

# On the performance of meta-models in building design optimization

A.Prada<sup>a,\*</sup>, A. Gasparella<sup>b</sup>, P. Baggio<sup>a</sup>

<sup>a</sup>*Department of Civil, Environmental and Mechanical Engineering, University of Trento,  
Italy*

<sup>b</sup>*Faculty of Science and Technology, Free University of Bozen-Bolzano, Italy*

---

## Abstract

Although evolutionary algorithms coupled with building simulation codes are often applied in academic research, this approach has a limited use for actual applications of building design due to the high number of expensive simulation runs. The use of a surrogate model can overcome this issue.

In the literature there are several functional approximation models that can emulate the building simulation during the optimization, thus increasing the process efficiency. However, there are no evidence-based studies comparing the performances of these methods for the building design optimization.

This study compares the efficiency, the efficacy and the quality of the Pareto solutions obtained by Polynomial, Kriging (*GRFM*), Radial-basis function networks (*RBFN*), Multivariate Adaptive Regression Splines (*MARS*) and support vector machines (*SVM*) functional approximations. The test bed of the comparison is the evaluation of the optimal refurbishment of three reference buildings for which the actual Pareto front is also obtained through

---

\*Corresponding author: via Mesiano 77, 38123 Trento, Italy.  
Tel.: +39 0461 282516  
Email address: [alessandro.prada@unitn.it](mailto:alessandro.prada@unitn.it) (A.Prada)

a brute-force approach.

The results show that the *MARS* method outperforms the other surrogate models both in terms of efficiency and effectiveness, and also by assessing the quality of the Pareto front.

*Keywords:* Multi-objective optimization, Surrogate Model, Efficient Global Optimization, Building Simulation, nZEB design

---

## 1. Introduction

The coming into force of the European Directive 2010/31/EU [1] guides the member states to the reduction of energy demand, and consequently carbon emissions, of the European building stock. Moreover the directive considers the economic effectiveness by means of the "cost-optimal approach" [2]. Optimization of the building and *HVAC* design and control becomes an essential tool in the design of new buildings and building refurbishments approaching the *nZEB* target [3–6]. Besides, when approaching the *nZEB* target while maintaining economical convenience, buildings might be easily subject to poor comfort conditions [7, 8]. Hence, the designers are always confronted with a multi-objective optimization problem with two or more conflicting goals. The overall gains in design quality as well as the cost reductions that can be achieved through a correct optimization process are high. For this reason, architects and engineers become increasingly aware of the potential benefits of applying building performance optimization (*BPO*) in the early stages of the design process, often coupling the optimization codes with dynamic building performance simulation (*BPS*) since it better describes the dynamic interactions between the building, energy systems,

occupants and the outdoor environment.

The use of gradient-based optimization or linear programming methods are not easily adapted for *BPO* since the relationship between design variables and cost functions can be non-convex, non-linear and the optimization problem can be subjected to non-linear constraints or to numerical approximations [9]. Thus, evolutionary algorithms (*EA*) are frequently adopted since they have less requirements on the problem characteristics. The popularity that *EA* are enjoying arises from the flexibility with which they can deal with various optimization problems including high dimensional problems, integer or real parameters as well as continuous or discrete variables, non-differentiable cost functions and so on [10]. The evolution of multi-objective optimization in building simulation is well documented in several works [7, 11, 12] and review papers illustrating the history and the current state of the art [13–15]. According to Hamdy et al. [12], genetic algorithms are to a considerable extent the most implemented algorithms in the literature dealing with building optimization and the NSGA-II [16] is probably the most popular.

The main challenge in the use of *EA* coupled with building simulation is that *EA* usually still needs a large number of cost function evaluations before a satisfying result can be obtained [17]. Moreover, when considering *BPS*, the evaluation of the cost functions is computationally very expensive and the time taken to perform a single *BPS* is of the order of minutes or hours, depending on the model complexity. This aspect reduces the effectiveness of the *BPO* and especially its diffusion in the professional practice [18]. For instance, if the simulation is used for a simulation predictive control in a

complex building, the time required for the *BPO* is not short enough to implement actions in the time frame of reliability of weather forecasts. For this reason, an approximation of the optimization problem is required to use *EA* efficiently. According to Jin [19] the main approximation strategies are:

- *Problem Approximation*: when the *BPS* is replaced by a model that is computationally less expensive. For example, when quasi-steady state methods or hourly lumped capacitance models are used in lieu of *BPS* [20–24].
- *Evolutionary Approximation*: when the fitness function evaluation of the offspring is estimated from the fitness values of their parents by means of the fitness inheritance approach [25].
- *Fitness Imitation*: when the alternative design solutions are clustered and the fitness function is evaluated only for the centroid. The objective functions of the other individuals are then estimated from the centroid response [26].
- *Functional Approximation*: when an explicit expression of the *BPS* is constructed starting from the building simulation results and used together with *EA* (a.k.a. meta-model or surrogate model approximation) [27–34].

The functional approximation approach is the most used in *BPS* to speed up the optimization process and this is the approach followed in this research. The meta-model can be used in the *EA* in three main ways. The first strategy is the direct use of a surrogate model in the optimization process [35].

For instance, Eisenhower et al. [30] optimized the energy consumption and the thermal comfort of an existing building by using a gradient-based optimizer on the surrogate model fitted on the EnergyPlus outcomes. However, the drawback of surrogate models can be the accuracy, since the meta-model contains an uncertainty in it. For instance Hopfe et al. [32] highlights the disadvantage of Kriging due to the limited number of design variables at which the meta-model still does the quality estimations. For this reason, the second strategy is the "generation-based control" [36]. In this approach, the surrogate model is firstly used in the *EA* code to find the optimal solutions. Following on from this point, the cost functions are evaluated for the optimal points by means of *BPS* and the surrogate model is then updated. Xu et al. [34] used a support vector machine algorithm (*SVM*) to fit a meta model. The regression model is then used to investigate the variable space with the purpose of selecting the best candidates for the cost function evaluation through *BPS*. Similarly, Brownlee and Wright [33] used the surrogate model to generate a surrogate population, to rank it and to select the best individuals for the actual fitness function evaluation.

Finally, a local meta-model can be used alternatively to guide the selection of new offspring (i.e. "individual-based control") since it is difficult to build a globally approximate model [37] especially in high dimensional problems. In one of the earliest works, Knowles [38] developed ParEGO (Pareto Efficient Global Optimization) which extended the local efficient global optimization algorithm *EGO* to the multi-objective optimization by converting the objective functions into a single objective through the augmented Chebyshev function. A similar study is proposed by Emmerich et al. [36], in which

the authors extend the local Gaussian Random Field meta-model (a.k.a. Kriging) to the multi-objective problems.

An efficient optimization with functional approximation is essential to find the trade-off solutions in building design and refurbishment. Moreover it can broaden the diffusion of the cost-optimal approach in real world applications overcoming the issues that limit the diffusion among architects and engineers. The performance of meta-models in emulating *BPS* have been recently tested in Østergård et al. [39] for an office and an educational building. The authors came to the conclusion that *GRFM* is more accurate but it is less efficient than *MARS* and Neural Networks. However, the performance of different meta-models on building optimization problems has not been investigated yet and no evidence-based studies are present in the literature.

For this reason, the performances in terms of efficiency, efficacy and solution quality of several meta-models with different sampling techniques are investigated in this paper (Figure 1). The efficiency index is meant to measure the resource level used by the algorithm whereas the efficacy is a measure of the distance between the predicted Pareto front and the true Pareto solution. Finally, the solution quality quantifies the Pareto front uniformity in the objective space. The refurbishment of three simplified reference buildings representative of a penthouse, a semi-detached house and an intermediate flat in an apartment building [40] are optimized by considering six types of energy saving measures (*ESM*). In this respect, mature and off-the-shelf technologies are considered such as the opaque envelope insulation, the replacement of the windows or the heating system and the installation of a mechanical ventilation with heat recovery.

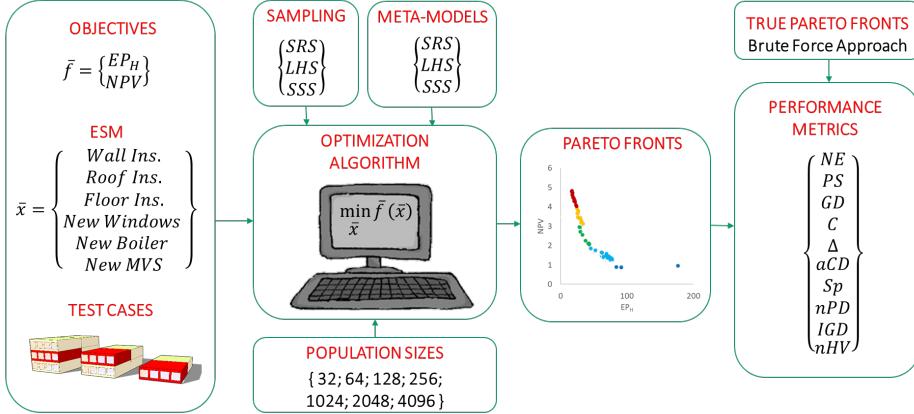


Figure 1: Scheme of the methodology

The integer optimization problem is solved by using the "generation-based control" approach with a customized algorithm developed in Matlab. The efficiency is computed by analyzing the number of expensive *BPS* runs required by polynomial, Kriging (*GRFM*) [41], Radial-basis function networks (*RBFN*) [42] , Multivariate Adaptive Regression Splines (*MARS*) [43] and support vector machines (*SVM*) [44] models. Moreover, different sampling techniques are considered to highlight the extent to which the selection of the initial sample of the energy saving measures (i.e. the initial population) can affect the efficiency of the optimization process. For this reason, simple random sampling (*SRS*), Latin hypercube (*LHS*) and Sobol sequence sampling (*SSS*) are implemented in the optimization algorithm. Moreover, since there is no guarantee that *EA* will find any optimal solution in a finite number of generations for the optimization problems, the efficacy is studied by comparing through some metrics the Pareto fronts against the brute-force solutions. In fact, the brute force solution is the exact solution of

the optimization problem since it considers all the *ESM* combinations due to the discrete nature of all the energy saving measures. Finally, the quality of the solution in terms of the diversity and uniformity of the Pareto front were also considered in order to maximize the information provided to the decision-maker.

## 2. Method

### 2.1. Optimization Algorithm

The multi-objective optimization problems have been solved using a customized algorithm developed in Matlab following the "generation-based control" approach (Figure 2). The algorithm firstly selects the initial population, i.e a sample of possible retrofit configurations each provided by a specific combination of energy saving measures (section 2.4), by a sampling technique. Once the cost function has been evaluated for the initial population through the *BPS*, the algorithm proceeds with the meta-model fitting in order to estimate the best set of meta-model parameters to emulate the input-output relationship. To this aim, several meta-models have been implemented, as reported in section 2.2. Following on from this point, the optimization problem is solved by using the *NSGA-II* [16] coupled with the meta-model in lieu of *BPS*, and the Pareto front is defined. The algorithm then proceeds to the evaluation of the real cost functions of the Pareto solutions highlighted by the *NSGA-II* and it saves the *BPS* results to an external dataset. In this way the expensive function evaluations are performed only once, since the dataset is checked before any *BPS* launch. After the *BPS* of all the Pareto solutions identified by the *NSGA-II* through the meta-model eval-

ulations, the code parses the external dataset with the results of the *BPS* and it identifies the non-dominated solutions, i.e. the Pareto front. If the Pareto front meets the stopping criterion, the algorithm finishes, otherwise it updates the meta-model, starting from all the solutions in the dataset, and then it returns to the *NSGA-II* optimization.

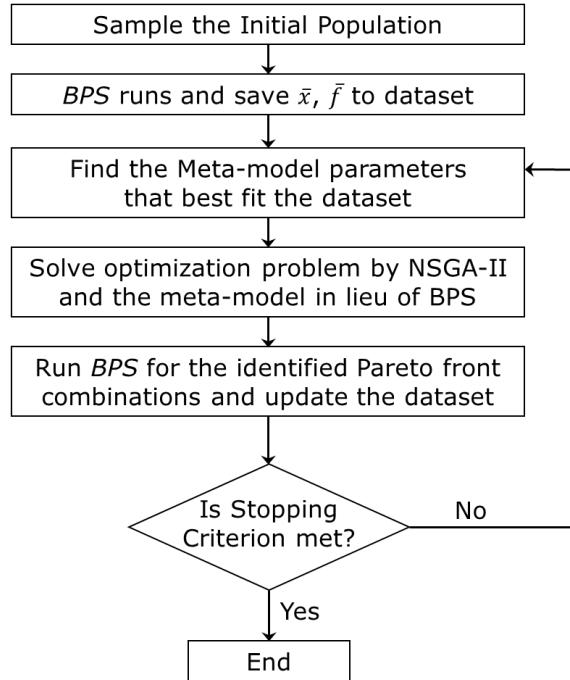


Figure 2: Flowchart of the developed optimization algorithm

According to the literature, the generation of the initial population can be performed using various procedures [45]. The optimization algorithm has been complemented with Sample Random Sampling, Latin Hypercube and Sobol Sequence Sampling in order to investigate the role of initial population in the algorithm performances. The *LHS* and the *SSS* are implemented

with the purpose of overcoming the clustering which can occur with random generator since the *SRS* could lead to an over-focus on the same regions of the hyperspace without sampling in others. The same initial population is adopted for all the meta-models for each of the optimization problems and sampling methods, in order to ensure the same starting point and therefore the result comparability. Furthermore, eight initial population sizes of the surrogate models are adopted in order to assess their impact on the performance. The population sizes are obtained as powers of 2 for all the sampling methods since the *SSS* method has this constraint and, especially, a unitary step of the exponents was adopted in the range 5 to 12.

The hypervolume measure, originally proposed by Zitzler and Thiele [46], is used as a stopping criterion. Although this has the drawback of the higher computational cost in the evaluation of the size of dominated space, the maximization of this index is the necessary and sufficient condition for the Pareto optimal solutions of a discrete optimization problem [47]. Moreover the hypervolume index is a unary metric able to quantify the solution accuracy and diversity [48]. In the code the threshold of  $10^{-3}$  in the variation of the normalized hypervolume between two consecutive generations was adopted for the convergence criterion.

## 2.2. Meta-models

The role of meta-models in the "generation-based control" is to filter out the variable domain regions with no eligible Pareto solutions. In the literature several functional approximation models have been presented [19, 49]. However, it is not reasonable that a complex function such as *BPS* can be predicted with sufficient accuracy by any meta-models. There should

exist a type of surrogate model that approximates the *BPS* with a better accuracy or that requires a lower number of *BPS* evaluations to achieve a good approximation. For this reason, in the analysis eight meta-models have been implemented and tested. In the following sections, the main features of all the surrogate models are described.

### 2.2.1. Polynomial approximation

The first typology of meta-models, little used in *BPS*, is the polynomial approximation of the cost functions [27]. Both full second ( $Pol^2$ ) and third order ( $Pol^3$ ) polynomial approximations have been developed in Matlab:

$$\begin{aligned}\bar{f} &= \beta_0 + \sum \beta_i \cdot x_i + \sum \beta_{ij} \cdot x_i x_j + \sum \beta_k \cdot x_k^2 \\ \bar{f} &= \beta_0 + \sum \beta_i \cdot x_i + \sum \beta_{ij} \cdot x_i x_j + \sum \beta_{ijk} \cdot x_i x_j x_k + \sum \beta_m \cdot x_m^3\end{aligned}\tag{1}$$

where  $f$  and  $x$  are respectively the cost functions and the optimization variables used in the *BPS* while  $\beta$  are the unknown polynomial coefficients that have to be estimated using a least square method.

### 2.2.2. Gaussian Random Field meta-model (GRFM)

This method, firstly proposed in the sixties by Krige [41], is popular in *BPS* applications [31, 32, 50, 51]. In this approach the cost function ( $f$ ) is approximated by a combination of a global regression model and a random deviation from it ( $e$ ) as shown in eq. 2.

$$\bar{f} = \sum g_i^T(\bar{x}) \cdot \beta_i + e(\bar{x})\tag{2}$$

where ( $g_i^T$ ) is the regression function and  $\beta_i$  are the unknown coefficients to be estimated. The random deviation has zero mean and a covariance matrix

obtained from the product of the cost function variances multiplied by the correlation model ( $r$ ), depending on unknown parameters determined by a maximum likelihood estimation. In this work the Matlab DACE toolbox [52] is used to estimate the *GRFM* model by assuming a linear regression ( $g_i^T$ ) with a cubic spline correlation model ( $r$ ).

### 2.2.3. Radial basis function networks (*RBFN*)

Neural networks have been widely used for function approximation [29, 53] and *RBFN* is one of the possible strategies [33]. Micchelli [42] in an early work proposed a mathematical formulation of *RBFN*. In this approach, each cost function can be approximated by a linear combination of unknown coefficients  $\beta_j$  multiplied by a set of radial-basis function  $\phi$ .

$$f_i = \sum \beta_j \phi(\|x - \mu^{(j)}\|) \quad (3)$$

where  $\|x - \mu^{(j)}\|$  denotes the Euclidean norm between the points and the center of the basis function that is one of the model unknowns. Several radial-basis functions have been proposed in the literature. In this work the Neural Network Toolbox of Matlab is used to approximate the *BPO* objectives by means of *RBFN*. In particular three different basis function sets (eq. 4) are used: linear ( $RBF^1$ ), cubic ( $RBF^3$ ) and thin-plate spline ( $RBF_{tps}$ ).

$$\begin{aligned} \phi &= d \\ \phi &= d^3 \\ \phi &= d^2 \cdot \ln(d) \end{aligned} \quad (4)$$

where  $d$  is the pairwise distances between the variable points already used in *BPS* and the new points to be evaluated.

#### 2.2.4. Multivariate Adaptive Regression Splines (*MARS*)

The fourth type of surrogate model implemented is based on the regression form introduced by Friedman [43] and it has been used in *BPS* in recent works [54, 55]. Similarly to *RBFN*, the cost functions are approximated by separate piecewise splines of differing slope. The cost functions are then an expansion of unknown coefficients ( $\beta_j$ ) and the  $N$  basis functions ( $B_j$ ) as shown by eq. (5).

$$f_i = \sum_1^N \beta_j B_j(x) \quad (5)$$

The coefficients  $\beta_j$  are adjusted to give the best fit to *BPO* actual objectives. Moreover, also the set of the  $N$  basis functions  $B_j$  and the slope change points (i.e. the knot) are automatically derived to increase the goodness of fit. In this work we used the ARESlab toolbox [56] to fit the *MARS* model to *BPS* data and to predict the cost functions in other regions of the space by setting the maximum number of basis function to 21 and using piecewise-cubic splines.

#### 2.2.5. Support Vector Machines (*SVM*)

The last surrogate model analyzed is a machine learning algorithm. It was first developed to classify models and then extended to regression through the minimization of an alternative loss function proposed by Vapnik [44]. *SVM* is fairly often used as a surrogate model in the energy simulations of buildings [30, 54]. The goal is to find a function that deviates from the *BPS* results in the data-set by a value no greater than a defined threshold for each training combination of *ESM*. This is computationally simpler to solve by using its Lagrange dual formulation. In this study the Matlab machine

learning toolbox was used for the *SVM* training and predicting by using the relation in eq. 6.

$$\bar{f} = \sum \gamma_j K(x, x_i) \quad (6)$$

where  $\gamma_j$  is the primal linear problem coefficient and  $K(x, x_i)$  is the kernel function. A cubic polynomial function is used as a kernel function in this study.

Eight meta-models are then implemented in the study. The models vary among them for the mathematical formulation and for the order of approximation as shown in the Table 1.

Table 1: Characteristics of the implemented meta-models

Symbol	Meta-model	Characteristics
$Pol^2$	Polynomial	Quadratic Order
$Pol^3$	Polynomial	Cubic Order
$GRFM$	Kriging	Linear Regression with Cubic Spline Correlation
$RBF^1$	Radial Basis Function Network	Linear Basis
$RBF^3$	Radial Basis Function Network	Quadratic Basis
$RBF_{tps}$	Radial Basis Function Network	Thin Plate Spline Basis
$MARS$	Multivariate Adaptive Regression	Piecewise-cubic splines with 21 Basis Functions
$SVM$	Support Vector Machines	Cubic polynomial kernel

### 2.3. Performance indicators

All computations ran on Windows computers having an Intel *i7* Quad-Core processor with  $4.0\text{ GHz}$  frequency and  $64\text{ GB}$  of RAM. On this hardware, each *BPS* run took up to 40 seconds to be performed. In order to compare to each other the surrogate models three different metric categories are defined as:

- **Efficiency:** the resource level used by the algorithm to converge;
- **Efficacy:** the distance of the predicted Pareto front from the true Pareto optimal;
- **Quality:** the ability to provide the decision maker with the maximum information about the possible solutions.

All metrics are evaluated starting from the objectives normalized with respect to the cost functions of the existing building, in order to avoid the risk that a different magnitude could undermine the metrics observations.

The efficiency is quantified by the adimensional metric  $NE$  defined as the ratio of costly  $BPS$  runs ( $n_{BPS}$ ) to get the Pareto solutions over the number of  $ESM$  combinations ( $n_{ESM}$ ) (eq. 7).

$$NE = \frac{n_{BPS}}{n_{ESM}} \quad (7)$$

This index ranges from 0 to 1 and it provides the same information as the CPU time since the time required for the meta-model estimations is negligible with respect to the computational time for the  $BPS$  runs. Furthermore, the  $NE$  is not affected by the computational efficiency of the Matlab codes we implemented for the meta-model estimations.

The efficacy of the optimization algorithm is evaluated by means of three different metrics. Firstly the percentage of the true Pareto solutions ( $PS$ ) actually found by each configuration of the optimization algorithm has been computed (eq. 8). This index ranges from 0 to 1 and it approaches 1 when all the Pareto solutions of the brute force ( $n_{Brute}$ ) are found by the optimization algorithm.

$$PS = \frac{n_{Par}}{n_{Brute}} \quad (8)$$

Additionally, the Generational Distance ( $GD$ ) [57] and the Coverage ( $C$ ) [58] were computed.  $GD$  evaluates the Euclidean distance between the front and the true Pareto, by using the cost functions as space coordinates. On the contrary, the  $C$ -metric computes the number of solutions in the front that are dominated by the real Pareto front.

Finally, the solution quality is quantified through some diversity metrics frequently used in the literature. In this regard, the  $\Delta$  indicator proposed by Deb et al. [16] is one of the most popular since it computes the the distance, in the cost function space, between extreme solutions and the distance from the nearest neighbor. Deb et al. [16] proposed also the average crowding distance ( $aCD$ ) to evaluate the solution density in the Pareto fronts. Another metric often implemented is the spacing ( $Sp$ ) [59].  $Sp$  assesses how evenly the members of the algorithm solutions are distributed and it approaches zero when the solutions are equidistant. In the same vein, the pure diversity ( $PD$ ) index [60] indicates the "biodiversity of two species" meant as Pareto solutions. Since the  $PD$  value depends also on the optimization problem, the normalized pure diversity  $nPD$  was defined as the ratio of the pure diversity over the pure diversity of the actual Pareto front.

Finally two metrics considering multiple aspects were computed. The inverted generational distance ( $IGD$ ) [61] computes the minimum Euclidean distance between the Pareto front identified by the optimization algorithm and the true Pareto front and, consequently, it accounts for quality and efficacy. Similarly, the hypervolume (a.k.a.  $S$ -metric) considers the efficacy and diversity since it measures the size of the objective space covered by the optimal solutions [46]. Since the hypervolume depends on the optimization

problem, the normalized hypervolume ( $nHV$ ) defined as the ratio of the  $S$ -metric over the  $S$ -metric of the true Pareto front was used.

#### 2.4. Optimization problems

The performances of the optimization algorithm with different meta models are evaluated by applying the algorithm to a refurbishment of three simplified buildings that are representative of an intermediate flat (*IF*) in a multi-storey building, a penthouse (*PH*) and a semi-detached house (*SD*)(Figure 3). The building characteristics are fully described in [40] and represent a

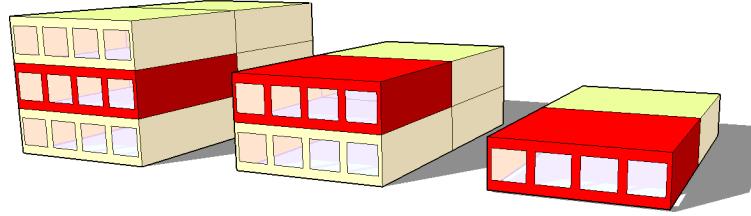


Figure 3: Reference buildings used in the optimization problems

typical configuration of houses built prior to the first Italian energy law and not renovated yet. The *BPS* are carried out in Trnsys simulation suite by considering the weather data of Milan, a city having a 4A climate according to ASHRAE classification [62]. A standard gas boiler coupled with radiators and on-off control system is the original HVAC system for all the test cases. The model developed by Haller et al. [63] computes the efficiency variation of the boiler due to the part load operation in the *BPS*. According to Kusuda [64], the boiler model is a detailed heat and mass balance model that has

already been validated in [65]. Six conventional retrofit solutions are considered for the enhancement of the energy performances and the initial cost of each retrofit solution is defined in [40].

- *external insulation* of the external envelope with an additional expanded polystyrene layer, with thicknesses ranging from 0 to 20 cm in steps of 1 cm. The insulation thickness is changed independently for vertical walls, roof and floor as well as different initial costs were considered;
- *windows replacement* with more efficient glazing systems namely double or triple pane with either high or low solar heat gain coefficient (*SHGC*);
- *boiler replacement* with either a modulating or a condensing boiler with an outdoor reset control;
- *mechanical ventilation system* installation with a cross-flow heat recovery system.

The energy efficiency enhancement and the minimization of the total cost are the objectives of the *BPO* according to the European framework of the cost optimality. These targets are quantified through the Primary Energy for Heating ( $EP_H$ ) and the Net Present Value ( $NPV$ ) indexes. The  $EP_H$  is computed in TRNSYS by summing the natural gas consumptions and the auxiliary power expressed in terms of primary energy by means of the primary energy factors, that are 1.05 for natural gas and 2.42 for power, according to the procedure defined in [66]. The economic evaluation was performed

according to the comparative framework methodology of cost-optimal level, proposed by the EU 244/2012 [67]. The  $NPV$  was calculated by summing the different cash flows, over a 30-year lifespan, related to each interventions following EN 15459 [68].

The total  $ESM$  combinations ( $n_{ESM}$ ) are 277830 for the semidetached house while, due to the adiabatic boundary conditions of roof and floor, they decrease to 13230 and 630 respectively for the penthouse and the intermediate flat (Figure 4).

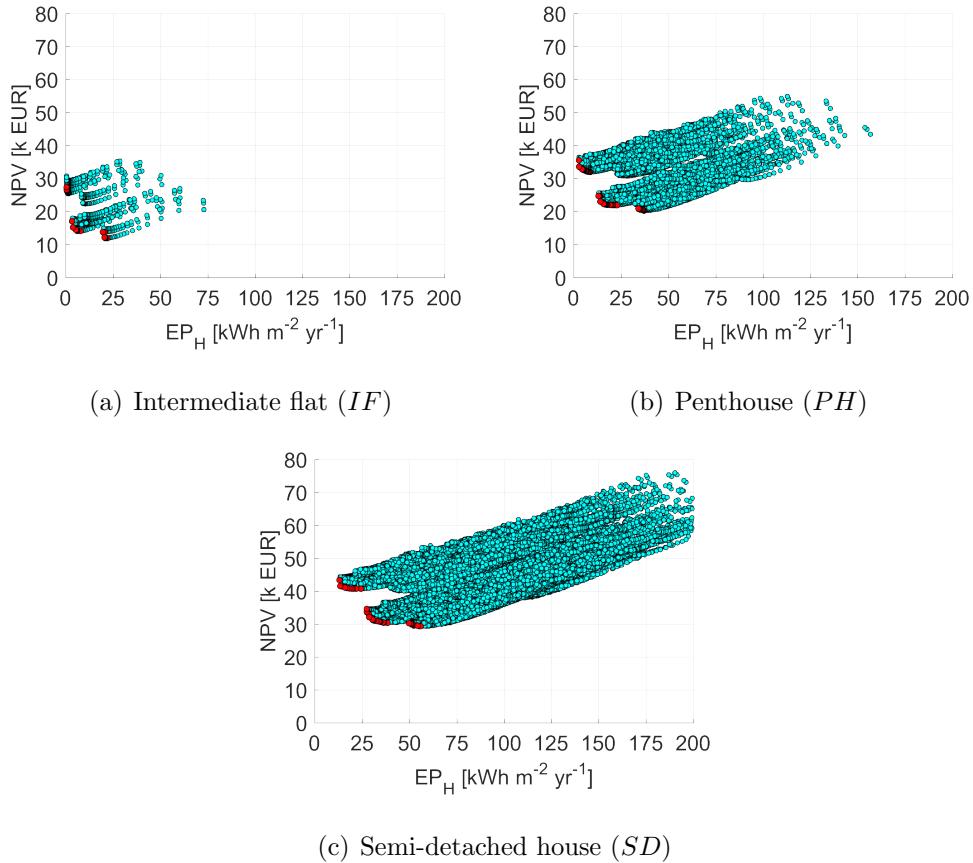


Figure 4: Actual Pareto front obtained by means of Brute force approach

Figure 4 highlights also the higher solution density around the Pareto front for the semi-detached house and the penthouse in addition to the largest number of *ESM* combinations. Hence, the three test cases also have different complexity in terms of the correct identification of the non-dominated solutions. For this reason they are not intended as three simple buildings, but rather as three optimization problems with different complexity levels.

### 3. Results

This study focuses on the analysis of the performance of eight surrogate models combined with three different sampling techniques, eight population sizes (reduced to 6 for the *IF*) and three test problems (i.e. test buildings), as highlighted in the previous sections. The normalized population ( $nP$ ) has been defined as the ratio of the population size over the number of *ESM* combinations ( $n_{ESM}$ ) for each optimization problem. In this way it is possible to simultaneously treat test problems with different characteristics in terms of the number and complexity of the solutions, in order to have global indications on the performance of the different algorithm configurations. The results are on the whole 528 approximated Pareto fronts obtained by means of each configuration of the optimization algorithms to be compared to the three true fronts obtained with the brute force approach. For this reason, the ten metrics, introduced in section 2.3, allow us to systematically compare and quantify the three different performances required to a *BPO* algorithm. Some of the metrics are related to each other and, consequently, they provide redundant information. However, the purpose of the metrics is to describe the performance space in a more independent and comprehensive way, which

mathematically translates into the search for linearly independent vectors. For this reason, an analysis of the correlations was conducted on the ten metrics in order to better identify the metrics triad able to better represent the three dimensional vector space. The Table 2 refers to figures of correlation

Table 2: Correlation matrix of performance metrics

	<b>NE</b>	<b>PS</b>	<b>GD</b>	<b>C</b>	<b>Δ</b>	<b>aCD</b>	<b>Sp</b>	<b>nPD</b>	<b>IGD</b>	<b>nHV</b>
<b>NE</b>	1	0.8239	<b>-0.3315</b>	-0.8208	-0.0505	0.1121	0.2863	<b>0.0383</b>	-0.4843	0.5017
<b>PS</b>	0.8239	1	-0.4366	-0.9987	-0.0434	0.0566	0.3139	0.2110	-0.6006	0.6360
<b>GD</b>	<b>-0.3315</b>	-0.4366	1	0.4472	0.2036	0.5497	0.3147	<b>-0.1643</b>	0.8363	-0.7616
<b>C</b>	-0.8208	-0.9987	0.4472	1	0.0482	-0.0418	-0.3060	-0.2021	0.6121	-0.6467
<b>Δ</b>	-0.0505	-0.0434	0.2036	0.0482	1	0.2271	0.6557	-0.0694	0.2096	-0.2328
<b>aCD</b>	0.1121	0.0566	0.5497	-0.0418	0.2271	1	0.6264	-0.0644	0.4899	-0.3225
<b>Sp</b>	0.2863	0.3139	0.3147	-0.3060	0.6557	0.6264	1	-0.0409	0.1769	-0.1088
<b>nPD</b>	<b>0.0383</b>	0.2110	<b>-0.1643</b>	-0.2021	-0.0694	-0.0644	-0.0409	1	-0.2618	0.3298
<b>IGD</b>	-0.4843	-0.6006	0.8363	0.6121	0.2096	0.4899	0.1769	-0.2618	1	-0.9292
<b>nHV</b>	0.5017	0.6360	-0.7616	-0.6467	-0.2328	-0.3225	-0.1088	0.3298	-0.9292	1

coefficient for the different metrics evaluated for the 528 optimization algorithm configurations. It shows that *NE*, *GD* and *nPD* are by far the metrics with a lower correlation per each of the three categories. Consequently, this triad is able to provide complementary information and to identify the best overall performance of an optimization algorithm configuration over the others. Nonetheless, in the following sections we also present the configurations of the optimization algorithm increasing either the efficiency and the efficacy, the metrics that analyze multiple aspects or the whole set of metrics calculated in this research in order to verify the extent to which these choices affect the optimal algorithm configurations.

### 3.1. Best overall performance

The algorithm configurations with the best overall performances are the non-dominated solutions in the vector space identified by the metric triad

$NE$ ,  $GD$  and  $nPD$ . The optimal  $BPO$  algorithm will be able to minimize  $NE$ ,  $GD$  and, at the same time, to maximize  $nPD$ .

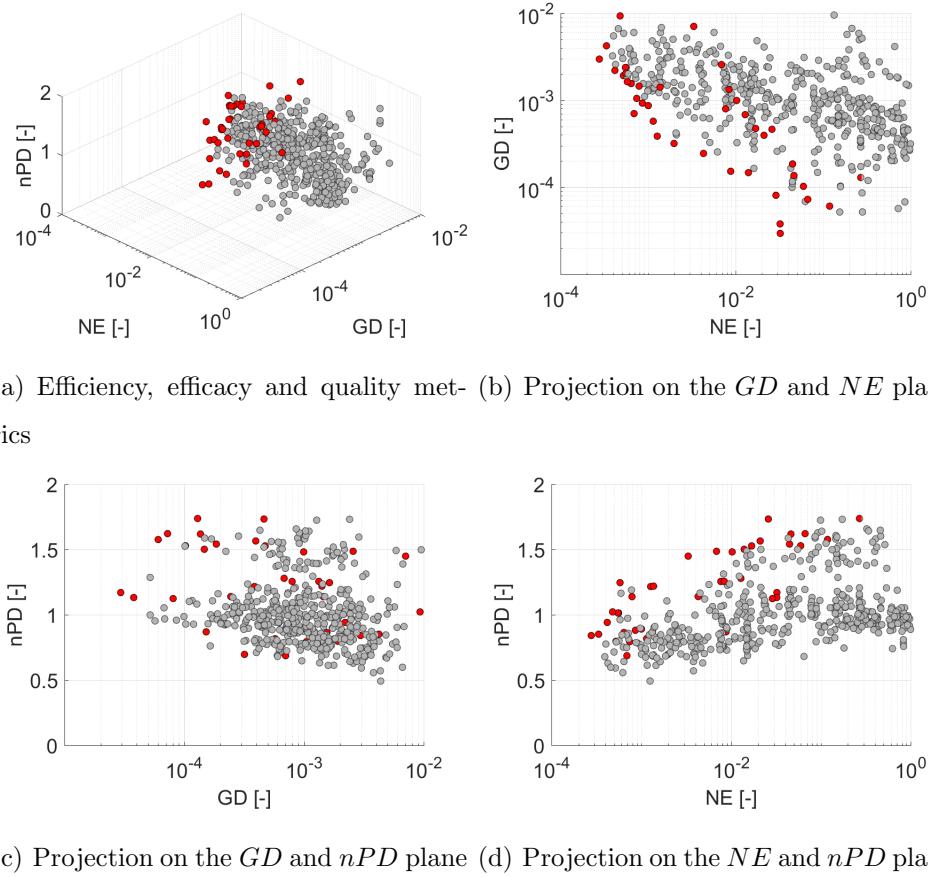


Figure 5: Identification of algorithms configurations with the best performance metrics

The graph in Figure 5a shows the set of metrics obtained for the 528 configurations of the investigated  $BPO$  algorithm. The red color has been used to indicate the non-dominated points. Then, the results have been projected on the different planes, in order to better visualize the trends (Figure 5b to 5d). Notice that there is not a single optimal configuration of the  $BPO$

algorithm in Figure 5a, since the three optimal performances are in conflict with each other. As can be seen from the chart in Figure 5b,  $GD$  and  $NE$  are conflicting goals since the convergence is improved at the cost of a high number of expensive  $BPS$  runs. Moreover, in the same graph we note that the presence of the  $nPD$  target leads to the identification of a second branch of non-dominated solutions for  $NE$  in the range from  $3 \cdot 10^{-3}$  to  $7 \cdot 10^{-2}$ . This result therefore indicates that, in this region, it is not possible to minimize  $GD$  and simultaneously maximize  $nPD$ . Hence, the minimization of the distance from the real Pareto front takes place at the expense of the  $nPD$  index. In fact, the  $nPD$  evaluates how solutions are spread on the Pareto front, namely their diversity and uniformity, but it also penalizes any repeated objective values that are redundant to decision markers. This aspect explains the red point distribution in the  $nPD$  and  $GD$  plane (Figure 5c). Also in this case two fronts are highlighted, due to the trade-off between the achievement of the efficiency, efficacy and quality targets. Figure 5c and 5d show how similar  $GD$  values can be achieved with few expensive simulations but with a reduced  $nPD$  metric, or with a greater  $nPD$  but with a higher computational cost in order to find the better combination of solutions that are not redundant for the decision maker.

Following on from this point, it becomes important to understand which  $BPO$  algorithm configurations produced the best overall performance, represented by red dots in Figure 5. Hence, the configuration was analyzed in terms of the sampling method, the population size, the meta-model and the optimization problems that produced each non-dominated solution in terms of the metric triad. The results were represented using a scatter-plot matrix

method (Figure 6). In this graph the diagonal sub-plots show the distribution of each parameters for the non dominated solutions, while the sub-plot in the  $(i, j)$  position is a scatter plot of parameter  $i$  against parameter  $j$ . The first result shown in the Figure 6 is related to the optimization problem

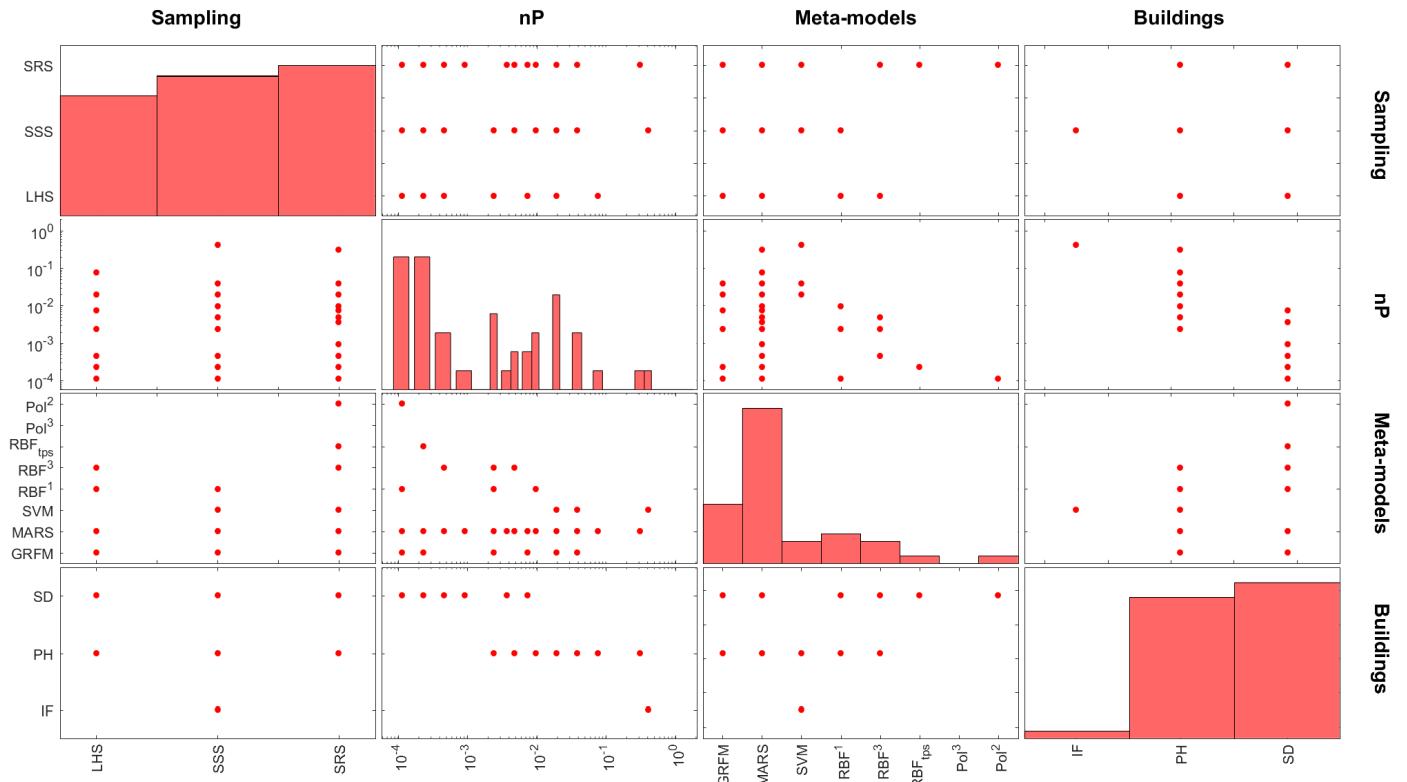


Figure 6: Algorithm configurations with the non-dominated  $NE$ ,  $GD$  and  $nPD$  metrics

distributions in the  $(4, 4)$  sub-plot position. Basically, the solutions with the non-dominated metrics were obtained mainly for cases  $PH$  and  $SD$  and only a limited number of non-dominated performance are obtained for the  $IF$

building. In detail, the *IF* cases are always combined with *SVM* method, Sobol sampling and a value of  $nP$  close to one as illustrated by the sub-plots in the fourth row in Figure 6. This result is mainly due to the simplicity of the optimization problem, for which the solution with an optimization algorithm is therefore not very efficient with respect to the brute force approach.

The graph in the (1, 1) position shows the distribution of sampling method in the non-dominated metric triads. As can be seen from this chart, the distribution is fairly uniform with a slight prevalence of the *SRS* method compared to *SSS* and *LHS*. The first column of the graph also shows how all sampling methods are substantially present with all the figures of the other parameters. This aspect therefore highlights the limited impact of sampling algorithms on the performance.

As can be seen from the chart in the position (2, 2), the  $nP$  varies according to a bimodal distribution in which two peaks emerge between  $1.15 \cdot 10^{-4}$  and  $2.30 \cdot 10^{-4}$  and for the value of  $1.93 \cdot 10^{-2}$ . This result highlights how two population sizes are more recurrent in the non-dominated algorithm configurations.

By analyzing the metrics of these two levels of  $nP$  in details, we note that the first peak is substantially related to the most efficient algorithm configurations. On the contrary, the algorithm configurations with the largest  $nP$  peak are those with the best performances in terms of *GD* and *nPD*. This result indicates that the pursuit of solution quality requires a greater number of expensive simulations and therefore a reduction in efficiency.

The most important result of this analysis is to a large extent emphasized by the distribution of surrogate models in solutions with the best metric

triads. As can be seen from the graph in the position (3,3), the *MARS* model is much more frequent than the others. This highlights the extent to which this surrogate model, coupled with the genetic optimization code, is able to improve the overall performance. It is also noted that the second most frequent model is the *GRFM* model while the other meta-models are significantly less present in configurations with optimal performance.

### 3.2. Best efficiency and efficacy performance

One of the possible strategies to increase the overall performance of the optimization algorithm is to obtain the solutions that are closer to the real front at the lowest computational cost and then operate by pruning the solutions through a post Pareto analysis in order to increase the *nPD* metric. For this reason, the same analysis conducted in the previous section was repeated by looking at the non-dominated solutions in terms of the efficiency (*NE*) and the efficacy (*GD*) metrics. The non-dominated solutions in terms of the *GD* and the *NE* can be obtained by applying the domination concept to the best solutions obtained in the previous section. Thus, the *GD* and *NE* non-dominated points are the lower red points in Figure 5b.

The scatter-plot matrix in Figure 7 shows the parameters that lead to the achievement of the best efficiency and efficacy. Renouncing the quality goal of identifying the optimal performance does not lead to relevant changes in the optimal configurations. There is a slight variation in the optimization problem distribution in Figure 7. The *IF* case is completely absent while the presence of *PH* is considerably reduced. The influence of the sampling method is again modest as shown by the uniform distribution in the upper left-hand sub-plot. Contrary to Figure 6, there is a slight prevalence of

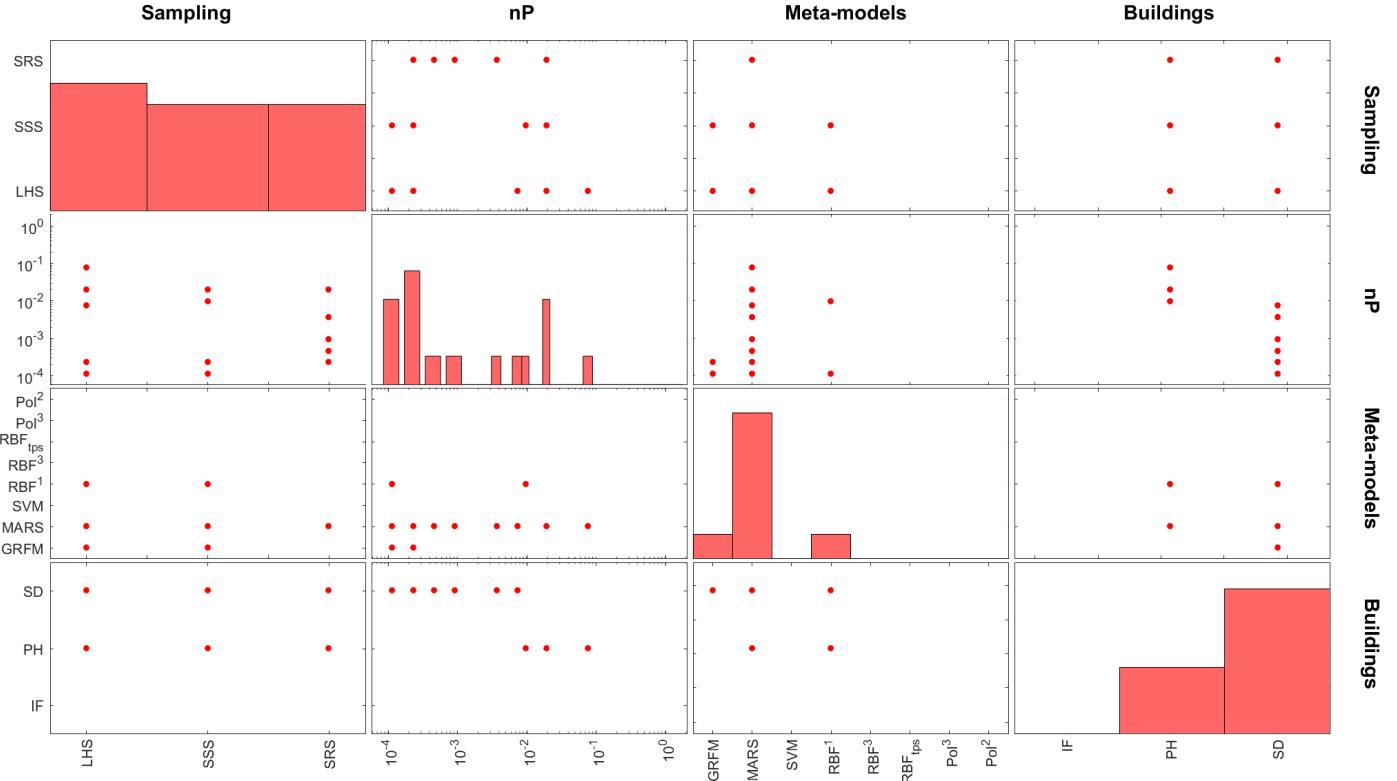


Figure 7: Algorithm configurations with the best  $NE$  and  $GD$  metrics

the *LHS* method compared to the other two. Nonetheless, an effect of the sampling method on the results is noted in Table 3. All the three sampling methods are present for the test *SD*, a  $nP$  of  $2.31 \cdot 10^{-4}$  and by using the *MARS* model. In these cases, enhanced sampling techniques (*LHS* or *SSS*) lead to a lower  $NE$ . However, this high efficiency is counterbalanced by the efficacy detriment (*LHS*) or by a lower solution quality (*SSS*) of the obtained front, indicated by the low  $nPD$  but not by the other quality metrics.

Table 3: Metrics of non-dominated solutions in terms of  $NE$  and  $GD$ 

NE	PS	GD	C	$\Delta$	aCD	Sp	nPD	IGD	nHV	Samp	nP	Model	Case
0.0140	0.256	$1.47 \cdot 10^{-4}$	0.744	0.641	0.031	0.0034	1.5018	$5.64 \cdot 10^{-3}$	0.9577	SSS	$9.68 \cdot 10^{-2}$	RBFlin	PH
0.0324	0.808	$2.95 \cdot 10^{-5}$	0.192	0.676	0.012	0.0056	1.1699	$5.49 \cdot 10^{-4}$	0.9986	LHS	$1.94 \cdot 10^{-2}$	MARS	PH
0.0322	0.846	$3.79 \cdot 10^{-5}$	0.154	0.636	0.011	0.0054	1.1318	$4.19 \cdot 10^{-4}$	0.9990	SSS	$1.94 \cdot 10^{-2}$	MARS	PH
0.0289	0.564	$8.08 \cdot 10^{-5}$	0.436	0.659	0.015	0.0062	1.1239	$2.09 \cdot 10^{-3}$	0.9954	SRS	$1.94 \cdot 10^{-2}$	MARS	PH
0.0831	0.756	0	0.244	0.717	0.013	0.0058	1.4864	$3.08 \cdot 10^{-3}$	0.9854	LHS	$7.74 \cdot 10^{-2}$	MARS	PH
0.0006	0.015	$1.55 \cdot 10^{-3}$	0.955	0.786	0.022	0.0059	0.8639	$5.46 \cdot 10^{-3}$	0.9651	LHS	$1.15 \cdot 10^{-4}$	GRFM	SD
0.0005	0.015	$1.92 \cdot 10^{-3}$	0.985	0.991	0.037	0.0115	0.7965	$1.16 \cdot 10^{-2}$	0.9353	LHS	$1.15 \cdot 10^{-4}$	RBFlin	SD
0.0004	0.000	$2.20 \cdot 10^{-3}$	0.985	0.858	0.032	0.0100	0.9411	$7.02 \cdot 10^{-3}$	0.9529	SSS	$1.15 \cdot 10^{-4}$	MARS	SD
0.0006	0.030	$1.63 \cdot 10^{-3}$	0.955	0.686	0.030	0.0079	1.2461	$4.67 \cdot 10^{-3}$	0.9518	LHS	$2.30 \cdot 10^{-4}$	MARS	SD
0.0003	0.000	$2.97 \cdot 10^{-3}$	1.000	1.053	0.107	0.0194	0.8409	$1.38 \cdot 10^{-2}$	0.8917	SSS	$2.30 \cdot 10^{-4}$	GRFM	SD
0.0011	0.119	$5.74 \cdot 10^{-4}$	0.836	0.604	0.015	0.0027	0.8161	$2.84 \cdot 10^{-3}$	0.9729	SSS	$2.30 \cdot 10^{-4}$	MARS	SD
0.0013	0.090	$3.86 \cdot 10^{-4}$	0.866	0.632	0.017	0.0043	1.2162	$2.34 \cdot 10^{-3}$	0.9874	SRS	$2.30 \cdot 10^{-4}$	MARS	SD
0.0007	0.045	$7.03 \cdot 10^{-4}$	0.940	0.909	0.025	0.0087	0.6860	$4.61 \cdot 10^{-3}$	0.9554	SRS	$4.61 \cdot 10^{-4}$	MARS	SD
0.0020	0.164	$3.18 \cdot 10^{-4}$	0.776	0.681	0.013	0.0019	0.6961	$2.62 \cdot 10^{-3}$	0.9828	SRS	$9.21 \cdot 10^{-4}$	MARS	SD
0.0043	0.149	$2.45 \cdot 10^{-4}$	0.761	0.614	0.014	0.0051	1.1394	$1.62 \cdot 10^{-3}$	0.9912	SRS	$3.69 \cdot 10^{-3}$	MARS	SD
0.0088	0.284	$1.52 \cdot 10^{-4}$	0.597	0.686	0.011	0.0047	0.8689	$9.42 \cdot 10^{-4}$	0.9941	LHS	$7.37 \cdot 10^{-3}$	MARS	SD

Similarly, in the  $PH$ , with a  $nP$  of  $1.94 \cdot 10^{-2}$  and the  $MARS$  model the  $SRS$  rewards efficiency at the expense of efficacy. A general trend is not found and, for this reason, the frequency of the three sampling algorithms in the solutions with non-dominated metrics is comparable. Rather, we can deduce a problem related to the lack of analysis of some regions of the variable space due to the choice of the initial population which can lead to an early algorithm convergence.

The most interesting results, however, are the confirmations of what emerged for the distributions of  $nP$  and of the surrogate models. A bimodal distribution is noted again, with the two peaks centered on the same  $nP$  values of those in Figure 6. Nonetheless, we note the absence of some solutions, and especially the one with high  $nP$ ,  $SVM$ , polynomial and high order  $RBF$  models, compared to the same subplot in Figure 6. This indicates again that the larger population is required above all to increase the solution quality rather than improving its efficacy. Furthermore, the best

performance of the *MARS* method is confirmed and it even increases the gap with the *GRFM* method whose frequency becomes comparable to that of the *RBF*<sup>1</sup>.

### 3.3. Best performance measured by multiple aspect metrics

As highlighted in the section 2.3, there are two metrics that consider efficacy and solution quality simultaneously. For this reason, solutions with the best performance in terms of efficiency (*NE*), efficacy and solution quality (*nHV* and *IGD*) were also analyzed.

The graphs in Figure 8 indicates the parameter distributions ensuring the best performance quantified by the *NE* and the multiple aspect metrics. With respect to the metric triad presented in Figure 6, there is a remarkable change in the distribution of the optimization problems and of the sampling methods. The *SD* case is to a considerable extent the most relevant but the other two buildings become comparable with each other in terms of frequency as highlighted in the sub-plot in the position (4, 4). Similarly, sampling according to the *LHS* method tends to be more frequent than the other two techniques. The multiple aspect metrics therefore are able to highlight the different performances obtainable for the *PH* and *IF* cases, since the *IF* metrics that are no longer dominated by the *PH* and *SD* cases. Some variations in the *nP* distribution are also noted. The *nP* distribution shows three peaks in lieu of the two obtained with the triad *NE*, *GD* and *nPD*. The first two peaks were also shown in Figure 6. Nonetheless, a new peak appears by looking at the optimal *NE*, *nHV* and *IGD*. The new frequency peak corresponds to *nP* equal to  $2.03 \cdot 10^{-1}$  and it is associated with the intermediate flat in the apartment block, for which a higher efficacy and

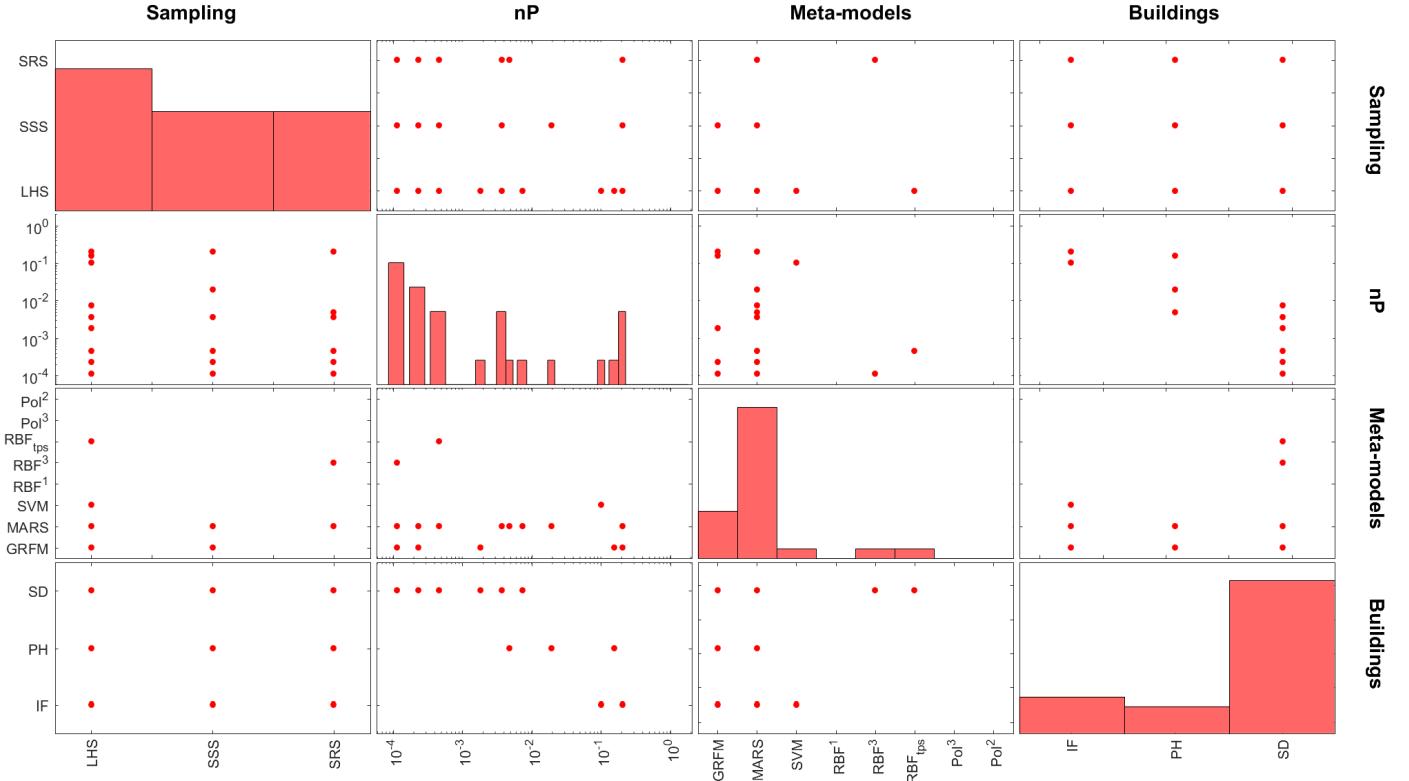


Figure 8: Algorithm configurations with the best  $NE$ ,  $nHV$  and  $IGD$  metrics

solution quality can be achieved. Though to be thorough, the use of  $GD$  and  $nPD$  does not emphasize this result since the  $IF$  metrics are dominated by those of  $PH$  and  $SD$ . The meta-model frequencies again show what has already emerged in the previous analyzes. The  $MARS$  model actually used continues to be by far the best performing surrogate model, while the  $GRFM$  and the  $RBFN$  are present only in a limited configuration number.

### *3.4. Trade-off among all the implemented performance indicators*

Finally, the analysis focuses on the solutions able to provide the trade-off among the ten metrics used in this research. In Table 2 there is not a perfect correlation between metrics that measure the same performance. Hence, all the metrics consider different aspects and they provide complementary pieces of information. In this section the extent to which this comprehensive information affects the optimal algorithm configurations is evaluated.

Note that a greater configuration number emerges in Figure 9 and the number of configurations with non-dominated metrics rises to 181 out of the 528 cases analyzed. This result shows how the same metric categories actually evaluate different aspects, which is very evident especially for the quality metrics. Analyzing the graphs, we note the substantially uniform distributions with regard to sampling techniques and the optimization problems.

The pursuit of the trade-off among all the quality metrics of the Pareto front leads to the increase of solutions with high  $nP$ . Similarly, even if the *MARS* model still outperforms the others, the other meta-models are more frequent than in previous cases.

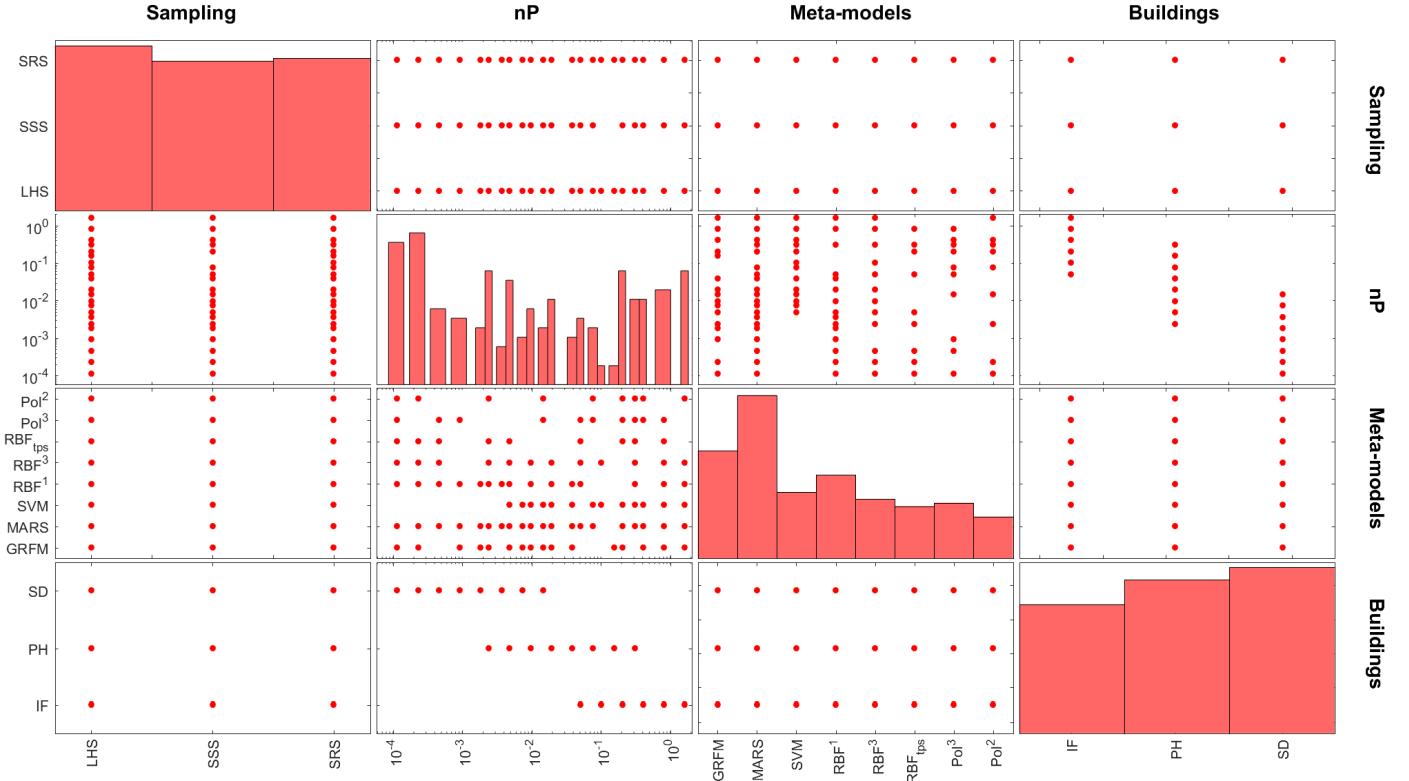


Figure 9: Algorithm configurations with the best 10 metrics

### 3.5. Comparison against *NSGA – II*

Finally, the benefits of the "generational-based approach" with respect to *EA* are quantified by comparing the computational times (Table 4). The time required by the different algorithm configurations (Figure 2) are compared against the time needed by the *NSGA – II* algorithm to solve the problem if the meta-models are not used to filter-out the unpromising space regions. All the other settings of the *NSGA – II*, such as the external

dataset, the sampling methods and the convergence criterion reflect what is implemented for the "generational-based approach" in order to guarantee the result comparability. Table 4 shows the time in seconds required by the *NSGA-II* to solve the optimization problems, starting from the initial population sampling, by the *LHS*, until convergence is reached. The other sampling methods provide similar results, due to the low impact of sampling techniques. The other columns highlight the percentage time savings obtained using the different meta-models, as the initial population size ( $nP$ ) and the test case (*Case*) vary. Negative values indicate a time saving and, consequently, an effective advantage in the use of meta-models.

Table 4: Comparison of optimization computational time in [s] of generational-based approach against *NSGA-II*

<b><math>nP</math></b>	<b>Case</b>	<b>NSGA - II</b>	<b>GRFM</b>	<b>MARS</b>	<b>SVM</b>	<b>RBF<sup>1</sup></b>	<b>RBF<sup>3</sup></b>	<b>RBF<sub>tps</sub></b>	<b>Pol<sup>2</sup></b>	<b>Pol<sup>3</sup></b>
0.2032	<i>IF</i>	13408	-22 %	-34 %	-25 %	-41 %	-45 %	-43 %	80 %	34 %
0.4063	<i>IF</i>	18033	-37 %	-46 %	-28 %	-39 %	-45 %	-43 %	44 %	35 %
0.8127	<i>IF</i>	22716	-27 %	-32 %	-23 %	-28 %	-30 %	-30 %	10 %	-34 %
0.0097	<i>PH</i>	27679	-71 %	-74 %	26 %	-73 %	-73 %	-70 %	754 %	-21 %
0.0193	<i>PH</i>	50926	-51 %	-64 %	-34 %	-71 %	-66 %	-76 %	5 %	-40 %
0.0387	<i>PH</i>	91611	-74 %	-73 %	-70 %	-73 %	-72 %	-71 %	171 %	-17 %
0.0774	<i>PH</i>	150745	-67 %	-70 %	-54 %	-72 %	-70 %	-71 %	21 %	-5 %
0.1548	<i>PH</i>	235836	-55 %	-67 %	-43 %	-66 %	-66 %	-66 %	22 %	5 %
0.3096	<i>PH</i>	336174	-57 %	-57 %	-30 %	-57 %	-54 %	-55 %	-3 %	14 %
0.0005	<i>SD</i>	35525	-56 %	-77 %	8 %	-63 %	-68 %	-69 %	-10 %	282 %
0.0009	<i>SD</i>	69981	-76 %	-82 %	-14 %	-60 %	-78 %	-64 %	19 %	106 %
0.0018	<i>SD</i>	134057	-69 %	-83 %	-75 %	-74 %	-80 %	-83 %	187 %	67 %
0.0037	<i>SD</i>	256222	-83 %	-81 %	-57 %	-78 %	-81 %	-82 %	-52 %	34 %
0.0074	<i>SD</i>	487812	-81 %	-78 %	-65 %	-81 %	-82 %	-78 %	-49 %	-49 %
0.0147	<i>SD</i>	865135	-76 %	-79 %	-76 %	-78 %	-79 %	-75 %	64 %	-80 %

Looking at the results, we notice some positive values for the polynomial models ( $Pol^2$  and  $Pol^3$ ). This is mainly due to the lack of accuracy with which these models can emulate *BPS*. As a result, many simulations are required before the convergence criterion is met, and the time required to estimate the polynomial coefficients can produce an increase in computational

time compared to the direct use of *BPS* in the *NSGA-II*. Furthermore, we note a great sensitivity of time savings to the value of  $nP$  for polynomial and *SVM* models. For these meta-models an evaluation of the optimal size of the initial population is required before proceeding to the optimization. This is not found instead for the other meta-models that are more robust to the  $nP$  choice.

There is a general trend of increasing saving percentages for more complex test cases (i.e. *PH* and *SD*). Hence, the benefit of the "generational-based approach" increases, when the number of *ESM* combinations grows.

The benefits in using meta-models increase when the possible combinations of *ESM* are high. The *MARS* is the model with the highest savings in 47% of cases in Table 4 and it is close to the greater savings in the other cases. This therefore demonstrates the best efficiency of *MARS* in speeding up the optimization process.

#### 4. Discussions

The results in section 3.5 clearly show the advantage of using surrogate models in the optimization process, especially when the complexity of the optimization problem increases. The lower the number of the possible *ESM* is (e.g. the *IF* building), the lower the performance improvement is when the functional approximation is used in the optimization algorithm. The *PS* values in Table 3 emphasizes the identification of 75% to 80% of actual optima with a *GD* distance in the order of  $10^{-5}$ , by simulating only 3% to 8% of the *ESM* combinations (i.e. *NE* values). It is possible to converge even simulating less than 1% of the combinations, by accepting an order of

magnitude more on the  $GD$  distance (i.e.  $10^{-4}$ ), but with a  $nHV$  close to one. This benefit is essentially due to the reduced number of evaluations in areas far from the Pareto front. In these areas it is therefore appropriate to have only a few assessments necessary to adapt the surrogate model to  $BPS$ . For this reason, it is important that the initial population is as uniform as possible in the variable space in order to reduce the risk of having little coincidence between the  $BPS$  and the meta-model outcomes. If the meta-model produces non-optimal cost functions in these areas, they will never be evaluated with the  $BPS$  code and, consequently, the meta-model will never be improved. For this reason a mutation operator can be useful in reducing this risk although this will reduce the algorithm efficiency.

Another consideration deals with the implemented metrics. The metrics belonging to the same categories have medium-large correlation coefficients in Table 2 except for the quality metrics. It is noted that  $nPD$  is poorly correlated with all the others quality metrics, likewise,  $aCD$  is poorly correlated with  $\Delta$ -metric. This testifies how quality metrics evaluate different aspects. Some metrics only evaluate the solution diversity, whereas the others take into account also the uniformity or the information redundancy ( $nPD$ ). The latter metric has the advantage of finding the subset of solutions capable of maximizing information for the decision maker but with a limited solution number. The use of  $nPD$  as a convergence criterion therefore would ensure the overcoming of another problem in the dissemination of  $BPO$  in the practice. For a designer it becomes difficult to present all the alternatives identified by the Pareto front and for a decision maker it becomes even more difficult to grasp the differences and make a choice. Therefore, the automatic

reduction of the redundant solutions simplifies the procedure. However, as shown by the results, the pursuit of a high  $nPD$  value leads to an increased number of expensive simulations and of the optimal size of the initial population. Therefore, the reduction of the solution number in the Pareto front to help decision maker, can be more efficiently done by means of a post Pareto analysis.

## 5. Conclusions

In this work we investigated the efficiency, the efficacy and the solution quality of several meta-models for speeding up the multi-objective optimization of three building refurbishments. The results confirm that the use of surrogate models can significantly speed up the optimization process leading to good results in terms of convergence to the true Pareto front with a limited number of evaluations of expensive cost functions.

The analysis showed how the metrics used to evaluate the solution quality actually highlight different aspects and provide complementary information. In this respect, the triad composed of the expensive simulation number ( $NE$ ), the distance from the true Pareto front ( $GD$ ) and the normalized pure diversity ( $nPD$ ) is able to provide complementary information on the algorithm performances and therefore to analyze in a more independent way the efficiency, the efficacy and the quality of the obtained fronts.

The optimal algorithm configurations point out that the sampling methods have a limited incidence on the algorithm performances. The sampling method operates mainly in the choice of the initial population, while the successive populations are guided by the Pareto front of the meta-model op-

timization. Nonetheless, a more uniform sampling allows a reduction in the risk of not analyzing some areas of the variable space and to better fit the meta-models to *BPS* outcomes.

The extent to which a greater optimal population is required in order to improve the quality of the Pareto front emerges from the analysis of the performance of the algorithm configurations. This is evident when the *nPD* is adopted as an optimization algorithm goal and, even more, when the solution quality is simultaneously quantified by means of several quality metrics. For this reason a post Pareto analysis, by means of result ranking or solution clustering, is more convenient to check the quality and the redundancy of the retrofit solutions to be presented to the decision maker.

The results highlight the surrogate model that better emulates the building dynamic simulation code. The *MARS* model is to a large extent the best meta-model in the different combinations of metrics and optimization problems analyzed in this research. Hence, the *MARS* model is the most suitable for the purpose of speeding up and improving the *BPO* process. The use of the *MARS* model with a piecewise-cubic function in the *BPO* is therefore a possible way to increase the dissemination in professional practice of the optimization of building refurbishment design.

## Nomenclature

$\beta$       Unknown coefficient of surrogate models

$\gamma$       Primal Linear problem coefficient of the *SVM* model

$\Phi$       Radial basis function of the *RBFN* model

- $\mu^{(j)}$  Center of the j-th radial basis function of the *RBFN* model
- $aCD$  Average Crowding Distance [-]
- $B_j$  j-th Basis Function in MARS model
- $BPO$  Building Performance Optimization
- $BPS$  Building Performance Simulation
- $EA$  Evolutionary Algorithms
- $EGO$  Efficient Global Optimization algorithm
- $EP_H$  Energy Performance for Heating [ $kWh\,m^{-2}\,y^{-1}$ ]
- $ESM$  Energy Saving Measures
- $f_i$  i-th cost function of the optimization problem
- $\bar{f}$  vector of the cost functions
- $g_i$  i-th regression function of the *GRFM*
- $GD$  Generational Distance [-]
- $GRFM$  Gaussian Random Field Model (a.k.a Kriging)
- $IF$  Intermediate Flat in an apartment building
- $IGD$  Inverted Generational Distance [-]
- $HV$  Hypervolume [-]
- $LHS$  Latin Hypercube Sampling

*MARS* Multivariate Adaptive Regression Spline model

*MVS* Mechanical Ventilation System

$n_{BPS}$  number of expensive *BPS* runs [−]

$n_{Brute}$  number of solutions in the true Pareto front obtained by the Brute Force [−]

$NE$  Normalized number of Expensive simulation runs [−]

$n_{ESM}$  number of *ESM* combinations [−]

$nHV$  Hypervolume normalized with respect to the *HV* of the brute force Pareto front [−]

$nP$  Population size of the optimization algorithm normalized with respect to  $n_{ESM}$  [−]

$n_{Par}$  number of solutions in the Pareto front [−]

$nPD$  Pure Diversity normalized with respect to the *PD* of the brute force Pareto front [−]

$NPV$  Net present value [*Eur*]

$PD$  Pure Diversity [−]

$PH$  Penthouse flat

$PS$  Percentage of true Pareto Solutions [−]

*RBFN* Radial Basis Function Network

*SD* Semi-Detached house

*SHGC* Solar heat gain coefficient [–]

*Sp* Spacing [–]

*SRS* Simple Random Sampling

*SSS* Sobol Sequence Sampling

*SVM* Support Vector Machines model

## References

- [1] Directive 2010/31/EU, of the European Parliament and of the council of 19 may 2010 on the energy performance of buildings OJ L 153/2010, 2010.
- [2] P. Brinks, O. Kornadt, R. Oly, Development of concepts for cost-optimal nearly zero-energy buildings for the industrial steel building sector, *Applied Energy* 173 (2016) 343–354.
- [3] F. P. Chantrelle, H. Lahmidi, W. Keilholz, M. E. Mankibi, P. Michel, Development of a multicriteria tool for optimizing the renovation of buildings, *Applied Energy* 88 (2011) 1386–1394.
- [4] J. Yao, Energy optimization of building design for different housing units in apartment buildings, *Applied Energy* 94 (2012) 330–337.

- [5] F. Ascione, R. F. De Masi, F. de Rossi, S. Ruggiero, G. P. Vanoli, Optimization of building envelope design for nZEBs in Mediterranean climate: Performance analysis of residential case study, *Applied Energy* 183 (2016) 938–957.
- [6] R. Wu, G. Mavromatidis, K. Orehounig, J. Carmeliet, Multi-objective optimisation of energy systems and building envelope retrofit in a residential community, *Applied Energy* 190 (2017) 634–649.
- [7] P. Penna, A. Prada, F. Cappelletti, A. Gasparella, Multi-objectives optimization of energy efficiency measures in existing buildings, *Energy and Buildings* 95 (2015) 57–69.
- [8] E. Carlon, M. Schwarz, L. Golicza, V. K. Verma, A. Prada, M. Baratieri, W. Haslinger, C. Schmidl, Efficiency and operational behavior of small-scale pellet boilers installed in residential buildings, *Applied Energy* 155 (2015) 854–865.
- [9] M. Wetter, J. Wright, A comparison of deterministic and probabilistic optimization algorithms for nonsmooth simulation-based optimization, *Building and Environment* 39 (2004) 989–999.
- [10] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*, John Wiley & Sons, Inc., New York, NY, USA, 2001.
- [11] F. Ascione, N. Bianco, C. D. Stasio, G. M. Mauro, G. P. Vanoli, Multi-stage and multi-objective optimization for energy

retrofitting a developed hospital reference building: A new approach to assess cost-optimality, *Applied Energy* 174 (2016) 37 – 68.

- [12] M. Hamdy, A.-T. Nguyen, J. L. Hensen, A performance comparison of multi-objective optimization algorithms for solving nearly-zero-energy-building design problems, *Energy and Buildings* 121 (2016) 57–71.
- [13] J. D. Knowles, H. Nakayama, Meta-Modeling in Multiobjective Optimization, in: Branke, Jürgen and Deb, Kalyanmoy and Miettinen, Kaisa and Słowiński, Roman (Ed.), *Multiobjective Optimization*, February, Springer-Verlag, Berlin, Heidelberg, 2008, pp. 245 – 284.
- [14] R. Evins, A review of computational optimisation methods applied to sustainable building design, *Renewable ans Sustainable Energy Reviews* 22 (2013) 230–245.
- [15] A.-T. Nguyen, S. Reiter, P. Rigo, A review on simulation-based optimization methods applied to building performance analysis, *Applied Energy* 113 (2014) 1043 – 1058.
- [16] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation* 6 (2002) 182–197.
- [17] M. Maaroufi, A. Sosa, B. Galvn, D. Greiner, G. Winter, M. Mendez, R. Aguasca, The role of artificial neural networks in

evolutionary optimisation: A review, in: D. Greiner, B. Galván, J. Périaux, N. Gauger, K. W. Giannakoglou (Eds.), *Advances in Evolutionary and Deterministic Methods for Design, Optimization and Control in Engineering and Sciences*, Springer, Cham, 2015, pp. 59–76.

- [18] S. Attia, M. Hamdy, W. O'Brien, S. Carlucci, Assessing gaps and needs for integrating building performance optimization tools in net zero energy buildings design, *Energy and Buildings* 60 (2013) 110–124.
- [19] Y. Jin, A Comprehensive Survey of Fitness Approximation in Evolutionary Computation, *Soft Computing* 9 (2005) 3–12.
- [20] J. Wang, Z. J. Zhai, Y. Jing, C. Zhang, Particle swarm optimization for redundant building cooling heating and power system, *Applied Energy* 87 (2010) 3668 – 3679.
- [21] V. Corrado, I. Ballarini, D. Dirutigliano, S. Paduos, Cost-optimal analysis of italian office buildings through the application of a quasi-steady state model validated by detailed dynamic simulation, in: *Proceedings of Building Simulation 2015 - 14<sup>th</sup> Conference of International Building Performance Simulation Association*, Hyderabad, India, pp. 2043–2050.
- [22] K. Negendahl, T. R. Nielsen, Building energy optimization in the early design stages: A simplified method, *Energy and Buildings* 105 (2015) 88 – 99.

- [23] Y. Fan, X. Xia, A multi-objective optimization model for energy-efficiency building envelope retrofitting plan with rooftop pv system installation and maintenance, *Applied Energy* 189 (2017) 327 – 335.
- [24] Y. Cascone, A. Capozzoli, M. Perino, Optimisation analysis of pcm-enhanced opaque building envelope components for the energy retrofitting of office buildings in mediterranean climates, *Applied Energy* 211 (2018) 929 – 953.
- [25] H. Sayyaadi, M. Babaie, M. R. Farmani, Implementing of the multi-objective particle swarm optimizer and fuzzy decision-maker in exergetic, exergoeconomic and environmental optimization of a benchmark cogeneration system, *Energy* 36 (2011) 4777 – 4789.
- [26] J. Lee, P. Hajela, Parallel genetic algorithm implementation in multidisciplinary rotor blade design, *Journal of Aircraft* 33 (1996) 962–969.
- [27] K. Klemm, W. Marks, A. J. Klemm, Multicriteria optimisation of the building arrangement with application of numerical simulation, *Building and Environment* 35 (2000) 537 – 544.
- [28] J. H. Lee, Optimization of indoor climate conditioning with passive and active methods using {GA} and {CFD}, *Building and Environment* 42 (2007) 3333 – 3340.

- [29] L. Magnier, F. Haghigat, Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and Artificial Neural Network, *Building and Environment* 45 (2010) 739–746.
- [30] B. Eisenhower, Z. O'Neill, S. Narayanan, V. A. Fonoberov, I. Mezić, A methodology for meta-model based optimization in building energy models, *Energy and Buildings* 47 (2012) 292–301.
- [31] E. Tresidder, Y. Zhang, A. I. J. Forrester, Acceleration of building design optimisation through the use of kriging surrogate models, in: BSO12 Proceedings of the 1<sup>st</sup> IBPSA-England Conference Building Simulation and Optimization, Loughborough, UK, pp. 118–125.
- [32] C. J. Hopfe, M. Emmerich, R. Marijt, J. L. M. Hensen, Robust multi-criteria design optimisation in building design, in: BSO12 Proceedings of the 1<sup>st</sup> IBPSA-England Conference Building Simulation and Optimization, Loughborough, UK, pp. 118–125.
- [33] A. E. Brownlee, J. A. Wright, Constrained, mixed-integer and multi-objective optimisation of building designs by NSGA-II with fitness approximation, *Applied Soft Computing* 33 (2015) 114–126.
- [34] W. Xu, A. Chong, O. T. Karaguzel, K. P. Lam, Improving evolutionary algorithm performance for integer type multi-objective

building system design optimization, *Energy and Buildings* 127 (2016) 714–729.

- [35] X. Chen, H. Yang, A multi-stage optimization of passively designed high-rise residential buildings in multiple building operation scenarios, *Applied Energy* 206 (2017) 541 – 557.
- [36] M. T. Emmerich, K. C. Giannakoglou, B. Naujoks, Single and multiobjective evolutionary optimization assisted by Gaussian random field metamodels, *IEEE Transactions on Evolutionary Computation* 10 (2006) 421–439.
- [37] Y. Jin, M. Olhofer, B. Sendhoff, A framework for evolutionary optimization with approximate fitness functions, *IEEE Transactions on Evolutionary Computation* 6 (2002) 481–494.
- [38] J. Knowles, ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems, *IEEE Transactions on Evolutionary Computation* 10 (2006) 50–66.
- [39] T. Østergård, R. L. Jensen, S. E. Maagaard, A comparison of six metamodeling techniques applied to building performance simulations, *Applied Energy* 211 (2018) 89 – 103.
- [40] P. Penna, A. Prada, F. Cappelletti, A. Gasparella, Multi-objective optimization for existing buildings retrofitting under government subsidization, *Science and Technology for the Built Environment* 21 (2015) 847–861.

- [41] D. G. Krige, A study of gold and uranium distribution patterns in the klerksdorp gold field, *Geoexploration* 4 (1966) 43 – 53.
- [42] C. A. Micchelli, Interpolation of scattered data: Distance matrices and conditionally positive definite functions, *Constructive Approximation* 2 (1986) 11–22.
- [43] J. H. Friedman, Multivariate Adaptive Regression Splines, *The Annals of Statistics* 19 (1991) 1–67.
- [44] V. N. Vapnik, Statistical Learning Theory, Wiley-Interscience, 1998.
- [45] A. Saltelli, T. Andres, F. Campolongo, J. Carboni, D. Gatelli, M. Saisana, S. Tarantola, Global sensitivity analysis, John Wiley & Sons, Chichester (UK), 2004.
- [46] E. Zitzler, L. Thiele, Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach, *IEEE Transactions on Evolutionary Computation* 3 (1999) 257–271.
- [47] M. Fleischer, The measure of pareto optima. applications to multi-objective metaheuristics, in: proceedings of EMO2003 - 2<sup>nd</sup> International conference on Evolutionary Multi-Criterion Optimization, Springer, Faro, Portugal, 2003, pp. 519–533.
- [48] N. Riquelme, C. V. Lucken, B. Baran, Performance metrics in multi-objective optimization, *2015 Latin American Computing Conference (CLEI)* 1 (2015) 1–11.

- [49] J. Knowles, H. Nakayama, Meta-Modeling in Multiobjective Optimization, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 245–284.
- [50] E. Gengembre, B. Ladevie, O. Fudym, A. Thuillier, A kriging constrained efficient global optimization approach applied to low-energy building design problems, *Inverse Problems in Science and Engineering* 20 (2012) 1101–1114.
- [51] J. Yuan, V. Nian, B. Su, Q. Meng, A simultaneous calibration and parameter ranking method for building energy models, *Applied Energy* 206 (2017) 657 – 666.
- [52] S. Lophaven, H. Nielsen, J. Søndergaard, DACE: A Matlab Kriging toolbox, Technical Report, Technical University of Denmark, Lyngby - Denmark, 2002.
- [53] L. Chen, Q. s. Fang, Z. y. Zhang, Research on the identification of temperature in intelligent building based on feed forward neural network and particle swarm optimization algorithm, in: 2010 Sixth International Conference on Natural Computation, volume 4, pp. 1816–1820.
- [54] A. Melo, R. Versage, G. Sawaya, R. Lamberts, A novel surrogate model to support building energy labelling system: A new approach to assess cooling energy demand in commercial buildings, *Energy and Buildings* 131 (2016) 233 – 247.

- [55] X. Chen, H. Yang, K. Sun, Developing a meta-model for sensitivity analyses and prediction of building performance for passively designed high-rise residential buildings, *Applied Energy* 194 (2017) 422 – 439.
- [56] G. Jekabsons, ARESLab - Adaptive Regression Splines toolbox for Matlab/Octave, Technical Report, Institute of Applied Computer Systems - Riga Technical University, Riga (Latvia), 2016.
- [57] D. A. Van Veldhuizen, G. B. Lamont, Evolutionary Computation and Convergence to a Pareto Front, Late Breaking Papers at the Genetic Programming 1998 Conference (1998) 221–228.
- [58] E. Zitzler, K. Deb, L. Thiele, Comparison of multiobjective evolutionary algorithms: empirical results., *Evolutionary computation* 8 (2000) 173–195.
- [59] J. R. Schott, Fault Tolerant Design Using Single and Multicriteria Genetic Algorithm Optimization., Master thesis, Massachusetts Institute of Technology, 1995.
- [60] H. Wang, Y. Jin, X. Yao, Diversity Assessment in Many-Objective Optimization, *IEEE Transactions on Cybernetics* 47 (2017) 1510–1522.
- [61] Q. Zhang, A. Zhou, S. Zhao, P. N. Suganthan, W. Liu, Multi-objective optimization Test Instances for the CEC 2009 Special Session and Competition, Technical Report (2009) 1–30.

- [62] ASHRAE 90.1, Energy standard for buildings except low rise residential buildings, Technical Report, ASHRAE - American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, Georgia, 2007.
- [63] M. Y. Haller, J. Paavilainen, L. Konersmann, R. Haberl, A. D. scher, E. Frank, C. Bales, W. Streicher, A unified model for the simulation of oil, gas and biomass space heating boilers for energy estimating purposes. part I: Model development, Journal of Building Performance Simulation 4 (2011) 1–18.
- [64] T. Kusuda, Standards criteria for hvac systems and equipment performance simulation procedures, ASHRAE Journal 23 (1981) 25–28.
- [65] E. Carlon, V. K. Verma, M. Schwarz, L. Golicza, A. Prada, M. Baratieri, W. Haslinger, C. Schmidl, Experimental validation of a thermodynamic boiler model under steady state and dynamic conditions, Applied Energy 138 (2015) 505 – 516.
- [66] UNI/TS 11300 Energy performance of buildings - Part 5: Evaluation of energy performance of the classification of building, Tech. Stand., UNI - Ente Nazionale Italiano di Unificazione, 2016.
- [67] Commission Delegated Regulation (EU) No 244/2012, of the european commissionof 16 january 2012 supplementing directive 2010/31/eu of the european parliament and of the council on the energy performance of buildings by establishing a compara-

tive methodology framework for calculating cost-optimal levels of minimum energy performance OJ L 81/18, 2012.

- [68] EN 15459 Energy performance of buildings Economic evaluation procedure for energy systems in buildings, Tech. Stand., CEN - European Committee for Standardization, 2009.