Computational time reduction using detailed building models with Typical Short Sequences

Hasan Sayegh, Antoine Leconte, Gilles Fraisse, Etienne Wurtz, Simon Rouchier

Version of Record: https://www.sciencedirect.com/science/article/pii/S0360544222000123
Manuscript_b291ab8407df2b8473616b0f233ea61c

Computational time reduction using detailed building models with Typical Short Sequences

Hasan Sayegh [a,b*], Antoine Leconte [a], Gilles Fraisse [b], Etienne Wurtz [a], Simon Rouchier [b]
*sayeghhasan@gmail.com
a Laboratoire d'Innovation pour les Technologies des Energies nouvelles et les Nanomatériaux (LITEN), CEA Tech, Chambéry, France
b Laboratoire Optimisation de la Conception et Ingénierie de l'Environnement (LOCIE), Chambéry, France

ABSTRACT

Comprehensive studies of detailed dynamic building models, which take into consideration both the envelope and the connected systems, yield more precise results compared with simplified ones, but at considerable computational expense. Aside from classical approaches that work on the model itself to accelerate the simulation process such as model reduction or metamodels, this paper focuses on the concept of applying reduced simulation sequences directly to detailed models to calculate annual results. The objective is to quickly and precisely reproduce the integrated annual profiles of predefined criteria of a computationally expensive reference model. After presenting and analyzing methods used in the literature to reduce weather data, we categorize the methods based on the type of data used and the nature of the process for selecting the typical days. Analysis of these methods led to the development of a new iterative approach with an embedded grouping algorithm. The method creates and iteratively enhances a short simulation sequence of typical days based on data reflecting the integrated annual profiles calculated using the detailed model. The reduced sequence led to much faster simulations while achieving profiles highly correlated with the reference integrated annual profiles. In addition, the last annual value, i.e., final annual sum, of each criterion extrapolated from a typical 12-day simulation differs little from the reference values (errors less than 1%). Moreover, the method was compared to two other clustering methods based on different types of selection data and an iterative method used in the literature. The results show that the classical method of day selection based only on weather data, typically used to generate Short Reference Years (SRYs), is in fact unable to accurately reproduce the annual reference profiles. Finally, the approach was also efficient when generalized, demonstrating its applicability to future optimization studies.

Keywords: Buildings, energy systems, short sequence, computation time reduction, optimization.

1 Introduction

Since the beginning of the 20th century, building energy demand has risen to the levels of transport and industry due to economic growth and increasing population. It currently accounts for about 40% of total energy consumption [1].

However, the building sector offers significant potential for improved energy efficiency with high-performing envelopes and energy-efficient systems. Yet, comprehensive studies which consider both the envelope and the systems result in complex models, especially when enlarging the scale to districts and heat networks, leading to high computational expense. Moreover, other approaches such as statistical analysis, life cycle assessment and optimization studies require repetitive simulations, taking even more time to perform. Simplifying complex models, such as by using analogical RC models [2]–[4] raises concerns about the accuracy of results due to the absence of a detailed reference model. On the other hand, more advanced approaches such as reduced order models [5]–[7] or metamodels (e.g., neural networks, kriging models) consider reference models and are therefore more precise but are sometimes difficult to apply, due to nonlinearity and control-command issues in the former or limited input parameters for the latter. Moreover, these approaches, which work on the model itself, are also less generalizable as each case study requires its own simplified twin model and therefore needs to be reapplied if the model undergoes modifications. There are several examples from the literature that work on simplifying the complex models such as the work of Eisenhower et al. [8] who used metamodels to study the thermal comfort and consumption of a two story building. Another is that of

Mancarella [9] who used spatial aggregation to reduce the number of nodes in an energy system network study, while Milan et al. [10] reduced nonlinearities and discontinuities to avoid non-convexity of the program.

On the other hand, there is another approach based on the reduction of input data profiles rather than the model itself. The objective is to quickly and precisely reproduce the annual dynamic of a complex model by executing it with reduced data files. This trend started in the 1970s with the development of methodologies for generating Typical Meteorological Years (TMYs), the term used in the USA, or Test Reference Years (TRYs), the term mainly used in Europe. Subsequently, computer time limitations led to the development of methodologies referred to as Short Reference Years (SRYs) to quickly reproduce the annual dynamic of a model.

The concept behind this approach is that the thermal performance of buildings and of other solar thermal and photovoltaic systems, depend on several meteorological parameters like diffuse horizontal solar radiation, relative humidity and wind velocity. A TRY provides a standard for meteorological parameters (hourly data) for a period of one year, therefore representing typical climatic conditions over a long time period by a single year. Most of the methodologies reported in the literature promote the idea of using sequences of real, measured data to compose a TRY, for example the Sandia National Laboratories method [11]–[17], the TMY2 method [18], the Danish method [19], [20] and the Festa-Ratto method [21]. Others such as the Design Reference Year method [22] employ adjusted data to give a monthly cumulative distribution similar to the cumulative distribution of the corresponding months in the original multi-year data set.

In contrast, SRYs include meteorological data covering only typical days of the year [20], [23], [24]. The approach evaluates monthly, seasonal, or annual performance starting from a short simulation sequence of typical days. Therefore, instead of simplifying the models and using TRYs, short sequences are used to reduce the computational expenses of a detailed fully dynamic simulation.

This work focuses on the concept of model simulation using short simulation sequences due to their applicability to dynamic simulations in different fields and scales. The main approaches for the selection of typical days are introduced, classified based on the type of data or process used in day selection and analyzed. We present a new approach which is then tested on a solar combisystem model and compared to other approaches used in the field of building performance analysis. Finally, a generalization study is implemented to evaluate the flexibility of the approach and its ability to be extended to other fields such as optimization which requires a stable level of accuracy despite the parametric modifications the model undergoes through the study.

2 Reducing the simulation period using short simulation sequences

2.1 Approaches to typical day selection

Several approaches are used in the literature for day selection based on weather data only, the outputs of a reference simulated model or both.

Reduced sequences determined from weather data only is the classical approach used with SRYs and simply calculates averages at each hour based on weather data. Although good results might be obtained with such a reference year in predicting building energy demands, the selection procedure is considered to be rather arbitrary [25]. This is true even with more advanced approaches using only weather data and is discussed in section 4.2.5. An example of employing only weather data to predict system performance is found in Petri et al. [26]. They predicted the monthly average useful energy gain of a collector and the delivered energy load of a solar energy system by replacing a specific month with a 5- to 10-day sequence that best represents the average monthly dry bulb temperature, wind speed and solar radiation. The daily solar radiation values are then broken down into hourly values using the method developed by Liu and Jordan [27]. A linear interpolation is performed on the maximum and minimum temperatures between their time of occurrence, and the average wind velocities are output at a constant velocity over a twenty-four hour period. After simulating the hourly recorded set of days, the reduced values are extrapolated to compare them with the reference months by applying a calendar adjustment factor (CAF). If a 6-day sequence is used to represent a 30-day

month, the CAF is 30/6 =5. The number is then multiplied by the simulation output to obtain the system performance for the month.

On the other hand, a reduced sequence determined from the outputs of the model simulation is adapted to the building. This works better than previous methods based only on fixed boundary conditions because it also works for several configurations of a building. This trend has been observed in the literature with the most recent papers using the second type of selection data (model outputs) while earlier studies were based on weather data only.

In addition to selection data, the approaches can be grouped in several categories based on the adopted selection process. The main approaches are stochastic, heuristic, iterative, and grouping algorithms.

Heuristic Approaches Heuristic methods are practical methods that directly select a set of typical days and are greatly influenced by the personal expertise or experience of the developer. The selection is quick but not guaranteed to be optimal. Belderbos and Delarue [28] and Haller et al. [29] used a simple heuristic approach to select a number of periods with different load and/or meteorological conditions to capture a variety of different events. Fripp [30] worked with investment periods optimized based on 12 days of sampled data: two for each even-numbered month. Hart and Jacobson [31] reduced energy generation data with variable renewables by selecting eight specific days containing hours with extreme meteorological/load events and 20 random days to characterize typical system behavior.

Stochastic Approaches To generate hourly sequences of weather data several statistical methods are presented in the literature. Autoregressive moving average (ARMA) and Markov probability transition matrix (MPTM) modeling are widely used black box methods. They are used for reducing specific weather data (most commonly solar radiation) or predicting the energy performance of building models. The ARMA model consists of a linear transfer function where a random distribution with a normal probability distribution function is applied. The method was used by Brinkworth [32] and Mustacchi et al. [33] to generate reduced insolation sequences from annual ones. On the other hand, the MPTM method is based in its simplest form on a transition function describing the probability of the occurrence of two consecutive hours, such as in Poggi et al. [34]. Other methods which have been applied to building energy simulation include the selection of real sequences of weather data. For example Hall et al. [35] based the selection on the deviation of the interval averages from the monthly averages of climate variables. They predicted the monthly averages of heating and cooling loads using the average of four day intervals for each month, i.e., 12 sets of four days each. The choice of the best combination of intervals was based on an empirical score that minimizes four weighted variables: deviation of dry bulb temperature, cloudy weather, wind speed-dry bulb temperature product and moisture ratio. Several assumptions were made in the empirical equation the key being that heating and cooling demands for a month are approximately a linear function of the monthly averages of certain weather statistics.

Iterative Approaches The iterative approach searches for the best solution after repeating the same action several times and comparing the quality of results in each iteration. There are many examples in the literature that use this approach for day selection, either directly by implementing iterations or indirectly through algorithms based on repetitive iterations. Ortiga et al. [36] used a graphical method of iteration while studying the optimization of cogeneration and tri-generation models for buildings. The results showed very good coherence between original and predicted cumulative energy demand profiles. In addition, they concluded that for optimization, a larger time sequence does not lead to better results. The French Atomic and Alternative Energies Commission (CEA) has developed an iterative approach for testing and characterizing solar combisystems that reduces a whole year to twelve days[37]. The test is called the Short Cycle System Performance Test (SCSPT) and selects a short sequence using empirical equations based on weather data, monthly backup energy values, energy stored in the tank and internal room temperature.

Grouping Algorithms More advanced approaches to select a representative set of historical periods employ grouping algorithms. Days with similar attributes are grouped into clusters followed by day selection of each group. Most of the studies relying on the clustering approach use *k*-means clustering

as their favored approach. Fazlollahi et al.[38] [39] used *k*-means clustering to perform a multi-objective optimization of district energy systems. Dominguez et al. [40] used the *k*-means approach while optimizing a CHP system. Kotzur et al. [41] tested the efficiency of hierarchical clustering in estimating and optimizing the performance of a residential energy supply system and compared it to other partitional clustering approaches. Menegon et al. [42] developed a new dynamic test procedure for laboratory characterization of energy systems using a *k*-means clustering algorithm. In his PhD thesis on multi-criteria optimization methods for urban densification projects, Ribault [43] divided the year into 14 clusters by *k*-means after which he started reducing each cluster using an iterative algorithm that deletes repetitive days with an aggregate weight function approach. Since this method yielded 83 days, the execution time was reduced but remained lengthy, so he recommended additional improvements.

In addition to clustering algorithms, discriminant analysis was used by Blachandra and Chandru [44] to reclassify days of the year comparing the monthly average load curves of electricity demand to the daily ones. The first days of the month found to be misclassified were grouped with the previous group (month) while those of the last days were grouped with the next, and the stray middle days were ignored.

2.2 Analysis of existing approaches

A detailed study by Sayegh et al. [45] analyzed and compared the performance of heuristic, iterative and grouping methods. There was a noticeable preference for clustering algorithms over other approaches, with a special focus in the *k*-means approach due to its good performance. In terms of flexibility, i.e., the ability to select specific predefined days, the heuristic approach is preferable. Moreover, heuristics are the simplest way of reducing algorithms and no difficult coding is required. However, expecting precise results from a heuristic method requires considerable experience on the part of the operator in quickly choosing an efficient sequence. This is in contrast to stochastic, clustering, and iterative cases where it is left to the computer to perform all the trials and ultimately select the sequence, with greater precision, and more quickly.

On the other hand, typical days are selected directly from previously calculated monthly, seasonal, or annual averaged/integrated data in most cases and therefore the dynamic behavior of the system is not taken into consideration during the selection process. Thus, the reduced profiles do not really generate the dynamic of the reference performances but rather their sums or averages. In addition, days are selected based on the outputs of a specific case study with fixed parametric configuration and boundary conditions and therefore the identified sequence is not effectively able to predict the model's performance after parametric modifications. This calls into question the reliability of the referenced studies whether optimization, characterization or other. Therefore, this paper presents a new approach termed the Typical Short Sequence (TypSS) Algorithm which takes into consideration the dynamic behavior of the model during the day selection process and which is adapted to work simultaneously on several cases within the same model to cover the possible parametric modifications it might undergo in subsequent studies.

3 Typical Short Sequence (TypSS) algorithm

The approach entails an iterative aspect with an embedded *k*-medoids grouping algorithm. It uses defined selection criteria chosen by the user to select the typical days. The TypSS approach employs averaged and cumulative values of selection criteria chosen by the user to evaluate both periodic values and annual sums as a complete simulation. The aim is to regenerate the annual integrated profiles of the system and its annual sums which are values used in characterization or optimization studies. There is no direct condition for choosing the type of criteria. However, it depends on the case study to be performed as discussed in the previous section. The algorithm is based on two functions, which work iteratively and separately. The first function divides the year into periods of different sizes enabling greater focus on periods with more changes in performance. The second function enhances the sequence found in the first by searching for days which are more representative of each period. The TypSS algorithm is described in Figure 1.

3.1 Description of the algorithm

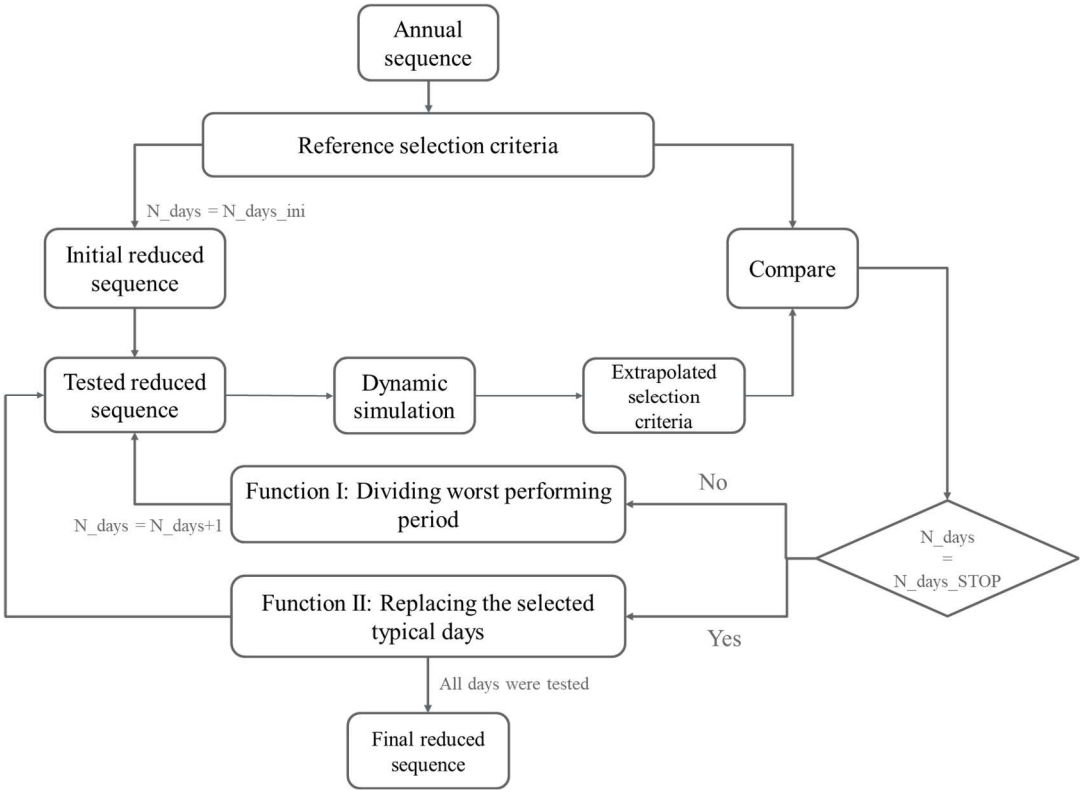


Figure 1. Scheme of the Typical Short Sequence (TypSS) Algorithm.

3.1.1 Preparation stage

The annual sequence is run once on the detailed model and the results are saved to serve as the reference selection criteria. These values are used within the algorithm itself. Running the detailed model fully is a very important step and gives the algorithm the ability, in every iteration, to decide whether the tested reduced sequence is suitable or needs to be modified in the following iteration. The length of the final reduced sequence, denoted *N_days_STOP*, is also specified manually before running the algorithm.

3.1.2 Initialization stage

The algorithm starts by generating an initial reduced sequence with a small number of typical days, dividing the year in equal parts. This initial number of typical days, denoted *N_days_ini*, can be changed. For example, if taken *N_days_ini* = 4, the first sequence roughly represents the four seasons as the starting point of the algorithm. It then searches for a representative day for each of the periods using *k*-medoids clustering. It uses the Euclidean distance between the value of a specific criterion recorded in a day and the value of the same criterion for all the other days in the same period and retains the day with minimal sum of Euclidean distances. An initial reduced sequence of *N_days_ini* days is thus generated.

3.1.3 Dynamic simulation and extrapolation of results

The detailed model is run on the reduced sequence. In order to initialize the values of the different components of the model (i.e. initial temperature of the storage tank, initial internal room temperature…), the simulation starts with the data from the last day of the sequence. The full reduced sequence is then considered and only the model performance of those days are retained for the following analysis. For example, if a sequence of *N_days* is considered, the simulation is applied on a sequence of *N_days+1* days where the last day is duplicated and the simulation starts with it. The values of this duplicated initial day are not considered for the following steps. The periodic values are

then predicted by extrapolating the criteria values of the selected days from the reduced simulation by the number of represented days in each period, i.e., the weight of the period.

3.1.4 Comparison with reference values

The extrapolated values are normalized and compared to their normalized values corresponding to the reference annual simulation using the same normalization parameters (min/max values) from the reference profiles. The aim of normalizing the values is to achieve global evaluation when considering several selection criteria of different scales. For example, if the selection criteria considered are internal room temperature and energy stored in the tank, the values of the two criteria are normalized by the algorithm to evaluate their influence simultaneously and equally.

3.1.5 Function 1: Dividing the worst performing period

After comparison with the reference values, the worst period showing the greatest difference in selection criteria prediction is detected. The algorithm divides this period into two equal parts and replaces the previously selected day of this period with two new ones representing the two new halves as previously explained in section 3.1.2, using *k*-medoids approach. Therefore, the initial sequence is now one day longer. The process is then repeated iteratively until finishing with a sequence of days, denoted *N_days*, equal to *N_days_STOP* as specified by the user.

3.1.6 Function 2: Replacing the selected typical days

Function 1 does not take into consideration global performance and influence of the periods on each other. For this, Function 2 is added which considers the entire performance of the sequence using two global values:

- The global coefficient of determination R^2_{Global} (eq. 1), the multiple of the coefficient of determination R^2_x of the selection criteria *x* (eq. 2), and
- The global annual sum error E_{Global} (eq. 3), the sum of the criteria annual sum errors E_x (eq. 4).

The annual sum error is directly linked to the main goal of the test sequence i.e., can the short sequence estimate the annual global performance of the system? However, this is not sufficient. So, the regression coefficient is used to express how well the short sequence results describe the reference results at each periodic time step for all criteria. Is the short sequence appropriate for each period?

$$R^2_{Global} = \prod_{x=1}^{n_{criteria}} R^2_x \quad (1)$$

Where:

R^2_x is the coefficient of determination of a selection criterion x

$n_{criteria}$ is the number of selection criteria

$$R^2_x = 1 - \frac{\sum_{i=1}^{365}(x_{TYPSS_i} - x_{YEAR_i})^2}{\sum_{i=1}^{365}(x_{YEAR_i} - \bar{x}_{YEAR})^2} \quad (2)$$

Where:

x_{TYPSS_i} is the typical daily value of criterion x obtained by the short sequence

x_{YEAR_i} is the daily value of criterion x obtained by the reference annual sequence

$\bar{x}_{YEAR}$ is the mean value of criterion x in the reference annual sequence

$$E_{Global} = \sum_{x=1}^{n_{criteria}} E_x \quad (3)$$

Where:

E_x is the annual sum error of a selection criterion x

$n_{criteria}$ is the number of selection criteria

$$E_x = \frac{|S_{TYPSS} - S_{YEAR}|}{S_{YEAR}} \times 100 \quad (4)$$

Where:
S_{TYPSS} is the annual sum of criterion x obtained by the short sequence
S_{YEAR} is the annual sum of criterion x obtained by the reference annual sequence

The function starts with the sequence of *N_days_STOP* days generated by Function 1. The algorithm replaces the typical days iteratively generating several *N_days_STOP* days sequences and calculating for each sequence the R^2_{Global} and the $E_{Global.}$ It is computationally expensive to test all possible combinations. So, starting from the first period of the sequence, a set of potential new typical days for period 1 while keeping the other days unmodified are tested. The set of potential days of the period are selected by dividing the period into clusters based on the used selection criteria and selecting the center of each cluster. The process is then repeated for the flowing periods (modifying the typical day of a single target period while keeping the others unmodified) until completing the entire sequence period by period. This method was adopted to reduce the number of tested sequence combinations while prioritizing the fact that there is an influence of the previous periods on the following periods. When all typical days are tested, the algorithm ultimately selects the sequence showing the highest R^2_{Global} and the lowest E_{Global} calculated in each iteration. This is the final short sequence.

3.2 Generalization of the algorithm

As explained in the introduction, the main aim of using short simulation sequences in a complex dynamic model is to accelerate the simulation process in studies requiring repetitive simulations without the need to manipulate the model. Model parameters are modified in life cycle assessment, statistical studies and optimization studies, which affects the output of the simulations. The outputs are later used to define the best performing model. However, when it comes to sequence reduction methods, these outputs are used as the selection criteria. Thus, a sequence which was generated based on specific output data of a certain parametric combination, might not generate other outputs of the same model with different parameters since the values of the initial selection criteria are different. Therefore, for the generated sequence to be applicable in such studies, it is essential that a single short simulation sequence is able to regenerate the performance functions of a large number of parametric modifications so that the results are reliable throughout the entire study and in all cases. Thus, the TypSS method was adapted to work on several *cases* at the same time and not just one. A *case* being the same case study but with unique parametric characteristics. TypSS then generates a single short simulation sequence for all cases.

Generating this sequence is achieved by simultaneously taking the output data of all cases and treating them at the same time. The algorithm then finds a single typical representative day for each period that would represent all output data for all cases. Therefore, the process is the same as a single case but this time treating the outputs of all cases simultaneously using global values. The values considered are the maximal total deviation sum Δ_{max} (eq. 6) of all data deviation sum Δ_{Total} (eq. 5) in Function 1 and the compromise between the total global coefficient of determination of all data points R^2_{Total} (eq. 7) and the highest individual annual sum error (eq. 8) in Function 2.

$$\Delta_{\mathrm{Total}} = \sum_{i=1}^{n_{\mathrm{criteria}}} \left| \hat{x}_{\mathrm{TYPSS_i}} - \hat{x}_{\mathrm{YEAR_i}} \right|_{\mathrm{all}} \quad (5)$$

Where:
$\hat{x}_{TYPSS_i}$ is the normalized period value obtained by the short sequence
$\hat{x}_{YEAR_i}$ is the normalized period value obtained by the reference annual simulation
$n_{criteria}$ is the number of selection criteria

$$\Delta_{\mathrm{max}} = \max(\Delta_{\mathrm{Total}})_j \quad j \in [1, n_{\mathrm{period}}] \quad (6)$$

Where:
n_{period} is the number of periods

$$R^2_{\mathrm{Total}} = \prod_{i=1}^{n_{\mathrm{criteria}}} R_{\mathrm{all}_i}^{\ 2} \quad (7)$$

Where:
R_{all}^2 is the coefficient of determination of a selection criterion of data for all cases at the same time
$n_{criteria}$ is the number of selection criteria

$$E_{Total} = \max(E_{Global})_i \quad i \in [1, n_{criteria}] \tag{8}$$

Where:

E_{Global} is the global annual sum error

$n_{criteria}$ is the number of selection criteria

4 Application of the TypSS algorithm

4.1 Case study

The case study is a building connected to a combined solar and heat pump system as shown in Figure 2. The model was developed in the European MacSheep project [46]. The project aimed to develop and evaluate compact combined renewable energy systems based on solar thermal and heat pump technology for space heating and hot water. The energy and economic performance of several systems were assessed and compared by means of a different model designed in Trnsys17. Most of the components involved in these models have been validated against experimental data from prototypes or commercial products as indicated in [47]. The reference system presented in Figure 2 was chosen to evaluate the TypSS algorithm due to its detailed holistic nature and as it was somewhat time consuming to evaluate in repetitive dynamic simulations (20 min for a single simulation) during system characterization.

The envelope is a two-story building with a net floor area of 140 m² (70m² for each floor) and insulated with a 12 cm EPS layer derived from the IEA SHC Task 44/HPP Annex 38 (T44A38) [47]. It is considered to be one common thermal zone. The internal thermal capacity of the inner air volume and walls is taken into account. The system is made up of:

- Solar thermal collectors with a surface of 9.28 m² represented by a multimode dynamic collector [48].
- A 0.763 m^3 storage tank considered to be a one-dimensional multinode model [49].
- A 5kW air source compression heat pump represented by a semi-physical model based on a calculation of the thermodynamic refrigerant cycle and the thermal properties of the refrigerant used [50] to ensure heating and domestic hot water supply in the case of poor solar supply.

The model is run using weather data for Chambery, France located near the French Alps, which has a moderate climate, cold winters (reaching an ambient temperature of -8°C) and relatively warm summers (reaching 28°C with occasional showers).

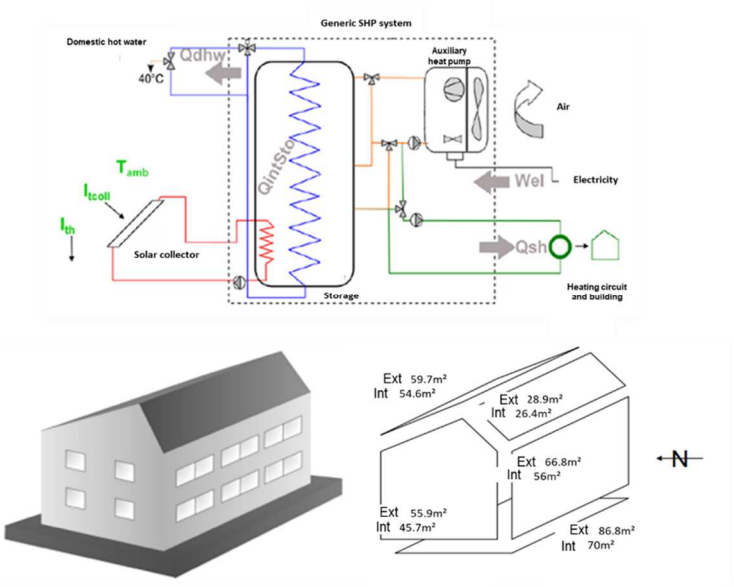


Figure 2. Case study: (Top) solar combisystem connected to a building, (bottom) envelope parts with areas of the interior and exterior facades [47], [51].

Three selection criteria were chosen as inputs for the time reduction algorithm TypSS: the daily-integrated backup energy (in kWh), the daily-integrated energy stored in the tank (in kWh) and the daily-averaged internal room temperature of the building (in °C). The backup energy is the electrical energy consumed by the heat pump to supply heating and domestic hot water. The energy stored in the tank provides an image of the energy content of the store based on its mean temperature. The choice of these criteria was based on the nature of the model and the aim of subsequent optimization of energy consumption taking into consideration the comfort of the occupants.

4.2 Algorithm output for a single case

4.2.1 Reduced sequence

Table 1 presents a 12-day short sequence in addition to the length of each period obtained by TypSS. It shows that the algorithm chose to break the inter-seasonal periods of the year into smaller ones therefore considering more days for these parts of the year. This is influenced by the nature of the selection criteria used, which represent high variations during this time of the year. This result appears in the table in periods 4,5,8 and 9 which had between 5 and 12 days only. In contrast, the seasonal periods, or the periods with stable hot or cold weather remained large, reaching up to 91 days. This was expected since the performance of the selection criteria during this time of the year have somewhat consistent profiles.

Table 1. Typical short sequence of 12 days and the number of days in each period.

Period	1	2	3	4	5	6	7	8	9	10	11	12
Number of days	91	22	23	11	12	23	91	5	6	12	23	46
From	Jan 1	April 1	April 23	May 16	May 27	June 8	July 1	Sep 30	Oct 5	Oct 11	Oct 23	Nov 15
To	Mar 31	April 22	May 15	May 26	June 7	June 30	Sep 29	Oct 4	Oct 10	Oct 22	Nov 14	Dec 31
Selected Day	Jan 27	April 12	April 23	May 24	June 1	June 26	July 10	Oct 3	Oct 6	Oct 15	Oct 31	Dec 26

Figure 3 shows the 12-day ambient temperature and horizontal radiation profiles (Figure 3(b)) compared to the annual reference profiles (Figure 3(a)) as an example of two of the data profiles that will be introduced to the simulation model. Figure 3(a) also shows the selected days in Table 1 as they are distributed in the year (in orange). The values are per hour, therefore, profiles on the left show 365x24=8760 data while those to the right show 12x24=288 data. The discontinuities between the selected days appear clearly in the reduced ambient temperature profile (right). Those discontinuities are more noticeable around the representative days of the larger periods, i.e., periods 1, 7, 11 and 12. The discontinuities are due to the days being taken from different parts of the year. Since the days are distinct as shown in Table 1, some climate characteristics, including the ambient temperature, will be discontinuous. This is not visible in the global horizontal profile in the same figure because the daily values of this characteristic always start and end with zero no matter the position during the year. The TypSS algorithm does not work on limiting these discontinuities but rather regenerates the performance of the model despite their presence.

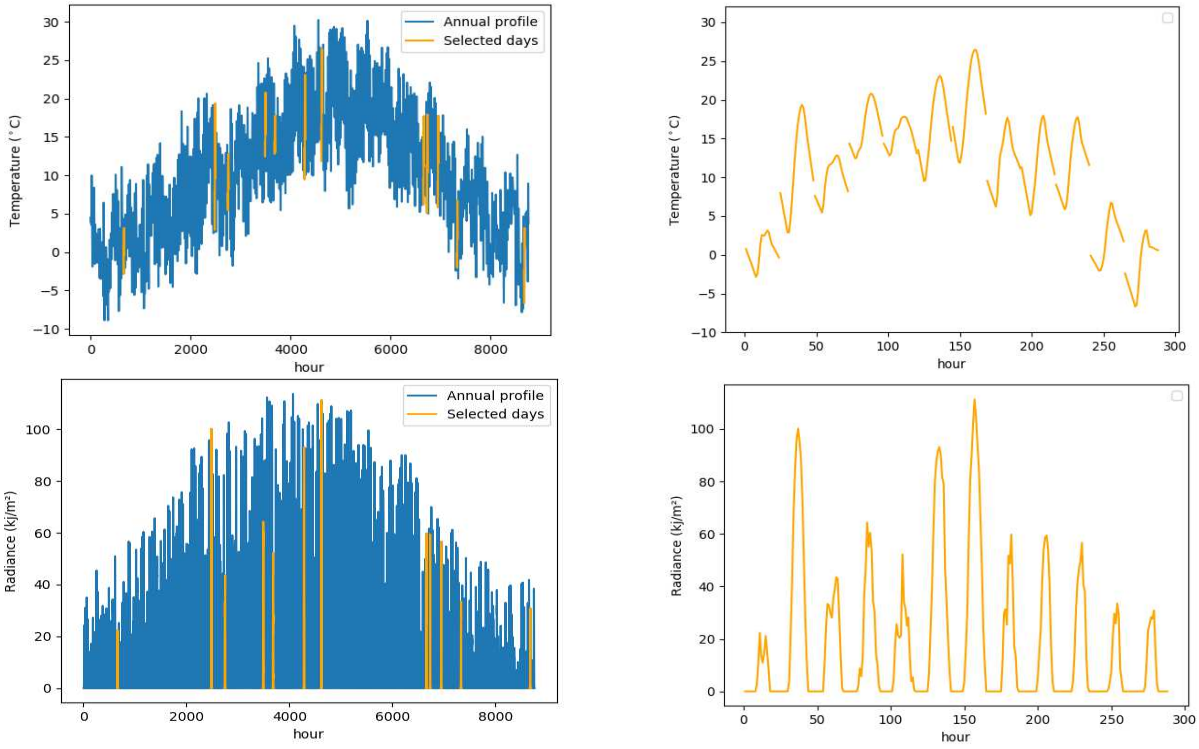
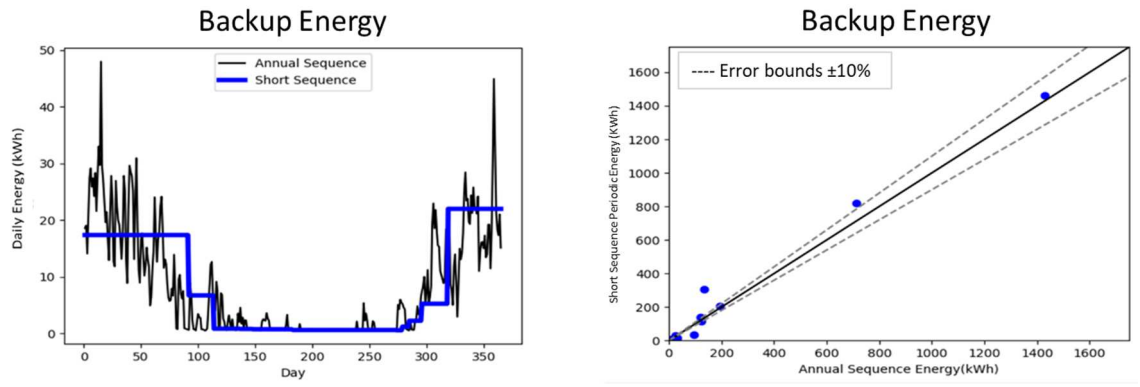


Figure 3. The hourly ambient temperature and global horizontal radiation profiles: (a) reference annual profile (in blue) and the 12 selected days (in orange), (b) 12 selected days profile.

4.2.2 Temporal profiles

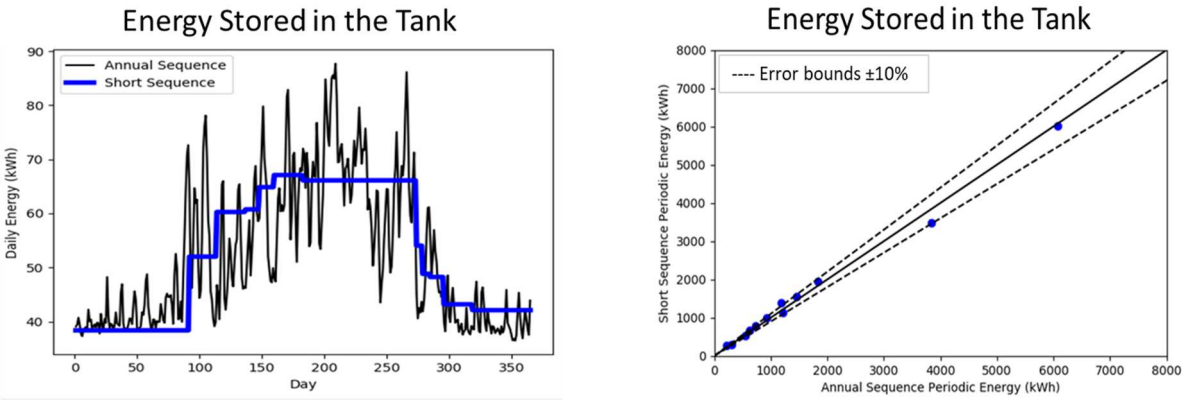
Simulating the model on the sequence of 12 days was about 25 times faster than the annual one. It took about 40 s for this case study while it takes 20 min for a full simulation with the complete sequence on the same computer configuration. Figures 4 to 6 show the results obtained for each period when simulating the model with the short sequence comparing them to the reference values obtained when running a full year simulation.

Regarding the temporal profiles, Figures 4(a), 5(a) and 6(a)) show profiles of daily values, and therefore the reduced simulation generates step-like profiles. Each step is the repetitive performance of the selected day through the period it is representing. The plots show that the output of the short simulation sequence is of the same profile as the annual profiles. The reduced backup energy curve decreases gradually with time until it reaches its minimum during the summer period before it starts increasing again following the same profile of the reference profile. This evolution is reversed in the case of the other two criteria. The curves start at their minimal values before they increase gradually through the year until reaching their maximum in summer after which they start decreasing gradually through the year; still following the evolution of the reference profiles.



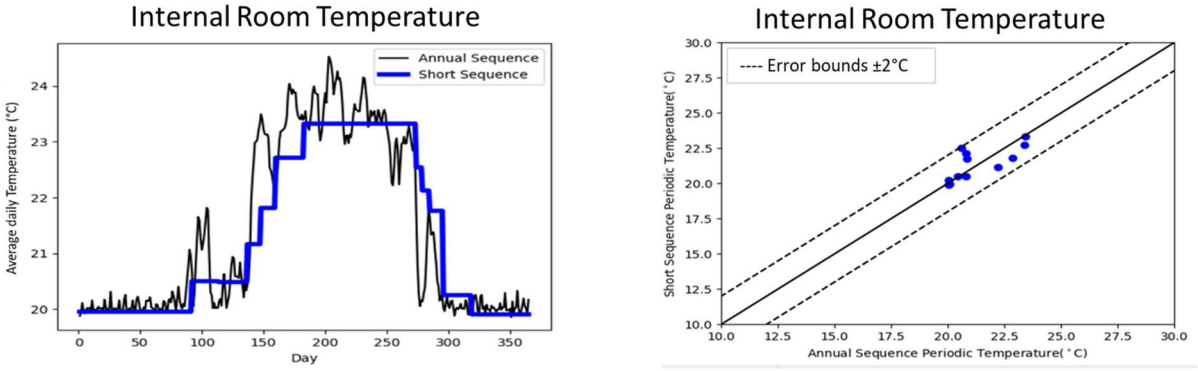
(a) *(b)*

Figure 4. Comparison between reference and extrapolated predicted backup energy: (a) temporal daily profile, (b) integrated values per period.



(a) *(b)*

Figure 5. Comparison between reference and extrapolated predicted energy stored in the tank: (a) temporal daily profile,(b) integrated values per period.



(a) *(b)*

Figure 6. Comparison between reference and extrapolated predicted internal room temperature: (a) temporal daily profile, (b) averaged values per period.

Moreover, the values recorded for the integrated and averaged values of the selection criteria per period are presented in Figures 4(b), 5(b) and 6(b). The x_axis represents the reference values per period while the y axis represents the extrapolated value of the typical days based on the number of days of their corresponding period. The identity line represents the precision of the predicted results, the closer the points are to this line the better the prediction is. The scatter points show high correlation with the identity lines for each criterion. The figures show that for most of the periods, values are within the 10% error limits for the energy criteria and the ±2°C limits for the internal room temperature. The statistical measure that enables this comparison to be quantified is the Coefficient of Variation of Root-Mean Squared Error or CV(RMSE). While the Root-Mean Square Error (RMSE) indicates the absolute fit of the model and shows how close the predicted values are to the actual data points giving an objective representation of the predictive accuracy of the model, CV(RMSE) (eq. 9) takes this metric one step further, by normalizing it using the average dependent variable value. According to ASHRAE Guideline 14, a CV(RMSE) of 25% is the maximum limit for compliance with the guideline [52]. For the periodic dataset given above, the CV(RMSE) was found to be 15.8% for backup energy, 6.2% for energy stored in the tank and 2.2% for internal room temperature, suggesting that the model is reliable.

$$\text{CV(RMSE)} = \frac{1}{\bar{x}_{YEAR}} \times \sqrt{\frac{\sum_{i=1}^{n_{period}} (x_{TYPSS_i} - x_{YEAR_i})^2}{n_{period}}} = \frac{1}{\bar{x}_{YEAR}} \times RMSE \quad (9)$$

Where:

n_{period} is the number of periods

x_{TYPSS_i} is the value of criterion x obtained by the short sequence for period i
x_{YEAR_i} is the value of criterion x obtained by the reference annual sequence for period i
$\overline{x}_{YEAR}$ is the mean value of criterion x in the reference annual sequence
RMSE is Root-Mean Square Error

4.2.3 Annual sum and cumulative profiles

In addition to temporal profiles, the annual sum of the criteria studied and the cumulative profiles are also important for any characterization or optimization study as they allow direct reading and comparison of the system performance through the year until reaching the final annual value. Therefore, it is important for the predicted curves to reflect the annual reference curves. Figure 7 shows the cumulative profiles of the backup energy (Figure 7(a)), the energy stored in the tank (Figure 7(b)) and the internal room temperature (Figure 7(c)) as obtained by both the reduced sequence (in blue) and the reference case (in black).

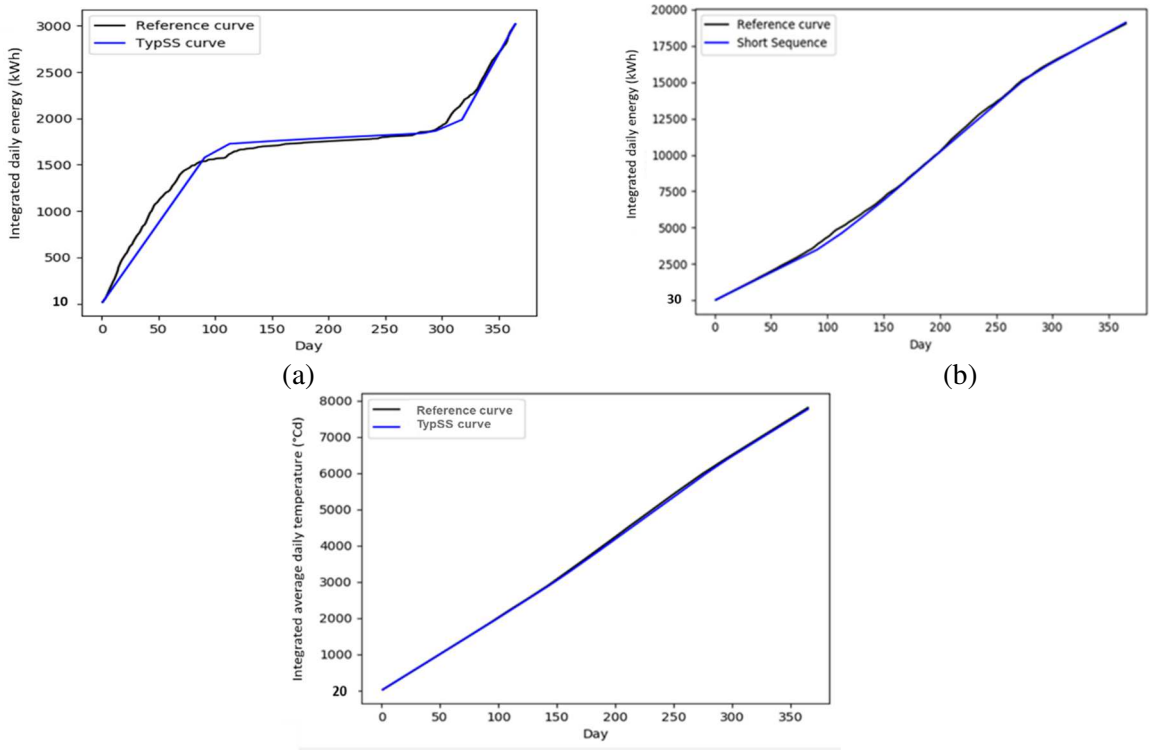


Figure 7. Annual and extrapolated reduced cumulative profiles of the selection criteria: (a) integrated backup energy, (b) integrated energy stored in the tank, (c) integrated internal room temperature.

With a minor deviation between the annual and reduced sequence curves, the figure shows that the curves have a high correlation with the reference curves; R^2 values for the backup energy, energy stored in the tank and internal room temperature are 0.97, 0.99 and 0.99 respectively. In addition to the coefficient of determination, Table 2 shows the CV(RMSE) with values recorded between 1.1 and 5.8%. Moreover, the final annual sum is estimated with an error of 0.1% for backup energy, 0.4% for energy stored in the tank and 0.5% for the internal room temperature as shown in Table 2.

Table 2. Reference and predicted annual sum of the selection criteria.

Criteria	Reference annual sum	Predicted annual sum	Error	CV(RMSE)
Backup energy	3017 kWh	3020 kWh	0.1%	5.8%
Energy stored in the tank	19011 kWh	19095 kWh	0.4%	1.8%
Internal room temperature	7804 °Cd	7763 °Cd	0.5%	1.1%

4.2.4 Impact of the number of typical days

Starting from an initial sequence of four typical days representing the four quarters of the year, seven typical day sequences were generated by TypSS. The sequences are of different lengths ranging between 6 and 30 typical days. The global performance values, i.e., the global R^2 and annual sum errors of each of the selection criteria have been analyzed and are shown in Figure 8. The curves show that good results are achieved even with very short sequences. However, increasing the number of typical days will help improve performance. Regarding backup energy, global R^2 increased to 0.99 with a 30-day sequence in contrast to 0.88 with a 6-day sequence. The annual sum error fluctuated between 0.02% and 2.5%. On the other hand, the performance of the other two criteria showed an almost stable recording for the global R^2 of around 0.99 and a decreasing annual sum error as the number of typical days increase. Finally, the performance of a 20-day sequence differed noticeably and did not follow the trend of the curves, performing worse than shorter sequences. This supports findings in the literature indicating that longer sequences do not always mean better performance. This could be reasoned to the fact that in longer sequences, the size of periods are generally smaller, therefore the possibility of catching specific extreme periodic instants (several consecutive days of heat wave for example), which are more difficult to predict when considering the proposed approach, is higher. Rather, the choice of sequence length is directly related to the case study and initial conditions.

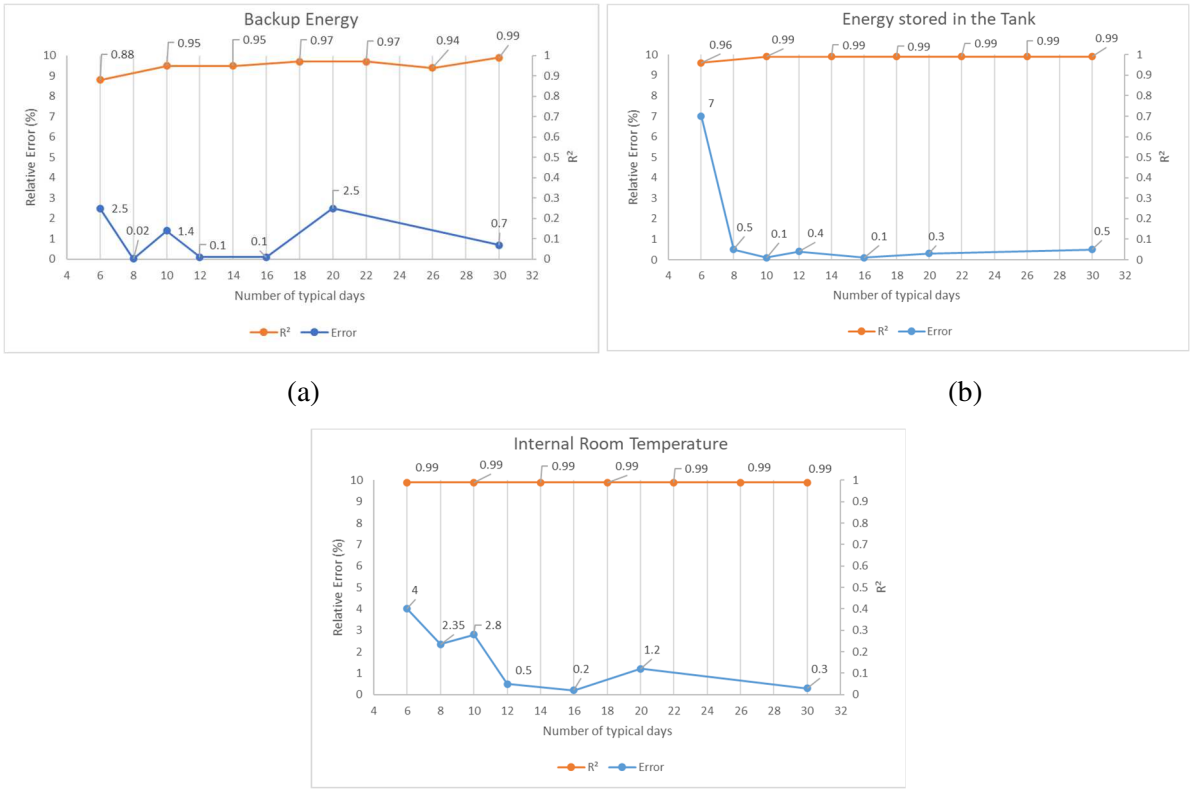


Figure 8. Performance of sequences of different sizes generated by TypSS for R^2 and relative error: (a) backup energy, (b) energy stored in the tank, (c) internal room temperature.

4.2.5 Comparison with other approaches

In order to compare the value of the new method with some commonly used methods in the literature the results in Section 4 were compared with two alternative approaches and applied to the same case study:

Clustering algorithm, *K*-medoids

K-medoids clustering algorithm is an exclusive algorithm which falls under the partitional clustering approaches. It divides data segments into a predetermined number of clusters in which the elements of

a cluster are unique and therefore not shared by other clusters. The difference between *K*-means and *K*-medoids clustering is that the former assigns a group center as the exact mean of the group, which might not be an existing group element, while the latter searches for an existing element closest to the mean and assigns it as the group center (Figure 9). This method of clustering is more realistic in the case of searching for an actual typical representative day. In this study, the *K*-medoids technique was applied twice to two different categories of attributes. The first are weather data values (more precisely ambient temperature and horizontal solar radiation) which are boundary condition values that influence the simulation but are independent of the model itself and the possible modifications it might undergo. This category of attributes, i.e., boundary conditions, is usually taken into account when developing SRYs in the literature. The second *k*-medoids study was applied to the performance of the model, i.e., the output of the simulation. Such values are dependent on the structure of the model and change upon parametric modifications. For consistency, the attributes used for the clustering algorithm were the previous selection criteria.

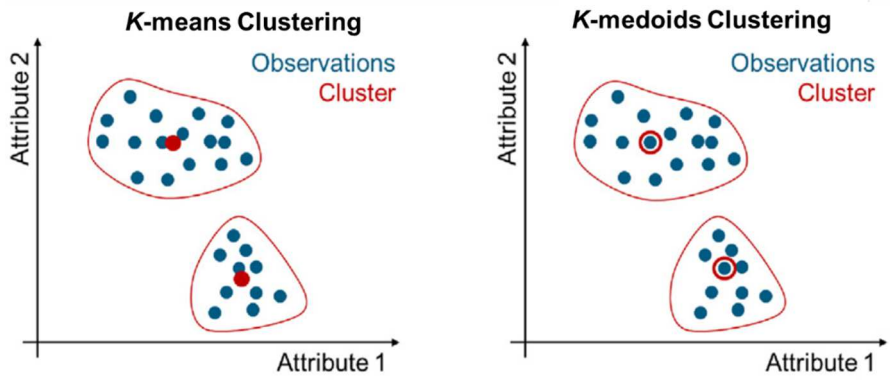


Figure 9. Principles of partitional clustering, Kotzur et al. [41]

Short Cycle System Performance Test (SCSPT)

In order to build a 12-day short sequence, the iterative method SCSPT uses monthly climate data in addition to several monthly performance criteria as attributes for the calculation of target selection criteria using empirical equations. The algorithm then starts from a random initial 12-day sequence and searches the typical days that would have the closest weather data characteristics to the three target weather criteria by calculating a global error and limiting it to a threshold [53].

The three sequence reduction methods are applied to the same solar combisystem model and therefore three sequences of 12-days are generated based on the same selection criteria used in the previous part. These criteria are used directly by the TypSS and clustering approaches and indirectly by the SCSPT approach through the empirical equations. The time consumed by each method to find its own sequence varies significantly. While the clustering algorithm was the fastest, taking a couple of minutes, it took about 2.5 hrs for the SCSPT method and 3 hrs for the TypSS algorithm to converge to their final sequences due to the repetitive simulation by the test sequences of the model. This significant difference in computational time will be compensated in following time-consuming studies such as parametric analysis or optimization studies. As mentioned previously, the aim of generating robust and generalized reduced sequences is to be used in studies with repetitive simulations that take tens and maybe hundreds of hours. Therefore, reducing the simulation time of a single simulation will lead to the reduction of the global time of heavy studies.

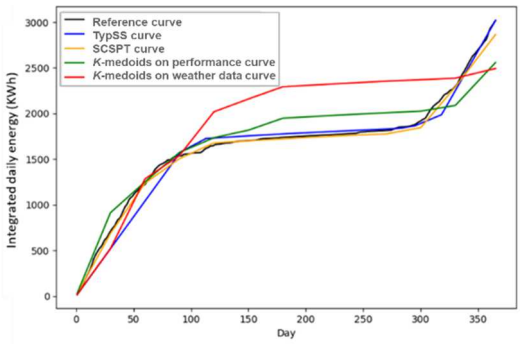
Simulating the model on the final sequences of each method gives the results in Figure 10 and detailed in Table 3.

Criterion	Backup Electrical Energy		Energy Stored in the Tank		Room Temperature	
Method	R^2	Annual Sum Error (%)	R^2	Annual Sum Error (%)	R^2	Annual Sum Error (%)
Weather data	0.51	17	0.97	1	0.99	0.6

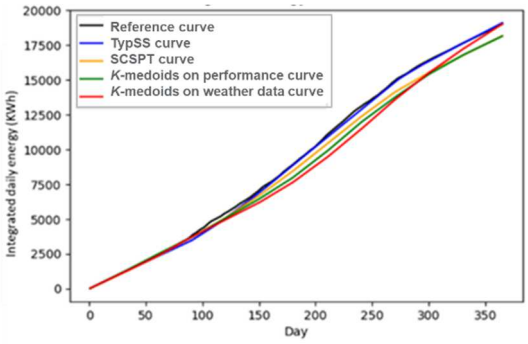
k-medoids Performance	0.88	15	0.98	4.5	0.99	0.6
k-medoids						
SCSPT	0.99	4.5	0.98	4	0.99	2.1
TypSS	0.97	0.1	0.99	0.4	0.99	0.5

Table 3. Comparison of the three time reduction method results.

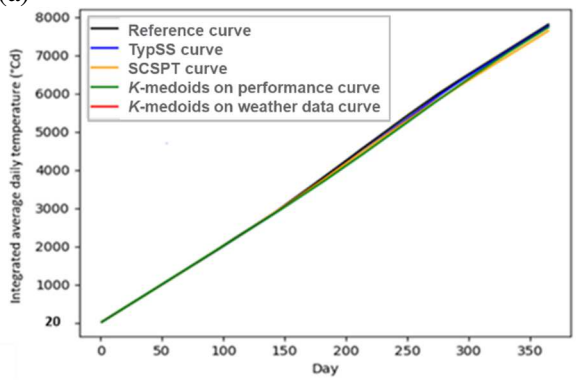
Figure 10 shows that the cumulative profiles were better generated globally using the TypSS approach compared with the other approaches. In the case of the backup energy curve (Figure 10(a)), the TypSS curve (in blue) followed with a good correlation with the reference curve (in black), recording an R^2 of 0.97. At the end of the winter period, where need for electricity decreases, the curve overestimates electricity consumption. However, this overestimation is then corrected at the beginning of the heating season, leading to a final value close to the reference value (0.1% difference as shown in Table 3). The SCSPT curve (in orange) showed the best cumulative profile of this criterion almost replicating the reference curve. However, the curve deviates at the end presenting an error of 4.5% in the annual sum estimation. This proximity in covering the variations of the relative curve is due to the length of the periods. The winter season is represented by only one day in the TypSS case (while three in the SCSPT case) which makes this period very sensitive to estimation error. However, these errors are then offset since more days are used to describe inter-seasonal periods. Finally, the clustering curves (in green and red) were the least efficient. They failed to attain the annual sum, with a 15% error for clustering on performance and 17% for clustering on weather data, but the evolution of the curves poorly reflected the reference curve, with R^2 of 0.88 and 0.51 respectively. Unlike the other two approaches, the clustering approach does not take into consideration the simulation process and the effect of representative days on each other when constructing the short sequence. This appears more clearly when comparing the *k*-medoids studies where using performance from the annual simulation was better than using weather data as these values were determined taking into consideration the reference simulation process. On the other hand, the evolution of the curves was better for the four studies for energy stored in the tank (Figure 10(b)) and internal room temperature (Figure 10(c)), with the TypSS method performing the best in terms of annual sum errors while worse for the basic clustering algorithm.



(a)



(b)



(c)

Figure 10. Annual and extrapolated reduced cumulative profiles as obtained by the three methods: (a) integrated backup energy, (b) integrated energy stored in the tank, (c) internal room temperature.

These results are encouraging and indicate that the TypSS algorithm is applicable to further studies requiring considerable computational time and that evaluate the annual sum, such as optimization or statistical studies. Its efficiency upon generalization is assessed in the following section.

4.3 Multiple cases

Modifying the surface of the solar collector (SSC), the volume of the storage tank (VST) of the solar combisystem and also the thickness of the insulating material (INS) of the building in the case study leads to different individual cases. Using Latin Hypercube Sampling (LHS), 50 individual cases of the solar combisystem presented previously in Figure 2 were generated with the SSC ranging between 6.5-25m², the VST ranging between 0.3-1m^3 and the INS ranging between 0.04-0.3m. Five distinct individual cases were then selected out of the 50 to be run by the TypSS algorithm and generate a single short simulation sequence. The sequence obtained appeared to be different from the sequence obtained with a single case.

4.3.1 Simulation output

Figure 11 shows the cumulative profiles of the selection criteria as obtained by the reduced sequence (dashed line) and the reference case (solid line). Each individual case is given a unique color for better visualization of the results. In addition, the coefficients of determination and CV(RMSE) of each curve are recorded in Table 4. The figure shows that despite using a single short sequence of 12 days for simulating five different cases each having unique parametric characteristics, the generated curves correlate very well with the reference annual curves, for all individual cases and all selection criteria. Regarding backup energy (Figure 11(a)), the sequence generated almost identical curves for cases 1 and 2, and very close curves with minor deviations for cases 3, 4 and 5.

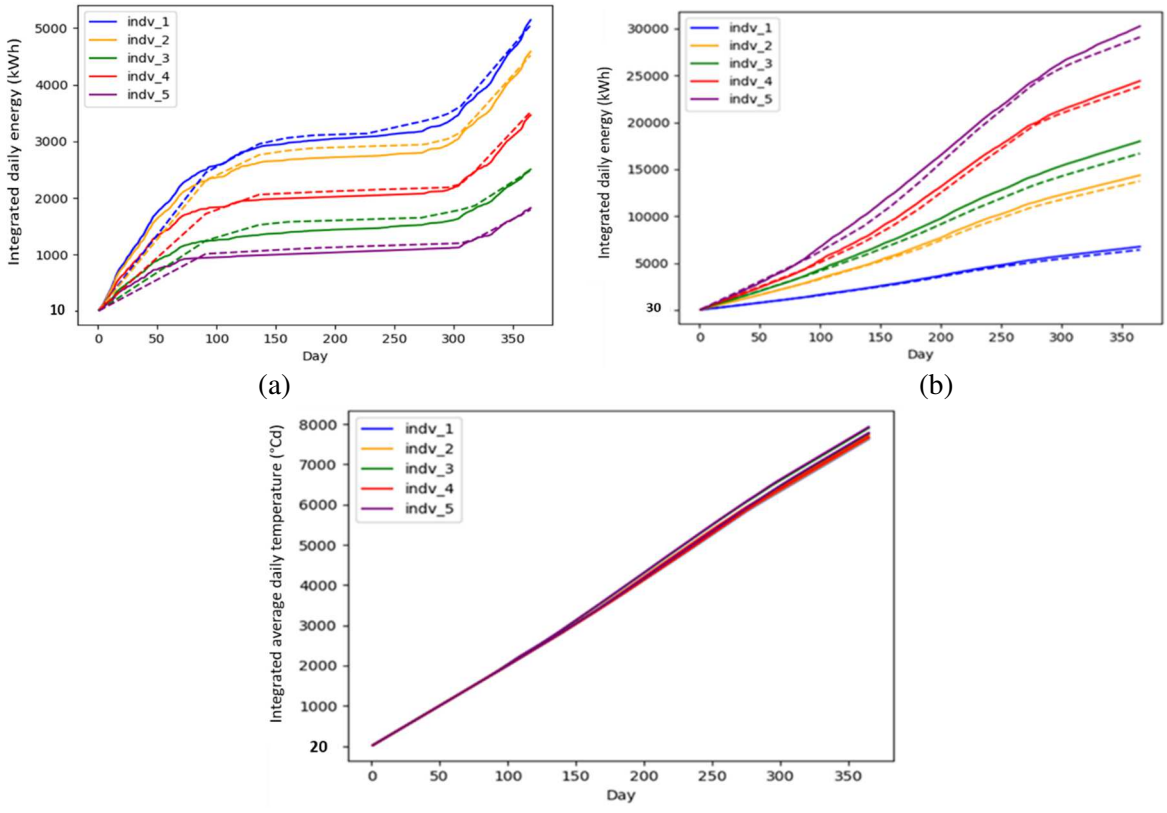


Figure 11. Annual (solid) and extrapolated reduced (dashed) cumulative profiles as obtained by the five individual cases: (a) backup energy, (b) energy stored in the tank, (c) internal room temperature.

Table 4 shows that the R^2 value for backup energy ranged between 0.92 and 0.97 and CV(RMSE) was less than 15.2%. Moreover, the curves generated for the other selection criteria, energy stored in the

tank (Figure 11(b)) and internal room temperature (Figure 11(c)), showed high correlation with the reference curves. The former showed R^2 between 0.98 and 0.99 and CV(RMSE) less than 10.3% and the latter R^2 equal to 0.99 and CV(RMSE) less than 3.5%, therefore leading to a global R^2 of 0.98 for all selection criteria in all of the cases.

Table 4. The global and individual coefficient of determination of the three selection criteria.

Criteria	**Backup Energy**		**Energy Stored in the Tank**		**Internal Room Temperature**	
Cases	R^2	CVRMSE	R^2	CVRMSE	R^2	CVRMSE
1	0.97	10.2	0.99	6.4	0.99	2.5
2	0.96	8.4	0.99	4.2	0.99	2.4
3	0.91	15.2	0.98	6.8	0.99	3.3
4	0.94	13.3	0.99	6.1	0.99	2.7
5	0.92	11.4	0.99	10.3	0.99	3.5
Global R^2	**0.98**					

Additionally, Table 5 shows that the annual sums of selection criteria were estimated with good precision and relative error not exceeding 2% for backup energy and internal room temperature, and 8% for energy stored in the tank between the reference values (AN) and the predicted values (TS).

Table 5. The reference and predicted annual sums and their relative errors of the selection criteria per case.

Criteria	**Backup Energy**			**Energy Stored in the Tank**			**Internal Room Temperature**		
Cases	AN (kWh)	TS (kWh)	Error (%)	AN (kWh)	TS (kWh)	Error (%)	AN (°Cd)	TS (°Cd)	Error (%)
1	5142	5052	1.7	6748	6414	4.9	7676	7633	1
2	4586	4518	1.5	14361	13746	4.3	7697	7648	1
3	2504	2486	1	17988	16685	7.2	7903	7770	1.6
4	3461	3520	1.7	24423	23824	2.4	7745	7676	1
5	1822	1793	1.6	30251	29085	3.9	7931	7785	1.8

.

4.3.2 Evaluation upon generalization

To assess its generalization quality, this sequence based on five cases is used for simulation of the 50 cases. The results appear in Figure 12 in comparison with the single case-based sequence. The curves show that while the relative errors did not exceed 10% for all selection criteria in the case of the five initial cases sequence, the backup energy was poorly estimated for most of the 50 cases when using the single initial case sequence recording up to a 45% error. On the other hand, while the influence of the number of initial cases did not seem to be noticeable in the case of energy stored in the tank, the performance of a single initial case was slightly better than that of five in the case of internal room temperature, with a minor difference up to 1%. Conversely, considering more initial cases in the day selection process would definitely increase the computational time taken by TypSS to converge. TypSS is a simulation-based algorithm so when considering more cases, more simulations are considered in each iteration which will lead to a higher simulation time. However, this time is offset when using the reduced sequence in intensive studies such as optimization, which would be faster due to considering a 12-day sequence instead of data for 365 days.

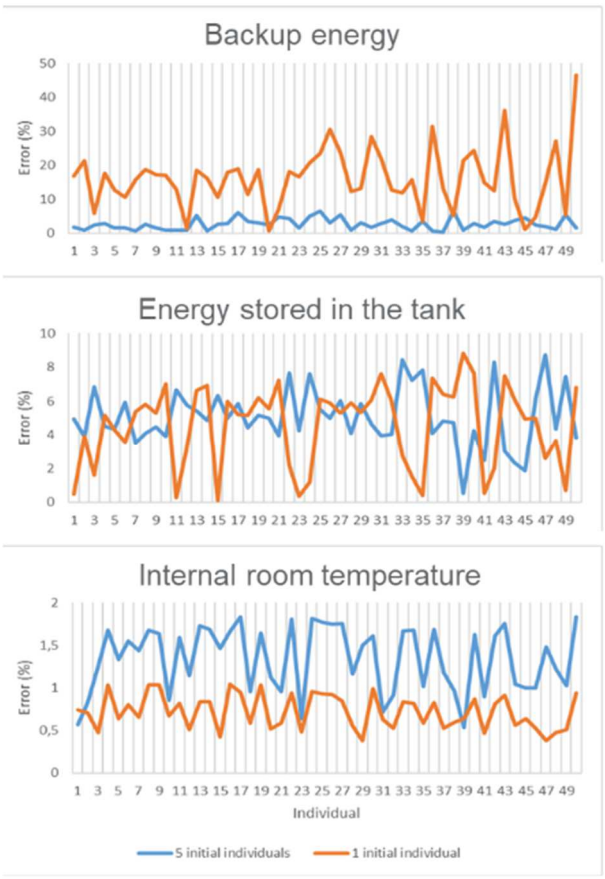


Figure 12. The relative percentage errors of the selection criteria of all 50 cases after simulation with the typical day sequences obtained with one initial case (orange) and five initial cases (blue).

5 Conclusions and perspectives

A new approach to building performance simulation was developed and tested in this work. The approach is iterative with an embedded grouping algorithm. It employs averaged and integrated values of selection criteria chosen by the user to generate reduced sequences used in detailed model studies. The sequence is obtained in a way to take into consideration the annual dynamic behavior of the system and possible parametric modifications the case might undergo in future studies.

Applied on a building model with a solar combisystem, several sequences of different lengths were generated. Recorded global performance values were very good even for very short sequences. The simulation of a sequence of 12 days was about 25 times faster than the annual one using the same computer configuration. In addition to the saved simulation time, results show that the periodic performance of the short simulation sequence are of high correlation with the reference ones recording coefficients of correlation R^2 superior to 0.97 in addition to final annual sum errors of each criterion not exceeding 1%.The analysis of the influence of the number of days on the reduced sequence showed that while increasing the number of days in the reduced sequence helps decrease annual sum errors, a longer sequence does not necessarily mean better results. The curves also showed that the choice of a 12-day sequence in this study is the best choice compared with longer or shorter sequences. The results were also compared with two approaches used in the literature and showed the best performance while the worst performance was for a method commonly used to generate Short Reference Years (SRYs).

Despite using a single short sequence of 12 days for simultaneously simulating five different solar combisystem cases each having a unique parametric configuration, the generated curves correlate well with the reference annual curves for all cases and all selection criteria. Moreover, the annual sums of the energy selection criteria were estimated with a high degree of precision and a relative error not

exceeding 2% for backup energy and internal room temperature and 8% for energy stored in the tank. The same sequence was then tested on 45 other cases not taken into consideration by the algorithm. The curves show that the sequence succeeded in predicting the annual performance of all selection criteria with relative errors not exceeding 10%. The curves also show that generating a sequence on more than a single case improves the quality of the results. This in turn supports the aim of developing a generalized sequence applicable to a wide parametric range.

In perspectives, the algorithm's sensitivity to input conditions such as number of initial periods, number of cases or number and type of selection criteria is currently being tested to evaluate the impact of their modification on the quality of the results. In addition, optimization studies will be implemented directly with complex models without simplification but using the short simulation sequence to assess its efficiency regarding both optimal solutions and computational time expenses. It is at this level of applications where the value of the proposed approach appears since despite the time it requires to generate the reduced sequence, once it is found and implemented to detailed models studies, it will speed up significantly each iteration and therefore the optimization process as a whole.

Acknowledgment

The authors would like to acknowledge the financial support provided by the Region Auvergne-Rhône-Alpes, France for the "Optimization of energy networks and energy producing buildings" OREBE project.

References

[1] L. Pérez-Lombard, J. Ortiz, et C. Pout, « A review on buildings energy consumption information », *Energy Build.*, vol. 40, nº 3, p. 394-398, janv. 2008, doi: 10.1016/j.enbuild.2007.03.007.

[2] J. Berger, S. Gasparin, D. Dutykh, et N. Mendes, « On the comparison of three numerical methods applied to building simulation: Finite-differences, RC circuit approximation and a spectral method », *Build. Simul.*, vol. 13, nº 1, p. 1-18, févr. 2020, doi: 10.1007/s12273-019-0555-z.

[3] M. J. N. Oliveira Panão, C. A. P. Santos, N. M. Mateus, et G. Carrilho da Graça, « Validation of a lumped RC model for thermal simulation of a double skin natural and mechanical ventilated test cell », *Energy Build.*, vol. 121, p. 92-103, juin 2016, doi: 10.1016/j.enbuild.2016.03.054.

[4] T. Weber et G. Jóhannesson, « An optimized RC-network for thermally activated building components », *Build. Environ.*, vol. 40, nº 1, p. 1-14, janv. 2005, doi: 10.1016/j.buildenv.2004.04.012.

[5] D. Kim et J. E. Braun, « A general approach for generating reduced-order models for large multi-zone buildings », *J. Build. Perform. Simul.*, vol. 8, nº 6, p. 435-448, nov. 2015, doi: 10.1080/19401493.2014.977952.

[6] J. Berger, W. Mazuroski, N. Mendes, S. Guernouti, et M. Woloszyn, « 2D whole-building hygrothermal simulation analysis based on a PGD reduced order model », *Energy Build.*, vol. 112, p. 49-61, janv. 2016, doi: 10.1016/j.enbuild.2015.11.023.

[7] S.-J. Cao et J. Meyers, « Fast prediction of indoor pollutant dispersion based on reduced-order ventilation models », *Build. Simul.*, vol. 8, nº 4, p. 415-420, août 2015, doi: 10.1007/s12273-015-0240-9.

[8] B. Eisenhower, Z. O'Neill, S. Narayanan, V. A. Fonoberov, et I. Mezić, « A methodology for meta-model based optimization in building energy models », *Energy Build.*, vol. 47, p. 292-301, avr. 2012, doi: 10.1016/j.enbuild.2011.12.001.

[9] P. Mancarella, « MES (multi-energy systems): An overview of concepts and evaluation models », *Energy*, vol. 65, p. 1-17, févr. 2014, doi: 10.1016/j.energy.2013.10.041.

[10] C. Milan, M. Stadler, G. Cardoso, et S. Mashayekh, « Modeling of non-linear CHP efficiency curves in distributed energy systems », *Appl. Energy*, vol. 148, p. 334-347, juin 2015, doi: 10.1016/j.apenergy.2015.03.053.

[11] I. J. Hall, R. R. Praine, H. E. Anderson, et E. C. Boes, « Generation of Typical Meteorological Years for 26 SOLMET stations. », Laboratories Report SAND 78-1601, Sandia Laboratories, Albuquerque, NM, 1978.

[12] D. L. Siurna, L. J. D'Andrea, et K. G. T. Hollands, « A Canadian representative meteorological year for solar system simulation », Proc. 10th Annual Conference of the Solar Energy Society of Canada (SESCI '84), August 1–6, 1984, Calgary, Alberta, Canada, 1984.
[13] R. Bahm, « TMY Data Handbook », Raymond Bahm and Associates, Albuquerque, NM, 1985.
[14] M. N. Bahadori et M. J. Chamberlain, « A simplification of weather data to evaluate daily and monthly energy needs of residential buildings », *Sol. Energy*, vol. 36, nº 6, p. 499-507, 1986, doi: 10.1016/0038-092X(86)90014-9.
[15] D. Pissimanis, G. Karras, V. Notaridou, et K. Gavra, « The generation of a "typical meteorological year" for the city of Athens », *Sol. Energy*, vol. 40, nº 5, p. 405-411, 1988, doi: 10.1016/0038-092X(88)90095-3.
[16] M. A. Mosalam Shaltout et M. T. Y. Tadros, « Typical solar radiation year for Egypt », *Renew. Energy*, vol. 4, nº 4, p. 387-393, juin 1994, doi: 10.1016/0960-1481(94)90045-0.
[17] M. Petrakis, S. Lykoudis, et P. Kassomenos, « A software tool for the creation of a typical meteorological year », *Environ. Softw.*, vol. 11, nº 4, p. 221-227, janv. 1996, doi: 10.1016/S0266-9838(96)00006-8.
[18] W. Marion et K. Urban, « User's Manual for TMY2s (Typical Meteorological Years) - Derived from the 1961-1990 National Solar Radiation Data Base », NREL/TP-463-7668, 87130, juin 1995. doi: 10.2172/87130.
[19] H. Lund et S. Eidorff, « Selection Methods for Production of Test Reference Years, Appendix D », Contract 284-77 ES DK, Final Report, EUR 7306 EN, 1980.
[20] H. Lund, « H. Lund Short Reference Years-Test Reference Years », Thermal Insulation Laboratory, Technical University of Denmark. Contract ESF-011-DK, Final Report, EUR 9402, 1983.
[21] R. Festa et C. F. Ratto, « Proposal of a numerical procedure to select Reference Years », *Sol. Energy*, vol. 50, nº 1, p. 9-17, janv. 1993, doi: 10.1016/0038-092X(93)90003-7.
[22] B. Anderson, S. Eidroff, H. Lund, E. Pedersen, S. Rosenorn, et O. Valbjorn, « Meteorological Data for Design of Building and Installation: A Reference Year », , (extract), Report No 66 (2nd edition), Thermal Insulation Laboratory, Denmark, 1977.
[23] S. H. Liem et A. H. C. Paassen, « Short description of the Dutch automatic procedure to generate Short Reference Years H. Lund (Ed.), Short Reference Years-Test Reference Years », Contract ESF-011-DK, Final Report, EUR 9402, Brussels, 1982.
[24] H. Lund, « Short Reference Years and Test Reference Years for EEC countries », Thermal Insulation Laboratory, Technical University of Denmark (1985) Contract ESF-029-DK (G), Final Report, EUR 10208 EN, 1985.
[25] A. H. C. Paassen, « A new approach to the calculation of the effect of the outdoor and indoor climate on the energy consumption in buildings based on methods of statistical analysis », 1981.
[26] W. R. Petrie et M. McClintock, « Determining typical weather for use in solar energy simulations », *Sol. Energy*, vol. 21, nº 1, p. 55-59, 1978, doi: 10.1016/0038-092X(78)90116-0.
[27] B. Y. H. Liu et R. C. Jordan, « The interrelationship and characteristic distribution of direct, diffuse and total solar radiation », *Sol. Energy*, vol. 4, nº 3, p. 1-19, juill. 1960, doi: 10.1016/0038-092X(60)90062-1.
[28] A. Belderbos et E. Delarue, « Accounting for flexibility in power system planning with renewables », *Int. J. Electr. Power Energy Syst.*, vol. 71, p. 33-41, oct. 2015, doi: 10.1016/j.ijepes.2015.02.033.
[29] M. Y. Haller, « Dynamic whole system testing of combined renewable heating systems – The current state of the art », *Energy Build.*, vol. 66, p. 667-677, nov. 2013, doi: 10.1016/j.enbuild.2013.07.052.
[30] M. Fripp, « Switch: A Planning Tool for Power Systems with Large Shares of Intermittent Renewable Energy », *Environ. Sci. Technol.*, vol. 46, nº 11, p. 6371-6378, juin 2012, doi: 10.1021/es204645c.
[31] E. K. Hart et M. Z. Jacobson, « A Monte Carlo approach to generator portfolio planning and carbon emissions assessments of systems with large penetrations of variable renewables », *Renew. Energy*, vol. 36, nº 8, p. 2278-2286, août 2011, doi: 10.1016/j.renene.2011.01.015.
[32] B. J. Brinkworth, « Autocorrelation and stochastic modelling of insolation sequences », *Sol. Energy*, vol. 19, nº 4, p. 343-347, 1977, doi: 10.1016/0038-092X(77)90004-4.

[33] C. Mustacchi, V. Cena, et M. Rocchi, « Stochastic simulation of hourly global radiation sequences », *Sol. Energy*, vol. 23, nº 1, p. 47-51, 1979, doi: 10.1016/0038-092X(79)90042-2.
[34] P. Poggi, G. Notton, M. Museli, et A. Louche, « Stochastic study of hourly total solar radiation in Corsica using a Markov model », *International journal of climatology*, p. 1843-1860, 2000.
[35] D. H. Nall et E. A. Arens, « Climate data abbreviation for the computerised calculation of heating and cooling requirements in buildings », *Energy Build.*, vol. 2, nº 2, p. 135-149, avr. 1979, doi: 10.1016/0378-7788(79)90029-X.
[36] J. Ortiga, J. C. Bruno, et A. Coronas, « Selection of typical days for the characterisation of energy demand in cogeneration and trigeneration optimisation models for buildings », *Energy Convers. Manag.*, vol. 52, nº 4, p. 1934-1942, avr. 2011, doi: 10.1016/j.enconman.2010.11.022.
[37] J. B. Barney, « Strategic Factor Markets: Expectations, Luck, and Business Strategy », *Manag. Sci.*, vol. 32, nº 10, p. 1231-1241, oct. 1986, doi: 10.1287/mnsc.32.10.1231.
[38] S. Fazlollahi, G. Becker, A. Ashouri, et F. Maréchal, « Multi-objective, multi-period optimization of district energy systems: IV – A case study », *Energy*, vol. 84, p. 365-381, mai 2015, doi: 10.1016/j.energy.2015.03.003.
[39] S. Fazlollahi, S. L. Bungener, P. Mandel, G. Becker, et F. Maréchal, « Multi-objectives, multi-period optimization of district energy systems: I. Selection of typical operating periods », *Comput. Chem. Eng.*, vol. 65, p. 54-66, juin 2014, doi: 10.1016/j.compchemeng.2014.03.005.
[40] F. Domínguez-Muñoz, J. M. Cejudo-López, A. Carrillo-Andrés, et M. Gallardo-Salazar, « Selection of typical demand days for CHP optimization », *Energy Build.*, vol. 43, nº 11, p. 3036-3043, nov. 2011, doi: 10.1016/j.enbuild.2011.07.024.
[41] L. Kotzur, P. Markewitz, M. Robinius, et D. Stolten, « Impact of different time series aggregation methods on optimal energy system design », *Renew. Energy*, vol. 117, p. 474-487, mars 2018, doi: 10.1016/j.renene.2017.10.017.
[42] D. Menegon, A. Soppelsa, et R. Fedrizzi, « Development of a new dynamic test procedure for the laboratory characterization of a whole heating and cooling system », *Appl. Energy*, vol. 205, p. 976-990, nov. 2017, doi: 10.1016/j.apenergy.2017.08.120.
[43] C. Ribault, « Multicriteria optimization method for design assistance of urban densification process. », Université de Lyon, 2017. [En ligne]. Disponible sur: https://tel.archives-ouvertes.fr/tel-02000276/.
[44] P. Balachandra et V. Chandru, « Modelling electricity demand with representative load curves », *Energy*, vol. 24, nº 3, p. 219-230, mars 1999, doi: 10.1016/S0360-5442(98)00096-6.
[45] H. Sayegh, A. Leconte, G. Fraisse, E. Wurtz, et S. Rouchier, « Determination of a Short Simulation Sequence for the Multi-Criteria Optimization of Buildings: Bibliographic Study and Perspectives », In the 8th Conference on Sustainable Development of Energy, Water and Environment Systems (SDEWES13), 2018.
[46] « MacSheep Project web page ». <https://www.macsheep.spf.ch/Project.17970.0.html.> [accessed 23.02.2020].
[47] R. Dott, M. Y. Haller, J. Ruschenburg, F. Ochs, et J. Bony, « The Reference Framework for System Simulations - Part B: Buildings and Space Heat Load. IEA SHC Task 44 / HPP Annex 38 », International Energy Agency., 2013.
[48] M. Y. Haller, B. Perers, C. Bales, J. paavilainen, A. Dalibard, et S. Fischer, « TRNSYS Type 832 v5. 00 "Dynamic collector Model by Bengt Perers" », 2012.
[49] H. Druck et T. Pauschinger, « Store-Model Type 340; 2006. », 2006.
[50] A. Heinz et M. Y. Haller, « Appendix A3-Description of TRNSYS Type 877 by IWT and SPF. Models of Sub-Components and Validation for the IEA SHC Task. », 2012.
[51] D. Cheze, A. Papillon, A. Leconte, M.Y. Haller, R. Habert, T. Persson, C. Bales, R. Habert, « Towards an harmonized whole system test method for combined renewable heating systems for houses », In Proc. of the EuroSun 2014 conference, pp. 16-19. 2014., 2014.
[52] J. S. Haberl, D. E. Claridge, et C. Culp, « ASHRAE's Guideline 14-2002 for Measurement of Energy and Demand Savings: How to Determine what was really saved by the retrofit. », 2005.
[53] H. Sayegh, G. Fraisse, A. Leconte, E. Wurtz, O. Ouvrier Bonaz, et S. Rouchier, « Determination Of A Short Simulation Sequence For The Multi-Criteria Optimization Of Buildings: A Case Study », Rome, Italy, p. 1280-1287. doi: 10.26868/25222708.2019.210214.