Surrogate Models for efficient implementation of Building Performance Analysis and Optimization

November 2025

Gonçalo Roque Matias Araújo PhD | Sustainable Energy Systems

fct Fundação para a Ciência e a Tecnologia

TÉCNICO LISBOA

IN+ CENTER for INNOVATION, TECHNOLOGY and POLICY RESEARCH

CERIS

Supervisors

Surrogate Models to improve Building Performance Analysis and Optimization

Paulo Manuel Cadete Ferrão
Instituto Superior Técnico
Supervisor
ferrao@tecnico.ulisboa.pt

Maria da Glória Gomes
Instituto Superior Técnico
Co-supervisor
maria.gloria.gomes@tecnico.ulisboa.pt

Manuel de Arriaga Brito Correia Guedes
Instituto Superior Técnico
Co-supervisor
manuel.guedes@tecnico.ulisboa.pt

Fundação para a Ciência e a Tecnologia
TÉCNICO LISBOA
CENTER for INNOVATION, TECHNOLOGY and POLICY RESEARCH

Structure

Surrogate Models to improve Building Performance Analysis and Optimization

Fundação para a Ciência e a Tecnologia

TÉCNICO LISBOA

CENTER for INNOVATION, TECHNOLOGY and POLICY RESEARCH

Renovate the built environment

Optimize building design and construction

Council Directive (EU) 2018/844 of the European Parliament and of the Council of 30 May [2018] amending Directive 2010/31/EU [2010] on the energy performance of buildings and Directive 2012/27/EU [2012] on energy efficiency OJ L156/75.

Renovate the built environment

Optimize building design and construction

Population welfare

Economic growth

Sustainable goals

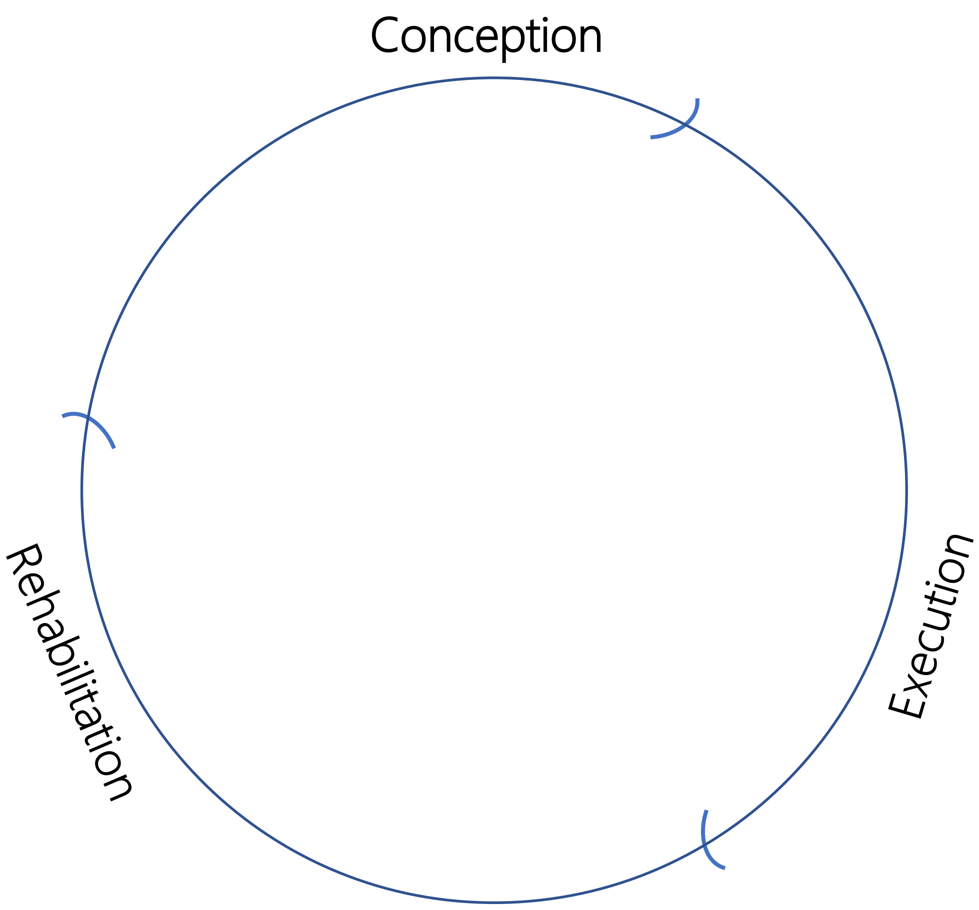
Council Directive (EU) 2018/844 of the European Parliament and of the Council of 30 May [2018] amending Directive 2010/31/EU [2010] on the energy performance of buildings and Directive 2012/27/EU [2012] on energy efficiency OJ L156/75.

Renovate the built environment

Optimize building design and construction

Population welfare	Economic growth	Sustainable goals
		
		

Council Directive (EU) 2018/844 of the European Parliament and of the Council of 30 May [2018] amending Directive 2010/31/EU [2010] on the energy performance of buildings and Directive 2012/27/EU [2012] on energy efficiency OJ L156/75.



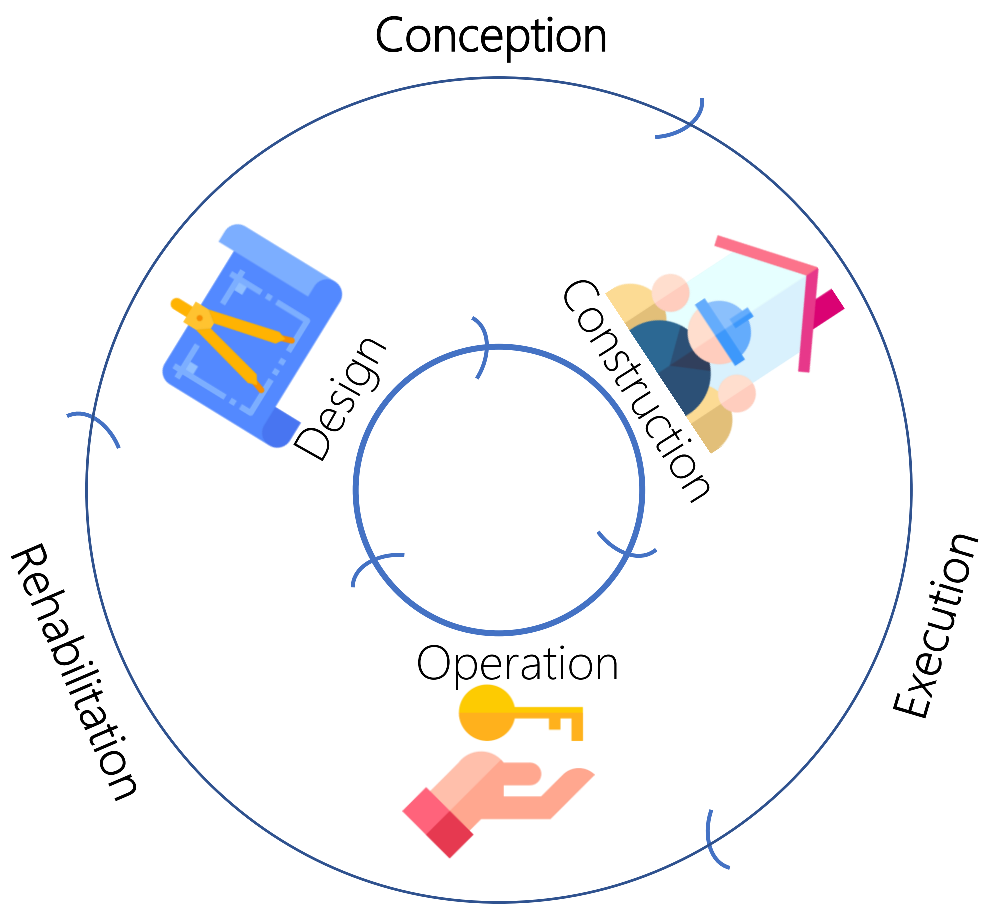
Renovate the built environment

Optimize building design and construction

Population welfare

Economic growth

Sustainable goals



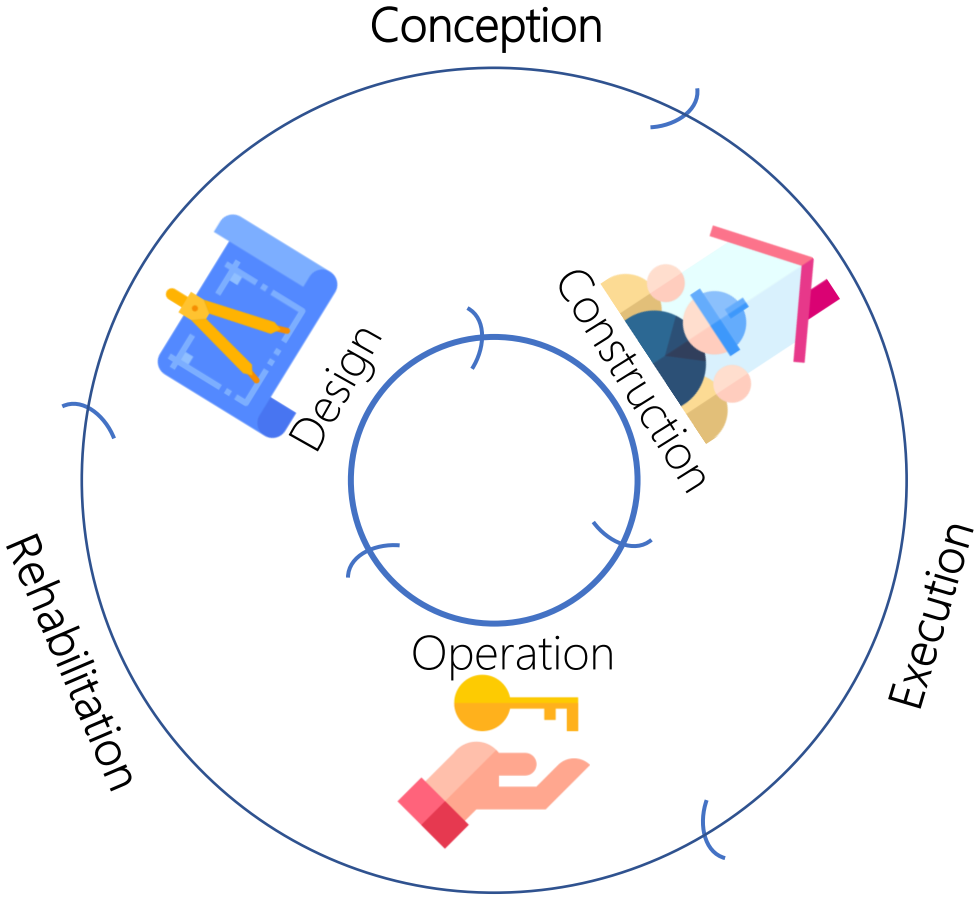
Renovate the built environment

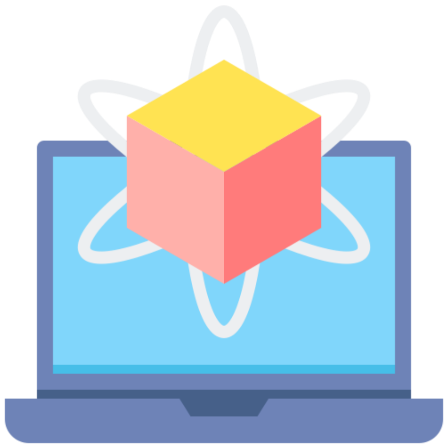
Optimize building design and construction

Population welfare

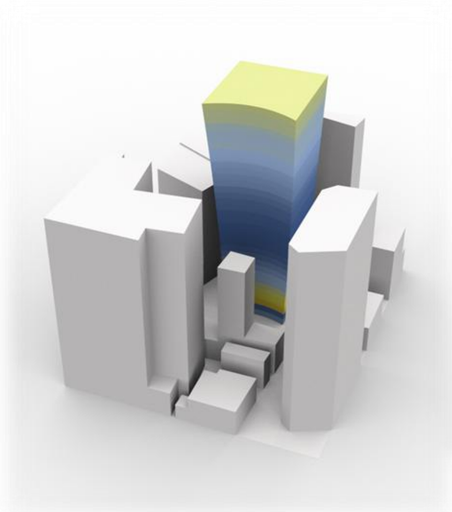
Economic growth

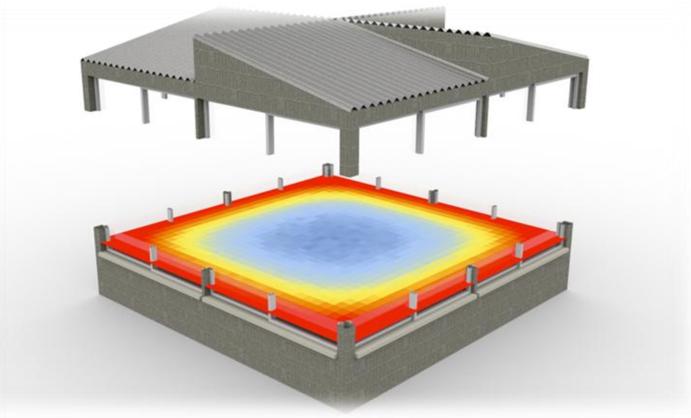
Sustainable goals



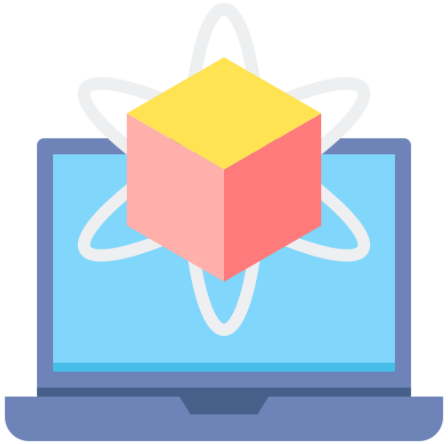


Building Performance Simulation tools

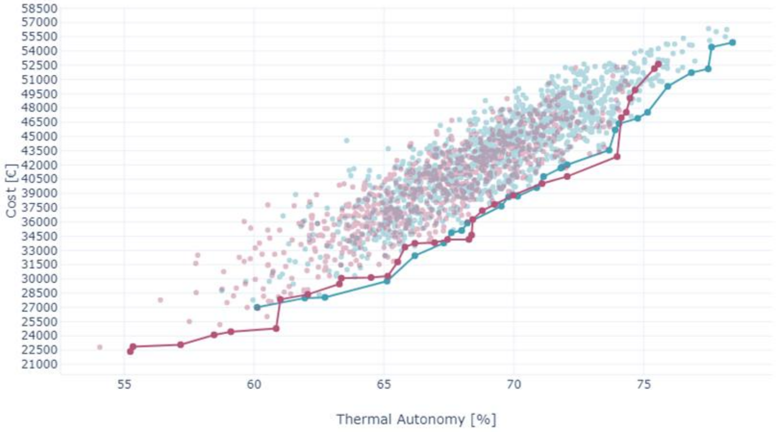


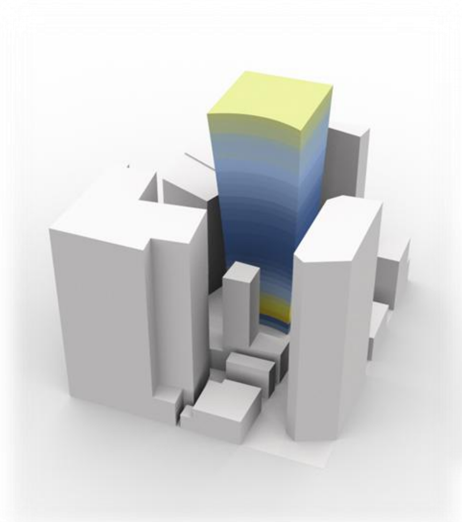


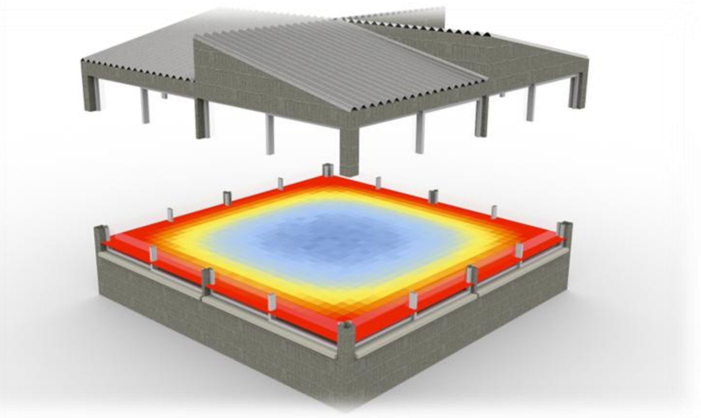
Analyses and Optimization Processes



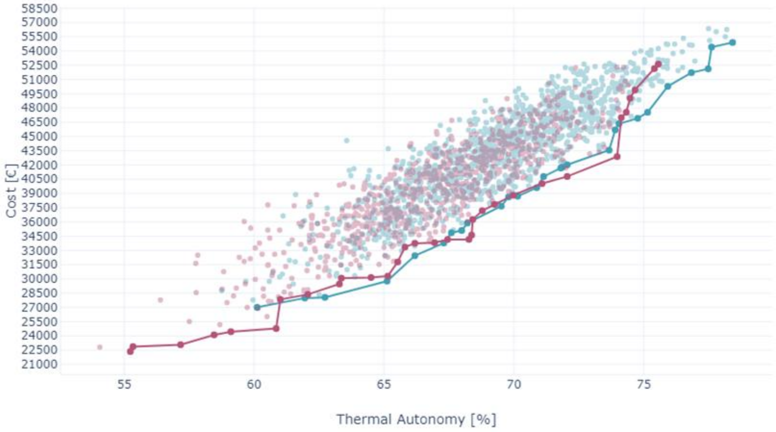
Building Performance Simulation tools

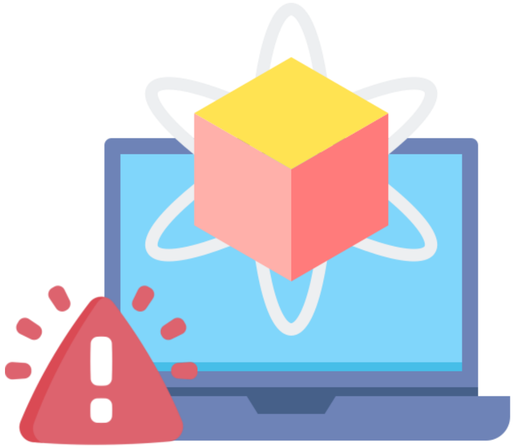






Analyses and Optimization Processes





Building Performance Simulation tools

Time consuming [1]

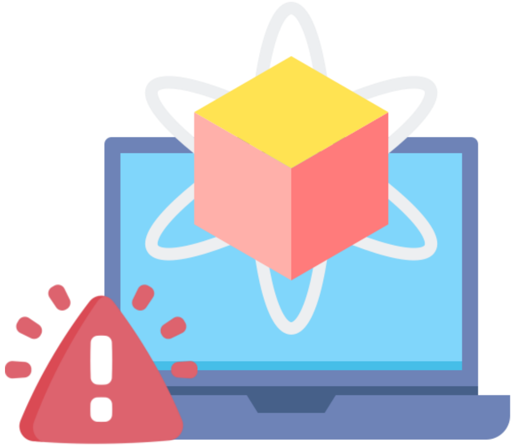
Portability [2]

Expertise [3]

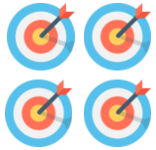
[1] Wei, T. (2013). A review of sensitivity analysis methods in building energy analysis. Renewable and Sustainable Energy Reviews, 20, 411–419.

[2] Crawley, D. B., Hand, J. W., Kummert, M., & Griffith, B. T. (2008). Contrasting the capabilities of building energy performance simulation programs. Building and Environment, 43(4), 661–673.

[3] Wang, H., & Zhai, Z. (John). (2016). Advances in building simulation and computational techniques: A review between 1987 and 2014. Energy and Buildings, 128, 319–335.



Analyses and Optimization Processes



Multiple Objectives [4, 5]

Multiple Algorithms [5, 6]

Building Performance Simulation tools

Time consuming [1]

Portability [2]

Expertise [3]

[4] Evins, R. (2013). A review of computational optimisation methods applied to sustainable building design. Renewable and Sustainable Energy Reviews, 22, 230–245.

[5] Nguyen, A., Reiter, S., & Rigo, P. (2014). A review on simulation-based optimization methods applied to building performance analysis. Applied Energy, 113, 1043–1058.

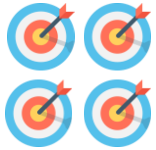
[6] Waibel, C., Wortmann, T., Evins, R., & Carmeliet, J. (2019). Building energy optimization: An extensive benchmark of global search algorithms. Energy and Buildings, 187, 218–240.

[1] Wei, T. (2013). A review of sensitivity analysis methods in building energy analysis. Renewable and Sustainable Energy Reviews, 20, 411–419.

[2] Crawley, D. B., Hand, J. W., Kummert, M., & Griffith, B. T. (2008). Contrasting the capabilities of building energy performance simulation programs. Building and Environment, 43(4), 661–673.

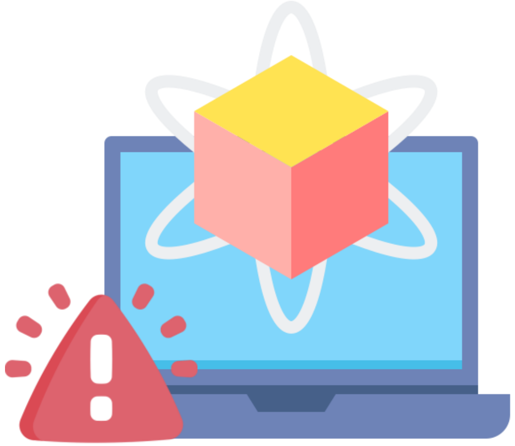
[3] Wang, H., & Zhai, Z. (John). (2016). Advances in building simulation and computational techniques: A review between 1987 and 2014. Energy and Buildings, 128, 319–335.

Analyses and Optimization Processes



Multiple Objectives [4, 5]

Multiple Algorithms [5, 6]



Building Performance Simulation tools

Time consuming [1]

Portability [2]

Expertise [3]

[4] Evins, R. (2013). A review of computational optimisation methods applied to sustainable building design. Renewable and Sustainable Energy Reviews, 22, 230–245.

[5] Nguyen, A., Reiter, S., & Rigo, P. (2014). A review on simulation-based optimization methods applied to building performance analysis. Applied Energy, 113, 1043–1058.

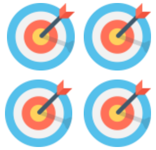
[6] Waibel, C., Wortmann, T., Evins, R., & Carmeliet, J. (2019). Building energy optimization: An extensive benchmark of global search algorithms. Energy and Buildings, 187, 218–240.

[1] Wei, T. (2013). A review of sensitivity analysis methods in building energy analysis. Renewable and Sustainable Energy Reviews, 20, 411–419.

[2] Crawley, D. B., Hand, J. W., Kummert, M., & Griffith, B. T. (2008). Contrasting the capabilities of building energy performance simulation programs. Building and Environment, 43(4), 661–673.

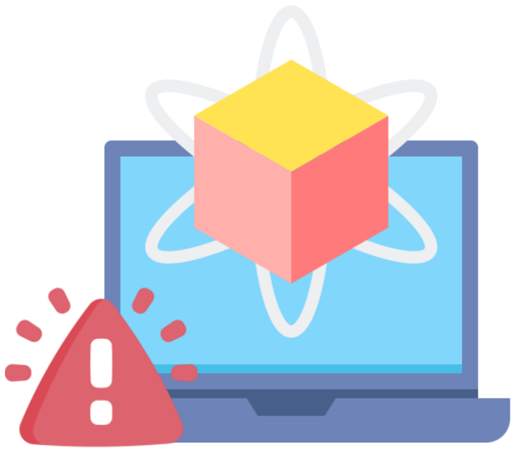
[3] Wang, H., & Zhai, Z. (John). (2016). Advances in building simulation and computational techniques: A review between 1987 and 2014. Energy and Buildings, 128, 319–335.

Analyses and Optimization Processes



Multiple Objectives [4, 5]

Multiple Algorithms [5, 6]



Building Performance Simulation tools

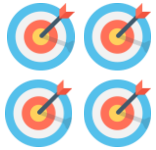
Time consuming [1]

Portability [2]

Expertise [3]

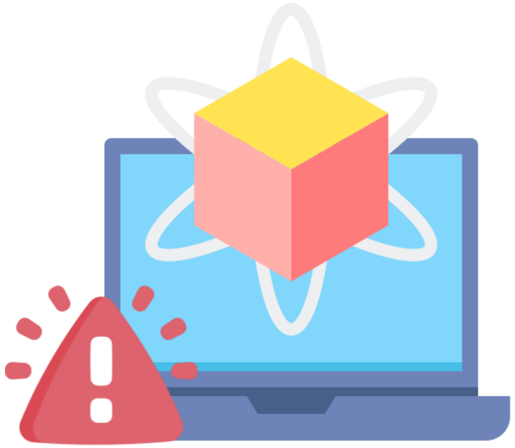
How to efficiently integrate AOP with building and urban projects?

Analyses and Optimization Processes



Multiple Objectives [4, 5]

Multiple Algorithms [5, 6]



Building Performance Simulation tools

Time consuming [1]

Portability [2]

Expertise [3]

How to efficiently integrate AOP with building and urban projects?

Make it quicker **Make it portable** **Make it easier**

Algorithmic Design and Analysis

Automates Design, Analysis, and Optimization processes. [7]

[7] Aguiar, R., Cardoso, C., & Leitão, A. (2017). Algorithmic design and analysis - fusing disciplines. Proceedings Catalog of the 37th Annual ACADIA 2017, 28–37.

[9] Araújo, G., Pereira, I., Leitão, A., & Correia Guedes, M. (2021). Conflicts in passive building performance: Retrofit and regulation of informal neighbourhoods. Frontiers of Architectural Research, 10(3), 625–638.

Algorithmic Design and Analysis

Automates Design, Analysis, and Optimization processes. [7]

BPS are still time-consuming for large models. [8]

Additional expertise to learn a programming language. [7]

[7] Aguiar, R., Cardoso, C., & Leitão, A. (2017). Algorithmic design and analysis - fusing disciplines. Proceedings Catalog of the 37th Annual ACADIA 2017, 28–37.

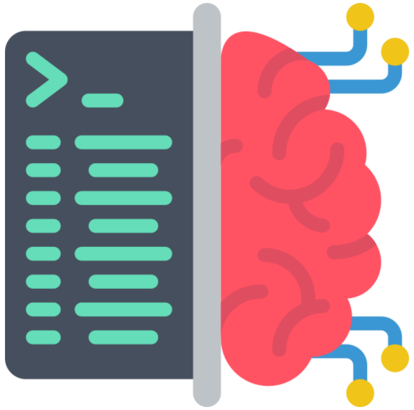
[9] Araújo, G., Pereira, I., Leitão, A., & Correia Guedes, M. (2021). Conflicts in passive building performance: Retrofit and regulation of informal neighbourhoods. Frontiers of Architectural Research, 10(3), 625–638.

Algorithmic Design and Analysis

Automates Design, Analysis, and Optimization processes. [7]

BPS are still time-consuming for large models. [8]

Additional expertise to learn a programming language. [7]



Surrogate Models

Quickly predict BPS outputs with fewer inputs. [10]

[7] Aguiar, R., Cardoso, C., & Leitão, A. (2017). Algorithmic design and analysis - fusing disciplines. Proceedings Catalog of the 37th Annual ACADIA 2017, 28–37.

[9] Araújo, G., Pereira, I., Leitão, A., & Correia Guedes, M. (2021). Conflicts in passive building performance: Retrofit and regulation of informal neighbourhoods. Frontiers of Architectural Research, 10(3), 625–638.

[10] 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, 29(4), 471–481.

[11] Alizadeh, R., Allen, J. K., & Mistree, F. (2020). Managing computational complexity using surrogate models: a critical review. Research in Engineering Design, 31(3), 275–298.

Algorithmic Design and Analysis

Automates Design, Analysis, and Optimization processes. [7]

BPS are still time-consuming for large models. [8]

Additional expertise to learn a programming language. [7]

Surrogate Models

Quickly predict BPS outputs with fewer inputs. [10]

Are usually case-specific and do not apply outside the study's boundaries. [11]

Reduces BPS expertise but requires expertise to develop and test the models. [10]

[7] Aguiar, R., Cardoso, C., & Leitão, A. (2017). Algorithmic design and analysis - fusing disciplines. Proceedings Catalog of the 37th Annual ACADIA 2017, 28–37.

[9] Araújo, G., Pereira, I., Leitão, A., & Correia Guedes, M. (2021). Conflicts in passive building performance: Retrofit and regulation of informal neighbourhoods. Frontiers of Architectural Research, 10(3), 625–638.

[10] 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, 29(4), 471–481.

[11] Alizadeh, R., Allen, J. K., & Mistree, F. (2020). Managing computational complexity using surrogate models: a critical review. Research in Engineering Design, 31(3), 275–298.

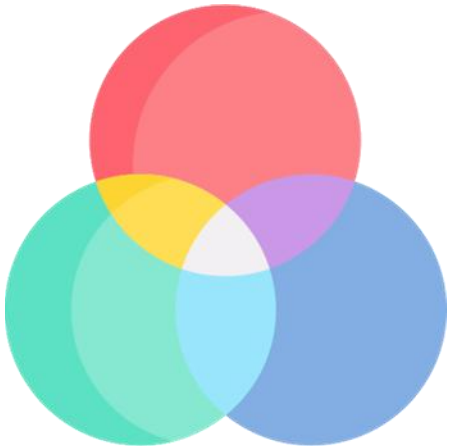
Flexible framework to develop Surrogate Models and integrate them with AOP.

Flexible framework to develop Surrogate Models and integrate them with AOP.

Synthetic

Iterative

Existing

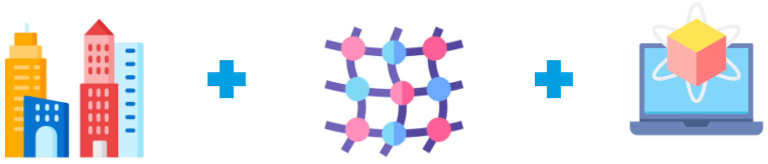


Synthetic

$$f(x_1, x_2, ..., x_n) = y$$

$$f\left(\begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix}\right) = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

where m represents combinations of the features' domain



Building Archetypes + Features' domain + Simulation

We can develop multiple building models that vary along specified domains of different features, simulate them, and build a database.

Comprehensive feedback of the features impact in simulation results.

Surrogate Models for multiple analysis and optimization problems.

Highly detailed database.

Simulation and computation times exponentially increase with the number of features and simulation types, to the point of being unfeasible.

Iterative

$$f(x_1, x_2, ..., x_i) = min(o_1, o_2, ..., o_j)$$

$$f\left(\begin{bmatrix} x_{11} & \cdots & x_{1i} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mi} \end{bmatrix}\right) = \begin{bmatrix} o_{11} & \cdots & o_{1j} \\ \vdots & \ddots & \vdots \\ o_{m1} & \cdots & o_{mj} \end{bmatrix}$$

where m represents the number of iterations in the optimization

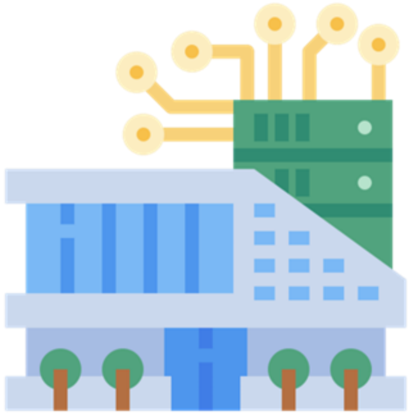
We can develop building optimization problems and build a database from the explored solutions' variables and objectives.

Less computational time than feature domains depending on the number of iterations.

Can support a high number of features.

Can support a high number of objectives.

Specific to the optimization problem at hand.



Existing

Extract

Cleaning

Analyses and Optimization

Extract building features from existing building databases and develop surrogate models for multiple AOP.

Database with real values.

Might not require simulations.

Surrogate Models for multiple analysis and optimization problems.

Noisy and imbalanced data.

Often does not feature many objectives.

Surrogate Models

Model type

Model tuning

Model deployment

Model type

Classification Models

Regression Models

Model type

Key Performance Indicators

Accuracy

Precision

Recall

F1-score

Classification Models

Regression Models

Models that predict **discrete** target outputs.

Neural Networks

Logistic regression

Ensemble models

Model type

Key Performance Indicators

Accuracy	Mean Absolute Error (MAE)
Precision	Mean Squared Error (MSE)
Recall	Mean Average Percent Error (MAPE)
F1-score	Coefficient of determination (R^2 score)

Classification Models

Models that predict **discrete** target outputs.

Neural Networks

Logistic regression

Ensemble models

Regression Models

Models that predict **continuous** target outputs.

Neural Networks

Linear regression

Ensemble models

Interpolation

Model tuning

Key Performance Indicators

Accuracy	Mean Absolute Error (MAE)
Precision	Mean Squared Error (MSE)
Recall	Mean Average Percent Error (MAPE)
F1-score	Coefficient of determination (R^2 score)

Model tuning

Key Performance Indicators

Accuracy	Mean Absolute Error (MAE)
Precision	Mean Squared Error (MSE)
Recall	Mean Average Percent Error (MAPE)
F1-score	Coefficient of determination (R^2 score)

Feature Engineering

Hyperparameter optimization

Model tuning

Key Performance Indicators

Accuracy	Mean Absolute Error (MAE)
Precision	Mean Squared Error (MSE)
Recall	Mean Average Percent Error (MAPE)
F1-score	Coefficient of determination (R^2 score)

Feature Engineering

Hyperparameter optimization

Clean, **change**, and **develop** features

Clean outliers

Create new features

Select relevant features

Model tuning

Key Performance Indicators

Accuracy	Mean Absolute Error (MAE)
Precision	Mean Squared Error (MSE)
Recall	Mean Average Percent Error (MAPE)
F1-score	Coefficient of determination (R^2 score)

Feature Engineering

Clean, **change**, and **develop** features

Clean outliers

Create new features

Select relevant features

Hyperparameter optimization

Optimize the model's **parameters** for best suitable **KPI**

Neural Networks

Solver

Learning rate

Hidden layer sizes

Linear regression

Polynomial degree

Ensemble models

Number of estimators

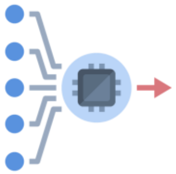
Depth

Samples split

Interpolation

Solver

Model deployment



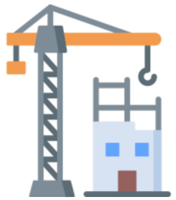
Inputs

Outputs

User interactions

Case studies with different **AOP** and **BID**

Design



Construction

Retrofit

Synthetic

Optimization

Existing

Model deployment

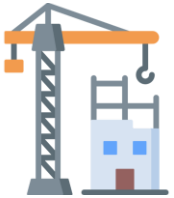
3D Models

Programming

Web App

Case studies with different AOP and BID

Design



Construction

Retrofit

Synthetic

Optimization

Existing

Case Studies

Surrogate Models to improve Building Performance Analysis and Optimization

Synthetic

Retrofit

Design

Fundação para a Ciência e a Tecnologia

TÉCNICO LISBOA

CENTER for INNOVATION, TECHNOLOGY and POLICY RESEARCH

Construction period	Number	Wall U_Value (kWh/m²)	Roof U_Value (W/m²·°C)	Floor U_Value (W/m²·°C)	Window U_Value (W/m²·°C)	Wall retrofit U-value (W/m²·°C)	Roof retrofit U-value (W/m²·°C)
<1919	1	2.78	1.99	1.80	2.69	0.61	0.63
1919-1945	2	2.78	1.99	1.80	2.69	0.61	0.63
1946-1960	3	1.49	1.99	1.80	2.69	0.57	0.63
1961-1970	4	1.08	1.99	3.03	2.69	0.49	0.63
1971-1980	5	1.26	1.99	3.03	2.69	0.53	0.63
1981-1990	6	0.50	1.99	3.03	2.69	0.32	0.63
1991-1995	7	0.49	1.99	3.03	2.69	0.32	0.63
1996-2000	8	0.46	1.99	2.31	2.69	0.29	0.63
2001-2005	9	0.25	1.99	2.31	2.69	0.19	0.63
>2006	10	0.25	1.99	2.31	2.69	0.19	0.63

Urban area in Lisbon

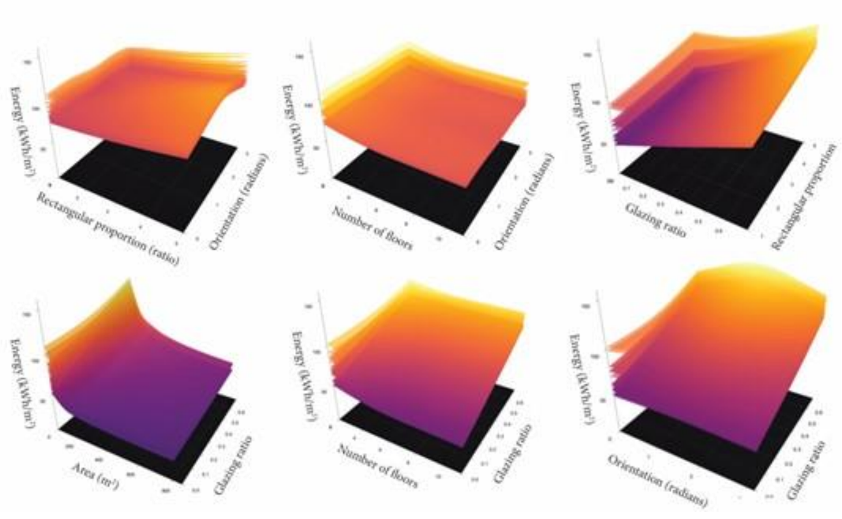
Data is retrieved from GIS file by extraction of the buildings':

Geometry

Number of Floors

Age

Glazing ratio

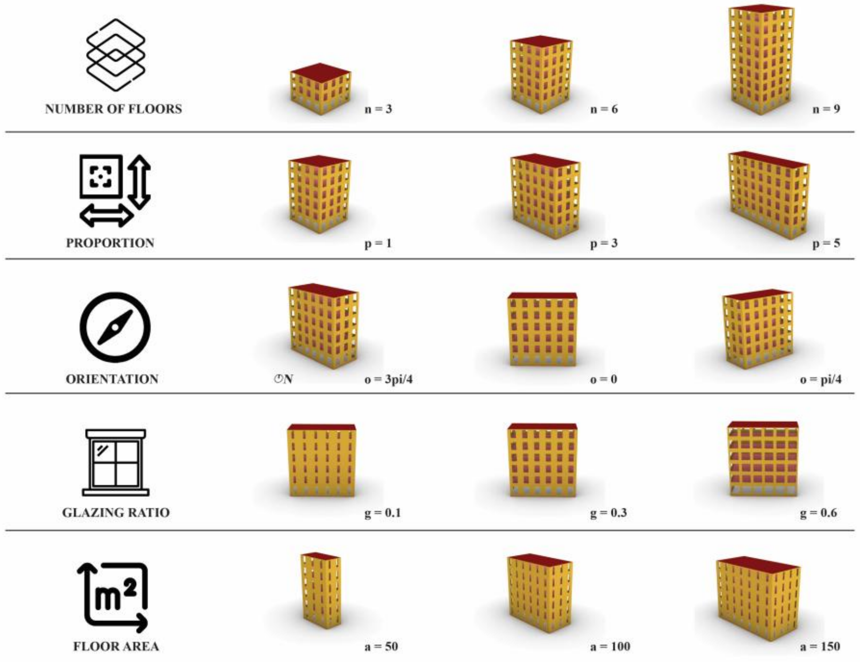


$$E(c, n, p, o, g, a) = Annual\ Energy\ Loads\ [kWh/m^2]$$

$$f\left(\begin{bmatrix} c_1 & n_1 & p_1 & o1 & g_1 & a_1 \\ c_1 & n_1 & p_1 & o1 & g_1 & a_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c_i & n_i & p_i & oi & g_i & a_i \end{bmatrix}\right) = \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_i \end{bmatrix}$$

Synthetic

Create a dataset that encompasses a grid-based set of feature values.



Linear regression

Interpolation

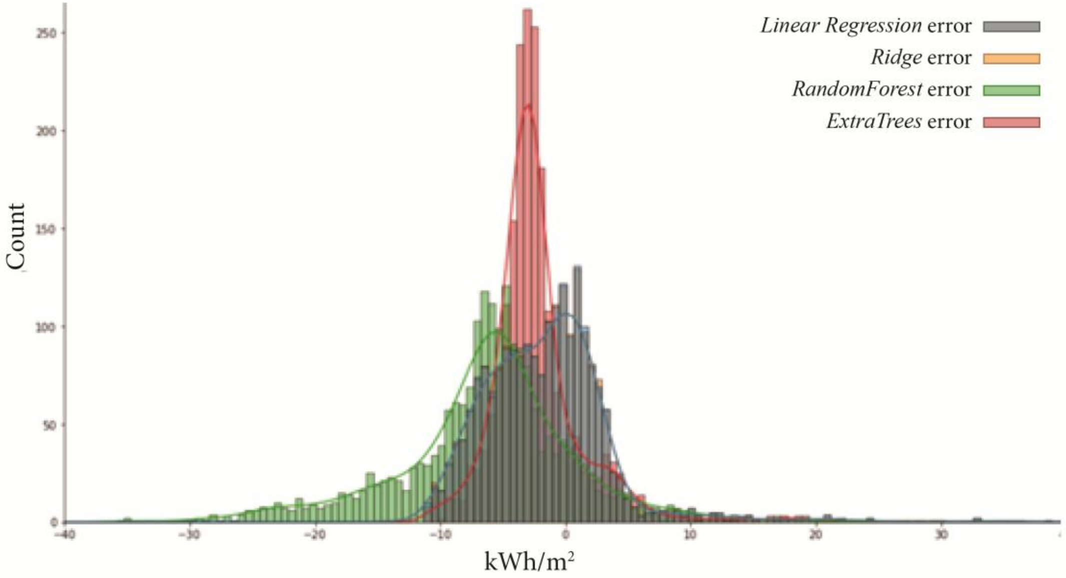
Ensemble models

	Linear Regression	Ridge	Random Forest	Extra Trees
Mean Error (kWh/m^2)	15.85	15.79	-6.06	-2.21
Root Mean Squared Error (kWh/m^2)	15.40	15.40	9.88	5.44
R^2 score	0.64	0.64	0.85	0.95

Regression Models

Results are used to train Surrogate models.

The case study simulation results are used to validate and select the best model:



Linear regression

Interpolation

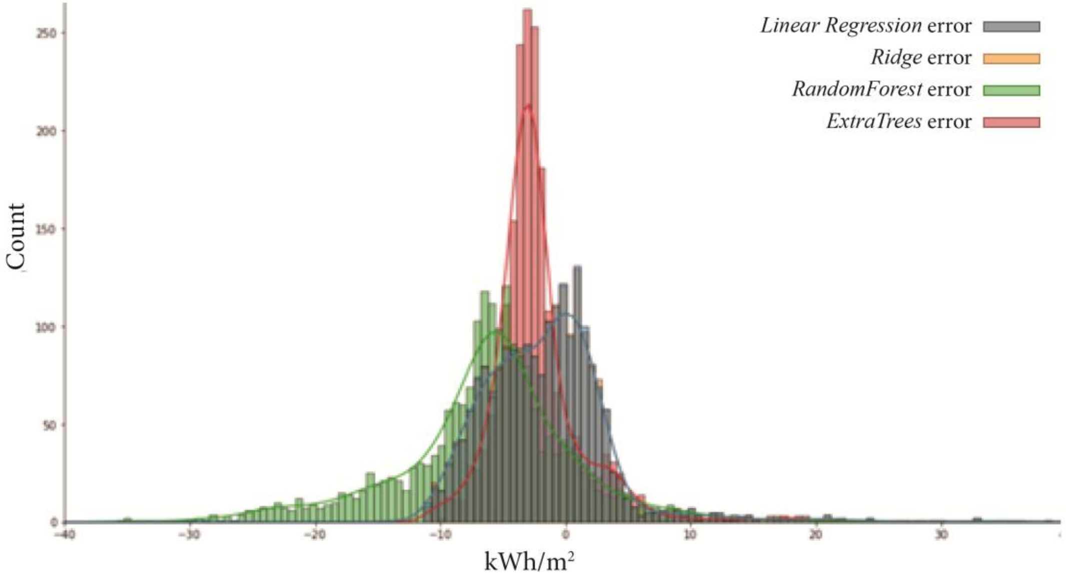
Ensemble models

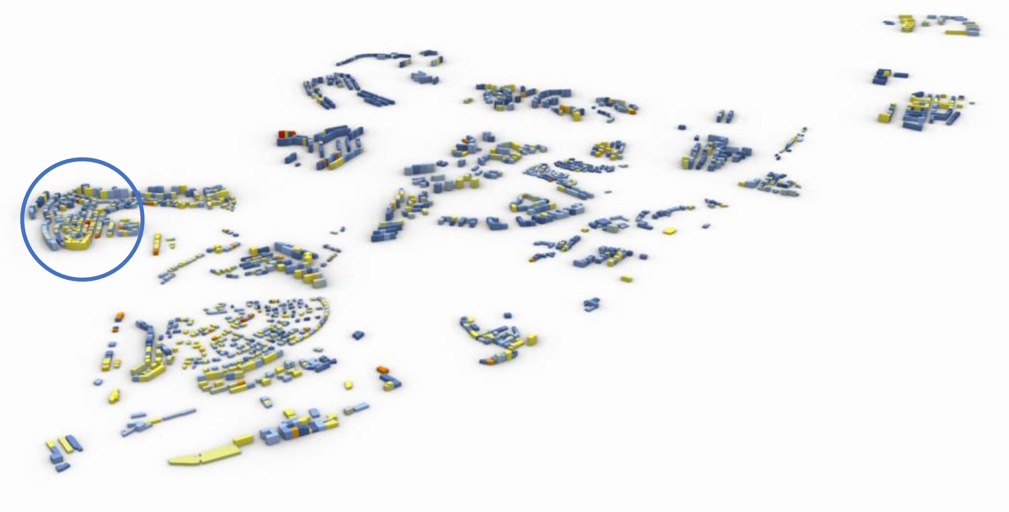
	Linear Regression	Ridge	Random Fores	Extra Trees
Mean Error (kWh/m^2)	15.85	15.79	-6.06	-2.21
Root Mean Squared Error (kWh/m^2)	15.40	15.40	9.88	5.44
R^2 score	0.64	0.64	0.85	0.95

Regression Models

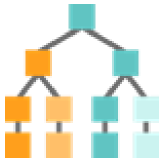
Results are used to train Surrogate models.

The case study simulation results are used to validate and select the best model:





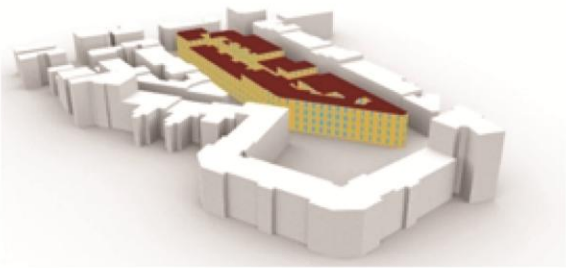
ExtraTrees regressor

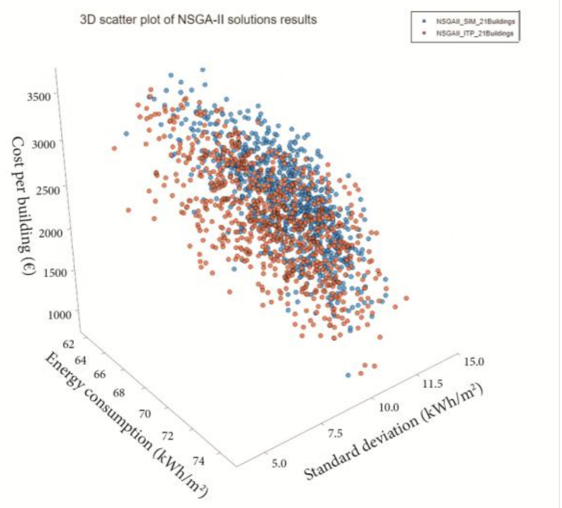


$$\frac{\sum_{i=1}^{n} annual\ loads_i}{n}\ [kWh/m^2]$$

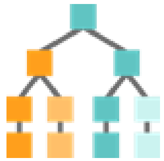
$$\sigma\left(\bigcup_{i=1}^{n} annual\ loads_i\right)\ [kWh/m^2]$$

$$\frac{\sum_{i=1}^{n} Cost_i}{n}\ [€]$$





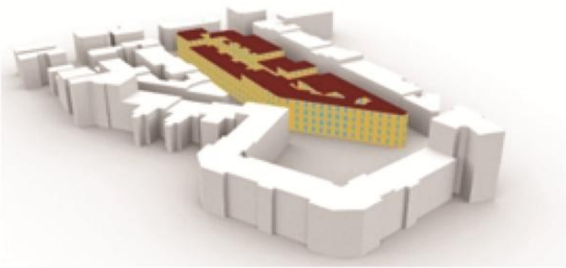
ExtraTrees regressor

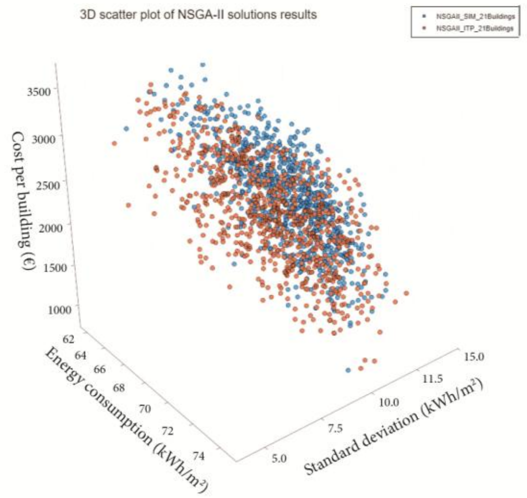


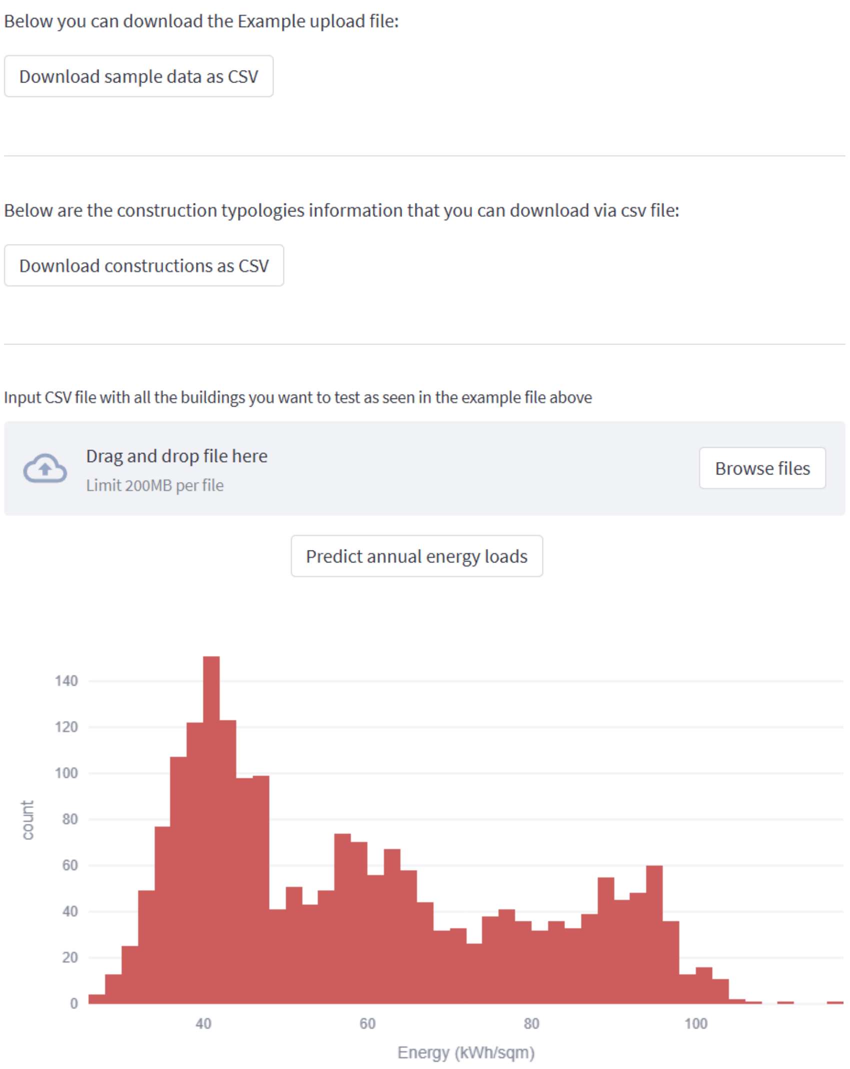
Speed up
factor of 85x

	elapsed time (seconds)	
	Dataset simulation	Optimization
Surrogate model	0.08	791.99
Simulation	5820.00	67516.70

$$\frac{\sum_{i=1}^{n} annual\ loads_i}{n}\ [kWh/m^2] \qquad \sigma\left(\bigcup_{i=1}^{n} annual\ loads_i\right)\ [kWh/m^2] \qquad \frac{\sum_{i=1}^{n} Cost_i}{n}\ [€]$$







ExtraTrees regressor

Model is deployed in a web app prototype

Prediction of input buildings' energy use

Optimization of building design

Building design Optimization

In this section you can select the design variables and their boundaries for a building and optimize its values for minimum annual energy loads

Select the design variables you wish to optimize:

Number of floors × Rectangular pro... × Orientation (Rad... × Floor area (sqm) ×

Number of floors - boundaries

Minimum number of floors:

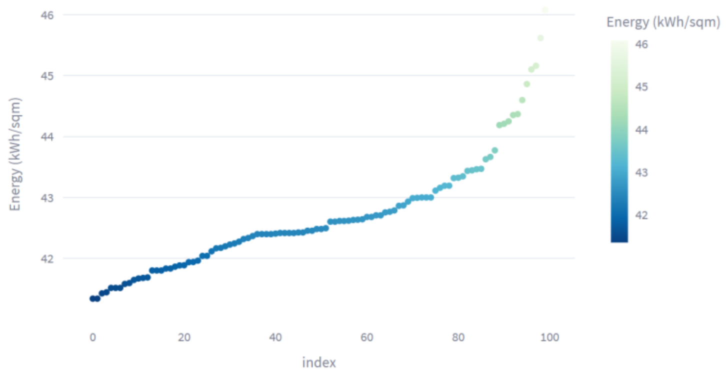
1

1 20

Maximum number of floors:

10

1 20



ExtraTrees regressor

Model is deployed in a web app prototype

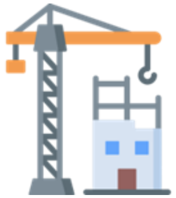
Prediction of input buildings' energy use

Optimization of building design

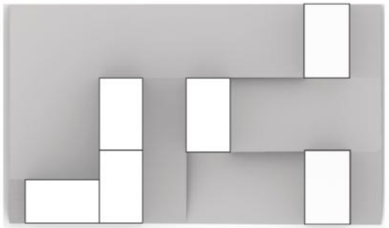
Results

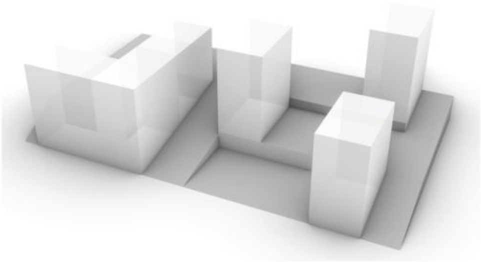
Surrogate Models to improve Building Performance Analysis and Optimization

Iterative



Construction





Optimization of construction materials for a 6 building block design

Iterative

Use previous iterations of a simulation-based optimization to train surrogate models capable of predicting the specified objectives

Variables:

$$(w, r, f, w_i) \in \{0, 1, 2\}$$

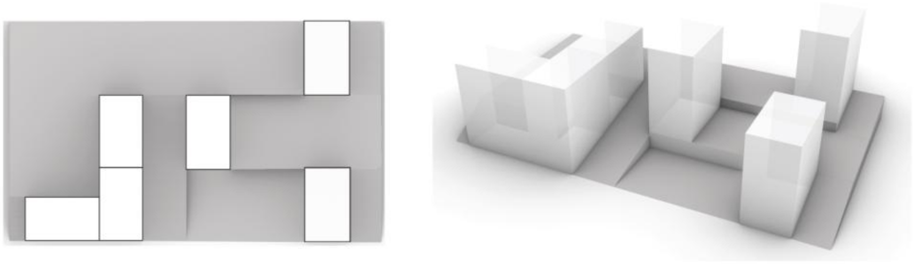
$w \rightarrow wall\ possible\ constructions$

$r \rightarrow roof\ possible\ constructions$

$f \rightarrow floor\ possible\ constructions$

$w_i \rightarrow window\ possible\ constructions$

$$g\left(\begin{bmatrix} (w,r,f,w_i)_{11} & \cdots & (w,r,f,w_i)_{1n} \\ \vdots & \ddots & \vdots \\ (w,r,f,w_i)_i & \cdots & (w,r,f,w_i)_{in} \end{bmatrix}\right) = \begin{bmatrix} f_{1_1} & f_{2_1} & f_{3_1} \\ \vdots & \vdots & \vdots \\ f_{1_i} & f_{2_i} & f_{3_i} \end{bmatrix}$$



Optimization of construction materials for a 6 building block design

Objectives:

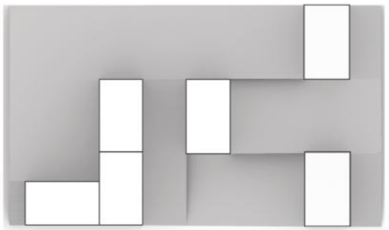
$$\boxed{\begin{aligned} f_1(w,r,f,w_i) &= \sum_{i=1}^{n} Heating_i + Cooling_i \\ f_2(w,r,f,w_i) &= \sigma\left(\bigcup_{i=1}^{n} Heating_i + Cooling_i\right) \end{aligned}}$$

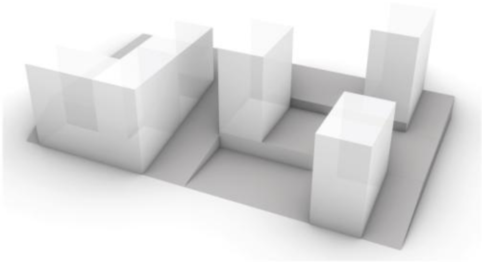
$$f_3(w,r,f,w_i) = \sum_{i=1}^{n} Cost_i$$

Iterative

Use previous iterations of a simulation-based optimization to train surrogate models capable of predicting the specified objectives

$$g\left(\begin{bmatrix} (w,r,f,w_i)_{11} & \cdots & (w,r,f,w_i)_{1n} \\ \vdots & \ddots & \vdots \\ (w,r,f,w_i)_i & \cdots & (w,r,f,w_i)_{in} \end{bmatrix}\right) = \begin{bmatrix} f_{1_1} & f_{2_1} & f_{3_1} \\ \vdots & \vdots & \vdots \\ f_{1_i} & f_{2_i} & f_{3_i} \end{bmatrix}$$



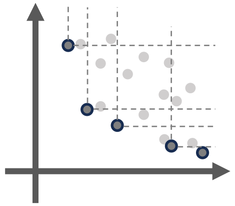


Optimization of construction materials for a 6 building block design

Objectives:

$$f_4\ (i_0, \ldots,\ i_n) = R^2(test,\ predictions)$$

$$i \in\ [6\ ,300]\ \text{—}Number\ of\ filters$$
$$n = Number\ of\ layers$$



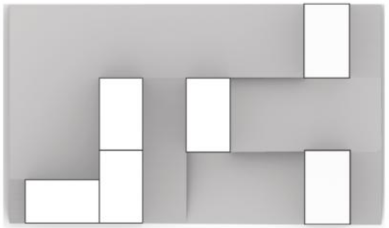
$$H(f_1, f_2, f_3) \qquad f_1\ \in [\min(f_1)\ ,\max(f_1)]$$
$$f_2\ \in [\min(f_2)\ ,\max(f_2)]$$
$$f_3\ \in [\min(f_3)\ ,\max(f_3)]$$

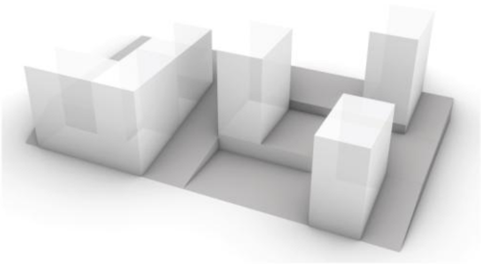
Hyperparameter optimization

Maximize a Neural Network's R^2 score and the optimization algorithms' Hypervolume of non-dominated solutions.

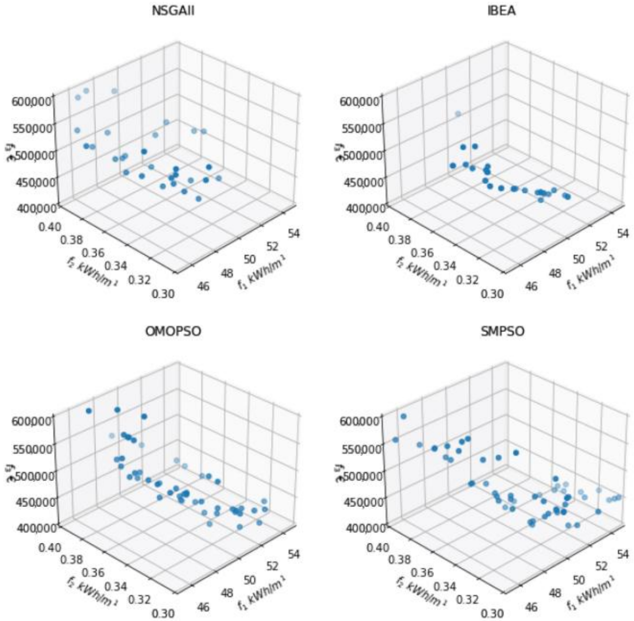
Neural Network Sequential model

Metaheuristics



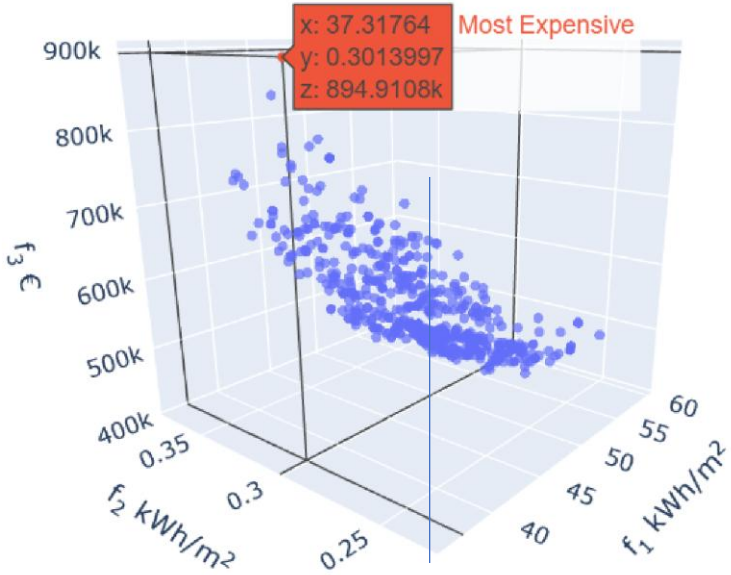


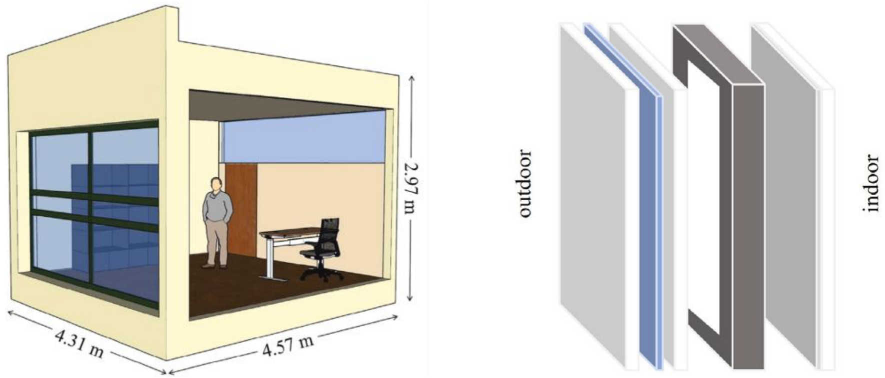
Optimization of construction materials for a 6 building block design



Neural Network Sequential model

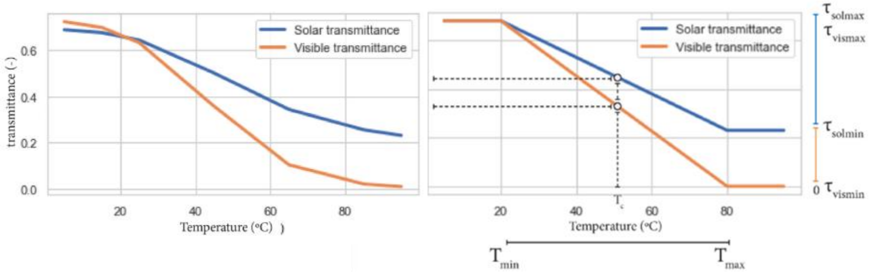
Model is deployed in a programming environment where a user can upload a building design solution, specify the variables and run the optimization with different algorithms





Optimization of Thermochromic glazing properties for an office space

Variables:

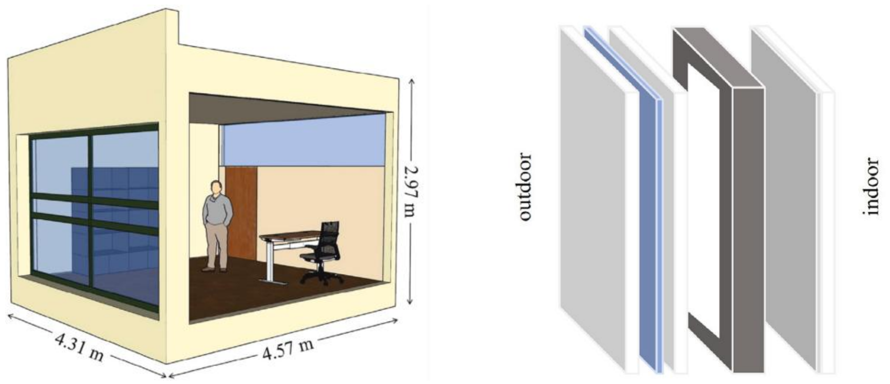


$$\tau_{sol}\ (T,\ T_{min},\ T_{max},\ \tau_{solmax}, \tau_{solmin}) = \\ = \begin{cases} \tau_{solmax} & \text{if } T \leq\ T_{min} \\ \tau_{solmax} - \frac{\tau_{solmax} - \tau_{solmin}}{(T_{max} - T_{min})} \times (T - T_{min}) & \text{if } T_{min} < T < T_{max} \\ \tau_{solmin} & \text{if } T \geq\ T_{max} \end{cases}$$

Iterative

Use previous iterations of a simulation-based optimization to train surrogate models capable of predicting the specified objectives

$$g\left(\begin{bmatrix} T_{max_1} & T_{min_1} & \tau_{min_1} & \tau_{max_1} \\ T_{max_2} & T_{min_2} & \tau_{min_2} & \tau_{max_2} \\ \vdots & \vdots & \vdots & \vdots \\ T_{max_i} & T_{min_i} & \tau_{min_i} & \tau_{max_i} \end{bmatrix}\right) = \begin{bmatrix} f_{11} & f_{21} \\ f_{12} & f_{22} \\ \vdots & \vdots \\ f_{1_i} & f_{2_i} \end{bmatrix}$$



Optimization of Thermochromic glazing properties for an office space

Objectives:

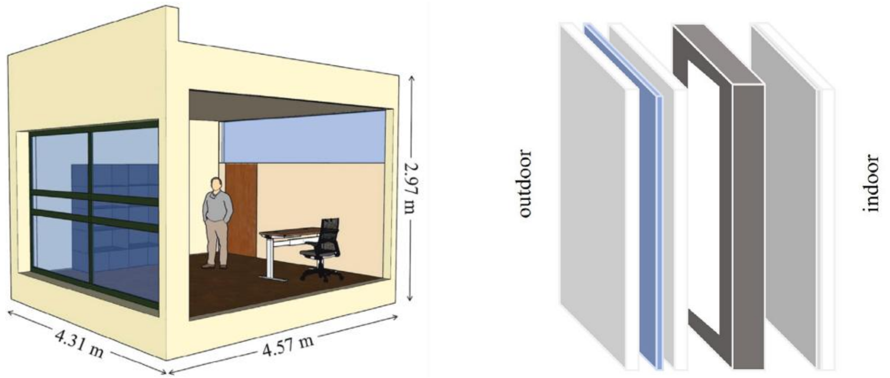
$$f_1(T_{min}T_{max}\tau_{min}\tau_{max}) = Heating + Cooling\ [kWh/m^2]$$

$$f_2(T_{min}T_{max}\tau_{min}\tau_{max}) = Lighting\ [kWh/m^2]$$

Iterative

Use previous iterations of a simulation-based optimization to train surrogate models capable of predicting the specified objectives

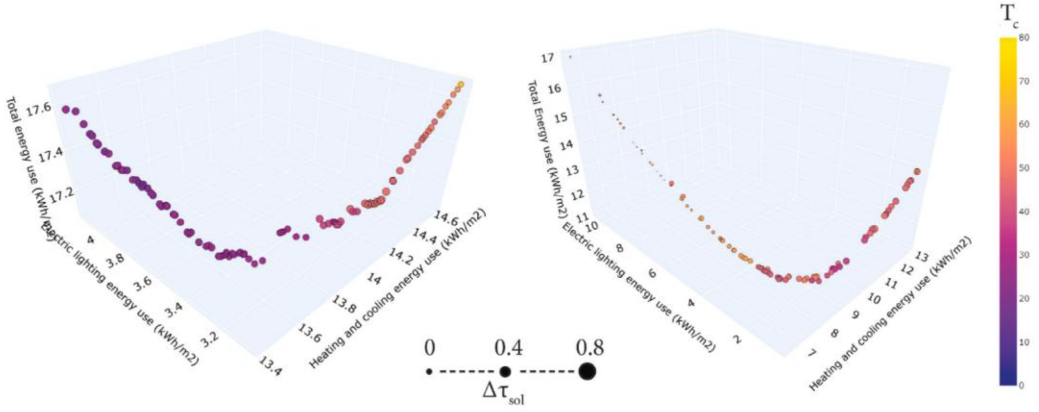
$$g\left(\begin{bmatrix} T_{max_1} & T_{min_1} & \tau_{min_1} & \tau_{max_1} \\ T_{max_2} & T_{min_2} & \tau_{min_2} & \tau_{max_2} \\ \vdots & \vdots & \vdots & \vdots \\ T_{max_i} & T_{min_i} & \tau_{min_i} & \tau_{max_i} \end{bmatrix}\right) = \begin{bmatrix} f_{11} & f_{21} \\ f_{12} & f_{22} \\ \vdots & \vdots \\ f_{1_i} & f_{2_i} \end{bmatrix}$$



Optimization of Thermochromic glazing properties for an office space

ExtraTrees regressor

Model is deployed in a programming environment where the user can upload weather file, specify variable boundaries and run the optimizations



$$g\left(\begin{bmatrix} T_{max_1} & T_{min_1} & \tau_{min_1} & \tau_{max_1} \\ T_{max_2} & T_{min_2} & \tau_{min_2} & \tau_{max_2} \\ \vdots & \vdots & \vdots & \vdots \\ T_{max_i} & T_{min_i} & \tau_{min_i} & \tau_{max_i} \end{bmatrix}\right) = \begin{bmatrix} f_{1_1} & f_{2_1} \\ f_{1_2} & f_{2_2} \\ \vdots & \vdots \\ f_{1_i} & f_{2_i} \end{bmatrix}$$

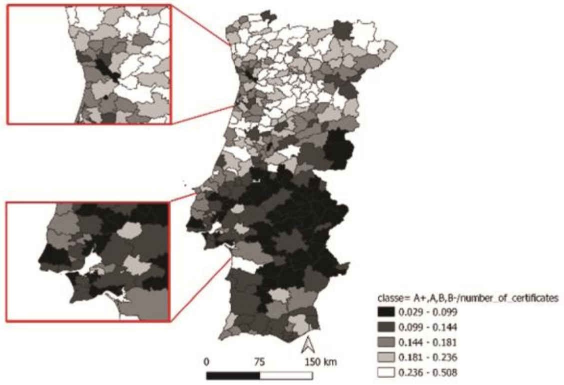
Results

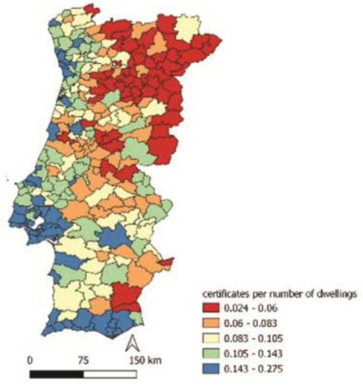
Surrogate Models to improve Building Performance Analysis and Optimization

Existing

Retrofit

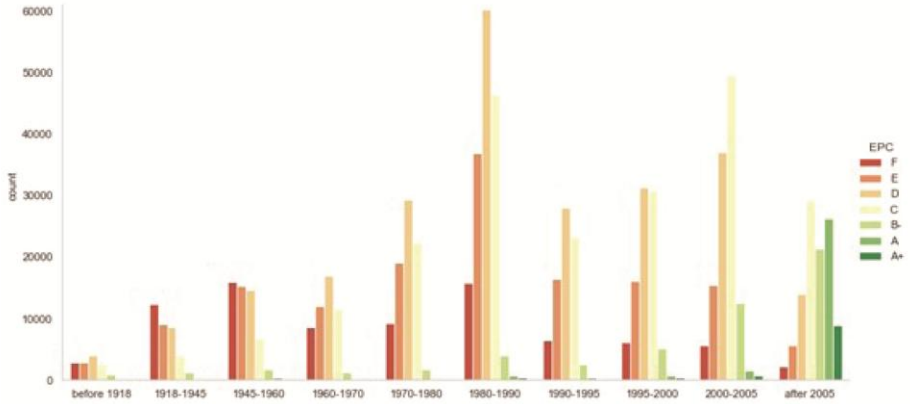
Design



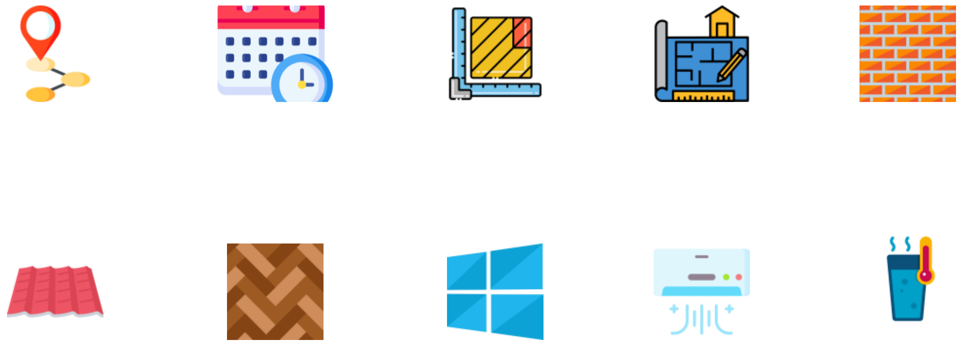


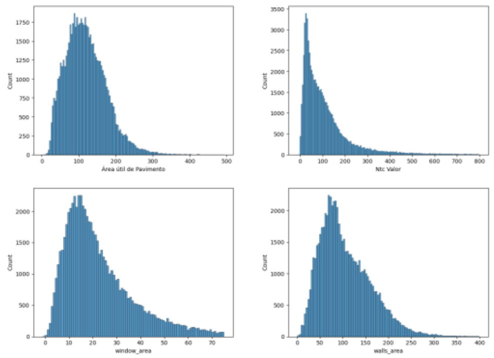
Portugal's Energy Performance Certificates (EPC)

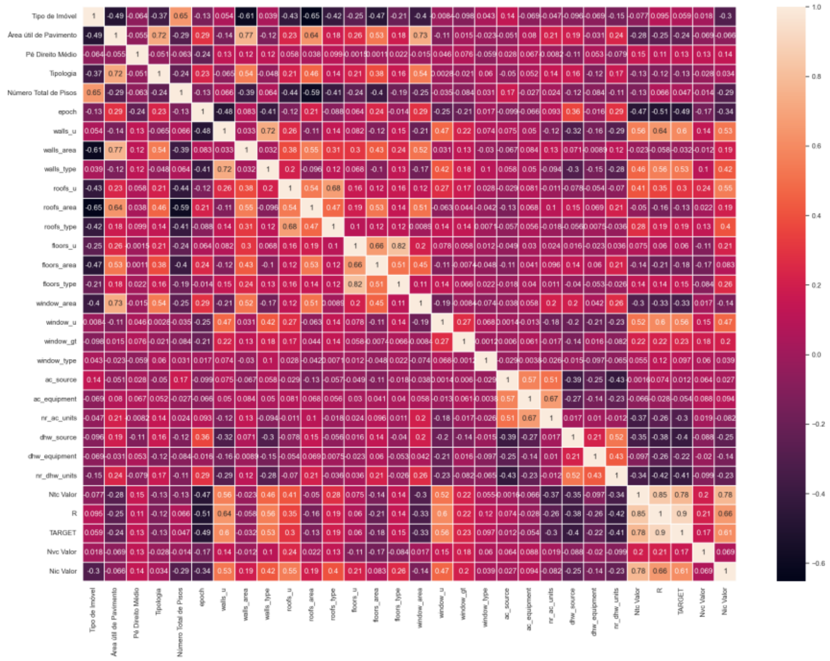
Data is retrieved from csv file by extraction of the certificates' features:



EPC labels histogram by construction period







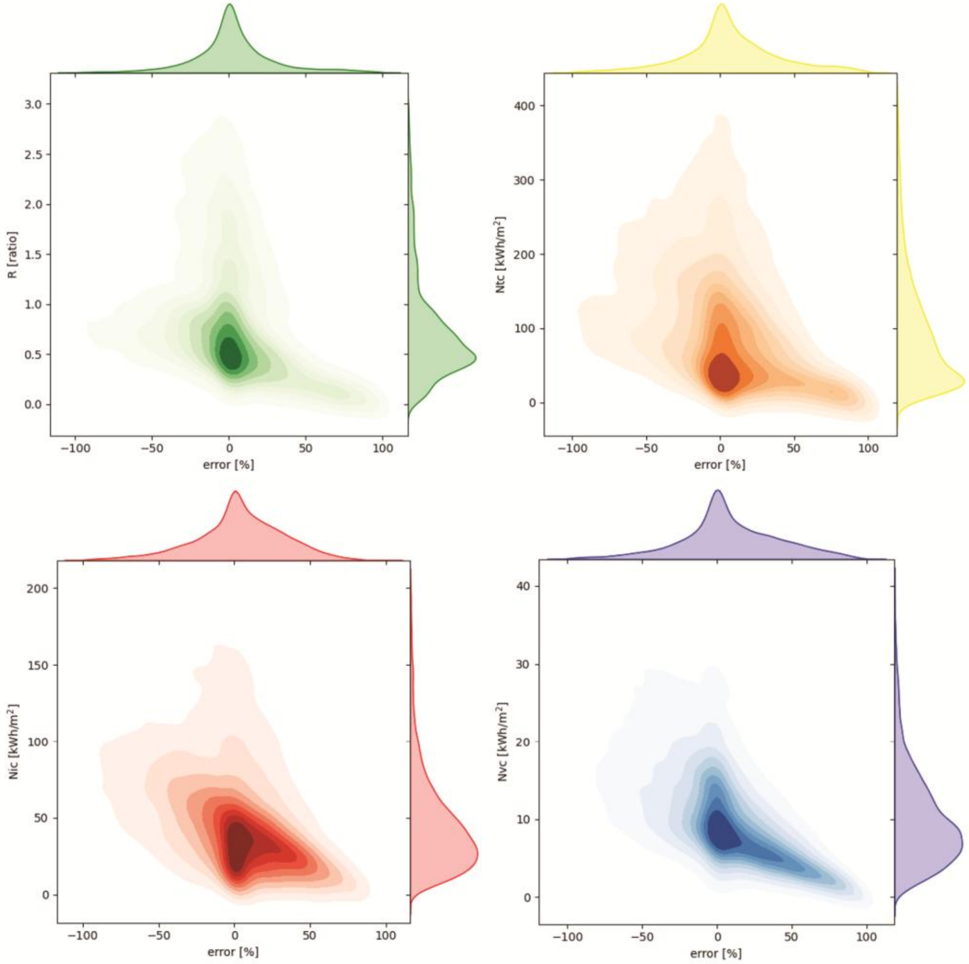
Feature Engineering

Apply feature engineering to obtain a balanced database for both features and prediction targets:

Create new features

Clean outliers

Select relevant features

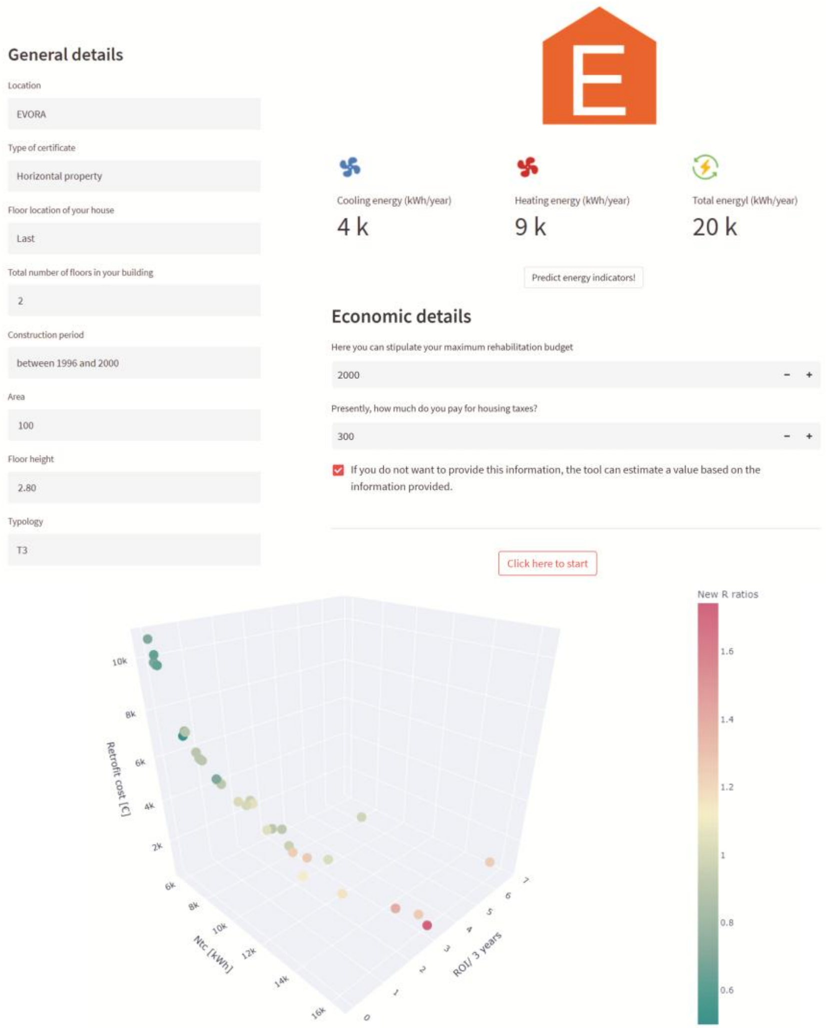


Regression Models

Results are used to train Surrogate models and select the best one between the lower number of features and higher R^2 score.

Model training and performance indicators results.

		R [ratio]		Ntc [kWh/m²]		Nic [kWh/m²]		Nvc [kWh/m²]	
k-best features	Model	R^2	$RMSE$	R^2	$RMSE$	R^2	$RMSE$	R^2	$RMSE$
10	ET	0.74	0.26	0.58	55.70	0.61	19.11	0.31	5.85
	MLP	0.71	0.27	0.51	60.22	0.55	20.46	0.16	6.43
	GB	0.69	0.28	0.50	60.98	0.53	20.89	0.14	6.51
15	ET	0.82	0.22	0.72	45.30	0.65	18.07	0.36	5.63
	MLP	0.77	0.24	0.66	49.91	0.57	20.05	0.19	6.33
	GB	0.76	0.25	0.65	51.07	0.57	20.01	0.17	6.39
20	ET	0.84	0.21	0.79	39.77	0.67	17.70	0.41	5.41
	MLP	0.78	0.24	0.72	45.29	0.58	19.77	0.23	6.15
	GB	0.78	0.24	0.73	44.87	0.57	19.96	0.24	6.14
25	ET	0.85	0.20	0.80	38.83	0.73	15.96	0.61	4.40
	MLP	0.80	0.23	0.74	44.33	0.67	17.66	0.47	5.11
	GB	0.80	0.23	0.75	43.19	0.68	17.22	0.33	5.75



ExtraTrees Regressor with 20 features

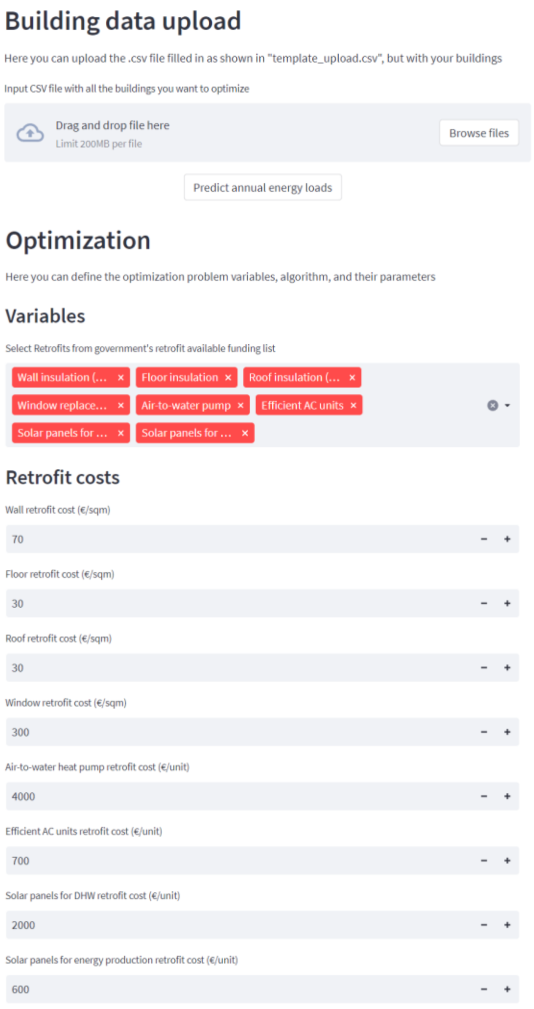
Model is deployed in a web app prototype for the optimization of building retrofit for:

Homeowners

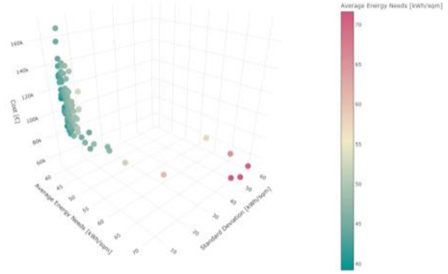
Policymaking

$$f\left(\begin{bmatrix} r_1 & \cdots & r_n \\ \vdots & \ddots & \vdots \\ r_i & \cdots & r_{in} \end{bmatrix}\right) = \begin{bmatrix} f_{1_1} & f_{2_1} & f_{3_1} \\ \vdots & \vdots & \vdots \\ f_{1_i} & f_{2_i} & f_{3_i} \end{bmatrix}$$

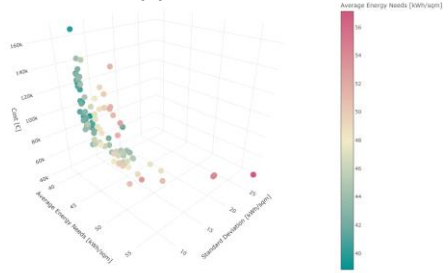
$f_1 = Ntc\ [kWh/m^2]$ $f_2 = Return\ on\ investment\ [ratio]$ $f_3 = Retrofit\ cost\ [€]$



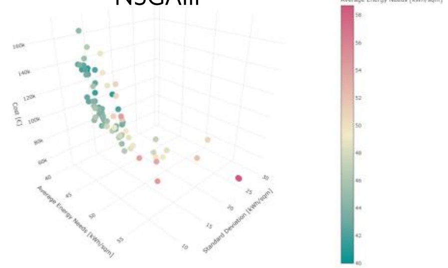
SPEA II



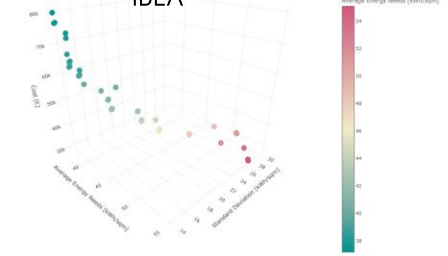
NSGAII



NSGAIII



IBEA



ExtraTrees Regressor with 20 features

Model is deployed in a web app prototype for the optimization of building retrofit for:

Homeowners

Policymaking

$$f\left(\begin{bmatrix}(r_1,\dots,r_n)_1 & \cdots & (r_1,\dots,r_n)_m \\ \vdots & \ddots & \vdots \\ (r_1,\dots,r_n)_i & \cdots & (r_1,\dots,r_n)_{im}\end{bmatrix}\right) = \begin{bmatrix}f_{1_1} & f_{2_1} & f_{3_1} \\ \vdots & \vdots & \vdots \\ f_{1_i} & f_{2_i} & f_{3_i}\end{bmatrix}$$

$$f_1 = \sum_{i=1}^{m} Ntc_i \; [kWh/m^2] \qquad f_2 = \sigma\left(\textstyle\bigcup_{i=1}^{m} Ntc_i\right) \; [kWh/m^2] \qquad f_3 = Retrofit\ cost\ [€]$$

Discussion

Surrogate Models to improve Building Performance Analysis and Optimization

Fundação para a Ciência e a Tecnologia

TÉCNICO LISBOA

CENTER for INNOVATION, TECHNOLOGY and POLICY RESEARCH

How to efficiently integrate AOP with building and urban projects?

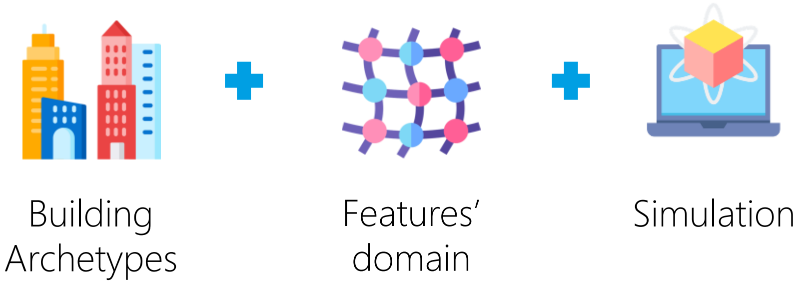
Flexible framework to develop Surrogate Models and integrate them with AOP.

Synthetic

Iterative

Existing

Model trained with a Synthetic BID.



"*AD based surrogate models for simulation and optimization of large urban areas*"

0.95 R^2 **score**, suitable for early design stages.

Simulation **time decreased** by **2 orders of magnitude**.

Small set of 6 inputs requires **minimum expertise**.

Versatile and **adaptable** to any building or urban project.

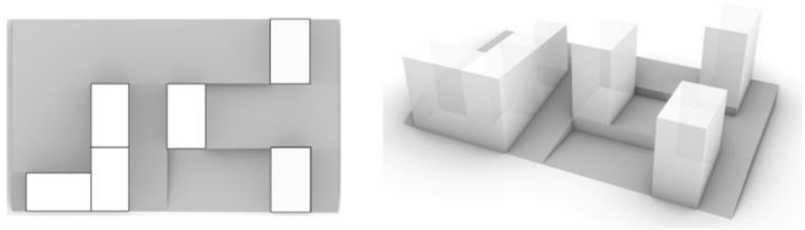
Small set of 6 inputs **decreases building details**.

High development time.

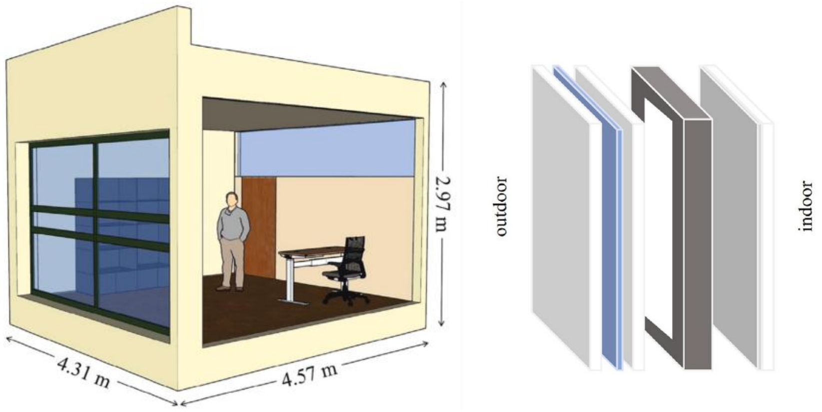
Increase in the **number of inputs exponentially increases** the model **development time**.

Process needs to be **repeated** for different simulation tools/outputs.

Model trained with BID obtained from optimization processes



"Surrogate Models for Efficient Multi-Objective Optimization of Building Performance"



"Multi-objective optimization of thermochromic glazing properties to enhance building energy performance"

Variables + Objectives + Simulation

0.97, 0.99 R^2 Score. Accuracy can be **improved** with **hyperparameter optimization.**

Lower development time than **Synthetic** databases.

Simulation **time decreased** by **2 orders of magnitude**.

Supports **any number of features.**

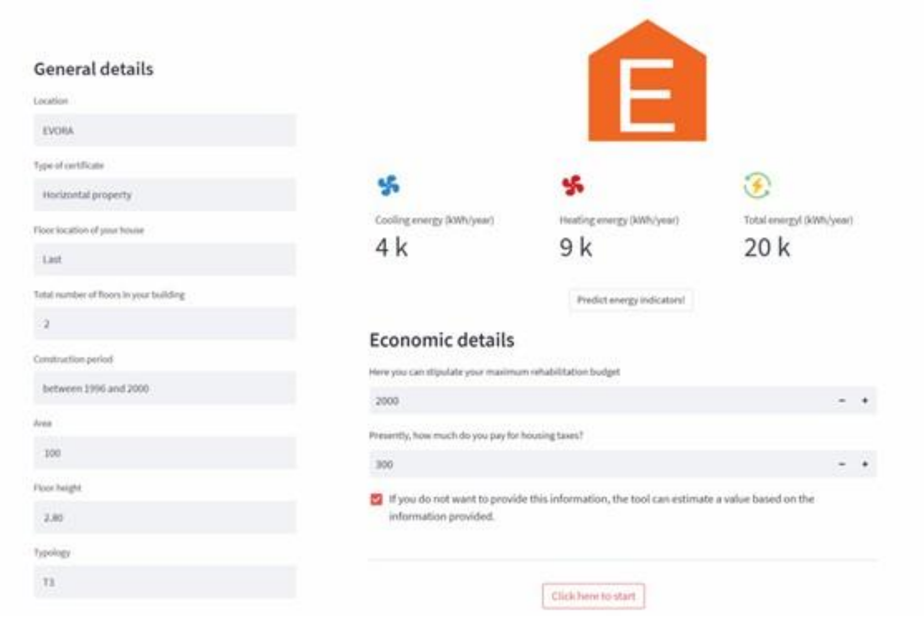
Supports **multiple objectives.**

Model is **case-specific** and not adaptable to problems outside the realms of the optimization.

Increase in the **number of inputs (variables)** also **increases** the model **development time**.

Less accurate than Synthetic databases in predicting **worse solutions.**

Model trained with BID obtained from existing databases



"Optimizing building retrofit through data analytics: A study of multi-objective optimization and surrogate models derived from energy performance certificates"

No simulations required.

Supports **any number of features.**

Supports **multiple objectives** if available in the data.

Lower development time than generating other BID.

Accuracy **highly dependable** on the existing **data's quality.**

Limited outputs and **prediction targets.**

0.84 R^2 **score, Less accurate** than **Synthetic** and **Optimization** (Can improve depending on the data).

Quantitative Comparison

	Problem	Scale	Objectives	Improvements	Inputs	Speed-up	R^2
	Retrofit	Urban	3	16%	6	≈85x	0.95
	Design	Building	1	8%			
	Material	Room	2	17%	4	≈200x	0.99
	Construction	Urban	3	22%	24	≈200x	0.97
	Retrofit	Building	3	60%	20	-	0.84/0.79
	Retrofit	Urban	3	25%		-	

Qualitative Comparison

	R^2	Development speed	Number of features	Number of objectives	Adaptability	Suitable for:
	★★☆	★☆☆	★☆☆	★☆☆	★★★	 Early-stage project studies with a low number of features
	★★★	★★☆	★★☆	★★☆	★☆☆	 Project construction and execution stages
	★☆☆	★★★	★★★	★★☆	★★☆	 Building operation and retrofit

Future projects

Explore new ways to develop a **smaller number of building samples** without loosing model accuracy and, therefore, be able to **increase the number of features and complexity** of a model.

Benchmark optimization and machine learning **algorithms** for multiple simulation outputs and analyses

Explore different feature engineering and selection techniques to improve the quality of the databases. **Experimental measurement** of building use and performance for data calibration.

Future applications

Assistance to field studies. Data feedback for field work.

Digital twin city models and databases. Enhance current data repositories.

Quick and easy Policymaking tools to achieve sustainable and economic goals.

On-the-fly assistance for design and execution projects. Enhanced collaborative work.

Research Outputs

Core Publications

Araujo, Gonçalo; Santos, Luís; Leitão, António & Gomes, R. (2022, April). *Ad based surrogate models for simulation and optimization of large urban areas*. In Proceedings of the 27th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA 2022), Sydney, Australia (pp. 9-15).

Araújo; G. R., Gomes;, R., Gomes;, M. G., Guedes;, M. C., & Ferrão, P. (2023). *Surrogate Models for Efficient Multi-Objective Optimization of Building Performance*. Energies, *16*, 4030. https://doi.org/https://doi.org/10.3390/en16104030

Araújo, G. R., Teixeira, H., Gomes, M. G., & Rodrigues, A. M. (2023). *Multi-objective optimization of thermochromic glazing properties to enhance building energy performance*. Solar Energy, *249*(October 2022), 446–456. https://doi.org/10.1016/j.solener.2022.11.043

Araújo, G., Gomes, R., Ferrão, P., & Gomes, M. da G. (2023). *A study of multi-objective optimization with surrogate models derived from energy performance certificates*. Energy and Built Environment.

Others

Araújo, Gonçalo; Teixeira, Henriqueta; Glória, G. M., & Moret, R. A (2022). Otimização de Envidraçados Termocrómicos para um Clima Mediterrânico. Congresso Construção 2022, 239.

Araújo, G., Pereira, I., Leitao, A., & Guedes, M. C. (2021). Conflicts in passive building performance: Retrofit and regulation of informal neighbourhoods. Frontiers of Architectural Research, 10(3), 625-638.

Aleixo, J., Araújo, G. R., & Guedes, M. C. (2021). Comparison of passive design strategies to improve living conditions: a study in Ondjiva, Southern Angola. Renewable Energy and Environmental Sustainability, 6, 21.

Guedes, M. C., Araújo, G., & Albuquerque, N. (2023). Thermal Comfort in Informal Settlements: Case Studies in Sub-Saharan Africa. In Climate Change and Sustainable Development (pp. 129-148). CRC Press.

Araújo, Gonçalo; Leitão, António, Inês Pereira; Gomes, Ricardo & Ferrão, Paulo (2021). *A non-linear surrogate model of building archetypes to speed up cities' adaptation to the post-carbon age*. In Congresso MITPortugal 2021.

Araújo, Gonçalo; Gomes, Ricardo & Ferrão, Paulo (2022). *Surrogate models for time-consuming building performance simulations and optimizations*. In Congresso MITPortugal 2022.

PhD Open Days

3rd place at the PhD Open Days 2021 pitch competition.

Surrogate Models to improve Building Performance Analysis and Optimization

Thank you for listening!

fct Fundação para a Ciência e a Tecnologia

TÉCNICO LISBOA

IN+ CENTER for INNOVATION, TECHNOLOGY and POLICY RESEARCH

CERIS