Journal of Building Performance Simulation

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tbps20

Surrogate optimization of energy retrofits in domestic building stocks using household carbon valuations

James Hey, Peer-Olaf Siebers, Paul Nathanail, Ender Ozcan & Darren Robinson

JOURNAL OF BUILDING PERFORMANCE SIMULATION
https://doi.org/10.1080/19401493.2022.2106309

Surrogate optimization of energy retrofits in domestic building stocks using household carbon valuations

James Hey[a], Peer-Olaf Siebers [a], Paul Nathanail[a], Ender Ozcan[b] and Darren Robinson[c]

[a]Computer Science, University of Nottingham, Nottingham, UK; [b]Land Quality Management Ltd, Nottingham, UK; [c]School of Architecture, The University of Sheffield, Sheffield, UK

ABSTRACT
Modelling energy retrofit adoption in domestic urban building stocks is vital for policymakers aiming to reduce emissions. The use of surrogate models to evaluate building performance combined with optimization procedures can optimize small building stocks but are insufficient at the urban scale. Recent methods train neural networks using samples of near-optimal solutions further decreasing the computational cost of optimization. However, these models do not make definitive predictions of decision makers with given environmental preferences. To rectify this, we extend the method by assigning a carbon valuation to households to derive their optimal retrofit solutions. By including the carbon valuation when training the predictive model, we can analyze the impact of households' changing attitudes to emissions. To demonstrate this method we construct an agent-based model of Nottingham, finding that simulated government campaigns to boost environmentalism improve both the number of retrofits performed and the mean emissions reduction of each installation.

ARTICLE HISTORY
Received 28 July 2021
Accepted 18 July 2022

KEYWORDS
Energy performance retrofits; building stock analysis; surrogate modelling; machine learning; optimization

List Of Acronyms

- ANN/DNN – Artificial/Deep Neural Network
- ABM – Agent-Based Model
- EWI – External Wall Insulation
- GA – Genetic Algorithm
- HCV – Household Carbon Valuation
- IWI – Internal Wall Insulation
- LCS – Lifecycle Carbon Savings
- NPV – Net Present Value
- SEPM – Surrogate Energy Performance Model
- SO – Surrogate Optimizer/Optimization
- WHRS – Whole House Retrofit Solution
- WTP – Willingness To Pay

1. Introduction

Domestic dwellings account for over a third of the national energy demand and approximately a quarter of total CO_2 emissions in the UK, with the majority of this energy demand being used for electric or gas-based space heating (Waters 2019; Office For National Statistics 2021). While significant progress has been made in the design of new buildings for efficient heat generation and retention, most of the existing stock will still be occupied by the UK government's deadline of 2050 for net zero emissions (The UK government 2008). Meeting this target will require strategies to transform the existing building stock with Whole House Retrofit Solutions (WHRSs). To analyze the impact of such policies an increased emphasis should be placed on robust tools for modelling and optimizing the policy decisions such as tax rate, public engagement, and methods of outreach. However, detailed dynamic analysis of building stock retrofit adoption is complex, as not only the physical properties of each dwelling differ, but each household has heterogeneous preferences in selecting a WHRS to meet their objectives.

Discovering near-optimal solutions to fit household preferences is computationally expensive at the urban scale. There have recently been some successful attempts to generate near-optimal sets of WHRSs using Deep Neural Network (DNN) based Surrogate Optimizers (SOs) (Hey et al. 2020; Thrampoulidis et al. 2021). These techniques use a data set of WHRSs generated using traditional methods of simulation, surrogate modelling, and optimization. These WHRSs are used to train a predictive model of household retrofit decisions which is then applied to the remaining buildings in the data set. While all surrogate models result in some loss of accuracy, they also greatly cut down the computational effort of finding potential

CONTACT James Hey james.m.hey@gmail.com

retrofits. Existing SOs take a preference-blind approach, building a front of potential retrofits for selection by a decision-maker. While this is useful when stakeholder engagement is possible, it is unsuitable for applications such as an agent-based model (ABM) in which the preference information of agents is available a priori, requiring a definitive solution for the preferences stated.

When attempting to include households in the building stock modelling scope, a form of synthetic population is required as well as a mechanism for realistic interaction between the physical and social components (Robinson 2019). We extend the surrogate optimization method by providing a mechanism for tailoring these retrofit solutions to a measurable and interpretable metric of environmentalism (willingness to pay for carbon mitigation) among the synthetic population. To create the SO, we use an established method of generating optimizations using a Surrogate Energy Performance Model (SEPM) in conjunction with a Genetic Algorithm (GA) to produce a data set of near-optimal WHRSs. The objective function used in this optimization is designed to include a Household Carbon Valuation (HCV) expressed in £/tCO2e. This allows household agents to transparently select the retrofit that most reflects their personal preference between economic and environmental considerations. This permits the modelling of heterogeneous emission-preferences among the synthetic agents, better capturing the diverse attitudes within the target population. When training the Surrogate Optimizer (SO), this HCV is provided as an input feature, allowing the SO to generate solutions for individual HCVs. We demonstrate a potential use of this method by integrating it with an ABM in which an agent representing the local government runs campaigns to boost environmentalism among homeowners. A model of this scale using a measure of households' willingness to pay for carbon mitigation would have been computationally infeasible using existing methods.

2. Related works

Surrogate Optimization (SO) is a recently developed method for building stock retrofit modelling. The SO method requires the simulation of a sample of building stock for the construction of a Surrogate Energy Performance Model (SEPM). The SEPM is used for retrofit optimization of another sample of the stock to create a data set of retrofit decisions for training the predictive SO. The SO can then be used for rapid evaluation of buildings at significantly reduced computational cost. Following is a brief summary of the related works covering the modelling of building stocks, energy simulation, surrogate energy performance models, and their optimization. Then the prior attempts at SO methods are considered which reduce the computational cost of retrofit evaluation using predictive models, but in which we find a lack of household preference transparency required for building stock models with an environmentally conscious synthetic population. Finally, a review of the works pertaining to Willingness To Pay (WTP) for carbon emissions mitigation is presented, demonstrating the measurability and interpretability of this metric, making it well-suited to stock modelling applications.

2.1. Building stock modelling

Building stock modelling can be approached from either a top-down or bottom-up perspective (Kavgic et al. 2010). Surrogate optimization is a bottom-up approach, as households are modelled individually with heterogeneous characteristics. This allows the analysis of policies enacted at the local level, which is difficult with top-down approaches. Broadly speaking, we can consider the computational cost of analysing a building stock from a bottom-up perspective as being determined by the size of the stock, and the level of detail in which each residence is modelled. Physical methods require the simulation of each building, and so generally scale linearly, making large stocks less feasible to model, while statistical methods rely on training a model using measured data (Reinhart and Davila 2019). When these models are trained using simulated data they are typically referred to as surrogate models, as the statistical model acts as a surrogate for the physical one. The use of statistical models greatly increases the scalability of the problem, as they can be evaluated in a fraction of the computational time physical models take.

One technique requires simulating a single representative member of each building archetype, then multiplying this energy usage by the total floor area this archetype represents in a given stock (Reinhart and Davila 2019). In the absence of a mapping between each archetype and each dwelling of that type, there is no direct representation of building-specific WHRSs at the scale of individual buildings. While this method is suitable for very large stocks, and this level of analysis can give estimates of total energy usage, the missing level of detail can be restrictive: it prevents us studying the impact that dynamic installations of retrofits have across the stock. A review of 29 national housing stock energy models found that the tools being utilized by researchers have 'limited scope and simplicity' (Sousa et al. 2017, 77). Models were criticized for overwhelming bias towards the use of monthly or annual energy balance calculations rather than dynamic simulation, resulting in oversimplification of physical phenomena. However, this approach is not suitable for explicit modelling of whole city housing stocks, which can be performed with the use of descriptive bottom-up approaches allowed by surrogate optimization. A review of bottom-up energy demand models by Kavgic et al. (2010) highlighted the importance of an energy models' ability to capture socio-technical behavioural elements as well as explore the impact of emission reduction strategies and emergent technologies, demonstrating the importance of developing robust bottom-up methodologies in solving building stock modelling problems.

When considering the optimization of building stocks, Sola et al. (2018) reviewed the available urban building stock optimization tools, including their individual drawbacks and benefits. This is generally done by extrapolating archetypes to an entire city as a building archetype is both a significant factor of commonality between buildings and often widely discernible using public data sets. A good example of this technique is presented by Wu et al. (2017), who performed retrofit optimization using greenhouse gas emissions and annualized costs as objective functions. They performed the optimization technique on representative buildings of a particular archetype then performed an upscaling methodology based upon the total floor area of that building type in the community. Another approach to urban building stock optimization is presented by Brownlee et al. (2020), who performed a case study by optimizing the distribution of a limited public investment for reducing a housing stock's energy demand. They used a process of sequential Pareto optimization, in which the whole stock optimization problem is encoded into sequential sub-problems, significantly reducing the search space compared with the naive approach of solving the solution as a global problem. While this process provides large computational improvements, it does require the generation of a front for each sub-problem which is still computationally expensive for large stocks.

2.2. Building energy simulation

Physical simulation of building energy involves the construction of a detailed three-dimensional representation of a target building, which is then simulated under given environmental and occupational conditions to determine energy flows. One of the most ubiquitous and full-featured tools available for building energy simulation is EnergyPlus, simulation software supported by the United States Department of Energy (Crawley et al. 2000). EnergyPlus is primarily designed for single building simulations and is feature-rich, supporting detailed HVAC modelling, detailed fenestration modelling capabilities, and a function mock-up interface to allow co-simulation of building occupant behaviour. EnergyPlus has been used to calculate energy demand in some building stock retrofit adoption models but, due to the computational cost of these calculations, the scope of the retrofit packages considered was kept limited (Wang et al. 2018).

There are alternative simulation software targeted specifically at urban simulation. CitySim, for example, creates three dimensional scenes up to the size of an entire city (Robinson et al. 2009; Zakhary et al. 2016). Using urban energy simulation tools is likely to be important when significant interaction occurs between buildings, such as shading effects from high-rise architecture. However, the simultaneous simulation of interacting buildings comes with significant resource costs compared with atomic simulation approaches.

2.3. Surrogate modelling and optimization

White box simulation, the attempt to capture the underlying processes in a high level of detail, can be very computationally expensive in optimization scenarios where lots of repetitions are required such as bottom-up building stock analysis. Surrogate modelling, also known as meta-modelling or response surface modelling, is the use of a faster but less accurate model to replace a slow process (Eisenhower et al. 2012). The use of SEPMs to reduce the computational cost of simulation stages is a common method in sustainable building design (Evins 2013; Melo et al. 2014). Any suitable regression modelling technique can be used for surrogate modelling, with both linear regression (Tian et al. 2015) and ANN-based surrogates being popular choices (Melo et al. 2014; Tseranidis, Brown, and Mueller 2016). The largest scale SEPM we have seen to date was presented by Edwards et al. (2017) who trained a Deep Neural Network (DNN) surrogate using a big data approach, achieving high levels of accuracy at hourly precision with errors of less than 5%. This level of precision is often unnecessary, with total energy demand calculations sufficient to analyze the impact of most retrofit installations, which they were able to calculate with errors of just 0.07% at greatly reduced computational time compared with traditional simulation.

A common use of SEPMs is their integration with an optimization method (Magnier and Haghighat 2010; Prada, Gasparella, and Baggio 2018; Waibel et al. 2019; Sharif and Hammad 2019). Due to the potentially large number of function calls required for optimization, the use of a surrogate greatly reduces the computational cost. Prada, Gasparella, and Baggio (2018) provide an analysis of the performance of different surrogate modelling techniques in the context of optimization. They used a Genetic Algorithm (GA) with alternate surrogate models and compare the results to a brute force optimization to determine the efficacy of alternative methods. They confirmed the practice as an acceptable way of optimizing, with the majority of Pareto solutions identified by simulating only 3–8% of the solution space for training. Sharif and Hammad (2019) consider this method from a different perspective, performing traditional simulation-based optimization to obtain a retrofit solution set, then training a DNN to predict the objective values associated with this set with good rates of accuracy. While more computationally expensive than training a SEPM, this method allows for more nuanced objective values, such as building specific lifecycle analysis, which may not be as simple to model as energy performance. The flexibility of DNNs, in particular, has made them a popular surrogate modelling technique for objective value estimation, with another good example by Ascione et al. who evaluate cost, energy savings, and thermal discomfort of solutions using trained DNNs (Ascione et al. 2017). They designed the system to be generic and applicable to any building type, although it was not targeted towards whole stock modelling.

Model-based optimization methods use a similar approach by attempting to construct a surrogate model of the fitness landscape during the optimization process in order to converge more efficiently to the optimum solution (Costa et al. 2018). These methods have been proposed as alternatives to more common metaheuristic approaches such as GA optimization, although they are still less common (Wortmann 2018). While deploying a similar concept to surrogate optimization, these algorithms are trained in real time to optimize a specific objective function; in contrast to a model trained to achieve fast optimizations on generic buildings from a data set.

2.4. Surrogate optimization in building stocks

There have been some recent attempts at surrogate retrofit optimizers of the kind we extend in this paper. A single objective whole city surrogate optimization was attempted by Hey et al. (2020), who trained a SO on a sample of buildings in Nottingham city in order to predict net present value optimized solutions for the remaining stock. A financially viable solution was identified for 16.7% of the whole building stock compared with 19.2% of the training sample, suggesting 87% of buildings with a viable solution were identified. Predicted solutions performed 11% worse than the sample solutions, making the significant computational savings of the approach a trade off with solution quality. While this paper evaluated the financial performance of alternative solutions well, simulated decision makers were not influenced by environmental preferences. Instead, only the lifecycle financial returns of the retrofit investment was considered. Thrampoulidis et al. (2021) use a hybrid classification and regression model to predict near-optimal retrofits using cost and emissions objectives. Their classification model used preset wall, window, and roof insulation settings to perform a binary classification, reducing the search space compared with our proposed method. However, they focussed on a wider range of energy-supply methods including renewable sources and energy storage. Their method generates a fixed-length Pareto front by first finding a solution which maximizes a single objective value, then making sequential calls to a model in which the output of the previous stage is used as input for the next. This method is efficient for generating a front of fixed size, but obfuscates the underlying trade off required to move from one point on the front to the next. Their work found surrogate optimization to be a convenient balance between computational cost and accuracy, as well as potentially more accessible for non-experts to use due to the diminished data requirements for using the trained model compared with traditional retrofit optimization approaches.

2.5. Willingness to pay for carbon mitigation

In economics, it is common to model the environment as a public good (Siebert 2007). Voluntary contributions, the mechanism through which socially responsible consumers interact with public goods, are introduced well by Brekke, Kverndokk, and Nyborg (2003). While early experimental data on public good contributions showed that, particularly in large groups, contributions were low (Isaac and Walker 1988), there are a range of circumstances in which contributions are frequent, such as under social pressure or with institutional intervention (Chaudhuri 2011; Zhang, An, and Dong 2021). We can consider these contributions towards emissions mitigation in terms of the Willingness To Pay (WTP) for a ton of emissions mitigated. A common method of deriving WTP estimations are discrete choice experiments, in which participants are offered a range of alternate proposals with different characteristics from which researchers can later estimate the significance of each factor. MacKerron et al. (2009) used a discrete choice experiment to ascertain WTP for emissions offset of passenger's airline flights. The derived WTP per tCO_2e ranged from £10.16 to £38.35 with a mean of £24.26. The most significant factor in offset decisions was independent certification of the offset, which nearly doubled the derived WTP. Values were negatively affected by the price of the offset and were lower for males. Rotaris, Giansoldati, and Scorrano (2020) performed a similar discrete choice experiment to determine the WTP of 1228 Italian airline passengers, determining a similar carbon value of between €12 and €38 per tCO_2e. Streimikiene et al. (april, 2019)

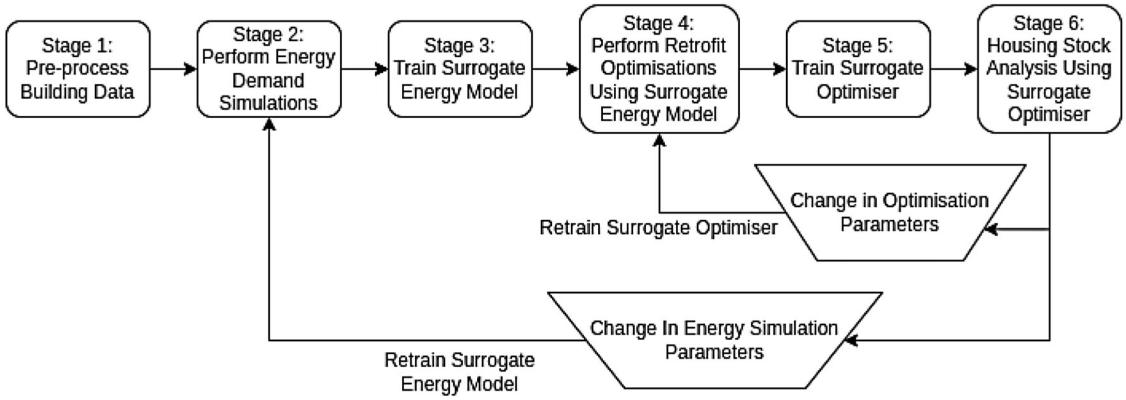

Figure 1. High level process flow diagram of method.

perform a review of WTP studies and commented on the wide range of factors which influence WTP, including income, gender, and country of origin while also noting that 'pro-environmental lifestyles are related to higher levels of WTP for energy conservation equipment' (Streimikiene et al. april, 2019, 1480).

When it comes to contributions in practice, a 2017 industry report on voluntary carbon offset markets shows an average carbon price of \$3/tCO_2e, although this varied significantly depending on the type of mitigation project and the project location (Ecosystem Marketplace 2017). The annual quantity of carbon offsetting is also highly volatile, suggesting an immature market. It is worth noting that real-world carbon offset prices are significantly lower than reported WTP from consumer studies, possibly indicating a gap between reported and actual decisions. Although this could also be explained by the high variance in offset prices, indicating consumer willingness to mitigate is much higher than found in industrial settings.

3. Method

A high level view of our proposed method can be seen in Figure 1. Stages 1–4 implement an established method of generating Whole House Retrofit Solutions (WHRS) by training an energy performance model from simulation data to be used in conjunction with a standard optimization method. Here we create Pareto solutions by explicitly defining a household carbon valuation (HCV) within the objective function, expressed in £/tCO_2e. We use these HCVs linked WHRSs to train the Surrogate Optimizer (SO) in stage 5, allowing the use of this tool for analysis (stage 6) of the wider housing stock, which we later demonstrate using an illustrative agent-based model.

3.1. Stage 1: data sources and pre-processing

Building footprints, eave heights, orientations, and locations were obtained from Ordnance Survey data for the city of Nottingham (Ordnance Survey 2018). These were transformed into three-dimensional models using the footprint and relative height data. Existing building fabric details were assigned stochastically by disaggregating English housing survey data, and stratified by building archetype. One thermal zone was modelled per floor, as this was possible without internal partition data while still approximating the separate living and sleeping zones used in models such as BREDEM (Henderson and Hart 2013). Flats and mixed use dwellings were excluded from the building sample due to insufficient data available to model the necessary internal partitions and communal areas. There are also concerns about socio-economic differences in retrofitting for communal and leasehold properties, which may require consensus or reclamation of funds through service charges which place them outside of the scope of the retrofit decision model used here.

3.2. Stage 2: energy demand simulation

3.2.1. Retrofit options

The retrofit options considered are shown in Table 1. These primarily cover the thickness and material of roof and wall insulation (internal or external), as well as the glazing type of windows. Thickness intervals were kept quite small and the range of thicknesses is large to ensure the method was robust over large optimization spaces. The space heating method is also included, although limited to gas central heating and electric space heating as they represent the most common heating methods as well as the widest fuel cost discrepancy. Considering these heating methods also allows for testing of an interesting hypothesis: in a scenario with a rapidly decarbonizing national grid, a strongly carbon averse and utility maximizing household with a long term perspective may retrofit electric heating to replace gas central heating despite significant cost increases. The physical properties

Table 1. High level description of retrofit options.

Component	Gene Name	Possible Values
External Wall Insulation Material	EWI_mat	None (uninsulated), XPS, EPS, PIR
External Wall Insulation Thickness	EWI_thick	30–150 mm in increments of 5 mm
Internal Wall Insulation Material	IWI_mat	None (uninsulated), XPS, EPS, PIR
External Wall Insulation Thickness	IWI_thick	30–150 mm in increments of 5 mm
Heating Method	Heating	Electric (Convection), Gas Central Heating
Roof Insulation Material	Roof_Mat	None (uninsulated), Mineral Wool
Roof Insulation Thickness	Roof_thick	50–400 mm in increments of 25 mm
Glazing Type	Glazing	Single Glazing, Double Glazing, Triple Glazing

Table 2. Insulation material properties.

Material Name	Thermal Conductivity (W/mK)	Density (kg/m^3)	Embodied Emissions (kgCO2e/m^3)	Sources
Expanded Polystyrene (EPS)	0.029	29	3.29	Parsons (2005) and Hammond and Jones (2008)
Extruded Polystyrene (XPS)	0.035	24	3.43	Parsons (2005) and Hammond and Jones (2008)
Polyisocyanurate (PIR)	0.025	24	5.40	Parsons (2005) and Boustead (2005)
Mineral Wool	0.040	20	1.12	Parsons (2005) and Hammond and Jones (2008)

and embodied carbon of the insulation materials used are shown in Table 2

3.2.2. Simulation using EnergyPlus

Simulations were performed using EnergyPlus,[1] as it is well established and considered reliable for evaluating delivered energy demand, and there exist tools to manipulate the input files programmatically. Simulations were performed atomically, each considering a potential state of a single domestic dwelling and thus excluding any shading effects, deemed reasonable given the low-rise nature of the data set used (average eave height 5.50m). Party walls of terrace houses are assumed adiabatic, based on the principle that temperature differences will be small and unknown when simulating individually. This also avoids any solar gains incorrectly attributed to the wall. Weather data was sourced from the met office weather center, located in Watnell (station id 03354), approximately 10 kilometres from the center of Nottingham (Met Office 2021) and generated using the UKCP09 tool (Eames, Kershaw, and Coley 2011).

To train the surrogate energy model, a broad range of input and output data was required. Simulations were performed using EnergyPlus, and so an EnergyPlus Input Data File (IDF) template was constructed.[2] Semantic attributes were drawn from survey data to generate a set of input combinations as shown in Figure 2, a process described in greater detail in Rosser et al. (2019). As such, the simulation of aggregate energy demand is expected to be plausible, but building-specific results may not be (in the absence of building-specific survey data). This is less likely to be significant in stock analysis as the buildings are being sampled so, provided unbiased attribution errors, the expected error will be minimized. When energy simulations failed, usually due to an issue with the dynamically constructed IDF, error reports are generated by EnergyPlus. Throughout the development process, all identified errors generated at this stage were corrected and the simulations re-run to ensure that no systematic bias in generated results was carried forward to SEPM training.

3.3. Stage 3: surrogate energy performance model training

The data flow of our Surrogate Energy Performance Model (SEPM) training process is shown in Figure 3, with candidate inputs for the surrogate model matching the retrofit variables used in the simulation process, as well as common geometric summary data such as glazing ratio, wall area, etc. A backwards feature selection process was performed, with features eliminated from the model systematically, and the model retrained. A single-tailed Student's t-test was performed on the r^2 values of a sample of 20 trained models, and features were eliminated if they had no effect on the model's performance.

The model training parameters were tuned using a grid search method, with 20 repetitions per settings combination, scored using the average r^2. The grid settings can be found in Table 3. The pre-processing methods, which functioned best during tuning, required continuous variables to be min-max normalized; while categorical variables such as material genes, were one-hot encoded (represented by one binary input node per category). The loss function, used to evaluate the model during training, was least squared errors (L2), as this is preferred for models not containing a large number of outliers. The Rectified Linear Unit (ReLU) activation function was used on all layers to provide the desired non-linearity

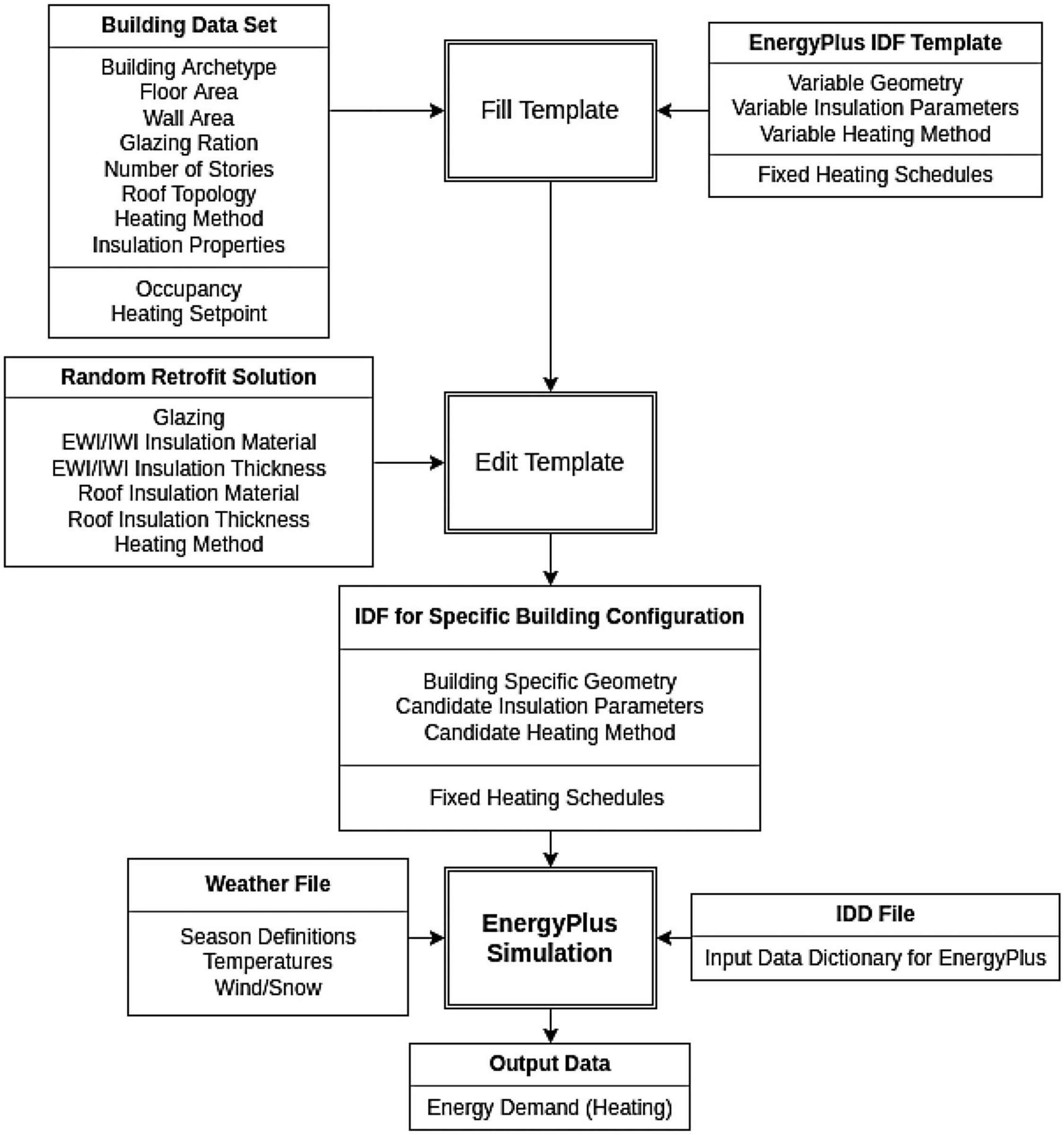

Figure 2. Conceptual data flow of single dwelling simulation during heating demand data generation stage (Stage 2).

in the neural network training. The best performing network topology comprized three hidden layers of diminishing size, and training batch sizes of 32 performed best, although the effect of this hyper-parameter was negligible.

3.4. Stage 4: optimization with genetic algorithm

A Genetic Algorithm (GA) was used to obtain the initial sample set of near-optimal retrofits, a process outlined in Figure 4. The method is population based, with parent solutions being selected, reproducing with mutation and then reevaluated to create the next generation. This required the selection of hyper-parameters, but also scenario settings that remain fixed for the purposes of optimization and would require re-performing if the optimization stage changed. These include the discounting rate, the prices (or dynamic price profiles) of labour, and fuel sources, as well as many more nuanced inputs, such as grid decarbonization strategies. As shown in Figure 1, this makes this stage a decision point and one that may need to be revisited cyclically should any of these scenario settings needed to be changed. If, for example, we wanted to consider a new scenario in which high inflation increases the discount rate, then the objective values at optimization time will be altered, rendering the sample of optimizations generated obsolete. As such, we will lay out the objective value calculations and how they were formed before discussing the hyper-parameters of the GA itself.

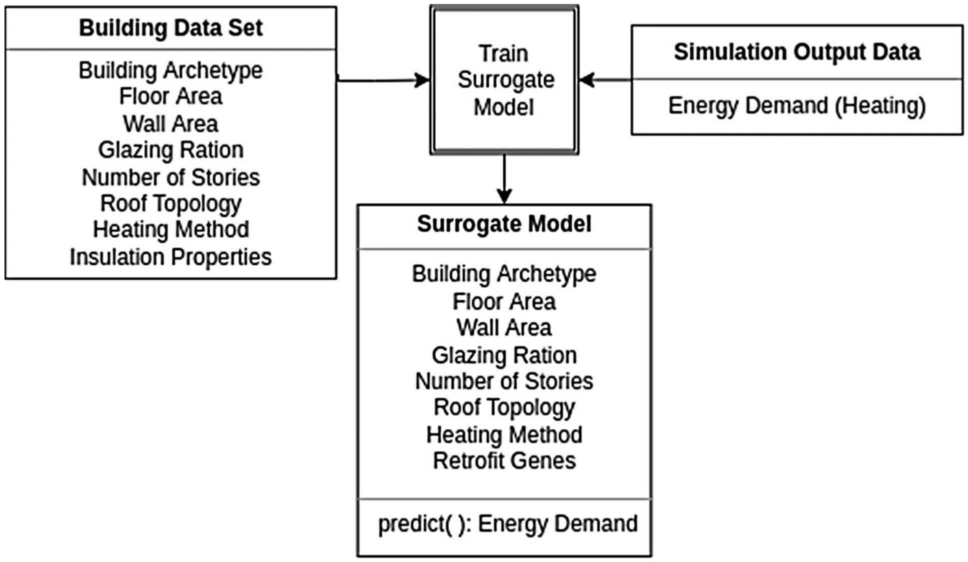

Figure 3. Conceptual data flow of surrogate energy performance model training (Stage 3).

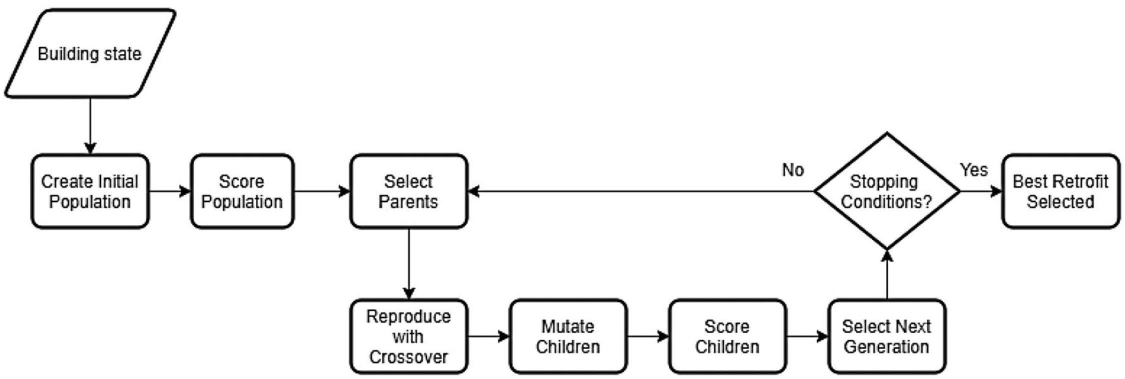

Figure 4. Process flow diagram of generic GA process.

Table 3. Settings grid used for artificial neural network surrogate model tuning, with selected settings highlighted in bold.

Setting	Energy performance model	Surrogate optimization model
Model type	Regression	Classification (materials) Regression (thickness)
Preproccessing	**Normalization**, Logerithic	**One hot encoding**, Binary encoding, Label encoding
Hidden Layers	1/ 16, 32, 64 2/ [64, 32], [32, 64] 3/**[64,32, 16]**, [128, 64, 32], [32, 64, 128]	1/ 16, 32, 64 2/ [64, 32], [32, 64] 3/**[64,32, 16]**, [128, 64, 32], [32, 64, 128]
Activation Function	**ReLU**	**ReLU + SoftMax (classification)** **ReLU (regression)**
Loss Function	**L2 loss**	**Cross Entropy (classificlation),** **L2 loss (regression)**
Batch Size	**32**, 16, 64	**32**, 16, 64
Training Function	**Adam**	**Adam**

3.4.1. Objective values

There were two primary objective values considered during the scoring stage of the GA. The first objective, Net Present Value (NPV), represents the value of financial savings achieved from a WHRS. The second objective, Lifecycle Carbon Savings (LCS), represents the CO_2e saved by energy demand mitigation, minus the embodied carbon cost of the retrofit. Both objectives consider lifetime values, including transportation, labour, and, where relevant, disposal of the materials used, as seemingly beneficial retrofits can become unfeasible if other lifecycle costs are considered (Gustafsson and Karlsson 1988).

NPV of savings is defined in Equation (1). This shows the NPV of a WHRS with an initial cost of C_0, a service period of p, providing an energy bill saving of s_t, in period t, and with maintenance costs of m_t. s_t includes savings from fuel source switching and can be negative if a solution adopts a more expensive fuel source. The discount rate of i is applied each year, representing the weighting towards the near future. In all cases, we applied a discount rate of 0.5%[3] A carbon tax at time period t is shown as g_t, applied to the emissions savings at period $t(e_t)$. This tax is also applied to the embodied carbon of the retrofit solution E_0. Where applicable life-cycle maintenance costs were rolled into upfront costs with appropriate discounting applied. While the assessment of the benefits of a retrofit assumes it will be operated for its full lifespan, there is no conceptual reason that a second retrofit could not be applied during this period, given the initial installation is a sunk cost. The NPV is expressed in present value currency terms, which in the case study considered are pounds sterling (£).

$$NPV = \sum_{t=1}^{t=p} \frac{s_t + g_t e_t - m_t}{(1+i)^t} - C_0 - E_0 g_0 \quad (1)$$

Equation 1: Net Present Value

LCS is represented by Equation (2). The value of emissions saving, e_t, will depend mostly upon the energy-saving properties of the retrofit. It will also be impacted by the energy carrier, as we model a decarbonizing grid resulting in decaying emissions contributions from electric heating methods over time. The lifetime embodied carbon is represented by E_0 and considers the cost of manufacturing, transporting, and disposing of building materials used in WHRSs.

The household's utility for a retrofit solution is shown in Equation (3), with the household's carbon valuation presented as v. Given that v is expressed in $£/tCO_2e$, the utility can be expressed in present value financial terms, allowing a direct trade off between NPV and emissions mitigation. The driver for v, which represents WTP for carbon mitigation, can be seen as the personal or social components that motivate households to be carbon conscious, and can be driven by a combination of altruism, social pressures, or self-preservation in the face of climate change.

$$LCS = \sum_{t=1}^{t=p} e_t - E_0 \quad (2)$$

Equation 2: Lifecycle Carbon Saving (LCS) calculation

$$U = NPV + vLCS \quad (3)$$

Equation 3: Utility function used in optimization

$$v_{u0} = \frac{-NVP}{LCS} \quad (4)$$

Equation 4: Minimum indifference carbon value

Equation (4) shows the minimum indifference carbon value. This is the minimum carbon value required for a household to be indifferent between inaction (their existing building state) and a given retrofit solution. If a household's individual carbon value exceeds this threshold, and the WHRS has a positive LCS, a utility maximizing household would benefit from adopting the solution. This value has some useful properties, being both unique for Pareto solutions and monotonically increasing on a concave Pareto front.

3.4.2. Embodied carbon

The inclusion of embodied carbon was deemed important as low carbon buildings tend to embody a greater share of whole life-cycle carbon in their construction, despite lower emissions overall (Su et al. 2020). This is particularly important when considering retrofits to existing stock that is not at end of life, as this carbon cost can be mitigated entirely through inaction. Embodied emissions values were taken from the Inventory of Carbon and Energy database where available, with interpolation used to obtain missing material thicknesses (Hammond and Jones 2008).

3.4.3. Grid decarbonization rate

It is the stated policy of the UK government to decarbonize the UK power grid by the year 2050 (Parliament of the United Kingdom 2008). Therefore, when considering the carbon emissions of power from the grid, a decay factor was included to ensure the full benefits of the decarbonized grid are factored in. This also means that optimizations under these scenarios have an a priori assumption that the national grid will meet its policy statements. To calculate the required decarbonization rate, the formula in Equation (5) is used. Given 2018 carbon levels (c_0) of 0.309 kg/kWh (Waters 2020), falling below 0.001 kg/kWh (c_t) by 2050 ($t = 32$) the required annual reduction must be 16.4%.[4] The decarbonization rate is highly relevant when considering extremely carbon-averse households as, when observing a highly decarbonizing grid, they would be more likely to consider using more expensive electric heating even when gas central heating is available.

$$r = 1 - \frac{c_t}{c_0}^{\frac{1}{t}} \quad (5)$$

Equation 5: Decarbonization Rate required to achieve CO2 levels of c_t after t years. An inversion of the compound annual-growth rate formula

3.4.4. GA parameter tuning

In order to tune our GA, we performed a parameter grid experiment with a variety of settings. The list of settings can be seen in Table 4. The soft elitist setting refers to a replacement method in which the single best performing parent is selected, with the rest of the population replaced by children. This is in contrast to pure elitism in which the best performing solutions, from both the parent and child population, are selected. In order to test both the single and multi-objective performance, a set of

Table 4. Settings grid used for genetic algorithm parameter-tuning grid.

Setting	Values
Stopping Condition	Max Calls
Max Calls	500
Mutation Method	Uniform, Random
Mate Selection Method	Tournement
Tournemenet Size	2
Recombination Method	2 Point, Uniform
Replacement Method	Pure Elitist, Soft Elitist
Mutation Rate	5%, 10%
Objective Function	NPV, Mixed, Carbon Savings
Population Size	8, 16, 32

objective functions were tested. The NPV and LCS objectives simulated a single-objective stakeholder wishing to either maximize their personal financial returns or minimize the life-cycle carbon cost of the building. The mixed objective simulated a stakeholder placing a value of £50 per tonne of relieved CO_2e.

The tuning was performed on a sample of 300 buildings selected using latin hypercube sampling across archetypes, size and initial state, with a grid size of 144 variable combinations. The optimization was replicated 30 times for each building-setting combination for a total of 1,296,000 optimization runs. The results were analyzed using the best of the 10 replications, with the least variant setting being selected where the same maximum was found. Universally the best performing optimizations had the higher population size of 32 and a higher mutation rate of 10%. However, mixed and NPV focussed objectives performed better with randomized mutations and a pure elitism selection method, while carbon focussed objective functions performed better with random mutation and soft elitism. As such, during front generation, GA settings were set in accordance with their objective function. Recombination method and mutation rate were found to be the least significant settings. The results also found that within 30 replications, most of the settings (with the exception of the small population sizes) were able to find a solution close to the maximum, however worse performing solutions had significant variance.

3.5. Front generation

In order to train the SO, a representative set of Household Carbon Value (HCV) linked Pareto fronts were generated. Pareto solutions were generated by altering the HCV provided to the objective function. HCV bounds were generated by using a NPV and carbon-only objective to find the extreme points of the front. A grid search was then performed between these bounds, and a local search performed around identified solutions. An example of the type of front generated by this method is shown in Figure 5. Solutions can be seen clustered around general intervention type combinations, with local non-dominated alternate solutions resulting from gradual variation in insulation thickness. Using this front generation technique, in contrast to a multi-objective genetic algorithm method, provides a one to one relationship between the HCV and optimized solutions for SO training.

3.6. Modelling scenario

Given the large number of inputs that the SO system has at this point, some structure is required to consider the parameters that can and can't be tested at different levels of computational cost. Model training occurs at two different stages in the process, first the SEPM then the SO itself. Any parameters set to generate training data, but not explicitly used in the model itself, become exogenous and fixed after the model is trained. This makes the surrogate model parameters very costly to sweep for urban-scale data sets as they require the retraining of both the SEPM and the SO for each value of each parameter, or for the parameter to be integrated into the model, which could potentially damage the accuracy of the model or require a significant increase in training set size.

Fortunately, the SO's parameters are less computationally expensive to sweep, as the first stage SEPM inputs can be held fixed. This allows for the reuse of the trained model. These are also the higher order parameters as instead of pertaining to the minutiae of individual buildings, the Surrogate Optimizer parameters relate to the environment in which retrofits are optimized. These areas, unlike the physical laws that dominate energy performance, are more likely to be economic or social in nature and can therefore be enshrined into policy.

Examples of interventions that could affect optimization parameters include:

(1) A carbon tax
(2) A subsidy on the cost of retrofits
(3) Direct provision of certain retrofit solutions
(4) Funding research and development into improved retrofit solutions
(5) Fines for poorly performing properties
(6) Restriction of renting or selling poorly performing properties
(7) National grid decarbonization

All of these proposed scenario spaces relate, with varying degrees of directness, to the inputs used by the GA when optimizing retrofits of each building. As such, they become parameters with which different SOs can be trained to evaluate the outcomes of such policies or interventions. There are also examples of non-policy factors that affect the outcome of the optimization phase:

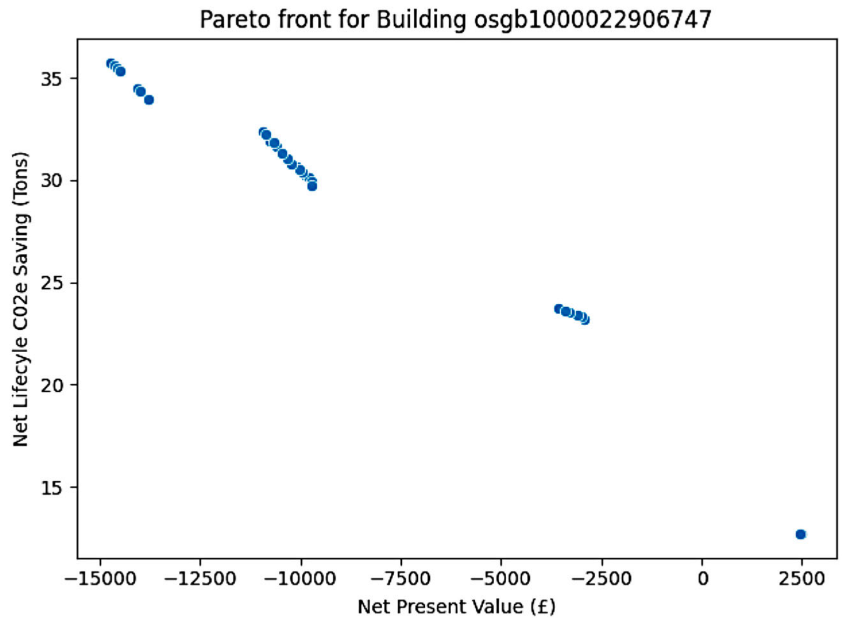

Figure 5. Sample of Pareto front generated with genetic algorithm.

(1) The market cost of materials and labour
(2) The market cost of fuel
(3) Changes in technology due to innovation

3.7. Stage 5: training of surrogate optimizer

The SO was trained using the data set of near-optimal WHRSs found using the SEPM-based optimization stage. A data flow showing the model features, drawn from building, household, and solution attributes are shown in Figure 6. The carbon value of a given household was paired to a given WHRS for that dwelling, accounting for the environmental preferences of the household as well as physical properties.

3.7.1. Model selection

As with the SEPM, the SO was parameter tuned using a grid search method, the setting combinations for which can be found in Table 3. The same pre-processing methods: min-max normalization and one-hot encoding were found optimal. Initial grid settings reflected the choices found optimal in related work, adjusted for input parameter size, and only minor differences were selected during the tuning process (Thrampoulidis et al. 2021). Sample size selection was kept more conservative than for the SEPM. Given the relative decline in marginal performance found in the model training as sample size increased, a set of approximately 20,000 total solutions was used, comprized of 1850 entire sets of non-dominated Pareto fronts.

3.8. Stage 6: housing stock analysis

Once the SO has been trained, near-optimal retrofits for a given HCV can be generated quickly and scored using the same objective values that were used for training, as laid out in Figure 7. Static analysis of households using a set HCV distribution can give a snapshot of how many positive utility retrofits exist in the chosen stock. The speed performance benefits of the carbon linked SO method are even more apparent when stochastic and temporal analysis of the stock is required, such as when coupled with an agent-based model as described below.

3.9. Agent-based model design

In order to demonstrate the value of exposing HCV as a model input, we have devised a simple, illustrative ABM to assess the impact of a local government campaigning for climate awareness on the quantity and quality of retrofits within the city. In the model, agents are endowed with a HCV $v_i \sim \mathcal{U}(0, 40)$ drawn uniformly between 0 and 50. If influenced by an environmental campaign, which occurs at a probability p per year, their HCV is increased by $n \sim \mathcal{U}(0, 15)$. These valuations represent approximations of the mean and variance of WTP found in carbon valuation studies (MacKerron et al. 2009; Rotaris, Giansoldati, and Scorrano 2020). Each year, agents evaluate their retrofit options at probability e. The household's near-optimal retrofit solutions are generated using the trained SO and scored as described in Figure 7. Solutions bearing positive utility are then retrofitted. The modelling scenarios give the local government two approaches to influence retrofits: Increasing the intensity of environmental

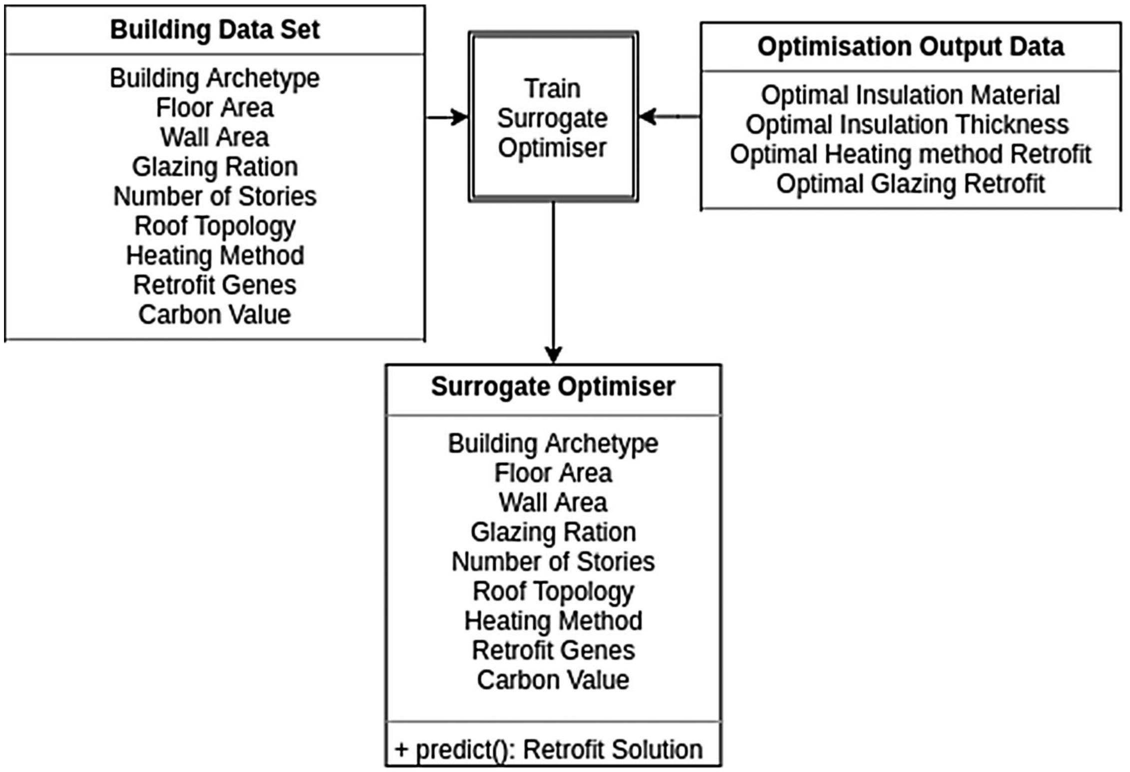

Figure 6. Conceptual data flow for training Surrogate Optimizer using solutions from the optimization stage (Stage 5).

campaigning, and direct outreach to encourage consideration of retrofits among households (through means such as payback finance structures, working with installers, and contacting households, which are abstracted into the evaluation rate). The scenarios modelled are shown in Table 5, and the state diagram representations of the agents are shown if Figure 8. With a domestic building stock of 95,500 dwellings, the outreach scenarios require an average of 47,750 optimizations per 25-year simulated run, representing a large computational cost that could become prohibitive using traditional optimization approaches.

Table 5. Intervention scenarios for agent based model.

Scenario	Evaluation rate (e)	Campaign penetration rate (p)
Baseline	0.01	0
Small campaign	0.01	0.05
Small campaign with outreach	0.02	0.05
Significant campaign	0.01	0.1
Significant campaign with outreach	0.02	0.1

4. Results

4.1. Energy simulation results

The purpose of running the EnergyPlus simulations is to generate a data set for SEPM training. The validity of these simulations, which is to say the closeness of the outcomes of the models to the target systems they represent, is of relatively minor importance for the methodological study provided the complexity of the simulation system is present. This is especially true as both the physical simulation process itself, and the generation of SEPMs from its results, are well established as valid within the literature. As such, a simple statistical validation technique was used to ensure that the distribution of simulation results broadly matched the distribution of heating demand values found for alternative sources.

The median national space heating energy consumption in the UK was reported at 10,118 kWh per household in 2019, an equivalent of 107 kWh per m^2 (Odysee Project 2019). These average values are slightly higher than the simulated, non-retrofitted simulation values of the urban building stock sample of 10,030.8 kWh per household and 103.8 kWh per m^2. Given the national figures were generated from a nationwide sampling methodology, rather than a bottom-up simulation methodology, this variation in median outcomes of approximately 1% and 3% per dwelling, and per unit area energy usage, was deemed to be within an acceptable margin. The distribution of heating energy demand values is in line with the most recent whole stock energy distribution data found, although the data is from an older 2009 survey in which the median energy consumption was higher (The UK government 2013). The distribution

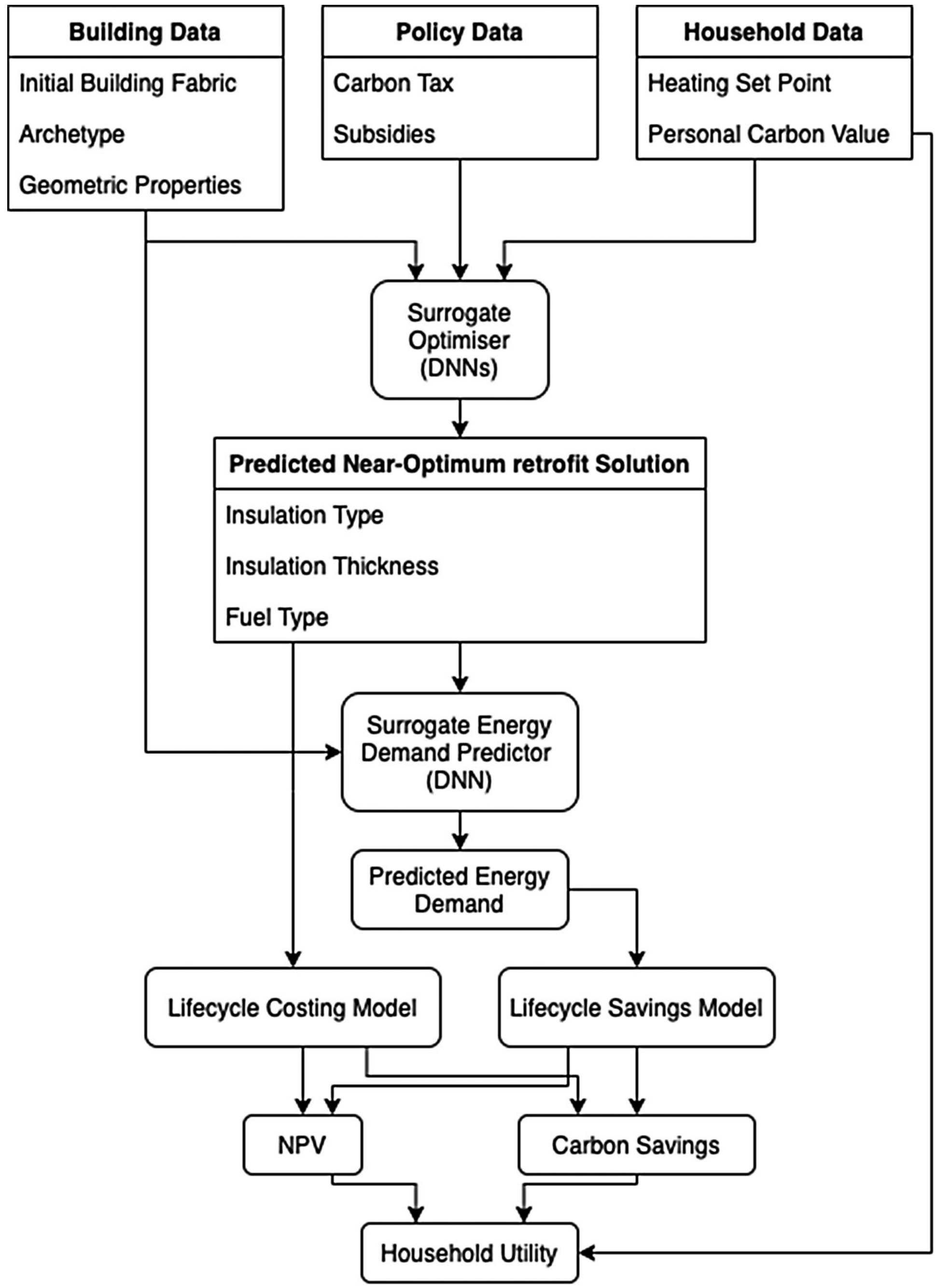

Figure 7. Process flow for generating and scoring retrofit solutions using trained surrogate optimizer.

of existing stock annual heating energy demand is shown in Figure 9.

4.2. Surrogate energy performance model results

The final SEPM was trained on a data set of 40,000 valid Whole House Retrofit Solutions (WHRSs), split into a training, validation, and test set of 70%, 15%, and 15% respectively. To prevent over-fitting, training was stopped after the validation loss stalled out, with a threshold of 0.001 over 50 epochs. The building and retrofit genes were selected using repeated random sampling. The mean absolute percentage error of the trained model when applied to the test data was 2.78%, with an adjusted r^2 value of 0.986 and a Mean Absolute Error (MAE) of 439 kWh per year. Model performance continued to increase with sample size until the full data set of 40,000 was used, suggesting performance could be increased further. However, training was stopped here as the improvement in MAE per 1000 additional samples fell below significance at a 1% level. The SEPM training process is visualized in Figure 10, showing the effect of

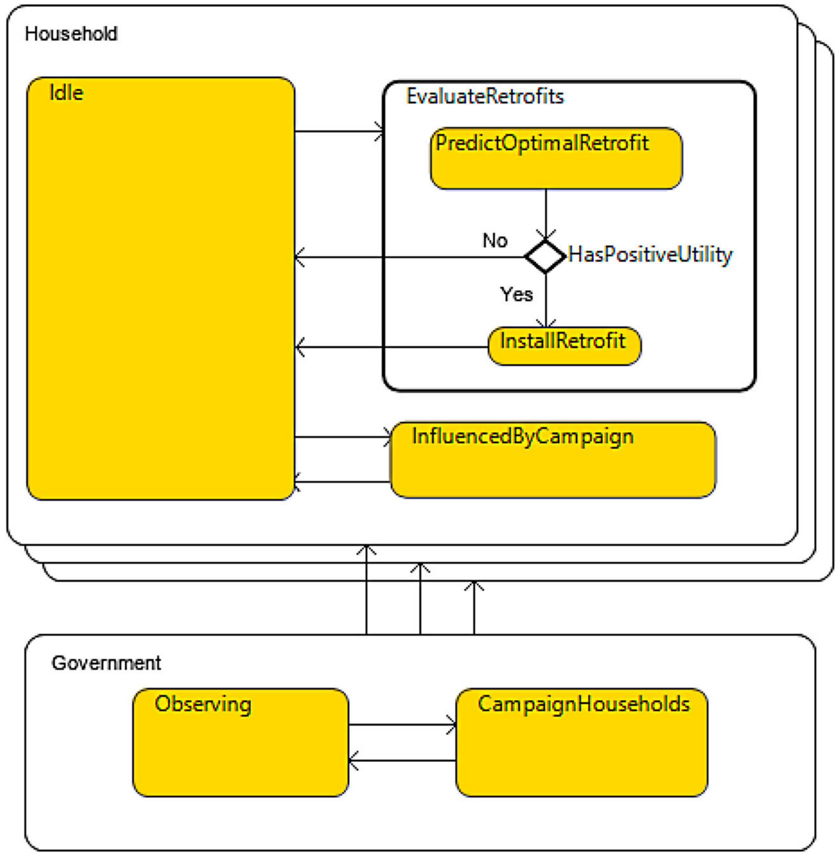

Figure 8. State chart representation of agents.

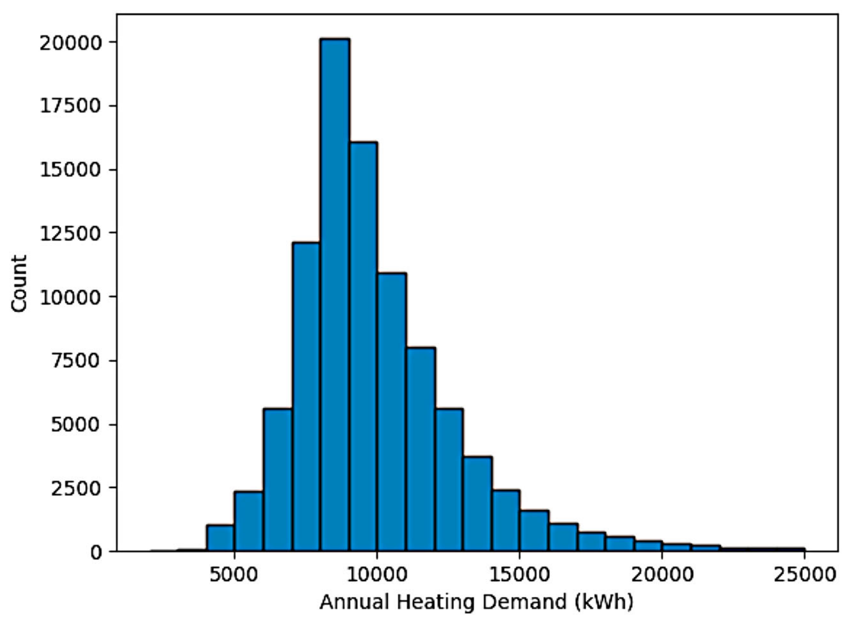

Figure 9. Distribution of simulated household annual space heating energy demand (kWh) for Nottingham building stock ($n = 95{,}500$).

increased sample size as well as training time on model performance. As shown in Figure 11, the residuals of the trained SEPM were normally distributed and centered around 0.

4.3. Generic algorithm results

The established optimization method, a genetic algorithm with an embedded SEPM, was used to create a

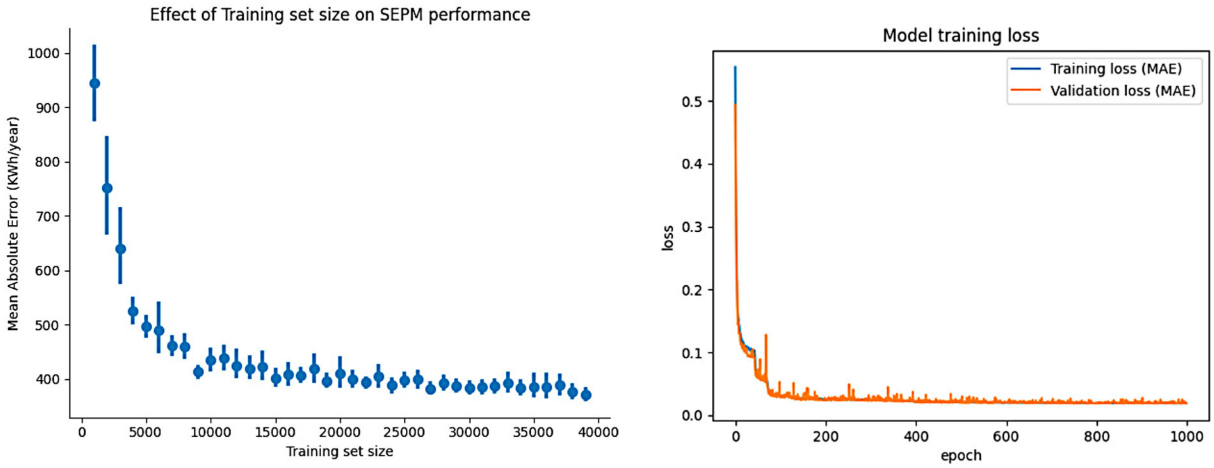

Figure 10. Performance of surrogate energy performance model during training process.

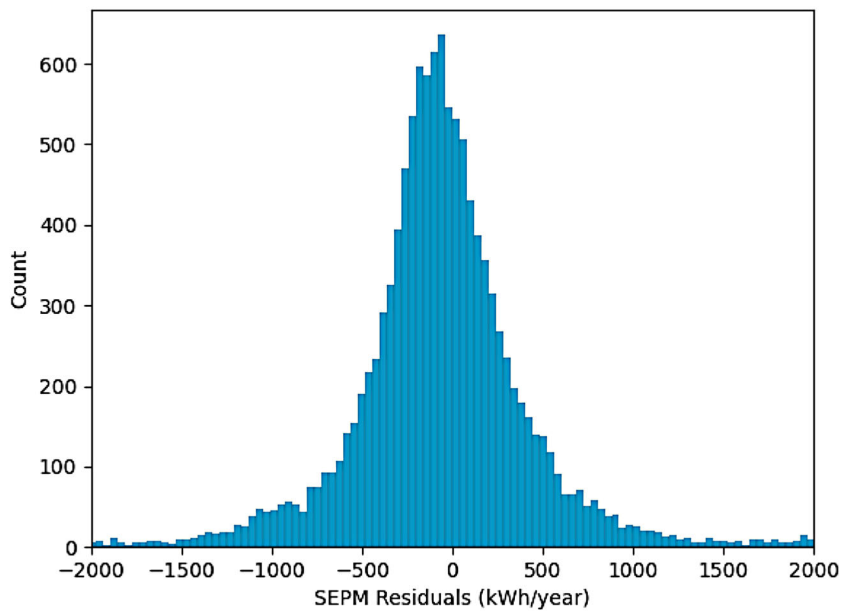

Figure 11. Surrogate energy performance model residuals are normally distributed and centered around 0.

data set of 1900 household Pareto fronts, with an average front size of 12.7 solutions for a total of 24,130 WHRSs. These solutions were categorized into three household scenarios for analysis: those with a HCV of £0/$tCO2_e$ (Max NPV), those greater than £0/$tCO2_e$ (Mixed Criteria), and a LCS maximizing (Max LCS) group. Figure 12 shows the range of WHRSs generated at this stage, coloured by the retrofit measures that form them. The most profitable retrofits involved fuel-type changes from electric space heating to the cheaper gas source. In contrast, more LCS-focussed solutions involved the opposite transition, due to the steep grid decarbonization targets considered in this scenario. Considering glazing, LCS-focussed solutions were divided between remaining double glazed and fitting triple glazed, suggesting that for some properties the embodied carbon of improved glazing would not be paid back within the lifecycle of the measure. Similarly, NPV maximizing solutions rarely included triple glazing insulation, suggesting this measure is also not financially viable for most households. Considering wall insulation, internal solutions formed of EPS with a mean thickness of 0.042 mm were favoured when NPV was a primary objective, suggesting a greater cost to performance ratio. When LCS was the household's priority, however, external solutions of maximum permitted thickness formed of PIR dominated, indicating that their performance increases continued to outpace embodied carbon in the range considered. The Cumulative Distributions of optimized insulation thickness, broken down by household preference scenario can be seen in Figure 13, showing that LCS-focussed objectives resulted in significantly more installations of maximum thickness. However, nearly half of

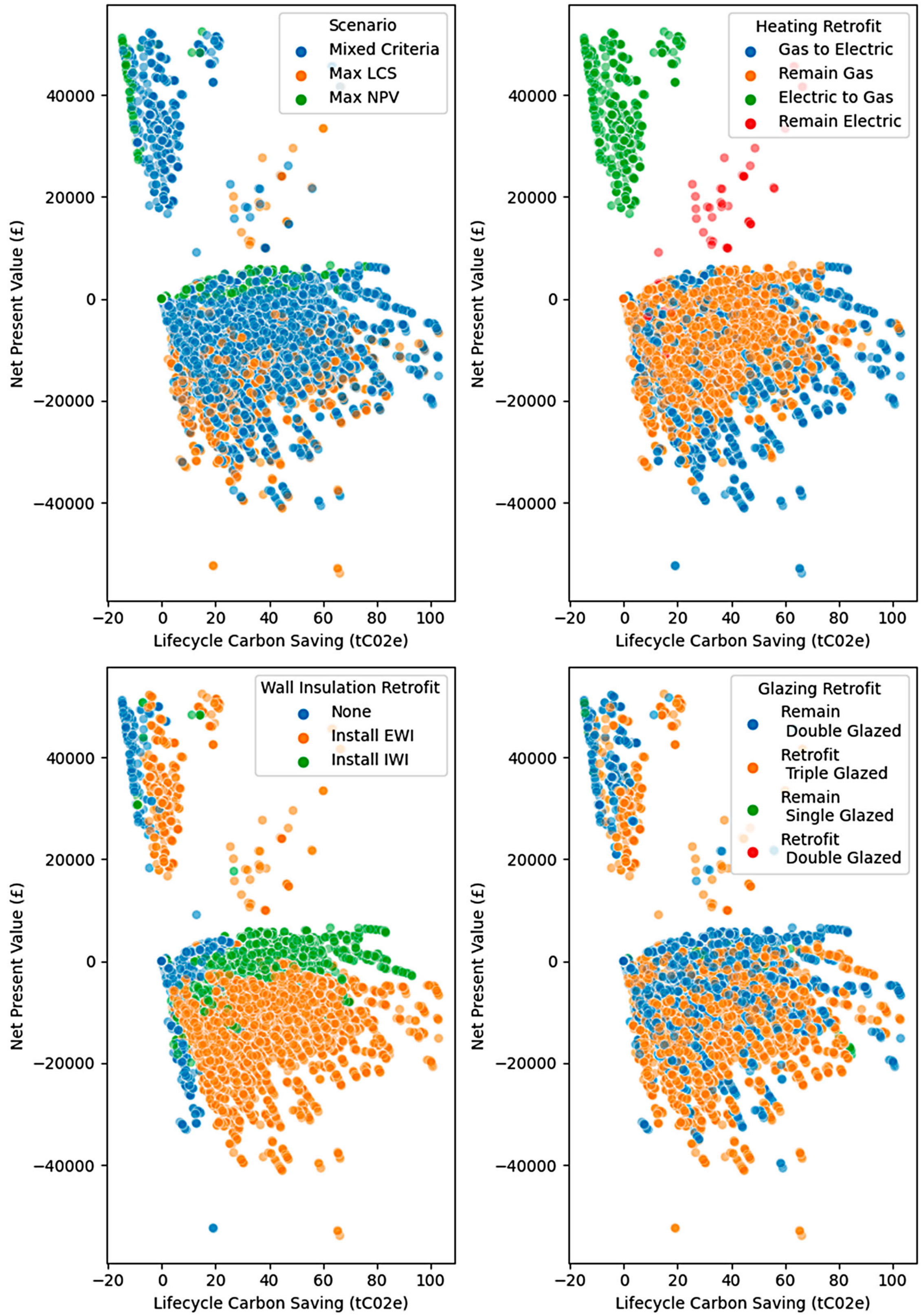

Figure 12. Trade-off solutions discovered using traditional optimization method on 1900 households.

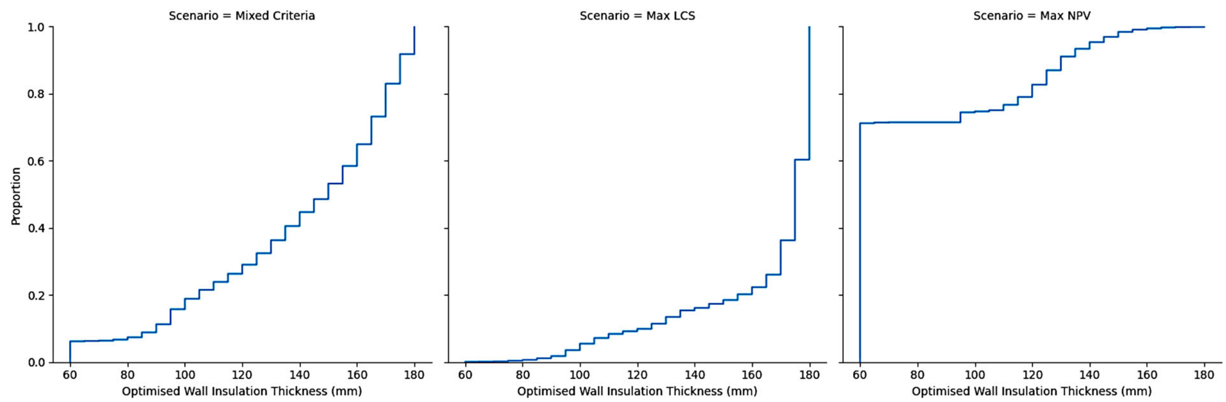

Figure 13. Cumulative distribution of optimized wall insulation thickness under different household scenarios.

installations were sub-maximal, suggesting the embedded carbon cost of thicker installations was eventually offset by the falling efficiency at higher thicknesses.

From this sample of 1900 dwellings, 763 (40%) of dwellings had a positive NPV solution, making the solution viable even for environmentally indifferent households. 712 (93%) of these resulted in a positive LCS, making them viable independent of household carbon preferences. Where a financially viable retrofit was found, the average NPV was £4399. In the Max LCS scenario, savings were positive for all dwellings, indicating embodied carbon was offset for at least one solution in all household cases. NPV maximizing solutions were most likely to contain a single measure, while most Max LCS solutions involved retrofits across all available measures.

4.4. Surrogate optimizer results

The raw SO results break down into the classification and regression model components. The classification models obtained adjusted f1 scores of 0.95, 0.93, 0.94, 0.95, and 0.79 for IWI material, EWI material, roof material, fuel source, and glazing respectively. This suggests the classification components function quite well. The poorer performance of glazing predictions predominately occurred at near LCS-maximizing solutions, which is unsurprizing given the traditional optimization phase was also inconsistent on whether triple glazing was carbon effective for the lifecycle considered. The material length genes were trained with regression output nodes (linear) and obtained mean absolute errors of 4.9 mm, 4.3 mm, and 7.6 mm for IWI, EWI, and roof insulation thicknesses respectively, placing the mean error within one thickness gene value of the test data.

We performed a static analysis of the entire building stock using the same scenario settings as were used in the traditional optimization method. In the NPV-maximizing scenario a positively performing retrofit was found for 35% of dwellings, compared with 42% of the traditionally optimized sample, suggesting 83% of these solutions were found successfully. The mean NPV of these solutions was £1780 greater, suggesting the 17% of unidentified solutions were of smaller NPV benefit, with the best performing solutions being identified more easily. The composition of solutions remained consistent in the mixed criteria and LCS-maximizing scenarios. LCS-maximizing solutions under-performed the carbon reduction of traditional optimization by only 6%, with most errors found in glazing classification (which were of minor effect).

Given that simulation time considerations are the primary benefit of a SO approach, it is worth comparing the run time between the SO output and the traditional optimization method. The original GA optimization took on average 25.3 cpu seconds to find a near-optimal WHRS for a given HCV scenario, making the evaluation of the entire stock take an estimated 671 cpu hours per HCV considered. In contrast, generating and scoring solutions for the entire building stock with the trained SO took 31.2 cpu seconds, allowing for rapid analysis of different HCVs and building scenarios without major computational cost. As with all surrogate modelling techniques, the loss of accuracy is traded off for this greater speed performance.

4.5. ABM results

The ABM utilized the ability of a trained SO to quickly evaluate the entire building stock of the city of Nottingham. The ABM was replicated 100 times per scenario. As shown in Figure 14, the campaigning to increase HCV of households increased the LCS of retrofit decisions made over time. The types of retrofits installed differed in the campaign scenarios, as shown in Figure 15, with increased campaigning resulting in a preference for more environmentally friendly retrofit options. This resulted in an average LCS of 25.51 tCO_2e per retrofit in the significant campaign scenario, contrasting with 25.21 tCO_2e in the

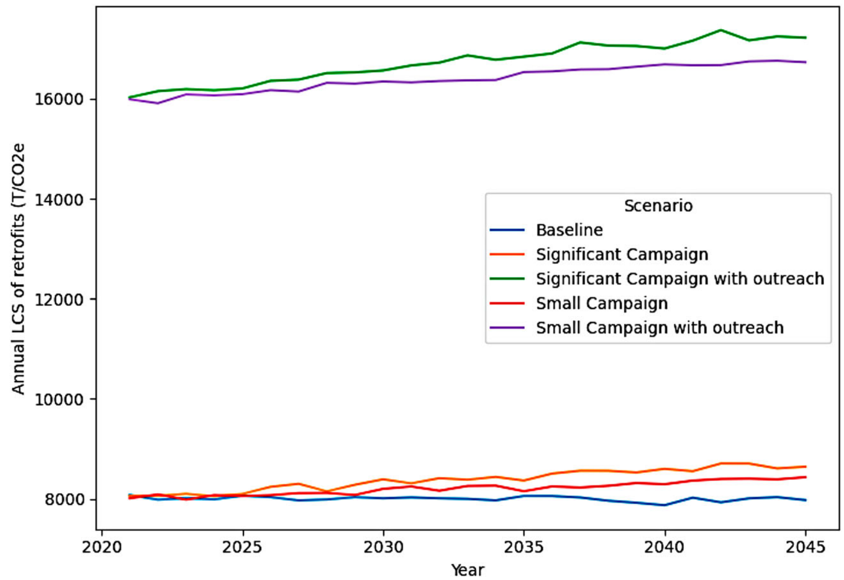

Figure 14. Lifecycle Carbon emissions savings of simulated retrofit decisions over time.

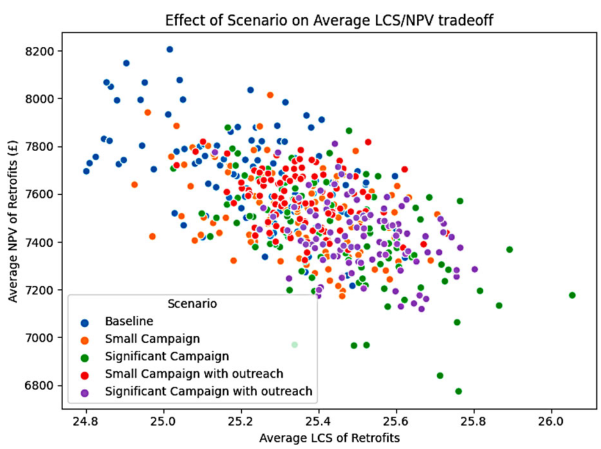

Figure 15. Average net present value and lifecycle carbon savings of retrofits in agent-based model under different scenarios.

baseline ($p < 0.01$). This came at the cost of a reduction in NPV from £7696 to £7429, showing a willingness of households with higher HCV to trade off financial value for more environmental retrofits.

In total, a significant campaign reduced net lifecycle emissions by 13.7 $ktCO2_e$ without the outreach programme, and 16.7 $ktCO2_e$ when outreach was included. The number of positive utility retrofits was also higher under campaign scenarios, with average increases of 2.9% and 6.3% for small and significant campaigns respectively. This indicates the installation of marginal retrofits which would not have been viable for financial reasons alone. While the effects are small on the household level (due to the relatively moderate changes in HCV) they show that on the urban scale the increased emissions awareness can have significant impacts over a long period. In this scenario, the evaluation rate of households, influenced by the outreach campaign, had the most significant effect on the total carbon reduction. This suggests that attempts to engage households to trigger a WHRS evaluation should be a priority ahead of campaigning, to increase their carbon valuation directly. However, the combined approach yielded the greatest reduction in total emissions.

5. Discussion

The objective of this research was to extend the recently developed SO method by adding an input feature based on the preferences of the decision-maker under consideration. This extension allows for the use of SOs in situations where preference information is important for definitive retrofit selection. Whilst this is applied to environmental preferences in the context of this work, the method could be applied to any retrofit outcomes for which willingness to pay data is available, such as measures of thermal comfort or degree of disruption caused by the installation. We demonstrate the utility of the method by first constructing the SO, then integrating it into a simple ABM, in which, in response to simulated environmental campaigns, agents' environmental preferences changed over time. This allowed for the assessment of such campaigns on the building stock which would not be computationally feasible without the developed method. The evaluation of the SO was found to be approximately 100,000 times quicker than using a traditional SEPM-based optimizer. This was of great benefit when running the ABM, as the 14,250,000 retrofit evaluations required for the scenarios would have taken an estimated time of 99,000 cpu hours using the GA and SEPM method, compared with just 1.25 cpu hours using the SO. This increase in computational feasibility for building stock models with explicit household carbon valuations was achieved with minimal loss of accuracy when predicting near-optimal retrofit choices, as demonstrated by the high classification performance of the SO on the test data. The worst performing SO retrofit category was the decision between double and single glazing, particularly for highly environmentally conscious agents. This is a reflection of the uncertainty of this decision under traditional optimization methods, as the environmental cost is close to the benefit in many cases, making it a marginal retrofit decision.

Limitations of this proof of principle study include simplifications and assumptions when implementing the SO method. Fuel sources were limited to electric and gas central heating and retrofit options excluded generative methods such as photovoltaic installations. These simplifications were made to reduce the complexity of the energy simulations due to data constraints, as estimation of factors such as roof tilt would introduce significant error at the simulation stage. These simplifications mean that a reduced set of retrofit options are presented to simulated householders compared with those available at the point of a real retrofit installation decision. A further limitation of the research was in the number of objective values, which were limited to NPV and LCS when in reality retrofit decision makers have a larger set of varying and contradictory objectives and constraints. This could introduce prediction error in instances where neither financial or environmental motives are of most significance to a given household. Increasing the number of objective values at the SO stage would pose a challenge, although this could be done by imposing constraints during the application phase, such as a retrofit trigger model used as part of the ABM design. The study also only considered heating demand due to the rarity of domestic cooling in the UK, but the implementation would need adapting to fit warmer climates at the simulation of energy demand phase. As a result of this limitation, the method as applied in this work is only sound in climates with limited risk of overheating and would need adapting to account for domestic cooling demand. Situations where the risk of both overheating and under-heating are present would also pose challenges, requiring a thermal comfort consideration beyond heating demand.

Future work should focus on integrating the trained SO into more sophisticated and realistic ABMs. In particular, this could be integrated with retrofit decision models, as the SO could generate solutions when a retrofit decision is likely to be triggered. These solutions would then be accepted or rejected by the retrofit decision model based on further household criteria. The method could also be extended to include a wider range of retrofit alternatives to allow more sophisticated installations to be considered, provided the data and computational capacity are available to perform the required energy evaluations. The integration of financial constraints into the model training would also remove the importance of post-processing solutions to ensure they meet a household's financial abilities.

6. Conclusion

In order to reduce the computational cost of domestic energy retrofit optimization, surrogate models have been used to evaluate the energy performance of buildings. Recent methods of training neural networks on a sample of near-optimal retrofit solutions, referred to in this paper as surrogate optimization, have been used to further increase the speed at which a large number of quality retrofit solutions can be discovered. However, these predictive models are not capable of making definitive predictions of the actions that decision makers with specific environmental preferences would make. This limits their utility in certain applications, such as embedding into agent-based models using dynamic synthetic populations. In this paper, we presented an extension to the emerging practice of surrogate optimization to find whole house retrofit solutions of building stock at the urban scale, performing the method in a way which exposes the household carbon valuations of the retrofit decision makers. This value, which represents the willingness to pay per tonne of $CO2_e$ emissions mitigated, was preserved through the optimization process and then used in conjunction with the other input variables to train a predictive surrogate optimizer capable of producing candidate near-optimal solutions. By including the household carbon valuation when training the predictive model, we are able to analyze the impact of households' changing attitudes to carbon mitigation. The exposure of this input variable allowed us to construct a simple agent-based model of a local government performing an environmental awareness campaign, which would not have been possible with prior methods. The effects of the campaign on households' carbon valuation affected both the quality and number of retrofits those households performed when evaluating retrofit decisions using the surrogate optimizer. Future work should focus on reducing some of the simplifications built into the model at the simulation phase by expanding the range of retrofits considered and integrating the method into a more sophisticated agent-based model to account for decision triggers and constraints not accounted for in the current model. This additional level of detail would allow for more confident retrofit adoption predictions concerning policies designed to increase environmentalism among households at the urban scale.

Notes

1. Version 9.0.1
2. in a similar manner to that used by the dynamic housing stock energy simulation platform *EnHub* (Sousa et al. 2018)
3. This rate was set to approximate the risk-free rate of return in the UK. While historically this has been above 2% (Siegel 1992), recent yields have plateaued to below 1% (Rachel and Smith 2017).
4. Using a reduction rate formula, a value of 0 would not be achieved in a finite time period, so 0.001 kg was used as an approximation

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by The Leverhulme Trust [RP 2013-SL-015].

ORCID

Peer-Olaf Siebers http://orcid.org/0000-0002-0603-5904

References

Ascione, Fabrizio, Nicola Bianco, Claudio De Stasio, Gerardo Maria Mauro, and Giuseppe Peter Vanoli. July, 2017. "CASA, Cost-optimal Analysis by Multi-objective Optimisation and Artificial Neural Networks: A New Framework for the Robust Assessment of Cost-optimal Energy Retrofit, Feasible for Any Building." *Energy and Buildings* 146: 200–219. doi:10.1016/j.enbuild.2017.04.069.

Boustead, I. 2005. *Eco-Profiles of the European Plastics Industry*. Technical Report. PlasticsEurope. www.plasticseurope.org.

Brekke, Kjell Arne, Snorre Kverndokk, and Karine Nyborg. September, 2003. "An Economic Model of Moral Motivation." *Journal of Public Economics* 87 (9–10): 1967–1983. doi:10.1016/S0047-2727(01)00222-5.

Brownlee, Alexander E. I., Jonathan A. Wright, Miaomiao He, Timothy Lee, and Paul McMenemy. November, 2020. "A Novel Encoding for Separable Large-scale Multi-objective Problems and Its Application to the Optimisation of Housing Stock Improvements." *Applied Soft Computing* 96: Article ID 106650. doi:10.1016/j.asoc.2020.106650.

Chaudhuri, A. 2011. "Sustaining Cooperation in Laboratory Public Goods Experiments: a Selective Survey of the Literature." *Experimental Economics* 14: 47–83.

Costa, Alberto, Giacomo Nannicini, A. Costa, Eth Zurich, and G. Nannicini. 2018. "RBFOpt: An Open-source Library for Black-box Optimization with Costly Function Evaluations." *Mathematical Programming Computation* 10: 597–629. doi:10.5281/zenodo.597767.

Crawley, Drury B., Curtis O. Pedersen, Linda K. Lawrie, and Frederick C. Winkelmann. 2000. "EnergyPlus: Energy Simulation Program." *ASHRAE Journal* 1 (Apr): 49–56. https://www.researchgate.net/profile/Drury-Crawley/publication/230606369_EnergyPlus_Energy_Simulation_Program/links/546a40d80cf2f5eb18077919/EnergyPlus-Energy-Simulation-Program.

Eames, M., T. Kershaw, and D. Coley. 2011. "On the Creation of Future Probabilistic Design Weather Years From UKCP09." *Building Services Engineering Research and Technology* 32 (2): 127–142. doi:10.1177/0143624410379934.

Ecosystem Marketplace. 2017. *Unlocking Carbon Market Potential in a Post-Paris World*. Technical Report. Ecosystem Marketplace. https://www.ecosystemmarketplace.com/articles/unlocking-carbon-market-potential-in-a-post-paris-world/.

Edwards, Richard E., Joshua New, Lynne E. Parker, Borui Cui, and Jin Dong. September, 2017. "Constructing Large Scale Surrogate Models From Big Data and Artificial Intelligence." *Applied Energy* 202: 685–699. doi:10.1016/j.apenergy.2017.05.155.

Eisenhower, Bryan, Zheng O'Neill, Satish Narayanan, Vladimir A. Fonoberov, and Igor Mezić. April, 2012. "A Methodology for Meta-model Based Optimization in Building Energy Models." *Energy and Buildings* 47: 292–301. doi:10.1016/j.enbuild.2011.12.001.

Evins, Ralph. 2013. "A Review of Computational Optimisation Methods Applied to Sustainable Building Design." *Renewable and Sustainable Energy Reviews* 22: 230–245. doi:10.1016/j.rser.2013.02.004.

Gustafsson, Stig-Inge, and Björn G. Karlsson. 1988. "Why is Life-cycle Costing Important when Retrofitting Buildings." *International Journal of Energy Research* 12 (2): 233–242. doi:10.1002/er.4440120206.

Hammond, Geoff, and Craig Jones. 2008. *Inventory of Carbon & Energy (ICE) Version 1.6a*. Technical Report. Bath: University of Bath. https://perigordvacance.typepad.com/files/inventoryofcarbonandenergy.pdf.

Henderson, J., and J. Hart. 2013. *BREDEM 2012-A Technical Description of the BRE Domestic Energy Model*. Technical Report. BRE. https://www.bre.co.uk/filelibrary/bredem/BREDEM-2012-specification.pdf.

Hey, James, Peer-Olaf Siebers, Paul Nathanail, Ender Özcan, and Darren Robinson. 2020. "Surrogate Optimisation of Housing Stock Retrofits using Deep Neural Networks". Proceedings of the 5th IBPSA England Conference on Building Simulation and Optimization (Virtual). Loughborough.

Isaac, R. Mark, and James M. Walker. February, 1988. "Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism." *The Quarterly Journal of Economics* 103 (1): 179. doi:10.2307/1882648.

Kavgic, M., A. Mavrogianni, D. Mumovic, A. Summerfield, Z. Stevanovic, and M. Djurovic-Petrovic. July, 2010. "A Review of Bottom-up Building Stock Models for Energy Consumption in the Residential Sector." *Building and Environment* 45 (7): 1683–1697. doi:10.1016/j.buildenv.2010.01.021.

MacKerron, George J., Catrin Egerton, Christopher Gaskell, Aimie Parpia, and Susana Mourato. April, 2009. "Willingness to Pay for Carbon Offset Certification and Co-benefits Among (high-)flying Young Adults in the UK." *Energy Policy* 37 (4): 1372–1381. doi:10.1016/j.enpol.2008.11.023.

Magnier, Laurent, and Fariborz Haghighat. March, 2010. "Multiobjective Optimization of Building Design Using TRNSYS Simulations, Genetic Algorithm, and Artificial Neural Network." *Building and Environment* 45 (3): 739–746. doi:10.1016/J.BUILDENV.2009.08.016.

Melo, A. P., D. Cóstola, R. Lamberts, and J. L. M. Hensen. 2014. "Development of Surrogate Models Using Artificial Neural Network for Building Shell Energy Labelling." *Energy Policy* 69: 457–466. doi:10.1016/j.enpol.2014.02.001.

Met Office. 2021. *Station History File for Watnall (Including Hucknall) (DCnn: 4206, Old Number: 316/614, WMO Number: 03354)*. Technical Report. National Meteorological Library & Archive. https://library.metoffice.gov.uk/portal/Default/en-GB/RecordView/Index/248997.

Odysee Project. 2019. *Sectoral Profile – Households Energy Consumption*. Technical Report. Odyssee-Mure.

Office For National Statistics. 2021. *Provisional UK Greenhouse Gas Emissions National Statistics*. https://data.gov.uk/dataset/9a1e58e5-d1b6-457d-a414-335ca546d52c/provisional-uk-greenhouse-gas-emissions-national-statistics.

Ordnance Survey. 2018. "OS Open TOID | OS Data Downloads | OS Data Hub." https://osdatahub.os.uk/downloads/open/OpenTOID.

Parliament of the United Kingdom. 2008. *Climate Change Act 2008*. Technical Report.

Parsons, Robert, ed. 2005. *Fundamentals: 2005 Ashrae Handbook*. Vol. 1, 1–42. Washington: American Society of Heating, Refrigerating and Air-Conditioning Engineers.

Prada, A., A. Gasparella, and P. Baggio. September, 2018. "On the Performance of Meta-models in Building Design Optimization." *Applied Energy* 225: 814–826. doi:10.1016/j.apenergy.2018.04.129.

Rachel, Lukasz, and Thomas D. Smith. August, 2017. "Are Low Real Interest Rates Here to Stay?." *International Journal of Central Banking* 50.

Reinhart, Christoph, and Carlos Cerezo Davila. 2019. "Urban Building Energy Modeling." In *Building Performance Simulation for Design and Operation*. 2nd ed., edited by Jan L. M. Hensen and Roberto Lamberts, 696–722. London: Routledge.

Robinson, Darren. 2019. "Integrated Resource Flow Modelling of the Urban Built Environment." In *Building Performance Simulation for Design and Operation*. 2nd ed., edited by Jan L. M. Hensen and Roberto Lamberts, 659–695. London: Routledge.

Robinson, Darren, Frédéric Haldi, Jérôme Henri Kämpf, P. Leroux, D. Perez, A. Rasheed, and U. Wilke. 2009. "CITYSIM: Comprehensive Micro-Simulation of Resource Flows for Sustainable Urban Planning." *International IBPSA Conference*. Glasgow.

Rosser, Julian F., Gavin Long, Sameh Zakhary, Doreen S. Boyd, Yong Mao, and Darren Robinson. March, 2019. "Modelling Urban Housing Stocks for Building Energy Simulation Using CityGML EnergyADE." *ISPRS International Journal of Geo-Information* 8 (4): 163. doi:10.3390/ijgi8040163.

Rotaris, Lucia, Marco Giansoldati, and Mariangela Scorrano. December, 2020. "Are Air Travellers Willing to Pay for Reducing Or Offsetting Carbon Emissions? Evidence From Italy." *Transportation Research Part A: Policy and Practice* 142: 71–84. doi:10.1016/j.tra.2020.10.014.

Sharif, Seyed Amirhosain, and Amin Hammad. September, 2019. "Developing Surrogate ANN for Selecting Near-optimal Building Energy Renovation Methods Considering Energy Consumption, LCC and LCA." *Journal of Building Engineering* 25: Article ID 100790. doi:10.1016/j.jobe.2019.100790.

Siebert, Horst. 2007. "Environmental Quality as a Public Good." In *Economics of the Environment*, 59–95. Berlin: Springer. https://link.springer.com/chapter/10.1007/978-3-540-73707-0_5.

Siegel, Jeremy J. April, 1992. "The Real Rate of Interest From 1800–1990. A Study of the U.S. and the U.K." *Journal of Monetary Economics* 29 (2): 227–252. doi:10.1016/0304-3932(92)90014-S.

Sola, Alaia, Cristina Corchero, Jaume Salom, and Manel Sanmarti. November, 2018. "Simulation Tools to Build Urban-Scale Energy Models: A Review." *Energies* 11 (12): 3269. doi:10.3390/en11123269.

Sousa, Gustavo, Benjamin M. Jones, Parham A. Mirzaei, and Darren Robinson. August, 2018. "An Open-source Simulation Platform to Support the Formulation of Housing Stock Decarbonisation Strategies." *Energy and Buildings* 172: 459–477. doi:10.1016/j.enbuild.2018.05.015.

Sousa, Gustavo, Benjamin M. Jones, Parham A. Mirzaei, and Darren Robinson. September, 2017. "A Review and Critique of UK Housing Stock Energy Models, Modelling Approaches and Data Sources." *Energy and Buildings* 151: 66–80. doi:10.1016/j.enbuild.2017.06.043.

Streimikiene, Dalia, Tomas Balezentis, Ilona Alisauskaite-Seskiene, Gintare Stankuniene, and Zaneta Simanaviciene. April, 2019. "A Review of Willingness to Pay Studies for Climate Change Mitigation in the Energy Sector." *Energies* 12 (8): 1481–1519. doi:10.3390/en12081481.

Sus, Xing, Shaochen Tian, Xiaolu Shao, and Xuan Zhao. September, 2020. "Embodied and Operational Energy and Carbon Emissions of Passive Building in HSCW Zone in China: A Case Study." *Energy and Buildings* 222: Article ID 110090. doi:10.1016/j.enbuild.2020.110090.

Thrampoulidis, Emmanouil, Georgios Mavromatidis, Aurelien Lucchi, and Kristina Orehounig. January, 2021. "A Machine Learning-based Surrogate Model to Approximate Optimal Building Retrofit Solutions." *Applied Energy* 281: Article ID 116024. doi:10.1016/j.apenergy.2020.116024.

Tian, Wei, Ruchi Choudhary, Godfried Augenbroe, and Sang Hoon Lee. 2015. "Importance Analysis and Meta-model Construction with Correlated Variables in Evaluation of Thermal Performance of Campus Buildings." *Building and Environment* 92: 61–74. doi:10.1016/j.buildenv.2015.04.021.

Tseranidis, Stavros, Nathan C. Brown, and Caitlin T. Mueller. December, 2016. "Data-driven Approximation Algorithms for Rapid Performance Evaluation and Optimization of Civil Structures." *Automation in Construction* 72: 279–293. doi:10.1016/J.AUTCON.2016.02.002.

The UK government. 2008. *Climate Change Act 2008*. https://www.legislation.gov.uk/ukpga/2008/27/section/2.

The UK government. 2013. *Report 4: Main Heating Systems*. Technical Report. Department of Energy and Climate Change. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/274772/4_Main_heating_systems.pdf.

Waibel, Christoph, Thomas Wortmann, Ralph Evins, and Jan Carmeliet. March, 2019. "Building Energy Optimization: An Extensive Benchmark of Global Search Algorithms." *Energy and Buildings* 187: 218–240. doi:10.1016/J.ENBUILD.2019.01.048.

Wang, Danhong, Jonas Landolt, Georgios Mavromatidis, Kristina Orehounig, and Jan Carmeliet. June, 2018. "CESAR: A Bottom-up Building Stock Modelling Tool for Switzerland to Address Sustainable Energy Transformation Strategies." *Energy and Buildings* 169: 9–26. doi:10.1016/J.ENBUILD.2018.03.020.

Waters, L.. 2020. *Energy Consumption in the UK (ECUK) 1970 to 2019*. Technical Report. Office for National Statistics. https://www.gov.uk/government/collections/digest-of-uk-energy-statistics-dukes.

Waters, Liz. July, 2019. *Energy Consumption in the UK (ECUK) 1970 to 2019*. Technical Report. Office of National Statistics.

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1020152/2020_Energy_Consumption_in_the_UK__ECUK_.pdf.

Wortmann, Thomas. 2018. "Genetic Evolution Vs. Function Approximation: Benchmarking Algorithms for Architectural Design Optimization." *Journal of Computational Design and Engineering* 6: 414–428. doi:10.1016/j.jcde.2018.09.001.

Wu, Raphael, Georgios Mavromatidis, Kristina Orehounig, and Jan Carmeliet. 2017. "Multiobjective Optimisation of Energy Systems and Building Envelope Retrofit in a Residential Community." *Applied Energy* 190: 634–649. doi:10.1016/j.apenergy.2016.12.161.

Zakhary, Sameh, Andrew Allen, Peer-Olaf Siebers, and Darren Robinson. October, 2016. "A Computational Workflow for Urban Micro-Simulation of Buildings' Energy Performance." In *Urban Transitions Global Summit*, Shanghai.

Zhang, Boyu, Xinmiao An, and Yali Dong. February, 2021. "Conditional Cooperator Enhances Institutional Punishment in Public Goods Game." *Applied Mathematics and Computation* 390: Article ID 125600. doi:10.1016/j.amc.2020.125600.