


INNOVATION
– PRESS –

Energy Use

Energy Use 3 (2026):100071

Contents lists available at INNOVATION PRESS

From Automation to Trustworthy Human-AI Collaboration in AI-Driven Building Energy Modeling

Haoyi Zhu¹ and Yueer He^{1,*}
¹School of Architecture and Urban Planning, Shenzhen University & State Key Laboratory of Subtropical Building and Urban Science, Shenzhen University, Shenzhen 518060, China

*Correspondence: heyueer@szu.edu.cn

Received: September 6, 2026; Accepted: September 20, 2026; Published Online: September 23, 2026; <https://doi.org/10.59717/ipj.energy-use.2026.100071>

© 2026 The Author(s). This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Zhu H. and He Y. (2026). From Automation to Trustworthy Human-AI Collaboration in AI-Driven Building Energy Modeling. *Energy Use* 2:100071.

Dear Editor,

Building energy modeling is undergoing a fundamental paradigm shift from conventional static, physics-based simulations toward intelligent, agentic workflows. By integrating large language models, model context protocol, and agent-to-agent protocol, AI-driven building energy modeling has the potential to enable real-time multidisciplinary interaction, automated tool invocation, and continuous closed-loop validation across heterogeneous software platforms. In this perspective, we argue that the significance of AI-driven building energy modeling lies not in replacing simulation experts, but in reorganizing how evidence is generated, communicated, and verified across design. Three boundaries will determine whether this transition becomes practically credible: the weak spatial grounding and lack of interactive design mechanisms, the need for auditable human-in-the-loop verification to control error propagation, and the computational energy and carbon cost and token overhead of persistent multi-agent inference. We therefore propose a research agenda centered on interactive design-simulation-feedback loops, traceable simulation justification, and explicit accounting of computational energy and token consumption, moving AI-driven building energy modeling from isolated automation toward trustworthy human-AI collaboration for evidence-based design.

BUILDING ENERGY MODELING IS ENTERING A NEW PARADIGM

The buildings and construction sector accounts for approximately 40% of global final energy demand, making it a critical target for decarbonization.¹ Driven by green building rating systems, energy efficiency codes, and low-carbon design paradigms, building energy modeling (BEM) has become an important means of testing energy-efficiency strategies, forecasting performance, and informing design decisions. Yet conventional BEM remain fragmented and sequential, with architects and engineers working in silos. This linear "design-simulate-revise" workflow limits early-stage simulation integration and cannot support rapid iteration or multidisciplinary collaboration, driving a shift toward real-time, agentic, and integrated BEM workflows.

Persistent challenges in BEM

Conventional BEM workflows impose several forms of friction (Figure 1).² First, practitioners must coordinate specialized knowledge across building physics, HVAC systems, thermodynamics, controls, and architectural design. Second, high-fidelity models depend on detailed and internally consistent inputs for geometry, envelopes, schedules, systems, and weather. Third, geometry generation, parameter mapping, model checking, and repeated simulation runs require substantial manual effort. Finally, limited interoperability among BIM platforms, parametric design tools, and simulation engines introduces recurring data translation and version-control problems. Beyond

software inconvenience, these frictions constrain BEM efficiency and scalability and gate energy evidence from early design. A useful AI intervention must therefore reduce coordination friction without obscuring the physical assumptions on which simulation results depend.

What agentic workflows change

Recent advances in generative AI have extended large language models (LLMs) from static content generation toward agentic systems that can plan tasks, call tools, and respond to execution feedback. Open protocols such as the Model Context Protocol³ (MCP) resolve cross-platform interoperability barriers, enabling agents to invoke domain-specific tools, share context continuously, and orchestrate complex tasks. An agentic workflow is a higher-level automated process built on such agentic systems, coordinating their planning, tool use, and feedback-driven self-correction into end-to-end task execution with continuous context exchange. This paradigm has been explored across diverse domains, shifting workflows from manual coordination and rigid scripts toward adaptive, intelligent, and collaborative environments, with representative end-to-end deployments already demonstrated in software engineering and beyond.⁴

This evolution creates new opportunities for BEM. In this perspective, AI-driven BEM refers specifically to the application of agentic workflows to building energy modeling, namely the orchestration of modeling and simulation through agents. Conventional BEM require architects and engineers to exchange models and parameters across platforms, manually collecting and validating data through prolonged cycles that delay feedback. In contrast, AI-driven BEM allows multiple agents to interface via MCP with BIM environments, parametric modeling tools, and energy simulation engines, autonomously executing model generation, data acquisition, simulation runs, and result interpretation. This streams performance feedback continuously to the design interface, forming an integrated collaborative loop among architects, engineers, and AI. The shift from retrospective "post-design verification" to real-time "continuous in-design validation" embeds energy modeling into early schematic generation, supporting evidence-based design (EBD).

EVOLUTION OF AGENTIC WORKFLOWS IN AI-DRIVEN BEM

Although AI-driven BEM remains at an early stage, but recent work traces a clear evolution from static multi-agent collaboration to dynamic, decentralized networks. In 2025, Zhang et al. proposed an LLM-based workflow with four specialized agents that automatically converted unstructured architectural descriptions into error-free EnergyPlus input data files, improving modeling efficiency and robustness.⁵ As an early implementation of AI-driven BEM, the framework adopted predefined agent roles and interactions, with tool access coordinated through customized scripts and APIs. This structured architecture provided a reliable foundation for task-oriented multi-

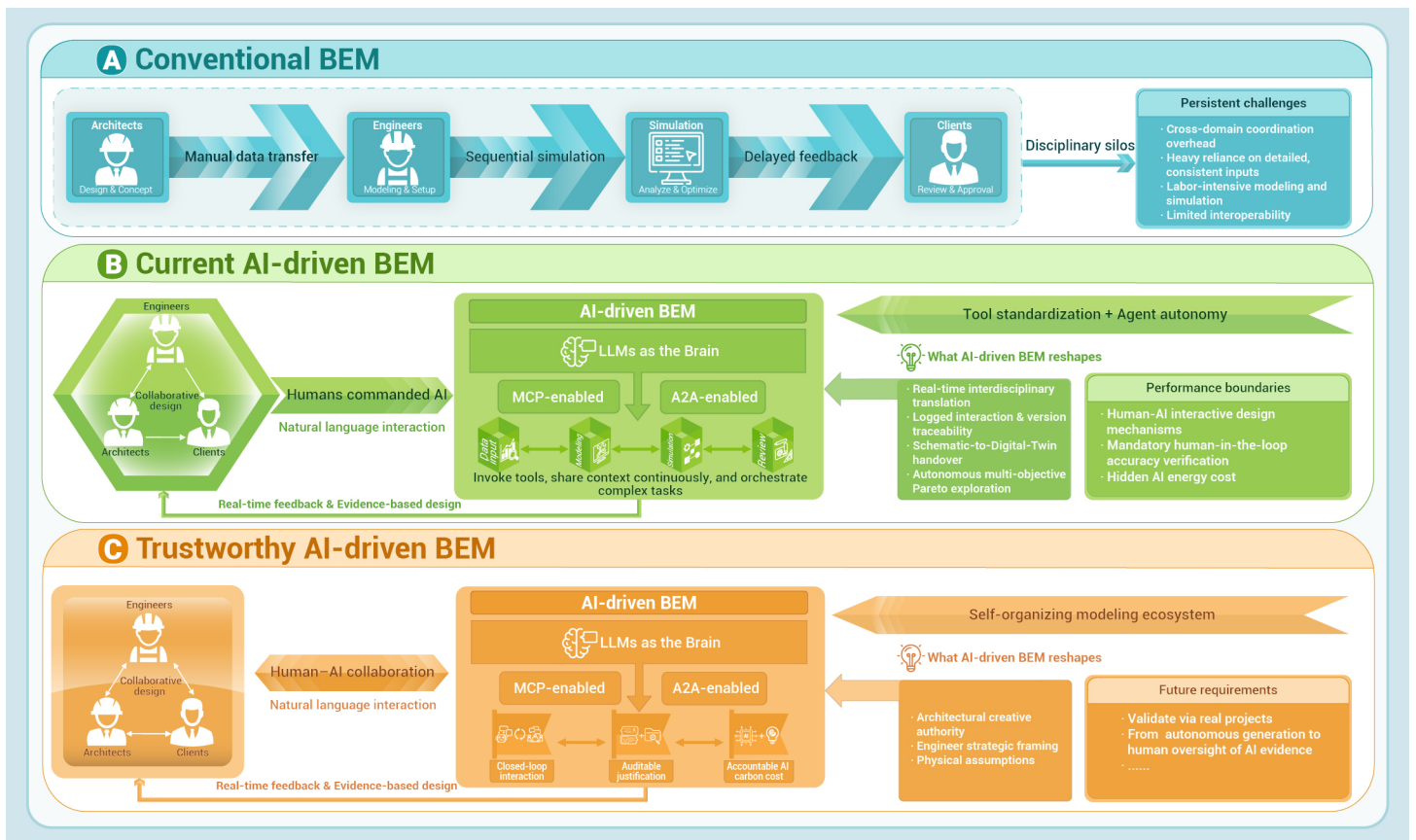


Figure 1. Evolution of BEM Workflows: From Static Manual Processes to AI-Driven Agentic Automation and Trustworthy Futures

agent collaboration, while also highlighting opportunities for subsequent frameworks to support more flexible agent coordination, standardized tool integration, and cross-platform scalability.

MCP marks a milestone for AI-driven BEM, advancing agentic workflows from the first step of specialized agent collaboration to the second step of standardized tool interconnection. Frameworks such as EnergyPlus-MCP⁶ wrap the EnergyPlus engine as a standardized service, enabling agents to manage models, edit parameters, run simulations, and analyze results via natural language prompts while preserving the physical rigor and reproducibility of EnergyPlus. Li et al. also further demonstrated two complementary paradigms for MCP-enabled AI-driven BEM: conversational integration, where a single agent iteratively collaborates with users for exploratory tasks; and agentic workflow integration, where multiple specialized agents autonomously coordinate for systematic automation.⁷ Both paradigms improve efficiency while maintaining transparency through visible tool invocations, with human oversight scaled to task needs.

Agent-to-Agent⁸ (A2A) protocol complements Anthropic's MCP. Released merely a few months after MCP, it resolves a limitation of the second step: while MCP broke down interoperability silos and standardized tool interconnection, agent coordination remained bound to predefined roles and static orchestration. A2A protocol enables agents to autonomously configure collaborations, allocate tasks, and synchronize context states based on real-time demands. Although a recent BEM case study offers preliminary evidence for dynamic multi-agent networks,⁹ the evidence base remains limited and doesn't yet demonstrate that A2A-enabled workflows are widely validated in building energy simulation. The third step therefore represents a potential trajectory toward more adaptive, decentralized agent collaboration, whose scalability, reliability, and practical value still require validation through further practical projects. Whether such autonomy is realized remains contingent on the trustworthy human-AI collaboration mechanisms developed below. For collaborative design, the main benefit is likely to lower coordination latency, not reduced headcount.

For collaborative design, the most consequential benefits are likely to come from reducing coordination latency rather than simply reducing head-

count or labor. As shown in Figure 1, AI-driven BEM reshape collaboration among architects, engineers, and clients in four ways. First, AI serves as a real-time interdisciplinary translator that aligns client goals, architectural intent, and engineering standards, cutting communication costs. Second, interaction logging and version traceability enable early-stage risk prediction, clarify liability allocation, and shorten schematic iterations. Third, high-fidelity schematic models transition directly to operational Digital Twins, streamlining the design-to-operations handover and reducing long-term commissioning and facility management costs. Fourth, AI-driven BEM explores multi-objective design spaces to reveal Pareto-optimal trade-offs among capital cost, energy, and aesthetics, cutting manual iteration.

THREE PERFORMANCE BOUNDARIES FOR TRUSTWORTHY AI-DRIVEN BEM

This trajectory described above shows that research on agentic workflows in AI-driven BEM is shifting from automating isolated tasks to orchestrating entire modeling processes. Yet building performance simulation remains physics-based, and autonomous AI agents do not eliminate the uncertainties inherent in conventional BEM. While current research focuses mainly on computational and workflow efficiency, discussion of the operational boundaries and practical applicability of AI in real-world engineering remains limited. The priority must therefore shift from asking "how to automate model generation" to asking "how to establish trustworthy, explainable, and sustainable modeling workflows". From this perspective, we identify three critical performance boundaries that require systematic attention.

Weak spatial grounding and lack of interactive design mechanisms

Interactive design mechanisms between AI-driven BEM and human practitioners remain underdeveloped. Although multimodal LLMs can read building images and generate coarse 3D geometry, BEM requires high-precision, editable models. Combined with the limited availability of high-fidelity 3D architectural datasets, current LLMs haven't yet attained a reliable, physics-grounded understanding of spatial topology for valid energy modeling, a capability they may develop as these limitations are addressed.¹⁰ In practice,

architects describe spatial form and program in natural language, but AI-driven BEM struggles to interpret the physical logic embedded in these qualitative intents. The resulting gap forces lengthy manual iteration to produce valid energy models and limits the system's use of conversational context to propose alternatives that balance engineering feasibility with architectural aesthetics.

Simulation Accuracy Still Mandates Human-in-the-Loop Verification

Despite these advances, human-in-the-loop verification remains essential. Building energy modeling is physics-based, and accuracy depends on careful calibration of boundary conditions such as HVAC configurations, occupancy schedules, and internal thermal loads. Because errors in BEM propagate, small input discrepancies compound across simulation steps and can lead to large prediction errors. Architects without training in building physics need expert help to evaluate such errors. Moreover, fully automated model generation by AI agents may hide modeling flaws within a black-box loop, amplifying downstream design and decision-making risks. For example, an automated agent may silently mis-map an occupancy schedule; EnergyPlus can still run without error, yet the output may over- or under-state loads by tens of percent—an error only an engineer's review can catch.

Is AI-driven BEM truly energy-efficient?

While literature emphasizes AI's contribution to reducing building energy use and design labor, the energy cost of AI itself is rarely examined. Training LLMs and running AI agents require high-performance computing and data centers, consuming significant electricity and generating carbon emissions. Token consumption, whose volume per agentic cycle scales with model size and task complexity, offers a more immediate and quantifiable proxy for this cost. Whether the computational energy embodied in persistent agentic workflows could offset a meaningful fraction of the operational energy savings achieved through optimization remains an open empirical question. This frames architectural energy optimization as a trade-off between building operational energy and computational energy expenditure.

A ROADMAP FOR TRUSTWORTHY AI-DRIVEN BEM

Although AI-driven BEM are maturing rapidly, the expertise and decision-making authority of architects and engineers remain irreplaceable. Architects hold subjective judgments about spatial experience, cultural meaning, and aesthetics that cannot be fully parameterized, while every project's unique microclimate, site conditions, and regulations require engineers to set strategic frameworks in advance. Also, current AI-driven BEM remains automation-oriented. Although it extends conventional BEM through stronger interoperability and agent autonomy, it still operates as a human-commands-AI workflow. Consequently, the future AI-driven BEM should be trustworthy AI-driven BEM, a form of human-AI collaboration in which human insight guides strategy, supervision, and value judgments.

Looking ahead, trustworthy AI-driven BEM should advance through three interconnected developments rather than pursue autonomous generation alone (Figure 1). First, to redress the weak spatial grounding and lack of interactive design mechanisms, trustworthy AI-driven BEM should evolve from passive command execution to interactive design–simulation–feedback loops that embed continuous energy validation into early schematic exploration, enabling real-time negotiation between spatial intent and environmental performance. The collaborative gains discussed above remain workflow hypotheses that future research must validate through real projects rather

than assume as consequences of automation. Second, to satisfy the need for auditable human-in-the-loop verification, the field should establish traceable "Simulation Justification Protocols" that document model assumptions, parameter provenance, and result confidence, ensuring every AI inference remains auditable and engineering-verifiable. Third, to confront the computational energy and carbon cost of persistent inference, the computational carbon of AI itself must be integrated into whole-life carbon assessments with AI computational energy reported alongside building operational energy as a measurable design variable. A practical mitigation lies in reducing per-cycle token consumption through a smaller or fine-tuned domain model, structured tool calls, and concise prompting, which narrows the computational-energy gap rather than eliminating it.

Together, these developments will move AI-driven BEM from isolated automation toward a sustainable modeling ecosystem. The next benchmark should therefore shift from measuring autonomous model generation to measuring how reliably humans can understand, verify, and govern the evidence produced by agentic workflows—ensuring they remain trustworthy, explainable instruments of evidence-based design that preserve human creative authority.

REFERENCES

1. UNEP and GlobalABC (2025). Global Status Report for Buildings and Construction 2024/2025: Not just another brick in the wall.
2. Hong T. Z., Langevin J. and Sun K. Y. (2018). Building simulation: Ten challenges. *Build Simul-China* **11**:871–898. DOI:10.1007/s12273-018-0444-x
3. Anthropic (2024). Introducing the Model Context Protocol. <https://www.anthropic.com/news/model-context-protocol>
4. Qian C., Liu W., Liu H., et al. (2024). Chatdev: Communicative agents for software development. Proceedings of the 62nd annual meeting of the association for computational linguistics (volume 1: Long papers).
5. Zhang L., Ford V., Chen Z. L., et al. (2025). Automatic building energy model development and debugging using large language models agentic workflow. *Energy Buildings* **327**. DOI:ARTN 115116 10.1016/j.enbuild.2024.115116.
6. Li H., Xu Y. J. and Hong T. Z. (2025). EnergyPlus-MCP: A model-context-protocol server for ai-driven building energy modeling. *Software* **32**. DOI:10.1016/j.softx.2025.102367
7. Li H., Zhang L., Zhou H. W., et al. (2026). MCP-enabled agentic AI workflow for building energy modelling: framework and use cases. *J Build Perform Simu*. DOI:10.1080/19401493.2026.2653969
8. Google (2025). A2A: A new era of agent interoperability. <https://developers.googleblog.com/en/a2a-a-new-era-of-agent-interoperability/>
9. Xu W., Wan H., Goel S., et al. (2025). Development of a dynamic multi-agent network for building energy modeling: A case study towards scalable and autonomous energy modeling. *Energy Buildings*:116712. <https://doi.org/10.1016/j.enbuild.2025.116712>.
10. Yamada Y., Bao Y., Lampinen A. K., et al. (2023). Evaluating spatial understanding of large language models. *arXiv preprint arXiv:2310.14540*

FUNDING AND ACKNOWLEDGMENTS

This research work is financially sponsored by the National Natural Science Foundation of China (grant number 52308105) and the State Key Laboratory of Subtropical Building and Urban Science (grant number 2024ZB15).

AUTHOR CONTRIBUTIONS

Haoyi Zhu: Conceptualization, Methodology, Visualization, Writing – original draft.
Yueer He: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

DECLARATION OF INTERESTS

The authors declare no competing interests.