# A CASE STUDY ON GENERATING ECO-CONSCIOUS OFFICE BUILDING DESIGNS USING A DATA-INFORMED OPTIMIZATION FRAMEWORK

Gen Karoji

Digital Design Lab
Nikken Sekkei
2-18-3 Iidabashi, Chiyoda-ku,
Tokyo, JAPAN
genkaroji@gmail.com

## ABSTRACT

The scrap-and-build design culture (the culture favoring knocking down and re-building buildings rather than remodeling, renovating, or re-using) has often been criticized in discussions about how to reduce the environmental impacts of office buildings. Therefore, we present a case study on generating eco-conscious office building designs, which will enable them to easily convert in the future, using our data-informed optimization framework. This framework was developed to improve the accuracy of the floor plan generation; the model that is integrated into the framework outputs the total cosine similarity between each generated floor plan and the referred plans dynamically selected from our dataset and the framework maximizes it. To generate more flexible floor plans for conversion, this paper elaborates on a way to create a dataset and formulation for optimization. Finally, we demonstrated the efficiency of the framework and described the optimized parameters that are important dimensions for eco-conscious office building designs.

**Keywords:** resilient design, floor plan generation, design optimization, association rule mining, cosine similarity.

## 1 INTRODUCTION

The scrap-and-build design culture is a unique architectural design culture in our country. This culture has been cultivated with belief that refers repeating a cycle of scraping buildings and building them is safer against natural disasters rather than repeating refurbishment and conversion to use buildings as long as possible. Recently, every building owner believes the scrap-and-build design culture is more economic. This is because they can rapidly adapt buildings for rising real estate prices and changes in demand for the building function. However, this scrap-and-build design culture has often been criticized in discussions about how to reduce the environmental impacts of buildings. In fact, according to reports from our city's government (Bureau of Environment Tokyo Metropolitan Government 2021), the office building function ranks number one regarding energy consumption and greenhouse gas emissions under its construction over many years. Additionally, a remote working paradigm shift and increasing conversion demand after the covid-19 pandemic revealed that the flexible design makes it possible to easily convert building functions. Therefore, there is no doubt creating opportunities to show eco-conscious office building designs which will enable them to easily convert in the future. Specifically, it could be effective evidence when architects persuade clients to stop fitting design (e.g. floor plan layouts, floor height, and/or hallway width) to an office function as much as possible.

To achieve the above generative design, there are three conceivable methods. One is design optimization using Multi-Objective Evolutionary Algorithm (MOEA) (Deb et al. 2002). It has still been studied, progressed well, and used to improve building performance (Wang et al. 2022). The second is consisting of fitting several rooms into an outline by using machine learning models. The third is combining both methods into one optimization framework. Firstly, design optimization using MOEA is effective in optimizing the dimensions and layout of the floor plan by maximizing building performance. In this case, algorithmic solutions of spatial allocation problems need to solve as the quadratic assignment problem using MOEA, which was studied by Jagielski and Gero, (Jagielski and Gero 1997). However, getting more accurate floor plans needs robust constraints, which could narrow diversity in solution space, or this forces users to manually explore accurate floor plans in a solution space. Secondly, the floor plan generation based on the rasterization technique is also a growing interest in applying machine learning to generative design. Chaillou tried to generate an apartment floor plan using a generative adversarial network (Chaillou 2019). This method is driven by its dataset and certain machine learning methods; it forces users to limit optimizing the building performance. In recent research, Doumpioti and Huang addressed this issue by creating a dataset composed of high performance whose data were calculated beforehand (Doumpioti and Huang 2022). However, they revealed the limitation of optimizing a design pattern that their dataset has not covered yet. Thirdly, combining into one optimization framework refers to incorporating machine learning model(s) into Multi-Objective Optimization (MOO). In this context, there are some studies for incorporating the prediction model of simulation results (e.g. daylight analyses and structure analyses). This model is generally called the surrogate model using synthetic datasets and its contribution has already been proved (Rahmani Asl et al. 2017). Additionally, there is a data-informed optimization framework that was developed by the author in another research whose main contribution about the method will not be included in this presentation, and is subject to its own publication. This framework integrates a custom-made model that contains association rule mining and the cosine similarity formula extracted from machine learning methods. To improve the accuracy in the floor plan generation, this model outputs the total cosine similarity between each generated floor plan and the referred plans in our dataset and the framework maximizes it. These referred data are dynamically selected considering features of a site configuration and site inventories (e.g. distance to neighbor buildings) using association rule mining results. Considering such situations and the building performance and a floor plan intimately relate, so combining the first and second ways into one optimization framework is a significant method to study the generation of eco-conscious office building designs. Hence, this paper demonstrated methodologies and insights by using this hybrid optimization framework.

## 2 LITERATURE REVIEWS

The current related research that referred to the development of a tool for eco-conscious building design has focused on during early design stages and it has been developed on the Building Information Modelling (BIM) environment. This is because the materials containing big impacts for costs and environmental impacts (e.g. the curtain wall as a building envelope) are generally decided by early design phases.

In a study using machine learning techniques, Budig's research (Budig et al. 2020) presents a way to predict material volumes and embodied carbon from generic design parameters in early design stages. This tool used a Bayesian neural network regression and built its model from data drawn from existing floor plans of residential buildings. They concluded that a machine learning model contributes to reducing calculation time and complementing uncertainties in the initial design. However, it has a limitation as a data-driven approach's pros and cons that do not easily explore design alternatives.

In terms of being possible to explore design alternatives, A study using MOEAs, a tool in Schwartz's research (Schwartz et al. 2021), shows an optimized set of building parameters in terms of life cycle carbon footprints and LCC. This result not only proved MOEA's contribution but also yielded a critical discussion against national compliance requirements for building performance evaluation. However, they mentioned a limitation about the response time of the optimization task that could take around 8–12h to run on a personal computer.

In the context of a tool on BIM, there are many tools developed for design optimization or simulation (Rahmani Asl et al. 2017). BIM is advertised as a coherent and comprehensive model that combines geometric and non-geometric information. As the adoption of BIM grows, its aspect of the database and infrastructure accommodating the whole architecture, engineering, construction, and operation will be more important and augmented (Koutamanis 2017). Moreover, the possibilities of integration of BIM-ready information from building product manufacturers and fabricators are shown by Afsar and Eastman (Afsari and Eastman 2014).

In this research, we adopted these ideas and followed the stream; our framework was developed on both Rhinoceros and Revit environments to be used in as many projects as possible. It addressed reducing response time to focus on early design phases and enable users to explore optimized design alternatives.

## 3 OUR FRAMEWORK

### 3.1 An Overview

This optimization framework is driven by MOEA on the MOO task. The term MOO refers to the task of searching for the best combination of inputted parameters. It is also called Combinational optimization which is oppositely distinguished from Mathematical optimization. Especially, MOO is often used in a problem whose solutions spread in several patterns. Most MOO algorithms (e.g. the Evolutionary algorithms, the Ant colony, and Simulated Annealing) use advanced stochastic search mechanisms in which some level of randomness is used. This enables the search mechanism to define the wide search space while trying to avoid “local optimum” solutions: confined areas, where no better solutions exist in immediate proximity, while better solutions may exist further away. Non-dominated Sorting Genetic Algorithm II (NSGA-II) is a well-known MOEA, widely used in many real-world applications (Deb et al. 2002). While today it can be considered as an outdated approach, NSGA-II has still great value in the purpose of maintaining diversity in solution space. Covering more than 200 optimization studies, Nguyen showed that the Genetic Algorithm (GA) is the most widely used optimization algorithm across the AEC discipline. They also concluded that among the different GA procedures, NSGA-II achieved more accurate solutions, faster, and more efficiently than the other optimization methods (Nguyen et al. 2014). In this framework, we adopted NSGA- II to maintain diversity in solutions. Regarding formulating the spatial allocation problems for optimization, a program to solve the spatial allocation problem is formulated by following a way used in the research by Jagielski and Gero (Jagielski and Gero 1997). To converge on the architecturally admissible floor plan, we modified their formula for the quadratic assignment problem. Specifically, we set constraints (e.g. maximum footprint area and rules about adjacency of rooms). In one example, if the building configuration crosses over site boundaries, the output returns an error and excludes its solution.

In terms of our custom-made model integrated into the framework, the cosine similarity formula in this model measures the total cosine similarity between each generated floor plan and the referred plans in our dataset. To promote the floor plan reproducibility and reduce response time simultaneously, then the framework maximizes this cosine similarity evaluation (CSE). In this context, the term, “reproducibility” is distinguished from the term, “accuracy” by its meaning of just reproducing a past plan. On the other hand, the term, “accuracy” refers to architectural reasonableness even applying to alternative plans which do not exist in the dataset. Cosine similarity is a measure of similarity between two sequences of numbers. This technique is often used by applications such as data mining and information retrieval in machine learning. A measured outcome is returned in positive space, where the outcome is neatly bounded in from 0 to 1. In addition, to more effectively promote floor plan reproducibility, a filtering system using association rule mining results filters these referenced data by the fitness of inputted site’s conditions. More specifically, filter item(s) (e.g. area, aspect ratio, and/or distance to neighbor buildings.) depending on the strong correlation each site inventory has are dynamically selected by taking advantage of association rule mining's findings about correlation. Association rule mining is an unsupervised method to find a frequent pattern in a dataset. This method is often employed in market basket analyses, but today this application area has been expanding into other fields. Interrogating the CumInCAD database, a search for "association rule" returned only 1 paper applying it to the digital heritage research field (Aydin et al. 2017). In this current situation, the contribution of association rule techniques to the computer-aided architectural design field has not been fully proven yet, and also this paper is among the first to investigate applying the association rule to the design optimization process.

## 3.2 Calculating Similarity

Cosine similarity is calculated as the cosine of the angle between two sequences of numbers which are viewed as two vectors in an inner product space. Hence, we need to rasterize floor plans drawn by line drawings into a grid divided by the same number of pixels and distribute numbers into each pixel based on room categories. A division of the grid should be decided by the size of a floor plan to maintain the resolution. In this case, the targeted range of floor area was 1,500–2,500 square meters, which is in the scope of a middle-scale office building. Therefore we rasterized them into 10 by 10 grids and then normalized them to a sequence of 100 numbers. For example, when an office room is the most overlapping room in a pixel, this pixel is distributed number 0. Similarly, 1 to stairs, 2 to hallways, 3 to WC (restrooms), and 4 to EV (elevators) are distributed. Note that the floor plans were aligned in the same direction (the top side is north, northwest, or northeast.) before rasterization. Sequences of numbers were created in the same way in all plans; it starts upward from the lower left and iterates this transition to the end of the upper right. Figure 1 shows a part of normalized floor plans in our dataset and a way to allocate the numbers of each room category.

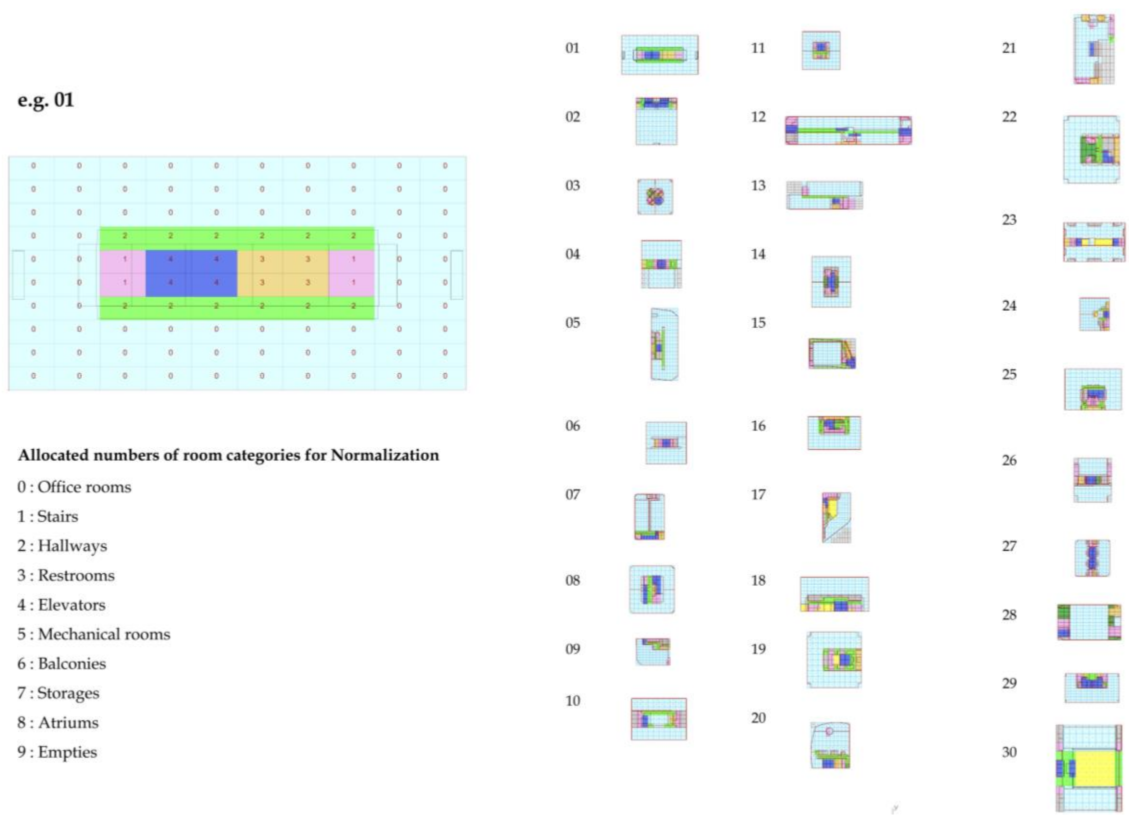



Figure 1: A part of normalized floor plans and allocated numbers of each room category.

### 3.3 Methods for Association Rule Mining

In terms of the algorithm, this research used the Apriori algorithm in association rule learning which is widely used and an easy-to-understand algorithm among association rule learning algorithms (Agrawal and Srikant 1994). The Apriori is slower than the other algorithms when the dataset is large. Although a threshold of a large dataset was 100 records in 2001 (Fernandez et al. 2001), today's experiments indicate that the Apriori is available in the range of small to large (100,000) records through subsequent improvements by Li and Sheu (Li and Sheu 2021). We assumed that the Apriori can be applied without some problems to our dataset which has currently a few hundred records but would be an even large dataset in the future. In terms of the structure of our dataset and data processing, to improve the accuracy of the floor plan, we especially focused on the location of a core that is a vertical space used for circulation and services including staircases, elevators, hallways, and restrooms. A core location is often determined by not only the size of the building area but also the surroundings to keep good orientation and views. In our country, it tends to be preferred that an office room is placed facing east or south and the core thereby faces west or north because of getting comfortable sunlight and reducing heating loads which are influenced by annual solar radiation. Therefore dataset items (in this paper the terms item and column name are used interchangeably.) are composed of horizontal projected building area, site's aspect ratio, distance to neighbor buildings of each orientation (south, north, east, and west) whose range is 0– 100 meters, and core style (central core location and side core location). Core style items took the Boolean datatype, but the others items that take the integer datatype will be converted to the Boolean datatype for association rule mining. Thus, the range, whether it will be true or false, is defined in advance. For one example, when the difference between the median and the boundary (lowest and greatest value) is defined as ±500 and the area of an inputted site is 2,000 square meters, the range returns true will be 1,500–2,500. In terms of the method of using mining results, when the data whose all data types are Boolean are prepared, the association rule mining can be executed. The results (association rules) are sorted by the lift value and then five rules from the top are used. The lift value is the terminology of an association rule and refers to the expected importance of the rule. The antecedent and the consequent are also its terminology. The antecedent is an item (column name) found within the data. The consequent is also an item found in combination with the antecedent. Both are composed of variable-length arrays of the item(s) and they are disjoint; they have no items in common. The filter item(s) that are common in the antecedents column of the above five rules are automatically selected for the filtering process. The data whose above filter item(s) are true are used in the next cosine similarity calculation.

## 4 A CASE STUDY ON GENERATING ECO-CONSCIOUS OFFICE BUILDING DESIGNS

Frameworks for evaluating the environmental impacts of buildings have been widely discussed in literature since the 1990s (Berardi, 2012). Improving a building's environmental efficiency is regarded as a challenging task that often involves an iterative process of modeling and simulation. Nevertheless, this iterative process is mostly carried out in a non-automated manner; sometimes limited time prevents architects from analyzing it more than once (Tsikos and Negendahl 2017). The improvement of the environmental efficiency of buildings is the result of the consideration of various building properties, such as the building geometry (e.g. spatial layout, aspect ratio, orientation), envelope characteristics (e.g. build-up of materials, window-to-wall ratio) or building systems (e.g. radiators, HVAC). Currently, these properties are optimized by several environmental metrics (e.g. Global Warming Potential (GWP), daylight metrics). A well-known daylight and lighting performance are mostly optimized by programming tools such as Grasshopper, and optimization and simulation plug-ins such as Ladybug and Honeybee. Hence, the suitable metric(s) for the optimization framework for building performance is often selected by the user following each problem. The economic value of building construction and operation is an important factor in decision-making processes in the Architecture, Engineering, and Construction (AEC) sector. Optimizing building designs not only by their life cycle environmental performance but also by their economic life cycle performance (multi-criteria optimization) is of great importance (Schwartz et al. 2021). In this case, we also focused on not only GWP but also Life Cycle Cost (LCC). These metrics are calculated following inputted materials that are selected by the users from our material database. Incidentally, there is research (Budig et al. 2020) that adopted the machine learning model to predict GWP. This is because GWP is difficult to quantify in the early design phases due to uncertain information regarding the resulting material volumes required.

### 4.1 Optimization Settings

In terms of MOEA settings, we defined appropriate parameters as follows through pilot studies. The population size is 15; the number of generations is 40; the mutation rate is 0.5; the crossover rate is 0.9. Figure 2 shows formulas of LCC and GWP as built-in objective functions. Figure 3 shows built-in system parameters (i.e. design variables) in the optimization process. These parameters were decided based on whether they interrelate with the office building life span, life cycle events, and costs to induce architects and clients to adopt the eco-conscious office building design. Each calculated building element was determined by whether it interrelates with the above built-in system parameters. Each room unit is assigned a specific element type (e.g. a double-skin curtain wall or Extruded Cement Panels (ECP)). Thus, the composition of building envelopes is determined by the allocated room unit type along perimeter zones. In addition, users can optionally change an element type into a specific type like ECP to Autoclaved lightweight concrete panels. Figure 4 shows typical examples of a combination of building elements and installations in a room unit. GWP, which each building element has, is calculated based on the Oekobaudat datasets. A cost of each building element and installation, which is part of LCC, is calculated based on cost estimation datasets that are monthly published by The National Construction Cost Research Institute.

**Objectives** If refurbishment hallway will be executed, n = 1. If not, n = 0

| | |
|---|---|
| ***LCCA*** | **--- Initial Cost** |
| • A suspended ceiling case : | CwC (Curtain wall costs)<br>EcC (ECP costs)<br>CeC I (Ceiling costs of an office room and a hallway)<br>CeC II (Ceiling costs of only a hallway)<br>PwC (Partition wall costs) |
| LCC difference = (CwC + EcC + PwC + CeC I ) + (ReC I * n ) + (AcC) | **--- Refurbishment and Renovation Cost** |
| • A open plenum case : | ReC I (Re-placing and new Partition wall costs)<br>ReC II (Re-placing, new Partition wall and new ceiling costs)<br>PaC (Painting Costs) |
| LCC difference = (CwC + EcC + PwC + CeC II ) + (ReC II * n + PaC * 3) + (AcC) | **--- Operating Cost**<br>AcC (Air Conditioning Costs) : |

***LCA*** ※ GWP includes Production, Transport, Installation, Waste Processing, Disposal and Recycling Potentials

| | |
|---|---|
| • A suspended ceiling case : | **--- Initial GWP**<br>CwP (Curtain wall's GWP)<br>EcP (ECP's GWP)<br>CeP I (Ceiling's GWP of an office room and a hallway)<br>CeP II (Ceiling's GWP of only a hallway)<br>PwP (Partition wall's GWP) |
| GWP difference = (CwP + EcP + PwP + CeP I ) + (ReP 1* n ) | |
| • A open plenum case : | **--- Refurbishment and Renovation GWP** |
| GWP difference = (CwP + EcP + PwP + CeP II ) + (ReP 2* n + PaP * 3) | ReP 1(new Partition wall's GWP)<br>ReP 2(new Partition wall's and new ceiling's GWP)<br>PaC (Painting GWP) |

Figure 2: Built-in system objective functions in the optimization process and its formulas.

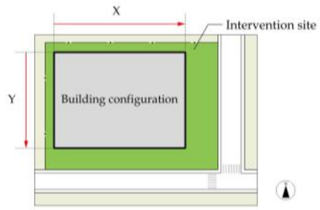
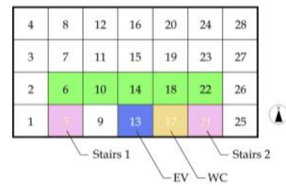
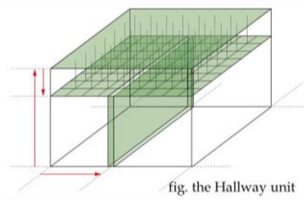
| Parameters | | |
|---|---|---|
| ***The building configuration*** | • X<br>• Y |   |
| ***The unit locations*** | • A Stairs 1 unit<br>• A Stairs 2 unit<br>• A WC unit<br>• A EV unit |   |
| ***The unit dimensions*** | • Floor Height<br>• Plenum Height<br>• Hallway Width |   |

Figure 3: Built-in system parameters in the optimization process.

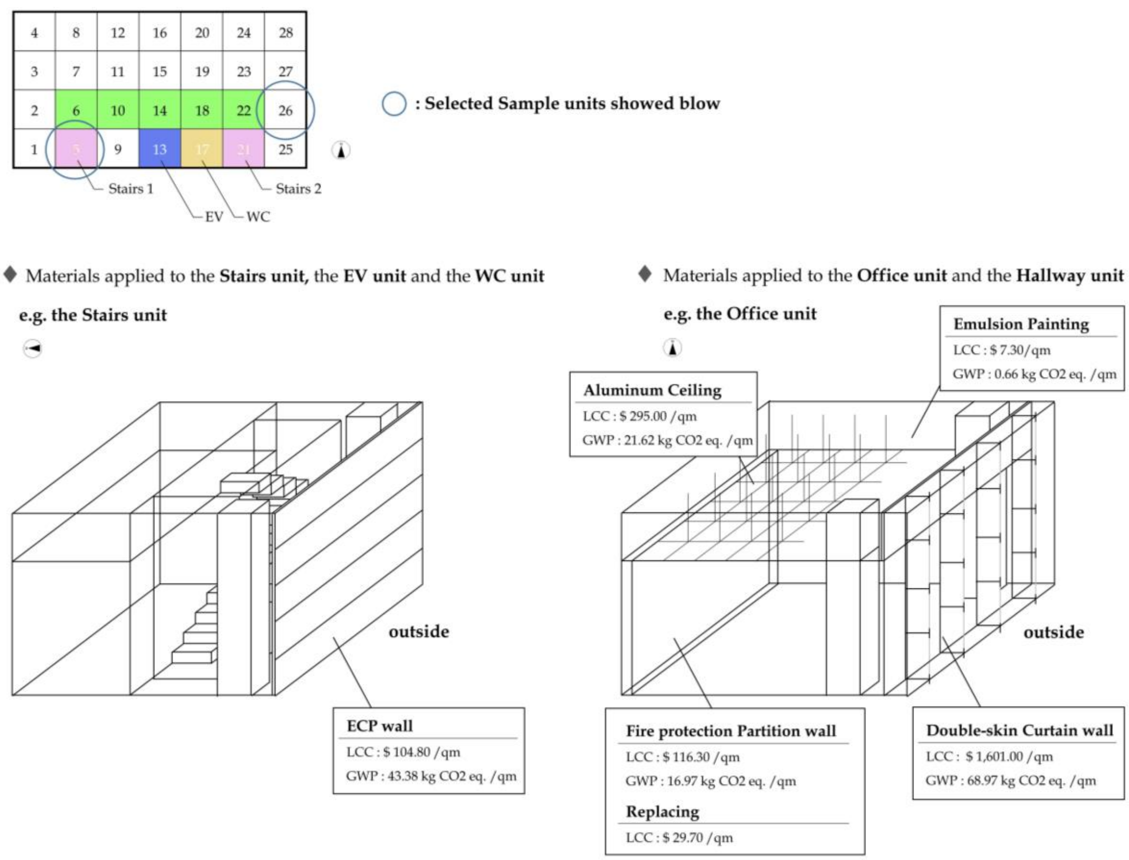



Figure 4: Examples of combination of building elements and installations in a room unit.

### 4.2 Detail of LCC Prediction

LCC contains not only initial construction costs but also refurbishment, cosmetic renovation, and operation costs. The building life span is primarily set as 60 years; users can change it within a range of 45–90 years. In terms of the conversion, whether the conversion will occur or not is predicted by our approximate model. This model was built using data from existing office building projects over the past 90 years. This data was provided by The Domestic Building Owners and Managers Association and the number of applied data was 322 records. Through building a model process, we found that floor height is an important parameter (i.e. feature) that has a strong correlation with whether a conversion is executed instead of the scrap-and-build design. Therefore, we simplified this model that returns minimum floor height as a threshold that is whether a conversion will occur in an inputted building life span. For example, if 60 years is inputted, this model returns 4.2 meters and a range of floor height parameter is automatically set as 4.2–6 meters. In terms of refurbishment, a common case of refurbishment in office buildings is widening the hallway due to changing it into a universal design. Based on the above-mentioned data, thresholds (i.e. singularities in mathematics) of width that is whether refurbishment will occur or not were defined as follows. If a hallway is straight and its width is lower than 1.8 meters, the refurbishment will be executed. Similarly, if a hallway is winding, this threshold is defined as 2.4 meters. In terms of cosmetic renovation, if the plenum height is lower than 0.5 meters, the open plenum is adopted instead of the suspended ceiling and it will be periodically repainted every 20 years to improve the aesthetics of old ducts and pipes. In terms of operation costs, it has been already revealed that the energy consumption of air conditioning interrelates with a volume of a room. Therefore, we targeted this air conditioning cost and make a formula (10.609 per square meter and year) for calculating its annual value using an average of consumed energy costs provided by The Office Management Research Lab. The life air conditioning operation cost estimation was calculated according to Equation 1 in Table 1. A variable h means ceiling height that is defined as floor height minus plenum height or floor height if their ceiling is the open plenum.

Table 1: Descriptions for estimating operation cost as a function c..

| Factor code | Description | Unit |
|---|---|---|
| h | Ceiling height | Millimeter (mm) |
| a | Office room area | Square meter (qm) |
| y | The building life span | Number of years |

$$\boldsymbol{C(h,a,y)} = \left(\frac{\boldsymbol{h}}{\mathbf{3200}}\right)^{\mathbf{0.89}} \mathbf{a10.609y} \quad \mathbf{(1)}$$

### 4.3 Association Rule Mining Settings

We set parameters (minimum lower bounds) to prune candidate rules. The support is the terminology of an association rule and refers to how frequently given item(s) appear in the dataset being mined. The term, confidence, is also its terminology and refers to a percentage value that shows how frequently the rule appears among all the rules. We set that the minimum support is 0.01 and the minimum confidence is 0.7.

### 4.4 Managing Dataset

The dataset was created using more than 300 office building projects containing life cycle information and as-built drawings. Each selected project received The Long Life Building Award in our country and was converted to some other functions. In our case study, to simply focus on the core in an office building plan, we calculated the CSE by turning the allocated numbers that are from 5 to 9 to 0 (we can see this in Figure 1.). Figure 5 shows a system diagram including the above flow and settings.

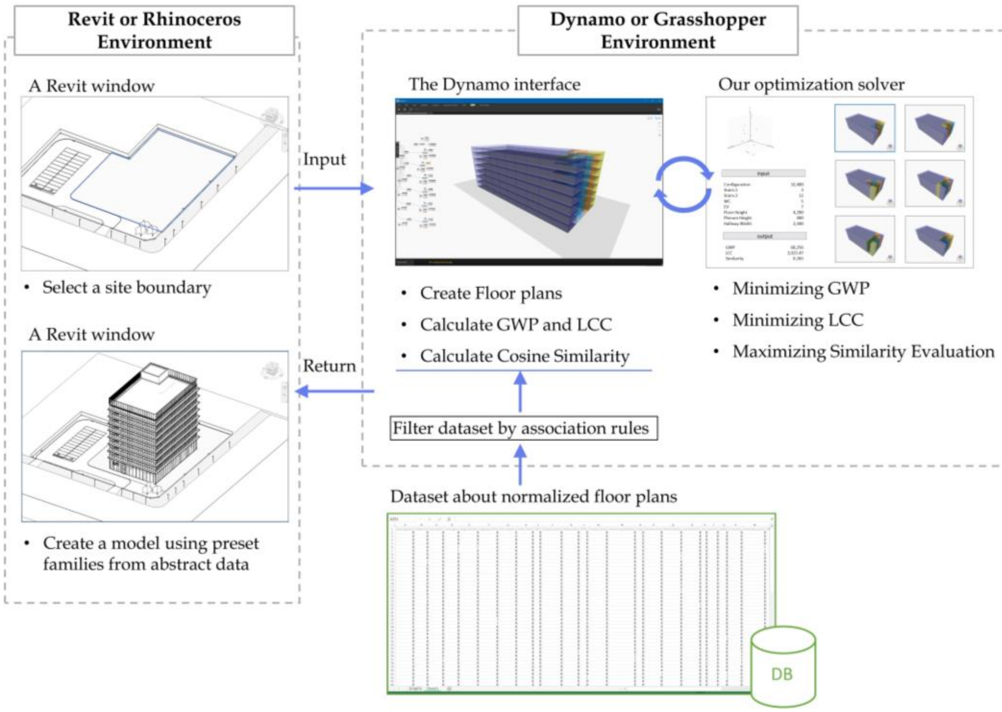



Figure 5: A system diagram including similarity evaluation and dataset filtering.

## 5 RESULTS

To prove the contribution of CSE in our custom-made model, our framework was applied to 20 case studies that have different site inventories. The performance of our model was calculated and compared through case studies.

Figure 6 shows a comparison between the most effective case (case A) and the ineffective case (case B) in terms of the CSE. Even though case B is relatively inferior to the high-quality floor plan, both cases succeeded to generate more plans that have appropriate core locations considering orientation and surroundings and also meet a regulation about maintaining the distance for evacuation than without the CSE. Specifically, the core of case A is located on the south side, not a general north side, due to existing neighbor buildings; in case B, the core located on the northwest side is generated more because the neighbor buildings exist in both north and south sides. These results show the contribution of the association rule-based filtering process. As shown by appearing central core style in solution A-1-1 in figure 6, it shows that referencing floor plans that have different core styles and locations improves not only the accuracy of floor plans but also the diversity of solutions. In contrast, the generated floor plans in case B that are not axially symmetry could have a defect for regulation about daylighting when they are converted into other building functions (the apartment, the hotel, etc.) in the future, although they offer an advantage for getting sunlight as an office function. This is because this case mainly referred to data whose core is split into north and south sides; additionally, they are composed of over three stairs and two hallways. This split floor plan seems to be excluded in a process of minimizing LCC and GWP.

In terms of the influence of the CSE, diversity, convergence, and CSE values in solution space are calculated as an arithmetic mean of all case studies. Although it is an obvious result, the mean CSE value was mathematically 5.3517 times higher than without the CSE. Diversity was calculated by the shortest Euclidean distance between each solution of a set on three-dimensional space of CSE, LCC, and GWP values. Convergence was calculated as a reciprocal in the same way, but it was on two-dimensional space excluding the CSE value from the above values. Additionally, the shortest Euclidean distance of each point in solution space was measured by using the Travelling Salesman Problem (TSP). TSP is an algorithmic challenge tasked with finding the shortest route between given a set of points. In general, it is impossible to precisely optimize all problems in one algorithm; it is necessary to select an appropriate algorithm for specific cases. In this case, a well-known brute-force algorithm completely solved our problems. As a result, diversity improved by 28.481 times higher, and convergence worsened by 0.32661 times lower than without the CSE.

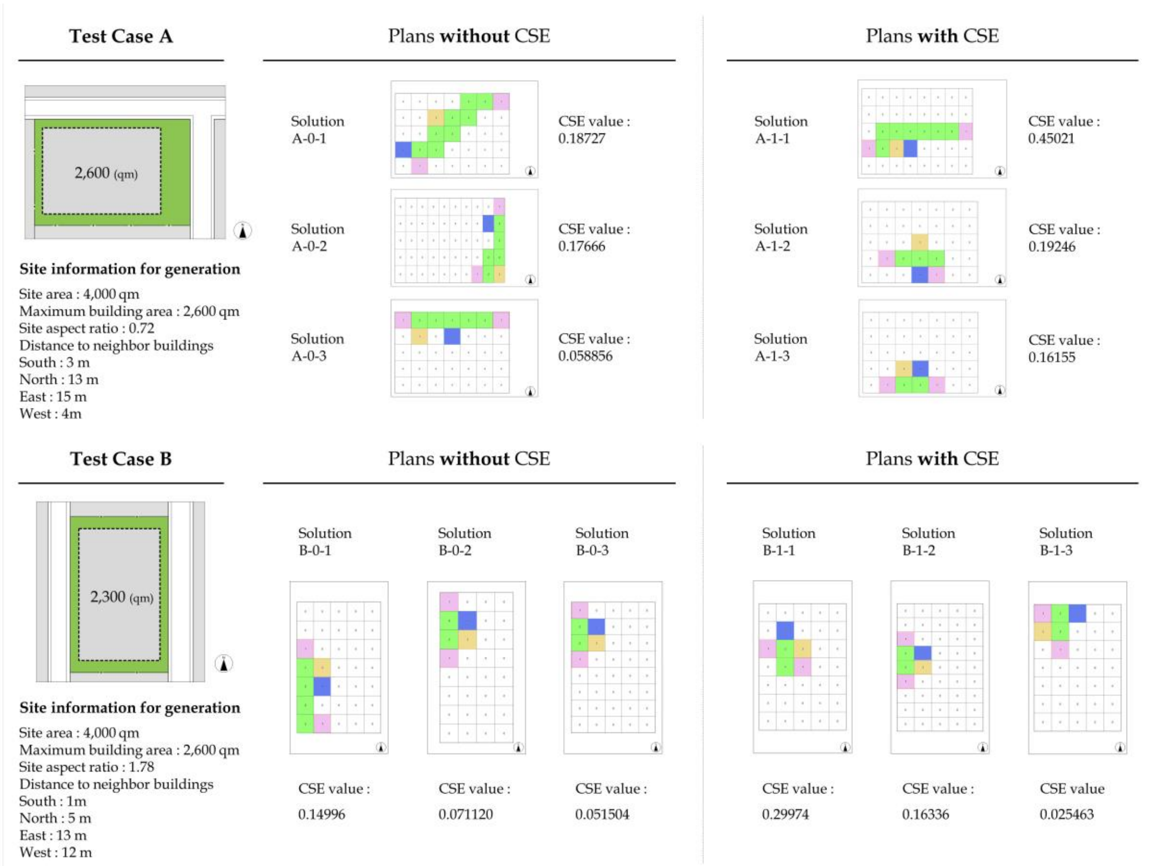



Figure 6: Comparison between the most effective case (case A) and ineffective case (case B).

Finally, we describe the results of optimized parameters that are important dimensions for eco-conscious office building designs. The floor height returned minimum values in the range determined by the building life span (if it is 60 years, floor height returns 4,200 millimeters. similarly 70 years returns 4,500 millimeters and 80 years returns 5,000 millimeters.). The hallway width returned within 2,600–3,000 millimeters in a straight hallway; if the hallway is winding, the width returned within 2,000–2,200 millimeters. The plenum height returned under 500 millimeters; the open plenum was adopted.

## 6 CONCLUSIONS AND DISCUSSION

This paper presented a way of generating eco-conscious office building designs in early design phases and especially shows a possibility that the method of our framework can be widely applied to any generative problems. This application will be possible by managing the dataset and changing design variables and objective functions. As a benefit of our custom-made model implementation, practicable initial office building design alternatives are available by running the optimization process once; total response time until reaching the end of the process (it no longer produces better solutions.) or reaching 40 generations is reduced to 5–6 minutes on a laptop personal computer. As future updates, a limitation of a way to reduce the floor plan to an orthogonal grid should be modified. To generate the floor plan on a non-orthogonal grid could be possible to optimize it by more detailed metrics like internal comfort. Additionally, we received the suggestion of adding the building life span to objective functions to maximize it. It has a trade-off relationship with LCC and GWP; it can be a more beneficial framework in terms of handling a complex problem. Finally, we expect that our framework and methods are widely used by architects and aid to clear a specific dilemma between sustainable architectural design and real estate investment.

## REFERENCES


Afsari, K., and C. Eastman. 2014. "Categorization of Building Product Models in BIM Content Library Portals". *Design in Freedom - Proceedings of the 18th International Conference on the Ibero-American Society of Digital Graphics (SiGraDi 2014)*, 1, pp. 370–374. Montevideo, Uruguay.

Agrawal, R., and R. Srikant. 1994. "Fast Algorithms for Mining Association Rules in Large Databases". *Proceedings of the 20th International Conference on Very Large Data Bases (VLDB'94), ed. Jorge B. Bocca, Matthias Jarke, and Carlo Zaniolo*, 487–499. Santiago de Chile, VLDB.

Aydin, S., A. Marc, and I. Sayah. 2017. "Association Rule Mining to Assess User - Generated Content in Digital Heritage: Participatory Content Making in 'The Museum of Gamers'". *Future Trajectories of Computation in Design. Proceedings of 17th International Conference CAAD Futures*, pp. 231–251. Istanbul, Turkey. https://doi.org/10.1007/978-981-10-5197-5_13.

Berardi, U. 2012. "Sustainability Assessment in the Construction Sector: Rating Systems and Rated Buildings". *Sustain Dev 2012*; 20(6). pp. 411–424.

Budig, M., O. Heckmann, A. Ng Qi Boon, M. Hudert, C. Lork, and L. Cheah. 2020. "Data-driven Embodied Carbon Evaluation of Early Building Design Iterations". *In CAADRIA 2020: RE: Anthropocene, Design in the Age of Humans. Proceedings of the 25th International Conference on Computer-Aided Architectural Design Research in Asia (CAADRIA)*, pp. 303–312. Bangkok, Thailand.

Bureau of Environment Tokyo Metropolitan Government. 2021. "Final Energy Consumption and Greenhouse Gas Emissions in Tokyo". https://www.kankyo.metro.tokyo.lg.jp/en/climate/index.files/Tokyo_GHG_2018.pdf. Accessed Mar. 25, 2023.

Chaillou, S. 2019. "Architecture & Style: A New Frontier for AI in Architecture". Available via https://towardsdatascience.com/architecturestyle-ded3a2c3998f. Accessed Mar. 25, 2023.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan. 2002. "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II." *IEEE transactions on evolutionary computation*, 6(2), pp. 182–197.

Doumpioti, C., and J. Huang. 2022. "Field Condition - Environmental Sensibility of Spatial Configurations with the Use of Machine Intelligence." *Co-creating the Future: Inclusion in and through Design - Proceedings of the 40th International Conference on Education and research in Computer Aided Architectural Design in Europe (eCAADe 2022)*, 2, pp. 67–74. Ghent, Belgium.

Fernandez, M., E. Menasalvas, Ó. Marbán, J. Peña, and S. Millán. 2001. "Minimal Decision Rules Based on the A Priori Algorithm." *International Journal of Applied Mathematics and Computer Science 2001* ,11(3), pp. 691–704.

Jagielski, R., and J. S. Gero. 1997. "A Genetic Programming Approach to the Space Layout Planning Problem." *Proceedings of 7th International Conference CAAD Futures*, pp. 875-884. München, Germany.

Koutamanis, A. 2017. "Briefing and Building Information Modelling: Potential for Integration.". *International journal of Architectural Computing*. 15. pp. 119–133.

Li, H., and C.-Y. Sheu Phillip. 2021. "A Scalable Association Rule Learning Heuristic for Large Datasets." *Journal of Big Data 8(1)*, pp. 1–32. https://doi.org/10.1186/s40537-021-00473-3

Nguyen, A. T., S. Reiter, and P. Rigo. 2014. "A review on simulation-based optimization methods applied to building performance analysis." *Applied Energy 2014*, 113, pp. 1043–1058. https://doi.org/10.1016/j.apenergy.2013.08.061.

Rahmani, A. M., S. Das, B. Tsai, L. Molloy, and A. Hauck. 2017. "Energy Model Machine (EMM)." *ShoCK! - Sharing Computational Knowledge! - Proceedings of the 35th International Conference on Education and research in Computer Aided Architectural Design in Europe (eCAADe 2017)*, 2, pp. 277–286. Sapienza University of Rome, Rome, Italy.

Schwartz, Y., R. Raslan, I. Korolija, and D. Mumovic. 2021. "A Decision Support Tool for Building Design: An Integrated Generative Design, Optimisation and Life Cycle Performance Approach." *International Journal of Architectural Computing* , 19(3): pp. 401–430. https://doi.org/10.1177/1478077121999802.

Tsikos, M., and K. Negendahl. 2017. "Sustainable Design with Respect to LCA Using Parametric Design and BIM Tools". *World Sustainable Built Environment Conference 2017*. Available via http://orbit.dtu.dk/files/133787517/Sustainable_Design_with_Respect_to_LCA_Using_Parametric_Design_and_BIM_Tools.pdf. Accessed Mar. 25, 2023.

Wang, L., P. Janssen, D. Bui, and K. W. Chen. 2022. "A Rapid Design Optimization Framework - Strategies for the Fast Evaluation of Design Options." *Co-creating the Future: Inclusion in and through Design - Proceedings of the 40th International Conference on Education and research in Computer Aided Architectural Design in Europe (eCAADe 2022)*, 1, pp. 619–628. Ghent, Belgium.

## AUTHOR BIOGRAPHIES

**GEN KAROJI** is a transdisciplinary researcher and architect at digital design lab, Nikken Sekkei. He holds a Master degree in Architecture from Keio University. His research interests lie in generative design, computer simulations for spatial performance and machine learning. His email address is genkaroji@gmail.com.