

Valuing urban building flexibility portfolios under carbon-constrained electricity dispatch: A Shenzhen case study

Kaixiang Zhang

Chinese University of Hong Kong, Shenzhen

Huanxin Liao

Chinese University of Hong Kong, Shenzhen

Yifan Zhu

Chinese University of Hong Kong, Shenzhen

Zhenshang Wang

Shenzhen Power Supply Co., Ltd.

Yuming Zhao

Shenzhen Power Supply Co., Ltd.

Jing Wang

Shenzhen Power Supply Co., Ltd.

Junyao Gao

Shenzhen Power Supply Co., Ltd.

Junhua Zhao

zhaojunhua@cuhk.edu.cn

Chinese University of Hong Kong, Shenzhen

Article

Keywords: Urban buildings, Building aggregation, Demand flexibility, Carbon-constrained dispatch, Emissions-budget dual, Shenzhen

Posted Date: August 25th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10523725/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Valuing urban building flexibility portfolios under carbon-constrained electricity dispatch: A Shenzhen case study

Kaixiang Zhang¹, Huanxin Liao¹, Yifan Zhu¹, Zhenshang Wang², Yuming Zhao², Jing Wang², Junyao Gao², and Junhua Zhao^{1,*}

¹The Chinese University of Hong Kong, Shenzhen, Shenzhen, China

²Shenzhen Power Supply Co., Ltd., Guangdong Provincial Key Laboratory of Source-Grid-Load-Storage Interactive Collaborative Technology (No. 2024B1212020004), Shenzhen, China

Kaixiang Zhang	222015008@link.cuhk.edu.cn
Huanxin Liao	huanxinliao@link.cuhk.edu.cn
Yifan Zhu	yifanzhu2@link.cuhk.edu.cn
Zhenshang Wang	wangzhenshang@sz.csg.cn
Yuming Zhao	zhaoyum97@sina.com
Jing Wang	wangjing1221@sz.csg.cn
Junyao Gao	gaojunyao@sz.csg.cn
Junhua Zhao	zhaojunhua@cuhk.edu.cn

*Correspondence: Junhua Zhao (zhaojunhua@cuhk.edu.cn)

Abstract

Urban building portfolios are often represented as aggregate demand, obscuring the different effects of temporal shifting and net energy reduction. We construct dispatch-oriented portfolios from hourly records for 143 screened Shenzhen buildings and evaluate them in a two-stage framework. A Guangdong DC optimal power flow supplies boundary prices and an offer-derived provincial average operational emissions proxy to a Shenzhen carbon-constrained dispatch model, which represents shifting and reduction with separate variables and costs. At the upper stress-test portfolio share of 0.60 and a 0.60% tighter emissions budget, energy-neutral shifting reduces modeled Shenzhen-subsystem operating cost by 0.389 million RMB (0.268%) while leaving the accounted-emissions-budget dual nearly unchanged. With a 2% daily limit and a 200 RMB/MWh reduction cost, 1,480 MWh of reduction is economic without the emissions constraint. Applying the tight budget raises reduction to 2,283 MWh and changes the budget dual from 242.5 to 88.4 RMB/tCO₂. Shifting therefore provides mainly temporal operating-cost value, whereas net reduction can additionally relieve a binding accounted-emissions budget. The paired cases support distinct product definitions for the two services.

Keywords: Urban buildings; Building aggregation; Demand flexibility; Carbon-constrained dispatch; Emissions-budget dual; Shenzhen

1 Introduction

Buildings account for roughly one-third of global final energy demand, and the buildings and construction sector produced 34% of energy-related CO₂ emissions in 2023 [1, 2]. Residential and commercial buildings are expected to contribute the largest absolute growth in global electricity demand between 2025 and 2030 [3]. In dense cities, this demand is both concentrated and heterogeneous: offices, retail buildings, hotels, public facilities, and mixed-use buildings differ in load magnitude, end-use composition, operating schedules, occupancy behavior, and flexibility shares. These loads could provide peak management, import scheduling, operating-cost reduction, and emissions-budget services if their flexibility is represented explicitly.

Demand response has long been studied as a smart-grid resource, with reviews emphasizing both its system value and the barriers to routine participation [4, 5]. Building efficiency and flexibility are also increasingly treated as grid resources rather than only as end-use conservation measures [6]. Demand-response models typically represent dispatchable load and operating costs, building-energy models resolve end uses and comfort, and carbon-accounting methods apply baselines and emission factors. These representations do not always distinguish an energy-neutral shift from a verified net reduction. The distinction matters because the first changes when electricity is consumed, whereas the second also changes how much electricity is consumed.

Load-pattern clustering supports customer segmentation, typical-curve extraction, anomaly detection, tariff design, and forecasting [7], while representative-period methods reduce temporal complexity in energy-system optimization [8–10]. For dispatch studies, statistical similarity alone is insufficient: the aggregated input must retain absolute scale, controllable end-use envelopes, adjustment limits, and exposure to time-varying price and emission signals. Clustering can organize heterogeneous profiles, but it does not create additional flexibility or establish its physical deliverability.

Building-flexibility research has advanced in HVAC control, thermal inertia, heat-pump aggregation, occupant comfort, and building-to-grid interaction [11–13]. Related work covers urban building energy modelling [14], coordinated building-cluster flexibility [15], local energy markets [16], emission-aware demand response [17], and electricity-carbon signal design [18]. A subsystem valuation must connect measured-building aggregation and explicit product definitions to operating-cost and emissions-budget outcomes.

We address this gap with a subsystem valuation framework that converts hourly building records into portfolios retaining total load and controllable AC load on their physical scale. A-type flexibility represents energy-neutral shifting, whereas B-type flexibility adds a separate, bounded net-reduction action. A provincial DC-OPF supplies fixed boundary signals to a city model that minimizes generation, import, flexibility, and load-shedding costs subject to an operational emissions budget.

Outputs are reported separately as subsystem operating-cost saving, accounted emissions, the emissions-budget dual, and reduction-availability values. Matched B-type cases with and without a binding emissions budget distinguish cost-driven reduction from the additional response induced by the budget. The three-archetype representation is benchmarked against a 143-building dispatch model. These comparisons quantify the operational consequences of treating temporal shifting and delivered energy reduction as separate demand-response products.

2 Methods

2.1 Overall portfolio-valuation framework

Fig. 1 summarizes the workflow from building records to subsystem valuation. Each portfolio retains total demand, the controllable AC envelope, adjustment limits, and action costs. Portfolio value is measured by changes in the Shenzhen-model objective and by the dual values of binding city constraints.

Panels (a) and (b) construct physical-scale typical-week portfolios. The provincial DC-OPF in panel (c) produces the two fixed boundary signals shown in panel (d): import price and the provincial average operational emissions proxy. These signals are held fixed while the Shenzhen model in panel (e) co-optimizes local supply, imports, and building flexibility, so city decisions do not feed back to provincial dispatch. Panel (f) defines the matched comparisons and reported outputs.

2.2 Building archetypes and scalable market portfolios

The dataset contains 148 annual hourly load files. Three files could not be matched to the building metadata and were excluded before feature construction. For the remaining 145 buildings, each file contains 8,760 hourly observations for lighting and plug loads, AC, motive power, special loads, and total demand. Let $p_{i,w,h,c}$ denote power in kW for building i , complete week w , hour h , and end-use category c , and let A_i be its floor area. The screening and clustering feature uses the area-normalized mean across 52 complete weeks:

$$\tilde{p}_{i,h,c} = \frac{1}{|\mathcal{W}_i|} \sum_{w \in \mathcal{W}_i} \frac{p_{i,w,h,c}}{A_i}, \quad h = 1, \dots, 168. \quad (1)$$

The remaining 24 annual hours are not used. The profile is vectorized across hours and the five end-use categories as

$$\mathbf{x}_i = \text{vec}(\tilde{p}_{i,h,c})_{h \in \mathcal{T}, c \in \mathcal{C}}, \quad (2)$$

and the feature matrix is standardized before clustering.

A k -means procedure is used to obtain representative building archetypes. With K archetypes, the assignment z_i and centroids $\mathbf{m}_k$ solve

$$\min_{z_i, \mathbf{m}_k} \sum_i \|\mathbf{x}_i - \mathbf{m}_{z_i}\|_2^2, \quad z_i \in \{1, \dots, K\}. \quad (3)$$

The silhouette score over $K = 2, \dots, 8$ is highest at $K = 5$, as shown in Supplementary Table S6. The $K = 5$ solution contains two singleton clusters: one with isolated extreme spikes and one with an all-zero AC channel. Removing these abnormal meter profiles leaves 143 buildings in three substantive groups of 96, 15, and 32 buildings. The diagnostic uses random seed 0 and 10 initializations; the final fit uses 20 initializations.

For the retained 143 buildings, the dispatch profile is recomputed as the median across the same 52 complete weeks:

$$\bar{p}_{i,h,c} = \text{median}_{w \in \mathcal{W}_i} \left(\frac{p_{i,w,h,c}}{A_i} \right). \quad (4)$$

This second step reduces the influence of remaining weekly spikes without changing the group assignments. The curves used for dispatch are reconstructed on the original physical scale rather than taken from standardized k -means centroids. With $\mathcal{S}_k = \{i : z_i = k\}$, let

$$A_k = \sum_{i \in \mathcal{S}_k} A_i, \quad \bar{q}_{k,t,c} = \frac{1}{A_k} \sum_{i \in \mathcal{S}_k} A_i \bar{p}_{i,t,c}, \quad (5)$$

where $\bar{q}_{k,t,c}$ is the area-weighted per-area profile of archetype k . The total-load and controllable-AC portfolio prototypes used in dispatch are therefore

$$\hat{L}_{k,t}^{\text{proto}} = 10^{-3} A_k \bar{q}_{k,t,\text{tot}}, \quad \hat{L}_{k,t}^{\text{ac,proto}} = 10^{-3} A_k \bar{q}_{k,t,\text{ac}}, \quad (6)$$

where 10^{-3} converts kW to MW. Fig. 2 displays the area-normalized profiles $\bar{q}_{k,t,c}$, whereas Eqs. (6)–(9) provide the physical-scale loads used in dispatch. The final input contains a 168-hour total-load profile and AC envelope for each archetype.

To examine system response at different portfolio scales, the archetype loads are multiplied by a scenario share ρ , defined as the share of weekly Shenzhen-subsystem electricity represented by the scaled building profiles. Let $\Delta t = 1$ h, let $E^{zone} = \sum_t D_t^{zone} \Delta t$, and let

$$E^{proto} = \sum_{t \in \mathcal{T}} \sum_{k=1}^K \hat{L}_{k,t}^{proto} \Delta t \quad (7)$$

be weekly archetype-sample energy. The scaling factor is

$$\beta(\rho) = \frac{\rho E^{zone}}{E^{proto}}. \quad (8)$$

The scaled subsystem-level baseline total load and AC load of archetype k are

$$L_{k,t}^0(\rho) = \beta(\rho) \hat{L}_{k,t}^{proto}, \quad L_{k,t}^{ac,0}(\rho) = \beta(\rho) \hat{L}_{k,t}^{ac,proto}. \quad (9)$$

To avoid double-counting subsystem electricity demand, the residual non-participating load is

$$D_t^{other}(\rho) = D_t^{zone} - \sum_{k=1}^K L_{k,t}^0(\rho). \quad (10)$$

The scenario share is limited so that $D_t^{other}(\rho)$ remains nonnegative at every hour:

$$\rho_{\max} = \min_t \frac{D_t^{zone} E^{proto}}{E^{zone} \sum_{k=1}^K \hat{L}_{k,t}^{proto}}. \quad (11)$$

In the case study, $\rho_{\max} = 0.603$ marks the residual-load non-negativity ceiling and defines the upper bound of the stress-test range.

2.3 A-type and B-type building flexibility

A-type flexibility represents energy-neutral shifting. Let $s_{k,t}^+ \geq 0$ and $s_{k,t}^- \geq 0$ denote upward and downward shifting power. B-type flexibility adds hourly net reduction $r_{k,t} \geq 0$. The net adjustment and realized load are

$$a_{k,t} = s_{k,t}^+ - s_{k,t}^- - r_{k,t}, \quad L_{k,t} = L_{k,t}^0 + a_{k,t} \geq 0. \quad (12)$$

Upward shifting is bounded by the AC envelope. Downward shifting and net reduction share the downward envelope so that the same controllable load cannot be used twice:

$$0 \leq s_{k,t}^+ \leq \alpha_k^{up} L_{k,t}^{ac,0}, \quad s_{k,t}^- + r_{k,t} \leq \alpha_k^{down} L_{k,t}^{ac,0}. \quad (13)$$

The realized adjustment is subject to a cyclic weekly ramp constraint,

$$-R_k \leq a_{k,t} - a_{k,t-} \leq R_k, \quad R_k = \kappa \max_t L_{k,t}^{ac,0}. \quad (14)$$

Here t^- is the preceding hour, with the first hour linked to the final hour of the 168-hour horizon. Shifted energy is conserved within each day:

$$\sum_{t \in \mathcal{T}_d} (s_{k,t}^+ - s_{k,t}^-) \Delta t = 0, \quad \forall k, d. \quad (15)$$

The amount moved between hours and its disutility cost are

$$E^{shift} = \frac{1}{2} \sum_{k,t} (s_{k,t}^+ + s_{k,t}^-) \Delta t, \quad C^{shift} = \frac{1}{2} \sum_{k,t} c_k^{shift} (s_{k,t}^+ + s_{k,t}^-) \Delta t. \quad (16)$$

The factor 1/2 counts one upward and one downward MWh as one MWh shifted. Thus c_k^{shift} is charged once per MWh transferred between hours. A-type operation fixes $r_{k,t} = 0$ and therefore changes timing without changing daily energy.

B-type operation allows net reduction in addition to energy-neutral shifting. Daily delivered reduction is

$$R_{k,d}^{red} = \sum_{t \in \mathcal{T}_d} r_{k,t} \Delta t, \quad (17)$$

and is limited as a fraction of daily AC baseline energy:

$$0 \leq R_{k,d}^{red} \leq \eta_k \sum_{t \in \mathcal{T}_d} L_{k,t}^{ac,0} \Delta t. \quad (18)$$

Its cost is separate from shifting cost:

$$C^{red} = \sum_k \sum_d c_k^{red} R_{k,d}^{red}. \quad (19)$$

A net reduction of 1 MWh is charged only through C^{red} at c_k^{red} . Equations (12)–(19) define aggregate market-product envelopes that enforce power, energy, and ramp limits. Their physical deliverability is assumed; indoor temperature and thermal dynamics are not modeled.

2.4 Provincial boundary-signal model

The upper layer extracts the external market conditions faced by Shenzhen. It is formulated as a provincial DC-OPF with nodal demand, dispatchable generation, fixed injections, line limits, and high-cost feasibility terms for data closure. These terms are penalized strongly and enter only when historical inputs are otherwise infeasible. The line-flow formulation follows the same base-MVA convention as the input network data; in the case study $S_{base} = 100$ MVA. The decision set is

$$\Omega^{prov} = \{P_{g,s,t}^{prov}, P_{g,t}^{prov}, \theta_{n,t}^{prov}, f_{\ell,t}^{prov}, \ell_{n,t}^{prov}, P_{n,t}^{bal}\}. \quad (20)$$

The compact upper-layer optimization problem is

$$\min_{\Omega^{prov}} \sum_{t,g,s} c_{g,s} P_{g,s,t}^{prov} \Delta t + VOLL \sum_{t,n} \ell_{n,t}^{prov} \Delta t + C^{bal} \sum_{t,n} P_{n,t}^{bal} \Delta t \quad (21a)$$

$$\text{s.t. } P_{g,t}^{prov} = \sum_s P_{g,s,t}^{prov}, \quad \forall g, t, \quad (21b)$$

$$0 \leq P_{g,s,t}^{prov} \leq \bar{P}_{g,s}^{prov}, \quad \forall g, s, t, \quad (21c)$$

$$f_{\ell,t}^{prov} = \frac{S_{base}}{x_{\ell}} \left(\theta_{i(\ell),t}^{prov} - \theta_{j(\ell),t}^{prov} \right), \quad \forall \ell, t, \quad (21d)$$

$$-\bar{F}_{\ell} \leq f_{\ell,t}^{prov} \leq \bar{F}_{\ell}, \quad \forall \ell, t, \quad (21e)$$

$$\begin{aligned} \sum_{g \in \mathcal{G}_n} P_{g,t}^{prov} + P_{n,t}^{fix,prov} + P_{n,t}^{bal} - (D_{n,t}^{prov} - \ell_{n,t}^{prov}) \\ = \sum_{m:(n,m) \in \mathcal{L}} f_{nm,t}^{prov} - \sum_{m:(m,n) \in \mathcal{L}} f_{mn,t}^{prov}, \end{aligned} \quad \forall n, t, \quad (21f)$$

$$0 \leq \ell_{n,t}^{prov} \leq D_{n,t}^{prov}, \quad \forall n, t, \quad (21g)$$

$$P_{n,t}^{bal} \geq 0, \quad \forall n, t, \quad (21h)$$

$$\theta_{n_0,t}^{prov} = 0, \quad \forall t. \quad (21i)$$

Equations (21b)–(21c) describe piecewise-linear dispatchable generation; Eqs. (21d)–(21e) describe the DC network; Eq. (21f) gives nodal balance; and Eqs. (21g)–(21i) close the feasibility and angle-reference conditions. The dual variables of the provincial nodal balances provide the nodal marginal prices used to build the Shenzhen import-price signal.

The first boundary signal is the import price,

$$c_t^{imp} = \frac{\sum_{n \in \mathcal{N}^{zone}} D_{n,t}^{zone} \lambda_{n,t}^{prov}}{\sum_{n \in \mathcal{N}^{zone}} D_{n,t}^{zone}}, \quad (22)$$

where $D_{n,t}^{zone}$ is Shenzhen-subsystem-related load at provincial node n . The second signal applied to imported electricity is

$$e_t^{imp} = \frac{\sum_g \gamma_g P_{g,t}^{prov} + E_t^{fix,prov} + \gamma^{bal} \sum_n P_{n,t}^{bal}}{\sum_n D_{n,t}^{prov} - \sum_n \ell_{n,t}^{prov}}. \quad (23)$$

Equation (23) computes an offer-derived provincial average operational emissions proxy per unit of served electricity. It is applied to Shenzhen imports as an accounting factor, not as a flow-tracing or marginal-emissions measure. Summary statistics for the two boundary signals are reported in Supplementary Table S8.

2.5 City-level carbon-constrained dispatch and marginal values

The city subsystem is a single-zone equivalent with local dispatchable generation, fixed injections, external imports, residual demand, and flexible building portfolios. No Type-A fixed injection is mapped to the Shenzhen subsystem in this case, so $P_t^{fix} = 0$ and $E_t^{fix} = 0$;

fixed-injection emissions remain part of the provincial proxy in Eq. (23). Given the fixed boundary signals, the decision set is

$$\Omega^{city} = \{P_{g,s,t}^{loc}, P_{g,t}^{loc}, P_t^{imp}, \ell_t, L_{k,t}, s_{k,t}^+, s_{k,t}^-, r_{k,t}, R_{k,d}^{red}, E^{var}, E^{tot}\}. \quad (24)$$

The dispatch problem is

$$\begin{aligned} \min_{\Omega^{city}} \quad & \sum_{t,g,s} c_{g,s} P_{g,s,t}^{loc} \Delta t + \sum_t c_t^{imp} P_t^{imp} \Delta t \\ & + \frac{1}{2} \sum_{k,t} c_k^{shift} (s_{k,t}^+ + s_{k,t}^-) \Delta t + \sum_{k,d} c_k^{red} R_{k,d}^{red} + VOLL \sum_t \ell_t \Delta t \end{aligned} \quad (25a)$$

$$\text{s.t.} \quad P_{g,t}^{loc} = \sum_s P_{g,s,t}^{loc}, \quad \forall g, t, \quad (25b)$$

$$0 \leq P_{g,s,t}^{loc} \leq \bar{P}_{g,s}^{loc}, \quad \forall g, s, t, \quad (25c)$$

$$\begin{aligned} \sum_g P_{g,t}^{loc} + P_t^{imp} + P_t^{fix} \\ = D_t^{other} + \sum_k L_{k,t} - \ell_t, \end{aligned} \quad \forall t, \quad (25d)$$

$$0 \leq P_t^{imp} \leq \bar{P}_t^{imp}, \quad \forall t, \quad (25e)$$

$$0 \leq \ell_t \leq D_t^{other} + \sum_k L_{k,t}, \quad \forall t, \quad (25f)$$

$$L_{k,t} = L_{k,t}^0 + s_{k,t}^+ - s_{k,t}^- - r_{k,t} \geq 0, \quad \forall k, t, \quad (25g)$$

$$0 \leq s_{k,t}^+ \leq \alpha_k^{up} L_{k,t}^{ac,0}, \quad \forall k, t, \quad (25h)$$

$$s_{k,t}^- \geq 0, \quad r_{k,t} \geq 0, \quad s_{k,t}^- + r_{k,t} \leq \alpha_k^{down} L_{k,t}^{ac,0}, \quad \forall k, t, \quad (25i)$$

$$-R_k \leq a_{k,t} - a_{k,t-} \leq R_k, \quad \forall k, t, \quad (25j)$$

$$\sum_{t \in \mathcal{T}_d} (s_{k,t}^+ - s_{k,t}^-) \Delta t = 0, \quad \forall k, d, \quad (25k)$$

$$R_{k,d}^{red} = \sum_{t \in \mathcal{T}_d} r_{k,t} \Delta t, \quad \forall k, d, \quad (25l)$$

$$0 \leq R_{k,d}^{red} \leq \eta_k \sum_{t \in \mathcal{T}_d} L_{k,t}^{ac,0} \Delta t : \nu_{k,d}^{red}, \quad \forall k, d, \quad (25m)$$

$$E^{var} = \sum_t \left(\sum_g \gamma_g P_{g,t}^{loc} + e_t^{imp} P_t^{imp} \right) \Delta t, \quad (25n)$$

$$E^{tot} = E^{fix} + E^{var}, \quad (25o)$$

$$E^{var} \leq cap^{tot}(\delta) - E^{fix} : \mu^{cap}. \quad (25p)$$

For A-type cases, $r_{k,t} = R_{k,d}^{red} = 0$; B-type cases use the stated value of η_k . The objective measures modeled Shenzhen-subsystem operating cost, combining local generation cost, import procurement cost, flexibility disutility, and any load-shedding penalty. Participant settlement and carbon-allowance transactions are outside the model.

The total cap is defined relative to the Base-case weekly emissions E_0 as

$$cap^{tot}(\delta) = (1 - \delta)E_0. \quad (26)$$

Let $E_0^{var} = E_0 - E^{fix}$ denote the baseline controllable-emission component. The effective tightening imposed on this component is

$$\delta^{var}(\delta) = 1 - \frac{cap^{tot}(\delta) - E^{fix}}{E_0^{var}}. \quad (27)$$

The accounted-emissions-budget dual is the marginal modeled subsystem operating cost of tightening Eq. (25p) by one additional ton of CO₂:

$$\pi_{CO_2} = -\mu^{cap}, \quad (28)$$

where the minus sign follows the solver convention. Similarly, when Eq. (25m) binds, $v_{k,d}^{red} = -\nu_{k,d}^{red}$ is the net decrease in optimized subsystem cost from one additional MWh of available daily reduction for that portfolio and day. The results report subsystem operating-cost saving, accounted emissions, π_{CO_2} , and $v_{k,d}^{red}$ separately.

Ethics statement

This study used pre-existing facility-level building energy-monitoring records and power-system operational data. No human participants were recruited, and no individual-level or identifiable information about building occupants was collected or analyzed. Ethics approval and informed consent were therefore not applicable. Data access and processing complied with the applicable confidentiality and third-party data-use agreements.

3 Case study and scenario design

3.1 Guangdong-Shenzhen test system and Shenzhen building data

The case study uses a Guangdong provincial power-system dataset to derive boundary signals for a Shenzhen urban subsystem that incorporates the 143-building sample.

The city subsystem is identified from Shenzhen-related buses and loads in the provincial data and represented as a single-zone equivalent. The resulting three archetypes are summarized in Supplementary Table S1. The 168-hour horizon provides a common basis for controlled comparisons across product definitions.

3.2 Accounting assumptions and scenario settings

Table 1 summarizes the implementation. The load-shedding penalty is 10,000 RMB/MWh and the flexibility ramp coefficient is $\kappa = 0.15$. Scenario shares are $\rho \in \{0.2, 0.4, 0.6\}$; the largest value approaches the residual-load ceiling and serves as the upper stress test.

Each flexibility case is compared with rigid demand under the same emissions budget. Base emissions, absolute cap levels, cap slack, solver status, and matched-reference outputs are retained as verification records and will be provided to editors and reviewers upon request.

The main Carbon-only scan uses 0.05%–0.60% tightening, with 0.10%, 0.20%, and 0.60% selected for flexibility comparisons. A diagnostic extension shows no load shedding through 1.00% tightening; at 1.20%, 587.8 MWh of shedding appears. The main range is therefore inside the no-shed region while spanning the rise in the budget dual from 24.6 to 242.5 RMB/tCO₂. B-type cases are run both without an emissions constraint and at the 0.60% budget to separate ordinary economic reduction from carbon-induced response. The import-ceiling binding-hour diagnostic for the main comparison cases is reported in Supplementary Table S7.

3.3 Validation checks

Validation covers four points: boundary-price signs, no-shed feasibility, local perturbation of the emissions-budget dual, and aggregation resolution. The latter compares $K = 1$, $K = 3$, and 143 separately modelled buildings under identical product parameters. Data and AC-envelope consistency checks are reported in Section 4.5.

3.4 Use of artificial-intelligence-assisted tools

OpenAI’s ChatGPT and Codex assisted with English-language editing and consistency checks. They were not used to generate data or conduct analyses. The authors checked the resulting edits and take responsibility for the manuscript’s content.

3.5 Code availability

The custom code used for data processing, clustering, dispatch optimization, and result generation is available from the corresponding author upon reasonable request for non-commercial verification.

4 Results

4.1 Building profiles and aggregation resolution

The three groups have different weekly schedules and AC shares, as summarized in Supplementary Table S1 and illustrated in Fig. 2. Cluster 0 contains 96 buildings and 76.95% of the used area; clusters 1 and 2 contain 15 and 32 buildings. They preserve the observed timing patterns and differences in controllable AC load.

Common flexibility bounds and costs make the three-cluster representation descriptive. The $K = 1$, $K = 3$, and 143-building comparison in Section 4.5 quantifies the aggregation error directly.

4.2 Rigid-demand response to the emissions budget

In the Base case (Table 2), weekly modeled subsystem operating cost is 144.652 million RMB, import energy is 723.52 GWh, and accounted emissions are 546.36 ktCO₂. At 0.60% tightening, cost increases by 0.258% and the budget dual is 242.51 RMB/tCO₂. All rows remain load-shedding-free.

The response is dispatch substitution (Fig. 3). Between 0.10% and 0.60% tightening, local dispatch decreases from 30.67 to 15.90 GWh/week and imports increase from 726.59 to 741.36 GWh/week. Under the adopted proxy, the model satisfies the accounted-emissions budget by replacing local dispatch with imports even though import-accounted emissions rise. This is an accounting response to the proxy, not evidence of a province-wide physical emissions change. The nonlinear dual curve reflects the piecewise-linear supply frontier.

4.3 A-type shifting: subsystem operating-cost effects

Fig. 4 compares A-type dispatch with rigid demand under the same emissions budget. At $\rho = 0.20$, weekly subsystem operating-cost savings are 0.130–0.132 million RMB across the selected budgets. At $\rho = 0.60$, they are 0.389–0.396 million RMB. The tight-budget saving is 0.268% of the rigid-demand subsystem cost.

Across the reported cases, the budget-dual change remains within 1.63 RMB/tCO₂. Shifting can change accounted emissions by moving demand across the hourly proxy, but it does not change daily electricity use. Under the study-week proxy, A-type shifting therefore primarily reduces cost. Its emissions effect may differ under a temporal marginal-emissions signal.

At a portfolio share of 0.60 and an emissions-budget tightening of 0.60%, broader bounds and lower shifting costs increase savings, as shown in Supplementary Fig. S1. With downward/upward bounds of 0.2/0.1, savings decrease from 0.443 to 0.376 million RMB as the shifting cost rises from 40 to 120 RMB/MWh moved. With bounds of 0.4/0.2, savings similarly decrease from 0.701 to 0.589 million RMB. The budget dual remains about 242.8 RMB per tonne across these settings.

4.4 B-type flexibility: economic and cap-induced reduction

B-type flexibility can respond to electricity cost even without an emissions constraint (Fig. 5). At a portfolio share of 0.60, a 2% daily reduction limit, and a 200 RMB/MWh reduction cost, the no-cap case delivers 1,480 MWh of net reduction and saves 0.830 million RMB (0.574%) relative to rigid Base demand. With the budget tightened by 0.60%, reduction increases to 2,283 MWh. The 803 MWh difference between the tight-cap and no-cap cases is the carbon-constraint-induced component of B-type response.

Table 3 separates subsystem operating-cost saving, the budget dual, dispatched reduction, and the dual value of reduction availability. For the tight budget and $c^{red} = 200$ RMB/MWh,

B-type flexibility saves 1.102 million RMB (0.760%) relative to rigid demand and changes the budget dual from 242.51 to 88.41 RMB/tCO₂.

The final column reports the mean and maximum $v_{k,d}^{red}$ among binding daily availability constraints. At $c^{red} = 200$ RMB/MWh under the tight budget, the mean is 128.3 RMB/MWh and the maximum is 459.8 RMB/MWh. These duals measure the net subsystem benefit of additional reduction availability after accounting for the modeled reduction cost. At $c^{red} = 500$ RMB/MWh, only 368 MWh is dispatched and the budget dual remains 237.52 RMB/tCO₂.

Provincial signals and all city cases are recomputed for wider coal-like and gas-like factor ranges. The resulting absolute accounted emissions and budget duals change materially, as shown in Supplementary Table S2. Lower emission factors yield lower accounted emissions but a higher RMB-per-ton dual under the same relative tightening. Under the uniform 80 RMB/MWh shifting cost and A-type downward/upward bounds of 0.2/0.1, A-type saving remains about 0.410 million RMB. B-type saving is 1.102–1.112 million RMB and its budget dual is 80.09–113.54 RMB per tonne. Fuel-classification uncertainty remains because offer prices serve as a proxy for generator type.

We compare one aggregate portfolio, the selected three groups, and 143 separately modelled buildings using a uniform 80 RMB/MWh shifting cost, A-type downward/upward bounds of 0.2/0.1, and B-type bounds of 0.3/0.2 with a 2% daily reduction limit and a 200 RMB/MWh reduction cost. Relative to the individual-building model, the $K = 3$ total objective differs by 0.009% for A-type and 0.011% for B-type operation, as shown in Supplementary Table S3. The matched savings are more sensitive to aggregation, differing by 3.28% and 1.41%, respectively. A-type shifted energy differs by 2.73%, and B-type reduction by 0.43%. Both $K = 1$ and $K = 3$ give similar results because all buildings use homogeneous linear product parameters. Group-specific parameters are needed for clustering to affect valuation.

The two products affect different constraints. In this case, A-type operation changes subsystem cost but has little effect on the budget dual, whereas B-type operation also lowers weekly energy and can relieve the binding accounted-emissions constraint.

4.5 Validation and scope of inference

An independently reconstructed aggregate with an additional local-median spike filter deviates from the dispatch portfolio, as illustrated in Supplementary Fig. S2. Because the unfiltered group profiles sum exactly to their member-building profiles, this comparison isolates sensitivity to the filter. Total-load and AC-envelope MAPE are 3.81% and 16.07%, respectively, as reported in Supplementary Table S4. The AC result combines level and temporal-shape changes; the 5% lower and 5% higher cases in Supplementary Table S3 instead test uniform AC-envelope scaling.

The provincial benchmark uses 228.17 MWh of emergency balancing energy, less than 0.002% of provincial weekly demand. Rigid city cases avoid load shedding through 1.00% budget tightening. At the reported 0.60% point, cap slack is below 5×10^{-9} tCO₂ and all solver terminations are optimal. The hourly import ceiling binds in 109 of 168 Base hours, 137 hours in the rigid $\delta = 0.60\%$ case, and 138 hours in the corresponding A-type case.

Perturbing the 0.60% cap by one tonne in each direction gives a central finite-difference slope of 242.508866 RMB per tonne, equal to the reported dual to six decimal places. The unchanged dual at both perturbation points indicates a stable local LP basis, as shown in Supplementary Table S5.

The main uncertainties concern the system boundary and flexibility deliverability. Offer-derived generator classes and provincial-average emission accounting determine the carbon signal, while fixed boundary prices omit city-to-province feedback. Because the import ceiling binds frequently, absolute local/import dispatch shares depend on this equivalent-system limit. The product envelopes omit thermal states, and the single typical week with scaled portfolio shares does not establish annual or citywide realizable potential.

5 Conclusions

We constructed urban building portfolios from hourly total-load and AC channels and evaluated them under carbon-constrained electricity dispatch. Provincial DC-OPF prices and an offer-derived provincial average operational emissions proxy enter the city model as fixed, one-way boundary signals. The model represents energy-neutral shifting and net reduction as separate products.

The Base case accounts for 546.36 ktCO₂ during the study week. Tightening the emissions budget from 0.05% to 0.60% raises the budget dual from 24.6 to 242.5 RMB/tCO₂. At $\rho = 0.60$, baseline A-type shifting saves 0.389 million RMB (0.268%) under $\delta = 0.60\%$ and leaves the dual nearly unchanged. B-type reduction at 200 RMB/MWh is partly economic: 1,480 MWh is dispatched without an emissions constraint. The tight budget raises reduction to 2,283 MWh, saves 1.102 million RMB (0.760%) relative to rigid demand under the same budget, and changes the budget dual to 88.4 RMB/tCO₂.

The $K = 3$ total objective differs by at most 0.011% from the 143-building model, while matched subsystem operating-cost savings differ by 3.28% for A-type and 1.41% for

B-type operation. The similar $K = 1$ result shows that clustering is descriptive when product parameters are homogeneous. Estimating group-specific flexibility, comfort, and cost parameters is necessary before attributing value differences to archetypes.

Product definition determines the observed source of subsystem value. Energy-neutral shifting primarily provides temporal dispatch value, whereas net reduction can also relieve a binding emissions budget. Building-flexibility valuation and product design should preserve this distinction.

Data Availability

The original building-level Shenzhen energy-monitoring records and provincial grid, generator, and load records are not publicly available because redistribution is prohibited by project-specific confidentiality and third-party data-use agreements. The de-identified 168-hour portfolio load profiles, derived boundary signals, scenario inputs and outputs, data-consistency summaries, and numerical values underlying the figures and tables are available from the corresponding author upon reasonable request for non-commercial verification.

References

- [1] United Nations Environment Programme & Global Alliance for Buildings and Construction. *Global Status Report for Buildings and Construction: Beyond Foundations—Mainstreaming Sustainable Solutions to Cut Emissions from the Buildings Sector* (UNEP, 2024). <https://www.unep.org/resources/report/global-status-report-buildings-and-construction-beyond-foundations-mainstreaming> (accessed 29 June 2026).
- [2] International Energy Agency. *Energy Efficiency 2025: Buildings* (IEA, 2025). <https://www.iea.org/reports/energy-efficiency-2025> (accessed 29 June 2026).
- [3] International Energy Agency. *Electricity 2026: Demand* (IEA, 2026). <https://www.iea.org/reports/electricity-2026> (accessed 29 June 2026).
- [4] Siano, P. Demand response and smart grids—A survey. *Renew. Sustain. Energy Rev.* **30**, 461–478 (2014). <https://doi.org/10.1016/j.rser.2013.10.022>.
- [5] Good, N., Ellis, K. A. & Mancarella, P. Review and classification of barriers and enablers of demand response in the smart grid. *Renew. Sustain. Energy Rev.* **72**, 57–72 (2017). <https://doi.org/10.1016/j.rser.2017.01.043>.
- [6] Langevin, J. *et al.* U.S. building energy efficiency and flexibility as an electric grid resource. *Joule* **5**, 2102–2128 (2021). <https://doi.org/10.1016/j.joule.2021.06.002>.

- [7] Chicco, G. Overview and performance assessment of the clustering methods for electrical load pattern grouping. *Energy* **42**, 68–80 (2012). <https://doi.org/10.1016/j.energy.2011.12.031>.
- [8] Nahmmacher, P., Schmid, E., Hirth, L. & Knopf, B. Carpe diem: A novel approach to select representative days for long-term power system modeling. *Energy* **112**, 430–442 (2016). <https://doi.org/10.1016/j.energy.2016.06.081>.
- [9] Kotzur, L., Markewitz, P., Robinius, M. & Stolten, D. Time series aggregation for energy system design: Modeling seasonal storage. *Appl. Energy* **213**, 123–135 (2018). <https://doi.org/10.1016/j.apenergy.2018.01.023>.
- [10] Teichgraeber, H. & Brandt, A. R. Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison. *Appl. Energy* **239**, 1283–1293 (2019). <https://doi.org/10.1016/j.apenergy.2019.01.169>.
- [11] Chen, Y., Xu, P., Gu, J., Schmidt, F. & Li, W. Measures to improve energy demand flexibility in buildings for demand response: A review. *Energy Build.* **177**, 125–139 (2018). <https://doi.org/10.1016/j.enbuild.2018.08.003>.
- [12] Zhang, L., Good, N. & Mancarella, P. Building-to-grid flexibility: Modelling and assessment metrics for residential demand response from heat pump aggregations. *Appl. Energy* **233–234**, 709–723 (2019). <https://doi.org/10.1016/j.apenergy.2018.10.103>.
- [13] Li, R. *et al.* Ten questions concerning energy flexibility in buildings. *Build. Environ.* **223**, 109461 (2022). <https://doi.org/10.1016/j.buildenv.2022.109461>.
- [14] Zhou, J. *et al.* State-of-the-art review of urban building energy modelling on supporting sustainable development goals. *Appl. Energy* **402**, 126924 (2025). <https://doi.org/10.1016/j.apenergy.2025.126924>.
- [15] Dong, Z., Zhang, X., Zhang, L., Giannelos, S. & Strbac, G. Flexibility enhancement of urban energy systems through coordinated space heating aggregation of numerous buildings. *Appl. Energy* **374**, 123971 (2024). <https://doi.org/10.1016/j.apenergy.2024.123971>.
- [16] You, Z., Lump, S. D., Doepfert, M., Tzscheutschler, P. & Goebel, C. Leveraging flexibility of residential heat pumps through local energy markets. *Appl. Energy* **355**, 122269 (2024). <https://doi.org/10.1016/j.apenergy.2023.122269>.
- [17] Nowak, C. & Bertsch, V. Emission-based demand response in energy system optimisations—A systematic literature review. *Appl. Energy* **401**, 126635 (2025). <https://doi.org/10.1016/j.apenergy.2025.126635>.

- [18] Liu, G. *et al.* Enhancing demand response in coupled electricity and carbon markets: Optimising market signals and dispatch configurations. *Energy* **338**, 138805 (2025). <https://doi.org/10.1016/j.energy.2025.138805>.

Acknowledgements

This work was supported by the project “Key Technologies, Market Mechanisms, and Demonstration Applications of Electricity-Carbon Coordination for Complex Building Clusters—Sub-project 4: Recommendations for Adjusting Market Mechanisms for Urban Energy-Supply Building Clusters” under Grant 090000KC23020082. The schematics and plots were prepared by the authors using Matplotlib 3.10.0 (<https://matplotlib.org/>).

Author contributions

Kaixiang Zhang: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Visualization, Writing—original draft, Writing—review and editing. Huanxin Liao: Methodology, Validation, Formal analysis, Writing—review and editing. Yifan Zhu: Validation, Visualization, Writing—review and editing. Zhenshang Wang: Resources, Data curation, Investigation, Writing—review and editing. Yuming Zhao: Resources, Data curation, Validation, Writing—review and editing. Jing Wang: Resources, Investigation, Validation, Writing—review and editing. Junyao Gao: Resources, Data curation, Project administration, Writing—review and editing. Junhua Zhao: Conceptualization, Methodology, Supervision, Project administration, Writing—review and editing. All authors reviewed and approved the manuscript.

Additional information

Competing interests

The authors declare no competing interests.

Figure legends

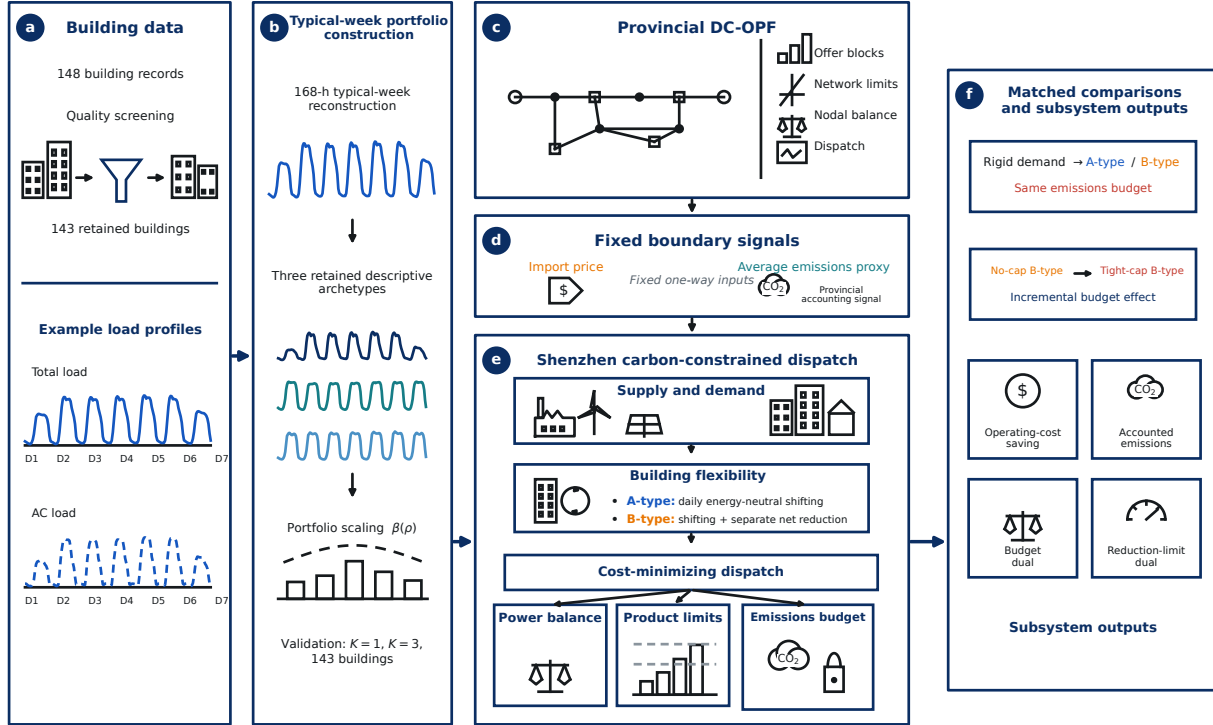


Figure 1: Workflow for constructing and valuing the building-flexibility portfolio. (a) Screening of 148 building records to retain 143 buildings. (b) Reconstruction of 168-hour typical-week profiles, descriptive archetypes, and portfolio scaling. (c) Provincial DC optimal power flow. (d) Extraction of fixed boundary signals for import price and the provincial average operational emissions proxy. (e) Shenzhen carbon-constrained dispatch with the A-type and B-type flexibility products. (f) Matched comparisons and the subsystem outputs reported in the analysis.

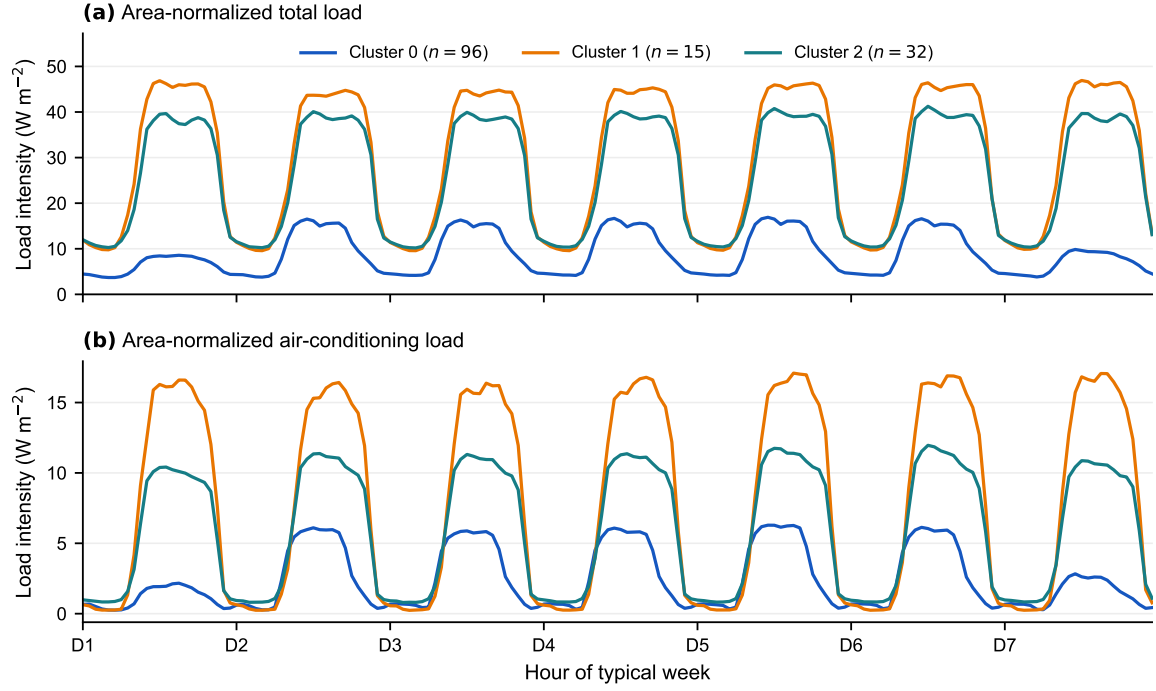


Figure 2: Area-normalized typical-week total-load and air-conditioning-load profiles of the three building archetypes. Cluster sizes are 96, 15, and 32 buildings, respectively; D1–D7 denote the seven sequential days in each 168-hour record.

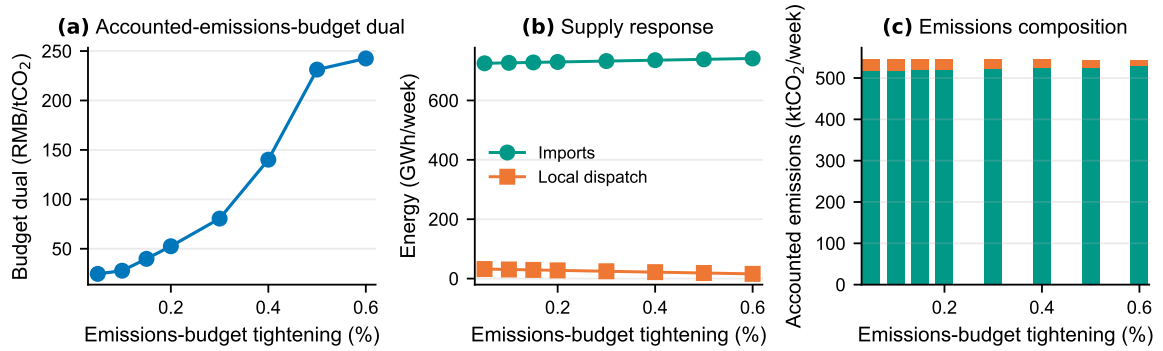


Figure 3: Rigid-demand results under emissions-budget tightening. (a) Accounted-emissions-budget dual. (b) Weekly import and local-dispatch energy. (c) Accounted import and local-dispatch emissions under the provincial average operational emissions proxy.

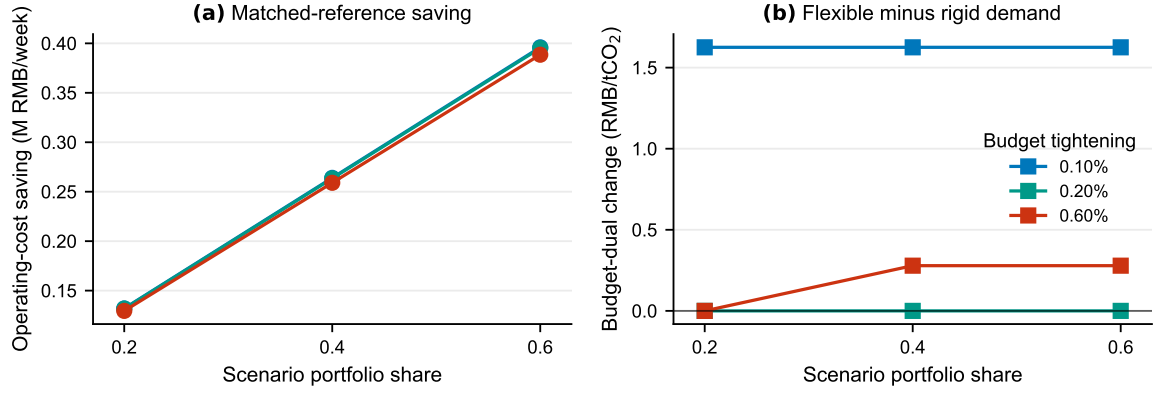


Figure 4: Baseline A-type results at selected emissions budgets. (a) Subsystem operating-cost saving relative to rigid demand under the same budget. (b) Signed budget-dual change, $\Delta\pi_{CO_2} = \pi_A - \pi_{rigid}$. The horizontal axis shows the scenario portfolio share.

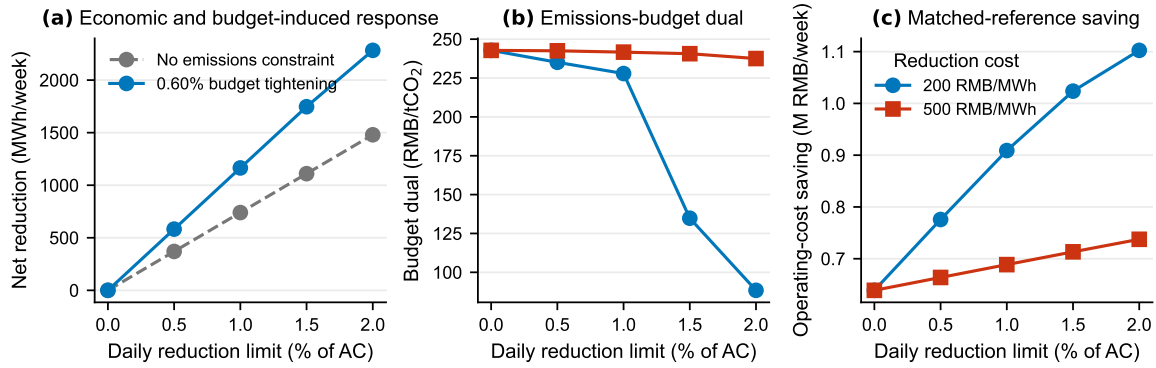


Figure 5: B-type sensitivity at $\rho = 0.60$. (a) Dispatched reduction with and without the emissions constraint at $c^{red} = 200$ RMB/MWh. (b) Accounted-emissions-budget dual. (c) Matched-reference subsystem operating-cost saving under $\delta = 0.60\%$. The daily limit is a share of baseline air-conditioning energy.

Tables

Table 1: Key data-processing, implementation, and carbon-accounting assumptions

Item	Implementation used in the case study
City identification	Shenzhen-related buses and loads identified from provincial records; 17 Shenzhen-related buses retained for boundary-price aggregation.
Building-data checks	Of 148 annual files, three lack a metadata match and two singleton profiles show abnormal metering patterns; 143 are retained. Retained files have no missing, non-finite, or negative values; AC is never above total load, and raw end-use channels sum to total load.
Clustering and typical weeks	Clustering uses area-normalized mean profiles over 52 complete weeks. The final dispatch profiles use the weekly median after screening. Independent channel medians produce a -0.40% aggregate end-use energy mismatch, so the measured total-load channel is used for balance and AC defines the flexibility envelope.
Import-emissions proxy	Equation (23) divides offer-derived province-wide operational emissions by served electricity and applies the result to Shenzhen imports.
Generator emission proxy	Fuel metadata are unavailable. Units with average offer prices below 50, from 50 to 300, and at or above 300 RMB/MWh are assigned zero, coal-like, and gas-like factors, respectively. Base factors are 0.90 and 0.40 tCO ₂ /MWh; low/high scenarios use 0.70/0.35 and 1.00/0.55.
Fixed injections	Provincial non-dispatchable injections are matched to generator names where possible; unmatched units use the matched-unit fallback mean of 0.420 tCO ₂ /MWh. No Type-A fixed injection maps to Shenzhen, so $P_t^{fix} = E^{fix} = 0$ in the city case.
Offer-block cleaning	Offer blocks exceeding installed capacity are scaled to capacity; missing residual capacity is filled with a fallback block at the last nonzero or average offer price.
Emergency balancing	Upper-layer data closure uses $C^{bal} = 9,000$ RMB/MWh and $\gamma^{bal} = 0.40$ tCO ₂ /MWh. Dispatched energy is 228.17 MWh over the study week and enters e_t^{imp} .
DC-flow convention	Provincial line flows use the S_{base}/x convention with $S_{base} = 100$ MVA, consistent with the network-data base.
Import-cap mode	The hourly limit is $\bar{F}_t^{imp} = D_t^{zone}$. It binds in 109 Base hours, 137 rigid tight-cap hours, and 138 A-type tight-cap hours; absolute dispatch shares are therefore import-limit-dependent.
Baseline A-type bounds	Main sweep uses $(\alpha^{down}, \alpha^{up}) = (0.2, 0.1)$ for all archetypes.
Baseline A-type cost	Cluster shifting costs are 100, 120, and 110 RMB per MWh moved between hours; shifted energy is $\frac{1}{2} \sum_{k,t} (s_{k,t}^+ + s_{k,t}^-) \Delta t$.
B-type sensitivity	Uses $(\alpha^{down}, \alpha^{up}) = (0.3, 0.2)$, shift cost of 80 RMB/MWh, $\eta \in \{0, 0.005, 0.01, 0.015, 0.02\}$, and reduction costs of 200 or 500 RMB/MWh.
Dual-sign convention	Imported LMP and budget-dual signs are normalized so positive values represent positive import cost and positive budget-tightening cost.
Model form and duals	The Pyomo linear programs are solved with HiGHS using primal and dual feasibility tolerances of 10^{-7} ; no unit-commitment binary variables are included.
Cap accounting	$E_0 = 546.364$ ktCO ₂ , $E^{fix} = 0$, and $E_0^{var} = 546.364$ ktCO ₂ in this case, so $\delta^{var} = \delta$.
Scenario-share ceiling	ρ_{max} is the hourly residual-load non-negativity ceiling used to define the upper stress-test range.

Table 2: Base and Carbon-only results for the Shenzhen subsystem under weekly emissions-budget tightening

Scenario	δ (%)	Objective (M RMB)	Cost change (%)	Budget dual (RMB/tCO ₂)	Import energy (GWh)	Emissions (ktCO ₂)
Base	0.00	144.652	0.000	—	723.52	546.36
Carbon-only	0.05	144.655	0.002	24.58	725.05	546.09
Carbon-only	0.10	144.662	0.007	27.79	726.59	545.82
Carbon-only	0.15	144.672	0.014	39.79	728.05	545.54
Carbon-only	0.20	144.683	0.021	52.53	729.55	545.27
Carbon-only	0.30	144.718	0.046	80.50	732.56	544.73
Carbon-only	0.40	144.779	0.088	140.12	735.52	544.18
Carbon-only	0.50	144.896	0.168	231.24	738.41	543.63
Carbon-only	0.60	145.025	0.258	242.51	741.36	543.09

Table 3: Matched-reference flexibility results at $\rho = 0.60$

Constraint	Product setting	Saving (M RMB/week)	Saving (%)	Budget dual (RMB/tCO ₂)	Net reduction (MWh/week)	Mean/max v^{red} (RMB/MWh)
None	B: $\eta = 2\%$, $c^{red} = 200$	0.830	0.574	—	1480	114.7/396.4
0.60%	A: baseline bounds/costs	0.389	0.268	242.79	0	—
0.60%	B: $\eta = 0$, shift cost 80	0.639	0.441	242.79	0	—
0.60%	B: $\eta = 2\%$, $c^{red} = 200$	1.102	0.760	88.41	2283	128.3/459.8
0.60%	B: $\eta = 2\%$, $c^{red} = 500$	0.738	0.509	237.52	368	257.1/266.9

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ZhangetalSupplementaryInformation.pdf](#)