Green Building Design Studio Game Development with Parametric Simulation and ML Prediction for Green Building Education in Rural Middle Schools

Jong Bum Kim[1], Jayedi Aman[2], Laura Zangori[3], Laura Cole[4], Heather Young[5], and Catherine Stimac[6]
[1,2,3]University of Missouri Columbia, [4] Colorado State University.
[5,6] Oregon Public Broadcasting.
[1]kimjongb@missouri.edu, ORCID: 0000-0001-8762-9554
[2]jayediaman@missouri.edu, ORCID: 0000-0002-6128-8293
[3]zangoril@missouri.edu, ORCID: 0000-0002-7512-5559
[4]laura.cole@colostate.edu ORCID: 0000-0001-5730-1881
[5]hyoung@opb.org, [6]cstimac@opb.org

Abstract. Green building education prioritizes workforce development to promote high-performing and net zero building adoptions. However, the concept and principles of net zero and building energy have rarely been reflected in the curriculum and instruction of K-12 science education in the United States. This research investigates the design and development of simulation game development paired with a science curriculum to teach green building design and energy principles in rural middle schools. This paper presents our education game development aligned with the newly developed curriculum unit that will be distributed to science classrooms. Green Building Design Studio game was developed from the following research phases: (i) Game scenario design, (ii) Energy simulation module creation, (iii) ML-prediction model development, and (iv) Cost estimation module creation. In ML prediction, the XGBoost algorithm demonstrated reliable performance and accuracy. The game was tested in a 3-day science immersion summer camp with twenty-seven middle school students in Missouri. The research team observed that the game enabled students to iterate de sign changes and promptly see the updated results from the dashboard. This paper describes the game development framework, methods and tools for energy simulation, ML prediction, and game development, as well as the findings and challenges.

Keywords. Green building education, game-based learning, energy simulation, ML prediction

1. Game-based Learning and Green Building Education

Green building education prioritizes workforce development to promote high-performing and net zero building adoptions. The U.S. Department of Energy (DOE) defines a green building as a high-performance building that has reduced greenhouse gas emissions from building operations using energy efficiency and clean energy.

ARCHITECTURAL INFORMATICS, Proceedings of the 30th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2025, Volume 1, 479-488.

However, the concept and principles of green building energy have rarely been reflected in science education within K-12 classrooms in the United States (Kandpal & Broman 2014).

Game-based learning and gamified training were studied to enhance knowledge and improve skills using serious games (Marsh, 2011). Serious games, built on digital game technologies, repurpose the application for education and training rather than entertainment (Dib, H., 2014). During the last decades, various serious games attempted to improve energy knowledge, such as ElectroCity (2007), Clim'way (2010), CityOne (2011), EnerCities (2011), EnergyVille (2011), Supergreen (2018), Powersaver (2019), etc. However, these technologies were not developed to be embedded within formal science curriculum materials.

The research investigates the design and development of simulation game development paired with a science curriculum to teach green building design and energy principles in rural middle schools in the United States. This paper presents our education game development aligned with the newly developed curriculums that will be distributed to the science classrooms. This paper also describes (i) background theory, (ii) curriculum design, (iii)the game development framework, (iv) methods and tools for energy simulation, ML prediction, and game development, and (v) the findings and challenges.

2. Integration Framework of Simulation, ML prediction, and Serious Game

2.1. OBJECTIVES

In this research, the simulation-based game was developed jointly with a newly developed curriculum to achieve learning objectives about complex energy concepts and green building principles. The research also prioritizes students' engagement by exploring various energy options in the serious game. The simulation-game development is an attempt to integrate physics-based simulation, ML predictions, and interactive game development to achieve the following objectives.

Realistic Datasets from Simulation: The first goal is to provide more accurate datasets for the curriculum and game using physics-based energy simulations and historic data-driven cost estimates. The research envisions that the simulation and estimation results allow students to explore various building configurations and energy options, establishing students' active learning in the science classrooms.

Problem-solving within the inverse relationships: Building construction and material costs are associated with their energy performance. In our research, cost estimation of building design and energy options aimed to contextualize the relationship between initial construction costs, long-term energy cost saving, and potential emission reduction. It will foster critical thinking about the trade-offs in green building design and renewable energy technologies throughout the building lifecycle.

More game and energy options for student engagement: The simulation-based ML predictions aim to drastically increase the number of what-if scenarios of energy options in the game. The simulation-based ML generate large datasets that can show the impact of various design and energy options. It can also increase students' engagement by exploring a range of design and energy options.

Game as an interactive learning environment: The game consists of interactive interfaces and dashboards for students' hands-on learning. By manipulating design and energy options, students can spontaneously compare and observe the impact of their design choices on energy performance, construction costs, operational costs, and emissions. This interactivity aims to promote systems thinking and iterative problem-solving in the context of green building and energy principles (Kim et al., 2023).

2.2. PROCESS AND METHODS

The simulation game development integrated physics-based building energy simulation, construction cost estimation, and ML predictions. The overarching research priority was to facilitate parallel and bi-directional synchronization and alignments among the three domains of curriculum development, game development, and building performance simulations. The overall development process is illustrated in Figure 1.

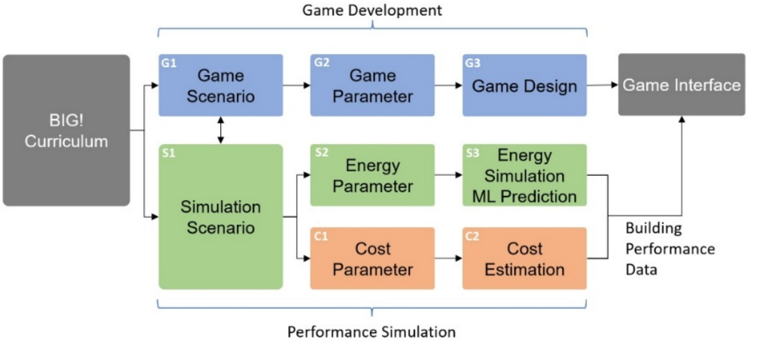

Figure 1 Simulation game development process

Build it Green! (BIG) is a newly developed science curriculum with embedded digital interactives and simulation game, funded by the NSF. BIG fosters place-based education by using local school buildings to enhance systems thinking about energy consumption and flow between buildings and Earth systems. Systems thinking is taught through the curricular framework in which students have ample opportunities to break large systems into smaller components to look closely at key components, causal interactions, and inputs/outputs within systems.

The simulation game (G1~G3 in Fig.1) is the key learning material in the last lesson that students can test their knowledge and skills gained from individual lessons. The formulation of learning objectives and game and simulation scenarios were presented (Kim et al., 2023). We designed the game scenario and interfaces to engage students with complex systems in more interactive and iterative ways. In the game, students manipulate the design and energy options, identify the building energy sources, flow, and consumption, and learn how individual element affects the overall building system.

Building performance simulation (S1~S3 and C1~C2 in Fig.1) addressed the following criteria:

- Physics-based energy simulations enable testing a wide range of building configurations of building design and energy systems. It includes building location, orientation, operation, dimensions, materials, mechanical and electrical equipment, etc. The use of validated simulations provides accuracy in the simulation results.
- ML predictions enable the production of many datasets according to the enormous building design options in the game. Despite their accuracy, physics-based simulations are knowledge-dependent and time-consuming, limiting their dissemination to the classroom setting. The ML model created from the physics-based simulation yields accuracy. Besides, the vast number of prediction results enables students to carry out manipulations and iterations of design and energy options.
- Building cost estimation using historical and location-specific data produces realistic projections of building construction costs. The building cost data is a baseline for understanding the potential cost of energy savings, energy generation, and greenhouse gas emission reduction throughout the building lifecycle.

3. Game and Simulation Development Framework

The game scenario was developed based on 13 lessons in the curriculum, teaching building energy systems such as building volumes, materials, façade and window configurations, daylighting, artificial lighting, HVAC systems, etc. The players gain knowledge in several ways: (i) manipulating the design options and the energy system configurations, (ii) reading the changes in building cost, energy cost, and gas emissions, (iii) reviewing the interactive feedback in the game dashboards, and (iv) iterating the design changes to improve energy performance, lower energy cost, and reduce gas emissions. Figure 2 shows the software development workflow for energy simulation, ML prediction, cost estimation, and performance data aggregation.

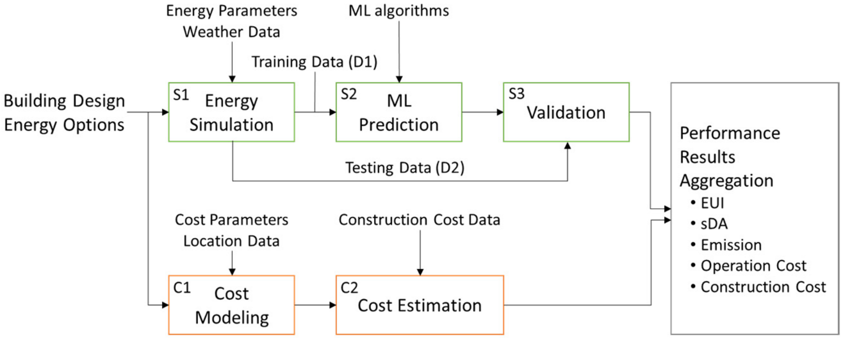

Figure 2 Software development workflow for game data creation

3.1. ENERGY SIMULATION

The game presents a baseline classroom of typical public schools in the United States. The energy model has a single thermal zone with several design options. In addition, simulation parameters include energy system options such as HVAC systems, lighting systems, solar panels, wind turbines, etc.

The simulation tested 15 parameter criteria, and 2,500 simulations were performed to create a dataset for ML model training and testing. Rhinoceros was used for energy modeling, Grasshopper visual scripting facilitated iterative simulation and data exchange, and energy simulation plug-ins calculated energy performance metrics.

Table 1 shows the example game and simulation parameters, including key design options such as building geometry, envelope properties, openings, and HVAC systems. A total of 139,968 cases were sent to the game, encompassing all parameter combinations.

Table 1 Example parameters in game and simulation

Category	Curriculum/ Game Parameters	Simulation Parameters
Room size*	Small, Medium, Large	15 ft x 15 ft, 10 ft x 20 ft, 25 ft x 25 ft,
Room shape*	Square, Rectangle	20 ft x 30 ft, 35 ft x 35 ft, 30 ft x 40 ft
Room Height	Standard, High	10 to 12
Room Orientation	North, West, South	0, 90, 180
Roof Type	Flat, Shed, Gable	Flat (0), Shed (1), Gable (2)
Exterior Wall Material*	Brick, CMU, Precast, Metal	R 2.5-7.5 m2K/W
Roof Material*	Membrane, Metal, Concrete	R 3-7 m2K/W
Window Numbers*	1,2,3	0%, 25%, 50%
Window Type*	Single Pane, Double Pane	4.8 W/m²·K, 1.2 W/m²·K
External Shading	No Shading, Shading	0, 1.5 ft
Door Numbers*	1,2	0.5 ACH
Door Material*	Wood, Glass, Metal	1.0 ACH
HVAC System	High, Medium, Low Efficiency	Electric Baseboard (1), Gas-fired Boiler (2), Ground Source Heat Pump (3)
Conditioning	Comfortable; Less Comfortable	S70F - W70 F; S75F - W65F
Total Iteration: 139,968 Scenarios		

3.2. ML-PREDICTION

The machine learning module was created using an ML library and Python scripting. The simulation results were split into training (70%), testing (20%), and validation (10%) datasets. The predictive model utilized the Extreme Gradient Boosting (XGBoost) algorithm and included outputs of spatial daylight autonomy (sDA), energy use intensity (EUI), and energy costs. Additionally, carbon emissions were calculated using Missouri's carbon factor, where the curriculum and game are disseminated. The prediction module processed all 139,968 simulation cases.

3.3. COST ESTIMATION

The cost estimation used 2023 Assemblies Costs with RSMeans data (Doheny, 2023). Assemblies Estimates use groups of individual unit cost lines combined for fast and effective estimation in the planning phase. After collecting the historical cost records for each building assembly, the calculation module combined the total costs and adjusted the total cost using the city cost indexes and the location factors according to the target school building location. The RSMeans data provides the national standard data and location factors for 730 cities in over 900 zip codes in the U.S. and Canada.

3.4. GAME AND SIMULATION COMMUNICATION

Designing an engaging game interface guided our methodology and process. We balanced the simulation accuracy with the limitations of a real classroom environment. The visual representation focused on middle school students experience and prior energy knowledge and the learning goals of the curriculum. The simulation generated JSON files to store data. The game then loads the appropriate JSON file based on a subset of student selections. This method was chosen for the efficient download and retrieval speed for typical school networks and student computers.

4. Results

Design Studio game resulted from the co-development of curriculum, energy simulation, and ML prediction. In this research, game scenarios, energy simulation, and curriculum lessons are fully aligned to facilitate students' learning.

4.1. SIMULATION RESULTS

The simulation results were compared with the ASHRAE Standards 140 (2020), a standard test method for building energy simulations. Figure 3 presents heating and cooling loads of the CASE 600 model from the validated simulation tools. The

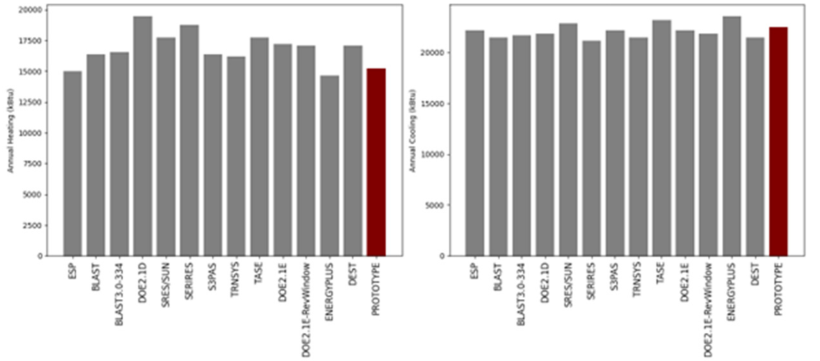

Figure 3 Comparison of heating load (left panel) and Cooling load (right panel)

prototype yielded 17,000 kBtu in heating load, which is similar to EnergyPlus (17,050 kBtu) and DEST (17,100 kBtu). The cooling load from the prototype is approximately 22,600 kBtu, slightly higher than EnergyPlus. It demonstrated validity and reliability.

4.2. ML PREDICTION RESULTS

To ensure the accuracy of the ML model, learning curves and regression were analyzed in Figure 4. Learning curves depict how the errors of the training set and validation set change as the size of the training data increases. It can evaluate the model's ability to generalize to unseen data.

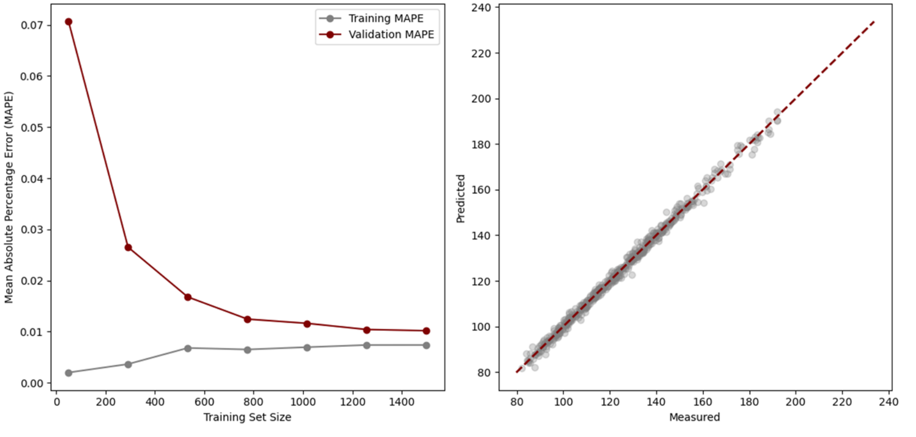

Figure 4 Learning curve (left panel), Regression plot (right panel)

In the left panel, the training error starts at a very low value of approximately 0.005 when the model is trained on a small subset of 50 samples. Then, the training error gradually stabilizes at about 0.01 when the training size becomes around 1,500 samples. This plateau suggests that adding more data has little effect on further reducing the training error.

On the other hand, the validation error starts at a higher value of approximately 0.07. This indicates that the model initially struggles to generalize to unseen data. However, as the training set grows, the validation error decreases steadily and approaches the training error, eventually stabilizing around 0.011 when the training size reaches 1,500 samples. This convergence reflects an improvement in the model's generalization. Overall, the current training data size appears sufficient for optimal model performance, and further increasing the dataset size is unlikely to result in significant improvements.

The regression plot in the right panel exhibits a strong correlation between the physics-based simulation and ML prediction values. Similar patterns were observed for both daylighting (sDA) and energy use intensity (EUI). The model achieved an R2 value of 0.97 for EUI, demonstrating its ability to reliably predict energy performance. Similarly, the results for sDA show a high degree of agreement, with deviations between predicted and simulated values consistently minimal.

Table 2 presents the deviation between the simulation and ML prediction results of the randomly selected samples. Most cases show minor deviation levels below 1%. In short, the ML model using XGBoost is applicable to the game with minimal error.

Table 2 Comparison of Simulation and ML Prediction Results

IDs	Simulation sDA (%)	ML sDA (%)	Deviation (%)	Simulation EUI (kBtu/ft²)	ML EUI (kBtu/ft²)	Deviation (%)
4204	15	14.95	0.05	120.83	119.53	1.08
4158	29.63	29.35	0.28	183.75	183.34	0.22
3166	13.70	13.42	0.28	138.85	139.78	0.67
3158	53.13	52.05	1.08	143.88	145.90	1.41
4237	10	10.10	0.10	95.11	96.28	1.23

4.3. GAME INTERFACE AND DASHBOARD

Figure 5 shows the interface of the resulting game. The game starts from the design studio, enabling students to choose various options for building design, electrical devices, and renewable energy. From the left-side panel, students can monitor the design changes, including building layout, opening, roof design, material colors, etc. Based on the building sizes, students read the available budget and required construction cost. After selections, they move to the electrical devices and renewable energy.

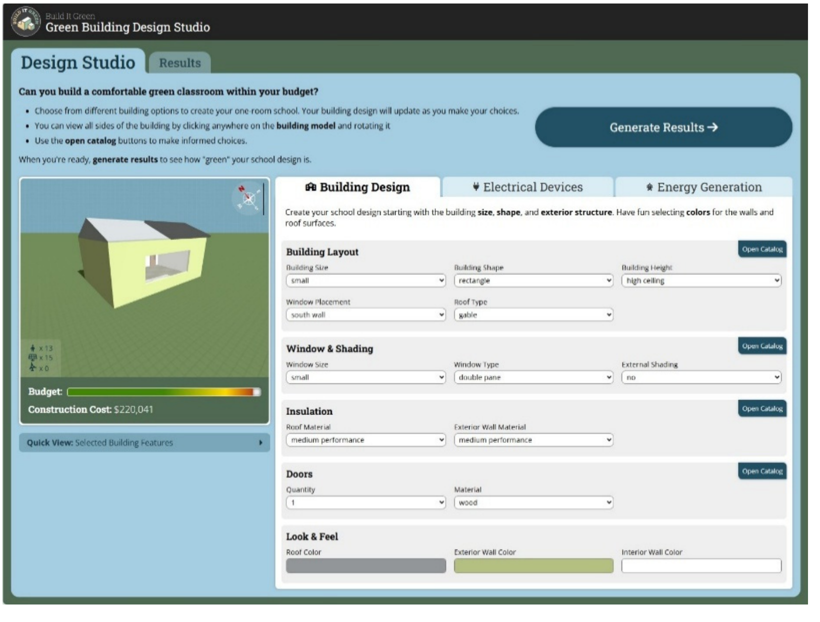

Figure 5 Design Studio Interface

Next, students can see the results from the Result dashboard in Figure 6. The dashboard provides multiple visualization formats to help students determine if their design meets the goal, including green rating leaves, smile icons, success rated badges, monthly plots, and numerical information. The results are displayed in three tabs for energy consumption, energy operating cost, and gas emissions.

Figure 6 Result Dashboard

In 2024 summer, the BIG curriculum and the game were tested at the science summer camp (Figure 7). Twenty-seven students from the local middle schools participated in the four-day camp. They learned the basic knowledge from individual lessons, visited solar houses equipped with solar panels and sustainable materials, designed the school building with paper models (middle and right), and played the simulation game (left). The research team observed that multi-layered learning opportunities engaged students in science learning. The game enabled students to iterate design changes and promptly see the updated results from the dashboard.

Figure 7 Simulation Game at the Science Summer Camp

5. Conclusion

The Build it Green (BIG) design studio game development laid a theoretical and empirical foundation for comprehensive science education curriculum for middle school aged students (12 - 14 years of age). The BIG game was our attempt to take advantage of design computing technologies for science education in rural school districts with underserved populations for STEM education. Within the U.S., most rural schools are underserved for STEM education, which includes less access to learning technologies, while the demand for workforce training is growing in the gas, oil, wind, and solar energy industries. The BIG curriculum and the game are being disseminated to these rural school districts. For broad dissemination, the game development targets tablet-based games without relying on data servers or internet access. When considering these limitations, further research is required to enhance communications and synchronization among physics-based simulation, ML predictions, and game development. Next research step is to improve accuracy of simulation results and game usability based on the feedback collected from the summer camp.

Acknowledgments

This material is based upon work supported by the National Science Foundation (Award No 2201204). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the NSF.

References

ANSI/ASHRAE Standard 140-2017, Standard Method of Test for the Evaluation of Building Energy Analysis Computer Programs

Ampatzidou, C., Constantinescu, T., Berger, M., Jauschneg, M., Gugerell, K., & Devisch, O. (2018). All work and no play? Facilitating serious games and gamified applications in participatory urban planning and governance. Urban Planning, 3(1), 34-46.

Dib H and Adamo-Villani N. Serious Sustainability Challenge Game to Promote Teaching and Learning of Building Sustainability. Journal of Computing in Civil Engineering 2014; 28(5). DOI: https://doi.org/10.1061/(ASCE)CP.1943-5487.0000357

Kim, J. B., Oprean D, Cole L, & Zangori L. (2023). Illumi's world: A mini-game development with parametric BIM-based simulations. International Journal of Architectural Computing. 2023;21(3). Sage Journals. doi: 10.1177/14780771231180256

Marsh T. Serious games continuum: between games for purpose and experiential environments for purpose. Entertainment Computing 2011; 29: 61–68.

Rossano, V., Roselli, T., & Calvano, G. (2017). A serious game to promote environmental attitude. In V. L. Uskov, R. J. Howlett, & L. C. Jain (Eds.) International Conference on Smart Education and Smart E-Learning (pp. 48-55). Switzerland: Sprinter International.

RSMeans, Assemblies Costs with RSMeans data. (2023). Norwell, MA:RSMeans.

Tien, P. W., Wei, S., Darkwa, J., Wood, C., & Calautit, J. K. (2022). Machine Learning and Deep Learning Methods for Enhancing Building Energy Efficiency and Indoor Environmental Quality – A Review. Energy and AI, 10, 100198.

Zangori, L., & Cole, L. (2019). Assessing the contributions of green building practices to ecological literacy in the elementary classroom: an exploratory study. Environmental Education Research, 25(11), 1674-1696.