



Delft University of Technology

## Improving the Calibration on Building Stock Level method by Comparing Objective Functions and Optimization Algorithms

van den Brom, P.I.; Smets, S.J.; Itard, L.C.M.

### DOI

[10.34641/clima.2022.420](https://doi.org/10.34641/clima.2022.420)

### Publication date

2022

### Document Version

Final published version

### Published in

CLIMA 2022 - 14th REHVA HVAC World Congress

### Citation (APA)

van den Brom, P. I., Smets, S. J., & Itard, L. C. M. (2022). Improving the Calibration on Building Stock Level method by Comparing Objective Functions and Optimization Algorithms. In *CLIMA 2022 - 14th REHVA HVAC World Congress: Eye on 2030, Towards digitalized, healthy, circular and energy efficient HVAC* Article 1482 TU Delft OPEN Publishing. <https://doi.org/10.34641/clima.2022.420>

### Important note

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

### Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

### Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# Improving the Calibration on Building Stock Level method by Comparing objective functions and optimization algorithms

Paula van den Brom <sup>a</sup>, Samuel Smets <sup>b</sup>, Laure Itard <sup>c</sup>

<sup>a</sup> Faculty of Architecture and the Built Environment, Delft University of Technology, Delft, the Netherlands, P.I.vandenBrom@tudelft.nl

<sup>b</sup> Faculty of Architecture and the Built Environment, Delft University of Technology, Delft, the Netherlands, L.C.M.Itard@tudelft.nl

<sup>c</sup> Faculty of Architecture and the Built Environment, Delft University of Technology, Delft, the Netherlands, S.J.Smets@tudelft.nl

**Abstract.** Many researchers have indicated the energy performance gap (difference between actual and predicted energy used in buildings), not only on an individual building level, but also on a building stock level. For policy makers it is important that predictions are correct on an building stock level to make them a useful tool to predict the effect of their proposed energy saving policies. Often not all input parameters for building energy simulations are known (e.g. insulation rates are often only possible to determine with destructive inspection or extensive measurements), therefore assumptions are made (e.g. assumptions for insulation rates are often made based on construction year). It is expected that a large part of the energy performance gap on building stock level are caused by incorrect assumptions of the unknown parameters in the building simulations. Previous research has shown that automated calibration of the assumptions on building stock level seems a promising method to reduce the energy performance gap and therewith make building energy simulations on building stock level a more reliable tool for policy makers. The previous research about calibration on building stock level was a proof of concept and still needs some improvements before it can be applied in practice. One of the aspects to improve the method is to determine the most suitable objective function and the most suitable optimization algorithm. In this paper we compare different objective functions (e.g. Root Mean Square Error, Mean Absolute Error, Sum of Absolute Errors). Next to that we compare different optimization algorithms (e.g. Genetic Algorithm, Particle Swarm and simulated Annealing Algorithm). For the comparison of the objective functions and the algorithms the former Dutch calculation method to determine the energy label in dwellings is used, in combination with the SHAERE database and data from the Dutch Statistics. The SHAERE database contains all input information on individual dwelling level to calculate the energy label of a dwelling of almost 2 million dwellings. The Dutch Statistics database contains the individual annual energy use of all dwelling of the Netherlands and can be linked to the SHAERE database.

**Keywords.** Energy Performance Gap, Calibration on building stock level, Optimization algorithms, measured data, Energy performance  
**DOI:** <https://doi.org/10.34641/clima.2022.420>

## 1. Introduction

Many researchers have indicated the energy performance gap (difference between actual and predicted energy used in buildings), not only on an individual building level, but also on a building stock level. For policy makers it is important that predictions are accurate on an building stock level to make them a useful tool to predict the effect of their proposed energy saving policies. Often not all input parameters for building energy simulations are known (e.g. insulation rates are often only possible

to determine with destructive inspection or extensive measurements), therefore assumptions are made (e.g. assumptions for insulation rates are often made based on construction year). It is expected that a large portion of the energy performance gap on building stock level are caused by incorrect assumptions of the unknown parameters in the building simulations. Previous research has shown that automated calibration of the assumptions on building stock level seems a promising method to reduce the energy performance gap and therewith make building energy simulations

on building stock level a more reliable tool for policy makers.

The previous research about calibration on building stock level was a proof of concept and still requires some improvements before it can be applied in practice. One of the aspects to improve the method is to determine the most suitable objective function and the most suitable optimization algorithm. In this paper we compare different objective functions (e.g. Root Mean Square Error, Mean Absolute Error, Sum of Absolute Errors). Next to that we compare different optimization algorithms (e.g. Genetic Algorithm, Particle Swarm and simulated Annealing Algorithm). For the comparison of the objective functions and the algorithms the former Dutch calculation method to determine the energy label in dwellings is used, in combination with the SHAERE database and data from the Dutch Statistics. The SHAERE database contains all input information on individual dwelling level to calculate the energy label of a dwelling of almost 2 million dwellings. The Dutch Statistics database contains the individual annual energy use of all dwelling of the Netherlands and can be linked to the SHAERE database.

This paper will first introduce the Energy performance gap (section 2). After that the calibration on building stock level will be explained further (section 3). In section 4 the research methods are explained, followed by the results in section 5. Finally the results are discussed in section 6 and conclusion are drawn in section 7.

## 2. The Energy Performance Gap

### 2.1 The Energy performance gap

Building energy simulation models are widely used to estimate the energy demand of a building. In Europe the EPBD (Energy Performance of Buildings Directive) demands all European countries to have a system that informs potential buyers and tenants of buildings about the energy performance of buildings. This is often done by an Energy Performance Certificate (or in for example the Netherlands also called Energy Label). The method used to calculate an Energy Performance Certificate is different per country but has to fulfil at least some minimal requirements as set by the European Union. Also important to mention is that the EPBD calculations use a standardized building use to make the results comparable. This makes that these calculation results are often not a realistic reflection of actual energy use in buildings, because all buildings are used differently. This means that there is a gap between actual and calculated energy use. However the gap is not only there on individual building level, but also on building stock level. This implies that there is a more structural problem. The gap between actual and calculated energy use in a building is called the Energy Performance Gap (EPG) and occurs on individual but also on building stock level.

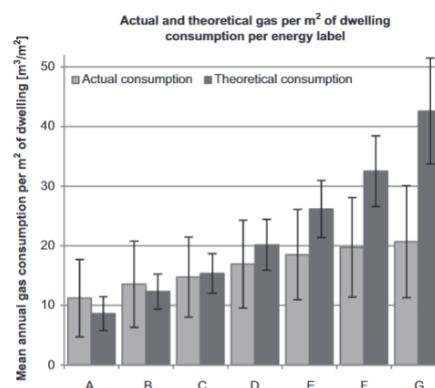


Fig. 0 – The energy performance gap source:[1]

### 2.2 National and international context

One of the first studies presented on the EPG on building stock level was presented based on Dutch data. Due to the strong increase on data availability also many other countries have found an EPG on building stock level e.g. Portugal [2], Switzerland [3, 4], Denmark [5], Ireland [6], United Kingdom [7], France [8], Germany [7]. This indicates that the EPG on building stock level is not just caused by the method used in the Netherlands, but there is a more structural cause of this gap. This is an issue because the Energy Performance Certificates (or Energy Labels) are increasingly used to develop Energy saving policies and to determine subsidies for Building Energy Renovation measures.

### 2.3 Causes of the Energy Performance Gap

The past years many researchers have studied the EPG and potential explanation of the EPG. Some of the explanations are:

1. Balance detail level simulation model and reliability of input parameters: If a simulation model is very detailed also many input parameters are required. The more input parameters are required the higher the probability that mistakes are made or assumptions have to be made because the input parameters are not known.
2. Occupant behaviour: on an individual level the occupant behaviour plays an important role. Every occupant uses her/his house differently. However we see also a gap on building stock level which might indicate that the assumptions we make for occupant behaviour (e.g. indoor temperature, ventilation rates) are not a realistic reflection of reality.
3. Mistakes in construction and/or installation process. In building energy simulation models we often assume that no construction or installation failures are present. However if something is not constructed properly or installed properly this could cause a difference between simulation and reality.
4. Difference in construction drawings and execution. Often construction drawings are used to make building energy simulation. If there is a difference

between the drawing and the real building this could cause a gap

5. Measurements are often considered as the solution for the energy performance gap. However if equipment is not calibrated or wrongly placed this could also provide wrong input.

### 3. Calibration on building stock level

Based on the previously presented findings an automatic calibration procedure on building stock level was developed aiming to reduce the EPG on building stock level. In this section the principle of the method is explained briefly. For a more detailed description of the method we refer to [9].

In section 2 we mentioned that many assumptions are made in Energy performance calculations which could be a cause of the EPG on a building stock level. The automated calibration method on building stock level aims to optimize the assumptions in such a way that the EPG will be reduced. The method makes use of existing data to optimize the assumptions. It is therefore only possible to execute this method if a representative dataset of buildings with actual energy use data is available including the input data for the energy calculation. Of which a significant part of the input should be measured and not be based on assumptions (we assume 30%, but more research should be done to determine the minimum required).

The method is based on a traditional calibration method. In a traditional calibration method the energy demand would be calculated and the unknown (assumed parameters) would be changed every optimization try until the optimal result is found (smallest gap between theoretical and actual energy use), see figure 1. Unfortunately this method won't be sufficient because the results will only be applicable for that individual building on which the occupant plays a very significant role. However, we have seen in the previous section that occupant behaviour should be averaged out if we look at the energy demand on a building stock level. Therefore the method suggest to optimize multiple buildings at the same time to find the smallest EPG on a building stock level. Since part of the buildings don't have assumed values but 'real' values the optimization could reduce the risk of assumed values 'compensating' for each other and therefore increase the probability of finding the 'real' assumed value. The results in the 'proof of concept study' have shown that this method has the potential to reduce the EPG significantly [9].

The parameters that are optimized are (see also appendix):

1. Rc-values (per construction period)
2. Air change rate ventilation
3. Indoor temperature settings
4. Domestic hot water consumption
5. Efficiency of heating system

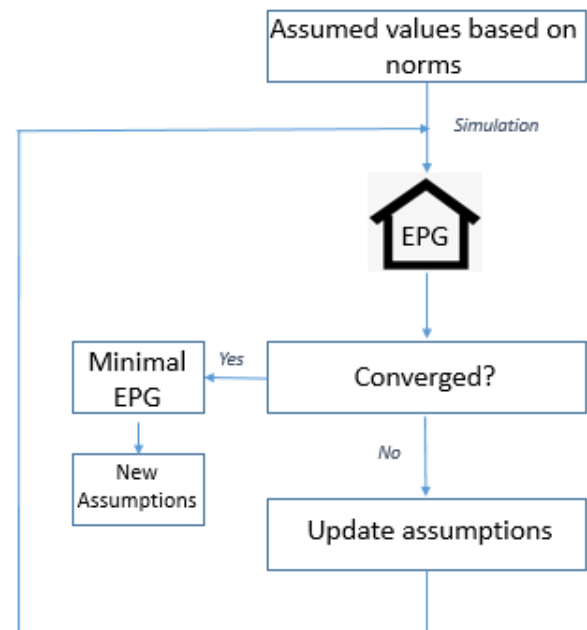


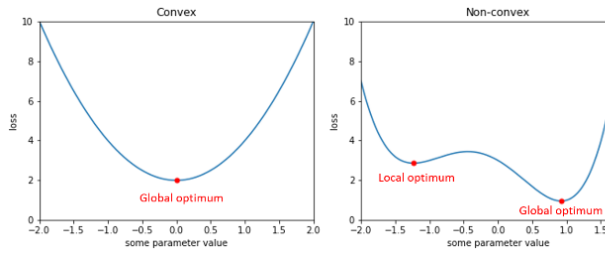
Fig. 1 – Calibration procedure

### 4. Method

This paper is based on the proof of concept method explained in section 3 and [9]. As written before the method is based on optimization. One of the remaining questions of the proof of concept paper was to determine which optimization method is the most suitable and which objective function is most suitable for this purpose. The three optimization algorithms that are applied are: Genetic Algorithm, Particle swarm optimization and Simulated Annealing. In this paper we compare also compare different objective functions (e.g. Root Mean Square Error, Mean Absolute Error, Sum of Absolute Errors). In this section we describe the different optimization algorithms, different objective functions and the data we use for the case study.

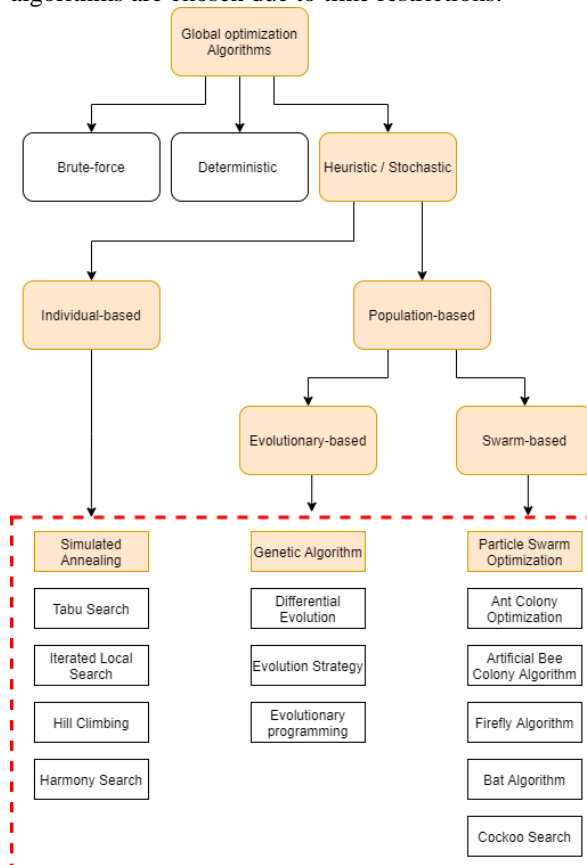
#### 4.1 Optimization algorithms

There are many optimization algorithms, and therefore it is important to explore the optimization algorithms out there and make a selection of suitable algorithms for the problem at hand. The most reliable optimization algorithm would be to simulate all possible values for all parameters (brute-force method) [10]. However this method is practical often not feasible due to the high computational power that is required. Because the optimization problem is a non-linear and non-convex problem a global optimization algorithm is required (fig 2) and special attention should be paid to the barrier settings on the optimization. Figure 3 shows an overview of several optimization algorithms.



**Fig. 2** - Examples of a convex and a non-convex problem. A non-convex problem has multiple sub-optima, of which the one with the lowest objective function loss (which is the objective function evaluation) is called the global optimum and the others local optima. (source: Thesis Samuel Smets)

The three algorithms indicated in orange (Simulated Annealing, Genetic Algorithm and Particle Swarm Optimization) are used for this paper. Only three algorithms are chosen due to time restrictions.



**Fig. 3** - Overview of the most well-known global optimization algorithms. The ones indicated in orange are used in this project and further explained in the literature review. (source: Thesis Samuel Smets)

Figure 3 shows that the three chosen optimization algorithms are all stochastic optimization algorithms. Stochastic optimization algorithms have the advantage of being easy implementable and have better potential for complex problems compared to deterministic algorithms, they are also relatively fast in terms of function evaluation [11]. The first algorithm which will be applied is Simulated annealing. This is an individual-based algorithm. The

advantage of individual-based algorithms is the lower number of computations per iteration, and therefore it usually has lower computation time until convergence. The disadvantage is that it is more prone to end up in local minima compared to population-based algorithms. Next to that, population-based algorithms usually explore the search space better, because multiple solution vectors are searching through the parameter space simultaneously [12].

Simulated Annealing (SA) is a global optimization algorithm that is inspired by the principles of the annealing process of metals. The annealing procedure of metals is defined as cooling down the matter slowly after being heated up to high temperatures, to get the optimal molecular arrangements of the metal particles, in which the energy of the system is minimized [13].

The second and the third optimization algorithms that will be applied are population-based algorithms. The principle of population-based optimization algorithms is that an initial set of parameter vectors are optimized every iteration until the global optimum is found. Population-based algorithms can be divided into evolutionary-based algorithms (also Evolutionary Algorithms) and swarm-based algorithms. The evolutionary-based algorithms use Darwinian evolution concepts and the swarm-based algorithms make use of specific movement patterns by the parameter solution vectors [12]. The genetic algorithm (evolutionary based) is the second algorithm that will be applied in this research. The principles of the GA are based on the natural selection processes of life in which new solution vectors generations are created from the previous solution vector generations [14]. Like in natural selection, the ‘genetic material’ of solution vectors that are the fittest will survive. ‘Genetic material’ refers in this case to the solution values of the vectors. The last algorithm that will be applied is the Particle swarm optimization, which is a swarm-based algorithm. Swarm-based optimization algorithms are based on swarm intelligence, which is the collective behaviour of an organized group of animals or insects [15]. Swarm-based algorithms are increasing in popularity, while the algorithms are flexible, versatile, adaptable to external variations and they have self-learning capabilities [15].

## 4.2 Objective functions

In optimization the optimal solution can be found by minimizing or maximizing the objective function. It can be seen as a score that evaluates the goodness of fit. There are different objective function. There is no ‘best’ objective functions, since it is dependent on the optimization problem [16].

The main choice one has to make choosing an objective function is whether to use squared errors, absolute errors or normal errors.



Square errors penalize strongly for outliers. If one wants a model that does not predict values too far from reality, this is a good objective function to use. If one wants to treat all samples the same in terms of distance of the output to real data, and one is not concerned about outliers, then the absolute error will be a better choice. If one does not care about the errors of single data samples, but one is only interested in the summed error of the complete dataset (in which positive and negative errors can cancel each other out), then the normal error will be the better choice.

The objective functions that will be analysed in this paper are:

Root mean squared error (RMSE)

$$RMSE = \sqrt{\frac{\sum (S-O)^2}{n}}$$

Mean absolute error (MAE)

$$MAE = \frac{1}{n} \sum |S - O|$$

Percent bias (PBIAS)

$$PBIAS = \frac{\sum (O - S)}{\sum O} 100(\%)$$

In these equations  $O$  stands for observations and  $\bar{O}$  stands for the mean of all observations.

The  $S$  stands for simulations and  $\bar{S}$  stands for the mean of all simulations. The  $n$  stands for number of samples.

### 4.3. Data

For this research we make use of the SHAERE database. This is a database from the Umbrella organisation of Dutch social housing companies (AEDES). They use this database to track the energy performance of their housing stock. In the database all input parameters to make the Energy performance calculation (based on ISO 82.3) are saved for all individual dwellings. This database is linked to the Dutch statistics database which has measured energy use data of almost every individual dwelling in the Netherlands. These data will be used to test the algorithms and the objective functions. The most important assumptions made in ISO 82.3 are presented in the appendix.

## 5. Results

In this section the results are presented. The experiments were run on a system with the following processor specifications: Intel(R) Xeon(R) Gold 6146 CPU, @ 3.20GHz, 4 Cores.

### 5.1 Objective function

The three different objective functions that are compared are: RMSE, MAE and PBIAS. All three calibrations were run over 100 generations, with a population size of 100. The loss shown in the minimal objective function loss found by the population so far. The minima of all objective functions were found

within the 100 generations. The calibration with the PBIAS objective function seemed to have reached the minima earlier than the RMSE and the MAE (~20 generations with respect to ~40 generations). The minimal objective function losses are not comparable between the different objective functions, so nothing can be said about that. The computation time is comparable for all of them, which is expected because the calibrations were run on the same dataset and the same algorithm with the same hyperparameters. Although PBIAS had its final drop in objective function loss at generation 98, the plot reveals that it was already close to this minimal objective function loss at around generation 20. Therefore the computation time at minimal objective function loss is also not very informative to look at. For the other model aspects these numbers will be important for determining which is the preferred option, but for selecting the objective function it is not of major importance for making a decision.

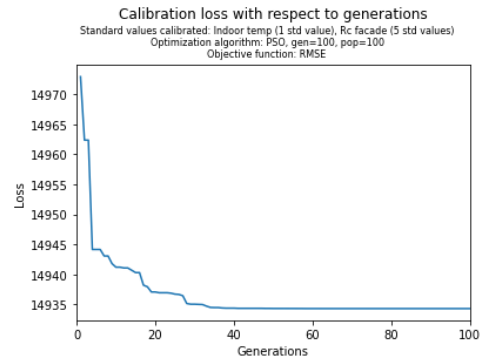


Fig. 4 - RMSE. (source: Thesis Samuel Smets)

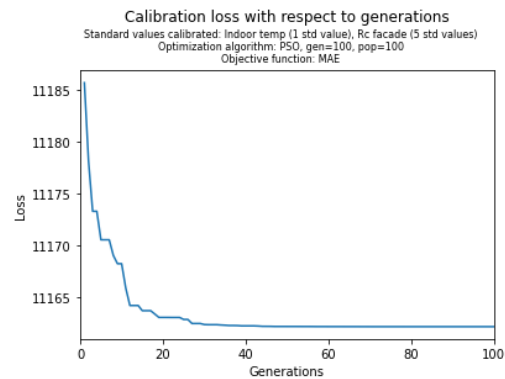
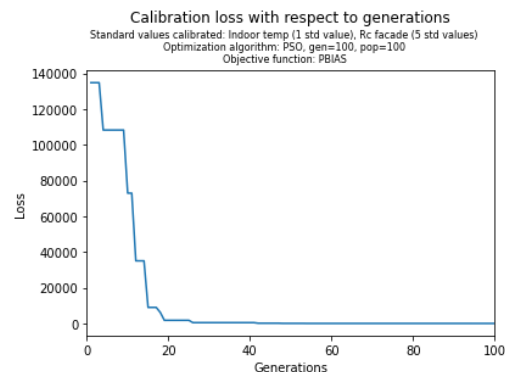
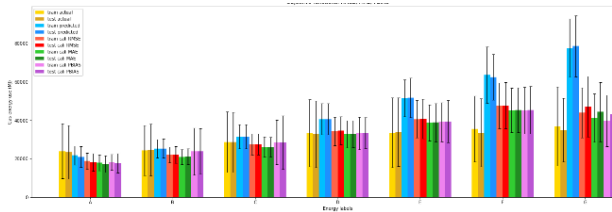


Fig. 5 - MAE. (source: Thesis Samuel Smets)



**Fig. 6 - PBIAS.** (source: Thesis Samuel Smets)

Table 1 shows the optimized standard values for the three different objective functions. However if we look at the results we see that in all three cases the EPG is reduced significantly (see fig 7)



**Fig. 7 - The average gas energy use (with standard deviation bars) for dwellings per energy label for both the training and the test dataset. In yellow one finds the actual energy use, in blue one finds the predicted energy use with the uncalibrated model, in red, green and purple one finds the predicted energy use with the calibrated model (with objective functions RMSE, MAE and PBIAS respectively).**

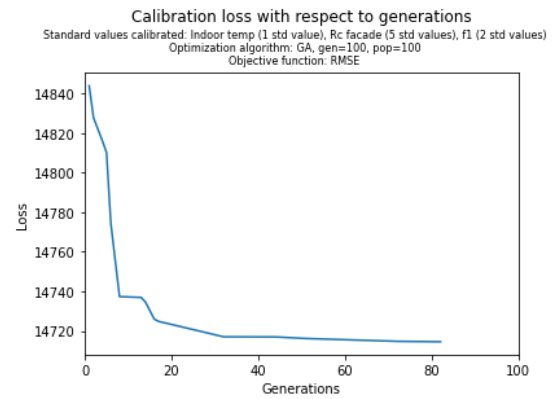
If we look at the optimized variables (Table 1) we see that the ones calculated with the PBIAS are the least realistic (e.g. the lowest Rc value is 0, which is not possible). Therefore it seems that this objective function is probably not suitable for this purpose. The differences between RMSE and MAE are only minimal and therefore we cannot conclude whether the RMSE or MAE is better.

**Tab. 1 - Calibration results of the objective function analysis. Minimal objective function loss reached during optimization over 100 generations with a population size of 100; Generation count at which the minimal objective function loss was reached; Computation time at which minimal objective function loss was reached..**

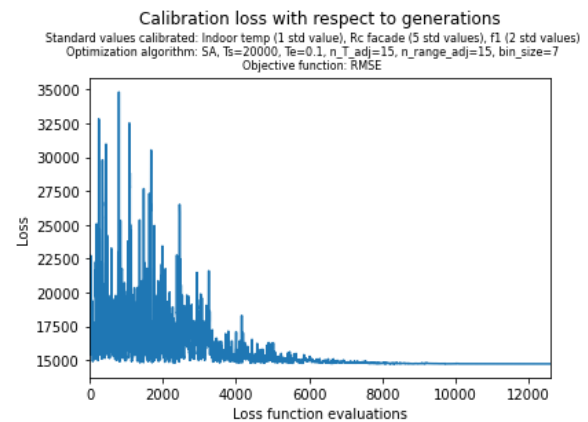
	Minimal subjective function loss	Computation time needed (hours)	Generations needed
RMSE	14934.32	9.44	58
MAE	11162.15	9.82	60
PBIAS	0.41	16.06	98

## 5.2 Optimization Algorithms

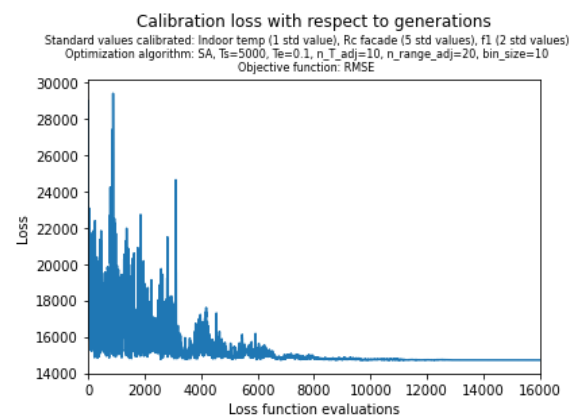
The figures below show the different optimization plots per optimization algorithm (using the RMSE objective function). The figures clearly show that the SA is an individual based algorithm and the other are population based algorithm. The individual based algorithm evaluates the function loss every single function evaluation.



**Fig. 8 - Genetic Algorithm with hyperparameter settings: mutation rate = 0.5, population size = 100, generations = 100**



**Fig. 9 - Simulated Annealing with hyperparameter settings: start temperature = 20000, end temperature = 0.1, temperature adjustments = 15, range adjustments = 15, number of cycles over all standard value solution dimensions per range = 7**



**Fig. 10 - Simulated Annealing with hyperparameter settings: start temperature = 5000, end temperature = 0.1, temperature adjustments = 10, range adjustments = 20, number of cycles over all standard value solution dimensions per range = 10.**

Table 2 shows the calibration results in numbers. The minimal objective function losses are similar to all

algorithms, as are the optimized values. One can also see that the Simulated Annealing algorithm has a longer computation time than the other two algorithms.

**Tab. 2** - Calibration results of the optimization algorithm analysis. Minimal objective function loss reached during optimization; Generation count (PSO and GA) and iterations (SA) at which the minimal objective function loss was reached; Computation time at which minimal objective function loss was reached.

	Minimal subjective function loss	Computation time needed (hours)	Generations needed
PSO	14713.10	13.47	84
GA	14714.49	13.47	82
SA	0.14713.39	19.38	12067

If we look at the validation function we see that there is no significant difference between the training and the test datasets and therefore there was no overfitting for any of the algorithms.

It becomes clear that all of the algorithms perform equally. The only difference is de computation time, which is longer for SA. The reason for this might be that finding the optimal hyperparameters of SA is known to be difficult. Therefore PSO and GA are the most preferred ones.

## 6. Conclusion and discussion

This paper is a follow up on the calibration on building stock level method from which the proof of concept is published in: [9]. It aimed to answer some of the remaining questions which are important to transform the proof of concept into a practical applicable method.

Therefore we investigated in this paper which objective function and optimization algorithms are the most suited for calibration on building stock level. Based on the analysis we can conclude that from the three objective functions (RMSE, MAE and PBIAS) which we analysed the RMSE is the most preferred one in terms of performance.

If we look at the three optimization algorithms (PSO, GA and SA) no clear conclusion could be made. However we saw that the SA method required significant longer computation time.

The examples again showed that it is possible to reduce the energy performance gap using automated calibration on building stock level and the two findings from this paper bring the applicability of the method again a step closer. However there are still several aspects that have to be investigated before the method will be applicable in practice. This paper also shows the importance of solving the EPG and it

showed that the EPG is an international problem.

## 7. Acknowledgement

This paper is based on the graduation thesis of Samuel Smets for the Master Industrial Ecology at Delft University of Technology.

## 8. Data access statement

The datasets used for this study are not public because they are owned by the umbrella organisation of the social housing companies in the Netherlands (Aedes) and are only allowed to be used with approval of Aedes.

## 9. References

- [1] Majcen, D., L.C.M. Itard, and H. Visscher, *Theoretical vs. actual energy consumption of labelled dwellings in the Netherlands: Discrepancies and policy implications*. Energy Policy, 2013. **54**: p. 125-136.
- [2] Palma, P., J.P. Gouveia, and S.G. Simoes, *Mapping the energy performance gap of dwelling stock at high-resolution scale: Implications for thermal comfort in Portuguese households*. Energy and Buildings, 2019. **190**: p. 246-261.
- [3] Cozza, S., J. Chambers, and M.K. Patel, *Measuring the thermal energy performance gap of labelled residential buildings in Switzerland*. Energy Policy, 2020. **137**: p. 111085.
- [4] Cozza, S., J. Chambers, C. Deb, J.-L. Scartezzini, A. Schlüter, and M.K. Patel, *Do energy performance certificates allow reliable predictions of actual energy consumption and savings? Learning from the Swiss national database*. Energy and Buildings, 2020. **224**: p. 110235.
- [5] Gram-Hanssen, K., S. Georg, E.T. Christiansen, and P.K. Heiselberg, *How building regulations ignore the use of buildings, what that means for energy consumption and what to do about it*, in *ECEEE 2017 Summer Study on energy efficiency*. 2017, European Council for an Energy Efficient Economy, ECEEE: Presqu'île de Giens.
- [6] Coyne, B. and E. Denny, *Mind the Energy Performance Gap: testing the accuracy of building Energy Performance Certificates in Ireland*. Energy Efficiency, 2021. **14**(6): p. 57.
- [7] Laurent, M.H., B. Allibe, T. Oreszcyn, I. Hamilton, R. Galvin, and C. Tigchelaar, *Back to reality: How domestic energy efficiency policies in four European countries can be improved by using empirical data instead of normative calculation.*, in *eceee Summer Study*



- proceedings, E.C.f.a.E.E.E. (ECEEE), Editor. 2013: Stockholm, Sweden.
- [8] Cayre, E., B. Allibe, M.H. Laurent, and D. Osso, *There are people in the house! How the results of purely technical analysis of residential energy consumption are misleading for energy policies*, in *ECEEE 2011 summer studies - Energy efficiency first: the foundation of a low-carbon society*, ECEEE, Editor. 2011.
- [9] Van den Brom, P., *Energy In Dwellings: A Comparison between Theory and Practice*, in *A+BE | Architecture and the Built Environment*. 2020, Delft University of Technology: Delft.
- [10] Lara, R.A., E. Naboni, G. Pernigotto, F. Cappelletti, Y. Zhang, F. Barzon, A. Gasparella, and P. Romagnoni, *Optimization Tools for Building Energy Model Calibration*. Energy Procedia, 2017. **111**: p. 1060-1069.
- [11] Lin, M.-H., J.-F. Tsai, and C.-S. Yu, *A Review of Deterministic Optimization Methods in Engineering and Management*. Mathematical Problems in Engineering, 2012. **2012**: p. 756023.
- [12] Mirjalili, S., *Genetic Algorithm*, in *Evolutionary Algorithms and Neural Networks: Theory and Applications*. 2019, Springer International Publishing: Cham. p. 43-55.
- [13] Eren, Y., İ.B. Küçükdemiral, and İ. Üstoglu, *Chapter 2 - Introduction to Optimization*, in *Optimization in Renewable Energy Systems*, O. Erdiñç, Editor. 2017, Butterworth-Heinemann: Boston. p. 27-74.
- [14] Arsenault, R., A. Poulin, P. Côté, and F. Brissette, *Comparison of Stochastic Optimization Algorithms in Hydrological Model Calibration*. Journal of Hydrologic Engineering, 2014. **19**(7): p. 1374-1384.
- [15] Chakraborty, A. and A.K. Kar, *Swarm Intelligence: A Review of Algorithms*, in *Nature-Inspired Computing and Optimization: Theory and Applications*, S. Patnaik, X.-S. Yang, and K. Nakamatsu, Editors. 2017, Springer International Publishing: Cham. p. 475-494.
- [16] Chai, T. and R.R. Draxler, *Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature*. Geosci. Model Dev., 2014. **7**(3): p. 1247-1250.

## 10. Appendix

Category	Assumptions
Façade insulation (Rc [m <sup>2</sup> K/W])	If the insulation is unknown and cannot be measured the assumed insulation level is based on construction year. ISSO 82.3 assumes the following values. Built before 1965 = 0.19 Built between 1965-1975 = 0.43 Built between 1975-1988 = 1.3 Built between 1998-1992 = 2 Built after 1992=2.3
Floor insulation (Rc [m <sup>2</sup> K/W])	If the insulation is unknown and cannot be measured the assumed insulation level is based on construction year. ISSO 82.3 assumes the following values.
Roof insulation (Rc [m <sup>2</sup> K/W])	If the insulation is unknown and cannot be measured the assumed insulation level is based on construction year. ISSO 82.3 assumes the following values.
Ventilation rate	Assumed ventilation rate is based on type of ventilation system (natural ventilation, mechanical exhaust ventilation, demand based mechanical exhaust ventilation, balanced ventilation with heat recovery) and minimum ventilation rate per m <sup>2</sup> floor are. Natural ventilation q=0.47; mechanical exhaust ventilation q=0.47; demand based ventilation q=0.29; balanced ventilation=0.47. If a heat recovery system is present q is multiplied by 1-efficiency of heat recovery system.
Infiltration rate	Assumed infiltration rate is based on floor area and type of building (detached, semidetached, terraced house, common staircase and galleries, common staircase no galleries and maisonettes). f <sub>2</sub> = air permeable factor based on ventilation system (0.12 for demand based else 0.13); The exact values of q <sub>inf,10</sub> can be found in table 10 of ISSO 82.3 (2011)
Indoor temperature	Assumed average constant indoor temperature of 18 °C
Domestic hot water consumption	Assumed amount for domestic hot water is based on number of occupants, which is based on floor area. Further it takes into account if a shower, bath and/or dishwasher is/are present and if water saving shower heads are installed.
Efficiency of heating system	The assumed efficiency of the heating system is based on the type of system, but also if the system is placed outside or within the thermal envelope of the building. Exact values can be found in table 19 ISSO 82.3 (2011)
Efficiency of domestic hot water system	The assumed efficiency of the domestic hot water system is based on the type of system, but also if the system is placed outside or within the thermal envelope of the build. The exact values can be found in table 24 of ISSO 82.3 (2011)