PERFORMANCE OF A SEQUENTIAL VERSUS HOLISTIC BUILDING DESIGN APPROACH USING MULTI-OBJECTIVE OPTIMIZATION

Richard Gagnon[1], Louis Gosselin[1], Stéphanie Armand Decker[2]

[1]Department of Mechanical Engineering, Université Laval, Québec City, Québec, Canada G1V 0A6

[2]Nobatek-INEF4, 67 Rue de Mirambeau 64600 Anglet, France

Article accepté pour publication dans: Journal of Building Engineering, Volume 26, Novembre 2019

Abstract

Integrated design processes are currently pushed forward in order to achieve net-zero energy building designs at affordable cost. Through a case study of a residential building, this paper compares a sequential versus a holistic design approach based on multi-objective optimization. In the holistic approach, 39 design variables related to the architecture and HVAC systems are simultaneously optimized. In the sequential approach, the architecture variables are first optimized; several optimal solutions are then selected for the second phase optimization of the heating system parameters. Carbon footprint, life cycle cost and thermal comfort are optimized by the algorithm NSGA-II. With only 100 computational hours, the holistic approach found 59% of the optimal solutions, whereas it took 765 hours to find 41% of the optimal solutions with the sequential approach. This comparison shows the negative effects of making irreversible variable selections in the early phase of a design process, as it reduces the ability to find optimal solutions in the end.

Keywords: integrated design; multi-objective optimization; life cycle cost; greenhouse gases emissions, thermal comfort, holistic approach

Nomenclature

Variables

A	Surface area [m^2]
A_{eff}	Effective area of the opening [m^2]
C	Set of centroids [-]
$C(P^{(t)})$	Metric of convergence [-]
c_j	j^{th} cluster centroid [-]
$CO2$	Greenhouse gases emissions during the life cycle of the building [ton of CO_{2eq}]
d_i	Euclidean distance [-]

D_{oh}	Depth of the solar overhang [m]
env	Wall envelope type [-]
E	Theoretical heating load [MWh]
f_k	Values of the normalized k^{th} objectives function [-]
$F^{(t)}$	Non-dominated set of solution from generation t [-]
g	Acceleration due to gravity [m/s^2]
Gen	Number of generations before stopping the evolution process [-]
G_{oh}	Gap between the overhang and the window [m]
H	Opening height [m]
HS	Heating system type [-]
k	Number of centroids [-]
LCC	Life cycle cost [k$]
LPD	Thermal discomfort [%]
$\dot{m}_w$	Maximum water flow rate in the water radiator [kg/hr]
P_c	Crossover probability [-]
P_m	Mutation probability [-]
$P^{(t)}$	Population from generation t [-]
$P*$	Pareto-optimal points [-]
PMV	Predicted mean vote [-]
Pop	Population size [-]
PPD	Predicted percentage of dissatisfied person [%]
q	Heating system rated capacity [kW]
t	Generation index [-]
T_a	Outdoor ambient temperature [°C]
T_{aux}	Set point temperature to activate heat pump auxiliary power [°C]
T_{av}	Mean internal and external air temperatures [K]
T_{occ_sum}	Summer set point temperature for the occupied hours (summer) [°C]
T_{occ_win}	Winter set point temperature for the occupied hours (winter) [°C]
T_{nv_in}	No. of °C higher than T_{occ_sum} to activate natural ventilation [°C]
T_{nv_out}	Minimum ambient temperature to activate natural ventilation [°C]
T_{unoc}	Winter set point temperature set point for the unoccupied hours [°C]
T_w	Water radiator inlet temperature [°C]
T_z	Zone temperature [°C]
$u1$	Unit cost of building parts or system [$/m^2 or $/unit]
$u2$	Unit environmental impact of building parts or system [kgCO2/m^2 or kgCO2./unit]
v_i	ith data vector [-]
v_{wind}	Wind speed [m/s]
Win	Windows type [-]
WWR	Windows to wall ratio [-]
X	Set of data vectors [-]
Z	Number of zones [-]
Δt	Time step duration of simulation [hr]
ΔT	Indoor-outdoor temperature difference [°C]

Greek letters

α	Water temperature increase rate at heat emitter inlet [°C/°C]

Subscripts

aux	Heat pump auxiliary unit
bo	Boiler
bm	Building material
E	East
EE	Embodied energy
er	Electric radiator
he	Heat emitter (radiator)
HO	Holistic optimization
hp	Heat pump
Mx	Medoid number x
N	North
OE	Operational energy
S	South
SO	Sequential optimization
W	West

Acronyms

BPO	Building performance optimization
CLT	Cross laminated timber
DSW	Double stud wall
EPS	Expended polystyrene
IAQ	Indoor air quality
IDP	Integrated design process
LFW	Light frame wooden wall
TDP	Traditional design process

1 Introduction

The building sector will face severe challenges over the next decades. Building performance needs to be improved in order to reduce our dependence to fossil fuels and the greenhouse gas emissions. In order to properly address these issues, the industry needs enhanced building design methodologies. Currently, a common practice in building design is to adopt a sequential approach where several design stages are generally addressed independently and successively. For example, architects first specify the building geometry, layout and envelope. Mechanical engineers will start from that point to evaluate the energy demand and to design a suitable HVAC system. The contractor then organizes the sequence of construction. Each stakeholder involved in the design process works separately to improve different performance criteria of the project such as life cycle

cost, thermal comfort, construction time and more. Such traditional design process (TDP) can lead to irreversible choices over the different steps of a project. Another approach has been proposed over the last decade, namely the integrated design process (IDP) where all specialists are teamed up in the design process thus avoiding the sequential phases. Recent studies have highlighted the correlation between the level of interaction in a building design process and the success of a project in term of energy efficiency, cost, comfort and environmental impact [1]–[3]. It appears that the more the team is integrated, the more the project is likely to be successful. A particularly attractive and useful tool for the design process is "building optimization", which has been gaining a lot of attention over the last decade, as reviewed by Evins [4]. From the 1990's until 2010, most of the scientific papers with a focus on sustainable building design were doing single objective optimization and since that time, it is becoming more mainstream to apply a multi-objective optimization for green building design. In doing so, it is recognized that building design actually involves a series of different objectives to be met such as reducing cost, energy consumption, greenhouse gases emissions, etc., increasing comfort (visual, thermal, acoustic), structural strength, air quality, practicability with respect to circulation, etc., while integrating subjective objectives such as aesthetics or integration into urban fabric. Current multi-objective techniques and the analysis of the results that they produce generally become less practical when more than a few objectives are considered simultaneously. As a result, the formulation of a multi-objective building design problem starts by choosing an appropriate, but limited set of objectives.

1.1 Influence of design process on building performance

The emergence of building performance optimization (BPO) has highlighted new ways to efficiently design buildings. Although many methodologies, software, algorithms have been developed to accomplish BPO, it is not a common practice to use it in the early phase of a design process. Hensen has reported that BPO was too often used in the final stage of the building design process rather than in the early stage [5]. He argued that considering the interaction between sub-systems is the most efficient way to design green buildings, rather than optimizing separately those sub-systems. In Ref. [6], Gagnon et al. have used sensitivity analysis techniques to observe the influence of design variables on building performance over the different stages of a design process. The results showed that fixing variables in the early stage reduces the probabilities of finding low-energy consumption designs in the end. In Ref. [7], the authors have used multi-criteria optimization for a real building

that was going to be LEED-certified. One of the main conclusions was that fixing variables too early reduces considerably the whole optimization potential. In Ref. [8], Negendahl reviews how designers such as architects and engineers use building performance simulation in the early design stage. One of the highlights of this work is that most tools used in the early stages are too simplified for getting precise results, but they are flexible enough to allow quick modifications during that design phase. The author discusses about the most three prevailing collaborative arrangements: The engineer as an assistant to the architect, the hybrid practitioner acting both as engineer and architect and the engineer as a partner to the architect. The first one is associated to the TDP. With that method, both parties control their own domain and so their own models. Feedback for building simulation becomes hard and is generally limited to human interactions and geometric data export to the building energy simulation environment.

1.2 Optimization of the architectural variables

In Ref. [9], Pal et al. studied the influence of the envelope variables on the operational energy (OE) and the embodied energy (EE) on building optimal designs. They compared the solutions obtained when one optimizes the sum OE+EE versus OE only. The first optimization has shown results with higher U-values for the building envelope compared to the optimal solutions considering only OE. Similarly, Schwartz et al. used a multi-objective optimization algorithm to minimize lifecycle carbon and lifecycle cost of a building refurbishment [10]. They focused on the envelope variables and looked at their influence on the heating load. They tested several scenarios with two types of heating systems and with several optimization objectives. The methodology was developed in order to simplify the decision-making process in early design and has demonstrated great potential for reduction of the life cycle environmental impact of buildings. In Ref. [11], Chardon et al. used an integrated approach to building envelope design. The main goal was to optimize the cost and energy demand. They developed a methodology of automated optimization extracting information from a real product database.

1.3 Optimization of the HVAC variables

Wright et al. studied the trade-off between the HVAC system energy cost and occupant thermal comfort. They put emphasis on the variables related to the control system set points and the size of HVAC components [12]. They used a genetic algorithm with success to find a set of trade-off

optimal solutions. In another work [13], the control strategies of a variable air volume (VAV) air conditioning system are optimized. Mossolly et al. proposed two control strategies based on the temperature set point and the indoor air quality or on the predicted mean vote (PMV) set point and the indoor air quality (IAQ). Variables such as flow rate of fresh air and supply air temperature are among the design variables according to the strategy. They intend to find the optimal controls in order to simultaneously achieve thermal comfort and IAQ while reducing the energy consumption.

1.4 Sequential optimization

In Ref. [14], Verbeeck et al. have optimized the energy use, environmental impact and financial costs over the lifecycle of a building (dwellings) in a two-step process. In the first step, the envelope-related variables were optimized with respect to 3 objectives. In the second step, the 44 optimal solutions were combined with system-related measures. Hamdy et al. have used a multi-stage optimization method in order to find cost-optimal and nearly net-zero energy building solutions [15]. In stage 1, the algorithm finds the optimal combinations of building envelope and heat recovery variables. In stage 2, the space-heating energy demand for each optimal solution is combined with the heating/cooling systems and new optimal designs are obtained. Finally, in stage 3, the economical and/or environmental viability is improved by adding renewable energy systems. In Refs. [14] and [15], the sequential optimization approach is not compared with a holistic optimization. In this paper, the term "holistic optimization" means that the designers/architects/engineers deal with all the variables at the same time and consider all the sub-systems of the building as a whole.

1.5 Holistic optimization

Bichiou et al. used 3 optimization algorithms for selecting envelope and HVAC systems for residential buildings [16]. They have compared 3 approaches: HVAC optimization only, full optimization (envelope and HVAC system) and a sequential optimization. In the latter, the envelope variables are first optimized and the optimal solution is used for the second stage optimization selecting the optimal HVAC systems. This single objective optimization minimized the life cycle cost. It was found that the full optimization leads to a slightly better solution but with more computational efforts. Evins has proposed a multi-level optimization framework where building related variables and those associated to the design of the energy supply system are optimized

simultaneously and after, the operational level variables [17]. The first optimization is done with a genetic algorithm in order to determine the optimal solutions with respect to the carbon emissions and the investment cost. It also serves to determine the building energy demands. The second one optimizes the energy system variables and is done with a mixed-integer linear programming. It was concluded that this holistic approach brings solutions that exploit synergies between different parts of the design (building variables, energy systems variables). In Ref. [18], Ferrara et al. applied a cost-optimal methodology on a low-consumption residential single-family house. One of their findings is that the envelope design should not be separated from the energy system selection. Cost optimal solutions required a proper balance between investment costs and operational cost.

Although the IDP is currently pushed forward, it is not systematically adopted and the outcomes are hard to prove. None of the studies presented above focused on an exhaustive comparison between design approaches. In Ref. [16], a comparison has been done between a full and a sequential optimization, but on a single objective function.

1.6 Objective and outline of this work

Based on the above literature review and on the current needs of the industry, the objective of the present work is to provide a comparison between a holistic optimization and a sequential optimization in a building design process. This paper compares the performance of two design approaches with respect to several objective functions. First, a holistic optimization is applied to 39 design variables defining a residential building. The NSGA-II multi-objective algorithm optimizes simultaneously 3 objectives functions, related to cost, thermal comfort and greenhouse gas emissions of the building. Second, a sequential optimization is applied, at first on the architectural variables of the building and finally on the HVAC variables. In Section 2, the building that is used as a test case is described. The specificities of the optimization methods are presented in Section 3. Finally, results are reported and analyzed in Section 4.

2 Description of the building for the test case

The building studied in this work is a typical single-family home of 100 m^2 per floor. The building energy model is inspired from the one described in Ref. [19]. The design variables characterizing the building are summarized in Tables 1 and 2, and include variables related to the envelope, windows, HVAC systems and operations. In Table 1, most of the variables refer to the architectural

variables whereas Table 2 shows variables that belong to the group of HVAC systems variables. For the sake of conciseness, only the features that have been added to the TRNSYS model described in Ref. [19] are detailed below. The building is located in Québec City (Canada). The set point temperature for the occupied hours changes according to the season. From October to April, it is defined by the variable T_{occ_win} and for the rest of the year, by T_{occ_sum}. In winter, the set point temperature during the unoccupied hours T_{unoc} is also a design variable and is fixed to 16°C during summer. Three types of heating system can be selected to satisfy the heating demand: electric radiators, gas boiler with water convectors and heat pump with water convectors. If the heat pump is selected, there is an option of adding an electrical auxiliary unit. The radiators rated capacity is a design variable for each zone. This variable is dependent of the heating system type. A heat recovery unit is integrated to the ventilation system for the supply of fresh air with a sensible effectiveness of 80%. The envelope composition is represented by two design variables: the envelope type (*env*) and the windows type (w*in*). Table 3 and Table 4 detail the composition of each possible envelope and window. An overhang has been added on the south façade which introduced two new design variables: the depth of the solar overhang (D_{oh}) and the gap between the overhang and the window (G_{oh}).

Natural ventilation has been included as a strategy to reduce overheating during summer since no cooling system is installed. It consists in the opening of side hung windows in each zone when required. It is assumed that half of the window area is operable with an angle of 50°. Prediction of ventilation air change rates is a complex topic and several approaches are proposed in literature such as Computational Fluid Dynamics (CFD) [20] or empirical relations [21]. In the present study, the approach of Warren [22] has been selected with:

$$q_{stack} = \frac{1}{3} A_{eff} C_d \sqrt{\frac{\Delta THg}{T_{av}}}\,, \tag{1}$$

$$q_{wind} = 0.025 A_{eff} v_{wind}\,, \tag{2}$$

and A_{eff} as described in [23]. In the model, windows are opened, and thus the natural ventilation model is activated, when the following three conditions are met:

1. Natural ventilation is only permitted from May to September.

2. $T_z > T_{occ\_sum} + T_{nv\_in}$, i.e. the temperature in the zone is higher than [20 – 27]°C. A range exists since T_{occ_sum} and T_{nv_in} are two design variables (see Table 1). T_z is the zone temperature.

3. $T_a > T_{nv\_out}$, i.e. the outdoor air is higher than [10 – 20]°C (T_{nv_out} is also a variable).

Table 1: List of design variables (first stage).

Variable Symbol	Description and units []	Range
T_{occ_win}	Winter set point temperature for the occupied hours (winter) [°C]	[19 – 22]
T_{occ_sum}	Summer set point temperature for the occupied hours (summer) [°C]	[16 – 19]
T_{unoc}	Winter set point temperature set point for the unoccupied hours [°C]	[15 – 21]
T_{nv_out}	Minimum ambient temperature to activate natural ventilation [°C]	[10 – 20]
T_{nv_in}	No. of °C higher than T_{occ_sum} to activate natural ventilation [°C]	[4 – 8]
WWR_N	Windows to wall ratio, North façade [-]	[0.1 – 0.3]
$WWR_{E\ or\ W}$	Windows to wall ratio, East or West façade [-]	[0.2 – 0.5]
WWR_S	Windows to wall ratio, South façade [-]	[0.3 – 0.7]
env	Wall envelope type [-]	{0,1,2,…18}
win	Windows type [-] (per façade)	{0,1,2,…8}
D_{oh}	Depth of solar overhang [m]	[0 – 0.5]
G_{oh}	Gap between the overhang and the window	[0 – 0.5]

Table 2: List of design variables (second stage).

Variable Symbol	Description and unit []	Range
HS	Heating system type (heat pump = 0, boiler = 1, electric rad. = 2)	{0,1,2}
$q_{hp,}$ $q_{bo,}$ $q_{er,}$	Heating system rated capacity [kW]	{2.5,5,10,15,20,25 30}
q_{he}[a]	Radiator rated capacity [kW]	{0.25,0.5,1.0,1.5,2.0,2.5,3.0, 3.5,4.0,4.5}
$\dot{m}_{w\_bo}$, $\dot{m}_{w\_hp}$	Maximum water flow rate in the water radiator [kg/hr]	[100 – 800]
T_{w_bo}, T_{w_hp}	Water radiator inlet water temperature [°C]	[35 – 65]
α_{bo}, α_{hp}[b]	Water temperature increase rate at heat emitter inlet [°C/°C]	[0 – 1]
q_{aux}	Heat pump auxiliary power (electric) [kW]	{0,2.5,5,7.5}
T_{aux}	Set point temperature to activate heat pump auxiliary power [°C]	[-30 – -15]

[a] *Note:* q_{he} is specified individually for each zone excepted for "corridor" that has no heat emitter. As an example, q_{he_liv} is the heater rated capacity in the living room.

[b] *Note:* Monthly water temperature increase rate for every degree less than 0°C (looking at the monthly average temperature).

Table 3: Unit cost and unit environmental impact for each envelope type.

No.	Composition of the envelope[a] *env* (simplified)	U-value [W/m^2-K]	$u1_{env}$ [\$/m^2]	$u2_{env}$ [kg $CO_{2eq.}$/m^2]
0	Concrete block, mineral wool	0.305	574.8	189.7
1	Concrete block, mineral wool	0.202	576.6	197.0
2	Concrete block, mineral wool	0.151	586.0	204.3
3	Concrete block, expanded polystyrene	0.271	587.7	183.4
4	Concrete block, expanded polystyrene	0.17	607.1	186.0
5	Concrete block, expanded polystyrene	0.131	641.1	188.7
6	LFW, mineral wool (between studs)	0.320	343.4	36.5
7	LFW, mineral wool (between studs)	0.231	350.9	40.1
8	LFW, mineral wool (between studs)	0.181	366.0	44.1
9	LFW, mineral wool (between studs), EPS (outside)	0.196	353.0	41.3
10	LFW, mineral wool (between studs), mineral wool (outside)	0.175	386.5	46.4
11	LFW, mineral wool (between studs), mineral wool (outside)	0.141	422.0	52.5
12	LFW, mineral wool (between studs), mineral wool (outside)	0.118	457.5	58.6
13	CLT, polyisocyanurate	0.167	472.9	-9.8
14	CLT, expanded polystyrene	0.187	462.6	-27.2
15	CLT, mineral wool	0.177	432.2	-15.5
16	DSW filled with cellulose, EPS (outside)	0.204	359.0	24.9
17	DSW filled with cellulose, EPS (outside)	0.166	371.6	22.8
18	DSW filled with cellulose, EPS (outside)	0.140	439.2	20.8

[a]Note that for every envelope, an interior gypsum board layer and an outdoor brick layer has been considered.

Table 4: Unit cost and unit environmental impact for each window type.

No.	Composition	U-value [W/m^2-K]	$u1_{win}$ [\$/m^2]	b [\$]	$u2_{win}$ [kg CO2/m^2]
0	Double glazed LowE Argon, Wooden frame	1.27	268	289	34.12
1	Double glazed LowE Argon, Aluminum frame	1.27	187	332	47.29
2	Double glazed LowE Argon, PVC frame	1.27	150	322	44.80
3	Triple glazed LowE Argon, Wooden frame	0.59	335	361	51.18
4	Triple glazed LowE Argon, Aluminum frame	0.59	234	415	87.81
5	Triple glazed LowE Argon, PVC frame	0.59	188	402	62.53
6	Double glazed LowE Argon & Low SHGC, Wooden frame	1.29	268	289	34.12
7	Double glazed LowE Argon & Low SHGC, Aluminum frame	1.29	187	332	47.29
8	Double glazed LowE Argon & Low SHGC, PVC frame	1.29	150	332	44.80

Table 5: Unit cost and unit environmental impact for each heating system unit.

q_{bo}, q_{hp} [kW]	$u1_{HS}$ [\$/unit]		q_{he} [kW]	$u1_{he}$ [\$/unit]		$u2_{he}$ [kgCO2/unit]	
	bo	hp	water rad.[a]	water rad.[a]	elec. rad.[b]	water rad.	elec. rad.
2.5	2532	6174	0.25	93	115	15.3	2.5
5	2896	8088	0.5	185	120	30.6	5.0
10	3624	11162	1.0	370	144	61.1	10
15	4353	14235	1.5	555	195	91.7	15
20	5081	17308	2.0	762	241	137.3	20
25	5809	20383	2.5	924	339	152.8	25
30	6537	23456	3.0	1109	390	183.4	30
			3.5	1316	436	229.0	35
			4.0	1523	482	274.7	40

		4.5	1686	580	290.2	45

[a] Extra cost of 5252$ for piping and pump
[b] Extra cost of 1646 for electric wire

3 Optimization methodologies

3.1 First approach: Holistic optimization

This first approach considers all the design variables at the same time and solves a multi-objective optimization problem. In this work, this is called the "holistic optimization approach". It can be associated to the integrated design process (IDP) in which different combinations of building system options are studied by a multidisciplinary design team. In the present work, the optimization algorithm searches simultaneously for the best values for all 39 design variables (see Table 1 and Table 2), based on the following three 3 simultaneous objectives: (i) minimization of the life cycle cost (*LCC*), (ii) minimization of the greenhouse gases emissions during the life cycle of the building (*CO2*) and (iii) minimization of the thermal discomfort (*LPD*). It is important to recall that a real building design process involves more objectives then these three, several of which have been mentioned in the introduction (e.g., structural strength, visual comfort, urban integration, etc.). However, in order for the problem to be solvable, only the three above-mentioned objectives were retrained and the other objectives and their related design variables have not been assessed. The decision to focus on cost, GHG emissions and discomfort is dictated by the fact that the present work is part of an initiative on the reduction of the environmental footprint of buildings (represented hereby by GHG emissions). Since cost is always part of the decision-making process, it was also integrated in the formulation of the problem. Finally, it is required to include comfort in the optimization. Otherwise, the optimization algorithm would point toward solutions where no heating is provided because of their low cost and GHG emissions, but these solutions would not be acceptable for occupants. In the end, optimizing this set of three objectives allowed to identify families of potential solutions, as will be explained below.

The first objective sums the building material/construction cost and the operation cost related to the energy consumption:

$$LCC = Cost_{BM} + Cost_{OE} \,. \tag{3}$$

The cost of building material included only the costs that were affected by the different design variables:

$$Cost_{BM} = u1_{env} A_{env} + (u1_{win} A_{win} + b) + u1_{HS} + u1_{he} + u1_{oh} A_{oh} . \quad (4)$$

The $u1$-values associated to the building materials are unit costs per m^2 that multiply the surface area of different components (opaque envelope, windows, shading system). $u1_{HS}$ and $u1_{he}$ change depending on the heating system type and size that are chosen. The $u1$-values include the material, the labor and are representative for the localization of the building (Quebec City, Canada). They have been determined from the "Residential Cost Data RSMeans online Advanced" database [24] and are reported in Tables 3, 4 and 5. The cost related to the heating energy demand depends on the energy source. To evaluate the cost of electricity, *Hydro-Québec* "Tarif D" is selected, which applies for domestic use [25]. For natural gas cost evaluation, *Gaz Métro* "Tarif D_1 is selected [26]. This tariff applies for general distribution, mainly used for residential buildings. For the purpose of this study, the unit price of natural gas has been set as shown in Section 15.2.2.2 of Ref. [26] as for the other cost components (distribution rate, transport rate, balance rate and adjustment to inventory rate). The energy consumption is evaluated over a period of 15 years. It is assumed that no maintenance is required for such a period with a stable price for the energy consumption.

The second objective function, *CO2*, accounts for the greenhouse gas emissions (in CO_{2eq}) due to the building materials, the building construction and the energy consumption over the same period of time (15 years). During this period, it is assumed that no replacement is required. The end-use of the building materials are not considered since it is difficult to predict what will happen in the future. Finally, the building material environmental impacts (cradle to building) are representative of what is found on the Quebec market. The function *CO2* is evaluated with the following equation:

$$CO2 = CO2_{BM} + CO2_{OE} , \quad (5)$$

where $CO2_{OE}$ is determined by multiplying the heating consumption over 15 years in kWh by the emission factor of the energy source in [$kgCO_{2eq}$/kWh]. The energy needs other than heating (e.g., lighting, plug loads, etc.) have not been considered since they are not affected by the selected design variables. The emission factor for electricity in Québec City is 0.02072 $kgCO_{2eq}$/kWh [27] and

0.22567 $kgCO_{2eq}$/kWh for natural gas[1]. Furthermore, the emissions related to the building materials are obtained from:

$$CO2_{BM} = u2_{env} A_{env} + u2_{win} A_{win} + u2_{HS} + u2_{rad} + u2_{shade} A_{shade} . \quad (6)$$

The unit environmental impacts ($u2$) are listed in Tables 3, 4 and 5 and are based on Ref. [19].

Finally, the third objective function, *LPD*, is computed as defined in Ref. [28] and represents the long-term percentage of dissatisfied person. The function is based on the well-known *PPD* index [29] and is also considering the weighted area of each thermal zone:

$$LPD = \frac{\sum_{t=1}^{T} \sum_{z=1}^{Z} (PPD_{z,t} A_z \Delta t)}{A_{tot} T \Delta t}, \quad (7)$$

where

$$PPD = 100 - 95^{-0.03353 PMV^4 - 0.2179 PMV^2} . \quad (8)$$

and z is the counter for the zones, Z is the total number of zones, t is the counter for the time step over the calculation period, T is the last counter, Δt is the duration of a time step, A_z is the zone surface area and A_{tot} is the total surface area. *PPD* is the predicted percentage of dissatisfied person and is determined from the *PMV*, i.e. the predicted mean vote which is an output already defined in TRNSYS as in Ref. [30].

As mentioned above, there are other objectives functions that could have been considered, but here, the environmental impact of building construction is the main focus bearing in mind that high thermal comfort and low cost are also important.

The NSGA-II algorithm is selected to solve this holistic optimization problem as its performance has been demonstrated in a number of studies [11], [26]–[28]. In this work, it was implemented on the DEAP[2] framework [19]. It handles discrete and continuous variables, constraints and multiple objectives. The optimization parameters (e.g., population size, crossover probability, mutation probability) can have an influence on the efficiency and the performance of the method. However, it is difficult to apply common rules to set the values of those parameters [34], [35], but guidelines

[1] Athena Impact Estimator for Buildings, Version 5.1.0102

[2] DEAP is an evolutionary framework freely available in python designed for a fast prototyping of new algorithms

are available in literature. The population size is among the most important parameter as the initial population should contain a rich set of solutions. In this study, the population size ($Pop_{HO} = 192$) has been set to 5 times the number of variables, rounded to a multiple of 4 (a requirement from the tournament selection operator). The crossover probability (P_c) and the mutation probability (P_m) are set to typical values, respectively 0.9 and 0.05. The number of generations before stopping the algorithm is 50.

3.2 Second approach: Sequential optimization

In practice, the holistic approach introduced in the previous section is rarely the selected method for designing buildings. The collaboration between design teams is complex and the combination of each model on one single platform is uncommon. Moreover, the high number of variables can lead to computational times that can be very high. As discussed earlier, it is also a common practice to first fix the architectural variables and after, to integrate the HVAC systems and control strategies. Therefore, the second optimization approach treated here is a so-called sequential optimization.

3.2.1 First stage

In the first stage, the architectural variables and a few control variables (Table 1) are optimized in order to minimize *LCC*, to minimize the total theoretical heating load (*E*) and to minimize *LPD*. Although T_{occ_win}, T_{occ_sum}, and T_{unoc} are not related to architecture, they are optimized in the first stage as they are required to evaluate *E*. The objective functions are slightly different than the ones introduced for the holistic approach of Section 3.1 since in that stage, not all the information is available to calculate them. *LCC* represents the cost related to the building materials variables only. The energy cost cannot be determined at this stage since the energy source is unknown. In the same manner, the carbon footprint of the energy consumption is unknown.

3.2.2 Second stage

Once the first stage optimization is completed, a sample of solutions can be selected on the Pareto front (see Section 3.2.3) for the second stage optimization. In theory, each solution on the Pareto front of the first optimization could be used for a second stage optimization. However, this would results in an unrealistic computational time. That is why first stage solutions were clustered as explained below and only one representative design from each cluster was used for the second stage optimization. Table 2 shows the design variables for the second stage optimization. The

optimization functions are identical to the ones described in Section 3.1 for the holistic approach. The algorithm NSGA-II is used to solve the multi-objective optimization problem, for both stages, with algorithm parameters as shown in Table 6.

Table 6: Algorithm parameters.

	First stage	Second stage
Gen	100	50
Pop	64	132
P_c	0.9	0.9
P_m	0.05	0.05

3.2.3 Selecting solutions on Pareto front

In the sequential optimization approach outlined above, it is required to select designs optimized in the first step, in order to continue the optimization in the second step. In a multi-objective optimization problem, the number of non-dominated optimal solutions can be quite large and the selection process might not be straightforward, in particular given that many optimal designs can turn out to be similar. A first possible approach consists in weighting the objectives and to sum them in order to identify one single "optimal" solution. Obviously, quantifying the weight of each objective function is subjective and the choice will affect the solution selection.

Another approach consists in analyzing the Pareto front. The main idea is to select a number of representative solutions that cover well the optimal design space. Clustering techniques [36] are particularly attractive for such a task. K-means[3] is one of those algorithms and aims to partition *n* observations into *k* clusters. It has been developed by MacQueen [37]. A cluster gathers solutions that have a high degree of similarity and solutions in other clusters should have high degree of dissimilarity. In this way, the number of solutions on the Pareto front can be reduced considerably by selecting a solution closed to the centroid of the cluster, namely the medoid. Further analysis (e.g. multi-criteria analysis) can be done to select one of them, when it is needed. In this study, each medoid will become the starting point of a second stage optimization. At this stage, the variables from the first optimization step are fixed to the optimal values defined by the medoid and the variables from the second set are optimized using the NSGA-II algorithm.

[3] http://scikit-learn.org/stable/documentation.html

In Ref. [38], k-means clustering technique has been used to assist decision-makers in the selection of a few solutions on the Pareto front. In a similar manner, Zitzler had used clustering to reduce the non-dominated set of solutions in multi-objective evolutionary algorithms without destroying their characteristics [39]. Yang et al. used cluster analysis in a multi-objective optimization problem as a strategy to group the optimal solutions into a quantity of manageable and meaningful clusters [40]. They optimized in two steps an indoor sports building with respect to the daylight quantity, the energy use intensity and the mass per square meter of the structure. With cluster analysis and parallel coordinate chart, they considerably reduced the design space with 20 clusters representative of all the solutions. This first step analysis was required to select the most important parameters for the second step optimization.

The *k-means* algorithm is designed has follows. In a first step, *k* centroids are randomly generated with respect to the objective functions. Each data vector is then associated to the closest centroid to form a cluster. From each cluster, a new centroid is determined and each observation is re-assigned to the closest centroids. This simple iterative algorithm tries to minimize an error function based on the Euclidean distance between a data vector and the centroid, often referred to as the inertia [41]:

$$KM(X,C) = \sum_{i=1}^{n} \min_{j \in \{1,\dots k\}} \left\| v_i - c_j \right\|^2 , \tag{9}$$

where $v_i = i^{\text{th}}$ data vector, $c_i = j^{\text{th}}$ cluster centroid, X = set of data vectors and C = set of centroids. From each final cluster, a medoid is determined, which represents the closest real solutions to the cluster centroid.

3.3 Comparison between both optimization approaches

In order to compare the capacity of each optimization approach in finding optimal solutions, the non-dominated fronts obtained with both approaches have to be compared with the "true" Pareto-optimal points P^*. Obtaining P^* can turn out to be quite hard for the kind of problems addressed in the present work. It would require evaluating all the solutions that could exist in the design space. In this work, all the non-dominated fronts were combined in order to get a new population P_{tot} from which a new non-dominated front, estimated as P^*, is determined:

$$P_{tot} = \bigcup F_{HO} \cup \bigcup_{x=0}^{n} F_{Mx}\,, \tag{10}$$

where F_{HO} represents the non-dominated front from the holistic optimization and F_{Mx} from the sequential optimization first stage medoid number *x*. Moreover, the convergence $C(P^{(t)})$ of each optimization run is evaluated [42]. It consists in a measure of how close is a set of non-dominated solutions $F^{(t)}$ to *P**, after each generation *t*, with *t* a value between 1 and *Gen*. For each point *i* in $F^{(t)}$, the normalized Euclidean distance d_i to P^* is computed as follows:

$$d_i = \min_{j=1}^{|P^*|} \sqrt{\sum_{k=1}^{M} \left(\frac{f_k(i) - f_k(j)}{f_k^{max} - f_k^{min}} \right)^2}\,. \tag{11}$$

$f_k^{\max}$ and $f_k^{\min}$ are the maximum and the minimum values of the k^{th} objectives function in P^*. $C(P^{(t)})$ is the average of d_i for all points in $F^{(t)}$:

$$C\left(P^{(t)}\right) = \frac{\sum_{i=1}^{|F^{(t)}|} d_i}{\left|F^{(t)}\right|}\,. \tag{12}$$

In order to get a metric within [0,1], $C(P^{(t)})$ values are normalized by dividing them by $C\left(P^{(t)}\right)_{\max}$. Here, the maximal convergence values corresponds to the maximal one found in all the optimisation approaches.

4 Results

4.1 Sequential optimization - first stage

The first stage optimization of the sequential optimization approach was used to find the optimal configurations of the architectural variables with respect to 3 objectives: *E*, *LCC*, and *LPD*. 6400 solutions have been evaluated by the genetic algorithm and 650 unique solutions are located on F_{SO}. It took 80 hours to get F_{SO} with an i7-6700 Intel processor and 16 Go of RAM memory. From the very beginning of the evolution process, the algorithm was already able to find non-dominated solutions that were found later on *P** as shown by the red line in Fig. 1. Over the generations, new

optimal solutions appear on the surface of *P** and fill the least crowded zones. The evolution of the convergence metric shows that the algorithm converged quickly in the first quarter of the evolution process and then oscillates around a value of ~0.4 (see Fig. 1).

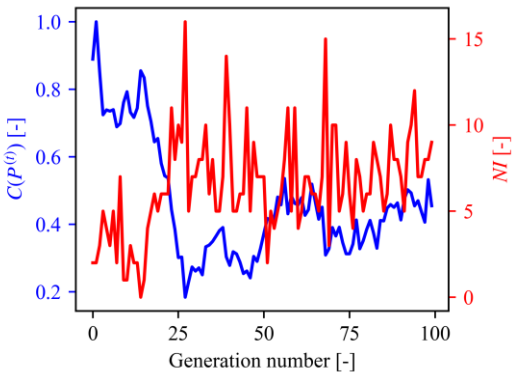


Figure 1: Convergence metric (blue) and number of new optimal solutions on the final Pareto front that have been found at each generation (in red).

Then, the k-means algorithm has been used to group those solutions into 10 clusters, each one represented by specific colors in Fig. 2. The medoids are shown by a star in the same figure.

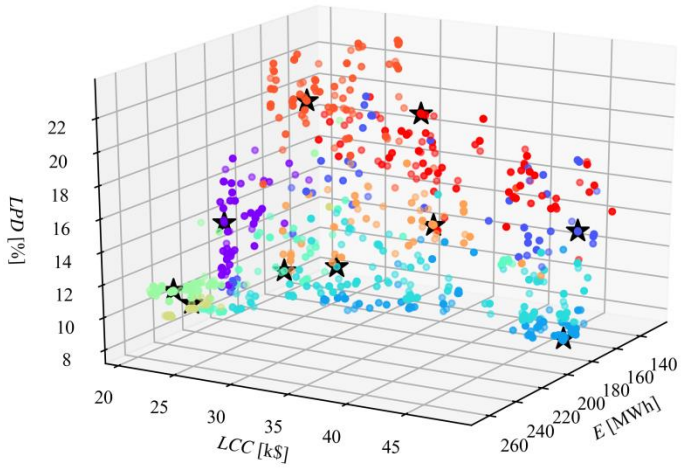


Figure 2: 10 clusters of the solutions on the Pareto front represented by a unique color with the corresponding medoid shown by a star.

Tables 7 and 8 show a detailed description of the 10 selected medoids. The analysis of the medoids shows that there is a good diversity, in terms design variables and of fitness function values. Material cost for the optimal solutions ranges from about 21.8 to 46.7 k$ and the heating energy need, from 137.3 to 243.1 MWh. Moreover, some solutions are very comfortable with an *LPD* as low as 8.8%

and sometimes, less comfortable with an *LPD* of 20.9%. Table 7 shows that different combinations of envelope, windows and *WWR* define the medoids. Some have deep overhangs and some not at all. The controls variables also show a wide diversity.

Table 7: Description of each medoid.

No.	T_{unoc}	T_{occ_sum}	T_{occ_win}	T_{nv_in}	T_{nv_out}	D_{oh}	G_{oh}	*env*	WWR_N	WWR_E	WWR_S	WWR_W	*win*
M0	16.5	17.2	20.6	4.4	16.1	0.4	1.7	18	0.24	0.41	0.43	0.26	6
M1	19.5	16.4	22.8	6.7	18.2	0.0	1.9	7	0.13	0.47	0.52	0.50	9
M2	19.0	16.4	22.8	7.3	18.2	0.4	0.0	5	0.15	0.47	0.31	0.22	4
M3	15.9	17.9	19.2	7.3	17.9	0.0	0.0	17	0.12	0.40	0.70	0.48	3
M4	16.6	16.1	19.0	7.2	18.2	1.3	1.8	5	0.11	0.47	0.70	0.21	6
M5	19	17.1	22.8	6.6	16.9	0.2	0.0	7	0.30	0.47	0.68	0.43	3
M6	19.1	16.4	19.1	5.0	16.9	0.0	0.0	11	0.14	0.45	0.53	0.49	6
M7	16.3	16.6	20.4	6.5	18.2	0.7	0.4	5	0.30	0.29	0.55	0.21	6
M8	18.9	16.2	22.9	7.2	15.6	0.4	0.0	9	0.15	0.44	0.31	0.48	6
M9	19.4	16.5	23.0	6.7	18.0	0.4	0.4	13	0.11	0.43	0.31	0.20	6

Table 8: Fitness values of each medoid.

No.	*E* [MWh]	*LCC* [k$]	*LPD* [%]
M0	149.1	31.1	12.8
M1	241.0	22.6	10.4
M2	169.4	46.7	8.8
M3	150.6	22.6	20.9
M4	243.1	21.8	11.1
M5	195.6	22.2	14.0
M6	137.3	27.5	17.2
M7	142.2	41.7	12.3
M8	185.7	26.0	9.9
M9	174.2	34.5	9.2

4.2 Comparison between the sequential and holistic optimization

The second stage optimization in the sequential optimizaion required a total of 765 hours of computation. It consists in 10 independent optimizations, each starting from one of the ten designs obtained in the first step of the optimization procedure (see Table 7). Each of these new optimization runs evaluated 6600 solutions.

Independently, the holistic optimization was also performed, varying all design variables at once. The holistic optimization evaluated 9600 solutions and required a total of 100 h of computational time. As mentioned above, the "ultimate" front *P** was achieved by combining the different solutions obtained by both optimization approaches and then, proceeding to a non-domination sorting.

The analysis of the 1008 solutions defining P^* shows that 40.8% of them were obtained by the sequential optimization and 59.2% by the holistic optimization. An animated figure shows a 3D view of the Pareto front P^* for a better understanding (see electronic annex). The color of each point is related to the optimization approach and the medoid number. Each point represents an optimal solution. Table 9 shows in detail the number of optimal solutions obtained for each independent optimization (including the holistic and the sequential optimization). In the end, some cases (i.e., $P^*_{M2}, P^*_{M4}, P^*_{M6}$ and P^*_{M7}) did not bring any of the optimal solutions. On the other hand, the medoid *M*3 was a first-stage optimized design that produced many optimal solution in the end: on *P**, 154 optimal solutions have been found from the variables selection proposed by *M*3.

Table 9: Number of optimal solutions obtained for each optimization and the required computational time.

	Holistic opt.	Sequential optimization									
	P^*_{HO}	P^*_{M0}	P^*_{M1}	P^*_{M2}	P^*_{M3}	P^*_{M4}	P^*_{M5}	P^*_{M6}	P^*_{M7}	P^*_{M8}	P^*_{M9}
No. of optimal solutions obtained for each optimization	597	48	20	0	154	0	58	0	0	71	60
Computation time [hr] (including 80 hr. for the first stage of the SO	100	156	142	154	168	162	150	166	153	155	163

The analysis of the convergence of each optimization with respect to *P** provides information about the evolution process (Fig. 3) and illustrates that the number of generation was sufficient to converge. The *M*0 optimization results show clearly that the algorithm has converged to optimal non-dominated values. The convergence stabilizes around 0.25 for many generations. The non-dominated solutions obtained from the holistic optimization are the ones giving the smallest convergence value at the end of the evolution process which is not surprising as it is the approach that has found the most optimal solutions on *P**. On the other hand, the solutions emerging from *M*2 moves away from *P** over the generations providing a convergence value around 0.75 at the end. In this case, it does not mean that the optimization has not converged, but rather that the solutions from this optimization converged in an area far from *P**. In other words, it was not a good choice to fix the variables values, for the second phase, to the values proposed by the medoid *M*2. The medoids that did not find any solution on *P** (*M*2, *M*4, *M*6 and *M*7) are the ones with the worst convergence values.

In Figs. 4, **Erreur ! Source du renvoi introuvable.** and **Erreur ! Source du renvoi introuvable.**, the non-dominated front F_{M2} (in green) is far away from the other fronts, which also confirms that

*M*2 would have been a bad selection. Note that the optimal solutions are colored based on the medoid from which they originate and are in black when they were obtained from holistic optimization.

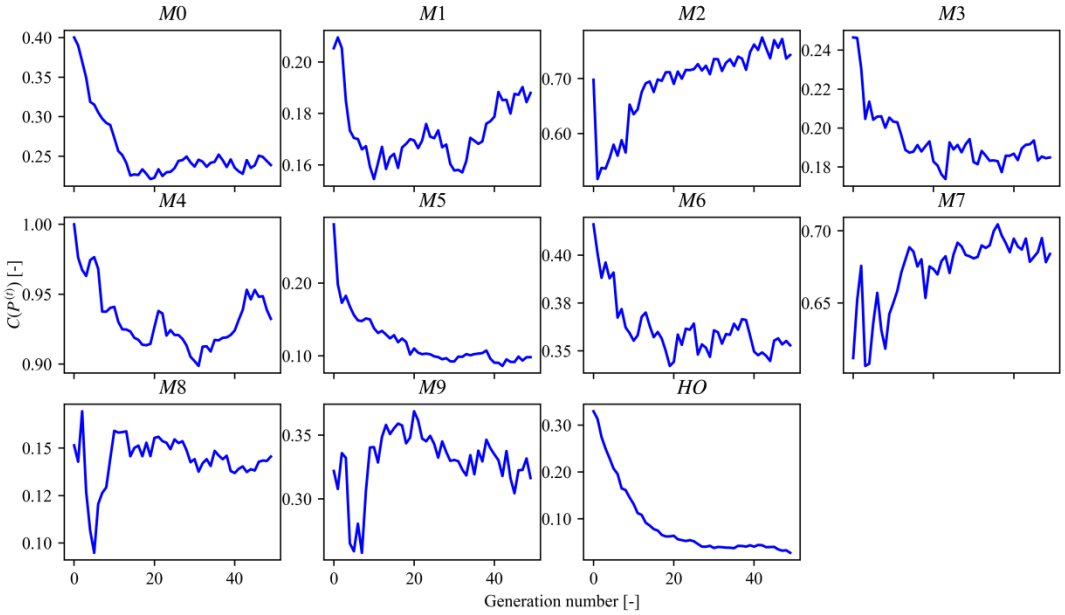


Figure 3: Convergence metric after each generation for the sequential and holistic optimization.

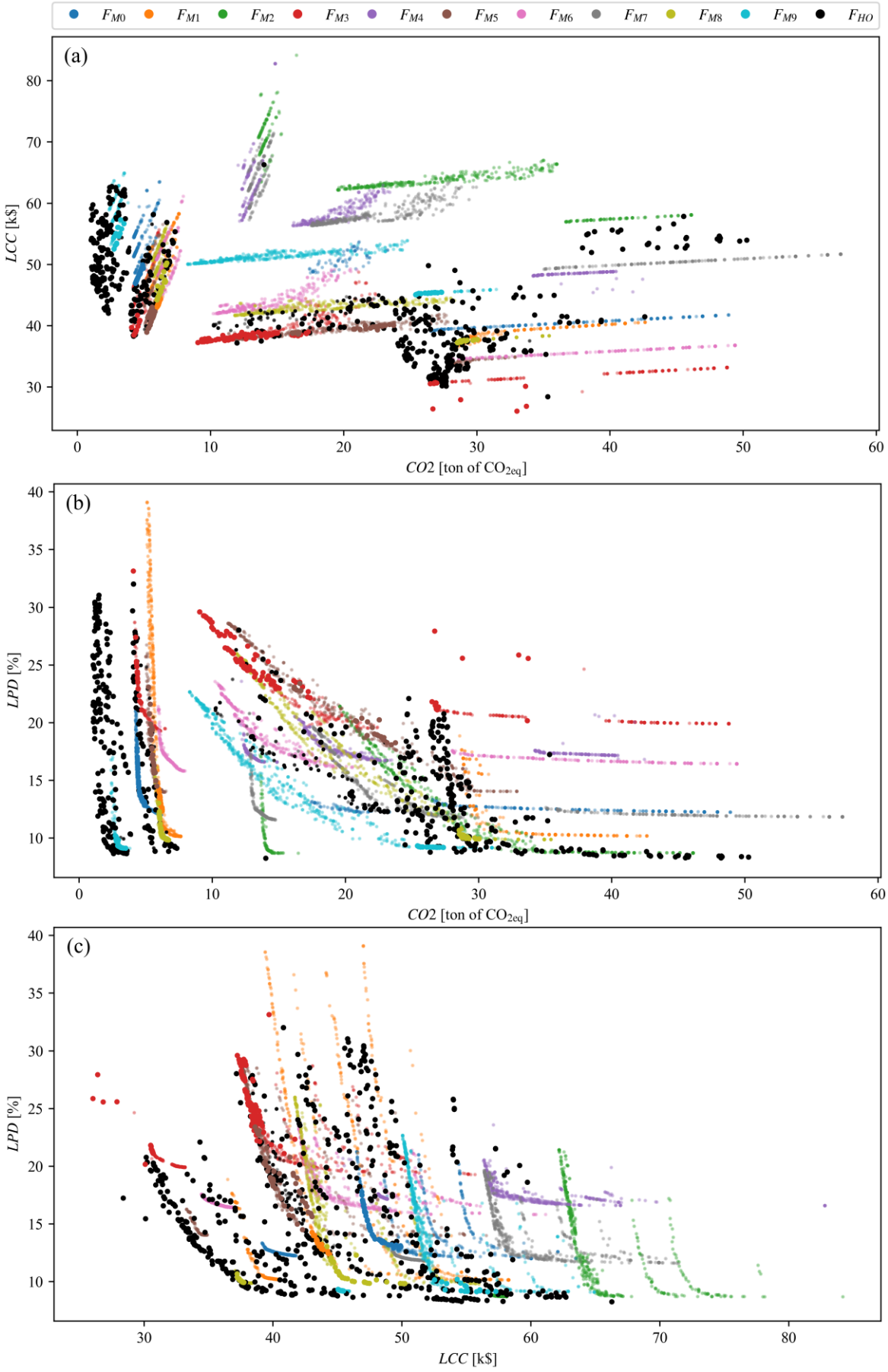


Figure 4: Non-dominated front for each optimization shown as a 2D projection for the objective functions (a) *CO2* and *LCC*, (b) *CO2* and *LPD* and (c) *LCC* and *LPD*. The bigger markers highlight the optimal solutions remaining on *P**.

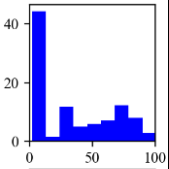
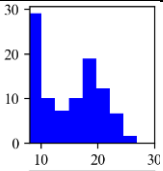
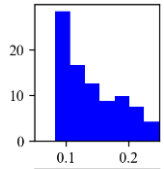
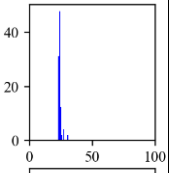
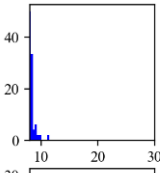
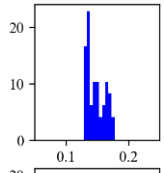
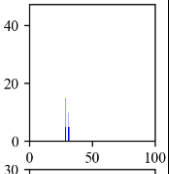
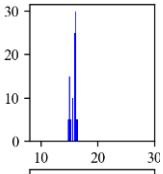
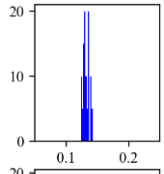
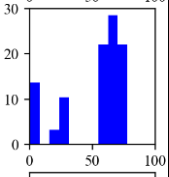
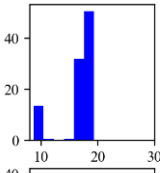
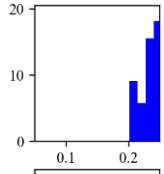
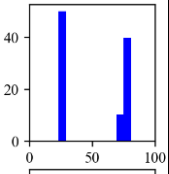
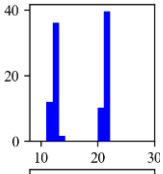
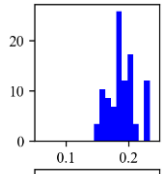
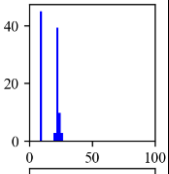
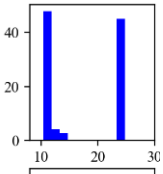
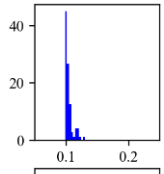
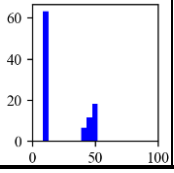
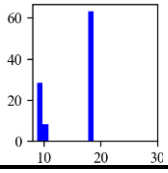
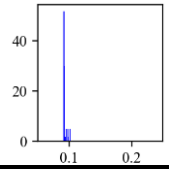
4.3 Analysis of optimal solutions

In Figs. 4a, 4b and 4c, one can observe the location of the optimal solutions obtain from the sequential optimization. They are not covering the entire design space and are "trapped" in specific areas. As opposed, the optimal solutions from the holistic optimization cover fairly well the design space (black markers). This is due to the fact that in the sequential optimization, variables have been fixed to specific values before applying the second stage optimization. On average, $CO2_{BM}$ (Eq. (6)) is responsible for 63.5% of the total CO2 emissions and $Cost_{BM}$ (Eq. (4)) for 85.1% of the life cycle cost. It means that among the optimal solutions on *P**, $CO2_{BM}$ and $CO2_{OE}$ are both important in the evaluation of *CO2*-values. It also means that the design variables with a strong influence on $Cost_{BM}$ have a lot of influence on *LCC* (e.g., envelope and window type). For example, *M*2 has selected the most expensive envelope type (641\$/m^2) with a high environmental impact (189 kg$CO_{2eq.}$/m^2). As opposed, *M*3 has an envelope type 17 with a cost of 372\$/m^2 and a very low environmental impact of 22 kg$CO_{2eq.}$/m^2. The position of those optimal solutions are in very distinct areas of the design space (green makers for *M*2 and red for *M*3). Medoid 4 and 7 also chose the envelope number 5 and Figs. 4(a),(b) and (c) show that this selection brings the optimal solutions to the highest *LCC*-values. As discussed earlier, none of the optimal solutions in $F_{M2,}$ F_{M4} F_{M6} and F_{M7} are on *P**, which means that they are all dominated by other solutions. In other words, fixing design variable values in the first phase can block the capacity to reach a certain level of optimality, with respect to the objective functions.

Table 10 shows in details the contribution of $CO2_{OE}$ and $CO2_{BM}$ to the total *CO2* as obtained with each optimization strategy that found solutions on *P**. The same analysis is done for *LCC*. Average values should be analysed carefully though, and therefore, a distribution of the frequency for the contribution of $CO2_{OE}$ to the total of *CO2* and of $Cost_{OE}$ to *LCC* is shown in the same table. The x-axis of the distribution figures has the same scale for a given objective function. It reveals that the sequential optimization produces solutions not as diverse as the holistic optimization. For example, on P^*_{M8}, almost all the solutions have a *CO2* that is mainly caused by $CO2_{BM}$ whereas on P^*_{M3}, *CO2* is mainly caused by $CO2_{OE}$.

The last column of Table 10 shows the average *LPD* with its distribution (discomfort index). There is a relation between T_{occ_win} and *LPD*. The solutions obtained from *M*1, *M*8 and *M*9 have a set point temperature during winter equal to 22.8, 22.9 and 23.0 °C respectively (the highest values found in the 10 medoids). The average *LPD* of those set of solutions is 13.2, 10.4 and 9.3% respectively. As opposed, *M*3 has T_{occ_win} = 19.2°C and the highest *LPD* with an average value of 25.6%. This relation has been confirmed after looking at a parallel coordinate plot for each solution (not shown here).

Table 10: Contribution of operational energy (heating) and building materials to *CO2* and *LCC*.

	Average contribution to the total *CO2* of:			Average contribution to the total *LCC* of:			Average of *LPD*	
	$CO2_{OE}$ [%]	$CO2_{BM}$ [%]	Distribution of $CO2_{OE}$ [%]	$Cost_{OE}$ [%]	$Cost_{BM}$ [%]	Distribution of $Cost_{OE}$ [%]	[%]	Distribution of *LPD*
P^*_{HO}	35.9	64.1		14.6	85.4		15.6	
P^*_{M0}	24.1	75.9		8.3	91.7		14.7	
P^*_{M1}	30.1	69.9		15.7	84.3		13.2	
P^*_{M3}	52.7	47.3		16.3	83.7		25.6	
P^*_{M5}	51.6	48.4		16.9	83.1		18.9	
P^*_{M8}	16.1	83.9		17.3	82.7		10.4	
P^*_{M9}	22.4	77.6		15.2	84.8		9.3	

Note: For each figure, the x-axis is related to the objective function and the y-axis is the frequency in [%].

A strong correlation exists between the heating system selection, *CO2* and *LCC*. For example, in Table 10, P^*_{M5} exhibits two distinct columns on the distribution plot. Each column gathered solutions that are defined by a specific heating system: heat pump or boiler. The solutions with a boiler increase the value of the contribution of $CO2_{OE}$ and also increase the value of *CO2* itself. In Fig. 3b, we can see that one set of solutions has a *CO2* in a range of [5.5 – 6] ton of CO_{2eq} and a second set in a range of [10 – 30] ton of CO_{2eq}. For the same set of solutions, the columns are also observed on the distribution plot of *LCC*.

Table 11 shows the number of optimal solutions using a specific heating system. Only the solutions on *P** were considered and they were traced back to the optimization that revealed them. The three types of heating systems can be observed among the optimal solutions: 18% of them use a gas boiler, 47%, a heat pump and 35%, an electric radiator.

Table 11: Number of heating systems on *P**, as a function of the origin of the optimal solutions.

	P^*_{HO}	P^*_{M0}	P^*_{M1}	P^*_{M3}	P^*_{M5}	P^*_{M8}	P^*_{M9}
Boiler	42	0	0	112	29	0	0
Heat Pump	292	48	20	22	29	39	22
E. Radiator	263	0	0	20	0	32	38

Another features of the solutions obtained with the holistic approach is the diversity of envelope selections. Thirteen (13) types of envelope are among the optimal solutions compared to 5 types of envelope achieved with the sequential optimization. Table 12 shows the number of optimal solutions that use a specific envelope number.

Table 12: Number of envelope on *P**, as a function of the origin of the optimal solutions.

env [-]	U-values [W/m^2-K]	P^*_{HO}	P^*_{M0}	P^*_{M1}	P^*_{M3}	P^*_{M5}	P^*_{M8}	P^*_{M9}	P^*

1	0.202	6	–	–	–	–	–	–	6
2	0.151	28	–	–	–	–	–	–	28
4	0.170	1	–	–	–	–	–	–	1
7	0.231	17	–	20	–	58	–	–	95
8	0.181	7	–	–	–	–	–	–	7
9	0.196	149	–	–	–	–	71	–	220
10	0.175	1	–	–	–	–	–	–	1
13	0.167	3	–	–	–	–	–	60	63
14	0.187	223	–	–	–	–	–	–	223
15	0.177	67	–	–	–	–	–	–	67
16	0.204	20	–	–	–	–	–	–	20
17	0.166	74	–	–	154	–	–	–	228
18	0.140	1	48	–	–	–	–	–	49

Except for the envelopes 1, 2 and 4 that were used in only 35 solutions, all the other selected envelopes are made from a structure using wood products. The most frequent selection is for the envelope #17 which is made from a double stud wall filled with cellulose. It offers a good balance between the price, a low U-value and mainly a low carbon footprint. Another popular envelope is #14, which is made of CLT with expanded polystyrene. The huge quantity of wood included in such a structure acts as a carbon sink, which explains the negative value in Table 3.

Figure 5 (parallel coordinates plot) highlights the relations between *env*, *HS*, q_{HS} and the objective functions. In this representation, each design is represented by a series of continuous line segments indicating the values of objective functions and design variables. In that case, a unique color is assigned for each envelope type. First, we see that the majority of the solutions are made with heating systems having a rated capacity lower than 10 kW. The solutions with a higher value of q_{HS} exist only in the presence of electric radiators. The solutions with the lowest *CO2* (~ 5 ton of CO_{2eq}) are the ones made from the envelopes 14 and 15 and using a heat pump as heating system. When the same envelopes are combined with electric radiators, *CO2* is ~ 25 ton of CO_{2eq}. The most costly solutions use a heat pump as heating system (*LCC* between 38 and 66 k$) combined with the envelope numbers 4, 14 or 15. Parallel coordinates with other design variables have been plotted with the aim of finding other relation between variables, but without success. The performance of each solution is the result of complex interactions between several variables. This highlights the pertinence of multi-objective optimization in presence of complex systems.

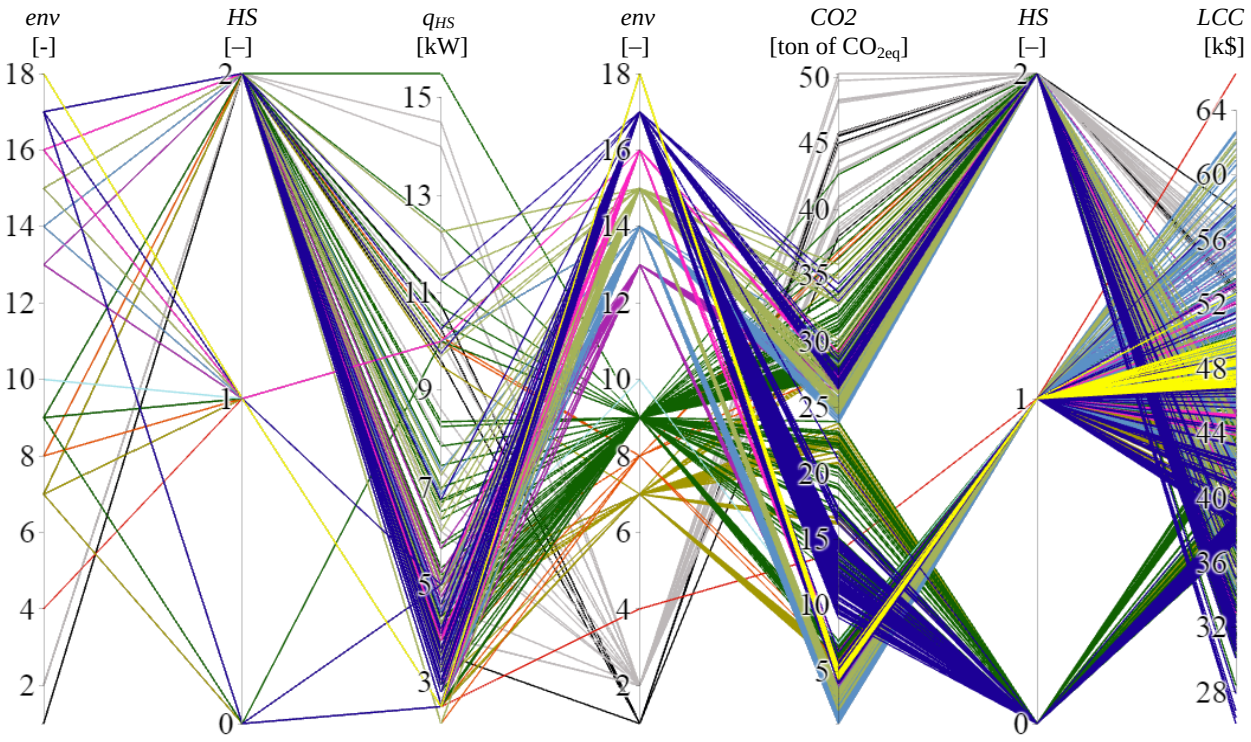


Figure 5: Parallel coordinates for all solutions on *P** with a color code related to the envelope type. Note that for clarity, *env* and HS are shown twice.

5 Conclusion

The present comparison successfully demonstrated that the sequential approach is not as performant as the holistic approach. This conclusion is valid for the computational effort, but also in terms of objective function values and diversity among the optimal solutions.

The holistic optimization succeeded to find 59.2% of the solutions on *P** in 100 h of computational time whereas the sequential optimization found 40.8% of the optimal solutions in 765 hours. The comparison shows that fixing variables in a first stage reduces the capacity of finding optimal solutions in a second stage, if the optimal solutions are compared with a holistic optimization. Moreover, the holistic optimization evaluated fewer solutions before converging to *P** (9,600 against 66,000) when compared to the sequential optimization.

Among the optimal solutions, most of them have a heat pump or electric radiators as heating system. When GHG emissions are considered as an objective function, it tends to favor heat pumps even though their initial costs are higher than those of boilers. Results are also explained by the low carbon intensity of the electricity in Québec. Different results would be observed in regions where

electricity relies on fossil fuels. Moreover, the wooden envelopes are often the best options for getting a low *CO2* value at an affordable price.

The selected period of time over which the energy consumption and greenhouse gas emissions are accounted (15 years in the present work) can affect the objective functions and different durations would highlight different optimal solutions by moving the spotlight to the initial cost/GHG (i.e., materials) or to the operational cost/GHG (i.e., energy consumption). One way to circumvent this issue could be to consider initial and operation costs as two different objectives. However, this would increase the number of dimensions of the design space and greatly complicate the analysis of the Pareto fronts.

The large amount of design parameters and design objectives directly impacts the computational time in multi-objective optimization. Although this paper demonstrates the relevance of considering all design parameters at once in a multi-objective optimization process, the complexity of this approach is obvious. One way to simplify the problem could be to perform a formal sensitivity analysis to eliminate the least influential parameters on the objective functions. Moreover, the software and tools used by different building professionals do not always communicate easily with each other, which can limit the possibility to implement a holistic approach in practice. For example, the life cycle environmental analysis platforms are rarely connected to the building simulation software. The same thing is true for the cost evaluation. Software developers should established collaborations to integrate modules to a main software that could realize effectively the optimization in practice. Different ways to use multi-objective building design optimization tools can be envisioned. On the short term, the results that they provide can be used to establish design recommendations, design guidelines, policies, etc. On the long term, research is also needed to develop effective and easy-to-use tool for architects and design engineers. In particular, simplified tools should be developed to support early adopters that would like to integrate multi-objective optimization in theirs design process (e.g., Metabuild[4]). Finally, the integrated design process is not easy to implement for other reasons: resistance to change, separation of budget between different 'hermetic' envelopes (e.g., construction, operation, etc.) and uncertainties in the early phases of the design. All those factors should be considered in order to facilitate the deployment of IDP.

[4] https://www.metabuild.io/

6 Acknowledgements

The authors are grateful to Natural Sciences and Engineering Research Council of Canada for the financial support through its ICP and CRD programs (IRCPJ 461745-12 and RDCPJ 445200-12) as well as the industrial partners of the NSERC industrial chair on eco-responsible wood construction (CIRCERB).

7 References

[1] R. Azari and Y.-W. Kim, "Development and Validation of a Framework for Evaluation of Integrated Design Teams of Sustainable High-Performance Buildings," in *Construction Research Congress 2014: Construction in a Global Network*, 2014, pp. 584–593.

[2] S. Mollaoglu-Korkmaz, L. Swarup, and D. Riley, "Delivering sustainable, high-performance buildings: Influence of project delivery methods on integration and project outcomes," *J. Manag. Eng.*, vol. 29, no. 1, pp. 71–78, 2011.

[3] S. Korkmaz, D. Riley, and M. Horman, "Piloting Evaluation Metrics for Sustainable High-Performance Building Project Delivery," *J. Manag. Eng.*, vol. 136, no. 8, pp. 877–885, 2010.

[4] R. Evins, "A review of computational optimisation methods applied to sustainable building design," *Renew. Sustain. Energy Rev.*, vol. 22, pp. 230–245, Jun. 2013.

[5] International Conference on Design & Decision Support Systems in Architecture and Urban Planning, J. P. Van Leeuwen, and H. J. P. Timmermans, Eds., *Developments in design & decision support systems in architecture and urban planning*. The Netherlands: Dept. of Architecture, Building and Planning, Eindhoven University of Technology, 2004.

[6] R. Gagnon, L. Gosselin, and S. Decker, "Sensitivity analysis of energy performance and thermal comfort throughout building design process," *Energy Build.*, vol. 164, pp. 278–294, Apr. 2018.

[7] M. Salminen, M. Palonen, and K. Sirén, "Combined energy simulation and multi-criteria optimisation of a leed-certified building," p. 6.

[8] K. Negendahl, "Building performance simulation in the early design stage: An introduction to integrated dynamic models," *Autom. Constr.*, vol. 54, pp. 39–53, Jun. 2015.

[9] S. K. Pal, A. Takano, K. Alanne, M. Palonen, and K. Siren, "A multi-objective life cycle approach for optimal building design: A case study in Finnish context," *J. Clean. Prod.*, vol. 143, pp. 1021–1035, Feb. 2017.

[10] Y. Schwartz, R. Raslan, and D. Mumovic, "Implementing multi objective genetic algorithm for life cycle carbon footprint and life cycle cost minimisation: A building refurbishment case study," *Energy*, vol. 97, pp. 58–68, Feb. 2016.

[11] S. Chardon, B. Brangeon, E. Bozonnet, and C. Inard, "Construction cost and energy performance of single family houses: From integrated design to automated optimization," *Autom. Constr.*, vol. 70, pp. 1–13, Oct. 2016.

[12] J. A. Wright, H. A. Loosemore, and R. Farmani, "Optimization of building thermal design and control by multi-criterion genetic algorithm," *Energy Build.*, vol. 34, no. 9, pp. 959–972, 2002.

[13] M. Mossolly, K. Ghali, and N. Ghaddar, "Optimal control strategy for a multi-zone air conditioning system using a genetic algorithm," *Energy*, vol. 34, no. 1, pp. 58–66, Jan. 2009.

[14] G. Verbeeck and H. Hens, "Life Cycle Optimization of Extremely Low Energy Dwellings," *J. Build. Phys.*, vol. 31, no. 2, pp. 143–177, Oct. 2007.

[15] M. Hamdy, A. Hasan, and K. Siren, “A multi-stage optimization method for cost-optimal and nearly-zero-energy building solutions in line with the EPBD-recast 2010,” *Energy Build.*, vol. 56, pp. 189–203, Jan. 2013.
[16] Y. Bichiou and M. Krarti, “Optimization of envelope and HVAC systems selection for residential buildings,” *Energy Build.*, vol. 43, no. 12, pp. 3373–3382, Dec. 2011.
[17] R. Evins, “Multi-level optimization of building design, energy system sizing and operation,” *Energy*, vol. 90, pp. 1775–1789, Oct. 2015.
[18] M. Ferrara, E. Fabrizio, J. Virgone, and M. Filippi, “Energy systems in cost-optimized design of nearly zero-energy buildings,” *Autom. Constr.*, vol. 70, pp. 109–127, Oct. 2016.
[19] R. Gagnon, L. Gosselin, S. Park, S. Stratbücker, and S. Decker, “Minimizing carbon footprint of energy and materials for a residential building: Comparison between two genetic algorithms and analysis of optimal set of solutions,” *J. Build. Perform. Simul.*, Feb. 2018.
[20] M. Regard-Alchakkif, “Ventilation naturelle d’un local par une grande ouverture ext&ieure : calcul A I’aide d’un code de champ,” p. 11.
[21] T. S. Larsen and P. Heiselberg, “Single-sided natural ventilation driven by a combination of wind pressure and temperature difference,” in *Proceedings III of 6th International Conference on Indoor Air Quality, Ventilation & Energy Conservation in Buildings: Sustainable Built Environment, Sendai, Japan*, 2007.
[22] P. R. Warren, “Ventilation through openings on one wall only,” presented at the Energy conservation in Heating, Cooling and Ventilating, 1977.
[23] M. Caciolo, P. Stabat, and D. Marchio, “Full scale experimental study of single-sided ventilation: Analysis of stack and wind effects,” *Energy Build.*, vol. 43, no. 7, pp. 1765–1773, Jul. 2011.
[24] “Residential Costs Online - RSMeans.” [Online]. Available: https://www.rsmeans.com/products/online/data-sets/residential-cost-data-online.aspx. [Accessed: 20-Mar-2018].
[25] Hydro-Québec, “Tarifs d’électricité en vigueur le 1er avril 2017.,” Hydro-Québec, Québec, 2017.
[26] Gaz Métro, “Conditions de service et Tarif au 1er octobre 2017,” Gaz Métro, Montréal, 2017.
[27] Hydro-Québec, “Taux d’émission de GES associés à l’électricité,” *Hydro-Québec*. [Online]. Available: http://www.hydroquebec.com/developpement-durable/centre-documentation/taux-emission-ges.html. [Accessed: 21-Mar-2018].
[28] S. Carlucci, G. Cattarin, F. Causone, and L. Pagliano, “Multi-objective optimization of a nearly zero-energy building based on thermal and visual discomfort minimization using a non-dominated sorting genetic algorithm (NSGA-II),” *Energy Build.*, vol. 104, pp. 378–394, Oct. 2015.
[29] P. O. Fanger, *Thermal comfort: Analysis and applications in environmental engineering*. Danish Technical Press, 1970.
[30] “TRNSYS 17 - Volume 5 Multizone building modeling with Type56 and TRNBuild.” Solar Energy Laboratory, 2012.
[31] D. Tuhus-Dubrow and M. Krarti, “Genetic-algorithm based approach to optimize building envelope design for residential buildings,” *Build. Environ.*, vol. 45, no. 7, pp. 1574–1581, Jul. 2010.
[32] E. Zitzler, K. Deb, and L. Thiele, “Comparison of multiobjective evolutionary algorithms: Empirical results,” *Evol. Comput.*, vol. 8, no. 2, pp. 173–195, 2000.

[33] A. E. I. Brownlee, J. A. Wright, and M. M. Mourshed, “A Multi-objective Window Optimisation Problem,” in *Proceedings of the 13th Annual Conference Companion on Genetic and Evolutionary Computation*, New York, NY, USA, 2011, pp. 89–90.
[34] A. Alajmi and J. Wright, “Selecting the most efficient genetic algorithm sets in solving unconstrained building optimization problem,” *Int. J. Sustain. Built Environ.*, vol. 3, no. 1, pp. 18–26, Jun. 2014.
[35] J. Wright and A. Alajmi, “Efficient Genetic Algorithm sets for optimizing constrained building design problem,” *Int. J. Sustain. Built Environ.*, vol. 5, no. 1, pp. 123–131, Jun. 2016.
[36] M. Cheikh, B. Jarboui, T. Loukil, and P. Siarry, “A Method for Selecting Pareto Optimal Solutions in Multiobjective Optimization,” p. 12.
[37] J. MacQueen, “Some methods for classification and analysis of multivariate observations,” presented at the Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1967.
[38] H. A. Taboada and D. W. Coit, “Data clustering of solutions for multiple objective system reliability optimization problems,” *Qual. Technol. Quant. Manag.*, vol. 4, no. 2, pp. 191–210, 2007.
[39] E. Zitzler and L. Thiele, “Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach,” *IEEE Trans. Evol. Comput.*, vol. 3, no. 4, pp. 257–271, 1999.
[40] D. Yang, S. Ren, M. Turrin, S. Sariyildiz, and Y. Sun, “Multi-disciplinary and multi-objective optimization problem re-formulation in computational design exploration: A case of conceptual sports building design,” *Autom. Constr.*, vol. 92, pp. 242–269, Aug. 2018.
[41] scikit-learn developers, “Scikit-learn user guide.” 2017.
[42] K. Deb and S. Jain, “Running performance Metrics for evolutionary multi-objective optimization,” Indian Institute of Technology Kanpur, Kanpur, India, Technical report 2002004, 2002.