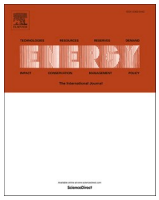
Energy 330 (2025) 136869

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy



# Informing heating electrification and prosumer-based solutions toward decarbonisation targets via a multi-objective optimization in grid-connected PV systems in rural buildings

Rundong Liao [a], Massimiliano Manfren [b], Benedetto Nastasi [c,*]

[a] College of Electrical Engineering, Sichuan University, Chengdu 610065, China
[b] Architecture, Built Environment and Construction Engineering A.B.C., Politecnico di Milano, Via Bonardi 9, 20133, Milano, Italy
[c] Department of Industrial Engineering, Tor Vergata University of Rome, Via del Politecnico 1, 00133, Rome, Italy

ARTICLE INFO

Handling Editor: Ruzhu Wang

*Keywords:*
Multi-objective optimisation
Heating electrification
Rural buildings
Photovoltaic systems
Heat pumps
Decarbonisation
Embodied carbon

ABSTRACT

In pursuit of rapid building stock decarbonisation, the electrification of heating through heat pumps, complemented by grid-connected photovoltaic (PV) systems and battery storage, has emerged as a viable solution. This study presents a concise, scalable energy system modelling framework that integrates a regression-based module for dynamic load estimation with an optimisation module to determine optimal system sizing and operation. By incorporating temperature-dependent heating and cooling demand alongside prosumer-based strategies, the framework provides insights into supply-demand balancing, cost considerations, and emissions reduction potential. For rural buildings in representative locations across China's diverse climate zones (Harbin, Beijing, Chengdu and Xiamen), the strategic integration of PV systems, heat pumps, and batteries has been shown to yield substantial carbon emissions reductions. Under scenarios achieving an optimal levelised cost of electricity (LCOE) below 0.04 $/kWh, all 4 regions can realise at least a 50 % reduction in emissions; with a moderate to medium increase in cost (24–60 %), reductions of up to 90 % become feasible. However, if the PV and battery systems' components are oversized, the carbon emissions produced during their manufacturing significantly reduce the benefits of displacing high-carbon grid electricity. Accounting for this embodied carbon burden, the maximum net reductions attainable are 79–85 % for Xiamen, 57–73 % for Chengdu, 75–84 % for Beijing and 72–82 % for Harbin. Overall, this research underscores the transformative potential of coupling electrification with prosumer-based approaches to meet ambitious decarbonisation targets and could be further developed with the assessment of climate-change impacts, dynamic electricity pricing, and community-scale solutions.

Nomenclature

| Symbol | Description | Unit |
|---|---|---|
| $A$ | Net area of conditioned area | $m^2$ |
| $A_{base}$ | Area of building base | $m^2$ |
| $A_{gi}$ | Surface area of each external window | $m^2$ |
| $A_r$ | Surface area of roof | $m^2$ |
| $A_{wi}$ | Surface area of each external wall | $m^2$ |
| $A^{PV}$ | Area of PV array | $m^2$ |
| $AM$ | Relative air mass | – |
| $AM_0$ | Reference air quality | – |
| $B$ | Blending factor | – |
| $B_b$ | Present values of battery replacement cost | $/kWh |
| $B_L$ | Lower blending limits | °C |
| $B_{max}$ | Maximum blending factor | – |

(*continued on next column*)

(*continued*)

| | | |
|---|---|---|
| $B_U$ | Upper blending limits | °C |
| $B_s$ | Shape coefficient of building | – |
| $C$ | Specific heat capacity of water | kJ/kg°C |
| $C_r$ | Heat loss coefficient of the hot water system | – |
| $C_e$ | Present discounted value of the total electricity bill from grid | $ |
| $Cost_{pv}$ | Investment cost of a single PV module | $/kW |
| $Cost_{bat\_energy}$ | Investment cost of a unit battery capacity | $/kWh |
| $Cost_{bat\_power}$ | Investment cost of the power components of the battery system | $/kW |
| $d$ | Eaves stick out | m |
| $d_f$ | Discount factor | – |
| $D_t$ | Time-varying load demand | kW |
| $e_s$ | Saturation vapour pressure | hPa |

(*continued on next page*)

* Corresponding author.
*E-mail address:* benedetto.nastasi@outlook.com (B. Nastasi).

https://doi.org/10.1016/j.energy.2025.136869
Received 25 March 2025; Received in revised form 15 May 2025; Accepted 26 May 2025
Available online 2 June 2025

(*continued*)

| | | |
|---|---|---|
| $EF$ | Emission factor of grid | kg $CO_2$/kWh |
| $F$ | Sky diffusion factor | – |
| $G$ | Total solar radiation | W/m$^2$ |
| $G_0$ | Reference solar radiation | W/m$^2$ |
| $h_r$ | Ross coefficient | – |
| $H$ | Specific humidity | g/kg |
| $H^*$ | Counterfactual of specific humidity | g/kg |
| $I$ | Initial investment cost | $ |
| $I_b$ | Direct beam radiation | W/m$^2$ |
| $I_d$ | Diffuse radiation from the sky | W/m$^2$ |
| $I_r$ | Reflected radiation from the ground | W/m$^2$ |
| $L_s$ | System loss of PV system | % |
| $n$ | Air change rate | h$^{-1}$ |
| $\eta^{PV}$ | Efficiency of the PV panels | – |
| $\eta_{ch}$ | Charging efficiency | – |
| $\eta_{dis}$ | Discharging efficiency | – |
| $P^{PV}$ | Power generation of solar PV panels | kW |
| $P_a$ | Atmospheric pressure | hPa |
| $P_b$ | Perimeter of the horizontal projection of the building | m |
| $P_{grid}$ | Purchase of power from the grid | kW |
| $P_{bat}^{in}$ | Charging operation of the battery | kW |
| $P_{bat}^{out}$ | Discharging operation of the battery | kW |
| $q_m$ | Number of people using hot water | – |
| $q_r$ | Hot water usage quote | L/person |
| $Q_d$ | Daily heat demand for DHW | kWh |
| $S$ | Global horizontal radiation | W/m$^2$ |
| $S^*$ | Counterfactuals of global horizontal radiation | W/m$^2$ |
| $t_l$ | Hot water design temperature | °C |
| $t_r$ | Cold water temperature | °C |
| $T$ | Ambient temperature | °C |
| $T^*$ | Counterfactuals of ambient temperature | °C |
| $U$ | Thermal transmittance of building envelope | W/(m$^2$K) |
| $V$ | Net volume of conditioned area | m$^3$ |
| $W$ | Wind speed | m/s |
| $W^*$ | Counterfactuals of wind speed | m/s |
| $W_c$ | Cell temperature | °C |
| $W_0$ | Reference temperature | °C |
| $x_{bat\_energy}$ | Size of the battery capacity | kWh |
| $x_{bat\_power}$ | Power rating of battery | kW |
| $x_{pv}$ | Number of PV panel | – |
| $\alpha$ | Perturbation factor | % |
| $\delta_d$ | Daily perturbations | % |
| $\delta_h$ | Hourly perturbations | % |
| $\sigma$ | Smoothing factor | – |
| $\Delta T$ | Temperature difference between the heat sink and the heat source | °C |
| $\rho$ | Reflectance of the ground | – |
| $\rho_h$ | Heat transfer coefficient | – |
| $\rho_r$ | Hot water density | kg/L |
| $\theta$ | Angle of incidence | ° |
| $\theta_r$ | Roof pitch angle | ° |
| $\theta_z$ | Zenith angles of the sun | ° |
| $\beta$ | Surface tilt angle | ° |
| $\phi_s$ | Solar azimuth angle | ° |
| $\phi_p$ | Surface azimuth angle | ° |
| $\omega_0$ | The fraction of the battery energy capacity for initial SOC | – |
| **Abbreviations** | **Description** | **Unit** |
| **ASHP** | Air source heat pump | – |
| **BAIT** | Building adjusted internal temperature | °C |
| **CDD** | Cooling degree days | – |
| **COP** | Coefficient of performance for heat pump | – |
| **DHI** | Direct horizontal irradiance | W/m$^2$ |
| **DNI** | Direct normal irradiance | W/m$^2$ |
| **GHI** | Global horizontal irradiance | W/m$^2$ |
| **HDD** | Heating degree days | – |
| **LCOE** | Levelized cost of energy | $ |
| **MILP** | Mixed integer linear programming | – |
| **NPC** | Net present value of cost | $ |
| **RH** | Relative humidity | % |
| **SOC** | Stage of charge | – |
| **WSHP** | Water source heat pump | – |

## 1. Introduction

Renewable energy applications are expanding worldwide as the global focus on energy sustainability and climate change mitigation grows [1]. The building sector has become a key area of focus for renewable energy applications in recent years, as it accounts for nearly 40 % of global energy demand and greenhouse gas emissions [2]. In China, the clean energy transition in rural areas is facing unprecedented opportunities and challenges in order to achieve the “double carbon” goal and promote the rural revitalisation strategy [3]. In 2022, the total energy consumption of China’s rural residential buildings amounted to 250 million tonnes of coal equivalent, accounting for 4.6 % of the nation’s total energy consumption [4]. This energy usage gave rise to 480 million tonnes of $CO_2$ emissions, representing 4.5 % of China’s energy-related carbon emissions [4]. Space heating constituted the principal energy demand, comprising approximately 44 % of primary energy consumption in rural areas [5], and was the primary source of associated carbon emissions [6]. Therefore, the electrification of heating and the adoption of clean energy sources in rural residential buildings play a critical role in supporting the transition of China’s rural energy structure. Photovoltaic (PV) has become an important option for distributed energy systems [7] among many renewable energy technologies due to its flexibility in installation, long service life, and ease of maintenance, as well as the falling prices in recent years [8]. Nonetheless, due to China’s vast territory and various climates, there are considerable disparities in solar energy resource availability, rural buildings’ characteristics, and prosumer energy consumption behaviours [9], which makes the identification of optimal configurations of PV systems challenging and intrinsically connected to the aspects previously mentioned.

A series of academic studies have been conducted in recent years on the optimal configuration of distributed PV systems [10–12], focused on economics, environmental benefits, and management strategies. Multi-objective optimisation techniques [13,14], including genetic algorithms [15], particle swarm optimisation [16], and mixed-integer linear programming (MILP) [17], have been widely applied to address trade-offs among economic, environmental, and technical goals in PV system design, i.e. adopting a techno-economic modelling perspective [18], informed by the use of multiple performance indicators. At the same time, accurate multi-scale load forecasting is crucial to optimising the results, as buildings’ characteristics, meteorological conditions, spatially heterogeneous prosumer behaviour can significantly affect the overall energy consumption of a building [19].

Despite significant progress in research on the optimal configuration of distributed PV systems, several notable shortcomings remain in practical applications. First, existing studies often treat supply-side and demand-side modelling separately or rely on proprietary software [20], lacking open-source software and models that enable a simple, transparent and understandable design optimisation process [21], incorporating both the PV supply side and detailed building energy demand simulations. Second, although archetype-based approaches [22] can be employed and validated tools such as ‘Demand Ninja’ [23] are available for hourly heating and cooling demand modelling at a global scale, their application to the analysis of electrification of end-uses (e.g., heating, cooling, and domestic hot water demands) in rural grid-connected buildings in multiple climatic regions in China requires further development and customisation of models. Furthermore, most current literature focuses on rural buildings in a single climate zone, resulting in an incomplete understanding of multi-scale decarbonisation trade-offs (from national carbon emission reduction targets to household-level self-consumption strategies). Finally, while multi-objective mixed-integer linear programming (MILP) methods are well established in microgrid design, their systematic application to balance economic, carbon emission, and reliability objectives across multiple scales remains underexplored.

This paper aims to create a clear and simple but flexible and scalable

modelling framework to address these shortcomings, and the innovative contributions are as follows:

1. Integration of dynamic demand and supply modelling

A modular framework combining a scalable regression-based demand estimation approach with energy supply optimisation, considering multiple services (electricity, heating, cooling, and domestic hot water), and enabling transparent PV system design across China's multi-climate zones.

2. Climate-aware prosumer behaviour modelling

A unified demand modelling method that links archetype-based building modelling to climate-dependent energy-use patterns, enhancing rural load profiling under diverse climatic conditions.

3. Multi-objective optimisation framework to support decarbonisation strategy testing

A multi-objective MILP framework to systematically identify the optimal configuration of PV panels, storage solutions, and operational strategies across multiple climate zones.

The article is structured as follows: Section 2 reviews the literature to identify knowledge gaps. Section 3 details the principles and methodology of the energy demand and energy supply modelling and the multi-objective MILP formulation. Section 4 presents the assumptions for the case study areas selected. Section 5 discusses the results obtained by applying the modelling framework to multiple Chinese climate zones, focusing on economic, environmental and reliability trade-offs. Finally, Section 6 summarises the main findings of this paper, discusses limitations, and suggests future research directions.

## 2. Literature review

This section discusses the basic principles required to achieve the modelling of the energy systems of this study, in relation to topics that have emerged in the recent scientific literature. Section 2.1 discusses the various modelling approaches used to simulate building energy demand, from white-box, physics-based methods to black-box, data-driven techniques, and highlights the importance of establishing modular and interpretable frameworks that clearly present the underlying assumptions, parameter sources, and calibration processes. Section 2.2 describes the methods involved in the design and optimisation of rural renewable energy systems, with a special focus on MILP technique of multi-objective optimisation, and the challenges emerging from existing literature. Finally, Section 2.3 summarises the main difficulties identified in the literature, particularly the gaps in creating systematic, understandable and interpretable modelling workflows.

### 2.1. *Interpretability and understandability in energy modelling*

In the existing literature, there has been a growing interest in both academia and industry regarding the interpretability and understandability of energy system modelling [24]. Building energy models can typically be subdivided into three categories: "white box" (based on detailed physics), "grey box" (hybrid physical-statistical), and "black box" (based on statistical or machine learning) [25], which involve machine learning (ML) applications [26] and physics-informed ML contexts [27]. Accurately estimating hourly or sub-hourly heating/cooling loads in buildings typically requires white-box models based on detailed physical principles [28], demanding significant modelling effort and time. In contrast, black-box techniques [29] like, for example, Deep Neural Networks (DNNs) [30], Gaussian Processes [31], and others, offer in many cases high predictive performance (depending however on the availability and quality of training data), but their underlying mechanisms may not be as easily understood by humans as other modelling techniques [32], i.e., they are not interpretable [24]. Enhancing the interpretability and simplicity of black-box models is an open challenge, with recently developed techniques like Kolmogorov-Arnold Networks (KAN) [33] demonstrating promising outcomes.

Among simpler and interpretable techniques, regression-based methods have gained prominence due to their approximate physics-aligned interpretability [34]. A notable implementation is the open-source tool 'Demand Ninja' [23], which employs piecewise linear regression models to simulate global hourly heating/cooling load profiles, and which has been widely used for measurement and verification (M&V) [35]. The tool employs the MERRA2 [36] climate database as input. Additionally, it is feasible to employ it for the simulation of future climate conditions by utilizing morphed meteorological data files [37], which are in accordance with climate change scenarios [38], because of the availability of open-source code. However, the tool does not provide a detailed explanation of the calculation of the model coefficients, which are reliant upon the physical characteristics and operation of the building [39], despite its interpretability.

Nonetheless, other authors have described the process in the context of detailed model calibration [40] and the estimation of energy balance components [41]. Simultaneously, the multipliers of the heating and cooling profiles are based on the average typical conditions for heating and cooling. For this reason, they may require modification to be calibrated to a specific case [42]. In general, using a piecewise linear energy signature-based approach [43] for the development of surrogate building models (when measured data are available) establishes a reliable connection between energy performance and local climate [39], which has been tested at scale by utilities [44,45] also with the inclusion of smart thermostats data [46] and prototype/reference buildings [47]. This approach can be employed in both the design and operation phases [48], and it is advisable that comfort settings [49] be considered, as they also affect base (balance-point) temperatures [50] and their variability [51], in addition to the other physical characteristics.

The accessibility of open-source models and open data information is particularly relevant to accelerate research and development, but the availability of code and data without a systematic explanation or validation process for model assumptions and internal mechanisms (logical blocks, type of simulation or optimisation algorithms, etc.) [21] may limit their effectiveness.

In fact, energy system models are not pure simulations or optimisations [52] of the objective world but contain a considerable number of value judgements and scenarios, and if this 'core' is not presented in a coherent and transparent way, it is difficult to be truly understood and trusted by external users (including policy makers or other research teams) [53]. Therefore, based on open-source models and data, it is also necessary to focus on the design goals of model understandability and interpretability, such as strengthening the description of the constituent modules, clarifying the sources of parameters and assumptions, and providing physical assumptions and algorithmic components that can be verified or calibrated. Only in this way can the researcher accurately grasp the modelling ideas and constraints when taking over or extending the model and ensure that the tool itself is 'highly interpretable' [54] and 'highly understandable' [21].

As specified before, the 'Demand Ninja' model input parameters can be specified based on physical assumptions, which can be derived by showing the correspondence between the building energy balance components [40,55]. In this way regression-based approaches [56] be effectively used for model calibration at the state-of-the-art [42]. This ability to potentially create calibrated models makes them suitable for Digital Twins (DT) [57], acting as a digital replica of the energy behaviour of the building (and physical assets more in general) in real-time with bi-directional flow of information, suitable for multiple potential applications in the built environment [58].

This is particularly important in relation to heating electrification by

means of air source heat pumps [59] due to the Coefficient of Performance (COP) variability [60] and its temperature dependence, which is essential both for analytical performance calculations [61] and dynamic control [62]. More generally, time and temperature dependence are essential for modelling energy demand at the building and community scale [63]; recent modelling advancements leveraging interpretable data-driven methods, derived from the Time of Week and Temperature (TOWT) algorithm [64], are founded on principles akin [65] to those employed in the previously mentioned 'Demand Ninja' tool [23], incorporating solar radiation [66] as an input and auto-regressive components [67].

### 2.2. Rural renewable energy system design and optimisation

Fundamentally, the optimisation of rural renewable energy systems is conceptually similar to the cost optimisation analysis in the building sector, where the latter is used to determine energy performance levels yielding minimal global costs over projected economic lifecycles [68], while the former takes into account the support of the rural energy transition and the climate treaty by extending the scope of optimisation to incorporate the environmental dimension, forming the dual objective of finding the global cost and $CO_2$ emissions over the lifecycle of the equipment. In addition, at the level of multi-objective optimisation [69,70], there is a lack of large-scale, systematic application of tools represented by the MILP methodology for multi-objective trade-offs (e.g., balancing economics, carbon emissions and reliability of energy supply) in rural distributed energy systems [14] considering different buildings and climate characteristics. On one hand, MILP can better model the complex decision-making strategies for prosumers, enhancing reliability and reducing overall costs [71]; on the other hand, MILP also facilitates extending algorithmic solution methods to collaborative optimisation at the level of user communities or building clusters [72], by means of scalable multi-commodity network [73,74] formulations or graph-based approaches [75,76]. Graph/network based models depict energy systems as interconnected nodes (e.g., buildings, energy conversion technologies) and edges (e.g., energy flows), rendering complex systems visually and conceptually intuitive [77], while also facilitating computational efficiency in analysis through appropriate formulations and simplifications [78,79].

Indeed, to truly implement this capability, it is necessary to strike a balance between model size and algorithm performance, and to incorporate more comprehensive design and discussion regarding both solution efficiency and interpretability. Existing research has demonstrated that by improving model structure and employing automated scale control, the intractability issues caused by excessively large model sizes can be alleviated to a certain extent [80]. However, without an understandable design of the model itself and a clear articulation of core parameters and constraints, it will still be impossible to effectively convey the decision-making logic to end users (including local governments, farmers, and enterprises).

Furthermore, it is important to continue with development and promotion of open-source tools in rural planning field [20]. However, discussions on rural building electrification remain fragmented. This is particularly evident in climatically diverse regions such as China, where significant uncertainty persists regarding the demand changes induced by complete electrification of rural buildings. Much of the existing literature concentrates on single regions or specific climate types, rather than offering a broader comparative perspective. For example, Li et al. [81] has only designed a hybrid renewable energy system for an electrification project in a rural village in western China, Ji et al. [82] has only designed a stand-alone microgrid and conducted sensitivity analyses for a rural village in Xinjiang, China, and Yuan et al. [83] has only designed a PV microgrid for a rural village in Guangdong, China. And most studies, including these literatures [84–86], do not provide an understandable and interpretable workflow, their demand is usually obtained based on estimates, and supply is usually modelled using commercial software. As a result, there is a limited understanding of the variations in load characteristics, building insulation levels, and new energy system configurations across different climatic zones. Without comprehensively accounting for these regional differences and incorporating them into scenario simulations and model calibrations, the conclusions drawn are often confined to particular geographical or climatic contexts, thereby hindering their generalisability to the entire country or a broader spectrum of rural areas.

Finally, one of the major added values of adopting PV systems in rural areas is their high efficiency in energy saving [87] and emissions reduction [88]. However, if only the operation phase is considered, obvious limitations arise, and a more comprehensive set of indicators may be considered [89]. When focusing primarily on energy and carbon emission, it is necessary to account for the embodied energy and embodied carbon emissions generated throughout the entire life cycle (including manufacturing, transportation, installation, and decommissioning), otherwise the overall environmental performance of the system might be overestimated [90]. Moreover, the values of embodied energy and embodied carbon are not fixed but dynamically evolve with the characteristics of the reference energy system and the various stages of the production process, with figures that can differ significantly among countries or regions [91], especially in China, where these quantities may be much higher than in other countries [90]. Therefore, incorporating embodied energy and embodied carbon is essential to provide a more scientifically reliable basis for configuration decisions during system optimisation.

### 2.3. Summary of literature review

In summary, the challenges of understandability [21] and interpretability [92] must be particularly emphasised in both energy system modelling [53] and building energy modelling [93]. Current literature predominantly focuses on demand-side load profiling or supply-side energy modelling but overlooks systematic investigation into methodological workflows incorporating interpretability and explainability features. Understandability can be enhanced through establishing modular logic blocks that enable systematic examination of individual model components, while interpretability involves selecting physics-informed data-driven techniques whose decision pathways remain human-comprehensible. Section 3 details the modelling process to achieve these goals. By building on logical blocks and adopting an open modelling approach with physical assumptions that can be easily calibrated on measured data, we aim to develop a versatile modelling framework. This framework aims to inform, by means of multi-objective optimisation, decarbonisation strategies involving the use of heat pumps (heating electrification), grid-connected PV and prosumer-based grid management practices. It is not just a traditional energy model, but a model that can be calibrated at different scales, ideally constituting the foundation of further applications based on the 'digital twin' [57,94] concept that overcomes the limitations of conventional simulation-based engineering methods (such as those used in building performance simulation) by demonstrating how processes operate in real-time (or near real-time) and realistic conditions.

## 3. Methodology

This study establishes a multi-purpose modelling framework based on modular logic blocks designed to minimise system complexity across diverse applications ranging from prosumer behaviour to investment optimisation. The methodology used in this study is divided into 3 main parts: energy demand modelling, energy supply modelling, and multi-objective optimisation, which will be explained in detail in Sections 3.1, 3.2, and 3.3 below, respectively. The overall workflow is shown in Fig. 1, linking the energy modelling process to open-source data and code according to the proposal in Ref. [53], allowing the user to evaluate the interpretation of the results.

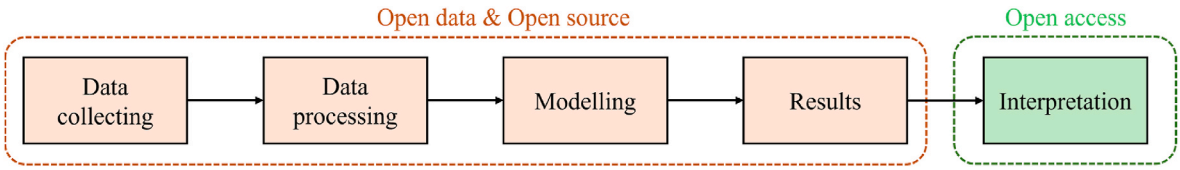



**Fig. 1.** The diagram of the modelling overall workflow.

### 3.1. Energy demand modelling

High-resolution (hourly) load inputs are required before PV system sizing and grid-connected scheduling optimisation; for a typical rural household, the loads to be modelled include base loads (lighting and appliances), space heating and cooling loads, and domestic hot water loads. Hereafter, we present a scalable approach to modelling heating, cooling, and base load profiles, as well as aggregating them for buildings. It stems from improvements to the 'Demand Ninja' model [23] and it includes additional details related to COP calculations for heat pumps, base load variability, etc. The main improvements in our method focus on explaining how to calculate the coefficients based on the buildings' physical features (like slopes and change points in the underlying regression model) and enhancing the statistical parameters (used in time-series model computation) to match the specific characteristics of the selected locations. Another improvement is related to a better definition of the diurnal profiles, making the time series more realistic and avoiding 'jumps' from one day to another. Finally, we verified the scalability of the model by simulating the total electricity demand of rural buildings in 4 climate zones in China.

The workflow of the energy demand model is shown in Fig. 2 and will be described in detail hereafter. Fig. 2a shows the comparison of the BAIT calculated using the 2009–2023 EPW weather file for Xiamen as parameters input with the original ambient temperatures in the file. Fig. 2b shows the relationship between HDDs (Heating Degree Days) and CDDs (Cooling Degree Days) (calculated from heating and cooling thresholds and BAIT), and heating/cooling energy demand (converted through HDDs and CDDs, with calculated heat transfer coefficients as slope), for a typical rural household in Xiamen as an example. Fig. 2c shows an example of heating/cooling electricity demand converted from energy demand by temperature-dependent COP modelling of heat pumps. Fig. 2d presents the domestic hot water electricity demand for a typical rural household, derived by coefficient calculations of the domestic hot water energy demand, with the addition of stochastic perturbations as well as modification with diurnal profiles, converted by a COP model for heat pumps. Fig. 2e shows the baseline electricity demand for a typical rural household, based on appliance power versus time of use, and stochastic perturbations and daily profile modelling. Fig. 2f gives the final electricity demand at the single building level for a typical rural building in Xiamen as an example, consists of baseline electricity demand, heating/cooling electricity demand, and domestic hot water electricity demand.

#### 3.1.1. Heating and cooling demand

The first step is to calculate the Building Adjusted Internal Temperature (BAIT) [23];BAIT is defined as the temperature a person would experience inside an unheated and uncooled building and takes into account the influence of outside air temperature, solar radiation, wind speed and humidity [23]. It is therefore a way to take these environmental parameters into account when calculating the heating and cooling demand of buildings. The coefficients used to calculate BAIT based on these parameters are derived from model fitting, as provided by the authors who tested them across a diverse array of buildings in various climates. Given the empirical nature of the model, coefficients can be adjusted to reflect the behaviours of buildings with specific characteristics and in specific climates.

BAIT is then subjected to a lag smoothing treatment to account for the thermal inertia of the insulated building and a blending treatment to account for occupant behaviour during hot weather for temporary ventilation (e.g. window opening). The above treatments are shown by Eqs. (1)–(6) (from Ref. [23]). It is worth noting that unlike the 'Demand Ninja' model, here the hourly BAIT is calculated directly from hourly data inputs to better capture temperature variations, rather than calculating the daily BAIT and then interpolating it to the hourly BAIT (this is because we have tested that in the Chinese region, the original model would cause the calculated BAIT to be significantly lower than the ambient temperature in summer, whereas a direct calculation of the BAIT at the hourly level is more in line with experience). In addition, the counterfactual values of the meteorological parameters in Eq. (2) were reformulated for the Chinese region, and the details of the determination are shown in Fig. 1 in Appendix A.

$$\mathrm{BAIT}_h = T + x(S - S^*) - y(W - W^*) + z(H - H^*)(T - T^*) \tag{1}$$

$$S^* = 130 + 3T, W^* = 2.52 - 0.003T, H^* = \mathrm{e}^{1.05+0.063T}, T^* = 16 \tag{2}$$

$$\mathrm{BAIT}_h = \frac{\mathrm{BAIT}_h + \sigma \mathrm{BAIT}_{h-1} + \sigma^2 \mathrm{BAIT}_{h-2}}{1 + \sigma + \sigma^2} \tag{3}$$

$$B = \frac{B_{max}}{1 + \mathrm{e}^{-B'}} \tag{4}$$

$$B' = (T - 0.5(B_\mathrm{U} + B_\mathrm{L}))\frac{10}{B_\mathrm{U} - B_\mathrm{L}} \tag{5}$$

$$\mathrm{BAIT}_h = (\mathrm{BAIT}_h(1 - B_h)) + (T_h B_h) \tag{6}$$

Where for each hour ($h$) BAIT is in °C, $S$ is global horizontal radiation (W/m$^2$) and $S^*$ is its counterfactual, $W$ is wind speed (m/s) and $W^*$ is its counterfactual, $H$ is specific humidity (g water per kg air) and $H^*$ is its counterfactual, $T$ is ambient temperature (°C) and $T^*$ is its counterfactual (The counterfactual temperature is defined as the temperature at which humidity exerts no influence on energy consumption, representing the hypothetical point at which neither heating nor cooling is required.), for each $h$, $\sigma$ is the smoothing factor range of [0–1], $B$ is the blending factor, $B_{max}$ is the maximum blending factor with default value of 0.5, $B_\mathrm{U}$ and $B_\mathrm{L}$ are the upper and lower blending limits, typical values are 23 °C and 15 °C, meaning that people start to open their windows in this temperature range, while above 23 °C the impact from the meteorological and smoothing parameters would be reduced by half.

The required meteorological data inputs can be provided by EPW files [95], which are openly available from EnergyPlus [96] or Climate.OneBuilding.Org website [97]. The Code for Thermal Design of Civil Buildings [98] divides China into 4 climatic zones, and 4 different locations have been selected accordingly: severe cold (Harbin), cold (Beijing), hot summer/cold winter (Chengdu), and hot summer/warm winter (Xiamen). This study addresses rural buildings using the selected locations because they can be considered representative of each one of the climatic zones indicate in the thermal design code. Initially, 70 weather files were collected with data for the last 15 years (2009–2023) in TMYx format, to account for potential changes in climate measured in recent years. It should be noted that the humidity data is only available as relative humidity in the raw EPW file, and it is necessary to first calculate the saturated vapour pressure based on Bolton's formula [99] (in Eq. (7)), and then calculate the specific humidity based on Eq. (8) (in

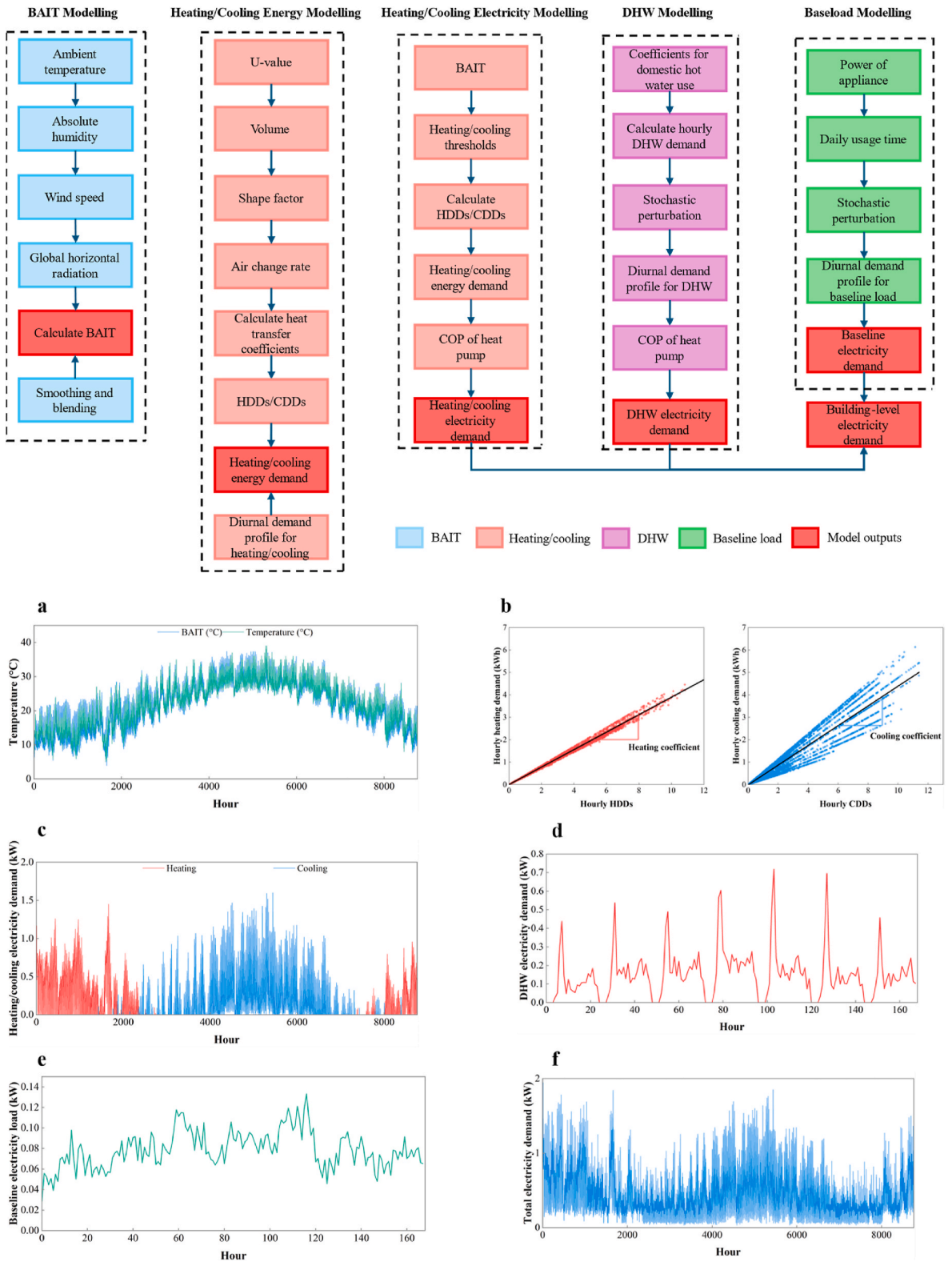



**Fig. 2.** The workflow of the demand-side model. The light blue part of the diagram above represents the BAIT modelling, pink is the heating/cooling demand modelling, purple is the domestic hot water demand modelling, green is the base load modelling, and red represents the output of the model results. The workflow steps are depicted graphically in picture a-f to ease interpretation.

Ref. [100]). The calculation can be carried out by calling the open-source Python library MetPy [101].

$$e_s = 6.112 \times \mathrm{e}^{\left(\frac{17.67 \cdot T}{T+243.5}\right)} \tag{7}$$

$$H = \frac{0.622 e_s \frac{RH}{100}}{P_a - 0.378 e_s \frac{RH}{100}} \times 1000 \tag{8}$$

Where, $H$ is specific humidity in g water per kg air, $RH$ is relative humidity (%), $e_s$ is saturation vapour pressure (hPa), $T$ is ambient temperature (°C), $P_a$ is atmospheric pressure (hPa).

HDDs and CDDs were then calculated using BAIT and a threshold based on [102–104] using 18 °C in winter and 26 °C in summer as the heating/cooling thresholds for severe cold, cold, and hot summer/cold winter regions, and 16 °C in winter and 26 °C in summer as the heating/cooling thresholds for hot summer/warm winter regions.

The physical parameterd representing heat transfer coefficients for the building are used next to convert HDDs and CDDs into daily and then hourly energy demand (i.e., replacing the empirical parameters in the original model formulation) by using representative heating/cooling daily profiles derived from Ref. [23]. The daily energy consumption is computed with a change-point (CP) model [105], presented in Fig. 3, which can be interpreted in this way: the weather-dependent component of energy consumption is proportional to HDDs and CDDs, respectively for heating and cooling, while the weather-independent component is assumed to be constant. The slopes of the regression model for heating and cooling are computed according to the physical parameters of the building, and this ensures physical interpretability when converting HDD and CDD into energy demand.

The calculation of the slopes in the change-point regression model assumes that they are approximately similar to the heat transfer coefficients for the building, as explained by Allinson and Lomas [107] and by Rasmussen et al. [108], and was carried out using the formulas in CIBSE Guide A, as shown in Eq. (9). The shape coefficient of the building is used to determine the surface area and volume of envelopes (Eqs. (10) and (11)).

$$\rho_h = \frac{\sum UA + \frac{1}{3} nV}{1000} \tag{9}$$

$$B_s = \frac{\left(A_r + \sum_{i=1}^{n} A_{w_i} + \sum_{i=1}^{m} A_{g_i}\right)}{V} \tag{10}$$

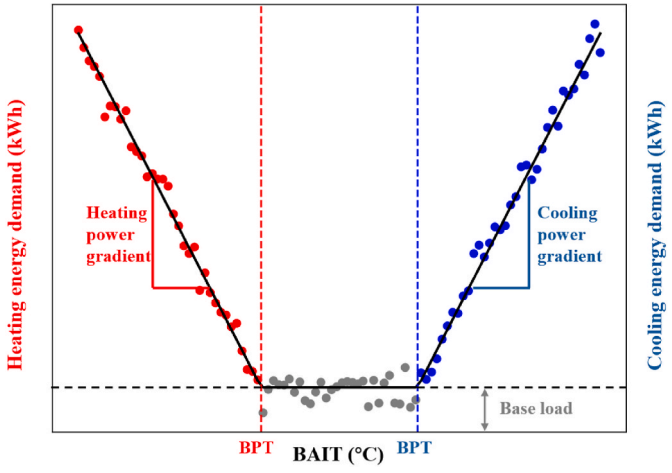



**Fig. 3.** Diagram of the principle of the change-point regression model. Source from Ref. [106].

$$A_r = \frac{A_{\text{base}} + P_b \cdot d + 4 \cdot d^2}{\cos(\theta_r)} \tag{11}$$

Where, $\rho_h$ is heat transfer coefficient, $U$ is the thermal transmittance of building envelope (W/(m$^2$K)), $A$ is envelope surface area (m$^2$), $n$ is the air change rate (h$^{-1}$), based on [102–104] taken as 0.5/h in severe cold regions and cold regions, and 1.0/h in the remaining regions, $V$ is the net volume of conditioned area (m$^3$). $B_s$ is the shape coefficient of building, based on [102–104] taken as 0.5 in severe cold, cold, and hot summer/cold winter regions, 0.35 in hot summer/warm winter regions. $A_r$ is the roof surface area (m$^2$), $A_{wi}$ is surface area of each external wall (m$^2$), $A_{gi}$ is surface area of each external window (m$^2$), $A_{\text{base}}$ is area of building base (m$^2$), $P_b$ is perimeter of the horizontal projection of the building (m), $d$ is eaves stick out (m) taken as typical value 0 m, $\theta_r$ is roof pitch angle (°) taken as typical value 30°.

Finally, heat pumps are used to meet the heating/cooling energy demand of buildings. Considering domestic hot water supply, air-source heat pump hot water supply system is not suitable for severe cold and cold regions [109], so water-source heat pump (WSHP) is used, and air-source heat pump (ASHP) is used in other regions. The COP models for WSHP and ASHP used in this study were summarised by Ruhnau et al. [110] based on the use of quadratic regression proposed by Fischer et al. [111] in conjunction with the fitting of data from suppliers, as shown in Eq. (12).

$$COP = \begin{cases} 6.08 - 0.09 \cdot \Delta T + 0.0005 \cdot \Delta T^2, ASHP \\ 9.97 - 0.20 \cdot \Delta T + 0.0012 \cdot \Delta T^2, WSHP \end{cases} \tag{12}$$

Where, $\Delta T$ is temperature difference between the heat sink and the heat source (°C), based on [109], the outlet temperature is set to 45 °C for heating and 7 °C for cooling with ASHP, and the source temperature is the ambient temperature, for WSHP the outlet temperature is set to 45 °C for heating and 10 °C for the heat source, and the outlet temperature is set to 15 °C for cooling and 7 °C for the cold source.

### 3.1.2. Domestic hot water demand

The daily heat demand for domestic hot water (DHW) was modelled according to Eq. (13), then a random perturbation (Eq. (14)) was added after uniformly distributing it as an hourly heat demand, followed by shape modification using diurnal profiles, which were developed based on the average daily DHW consumption profiles for dwellings in different countries reviewed in Ref. [112]. Finally, converting the heat demand of DHW to electricity demand using Eq. (12).

$$Q_d = \frac{q_m \cdot q_r \cdot C(t_r - t_l)\rho_r \cdot C_r}{3600} \tag{13}$$

Where, $Q_d$ is daily heat demand for DHW (kWh), $q_m$ is number of people using hot water, $q_r$ is hot water usage quota (L/person) taken as 50L/person per day based on [113,114], $C$ is specific heat capacity of water (kJ/kg°C), $t_r$ and $t_l$ are hot water design temperature and cold water temperature (°C), which are set to 60 °C and 10 °C respectively according to Ref. [109], $\rho_r$ is hot water density in kg/L, $C_r$ is heat loss coefficient of the hot water system.

### 3.1.3. Baseline electricity demand

Baseline electricity demand refers to the power drawn by appliances other than those used for heating, cooling and DHW. The total electricity demand is calculated from each appliance's power rating and usage duration, then distributed evenly across all hours of the day. Stochastic perturbations (Eq. (14)) are added, and the resulting profile is adjusted using province-specific diurnal curves derived from the typical power curves published for each city.

$$\alpha = 1 + \delta_d + \delta_h \tag{14}$$

Where, $\alpha$ is perturbation factor and at each time step, the value of that

time step is multiplied by the perturbation factor to add a stochastic perturbation for that time series, $\delta_d$ is daily perturbations that are generated by randomly sampling one day's perturbation values from a normal distribution with mean zero and standard deviation of 'input values', $\delta_h$ is hourly perturbations that are generated by randomly sampling each hour's perturbation values from a normal distribution with mean zero and standard deviation of 'input values', The standard deviation values input here are both 20 %.

### 3.2. Energy supply modelling

The energy supply in this study is mainly from PV power generation, with the grid and batteries as back-up energy sources. In general, the power generation of solar PV panels ($P^{PV}$) can be calculated based on the total solar radiation ($G$) on the tilted PV module, the efficiency of the PV panels ($\eta^{PV}$), the area ($A^{PV}$) of the installed PV panel array [115], and the system loss ($L_s$), i.e., the following Eq. (15).

$$P^{PV} = G\eta^{PV}A^{PV}(1 - L_s) \tag{15}$$

The total radiation on the inclined plane consists of three components: direct beam radiation ($I_b$), diffuse radiation from the sky ($I_d$) and reflected radiation from the ground ($I_r$), as shown in Fig. 4, with the following mathematical formulate Eq. (16).

The sky diffuse radiation on the inclined surface is calculated based on the Klucher model [116] (Eq. (17)–(19)), which assumes that the sky diffuse radiation is not perfectly homogeneous but is related to the surface tilt ($\beta$), angle of incidence ($\theta$) and solar zenith angle ($\theta_z$). Beam radiation is calculated from the angular relationship with the tilted plane, and reflected solar radiation is determined by the reflectance of the ground ($\rho$) and the geometrical properties of the tilted plane, based on Eq. (20) and Eq. (21), respectively.

$$G = I_b + I_d + I_r \tag{16}$$

$$I_d = DHI \cdot \frac{1+\cos\beta}{2} \cdot \left(1 + F \cdot \sin^3\left(\frac{\beta}{2}\right)\right) \cdot \left(1 + F \cdot \cos^2\theta \cdot \sin^3\theta_z\right) \tag{17}$$

$$\cos\theta = \cos\theta_z \cos\beta + \sin\theta_z \sin\beta \cos\left(\phi_s - \phi_p\right) \tag{18}$$

$$F = 1 - \left(\frac{DHI}{GHI}\right)^2 \tag{19}$$

$$I_b = DNI \cdot \cos(\theta) \tag{20}$$

$$I_r = GHI \cdot \rho \cdot \frac{1-\cos(\beta)}{2} \tag{21}$$

Based on the total radiation in the tilted plane, the output power of the PV module is calculated by means of the semi-empirical module efficiency model of Durisch et al. [117] (Eq. (22)), which is suitable to represent most module types.

$$\eta^{PV} = p \cdot \left(q \cdot \frac{G}{G_0} + \left(\frac{G}{G_0}\right)^m\right) \cdot \left(1 + r \cdot \frac{W_c}{W_0} + s \cdot \frac{AM}{AM_0} + \left(\frac{AM}{AM_0}\right)^u\right) \tag{22}$$

Where, $G_0$, $W_0$, $AM_0$: reference radiation, reference temperature and reference air mass are taken as 1000 W/m$^2$, 25 °C and 1.5 respectively. $p$, $q$, $m$, $r$, $s$, $u$: module-specific parameters, the parameters for the different technology modules can be referred to the values in Ref. [117]. The cell temperature ($W_c$) can be calculated from the ambient temperature ($T$) and radiation as well as the Ross coefficient ($h_r$) (Eq. (23)). Relative air mass ($AM$) is calculated based on the Kasten and Young model [118] according to the solar zenith angle as in Eq. (24).

$$W_c = T + h_r \cdot G \tag{23}$$

$$AM = \frac{1}{\cos(\theta_z) + 0.50572 \cdot (96.07995 - \theta_z)^{-1.6364}} \tag{24}$$

All the above calculations are managed in the Python environment and, except for Eq. (16) and Eq. (22), the open-source Python library pvlib [119] can be called for data acquisition and calculations. Besides that, it is also possible to obtain PV generation data directly from the open-source platform PVGIS [120] based on location. However, it should be noted that since PVGIS uses the European time zone, the data timestamps need to be lagged.

### 3.3. Multi-objective optimisation

The optimisation model presented in this paper is based on mixed-integer linear programming (MILP), aims to determine the optimal design and operation of a grid-connected PV system and provides a refined representation of the battery storage system, distinguishing between the energy capacity and the power rating of the battery. Fig. 5 shows a high-level schematic of the components and logic blocks of the optimisation model.

The net present value of cost (NPC) and annual $CO_2$ emissions were used as the objective function. And for each discrete period ($t$) to determine the charging ($P^{in}_{bat,t}$) and discharging ($P^{out}_{bat,t}$) operation of the battery as well as the purchase of power from the grid ($P_{grid,t}$), to meet the time-varying load demand ($D_t$) whilst minimising the weighted sum of the economic cost and the environmental emissions. To be specific, the decision variables include $x_{pv}$ (installed PV capacity in 1 kW), $x_{bat\_energy}$ (battery energy capacity in kWh), $x_{bat\_power}$ (battery power rating) and operational variables during discrete periods ($t = 0,1,\ldots,T-1$): $P^{in}_{bat,t} \geq 0$, $P^{out}_{bat,t} \geq 0$, $P_{grid,t} \geq 0$, curtailment of PV generation $P_{curtail,t} \geq 0$, and state of charge (SOC) $SOC_t \geq 0$. The PV generating capacity of the system is described by the output of each 1kWp module $PV_t$ (in kW) during the period $t$, while the load demand is $D_t$ (in kW). To ensure the healthy operation of the battery, it is required that its SOC always meets the safety interval, i.e. $0.2x_{bat\_energy} \leq \mathrm{SOC}_t \leq x_{bat\_energy}$ (initial state is set to $SOC_0 = \omega_0 x_{bat\_energy}$ indicate the initial SOC as a fraction of battery energy capacity), and the dynamic update relationship of the battery is shown in Eq. (25).

$$\mathrm{SOC}_t = \mathrm{SOC}_{t-1} + \eta_{ch} P^{in}_{bat,t-1} \cdot \Delta T - \frac{1}{\eta_{dis}} P^{out}_{bat,t-1} \cdot \Delta T, t = 1, \ldots, T-1 \tag{25}$$

Where the charging efficiency $\eta_{ch} = 0.95$ and discharging efficiency $\eta_{dis} = 0.95$. In each period, the following constraints are imposed:

$$P^{in}_{bat,t} \leq x_{bat\_power}, P^{out}_{bat,t} \leq x_{bat\_power}, \forall t. \tag{26}$$

And the conditions for balancing the power supply must be fulfilled (in Eq. (27)).

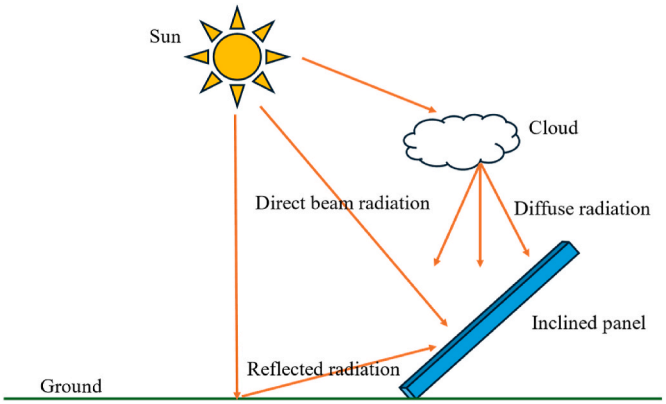



**Fig. 4.** Diagram of solar radiation received on an inclined panel.

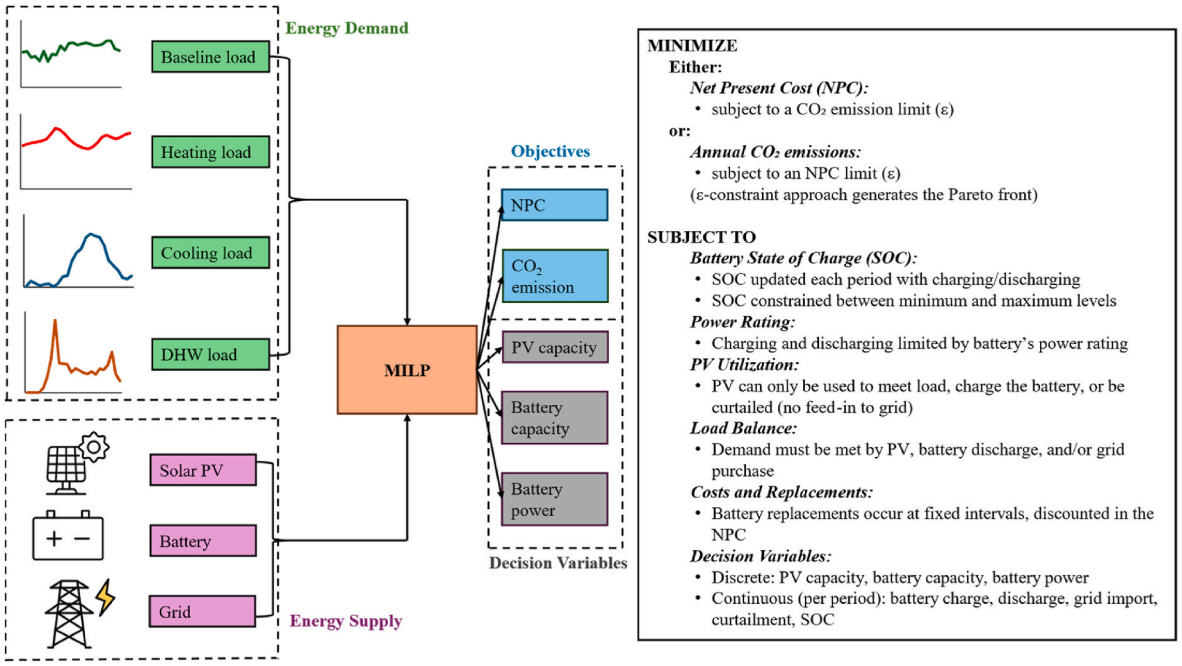



**Fig. 5.** A high-level schematic of the main components of the model and the logic blocks.

$$D_t = x_{pv} PV_t - P_{curtail,t} + P^{out}_{bat,t} + P_{grid,t} - P^{in}_{bat,t}, \forall t. \tag{27}$$

Meanwhile, since PV generation not directly used to meet the load can only be allocated to battery charging or curtailed (given that PV generation is not sold to the grid due to low feed-in tariffs), we require:

$$P^{in}_{bat,t} + P_{curtail,t} \le x_{pv} PV_t, \forall t. \tag{28}$$

In terms of economic costs, the initial investment costs are simply defined as:

$$I = Cos\ t_{pv} x_{pv} + Cos\ t_{bat\_energy} x_{bat\_energy} + Cos\ t_{bat\_power} x_{bat\_power} \tag{29}$$

where $Cost_{pv}$ and $Cost_{bat}$ are the investment costs of a single PV module (taken as \$100/kW) and a unit of battery capacity (taken as \$53/kWh), respectively. The life cycle of the entire system is considered to be 20 years, of which the battery replacement needs to occur in years 8 and 16, with a discounted cost:

$$B_b = \frac{Cos\ t_{bat\_energy} x_{bat\_energy} + Cos\ t_{bat\_power} x_{bat\_power}}{(1+r)^8} + \frac{Cos\ t_{bat\_energy} x_{bat\_energy} + Cos\ t_{bat\_power} x_{bat\_power}}{(1+r)^{16}} \tag{30}$$

Where $r = 0.04$ is the discount rate. In addition, operation and maintenance costs are calculated at 1 % of the initial investment cost and are discounted by a discount factor $d_f$ ($d_f = \sum_{k=1}^{20} \frac{1}{(1+r)^k}$) and recorded as $0.01\ I\alpha$. The cost of purchasing electricity from the grid is then priced at $\pi_e$ (in \$/kWh), and its discounted total cost is $C_e$ ($C_e = \pi_e d_f \sum_{t=0}^{T-1} P_{grid,t}$).

The NPC (in \$) can be expressed as Eq. (31). Assuming that the emissions from electricity purchases per kWh of grid electricity are $EF$ (in kg $CO_2$/kWh), the $CO_2$ emissions are as shown in Eq. (32).

$$\mathrm{NPC} = I + B_b + 0.01 I d_f + C_e \tag{31}$$

$$\mathrm{CO_2} = \mathrm{EF} \sum_{t=0}^{T-1} P_{grid,t} \tag{32}$$

In summary, the multi-objective optimisation model for this grid-connected PV system design problem as follows:

The optimised result is given in the form of a Pareto front [121], representing a set of optimised solutions that do not dominate each other in the search space but outperform other results, which reflects trade-offs between economic costs and environmental emissions. The objective function is optimised and solved using the commercial solver CPLEX. The parameter inputs and flowchart illustrating the creation of the optimisation model are provided in Appendix B.

$$\min_{\substack{x_{pv}, x_{bat\ energy}, x_{bat\ power}, \\ \{P^{in}_{bat,t}, P^{out}_{bat,t}, P_{grid,t}, SOC_t, P_{curtail,t}\}_{t=0}^{T-1}}} \{NPC, CO_2\}$$

$$\begin{aligned}
& SOC_0 = w_0 x_{bat_energy}, \\
& SOC_t = SOC_{t-1} + \eta_{ch} P^{in}_{bat,t-1} \Delta T - \frac{P^{out}_{bat,t-1} \Delta T}{\eta_{dis}}, t = 1, \ldots, T-1, \\
& 0.2 x_{bat_energy} \le SOC_t \le x_{bat_energy}, \forall t, \\
& P^{in}_{bat,t} \le x_{bat_power}, P^{out}_{bat,t} \le x_{bat_power}, \forall t, \\
& P^{in}_{bat,t} + P_{curtail,t} \le x_{pv} PV_t, \forall t, \\
& D_t = x_{pv} PV_t - P_{curtail,t} + P^{out}_{bat,t} + P_{grid,t} - P^{in}_{bat,t}, \forall t, \\
& x_{pv} \in \mathbb{Z}_{\ge 0}, x_{bat_energy} \in \mathbb{Z}_{\ge 0}, x_{bat_power} \in \mathbb{Z}_{\ge 0}, \\
& P^{in}_{bat,t}, P^{out}_{bat,t}, P_{grid,t}, SOC_t, P_{curtail,t} \ge 0, \forall t.
\end{aligned} \tag{33}$$

## 4. Case study description

As described in Section 3.1.1, weather data spanning the past 15 years are available for 70 locations across 4 climatic zones: 22 in the severe cold region, 14 in the cold region, 22 in the hot-summer/cold-winter region, and 12 in the hot-summer/warm-winter region. Initially, the U-value of the envelope necessary for modelling heating and cooling demand is established based on the standard limits delineated for each climatic zone in the Design Standard for Energy Efficiency of Rural Residential Buildings [122]. The building volume range is defined using an average rural household occupancy of 2–4 people per household, a storey height of 3.3m and a per capita floor area of 48.9 $m^2$. Surface areas are then calculated based on building shape coefficients (Eqs. (10) and (11)). The resulting envelope parameters for the 4 climate zones are shown in Table 1.

The heat loss coefficients for the climate zones are then calculated based on Eq. (9), where uncertainty ranges such as U-values, and ranges of volumes are randomly assigned for each zone using the latin hypercube sampling (LHS) [123] to generate samples. In modelling the basic energy needs, Table 2 shows the main appliances and how long they are used each day based on earlier surveys, and LHS was used within these uncertainties to create random samples for different areas.

## 5. Results and discussion

This section presents and discusses the results of optimally modelling grid-connected PV systems for rural buildings in 4 representative climatic zones in China from both a numerical and a visual point of view. Section 5.1 will present the results of the energy demand modelling for rural buildings in these 4 typical representative areas, and the energy supply modelling, involving the optimisation of the grid-connected PV system, will be presented in Section 5.2.

### 5.1. Results of electricity demand modelling for four climate zones

In this section, the simulated electricity demand generated for rural buildings in representative areas of 4 climate zones in China using the above modelling framework is presented. For better presentation, one location in each climate zone is selected as a representative display in the main text, namely, Harbin in the severe cold region, Beijing in the cold region, Chengdu in the hot-summer/cold-winter region, and Xiamen in the hot-summer/warm-winter region. Fig. 6 shows the global horizontal solar radiation and calculated BAIT for these 4 regions. Their BAIT is closely aligned with their climate characteristics, with Harbin having the lowest average and Xiamen the highest. In contrast, global horizontal solar radiation, affected by factors such as latitude, atmospheric conditions, topography, and seasonal changes, displays a pattern significantly different from BAIT, with Beijing averaging the highest and Chengdu the lowest.

Table 3 summarises the comparison between heating and cooling energy demand and electricity demand for rural buildings in these 4 representative locations, scaled by net floor area.

It can be seen that there are significant differences in the energy consumption characteristics of rural buildings in the 4 regions: Harbin has the highest total electricity consumption (52.9 kWh/$m^2$), Xiamen has the lowest (26.1 kWh/$m^2$), and Beijing and Chengdu are in the middle (35.3 and 48.4 kWh/$m^2$, respectively). This disparity is mainly due to the differences in heating and cooling demand caused by the climatic conditions of each region: heating loads are dominant in severe cold regions (e.g., Harbin), resulting in much higher energy consumption per unit area than in regions with hot summers and warm winters (e.g., Xiamen), while Beijing and Chengdu have a relatively balanced heating and cooling demand in both the winter and summer seasons. In addition, it is interesting to note that rural buildings in Chengdu have a higher heating energy demand per unit area than Beijing, but in fact their calculated HDDs are only about half of those in Beijing, so this is caused by the difference in the insulation capacity of the houses; in this case, the heat transfer coefficient is calculated to be 0.36 in Chengdu and 0.25 in Beijing. This result side-by-side reinforces the physical interpretability of the calculations using archetypes combined with piecewise models in this study.

Fig. 7 shows the results of modelling the total electricity demand for four different locations, making it easy to see the big differences in electricity needs for each of the four climate zones in China. The total electricity demand is characterised as a function of BAIT to better observe the pattern of electricity demand changes with temperature, where heating and cooling electricity demand are separately characterised as functions of HDDs and CDDs, respectively, to visualise the sensitivity of these two components of electricity demand to change in degree-days. When calculated using the standard design degree-day thresholds of 26 °C in summer and 18 °C in winter (16 °C in Xiamen), the annual electricity demand for rural buildings in Harbin, Beijing, and Chengdu accounted for most of the heating electricity demand, while in Xiamen the heating electricity demand for rural buildings was almost equal to the cooling electricity demand. The total electricity demand is described as a function of BAIT to clearly show how electricity demand changes with temperature, where heating and cooling electricity demand are separately described as functions of HDDs and CDDs, respectively, to highlight how sensitive these two parts of electricity demand are to changes in degree-days.

When calculated using the standard design degree-day thresholds of 26 °C in summer and 18 °C in winter (16 °C in Xiamen), the annual electricity demand for rural buildings in Harbin, Beijing, and Chengdu accounted for most of the heating electricity demand, while in Xiamen the heating electricity demand for rural buildings was almost equal to the cooling electricity demand. The figure illustrates that one of the challenges of implementing renewable energy electrification projects in rural China is the need to meet high electricity loads during the winter months when PV generation is low; however, this trend may change as a result of climate change in the future, where global warming and economic conditions that support the pursuit of comfort may bring about a drop in cooling temperature thresholds, leading to more electricity demand during hot weather.

**Table 2**
The data range of baseline loads.

| Appliance | Power range (W) | Range of time used (h) |
|---|---|---|
| Lamps | 40–80 | 4–8 |
| Fridge | 40–80 | 24–24 |
| TV | 50–100 | 2–6 |
| Cooker | 400–600 | 0–0.8 |
| Electric kettle | 1000–1200 | 0–0.5 |
| Hair dryer | 300–500 | 0–0.1 |
| Washing machine | 200–300 | 0–0.5 |

**Table 1**
The range of U-values of envelopes in 4 climate zones for rural buildings.

| Climate zone | U-values of external wall (W/($m^2$K)) | U-values of window (W/($m^2$K)) | U-values of roof (W/($m^2$K)) | Volume ($m^3$) | Shape coefficient |
|---|---|---|---|---|---|
| Severe cold | 0.4–0.5 | 2.0–2.2 | 0.3–0.4 | 274–548 | 0.5 |
| Cold | 0.45–0.65 | 1.6–2.8 | 0.42–0.5 | 274–548 | 0.5 |
| Hot-summer/cold-winter | 1.0–1.8 | 2.8–3.2 | 0.5–1.0 | 274–548 | 0.5 |
| Hot-summer/warm-winter | 1.5–2.0 | 4.0–4.0 | 0.35–1.0 | 274–548 | 0.35 |

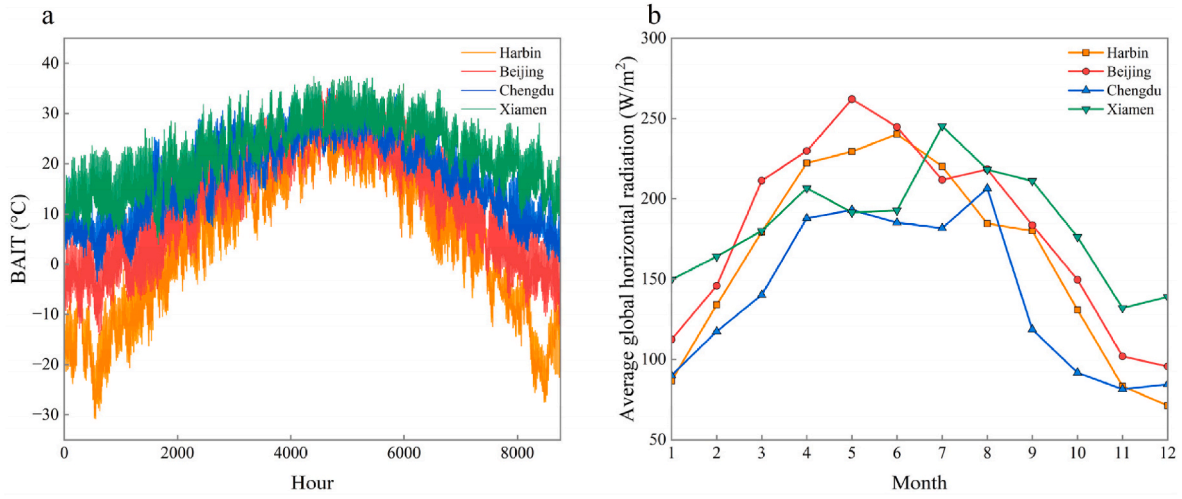



**Fig. 6.** **a** Calculated BAIT for rural buildings in Harbin, Beijing, Chengdu, Xiamen, **b** Monthly mean global horizontal solar radiation for these 4 areas.

**Table 3**
Summary of the results of demand modelling at 4 representative locations.

| Parameters | Location | | | |
|---|---|---|---|---|
| | Harbin | Beijing | Chengdu | Xiamen |
| Net building area ($m^2$) | 165 | 184 | 127 | 148 |
| Heating energy demand (kWh/$m^2$) | 170.7 | 93.1 | 96.0 | 20.9 |
| Cooling energy demand (kWh/$m^2$) | 1.8 | 5.1 | 14.9 | 30.0 |
| Heating electricity demand (kWh/$m^2$) | 38.5 | 21.0 | 27.7 | 5.7 |
| Cooling electricity demand (kWh/$m^2$) | 0.2 | 0.6 | 3.5 | 7.2 |
| Total electricity demand (kWh/$m^2$) | 52.9 | 35.3 | 48.4 | 26.1 |

As forecasted in Ref. [124], with future climate change and economic growth, heating demand in China is expected to drop significantly and cooling demand to rise significantly, and the temperature thresholds for heating and cooling are likely to be close to the 22 °C in summer and 18 °C in winter currently used in economically developed countries.

### 5.2. *Results of multi-objective optimisation of grid-connected PV*

Similarly, this section will show the results of optimising grid-connected PV systems for typical rural buildings in Harbin, Beijing, Chengdu and Xiamen. Since there is no initiative to sell electricity from PV to the grid, the battery is added to the system's dispatch to explore whether it is a better option than just PV and the grid. Both the grid and the batteries act as backups to provide reliability for electricity consumption.

PV power generation in different regions is modelled using the energy supply model introduced in Section 3.2, and Fig. 8 shows the seasonal and daily output models for 1kWp PV power generation in these 4 regions. The seasons are defined in this way: Spring from March to May, Summer from June to August, Autumn from September to November, and Winter from December to February.

Interestingly, the PV power generation curve from the tilted-plane total irradiance model shows seasonal patterns that are different from the original horizontal global irradiance; instead of having a clear peak in the summer and a noticeable drop in the winter, the curve is more even throughout the year. This phenomenon, in addition to being specific to China's climatic conditions, can be attributed to the installation of PV modules with a south-facing orientation and a tilt angle equal to the local latitude in this study, which is intended to maximise power output, as well as to the specific solar radiation data patterns for the location. As illustrated in Fig. 4, the total solar irradiance on a tilted surface comprises three components: direct beam radiation, diffuse radiation from the sky, and reflected radiation from the ground.

Looking at the radiation data for Beijing's Typical Meteorological Year (TMY) [97], as shown in Fig. 9a, we can see that the direct beam radiation is higher in winter and spring than in summer. This difference in radiation throughout the seasons directly affects the total solar radiation amount on the tilted surface, as shown by the total radiation values in Fig. 9d. This pattern of radiation matches the changes in solar power production in Beijing in Fig. 8, where the average power output is much higher in Spring and Winter compared to Summer. To validate this conclusion, the optimal 1 $kW_p$ PV power generation of these 4 regions' estimates obtained from PVGIS was compared, as shown in Fig. 10. Even though there are small differences in the actual energy output because of variations in PV panel features and weather information, the general patterns match, which supports that the energy supply model employed in this study is correct.

Furthermore, Fig. 10 also presents a comparison of the 1 $kW_p$ PV generation for the Southampton area in the UK, calculated by our model with the estimates provided by PVGIS. Comparison with Fig. 8 also underscores the unique characteristics of rooftop PV power generation in China, attributable to its distinct geographical and climatic conditions.

Fig. 11 depicts the optimisation process of the Pareto frontier derived with NPC and annual $CO_2$ emissions over the life cycle as objective functions, with 59 sets of non-dominated solutions for Harbin, 45 sets for Beijing, 61 sets for Chengdu, and 38 sets for Xiamen obtained under the setup restriction of a maximum of 75 solutions. The reduction of annual carbon emissions (moving leftward along the horizontal axis) corresponds to a substantial increase in net present cost (NPC) (moving upward along the vertical axis). The horizontal and vertical distances between the "lowest-cost" scenario (grey marker) and the "lowest-emission" scenario (black marker) directly illustrate the implicit cost–abatement trade-off. Consequently, in high-cost regions such as Harbin and Chengdu, higher carbon prices or sector-specific subsidies are required to offset their elevated marginal abatement costs; by contrast, in Xiamen, significant emission reductions can be realised with relatively modest incentives.

The quartiles of the optimal solutions were chosen to enable a closer look at results, based on annual $CO_2$ emissions, meaning the groups of solutions at the 25th, 50th, and 75th percentiles for annual $CO_2$ emissions were organised from lowest to highest for each location, as shown in Table 4. It can be seen that for these 4 regions, benefiting from today's falling PV and battery prices, reducing reliance on the grid will only increase costs by a small amount while resulting in a significant reduction in $CO_2$ emissions, so the continued fall in PV and battery prices and overall economic growth in the future are likely to make PV systems extremely effective for carbon emission reduction in rural areas.

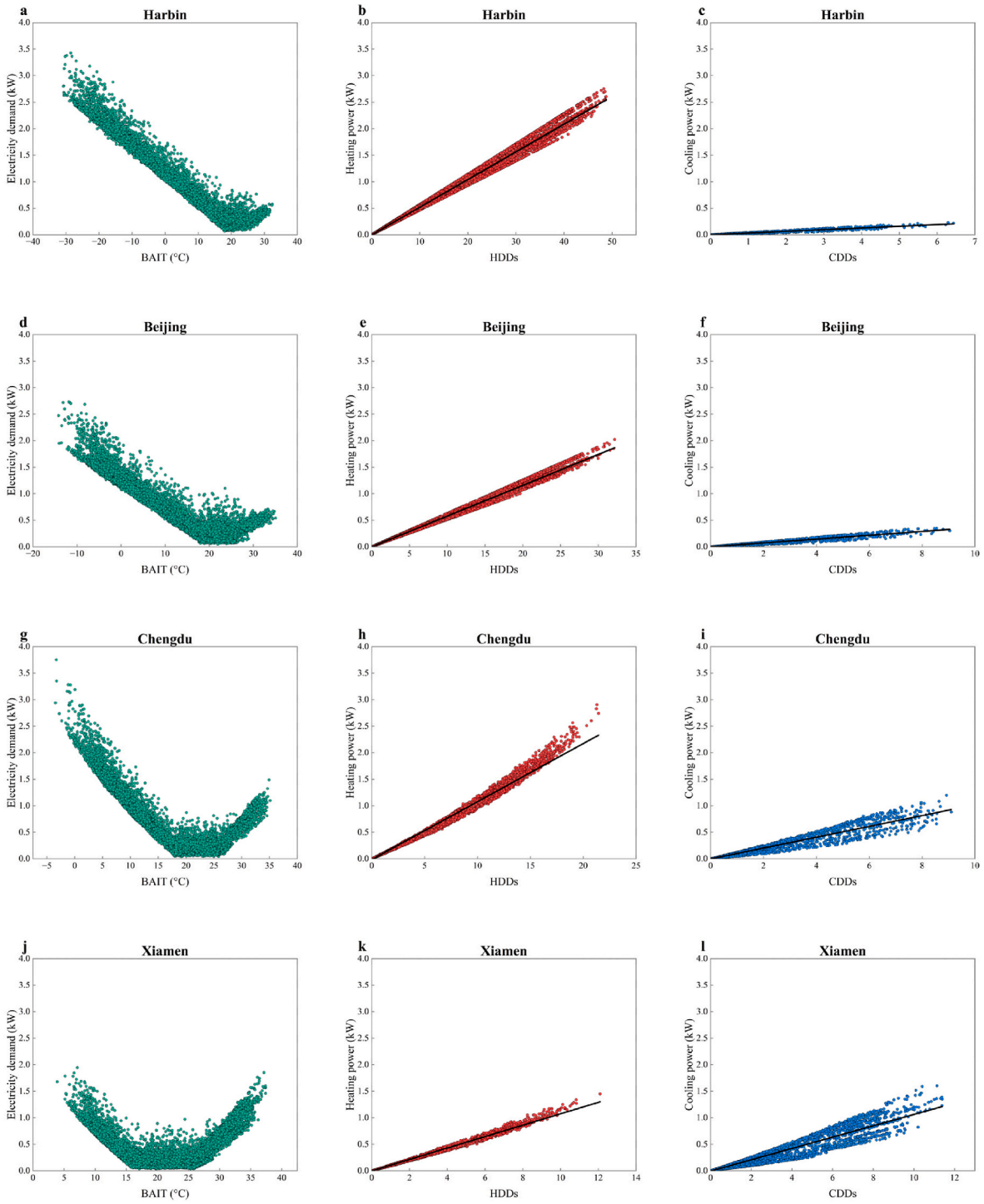



**Fig. 7.** Summary of electricity demand modelling for rural buildings in representative locations in 4 climate zones in China. **a,b,c** represent Harbin in the severe cold region, **d,e,f** represent Beijing in the cold region, **g,h,i** represents Chengdu in the hot summer and cold winter region, and **j,k,l** represents Xiamen in the hot summer and warm winter region.

Although the current NPC of the system is still high for most of the regions, which may be a major obstacle for investors, the levelised cost of electricity (LCOE) of the system is better than the national average minimum tariff of $0.07/kWh, which implies that the project is already viable under these assumptions. For simplicity, the LCOE here is calculated using the NPC divided by the total lifecycle electricity demand in the different quantiles.

Fig. 12 presents the Pareto-optimal solutions for the case-study building's LCOE versus its carbon emission factor. From the "Lower than grid price" region, it is evident that the vast majority of solutions achieve an LCOE below the national minimum electricity price of $0.07/kWh. Moreover, across all 4 regions, the optimal LCOE scenarios for the optimised solutions can already achieve at least approximately 50 % carbon emission reduction (relative to the grid emission factor of 0.6205 $kgCO_2/kWh$). Furthermore, with moderate cost increases, they can meet 80 % or even 90 % carbon reduction targets. The "Cost increase" region quantifies the additional LCOE required for the building to move from an 80 % to a 90 % carbon reduction target. Table 5 reports the cost

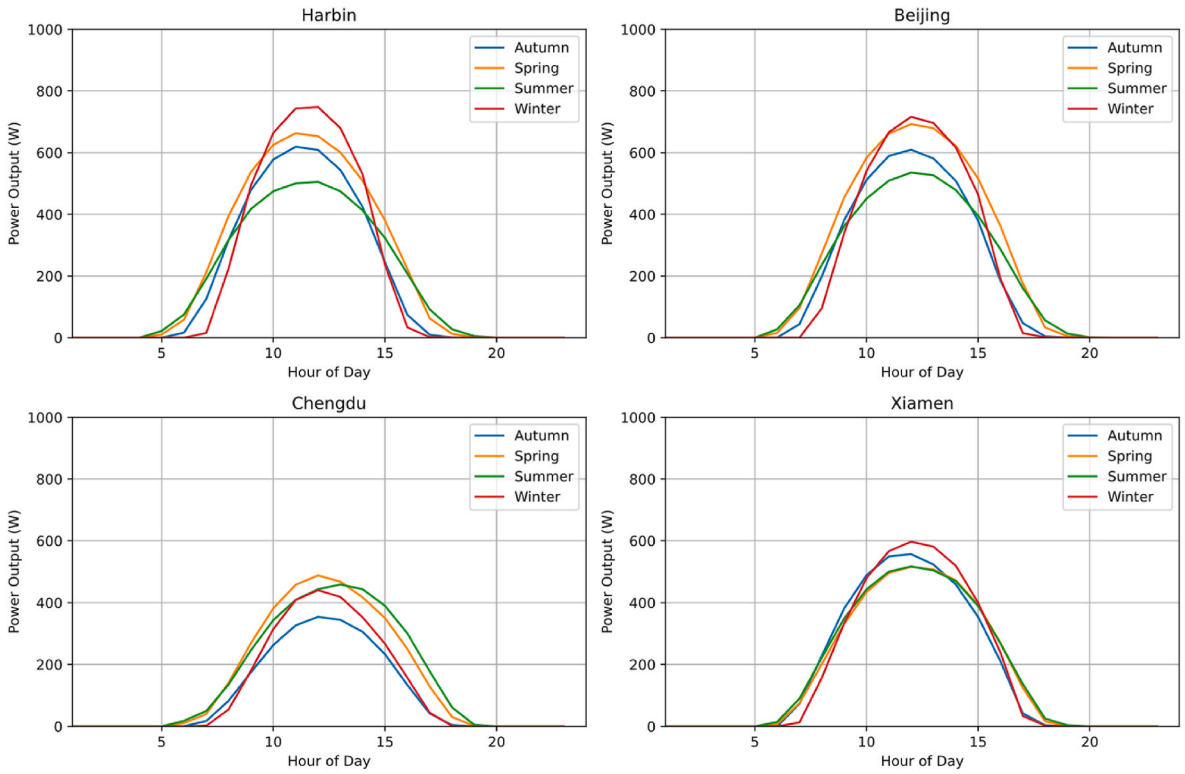



**Fig. 8.** Average daily power generation patterns by season for 1 kWp PV in Harbin, Beijing, Chengdu and Xiamen from the energy supply model in this study.

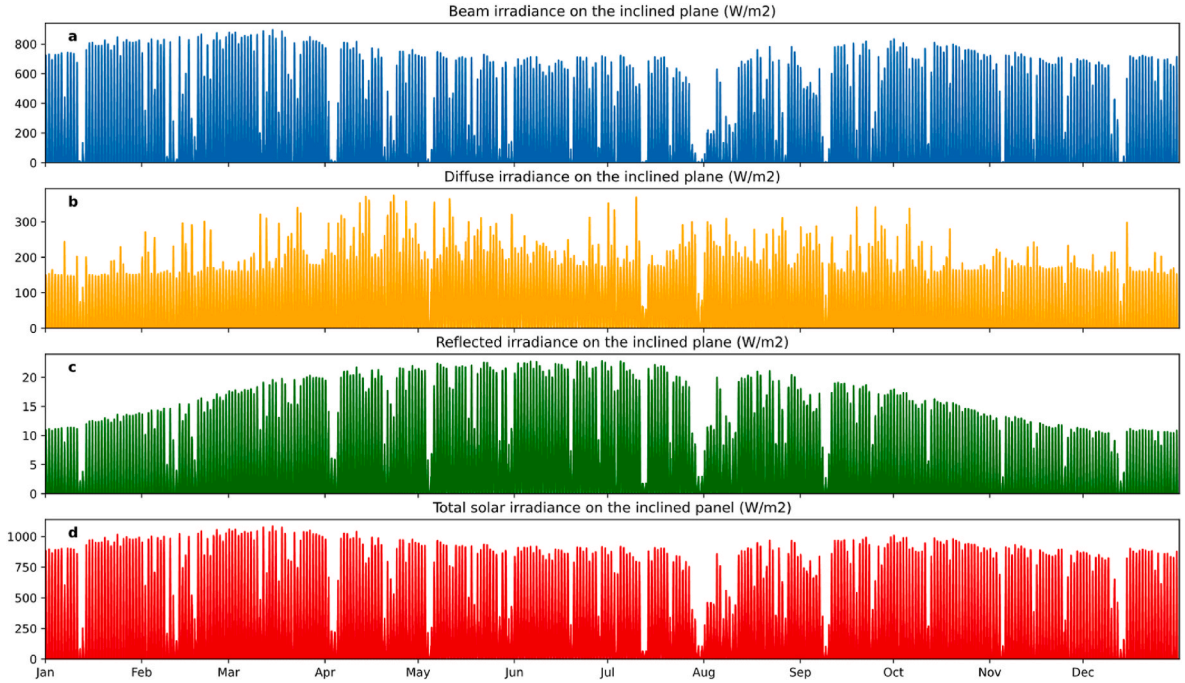



**Fig. 9.** Calculated solar radiation received on a 1 kWp inclined 42° south-facing PV panel in Beijing.

increases required for these 4 cases to attain 80 % and 90 % carbon emission reductions.

For rural buildings in Harbin, Beijing, Chengdu, and Xiamen, achieving an 80 % carbon emission reduction requires an approximate cost increase of 15 %, 16 %, 34 %, and 4 %, respectively, compared to their optimal LCOE scenarios. To further increase the carbon reduction target to 90 %, the 4 regions only need to increase the cost by 24 %–60 %, and the LCOE is lower than the grid price.

Fig. 13, using Beijing as a representative example, presents a visual representation of the optimisation results of its optimal NPC solution. It ideally captures the time-series behaviour of the operation scheduling for PV/battery/grid over a typical week in spring, summer, autumn, and winter, and it can be intuitively determined that the system has power supply reliability. Furthermore, an analysis of the seasonal operating

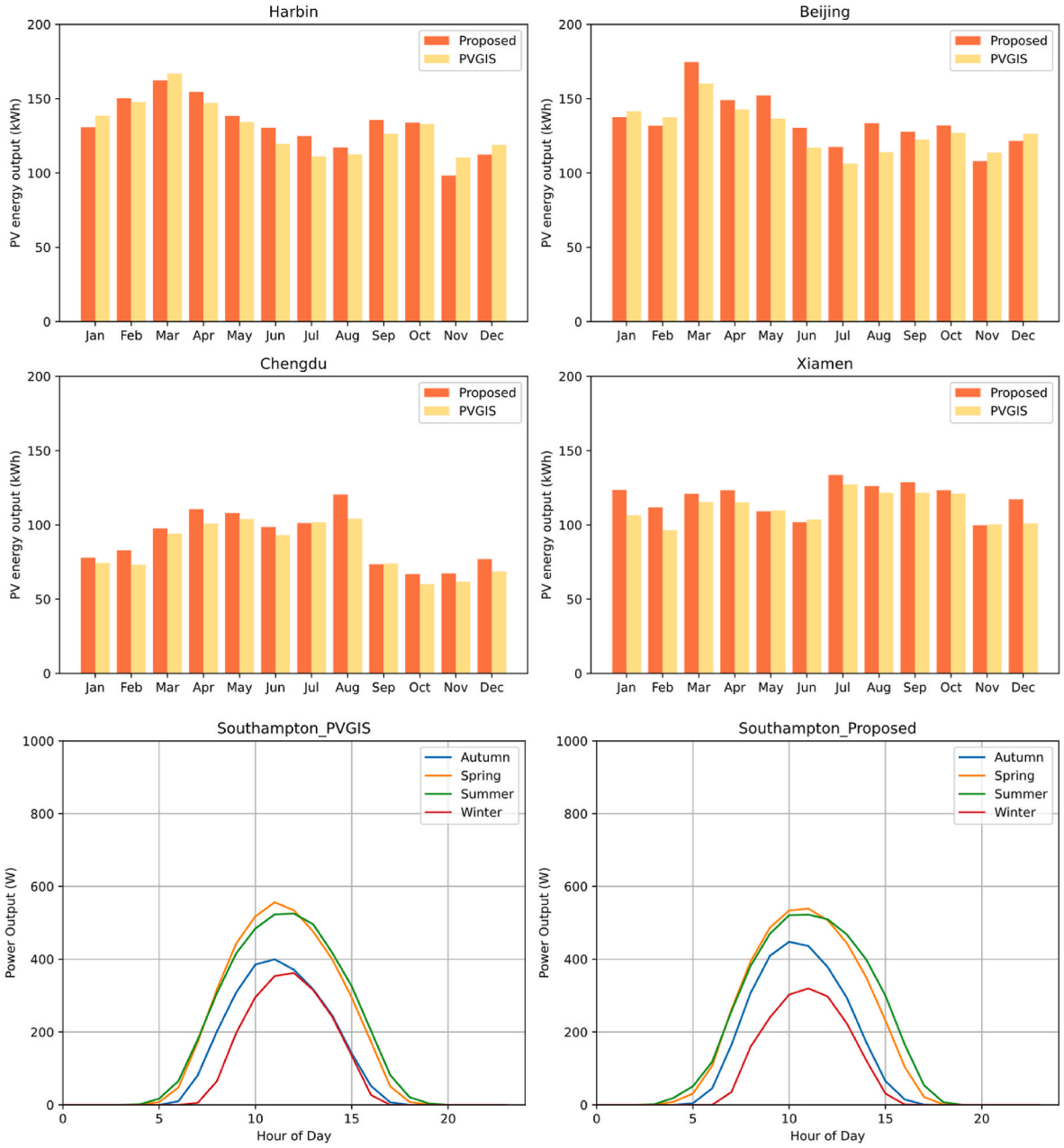



**Fig. 10.** Optimal 1kWp PV monthly energy output for Harbin, Beijing, Chengdu, Xiamen calculated by proposed model and PVGIS (above). Comparison of the proposed model and PVGIS for estimating 1kWp PV energy generation in Southampton, UK (lower).

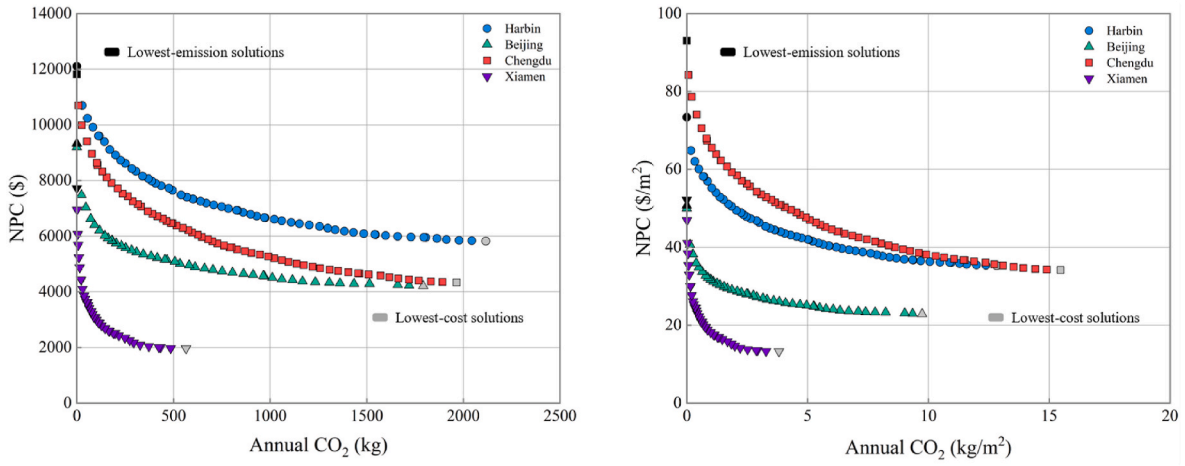



**Fig. 11.** The two-objective pareto-optimal frontier of typical rural buildings in Harbin, Beijing, Chengdu and Xiamen.

**Table 4**
The quartiles of optimisation results for grid-connected PV systems.

| Quantile | Location | System | | | | Emissions | Cost | |
|---|---|---|---|---|---|---|---|---|
| | | PV (kW) | Battery capacity (kWh) | Battery power (kW) | Grid (kWh/m$^2$) | $CO_2$ (kg/m$^2$) | NPC ($) | LCOE ($/kWh) |
| 25th | Harbin | 16 | 39 | 6 | 3.6 | 2.2 | 8067 | 0.05 |
| 50th | Harbin | 12 | 30 | 4 | 8.1 | 5.0 | 6928 | 0.04 |
| 75th | Harbin | 9 | 22 | 3 | 13.1 | 8.1 | 6230 | 0.04 |
| 25th | Beijing | 11 | 30 | 4 | 1.6 | 1.0 | 5829 | 0.05 |
| 50th | Beijing | 8 | 23 | 3 | 4.7 | 2.9 | 5013 | 0.04 |
| 75th | Beijing | 6 | 16 | 2 | 8.9 | 5.5 | 4508 | 0.04 |
| 25th | Chengdu | 22 | 29 | 4 | 4.0 | 2.5 | 7153 | 0.06 |
| 50th | Chengdu | 15 | 21 | 3 | 8.9 | 5.5 | 5811 | 0.05 |
| 75th | Chengdu | 10 | 13 | 2 | 15.5 | 9.6 | 4899 | 0.04 |
| 25th | Xiamen | 9 | 20 | 2 | 0.4 | 0.3 | 3865 | 0.05 |
| 50th | Xiamen | 7 | 15 | 2 | 1.0 | 0.6 | 3089 | 0.04 |
| 75th | Xiamen | 4 | 11 | 2 | 2.4 | 1.5 | 2427 | 0.03 |

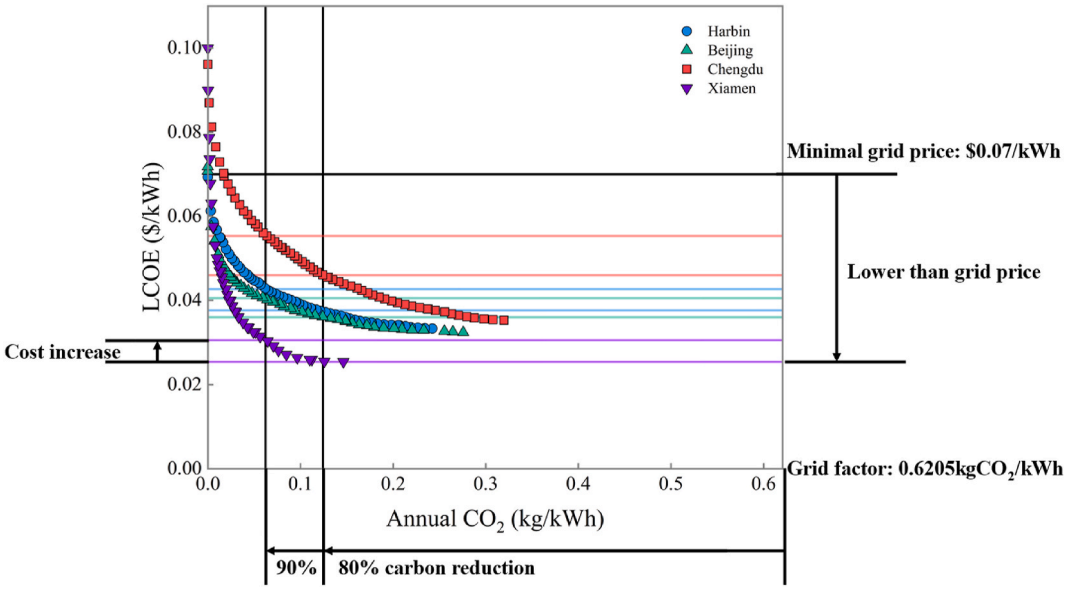



**Fig. 12.** The optimal solution of LCOE and carbon emission factors of the case building.

**Table 5**
The cost increase to achieve 80 % and 90 % carbon reductions in the optimised solution.

| Location | LCOE ($/kWh) | Cost increase ratio (%) | Annual $CO_2$ (kg/kWh) | Carbon reduction (%) |
|---|---|---|---|---|
| Harbin | 0.033 | – | 0.242 | 61 |
| Harbin | 0.038 | 15.2 | 0.121 | 81 |
| Harbin | 0.043 | 30.0 | 0.062 | 90 |
| Beijing | 0.032 | – | 0.276 | 56 |
| Beijing | 0.037 | 15.6 | 0.116 | 81 |
| Beijing | 0.041 | 28.1 | 0.059 | 91 |
| Chengdu | 0.035 | – | 0.320 | 48 |
| Chengdu | 0.047 | 34.3 | 0.118 | 81 |
| Chengdu | 0.056 | 60 | 0.060 | 90 |
| Xiamen | 0.025 | – | 0.146 | 77 |
| Xiamen | 0.026 | 4 | 0.112 | 82 |
| Xiamen | 0.031 | 24 | 0.056 | 91 |

patterns reveals that the PV generation curve is consistent with the generation pattern shown in Fig. 8. The grid participates during spring, autumn, and winter to meet load demand, while in summer, due to lower load demand, priority is given to PV and batteries.

Although the model demonstrates that optimised systems can achieve significant carbon reduction with a LCOE lower than grid prices, benefiting from the sharp decline in photovoltaic and battery prices in recent years and the strong decarbonisation capabilities enabled by integrated configurations utilizing solar power, it is crucial not to overlook the impacts of embodied energy and embodied carbon. These factors must be incorporated into a comprehensive life cycle analysis to revisit the environmental benefits. It is worth pointing out that, due to the uncertainties involved in the assessment of embodied energy and embodied carbon at multiple points of production, transport, installation and decommissioning, their values should be presented as intervals rather than as a single fixed value. Therefore, this study adopts data from prior research on embodied energy and embodied carbon related to residential PV and batteries [90,125–128], along with findings from institutional reports, to establish the data used for the calculations in this study, as detailed in Table 6.

Fig. 14 illustrates the optimal Pareto frontier with NPC and annual $CO_2$ emissions as objective functions, which considers the embodied carbon from PV and battery. The grey scatter points indicate that, after incorporating embodied carbon, the original configurations are no longer optimal according to Pareto criterion. Specifically, in the context of full lifecycle embodied carbon emissions, when the size of PV panels and batteries exceeds a certain critical threshold, their embodied carbon burden exceeds the carbon reduction benefits gained from displacing high-carbon grid electricity, thereby triggering a nonlinear decline in overall $CO_2$ reduction performance and an inflection point on the Pareto frontier (representing the optimal configurations according to the two objectives, NPC and $CO_2$ emissions).

This phenomenon highlights that, in renewable energy system planning and applications, it is essential to appropriately size technologies and to account for full lifecycle embodied carbon in order to achieve a synergistic optimisation of carbon reduction benefits and economic efficiency.

After eliminating suboptimal solutions, we identified potential options and demonstrated the impact of embodied carbon by calculating

**Table 6**
The uncertain range of embodied energy and carbon emission from PV and batteries.

| Components | Embodied energy | Embodied carbon |
|---|---|---|
| PV | 3000-7000 kWh/kWp | 500-1000 $kgCO_2$/kWp |
| Battery | 1100-1400 kWh/kWh | 150-225 $kgCO_2$/kWh |

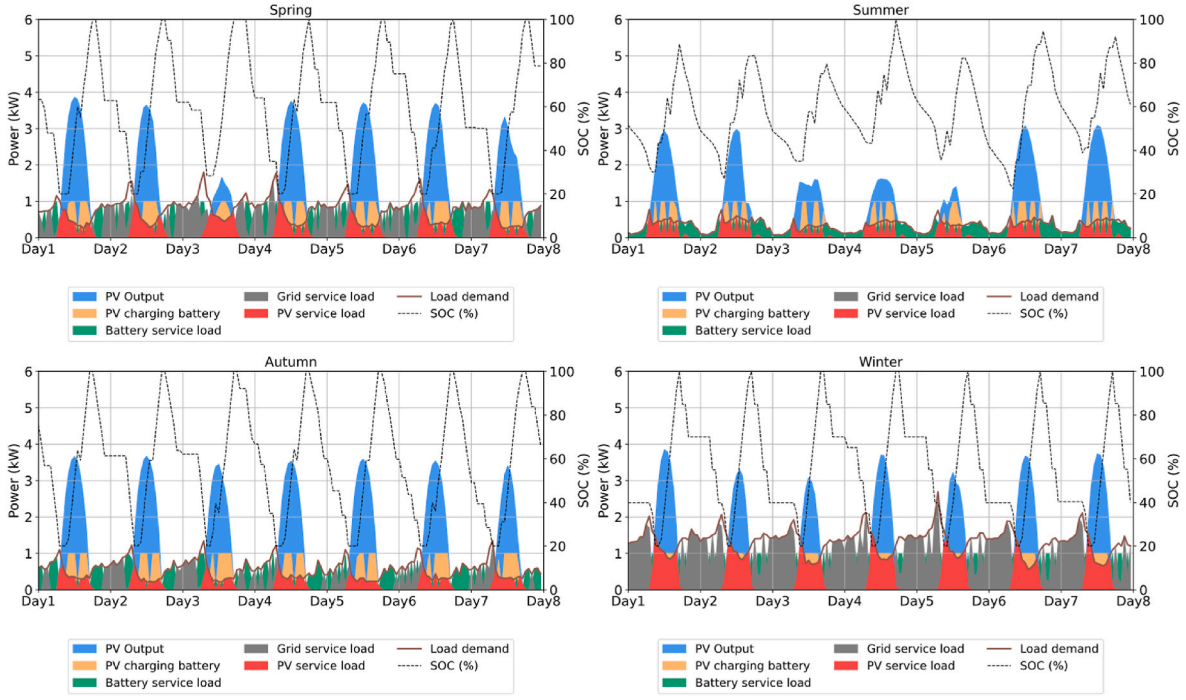



**Fig. 13.** Time series behaviour of a typical week of PV/battery/grid operation scheduling during summer/spring/winter period of the year in Beijing.

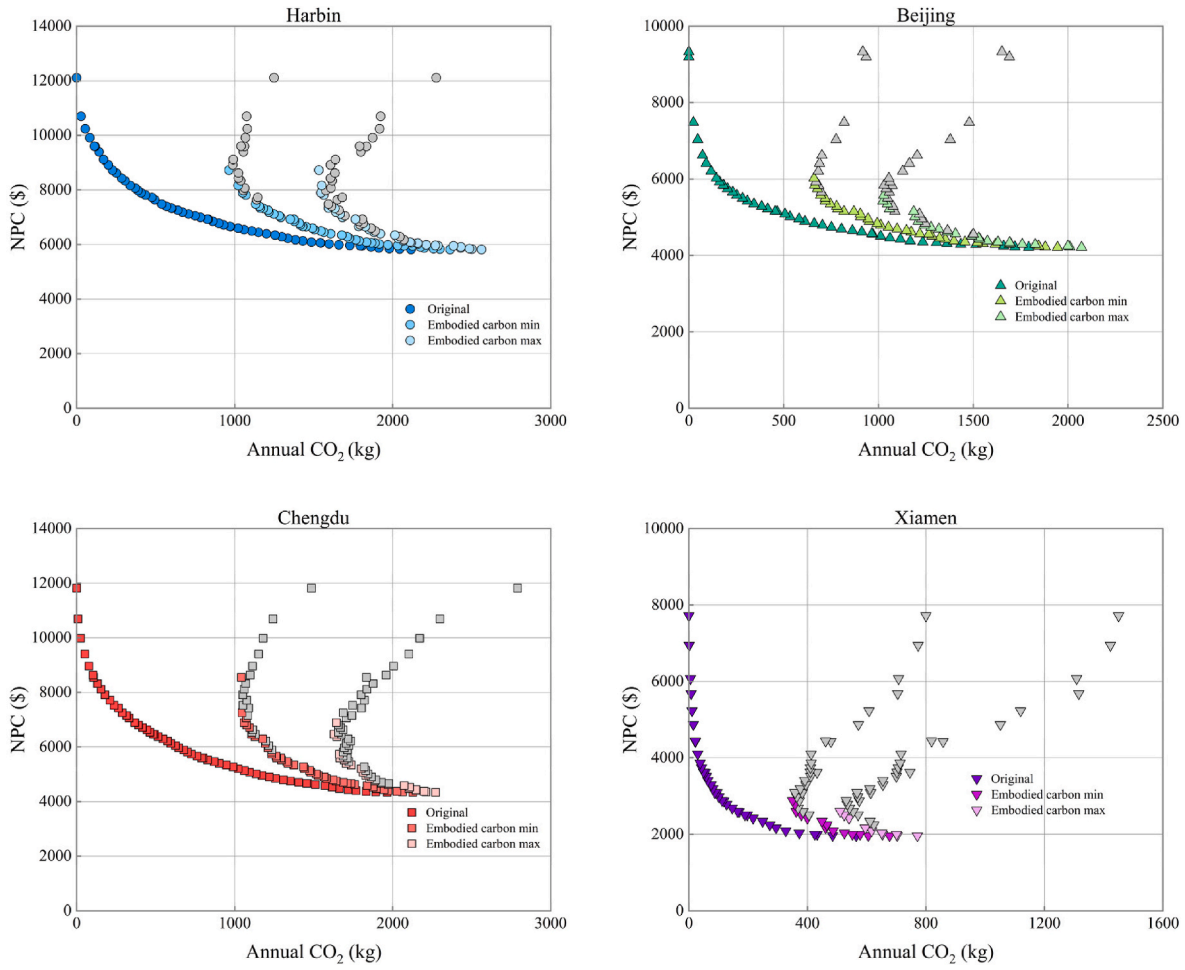



**Fig. 14.** The two-objective pareto-optimal frontiers for 4 regions considering embodied carbon.

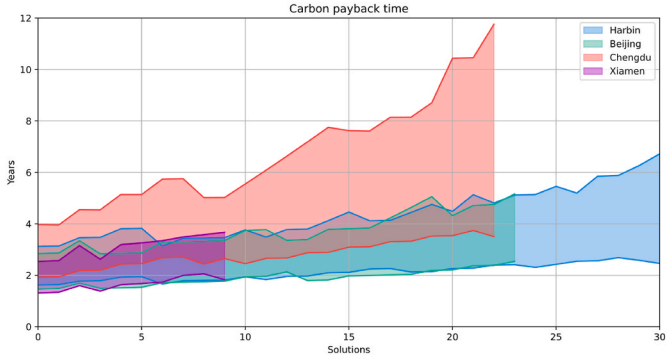



**Fig. 15.** The CPT estimation of PV plus battery system in the 4 regions. The x-axis represents configuration solutions of different scales, arranged from smallest to largest.

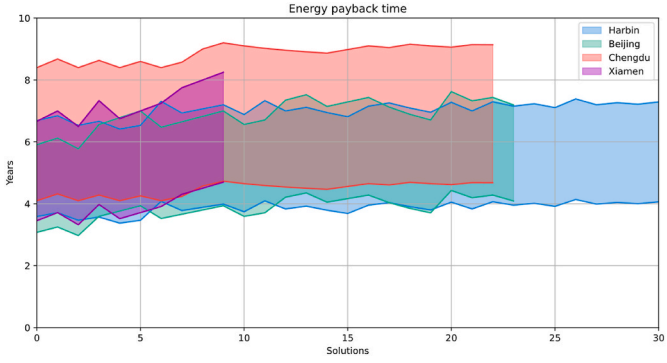



**Fig. 16.** The EBPT estimation for PV plus battery system in the 4 regions. The x-axis represents configuration solutions of different scales, arranged from smallest to largest.

the system's carbon payback time (CPT). CPT is typically defined as the ratio of the total embodied carbon over the system's lifecycle to the annual net carbon reduction achieved by offsetting fossil fuel-based generation through electricity production, with the results shown in Fig. 15. Due to the significant uncertainties in embodied carbon estimates, the CPT of optimised solutions across different regions exhibits considerable variation, from approximately 2 years–12 years. This finding emphasizes that, when evaluating system performance, it is crucial to consider not only the operational carbon savings but also the substantial effect of embodied carbon in delaying net carbon benefits, which is essential for informed system design and deployment. It further suggests that pursuing excessively large configurations to achieve high carbon reductions may not be necessary, especially given the potential for further declines in grid carbon emission factors in the future.

Moreover, the energy payback time (EPBT) is estimated based on the energy consumed throughout the system's lifecycle and the electricity generated. It is typically defined as the amount of time required for an energy system to "repay" the total energy consumed during its manufacturing, transportation, installation, operation, and decommissioning processes with the energy it generates during its lifetime, to reflect the impact of the embodied energy of the components.

Fig. 16 presents the EPBT results for photovoltaic-plus-battery systems across the 4 case study regions. The substantial uncertainty in the implicit energy of both the PV panels and batteries results in significant variability in the EPBT, ranging from a minimum of 3 years in Beijing to a maximum of 9 years in Chengdu, with differences exceedingly twofold even within the same region.

In summary, regional characteristics and the uncertainties in embodied energy and embodied carbon must be fully taken into account during system design and deployment to accurately assess the overall energy efficiency and environmental sustainability of the system, and to provide proper guidance on the optimal system configuration.

## 6. Conclusion

This study set out to address the issues associated with designing and operating grid-connected PV systems in rural buildings across China's diverse climate zones. By integrating open-source, modular methods for demand and supply modelling, knowledge gaps identified in literature have been addressed by merging a regression-based load estimation approach with a multi-objective optimisation model, underscoring the importance of simultaneous consideration of supply- and demand-side factors to inform investments and policy related decisions toward decarbonisation targets.

Key findings highlight that:

- optimised system sizing and operation are strongly influenced by local building attributes and climate-specific conditions. Simplifying heating, cooling and base load modelling at hourly resolution using physics-informed statistical and machine learning tools is crucial to support the optimisation process.
- creating models that are composed by logical blocks and are relatively simple to calibrate on measured data enables the use of the tool not only in design but also in operation.
- around 80 % and 90 % reduction compared to grid carbon emissions can be achieved with an increase of the LCOE in the range between 4 % and 60 %.
- the carbon payback time for systems considering embodied carbon varies from about 2 to 12 years, depending on location and assumptions, due to uncertainty.

These results indicate substantial near-term decarbonisation potential for rural grid-connected buildings under current technology costs and grid carbon intensities. Moreover, the significance of tailoring systems to local contexts is evident, as climate variability and occupant behaviour patterns shape the solutions' technical and economic viability. By combining physics-informed statistical approaches and machine learning algorithms for demand forecasting with robust multi-objective optimisation, this work provides a scalable method to tackle decarbonisation problem in rural grid-connected building. The framework can be extended to large-scale applications targeting building stock decarbonisation while being transparent and adaptable. Open-source modelling also ensures broader accessibility and fosters collaborative development, promoting continual refinement as technologies and policy frameworks evolve.

Looking ahead, additional research may investigate numerous potential directions. First, the modelling framework's flexibility makes it well suited for global applications where open data is available, with minimal adjustments for local climate conditions, building materials, or occupant behaviour.

Second, the model's scalability, within the context of evolving decarbonisation pathways for energy systems, provides the possibility to conduct more sophisticated scenario analyses. For example, while considering the installation of thermal energy storage for heating and domestic hot water (DHW) as well as cold energy storage for cooling, it is important to note that these measures will add complexity to the design and operation of such systems; however, the resulting lower $CO_2$ emissions may provide greater environmental and economic benefits. Such an approach would enable stakeholders to quantify the cost-effectiveness of solutions under uncertain and evolving conditions. Additionally, expanding the scope to incorporate dynamic pricing mechanisms, market interactions, and grid constraints could offer deeper insights into system-wide impacts.

Overall, the findings of this research can serve as a foundation for both researchers and practitioners to further refine methods, calibrate tools with real-world data, and explore emerging solutions in energy storage, demand response, and heat pump technologies. By simplifying complex modelling processes and ensuring transparency, our framework aims to accelerate the adoption of decentralised renewable energy systems and support the transition toward low-carbon, cost-effective rural energy infrastructures.

## CRediT authorship contribution statement

**Rundong Liao:** Writing – original draft, Visualization, Validation, Software, Conceptualization. **Massimiliano Manfren:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Benedetto Nastasi:** Writing – review & editing, Supervision, Project administration.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A

To eliminate the bias introduced by seasonal fluctuations in solar irradiance, wind speed and humidity, we perform a joint regression of these meteorological variables against ambient temperature using data from all study regions across China's various climate zones. These functions are then used as "counterfactual" meteorological inputs in the calculation of BAIT. This approach is in line with the one proposed by Staffel et al. [23] for the

'Demand Ninja' tool. Consequently, whether actual insolation diminishes in winter or humidity rises in summer, the generated values follow this "purified" curve, ensuring that the influence of meteorological factors on the energy-consumption model reflects temperature alone, without amplification or attenuation by seasonal extremes.

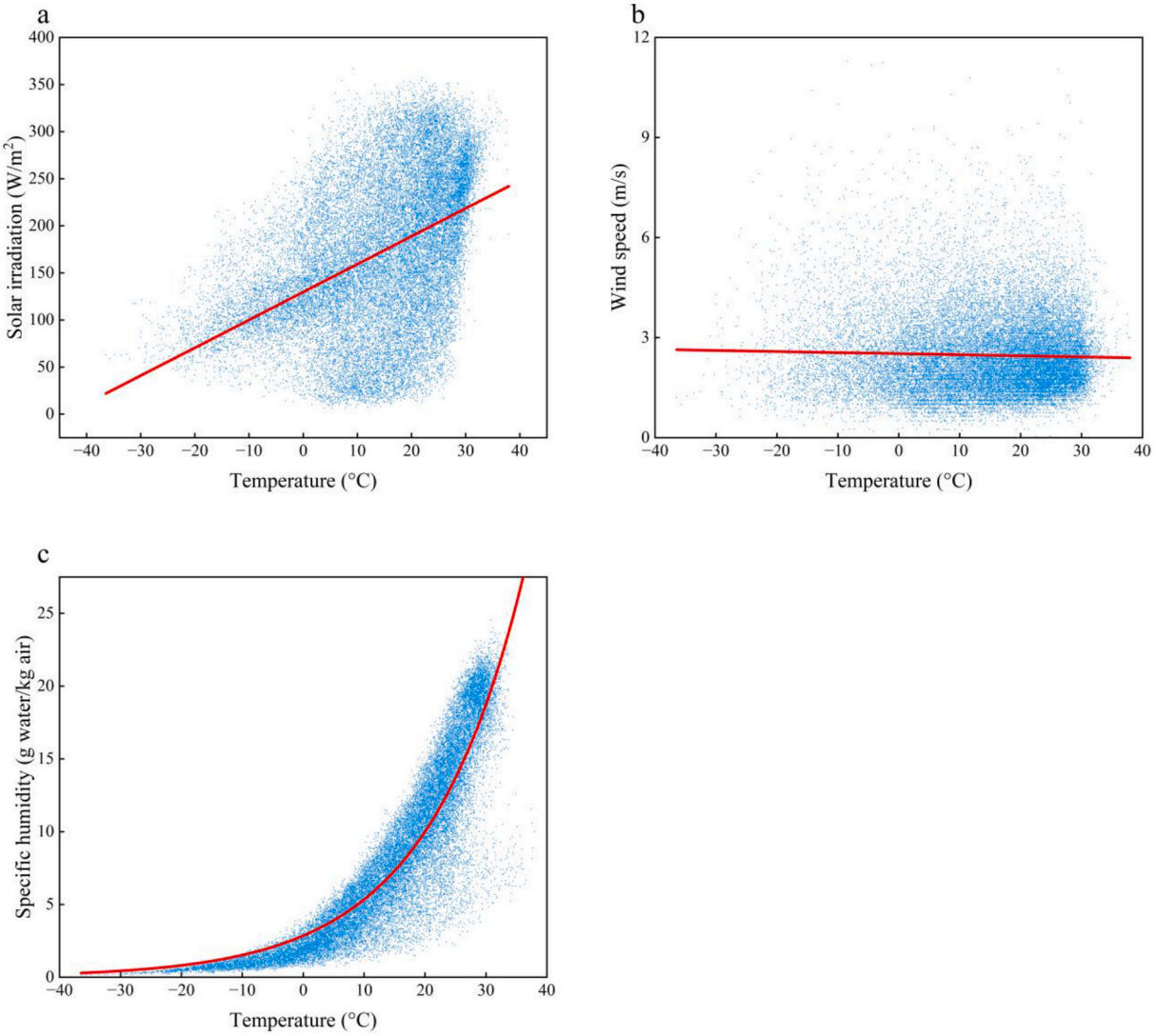



**Fig. 1.** Correlation regressions of temperature with other weather variables using representative locations from each of the climate zones in China. **a** Considering solar irradiance, **b** Considering wind speed, simple linear regressions are indicated by the red line. **c** Considering specific humidity, log-linear regressions are indicated by the red line. All red lines are used to formulate the relationships of the weather counterfactuals in this study.

## Appendix B

**Table 1**
Optimisation parameters used in this paper

| Parameter | Input value | Units | Description |
|---|---|---|---|
| Discount rate | 0.04 | – | Annual discount rate for lifecycle cost |
| Analysis horizon | 20 | Years | Total number of years for present-value calculation |
| PV capital cost | 100 | $/kW | Investment cost per kW of PV |
| Battery energy cost | 53 | $/kWh | Investment cost per kWh of battery energy capacity |
| Battery power cost | 50 | $/kW | Investment cost per kW of battery power rating |
| Electricity price | 0.07 | $/kWh | Grid electricity purchase price |
| Emission factor | 0.6205 | kg $CO_2$/kWh | $CO_2$ emission factor for grid electricity |
| Initial SOC fraction | 0.5 | – | State–of–charge at time zero (fraction of energy capacity) |
| Charge/discharge efficiency | 0.95/0.95 | – | Battery round-trip efficiencies |
| Minimum SOC | 0.2 | – | Lower SOC limit (20 % of energy capacity) |
| Battery replacement | True | | Whether to include replacements at year 8 and 16 |
| CPLEX time limit | 600 | seconds | Global time limit per solve call (10 min) |
| CPLEX MIP gap tolerance | 0.001 | – | Relative MIP optimality gap tolerance (0.1 %) |

(*continued on next page*)

**Table 1** (*continued*)

| Parameter | Input value | Units | Description |
|---|---|---|---|
| CPLEX thread count | 4 | – | Number of parallel threads to use |
| CPLEX node display | 2 | – | Display level for node log (0 = none, 2 = summary every 100 nodes) |
| CPLEX emphasis | 4 | – | MIP emphasis (4 = optimality over feasibility) |

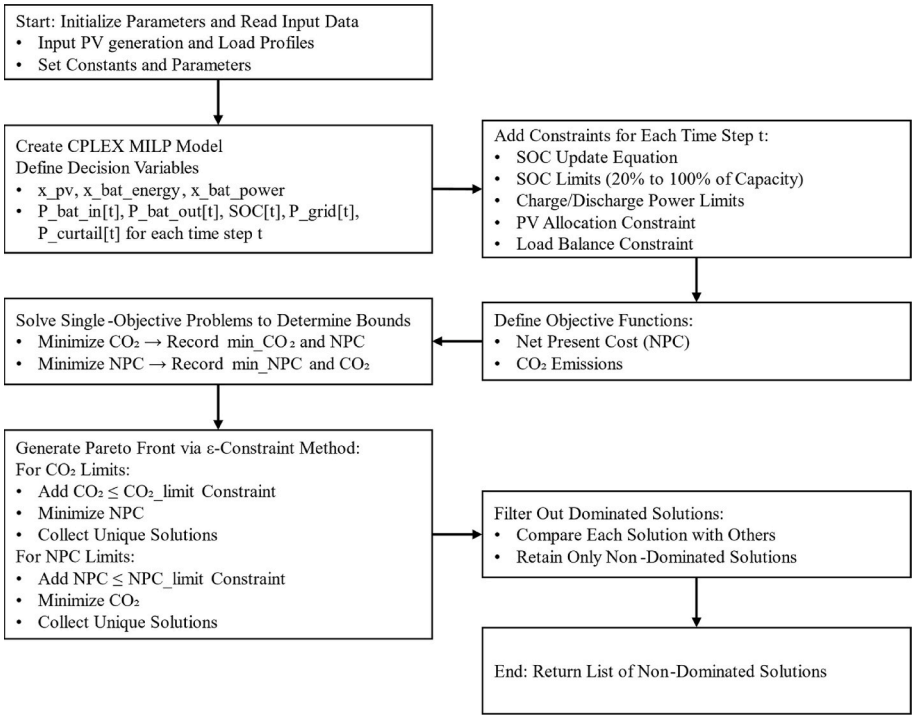



**Fig. 2.** Flowchart of the optimisation steps.

## Data availability

Data will be made available on request.

## References


[1] Østergaard PA, Duic N, Noorollahi Y, Kalogirou S. Advances in renewable energy for sustainable development. Renew Energy 2023;219:119377. https://doi.org/10.1016/j.renene.2023.119377.

[2] Seyedabadi MR, Karrabi M, Shariati M, Karimi S, Maghrebi M, Eicker U. Global building life cycle assessment: comparative study of steel and concrete frames across European Union, USA, Canada, and Australia building codes. Energy Build 2024;304:113875. https://doi.org/10.1016/j.enbuild.2023.113875.

[3] Wang Y, Wang Z, Shuai J, Shuai C. Can digitalization alleviate multidimensional energy poverty in rural China? Designing a policy framework for achieving the sustainable development goals. Sustain Prod Consum 2023;39:466–79. https://doi.org/10.1016/j.spc.2023.05.031.

[4] China Association of Building Energy Efficiency. Research report on carbon emissions in China's urban and rural construction sector in 2024. 2025.

[5] Wang S, de Souza CB, Golubchikov O. Recent advances in decarbonising heating in rural China: a review. Renew Sustain Energy Rev 2025;210:115282. https://doi.org/10.1016/j.rser.2024.115282.

[6] Wu S, Zheng X, You C, Wei C. Household energy consumption in rural China: historical development, present pattern and policy implication. J Clean Prod 2019;211:981–91. https://doi.org/10.1016/j.jclepro.2018.11.265.

[7] Wen Q, Liu G, Rao Z, Liao S. Applications, evaluations and supportive strategies of distributed energy systems: a review. Energy Build 2020;225:110314. https://doi.org/10.1016/j.enbuild.2020.110314.

[8] Creutzig F, Agoston P, Goldschmidt JC, Luderer G, Nemet G, Pietzcker RC. The underestimated potential of solar energy to mitigate climate change. Nat Energy 2017;2:17140. https://doi.org/10.1038/nenergy.2017.140.

[9] Luo X, Yang Y, Liu Y, Zhao T. Classification of energy use patterns and multi-objective optimal scheduling of flexible loads in rural households. Energy Build 2023;283:112811. https://doi.org/10.1016/j.enbuild.2023.112811.

[10] Topić D, Knežević G, Fekete K. The mathematical model for finding an optimal PV system configuration for the given installation area providing a maximal lifetime profit. Sol Energy 2017;144:750–7. https://doi.org/10.1016/j.solener.2017.02.011.

[11] Barakat S, Ibrahim H, Elbaset AA. Multi-objective optimization of grid-connected PV-wind hybrid system considering reliability, cost, and environmental aspects. Sustain Cities Soc 2020;60:102178. https://doi.org/10.1016/j.scs.2020.102178.

[12] Luo Z, Yang S, Xie N, Xie W, Liu J, Souley Agbodjan Y, et al. Multi-objective capacity optimization of a distributed energy system considering economy, environment and energy. Energy Convers Manag 2019;200:112081. https://doi.org/10.1016/j.enconman.2019.112081.

[13] Al-Shahri OA, Ismail FB, Hannan MA, Lipu MSH, Al-Shetwi AQ, Begum RA, et al. Solar photovoltaic energy optimization methods, challenges and issues: a comprehensive review. J Clean Prod 2021;284:125465. https://doi.org/10.1016/j.jclepro.2020.125465.

[14] Ridha HM, Gomes C, Hizam H, Ahmadipour M, Heidari AA, Chen H. Multi-objective optimization and multi-criteria decision-making methods for optimal design of standalone photovoltaic system: a comprehensive review. Renew Sustain Energy Rev 2021;135:110202. https://doi.org/10.1016/j.rser.2020.110202.

[15] Starke AR, Cardemil JM, Escobar R, Colle S. Multi-objective optimization of hybrid CSP+PV system using genetic algorithm. Energy (Calg) 2018;147:490–503. https://doi.org/10.1016/j.energy.2017.12.116.

[16] Sharafi M, Elmekkawy TY. Multi-objective optimal design of hybrid renewable energy systems using PSO-simulation based approach. Renew Energy 2014;68:67–79. https://doi.org/10.1016/j.renene.2014.01.011.

[17] Attia AM, Al Hanbali A, Saleh HH, Alsawafy OG, Ghaithan AM, Mohammed A. A multi-objective optimization model for sizing decisions of a grid-connected photovoltaic system. Energy (Calg) 2021;229:120730. https://doi.org/10.1016/j.energy.2021.120730.

[18] Wankouo Ngouleu CA, Koholé YW, Fohagui FCV, Tchuen G. Techno-economic analysis and optimal sizing of a battery-based and hydrogen-based standalone photovoltaic/wind hybrid system for rural electrification in Cameroon based on meta-heuristic techniques. Energy Convers Manag 2023;280:116794. https://doi.org/10.1016/j.enconman.2023.116794.

[19] Zhao Y, Ding X, Wu Z, Yin S, Fan Y, Ge J. Impact of urban form on building energy consumption in different climate zones of China. Energy Build 2024;320:114579. https://doi.org/10.1016/j.enbuild.2024.114579.

[20] Silinto BF, van der Laag Yamu C, Zuidema C, Faaij APC. Hybrid renewable energy systems for rural electrification in developing countries: a review on energy

system models and spatial explicit modelling tools. Renew Sustain Energy Rev 2025;207:114916. https://doi.org/10.1016/j.rser.2024.114916.
[21] Pfenninger S. Open code and data are not enough: understandability as design goal for energy system models. Progress in Energy 2024;6:033002. https://doi.org/10.1088/2516-1083/ad371e.
[22] Pasichnyi O, Wallin J, Kordas O. Data-driven building archetypes for urban building energy modelling. Energy (Calg) 2019;181:360–77. https://doi.org/10.1016/j.energy.2019.04.197.
[23] Staffell I, Pfenninger S, Johnson N. A global model of hourly space heating and cooling demand at multiple spatial scales. Nat Energy 2023;8:1328–44. https://doi.org/10.1038/s41560-023-01341-5.
[24] Manfren M, James PAB, Tronchin L. Data-driven building energy modelling – an analysis of the potential for generalisation through interpretable machine learning. Renew Sustain Energy Rev 2022;167:112686. https://doi.org/10.1016/j.rser.2022.112686.
[25] Chen Y, Guo M, Chen Z, Chen Z, Ji Y. Physical energy and data-driven models in building energy prediction: a review. Energy Rep 2022;8:2656–71. https://doi.org/10.1016/j.egyr.2022.01.162.
[26] Das HP, Lin Y-W, Agwan U, Spangher L, Devonport A, Yang Y, et al. Machine learning for smart and energy-efficient buildings. Environmental Data Science 2024;3:e1. https://doi.org/10.1017/eds.2023.43.
[27] Ma Z, Jiang G, Hu Y, Chen J. A review of physics-informed machine learning for building energy modeling. Appl Energy 2025;381:125169. https://doi.org/10.1016/j.apenergy.2024.125169.
[28] Magni M, Ochs F, de Vries S, Maccarini A, Sigg F. Detailed cross comparison of building energy simulation tools results using a reference office building as a case study. Energy Build 2021;250:111260. https://doi.org/10.1016/j.enbuild.2021.111260.
[29] Flora Montgomery, Potvin Corey, McGovern Amy, Handler Shawn. Comparing explanation methods for traditional machine learning models part 1: an overview of current methods and quantifying their disagreement. 2022.
[30] Olu-Ajayi R, Alaka H, Sulaimon I, Sunmola F, Ajayi S. Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques. J Build Eng 2022;45:103406. https://doi.org/10.1016/j.jobe.2021.103406.
[31] Zhang W, Cai Y, Zhan H, Yang M, Zhang W. Multi-energy load forecasting for small-sample integrated energy systems based on neural network Gaussian process and multi-task learning. Energy Convers Manag 2024;321:119027. https://doi.org/10.1016/j.enconman.2024.119027.
[32] Rudin C, Chen C, Chen Z, Huang H, Semenova L, Zhong C. Interpretable machine learning: fundamental principles and 10 grand challenges. Stat Surv 2022;16:1–85. https://doi.org/10.1214/21-SS133.
[33] Liu X, Yang Z, Guo Y, Li Z, Xu X. A novel correlation feature self-assigned Kolmogorov-arnold networks for multi-energy load forecasting in integrated energy systems. Energy Convers Manag 2025;325:119388. https://doi.org/10.1016/j.enconman.2024.119388.
[34] Manfren M, Nastasi B. From in-situ measurement to regression and time series models: an overview of trends and prospects for building performance modelling. AIP Conf Proc 2019;2123:020100. https://doi.org/10.1063/1.5117027.
[35] ASHRAE. ASHRAE guideline 14-2014: measurement of energy, demand, and water savings. Atlanta, GA, USA: American Society of Heating, Refrigerating and Air-Conditioning Engineers; 2014.
[36] Gelaro R, McCarty W, Suárez MJ, Todling R, Molod A, Takacs L, et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J Clim 2017;30:5419–54. https://doi.org/10.1175/JCLI-D-16-0758.1.
[37] Rodrigues E, Fernandes MS, Carvalho D. Future weather generator for building performance research: an open-source morphing tool and an application. Build Environ 2023;233:110104. https://doi.org/10.1016/j.buildenv.2023.110104.
[38] Yan D, Wu Y, Malik J, Hong T. Ten questions on future and extreme weather data for building simulation and analysis in a changing climate. Build Environ 2025;269:112461. https://doi.org/10.1016/j.buildenv.2024.112461.
[39] Reinhart C. Linking energy use to local climate. Nat Energy 2023;8:1311–2. https://doi.org/10.1038/s41560-023-01407-4.
[40] Vesterberg J, Andersson S, Olofsson T. Robustness of a regression approach, aimed for calibration of whole building energy simulation tools. Energy Build 2014;81:430–4. https://doi.org/10.1016/j.enbuild.2014.06.035.
[41] Vesterberg J, Andersson S, Olofsson T. Calibration of low-rise multifamily residential simulation models using regressed estimations of transmission losses. J Build Perform Simul 2016;9:304–15. https://doi.org/10.1080/19401493.2015.1067257.
[42] Chong A, Gu Y, Jia H. Calibrating building energy simulation models: a review of the basics to guide future work. Energy Build 2021;253:111533. https://doi.org/10.1016/j.enbuild.2021.111533.
[43] ISO 16346:2013. Energy performance of buildings — assessment of overall energy performance. 2013.
[44] Oh S, Gardner JF. Large scale energy signature analysis: tools for utility managers and planners. Sustainability (Basel) 2022;14. https://doi.org/10.3390/su14148649.
[45] Acquaviva A, Apiletti D, Attanasio A, Baralis E, Bottaccioli L, Castagnetti FB, et al. Energy signature analysis: knowledge at your fingertips. Big data (BigData Congress), 2015. IEEE: IEEE International Congress on; 2015. p. 543–50.
[46] Kim Y-J, Waegel A, Hakkarainen M, Yi YK, Braham W. Understanding HVAC system runtime of U.S. homes: an energy signature analysis using smart thermostat data. Build Simulat 2025;18:235–58. https://doi.org/10.1007/s12273-024-1203-9.
[47] Eggimann S, Fiorentini M. Transferring energy signatures across space and time to assess their viability for rapid urban energy demand estimation. Energy Build 2024;316:114348. https://doi.org/10.1016/j.enbuild.2024.114348.
[48] Manfren M, Nastasi B. Parametric performance analysis and energy model calibration workflow integration—A scalable approach for buildings. Energies 2020;13. https://doi.org/10.3390/en13030621.
[49] Manfren M, Nastasi B, Piana E, Tronchin L. On the link between energy performance of building and thermal comfort: an example. AIP Conf Proc 2019;2123:20066. https://doi.org/10.1063/1.5116993.
[50] Lee K, Baek H-J, Cho C. The estimation of base temperature for heating and cooling degree-days for South Korea. J Appl Meteorol Climatol 2014;53:300–9.
[51] Meng Q, Mourshed M. Degree-day based non-domestic building energy analytics and modelling should use building and type specific base temperatures. Energy Build 2017;155:260–8. https://doi.org/10.1016/j.enbuild.2017.09.034.
[52] Lund H, Arler F, Østergaard P, Hvelplund F, Connolly D, Mathiesen B, et al. Simulation versus optimisation: theoretical positions in energy system modelling. Energies 2017;10:840.
[53] Pfenninger S, Hirth L, Schlecht I, Schmid E, Wiese F, Brown T, et al. Opening the black box of energy modelling: strategies and lessons learned. Energy Strategy Rev 2018;19:63–71. https://doi.org/10.1016/j.esr.2017.12.002.
[54] Manfren M, Gonzalez-Carreon KM, James PAB. Interpretable data-driven methods for building energy modelling - a review of critical connections and gaps. Energies 2024;17. https://doi.org/10.3390/en17040881.
[55] Tronchin L, Manfren M, Nastasi B. Energy analytics for supporting built environment decarbonisation. Energy Proc 2019;157:1486–93. https://doi.org/10.1016/j.egypro.2018.11.313.
[56] Fumo N, Torres MJ, Broomfield K. A multiple regression approach for calibration of residential building energy models. J Build Eng 2021;43:102874. https://doi.org/10.1016/j.jobe.2021.102874.
[57] Abdelrahman M, Macatulad E, Lei B, Quintana M, Miller C, Biljecki F. What is a digital twin anyway? Deriving the definition for the built environment from over 15,000 scientific publications. Build Environ 2025;274:112748. https://doi.org/10.1016/j.buildenv.2025.112748.
[58] de Wilde P. Building performance simulation in the brave new world of artificial intelligence and digital twins: a systematic review. Energy Build 2023;292:113171. https://doi.org/10.1016/j.enbuild.2023.113171.
[59] Wang Y, Wang J, He W. Development of efficient, flexible and affordable heat pumps for supporting heat and power decarbonisation in the UK and beyond: review and perspectives. Renew Sustain Energy Rev 2022;154:111747. https://doi.org/10.1016/j.rser.2021.111747.
[60] Lazzarin R, Busato F, Noro M. Heat pumps in refurbishment of existing buildings. J Fed Eur Heat Vent Air Cond Assoc REHVA 2012;6:45–9.
[61] Walden JVM, Padullés R. An analytical solution to optimal heat pump integration. Energy Convers Manag 2024;320:118983. https://doi.org/10.1016/j.enconman.2024.118983.
[62] Bizzarri M, Conti P, Schito E, Testi D. Improving energy efficiency through forecast-driven control in hybrid heat pumps. Energy Convers Manag 2025;332:119737. https://doi.org/10.1016/j.enconman.2025.119737.
[63] Kazmi H, Fu C, Miller C. Ten questions concerning data-driven modelling and forecasting of operational energy demand at building and urban scale. Build Environ 2023;239:110407. https://doi.org/10.1016/j.buildenv.2023.110407.
[64] Mathieu JL, Price PN, Kiliccote S, Piette MA. Quantifying changes in building electricity use, with application to demand response. IEEE Trans Smart Grid 2011;2:507–18.
[65] Manfren M, James PAB, Aragon V, Tronchin L. Lean and interpretable digital twins for building energy monitoring – a case study with smart thermostatic radiator valves and gas absorption heat pumps. Energy and AI 2023;14:100304. https://doi.org/10.1016/j.egyai.2023.100304.
[66] Mirfin A, Xiao X, Jack MW. TOWST: a physics-informed statistical model for building energy consumption with solar gain. Appl Energy 2024;369:123488. https://doi.org/10.1016/j.apenergy.2024.123488.
[67] Lopez-Villamor I, Eguiarte O, Arregi B, Garay-Martinez R, Garrido-Marijuan A. Time of the week AutoRegressive eXogenous (TOW-ARX) model to predict thermal consumption in a large commercial mall. Energy Convers Manag X 2024;24:100777. https://doi.org/10.1016/j.ecmx.2024.100777.
[68] Ferrara M, Monetti V, Fabrizio E. Cost-optimal analysis for nearly zero energy buildings design and optimization: a critical review. Energies 2018;11:1478. https://doi.org/10.3390/en11061478.
[69] Chen M, Wei J, Yang X, Fu Q, Wang Q, Qiao S. Multi-objective optimization of multi-energy complementary systems integrated biomass-solar-wind energy utilization in rural areas. Energy Convers Manag 2025;323:119241. https://doi.org/10.1016/j.enconman.2024.119241.
[70] Gao J, Meng Q, Liu J, Yan Y, Wu H. Multi-energy cooperative optimal scheduling of rural virtual power plant considering flexible dual-response of supply and demand and wind-photovoltaic uncertainty. Energy Convers Manag 2024;320:118990. https://doi.org/10.1016/j.enconman.2024.118990.
[71] Ashrafi M, Abbaspour-Tehranifard A, Fotuhi-Firuzabad M, Fattaheian-Dehkordi S, Bacha S, Caire R. MILP-Based service restoration in prosumer-based active distribution networks. 2023 IEEE PES innovative smart grid technologies Europe (ISGT EUROPE). 2023. p. 1–6. https://doi.org/10.1109/ISGTEUROPE56780.2023.10408426.
[72] Dukovska I, Slootweg JG, Paterakis NG. Decentralized coordination of a community of electricity prosumers via distributed MILP. IEEE Trans Power Syst 2021;36:5578–89. https://doi.org/10.1109/TPWRS.2021.3073039.

[73] Manfren M. Multi-commodity network flow models for dynamic energy management – mathematical formulation. Energy Proc 2012;14:1380–5. https://doi.org/10.1016/j.egypro.2011.12.1105.

[74] Adhikari RS, Aste N, Manfren M. Multi-commodity network flow models for dynamic energy management – smart grid applications. Energy Proc 2012;14:1374–9. https://doi.org/10.1016/j.egypro.2011.12.1104.

[75] Shin S, Coffrin C, Sundar K, Zavala VM. Graph-based modeling and decomposition of energy infrastructures. IFAC-PapersOnLine 2021;54:693–8. https://doi.org/10.1016/j.ifacol.2021.08.322.

[76] Markensteijn AS, Romate JE, Vuik C. A graph-based model framework for steady-state load flow problems of general multi-carrier energy systems. Appl Energy 2020;280:115286. https://doi.org/10.1016/j.apenergy.2020.115286.

[77] Liu X, Mancarella P. Modelling, assessment and sankey diagrams of integrated electricity-heat-gas networks in multi-vector district energy systems. Appl Energy 2016;167:336–52. https://doi.org/10.1016/j.apenergy.2015.08.089.

[78] Marquant JF, Evins R, Bollinger LA, Carmeliet J. A holarchic approach for multi-scale distributed energy system optimisation. Appl Energy 2017;208:935–53. https://doi.org/10.1016/j.apenergy.2017.09.057.

[79] Reynolds J, Ahmad MW, Rezgui Y. Holistic modelling techniques for the operational optimisation of multi-vector energy systems. Energy Buildn.d. https://doi.org/10.1016/j.enbuild.2018.03.065..

[80] Bröchin M, Pickering B, Tröndle T, Pfenninger S. Harder, better, faster, stronger: understanding and improving the tractability of large energy system models. Energy Sustain Soc 2024;14:27. https://doi.org/10.1186/s13705-024-00458-z.

[81] Li J, Liu P, Li Z. Optimal design and techno-economic analysis of a solar-wind-biomass off-grid hybrid power system for remote rural electrification: a case study of west China. Energy (Calg) 2020;208:118387. https://doi.org/10.1016/j.energy.2020.118387.

[82] Ji L, Liang X, Xie Y, Huang G, Wang B. Optimal design and sensitivity analysis of the stand-alone hybrid energy system with PV and biomass-CHP for remote villages. Energy (Calg) 2021;225:120323. https://doi.org/10.1016/j.energy.2021.120323.

[83] Yuan H, Ye H, Chen Y, Deng W. Research on the optimal configuration of photovoltaic and energy storage in rural microgrid. Energy Rep 2022;8:1285–93. https://doi.org/10.1016/j.egyr.2022.08.115.

[84] Li J, Liu P, Li Z. Optimal design and techno-economic analysis of off-grid hybrid renewable energy system for remote rural electrification: a case study of southwest China. CET Journal-Chemical Engineering Transactions 2020;81.

[85] Li J, Liu P, Li Z. Optimal design and techno-economic analysis of a hybrid renewable energy system for off-grid power supply and hydrogen production: a case study of West China. Chem Eng Res Des 2022;177:604–14. https://doi.org/10.1016/j.cherd.2021.11.014.

[86] Li J, Liu P, Li Z. Optimal design of a hybrid renewable energy system with grid connection and comparison of techno-economic performances with an off-grid system: a case study of West China. Comput Chem Eng 2022;159:107657. https://doi.org/10.1016/j.compchemeng.2022.107657.

[87] Bahaj AS, James PAB. Urban energy generation: the added value of photovoltaics in social housing. Renew Sustain Energy Rev 2007;11:2121–36. https://doi.org/10.1016/j.rser.2006.03.007.

[88] Breyer C, Koskinen O, Blechinger P. Profitable climate change mitigation: the case of greenhouse gas emission reduction benefits enabled by solar photovoltaic systems. Renew Sustain Energy Rev 2015;49:610–28. https://doi.org/10.1016/j.rser.2015.04.061.

[89] Han Z, Han W, Song X, Lv L, Zhang N, Sui J. A new multi-objective optimization model for an integrated energy system based on life-cycle composite technical, economic and environmental indices. Energy Convers Manag 2025;327:119532. https://doi.org/10.1016/j.enconman.2025.119532.

[90] Wikoff HM, Reese SB, Reese MO. Embodied energy and carbon from the manufacture of cadmium telluride and silicon photovoltaics. Joule 2022;6:1710–25. https://doi.org/10.1016/j.joule.2022.06.006.

[91] Peng J, Lu L, Yang H. Review on life cycle assessment of energy payback and greenhouse gas emission of solar photovoltaic systems. Renew Sustain Energy Rev 2013;19:255–74. https://doi.org/10.1016/j.rser.2012.11.035.

[92] Manfren M, Gonzalez-Carreon KM, James PAB. Interpretable data-driven methods for building energy modelling—A review of critical connections and gaps. Energies 2024;17:881. https://doi.org/10.3390/en17040881.

[93] Liguori A, Quintana M, Fu C, Miller C, Frisch J, van Treeck C. Opening the black box: towards inherently interpretable energy data imputation models using building physics insight. Energy Build 2024;310:114071. https://doi.org/10.1016/j.enbuild.2024.114071.

[94] Ghenai C, Husein LA, Al Nahlawi M, Hamid AK, Bettayeb M. Recent trends of digital twin technologies in the energy sector: a comprehensive review. Sustain Energy Technol Assessments 2022;54:102837. https://doi.org/10.1016/j.seta.2022.102837.

[95] EnergyPlus weather data n.d.https://energyplus.net/weather (accessed December 16, 2024)..

[96] Crawley DB, Lawrie LK, Winkelmann FC, Buhl WF, Huang YJ, Pedersen CO, et al. EnergyPlus: creating a new-generation building energy simulation program. Energy Build 2001;33:319–31. https://doi.org/10.1016/S0378-7788(00)00114-6.

[97] Climate.OneBuilding.Org. Repository of building simulation climate data n.d. https://climate.onebuilding.org/default.html (accessed December 16, 2024)..

[98] CHN GB50176. Code for thermal design of civil building. 2016.

[99] Bolton D. The computation of equivalent potential temperature. Mon Weather Rev 1980;108:1046–53. https://doi.org/10.1175/1520-0493(1980)108<1046:TCOEPT>2.0.CO. 2.

[100] Salby ML. Fundamentals of atmospheric physics. 1996.

[101] May RM, Goebbert KH, Thielen JE, Leeman JR, Camron MD, Bruick Z, et al. MetPy: a meteorological python library for data analysis and visualization. Bull Am Meteorol Soc 2022;103:E2273–84. https://doi.org/10.1175/BAMS-D-21-0125.1.

[102] CHN JGJ 26. Energy efficiency design standards for residential buildings in severe cold and cold areas. 2018.

[103] CHN JGJ 134. Energy efficiency design standards for residential buildings in hot summer and cold winter areas. 2010.

[104] CHN JGJ 75. Energy efficiency design standard for residential buildings in hot summer and warm winter areas. 2012.

[105] Jensen U, Lütkebohmert C. Change-point models. Encyclopedia of statistics in quality and reliability. 2007. https://doi.org/10.1002/9780470061572.eqr076.

[106] Liao R, Manfren M, Nastasi B. Off-grid PV systems modelling and optimisation for rural communities - leveraging understandability and interpretability of modelling tools. Energy (Calg) 2025;324:135948. https://doi.org/10.1016/j.energy.2025.135948.

[107] Li M, Allinson D, Lomas K. Estimation of building heat transfer coefficients from in-use data. Int J Build Pathol Adapt 2020;38:38–50. https://doi.org/10.1108/IJBPA-02-2019-0022.

[108] Rasmussen C, Bacher P, Calì D, Nielsen HA, Madsen H. Method for scalable and automatised thermal building performance documentation and screening. Energies 2020;13. https://doi.org/10.3390/en13153866.

[109] GB 50015. Standard for design of building water supply and drainage. 2019.

[110] Ruhnau O, Hirth L, Praktiknjo A. Time series of heat demand and heat pump efficiency for energy system modeling. Sci Data 2019;6:189. https://doi.org/10.1038/s41597-019-0199-y.

[111] Fischer D, Wolf T, Wapler J, Hollinger R, Madani H. Model-based flexibility assessment of a residential heat pump pool. Energy (Calg) 2017;118:853–64. https://doi.org/10.1016/j.energy.2016.10.111.

[112] Fuentes E, Arce L, Salom J. A review of domestic hot water consumption profiles for application in systems and buildings energy performance analysis. Renew Sustain Energy Rev 2018;81:1530–47. https://doi.org/10.1016/j.rser.2017.05.229.

[113] GB 50555. Standard for water saving design in civil building. 2010.

[114] Liao R. Water scarcity assessment index from the realistic perspective of human basic water requirements. Environmental and Sustainability Indicators 2024;22:100404. https://doi.org/10.1016/j.indic.2024.100404.

[115] Perera ATD, Wickramasinghe PU, Nik VM, Scartezzini J-L. Machine learning methods to assist energy system optimization. Appl Energy 2019;243:191–205. https://doi.org/10.1016/j.apenergy.2019.03.202.

[116] Klucher TM. Evaluation of models to predict insolation on tilted surfaces. Sol Energy 1979;23:111–4. https://doi.org/10.1016/0038-092X(79)90110-5.

[117] Durisch W, Bitnar B, Mayor J-C, Kiess H, Lam K, Close J. Efficiency model for photovoltaic modules and demonstration of its application to energy yield estimation. Sol Energy Mater Sol Cell 2007;91:79–84. https://doi.org/10.1016/j.solmat.2006.05.011.

[118] Kasten F, Young AT. Revised optical air mass tables and approximation formula. Appl Opt 1989;28:4735–8. https://doi.org/10.1364/AO.28.004735.

[119] Anderson KS, Hansen CW, Holmgren WF, Jensen AR, Mikofski MA, Driesse A. Pvlib python: 2023 project update. J Open Source Softw 2023;8:5994. https://doi.org/10.21105/joss.05994.

[120] Šúri M, Huld TA, Dunlop ED. PV-GIS: a web-based solar radiation database for the calculation of PV potential in Europe. Int J Sustain Energy 2005;24:55–67. https://doi.org/10.1080/14786450512331329556.

[121] Van Veldhuizen David A, Lamont Gary B. Evolutionary computation and convergence to a pareto front. Late breaking papers at the genetic programming 1998 conference. 1998. p. 221–8.

[122] GB/T 50824. Design standard for energy efficiency of rural residential buildings. 2013.

[123] Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliab Eng Syst Saf 2003;81:23–69. https://doi.org/10.1016/S0951-8320(03)00058-9.

[124] Shi Y, Gao X, Xu Y, Giorgi F, Chen D. Effects of climate change on heating and cooling degree days and potential energy demand in the household sector of China. Clim Res 2016;67:135–49.

[125] Fthenakis V, Leccisi E. Updated sustainability status of crystalline silicon-based photovoltaic systems: life-cycle energy and environmental impact reduction trends. Prog Photovoltaics Res Appl 2021;29:1068–77. https://doi.org/10.1002/pip.3441.

[126] Bhandari KP, Collier JM, Ellingson RJ, Apul DS. Energy payback time (EPBT) and energy return on energy invested (EROI) of solar photovoltaic systems: a systematic review and meta-analysis. Renew Sustain Energy Rev 2015;47:133–41. https://doi.org/10.1016/j.rser.2015.02.057.

[127] Song A, Zhou Y. An integrative lifecycle design approach based on carbon intensity for renewable-battery-consumer energy systems. Commun Eng 2025;4:14. https://doi.org/10.1038/s44172-024-00339-5.

[128] Le Varlet T, Schmidt O, Gambhir A, Few S, Staffell I. Comparative life cycle assessment of lithium-ion battery chemistries for residential storage. J Energy Storage 2020;28:101230. https://doi.org/10.1016/j.est.2020.101230.