Integrating Surrogate Modeling and Design Analytics for Data-informed Exploration in the Early Phases of Building Design

Esmaeil Mottaghi[1], Halil Erhan[2], Ahmed M. Abuzuraiq[3], Victor Okhoya[4], Spyridon Ampanavos[5], Marcelo Bernal[6], Cheney Chen[7] and Yehia Madkour[8]
[1,2,3] Simon Fraser University, [3,4,5,6]Perkins&Will
[1,2,3]{ema63|herhan|aabuzurua}@sfu.ca
[4,5,6,7,8]{Victor.Okhoya|Spyridon.Ampanavos|Marcelo.Bernal|Cheny.Chen|Yehia.Madkour}@perkinswill.com

Building design workflows have been influenced by incorporating data-driven decisions, algorithmic content generation, and performance analysis support offered by machine learning methods. The challenge lies in bringing these seemingly isolated but highly related design tasks together in an engaging design environment that can maintain designers' creative decision-making flow while reducing system interferences' complexity. This paper presents D-Predict.v2 as an experimental prototype we developed following a design study methodology to tackle this challenge. As a contribution, we propose a design workflow and a system that can be adapted to support building performance prediction using surrogate models in direct or parametric design modelling driven by interactive data visualizations. Our initial findings demonstrated that expert designers welcomed the proposed workflow with caution and recommendations. Our future work will focus on conducting ecologically valid case studies and rebuilding the system by addressing the concerns raised by the expert designers.

Keywords: *Generative Design, Building Performance Assessment, Surrogate Modelling, Design Analytics.*

INTRODUCTION

We present a design study research project integrating generative design with performance assessment using Machine-Learning-based Surrogate Models (ML-SM) and Design Analytics interfaces. Building performance assessment involves complex tasks ranging from model preparation to performance simulation and data analysis. The complexity of the process is emphasized when making trade-off decisions between different performance metrics, such as form, view quality, solar gain, and energy load. Our research questions concerned whether and how design modelling tasks can be informed by data rapidly generated using ML-SM techniques and design analytics dashboards. We introduce a fully functional version of D-Predict.v2 as an experimental prototype system we built on the Rhino3D environment by questioning (a) if ML-SM can substitute labour-intensive and computationally intense performance simulations in concept development, (b) how quickly designers can generate performance data and (c) how such data can be revealed to the designers to support data-

informed decision-making. Given the computational cost of performing simulations, using surrogate models for performance prediction is a viable solution (Westermann & Evins, 2019; Yousif et al., 2022). We aimed to leverage design data visualizations to make sense of performance data and reduce the designers' choice overload (Erhan et al., 2017).
Our research goal is to combine performance prediction and data analytics, typically practiced discretely in the current design workflows. By integrating modelling and visual data analytics, D-Predict.v2 explores the role of surrogate modelling in the design process.

We follow two concurrent methodologies: a design study methodology (Sedlmair et al., 2012) for problem characterization and agile system development methods (Turk et al., 2003) for system development based on problem characterization. By collaborating with design practitioners, we elicit requirements as use cases and develop and test interaction features within design scenarios. D-Predict.v2 includes interactive design data visualizations for coordinating design and performance prediction using ML-SL. Its functional prototype supports (a) direct or parametric design specification and generation, (b) assessment of design models using interactive visualizations, and (c) connecting with ML-SL for performance prediction. We aim to maintain the task flow between generating alternatives, predicting their performance, and managing trade-offs between alternative solutions.

Our initial findings from a formative expert review evaluate the D-Predict.v2's utility, usability, and adoptability. The preliminary findings indicate that the proposed integration aligns with existing design exploration processes, helping the identification of high-performing designs concerning daylight and energy load metrics. While the use of surrogate models accelerates this process compared to simulation methods, challenges emerged in accommodating architectural design styles, model resolution, and the generality of trained surrogate models. Future research will focus on experimenting with real-world cases to refine the system features with broader performance concerns.

BACKGROUND

By utilizing generative design methods, designers can explore a larger number of design alternatives, potentially leading to more diverse and better-performing designs than conventional methods. Parametric design models are used in these methods combined with exploration tools that range from interactive (e.g., mixed-initiative interfaces) to fully autonomous (e.g., optimization). Similarly, performance evaluation ranges from expertise-driven assessment to ML-driven performance simulation and prediction. The goal of incorporating performance analysis is to identify and address potential issues early in the design process, improve the cost-effectiveness of built environments, and meet project constraints while controlling performance goals. Through this, designers can optimize building performance, enhance energy efficiency, and reduce environmental impact, ultimately advancing the sustainability and resilience of built environments. However, generative design and performance analysis rely on fundamentally distinct toolsets. The challenge lies in bringing these seemingly isolated but highly related design tasks together in an engaging design environment that can maintain designers' flow in creative decision-making while reducing the design systems' interferences.

Design Analytics offers a set of methods that can potentially integrate design exploration and performance assessment tasks (Erhan et al., 2020). Through interactive spatial and non-spatial design data visualizations, designers can test design decisions and make sense of the relevant data. It can enable the analysis of competing alternatives and the choice of several competing ones. We aim to address this challenge and explore how to support design generation and evaluation to improve design outcomes with Design Analytics.

While general frameworks and descriptive accounts provide a foundation for understanding design, we focus on enabling designers with new tools for high cohesion and low coupling to support creative flow. Research on this issue mainly involves case studies of specific domains such as programming, data analysis, or building design. However, this focus on domains can lead to the invention of specific interfaces that limit the consideration of the entire problem.

Parametric [Generative] Design Considering Performance Metrics

A typical parametric design and performance assessment process includes specifying, synthesizing, and evaluating workflows (table 1). In parametric design, designers explore solutions by specifying design models parametrically, generating designs manually, changing parameters or using algorithms, and evaluating the solutions using data visualizations or conventional methods; designers search the parts of design space otherwise inaccessible by traditional methods. As an extension to parametric design, performance assessment aims to experiment if and how a built environment would perform in the real world by simulating its model against various design criteria such as energy loads, structural behaviour, and artificial or daylighting (Westermann and Evins 2019; Westermann et al. 2021) to find *satisficing* alternatives (Simon 1969). In performance assessment, models are specified at an abstraction level suitable for testing design concerns, computing the model's behaviour under conditions that can simulate real-world factors, and evaluating the computed data to assess the conformance of the designs against the requirements. Performance simulations help optimize and improve energy efficiency and reduce environmental impact, promoting built environments' sustainability and resilience. However, simulations can be computationally expensive. ML-SL allows exploring the effect of variable changes on performance early in the design process rather than running time-intensive simulations later, where they have a diminished impact on design outcomes.

ML-SL for Performance Analysis

Researchers and practitioners have experimented with surrogate modelling using deep-learning techniques to increase the efficiency of analyzing a built environment's performance or its environmental impact (Westermann and Evins 2019; Westermann et al. 2021; Yousif and Bolojan 2021). ML-SL has proven to predict building performance metrics, including thermal comfort and energy use (Chegari, 2022). Researchers at Perkins and Will recently introduced a general-purpose surrogate model for spatial daylight prediction. Using a raytraced-based encoding and a deep-learning model, the method predicts the distribution of daylight for any shaped spaces. Such models can reduce the need for labour-intensive and computationally expensive simulation-based analysis or time-consuming physical prototyping (Wortmann et al. 2015). Therefore, this technique can be integrated into the generative design workflows to evaluate design alternatives that are explored rapidly.

Yousif et al. (2022) proposed an automated performance-driven generative design approach using surrogate daylighting models for creative floor layout exploration. Zorn et al. (2022) proposed integrating surrogate models for early-stage design evaluation into a real-time dashboard. Westermann et al. (2020) developed the Net-Zero Navigator

Table 1
Three high-level design tasks: specify, generate, and evaluate alternatives.

Data	Specify	Generate	Evaluate
Design Form	Create parametric models	Generate parametric variations/alternatives	Form conformance considering design requirements
Design Performance	Build or select simulation or surrogate models concerning design considerations	Compute design performance assessment data for each alternative using surrogate models	Performance conformance considering design requirements,

platform, featuring accurate surrogate models for energy and interactive data visualizations. However, some of these approaches lack interactive design exploration interfaces or neglect to integrate creativity support features.

Challenges to using surrogate modelling

Generative design and ML-SM-driven performance evaluations can be combined to understand and solve complex design problems (Wortmann et al. 2015); however, they are seen as two distinct, isolated tasks in practice. Two main challenges must be addressed to ensure the success of combining generative design with performance assessment. The first challenge is creating accurate surrogate models accessible to designers. Since surrogate modelling abstracts or approximates more complex designs, the effectiveness and accuracy of this technique depend on the model's quality and the data used to train it. The second is exploring the numerous outputs of generative design in a way that considers both quantifiable (performance) and unquantifiable (e.g., form composition and aesthetics) design concerns.

The quality of surrogate models depends on multiple factors, including the choice of ML techniques (e.g., neural networks, support vector machine), the training process (to fine-tune the model and avoid overfitting and loss of generalization), and the training dataset. The training dataset should ideally include diverse and distinct design situations to improve the model's generalizability. Creating custom surrogate models for each design project is often more practical than training globally applicable models that can adapt to changing contexts. Proper tools for testing the validity of these models and creating the training datasets are necessary, and the machine-learning community provides extensive guidance and tools for this. However, surrogate models may only be useful to designers if they are a seamless part of a design workflow.

The second challenge is exploring the numerous outputs of generative designs despite the cognitive choice overload they create (Erhan et al., 2017). Multiple data analysis and visualization approaches have been proposed to address this challenge, referred to as design space navigation tools. These tools can filter, sort, and group-generated designs while considering both the performance and the visual forms of these designs. Examples of design space navigation tools include DreamLens by Autodesk Research (Matejka et al., 2018), Design Explorer and Thread by Thornton Tomasetti CORE Studio (Tomasetti, 2022), the Design Space Explorer by Fuchkina et al. (2018), and the thesis work by Beham et al. (2014), van Kastel (2018) and Abu Zuraiq (2020).

PROPOSED DESIGN ANALYTICS WORKFLOW WITH ML-SL

Earlier, we proposed design analytics interfaces combining design and performance prediction using data visualizations central to an integrated workflow (Mottaghi, Abuzuraiq, and Erhan, 2024).
The performance prediction layer generates data by selecting surrogate models for performance concerns. After merging data from a parametric model and performance prediction, designers explore the potential of each solution or set of solutions using coordinated views (figure 1).

High-level system requirements

In this study, we employ a design study methodology (Sedlmair et al., 2012), which builds on the literature review above and involves architectural design practitioners as collaborators in the research process. We gathered a set of high-level requirements for design analytics systems for performance evaluation in parametric design based on a literature review and continuous interaction with architects with computational design expertise from the building industry. We summarize them in two different sets. The first set classifies such systems

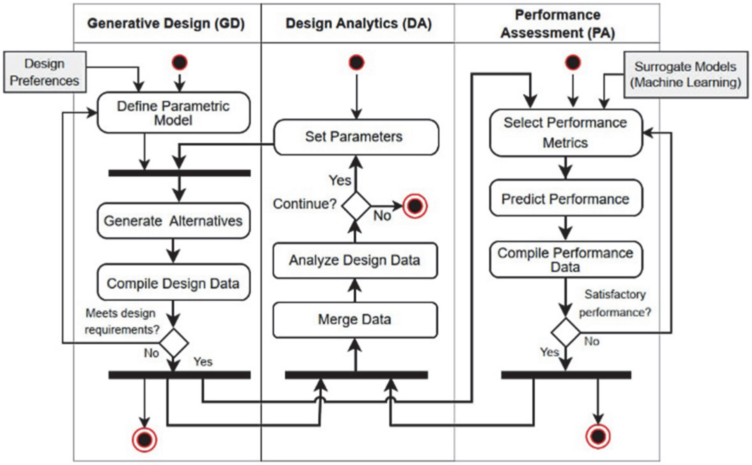

Figure 1
The proposed workflow integrates two related design tasks, design generation and performance assessment, with data visualizations to support data-informed decision-making (Mottaghi, Abuzuraiq, and Erhan, 2024)

as creativity support tools (Shneiderman 2007). They should:

- R1. Support exploration design alternatives.
- R2. Support collecting and retrieving design.
- R3. Keep a history of choices.
- R4. Support the different exploration methods and styles ("wide walls").

The second set focuses on design:

- R5. Facilitate simple and engaging interactions with design models and surrogate models for performance analysis, catering to designers with varying levels of expertise.
- R6. Guide selecting parameters and make their sensitivity on design generation transparent (Hamby 1994; Bernal et al. 2020).
- R7. Provide advanced data analysis visualization for those who can use them (Abu Zuraiq 2020).

The high-fidelity version of D-Predict.v2 offers design generation, exploration, and comparison capabilities. It aims to provide insights into design exploration by analyzing alternatives using surrogate models and their comparative evaluation of incremental design decisions or stored alternatives. The prototype is an integral part of the design exploration process as a usable system.

D-PREDICT.V2 SYSTEM DESIGN

Building on the high-level system design requirements and incorporating constant feedback from computational designers, we composed D-Predict.v2 interfaces in four main components (figure 2): (a) modelling (C); (b) parametric controllers for form, material, zoning, shading device adjustments; (c) data visualization view for daylight (A, B, E) and energy load (F) for reporting, and (d) comparative analysis of alternative design solutions. In this version, the surrogate model is used to predict daylight performance, and the energy load is computed using the Ladybug.

Modelling. In various phases of a building design, modelling tasks continue based on the decisions made using the feedback received or better solutions are discovered. In such cases, models can be stable, i.e., form decisions are made, but the material may change, or both form and material choices can be amenable to limited. changes or form and material choices are open to any change, as in the early phases of design.

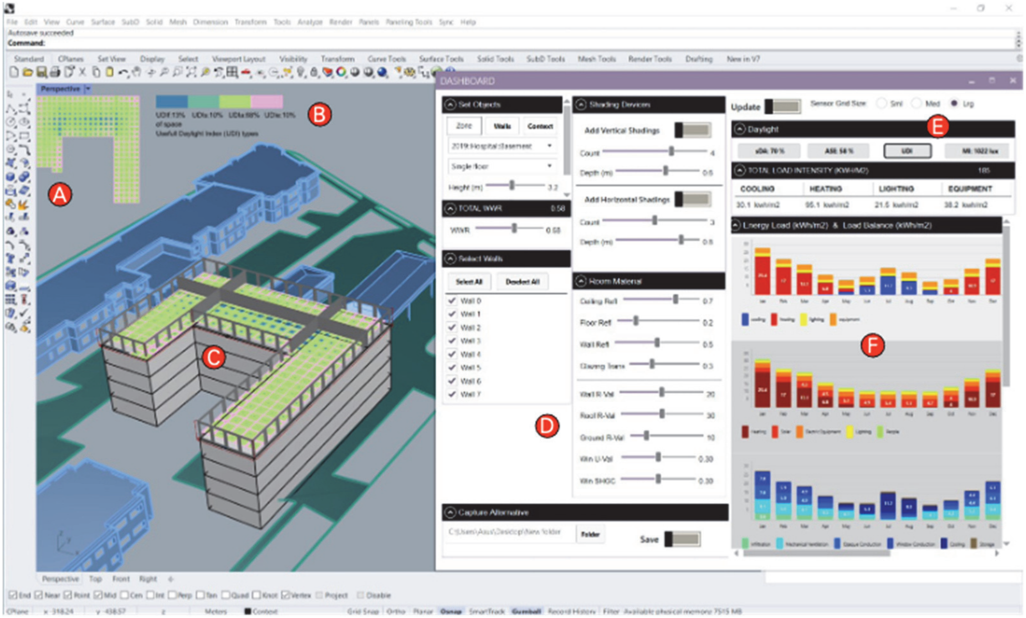

Figure 2
D-Predict.v2 main interface with model-agnostic parametric controllers and data analysis panels.

D-Predict.v2 allows designers to experiment with form and material choices, mainly in the exploration phase, but can also support direct or parametric modelling in the later phases. A typical modelling workflow in D-Predict.v2 starts with creating or selecting a building boundary, including elements that define interior zones. The system generates walls or panels and inserts openings using the Window-to-Wall ratio set by designers (figure 4 A). Designers define zones by selecting partitioning elements for energy load simulation and select context elements, such as buildings, for daylight prediction. The boundary elements can be assigned different material properties for each room or zone, such as surface reflectivity, R values, or U-values (figure 3 D).

Parametric Controllers. Designers can edit windows, shading devices, and (window) glazing parameters on select exterior walls (figure 3 B). The system detects and shows the walls for individual or group selection (figure 3 B). Shading devices are parametrically generated on each opening. The number of horizontal and vertical parts with their depths is set on the interface (figure 3 C). D-Predict.v2 uses local weather data when computing daylight and energy loads.

Data visualizations. The daylight metrics are viewed in two different locations: (a) on the model and in the peripheral view near the model as heatmap visualization (figure 2 A, C) and (b) in the tabular view on the general visualization panel (figure 2 E). Does the surrogate model generate data on uniformly distributed sensors that are used to compute four daylight metrics: Spatial Daylight Autonomy (sDA), Annual Sunlight Exposure (ASE), Useful Daylight Illuminance (UDI), and Mean Illimumance (MI) (figure 4).

Energy load metrics for annual and monthly cooling, heating, and equipment load intensity are viewed and visualized on a stack bar chart.

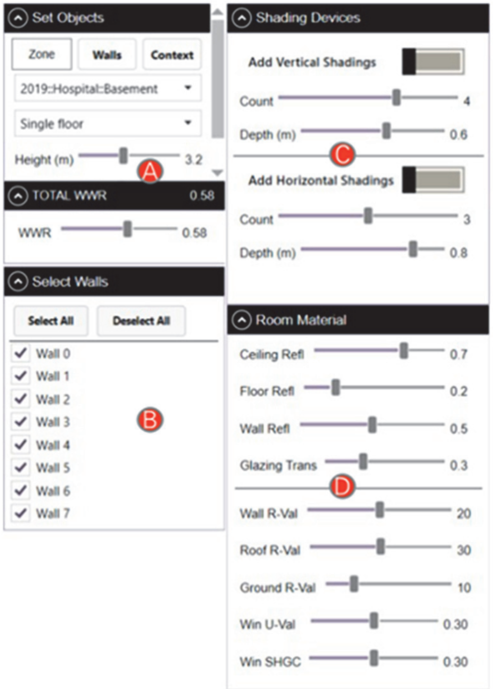

Figure 3
Parametric controllers (A) define zones and contexts; (B) select walls to apply material or openings; (C) set up Horizontal and vertical shading devices; and (D) assign material to boundary elements, such as ceilings, floors, walls, or windows.

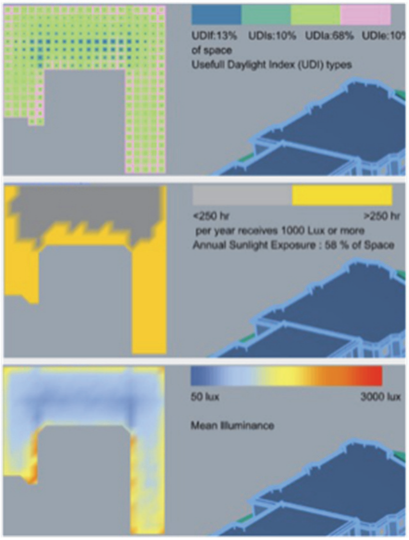

Figure 4
Peripheral daylight metrics (UDI, ASE, and MI) visualizations with benchmark references.

Similarly, energy gain and loss are visualized on two stack bar charts. Although normally viewed on the same chart, two different charts were used due to the limitation of HumanUI (figure 5).

Comparing alternative solutions. One of the main contributions of the D-Predict.V2 is its ability to store and retrieve design alternatives as the designers explore different configurations. The stored alternatives can be viewed and compared in the same data visualization views (figure 6). Designers can select an alternative for further inspection and choose more than two alternatives to compare their performance values, including the images captured from the 3D models and daylight visualizations. We aim to improve this feature in the next version by introducing interfaces that enable interactive comparison features, as in DesignSense (Abu Zuraiq, 2020)

System components and implementation

D-Predict.v2 is developed on the Rhino3D (ref) platform using Grasshoppers and add-ons. A relatively complex system similar to D-Predict.v2 is developed in an ideal scenario following the separation of concerns. That is, (design) model and data, the logic such as operations on data, and the visualizations are decoupled in respective packages. The selection of the Rhino3D for proof of concept implementation and practical reasons constrained us from achieving complete decoupling. However, this trade-off provided a rapid prototyping opportunity with some degree of modularity in the parametric definition.

Each review session took about 1-1/2 hours and had four parts: completing consent and pre-test questionnaire to verify the participant qualifications (5-10 min), showing the system features and workflow (15 min), performing design tasks (40-60 min), and completing a semi-structured post-task interview (20 min).

Figure 5
Energy load visualization by predicting potential energy sources and consumption.

Focusing on two different performance metrics, daylight and energy load, is a deliberate choice. Our goal in this version is to establish a modular system design that allows adding or removing performance assessment modules as needed. This limited selection also helped us in the system's usability and utility evaluation by computational designers to clearly show the purpose of the system.

The daylight prediction module relies on an ML-SM we developed as a custom component and is integrated into the same parametric definition with other features. The energy load surrogate module was not ready to be added to the definition; therefore, we opted to simulate it using the Ladybug add-on. As part of the modular system architecture, other system components include shading device generation, zone definitions, material assignment, and interaction with the building elements.

EXPERT REVIEW: FORMATIVE EVALUATION OF D-PREDICT.V2

As a formative qualitative study, we evaluated the utility, usability, and adoptability of the D-Predict.v2. For utility, we tested whether it could enable designers to make data-informed decisions in the early design phases, especially regarding daylighting and energy performance. We also tailored the study to evaluate the system's adoptability by asking the expert designers if it adds value, fits the needs, improves on the current tool(s), and integrates into their current processes. We also assessed if the system is easy to use and learn.

We recruited seven volunteer expert architectural designers (with 5+ years of experience) who use parametric CAD, have sufficient data literacy to read graphs and charts and know performance metrics for daylighting and energy loads.

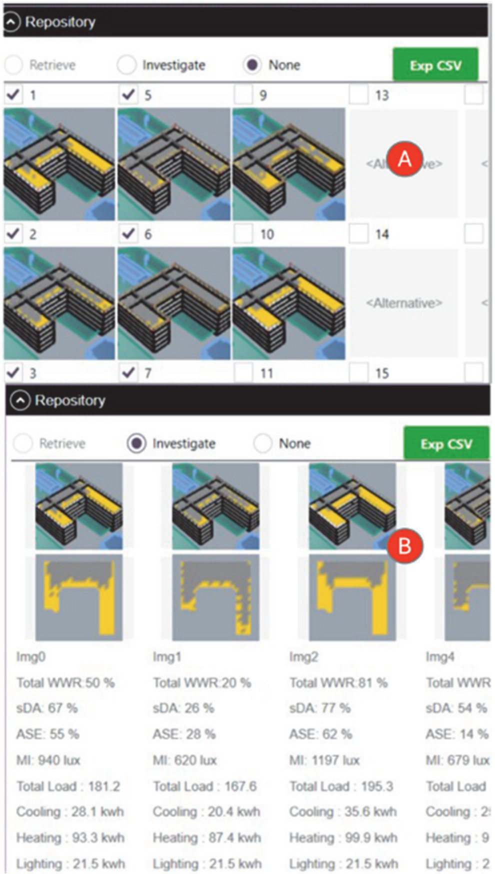

Figure 6
(A) Alternatives stored and restored persistently or (B) saved alternatives can be compared against each other for further exploration.

Participants generated about one alternative every 6 min. and saved an average of 2.4 in the repository for further review. Due to space limitations, we introduce only our initial observations below and defer sharing the details to a more comprehensive report.

Ease of Use and Speed: 7 out of 7 participants expressed positive sentiments about the tool's speed and ease of use.

- *"… I love about this is how fast this generates the data we are looking for and fast results."*
- *"The output is very well organized and very useful…simple and easy to use, too."*
- *"I appreciate how quickly it was able to generate the simulations."*
- *"…it's really fantastic to have all the data comparison and iterations all at once."*

Learnability: 5 out of 7 participants commented on the tool's clarity and learning curve.

- *"At first look, it feels a little bit overwhelming in terms of the parameters."*
- *"…after the demo, it's pretty straightforward. It doesn't have a steep learning curve."*
- *"I think [data views] pretty straightforward, like the [general] user interface."*

Tool Features: 3 out of 7 participants asked for enhanced data analysis features or improvements, e.g. adding sensitivity analysis. 4 out of 7 participants suggested more interactive data visualizations.

- *"I think the comparison is really interesting. You can see how the numbers change based on your design moves. But… It is hard to tell now which parameter will impact the result the most… quickly enough for comparisons to be made…[T]o see the impact of each parameter, [I] change one parameter at a time."*
- *"I think it was just that selecting…trying to identify which wall I wanted to look at was a bit difficult. The wall numbers were behind some geometry and were difficult to read."*
- *"[show] the result in more [interactive] graphics... there might also be…[for] analysis."*

Adaptability: The participants expressed the need to design at different levels of detail.

- *"…I think having the option to bring it into another platform would help to assess a more granular scale…Because there are so many different factors that impact design decisions, having more design freedom in the design expression would be a big factor in later early design stages."*

This note highlights that considering multiple factors is important for predicting the effect of early decisions. The current version is capable of testing a conceptual design. The designers who switch between tools asked for the ability to switch between tools:

- *"…I guess the resolution of the information is appropriate for that stage. Then, more reliable and computationally heavy analysis can be [possible] later in the process."*
- *"We use Rhino in [the] early design…and then we bring the design to Revit...Sometimes it happens side by side…in both Rhino and Revit…[This] [should] come into play."*

CONCLUSION

D-Predict.v2 is an experimental system adapting a computational design workflow we proposed combining generation and performance assessment of design alternatives using design analytics interfaces. Therefore, our main contribution is a plausible data-informed design workflow and its realization in a functional system. D-Predict.v2 is built on a modular structure that includes an ML-SM for predicting daylight analysis and the visualization of relevant data on multiple modes. Ideally, this modularity allows additional performance criteria to be added as more ML-SMs are developed. However, we have yet to know if this approach is feasible or adaptable in current practices. As a formative evaluation study, we found promising comments and feedback from expert designers. To investigate the proposed workflow further and its potential effect on building design, we will conduct real-world case studies and revisit the design of D-Predict.v2.

ACKNOWLEDMENT

This research is funded by the NSERC Discovery Grant (RGPIN-2022-03108) and MITACS Accelerate Program (IT36604) of Canada.

REFERENCES

Abu Zuraiq, A. (2020). DesignSense: A visual analytics interface for navigating generated design spaces. Master's thesis, Simon Fraser University, SIAT, Surrey, BC, Canada.

Beham, M., Herzner, W., Eduard M.G., and Kehrer, J. (2014). Cupid: Cluster-Based Exploration of Geometry Generators with Parallel Coordinates and Radial Trees. *IEEE Trans. on Vis. and Comp. Graphics, 20*(12):1693–1702.

Chegari, B., Tabaa, M., Simeu, E., Moutaouakkil, F., & Medromi, H. (2022). An optimal surrogate-model-based approach to support comfortable and nearly zero energy buildings design. *Energy, 248*, 123584.

Erhan, H., Chan, J., Fung, G., Shireen, N., & Wang, I. (2017). Understanding Cognitive Overload in Generative Design - An Epistemic Action Analysis. CAADRIA 2017

Erhan, H., Abuzuraiq, A. M., Zarei, M., Alsalman, O., Woodbury, R., & Dill, J. (2020). What do design data say about your model? A case study on reliability and validity. In *Proc. of the 25th CAADRIA* (Vol. 1, pp. 557-567).

Mottaghi, E., Abuzuraiq, A. M., & Erhan, H. (2024). D-Predict: Integrating generative design and surrogate modelling with design analytics. 29th CAADRIA, Singapore.

Fuchkina, E., Schneider, S., Bertel, S., & Osintseva, I. (2018). Design Space Exploration Framework: A modular approach to flexibly explore large sets of design variants of parametric models within a single environment. In *Proc. of 36th eCAADe 2018*.

Herbert A Simon. (1969) *The sciences of the artificial*. MIT Press.

Kastel, J. van. (2018). Visual Analytics for Generative Design Exploration: An interactive 3D data environment for a computational design system facilitating the performance-driven design process of a nearly Zero-Energy sports hall. Master's thesis, TUDelft.

Matejka, J., Glueck, M., Bradner, E., Hashemi, A., Grossman, T., & Fitzmaurice, G. (2018). Dream Lens: Exploration and Visualization of Large-Scale Generative Design Datasets. In *Proceedings of the 2018 Conference on Human Factors in Computing Systems*.

Sedlmair, M., Meyer, M., & Munzner, T. (2012). Design study methodology: Reflections from the trenches and the stacks. *IEEE Trans. on Vis. and Comp. Graphics, 18*(12), 2431-2440.

Thornton Tomasetti *Design Explorer 2*. Accessed 8 May 2023 from https://tt-acm.github.io/DesignExplorer/

Westermann, P., & Evins, R. (2019). Surrogate modelling for sustainable building design–A review. *Energy and Buildings, 198*, 170-186.

Westermann, P., Christiaanse, T.V., Beckett, W., Kovacs, P., & Evins, R. (2021). BESOS: Building and Energy Simulation, Optimization and Surrogate Modelling. *J. Open Source Softw*.

Wortmann, T., Costa, A., Nannicini, G., & Schroepfer, T. (2015). Advantages of surrogate models for architectural design optimization. *AI EDAM, 29*(4), 471-481.

Yousif, S., & Bolojan, D. (2021). Deep-Performance - Incorporating Deep Learning for Automating Building Performance Simulation in Generative Systems. In *Proc. Of 26th CAADRIA*.