JANUSBM: A DUAL-FIDELITY MULTI-ZONE WHITE-BOX BUILDING MODELING FRAMEWORK

A PREPRINT

Haozhen Cheng
Karlsruhe Institute of Technology
76344 Eggenstein-Leopoldshafen
haozhen.cheng@kit.edu

Hüseyin K. Çakmak
Karlsruhe Institute of Technology
76344 Eggenstein-Leopoldshafen
hueseyin.cakmak@kit.edu

Veit Hagenmeyer
Karlsruhe Institute of Technology
76344 Eggenstein-Leopoldshafen
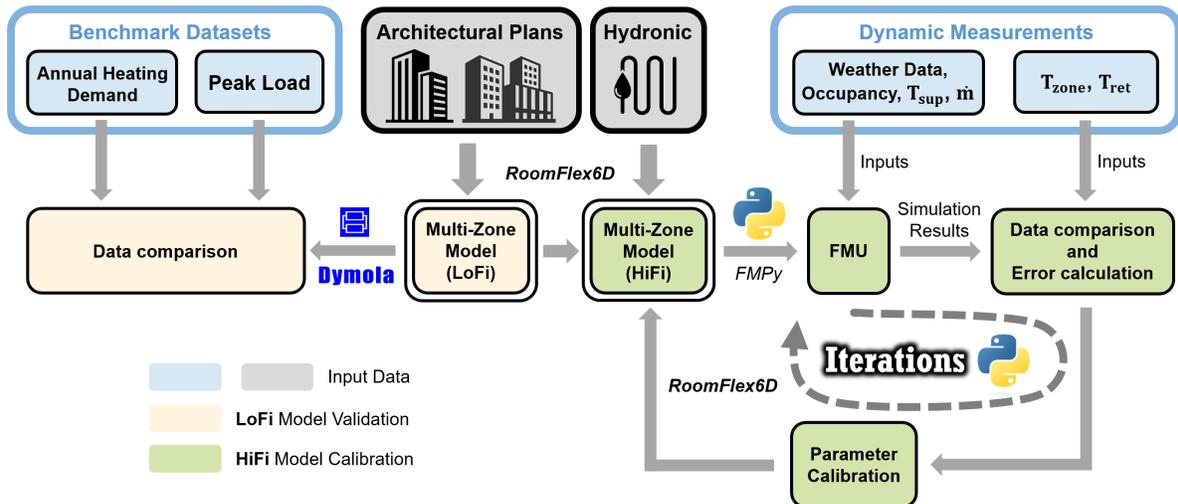veit.hagenmeyer@kit.edu

March 25, 2026

Figure 1: Overview of the structured dual-fidelity modeling framework JanusBM, including data input, model generation, validation, and iterative calibration workflow.

ABSTRACT

Accurate building energy models are crucial for analyzing sector-coupled energy systems, where buildings interact with electrified heating, energy storage, and advanced control across various scenarios. High-fidelity (HiFi) white-box models that resolve hydronic distribution and emitter dynamics can capture short-term transients, yet their numerical stiffness and computational burden limit long-term simulations and large-scale scenario exploration. Conversely, reduced-order low-fidelity (LoFi) representations enable rapid annual assessments but may fail to capture the hydronic- and control-induced dynamics that govern transient and peak behavior. This paper proposes a dual-fidelity, multi-zone white-box building modeling framework, which is called JanusBM, built on a novel topology-driven modeling tool RoomFlex6D, coupling a HiFi hydronic model and a LoFi ideal-load surrogate that removes explicit hydronic states in Modelica. To ensure applicability and physical consistency across time scales, we introduce a two-stage hybrid validation and calibration pipeline that uses complementary data: the IEA EBC Annex 60 benchmark for energy-scale validation and time-series measurements from real-world experimental buildings for hydronic dynamics-scale calibration. Results show that the generated LoFi models achieve a high degree of consistency with Annex 60 benchmark on the energy scale, and the proposed calibration workflow robustly improves

loop-level return water temperature transients and zone-level temperature dynamics. Moreover, the LoFi model achieves orders-of-magnitude faster simulations suited to annual energy analyses, whereas the HiFi model becomes necessary when the required heat differs from the actual delivered heat due to distribution and control limitations, especially in transient and peak-oriented assessments.

Keywords Building · Multi-zone · White-box · Modelica · Modeling · Simulation · Hydronic · Validation · Calibration

1 Introduction

As a central element in sector-coupled energy systems, buildings serve as the primary interface between various energy infrastructures, including heat and electricity [Gea-Bermúdez et al., 2021]. As electrified heating technologies (e.g., heat pumps), distributed photovoltaics, energy storage, and flexible loads become increasingly prevalent, system-level studies are no longer limited to estimating annual energy demand [Jing et al., 2022, Rinaldi et al., 2022]; instead, they must evaluate buildings across a wide range of design and operational scenarios, where short-term dynamics and operational constraints can play a decisive role [Drgoňa et al., 2020, Rinaldi et al., 2021, Vera-Piazzini and Scarpa, 2024].

High-Fidelity (HiFi) multi-zone white-box models, typically implemented in physics-based environments such as EnergyPlus [Crawley et al., 2001] and TRNSYS [Klein, 1988], that explicitly characterize building envelopes, thermal mass, and hydronic heating systems, can reproduce transient behavior including water-side dynamics, emitter inertia, and their interaction with zone temperatures. However, as zoning resolution increases, the hydronic networks and control loops introduce numerical stiffness and rapidly increase the system dimension, making simulation costly and sensitive to solver settings [Blum et al., 2021]. Consequently, HiFi representations are often impractical for long-term simulations, large-scale scenario scans, or optimization workflows that require thousands of evaluations [Bhattacharya et al., 2020].

Conversely, reduced-order or Low-Fidelity (LoFi) models enable fast annual simulations and are widely used for convenient parameter configuration of Building Energy Models (BEMs) and long-term planning studies, but they often neglect hydrodynamics constraints and detailed control interactions. This simplification can systematically fail to represent short-term transients and peak-oriented indices that are critical in sector-coupled applications, thereby subsequent conclusions regarding scale margin, flexibility assessments, and controller robustness can be biased [ISO, 2017]. To accurately represent building systems, multi-zone BEMs have been extensively studied and are becoming increasingly mature. However, researchers often either over-refine zoning in expensive HiFi environments, making large-scale scenario exploration infeasible, or oversimplify accuracy for speed, leading to systematic biases in peak-oriented conclusions and incorrect controller evaluations.

1.1 Contributions

To make model-fidelity choices explicit and comparable, we propose the unified dual-fidelity, multi-zone white-box building modeling framework JanusBM, which is shown in Figure 1, enabling controlled and quantitative comparisons between zoning resolution, model fidelity, and computational cost under consistent envelope assumptions. Specifically, we

- introduce the topology-driven modeling tool RoomFlex6D, to automatically generate zoning-consistent modular multi-zone envelope models, and to configure hydronic loop topologies and their couplings to zones in a systematic behavior;

- integrate two complementary representations HiFi and LoFi derived from the same zoning and envelope description into the proposed dual-fidelity framework JanusBM, ensuring that fidelity comparisons are not affected by envelope modeling differences;

- develop a two-step validation-calibration workflow for the proposed framework using complementary datasets, to validate the LoFi model at energy scale, and to calibrate the HiFi model in terms of dynamic performance using a closed-loop RoomFlex6D–Functional Mock-up Unit (FMU)–Python workflow; and

- demonstrate, how fidelity choices impact computational burden, thereby providing practical evidence to guide fidelity selection in sector-coupled studies.

1.2 Related Work

Multi-zone white-box building energy modeling is a mature research area. Compared with single-zone representations, multi-zone models indicate inter-zonal heat transfer paths and system-building interactions, such as solar and internal

gains, and zone-specific Heating, Ventilation and Air Conditioning (HVAC) interactions, which are essential for capturing dynamic responses, peak loads, and system dynamics under hydronic heating operation [Zhang et al., 2022]. Existing studies related to this topic can be broadly grouped into four categories: (a) automatic zoning strategies, that balance zoning resolution against computational burden [Jansen et al., 2021, Shin and Haberl, 2022], (b) LoFi or reduced-order modeling toolchains, including ideal-load surrogates and Reduced Order Model (ROM)-based workflows, which enable fast simulations for large-scale simulations for urban planning and scenario analysis [Belic et al., 2021, Jansen et al., 2021], (c) interoperability and standardized Modelica-based workflows, where detailed libraries establish modeling conventions and support Building Information Modeling (BIM)-to-Modelica transformations [Wetter et al., 2021a, 2019] and Functional Mock-up Interface (FMI)-based co-simulation [Wetter et al., 2021b] for multiphysics building and HVAC modeling, and (d) validation and calibration practices, in which benchmark datasets and real-world measurements [Langner et al., 2025] provide complementary—yet asymmetric—evidence of model credibility across different scales [Ohlsson and Olofsson, 2021].

Zoning strategy and structured simplification. Numerous studies have investigated zoning definition and zoning strategies based on geometry, usage patterns, and facade characteristics, as well as structured simplification methods. Georgescu and Mezić [2015] proposed a structured model simplification technique that couples zoning structure with simulation accuracy and complexity, arguing that physically structured zoning is preferable to arbitrary subdivision. Chen and Hong [2018] introduced a three-zone approach guided by ASHRAE 90.1 Appendix G guidelines [ASHRAE, 2026], combined with floor or zone multipliers, to accelerate simulation while maintaining acceptable accuracy for annual assessments. Shin and Haberl [2019] provided a comprehensive review of zoning methods and highlighted the recurring trade-off between zoning resolution and computational burden. In practical building energy system studies, however, building geometry often varies from project to project, and the simulation workflow requires repeated model generation under different zoning schemes, envelope configurations, and system layouts. Traditional geometry-driven methods based on detailed 3D model libraries or BIM data may not be flexible enough in this context, as even minor changes in zoning or topology often require significant manual rework and consistency checks. Another related modeling framework, ALICE/ALICE2Modelica, was developed by Mork et al. [2024] to support automated generation of large-scale, multi-zone building models. In its current workflow, building orientation, adjacency relationships, and interior/exterior wall identification are derived from per-floor Scalable Vector Graphics (SVG) plans. This dependency means that changes in zoning or building layout typically may require manual updates to the floor-plan inputs, limiting rapid re-zoning and structural reconfiguration. Motivated by this limitation, a topology-driven Lego®-like modeling tool RoomFlex6D is introduced by the present paper, enabling more flexible and reconfigurable multi-zone model construction.

Reduced-order toolchains. To support large-scale scenario exploration and planning studies, many approaches employ reduced-order or LoFi representations, including ideal-load surrogates and ROM-based workflows used in urban simulations. These methods are highly scalable, and are convenient for parameter configuration and long-term analyses. However, early work in this direction—for example, TEASER [Remmen et al., 2018] used by Modelica AixLib, and CityBES [Chen et al., 2017]—already recognized that single-zone or lumped models may fail to accurately reflect the spatial distribution of temperature and loads, particularly when facade exposure, occupancy patterns, and equipment operation differ across zones, thereby limiting the credibility of both comfort assessment and control-oriented analysis [Johari et al., 2020].

Modelica interoperability workflows. The Modelica-based building and HVAC libraries, including International Building Performance Simulation Association (IBPSA)-related [International Building Performance Simulation Association, 2026] Modelica libraries [Jorissen et al., 2018, Maier et al., 2023, Nytsch-Geusen et al., 2012, Wetter et al., 2014], provide a robust foundation for multiphysics, multiscale simulations, enabling explicit representation of building envelopes, thermal mass, fluid networks, and control systems. In parallel, interoperability efforts such as the IBPSA activity Annex 60 [Wetter and van Treeck, 2017] demonstrated the potential of modular, tool-independent simulation chains. Nevertheless, these efforts rarely support controlled cross-fidelity comparisons, making it difficult to validate the model correctness across time scales and accuracy levels.

Validation and calibration across time/accuracy scales. A further obstacle is the asymmetry in the availability of data for validation and calibration: benchmark datasets typically offer standardized annual indicators but lack time-resolved measurements, while measured time-series data [Campodonico-Avendano et al., 2026, Tajalli-Ardekani et al., 2025] provide rich, detailed signals but are rarely suitable for reliable annual energy calibration under controlled conditions [Ohlsson and Olofsson, 2021]. This mismatch becomes particularly critical in the context of **digital twins**, which are not just standalone models, but also executable representations of as-built assets that are continuously informed by an evolving stream of operational data, and require mechanisms to dynamically update model parameters to remain consistent with their physical counterparts over time [Wright and Davidson, 2020]. Consequently, establishing a credible building digital twin necessitates systematic model calibration rather than one-off validation: energy-scale

benchmarks can constrain long-term consistency, while dynamics-scale measurements enable parameter updating for short-term behaviors and peak conditions.

1.3   Outline

Building on the identified research gaps, this paper proposes a dual-fidelity white-box building modeling framework, including the topology-driven RoomFlex6D tool and its Modelica-based generation of LoFi and HiFi model instances under consistent envelope assumptions, as presented in Section 2. Section 3 details the two-stage validation–calibration pipeline, while Section 4 introduces the zoning strategy and explains how zoning configurations are compared in a controlled manner. Section 5 reports the evaluation results, covering LoFi validation and HiFi calibration, together with the resulting accuracy–runtime trade-off, and finally, Section 6 concludes the paper and outlines future work.

2   Model Construction

A novel Lego®-like multi-zone white-box building modeling tool, that can generate both HiFi and LoFi models with the same building envelope configuration, is introduced in the Sections below.

2.1   RoomFlex6D Modeling Tool

To describe building instances in a way that directly supports rapid multi-zone model generation and systematic zoning studies, the present paper proposes a novel, lightweight, and simulation-oriented modeling tool, RoomFlex6D. Unlike general semantic building standards such as CityGML [Open Geospatial Consortium, 2026], IFC [buildingS-MART, 2026], or gbXML[gbXML, 2026], which enable rich geometric and semantic interoperability across domains, RoomFlex6D can generate a minimally sufficient description for thermal simulation that can be edited and repeatedly regenerated, and is highly comparable across different zoning schemes. Specifically, RoomFlex6D encodes the building as a topology-driven, face-level adjacency graph (zone, orientation, boundary type, and neighbor) that automatically maintains parameter consistency between adjacent faces of adjacent thermal zones.

Furthermore, the presented tool allows for the flexible construction of hydronic loop topologies and their coupling with zones and embedded/non-embedded heating elements, which are crucial for transient hydronic dynamics and controller interactions, which cannot be implemented in a stable and standardized manner in typical semantic models [Luo et al., 2021].

RoomFlex6D provides a Python-based topology-driven Graphical User Interface (GUI) in Figure 2 that encodes a multi-zone building as a set of orthogonal zone volumes, and exports a face-level parameter table for automated Modelica model generation. Each zone is connected to its neighbors (or the ambient) through six directional faces (N/E/S/W/U/D), where each face is assigned an adjacent-face boundary type (Table 1 based on Wetter [2006]). This structured Lego®-like representation ensures the uniqueness of adjacency relationships while avoiding ambiguous or non-manifold connections, enabling controlled zoning changes while keeping envelope assumptions comparable across model instances. The exported parameter table serves as the source for generating and updating Modelica models (see Figure 3 for an example). Based on the generated structural parameter table (Table 2), the heat transfer mechanisms achievable through the RoomFlex6D modeling tool include:

- conductive heat transfer from walls, floors, and roofs,
- convective heat exchange between zoned air and interior surfaces,
- convective heat exchange on exterior surfaces affected by external temperature and wind speed,
- solar radiation through transparent components, and the heating of the external surface by solar radiation,
- internal heat generated by people and equipment, and
- heat loss due to infiltration and ventilation.

2.2   HiFi Hydronic Heating System Model

As described in Table 2, in addition to the envelope and multi-zone structure, RoomFlex6D supports an explicit representation of the building heating hydronic and its coupling to zone heat transfer. Typical configurations such as radiator heating and Underfloor Heating (UFH) are modeled as dynamic heaters with finite thermal capacity, exchanging heat with zone air (or embedded constructions) via convection and radiation. The hydronic network comprises supply/return piping, hydraulic resistances, circulation pumps, and control elements (e.g., control and mixing valves),
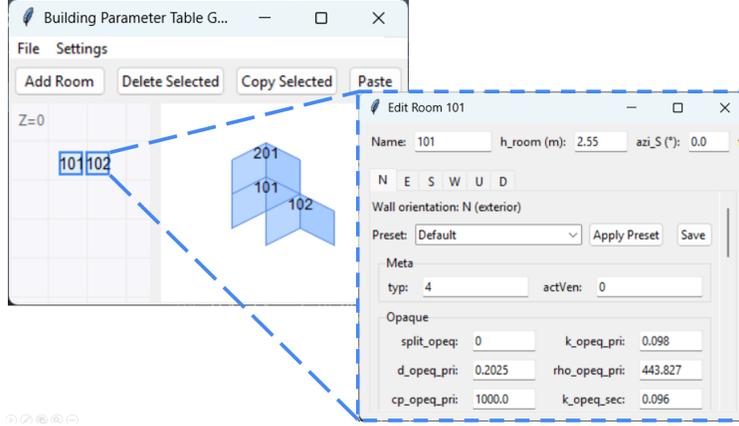
Figure 2: RoomFlex6D GUI for topology-driven zoning and face-level boundary assignment: drag-and-drop zone layout (left), per-zone face properties and hydronic mapping (right), exporting a reproducible parameter table. See Appendix A for more details.

Table 1: Face boundary-type codes used in RoomFlex6D to parameterize openings/ground coupling for each zone face.

Boundary Type	Comments
1	There is an opening that cannot be closed.
2	There is an operable window(s).
3	There is an operable door(s).
4	There is no operable opening, and the exterior is not ground.
5	There is no operable opening, and the exterior is ground.

enabling simulation of operational dynamics, such as switching cycles and transient responses to boundary-condition changes.

Based on the generated zone structure, the HiFi model integrates these hydronic components with multi-zone heat transfer, solar gains, internal loads, infiltration, and building thermal mass, as shown in Figure 4. However, this coupling introduces numerical stiffness due to multiple timescales — from rapid hydronic control to the slow thermal response of building structures. As the number of thermal zones increases, the system dimensionality grows rapidly, leading to increased computational costs and greater sensitivity to solver settings (e.g., time step and tolerance). This motivates the complementary LoFi ideal-load model used for large-scale and year-round studies within the JanusBM framework.

2.3 LoFi Ideal-Load Building Model

To enable rapid simulations suitable for year-round analysis and large-scale scenario exploration, a reduced-order ideal-load multi-zone building LoFi model is derived as a surrogate representation of the HiFi model. This model retains the same multi-zone structure, envelope description, and boundary conditions generated by RoomFlex6D, while removing explicit representations of hydronic components and fluid dynamics. It provides heat to each zone via an idealized heat input that meets the heating requirements to maintain a preset zone temperature. Therefore, the LoFi model analyzes the thermodynamics of the building envelope and zone air volumes while simplifying the behavior of the internal hydronic system. Consequently, the solver burden caused by rapid hydronic timescales is eliminated, which is a major source of numerical stiffness in HiFi models.

Despite the simplifications, key physical properties relevant to system-level analysis are preserved in the LoFi model, including explicit simulations of multi-zone heat transfer, solar radiation gain, internal loads, infiltration, and building thermal mass, thus capturing the overall building thermal behavior with reasonable accuracy. This makes the simplified model particularly suitable for year-round simulations, extensive parameter scans, and scenario sets involving multiple energy system configurations, to support tasks such as annual energy assessments, scenario comparisons, and preliminary system design.

In summary, within the proposed dual-fidelity framework, the simplified ideal load model is not intended to replace the full hydronic model. Instead, it serves as a computationally efficient alternative model that can be calibrated and validated against high-fidelity models and benchmark data.
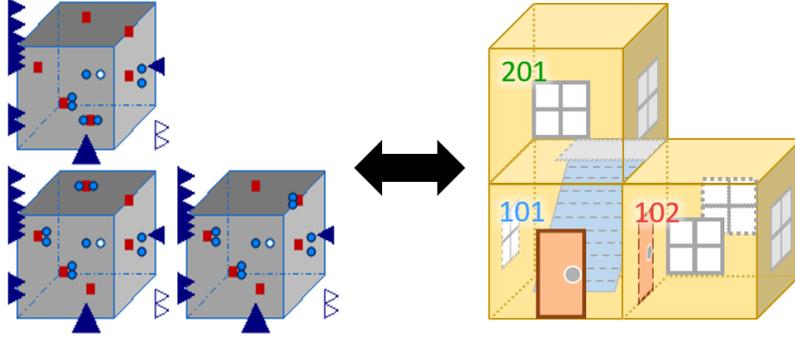
Figure 3: Topology-to-Modelica translation example: a 3-zone model instance generated from the RoomFlex6D parameter table (left) and its schematic graph representation (right).

Table 2: Face-level parameter-table minimal schema enabling zoning-consistent envelope and hydronic coupling (each row = one zone face). Full list in Appendix D.

Parameter	Symbol
Primary zone ID	z_{pri}
Face orientation	ori
Adjacent zone ID	z_{adj}
Boundary type	typ
Ventilation flag	$actVen$
Split flag	$split$
Thermal parameter	$k, d, \rho, c_p, A_{\mathrm{opeq,fra,gla,open,RS}}$
Hydronic	$heat_{\mathrm{type,loop,order,ctrl}}$
Building orientation	azi_{S}

3   Validation and Calibration Pipeline

To ensure physical consistency and applicability across different model accuracies, the present paper introduces a structured validation and calibration process.

The Annex 60 benchmark dataset provides standardized building descriptions and annual heating demand references, enabling energy-scale validation under well-defined boundary conditions [Wetter and van Treeck, 2017]. However, Annex 60 does not provide time-resolved measurements, thus failing to support validation of transient system behavior.

The experimental building dataset from the Living Lab at Karlsruhe Institute of Technology (KIT) Campus North is available [Tajalli-Ardekani et al., 2025]. Conversely, this dataset contains local weather data, time-series measurements of zone temperature, supply and return water temperatures, and mass flow rates in each of the hydronic heating loops (UFH), and is stored in a time series database with a Grafana (grafana.com) interface for further data analysis and export [Hagenmeyer et al., 2016]. While these measurements capture detailed dynamic behavior, the experimental building operates under user-driven conditions and cannot provide reliable annual heating demand values suitable for direct energy-scale calibration.

This asymmetry motivates a two-step validation and calibration strategy described below, leveraging the strengths of each dataset while avoiding their respective limitations.

Workflow of Pipeline. In the present paper, these two types of data are used by the LoFi and HiFi models, respectively, to validate the model reliability at energy scales, and to calibrate model parameters at dynamic scales by using a novel iterative calibration workflow, as shown in the bottom right corner of Figure 1. It outlines the closed-loop parameter calibration process based on an iterative RoomFlex6D–FMU–Python workflow:

- **data input:** Building plans and geographic information (such as architectural drawings or OpenStreetMap (OSM) [OSMF, 2026]) serve as structured inputs, and during simulation, the model receives external boundary and system inputs from Living Lab experimental data [Tajalli-Ardekani et al., 2025], including weather data, occupancy-related signals, and UFH information;
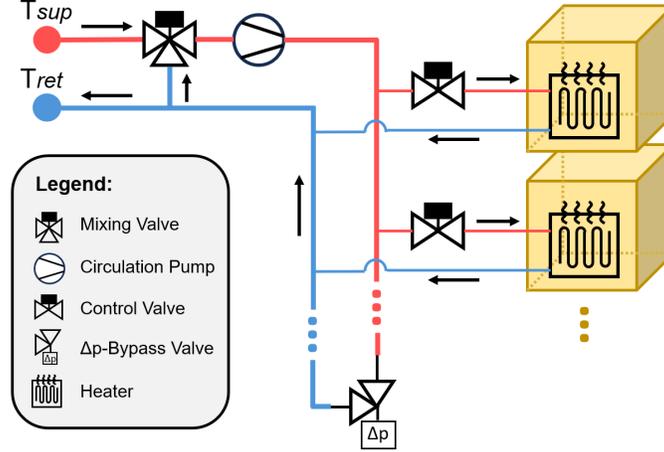
Figure 4: HiFi model structure generated by RoomFlex6D: explicit hydronic network, emitters, and zone coupling: source of stiffness but required for transient loop dynamics.

- **model generation and simulation:** For the HiFi models calibration, RoomFlex6D automatically generates multi-zone white-box building models from the inputs mentioned above, which are then exported as executable FMUs, and the Python package FMPy handles batch simulation iterations;

- **loss calculation and parameter calibration:** The simulation outputs are aligned and compared with synchronously measured data from the Living Lab as time series to calculate errors, and construct a calibration objective function, which drives the parameter update module to iteratively correct the parameters to be calibrated, and then the updated parameters are written back to the model generation-simulation link for the next round of simulation.

Algorithm: Specifically, within a selected calibration time window, the observed signal set is defined as $\mathcal{S} = \{T_{\mathrm{zone}}, T_{\mathrm{ret}}\}$, where T_{zone} the zone air temperature and T_{ret} denotes the return water temperature of each hydronic heating loop. The parameter vector to be calibrated is $\Phi = \{dis_{\mathrm{pip}}, h_{\mathrm{int}}\}$, where dis_{pip} represents the pipe spacing distance of the embedded floor-heating emitters and h_{int} denotes the effective internal convective heat-transfer coefficient between the floor-heating surface and the surrounding zone air. In the implementation, dis_{pip} is treated as a global (shared) parameter across all loops to meet real-world physics, while h_{int} is updated per loop. Both are optimized in log-space to enforce positivity and improve numerical stability.

- **Flow-gating:** For each iteration, simulation outputs are aligned to the measurement time grid within the calibration time window. To prevent non-informative segments from distorting the calibration results, the return water channels are evaluated only during time steps with sufficient measured mass flow rate, i.e., $\dot{m}(t) \geq \dot{m}_{\mathrm{min}}$, which means that periods with $\dot{m} \approx 0$ do not contribute to the return temperature residuals. This is because in the one-dimensional fluid network model of Modelica, the convection processes become numerically ill-conditioned when $\dot{m}$ approaches 0, and temperature states within volume elements can be affected by fluid mixing, leading to distorted return temperature signals that fail to represent the true heat transfer physics.

- **Bias-driven update:** The scalar objective is constructed as a robust, weighted aggregation of normalized residuals across all channels. For channel $c \in \mathcal{S}$, the normalized residual is $r_c(t) = [y_c^{\mathrm{sim}}(t) - y_c^{\mathrm{meas}}(t)]/s_c$, where s_c is a channel-dependent normalization factor, and separately configured for zone temperatures and return temperatures. A Huber penalty ρ_δ is applied to reduce the influence of outliers and sporadic numerical spikes in the residuals. The total loss is then computed as

$$\mathcal{J}(\mathbf{x}) = \frac{\sum_{c\in\mathcal{S}} w_c \sum_{t\in\mathcal{T}_c} \rho_\delta\big(r_c(t)\big)}{\sum_{c\in\mathcal{S}} w_c |\mathcal{T}_c|}, \tag{1}$$

where w_c is the channel weight reflecting relative importance, and $\mathcal{T}_c$ is the set of valid time indices for channel c, in which the flow-gating for return temperature channels is also included.

After obtaining the total loss, the parameter update in the present paper adopts a bias-driven iterative scheme that uses the signed mean deviation of each channel. For a zone z, a bias statistic is computed as $b_{\mathrm{zone},z} = \mathrm{mean}(T_{\mathrm{zone},z}^{\mathrm{sim}} - T_{\mathrm{zone},z}^{\mathrm{meas}})$. For a loop L, a flow-gated bias statistic is computed as $b_{\mathrm{ret},L} = \mathrm{mean}(T_{\mathrm{ret},L}^{\mathrm{sim}} -$

$T^{\mathrm{meas}}_{\mathrm{ret},L}$) evaluated only on $\dot{m}_L \geq \dot{m}_{\mathrm{min}}$. These bias terms are combined into a signed direction indicator d_L that indicates whether the loop-to-zone heat exchange should be strengthened or weakened, using separate scaling for zone and return signals. The per-loop updated parameter h_{int} is updated in log-space with step size limiting and per-loop importance scaling by using the Python method clip:

$$\log h^{(k+1)}_{\mathrm{int},L} = \log h^{(k)}_{\mathrm{int},L} + \mathrm{clip}(\alpha_h \, \mathrm{imp}_L \, d_L, \, -\Delta_h, \, \Delta_h), \tag{2}$$

where α_h is the nominal step size, Δ_h is the maximum per-iteration log-step, and imp_L is an error-derived importance factor that increases the update magnitude for loops with larger loss contributions. The global pipe spacing parameter dis_{pip} is updated synchronously based on the aggregated direction across loops:

$$\log dis^{(k+1)}_{\mathrm{pip}} = \log dis^{(k)}_{\mathrm{pip}} + \mathrm{clip}(-\alpha_d \, d_{\mathrm{g}}, \, -\Delta_d, \, \Delta_d), \tag{3}$$

with

$$d_{\mathrm{g}} = \mathrm{average}(d_L; \, \mathrm{weights} = \mathrm{imp}_L), \tag{4}$$

where the presence of the negative sign allows the pipe spacing to be reduced when the system is biased towards insufficient heating, thereby increasing effective heat transfer. To ensure physically plausible values and stable convergence, the updated parameters are constrained by predefined bounds $dis_{\mathrm{pip}} \in [dis_{\mathrm{min}}, dis_{\mathrm{max}}]$ and $h_{\mathrm{int},L} \in [h_{\mathrm{min}}, h_{\mathrm{max}}]$.

- **Damping:** A damping mechanism integrates the previous iterate with the proposed update in log-space to mitigate oscillations under strong transients:

$$\mathbf{x}_{\mathrm{use}} = (1 - \lambda)\mathbf{x}^{(k)} + \lambda \mathbf{x}^{(k+1)}. \tag{5}$$

Finally, to improve efficiency across different channels, each channel's loss contribution is separately computed: $\mathrm{contrib}_c = w_c \sum_{t \in \mathcal{T}_c} \rho_\delta(r_c(t))$, and the step sizes (α_h, α_d) and damping factor λ based on the contribution ranking are adapted to emphasize dominant error sources. This closed-loop design couples a robust objective with a physically informed update rule, allowing the calibration to progressively reduce dynamic-scale errors while preserving feasibility and numerical stability.

4 Zoning Strategy

Zoning resolution is a key design dimension for multi-zone building energy consumption models, directly impacting physical accuracy and computational performance [Pan et al., 2023]. Based on the structured zoning of building models generated using RoomFlex6D, the building is progressively subdivided into coarser and finer multi-zone configurations while maintaining consistent envelope parameters, orientation, exposure, and adjacency relationships. Figure 16 in Appendix C shows the district studied in the Annex 60 project [Wetter and van Treeck, 2017], while Figure 5 shows architectural floor plans of five different building topologies within the district, arranged vertically from bottom to top, corresponding to the ground floor to the top floor. Each colored block represents a divided thermal zone. Further details are discussed in Section 5.1.

Figure 5: Zoning configurations per Annex 60 typology used in LoFi validation (from left to right: D, S, T, A/O; color blocks denote thermal zones).

5 Evaluation

Based on the modeling tool and verification-calibration pipeline proposed in Section 2 and 3, this Section first verifies the LoFi model on the IEA EBC Annex 60 benchmark to confirm the applicability of the model generation method and

zoning settings to annual and peak load indicators. Based on this, explicit hydronic loops and terminal heat exchange process are then introduced to construct a HiFi model. Dynamic response analysis and iterative parameter calibration are carried out in combination with real-world laboratory time series measurement data, focusing on evaluating the transient consistency and improvement of the loop return water temperature and the thermal zone temperature behavior.

5.1 Energy-Scale Validation of LoFi Using the Annex 60 Benchmark

The LoFi model is first validated on the Annex 60 district benchmark using the prescribed boundary conditions and inputs, while keeping the envelope assumptions consistent across zoning configurations. For each building, (a) the annual heating demand and (b) the peak heating load are evaluated, since the former reflects energy-scale consistency, while the latter is critical for short-term and peak-oriented sector-coupled analyses. To ensure comparability with the reference study, all simulations in the present paper strictly adhere to the same boundary conditions and modeling assumptions specified in the Annex 60 final report as described in Appendix C.

Annex 60 has demonstrated that the number of thermal zones is not the primary source of differences in annual and peak heating demand across different modeling methods. Therefore, the zoning resolution of the selected zoning scheme is intentionally kept at a moderate level in Figure 5 to avoid introducing unnecessary modeling complexity. Apartments (A) and office buildings (O) are represented as single-story structures, consistent with the baseline definition. Furthermore, for buildings of the same type in different locations (e.g. Building 3 and 19 in Figure 16 in Appendix C), the models can be fine-tuned using the orientation parameter azi_S and adjacency settings in the RoomFlex6D modeling tool to adapt to changes in energy consumption due to variations in sunlight and wind speed, ensuring that differences in simulated heating demand stem from the modeling framework, rather than differences in building definitions.

Annual heating demand. The energy-scale evaluation focuses on annual heating demand and peak heating load, which are the primary performance indicators defined in the Annex 60 neighborhood benchmark. Figure 6a compares the annual heating demand of individual buildings obtained from the proposed framework with the Annex 60 reference envelope. For each building or building typology, the vertical error indicators in Figure 6b indicate the minimum–maximum range reported across five Annex 60 participants, the shaded bands represent errors of ±20% based on the average of the reference values, while the red dots on the indicators represent results obtained using the presented LoFi models. Across all buildings, the simulated annual heating demand lies within or very close to the reference error range. Differences between zoning resolutions remain limited at the annual energy scale, and no systematic overestimation or underestimation is observed for any specific building typology.

Peak heating load. While the simulation results show a high degree of agreement with the Annex 60 benchmark in terms of annual heating demand, significant deviations are observed for peak heating load, which are shown in Figure 7. In particular, for buildings with relatively high reference values, the simulated peak loads often fall below the lower limit of the Annex 60 reference range. This behavior is observed in multiple high-demand buildings, indicating a systematic underestimation of extreme peak conditions. This deviation can be explained by the inherent sensitivity of peak load indicators to short-term dynamics and modeling assumptions. In the Annex 60 benchmark, peak loads are strongly influenced by the simulation time resolution, the control strategy, and assumptions about system response under extreme outdoor conditions. In contrast, the LoFi models used for energy-scale validation intentionally smooth fast transients and exclude explicit system-level dynamics, resulting in a more moderate peak response.

Importantly, this underestimation does not affect the consistency of annual heating demand or the overall energy balance of the studied neighborhood, as shown in Figure 8, which remains well consistent with the Annex 60 reference values. The observed differences reflect a known limitation of energy-scale surrogate models when applied to peak-oriented indices, motivating subsequent analysis using higher-fidelity models to investigate dynamic behavior and peak load characteristics in more detail.

5.2 Dynamic-Scale Calibration of HiFi Using Experimental Building Data

The HiFi case study is based on the Living Lab test building B665 [Langner et al., 2025] and is implemented using the RoomFlex6D modeling tool. The building is represented using a LoFi envelope representation, divided into a 12-zone thermal network (six zones per floor; the zoning layout and thermal zone codes are shown in Figure 9). An explicit hydronic subsystem is added, containing a total of 10 UFH loops, that is, each thermal zone, except for zones 102 and 204, contains one UFH loop.

To validate and calibrate the HiFi model, all boundary conditions, control actions, and validation objectives are taken from the Living Lab measurements over December 30, 2024 to January 3, 2025 with a 1 min sampling interval [Tajalli-Ardekani et al., 2025]. The calibration focuses on two UFH-related parameters per loop: pipe spacing distance dis_{pip} and the effective convective heat transfer coefficient h_{int} between the embedded UFH and the zone air, as discussed in
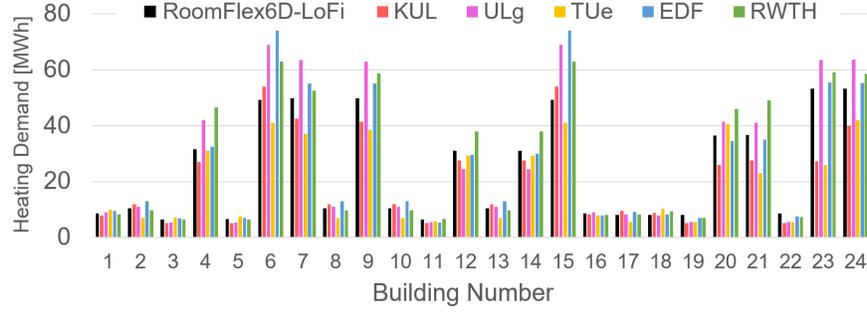
(a) Comparison of LoFi simulation results across 24 studied buildings with results from each of the 5 participants.
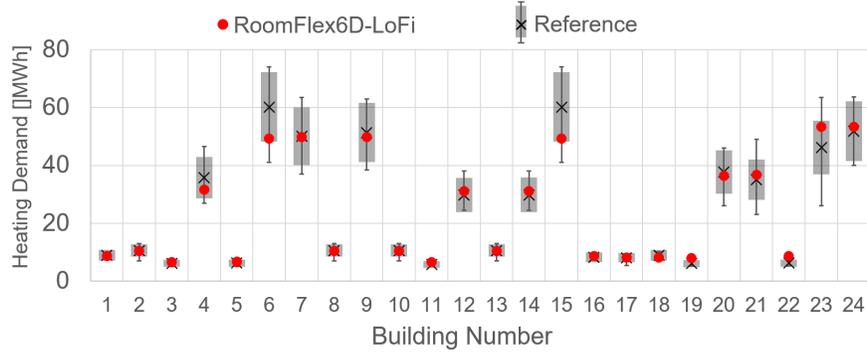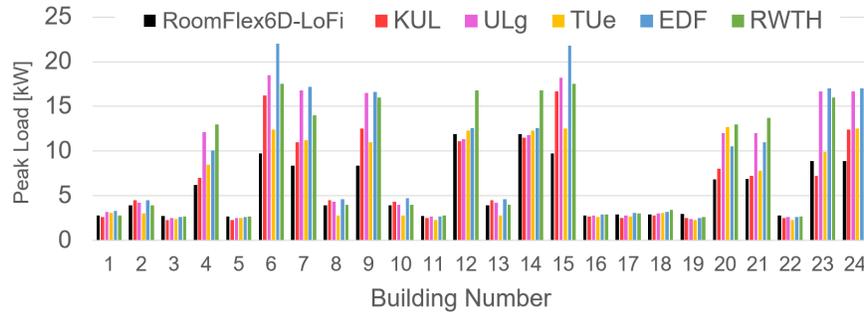

(b) Comparison of LoFi simulation results across 24 studied buildings with the average results for all participants.

Figure 6: Building-level annual heating demand: LoFi vs Annex 60 participant ranges-agreement within benchmark across 24 buildings.

Section 3. In contrast, the envelope thermal parameters mainly affect the slower zone-level drifts and are difficult to identify without additional constraints. To avoid confusion and parameter non-identification, envelope parameters (e.g., k, d, ρ, c_p from Table 2) are not tuned at this stage. Furthermore, to resolve key hydronic and control-related transients, the HiFi model simulations employ a fixed output step size of 30 s.

Calibration convergence. During the calibration of the HiFi model, the authors weigh the diminishing returns against the computational cost of repeated HiFi simulations, running 23 iterations before terminating after the target value reached significant stability: the total loss decreases from 0.64 to approximately 0.32, with the main improvement occurring in the first 8–10 iterations, followed by only minor improvements, as shown in Figure 10a, which indicates that improvements under hydronic and thermal conditions reach a balance. It is important to note that due to the different number of signals and effective samples aggregated in the two groups of contributions (e.g., $T_{\rm ret}$ only applies when mass flow rate $\dot{m} > \dot{m}_{min}$) in Figure 10b, their absolute values cannot be directly compared. Nevertheless, the convergence behavior after iterations and the trend of relative loss reduction are successfully evaluated.

$T_{\rm ret}$ **evaluation.** Based on the convergence trend in Figure 10, it is evaluated whether the calibrated parameter set could improve the practically meaningful loop-level dynamics. Therefore, Figure 11 presents a time-series comparison of two UFH loops located in different zones (104 and 206) with different operating modes within the studied measurement window, where loop 104 exhibits frequent and irregular flow switching events and a significant transient ramp, while loop 206 contains regular switching events and a longer quasi-steady-state period interspersed with switching events, serving as a supplementary case for evaluating steady-state offset. In each subplot, the black curve represents the measured return water temperature $T_{\rm ret}$; the orange and blue curves represent the simulated $T_{\rm ret}$ before and after calibration, respectively. The gray-shaded area corresponds to low-flow conditions ($\dot{m} = 0 \ kg/s$ marked by the gray dashed line on the right axis), where the return water temperature residuals are not considered in the total loss. Consistent with this logic, the largest improvements after calibration occur during periods of active circulation, where the post-calibration model better reproduces both the transient ramp shapes and the timing of return temperature responses to flow switching.

(a) Comparison of LoFi simulation results across 24 studied buildings with results from each of the 5 participants.
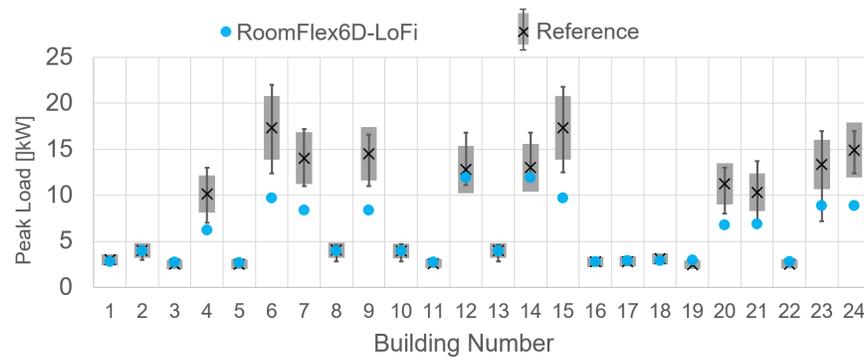
(b) Comparison of LoFi simulation results across 24 studied buildings with the average results for all participants.

Figure 7: Building-level peak heating load: LoFi systematically underestimates high-demand peaks relative to benchmark ranges: limitation of ideal-load surrogates.

Figure 8: Neighborhood aggregation for Annex 60 benchmark: annual demand remains consistent despite peak-load underestimation at the building level.

$T_{\mathbf{zone}}$ **evaluation.** To prevent the improved UFH heat exchange parameters from introducing indoor temperature compensation errors, the iterative results at the zone level are evaluated in Figure 12. To concisely and effectively examine the performance of different zones within the same calibration time window, zone 106, which shows the most significant contribution from the calibrated residual temperature, and zone 104, which operates stably and has relatively low residuals, are selected. In both zones, the calibrated trajectories are closer to the measured values, confirming that hydronic tuning does not reduce the indoor temperature fit.

For zone 106, a residual bias of up to $1.1\,°\mathrm{C}$ within the observation time window remains after calibration, although calibration significantly reduces the bias. The residual exhibits a slow drift across the entire time window rather
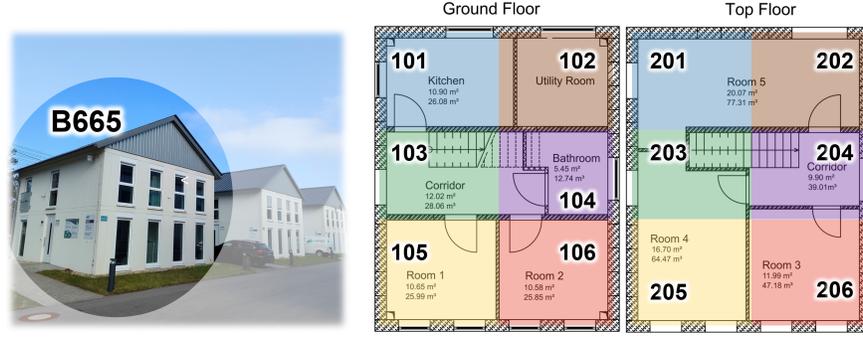
Figure 9: Living Lab [Wiegel et al., 2022] B665 case: building context and 12-zone layout and IDs used for HiFi calibration (10 UFH loops mapped to zones).
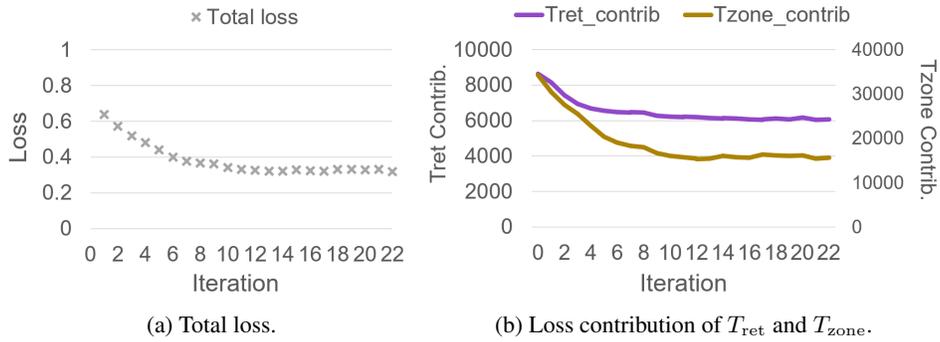

(a) Total loss.

(b) Loss contribution of T_{ret} and T_{zone}.

Figure 10: HiFi calibration convergence over 23 iterations: evolution of total loss and grouped contributions (T_{ret} gated by mass flow vs T_{zone}). Main gains occur in the first 8–10 iterations.

than being limited to a single heating event, suggesting that the residual is primarily affected by zone-level structural uncertainties, which are weakly related to the calibrated UFH parameters (dis_{pip} and h_{int}). This may be attributed to

- **air node simplification:** the proposed model simplifies the thermal zone space into a well-mixed air node, and the sensors may measure sub-zones colder than the model's average air temperature;

- **interior decorations:** actual interior modifications, such as carpets, increase the effective thermal resistance between the UFH and the zone air, causing the model to overestimate the delivered heat;

- **neglected thermal mass:** unmeasured indoor thermal mass (e.g., furniture and other items) alters the effective thermal buffer and time constant, potentially resulting in an overestimation of the simulated temperature gradient; and

- **air infiltration:** air exchange through door and window leaks could also increase real heat losses relative to the modeled assumptions.

Unlike the T_{ret}-model, temperatures in thermal zones are affected by more disturbances. However, the hydronic calibration of the HiFi model in the presented work improves loop-level dynamics and yields uncompensated improvements at the zone-level, providing guidance for future modeling improvements.

Global consistency and Median Root Mean Square Error (median RMSE). Figure 13 supplements the convergence evidence in Figure 10, demonstrating the consistency of all hydronic loops and thermal zones on a sample-by-sample basis from a global perspective. The left subplot summarizes all aligned T_{ret} samples and colors them by $\dot{m}$ level, which is distinguished by $0.015\,\mathrm{kg\,s^{-1}}$ and $0.03\,\mathrm{kg\,s^{-1}}$, thus extending the loop-level observations in Figure 11 beyond the two selected loops. The scatter points are concentrated near the fitted lines (marked with grey color), while the degree of dispersion clearly depends on the operating state, that is, points at high flow rates form narrower bands, while points at low flow rates exhibit greater dispersion, which is consistent with the evaluation in Figure 11. The right subplot also summarizes all zonal temperature samples, providing a concise correspondence to the representative temperature trajectories in Figure 12, showing that residual errors are dominated by structural uncertainties rather than insufficient

(a) T_{ret} and $\dot{m}$ of the UFH loop of zone 104.

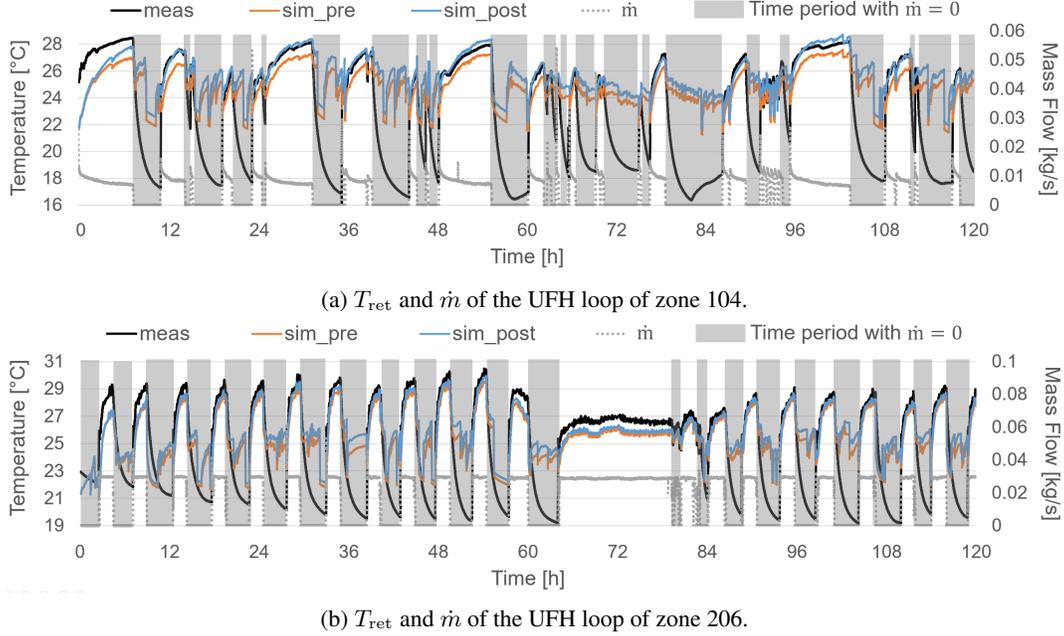(b) T_{ret} and $\dot{m}$ of the UFH loop of zone 206.

Figure 11: HiFi loop-level transient before/after calibration (zones 104 & 206): improved ramp shape and switching response during active circulation. Gray shading indicates zero-flow intervals excluded from T_{ret} loss.

(a) T_{zone} of zone 104.

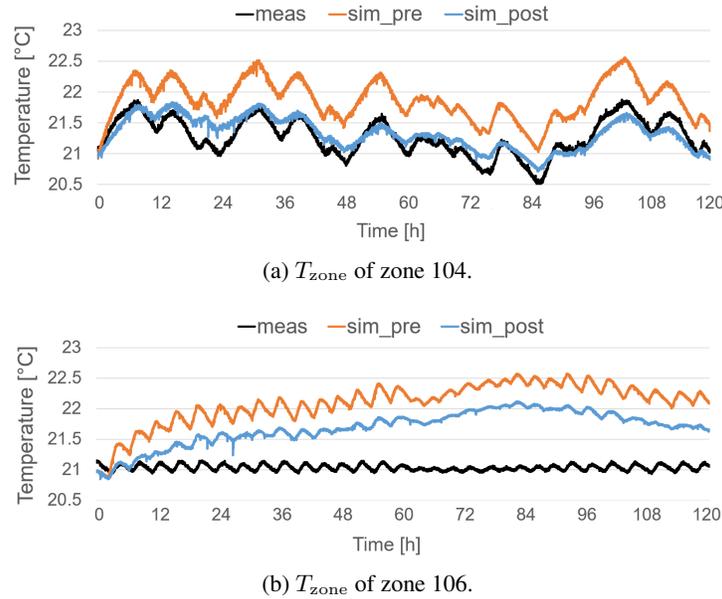
(b) T_{zone} of zone 106.

Figure 12: HiFi zone-level temperatures (104 & 106) before/after calibration: hydronic tuning improves fit without compensating errors; residual drift indicates envelope and structural uncertainty.

convergence of the hydronic calibration. In addition to the visual consistency based on the scatter plot, Figure 13 also reports the quantitative improvement through the median RMSE of all signals within each group.

5.3   Discussion

Increasing zonal resolution and model complexity can enhance the model's representational capabilities, but it also significantly increases simulation costs. For the LoFi model with ideal loads, the annual simulation runtime increased from approximately $10\,\mathrm{s}$ (4-zones) to $47.3\,\mathrm{s}$ (16-zones), showing a quasi-linear increase with the number of zones (see

13

(a) Scatter plot of T_{ret} colored by flow level.
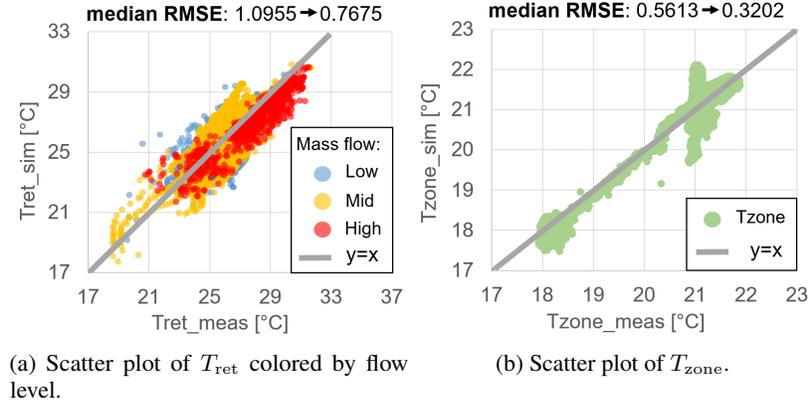
(b) Scatter plot of T_{zone}.

Figure 13: HiFi post-calibration global consistency across all samples, identity line y=x, and median RMSE improvement.

Figure 14 and Table 3). In contrast, at the same 16 zoning resolution, upgrading the LoFi model to a high-precision HiFi model that includes a hydronic distribution network increases the single-simulation time to approximately $350\,\mathrm{s}$ for only 5 simulated days. The main reason for this significant difference is not only the increased computational cost due to model stiffness and complexity, but also the smaller step size set in HiFi models to achieve transient observation of dynamic behavior. Despite the large increase in cost, the HiFi model can reproduce the dynamic changes in the loop-level return water temperature and zone-level temperature behavior under varying mass flow, which cannot be expressed by the ideal load assumption.
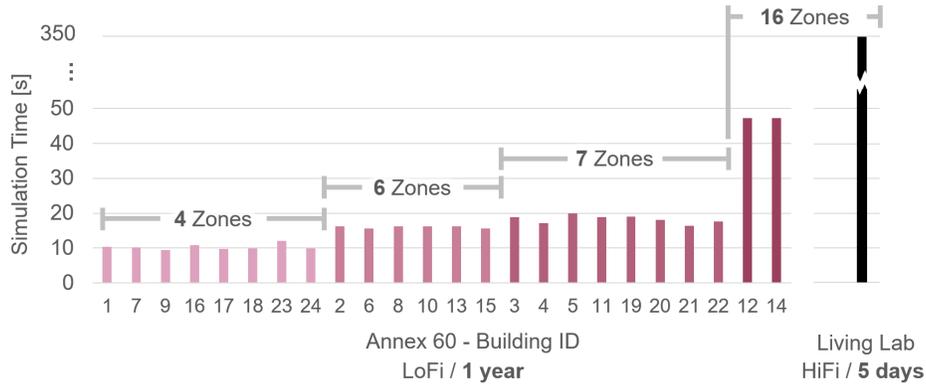

Figure 14: Runtime scaling summary: LoFi annual Annex 60 simulations vs HiFi 5-day Living Lab simulation: quantifying the accuracy-speed trade-off motivating fidelity choice.

Table 3: Relative runtime per simulated day vs zoning and model fidelity.

Zoning	LoFi (4)	LoFi (6)	LoFi (7)	LoFi (16)	HiFi (16)
Cost per simulated day	×1.0	×1.6	×1.8	×4.7	×2555

From an application perspective, the presented modeling framework JanusBM provides a practical guideline for modeling accuracy:

- ideal-load LoFi is sufficient for energy-centric studies where the decision variables depend mainly on seasonal or annual heating demand, or on average indoor temperature regulation, and where distribution constraints do not affect the results (e.g., floor-level scenario screening, studies of building renovation comparison, and long-term sector coupling planning);

- when temperature or power trajectories are affected by hydronic loop dynamics, hydronic HiFi becomes necessary when required heat and deliverable heat differ due to distribution and control constraints, preferably with a zoning that maintains the loop-to-zone mapping. This includes analysis, optimization, and control

14

design for short-term dynamics, such as peak loads and ramp rates, or studies requiring flow-related heat transfer, or loop-level performance assessments.

6   Conclusion and Outlook

This paper presents the dual-fidelity, multi-zone white-box building modeling framework JanusBM that enables consistent analysis across annual and transient time scales by coupling a HiFi hydronic model with a LoFi ideal-load surrogate, both derived from the same zoning and envelope description generated by the topology-driven RoomFlex6D modeling tool. To ensure physical applicability under complementary data availability, the two-stage hybrid validation and iterative RoomFlex6D–FMU–Python calibration pipeline using two complementary dataset is introduced. **On the energy scale**, the LoFi models achieve a high degree of consistency with Annex 60 in terms of annual heating demand at both building and neighborhood aggregation levels, while peak heating load is systematically underestimated in high-demand buildings. **On the dynamic scale**, the proposed iterative calibration workflow improves the transient behavior of loop-level return water temperature during active circulation periods, achieving uncompensated improvements at the zone temperature level within the same calibration window. Finally, the framework quantitatively characterizes the accuracy–speed trade-off: LoFi simulation cost increases quasi-linearly with zoning, while introducing explicit hydronic dynamics at comparable zoning led to orders-of-magnitude higher cost, motivating a fidelity choice guided by whether required and deliverable heat are biased due to thermal distribution and control constraints.

Several limitations point the way for future work: First, the demonstrated peak load deviation in energy-scale LoFi validation motivates extending the surrogate to include controlled transient capability when peak metrics are observed. Then, the current HiFi calibration intentionally isolates hydronic impacts by tuning UFH-related parameters while keeping building envelope parameters fixed to avoid non-identifiability, and additional constraints will be required when extending the calibration to include envelope properties. In addition, remaining structured residuals at the zone level indicate structural uncertainties beyond hydronic heat exchange, such as sensor placement, interior modifications, and air leakage, suggesting that future model extensions should account for these effects, or incorporate them with suitable grey-box representations. Finally in future work on model generation, the CityGML, IFC and gbXML source can be used as upstream providers of building geometry and attributes [Benner, 2026], which would then be compiled into the RoomFlex6D topology and parameter tables to maintain both interoperability and the efficiency required for large-scale scenario exploration.

Acknowledgments

References

Juan Gea-Bermúdez, Ida Græsted Jensen, Marie Münster, Matti Koivisto, Jon Gustav Kirkerud, Yi kuang Chen, and Hans Ravn. The role of sector coupling in the green transition: A least-cost energy system development in northern-central europe towards 2050. *Applied Energy*, 289:116685, 2021. doi:10.1016/j.apenergy.2021.116685. URL https://www.sciencedirect.com/science/article/pii/S0306261921002130.

Rui Jing, Yue Zhou, and Jianzhong Wu. Electrification with flexibility towards local energy decarbonization. *Advances in Applied Energy*, 5:100088, 2022. ISSN 2666-7924. doi:https://doi.org/10.1016/j.adapen.2022.100088. URL https://www.sciencedirect.com/science/article/pii/S2666792422000063.

Arthur Rinaldi, Selin Yilmaz, Martin K. Patel, and David Parra. What adds more flexibility? an energy system analysis of storage, demand-side response, heating electrification, and distribution reinforcement. *Renewable and Sustainable Energy Reviews*, 167:112696, 2022. ISSN 1364-0321. doi:https://doi.org/10.1016/j.rser.2022.112696. URL https://www.sciencedirect.com/science/article/pii/S1364032122005858.

Ján Drgoňa, Javier Arroyo, Iago Cupeiro Figueroa, David Blum, Krzysztof Arendt, Donghun Kim, Enric Perarnau Ollé, Juraj Oravec, Michael Wetter, Draguna L. Vrabie, and Lieve Helsen. All you need to know about model predictive control for buildings. *Annual Reviews in Control*, 50:190–232, 2020. ISSN 1367-5788. doi:https://doi.org/10.1016/j.arcontrol.2020.09.001. URL https://www.sciencedirect.com/science/article/pii/S1367578820300584.

Arthur Rinaldi, Martin Christoph Soini, Kai Streicher, Martin K. Patel, and David Parra. Decarbonising heat with optimal pv and storage investments: A detailed sector coupling modelling framework with flexible heat pump

operation. *Applied Energy*, 282:116110, 2021. ISSN 0306-2619. doi:https://doi.org/10.1016/j.apenergy.2020.116110. URL https://www.sciencedirect.com/science/article/pii/S0306261920315282.

Ofelia Vera-Piazzini and Massimiliano Scarpa. Building energy model calibration: A review of the state of the art in approaches, methods, and tools. *Journal of Building Engineering*, 86:108287, 2024. ISSN 2352-7102. doi:https://doi.org/10.1016/j.jobe.2023.108287. URL https://www.sciencedirect.com/science/article/pii/S2352710223024701.

Drury B. Crawley, Linda K. Lawrie, Frederick C. Winkelmann, W.F. Buhl, Y.Joe Huang, Curtis O. Pedersen, Richard K. Strand, Richard J. Liesen, Daniel E. Fisher, Michael J. Witte, and Jason Glazer. Energyplus: creating a new-generation building energy simulation program. *Energy and Buildings*, 33(4):319–331, 2001. ISSN 0378-7788. doi:https://doi.org/10.1016/S0378-7788(00)00114-6. URL https://www.sciencedirect.com/science/article/pii/S0378778800001146. Special Issue: BUILDING SIMULATION'99.

Sanford A Klein. Trnsys-a transient system simulation program. *University of Wisconsin-Madison, Engineering Experiment Station Report*, pages 38–12, 1988.

David Blum, Javier Arroyo, Sen Huang, Ján Drgoňa, Filip Jorissen, Harald Walnum, Yan Chen, Kyle Benne, Draguna Vrabie, Michael Wetter, and Lieve Helsen. Building optimization testing framework (boptest) for simulation-based benchmarking of control strategies in buildings. *Journal of Building Performance Simulation*, 14:586–610, 10 2021. doi:10.1080/19401493.2021.1986574.

Saptarshi Bhattacharya, Yan Chen, Sen Huang, and Draguna Vrabie. A learning-based time-efficient framework for building energy performance evaluation. *Energy and Buildings*, 228:110411, 2020. ISSN 0378-7788. doi:https://doi.org/10.1016/j.enbuild.2020.110411. URL https://www.sciencedirect.com/science/article/pii/S0378778820310926.

ISO. Energy performance of buildings — energy needs for heating and cooling, internal temperatures and sensible and latent heat loads, 2017. URL https://www.iso.org/standard/65696.html.

Kun Zhang, David Blum, Hwakong Cheng, Gwelen Paliaga, Michael Wetter, and Jessica Granderson. Estimating ashrae guideline 36 energy savings for multi-zone variable air volume systems using spawn of energyplus. *Journal of Building Performance Simulation*, 15(2):215–236, 2022. doi:10.1080/19401493.2021.2021286. URL https://doi.org/10.1080/19401493.2021.2021286.

David Jansen, Veronika Richter, Diego Cordoba Lopez, Philipp Mehrfeld, Jérôme Frisch, Dirk Müller, and Christoph van Treeck. Examination of reduced order building models with different zoning strategies to simulate larger non-residential buildings based on bim as single source of truth. *Modelica Conferences*, 2021. doi:https://doi.org/10.3384/ecp21181665. URL https://ecp.ep.liu.se/index.php/modelica/article/view/246.

Minjae Shin and Jeff S. Haberl. A procedure for automating thermal zoning for building energy simulation. *Journal of Building Engineering*, 46:103780, 2022. ISSN 2352-7102. doi:https://doi.org/10.1016/j.jobe.2021.103780. URL https://www.sciencedirect.com/science/article/pii/S2352710221016387.

Filip Belic, Dražen Slišković, and Zeljko Hocenski. Detailed thermodynamic modeling of multi-zone buildings with resistive-capacitive method. *Energies*, 14:7051, 10 2021. doi:10.3390/en14217051.

Michael Wetter, Jianjun Hu, Anand Krishnan Prakash, Paul Ehrlich, and Gabe Fierro. Modelica-json: Transforming energy models to digitize the control delivery process. 09 2021a. doi:10.26868/25222708.2021.30141.

M Wetter, C van Treeck, L Helsen, A Maccarini, D Saelens, D Robinson, and G Schweiger. Ibpsa project 1: Bim/gis and modelica framework for building and community energy system design and operation – ongoing developments, lessons learned and challenges. *IOP Conference Series: Earth and Environmental Science*, 323(1):012114, aug 2019. doi:10.1088/1755-1315/323/1/012114. URL https://doi.org/10.1088/1755-1315/323/1/012114.

Michael Wetter, Kyle Benne, and Baptiste Ravache. Software architecture and implementation of modelica buildings library coupling for spawn of energyplus. *Modelica Conferences*, pages 325–334, 09 2021b. doi:10.3384/ecp21181325.

Felix Langner, Jovana Kovačević, Luigi Spatafora, Stefan Dietze, Simon Waczowicz, Hüseyin K. Çakmak, Jörg Matthes, and Veit Hagenmeyer. Experimental evaluation of model predictive control and fuzzy logic control for demand response in buildings. *Applied Energy*, 401:126666, 2025. ISSN 0306-2619. doi:https://doi.org/10.1016/j.apenergy.2025.126666. URL https://www.sciencedirect.com/science/article/pii/S0306261925013960.

K.E. Anders Ohlsson and Thomas Olofsson. Benchmarking the practice of validation and uncertainty analysis of building energy models. *Renewable and Sustainable Energy Reviews*, 142:110842, 2021. ISSN 1364-0321. doi:https://doi.org/10.1016/j.rser.2021.110842. URL https://www.sciencedirect.com/science/article/pii/S1364032121001362.

Michael Georgescu and Igor Mezić. Building energy modeling: A systematic approach to zoning and model reduction using koopman mode analysis. *Energy and Buildings*, 86:794–802, 2015. ISSN 0378-7788. doi:https://doi.org/10.1016/j.enbuild.2014.10.046. URL https://www.sciencedirect.com/science/article/pii/S0378778814008913.

Yixing Chen and Tianzhen Hong. Impacts of building geometry modeling methods on the simulation results of urban building energy models. *Applied Energy*, 215:717–735, 2018. ISSN 0306-2619. doi:https://doi.org/10.1016/j.apenergy.2018.02.073. URL https://www.sciencedirect.com/science/article/pii/S0306261918301983.

ASHRAE. American Society of Heating, Refrigerating and Air-Conditioning Engineers, 2026. URL https://www.ashrae.org/.

Minjae Shin and Jeff S. Haberl. Thermal zoning for building hvac design and energy simulation: A literature review. *Energy and Buildings*, 203:109429, 2019. ISSN 0378-7788. doi:https://doi.org/10.1016/j.enbuild.2019.109429. URL https://www.sciencedirect.com/science/article/pii/S0378778819312769.

Maximilian Mork, Eziama Ubachukwu, Jakob Benz, Philipp Althaus, André Xhonneux, and Dirk Müller. Alice2modelica - automated building model generation for building control and simulation. In *2024 Open Source Modelling and Simulation of Energy Systems (OSMSES)*, pages 1–7, 2024. doi:10.1109/OSMSES62085.2024.10668986.

Peter Remmen, Moritz Lauster, Michael Mans, Marcus Fuchs, Tanja Osterhage, and Dirk Müller. Teaser: an open tool for urban energy modelling of building stocks. *Journal of Building Performance Simulation*, 11(1):84–98, 2018. doi:10.1080/19401493.2017.1283539. URL https://doi.org/10.1080/19401493.2017.1283539.

Yixing Chen, Tianzhen Hong, and Mary Ann Piette. Automatic generation and simulation of urban building energy models based on city datasets for city-scale building retrofit analysis. *Applied Energy*, 205:323–335, 2017. ISSN 0306-2619. doi:https://doi.org/10.1016/j.apenergy.2017.07.128. URL https://www.sciencedirect.com/science/article/pii/S0306261917310024.

F. Johari, G. Peronato, P. Sadeghian, X. Zhao, and J. Widén. Urban building energy modeling: State of the art and future prospects. *Renewable and Sustainable Energy Reviews*, 128:109902, 2020. ISSN 1364-0321. doi:https://doi.org/10.1016/j.rser.2020.109902. URL https://www.sciencedirect.com/science/article/pii/S1364032120301933.

International Building Performance Simulation Association. IBPSA, 2026. URL https://www.ibpsa.org/.

Filip Jorissen, Glenn Reynders, Ruben Baetens, Damien Picard, Dirk Saelens, and Lieve Helsen. Implementation and Verification of the IDEAS Building Energy Simulation Library. *Journal of Building Performance Simulation*, 11:669–688, 2018. doi:10.1080/19401493.2018.1428361.

Laura Maier, David Jansen, Fabian Wüllhorst, Martin Kremer, Alexander Kümpel, Tobias Blacha, and Dirk Müller. Aixlib: an open-source modelica library for compound building energy systems from component to district level with automated quality management. *Journal of Building Performance Simulation*, 17(2):196–219, 2023. doi:10.1080/19401493.2023.2250521. URL https://doi.org/10.1080/19401493.2023.2250521.

Christoph Nytsch-Geusen, Jörg Huber, Manuel Ljubijankic, and Jörg Rädler. Modelica buildingsystems - eine modellbibliothek zur simulation komplexer energietechnischer gebäudesysteme. *BAUSIM 2012 IBPSA Germany*, 2012. doi:10.1002/bapi.201310045.

Michael Wetter, Wangda Zuo, Thierry S. Nouidui, and Xiufeng Pang. Modelica Buildings library. *Journal of Building Performance Simulation*, 7(4):253–270, 2014. doi:10.1080/19401493.2013.765506. URL https://doi.org/10.1080/19401493.2013.765506.

Michael Wetter and Christoph van Treeck. *IEA EBC Annex 60: New Generation Computing Tools for Building and Community Energy Systems*. September 2017. ISBN 978-0-692-89748-5. URL https://www.iea-annex60.org/pubs.html.

Italo Aldo Campodonico-Avendano, Silvia Erba, Panayiotis Papadopoulos, Salvatore Carlucci, Antonio Luparelli, Amedeo Ingrosso, Greta Tresoldi, Muhammad Salman Shahid, Frederic Wurtz, Benoit Delinchant, Per Martin Leinan, Stefano Cera, Peter Riederer, Runar Solli, Amin Moazami, and Mohammadreza Aghaei. Collectief dataset: A high-resolution indoor environmental dataset from european buildings across diverse climates supporting thermal, air-quality, and visual-comfort assessments. *Data in Brief*, page 112486, 2026. ISSN 2352-3409. doi:https://doi.org/10.1016/j.dib.2026.112486. URL https://www.sciencedirect.com/science/article/pii/S2352340926000399.

Erfan Tajalli-Ardekani, Felix Langner, Jovana Kovačević, Stefan Dietze, Malte Holzhaeuer, Tobias Moser, Abdulrahman Seddiqi, Luigi Spatafora, Simon Waczowicz, Hüseyin Kemâl Çakmak, Jörg Matthes, and Veit Hagenmeyer. Experimental data of three identical residential buildings controlled with different controllers for heating demand response, October 2025. URL https://doi.org/10.5281/zenodo.17455075.

Louise Wright and Stuart Davidson. How to tell the difference between a model and a digital twin. *Advanced Modeling and Simulation in Engineering Sciences*, 7, 12 2020. doi:10.1186/s40323-020-00147-4.

Open Geospatial Consortium. CityGML, 2026. URL `https://www.ogc.org/standards/citygml/`.

buildingSMART. IFC, 2026. URL `https://www.buildingsmart.org/standards/bsi-standards/industry-foundation-classes/`.

gbXML. About gbXML, 2026. URL `https://www.gbxml.org/About_GreenBuildingXML_gbXML`.

Na Luo, Marco Pritoni, and Tianzhen Hong. An overview of data tools for representing and managing building information and performance data. *Renewable and Sustainable Energy Reviews*, 147:111224, 2021. ISSN 1364-0321. doi:https://doi.org/10.1016/j.rser.2021.111224. URL `https://www.sciencedirect.com/science/article/pii/S1364032121005116`.

Michael Wetter. Multizone airflow model in modelica. *Proceedings of the 5th International Modelica Conference*, 2: 431–440, 01 2006.

Veit Hagenmeyer, Hüseyin K. Çakmak, Clemens Düpmeier, Timm Faulwasser, Jörg Isele, Hubert B. Keller, et al. Information and communication technology in energy lab 2.0: Smart energies system simulation and control center with an open-street-map-based power flow simulation example. *Energy Technology*, 4(1):145–162, January 2016. doi:10.1002/ente.201500304. URL `https://doi.org/10.1002/ente.201500304`.

OSMF. OSM, 2026. URL `https://www.openstreetmap.org/`.

Yiqun Pan, Mingya Zhu, Yan Lv, Yikun Yang, Yumin Liang, Ruxin Yin, Yiting Yang, Xiaoyu Jia, Xi Wang, Fei Zeng, Seng Huang, Danlin Hou, Lei Xu, Rongxin Yin, and Xiaolei Yuan. Building energy simulation and its application for building performance optimization: A review of methods, tools, and case studies. *Advances in Applied Energy*, 10:100135, 2023. ISSN 2666-7924. doi:https://doi.org/10.1016/j.adapen.2023.100135. URL `https://www.sciencedirect.com/science/article/pii/S2666792423000148`.

Friedrich Wiegel, Jan Wachter, Michael Kyesswa, Ralf Mikut, Simon Waczowicz, and Veit Hagenmeyer. Smart energy system control laboratory – a fully-automated and user-oriented research infrastructure for controlling and operating smart energy systems. *at - Automatisierungstechnik*, 70(12):1116–1133, 2022. doi:doi:10.1515/auto-2022-0018. URL `https://doi.org/10.1515/auto-2022-0018`.

Joachim Benner. GML-ToolboxApp, 2026. URL `https://github.com/KIT-IAI/GMLToolbox/`.

A   RoomFlex6D GUI Details

Figure 15 illustrates the Python-based GUI used to construct the zone topology and to edit the face-level parameter table exported by RoomFlex6D. Users need to specify adjacency relationships and boundary conditions in a structured manner before generating the Modelica model.
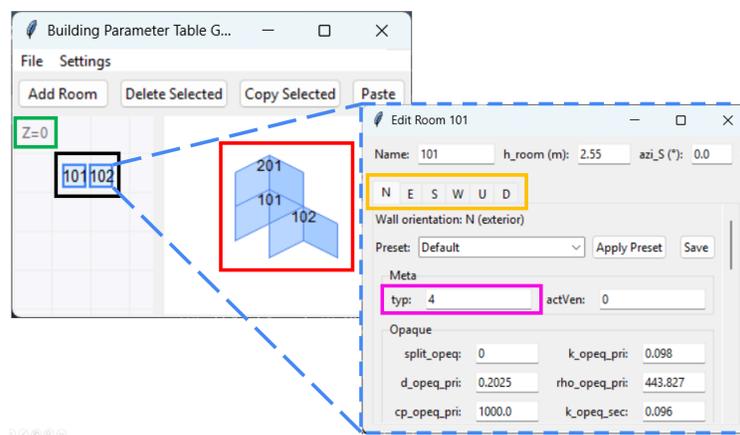

Figure 15: RoomFlex6D GUI—Further explanation for Figure 2 with colored blocks.

A.1   Topology Configuration

The thermal zone topology configuration page is shown on the left side of the Figure 15, where each zone is represented by an orthogonal volume (marked with a black box). The corresponding layer number of the thermal zone is displayed in the upper-left corner of the configuration page (marked with green boxes), allowing users to modify the relative placement of zones via drag-and-drop. A lightweight 3D preview area marked in red is a 3D preview, provides immediate feedback on the configured topology.

A.2   Zone Parameter Dialog

By double-clicking a single thermal zone, users can configure zone-level parameters and face-level properties, as shown in the configuration window on the right side of the Figure 15. The dialog exposes the six directional faces N/E/S/W/U/D (marked with an orange box), which are the interfaces for thermal interaction with adjacent zones or the exterior.

A.3   Face-level Boundary Specification

The boundary conditions of each face are specified as one of five internal/external boundary types connecting to a unique adjacent area (marked with a pink box), as defined in Table 1, which distinguishes non-closable openings, operable windows, operable doors, opaque exterior faces, and ground-coupled exterior faces.

A.4   Export and Model Generation

The GUI exports a structured face-level parameter table (shown in Table 2 or Table 4 in Appendix D) that is used for the subsequent automatic Modelica model generation (Figure 3 for an Example). Therefore, updating the topology or parameters only requires regenerating the table and re-instantiating the Modelica model, supporting fast and repeatable zoning and structural modifications.

B   HiFi Hydronic Subsystem Specification and Configuration in RoomFlex6D

RoomFlex6D allows users to configure a building hydronic heating subsystem and couple it to the generated multi-zone envelope structure via the structured parameter table in Table 2 or Table 4 in Appendix D. The HiFi model addresses transient and control-relevant behaviors by explicitly representing both the heater-side thermodynamics and the network-side hydronic dynamics.

B.1   Heater Models

Radiator and UFH systems are supported as typical configurations in the proposed framework JanusBM. In both cases, heaters are considered as dynamic components with finite heat capacity. The heat exchange of a building is modeled through convection and radiation to the zone air, or through embedded constructions in the case of UFH, which enables delayed heat release effects caused by construction thermal inertia.

B.2   Hydronic Network Components

The network comprises supply and return pipes, hydraulic flow resistances, circulation pumps, and control elements including control valves, bypass, and mixing valves. Mass flow rate and pressure drop are computed based on conservation laws by the Modelica algorithm. Meanwhile, the state variables include the supply water temperature and the opening state of control valves, which can be driven by zone-temperature feedback within a closed-loop control, or prescribed as external inputs to represent scenario- or controller-dependent operation.

B.3   Configuration Parameters by the GUI

The GUI supports configuring (a) heater type (radiator vs UFH), (b) the loop structure and the sequence through which hot water flows across heaters, and (c) UFH-specific parameters such as embedding depth and other settings that influence heating dynamics. These settings are exported in the hydronic-related columns of Table 2 or Table 4 in Appendix D, including heater type, loop identifier, flow order, and control mapping, providing a reproducible interface between topology definition and Modelica model generation.

C   Background Annex 60 Benchmark

The studied Annex 60 neighborhood (shown in Figure 16) consists of 24 buildings with five building types: detached houses (D), semi-detached buildings (S), terraced buildings (T), apartment (A), and office buildings (O), where two different thermal insulation stages of building envelopes were represented by number 1 (better) and 2 (worse). Five modeling teams (TUe, EDF, KUL, ULg and RWTH Aachen) in Annex 60 benchmark were required to use the Modelica modeling language and adhere the same boundary conditions for building simulation, including:

- a reference climate in Uccle, Belgium,

- a constant heating setpoint of 21°C,

- zero internal heat gain,

- no shading, and

- a baseline-defined heating season (from October 10th to May 19th of the following year).
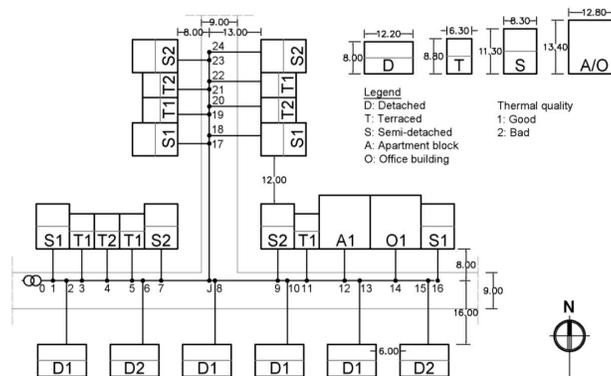

Figure 16: Annex 60 district layout used for LoFi validation (24 buildings, five typologies).

Nevertheless, these teams had the flexibility to choose their own modeling methods regarding model complexity and libraries. Zoning strategies of these participants ranged from single-zone representations to multi-zone configurations with one or more zones per floor, depending on building type and facade orientation. Some participants explicitly simulated heat transfer between adjacent buildings, while others neglected inter-building thermal coupling. This diversity in modeling choices resulted in significant differences in the predicted annual and peak heating demands, as described in the final report of Annex 60.

Despite these unified boundary conditions, Annex 60 was reported as participant result ranges rather than a unique reference value. Common sources of variation include (but are not limited to):

- **inter-building coupling assumptions:** whether and how thermal interactions with adjacent buildings or shared boundaries are represented;

- **envelope and boundary idealizations:** e.g., discretization of envelope elements, treatment of thermal bridges, and the handling of ground-coupled surfaces and adjacent areas; and

- **numerical settings and solver-related effects:** the settings of time step and convergence tolerances can influence peak metrics more strongly than annual energy metrics.

D   Parameter Table Structured by RoomFlex6D Modeling Tool

Table 4: Full version of face-level parameter-table schema generated by RoomFlex6D for model generation.

Parameter	Symbol	Description	Type	Example
Primary zone ID	z_{pri}	Index of the current thermal zone	Integer	101
Face orientation	ori	Face direction (1–6 corresponding to N/E/S/W/U/D)	Integer	2
Adjacent zone ID	z_{adj}	Identifier of the neighboring zone (0 denotes an external boundary)	Integer	102
Boundary type	typ	(See Table 1)	Integer	3
Ventilation flag	$actVen$	Indicates whether active ventilation is applied on this face	Integer	0
Split flag	$split$	Indicates whether UFH is NOT installed	Integer	1
Thermal parameter	$k, d, \rho, c_p, A_{\text{opeq,fra,gla,open,RS}}$	Thermodynamic parameters and area of walls, door and window frames, glass, openings, and roller shutters.	Real	0.70
Hydronic	$heat_{\text{type,loop,order,ctrl}}$	Types, numbering, sequence, and control schemes of heating loops.	Integer	1
Building orientation	azi_S	The angle in degrees between the user-defined south-facing surface and true south, used to calculate the impact of solar radiation and wind speed.	Real	2.3