University of Tartu
Faculty of Science and Technology
Institute of Ecology and Earth Sciences
Department of Geography

Master’s thesis in Geoinformatics for Urbanised Society (30 ECTS)

Patterns of inequalities in housing energy efficiency and links with population risk factors in Tallinn, Riga and Vilnius

Kirils Gončarovs

Supervisor: Associate Professor Anneli Kährik

Tartu, 2025

Abstract

Patterns of inequalities in housing energy efficiency and links with population risk factors in Tallinn, Riga and Vilnius

The introduction of the EU Emission Trading System in the residential building sector poses a risk of increased energy poverty, particularly during the energy crisis. Energy poverty disproportionally affects different population groups due to the vulnerabilities to price changes and underlying inequalities of access to energy-efficient housing. Thus, monitoring the inequalities associated with housing energy efficiency is important for facilitating just energy transition. To evaluate the current patterns of inequalities in access to energy-efficient housing, the study develops a machine learning-based modelling framework to mitigate the limitations of the existing data availability on building-level energy performance. The study applied the developed methodology to assess the building energy performance in three cities – Tallinn, Riga, and Vilnius – to further identify the inequalities associated with access to energy-efficient housing between the different population groups. The study identified that within the context of the three capital cities, the existing inequalities related to access to energy-efficient housing can be understood through the prism of the spatial unequal distribution of occupational groups, particularly in relation to low occupational groups, and the housing market segmentation.

Keywords: Energy efficiency, housing renovation, residential inequalities

CERCS code: S230 Social Geography

Annotatsioon

Eluasemete energiatõhususe ebavõrdsuse mustrid ja seosed elanikkonna riskiteguritega Tallinnas, Riias ja Vilniuses

ELi heitkogustega kauplemise süsteemi kasutuselevõtt elamumajanduse sektoris kujutab endast energiavaesuse suurenemise ohtu, eriti energiakriisi ajal. Energiavaesus mõjutab erinevaid elanikkonnarühmi ebaproportsionaalselt, kuna nad on tundlikud hinnamuutuste suhtes ja neil on ebavõrdne juurdepääs energiatõhusale eluasemele. Seega on eluasemete energiatõhususega seotud ebavõrdsuse jälgimine oluline õiglase energia ülemineku hõlbustamiseks. Ebavõrdse energiatõhusale eluasemele juurdepääsu praeguste mustrite hindamiseks töötatakse käesolevas uurimistöös välja masinõppel põhinev modelleerimisraamistik, et leevendada olemasolevate hoonete energiatõhususe andmete kättesaadavust. Töös kasutati hoonete energiatõhususe hindamiseks väljatöötatud metoodikat kolmes linnas: Tallinnas, Riias ja Vilniuses, et teha kindlaks eri elanikkonnarühmade vahelised ebavõrdsused seoses juurdepääsuga energiatõhusale eluasemele. Töös leiti, et kolme pealinna andmete põhjal olemasolev ebavõrdsus seoses energiatõhusate eluruumide kättesaadavusega on seotud ametirühmade, eelkõige madalate ametirühmade ruumilise ebavõrdsuse ja eluasemeturu segmenteerituse kaudu.

Märksõnad: Energiatõhusus, eluhoonete renoveerimine, elukoha ebavõrdus

CERCS valdkond: S230 Sotsiaalne geograafia

Table of Contents

1. Introduction

In *Energy For A Sustainable World* (1987), energy is identified as one of the most important issues intertwined and connected with other global problems (Goldemberg *et al.*, 1987, p. 35). As buildings represent 40% of the final energy use in the EU (European Parliament, 2012), the sector represents an important challenge in the facilitation of just energy transition.

The three main tenets within the energy justice framework are distributional, procedural, and recognition justice (McCauley *et al.*, 2014). Particularly within distributional justice falls the question of energy poverty (Jenkins *et al.*, 2016), which describes the lack of accessible essential energy services (Reddy *et al.*, 2000, p. 44). The spatiality of energy poverty, however, has not been sufficiently examined (Bouzarovski and Simcock, 2017). Furthermore, the existing inequalities in income (Bouzarovski and Simcock 2017), gender (Petrova and Simcock, 2021), levels of urbanisation (Aristondo and Onaindia, 2018), ownership status (Aristondo and Onaindia, 2018), race and ethnicity (Dogan *et al.*, 2022), and household composition (Boemi and Papadopoulos, 2019) further exacerbate energy poverty and inequalities related to energy consumption and production.

The topics of inequalities in energy distribution and energy poverty are significantly understudied in the contexts of Estonia, Latvia, and Lithuania. This is particularly important for the policy context of Estonia, as energy poverty is not mentioned in the National Energy and Climate Plan as a significant issue (Directorate of European Commission, 2023), even though in 2020, 22.6% of the population experienced a high share of energy expenditure in their income (Energy Poverty Advisory Hub, 2022). And even as in the case of Latvia and Lithuania, where energy poverty is considered an important energy issue in the existing planning documents, the measures of its alleviation are insufficient in addressing it. As housing represents the primary area of interactions with energy poverty, energy inequalities in the study are addressed through the inequalities of access to energy-efficient housing.

Energy performance certificates represent a unified framework for measuring the energy efficiency of a building, developed by the European Parliament and implemented at the national level (European Parliament, 2003). However, the existing data from energy performance certificates are not available for the substantial segment of the building stock in Estonia, Latvia, and Lithuania. Different methodological frameworks were developed to model building energy performance using machine-learning-based techniques (Seyedzadeh *et al.*, 2018), yet not applied for the urban-level energy performance predictions in the context of the Baltic counties. Thus, methodologically, the work aims to develop a machine-learning-based methodology for urban-level building energy performance prediction using various open data sources based on energy performance certificate data. Furthermore, the study aims to identify the existing demographic, ethnic, and occupational inequalities related to the energy efficiency of housing based on the predicted and actual data for Tallinn, Vilnius, and Riga.

The work defines the following research questions:

1. Which algorithm is the most suitable for modelling the energy performance of buildings based on the existing data availability and limitations? How does the application of the developed machine-learning-based modelling framework compare between the cities?
2. What distributional patterns of energy-efficient housing can be observed between the population groups in the three capital cities within the given housing and urban contexts? Are there only cases of identified inequalities in particular country contexts, or are the patterns shared between all three capital cities?
3. How are the specific identified distributional patterns addressed in the existing national policies of energy efficiency and energy poverty? If not sufficiently addressed, what types of policy actions can be used to mitigate the identified trends of inequalities in the energy efficiency of housing?

2. Literature review

2.1. Energy poverty

Six primary energy vulnerability factors contribute to energy poverty: access, affordability, flexibility, energy efficiency, household needs, and existing practices (Bouzarovski and Petrova, 2015). All the factors are encompassed in its definition. As an overarching concept, energy poverty has been defined as ‘the absence of sufficient choice in accessing adequate, affordable, reliable, high-quality, safe, and environmentally benign energy services to support economic and human development’ (Reddy et al. 2000, 44 p). The shift of primary focus changes from the challenges of energy accessibility in the developing world to more prominent affordability issues in the developed world (Reddy et al., 2000).

Within the context of the European Union (EU), energy poverty is described in the Energy Efficiency Directive as ‘a household’s lack of access to essential energy services, where such services provide basic levels and decent standards of living and health, including adequate heating, hot water, cooling, lighting, and energy to power appliances… caused by a combination of factors, including at least non-affordability, insufficient disposable income, high energy expenditure and poor energy efficiency of homes’ (European Parliament, 2023a). Even though both definitions are focused on the absence of adequate energy services, the legal definition in the European Union is centred around the affordability of energy services at the household level.

Some scholars propose applying the energy poverty definition to include only low-income households (Mulder et al., 2023). However, other studies point out that middle-income households also experience a relatively high incidence of energy poverty, with a significant variation in incidence values between the indicators used (Maier and Dreoni, 2024). This reveals the overall heterogeneity in energy poverty measurement approaches, requiring a more in-depth assessment of the indicators used.

2.1.1. Energy poverty indicators

The multi-dimensionality of energy poverty results in three indicators based on the method: expenditure, consensual, and direct measurement approaches (Thomson et al., 2017). Expenditure approaches consider criteria such as income, housing, or energy costs (Rodriguez-Alvarez et al., 2021). Consensual approaches evaluate the self-reported perception of energy poverty by indicators such as indoor housing conditions (Thomson et al., 2017). Direct measurement approaches are characterised by evaluating the actual energy use of energy services (Kahouli and Okushima, 2021). The following section will address the examples of three indicator categories. Even though each indicator category is analysed separately, it is important to note that the concentration on a single indicator of energy poverty results in exclusion, as many indicators do not significantly overlap between the households identified as energy-poor (Deller et al., 2021). Furthermore, focusing on one energy poverty indicator over another can result in significantly different budgetary and distributional implications in targeted policy actions (Maier and Dreoni, 2024), as the different indicators can highlight different population and housing segments (Barrett et al., 2022).

Expenditure approaches for energy poverty measurement

As mentioned above, expenditure energy poverty indicators compare housing and energy costs to income (Oliveira Panão, 2021; Rodriguez-Alvarez et al., 2021). According to Liddell et al. (2012), the first definitions of energy poverty primarily examined energy expenditure rates. This included the 10% indicator created by Brenda Boardman in the 1991 book 'Fuel Poverty' (Liddell et al., 2012), which takes the double median of weekly expenses on energy based on the 1988 Family Expenditure Survey for UK households (Liddell et al., 2012; Herrero, 2017). The 10% indicator is often criticised for not properly reflecting energy poverty rates (Yip et al., 2020), especially because of its high sensitivity to energy price fluctuations (Hills, 2012, p. 30). The use of the double median approach, however, mitigates the issues related to the use of the 10% indicator as the indicator becomes relative to the current situation in the energy market (Croon *et al.*, 2023).

In 2012, a new indicator for energy poverty evaluation – Low Income and High Costs (LIHC) – was proposed to mitigate the challenges created using the 10% indicator (Hills, 2012, p. 32). The main difference between the indicators lies in the existence of two energy poverty gaps: energy cost and income levels (Hills, 2012, p. 34). To count a household energy-poor, their energy expenditure must be higher than the national median, and their total available income without energy costs must be lower than 60% of the national median (Costa-Campi et al., 2019). Even though it is based on the context of the UK, the use of LIHC indicator has expanded to study energy poverty in different countries and regions (Costa-Campi et al., 2019; Oliveira Panão, 2021; Dogan et al., 2022; Kalinowski et al., 2024). The next iteration of the indicators applied in evaluating energy poverty rates in the UK focused on the energy efficiency of housing instead of energy costs (Croon et al., 2023).

Consensual approaches for energy poverty measurement

Consensual approaches to energy poverty measurement focus on asking households whether the energy is affordable and whether housing conditions are benign (Thomson et al., 2017). Furthermore, consensual approaches are used more for cross-country energy poverty assessments and comparisons (Thomson et al., 2017), particularly by utilising data from the EU statistics on income and living conditions (EU-SILC) as the basis for comparative studies in Europe (Makridou et al., 2024; Śmiech et al., 2025). The three primary indicators for energy poverty in EU-SILC data include the 'inability to keep home adequately warm', 'arrears on utility bills', and 'presence of leak, damp, rot in the dwelling' (Halkos and Kostakis, 2023). For instance, the indicator of 'inability to keep home adequately warm' was used by Makridou et al. (2024) in a pan-European longitudinal study on factors associated with energy poverty, resulting in the identification of positive correlations between energy poverty and levels of urbanisation, overcrowding, electricity prices, and gross domestic product value on the national level.

Other studies employ own surveys and energy poverty indicators, adapting to the local context (Ntaintasis et al., 2019). The survey developed by Ntaintasis et al. (2019) for the Attica region in Greece utilised additional perceptual energy poverty indicators such as ‘inability to keep home adequately cool during summer’, ‘health problems that can be attributed to poor indoor conditions’, and ‘restriction of other essential needs of the household due to the expenditures required for adequately heating the residence’. Another study in Nova Scotia, Canada, used additional indicators, such as the ability to control indoor temperature during winter and summer and having difficulty sleeping because the dwelling was too cold, to evaluate energy poverty and calculate expenditure-based indicators (Ntaintasis et al., 2019; Riva et al., 2024).

Direct measurement approaches for energy poverty measurement

Direct measurement approaches attempt to measure the sufficiency of energy services through indoor building conditions such as indoor temperature measurements (Thomson et al., 2017). Yun et al. (2024) propose an energy poverty monitoring system, which includes sensor nodes for dry-bulb temperature, relative humidity, air velocity, and black bulb temperature to monitor the thermal indoor environment for energy-poor households in South Korea. A similar system with a continuous real-life measurement system was developed and tested in Getafe, Spain (López-Vargas and Ledezma-Espino, 2023). The justification for applying frameworks measuring energy poverty through thermal comfort and, subsequently, indoor temperature stems from the existing guidance of the World Health Organisation on the home temperature range (Ormandy and Ronique Ezratty, 2011).

Thomson et al. (2017) note the strong lack of direct measurement data at the European level and its national-level scarcity. Furthermore, within the context of Central and Eastern Europe, the primary pillar of energy poverty is associated with energy costs rather than indoor temperatures due to the use of district heating (Tirado Herrero and Ürge-Vorsatz, 2012). Additionally, the notion of thermal comfort has significant variation between the different countries and households (Nicol et al., 1999; Nicol and Roaf, 2017), increasing the overall challenge of evaluation of energy poverty through indoor temperatures.

2.1.2. Hidden energy poverty

In the conventional indicators related to energy poverty, an often-overlooked issue is the limitation of household energy consumption to reduce costs (Cong et al., 2022). The practice of restricting one's energy consumption below the level of basic household needs is called ‘hidden energy poverty’ (Meyer et al., 2018). Crucially, many households practising restrictive behaviour are usually not considered energy-poor by expenditure-related metrics and, thus, not eligible for some of the energy poverty alleviation measures (Cong et al., 2022; Eisfeld and Seebauer, 2022). Within the context of Central and Eastern Europe, however, a direct link between income poverty and exposure to hidden energy poverty was identified (Karpinska and Śmiech, 2020b). A variety of different metrics are used to measure hidden energy poverty, which are summarised in Table 1.

Table 1. Methodologies and indicators used to evaluate hidden energy poverty.

Country	Approach	Indicators	Study
Australia	Consensual	Underconsumption entailing privation Incidental masking Disguise by coping mechanisms Intentional concealment Failure to recognise the health risks of cold homes Ignored energy vulnerability	Willand et al. (2023)
Austria	Consensual	Heating up to a comfortable temperature without paying attention to the costs. Sitting close to the radiator to keep warm. Putting on a pullover first instead of turning on the heating. Turning off the heating when leaving the flat. Closing doors between heated and not heated rooms.	Eisfeld and Seebauer (2022)
Italy	Expenditure	Combination of energy expenditure and poverty rates	Betto et al. (2020)
Poland	Expenditure	Multiple linear regression model with categorical values for the living environment, housing conditions, household structure, and socio-economic situation.	Karpinska and Śmiech (2020a)
Spain	Expenditure	Required energy expenditure Low absolute energy expenditure indicator	Barrella et al. (2022)
USA	Direct and consensual	Energy equity gap (based on the inflection temperature)	Cong et al. (2022)

2.1.3. Characteristics of energy poverty

As stated by Bouzarovski and Simcock (2017), domestic energy deprivation is strongly impacted by household incomes. In Ireland, income poverty was strongly associated with measured energy poverty indicators (Barrett et al., 2022). Furthermore, a strong link between regional income inequality and energy poverty was observed in Italy, which further drove variation in the experience of energy poverty based on household income levels (Bardazzi et al., 2021). A similar outcome was identified in the global energy poverty study, which identified national income inequality as an aggravating factor for energy poverty within all population groups (Igawa and Managi, 2022). In the case of Poland, the experience of energy poverty was linked to increased chances of poverty, with a 15.1% chance of becoming poor after experiencing energy poverty and an 11.7% chance of becoming severely poor (Karpinska and Śmiech, 2021). A similar outcome was also identified in Germany (Drescher and Janzen, 2021). The different patterns of energy vulnerability are further replicated through the adaptive behaviour to energy poverty, facilitating the vicious circle of energy vulnerability (Bouzarovski and Simcock, 2017).

The urbanisation rate has a negative impact on energy poverty rates. (Roberts et al., 2015; Bouzarovski and Tirado Herrero, 2017; Aristondo and Onaindia, 2018; Drescher and Janzen, 2021; Karpinska and Śmiech, 2021) The study of the distribution of energy-poor households in Poland, Czechia, and Hungary found that the highest proportions of energy-poor households were in small to medium-sized towns due to the lower accessibility of building renovation programmes yet facing the same legal and technical issues of the building stock (Bouzarovski and Tirado Herrero, 2017). In another study in Poland, one of the key energy-poor profiles was of medium-sized families living in detached housing in rural areas (Karpinska and Śmiech, 2021). A similar pattern was observed in Spain, where lower population density areas have a higher proportion of energy-poor populations (Aristondo and Onaindia, 2018). In the study by Drescher & Janzen, 2021, rural households' increased chance of energy poverty was further attributed to the differences in grid access fees compared to their urban counterparts. Additionally, the lack of diverse heating and fuel supplier options significantly increases energy prices' effect on the rural population's energy poverty rates (Roberts et al., 2015).

Women take a disproportionate amount of responsibility in managing the consequences of energy poverty, which results in worse well-being and the need for additional labour (Petrova and Simcock, 2021). Furthermore, the evidence from Spain shows that women-led households have a higher chance of being energy-poor (Aristondo and Onaindia, 2018). Robinson (2019) notes five dimensions of gendered vulnerability to energy poverty, which include exclusion from a productive economy, susceptibility to negative mental health outcomes, a lack of social protection, coping and helping others to cope, and unpaid caring or domestic roles. Furthermore, these dimensions significantly overlap with experiences of income poverty (Robinson, 2019). Overall, the assessment of Nguyen and Su (2021) in 51 developing countries concluded that reducing energy poverty could significantly improve women's socio-economic rights. However, as Listo

(2018) states, the discourse on gender and energy poverty continues to construct women and gender equality in a problematic way, particularly in Global South, and ignores queer or transgender identities or gendered inequalities between women.

In Central and Eastern Europe, one-person households dominate in assessing energy poverty profile distribution (Karpinska and Śmiech, 2023), as well as in other countries such as Greece (Boemi and Papadopoulos, 2019), Germany (Drescher and Janzen, 2021), and Canada (Riva et al., 2021). Furthermore, other research has noted that single-parent households (Drescher and Janzen, 2021; Mulder *et al*., 2023) or single pensioner households (Stojilovska et al., 2021) were even more prone to energy poverty. In the study of energy poverty in one-person households in Poland, the income level was the most significant factor in assessing energy poverty (Piekut, 2020). Furthermore, the thermal comforts of a one-person household vary greatly based on the occupational status and level of education (Piekut, 2020). Drescher and Janzen (2021) state that higher energy poverty rates in one-person households could be attributed to lower household income possibilities, lack of cost sharing, and economic scale in domestic energy services.

Tenants are more prone to energy poverty than homeowners (Aristondo and Onaindia, 2018; Taltavull de La Paz et al., 2022). In Spain, the risk of being energy-poor is almost twice as high as for homeowners (Taltavull de La Paz et al., 2022). Similarly, in Australia, the neighbourhoods with the higher share of rental properties show higher risks of energy stress (Willand et al., 2020). Furthermore, in France, the chance to live in an energy-inefficient dwelling was significantly higher for tenants than for owners occupying their own dwellings, and their participation in building renovation projects was lower (Charlier, 2014). The potential difference between the willingness to invest in energy efficiency improvements lies in the 'landlord-tenant problem' (Ambrose, 2015; Petrov and Ryan, 2021), which is further characterised by split incentives between the two groups (International Energy Agency, 2007). In building renovation and energy efficiency, split incentives between the landlords and tenants are characterised by the willingness of landlords to minimise capital costs and the willingness of tenants to maximise energy efficiency to save on energy costs (International Energy Agency, 2007). Furthermore, in Ireland, a larger difference in energy efficiency levels between rental and non-rental properties was observed in areas with a bigger scarcity of rental properties compared to the rest of the country (Petrov and Ryan, 2021).

In the context of the USA, a link between race and energy poverty was identified, with African-American households having a higher chance of being energy-poor (Dogan et al., 2022). Additionally, indigenous populations in Mexico experience higher energy poverty rates compared to non-indigenous households (Guzmán-Rosas, 2022). A similar phenomenon was observed in Kansas City, Missouri, between the racial and ethnic minority-headed households, and increased energy use intensity, describing lower energy efficiency (Reames, 2016). Furthermore, in Australia, neighbourhood-level ethnic diversity was positively associated with energy poverty, and social inclusion policies were outlined as the potential mitigation strategy for energy poverty

(Awaworyi Churchill and Smyth, 2020). The study by Reames (2016) notes that the past institutionalised residential segregation is further translated into current energy-related disparities. The graphical summary of the identified characteristics associated with increased energy poverty is demonstrated in *Figure 1*.

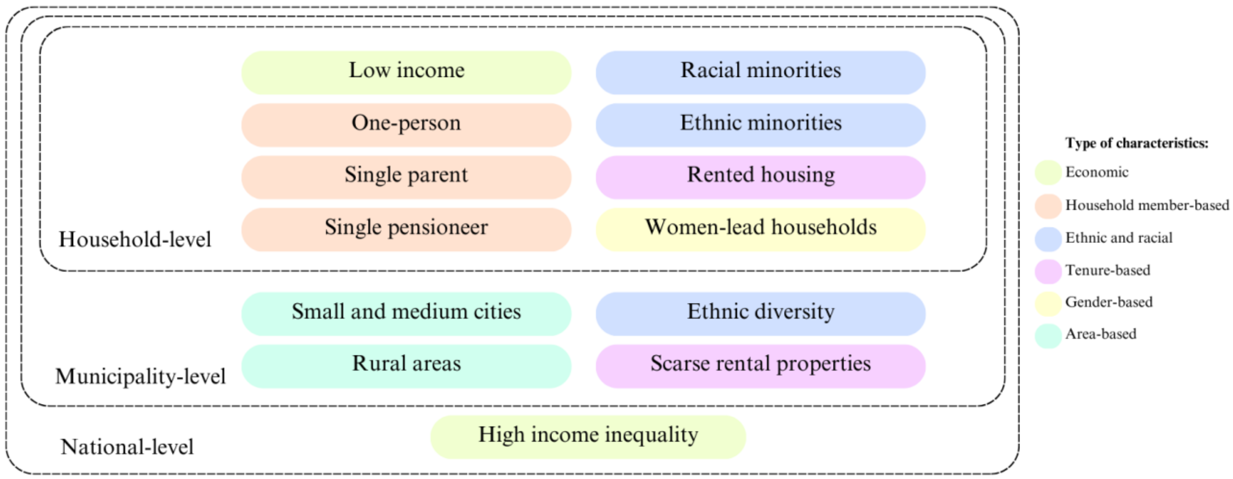

Figure 1. Characteristics associated with an increased chance of energy poverty based on the literature review.

2.1.4. Tackling energy poverty

In the energy policy of many European countries, energy poverty is not considered an issue in the energy field and is tackled exclusively through social policy linked to general income poverty (Bouzarovski et al., 2021). This further creates confusion about the primary responsible governmental body to oversee the issue, as observed in many European countries (Bouzarovski et al., 2021; Feenstra et al., 2021; Koďousková and Bořuta, 2022). The confusion reduces the overall potential of action as some of the key policy mitigation strategies for energy poverty include increasing energy efficiency and regulating energy prices, which have a significantly positive impact on reducing energy poverty (Rodriguez-Alvarez et al., 2021). Public bodies at different levels are responsible for actions related to energy poverty, which further increases the challenges of the implementation of an integrated policy approach (Seebauer et al., 2019).

The social policy approaches focus on general income redistribution and include the support of mitigation of energy costs (Stojilovska et al., 2023). Within the regional Italian context, the overall income redistribution policies targeting income inequality alleviation were suggested as a possible energy poverty alleviation method (Bardazzi et al., 2021). Furthermore, Stojilovska et al. (2023) note the evermore-increasing role of different strategies in the labour market, social protection, and health fields to mitigate the health and economic impacts of energy poverty in the world of severe climate change impact.

Energy policy approaches focus on technical support to mitigate energy poverty (Stojilovska et al., 2023). For instance, the study of Simionescu et al. (2023) confirms that, in the long run, the increase in renewable energy consumption reduces the arrears of utility bills and thus contributes to reducing energy poverty. However, the initial accessibility of such interventions is low for low-income households due to the initial investments (Simionescu et al., 2023), and other studies point out that renewable energy systems currently do not mitigate energy poverty (Makridou et al., 2024). Furthermore, the assessment of low-income households in Greece revealed that 46.7% of households could not afford any energy efficiency measures to facilitate minimal thermal comfort (Boemi and Papadopoulos, 2019).

Overall, most current energy policy approaches to tackling energy poverty are largely inaccessible to energy-poor households (Koďousková and Bořuta, 2022). Property prices rising due to the improvements in housing's energy efficiency can make it unaffordable for low-income residents and renters (Seebauer et al., 2019). The existing notions of energy poverty-related stigmatisations and political non-recognition significantly affect the reproduction of the distributive inequalities in energy affordability and further alienation from getting support (Bouzarovski and Simcock, 2017). This describes the overall emerging issues in the field of energy transition, as it creates several issues associated with distributive energy justice.

2.1.5. Emerging issues

In many instances, the facilitation of just energy transition was connected to increased inequalities. From the point of view of benefitting from the support mechanisms, wealthier households and regions received more financial support to improve energy efficiency and increase the capacities of renewable energy generation. This, in return, enables further facilitation of gentrification caused by investments into sustainable urban development and strengthens the inaccessibility of energy transition for underserved individuals and communities.

Who is the main beneficiary?

The facilitation of building renovation activities creates additional difficulties for low-income households, potentially reducing their overall participation in climate and energy transition (Grossmann, 2019; Umit et al., 2019; Albrecht and Hamels, 2021; Bardazzi et al., 2021). An assessment of 22 European countries by Umit et al. (2019) identified a strong link between income and investment in energy efficiency technologies. Wealthier individuals were more likely to invest than low-income individuals, who tend to save energy through less use of existing technologies (Umit et al., 2019). Additionally, the use of energy efficiency subsidies by wealthier households was observed in the context of Italy (Bardazzi et al., 2021). A similar notion of the reproduction of the existing income inequalities through the distribution of energy efficiency subsidies was observed in Australia (Willand et al., 2020). This further adds to the overall notion of building renovation's unaffordability, as described by the analysis of Albrecht and Hamels (2021), which found that approximately 50% of homeowners cannot afford renovation measures to achieve 2050 climate policy goals in the Flemish region of Belgium. As renovation activities are often economically net-negative for the residents, the overall affordability of housing, especially for low-income households, is threatened, requiring economising on other vital human activities (Grossmann, 2019).

Particularly in the context of Central and Eastern Europe, the allocation of building renovation grants followed the existing regional disparities (Turcu, 2017; Lihtmaa et al., 2018; Frantál and Dvořák, 2022). The analysis of Estonia's building renovation grant programme revealed a strong relationship between regional development indicators, particularly higher local real estate prices and grant distribution, exacerbating existing regional inequalities (Lihtmaa et al., 2018). Similarly, a significantly higher share of building renovation grants was allocated to Bucharest compared with the rest of Romania, even though most of the housing in-need of energy efficiency improvement is outside the capital city (Turcu, 2017). Frantál and Dvořák (2022) note a stark difference in the New Green Savings Programme use between the different districts in Czechia, with a positive correlation between the subsidy allocation and indicators of territorial stability and socio-cultural capital. Furthermore, the study noted the potential misuse of the programme's funding in more affluent areas for building and purchasing new housing (Frantál and Dvořák, 2022).

Building renovation in rental housing posits additional risks associated with housing affordability (Polanska et al., 2024). In an assessment by von Platten et al. (2021), the existing notions of the Renovation Wave, particularly regarding targeting the worst-performing buildings, were challenged. The authors argue that facilitating renovation activities in this housing segment can threaten the already depleted stock of affordable housing, as the two overlap significantly (von Platten *et al.*, 2021). Furthermore, the notion of renovictions – eviction of tenants to perform renovations – has been observed in the different contexts of residential energy efficiency improvements as a practice of subsequent rent increases (Grossmann and Huning, 2016; Woodhall-Melnik, et al., 2025). Polanska and Richard (2021) note the three primary practices of renovictions by housing companies in Sweden – silencing, surveilling, and dividing collective demands – aimed at nullifying collective resistance and neglecting tenants' participation in the renovation projects' decision-making. The existing notions further highlight the two-edge issues of renters concerning increased risk of displacement due to renovations and benefitting less from the existing energy efficiency improvements, resulting in higher energy consumption and costs, as highlighted by Willand et al. (2020).

Green gentrification

Sustainable development of urban environments has been linked to gentrification and displacement (Dooling, 2009; Checker, 2011). Quinton and Nesbitt (2024) state that even though many terms, such as green, environmental, ecological, climate, carbon, or resilience gentrification, exist, they, in practice, describe different types of interventions in the built environment associated with climate impact mitigation or adaptation. Rice et al. (2020) note that this form of gentrification is an outcome of urban carbon politics, which results in the rejection of classic forms of suburbanisation by middle- and upper-income residents in favour of urban areas with access to low-carbon infrastructure.

The case study of the Letnica district urban regeneration programme in Gdansk, Poland, shows that even by itself, potential energy savings from building renovation activities can become a financial accumulation strategy benefiting both owners and housing investors and developers driving gentrification in the area (Bouzarovski et al., 2018). Similarly, Gould and Lewis (2021) describe resilience in coastal areas as the financial accumulation strategy in the communities affected by hurricanes, where the level of wealth determines the possibilities of resilient building (re)construction. They further state that employing private solutions to a public problem makes resilience a privilege (Gould and Lewis, 2021). Blok (2020) compares green gentrification in Copenhagen, Denmark, and Surat, India, and further highlights the underlying local and global inequalities.

However, the notions of gentrification can also stem from the interactions between energy infrastructure, residents, and emerging other forms of energy users. In the Bates et al. (2024) study, the residents of the Holyoke, Massachusetts community raised concerns associated with renewable energy projects considering high-performance computer centre facility development in the area without direct benefits to them and risks associated with electricity cost increase. In this light, Libertson et al. (2021) describe energy gentrification as a form of gentrification where competition over land is replaced with competition over energy, including individuals, industries, and other energy-intensive businesses. In the context of ever-increasing reliance on data centres for hi-tech services, the development of digital infrastructure can further exacerbate residential inequalities and cause displacement (Baumann *et al.*, 2024).

2.2. Policy context

2.2.1. EU policy framework

Within the policy framework of the European Union (EU), the Energy Efficiency Directive (EED) and the Energy Performance of Buildings Directive (EPBD) are two cornerstones for actions in the field of building renovation and assessment of energy poverty. Additional instruments, such as the Social Climate Fund, were developed due to the ever-increasing need to enable the participation of vulnerable population groups in just energy transition, particularly in light of the EU Emission Trading System expansion to the residential sector (European Parliament, 2023b).

The EED is a key policy document establishing the overall energy efficiency framework in the EU. EED is the formalisation of the Presidency Conclusions of the Brussels European Council in

2007, which established the 20% reduction goal in greenhouse gas (GHG) emissions and energy consumption of the Union (Council of the European Union, 2007). The process resulted in the facilitation of EED in 2012. As a key implementation step of EED, the Member States were obligated to create and submit the National Energy Efficiency Action Plans to achieve a cumulative reduction in energy consumption (European Parliament, 2012). After the Paris Agreement on climate change within the 'Clean Energy for all Europeans' package', EED was revised to facilitate a binding energy efficiency target for 2030 and the requirement for the Member States to submit National Energy and Climate Plans for 2030 (European Parliament, 2018b). Within the EU Green Deal and REPowerEU plan, the EED text was revised as a reaction to the need for more rapid climate action and reduced fossil fuel dependency from Russia (European Parliament, 2023a). The newest revision of EED formalised the 'energy efficiency first' principle, which mandates taking energy efficiency as the first option in any decision (European Parliament, 2023a). The facilitation of EED goals in the residential building efficiency sector is supported via the EPBD.

As the overarching building energy efficiency framework facilitation document, EPBD emerged in 2002, laying the foundation for energy performance certificates (EPC) and minimum performance requirements for new and existing buildings. Crucially, Article 7(1) of EPBD required the availability of EPC to the end-users (European Parliament, 2003). The Presidency Conclusions of the Brussels European Council in 2007 facilitated the preparation of the recast of EPBD in 2010. Key additions to the building energy performance framework consisted of conceptualising nearly zero-energy buildings, identifying and providing financial incentives for building renovation, and reducing market barriers (European Parliament, 2010). In the EPBD amendment of 2018, long-term renovation strategies were established for the Member States in Article 2(1) (European Parliament, 2018a). Article 2(1) of 2018 EPBD recast formalised the requirements for energy poverty alleviation by the Member States via long-term renovation strategies (European Parliament, 2018a). To reach the objectives of the European Green Deal facilitated via the Renovation Wave strategy, the 'Fit For 55' legislative package, and the REPowerEU plan, the 2024 EPBD recast proposal was presented to accelerate the reduction of GHG and energy poverty in the region by achieving a zero-emission building stock by 2050. The 2024 EPBD recast expanded the use of renewable energy in buildings, improved the framework of national building renovation plans, set the monitoring framework, set the minimum energy performance requirements for buildings, and expanded the concept of one-stop shops for building renovation (European Parliament, 2024). The newest EPBD recast integrated more aspects of energy poverty alleviation.

The 2024 EPBD recast has introduced several elements targeted at energy poverty. Article 3 on the National Building Renovation plans was updated to include vulnerable households and goals for reducing energy poverty (European Parliament, 2024). Furthermore, the Commission will evaluate the National Building Renovation plans based on their ability to reduce energy poverty to the levels set in the planning documents. Article 9 formalised the need to increase the technical support for vulnerable households to meet the minimum energy performance standards (European Parliament, 2024). Article 17 mandates the creation of a link between the financial measures for renovation and vulnerability (European Parliament, 2024). Furthermore, it formalises the prioritisation of energy-poor and vulnerable populations as the beneficiaries of financial incentives created for the renovation. Within Article 18, the one-stop shops shall offer a dedicated service for vulnerable households (European Parliament, 2024). The overall framework of EPBD recast integrates different criteria and parameters for energy poverty evaluation. However, the current design puts the obligation of designing strategies for energy poverty alleviation, which will result in an asymmetry of action between the Member States as there is no binding goal for energy poverty reduction.

2.2.2. National policy frameworks

Estonia

On August 17, 2023, the updated Estonian National Energy and Climate Plan (NECP) for 2030 was submitted. The development process included the Ministry of Economic Affairs and Communication, the Ministry of Environment, and the Ministry of Rural Affairs, with 106 measures, out of which 44 were related to the energy and building sectors (Directorate of European Commission, 2023).

In the building renovation, the plan addresses the target goals of the Estonian Long-term Strategy for the Renovation of Buildings (REKS) and the Estonian Energy Policy Development Plan (ENMAK) (Directorate of European Commission, 2023). The REKS goals are set as the square metres of the area of renovated buildings with preliminary goals for each building type and a 5-year period between 2021 and 2050. The overall notion of the renovation volumes proposed by REKS suggests an immediate acceleration of retrofitting activities from 2021 onward for multi-apartment buildings. At the same time, a similar acceleration pattern is noticed only after 2031 for both single-family housing and private non-residential buildings (Majandus- Ja Kommunikatsiooni Ministeerium, 2020).

REKS mentions energy poverty as not a widespread problem in Estonia, yet acknowledges the need for additional support to vulnerable households in participating in building renovation projects. The timing of REKS preparation is important, as it was published in 2019. Further pandemic and attempted invasion of Ukraine increased energy poverty rates by increasing energy consumption, disrupting energy supply chains, and due to the need to change the fossil fuel import structure (Bórawski et al., 2022). With this context, the Estonian Resilience and Recovery Plan has proposed an additional Housing Investment Fund (HIF) to address the issues that vulnerable households face in participating in the renovation projects. HIF will combine public and private financing sources to bring renovation loans to the residents of multi-apartment buildings with low property value and limited capacity for action (Directorate of European Commission, 2023). Due to the long loan period of 30 years and low interest rates at around 2-3%, it is envisaged as a suitable solution for lower socio-economic classes (Riigi Tugiteenuste Keskus, 2023). No explicit mentions of social vulnerability were observed outside the proposed financing solutions within Estonia's planning documents. Furthermore, the Estonian NECP does not mention energy poverty as an issue.

Latvia

Compared to Estonia and Lithuania, Latvia's national government did not submit an updated NECP draft in 2023 as envisioned in the 2023 recast of EED and developed it only by July 2024 (Latvijas Republikas Klimata un enerģētikas ministrija, 2024). The Latvian NECP was finalised on May 7, 2019. The Ministry of Economy led the development of the planning document together with the Ministry of Regional Development and Environmental Protection, Ministry of Transportation, and Ministry of Agriculture (Ministru Kabinets, 2019). In contrast, the update was led by the Ministry of Climate and Energy (Latvijas Republikas Klimata un enerģētikas ministrija, 2024).

In the Latvian NECP, three strategic pathways are envisaged: reduction of energy resource consumption, energy efficiency improvements in housing and small housing complexes, and the creation of long-term solutions for Latvian housing estate improvements and the integration of additional financing. In all the goals, a strong focus on ESCO service development and the need for additional refinancing solution integration was stated, even though the market of ESCO services in Latvia is underdeveloped, with only one organisation performing such work. The support of existing instruments, such as the ALTUM grant programme for building renovation, was stated within the Long-term renovation strategy (Ministru kabinets, 2020).

Energy poverty is defined in Article $1(10^1)$ of the Republic of Latvia Energy Law since February 2, 2021. Furthermore, Section 17 of the Energy Law is dedicated to energy poverty. It defines energy-poor households as “households that are considered poor or low-income that receive financial support for housing expenses” or households that rent social housing (Latvijas Vēstnesis, 2024). Article 121 further requires the national agencies and legislative bodies to prioritise and consider energy-poor populations in energy efficiency policies. Yet the legal framework does not support regional and local level governance, which operates with energy poverty and energy-poor populations more directly. The 2019 Latvian NECP has already addressed energy poverty by further estimating the energy-poor population numbers. It further establishes the goal of reducing the proportion of energy-poor households from 9.8% to 7.5% by 2030. However, the only planned activity to support energy poverty alleviation is introducing the concept in the national policy, which was performed in 2021 in the Energy Law (Latvijas Republikas Ekonomikas Ministrija, 2019; Latvijas Vēstnesis, 2024). The updated NECP introduces an additional subsidy support mechanism that increases investment affordability in energy efficiency improvements of buildings and equipment for energy-poor households. The mechanism aims to support 2017 energy-poor households by 2030 (Latvijas Republikas Klimata un enerģētikas ministrija, 2024).

Lithuania

The updated draft of the Integrated NECP of the Republic of Lithuania was submitted on July 24, 2023, by the Ministry of Energy, with 209 implemented or planned measures to address the emission reduction and challenges in the energy sector.

In building renovation, the plan addresses both the existing and planned measures. The existing multi-apartment building renovation programme is being implemented between 2021 and 2026, with a total renovation rate of 3267 buildings by 2027 (The Ministry of Energy of the Republic of Lithuania, 2023). Within the Lithuanian NECP, the existing building renovation programme is extended until 2030. Even though the existing measures will continue, Lithuania's long-term renovation strategy acknowledges the need to improve and scale up the existing funding measures to meet new objectives (Government of the Republic of Lithuania, 2021).

The Lithuanian NECP identifies energy poverty as a primary social context for the plan. It further states that Lithuania is one of the countries most affected by it in the EU (The Ministry of Energy of the Republic of Lithuania, 2023). The plan contains the goals of the National Progress Programme 2021-2030 to reduce the share of the population unable to keep their home adequately warm from 28.0% in 2018 to 17.0% in 2030 and the share of households that spend two times the median of the household national median energy costs from 17.1% to 10.0% (The Ministry of Energy of the Republic of Lithuania, 2023). Additionally, five measures are represented in the plan, with two new measures planned: more physical interaction with hard-to-reach consumers and energy efficiency information hub creation.

2.3. Supervised machine learning for building energy performance prediction

Seyedzadeh et al. (2018) define four main categories of building energy assessments: engineering calculations, simulation model-based benchmarking, statistical modelling, and machine learning (ML). The supervised ML is defined as a function approximation problem, where the training data takes the form of predictors and a predicted variable to produce a new predicted variable from the query of predictors (Jordan and Mitchell, 2015). Two primary categories of supervised ML problems are classification, where the predicted variable is represented by a label, and regression, represented by a continuous value (Jordan and Mitchell, 2015; Gianey and Choudhary, 2017). In classifying multiclass labels, the k-nearest neighbour (kNN) algorithm is deemed the simplest, whereas linear regression (LR) is the simplest algorithm for regression problems (Gianey and Choudhary, 2017). The Seyedzadeh et al. (2018) review further defines two primary supervised ML algorithms for building energy predictions and modelling in a context of lack of uncertainty in data: support vector machines (SVM) and artificial neural networks (ANN). Other studies have pointed out that ensemble-based methods, such as random forest (RF) and eXtreme gradient boosting (XGBoost), are also suitable for building energy performance evaluation (Tsanas and Xifara, 2012; Pham et al., 2020; Ali et al., 2024)

2.3.1. Support vector machines

SVM describes algorithms of distribution-free learning for non-linear dependencies in regression and classification problems between the vector inputs and the dependent variable (Kecman, 2005). Methodologically, SVM is defined as identifying the dependency function of the weights, which are the subject of learning (Samardzioska et al., 2021). SVM is particularly useful in the context of multi-dimensional problems, as the model complexity is not dependent on the dimensionality of space (Vapnik et al., 1996).

Samardzioska et al. (2021) applied SVM to develop an annual building energy consumption prediction model, utilising data for sixty renovated buildings in North Macedonia. A building survey was performed for every building to evaluate the thermal conductivity of different surfaces and gather data on building geometries. The developed model's mean absolute percentage error (MAPE) was 2.44%, whereas the MAPE of LR was 8.37% (Samardzioska et al., 2021). In a different context, Dong et al. (2005) applied SVM for the estimation of the landlord energy consumption for four commercial buildings in the Central Business District of Singapore. Using longitudinal and weather data for energy consumption, the authors developed an SVM model for predicting energy consumption with error values between -2.72% and 3.44% for different buildings in the study. However, the authors note that the algorithm's performance can be attributed to the small data pool with few abnormalities (Dong et al., 2005).

2.3.2. Artificial neural networks

As a method, ANN represent a series of connected basic computing units for local information transmissions resembling the operations of biological neurons (Yegnanarayana, 2004). The model consists of three primary layers – input, output, and hidden layers – which are interconnected, and the signal is transferred from one layer to another, impacted by the applied transfer and activation functions (Tsoka et al., 2022).

Tsoka et al. (2022) applied ANN for the building energy label classification based on the two methodologies of residential building EPCs in the Lombardy region of Italy. Additionally, the study utilised other input datasets associated with building characteristics, environmental conditions, and social factors. The developed algorithms achieved 93.10% for the before-2015 EPC labelling, and 89.62% for the post-2015 EPC labelling system in Italy (Tsoka et al., 2022). Chari and Christodoulou (2017) used ANN to predict building energy ratings based on various dwelling configurations created using the Irish 'Dwelling Energy Assessment Procedure'. 68 energy-related factors, including external factors such as weather conditions, dwelling location, building characteristics, and thermal properties, were used for the model development, which were further divided into models with fewer associated input factors. The study found that increasing input factors reduces prediction accuracy while simultaneously decreasing the variance of building energy ratings (Chari and Christodoulou, 2017). Therefore, finding an optimal balance between accuracy and variance is necessary (Chari and Christodoulou, 2017).

2.3.3. Random forest

RF represents an ensemble-based ML algorithm. Ensemble algorithms are a set of algorithms whose individual decisions are combined through weighted or unweighted voting as a way of performance improvement (Dietterich, 2000). RF is one of the primary algorithms for classification and regression problems, within which many decision trees are generated (Breiman, 2001). For performance improvement, randomness is integrated as random feature selection to reduce overfitting, strengthening the model performance over other bagging algorithms (Breiman, 2001).

Tsanas and Xifara (2012) applied an RF algorithm to predict heating and cooling loads based on 768 developed building models created based on different combinations of building orientations, forms, glazing areas, and their distribution theoretically located in Athens, Greece. The study concluded that applying the RF for building energy load prediction is a suitable method due to the low variations between the ground truth and outputs of the ML model (Tsanas and Xifara, 2012). Furthermore, the authors stress that the linear techniques are inappropriate for building energy load modelling. Wang et al. (2018) also used RF to make a regression predictive algorithm for hourly electricity consumption prediction based on two buildings' one-year historical energy consumption data. The data comprised the weather variables, the number of occupants, and temporal characteristics. The study concluded that applying random forest algorithms is more suitable for electricity consumption estimation than a regression tree and support vector regression (SVR) (Wang et al., 2018).

2.3.4. eXtreme gradient boosting

As with RF, XGBoost is an ensemble-based algorithm. However, XGBoost is based on gradient tree boosting, which includes functions as parameters and requires training in an additive manner with smaller trees compared to RF (Chen and Guestrin, 2016; Seyedzadeh et al., 2020). Furthermore, the algorithm authors note that XGBoost allows for parallel and out-of-core computations in cash- and sparsity-aware learning (Chen and Guestrin, 2016).

Ali et al. (2024) applied different ensemble ML algorithms, including XGBoost and RF, and the end-use demand segregation methods to predict the building energy performance of a set of building archetypes based on the entire building stock of Ireland. XGBoost performed best, whereas LR, kNN, and SVM models performed the worst based on root mean square error, mean absolute error, and accuracy (Ali et al., 2024). The methodology was applied to evaluate different building renovation scenarios associated with implementing Ireland’s National Climate Action Plan 2023. The dataset developed by Tsanas and Xifara (2012) was used by Alawi, Kamar and

Yaseen (2024) to compare the performance of different machine learning algorithms, including SVR, kNN, RF, XGBoost, multi-layer perception, and gradient boosting algorithms. As with the Tsanas and Xifara (2012) study, the modelled heating and cooling loads for the developed building models based on the set of performance metrics for different numbers of input variable scenarios. The study identified RF models deemed overall the best for heating and cooling load predictions, whereas XGBoost showed particularly good results in cooling load predictions (Alawi et al., 2024).

3. Methodology

The overall methodological framework of the study comprises four core elements: data preparation, model selection process based on the selected error metrics, application of the selected model for all three cities, and multi-level statistical analysis of building- and neighbourhood-level characteristics correlated with residential building energy performance. The methodology for the study is summarised in Figure 4. The model selection step was performed only for the Tallinn building and energy data. The most suitable model based on the selected error metric performance was replicated for Riga and Vilnius building and energy data. The described workflow is performed in the R software.

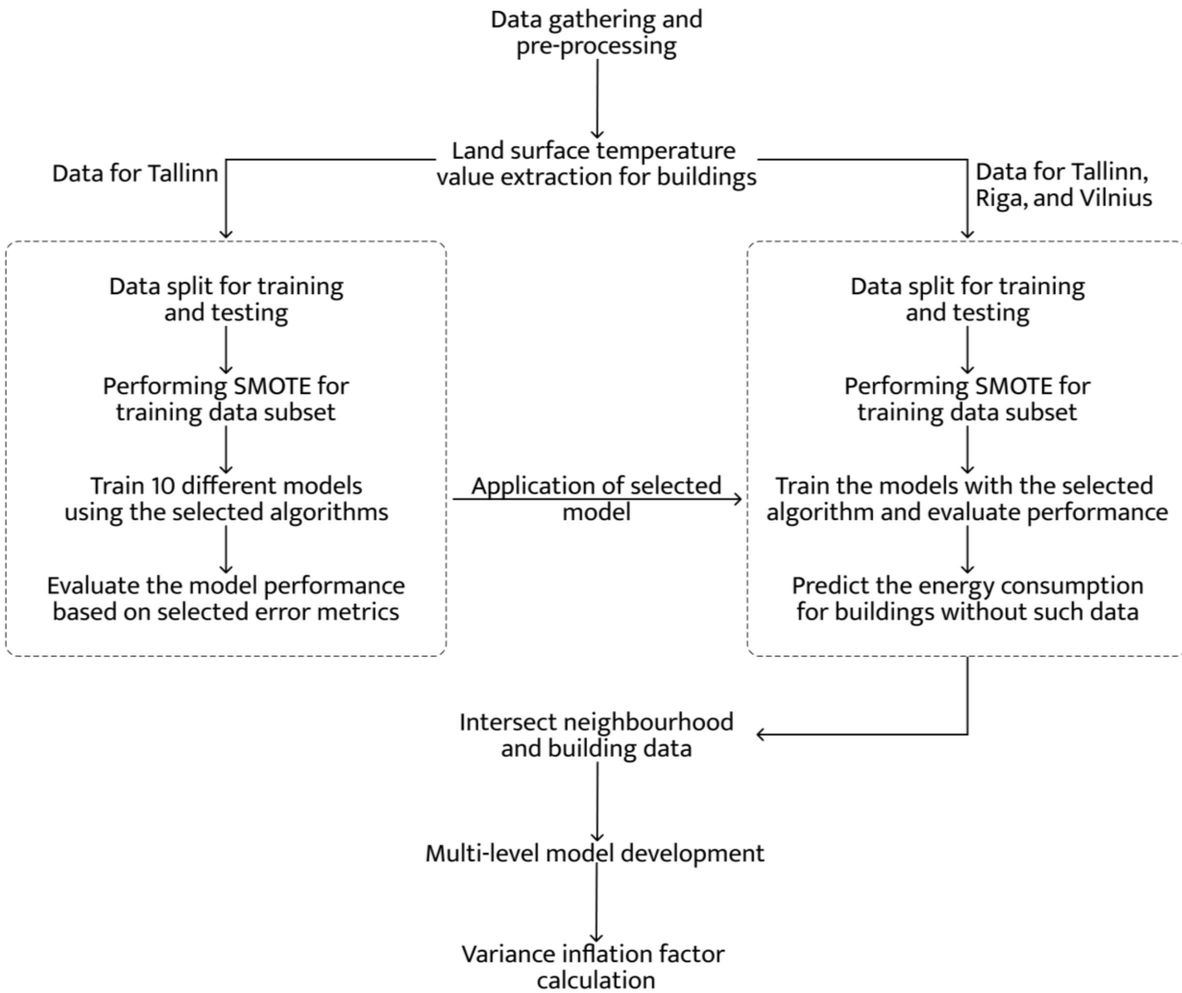

Figure 4. The graphical summary of the study methodology.

3.1. Study areas

3.1.1. Tallinn, Estonia

Tallinn was founded in the thirteenth century, with a rapid city development starting in the 1860s due to rapid industrialisation. During the Soviet Union occupation, the city significantly expanded, predominantly due to external migration (Ruoppila and Kährik, 2003). To meet the housing demand, large housing estates such as Mustamäe, Väike-Õismäe (located in the district of Haabersti), and Lasnamäe were built to meet the housing demand. Differentiation of housing types and districts was predominantly observed from the ethnic perspective, with the non-Estonian population being overrepresented in the housing estates, whereas the socio-economic segregations remained relatively low until the 2000s (Ruoppila and Kährik, 2003). The post-socialist housing of Tallinn has been shaped by privatisation and suburbanisation (Ruoppila, 2007), which increased occupational and income disparities (Tammaru et al., 2015). The inner city, which comprises the neighbourhoods in Kesklinn, Põhja-Tallinn, and Kristiine districts, has been restructured economically and socially, and gentrification is taking place (Pastak and Kährik, 2021; Maloutas and Karadimitriou, 2022).

The distribution of residential buildings based on the period of construction is summarised in *Figure 2*. The neighbourhoods of the inner city predominantly observe pre-socialist housing with a mix of both socialist-era and post-socialist housing. The different neighbourhoods of Lasnamäe and Mustamäe, as well as parts of Haabersti and Põhja-Tallinn, are overrepresented by socialist-era multi-apartment housing with smaller proportions of post-socialist housing developed in between the housing estates. In the case of Tallinn, three distinct areas of higher representation of detached dwellings are observed. They consist of Northern Pirita and North-Western Haabersti, which are dominated by detached post-socialist housing, as well as a more mixed district of Nõmme, where a mix of pre-socialist and socialist detached housing and small apartment buildings is observed.

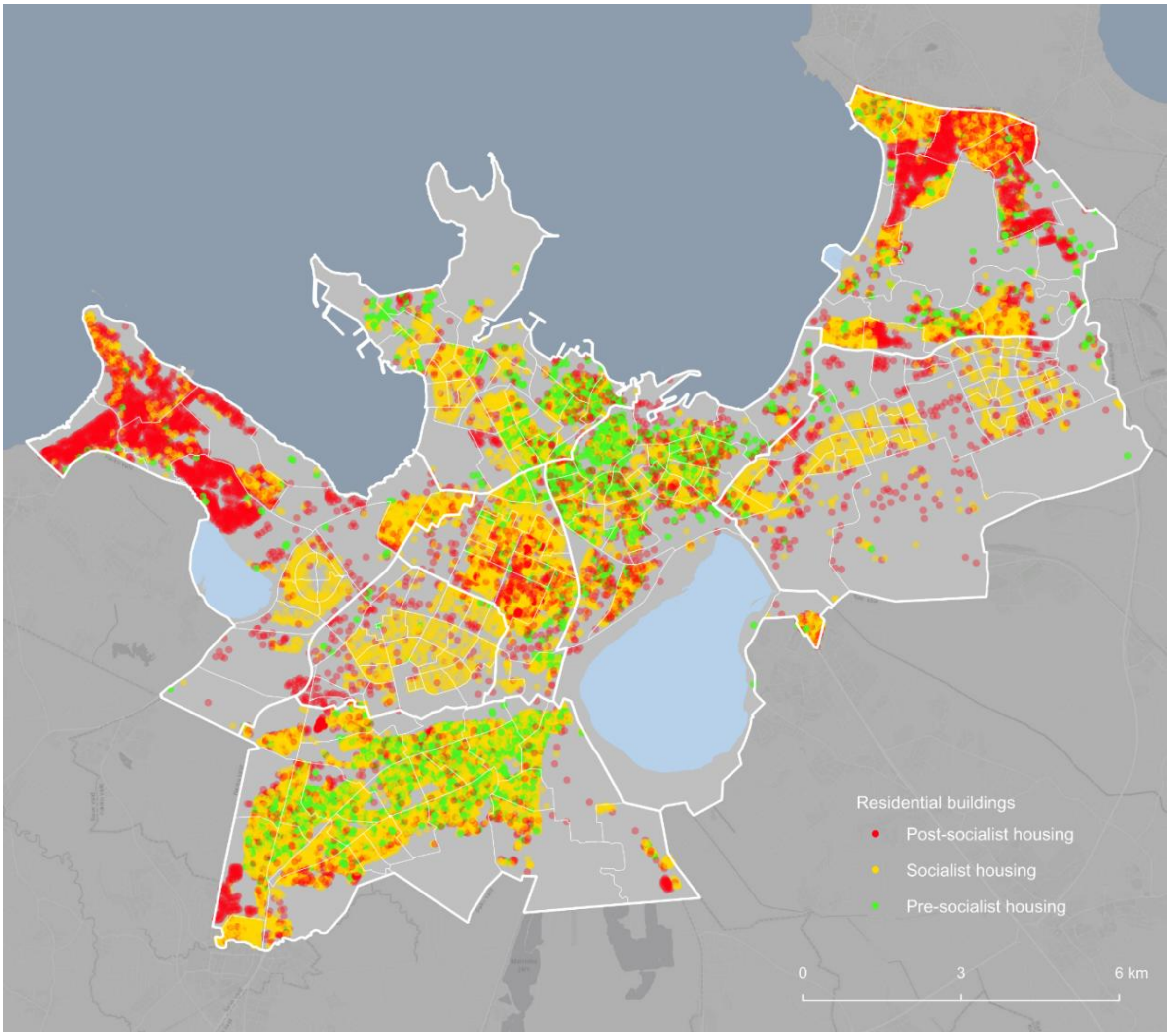

Figure 2. Distribution of residential buildings based on the period of construction in Tallinn.

3.1.2. Riga, Latvia

Riga was founded at the beginning of the thirteenth century, and, similarly to Tallinn, started rapid development in the late nineteenth century. At the beginning of the twentieth century, Riga was the second largest city in the western part of the Russian Empire after St. Petersburg (Krišjāne and Bērziņš, 2014). Housing estates were built in Riga to meet the housing demands created by industrialisation and urbanisation policies of the Soviet Union. The current housing market in Riga was significantly shaped by neoliberal housing management policies characterised by privatisation, deregulation, and reduced state intervention stemming from the dissolution of the Soviet Union in the early 1990s (Lulle, 2024). Furthermore, a rapid population decrease was observed in Riga, with a simultaneous growth of the metropolitan area due to low-density suburbanisation processes in neighbouring municipalities (Freimane, 2020).

The distribution of housing based on the period of construction is illustrated in *Figure 3*. The neighbourhoods of the inner city and its peripheral neighbourhoods are overrepresented by pre-socialist multi-apartment housing. The distribution of newer housing is very sparse in the areas of the inner city. Outside of the inner city, many socialist housing estates were developed, including the neighbourhoods of Iļģuciems, Imanta, Purvciems, Pļavnieki, and Ķengarags. Since then, many of these neighbourhoods have also seen the development of post-socialist housing. The peripheries of Riga are predominantly represented by post-socialist period detached housing, representing the suburbanisation processes described by Freimane (2020).

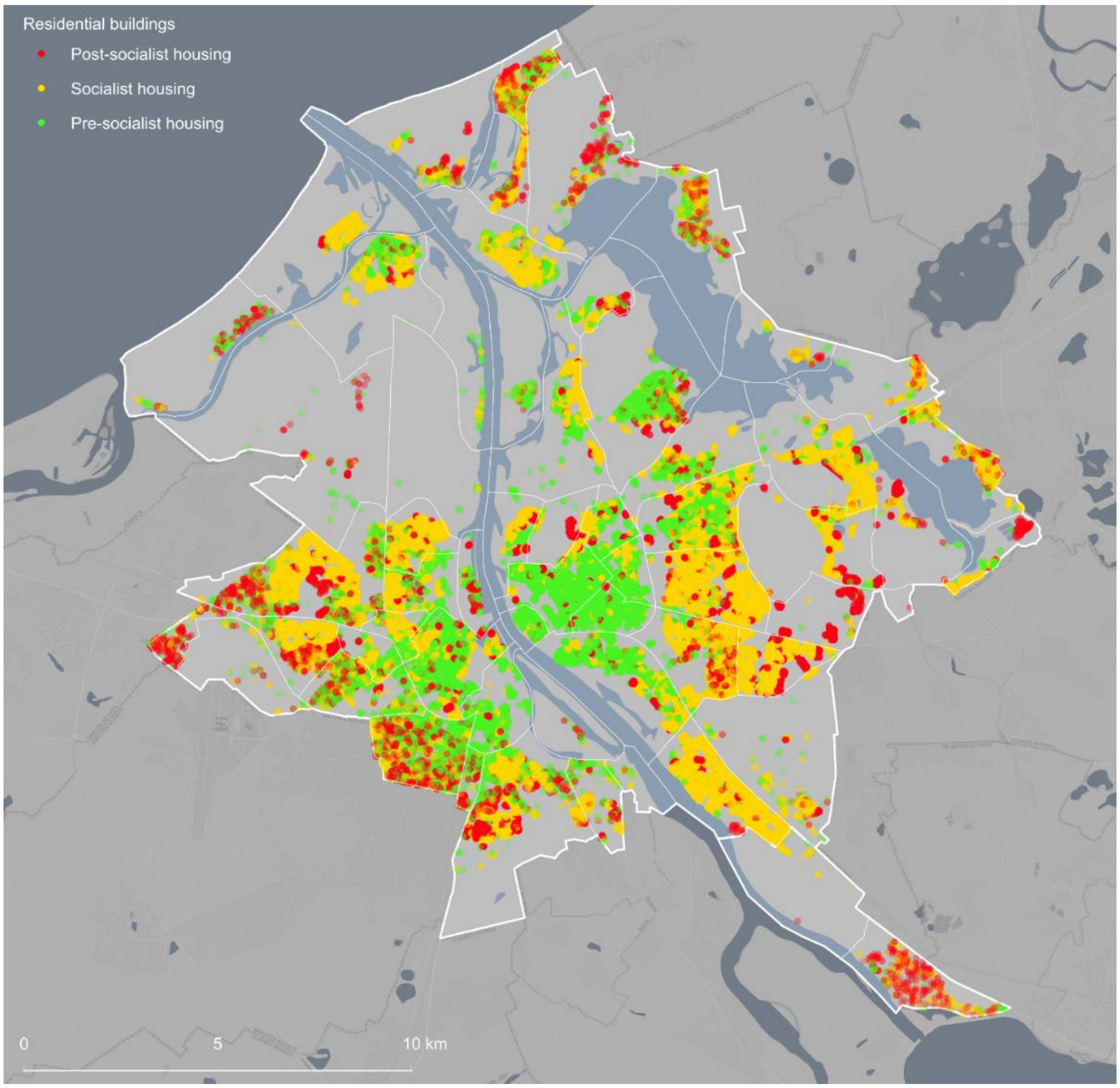

Figure 3. Distribution of residential buildings based on the period of construction in Riga.

3.1.3. Vilnius, Lithuania

Vilnius was founded in the fourteenth century. However, Vilnius had significantly slower urban development in the nineteenth century compared to Tallinn and Riga due to the lack of significant industrial expansion (Samalavičius et al., 2024). A rapid industrialisation and development of housing estates in Vilnius started in the 1960s, with Lazdynai neighbourhood becoming an emerging symbol of socialist ideas in the cityscape (Mikailiene, 2010). In the post-socialist transition, the housing privatisation and the spread of the service economy were the key drivers of new urban developments in Vilnius, characterised by fragmentation (Burneika, 2008). Ethnic composition and intensive suburbanisation were identified as key drivers of segregation levels and patterns in Vilnius (Burneika et al., 2019; Ubarevičienė and Burneika, 2020).

Figure 4 shows the distribution of residential housing divided by construction period. In the case of Vilnius, the socialist-era built housing overall dominates in many of the existing neighbourhoods, with post-socialist neighbourhoods being developed in neighbourhoods further away from the inner city. In the inner city, more variability in housing based on the construction period is observed. The observed distribution of residential housing follows the slow urban development rates during the pre-socialist period described by Samalavičius et al. (2024).

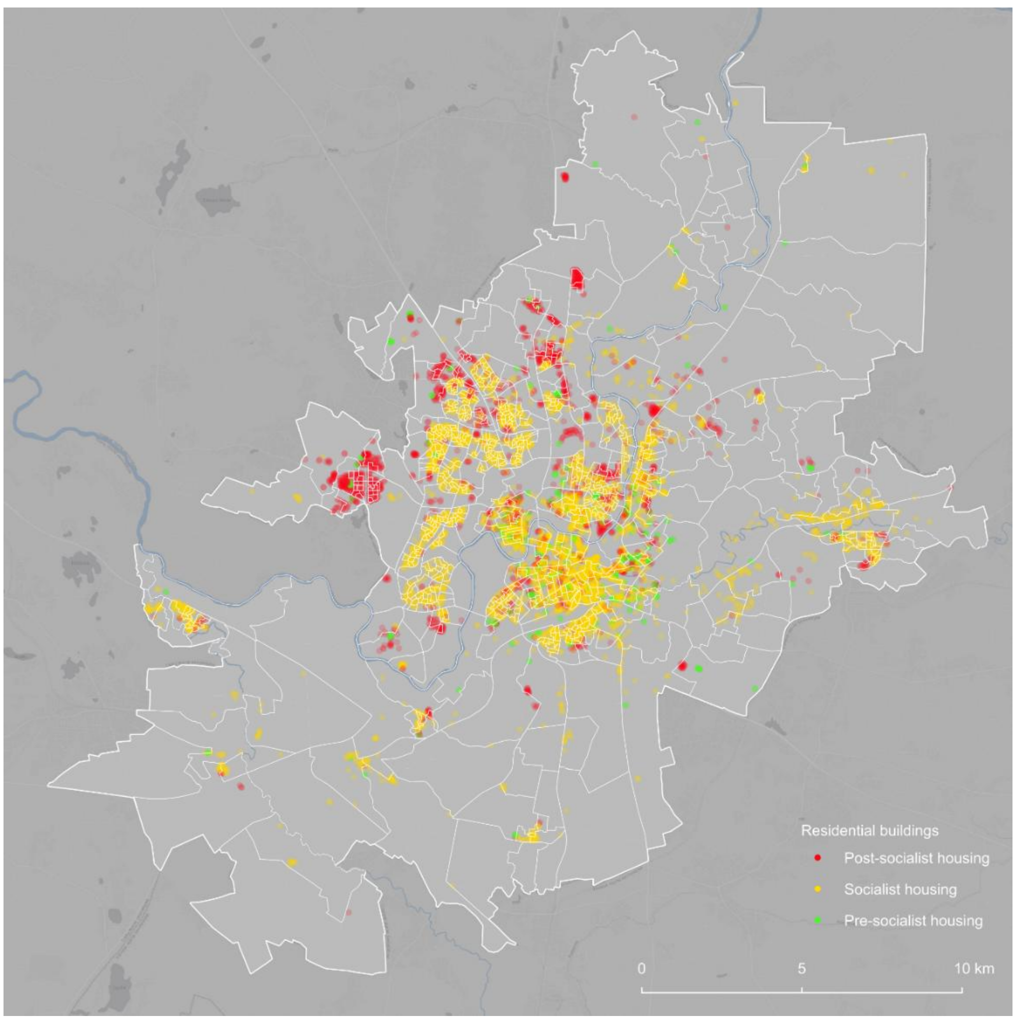

Figure 4. Distribution of residential buildings based on the period of construction in Vilnius.

3.2. Data

Open-access, publicly available data was used to develop the machine-learning-based model for building energy performance prediction. Dataset typologies are summarised in *Table 2*, and the detailed dataset description is provided in *Annexe I*.

Table 2. Summary of the datasets used in building an energy performance predictive machine-learning-based model.

Dataset	Type	Described time	Variables used
Building topographical data	V	Estonia: 05.11.2024	*Estonia:* Building code, year of building construction, building height (m), total projected building surface (m^2)
		Latvia: 10.05.2024	*Latvia:* Building code, year of building construction, number of building floors, building construction area (m^2)
		Lithuania: 11.11.2024	*Lithuania:* Building address, year of building construction, total heated area (m^2), recorded building renovation activities
Energy performance certificates	A	Estonia: 05.11.2024	Estonia: Building code, energy label, Total annual energy consumption (kWh/m^2)
		Latvia: 06.10.2024	Latvia: Building code, energy label, Total annual heat energy consumption (kWh/m^2)
		Lithuania: 02.01.2025	Lithuania: Building address, total annual energy consumption, energy label
Land surface temperature	R	Estonia: 06.03.2024 Latvia: 06.03.2024 Lithuania: 07.03.2024	Land surface temperature (°C)
Building renovation grants	A	Estonia: 31.12.2023 Latvia: 31.12.2023	Building code

Note: Type denotes the dataset typology: 'V' stands for a vector dataset, 'R' stands for a raster dataset, and 'A' stands for an attribute tabular dataset.

The study uses EPCs as the primary energy data source on building levels. However, several key differences exist between the EPCS in Estonia, Latvia, and Lithuania, as described in the following subsection. To obtain the 10-metre resolution land surface temperature raster datasets, an algorithm developed by Onačillová et al. (2022) combining Landsat-8 and Sentinel-2 raster data was used. Based on Landsat-8 data, the authors defined a multiple linear regression model to represent the relationship between spectral indices – normalised difference vegetation, built-up, and water indices - and land surface temperature. This model, which was made available in Google Earth Engine, enabled the calculation of land surface temperature at a 10-metre resolution using the same indices (Onačillová et al., 2022). The date of measurement of satellite data was used as the primary data filter, excluding newer data from the study.

The population data for Tallinn, Vilnius, and Riga were provided to the author by partners of the Centre for Migration and Urban Studies of the University of Tartu in Latvia and Lithuania. In all three cases, the data used was the 2021 population census. Its aggregation level was the 1 km^2 grid cells for Riga, Latvia, and to the level of small statistical neighbourhoods in Vilnius, Lithuania and Tallinn, Estonia. The occupational status of workers and population age groups were synchronised between all the datasets. However, the variable describing the population's ethnicity varied between the capital cities: the declared ethnicity was used for Riga, with four categories being Latvian, Russian, other, and non-declared; The primary household language was used for Tallinn, with three categories being Estonian, Russian, and other; and the mother tongue of a person was used for Vilnius, with four categories being Lithuanian, Polish, Russian, and other.

3.2.1. Building energy performance certificates

Building EPCs were introduced in the original EPBD text in 2002 as the legally recognised certificate to describe the energy performance of a building based on the established reference values and benchmarks for cross-comparison (European Parliament, 2003). However, the transposition of the EPBD and following recasts allowed for the variation in the definition of methodologies for EPCs, there are several core differences in its interpretation between Estonia, Latvia, and Lithuania. In Estonia, the energy label is determined based on the total energy demand associated with building energy use, which varies across different building typologies and sizes (Ruggieri et al., 2023). In Latvia, the energy label is determined based on actual building energy consumption for heating, cooling, ventilation, hot water preparation, and lighting, with variability between different sizes of residential and non-residential buildings and reference values for total energy consumption and energy consumption for heating (Ministru kabinets, 2021). In Lithuania, the energy label is determined based on the thermal transmission characteristics of the building and the efficiency of using primary non-renewable energy (Ruggieri et al., 2023). Furthermore, no national legal benchmarks and reference values for energy consumption exist (Ruggieri et al., 2023).

The described differences between the EPCs in the three countries required the creation of an additional variation of the methodology based on the described. In Estonia, the reference values for the energy labels were used with the variation between the different sizes of buildings, with three main categories being below 120 m^2, between 120 and 220 m^2, and above 220 m^2 (Majandusja taristuministri, 2019). Furthermore, separate ranges were designed for multifamily buildings (Majandus- ja taristuministri, 2019). In Latvia, a single legally defined energy label category was used for all buildings (Ministru kabinets, 2021). The means were calculated from the pre-defined energy consumption values. However, as the Lithuanian energy performance certificate regulation does not contain the reference energy consumption values per energy label, the study author computed and applied the median values of total energy demand per energy label for Vilnius. The final values used in the study are described in Table 3.

Table 3. Values of energy performance labels used in the study.

Label	**Estonia-120 (kWh/m^2)**	**Estonia-120-220 (kWh/m^2)**	**Estonia-mf (kWh/m^2)**	**Latvia (kWh/m^2)**	**Lithuania (kWh/m^2)**
A++	-	-	-	-	12.6
A+	-	-	-	30	31.3
A	145	120	105	30-40	38.0
B	146-165	121-140	106-125	40-60	73.4
C	166-185	141-160	126-150	60-80	126.2
D	186-235	161-210	151-180	80-100	166.8
E	236-285	211-260	181-220	100-125	285.6
F	286-350	261-330	221-280	125	290.0
G	351-420	331-400	281-340	-	360.6
H	421	401	341	-	-
Source:	Majandus- ja taristuministri (2019)	Majandus- ja taristuministri (2019)	Majandus- ja taristuministri (2019)	Ministru kabinets (2021)	*Own calculation based on Energy Performance Certificate data*

Note: ‘mf’ denotes multifamily housing.

3.3. Methods

Data pre-processing and augmentation

The work with the vector data was performed using the sf package (Pebesma, 2018; Pebesma and Bivand, 2023), whereas the raster data was manipulated with the *terra* package (Hijmans, 2025). The work was initiated by sampling the land surface temperatures for every building, which was done by calculating the mean value for the land surface temperature raster centroid points within the boundaries of the vectors. Then, the EPC data was merged with the vector building dataset using the building code or address. The EPCs were filtered out to include only the ones active during the selected date for the land surface temperature raster data. The longitude and latitude of building area polygons were added as additional input variables for the predictive machine learning models. The final datasets used for the model training were attribute tables.

A testing and training data split was performed to preserve a part of the dataset and calculate the selected metrics for errors. 10% of the available data remained to be used only for the performance indicator calculations. In the model selection process and the final model development without the data split, the synthetic minority over-sampling technique (SMOTE) was used. As the building energy labels were unequally distributed between the classes, SMOTE allowed the introduction of synthetic data generated along the line segments for a selected number of neighbours (Chawla et al., 2002). In this study, SMOTE was performed by using the *smotefamily* package (Siriseriwan, 2024). For the hyperparameter optimisation for the selected predictive models, repeated cross-validation was performed with the *trainControl* function in the *caret* package, with five folds repeated three times (Kuhn, 2024).

3.3.1. Predictive model selection and application

The predictive model selection process was performed based on Tallinn building and energy data. A set of criteria was selected to evaluate and compare the performance of the models, with the best-performing modelling framework replicated for Riga and Vilnius. The final output of this work section included three predictive machine-learning-based models for every selected capital city. *Equations 1, 2,* and *3* describe the input models for the different machine-learning model types used in the study for Tallinn, Riga, and Vilnius.

$$Y_{Tallinn} \sim YC + Long + Lat + LST + S + H + V \qquad (1),$$

where

Y – predicted energy variable,

YC – year of building construction,

Long – the centroid points’ longitude,

Lat – the centroid points’ latitude,

LST – mean land surface temperature (°C),

S – building projection area (m^2),
H – building height (m),
V - building volume (m^3).

Due to data availability on buildings in Riga, the height of the building was substituted by the number of building floors in the model. Furthermore, the lack of a height variable did not allow for the calculation of the total building volume used for the Tallinn machine-learning model.

$$Y_{Riga} \sim YC + Long + Lat + LST + S_c + F \quad (2),$$

where
S_c – building construction area (m^2),
F – number of building floors.

The building height and volume variables were unavailable within the data available for the Vilnius building stock. In this instance, it has been substituted by the total heated area, which depends on the number of floors.

$$Y_{Vilnius} \sim YC + Long + Lat + LST + S_h + S \quad (3),$$

where
S_c – total heated area (m^2).

Selected models

In total, ten models were selected for the predictive model development. The tested models are described in *Table 4*. The models represent classification and regression problems predicting either the energy label or the total annual energy consumption per square metre. Only the training and testing data subset with the energy label and total annual energy consumption values were used to facilitate comparability between the classification and regression models. Furthermore, the values of energy labels were transformed into the total annual energy consumption values to compare the results based on the selected performance indicators.

Table 4. Summaries of the selected algorithms for model development.

Model	**Abbreviation**	**Type of model**	**Package**	**Reference**
Naïve Bayes	NB	Classification	*klaR*	Weihs et al. (2005)
Support Vector Machine with Class Weights	SVM	Classification	*kernlab*	Karatzoglou et al. (2004)
k-Nearest Neighbour	KNN	Classification	Base R	R Team (2014)

Random Forest	RF	Classification, Regression	*randomForest*	Liaw and Wiener (2002)
eXtreme Gradient Boosting	XGBoost	Classification, Regression	*xgboost, plyr*	Chen and He (2024), Wickham (2011)
Linear Regression	LR	Regression	Base R	R Team (2014)
General Additive Model	GAM	Regression	*mgcv*	Wood (2011)
Bayesian Regularized Neural Networks	BRNN	Regression	*brnn*	Rodriguez et al. (2023)

The instances of application of SVM, kNN, RF, XGBoost, LR, and neural networks were already identified and described in the existing literature on machine-learning algorithm applications to develop models for energy performance evaluation. Two additional statistical modelling algorithms were added for the selection process: naïve Bayes (NB) and general additive model (GAM).

NB in the literature was described as one of the most effective classifiers (Friedman et al., 1997), suitable for large data handling (Wu et al., 2008). During the training phase, the classifier learns the conditional probability of each attribute for each class label (Friedman et al., 1997). During the classification, Bayes' rule is applied to compute the class probability of the attribute combination and the prediction is made based on the highest posterior probability (Friedman et al., 1997). One of the key features of the NB is the probabilistic independence of attributes (Friedman et al., 1997; Wu et al., 2008).

GAMs are a likelihood-based regression model which utilises scatterplot smoothers to generalise Fisher scoring procedures, first conceptualised by Hastie and Tibshirani (1986) as a flexible method for nonlinear covariate effect identification. In GAMs, the selection of the smoothing function is the primary consideration for the model development (Wood, 2025). Limited applications of GAMs in building energy modelling were also identified (Kilis et al., 2025). However, even as GAMs are suitable for large data handling, one of the key issues, particularly in high-dimensional space, is the computational capacity of such algorithms (Wood, 2025).

Selection parameters for model performance

A set of three primary indicators, mean absolute error (MAE), root mean squared error (RMSE), and median absolute deviation (MAD), was used to compare the different developed models based on their performance for the testing subset of data. These three indicators were calculated globally for all the outputs for the testing subset of data. The error curves between the predicted and actual energy consumption values were plotted to assess the model performance further, as within the existing literature on global metric selection, there exists a consistent debate on the application of the different indicators (Willmott and Matsuura, 2005; Chai and Draxler, 2014; Hodson, 2022).

MAE describes the absolute difference between the predicted and actual value, calculated with *Equation 4*. MAE is a suitable global model evaluation parameter for exponentially distributed variables and Laplacian-like error distributions (Hodson, 2022). Importantly, MAE gives the same weight to all errors compared to RMSE (Chai and Draxler, 2014), and preserves the unit of data (Hodson, 2022).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|EC_{actual} - EC_{predicted}\right| \qquad (4),$$

where
EC_{actual} – actual building energy consumption, kWh/m^2,
$EC_{predicted}$ – predicted building energy consumption, kWh/m2.

RMSE describes the square root of the averaged squared difference between the actual and predicted value, described by *Equation 5*. As an indicator, RMSE was shown to be more sensitive towards outliers, especially when the distribution of error magnitudes is highly variable (Willmott and Matsuura, 2005). However, Chai and Draxler (2014) have shown that RMSE is more appropriate for normally distributed errors and described that the use of absolute values, such as within MAE, should be avoided for many mathematical calculations.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(EC_{actual} - EC_{predicted}\right)^2} \qquad (5)$$

As the application of RMSE or MAE has been shown to have associated proponents, the study utilises MAD as an additional metric of model performance. It is described as the median value of the absolute predicted values and the actual global median, calculated by *Equation 6*. As with the MAE, MAD preserves the data unit, yet is described as the more robust metric for the error distribution, deviating from the normal distribution (Hodson, 2022).

$$MAD = 1.483 * median_{i=1}^{n}\left|EC_{predicted} - median(EC_{actual})\right| \qquad (6)$$

The final model was selected based on the best performance on the metric, taking into consideration the error distribution curves. After the model selection process for the Tallinn building and energy data, the selected model was applied to Riga and Vilnius. Following the argument of Hodson (2022) on the application of mean values for error distribution different from Gaussian or Laplacian, the study uses median absolute percentage error (MAPE) to evaluate the final model performance using 10% testing data. *Equation 7* summarises the calculation of median absolute percentage error.

$$MAPE = 100\% * median_{i=1}^{n} \left| \frac{(EC_{actual} - EC_{predicted})}{(EC_{actual}} \right| \quad (7)$$

The datasets with the actual energy consumption were merged with the machine learning model's predicted outputs. For multifamily buildings, the calculated energy consumption was multiplied by the average size of dwellings in urban areas per country. In contrast, the existing area was used for the detached housing units. When considering the energy prices, electricity and heating prices were considered for Estonia and Lithuania, whereas only the heating price was used for Riga. This stems from the differences in the EPCs definitions and available reference values per energy class.

3.3.2. Neighbourhood-level analysis

A multilevel model was developed for every capital city. Multilevel or hierarchical models represent an established modelling framework in social sciences to simultaneously study group-level and individual-level effects (Greenland, 2000). The study implemented the models using the *lme4* package in R, which was developed to create and analyse mixed-effect models (Bates et al., 2015). In the models, the dependent variable was the predicted or already available energy consumption value per square metre, whereas the independent variables described the building characteristics and neighbourhood composition. The building's heated area variable was logarithmically transformed, and it was scaled together with building age. For neighbourhood-level variables, the proportional data was calculated and scaled. Furthermore, in the multi-level statistical models, one variable per category was excluded to be the reference group. For occupational status, the reference variable is the high occupational groups; for mother tongue or ethnicity, the reference variable is the ethnic or linguistic majority group; and for population age groups, it is the working age adult group. *Equations 8, 9,* and *10* describe the developed models for Tallinn, Riga, and Vilnius.

$$EC_{nb} = \beta_0 + \beta_1 L\_RU_n + \beta_2 L_{OTH_n} + \beta_3 O\_M_n + \beta_4 O\_L_n + \beta_5 A\_C_n + \beta_6 A\_E_n + \gamma_n + \beta_7 BA_b + \beta_8 I(RG)_b + \beta_9 I(MF)_b + \beta_{10} BHA_b + \varepsilon_{nb} \quad (8),$$

where

n – neighbourhood-level variable.

b – building-level variable.

EC_{nb} – energy consumption per square metre (kWh/m^2/a).

$\beta_0 - \beta_{10}$ – estimate values.

$L\_RU_n,\ L\_OTH_n$ – primary languages used in households (Russian, Other).

$O\_M_n,\ O\_L_n$ – working individuals occupational status groups (Medium, Low).

$A\_C_n,\ A\_E_n$ – individuals by age groups (C: 0-14, E: 65+).

BA_b – building age.
$I(RG)_b$ – building renovation grant allocation.
$I(MF)_b$ – multifamily building.
BHA_b – building's heated area.
γ_n – random neighbourhood effect.
ε_{nb} – random building effect.

The primary difference between the models lies in ethnic variables. In the case of Latvia, the 2021 population census utilises the declared ethnicity of a person rather than the mother tongue or the primary language used in a household. Additionally, the dependent variable for Riga includes only heat energy costs rather than the total energy costs due to the difference in energy labels between the countries.

$$EC_{nb} = \beta_0 + \beta_1 L\_RU_n + \beta_2 L\_OTH_n + \beta_3 L\_NA_n + \beta_4 O\_M_n + \beta_5 O\_L_n + \beta_6 A\_C_n + \beta_7 A\_E_n + \gamma_n + \beta_8 BA_b + \beta_9 I(RG)_b + \beta_{10} I(MF)_b + \beta_{11} BHA_b + \varepsilon_{nb} \quad (9),$$

where
$L\_RU_n,\ L\_OTH_n, L\_NA_n$ – declared ethnicity of a person (Russian, Other, Non-declared)

For the model for Vilnius, four ethnic groups based on the mother tongue are present. These include individuals with Lithuanian, Russian, Polish, and other languages as their mother tongue. As with the Tallinn multilevel model, the dependent variable is the total annual energy costs.

$$EC_{nb} = \beta_0 + \beta_1 L\_RU_n + \beta_2 L\_PL_n + \beta_3 L\_OTH_n + \beta_4 O\_M_n + \beta_5 O\_L_n + \beta_6 A\_C_n + + \beta_7 A\_E_n + \gamma_n + \beta_8 BA_b + \beta_9 I(RG)_b + \beta_{10} I(MF)_b + \beta_{11} BHA_b + \varepsilon_{nb} \quad (10),$$

where
$L\_RU_n,\ L\_PL_n, L\_OTH_n$ – the mother tongue of a person (Russian, Polish, Other)

After developing the multi-level models, the variance inflation factors (VIFs) were calculated to determine the multicollinearity between the variables with the *performance* package in R (Lüdecke et al., 2021). VIF describes how much the estimated variance of the variable's regression coefficient is increased above what it would be if the coefficient of determination for the same variable were equal to zero (O'Brien, 2007). The upper threshold of the VIF used in the study is 10, being one of the most common thresholds (O'Brien, 2007). Afterwards, the identified patterns of inequality were assessed based on the existing literature on the unequal impacts of energy poverty between the different residential groups and emerging characterisations of unequal distributional impacts of energy efficiency and climate policies. Taking into consideration the existing policy backgrounds of the countries on energy efficiency and energy poverty, the identified patterns were analysed based on the existence of existing mitigation strategies.

4. Results

4.1. Machine learning model selection

In total, the data of 3008 buildings in Tallinn were used for the 10-model cross-comparison, out of which the data of 297 buildings were used for model testing. The dataset represented a subset of the buildings with all non-zero variables, including total annual energy consumption per square metre and energy label. The performance of SMOTE increased the total number of data points used for training purposes from 2711 to 3989. After training every model with selected classification and regression algorithms, the selection performance indicators were calculated for the testing subset of data. The energy label values were transformed for the classification models based on Estonia's defined thresholds of EPCs. The outputs of performance metric calculations are summarised in *Table 5*.

Table 5. Summary of the model selection performance indicators for Tallinn building energy performance modelling.

Model	**Type**	**MAE**	**RMSE**	**MAD**
NB	Classification	51.57	73.84	49.68
SVM	Classification	54.57	65.28	49.68
KNN	Classification	47.28	31.00	49.68
RF	Classification	37.69	54.25	24.47
XGBoost	Classification	37.88	54.36	24.47
RF	Regression	34.96	284.61	40.22
XGBoost	Regression	39.38	311.83	45.84
LR	Regression	48.51	340.64	58.12
GAM	Regression	43.91	335.94	49.70
BRNN	Regression	42.89	33.42	49.70

The results of the global metric calculations for different models show highly variable results between the different performance indicators. Based on the MAE, the RF regression model has shown the best results, whereas RMSE showed that KNN and BRNN have a better performance compared to the rest of the models. However, when considering MAD as an indicator, the performance of classification RF and XGBoost is substantially better compared to the rest of the models. The variability of the results between the different indicators required to evaluate the error distribution curves, as, based on the literature on MAE, RMSE, and MAD, different indicators are better equipped for different error distributions (Hodson, 2022). The results are displayed in *Figure 5*. Based on the distribution, the error curves for the testing data predictions are neither Gaussian nor Laplacian, which resulted in the prioritisation of the MAD as an indicative parameter of model performance. Furthermore, even though the performance between the two models was the same based on MAD, the RF classification model showed slightly improved performance with other metrics, such as MAE and RMSE, compared to the XGBoost classification model. Hence, the following developed model for all three cities is the RF classification model for energy labels, which are transformed to standardised energy demand based on the values in Table 3.

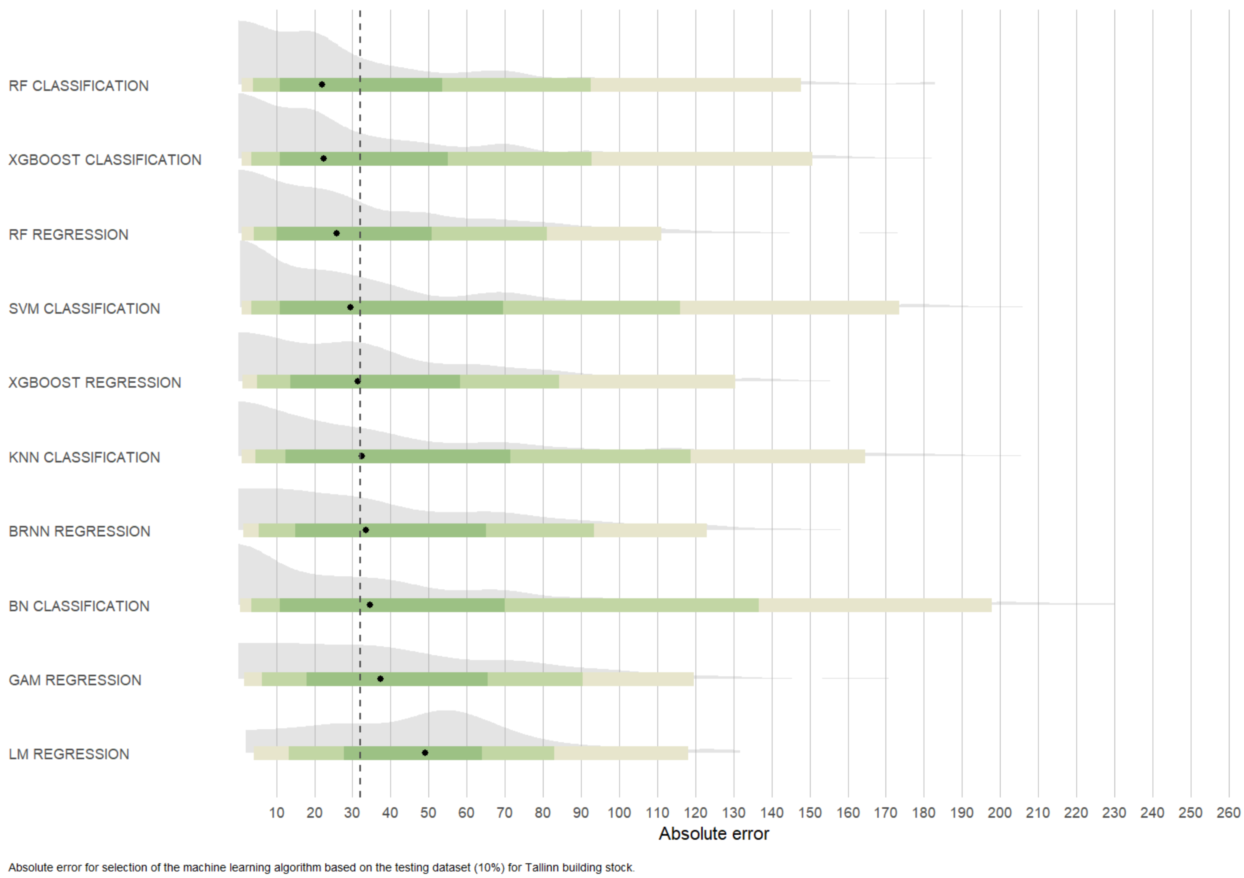

Figure 5. The absolute error distribution curves and error distribution confidence intervals. *Note:* The dot denotes the median absolute error for the predicted energy consumption on testing data. The green colour represents the 50% confidence interval range, pale green represents the 80% confidence interval range, and beige represents 95% of the confidence interval range. The stroked line represents the global median absolute error range between all the models.

4.2. Energy performance prediction

After the model selection process was facilitated, three RF classification models were developed for all the available building data without performing the data split for performance metric calculation. In total, 5944 individual building data points were used for Tallinn, 1192 for Riga, and 802 for Vilnius. After applying SMOTE, the total number of used data points increased to 7139 in Tallinn, 2850 in Riga, and 2061 in Vilnius. *Table 6* summarises the variable importance in every model based on the *varImp* method in the *caret* package (Kuhn, 2024).

Table 6. The variable importance evaluation for the developed RF classification model for building energy label performance in three capital cities.

Tallinn, Estonia		**Riga, Latvia**		**Vilnius, Lithuania**	
YC	1094.73	*YC*	447.90	*Lon*	320.97
Lat	907.04	*Lat*	442.16	S_h	293.23
Lon	890.67	*Lon*	410.93	*LST*	263.52
V	872.16	S_c	385.65	*Lat*	257.93
S	842.15	*LST*	359.82	*YC*	256.32
LST	823.42	*F*	293.04	*S*	247.13
H	632.90				

The modelling frameworks for Tallinn and Riga show the same patterns of variable importance, with year of building construction and location-related variables being the most important. A decrease followed the dimension-related variables in variable importance based on the level of spatial dimension described. In both instances, the building height-related variables were the least important. The mean land surface temperature of a building had a low level of importance in both instances. However, the models for Vilnius had several important changes compared to Riga and Tallinn: the year of building construction had one of the lowest variable importance, whereas the mean land surface temperature value had a higher level of importance. The role of the land surface temperatures in the models reflects the differences in the EPC frameworks, where, only in Lithuania, the building energy label is assigned based on the energy transmission values of different building elements, rather than the energy consumption values.

10% of the training data before the application of SMOTE was used for final tests of model performance. It represented 656 data points for Tallinn, 130 for Riga, and 85 for Vilnius. The results can be observed in Table 7. In the case of Tallinn and Riga, the values of MAE and RMSE show similar performance, with the overall decrease of the values being observed. However, the

MAD value, an initial selection indicator for applying the RF classification algorithm for the developed machine-learning-based predictive model, increased. In the case of the developed model for Vilnius, the MAE and RMSE indicators increased, whereas the MAD indicator decreased drastically. As MAD is the only indicator using the median values rather than the mean, this potentially indicates a low number of substantially misclassified testing data points, drastically increasing the values of MAE and RMSE, which can be attributed to the overall small number of points used for testing purposes.

Table 7. Summary of the performance indicators for the final machine-learning-based models.

Model	MAE	RMSE	MAD	MAPE
Tallinn, Estonia	33.78	53.13	40.78	17.46%
Riga, Latvia	33.29	47.76	41.45	30.80%
Vilnius, Lithuania	52.95	82.52	6.53	1.54%

When comparing the MAPE between the different developed models, a substantial variation is observed between the cities. Overall, the lowest error rate is observed in the model for Vilnius, which follows the MAD results. This deems the approach to be most suitable in the context of building energy efficiency labels associated with the heat transmission values, which is more directly related to the land surface temperatures of buildings. Even though the model performance indicator values for Riga and Tallinn have low variance, the difference in MAPE is drastic due to the differences in the described energy performance thresholds of associated energy labels. In further work, especially on the discussion stage, the results associated with Riga energy efficiency of building modelling will be used more cautiously due to the highest error rate.

Tallinn, Estonia

After applying the RF classification model to the buildings without an energy label in Tallinn, the total number of buildings with it increased from 6677 to 23473, describing 85.28% of the total residential building stock. The mapped results for the level of statistical neighbourhoods of Tallinn are summarised in *Figure 6*. Overall, different neighbourhoods of Kesklinn and Põhja-Tallinn, predominantly situated in Tallinn's inner city, observe the lowest levels of energy efficiency in housing, particularly Vanalinn and Kadriorg in Kesklinn, and Kalamaja, Pelgulinn, and Kopli in Põhja-Tallinn. From the point of view of housing, the following neighbourhoods are represented by higher shares of pre-socialist housing with substantial proportions of newer housing. In case of Kopli, the neighbourhood is predominantly represented by early socialist-era multi-apartment buildings. The energy consumption of residential buildings in other neighbourhoods and districts is lower. However, it is important to note that in the case of Estonia, the energy label and thus the associated energy consumption value used in EPCs describes the overall energy demand of a building, which can potentially indicate the increased electricity consumption due to the mixed use of buildings, particularly in the inner city.

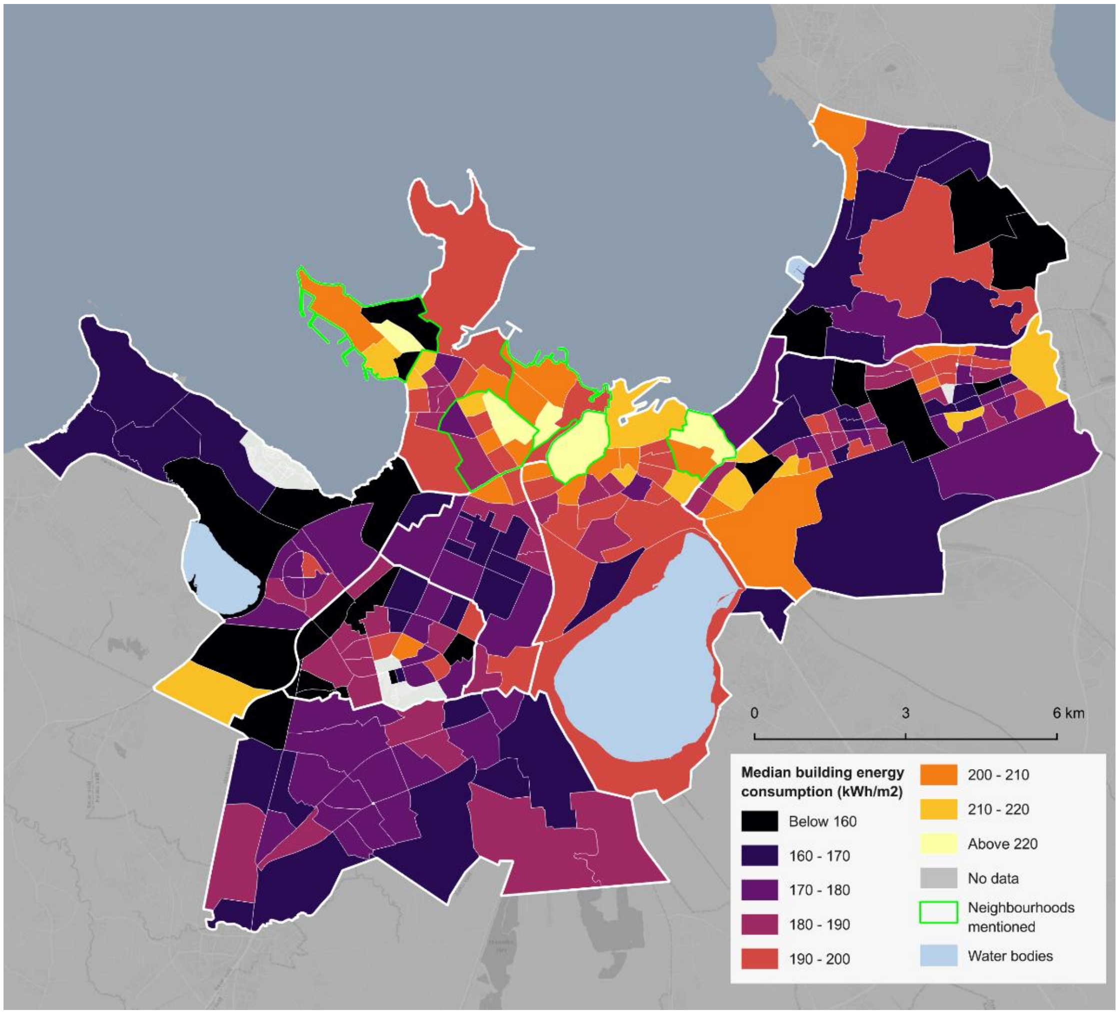

Figure 6. Median building energy consumption based on the existing and modelled data in Tallinn, divided by neighbourhoods.

Riga, Latvia

After using the RF classification algorithm, the total number of buildings with an energy label increased from 1322 to 24616, describing 97.92% of the residential buildings in Riga. The modelling results aggregated to the level of statistical grids are summarised in *Figure 7*. Several neighbourhoods are highlighted when overlaying Riga's neighbourhoods with the grid level data. Grid units of Avoti, Latgale, and Pētersala-Andrejsala neighbourhoods, which are associated predominantly with pre-socialist housing built at the beginning of the twentieth century in the inner city, have a higher median heat consumption. However, typologically, higher values of heat energy consumption were also observed in neighbourhoods with predominantly prefabricated multi-apartment buildings, such as Iļģuciems, Dzirciems, Čiekurkalns, and Teika, which are associated with older socialist-era housing. Furthermore, with the increase of distance from the downtown, the overall energy efficiency of housing observed is improved, compared to the neighbourhoods surrounding it.

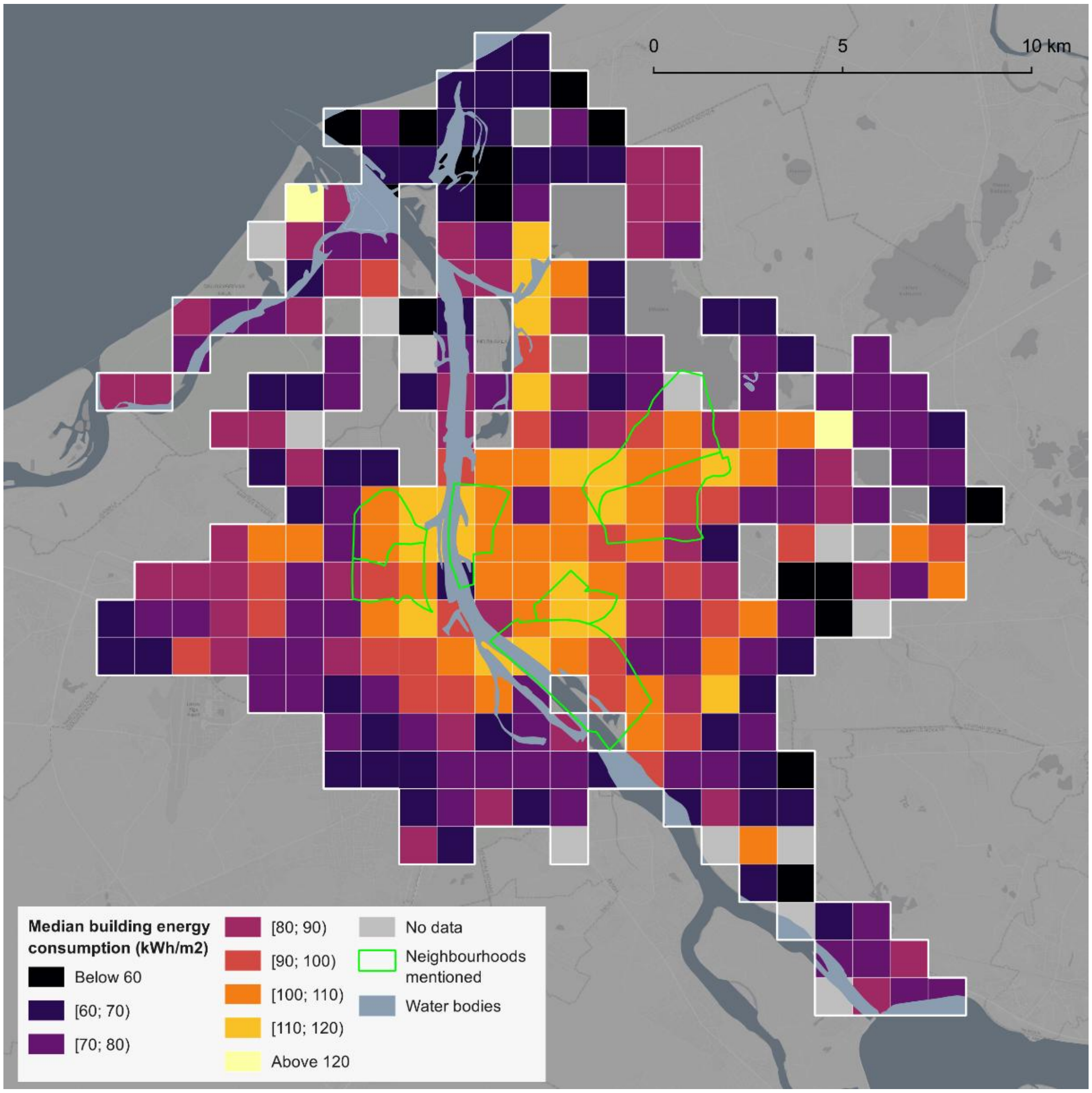

Figure 7. Median building heat energy consumption based on the existing and modelled data in Riga, divided by statistical 1 km^2 grid cells.

Vilnius, Lithuania

For Vilnius, the total number of buildings with an energy label increased from 911 to 3993, representing 49.04% of the residential building stock in the used dataset. *Figure 8* describes the outputs of the machine learning algorithm application for Vilnius. Overall, the neighbourhoods of Karoliniškės eldership, which are characterised by Soviet-era-built housing, have the highest mean building energy consumption. Additionally, high variabilities are observed in the Centras eldership with various residential building typologies. The elderhips of Justiniškės and Pašilaičiai in the east of Vilnius have lower median building energy consumption. These elderships and their neighbourhoods are characterised by various typologies of prefabricated multi-apartment buildings. Due to the low number of classified pre-socialist residential buildings, the direct link between the worse energy performance and pre-socialist housing cannot be observed from mapping the modelling results, as done for Riga and Tallinn.

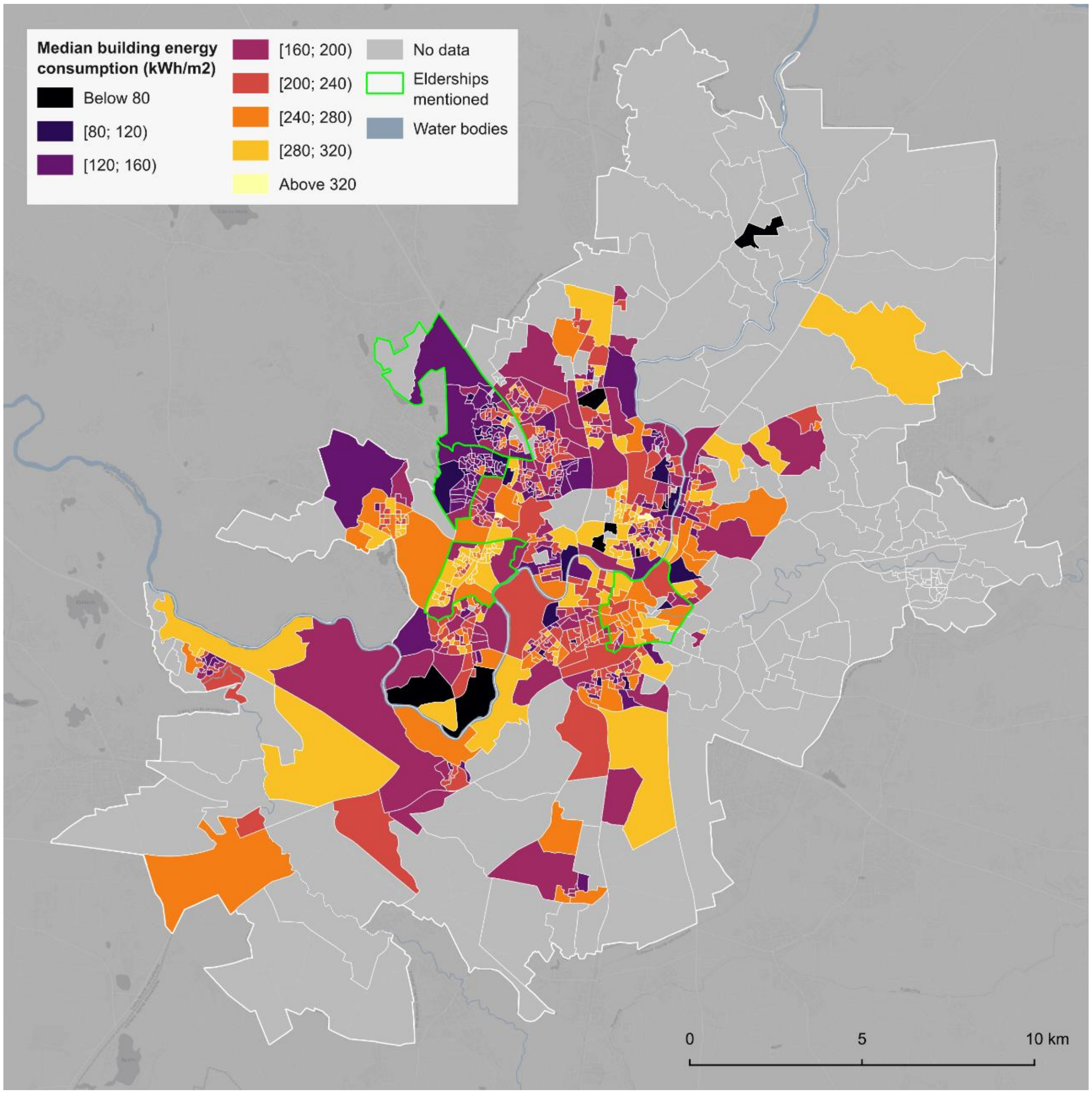

Figure 8. Median building energy consumption based on the existing and modelled data in Vilnius, divided by neighbourhoods.

4.3. Neighbourhood-level analysis

After performing the building energy performance and associated dwelling-level energy costs, the socio-economic, demographic, and ethnic datasets were used to assess the trends associated with the variable energy costs between the different population groups. The following subchapters describe the results per city.

4.3.1. Tallinn, Estonia

Table 7 summarises the results of the multi-level model for Tallinn. Building-level characteristics have several dynamics. Building age positively impacts the building energy consumption per square metre, describing poorer energy performance being observed in older buildings. Overall, multi-apartment buildings are associated with better energy performance compared to detached dwellings. Simultaneously, the total building heated area has a positive impact on the building energy consumption values. The building renovation grants have a positive impact on building energy consumption, potentially being associated with the increased electrification of the building elements, such as ventilation, as well as not considering the changes to energy systems, which reduce the climate impact of the building.

Table 7. The results of the multilevel model for total annual building energy consumption per square metre for Tallinn, Estonia.

	Estimate	Std. Error	p-value	Significance	VIF
Intercept	181.42	1.31	$< 2.00 * 10^{-16}$	***	-
Building age	12.28	0.42	$< 2.00 * 10^{-16}$	***	1.06
Renovation grants (YES = 1)	8.21	2.77	$3.02 * 10^{-3}$	**	1.02
Multi-apartment building (YES = 1)	-4.96	1.03	$1.46 * 10^{-6}$	***	1.73
Total heated area	6.54	0.65	$< 2.00 * 10^{-16}$	***	2.01
Russian-speaking population (n)	-0.07	1.30	0.04	-	3.28
Population with other mother tongues (n)	-0.39	1.19	0.89	-	3.11
Middle occupational groups (n)	-1.67	0.81	0.74	*	1.41
Low occupational groups (n)	-0.23	1.58	0.96	-	4.72
People aged 0-14 (n)	-2.18	1.61	0.17	-	3.75
People above 65 years of age (n)	-2.07	1.40	0.14	-	3.07

Note: Significance codes are denoted as "***" for the p-values between 0 and $1*10^{-3}$; "**" for the p-values between $1*10^{-3}$ and 0.01; "*" for the values between 0.01 and 0.05; '.' for the p-values between 0.05 and 0.10; and '-' for p-values above 0.10.

When considering neighbourhood-level variables, only one variable had a statistically significant impact: middle occupational groups. Compared to high occupational groups, a decrease in building energy consumption per square metre is observed when the proportion of the middle occupational groups increases from the total share of workers on a neighbourhood level. As the inner city neighbourhoods observed the highest median building energy consumption values, the results can be attributed to the overrepresentation of the high occupational groups in the inner city. However, it is important to note the lack of statistical significance of low occupational groups compared to high occupational groups in relation to building energy efficiency. This observation is further important when considering the existing instruments of energy transition. As the two categories do not exude variation, the observed energy efficiency values for the buildings are more homogeneous. Furthermore, the statistically significant variations of energy efficiency levels were not observed based on the ethnic variables used in the study. All the variables are below the set threshold of VIF values, describing the suitable levels of collinearity between the independent variables of the model.

4.3.2. Riga, Latvia

Table 8 refers to the multilevel modelling result for Riga. When considering the multilevel model developed for Riga, the effects of the building-level variables on building energy consumption are similar, with one distinction: impacts of building being classified as a multi-apartment building. In this case, the multi-apartment buildings were associated with higher heat energy consumption per square metre and thus poorer energy performance compared to the detached housing. This observation is also further confirmed by the substantially lower values of median building heat energy consumption in the city outskirt grid cells, where detached housing is more dominant housing typology, compared to the rest of Riga (see *Figure 7*). The root cause of the observation is a substantially higher share of pre-socialist multi-apartment buildings from the total building stock.

Table 8. The results of the multilevel model for annual heat energy consumption per square metre for Riga, Latvia.

	Estimate	Std. Error	p-value	Significance	VIF
Intercept	27.48	1.22	$< 2.00 * 10^{-16}$	***	-
Building age	10.10	0.23	$< 2.00 * 10^{-16}$	***	1.15
Renovation grants (YES = 1)	8.38	2.51	$8.20 * 10^{-4}$	***	1.54
Multi-apartment building (YES = 1)	5.83	0.55	$< 2.00 * 10^{-16}$	***	1.00
Total heated area	20.42	0.46	$< 2.00 * 10^{-16}$	***	1.69
Ethnicity: Russian (n)	-0.52	0.78	0.50	-	1.83
Ethnicity: Other (n)	-0.32	0.53	0.55	-	1.45
Ethnicity: Non-declared (n)	-0.02	0.56	0.98	-	1.66

Middle occupational groups (n)	0.81	0.39	0.04	*	1.24
Low occupational groups (n)	1.79	0.62	$4.29 * 10^{-3}$	**	1.53
People aged 0-14 (n)	-0.99	0.56	0.08	.	2.06
People above 65 years of age (n)	0.84	0.50	0.10	.	1.62

Note: Significance codes are denoted as "***" for the p-values between 0 and $1*10^{-3}$; "**" for the p-values between $1*10^{-3}$ and 0.01; "*" for the values between 0.01 and 0.05; '.' for the p-values between 0.05 and 0.10; and '-' for p-values above 0.10.

When considering the neighbourhood level effects, an association between the occupational status and building energy efficiency levels emerges. Compared to the proportion of the high occupational groups, the proportional increase of either middle or low occupational groups is correlated with an increase in heat energy consumption, indicating poorer energy performance of buildings. Furthermore, between the middle and low occupational groups, the proportional increase of the low occupational groups is associated with a bigger increase in heat energy consumption compared to the middle occupational groups. These observations further strengthen the notions described by Bouzarovski and Simcock (2017) on domestic energy deprivation being strongly associated with household income, as in Riga, more vulnerable socio-economic groups experience worse building energy performance. However, as the predictive model exhibited the highest error values, these results will be considered only in combination with other areas. All the variables are below the set threshold of VIF values, describing the suitable levels of collinearity between the independent variables of the model.

4.3.3. Vilnius, Lithuania

The results of the multilevel model for building energy consumption per square metre for Vilnius are summarised in *Table 9*. On the building level, as with the previous two models, the building age increase is strongly associated with an increase in building energy consumption. However, the dynamics of other building-level variables are different compared to the previous two models. First of all, in the case of the Vilnius model, the differentiation between the types of housing is not statistically significant. Furthermore, as the energy labels in Lithuania, out of which the consumption values were computed, are strongly associated with variations of building element thermal transmission, the renovation grants are associated with a significant decrease in building energy consumption. Finally, the building's heated area is negatively associated with building energy consumption, describing smaller buildings overall to have worse energy performance.

Table 9. The results of the multilevel model for total annual energy consumption per square metre for Vilnius, Lithuania.

	Estimate	Std. Error	p-value	Significance	VIF

Intercept	185.80	23.50	$3.49 * 10^{-15}$	-	-
Building age	20.93	1.85	$< 2.00 * 10^{-16}$	***	1.37
Renovation grants (YES = 1)	-41.72	3.67	$< 2.00 * 10^{-16}$	***	1.02
Multi-apartment building (YES = 1)	30.33	23.50	0.20	-	1.05
Total heated area	-4.91	1.69	$2.78 * 10^{-3}$	**	1.29
Russian-speaking population (n)	-2.04	2.81	0.48	-	2.07
Polish-speaking population (n)	-11.48	3.85	$2.95 * 10^{-3}$	**	3.70
Population with other mother tongues (n)	-8.65	2.54	$7.10 * 10^{-4}$	***	1.79
Middle occupational groups (n)	-0.23	2.09	0.91	-	1.08
Low occupational groups (n)	14.59	4.88	$2.88 * 10^{-3}$	**	5.77
People aged 0-14 (n)	-8.63	3.37	0.01	*	2.83
People above 65 years of age (n)	-8.29	2.86	$3.88 * 10^{-3}$	**	2.27

Note: Significance codes are denoted as "***" for the p-values between 0 and $1*10^{-3}$; "**" for the p-values between $1*10^{-3}$ and 0.01; "*" for the values between 0.01 and 0.05; '.' for the p-values between 0.05 and 0.10; and '-' for p-values above 0.10.

The neighbourhood-level variables describe several trends that potentially impact the energy efficiency of housing. The increase in the proportion of low occupational groups compared to the high occupational groups is associated with higher total annual building energy consumption and thus worse energy performance. This, together with the impacts of building renovation activities, can potentially hinder the overall inaccessibility of climate and energy transition interventions for lower socio-economic groups compared to high socio-economic groups *(*Grossmann, 2019; Umit *et al.*, 2019; Albrecht and Hamels, 2021; Bardazzi *et al.*, 2021). When considering the ethnic neighbourhood variables, an increase in the proportion of people with Polish or languages other than Lithuanian, Russian, and Polish is associated with a lower building-level energy consumption compared to the proportion of people with Lithuanian as their mother tongue. This, however, can be partially attributed to the higher shares of individuals with Lithuanian as their mother tongue living in the Vilnius inner city (Valatka *et al.*, 2016), which, based on Figure 8, has a higher median building energy consumption value.

5. Discussion and conclusions

A set of classification and regression machine-learning algorithms was tested to create a unified modelling framework for building energy performance prediction at the urban level. Working with the existing availability of data, the study identified the random forest classification algorithm to be the most suitable for the development of the modelling framework, primarily due to the lowest median absolute deviation values and error distribution curves. Even as XGBoost for the classification tasks performed with the same median absolute deviation values, its performance based on mean absolute error and root mean squared error was worse by 0.19 and 0.11, respectively. When adapting the modelling framework from Tallinn to Riga and Vilnius, substantial variations of the performance indicators were observed. From the perspective of median absolute deviation and median absolute percentage error, the adapted modelling framework for Vilnius residential building stock performed the best, whereas the values of mean absolute error and root mean squared error were lowest for Riga. Importantly, however, the highest median absolute percentage error and median absolute deviation values were observed for the modelling framework application for Riga. Thus, when considering the following trends and patterns from the multilevel models, the results of Riga were used as accompanying, rather than distinctive, results.

When considering the identified building characteristics of low energy efficiency based on the results of mapping and multilevel models, an overlapping pattern emerges. Building age positively correlates with annual energy consumption, reflecting lower energy efficiency of older building stock. The observation can be seen from the mapping of results for Tallinn and Riga, where the neighbourhoods with higher shares of pre-socialist buildings are also associated with higher median energy consumption values per square meter, as well as the multilevel model results of all three cities, where the building age had a significant positive relationship with building energy consumption values. In the case of Vilnius, the mapping of the existing and predicted energy consumption values did not identify such a relationship, as the neighbourhoods of the worst energy efficiency were quite diverse from the point of view of housing structures. It can potentially be attributed to the underrepresentation of the pre-socialist housing in Vilnius and the associated dataset used in the study. However, from other trends observed in the study, the pre-socialist housing of the inner cities can be seen as the most inefficient from an energy consumption perspective. Furthermore, in all three cities, the neighbourhoods of post-socialist housing developments have the lowest median energy consumption values and consequently better energy efficiency.

The results of the study associated with housing typologies are more heterogeneous across the three cities. In the case of Tallinn, even as some of the neighbourhoods of predominantly detached dwellings, particularly in the case of Haabersti and Pirita, exude the lowest energy performance values, the results of the multilevel model show slightly increased energy performance values overall for the detached dwellings compared to multi-apartment buildings. This potentially can be attributed to the higher median building energy consumption values in the neighbourhoods of Nõmme and the developed pre-socialist detached housing stock. In the case of Riga, however, the multi-apartment buildings are associated with higher energy consumption values and thus, worse energy efficiency, potentially stemming from a substantial share of pre-socialist housing in the multi-apartment building stock. Here, the interactions between the building age and housing type play a stronger role, compared to the housing types individually. Finally, the analysis of the association between housing type and energy consumption in Vilnius did not yield a statistically significant association. Thus, the housing type cannot be considered as an overarching determinant of energy efficiency between all three capital cities, based on substantial differences in the results.

When considering the neighbourhood composition effects, the occupational group distribution on the neighbourhood level had the biggest impact on the housing energy efficiency. Between the cities, in the case of Vilnius and Riga, a common dynamic was established: the increase of the proportion of low occupational groups on the neighbourhood level is correlated with an increase in building energy consumption within the neighbourhood, meaning that lower socio-economic status groups tend to live in the housing with worse energy efficiency. This, furthermore, posits an additional risk associated with an increased susceptibility to energy poverty (Bouzarovski and Simcock, 2017; Barrett et al., 2022), as well as describes the overall inability to participate in the energy transition (Koďousková and Bořuta, 2022). Simultaneously, the observed trends describe an overall association between the higher shares of high occupational groups on the neighbourhood level being associated with better building energy performance. It signifies the overrepresentation of high occupational groups in energy-efficient housing, either through living in newer or renovated housing. From this perspective, the case of Tallinn is an outlier based on the statistically significant variable interactions. Only the proportional increase of middle occupational groups in the neighbourhood has been identified to lower the building-level energy consumption. Here, however, the simultaneous overrepresentation of high occupational groups in the inner city with pre-socialist housing and in the neighbourhoods of newer detached housing can be attributed to the identified heterogeneous patterns. Importantly, no overarching patterns associated with age or ethnic population distribution were observed in the three capital cities, with only in the exception of Vilnius.

Overall, the identified occupational inequalities in access to energy-efficient housing are not addressed in the policy frameworks of energy poverty and energy efficiency. From the perspective of three policy frameworks, the case of Vilnius and Lithuania shows the biggest emphasis on the reduction of inequalities associated with energy efficiency and costs, yet they insufficiently address the minimisation of energy efficiency gaps between the occupational groups. In *Annexe II*, two mechanisms of energy poverty reduction were identified in Lithuania, which were the provision of the heating allowance and the additional support for the beneficiaries of the heating allowance in the facilitation of multi-apartment building renovation projects. However, the impacts of such policies are not observed on the urban level, as the increase of low occupational groups in the neighbourhood was associated with worse building-level energy performance. However, the question of recognition justice has a bigger role in the outcomes associated with the inequality reduction. The definitions of energy poverty are linked to receiving a heating allowance across the countries. For instance, the developed legal definition of energy-poor households in Latvia focuses only on low-income households receiving additional support for household expense coverage (Latvijas Vēstnesis, 2024). Such definitions substantially limit the share of the population eligible for the instruments associated with energy inequality reduction, as, based on the results of the study, much larger shares of the population experience inequalities associated with energy-efficient housing access and subsequent risks of energy poverty stemming from it.

When comparing the approaches, policy frameworks in Lithuania focus on distributive justice facilitation through a set of activities and financial mechanisms of energy poverty reduction. In contrast, within the context of Latvia, the planning documents primarily highlight the need for recognition justice by adoption of a legal definition of energy poverty, with lower emphasis on the instruments associated with distributive energy justice. Only one instrument of support – the facilitation of the renovation of the 2017 dwellings for energy-poor households – was identified to reduce the observed inequalities in Latvia. Here, the case of the Estonian policy framework is an outlier, as, from the point of view of recognition justice, energy poverty was not deemed as a relevant issue, yet simultaneously, distributive justice facilitation mechanisms were identified. Within the national framework of Estonia, the development of the Housing Investment Fund, as envisioned within the Estonian Resilience and Recovery Plan, can potentially enable a more just energy transition to target not only more vulnerable areas but also the vulnerable households more directly. However, between all three countries, the EU-level initiatives and policies addressing energy inequalities are far stronger compared to the national ones. Here, the updated National Building Renovation Plans with stronger inclusion of inequality alleviation measures will provide a better policy action basis for facilitating energy justice, compared to the existing efforts, if addressed by the Member States in a way that is requested within the 2024 recast of EPBD.

Several examples of European policy instruments concerning energy poverty and energy justice alleviation measures exist. A summary overview of some of the policies that have been identified and implemented is provided in *Annexe II*. As the assessment results identified socio-economic status as a determinant of access to energy efficiency measures, providing targeted individual-level support for low-income households is necessary to reduce the inequalities between the groups. In France, for instance, several instruments are already established: low-income households receive energy checks that can be used to either pay for energy bills or renovation activities, the building renovation programme's support is varied by the household income levels, as well as the energy suppliers are obligated to work with energy-poor populations to reduce their energy consumption as the part of the national energy trading system. The identified patterns of inequality require more consideration and potential change of the structures associated with the receipt of building renovation instrument support. Within the context of the three countries, the support is currently provided at the building or association of apartment owners level rather than the household level. Here, more direct support to socio-economically vulnerable households, such as energy checks for energy efficiency improvement, can be used as the energy justice facilitation instrument within the context of all three countries.

The results of the study, furthermore, revealed the need for a stronger focus on pre-socialist housing when considering the designs of the programmes of building renovation due to the biggest potential for energy consumption savings. Here, however, the notions of just transition of this housing segment are even more important, as the potential economic burden of lower socio-economic groups is higher due to the methods of renovation of buildings with architectural value. The observed dynamics must require a design framework of a support instrument to simultaneously target housing and household segments, as the targeting of only the housing segment can result in a substantial overrepresentation of high socio-economic groups due to their presence in pre-socialist housing. This, potentially, can result in the design of renovation subsidies in a way where individual households receive various levels of support based on their socio-economic status, similarly to France (*refer to Annexe II*), as well as the housing segment. The proposed approach simultaneously tackles the gaps observed in access to energy-efficient housing and its distribution between the different population groups with the housing market segmentation from the point of view of energy efficiency.

There are three overall limitations to the presented study. First, the structures of the energy performance certificate and the described energy variables differ between the countries and cities of assessment. Subsequently, if the goal of the work had not been the creation of a unified methodology for building energy performance assessment, the use of different machine learning algorithms per the city of action for the model development could have been performed. Second, the structures of the building registry datasets between the three countries are varied and different variables are used. This resulted in slight variations within the final models based on the data availability. Furthermore, the lack of publicly available building registry data for residential buildings in Vilnius required the use of a dataset, where a very small sample of detached housing was present, thus reducing the ability to facilitate the analysis of all housing segments. Finally, the selected level of aggregation in all the cities was used due to data availability. Even though the multilevel model is a suitable method for statistical analysis and inequality pattern identification, applying dwelling or building-level data could provide additional insights into the variabilities and trends associated with energy inequalities. However, even as a set of limitations is presented, the work created a unified methodology for building energy performance prediction and subsequent application for assessing urban-level energy inequalities in three national contexts.

Patterns of inequalities in housing energy efficiency and links with population risk factors in Tallinn, Riga and Vilnius

Kirils Gončarovs

Summary

The inequalities related to access to energy-efficient housing are not sufficiently addressed within the context of the Baltic countries. To minimise the knowledge gap, the work aimed to identify the underlying population inequalities associated with access to energy-efficient housing based on the energy performance certificate data. However, as the certificate data are scarce, the study methodologically aimed to develop a machine-learning-based modelling framework to predict building energy performance at the urban level. The machine-learning-based modelling framework was applied in the context of open-source data availability on the building stock in Tallinn, Riga, and Vilnius.

To facilitate the development of the modelling framework, ten different classification and regression algorithms were compared based on the set of global indicators for the model for Tallinn, Estonia. The selection process resulted in identifying a random forest classification algorithm for building energy labels as the most suitable solution. The approach was replicated for Vilnius and Riga, and the predicted building energy efficiency datasets were generated. However, the variability of the global metrics associated with the model performance varied substantially across the models, with the best performance being observed for the application of the modelling framework for Vilnius, and the worst performance being observed for Riga.

After the application of the modelling framework for the capital cities, the neighbourhood or grid-level population datasets were applied together with the building-level characteristics to evaluate dwelling-level standardised building energy performance metrics in a series of multilevel models to assess the variations of energy costs based on population variables. The study revealed a strong association between the occupational status and access to energy-efficient housing: both in Riga and Vilnius, a statistically significant correlation between the proportion of low occupational groups on the neighbourhood level and the energy performance of buildings was observed, with higher proportions being associated with worse performance. However, other overlapping patterns of inequalities between the three countries were not observed. The work concluded that the existing inequalities of access to energy-efficient housing within the context of the capitals in the Baltic countries should be primarily understood through the prism of socio-economic and occupational inequalities, which are not sufficiently addressed within the current policy frameworks associated with energy efficiency and energy poverty.

Eluasemete energiatõhususe ebavõrdsuse mustrid ja seosed elanikkonna riskiteguritega Tallinnas, Riias ja Vilniuses

Kirils Gončarovs

Kokkuvõte

Ebavõrdsust, mis on seotud juurdepääsuga energiatõhusale eluasemele, ei ole Balti riikide kontekstis piisavalt käsitletud. Teadmiste puudujäägi vähendamiseks oli töö eesmärk kindlaks teha energiatõhusate eluruumide kättesaadavusega seotud ebavõrdsus, mis põhineb energiamärgise andmetel. Kuna aga sertifikaadiandmeid on vähe, oli uuringu metodoloogiline eesmärk töötada välja masinõppepõhine modelleerimisraamistik, et prognoosida hoonete energiatõhusust linnatasandil. Masinõppel põhinevat modelleerimisraamistikku rakendati Tallinna, Riia ja Vilniuse hoonete kohta kättesaadavate avatud andmete kontekstis.

Modelleerimisraamistiku väljatöötamise hõlbustamiseks võrreldi kümmet erinevat klassifitseerimis- ja regressioonalgoritmi, mis põhinesid Tallinna, Eesti üldiste näitajate andmetel loodud mudelil. Valikuprotsessi tulemusel tuvastati kõige sobivamaks lahenduseks hoonete energiamärgiste jaoks otsustusmetsa (*Random Forest*) klassifitseerimise algoritm. Lähenemist korrati Vilniuse ja Riia puhul ning genereeriti prognoositud hoonete energiatõhususe andmekogumid. Mudelite tulemuslikkusega seotud globaalsete mõõdikute varieeruvus varieerus siiski oluliselt, kusjuures parimad tulemused täheldati Vilniuse puhul modelleerimisraamistiku rakendamisel ja halvimad tulemused Riia puhul.

Pärast mudeli rakendamist pealinnade andmetel kasutati naabruskonna või võrgustiku tasandi elanikkonna andmekogumeid koos hoonete tasandi näitajatega, et hinnata eluruumide tasandil standardiseeritud hoone energiatõhususe näitajaid mitmetasandiliste mudelite abil, et hinnata energiakulude varieerumist sõltuvalt elanikkonna muutujatest. Uuring näitas tugevat seost ameti ja energiatõhusate eluruumide kättesaadavuse vahel: nii Riias kui ka Vilniuses täheldati statistiliselt olulist korrelatsiooni madala kutsestatuse taseme ja hoonete energiatõhususe vahel, kusjuures suurem osakaal oli seotud kehvema energiatõhususega. Muid kattuvaid ebavõrdsuse mustreid kolme riigi vahel siiski ei täheldatud. Töös jõuti järeldusele, et olemasolevat ebavõrdsust energiatõhusate eluasemete kättesaadavuse osas Balti riikide pealinnade kontekstis tuleks mõista eelkõige sotsiaalmajandusliku ja kutsealase ebavõrdsuse kaudu, mida ei ole praegustes energiatõhususe ja energiavaesusega seotud poliitikaraamistikes piisavalt käsitletud.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Associate Professor Anneli Kährik, for supporting me through the entire academic journey at the University of Tartu. I am particularly thankful for your willingness to explore uncharted territory with me throughout this research, ultimately arriving at new insights into the subject matter.

I would also like to thank the review committee for my thesis, Associate Professor Age Poom and MSc. Marta Jemeļjanova, for their insightful comments and support in improving the quality of this work. Their challenging questions and thoughtful suggestions opened a new perspective I did not consider beforehand.

I would also like to thank Dr. Rūta Ubarevičienė from the Lithuanian Centre for Social Sciences and Associate Professor Māris Bērziņš from the University of Latvia for providing crucial data on the population in Rīga and Vilnius. It would be impossible to conduct this comparative research without your support.

The thesis was conducted with the financial support of the Centre of Excellence in Energy Efficiency (TK230U7) project, as well as the Tallinn Strategic Centre. The financing received supported my day-to-day activities and allowed me to achieve substantially better results.

I would like to thank all the ‘K’ in my life for constantly pushing the boundaries of myself and supporting me in this journey of many transitions – personal, professional, and academic. I will cherish your support for all my life.

References

Alawi, O.A., Kamar, H.M. and Yaseen, Z.M. (2024) ‘Optimizing building energy performance predictions: A comparative study of artificial intelligence models’, *Journal of Building Engineering*, 88, p. 109247. Available at: https://doi.org/10.1016/J.JOBE.2024.109247.

Albrecht, J. and Hamels, S. (2021) ‘The financial barrier for renovation investments towards a carbon neutral building stock – An assessment for the Flemish region in Belgium’, *Energy and Buildings*, 248, p. 111177. Available at: https://doi.org/10.1016/J.ENBUILD.2021.111177.

Ali, U. *et al.* (2024) ‘Urban building energy performance prediction and retrofit analysis using data-driven machine learning approach’, *Energy and Buildings*, 303, p. 113768. Available at: https://doi.org/10.1016/J.ENBUILD.2023.113768.

Ambrose, A.R. (2015) ‘Improving energy efficiency in private rented housing: Why don’t landlords act?’, *Indoor and Built Environment*, 24(7), pp. 913–924. Available at: https://doi.org/10.1177/1420326X15598821/FORMAT/EPUB.

Aristondo, O. and Onaindia, E. (2018) ‘Inequality of energy poverty between groups in Spain’, *Energy*, 153, pp. 431–442. Available at: https://doi.org/10.1016/J.ENERGY.2018.04.029.

Awaworyi Churchill, S. and Smyth, R. (2020) ‘Ethnic diversity, energy poverty and the mediating role of trust: Evidence from household panel data for Australia’, *Energy Economics*, 86, p. 104663. Available at: https://doi.org/10.1016/J.ENECO.2020.104663.

Bardazzi, R., Bortolotti, L. and Pazienza, M.G. (2021) ‘To eat and not to heat? Energy poverty and income inequality in Italian regions’, *Energy Research & Social Science*, 73, p. 101946. Available at: https://doi.org/10.1016/J.ERSS.2021.101946.

Barrella, R. *et al.* (2022) ‘The dark side of energy poverty: Who is underconsuming in Spain and why?’, *Energy Research & Social Science*, 86, p. 102428. Available at: https://doi.org/10.1016/J.ERSS.2021.102428.

Barrett, M.; *et al.* (2022) ‘Energy poverty and deprivation in Ireland’. Available at: https://doi.org/10.26504/RS144.

Bates, A. *et al.* (2024) ‘Whose low-carbon future? Community perceptions and expectations on the renewable energy transition in a post-industrial city’, *Energy Research & Social Science*, 118, p. 103781. Available at: https://doi.org/10.1016/J.ERSS.2024.103781.

Bates, D. *et al.* (2015) ‘Fitting Linear Mixed-Effects Models Using lme4’, *Journal of Statistical Software*, 67(1), pp. 1–48. Available at: https://doi.org/10.18637/JSS.V067.I01.

Baumann, Y. *et al.* (2024) ‘Data and Technological Spatial Politics’, in *Dialogues in Data Power*. Bristol University Press, pp. 159–185. Available at: https://doi.org/10.51952/9781529238327.CH008.

Betto, F., Garengo, P. and Lorenzoni, A. (2020) ‘A new measure of Italian hidden energy poverty’, *Energy Policy*, 138, p. 111237. Available at: https://doi.org/10.1016/J.ENPOL.2019.111237.

Blok, A. (2020) ‘Urban green gentrification in an unequal world of climate change’, *Urban Studies*, 57(14), pp. 2803–2816. Available at: https://doi.org/10.1177/0042098019891050.

Boemi, S.N. and Papadopoulos, A.M. (2019) ‘Energy poverty and energy efficiency improvements: A longitudinal approach of the Hellenic households’, *Energy and Buildings*, 197, pp. 242–250. Available at: https://doi.org/10.1016/J.ENBUILD.2019.05.027.

Bórawski, P., Saniuk, S. and Siksnelyte-Butkiene, I. (2022) ‘Combating Energy Poverty in the Face of the COVID-19 Pandemic and the Global Economic Uncertainty’, *Energies 2022, Vol. 15, Page 3649*, 15(10), p. 3649. Available at: https://doi.org/10.3390/EN15103649.

Bouzarovski, S., Frankowski, J. and Tirado Herrero, S. (2018) ‘Low-Carbon Gentrification: When Climate Change Encounters Residential Displacement’, *International Journal of Urban and Regional Research*, 42(5), pp. 845–863. Available at: https://doi.org/10.1111/1468-2427.12634.

Bouzarovski, S. and Petrova, S. (2015) ‘A global perspective on domestic energy deprivation: Overcoming the energy poverty–fuel poverty binary’, *Energy Research & Social Science*, 10, pp. 31–40. Available at: https://doi.org/10.1016/J.ERSS.2015.06.007.

Bouzarovski, S. and Simcock, N. (2017) ‘Spatializing energy justice’, *Energy Policy*, 107, pp. 640–648. Available at: https://doi.org/10.1016/J.ENPOL.2017.03.064.

Bouzarovski, S., Thomson, H. and Cornelis, M. (2021) ‘Confronting Energy Poverty in Europe: A Research and Policy Agenda’, *Energies 2021, Vol. 14, Page 858*, 14(4), p. 858. Available at: https://doi.org/10.3390/EN14040858.

Bouzarovski, S. and Tirado Herrero, S. (2017) ‘Geographies of injustice: the socio-spatial determinants of energy poverty in Poland, the Czech Republic and Hungary’, *Post-Communist Economies*, 29(1), pp. 27–50. Available at: https://doi.org/10.1080/14631377.2016.1242257.

Breiman, L. (2001) ‘Random forests’, *Machine Learning*, 45(1), pp. 5–32. Available at: https://doi.org/10.1023/A:1010933404324/METRICS.

Burneika, D. (2008) ‘POST-SOVIET TRANSFORMATIONS OF URBAN SPACE IN VILNIUS’, *Annales Geographicae*, 41(t).

Burneika, D., Ubarevičienė, R. and Baranuskaitė, A. (2019) ‘Soviet housing estates in Vilnius, Lithuania: Socio-ethnic structure and future(-less?) perspectives’, *Urban Book Series*, pp. 247–270. Available at: https://doi.org/10.1007/978-3-030-23392-1_12.

Chai, T. and Draxler, R.R. (2014) ‘Root mean square error (RMSE) or mean absolute error (MAE)?-Arguments against avoiding RMSE in the literature’, *Geosci. Model Dev*, 7, pp. 1247–1250. Available at: https://doi.org/10.5194/gmd-7-1247-2014.

Chari, A. and Christodoulou, S. (2017) ‘Building energy performance prediction using neural networks’, *Energy Efficiency*, 10(5), pp. 1315–1327. Available at: https://doi.org/10.1007/S12053-017-9524-5/FIGURES/7.

Charlier, D. (2014) ‘Split Incentives and Energy Efficiency: Empirical Analysis and Policy Options’, *Document de travail ART-Dev 2014-07* [Preprint].

Chawla, N. V. *et al.* (2002) ‘SMOTE: Synthetic Minority Over-sampling Technique’, *Journal of Artificial Intelligence Research*, 16, pp. 321–357. Available at: https://doi.org/10.1613/JAIR.953.

Checker, M. (2011) ‘Wiped Out by the “Greenwave”: Environmental Gentrification and the Paradoxical Politics of Urban Sustainability’, *City & Society*, 23(2), pp. 210–229. Available at: https://doi.org/10.1111/J.1548-744X.2011.01063.X.

Chen, T. and Guestrin, C. (2016) ‘XGBoost: A scalable tree boosting system’, *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 13-17-August-2016, pp. 785–794. Available at: https://doi.org/10.1145/2939672.2939785/SUPPL_FILE/KDD2016_CHEN_BOOSTING_SYSTEM_01-ACM.MP4.

Chen, T. and He, T. (2024) ‘xgboost: eXtreme Gradient Boosting’.

Cong, S. *et al.* (2022) ‘Unveiling hidden energy poverty using the energy equity gap’, *Nature Communications 2022 13:1*, 13(1), pp. 1–12. Available at: https://doi.org/10.1038/s41467-022-30146-5.

Costa-Campi, M.T., Jové-Llopis, E. and Trujillo-Baute, E. (2019) ‘Energy poverty in Spain: an income approach analysis’, *Energy Sources, Part B: Economics, Planning, and Policy*, 14(7–9), pp. 327–340. Available at: https://doi.org/10.1080/15567249.2019.1710624.

Council of the European Union (2007) *Presidency conclusions of the Brussels European Council of 8/9 March 2007*. Brussels, Belgium: Council of the European Union. Available at: https://www.consilium.europa.eu/ueDocs/cms_Data/docs/pressData/en/ec/93135.pdf (Accessed: 5 February 2024).

Croon, T.M. *et al.* (2023) ‘Beyond headcount statistics: Exploring the utility of energy poverty gap indices in policy design’, *Energy Policy*, 177, p. 113579. Available at: https://doi.org/10.1016/J.ENPOL.2023.113579.

Deller, D., Turner, G. and Waddams Price, C. (2021) ‘Energy poverty indicators: Inconsistencies, implications and where next?’, *Energy Economics*, 103, p. 105551. Available at: https://doi.org/10.1016/J.ENECO.2021.105551.

Dietterich, T.G. (2000) ‘Ensemble Methods in Machine Learning’, *International workshop on multiple classifier systems*, 1857 LNCS, pp. 1–15. Available at: https://doi.org/10.1007/3-540-45014-9_1.

Directorate of European Commission (2023) *Draft update of Estonia's National Energy and Climate Plan for 2030*. Brussels: European Commission.

Dogan, E. *et al.* (2022) 'Race and energy poverty: Evidence from African-American households', *Energy Economics*, 108, p. 105908. Available at: https://doi.org/10.1016/J.ENECO.2022.105908.

Dong, B., Cao, C. and Lee, S.E. (2005) 'Applying support vector machines to predict building energy consumption in tropical region', *Energy and Buildings*, 37(5), pp. 545–553. Available at: https://doi.org/10.1016/J.ENBUILD.2004.09.009.

Dooling, S. (2009) 'Ecological Gentrification: A Research Agenda Exploring Justice in the City', *International Journal of Urban and Regional Research*, 33(3), pp. 621–639. Available at: https://doi.org/10.1111/J.1468-2427.2009.00860.X.

Drescher, K. and Janzen, B. (2021) 'Determinants, persistence, and dynamics of energy poverty: An empirical assessment using German household survey data', *Energy Economics*, 102, p. 105433. Available at: https://doi.org/10.1016/J.ENECO.2021.105433.

Eisfeld, K. and Seebauer, S. (2022) 'The energy austerity pitfall: Linking hidden energy poverty with self-restriction in household use in Austria', *Energy Research & Social Science*, 84, p. 102427. Available at: https://doi.org/10.1016/J.ERSS.2021.102427.

Energy Poverty Advisory Hub (2022) *Energy Poverty National Indicators: Insights for a more effective measuring - European Commission*. Available at: https://energy-poverty.ec.europa.eu/discover/publications/publications/energy-poverty-national-indicators-insights-more-effective-measuring_en (Accessed: 21 May 2024).

European Parliament, C. of the E.U. (2003) *Directive 2002/91/EC of the European Parliament and of the Council of 16 December 2002 on the energy performance of buildings*. Brussels: European Parliament. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2003.001.01.0065.01.ENG&toc=OJ%3AL%3A2003%3A001%3ATOC (Accessed: 7 February 2024).

European Parliament, C. of the E.U. (2010) *Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings (recast)*. Brussels: European Parliament. Available at: https://eur-lex.europa.eu/legal-content/EN/ALL/;ELX_SESSIONID=FZMjThLLzfxmmMCQGp2Y1s2d3TjwtD8QS3pqdkhXZbwqGwlgY9KN!2064651424?uri=CELEX:32010L0031#ntr4-L_2010153EN.01001301-E0004 (Accessed: 7 February 2024).

European Parliament, C. of the E.U. (2012) *Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012 on energy efficiency, amending Directives 2009/125/EC and 2010/30/EU and repealing Directives 2004/8/EC and 2006/32/EC (Text with EEA relevance)*. Brussels: European Parliament. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02012L0027-20130701&qid=1707484427042 (Accessed: 9 February 2024).

European Parliament, C. of the E.U. (2018a) *Directive (EU) 2018/844 of the European Parliament and of the Council of 30 May 2018 amending Directive 2010/31/EU on the energy performance of buildings and Directive 2012/27/EU on energy efficiency (Text with EEA relevance)*. Brussels: European Parliament. Available at: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=uriserv:OJ.L_.2018.156.01.0075.01.ENG (Accessed: 7 February 2024).

European Parliament, C. of the E.U. (2018b) *Directive (EU) 2018/2002 of the European Parliament and of the Council of 11 December 2018 amending Directive 2012/27/EU on energy efficiency (Text with EEA relevance.)*. Brussels: European Parliament. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2018.328.01.0210.01.ENG (Accessed: 9 February 2024).

European Parliament, C. of the E.U. (2023a) *Directive (EU) 2023/1791 of the European Parliament and of the Council of 13 September 2023 on energy efficiency and amending Regulation (EU) 2023/955 (recast) (Text with EEA relevance)*. Brussels: European Parliament. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ%3AJOL_2023_231_R_0001&qid=1695186598766 (Accessed: 9 February 2024).

European Parliament, C. of the E.U. (2023b) *Regulation (EU) 2023/955 of the European Parliament and of the Council of 10 May 2023 establishing a Social Climate Fund and amending Regulation (EU) 2021/1060*. Strastbourg: European Parliament. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32023R0955 (Accessed: 19 February 2025).

European Parliament, C. of the E.U. (2024) *Directive (EU) 2024/1275 of the European Parliament and of the Council of 24 April 2024 on the energy performance of buildings (recast) (Text with EEA relevance)*. Strastbourg: European Parliament. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401275&pk_keyword=Energy&pk_content=Directive (Accessed: 19 February 2025).

Feenstra, M. *et al.* (2021) 'Humanising the Energy Transition: Towards a National Policy on Energy Poverty in the Netherlands', *Frontiers in Sustainable Cities*, 3, p. 645624. Available at: https://doi.org/10.3389/FRSC.2021.645624/BIBTEX.

Frantál, B. and Dvořák, P. (2022) 'Reducing energy poverty in deprived regions or supporting new developments in metropolitan suburbs? Regional differences in the use of subsidies for home energy efficiency renovations', *Energy Policy*, 171, p. 113250. Available at: https://doi.org/10.1016/J.ENPOL.2022.113250.

Freimane, S. (2020) ‘Inclusion and Equality in Contemporary Urban Housing: The Case of Riga’, *European Journal of Sustainable Development*, 9(3), pp. 27–27. Available at: https://doi.org/10.14207/EJSD.2020.V9N3P27.

Friedman, N., Geiger, D. and Goldszmidt, M. (1997) ‘Bayesian Network Classifiers’, *Machine Learning*, 29(2–3), pp. 131–163. Available at: https://doi.org/10.1023/A:1007465528199/METRICS.

Gianey, H.K. and Choudhary, R. (2017) ‘Comprehensive Review On Supervised Machine Learning Algorithms’, *Proceedings - 2017 International Conference on Machine Learning and Data Science, MLDS 2017*, 2018-January, pp. 38–43. Available at: https://doi.org/10.1109/MLDS.2017.11.

Goldemberg, J. *et al.* (1987) *Energy For A Sustainable World*.

Gould, K.A. and Lewis, T.L. (2021) ‘Resilience Gentrification: Environmental Privilege in an Age of Coastal Climate Disasters’, *Frontiers in Sustainable Cities*, 3, p. 687670. Available at: https://doi.org/10.3389/FRSC.2021.687670/BIBTEX.

Government of the Republic of Lithuania (2021) *LONG-TERM RENOVATION STRATEGY OF LITHUANIA*. Vilnius: Government of the Republic of Lithuania.

Greenland, S. (2000) ‘Principles of multilevel modelling’, *International Journal of Epidemiology*, 29(1), pp. 158–167. Available at: https://doi.org/10.1093/IJE/29.1.158.

Grossmann, K. (2019) ‘Energy efficiency for whom? : a conceptual view on retrofitting, residential segregation and the housing market’, *Sociologia urbana e rurale : XLI, 119, 2019*, (119), pp. 78–95. Available at: https://doi.org/10.3280/SUR2019-119006.

Grossmann, K. and Huning, S. (2016) ‘Displacement through energy-efficient retrofitting? Impacts, actors and policies’, *Contested cities. Congreso Internacional Madrid 2016*, pp. 1–12. Available at: http://www.zeit.de/news/2014-03/22/wohnen-mieterbundvermieter-nutzen-sanierung-immer-oefter-als-waffe- (Accessed: 27 February 2025).

Guzmán-Rosas, S.C. (2022) ‘Ethnicity as a social determinant of energy poverty: the case of Mexican indigenous population’, *Local Environment*, 27(9), pp. 1075–1101. Available at: https://doi.org/10.1080/13549839.2022.2100879.

Halkos, G. and Kostakis, I. (2023) ‘Exploring the persistence and transience of energy poverty: evidence from a Greek household survey’, *Energy Efficiency*, 16(6), pp. 1–18. Available at: https://doi.org/10.1007/S12053-023-10137-1/FIGURES/1.

Hastie, T. and Tibshirani, R. (1986) ‘Generalized Additive Models’, *https://doi.org/10.1214/ss/1177013604*, 1(3), pp. 297–310. Available at: https://doi.org/10.1214/SS/1177013604.

Herrero, S.T. (2017) ‘Energy poverty indicators: A critical review of methods’, *Indoor and Built Environment*, 26(7), pp. 1018–1031. Available at: https://doi.org/10.1177/1420326X17718054/ASSET/IMAGES/LARGE/10.1177_1420326X17718054-FIG2.JPEG.

Hijmans, R.J. (2025) ‘Package “terra”’. Available at: https://rspatial.github.io/terra/ (Accessed: 11 March 2025).

Hills, J. (2012) ‘Getting the measure of fuel poverty: final report of the Fuel Poverty Review’. Available at: http://sticerd.lse.ac.uk/case/ (Accessed: 12 May 2024).

Hodson, T.O. (2022) ‘Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not’, *Geoscientific Model Development*, 15(14), pp. 5481–5487. Available at: https://doi.org/10.5194/GMD-15-5481-2022.

Igawa, M. and Managi, S. (2022) ‘Energy poverty and income inequality: An economic analysis of 37 countries’, *Applied Energy*, 306, p. 118076. Available at: https://doi.org/10.1016/J.APENERGY.2021.118076.

International Energy Agency (2007) *Mind the gap: Quantifying principal-agent problems in energy efficiency*. Organisation for Economic Cooperation and Development (OECD). Available at: https://doi.org/10.1787/9789264038950-EN.

Jenkins, K. *et al.* (2016) ‘Energy justice: A conceptual review’, *Energy Research & Social Science*, 11, pp. 174–182. Available at: https://doi.org/10.1016/J.ERSS.2015.10.004.

Jordan, M.I. and Mitchell, T.M. (2015) ‘Machine learning: Trends, perspectives, and prospects’, *Science*, 349(6245), pp. 255–260. Available at: https://doi.org/10.1126/SCIENCE.AAA8415/ASSET/AB2EF18A-576D-464D-B1B6-1301159EE29A/ASSETS/GRAPHIC/349_255_F5.JPEG.

Kahouli, S. and Okushima, S. (2021) ‘Regional energy poverty reevaluated: A direct measurement approach applied to France and Japan’, *Energy Economics*, 102, p. 105491. Available at: https://doi.org/10.1016/J.ENECO.2021.105491.

Kalinowski, S. *et al.* (2024) ‘To heat or not to heat? On (in)perceptible energy poverty among the poor’, *Annals of the Polish Association of Agricultural and Agribusiness Economists*, XXVI(1), pp. 124–138. Available at: https://doi.org/10.5604/01.3001.0054.3270.

Karatzoglou, A. *et al.* (2004) ‘kernlab - An S4 Package for Kernel Methods in R’, *Journal of Statistical Software*, 11, pp. 1–20. Available at: https://doi.org/10.18637/JSS.V011.I09.

Karpinska, L. and Śmiech, S. (2020a) ‘Conceptualising housing costs: The hidden face of energy poverty in Poland’, *Energy Policy*, 147, p. 111819. Available at: https://doi.org/10.1016/J.ENPOL.2020.111819.

Karpinska, L. and Śmiech, S. (2020b) 'Invisible energy poverty? Analysing housing costs in Central and Eastern Europe', *Energy Research & Social Science*, 70, p. 101670. Available at: https://doi.org/10.1016/J.ERSS.2020.101670.

Karpinska, L. and Śmiech, S. (2021) 'Breaking the cycle of energy poverty. Will Poland make it?', *Energy Economics*, 94, p. 105063. Available at: https://doi.org/10.1016/J.ENECO.2020.105063.

Karpinska, L. and Śmiech, S. (2023) 'Multiple faces of poverty. Exploring housing-costs-induced energy poverty in Central and Eastern Europe', *Energy Research & Social Science*, 105, p. 103273. Available at: https://doi.org/10.1016/J.ERSS.2023.103273.

Kecman, V. (2005) 'Support Vector Machines – An Introduction', in V. Kecman (ed.) *Support vector machines–an introduction*. Berlin: Springer, Berlin, Heidelberg, pp. 1–47. Available at: https://doi.org/10.1007/10984697_1.

Kilis, V. *et al.* (2025) 'Optimization of Renewable-Based Multi-Energy Systems in Residential Building Design', *Energies*, 18(6). Available at: https://doi.org/10.3390/EN18061541.

Koďousková, H. and Bořuta, D. (2022) 'Energy poverty in Slovakia: Officially defined, but misrepresented in major policies', *Energy Policy*, 168, p. 113095. Available at: https://doi.org/10.1016/J.ENPOL.2022.113095.

Krišjāne, Z. and Bērziņš, M. (2014) 'Intra-urban residential differentiation in the post-Soviet city: the case of Riga, Latvia', *Hungarian Geographical Bulletin*, 63(3), pp. 235–253. Available at: https://doi.org/10.15201/HUNGEOBULL.63.3.1.

Kuhn, M. (2024) 'Package "caret"'. Available at: https://github.com/topepo/caret/ (Accessed: 11 March 2025).

Latvijas Republikas Ekonomikas Ministrija (2019) *4. pielikums Latvijas Nationālais enerģētikas un klimata plānām 2030. gadam*. Rīga: Latvijas Republikas Ekonomikas Ministrija.

Latvijas Republikas Klimata un enerģētikas ministrija (2024) *Latvijas Nacionālais enerģētikas un klimata plāns 2021.-2030. gadam*. Rīga. Available at: https://www.kem.gov.lv/lv/nacionalais-energetikas-un-klimata-plans-2021-2030-gadam (Accessed: 2 March 2025).

Latvijas Vēstnesis (2024) *Enerģētikas likums*. Rīga: Latvijas Republikas Saeimas un Ministru Kabineta Ziņotājs. Available at: https://likumi.lv/ta/id/49833-energetikas-likums (Accessed: 28 February 2024).

Liaw, A. and Wiener, M. (2002) 'Classification and Regression by randomForest', 2(3). Available at: http://www.stat.berkeley.edu/ (Accessed: 24 February 2025).

Libertson, F., Velkova, J. and Palm, J. (2021) 'Data-center infrastructure and energy gentrification: perspectives from Sweden', *Sustainability: Science, Practice and Policy*, 17(1), pp. 153–162. Available at: https://doi.org/10.1080/15487733.2021.1901428.

Liddell, C. *et al.* (2012) ‘Measuring and monitoring fuel poverty in the UK: National and regional perspectives’, *Energy Policy*, 49, pp. 27–32. Available at: https://doi.org/10.1016/J.ENPOL.2012.02.029.

Lihtmaa, L., Hess, D.B. and Leetmaa, K. (2018) ‘Intersection of the global climate agenda with regional development: Unequal distribution of energy efficiency-based renovation subsidies for apartment buildings’, *Energy Policy*, 119, pp. 327–338. Available at: https://doi.org/10.1016/J.ENPOL.2018.04.013.

Listo, R. (2018) ‘Gender myths in energy poverty literature: A Critical Discourse Analysis’, *Energy Research & Social Science*, 38, pp. 9–18. Available at: https://doi.org/10.1016/J.ERSS.2018.01.010.

López-Vargas, A. and Ledezma-Espino, A. (2023) ‘Iot application for energy poverty detection based on thermal comfort monitoring’, *Heliyon*, 9(1), pp. 2405–8440. Available at: https://doi.org/10.1016/j.heliyon.2023.e12943.

Lüdecke, D. *et al.* (2021) ‘performance: An R Package for Assessment, Comparison and Testing of Statistical Models’, *Journal of Open Source Software*, 6(60), p. 3139. Available at: https://doi.org/10.21105/JOSS.03139.

Lulle, A. (2024) ‘(Post-)Socialist Housing and Aging in Neoliberal Riga’, *Urban Planning*, 9. Available at: https://doi.org/10.17645/UP.7705.

Maier, S. and Dreoni, I. (2024) ‘Who is “energy poor” in the EU?’, *JRC Working Papers on Taxation and Structural Reforms* [Preprint]. Available at: https://www.econstor.eu/handle/10419/306591 (Accessed: 12 February 2025).

Majandus- Ja Kommunikatsiooni Ministeerium (2020) *Long-term strategy for building renovation*. Tallinn: Majandus- Ja Kommunikatsiooni Ministeerium. Available at: https://www.kul.ee/sites/kulminn/files/lisa_3_-_kvaliteetne_ruum_aluspohimotted.pdf (Accessed: 18 February 2024).

Majandus- ja taristuministri (2019) *Majandus- ja taristuministri 30.04.2015. a määrus nr 36 „Nõuded energiamärgise andmisele ja energiamärgisele“ Lisa 3: Kaalutud energiaerikasutuse ja energiatõhususarvu klassi määramine*. Available at: https://www.riigiteataja.ee/aktilisa/1190/3202/4003/Lisa_3.pdf# (Accessed: 11 March 2025).

Makridou, G., Matsumoto, K. and Doumpos, M. (2024) ‘Evaluating the energy poverty in the EU countries’, *Energy Economics*, 140, p. 108020. Available at: https://doi.org/10.1016/J.ENECO.2024.108020.

Maloutas, T. and Karadimitriou, N. (2022) ‘Vertical Cities: Micro-segregation, social mix and urban housing markets’, *Vertical Cities: Micro-segregation, Social Mix and Urban Housing Markets*, pp. 1–357. Available at: https://doi.org/10.4337/9781800886391.

McCauley, D.A. *et al.* (2014) ‘Advancing Energy Justice: The Triumvirate of Tenets’, *International Energy Law Review*, 32(3), pp. 107–110. Available at: http://dspace.stir.ac.uk/handle/1893/18349 (Accessed: 20 May 2024).

Meyer, S. *et al.* (2018) ‘Capturing the multifaceted nature of energy poverty: Lessons from Belgium’, *Energy Research & Social Science*, 40, pp. 273–283. Available at: https://doi.org/10.1016/J.ERSS.2018.01.017.

Mikailiene, Ž. (2010) ‘Soviet Vilnius: Ideology and the Formation of Identity’, *Lithuanian Historical Studies*, 15(1), pp. 171–189. Available at: https://doi.org/10.30965/25386565-01501010.

Ministru Kabinets (2019) *Latvijas nacionālais enerģētikas un klimata plāns*. Rīga: Ministru Kabinets.

Ministru kabinets (2020) *Ēku atjaunošanas ilgtermiņa stratēģija*. Rīga: Ministru kabinets.

Ministru kabinets (2021) *Ministru kabineta 2021. gada 8. aprīļa noteikumi Nr. 222 ‘Ēku energoefektivitātes aprēķina metodes un ēku energosertifikācijas noteikumi’*. Rīga: Ministru kabinets. Available at: https://likumi.lv/ta/id/322436-eku-energoefektivitates-aprekina-metodes-un-eku-energosertifikacijas-noteikumi (Accessed: 11 March 2025).

Mulder, P., Dalla Longa, F. and Straver, K. (2023) ‘Energy poverty in the Netherlands at the national and local level: A multi-dimensional spatial analysis’, *Energy Research & Social Science*, 96, p. 102892. Available at: https://doi.org/10.1016/J.ERSS.2022.102892.

Nguyen, C.P. and Su, T.D. (2021) ‘Does energy poverty matter for gender inequality? Global evidence’, *Energy for Sustainable Development*, 64, pp. 35–45. Available at: https://doi.org/10.1016/J.ESD.2021.07.003.

Nicol, J.F. *et al.* (1999) ‘Climatic variations in comfortable temperatures: the Pakistan projects’, *Energy and Buildings*, 30(3), pp. 261–279. Available at: https://doi.org/10.1016/S0378-7788(99)00011-0.

Nicol, J.F. and Roaf, S. (2017) ‘Rethinking thermal comfort’, *Building Research & Information*, 45(7), pp. 711–716. Available at: https://doi.org/10.1080/09613218.2017.1301698.

Ntaintasis, E., Mirasgedis, S. and Tourkolias, C. (2019) ‘Comparing different methodological approaches for measuring energy poverty: Evidence from a survey in the region of Attika, Greece’, *Energy Policy*, 125, pp. 160–169. Available at: https://doi.org/10.1016/J.ENPOL.2018.10.048.

O’Brien, R.M. (2007) ‘A caution regarding rules of thumb for variance inflation factors’, *Quality and Quantity*, 41(5), pp. 673–690. Available at: https://doi.org/10.1007/S11135-006-9018-6/METRICS.

Oliveira Panão, M.J.N. (2021) ‘Lessons learnt from using energy poverty expenditure-based indicators in a mild winter climate’, *Energy and Buildings*, 242, p. 110936. Available at: https://doi.org/10.1016/J.ENBUILD.2021.110936.

Onačillová, K. *et al.* (2022) ‘Combining Landsat 8 and Sentinel-2 Data in Google Earth Engine to Derive Higher Resolution Land Surface Temperature Maps in Urban Environment’, *Remote Sensing 2022, Vol. 14, Page 4076*, 14(16), p. 4076. Available at: https://doi.org/10.3390/RS14164076.

Ormandy, D. and Ronique Ezratty, V. (2011) ‘Health and thermal comfort: From WHO guidance to housing strategies’. Available at: https://doi.org/10.1016/j.enpol.2011.09.003.

Pastak, I. and Kährik, A. (2021) ‘SYMBOLIC DISPLACEMENT REVISITED: Place-making Narratives in Gentrifying Neighbourhoods of Tallinn’, *International Journal of Urban and Regional Research*, 45(5), pp. 814–834. Available at: https://doi.org/10.1111/1468-2427.13054.

Petrov, I. and Ryan, L. (2021) ‘The landlord-tenant problem and energy efficiency in the residential rental market’, *Energy Policy*, 157, p. 112458. Available at: https://doi.org/10.1016/J.ENPOL.2021.112458.

Petrova, S. and Simcock, N. (2021) ‘Gender and energy: domestic inequities reconsidered’, *Social & Cultural Geography*, 22(6), pp. 849–867. Available at: https://doi.org/10.1080/14649365.2019.1645200.

Pham, A.D. *et al.* (2020) ‘Predicting energy consumption in multiple buildings using machine learning for improving energy efficiency and sustainability’, *Journal of Cleaner Production*, 260, p. 121082. Available at: https://doi.org/10.1016/J.JCLEPRO.2020.121082.

Piekut, M. (2020) ‘Patterns of Energy Consumption in Polish One-Person Households’, *Energies 2020, Vol. 13, Page 5699*, 13(21), p. 5699. Available at: https://doi.org/10.3390/EN13215699.

von Platten, J. *et al.* (2021) ‘Renovating on Unequal Premises: A Normative Framework for a Just Renovation Wave in Swedish Multifamily Housing’, *Energies 2021, Vol. 14, Page 6054*, 14(19), p. 6054. Available at: https://doi.org/10.3390/EN14196054.

Polanska, D. V. *et al.* (2024) ‘Predatory commodification and housing renovation’, *Journal of Urban Affairs*, 46(4), pp. 682–700. Available at: https://doi.org/10.1080/07352166.2022.2088375.

Polanska, D. V and Richard, Å. (2021) ‘Resisting renovictions: Tenants organizing against housing companies’ renewal practices in Sweden Introduction: Tenants as active subjects’, *Radical Housing Journal*, 3(1), p. 188. Available at: www.dominikavpolanska.se (Accessed: 27 February 2025).

Quinton, J. and Nesbitt, L. (2024) ‘Different names for the same thing? A systematic review of green, environmental, eco-, ecological, climate, carbon, and resilience gentrification’, *Cities*, 151, p. 105107. Available at: https://doi.org/10.1016/J.CITIES.2024.105107.

R Team (2014) ‘R: A language and environment for statistical computing.’, *MSOR connections* [Preprint].

Reames, T.G. (2016) ‘Targeting energy justice: Exploring spatial, racial/ethnic and socio-economic disparities in urban residential heating energy efficiency’, *Energy Policy*, 97, pp. 549–558. Available at: https://doi.org/10.1016/J.ENPOL.2016.07.048.

Reddy, A.K.N. *et al.* (2000) ‘Energy and Social Issues’, in *World Energy Assessment*, pp. 39–60.

Rice, J.L. *et al.* (2020) ‘Contradictions of the Climate-Friendly City: New Perspectives on Eco-Gentrification and Housing Justice’, *International Journal of Urban and Regional Research*, 44(1), pp. 145–165. Available at: https://doi.org/10.1111/1468-2427.12740.

Riigi Tugiteenuste Keskus (2023) *TAASTE- ja VASTUPIDAVUSKAVA*. Tallinn: Riigi Tugiteenuste Keskus. Available at: https://pilv.rtk.ee/s/kwzABsf9cJY4mmb (Accessed: 24 February 2024).

Riva, M. *et al.* (2021) ‘Energy poverty in Canada: Prevalence, social and spatial distribution, and implications for research and policy’, *Energy Research & Social Science*, 81, p. 102237. Available at: https://doi.org/10.1016/J.ERSS.2021.102237.

Riva, M. *et al.* (2024) ‘Energy poverty and well-being at the local level: Insights from a community-wide survey in Atlantic Canada’, *Energy Research & Social Science*, 117, p. 103709. Available at: https://doi.org/10.1016/J.ERSS.2024.103709.

Roberts, D., Vera-Toscano, E. and Phimister, E. (2015) ‘Fuel poverty in the UK: Is there a difference between rural and urban areas?’ Available at: https://doi.org/10.1016/j.enpol.2015.08.034.

Robinson, C. (2019) ‘Energy poverty and gender in England: A spatial perspective’, *Geoforum*, 104, pp. 222–233. Available at: https://doi.org/10.1016/J.GEOFORUM.2019.05.001.

Rodriguez, P., Gianola, D. and Paulino, M. (2023) ‘Bayesian Regularization for Feed-Forward Neural Networks’.

Rodriguez-Alvarez, A., Llorca, M. and Jamasb, T. (2021) ‘Alleviating energy poverty in Europe: Front-runners and laggards’, *Energy Economics*, 103, p. 105575. Available at: https://doi.org/10.1016/J.ENECO.2021.105575.

Ruggieri, G., Maduta, C. and Melica, G. (2023) ‘Progress on the implementation of Energy Performance Certificates in EU’, *JRC TECHNICAL REPORT*, 31730 EN. Available at: https://doi.org/10.2760/522233.

Ruoppila, S. (2007) ‘Establishing a Market-orientated Urban Planning System after State Socialism: The Case of Tallinn’, *European Planning Studies*, 15(3), pp. 405–427. Available at: https://doi.org/10.1080/09654310601017117.

Ruoppila, S. and Kährik, A. (2003) ‘Socio-economic residential differentiation in post-socialist Tallinn’, *Journal of Housing and the Built Environment*, 18(1), pp. 49–73. Available at: https://doi.org/10.1023/A:1022435000258/METRICS.

Samalavičius, A., Gabrėnas, A. and Gabrėnienė, A. (2024) ‘The new center of Vilnius: spatial transformation and challenges of genius loci’, *Journal of Baltic Studies*, 55(2), pp. 291–308. Available at: https://doi.org/10.1080/01629778.2023.2213209/ASSET/F9827FF9-0827-4A52-A2C2-3D81D78E5B08/ASSETS/GRAPHIC/RBAL_A_2213209_F0007_OC.JPG.

Samardzioska, T. *et al.* (2021) ‘Prediction of Energy Consumption in Buildings Using Support Vector Machine’, *Tehnički vjesnik*, 28(2), pp. 649–656. Available at: https://doi.org/10.17559/TV-20190822153751.

Seebauer, S., Friesenecker, M. and Eisfeld, K. (2019) ‘Integrating climate and social housing policy to alleviate energy poverty: an analysis of targets and instruments in Austria’, *Energy Sources, Part B: Economics, Planning, and Policy*, 14(7–9), pp. 304–326. Available at: https://doi.org/10.1080/15567249.2019.1693665.

Seyedzadeh, S. *et al.* (2018) ‘Machine learning for estimation of building energy consumption and performance: a review’, *Visualization in Engineering 2018 6:1*, 6(1), pp. 1–20. Available at: https://doi.org/10.1186/S40327-018-0064-7.

Seyedzadeh, S. *et al.* (2020) ‘Machine learning modelling for predicting non-domestic buildings energy performance: A model to support deep energy retrofit decision-making’, *Applied Energy*, 279, p. 115908. Available at: https://doi.org/10.1016/J.APENERGY.2020.115908.

Simionescu, M. *et al.* (2023) ‘The role of renewable energy policies in TACKLING energy poverty in the European UNION’, *Energy Policy*, 183, p. 113826. Available at: https://doi.org/10.1016/J.ENPOL.2023.113826.

Siriseriwan, W. (2024) ‘Package “smotefamily”’.

Śmiech, S., Karpinska, L. and Bouzarovski, S. (2025) ‘Impact of energy transitions on energy poverty in the European Union’, *Renewable and Sustainable Energy Reviews*, 211, p. 115311. Available at: https://doi.org/10.1016/J.RSER.2024.115311.

Stojilovska, A., Thomson, H. and Mejía-Montero, A. (2023) ‘Making a Case for Centring Energy Poverty in Social Policy in Light of the Climate Emergency: A Global Integrative Review’, *Social Policy and Society*, 22(4), pp. 715–729. Available at: https://doi.org/10.1017/S1474746423000209.

Stojilovska, A., Yoon, H. and Robert, C. (2021) ‘Out of the margins, into the light: Exploring energy poverty and household coping strategies in Austria, North Macedonia, France, and Spain’, *Energy Research & Social Science*, 82, p. 102279. Available at: https://doi.org/10.1016/J.ERSS.2021.102279.

Taltavull de La Paz, P. *et al.* (2022) ‘Sources of Energy Poverty: A Factor Analysis Approach for Spain’, *Frontiers in Energy Research*, 10, p. 847845. Available at: https://doi.org/10.3389/FENRG.2022.847845/BIBTEX.

Tammaru, T. *et al.* (2015) ‘The “market experiment”: Increasing socio-economic segregation in the inherited bi-ethnic context of Tallinn’, *Socio-Economic Segregation in European Capital Cities – East Meets West*, (683680), pp. 333–357. Available at: https://doi.org/10.4324/9781315758879-22.

The Ministry of Energy of the Republic of Lithuania (2023) *ON THE DRAFT UPDATE OF THE NATIONAL ENERGY AND CLIMATE PLAN OF THE REPUBLIC OF LITHUANIA 2021-2030*. Vilnius: The Ministry of Energy of the Republic of Lithuania.

Thomson, H., Bouzarovski, S. and Snell, C. (2017) ‘Rethinking the measurement of energy poverty in Europe: A critical analysis of indicators and data’, *Indoor and Built Environment*, 26(7), pp. 879–901. Available at: https://doi.org/10.1177/1420326X17699260.

Tirado Herrero, S. and Ürge-Vorsatz, D. (2012) ‘Trapped in the heat: A post-communist type of fuel poverty’, *Energy Policy*, 49, pp. 60–68. Available at: https://doi.org/10.1016/J.ENPOL.2011.08.067.

Tsanas, A. and Xifara, A. (2012) ‘Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools’, *Energy and Buildings*, 49, pp. 560–567. Available at: https://doi.org/10.1016/J.ENBUILD.2012.03.003.

Tsoka, T. *et al.* (2022) ‘Explainable artificial intelligence for building energy performance certificate labelling classification’, *Journal of Cleaner Production*, 355, p. 131626. Available at: https://doi.org/10.1016/J.JCLEPRO.2022.131626.

Turcu, C. (2017) ‘Unequal spatial distribution of retrofits in Bucharest’s apartment buildings’, *Building Research & Information*, 45(8), pp. 892–909. Available at: https://doi.org/10.1080/09613218.2016.1229894.

Ubarevičienė, R. and Burneika, D. (2020) ‘Environmental & Socio-economic Studies Fast and uncoordinated suburbanization of Vilnius in the context of depopulation in Lithuania’, *Environ. Socio.-econ. Stud*, 8, pp. 44–56. Available at: https://doi.org/10.2478/environ-2020-0022.

Umit, R. *et al.* (2019) ‘The role of income in energy efficiency and curtailment behaviours: Findings from 22 European countries’, *Energy Research & Social Science*, 53, pp. 206–214. Available at: https://doi.org/10.1016/J.ERSS.2019.02.025.

Vapnik, V., Golowich, S. and Smola, A. (1996) ‘Support Vector Method for Function Approximation, Regression Estimation and Signal Processing’, *Advances in Neural Information Processing Systems*, 9.

Wang, Z. *et al.* (2018) ‘Random Forest based hourly building energy prediction’, *Energy and Buildings*, 171, pp. 11–25. Available at: https://doi.org/10.1016/J.ENBUILD.2018.04.008.

Weihs, C. *et al.* (2005) ‘klaR Analyzing German Business Cycles’, *Data Analysis and Decision Support*, pp. 335–343. Available at: https://doi.org/10.1007/3-540-28397-8_36.

Wickham, H. (2011) ‘The Split-Apply-Combine Strategy for Data Analysis’, *Journal of Statistical Software*, 40(1), pp. 1–29. Available at: https://doi.org/10.18637/JSS.V040.I01.

Willand, N. *et al.* (2020) ‘Retrofit Poverty: Socio-economic Spatial Disparities in Retrofit Subsidies Uptake’, *Buildings & Cities*, 1(1), pp. 14–35. Available at: https://doi.org/10.5334/bc.13.

Willand, N., Torabi, N. and Horne, R. (2023) ‘Recognition justice in Australia: Hidden energy vulnerability through the experiences of intermediaries’, *Energy Research & Social Science*, 98, p. 103013. Available at: https://doi.org/10.1016/J.ERSS.2023.103013.

Willmott, C.J. and Matsuura, K. (2005) ‘Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance’, *Climate Research*, 30(1), pp. 79–82. Available at: https://doi.org/10.3354/CR030079.

Wood, S.N. (2011) ‘Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models’, *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 73(1), pp. 3–36. Available at: https://doi.org/10.1111/J.1467-9868.2010.00749.X.

Wood, S.N. (2025) ‘Generalized Additive Models’, *Annual Review of Statistics and Its Application*, 12(Volume 12, 2025), pp. 497–526. Available at: https://doi.org/10.1146/ANNUREV-STATISTICS-112723-034249.

Woodhall-Melnik, J., LeBlanc Haley, T. and Reiser, C. (2025) ‘Renovate to evict: an analysis of media discussions of renoviction’, *Housing Studies* [Preprint]. Available at: https://doi.org/10.1080/02673037.2024.2446716.

Wu, X. *et al.* (2008) ‘Top 10 algorithms in data mining’, *Knowledge and Information Systems*, 14(1), pp. 1–37. Available at: https://doi.org/10.1007/S10115-007-0114-2/METRICS.

Yegnanarayana, B. (2004) *Artificial Neural Networks*. PHI Learning Pvt. Ltd.

Yip, A.O. ni, Mah, D.N. yin and Barber, L.B. (2020) ‘Revealing hidden energy poverty in Hong Kong: a multi-dimensional framework for examining and understanding energy poverty’, *Local Environment*, 25(7), pp. 473–491. Available at: https://doi.org/10.1080/13549839.2020.1778661.

Yun, W.S. *et al.* (2024) ‘IoT-Based Indoor Thermal Environment and Occupancy Monitoring for Energy Poverty Care’, *Energies 2024, Vol. 17, Page 326*, 17(2), p. 326. Available at: https://doi.org/10.3390/EN17020326.

Annexes

Annexe I. Datasets used in the study

Tallinn

Name:	Buildings from the Estonian Topographical Database with Building Registry data
File format:	.GPKG
Description:	A dataset containing the combined information on the buildings from the Estonian Topographical Database and Building Registry. Energy performance certificate information is included in the Building Registry.
Type:	Vector (Polygon)
Temporal extent:	Last updated: 02/11/2024
Spatial extent:	N: 6633595.3269999995827675 S: 6377179.0940000005066395 W: 370011.2199999988079071 E: 738969.5490000024437904 CRS: EPSG:3301
Provider:	Maa- Ja Ruumiamet, Kliimaministeerium
Date of acquisition:	05/11/2024
Access level:	Publicly-available
Link:	Hyperlink to the source

Name:	KredEx building renovation grant utilisation
File format:	.CSV
Description:	A temporal dataset of KredEx renovation grant distribution for multi-apartment buildings.
Type:	Attribute table
Temporal extent:	09/30/2010 – 04/10/2023

Spatial extent:	*NA*
Provider:	Kliimaministeerium
Date of acquisition:	17/05/2024
Access level:	Restricted access
Link:	*NA*

Name:	Statistical neighbourhoods of Tallinn with population variables
File format:	.SHP; .CSV
Description:	A dataset of Tallinn divided into 230 statistical neighbourhoods with three categories of variables:
Type:	Vector (Polygons)
Temporal extent:	31/12/2021
Spatial extent:	N: 6606228.2819000007584691 S: 6579478.3634999990463257 W: 531218.1688999980688095 E: 552564.0905000014463440 CRS: EPSG:3301
Provider:	Anneli Kährik, University of Tartu
Date of acquisition:	08/06/2024
Access level:	Restricted access
Link:	*NA*

Riga

Name:	Building cadastre attribute data
File format:	.XML
Description:	A collection of XML files for the attribute data for the buildings. XML files were converted to CSV with a Python script. From the files, the building cadastre code, name, construction area, floors, and year of exploitation were exported during the file format change.
Type:	Attribute data

Temporal extent:	Last updated: 29/09/2024
Spatial extent:	*NA*
Provider:	Valsts zemes dienests
Date of acquisition:	05/10/2024
Access level:	Publicly-available
Link:	Hyperlink to the source

Name:	Building cadastre spatial data
File format:	.SHP
Description:	Building cadastre spatial data contains information about the spatiality and parcel codes for Riga. The dataset is split into 128 cadastral groups in Riga. Parcel codes were used to link spatial and attribute cadastre data for buildings.
Type:	Vector (Polygon)
Temporal extent:	Last updated: 29/09/2024
Spatial extent:	N: 328540.6270000003278255 S: 301544.7369999997317791 W: 494913.3839999996125698 E: 519729.6189999999478459 CRS: EPSG:3059
Provider:	Valsts zemes dienests
Date of acquisition:	05/10/2024
Access level:	Publicly-available
Link:	Hyperlink to the source

Name:	Energy performance certificate data
File format:	.CSV
Description:	A database of all the available attribute variables of the energy performance, without the associated building identification code.
Type:	Attribute data

Temporal extent:	Last updated: 24/11/2023
Spatial extent:	*NA*
Provider:	Būvniecības valsts kontroles birojs
Date of acquisition:	25/11/2023
Access level:	Publicly-available
Link:	Hyperlink to the source

Name:	Energy performance certificate cadastre objects
File format:	.CSV
Description:	A database of the cadastre object identification data and associated energy performance certificate data. The cadastre object and energy-related data were linked through the number of the energy performance certificate number.
Type:	Attribute data
Temporal extent:	Last updated: 24/11/2023
Spatial extent:	*NA*
Provider:	Būvniecības valsts kontroles birojs
Date of acquisition:	25/11/2023
Access level:	Publicly-available
Link:	Hyperlink to the source

Name:	Building renovation projects financed by ALTUM
File format:	.XLSX
Description:	A spreadsheet containing the information about the acquisition of the building renovation grants in Latvia for the 2016-2023 ALTUM programme. Connected to the other data through the building address.
Type:	Attribute data
Temporal extent:	Last updated: 31/12/2023
Spatial extent:	*NA*

Provider:	ALTUM
Date of acquisition:	11/02/2025
Access level:	Publicly-available
Link:	Hyperlink to the source

Name:	Building address data
File format:	.SHP
Description:	Dataset of all building addresses used to the extent of Riga. The dataset used to spatially link ALTUM building renovation grants and building cadastre data.
Type:	Vector (Points)
Temporal extent:	Last updated: 22/02/2025
Spatial extent:	N: 326568.0000000000000000 S: 301566.6000000000349246 W: 496251.7820000000065193 E: 519529.5706180345732719 CRS: EPSG:3059
Provider:	ALTUM
Date of acquisition:	22/02/2025
Access level:	Publicly-available
Link:	Hyperlink to the source

Name:	Riga population characteristics divided into grid cells
File format:	.SHP
Description:	Riga population characteristics divided into grid cells of a size of 1 square kilometre. In total, 280 units are represented. The population characteristics include the largest ethnic groups, the division of workers into three occupational groups, and three age groups.
Type:	Vector (Polygons)
Temporal extent:	31/12/2021

Spatial extent:	N: 327000.0000000009313226 S: 300999.9999999990686774 W: 495999.9999999998835847 E: 520000.0000000000000000 CRS: EPSG:3059
Provider:	Māris Bērziņš, University of Latvia
Date of acquisition:	30/01/2025
Access level:	Restricted access
Link:	*NA*

Vilnius

Name:	Buildings with associated heating indicators
File format:	.SHP
Description:	A building dataset, which includes all structural and heating information about the buildings in Vilnius, including information about the facilitation of renovation activities. The dataset is primarily for multi-apartment buildings. Does not include energy performance certificate data.
Type:	Vector (Polygons)
Temporal extent:	Last updated: 06/11/2024
Spatial extent:	N: 6075900.7651000004261732 S: 6050592.4375000018626451 W: 569237.5830000005662441 E: 594353.1001000003889203 CRS: EPSG:3346
Provider:	Vilniaus miesto duomenų centras|
Date of acquisition:	31/12/2024
Access level:	Publicly-available
Link:	Hyperlink to the source

Name:	Energy performance certificate data

File format:	.CSV
Description:	A dataset containing all the data related to the energy performance certification of a building. Linked to the building dataset through the address.
Type:	Attribute data
Temporal extent:	Last updated: 01/01/2025
Spatial extent:	*NA*
Provider:	Statybos sektoriaus vystymo agentūra
Date of acquisition:	02/01/2025
Access level:	Publicly-available
Link:	Hyperlink to the source

Name:	Vilnius neighbourhood population characteristics
File format:	.SHP; .XLSX
Description:	A dataset containing the population data of Vilnius, divided into 872 neighbourhoods in the city. The data contains all the associated ethnic, age group, and occupational variables relevant to the study.
Type:	Vector (Polygons)
Temporal extent:	31/12/2021
Spatial extent:	N: 6078140.3430000003427267 S: 6048670.1071000006049871 W: 566166.9610000001266599 E: 595515.8850000001257285 CRS: EPSG:3346
Provider:	Rūta Ubarevičienė, The Lithuanian Centre for Social Sciences
Date of acquisition:	22/10/2024
Access level:	Restricted access
Link:	*NA*

Annexe II. Gathered policy action examples in Europe

Country	Type of measure	Measure	Reference	Comment
Lithuania	Direct financial	Compensation for heating costs of the dwelling	Seimas of the Republic of Lithuania, Law on Cash Social Assistance for Low-Income Families (Single Residents) (Law No. IX-1675), (2003)	During the heating season, low-income households are eligible for partial compensation of the heating expenses. It is constructed of the difference in the state-supported income per family and the total costs of heating exceeding a certain threshold (25%) of overall expenses.
Catalonia (Spain)	Administrative	Security from energy supply cuts for energy-poor households	Comunidad Autónoma de Cataluña, BOE-A-2015-9725 Ley 24/2015, de 29 de julio, de medidas urgentes para afrontar la emergencia en el ámbito de la vivienda y la pobreza energética, (2015).	For the households considered at risk of residential exclusion, the right of access to basic supplies of drinking water, gas, and electricity is guaranteed.
France	Direct financial	Provision of energy checks to the low-income households	Government of the French Republic, Chapter IV: Protection of consumers in situations of energy poverty (Articles L124-1 to L124-5), (2024).	Households falling under a certain income threshold are provided with energy checks, which can be used for the payment of energy bills or renovation activities
France	Energy Efficiency	Building renovation subsidies	Gouvernement de la République française, Décret n° 2020-26 du 14 janvier 2020 relatif à la prime de transition énergétique - Légifrance, (2024)	The national subsidy for the building renovation programme, in which the amount of funding is varied by the income category of a household, providing the highest benefits to the low-income households
Denmark	Direct financial	Heating supplements for pensioners	Beskæftigelsesministeriet, Bekendtgørelse om social pension (førtidspension, seniorpension og folkepension), (2023)	The heating allowance provides direct support to the pensioners to pay for their energy bills, if they exceed an annual expense threshold
Ireland	Direct financial	Fuel Allowance Scheme	Department of Social Protection of the Republic of Ireland. (2024, February 15). Operational Guidelines: Fuel Allowance Scheme. GOV.IE.	Fuel Allowance is a type of payment for low-income households or households receiving other types of state support. The payments to households are performed either on a weekly basis or bi-seasonally between the end of September and April (28 weeks total). In the 2023-2024 season, the weekly allowance rate was 33 EUR.
The Netherlands	Energy Efficieny	Nationaal Warmtefonds zero-interest rate loans for building renovation	Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. (2022, September 14). Geen rente voor mensen met kleinere portemonnee die huis	In 2022, the Nationaal Warmetefonds started providing zero-interest loans for building renovation for low-income households.

			willen verduurzamen, Rijksoverheid.	
France	Energy Efficieny	White Certificate Scheme	Ministère de la Transition écologique et de la Cohésion des territoires, Code de l'énergie Partie législative Chapitre Ier : Le dispositif des certificats d'économies d'énergie, (2021)	White certificate scheme operates as an Energy Trading System for French energy providers. It distributes a certain amount of certificates, which correspond with the maximum energy production capacity in order to decrease energy production and motivate the energy provider companies to facilitate the energy efficiency improvements in different sectors. In Article L. 21-1-1, a framework of facilitation of energy obligations via the work with energy-poor populations for their benefit is stated.
Portugal	Direct financial	Social energy tariff	DA INOVAÇÃO E DO DESENVOLVIMENTO MINISTÉRIO DA ECONOMIA, Decreto-Lei n. 138-A/2010 de 28 de Dezembro, (2010)	To ensure that all consumers have access to the electricity supply, the social electricity tariff is established to create a stable electricity tariff with lower annual inflation (1%). The social electricity tariff is a part of the social security system of Portugal
United Kingdom	Direct financial	Warm Home Discount Scheme	Department for Business, Energy & Industrial Strategy. Warm Home Discount: Better targeted support from 2022, 2021	The Warm Home Discount is a key policy in the UK combating energy poverty since 2011. It operates as electricity bill rebates for different core groups of households, which receive social benefits. The Warm Home Discount scheme has additional targets for energy-poor populations.
United Kingdom	Direct financial	Cold Weather Payment	ADM Chapter L4: Cold Weather Payments , Government of the United Kingdom (2020).	Between November 1 and March 31, if the average temperature in the area is below 0°C, the households receiving social benefits will receive a 25£/week payment.
United Kingdom	Direct financial	Winter Fuel Payment	Thurley, D., Mcinnes, R., & Kennedy, S. (2019). Winter Fuel Payments update (CBP-6019).	Winter Fuel Payment is a one-time payment per heating season for the population older than 66 years.
United Kingdom	Energy Efficiency	Better Energy Warmer Homes Scheme	Sustainable Energy Authority of Ireland. (2023). Better Energy Warmer Homes Scheme Scheme and Application Guidelines	The Scheme provides free energy efficiency updates for eligible homes, including attic insulation, cavity, external, and internal wall insulation, window replacement, heating system and ventilation updates, and lighting upgrades. The beneficiaries of the scheme include households receiving fuel allowance or other types of social benefits

Northern Ireland (United Kingdom)	Energy Efficiency	Affordable Warmth Scheme	Northern Ireland Housing Executive. Affordable Warmth Scheme. Housing Executive. Retrieved March 23, 2024, from https://www.nihe.gov.uk/housing-help/affordable-warmth-boiler-replacement/affordable-warmth-scheme	The scheme is focused on energy poverty alleviation in private housing (social housing is not eligible) with a grant of to 10,000.00£ for energy efficiency measures divided into four levels of priority: insulation/ventilation/draught-proofing (focusing on the roof and basement); replacement of the heating system; replacement of windows; and solid wall insulation (the last level priority due to the expenses)
Ireland	Energy Efficiency	Housing Aid for Older People	Department of Housing, Local Government and Heritage of the Republic of Ireland (2023, December 19). Housing Aid for Older People GranT. Gov.Ie. https://www.gov.ie/en/service/1ca60-housing-aid-for-older-people-grant/	The scheme focuses on providing support for renovating the building for the elderly population, prioritising the elderly population with medical needs and low income. The maximum sum of the grant is 8,000.00 EUR, which may cover up to 95% of the approved costs. The grant can be for different types of essential repairs including door and window replacement, central heating system repairments, or roofing updates.
Finland	Direct financial	General Housing Allowance	Kansaneläkelaitos. General housing allowance. Kela. Retrieved March 23, 2024, from https://www.kela.fi/general-housing-allowance	The General Housing Allowance is provided for low-income households to cover up to 80% (from April 1, 2024, to 70%) of housing expenses. In addition, pensioners with a low income can receive a Housing allowance for pensioners, which equals 85% of the total housing costs.
Lithuania	Energy Efficiency	Support of vulnerable populations in building renovation	Lietuvos Respublikos Aukščiausioji Taryba - Atkuriamasis Seima (2024) I-2455 Lietuvos Respublikos valstybės paramos daugiabučiams namams atnaujinti (modernizuoti) įstatymas. Vilnius: Lietuvos Respublikos Aukščiausioji Taryba - Atkuriamasis Seima. Available at: https://www.e-tar.lt/portal/en/legalAct/TAR.9D04F98F7C14/ewjiwSYIND (Accessed: 24 March 2025).	Within the Multi-Apartment Building Modernisation Programme, households eligible for the heating allowance can apply for public support to cover the remaining housing renovation costs. The household receives the full cost subsidy for the renovation activities if approved.

Non-exclusive licence to reproduce the thesis and make the thesis public

I, Kirils Gončarovs,

1. grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright, my thesis

   Patterns of inequalities in housing energy efficiency and links with population risk factors in Tallinn, Riga and Vilnius,

   supervised by Anneli Kährik.

2. I grant the University of Tartu a permit to make the thesis specified in point 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 4.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.
3. I am aware of the fact that the author retains the rights specified in points 1 and 2.
4. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Kirils Gončarovs

24/05/2025