Energy performance certificates in the USA and in France—a case study of multifamily housing

Berangere Lartigue · Laura Biewesch · Flore Marion · Erica Cochran · Françoise Thellier

Received: 13 April 2021 / Accepted: 24 April 2022 / Published online: 3 May 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract In the USA and the European Union, buildings account for more than 40% of total energy use and a large proportion of buildings are energy inefficient. Countries address these inefficiency challenges with various initiatives and strategies. One of them relies on rating buildings with energy performance certificates, with the goal that awareness on energy consumption would lead to an efficient retrofit. In this article, we analyze the different methods chosen by the USA and France to rate multifamily buildings, i.e., the Energy Star score and the Diagnostic de Performance Energétique. We conduct a case study of a multifamily housing using a Design of Experiments to determine what inputs are the most influential on the output. In the French certificate, the results show that the climate, ventilation system, and building envelope are the most influential inputs on the energy consumption. In the USA certificate, the actual energy consumption and the climate are the most influential factors on the building score. We then discuss the significant differences in the two approaches, and the consequences in terms of accuracy, as well as how the DPE and ES scores are used as a tool in public energy policy to propose energy conservation measures and reduce energy consumption.

Keywords Energy performance certificate · Multifamily housing · Design of Experiments · Energy policy · Energy Star score · Diagnostic de Performance Energétique

Abbreviations

- ASEUI: Actual Source Energy Use Intensity
- CDD: Cooling Degree Days
- DoE: Design of Experiments
- DPE: Diagnostic de Performance Energétique
- ENV: Envelope of the building
- EPC: Energy Performance Certificate
- ES: Energy Star
- EU: European Union
- GFA: Gross floor area
- HDD: Heating degree days
- H-HW: Heating and hot water systems
- NFA: Net floor area
- PSEUI: Predicted Source Energy Use Intensity
- TSEU: Total source energy use
- USA: United States of America
- VENT: Ventilation system

B. Lartigue (✉) · L. Biewesch · F. Thellier
Laboratoire Matériaux et Durabilité des Constructions, Université Toulouse III–Paul Sabatier, 118, route de Narbonne, 31062 Toulouse, France
e-mail: berangere.lartigue@univ-tlse3.fr

B. Lartigue · L. Biewesch · F. Marion · E. Cochran
Center for Building Performance and Diagnostics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

© Springer
Introduction

Background

In the USA and the European Union, buildings account for approximately 40% of total energy use (US Energy Information Administration, 2021; European Commission, 2021). Improving their energy performance is a goal in mitigating the climate change. The USA and the EU address this goal with various initiatives and strategies. One common strategy relies on rating buildings: governments provide energy tracking and benchmarking tools, with the objective to help building owners assess the energy performance, establish baselines, and set performance improvement goals. Yet, these tools may differ a lot in their own structure, the input they require, their calculation core, and their output.

For example, the US Environmental Protection Agency (EPA) developed Energy Star Portfolio Manager (ES-PM), an operational rating tool available online, for building owners to voluntarily evaluate energy use in buildings. Currently, 18 categories of buildings (such as education, office, or retail) which include more than 80 property types (such as K-12 school, or medical office) can receive a 1–100 Energy Star (ES) score. This score compares the building under investigation to other buildings nationwide that have the same property type (US Environmental Protection Agency, How the 1–100 Energy Star score is calculated? n.d.a). As an example, a score of 50 points represents a median energy performance, while a score of 75 points indicates that the building performs better than 75% of its peer group. Only buildings scoring more than 75 may be eligible for Energy Star certification. Since the launch of ES-PM in 1992, by the end of 2015, more than 450,000 buildings (all types) were benchmarked with the tool. Since 1992, nearly 38,000 commercial buildings have been Energy Star certified, meaning that their score is above 75 (US Environmental Protection Agency, 2020).

In the European Union, the Energy Performance of Buildings Directive required since 2002 all member states to implement Energy Performance Certificates of new and existing buildings (Official Journal of the European Union, 2018). One of the goals of this directive is to increase the energy efficiency of the building stock. Yet, specifics of the method of certification are left to each European country to define. One representative example of an Energy Performance Certificate is the French ‘Diagnostic de Performance Énergétique’ (DPE), implemented in France. In 2022, more than 11 million of French buildings (all types) have been certified by a DPE (Agence de l’Environnement et de la Maîtrise de l’Energie, 2022). This certification is mandatory when selling or changing tenants.

The French DPE and the American ES differ in their approach; the DPE is an asset-rating tool while ES-PM is an operational rating tool. An asset rating looks at the building performance independently of its occupants, based on a model of the building itself, while an operational rating focuses on the actual energy use (Goldstein & Eley, 2014). According to Goldstein and Eley (2014), an asset rating is more valuable when wanting to know how efficient a building is, while an operational rating is more valuable for benchmarking against peers.

Pérez-Lombard et al. (2009) reviewed the origin and the historic development of energy certification schemes in buildings. They underlined that energy certification schemes for buildings emerged in the early 1990s as an essential method for improving energy efficiency, minimizing energy consumption, and enabling greater transparency with regards to the use of energy in buildings.

Some recent reviews by Li et al. (2019), Economidou et al. (2020), or Semple and Jenkins (2020) summarize the abundant literature that covers studies conducted on EPC in different countries, especially in Europe. Taranu and Verbeeck (2018) made a detailed review of the various Energy Performance Certificates (EPC) that are used in European countries. They analyzed how the calculation methods and the technical required information are different, as well as the objectives of the EPC: refurbishment, information, etc.

Less studies are available on the US Energy Star certification. Arjunan et al. (2020) proposed ways to improve the accuracy by adding some models to the existing calculation method. Most of the work on Energy Star-Portfolio Manager is focused on comparing it to other domestic tools. In 2015, Pacific North National Laboratory developed an overview of the 44 energy data tools available in the USA and assessed if they provided benchmarking, data display, and various levels of energy analysis (Henderson & Fowler, 2018).
This study includes an asset-rating tool, called Asset Score, which was launched in 2015. The Asset Score complements the operational rating ES score with an asset rating that takes into consideration the types of systems (Office of Energy Efficiency and Renewable Energy, n.d.a). The Asset Score provides a targeted energy efficiency evaluation of a building’s envelope (its roof, walls and windows) and its major energy-related systems (such as lighting, hot water, and HVAC systems). These physical and structural elements have a significant impact on how efficiently energy is used within a building regardless of how the building is operated or the behavior of its occupants. The Asset Score normalizes for operational and occupancy variables and does not require users to collect energy consumption data. In that way, it is more similar to the French DPE. Yet, our study focuses on analyzing the buildings rating tools that were officially chosen by the respective governments.

Aim and scope

In this paper, we investigate the two approaches chosen by France and the USA to estimate the energy performance of their buildings. If the French certificate is close to other European certificates, the USA one is significantly different. Our aims are the following:

- To analyze the calculation methods and quantify the impact of the factors on the response for each method in the case of a multifamily building,
- To discuss the consequences of these approaches on the accuracy of the results, and their efficiency in terms of reducing energy consumption.

To do so, we use a Design of Experiments (DoE) that quantifies the impact of the input variables on the output (energy consumption or score) for each energy performance evaluation tools. The DoE is presented in “Methods” section. Such design is a series of tests in which purposeful changes are made to the input variables of a system or process, and the effects on the output variables are measured (Telford, 2007).

“Presentation of the French Diagnostics de Performance Energetique (DPE) calculation method” and “Design of experiments for the US Energy Star score” sections present the American ES score and the results of the DoE for the ES score. “Discussion on the two methods” section discusses the two tools based on the conducted design of experiments, and expands the scope to the efficiency of the tools.

Both tools offer similar calculations of greenhouse gases (GHG). This study only focuses on the building energy. The type of building chosen to conduct the study is a multifamily housing. As the two tools do not use the same units (kWh or kBtu, m² or ft²) and in order to simplify our analysis, the results of this paper are given in the International System of Units.

Method

Designs of Experiments (DoE) identify the most influencing factors on the efficiency of a process and enable finding improvement levers.

In our work, we conduct two Designs of Experiments, one for the French diagnostics DPE and one for the US Energy Star, in order to determine the quantitative influence of building’s variables on the final response of each method.

The DoE gives the impact of the input variables (called factors) on the output (called the response) of a process. A low and high value (called levels and coded as (− 1) and (+ 1), respectively) is defined for each factor (Bonte et al., 2014). A general recommendation for choosing the factors ranges is to set the levels far enough apart so that one would expect to see a difference in the response, but not so far apart as to be out of the likely operating range (Lundstedt et al., 1998).

A full factorial design contains all possible combinations of low and high levels for all the factors. This type of experimental design has 2^n numbers of experiments; 2 being the number of levels assigned, n the number of factors.

The Yates’ algorithm exploits the special structure of these designs to estimate the Main Effect (ME) of one factor on the response when its level changes from (− 1) to (+ 1), ignoring the effects of all other inputs (Yates, 1937). To do so, a matrix called “design matrix” is built. This design matrix collects all the experiments that are performed. Each experiment is characterized by a unique combination of n
factors on low level and high level. The design matrix has \((n + r)\) columns (where \(n\) is the number of factors and \(r\) the number of responses), and \(2n\) rows as factors vary on two levels. The \(n\) first columns are filled with combinations of \(-1\) and \(+1\). Yates’s algorithm needs the combinations to be in the Yates order, which is defined by:

- First row is filled with \(-1\),
- \(-1\) and \(+1\) values switch on the rows \(2n - 1\) in column \(n\).

The responses for each combination of factors are written in columns \(n + k\) \((1 \leq k \leq r)\). Once the design matrix is established, the main effects of the factors are evaluated with the Yates algorithm (Yates, 1937) with:

\[
\text{ME}_{\text{factor}} = \frac{1}{2n-1} \sum_{n} (\text{level}_{\text{factor}} \times \text{response})
\]  

(1)

In each DoE, the main effects \(\text{ME}\) are calculated for each factor, quantifying their impact on the response.

### Presentation of the French Diagnostics de Performance Energetique (DPE) calculation method

#### Legislative context

The European Union established the first version of the Energy Performance of Buildings Directive in 2002, and revised it in 2010 and 2018 (Official Journal of the European Union, 2018). Each European country was mandated to develop a national law requiring the evaluation of the energy consumption and greenhouse gas emission of buildings, in order to promote energy efficiency and boost building retrofits.

France developed the Diagnostic de Performance Energétique (DPE), which is an evaluation of the energy consumption of a building and its greenhouse gas emission. The DPE gives a label based on these values. It follows calculation methods that have been defined in a decree (Legifrance, 2021b). Conducting a DPE in France has been mandatory since 2006 in any building for sell, and since 2007 in any building for rent. The DPE is valid for 10 years.

DPE are conducted by technicians who must be certified by an accredited certification body, to prove that they have the knowledge, experience, and skills to perform DPE in buildings. These technicians are accredited for 5 years.

#### 3CL-DPE method

The decree mandates the establishment of a DPE for all existing buildings that have a heating system. The method used to calculate the building’s energy consumption is the Consumptions Calculation method (called 3CL for Calcul de la Consommation Conventionelle des Logements).

The 3CL-DPE calculation method gives an estimation of the energy consumption, based on the average use of a building under its climatic conditions. The method is based on assumptions of a conventional scenario of occupancy and occupants’ behavior. The energy consumption calculated with this conventional method may be different from the actual energy consumption, due for example to the occupants’ behaviors and their actions on the building. The main assumptions are:

- If there is a heating system (other than a mobile system or open fireplace), all the dwelling’s surface is considered as a heated area during the heating period.
- The energy consumption required to heat a building is calculated with the average values of the building location’s heating degree hours (HDH). HDH are a measure of how much (in °C), and for how long (in hours), the outside air temperature is lower than a specific “base temperature”. The base temperature used in the method is 19 °C, which means that the building’s heating system can maintain a temperature of 19 °C in the building. A standard inoccupation period of 1 week in December is considered, as well as a set point temperature of 16 °C during the day, on weekdays.
- Hot water demands are fixed according to the dwelling’s surface (through the number of occupants) and its location. Indeed, the temperature of the water entering the building depends on the location. Cold regions have a lower water temperature and therefore a higher energy load to heat it.
An average value of water demand of 56 L per day and per person is considered.

- If the building has an AC system, then the cooling degree hours are calculated for an outside temperature higher than 28°C.

Five usages are considered for the DPE. They are the heating, domestic water heating, cooling, lighting, and auxiliaries (i.e., the equipment for energy production and distribution, such as pumps or fans). Among the 5 usages, heating is the most energy consuming.

The key takeaway is that the energy demands of the 5 usages are calculated first, then transformed into energy consumptions taking into account the energy efficiency of the systems. Finally, the total energy consumption is calculated by summing up the 5 consumptions of each usage and transformed into source energy based on the type of energy used. The energy consumption is divided by the net floor area and expressed in kWh Source Energy per m² per year.

The calculation of the building’s energy consumption is incremental. The method is detailed in (Legifrance, 2021b). This document of 147 pages gives all the steps for the calculation, as well as the discrete values of parameters to choose in tables, such as the materials characteristics, energy systems efficiencies, air exchange rate, based on the type of system and its age.

For each building, a report is provided which includes the following:

- The energy consumption expressed in kWh Source Energy/m²·year or kWhSE/(m²·year) and the greenhouse gas emission (in kg CO₂/(m²·year)), classified on two sliding scales running from A to G (Fig. 1),
- The annual energy expenses in euros,
- Factors representative of summer comfort for passive cooling; of the building envelope performance; and of the ventilation system performance,
- The expert’s recommendations to improve the efficiency of the building.

Calculation in practice

The 3CL-DPE method calculates the energy consumption of the building, including heating, cooling, domestic water heating, lighting, and auxiliaries, as a function of a very large amount of data. To make the calculation of the energy consumption easier, the French government certified several software (RT-Bâtiment, 2022). In this study, we have selected the software “LICIEL V4” developed by a French company called “LICIEL Environnement” because of its user interface, the accuracy of the results, and the clarity of the display of the results.

![Energy-efficient building](image1)

![Low GHG emissions](image2)

![Energy-intensive building](image3)

![High GHG emissions](image4)

**Fig. 1** French DPE’s sliding scales running from A (best) to G (worst) for energy source consumption (left) and greenhouse gas emissions (right)
The calculation of the heating consumption is the longest and most complicated to perform among the 5 usages. As an example, the heating demand depends on a large variety of parameters:

- The heat losses through the building envelop (called GV): all walls and roofs must be considered. This is the longest part of the calculation. To point out the parameters involved, the calculation of GV is given in the Annex,
- The heat transfer due to air exchange,
- The energy consumption of the ventilation auxiliaries,
- The solar gains,
- The inertia of the building,
- The intermittence (possibility of regulation).

The heating consumption is calculated from the heating demand by taking into account the power efficiency of the heating system.

To give an idea of the complexity of the calculation of the heating consumption, the decree dedicates 61 pages to describe it, whereas the calculation of the cooling consumption takes 3 pages, the water consumption 5 pages, the auxiliaries 5 pages, and the lighting 1 page.

The French diagnostics DPE also provides the GHG emission. The calculation method can be found in Legifrance, 2021b. It is not presented in this article as we only focus on the building energy.

Design of experiments for the French DPE

In this part of the study, we aim to estimate the influence of the different inputs required by the French diagnostics DPE on the output, i.e. the building energy consumption, calculated with the 3CL-DPE method.

Factors and response of the design of experiments for the French diagnostics DPE

As presented above, the calculation method for the energy consumption is a long process detailed in Legifrance, 2021b. A lot of parameters are considered. To conduct our design of experiments, considering all the parameters in the calculation would have been prohibitively complex. Moreover, their importance in the calculation is unequal. Therefore, we have defined 5 main factors that directly impact the energy consumption, some covering several parameters. They are the following:

- Net floor area (NFA)
- Envelope of the building (ENV): this factor characterizes the thermal performance of the envelope. The factor ENV covers a lot of parameters that intervene in the calculation of the thermal heat loss GV of the building (cf. Annex), such as the surfaces of the walls, roof, and floors, and their corresponding $U$ values.
- Ventilation system (VENT): this factor characterizes the type and efficiency of the ventilation system and aims to quantify its impact on the energy consumption. VENT covers several parameters linked to the air change, such as the intake and exhaust air flow.
- Heating and hot water systems (H-HW): this factor characterizes the type and efficiency of the heating and hot water systems. It covers all the parameters required for the calculation of the energy consumption, linked to the HVAC system.
- Heating degree days (HDD): this factor characterizes the climate. Calculations of the energy consumption are performed using the heating degree hours of each location, as per the decree (Legifrance, 2021b). Yet, to display the results we chose to characterize the climate using the heating degree days, to compare the four cities (two French and two US) investigated in the case study.

The response in the design of experiments is the total energy consumption expressed in kWhSE/(m$^2$.year). This value is linked to the scale A (best) to G (worst), as seen in Fig. 1.

The auxiliaries and lighting are calculated using the conventional method and are included in the total energy consumption. We chose not to consider them as factors in the Design of Experiments because their values are constrained by the calculation method and do not vary significantly. The cooling system is not considered because we chose to investigate a multifamily building without an AC system, which is common in France.

The response in the design of experiments is the total energy consumption expressed in kWhSE/(m$^2$.year). This value is linked to the scale A (best) to G (worst), as seen in Fig. 1.

As we want to analyze the sensitivity of these 5 factors on the response, a DoE with $2^5=32$ experiments has been conducted. The experiments have
been implemented in the software LICIEL and the corresponding responses calculated.

Building settings for the design of experiments for the French diagnostics DPE

The building under investigation is a multifamily building, with several floors and apartments per floor. A baseline building has been defined, and from this reference, the low and high levels of the factors have been chosen. Table 1 lists the 5 factors and the values of their low and high levels and their corresponding responses. The factors’ low levels (−1) are characteristic of a poor energy performance multifamily building in France, whereas the factors’ high levels (+1) represent a high-energy performance building.

- **Net floor area (NFA)**
  The low level’s building has 5 floors and a net floor area (NFA) of 1308 m². For the high level, we chose to increase this value by 20% to 1572 m².

- **Envelope of the building (ENV)**
  The envelope (ENV) of the low level’s building is not well insulated (high U values in Table 1), whereas the high level’s building has a high thermal insulation performance.

- **Ventilation system (VENT)**
  The ventilation system has an impact on the heat loss due to air exchange. The decree provides equations to calculate the heat loss as a function of the type and age of the ventilation system, and of the orientation of the building due to the impact of the wind. The older the system, the higher the heat loss. Recent systems are more efficient leading to a lower impact on the heat loss, and therefore the heating consumption. The controlled mechanical ventilation system (VENT) defined for the low level is chosen to be old (before 1982 in the decree) and leads to high cold air flow to heat,

| Table 1 | Factors and responses of the design of experiments for the French DPE |
|---------|-----------------------------------------------------------------|
| Factors | Unit | Low level (−1) | High level (+1) |
| NFA     | m²   | 1308           | 1572            |
| ENV     | W/(m².K) | U_{wall}=1.17 | U_{wall}=0.33 |
|         |       | U_{ceil}=1.11  | U_{ceil}=0.53  |
|         |       | U_{floor}=2.00 | U_{floor}=0.69 |
|         | Walls: 20 cm concrete, 2 cm interior insulation | Walls: 37.5 cm terracotta bricks, 2 cm exterior insulation |
|         | Ceiling: under terrace, 2 cm interior insulation | Ceiling: under terrace, 6 cm exterior insulation |
|         | Floor: 4 cm concrete floor without insulation | Floor: 4 cm concrete floor with ext. insulation |
| VENT    | Controlled mechanical ventilation installed before 1982 | Controlled mechanical ventilation with bypass and heat exchanger |
| H-HW    | Heating system: individual wood stove | Heating and domestic hot water system: individual wall-mounted natural gas condensing boiler, with programmers and pilot burner |
|         | Domestic hot water heating system: electric water heater older than 5 years | Ducting: low temperature heating insulated ducts and located in the heated volume; regulated underfloor heating |
|         | Ducting: high temperature heating non-insulated ducts, located in the heated volume; heaters without thermostatic valves | |
| HDD     | 5459 (Mulhouse, France) | 3061 (Marseille, France) |
|         | (For information CDD = 536) | (For information CDD = 1306) |
| Response | Energy consumption kWhSE/(m².year) | 236 | 31 |
| DPE label | D | A |
whereas the high level has a heat recovery system. The choice of these systems leads to different energy consumptions due to air exchange, a high value for the low-level system ($\text{VENT} = -1$), and a low value for the high level ($\text{VENT} = +1$).

- **Heating and hot water systems (H-HW)**

  Similarly to the ventilation system, the decree provides equations to calculate the heating and hot water consumption as a function of the type of system and its age, impacting its efficiency and therefore its energy consumption. The low level’s heating system is chosen to be an individual wood stove, with a high-temperature water distribution and radiators; the hot water system is electrical and more than 5 years old. On the other hand, the high level’s heating system is a low-temperature gas condensing boiler, with an insulated distribution, and a temperature-controlled underfloor heating; the hot water system is an individual condensation boiler with a regulation system.

- **Annual heating degree days (HDD)**

  Regarding the climate, we chose cities that are representative of the French climate. Mulhouse (MUL) is the low level, the city is in the North-East of France and has a semi-continental climate with cold winters. For the high level, the building is located in Marseille (MAR), that has a Mediterranean climate with hot summers and mild winters (Peel et al., 2007).

Results of the design of experiments for the French diagnostics DPE

The main effects of the factors are presented in Fig. 2. The main effects being negative indicate that passing from the low level ($-1$) of the factors to the high level ($+1$) (cf. Table 1) results in decreasing the energy consumption, which is favorable.

The results of the experimental designs indicate that the heating degree days (HDD) have the most important effect on the energy consumption: the multifamily building in this analysis consumes in average 84.6 kWhSE/(m$^2$.year) less energy in Marseille (MAR) than in Mulhouse (MUL). The Cooling Degree Days (CDD) are not part of the DPE’s inputs, as the climate in France is not as warm and humid as in some regions in the USA, and most residential buildings are not air-conditioned. Nevertheless, if the building is cooled, the cooling energy consumption must be calculated. Also, the heating and domestic water systems (H-HW) have an important impact on the building energy consumption: changing poorly efficient systems to more efficient ones can decrease the energy consumption by 54 kWhSE/(m$^2$.year). Additionally, the improvement of the ventilation system (VENT) or the building envelope (ENV) can lead to a reduction of the energy consumption by 29.8 kWhSE/(m$^2$.year) and 26.9 kWhSE/(m$^2$.year), respectively. Finally, increasing the net floor area (NFA) does not have a big impact on the energy consumption: it is the less important factor of the five factors analyzed in this study.

![Fig. 2 Main effects of the 5 factors on the energy consumption for the French DPE](image-url)
Presentation of the US Energy Star calculation method

Energy Star Portfolio Manager (ES-PM) is an online tool that is used to measure and manage the energy and water consumption of almost any building type (US Environmental Protection Agency, *Energy Star, Use Portfolio Manager*, n.d.b). Users are required to enter information about one building or a whole portfolio of buildings. This information includes consumption data, cost information and operational use details. Then, the user can track key performance indicators such as energy use and costs, water use and costs, and greenhouse gas emissions. Portfolio Manager can thus help the user implement an energy management program for the categories of buildings that can currently receive an ES-score. It is important to highlight that the ES tool uses the building energy consumption as an indirect input through the utility bills and does not calculate it, while the French DPE calculates the building’s energy consumption. We need to keep it in mind when comparing the inputs of each method: the energy consumption plays a very important role in the ES score, but not in the DPE, as it is the output.

The ES score is based on the analysis of different surveys’ data depending on the building’s property type. For multifamily properties, the reference data used to establish the peer-building group is an industry survey (Fannie Mae’s Multifamily Energy and Water Market Research Survey) conducted by the Federal National Mortgage Association. In that survey, there were 350 observations that provided complete whole building energy data; the Environmental Protection Agency applied additional filters so the resulting number of properties in the data set is 322 (US Environmental Protection Agency, 2018). For example, one of the filters is to eliminate buildings with fewer than 20 units, because analysis could not model behavior for these buildings due to limited data. Thus, the ES score only applies to multifamily housing properties that are at least 20 units in size.

The Energy Star calculation method includes all the energy usages. The electricity bill is entered, reflecting the occupants’ actions, as well as the bills for all the other types of energy. This is a significant difference with the French method which does not take appliances into account.

For the multifamily housing analysis, the ES score is calculated in several steps:

- Calculation of the actual source energy use intensity (ASEUI)

\[
ASEUI = \frac{TSEU}{GFA} \tag{2}
\]

with

- \( ASEUI \) the actual source energy use intensity \([\text{kWhSE}/(\text{m}^2\cdot\text{year})]\), also called as the “dependent” variable (US Environmental Protection Agency, 2018)
- \( TSEU \) the building’s total source energy use \([\text{kWhSE/year}] \) is based on the actual monthly utility bills
- \( GFA \) the building’s gross floor area \([\text{m}^2]\)

- Calculation of the predicted source energy use intensity (PSEUI)

The predicted source energy use intensity (PSEUI) is the mean energy use intensity (EUI) for a hypothetical population of buildings that share the same values for each of some relevant variables (described below). That is the mean energy use for a building that operates just like the building that is being analyzed (US Environmental Protection Agency, 2018). It is calculated using a weighted ordinary least squares regression equation across the filtered data set of 322 observations, based on 5 so-called “independent” variables, as opposed to the dependent variable which is the Actual Source Energy Use Intensity seen above. The 5 independent variables are the following:

- Number of units per 1000 square feet,
- Number of bedrooms per unit,
- Low-rise building \([\text{yes} (=1) \text{ or no} (=0)]\). A building is low-rise if it has 4 stories or fewer,
- Heating degree days (HDD),
- Cooling degree days (CDD).

It is worth noting that 27 independent variables were actually analyzed, but only these 5 variables remain due to the lack of data (in most cases) (US Environmental Protection Agency, 2018).
Additionally, there is no variable about the building envelope, the HVAC systems or the occupancy of the building.

The predicted source energy use intensity (PSEUI) is obtained according to Eq. 3 below:

\[
PSEUI = C + \sum_{n=1}^{5} (x_n - \bar{x}_n) \cdot xcf_n
\]

with

- \(PSEUI\): Predicted source energy use intensity [kWhSE/(m².year)]
- \(C\): a constant equal to \(C = 130.7\) kWhSE/(m².year)
- \(x_n\): the 5 independent variables listed above
- \(\bar{x}_n\): the mean values of the 5 independent variables listed above
- \(xcf_n\): the coefficients of the regression equation for each variable

The values of the mean values \(\bar{x}_n\) and the coefficients \(xcf_n\) of the 5 independent variables are given in (US Environmental Protection Agency, 2018) and reported in Table 2.

- Calculation of the energy efficiency ratio (EERATIO)

\[
EERATIO = \frac{ASEUI}{PSEUI}
\]

A lower EERATIO indicates that the building is more efficient because it uses less energy than predicted. A lower EERATIO is thus related to a higher ES score.

- Determination of the Energy Star score via a lookup table

The EERATIO of the 322 observations were calculated and were used to create a cumulative distribution (US Environmental Protection Agency, 2018). The curve was fitted to the data using a two-parameter gamma distribution: a shape parameter \((\alpha = 15.13)\) and a scale parameter \((\beta = 0.06561)\). We created a lookup table (Table 3) to model the gamma function. That function gives the exact value of the inferior and superior EERATIO to get each score. For example, if the ratio of a building is between 0.5709 and 0.5941 (Table 3), the score is 97, which means that 97% of the survey’s buildings perform worse than that building. The calculation method of the ES score is identical for each property type, but with different input variables and different surveys (i.e., different regression equations and different gamma distributions). Also, it is important to remind that the method is based on the analysis of a survey data, and not on the data of buildings previously entered into Portfolio Manager.

### Design of experiments for the US Energy Star score

In this part of the study, our goal is to know which of the inputs of the Energy Star calculation method are the most influential on the final score.

Factors and response of the US Energy Star score

The factors chosen for the DoE are not the same as for the DPE, because the calculation methodology is different. Based on the calculation of the ES score

| Table 2 | Variables for Eq. (3) for multifamily housing (US Environmental Protection Agency, 2018) |
|---------|---------------------------------|
| n       | \(x_n\) | \(\bar{x}_n\) | \(xcf_n\) |
| 1       | Number of units /1000 ft²       | 1.215 | 48.01 |
| 2       | Number of bedrooms per unit     | 1.238 | 22.64 |
| 3       | Percentage of units located in low-rise buildings | 0.4867 | -19 |
| 4       | Heating Degree Days (HDD)       | 4.233 | 0.008989 |
| 5       | Cooling Degree Days (CDD)       | 1.364 | 0.01406 |

| Table 3 | Extract of the Energy Star score lookup table for multifamily housing (values of EERATIO giving a score from 95 to 100) |
|---------|-------------------------------------------------|
| Energy Star score | Cumulative percent | EERATIO |
| 100      | 0%      | 0.0000 | ≥ |
| 99       | 1%      | 0.4965 | 0.4965 |
| 98       | 2%      | 0.5412 | 0.5412 |
| 97       | 3%      | 0.5709 | 0.5709 |
| 96       | 4%      | 0.5941 | 0.5941 |
| 95       | 5%      | 0.6134 | 0.6134 |

 Springer
developed in the previous section, all the inputs have been used as factors in the DoE. Yet, among the 5 “independent” variables in Eq. (3), we have excluded the input “Low-Rise building” (LR). Indeed, we have conducted the DoE with a multi-storey building with more than 4 floors, leading to a null value of LR.

Finally, the 6 factors used for the design of experiments are the following:

- Gross floor area (GFA)
- Total source energy use (TSEU)
- Total number of units (UNIT)
- Total number of bedrooms (BED)
- Annual heating degree days (HDD)
- Annual cooling degree days (CDD)

The HDD and CDD are calculated in Portfolio Manager for a base temperature of 65°F (18 °C) for both heating and cooling. The CDD are a factor in this design of experiments. Indeed, the variety of climates in the USA, including hot and humid regions, requires considering this parameter. In France, there are no such areas, and very few multifamily buildings have air-conditioning, so CDD is not a relevant parameter for the DPE.

The response of the Design of Experiments is the ES score. Since there are 6 factors, $2^6 = 64$ experiments have been conducted and the corresponding responses calculated. We did not use the tool Portfolio Manager. The calculation method of the ES score being simpler than the DPE, we modeled Eqs. (2), (3), and (4), as well as the gamma function to calculate the score. The objective of computing the ES score instead of using the tool Portfolio Manager was to be able to control the factors HDD and CDD independently, which would not have been possible otherwise. Indeed, as HDD and CDD are factors in the DoE, their impact has to be studied independently.

### Building settings for the design of experiments for the US Energy Star score

The high and low levels assigned to each factor and the corresponding scores are indicated in Table 4. As there was insufficient accurate data for US multifamily buildings, we used the EPA’s guide for multifamily housing for reference, which gives an example of an American multifamily property (US Environmental Protection Agency, 2018). It is worth reminding that the calculation process of each method being different as described above, the inputs are different according to the certificate. Yet, all the inputs have been chosen with an identically sized multifamily building in mind.

- Gross floor area (GFA)

  We chose to study a building of the same size (same GFA and NFA) in the two DoE. The building investigated in the DoE of the French DPE has a net floor area of 1308 m² for the low level. A ratio of 1.25 has been chosen between the GFA and the NFA, leading to a GFA of 1635 m² for the low level. The high level is increased by 20% to reach 1962 m².

  The number of stories of the building sets the value of the independent variable low rise. The buildings studied are considered to have more than 4 stories, and therefore the variable low-rise building is set to NO, or 0. This variable is eliminated in Eq. (3).

| Table 4 | Factors and response of the design of experiments for the US Energy Star score |
|---------|--------------------------------------------------------------------------------|
| Factors | Unit | Low level (−1)                                                                 | High level (+1)                                                                 |
| GFA     | m²   | 1 635 (NFA = 1308 m²)                                                          | 1 962 (NFA = 1572 m²)                                                          |
| TSEU    | kWhSE | 1,003,810                                                                      | 836 508                                                                        |
| UNIT    |      | 42                                                                            | 51                                                                             |
| BED     |      | 62                                                                            | 74                                                                             |
| HDD     |      | 7 765 (Minneapolis, USA)                                                       | 3 238 (Atlanta, USA)                                                           |
| CDD     |      | 994 (Minneapolis, USA)                                                         | 1 940 (Atlanta, USA)                                                           |
| Response| ES score | 63                                                                           | 88                                                                             |
• Total source energy use (TSEU), number of units (UNIT) and number of bedrooms (BED)

The EPA guide’s building example has a gross floor area (GFA) equal to 120,000 ft² (i.e. 11,148 m²), 300 units and 440 bedrooms. By keeping these ratios, the low-level factors are set to:

Number of units UNIT = 42
Number of bedrooms BED = 62

The high level’s values of the factors UNIT and BED have been increased by 20% to 51 and 74, respectively.

We adjusted the value of TSEU to obtain for the low level building a score of 63, as in the example. To do so, the TSEU has been set to 3,425,000 kBtu, i.e., 1,003,810 kWh for the low level, and decreased by 20% for the high level.

• Heating degree days (HDD) and cooling degree days (CDD)

Regarding the climate conditions, we have chosen Minneapolis in the Midwest of the USA, for the location of the low-level’s building, as it has a temperate and humid continental climate with cold winters. The high-level’s building is located in Atlanta in the South, with a warm oceanic climate with warm summers (Peel et al., 2007).

Results of the design of experiments for the US Energy Star score

The factors’ main effects on the ES score are reported on Fig. 3. It is interesting to observe that some factors have a positive effect (i.e., they increase the score), and some have a negative one (i.e. they decrease the score). The climate through HDD and CDD and the energy consumption (TSEU) are the main impacting factors. The number of units (UNIT) and the gross floor area (GFA) have a close impact on the score. The number of bedrooms (BED) has the least influence on the score.

Unsurprisingly, a decrease in the energy consumption has a positive effect on the ES score because a building performs better if its energy consumption is lowered. For example, reducing the energy consumption by 20% (from the value of TSEU at the level −1 to the value at level +1, in Table 4) increases the score by 19 points. When the gross floor area (GFA) is higher, the score is also better. This can be easily explained with the EERATIO’s formula (Eq. 4): if GFA increases, ASEUI decreases (Eq. 2), and thus EERATIO decreases, leading to a higher score.

To understand how the other factors (HDD, CDD, UNIT, and BED) impact the score, it is important to remember that the regression equation Eq. (3) has been established so that some parameters do
not penalize buildings in “unfavorable situations” which could increase their energy consumption. These buildings are the ones that meet one or several of the following criteria (US Environmental Protection Agency, 2018):

- Number of units per 92.90 m² (equivalent to 1000 ft²) superior to 1.215 (x₁)
- Number of bedrooms per apartment superior to 1.238 (x₂)
- HDD > 4 233 (x₄)
- CDD > 1 364 (x₅)

Table 2 lists the values of the variables (xₙ and xcfₙ) used in Eq. (3). All these variables are positive, meaning that if the factors HDD, CDD, UNIT, and BED, are lower than the mean value xₙ, therefore PSEUI decreases, EERATIO increases, and the score decreases. On the other hand, if the factors are higher than the mean value, therefore it means it is an unfavorable situation, PSEUI increases, as the score.

As an example, the mean effect of HDD on the score is negative (Fig. 3). It means that passing from the low level (−1) to the high level (+1), corresponding to a decrease of the HDD (Table 4) leads to a decrease of the score by 20.8. In other words, the HDD of Minneapolis, which is higher than the mean value \( x₄ = 4.233 \), leads to an increase of the score, to not penalize a building located in a cold city. In the same way, the mean effect of the CDD on the score is positive, meaning that an increase of the CDD leads to an increase of the score. That method allows the comparison of the same types of building in the whole country, considering the wide range of climates.

Finally, the mean effects of the factors UNIT and BED on the score are also positive, as shown on Fig. 3. The denser and more occupied the building is, higher is the score.

It is also interesting to note that for the French DPE, the extreme cases (−1) and (+1) presented in Table 1 correspond to the worst and best values and labels among the 32 cases tested in the design of experiments. For the US ES score, 64 experiments have been computed and the worst and best situations do not occur when the factors are all set at their lowest and highest levels, respectively (Table 4). Indeed, due to the additional factor CDD, the combinations of HDD and CDD leads to test some cases that do not correspond to an existing location. Yet, in order to see the mathematical impact of each factor on the score, HDD and CDD have to be considered independently. Table 5 shows an extract of the design of experiments giving factors leading to the worst and best score. Thus, the worst score occurs for all the factors set at −1 except HDD at +1 (experiment n°17 among the 64 experiments computed), leading to a score equal to 32. That case corresponds to an energy-intensive building (high TSEU), located in an area with mild winter (low HDD) and hot summer (high CDD). Compared to peer buildings located in a similar climate, it has a lower score due to a too high TSEU for such a climate. The best-case scenario occurs for all the factors set at +1 except HDD at −1, leading to a high score of 97 (Table 5, Exp. 48). That corresponds to a high-performance building (low TSEU), located in an area with rigorous winter (high HDD) and hot summer (high CDD). The calculation method favors a building that performs well in stringent weather conditions.

| Exp. #  | GFA | TSEU | UNIT | BED | HDD | CDD | ES score |
|---------|-----|------|------|-----|-----|-----|----------|
| Exp. 17 | −1  | −1   | −1   | −1  | +1  | −1  | 32       |
| Exp. 48 | +1  | +1   | +1   | +1  | −1  | +1  | 97       |
| Exp. 1  | −1  | −1   | −1   | −1  | −1  | −1  | 63       |
| Exp. 49 | −1  | −1   | −1   | −1  | +1  | +1  | 43       |
Discussion on the two methods

The French DPE and the US ES score are the building certifications chosen in these countries to label the energy performance of buildings. In this section, we discuss some consequences of their different approaches.

Significant differences in the approaches

The French DPE is similar to the majority of other European certificates (Taranu & Verbeeck, 2018), emphasizing the performance of the building with a complex calculation of the heat transfer through the buildings features (envelope, of the HVAC systems, ventilation, etc.). The building’s energy consumption is calculated for an average usage, and a letter grade is given to help building owners know how their buildings perform compared to other French buildings. In the DPE report, some improvements are also suggested to make the building more energy efficient. The DPE’s label depends on the building’s features and its geographical location, but is not dependent on the actual occupants’ behavior or even occupancy. Consequently, unless retrofitting is being undergone, the DPE is not subject to change over years. The DPE is mandatory for any building transaction and has a validity of 10 years.

The US Energy Star score lays on a different approach by considering the actual utility bill of the building and by comparing the building to its peers. Each building receives a 1–100 score which indicates to building owners how their facilities perform compared to a population of buildings of the same type. Thus, a score for a same building would change with a different technical management of the building or different occupancy. A striking example of this possible variability has been shown during the COVID-19 pandemic lockdown in 2020, where the scores of some buildings such as hotels have been significantly increased due to the decrease of their utility bills. The US ES score is voluntary and should be done every year.

Different impact of the climate on the response

The impact of the climate (through the HDD and CDD) on the score is different according to the method. In France, a building located in a cold-winter region (especially in the North-East) receives a lower DPE’s label than the same building located in a warmer region because it uses more energy for heating. Table 6 shows the energy consumptions of an energy-intensive building (all factors set at −1) located in Mulhouse (Exp. 1 where HDD = −1) and in Marseille (Exp. 2, where HDD = +1). The building consumes 236 kWhSE/(m²·year) leading to a D label in Mulhouse, and 107 kWhSE/(m²·year) leading to a B label in Marseille. From a mathematical point of view, the DoE emphasizes the strong impact of the climate in the calculation method.

The French National Observatory for Energy Retrofitting issued a report on the building stock based on the DPE results (Observatoire National de la Rénovation Energétique, 2020). Among other studies, the dependency of the DPE on the geographical zone has been investigated. The study shows the proportion of high energy-consuming buildings labeled F and G as a function of the geography. It appears that the proportion of such buildings is higher in the cold areas of the east of France (where Mulhouse is located) and in the mountains than on the Mediterranean arc (Marseille): 30% of the buildings are labeled F or G in cold areas whereas 5% are in warm areas. This key takeaway shows that the DPE is more sensitive to the climate than the ES score.

Rating buildings energy consumption on the same scale is possible in France because the climate range is not as wide as in the USA. In the USA, there is no other choice than comparing buildings located in the same climate; otherwise, the difference of energy consumption would be too extreme. Indeed, a building located in Atlanta consumes in average less energy in winter than a similar building in Minneapolis. That is the reason why the US Environmental Protection Agency chose to design the Energy Star’s score

Table 6 Select cases of the French DPE (extract from the design of experiments for the DPE)

| Exp. # | NFA | ENV | VENT | H-HW | HDD | Energy consumption (kWhSE/(m²·year)) |
|--------|-----|-----|------|------|-----|------------------------------------|
| Exp. 1 | −1  | −1  | −1   | −1   | −1  | 236                                |
| Exp. 2 | −1  | −1  | −1   | −1   | +1  | 107                                |
method by comparing the buildings with their peers in similar weather conditions. That method helps some buildings not to receive a low ES score due to unfavorable conditions (such as buildings located in very cold cities in winter and/or very warm cities in summer, high and/or dense buildings) that increase their Actual Source Energy Use Intensity (ASEUI). Thanks to the equation regression coefficients, the calculated Predicted Source EUI (PSEUI) of these buildings increases, and this leads to an increase of the score. For example, Table 5 shows that the same energy-intensive building (high TSEU) located in Minneapolis and scoring 63 (Exp. 1), receives a score of only 43 in Atlanta (Exp. 49). In these two experiments, only the location, through the pair (HDD and CDD), changes. The score of this energy-intensive building is degraded in a favorable climate. In a large country like the USA, this is a good way to compare similar categories of buildings that are located in very different climates and with different characteristics.

Impact of the technicians on the European EPCs

We have conducted a study to quantify the impact of physical factors (such as NFA, climate, HVAC systems) on the response. It is worth mentioning that some studies investigated the impact of the technicians who perform the certificates. The ES score is not impacted by the person who conducts the calculation, as the inputs are much less numerous and more easily accessible. In contrast, the European certificates, such as the French DPE, require a lot more data that are not easily knowable. For example, the composition of the envelope, with all its layers, is not easy to obtain due to the lack of information; to the way it was built at the construction stage; and to the way it ages and impacts the thermal properties. Therefore, the choice made by the technicians has an impact on the score, and the evaluation of a same building by different technicians can lead to different scores.

Tronchin and Fabbri (2012) investigated the impact of these technicians’ analyses on the results of the Italian energy performance certificate, equivalent to the French DPE. In a study conducted in Ireland, Schuitema et al. (2020) showed that the trust in the assessors who conducted the Irish energy performance certificates has an important impact on the people’s responses to the certificates.

Limitations of the tools

Case of the DPE

The main limitation of the DPE, as well as of all the European Energy Performance is the accuracy of the calculated energy consumption. Several studies have been conducted in Europe on the difference between the actual energy consumption and the calculated one using Energy Performance Certificates (Sunikka-Blank & Galvin, 2012; de Wilde, 2014; Burman et al., 2014). This difference is called the performance gap. There is a broad consensus in the literature that the two main causes of the performance gap are the “prebound effect” and “rebound effect” (Sunikka-Blank & Galvin, 2012). As van der Bent et al. (2021) summarized: “the prebound effect means a lower energy consumption than theoretically assumed in buildings with a poor energy performance because inhabitants do not heat the whole dwelling. The rebound effect means that dwellings with a high energy performance use more energy than theoretically assumed, because inhabitants think that the dwelling is energy efficient”.

Cozza et al. (2020a) discuss the many factors that could explain the gap due to the calculation method. Their review identifies three main inaccuracies: the expected indoor air temperature, the $U$ values assumed for building façade elements, and the expected air change. Even though this list has been established over different studies in several countries with different energy performance certificates, inaccuracies in these data appear to overestimate the energy consumption, especially in energy-intensive buildings (Sunikka-Blank & Galvin, 2012). Our Design of Experiments conducted for the DPE confirms the impact on the calculated energy consumption of the envelope characteristics and the air change, through the factors ENV and VENT, and quantifies how much a variation in the input affects the result.

Case of the ES score

On the other hand, the USA score considers the actual energy consumption, so the real operation of the building, with its real occupancy. In this sense, it is more accurate. The energy performance of the building itself is indirectly evaluated, as the operation and occupants’ actions have a significant impact on the
Total Source Energy Use (TSEU). The ES score is not mandatory, which is a limitation to initiate awareness and energy savings.

Impact of the tools to reduce energy consumption

**Case of the DPE**

In a study conducted in Sweden, Von Platten et al. (2019) investigated the evolution over time of the Swedish EPC, close to the DPE. They showed that the second generation of EPC conducted around 2018, 10 years after the first one, has an overall lower energy consumption than the first one, regardless the renovation degree.

Yet, the impact of the European EPC on the decrease of the energy consumption is not easy to quantify. One of the criteria to estimate the impact depends on how the savings are calculated. Indeed, the performance gap described above leads to another gap between actual and predicted energy savings after energy renovations (van der Brom et al., 2019). Filippidou et al. (2019) have conducted a study in the Netherlands, where the EPC is close to the French DPE. They have shown that the effectiveness of energy renovations is lower when the renovation is undergone based on the EPC rather than on the actual consumption. Cozza et al. (2020b) reached similar conclusions in a study conducted in Switzerland.

In France, the “Climate and resilience” law, adopted in 2021, aims to mitigate climate change across different sectors (Legifrance, 2021a). A third of the law’s 305 articles target the housing sector, and the law has adopted the DPE as the performance metrics to guide the retrofitting policies. For instance, buildings labeled F and G according to the DPE are defined as very poorly insulated buildings, and incentive and mandatory measures have been launched to eradicate them.

**Case of the ES score**

Studies like the European ones on the accuracy of the score calculation have not been performed in the USA as the score is based on the actual consumption. Yet, the policy impact is investigated through building benchmarking. Benchmarking is the tracking of the energy use of a building, to compare it to its peers, established norms, or its past performance, with the goal of informing and motivating performance improvement (Office of Energy Efficiency and Renewable Energy, n.d.b). The Environmental Protection Agency reported that the 35,000 buildings benchmarked over the period from 2008 to 2011 saved 7% of energy and had a 6-point increase of their ES scores (US Environmental Protection Agency, 2012). However, Palmer and Walls (2015) pointed out the difficulty to evaluate accurately the effectiveness of these benchmarking policies, the energy use in buildings susceptible to be affected by confounding factors such as the economy, local markets, or the energy sector.

Currently, 27 US cities have implemented building benchmarking ordinances (Better buildings–US Department of Energy, 2019). These ordinances require the buildings’ owners to measure and disclose their energy use information. ES Portfolio Manager has been chosen as the tool to collect these data. Meng et al. (2017) investigated the specific case of New York City, where a benchmarking and disclosure ordinance for commercial and multifamily buildings was implemented in 2009. They found that this policy saved energy among individual buildings in NYC by about 6% three years after implementation, and by 14% 4 years after the policy took effect. They also found that disclosure of Energy Star scores is a significant factor in this reduction.

On a further step of benchmarking energy data, some cities and local governments implement Building Performance Standards (BPS) (Institute for Market Transformation, 2021). These emerging policies, adopted locally, require building owners to meet performance targets by actively improving their buildings over time. The performance metrics used is the Energy Star score. They are applied to existing commercial and multifamily buildings (US Environmental Protection Agency, 2021).

**Improvements made to the tools over time**

**Improvements made to the DPE**

The DPE’s method has been criticized when first introduced and has been improved several times, in 2012 and 2021. Conventional Consumptions Calculation method (3CL-DPE) presented in this article was initially used for some residential buildings: houses, apartments, or multifamily housing, built after 1948...
and with individual heating systems. Other buildings (commercial buildings, public buildings, etc.) were evaluated according to the average energy consumption of the last three years’ bills, which was converted in Source Energy. This evaluation has been removed in the new version of the DPE that was released in July 2021. The 3CL method is therefore the only remaining one to establish the DPE.

Recently, Taranu et al. (2020) used behavioral sciences to propose an improved version of the Flamish Energy Performance Certificate, close to the French DPE. Their objective was to help the owners understanding the metrics of the EPC, and the ensuing actions in renovation they could lead. On a wider scale, as all the member states of the EU have implemented their EPC for approximately 10 years, the European EPC database is now substantial and can be used at urban scale for energy planning and building renovation planning (Li et al., 2019).

**Improvements made to the ES score**

In the USA, Scofield (2014) focused on the evolution of the ES score’s calculation method, and especially on what changed after its revision in 2007. According to his analysis, ES scores’ calculation methods and the data on which they are based need to be improved. His results demonstrate that ES scores are found to be uncertain by ±35 points. He concluded in his study that the Environmental Protection Agency should reuse the calculation method with the logarithm \( \ln(EUI) \) like in the previous calculation method in 2007, as it would help to get more accurate scores. Hsu (2014) also concludes that there is a wide uncertainty in the score calculations and suggests using the buildings data collected with Portfolio Manager to improve the methods to calculate the score, instead of the survey data from the Environmental Protection Agency.

EPA significantly updated performance metrics for commercial buildings when the most recent Commercial Buildings Energy Consumption Survey (CBECS) data became available and was rolled out in the fall 2018 (US Environmental Protection Agency, 2022). So far, the calculation method for multifamily housing has not been modified. While the coefficients of the gamma function (used to calculate the score as a function of EERATIO) have slightly changed in 2019, the resulting scores of the present study (made with the current coefficients) are varying by 1 point or less comparing the former and new coefficients.

**Limitations of our case study**

**Choice of the factors in the designs of experiments**

The design of experiments that we have conducted quantifies the impact of the factors on the response (the energy consumption or the ES score). In the case of the DPE, the complexity of the calculation method imposes the definition of factors that cover several inputs. Our results are therefore a function of the choice of factors we made. In addition, some factors can be linked in practice. For example, the buildings’ envelopes are adapted to the climate and lead to a lower energy consumption. The choice of factors has a lesser impact on the ES score for which we considered all the inputs as factors.

**Values of the factors**

Our study depends on the low and high values of the factors that we defined. Our goal was to investigate a low and a high energy efficient building, but a large set of different features could have been chosen to define the low and high levels. The absolute results of the design of experiments are strongly correlated to the values of the factors we have chosen. What we wanted to emphasize was not the values of the responses, but rather the relative impact of the factors on the response.

**Conclusion**

Our analysis shows the strong difference of approach in labeling the energy consumption of the building stock, in France and in the USA. The French diagnostics DPE, similar to other European certificates, computes the energy consumption, based on the features of the building for a standard occupancy and use of the building. The DPE gives a score that indicates whether the building is energy intensive. The US ES score is based on the actual total energy consumption of the occupied building, and it indicates how the building performs compared to its peers.
Even if the approaches are different, we determined for each of them the main impacting factors. To do so, we conducted a case study of a multifamily housing, using a design of experiments to quantify the impact of the factors on the response. Regarding the complexity of the French DPE and the quantity of data required, we included several parameters into a factor. For example, the factor ENV characterizing the whole envelope of the building actually depends on several data such as the façades surfaces and $U$ values. The results of the DoE show that the most impacting factors for the DPE are the climate through the HDD, followed by the heating and hot water systems H-HW, the ventilation and the envelope.

To conduct the DoE of the US ES score, we used all the factors required for the calculation, since the calculation method is simpler. We quantified the impact of the factors and found out that the main influencing factors are the HDD and the Total Source Energy Use TSEU. Through the factor TSEU, the score is extremely linked to the occupants’ behavior and how the building is operated. Regarding the factors HDD and CDD, the DoE also shows how the equation significantly modifies the score to allow locations with extreme weather to obtain score that can compete with buildings in milder climate.

The discussion underlines the accuracy of the ES score method compared to the DPE method, which shows a variability in the energy consumption calculated, due the complexity of the calculation and the very large amount of data needed. Regular updates in the DPE calculation method have been implemented over time to improve the estimations’ accuracy, the last major update being performed in 2021. In the same way, the regression equation used to compare the building tested with its peers in the ES score has been modified several times, with a last update in 2018.

At the building scale, the purpose of these EPC, both in France and the USA, is to raise awareness about the energy consumption and to encourage the building’s owners to retrofit. These certificates lead to more and more incentive policies: in France, the buildings labeled F and G could not be rented as of 2025, requiring owners to conduct retrofit; in the USA, the public benchmarking of the ES scores has proven its efficacy in decreasing the energy consumption of concerned buildings over time.

At a larger scale, both tools have become the metrics guiding buildings retrofitting policy. Despite the limitation in terms of results variability, the longitudinal studies of the DPE provide insights into the selection of best practices and implementation of energy conservation measures. In the USA, the ES scores are used as performance metrics to achieve mandatory performance targets, following the building performance standards.

**Authors’ contributions** The contribution of the authors is as follows: Berangere Lartigue: literature review, conceptualization, methodology, software, data analysis, formal analysis, writing—original draft and revised version. Laura Biewescht: literature review, conceptualization, methodology, software, data analysis, formal analysis, writing—original draft, review. Flore Marion: literature review, conceptualization, methodology, data analysis, formal analysis, writing—original draft, review. Erica Cochran Hameen and Françoise Thellier: supervision, writing—original draft, review. All authors read and approved the final manuscript.

**Declarations**

**Competing interests** The authors declare that they have no competing interests.

**Annex**

Calculation method for the thermal heat loss GV of the building:

\[
GV = DP_{\text{walls}} + DP_{\text{roofs}} + DP_{\text{floors}} + DP_{\text{windows}} + DP_{\text{doors}} + PT + DR
\]

where

\[
DP_{\text{walls}} = \sum_{i=1}^{n} (b_i \times S_{\text{wall}} \times U_{\text{wall}})
\]

is the heat loss of the $n$ walls, in W/K

\[
DP_{\text{roofs}} = \sum_{i=1}^{n} (b_i \times S_{\text{roof}} \times U_{\text{roof}})
\]

is the heat loss of the $n$ roofs, in W/K

\[
DP_{\text{floors}} = \sum_{i=1}^{n} (b_i \times S_{\text{floor}} \times U_{\text{floor}})
\]

is the heat loss of the $n$ floors, in W/K

\[
DP_{\text{windows}} = \sum_{i=1}^{n} (b_i \times S_{\text{window}} \times U_{\text{window}})
\]

is the heat loss of the $n$ windows, in W/K

\[
DP_{\text{doors}} = \sum_{i=1}^{n} (b_i \times S_{\text{door}} \times U_{\text{door}})
\]

is the heat loss of the $n$ doors, in W/K
$b_i$ are the heat loss reduction's coefficients of the walls, roofs, floors, windows or doors, dimensionless (Legifrance, 2021b)

$S_i$ are the surfaces of each wall, roof, floor, window or door, in m²

$U_i$ are the coefficients of heat transmission of the walls, roof, floor, windows or doors in W/(m².K)

PT is the heat loss due to the thermal bridges, in W/K

DR is the heat loss due to air exchange, in W/K.

References

Agence de l’Environnement et de la Maîtrise de l’Energie. (2022). Diagnostic de Performance Énergétique. Retrieved March 3, 2022, from https://www.observatoire-dpe.fr/index.php/statistique. Accessed 29 Apr 2022.

Arjunan, P., Poolla, K., & Miller, C. (2020). EnergyStar++: Towards more accurate and explanatory building energy benchmarking. *Applied Energy*, 276, 115413. https://doi.org/10.1016/j.apenergy.2020.115413

Better buildings – US Department of Energy. (2019). Benchmarking and transparency: resources for state and local leaders. Retrieved March 3, 2022, from https://betterbuildingsolutioncenter.energy.gov/sites/default/files/attachments/Benchmarking_Transparency_Resource_PDF_Final_2.14.pdf. Accessed 29 Apr 2022.

Bonte, M., Thellier, F., & Lartigue, B. (2014). Impact of occupant’s actions on energy building. *Energy and Buildings*, 76, 219–227. https://doi.org/10.1016/j.enbuild.2014.02.068

Burman, E., Mumovic, D., & Kimpian, J. (2014). Towards measurement and verification of energy performance under the framework of the European directive for energy performance of buildings. *Energy*, 77, 153–163. https://doi.org/10.1016/j.energy.2014.05.102

Cozza, S., Chambers, J., & Patel, M. K. (2020a). Measuring the thermal energy performance gap of labelled residential buildings in Switzerland. *Energy Policy*, 137, 111085. https://doi.org/10.1016/j.enpol.2019.111085

Cozza, S., Chambers, J., Deb, C., Scartezzini, J.-L., Schlüter, A., & Patel, M. K. (2020b). Do energy performance certificates allow reliable predictions of actual energy consumption and savings? Learning from the Swiss national database. *Energy and Buildings*, 224, 110235. https://doi.org/10.1016/j.enbuild.2020.110235

De Wilde, P. (2014). The gap between predicted and measured energy performance of buildings: A framework for investigation. *Automation in Construction*, 41, 40–49. https://doi.org/10.1016/j.autcon.2014.02.009

Economidou, M., Todeschi, V., Bertoldi, P., D’Agostino, D., Zangheri, P., & Castellazzi, L. (2020). Review of 50 years of EU energy efficiency policies for buildings. *Energy and Buildings*, 225, 110322. https://doi.org/10.1016/j.enbuild.2020.110322

European Commission. (2021). Energy performance of buildings directive, facts and figures. Retrieved April 2, 2021, from https://ec.europa.eu/energy/topics/energy-efficiency/energy-efficient-buildings/energy-performance-buildings-directive_en#facts-and-figures. Accessed 29 Apr 2022.

Filippidou, F., Nieboer, N., & Visscher, H. (2019). Effectiveness of energy renovations: A reassessment based on actual consumption savings. *Energy Efficiency*, 12, 19–35. https://doi.org/10.1007/s12053-018-9634-8

Goldstein, D. B., & Eley, C. (2014). A classification of building energy performance indices. *Energy Efficiency*, 7, 353–375. https://doi.org/10.1007/s12053-013-9248-0

Henderson, J. W., & Fowler, K. M. (2015). Federal Metering Data Analysis Needs and Existing Tools (Report No. 24191). Pacific North National Laboratory, US Department of Energy. Retrieved March 3, 2022, from https://www.pnl.gov/main/publications/external/technical_reportts/PNNL-24191.pdf. Accessed 29 Apr 2022.

Hsu, D. (2014). Improving energy benchmarking with self-reported data. *Building Research & Information*, 42(5), 64–656. https://doi.org/10.1080/09613218.2014.887612

Institute for Market Transformation. (2021). The growing impact of building performance standard. Retrieved March 3, 2022, from https://www.imt.org/the-growing-impact-of-building-performance-standards/. Accessed 29 Apr 2022.

Legifrance. (2021a). Loi n° 2021a-1104 du 22 août 2021a portant lutte contre le dérèglement climatique et renforcement de la résilience face à ses effets. Retrieved March 3, 2022, from https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000043956924. Accessed 29 Apr 2022.

Legifrance. (2021b). Arrêté du 8 octobre 2021b modifiant la méthode de calcul et les modalités d’établissement du certificat de performance énergétique. Retrieved March 3, 2022, from https://www.legifrance.gouv.fr/JORFT0043956924/docstart. Accessed 29 Apr 2022.

Li, Y., Kubicki, S., Guerriero, A., & Rezgui, Y. (2019). Review of building energy performance certification schemes towards future improvement. *Renewable and Sustainable Energy Reviews*, 113, 109244. https://doi.org/10.1016/j.rser.2019.109244

Lundstedt, S., Seifert, E., Abramo, L., Thelin, B., Nystrom, A., Pettersen, J., & Bergman, R. (1998). Experimental design and optimization. *Chemometrics and Intelligent Laboratory Systems*, 42(1–2), 3–40. https://doi.org/10.1016/S0169-7439(98)00065-3

Meng, T., Hsu, D., & Han, A. (2017). Estimating energy savings from benchmarking policies in New York City. *Energy*, 133, 415–423. https://doi.org/10.1016/j.energy.2017.05.148

Observatoire National de la Rénovation Énergétique. (2020). Le parc de logements par classe de consommation.
énergétique. Retrieved March 5, 2022, from https://www.statistiques.developpement-durable.gouv.fr/sites/default/files/2021-12/document_travail_49_parc_logements_consommation_energie_septembre2020.pdf (in French). Accessed 29 Apr 2022.

Office of Energy Efficiency & Renewable Energy. (n.d.a). Building energy asset score. Retrieved March 3, 2022, from https://www.energy.gov/eere/buildings/building-energy-asset-score. Accessed 29 Apr 2022.

Office of Energy Efficiency & Renewable Energy. (n.d.b). Building energy use benchmarking. Retrieved March 3, 2022, from https://www.energy.gov/eere/slc/building-energy-use-benchmarking. Accessed 29 Apr 2022.

Official Journal of the European Union. (2018). Directive 2018/844 of the European Parliament and of the Council of 30 May 2018 on the energy performance of buildings. Retrieved March 5, 2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2018.156.01.0075.01.ENG. Accessed 29 Apr 2022.

Palmer, K.L., & Walls M. (2015). Can benchmarking and disclosure laws provide incentives for energy efficiency improvements in buildings? Resources for the Future, Discussion paper 15-09. Retrieved October 13, 2021, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2564251. Accessed 29 Apr 2022.

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Koppen-Geiger climate classification. Hydrology and Earth System Sciences, 11, 1633–1644. https://doi.org/10.5194/hess-11-1633-2007

Pérez-Lombard, L., Ortiz, J., González, R., & Maestre, I. R. (2009). A review of benchmarking, rating and labelling concepts within the framework of building energy certification schemes. Energy and Buildings, 41(3), 272–278. https://doi.org/10.1016/j.enbuild.2008.10.004

RT-Bâtiment. (2022). Evaluation des logiciels. Retrieved March 3, 2022, from http://www.rt-batiment.fr/evaluation-des-logiciels-a50.html. Accessed 29 Apr 2022.

Schiutema, G., Aravena, C., & Denny, E. (2020). The psychology of energy efficiency labels: Trust, involvement, and attitudes towards energy performance certificates in Ireland. Energy Research & Social Science, 59, 101301. https://doi.org/10.1016/j.erss.2019.101301

 Scofield, J. H. (2014). Energy star building benchmarking scores: good idea, bad science. Proceedings of ACEEE Summer Study on Energy Efficiency in Buildings, Pacific Grove, USA. Retrieved April 2, 2021, from http://aceee.org/files/proceedings/2014/data/index.htm. Accessed 29 Apr 2022.

Semple, S., & Jenkins, D. (2020). Variation of energy performance certificate assessments in the European Union. Energy Policy, 137, 111127. https://doi.org/10.1016/j.enpol.2019.111127

Sunikka-Blank, M., & Galvin, R. (2012). Introducing the prebound effect: The gap between performance and actual energy consumption. Building Research and Information, 40, 260–273. https://doi.org/10.1080/09613218.2012.690952

Taranu, V., & Verbeeck, G. (2018). A closer look into the European Energy Performance Certificates under the lenses of behavioural insights – a comparative analysis. Energy Efficiency, 11, 1745–1761. https://doi.org/10.1007/s12053-017-9576-6

Taranu, V., Verbeeck, G., & Nuysts, E. (2020). Upgrading the energy label for dwellings in Flanders: An example of a behaviourally informed policy tool. Building Research & Information, 48(1), 18–33. https://doi.org/10.1080/09613218.2019.1661763

Telford J. K. (2007). A brief introduction to design of experiments. Johns Hopkins APL Technical Digest, 27(3), 224–232. Retrieved March 3, 2022, from https://www.jhuapl.edu/Content/techdigest/pdf/V27-N03/27-03-Telford.pdf. Accessed 29 Apr 2022.

Tronchin, L., & Fabbri, K. (2012). Energy Performance Certificate of building and confidence interval in assessment: An Italian case study. Energy Policy, 48, 176–184. https://doi.org/10.1016/j.enpol.2012.05.011

US Energy Information Administration. (2021). How much energy is consumed in US buildings? Retrieved March 3, 2022, from http://www.eia.gov/tools/faqs/faq.cfm?id=86&t=1. Accessed 29 Apr 2022.

US Environmental Protection Agency. (n.d.a). Energy Star, How the 1–100 Energy Star score is calculated? Retrieved March 3, 2022, from http://www.energystar.gov/buildings/facility-owners-and-managers/existing-buildings/use-portfolio-manager/understand-metrics/how-1-100. Accessed 29 Apr 2022.

US Environmental Protection Agency. (n.d.b.). Benchmark your building using Energy Star Portfolio Manager. From https://www.energystar.gov/buildings/benchmark. Accessed 29 Apr 2022.

US Environmental Protection Agency. (2012). Benchmarking and energy savings. Retrieved March 3, 2022, from https://www.energystar.gov/sites/default/files/buildings/tools/DataTrends_Savings_20121002.pdf. Accessed 29 Apr 2022.

US Environmental Protection Agency. (2018). Energy Star, Energy Star Score for Multifamily Housing in the United States. Retrieved March 3, 2022, from https://www.energystar.gov/sites/default/files/tools/Multifamily_August_2018_EN_508.pdf. Accessed 29 Apr 2022.

US Environmental Protection Agency. (2020). Energy star, facts and stats. Retrieved March 3, 2022, from https://www.energystar.gov/buildings/about-us/facts-and-stats. Accessed 29 Apr 2022.

US Environmental Protection Agency. (2021). Building performance standards: overview for state and local decision makers. Retrieved March 3, 2022, from https://www.epa.gov/sites/default/files/2021-02/documents/benchmarking_building_performance_standards_section2.pdf. Accessed 29 Apr 2022.

US Environmental Protection Agency. (2022). Updates to energy star scores. From https://www.energystar.gov/buildings/benchmark/understand_metrics/score_updates. Accessed 29 Apr 2022.

van der Bent, H. S., van der Brom, P. I., Visscher, H. J., Meijer, A., & Mouter, N. (2021). The energy performance of dwellings of Dutch non-profit housing associations: Modelling actual energy consumption. Energy and Buildings, 253, 11486. https://doi.org/10.1016/j.enbuild.2021.111486
van den Brom, P., Meijer, A., & Visscher, H. (2019). Actual energy saving effects of thermal renovations in dwellings—longitudinal data analysis including building and occupant characteristics. *Energy and Buildings, 182*, 251–263. https://doi.org/10.1016/j.enbuild.2018.10.025

Von Platten, J., Holmberg, C., Mangold, M., Johansson, T., & Mjörnell, K. (2019). The renewing of energy performance certificates – reaching comparability between decade-apart energy records. *Applied Energy, 255*, 113902. https://doi.org/10.1016/j.apenergy.2019.113902

Yates F. (1937). The design and analysis of factorial experiments. *Technical Communication of the Commonwealth Bureau of Soils 35 Commonwealth Agricultural Bureaux, Farnham Royal.*

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.