

# P5 Report

## Humble Giants

Computational Intelligence for designing more Sustainable High-rise Buildings



Delft University of Technology  
Faculty of Architecture and the Built Environment  
MSc Track Building Technology  
Sustainable Design Graduation Studio

Tutors:	Dr. Michela Turrin	Design Informatics
	Msc. Berk Ekici	Design Informatics
	Dr. Regina Bokel	Climate Design
Delegate Examiner:	Dr. Harry Boumeester	

**Personal Information:**

Name: Fredy Fortich

Student number: 4821858

Telephone number: +

E-mail address:

## Contents

1.0 Introduction.....	6
1.1 Scope.....	7
1.2 Methodology.....	7
2.0 Literature Review.....	9
2.1 High-Rise Typology.....	9
2.2 Building Physics.....	17
2.3 Performance-Based Computational Design.....	27
2.4 Surrogate Modeling & Machine- Learning.....	34
3.0 Method.....	42
3.1 General Workflow.....	43
3.2 Overview of the Parametric Model.....	43
3.3 Validation of Simulation Model.....	54
3.4 Collecting samples.....	61
3.5 Training the Surrogate model.....	65
3.6 Evaluation of the Surrogate Model.....	66
3.7 Computational Optimization.....	75
3.8 Surrogate model in the Design Process.....	95
3.9 Discussion.....	96
4.0 Showcase of a design solution.....	99
4.4 Proposal.....	99
5.0 Conclusions.....	106
6.0 Reflection.....	112
6.1 Main topics.....	113
6.2 Aim.....	113
6.3 Limitations.....	114
6.4 Further Steps.....	114
7.0 Bibliography.....	115
8.0 Appendix.....	121

## Abstract

As urbanization increases around the world, high-rise buildings will continue to become a more prevailing typology, nonetheless, due in part to cumbersome computational simulations, rarely do designers have enough information during the early stages of design, which is the time when their choices affect the most the efficiency of their building. Surrogate models, aka meta-models that predict how the original simulation models behave offer a clear advantage in terms of speed of the results. This study delves into performance-based design using surrogate models to give the designer a tool to quickly understand the variables that will affect its efficiency. Looking specifically to improve four (4) results: energy consumption, natural daylight, comfort, and floor area. This study contemplates 16 unique variables ranging from effects of the Context (1), general building shape & orientation (6) to façade variables (9). The energy results are validated in DesignBuilder software before proceeding to collect 500 samples for two different locations: Bogotá and Amsterdam. This data is then run through three machine learning methods, Multilinear Regression, Non-linear Regression, and ANN. Next, the chosen ANN-based surrogate models for each of the outcomes are trained and hyperparameters finetuned to increase their R2 value and reduce their standard error (MSE) and mean absolute error (MAE). Finally, the generic surrogate models are run and compared through various optimization algorithms to determine Pareto-frontier options that ultimately improve the energy performance of a solution with the daylight, comfort, and floor area as design constraints or goals. A time reduction of up to 99.96% was achieved to collect another 500 samples. Finally, the final model also serves as an aid for visualization of the design space by allowing near-real-time (6 seconds) to generate the form of each design solution

## Keywords

High-Rise, Performance-based design, Surrogate Model, Energy Efficiency, Energy Use Intensity (EUI), Daylight, Useful Daylight Illuminance (UDI), Comfort, Machine Learning, Computational Optimization, Artificial Neural Networks (ANN), Uniform Latin Hypercube sampling (ULH)

## Acknowledgments

I would like to express my greatest gratitude to my mentor Berk Ekici who was constantly available for both technical support and personal encouragement. His recommendations served as a clear path to learn about surrogate models and performance-based design. I would also like to show deep appreciation to Dr. Michela Turin who shared her in-depth knowledge of the computational realm and gave me key recommendations on the methodology and framework for the entire research. Likewise, my mentor, Dr. Regina Bokel provided great contributions to the building physics aspect of this thesis through precise feedback on Energy modeling and the overall structure. Also, Dr. Harry Boumeester feedback as a delegate board of examiners gave great insight from a more overarching approach.

I would like to equally express my deep gratitude for the Delft University of Technology represented by the Education and Student affairs that granted me the possibility of coming to this institution through the scholarships received at the beginning of this academic endeavor. I would not be here if it were not for this wonderful opportunity.

The great friendships that I made with my colleagues from the Building Technology Track not only provided motivation but also an inspiration to work hard through difficult times.

Finally, I would like to express my deepest love to my parents Martha and Remberto, who supported me unconditionally during this whole journey. Without their words and care, this would not have been possible.

I will always be grateful and indebted to the people that made this possible.

## 1.0 Introduction

Our future will be marked by scarcity of resources and fossil fuels (IPTOES, 2010) therefore, now more than ever, our buildings should aim to become more efficient in their use of energy and resources. Andy van den Dobbelsteen makes a great point by highlighting that the price of oil is embedded in everything that we consume. (Dobbelsteen, 2012) He stresses that today, it is critical to breaking off from our fossil fuel energy dependency. Our new buildings need to be energy conscious by allowing for the lowest possible environmental footprint but equally maintaining the essential reasons for architecture: shelter and comfort.

Most current high-rise buildings reflect the architectural language of the High-tech movement. The High-tech movement, the predecessor of today's most modern high-rise buildings, was born in the ecological techno-optimism of the 1960s by Frei Otto and Buckminster Fuller. They encouraged the conscious use of materials and methods due to our limited resources on earth. Fuller's "Spaceship Earth" concept stated that we live on a finite planet; therefore, we should use all tools and resources as efficiently as possible. However, current high-rise buildings rarely follow their original idealistic principles: sustainability, efficiency, ecological and humanistic responsibilities. (Harper, 2019)

The goal of this research is to provide architects and engineers with an efficient tool of design for quick insight into their proposal. Ideally, High-rises should be designed contemplating **locality in design** during the **early concept stage** of the design process. **Computational** simulations and optimization can help the designer to identify promising design solutions, however, existing simulations are computationally expensive and thus cumbersome during this initial stage, therefore designers rarely design with enough contextual information. Surrogate models have the potential to provide performance assessments without the need to run all related performance simulations, offering a faster alternative to the use of simulations only. This leads to the following research question:

*Research Question:*

*"How can a computational method using surrogate modeling be used to quickly identify, and optimize the most influential factors and their combinations for context-based passive solutions of sustainable High-Rise office buildings during their initial design phase?"*

with the following sub-questions:

- How do different locations/climates dictate the building's ideal shape and orientation?
- How does the volumetric context of surrounding buildings affect its shape and façade parameters?
- How does the position of its Core affect a building's energy performance concerning its location/climate?
- What is the Pareto-optimality of these parameters when simultaneously seeking energy consumption, thermal comfort, and natural daylight optimums?
- What are the key validation metrics and features necessary for obtaining suitable surrogate models?

The goal of this research is to ultimately provide architects and engineers with an alternative efficient tool of design for quick insight into their proposal. This, in turn, would aid the designer to rapidly optimize their project through parametric optimization tools as well as allow them to explore their design choices effectively.

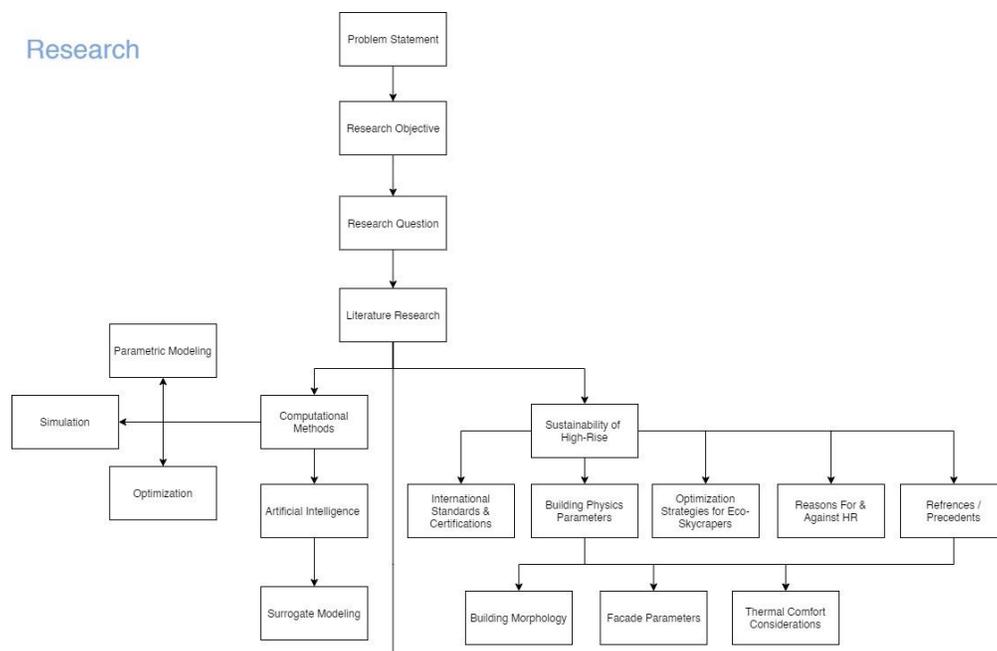
## 1.1 Scope

The research aims at understanding how a computational method can using surrogate modeling be used to look at how shape and façade variables affect building energy, daylight, and comfort performance on high rise building typology. It also provides results regarding how shape and façade variables affect building energy, daylight, and comfort performance on high rise building typology. Due to the implicit cost of building high-rises, the usable floor area is also a key metric to consider, therefore it is similarly monitored. Likewise, a good high-rise typology can vary drastically depending on the location it is implanted; hence, location-specific design and differences will be addressed.

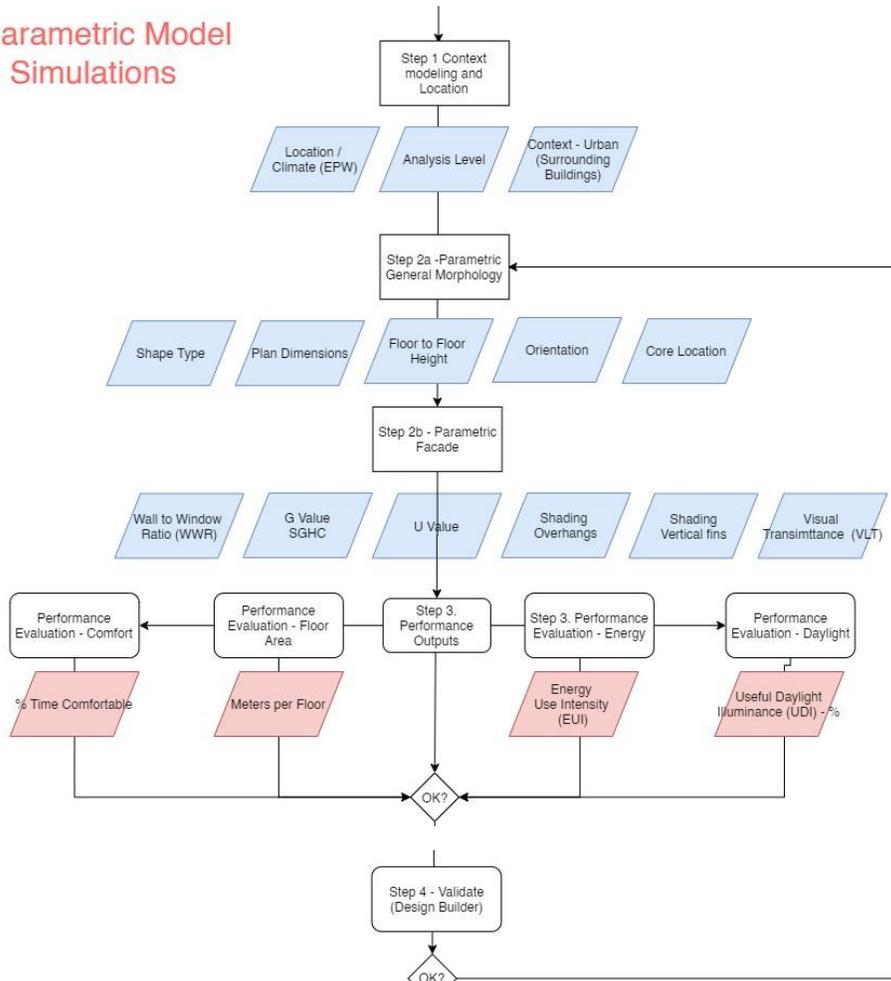
Although HVAC settings are important on the performance outputs, these parameters will be defined but kept constant, thus falling out of the scope of this research. Production of energy within the site was another possible research theme yet was not included due to time constraints. Similarly, the life cycle costs of a building are another theme that is excluded within the scope.

## 1.2 Methodology

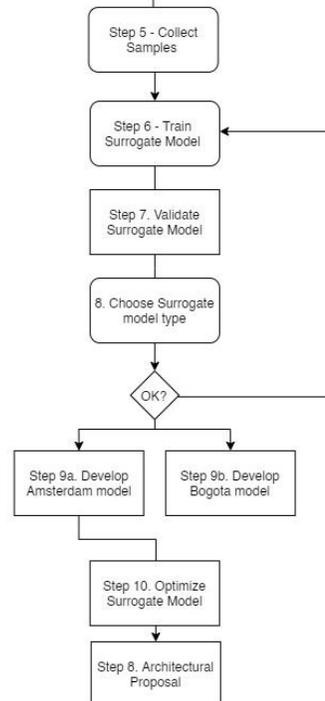
In terms of the computational realm, this research aims to shed light on parametric modeling, simulation and optimization applied to the architectural realm. It will look at data collection (design of experiments) and parametric optimization using Galapagos (Rutten, 2019) as the main Graphic User Interface and Octopus (Vierlinger, 2018) and Optimus (Cubukcuoglu, et al., 2019) as the optimization software plugins. Also, the application of various Machine learning methods for creating surrogate models in Grasshopper such as Ant (Rahman, 2017) Dodo (Lorenzo, 2019) and Lunchbox (Miller, 2018) is used. The surrogate models and their “power to predict” will be addressed, thus looking at their performance and precision metrics. For the energy review, the Grasshopper simulations Ladybug and Honeybee (Sadeghipour & Mackey, 2017-2020) are compared with DesignBuilder software (DesignBuilder, 2020) to validate the Grasshopper model. Finally, a redesign is made of one of the generic proposals into a basic scheme architectural proposal in the form of 3D Diagrams, Plan, and Views.



## Parametric Model & Simulations



## Surrogate Model & Optimization



## 2.0 Literature Review

### 2.1 High-Rise Typology

Firstly, it is important to define what constitutes a High-Rise building. The Council on Tall Buildings and Urban Habitat (CTBUH) explains that the definition of “tall” can vary depending on three factors: Its context, its proportions, and its technologies.

- Context: If the building is taller compared to its neighbors, it could be considered a high-rise.
- Proportion: If the proportions make it a slender construction, in contrast with a building that has the same height but its large footprint unqualifies it from being considered tall.
- Technology: If the building uses structural or vertical transportation technologies on tall buildings

On the other hand, according to Emporis a high-rise building is a structure between 35-100 meters or has between 12-39 floors. Above 100 meters it is considered as a skyscraper. (Emporis, 2020)

Due to the imprecise definition of a high-rise, for this study, a High-Rise building will be defined as one with at least 20 floors ( $\approx 85$  m at FFH 4.2). The parametric model contemplates a height of up to 30 floors ( $\approx 300$  m at FFH of 4.2 m) as CTBUH identifies most high rises built in the last century falling within this category (see blue dots).

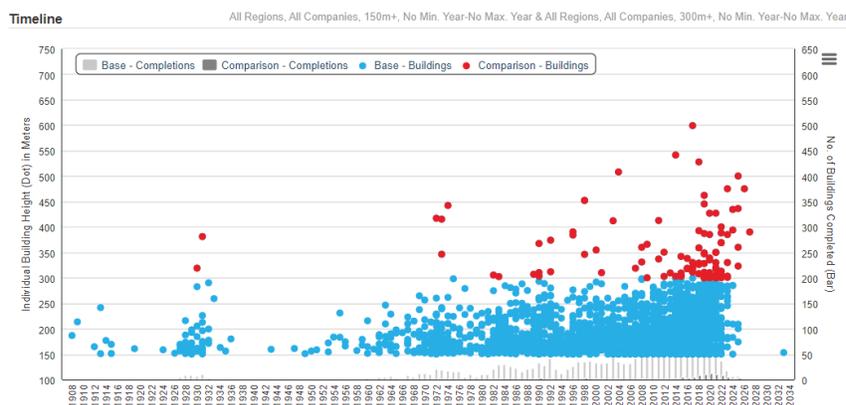


Table 1 High-Rises between 150 and 300 m (CTBUH, 2019)

With this definition in mind, a quick exercise was done to classify the three main building typologies and some significant examples of each:

Type	Subtype	Example	Author	Date	Page
Office	Modernist	Seagram Building	Mies van der Rohe	1958	-
	High-Tech	HSBC	Norman Foster	1985	-
	Landmark (proposal)	Doric Column Chicago Tribune	Adolf Loos	1922	33
Residential	Modernist	Unité d'Habitation	LeCorbusier	1952	45
	Landmark - Tallest	432 Park Avenue	Rafael Viñoly	2015	-
Mixed-use	Landmark - Tallest	Burj Khalifa	SOM	2009	38
	The Hyper Building	CCTV	OMA	2008	39
	Quasi-Public spaces	Tour Signal	Jean Nouvel	2008	39

Table 2 Tall building Typologies - Own production. Info from: (Zandbelt, 2012)

These different typologies have a significant impact on energy consumption, ideal daylighting and thermal comfort optimums, thus only one typology was chosen: the office building. Office buildings are widely related to the high-rise typology due to many of the reasons in the following segment.

High-rise constructions are an essentially unique building type in modern architecture, yet when it comes to categorizing differences, they are significantly different from other building typologies mainly in the following two categories:

### 2.1.1 Façade technologies

High rise envelopes need to account for more extreme weather conditions than traditional windows (such as higher wind loads and UV radiation). This, in turn, leads to more specialized façade technologies and even new ways of constructing. The curtain wall system, for instance, was born from the high-rise typology. This high level of specialization of the facades also means that applying dedicated glazing and insulation parameters are much more appealing.

### 2.1.2 Vertical Core

The core is the means of transportation of the building. It is employed not only to transport people but also the Mechanical Electrical and Plumbing (MEP) fixtures. As the building grows, the cores need to grow as well to provide for vertical lift shafts, emergency stairs, and all the MEP ducts that allow for the building to function properly. This means that the core size depends largely on architectural design, vertical transportation methods, and the ducts & equipment. Generally, the area dedicated to the core is from 15-25% of the Net Floor Area (NFA)

Also, the space efficiency of a building is directly related to the core. This is because, the higher the building, the larger the structural elements required to support it. Consequently, the larger the core, the lower the efficiency. This trend can be seen in the following table.

Number of Stories	Efficiency (%)
Two to four	83-86
Five to nine	79-83
10 to 19	72-80
20 to 29	70-78
30 to 39	69-75
40 +	68-73

Table 3 Efficiency of High-rises related to their height (ÖZGEN, 2009)

For this reason, the model that will be developed should consider this loss of efficiency as the tower height increases. See section 3.2.1 Core Type.

Efficient High-rise buildings have employed many different core strategies, but they can be mainly categorized into two types: Central core and Perimetral core.

### Central Core

In 2009, 9 of the tallest buildings in the world all employed a variation of the central core typology.

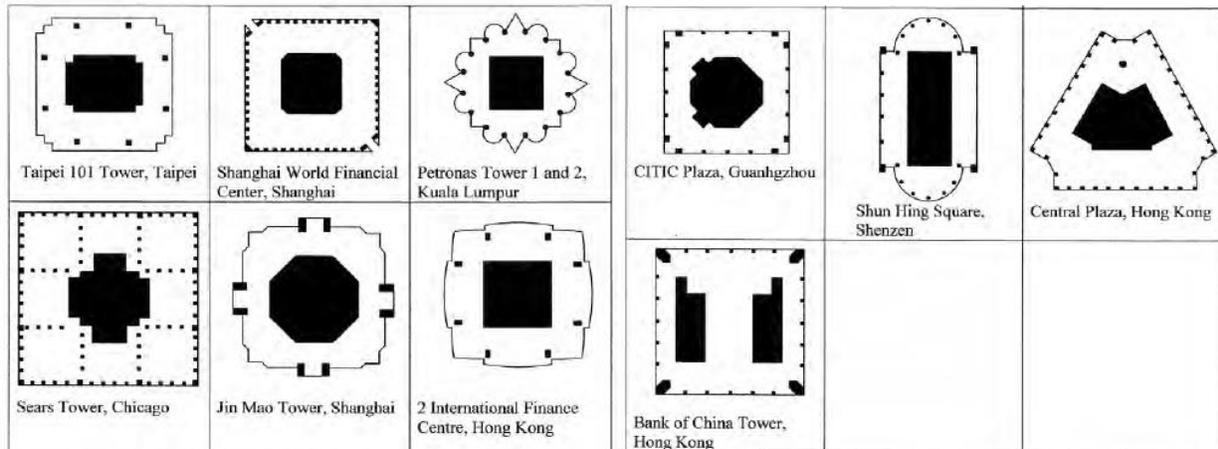


Figure 1 Central Core Typology (ÖZGEN, 2009)

This typology is very efficient in terms of lease span, which is a definition that expresses the distances between the façade of the building and the fixed interior divisions, usually the core walls. When lease spans are kept between the ideal 6-8 m, good daylight illumination levels are achieved. Also, this configuration permits buildings to have high relative compactness, thus be good at saving energy consumption.

### Perimetral/Lateral cores

This does not mean that lateral cores are not good design. There are many examples of well-designed skyscrapers with lateral cores and irregular shapes:

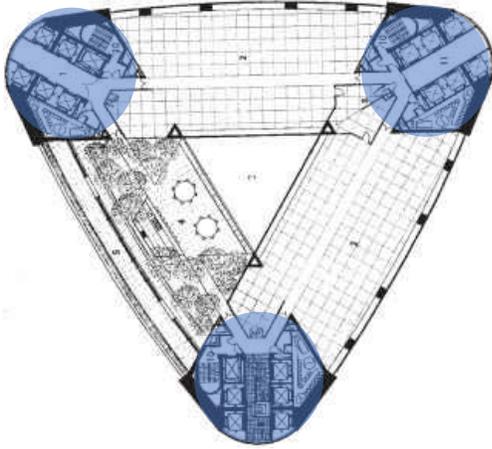


Figure 2 Fosters & Partners Commerzbank – Frankfurt (Buchanan, 1998)

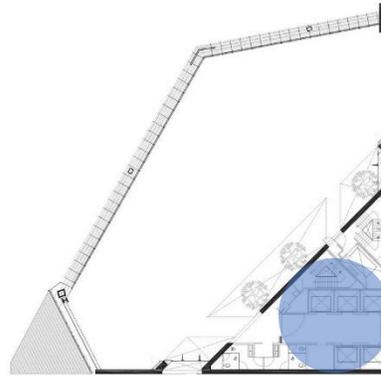


Figure 3 LBR - Torre Reforma – Mexico (Romano, 2016)

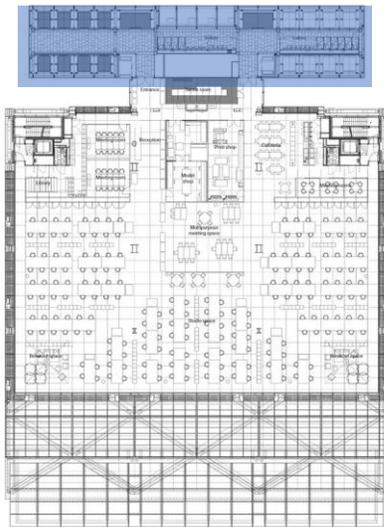


Figure 4 Rogers & Partners – Leadenhall – London (Rogers, 2014)

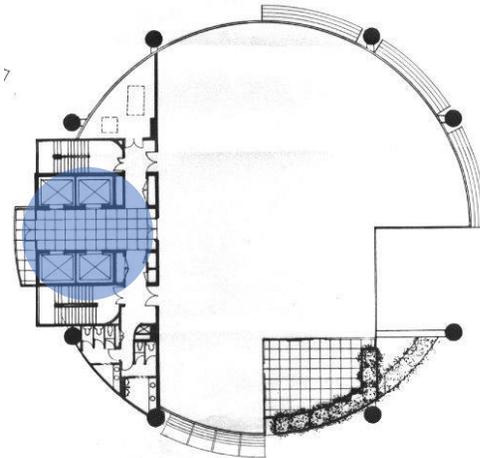


Figure 5 Ken Yeang - Menara Mesiniaga - Petaling Jaya (Yeang Sdn. & Bhd, 1992)

In the last example, Ken Yeang used the atrium as well as the core location to improve the passive sustainable strategies in his buildings. The location of the core depended on the climate type. He placed the cores on the hot sides of the buildings to act as thermal buffers, thus reducing the insolation. Exploring the core location parameter depending on the climate type could lead to even smarter eco-skyscraper design.

The core of a high-rise is a versatile and vital element to this typology yet there is little research on its effects on a building's energy, daylight, or thermal performance. The vertical cores show great potential to improve this by serving as a thermal shield, thermal mass, or daylight reflection mechanism, therefore, it was considered as an important variable for this research.

### 2.1.3 Advantages and Disadvantages of High-Rises

A general literature review on the reasons For and Against the construction of high-rises was the starting point to understand what a “sustainable high-rise” meant. The main identified topics regarding High-Rise sustainability were: Power & Value, Urban connections, Density, Vitality/Urbanity, Economy, Context, and Energy/Resources. (See Summary Table 3) All these topics fall into two main categories: the building shape or the building context. A frequently recurring concern is that high-rises usually suffer from lack of context specificity: most high-rise buildings designed today are rarely site or location-specific. Metaphorically, they are imposing, self-centered giants. And although this site-specificity can be viewed from the different views of the topics previously identified; such as the vitality, density, or social aspects, this thesis chose only to focus on energy(resource) performance based on its urban context and climate. If any of these topics can be improved, then the High-Rise typology will be much more defensible within the future of our cities.

<b>Economy</b>			
To create housing in areas with lower real estate value	18	Pressure for higher densities and higher profit	45
Can create a unique selling point in a competitive real estate market	39	Little "post occupancy evaluation to asses long term value	45
Investment advantages at locations that would otherwise not attract investment	50	Narrow sector demand - low suitability of plan and affordability	
Desire - pressure by "stackholder" groups	50	High-Rise does not allow phased construction, thus large investment thus high financial risk / burder	72
Global / Economic positioning	50		
Corporate culture - Tall office buildings refere more and more to financial and service industry			56

<b>Context</b>			
Can symbolize a new stage in urban development and enhance legibility or cohesion of a city as a whole.	75	Unplanned growth - "Cities like London continue to "judge" each planning application on the basis of its merit and impact on its surroundings with no assumptions of adjacent similar projects"	47
Can provide fantastic views	75	Cities can loose UNESCO status	48
GAP- Green Area Preserved – open space beyond the city preserved through building more intensively than on average within the city	128	Building aesthetics seems to be more and more generic and non-place specific, "with a total lack of notional references to use, locations, ownership or cultural context."	64
		European cities struggle with relationship with high-rise and built heritage	105

<b>Energy &amp; Resources</b>			
Bigger opportunities to apply a large scale solution when efficient proposals are considered.		Skyscrapers generally perform worst than conventional buildings - high construction costs to improve it	50
Energy Efficient mobility - More density = more efficiency in energy consumption for travel	79 / 129	Buildings need to be spaced out to allow for good energy performance	53
Space - Less land use	126	Transport and buildings are responsible for some two thirds of CO2 emmissions in rich countries (Neale, 2008)	81
Time - Longer life of service	126	Foundations and superstructure consume more resources for construction compared to other typologies	126
Potential for "synergetic buildings" that uses the waste of one function as the resource of another / vice versa	135	Requires greater material for structural system, greater energy demands to transport and pump materials, services and people against gravity	167

For	Pg	Against	Pg
<b>Power &amp; Value</b>			
To express power and strength - San Gimignano and Bologna	9	Chosen as a Symbols, yet "the connection between signifier and signified is arbitrary" - Ferdinand de Saussure	26
Military height advantage	9		
"The Pig and the skyscraper - A history of our Future" - Marco d'Eramo – Emergence of Capitalism - Due to a mix of scarcity of land, Technology and Ego			32
In European cities - As an axis / icon	33	In American cities - Hierarchy achieved only by huge scale	33
Project an "International Image"	44	Real value to a city usually comes from infrastructure and policies rather than tall buildings	55
Gave Rotterdam a proud image of a young high-rise city	157	Cities start becoming "like everywhere else"	55
		Clinging to old idealism of the future metropolis	150
<b>Urban connections</b>			
To reduce excessive travel to alternative suburban areas	18	Negative impact on urban character - extensive black walls, service entrances and single front entrance	50
To save on urban infrastructure - public transportation	21	Problems at ground floor - Blind street level facades, gaping holes or gates for parkin entrances, messy facilities and container areas.	95
High-rises allow us to avoid excessive land claims outside the city and produce a more sustainable use of transport	144	In purely high rise (no plinth) Commercial and public uses have diminished potential	90
<b>Density</b>			
Intensification of space and use where there is/will be existing infrastructure	50	With tower separation rules for residential towers, high-rise structures do NOT accomodate higher density	25
In large cities their ideal locations mean that they offer urban setting and lifestyle, services and facilities	74	Medium cities dont offer enough public services to advocate for High-Rises	74
With similar density, people have access to daylight and to open space (OSR) for playing and growing food	89	Inefficiency of building type (bad ratio of core vs building footprint)	50
Intensive use of surface area, intensive use of land (Stacking), multiple or mixed use of surface (mixing)	128	Does not have a significant impact on capacity of city block	53
Reduction consumption of fuel for transport (Graph)	129	Height does not necessarily translate into density	70
"...Comprehensive management of height and choice of morphological models are decisive in achieveing higher capacities"			54
<b>Vitality / "Urbanity"</b>			
When carefully considered, it can activate public life in cities - Example: Rockefeller Center NYC	25	Many projects built on LeCorbusiers principles ended up as territories of social despair and decay & demolished a few decades later	10
Significant corporate presence in highly accessible locations in cowntown of cities	18	Limit to impact of mixed use and tenure- Limiting characteristics of space and geometry - high rental and maintenance costs	50
On the principles of modernist "social reconstruction"	45		
Regeneration of places - Signposting	50		
Diversification of tenancies & encouragement of sectors outside existing demand	50	Comprmise of vitality and activity - Buildings "suck" activity inside rather than inject vitality into public space	50
Larger open floor plates and shapes can allow for seamless communication, mix of functions and flexiblity of operations	56	Lack of development flexibility - phasing, future remodeling or redevelopment	50

Table 2: Reasons for High-Rises. Own production. Info from: (Zandbelt & Mayer, 2012)

#### 2.1.4 Sustainable High-Rises

As previously stated, one important way to achieve sustainability is by designing locally. Possible local design factors include:

- Urban Context
- Social Context
- Historical Context
- Local Culture
- Climate
- Economy
- Availability of Resources
- Life Cycle analysis

Due to the broad subject and limited time, only Urban context and Climate were chosen as the main topics due to their possible applicability to the computational realm. An important point to consider with the urban context is that tall buildings receive much more direct solar radiation, yet this can vary in areas with other tall buildings. (Elotefya, et al., 2015) For this reason, the shadows created by the surrounding buildings were an important element to be considered.

Another way to improve a building's sustainability is by applying optimization strategies on buildings. Ken Yeang, creator of the self-proclaimed "eco-skyscraper" shows many great strategies to achieve this through many of his designs. (See table 4). The reason for this table was to identify all the possible optimization strategies applied to high-rise buildings. The recurring theme of this table also highlights the importance of designing locally.

Again, aspects like social vitality or wind behavior on buildings are important elements to consider when designing tall, yet a conscious decision was made to not include them in the scope of this thesis. Social vitality is something that relates to human relationships and community, therefore computational tools seem unfitting for its study.

Wind behavior, on the other hand, is a subject that has been studied extensively in other fields such as Aerospace Engineering that study aerodynamics with the use of Computational Fluid Dynamic (CFD) software. This part of this subject was consuls left out to focus on lighting and energy factors that play an equally important role in high-rise design. Likewise, the subject on CFD is a very specialized one, requiring much more research to master and use properly.

Finally, another theme mentioned in the table is green technologies such as vertical landscapes or agricultural production along the building's facade. Although there are many examples of green filled proposals in renders and competition entries, actual successful real-life projects currently applying this technology are few and far between. Successful examples of this are Bosco Vertical in Milan by Stephano Boeri and Oasia Hotel Downtown Singapore by WoHa. These projects have both overcome many challenges like water supply and drainage, additional weight from planters, redundant safety systems to keep the plants in place, costly maintenance, and plant replacement, yet green facades offer little improvement compared to traditional shading devices for internal temperature. (Phan Anh Nguyen, 2019)

That is not to say that planting technology is a negative solution; it offers many other advantages such as reducing heat island effect, purifying air, noise cancelation, and creating relaxing atmospheres for the user. This technology seems particularly desirable in low/mid-rise building typologies that have lower wind velocities hitting the plants and potentially damaging them. Hence, green technology is a potentially important subject for a designer to consider but for now, they seem like few niche solutions on high-end skyscrapers, therefore they were not contemplated for this thesis.

Type	1	Daylight	Description	Limitations / Comments	Climate Type	Example	Pg
Use	1.1	Solar Access	Public places designated to have full sunshine "sun Spots" during their most intensive use or high frequency periods	Requires proactive rather than reactive city planning.	Temperate	Rotterdam City Planning	99
Use	1.2a	Solar Orientation	Orientation of the tower specific to the sun path of tropical climate	building cores are placed on the hot east and west facades	Tropical	IBM Plaza - Kuala Lumpur	170
Use	1.2b	Solar Orientation	Same principles, temperate climate - Maximize solar gain into interior spaces in winter and maximize solar shading in summer	Peripheral apartments enclose an internal atrium	Temperate	Bishopgate Tower - London	174
Use	1.3	Diagonal Lighthshaft	Atrium and Diagonal lightshaft - diagonal opening through the section of the building	Allows for more natural sunlight to enter deepest sections of the building	Tropical	Solaris Fusionopolis - Singapur	182
Use	1.4	Ecocells	Sun wells that allow daylight to penetrate into the basement levels	Allows for natural sunlight and ventilation	Tropical	Spire Edge - India Manesar	180
Use	1.5	Horizontal Louvers	The size and depth of the louvers was determined by the sunpath	Louvers also serve as light shelves	Tropical	Spire Edge - India Manesar	180
Type	2	Facade Tech	Description	Limitations / Comments	Climate Type	Example	Pg
Prod	2.1	PV Panels on facades	Vertical surface solar energy generation - The farther north / south hemispheres, the more productive the facade can be	Amsterdam: Average maximal Solar angle of 37 (more vertical than horizontal)	Temperate	Amsterdam	
Prod	2.2	Solar Collectors	Use of limited roof space for parabolic solar collectors	Although High-rises do not have extensive roof space	Temperate	Villa Flora	139
Use	2.3	Louvered Facades	Louvered panels protect the facade from solar gain	Design determined by orientation	Tropical	Menara Mesiniaga	171
Use	2.4	Balconies	Deeply recessed balconies and planter boxes	Provide sunshading	Tropical	Central Plaza Tower	172
Use	2.5	Recessed Windows	Deeply recessed windows	Solar heat gain protection	Tropical	Menara Umno Penang	172
Use	2.6	Buffer Cores	Cores serve as thermal buffers	Located in zones with highest solar heat gain	Tropical	Singapore National Library	176
Use	2.7	Glazing	Low-E	low external thermal transfer value (ETTV) - 39 watts /m2	Tropical	Solaris Fusionopolis - Singapur	184
			Double glazing		Gen	General	
			Double glazing Argon-filled cavities		Gen	General	
			Triple Glazing		Gen	General	
Type	3	Vitality / Social	Description	Limitations / Comments	Climate Type	Example	Pg
Use	3.1	Ecomimesis	State of stasis, the built environment that imitates an ecosystem that recycles and reuses byproducts to produce zero waste	Is not the result of adding more and more technological gadgetry	Gen	General	168
Use	3.2	Sky Courts / Gardens	Spiraling green garden connected to the surrounding landscape and sky courts and sky gardens	Culminate in rooftop swimming pool	Tropical	Menara Mesiniaga	171
Use	3.3	Vertical spatial continuity	Continuous landscaped ramp through the tower	Spacial continuity between street and the tower	Tropical	EDITT Tower - Singapore	173
Use	3.4	Comunal Skycourt	Sky gardens every 5th level to provide landscaped communal sky courts		Tropical	Bombay Glassworks tower - Mumbai	185
Type	4	Green / Water	Description	Limitations / Comments	Climate Type	Example	Pg
Use	4.1	Vertical Green landscapes	Provides shading, evaporative cooling and micro-climate improvements	High maintenance of Vegetation	Tropical	EDITT Tower - Singapore	173
Prod	4.2	Greenhouse	Vertical glazed greenhouse could be used for agriculture	Cannot yield as much as commercial farming, but provide locally grown fruit, vegetable, herbs and spice	Temperate	Villa Flora / Harvest Tower	140
Prod	4.3	Rainwater reuse	Rain water catchment, retention, storage and recycling		Tropical	Spire Edge - India Manesar	180
Use	4.4	Efficient Water Fixtures	low-flow rate fixtures and gray water flushing		Temperate	Leza Soho - Hangzhou, China	
Type	5	Wind	Description	Limitations / Comments	Climate Type	Example	Pg
Use	5.1	General Shape & Texture	Iso Standard - wind study to measure effects of building shape, texture and facade during preliminary design stage	Requires existing context, can change with future proposals		Generic	
Use	5.2	Aerodynamic Shape	Structure that promotes air acceleration through building, thus cooling interior by few degrees	Requires computational fluid dynamics (CFD software) or wind tunnel tests	Tropical	Singapore National Library	
Use	5.3	Aerodynamic Shape	Oval Shape	Reduces stress on superstructure generated by wind loads	Tropical	EDITT Tower - Singapore	174
Prod	5.4	Wind Turbines	Turbines in facade on rooftop level		Tropical	Miami COR building	
Use	5.5	Wind wing Wall	A 21 story high vertical project wall directed towards the prevailing wind	Creates positive and negative pressure zones that are very effective in providing natural ventilation through the common areas	Tropical	Menara Umno Penang	172
Type	6	Natural Ventilation	Description	Limitations / Comments	Climate Type	Example	Pg
Use	6.1	External Atrium	External atrium	Creates a wind shaft to cool the offices facing the atrium	Tropical	Kuala Lumpur Tower - Plaza Atrium	170
Use	6.2	Central Atrium	Central atrium open at the base	Creates a cool microclimate while allowing natural daylight to reach the circulation areas within the building	Tropical	Library of Singapore	176
Use	6.3	Vertical Atria Gardens	Vertical atria gardens / airwells and public spaces	Provide passive cooling	Dry	Ecobay Complex - Abu Dhabi	178
Use	6.4	Service Core Placement	Service core located on the hot side of the building	Naturally ventilated to reduce the air conditioning load on the offices	Tropical	Menara Boustead - Kuala Lumpur	171
Use	6.6	Cooridors	Naturally ventilated single-loaded corridors face the hot wester side of the site	Act as buffer, reducing insolation	Tropical	Casa del Sol - Kuala Lumpur	172

Use Effective Usage  
Prod Production

Table 3 Optimization Strategies for High-Rises- Own Production, info from (Zandbelt & Mayer, 2012)

## 2.2 Building Physics

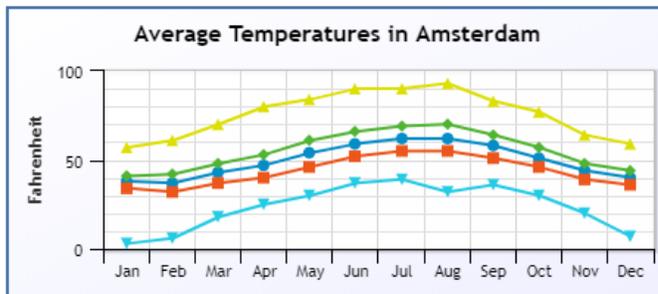
### 2.2.1 Climate Zones

The Köppen climate classification divides climates into five main climate groups:

A (tropical), B (dry), C (temperate), D (continental), and E (polar), these are subdivided with a second later that indicates seasonal precipitation and temperature patterns, while the third letter indicates the level of heat. These climatic zones affect the outdoor temperature, relative humidity, solar radiation, and wind speed and thus have a great impact on the configuration of a building's passive design and control strategies. It is crucial to design according to its specific climatic context.

As defined by the research question, the idea is to generate a tool that contemplates different climate factors. Preferably, various climate types and locations would be reviewed, but at least 2 locations should prove the methodology. Weather zones with larger seasonal variations offer more complexity when designing because while some solutions might work well in the warm summers, they might not be as effective in cold winters.

#### Amsterdam, Netherlands



Elevation: 2 meters

Latitude: 52 18N Longitude: 004 46E

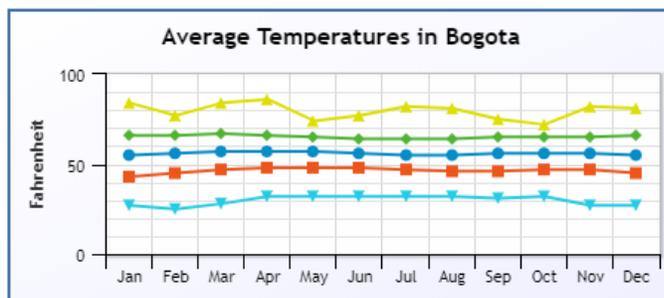
Köppen Classification:

(Cfb) Marine West Coast Climate

*Figure 6 Climate Annual Summary (Weather Base, 2020)*

The reason for choosing Amsterdam is to start with a control “local” location that will be able to be compared to other locations worldwide. Also, Amsterdam has some seasonal variation that will show the tradeoffs when optimizing the three intended outputs.

#### Bogota, Colombia



Elevation: 2547 meters

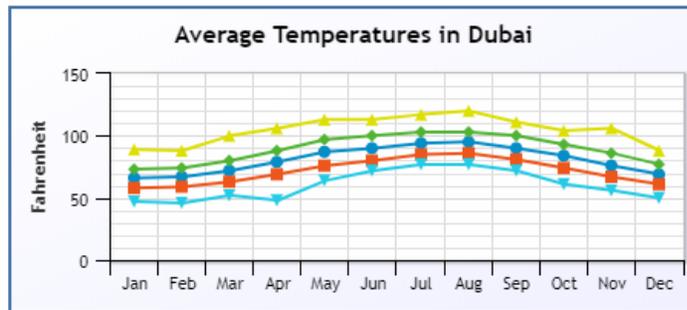
Latitude: 04 42N Longitude: 074 08W

Köppen Classification: (Cfb) Marine West Coast Climate

*Figure 7 Climate Annual Summary (Weather Base,*

Bogotá has the same climate classification of Amsterdam yet because of its location near the equator, the climate is much more constant throughout the year, showing barely any seasonal variation. There is a marked difference between summer or winter season. Most days are cloudy, but when the sun shines there is a high radiant load that should be protected.

## Dubai, United Arab Emirates



Elevation: 4 meters    Latitude: 25 15N  
Longitude: 055 20E

Köppen Classification: (BwH) Tropical and Subtropical Desert Climate

*Figure 8 Climate Annual Summary (Weather Base, 2020)*

The reason for choosing Dubai is because today it is the city with the highest skyscraper, the Burj Khalifa in a harsh desert environment. Although these buildings undoubtedly showcase a feat of human engineering, research on the basics of building physics makes it seem completely out of place with large glazing areas and non-existent shading strategies. This location would thus be an ideal extreme desert climate study.

### 2.2.2 Building Energy Performance

Many regulations, codes, and standards around the world are in place for classifying energy-efficient buildings. Next is a summary of the most recognized international standards reviewed for this report.

#### **International Green Construction Code and ASHRAE Standard 189**

This is a general code that covers land use, water conservation energy conservation, and other criteria. It was established by ICC, AIA, ASTM International, ASHRAE, US Green Building Council, and IES with the idea to provide baselines for green construction. (Efficient Windows Collaborative, 2000-2018)

#### **ENERGY STAR certification**

In the United States, the Environmental Protection Agency (EPA) and the US Department of Energy (DOE) developed the ENERGY STAR certification for products that meet certain energy performance criteria.

Windows have labels that show the zones in which they are qualified. The building envelope performance should vary by climate therefore the recommendations are given for four distinct U.S. climate zones:

- Northern Zone Required Properties (mostly heating)
- North/Central Zone Required Properties (heating & cooling)
- South/Central Zone Required Properties (heating & cooling)
- Southern Zone Required Properties (mostly cooling)

This regional climatic specification is specific to the US but the principle of differentiation of building components can be broadened for the whole world. (Efficient Windows Collaborative, 2000-2018) The two most common international certifications on the sustainability of constructions are LEED (US) and BREEM (EU), each has its chapters on energy efficiency and reduction.

## Dutch building degree 2012 (Bouwbesluit)

This is the Dutch building degree that determines a series of parameters that must be met depending on building type and stage (new or existing). This document has all the legislations that a building in the Netherlands must comply with. Section 3.11 specifies the daylight and Article 3.75 about the daylight surface. Chapter 5 specifies the technical building regulations from the energy efficiency and environmental performance. (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 2012)

## Heat Balance and Energy demands

The fundamental principle of energy modeling is based on the first law of thermodynamics: All energy is transformed, never destroyed. Energy In = Energy Out

Thus, any space has thermal balance when, under stationary conditions, the internal and external energy achieve balance, thus  $Q_{in} = Q_{out}$ .

The energy balance formula is:

$$\pm Q_{transmission} \pm Q_{ventilation} - Q_{infiltration} + Q_{sun} + Q_{internal} + Q_{energyuse} = 0$$

Modifying any of these components will affect in turn the final energy use.

$$\pm Q_{transmission} \pm Q_{ventilation} - Q_{infiltration} + Q_{sun} + Q_{internal} = -Q_{energyuse}$$

Therefore, by varying the variables of the façade, like the transmission or  $Q_{sun}$ , the final energy consumption can be lowered.

Looking at each independently:

**Q<sub>transmission</sub>** is the energy that is transmitted through the façade, thus dependent on envelope parameters and the climate conditions.

$$Q_{transmission} = \sum U * A * (T_e - T_i)$$

*U = the U value*

*A = the area of the façade*

*T<sub>e</sub> = the temperature outside (climate dependent)*

*T<sub>i</sub> = the temperature inside (comfort dependent)*

**Q<sub>ventilation</sub>** is the energy from the Natural Ventilation or HVAC system

$$Q_{ventilation} = V_{vent} * \rho * C_p * (T_e - T_i)$$

*V<sub>vent</sub> = Ventilation flow*

*ρ = air density = 1.2*

*C<sub>p</sub> = air heat capacity*

*T<sub>e</sub> = temperature outside (climate dependent)*

*T<sub>i</sub> = temperature inside (comfort dependent)*

**Qsun** is the energy input from the sun.

$$Q_{\text{sun}} = A_{\text{glass}} * q_{\text{sun}} * g$$

*A<sub>glass</sub>* = area of the glazing

*q<sub>sun</sub>* = Intensity of the solar radiation on the glass (climate dependant)

*g* = g-value (solar transmittance factor of the glass) (see SHGC)

**Qinfiltration** is the outdoor air filtered through cracks and that was not intended to enter the space

$$Q_{\text{infiltration}} = V_i * \rho * C_p * (T_e - T_i)$$

*V<sub>i</sub>* = infiltration flow (0.2)

*ρ* = air density = 1.2

*C<sub>p</sub>* = air heat capacity

*T<sub>e</sub>* = temperature outside (climate dependent)

*T<sub>i</sub>* = temperature inside (comfort dependent)

**Qinternal** is the energy that is added to the building through lighting, people, and equipment.

$$Q_{\text{int}} = Q_{\text{people}} + Q_{\text{light}} + Q_{\text{equipment}}$$

$$Q_t + Q_v + Q_i + Q_{\text{sun}} + Q_{\text{int}} = -Q_{\text{energyuse}}$$

$Q_t = \sum U * A * (T_e - T_i)$	$Q_v = V_{\text{vent}} * \rho * C_p * (T_e - T_i)$	$Q_i = V_i * \rho * C_p * (T_e - T_i)$	$Q_{\text{sun}} = A_{\text{glass}} * q_{\text{sun}} * g$	$Q_{\text{int}} =$
<i>U</i> = U-value <i>A</i> = area of the façade	<i>V<sub>vent</sub></i> = Ventilation flow <i>ρ</i> = air density = 1.2 <i>C<sub>p</sub></i> = air heat capacity	<i>V<sub>i</sub></i> = infiltration flow (0.2) <i>ρ</i> = air density = 1.2 <i>C<sub>p</sub></i> = air heat capacity	<i>A<sub>glass</sub></i> = glazing area <i>g</i> = g-value	<i>Q<sub>people</sub></i> + <i>Q<sub>light</sub></i> + <i>Q<sub>equipment</sub></i>
<i>T<sub>e</sub></i> = temperature outside (climate dependent)	<i>T<sub>e</sub></i> = temperature outside (climate dependent)	<i>T<sub>e</sub></i> = temperature outside (climate dependent)	<i>q<sub>sun</sub></i> = Radiation on glass (climate dependent)	
<i>T<sub>i</sub></i> = temperature inside (comfort dependent)	<i>T<sub>i</sub></i> = temperature inside (comfort dependent)	<i>T<sub>i</sub></i> = temperature inside (comfort dependent)		

Figure 9 Summary Heat & Energy Balance (source: own)

This table shows a summary of all the elements that come into play. In **orange** the envelope dependent parameters, in **blue** the climate/location-dependent parameters and **green** the comfort dependent parameters. Consequently, if these parameters are carefully modified and enhanced, the final energy use of a building will be improved. The final chosen parameters are described in chapter 3.2

## Energy Modeling

The main steps to generate an energy model are to define the following variables:

1. Envelope Geometry + Materials (Insulation)
2. Program / Uses
3. Internal Mass
4. Internal Loads
5. Schedules
6. Natural Ventilation
7. HVAC System

(Levitt, 2015)

Ultimately, energy consumption is a metric that depends on various factors such as window properties, people density, typology or use, HVAC system type, ideal user comfort, occupancy schedules, peak electricity demand. The heat lost or gained through its envelope represents a significant portion of the building's heating and cooling loads. (Efficient Windows, 2011-2020)

### Energy Modeling Performance Metrics

The final step in achieving a trustworthy energy model is choosing the right output to review and optimize. Among the possible outputs are Cooling, Heating, Equipment, and Lighting loads, but these are 4 separate output variables measured in KW h / m<sup>2</sup>. A way to unify them into a single output is, to sum up, the Total Annual energy demand measured in Kw H /m<sup>2</sup> / yr. Even further, by looking at this yearly output and then normalizing the results by floor area the site Energy Use Intensity (EUI) is defined.

### EUI

Site energy EUI is used to understand the energy use of that individual building. Source energy, on the other hand, incorporates all the energy required for the building including transmission, delivery, and production. This measurement provides a complete assessment of the energy efficiency of the building by contemplating all possible losses and forms of energy. (EnergyStar, 2020)

Nonetheless, due to the uncontrollable variable of different energy generation sources per location, it makes it impossible to equate between generic buildings across different locations. For comparison purposes, only site energy EUI will be used for this paper because it accounts for only the energy required to run the building. Energy production or HVAC system efficiency would add yet another variable to the comparison. Looking only at site Energy makes the comparison simpler. Reducing this base energy consumption would result in a more efficient building, independently of the further improvements or internal energy production that would be a supplementary improvement for this benchmark.

Location-specific EUI means that climate change has a significant impact on EUI due to the variations in heating and cooling costs per country. For this reason, sometimes EUI values are "weather-normalized" to compare buildings in different climates and regions. This method, however, was not applied to the models in this paper.

The building use is also important for EUI; Hospitals, for example, use more energy because of their specialized equipment, it is, therefore, it is crucial to compare buildings with similar uses to each other. In this research, an only open office building typology is used and compared.

### 2.2.3 Natural Daylight

There are many ways to measure the daylight on a building, including Daylight uniformity that is necessary for stadiums and conference rooms, Daylight access that is especially useful in public spaces, Daylight factor that is very applicable in places with substantial periods of overcast skies where daylight is sought after or Spatial Daylight Autonomy (sDA) that is used on LEED and BREEAM certifications and considers geographical location and specific weather information on an annual basis. (Advanced Buildings, 2020) Nonetheless, the metric that was used for this study Useful Daylight Illuminance (UDI) was used.

#### **Useful Daylight Illuminance UDI**

Useful daylight illuminance (UDI) is a relatively new paradigm created by Mardaljevic and Nabil in 2005 to measure annual daylight over the work plane that users would consider “useful”. The graphical percent values represent the percentage of the floor area that meets the UDI criteria at least 50% of the time. The author argues that *“UDI also preserves the interpretive simplicity of the conventional daylight factor approach. The degree to which UDI is not achieved because illuminances exceed the upper limit is indicative of the potential for occupant discomfort”* (Mardaljevic, 2005)

Consequently, this single value is especially convenient because it contemplates useful special daylight autonomy while seeking to limit possible glare occurrences. In other words, only by counting lux values below 2000 it implicitly avoids solutions with intolerable glare or excessive warmth from solar radiation. (Thomas Wortmann, 2015) The widely accepted 100-2000 lux range was used for the parametric model. (Mackey & Sadeghipour, 2019) Finally, it must be noted that this metric should be taken critically as it also depends on the depth of the floorplate; meaning that places near the windows could present good UDI while place deep in the floorplate would fair much worst.

### 2.2.4 Thermal Comfort

Even though the energy model can predict an indoor temperature, it is then important to review if this is a temperature that the users will find comfortable. We must differentiate the various temperature measurements to know which one to review and how they are related to each other:

**Air Temperature** is the average temperature of the air inside the building.

**Dry-Bulb Temperature** is the air temperature outside the building.

**Mean Radiant temperature** is the mean of the mean internal air and mean radiant temperatures of all the surrounding objects. It is the temperature of all surrounding surfaces multiplied by the angle and divided by 360°, meaning that the material’s emissivity plays a role.

**Operative (comfort) temperature** is based on both the air and radiant temperatures. It is the measurement of the average perceived temperature of each zone, thus the measurement that is affected by the other temperatures and that which is sensed by the building’s users; hence, it will be used to review comfort. (DesignBuilder, 2020)

#### **The static PMV / PPD method**

Thermal comfort is a measure to express the level of satisfaction of the user in their environment. The HVAC system’s main goal is to maintain a comfortable level of thermal comfort. When climate permits, buildings should seek natural ventilation as they typically use about half of the ones which air conditioning. (M.A Humphreys, 2002)

Through field surveys, users are asked to fill in a “comfort vote” on a scale. The scale is then categorized between 7 points, with the middle of the scale being the ideal 0 or neutral comfort. The two classification methods are ASHRAE and Bedford, but they offer similar classification.

Descriptors for the ASHRAE and Bedford scales

ASHRAE descriptor	Numerical equivalent	Bedford descriptor
Hot	3	Much too hot
Warm	2	Too hot
Slightly warm	1	Comfortably warm
Neutral	0	Comfortable
Slightly cool	-1	Comfortably cool
Cool	-2	Too cool
Cold	-3	Much too cool

Table 4 (Humphreys & Nicol, 2002)

The main issue with thermal comfort is that because of the two user variables (metabolic rate and clothing) are impossible to define and that field surveys are conducted differently, there is usually a discrepancy between rational indices and field measurements. (Humphreys & Nicol, 2002)

This method is one of the most recognized comfort models, but it is not necessarily the ideal one. This model is commonly used in buildings that utilize HVAC systems. It was developed by P.O. Fanger in the 1960s in a room under strictly controlled climate conditions. It plots the predicted percentage of dissatisfaction against the predicted mean vote. The range of international standard for the thermal limit is between 10% of people dissatisfied, this correlates between -0.5 to 0.5 PMV

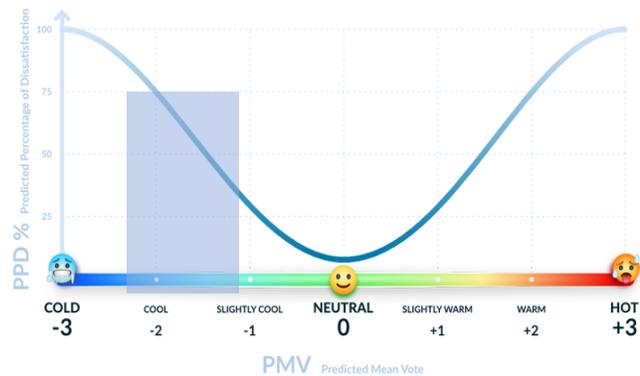


Figure 10 (SIMSCALE, 2019)

This complies with ASHRAE 55-2017. This standard is met by 80% user satisfaction and would fall within the blue rectangle like so:

For naturally ventilated or hybrid buildings other methods are more useful. This method has the potential of being dangerous for sustainability because most HVAC engineers apply this same static method in developing tropical and arid regions. Chris Mackey notes that this cycle should be broken, to avoid excessively air-conditioned buildings that by having higher cooling demands, burn more fossil fuel. Instead, passive design and natural ventilation should be sought after. (Mackey & Sadeghipour, 2019)

### Adaptive Comfort Standards

More recent adaptive comfort paradigms have much higher energy-saving potential because they assume that users tend to shift their comfort range depending on the outside temperature. This high correlation between the average outdoor temperature translates to more energy-efficient buildings because the indoor temperature setpoints adapt to the season. Although the assumptions for these set points are still relatively crude and there is still much research potential in this area. (Mackey & Sadeghipour, 2019)

According to literature, there are currently 3 adaptive standards: the American ASHRAE 55 2010, the ISO 7730, and the European EN 15251 with an updated version: NEN-EN 16798-1 for the Netherlands. They

are based on Humphreys Adaptive method that was developed using collected data from field studies where the users interact within the building.

The adaptive method assumes that if a change occurs in the environment to produce discomfort, people will react in ways that tend to restore their comfort. This implies that the comfort temperature is continually changing. In free-running buildings this comfort temperature is closely related to the outdoor temperature and can be simplified to the following formula:

$$T_c = 13.5 + 0.54 T_o$$

$T_c$  = Comfort temperature

$T_o$  = monthly mean of the outdoor air temperature

For heated or cooled buildings, the relationship is more complex because the user presence starts to play a role. In these types of buildings, consequently, the indoor comfort temperature will change with the seasons as the users adjust their clothing to the weather, thus an “adaptive algorithm” is proposed by the authors. It is a more precise method than the ASHRAE adaptive as it relates the set-point directly to the running mean of the outdoor air temperature.

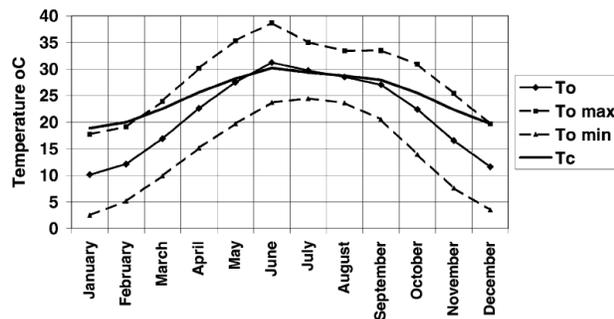


Table 5. Example of interior comfort temperature related to exterior mean temperature (Humphreys & Nicol, 2002)

Recent research suggests that this indoor variable does not increase occupant discomfort yet reduces energy demand. With this proposed method, up to 18% savings are expected for the European Union (Humphreys & Nicol, 2002)

Later, this study is even more radical. It asks if buildings should need to specify indoor climate. It argues that the characteristics of a building’s controls and building management concerning the local climate are more meaningful.

This study seeks to improve passive building parameters to improve thermal comfort, however, this does not mean that it should stop there. Humphrey et al conclude that when users are given other means of control to suit their conditions their thermal discomfort is lowered; thus, traditional curtains or window blinders continue to be just as important today.

The thermal comfort field of study is still being actively updated and developed nowadays, as there exists no unifying model for hybrid passive buildings with AC so there is a great potential for additional research in this field. On the other hand, this adaptive method is not a good predictor for calculating outdoor comfort, rather the UTCI is recommended.

## 2.2.5 Conclusions of Building Physics

The following are general conclusions found in literature about efficiency in high-rises, these findings will eventually be compared to the results given by the surrogate model to verify its reliability: Having the spaces that need more natural light oriented towards the north direction in the northern hemisphere leads to a decrease in the energy consumption required from artificial lighting. (Dobbelsteen, et al., 2007)

There is a strong connection between energy consumption and the shape coefficient (60) meaning that buildings with higher relative compactness, have lower perimetral façade area, thus less energy is lost for cooling or heating. (Raof, 2017) For this reason, for energy purposes, the ellipse is usually the most efficient in all climates and the Y shaped plan is the least efficient in all climates. Likewise, a 0°-degree rotation from the north tends to be the most efficient for all climate types. (Raji, et al., 2017) The effect of plan shape on building energy consumption is the highest in the sub-tropical climate (15.7%) and is lowest in the temperate climate (12.8%) and tropical climate (11.0%). (Raji, et al., 2017).

Finally, Babak et Al made an extremely useful general table that shows the improvement of energy savings by summarizing the effects of various building ratios, shapes, and orientations regions:

Plan shape	Temperate				Sub-Tropical				Tropical			
	A	B	C	D	A	B	C	D	A	B	C	D
<1%	Ellipse	+	+	+	Ellipse	+	+	+	Octagon	+	+	+
	Octagon	+	+	+					Ellipse	+	+	+
									Circle	+	+	+
1-5%	Circle	+	+	+	Rectangle	+	+	+	Square	+	+	+
	Square	+	+	+	Octagon	+	+	+	Rectangle	+	+	+
	Rectangle	+	+	+	Circle	+	+	+	+ shape	+	+	+
5-10%	Triangle	+	+	+	Square	+	+	+	Triangle	+	+	+
	Atrium	+	+	+	Z shape	+	+	+	Courtyard	+	+	+
	U shape	+	+	+	Courtyard	+	+	+	Z shape	+	+	+
	+ shape	+	+	+	H shape	+	+	+	H shape	+	+	+
>10%	H shape	+	+	+	U shape	+	+	+	U shape	+	+	+
	Z shape	+	+	+	Triangle	+	+	+	Y shape	+	+	+
	Y shape	+	+	+	+ shape	+	+	+				
					Y shape	+	+	+				
MD (%)	12.8				15.7				11.0			
<b>Plan aspect ratio</b>												
<1%	1:1, 2:1, 3:1				3:1, 4:1				1:1, 2:1, 3:1			
1-5%	4:1, 5:1				1:1, 2:1, 5:1, 8:1				4:1, 5:1			
5-10%	8:1				10:1				8:1, 10:1			
>10%	10:1				---				---			
MD (%)	12.4				6.0				8.8			
<b>Plan orientation</b>												
	1:1	3:1	5:1	10:1	1:1	3:1	5:1	10:1	1:1	3:1	5:1	10:1
<1%	0°	0°	0°	0°	0°	0°	0°	0°	45°	45°	0°	0°
1-5%	45°	135° 45°	135°		45°	---	---	---	---	135° 45°	45° 135° 90°	45°
5-10%	---	90°	45° 90°	135°	---	45° 135°	---	---	---	---	---	135° 90°
>10%	---	---	---	45° 90°	---	90°	45° 135° 90°	45° 135° 90°	---	---	---	---
MD (%)	1.2	5.6	8.4	15.1	2.1	12.3	20.4	32.0	0.7	2.8	4.7	7.9
<b>WWR (%): deep plan (1:1)</b>												
	N	E	S	W	N	E	S	W	N	E	S	W
<1%	10-90	35-60	65-75	10-15	10-15	10-20	10-70	10-20	10-50	10-20	10-80	10-20
1-5%	---	10-35 60-90	10-65 75-90	15-90	15-50	20-90	70-90	20-90	50-90	20-90	80-90	20-90
5-10%	---	---	---	---	50-80	---	---	---	---	---	---	---
>10%	---	---	---	---	80-90	---	---	---	---	---	---	---
MD (%)	0.5	2.8	1.8	4.5	11.3	2.9	1.1	3.1	2.9	3.3	1.1	3.0
<b>WWR (%): narrow plan (5:1)</b>												
	N	E	S	W	N	E	S	W	N	E	S	W
<1%	10-70	---	25-35	---	15-40	---	10-40	---	10-35	---	10-55	---
1-5%	70-90	---	10-25 35-55	---	10-15 40-75	---	40-90	---	35-70	---	55-90	---
5-10%	---	---	55-85	---	75-90	---	---	---	70-90	---	---	---
>10%	---	---	85-90	---	---	---	---	---	---	---	---	---
MD (%)	3.0	---	10.3	---	6.8	---	5.2	---	8.6	---	3.2	---

Energy efficiency of design options: <1% (remarkable energy saving); 1-5% (average energy saving); 5-10% (low energy saving); >10% (not recommended). A: High space efficiency; B: Aerodynamic form; C: Narrow plan (NV & daylight access); D: Less material use for external envelope; MD: Maximum deviation; N: North orientation; E: East orientation; S: South orientation; W: West orientation.

Table 6 – Early-stage design considerations for energy efficiency of high-rise office buildings (Raji, et al., 2017)

## Envelope

The solar heat gain coefficient is negatively related to the indoor comfort index, therefore a careful balance must be found. The visible light transmittance (VLT) must be maximized to reduce the daylight of unsatisfied time (DUT) (Chen, et al., 2016)

In general, temperate climates have the most complex solutions because of the complex trade-offs between energy demands (cooling, heating, lighting). Very hot climates have the simplest solutions because the main objective is to reduce solar gain. (Evins, et al., 2012)

Shading elements / Overhangs should be designed in conjunction with the glazing area as they are closely related. (Evins, et al., 2012)

This table shows the relationship with the parameters and what outcomes they affect:

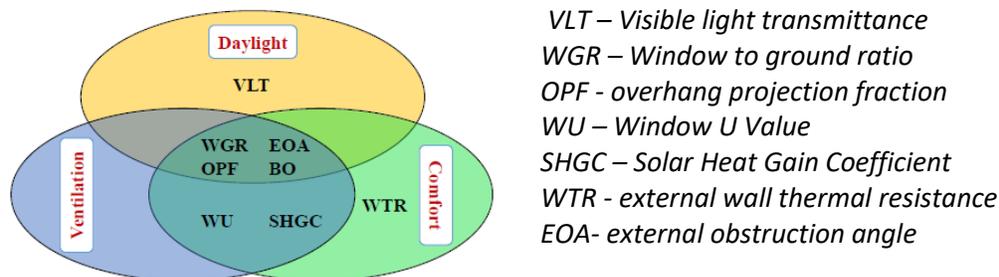


Figure 11- The contribution of design variables to indoor assessment indices (Chen, et al., 2016)

Babak Raji et Al performed a sensibility analysis concluding four of the most influential envelop parameters that affect the buildings energy performance: their chosen parameters were glazing type, window-to-wall ratio, sun shading and roof strategies (Raji, et al., 2016) As high rise buildings typically have a low roof to façade ratio, roof strategies will not be considered in this study.

For their case study of a typical residential building in Hong Kong: “As the building orientation turns from north to south, the thermal comfort time decreases whereas the daylight and ventilation performance is enhanced.” It is important to note that, overhangs, wall thermal resistance (WTR), infiltration air mass flow coefficient (IAMFC), and wall specific heat (WSH), had a weak influence over most performance indices (Chen, et al., 2016). For this reason, none of these parameters were considered except for the overhangs as they play a big role in the architectural expression of high-rises.

In a more extreme matter regarding overhangs, in a case study for a high-rise building in the Netherlands, the overhangs surprisingly had a minimal detrimental effect on the total energy consumption; yet it is important to note that it could also affect discomfort such as glare for the occupants. (Raji, et al., 2016) This means that overhangs/shading devices are elements that should be considered very carefully before implementing due to architectural aesthetic tastes or preconceived designers’ ideas. Careful consideration that can only come through sensitive site condition analysis.

Later though, Babak Raji et Al state that external shading (such as outdoor blinds) generally performed better in terms of energy-saving and solar control in all climates. In northern hemispheres, a southern facade requires overhangs or fixed (stable) blinds, whereas east or west facades require more dynamic shading due to the critical low sun angle in early and late day times. (Raji, et al., 2017)

Noteworthy is that high-performance design solutions produce considerable energy savings: 42% total energy, 64% for heating energy, and 34% for electric lighting energy. (Raji, et al., 2016)

In a subsequent study by the same authors, their general conclusion is that the optimal WWR range of the window-to-wall ratio is **20–30% in the temperate climate**, 35–45% in the sub-tropical climate, and 30–40% in the tropical climate, 50% in a hot climate. (Raji, et al., 2017)

It is important to note, however, that this study also includes a table where they specify the ideal WWR for each cardinal direction and climate location:

Climate Type/Plan Aspect Ratio		Temperate		Sub-Tropical		Tropical	
		1:1	5:1	1:1	5:1	1:1	5:1
Recommended WWR value (%)	North	10–90	10–70	10–15	15–40	10–50	10–35
	East	35–60	No glazing	10–20	No glazing	10–20	No glazing
	South	65–75	25–35	10–70	10–40	10–80	10–55
	West	10–15	No glazing	10–20	No glazing	10–20	No glazing

Table 7 - Recommended WWR value for different orientations and climates. (Raji, et al., 2017)

All in all, these general findings and summary tables give a good background knowledge of the overall expected results. These tables give general ranges during early-stage design, yet the designer is not able to visualize immediately the impact of each of their decisions on the building efficiency, thus the reason to develop this thesis. Similarly, these solutions are mainly aimed at solely lowering building energy consumption, yet daylight and thermal comfort are not the focus, further reaffirming the need for an interactive and optimizable High-rise tool.

### 2.3 Performance-Based Computational Design

One approach of performance-based computational design used for this thesis is called the performative computational architecture (PCA) framework (Ekici, et al., 2019), focusing on the initial stages of the architectural design and being composed of main 3 steps:

1. Form Generation / Parametric modeling

Is the process where virtual models are defined by parameters that shall be analyzed in the following step. Here, the solution space is determined. The number of variables and complexity of each must be carefully considered to avoid an excessively high computational cost during step 2.

2. Performance Assessment / Building Performance Simulations (BPS)

Is the step where software mimics a building's behavior to simulate how it would behave in real-life. Once this has run, the performance of one iteration can be evaluated. The validation of this model and the choice of indicators is key to meaningful results.

3. Computational optimization

In this step, algorithms are used to find a solution that better fits within the defined objective(s), whether it means to maximize an ideal parameter or to minimize a negative one. It could also imply that the optimization is carried out within a certain set of objective constraints

### 2.3.1 Form Generation

To achieve better buildings, that reduce energy consumption and their greenhouse gas emissions, it is necessary to foresee a building's physics and performance; for this, Performance-Based computational design is a field of study that has been growing in recent years.

Conversely, when looking at the whole design, the complexity of a real-world problem is enormous and can entail many variables presented in section 2.1.4 such as social, historical, and economic factors that are difficult to define objectively in the computational realm. For this reason, the basis of design theory must be addressed.

The Co-Evolution paradigm states that architectural design is too difficult to “solve” objectively because of its subjective and intuitive nature. On the other hand, the opposing Generate-and-Test paradigm defines architectural design as well defined, rule-based and structured process. Within this thought process, a set of potential design candidates for a given problem definition is known as the “design space”. This design space provides the architect with a framework where multiple design solutions can be explored until a suitable design is recognized but also where the design problem can be better understood by its user. (Wortmann, 2018). In this sense, performance-based design recognized the complexity of the Co-evolution paradigm yet seeks to serve as an additional tool for the architect or decision-maker to develop a more informed choice while considering the complete problem at hand.

In the PCA framework firstly, the design options must be generated automatically, usually through parametric design, scripting, and/or programing, only then can these options be run through the necessary computational optimization process later. This first step defines the basic rules for the model and limits the design space. The solutions are therefore limited by the initial set of rules defined by the computational designer and/or architect. These “rules” are called the decision variables that can be continuous, such as a number that gradually increases or discrete, such as a set choice of the parameters. Similarly, the outputs must also be defined, this is called the problem formulation, in which the objectives such as minimizing or maximizing the outputs are determined from the start.

#### **Complexity**

A term coined “curse of dimensionality” explains that there is a frequently reported problem: Inputs and outputs must be limited in to manage the computational cost. (Westermann & Evins, 2019) This “curse” is a recurring subject in the third step of the computational optimization process therefore the complexity of the problem must be considered initially in the form generation step.

Being this the first stage of the workflow, the time that will be required to generate the model and simulate its behavior remains a critical aspect. The “cost” of the following steps: performance assessment and optimization, is a factor to consider at this stage. An evaluation's cost refers to how many resources are needed to evaluate the function. A cheap function can be evaluated thousands of times, while a costly one can be evaluated fewer times. When the evaluation of the objective function is a time-expensive simulation, the next steps would suffer significantly. A sufficiently simple yet useful model is therefore key in form generation. The model must be carefully set up because there will always be trade-offs between speed, accuracy, and complexity.

For this reason, discrete parameters were chosen for this thesis, showing the overall behavior of the model with sufficiently distanced steps between the decision variables.

## Sensitivity Analysis

A sensitivity analysis is typically a method used in early-stage design to define which parameters influence most of the outcome. This is especially important to identify the parameters that play the biggest role in the results. The parameters can be ranked in the level of influence; thus, the least significant parameters could be ignored in further, more refined models. By reducing the variables, the model becomes much less complex and the simulation can, therefore, be less computationally expensive. The study of multiple parameters is Total Sensitivity. Although a sensitivity analysis was carried out for this work, the use of surrogate models made it less critical, because the computational cost of running a surrogate model is much less than that of a simulation-based model.

### 2.3.2 Performance Assessment

Once the form is generated, the next step is to test its simulated performance through Building Performance Simulation (BPS) (Brea, et al., 2020) to then review its accuracy. BPS uses a physics-based engine that replicates how a building would behave in the real world and predicts the results, nonetheless, the input parameters must be carefully finetuned to achieve coherent results, otherwise, the “garbage in / garbage out” paradigm arises, meaning that if incorrect information is used as input, incorrect and unrealistic building simulations will be generated. For this reason, the validation of the simulation model was a crucial step, see section 3.3.

### Physics-based Engines

For the energy and comfort simulations, Energyplus (DOE, BTO, NREL, 2020) is used. EnergyPlus is a building energy analysis engine developed by the US Department of Energy. It is a simulation engine that has been around for many years and showed up extensively on scientific papers for research, making it a trustworthy option (File type: .idf). For daylight, DAYSIM is used (MIT, 2012). DAYSIM is a validated, RADIANCE-based daylighting analysis software that models the annual amount of daylight in and around buildings. (MIT, 2012) Although these tools operate independently, they can be used within the Rhino/Grasshopper environment through additional plugin components called Ladybug/Honeybee tools, the advantage of this coupling is that they can be linked to the Parametric model from the previous section, thus allowing testing the model’s various alternatives and subsequently setting up the stage for the next phase: the optimization. Honeybee uses Openstudio® ( NREL, ANL, LBNL, ORNL & PNNL, 2020) API (application programming interface) to connect with EnergyPlus (DOE, BTO, NREL, 2020) engine and run the simulations. (File type: .osm).

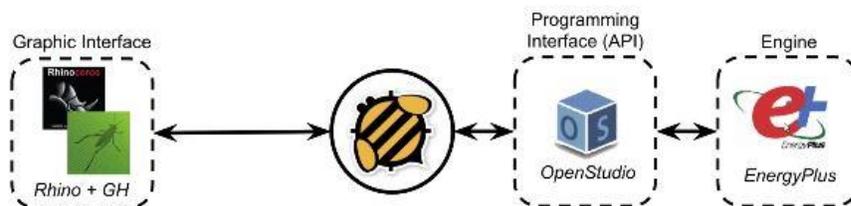


Figure 12 Tool Interoperability (Mackey, 2019)

### 2.3.3 Computational Optimization

Following the automated form generation and the simulation of the designs, theoretical and practical knowledge of the optimization process is also required. Optimization in building design has been used in many different architectural domains such as acoustic, structural, space layout, construction, HVAC systems, control designs, etc. (Kheiri, 2018); nonetheless, this thesis focuses on the optimization quest for energy-efficient buildings within the daylight and comfort framework.

During the literature review, the two encountered search strategies are single (SOO) vs multi-objective optimization (MOO) (Wortmann, 2018). Single-objective problems are much simpler because they only seek to improve one goal. Conversely, MOO problems, as their name indicates, must consider multiple objectives simultaneously, drastically increasing the complexity of the problem.

Mathematical optimization or mathematical programming is simply the selection of the best element by meeting a criterion from a set of alternatives. The most basic optimization problem consists of maximizing or minimizing a real function; nonetheless, due to the complexity of the model or the high cost associated with running it, usually for the designer it is more advantageous to find a “sufficiently good” solution rather than the best one, thus metaheuristic optimization is used. (Wortmann, 2018) Simply put, metaheuristic is a set of clever strategies to improve the search for acceptably good solutions.

Multi-objective optimization typically uses heuristic algorithms to explore near-optimal design options for different, sometimes even conflicting objectives. As these objectives are defined from step 1, it is necessary to have some criteria to evaluate it, the criteria are expressed as functions of the decision variables, also called objective or fitness functions. Only then can a design solution be classified as good or better.

A building optimization for a multi-objective problem can be expressed mathematically as:

$$\min f_m(x) \quad m = x_1, x_2, \dots, x_M;$$

$f_m$  denotes the objective and  $x$  is the set of  $n$  building design variables. With multiple objectives, the solution of the multi-objective problem means that there will be a trade-off between objectives. When these objectives are conflicting, they make up a multidimensional space  $Z$  (in contrast to a single-objective optimization), this means that there is no single optimal solution. (Brea, et al., 2020)

In multi-objective optimization the Pareto-Front (Cenaero, 2018) is the set of optimal design solutions that are not dominated by other solutions, meaning that, in what's known as the “Pareto-front”, one objective cannot be improved further without reducing or deteriorating the other. The Pareto frontier will always seek to approximate to a utopian objective vector  $z$  where both objectives are improved.

The optimization aims to aid the designer to learn and understand the trade-offs between the objectives and find an acceptable “balanced” solution.

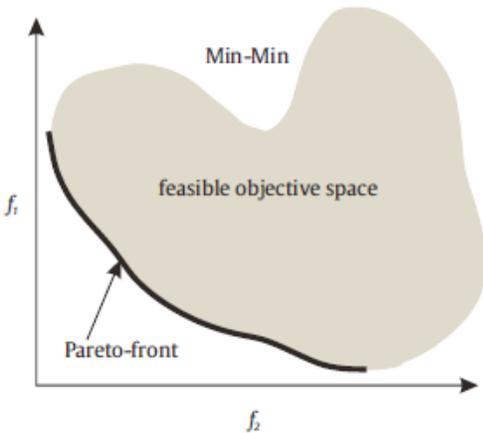


Figure 13 Example of Pareto front on a bi-objective optimization problem (Facundo Brea, 2020)

One of the most used performance indicators for the optimization algorithm in architecture is the resulting area between the Pareto frontier and the utopian objective, this area is known as the hypervolume. (Auger, et al., 2012)

### Meta-Heuristic Optimization algorithms

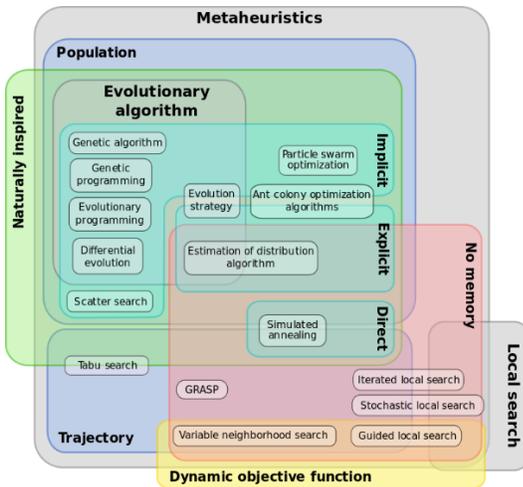


Figure 14 (Nojhan, 2007)

As seen in the image above, there are many types of optimization algorithms, yet according to a review of optimization in the architectural domain, the most common ones used are evolutionary computation (EC) and swarm intelligence (SI) (Ekici, et al., 2019)

### Constraints

In all optimization problems, there are restrictions imposed by the characteristics of the environment, available resources, physical limitations, time restrictions, etc. These restrictions must be satisfied to consider a certain solution acceptable. The formulation of the problem in form generation is already the first constraint because only the solutions that the parametric model can recreate will be considered. Thus, the parameters become the first constraint. The second constraint is when, within the design space the decision-maker decides to narrow the search by limiting the preferred solutions within the objectives he deems acceptable, such as comfort above a certain benchmark.

## Local and Global Optima

In any optimization problem, four important terms were key to understanding the search for an optimal design: Local vs Global optima and Robustness vs Accuracy.

In the search for optimal solutions for a problem algorithms risk on remaining to look for optimal solutions in a design space known as the local optima. This might seem like the ideal solution, but other, better solutions might be available in what is called the global optima, which is the definitive goal of any search algorithm.

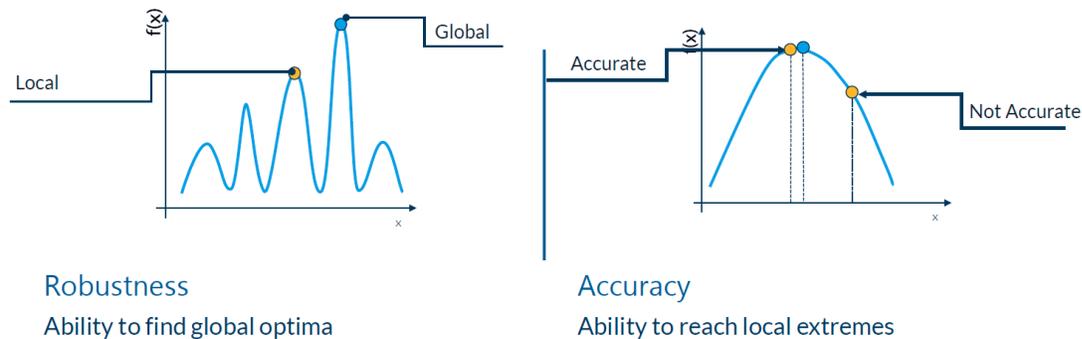


Figure 15 Robustness vs Accuracy (ESTECO, 2020)

Robustness vs accuracy is contradictory terms when looking for this optimum, as some robust algorithms will find inaccurate options but identify the global optima while others might find accurate options in local extremes. (ESTECO, 2020) This tradeoff is inherent in all search algorithms. Although robustness and accuracy are key during optimizing, the third important criteria to consider is an efficient computational cost.

## Optimization Platforms

There are many optimization platforms: free academic versions are AMPL, GAMS, MPL, AIMMS. Other platforms work from a server: NEOS or are open-sourced: COIN-OR or GNU Octave (Hakanen, 2015). Yet more options were encountered during the research including Matlab, Scilab, SciPy, Wolfram Alpha, Lingo. Other optimization methods come as plugins within Grasshopper environment such as Galapagos, Octopus, Dodo, Wallecei among others. Finally, process integration platforms that enable various modeling/simulation and optimization platforms in a single platform are Optimus, ModeFRONTIER, BOSS Quattro, OptiSLang, and Heeds (Hakanen, 2015).

#### 2.3.4 Relevant Precedents

Ayca Kirmat et al studied the effects of amorphous shading devices in buildings using multiple objectives such as lowering the Total Energy Consumption while maximizing the Useful Daylight Illuminance. This study compared two different algorithms NSGA II against JcGA-DE, an energy saving of up to 14% was achieved while maintaining daylight availability above 50%. (Kirmat, et al., 2019)

Although relevant, this study only looked at the shape of the shading devices towards the south façade and did not consider possible additional savings by reviewing the general building shape or orientation.

Farshad Kheiri published an extensive review of the optimization methods for energy-efficient buildings where the objective was to lower building energy while considering both envelope parameters and geometric configurations. It gave a good overview of various optimization algorithms. They briefly mention the use of surrogate models and a concept in Machine learning coined “neuro-evolution”, that describes the use of evolutionary algorithms to train an artificial neural network or similar ML method. EnergyPlus simulation tool was used most frequently and more than half of the studies that used visualization tools used Rhino and the next widely used was Design builder, all software that was used for this thesis. They conclude exactly what this thesis seeks to explore further:

“The trade-off between accuracy and time is still a challenge. Implementation of different machine learning techniques has assisted accelerating the optimization process whereas the applicability of many meta-models is limited by the utilized trainset. Future research can improve creating more robust models to estimate building energy performance with diverse configurations.” (Kheiri, 2018)

Just last year, researchers from TUDelft conducted a comprehensive review of the use of swarm and evolutionary algorithms applying optimization on the built environment. This review served as the basis to review what had been done in the building optimization realm and what could be further explored. Important conclusions from this article state that “only three sources focused on different building topics as part of the same optimization problem such as the skin and layout” thus presenting promising potential. Furthermore, it concluded that due to the expensive computational time, limited studies included tall buildings yet “objective functions based on ANN could be an effective solution”. (Ekici, et al., 2019) This reinforced the idea of exploring the use of surrogate modeling using envelope and shape parameters simultaneously as a topic for this research.

This year, Evangelia et al aimed to reduce the energy consumption on a typical central-core, open plan high-rise office building; in this case, for the Mediterranean climate (CsA) of Athens, Greece. Similarly, energy simulations with DesignBuilder were used as a benchmark for the Honeybee/Ladybug Grasshopper component. Of the optimized variables, WWR, U-value, and external shading area, g-value were also parameters reviewed. On the other hand, it also included airtightness of the façade, cooling setpoint, and PV surface area. Nonetheless, designs were simulated in two distinct optimization rounds, and limiting the search space to 1000 designs. Likewise, the shape was limited to four options: 2 rectangles (3:1, 1.5:1), 1 square (1:1), and an octagon (1:1) while the orientation to 4: East/West, NW/SE, North/South, NE/SW. This paper focused on one specific location and did not contemplate the effects of the urban context. Also, differently from this work, more than one optimization algorithm could be compared because a lighter surrogate model didn't render the optimization too computationally expensive. (Evangelia Despoina Giouri, 2020)

## 2.4 Surrogate Modeling & Machine- Learning

This thesis aims to predict real numbers based on collected data from a simulation model, therefore the supervised regression type of machine learning is utilized. Through machine learning methods, surrogate models are developed.

On the other hand, this thesis does not focus on the accuracy of the models but rather discusses its applicability in the context of simulation-based optimization. And although the use of the validation metrics presented in section 2.4.4 is necessary to develop trusted models, the method and approach of using these models for design exploration and optimization is the main goal of this paper.

### 2.4.1 Surrogate models in architecture

In the frame of this thesis, a surrogate model is as meta-model of the original simulation model, simply put, a model of a model. As they seek to approximate or mimic the original results, their main challenge lies in achieving this behavior as closely as possible and representing the full spectrum of the design options. For this reason, the sampling method, sampling count, and validation metrics are vital for a good surrogate model.

As a workflow, the surrogate model is generated through a machine learning method that reads the results from the simulations to build a statistical model. This simplified version of the model can be then used at a fraction of the computational cost to predict results, and thus bypassing the traditional simulation step:

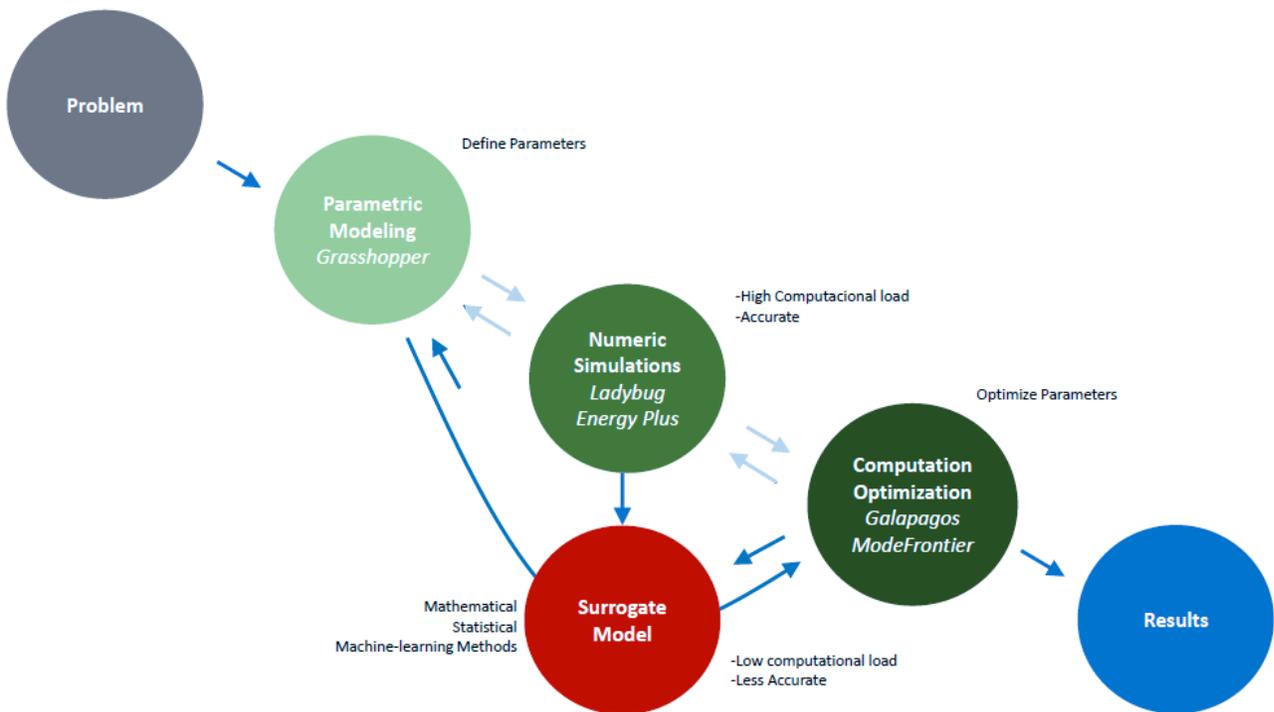


Figure 16 Simplification of a surrogate model workflow (source: own)

Through the literature study, various surrogate model types were identified: ANN (Artificial Neural Networks), RBF (Radial Basis Function), Gaussian Processes, Linear and Non-Linear Regression, MARS, SVM (Support Vector Machines), Random Forests, Ensembles, among others.

The main advantage of using this Meta-model approach lies in the time and computational cost that can be potentially saved with its use. Facundo Brea et al show a diagram of how, even though the first steps of Form generation, optimization algorithms, and Physics-based BPS (Simulations) (Wortmann, et al., 2015) are the same, the surrogate model can produce results faster because fewer total simulations are required. Albeit, two additional, speedier steps are still necessary: Training and Evaluating. (Brea, et al., 2020)

Training the Meta-model model is optimizing its hyperparameters in a way that the predicted results are closer to the results the simulations would provide.

Evaluating the metamodel is done through reviewing and optimizing its validation metrics presented in section 2.4.4.

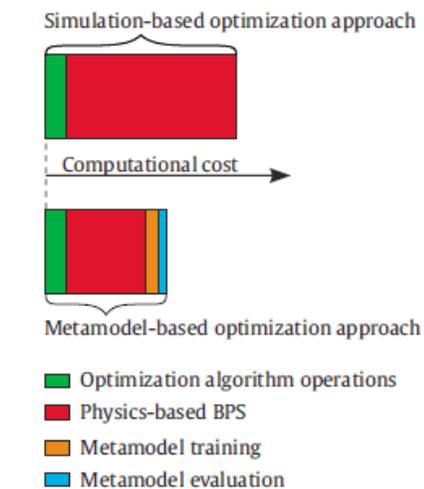


Figure 17 Computational cost: simulation vs metamodel approach (Brea, et al., 2020)

As surrogate models handle such large amounts of data, this information must be made to be visually understandable images to humans. Thomas Wortmann presents a review of visualization graphs to better understand the design space and potential solutions. He concludes that design optimization should be viewed in a wider context, not simply as a methodology to find the “best” performing solution but a way to better understand the relationships between design variables and the performance criteria, therefore a graphical and visual representation of the design options is a critical tool to comprehend these relationships. (Wortmann, 2017). : “Further avenues of development are a user-friendly implementation and a visual interface that facilitates the exploration of the (space of) solution spaces.” (Wortmann, et al., 2015)

### Developing a Surrogate model

An overview of the steps required to create a surrogate model are:

1. Problem definition
2. Implementation of a building model
3. Simulation of the samples based on a sampling strategy
4. Collection of samples to a dataset
5. The Surrogate model is trained (fitted)
6. Surrogate model is evaluated / validated (Westermann & Evins, 2019)

### Potential Time Savings

The use of machine learning methods to develop surrogate models shows very promising results in reducing the amount of computation time needed, thus significantly improving the information available in early-stage design. Through the literature review, studies mention the savings compared to a simulation-based approach:

- From more than seven hours to a few seconds (Chen & Yang, 2017)
- Saving up to 90 days (Ekici, et al., 2019)
- Saving up to 75% of the simulations. (Brea, et al., 2020)

### 2.4.2 Design of experiments

Design of experiments (DoE) is an analysis tool aimed at organizing the information that is collected to study the relationships between the input variables and the output aka dependent variables, simply put, DoE is a sampling strategy. Literature shows that there are two strategies for collecting data: sequential vs iterative (Westermann & Evins, 2019). In 2017 researchers present a general review of the Design of computer experiments applied to a wide range of subjects, not only the built environment. (Garuda, et al., 2017) Similarly, to sequential vs iterative, they classified DoE into two broad categories: Static vs Adaptive. Static / Sequential sampling first collects defined by a sampling strategy and uses this dataset to train the model; it is more widely used nowadays. Adaptive / Iterative sampling picks samples by identifying parts of the design space that need further accuracy. (Westermann & Evins, 2019)

### 2.4.2 Plugins for Machine Learning / Surrogate Modeling

The following plugins for Grasshopper include Machine learning components for developing the intended surrogate models for supervised regression:

#### *Octopus*

Octopus was initially created as a Multi-Objective Evolutionary Optimization; it is based on SPEA-2 and HypE optimization algorithm from ETH Zurich. It also includes Support Vector Machines from statistical mathematics, using the 'kernel trick' for dimensionality reduction. Likewise, it includes the Artificial Networks component by a multi-core resilient propagation algorithm, using the Encog library by Jeff Heaton. Only the original multi-objective search and optimization module was utilized for this thesis. (Vierlinger, 2018)

#### *ANT*

ANT is a plugin that takes advantage of the "Skicit Learn" python module. This module includes many machine learning algorithms including Lasso (LARS, Elastic Net, k-Nearest Neighbors, PCA, linear and logistic regression, stochastic gradient descent as well as Support vector classification and regression. (Rahman, 2017)

#### *Lunchbox*

The Lunchbox plugin counts with many tools for paneling and shapes as well as machine learning algorithms such as regression analysis, clustering, and networks using the Accord.NET framework. (Miller, 2018). This tool proved the most useful due to its widely available information, tutorials, and examples. It was used for the validation and linear and non-linear regression sections of this research. (Miller, 2019)

#### *Dodo*

"Dodo is a collection of tools for machine learning, optimization, and geometry manipulation Regarding AI, it features neural networks, gradient descent, stochastic gradient descent, and swarm optimization. Dodo has tools for scalar, vector, and tensor fields manipulation which can be visualized using isosurfaces." (Lorenzo, 2019) Ultimately, the ANN component of this plugin was the one chosen for the final surrogate models because of its speed and predictive results once properly trained. See Section 3.4.

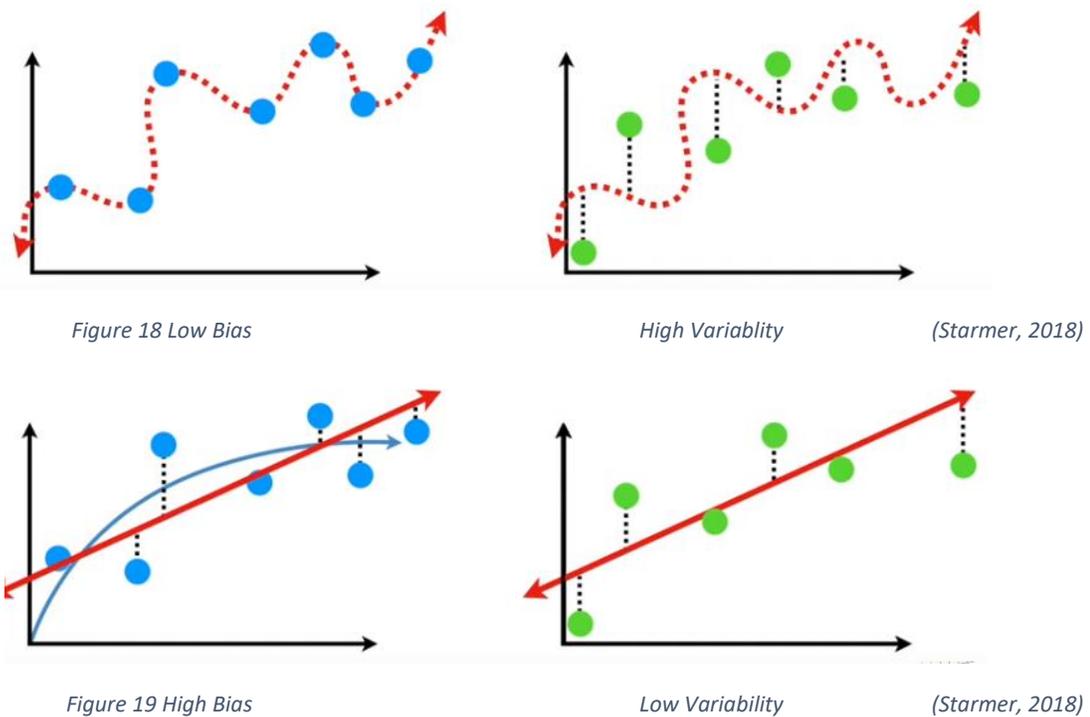
### 2.4.3 Concepts

#### Overfitting and Underfitting

Overfitting is when the model is too closely related to the training data, resulting in less predictable results from new data. (Oxford, 2020) Underfitting, conversely, occurs when a model cannot predict correctly the underlying structure of the data, also failing to give more accurate results. A good surrogate model is when there is a balance between the “tendency” of the model with fewer errors. To avoid this, overtraining or undertraining must be evaded. Overtraining happens when the predictive “line” is forced to follow too closely all the data points, while undertraining usually occurs when there are not enough samples therefore the model is not precise enough. Likewise, trying to fit a linear model to an exponential curve, for example, is also a case of underfitting because although the data is enough in this case, a more accurate non-linear curve would better predict the results. (Wikipedia, 2020)

#### Bias vs Variability

In general, Machine learning is about making predictions. It requires input data, aka training data to “learn” what predictions to make. Then, testing data is used to compare the predictions to the generated predictive model. The training data will not always fit perfectly because it must also account for the predicted data. The resulting poor predictions are called the Bias-Variance Trade-off. Although the data used for training should not be the same used for testing because, otherwise, there would be no way to know if the predictions are working, for this reason, the data must be cross-validated. In machine learning lingo, the difference in fits between the data is called variance. This variance is critical for understanding the behavior and precision of the machine learning method. (Starmer, 2018)



#### 2.4.4 Validation Metrics

To achieve a precise and statistically significant surrogate model, it is essential to understand the metrics to validate the outputs of the machine learning method. First, commonly used terminology:

- Inputs = X = independent variables
- Outputs = Y = dependent Variables
- Predicted (aka Fit) Outputs by the machine learning method = Y hat or  $\hat{Y}$

Next is a brief description of the evaluation metrics used to determine how well the predicted outputs (Y hat) match the originally measured outputs (y). The most important terms are:

#### **R / Pearson coefficient**

The most commonly used coefficient of correlation (R) also known as the Pearson coefficient shows how strong a relationship is between two variables. A negative 1 shows a strong negative correlation and a positive 1 shows a strong positive correlation. R is not as intuitive as  $R^2$  that shows the percentage of relationships that can be explained. (Starmer, 2018)

#### **$R^2$ and Adjusted $R^2$**

Simply squaring the value of R gives the  $R^2$  Value, also known as the coefficient of determination. Because squared values cannot be negative, it no longer explains the direction of the correlation, nonetheless, this single value explains the percentage variance of the results that can be explained by the independent variables, in other words, its “predictive power”. (Starmer, 2018) Example:  $R^2$  80% would mean that 80% of the results can be explained by the surrogate model. Therefore, the closer to 100% the more precise. Attentively it is presented as a decimal. 80% = 0.8

The way to compute the  $R^2$  value is by calculating the variance of the trained inputs Y as well as that for the predicted (fit)  $\hat{Y}$ . (Anderson, 2020) The formula for variation is:

$$\sum_{i=1}^n (Y_i - \bar{Y})^2$$

$\bar{Y}$  = mean of the samples

$\hat{Y}$  = predicted data

$Y_i$  = training data

The Predicted  $\hat{Y}$  variance is then divided by the measured Y variance. This is the  $R^2$ .

Alternatively, R squared can be calculated using the sum of squares (SS) looking at either the RR Regression or SS Residuals divided by the SS total: (Wikipedia, 2020)

$$\frac{SS_{reg}}{SS_{tot}} \quad \text{or} \quad 1 - \frac{SS_{res}}{SS_{tot}}$$

## Adjusted R<sup>2</sup>

Once the R<sup>2</sup> is validated, the Adjusted R<sup>2</sup> metric is necessary when multiple variables are considered, it is adjusted to the sample size. The more samples, the better results; and the more variables, the higher the penalty. When comparing 2 different models, the Adjusted R<sup>2</sup>, represented as ( $\bar{R}^2$ ) is a more accurate measurement. (Anderson, 2020)

$$\bar{R}^2 = 1 - (1 - R^2) \left[ \frac{n - 1}{n - (k + 1)} \right]$$

*N = the sample size (500)*

*k = the number of independent variables*

## MSE / RMSE

This metric is frequently confused therefore another terminology clarification must be made:

- Mean square error (MSE) = Mean squared deviation (MSD) = Standard Error of Regression (SDR)

Similarly, related, but not the same is:

- Root mean square error (RMSE) = Root Mean square deviation (RMSD) aka Standard error of the mean (SEM) = and sometimes only referred to as the “Standard Error”

The MSE tells us on average how much each observation is missing the prediction. The higher, the worst fitting the model, therefore values closer to 0 is always preferred.

The variance is a squared metric so it cannot be plotted to the same histogram as the data so the new term, the (SDR) is required. MSE measures the quality of an estimator as so:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

As MSE and RMSE are means, they depend on the number of samples (df Total) and the number of variables, also known as the Degrees of Freedom (df). Subtracting these (n-1) gives the degrees of freedom of the regression, used to compute MSE. For this thesis, this was also computed with a custom GH script (see method). The RMSE is simply the squared root of the MSE, thus normalizing the calculation with the data. (Dekking, 2005)

## MAE - Mean absolute error

The mean absolute value is a similar output to MSE and is also used to validate the accuracy of the model. Instead of squaring the variance and then calculating its squared root, MAE simply finds the absolute value of the results. (Willmott & Matsuura, 2005)

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{Y}_i - Y_i|$$

## P-Value

Once  $R^2$  has been determined, it must also be reviewed that this information is statistically significant, meaning that the results simply did not happen by chance. (Koehrsen, 2018)

Any experiment starts with a null hypothesis ( $H_0$ ) meaning that one assumes that there is no significant difference between the data. Only if at least one of the variables affects the data, the Null Hypothesis ( $H_0$ ) can be rejected thus we can infer that a variable affects the outcome: (Koehrsen, 2018)

- $H_0 = \text{Var}1 = \text{Var}2 = \text{Var}3 = 0 \dots$  means there is no useful linear relationship between features and outputs because the slope of the regression line is 0.
- $H_1 = \text{Var}(i) \neq 0 \dots$  If it's not a 0 however, at least one of the variables affects the data, and a relationship can be inferred, and that one variable is significant.

This value is the P-value: the smaller the values for P, the higher the confidence, in other words, the more likely the parameters affect the results. P-Value is, therefore, how much confidence can be given to the parameters affecting the results. A commonly used threshold is 0.05, meaning that only 5% of the results could be incorrect. This threshold is referred to as Alpha. Thus, it is commonly accepted that P-values  $< 0.05$  mean that inputs are statistically significant. Usually trying to reduce the number of False positives below 5% is not worth it because of the high costs. Likewise, lower P-value seeks to avoid False Positives or Negatives, meaning when the algorithm predicts classification problems incorrectly. (Starmer, 2018)

Finally, the F Statistic or Significance F is a test statistic that shows if a group of variables coupled together are jointly significant. (Glen, 2020). Therefore, the F statistic must be used together with the P-value to validate the probability significance of the results.

### 2.4.5 Relevant Precedents

As early as 2015, which in computational terms is a lot, Thomas Wortman used an optimization problem for a performative façade to assess the quality of lighting conditions (UDI). He compared the best solution found by the RBFopt surrogate model versus another solution found by more traditional Differential Evolution algorithms in Galapagos and found that RBFopt outperformed the popular genetic algorithm by almost two orders of magnitude. This study showed the proof of concept of using surrogate models to provide good solutions and contribute to a better understanding of the problem, yet it was limited to 15-panel variables with different louver angles. (Wortmann, et al., 2015)

In 2016 Simong Fong and Zhonghuan Tian wrote a chapter on applying the meta-heuristic algorithms on neural network training. This was of interest as it delved into the subject of deep learning, using the metaheuristic optimization methods learned for solving an architectural problem to rather improve the training of the surrogate model developed by the artificial neural network. It gave a general overview of the use of metaheuristics implemented in neural network training. It concluded that there is a "...high possibility of applying meta-heuristic in DL to speed up training without declining performance. However, relevant publications in this direction are still rare." (Fong, 2016) It did not delve, however into optimizing building components but rather focused only on the theory and applicability of this methodology.

In 2017, Xi Cheng and Hongxing Yang published a research paper on optimizing passive strategies of a high-rise residential building with multiple building operation scenarios. They also used machine learning methods to develop surrogate models to reduce the computational load. For optimization, they used the NSGA-II algorithm. Similarly, to this research, they choose two different representative locations: Hong Kong and Los Angeles. The same input variables where: Building Orientation, U-Value, Overhang size (overhang projection fraction), window SHGC, and VLT with EnergyPlus as the simulation engine also. They also looked at different ventilation strategies like single-sided vs cross ventilation, infiltration air mass flow coefficient and analyzed Window to Ground Ratio rather than Window to Wall Ratio. Rather the looking to improve the comfort levels, they developed a hybrid ventilation control algorithm to determine when to use HVAC or natural ventilation. Their main objective was also to reduce the total energy demand (kW/m<sup>2</sup>). Contrary to this research, however, they used a simplified approach on daylight optimization by calculating required lighting as a fraction of the rated power according to continuous/off dimming control method and where the daylight illuminance level had to simply meet a threshold of 150 Luxes rather than optimizing UDI as done in this paper. Also differently, this paper will look at an office building rather than a residential one and will include a parameter that contemplates the effects of surrounding building on the energy and daylight performance. (Chen & Yang, 2017)

Last year, Berk Ekici et Al developed a surrogate model for a high rise building subdivided into 5 different zones, thus with 5 optimization problems. The variables were overhang length and glazing types to optimize the spatial Daylight Autonomy and Annual sunlight exposure. This study proved once again the time efficiency of the surrogate model saving up to 90 days compared to the metaheuristic simulation-based approach. (Ekici, et al., 2019) Although this study investigates façade shading devices and investigates the effects of the urban context, it does not simultaneously investigate the effects of the shape or the location which was also considered in this thesis.

Just this year, researchers from Argentina and Uruguay performed multi-objective optimization using NSGA-2 coupled with ANN-based metamodels, also trained using EnergyPlus software. It showed the potential of this methodology to improve a building's energy efficiency and thermal comfort on a dwelling. It investigated 12 variables such as roof type, external and internal wall types, solar orientation, solar absorptance, size and type of windows, and area of external shadings. It concluded that up to 75% of the building's energy simulations could be reduced to find Pareto optimal designs. It reviewed the convergence of MSE based on the sample size to define better criteria on how many samples should be collected to achieve accurate results of the metamodel. (Brea, et al., 2020)

From this research, it can be concluded that some investigation has been done on building optimization using surrogate models, yet most existing research does not explore the effects of the context and climate location neither does it use the full potential of the lightness of the model to empower the designer to “play” with the model or to compare various optimization algorithms. This is a relatively new area of study that shows great potential for further improving the behavior of our future designs.

### 3.0 Method

As presented during the literature review, the original performative computational architecture (PCA) framework (Ekici, et al., 2019), can be divided into three parts:

1. Form Generation / Parametric Modeling (PM)
2. Performance Assessment / Building Performance Simulations (BPS)
3. Computational Optimization (CO)

Within the Performance Assessment, relevant previous showed how using DesignBuilder software (DesignBuilder, 2020) as a benchmark for the Honeybee/Ladybug Grasshopper component (Evangelia Despoina Giouri, 2020) gave higher reliability to the energy, daylight and comfort calculations. For this reason, this step was also considered during the performance assessment and before computational optimization. Moreover, having identified the potential advantages of speed and computational efficiency of using surrogate models, the steps in the literature review related to generating surrogate models were included between 2 and 3 to reduce the total number of BPS and improve the speed of the computational optimization. Also as identified in the literature review, an additional step: in the form of a visual interface tool, was necessary to facilitate the review and exploration of the optimized results. Finally, it was deemed important for this new tool to showcase its applicability, for this reason, one of the optimal solutions was chosen as a “showcase” example to demonstrate how the tool could be integrated into the whole PCA workflow when working on early stages of design.

The new methodology is thus composed of the following steps:

1. Problem definition
2. Simulation of the Parametric Modeling (PM) / Form Generation 
3. Building Performance Simulations (BPS) / Performance assessment 
4. Validation of Simulation Model in Design Builder 
5. Data Collection based on the sampling strategy 
6. Training of the Surrogate model 
7. Evaluation of Surrogate model 
8. Computational Optimization (CO) 
9. Surrogate model in Design Process 
10. A showcase of a design Solution 

The following figure shows a summary of the tools used for each of the steps defined in the general method defined above.

### 3.1 General Workflow

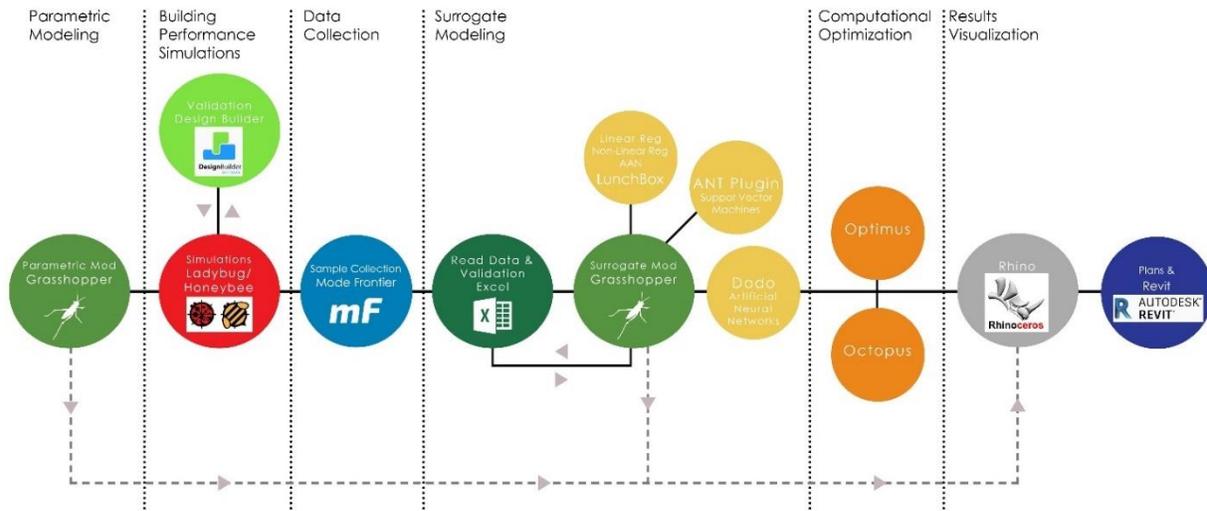


Figure 20 Final Workflow

The parametric model is built in Rhino/Grasshopper with Ladybug/Honeybee (Daysim/Energyplus) as the building physics simulation engine, then ModeFRONTIER (ESTECO, 2020) is utilized as data collection software. Following, this data is saved as an excel sheet and read back into a new Grasshopper environment to be run through various plugins with Machine learning methods. The best performing Surrogate models were trained and chosen.

Next, the original parametric form generation model was simplified by stripping it from its Ladybug/Honeybee simulation components and then coupled with the trained surrogate models. The form generation model was given the option to be turned off, this allowed for the optimization process to run extremely quickly by using only the trained surrogate models. Once an interesting solution was found, the model could be turned back on to immediately review the results, thus making it more useful in the design process.

Finally, the generic Rhino geometry was translated through the Rhino.Inside Beta ( Robert McNeel & Associates, 2020) plugin into Revit Architecture (Autodesk, 2020). This would thus allow the designer to quickly generate plans, sections, elevations, and 3D images to showcase any of the Pareto-optimal solutions.

### 3.2 Overview of the Parametric Model

Many factors play a critical role in improving the energy, daylight, and comfort of a building, such as the HVAC system type or dynamic façade technologies. Nonetheless, the following parameters were chosen to be studied for this research because of their high influence according to the literature review and because of the effects they have on the architectural image of a high-rise.

### 3.2.1 General Parameters

Table A

#### *General Parameters*

Building Parameter	ID	Description	Units	Range Min	Range Max	Data Interval	# of Inputs
x6 Orientation	BO	The angle to the north	°	0	315	45	7
x9 Shape	BS	General shape: Triangle, Parallelogram, Octagon, Ellipse	# of Segments	3, 4, 8, 32		-	4
x4 Plan Length	BL	Length of plan	m	20	80	20	4
x16 Plan Width	BW	Width of plan	m	20	60	20	3
x1 Analysis Level	AL	Building where building level is performed	m	0	30	15	3
x4 Floor to floor Height	FFH	Distance from floor to next floor	m	3	4.5	0.5	4
x2 Core Type	Core	Position of the core to the plan	-	Central Core Lateral Core		-	2
Geographic Location	Loc	Options for 4 different locations and climates	-	Amsterdam Bogota, Dubai		-	3

#### **Orientation**

The orientation of a building and its facades play a crucial role in the 3 performance parameters. Differences in annual electricity use between orientations are driven by the solar gains that affect the cooling or heating or natural daylight. (Efficient Windows, 2011-2020)

In the northern hemisphere, a southern orientation is usually preferred because it is possible to shade from the sun but also capture daylight (through light-shelves for example) to reduce lighting energy load.

In contrast, northern oriented facades receive good ambient and indirect daylight, therefore glare is rarely an issue. The hardest orientations to shade are the eastern and western ones because of the low position of the sun, thus vertical sun-shading can be good passive strategies on these sides.

The model was developed to rotate 315 degrees with steps of 45° to keep the variables to a minimum of 7. Even though a 180° rotation would result in the same option for the central core, this same 180° rotation would result in a different mirrored version for the lateral core, making it a significant option to be reviewed.

## Shape

For the shape parameter, four (4) different shapes were chosen depending on the number of segments of the polygon in plan view: Triangle, Parallelogram, and Octagon had their respective 3, 4, and 8 sides. The ellipse option was simplified to have only 32 sides to keep an acceptable model complexity, so although it is not a perfect ellipse, it is categorized as such for this exercise.

## Length and Width / Usable Area

The width and length of a building are critical variables to consider when seeking to optimize the interior floor area of a high rise, being this a key measurement within this typology, as every squared meter of usable floor area is particularly more expensive when building a Highrise. For this reason, the interior usable floor area is calculated as an additional performance metric.

Likewise, the length and the width of a typical floor plate of a Highrise can drastically influence the depth of the plan, therefore it was expected that these parameters would have a high impact on the daylight penetration into the deeper parts of the floorplate, ultimately affecting UDI.

These parameters are measured in meters and limit the external perimeter of the shape, the smallest polygon is limited to 20 m x 20 m for a triangle with a floor area of only 130 m<sup>2</sup>; lower than this is considered economically inviable because of the high cost vs usable area ratio, making it a good minimum range. The largest external perimeter is 60 m x 80 m for the ellipse with a floor area of up to 4300 m<sup>2</sup>; rarely was a floor area for a high rise found in the literature review making this a good maximum range.

## Analysis Level

The analysis level parameter was defined as the position where the analysis was taking place within the height of the building: the first, middle, or top floors. While the context remained static, the level of analysis could give insight into the effects of different urban contexts. The top floor could be interpreted as the lower level of a building with no context (such as the countryside), the middle level could be interpreted as a building with medium height context, and the lower level could be interpreted as a building with a densely packed high-rise context (such as Manhattan).

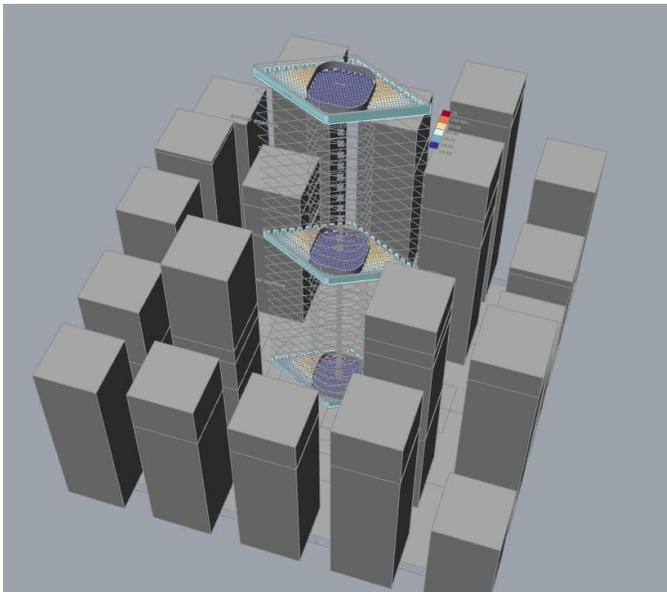


Figure 21 -Parametric grasshopper model with Analysis Levels: N00, Mid-Level and Top Level (source: own)

## Floor to Floor Height

The floor to floor height affects the daylight and energy demand as a higher floor to floor height allow for deeper penetration of natural daylight into the floorplate, thus potentially improving UDI but also creating larger surface area towards the outside, affecting the energy consumption. Hence, floor to floor height variable was considered in the model.

To simplify the model for efficiency, the total number of floors was kept at 30 floors while the context building height was kept between 15 to 20 floors. This, however, did not affect the ability to modify the total height of the building as the Floor to Floor Height (FFH) parameter was kept. According to research, older office buildings from the 1980s have an FFH of at least 3 m while more modern ones offer 4 m and even 4.5 m to allow for MEP ducts flexibility and more daylight, therefore the chosen range was between 3 and 4.5 m

## Core Type

As seen from the research from section 2.1.2, efficient high-rise buildings can be designed with central or lateral cores. The parametric model accounted for two options by allowing for a 1:1 ratio central core versus a 2:1 ratio for the lateral one (inspired by the Leadenhall building lateral core in the literature review). Both core options used the following formula to calculate its size to the floorplate:

$$\text{Core size} = (\text{GFA} * 0.2) + (\text{TotH} * 0.02)$$

*GFA = Gross floor area*

*TotH = Total Height*

An average of 20% was used. Similarly, it seeks to consider that as high-rises increase in height, their core must expand to accommodate additional MEP ducts and lifts required to service its new height. The total height is therefore multiplied by 0.2 to account for this increase. This formula was custom developed for this model as a hypothesis, yet it could be further refined with additional information on high-rise/core ratios, which was not the main objective of this paper.

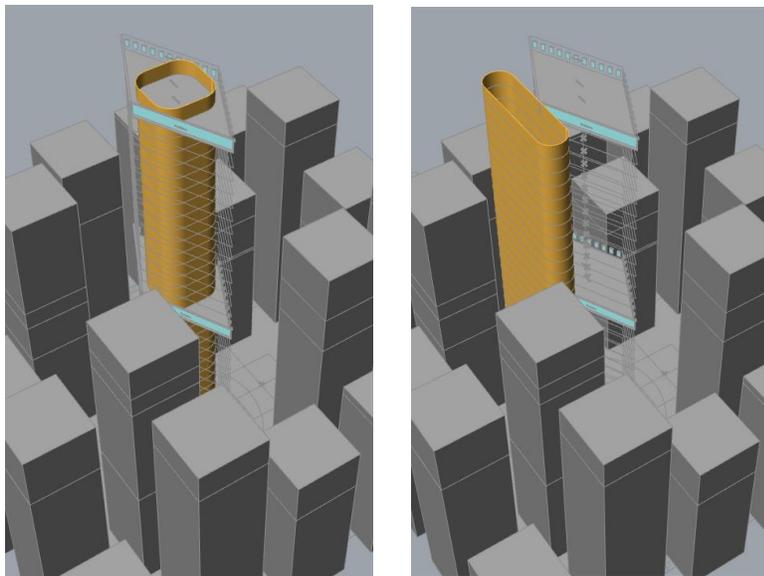


Figure 22 Core Options: Central (Internal) versus Lateral (External) (source: own)

Another essential aspect of the parametric model was assigning the core as a Thermal Mass component. When the central core type was chosen, it was important to realistically model the effects of a solid concrete core within the floorplate, possibly affecting the temperature inside. Its walls were assigned as adiabatic surfaces to restrict the energy/heat exchange from the floorplate to the core. This option could have significant advantages in passive design.

Inversely, when the option for the exterior lateral core was chosen, the script automatically updated the core walls to react as exterior walls. Then, depending on the orientation of the building, the core could theoretically serve as a thermal shield from direct sun.

Although more core options are possible, like 3 cores from the Commerzbank building, for simplicity purposes, only these two options were considered. The goal of this variable was to see the general effects of the core position rather than to do an in-depth analysis of high-rise core positioning. Further studies could further investigate this subject.

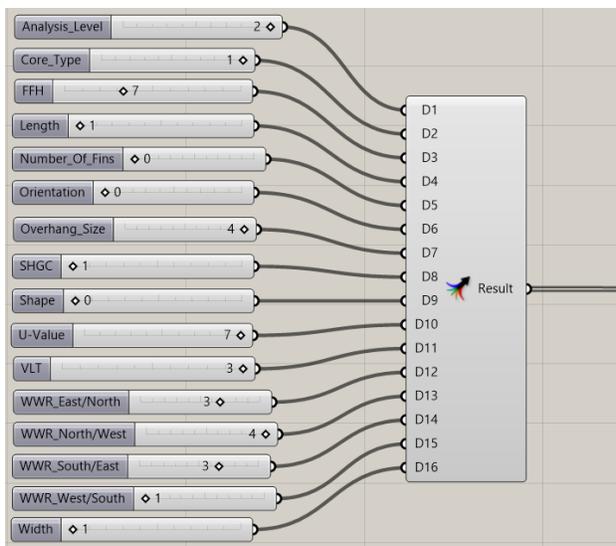


Figure 23 Example of the 16 parameters set up in Grasshopper (source: own)

## Geographic Location

The three chosen locations were: Amsterdam, Bogotá, and Dubai.

500 samples were collected for each location. Because the geographic location of a simulation naturally changes the weather and sun path, this one variable would drastically change the outcomes of each of the variables, further complicating the training phase of the Surrogate model. To keep it simple, initially, three different surrogate models were developed for each location rather than one that contemplated all 4 locations.

Once acceptable surrogate models were developed for each location, the 500 results from each of the locations were simply compiled in a master 1500 sample table that contained the four locations, and the meta-model was trained.

### 3.2.2 Envelope Parameters

Table B

#### Envelope Parameters

Building Parameter	ID	Description	Units	Range Min	Range Max	Data Interval	# of Inputs
x8 Solar Heat Gain Coefficient / g value (0-1)	SHGC	Solar radiation through transparent materials	-	0.2	0.8	0.3	3
x13 WWR_North/West	Nw	Aglass/ Awall	ratio	0.2	0.8	0.2	4
x15 WWR_West/South	Sw	Aglass/ Awall	ratio	0.2	0.8	0.2	4
x14 WWR_South/East	Se	Aglass/ Awall	ratio	0.2	0.8	0.2	4
x12 WWR_East/North	En	Aglass/ Awall	ratio	0.2	0.8	0.2 <td 4	
x11 Visible Light Transmittance	VLT / VT	Glazing	%	30	90	30	3
X7 Overhang_Size	HS	Overhang length	m	0.4	1.6	0.4	4
x5 Number of Fins	#Fin	Number of Vertical Fins per window		1	4	1	4
x10 U value of total window assembly	U Value / $\alpha$	Rate of heat flow	W/m <sup>2</sup> K	0.7	4.9	0.7	7

#### Solar Heat Gain Coefficient

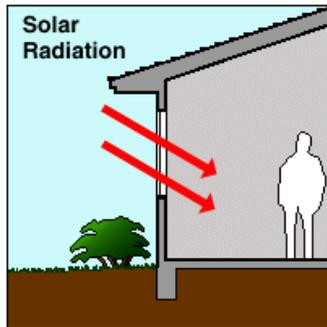


Figure 24 SGHC (Efficient Windows Collaborative, 2000-2018)

This coefficient is a glazing only parameter and measures the amount of radiation that the glazing permits into the building.

A low SHGC is the most important window property in warm climates. To improve energy performance, windows should have an SHGC of 0.25 or less. As a means of example, Triple vacuum Low-E glazing has an SHGC of 0.1 while traditional single clear glazing has an SHGC 0.9

SHGC is very similar to the g-value and is usually presented between values of 0 to 1.

For this model, the SHGC also accounted for the frames in the construction, not just the glass, thus representing the entire window assembly, as recognized by the National Fenestration Rating Council (NFRC) The whole window SHGC is typically 0.8 times lower than glass-only (Efficient Windows, 2019)

## Wall to Window Ratio

Window area or the window-to-wall ratio (WWR) is an important variable affecting energy performance in a building. This parameter is the percentage of opaque vs glazed areas of the facade. It essentially affects energy performance in two ways.

The first is that opaque areas of a façade tend to have much better insulation, therefore less energy is lost. The lower ratios mean smaller glazed areas, therefore more insulation. On the other hand, higher ratios mean more glazing, this usually means that the glazing will be the worst insulator than the opaque zone, therefore more energy is lost. It is worth mentioning that highly insulated glazing does exist, but it comes as a high price, hence not frequently used.

The other way that it affects the energy performance is by allowing more daylight in through the glazing area. This results in less lighting energy consumption. This means that the WWR is a parameter that must be carefully chosen by contemplating the tradeoffs of energy lost but daylight gain through the glazing.

Finally, another important factor is the views because a lower WWR allows for more broad views of the surrounding, thus affecting positively the human experience. Views, however, will not be a subject of this thesis because it depends specifically on the chosen site instead of a generic one. Window to wall ratio is a parameter that is equally linked to all three optimization goals as it affects the insulation level of the façade, the amount of daylight that enters the building, and the user experience.

## Visual Light Transmittance

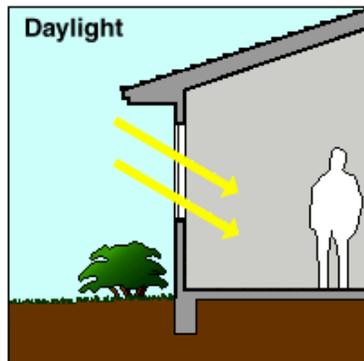


Figure 25 VLT (Efficient Windows Collaborative, 2000-2018)

VLT is an optical property of the glazing that indicates the fraction of visible light that can pass through it. In simple terms, VLT is the measure of the transparency of the glazing. The higher the measure, the more you can see through the window, likewise, the more natural daylight will permeate the interior space. VLT is usually measured within a range of values of 0 and 1, although on other occasions it is presented as a percentage that ranges from 0 to 100% which is essentially the same.

As a means of example, the range varies from above 90% for uncoated water-white clear glass to less than 10% for highly reflective coatings on tinted glass. A typical double-pane IGU has a VLT of around 78%

## Shading Elements

The Center for Window and Cladding predicts that that global warming will lead to an increase of several degrees in summertime temperatures over the next few decades. This means that we must seek to reduce summertime gains. The simplest and most efficient way to achieve this is by using shading devices. These are efficient methods to reduce solar heat gain while maintaining daylight. (CWCT, 2000) There are many types of shading elements such as blinds, brise soleils, awnings, egg crates, louvers, and overhangs. Also, external shading is much more effective than internal shading (Givoni, 1998), thus external was used. This thesis focused only on a simple extrusion of the Overhang size and the number of fins because the goal for this parametric model is to generate a more global understanding of the general effects of these elements. Once this is understood, further refinements or different shading types could be developed during the later stages of the design process.

### Overhang size

Overhang size is simply the length of the overhang above the window. This parameter is dynamically dependent on the WWR, so when the window changed width, did the width of the overhang. This parameter is of special importance on the southern facades (in the Northern Hemisphere) because it can easily protect the façade from unwanted gains from the high positioned sun on the skydome during summer months. Inversely, overhang size is very important on northern facades in the Southern Hemisphere. (Givoni, 1998)

### Number of fins

The “fins” are defined for this exercise as vertical elements attached to the façade. This parameter is also dynamically dependent on the WWR, so when the window changed height, did the length of the fins. This parameter was chosen due to its energy-saving potential on eastern and western facades. (Givoni, 1998). From this literature review, it was suspected that multiple fins would better shield the low sun rather than longer ones. Initially, the fin length was also included as one of the parameters to review but was discarded after reviewing the isolated P-Values and discovering that the fin size was barely playing a significant role in the results of energy, daylight of comfort.

### U Value / thermal resistance

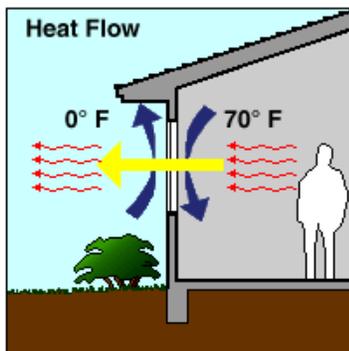


Figure 26 U-Value (Efficient Windows Collaborative, 2000-2018)

U value measures the rate of which heat is transferred through a building's envelope. Unlike SHGC that measures radiation only transfer through glazing, this measure contemplates the three basic heat transfer mechanisms: conduction, convection, and radiation. It can also be used to measure the insulating values of the opaque zones of the façade.

The lower the U-value of a façade, the more slowly heat will be transmitted through it, and so the better it will work as an insulator. So, the lower the U-value, the less energy will be required to keep the building under comfortable conditions.

U values can be measured in two ways, the first is the Overall U factor that considers the façade as a whole component (including the edge of the glass and the frame) and the second is by measuring the U-factor from the center of glazing assuming perpendicular heat flows. For this

As an example, traditional single glass has a U factor of 6.0, High-performance double-pane windows, and its framing can have U-factors of 0.70 or lower, while some of the most expensive triple-pane windows can achieve rare U-factors as low as 0.2. For this exercise, more common and accepted values between 0.7 and 4.9 were selected within the range of parameters to optimize. In contrast, the R-value is the reciprocal of the U-Value ( $R=1/U$ ) The higher the R-value of a material, the higher the insulating value; while the lower the U-factor, the lower the rate of heat flow.

### 3.2.3 Out of scope Parameters

#### **WWR Window Position**

A noteworthy fact is that even though the WWR is important for the energy calculations, the position of the window is also decisive. Two facades with the same WWR could behave differently for the daylight and energy simulations depending on the window spacing or if its located higher or lower regarding the internal space. This exercise does not aim to optimize this parameter; therefore, for both GH and DB models, all window location parameters were kept as similar as possible, with the windows originating from the center of the façade and at a preferred spacing of 2 m.

#### **Light-to-Solar-Gain Ratio**

This is a measure that contemplates both the VLT and the SGHC. Previously, windows depended on tints to reduce the radiative heat gains and therefore reduced also the light transmittance. Now, high-performance tinted glazing or low-E coatings allow reduced solar heat gain with little reduction in visible transmittance. For cold weather, this is overall a positive ratio to increase, as it allows light to enter the building without causing overheating. Nonetheless, the proposed model keeps both VLT and SHGC as separate parameters to specify their correlation with the outputs.

#### **Automated Façade Technologies**

Newer façade systems such as motorized shades, switchable electrochromic/gas chromic coatings, or variable double/triple chamber optical and thermal properties are all state-of-the-art technologies that allow for dynamic shading. Although all these can be additional improvements once a good performing passive design is defined.

#### **Façade Area \* Glass% \* g-value**

Another efficient way to look at the envelope parameters is by combining multiple variables. Combining Façade Area, glass percentage, and g-value makes it possible to reduce these three variables into one. Although this can be very efficient when simulation times are critical, this approach was not taken for the following two reasons:

1. The methodology using surrogate modeling allowed for a much larger number of variables than traditional simulated modeling techniques because once the model was set up correctly, the results of any combination of variables could be assessed almost immediately through the surrogate model. Meanwhile, the computer-intensive simulations were limited to the 500 samples, approximately 150 hrs.
2. The goal of the thesis is to allow for the designer/engineer to use the surrogate model to explore how changing each parameter affects the outcome performance metrics; this means that keeping the parameters separate allowed for each input to be tweaked independently, allowing for reflection and understanding of the potential gains or losses of their choice.

Nonetheless, this method continues to be as a valid analysis tool for quick hand calculation analysis

### 3.2.4 Chosen Parameters and Objectives

A total of 16 independent parameters were chosen, including effects of the Context (1), general building shape & orientation (6) to façade variables (9). These parameters were chosen because of the potential they showed on affecting the energy, daylight, and comfort during the literature review. Likewise, they were chosen due to their ease of application to the parametric model and their high effect on the final “look” of any initial architectural design scheme. While other parameters, such as the HVAC system or the infiltration rate certainly affect the energy demand, they were not included in this optimization because they would not significantly affect the architectural definition of the proposal during the early stages of design.

Table C

	Topic	Building Parameter	ID	Description	Units
y1	Energy	Energy Use Intensity (site)	EUI	AED / GBA - Annual total energy consumption related to the area and use	kW h /m2 yr
y2	Daylight	Useful Daylight Illuminance	UDI	Useful daylight illuminance - the annual occurrence of illuminances across the work plane that are within a range considered “useful” by occupants (9).	%
y3	Comfort	% of Time Comfortable	Com	As defined by EN 15251	%
y4	Area	Area of Floorplate	Area	The useful floor area of one floorplate	m <sup>2</sup>

The four chosen dependent variables where: Energy, Daylight, Comfort, and Area. The choice and reasoning for the validation metrics can be found in section 2.2 of the literature review. The main objective was to lower the energy of a building concerning its floorplate, for this reason, Energy Use intensity (EUI) was chosen. Next, to balance the daylight levels within the office space, improving the Useful Daylight Illuminance (UDI) (100-2000 lux) was the chosen metric. Finally, a minimal comfort range is desired, thus utilizing the adaptive comfort metric as determined by the EN 15251; this single metric for comfort, only considered the times when the temperature would not fall above or below the adaptive comfort range, thus avoiding the times when the user might consider the temperature too hot or too cold. Finally, because of the various shapes and building dimensions, the area of the floorplate was deemed critical to comprehend the results within a useful and comparable framework.

In conclusion, the 16 variables (reordered according to the dataset collected) are x1 Analysis Level, x2 Core Type, x3 Floor to Floor Height (FF), x4 Floor Length, x5 Number of Fins, x6 Orientation, x7 Overhang Size, x8 Solar Height Gain Coefficient (SHGC), x9 Shape, x10 U value of the total window assembly, x11 Visible Light Transmittance (VLT), x12 WWR\_East/North, x13 WWR\_North/West, x14 WWR\_South/East, x15 WWR\_West/South, x16 Plan Width. The output objectives are 4: y1 Energy Use Intensity (EUI), y2 Useful Daylight Illuminance (UDI), y3 Percentage of time Comfortable, and y4 Area of the Floorplate.

The variables were modeled as discrete parameters rather than continuous, the reason for this was to limit the number of options, reducing the complexity significantly. Due to the way the variables were modeled for the form generation in Grasshopper (Rutten, 2015), the discrete variables were translated into integers. Below is a table that serves as a legend for the parameters.

Table D

Parameters Legend									
Analysis Level (m)	30	15	0						
	0	1	2						
Core Type	Central		Lateral						
	0		1						
FFH (m)	3	3.5	4	4.5					
	6	7	8	9					
Length (m)	20	40	60	80					
	1	2	3	4					
Number_Of_Fins	1	2	3	4					
	1	2	3	4					
Orientation (°)	0	45	90	135	180	225	270	315	
	0	1	2	3	4	5	6	7	
Overhang_Size (m)	0.4	0.8	1.2	1.6					
	1	2	3	4					
SHGC	0.2	0.5	0.8						
	1	2	3						
Shape	Triangle	Square	Octagon	Ellipse					
	0	1	2	3					
U_Value (W/m <sup>2</sup> K)	0.7	1.4	2.1	2.8	3.5	4.2	4.9		
	1	2	3	4	5	6	7		
VLT (%)	30	60	90						
	1	2	3						
WWR_East_North (%)	20	40	60	80					
	1	2	3	4					
WWR_North_West (%)	20	40	60	80					
	1	2	3	4					
WWR_South_East (%)	20	40	60	80					
	1	2	3	4					
WWR_West_South (%)	20	40	60	80					
	1	2	3	4					
Width (m)	20	40	60						
	1	2	3						

### 3.3 Validation of Simulation Model

After the parametric model was created in the Grasshopper (Rutten, 2015) environment, initial simulations were run through the Ladybug/Honeybee plugins (Sadeghipour & Mackey, 2017-2020), these results were then compared with a widely accepted energy modeling software known DesignBuilder (Version 5.5.2.007) (DesignBuilder, 2020) to ensure the consistency and validity in the results. One random control option of the many possible iterations was chosen to be represented in both software.

Two methods of modeling validation were tested, the first was to use a ladybug node to export the gbXML file from grasshopper. This proved very efficient to transfer the general geometry. It correctly transferred the space dimensions, walls, and shading surfaces. Likewise, this also served as a valuable method to identify the errors of space definition or geometry that would have been ignored in the grasshopper environment.

The second method was to create a model from an exported 2D .dwg file from Grasshopper (GH)/Rhino. This meant that all the objects, walls, and openings were re-modeled in Design-Build (DB). Although this proved more time consuming, it ensured that all the parameters from both models matched as closely as possible. The results of both independent (GH vs DB) models were then compared.

The examination and definition of each of all these steps was a critical part of validating the two independent energy models (Grasshopper vs DesignBuilder) to ensure that the results made sense and were aligned to a hypothetical real case scenario.

At first instance, only Level 00 was validated, once this provided similar results, three levels at higher altitudes were also reviewed.

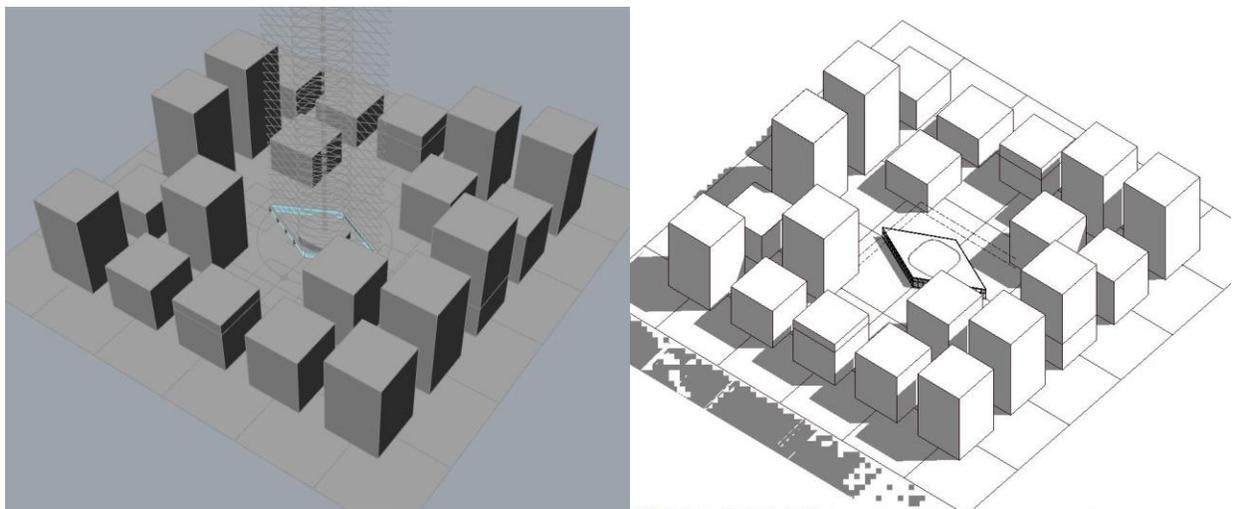


Figure 27 - Grasshopper model Level 00

DesignBuilder model Level 00 (source: own)

The first runs the energy results from the Ladybug component were off the charts: The energy consumption was extremely high, not in line with the results from the DesignBuilder. As this software has been designed from the ground up for simplified energy modeling with a controlled and straightforward graphical user interface, the most trusted results were those provided by DesignBuilder.

The grasshopper model was carefully compared with the Design builder model to review the most significant variables to achieve similar results. Next is the complete list of variables that were considered:

### Chosen Parameters for cross model validation

	Grasshopper Model		Design Builder		Units
	Floor 01	Core 01	Floor 01	Core 01	
Shape	Parallelogram	Parallelogram	Parallelogram	Parallelogram	-
Length	40	40	40	40	m
Width	20	20	20	20	m
Floor to Floor Height	3.5	3.5	3.5	3.5	m
No. of Floors	30	30	30	30	-
N0 Level	0	0	0	0	m
N Mid Level	52.5	52.5	52.5	52.5	m
N Top Level	105	105	105	105	m
Context Type	Mid rise	Mid rise	Mid rise	Mid rise	-
<b>Exterior Wall</b>	ASHRAE 90.1-2010 Ext Wall Metal ClimateZone 7-8	none	Semi-exposed wall Typical reference - Lightweight	none	-
R-Value	2.939	none	2.881	none	m2-K/w
U-Value	0.34	none	0.347	none	w/m2-K
<b>Internal wall</b>	none	M11 100 mm Lightweight Concrete	none	Project Internal Mass Concrete -100 mm	-
U-Value	2.6	2.6	2.6	2.6	w/m2-K
<b>Glazing:</b>	Custom		Project external Glazing		-
U-Value	2.08	none	2.08	none	w/m2-K
SHGC	0.691	none	0.691	none	-
VLT	0.74	none	0.744	none	-
<b>Openings</b>					
Window Height	1.5	none	1.5	none	m
Window Spacing	2	none	2	none	m
Sill Height	0.8	none	0.8	none	m
WWR North	20	none	20	none	%
WWR South	80	none	80	none	%
<b>Roof</b>	ASHRAE 90.1-2010 Ext Roof Metal Climate Zone 7-8	ASHRAE 90.1-2010 Ext Roof Metal Climate Zone 7-8	Flat roof - Typical Ref - Lightweight	Flat roof - Typical Ref - Lightweight	-
R-Value	3.45	3.45	3.5	3.5	m2-K/w
U-Value	0.29	0.29	0.286		w/m2-K
<b>HVAC</b>	Ideal Air Loads	Ideal Air Loads	Air to Water Heat Pump (ASHP), Convectors, Nat Vent	Air to Water Heat Pump (ASHP), Convectors, Nat Vent	-
Mechanical Vent					
Heating CoP	Ideal Loads	Ideal Loads	1	1	-
Cooling CoP	Ideal Loads	Ideal Loads	1	1	-

<b>Core Type</b>		Internal	-	Internal	-	-
<b>Activity</b>		Open Office	Office: Corridor	Generic Office Area	Circulation Area	-
Occupancy density		0.0565	0.1173	0.0565	0.1173	pp/m2
Equipment Load		11.77	1.85	11.77	1.85	W/m2
Heating Set Point		21	20	21	20	°C
Heating set Back		12	12	28	12	°C
Cooling Set Point		25	23	25	26	°C
Cooling Set back		28	28	28	28	°C
FractionGlz Operable / Free Aperture		25	-	25	-	%
Min Temp Natural Ventilation Setpoint		24	-	-	-	°C
Min Temp Outside Natural Ventilation		14	-	-	-	°C
Airtightness		0.2	-	0.2	-	ac/h
Heating Sizing Factor		1.25	1.25	1.25	1.25	
Cooling Sizing Factor		1.15	1.15	1.15	1.15	
Timestep per hour		4	4	4	4	
<b>Lighting controls</b>						
Working plane height:		0.75 m	0.75 m	0.75 m	0.75 m	
Normalised power density		5	5	5	5	W/m2-100lux
Control Type		Autodimming with switch off occupancy sensor		3 Stepped		
Lighting Set point		300	300	100	100	

Table 8 Chosen Parameters for Cross Model Validation (source: own)

The calibration process was a back and forth process that required many runs of building energy simulations. On each run, the results from each of the independent models were reviewed

### 3.3.1 Fine-tuned Parameters

#### Mechanical Ventilation (COP)

This coefficient is used to calculate the fuel required to meet the heating or cooling demand on an HVAC (Heating, Ventilation, and Air Conditioning) system. This value represents the total seasonal efficiency of the entire system and includes the outcome of the complete HVAC equipment (DesignBuilder, 2020)

Since this thesis does not aim to investigate the efficiency of the HVAC variable, a CoP of 1.0 was chosen in the DesignBuilder software which in turn reflects most closely to the Ideal Air Loads parameter chosen in the Grasshopper model. Heated and Cooled spaces were set to a CoP of 1 as per building physics expert recommendation. Natural Ventilation was set to similar air exchanges per hr. of 5.0.

A discrepancy arose when Mechanical Ventilation had not been included in the DB model, once corrected and coordinated on both models with an equal minimal fresh air per person (0.01 m<sup>2</sup>/s) the results were much more equivalent.

## Natural Ventilation

Next brief explanations of the setpoints defined for natural ventilation the model:

- **Minimal indoor temperature:** the *temperature at which people start opening windows, to cool off space. Usually between 22-24°C. Set at 24°C in the model*
- **Minimal outdoor temperature:** the *temperature at which it is much too cold outside for natural ventilation. Usually 12-15°C. Set at 14 °C in the model.*

These natural ventilation set points played a significant role in energy loads, therefore getting these parameters as equivalent as possible was a crucial part of the cross-model validation.

The height of the window is also important for airflow because of the temperature difference or buoyancy. Warm air tends to rise; therefore, the warm air exits from the top of the windows while the cool air enters through the bottom. There is more flow the bigger the temperature difference from indoors and outdoors but also the height affects how much warm air escapes through the top. (Mackey & Sadeghipour, 2019). For this reason, the position and spacing of the windows were also modeled as similarly as possible in both models.

## Temperature Control

Similarly to natural ventilation, temperature control was critical in achieving similar results in both models. Heating and cooling setpoints were carefully defined and matched so that both models would be equivalent and thus achieve similar results.

## Usage Density

For this exercise, the Building model will be simplified by defining one zone /activity for the entire floor area and occupancy density used was 0.11 people per m<sup>2</sup>, an acceptable density for office space.

## Artificial Lighting

The Grasshopper model was developed so that the daylight simulation would automatically update the lighting schedule. This meant that the lights would automatically be turned off when enough natural light was available. This automatic lighting control significantly reduced the lighting loads. This also meant that automatic lighting would reduce the lighting loads in the building and ultimately the final output to optimize: EUI.

This image shows how the updated lighting schedule significantly reduces the time the lights are turned on during highly lite summer months.

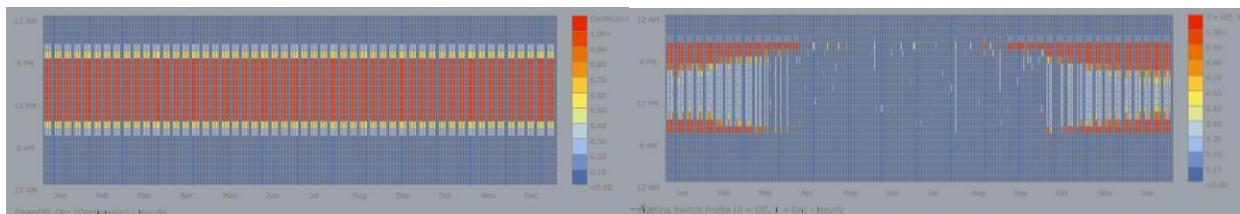


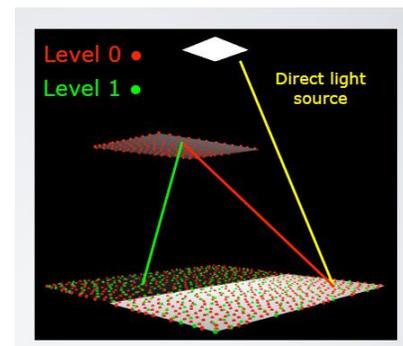
Figure 28 Updated Lighting Annual Schedule: No light control / light control (source: own)

## Daylight Simulation

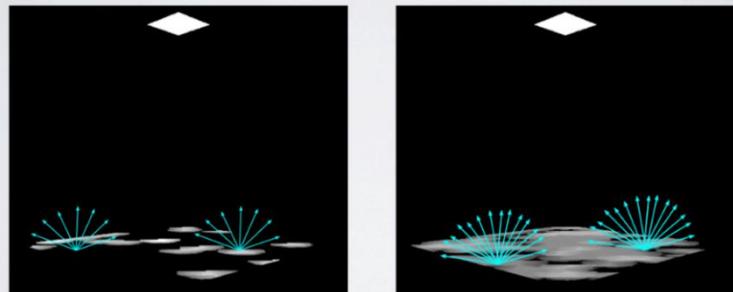
The backward raytracing engine used by Ladybug is Radiance, it uses a stochastic sampling for calculation, and as such a 5-10% difference is expected between runs. For testing daylight, the test surfaces on each model were placed at 75 cm from the floor. Like in all computational simulations, a balance of accuracy vs the simulation time must be found; thus, various test runs were made to find a balance between quality results and acceptable simulation times. The most important part to determine these definitions was to review how they would affect the energy analysis therefore a complete day of runs was employed to check their influence on the final energy output. The results show that the following recipe was good enough and did not have a significant impact on the energy outcome.

**Ambient bounces (AB):** Amount of bounces of the light rays within the space. When shading is present at least 2 bounces must be used to ensure the correct effect of the bouncing light rays.

These rays overestimating the effect of the shading because it does not bounce enough times thus a balance must be found between the number of bounces and speed of the simulation. (Mardaljevic, 2011)

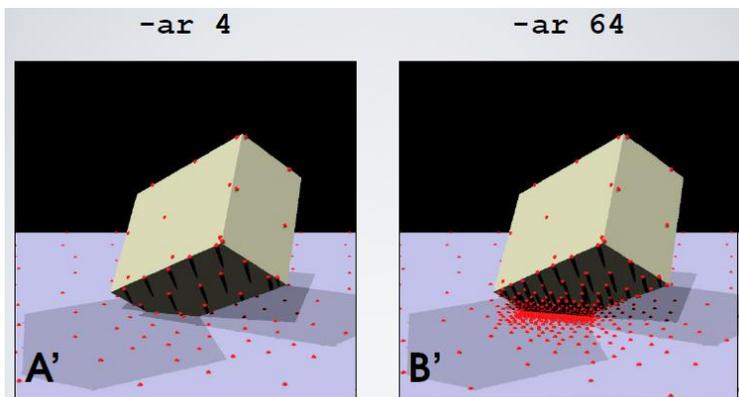


**Ambient Accuracy (AA):** This is how much accuracy is expected in the model. In terms of percentage of error, therefore  $0.1 = 10\%$  error. (Mardaljevic, 2011)



**Ambient Division (AD):** Number of rays that are divided when the light ray hits a surface. The more rays the more accurate the results. More rays take longer to calculate. (Mardaljevic, 2011)

**Ambient resolution (AR):** Number of points used to understand the scene. For this reason, if the scene has many small details, the ambient resolution must be higher. (Mardaljevic, 2011)



## RAD Parameters

This study also looked at existing literature to compare and define the radiance parameters for this study:

Table E

Radiance ambient parameters

	Ambient bounces	Ambient accuracy	Ambien resolution	Ambien divisions	Ambien super-samples
<i>This study</i>	2	0.1	300	1000	20
Ayca Kirimtata et al (2019)	2	0.15	300	1000	20
Wagdy et al. (2017)	6	0.1	300	1000	20
Sherif et al. (2016)	6	0.1	300	1000	20
Mohsenin and Hu (2015)	6	0.1	300	1000	256

Table 9 source: own & Ayca Kirimtata, 2019

## Differences in Activity Schedules:

Initially similar “Open Office Space” activity and lighting schedules were chosen on both models, thus giving similar uses and occupied times. Upon more detailed inspection, however, each software had slightly different expected occupations, and this discrepancy built up throughout the year. To eliminate this variable, a custom schedule for Ladybug component was created so that it reflected the exact occupancy as that used in DesignBuilder.

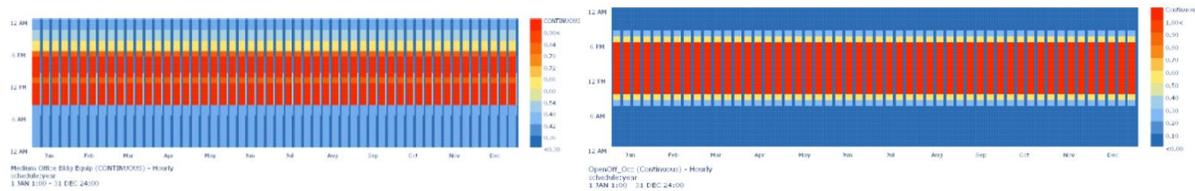


Figure 29 Updating Activity Schedule for Ladybug to match Schedule of DesignBuilder:

## Differences in Core geometry:

Likewise, the models have a slightly different core geometry but the area they occupy is equivalent. This proved to not be a significant difference as the final energy calculations were similar without updating this aspect, therefore the simplified version of the design Builder model was kept.

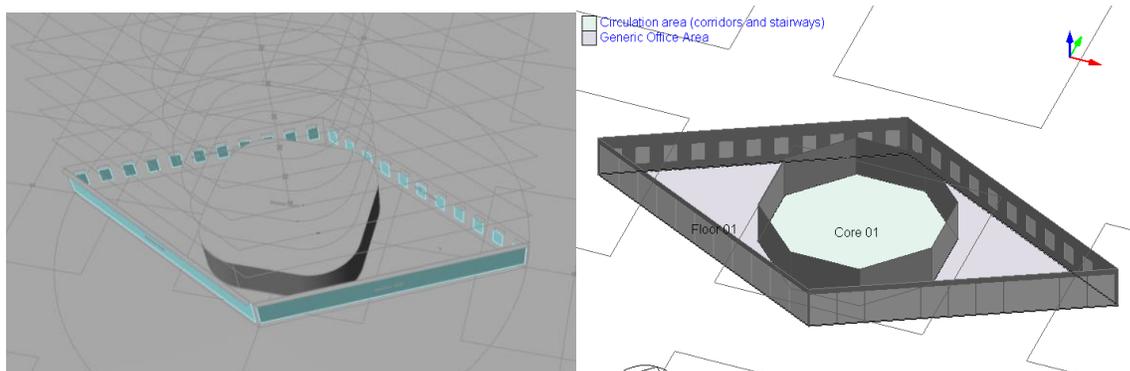


Figure 30 - Grasshopper model Level 00

DesignBuilder model Level 00

### 3.3.2 Results from Model Validation

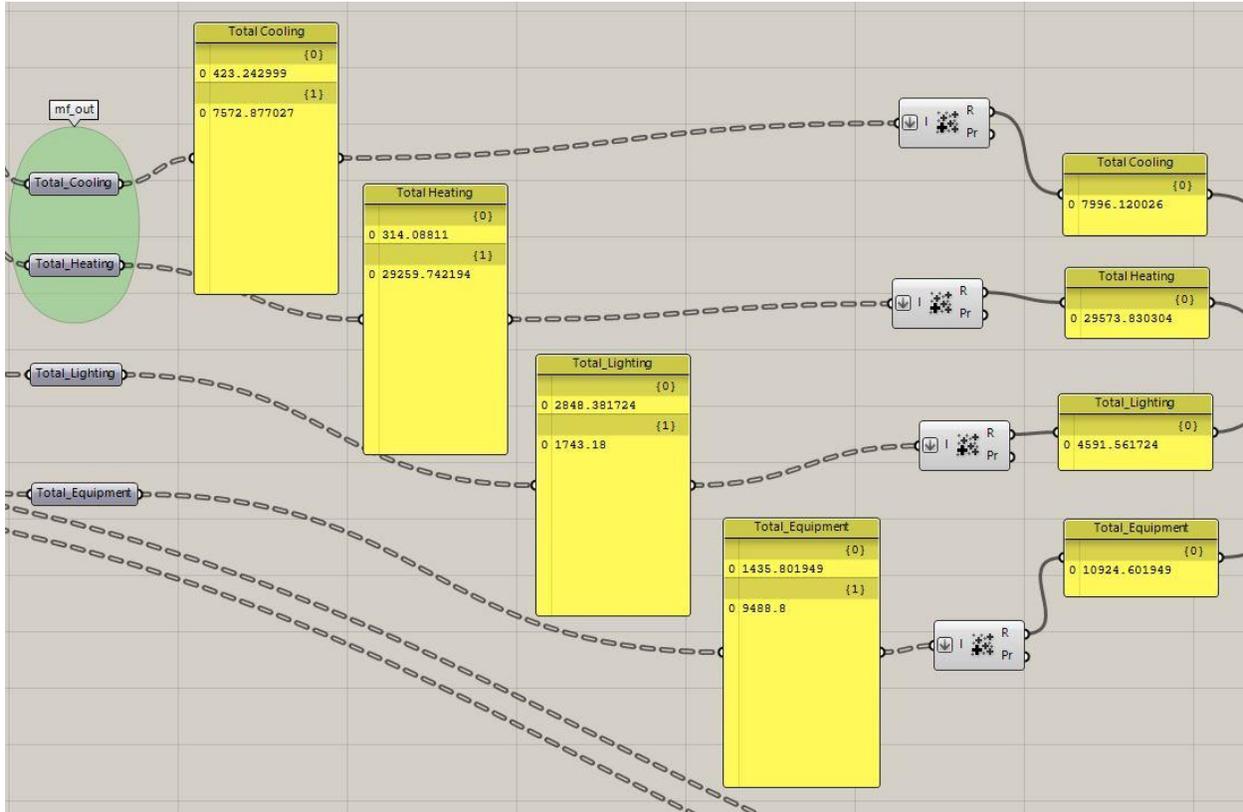


Figure 31 Energy Calculations for level 1 in Grasshopper Honeybee/Ladybug (source: own)

	Electricity [kWh]	Natural Gas [kWh]	Additional Fuel [kWh]	District Cooling [kWh]	District Heating [kWh]	Water [m3]
Heating	0.00	0.00	0.00	0.00	27524.57	0.00
Cooling	0.00	0.00	0.00	8364.10	0.00	0.00
Interior Lighting	4708.30	0.00	0.00	0.00	0.00	0.00
Exterior Lighting	0.00	0.00	0.00	0.00	0.00	0.00
Interior Equipment	11494.11	0.00	0.00	0.00	0.00	0.00
Exterior Equipment	0.00	0.00	0.00	0.00	0.00	0.00
Fans	0.00	0.00	0.00	0.00	0.00	0.00
Pumps	0.00	0.00	0.00	0.00	0.00	0.00
Heat Rejection	0.00	0.00	0.00	0.00	0.00	0.00
Humidification	0.00	0.00	0.00	0.00	0.00	0.00
Heat Recovery	0.00	0.00	0.00	0.00	0.00	0.00
Water Systems	0.00	0.00	0.00	0.00	0.00	0.00
Refrigeration	0.00	0.00	0.00	0.00	0.00	0.00
Generators	0.00	0.00	0.00	0.00	0.00	0.00
Total End Uses	16202.41	0.00	0.00	8364.10	27524.57	0.00

Figure 32 Energy Calculations for level 1 in DesignBuilder (source: own)

Table F

Results from Model Validation

Results:	Grasshopper Model (kWh)		DesignBuilder (kWh)		Difference %		
	N00	All 3 Levels	N00	All 3 Levels	N00	All 3 Levels	3 Levels
Cooling	7996	24787.6	8364	25696	-4.6	-3.7	
Heating	29573	91676.3	27542	96698	6.9	-5.5	
Lighting	4591	13773	4708	14357	-2.5	-4.2	
Equipment	10924	32772	11494	34482	-5.2	-5.2	
<b>Totals</b>	<b>53084</b>	<b>163008.9</b>	<b>52108</b>	<b>171233</b>	<b>1.8</b>	<b>-5.0</b>	

*Table 10 Summary of the GH/DB Validation Results (source: own)*

The above table shows how the results between the Ladybug/Honeybee and DesignBuilder models were finally calibrated within a +/- 5% margin for all three levels, which was considered an acceptable margin of error for this exercise. This concludes that once fine-tuned, the parametric model could be trusted because the results were sufficiently close to those obtained from the independent DesignBuilder software.

### 3.4 Collecting samples

As identified in the literature review, the two approaches for collecting samples for a surrogate model are static (sequential) vs adaptive (iterative). (Westermann & Evins, 2019) Although adaptive sampling strategies show promising results and have successfully shown to require fewer simulation runs to achieve better surrogate accuracy, this technique was not utilized due to time constraints related to the complexity of this subject. Further research could investigate the potential of this topic.

Eventually, Uniform Latin Hypercube (ULH) sampling method was chosen due to its frequent use in existing literature (17 out of the 57 articles from the review performed by Westermann & Evins), this strategy also has advantages “given its simple concept, ease of implementation, and efficient stratification” (Chen & Yang, 2017). However, because the numbers of samples required are problem-dependent (Brea, et al., 2020), it was difficult to accurately define the number of samples needed thus the simulation time was another important variable to consider

#### Simulation times

Each simulation took an average of 16 minutes, therefore only hundreds rather than thousands of samples where possible. Without considering unintended computer malfunctions, approximately 3.4 samples per hour could be collected using a Desktop computer with the following specifications: Intel® Core™ i7-5820K CPU @ 3.3 Mhz, 6 Cores, 12 Logical Processors, x64bit Windows 10 Pro, RAM: 16 GB.

Ultimately, 500 samples were chosen to be collected for each location. 500 samples took 147 hrs. per location (approximately 6 days) with the aforementioned PC. Although, it is worth mentioning that this time could be significantly reduced if the computational load for the daylight simulation was divided between multiple CPUs. 500 ULH samples were collected for each location: Amsterdam, Bogotá, and Dubai for a total of 1500 samples.

### 3.4.1 Choosing the Sampling Strategy

Before the 1500 samples using ULH sampling strategy, however, the choice for the right sampling strategy was unclear, for this reason, a Quasi-Random sampling strategy Sobol was used to collecting an initial batch of 500 samples. The reason for this is that literature showed promising outcomes compared to ULH: “Sobol’s sampling was used for Monte Carlo simulation and provided more precise and robust output distributions than Latin-hypercube and random sampling” (Westermann & Evins, 2019)

Hence, 500 samples were collected using the Sobol sequence and compared to the 500 samples using ULH; both for the same Amsterdam Location. This served as a tool of comparison to review the difference, if any, between a near-random sampling strategy (ULH) versus a quasi-random sampling method (Sobol) (Westermann & Evins, 2019)

ModeFRONTIER software (ESTECO, 2020) was used as the tool to generate the DoE and collect the samples. With this dataset, the results from each sampling method were run through Linear Regression data analysis to compare the usefulness of the information that could be extracted from each.

Table G

Sobol vs ULH sampling method

	Sobol			ULH		
	EUI	Comfort	UDI	EUI	Comfort	UDI
R <sup>2</sup>	0.83434507	0.99935141	0.36186573	0.84796182	0.99945149	0.35780101
Standard Error	11.0616862	1.04635648	11.0250236	9.74086331	0.96024303	10.7197914
P-Value	2.133E-153	0	2.4923E-32	9.339E-186	0	3.0092E-37
Significant Variables	12	<b>3</b>	<b>7</b>	12	<b>6</b>	<b>11</b>

Table 11 Sobol Vs ULH Sampling methods (source: own)

The significant variables count the number of variables that reject the Null hypothesis by complying with the  $p < 0.05$  alpha threshold. ULH showed more significant variables than Sobol, meaning that the sampling method might be finding more statistically significant relationships between variables and outputs, making ULH slightly more interesting to utilize. As for the rest of the results, both ULH and SOBOL have very similar R2, Standard error, and P-Values, making the difference in these results insignificant.

Ultimately the Uniform Latin hypercube technique was chosen to collect samples for Amsterdam, Bogota, and Dubai as its relatively uniform distribution showed more statistical relationships (p-value) than Sobol.

### 3.4.2 Verifying Statistical Significance

Once a total of 1500 samples using ULH were collected amongst the three locations: Amsterdam, Bogota, and Dubai, it was considered important to also review the statistical significance of the results. For this reason, an Analysis of variance (ANOVA) analysis was carried out with Data Analysis Multilinear Regression tool using Excel. The Significance F proved the overall statistical significance of the variables over each of the outcome while the p-value for each variable showed which outcome they were affecting.

Next a summary table for each of the 500 samples for each of the locations. See Appendix for complete tables.

Multilinear Regression  
Summary H

Samples	Amsterdam				Bogota				Dubai			
	500	500	500	500	500	500	500	500	500	500	500	500
	EUI	Area	Comfort	UDI	EUI	Area	Comfort	UDI	EUI	Area	Comfort	UDI
R Square	0.8329	0.8509	0.8129	0.3722	0.7315	0.8508	0.8218	0.3937	0.7315	0.8508	0.8218	0.3937
Adjusted R Square	0.8273	0.8459	0.8067	0.3514	0.7226	0.8458	0.8159	0.3736	0.7226	0.8458	0.8159	0.3736
Standard Error	12.404	377.93	5.25	10.844	6.5352	378.03	7.2923	13.374	6.5352	378.03	7.2923	13.374
Significance F	2E-175	2E-187	9E-164	2E-39	3E-126	3E-187	8E-169	7E-43	3E-126	3E-187	8E-169	7E-43

Firstly, this table shows that for all outputs and all locations, the significance F is below the 0.05 Alpha threshold, meaning that indeed the inputs are affecting the outputs, thus the collected data is behaving as expected.

This summary also shows that for the three locations, EUI and Comfort could be adequately predicted by using even the simplest of machine learning methods: Multilinear regression. This means that the data for EUI and Comfort behaves more linearly and the relationships are less complex. Conversely, for the area output, although the R squared value is somewhat acceptable (84%) with this method, the standard error is very high, meaning that its predictions could fall very far from the real data. Although the floor area depends only on 3 of the 16 variables: length, width, and shape; the shape variable changes drastically the area outputs whether it is a triangle, rectangle, octagon, or ellipse; thus, possibly resulting in a higher standard error. Independently, the area was simply used as a proof-of-concept to validate that only these three variables were statistically significant. The area in the final model was simply calculated geometrically rather than by predicting it through a surrogate model.

Moreover, the UDI output has a more complex relationship with the variables therefore the R squared value is much lower when using simple Multilinear regression. For this reason, to have better surrogate models, other machine learning methods were explored: Non-linear Regression and ANN. Using more complex machine learning methods also meant that the EUI and Comfort could also be improved to better fit the data.

Table I  
MLRegression Individual P-Values

	Amsterdam - P-Values				Bogota - P-Values				Dubai P-Values			
	EUI	Area	Comfort	UDI	EUI	Area	Comfrt	UDI	EUI	Area	Comfrt	UDI
Analysis Level	2E-13	0.1872	2E-49	0.0736	2E-06	0.1951	2E-89	0.1085	1E-97	0.2017	7E-109	0.1041
Core_Type	5E-57	0.422	0.2164	4E-38	2E-30	0.4055	0.0567	4E-42	0.3377	0.3845	5E-10	7E-42
FFH	2E-43	0.5856	1E-05	0.0109	2E-09	0.5645	0.1502	0.9103	3E-30	0.5888	2E-10	0.9212
Length	8E-76	9E-132	4E-18	2E-08	8E-51	1E-131	0.6818	0.1907	2E-49	131	0.001	0.3912
Number Of_Fins	0.1779	0.905	0.0073	0.3157	0.2657	0.8941	3E-07	0.6192	0.0403	0.9488	2E-05	0.6614
Orientation	0.1596	0.256	0.803	0.3333	0.0503	0.2663	0.6743	0.2783	0.9711	0.3065	0.8088	0.2329
Overhang Size	0.0726	0.6701	9E-06	0.3183	0.0042	0.6896	2E-13	0.0209	0.0082	0.7223	2E-10	0.0127
SHGC	0.0393	0.5282	3E-50	0.9383	2E-10	0.5488	5E-97	0.7809	0.4937	0.5583	1E-73	0.8735
Shape	1E-42	2E-87	0.0014	0.3272	6E-27	4E-87	6E-09	4E-06	3E-38	3E-87	8E-43	4E-06
U_Value	2E-69	0.433	8E-135	0.0851	5E-62	0.4469	2E-97	0.1013	0.0127	0.4687	7E-21	0.1442
VLT	0.0024	0.6506	0.1391	0.0299	0.0137	0.6572	0.0331	4E-10	0.2076	0.6426	0.0384	4E-10
WWR East_North	1E-08	0.9275	0.0455	0.0015	0.0108	0.9608	0.0002	0.1071	1E-12	0.921	1E-06	0.3347
WWR North_West	7E-11	0.8315	3E-07	0.3713	2E-05	0.8223	0.457	0.2038	1E-11	0.7862	0.0145	0.1537
WWR South_East	1E-09	0.0618	9E-06	0.3744	0.0016	0.0583	0.2716	0.0194	8E-15	0.0575	0.0225	0.0101
WWR West_South	8E-13	0.8455	8E-08	0.1799	0.0042	0.8212	0.2138	0.4956	3E-21	0.8491	0.0038	0.3013
Width	2E-60	1E-114	9E-10	8E-05	8E-38	2E-114	0.0185	0.1402	1E-49	6E-114	1E-23	0.463

Table 12 Individual P-Values per Location (source: own)

This table shows the P-values for each variable in each location. Only the P-values lower than the 5% Alpha threshold are identified in green. This information is especially useful to identify which variable is statistically significant for each output, meaning that it affects it. Variables that are not colored in green mean that they do not affect the output for that location. It is noteworthy to see that even though most variables have similar statistical significance, in some locations, more variables affect the outputs more than in others. For example, all Wall to Window Ratios orientations, are more important for Comfort in Amsterdam than in the other locations presumably because of more extreme sun positions throughout the year due to it being farther from the equator. Also noteworthy is that the core type affects comfort, presumably because the core can serve as a shield from the hot summers of Dubai. The analysis level (effects of the context) does indeed affect the EUI and Comfort, meaning that this is an important variable to consider when designing High Rises. Finally, as expected, the floor area depends equally on the length, width, and shape similarly on any location, proving that the P-values are showing accurate results.

### 3.5 Training the Surrogate model

To properly train a surrogate model, it was critical to separate the collected samples into 2 main datasets: the training data and the test data to cross-validate the results.

#### Cross-Validation of the Data



This diagram shows how the cross-validation method was set up to train and test the surrogate model. According to research, a division of 80% for the training and 20% for the testing data is adequate and accepted. (Westermann & Evins, 2019) A custom-cross validation Grasshopper script was developed with a split component that separated the data randomly by changing a seed slider. This allowed the various samples to be mixed randomly from the original 500-sample batch; resulting in a robust cross-validation method.

Next, these two datasets are separated further into two subgroups: the inputs (independent variables) and the outputs (dependent variables). To achieve this, the custom script also separated the 16 input columns from the 4 output columns. In the end, by subdividing the data into 4 parts for cross-validation, the predicted results from the trained surrogate model can be compared with an independent 20% of actual measured results, ensuring the trustworthiness of the predicted results of the trained surrogate model.

As explained, the data is thus separated into four groups:

80% X (training)

80% Y (training)

20% Cross-Validation Y (testing input)

20% Cross-Validation Y (for testing against predicted  $\hat{Y}$  values)

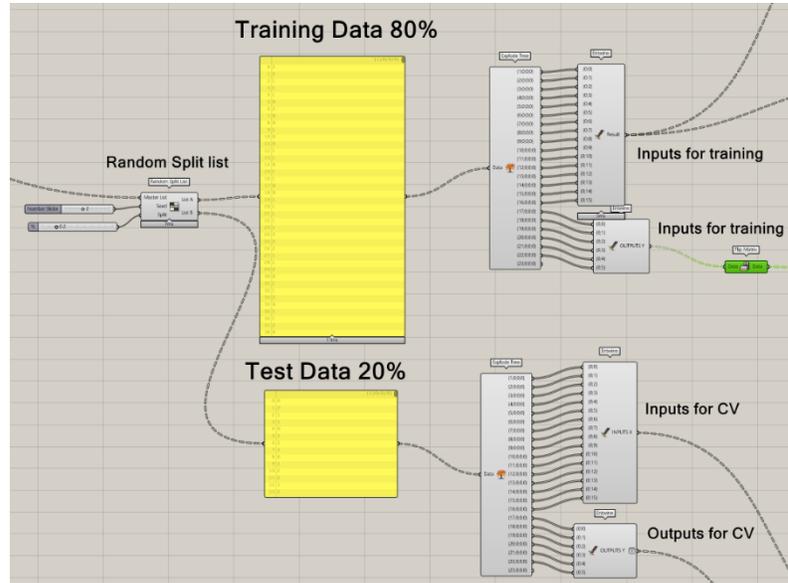


Figure 33 Custom Grasshopper Script for data Cross-Validation (source: own)

Once the surrogate model was trained, the data could be randomly modified using the seed slider to test its validity on any other random 80/20 samples. New datasets meant that the R and Standard Errors changed slightly from seed to seed but the overall validation metrics remained similar.

### 3.6 Evaluation of the Surrogate Model

To achieve good results from a surrogate model, the validation metrics explained in section 2.4.4 must be validated and finetuned. The goal for improving the predictive power of the model was: Increase R,  $R^2$ , Adjusted  $R^2$  while simultaneously reduce the MAE, MSE, or Standard Error. Likewise, the P-value results should be lower than the accepted 0.05 alpha threshold for the data to be deemed statistically significant. The formulas for  $R^2$ , Adjusted  $R^2$ , and MSE were converted into a custom grasshopper script that allowed immediate calculation of these metrics, independent of the machine learning method used.

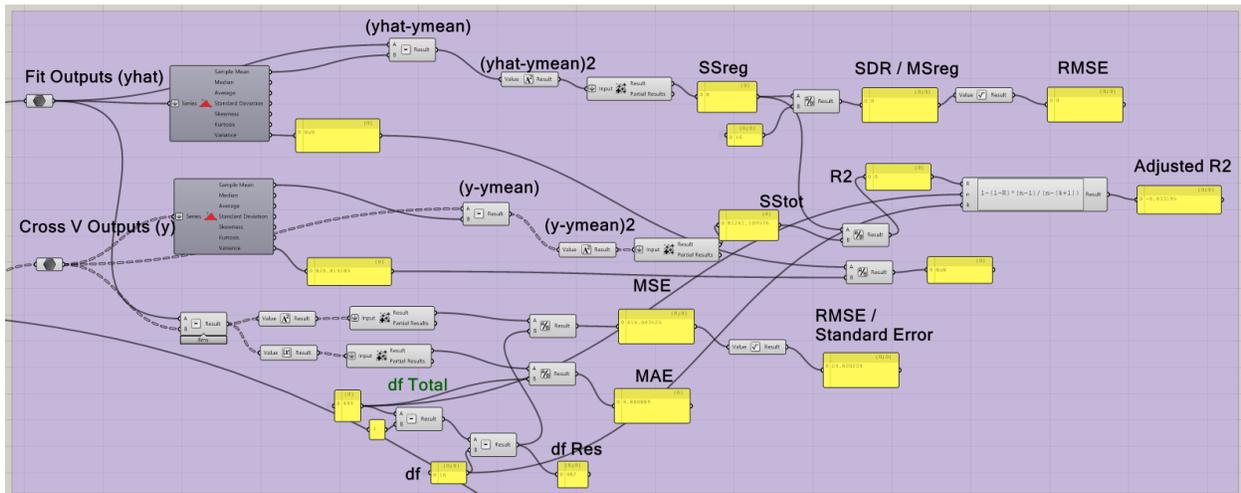


Figure 34 Custom Grasshopper Script for Calculation of MSE/RMSE/MAE (source: own)

This script, however, could be prone to error, therefore an initial test was carried out to compare the Validation outputs from the custom script with those calculated using an external Data Analysis method. In this case, the Excel Data Analysis tool was employed to compare the results using the most basic of the machine learning methods: Linear Regression. The following image shows that the linear regression calculations matched exactly (0.000001) with those from the script. See the image below.

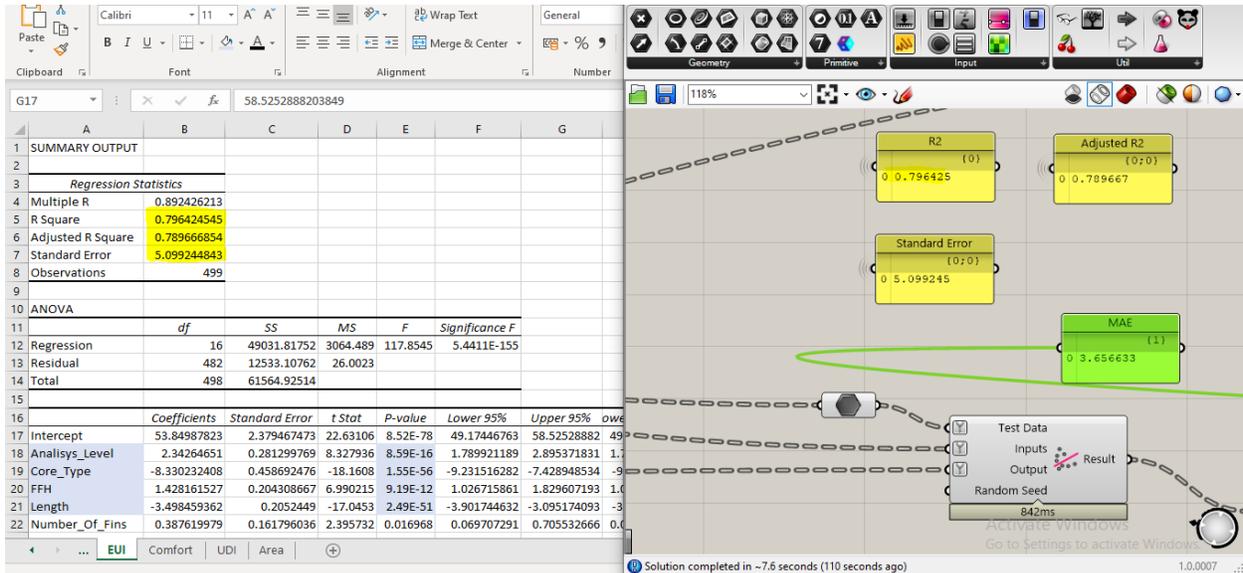


Table 13 Cross-validation: Excel Data Regression vs Lunchbox Linear Regression Component (Source: own)

As a conclusion, this meant that the custom script developed could be trusted. Once it could validate Multilinear regression, it was deduced that further calculations of R2, MSE, Standard Error would be accurate for any other Machine Learning method because any of the predicted  $\hat{Y}$  outputs could be compared to the originally known Y values. This script would permit us to quickly identify which surrogate model method and what parameters should be tuned to achieve better predictive results. The P values however were not calculated in the custom script because the statistical significance of the variables was proven already in section 3.4.2.

### 3.5.1 Choosing the Surrogate model

All in all, 5 Machine Learning methods were explored in the search of the ideal surrogate model.

Firstly, Support Vector Machines was explored with a plugin named ANT (Rahman, 2017) using the Scikit Learn Python module. Although promising at first, this plugin showed little information for validating the results (score logs); a lack of community support for solving errors and the limited method to read the data proved to be a cumbersome process. Multiple tests were run to read the results correctly, yet ultimately this method was abandoned.

The Lunchbox plugin (Miller, 2018), using the Accord.NET framework was used for the MLR and NLR machine learning methods. Rather than using the same ANN component from the Lunchbox plugin, the plugin Dodo (Greco, 2020) was utilized for ANN because it proved to be much quicker to train and to generate results. Ultimately, three machine learning types were chosen. This research tested results from Multilinear Regression (MLR), Non-Linear Regression (NLR), and Artificial Neural Networks (ANN).

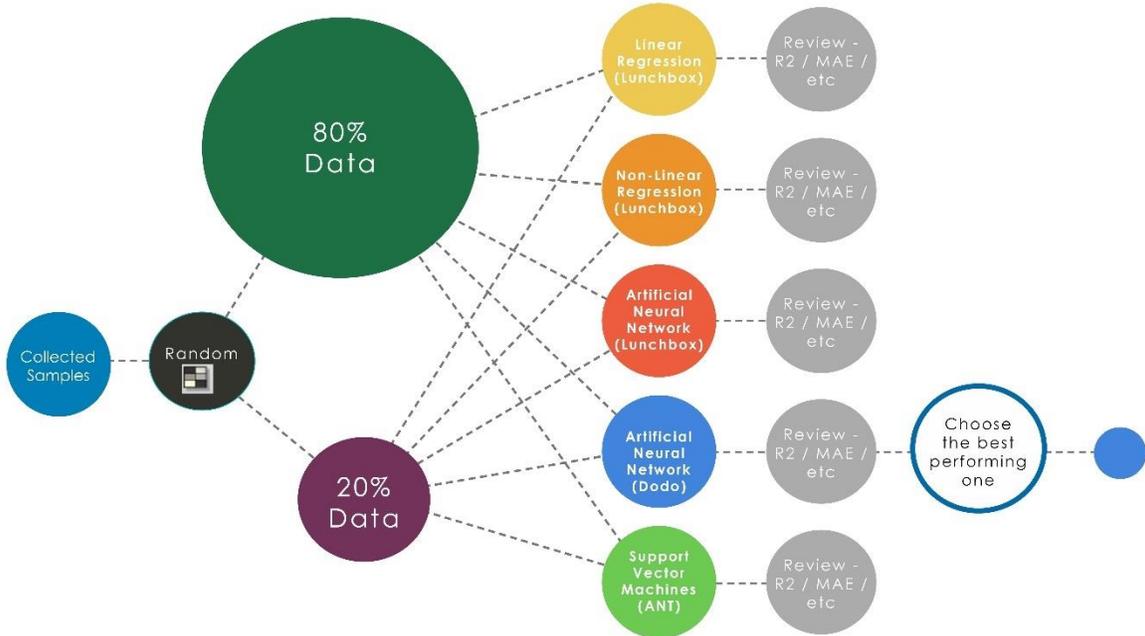


Figure 35 Machine learning workflow (source: own)

Using the custom script for the validation metrics, the 5 different machine learning methods could be compared and the best performing one could be chosen. This script serves as a tool to identify the ideal method as well as a tool to play with the random seeds and finetune the parameters of the ML methods.

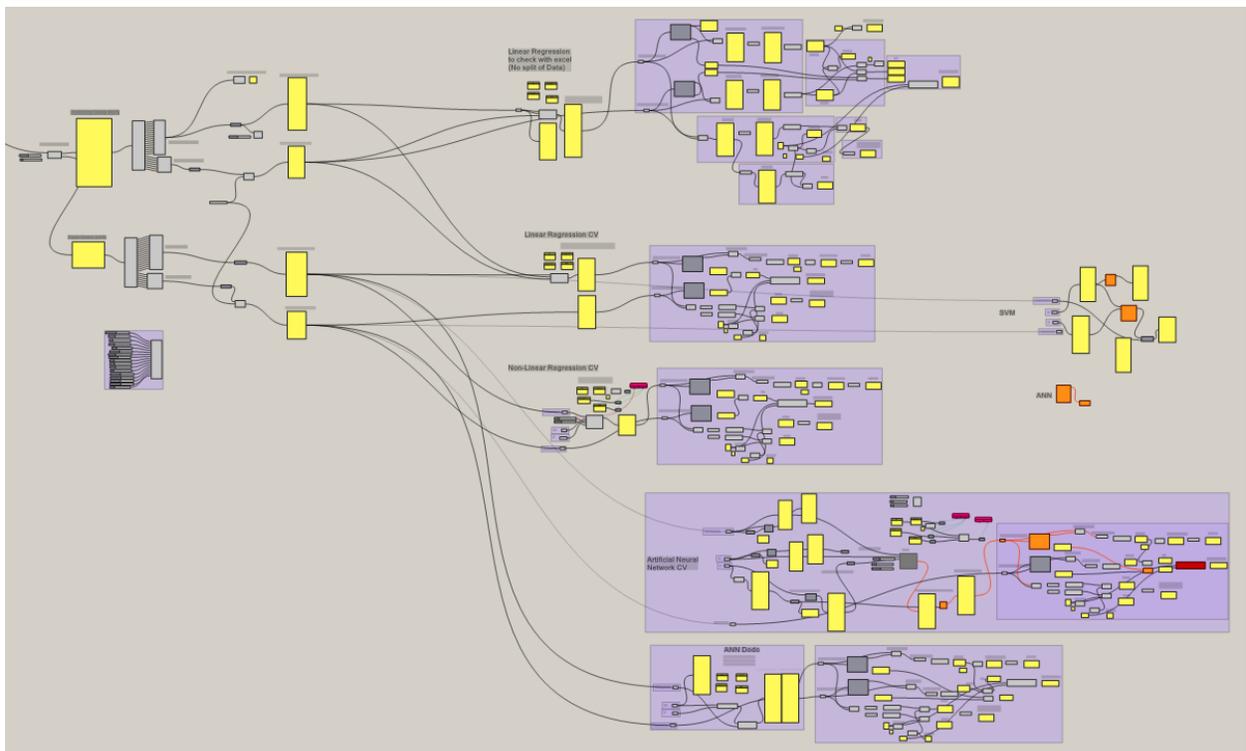


Figure 36 Custom Grasshopper Script for choosing the best performing ML method (source: own)

## Hyperparameter Tuning

### Non-Linear Regression

Having the custom script set up, it was possible to run a quick multi-objective optimization to adjust the Sigma and Complexity parameters of one of the ML methods: Nonlinear regression. Adjusted  $R^2$  was multiplied by -1 to trick the Octopus Hype Reduction algorithm into maximizing rather than minimizing, alternatively, Standard error and MAE were intended to be minimized. The results of this optimization slightly improved the Adjusted  $R^2$  to 0.48, still significantly low and resulting in an unwanted high standard error of 77 and as well as a high MAE of 276, thus still not good enough.

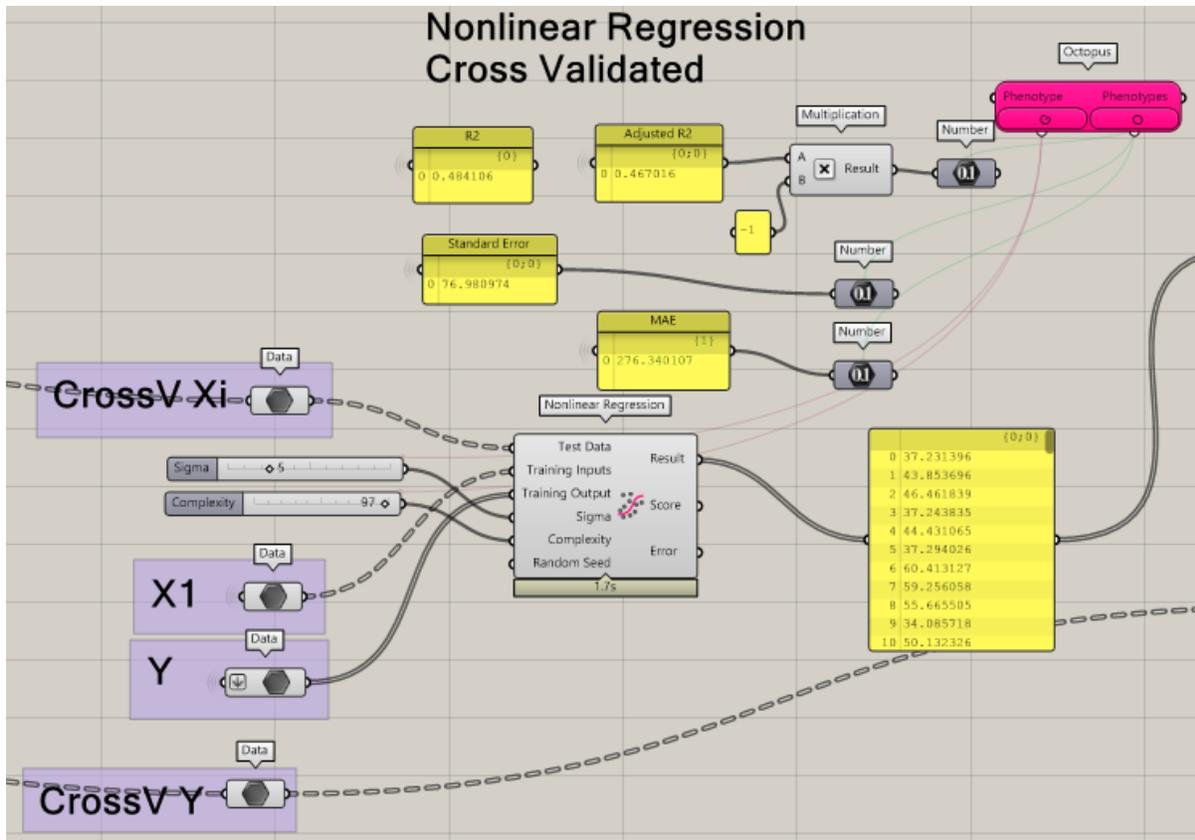


Figure 40 Non-Linear ML Method for UDI (source: own)

As literature showed, the methodology of optimizing the hyperparameters of a Machine Learning method has been previously explored. (Brea, et al., 2020). Rather than HYPE reduction and Non-linear Regression however, the researchers used multi-objective Non-dominated Sorting Genetic Algorithm II (NSGA-II) coupled to the ANN metamodels (Brea, et al., 2020). This step was not further explored to improve the final surrogate models because of time constraints and because the chosen Dodo (Greco, 2020) interface functions with a separate pop up window rather than optimizable sliders within the Grasshopper environment. Instead, the final chosen ANN model was iteratively fine-tuned to achieve acceptable results. This methodology, however, shows great potential to improve the accuracy of surrogate models for future research.

First, Artificial Neural Networks from Lunchbox was explored. For ANN the data was required to be normalized between 0 and 1 and then “denormalized” again into the predicted UDI values. A similar optimization approach using ANN variables (Hidden layers, Alpha and Max Iterations) was tested yet the results did not pass the validation criteria: 80-95% Adjusted R<sup>2</sup> / Standard Error and MAE > 10.

For this reason, ANN on Dodo was also tested. (See figures 41 & 42)

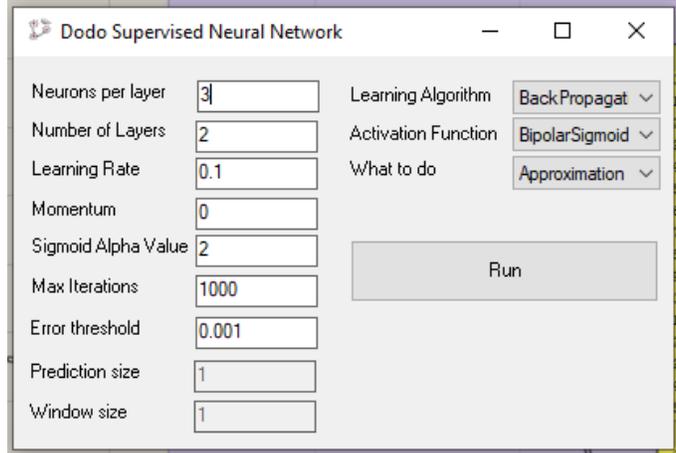


Figure 41 Dodo ANN Hyperparameter Optimization (source: own)

Due to its window-based interface, the ANN variables (Neurons, Hidden Layers, Learning Rate, Max Iterations, Sigmoid Alpha, etc.) could not be run through a slider optimization plugin, hence it was modified by a trial and error method. After some manual iterations, results achieved an average **Adjusted R<sup>2</sup> of 90%** with a low **Standard Error of 2.4** and **MAE of less than 1**, as presented in the image below:

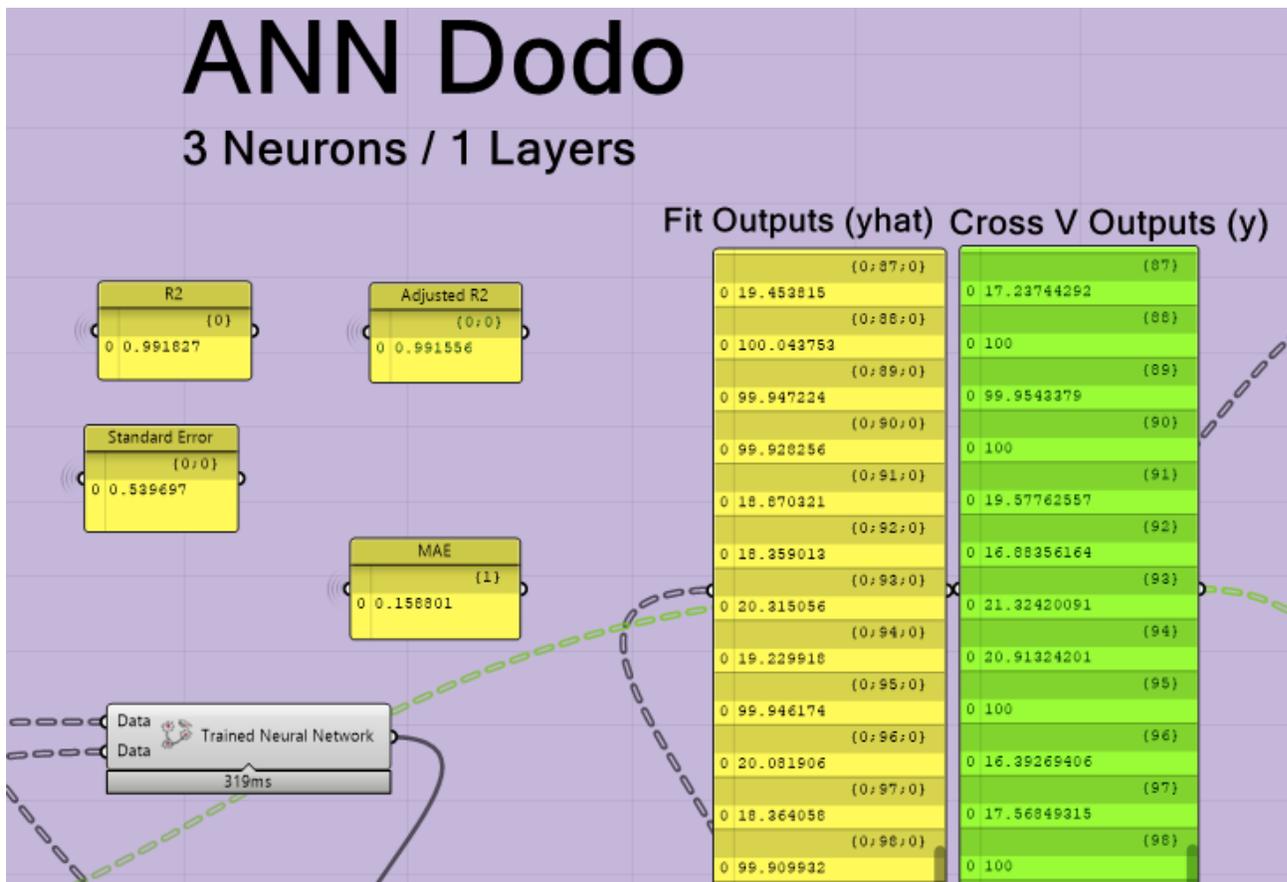
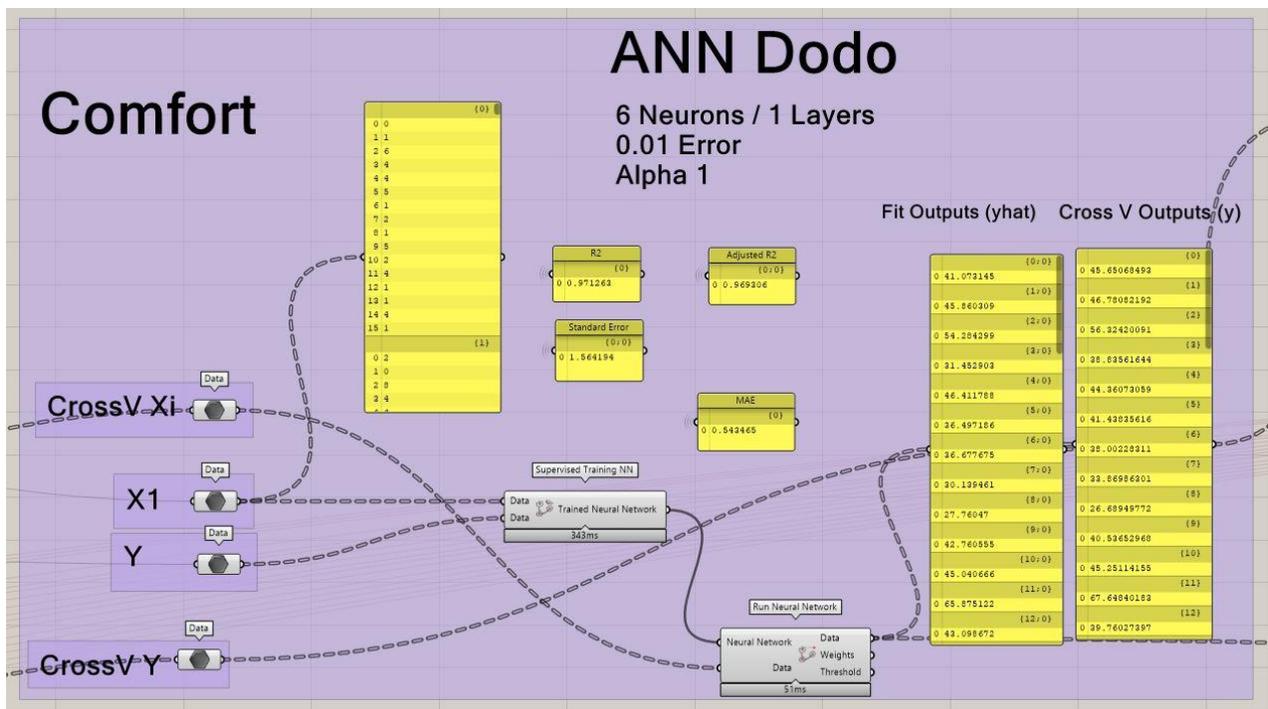


Figure 37 Results of Tuned ANN / Predicted Outputs (yhat) vs Cross-Validation Outputs (y) (Source: own)

The side by side table above shows a set of random 12 Y output values. The outputs (Y) from the independent 20% cross-validation test data show very close rapport to the simulation outputs ( $\hat{Y}$ ) of the ANN model. Even when switching the dataset to other random seeds, the predictive power remained above 80% with the standard error below 3, suggesting a good surrogate model.

For ANN Dodo, the training phase seemed a lot more exploratory than initially suspected because the choices made seemed to affect the model randomly. The improved trial an error manual method consisted of changing one variable at a time and reviewing if the  $R^2$  increased while MSA and MAE errors decreased. Once an improvement was identified, the last choice was kept and the next hyperparameter was modified. This ensured steady yet slow progress of the ANN-based surrogate model training.

Although, for predicting UDI, it is worth noting that in some rare runs the calculated  $R^2$  value was resulting in values slightly higher than 1 (See Seed 3 ANN Table J Below), meaning that a higher than 100% explained relationship. This occurrence makes no logical sense and upon further research, it was concluded that the model UDI behavior might be slightly too complex in some cases for the model to accurately predict its behavior. (Frost, 2020). This could mean that the original 500 sample size for the 16 independent variables (inputs) vs its 4 dependent variables (outcomes) might have fallen slightly short and more samples could improve this slight error.



Example of the final ANN Dodo model with tuned Hyperparameters:  
6 Neurons / 1 Layer / 0.01 Error threshold / Sigmoid Alpha Value 1

## Simulation BPS Model vs Surrogate model

Yet another step to review that the results from the surrogate were significant was to use the original “control” parametric model to run a daylight and energy simulation. Then these results were compared with those from the final surrogate model:

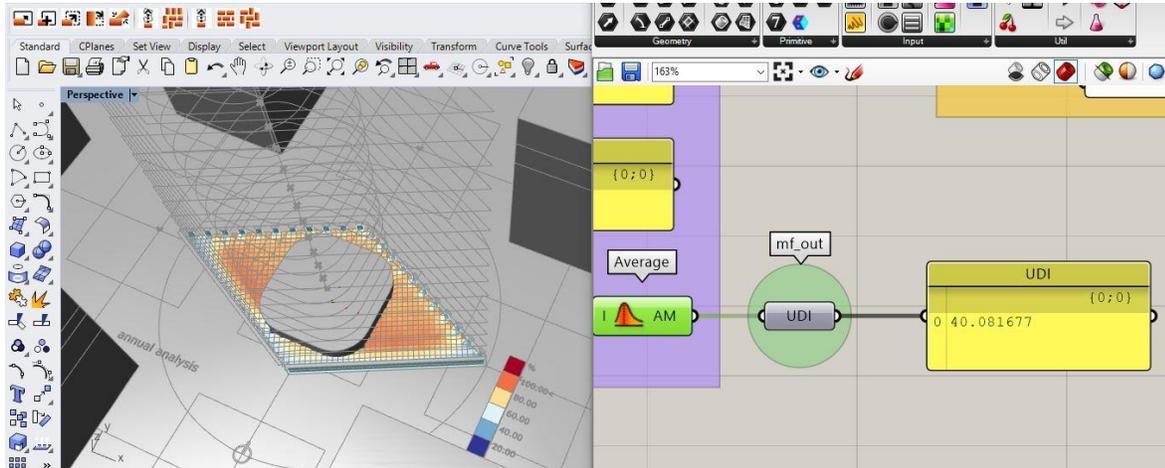


Figure 38 Original Simulation Control model UDI: 40.1 (source: own)

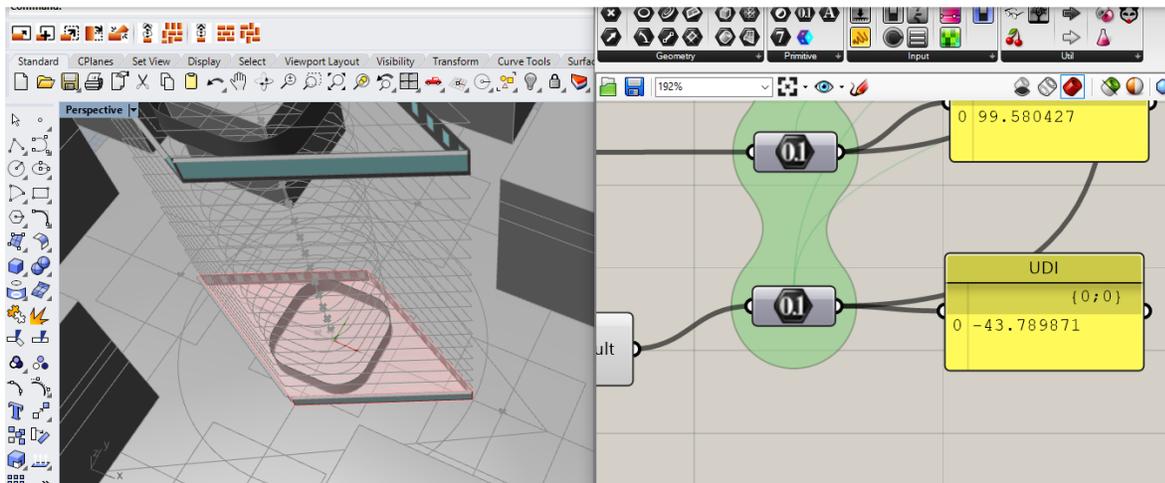


Figure 39 Final Surrogate Model UDI:43.8 (source: own)

As can be observed from the table 15 and figures 43 and 44, the results from the Surrogate Model were inline (within 10%) with the results from the simulated control model; this gave a further reassurance that the methodology and workflow were performed correctly.

Table J

Comparison of results from Simulation

	Original BPS Control Model	Final Surrogate Model	% Difference
UDI	40.1	43.8	8.4
EUI	95.5	100.2	4.7

Table 14 Results from Simulation vs Surrogate Model for Amsterdam Location (source: own)

## Comparison of the Machine Learning Methods

Table K

			R Square	Adjusted R Square	Standard Error	Mean Absolute Error
Multilinear Regression	EUI	Seed 1	0.718145	0.708789	5.980838	1.686682
		Seed 2	0.877971	0.87392	5.098759	1.791392
		Seed 3	0.969919	0.96892	5.141372	1.784924
		Seed 4	0.98683	0.986393	5.323705	1.86978
		Seed 5	0.875876	0.871756	6.217482	1.971926
		Average	0.8857482	0.8819556	5.5524312	1.8209408
	Comfort	Seed 1	0.729863	0.720896	2.624437	0.886089
		Seed 2	0.71456	0.705085	0.80835	2.386
		Seed 3	0.842415	0.837184	2.474655	0.824494
		Seed 4	0.96553	0.964386	2.231078	0.730064
		Seed 5	0.962526	0.961282	2.757766	0.99436
		Average	0.8429788	0.8377666	2.1792572	1.1642014
	UDI	Seed 1	0.438801	0.420172	5.044999	1.926218
		Seed 2	0.230504	0.204961	4.951339	1.834723
		Seed 3	0.346347	0.324649	4.78922	1.634628
		Seed 4	0.387203	0.366861	4.680839	1.621017
		Seed 5	0.449362	0.431084	5.576187	1.968647
		Average	0.3704434	0.3495454	5.0085168	1.7970466
Non-Linear Regression	EUI	Seed 1	0.62267	0.610145	177.78455	544.72342
		Seed 2	0.73601	0.727247	171.1969	599.72839
		Seed 3	0.730072	0.721112	149.55488	507.67127
		Seed 4	0.729833	0.720865	153.29181	500.8856
		Seed 5	0.651518	0.63995	190.65296	648.33496
		Average	0.6940206	0.6838638	168.49622	560.26872
	Comfort	Seed 1	0.810935	0.804659	83.224339	310.44516
		Seed 2	0.774355	0.766865	74.773485	259.47391
		Seed 3	0.951593	0.949986	72.030026	248.24983
		Seed 4	0.869526	0.865195	66.113987	218.64417
		Seed 5	0.865054	0.860574	78.121753	283.83417
		Average	0.8542926	0.8494558	74.852718	264.12945
	UDI	Seed 1	0.49283	0.475994	78.454722	298.25815
		Seed 2	0.251743	0.226905	72.929642	263.21719
		Seed 3	0.355431	0.334035	70.662601	244.83104
		Seed 4	0.38354	0.363077	67.68632	228.81792
		Seed 5	0.429776	0.410847	77.260067	285.12267
		Average	0.382664	0.3621716	73.39867	264.04939

Artificial Neural Network						
Artificial Neural Network	EUI	Seed 1	0.983539	0.982993	7.590498	2.129707
		Seed 2	0.963922	0.962724	6.488874	1.928935
		Seed 3	1.085034	1.087857	5.802253	1.805024
		Seed 4	0.933024	0.930801	4.532912	1.550273
		Seed 5	0.945534	0.943726	4.19745	1.445061
		Average	0.9822106	0.9816202	5.7223974	1.7718
	Comfort	Seed 1	0.887935	0.884215	1.388643	0.491288
		Seed 2	0.870895	0.866609	1.484493	0.506394
		Seed 3	0.957552	0.956143	1.233582	0.436458
		Seed 4	0.940751	0.938784	1.304237	0.446558
		Seed 5	0.842245	0.837008	1.951431	0.67656
		Average	0.8998756	0.8965518	1.4724772	0.5114516
	UDI	Seed 1	0.951914	0.950318	4.104351	1.486348
		Seed 2	0.847411	0.842346	2.072161	0.653077
		Seed 3	0.962942	0.961712	2.037623	0.621746
		Seed 4	0.825666	0.819879	2.190636	0.710077
Seed 5		0.915687	0.912888	1.899354	0.647612	
Average		0.900724	0.8974286	2.460825	0.823772	

In the final step to review and to choose the ideal Surrogate model, 5 random seeds were generated and used to compare the three selected ML methods, allowing them to identify the one with the best predictive power and accuracy.

As seen from the table above, the ANN model performed best for predicting EUI, Comfort, and UDI. With a sufficiently high average Adjusted R Squared of 98% for EUI, 87% for Comfort, and 90% for UDI. Likewise, a Relatively low standard error of 5.7 and 1.8 MAE for EUI, 1.4 and 0.5 MAE for Comfort, and 2.5 and 0.8 MAE for UDI. In conclusion, after tuning the hyperparameters of the ANN-based metamodel, this proved to be the most accurate among the three ML methods explored, therefore it was used for the remainder of this thesis.

### 3.7 Computational Optimization

The main objective of this research paper was to lower the energy consumption of a high-rise building by reducing the EUI while simultaneously maintaining daylight, comfort, and floor area as design constraints or goals. Firstly, this purpose was defined as a single objective optimization problem. Secondly, it was also viewed as a multi-objective optimization problem. Both optimization strategies were then compared.

#### 3.7.1 Single-objective optimization (SOO)

A fitness function or objective function was defined to simplify the search space. This function served to define a merit criterion. The single objective was to minimize the EUI. Moreover, UDI, Comfort, and Floorplate area became the constraints, essentially creating a boundary value problem. Choosing these benchmarks, proved critical for any of the subsequent optimization runs.

The constraint on thermal comfort was set to a **comfortable at least 60% of the time** considering that additional steps using an ideal HVAC system would have to be taken to improve it further.

There is no single consensus on the ideal UDI daylight for office buildings, but on average, the measured **UDI (100-2000 lux) for at least 50% of the time** is considered a sufficiently good boundary condition (Umberto Berardi, 2015)

As for the floor plate area. An office building “Atrio” built-in Bogotá in 2019 was used as a reference, this building has a central concrete core and a Gross floor area of approximately 1200 m<sup>2</sup>. A search around this global area proved as a good reference for a state-of-the-art and marketable high-rise office building. The benchmark was therefore expanded to look within a range of +/- 700 m<sup>2</sup> thus constrained valid solutions between **500 and 2000 m<sup>2</sup>**.

A C# component was used as a penalty function with an “If. Else” gate to filter the outputs of the surrogate model with these conditions:



Figure 40 C# component for single Objective optimization (source: own)

If any of the results fell outside these conditions, the four outputs would be multiplied together, giving an impractically high number. Hence, the ideal EUI minimization solutions would always fall within these limits.

The penalty function was then run through three different single-objective optimization components and four different optimization algorithms: Galapagos running Evolutionary Algorithm, Opossum running RBFOpt, Opossum running CMAES (with a random start) and Optimus running jEDE.

As in previous research, to understand the performance of the optimization tools, the lowest value for the fitness function EUI ( $f_{x\_min}$ ) was recorded. The time limit was set to 20 minutes (1200) and the maximum number of fitness evaluations (FES) were recorded to identify the tool that performed the fastest. (Cubukcuoglu, et al., 2019) . For each, at least 5 runs were performed to ensure that the algorithm was consistent with these results.

The optimization problem ran on a computer with the following specifications: Intel® Core™ i7-3630QM CPE @ 2.40 GHz, x64bit Windows 10 Home, RAM 8 GB

As a proof-of-concept, the Amsterdam location was used for these single-objective optimizations, yet the methodology could be applied equally to Bogota and Dubai locations

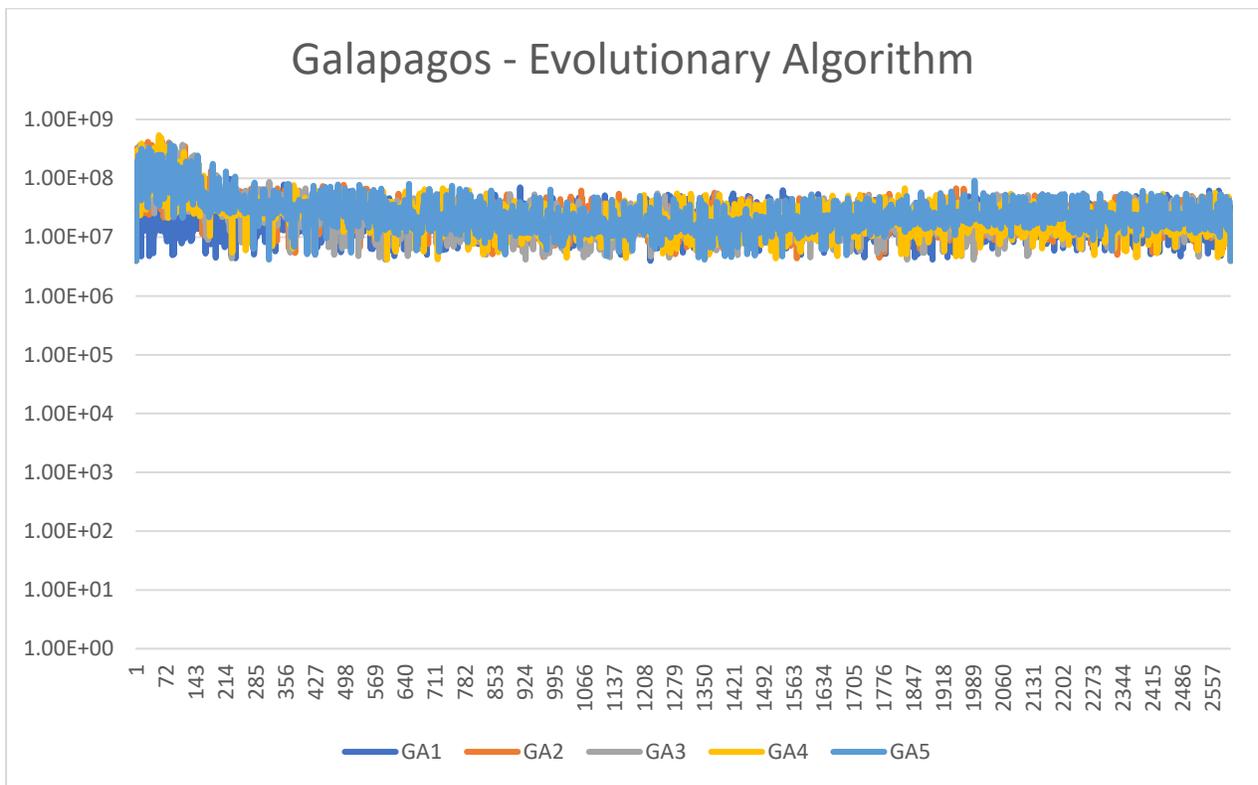


Figure 41 Convergence Graph - Galapagos GA: FES in 20 minutes (source: own)

For Galapagos, after 1200 seconds and approximately 2605 FES, the  $F(x)_{min}$  AVG remained very high (1.21E+08) and did not converge for any of the 5 runs. The input variables for each of the solutions were also recorded, yet not further explored because the solutions were not deemed valuable as they fell outside the ideal constraints.

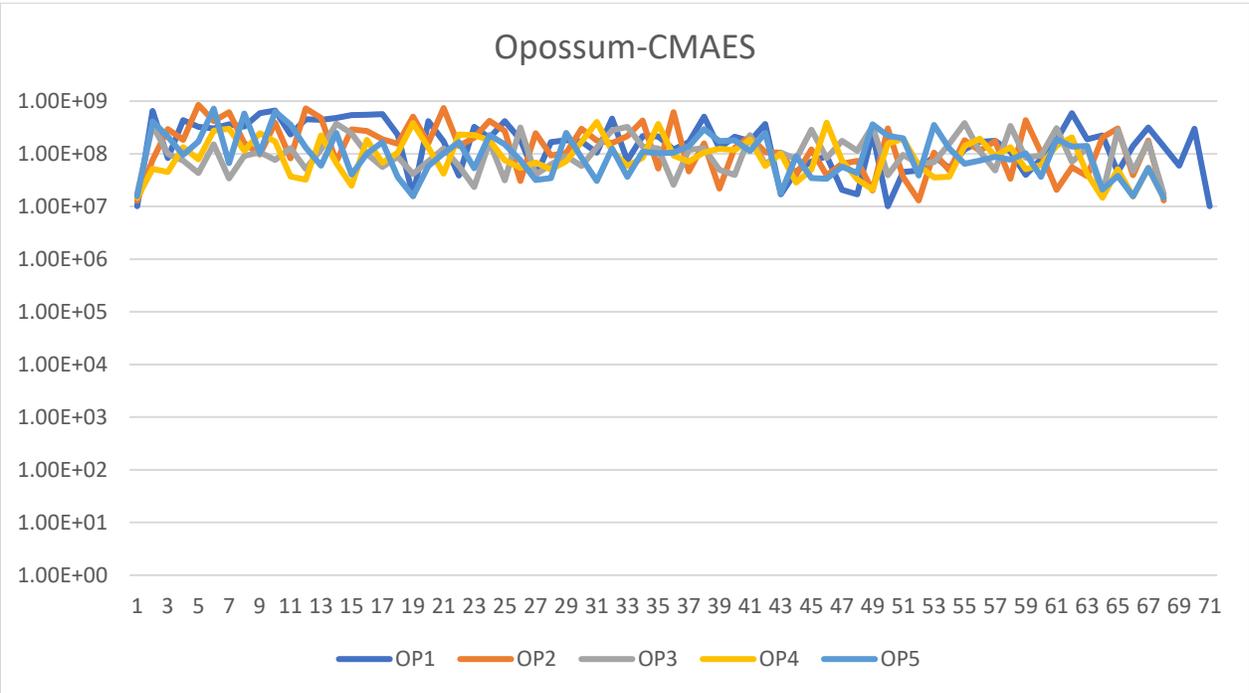
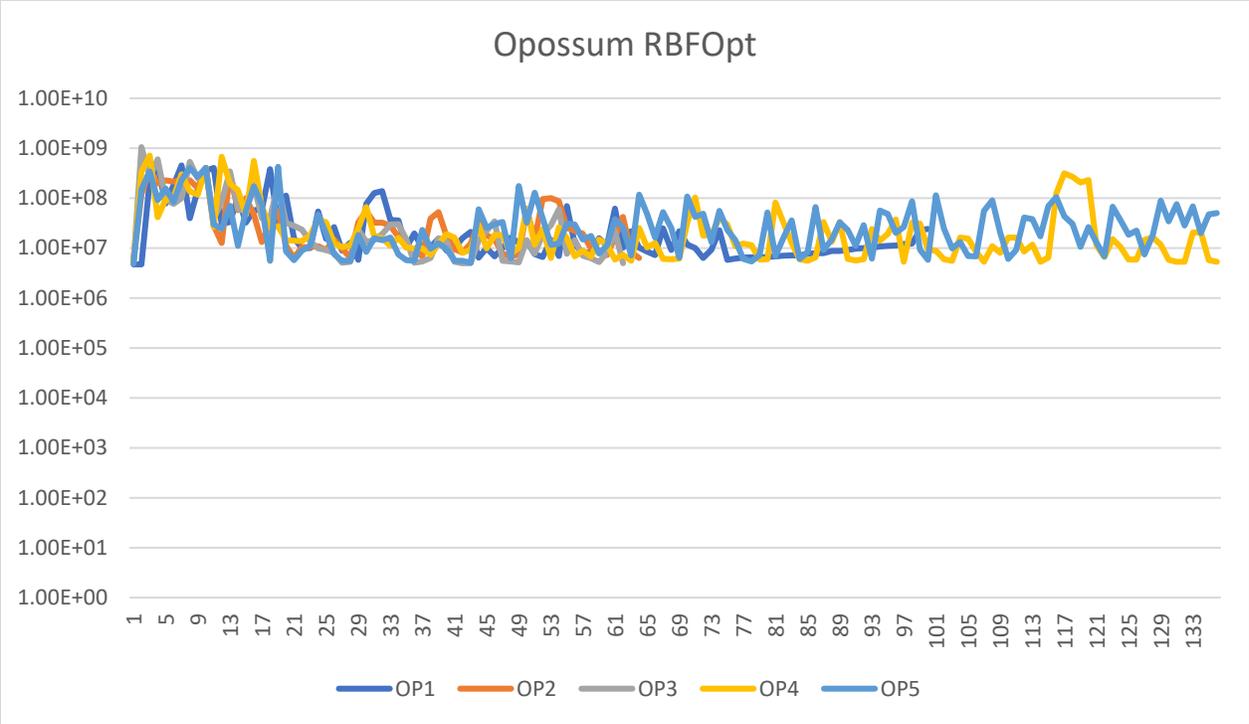


Figure 42 Convergence Graph - Opossum RBFOpt & CMAES: FES in 20 minutes (source: own)

Similarly, for both optimization algorithms used by Opossum, neither RBFOpt nor CMAES algorithms were able to achieve a feasible convergence of the fitness function within the allotted time of 1200 seconds. The lowest values for EUI remained very high  $F(x)_{\min}$  AVG of  $5.28E+06$  and  $1.41E+07$ , meaning that the results must have been stuck in a local optimum. No solution that would contemplate all constraints was found. The input variables were saved yet not further results for comfort or UDI were explored.

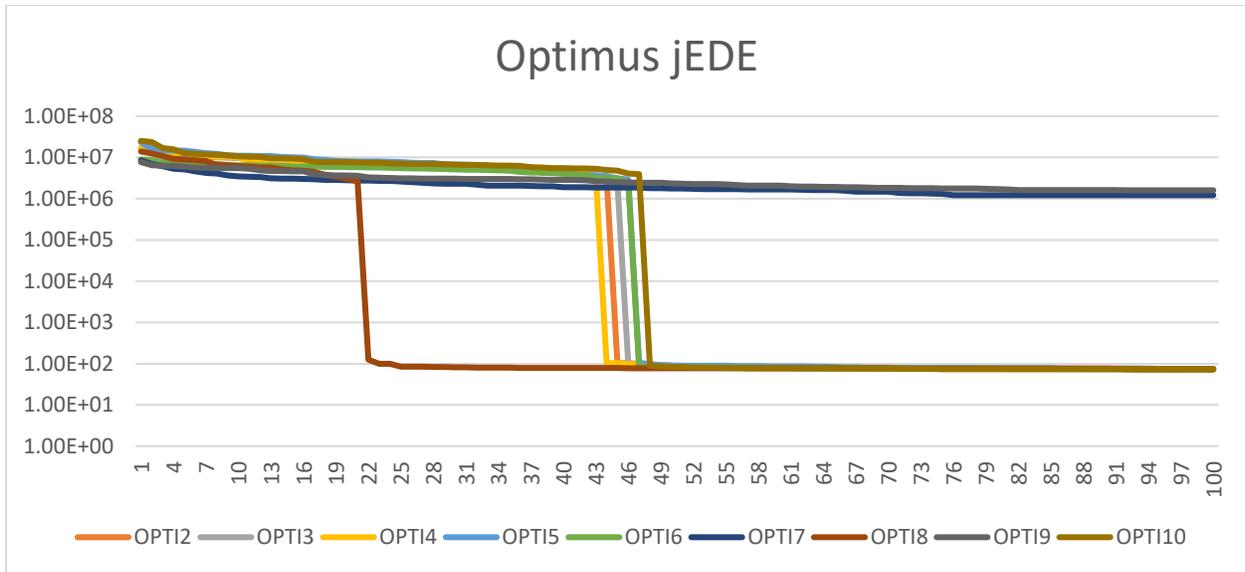


Figure 43 Convergence Graph – Optimus jEDE: FES in 105 seconds (source: own)

Differently from the previous 3 single-objective optimization algorithms, Optimus jEDE did converge in 8 of the 10 performed tests. More impressively, was the fact that in conducted 5000 FES in 105 seconds or less. This speed allowed us to perform 10 rather than 5 runs. While all other optimization tools took a total of 6000 seconds for the 5 runs (100 minutes / 1.67 hours), on Optimus running twice as many runs (10) took a total of 1050 seconds (16.67 minutes). Although in some cases, like a test run 7 and 9, the fitness function did not converge, the speed of the tool makes up for this by permitting many more runs in total and by allowing the designer to explore various FES settings quickly and with a low computational load.

The following table shows how the optimization is repeatedly converging to the same or similar solutions for each x variable. In green, the variables that constantly show up for the 8 valid iterations, while in yellow the variables that are also repeated, albeit less frequently. The recurrence of these variables means that certain ideal solutions for the Amsterdam location can be discerned for each option.

(To read the results correctly, refer to the Legend table D in section 3.2.4)

## Optimus – jEDE – Amsterdam location

105	Seconds	50 Generation					100 Population Size				
FES 5000	Opti 1	Opti 2	Opti 3	Opti 4	Opti 5	Opti 6	Opti 7	Opti 8	Opti 9	Opti 10	
	Input	Input	Input	Input	Input	Input	Input	Input	Input	Input	
Analysis_Level	0	0	1	0	0	0	2	0	2	0	
Core_Type	1	1	1	1	1	1	1	1	1	1	
FFH	6	6	6	6	6	6	6	6	7	6	
Length	4	4	4	1	1	1	1	4	1	3	
Number_Of_Fins	0	0	0	0	1	3	2	0	0	2	
Orientation	0	7	7	7	6	1	7	7	7	6	
Overhang_Size	4	4	2	4	1	4	1	1	1	1	
SHGC	3	3	3	3	3	3	1	3	2	3	
Shape	3	3	0	3	3	3	0	0	0	0	
U_Value	1	1	1	1	1	1	7	1	7	1	
VLT	3	3	3	3	2	2	3	3	3	3	
WWR_East_North	1	1	1	2	1	1	2	2	1	1	
WWR_North_West	4	4	4	4	4	4	4	4	2	4	
WWR_South_East	1	1	1	1	1	2	1	1	2	1	
WWR_West_South	4	2	1	1	4	2	2	1	2	1	
Width	2	1	2	3	3	3	1	2	1	3	
<b>EUI</b>	72.34	73.23	73.19	73.65	74.80	74.20	1.22E06	73.78	1.60E06	72.57	
<b>Comfort</b>	61.68	61.87	61.88	61.45	69.54	59.58		69.54		66.53	
<b>UDI</b>	60.37	53.63	62.27	54.77	67.99	61.33		63.41		64.47	
<b>Floor Area</b>	2710	1368	1480	1016	1016	1016		1480		1703	

- Analysis level: 0 means that the highest level 30 m is preferred, for balancing these objectives, a high-rise would do better without urban context.
- Core Type: a central core type is consistently better when lowering EUI is the main goal.
- FFH: The lowest floor to floor height of 3 m is the ideal for the lowest EUI.
- The number of vertical fins surprisingly shows 0 repeatedly, their effects on energy consumption might not be as significant on this location.
- SHGC: Solar heat gain coefficient is usually ideal using the highest setting on this model (0.8)
- Shape: Elliptical shape (3) shows up repeatedly yet the triangle (0) also seems to give good EUI
- U-Value is unsurprisingly ideal when the lowest, meaning there should be less heat transfer through the façade to save on EUI.
- VLT: Visible light transmittance is usually ideal when highest (90%)
- WWR North is consistently the highest 80%, which is as expected because only useful indirect sunlight penetrates through the northern side.

- Conversely, WWR South is consistently lowest (20%), because direct sunlight should be avoided to shield the floorplate from overheating during the summer, avoiding in higher cooling loads.

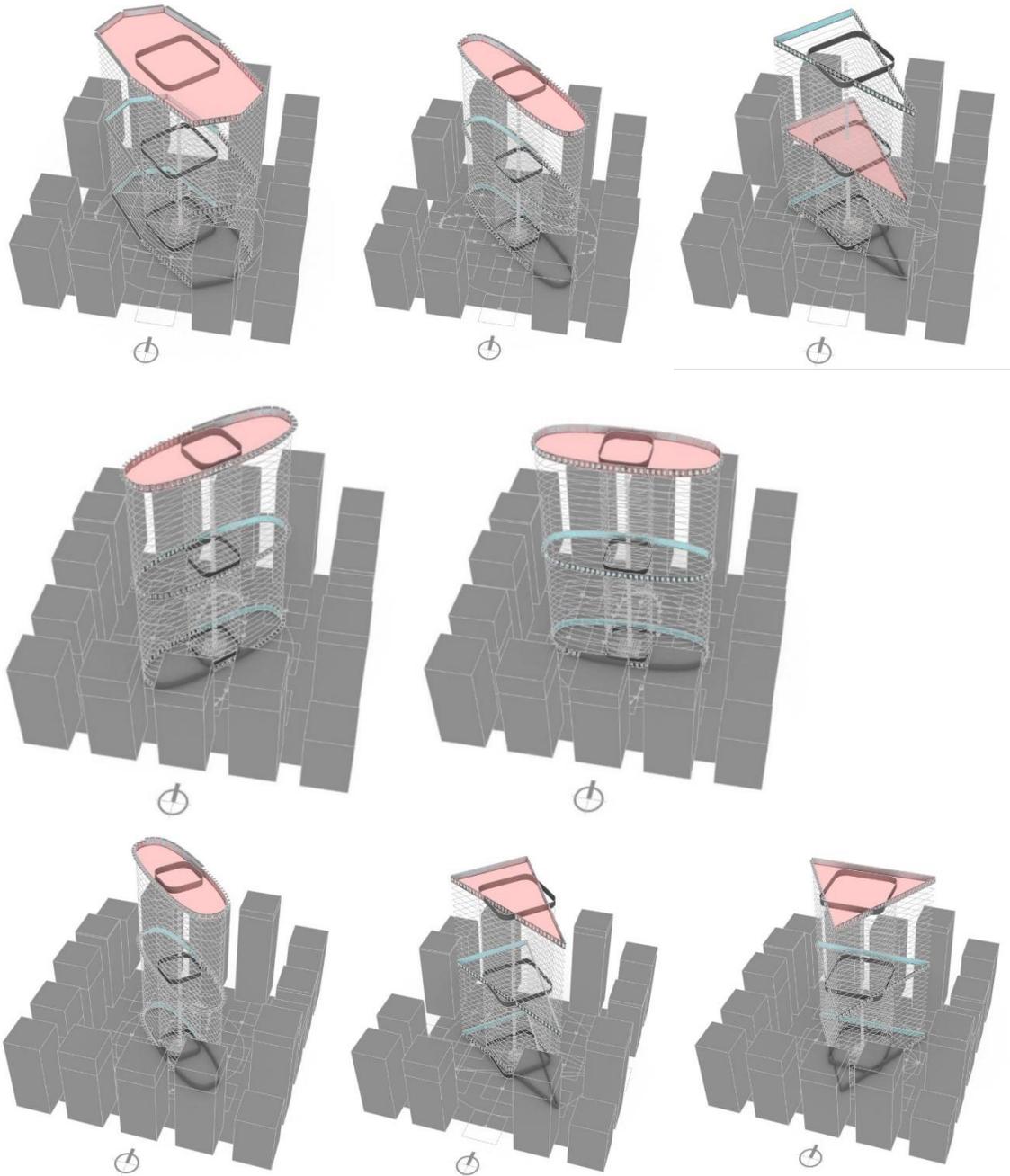


Figure 44 Optimus jEDE - Solutions 1-6, 8 & 10 (source: own)

### 3.7.2 Multi-Objective Optimization (MOO)

For multi-objective optimization, the grasshopper plugin Octopus (Vierlinger, 2018) was used. A three objective optimization using SPEA 2 Reduction algorithm was performed. The objectives optimized for this run were: EUI, UDI, and Comfort, while the Floor area was not optimized but still recorded to allow for a more complete exploration of the possible design spectrum.

Being a multi-objective optimization, there was no need to use the C# component with the penalty function. Rather than limiting the search by constraints, the objectives of EUI, Comfort, and UDI were explored simultaneously from the start of the run; this meant that the algorithm began the search space within already acceptable margins and improved each design within the Pareto-front (Cenaero, 2018). A Pareto-front between these outputs was quickly established and various options within this non-dominated boundary were explored.

Similarly, to the single-objective optimization, the tool was run 10 different times and allowed to optimize for 1200 seconds (20 min). After the 10 runs for a total time of 12000 seconds (200 minutes / 3.3 hours) the results were recorded. As an example, one of the results is presented next.

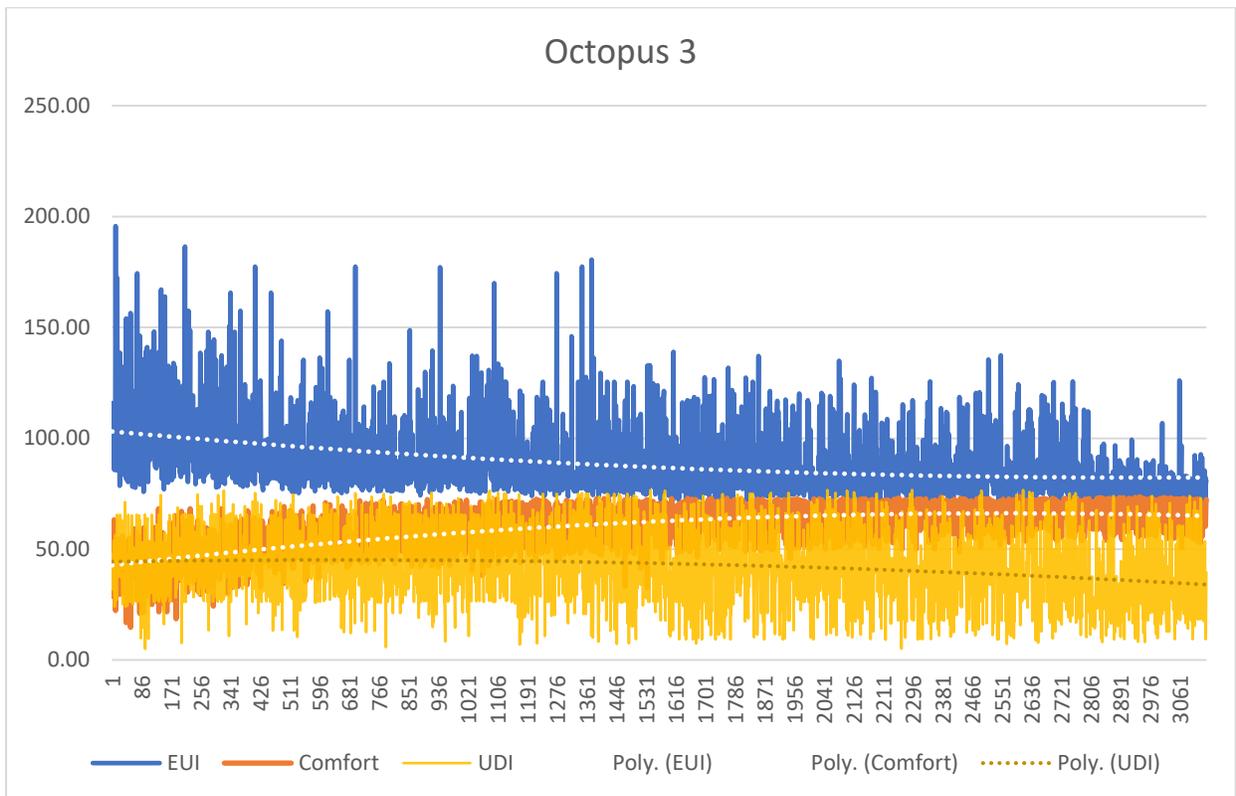


Figure 45 Convergence Graph: Octopus run 3: FES in 20 minutes (source: own)

The above chart shows the third of the 10 optimization runs (see appendix for others) to exemplify the progressive reduction of the EUI and comfort. Conversely, UDI is also slightly reduced when searching for Pareto-optimal solutions. This trend could be explained because of the inherent tradeoffs of a multi-objective optimization with conflicting goals, meaning that by lowering the EUI while improving comfort,

UDI might slightly decrease also. Nonetheless, solutions in the Pareto front were still high in all three objectives as seen from the collected samples.

### Iterations within the solution space

The following figures show the complete solution space for one of the runs performed in Octopus, each cube representing an iteration that the algorithm reviewed. The more transparent cubes are the history of the optimization process, meaning that better solutions non-dominated solutions were found, these solutions are represented in solid colors. The sizes of the cubes represent in an abstract manner the relative size of the floor plate, allowing the designer to intuitively identify which solutions have higher or lower floor areas. A Pareto-optimal line can be identified with each progressive generation. By identifying a larger hypervolume, EUI and UDI tend to show that they are indeed conflicting objectives

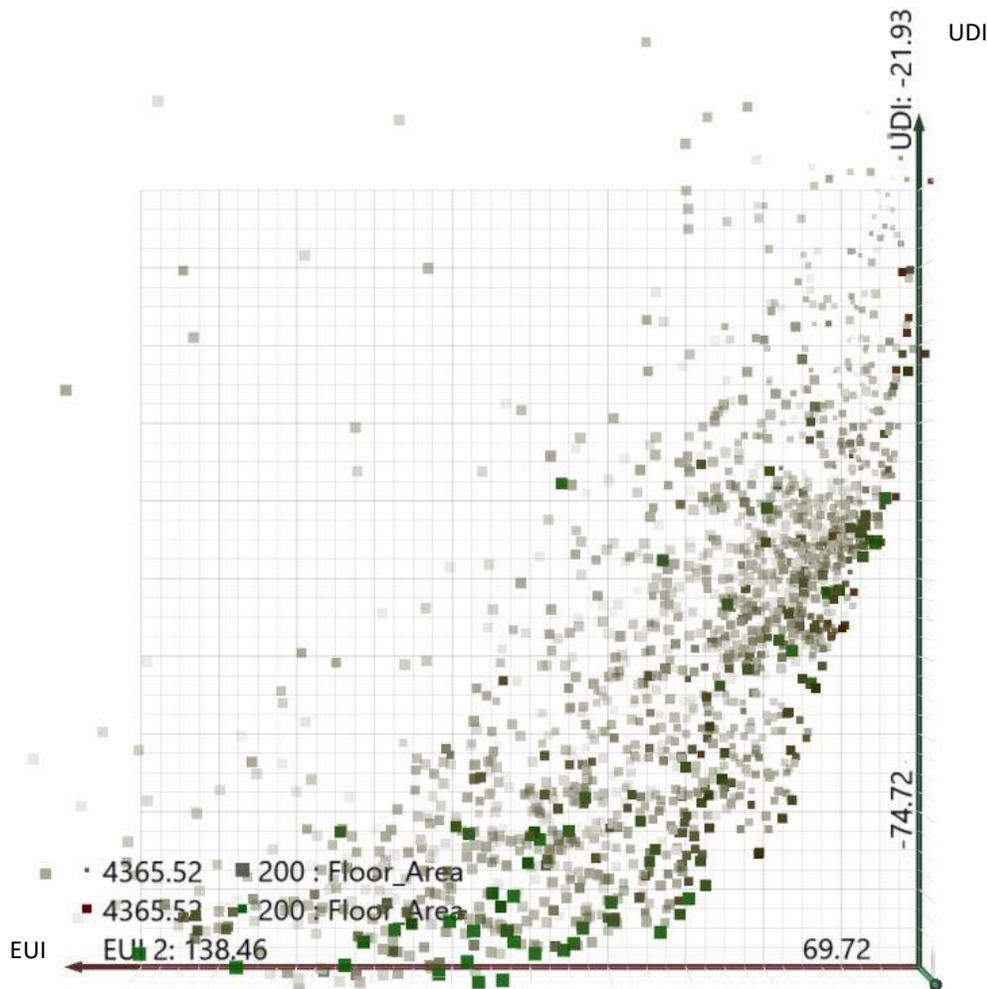


Figure 46 Multi-objective design space – EUI (x) / UDI (y) (source: own)

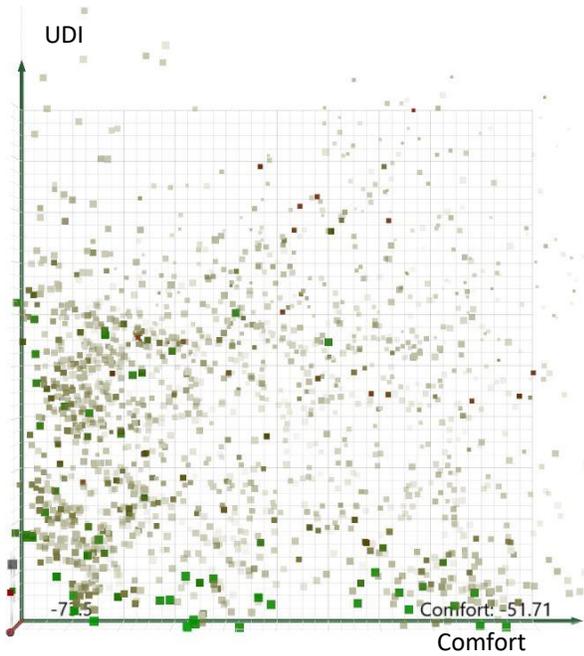


Figure 48 Multi-objective design space - Comfort(x) / UDI(y)  
(source: own)

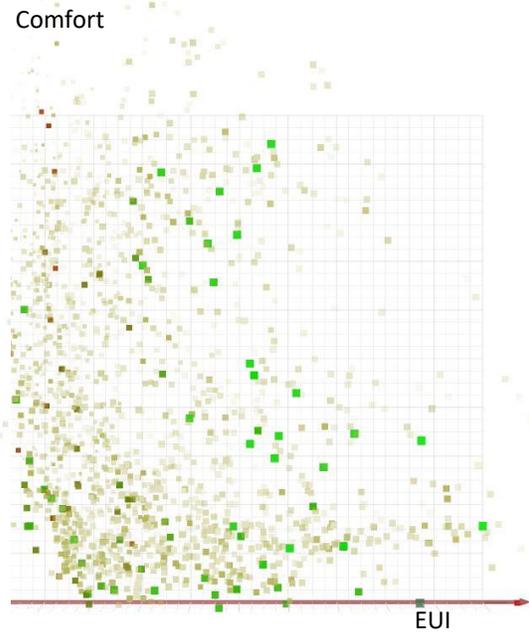


Figure 47 Multi-objective design space - EUI (x) vs Comfort (y)  
(source: own)

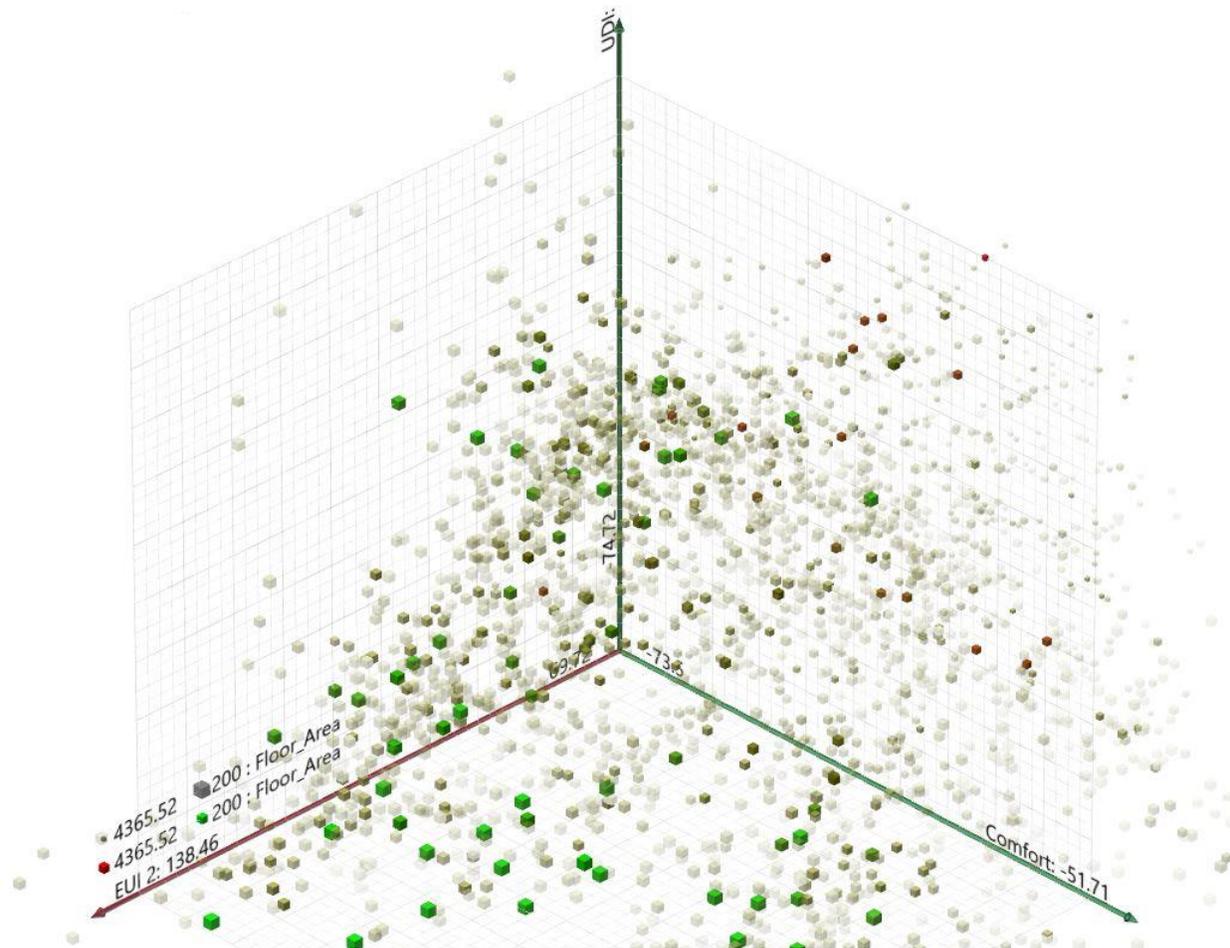


Figure 49 Multi-objective design space - EUI (x) vs Comfort (y) vs UDI (z) (source: own)

1200 FES 3100	Seconds	36 Generation					100 Population Size				
	Oct1	Oct2	Oct3	Oct4	Oct5	Oct6	Oct7	Oct8	Oct9	Oct10	
	Inputs	Inputs	Inputs	Inputs	Inputs	Inputs	Inputs	Inputs	Inputs	Inputs	
Analysis_Level	0	0	1	0	0	0	0	0	0	1	
Core_Type	1	1	0	1	1	1	0	0	0	0	
FFH	6	8	6	7	6	6	6	6	7	7	
Length	3	1	1	3	3	3	4	2	4	2	
Number_Of_Fins	4	2	1	0	0	2	0	0	2	2	
Orientation	2	5	1	3	7	6	6	5	0	4	
Overhang_Size	2	3	1	3	3	1	3	3	1	2	
SHGC	3	3	3	3	3	3	3	3	3	3	
Shape	2	3	0	1	2	0	0	1	3	1	
U_Value	1	1	1	1	1	1	1	1	1	1	
VLT	1	2	2	2	2	3	3	2	2	1	
WWR_East_North	1	4	4	1	1	1	1	4	3	2	
WWR_North_West	4	4	3	1	4	4	4	2	3	3	
WWR_South_East	1	4	2	4	2	1	3	3	3	3	
WWR_West_South	4	1	1	4	2	1	4	4	1	3	
Width	2	3	3	2	2	3	3	3	1	2	
EUI	79.23	84.24	99.10	83.91	74.23	87.30	81.31	93.61	86.18	100.48	
Comfort	67.60	68.15	64.64	70.18	67.00	68.20	64.84	69.86	66.16	63.72	
UDI	69.01	53.46	68.87	70.08	70.84	65.56	59.41	66.03	59.41	70.11	
Floor Area	2010.3	1016.0	470.9	1560.0	2010.3	1703.4	2319.6	1560.0	1367.9	1005.0	

In the same way as the table developed for the Optimus jEDE solutions, this table also shows frequent solutions that further reinforce the variables that are important when considering efficient high-rises that contemplate EUI, Comfort, and UDI. The conclusions for the Amsterdam Location are as follows:

- Analysis Level: Similarly, to single-objective optimization (SOO) the consistent choice (0) is measured at +30m meaning that for Amsterdam, an urban context is counterproductive, possibly due to higher cooling loads associated during the summer.
- Core Type: In this optimization, the last 4 solutions show that a lateral core type can also give good results, albeit the EUI is slightly higher in all these cases.
- FFH: The recurrence of the lowest (3m) and second-lowest (3.5m) floor to floor heights shows that a lower FFH is ideal. Even though modern high-rises are starting to use higher FFH for “modern-look” of flexibility, this new policy should be revisited if improving the EUI, Comfort, and UDI has sought after.
- Number of Fins: There is no clear pattern on the number of fins, probably signifying again that the fins do not play a major role in this location

- Orientation: Like the previous conclusion, there is no clear pattern for the orientation of the building in neither optimization runs. The reason for this could be because most of the ideal solutions found have a 1:1 plan ratio therefore the orientation is inconsequential. In options with longer lengths vs widths, the orientation is indeed more important.
- Overhang size: Likewise, there is no recurring solution for overhang size, probably explained because the sun varies drastically throughout the year, therefore there is no ideal passive solution for this criterion, and simply does not affect EUI, Comfort, and UDI as much as the other analyzed variables.
- SHGC: Also like the solutions from SOO, SHGC should be highest (0.8)
- Shape: Differently from the SOO shape when conducting a MOO, the shape did not have a clear pattern because it was a choice to choose from the many design options. SOO showed that when the sole objective is to lower the EUI, then the ellipse is the preferred option; just as was identified in the literature review. On the other hand, when other objectives like comfort and UDI are considered, the choice expands.
- U-Value is regularly ideal when lowest, this comes as no surprise and when design and cost permits, U value should be lowest in this location.
- VLT: In MOO, is not as defined, while in SOO VLT should be highest, here the choice is less defined.
- WWR North is also rather repeatedly set as highest (80%) with a few exceptions, taking advantage of the indirect northern light in the Northern Hemisphere.
- WWR South. Various the solutions for this side are “incorrect” because they do not follow the typical low (20%) WWR that SOO showed; nonetheless, upon further inspection, it was discovered that cleverly, the algorithm found interesting solutions (solutions 6, 8 and 10) to use the lateral core as a shield from the sun. Therefore, the southern windows can be larger while still protected most of the time during summer and receiving indirect sunlight during winter.
- Length and Width variables depend more on the choice of the final floor area, therefore there is no “optimal” choice but rather it allows the designer to tweak their design.

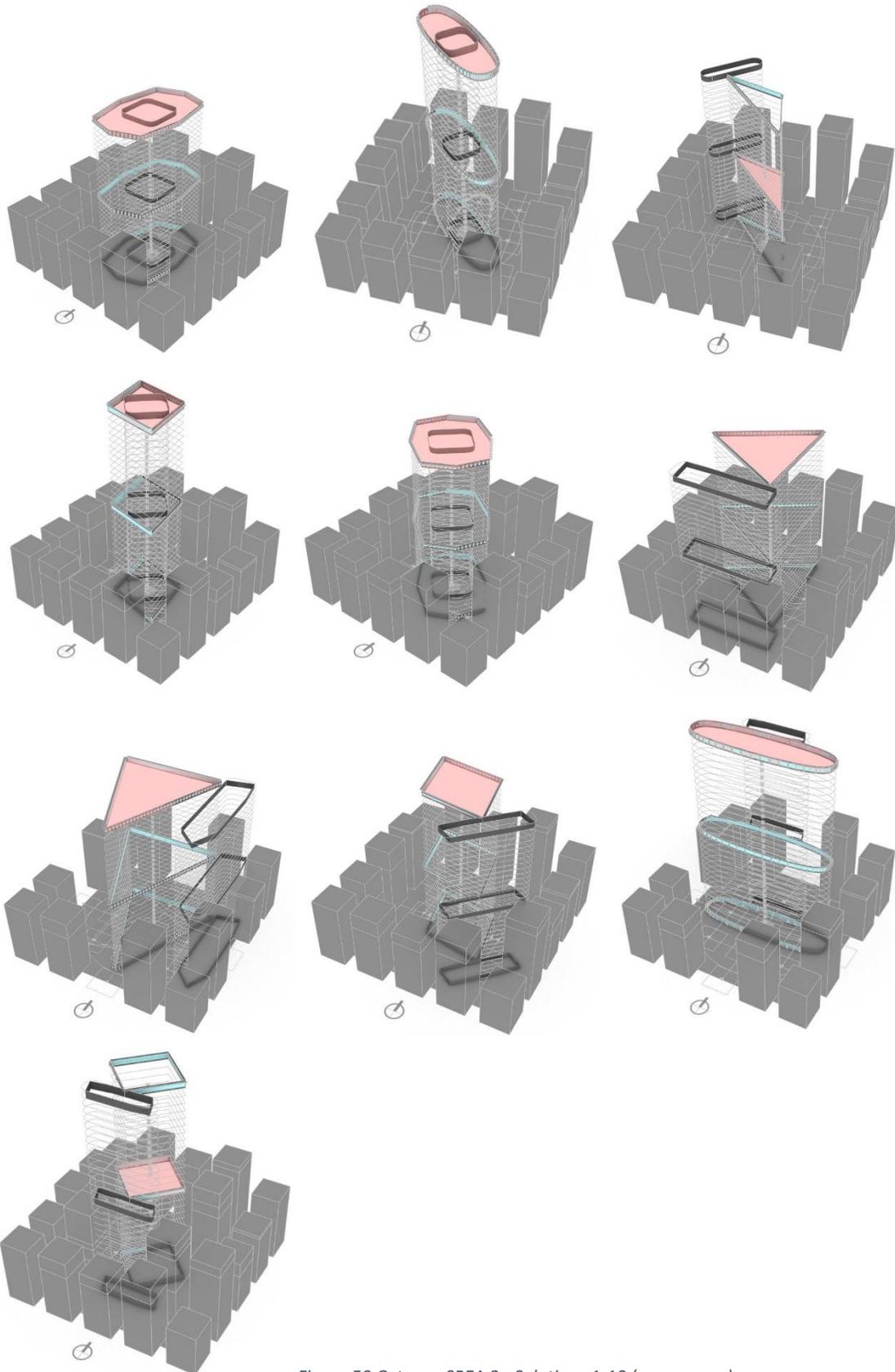


Figure 50 Octopus SPEA 2 - Solutions 1-10 (source:own)

This thesis-based its results as presented by the authors from the Optimus journal article (Cubukcuoglu, et al., 2019) To evaluate the performance of the optimization tools, EUJ is reported as  $f(x)_{\min}$ , signifying the minimum fitness value of this design problem.

The following table shows a summary of the 5 different optimization tools and algorithms used: 4 single objective and 1 multi-objective (Octopus SPEA 2). For clarity, only 1 of the 5 or 10 runs for each algorithm displayed, yet the overall convergence and FES are recurrent in each of the runs of these optimization methods.

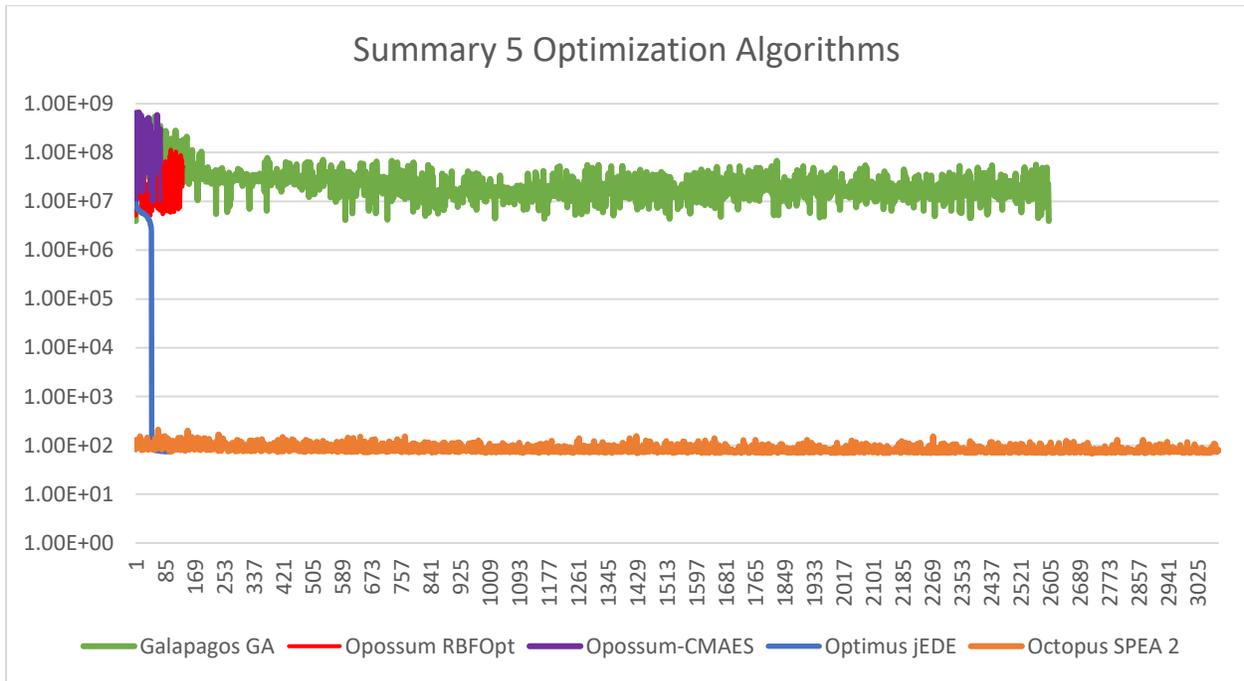


Figure 51 Convergence Graph- Summary of 5 Optimization Algorithms: FES in 20 minutes (source: own)

This table shows that Galapagos GA, Opossum RBFOpt, and Opossum CMAES failed to find a feasible solution for the penalty function aimed to reduce EUJ while maintaining the other objectives as constraints. Instead, Optimus jEDE found feasible solutions in 8 out of the 10 runs in a fraction of the time (105 seconds instead of 20 minutes). Likewise, Octopus SPEA 2 also proved useful by identifying many optimal Pareto solutions within the 20 min time limit.

	Galapagos GA	Opossum RBFOpt	Opossum-CMAES	Optimus jEDE	Octopus SPEA 2
F(x) <sub>min</sub> _AVG	6.75E+06	5.28E+06	1.41E+07	<b>73.47</b>	<b>82.74</b>
FES	2601	136	76	<b>5000</b>	<b>3186</b>
Time (s)	1200	1200	1200	<b>105</b>	1200

Having run multiple tests on various optimization tools, Optimus jEDE proved the fastest, carried out the maximum average number of fitness evaluations 5000 (FES) in only 105 seconds, and achieved the lowest average EUJ 73.47 kW h /m2 yr ( $F(x)_{\min\_AVG}$ ) within the constraints. Octopus SPEA 2 on the other hand, performed second best while allowing for a good understanding of the tradeoffs, exhibiting the wide spectrum of optimal design solutions with the use of its useful graphical user interface (GUI).

To better review, the results, the unsuccessful and successful optimization tools were further grouped in independent graphs while showing the same information from the summary graph above.

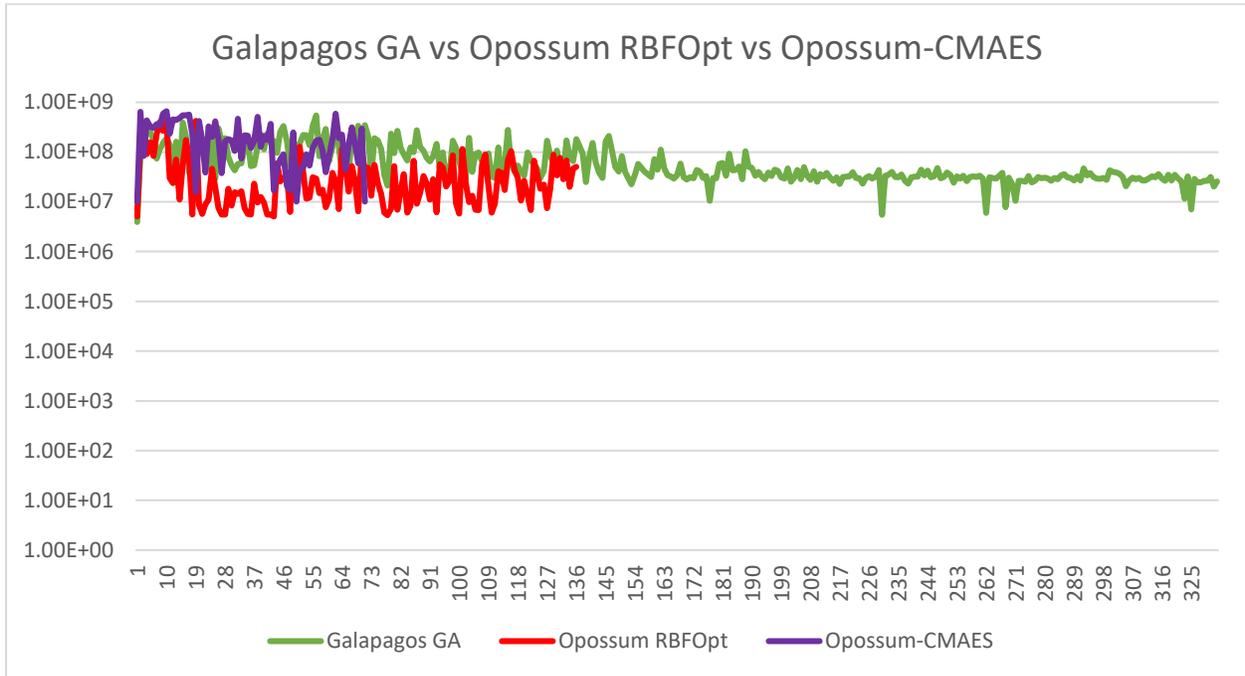


Figure 52 Convergence Graph- Summary of 3 Unsuccessful Algorithms: FES in 20 minutes (source: own)

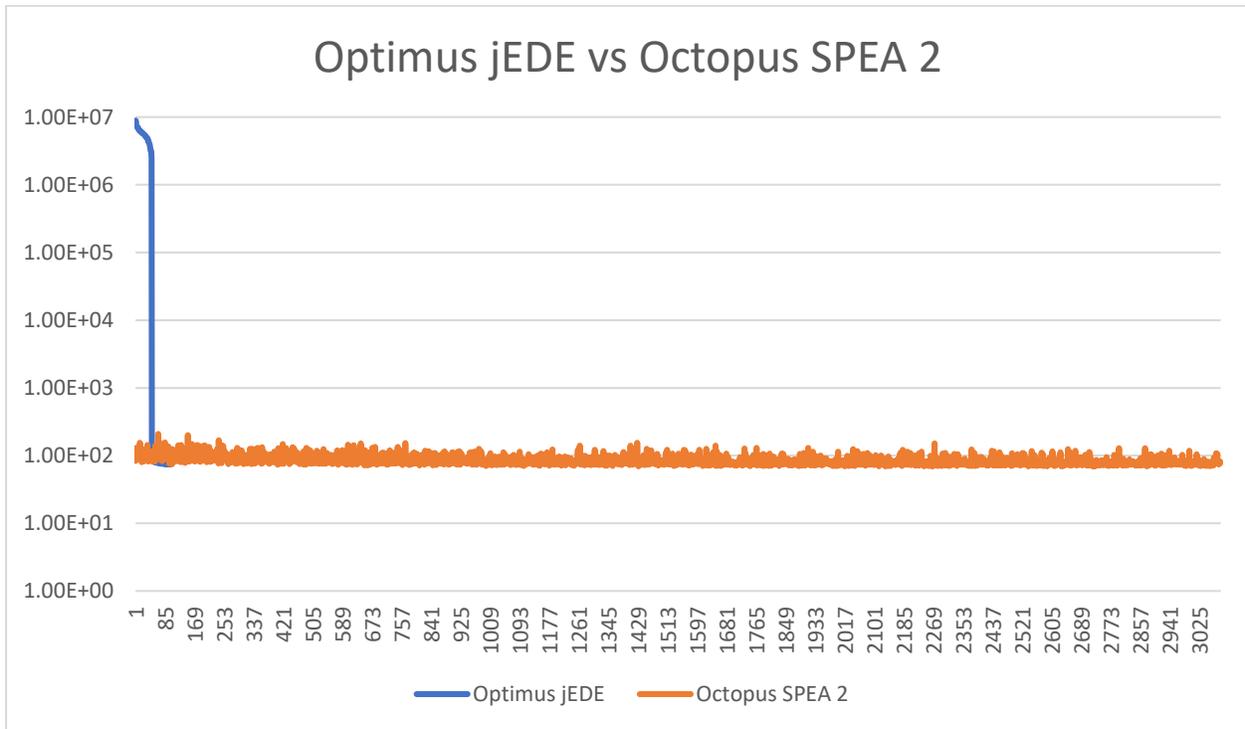


Figure 53 Convergence Graph- Summary of 2 Successful Algorithms: FES in 20 minutes (source: own)

### 3.7.3 Time saved

In terms of computational load, the speed of the optimization using the surrogate models was significantly improved when compared to the original BPS model. Further, when the parametric geometry generation was turned off, the results were calculated immediately, therefore allowing for multiple optimization runs and tests with 5 different tools.

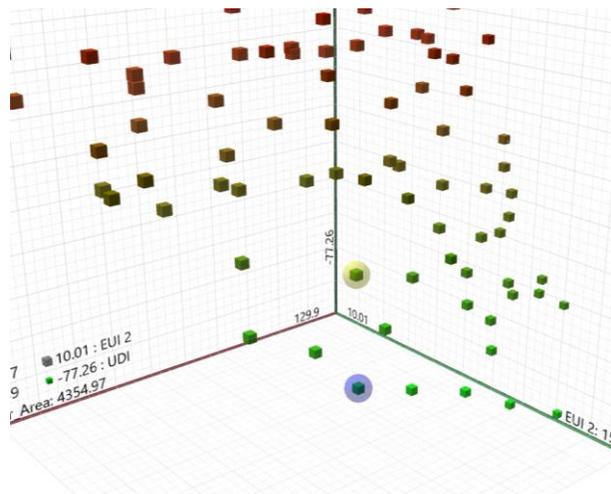
The newly trained surrogate models took only a fraction of the original time. While for ladybug simulation model was able to simulate approximately 3.4 simulations per hour (147 hrs total for 500 samples), the surrogate model took around 30 seconds per simulation, including the form generation. Therefore, rerunning an entire optimization of 500 samples would take approximately 4 hrs. and 15 minutes. The simulation time for 500 samples went from 8820 minutes (100%) down to 250 minutes (2.8%), this is a 97.2% reduction of the computational load. Even better, by turning off the form generation aspect of the script, the surrogate model was even faster. Using this method, 500 samples were generated in Optimus in 189 seconds (0.036%) meaning that it could generate an option in around 0.4 seconds; thus reducing the total computational load by 99.96%.

### 3.7.4 Amsterdam Vs Bogotá Multi-Objective Optimization

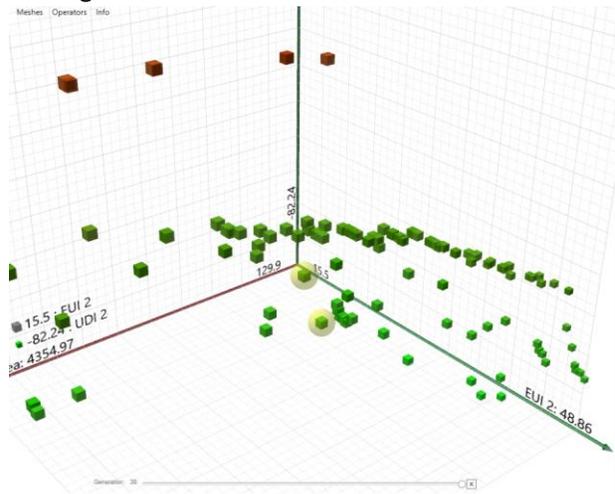
Before this case study, a quick study was carried out to compare two of the three metamodel locations, for both Amsterdam and Bogotá, a three objective optimization also using the Octopus SPEA 2 Reduction algorithm was performed. The parameters explored were: EUI, UDI, and Floor area.

Ultimately two significant designs were chosen to make observations & conclusions about the characteristics of the design solution for each location. The designs were chosen for their balance of EUI, UDI, and floor area: closest to the utopian objective but also for the notable conclusions that could be inferred from their architectural parameters.

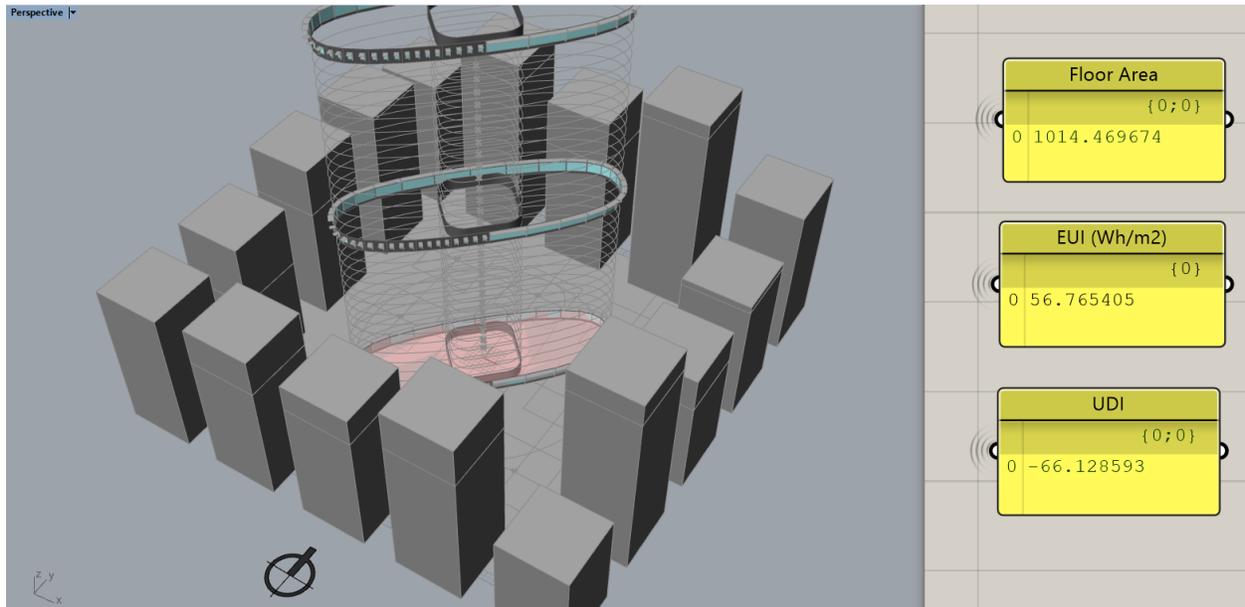
Amsterdam



Bogota



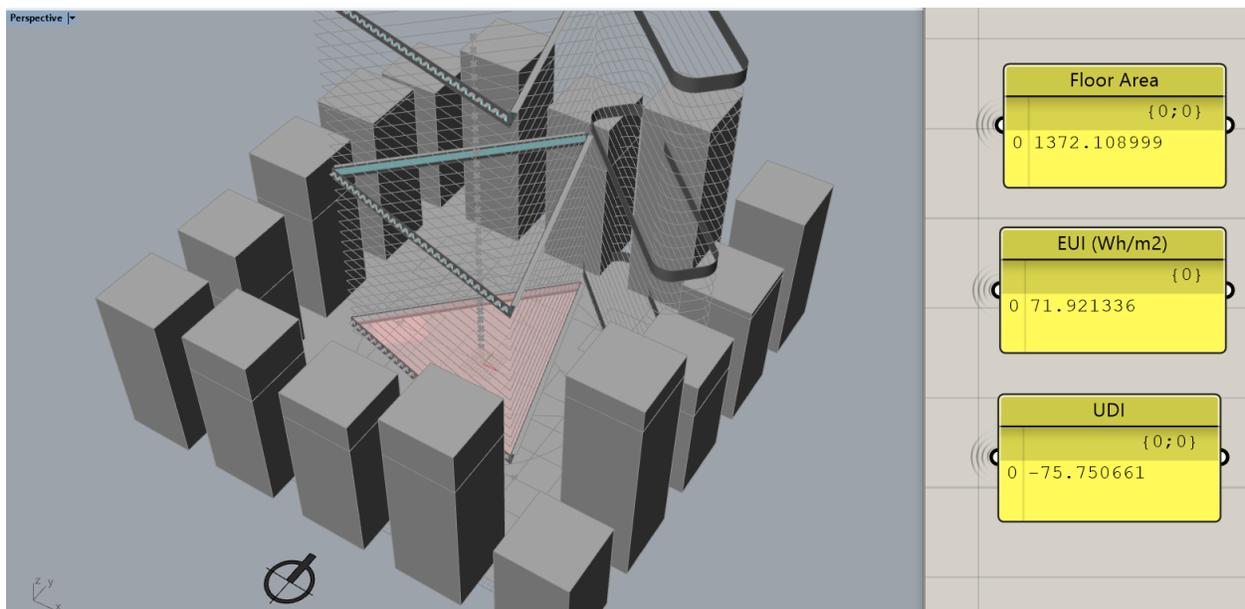
## Amsterdam Solution 1



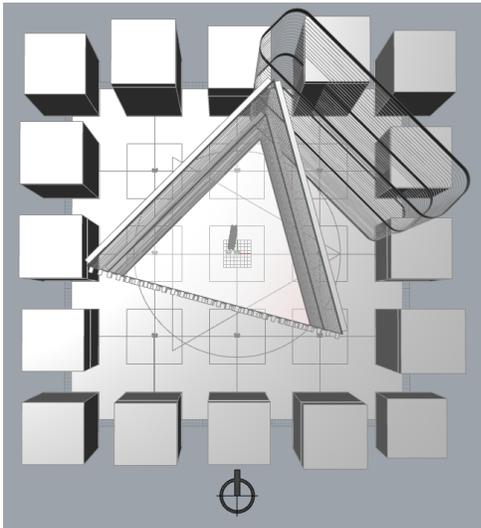
The first chosen Pareto optimal solution for Amsterdam shows the following interesting characteristics:

1. The analysis level is lowest at N00, meaning that an urban context could indeed help reduce energy costs while not being too counterproductive on UDI.
2. The optimization found a way to orient 1 side of the building to the south and reduce the WWR on that side to protect it from potential high cooling energy demand during the summer.
3. The rest of the façade orientations are set to the highest WWR (80%) thus improving UDI.
4. Large overhangs are preferred to protect from the high sun path that could cause overheating during summer.

## Amsterdam Solution 2

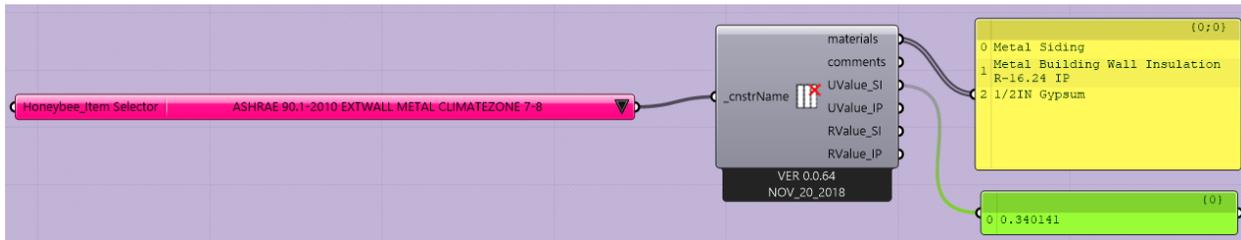


This solution is quite interesting as it gives relatively good results with a very unexpected shape:



1. The triangle has its southern oriented side with a low WWR (20%) to shield it from direct sun.
2. The northwestern façade is proposed at WWR 60% as a balance between and eastern facades are
3. It is suspected that the core could be acting as a reflection surface and redirecting indirect UDI into the floorplate, thus improving the UDI while not reducing significantly the EUI from direct sunlight penetration.

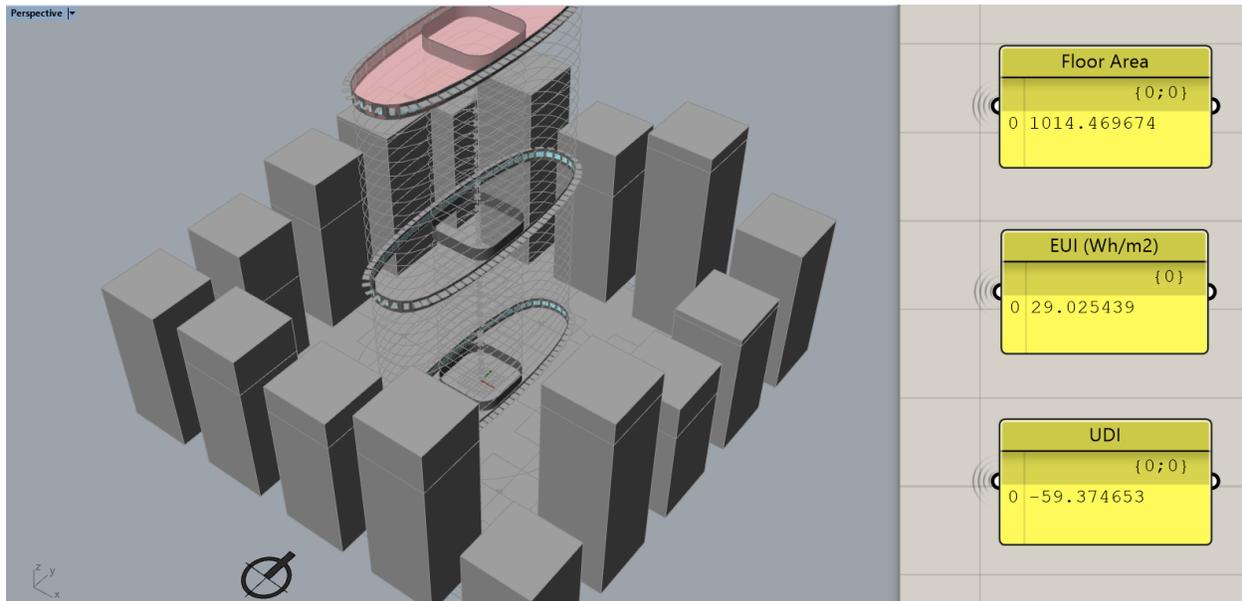
To check this hypothesis, the original simulation model was reviewed. Upon inspection, the wall type chosen for the lateral variant of the core walls was effectively an ASHRAE 90.1-201 Exterior Wall **Metal**, thus reflective. This was initially chosen to mimic stainless steel finishes that are seen on multiple high-rise solutions. This material was also chosen due to its similar U value with a wall in DesignBuilder during the validation process, yet its reflectance was not considered at the time. This further reinforced the theory of the reflecting core effect.



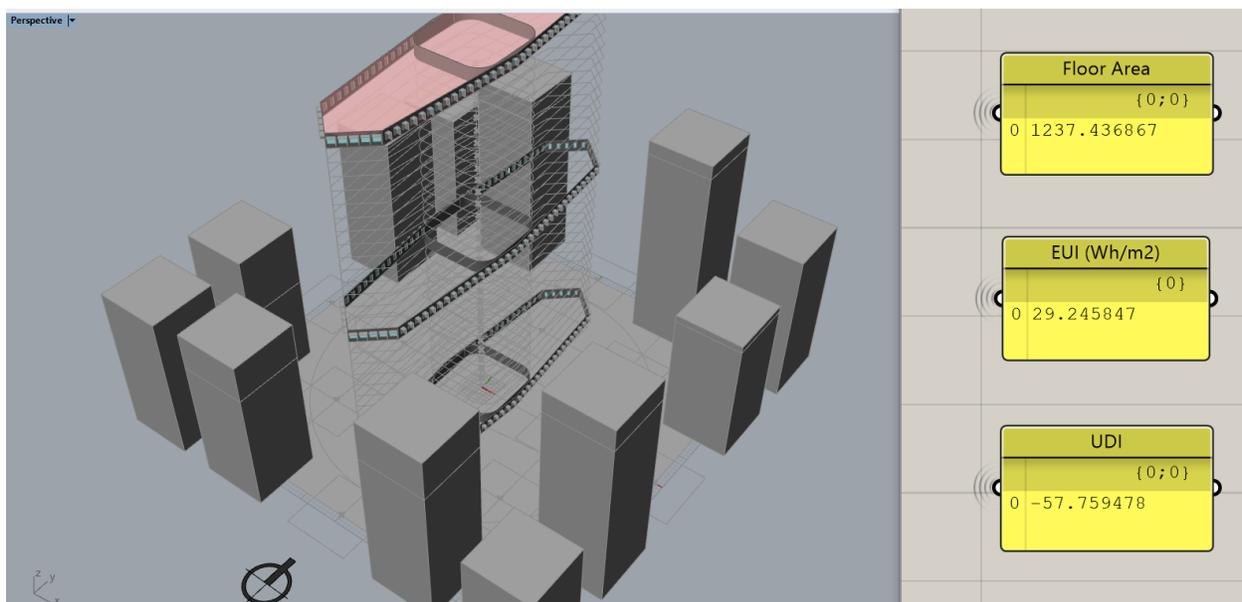
Characteristics for both chosen Amsterdam Solutions:

1. VLT is kept at a middle range of 60%
2. SHGC is kept to the highest to increase solar radiation gains that would be ideal for heating the building during the winter, thus reducing heating loads.

## Bogotá Solution 1



## Bogotá Solution 2



In the case of Pareto optimal options for Bogotá, both solutions are similar variants in shape, one is the ellipsoid and the second one an octagon, with very similar energy and daylight performance yet the intrinsic efficiency of the ellipsoid achieved approximately +200 more squared meters of area. Other important characteristics of these two solutions are:

1. The analysis level chosen is the highest, meaning that for Bogota context shading is ideally avoided, this could be due to the vertical sun path of this location.
2. The orientation of the floorplan is rotated 90° from the north as the eastern and western facades allow for high UDI while needing to be protected from direct sunlight.

3. For this reason, the East and West facades required
4. Ideal overhang size is somewhere in the middle: 0,8 m
5. Fins do not change considerably the EUI or UDI, therefore, are not necessary.

### General conclusions for both locations:

Because these solutions are searching for a balance for UDI and EUI and floor area, they will not fit the general recommendations from the table from chapter 8.1 that discusses solely the ideal shapes and orientations for energy efficiency.

1. Although Floor to Floor Height is important to improve UDI, it seems to not be worth increasing when considering EUI because all solutions opt for the lowest FFH.
2. A narrow plan of 80 x 20 m or 4:1 ratio is ideal to substantially improve UDI even though it would also increase energy consumption when compared to a 1:1 ratio.
3. A low façade U-value is always ideal for reducing energy consumption

### Initial Comparison to Baseline Model

Yet another advantage of exploring the whole spectrum of optimal designs was that solutions comparable to the control model used for calibration with DB could be performed. A Pareto front solution with a similar area as the original model (400) was chosen and compared. Once again, the triangular/lateral core solution seemed ideal. This solution also made sense, because the southern oriented façade had reduced WWR while the other two facades opened (WWR) 80% to allow for indirect to improve UDI; also similar to Amsterdam solution 2 form the previous point.

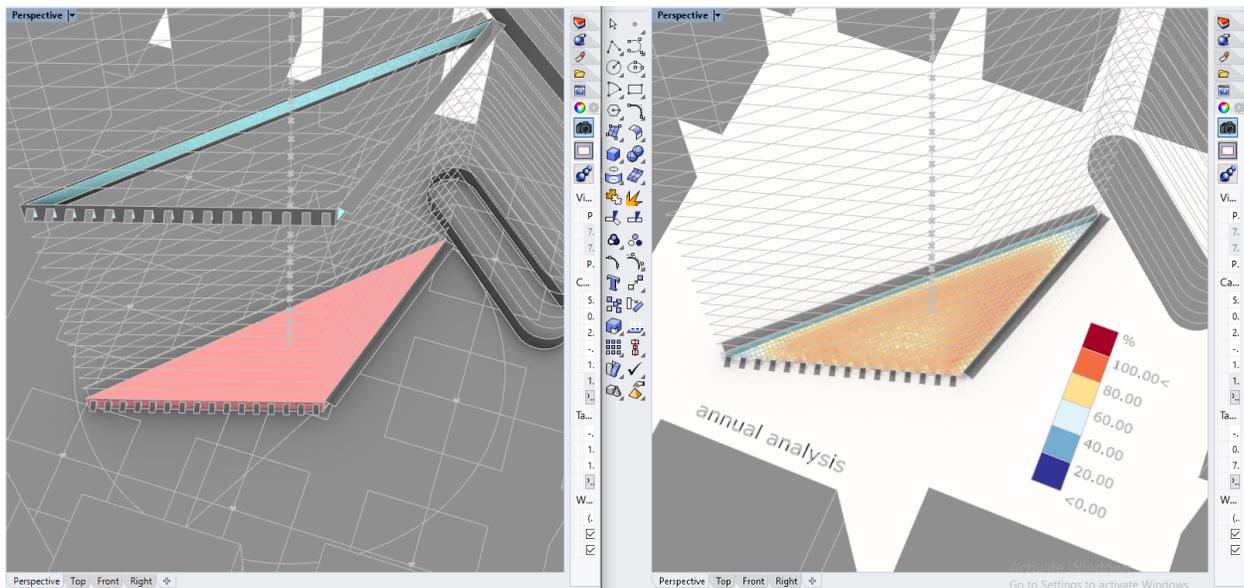


Figure 54 Surrogate model vs Simulated model of predicted improvement (source: own)

To further confirm the validity of this option, the original parametric model was updated to these inputs and then simulated again through Ladybug/Honeybee (Mackey & Sadeghipour, 2019). The results from the simulation were in line with what was predicted by the surrogate model with some caveats.

Table L

	EUI	Floor_Area	UDI
<b>Control</b>	95.41	400	40.081677
Opt 422 (Surrogate)	91.123	422.18	77.02
<b>Opt 422 (Simulation)</b>	107.41	422.18	72.47
% Difference (Sur)	4.49	5.55	92.16
<b>% Difference (Sim)</b>	-12.58	5.55	80.81

This table compares three values:

1. The **control model**'s original simulated results without any optimization
2. The **Surrogate model** predicted results of an optimized solution
3. The **Simulated model** of the optimized solution found through with the surrogate model

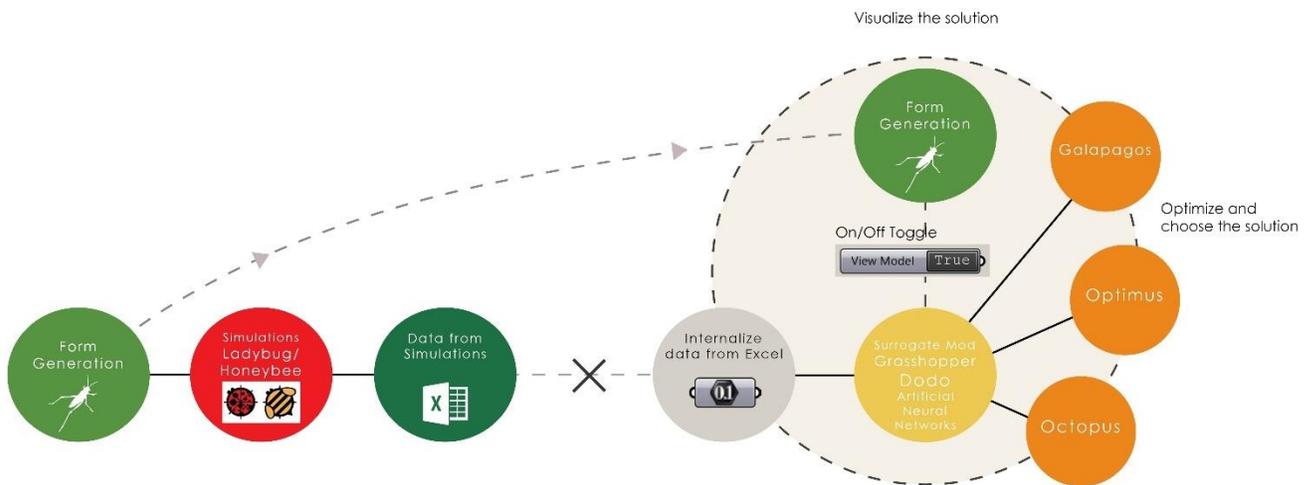
These results show that the surrogate model effectively served to locate a more optimal solution, nonetheless an **Error** in the calculation is evident. It predicted a lower EUI of 91.1 107.4 k Wh/m<sup>2</sup> when in fact the simulation gave a higher 107.4 k Wh/m<sup>2</sup>. This could be explained because the surrogate model for EUI was the most basic type of machine learning method: linear regression with an Adjusted R<sup>2</sup> of 84%, Standard Error of 9.74, and MAE of 7.28.

In contrast, the predicted result for the UDI was much closer to the simulated result: 77.0 Predicted vs 72.5 simulated. This can be explained because, for UDI, a more complex and precise ANN meta-model was used.

Ultimately, what this means is that the surrogate model predicted improvements of ≈5% for EUI and 92% for UDI while the simulated version showed a decrease of ≈12% for EUI and an improvement of ≈80% for UDI . This measurement was made early in the design process and the ANN-based metamodel had not been finely finetuned, and even though the model served to locate a better solution, it is critical to review the chosen design again with traditional simulation tools to avoid misinterpreting the data. Also, calibrating the validations metrics from the start is critical for the trustworthy results of the metamodel.

### 3.8 Surrogate model in the Design Process

The final step in the proposed methodology of this thesis permits the designer an effective use of the surrogate model by allowing quick visualization of the results. As described briefly in the introduction to the Method, the stripped-down initial Grasshopper parametric model was merged with the validated Surrogate model to permit visualization of the geometry and its predictions simultaneously. For the optimization runs, the geometry could be turned off to generate quicker results; once an interesting solution was identified the geometry could be turned back on for further inspection. This is particularly valuable because the human operator is not interested in the whole design space that the algorithm is searching but rather only in ideal Pareto-optimal solutions. This improves the already enhanced process of using meta-models in the design process because it takes advantage of the speed of the computer for the optimization but also keeps the interactive component that is required by the designer or architect.



This image illustrates the evolution of the tool. Once the form generation segment collected the 500 samples through the **simulation software**, this could then be internalized into a model coupled with the final surrogate model. The coupled model is composed of 3 main parts: The **form generation**, the **surrogate model**, and the **optimization tools**. The link between variables and the surrogate model could be activated and deactivated at will.

The data from any of the simulated locations (Amsterdam, Bogota, and Dubai) could be read from the excel files and used to train the surrogate model. Once the model was trained, the variables could be entered manually, slider by slider, or could be read like a list set with the 16 independent variables x1..x16.

This method thus incorporates the multi-objective optimization tool to accelerate the iteration process but also serves as an efficient design tool for decision-making in the early stages of design because of the possibilities it allows for quick design exploration. Moreover, although the surrogate model has shown advantage by incorporating it into a single and multi-objective optimization process, this tool can also be a very useful as a standalone tool that allows the designer to explore iteratively the consequences of changing one variable at a time.

### 3.9 Discussion

From a pure engineering standpoint, it could be argued that single and multi-objective optimizations are completely different problems therefore they should not be compared. Nonetheless, the comparison between the results from Optimus jEDE and Octopus SPEA2 served as a validation that indeed the minimal EUI was around 70 kW h /m<sup>2</sup> yr. for Amsterdam. Differently than from the engineering realm, in the architectural optimization realm, it is useful to have various methods to solve a unique problem.

Optimus jEDE proved to be by far the fastest optimization tool, therefore it is recommended as an initial step to run many quick simulations and get a general idea of the Minimal or Maximal values of an architectural solution. Likewise, its light software architecture never seemed to slow down the computer in which it was being run, lightening the computational load considerably. Even the script could be reviewed simultaneously, although not recommended. This makes Optimus jEDE an ideal tool for early-stage design when time is of the essence and quick optimizations are critical. Although Galapagos GA and Opossum RBFOpt or MAES might be efficient tools in other optimization problems, for this one they proved slow and ineffective, taking up time and computational calculations, slowing the computer down considerably and eventually not being able to find a feasible solution to the fitness function. Although, the possibility still exists that after longer runs these SOO algorithms could eventually converge, but because the whole point of doing this thesis is to lighten the computational burden, extremely long computational optimization runs would be counterproductive.

Therefore, my recommendation for SOO problems is to first try Optimus jEDE and later the other options. On the other hand, its nature to solve a single objective limits its applicability because most architectural solutions should hardly be reduced to a single objective but rather be reviewed with an overarching assessment in a range of options.

For the reasons mentioned above, was deemed that using a Multi-Objective Optimization tool could be more useful in this multi-criteria scenario. For this exercise, being able to visualize the 3-Dimensional design space allowed not only to identify good solutions but also to understand better the compromises that had to be made. The Octopus Multi-Objective visual tool exposed more clearly that good energy efficiency (lower EUI) signified an inherent tradeoff to good daylight (lower UDI), therefore it was more useful to view the design space as a large set of possible solutions rather than one.

Although the SPEA 2 algorithm was left to run for 20 minutes and produced the second-highest FES, ( $\approx 3200$  rather than 5000 from jEDE) it is important to note that by using the MOO tool from Octopus, it was possible to get a general overview of the lowest and highest range values for each output. This meant that as soon as the optimization was started, the 3-Dimensional graph was aware of the general overview of ideal values. As more non-dominated solutions were found, the minimal fitness functions were further refined and decreased or increased slightly, yet the overall ranges were easily identified from the start.

At a point during the research, modeFRONTIER (ESTECO, 2020) was used as the optimization tool to find good performing designs, and although it provides many optimization algorithms that could be extremely useful for many design problems, the process of reviewing the results and visualizing them proved to be someone time-consuming. For this, a segment of the custom-script was dedicated to reading properly the output data form ModeFrontier. This back and forth, exporting and importing was possible but not ideal. Valuable time was lost in creating this custom component and later identifying the design options that resulted in the optimal design solution. An alternative to this limitation could be for the software developer ESTECO, to develop a custom Grasshopper component that could read the choices of design options in real-time within the Grasshopper Environment, thus facilitating greatly the design exploration aspect.

In conclusion, my recommendation for more complex architectural problems that manage multiple objectives is to use Octopus all within the Grasshopper environment. Currently, it is limited to the following optimization Algorithgms: SPEA 2, HypE Reduction, Polynomial Mutate, Alt Polyn. Mutate and HypE Mutation (Vierlinger, 2018). Further optimization algorithms would ideally be added to the Octopus library, making it a powerful tool in architectural design optimization.

## Comfort Error Output

During the process, by having performed a statistical significance F test, I was able to identify an error with one of the outputs. Initially, the value for F-significance of the comfort outputs was 0, meaning that the prediction was not statistically significant. This led to an investigation of the original simulation model. The solution was identified as a simple error in reading the list: rather than reading the comfort inside the core area, the script was mistakenly reading the comfort inside the core area, meaning that none of the inputs were significantly affecting the comfort output. This was corrected in the original simulation model and the samples were collected once more.



Figure 55 Comfort Error Core vs Floorplate (source: own)

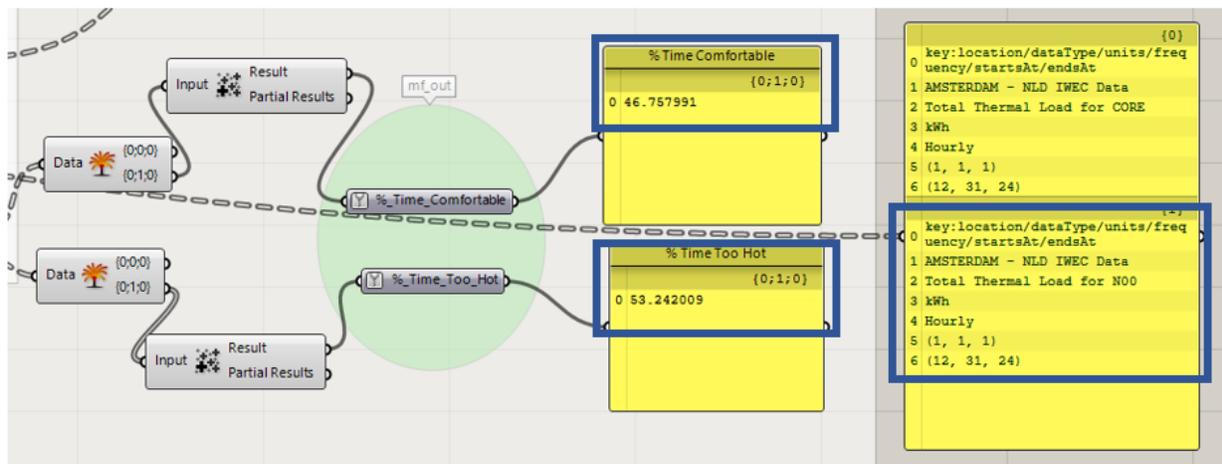


Figure 56 Comfort Solution Core vs Floorplate (source: own)

## 4.0 Showcase of a design solution

### 4.4 Proposal

Considering the wide spectrum of options that the optimization tools could generate, it was a daunting task to choose a single architectural solution. The final solution was solution 2 from Amsterdam vs Bogota Multi-Objective comparison study. It was chosen because even though it lay on the Pareto-frontier line, was a very atypical architectural solution compared to the expected ones. It was chosen because of this nonconforming image while still showcasing that it was an efficient option in the three objectives searched on this thesis paper. This uncommon solution also exemplified the power of a computational methodology that could suggest uncharacteristic solutions that would have otherwise never have been explored by the human designer.

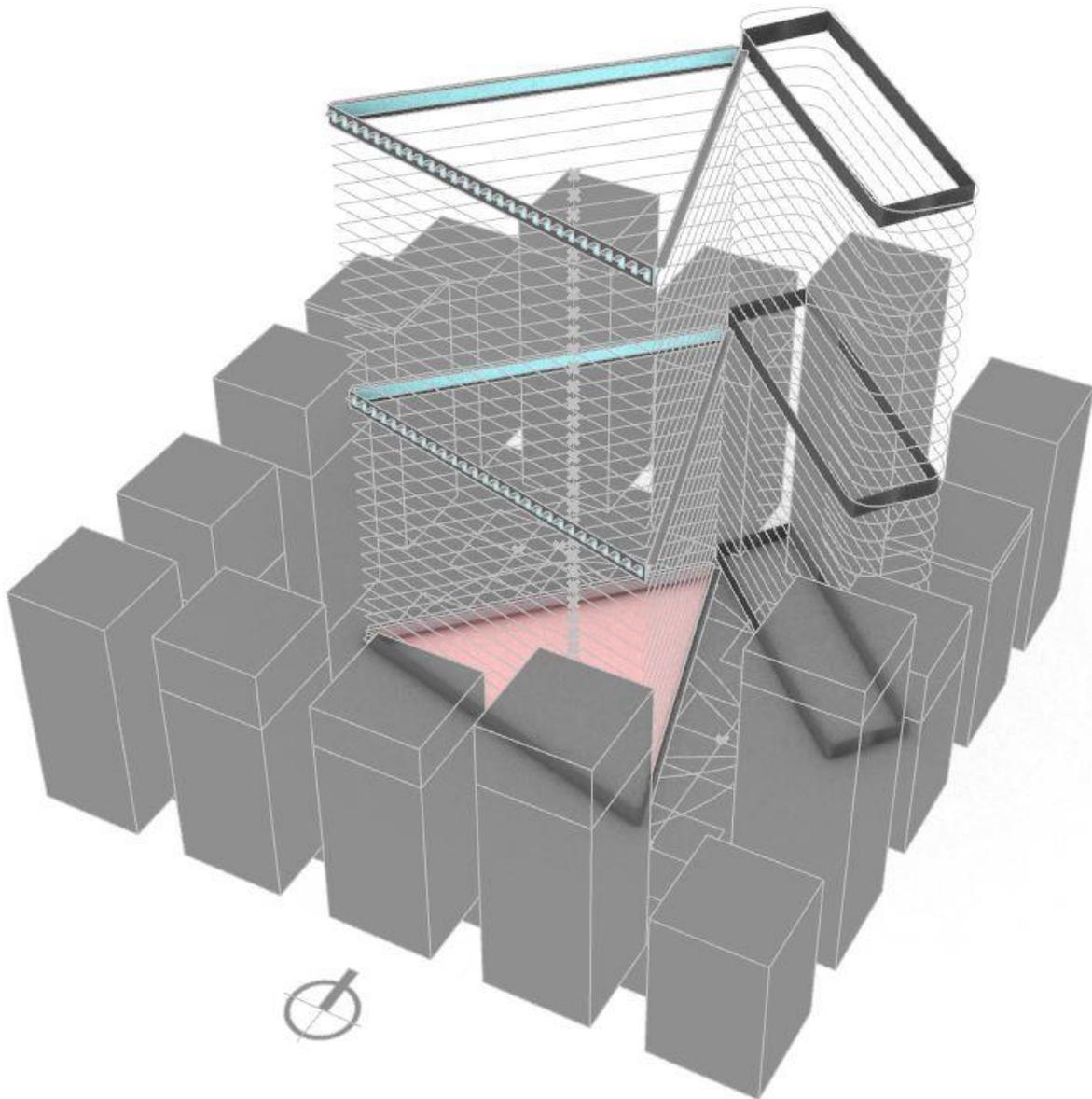


Figure 57 Final solution to showcase the design (source: own)

To double-check the accuracy of this option, a final simulation was performed with the original BPS simulation script in Grasshopper. Although it did not have the highest UDI as compared with other optimal solutions, it indeed showed that it was a valid option to develop into a complete architectural solution. Further refinements, like adding an Atrium in the center, or service-based program that does not require daylight, could counter the lower UDI towards the center of this proposal.

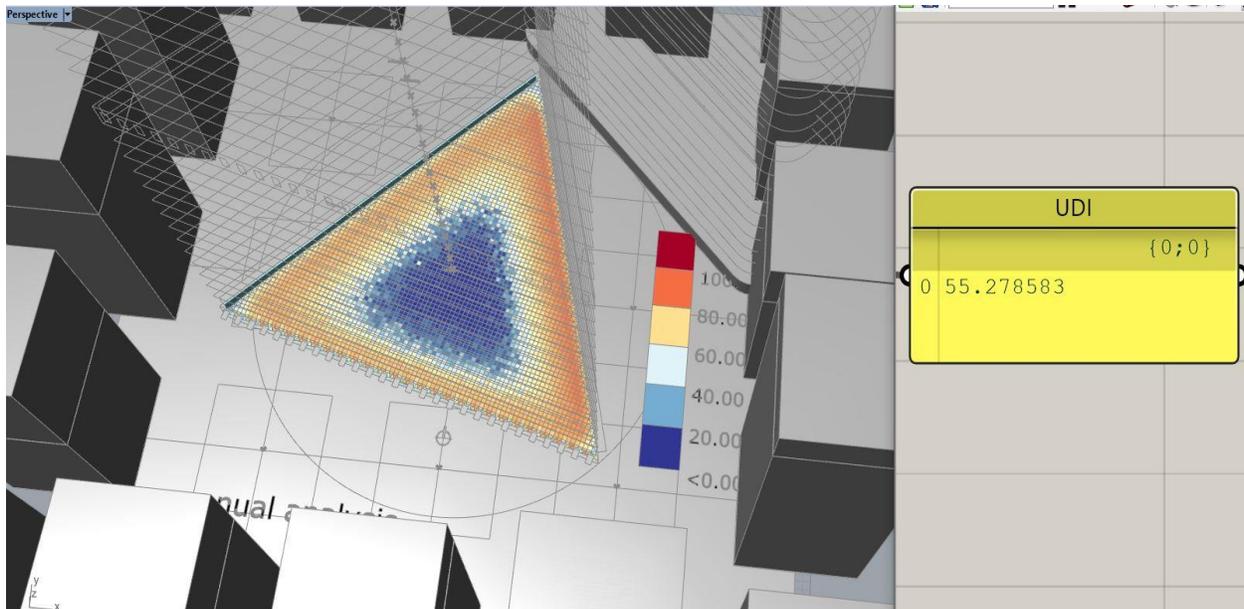


Figure 58 PBS Simulation of the final design solution (source: own)

### Final Architectural Solution

	Inputs
Analysis_Level	2
Core_Type	0
FFH	6
Length	3
Number_Of_Fins	0
Orientation	7
Overhang_Size	3
SHGC	3
Shape	0
U_Value	1
VLT	2
WWR_East_North	3
WWR_North_West	3
WWR_South_East	1
WWR_West_South	4
Width	3

<b>EUI</b>	85.97
<b>Comfort</b>	56.41
<b>UDI</b>	61.84
<b>Floor Area</b>	1703.4

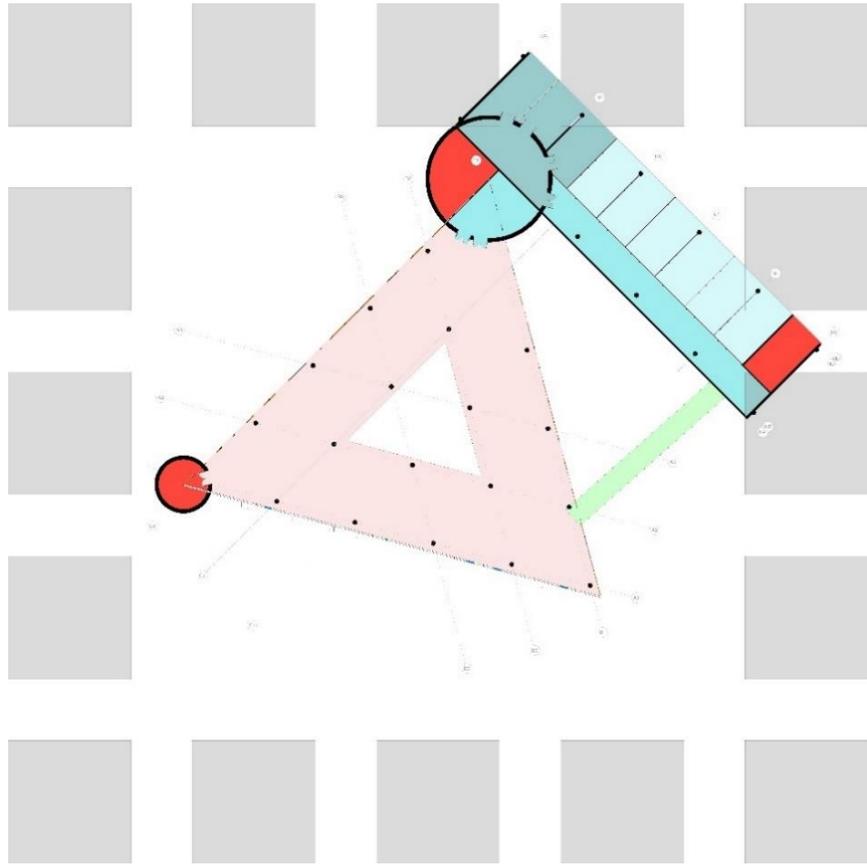


Figure 59 Diagrammatic planview of chosen design solution (source: own)

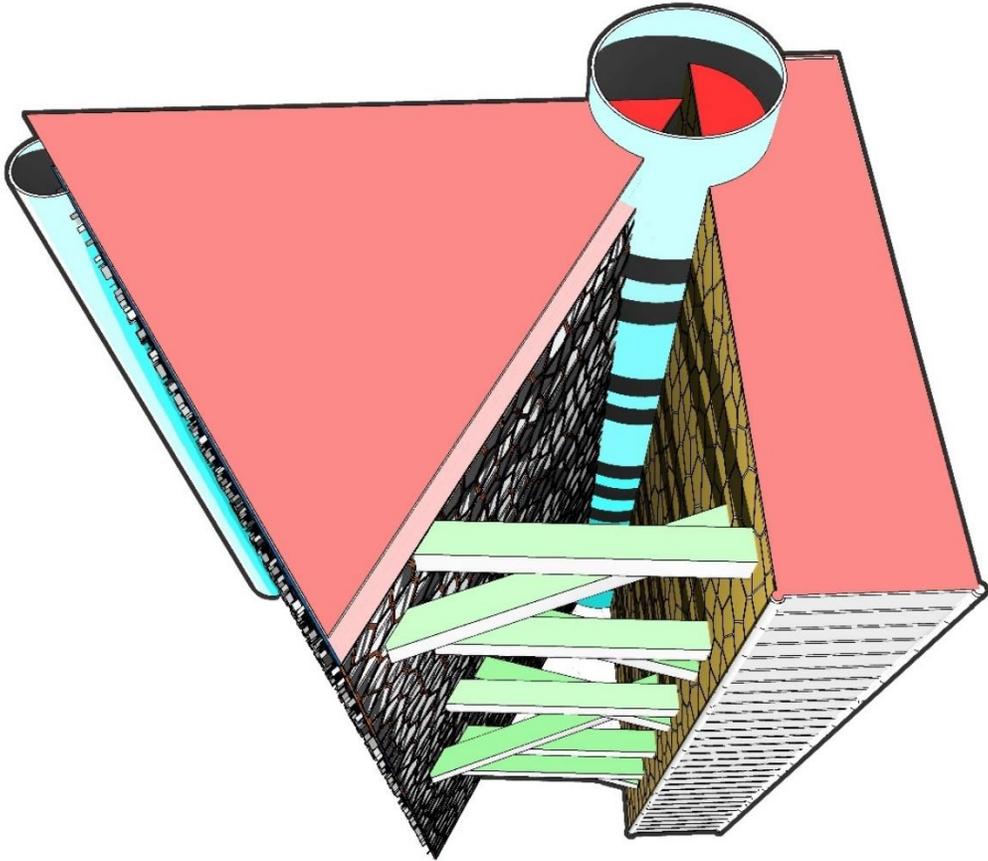


Figure 60 Top view of chosen design solution (source: own)

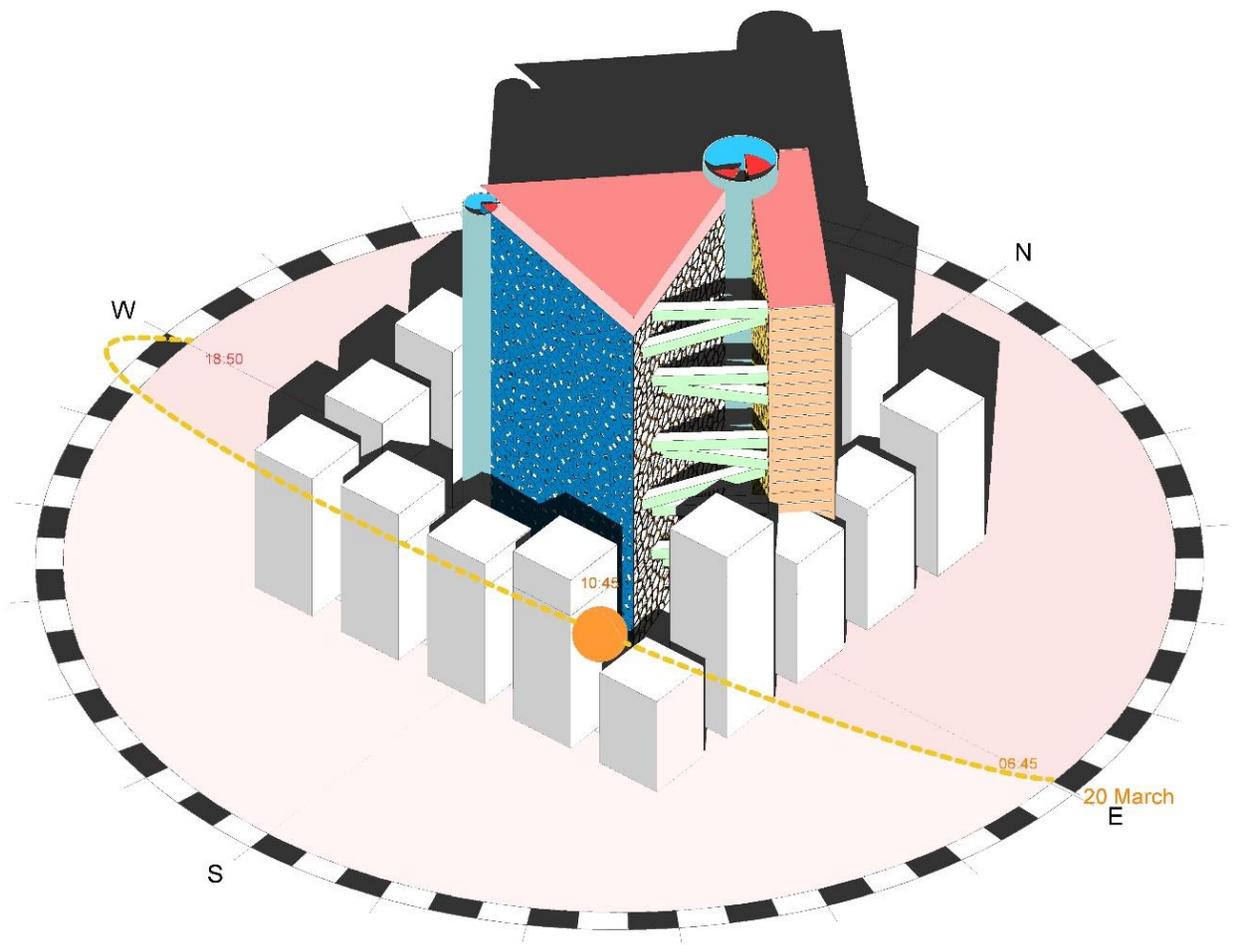


Figure 61 Axonometric diagram of chosen design solution (source: own)

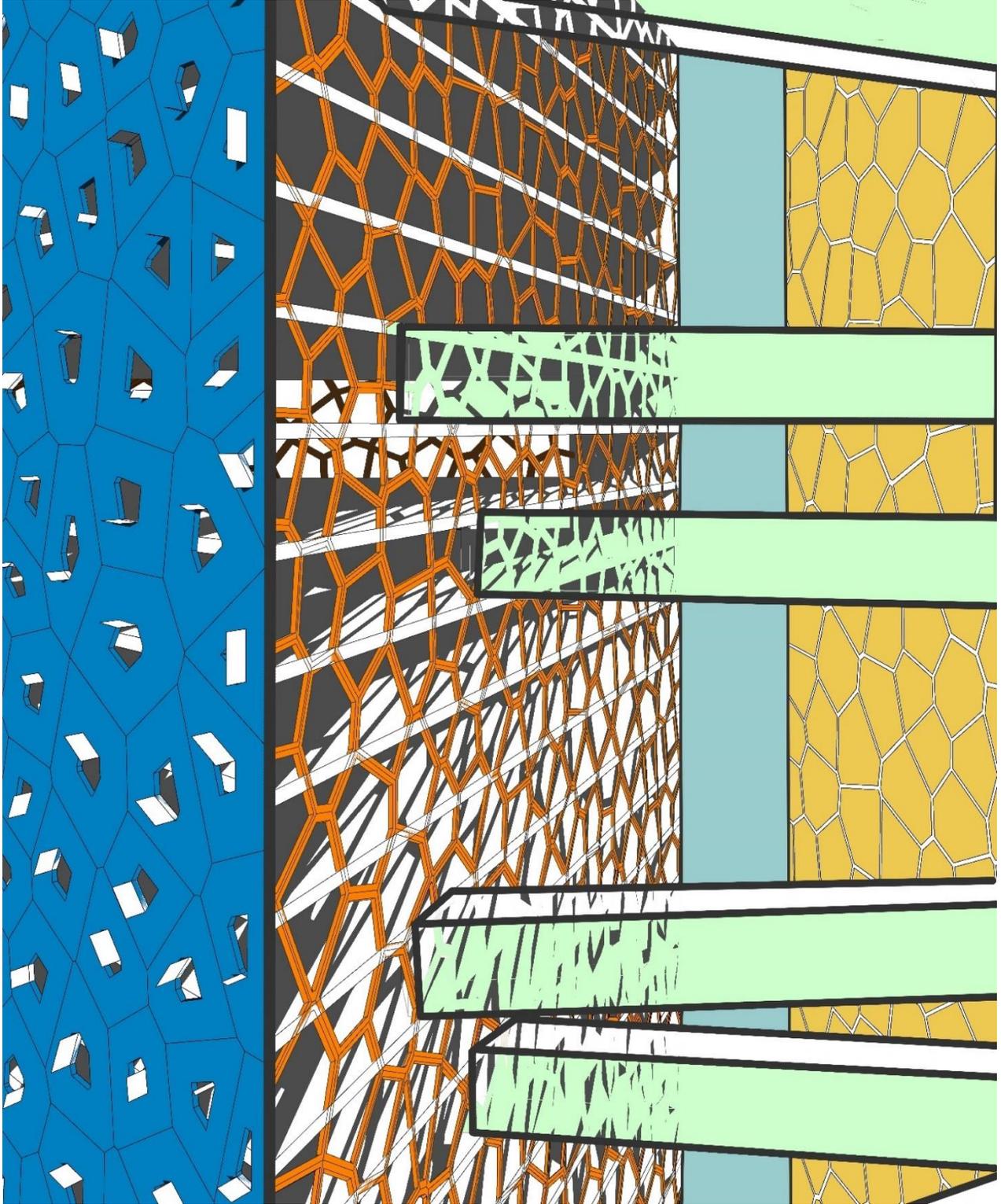
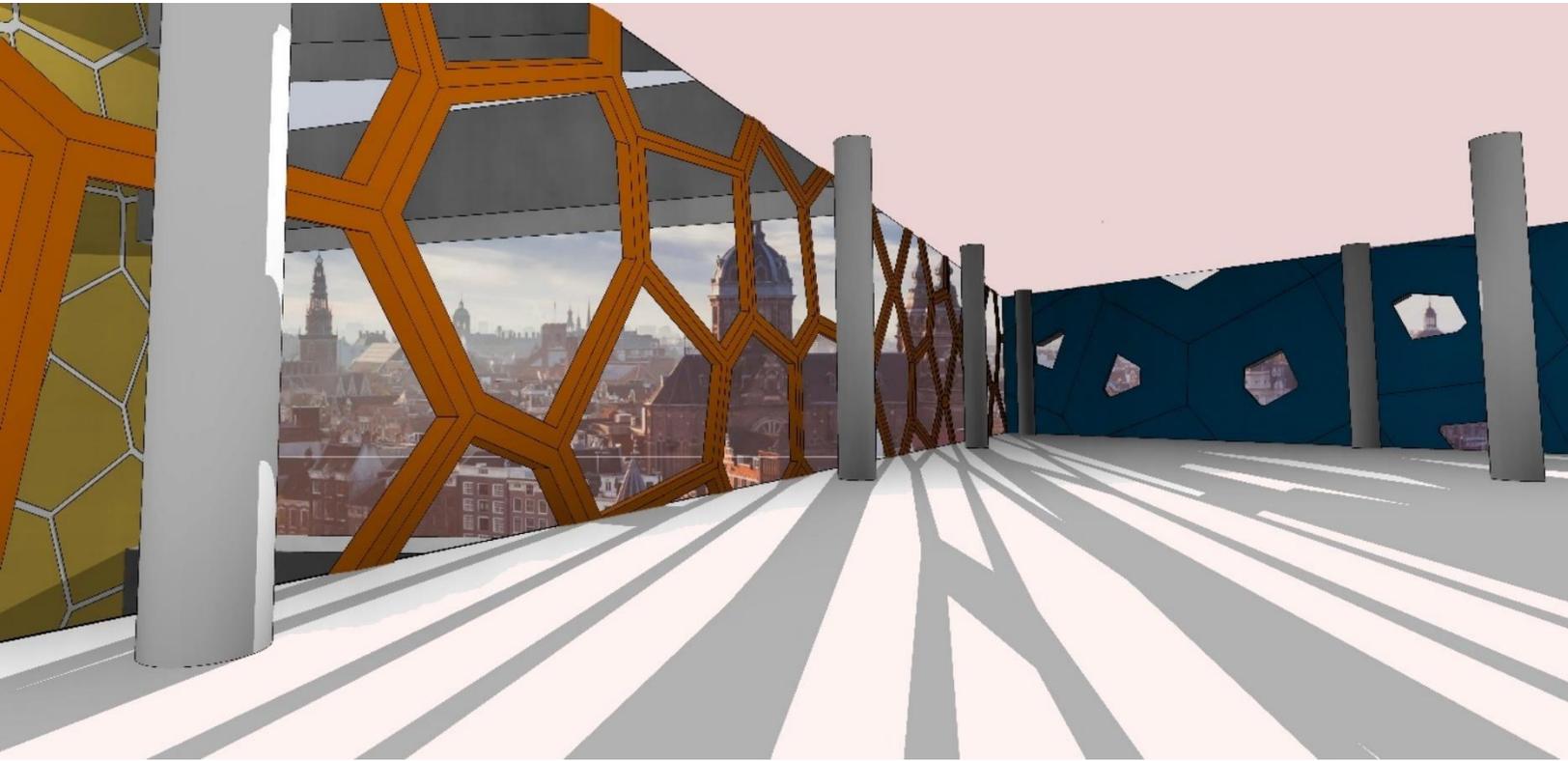


Figure 62Diagram "bridges" of chosen design solution (source: own)



*Figure 63 Interior concept view of chosen design solution (source: own)*

## 5.0 Conclusions

The initial research question of this paper was:

*“How can a computational method using surrogate modeling be used to quickly identify, and optimize the most influential factors and their combinations for context-based passive solutions of sustainable High-Rise office buildings during their initial design phase?”*

The main research question of this paper was to understand how a computational method could be used to review how shape and façade variables would affect building energy, daylight, and comfort performance on a high-building typology. For this, research was conducted on creating surrogate models using various machine learning methods. The methods used in this paper were Linear Regression, Non-Linear Regression, and Artificial Neural Networks. Ultimately an ANN-based surrogate model was developed for the three different locations, Amsterdam, Bogotá, and Dubai. Later, as a proof of concept, the Amsterdam-based ANN surrogate model was used to review various optimization algorithms, including Evolutionary Algorithm, RBFOP, CMAES, jEDE, and SPEA2. Using Surrogate model-based optimization, the speed of reviewing 500 samples was reduced from 8820 minutes (147 hrs / 6.125 days) down to 189 seconds using jEDE, the fastest performing algorithm reviewed. This accounted for a time saving of up to 99.96%.

Likewise, this paper presented a new optimization workflow which considered the following steps in the architectural design process:

1. Problem definition
2. Simulation of the Parametric Modeling (PM) / Form Generation
3. Building Performance Simulations (BPS) / Performance assessment
4. Validation of Simulation Model in Design Builder
5. Data Collection based on the sampling strategy
6. Training of the Surrogate model
7. Evaluation of Surrogate model
8. Computational Optimization (CO)
9. Surrogate model in Design Process
10. A showcase of a design Solution

The final steps 9 and 10 show how the early-stage design exploration phase can be coupled using a custom-made script that includes not only the surrogate model and the optimization algorithms but also the visualization components to allow for easy exploration of the design solutions and their individual EUI / Comfort / UDI and Floor area outputs.

The main objective in terms of the building physics aspect of this research was to view the effects of orientation, shape, core, context, and façade parameters and their effects on Thermal Comfort, Energy Consumption, and Natural daylight of a building.

The sub-question for this paper where:

- a. How do different locations/climates dictate the building's ideal shape and orientation?

As expected from the initial literature review, when only seeking to lower the EUI, the ideal shape of the floorplan of a high-rise tended to be the ellipsoid, because a circle has the shortest perimeter related to its area. In contrast, when EUI / Comfort and UDI were simultaneously optimized, the other options for the building shape (triangle, square, and octagon) were also "optimal", laying also on the Pareto-frontier, particularly when balanced correctly with the WWR for each of the building's orientation.

In the sensitivity analysis, the orientation seemed to have a very low effect on the outputs. This can be explained because the width and length can essentially change the "orientation" by 90° simply by changing the ratio of these two inputs; for example, (length 60 m x width 20 m) would essentially be same as (length 20 m width 60 m) with a 90° orientation. Likewise, the orientation was modeled in large steps of 45°; more subtle variations could show a more important role of this parameter.

Although the Pareto front discovered optimal cylindrical floorplan solutions. The ideal Amsterdam orientation tended to converge to 0° from the center bearing in mind-independent WWR solutions were defined for the southern direct sunlight facade (low WWR) versus the northern façade (High WWR). On the contrary, the cylindrical floorplan was rotated 90° for the Bogota location to account for the vertical distribution of the sun path throughout the year (no seasons). These eastern/western facades, however, required low WWR to limit excessive energy gains from the sun.

Finally, with the lateral core iterations, the orientation of the building was crucial. In some optimal solutions for Amsterdam, the core served as a shield or even a reflective element that improved the UDI considerably by allowing for higher than expected wall to window ratios.

- b. How does the volumetric context of surrounding buildings affect its shape and façade parameters?

The context inputs were refined into a single variable that allowed to view the effect of high, low or medium context buildings related to the studied geometry, this input was labeled Analysis Level and noted the outputs depending on the level: Ground floor (N00), Central floor (N15) or Top Floor (N30). Further parameters of block size, block separation, and proximity of the blocks were created as parametric variables, yet they were not included in this research to simplify the number of variables.

During the initial simulation runs using a non-optimized ANN-based surrogate model, the low ground analysis level seemed to be the optimal solution, meaning that indeed the urban context was helping improve the sought-after objectives of this paper. Nonetheless, when the ANN was improved and run through both jEDE and SPEA2 optimization algorithms with 10 runs each, it showed that the preferred solution for almost all Pareto-optimal designs was the topmost analysis level, meaning that context was playing a detrimental effect to the objectives. When the lowest analysis level was chosen, the urban context did indeed slightly lower the EUI (example: from 89.6 to 85.9 (kWh/m<sup>2</sup> yr) but also reduced the Thermal comfort (from 68.2% to 56.4%) and the UDI (from 64.1 to 61.8). So, although the context plays a role in the outcome, its effects are negatively affecting the overall performance of the optimized Amsterdam Location.

- c. How does the position of its Core affect a building's energy performance concerning its location/climate?

Regarding the core, it was discovered that its position played a significant role in the Useful Daylight Illuminance of the building. Interestingly, the core gave rise to some unexpected Pareto optimal solutions such as a triangular floor plate in Amsterdam with the core being suspected to serve as a sort of daylight reflection element. Also, some solutions used the core as a shield from the most unwanted southern sun, allowing at the same time to increase the WWR on the southern façade. Moreover, on the central core variants with smaller areas, such as the control option, its location within the floorplate decreased the UDI by acting as a barrier within, therefore the core size concerning its usable floor plate should be carefully considered when designing a high-rise.

- d. What is the Pareto-optimality of these parameters when simultaneously seeking energy consumption, thermal comfort, and natural daylight optimums?

The Pareto optimality was reviewed thoroughly in the MOO section of this paper. In conclusion, many design options with widely different architectural expressions all fall within this Pareto-optimality. It is therefore up to the designer to choose the one that best suits its intention and ideal Floorplate area. For that, the Octopus Tool is recommended because of its ease-of-use and efficiency of its SPEA 2 optimization algorithm.

- e. What are the key validation metrics and features necessary for obtaining suitable surrogate models?

To create precise surrogate models, it was critical to understand the validation metrics that define any Machine Learning method. Once these metrics, such as  $R^2$ , P-Value, Standard Error (RMSE), or Mean Absolute Errors (MAE) could be efficiently calculated and reviewed, the definition of an ideal meta-model was more easily attainable. Although this process still needed optimization or a manual iterative process to perfect, the procedure became much more streamlined once a general "feel" for improving the ML methods was developed. To do this, each of the hyperparameters was tuned gradually until the validation metrics were within acceptable margins.

The key validation metrics for obtaining a surrogate model were explained in the literature review section of this paper and later applied for the training of the surrogate model. The metrics are  $R^2$ , Adjusted  $R^2$ , MAE, RMSE (standard error), and the P-Values. To show the correlation between the input variables and the outputs a commonly used metric is the  $R^2$  and better yet, the adjusted  $R^2$ ; which accounts for the bias by considering the number of variables thus reducing slightly the  $R^2$  (Frost, 2020). The MAE and RMSE are metrics to predict the deviation of the predictions from the results, they are used to validate the accuracy of the model. Finally, the P-Values are used to review the statistical significance of the results, if they do not

The examination of individual p-values of the inputs uncovered more subtleties in understanding how each parameter was affecting the dependent variables (outputs) based on the chosen geographic locations: Amsterdam, Bogotá, or Dubai.

## **Conclusions On Validation**

To validate the collected samples from the simulations, a separate building physics software was utilized: DesignBuilder. (DesignBuilder, 2020) Through the comparison of the results from the original parametric simulation model with the results from DesignBuilder, a more global understanding of the variables that affected the Energy Balance was achieved. Likewise, it ensured that the collected samples were more trustworthy.

Daylight parameters were carefully finetuned to find a proper balance between the long simulation time versus an acceptable daylight result. This, in turn, was reviewed in conjunction with the energy simulation to achieve results similar to those from the DesignBuilder software.

It is also vital to note the importance of the occupancy and lighting schedules for the total energy performance of a building, this was an essential step in achieving similar results.

Yet another important definition that was identified the simulation model validation was natural ventilation. The results from both models varied significantly when natural ventilation was introduced in the energy calculation, for this reason, this aspect was carefully fine-tuned to achieve congruent results between both models.

## **On Envelope Parameters**

Firstly, by reviewing independent P-values, it is concluded that although the envelope parameters are important, the shape parameters play a much more important role on the optimized outputs, thus reinforcing the idea that good information for good initial design is critical in the design process; finetuning the envelope parameters only should come after a good general design is found. On that note, the most important parameter from the envelope perspective on energy performance and natural daylight was the WWR.

The presented optimized options offer solutions that apply to the entirety of the building, ideally, though each façade orientation would have optimized characteristics to respond in the most precise manner for each orientation and building floor height. This would have to be modeled so that each façade parameter (and not only the WWR) responds independently from the other.

The optimization presented here consists solely of passive design solutions, active solutions would always improve results as they can adapt to the ever-changing daylight or seasons, therefore these recommendations are the simple first step in initial design stages rather than finalized set of values.

## On Computational method

Although the Honeybee/Ladybug (Sadeghipour & Mackey, 2017-2020) components in Grasshopper (Rutten, 2015) using the EnergyPlus Engine (DOE, BTO, NREL, 2020) for daylight and energy simulation are extremely powerful tools for building analysis, the initial results from the runs must always be taken with a highly critical approach. Once the model was set up, the inputs and assumptions from the software were carefully fine-tuned with the DesignBuilder (DesignBuilder, 2020) software to achieve a model with more reliable results.

Three independent Surrogate models were developed: one for Amsterdam, Bogotá, and Dubai. Although both models for Amsterdam and Bogotá have the same Köppen Classification of (Cfb) Marine West Coast Climate, results show that their optimal solutions are similar in some respects such as shape and FFH, yet different in others such as context (analysis level), orientation or WWR.

Three independent models were trained because it made them easier to train. By trying to train 3 simultaneous locations in a single ANN model, the data outputs would vary greatly from location to location, meaning that the ANN would have to be more complex. This complexity was deemed unnecessary although theoretically it could be done with enough time to train the model and samples to ensure that the R squared values and MAE, MSE is within acceptable margins. It is also worth noting that once trained, this location variable should not be optimized by any of the optimization algorithms, the variable would instead serve as a choice for the designer to review their building in a different location.

These surrogate models offer the advantage of a design space with which the designer can experiment with. Traditional simulation-based optimization usually relies on Genetic algorithms or other nature-inspired algorithms to find more “ideal” or fitting solutions. Nonetheless, the nature of the problem is much more complex. By identifying the P-values and Correlations of the variables to the outputs the designer can see to which choices he should give more importance to. By identifying the points that fall within the Pareto optimality of the design space the designer can also recognize patterns and their inherent tradeoffs by contradicting objectives. The response surface represents a more complete design space on which the designer can immediately experiment.

Similarly, even though surrogate models can take many simulations to set up, once they are calculated and validated as more optimization experiments are required, they become gradually more attractive, especially when the entire design cycle is considered. (Eisenhower, et al., 2012) For this reason, an organized and easily modifiable original parametric model is key, so if a new surrogate model must be derived, quick changes or improvements could then open an easy way for a new batch of collected samples.

The advantage of metamodeling is that once it is generated, many additional options can be explored, and single or multiple objective optimization algorithms can be utilized without spending time on new computationally expensive daylight or energy simulations. Moreover, some optimized solutions can broaden the spectrum of intelligent design possibilities by presenting options that would have otherwise never have been imagined by the designer; a perfect example being the triangular, lateral core showcase example for Amsterdam.

Nevertheless, although surrogate models can serve as efficient tools for exploring either the design space or finding optimal design options, once these solutions have been identified and will be further developed, they should be verified to ensure that the surrogate model has indeed predicted the results correctly. For this reason, Validation metrics are crucial to use surrogate models with more confidence.

The computational goal of the research: to develop surrogate models to aid the architect in understating their design choices, this objective was achieved by creating two custom surrogate models for Amsterdam and Bogota that reflect how a high-rise building would behave at these locations. These models can be interactively explored with virtually no waiting time when only the inputs and outputs are reviewed. When the model is intended for output review as well as for visualization, the waiting time increases to approximately 5-6 seconds which is the time it takes to generate the new parametric geometry.

## 6.0 Reflection

### **Background on Artificial Intelligence**

Artificial Intelligence (AI) is already permeating all levels of human organization: medicine, education, military, from artistic endeavors such as surrealist art to things as mundane as Netflix/Social media recommendations (Knight, 2019)

#### *Healthcare:*

Researches from Pennsylvania Healthcare provided AI with 1.77 million electrocardiograms (ECG) logs for it to detect patterns. AI predicted better than the existing techniques in which patients were at greater risk of premature death. This means that “AI can potentially teach us things that we’ve been maybe misinterpreting for decades.” (Tangermann, 2019)

#### *Aviation:*

So far, Airbus has used generative design to develop a 3D-printed “bionic partition” for airplane cabins that is 45% lighter yet 8% stronger than anything it’s used to date. (Wilson, 2019) Bionic Partition (Nagy, 2016). Computational Intelligence can help us design unconventional and more efficient options.

#### *Design:*

Philip Stark partnered with Kartell to design the first commercial product with the help of an algorithm. They developed an experimental algorithm called Deep style that even seeks to mimic a company’s design ethos (Schwab, 2019)

AI could reinforce our design ethos, although it should not replace human thinking but rather serve as a tool to aid our design process to make more conscious and informed decisions. There should always be a collaboration loop between the computational methods and the designer (Knight, 2019)

Technology Transfer is the process of using techniques or material from one creative field or industry into another (Pawley, 1990) We need this technology transfer from the computational field and into architectural design.

Therefore, as a general scope, we could use computational intelligence to eliminate our human bias, speed our comprehension and problem-solving skills, and enrichen our capabilities as architectural designers.

### **Application to the built Environment**

Today, the evolution of High-Tech: Neo-futurism with examples such as Zaha Hadid Architects or UN Studio now deeply rely on computational modeling tools to achieve flowing shapes and aesthetically interesting morphologies. Yet simply complex computational shapes do not reflect the full potential of computational tools.

The next step is to improve the parameters that constitute the shapes and characteristics to achieve solutions that minimize or maximized ideal criteria, such as: reduce energy performance, improve structural performance, daylight, wind flow, or increase thermal comfort for the building’s users. This opens a great field of study known as Parametric Optimization.

Nonetheless, this exciting new field comes with limitations. First, the computationally heavy simulations require much time to achieve results, therefore these simulations usually are not used to understand the direct effects on the building's performance. Secondly, the parametric models are usually project/site-specific, therefore the results extracted from the optimization can only be applied to that case.

The third application of the computational realm to the Built Environment in the field of Artificial intelligence or Computational Intelligence. Surrogate models represent a branch of this field, they intend to solve the first limitation of parametric optimization, which is to reduce the computational cost.

## 6.1 Main topics

The two main categories for research conducted were: Building Physics and Computational Methods, under these the subcategories are as follows:

### Building Physics

- Climate Zones
- Energy Efficiency and Standards
- Energy Modeling
- Building Performance Parameters
- Comfort Metrics
- Daylight Analysis

### Computational:

- Grasshopper interface & Use
- Ladybug & Honeybee Software for Energy Modeling
- DesignBuilder Software for Energy Modeling
- ModeFrontier Software for data collection & Sensitivity Analysis
- Surrogate Modeling

## 6.2 Aim

The general aim of this thesis is not solely to identify the best performing parameters but rather to help designers understand the effect of their choices on energy performance, daylight, and comfort results.

Therefore, the specific aim is the development of a tool, in the form of a responsive surrogate model, for architects to play with and subsequently develop a better intuition of their choices. It is not intended to model closely a specific site, but rather to represent a generalizable surrogate model of a high-rise office building. This is an exploration of the possibilities of surrogate modeling applied to the built environment. Ideally, a generic surrogate model would fully replace building simulation tools for the most common types of building projects and locations.

### 6.3 Limitations

Energy simulation depends greatly not only on building-related variables but also on its use by its occupants, this simulation considers only building energy and would fall short in predicting this second human variable. Likewise, the HVAC system affects immensely the energy consumption of a building, the variables studied here focuses solely not the architectural element and choices and not those done by the MEP engineers.

Simulating building energy, daylight and comfort is a widely used approach in parametric optimization is a widely used approach despite that it is not possible to guarantee the perfect representation of real-world occurrences. Although these simulations provide a good indication of the underlying percentages of results and trends. The methodology, nevertheless, remains valid.

This Parametric / Surrogate model is meant to be used as a broad tool by allowing a designer to “play” to understand the general effects of 16 parameters on the 4 objectives, it is not however meant to be used as a precise tool to predict the energy performance, daylight or comfort for a particular building in a particular location of either Amsterdam, Bogotá, and Dubai.

Regarding the collected data for each location, it must be clarified that the validation of the simulation model between Ladybug/Honeybee and DesignBuilder was performed with 1 design option in 1 location (Amsterdam) It was then assumed that it would correctly simulate the other solutions in any of the two other locations: Bogotá and Dubai by using the EPW files from Ladybug tools.

### 6.4 Further Steps

A further step for this research could be to run similar MOO optimizations both for Bogotá and Dubai locations. Amsterdam and Bogotá solutions where compared yet further exploration amongst these three locations could show interesting conclusions from the Building Physics standpoint. Due to time limitations, only Amsterdam optimization was explored in depth.

Additional steps could also be to collect more samples from other locations to be compared with the Amsterdam, Bogota, and Dubai models. Although through the development of this thesis, the predictability of the ANN-based surrogate model proved sufficiently good, on rare occasions, for predicting the UDI, the R squared value was slightly above 100%, making it unreliable. It is suggested that to improve this, more samples for each location could be collected since literature shows that more samples could lead to better accuracy of the surrogate model.

Although the workflow for creating surrogate models was investigated, little research was done on the number of samples required to achieve a good model. This thesis used Uniform Latin Hypercube (ULH) static sampling method yet investigation on the adaptive sampling methods show potential from the literature review and could be an interesting point for further investigation.

In the conclusions, it was mentioned that three independent metamodels were developed for each of the three locations. A further step could be to try to use the same dataset of 1500 samples to train a single ANN-based model.

## 7.0 Bibliography

NREL, ANL, LBNL, ORNL & PNNL, 2020. *OpenStudio*. [Online]

Available at: <https://www.openstudio.net/>

Robert McNeel & Associates, 2020. *Rhino.Inside*. [Online]

Available at: <https://www.rhino3d.com/inside>

[Accessed 30 06 2020].

Advanced Buildings, 2020. *Daylighting Pattern Guide*. [Online]

Available at: <https://patternguide.advancedbuildings.net/using-this-guide/analysis-methods/daylight-factor>

[Accessed 29 01 2020].

Anderson, A., 2020. *How to Calculate the Adjusted Coefficient of Determination*. [Online]

Available at: <https://www.dummies.com/education/math/business-statistics/how-to-calculate-the-adjusted-coefficient-of-determination/>

[Accessed 20 04 2020].

Auger, A., Bader, J., Brockhoff, D. & Zitzler, E., 2012. Hypervolume-based multiobjective optimization: Theoretical foundations and practical implications. *Theoretical Computer Science*, Volume 425, pp. 75-103.

Autodesk, 2020. *Revit Architecture*. [Online]

Available at: <https://www.autodesk.com/products/revit/overview>

[Accessed 30 05 2020].

Brea, F., Romana, N. & Fachinotti, V. D., 2020. An efficient metamodel-based method to carry out multi/objective building performance optimizations. *Energy and Buildings*.

Buchanan, P., 1998. In: *Reinventing the Skyscraper*. s.l.:A and U. no 329,, pp. 30-67.

Cenaero, 2018. *proximedia*. [Online]

Available at: <http://www.cenaero.be/>

[Accessed 20 02 2020].

Chen, X. & Yang, H., 2017. A multi-stage optimization of passively designed high-rise residential buildings in multiple building operation scenarios. *Applied Energy*, Volume 206, pp. 541-557.

Chen, X., Yang, H. & Sun, K., 2016. A holistic passive design approach to optimize indoor environmental quality of a typical residential building in Hong Kong. *Energy*, Volume 113, pp. 267-281.

Cubukcuoglu, C., Ekici, B., Tasgetiren, M. F. & Sariyildiz, S., 2019. Optimus: Self-Adaptive Differential Evolution with Ensemble of Mutation Strategies for Grasshopper Algorithmic Modeling. *Algorithms*, Volume 12.

CWCT, 2000. *Centre for Window and Cladding Technology*. [Online]

Available at: <https://www.cwct.co.uk/performance/thermal.htm>

[Accessed 30 01 2020].

Daniels, M., 2019. *Human Terrain*. [Online]

Available at: [https://pudding.cool/2018/10/city\\_3d/](https://pudding.cool/2018/10/city_3d/)

[Accessed 02 02 2020].

Dekking, M., 2005. *A modern introduction to probability and statistics: understanding why and how..* London: Springer.

DesignBuilder, 2020. *DesignBuilder Version 5.5.2.007*. [Online]  
Available at: <https://designbuilder.co.uk/helpv5.5/>  
[Accessed 25 03 2020].

Dobbelsteen, A. v. d., 2012. High-rise Buildings: A Contribution to Sustainable Construction in the City?. In: D. / . M. H. Zandbelt, ed. *High-Rise and the Sustainable City*. Amsterdam: Techne Press, pp. 120-147.

Dobbelsteen, A. v. d., Thijssen, S., Colello, V. & Metz, T., 2007. *Ecology of Building Geometry - Environmental Performance of Different Building Shapes*. s.l., CIB World Building Congress.

DOE, BTO, NREL, 2020. *EnergyPlus 9.3.0*. [Online]  
Available at: <https://energyplus.net/>  
[Accessed 01 04 2020].

Efficient Windows Collaborative, 2000-2018. *Efficient Windows Collaborative*. [Online]  
Available at: <https://www.efficientwindows.org/>  
[Accessed 25 01 2020].

Efficient Windows, 2011-2020. *Windows for high-performance commercial buildings*. [Online]  
Available at: <https://www.commercialwindows.org/index.php>  
[Accessed 10 01 2020].

Efficient Windows, 2019. *Energy Star Windows*. [Online]  
Available at: <https://www.efficientwindows.org/energystar.php>  
[Accessed 28 01 2020].

Eisenhower, B. et al., 2012. A methodology for meta-model based optimization in building energy models. *Energy and Buildings*, Volume 47, pp. 292-301.

Ekici, B., Cubukcuoglu, C., Turrin, M. & Sariyildiza, I. S., 2019. Performative computational architecture using swarm and evolutionary optimization: A review. *Building and Environment*, Volume 147, pp. 356-371.

Ekici, B. et al., 2019. *A Methodology for daylight optimization of high-rise buildings in the dense urban district using overhang length and glazing type variables with surrogate modeling*. s.l., s.n.

Elotefya, H., Abdelmagidab, K. S., Morghanya, E. & Ahmed, T. M., 2015. Energy-efficient Tall buildings design strategies: A holistic approach. *Energy Procedia* 74, pp. 1358-1369.

Emporis, 2020. *multi-story building (ESN 51477)*. [Online]  
Available at: <https://www.emporis.com/building/standard/66/multi-story-building>  
[Accessed 10 01 2020].

EnergyStar, 2020. *Energy Star*. [Online]  
Available at: <https://www.energystar.gov/buildings/facility-owners-and-managers/existing-buildings/use-portfolio-manager/understand-metrics/difference>  
[Accessed 25 03 2020].

ESTECO, 2020. *Getting Started with modeFRONTIER*, Trieste: s.n.

ESTECO, 2020. *ModeFRONTIER*. [Online]  
Available at: <https://www.esteco.com/modelfrontier>  
[Accessed 31 02 2020].

Evangelia Despoina Giouri, M. T. ., M. T., 2020. Zero energy potential of a high-rise office building in a Mediterranean climate. *Energy & Buildings*, Volume 209, p. 109666.

- Evins, R., Pointer, P. & Burgess, S., 2012. *MULTI-OBJECTIVE OPTIMISATION OF A MODULAR BUILDING*. Loughborough, UK, Copyright IBPSA-England,.
- Fong, Z. T. a. S., 2016. Chapter 9. In: *Survey of Meta-Heuristic Algorithms for Deep Learning*. s.l.:s.n.
- Frost, J., 2020. *Five Reasons Why Your R-squared can be Too High*. [Online]  
Available at: <https://statisticsbyjim.com/regression/r-squared-too-high/>  
[Accessed 01 07 2020].
- Garuda, S. S., Karimia, I. A. & Kraft, M., 2017. Design of computer experiments: A review. *Computers and Chemical Engineering*, Volume 106, pp. 71-95.
- Givoni, B., 1998. *Climate Considerations in Building and Urban Design*. New York: John Wiley & Sons, Inc.
- Glen, S., 2020. *F Value: Simple Definition and Interpretation*. [Online]  
Available at: <https://www.statisticshowto.com/probability-and-statistics/f-statistic-value-test>  
[Accessed 15 06 2020].
- Greco, L., 2020. *Dodo*. [Online]  
Available at: <https://www.food4rhino.com/app/dodo>  
[Accessed 31 05 2020].
- Hakanen, J., 2015. *Optimization Software*, s.l.: The University of Jyvaskyla.
- Harper, P., 2019. *Dezeen*. [Online]  
Available at: <https://www.dezeen.com/2019/11/21/high-tech-architecture-environment-sustainability/>  
[Accessed 26 11 2019].
- Humphreys, M. & Nicol, J., 2002. Adaptive thermal comfort and sustainable thermal standards for buildings. *Energy and Buildings*, Volume 34, pp. 563-572.
- IPTOES, 2010. [Online]  
Available at: <http://peakoiltaskforce.net>  
[Accessed 27 12 2019].
- Kheiri, F., 2018. A review of optimization methods applied in energy-efficient buildings. *Renewable and Sustainable Energy Reviews*, Volume 92, p. 897/920.
- Kheiri, F., 2018. A review of optimization methods applied in energy-efficient building geometry and envelope design. *Renewable and Sustainable Energy Reviews*, Volume 92, pp. 897-920.
- Kirimtata, A., Krejcar, O., Ekici, B. & Tasgetiren, M. F., 2019. Multi-objective energy and daylight optimization of amorphous shading devices in buildings. *Solar Energy*, Volume 185, pp. 100-111.
- Koehrsen, W., 2018. *Statistical Significance Explained*. [Online]  
Available at: <https://towardsdatascience.com/statistical-significance-hypothesis-testing-the-normal-curve-and-p-values-93274fa32687>  
[Accessed 15 05 2020].
- Köppen, W. / R. G., 2020. *Köppen climate classification*. [Online]  
Available at: <https://www.weatherbase.com/gr/koppen.png>  
[Accessed 31 01 2020].

- Levitt, B., 2015. *Honeybee Quickstart*. [Online]  
Available at: <https://www.youtube.com/watch?v=j9d6ldaDox8>  
[Accessed 15 02 2020].
- Lorenzo, 2019. *Dodo v.03*. [Online]  
Available at: <https://www.food4rhino.com/app/dodo>  
[Accessed 20 04 2020].
- Mackey, C. & Sadeghipour, R. M., 2019. *#3 Energy and Comfort Modeling for Passive Design*. [Online]  
Available at: <https://vimeo.com/ondemand/3hbenergy?autoplay=1>  
[Accessed 30 03 2020].
- Mardaljevic, A. N. a. J., 2005. Useful daylight illuminance: A new paradigm for assessing daylight in buildings. *Lighting Research and Technology*.
- Mardaljevic, J., 2011. *Ambient Calculation: Crash Course*. Berkely , International Radiance Workshop 2011.
- Miller, N., 2018. *LunchBox 2018.11.16*. [Online]  
Available at: <https://www.food4rhino.com/app/lunchbox>  
[Accessed 05 04 2020].
- Miller, N., 2019. *Free Generative Design – A brief overview of tools created by the Grasshopper community*. [Online]  
Available at: <https://provingground.io/2019/11/19/free-generative-design-a-brief-overview-of-tools-created-by-the-grasshopper-community/>  
[Accessed 01 05 2020].
- Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 2012. *Bouwbesluit 2012*. [Online]  
Available at: <https://rijksoverheid.bouwbesluit.com/>  
[Accessed 15 12 2019].
- Ministry of the Interior and Kingdom Relations, Netherlands, 2012. *Building Decree Online 2012*. [Online]  
Available at: <https://rijksoverheid.bouwbesluit.com/Inhoud>  
[Accessed 29 01 2020].
- MIT, 2012. *MIT Sustainable Design Lab / Daysim*. [Online]  
Available at: <https://github.com/MITSustainableDesignLab/Daysim>  
[Accessed 18 06 2020].
- Nojhan, 2007. *Classification of metaheuristics*. [Online]  
Available at: <http://metah.nojhan.net/post/2007/10/12/Classification-of-metaheuristics>  
[Accessed 20 05 2020].
- Oxford, 2020. *Lexico powered by Oxford*. [Online]  
Available at: <https://www.lexico.com/definition/overfitting>  
[Accessed 30 06 2020].
- Phan Anh Nguyen, R. B. A. v. d. D., 2019. Effects of a Vertical Green Façade on the Thermal Performance and Cooling Demand. *JOURNAL OF FACADE DESIGN & ENGINEERING*, 7(2), p. 0.
- Rahman, M. A., 2017. *ANT*. [Online]  
Available at: <https://github.com/MahmoudAbdelRahman/-ANT->  
[Accessed 01 04 2020].

Raji, B., Tenpierik, M. J. & Dobbelsteen, A. v. d., 2016. An assessment of energy-saving solutions for the envelope design of high-rise buildings in temperate climates. *Energy and Buildings*, Volume 124, pp. 210-221.

Raji, B., Tenpierik, M. J. & Dobbelsteen, A. v. d., 2017. Early-Stage Design Considerations for the Energy-Efficiency of High-Rise Office Buildings. *Sustainability*, 623(9).

Raof, B. Y., 2017. The correlation between Building Shape and Building Energy Performance. *International Journal of Advanced Research*, 5(5), pp. 552-561.

Rogers, S. H. +. P., 2014. *The Leadenhall Building*. [Online]  
Available at: <https://www.rsh-p.com/projects/the-leadenhall-building/>  
[Accessed 30 06 2020].

Romano, B., 2016. *Torre Reforma LBR&A Arquitectos*. [Online]  
Available at: <http://www.lbrarquitectos.com/portfolio/torre-reforma/>  
[Accessed 30 06 2020].

Rutten, D., 2015. *Grasshopper: Algorithmic Modeling for Rhino*. [Online]  
Available at: <https://www.grasshopper3d.com/>  
[Accessed 01 12 2019].

Rutten, D., 2019. *Galapagos*. [Online]  
Available at: <http://grasshopperdocs.com/addons/galapagos.html>  
[Accessed 25 06 2020].

Sadeghipour, M. & Mackey, C., 2017-2020. *Ladybug Tools*. [Online]  
Available at: <https://www.ladybug.tools/>  
[Accessed 30 11 2019].

SIMSCALE, 2019. *What Is PMV? What Is PPD? The Basics of Thermal Comfort*. [Online]  
Available at: <https://www.simscale.com/blog/2019/09/what-is-pmv-ppd/>  
[Accessed 26 01 2020].

Starmer, J., 2018. *Introduction to Machine Learning*. [Online]  
Available at: <https://statquest.org/video-index/>  
[Accessed 01 04 2020].

Thomas Wortmann, A. C. N., 2015. Advantages of surrogate models for architectural. *Artificial Intelligence for Engineering Design*, Issue 29, p. 471–481.

Umberto Berardi, H. K. A., 2015. Analysis of the Impacts of Light Shelves on the Useful Daylight. *Energy Procedia*, Issue 78, p. 1793 – 1798.

USGBC, 2019. *Good to know: Green building incentive strategies*. [Online]  
Available at: <https://www.usgbc.org/articles/good-know-green-building-incentive-strategies-0>  
[Accessed 09 01 2020].

Vierlinger, R., 2018. *Octopus*. [Online]  
Available at: <http://grasshopperdocs.com/addons/octopus.html>  
[Accessed 25 06 2020].

Vierlinger, R., 2018. *Octopus 0.4*. [Online]  
Available at: <https://www.food4rhino.com/app/octopus>  
[Accessed 20 04 2020].

Westermann, P. & Evins, R., 2019. Surrogate modeling for sustainable building design – A review. *Energy & Buildings*, Issue 198, pp. 170-186.

Westermann, P. & Evins, R., 2019. Surrogate modeling for sustainable building design – A review. *Energy & Buildings*, Issue 198, pp. 170-186.

Wikipedia, 2020. *Coefficient of Determination*. [Online]  
Available at: [https://en.wikipedia.org/wiki/Coefficient\\_of\\_determination](https://en.wikipedia.org/wiki/Coefficient_of_determination)  
[Accessed 14 05 2020].

Wikipedia, 2020. *Mean Squared error*. [Online]  
Available at: [https://en.wikipedia.org/wiki/Mean\\_squared\\_error](https://en.wikipedia.org/wiki/Mean_squared_error)  
[Accessed 20 04 2020].

Wikipedia, 2020. *Overfitting*. [Online]  
Available at: [https://en.wikipedia.org/wiki/Overfitting#cite\\_note-1](https://en.wikipedia.org/wiki/Overfitting#cite_note-1)  
[Accessed 30 06 2020].

Willmott, C. J. & Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, Volume 30, pp. 79-82.

Wortmann, T., 2017. Surveying design spaces with performance maps: A multivariate visualization method for parametric design and architectural design optimization. *International Journal of Architectural Computing*, Volume 15, p. 38.53.

Wortmann, T., 2018. *Efficient, Visual, and Interactive Architectural Design Optimization with Model-based Methods*, Singapore: Singapore University of Technology and Design.

Wortmann, T., Costa, A., Nannicini, G. & Schroepfer, T., 2015. Advantages of surrogate models for architectural design Optimization. *Artificial Intelligence for Engineering Design, Analysis, and Manufacturing*, Volume 29, p. 471–481.

Yeang Sdn. & Bhd, T. R. H. &, 1992. *Menara Mesiniaga*. Petaling Jaya, Malasia: s.n.

Zandbelt, D. & Mayer, H., 2012. *High-rise and the Sustainable City*. Amsterdam: Techne Press.

## 8.0 Appendix

Amsterdam EUI

<i>Regression Statistics</i>	
Multiple R	0.912613798
R Square	0.832863943
Adjusted R Square	0.827315859
Standard Error	12.40358111
Observations	499

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	16	369526.1096	23095.38185	150.1173763	1.6908E-175
Residual	482	74155.13331	153.8488243		
Total	498	443681.2429			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	83.00112601	5.734355313	14.4743605	1.09848E-39	71.73370341	94.26854861
	-		-			-
Analysis_Level	5.151540588	0.682665018	7.546220259	2.25415E-13	-6.492907635	3.810173541
	-		-			-
Core_Type	20.40124715	1.116594105	18.27096083	4.69909E-57	-22.59524054	18.20725376
FFH	7.624224224	0.498231517	15.30257312	2.27374E-43	6.645250171	8.603198276
	-		-			-
Length	11.03726107	0.496868448	22.21364851	8.38762E-76	-12.01355683	10.06096531
Number_Of_Fins	0.530092675	0.392882744	1.349238884	0.177893689	-0.241881796	1.302067147
	-		-			-
Orientation	0.341681141	0.242554461	1.408678033	0.159575129	-0.818275887	0.134913605
Overhang_Size	0.896656559	0.498429318	1.798964319	0.072649745	-0.082706152	1.876019271
	-		-			-
SHGC	1.410926812	0.682665953	2.066789483	0.039287429	-2.752295697	0.069557928
	-		-			-
Shape	-7.51423599	0.496750367	15.12678499	1.39564E-42	-8.490299733	6.538172247
U_Value	5.817990039	0.278459064	20.89352003	1.69253E-69	5.270846414	6.365133665
	-		-			-
VLT	2.082622714	0.681429487	3.056255644	0.002365565	-3.421562068	-0.74368336
WWR_East_North	2.905349995	0.498767216	5.825062072	1.04459E-08	1.925323348	3.885376642
WWR_North_West	3.326750141	0.49858351	6.672403063	6.92711E-11	2.347084458	4.306415824
WWR_South_East	3.08149219	0.498685401	6.179230801	1.37211E-09	2.101626302	4.061358078
WWR_West_South	3.662798118	0.497880659	7.356779288	8.1704E-13	2.684513466	4.641082769
	-		-			-
Width	12.95817877	0.682355311	18.99036845	1.8994E-60	-14.29893727	11.61742026

Amsterdam Floor  
Area

<i>Regression Statistics</i>	
Multiple R	0.922418729
R Square	0.850856311
Adjusted R Square	0.845905483
Standard Error	377.9269054
Observations	499

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	16	392748020	24546751	171.8614213	2.3529E-187
Residual	482	68843455.47	142828.7		
Total	498	461591475.5			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	-1939.46836	174.7210857	-11.1004	1.19103E-25	-2282.777451	-1596.16
-	-	-	-	-	-	-
Analysis_Level	27.47484614	20.80024111	-1.32089	0.187164553	-68.34519581	13.3955
Core_Type	27.34094818	34.02170315	0.803633	0.422005353	-39.50822425	94.19012
FFH	8.281998263	15.18070417	0.545561	0.585620203	-21.54653519	38.11053
Length	520.5483095	15.13917259	34.3842	8.5761E-132	490.8013813	550.2952
-	-	-	-	-	-	-
Number_Of_Fins	1.430045251	11.97081378	-0.11946	0.904959919	-24.95147189	22.09138
Orientation	-8.40541081	7.390434745	-1.13734	0.255962729	-22.92686038	6.116039
Overhang_Size	6.474224444	15.18673101	0.426308	0.670073772	-23.36615114	36.3146
-	-	-	-	-	-	-
SHGC	13.12856821	20.8002696	-0.63117	0.528226611	-53.99897387	27.74184
Shape	373.154757	15.13557473	24.65415	1.98342E-87	343.4148982	402.8946
U_Value	6.657934899	8.484418455	0.784725	0.433000354	-10.01308094	23.32895
-	-	-	-	-	-	-
VLT	9.410443682	20.76259548	-0.45324	0.650579729	-50.20682355	31.38594
WWR_East_North	1.384381769	15.19702648	0.091096	0.927454498	-28.47622337	31.24499
-	-	-	-	-	-	-
WWR_North_West	3.234253078	15.19142911	-0.2129	0.831495086	-33.08385994	26.61535
WWR_South_East	28.44479005	15.19453364	1.872041	0.06180625	-1.410916893	58.3005
WWR_West_South	2.956937305	15.1700138	0.19492	0.845537802	-26.85059067	32.76447
Width	634.5651701	20.79080458	30.52143	1.094E-114	593.7133623	675.417

Amsterdam Comfort

<i>Regression Statistics</i>	
Multiple R	0.901600646
R Square	0.812883725
Adjusted R Square	0.806672396
Standard Error	5.249953174
Observations	499

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	16	57713.14774	3607.071734	130.8711503	9.4106E-164
Residual	482	13284.88802	27.56200833		
Total	498	70998.03576			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	65.20171225	2.427129441	26.86371446	7.65009E-98	60.43265076	69.97077375
	-	-	-	-	-	-
Analysis_Level	4.807662163	0.288945535	16.63864495	1.88414E-49	-5.375410633	4.239913694
	-	-	-	-	-	-
Core_Type	0.585040019	0.47261083	1.237889574	0.216359638	-1.513672038	0.343591999
	-	-	-	-	-	-
FFH	0.932576173	0.210882012	4.422265159	1.20819E-05	-1.34693779	0.518214555
Length	1.902422654	0.210305078	9.046013893	3.64528E-18	1.489194653	2.315650656
	-	-	-	-	-	-
Number_Of_Fins	0.448303564	0.166291976	2.695882118	0.007265774	-0.775050315	0.121556813
Orientation	0.025625642	0.102663864	0.249607219	0.802997454	-0.176098367	0.22734965
	-	-	-	-	-	-
Overhang_Size	0.947699221	0.210965733	4.492195045	8.83361E-06	-1.362225343	0.533173099
SHGC	4.854668839	0.288945931	16.80130544	3.34898E-50	4.286919592	5.422418086
Shape	0.675915031	0.210255098	3.214737892	0.001393238	0.262785234	1.089044828
	-	-	-	-	-	-
U_Value	4.135664799	0.117860885	35.08937501	7.8387E-135	-4.367249403	3.904080196
	-	-	-	-	-	-
VLT	0.427374955	0.288422583	1.481766618	0.139055845	-0.994095876	0.139345967
	-	-	-	-	-	-
WWR_East_North	0.423331242	0.211108752	2.005275657	0.045492142	-0.838138382	0.008524102
	-	-	-	-	-	-
WWR_North_West	1.090778042	0.211030997	5.168804865	3.45769E-07	-1.5054324	0.676123684
	-	-	-	-	-	-
WWR_South_East	0.945477036	0.211074123	4.479360247	9.35907E-06	-1.360216133	0.530737939
	-	-	-	-	-	-
WWR_West_South	1.147249889	0.210733507	5.444079141	8.30836E-08	-1.56131971	0.733180068
Width	1.805474748	0.288814448	6.251331126	8.9697E-10	1.237983851	2.372965644

Amsterdam UDI

<i>Regression Statistics</i>	
Multiple R	0.610098731
R Square	0.372220462
Adjusted R Square	0.351381307
Standard Error	10.84377036
Observations	499

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	16	33604.7563	2100.297268	17.86159112	2.04321E-39
Residual	482	56677.10544	117.5873557		
Total	498	90281.86173			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	53.59981554	5.013232199	10.69166825	4.36947E-24	43.74932622	63.45030486
Analysis_Level	1.070135069	0.596816566	1.793071991	0.073588258	-0.102548538	2.242818676
-	-	-	-	-	-	-
Core_Type	13.79053094	0.976176957	-14.127081	3.65082E-38	-15.70861897	11.87244291
FFH	1.113110736	0.435576477	2.555488633	0.010909747	0.257247439	1.968974032
-	-	-	-	-	-	-
Length	2.465844269	0.434384821	5.676635443	2.37542E-08	-3.319366083	1.612322455
-	-	-	-	-	-	-
Number_Of_Fins	0.344975188	0.343475825	1.004365265	0.315706524	-1.019870109	0.329919733
-	-	-	-	-	-	-
Orientation	0.205371408	0.212052056	0.968495245	0.333282805	-0.622032043	0.211289227
Overhang_Size	0.435289153	0.435749404	0.998943773	0.318323153	-0.420913927	1.291492233
-	-	-	-	-	-	-
SHGC	0.046218339	0.596817384	0.077441341	0.93830459	-1.218903552	1.126466874
Shape	0.425895428	0.434281588	0.980689578	0.327237821	-0.427423545	1.2792144
-	-	-	-	-	-	-
U_Value	0.420005261	0.243441481	-1.72528223	0.085117591	-0.89834291	0.058332389
-	-	-	-	-	-	-
VLT	1.297544435	0.595736409	2.178051256	0.029886243	-2.468105643	0.126983226
WWR_East_North	1.394643046	0.436044809	3.198393871	0.001472908	0.537859524	2.251426567
WWR_North_West	0.390053288	0.435884205	0.894855292	0.371311082	-0.466414663	1.246521239
-	-	-	-	-	-	-
WWR_South_East	0.387596305	0.435973283	0.889036829	0.374426755	-1.244239284	0.469046674
WWR_West_South	0.584637404	0.435269741	1.343161146	0.179851877	-0.270623187	1.439897995
-	-	-	-	-	-	-
Width	2.381905808	0.596545807	-3.99282969	7.54714E-05	-3.554057399	1.209754216

Bogota EUI

<i>Regression Statistics</i>	
Multiple R	0.855299342
R Square	0.731536964
Adjusted R Square	0.722625328
Standard Error	6.535181004
Observations	499

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	16	56093.69606	3505.856004	82.08784094	2.7357E-126
Residual	482	20585.54074	42.70859075		
Total	498	76679.23681			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	54.0031786	3.020531321	17.87870174	3.25663E-55	48.06814301	59.9382142
Analysis_Level	1.748643409	0.359740648	4.860844662	1.58408E-06	1.041789772	2.455497046
	-	-	-	-	-	-
Core_Type	7.243715247	0.588486851	12.30905198	1.81689E-30	-8.400031814	-6.08739868
FFH	1.608893726	0.262075044	6.139057356	1.7358E-09	1.093943027	2.123844424
	-	-	-	-	-	-
Length	4.435676381	0.261828451	16.94115506	7.55847E-51	-4.950142549	3.921210213
Number_Of_Fins	0.230676888	0.206995249	1.114406679	0.265660125	-0.176047638	0.637401415
	-	-	-	-	-	-
Orientation	0.251328922	0.128071025	1.962418289	0.05028866	-0.502975408	0.000317564
Overhang_Size	0.755363433	0.26257245	2.876780995	0.004195674	0.239435383	1.271291483
	-	-	-	-	-	-
SHGC	2.322381956	0.359050612	-6.46811863	2.4399E-10	-3.027879743	1.616884168
	-	-	-	-	-	-
Shape	2.997564243	0.262167412	11.43377897	5.95167E-27	-3.512696435	-2.48243205
U_Value	2.833474259	0.146563241	19.33277563	4.53831E-62	2.545492457	3.121456061
	-	-	-	-	-	-
VLT	-0.89006372	0.359648186	2.474817765	0.01367337	-1.596735679	0.183391762
WWR_East_North	0.672120007	0.262709168	2.558418543	0.010819491	0.155923319	1.188316696
WWR_North_West	1.120477004	0.262679753	4.265562877	2.40083E-05	0.604338114	1.636615895
WWR_South_East	0.83561019	0.262715329	3.180667814	0.00156406	0.319401398	1.351818983
WWR_West_South	0.75311329	0.261884521	2.875745723	0.004209222	0.23853695	1.267689629
	-	-	-	-	-	-
Width	5.051054792	0.359636283	14.04489767	8.3235E-38	-5.757703362	4.344406221

Bogota Floor Area

<i>Regression Statistics</i>	
Multiple R	0.922379606
R Square	0.850784138
Adjusted R Square	0.845830915
Standard Error	378.0341565
Observations	499

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	16	392747577	24546723.56	171.7637246	2.6423E-187
Residual	482	68882534.91	142909.8235		
Total	498	461630111.9			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	1942.286777	174.72569	11.11620607	1.0342E-25	-2285.604915	1598.968639
Analysis_Level	26.99901078	20.80956172	1.297432937	0.195102884	-67.88767451	13.88965295
Core_Type	28.34389967	34.04161727	0.832624944	0.405468641	-38.54440196	95.2322013
FFH	8.740402794	15.15999604	0.576543871	0.564516761	-21.0474413	38.52824689
Length	520.3579703	15.14573161	34.35674049	1.1276E-131	490.5981544	550.1177863
Number_Of_Fins	1.594224257	11.97384961	0.133142165	0.894136516	-25.12161597	21.93316746
Orientation	-8.24535685	7.408398019	1.112974334	0.266274199	-22.80210242	6.311388718
Overhang_Size	6.069713054	15.18876898	0.399618499	0.689614547	-23.77466692	35.91409303
SHGC	12.46052244	20.76964591	0.599939089	0.548828763	-53.27075568	28.3497108
Shape	373.0839665	15.16533918	24.6010961	3.5408E-87	343.2856236	402.8823093
U_Value	6.454184138	8.47809896	0.761277283	0.446863754	-10.20441454	23.11278282
VLT	9.238768729	20.80421316	0.444081622	0.657182874	-50.11692309	31.63938563
WWR_East_North	0.747658561	15.19667761	0.049198817	0.960781234	-29.11226107	30.60757819
WWR_North_West	3.414260535	15.19497607	0.224696671	0.822310477	-33.27083681	26.44231574
WWR_South_East	28.84285083	15.19703397	1.897926325	0.058302972	-1.017769005	58.70347067
WWR_West_South	3.424950161	15.14897503	0.226084613	0.82123145	-26.34123877	33.19113909
Width	634.1191952	20.80352462	30.48133462	1.6615E-114	593.2423937	674.9959966

Bogota Comfort

<i>Regression Statistics</i>	
Multiple R	0.906522641
R Square	0.821783298
Adjusted R Square	0.815867391
Standard Error	7.292327764
Observations	499

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	16	118192.0615	7387.003846	138.9107846	8.058E-169
Residual	482	25631.81731	53.17804422		
Total	498	143823.8788			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	45.61585899	3.370481155	13.53393088	1.33628E-35	38.99320771	52.23851027
	-	-	-	-	-	-
Analysis_Level	10.06091147	0.401419137	25.06335786	2.28271E-89	-10.84965908	9.272163863
Core_Type	1.254619356	0.656667199	1.910586302	0.056650793	-0.035664632	2.544903344
FFH	0.421393361	0.292438284	1.4409651	0.150243838	-0.153218003	0.996004726
Length	0.119850238	0.292163121	0.410216859	0.681829147	-0.45422046	0.693920936
	-	-	-	-	-	-
Number_Of_Fins	1.203026754	0.230977107	5.208424193	2.82683E-07	-1.656873182	0.749180325
	-	-	-	-	-	-
Orientation	0.060097301	0.142908956	0.420528587	0.67428677	-0.340898806	0.220704203
	-	-	-	-	-	-
Overhang_Size	2.216569069	0.292993318	7.565254684	1.97786E-13	-2.792271018	-1.64086712
SHGC	10.69458095	0.400649156	26.69313238	4.81077E-97	9.90734627	11.48181562
	-	-	-	-	-	-
Shape	1.729294113	0.292541354	5.911280886	6.43036E-09	-2.304107999	1.154480227
	-	-	-	-	-	-
U_Value	4.376337093	0.163543625	26.75944773	2.35316E-97	-4.697683616	-4.05499057
	-	-	-	-	-	-
VLT	0.857795548	0.401315962	2.137456838	0.033063386	-1.646340432	0.069250664
WWR_East_North	1.112290597	0.293145876	3.794324557	0.000166888	0.536288886	1.688292308
WWR_North_West	0.218174829	0.293113053	0.744336789	0.457035475	-0.357762388	0.794112046
WWR_South_East	0.322675397	0.29315275	1.100707384	0.271573419	-0.253339821	0.898690615
WWR_West_South	0.363787955	0.292225687	1.244886985	0.213778212	-0.210405679	0.937981589
Width	0.948804327	0.40130268	2.364310965	0.018459126	0.160285541	1.737323113

Bogota UDI

<i>Regression Statistics</i>	
Multiple R	0.627443542
R Square	0.393685398
Adjusted R Square	0.373558773
Standard Error	13.3738573
Observations	499

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	16	55977.26668	3498.579168	19.56042721	6.87948E-43
Residual	482	86210.5485	178.8600591		
Total	498	142187.8152			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	69.00950231	6.181336806	11.16417119	6.7391E-26	56.86380671	81.1551979
Analysis_Level	1.183684727	0.736187734	1.60785717	0.108521266	-0.262848992	2.630218445
Core_Type	-18.0903395	1.204303166	15.02141654	4.12675E-42	-20.45667224	15.72400676
FFH	0.060482813	0.536320912	0.11277355	0.910257041	-0.993333011	1.114298638
Length	0.702144987	0.535816274	1.310421167	0.190677483	-1.754969249	0.350679274
Number_Of_Fins	0.210636869	0.423603405	0.49725018	0.619239382	-0.621700564	1.042974303
Orientation	0.284442548	0.262089698	1.085287022	0.278337056	-0.79942204	0.230536945
Overhang_Size	1.244776517	0.537338824	2.316557938	0.020946257	0.188960601	2.300592433
SHGC	0.204473854	0.734775617	0.278280674	0.780916247	-1.648232907	1.2392852
Shape	2.506976091	0.536509939	4.67274865	3.86147E-06	1.45278885	3.561163333
U_Value	0.492436917	0.299932913	-1.64182354	0.101278832	-1.081774461	0.096900628
VLT	-4.69721873	0.735998516	6.382103534	4.10729E-10	-6.143380654	3.251056805
WWR_East_North	0.86794143	0.53761861	1.614418499	0.107091086	-0.188424238	1.924307098
WWR_North_West	0.684129764	0.537558414	1.272661252	0.203751817	-1.740377153	0.372117624
WWR_South_East	1.261275039	0.537631217	2.345985498	0.019381133	-2.317665478	0.204884599
WWR_West_South	0.365482246	0.535931018	0.681957628	0.495593305	-1.418531968	0.687567475
Width	1.087256664	0.735974157	1.477302774	0.14024757	-2.533370725	0.358857397

Dubai EUI

<i>Regression Statistics</i>	
Multiple R	0.894235035
R Square	0.799656297
Adjusted R Square	0.792992058
Standard Error	1300.224527
Observations	498

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	16	3245708031	2.03E+08	119.9921291	2.6059E-156
Residual	481	813170817.3	1690584		
Total	497	4058878849			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	1852.381805	601.9104099	3.077504	0.002206539	669.6831267	3035.080483
	-					
Analysis_Level	1923.719593	71.69948636	-26.8303	1.28424E-97	-2064.602499	1782.836687
Core_Type	112.4775524	117.211247	0.959614	0.337731635	-117.8317837	342.7868885
FFH	640.9934661	52.35381974	12.24349	3.42577E-30	538.1230183	743.863914
	-					
Length	866.6696836	52.10279134	-16.6338	2.07367E-49	-969.0468837	764.2924835
Number_Of_Fins	84.74920825	41.21479962	2.056281	0.040294417	3.765911997	165.7325045
Orientation	0.924877916	25.5029548	0.036266	0.971085688	-49.18608618	51.03584201
Overhang_Size	138.61488	52.17591883	2.656683	0.008153647	36.09399112	241.1357688
SHGC	49.00255446	71.54618423	0.684908	0.493731876	-91.57912702	189.5842359
	-					
Shape	739.1487055	52.17786148	-14.1659	2.53948E-38	-841.6734115	636.6239995
U_Value	73.07476586	29.19694486	2.502822	0.012651347	15.70545055	130.4440812
	-					
VLT	90.60131096	71.80044866	-1.26185	0.207614734	-231.6825987	50.47997678
WWR_East_North	381.6181481	52.44957301	7.275906	1.40887E-12	278.5595539	484.6767423
WWR_North_West	362.5399202	52.35819081	6.924226	1.40949E-11	259.6608836	465.4189568
WWR_South_East	420.1167738	52.37710098	8.021001	8.06001E-15	317.2005805	523.0329671
WWR_West_South	517.9836288	52.21712521	9.919804	3.16557E-21	415.3817732	620.5854845
	-					
Width	1198.055797	71.7396235	-16.7001	1.02764E-49	-1339.017569	1057.094025

## Dubai Floor Area

<i>Regression Statistics</i>	
Multiple R	0.922117627
R Square	0.850300917
Adjusted R Square	0.845321322
Standard Error	378.4754367
Observations	498

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	16	391357790	24459861.88	170.757034	1.4617E-186
Residual	481	68900198.63	143243.6562		
Total	497	460257988.6			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	-1942.98299	175.2068974	11.08964897	1.32647E-25	-2287.248453	1598.717527
Analysis_Level	26.68434151	20.87062185	1.278559964	0.201668617	-67.69319677	14.32451375
Core_Type	29.69580263	34.11839799	0.870375058	0.384529629	-37.34371613	96.73532138
FFH	8.243623017	15.23939472	0.54094163	0.588798236	-21.70038809	38.18763412
Length	520.6634087	15.1663242	34.33023069	1.9137E-131	490.8629744	550.4638429
Number_Of_Fins	0.770370977	11.99699665	0.064213653	0.948826784	-24.34336763	22.80262568
Orientation	7.599990933	7.423519369	1.023771954	0.306457598	-22.18652475	6.986542879
Overhang_Size	5.400550967	15.18761049	0.355589246	0.722304086	-24.44170888	35.24281081
SHGC	12.19964074	20.82599794	0.585789011	0.558292057	-53.12081413	28.72153264
Shape	373.9682372	15.18817597	24.62232713	3.19574E-87	344.1248663	403.8116082
U_Value	6.163642698	8.498783272	0.725238249	0.468658555	-10.53568594	22.86297133
VLT	9.705110443	20.90001042	-0.46435912	0.642600576	-50.77171155	31.36149067
WWR_East_North	1.514025022	15.26726703	0.099168045	0.921046161	-28.48475261	31.51280265
WWR_North_West	-4.13564896	15.24066707	0.271356164	0.786233532	-34.08216011	25.81086219
WWR_South_East	29.02489191	15.24617154	1.903749531	0.057539337	-0.932435009	58.98221883
WWR_West_South	2.894435655	15.19960504	0.190428347	0.849053829	-26.97139238	32.76026369
Width	634.3220388	20.88230515	30.37605447	6.1285E-114	593.2902269	675.3538507

## Dubai Comfort

<i>Regression Statistics</i>	
Multiple R	0.897308758
R Square	0.805163007
Adjusted R Square	0.798681943
Standard Error	1.886196471
Observations	498

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	16	7071.822102	441.9888814	124.2331475	3.3546E-159
Residual	481	1711.271558	3.557737128		
Total	497	8783.093661			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	87.3033705	0.873173262	99.98401728	0	85.58766523	89.01907577
	-	-	-	-	-	-
Analysis_Level	3.043764873	0.10401228	29.26351445	7.0897E-109	-3.248139452	2.839390294
Core_Type	1.077027608	0.170034818	6.334159216	5.48641E-10	0.742924806	1.411130409
FFH	0.492579046	0.075948106	6.485731809	2.19551E-10	0.343347993	0.6418101
	-	-	-	-	-	-
Length	0.249354453	0.075583947	3.299039866	0.001042119	-0.397869967	0.100838939
	-	-	-	-	-	-
Number_Of_Fins	0.255382663	0.059789066	4.271394111	2.34202E-05	-0.372862686	-0.13790264
Orientation	0.008956478	0.036996367	0.24209075	0.808813027	-0.063737985	0.081650941
	-	-	-	-	-	-
Overhang_Size	0.494542903	0.07569003	6.533791833	1.63626E-10	-0.643266861	0.345818944
SHGC	2.257907785	0.10378989	21.75460242	1.43801E-73	2.053970184	2.461845387
	-	-	-	-	-	-
Shape	1.148948778	0.075692849	15.17909288	8.42153E-43	-1.297678274	1.000219282
	-	-	-	-	-	-
U_Value	0.416346836	0.042355127	9.829904186	6.67141E-21	-0.49957077	0.333122902
	-	-	-	-	-	-
VLT	0.216220825	0.104158743	2.075877813	0.038435927	-0.42088319	0.011558459
WWR_East_North	0.375285591	0.076087012	4.932321299	1.12115E-06	0.2257816	0.524789583
WWR_North_West	0.186285616	0.075954447	2.452596576	0.014537062	0.037042104	0.335529129
WWR_South_East	0.17393184	0.075981879	2.289122643	0.022504725	0.024634425	0.323229255
WWR_West_South	0.220360041	0.075749807	2.909050843	0.003793116	0.071518626	0.369201455
	-	-	-	-	-	-
Width	1.098269217	0.104070506	-10.5531265	1.46989E-23	-1.302758204	-0.89378023

Dubai UDI

<i>Regression Statistics</i>	
Multiple R	0.627799343
R Square	0.394132015
Adjusted R Square	0.373978403
Standard Error	14.07329865
Observations	498

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	16	61972.72523	3873.295327	19.55639514	7.35327E-43
Residual	481	95265.77054	198.057735		
Total	497	157238.4958			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	72.19736354	6.514924759	11.0818415	1.42192E-25	59.3961347	84.99859239
Analysis_Level	1.263623293	0.776056953	1.628261029	0.104124323	-0.26125735	2.788503937
Core_Type	19.00169314	1.268664642	14.97771161	6.67947E-42	-21.49450266	16.50888363
FFH	0.056119314	0.566664392	0.099034481	0.921152152	-1.169562804	1.057324176
Length	0.484022671	0.563947325	0.858276385	0.391167342	-1.592127373	0.624082031
Number_Of_Fins	0.195509947	0.446098479	0.438266337	0.661389985	-0.681032597	1.07205249
Orientation	0.329706939	0.276037478	1.194428165	0.232899187	-0.872095233	0.212681356
Overhang_Size	1.412896081	0.564738838	2.501857471	0.012685422	0.303236127	2.522556034
SHGC	0.123368411	0.77439765	0.159308865	0.873492407	-1.644988676	1.398251854
Shape	2.637580038	0.564759865	4.670268203	3.9084E-06	1.52787877	3.747281307
U_Value	-0.46219315	0.316020284	1.462542667	0.144245629	-1.083143987	0.158757687
VLT	4.979836813	0.777149744	6.407821471	3.52309E-10	-6.506864689	3.452808937
WWR_East_North	0.548174104	0.567700801	0.965603894	0.33472765	-0.567305834	1.663654043
WWR_North_West	0.809771995	0.566711703	1.428895839	0.153683242	-1.923308447	0.303764458
WWR_South_East	1.464534792	0.566916382	2.583334753	0.010079071	-2.57847342	0.350596165
WWR_West_South	0.584846194	0.565184845	1.034787467	0.30128814	-1.695382511	0.525690122
Width	0.570372622	0.776491388	0.734551125	0.462970855	-2.096106889	0.955361644

## Galapagos – Evolutionary Algorithm

	Galapagos 1	Galapagos 2	Galapagos 3	Galapagos 4	Galapagos 5	F(x)_min AVG
F(x)_min	EUI	EUI	EUI	EUI	EUI	
	1.65E+07	1.70E+08	1.57E+08	9.20E+07	1.72E+08	1.21E+08
	1.99E+07	3.93E+06	3.93E+06	3.93E+06	3.93E+06	
	1.65E+07	1.22E+08	2.12E+08	1.20E+08	1.48E+08	
	2.06E+07	1.16E+08	2.14E+08	3.01E+08	1.95E+08	
	1.81E+07	3.33E+08	7.76E+07	2.49E+08	1.06E+08	
	1.79E+07	1.08E+08	1.25E+08	1.71E+08	7.86E+07	
	1.92E+07	1.51E+08	1.33E+08	7.40E+07	2.48E+08	
	2.14E+07	2.77E+08	7.63E+07	1.14E+08	1.39E+08	
	2.11E+07	1.86E+08	3.12E+08	1.53E+08	1.17E+08	
	2.31E+07	5.75E+07	3.64E+08	1.87E+08	1.45E+08	
	2.53E+07	9.36E+07	1.21E+08	1.41E+08	1.73E+08	
	2.20E+07	9.76E+07	2.04E+08	2.34E+07	1.83E+08	
	2.64E+07	3.20E+08	1.06E+08	1.61E+08	7.49E+07	
	4.68E+06	3.53E+07	2.34E+08	3.67E+07	4.20E+07	
	4.95E+07	1.36E+08	1.05E+08	3.91E+08	3.19E+08	
	1.88E+07	7.08E+07	2.34E+08	2.10E+08	1.15E+08	
	2.40E+07	4.26E+07	1.61E+08	3.81E+07	2.05E+08	
	2.27E+07	1.54E+08	1.78E+08	9.41E+07	1.67E+08	
	2.29E+07	1.56E+08	1.40E+08	3.69E+07	7.98E+07	
	1.28E+07	7.13E+07	1.09E+08	1.55E+08	1.21E+08	
	2.10E+07	1.74E+08	2.94E+08	1.80E+08	1.33E+08	
	1.73E+07	9.03E+07	5.27E+07	6.50E+07	5.04E+07	
	1.50E+07	3.48E+08	1.16E+08	4.81E+07	1.58E+08	
	1.69E+07	2.55E+07	3.37E+08	1.84E+08	1.26E+08	
	2.15E+07	1.16E+08	1.86E+08	2.94E+07	1.24E+08	
	1.60E+07	8.14E+07	1.99E+08	3.01E+08	1.06E+08	
	...	...	...	...	...	

## Galapagos - Inputs

	1200 2605 FES	Seconds	2602 generations	1 Population Size		
	Galapagos 1	Galapagos 2	Galapagos 3	Galapagos 4	Galapagos 5	
	Inputs	Inputs	Inputs	Inputs	Inputs	
Analysis_Level	1	0	0	1	1	
Core_Type	0	1	1	0	0	
FFH	8	6	6	8	8	
Length	1	4	2	1	1	
Number_Of_Fins	0	0	1	0	0	
Orientation	4	7	7	4	4	
Overhang_Size	4	2	3	4	4	
SHGC	1	3	3	1	1	
Shape	0	1	2	0	0	
U_Value	7	1	1	7	7	
VLT	3	3	3	3	3	
WWR_East_North	4	1	1	4	4	
WWR_North_West	4	4	4	4	4	
WWR_South_East	4	1	1	4	4	
WWR_West_South	1	1	1	1	1	
Width	1	3	3	1	1	
<b>EUI</b>	<b>1.65E+07</b>	<b>1.70E+08</b>	<b>1.57E+08</b>	<b>9.20E+07</b>	<b>1.72E+08</b>	

## Opossum - RBFOpt

	Opossum 1	Opossum 2	Opossum 3	Opossum 4	Opossum 5	
	Results	Results	Results	Results	Results	F(x)_min AVG
F(x)_min	EUI	EUI	EUI	EUI	EUI	
	4.71E+06	6.34E+06	4.97E+06	5.31E+06	5.07E+06	5.28E+06
	4.71E+06	1.19E+08	1.05E+09	3.39E+08	1.48E+08	
	2.21E+08	3.65E+08	2.65E+08	7.11E+08	3.43E+08	
	4.34E+08	1.96E+08	5.99E+08	4.18E+07	9.10E+07	
	7.74E+07	2.26E+08	9.47E+07	9.37E+07	1.56E+08	
	1.81E+08	2.14E+08	7.68E+07	1.07E+08	8.39E+07	
	4.52E+08	2.29E+08	1.00E+08	2.99E+08	2.29E+08	
	3.94E+07	2.28E+08	5.30E+08	1.37E+08	4.01E+08	
	1.46E+08	1.60E+08	2.50E+08	1.18E+08	2.68E+08	
	3.51E+08	4.05E+08	4.05E+08	4.05E+08	4.05E+08	
	4.05E+08	2.81E+07	3.88E+07	2.62E+07	2.97E+07	
	3.03E+07	1.26E+07	8.62E+07	6.72E+08	2.43E+07	
	3.34E+07	3.16E+08	3.45E+08	1.95E+08	7.06E+07	
	1.04E+08	5.70E+07	6.15E+07	1.45E+08	1.11E+07	
	3.23E+07	1.00E+08	5.22E+07	4.64E+07	5.53E+07	
	5.76E+07	4.76E+07	1.91E+08	5.54E+08	1.73E+08	
	5.93E+07	1.32E+07	3.96E+07	6.98E+07	7.50E+07	
	3.83E+08	3.02E+07	4.05E+07	2.51E+07	5.55E+06	
	3.13E+07	5.63E+07	1.92E+08	2.77E+07	4.24E+08	
	1.13E+08	1.12E+07	3.20E+07	1.45E+07	8.64E+06	
	1.57E+07	6.34E+06	2.81E+07	1.41E+07	5.69E+06	
	9.98E+06	1.04E+07	2.33E+07	1.38E+07	8.86E+06	
	1.00E+07	1.05E+07	1.37E+07	1.79E+07	1.11E+07	
	5.41E+07	1.10E+07	9.89E+06	3.74E+07	4.63E+07	
	1.52E+07	9.14E+06	9.12E+06	3.27E+07	1.94E+07	
	2.64E+07	1.49E+07	7.99E+06	1.24E+07	7.60E+06	
	8.88E+06	9.75E+06	5.07E+06	1.04E+07	5.52E+06	
	1.29E+07	6.42E+06	5.34E+06	1.12E+07	5.53E+06	
	5.78E+06	3.26E+07	2.33E+07	2.49E+07	1.83E+07	
	7.72E+07	5.49E+07	1.34E+07	6.64E+07	8.39E+06	
	1.27E+08	3.23E+07	1.43E+07	1.71E+07	1.56E+07	
	1.38E+08	3.21E+07	1.92E+07	1.57E+07	1.44E+07	
	3.62E+07	2.81E+07	2.95E+07	1.11E+07	1.63E+07	
	3.60E+07	1.66E+07	2.98E+07	1.60E+07	7.43E+06	
	1.03E+07	1.09E+07	1.49E+07	1.03E+07	5.76E+06	
	1.96E+07	9.13E+06	5.11E+06	9.92E+06	5.57E+06	
	...	...	...	...	...	

### Opossum RBFOpt - Inputs

1200 FES 136	Seconds	140 Generations				1 Population Size
	Opossum 1	Opossum 2	Opossum 3	Opossum 4	Opossum 5	
	Inputs	Inputs	Inputs	Inputs	Inputs	
Analysis_Level	0	2	2	2	0	
Core_Type	1	1	1	1	1	
FFH	6	6	6	7	6	
Length	4	1	1	1	1	
Number_Of_Fins	0	4	0	0	4	
Orientation	7	0	7	4	6	
Overhang_Size	4	2	4	4	4	
SHGC	3	1	1	2	1	
Shape	3	0	0	0	0	
U_Value	1	7	7	7	2	
VLT	3	1	1	3	3	
WWR_East_North	1	3	4	2	1	
WWR_North_West	4	4	4	4	4	
WWR_South_East	1	3	4	4	1	
WWR_West_South	2	3	2	1	4	
Width	1	1	1	1	1	
EUI	4.71E+06	6.34E+06	4.97E+06	5.31E+06	5.07E+06	

## Opossum – CMAES

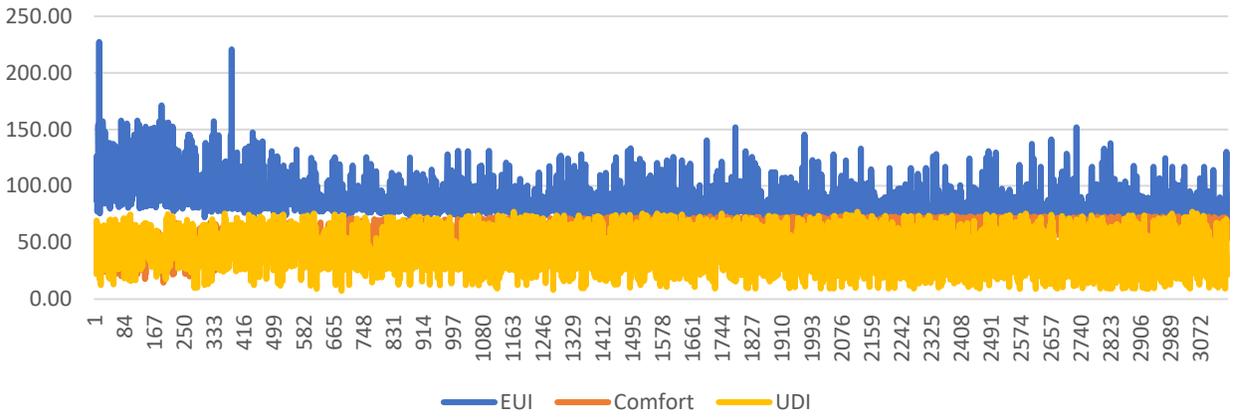
	Opossum 1 Results	Opossum 2 Results	Opossum 3 Results	Opossum 4 Results	Opossum 5 Results	F(x)_min AVG
F(x)_min	EUI	EUI	EUI	EUI	EUI	
	1.01E+07	1.30E+07	1.72E+07	1.46E+07	1.57E+07	1.41E+07
	6.47E+08	7.59E+07	3.47E+08	5.21E+07	4.14E+08	
	8.34E+07	2.95E+08	9.89E+07	4.51E+07	2.12E+08	
	4.36E+08	1.83E+08	7.42E+07	1.34E+08	9.56E+07	
	3.24E+08	8.49E+08	4.33E+07	7.87E+07	1.70E+08	
	3.05E+08	4.17E+08	1.50E+08	2.77E+08	7.19E+08	
	3.65E+08	6.18E+08	3.42E+07	3.00E+08	6.65E+07	
	3.31E+08	1.59E+08	9.11E+07	1.14E+08	5.83E+08	
	5.87E+08	9.65E+07	1.06E+08	2.46E+08	1.04E+08	
	6.62E+08	3.88E+08	7.66E+07	1.71E+08	6.35E+08	
	2.34E+08	8.14E+07	1.28E+08	3.72E+07	3.57E+08	
	4.52E+08	7.24E+08	5.20E+07	3.24E+07	1.31E+08	
	4.40E+08	4.78E+08	1.20E+08	2.22E+08	5.94E+07	
	4.78E+08	7.13E+07	3.74E+08	6.66E+07	2.51E+08	
	5.46E+08	2.90E+08	2.44E+08	2.46E+07	4.04E+07	
	5.48E+08	2.73E+08	1.00E+08	1.84E+08	1.00E+08	
	5.68E+08	1.88E+08	5.62E+07	6.78E+07	1.67E+08	
	2.30E+08	1.55E+08	8.78E+07	1.05E+08	3.60E+07	
	1.59E+07	5.04E+08	4.17E+07	3.84E+08	1.57E+07	
	4.18E+08	1.60E+08	7.28E+07	1.31E+08	5.87E+07	
	1.71E+08	7.37E+08	1.26E+08	4.20E+07	1.03E+08	
	3.87E+07	1.41E+08	5.89E+07	2.33E+08	1.67E+08	
	3.26E+08	2.02E+08	2.33E+07	2.25E+08	5.30E+07	
	2.01E+08	4.23E+08	1.70E+08	1.82E+08	2.25E+08	
	4.17E+08	2.70E+08	3.13E+07	7.53E+07	1.51E+08	
	1.86E+08	3.06E+07	3.16E+08	5.36E+07	7.55E+07	
	3.78E+07	2.47E+08	4.05E+07	6.79E+07	3.18E+07	
	...	...	...	...	...	

## Opossum – CMAES Inputs

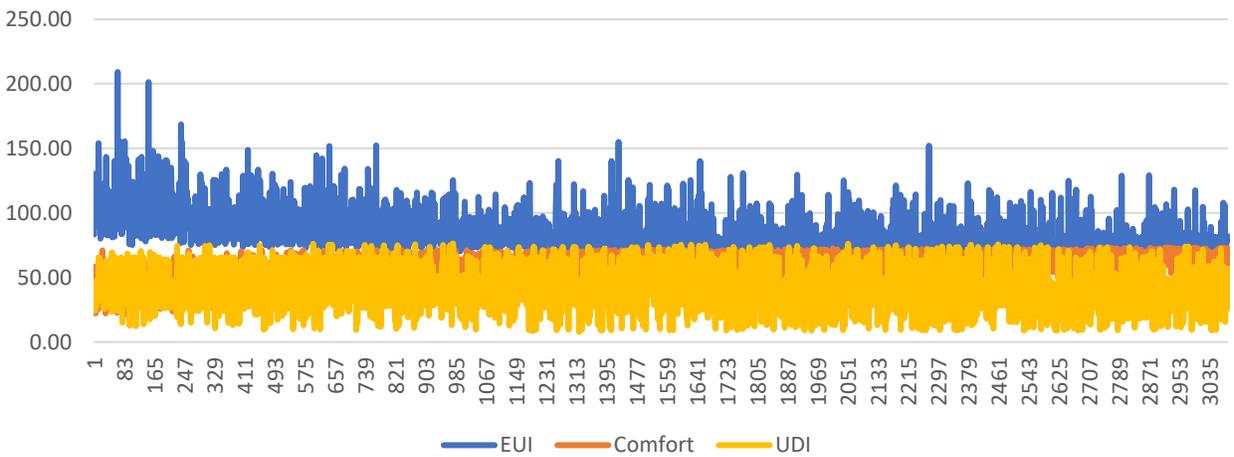
	1200 FES 71	Seconds	140 Generations			1 Population Size
	Opossum 1	Opossum 2	Opossum 3	Opossum 4	Opossum 5	
	Inputs	Inputs	Inputs	Inputs	Inputs	
Analysis_Level	1	2	1	1	1	
Core_Type	1	1	1	0	0	
FFH	7	6	8	7	7	
Length	1	1	2	1	1	
Number_Of_Fins	1	2	2	1	3	
Orientation	5	6	3	2	0	
Overhang_Size	3	1	2	3	2	
SHGC	2	2	2	1	1	
Shape	0	0	0	0	0	
U_Value	5	6	6	7	7	
VLТ	3	3	2	3	3	
WWR_East_North	3	4	1	1	4	
WWR_North_West	3	3	1	3	2	
WWR_South_East	4	1	2	2	4	
WWR_West_South	3	1	4	3	2	
Width	1	1	1	1	1	
EUI	1.01E+07	1.30E+07	1.72E+07	1.46E+07	1.57E+07	



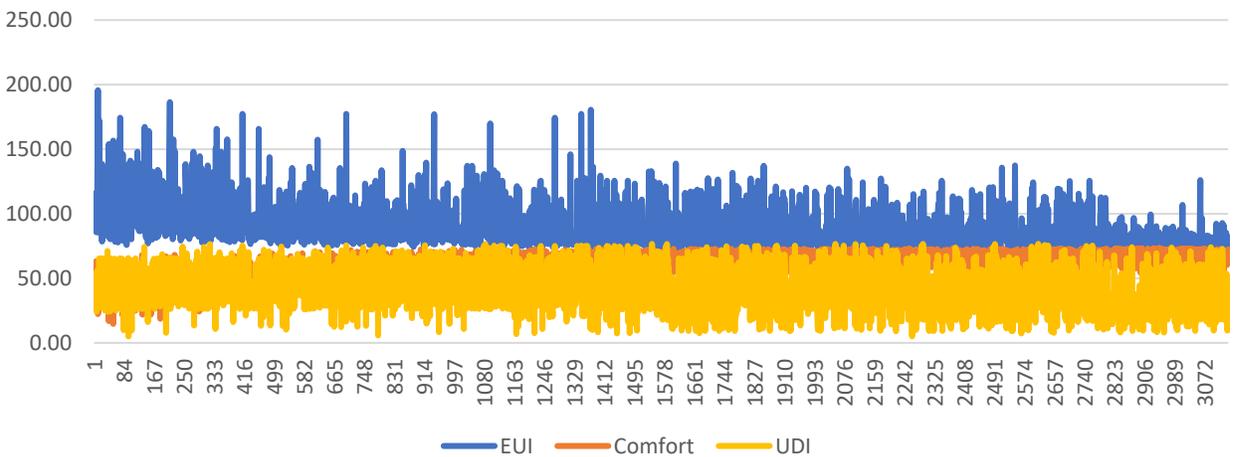
### Octopus 1



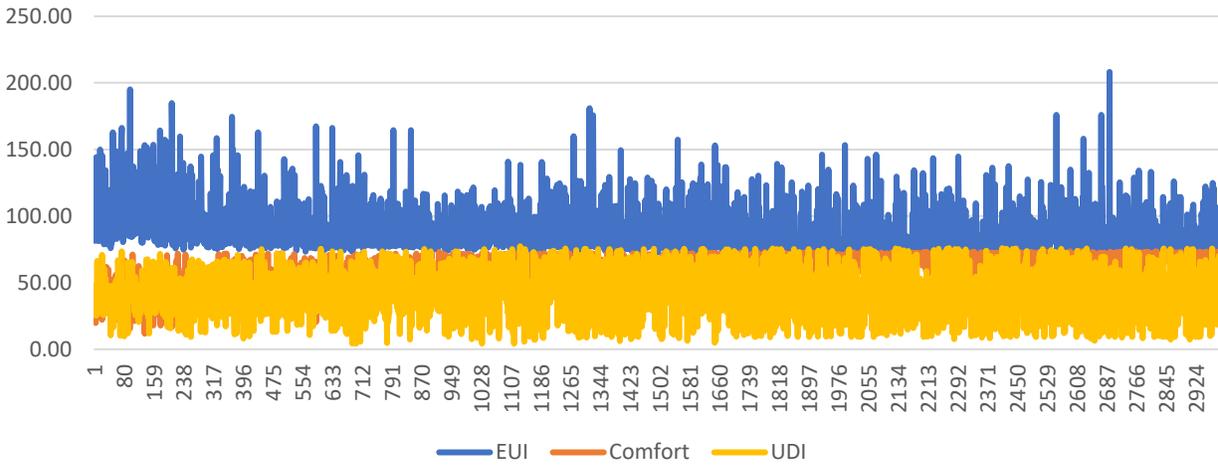
### Octopus 2



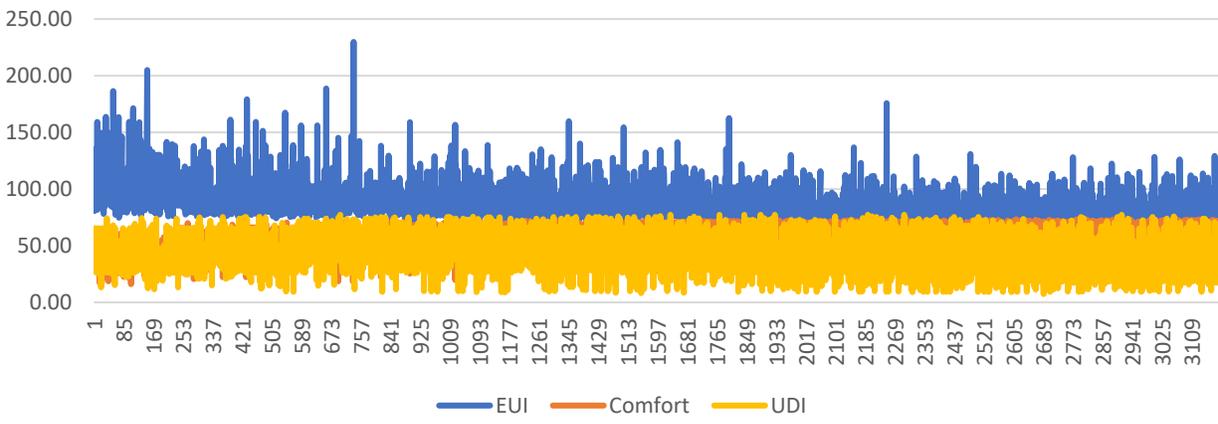
### Octopus 3



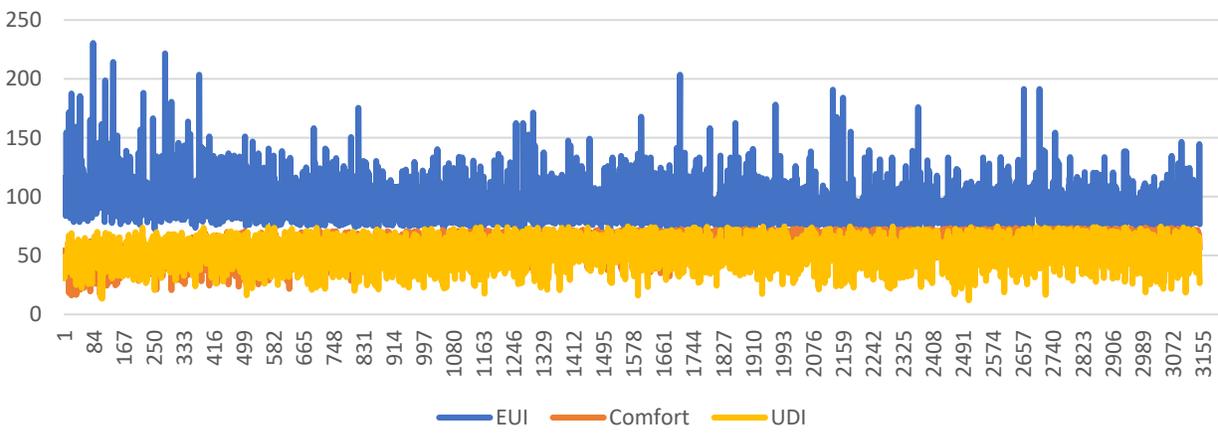
### Octopus 4



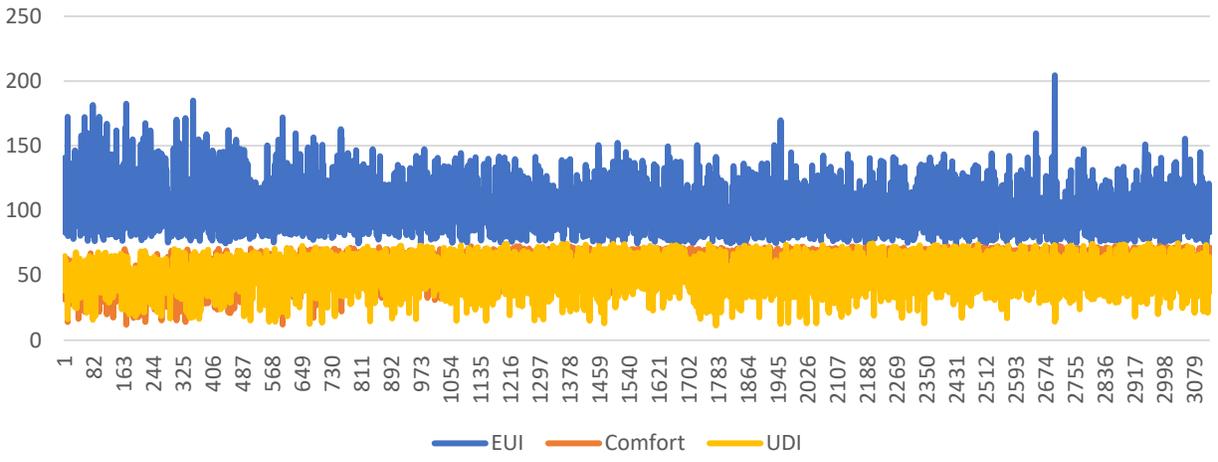
### Octopus 5



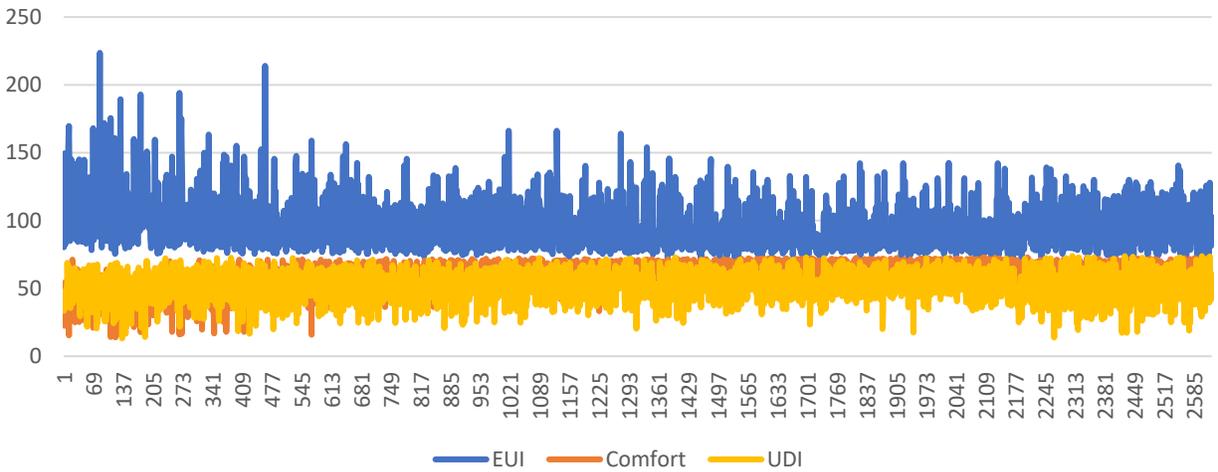
### Octopus 6



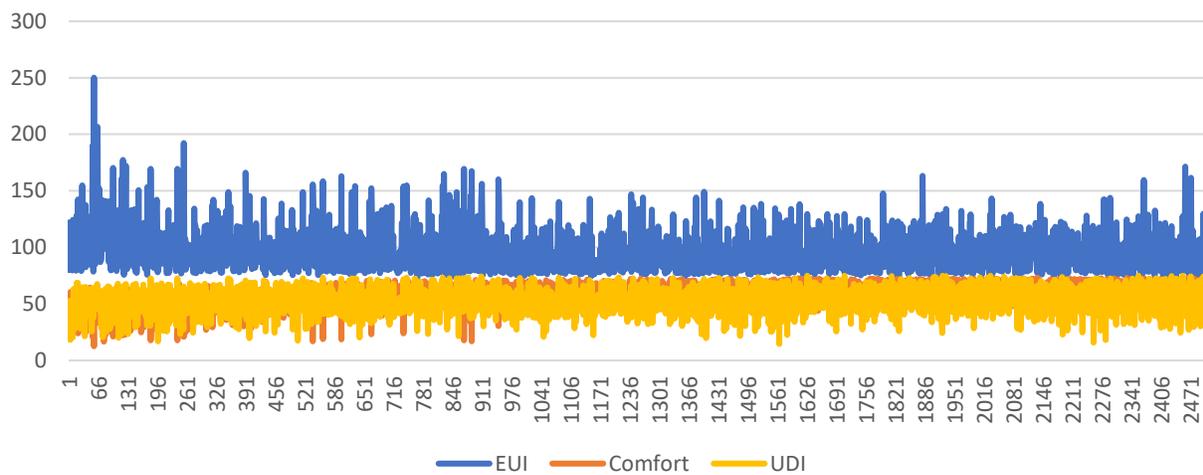
### Octopus 7



### Octopus 8



### Octopus 9



### Octopus 10

