Published in the Proceedings of 35th Passive and Low Energy Architecture PLEA 2020 Conference, 1st-3rd September 2020, Coruña, Spain, which should be cited to refer to this work.

Analysis and Visualisation of Decision Paths towards Reaching Environmental Impact Targets at Early-Design Stage

E. NAULT[12], S. AGUACIL MORENO[12], T. JUSSELME[3]

[1]Building2050 Research Group, École polytechnique fédérale de Lausanne (EPFL), Fribourg, Switzerland.
[2]CSD ENGINEERS, Lausanne, Switzerland.
[3]Energy Institute, University of Applied Science of Western Switzerland (HEIA-FR, HES-SO), Fribourg, Switzerland.

ABSTRACT: Environmental impact objectives are commonly found in building performance labels and rating schemes. Anticipating a building's impact from the conceptual design stage and identifying decisions that do not compromise its chances of reaching these targets is therefore crucial. Yet, few methods and tools are able to provide tangible decision support through a context-specific and early-stage-oriented approach. This paper proposes a workflow to do so based on a generative approach and interactive decision trees. Illustrated on a case study, the approach consists in generating building scenarios by varying parameters not yet fixed at the early stage, including geometrical (e.g. building shape and height), architectural (e.g. façade opening ratio) and technical (e.g. heating system) parameters. The series of scenarios are evaluated in terms of their greenhouse gas (GHG) emissions over their life cycle (including construction and operation), as well as from building-induced mobility. The effects of filtering this database according to a given impact target are explored using a classification algorithm that produces a decision tree showing the proportion of target-complying and non-complying scenarios, as well as the (un)favourable decision pathways. Stakeholders of the planning and design process can therefore get insights into the implications of a given string of decisions.
KEYWORDS: Environmental impact target, decision tree, early design, decision support, life cycle assessment

1. INTRODUCTION

Performance targets related to the environmental impact of buildings over their life cycle can increasingly be found in labels and rating schemes, such as DGNB (developed in Germany) [1], LEED (that originated in the US) [2], and the Swiss SNBS [3] and 2000-Watt Site [4] certifications, the latter based on the intermediate targets for 2050 from the 2000-Watt Society concept [5, 6].

Anticipating the impact of a building from the conceptual design stage and identifying decisions that do not compromise the chance of reaching these targets is therefore crucial. The ENV1.1 criterion of DGNB for instance begins with Indicator 1: LCA [life cycle assessment] in planning, consisting in comparing the *"most likely / preferred building variants"* in terms of their potential environmental impacts over the construction and use phases [7]. This task could greatly be facilitated not only by a workflow that allows comparing building variants, which are automatically generated and assessed, but that also provides information on the implications of each design decision with respect to the environmental impact of the building.

However, assessing the life cycle performance is highly demanding in terms of time and high-resolution details, which are both limited at the early design stage [8]. In practice, few methods and tools are able to provide tangible decision support through a context-specific and early-stage-oriented approach. A majority of the existing tools, ranging from simple spreadsheets to advanced dynamic simulation software (e.g. [9, 10]), are essentially made to assess one design, as opposed to providing support towards defining this one design. They are therefore more useful at a stage in the design process where at least some of the decisions have been made, e.g. building shape and dimensions and façade details. Although results can be compared to a performance objective value (e.g. related to a given certification scheme), no information is gained on the influential design parameters or on which changes could ensure compliance.

At the European scale, the latest research suggests two promising paths towards increasing the usability of LCA in early design. First, based on parametric LCA [11], the German Caala software makes it possible to conduct a real-time assessment based on a simplified 3D model [12]. This method attributes pre-defined building components to reach a level of detail compatible with a LCA.

Another approach consists in applying LCA within a generative data-driven method to construct a knowledge database of multiple design alternatives [13]. This approach was implemented into the French Vizcab software [14], which supports multi-dimensional data exploration during the design process through parallel coordinate plots enabling a multi-dimensional exploration process [15].

Indeed, to bring insights to decision makers through a data-driven approach, not only is it necessary to generate a set of design alternatives, but this data must also be presented through meaningful and intuitive visuals. In a comparison of different data visualisation techniques applied to building performance simulation data, [16] stated that *"The Decision Tree is […] the most suitable to show the impact of parameters."*

This paper is a first attempt to use this type of visualisation, to represent the results from a workflow for supporting and informing early-stage decisions among various project stakeholders. The workflow is based on a generative approach and interactive decision trees. The objectives are to identify influential parameters and favourable versus constraining architectural and technical choices, towards supporting the elaboration of an environmental strategy from the planning stage, and the formalization of specifications, for example in the context of a master plan or architectural competition.

2. METHODOLOGY

The method consists in: 1) generating, for a case study piece of land (or lot), a series of building scenarios from a set of variable parameters (e.g. window-to-wall ratio), 2) evaluate their environmental impact, i.e. global warming potential (GWP) expressed in terms of greenhouse gas (GHG) emissions, and 3) investigate the implications of applying different GWP targets on the (un)favourable variable parameter options, through a decision tree-type of visualisation. We briefly present these three phases, with a focus on step 3. Further details on steps 1 and 2 can be found in [17], which describes the overarching applied-research project named SETUP (*Specific Environmentally-conscious Targets for Urban Planning*). The whole process is semi-automated by streamlining the modelling, simulation, and data processing steps using various software that are introduced in the following sections.

2.2 Generation of scenarios

The case study lot is located on an existing industrial site under development, located in Fribourg, Switzerland. At this stage of the project, only the maximum constructible height (of 19 m) and the building program (office) have been defined in the masterplan. To generate building scenarios on the lot, 16 variable parameters are considered in this study, including geometrical (building shape, depth and height), construction (window-to-wall ratio, glazing type, insulation level, etc.) and technical (e.g. heating system, surface of photovoltaic panels on roof) parameters. For each parameter, a set of at least two options are defined (e.g. window-to-wall ratio of 0.4, 0.65 or 0.9).

Saltelli's extension of the SOBOL sequence [18] is used to sample the space of possible scenarios; a total of 13,177 scenarios are thus generated and evaluated. Fig. 2 illustrates four example scenarios showing variations, among others, in terms of window-to-wall ratio, building dimensions, and height of the surrounding buildings (context height being one variable parameter).

A 3D model of each scenario is automatically drawn through a parametric modelling script developed in the 3D modelling software Rhinoceros (Rhino) [19] and associated graphical algorithm editor Grasshopper (GH) [20].

2.3 Evaluation of environmental impact

Impacts from the building construction and operation phases are accounted for in the calculation of its GWP. We also consider the building-induced mobility, that is, the impact from the traffic that can be attributed to the building's existence [21].

The embodied GHG emissions from materials used in the construction phase, as well as related to the systems, are estimated using data (i.e. GHG emission factors and building and component lifetime) from [22] and [23]. Emissions factors are multiplied by the corresponding element's surface, volume, or number of units, considering also replacements during the building lifetime. These simple mathematical operations are done directly in GH.

Emissions related to the energy consumption for space and water heating and for electricity (lighting, ventilation, etc.) during the operation phase are calculated through dynamic hourly simulations, via GH plugins Ladybug+Honeybee [24], which run through OpenStudio [25] and EnergyPlus [26].

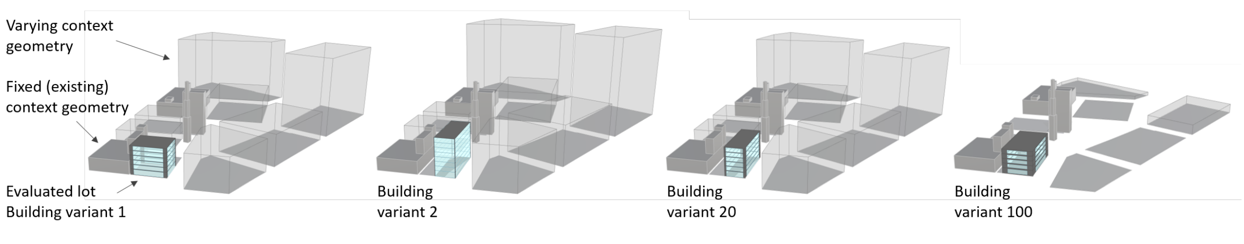


Figure 2: 3D model of case study with location of evaluated lot within the industrial site under development, and examples of the generated building design variants.

Occupancy and load profiles corresponding to an office building are mainly defined based on [27]. As mentioned, further information on simulation hypotheses can be found in [17].

Finally, the GHG emissions associated to building-induced mobility are estimated following the calculation method found in [21], directly coded in GH. Although this mobility impact here remains constant across scenarios, it is included so as to allow comparison of the total results to the Swiss SIA 2040 targets [6], which encompass all three domains (construction, operation and building-induced mobility).

The generation and evaluation process lasts in the order of a few hours for the 13,177 scenarios, the most computationally-demanding step being the dynamic hourly simulations.

2.4 Classification for decision tree

To investigate the implications of imposing a given GWP target, e.g. related to a certain building performance label, the results database is processed through a classification algorithm used to produce an interactive decision tree. Fig. 1 illustrates the structure of an example decision tree [28]. *Instances* (in our case, scenarios) are classified based on their *class* and their *feature values.* The *class* is here equivalent to the scenario's compliance to the GWP target; yes: below target, no: above target. A *feature* here corresponds to a variable parameter.

In Fig. 1, the features and their values are respectively represented by nodes (Nx) and branches (bx-y), where y is a value for parameter x.

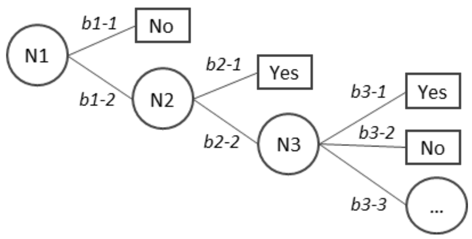


Figure 1: Decision tree structure with nodes (Nx), branches (bx-y) and resulting class (yes/no). Schema adapted from [28].

The C4.5 algorithm is here employed through the 'LearningJS' library [29], with the information gain as a splitting criterion [30]. Information gain is a measure of the level of 'purity' (same class) of the data subsets corresponding to a given feature's values. It is used recursively to determine the hierarchy of parameters that can best discriminate between complying and non-complying scenarios and that lead to the most compact tree. This visualization thus allows seeing, for the most important parameters, which of their option(s) lead to (un)favourable scenarios with respect to a given indicator target.

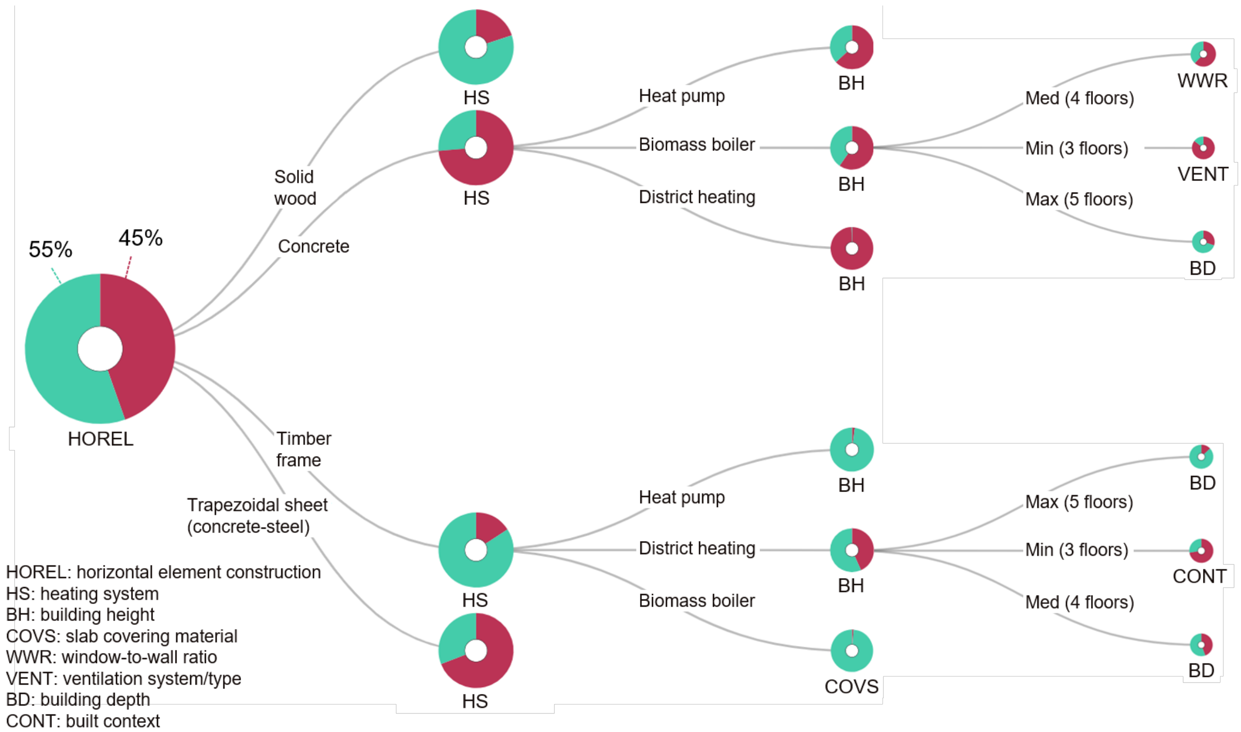
It is to note that the ordering of the features, i.e. the relative positioning of the variable parameters in the sequence of nodes, is not fixed. Indeed, as mentioned, the classification algorithm determines the hierarchy of the features and it is therefore not possible to impose a certain parameter order. As such, the resulting sequence will not necessarily follow a logical or chronological project decision stream, as will be seen in the next section.

A dedicated web app, developed in the context of the SETUP project [17], generates the decision tree (i.e. runs the classification algorithm) and produces an interactive visualization from an uploaded JSON file. This file is previously automatically built and exported by an Excel-based tool using the VBA-JSON library [31]. The web app enables users to open/close tree branches and view data on the number of (non-) complying scenarios at each node.

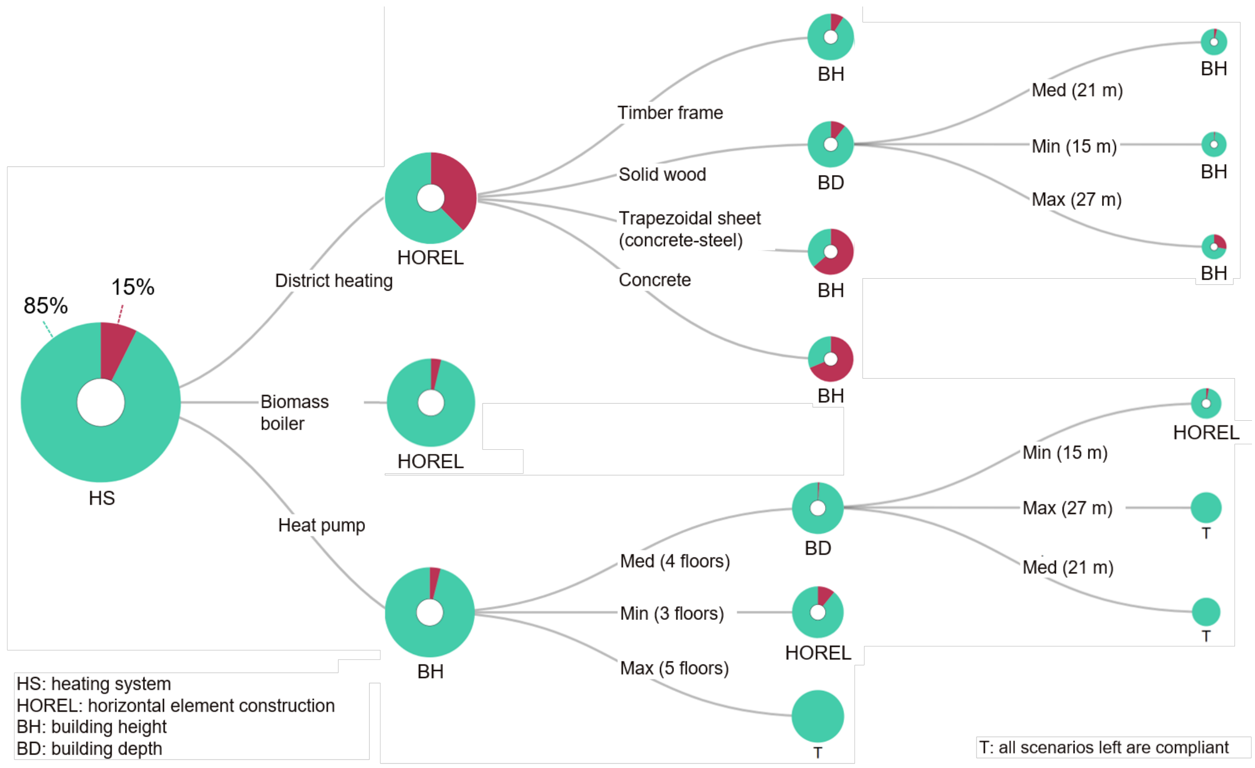
3. RESULTS AND DISCUSSION

The first few nodes and example branches of the decision tree generated from the database are shown in Fig. 3, for a GWP target of (a) 20 $kgCO_{2\text{-}eq}/m^2yr$ (corresponding to the SIA 2040 target for office buildings [6]), and for a slightly higher value of (b) 22 $kgCO_{2\text{-}eq}/m^2yr$. Each node is represented as a pie chart showing the share of complying (green) and non-complying (red) scenarios. The overall database distribution, found in the first node (root of the tree), indicates that by increasing the target by 2 $kgCO_{2\text{-}eq}/m^2yr$ (i.e. by 10%), the percentage of complying scenarios goes from 55% to 85%. This first piece of information provides insights into the feasibility of aiming for a given performance target. We could also imagine progressively reducing the target to see what GWP value could still be reachable and by which string of decisions. Although not illustrated here, feasibility results drop to 24% and 3% when decreasing the target to 18 and 16 $kgCO_{2\text{-}eq}/m^2yr$ respectively. At 15 $kgCO_{2\text{-}eq}/m^2yr$, there are no more complying scenarios (0% feasibility). It must be kept in mind that this result depends on the characteristics of the case study, including the options considered for the different variable parameters.

By comparing the two trees, we also observe that the first parameter is different. In Fig. 3(a), the type of construction for horizontal elements (roof, slabs) is at the root, whereas the heating system is found to be the most efficient classifier for tree (b). According to the target, different parameters are thus brought to the attention of the user, who is informed on the strongly influential decisions. These are captured by the tree root and following nodes and their corresponding branches.



(a) 20 kgCO$_{2\text{-}eq}$/m^{2}yr target



(b) 22 kgCO$_{2\text{-}eq}$/m^{2}yr target

Figure 3: Decision tree (root and example branches) showing at each node (option/parameter) the proportion of scenarios that have an impact below (green) and above (red) the GWP target of (a) 20 and (b) 22 kgCO$_{2\text{-}eq}$/m^{2}yr.

As mentioned, the order of the parameters is dictated by the classification algorithm and thus leads, for the presented examples, to a sequence that does not follow a standard decision flow. To allow the user to impose a certain parameter order would require a complementary visualization technique. However, this integration of parameters within a single graph provides an overview of the interplay between parameters that are typically rather treated separately and at different design stages.

In the first tree, the user has opened up the branches corresponding to a concrete construction type, where about 3/4 of scenarios are non-compliant, and a timber frame construction, for which a majority of scenarios are compliant. For both options, the second most important parameter is the heating system. However, we can see that by choosing a timber frame construction, the choice of the heating system is less constrained, as the majority of scenarios remain compliant for all options. The next node for almost all HVAC options is the building height, which, when unfolded, leads to a more diverse set of parameters (fourth tree level). The structure of the second tree shows a different amount and order of parameters.

Out of the visible branches, we can start identifying the combinations of decisions that represent favourable pathways toward achieving each target. In Fig. 3(b), we even see some paths that lead to 'pure' nodes for which all remaining scenarios are target-compliant. This is the case for instance for all 5-floors building scenarios that have a heat pump as the heating system.

The decision tree visualisation integrates diverse parameters that usually refer to different professions; urban design features related to building dimensions, architectural elements such as window-to-wall ratio, and engineering components such as heating or ventilation system. As such, it can be used as a communication support, providing an overview of the linkages between these parameters, and highlighting the importance of making coordinated decisions. This mixing of parameters that are usually addressed at different stages can however also be perceived as a limitation of the method, since the parameter order cannot be imposed.

Although no survey of statistical significance has yet been done to gather the opinion of professionals on the interactive decision tree, we can here report some preliminary qualitative feedback. The practitioners involved in the working group around the SETUP project, which included among others an architect/urban designer and an engineer, have unanimously emphasized their appreciation for the decision tree, that was singled out as particularly didactic among the various graphical outputs of the SETUP prototype tool (some of which can be seen in [17]).

4. CONCLUSION

This paper presents an interactive decision-tree visualization that maps (un)favourable decision pathways with respect to a given environmental impact target. The tree is generated via a classification algorithm applied on a pre-simulated database of building design scenarios, distinct in terms of a set of parameters that are varied across conceivable options, such as building height, type of ventilation system, and glazing type.

Producing such a database enables exploring multiple alternatives without having to go through manual iterations over the design. The only iteration is on the performance target, that can be varied to see how it affects the compliant and non-compliant decision pathways made visible by the tree.

The decision tree exploits the rich information contained in the database by offering a dynamic, intuitive and simple communication instruments that can be used within an interdisciplinary planning and design team.

ACKNOWLEDGEMENTS

This research was conducted at the Ecole polytechnique fédérale de Lausanne and co-financed by the NPR program of the State of Fribourg (project number PC-2017-03), the COGENER SIG fund (Geneva Industrial Services), and the industrial partners of the SETUP project (blueFACTORY Fribourg-Freiburg, Climate Services, CSD Engineers, Implenia and Urbaplan). We would like to thank the partners for their committed participation to this work, and Ogier Maitre, Sylvain Tauxe, Pedro Antunes and Katarina Slavkovic for their contributions.

REFERENCES

1. DGNB GmbH, (2020). DGNB System, [Online], Available: https://www.dgnb-system.de/en/ [23 April 2020].
2. U.S. Green Building Council, (2020). LEED rating system, [Online], Available: https://www.usgbc.org/leed [23 April 2020].
3. NNBS, (2020). Standard Construction durable Suisse Bâtiment (SNBS), [Online], Available: https://www.nnbs.ch/fr/snbs-batiment [23 April 2020].
4. SuisseEnergie, (2020). Site 2000 watts, [Online], Available: https://www.2000watt.swiss/fr/ [23 April 2020].
5. Frischknecht, R., Wyss, F., (2014). Concept pour l'établissement du bilan de la société à 2000 watts. SuisseEnergie, Ville de Zurich, SIA.
6. SIA, (2017). SIA 2040:2017, La voie SIA vers l'efficacité énergétique. SIA Zurich.
7. DGNB GmbH, (2018). Environmental quality ENV1.1 Building life cycle assessment. *DGNB system – New building criteria set, Version 2018*.
8. Marsh, R., (2016). LCA profiles for building components: Strategies for the early design process. *Building Research and Information*, 44: p. 358-375.

9. E4tech Software, (2020). Lesosai. http://www.lesosai.com/en/ [23 April 2020].
10. DesignBuilder Software, (2020). DesignBuilder. https://designbuilder.co.uk/ [23 April 2020].
11. Hollberg, A., (2017). A parametric method for building design optimization based on Life Cycle Assessment. Doctoral Thesis. Bauhaus-Universität Weimar.
12. Caala GmbH, (2020). CAALA Software. https://caala.de/ [23 April 2020].
13. Jusselme, T., (2020). Data-driven method for low-carbon building design at early stages. Doctoral Thesis. EPFL, Lausanne.
14. Combo Solutions, (2020). Vizcab Software. https://vizcab.io/ [23 April 2020].
15. Inselberg, A., Dimsdale, B., (1990). Parallel coordinates: a tool for visualizing multi-dimensional geometry. In *IEEE Conference on Visualization: Visualization '90*. San Francisco, USA, Oct. 23-26.
16. Jusselme, T., Tuor, T., Lalanne, D., Rey, E., Andersen, M., (2017). Visualization techniques for heterogeneous and multidimensional simulated building performance data sets. In *International Conference for Sustainable Design of the Built Environment*. London, UK, Dec. 20-21.
17. Nault, E., Jusselme, T., Aguacil, S., Andersen, M., (2020). Strategic environmental urban planning – A contextual approach for defining performance goals and informing decision-making. *Building and Environment*, 168.
18. Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M. and Tarantola, S., (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, 181 (2): p. 259-270.
19. McNeel, (2015). Rhinoceros (version 5). https://www.rhino3d.com/ [10 April 2020].
20. McNeel, (2015). Grasshopper (version 0.9.0076). https://www.grasshopper3d.com/ [10 April 2020].
21. SIA, (2016). SIA 2039:2016, Mobilité - Consommation énergétique des bâtiments en fonction de leur localisation. SIA Zurich.
22. KBOB, eco-bau, IPB, (2016). Données des écobilans dans la construction 2009/1:2016.
23. SIA, (2010). SIA 2032:2010, L'énergie grise des bâtiments. SIA Zurich.
24. Roudsari, M. S., Pak, M., Smith, A., Gill, G., (2013). Ladybug: a parametric environmental plugin for grasshopper to help designers create an environmentally-conscious design. In *13th Conference of International Building Performance Simulation Association.* Chambéry, France, Aug. 26-28.
25. Guglielmetti, R., Macumber, D., Long, N., (2011). OpenStudio: an open source integrated analysis platform. In *12th Conference of International Building Performance Simulation Association.* Sydney, Australia, Nov. 14-16.
26. Crawley, D. B., Pedersen, C. O., Lawrie, L. K., Winkelmann, F. C., (2000). EnergyPlus: Energy Simulation Program. *ASHRAE Journal* 42.
27. SIA, (2015). SIA 2024:2015, Données d'utilisation des locaux pour l'énergie et les installations du bâtiment. SIA Zurich.
28. Kotsiantis, S. B., (2007). Supervised Machine Learning: A Review of Classification Techniques. *Informatica* 31.
29. Liu, Y., (2014). LearningJS: A Javascript Implementation of Logistic Regression and C4.5 Decision Tree Algorithms, [Online], Available: https://github.com/yandongliu/learningjs [10 April 2020].
30. Quinlan, J. R., (1993). C4.5: Programs for machine learning, Morgan Kaufmann Publishers, San Mateo, California.
31. Hall, T., (2018). JSON conversion and parsing for VBA, [Online], Available: https://github.com/VBA-tools/VBA-JSON/ [10 April 2020].