Please cite this paper as follows:

Benedetto Grillone, Stoyan Danov, Andreas Sumper, Jordi Cipriano, Gerard Mor,
A review of deterministic and data-driven methods to quantify energy efficiency savings and to predict retrofitting scenarios in buildings,
Renewable and Sustainable Energy Reviews,
Volume 131,
2020,
110027,
ISSN 1364-0321,
https://doi.org/10.1016/j.rser.2020.110027.
(http://www.sciencedirect.com/science/article/pii/S136403212030318X)

A review of deterministic and data-driven methods to quantify energy efficiency savings and to predict retrofitting scenarios in buildings

Benedetto Grillone[a,*], Stoyan Danov[a], Andreas Sumper[b], Jordi Cipriano[c,d], Gerard Mor[c]

[a]*Centre Internacional de Mètodes Numèrics a l'Enginyeria. Building Energy and Environment Group. CIMNE - UPC Campus Terrassa Edifici GAIA, C. Rambla Sant Nebridi 22, 08222 Terrassa, Barcelona, Spain*

[b]*Centre d'Innovació Tecnològica en Convertidors Estàtics i Accionaments (CITCEA-UPC), Departament d'Enginyeria Elèctrica, Universitat Politècnica de Catalunya, ETS d'Enginyeria Industrial de Barcelona, Av. Diagonal 647, Pl. 2, 08028 Barcelona, Spain*

[c]*Centre Internacional de Mètodes Numèrics a l'Enginyeria. Building Energy and Environment Group. CIMNE-Lleida. Pere de Cabrera 16 2. Office G., 25002 Lleida, Spain*

[d]*Department of Environmental and Soil Sciences, INSPIRES, University of Lleida, Rovira Roure 191, 25198 Lleida, Spain*

Abstract

Increasing the energy efficiency of the built environment has become a priority worldwide and especially in Europe. Because of the relatively low turn-over rate of the existing built environment, energy efficiency retrofitting appears to be a fundamental step in reducing its energy consumption. Last experiences have shown that there is a vast energy efficiency potential lying in the building stock, and it is mainly untapped. One of the reasons is a lack of robust methodologies for planning and evaluation of building energy retrofitting strategies. Nowadays, dynamic measured data coming from automated metering infrastructure provides valuable information to evaluate the effect of energy conservation strategies. For this reason, energy performance modelling and assessment methods based on this data are starting to play a major role. In this paper, several methodologies for the measurement and verification of energy savings, and for the prediction and recommendation of energy retrofitting strategies, are analysed in detail. Practitioners looking at different options for these two processes, will find in this review a thorough and detailed overview of the different methods that can be used. Guidance is also provided to determine which

*Corresponding author. *E-mail address*: bgrillone@cimne.upc.edu (B. Grillone)

method could work best depending on the specific case under analysis. The reviewed approaches include statistical learning models, machine learning models, Bayesian methods, deterministic approaches, and hybrid techniques that combine deterministic and data-driven modeling. Existing research gaps are identified and prospects for future investigation are presented within the main conclusions of this research work.

Keywords: building energy retrofitting, energy savings evaluation, data driven approach, measurement and verification, retrofitting decision support, energy performance improvement

---

Abbreviations

ANN Artificial Neural Network

ASHRAE American Society of Heating, Refrigerating and Air-Conditioning Engineers

BART Bayesian Additive Regression Trees

BES Building Energy Simulation

BPD Building Performance Database

CDD Cooling Degree Days

CV(RMSE) Coefficient of Variation of the Root Mean Square Error

EEM Energy Efficiency Measure

EUI Energy Usage Intensity

FRL Fallen Rule List

GA Genetic Algorithm

GBM Gradient Boosting Machine

GMR Gaussian Mixture Regression

GP Gaussian Process

HDD Heating Degree Days

HVAC Heating, Ventilating and Air Conditioning

IPMVP International Performance Measurement and Verification Protocol

MCEM Monte Carlo Expectation Maximization

M&V Measurement and Verification

NMBE Normalized Mean Bias Error

NRE Non-Routine Event

NSGA Non Sorted Genetic Algorithm

NZEB Nearly Zero Energy Building

PDF Probability Density Function

RMSE Root Mean Square Error

SVM Support Vector Machines

Contents

1. Introduction

Low energy performance of the built environment is one of the main barriers to reach the 2030 European energy efficiency targets [1]. One of the most successful ways to address low building energy efficiency is a massive and affordable implementation of energy renovation strategies [2, 3]. However, at present, there are still several barriers hindering the adoption of procedures and technologies that improve energy efficiency, and limiting the investments in this field. Tuominen et al. [4] found a low impact of renovations on on property prices, lack of trusted information, and small prioritization for energy performance improvements, to be frequently cited as the main barriers, in the case of privately owned residential buildings. On the other hand, Kontokosta [5] identified information asymmetry between project partners, uncertainty over expected savings, and shortage of expertise in energy technologies, as the main obstacles in the retrofitting decision making process for commercial office buildings. In the latter case, the author also highlights that these issues have been worsened by case-study oriented approaches, many times because of lack of extensive data and comprehensive pre/post analyses of load profiles following an energy efficiency measure (EEM) implementation.

For commercial and public buildings, applied EEMs can have a significant impact, but the evaluation of this impact with certainty and reliability is no easy task. At the same time, no consolidated framework exists to evaluate ex-ante the effect of different energy retrofitting strategies over buildings. Several techniques to find the most cost-efficient set of measures for a particular building have been developed [6], but scaling up such methods proves to be a major technical challenge, since the effectiveness of retrofitting actions depends on many parameters and this is a clear constraint for any evaluation method.

The objective of this review paper is to establish the state of knowledge related with the modeling-based approaches used to support the planning and evaluation of building energy retrofitting strategies. More specifically, the paper aims at reviewing methods, as well as tools, to:

- determine the energy savings obtained through an energy retrofitting program (commonly referred to as measurement and verification).
- support the process of identification of the most appropriate energy renovation action according to the specific features of the analyzed building (in this paper referred to as prediction and recommendation).

Although few reviews already exist, partially covering the topics addressed in this article, to the best of the authors' knowledge no published review provides an indepth and comprehensive analysis such as the one presented here. In no other review work the measurement and verification, and the prediction and recommendation processes are analysed together and in a structured way as in the present article. The details of this analysis are described in Chapter 3. This review work focuses mainly on data-driven methods, although some deterministic and hybrid methods are also analysed. The reason for this is that, in the last years, a surge in the number of smart energy monitoring devices has significantly increased the amount of building energy performance data available. This made possible the setting up of many publicly available databases containing energy consumption data and building characteristics of hundreds of thousands of buildings. Data-driven methods are hence becoming of increasing interest, as they are able to harness such huge amount of information for both evaluating the applied energy retrofitting measures and predicting the energy savings potential of new EEMs [7]. Moreover, traditional deterministic methods not based on data have to face an important issue related with their scalability, since the results obtained are usually only valid for the specific building under analysis. This means that using these methods to develop large scale retrofitting strategies can be a major challenge [8]. It's also important to point out that data-driven techniques are being already widely employed in building energy efficiency, and several interesting applications are arising, such as control optimization in demand response, efficiency improvement of HVAC systems, energy efficient operation of different types of buildings, and more [9–11].

The article is organized as follows. Section 2 introduces the reader to different key concepts and how they are used in the context of this review work: the measurement and verification process, the prediction and recommendation process, and the distinction between data-driven and deterministic models. In Section 3, a concise overview of previous studies focused on building energy consumption modeling and forecasting techniques is provided. In Section 4, a review of existing M&V protocols, as well as data-driven energy baseline estimation methods is presented. State-of-the-art techniques for non-routine event detection and uncertainty estimation are also reviewed in that section. Section 5 includes a detailed review of methods to predict the effect of energy efficiency measures and to plan energy retrofitting strategies. Finally, in Sections 6 and 7 the discussion and conclusions of this review work are outlined. The structure of the paper is also illustrated in Fig. 1, where the two main processes reviewed in this article are highlighted, together with the different applications studied in each case.

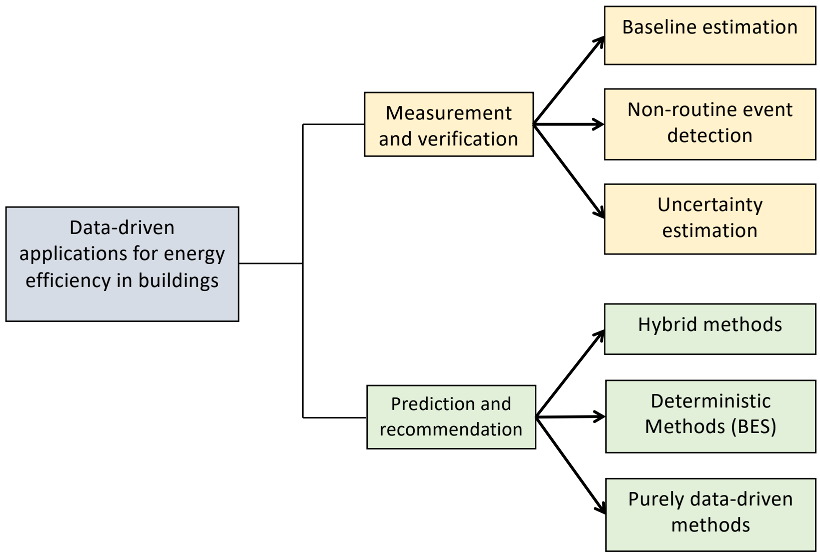

Figure 1: Illustration of paper structure

2. Background

In this section, some concepts which can help to better understand the full content of the review, are introduced, namely: the measurement and verification process, the prediction and recommendation process, and the difference between data-driven methods and deterministic methods.

2.1. The measurement and verification process

Measurement and verification (M&V) is the process of using measurements to accurately estimate real savings generated in a facility thanks to the implementation of an energy management strategy [12].

2.1.1. Baseline modeling

Since savings can't be directly measured, as they represent the absence of energy usage, they are determined by comparing measured energy consumption before and after the implementation of a retrofit measure, considering the relevant adjustments for changes in conditions. In order to carry out a comparison between the energy usage before and after the EEM application, a model of the consumption prior to the implementation of the measures needs to be developed. This is model is called the baseline energy model. The baseline model can be defined as the energy characterization of the starting situation and has a fundamental role in the determination of energy savings. In fact, the baseline model allows to isolate the effects of a retrofit intervention from the effects of other parameters that can simultaneously affect the energy consumption, therefore reducing the uncertainty with which savings are estimated. In this article, the most common data-driven methods used to develop baseline models are reviewed.

2.1.2. Advanced measurement and verification (M&V 2.0)

In recent years, M&V has been transitioning to a new state, known in the field as "advanced measurement and verification" (or M&V 2.0). This new form of M&V is a result of the breakthroughs in advanced metering infrastructure systems and automated analytics techniques. In M&V 2.0, high granularity datasets with increased sampling frequency, volume, and resolution, are analysed, in order to perform an estimation of energy efficiency savings which is almost in real-time [13]. This is enabling M&V to advance from a static and cumbersome process to a more dynamic one, that translates into hourly energy insights, maximized savings and great benefit for all the parts involved in the energy retrofitting programs [14]. One of the main drivers of M&V 2.0 is the development of accurate baseline models for real-time savings estimation, through the application of advanced statistical and machine learning techniques. The new features of M&V 2.0 are not only limited to savings evaluation, in fact, most of the advanced M&V tools currently on the market also provide a range of different services, such as analysis and visualization of energy monitoring data, system-level fault detection and diagnostics, and building energy benchmarking [15].

2.2. The prediction and recommendation process

The term *prediction* refers to to a group of techniques used to predict the effect of an hypothetical EEM application on an individual building or facility. The prediction results are then used to *recommend* the application of specific EEMs over others, and to plan optimal energy retrofitting scenarios. Thanks to prediction and recommendation techniques, it's possible to answer many different questions, such as: "What is the return on investment for a specific EEM?", "Which EEM would perform best in the selected building, given its characteristics?", "Which low capital cost measures can be applied to increase the energy performance of the selected building?", "Of all the buildings belonging to the considered stock, which ones would benefit the most from an energy renovation program?", "Which EEM would yield the highest energy savings, in a 30 years time span?" etc. All these questions are commonly answered by an engineer, after performing a building energy audit, although the results obtained with the audit can be very uncertain. In Section 5, an overview is provided of practical data-driven and deterministic methods to predict EEM impact and plan energy retrofitting strategies for an individual facility or a group of buildings.

2.3. Data-driven models and deterministic models

Having clear the goals of the two main processes that are going to be studied in this review, let's now define the two categories of methods under analysis: data-driven models and deterministic models.

Data-driven models are statistical models that find relationships between state variables of the analyzed system (inputs and outputs) without explicit or detailed knowledge of its physical behaviour. In the case of models built for M&V, for example, typical input variables can be external air temperature, wind speed and direction, solar irradiance, building occupancy rate, while typical output variables can be the total electrical or thermal load of the building. Depending on the level of physical significance of the parameters used, these models are usually referred to as *grey-box* or *black-box* models.

The other class of methods reviewed in this article are deterministic methods: detailed building energy simulation models based on the differential equations of the energy transfer flows occurring in the control volumes (rooms or spaces) of the buildings. These physics-based models are usually referred to as *white-box* models.

While for the measurement and verification process, the methods reviewed in this article are exclusively data-driven, in the prediction and recommendation section, both data-driven and deterministic models are analysed, as well as "hybrid" models, in various which data-driven techniques are used to analyse results obtained with deterministic methods.

3. Existing review studies

This paragraph gives a concise but complete overview of previously published review works regarding the different topics treated in this article. To the best of the authors' knowledge, there is no published review that addresses the same topics presented in this article, that is: an up-to-date and detailed analysis of data-driven and deterministic methodologies used to verify the effect of EEMs in buildings and to predict the impact of future energy retrofitting strategies. The existing data-driven and machine learning techniques used to model and forecast building energy consumption have been thoroughly analysed in a wide range of reviews published over the last years: [16–23].

Deb et al. [24] divided state-of-the-art forecasting methods in nine different categories and compared them in terms of length of training, data needed, accuracy, and computation time required for the estimation. Wei et al. [25] extended this analysis to other applications, such as energy pattern profile identification, energy-usage mapping, benchmarking of the building stock, and the definition of extensive retrofitting plans. Data-driven techniques related to the development of retrofitting strategies were also studied in the same review(artificial neural networks, genetic algorithms, and clustering techniques).

Harish and Kumar [26] carried out an analysis of different approaches to model and simulate building energy systems and to evaluate the impact of energy retrofitting strategies. Different dynamic modeling techniques were reviewed, including the forward approach (white-box), the data-driven approach (black-box) and the hybrid grey-box approach. The different methods were then classified according to the model type, the parameters used, the simulation period and, the method of validating the results. A list of building energy simulation software, together with their strengths and limitations is also presented in that paper.

Lee et al. [27] reviewed retrofitting analysis toolkits for commercial buildings, classifying them in 3 main categories: toolkits using data-driven methods, toolkits using normative calculations, and toolkits using physics-based energy models. From the analysis, it appears that there is still room for improvement of these methods, especially regarding: (i) mitigation of the high degree of uncertainty associated with these tools, (ii) interoperability between the different tools, (iii) incorporation of human behaviour in the models, (iv) extension of output parameters. An overview of the current state of advanced measurement and verification tools was also provided by Granderson and Fernandes [15]. The authors reviewed sixteen different commercially available tools and classified them according to various criteria: the standard protocol employed, the type of baseline models used, the input data granularity required, the possibility to provide uncertainty estimates, and more. Granderson et al. [28] also compared the accuracy of ten different baseline energy use models for automated measurement and verification of energy savings. The techniques were tested on 537 commercial buildings in the US using training periods of different lengths and without any non-routine adjustment. Two different error metrics: normalized mean bias error (NMBE) and coefficient of variation of the root mean squared error (CV(RMSE)) were calculated and compared, showing similar performances for the ten models. Results of this analysis showed that data-driven statistical techniques are better candidates for scaling up the adoption of whole-building energy savings evaluations using advanced metering infrastructure. In a subsequent publication [29], the same authors applied one of the ten methods (the *time of the week and temperature* baseline model) on a set of 84 buildings, in an attempt to test the applicability of these M&V approaches on a larger scale. It was found that 70% of the buildings of the data set were well fit to be analysed with the automated approach, and in 80% of the cases savings and uncertainties were quantified to levels above the minimum acceptable thresholds defined by the ASHRAE Guideline 14 [30], a standard protocol used for M&V.

Although the presented review works are of great importance, there is still a shortage of studies covering specifically, and in detail, the processes of measurement and verification of energy savings, and of energy retrofitting planning. Practitioners looking at different options for these two processes, will find in this review a thorough, as well as detailed, overview of the different methods that can be used. Guidance is also provided to determine which method could work best depending on the specific case under analysis. At the same time, it's important to highlight how this review work is mainly focused on data-driven approaches. Considering the growing attention that statistical and machine learning techniques are now receiving in the field of building energy performance analysis, such a study appears essential to identify the research gaps and to highlight future research lines.

4. Measurement & Verification: review of methods and data-driven applications

This section aims at reviewing the most popular M&V methods currently in use, with special focus on the data-driven techniques used to estimate baseline energy models. In the first part of the section, four frequently employed M&V protocols are introduced. Following, a review of state-of-the-art data-driven techniques to develop baseline energy models and estimate retrofit savings is presented. The last paragraphs of the section present a review of data-driven approaches to the problems of non-routine event detection and savings uncertainty estimation.

4.1. Measurement & Verification protocols

M&V is an evolving science and various methods and best practices were drawn up and documented in different guidelines. Attempts have been made to create a unique standard for the M&V process, but depending on the analysed facility's geographical location, principal use (residential, commercial, industrial, etc.), and type of metering data available, practitioners still employ different protocols. The optimal degree of standardization that will ultimately be required for advanced M&V is an open issue and currently under discussion among stakeholder groups [15].

4.1.1. International Performance Measurement and Verification Protocol (IPMVP)

The International Performance Measurement and Verification Protocol [12], proposed by Efficiency Valuation Organization (EVO), defines standard terms and suggests best practices to quantify energy savings following the application of one or more energy efficiency measures. According to this protocol, four different options are available to determine energy efficiency savings:

- Option A: Partially Measured Retrofit Isolation. This option involves the use of measurement instruments to monitor the consumption of the equipment affected by the applied EEM, isolated from the energy usage of the rest of the building. In this option, only partial measurement is used, meaning that some parameter(s) are estimated rather than measured.

- Option B: Retrofit Isolation. This case is equivalent to option A, with the exception that no estimations are allowed and full measurement of all the relevant parameters is required.

- Option C: Whole Building. In this approach, utility meters are used to evaluate the energy performance of the whole building. Option C determines the total savings of all implemented EEMs and is only applicable in projects where savings are expected to have a substantial impact, making them distinguishable from energy variations unrelated to the applied measures.

- Option D: Calibrated Simulation. This option involves using building energy modeling software that allows the prediction of energy consumption in different scenarios. The models used for this scope are first calibrated, making sure that the predicted energy load of the building matches the real (metered) data.

4.1.2. ASHRAE Guideline 14

The ASHRAE Guideline 14 for measurement of Energy, Demand and Water Savings [30], published by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE), also specifies three different approaches to determine energy savings:

- Retrofit Isolation Approach, similar to IPMVP option B
- Whole Facility Approach, similar to IPMVP option C
- Whole Building Calibrated Simulation Approach, similar to IPMVP option D

Furthermore, the ASHRAE guideline provides different metrics to evaluate the validity of the applied models, such as thresholds for net determination bias or the maximum acceptable uncertainty of the estimated savings.

4.1.3. DOE Uniform Methods Project

The US Department of Energy (DOE), is also building a set of protocols to assess savings due to energy renovation programs. These protocols, joined together under the name Uniform Methods Project [31], provide a simple and clear method to determine energy savings for residential, industrial, and commercial buildings. The protocols are based on IPMVP, but supplementary practices are included, that can be used to aggregate savings from single retrofitting actions and assess program-wide effects.

4.1.4. CalTRACK

CalTRACK [32] is a protocol that was born from the efforts of the California Energy Commission and the California Public Utilities Commission to have a standardized protocol for the evaluation of energy savings in the residential sector. CalTRACK specifies a set of methods to measure and report changes in the energy consumption of a building following the application of an EEM. These methods have the goal of estimating the energy that would have been consumed in the building if the intervention had not taken place. The techniques implemented have been empirically tested by a technical team with several different stakeholders and developed under an open-source license model. The data required to apply the CalTRACK methods includes one full year of consumption data before the EEM application, local weather data, and the date of implementation of the measure.

4.2. Data-driven baseline estimation methods

Several baseline energy modeling approaches, using both monthly billing and interval meter data, are presented in the next paragraphs. The reviewed methods are classified into statistical learning, machine learning, and Bayesian techniques, Fig.

2 shows an overview of how this section is structured, and Table 1 summarizes the characteristics of all the models analysed.

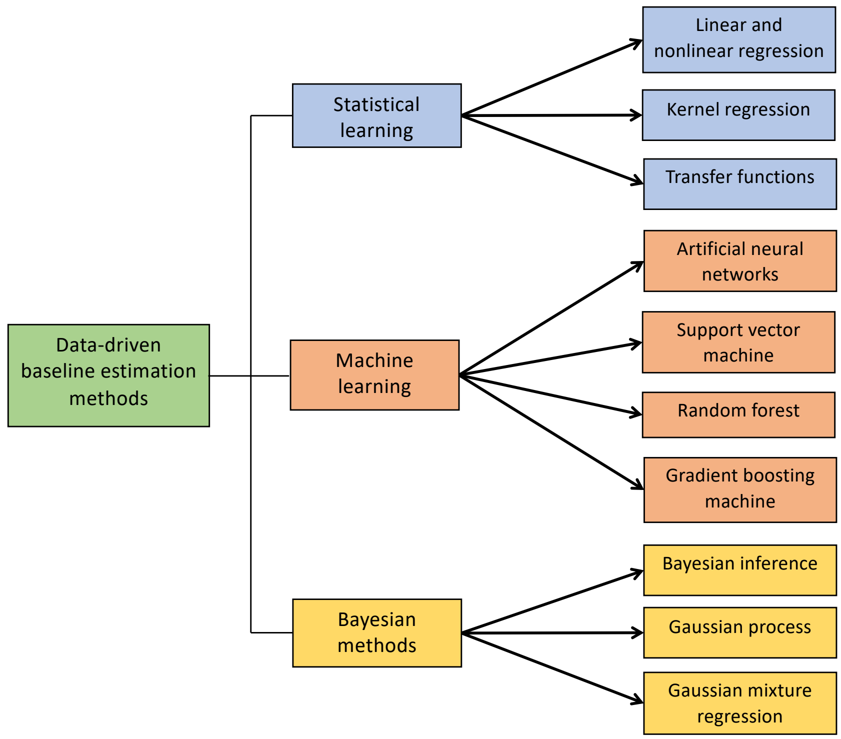

Figure 2: Load profiles identified for the office building

4.2.1. Statistical learning techniques

Statistical learning is a branch of data-driven modeling that is based on building a statistical model by inferring relationships between different variables in the analysed dataset. This model is then used to make predictions on other datasets supposed to be similar to the one used to build the model.

4.2.1.1. Linear and nonlinear regression.
Regression analysis has been the first implemented statistical method for the evaluation of energy savings in buildings. Its origins can be traced back to the development

of the PRInceton Scorekeeping Method (PRISM) [33], a statistical procedure formulated to include weather normalization in the estimation (scorekeeping) of energy savings. This model is obtained by applying a regression technique that takes into account different variables frequently having an impact on energy usage, such as occupancy, climate, and equipment operation. Common variables chosen for the regression can be: average outdoor temperature, relative humidity, cooling degree days (CDD), heating degree days (HDD), building occupancy and building working days.

Mathieu et al. [34] used linear regression to estimate building energy baselines using high granularity (15-min-interval) consumption data. The model proposed includes an indicator variable that marks the hour of the week and a piecewise-linear temperature regressor having fixed change points. In addition, two different regression models are fit for when the building is considered occupied or unoccupied. This method has been shown to be highly accurate [28] and has been used as a benchmark model in several recent publications regarding measurement and verification methods [35][36][37].

Mohd et al. [38] also tested a linear regression approach to evaluate the effect of an EEM over the HVAC system in an office complex in Malaysia. Both single variable and multivariate linear regressions were fitted, using monthly billing data, temperature readings, and occupancy details. A similar approach was followed by Wang et al. [39] , who tested different linear and nonlinear regression models to assess the energy savings caused by a mechanical system retrofitting in a healthcare facility in Dallas, Texas. The models were fitted with electricity and gas monthly billing data and using average outdoor temperature and degree-day as independent variables. The regression model approach was also tested in the industrial sector: Kissock and Eger [40] built a baseline energy model with multivariable piece-wise linear regression, that was used to disaggregate savings in an industrial facility. The facility's consumption was supposed linearly dependent on its production and on the outdoor air temperature.

Regression analysis is appealing for its simplicity and the possibility of applying it even when low resolution data is available. On the other hand, the linear approach can sometimes be too simple to capture complex relationships between variables.

4.2.1.2. Kernel regression.
Kernel regression belongs to a special class of regression models, called time-varying coefficient models, where the regressors are not considered constant, but dynamically changing over time. The use of kernel regression to estimate building energy consumption baselines was first proposed by Brown et al. [41], with the goal of improving the predictive accuracy of standard linear regression models. The idea behind kernel regression is that the regressors are not estimated using the whole historical dataset.

Instead, the regressors are evaluated for each timestep, by estimating a weighted average of all the timesteps with the nearest values of the regression parameters (e.g. weather conditions, time of the day, etc.). The main advantages of kernel regression are an increased estimation accuracy, compared to standard linear regression, and the ability to provide robust and reasonable results even in case of small training sets. On the other hand, since the coefficients are evaluated for a rolling time window and not considering the whole timeseries dataset, when making predictions for longer time frames (e.g. one year or more) the model might not be able characterize the existing seasonal variations and generalize properly .

4.2.1.3. Transfer functions.
Transfer functions have been shown to be capable of accurately estimating the thermal parameters of buildings [42–45], their application for verification of energy savings is now also being tested. The great advantage of transfer functions is the possibility to take into account the building dynamics connected to its thermal inertia. Furthermore, the coefficients of the transfer function model are coupled with the features of the building, thus avoiding the requirement of large amounts of data to obtain reliable results. One of the drawbacks of the method is that the calculations are based on the internal temperature of the building, which is not always known when performing M&V. This baseline estimation methodology was first suggested by Díaz et al. [46], who combined two transfer function models to assess energy efficiency savings in a building of the University of Granada.

4.2.2. Machine learning techniques

The term machine learning (ML) identifies algorithms that make use of statistical models in order learn from data without any specifically programmed instruction. ML algorithms identify patterns in the dataset through iteration and are then able to harness the gained information to make predictions.

4.2.2.1. Artificial neural networks.
Artificial neural networks (ANN) have been applied in several cases to develop baseline energy models [17–20]. The black-box nature of these models makes them very popular, since they can be easily applied to many different problems after just a quick data pre-processing phase. But their simplicity comes at the expense of feature interpretability, making the process of debugging and model improvement considerably more difficult. Low model interpretability and the need for large amounts of training data are the main drawbacks of ANNs. Yalcintas [47] tested ANN models using Levenberg-Marquardt back-propagation to evaluate energy retrofitting savings

in two hotel buildings. Adnan et al. [48] used an Hybrid Artificial Neural Network, in combination with Evolutionary Programming, to quantify the savings achieved for a chiller unit in Malaysia, using three different inputs: operating time, refrigerant tonnage and differential temperature. Chang et al. [49] also assessed post retrofit energy savings for an air conditioning system, using ANNs and an energy saving evaluation model based on a parameter named *Refrigeration Operation Energy saving Effect Ratio* (ROEER).

4.2.2.2. Support vector machine.
Support vector machine (SVM) was first applied to estimate building energy baselines by Dong et al. [50]. This machine learning approach is usually preferred when the training data available is small, since it proves to be very powerful in solving problems with non-linear formulations, even with small training datasets. The training time of this technique scales cubically with the size of the dataset [51], making SVM not ideal when dealing with large-size problems. In [25], an overview of the most recent applications of SVM to building energy consumption prediction is presented.

4.2.2.3. Random forest.
Random forest is an ensemble learning algorithm that constructs several decision trees and then outputs the mean of their prediction, in order to correct for the individual trees' tendency to overfit the data. This powerful methodology has been used for several applications in the domain of building energy prediction. Ahmad et al. [52] used random forests to predict hourly HVAC energy consumption, while Araya et al. [53] proposed their use for fault detection and diagnosis. In the measurement and verification framework, the use of random forests was outlined both in [28] and [15]. Random forests prove to be very accurate in the prediction of building energy usage, although the black-box nature of this algorithm means that the computational time associated with this calculation is quite high, due to the necessity of optimizing the hyper-parameters and performing cross-validation to avoid overfitting.

4.2.2.4. Gradient boosting machine.
Similar to random forest, the gradient boosting machine (GBM) is a powerful machine learning algorithm based on the concept that a "strong learner", having high prediction accuracy, can be obtained by iteratively combining several less complex models, called "weak learners". Touzani et al. [36] used this approach to build an energy consumption baseline model that can be applied for energy savings estimation. The algorithm has four hyper-parameters that were optimized using grid search with 5-fold block cross-validation. The results of the GBM method were compared to the ones obtained with a piecewise linear regression model and a random forest

algorithm. This analysis showed that the GBM was able to improve both R^2 prediction accuracy and CV(RMSE) in most of the analysed cases.

4.2.3. Bayesian methods

As an alternative to the more traditional frequentist approach, several researchers studied the application of the Bayesian paradigm to the measurement and verification process. In Bayesian statistics, a probability model is fit to a dataset, with the goal of obtaining a probability distribution on the model parameters and on other values, like predictions for unobserved data [54]. Then, as new data becomes available, Bayes' theorem is used to update these probability distributions. Among the advantages of Bayesian methods, authors list: the possibility of automatically and exactly quantifying the uncertainty of the models (including different sources of uncertainty, like measurement errors and weather variability), lower sensitivity to outliers, the possibility to have real-time updates of the estimates, and more [55][56].

4.2.3.1. Bayesian paramaters inference.

Lindelöf et al. [57] applied Bayesian inference to analyse energy invoices and climate data to estimate the impact of the installation of a model-predictive controller for a heating system in an office building in Switzerland. The approach tries to estimate the probability density function (PDF) of three parameters: the building's heat-loss coefficient, the building's balance temperature, and the stochastic variations of the heating demand, conditioned on the information contained in the utility invoices. The impact of the EEM is assessed by estimating the variations of the heat loss coefficient, through the analysis of a PDF obtained by fitting a Bayesian model to the billing data before and after the EEM application. The Bayesian approach allows to extract high amounts of information from the data and proves to be especially useful in the case of data with monthly granularity. One of the main challenges of this method is that the first probability model, called prior, is often not easy to find and justify, and can be a major source of inaccuracy.

4.2.3.2. Gaussian process.

The application of Gaussian processes (GP) in the M&V process was first proposed by Heo and Zavala [58], with the goal of solving certain limitations of the linear regression method. The Gaussian process approach is non-parametric, since its aim is not finding the parameters of a given function that can best fit the data, but to look for a distribution over the functions $f(x)$ potentially consistent with the observations. GPs can capture complex building energy behaviour, such as nonlinear trends, multivariable interactions and time correlations. At the same time, since GPs belong to the framework of Bayesian statistics, this method allows the savings' uncertainties to be quantified thoroughly. Burkhart et al. [59] suggested the use of Monte Carlo expectation maximization (MCEM) to enhance GP modeling and grant more accurate predictions in case of uncertain input data. Maritz et al. [60] published a guideline to perform M&V using GPs, with special emphasis on the process of kernel selection. The approach is described step by step and then applied to adjust the baseline consumption of an academic facility. A two-stage grid search technique is used to determine the best fit coefficients for the model, which is then applied to calculate savings in two different case studies. One of the main issues associated with this method is its computational and memory complexity, that increases cubically with the size of the training dataset.

4.2.3.3. Gaussian mixture regression.

Srivastav et al. [61] tested the performance of Gaussian mixture regression (GMR) for building baseline energy prediction. The approach was tested on both simulated data from the US Department of Energy and on real data from a commercial building in California, accuracy was compared with a linear regression model. The model showed an estimation accuracy comparable with the multivariate regression approach in both cases, although GMR has the key advantage of allowing the computation of confidence intervals that adapt locally for different circumstances, according to the uncertainty of training data. At the same time, GMR seems to be less sensitive to data sparsity and to regressors correlation. Similarly to GPs, the main challenges of GMR are linked to its long computational time.

4.3. Non-routine event detection

The issue of non-routine event detection is a known challenge in the M&V research community and is common to all the previously introduced baseline estimation methods. Non-routine events (NREs) are defined as fluctuations in the energy usage of a building that are not caused by any variation of the explanatory variables of the baseline model, and that are not attributable to the applied measure itself. In order to achieve a precise evaluation of the energy savings, non-routine events must be detected, and accounted for as non-routine adjustments in the estimation of avoided energy use. This process is usually performed manually and, depending on the kind of event, it might require some engineering expertise and knowledge of what the

NRE was [29]. This is a considerable issue in automated M&V, as failing to identify such events could lead to an over (or under) estimation of the savings. Recently, Touzani et al. [62] proposed an automated technique, based on statistical change point detection, to identify non-routine events and adjust the savings calculations. The preliminary results of this study, carried out on a set of synthetic data created using energy simulation software EnergyPlus, show a high identification rate for true positives, as well as for false positives, suggesting that the algorithm might still be improved to achieve better results.

4.4. Uncertainty estimation

In the M&V context, determining the uncertainty of the obtained results proves to be an issue of major importance. Providing a range of uncertainty, together with the point estimate result, can help establishing the amount of risk associated with a given investment, and support stakeholders in making more informed decisions [63]. Energy savings estimates usually provide results in form of a single point value, the uncertainty can then be interpreted as the interval of doubt around this estimate [37]. According to the IPMVP, when dealing with energy savings, three kinds of quantifiable uncertainties are identified: sampling uncertainty, arising from the fact that in some projects not all the devices can be monitored, hence sampling techniques are used, measurement uncertainty, related to the accuracy of the monitoring infrastructure used to measure the energy consumption, and modeling uncertainty, related to the errors of the baseline models used to estimate the savings.

Reddy and Claridge [64] argued that the uncertainty in the consumption baseline model is the key factor in determining the uncertainty in the measured savings and proposed a formula to estimate it taking into account the CV(RMSE) of the employed statistical model and the relative influence of the EEM on the baseline energy consumption. Koran et al. [65] compared four different methods to calculate the uncertainty of energy efficiency savings estimated using metering data: a formula found in the ASHRAE Guideline 14 [30], an improved version of the ASHRAE formula, an exact formula that can be used in the case of ordinary least square regression, and a bootstrapping technique. All the four methods presented provided reasonable results, although the accuracy of the methods was not evaluated. Subsequently, a work by Touzani et al. [37] compared the accuracy of two different approaches to determine the uncertainty of energy efficiency savings estimations. Four different baseline models were applied: two hourly models and two daily ones. The uncertainty of the model estimates were then analysed using two methods: the ASHRAE Guideline 14 approach and the k-fold cross-validation approach, a method to assess model accuracy commonly utilized in the machine learning community. The study was carried out on a dataset comprising whole-building electricity consumption data, sampled every 15 minutes, from 69 commercial buildings located in Central California, Northern California, and Washington DC. The results showed that both methods underestimated the uncertainty of all the four baseline models tested, although the underestimation proved to be stronger for hourly models, probably due to higher autocorrelation of residuals.

Among the few authors to take into account other uncertainties than the modelling one, Olinga et al. [66] proposed a method to optimally allocate budget and effort in M&V while handling both sampling and modeling uncertainties. The results of their case study show a 42 % reduction of the sampling cost and an 11 % reduction of the total M&V cost thanks to the implementation of the proposed approach.

Table 1: Characteristics of the analysed baseline estimation methods for M&V

Model	**Advantages**	**Limitations**	**Explanatory variables used in the referenced articles**	**References**
Linear and nonlinear regression	Easy to interpret and explain	Sometimes too simple to capture complex relationships	Indoor air temperature, outdoor air temperature, HVAC schedule	[33, 34, 38–40]
Kernel regression	Better fitting than traditional regression	Not ideal to predict long time intervals	Indoor air temperature, outdoor air temperature, HVAC schedule	[41]
Transfer functions	Can model dynamic effects caused by thermal inertia	Requires indoor temperature data	Indoor air temperature, outdoor air temperature, solar radiation, HVAC schedule	[46]
Artificial neural networks	Performs well with non-linear timeseries	Requires large amounts of data, tends to overfit, slow to train	Outdoor air temperature, wind speed and direction, visibility, air pressure, operating time, refrigerant tonnage, running time of the system, refrigerating capacity, power rating of water pumps, differential temperature	[47–49]
Support vector machine	Performs well even with small training datasets	Long computational time for large datasets	Outdoor air temperature, relative humidity, global solar radiation	[25, 50]
Random forest	High predictive accuracy	Hyper-parameters optimization and cross-validation are needed to avoid overfitting	Outdoor air and dew point temperatures, relative humidity, hour of the day, day of the week, number of occupants booked in the hotel, energy consumption of previous hour	[52, 53]
Gradient boosting machine	Higher predictive accuracy than random forest	Hyper-parameters optimization and cross-validation are needed to avoid overfitting	Outdoor air temperature, time of the week, U.S. federal holidays	[36]
Bayesian inference	Accurate uncertainty estimation	Priors are often difficult to justify and can be a major source of inaccuracy	Cooling Degree Days	[57]
Gaussian processes	Able to capture complex (nonlinear) building energy behaviour	Computational and memory complexity	Outdoor air temperature, HVAC supply temperature, occupancy, relative humidity	[58–60]
Gaussian mixture regression	Dynamic confidence intervals	The optimization problem is not trivial to solve, long computation time	Outdoor air temperature, solar radiation, outdoor humidity	[61]

5. Prediction and recommendation: review of the methods

In this section, various techniques to predict the effect of energy efficiency measures and to plan energy retrofitting strategies for specific buildings or groups of buildings, are analysed. As many methods are involving the combination of deterministic models based on simulations and data-driven approaches, this section of the review presents three different categories of methods: deterministic, hybrid, and purely data-driven. Table 2 shows an overview of the methods discussed this section; the type of buildings where they were applied and the categories of the analysed retrofitting measures are also schematized.

5.1. Deterministic methods (Building Energy Simulation)

The approaches presented here are based on the application of building energy simulation (BES) to predict the energy performance of buildings in different scenarios.

5.1.1. BES models for retrofit and NZEB comparative analysis

Zangheri et al. [67] used building energy modeling software EnergyPlus [68] to identify which would be the most cost-optimal retrofit combination to reach nearly zero-energy building (NZEB) levels in different building/climate combinations. The study analyzes four different building typologies of 60s-70s and ten different climate areas within the European Union. In order to perform the study, first a "base refurbishment level" was defined, as the minimum possible level of refurbishment to which compare the deeper ones. The base refurbishment level was defined following the assumption that it is not possible to not intervene at all on a building older than 40 years, and includes the rehabilitation of the building envelope, and the substitution of the old heating or cooling systems with comparable equipment. It was found that cost-optimal and NZEB scenarios are characterized by an average increased investment cost, with respect to the base refurbishment level, of 50% and 115 % respectively. The energy efficiency potential of the cost-optimal cases proved to be substantial (between 36 % and 88% primary energy savings), with associated 30 years global costs many times lower than their respective base refurbishment levels.

Similarly, Rysanek and Choudhary [69] used TRNSYS [70], a simulation tool for transient systems, to analyse different energy retrofitting scenarios for a mid-sized office building in Cambridge (UK), while taking into consideration both technical and economic uncertainty. The authors also provide an analysis of how relevant the approach is to real-world contexts. TRNSYS was also used by Valdiserri et al. [71] to evaluate the thermal demand reduction of a tertiary building in Italy, due to an improvement of the thermal envelope and installation of high efficiency windows. An investment cost analysis was also performed, using the Net Present Value (NPV) method.

5.1.2. BES combined with data collected from bills and questionnaires

Another frequently applied method to predict the energy savings of specific energy efficiency measures is to use building energy simulation tools and compare the simulation with the real consumption obtained from metering or energy bills. Suastegui et al. [72] used this method to evaluate potential savings in the residential sector in Mexico due to replacement of oversized HVAC units. A sample of 300 houses was analysed and questionnaires were used to gather data about the households size and HVAC units capacity. An energy simulation of these buildings was then run using a model based on the Transfer Function Method. The model provides the optimal HVAC sizing for the analysed households, which is then used to calculate the kWh that could be saved in these households by replacing oversized units.

5.2. Hybrid methods

Hybrid methods make use of data-driven techniques to optimize the results obtained with deterministic methods. The reviewed approaches involve the use of different data-driven algorithms to scale up the results obtained to a higher amount of buildings, or to find the optimal solution, within the BES results, according to a given cost function.

5.2.1. BES combined with Artificial Neural Networks

This method, presented by Ascione et al. [73], proposes the use of EnergyPlus simulations and artificial neural networks to predict building energy retrofitting effects and evaluate different renovation scenarios. The approach takes advantage of the reliable and rigorous assessment of EEM impact granted by building energy simulation software and scales the results obtained to a large number of buildings, through the application of artificial neural networks. This combination grants high accuracy of results, while keeping the computational times reasonably low. The method employs two different families of ANNs, one trained with pre-retrofit building simulation data and one with post-retrofit building simulation data, the difference between the outputs is considered as the improvement due to the implemented energy retrofit. The approach was tested on office buildings built in Southern Italy in the period between 1920 and 1970, about 8800 units, representing approximately 13% of the office buildings in Italy. Three independent networks are modeled in the first family (pre-retrofit), each of them having a different output: primary energy demand for heating, primary energy demand for cooling, and percentage of annual discomfort hours. The second category of neural networks (targeting the refurbished building stock), consists of four ANNs with single output: the three networks introduced for the pre-retrofit case, plus a new network included to predict the electricity produced by photo-voltaic panels and used in the building. The accuracy of the ANNs were assessed by analysing regressions and distributions of relative error between the networks' outputs and the results obtained with EnergyPlus models. In both cases (pre and post-retrofit), the accuracy of the models showed to be quite high, with the average absolute value of relative errors ranging between 6.1% and 11%.

5.2.2. Multi-objective and multi-criteria optimization of BES data using Genetic Algorithms

In the framework of decision aid systems for energy retrofitting strategies, two very popular solutions are multi-objective and multi-criteria optimizations. Asadi et al. [74] wrote a detailed review on the topic, explaining also the conceptual distinction between multi criteria and multi objective models: in multi-criteria optimization, the group of possible alternatives is finite and explicitly known a priori, to be evaluated according to multiple criteria, while in multi-objective optimization models, the potential solutions are implicitly determined by the optimization variables and constraints. A very popular technique, frequently used by scientists in both these cases, is the genetic algorithm (GA). Following, different applications of genetic algorithms in the building energy retrofitting field are presented.

Siddharth et al. [75] built an IT tool that uses GAs to create several combinations of building variables correlated with energy consumption. For each of these combinations, the energy consumption of the building is simulated and a nonlinear regression model is fit between the system characteristics and the annual energy demand of the building. In this way, different system configurations are determined, allowing the evaluation of hypothetical energy efficiency measures. The tool was successfully tested in three different climate zones in India and the US. Genetic algorithms and other optimization techniques, such as particle swarm optimization and sequential search, were also applied by Bichiou and Krarti [76] to optimize the selection of building envelopes and HVAC systems for houses in five different US cities, with the goal of minimizing their operating costs. The comparative analysis showed that savings in computational effort could be as high as 70% when using genetic algorithms in place of particle swarm or sequential search.

Ascione et al. [77–79] also used GAs to analyse EnergyPlus simulation data in both multi-objective and multi-criteria analyses. The approach was successfully used first to determine the optimal renewable energy mix in a building and then to identify optimal energy retrofitting strategies in typical hospital and office reference buildings.

5.2.3. Multi objective optimization of BES data using NSGA-II

Chantrelle et al. [80] developed MultiOpt, a multi-criteria tool that uses NSGA-II (a non-dominated sorting genetic algorithm) [81] coupled with environmental databases and assessment software (TRNSYS), to optimize the retrofitting process of buildings across a variety of different objectives. NSGA-II was also used by Delgarm et al. [82], in combination with EnergyPlus, to analyse how different architectural parameters affect the energy consumption of a building in four different climate regions of Iran. The analysis shows that the optimization process could decrease the building's energy consumption by up to 42.2 %.

5.2.4. Multi-objective optimization of BES data using Genetic Algorithms and Artificial Neural Networks

This optimization methodology, that combines different approaches introduced in the previous paragraphs, was used by Magnier and Haghighat [83] to reduce the energy usage while keeping the optimal thermal comfort in a residential building. The approach features the use of NSGA-II to solve the optimization problem and a multilayer feed-forward ANN to reduce the time of computation required by the analysis.

More recently, Asadi et al. [84] used a similar technique to analyze TRNSYS data and identify optimal building energy retrofitting strategies. The set of possible retrofitting actions was summarized in five decision variables introduced as inputs for the ANN: external wall insulation materials, roof insulation materials, window types, solar collector types, HVAC system. The ANN, trained with building simulation results, had four different outputs: total percentage of discomfort hours, and energy demands for space heating, space coolings and sanitary hot water. A multi-objective GA was then applied to analyze the results of the ANN analysis and find the optimal solutions in terms of energy usage, renovation cost, and thermal discomfort hours.

5.2.5. Mixed-Integer Linear Programming

Iturriaga et al. [85] used a Mixed-Integer Linear Programming model to design the energy renovation of an existing building, with the goal of achieving the nearly Zero Energy Building standard. The proposed approach attempts to model the energy demand of the building through a linear model, introducing the EEMs as virtual energy sources that produce, at specific points in time, the energy that would be saved. To calculate the exact demand reduction corresponding to each EEM, dynamic TRNSYS simulations are run. The linear programming approach is then used to optimize the obtained results for the optimal cost case and the Zero Energy Building case. The method was successfully implemented to obtain the system configuration that minimizes the annual net costs for a real building located in the city of Bilbao (Spain).

5.3. Data-driven methods

The data-driven methods analysed in this section have the goal of providing recommendations for building energy retrofit by drawing conclusions based on the analysis of collected data from real use-cases.

5.3.1. User-facing Fallen Rule List using audit data

This method, presented by Marasco and Kotokosta [86], proposes the application of a fallen rule list classifier to how different building would react to different groups of EEMs. The classifier uses binary features obtained from energy audit data for over 1000 buildings in the city of New York and has the goal of providing a tool for decision-makers with the capability of either supporting, or potentially replacing, a complete energy audit. The classifier analyzes the correlation between building specific data and the EEM recommended by energy consultants after performing a building audit. The model was trained on 764 buildings and then tested on 192 buildings, showing a good overall performance for predicting the EEMs of the following categories: cooling system, distribution system, domestic hot water, fuel switching, lighting and motors, representing collectively 62% of EEMs analysed in this study.

5.3.2. Artificial Neural Networks using audit data

Beccali et al. [87] implemented artificial neural networks to create a decision aid tool able to evaluate energy performance and possible refurbishment strategies for tertiary buildings in Southern Italy. The networks were trained using audit data from 151 non-residential buildings, located in different regions of Southern Italy. The audits collected information about the buildings' geometric and equipment characteristics, as well as data about ten different proposed retrofitting actions. This data was employed to determine the ideal architecture configuration for two ANNs and for their subsequent training. One of the networks estimates the effective energy performance of any building, while the other assesses key economic indicators, allowing users to gain information about possible energy savings, payback time and investment costs per kWh saved.

5.3.3. Clustering techniques

This method is based on the assumption that clustering techniques can help in the development of renovation plans for groups of buildings that respond similarly to the application of EEMs. Geyer et al. [8] tested the application of clustering algorithms using performance-based indicators of the impact of applied measures. The impact of the measure on the considered building is described by a parameter equal to the quotient of the emission reduction caused by the measure, and the investment costs. To assess the impact of an applied EEM, different calculation methods are applied: simplified estimations, monthly sums, dynamic simulations or building energy simulations. Two different clustering methodologies are tested: hierarchical clustering and partitioning k-means clustering. A set of six different retrofit measures, as well as their combination, was simulated. To estimate their effect, simplified calculations using monitored energy consumption and geometric information about the buildings were realized. This method allows the evaluation of how buildings with different characteristics react to applied EEMs and to identify the clusters (groups of buildings) with highest priority for action. Salvalai et al. [88] also investigated the combination of clustering algorithms and building energy simulation, to evaluate optimal renovation strategies for a sample of school buildings in Northern Italy.

5.3.4. Linear regression

Walter and Sohn [89] trained a multivariate linear regression model using data contained in a large building energy database, to estimate energy savings due to the implementation of particular retrofits. The model's input parameters are both categorical and numerical variables, while the response variable is the annual source energy usage intensity (EUI). Through this method, it's possible to analyse the impact of specific building properties and installed systems on the EUI, predict for possible combinations of explanatory variables not included in the database and yield predictions that have clear and well-known statistical properties. The predictors chosen include the majority of the fields in the US Building Performance Database[90], in case of highly correlated fields only one of them is chosen. This method proves to be highly effective, as the data required to perform this type of analysis is generally affordable and easy to obtain, making this approach cheaper and faster than other methods that involve the creation of building energy simulation models.

5.3.5. Genetic algorithm combined with A graph search*

This method was examined by Yi-Kai et al. [91], with the goal of analysing all possible retrofitting actions, and their trade-offs, to identify optimal solutions. Six experienced building renovation stakeholders were interviewed to determine the assessment scores of different renovation actions, as well as the cost information for each action. Based on this data, a two-stage hybrid GAA* algorithm (combination of Genetic Algorithm and the best first (A*) algorithm) was used to test all the possible scenarios and identify the optimal solutions. This approach was compared to two commonly adopted methods: zero-one goal programming (ZOPG) and Genetic Algorithm (GA) proving be better than either of them alone.

Table 2: Characteristics of the analysed prediction and recommendation methods

Method	Type	Test buildings	Measure categories	References
BES	Deterministic	Single family houses, offices, schools, apartment blocks	Building envelope, HVAC systems, domestic hot water, PV and solar system installation, lighting system	[67, 69, 71]
BES + data from bills and questionnaires	Deterministic	Single family houses, apartment blocks	Building envelope, cooling system	[72]
BES + ANN	Hybrid	Offices	Building envelope, solar shading, ventilation system, heating system, cooling system, PV system installation	[73]
BES + GA	Hybrid	Hospitals, offices, single family houses, apartment blocks	Heating, cooling and ventilation systems, building envelope, domestic hot water, PV system installation	[75–79]
BES + NSGA-II	Hybrid	Offices, schools	Building envelope	[80, 82]
BES + GA + ANN	Hybrid	Schools, apartment blocks	Building envelope, HVAC systems, solar collectors installation	[83, 84]
Mixed integer linear programming	Hybrid	Apartment blocks	Building envelope	[85]
Linear regression	Data-driven	Commercial buildings	Building envelope, HVAC systems, lighting system	[89]
ANN with audit data	Data-driven	Schools, sport buildings, libraries, offices	Building envelope, HVAC systems, management system, lighting system, solar thermal installation, water pumping system	[87]
Clustering	Data-driven	Schools, apartment blocks	Building envelope, heating system, PV and solar thermal installation	[8, 88]
FRL with audit data	Data-driven	Apartment blocks, offices, hotels	Building envelope, HVAC systems, DHW systems, PV and solar thermal installation, lighting system, building management system	[86]
GA + A* graph search	Data-driven	Offices	Building envelope, solar shadings, water management, HVAC systems, lighting system, building management system	[91]

6. Discussion

In this article, two fundamental processes required for the improvement of building energy performance have been studied: the measurement and verification process, and the prediction and recommendation process. After describing their goals and main challenges, different methods found in literature were reviewed. The analysis was focused mainly on data-driven approaches, although for the prediction and recommendation process, deterministic methods were also considered, since their combination with data-driven techniques is becoming of increasing interest.

In the first part of the article, different methods for energy baseline estimation were reviewed. For every method, advantages and limitations were examined. The reviewed articles show that more complex methods generally provide more accurate estimations, although the bias-variance trade-off should be always kept in mind: as the models' complexity increases, they can become more accurate, but also more likely to overfit (fail to properly fit additional data, as new observations are added to the dataset) [92]. This said, it was found that different models still have different specific cases where they work best, regardless of their level of complexity. Another interesting insight that emerged from the review is that, when comparing different methods, being able to accurately determine the uncertainty of the results obtained is a very valuable feature. If the main concern of the M&V practitioner is to obtain the best possible estimation of model uncertainty, Bayesian methods seem to be the most optimal choice, as they provide accurate uncertainty estimations without assuming normally distributed errors. On the other hand, statistical learning techniques seem to be favoured when the main concern is the interpretability of the model, and machine learning techniques are most frequently employed when large amounts of data are available and the practitioner is interested in optimizing the model's predictive accuracy.

In the second part of this review, several deterministic and data-driven methods to predict the effect of energy retrofitting actions on buildings were analysed and presented. Although many of the methods reviewed use deterministic building energy simulations for this task, the analysis of the simulations' results is often performed with data-driven techniques. These approaches, classified as "hybrid" models, appear to be quite popular because of the possibility to combine the accuracy of deterministic methods and the computational efficiency of large scale optimization techniques.

In conclusion, it's important to remark that the comparison of the presented methods is no trivial work, as they were all applied in different use cases, with data of different granularity, and not using the same explanatory variables. This issue is a known problem in the building performance research community and was already pointed out by Miller [93], who proposed and worked on the creation of a public dataset from electricity meters of non-residential buildings, to test and compare prediction algorithms and feature extraction techniques [35][94].

7. Conclusions and future work

In order to improve the energy performance of the current building stock, it is essential to implement energy renovation programs. One of the main barriers to the widespread application of such programs is the lack of information regarding the impact of retrofitting actions. It appears clear that quantifying energy savings from implemented measures and determining the uncertainty of the obtained results, are two key steps towards the achievement of a more efficient built environment. The set of calculations performed to collect this data, is often referred to as the measurement and verification process. At the same time, another major task is to be able to find tailored effective renovation strategies for specific buildings or groups of buildings, in the article this process was referred to as the prediction and recommendation process.

In this review, the main methods currently utilized for these two processes were studied, with a special focus on data-driven approaches, as they are innovative techniques proving to be more effective and scalable than other traditional methods [16][20]. All of the reviewed techniques have different characteristics and have been applied in some specific cases, their characteristics were discussed in detail and then schematized in Tables 1 and 2. State-of-the-art methods to identify non-routine events and estimate uncertainty in M&V were also reviewed. Thanks to the additional analysis provided by these methods, it's possible to obtain more accurate estimates of the calculated savings and of their uncertainty.

From the review work, it was also seen that, while the M&V process seems to have a well defined structure, with different established standardization protocols and a range of published scientific articles addressing the topic . The prediction and recommendation process seems to lack such a structure and a considerable standardization effort would be needed in order to establish metrics of comparison and standardized approaches for the different methods currently in use.

Finally, it appears clear that, with more and more data being collected by automated metering infrastructure, data-driven methods are becoming a fundamental tool to plan effective strategies for the energy demand reduction of the existing building stock. For this reason, it is essential that governments and institutions quickly operate to develop policies that can facilitate the collection and analysis of building energy data.

8. Acknowledgements

This work was supported by the European Commission through the Interreg MED project SHERPA [grant number 581] and H2020 project SENSEI [grant number 847066].

References

[1] European Commission. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions-A policy framework for climate and energy in the period from 2020 to 2030, COM (2014) 15 final, Brussels.* Tech. rep. 2014.

[2] Pieter de Wilde. "The gap between predicted and measured energy performance of buildings: A framework for investigation". In: *Automation in Construction* 41 (2014), pp. 40–49.

[3] Ray Galvin. "Making the rebound effect more useful for performance evaluation of thermal retrofits of existing homes: Defining the energy savings deficit and the energy performance gap". In: *Energy and Buildings* 69 (2014), pp. 515–524.

[4] Pekka Tuominen, Krzysztof Klobut, Anne Tolman, Afi Adjei, and Marjolein de Best-Waldhober. "Energy savings potential in buildings and overcoming market barriers in member states of the European Union". In: *Energy and Buildings* 51 (2012), pp. 48–55.

[5] Constantine E. Kontokosta. "Modeling the energy retrofit decision in commercial office buildings". In: *Energy and Buildings* 131 (2016), pp. 1–20.

[6] Zhenjun Ma, Paul Cooper, Daniel Daly, and Laia Ledo. "Existing building retrofits: Methodology and state-of-the-art". In: *Energy and Buildings* 55 (2012), pp. 889–902.

[7] Paul A. Mathew, Laurel N. Dunn, Michael D. Sohn, Andrea Mercado, Claudine Custudio, and Travis Walter. "Big-data for building energy performance: Lessons from assembling a very large national database of building energy use". In: *Applied Energy* 140 (2015), pp. 85–93.

[8] Philipp Geyer, Arno Schlter, and Sasha Cisar. "Application of clustering for the development of retrofit strategies for large building stocks". In: *Advanced Engineering Informatics* 31 (2017), pp. 32–47.

[9] Wenzhuo Li, Choongwan Koo, Taehoon Hong, Jeongyoon Oh, Seung Hyun Cha, and Shengwei Wang. “A novel operation approach for the energy efficiency improvement of the HVAC system in office spaces through real-time big data analytics”. In: *Renewable and Sustainable Energy Reviews* 127 (2020), p. 109885.

[10] E. Iddio, L. Wang, Y. Thomas, G. McMorrow, and A. Denzer. “Energy efficient operation and modeling for greenhouses: A literature review”. In: *Renewable and Sustainable Energy Reviews* 117 (2020), p. 109480.

[11] Fabiano Pallonetto, Mattia De Rosa, Francesco DEttorre, and Donal P. Finn. “On the assessment and control optimisation of demand response programs in residential buildings”. In: *Renewable and Sustainable Energy Reviews* 127 (2020), p. 109861.

[12] EVO. *International Performance Measurement and Verification Protocol*. Tech. rep. 2017.

[13] Ellen Franconi, Matt Gee, Miriam Goldberg, Jessica Granderson, Tim Guiterman, Michael Li, and Brian Arthur Smith. *The Status and Promise of Advanced M&V: An Overview of M&V 2.0 Methods, Tools, and Applications*. Tech. rep. LBNL–1007125, 1350974. 2017.

[14] Colm V. Gallagher, Kevin Leahy, Peter ODonovan, Ken Bruton, and Dominic T.J. OSullivan. “Development and application of a machine learning supported methodology for measurement and verification (M&V) 2.0”. In: *Energy and Buildings* 167 (2018), pp. 8–22.

[15] Jessica Granderson and Samuel Fernandes. “The state of advanced measurement and verification technology and industry application”. In: *The Electricity Journal* 30.8 (2017), pp. 8–16.

[16] Aurlie Foucquier, Sylvain Robert, Frdric Suard, Louis Stphan, and Arnaud Jay. “State of the art in building modelling and energy performances prediction: A review”. In: *Renewable and Sustainable Energy Reviews* 23 (2013), pp. 272–288.

[17] Hai-xiang Zhao and Frdric Magouls. “A review on the prediction of building energy consumption”. In: *Renewable and Sustainable Energy Reviews* 16.6 (2012), pp. 3586–3592.

[18] A.S. Ahmad, M.Y. Hassan, M.P. Abdullah, H.A. Rahman, F. Hussin, H. Abdullah, and R. Saidur. “A review on applications of ANN and SVM for building electrical energy consumption forecasting”. In: *Renewable and Sustainable Energy Reviews* 33 (2014), pp. 102–109.

[19] Muhammad Qamar Raza and Abbas Khosravi. “A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings”. In: *Renewable and Sustainable Energy Reviews* 50 (2015), pp. 1352–1372.

[20] Zeyu Wang and Ravi S. Srinivasan. “A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models”. In: *Renewable and Sustainable Energy Reviews* 75 (2017), pp. 796–808.

[21] B. Yildiz, J.I. Bilbao, and A.B. Sproul. “A review and analysis of regression and machine learning models on commercial building electricity load forecasting”. In: *Renewable and Sustainable Energy Reviews* 73 (2017), pp. 1104–1122.

[22] Kadir Amasyali and Nora M. El-Gohary. “A review of data-driven building energy consumption prediction studies”. In: *Renewable and Sustainable Energy Reviews* 81 (2018), pp. 1192–1205.

[23] Tanveer Ahmad, Huanxin Chen, Yabin Guo, and Jiangyu Wang. “A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review”. In: *Energy and Buildings* 165 (2018), pp. 301–320.

[24] Chirag Deb, Fan Zhang, Junjing Yang, Siew Eang Lee, and Kwok Wei Shah. “A review on time series forecasting techniques for building energy consumption”. In: *Renewable and Sustainable Energy Reviews* 74 (2017), pp. 902–924.

[25] Yixuan Wei, Xingxing Zhang, Yong Shi, Liang Xia, Song Pan, Jinshun Wu, Mengjie Han, and Xiaoyun Zhao. “A review of data-driven approaches for prediction and classification of building energy consumption”. In: *Renewable and Sustainable Energy Reviews* 82 (2018), pp. 1027–1047.

[26] V.S.K.V. Harish and Arun Kumar. “A review on modeling and simulation of building energy systems”. In: *Renewable and Sustainable Energy Reviews* 56 (2016), pp. 1272–1292.

[27] Sang Hoon Lee, Tianzhen Hong, Mary Ann Piette, and Sarah C. Taylor-Lange. “Energy retrofit analysis toolkits for commercial buildings: A review”. In: *Energy* 89 (2015), pp. 1087–1100.

[28] Jessica Granderson, Samir Touzani, Claudine Custodio, Michael D. Sohn, David Jump, and Samuel Fernandes. “Accuracy of automated measurement and verification (M&V) techniques for energy savings in commercial buildings”. In: *Applied Energy* 173 (2016), pp. 296–308.

[29] Jessica Granderson, Samir Touzani, Samuel Fernandes, and Cody Taylor. "Application of automated measurement and verification to utility energy efficiency program data". In: *Energy and Buildings* 142 (2017), pp. 191–199.

[30] ASHRAE. *ASHRAE Guideline 142014, Measurement of Energy, Demand, and Water Savings*. ASHRAE Atlanta, 2014.

[31] Charles W. Kurnik, Michael Li, Hossein Haeri, and Arlis Reynolds. *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures: January 2012 - September 2016*. Tech. rep. NREL/SR7A40-70472, 1463449. 2018.

[32] CalTRACK Initiative. *CalTRACK*.

[33] Margaret F. Fels. "PRISM: An introduction". In: *Energy and Buildings* 9.1-2 (1986), pp. 5–18.

[34] Johanna L. Mathieu, Phillip N. Price, Sila Kiliccote, and Mary Ann Piette. "Quantifying Changes in Building Electricity Use, With Application to Demand Response". In: *IEEE Transactions on Smart Grid* 2.3 (2011), pp. 507–518.

[35] Clayton Miller and Forrest Meggers. "The Building Data Genome Project: An open, public data set from non-residential building electrical meters". In: *Energy Procedia* 122 (2017), pp. 439–444.

[36] Samir Touzani, Jessica Granderson, and Samuel Fernandes. "Gradient boosting machine for modeling the energy consumption of commercial buildings". In: *Energy and Buildings* 158 (2018), pp. 1533–1543.

[37] Samir Touzani, Jessica Granderson, David Jump, and Derrick Rebello. "Evaluation of methods to assess the uncertainty in estimated energy savings". In: *Energy and Buildings* 193 (2019), pp. 216–225.

[38] Suhaidi Mohd Aris, Nofri Yenita Dahlan, Mohd Nasrun Mohd Nawi, Tengku Ahmad Nizam, and Mohamad Zamhari Tahir. "Quantifying energy savings for retrofit centralized HVAC systems at Selangor state secretary complex". In: *Jurnal Teknologi* 77.5 (2015).

[39] Haidong Wang, Yuantao Xue, and Yu Mu. "Assessment of energy savings by mechanical system retrofit of existing buildings". In: *Procedia Engineering* 205 (2017), pp. 2370–2377.

[40] J. Kelly Kissock and Carl Eger. "Measuring industrial energy savings". In: *Applied Energy* 85.5 (2008), pp. 347–361.

[41] Matthew Brown, Chris Barrington-Leigh, and Zosia Brown. “Kernel regression for real-time building energy analysis”. In: *Journal of Building Performance Simulation* 5.4 (2012), pp. 263–276.

[42] M.J. Jimnez and M.R. Heras. “Application of multi-output ARX models for estimation of the U and g values of building components in outdoor testing”. In: *Solar Energy* 79.3 (2005), pp. 302–310.

[43] M.J. Jimnez, H. Madsen, and K.K. Andersen. “Identification of the main thermal characteristics of building components using MATLAB”. In: *Building and Environment* 43.2 (2008), pp. 170–180.

[44] M.J. Jimnez, B. Porcar, and M.R. Heras. “Estimation of building component UA and gA from outdoor tests in warm and moderate weather conditions”. In: *Solar Energy* 82.7 (2008), pp. 573–587.

[45] M.J. Jimnez, B. Porcar, and M.R. Heras. “Application of different dynamic analysis approaches to the estimation of the building component U value”. In: *Building and Environment* 44.2 (2009), pp. 361–367.

[46] Julin Arco Daz, Jos Snchez Ramos, M. Carmen Guerrero Delgado, David Hidalgo Garca, Francisco Gil Montoya, and Servando lvarez Domnguez. “A daily baseline model based on transfer functions for the verification of energy saving. A case study of the administration room at the Palacio de la Madraza, Granada”. In: *Applied Energy* 224 (2018), pp. 538–549.

[47] Melek Yalcintas. “Energy-savings predictions for building-equipment retrofits”. In: *Energy and Buildings* 40.12 (2008), pp. 2111–2120.

[48] Wan n Nazirah Wan Md Adna, Nofri Yenita Dahlan, and Ismail Musirin. “Development of Hybrid Artificial Neural Network for Quantifying Energy Saving using Measurement and Verification”. In: *Indonesian Journal of Electrical Engineering and Computer Science* 8.1 (2017), p. 137.

[49] Chenchen Chang, Jing Zhao, and Neng Zhu. “Energy saving effect prediction and post evaluation of air-conditioning system in public buildings”. In: *Energy and Buildings* 43.11 (2011), pp. 3243–3249.

[50] Bing Dong, Cheng Cao, and Siew Eang Lee. “Applying support vector machines to predict building energy consumption in tropical region”. In: *Energy and Buildings* 37.5 (2005), pp. 545–553.

[51] Abdiansah Abdiansah and Retantyo Wardoyo. “Time Complexity Analysis of Support Vector Machines (SVM) in LibSVM”. In: *International Journal of Computer Applications* 128.3 (2015), pp. 28–34.

[52] Muhammad Waseem Ahmad, Monjur Mourshed, and Yacine Rezgui. “Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption”. In: *Energy and Buildings* 147 (2017), pp. 77–89.

[53] Daniel B. Araya, Katarina Grolinger, Hany F. ElYamany, Miriam A.M. Capretz, and Girma Bitsuamlak. “An ensemble learning framework for anomaly detection in building energy consumption”. In: *Energy and Buildings* 144 (2017), pp. 191–206.

[54] Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian Data Analysis, Third Edition*. 3rd. Chapman & Hall, 2014.

[55] Herman Carstens, Xiaohua Xia, and Sarma Yadavalli. “Bayesian Energy Measurement and Verification Analysis”. In: *Energies* 11.2 (2018), p. 380.

[56] John A Shonder and Piljae Im. “Bayesian Analysis of Savings from Retrofit Projects”. In: *ASHRAE Transactions* (2012), p. 14.

[57] David Lindelf, Mohammad Alisafaee, Pierluca Bors, Christian Grigis, and Jean Viaene. “Bayesian verification of an energy conservation measure”. In: *Energy and Buildings* 171 (2018), pp. 1–10.

[58] Yeonsook Heo and Victor M. Zavala. “Gaussian process modeling for measurement and verification of building energy savings”. In: *Energy and Buildings* 53 (2012), pp. 7–18.

[59] Michael C. Burkhart, Yeonsook Heo, and Victor M. Zavala. “Measurement and verification of building systems under uncertain data: A Gaussian process modeling approach”. In: *Energy and Buildings* 75 (2014), pp. 189–198.

[60] Jacques Maritz, Foster Lubbe, and Louis Lagrange. “A Practical Guide to Gaussian Process Regression for Energy Measurement and Verification within the Bayesian Framework”. In: *Energies* 11.4 (2018), p. 935.

[61] Abhishek Srivastav, Ashutosh Tewari, and Bing Dong. “Baseline building energy modeling and localized uncertainty quantification using Gaussian mixture models”. In: *Energy and Buildings* 65 (2013), pp. 438–447.

[62] Samir Touzani, Baptiste Ravache, Eliot Crowe, and Jessica Granderson. “Statistical change detection of building energy consumption: Applications to savings estimation”. In: *Energy and Buildings* 185 (2019), pp. 123–136.

[63] Wei Tian, Yeonsook Heo, Pieter de Wilde, Zhanyong Li, Da Yan, Cheol Soo Park, Xiaohang Feng, and Godfried Augenbroe. “A review of uncertainty analysis in building energy assessment”. In: *Renewable and Sustainable Energy Reviews* 93 (2018), pp. 285–301.

[64] T. Agami Reddy and David Claridge. “Uncertainty of Measured Energy Savings from Statistical Baseline Models”. In: *HVAC&R Research* 6.1 (2000), pp. 3–20.

[65] Bill Koran, Erik Boyer, M. Sami Khawaja, Josh Rushton, and Jim Stewart. “A Comparison of Approaches to Estimating the Time-Aggregated Uncertainty of Savings Estimated from Meter Data”. In: (2017), p. 13.

[66] Zadok Olinga, Xiaohua Xia, and Xianming Ye. “A cost-effective approach to handle measurement and verification uncertainties of energy savings”. In: *Energy* 141 (2017), pp. 1600–1609.

[67] Paolo Zangheri, Roberto Armani, Marco Pietrobon, and Lorenzo Pagliano. “Identification of cost-optimal and NZEB refurbishment levels for representative climates and building typologies across Europe”. In: *Energy Efficiency* 11.2 (2018), pp. 337–369.

[68] EnergyPlus Development Team. “EnergyPlus engineering reference: The reference to EnergyPlus calculations”. In: *EnergyPlus Version 9.1.0 Documentation* (2019), p. 1748.

[69] A.M. Rysanek and R. Choudhary. “Optimum building energy retrofits under technical and economic uncertainty”. In: *Energy and Buildings* 57 (2013), pp. 324–337.

[70] University of Wisconsin–Madison. Solar Energy Laboratory. *TRNSYS, a transient simulation program.* Madison, Wis. : The Laboratory, 1975., 1975.

[71] Paolo Valdiserri, Cesare Biserni, Giacomo Tosi, and Massimo Garai. “Retrofit Strategies Applied to a Tertiary Building Assisted by Trnsys Energy Simulation Tool”. In: *Energy Procedia* 78 (2015), pp. 765–770.

[72] J.A. Sustegui Macas, C. Prez Tello, A. Acua Ramrez, A.A. Lambert Arista, H.D. Magaa Almaguer, P.F. Rosales Escobedo, and A.H. Ruelas Puente. “Assessment of electrical saving from energy efficiency programs in the residential sector in Mexicali, Mexico”. In: *Sustainable Cities and Society* 38 (2018), pp. 795–805.

[73] Fabrizio Ascione, Nicola Bianco, Claudio De Stasio, Gerardo Maria Mauro, and Giuseppe Peter Vanoli. “Artificial neural networks to predict energy performance and retrofit scenarios for any member of a building category: A novel approach”. In: *Energy* 118 (2017), pp. 999–1017.

[74] Ehsan Asadi, M. Gameiro da Silva, C. Henggeler Antunes, and Lus Dias. “State of the Art on Retrofit Strategies Selection Using Multi-objective Optimization and Genetic Algorithms”. In: *Nearly Zero Energy Building Refurbishment*. London: Springer London, 2013, pp. 279–297.

[75] V. Siddharth, P.V. Ramakrishna, T. Geetha, and Anand Sivasubramaniam. “Automatic generation of energy conservation measures in buildings using genetic algorithms”. In: *Energy and Buildings* 43.10 (2011), pp. 2718–2726.

[76] Youssef Bichiou and Moncef Krarti. “Optimization of envelope and HVAC systems selection for residential buildings”. In: *Energy and Buildings* 43.12 (2011), pp. 3373–3382.

[77] Fabrizio Ascione, Nicola Bianco, Rosa Francesca De Masi, Claudio De Stasio, Gerardo Maria Mauro, and Giuseppe Peter Vanoli. “Multi-objective optimization of the renewable energy mix for a building”. In: *Applied Thermal Engineering* 101 (2016), pp. 612–621.

[78] Fabrizio Ascione, Nicola Bianco, Claudio De Stasio, Gerardo Maria Mauro, and Giuseppe Peter Vanoli. “Multi-stage and multi-objective optimization for energy retrofitting a developed hospital reference building: A new approach to assess cost-optimality”. In: *Applied Energy* 174 (2016), pp. 37–68.

[79] Fabrizio Ascione, Nicola Bianco, Gerardo Mauro, Davide Napolitano, and Giuseppe Vanoli. “A Multi-Criteria Approach to Achieve Constrained Cost-Optimal Energy Retrofits of Buildings by Mitigating Climate Change and Urban Overheating”. In: *Climate* 6.2 (2018), p. 37.

[80] Fanny Pernodet Chantrelle, Hicham Lahmidi, Werner Keilholz, Mohamed El Mankibi, and Pierre Michel. “Development of a multicriteria tool for optimizing the renovation of buildings”. In: *Applied Energy* 88.4 (2011), pp. 1386–1394.

[81] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. “A fast and elitist multiobjective genetic algorithm: NSGA-II”. In: *IEEE Transactions on Evolutionary Computation* 6.2 (2002), pp. 182–197.

[82] Navid Delgarm, Behrang Sajadi, Saeed Delgarm, and Farshad Kowsary. “A novel approach for the simulation-based optimization of the buildings energy consumption using NSGA-II: Case study in Iran”. In: *Energy and Buildings* 127 (2016), pp. 552–560.

[83] Laurent Magnier and Fariborz Haghighat. “Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and Artificial Neural Network”. In: *Building and Environment* 45.3 (2010), pp. 739–746.

[84] Ehsan Asadi, Manuel Gameiro da Silva, Carlos Henggeler Antunes, Lus Dias, and Leon Glicksman. “Multi-objective optimization for building retrofit: A model using genetic algorithm and artificial neural network and an application”. In: *Energy and Buildings* 81 (2014), pp. 444–456.

[85] E. Iturriaga, U. Aldasoro, J. Ters-Zubiaga, and A. Campos-Celador. “Optimal renovation of buildings towards the nearly Zero Energy Building standard”. In: *Energy* 160 (2018), pp. 1101–1114.

[86] Daniel E. Marasco and Constantine E. Kontokosta. “Applications of machine learning methods to identifying and predicting building retrofit opportunities”. In: *Energy and Buildings* 128 (2016), pp. 431–441.

[87] Marco Beccali, Giuseppina Ciulla, Valerio Lo Brano, Alessandra Galatioto, and Marina Bonomolo. “Artificial neural network decision support tool for assessment of the energy performance and the refurbishment actions for the non-residential building stock in Southern Italy”. In: *Energy* 137 (2017), pp. 1201–1218.

[88] Graziano Salvalai, Laura Elisabetta Malighetti, Leopoldo Luchini, and Sara Girola. “Analysis of different energy conservation strategies on existing school buildings in a Pre-Alpine Region”. In: *Energy and Buildings* 145 (2017), pp. 92–106.

[89] Travis Walter and Michael D. Sohn. “A regression-based approach to estimating retrofit savings using the Building Performance Database”. In: *Applied Energy* 179 (2016), pp. 996–1005.

[90] US DOE and Lawrence Berkeley National Laboratory. *Building Performance Database (BPD) | Department of Energy.*

[91] Yi-Kai Juan, Peng Gao, and Jie Wang. “A hybrid decision support system for sustainable office building renovation and energy performance improvement”. In: *Energy and Buildings* 42.3 (2010), pp. 290–297.

[92] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, eds. *An introduction to statistical learning: with applications in R.* Springer texts in statistics 103. OCLC: ocn828488009. New York: Springer, 2013.

[93] Clayton Miller. "More Buildings Make More Generalizable ModelsBenchmarking Prediction Methods on Open Electrical Meter Data". In: *Machine Learning and Knowledge Extraction* 1.3 (2019), pp. 974–993.

[94] Clayton Miller and Forrest Meggers. "Mining electrical meter data to predict principal building use, performance class, and operations strategy for hundreds of non-residential buildings". In: *Energy and Buildings* 156 (2017), pp. 360–373.