Preprint Cambridge Centre for Computational Chemical Engineering ISSN 1473 – 4273

Transforming building retrofits: Linking energy, equity, and health insights from The World Avatar

Jiying Chen[1], Jiaru Bai[1], Jieyang Xu[1], Feroz Farazi[1,2], Sebastian Mosbach[1,2,3], Jethro Akroyd[1,2,3], Markus Kraft[1,2,3]

released: November 18, 2024

[1] Department of Chemical Engineering and Biotechnology
University of Cambridge
Philippa Fawcett Drive
Cambridge, CB3 0AS
United Kingdom

[2] CARES
Cambridge Centre for Advanced Research and Education in Singapore
1 Create Way
CREATE Tower, #05-05
Singapore, 138602

[3] CMCL
No. 9, Journey Campus
Castle Park
Cambridge
CB3 0AX
United Kingdom

Preprint No. 331

UNIVERSITY OF CAMBRIDGE

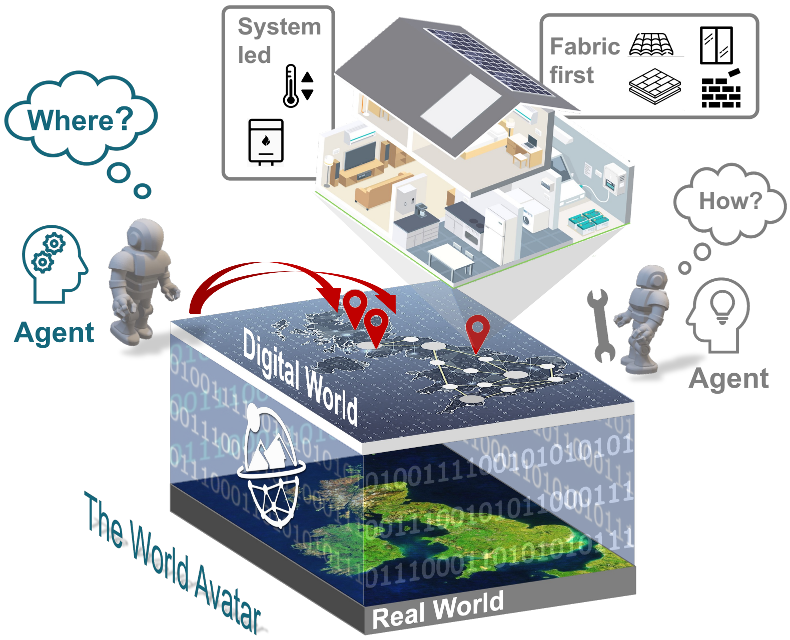
Keywords: building energy retrofit, multi-scale strategy, policy equity, interoperability

Edited by
Computational Modelling Group
Department of Chemical Engineering and Biotechnology
University of Cambridge
Philippa Fawcett Drive
Cambridge, CB3 0AS
United Kingdom

E-Mail: mk306@cam.ac.uk
World Wide Web: https://como.ceb.cam.ac.uk/

Abstract

The upgrading of energy-inefficient buildings is a critical part of the energy transition. Holistic analyses that foster informed and equitable policy interventions require interoperable data. We apply a principled approach that leverages The World Avatar to create a virtual knowledge graph underpinned by machine-understandable data representations. This approach provides a common terminology to integrate heterogeneous data sources to support multi-scale analysis of building energy retrofit options. We consider a case study in the UK based on the holistic analysis of household-level energy performance data, public health statistics and socio-economic metrics across geographic hierarchies. The analysis identifies regions with critical retrofit necessities, revealing disparities between these imperatives and extant policy levers. Granular retrofit targets are proposed to optimise resource allocation to the most vulnerable areas. Bespoke retrofit strategies are developed for 14.4 million households in the UK, providing actionable insights to support the targeted application of 'fabric-first' or 'system-led' retrofit pathways.



Highlights

- Machine-understandable representation of energy and environmental data
- Implemented multi-criteria decision-making framework
- Formulated national building retrofit strategy and analysed policy disparities
- Household-level 'fabric first' vs 'system-led' retrofit recommendations
- Presented insights at county, local authority, and constituency levels

Contents

1 Introduction

Climate change and the need to achieve sustainability is a worldwide cause for concern. The buildings sector is a major contributor to climate change [64], responsible for 37% of global energy-related carbon emissions [59]. Of this total, emissions from energy for heating, cooling, hot water and other operational needs, account for about 70% of total emissions of buildings [38]. This highlights the potential for immediate and notable progress in decarbonisation by reducing the energy consumption of buildings.

The urgency of reducing emissions from buildings is particularly acute in the UK. An estimated 9.6 million homes require retrofitting to meet energy efficiency standards and align with climate goals [27]. Two solutions are widely discussed: a *fabric-first* approach, which focuses on enhancing structural insulation [19, 29], and a *system-led* approach, which aims to improve heating and cooling systems [58]. The effectiveness of one versus the other depends on specific building characteristics like type, size, and infrastructure.

The UK government has initiated programs to reduce operational carbon. The Energy Company Obligation (ECO) scheme [33] was launched in 2013 as a flagship policy to subsidise the cost of retrofitting homes with improved insulation and energy-efficient heating systems. However, significant gaps remain in the efficacy of the policy and regional equity [32, 34]. Its implementation has been criticised for uneven reach, particularly in rural and remote areas [35]. These shortcomings motivate the need for tailored retrofitting strategies that consider regional and stakeholder-specific challenges [2], optimising resource allocation to emphasise energy justice [32].

Although the need for tailored strategies is increasingly recognised [18], determining what actions to take and in what priority order remains challenging because it requires analysis of heterogeneous data describing multiple factors. Energy Performance Certificates (EPC) are widely used as a source of such data. In the UK, an EPC is issued when selling, renting, building and sometimes modifying a property [42]. Each EPC provides a rating that condenses energy-related data – including information about the structure, insulation quality and heating system in a property – into a single score. At a household level, individual EPC data has been employed as the input to optimise retrofitting strategies for residential buildings and clusters of similar properties through numerical and machine learning techniques [3, 51, 62]. At a wider scale, EPC data has been used to assess decarbonisation potential and formulate tailored efficiency-oriented retrofit strategies at district and city scales in Stockholm and Oslo [48, 50]. However, efficiency-oriented strategies struggle to capture the diverse socio-economic contexts and energy use patterns that are crucial for comprehensive assessment.

Acknowledging the limitations of efficiency-oriented strategies, policymakers are adopting the alleviation of fuel poverty as an alternative guiding principle. In the UK, a household is considered to be in fuel poverty if it has an EPC score less than 69 (*i.e.*, band D or below), and the householders are left with a residual income below the official poverty line after heating the property [45]. The consideration of fuel poverty in retrofit analyses provides a broader perspective on energy vulnerability and has gained traction internationally. Researchers have emphasised systematic interventions at the national scale to address energy poverty in the USA [9]. While in the UK and Ireland, researchers have commented on retrofits as a policy tool to mitigate fuel poverty [39, 52], assessed the impact of adopting heat pumps on inequality [54]. More locally, regional-scale studies have considered the design and implementation of retrofits such as improving insulation, integrating renewable energy and upgrading energy-efficient heating systems to reduce fuel poverty [12, 26, 61].

Despite the adoption of fuel poverty-oriented strategies, the sole reliance on fuel poverty overlooks regional differences in climate conditions, housing stock and demographics. Areas with similar fuel poverty rates may face different challenges, resulting in different urgencies for energy upgrades. For instance, cold home environments in regions with pronounced ageing populations [30, 37] significantly exacerbate chronic obstructive pulmonary disease (COPD) [60] and cardiovascular disease (CVD) [36], compared to younger communities. Several studies have shown that improvements in energy efficiency also have co-benefits for health, and have considered health-oriented analyses. For example, researchers have shown that retrofits can reduce cold-related mortality and medical burdens at a city-scale [6] and that sustainable buildings improve health for vulnerable groups and reduce social inequalities [49], and have advocated the prioritised use of public health evidence in strategy formulation [14]. It is clear that a more nuanced approach to energy retrofit strategies has the potential to achieve better outcomes beyond just reduced energy use. However, most studies focus on building, community and urban scales, while national-level analyses remain sparse and typically consider factors in isolation rather than adopting a multi-criteria approach.

One key barrier to a multi-criteria approach is the interoperability of data. Inconsistencies in data completeness, update frequencies, and formats impede large-scale automated analysis, necessitating substantial manual efforts [41]. One means to enhance interoperability and scalability is to use knowledge models (*i.e.*, ontologies) and technologies from the Semantic Web to provide a general and scalable way to connect disparate data sources and offer a uniform machine-understandable representation of the data. This idea forms a key element of the World Avatar (TWA) approach [1, 40], which aims to enable a digital ecosystem equipped with machine-understandable data representations and autonomous computational agents. A provenance framework can be used to semantically annotate inputs and outputs from agents to form chains of dependent information, enabling automated information cascading in response to new information [7]. The underlying knowledge models facilitate interoperability and knowledge retention by explicitly codifying domain expertise.

The **purpose of this paper** is to develop a framework for the multi-criteria assessment of what retrofit actions to take and in what priority order. The study leverages ideas and developments from TWA to create a virtual knowledge graph supported by machine-understandable data representations that integrate disparate data sources, facilitating a holistic approach. Energy consumption, performance efficiency, fuel poverty, and public health data from the UK are analysed as a case study. The analysis shows a disconnect between household-level retrofit priority and current policy actions. The analysis includes an assessment of the relative benefits of adopting a 'fabric-first' or 'system-led' approach for each of the 14.4 million homes with EPC data in England, with output provided in the form of aggregated insights to support prioritised and equitable resource allocation.

2 Method

2.1 Knowledge-based data representation

Developing a comprehensive building retrofit strategy necessitates the integration of diverse data sources encompassing energy performance, energy consumption, geospatial, public health, and socio-economic domains. This section elucidates how these disparate national-scale data sources were unified in a machine-understandable format, enabling autonomous agents to perform seamless extraction and computation, supporting informed decision-making for building retrofit initiatives in UK. The data sources used in this study are summarised in **Table 1**. All data used in this work, except for the prevalence of COPD and CVD, are obtained from governmental sources as of December 2023. The COPD and CVD data are sourced from GP practice records updated in April 2024. Figures showing data are provided in Appendix A.1.

Table 1: *Summary of data sources represented in this work.*

Name	Description	Format	Frequency
Building location and geometry [28]	Includes Topographic Identifier (TOID), footprint polygons and building height.	.shp	Biennial
EPC [25]	Assessments of heating, insulation, lighting and hot water efficiency.	API	Biannual
Electricity consumption [23]	Average annual per household at the Lower Layer Super Output Area (LSOA) level.	.xlsx	Annual
Gas consumption [24]	Average annual per household at the LSOA level.	.xlsx	Annual
Fuel poverty rate [22]	Proportion of households unable to afford adequate heating at the LSOA level.	.xlsx	Annual
Prevalence of COPD [8]	Record from GP practices published by NHS Digital at the Middle Super Output Areas (MSOA) level.	.xlsx	Annual
Prevalence of CVD [8]	Record from GP practices published by NHS Digital at the MSOA level.	.xlsx	Annual
ECO policy strength [21]	Count of government-subsidised retrofit measures at the local authority level.	.xlsx	Annual
LSOA boundaries [44]	Digital boundaries of LSOAs.	.shp	10 years

These data reside in silos, with diverse formats, granularities, and limited interoperability. To achieve seamless data integration, enable machine discoverability and autonomous processing by computational agents, we adopt ontology-based representations. An ontology provides a high-level knowledge-based conceptual model that defines relationships between entities, providing a standardised, machine-readable approach to represent and connect data within and across domains.

Figure 1 summarises the ontologies used to integrate data across energy, geospatial, public health, and socio-economic domains in this study. In order to maximise interoper-

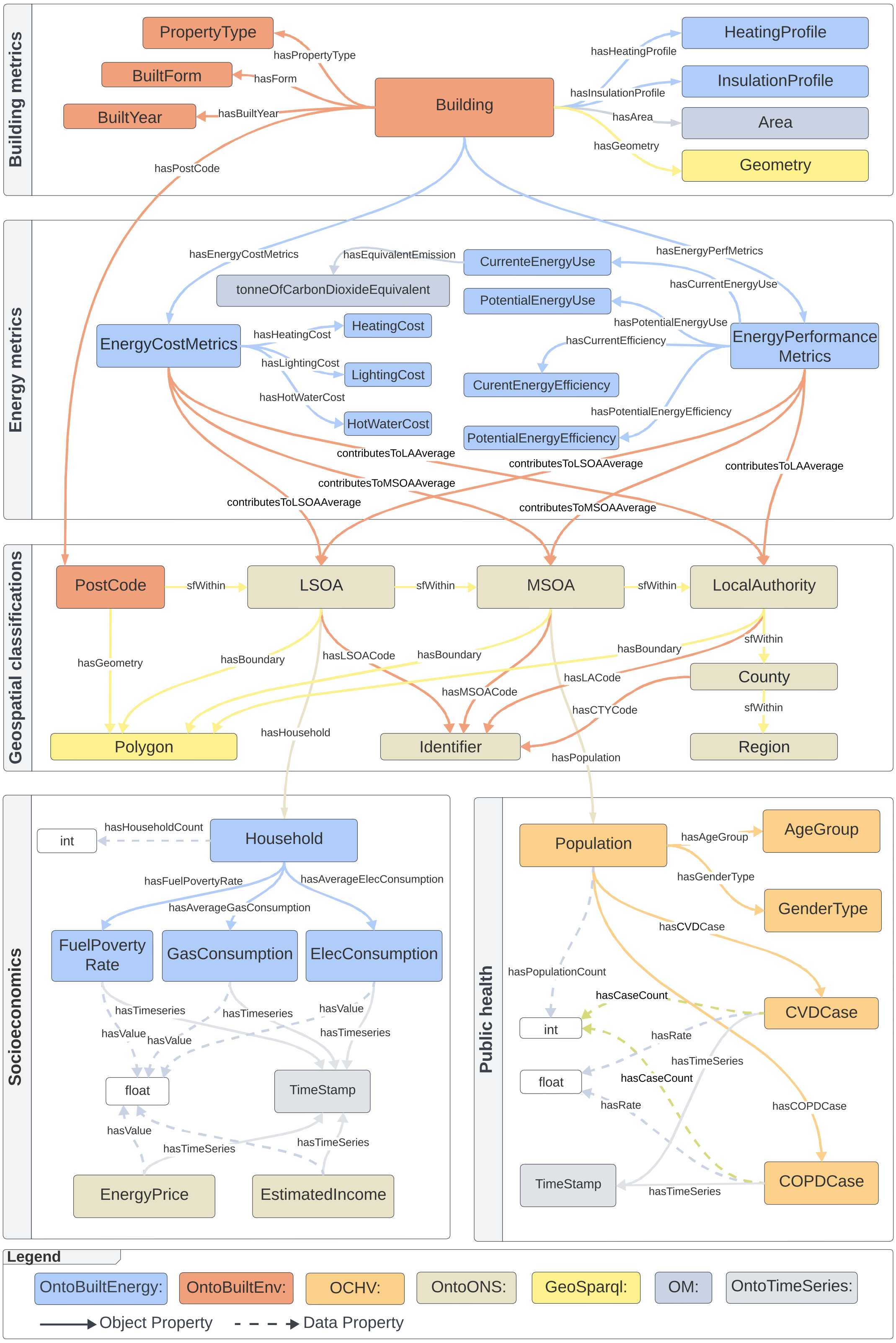


Figure 1: *Ontological framework used to integrate building, energy, geospatial, socioeconomic and public health data in the multi-criteria analysis.*

ability, the entities are instantiated using existing ontologies wherever possible, such as GeoSPARQL [10] and ONS Statistical Entity Ontology (OntoONS) [46] for geospatial data, and the Ontology of Consumer Health Vocabulary (OCHV) [4] for public health statistics. Additionally, two new domain-specific ontologies, OntoBuiltEnergy and OntoBuiltEnv are introduced to provide consistent semantic representations for building energy consumption and built environment data, respectively. The alignment between the ontologies ensures that the data is represented in a standardised way, facilitating seamless integration and analysis. The namespace references for these ontologies, which establish unique web-based URIs to ensure consistent referencing of terms across the World Wide Web, are listed in Appendix A.2.

An Ontology-Based Data Access (OBDA) solution [63] was used to represent the data as a virtual knowledge graph sitting on top of the raw data. The virtual knowledge graph maps the heterogeneous data to the ontology concepts defined in Figure 1. This mapping process enables querying through machine-readable semantic descriptions, simplifying the process of navigating the complex relationships in the underlying databases and facilitating interoperable data exchanges and intelligent decision-making [43]. The details of OBDA declarations and example semantic queries can be found in Appendix A.3. In addition to creating machine-readable representations, the completed interoperable ecosystem involves digesting heterogeneous data, providing querying endpoints, and enabling collaborative computation. The workflow and associated implementation tools are detailed in Appendix A.4.

2.2 Multi-criteria analysis framework

We propose a multi-criteria analysis framework that uses the knowledge graph to evaluate and rank areas by calculating a Building Retrofit Priority Index (BRPI) for retrofit prioritisation. **Figure 2** illustrates the four preparatory steps for calculating the BRPI.

The first step is to extract and align data to support analysis of whether retrofitting is required. The data listed in **Table 2** are retrieved from the knowledge graph described in Section 2.1 using cross-domain queries that retrieve data represented using the different ontologies. Building-level data are aligned with broader geographic regions through geospatial relationships defined in the ontology. Each region is associated with a unique identifier that is used to enable links to other concepts such as energy consumption and COPD records. These concepts are further connected to specific numerical values through data properties, facilitating the integration and access of different concept data. When querying a building's energy performance using its identifier, the system can simultaneously retrieve its geometric shape and administrative boundaries from the knowledge graph. This method ensures that cross-scale and cross-domain data are accessed and analysed coherently and comprehensively.

The second step is to standardise the numerical values of the extracted criteria, which are then classified into two categories: benefit and cost. Benefit criteria reflect positive outcomes where higher values are preferable, such as improved energy performance. In contrast, cost criteria represent factors where lower values are desirable, indicating reduced burdens like lower energy consumption or decreased fuel poverty rates.

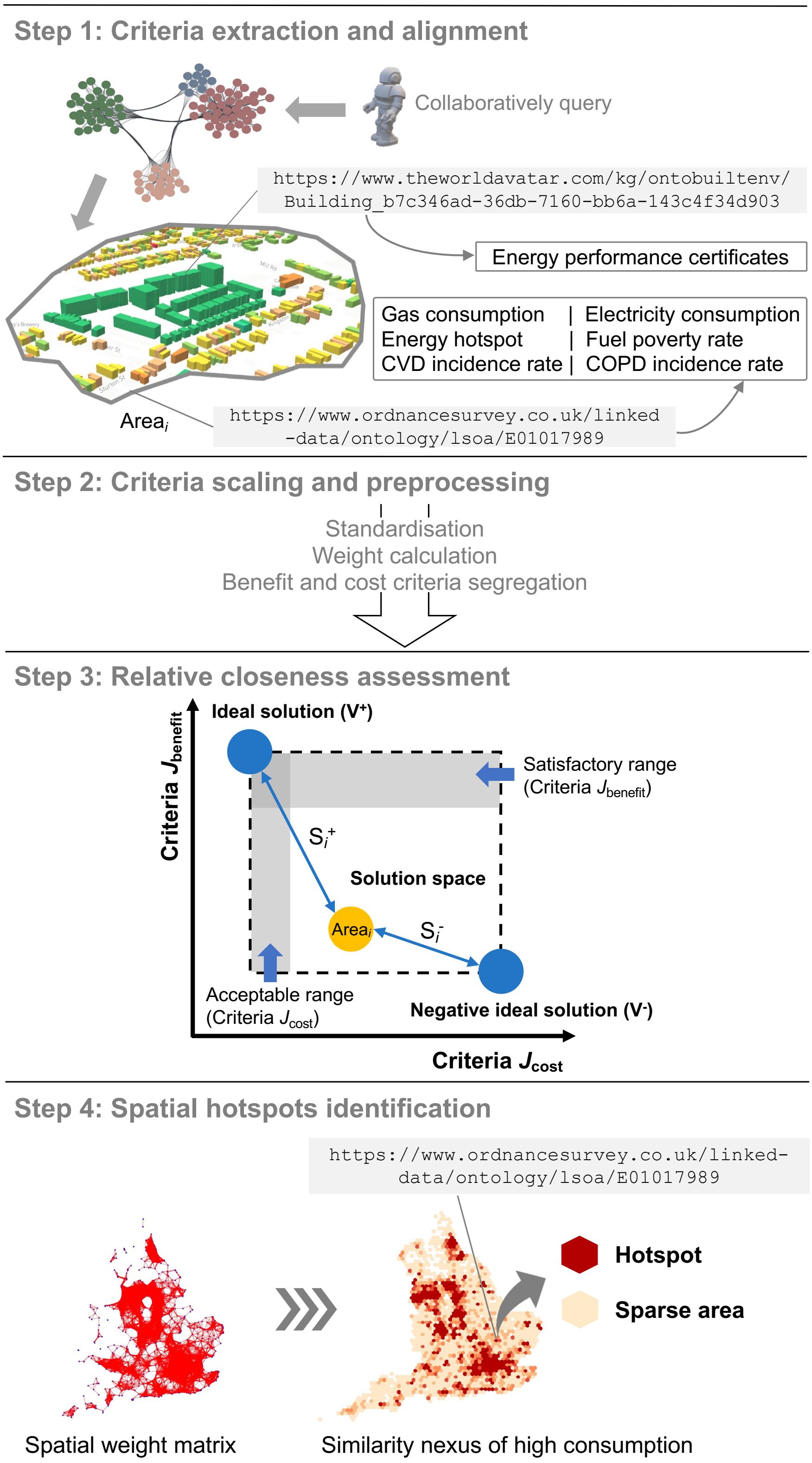


Figure 2: *Four steps in the framework of energy retrofit prioritisation: entity extraction and granularity alignment, criteria processing, assessment of relative closeness to ideal scenarios, and geospatial hot spot identification.*

Table 2: *Criteria and matching attributes used in the multi-criteria analysis framework.*

Name	Symbol	Namespace	Granularity	Type
Aggregated energy performance efficiency	E_{eff}	OntoBuiltEnergy	LSOA	Criteria
Energy consumption	E_{cons}	OntoBuiltEnergy	LSOA	Criteria
High energy consumption spatial hot spot	χ_i	OntoBuiltEnergy	LSOA	Criteria
Fuel poverty rate	R_{FP}	OntoBuiltEnergy	LSOA	Criteria
Disaggregated COPD prevalence	R_{COPD}	OCHV	LSOA	Criteria
Disaggregated CVD prevalence	R_{CVD}	OCHV	LSOA	Criteria
Dwelling energy performance efficiency	$\hat{E}$	OntoBuiltEnergy	Dwelling	Attribute
Dwelling geometry	G_{dwelling}	GeoSPARQL	Dwelling	Attribute
Postcode	P_{dwelling}	OntoBuiltEnv	Dwelling	Attribute
LSOA boundary	G_{LSOA}	GeoSPARQL	LSOA	Attribute
LSOA code	C_{LSOA}	OntoONS	LSOA	Attribute
MSOA boundary	G_{MSOA}	GeoSPARQL	MSOA	Attribute
MSOA code	C_{MSOA}	OntoONS	MSOA	Attribute

The third step considers the average energy performance efficiency, energy consumption, the rate of fuel poverty, the rate of cardiovascular disease, and the rate of chronic obstructive pulmonary disease in each LSOA. This step determines the ideal and negative-ideal solutions and quantifies the relative closeness of each geographic area (in this case each LSOA) to these scenarios. The ideal solution is defined as the combination of the maximum observed benefit criteria and minimum observed cost criteria, while the negative-ideal solution is defined as the opposite. The ideal solution presents a situation where energy performance is maximised, while energy consumption, fuel poverty, and the prevalence of cardiovascular and respiratory diseases are minimised. In contrast, the negative-ideal solution reflects the worst-case conditions, characterised by inefficiency and high burdens across these criteria. The relative closeness of each area to the ideal and negative-ideal solutions is calculated. Full details of the calculation methods are provided in Appendix A.5.1.

The fourth step identifies energy consumption hot spots. These are defined as areas with significant spatial clustering of high energy (gas and electricity) usage that correlate with carbon emissions. Targeting these clusters is essential for the consistent and effective implementation of retrofit strategies across affected regions. The hot spot identification is performed using Local Indicators of Spatial Association (LISA) [31], which measures spatial auto-correlation to detect clusters of high energy consumption. The mathematical formulation and significance testing process are detailed in Appendix A.5.2.

Finally, the BRPI is calculated by combining the relative closeness to the negative-ideal solution with a weighted term that reflects whether an area has also been identified as an energy consumption hot spot. See Appendix A.5.3. High values of the BRPI indicate higher priority for retrofit. In this manner, areas that are identified as hot spots are assigned higher retrofit priority, even if their relative closeness to the ideal solution is similar to that of other areas. This adjustment ensures that high-consumption clusters are strategically prioritised to achieve energy efficiency goals.

2.3 Decision-making for household retrofit measures

The final step in the analysis is to assess the effectiveness of alternative household-level retrofit options to improve energy performance. Case-Based Reasoning (CBR) [55] and Incremental Analysis [57] are used to suggest the most appropriate strategy for each candidate dwelling. The assessment considers two alternative approaches: the *fabric-first* strategy, which aims to enhance the thermal efficiency of the envelope (walls, windows, roof, and floor) of a building, and the *system-led* strategy, which seeks to improve the efficiency of the systems (heating, hot water, and lighting) within a building. Each strategy is associated with specific performance indicators that are described in Appendix A.5.4.

The inputs to the assessment are retrieved from the knowledge graph described in Section 2.1. They comprise quantitative values, such as energy consumption, thermal efficiency, and the number of heated rooms, alongside qualitative descriptions such as the efficiency evaluation of the walls, windows and roof, which can be converted into numerical indicators for analysis.

The CBR is applied to each building with an energy performance rating below 69 (*i.e.*, band D or below), which is the threshold applied when assessing fuel poverty in the UK [45]. For each such building, the CBR identifies similar exemplar buildings that have better energy performance. The features used to determine similarity include the building form (*e.g.*, structure types like detached or semi-detached), the building usage (such as residential or commercial), the number of heated rooms, and the total floor area. The CBR evaluates whether the differences in the fabric or system of the exemplar building are sufficient to justify a particular retrofit approach. A fabric-first strategy is typically selected when the envelope of the candidate building significantly lags behind that of the exemplar building, while a system-led approach is chosen if the efficiencies of the systems are notably lower than for the exemplar building. Full details of the criteria and decision-making process are given in Appendix A.5.4.

If the CBR fails to indicate a definitive strategy, a decision tree model [17] is used to establish a mapping between the current energy efficiency and the combined feature set, comprising fabric, system, and metadata features. Details of the implementation are given in Appendix A.5.4. Subsequently, taking the current insulation and system conditions of each retrofit candidate as the baseline, separate adjustments are made to both the fabric and system features of the retrofit candidate. The adjusted conditions are input into an energy efficiency model to determine the incremental impact on the overall efficiency resulting from the improvements in fabric and system features, respectively. The implementation of the energy efficiency model is summarised in Appendix A.5.4. This what-if scenario analysis helps to identify the marginal benefits of enhancing the fabric and system performance based on the state of each building, enabling automated bespoke building-level identification of the retrofit pathway that would deliver the most substantial improvement.

3 Results and Discussions

The results of the BRPI calculations are compared with the outcomes achieved via the current implementation intensity of the Energy Company Obligation (ECO) policy. The comparison identifies the areas where retrofit actions are most urgent but currently lack sufficient policy intervention. The size of the implementation gap is quantified in terms of the number of additional installations that are required in vulnerable areas to achieve policy equity. The effectiveness of fabric-first and system-led measures is assessed at the household level, and the proportion of dwellings suitable for each approach summarised across different regions, different local authorities and different electoral constituencies.

3.1 National building retrofit strategy

The Energy Company Obligation (ECO) is a government scheme that mandates energy suppliers to deliver energy efficiency and heating measures to homes in Great Britain [21]. Key sub-schemes include the Home Heating Cost Reduction Obligation (HHCRO) [53], Carbon Emissions Reduction Obligation (CERO), and Carbon Saving Community Obligation (CSCO) [13]. Since its inception in January 2013, the scheme has undergone several phases: ECO1, ECO2, ECO HTH, ECO3, and ECO4. ECO4, running from 2022 to 2026, aims to reduce heating costs, alleviate fuel poverty, and contribute to carbon reduction targets [20]. The scheme has consistently invested more than £1 billion annually to support low-income and vulnerable households.

Figure 3 shows the regional discrepancies between the retrofit urgency described by the BRPI and the number of interventions made under the current implementation of the ECO. The upper map shows the installation density of ECO-supported energy retrofits at the local authority level, based on local council records [21]. The lower map shows the BRPI in each LSOA, with a resolution of approximately 1,000 households per LSOA. The values of the BRPI are aggregated to the local authority scale for comparison with the ECO statistics. The paired plots surrounding the maps show the BRPI (yellow markers) and normalised ECO installation intensity (blue markers) in each local authority, with grey lines connecting pairs of markers to indicate that they belong to the same local authority.

Regions with higher BRPI values indicate a significant urgency for building retrofits. Such regions are characterised by high household energy consumption, low building efficiency, and elevated fuel poverty rates, along with high incidences of respiratory and cardiovascular diseases. Examples include the West Midlands, South West, East of England, and South East. Notably, in the East of England and the South East, the installation intensity of ECO-supported energy upgrades for local authorities within these regions is about 20–25% of the national maximum. This mismatch suggests that regions in significant need of retrofits, as identified by energy efficiency, socioeconomic, and public health data, may not be receiving sufficient policy attention.

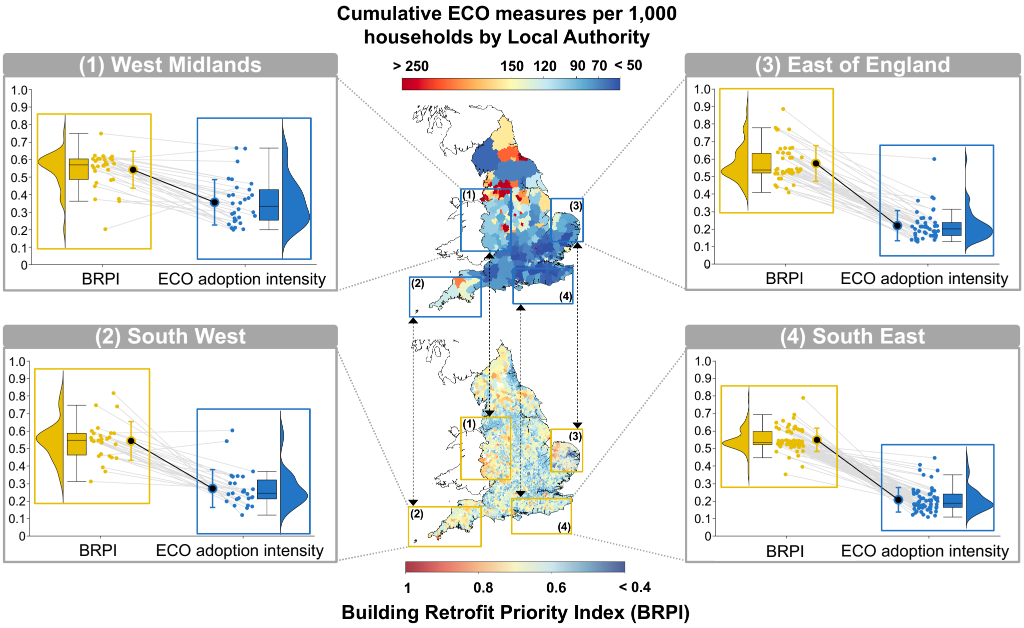


Figure 3: *Regional discrepancies between the Building Retrofit Priority Index (BRPI) and the energy policy implementation intensity.*

3.2 Regional variations in policy effects

The trajectory of retrofit installations under the ECO scheme is shown in **Figure 4**. The top panel shows that ECO4 has achieved retrofits at a rate of approximately 10 installations per 1,000 households annually. Although this rate has been maintained with relative consistency, achieving comprehensive retrofit coverage across the UK necessitates a carefully staged approach. It is proposed that the initial emphasis should be on addressing regional disparities in policy implementation and retrofit urgency, as identified in prior analyses. Such a strategy should systematically target the most pressing needs while ensuring an equitable distribution of resources across all regions.

A "what-if" scenario was developed to assess potential targeted interventions. Figure 4(b) shows the current situation. Figure 4(c) shows the impact of a scenario that maintains an average rate of 10 installations per 1000 households while increasing the rate of retrofit measures in vulnerable areas. The results indicate that an additional 40 energy retrofit measures per 1,000 households would be required in vulnerable regions to create a more balanced alignment between the intensity of national energy support policies and the spatial distribution of retrofit urgency, helping to mitigate existing regional disparities.

3.3 Household-level insights

Figure 5 shows the results of applying Case-Based Reasoning and Incremental Analysis to quantify the priority of system-led versus fabric-first retrofit approaches for households across England that have an EPC rating of D or below. In total, the analysis considered approximately 14.4 million households.

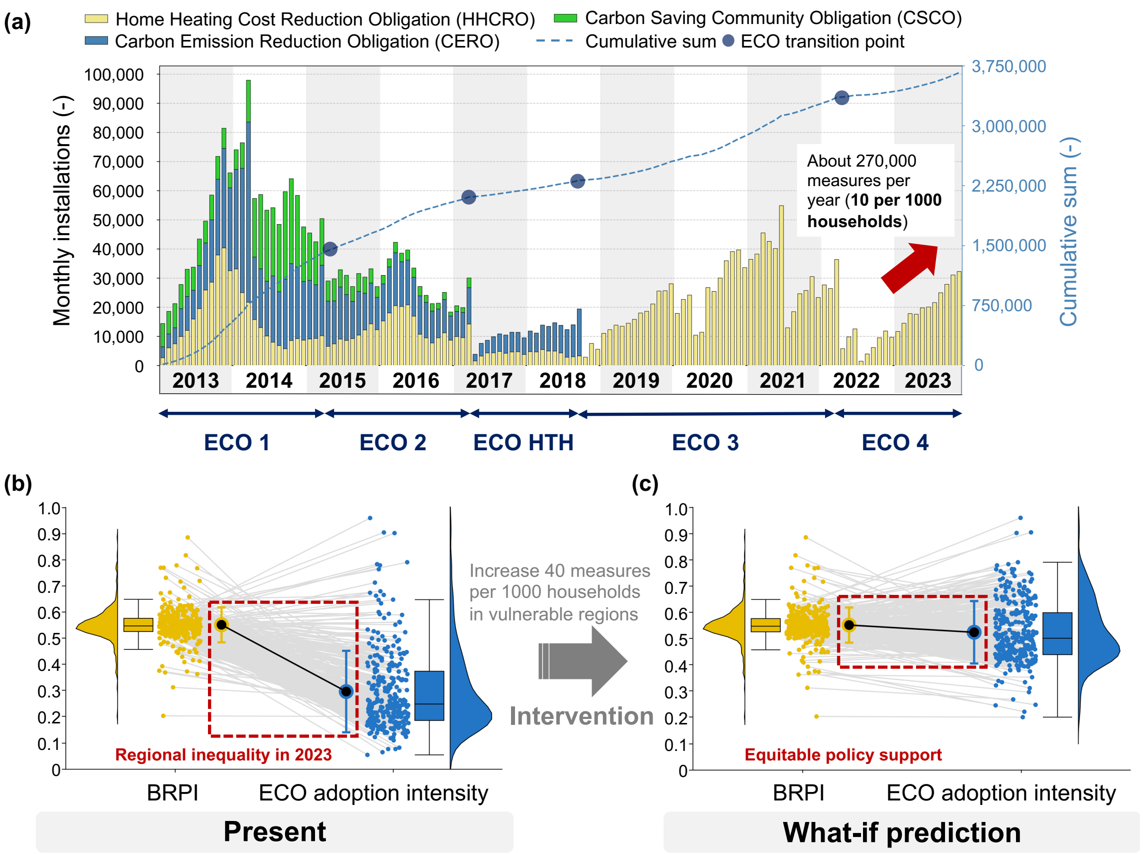


Figure 4: *(a) Monthly trends and cumulative sum of retrofit measures over the past 10 years, supported by the ECO scheme. Note that if a household receives retrofit measures in different months, they are counted in each respective month. (b) The current necessity for building retrofits, as quantified by the BRPI for all local authorities in England, compared with the normalised number of energy retrofit installations (where 1 represents the highest regional density of measures nationally). (c) The projected outcome of additional installation efforts in vulnerable regions to achieve equitable policy support.*

In the case of households with similar EPC ratings for their heating systems and fabric features, it was found that medium-sized houses tend to exhibit higher marginal benefits from fabric-first improvements, as insulation and air-tightness upgrades can significantly enhance energy efficiency. In contrast, in larger detached houses with more than 3 heated rooms and a floor area exceeding 150 m^2, it was found that system-led upgrades (such as installing heat pumps or more efficient heating systems) generally provided greater cost-effectiveness due to the increased heating demand. The findings for small homes and flats were more variable, with the optimal approach (fabric-first or system-led) depending on the specific efficiency characteristics of the building.

Across the country, our analysis revealed that 52.6% of retrofit candidates – approximately 7.6 million households –would achieve more effective energy performance improvements through fabric-first measures. This proportion increases in certain regions, rising to 55%

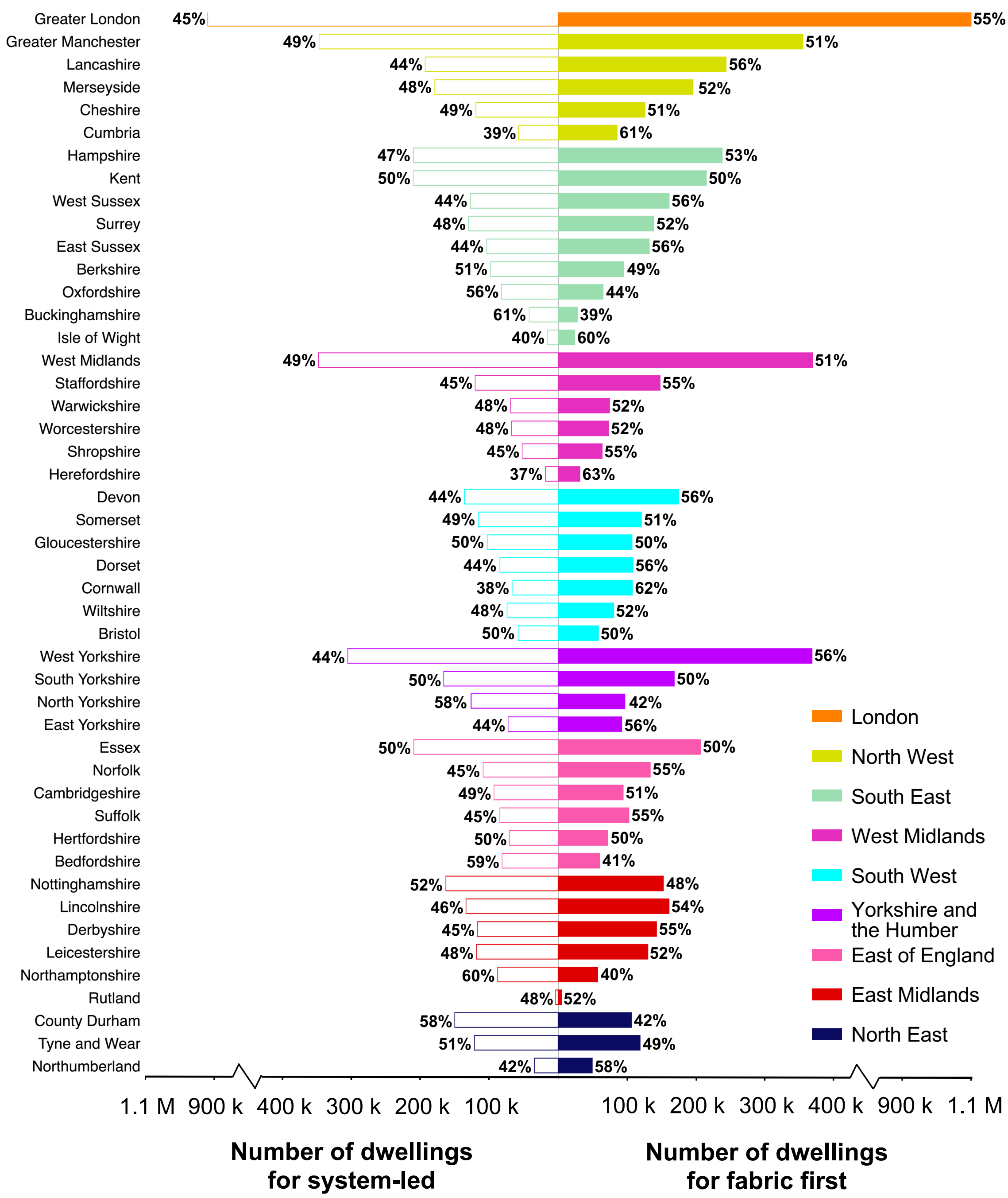


Figure 5: *Number of dwellings with an EPC rating of D or below that were found to be more suitable for 'fabric first' versus 'system-led' retrofit approaches in each county in England.*

among 2 million homes with an EPC rating of D or below in Greater London, and approaching two-thirds in coastal counties such as Cornwall and the Isle of Wight, and rural counties like Herefordshire. Notably, regions with a higher concentration of fabric-first candidates were also identified as having high retrofit urgency but lacking sufficient policy support, suggesting that these areas could benefit from increased government subsidies specifically aimed at fabric-first retrofits. These findings emphasise the importance of tailored retrofit strategies that address the unique requirements of each region, ensuring that funding for retrofits is both cost-effective and impactful across various parts of the UK.

Appendix B provides tabular summaries of the urgency of taking retrofit measures and the proportion of households that would benefit from fabric-first versus system-led interventions at a local authority and electoral constituency level. These findings provide evidence of the need to allocate tailored retrofit targets to local authorities, strengthening building energy-saving measures. The targets should align with the UK's governance framework to ensure efficient resource use. By prioritising regions with the greatest need, the national retrofit strategy can be deployed in phases, ensuring a fairer transition that not only improves long-term energy efficiency and reduces carbon emissions but also considers public well-being while achieving decarbonisation progress.

4 Conclusions

This study presents a machine-readable framework as part of The World Avatar to establish a common terminology for integrating cross-domain concepts and data on building-level energy performance, energy consumption, fuel poverty rate and public health records, enabling scalable and interoperable analysis to support informed decision-making in building energy retrofitting. Considering the UK as a case study, this system identifies spatial hot spots of high building energy use and quantifies energy retrofit urgency by assessing holistic vulnerability across energy, socioeconomic, and health dimensions. The resulting urgency map is compared with the intensity of current policy interventions to examine the equity of their action. Case-based reasoning and incremental analysis are applied to recommend the appropriate retrofit route for 14.4 million households by simulating the effectiveness of fabric-first and system-led measures in improving overall energy efficiency ratings. The recommended retrofit pathways are reported at aggregated local authority and electoral constituency levels to provide actionable insights.

The analysis shows that urgent retrofit needs span diverse regional contexts, including urban centres (*e.g.*, Birmingham Selly Oak and London-adjacent constituencies, Bexleyheath and Crayford) and sparsely populated areas (*e.g.*, Bexhill and Battle in East Sussex). Regions like the West Midlands, East of England, South East, and South West are shown to face acute retrofit necessity, yet only receive about 25% of the policy-support, measured as the number of energy-upgrade installations per 1,000 households, compared to the best-supported areas of the UK. This misalignment between retrofit demand and intervention underscores the need to strengthen region-specific policy initiatives.

Across the 14.4 million individual households considered by the analysis, fabric-first are recommended for 55% of Greater London retrofit candidates and up to two-thirds of candidates in coastal areas (*e.g.*, Cornwall, Devon) and rural counties (*e.g.*, Herefordshire, Cumbria). Conversely, system-led retrofits are more suited to newer housing, which is recommended for over 60% of retrofit candidates in commuter counties like Bedfordshire and Buckinghamshire, and industrial transition areas like Northamptonshire.

Ensuring an up-to-date and comprehensive coverage of high-quality data is critical for an accurate understanding of building energy efficiency and improvement pathways. The case study in this paper illustrates the capability of The World Avatar to provide a holistic framework that uses linked data protocols to enable computational interoperability between concepts and data related to geographical hierarchies and public statistics. This interoperability opens up the possibility of more comprehensive decision-making. The approach can be easily extended to integrate additional data as it becomes available, thereby further enhancing the understanding of real-world events, offering the potential for countries striving for scalable solutions to retrofit buildings while addressing intersecting social, environmental and health vulnerabilities.

Beyond energy retrofit strategies, this study lays the groundwork for future research that could further integrate technologies such as IoT sensor networks and generative AI with urban spatial cognition into The World Avatar. Future investigations could use scalable real-time models to enable dynamic simulation and evaluation of scenarios that describe the installation and configuration of energy infrastructure beyond buildings. These advancements would foster a more responsive approach to urban sustainability, promoting resilient and inclusive smart cities.

Acknowledgements

This research was supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme. Elements of this work were also supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/Y016076/1.

Nomenclature

AI Artificial Intelligence

BRPI Building Retrofit Priority Index

CBR Case-Based Reasoning

CERO Carbon Emissions Reduction Obligation

CityGML City Geography Markup Language

COPD Chronic Obstructive Pulmonary Disease

CSCO Carbon Saving Community Obligation

CVD Cardiovascular Disease

ECO Energy Company Obligation

EPC Energy Performance Certificate

GeoSPARQL Geographic Query Language for RDF Data

HHCRO Home Heating Cost Reduction Obligation

IoT Internet of Things

IRI Internationalised Resource Identifier

LOD Linked Open Data

LSOA Lower Layer Super Output Area

MSOA Middle Layer Super Output Area

OBDA Ontology-Based Data Access

OCHV Ontology of Consumer Health Vocabulary

OGC Open Geospatial Consortium

OM Ontology of units of Measure

ONS Office for National Statistics

OntoBuiltEnergy Ontology for Building Energy Consumption

OntoBuiltEnv Ontology for Built Environment Data

OntoONS ONS Statistical Entity Ontology

OntoTimeSeries Ontology of Time Series

OS Ordnance Survey

SPARQL SPARQL Protocol and RDF Query Language

SQL Structured Query Language

TOID Topographic Identifier

TWA The World Avatar

UPRN Unique Property Reference Number

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT-4 in order to enhance the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Data and code availability

The codes developed for this work are available under an open-source licence on GitHub in The World Avatar repository https://github.com/cambridge-cares/TheWorldAvatar. The datasets used in the work are freely available for download as per the references in the paper.

A Implementation details

A.1 Spatial and statistical overview of key data sources

Figure A.1 presents the average energy efficiency rating at the LSOA level, derived from aggregated household-level EPC data [25]. This aggregation is achieved through geographic mapping relationships defined in the knowledge graph, enabling the interoperable matching of data across different geographic hierarchies.

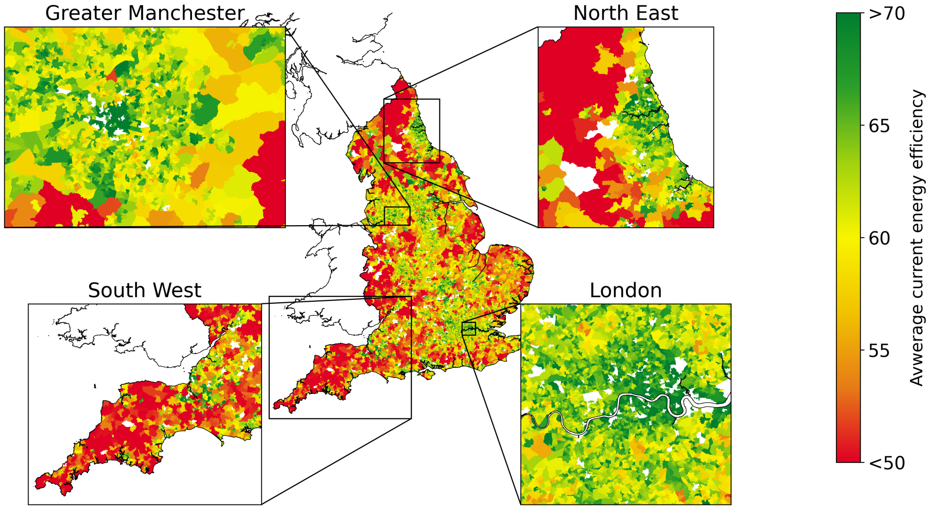


Figure A.1: *Average energy efficiency derived from the aggregation of energy performance certificates from individual properties.*

Figure A.2 shows the proportion of households experiencing fuel poverty in England in 2023. This data is directly sourced from the Department for Energy Security and Net Zero [22], providing insight into the distribution of energy affordability challenges.

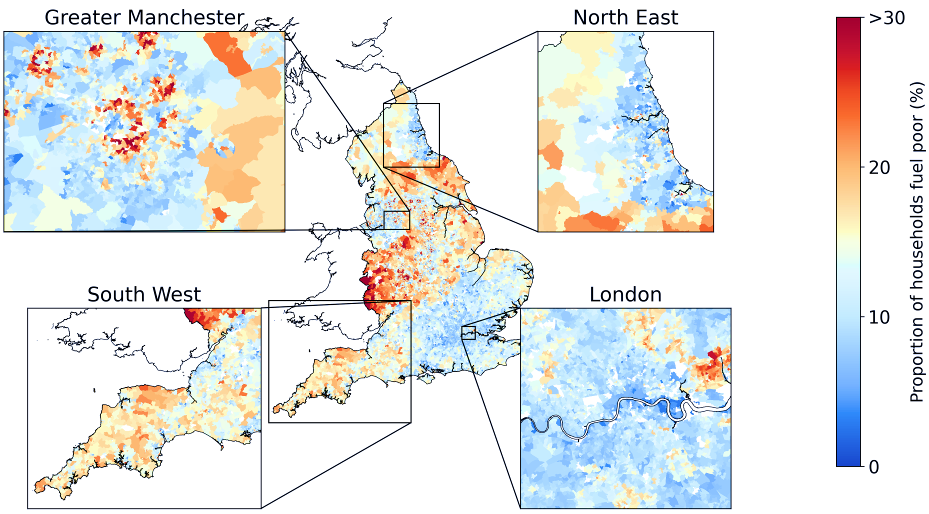


Figure A.2: *Proportion of households in fuel poverty across LSOAs in 2023.*

Figure A.3 shows the distribution of average annual electricity and gas consumption per household across the UK for 2022–2023. This data is sourced from the Department for Energy Security and Net Zero [23, 24], reflecting national energy usage patterns.

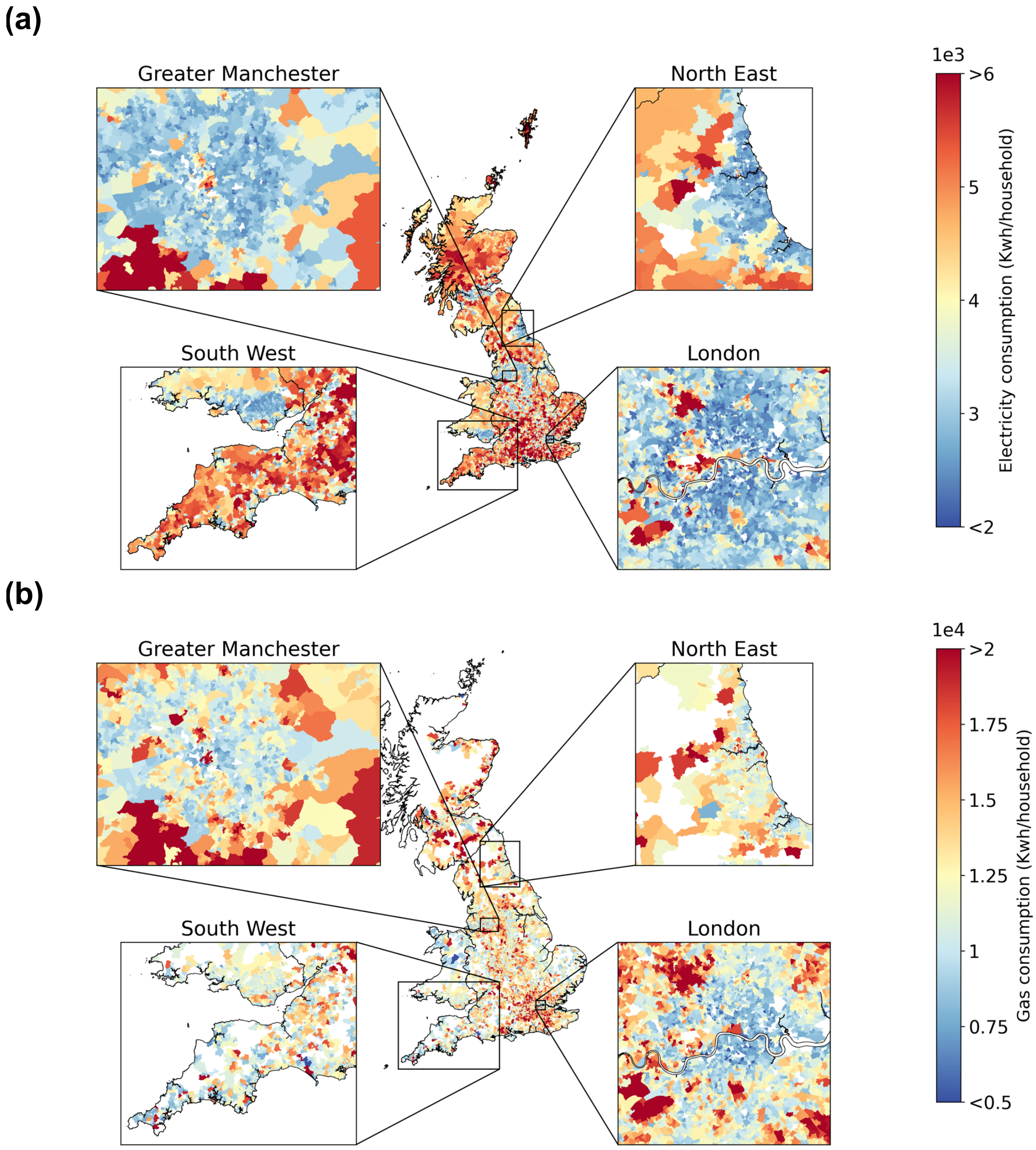


Figure A.3: *Annual average energy consumption per household (2022–2023): (a) electricity consumption (kWh/household); (b) gas consumption (kWh/household).*

Figure A.4 shows the cumulative probability and spatial distribution of incidence rates for COPD and the main types of CVD across the UK. These health metrics are sourced from GP practice records, as published by NHS Digital [8].

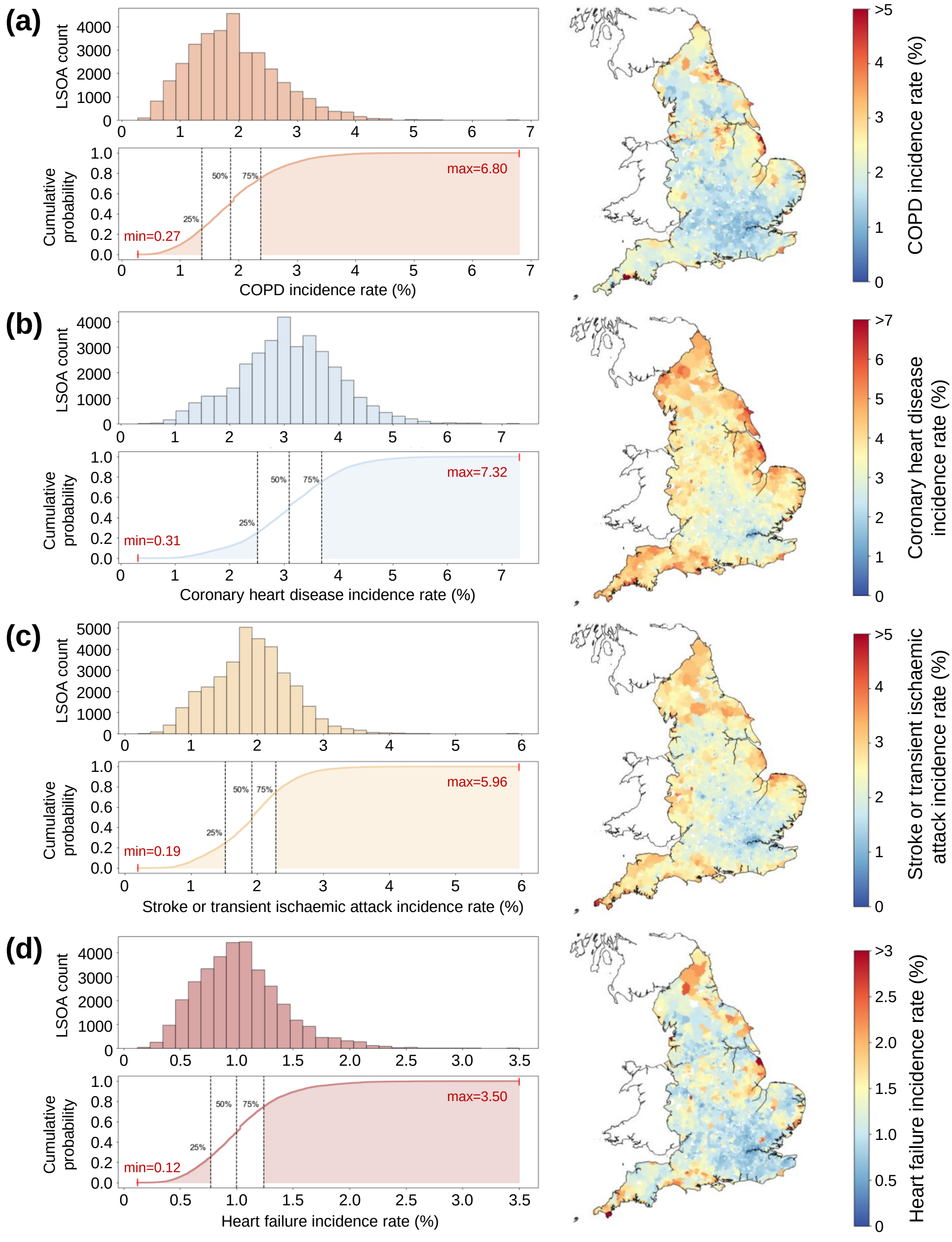


Figure A.4: *Cumulative probability and spatial distribution of incidence rates for COPD and CVD: (a) COPD incidence rate (%, range: 0.27%-6.80%); (b) CVD: Coronary heart diseaseincidence rate (%, range: 0.31%-7.32%); (c) CVD: Stroke or transient ischaemic attack incidence rate (%, range: 0.19%-5.96%); (d) CVD: Heart failure incidence rate (%, range: 0.12%-3.50%). Subfigures (b), (c), and (d) represent different forms of CVD.*

A.2 Namespaces

The following list defines the namespaces referenced in Figure 1 in the main text.

OntoBuiltEnv: <https://www.theworldavatar.com/kg/ontobuiltenv/>
OntoBuiltEnergy: <https://www.theworldavatar.com/kg/ontobuiltenergy/>
OCHV: <http://sbmi.uth.tmc.edu/ontology/ochv#CHV_Concept/>
OntoTimeSeries: <https://www.theworldavatar.com/kg/ontotimeseries/>
OntoONS: <https://www.ordnancesurvey.co.uk/linked-data/ontology/>
GeoSparql: <http://www.opengis.net/ont/geosparql#>
OM: <http://www.ontology-of-units-of-measure.org/resource/om-2/>
owl: <http://www.w3.org/2002/07/owl#>
rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
rdfs: <http://www.w3.org/2000/01/rdf-schema#>
time: <http://www.w3.org/2006/time#>
xsd: <http://www.w3.org/2001/XMLSchema#>

A.3 OBDA mapings

While ontologies provide a semantic layer for data representation, a lot of existing data is already available in formats designed for use with relational databases. Such relational databases are optimized for storage and efficient querying but contribute to the development of data silos because they lack the semantic interoperability necessary for integrated analyses across multiple datasets. Ontology-Based Data Access (OBDA) offers a solution to this limitation by providing a semantic interface to relational databases, enhancing interoperability and enabling more intuitive data access. An OBDA mapping typically comprises three key components: an ontology, mappings, and queries.

- The **ontology** represents the concepts and relationships that span the data, such as buildings, energy efficiency ratings, fuel poverty, and health conditions. These abstract concepts and relationships form the foundation for semantic descriptions, facilitating a more integrated and meaningful understanding of the data.
- The **mappings** define how these concepts correspond to the tables in the relational databases that store data.
- The **queries** allow questions structured using the semantic layer defined by the ontology to be answered using the underlying relational databases. For example, to identify buildings with poor energy efficiency located in areas with high fuel poverty and significant health issues related to cold homes.

Listing 1 shows the OBDA mapping used in this study. Listings 2 and 3 show example semantic SPARQL queries. Listing 4 shows an SQL equivalent to the query in Listing 3.

Listing 1: *OBDA mapping.*

```
[PrefixDeclaration]
twa:        https://www.theworldavatar.com/kg/
building:   https://www.theworldavatar.com/kg/ontobuiltenv/building
ontobuiltenergy:  https://www.theworldavatar.com/kg/ontobuiltenergy/
ocgml:      http://www.theworldavatar.com/ontology/ontocitygml/citieskg/OntoCityGML.owl#
os:         http://data.ordnancesurvey.co.uk/ontology/spatialrelations/
owl:        http://www.w3.org/2002/07/owl#
rdf:        http://www.w3.org/1999/02/22-rdf-syntax-ns#
xml:        http://www.w3.org/XML/1998/namespace
xsd:        http://www.w3.org/2001/XMLSchema#
obda:       https://w3id.org/obda/vocabulary#
rdfs:       http://www.w3.org/2000/01/rdf-schema#
geo:        http://www.opengis.net/ont/geosparql#

[MappingDeclaration] @collection [[
mappingId       Building-Class-Declaration
target      <https://www.theworldavatar.com/kg/Building> rdf:type owl:Class .
source      SELECT 1

mappingId       Asset-mapping-UUID-UPRN-via-Standard-Table
target      building:{uuid} a <https://www.theworldavatar.com/kg/Building> ; ontobuiltenv:hasUPRN
↪  {uprn}^^xsd:string ; ontobuiltenv:hasTOID {os_topo_toid}^^xsd:string .
source          SELECT ga_uuid.strval as uuid, mt."IDENTIFIER_1" as uprn, mt."IDENTIFIER_2" AS
↪  os_topo_toid
                FROM "citydb"."cityobject_genericattrib" ga_uuid
                JOIN "citydb"."cityobject_genericattrib" ga_toid ON ga_uuid.cityobject_id =
                ↪  ga_toid.cityobject_id
                                JOIN "public"."TOID_UPRN_MatchingTable" mt ON ga_toid.strval =
                                ↪  mt."IDENTIFIER_2"
                WHERE ga_uuid.attrname = 'uuid'
                AND ga_toid.attrname = 'os_topo_toid'

mappingId       Asset-mapping-UUID-Geometry
target      building:{uuid} a <https://www.theworldavatar.com/kg/Building> ; geo:hasGeometry
↪  "{wkt}"^^geo:wktLiteral .
source          SELECT ga_uuid.strval as uuid, public.ST_AsText(sg.geometry) as wkt
                FROM "citydb".cityobject_genericattrib ga_uuid
                JOIN "citydb".surface_geometry sg ON ga_uuid.cityobject_id = sg.cityobject_id
                WHERE ga_uuid.attrname = 'uuid'
                AND sg.parent_id IS NOT NULL;

mappingId       Asset-mapping-Building-EnergyProperty
target      building:{uuid} a <https://www.theworldavatar.com/kg/Building> ;
↪  building:hasEnergyEfficiency {CURRENT_ENERGY_EFFICIENCY}^^xsd:string ; building:hasEPCrating
↪  {CURRENT_ENERGY_RATING}^^xsd:string .
source          SELECT ga_uuid.strval as uuid, dep."CURRENT_ENERGY_EFFICIENCY",
↪  dep."CURRENT_ENERGY_RATING"
                FROM "citydb"."cityobject_genericattrib" ga_uuid
                JOIN "citydb"."cityobject_genericattrib" ga_toid ON ga_uuid.cityobject_id =
                ↪  ga_toid.cityobject_id
                                JOIN "public"."TOID_UPRN_MatchingTable" mt ON ga_toid.strval =
                                ↪  mt."IDENTIFIER_2"
                                JOIN "public"."Domestic EPC" dep ON mt."IDENTIFIER_1" = dep."UPRN"
                                WHERE ga_uuid.attrname = 'uuid'
                AND ga_toid.attrname = 'os_topo_toid'
```

Listing 1: *OBDA mapping (continued).*

```
mappingId       Asset-mapping-GenericBuildingDescription
target      building:{uuid} a <https://www.theworldavatar.com/kg/Building> ;
↪  ontobuiltenv:hasPostcode {POSTCODE}^^xsd:string ; ontobuiltenv:hasAddress1
↪  {ADDRESS1}^^xsd:string ; ontobuiltenv:hasPropertyType {PROPERTY_TYPE}^^xsd:string ;
↪  ontobuiltenv:hasBuiltForm {BUILT_FORM}^^xsd:string .
source      SELECT ga_uuid.strval as uuid, "public"."Domestic EPC"."POSTCODE", "public"."Domestic
↪  EPC"."ADDRESS1", "public"."Domestic EPC"."PROPERTY_TYPE", "public"."Domestic EPC"."BUILT_FORM"
            FROM "citydb"."cityobject_genericattrib" ga_uuid
            JOIN "citydb"."cityobject_genericattrib" ga_toid ON ga_uuid.cityobject_id =
            ↪  ga_toid.cityobject_id
            JOIN "public"."TOID_UPRN_MatchingTable" mt ON ga_toid.strval = mt."IDENTIFIER_2"
            JOIN "public"."Domestic EPC" ON mt."IDENTIFIER_1" = "public"."Domestic EPC"."UPRN"
            WHERE ga_uuid.attrname = 'uuid'
            AND ga_toid.attrname = 'os_topo_toid'

mappingId       Asset-mapping-EnergyCost
target      building:{uuid} a <https://www.theworldavatar.com/kg/Building> ;
↪  building:hasCO2EmissionsCurrent {CO2_EMISSIONS_CURRENT}^^xsd:string ;
↪  building:hasHeatingCostCurrent {HEATING_COST_CURRENT}^^xsd:string ;
↪  building:hasHotWaterCostCurrent {HOT_WATER_COST_CURRENT}^^xsd:string ;
↪  building:hasLightingCostCurrent {LIGHTING_COST_CURRENT}^^xsd:string .
source      SELECT ga_uuid.strval as uuid, "public"."Domestic EPC"."CO2_EMISSIONS_CURRENT",
↪  "public"."Domestic EPC"."HEATING_COST_CURRENT", "public"."Domestic
↪  EPC"."HOT_WATER_COST_CURRENT", "public"."Domestic EPC"."LIGHTING_COST_CURRENT"
            FROM "citydb"."cityobject_genericattrib" ga_uuid
            JOIN "citydb"."cityobject_genericattrib" ga_toid ON ga_uuid.cityobject_id =
            ↪  ga_toid.cityobject_id
            JOIN "public"."TOID_UPRN_MatchingTable" mt ON ga_toid.strval = mt."IDENTIFIER_2"
            JOIN "public"."Domestic EPC" ON mt."IDENTIFIER_1" = "public"."Domestic EPC"."UPRN"
            WHERE ga_uuid.attrname = 'uuid'
            AND ga_toid.attrname = 'os_topo_toid'

mappingId       Asset-mapping-StructureProperty
target      building:{uuid} a <https://www.theworldavatar.com/kg/Building> ;
↪  building:hasGlazedType {GLAZED_TYPE}^^xsd:string ; building:hasFloorDescription
↪  {FLOOR_DESCRIPTION}^^xsd:string ; building:hasWindowsDescription
↪  {WINDOWS_DESCRIPTION}^^xsd:string ; building:hasWallsDescription
↪  {WALLS_DESCRIPTION}^^xsd:string ; building:hasSecondHeatDescription
↪  {SECONDHEAT_DESCRIPTION}^^xsd:string ; building:hasRoofDescription
↪  {ROOF_DESCRIPTION}^^xsd:string .
source      SELECT ga_uuid.strval as uuid, "public"."Domestic EPC"."GLAZED_TYPE",
↪  "public"."Domestic EPC"."FLOOR_DESCRIPTION", "public"."Domestic EPC"."WINDOWS_DESCRIPTION",
↪  "public"."Domestic EPC"."WALLS_DESCRIPTION", "public"."Domestic EPC"."SECONDHEAT_DESCRIPTION",
↪  "public"."Domestic EPC"."ROOF_DESCRIPTION"
            FROM "citydb"."cityobject_genericattrib" ga_uuid
            JOIN "citydb"."cityobject_genericattrib" ga_toid ON ga_uuid.cityobject_id =
            ↪  ga_toid.cityobject_id
            JOIN "public"."TOID_UPRN_MatchingTable" mt ON ga_toid.strval = mt."IDENTIFIER_2"
            JOIN "public"."Domestic EPC" ON mt."IDENTIFIER_1" = "public"."Domestic EPC"."UPRN"
            WHERE ga_uuid.attrname = 'uuid'
            AND ga_toid.attrname = 'os_topo_toid'
]]
```

Listing 2: *SPARQL query to extract detailed building information, including energy performance and structural attributes.*

```
PREFIX rdf:  <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX building: <https://www.theworldavatar.com/kg/ontobuiltenv/building/>
PREFIX os: <http://data.ordnancesurvey.co.uk/ontology/spatialrelations/>

SELECT ?Property (GROUP_CONCAT(?tmp; separator=", ") AS ?Value) WHERE {
    SERVICE <http://174.138.27.240:3838/ontop/sparql> {
        {
            BIND("Building Instance" AS ?Property)
            BIND(<https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> AS ?tmp)
        } UNION {
            BIND("UPRN" AS ?Property)
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> os:hasUPRN ?tmp .
        } UNION {
            BIND("Energy Efficiency" AS ?Property )
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasEnergyEfficiency ?tmp .
        } UNION {
            BIND("EPC Rating" AS ?Property )
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasEPCrating ?tmp .
        } UNION {
            BIND("Postcode" AS ?Property)
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasPostcode ?tmp .
        } UNION {
            BIND("Address1" AS ?Property)
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasAddress1 ?tmp .
        } UNION {
            BIND("Property Type" AS ?Property)
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasPropertyType ?tmp .
        } UNION {
            BIND("Built Form" AS ?Property)
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasBuiltForm ?tmp .
        } UNION {
            BIND("CO2 Emissions Current" AS ?Property)
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasCO2EmissionsCurrent ?tmp
            ↪ .
        } UNION {
            BIND("Heating Cost Current" AS ?Property)
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasHeatingCostCurrent ?tmp .
        } UNION {
            BIND("Hot Water Cost Current" AS ?Property)
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasHotWaterCostCurrent ?tmp
            ↪ .
        } UNION {
            BIND("Lighting Cost Current" AS ?Property)
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasLightingCostCurrent ?tmp
            ↪ .
        } UNION {
            BIND("Glazed Type" AS ?Property)
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasGlazedType ?tmp .
        } UNION {
            BIND("Floor Description" AS ?Property)
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasFloorDescription ?tmp .
        } UNION {
            BIND("Windows Description" AS ?Property)
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasWindowsDescription ?tmp .
        } UNION {
            BIND("Walls Description" AS ?Property)
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasWallsDescription ?tmp .
        } UNION {
            BIND("Second Heat Description" AS ?Property)
            <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasSecondHeatDescription
            ↪ ?tmp .
        } UNION {
            BIND("Roof Description" AS ?Property)
            ↪ <https://www.theworldavatar.com/kg/Building/760a47c9-71de-45c4-a0e5-d0c8314eca30> building:hasRoofDescription ?tmp
            ↪ .
        }
    }
} GROUP BY ?Property
```

Listing 3: *Example SPARQL query to calculate the average EPC efficiency rating for buildings in a given LSOA and retrieve LSOA statistics such as fuel poverty and total electricity and gas consumption for the year 2022.*

```
[PrefixDeclaration]
twa:         https://www.theworldavatar.com/kg/
building:    https://www.theworldavatar.com/kg/ontobuiltenv/building
ontobuiltenergy:  https://www.theworldavatar.com/kg/ontobuiltenergy/
os:          http://data.ordnancesurvey.co.uk/ontology/spatialrelations/
rdf:         http://www.w3.org/1999/02/22-rdf-syntax-ns#

SELECT (AVG(?epcRating) AS ?avgEpcRating) ?fuelPoverty
       ?totalElectricity ?totalGas
WHERE {
  ?building a building:Building ;
            building:hasEpcRating ?epcRating ;
            building:belongsToPostcode ?postcode .
  ?postcode os:belongsToLsoa ?lsoa .
  ?lsoa os:hasCode "E01017943" ;
        ontobuiltenergy:hasFuelPoverty ?fuelPoverty ;
        ontobuiltenergy:hasElectricityConsumption ?totalElectricity ;
        ontobuiltenergy:hasGasConsumption ?totalGas ;
        rdf:recordedInYear "2022" .
}
GROUP BY ?fuelPoverty ?totalElectricity ?totalGas
```

Listing 4: *Equivalent SQL query to the SPARQL query in Listing 3.*

```
SELECT
    AVG(e.rating) AS avg_epc_rating,
    lsoa.fuel_poverty AS lsoa_fuel_poverty_rate,
    lsoa.total_electricity_consumption AS lsoa_electricity_consumption_kwh,
    lsoa.total_gas_consumption AS lsoa_gas_consumption_kwh,
    COUNT(DISTINCT b.id) AS total_buildings_in_lsoa,
    COUNT(e.uprn) AS buildings_with_epc,
    (SELECT COUNT(DISTINCT x.uprn)
     FROM EPCData x
     WHERE x.uprn IN (
         SELECT b.uprn
         FROM Buildings b
         JOIN GeographyMatching g ON b.postcode = g.postcode
         WHERE g.lsoa_code = 'E01017943'
     )) AS total_high_epc_buildings
FROM Buildings b
JOIN GeographyMatching g ON b.postcode = g.postcode
JOIN EPCData e ON b.uprn = e.uprn
JOIN LSOAData lsoa ON g.lsoa_code = lsoa.lsoa_code
LEFT JOIN ElectricityConsumption ec ON lsoa.lsoa_code = ec.lsoa_code
LEFT JOIN GasConsumption gc ON lsoa.lsoa_code = gc.lsoa_code
WHERE lsoa.lsoa_code = 'E01017943'
  AND lsoa.year = '2022'
GROUP BY lsoa.fuel_poverty,
         lsoa.total_electricity_consumption,
         lsoa.total_gas_consumption
ORDER BY avg_epc_rating DESC;
```

The OBDA mapping in Listing 1 instantiates buildings with attributes such as Topographic Identifier (TOID), geometry, location, building function, footprint, and elevation, as required by the CityGML ontology. The instantiated buildings are linked to Unique Property Reference Numbers (UPRN). The `hasUPRN` relation is used to link the buildings to energy characteristics from the EPC source data. It is worth noting that for standalone houses, each building has a single UPRN and a distinct EPC assessment result. For flats or apartments, each individual dwelling has its own EPC, meaning a building may have multiple `hasEnergyEfficiency` relations, each corresponding to a separate dwelling. This is referred to as a parent building. In both cases, we use the term "building" to represent a more general concept, rather than explicitly distinguishing between parent buildings and individual dwellings. This approach streamlines the representation and querying process for large datasets. Should there be a specific need to determine whether a building contains only one EPC, this can be addressed through customised queries and counting the returned results.

In the SQL query in Listing 4, the `Buildings` table contains building-specific data, identified by UPRNs and postcodes. The `GeographyMatching` table links each building's postcode to its corresponding LSOA, establishing the geographical relationships necessary for the query. The `EPCData` table stores the EPC ratings for buildings, which are matched to the building records via the UPRN. Additionally, the `LSOAData` table holds LSOA-level statistics such as fuel poverty and total energy consumption, indexed by the LSOA code. This structure allows the query to aggregate and analyze data across multiple tables, providing both building-specific details and broader LSOA-level insights.

In comparison with Listing 4, the OBDA solution exemplified in Listing 3 simplifies queries by enabling machines to use semantic concepts and relations defined in an ontology, rather than relying on the cumbersome and manual process of interacting directly with database schema details.

A.4 Workflow and UML Diagram

Figure A.5 illustrates the data processing workflow used in the study. The workflow commences with the collation of data sources, including vector, raster, and tabular datasets. The data are loaded into a relational database are accessed via the Ontop Ontology-Based Data Access (OBDA) framework [11]. The RDF triples, generated from tabular, time-series, and relational datasets, represent geospatial and energy-related attributes of the built environment. These triples are then integrated into a virtual knowledge graph, which supports semantic querying and the unification of data across different granularities. The virtual knowledge graph ensures interoperability across datasets, such as energy performance, public health metrics, and socioeconomic indicators, enabling coherent access and analysis at the required levels of detail.

At the bottom of Figure A.5, a visualisation presents an interface powered by the virtual knowledge graph, integrating building location, geometry, and energy performance ratings. Buildings are rendered in 3D, with colours representing the matched EPC ratings – green for the highest efficiency and red for the lowest – providing a clear display of energy performance across the Cambridge area. The virtual knowledge graph facilitates seamless integration and querying, providing an intuitive display of energy performance disparities across the region, while remaining interoperable with other socioeconomic and public health statistics.

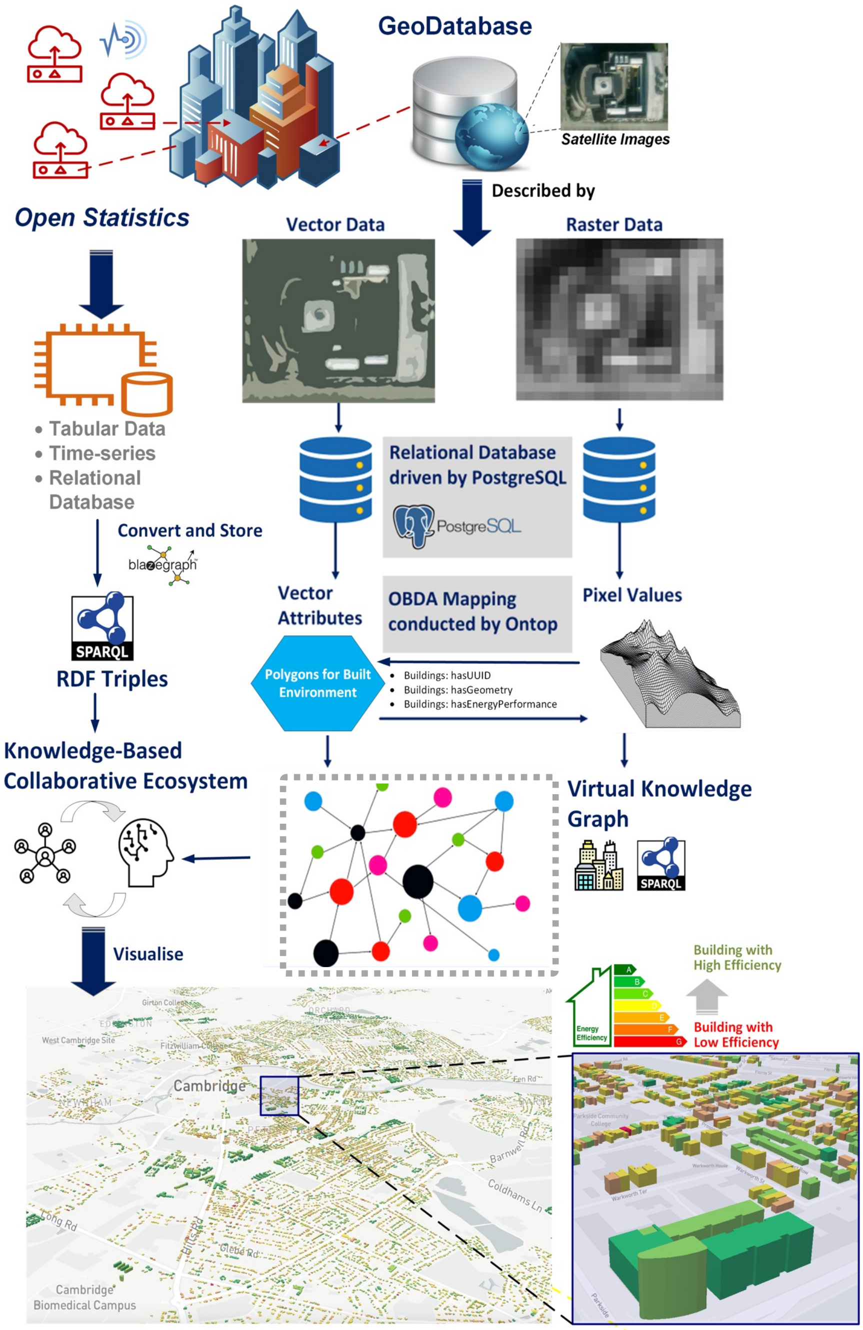


Figure A.5: *Workflow provided via The World Avatar: knowledge graph generation from vector, raster, and tabular data via ontology mapping; containerized processing for querying, computation, and multi-scale visualisation*

The workflow was implemented as part of The World Avatar via a containerised stack and a set of computational agents. Figure A.6 presents a sequence diagram showing the interactions between the different components of the system. The process begins with an actor initiating the stack via a Stack Manager. This triggers the creation of containers for essential services such as PostGIS, CityTiler, and Ontop. Data ingestion is handled by the TWA Stack Data Uploader. This loads satellite images and building investigation data, creates and populates the databases within the stack in OntoCityGML format, generates IRIs (internationalized resource identifiers) for each building, and produces visualisation tiles. An Energy Performance Agent and a Feature Information Agent interact with the knowledge graph, managing EPC data and building metadata respectively. The TWA Visualisation Framework allows users to interact with the integrated information. This orchestrated system of agents and containers enables the seamless transformation of diverse data sources into a cohesive, queryable knowledge graph, supporting multi-scale visualisation and analysis of urban environments.

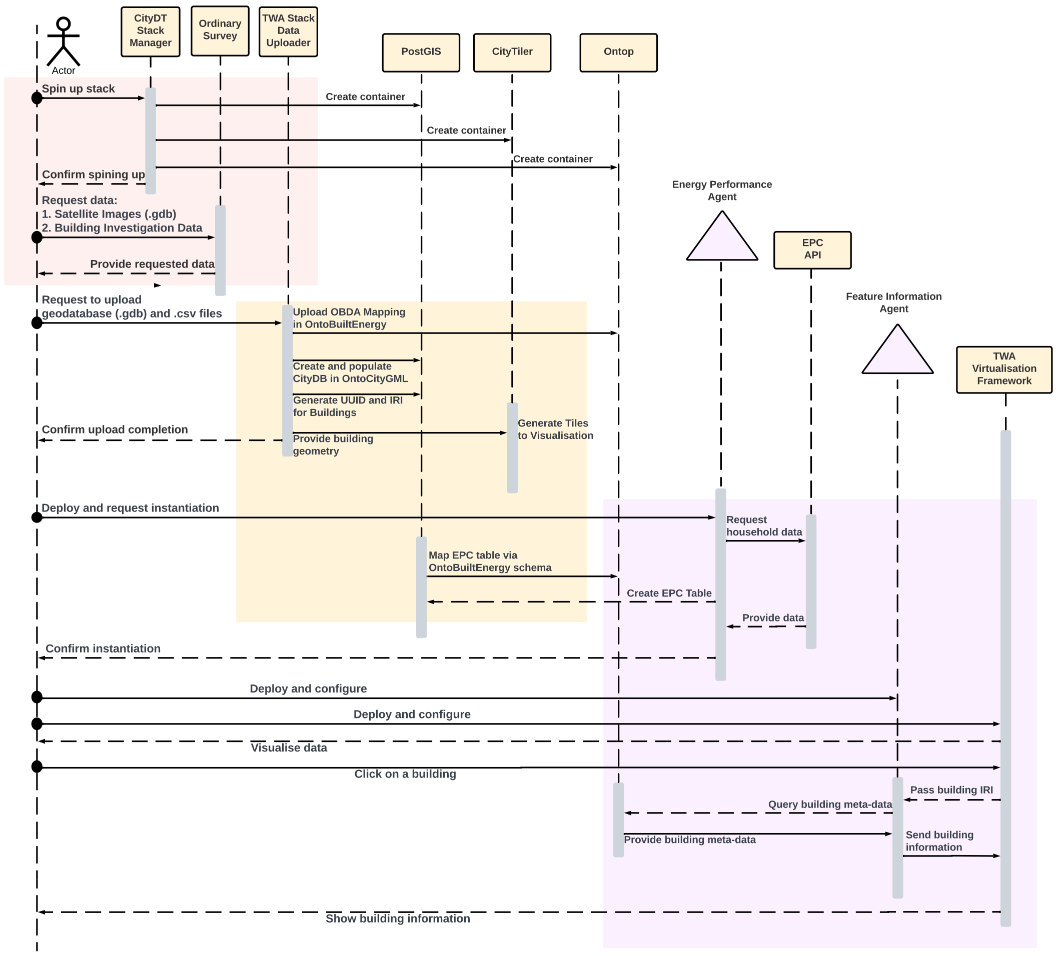


Figure A.6: *UML diagram for workflow deployment.*

A.5 Analysis of source data

A.5.1 Relative closeness to ideal and negative-ideal scenarios

In this study, the criteria associated with each geographic unit collectively define a solution space, within which each area represents a point. These criteria are E_{eff} (Energy Performance Efficiency), E_{cons} (Energy Consumption), R_{FP} (the Rate of Fuel Poverty), R_{CVD} (the Rate of Cardiovascular Disease), and R_{COPD} (the Rate of Chronic Obstructive Pulmonary Disease). Let x_{ij} represent the raw value of area i with respect to criterion $j \in J$, where J is the set of all criteria

$$J = \{E_{\text{eff}},\, E_{\text{cons}},\, R_{\text{FP}},\, R_{\text{CVD}},\, R_{\text{COPD}}\}. \tag{A.1}$$

To facilitate comparison among the areas, the raw values of each criterion x_{ij} are normalised to obtain a normalised decision matrix $R = [r_{ij}]$. The normalised value r_{ij} for each criterion j of area i is computed as follows

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{n} x_{ij}^2}}, \tag{A.2}$$

where n is the number of areas under consideration.

Weights w_j are assigned to each criterion j to reflect their relative importance. In this study, the entropy weight method [15] is selected to determine the objective weights based on the degree of dispersion of each criterion. The entropy weight method assigns higher weights to criteria with greater variability in their data distribution, indicating that they provide more information for distinguishing among the areas. Conversely, criteria with uniform values across all areas have lower information entropy and thus receive weights approaching zero, as they contribute less to the decision-making process. The proportion p_{ij} for each criterion j is expressed as follows

$$p_{ij} = \frac{r_{ij}}{\sum_{i=1}^{n} r_{ij}}. \tag{A.3}$$

The information entropy H_j for each criterion j is calculated as:

$$H_j = -\frac{1}{\ln n} \sum_{i=1}^{n} p_{ij} \ln p_{ij}. \tag{A.4}$$

To avoid undefined expressions when $p_{ij} = 0$, we define $p_{ij} \ln p_{ij} = 0$, which is justified mathematically by the concept of limits. The weight w_j for each criterion j is then given by

$$w_j = \frac{1 - H_j}{\sum_{j=1}^{m} 1 - H_j}, \tag{A.5}$$

where m is the number of criteria.

A weighted normalised decision matrix $V = [v_{ij}]$ is computed by multiplying each normalised value r_{ij} by its corresponding weight w_j

$$v_{ij} = w_j \cdot r_{ij}. \tag{A.6}$$

The Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS) method [47] is applied to evaluate the relative closeness of each area to ideal and negative-ideal scenarios. In this context, it serves to quantify the urgency for implementing energy retrofit measures for each analysed entity. The TOPSIS determines the ideal solution (v_j^+) and a negative-ideal solution (v_j^-) based on two types of criteria: benefit criteria (J_{benefit}), where higher values are preferred (such as efficiency), and cost criteria (J_{cost}), where lower values are preferred (such as consumption or risk). Let i index the areas under consideration. We partition the criteria set J into benefit criteria J_{benefit} and cost criteria J_{cost}

$$\begin{cases} J_{\text{benefit}} &= \{E_{\text{eff}}\}, \\ J_{\text{cost}} &= \{E_{\text{cons}},\ R_{\text{FP}},\ R_{\text{CVD}},\ R_{\text{COPD}}\}. \end{cases} \tag{A.7}$$

The ideal and negative-ideal solutions for the areas under study are defined as follows.

$$\begin{aligned} v_j^+ &= \left(\max_i v_{ij} \mid j \in J_{\text{benefit}}\right) \cup \left(\min_i v_{ij} \mid j \in J_{\text{cost}}\right) \\ &= \left(\max_i E_{\text{eff}},\ \min_i E_{\text{cons}},\ \min_i R_{\text{FP}},\ \min_i R_{\text{CVD}},\ \min_i R_{\text{COPD}}\right), \\ v_j^- &= \left(\min_i v_{ij} \mid j \in J_{\text{benefit}}\right) \cup \left(\max_i v_{ij} \mid j \in J_{\text{cost}}\right) \\ &= \left(\min_i E_{\text{eff}},\ \max_i E_{\text{cons}},\ \max_i R_{\text{FP}},\ \max_i R_{\text{CVD}},\ \max_i R_{\text{COPD}}\right), \end{aligned} \tag{A.8}$$

The Euclidean distances from each area to both the ideal and negative-ideal solutions are calculated by

$$\begin{aligned} S_i^+ &= \sqrt{\sum_{j=1}^{m} (v_{ij} - v_j^+)^2}, \\ S_i^- &= \sqrt{\sum_{j=1}^{m} (v_{ij} - v_j^-)^2}. \end{aligned} \tag{A.9}$$

The relative closeness C_i of each area i to the negative-ideal solution is then computed

$$C_i = \frac{S_i^+}{S_i^+ + S_i^-}, \tag{A.10}$$

where higher values of C_i indicate areas that are closer to the negative-ideal solution and hence, are higher in priority for retrofitting.

A.5.2 Energy consumption hot spots

The analysis in the main text used a quantitative description of the spatial clustering characteristics of energy consumption to identify high energy consumption clusters. Large, continuous clusters of high energy consumption are considered priority areas for policy implementation because their spatial proximity facilitates more efficient resource allocation and coordinated policy enforcement.

Local Moran's I Local Moran's I [5] is used to identify spatial clusters and outliers. The formula for calculating Local Moran's I for location i is

$$I_i = \left(\frac{O_i - \bar{O}}{s}\right) \sum_{j=1}^{N} \lambda_{ij} \left(\frac{O_j - \bar{O}}{s}\right), \tag{A.11}$$

where the standard deviation of the observations across all locations

$$s = \sqrt{\frac{1}{N-1} \sum_{k=1}^{N} (O_k - \bar{O})^2}, \tag{A.12}$$

and λ_{ij} represents the spatial weight between locations i and j. O_i is the value of the observation at location i, $\bar{O}$ is the mean observation value across all locations, and N is the total number of locations. In this study, the observations were based on data reported for the energy consumption E_{cons} in each LSOA. There are over 30,000 LSOAs across the UK. Each contains approximately 1000 households, and has only a few direct neighbors. The mean $\bar{O}$ was calculated as the arithmetic (non-area weighted) mean over the set of observations, and it was chosen to define the spatial weights using a Gaussian smoothing approach

$$\lambda_{ij} = \exp\left(-\frac{d_{ij}^2}{2h^2}\right), \tag{A.13}$$

where d_{ij} represents the Euclidean distance between LSOA i and j, and h is the mean of the pairwise distances between all LSOAs. The Gaussian smoothing ensures that all spatial points exert some level of influence on each other, with closer points having a stronger influence, and provides a more continuous and flexible measure of spatial relationships, which is particularly useful when dealing with large datasets.

A positive of Local Moran's I, $I_i > 0$, indicates positive spatial autocorrelation, meaning that locations with similar values are clustered together. Conversely, $I_i < 0$ suggests negative spatial autocorrelation, indicating that dissimilar values are more likely to be found in proximity to one another. However, significance testing is required to validate the statistical reliability of the autocorrelations to ensure that the observed spatial patterns are not due to random chance.

Significance testing We perform significance testing to determine the significance of the observed Local Moran's I values by comparing the observed values to a distribution of Local Moran's I values derived from random permutations of the data. First, we calculate the observed Local Moran's I value I_i for each location i. Then, we generate M random permutations of the observation data and compute the Local Moran's I values for each permutation $I_i^{(k)}$ for $k = 1, 2, \ldots, M$. The p-value for each location i is

$$p_i = \frac{1}{M} \sum_{k=1}^{M} \mathbb{I}(I_i^{(k)} \geq I_i), \tag{A.14}$$

where $\mathbb{I}$ is an indicator function that equals 1 if $I_i^{(k)} \geq I_i$ and 0 otherwise. The observed Local Moran's I value was considered to be statistically significant if the p-value is less than a chosen significance level, in this case, $p_i < 0.05$.

Classification of high energy consumption hot spots The following Boolean expressions were used to classify the energy consumption hot spots

$$\begin{cases} \text{Hot spot} : (I_i > 0) \wedge (p_i < \alpha) \wedge (E_i > \bar{E}) \\ \text{Cold spot} : (I_i > 0) \wedge (p_i < \alpha) \wedge (E_i < \bar{E}) \\ \text{Spatial outliers} : (I_i < 0) \wedge (p_i < \alpha) \wedge (E_i < \bar{E}) \\ \text{Spatial outliers} : (I_i < 0) \wedge (p_i < \alpha) \wedge (E_i > \bar{E}) \end{cases} \tag{A.15}$$

where I_i was calculated in terms of the energy consumption at location E_i, $\bar{E}$ is the mean energy consumption and $\alpha = 0.05$ is the significance level. Hot spots are defined as locations with significant positive spatial autocorrelation and above-average energy consumption, and cold spots as locations with significant positive spatial autocorrelation and below-average energy consumption. The remaining combinations represent spatial outliers, where high-energy consumption locations are interspersed with low-energy locations, or vice versa.

Figure A.7 shows the distribution of hot spots in England. Beyond London, energy hot spots are predominantly concentrated in central and western England, reflecting spatially clustered regions with high energy demands.

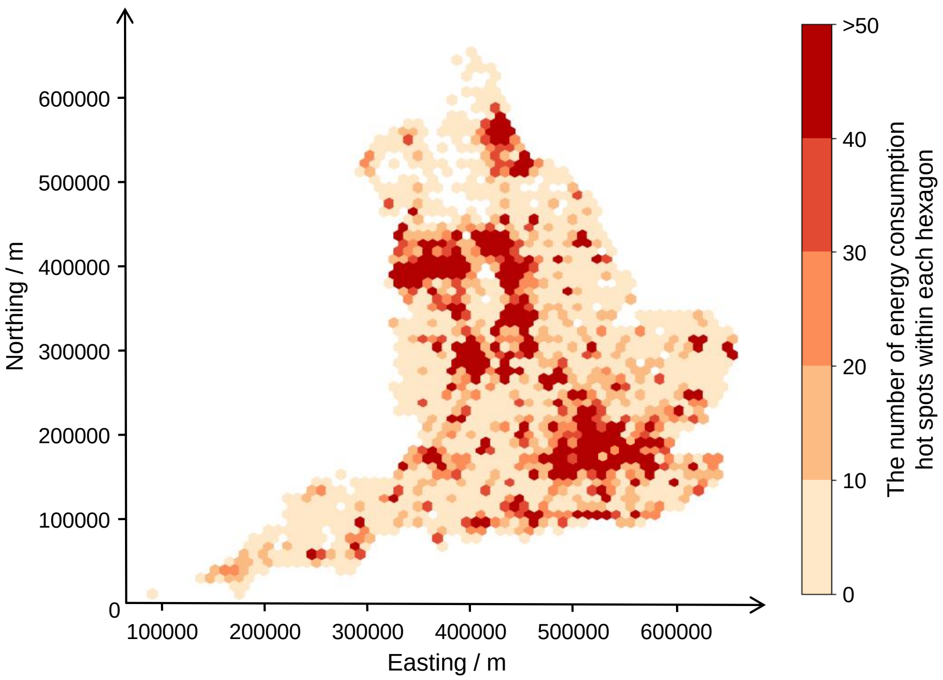


Figure A.7: *Energy consumption hot spot in England: Each hexagon shows the number of identified hot spots. The map is projected using the British National Grid coordinate system (EPSG: 27700), where the x-axis and y-axis represent Eastings and Northings, respectively, measured in metres.*

A.5.3 Building retrofit priority index

After calculating the relative closeness to the ideal and negative-ideal solution and identifying energy consumption hot spots, we propose a Building Retrofit Priority Index (BRPI) to consolidate these findings. An adjustment to the TOPSIS score is made via the Building Retrofit Priority Index (BRPI) to incorporate the significance of spatial autocorrelation in energy consumption. The BRPI modifies the relative closeness based on whether an area is identified as an energy consumption hot spot

$$\mathrm{BRPI}_i = C_i + \beta \cdot \chi_i \cdot (1 - C_i), \tag{A.16}$$

where

$$\chi_i = \begin{cases} 1 & \text{if area } i \text{ is an energy consumption hot spot,} \\ 0 & \text{otherwise,} \end{cases} \tag{A.17}$$

$\beta = 0.1$ is an adjustment coefficient that controls the influence of the hot spot status, and χ_i indicates whether area i is classified as a hot spot as per Equation A.15.

A larger value of BRPI_i indicates a higher priority for intervention. The BRPI maintains the integrity of the original TOPSIS scores while ensuring that regions identified as high energy consumption hot spots receive an appropriately higher priority. This balances the objective assessment of retrofit necessity based on multiple criteria with the strategic prioritization of regions where interventions can yield the most significant impact.

A.5.4 Case-based reasoning and incremental analysis

The main text combines case-based reasoning (CBR) with an incremental analysis approach to evaluate the marginal benefits of fabric-first and system-led retrofit strategies to improve the energy efficiency of buildings.

Indicators for fabric-first and system-led strategy We define the following indicators to evaluate the impact of each retrofit strategy.

The **fabric-first strategy** focuses on improving the thermal efficiency of the building's envelope. The indicators used in the analysis are

$$\mathbb{F}_i = \{W_{\mathrm{a}i}, W_{\mathrm{g}i}, W_{\mathrm{r}i}, W_{\mathrm{f}i}\}, \tag{A.18}$$

where $W_{\mathrm{a}i}$, $W_{\mathrm{g}i}$, $W_{\mathrm{r}i}$, and $W_{\mathrm{f}i}$ represent the thermal efficiency ratings of the wall, window, roof, and floor, respectively, for retrofit candidate i.

The **system-led strategy** targets the efficiency of the building's systems, including heating, hot water, and lighting. The indicators used in the analysis are

$$\mathbb{S}_i = \{S_{\mathrm{h}i}, S_{\mathrm{c}i}, S_{\mathrm{w}i}, S_{\mathrm{l}i}\}, \tag{A.19}$$

where $S_{\mathrm{h}i}$, $S_{\mathrm{c}i}$, $S_{\mathrm{w}i}$, and $S_{\mathrm{l}i}$ represent the efficiency ratings of the main heating system, main heating control system, hot water system, and lighting, respectively, for retrofit candidate i.

In the raw EPC dataset, the efficiency of fabric and system-related attributes are provided as categorical ratings assigned by energy assessors during on-site inspections. For quantitative analysis, we convert the categories into numerical values as follows

$$(\textit{Very Poor, Poor, Average, Good, Very Good}) \mapsto (0, 0.25, 0.5, 0.75, 1). \tag{A.20}$$

Case-based reasoning for retrofit strategy selection The following energy-related attributes are considered by the case-based reasoning used to select a retrofit strategy

$$\mathbb{R}_i = \{R_{\mathrm{b}i}, R_{\mathrm{p}i}, R_{\mathrm{h}i}, R_{\mathrm{a}i}\}, \tag{A.21}$$

where i indexes the building and $R_{\mathrm{b}i}, R_{\mathrm{p}i}, R_{\mathrm{h}i}, R_{\mathrm{a}i}$ represent the built form, property type, number of heated rooms, and total floor area of building i respectively.

We define matching functions that indicate whether the properties for two buildings match, including a 20% tolerance when comparing areas

$$\delta(R_{\alpha i}, R_{\alpha j}) = \begin{cases} 1 & \text{if } R_{\alpha i} = R_{\alpha j},\ \alpha \in \{\text{b, p, h}\} \\ 0 & \text{otherwise,} \end{cases} \tag{A.22}$$

$$\delta_{\text{area}}(R_{\mathrm{a}i}, R_{\mathrm{a}j}) = \begin{cases} 1 & \text{if } 0.8\ R_{\mathrm{a}i} \leq R_{\mathrm{a}j} \leq 1.2\ R_{\mathrm{a}i} \\ 0 & \text{otherwise.} \end{cases} \tag{A.23}$$

Building i and j are considered to match when

$$\operatorname{diag}\left(\delta(R_{\mathrm{b}i}, R_{\mathrm{b}j}), \delta(R_{\mathrm{p}i}, R_{\mathrm{p}j}), \delta(R_{\mathrm{h}i}, R_{\mathrm{h}j}), \delta_{\text{area}}(R_{\mathrm{a}i}, R_{\mathrm{a}j})\right) = \mathbf{I}$$

where $\mathbf{I}$ is the identity matrix. For each building i requiring retrofit (*i.e.*, with an EPC rating below 69), we identify matching buildings with EPC rating above 69. We compare the differences in the fabric and system attributes between the retrofit candidate and first $k = 10$ exemplar cases.

For each fabric-related attribute $f \in \mathbb{F}$ and system-related attribute $s \in \mathbb{S}$, the following metrics are calculated

$$\overline{\Delta\mathbb{F}_i} = \frac{1}{k} \sum_{j=1}^{k} \sum_{f \in \mathbb{F}} |f_j - f_i|, \tag{A.24}$$

$$\overline{\Delta\mathbb{S}_i} = \frac{1}{k} \sum_{j=1}^{k} \sum_{s \in \mathbb{S}} |s_j - s_i|, \tag{A.25}$$

where j indexes the exemplars and i indexes the each retrofit candidate.

Case-based decision The decision criteria for choosing a retrofit strategy are as follows:

Fabric-first is chosen when the average difference in system-related attributes is below a threshold ε, while the average difference in fabric-related attributes exceeds ε

$$\left(\overline{\Delta\mathbb{S}_i} < \varepsilon\right) \wedge \left(\overline{\Delta\mathbb{F}_i} > \varepsilon\right) \tag{A.26}$$

System-led is chosen when the average difference in fabric-related attributes is below the threshold ε, while the average difference in system-related attributes exceeds ε

$$\left(\overline{\Delta\mathbb{F}_i} < \varepsilon\right) \wedge \left(\overline{\Delta\mathbb{S}_i} > \varepsilon\right) \tag{A.27}$$

Incremental analysis is applied when both the average differences in fabric and system attributes exceed the threshold ε, or when no sufficiently similar exemplar cases are found. This occurs when

$$\left(\overline{\Delta\mathbb{F}_i} > \varepsilon\right) \wedge \left(\overline{\Delta\mathbb{S}_i} > \varepsilon\right) \tag{A.28}$$

In this analysis, $\varepsilon = 0.25$, representing a difference of one EPC efficiency rating level.

Incremental analysis In the cases where the CBR did not yield a clear recommendation, an incremental analysis was performed using an XGBoost model [16]. XGBoost builds an ensemble of decision trees sequentially, where each new tree is added to correct the errors of the existing ensemble. The XGBoost used in this study is trained to predict the energy efficiency score for each retrofit candidate i based on a set of features X_i. These features include all the indicators within the sets $\mathbb{F}_i$ (fabric-related indicators), $\mathbb{S}_i$ (system-related indicators), and the descriptive metadata ($\mathbb{R}_i$)

$$X_i = \mathbb{F}_i \cup \mathbb{S}_i \cup \mathbb{R}_i. \tag{A.29}$$

The training process is as follows. $\hat{y}_i$ represents the model output (*i.e.*, the estimated energy efficiency score). The initial prediction for building i is taken as the mean energy efficiency score

$$\hat{y}_i^{(0)} = \overline{E_{\text{eff}}}. \tag{A.30}$$

At each iteration t, a new decision tree $T_t(x)$ is trained to fit the residuals. The residuals (negative gradients) are calculated based on the current model predictions

$$g_i^{(t)} = \frac{\partial L(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}}, \tag{A.31}$$

where $L(y_i, \hat{y}_i)$ is the loss function employed in the XGBoost model. In this study, the loss function is defined as an exponential penalty on the prediction error

$$L(y_i, \hat{y}_i) = \exp\left((y_i - \hat{y}_i)^2\right) - 1. \tag{A.32}$$

The exponential loss increases the penalty for larger errors, ensuring that the model places a higher emphasis on minimizing significant deviations between the actual value y_i and the predicted value $\hat{y}_i$. The decision tree $T_t(x)$ is trained

$$T_t(x) = \arg\min_T \sum_{i=1}^{N} L\left(y_i, \hat{y}_i^{(t-1)} + T(x_i)\right). \tag{A.33}$$

The putout of the model is updated by adding the contribution of the new tree, scaled by a learning rate η

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta T_t(x_i). \tag{A.34}$$

In our implementation, the learning rate is initially set to 0.1 and further refined using Bayesian optimisation [56]. The search space for the learning rate is defined between 0.01 and 0.3. Through this iterative process, the tree model optimizes the prediction accuracy, refining the model with each new tree added to the ensemble.

Marginal benefit assessment For each building, the marginal benefit of upgrading two categories of features – fabric (e.g., walls, roof, windows) and systems (e.g., main heating, hot water) – is assessed. To achieve this, improvements are simulated by upgrading either fabric-related or system-related features to the next level prescribed by the EPC categories (*e.g.* from *average* to *good*) while keeping all other features unchanged.

Let X_i denote the feature vector of building i defined as per equation (A.29). When all the fabric-related features $\mathbb{F}_i$ are upgraded to $\mathbb{F}_i'$, the resulting feature vector is given

$$X_i^{(\mathbb{F})} = (X_i \setminus \mathbb{F}_i) \cup \mathbb{F}_i'. \tag{A.35}$$

Likewise, when all the system-related features $\mathbb{S}_i$ are upgraded to $\mathbb{S}_i'$

$$X_i^{(\mathbb{S})} = (X_i \setminus \mathbb{S}_i) \cup \mathbb{S}_i'. \tag{A.36}$$

The marginal benefit of upgrading all features in a category is calculated

$$\mu_{\mathbb{F}_i} = \hat{y}_i(X_i^{(\mathbb{F})}) - E_{\mathrm{eff},i} \tag{A.37}$$

$$\mu_{\mathbb{S}_i} = \hat{y}_i(X_i^{(\mathbb{S})}) - E_{\mathrm{eff},i} \tag{A.38}$$

where $\hat{y}_i(X_i^{(\mathbb{F})})$ and $\hat{y}_i(X_i^{(\mathbb{S})})$ are the predicted energy efficiency scores for building i after incrementally upgrading the fabric-related and system-related features, and $E_{\mathrm{eff},i}$ is the current energy efficiency score from the EPC of building i. For each building i, the marginal benefits of the fabric-first and system-led strategies are compared

$$\text{Retrofit Decision} = \begin{cases} \text{Fabric-first} & \text{if } \mu_{\mathbb{F}_i} > \mu_{\mathbb{S}_i}, \\ \text{System-led} & \text{if } \mu_{\mathbb{S}_i} > \mu_{\mathbb{F}_i}. \end{cases} \tag{A.39}$$

In the event that the marginal benefits are equal ($\mu_{\mathbb{F}_i} = \mu_{\mathbb{S}_i}$), a secondary criterion is used

$$\mu_{ij} = \hat{y}_i(X_i^{(j)}) - E_{\mathrm{eff},i}, \tag{A.40}$$

where $X_i^{(j)}$ represents the feature vector of building i with only a feature j upgraded to the next level. We compare the largest individual marginal improvements

$$\text{Retrofit Decision} = \begin{cases} \text{Fabric-first} & \text{if } \max\limits_{j\in\mathbb{F}} \mu_{ij} > \max\limits_{j\in\mathbb{S}} \mu_{ij}, \\ \text{System-led} & \text{if } \max\limits_{j\in\mathbb{S}} \mu_{ij} \geq \max\limits_{j\in\mathbb{F}} \mu_{ij}. \end{cases} \tag{A.41}$$

In the context of the combined CBR and XGBoost model, this analysis allows us to identify which strategy yields the greatest marginal improvements in energy efficiency, facilitating targeted and effective retrofit interventions. The model considers both the current state of each building and projections based on similar cases, alongside the benefits of incremental improvements towards reaching the energy efficiency goal of 69+.

B Tabular summary data

Table B.1: *Local authority-level breakdown of the number of dwellings requiring retrofits (EPC < 69), BRPI values (for England only, indicated as "-" for Wales authorities due to unavailable fuel poverty components), and the estimated proportion of dwellings benefiting from fabric-first vs system-led measures across England and Wales.*

LA code	LA name	Number of houses with EPC (-)	Number of houses with EPC below 69 (-)	Average EPC rating (-)	BRPI (-)	Fabric-first (%)	System-led (%)
E07000223	Adur	23888	15341	63	0.524	59.3	40.7
E07000032	Amber-Valley	48543	31427	62	0.578	59.2	40.8
E07000224	Arun	72336	42531	64	0.522	54.1	45.9
E07000170	Ashfield	51483	30483	64	0.559	43	57
E07000105	Ashford	52046	25687	67	0.527	45.5	54.5
E07000200	Babergh	36500	21857	63	0.632	57.9	42.1
E09000002	Barking-and-Dagenham	80989	46249	67	0.490	58	42
E09000003	Barnet	158618	87069	66	0.565	54.2	45.8
E08000016	Barnsley	100796	60467	65	0.554	46.5	53.5
E07000066	Basildon	64686	35747	66	0.441	44.5	55.5
E07000084	Basingstoke-and-Deane	76605	32862	69	0.482	41.5	58.5
E07000171	Bassetlaw	53863	33227	63	0.577	57.1	42.9
E06000022	Bath-and-North-East-Somerset	76747	47088	63	0.561	47.4	52.6
E06000055	Bedford	76056	37026	67	0.521	48.2	51.8
E09000004	Bexley	93193	57653	64	0.731	53.8	46.2
E08000025	Birmingham	453395	302147	62	0.748	51.2	48.8
E07000129	Blaby	33898	20546	65	0.549	41.1	58.9
E06000008	Blackburn-with-Darwen	63745	41594	62	0.591	57.6	42.4
E06000009	Blackpool	80602	61242	58	0.611	58.4	41.6
W06000019	Blaenau-Gwent	26276	18234	62	—	61.6	38.4
E07000033	Bolsover	32574	18929	64	0.567	57.2	42.8
E08000001	Bolton	114724	69348	64	0.560	48.4	51.6
E07000136	Boston	31613	19091	62	0.577	60.4	39.6
E06000058	Bournemouth-Christchurch-and-Poole	179654	102307	65	0.423	59.6	40.4
E06000036	Bracknell-Forest	48815	23627	68	0.458	43.4	56.6
E08000032	Bradford	238814	170245	60	0.626	60.5	39.5
E07000067	Braintree	58339	32126	65	0.537	53.3	46.7
E07000143	Breckland	58618	34715	64	0.542	58.3	41.7
E09000005	Brent	126490	65156	67	0.428	58.3	41.7
E07000068	Brentwood	29448	17628	64	0.632	49.6	50.4
W06000013	Bridgend	54047	31600	64	—	53.9	46.1
E06000043	Brighton-and-Hove	141715	87362	63	0.547	61	39
E06000023	Bristol-City-of	202100	117555	64	0.536	49.8	50.2
E07000144	Broadland	48926	29978	64	0.538	57.1	42.9
E09000006	Bromley	124623	78331	63	0.671	55.3	44.7
E07000234	Bromsgrove	35349	21265	64	0.486	52.1	47.9
E07000095	Broxbourne	34955	19005	65	0.520	51.1	48.9
E07000172	Broxtowe	40368	27963	61	0.569	58.5	41.5
E06000060	Buckinghamshire	48844	21146	68	0.533	56.6	43.4
E07000117	Burnley	45349	33948	59	0.606	61.2	38.8
E08000002	Bury	75251	49988	63	0.560	49.4	50.6
W06000018	Caerphilly	65640	40090	64	—	57.6	42.4
E08000033	Calderdale	92902	65056	60	0.602	60.1	39.9
E07000008	Cambridge	57177	27066	68	0.450	55.4	44.6
E09000007	Camden	115448	58414	67	0.408	59.6	40.4
E07000192	Cannock-Chase	39445	23492	65	0.564	53.7	46.3
E07000106	Canterbury	66533	38122	65	0.540	52.8	47.2
W06000015	Cardiff	160849	88900	66	—	51.1	48.9
W06000010	Carmarthenshire	73021	51996	57	—	63.8	36.2
E07000069	Castle-Point	28463	21607	60	0.638	50.5	49.5
E06000056	Central-Bedfordshire	116279	56095	68	0.431	42.4	57.6
W06000008	Ceredigion	33737	25378	53	—	70.2	29.8
E07000130	Charnwood	68096	38062	65	0.554	51	49
E07000070	Chelmsford	67496	36652	66	0.523	45.5	54.5
E07000078	Cheltenham	54337	31455	64	0.524	50.7	49.3
E07000177	Cherwell	63557	31516	67	0.520	39.9	60.1
E06000049	Cheshire-East	165170	97703	64	0.552	51.9	48.1
E06000050	Cheshire-West-and-Chester	138786	78855	65	0.535	54	46

Table B.1: *(Continued)*

LA code	LA name	Number of houses with EPC (-)	Number of houses with EPC below 69 (-)	Average EPC rating (-)	BRPI (-)	Fabric-first (%)	System-led (%)
E07000034	Chesterfield	44571	26887	65	0.549	47.3	52.7
E07000225	Chichester	54260	31418	63	0.621	52	48
E07000118	Chorley	44671	24166	66	0.529	47.4	52.6
E09000001	City-of-London	7807	2880	69	0.203	41.7	58.3
E07000071	Colchester	83588	41169	67	0.516	47.3	52.7
W06000003	Conwy	54606	38266	59	—	60.1	39.9
E06000052	Cornwall	267052	173902	58	0.747	61.8	38.2
E07000079	Cotswold	41668	25899	61	0.740	57.5	42.5
E06000047	County-Durham	254406	150076	65	0.529	43.4	56.6
E08000026	Coventry	144194	91718	64	0.603	50.6	49.4
E07000226	Crawley	38354	18146	68	0.412	47.8	52.2
E09000008	Croydon	154735	91970	64	0.550	55.3	44.7
E06000063	Cumberland	1898	1111	65	0.558	56.4	43.6
E07000096	Dacorum	60416	33477	65	0.531	45.9	54.1
E06000005	Darlington	50203	30878	63	0.543	41.3	58.7
E07000107	Dartford	42742	19728	69	0.507	50.5	49.5
W06000004	Denbighshire	42909	31025	58	—	61.9	38.1
E06000015	Derby	110054	68928	63	0.578	52.8	47.2
E07000035	Derbyshire-Dales	30189	20344	60	0.630	60.6	39.4
E08000017	Doncaster	131147	85174	63	0.571	58.2	41.8
E06000059	Dorset	155438	92052	63	0.547	52	48
E07000108	Dover	50328	29948	64	0.547	50.4	49.6
E08000027	Dudley	108057	71855	62	0.583	54.4	45.6
E09000009	Ealing	143061	79033	66	0.525	57.2	42.8
E07000061	Eastbourne	48949	27576	65	0.514	59.3	40.7
E07000009	East-Cambridgeshire	36708	20393	64	0.535	53.3	46.7
E07000040	East-Devon	66164	39712	63	0.546	49.5	50.5
E07000085	East-Hampshire	48461	25866	66	0.543	43.7	56.3
E07000242	East-Hertfordshire	56165	29783	66	0.537	43.8	56.2
E07000086	Eastleigh	52735	25377	68	0.415	47.6	52.4
E07000137	East-Lindsey	68769	47981	59	0.615	62.4	37.6
E06000011	East-Riding-of-Yorkshire	140941	90023	62	0.577	57.8	42.2
E07000193	East-Staffordshire	50512	29563	64	0.606	53.8	46.2
E07000244	East-Suffolk	105471	66028	62	0.609	57.8	42.2
E07000207	Elmbridge	56672	33013	64	0.648	51.6	48.4
E09000010	Enfield	125464	81027	63	0.560	54.9	45.1
E07000072	Epping-Forest	49704	29538	64	0.661	48.8	51.2
E07000208	Epsom-and-Ewell	28481	16904	64	0.613	56.7	43.3
E07000036	Erewash	43459	28426	62	0.567	58.3	41.7
E07000041	Exeter	53974	27455	67	0.312	55.2	44.8
E07000087	Fareham	39882	22403	66	0.485	48.2	51.8
E07000010	Fenland	42609	24736	64	0.539	54.1	45.9
W06000005	Flintshire	57352	35436	62	—	60	40
E07000112	Folkestone-and-Hythe	52103	31384	63	0.599	53.2	46.8
E07000080	Forest-of-Dean	33106	21345	60	0.600	62.2	37.8
E07000119	Fylde	40243	25367	63	0.572	53.5	46.5
E08000037	Gateshead	81557	47087	65	0.513	44	56
E07000173	Gedling	50672	31921	63	0.561	47	53
E07000081	Gloucester	54473	29944	66	0.419	48.6	51.4
E07000088	Gosport	35801	19353	65	0.482	55	45
E07000109	Gravesham	38860	23159	64	0.544	49.9	50.1
E07000145	Great-Yarmouth	43445	27022	61	0.635	60	40
E09000011	Greenwich	121785	53851	69	0.500	54	46
E07000209	Guildford	55502	33221	64	0.612	51.7	48.3
W06000002	Gwynedd	48070	37187	52	—	68.9	31.1
E09000012	Hackney	117791	48154	69	0.489	63.2	36.8
E06000006	Halton	49234	25944	66	0.513	47	53
E09000013	Hammersmith-and-Fulham	102067	53375	67	0.507	57.5	42.5
E07000131	Harborough	37462	19982	66	0.547	52.4	47.6
E09000014	Haringey	124590	72665	65	0.545	56.4	43.6
E07000073	Harlow	39020	21046	67	0.488	50.2	49.8
E09000015	Harrow	84874	51184	64	0.574	54.1	45.9
E07000089	Hart	37102	18375	67	0.525	42.9	57.1
E06000001	Hartlepool	46082	27425	65	0.529	52.4	47.6
E07000062	Hastings	50047	30351	63	0.598	55.8	44.2
E07000090	Havant	48300	27732	65	0.511	53.3	46.7
E09000016	Havering	88089	56369	64	0.544	49.6	50.4
E06000019	Herefordshire-County-of	77500	50848	60	0.618	62.2	37.8
E07000098	Hertsmere	37623	20603	66	0.606	52.8	47.2

Table B.1: *(Continued)*

LA code	LA name	Number of houses with EPC (-)	Number of houses with EPC below 69 (-)	Average EPC rating (-)	BRPI (-)	Fabric-first (%)	System-led (%)
E07000037	High-Peak	33943	20902	63	0.575	53.2	46.8
E09000017	Hillingdon	108093	63760	65	0.540	55.7	44.3
E07000132	Hinckley-and-Bosworth	44722	25799	65	0.544	52.6	47.4
E07000227	Horsham	57436	31662	65	0.545	46.8	53.2
E09000018	Hounslow	97964	50777	67	0.530	53.9	46.1
E07000011	Huntingdonshire	73619	38895	66	0.516	52.7	47.3
E07000120	Hyndburn	36022	26746	60	0.597	60.1	39.9
E07000202	Ipswich	59688	33520	65	0.530	47.5	52.5
W06000001	Isle-of-Anglesey	33648	24351	55	—	60.1	39.9
E06000046	Isle-of-Wight	66331	40384	62	0.596	59.7	40.3
E06000053	Isles-of-Scilly	962	820	47	0.816	60.2	39.8
E09000019	Islington	113101	50510	68	0.476	63.4	36.6
E09000020	Kensington-and-Chelsea	104736	59242	64	0.539	57.3	42.7
E07000146	King-s-Lynn-and-West-Norfolk	72972	48839	60	0.778	60.5	39.5
E06000010	Kingston-upon-Hull-City-of	123044	76062	64	0.573	53.2	46.8
E09000021	Kingston-upon-Thames	71239	44212	63	0.555	57.3	42.7
E08000034	Kirklees	163669	105628	63	0.585	54.1	45.9
E08000011	Knowsley	71598	32005	67	0.509	46.2	53.8
E09000022	Lambeth	172157	89038	67	0.500	61.2	38.8
E07000121	Lancaster	59632	39114	62	0.571	49.2	50.8
E08000035	Leeds	366818	226070	63	0.577	48.6	51.4
E06000016	Leicester	144195	91806	63	0.608	52.7	47.3
E07000063	Lewes	41259	25178	63	0.545	52.1	47.9
E09000023	Lewisham	132879	72323	66	0.516	59.5	40.5
E07000194	Lichfield	39193	22282	65	0.564	53.1	46.9
E07000138	Lincoln	44352	25507	65	0.549	41	59
E08000012	Liverpool	234904	133305	64	0.565	53.1	46.9
E06000032	Luton	80090	49557	64	0.662	47.8	52.2
E07000110	Maidstone	72679	36389	67	0.522	48.5	51.5
E07000074	Maldon	23657	15419	61	0.769	58.7	41.3
E07000235	Malvern-Hills	33562	20858	61	0.447	59.2	40.8
E08000003	Manchester	302316	147700	67	0.554	53.8	46.2
E07000174	Mansfield	43660	26083	64	0.560	52.5	47.5
E06000035	Medway	110176	62997	65	0.525	50.6	49.4
E07000133	Melton	20806	13478	62	0.588	59.2	40.8
W06000024	Merthyr-Tydfil	28083	17287	64	—	40.3	59.7
E09000024	Merton	87153	51653	64	0.539	52.1	47.9
E07000042	Mid-Devon	33317	20464	61	0.587	60.4	39.6
E06000002	Middlesbrough	68018	38755	65	0.537	53.6	46.4
E07000203	Mid-Suffolk	40528	22768	64	0.607	57.4	42.6
E07000228	Mid-Sussex	60972	31231	67	0.523	53.2	46.8
E06000042	Milton-Keynes	120039	50051	70	0.447	40.2	59.8
E07000210	Mole-Valley	32804	20949	62	0.719	49.5	50.5
W06000021	Monmouthshire	37389	20849	64	—	53.4	46.6
W06000012	Neath-Port-Talbot	64444	41202	62	—	53.8	46.2
E07000175	Newark-and-Sherwood	52145	30938	64	0.566	40.6	59.4
E07000195	Newcastle-under-Lyme	45440	28987	63	0.591	48	52
E08000021	Newcastle-upon-Tyne	128707	71040	66	0.524	51.1	48.9
E07000091	New-Forest	68122	41814	63	0.540	49.2	50.8
E09000025	Newham	134752	65015	68	0.530	52.1	47.9
W06000022	Newport	66398	35425	66	—	49.3	50.7
E07000043	North-Devon	45042	28022	61	0.573	60	40
E07000038	North-East-Derbyshire	39134	25046	64	0.561	53.8	46.2
E06000012	North-East-Lincolnshire	66787	45533	62	0.587	56.6	43.4
E07000099	North-Hertfordshire	54488	29913	65	0.532	50.8	49.2
E07000139	North-Kesteven	48721	28054	64	0.544	37.7	62.3
E06000013	North-Lincolnshire	65255	41078	63	0.570	54.9	45.1
E07000147	North-Norfolk	50511	35662	59	0.885	62.1	37.9
E06000061	North-Northamptonshire	21176	8295	70	0.517	36.3	63.7
E06000024	North-Somerset	93441	50054	65	0.423	46.9	53.1
E08000022	North-Tyneside	87040	46309	66	0.498	46.2	53.8
E06000057	Northumberland	140408	84854	63	0.549	58.5	41.5
E07000218	North-Warwickshire	27961	17519	63	0.600	59.5	40.5
E07000134	North-West-Leicestershire	45368	26704	64	0.570	53.9	46.1
E06000065	North-Yorkshire	4699	2426	66	0.568	60.7	39.3
E07000148	Norwich	70959	36727	66	0.409	45.8	54.2
E06000018	Nottingham	181682	108993	64	0.586	44.5	55.5
E07000219	Nuneaton-and-Bedworth	50151	31288	64	0.565	54	46
E07000135	Oadby-and-Wigston	19977	13523	63	0.564	57.5	42.5

Table B.1: *(Continued)*

LA code	LA name	Number of houses with EPC (-)	Number of houses with EPC below 69 (-)	Average EPC rating (-)	BRPI (-)	Fabric-first (%)	System-led (%)
E08000004	Oldham	94824	61277	63	0.565	57	43
E07000178	Oxford	65961	34902	66	0.540	50.7	49.3
W06000009	Pembrokeshire	51438	34526	59	—	69.7	30.3
E07000122	Pendle	42120	32677	58	0.626	61.6	38.4
E06000031	Peterborough	94754	43368	68	0.450	40.1	59.9
E06000026	Plymouth	121630	68220	65	0.420	54.4	45.6
E06000044	Portsmouth	95970	58442	64	0.531	64.7	35.3
W06000023	Powys	54472	38613	56	—	63.9	36.1
E07000123	Preston	66806	37736	65	0.554	49.1	50.9
E06000038	Reading	74990	41852	65	0.531	52.2	47.8
E09000026	Redbridge	98115	63207	63	0.576	49.6	50.4
E06000003	Redcar-and-Cleveland	57718	36098	63	0.539	56.6	43.4
E07000236	Redditch	31061	17362	65	0.202	54.7	45.3
E07000211	Reigate-and-Banstead	61127	31708	66	0.541	51.7	48.3
W06000016	Rhondda-Cynon-Taf	91640	63270	62	—	61.4	38.6
E07000124	Ribble-Valley	23907	14187	64	0.588	57.3	42.7
E09000027	Richmond-upon-Thames	80668	51734	63	0.556	50.9	49.1
E08000005	Rochdale	86509	50528	65	0.553	48.1	51.9
E07000075	Rochford	28107	17403	64	0.544	53	47
E07000125	Rossendale	28707	19631	61	0.585	58.9	41.1
E07000064	Rother	43578	28629	61	0.674	50.6	49.4
E08000018	Rotherham	100197	60050	65	0.551	54.1	45.9
E07000220	Rugby	44646	22783	67	0.552	44.8	55.2
E07000212	Runnymede	35414	19961	65	0.606	50.9	49.1
E07000176	Rushcliffe	46707	27405	64	0.566	53.7	46.3
E07000092	Rushmoor	40860	21009	67	0.506	50	50
E06000017	Rutland	15491	9517	63	0.585	52.5	47.5
E08000006	Salford	147717	60459	69	0.519	36.4	63.6
E08000028	Sandwell	117233	73694	63	0.603	52.1	47.9
E08000014	Sefton	120280	77892	62	0.562	51.2	48.8
E07000111	Sevenoaks	40705	26403	62	0.637	48.1	51.9
E08000019	Sheffield	220430	130280	64	0.568	45	55
E06000051	Shropshire	131020	82089	61	0.604	60.8	39.2
E06000039	Slough	54030	28036	67	0.521	56.3	43.7
E08000029	Solihull	80042	47894	64	0.561	51.2	48.8
E06000066	Somerset	3892	1864	67	0.533	57.2	42.8
E06000045	Southampton	115281	65248	65	0.526	63	37
E07000012	South-Cambridgeshire	67271	34238	67	0.530	50.8	49.2
E07000039	South-Derbyshire	40641	20464	67	0.539	53.5	46.5
E06000033	Southend-on-Sea	74605	53275	60	0.769	50	50
E06000025	South-Gloucestershire	104626	55429	66	0.310	48.5	51.5
E07000044	South-Hams	42241	28693	59	0.586	61.1	38.9
E07000140	South-Holland	38273	24925	62	0.572	60.2	39.8
E07000141	South-Kesteven	64128	38293	63	0.546	53.3	46.7
E07000149	South-Norfolk	56010	30854	65	0.544	38.3	61.7
E07000179	South-Oxfordshire	54793	31922	64	0.606	43.1	56.9
E07000126	South-Ribble	40627	23814	65	0.526	52.6	47.4
E07000196	South-Staffordshire	35502	22335	63	0.573	53.9	46.1
E08000023	South-Tyneside	65431	37467	66	0.503	47.4	52.6
E09000028	Southwark	146926	62196	69	0.476	57.2	42.8
E07000213	Spelthorne	39114	23481	64	0.526	54.8	45.2
E07000197	Stafford	53598	31363	64	0.576	53.8	46.2
E07000198	Staffordshire-Moorlands	35284	25354	60	0.622	61.5	38.5
E07000240	St-Albans	57466	32532	65	0.604	47.8	52.2
E07000243	Stevenage	33141	17235	67	0.488	45.3	54.7
E08000013	St-Helens	69223	41545	64	0.537	49.1	50.9
E08000007	Stockport	110321	73203	63	0.559	54.2	45.8
E06000004	Stockton-on-Tees	88635	49333	66	0.516	36.2	63.8
E06000021	Stoke-on-Trent	106952	70695	63	0.617	57.9	42.1
E07000221	Stratford-on-Avon	68115	36071	65	0.482	54.1	45.9
E07000082	Stroud	47859	28865	63	0.564	45.7	54.3
E08000024	Sunderland	153320	86878	66	0.513	50.2	49.8
E07000214	Surrey-Heath	33044	19193	65	0.590	47.2	52.8
E09000029	Sutton	78984	46006	64	0.541	58.9	41.1
E07000113	Swale	57947	32881	65	0.533	47.9	52.1
W06000011	Swansea	94000	59217	62	—	57.3	42.7
E06000030	Swindon	89807	42606	68	0.297	45.2	54.8
E08000008	Tameside	99557	57733	65	0.536	53.7	46.3
E07000199	Tamworth	27381	15533	66	0.541	54.7	45.3

Table B.1: *(Continued)*

LA code	LA name	Number of houses with EPC (-)	Number of houses with EPC below 69 (-)	Average EPC rating (-)	BRPI (-)	Fabric-first (%)	System-led (%)
E07000215	Tandridge	32069	19081	64	0.659	51	49
E07000045	Teignbridge	57485	36033	62	0.554	54	46
E06000020	Telford-and-Wrekin	74833	35101	68	0.534	40.2	59.8
E07000076	Tendring	65506	44045	61	0.767	50	50
E07000093	Test-Valley	45536	24632	65	0.511	47.8	52.2
E07000083	Tewkesbury	36479	18305	67	0.413	47.1	52.9
E07000114	Thanet	74017	44578	63	0.547	55.2	44.8
E07000102	Three-Rivers	32957	19549	64	0.664	51.4	48.6
E06000034	Thurrock	62221	34536	66	0.517	51.3	48.7
E07000115	Tonbridge-and-Malling	47312	25348	66	0.525	46.2	53.8
E06000027	Torbay	63798	42211	62	0.555	56.9	43.1
W06000020	Torfaen	35742	19237	66	—	46.1	53.9
E07000046	Torridge	32046	21236	58	0.588	61.4	38.6
E09000030	Tower-Hamlets	162609	42883	74	0.443	64.5	35.5
E08000009	Trafford	88199	56651	63	0.563	52.5	47.5
E07000116	Tunbridge-Wells	48827	30234	63	0.626	53	47
E07000077	Uttlesford	33636	17204	66	0.658	48.7	51.3
W06000014	Vale-of-Glamorgan	54426	30667	65	—	47.8	52.2
E07000180	Vale-of-White-Horse	57209	26396	68	0.516	44.3	55.7
E08000036	Wakefield	178347	108436	66	0.546	55	45
E08000030	Walsall	108747	67981	63	0.598	52	48
E09000031	Waltham-Forest	116083	68436	65	0.542	54.3	45.7
E09000032	Wandsworth	171372	85754	67	0.507	32.7	67.3
E06000007	Warrington	82158	44122	66	0.518	46.9	53.1
E07000222	Warwick	65957	37088	65	0.467	47.8	52.2
E07000103	Watford	40664	21221	67	0.527	51	49
E07000216	Waverley	48387	29433	63	0.693	49.5	50.5
E07000065	Wealden	62505	38167	63	0.623	47.8	52.2
E07000241	Welwyn-Hatfield	42221	20485	67	0.518	49	51
E06000037	West-Berkshire	62720	37139	63	0.545	53.5	46.5
E07000047	West-Devon	24290	16193	58	0.617	61.6	38.4
E07000127	West-Lancashire	45474	26758	64	0.560	52.3	47.7
E07000142	West-Lindsey	42634	25513	63	0.581	58.5	41.5
E09000033	Westminster	150954	71056	67	0.503	61.4	38.6
E06000064	Westmorland-and-Furness	1987	1102	64	0.564	61.9	38.1
E06000062	West-Northamptonshire	24712	10184	69	0.521	47.5	52.5
E07000181	West-Oxfordshire	46834	24092	65	0.541	52.6	47.4
E07000245	West-Suffolk	79329	44302	65	0.527	52.1	47.9
E08000010	Wigan	125952	76995	64	0.537	47.1	52.9
E06000054	Wiltshire	202106	113288	65	0.541	54.3	45.7
E07000094	Winchester	51892	26029	67	0.532	44.3	55.7
E06000040	Windsor-and-Maidenhead	60500	35860	64	0.610	49.1	50.9
E08000015	Wirral	137947	91161	62	0.567	54.8	45.2
E07000217	Woking	43757	23467	66	0.533	52.5	47.5
E06000041	Wokingham	63271	31863	67	0.520	43	57
E08000031	Wolverhampton	101199	63540	63	0.619	49.6	50.4
E07000237	Worcester	45232	27175	64	0.362	47.2	52.8
E07000229	Worthing	47996	32079	62	0.531	60.7	39.3
W06000006	Wrexham	47362	28170	63	—	53	47
E07000238	Wychavon	55293	30084	65	0.372	44.3	55.7
E07000128	Wyre	47294	31329	62	0.563	58.7	41.3
E07000239	Wyre-Forest	38360	25021	62	0.375	56.7	43.3
E06000014	York	77555	46386	65	0.545	41.9	58.1

Table B.2: *Electoral constituency-level breakdown of the number of dwellings requiring retrofits (EPC < 69), BRPI values (for England only, indicated as "-" for Wales authorities due to unavailable fuel poverty components), and the estimated proportion of dwellings benefiting from fabric-first vs system-led measures across England and Wales.*

PCON code	Constituency name	Number of houses with EPC (-)	Number of houses with EPC below 69 (-)	Average EPC rating (-)	BRPI (-)	Fabric-first (%)	System-led (%)
E14001063	Aldershot	47785	24230	67	0.391	46.7	53.3
E14001064	Aldridge-Brownhills	28016	17437	64	0.568	52.0	48.0
E14001065	Altrincham and Sale West	37817	24109	63	0.575	52.5	47.5
E14001066	Amber Valley	37619	23814	63	0.580	59.2	40.8
E14001067	Arundel and South Downs	45194	27156	63	0.596	51.3	48.7
E14001068	Ashfield	46164	27988	63	0.571	47.4	52.6
E14001069	Ashford	54365	26591	67	0.495	49.7	50.3
E14001070	Ashton-under-Lyne	42336	25247	64	0.538	53.7	46.3
E14001071	Aylesbury	50040	23459	68	0.470	56.6	43.4
E14001072	Banbury	57351	27088	68	0.542	45.4	54.6
E14001073	Barking	57727	30379	68	0.350	58.0	42.0
E14001074	Barnsley North	41678	25112	65	0.547	46.5	53.5
E14001075	Barnsley South	42378	24917	65	0.560	46.5	53.5
E14001076	Barrow and Furness	40076	28288	60	0.547	55.3	44.7
E14001077	Basildon and Billericay	36425	19068	67	0.483	44.5	55.5
E14001078	Basingstoke	49494	19423	70	0.482	41.5	58.5
E14001079	Bassetlaw	53437	31283	64	0.564	57.1	42.9
E14001080	Bath	47118	28709	64	0.567	47.4	52.6
E14001081	Battersea	76620	32291	70	0.482	32.7	67.3
E14001082	Beaconsfield	45558	26370	64	0.541	56.6	43.4
E14001083	Beckenham and Penge	37510	22265	64	0.553	55.3	44.7
E14001084	Bedford	49614	26344	66	0.528	48.2	51.8
E14001085	Bermondsey and Old Southwark	80832	24765	72	0.443	57.2	42.8
E14001086	Bethnal Green and Stepney	71064	23663	71	0.460	64.5	35.5
E14001087	Beverley and Holderness	43172	27530	62	0.574	57.8	42.2
E14001088	Bexhill and Battle	50314	31463	63	0.700	49.0	51.0
E14001089	Bexleyheath and Crayford	38711	23116	65	0.769	53.8	46.2
E14001090	Bicester and Woodstock	42514	20702	67	0.518	45.4	54.6
E14001091	Birkenhead	47778	28475	64	0.571	54.8	45.2
E14001092	Birmingham Edgbaston	49347	30829	63	0.626	51.2	48.8
E14001093	Birmingham Erdington	45607	30764	62	0.622	51.0	49.0
E14001094	Birmingham Hall Green and Moseley	48359	36029	59	0.721	61.2	38.8
E14001095	Birmingham Hodge Hill and Solihull North	45788	35122	59	0.628	58.4	41.6
E14001096	Birmingham Ladywood	89897	37191	69	0.641	46.1	53.9
E14001097	Birmingham Northfield	42985	26015	64	0.571	44.3	55.7
E14001098	Birmingham Perry Barr	47048	35331	60	0.735	68.2	31.8
E14001099	Birmingham Selly Oak	45230	31526	61	0.740	65.0	35.0
E14001100	Birmingham Yardley	44343	33664	60	0.684	48.2	51.8
E14001101	Bishop Auckland	50845	29313	64	0.566	43.4	56.6
E14001102	Blackburn	50210	31366	63	0.595	57.6	42.4
E14001103	Blackley and Middleton South	61773	30075	66	0.550	52.3	47.7
E14001104	Blackpool North and Fleetwood	44815	32285	60	0.562	58.5	41.5
E14001105	Blackpool South	51190	38677	58	0.624	58.4	41.6
E14001106	Blaydon and Consett	34343	19929	65	0.524	43.5	56.5
E14001107	Blyth and Ashington	36655	18012	68	0.483	58.5	41.5
E14001108	Bognor Regis and Littlehampton	51154	29593	65	0.515	54.1	45.9
E14001109	Bolsover	43425	25202	65	0.568	55.3	44.7
E14001110	Bolton North East	43600	25725	64	0.559	48.4	51.6
E14001111	Bolton South and Walkden	43478	25647	64	0.552	42.8	57.2
E14001112	Bolton West	40775	24042	65	0.549	48.4	51.6
E14001113	Bootle	49950	29201	64	0.550	51.2	48.8
E14001114	Boston and Skegness	55303	34960	62	0.606	61.8	38.2
E14001115	Bournemouth East	56204	32173	64	0.469	59.6	40.4
E14001116	Bournemouth West	53355	30608	64	0.469	55.3	44.7
E14001117	Bracknell	46982	22542	68	0.458	43.4	56.6
E14001118	Bradford East	60569	43291	60	0.652	60.5	39.5
E14001119	Bradford South	46421	31644	62	0.596	57.6	42.4
E14001120	Bradford West	60464	43822	59	0.695	65.7	34.3
E14001121	Braintree	41968	23324	65	0.546	51.7	48.3
E14001122	Brent East	67847	33521	68	0.332	58.3	41.7
E14001123	Brent West	45708	23696	67	0.348	58.3	41.7
E14001124	Brentford and Isleworth	61331	29615	68	0.529	52.4	47.6

Table B.2: *(Continued)*

PCON code	Constituency	Number of houses with EPC (-)	Number of houses with EPC below 69 (-)	Average EPC rating (-)	BRPI (-)	Fabric-first (%)	System-led (%)
E14001125	Brentwood and Ongar	39888	23727	64	0.558	49.1	50.9
E14001126	Bridgwater	58732	35037	62	0.495	52.1	47.9
E14001127	Bridlington and The Wolds	46946	30448	61	0.596	57.8	42.2
E14001128	Brigg and Immingham	37217	24902	61	0.551	55.8	44.2
E14001129	Brighton Kemptown and Peacehaven	46274	25106	65	0.520	59.0	41.0
E14001130	Brighton Pavilion	58128	35799	63	0.548	61.0	39.0
E14001131	Bristol Central	56693	32314	64	0.556	49.0	51.0
E14001132	Bristol East	43980	25157	65	0.528	49.8	50.2
E14001133	Bristol North East	40132	22044	65	0.455	46.4	53.6
E14001134	Bristol North West	42282	24090	64	0.542	54.8	45.2
E14001135	Bristol South	49183	27258	66	0.520	49.8	50.2
E14001136	Broadland and Fakenham	41711	26466	63	0.559	59.8	40.2
E14001137	Bromley and Biggin Hill	38972	24619	63	0.655	55.3	44.7
E14001138	Bromsgrove	37280	22106	64	0.595	52.1	47.9
E14001139	Broxbourne	39243	21108	66	0.523	46.6	53.4
E14001140	Broxtowe	38494	25895	62	0.569	58.5	41.5
E14001141	Buckingham and Bletchley	49310	25551	66	0.484	45.1	54.9
E14001142	Burnley	48343	35053	60	0.612	61.4	38.6
E14001143	Burton and Uttoxeter	49964	28043	65	0.600	53.8	46.2
E14001144	Bury North	38953	25307	63	0.556	49.4	50.6
E14001145	Bury South	43396	27841	63	0.568	42.3	57.7
E14001146	Bury St Edmunds and Stowmarket	54496	26962	67	0.495	53.9	46.1
E14001147	Calder Valley	45589	30774	61	0.600	60.1	39.9
E14001148	Camborne and Redruth	44430	27435	60	0.598	61.8	38.2
E14001149	Cambridge	57873	26603	68	0.480	55.4	44.6
E14001150	Cannock Chase	41705	24492	65	0.564	53.7	46.3
E14001151	Canterbury	52019	28365	66	0.541	52.8	47.2
E14001152	Carlisle	38904	22348	65	0.539	53.9	46.1
E14001153	Carshalton and Wallington	39730	22528	65	0.538	58.9	41.1
E14001154	Castle Point	30467	22707	60	0.559	46.8	53.2
E14001155	Central Devon	42816	27676	60	0.611	57.5	42.5
E14001156	Central Suffolk and North Ipswich	40360	23887	63	0.532	54.9	45.1
E14001157	Chatham and Aylesford	42739	23767	66	0.495	49.3	50.7
E14001158	Cheadle	34659	23208	63	0.572	54.2	45.8
E14001159	Chelmsford	47689	24137	67	0.483	45.5	54.5
E14001160	Chelsea and Fulham	73168	38681	66	0.520	57.4	42.6
E14001161	Cheltenham	53554	30384	65	0.528	50.7	49.3
E14001162	Chesham and Amersham	36237	22559	63	0.561	56.6	43.4
E14001163	Chester North and Neston	34994	18904	66	0.535	54.0	46.0
E14001164	Chester South and Eddisbury	41076	23214	64	0.578	52.8	47.2
E14001165	Chesterfield	42462	25113	65	0.548	47.3	52.7
E14001166	Chichester	54224	29955	64	0.599	53.2	46.8
E14001167	Chingford and Woodford Green	36187	22211	64	0.550	52.0	48.0
E14001168	Chippenham	43737	22193	67	0.495	50.3	49.7
E14001169	Chipping Barnet	48602	28442	65	0.672	58.2	41.8
E14001170	Chorley	43527	22659	67	0.522	47.4	52.6
E14001171	Christchurch	38957	22448	65	0.482	56.0	44.0
E14001172	Cities of London and Westminster	100002	43196	68	0.340	60.6	39.4
E14001173	City of Durham	49830	27155	67	0.528	43.4	56.6
E14001174	Clacton	46129	31676	61	0.576	50.0	50.0
E14001175	Clapham and Brixton Hill	41396	26593	63	0.482	61.2	38.8
E14001176	Colchester	59658	27308	68	0.485	47.3	52.7
E14001177	Colne Valley	45903	30047	63	0.567	54.1	45.9
E14001178	Congleton	43920	24254	66	0.535	51.9	48.1
E14001179	Corby and East Northamptonshire	52787	25640	67	0.516	36.3	63.7
E14001180	Coventry East	57209	35280	64	0.605	52.9	47.1
E14001181	Coventry North West	46693	30750	63	0.596	50.3	49.7
E14001182	Coventry South	50558	29926	64	0.617	51.6	48.4
E14001183	Cramlington and Killingworth	41044	19200	69	0.485	53.1	46.9
E14001184	Crawley	40893	18817	68	0.392	47.8	52.2
E14001185	Crewe and Nantwich	51001	28685	66	0.536	51.9	48.1
E14001186	Croydon East	42731	26846	64	0.663	58.4	41.6
E14001187	Croydon South	46705	29064	64	0.625	55.1	44.9
E14001188	Croydon West	58624	30179	67	0.427	52.3	47.7
E14001189	Dagenham and Rainham	45388	26680	66	0.533	53.4	46.6
E14001190	Darlington	46299	28838	63	0.543	41.3	58.7
E14001191	Dartford	47991	22024	69	0.482	50.5	49.5
E14001192	Daventry	41872	22842	65	0.551	42.5	57.5
E14001193	Derby North	48438	28634	64	0.570	52.8	47.2

Table B.2: *(Continued)*

PCON code	Constituency	Number of houses with EPC (-)	Number of houses with EPC below 69 (-)	Average EPC rating (-)	BRPI (-)	Fabric-first (%)	System-led (%)
E14001194	Derby South	53545	33061	63	0.600	52.8	47.2
E14001195	Derbyshire Dales	36405	23592	61	0.617	58.0	42.0
E14001196	Dewsbury and Batley	41592	25374	64	0.606	54.1	45.9
E14001197	Didcot and Wantage	46787	21086	69	0.563	43.6	56.4
E14001198	Doncaster Central	51803	32179	64	0.577	58.2	41.8
E14001199	Doncaster East and the Isle of Axholme	37517	22731	64	0.555	57.1	42.9
E14001200	Doncaster North	45884	31533	62	0.583	58.2	41.8
E14001201	Dorking and Horley	37340	22134	64	0.649	50.3	49.7
E14001202	Dover and Deal	47969	27818	64	0.533	50.4	49.6
E14001203	Droitwich and Evesham	43797	23694	65	0.372	44.3	55.7
E14001204	Dudley	43584	25933	64	0.598	54.4	45.6
E14001205	Dulwich and West Norwood	58812	34831	64	0.529	59.6	40.4
E14001206	Dunstable and Leighton Buzzard	48251	22259	68	0.483	42.4	57.6
E14001207	Ealing Central and Acton	69375	36310	67	0.348	57.3	42.7
E14001208	Ealing North	49087	26374	66	0.554	57.2	42.8
E14001209	Ealing Southall	36697	20469	67	0.580	57.2	42.8
E14001210	Earley and Woodley	36771	17645	68	0.495	48.2	51.8
E14001211	Easington	48508	26656	66	0.517	43.4	56.6
E14001212	East Grinstead and Uckfield	51996	27614	66	0.574	50.7	49.3
E14001213	East Ham	63463	34415	67	0.598	52.1	47.9
E14001214	East Hampshire	43460	22777	66	0.536	42.5	57.5
E14001215	East Surrey	45874	25303	65	0.645	51.4	48.6
E14001216	East Thanet	60762	35730	64	0.550	55.2	44.8
E14001217	East Wiltshire	41229	23178	64	0.516	51.8	48.2
E14001218	East Worthing and Shoreham	40084	26460	63	0.533	60.2	39.8
E14001219	Eastbourne	55003	30894	65	0.495	59.3	40.7
E14001220	Eastleigh	48084	21535	69	0.420	47.7	52.3
E14001221	Edmonton and Winchmore Hill	48839	32402	63	0.544	54.9	45.1
E14001222	Ellesmere Port and Bromborough	37744	20484	66	0.516	54.4	45.6
E14001223	Eltham and Chislehurst	40210	23887	64	0.690	54.8	45.2
E14001224	Ely and East Cambridgeshire	47129	25463	65	0.532	51.7	48.3
E14001225	Enfield North	47064	29546	63	0.566	54.9	45.1
E14001226	Epping Forest	39535	23650	64	0.555	48.8	51.2
E14001227	Epsom and Ewell	41304	24624	64	0.642	52.7	47.3
E14001228	Erewash	40798	26217	62	0.566	58.3	41.7
E14001229	Erith and Thamesmead	49888	25551	67	0.544	53.9	46.1
E14001230	Esher and Walton	48281	27884	64	0.648	51.6	48.4
E14001231	Exeter	51687	26484	67	0.392	55.2	44.8
E14001232	Exmouth and Exeter East	46966	22002	68	0.480	51.8	48.2
E14001233	Fareham and Waterlooville	38553	20935	66	0.484	48.7	51.3
E14001234	Farnham and Bordon	41384	22812	65	0.633	46.8	53.2
E14001235	Faversham and Mid Kent	43009	21852	67	0.526	48.2	51.8
E14001236	Feltham and Heston	44794	23759	67	0.530	53.9	46.1
E14001237	Filton and Bradley Stoke	44638	21867	68	0.368	48.5	51.5
E14001238	Finchley and Golders Green	58496	33928	64	0.515	54.2	45.8
E14001239	Folkestone and Hythe	59385	34584	64	0.599	53.2	46.8
E14001240	Forest of Dean	37058	23566	61	0.598	55.2	44.8
E14001241	Frome and East Somerset	35159	20934	63	0.541	49.8	50.2
E14001242	Fylde	45195	27705	64	0.569	56.4	43.6
E14001243	Gainsborough	47111	27859	63	0.581	58.5	41.5
E14001244	Gateshead Central and Whickham	51996	27614	66	0.517	44.0	56.0
E14001245	Gedling	44212	28452	63	0.562	47.0	53.0
E14001246	Gillingham and Rainham	41117	25439	64	0.541	50.6	49.4
E14001247	Glastonbury and Somerton	37262	22072	63	0.556	54.1	45.9
E14001248	Gloucester	55460	29303	66	0.448	48.6	51.4
E14001249	Godalming and Ash	39850	23593	64	0.665	50.7	49.3
E14001250	Goole and Pocklington	44106	25642	64	0.571	57.8	42.2
E14001251	Gorton and Denton	52377	34343	63	0.598	53.8	46.2
E14001252	Gosport	42940	23063	65	0.483	51.4	48.6
E14001253	Grantham and Bourne	47931	28020	64	0.550	46.7	53.3
E14001254	Gravesham	43051	24264	65	0.544	49.9	50.1
E14001255	Great Grimsby and Cleethorpes	53127	37533	61	0.608	56.6	43.4
E14001256	Great Yarmouth	46694	28632	62	0.555	60.0	40.0
E14001257	Greenwich and Woolwich	73268	26118	72	0.469	54.0	46.0
E14001258	Guildford	46953	27437	64	0.612	51.7	48.3
E14001259	Hackney North and Stoke Newington	63134	28521	68	0.518	63.2	36.8
E14001260	Hackney South and Shoreditch	63107	21802	71	0.465	63.2	36.8
E14001261	Halesowen	37220	23972	63	0.584	53.2	46.8
E14001262	Halifax	54113	37293	60	0.610	60.1	39.9

Table B.2: *(Continued)*

PCON code	Constituency	Number of houses with EPC (-)	Number of houses with EPC below 69 (-)	Average EPC rating (-)	BRPI (-)	Fabric-first (%)	System-led (%)
E14001263	Hamble Valley	42642	18909	70	0.454	46.6	53.4
E14001264	Hammersmith and Chiswick	66727	34556	66	0.518	55.7	44.3
E14001265	Hampstead and Highgate	67508	38418	65	0.422	57.8	42.2
E14001266	Harborough, Oadby and Wigston	40213	23343	65	0.553	54.5	45.5
E14001267	Harlow	49326	24736	68	0.482	49.2	50.8
E14001268	Harpenden and Berkhamsted	38306	22603	64	0.568	46.8	53.2
E14001269	Harrogate and Knaresborough	49030	30403	64	0.541	52.3	47.7
E14001270	Harrow East	34889	21704	64	0.597	56.5	43.5
E14001271	Harrow West	45111	24499	66	0.548	54.1	45.9
E14001272	Hartlepool	50223	28993	65	0.529	52.4	47.6
E14001273	Harwich and North Essex	41324	24272	64	0.553	48.7	51.3
E14001274	Hastings and Rye	62057	37771	63	0.609	53.3	46.7
E14001275	Havant	40020	22559	65	0.495	53.3	46.7
E14001276	Hayes and Harlington	47065	26035	66	0.536	55.7	44.3
E14001277	Hazel Grove	29767	19556	63	0.549	54.2	45.8
E14001278	Hemel Hempstead	46334	24347	66	0.495	45.9	54.1
E14001279	Hendon	63908	29576	69	0.518	54.2	45.8
E14001280	Henley and Thame	37738	22744	63	0.606	43.1	56.9
E14001281	Hereford and South Herefordshire	44725	27197	62	0.595	62.2	37.8
E14001282	Herne Bay and Sandwich	46463	27657	64	0.543	53.1	46.9
E14001283	Hertford and Stortford	47453	23934	67	0.518	43.8	56.2
E14001284	Hertsmere	40623	21598	66	0.551	50.9	49.1
E14001285	Hexham	33205	22693	59	0.603	55.1	44.9
E14001286	Heywood and Middleton North	44120	24106	65	0.545	48.1	51.9
E14001287	High Peak	36255	21998	64	0.575	53.2	46.8
E14001288	Hinckley and Bosworth	42574	24794	64	0.551	53.3	46.7
E14001289	Hitchin	43994	19987	68	0.519	45.3	54.7
E14001290	Holborn and St Pancras	72098	32547	68	0.312	59.6	40.4
E14001291	Honiton and Sidmouth	44697	27265	62	0.559	53.2	46.8
E14001292	Hornchurch and Upminster	38995	25491	63	0.547	49.6	50.4
E14001293	Hornsey and Friern Barnet	52612	31728	64	0.545	55.2	44.8
E14001294	Horsham	49982	24407	67	0.519	46.8	53.2
E14001295	Houghton and Sunderland South	50137	27926	67	0.495	50.2	49.8
E14001296	Hove and Portslade	57816	36237	63	0.537	61.0	39.0
E14001297	Huddersfield	42803	28217	62	0.612	54.1	45.9
E14001298	Huntingdon	54330	25408	68	0.524	52.7	47.3
E14001299	Hyndburn	43437	31640	60	0.597	59.6	40.4
E14001300	Ilford North	37675	24483	64	0.580	49.6	50.4
E14001301	Ilford South	48264	29775	64	0.576	53.1	46.9
E14001302	Ipswich	55464	30286	65	0.534	47.5	52.5
E14001303	Isle of Wight East	37018	23212	62	0.596	59.7	40.3
E14001304	Isle of Wight West	34557	19703	63	0.596	59.7	40.3
E14001305	Islington North	58386	27870	68	0.480	63.4	36.6
E14001306	Islington South and Finsbury	63447	24954	70	0.466	63.3	36.7
E14001307	Jarrow and Gateshead East	41966	23653	66	0.483	45.5	54.5
E14001308	Keighley and Ilkley	44329	31186	60	0.611	60.5	39.5
E14001309	Kenilworth and Southam	41509	23263	65	0.598	49.5	50.5
E14001310	Kensington and Bayswater	105323	57218	65	0.533	59.5	40.5
E14001311	Kettering	46774	24273	67	0.521	36.3	63.7
E14001312	Kingston and Surbiton	59385	35890	63	0.547	57.3	42.7
E14001313	Kingston upon Hull East	42001	25316	65	0.552	53.2	46.8
E14001314	Kingston upon Hull North and Cottingham	55929	34617	64	0.625	62.7	37.3
E14001315	Kingston upon Hull West and Haltemprice	52298	32041	64	0.576	50.7	49.3
E14001316	Kingswinford and South Staffordshire	30981	20189	63	0.566	54.3	45.7
E14001317	Knowsley	56442	24105	68	0.485	46.2	53.8
E14001318	Lancaster and Wyre	40588	23995	64	0.562	53.4	46.6
E14001319	Leeds Central and Headingley	64335	36495	64	0.634	48.6	51.4
E14001320	Leeds East	49173	31247	63	0.595	50.7	49.3
E14001321	Leeds North East	43900	29271	62	0.596	49.6	50.4
E14001322	Leeds North West	41644	28012	62	0.565	44.6	55.4
E14001323	Leeds South	76738	41408	64	0.596	58.5	41.5
E14001324	Leeds South West and Morley	40683	24000	65	0.533	44.6	55.4
E14001325	Leeds West and Pudsey	46700	29879	62	0.556	44.8	55.2
E14001326	Leicester East	43540	27286	64	0.600	52.7	47.3
E14001327	Leicester South	59395	37907	62	0.647	59.1	40.9
E14001328	Leicester West	50012	31044	63	0.581	51.5	48.5
E14001329	Leigh and Atherton	46403	26616	65	0.533	47.1	52.9
E14001330	Lewes	39001	24266	63	0.562	49.5	50.5
E14001331	Lewisham East	49518	29370	65	0.537	59.0	41.0

Table B.2: *(Continued)*

PCON code	Constituency	Number of houses with EPC (-)	Number of houses with EPC below 69 (-)	Average EPC rating (-)	BRPI (-)	Fabric-first (%)	System-led (%)
E14001332	Lewisham North	59040	25432	69	0.484	59.5	40.5
E14001333	Lewisham West and East Dulwich	48843	31005	64	0.528	58.4	41.6
E14001334	Leyton and Wanstead	50555	29524	65	0.553	52.0	48.0
E14001335	Lichfield	38967	21454	66	0.560	53.5	46.5
E14001336	Lincoln	52572	29664	65	0.545	39.3	60.7
E14001337	Liverpool Garston	41864	22706	65	0.544	53.1	46.9
E14001338	Liverpool Riverside	82048	35298	68	0.538	49.2	50.8
E14001339	Liverpool Walton	52067	31236	63	0.558	52.4	47.6
E14001340	Liverpool Wavertree	48881	32684	61	0.620	53.1	46.9
E14001341	Liverpool West Derby	41255	23340	64	0.573	51.8	48.2
E14001342	Loughborough	46129	24418	66	0.555	51.0	49.0
E14001343	Louth and Horncastle	51097	34865	60	0.610	62.4	37.6
E14001344	Lowestoft	46096	28144	63	0.544	57.8	42.2
E14001345	Luton North	35854	21739	64	0.545	47.8	52.2
E14001346	Luton South and South Bedfordshire	52675	31525	64	0.601	44.9	55.1
E14001347	Macclesfield	42562	26161	64	0.558	51.9	48.1
E14001348	Maidenhead	44330	25466	65	0.574	46.8	53.2
E14001349	Maidstone and Malling	49526	22755	68	0.485	47.6	52.4
E14001350	Makerfield	36616	22878	64	0.535	47.1	52.9
E14001351	Maldon	36298	21844	64	0.550	49.4	50.6
E14001352	Manchester Central	100016	32832	71	0.534	54.7	45.3
E14001353	Manchester Rusholme	56166	26526	67	0.564	53.8	46.2
E14001354	Manchester Withington	50774	30755	64	0.564	53.8	46.2
E14001355	Mansfield	46602	27289	64	0.559	52.5	47.5
E14001356	Melksham and Devizes	38800	21440	65	0.538	54.3	45.7
E14001357	Melton and Syston	38309	23016	63	0.570	53.1	46.9
E14001358	Meriden and Solihull East	38825	22315	65	0.562	51.2	48.8
E14001359	Mid Bedfordshire	51139	22612	69	0.495	44.7	55.3
E14001360	Mid Buckinghamshire	11290	5083	68	0.551	56.6	43.4
E14001361	Mid Cheshire	40508	21308	67	0.519	52.8	47.2
E14001362	Mid Derbyshire	32296	21192	63	0.550	55.6	44.4
E14001363	Mid Dorset and North Poole	31530	17911	66	0.483	56.0	44.0
E14001364	Mid Leicestershire	36593	21763	65	0.554	49.1	50.9
E14001365	Mid Norfolk	48205	27666	65	0.557	48.9	51.1
E14001366	Mid Sussex	49666	24813	67	0.495	53.2	46.8
E14001367	Middlesbrough and Thornaby East	60322	35948	64	0.560	43.9	56.1
E14001368	Middlesbrough South and East Cleveland	42885	23558	65	0.517	55.0	45.0
E14001369	Milton Keynes Central	58990	20106	72	0.447	40.2	59.8
E14001370	Milton Keynes North	64379	26356	70	0.447	40.2	59.8
E14001371	Mitcham and Morden	44181	25332	65	0.530	52.1	47.9
E14001372	Morecambe and Lunesdale	41169	27989	61	0.579	53.8	46.2
E14001373	New Forest East	34081	20747	63	0.529	49.2	50.8
E14001374	New Forest West	39255	23487	64	0.539	49.2	50.8
E14001375	Newark	46574	28561	63	0.584	50.5	49.5
E14001376	Newbury	50237	28907	64	0.534	53.5	46.5
E14001377	Newcastle upon Tyne Central and West	57653	30346	66	0.524	51.1	48.9
E14001378	Newcastle upon Tyne East and Wallsend	54688	29103	66	0.483	49.2	50.8
E14001379	Newcastle upon Tyne North	42731	20214	68	0.531	49.2	50.8
E14001380	Newcastle-under-Lyme	36304	22049	64	0.598	48.0	52.0
E14001381	Newton Abbot	42804	25849	64	0.535	54.0	46.0
E14001382	Newton Aycliffe and Spennymoor	54659	30510	66	0.514	43.4	56.6
E14001383	Normanton and Hemsworth	54859	34206	66	0.552	55.0	45.0
E14001384	North Bedfordshire	42188	20068	67	0.495	44.7	55.3
E14001385	North Cornwall	50117	32021	59	0.601	61.8	38.2
E14001386	North Cotswolds	42424	25246	62	0.564	50.2	49.8
E14001387	North Devon	47798	29517	61	0.573	60.0	40.0
E14001388	North Dorset	42341	24800	63	0.553	52.0	48.0
E14001389	North Durham	39379	23098	65	0.523	43.4	56.6
E14001390	North East Cambridgeshire	54688	30424	65	0.539	54.1	45.9
E14001391	North East Derbyshire	38142	23566	64	0.558	50.4	49.6
E14001392	North East Hampshire	43813	20845	68	0.516	42.0	58.0
E14001393	North East Hertfordshire	42572	23865	65	0.546	47.3	52.7
E14001394	North East Somerset and Hanham	34211	19635	65	0.475	48.0	52.0
E14001395	North Herefordshire	38284	26264	59	0.647	62.2	37.8
E14001396	North Norfolk	45406	31540	59	0.596	62.1	37.9
E14001397	North Northumberland	48509	30040	61	0.561	58.5	41.5
E14001398	North Shropshire	47422	29235	62	0.607	60.8	39.2
E14001399	North Somerset	43952	23159	66	0.457	46.9	53.1
E14001400	North Warwickshire and Bedworth	38762	25085	63	0.585	56.0	44.0

Table B.2: *(Continued)*

PCON code	Constituency	Number of houses with EPC (-)	Number of houses with EPC below 69 (-)	Average EPC rating (-)	BRPI (-)	Fabric-first (%)	System-led (%)
E14001401	North West Cambridgeshire	64080	29620	68	0.495	46.1	53.9
E14001402	North West Essex	42195	20570	67	0.551	46.5	53.5
E14001403	North West Hampshire	47405	22984	67	0.474	44.2	55.8
E14001404	North West Leicestershire	49307	28129	65	0.566	53.9	46.1
E14001405	North West Norfolk	52279	33181	62	0.568	60.5	39.5
E14001406	Northampton North	36996	22327	64	0.523	47.5	52.5
E14001407	Northampton South	49396	25349	67	0.480	47.5	52.5
E14001408	Norwich North	39881	22650	66	0.495	50.9	49.1
E14001409	Norwich South	56683	28226	67	0.495	42.4	57.6
E14001410	Nottingham East	63877	39779	62	0.616	44.5	55.5
E14001411	Nottingham North and Kimberley	55514	33592	64	0.558	47.4	52.6
E14001412	Nottingham South	69051	38410	65	0.581	44.5	55.5
E14001413	Nuneaton	41981	24415	65	0.569	56.0	44.0
E14001414	Old Bexley and Sidcup	31513	21495	62	0.744	53.8	46.2
E14001415	Oldham East and Saddleworth	44456	29243	63	0.574	59.0	41.0
E14001416	Oldham West, Chadderton and Royton	46950	28084	64	0.560	57.0	43.0
E14001417	Orpington	32588	21392	63	0.779	55.3	44.7
E14001418	Ossett and Denby Dale	41044	26196	64	0.560	54.6	45.4
E14001419	Oxford East	54440	27906	66	0.525	50.7	49.3
E14001420	Oxford West and Abingdon	47594	24979	66	0.527	47.9	52.1
E14001421	Peckham	51957	21341	70	0.473	57.2	42.8
E14001422	Pendle and Clitheroe	50536	35127	61	0.605	60.3	39.7
E14001423	Penistone and Stocksbridge	33003	20290	64	0.546	45.5	54.5
E14001424	Penrith and Solway	41550	28249	59	0.582	59.5	40.5
E14001425	Peterborough	60987	30633	66	0.520	40.1	59.9
E14001426	Plymouth Moor View	43707	21794	67	0.456	54.4	45.6
E14001427	Plymouth Sutton and Devonport	62853	36912	64	0.456	54.4	45.6
E14001428	Pontefract, Castleford and Knottingley	57110	32076	67	0.531	55.0	45.0
E14001429	Poole	49117	25145	66	0.469	59.6	40.4
E14001430	Poplar and Limehouse	102390	20904	76	0.423	64.5	35.5
E14001431	Portsmouth North	9702	4759	67	0.523	66.7	33.3
E14001432	Portsmouth South	60414	34825	65	0.539	62.8	37.2
E14001433	Preston	50853	28402	65	0.556	49.1	50.9
E14001434	Putney	50162	24802	67	0.483	32.7	67.3
E14001435	Queen's Park and Maida Vale	63080	31986	67	0.452	59.9	40.1
E14001436	Rawmarsh and Conisbrough	39481	23791	65	0.551	56.5	43.5
E14001437	Rayleigh and Wickford	33853	20732	64	0.529	47.3	52.7
E14001438	Reading Central	59149	32890	65	0.540	52.2	47.8
E14001439	Reading West and Mid Berkshire	35674	22599	63	0.548	52.8	47.2
E14001440	Redcar	38863	23118	64	0.530	56.6	43.4
E14001441	Redditch	35148	19442	65	0.495	48.1	51.9
E14001442	Reigate	47175	24358	66	0.553	51.7	48.3
E14001443	Ribble Valley	41695	24343	65	0.565	51.7	48.3
E14001444	Richmond and Northallerton	39454	27022	60	0.620	62.2	37.8
E14001445	Richmond Park	56310	34714	63	0.566	53.8	46.2
E14001446	Rochdale	48794	27452	65	0.568	48.1	51.9
E14001447	Rochester and Strood	50557	24900	67	0.495	50.6	49.4
E14001448	Romford	39934	25099	64	0.548	49.6	50.4
E14001449	Romsey and Southampton North	36519	21249	64	0.552	58.8	41.2
E14001450	Rossendale and Darwen	44379	29193	62	0.582	58.0	42.0
E14001451	Rother Valley	37600	23383	64	0.555	54.1	45.9
E14001452	Rotherham	38713	21551	65	0.551	53.8	46.2
E14001453	Rugby	46091	23337	67	0.549	50.1	49.9
E14001454	Ruislip, Northwood and Pinner	38045	23597	63	0.577	55.0	45.0
E14001455	Runcorn and Helsby	38435	20301	66	0.515	52.3	47.7
E14001456	Runnymede and Weybridge	50004	27488	65	0.623	51.3	48.7
E14001457	Rushcliffe	42880	24898	65	0.568	53.7	46.3
E14001458	Rutland and Stamford	41508	24962	63	0.575	52.9	47.1
E14001459	Salford	83503	26993	71	0.495	36.4	63.6
E14001460	Salisbury	41261	21841	66	0.550	54.3	45.7
E14001461	Scarborough and Whitby	51625	34137	61	0.578	57.3	42.7
E14001462	Scunthorpe	37044	21968	64	0.564	54.9	45.1
E14001463	Sefton Central	30755	19711	63	0.550	51.2	48.8
E14001464	Selby	38450	21622	65	0.522	48.9	51.1
E14001465	Sevenoaks	35145	22493	62	0.629	49.1	50.9
E14001466	Sheffield Brightside and Hillsborough	42716	26928	63	0.584	45.0	55.0
E14001467	Sheffield Central	67643	31928	67	0.619	44.0	56.0
E14001468	Sheffield Hallam	33308	22686	62	0.596	47.0	53.0
E14001469	Sheffield Heeley	38570	22800	64	0.542	45.5	54.5

Table B.2: *(Continued)*

PCON code	Constituency	Number of houses with EPC (-)	Number of houses with EPC below 69 (-)	Average EPC rating (-)	BRPI (-)	Fabric-first (%)	System-led (%)
E14001470	Sheffield South East	34474	20206	65	0.538	45.0	55.0
E14001471	Sherwood Forest	45278	24190	66	0.544	43.6	56.4
E14001472	Shipley	40844	26583	63	0.566	60.5	39.5
E14001473	Shrewsbury	42415	22924	65	0.555	60.8	39.2
E14001474	Sittingbourne and Sheppey	49018	26833	66	0.525	47.9	52.1
E14001475	Skipton and Ripon	45420	29610	61	0.598	54.5	45.5
E14001476	Sleaford and North Hykeham	53539	31238	64	0.544	37.7	62.3
E14001477	Slough	55239	27324	67	0.524	56.3	43.7
E14001478	Smethwick	44693	30917	61	0.648	52.1	47.9
E14001479	Solihull West and Shirley	37028	22690	64	0.565	51.2	48.8
E14001480	South Basildon and East Thurrock	37342	21607	65	0.519	47.8	52.2
E14001481	South Cambridgeshire	52853	26294	67	0.556	52.8	47.2
E14001482	South Cotswolds	41602	24747	63	0.571	53.3	46.7
E14001483	South Derbyshire	43768	21530	67	0.542	53.5	46.5
E14001484	South Devon	47597	32641	59	0.579	58.6	41.4
E14001485	South Dorset	43985	25368	64	0.536	52.0	48.0
E14001486	South East Cornwall	43810	29246	58	0.596	61.8	38.2
E14001487	South Holland and The Deepings	48344	30101	63	0.563	56.0	44.0
E14001488	South Leicestershire	40565	22900	66	0.543	46.7	53.3
E14001489	South Norfolk	47481	26449	65	0.524	38.3	61.7
E14001490	South Northamptonshire	50988	23814	68	0.542	42.5	57.5
E14001491	South Ribble	37403	21833	65	0.532	50.0	50.0
E14001492	South Shields	40977	23138	66	0.495	47.4	52.6
E14001493	South Shropshire	41585	29641	57	0.657	60.8	39.2
E14001494	South Suffolk	40911	24170	63	0.552	57.9	42.1
E14001495	South West Devon	38505	22053	65	0.485	57.1	42.9
E14001496	South West Hertfordshire	43422	24825	65	0.565	47.9	52.1
E14001497	South West Norfolk	52257	32848	62	0.564	59.6	40.4
E14001498	South West Wiltshire	44140	25683	63	0.533	54.3	45.7
E14001499	Southampton Itchen	53910	28871	65	0.495	63.0	37.0
E14001500	Southampton Test	56557	32135	64	0.535	63.0	37.0
E14001501	Southend East and Rochford	46132	28936	63	0.560	50.7	49.3
E14001502	Southend West and Leigh	42235	31702	59	0.576	50.0	50.0
E14001503	Southgate and Wood Green	55337	33781	64	0.567	55.6	44.4
E14001504	Southport	46580	31645	60	0.598	51.5	48.5
E14001505	Spelthorne	44110	25117	65	0.526	54.8	45.2
E14001506	Spen Valley	36874	22226	64	0.551	54.1	45.9
E14001507	St Albans	44055	23526	66	0.532	47.8	52.2
E14001508	St Austell and Newquay	55497	33128	62	0.580	61.8	38.2
E14001509	St Helens North	43006	25862	65	0.537	49.1	50.9
E14001510	St Helens South and Whiston	50856	26360	66	0.534	47.8	52.2
E14001511	St Ives	47260	33783	56	0.646	61.8	38.2
E14001512	St Neots and Mid Cambridgeshire	46884	18218	70	0.483	51.8	48.2
E14001513	Stafford	41005	23519	65	0.581	51.0	49.0
E14001514	Staffordshire Moorlands	29546	21420	59	0.623	61.5	38.5
E14001515	Stalybridge and Hyde	45850	24734	66	0.534	53.7	46.3
E14001516	Stevenage	41563	20786	67	0.482	46.9	53.1
E14001517	Stockport	40421	25211	63	0.555	54.2	45.8
E14001518	Stockton North	48653	26488	66	0.520	36.2	63.8
E14001519	Stockton West	45738	24812	66	0.495	38.2	61.8
E14001520	Stoke-on-Trent Central	42652	27165	63	0.700	57.9	42.1
E14001521	Stoke-on-Trent North	43322	27523	63	0.604	55.0	45.0
E14001522	Stoke-on-Trent South	35943	24044	63	0.601	57.6	42.4
E14001523	Stone, Great Wyrley and Penkridge	35197	21765	63	0.569	53.8	46.2
E14001524	Stourbridge	36184	23279	63	0.597	54.4	45.6
E14001525	Stratford and Bow	62305	22303	71	0.483	57.0	43.0
E14001526	Stratford-on-Avon	52455	27397	65	0.579	54.1	45.9
E14001527	Streatham and Croydon North	62530	36519	65	0.483	58.2	41.8
E14001528	Stretford and Urmston	42463	26022	64	0.555	52.5	47.5
E14001529	Stroud	45849	26865	64	0.559	45.7	54.3
E14001530	Suffolk Coastal	50402	31015	63	0.552	57.8	42.2
E14001531	Sunderland Central	58864	36770	64	0.554	50.2	49.8
E14001532	Surrey Heath	44041	24502	66	0.597	50.1	49.9
E14001533	Sussex Weald	38462	22184	64	0.623	47.8	52.2
E14001534	Sutton and Cheam	44331	25623	64	0.544	58.9	41.1
E14001535	Sutton Coldfield	35944	24021	62	0.585	51.2	48.8
E14001536	Swindon North	42861	17664	69	0.312	45.2	54.8
E14001537	Swindon South	48258	24360	67	0.312	45.2	54.8
E14001538	Tamworth	36659	20926	65	0.560	53.8	46.2

Table B.2: *(Continued)*

PCON code	Constituency	Number of houses with EPC (-)	Number of houses with EPC below 69 (-)	Average EPC rating (-)	BRPI (-)	Fabric-first (%)	System-led (%)
E14001539	Tatton	39529	23742	64	0.578	51.7	48.3
E14001540	Taunton and Wellington	57190	29898	66	0.483	53.4	46.6
E14001541	Telford	46987	19913	69	0.526	40.2	59.8
E14001542	Tewkesbury	45469	22854	67	0.457	49.1	50.9
E14001543	The Wrekin	42096	21503	67	0.556	54.6	45.4
E14001544	Thirsk and Malton	51761	33324	61	0.582	59.8	40.2
E14001545	Thornbury and Yate	34068	18217	66	0.368	48.5	51.5
E14001546	Thurrock	51328	26841	67	0.495	51.3	48.7
E14001547	Tipton and Wednesbury	40945	22596	65	0.576	53.2	46.8
E14001548	Tiverton and Minehead	49310	29693	62	0.595	55.8	44.2
E14001549	Tonbridge	40688	22284	65	0.582	47.2	52.8
E14001550	Tooting	56733	31930	66	0.530	32.7	67.3
E14001551	Torbay	52012	33776	62	0.560	56.9	43.1
E14001552	Torridge and Tavistock	49405	32750	59	0.595	61.5	38.5
E14001553	Tottenham	68380	36974	66	0.527	59.1	40.9
E14001554	Truro and Falmouth	46434	29276	60	0.572	61.8	38.2
E14001555	Tunbridge Wells	49509	29697	63	0.626	53.0	47.0
E14001556	Twickenham	49247	30783	63	0.545	50.9	49.1
E14001557	Tynemouth	43300	24108	66	0.520	46.2	53.8
E14001558	Uxbridge and South Ruislip	42635	25118	65	0.533	55.7	44.3
E14001559	Vauxhall and Camberwell Green	78594	33062	69	0.466	59.6	40.4
E14001560	Wakefield and Rothwell	54243	33423	65	0.547	50.7	49.3
E14001561	Wallasey	42401	28244	62	0.575	54.8	45.2
E14001562	Walsall and Bloxwich	86883	53292	63	0.640	52.0	48.0
E14001563	Walthamstow	55657	31404	66	0.537	54.3	45.7
E14001564	Warrington North	43108	22449	67	0.518	46.9	53.1
E14001565	Warrington South	50595	25433	67	0.515	46.9	53.1
E14001566	Warwick and Leamington	52689	27955	66	0.550	47.8	52.2
E14001567	Washington and Gateshead South	50104	24515	68	0.483	48.0	52.0
E14001568	Watford	52980	28054	67	0.528	51.9	48.1
E14001569	Waveney Valley	45156	26365	64	0.584	52.7	47.3
E14001570	Weald of Kent	37214	20076	66	0.598	49.1	50.9
E14001571	Wellingborough and Rushden	48861	24697	67	0.495	36.3	63.7
E14001572	Wells and Mendip Hills	49301	28045	64	0.522	50.2	49.8
E14001573	Welwyn Hatfield	43501	20117	68	0.495	49.0	51.0
E14001574	West Bromwich	40371	24179	64	0.584	52.1	47.9
E14001575	West Dorset	48242	28474	62	0.567	52.0	48.0
E14001576	West Ham and Beckton	83738	34731	70	0.482	52.1	47.9
E14001577	West Lancashire	42068	24095	64	0.559	52.3	47.7
E14001578	West Suffolk	57091	32734	64	0.539	52.1	47.9
E14001579	West Worcestershire	44232	26884	62	0.431	50.4	49.6
E14001580	Westmorland and Lonsdale	42037	28000	60	0.583	61.0	39.0
E14001581	Weston-super-Mare	56040	29411	66	0.457	46.9	53.1
E14001582	Wetherby and Easingwold	35914	23170	62	0.584	50.5	49.5
E14001583	Whitehaven and Workington	37054	24881	62	0.554	57.6	42.4
E14001584	Widnes and Halewood	37249	18991	66	0.495	46.6	53.4
E14001585	Wigan	44812	27176	64	0.544	47.1	52.9
E14001586	Wimbledon	48620	28642	64	0.559	54.5	45.5
E14001587	Winchester	42738	22650	66	0.538	44.3	55.7
E14001588	Windsor	46831	26341	65	0.598	51.9	48.1
E14001589	Wirral West	29750	19542	63	0.562	54.8	45.2
E14001590	Witham	37197	19960	65	0.546	51.5	48.5
E14001591	Witney	50460	25704	66	0.525	48.3	51.7
E14001592	Woking	48656	26019	66	0.533	52.5	47.5
E14001593	Wokingham	47879	22071	68	0.530	43.0	57.0
E14001594	Wolverhampton North East	35061	19971	65	0.596	50.8	49.2
E14001595	Wolverhampton South East	38536	22043	65	0.598	50.8	49.2
E14001596	Wolverhampton West	38878	27412	60	0.645	49.6	50.4
E14001597	Worcester	48405	28191	64	0.362	47.2	52.8
E14001598	Worsley and Eccles	53239	25295	67	0.527	42.4	57.6
E14001599	Worthing West	49518	31497	63	0.528	56.9	43.1
E14001600	Wycombe	51763	28444	65	0.516	56.6	43.4
E14001601	Wyre Forest	41613	26592	62	0.375	56.7	43.3
E14001602	Wythenshawe and Sale East	61144	31342	66	0.516	53.4	46.6
E14001603	Yeovil	51894	28554	64	0.524	54.1	45.9
E14001604	York Central	50900	28509	66	0.553	46.9	53.1
E14001605	York Outer	33811	20650	65	0.536	37.7	62.3

References

[1] J. Akroyd, S. Mosbach, A. Bhave, and M. Kraft. Universal digital twin – a dynamic knowledge graph. *Data-Centric Engineering*, 2:e14, 2021. doi:10.1017/dce.2021.10.

[2] J. Alabid, A. Bennadji, and M. Seddiki. A review on the energy retrofit policies and improvements of the UK existing buildings, challenges and benefits. *Renewable and Sustainable Energy Reviews*, 159:112161, 2022. doi:10.1016/j.rser.2022.112161.

[3] U. Ali, M. H. Shamsi, M. Bohacek, C. Hoare, K. Purcell, E. Mangina, and J. O'Donnell. A data-driven approach to optimize urban scale energy retrofit decisions for residential buildings. *Applied Energy*, 267:114861, 2020. doi:10.1016/j.apenergy.2020.114861.

[4] M. Amith, L. Cui, K. Roberts, H. Xu, and C. Tao. Ontology of consumer health vocabulary: providing a formal and interoperable semantic resource for linking lay language and medical terminology. In *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2019. doi:10.1109/bibm47256.2019.8983220.

[5] L. Anselin. Local indicators of spatial association—lisa. *Geographical Analysis*, 27 (2):93–115, 1995. doi:10.1111/j.1538-4632.1995.tb00338.x.

[6] M. Avanzini, M. D. Pinheiro, R. Gomes, and C. Rolim. Energy retrofit as an answer to public health costs of fuel poverty in lisbon social housing. *Energy Policy*, 160: 112658, 2022. doi:10.1016/j.enpol.2021.112658.

[7] J. Bai, S. Mosbach, C. J. Taylor, D. Karan, K. F. Lee, S. D. Rihm, J. Akroyd, A. A. Lapkin, and M. Kraft. A dynamic knowledge graph approach to distributed self-driving laboratories. *Nature Communications*, 15(1), Jan. 2024. doi:10.1038/s41467-023-44599-9.

[8] C. Baker. Constituency data: How healthy is your area?, 7 2024. URL https://commonslibrary.parliament.uk/constituency-data-how-healthy-is-your-area/. Accessed Aug 2024.

[9] K. J. Baker, R. Mould, and S. Restrick. Rethink fuel poverty as a complex problem. *Nature Energy*, 3(8):610–612, 2018. doi:10.1038/s41560-018-0204-2.

[10] R. Battle and D. Kolas. Enabling the geospatial semantic web with parliament and geosparql. *Semantic Web*, 3(4):355–370, 2012. doi:10.3233/sw-2012-0065.

[11] K. Bereta, G. Xiao, and M. Koubarakis. Ontop-spatial: Ontop of geospatial databases. *Journal of Web Semantics*, 58:100514, 2019. doi:10.1016/j.websem.2019.100514.

[12] D. Bienvenido-Huertas, D. Sánchez-García, and C. Rubio-Bellido. Analysing natural ventilation to reduce the cooling energy consumption and the fuel poverty of social dwellings in coastal zones. *Applied Energy*, 279:115845, 2020. doi:10.1016/j.apenergy.2020.115845.

[13] P. Bridgen and C. Robinson. A decade of fuel poverty in england: A spatio-temporal analysis of needs-based targeting of domestic energy efficiency obligations. *Energy Research & Social Science*, 101:103139, 2023. doi:10.1016/j.erss.2023.103139.

[14] L. Carmichael, E. Prestwood, R. Marsh, J. Ige, B. Williams, P. Pilkington, E. Eaton, and A. Michalec. Healthy buildings for a healthy city: Is the public health evidence base informing current building policies? *Science of The Total Environment*, 719:137146, 2020. doi:10.1016/j.scitotenv.2020.137146.

[15] P. Chen. Effects of the entropy weight on topsis. *Expert Systems with Applications*, 168:114186, 2021. doi:10.1016/j.eswa.2020.114186.

[16] T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, volume 11 of *KDD '16*, page 785–794. ACM, 2016. doi:10.1145/2939672.2939785.

[17] T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, page 785–794, 2016. doi:10.1145/2939672.2939785.

[18] J. A. Dauda and S. O. Ajayi. Understanding the impediments to sustainable structural retrofit of UK's school building stock: A systematic review. *Journal of Building Engineering*, 60:105168, 2022. doi:10.1016/j.jobe.2022.105168.

[19] D. Densley Tingley. Embed circular economy thinking into building retrofit. *Communications Engineering*, 1(1):28, 2022. doi:10.1038/s44172-022-00027-2.

[20] Department for Energy Security & Net Zero. Energy company obligation (eco) - great british insulation scheme (2023–2026) & amendments to eco4 regulations. Technical report, Department for Energy Security & Net Zero, London, March 2023. URL https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1152132/gb-insulation-scheme-government-response.pdf.

[21] Department for Energy Security and Net Zero. Household energy efficiency, 8 2024. URL https://www.data.gov.uk/dataset/1656fb7d-1ca3-462d-a11b-8078acc33275/household-energy-efficiency. Accessed Aug 2024.

[22] Department for Energy Security and Net Zero. Fuel poverty in england: sub-regional, 4 2024. URL https://www.data.gov.uk/dataset/f3009590-2bc9-40d9-8dc3-571e6fddae45/fuel-poverty-in-england-sub-regional. Accessed Aug 2024.

[23] Department for Energy Security and Net Zero. Lower and middle super output areas electricity consumption, 1 2024. URL https://www.gov.uk/government/statistics/lower-and-middle-super-output-areas-electricity-consumption. Accessed Aug 2024.

[24] Department for Energy Security and Net Zero. Sub-national gas consumption data, 3 2024. URL https://www.gov.uk/government/collections/sub-national-gas-consumption-data. Accessed Aug 2024.

[25] Department for Levelling Up, Housing & Communities. Energy performance of buildings data: England and wales, 7 2024. URL https://epc.opendatacommunities.org. Accessed July 2024.

[26] L. Desvallées. Low-carbon retrofits in social housing: Energy efficiency, multidimensional energy poverty, and domestic comfort strategies in southern europe. *Energy Research & Social Science*, 85:102413, 2022. doi:10.1016/j.erss.2021.102413.

[27] A. Donkin and M. Marmot. Left out in the cold: The hidden health costs of cold homes. Technical report, UCL Institute of Health Equity, 2024. URL https://www.instituteofhealthequity.org/resources-reports/left-out-in-the-cold-the-hidden-impact-of-cold-homes/copy-of-read-the-report.pdf. Accessed Aug 2024.

[28] EDINA. Digimap discovery service, 2024. URL https://digimap.edina.ac.uk/roam/map/os. Accessed Aug 2024.

[29] A. Galimshina, M. Moustapha, A. Hollberg, S. Lasvaux, B. Sudret, and G. Habert. Strategies for robust renovation of residential buildings in Switzerland. *Nature Communications*, 15(1), 2024. doi:10.1038/s41467-024-46305-9.

[30] A. Gasparrini, P. Masselot, M. Scortichini, R. Schneider, M. N. Mistry, F. Sera, H. L. Macintyre, R. Phalkey, and A. M. Vicedo-Cabrera. Small-area assessment of temperature-related mortality risks in england and wales: a case time series analysis. *The Lancet Planetary Health*, 6(7):e557–e564, 2022. doi:10.1016/s2542-5196(22)00138-3.

[31] A. Getis and J. K. Ord. The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24(3):189–206, 1992. doi:10.1111/j.1538-4632.1992.tb00261.x.

[32] R. Gillard, C. Snell, and M. Bevan. Advancing an energy justice perspective of fuel poverty: Household vulnerability and domestic retrofit policy in the United Kingdom. *Energy Research & Social Science*, 49:7–15, 2019. doi:10.1016/j.erss.2017.05.012.

[33] A. Katris and K. Turner. Can different approaches to funding household energy efficiency deliver on economic and social policy objectives? ECO and alternatives in the uk. *Energy Policy*, 155:112375, 2021. doi:10.1016/j.enpol.2021.112375.

[34] F. Kern, P. Kivimaa, and M. Martiskainen. Policy packaging or policy patching? the development of complex energy efficiency policy mixes. *Energy Research & Social Science*, 23:11–25, 2017. doi:10.1016/j.erss.2016.11.002.

[35] N. Kerr, A. Gouldson, and J. Barrett. The rationale for energy efficiency policy: Assessing the recognition of the multiple benefits of energy efficiency retrofit policy. *Energy Policy*, 106:212–221, 2017. doi:10.1016/j.enpol.2017.03.053.

[36] H. Khraishah, B. Alahmad, R. L. Ostergard, A. AlAshqar, M. Albaghdadi, N. Vellanki, M. M. Chowdhury, S. G. Al-Kindi, A. Zanobetti, A. Gasparrini, and S. Rajagopalan. Climate change and cardiovascular disease: implications for global

health. *Nature Reviews Cardiology*, 19(12):798–812, 2022. doi:10.1038/s41569-022-00720-x.

[37] K. Lazo Green, M. M. C. Tan, E. E. Johnson, N. Ahmed, C. Eastaugh, F. Beyer, D. Craig, G. F. Spiers, and B. Hanratty. Interventions for cold homes: a rapid review of the health impacts. *European Journal of Public Health*, 34(4):682–695, 2024. doi:10.1093/eurpub/ckae058.

[38] Y. Liang, Y. Pan, X. Yuan, Y. Yang, L. Fu, J. Li, T. Sun, Z. Huang, and R. Kosonen. Assessment of operational carbon emission reduction of energy conservation measures for commercial buildings: Model development. *Energy and Buildings*, 268: 112189, 2022. doi:10.1016/j.enbuild.2022.112189.

[39] C. Liddell, C. Morris, S. McKenzie, and G. Rae. Measuring and monitoring fuel poverty in the uk: National and regional perspectives. *Energy Policy*, 49:27–32, 2012. doi:10.1016/j.enpol.2012.02.029.

[40] M. Q. Lim, X. Wang, O. Inderwildi, and M. Kraft. *The World Avatar—A World Model for Facilitating Interoperability*, page 39–53. Springer International Publishing, 2022. doi:10.1007/978-3-030-86215-2_4.

[41] A. Luers, L. Yona, C. B. Field, R. B. Jackson, K. J. Mach, B. W. Cashore, C. Elliott, L. Gifford, C. Honigsberg, L. Klaassen, H. D. Matthews, A. Peng, C. Stoll, M. Van Pelt, R. A. Virginia, and L. Joppa. Make greenhouse-gas accounting reliable — build interoperable systems. *Nature*, 607(7920):653–656, 2022. doi:10.1038/d41586-022-02033-y.

[42] Ministry of Housing, Communities and Local Government (UK). A guide to Energy Performance Certificates for the marketing, sale and let of dwellings, 2017. URL https://www.gov.uk/government/publications/energy-performance-certificates-for-the-construction-sale-and-let-of-dwellings/a-guide-to-energy-performance-certificates-for-the-marketing-sale-and-let-of-dwellings. Accessed Oct 2024.

[43] C. Nikolaou, E. V. Kostylev, G. Konstantinidis, M. Kaminski, B. Cuenca Grau, and I. Horrocks. Foundations of ontology-based data access under bag semantics. *Artificial Intelligence*, 274:91–132, 2019. doi:10.1016/j.artint.2019.02.003.

[44] Office for National Statistics. Lower Layer Super Output Area (LSOA) Boundaries, 2016. URL https://www.data.gov.uk/dataset/fa883558-22fb-4a1a-8529-cffdee47d500/lower-layer-super-output-area-lsoa-boundaries. Accessed June 2024.

[45] Office for National Statistics. How fuel poverty is measured in the UK: March 2023, 2023. URL https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/howfuelpovertyismeasuredintheuk/march2023. Accessed Oct 2024.

[46] Office for National Statistics. Ons statistical entity ontology, 2023. URL https://statistics.data.gov.uk/resource?uri=http%3A%2F%2Fstatistics.data.gov.uk%2Fdef%2Fstatistical-entity. Accessed Aug 2024.

[47] S. Opricovic and G.-H. Tzeng. Compromise solution by mcdm methods: A comparative analysis of vikor and topsis. *European Journal of Operational Research*, 156 (2):445–455, 2004. doi:10.1016/s0377-2217(03)00020-1.

[48] O. Pasichnyi, F. Levihn, H. Shahrokni, J. Wallin, and O. Kordas. Data-driven strategic planning of building energy retrofitting: The case of stockholm. *Journal of Cleaner Production*, 233:546–560, 2019. doi:10.1016/j.jclepro.2019.05.373.

[49] H. Pearsall, V. H. Gutierrez-Velez, M. R. Gilbert, S. Hoque, H. Eakin, E. S. Brondizio, W. Solecki, L. Toran, J. E. Baka, J. E. Behm, C. Brelsford, C. Hinrichs, K. A. Henry, J. Mennis, L. A. Roman, C. Rosan, E. C. South, and R. D. Valletta. Advancing equitable health and well-being across urban–rural sustainable infrastructure systems. *npj Urban Sustainability*, 1(1), 2021. doi:10.1038/s42949-021-00028-8.

[50] K. Qu, X. Chen, A. Ekambaram, Y. Cui, G. Gan, A. Økland, and S. Riffat. A novel holistic epc related retrofit approach for residential apartment building renovation in norway. *Sustainable Cities and Society*, 54:101975, 2020. doi:10.1016/j.scs.2019.101975.

[51] F. Re Cecconi, A. Khodabakhshian, and L. Rampini. Data-driven decision support system for building stocks energy retrofit policy. *Journal of Building Engineering*, 54:104633, 2022. doi:10.1016/j.jobe.2022.104633.

[52] M. A. T. Reaños, J. Curtis, D. Meier, and A. Pillai. Looking beyond energy efficiency and the role of financial literacy in experiencing fuel poverty: Evidence from irish homeowners. *Energy Research & Social Science*, 112:103515, 2024. doi:10.1016/j.erss.2024.103515.

[53] J. Rosenow, R. Platt, and B. Flanagan. Fuel poverty and energy efficiency obligations – a critical assessment of the supplier obligation in the uk. *Energy Policy*, 62: 1194–1203, 2013. doi:10.1016/j.enpol.2013.07.103.

[54] T. Savage, J. Akroyd, S. Mosbach, M. Hillman, F. Sielker, and M. Kraft. Universal digital twin – the impact of heat pumps on social inequality. *Advances in Applied Energy*, 5:100079, 2022. doi:10.1016/j.adapen.2021.100079.

[55] A. Schultheis, C. Zeyen, and R. Bergmann. *An Overview and Comparison of Case-Based Reasoning Frameworks*, page 327–343. Springer Nature Switzerland, 2023. doi:10.1007/978-3-031-40177-0_21.

[56] M. E. Scyphers, J. E. Missik, H. Kujawa, J. A. Paulson, and G. Bohrer. Bayesian optimization for anything (boa): An open-source framework for accessible, user-friendly bayesian optimization. *Environmental Modelling & Software*, 182:106191, 2024. doi:10.1016/j.envsoft.2024.106191.

[57] H.-S. Shih. Incremental analysis for mcdm with an application to group topsis. *European Journal of Operational Research*, 186(2):720–734, Apr. 2008. doi:10.1016/j.ejor.2007.02.012.

[58] P. Sławomir, S. Anweiler, M. Tańczuk, I. Klementowski, D. Przysiężniuk, Ł. Adrian, G. McNamara, and Ž. Stevanović. Energy source impact on the economic and environmental effects of retrofitting a heritage building with a heat pump system. *Energy*, 278:128037, 2023. doi:10.1016/j.energy.2023.128037.

[59] United Nations Environment Programme, Yale Center for Ecosystems + Architecture. Building materials and climate: Constructing a new future, 2023. URL https://wedocs.unep.org/20.500.11822/43293. Accessed Aug 2024.

[60] G. Viegi, F. Pistelli, D. L. Sherrill, S. Maio, S. Baldacci, and L. Carrozzi. Definition, epidemiology and natural history of copd. *European Respiratory Journal*, 30(5): 993–1013, 2007. doi:10.1183/09031936.00082507.

[61] A. Vilches, A. Barrios Padura, and M. Molina Huelva. Retrofitting of homes for people in fuel poverty: Approach based on household thermal comfort. *Energy Policy*, 100:283–291, 2017. doi:10.1016/j.enpol.2016.10.016.

[62] L. Wang, J.-j. Peng, and J.-q. Wang. A multi-criteria decision-making framework for risk ranking of energy performance contracting project under picture fuzzy environment. *Journal of Cleaner Production*, 191:105–118, 2018. doi:10.1016/j.jclepro.2018.04.169.

[63] G. Xiao, D. Calvanese, R. Kontchakov, D. Lembo, A. Poggi, R. Rosati, and M. Zakharyaschev. Ontology-based data access: A survey. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, IJCAI-2018. International Joint Conferences on Artificial Intelligence Organization, 2018. doi:10.24963/ijcai.2018/777.

[64] X. Zhong, M. Hu, S. Deetman, B. Steubing, H. X. Lin, G. A. Hernandez, C. Harpprecht, C. Zhang, A. Tukker, and P. Behrens. Global greenhouse gas emissions from residential and commercial building materials and mitigation strategies to 2060. *Nature Communications*, 12(1):6126, 2021. doi:10.1038/s41467-021-26212-z.