István Kistelegdi, Zsolt Ercsey, Tamás Stortz, Anthony John Gall

Sensitivity analysis of building comfort and energy design variables

Background in the field of building comfort, energy and environmental optimization

First building comfort, energy and environmental design optimization (BECEDO) research efforts are dated back to the early 70's[1], however most in-depth analysis is carried out in non-architectural fields (IT, mathematics and operation technologies). Among building design optimization studies, most attention is focussed on the application of active (mechanical) design variables (HVAC systems), renewable energy harvesting and supply system variables and the combination of active and passive (architectural) design input variables[2]. Regarding the passive design variables, which should be determined for an optimal comfort-energy, as well as environmental impact balance performance, only the numerically easy to be-parametrized design variables are taken into consideration, such as opaque and transparent envelope structures and materials, e.g., the thicknesses and thermal properties of the insulation and walls, as well as wall-window ratios (WWR), orientation (ORI), materials, structures (STR) and shading for instance[3]. Though building shape has significant impact on building operation cost[4], i.e., up to 60-80% energy conservation[5][6] and up to 80% LCA savings are possible, the investigations dealing with building geometry as a design variable (BGDV) in the BECEDO process is still in its infancy[7]. Estimations[8] predict 60-70% energy consumptions reduction in HVAC and artificial lighting system improvements and up to 20% savings by using intelligent automation systems, however the energy saving potential of optimized space organization and complete building shape design is still missing. Another issue evolves after analysis of the existing BECEDO literature: the stochastic behaviour of the most most frequently applied evolutionary technique (generic algorithms GA) randomly moves in the search space and only near optimum solutions delivers.

Therefore, current research framework intends to elaborate a BECEDO methodology that incorporates the intensive analysis of BGDV-s in the optimization process and aims to ensure guaranteed solution(s) due to 'scanning' the total search space, made with help of carefully predefined, comprehensive and strict rules.

Research framework

The fundament of the new design method to be developed is represented by the multiple award winning and patented EnergiaDesign energy-positive building planning method[9], using complex CFD, thermal and daylight simulations as a design support.

As a first step, a modular space arrangement system was created, complemented by a series of building geometry producing rules to select only those building geometry configurations, which meet the space generation rules. 176 family house shape configurations were selected from a total of 201,359,550 possible geometries. Applying 3 typical WWR (30%, 60%, 90%) in the main façade and 5 ORI from (E, SE, S, SW, W), as well as 2 different structures (U_{wall} = 0,24 W/m^{2}K and 0,11 W/m^{2}K; U_{floor} = 0,28 W/m^{2}K and 0,17 W/m^{2}K; U_{roof} = 0,17 W/m^{2}K and 0,14 W/m^{2}K; $U_{glazing}$ = 1 W/m^{2}K and 0,7 W/m^{2}K;). The combination of these design input variables resulted 5,010 building cases, which were calculated in a dynamic thermal simulation framework (IDA ICE) to gain a database of input and output design variables.

---

[1] C. L. Gupta, "A systematic approach to optimum thermal design," *Build. Sci.*, vol. 5, no. 3–4, pp. 165–173, 1970, doi: 10.1016/0007-3628(70)90006-X.

[2] Z. Yu, Z. Gou, F. Qian, J. Fu, and Y. Tao, "Towards an optimized zero energy solar house: A critical analysis of passive and active design strategies used in Solar Decathlon Europe in Madrid," *J. Clean. Prod.*, vol. 236, p. 117646, 2019, doi: 10.1016/j.jclepro.2019.117646.

3 S. N. Al-Saadi and K. S. Al-Jabri, "Optimization of envelope design for housing in hot climates using a genetic algorithm (GA) computational approach," J. Build. Eng., vol. 32, no. May, 2020, doi: 10.1016/j.jobe.2020.101712.

4 Y. Fang and S. Cho, "Design optimization of building geometry and fenestration for daylighting and energy performance," Sol. Energy, vol. 191, no. August, pp. 7–18, 2019, doi: 10.1016/j.solener.2019.08.039.

5 Gerhard Hausladen, Michael Saldanha, Petra Liedl, Christina Sager: ClimateDesign: Solutions for Buildings that Can Do More with Less Technology, Published by Birkhäuser Architecture, 2005, ISBN 10: 3764372443ISBN 13: 9783764372446

6 B. Kiss and Z. Szalay, "Modular approach to multi-objective environmental optimization of buildings," Autom. Constr., vol. 111, no. November 2019, p. 103044, 2020, doi: 10.1016/j.autcon.2019.103044.

7 Kistelegdi, I.; Horváth, K.R.; Storcz, T.; Ercsey, Z. Building Geometry as a Variable in Energy, Comfort, and Environmental Design Optimization—A Review from the Perspective of Architects. Buildings 2022, 12, 69. https://doi.org/10.3390/buildings12010069

8 Shi, X.; Tian, Z.; Chen, W.; Si, B.; Jin, X. A review on building energy efficient design optimization rom the perspective of architects. Renew. Sustain. Energy Rev. 2016, 65, 872–884.

9 Balint BARANYAI, 2Balint PÓTH, 3István KISTELEGDI Jr: PLANNING AND RESEARCH OF SMART BUILDINGS AND CONSTRUCTIONS WITH THE 'ENERGYDESIGN ROADMAP' METHOD, POLLACK PERIODICA An International Journal for Engineering and Information Sciences DOI: 10.1556/Pollack.8.2013.3.2 Vol. 8, No. 3, pp. 15–26 (2013)

As a second step, different regression models were created to substitute modeling and calculation time capacity intensive white-box simulations. These prediction models should support an automated decision and evaluation algorithm system to produce the optimized buildings. Beside the regression analysis, sensitivity analysis provides information of critical importance about the impact of input design variables on the comfort and energy performance of the building cases. In the following sections the sensitivity measure of the input variables in two different regression model system were investigated.

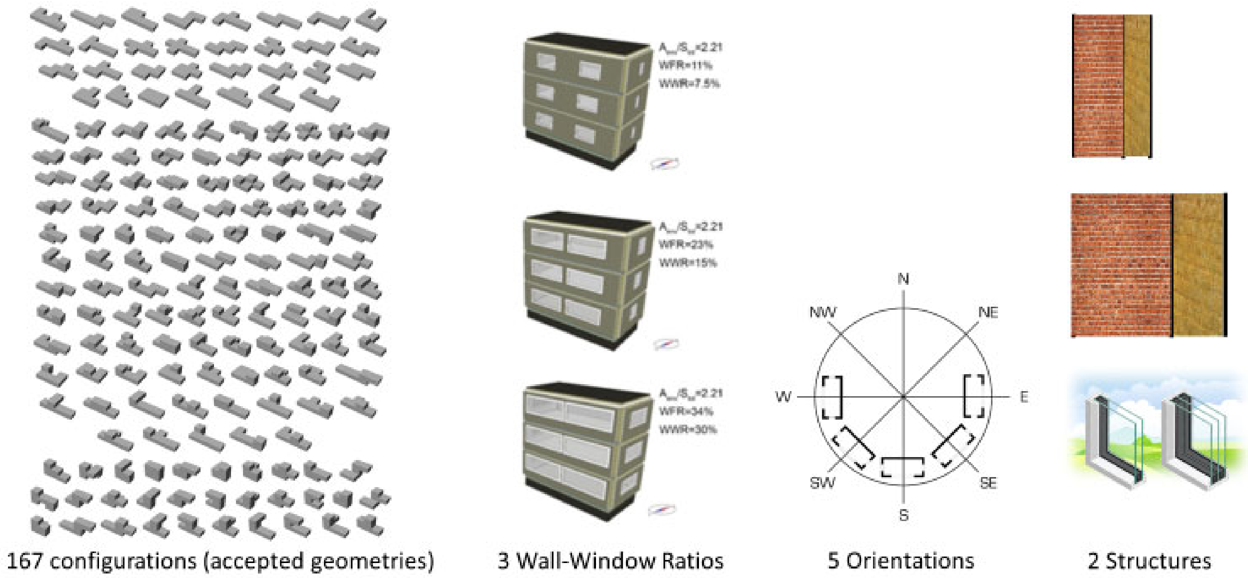

Fig.1.: Simulation database consisting of 167 shapes, 3 WWR, 5 ORI and 2 structural variations

Feature importance by decision tree regressor

Decision trees are effective decision support tools; where the decision tree is a special graph, i.e. a tree, representing the decisions together with their possible consequences. They are widely used because they are simple to understand and interpret. Hereinafter, a binary decision tree is used, where the internal nodes of the tree represent simple tests related to the value of the considered input variable. There are exactly two successor children of each node. Should the value of the variable be below a given threshold then the decision proceeds towards one of its children, otherwise the other child is selected and the decision proceeds towards that branch. The leaves, i.e. those nodes without children, represent the decisions. When generating the decision tree, it is important to select in the right sequence the considered input variables together with their thresholds. This is usually performed via minimizing the variance sum of the subtrees.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad \text{Equation 1.}$$

To select the decisions applied at the nodes of the decision tree, the Mean Squared Error method is used to evaluate the gain of the variance, see Equation 1. The regression model based on this decision tree is built. The importance of the input variables based on their probable influence on the output related to the decision trees is given in the following table and the relative difference of the influence of the input variables is depicted in the following graph (Fig. 2.). It is obvious, that considering the annual heating energy demand (HE) together with the sum of energy demand (SE) as considered outputs, the most influential variable is the structure (Struct), i.e. the thermal insulation and mass (0.5+) marked with green color, while in the other cases (cooling energy CE, lighting energy LE, thermal comfort TE, daylight factor DF) the most influential variable is the wall-window ratio (WWR), namely 0.6-0.9. HE and SE is also influenced by orientation (Ori) and A/S (external envelope surface to floor space ratio) with a medium intensity. The figure displays as well, that for the annual heating energy (HE), annual cooling energy (CE) and sum of energy (SE) the influence is strongly based on the structure. On the other hand, considering the lighting energy and comfort related outputs (TC, DF) the most important if not the only determinative variable is the wall window ratio.

R^2	Output	A/S	Struct	WWR	Ori
0,968	HE	0,240	0,524	0,028	0,208
0,889	CE	0,093	0,189	0,646	0,072
0,963	LE	0,043	0,000	0,935	0,022
0,921	TC	0,089	0,009	0,840	0,062

0,963	DF	0,034	0,001	0,899	0,066
0,952	SE	0,149	0,592	0,050	0,209

Table.1.: Feature importance of input variables in decision trees

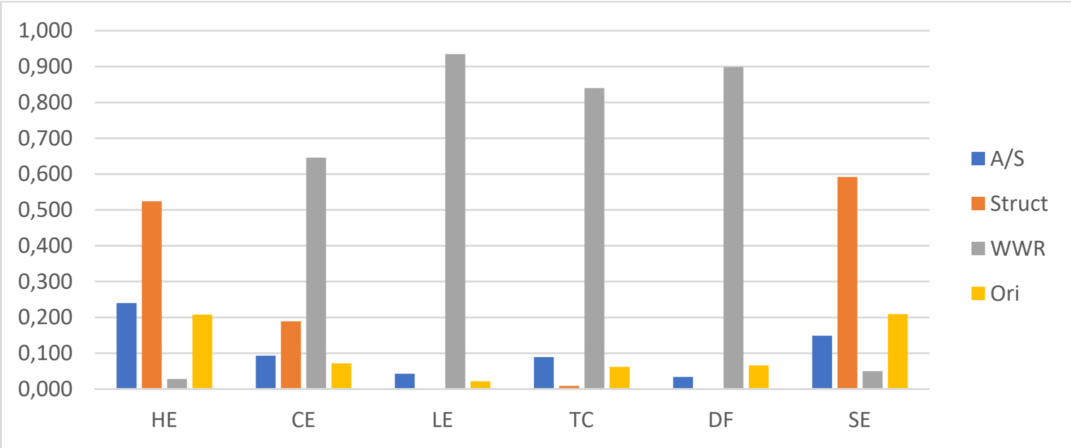

Fig.2.: Relative difference of the input variables' feature importance in decision trees

Normalized absolute feature weights of linear regression

Linear regression models the dependent variable as a linear combination of the descriptive variables. When generating the linear combination model, the coefficients of the descriptive variables is sought. To support the nonlinearity of the functions, new input variables are added with multiplicative combinations of the original inputs, namely quadratic (secondary) and tertiary (3rd power) input variables are introduced.

When the original 4 descriptive variables are used the most influential input variable is the A/S (0.8-0.9), nevertheless the accuracy of the approximation is not appropriate, i.e. R^2=0.7-0.8. This case is given in the following table. For the calculation of the accuracy please consider Equation 2.

R^2	Output	A/S	Struct	WWR	Ori
0,721	HE	0,906	0,423	0,000	0,000
0,771	CE	0,832	0,555	0,022	0,001
0,833	LE	0,990	0,099	0,104	0,003
0,831	TC	1,000	0,015	0,026	0,001
0,842	DF	0,989	0,132	0,068	0,002
0,691	SE	0,847	0,531	0,001	0,000

Table.2.: Normalized absolute feature weights of input variables in linear regression

$$R^2 = 1 - \frac{\sum(y - \tilde{y})^2}{\sum(y - \bar{y})^2} \qquad \text{Equation 2.}$$

When the original descriptive variables are used together with multiplicative combinations of the inputs up to the quadratic power, the single complex descriptor is the most influential (between 0.85-0.99 depending on the estimated output). Please consider the table below. The next graph depicts the influential ratios of the input variables. In both six output cases the most influential variable is the complex descriptor of the building configuration, the A/S ratio (compactness of the building body shape).

R^2	Out	A/S	Struct	WWR	Ori	A/S^2	A/S * Struct	A/S * WWR	A/S * Ori	$Struc^2$	Struct* WWR	Struct* Ori	WWR^2	WWR* Ori	Ori^2
0,93	HE	0,99	0,05	0,01	0,01	0,00	0,03	0,00	0,00	0,14	0,00	0,00	0,00	0,00	0,00
0,86	CE	0,98	0,02	0,01	0,00	0,20	0,06	0,00	0,00	0,05	0,00	0,00	0,00	0,00	0,00
0,92	LE	0,85	0,02	0,32	0,03	0,42	0,08	0,04	0,00	0,05	0,00	0,00	0,00	0,00	0,00
0,89	TC	0,98	0,01	0,01	0,00	0,20	0,01	0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,00
0,91	DF	0,98	0,00	0,01	0,01	0,21	0,02	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00
0,92	SE	0,99	0,04	0,00	0,01	0,06	0,02	0,00	0,00	0,12	0,00	0,00	0,00	0,00	0,00

Table.3.: Normalized absolute feature weights of input variables in linear regression – 2nd power

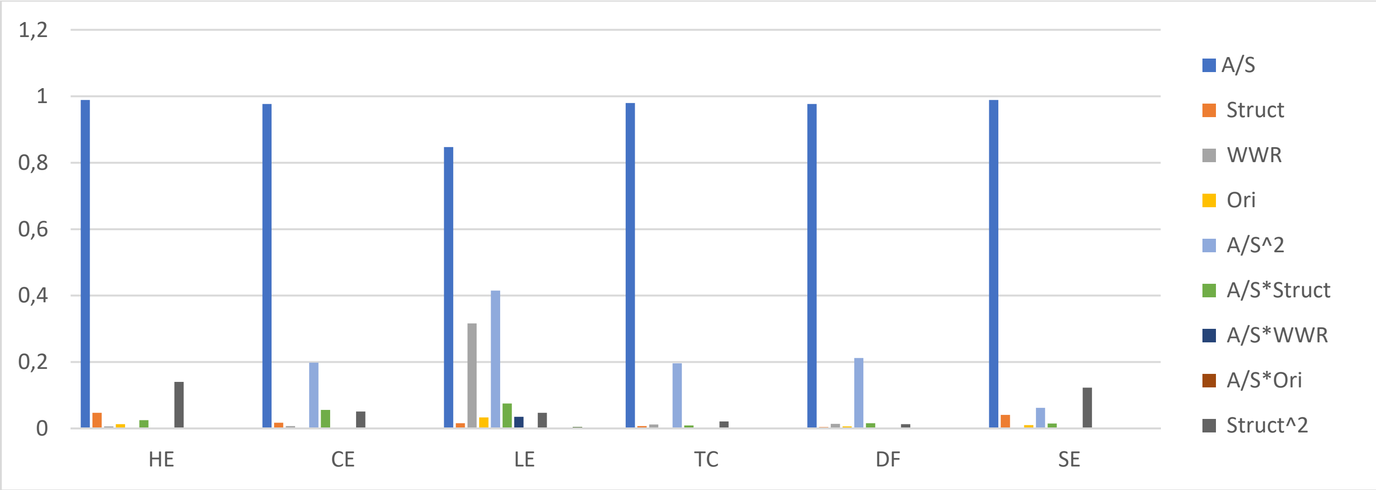

Fig.3.: Normalized relative feature weights of input variables in linear regression – 2nd power

When the original descriptive variables are used together with multiplicative combinations of the inputs up to the third power, the number of single variables is 34, therefore it is problematic to depict them or to summarize them in a table. However, during the investigations, it can be seen that the product of the complex descriptor and the structure is the most influential (`A/S*Struct`: 0,802), the complex descriptor is the second most influential (`A/S`: 0.534), while the third most decisive is the product of the complex descriptor and the square of the insulation ($\mathtt{A/S*Struct}^2$: 0.267); moreover, the other input variables have relatively small influence.

Discussion

Important to mention that the only shape expressing variable in particular study is represented by the A/S ratio that is responsible to provide information about the compactness of the building shape. By representing the ratio between the external envelope surface (heat loss surface) and the conditioned indoor floor space area, this variable is a complex and indirect descriptor of the geometry. None of the remaining design input variables describe the shape of the geometry.

In the decision tree evaluation structure dominates the influence as function of the transmission and thermal bridge-based heat loss, further, the WWR gains importance as the function of winter solar heat gains, transmission heat losses and summer solar load. The A/S descriptor possesses no significance.

In the linear regression investigation, the A/S input variable gains on influence in both secondary and tertiary power for all output result parameters. On the one hand, this effect is caused by the various compactness of the form (affects heating demand and thermal comfort), on the other hand this effect is due to the changing depth of the indoor space (modifies solar gains and load, therefore heating, cooling, and lighting demand, as well as thermal and visual comfort). This would mean that the geometry as design variable has an outmost importance on comfort and energy performance of buildings. The accuracy of the approximation is appropriate in both decision tree as well as linear egression (2^{nd} and 3^{rd} power). Nevertheless, this conclusion is regarded only as a potential main impact variable for the time being since the scale of the input variables can modify the and hence distort the feature weights.

Outlook

In further research steps more sensitivity analysis methods (e.g., Morris Elementary-Effect), as well as regression analysis (e.g., dense neural network) are to be carried out to broaden the insights about the measure of input variables' influence. Further analysis is required to examine the robustness of the 2^{nd} and 3^{rd} power linear regression results, mainly to handle the different input variable value-scales. The insights provide the basis for the optimized building design algorithm.