Achievable Energy Flexibility Forecasting of Buildings Equipped With Integrated Energy Management System

PENG ZHANG¹, XIAOXING LU², (Student Member, IEEE), AND KANGPING LI³, (Member, IEEE)

¹China Southern Power Grid, Guangzhou Branch, EHV Power Transmissions Company, Guangzhou 510663, China
²Department of Electrical Engineering, North China Electric Power University, Baoding 071003, China
³Department of Electrical Engineering, Tsinghua University, Beijing 100084, China

Corresponding authors: Peng Zhang (pengzhang@ncepu.edu.cn), Xiaoxing Lu (xiaoxinglu_xxl@ncepu.edu.cn), and Kangping Li (kangpingli@tsinghua.edu.cn)

ABSTRACT Buildings’ achievable energy flexibility refers to the real load reduction amount in an incentive-based demand response (DR) event, which presents dynamic, subjective, and uncertain characteristics. It is different from the buildings’ theoretical energy flexibility, which refers to its physical load reduction potential during a certain time period and is static and certain. The former serves as a foundation in a DR aggregator’s market transaction strategy formulation process and thereby calls on the necessity of its accurate forecasting. However, most of the existing literature focuses on the latter one, which is not necessarily equal to the real achievable flexibility. Therefore, to help DR aggregators bid accurately in the ancillary service market, this paper proposes an achievable energy flexibility forecasting method for a building equipped with an integrated energy management system based on the decision tree model. The impact of gas-fired equipment on buildings’ achievable-energy-flexibility is also taken into account in the proposed method. The case study indicates that the proposed method exhibits promising performance in forecasting the achievable energy flexibility of a building.

INDEX TERMS Demand response aggregators, achievable energy flexibility, theoretical energy flexibility, forecasting, integrated energy management system.

Abbreviations:
DR Demand response.
IBDR Incentive-based DR.
ASM Ancillary service market.
CCHP Combined cooling, heating, and power systems.

Indexes:
l Index of samples in the testing set.
k Index of IBDR events.
t Index of timeslots.

Parameters:
r^e The electricity price.
r^g The natural gas price.

Inc The economic incentive of unit load reduction in IBDR events.
P^t_{grid,baseline} The baseline load of the building.
\eta^e_{CCHP} The efficiency of the CCHP to produce electricity.
\eta^h_{CCHP} The efficiency of the CCHP to produce heat.
G^\text{min}_{CCHP} The minimum gas consumption of the CCHP.
G^\text{max}_{CCHP} The maximum gas consumption of the CCHP.
\eta_{EC} Heat production efficiency of the electric air conditioner.
P_{\text{min EC}} P_{\text{max EC}} The minimum and maximum power of the electric air conditioner.
\eta^f_{B, ch} Charging efficiency of the battery.
The heat conduction area of the building.

variables:

- $\eta_{B,\text{dch}}$: Discharging efficiency of the battery.
- $\text{Cap}_B$: The capacity of the battery.
- $\text{SOC}^\text{min}_B$: The minimum SOC of the battery.
- $\text{SOC}^\text{max}_B$: The maximum SOC of the battery.
- $P_{\text{min},\text{dch}}$: The minimum discharging power of the battery.
- $P_{\text{max},\text{dch}}$: The maximum discharging power of the battery.
- $P_{\text{min},\text{ch}}$: The minimum charging power of the battery.
- $P_{\text{max},\text{ch}}$: The maximum charging power of the battery.
- $\eta^\text{t}_\text{tan,}\text{ch}\eta^\text{t}_\text{tan,}\text{dch}$: Heat storage and release efficiency of the thermal storage tank.
- $\text{Cap}_\text{tan}_k$: The capacity of the thermal storage tank.
- $\text{SOC}^\text{min}_\text{tan}_k$: The minimum heat storage state of the thermal storage tank.
- $\text{SOC}^\text{max}_\text{tan}_k$: The maximum heat storage state of the thermal storage tank.
- $Q_{\text{min},\text{tan,}\text{ch}}$: The minimum thermal power storage of the thermal storage tank.
- $Q_{\text{max},\text{tan,}\text{ch}}$: The maximum thermal power storage of the thermal storage tank.
- $Q_{\text{min},\text{tan,}\text{dch}}$: The minimum thermal release power of the thermal storage tank.
- $Q_{\text{max},\text{tan,}\text{dch}}$: The maximum thermal release power of the thermal storage tank.
- $M_{\text{air}}$: The mass of the air in the building.
- $C_{\text{Pair}}$: The specific heat capacity of the air in the building.
- $R$: The equivalent thermal resistance of the building.
- $S$: The heat conduction area of the building.
- $\theta^\text{t}_\text{out}$: The outdoor temperature of the building.
- $P_{\text{grid}}^t$: The actual electrical load of the building.
- $G_{\text{CCHP}}^t$: The gas consumption of the CCHP.
- $P_{\text{CCHP}}^t$: The electricity generation of the CCHP.
- $Q_{\text{CCHP}}^t$: The heat generation of the CCHP.
- $X^t_{\text{CCHP}}$: A binary variable that indicates the CCHP is on or off.
- $Q_{\text{EC}}^t$: Heat generated by the electric air conditioner.
- $P_{\text{EC}}^t$: The power of the electric air conditioner.
- $X^t_{\text{EC}}$: A binary variable that indicates the electric air conditioner is on or off.
- $\text{SOC}_B^t$: The battery’s State-of-Charge.
- $P_{\text{B,}\text{ch}}^t$: The power of the battery charging.
- $P_{\text{B,}\text{dch}}^t$: The power of the battery discharging.
- $\text{SOC}^\text{t}_\text{tan}$: Heat storage state of the thermal storage tank.
- $Q_{\text{tan,}\text{ch},\text{ch}}^t$: Heat storage power of the thermal storage tank.
- $Q_{\text{tan,}\text{dch},\text{dch}}^t$: Heat release power of the thermal storage tank.
- $\theta^\text{t}_\text{in}$: The indoor temperature of the building.
- $Q^t$: The heat demand of the HVAC system.

I. INTRODUCTION

A. BACKGROUND AND MOTIVATION

Demand response (DR) takes advantage of variable electricity prices (price-based DR) or economic incentives (incentive-based DR, IBDR) to motivate changes in electricity usage of demand-side flexible resources [1], [2]. It could achieve a similar effect in maintaining the dynamic supply and demand balance with respect to the generation side but is more economical, rapid, and environmentally friendly. Therefore, DR has gained increasing attention in an environment that the generation side becomes less stable and predictable due to the large-scale integration of the variable renewable energy sources.

With the deepening reform of the electricity market, demand-side flexible resources, including industrial, commercial [3], [4], and residential customers [5], [6], are endowed with the ability to participate in the market transaction process, for example, the ancillary service market (ASM) [7], [8]. Among all the flexible providers, buildings take an indispensable part, which accounts for more than 70% of the total electricity consumption in the United States and 90% in Hong Kong [9]. Buildings could achieve electricity consumption reduction through the utilization of structural thermal mass [10], [11], heat pumps [12], [13], electric heaters [14], thermal energy storage tanks [15], batteries [16], electric vehicles [17], and so on. In addition, buildings equipped with an integrated energy management systems can reduce their electricity consumption by the utilization of gas-fired equipment (such as combined cooling, heating, and power systems, CCHP). According to a DR operation report of PJM, the contribution of buildings to load reduction in IBDR events is 53 times more than that of residences [18].

However, a single building is not eligible to directly participate in the ASM because of its relatively small flexibility compared to the large market access threshold [19], [20]. This reality gives birth to the DR aggregator, who gathers small-scale flexibility from the demand-side resources and provides them access to participate in the ASM mediately [21]. The whole transaction process of the DR aggregator is shown in Fig. 1, which can be summarized as follows: In an IBDR program, the DR aggregator firstly aggregates
multiple individual buildings by signing IBDR contracts with them. Then, it will estimate the flexibility these customers (i.e., the buildings they aggregate) could provide, based on which it submits a bid (load reduction commitments) in the ASM. If the bid is accepted and cleared by the market operator, the DR aggregator needs to implement an IBDR program to reduce its customers’ electricity consumption, for achieving the promised load reduction amount in its bid. If the real load reduction in the IBDR event is close to the promised DR amount in its bid, it will be rewarded by the ASM operator for the provided DR service. On the contrary, if the DR provision could not be fulfilled, the DR aggregator would face a penalty. Besides, if the actual DR amount is much larger than the bidding amount, the excess would not be compensated and the profit of the DR aggregator will also be affected [22], [23]. Therefore, it is necessary for a DR aggregator to accurately forecast the flexibility of its customers in an IBDR event before bidding in the ASM.

**FIGURE 1.** The schematic diagram of DR aggregators.

It is worth emphasizing that the aforementioned ‘flexibility’ refers to the buildings’ achievable energy flexibility. It is different from the theoretical energy flexibility, which comprises the overall (i.e., all facilities and devices) consumers’ potential suitable for DR and can be obtained by analyzing the technical constraints of the building. It is static and deterministic [24]. Nevertheless, the buildings’ achievable energy flexibility is defined as the achievable DR potential that takes into account the technical restrictions (e.g., the shifting time, duration, and the number of interruptions), the economic factors (e.g., DR investments and operational cost), and also psychological factors (e.g., consumers’ attitude toward load reduction and financial reward) [25]. In other words, it is the real load reduction amount in an incentive-based DR event and presents a dynamic, subjective, and uncertain characteristic due to the various willingness of the customers [26]. The achievable energy flexibility is not necessarily equal to the building’s theoretical energy flexibility. For example, the achievable energy flexibility of a building may be zero when the customer is unwilling to respond, even if the theoretical energy flexibility of the building is greater than zero [27]. Hence, it is the achievable energy flexibility that lays the foundation of the aggregator’s bidding process. Therefore, this paper aims at the accurate forecast of the buildings’ achievable energy flexibility and provides a reference for the aggregator when bidding in the ASM.

**B. RELATED WORK**

Literature review about the DR flexibility evaluation has been extensively performed along with the last years addressing different aspects of the field. Some studies focus on quantifying the energy flexibility of different types of buildings that are equipped with different types of energy components. For instance, research [28] introduced a method that allows the assessment of the energy flexibility performance of buildings equipped with electrically driven heating and cooling systems. The research in [29] introduced an energy flexibility quantification framework applicable to various energy systems commonly found in office buildings (i.e., lights, heating, ventilation, and air conditioning systems). In [30], a methodology to assess the energy flexibility of residential buildings equipped with heat pumps, renewables, thermal and electrical storage systems was presented. Authors in [31] reviewed the relevant literature at the time and compared the robustness and applicability of different energy flexibility indicators in literature, and similar work was also done in literature [32]. Unlike most studies that focus on individual buildings, the energy flexibility potential of apartment building clusters was estimated by literature [33].

Moreover, some research works put forward the quantification methods of buildings’ energy flexibility and further carry out the sensitivity analysis of the quantification results. For instance, research [34] quantified the energy flexibility of buildings during a certain time period under different scenarios (different starting times, duration, building characteristics, and weather conditions). The research in [35] shows that the reduction and increase capacity of the cooling load in buildings relies heavily on the climate zone, the length of the pre-cooling period, and the pre-cooling set point. In addition, the effect of modulating the indoor temperature set point on the buildings’ operational flexibility is investigated in [36], [37]. The influence of different electricity prices on the buildings’ energy flexibility quantification is investigated in [38], [39], among which the former focuses on an individual building while the latter concentrates on the building-group-level.

Another aspect is, building-integrated energy management systems [40], [41] that use mixed input energy sources such as natural gas and electricity to satisfy the buildings’ multi-energy needs (electricity, heat, cooling, etc.) have been widely employed in various types of buildings in recent years. The energy flexibility of these buildings depends not only on the electrical components but also on the gas components, which has been overlooked in the previous literature.

In summary, most of the existing literature focuses on quantifying the theoretical energy flexibility of a building’s electrical sector. These studies are of great significance for DR aggregators to identify and recruit buildings with high...
energy flexibility, but they are not competent for assisting the DR aggregator in accurately bid in the ASM and the flexibility from the gas sector is neglected. Since what the DR aggregator need is the information of the building’s achievable energy flexibility rather than its theoretical energy flexibility. Therefore, it is necessary to carry out works focusing on the achievable energy flexibility forecasting of a building with an integrated energy management system so as to better assist the DR aggregator in its transaction process.

C. CONTRIBUTIONS AND PAPER FRAMEWORK
Based on this context, this paper tries to answer the following questions: 1) How would the electricity and gas system couple with each other in an IBDR event? 2) How can the achievable energy flexibility of a building with an integrated energy management system be accessed? To answer these questions, a more comprehensive analysis is necessary, which not only helps to evaluate the extent to which the electricity- and gas- supply loads contribute in an IBDR event, but also to determine the building’s achievable energy flexibility that serves as a significant reference for the aggregator. Overall, the contributions of this paper can be summarized as follows:

(1) Aiming at a building integrated with an integrated energy management system, an achievable energy flexibility forecasting method is proposed based on feature selection and the decision tree algorithm, which takes into account the impact of gas-fired equipment (such as combined cooling, heating, CCHP, etc.) on buildings’ achievable energy flexibility.

(2) The achievable energy flexibility is distinguished from the theoretical energy flexibility in terms of definition and their role in the aggregator’s bidding process.

(3) The performance of the proposed methodology is verified by comparing it with the other five methods under different scenarios, including different baseline load estimation methods and different historical data lengths.

The remainder of this paper is structured as follows. Section II mathematically describes the proposed achievable energy flexibility forecasting method and presents its overall framework and implementing process. Case studies are carried out in Section III, followed by the sensitivity analysis. Finally, Section IV draws together the main conclusions and key findings of this paper.

II. METHODOLOGY
A. OVERVIEW OF THE METHODOLOGY
The schematic diagram of the achievable energy flexibility (hereinafter referred to as energy flexibility) of a building studied in this paper is illustrated as Fig. 2. Calculating the energy flexibility of a building under an IBDR event requires the information that the actual electricity consumption of the building in this IBDR event and the electricity that would have been consumed by the building in the absence of this IBDR event (which is generally referred to as baseline load) [42], [43]. When the baseline load is defined, the energy flexibility of a building studied in this paper can also be defined as the deviation between the building’s actual load and the baseline load under an IBDR.

To forecast the energy flexibility of a building, the suggested methodology is illustrated in Fig. 3. The methodology starts from collecting the historical electricity load data of a building which has been exposed to a variety of IBDR...
events (different days, different incentives). The building is equipped with an integrated energy management controller. More on this step are given in Subsection B of Section 2.

The second step is to divide the electricity load data of the building into the load data under IBDR events and the load data without IBDR events. Then, a baseline estimation method is adopted to estimate the building’s baseline load based on the load data without IBDR events. More on baseline estimation can be found in Subsection C of Section 2.

Once the baseline profiles and the actual load profiles of the building in each historical IBDR event are obtained, a set of energy flexibility data of the building can be computed according to the formula in Subsection D of Section 2. Then, it is necessary to collect the data of features that affect the energy flexibility of the building, and establish the one-to-one correspondence between the energy flexibility data and the feature data. The details of the features are described in the Subsection E of Section 2.

Finally, a machine learning algorithm is employed as the forecasting algorithm to structure a building energy flexibility forecasting model. The feature data and the energy flexibility data are used as the input and output of the machine learning algorithm for model training and testing, respectively. The details of this step are given in Subsection F of Section 2.

B. LOAD DATA COLLECTION

For the targeted building which is equipped with an integrated energy management controller, its electricity consumption is determined by the controllers. Therefore, a typical building integrated energy management model [18] is adopted in this paper to simulate the electricity load data of the targeted building. The energy flow of the targeted building in winter is presented in Fig. 4.

1) OBJECTIVE FUNCTION

The optimization objective of the building integrated energy management model is to minimize the building’s energy costs, which comprises the electricity bills, gas bills and the economic rewards obtain from IBDR events, as shown in Eq. (1).

$$
\min C = \sum_{t \in T} P_{grid}^t \cdot \Delta t \cdot r_e^t + \sum_{t \in T} G_{CCHP}^t \cdot \Delta t \cdot r_g^t - \sum_{t \in TD} (P_{grid,baseline}^t - P_{grid}^t) \cdot \Delta t \cdot Inc
$$

where $P_{grid}^t$, $G_{CCHP}^t$, and $P_{grid,baseline}^t$ represent the actual electrical load of the building, the gas consumption of the CCHP, and the baseline load of the building, respectively. $r_e^t$, $r_g^t$ and Inc denote the electricity prices, nature gas prices and the economic incentive of unit load reduction in IBDR events. $T = \{t|1, 2, \ldots, T\}$ is the set of timeslots in a day, and $TD = \{t|\text{start}, \ldots, \text{end}\}$ is the set of timeslots in an IBDR event.

To reduce the complexity of describing the building integrated energy management model, for each component, a black-box approach is used to capture its energy conversion relationship and energy conversion efficiency. In this paper, the units of electricity, gas and heat are all kW.

2) CONSTRAINTS OF THE CCHP

The electricity ($P_{CCHP}^t$) and heat ($Q_{CCHP}^t$) generation of the CCHP are computed as Eq. (2)-(3), where $G_{CCHP}^t$ is the is gas consumed by the CCHP at timeslot $t$, $\eta_{CCHP}^e$ and $\eta_{CCHP}^g$ are the heat and electricity production efficiency of the CCHP.

The operation range of the CCHP is shown in Eq. (4), where $X_{CCHP}^t$ is a binary variable that represents the on/off state of the CCHP. It indicates that the CCHP is running and 0 otherwise.

$$
P_{CCHP}^t = G_{CCHP}^t \cdot \eta_{CCHP}^e \forall t
$$

$$
Q_{CCHP}^t = G_{CCHP}^t \cdot \eta_{CCHP}^g \forall t
$$

$$
X_{CCHP}^t \cdot G_{CCHP}^t \leq X_{CCHP}^t \cdot G_{CCHP}^t \leq X_{CCHP}^t \cdot G_{CCHP}^\text{max} \quad \forall t
$$

3) CONSTRAINTS OF THE ELECTRICITY AIR CONDITION

It is assumed that the air conditioner in the building works at the heating mode in winter. The heat ($Q_{EC}^t$) generated by the air conditioner could be calculated based on the electricity consumption ($P_{EC}^t$) and the heat production efficiency ($\eta_{EC}$), as shown in Eq. (5). The power range of the air conditioner is shown in Eq. (6), where $X_{EC}^t$ is a binary variable that indicates the on/off state of the air conditioner. The maximum and minimum power of the air conditioner are represented by $P_{EC}^\text{max}$ and $P_{EC}^\text{min}$.

$$
Q_{EC}^t = P_{EC}^t \cdot \eta_{EC} \quad \forall t
$$

$$
X_{EC}^t \cdot P_{EC}^\text{min} \leq X_{EC}^t \cdot P_{EC}^t \leq X_{EC}^t \cdot P_{EC}^\text{max} \quad \forall t
$$

4) CONSTRAINTS OF THE BATTERY

Let $P_{B,ch}^t$ and $P_{B,dch}^t$ represent the charging and discharging power of the building’s battery at timeslot $t$, and SOC$_B^t$ denotes the battery’s State of Charge (SOC). The constraints of the battery include SOC constraints shown in Eq. (7)-(9) and charge/discharge power constraints shown in Eq. (10)-(12), where $Cap_B$, $\eta_{B,ch}^t$, $\eta_{B,dch}^t$, $Cap_B$ are the capacity, charging efficiency, discharging efficiency of the battery. Equation (7) represents the coupling relationship of SOC in consecutive periods. Equation (8) limits the discharge depth and guarantees that the battery is not overcharged. Constraint (9) enforces same SOC state at the end of a scheduling cycle as the initial value to facilitate daily operation. Equation (10)-(11) bound the power range that the batteries could be charged/discharged. Equation (12) means that the battery cannot be charged and discharged at the same time.

$$
SOC_{B}^{t+1} = SOC_{B}^t + \frac{P_{B,ch}^t \cdot \eta_{B,ch}^t \cdot \Delta t}{Cap_B} - \frac{P_{B,dch}^t \cdot \Delta t}{\eta_{B,dch}^t \cdot Cap_B} \quad \forall t
$$

$$
SOC_{B}^{\text{min}} \leq SOC_{B}^t \leq SOC_{B}^\text{max} \quad \forall t
$$

$$
SOC_{B}^t = SOC_{B}^1 \quad \forall t
$$
5) CONSTRAINTS OF THE THERMAL STORAGE TANK

The thermal storage tanks use water as a medium for heat. The constraints of the building’s thermal storage tank are similar to those of batteries, as shown in Eq. (13)-(17), where \( Q_{ch}^t \) and \( Q_{dch}^t \) represent the heat of the tank charging/discharging at timeslot \( t \). The capacity and state of the tank are represented by \( \text{Cap}_{tan}^t \) and \( \text{SOC}_{tan}^t \).

\[
\begin{align*}
\text{SOC}_{tan}^{t+1} &= \text{SOC}_{tan}^t + \frac{Q_{tan,k,\text{ch}}^t \cdot \eta_{tan,k,\text{ch}}^t \cdot \Delta t}{\text{Cap}_{tan}^t} \\
&\quad - \frac{Q_{tan,k,dch}^t \cdot \Delta t}{\eta_{tan,k,dch}^t \cdot \text{Cap}_{tan}^t} \quad \forall t \tag{13}
\end{align*}
\]

\[
\begin{align*}
\text{SOC}_{min}^t \leq \text{SOC}_{tan}^t \leq \text{SOC}_{max}^t \quad \forall t \tag{14}
\end{align*}
\]

\[
\begin{align*}
Q_{\text{min}}^t \leq Q_{tan,k,\text{ch}}^t \leq Q_{\text{max}}^t \quad \forall t \tag{15}
\end{align*}
\]

\[
\begin{align*}
Q_{\text{min}}^t \leq Q_{tan,k,dch}^t \leq Q_{\text{max}}^t \quad \forall t \tag{16}
\end{align*}
\]

\[
\begin{align*}
Q_{\text{ch}}^t \cdot Q_{\text{dch}}^t = 0 \quad \forall t \tag{17}
\end{align*}
\]

6) CONSTRAINTS OF THE HVAC SYSTEM

The heat demand of the heating, ventilation, and air conditioning system (HVAC) in the building can be adjusted in order to reduce electricity consumption in an IBDR event. This paper employs a linearized model [44] to describe the HVAC system, as formulated in (18)-(20), where \( \theta_{\text{in}}^t \) and \( \theta_{\text{out}}^t \) are the indoor and outdoor temperature of the building, \( M_{\text{air}}^t \) and \( C_{\text{air}}^t \) denote the mass and the specific heat capacity of the air in the building, \( R \) and \( S \) represent the equivalent thermal resistance and the heat conduction area of the building. The heat demand \( \left( Q' \right) \) of the HVAC system can be supplied by the CCHP, the air conditioner and the thermal storage tank, which could be shown in Eq. (19). Inequality Eq. (20) ensures that the indoor temperature is within the occupants’ acceptable range when adjusting the heat demand. The maximum and minimum allowable indoor temperature \( \theta_{\text{in}}^\text{max} \), \( \theta_{\text{in}}^\text{min} \) depend on the habits and response willingness of the occupants.

\[
\begin{align*}
\theta_{\text{in}}^{t+1} &= \theta_{\text{in}}^t + \frac{S \cdot (\theta_{\text{in}}^t - \theta_{\text{out}}^t) + Q' \cdot \Delta t}{M_{\text{air}}^t \cdot C_{\text{air}}^t} \quad \forall t \tag{18}
\end{align*}
\]

\[
\begin{align*}
Q' &= Q_{\text{CCHP}}^t + Q_{\text{EC}}^t + Q_{tan,k}^t \quad \forall t \tag{19}
\end{align*}
\]

\[
\begin{align*}
\theta_{\text{in}}^\text{min} \leq \theta_{\text{in}}^t \leq \theta_{\text{in}}^\text{max} \quad \forall t \tag{20}
\end{align*}
\]

7) ELECTRICITY BALANCE OF THE BUILDING

The electricity consumed by the air conditioner, the battery, and the other uncontrollable electricity load should be equal to the electricity provided by the electricity retailer, the CCHP and the battery, as shown in Eq. (21).

\[
\begin{align*}
P_{\text{grid}}^t + P_{\text{CCHP}}^t + P_{\text{B,\text{dch}}}^t &= P_{\text{B,\text{ch}}}^t + P_{\text{EC}}^t + P_{\text{UNC}}^t \quad \forall t \tag{21}
\end{align*}
\]

C. BASELINE LOAD ESTIMATION

In order to obtain the energy flexibility data of the targeted building in historical IBDR events, it is necessary to get the baseline load of the targeted building in these IBDR events. However, once the targeted building participates in an IBDR event, its baseline load would no longer exists in reality and thus cannot be obtained by smart meter measurement. Therefore, its baseline load under each IBDR event should be accurately estimated.

In recent years, machine learning algorithms such as Multiple Linear Regression, Support Vector Regression, Artificial Neural Network, Random Forests and so on have been increasingly used in baseline load estimation. The classical artificial neural network is employed in this article to estimate the baseline load of the targeted building due to its excellent performance in many applications [45], as shown in Fig. 5. Firstly, the neural network model is trained by multiple groups of input data and output data which are constructed by the load data of the building without DR, so that the neural network model can learn the load change rule of the building without DR. When estimating the baseline load of the building at timeslot \( t \) on a certain day (denoted as day \( Y \)), the non-DR load data of the building at timeslot \( t \) in the previous \( k \) days is used as input data, and the trained model can give the estimated non-DR load at timeslot \( t \) on day \( Y \), that is, the estimated value of the baseline load of the building at timeslot \( t \) on day \( Y \).

D. ENERGY FLEXIBILITY CALCULATION

After obtaining the targeted building’s baseline load and actual load at each time point of each DR event, a series of energy flexibility data of the building can be calculated,
as shown in (22).
\[
\text{Flex}^{k,t} = P_{\text{grid,baseline}}^{k,t} - P_{\text{grid}}^{k,t} \quad \forall t \in TD_k, \forall k \in K
\]  

(22)

Let \( K = \{k|1, 2, \ldots, K\} \) denote the set of IBDR events that the targeted building participated, and \( TD_k = \{t|t_{k,\text{start}}, \ldots, t_{k,\text{end}}\} \) is the set of timeslots for IBDR event \( k \), where \( \text{Flex}^{k,t}, P_{\text{grid,baseline}}, P_{\text{grid}} \) are the building’s energy flexibility, baseline load, actual load at timeslot \( t \) of IBDR event \( k \). The total amount of energy flexibility data of the building is \( |TD_1 \cup TD_2 \ldots \cup TD_K| \).

E. FEATURE SELECTION

The feature data will be used as the input to forecast the building’s energy flexibility. Therefore, the selection of proper features that can affect the energy flexibility of the building is of great significance to realize an accurate energy flexibility forecasting. The following requirements should be taken into account when selecting features:

1) The features should affect the energy flexibility of the building
2) The data of the features should be easy enough for DR aggregators to obtain in practice.
3) The impact of gas consumption on the building’s energy flexibility should be taken into account.

Therefore, in order to forecast the building’s energy flexibility at DR timeslot \( t \), the following features are selected as the input features of the forecasting model. 1) Prediction time, that is timeslot \( t \). 2) The building’s baseline load at timeslot \( t \). 3) Outdoor temperature at timeslot \( t \). 4) Average outdoor temperature of the day. 5) Electricity price at timeslot \( t \). 6) Natural gas price at timeslot \( t \). 7) The incentive of the DR event. 8) The duration of this DR event.

F. BUILDING ENERGY FLEXIBILITY FORECASTING MODEL

After the acquisition of the energy flexibility dataset and corresponding features’ dataset, they are used as output data and input data respectively to train a building energy flexibility forecasting model. The complex nonlinear mapping relationship between the energy flexibility of the building and the corresponding features will be established by the forecasting model during the training process. After the forecasting model is trained, as long as the feature data of the predicted time point is input, the forecasting model can output the building’s energy flexibility of the predicted time point. The overall forecasting framework is shown in Fig. 6.

Decision tree model [46] is one of the machine learning algorithms. In this paper, a decision tree is used as the forecasting algorithm to structure the building energy flexibility forecasting model. Fig. 7 shows an example of a trained decision tree model. As shown in Fig. 7, a trained decision tree model consists of a root node (orange box), several internal nodes (blue boxes) and several leaf nodes (green boxes). Each root node and internal node corresponds to a question, and each leaf node contains a decision result. For a given input instance, the trained decision tree model will use the questions of the root node or internal nodes to test the input instance, and then assign the input instance to the internal nodes of the next layer based on the test result. This process is repeated until the input instance reaches a leaf node, and the decision result contained in this leaf node is the output of the decision tree model.

When building owners participate in IBDR, various questions will arise in their minds to judge the external information / conditions and then they make final decisions based on the judgment results. In other words, the mapping relationship between the features and the energy flexibility of buildings is constituted by a series of test questions of the building owners. Decision trees establish the complex mapping
relationship between input and output by constructing a series of test questions. Therefore, decision trees are suitable for learning the mapping relationship between the features and the energy flexibility of buildings.

In this paper, four common error metrics, i.e., absolute percent error (APE), mean absolute percent error (MAPE), mean absolute error (MAE), and root mean square error (RMSE) are used to evaluate the forecasting accuracy of the building energy flexibility forecasting model structured in this paper, which are defined as follows.

\[
\text{APE} = \frac{|F_{\text{actual}}^l - F_{\text{forecasting}}^l|}{F_{\text{actual}}^l} \times 100\% \tag{23}
\]

\[
\text{MAPE} = \frac{1}{L} \sum_{l=1}^{L} \left( \frac{|F_{\text{actual}}^l - F_{\text{forecasting}}^l|}{F_{\text{actual}}^l} \right) \times 100\% \tag{24}
\]

\[
\text{MAE} = \frac{1}{L} \sum_{l=1}^{L} |F_{\text{actual}}^l - F_{\text{forecasting}}^l| \tag{25}
\]

\[
\text{RMSE} = \sqrt{\frac{1}{L} \sum_{l=1}^{L} \left( \frac{F_{\text{actual}}^l - F_{\text{forecasting}}^l}{F_{\text{actual}}^l} \right)^2} \tag{26}
\]

where \(F_{\text{actual}}^l\) and \(F_{\text{forecasting}}^l\) are the actual and the forecasted energy flexibility of the studied building, respectively. \(L\) is the number of samples in the testing set. The smaller the value of the error metrics, the higher the accuracy of the forecasting model.

### III. CASE STUDY

#### A. SIMULATION SETTINGS

A typical building located in Lanzhou, Gansu Province, China, is chosen as the target of this paper to validate the proposed methodology. The parameters of the building are listed in Table 1. The proposed methodology begins with the electricity load data collecting, including the load data with and without IBDR events. The non-DR load data of the building in two years (2018 to 2019) is collected firstly. Then, 120 virtual IBDR events which are randomly posted on different days in these two years are set up for the building. The load data of the building in IBDR events is speculated by the building integrated energy management model. The start time, duration, and incentive of each IBDR event are random but within a certain range, as shown in Table 2.

It is worth mentioning that this step is unnecessary if the building’s real-world load data under IBDR events can be obtained. The electricity and gas price are 0.029$/kWh and 0.113$/kWh in 2018, and changed to 0.042$/kWh and 0.106$/kWh in 2019 [47]–[49].

In order to prove the superiority of the proposed method, five benchmark methods are introduced to make a comparison with the proposed method, listed as follows.

1) **Proposed Method**: The decision tree method is used as the forecasting algorithm. The input features are shown in Section II.

2) **Benchmark A1**: A typical nonlinear regression algorithm, neural network, is used as the forecasting algorithm. The input features are the same as those in A0.

3) **Benchmark A2**: A typical linear regression algorithm, ridge regression, is used as the forecasting algorithm. The input features are the same as those in A0.

4) **Benchmark B0**: A decision tree method that has the same parameters as that in A0 is employed, while the natural gas prices are not included in its input features.

5) **Benchmark B1**: A neural network that has the same parameters as that in A1 is employed, while the natural gas prices are not included in its input features.

6) **Benchmark B2**: A ridge regression that has the same parameters as that in A2 is employed, while the natural gas prices are not included in its input features.

#### B. SIMULATION RESULTS

Table 3 presents the performance of the proposed energy flexibility forecasting models compared with the other five benchmark methods, and the best results are marked as bold. Comparing the proposed method with the benchmark method A1 and A2, it can be found that the values of all error metrics of the proposed method are smaller, that is, the forecasting accuracy of the proposed method is higher. The comparison results prove that the decision tree algorithm is more suitable to forecast the building energy flexibility than the neural

| Parameter | Value | Parameter | Value |
|-----------|-------|-----------|-------|
| \(\eta_{\text{GB}}\) | 0.3 | \(\eta_{\text{m,k,1h,11h}}\) | 0.95/0.95 |
| \(\eta_{\text{GB}}\) | 0.35 | \(\eta_{\text{m,k,1h,11h}}\) | 6000(kWh) |
| \(G_{\text{GB}}^{\text{total,max}}\) | 300/2000(kW) | \(SOC_{\text{GB}}^{\text{total,max}}\) | 0/1 |
| \(\eta_{\text{EC}}\) | 3.5 | \(G_{\text{EC}}^{\text{total,min}}\) | 0/1800(kW) |
| \(I_{\text{EC}}\) | 0/2000(kW) | \(Q_{\text{EC}}^{\text{total.min}}\) | 0/1800(kW) |
| \(\eta_{\text{b,1h,11h}}\) | 0.99/0.99 | \(M_{\text{ap}}\) | 414(t) |
| \(C_{\text{ap}}\) | 1000(kWh) | \(C_{\text{ap}}\) | 1000 |
| \(SOC_{\text{GB}}\) | 0.1/1 | \(CP_{\text{ap}}\) | (kJ/kg*K) |
| \(I_{\text{GB}}\) | 0/200(kW) | \(R\) | 0.3 |
| \(I_{\text{GB}}\) | 0/200(kW) | \(S\) | 29600(m²) |

| Method | MAPE(%) | MAE(kWh) | RMSE(kWh) |
|-------|---------|----------|-----------|
| Proposed method | 8.2 | 55.717 | 139.211 |
| Benchmark A1 | 16.4 | 127.561 | 173.052 |
| Benchmark A2 | 15.2 | 109.134 | 141.614 |
| Benchmark B0 | 11.2 | 70.741 | 169.175 |
| Benchmark B1 | 17.4 | 139.187 | 177.584 |
| Benchmark B2 | 15.9 | 125.558 | 178.723 |

| TABLE 1. Parameters of the building. |
|---|
| **Parameter** | **Value** | **Parameter** | **Value** |
| \(\eta_{\text{GB}}\) | 0.3 | \(\eta_{\text{m,k,1h,11h}}\) | 0.95/0.95 |
| \(\eta_{\text{GB}}\) | 0.35 | \(\eta_{\text{m,k,1h,11h}}\) | 6000(kWh) |
| \(G_{\text{GB}}^{\text{total,max}}\) | 300/2000(kW) | \(SOC_{\text{GB}}^{\text{total,max}}\) | 0/1 |
| \(\eta_{\text{EC}}\) | 3.5 | \(G_{\text{EC}}^{\text{total,min}}\) | 0/1800(kW) |
| \(I_{\text{EC}}\) | 0/2000(kW) | \(Q_{\text{EC}}^{\text{total.min}}\) | 0/1800(kW) |
| \(\eta_{\text{b,1h,11h}}\) | 0.99/0.99 | \(M_{\text{ap}}\) | 414(t) |
| \(C_{\text{ap}}\) | 1000(kWh) | \(C_{\text{ap}}\) | 1000 |
| \(SOC_{\text{GB}}\) | 0.1/1 | \(CP_{\text{ap}}\) | (kJ/kg*K) |
| \(I_{\text{GB}}\) | 0/200(kW) | \(R\) | 0.3 |
| \(I_{\text{GB}}\) | 0/200(kW) | \(S\) | 29600(m²) |

| TABLE 2. Parameters of the DR events. |
|---|
| **Incentive ($/kWh)** | **Time** |
| 0.3, 0.4, 0.5 | 17:00-18:00; 18:00-19:00-017; 00-19:00 |

| TABLE 3. The forecasting errors of different methods. |
|---|
| **Method** | **MAPE(%)** | **MAE(kWh)** | **RMSE(kWh)** |
| Proposed method | 8.2 | 55.717 | 139.211 |
| Benchmark A1 | 16.4 | 127.561 | 173.052 |
| Benchmark A2 | 15.2 | 109.134 | 141.614 |
| Benchmark B0 | 11.2 | 70.741 | 169.175 |
| Benchmark B1 | 17.4 | 139.187 | 177.584 |
| Benchmark B2 | 15.9 | 125.558 | 178.723 |
P. Zhang et al.: Achievable Energy Flexibility Forecasting of Buildings Equipped

FIGURE 8. Comparison of the forecasted and actual values. (a) proposed method. (b) Benchmark method A1. (c) Benchmark method A2.

FIGURE 9. The distribution of APE of the proposed method, the Benchmark method A1, and the Benchmark method A2.

network algorithm and ridge regression algorithm since the training and test data sets of the three cases are the same.

In addition, by comparing the error metrics of the proposed method and the benchmark method B0, it can be seen that the forecasting accuracy of the proposed method is higher than that of benchmark method B0. Although the same decision tree algorithm is used as the forecasting algorithms in these two scenarios, the natural gas prices are not included in the input features of benchmark method B0. For buildings with an integrated energy management system, when natural gas prices are relatively lower than that of electricity, they could choose to use CCHP for electricity supplement in IBDR events, and thus reducing the necessary purchase of electricity from the market. When natural gas prices are relatively high, they will not use CCHP during IBDR events. That is, the prices of natural gas will affect the energy flexibility of these buildings. Therefore, if the natural gas information is not contained in the energy flexibility forecasting process, the accuracy of the forecasting models will be affected. The same conclusion can be also obtained from the comparison between benchmark method A1 and B1, A2 and B2.

In order to describe the difference between the proposed method and the benchmark method A1 and A2 more intuitively, the comparison between these forecasted results is presented in the form of the line chart, as shown in Fig. 8. Furthermore, the distribution of the APE of these three methods’ is illustrated in Fig. 9, where the red, green, and orange dots represent the APE of proposed method A0, benchmark method A1 and A2, respectively. By comparing the three subgraphs in Fig. 8, one could find that the dashed blue line and the solid red line presents the most similar trend in subgraph(a), which means that the forecasted values of the proposed method are very close to the real values and outperforms the other two to a large extent. In addition, the distribution of APE in Fig. 9 also shows that the APE value of the proposed method is mostly lower than that of the benchmark method A1 and A2, and the latter two presents a quite similar trend. This result is consistent with the MAPE performance presented in Table 3, and further illustrates the validity of the proposed method.

C. DISCUSSIONS

This section will investigate how the forecasting performance of the building energy flexibility forecasting model proposed in this paper is affected by 1) different baseline estimation strategies and 2) the length of historical data.

1) THE IMPACT OF BASELINE ESTIMATION STRATEGIES

As defined in Section II, the energy flexibility of a building studied in this paper is the deviation between the building’s actual load and its baseline load in an IBDR event. The energy flexibility data of a building would be slightly different under different baseline load estimation strategies, which will in turn affect the performance of the building energy flexibility forecasting models. Here, three different baseline load estimation strategies, including neural network regression, high 4 of 5, exponential moving average [50], are introduced to investigate its influence on the forecasting performance. The corresponding results are listed in Table 4. As shown in Table 4, different MAPE could be obtained using different baseline load estimation strategies, where the neural network baseline estimation method outperforms the other

TABLE 4. The MAPE of three methods under different baseline load estimation methods.

| Methods          | Baseline estimation methods |
|------------------|-----------------------------|
|                  | Neural network | High 4 of 5 | Exponential moving average |
| Proposed method  | 0.082          | 0.096       | 0.110                     |
| Benchmark method A1 | 0.164        | 0.192       | 0.198                     |
| Benchmark method A2 | 0.152        | 0.163       | 0.170                     |

FIGURE 10. The MAPE of the proposed method, the benchmark method A1 and the benchmark method A2 under various lengths of historical data.
two. In addition, it is worth mentioning that the performance of the proposed method is still superior to the two benchmark methods under all three baseline load estimation strategies. The results on the one hand verify the applicability of the proposed method under different scenarios, and on the other hand, implicates the significance of the accurate baseline load estimation for flexibility forecasting.

2) THE IMPACT OF HISTORICAL DATA LENGTH

The length of available historical data may affect the number of energy flexibility data obtained, and thereby affect the accuracy of the building energy flexibility forecasting models. The performance of the building energy flexibility forecasting models in three methods is evaluated using various lengths of historical data including 6 months, 12 months, 18 months, and 24 months. The evaluated results are shown in Fig. 10. It can be seen that the MAPE of the three methods presents a decreasing trend with the increase of historical data length. This is because the increase in the length of historical data allows the forecasting models to learn more, thus improving the generalization ability of the forecasting model. Of course, the impact of historical data length on the performance of the energy flexibility forecasting models also depends on the frequency of DR events (that is the number of DR events per unit historical data length). Since it is set to be the same for these three cases, the final results would not be influenced.

IV. CONCLUSION

This paper proposes an achievable energy flexibility forecasting method for buildings equipped with an integrated energy management system. The proposed methodology first collects a large amount of achievable energy flexibility data of a building and selects corresponding data of features that will affect it. Then, a decision tree algorithm is adopted to establish the building energy flexibility forecasting model, which could fit the mapping relationship between the building’s achievable energy flexibility and these features. The simulation results manifest that the proposed method shows promising performance. Furthermore, the impacts of different baseline estimation strategies and the length of historical data on the performance of the proposed methodology are discussed, where the proposed method also performs the best. In conclusion, the proposed building energy flexibility forecasting model could achieve relatively high reliability, which would be helpful for the DR aggregator to realize the DR potential assessment prior to the implementation of IBDR events using available feature data, and also could achieve more precise and profitable bidding in the ancillary services market. Furthermore, with the widespread installation of distributed PV in buildings, the uncertainty of PV output will significantly enhance the difficulty of forecasting the buildings’ achievable energy flexibility. We plan to improve our methods in future work to adapt to such case.

REFERENCES

[1] Y. Liu, L. Xiao, G. Yao, and S. Bu, “Pricing-based demand response for a smart home with various types of household appliances considering customer satisfaction,” IEEE Access, vol. 7, pp. 86463–86472, 2019, doi: 10.1109/ACCESS.2019.2924410.
[2] D. Liu, Y. Sun, Y. Qu, B. Li, and Y. Xu, “Analysis and accurate prediction of user’s response behavior in incentive-based demand response,” IEEE Access, vol. 7, pp. 3170–3180, 2018, doi: 10.1109/ACCESS.2018.2889500.
[3] F. Wang, L. Zhou, H. Ren, X. Liu, M. Shafie-Khah, and J. P. S. Catalao, “Multi-objective optimization model of source-load-storage synergetic dispatch for a building energy management system based on TOU price demand response,” IEEE Trans. Ind. Appl., vol. 54, no. 2, pp. 1017–1028, Mar./Apr. 2018.
[4] A. Khalid, N. Javid, M. Gaizani, M. Alhussien, K. Auranzegb, and M. Ilahi, “Towards dynamic coordination among home appliances using multi-objective energy optimization for demand side management in smart buildings,” IEEE Access, vol. 6, pp. 19509–19529, 2018.
[5] I.-Y. Joo and D.-H. Choi, “Distributed optimization framework for energy management of multiple smart homes with distributed energy resources,” IEEE Access, vol. 5, pp. 15551–15560, 2017.
[6] N. Javid, I. Ullah, M. Akbar, Z. Iqbal, F. A. Khan, N. Allrje, and M. S. Alabed, “An intelligent load management system with renewable energy integration for smart homes,” IEEE Access, vol. 5, pp. 13587–13600, 2017.
[7] Z. Xiang, C. Guo, H. Wu, J. Zhang, C. Shao, and Y. Ding, “The design of provincial ancillary service markets in China under the new-round electric power industry reform,” IOP Conf. Ser., Earth Environ. Sci., vol. 227, no. 3, Mar. 2019, Art. no. 032011.
[8] Transaction Rules of Adjustable Load Participating in Ancillary Service Market. Accessed: Mar. 20, 2021. [Online]. Available: https://www.163.com/dy/article/FR3F95QN05509P99.html.
[9] X. Cao, X. Dai, and J. Liu, “Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade,” Energy Buildings, vol. 128, pp. 198–213, Sep. 2016.
[10] M. Liu and P. Heiselberg, “Energy flexibility of a nearly-zero-energy building with weather predictive control on a convective building energy system and evaluated with different metrics,” Appl. Energy, vols. 233–234, pp. 764–775, Jan. 2019.
[11] H. Johra, P. Heiselberg, and I. L. Dréau, “Influence of envelope, structural thermal mass and indoor content on the building heating energy flexibility,” Energy Buildings, vol. 183, pp. 325–339, Jan. 2019.
[12] J. Clauss and L. Georges, “Model complexity of heat pump systems to investigate the building energy flexibility and guidelines for model implementation,” Appl. Energy, vol. 255, Dec. 2019, Art. no. 113847.
[13] D. Patteweuw, G. P. Henze, and L. Helsen, “Comparison of load shifting incentives for low-energy buildings with heat pumps to attain grid flexibility benefits,” Appl. Energy, vol. 167, pp. 80–92, Apr. 2016.
[14] C. Finck, R. Li, R. Kramer, and W. Zeiler, “Quantifying demand flexibility of power-to-heat and thermal energy storage in the control of building heating systems,” Appl. Energy, vol. 209, pp. 409–425, Jan. 2018.
[15] S. Stinner, K. Huchtemann, and D. Müller, “Quantifying the operational flexibility of building energy systems with thermal energy storages,” Appl. Energy, vol. 181, pp. 140–154, Nov. 2016.
[16] Y. Zhou and S. Cao, “Quantification of energy flexibility of residential net-zero-energy buildings involved with dynamic operations of hybrid energy storages and diversified energy conversion strategies,” Sustain. Energy, Grids Netw., vol. 21, Mar. 2020, Art. no. 100304.
[17] Y. Zhou and S. Cao, “Energy flexibility investigation of advanced grid-responsive energy control strategies with the static battery and electric vehicles: A case study of a high-rise office building in Hong Kong,” Energy Convers. Manage., vol. 199, Nov. 2019, Art. no. 118888.
[18] (Apr. 2021). 2021 Demand Response Operations Markets Activity Report. [Online]. Available: https://www.ijpm.com/markets-and-operations/demand-response.
