www.fig.net/resources/proceedings/2008/verona_am_2008_comm7/papers/12_sept/7_2_grover.pdf.

Hacar, M. (2022) 'Analyzing the Behaviors of OpenStreetMap Volunteers in Mapping Building Polygons Using a Machine Learning Approach', *ISPRS International Journal of Geo-Information*, 11(1), p. 70. Available at: https://doi.org/10.3390/ijgi11010070.

Haghighi, S. *et al.* (2018) 'PyCM: Multiclass confusion matrix library in Python', *Journal of Open Source Software*, 3(25), p. 729. Available at: https://doi.org/10.21105/joss.00729.

Hanus, P., Pęska-Siwik, A. and Szewczyk, R. (2018) 'Spatial analysis of the accuracy of the cadastral parcel boundaries', *Computers and Electronics in Agriculture*, 144, pp. 9–15. Available at: https://doi.org/10.1016/j.compag.2017.11.031.

Henssen, J. (1995) 'Cadastre and its legal aspects'. *'From Research to Application through Cooperation' of the Joint European Conference and Exhibition on Geographical Information*, The Hague, The Netherlands, 26 March. Available at: https://www.fig.net/organisation/comm/7/Library/reports/events/Delft_seminar_95/paper2.html.

Heydarian, M., Doyle, T.E. and Samavi, R. (2022) 'MLCM: Multi-Label Confusion Matrix', *IEEE Access*, 10, pp. 19083–19095. Available at: https://doi.org/10.1109/ACCESS.2022.3151048.

Historic England (2019) 'A Brief Introduction to Terraced Housing', *The Historic England Blog*, 15 November. Available at: https://heritagecalling.com/2019/11/15/a-brief-introduction-to-terraced-housing/.

HM Government (2013) *Construction*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/210099/bis-13-955-construction-2025-industrial-strategy.pdf.

HM Land Registry (2023) ‘Get information about property and land’. Available at: https://www.gov.uk/get-information-about-property-and-land/copies-of-deeds.

ISO (2012) ‘ISO 19152:2012 Geographic information — Land Administration Domain Model (LADM)’. Available at: https://www.iso.org/obp/ui/en/#iso:std:iso:19152:ed-1:v1:en.

Ji, Y., Lee, A. and Swan, W. (2019) ‘Retrofit modelling of existing dwellings in the UK: the Salford Energy House case study’, *International Journal of Building Pathology and Adaptation*, 37(3), pp. 344–360. Available at: https://doi.org/10.1108/IJBPA-12-2018-0106.

Johnson, P.A. *et al.* (2017) ‘The Cost(s) of Geospatial Open Data: JOHNSON et al.’, *Transactions in GIS*, 21(3), pp. 434–445. Available at: https://doi.org/10.1111/tgis.12283.

Jones, B. (2023) ‘Mapping a Sustainable Future in This New Era of Understanding (12302) FIG Working Week 2023’. *FIG Working Week 2023*, Orlando, Florida, USA, 28 June. Available at: https://fig.net/resources/proceedings/fig_proceedings/fig2023/papers/ps02/PS02_jones_12302_abs.pdf.

Jones, O. (2015) *The establishment: and how they get away with it*. London: Penguin Books.

Kain, R.J.P. (1975) ‘R. K. Dawson’s Proposal in 1836 for a Cadastral Survey of England and Wales’, *The Cartographic Journal*, 12(2), pp. 81–88. Available at: https://doi.org/10.1179/caj.1975.12.2.81.

Kearns, A., Kearns, O. and Lawson, L. (2013) ‘Notorious Places: Image, Reputation, Stigma. The Role of Newspapers in Area Reputations for Social Housing Estates’, *Housing Studies*, 28(4), pp. 579–598. Available at: https://doi.org/10.1080/02673037.2013.759546.

Koeva, M. *et al.* (2020) ‘Innovative Remote Sensing Methodologies for Kenyan Land Tenure Mapping’, *Remote Sensing*, 12(2), p. 273. Available at: https://doi.org/10.3390/rs12020273.

Krigsholm, P. *et al.* (2017) ‘Understanding the future of the Finnish cadastral system – A Delphi study’, *Land Use Policy*, 68, pp. 133–140. Available at: https://doi.org/10.1016/j.landusepol.2017.07.032.

Lakomaa, E. and Kallberg, J. (2013) ‘Open Data as a Foundation for Innovation: The Enabling Effect of Free Public Sector Information for Entrepreneurs’, *IEEE Access*, 1, pp. 558–563. Available at: https://doi.org/10.1109/ACCESS.2013.2279164.

Land Registry (2019) ‘Registering land for more than 150 years’. Available at: https://hmlandregistry.blog.gov.uk/2019/05/30/registering-land-for-more-than-150-years/.

Lomax, S. (2013) *Nottingham: The Buried Past of a Historic City Revealed*. South Yorkshire: Pen & Sword Archaeology.

Martyn, A. (2018) ‘The Concept of Land Plot as a Combination of Smart Contracts: a Vision for Creating Blockchain Cadastre’, *Baltic Surveying*, 8, pp. 68–73. Available at: https://doi.org/10.22616/j.balticsurveying.2018.009.

Medda, F. (2012) ‘Land value capture finance for transport accessibility: a review’, *Journal of Transport Geography*, 25, pp. 154–161. Available at: https://doi.org/10.1016/j.jtrangeo.2012.07.013.

Milojevic-Dupont, N. *et al.* (2020) ‘Learning from urban form to predict building heights’, *PLOS ONE*. Edited by F. Biljecki, 15(12), p. e0242010. Available at: https://doi.org/10.1371/journal.pone.0242010.

Ministry of Public Building and Works (1965) *The Building Regulations*.

Minn, M. (2021) 'MMQGIS'. Available at: https://plugins.qgis.org/plugins/mmqgis/.

Mohammadpourkarbasi, H. and Sharples, S. (2013) 'Eco-Retrofitting Very Old Dwellings: Current and Future Energy and Carbon Performance for Two UK Cities', *29th Conference, Sustainable Architecture for a Renewable Future*, Munich, Germany 10-12 September 2013.

van der Molen, P. (2014) '"Cadastre 2014": a Beacon in Turbulent Times', in *FIG Congress 2014 Engaging the Challenges – Enhancing the Relevance Kuala Lumpur, Malaysia*. FIG Publications. Available at: https://www.fig.net/resources/proceedings/fig_proceedings/fig2014/papers/ss27/SS27_van_der_molen_7119.pdf.

Moloney, S., Fünfgeld, H. and Granberg, M. (eds) (2018) *Local action on climate change: opportunities and constraints*. London ; New York: Routledge, Taylor & Francis Group (Routledge advances in climate change research).

Nachtigall, F. *et al.* (2023) 'Predicting building age from urban form at large scale', *Computers, Environment and Urban Systems*, 105, p. 102010. Available at: https://doi.org/10.1016/j.compenvurbsys.2023.102010.

Nejat, P. *et al.* (2015) 'A global review of energy consumption, CO 2 emissions and policy in the residential sector (with an overview of the top ten CO 2 emitting countries)', *Renewable and Sustainable Energy Reviews*, 43, pp. 843–862. Available at: https://doi.org/10.1016/j.rser.2014.11.066.

Notitngham City Council (2023) 'Public Register of Selective Licences'. Available at: https://www.nottinghamcity.gov.uk/media/3rlpnrr3/9-sl-1st-sept-2023-public-register-redacted.xlsx.

Nottingham City Council (2020) *Old Meadows: Conservation Area Character Appraisal and Management Plan*. Available at: https://www.nottinghamcity.gov.uk/information-for-business/planning-and-building-control/planning-policy/conservation-areas-and-listed-buildings/old-meadows-conservation-area/.

Nottingham Climate Assembly (2021) 'The Green Meadows Project'. Available at: https://nottinghamclimateassembly.co.uk/2021/04/16/the-green-meadows-project/.

Office for National Statistics (2022) 'Population estimates for the UK, England, Wales, Scotland and Northern Ireland: mid-2021'. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2021.

Ogunleye, A. and Wang, Q.-G. (2020) 'XGBoost Model for Chronic Kidney Disease Diagnosis', *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 17(6), pp. 2131–2140. Available at: https://doi.org/10.1109/TCBB.2019.2911071.

OpenStreetMap contributors (2017) 'Planet dump retrieved from https://planet.osm.org'.

Ordnance Survey (2023) 'Ordnance Survey history'. Available at: https://www.ordnancesurvey.co.uk/about/history.

Österberg, T. (2001) 'Cadastre as a Tool for Implementing Land Policy', in. *FIG Working Week 2001*. Available at: https://www.fig.net/resources/proceedings/fig_proceedings/korea/full-papers/pdf/session10/osterberg.pdf.

Pedregosa, F. *et al.* (2012) ‘Scikit-learn: Machine Learning in Python’, *Journal of Machine Learning Research*, 12.

Peters, M., Fudge, S. and Sinclair, P. (2010) ‘Mobilising community action towards a low-carbon future: Opportunities and challenges for local government in the UK’, *Energy Policy*, 38(12), pp. 7596–7603. Available at: https://doi.org/10.1016/j.enpol.2010.01.044.

Pietro, C., Marco, D. and Francesca, D.P. (2023) ‘Automatic identification of residential building features using machine learning techniques’, *Procedia Structural Integrity*, 44, pp. 1980–1987. Available at: https://doi.org/10.1016/j.prostr.2023.01.253.

Powell-Smith, A. (2016) ‘How to use Land Registry data to explore land ownership near you’, 14 March. Available at: https://anna.ps/blog/how-to-use-land-registry-data-to-explore-land-ownership-near-you.

Pullar, D. and Donaldson, S. (2022) ‘Accuracy Issues for Spatial Update of Digital Cadastral Maps’, *ISPRS International Journal of Geo-Information*, 11(4), p. 221. Available at: https://doi.org/10.3390/ijgi11040221.

QGIS Development Team (2023) *QGIS Geographic Information System*. QGIS Association. Available at: https://www.qgis.org.

Rahmatizadeh, S. *et al.* (2018) ‘A framework for selecting a fit-for-purpose data collection method in land administration’, *Land Use Policy*, 70, pp. 162–171. Available at: https://doi.org/10.1016/j.landusepol.2017.10.034.

Regnier, C. *et al.* (2018) ‘Quantifying the benefits of a building retrofit using an integrated system approach: A case study’, *Energy and Buildings*, 159, pp. 332–345. Available at: https://doi.org/10.1016/j.enbuild.2017.10.090.

Ren, C. and Bloemraad, I. (2022) ‘New Methods and the Study of Vulnerable Groups: Using Machine Learning to Identify Immigrant-Oriented Nonprofit Organizations’, *Socius: Sociological Research for a Dynamic World*, 8, p. 237802312210769. Available at: https://doi.org/10.1177/23780231221076992.

Segal, M. and Xiao, Y. (2011) ‘Multivariate random forests’, *WIREs Data Mining and Knowledge Discovery*, 1(1), pp. 80–87. Available at: https://doi.org/10.1002/widm.12.

Seyfang, G. (2010) ‘Community action for sustainable housing: Building a low-carbon future’, *Energy Policy*, 38(12), pp. 7624–7633. Available at: https://doi.org/10.1016/j.enpol.2009.10.027.

Sherman, G.E. (2008) *Desktop GIS: mapping the planet with open source tools*. Raleigh, N.C. ; Dallas, TX: Pragmatic Bookshelf (Pragmatic programmers).

Shrubsole, G. (2019a) *Who owns England? how we lost our green & pleasant land & how to take it back*. London: William Collins.

Shrubsole, G. (2019b) ‘Who owns the country? The secretive companies hoarding England’s land’. Available at: https://www.theguardian.com/environment/2019/apr/19/who-owns-england-secretive-companies-hoarding-land.

Slack, E. (2018) ‘Property Taxes’, *Societal Impacts of Tax*. Available at: https://responsibletax.kpmg.com/article/property-taxes#:~:text=In%20both%20developed%20and%20developing,value%20arising%20from%20government%20investment.

Stanley, V. (2023) 'How Does the World Bank See Land as Integral to Addressing Climate Change (12301) FIG Working Week 2023'. *c*, Orlando, Florida, USA, 28 June. Available at: https://fig.net/resources/proceedings/fig_proceedings/fig2023/papers/ps02/PS02_stanley_12301_abs.pdf.

Stapleton, M. (2010) 'Climate Change and the Cadastral Surveyor', in. FIG Publications. Available at: https://www.oicrf.org/documents/40950/43224/Climate+change+and+the+cadastral+surveyor.pdf/096d86a4-b3b4-4a84-0b4b-e8014a53ccbb?t=1510220145868.

Sturrock, H.J.W. *et al.* (2018) 'Predicting residential structures from open source remotely enumerated data using machine learning', *PLOS ONE*. Edited by Q. Gu, 13(9), p. e0204399. Available at: https://doi.org/10.1371/journal.pone.0204399.

Sugiyama, N. and Takeuchi, T. (2008) 'Local Policies for Climate Change in Japan', *The Journal of Environment & Development*, 17(4), pp. 424–441. Available at: https://doi.org/10.1177/1070496508326128.

Technology Strategy Board (2013) *Scaling up retrofit of the nation's homes: Competition for collaborative R & D funding*. London, UK.

The Guardian (2006) 'UK fights against tide on data directive', *The Guardian*, 27 July. Available at: https://www.theguardian.com/technology/2006/jul/27/epublic.guardianweeklytechnologysection.

Toms, K.N. and Lewis, S. (1974) 'Cadastral surveying and land registration: A comparative analysis of the English and Tasmanian systems', *Australian Surveyor*, 26(4), pp. 243–264. Available at: https://doi.org/10.1080/00050326.1974.10440762.

UK Parliament (1845) *Inclosure Act 1845*. Available at: https://www.legislation.gov.uk/ukpga/Vict/8-9/118/contents.

UK Parliament (2008) *The Climate Change Act 2008*. Available at: http://www.legislation.gov.uk/ukpga/2008/27/contents.

UK Parliament (2018a) *European Union (Withdrawal) Act 2018*. Available at: https://www.legislation.gov.uk/ukpga/2018/16/contents/enacted.

UK Parliament (2018b) *The INSPIRE (Amendment) (EU Exit) Regulations 2018*. Available at: https://www.gov.uk/eu-withdrawal-act-2018-statutory-instruments/the-inspire-amendment-eu-exit-regulations-2018#:~:text=This%20instrument%20amends%20The%20INSPIRE,states%20to%20operate%20national%20SDIs.

Ürge-Vorsatz, D. and Metz, B. (2009) 'Energy efficiency: how far does it get us in controlling climate change?', *Energy Efficiency*, 2(2), pp. 87–94. Available at: https://doi.org/10.1007/s12053-009-9049-7.

Van Bronkhorst, B. (2023) 'Land', *Understanding Poverty*. Available at: https://www.worldbank.org/en/topic/land#:~:text=The%20World%20Bank%20invests%20in,develop%20digital%20land%20administration%20services.

Van Rossum, G. and Drake, F.L. (2009) *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.

Vejchodská, E. *et al.* (2022) 'Bridging land value capture with land rent narratives', *Land Use Policy*, 114, p. 105956. Available at: https://doi.org/10.1016/j.landusepol.2021.105956.

Viscusi, G., Castelli, M. and Batini, C. (2014) ‘Assessing Social Value in Open Data Initiatives: A Framework’, *Future Internet*, 6(3), pp. 498–517. Available at: https://doi.org/10.3390/fi6030498.

Williamson, I.P. (1985) ‘Cadastres and Land Information Systems in Common Law Jurisdictions’, *Survey Review*, 28(217), pp. 114–129. Available at: https://doi.org/10.1179/sre.1985.28.217.114.

Williamson, I.P. (ed.) (2010) *Land administration for sustainable development*. 1st ed. Redlands, Calif. : [LaVergne, TN]: ESRI Press Academic ; Distributed to the trade in North America [by] Ingram Publisher Services.

Winkler, T. (2013) ‘At the Coalface: Community–University Engagements and Planning Education’, *Journal of Planning Education and Research*, 33(2), pp. 215–227. Available at: https://doi.org/10.1177/0739456X12474312.

‘Written Evidence (LUE0027) to the House of Lords’ (2022). House of Lords. Available at: https://committees.parliament.uk/writtenevidence/108111/html/.

Zadik, M. (2013) ‘The Nottingham Inclosure Act 1845: Why it is important today’. Available at: http://www.ng-spaces.org.uk/the-nottingham-inclosure-act-1845-why-it-is-important-today/.

Zevenbergen, J. *et al.* (2013) ‘Pro-poor land administration: Principles for recording the land rights of the underrepresented’, *Land Use Policy*, 31, pp. 595–604. Available at: https://doi.org/10.1016/j.landusepol.2012.09.005.