IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 15 December 2023, accepted 22 January 2024, date of publication 25 January 2024, date of current version 1 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3358404

RESEARCH ARTICLE

Optimization of Residential Demand Response Program Cost With Consideration for Occupants Thermal Comfort and Privacy

REZA NEMATIRAD[1], (Graduate Student Member, IEEE), M. M. ARDEHALI[2], AMIR KHORSANDI[2], AND ARASH MAHMOUDIAN[3], (Student Member, IEEE)

[1]Electrical and Computer Engineering Department, Kansas State University, Manhattan, KS 66506, USA
[2]Department of Electrical Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran 1591634311, Iran
[3]College of Science and Technology, Bellevue University, Bellevue, NE 68005, USA

Corresponding author: Amir Khorsandi (a_khorsandi@aut.ac.ir)

ABSTRACT Residential consumers can optimize their participation in demand response programs (DRPs) using home energy management systems (HEMS). By automatically adjusting air conditioning (AC) setpoints and shifting certain appliances to off-peak hours, HEMS can lead to significant cost reductions. While HEMS aims to adjust AC temperature setpoints, it is important to consider the occupants thermal comfort. This study aims to develop a multi-objective model for the application of DRPs in a smart residential house. The objectives of the model are to achieve (a) reduction in electrical load demand, (b) adjustment in thermal comfort temperature setpoints, and (c) minimization of consumer costs subject to the related constraints. Determining occupancy status through HEMS provides more economic benefits and thermal comfort for consumers. However, traditional methods such as direct occupancy monitoring are often costly, inaccurate, and can intrusively collect data on residents' activities, locations, and routines, compromising their privacy. To tackle these challenges, this study introduces advanced forecasting algorithms such as random forest, light gradient boosting machine, and multilayer perception artificial neural networks to predict occupancy by utilizing indirect data sources, such as energy consumption patterns. This approach enables the prediction of residential presence without direct monitoring. However, inherent uncertainty associated with predicted parameters can compromise the effectiveness of DRPs, and potentially lead to non-optimal energy savings, jeopardizing consumer comfort, and even system instability. To address these uncertainties, this study integrates robust counterpart optimization techniques, augmented with uncertainty budgets to control uncertainty variation making them less conservative. Simulations show uncertainty increases costs by 36% and reduces AC temperature setpoints. Further, implementing DRPs reduces demand by shifting some appliances to off-peak hours and lowers costs by 13.2%.

INDEX TERMS Demand response program, thermal comfort, occupant privacy, occupancy forecasting, learning algorithms, robust counterpart optimization, multi-objective genetic algorithm.

NOMENCLATURE

IDENTIFIERS AND VARIABLES

h_i	Index of time, from 1 to H.
H_s^{start}	Start of next interval.
H_s^{end}	End of a period.
K	Index used to represent time lags in the ARX model.
S	Index used for the load points (customers) connected to each node.

The associate editor coordinating the review of this manuscript and approving it for publication was Akin Tascikaraoglu.

$\boldsymbol{A}$	Matrix of coefficients.
b	Right side vector of inequality.
c	Linear programming vector.
d	Linear programming vector.
c_c	Penalty or reward price at time h (\$/kW).
L	Number of uncertain data.
N	Set of nodes in the community circuit.
M	Set of exogenous inputs to the ARX model.
O	Objective function.
Occ	Occupancy level.
p_h^d	Consumers demand (kW).
p_h^{shift}	Shiftable loads (kW).
$p_h^{n\_shift}$	Non-shiftable loads (kW).
p_h^{mis}	Miscellaneous electrical loads (kW).
$p_h^{d'}$	Internal demand except for air conditioning power(kW).
p_h^u	Purchased power from the utility (kW).
$p_{i,j,h}^{AC}$	Air conditioning power consumption (kW).
$p_h^{d,des}$	Desired demand at time h (kW).
$T^{AC,set}$	Air conditioning temperature setpoints (°C).
$T^{AC,des}$	Air conditioning desired setpoints (°C).
$u_{s,h}$	Binary variable indicating if the shiftable load s is ON at time h(= 1: ON, 0: OFF).
U	Set of uncertain parameters.
c_h^u	Cost of electricity at time h (\$/kW).
γ	Uncertainty budget.
γ'	Uncertainty budget.
Γ	Auxiliary variable.
ς	Uncertainty adjustment factor.
α	Coefficient of the ARX model.
β	Coefficient of the ARX model.

SUPERSCRIPTS

$\sim$	Uncertain form of a certain parameter.
$\circ$	Nominal or given value of a parameter.
$\frown$	Basic shifts of uncertain parameters
d	Demand.
O	Objective function.
Ro	Robust optimization.
T	Transposed of vectors.

I. INTRODUCTION

Electricity demand has been experiencing significant growth over the past few decades, resulting in a wider supply-demand gap. To address this issue, utilities have introduced demand response programs (DRPs) [1]. The primary goal of these programs is to provide customers with opportunities to reduce their demands and electricity costs. DRPs require consumers to manage and monitor their power consumption, which is unlikely to be accomplished manually [2]. So, home energy management systems (HEMSs) have been developed to successfully implement DRPs into smart buildings. HEMS is responsible for scheduling or shifting the consumption of some appliances into off-peak hours to reduce consumer energy costs [3]. They also adjust the air conditioning (AC) temperature setpoints corresponding with DRP schemes [4]. While HEMS aims to adjust AC temperature setpoints to reduce demand, it is important to consider the occupants thermal comfort. In the existing literature, various studies have been conducted on the implementation of DRPs in smart buildings through HEMS with a particular focus on preserving occupants thermal comfort.

A. OCCUPANTS THERMAL COMFORT

Several studies have incorporated occupants thermal comfort into their optimization problems. For example, a multi-objective optimization approach is proposed in [5] for residential load scheduling in a smart grid, focusing on reducing peak load and costs while maintaining occupant thermal comfort. Another study proposes a multi-objective optimization methodology in individual neighborhoods to maximize consumer thermal comfort, reduce demand, and minimize battery operational costs [6]. The paper [7] develops a multi-objective optimization model that balances user satisfaction with various demand response types in a dynamic electricity pricing environment, for optimizing efficient energy consumption across different user categories. A multi-objective optimization is developed in [8] to enhance economic efficiency and reliability in smart integrated energy systems while addressing demand responses and consumer comfort. A proportional response strategy is used in [9] to optimize residential thermal comfort and energy costs as a result of identifying the outdoor temperature and the user preferences.

While valuable studies have been conducted to consider the occupants thermal comfort through HEMS, a significant gap remains in the literature regarding occupancy detection. For HEMS to effectively leverage DRPs to minimize costs and energy consumption, while simultaneously preserving occupants thermal comfort, HEMS must be aware of occupancy levels.

B. OCCUPANCY FORECASTING

DRP capabilities can be dynamically influenced by occupancy status [10]. Occupancy forecasting specifically refers to the advanced prediction of how many people will be present in a building and at what times. Various technologies and devices such as video cameras, infrared cameras, passive infrared sensors, and break-beam sensors can detect occupancy levels. For example, a synchronized low-energy electronically chopped passive infrared sensor is used [11] to enhance occupancy detection accuracy in indoor environments. In [12], a thermal sensor array is employed to estimate the number of people in a space. To better monitor occupancy changes, cloud-based technologies are used [13]. The study [14] introduces a new approach for estimating CO_2 levels using indoor temperature and humidity data, applied to monitoring occupancy within a residential setting. As proved in [15], Wi-Fi connection data results are more accurate than CO_2 sensors for estimating occupancy levels.

While these technologies and devices that directly monitor occupancy levels can provide valuable data for building energy management systems, they have several drawbacks. Direct monitoring technologies could potentially infer sensitive personal information about occupants activities, locations, and routines. Occupants often lack awareness and control over what data are being collected and how they are being used, leading to privacy violations [16]. Additionally, such technologies can be expensive to install and maintain, and they are limited in their ability to only estimate the current occupancy state [17]. To address these challenges, learning algorithms present an alternative to occupancy forecasting. First, they are non-intrusive and privacy-aware, as they do not directly monitor occupants. Second, these algorithms leverage historical data and other types of indirect information, such as power consumption, to which utilities have access. The approach mitigates data exchange risks since explicit occupancy information is not shared. Third, the use of learning algorithms eliminates the need for additional devices like sensors, making the solution cost-effective. Lastly, unlike direct monitoring methods, which can only estimate current occupancy levels, learning algorithms can forecast future occupancy levels, providing a more dynamic and adaptable solution.

The effectiveness of machine learning techniques has been demonstrated in several power system applications. For example, short-term photovoltaic generation forecasting is discussed in [18] based on long short-term memory (LSTM), convolutional neural networks, random forests, etc. An innovative machine learning-based combined bootstrap and cumulant method is proposed in [19] for forecasting wind turbine power outputs. Authors of [20] utilize deep learning frameworks to identify collective human behaviors in smart cities. Besides, multilayer perception artificial neural networks (MLP-ANN) with feature selection techniques are utilized in [21] for solar irradiance forecasting. However, few studies have applied these approaches to occupancy forecasting. For example, the authors in [22] propose a statistical method for occupancy detection that estimates occupancy levels. It requires historical data, such as energy consumption, temperature and humidity, modules, and various sensors, and can only detect the current occupancy status. Also, traditional machine learning algorithms such as support vector machines are used in [23] for the analysis and prediction of binary occupancy levels for residential consumers. However, these algorithms cannot deal with imbalanced data [24]. Therefore, powerful techniques such as next-generation machine learning algorithms and deep learning are required to provide a non-intrusive, accurate, and cost-effective solution to forecast occupancy levels for DRP applications.

Occupancy and load demand serve as crucial input parameters for HEMS, and errors in their estimation can significantly impact the system performance. Failing to account for such uncertainties can compromise the effectiveness of DRPs, and potentially lead to non-optimal energy savings, compromised consumer comfort, and even system instability. Therefore, it is critical to consider the associated uncertainties in the predicted parameters to develop more realistic and effective solutions.

C. UNCERTAINTY ANALYSIS

In real-world problems, changing even one data point can violate multiple constraints, resulting in a non-optimal or even impossible solution [25]. Traditionally, optimization problems have been solved by assuming the data are deterministic, despite the fact that most real-life data are uncertain. For example, occupancy uncertainty accounts for up to 30% variation in the building energy [26]. The approaches to model uncertainty can be divided into probabilistic and non-probabilistic models. Probabilistic models quantify uncertainty by assigning probability distributions to uncertain parameters. Monte Carlo simulation, stochastic programming, and Bayesian networks are among the most commonly used probabilistic techniques. These methods often require large numbers of the data and fitting the data into a known probability distribution function. On the other hand, non-probabilistic models, such as robust optimization and fuzzy logic address uncertainty by establishing parameter bounds. They are particularly useful when data is insufficient, or when probability distributions are either unknown or fitting them are statistically insignificant [27]. Due to inadequate data to fit a proper probability and avoid expensive computation, this study utilizes robust counterpart optimization methods to model the uncertain parameters. The primary approach to dealing with uncertainty with robust optimization is to evaluate the worst-case scenario. However, using the worst-case scenario approach without considering uncertainty budgets may lead to overly conservative solutions. This is because uncertain data rarely hold their worst values at the same time. In optimization problems, uncertainty budgets are used to quantify the uncertainty in the data and determine the acceptable level (not simply worst values).

D. CONTRIBUTIONS

The contribution of this study can be summarized as follows:

1- Advanced occupancy detection: This study recognized occupancy levels as a key factor in the effective deployment of DRPs via HEMS. The next-generation learning algorithms are employed to provide a non-intrusive, cost-effective, and precise forecasting method that equips HEMS with future occupancy information.

2- Optimal thermal comfort: As HEMS adjusts AC set-point temperatures it is essential to maintain occupants thermal comfort. This study elevates the importance of this aspect by explicitly defining occupants thermal comfort as a primary objective function to strike a balance between energy conservation and the thermal comfort of residents.

3- Tackling predictive variabilities: Inherent uncertainties in forecasted data can greatly influence the effectiveness of HEMS. Recognizing this challenge, this study incorporates robust counterpart optimization techniques that not only account for uncertainties but also effectively quantify them.

The summary of studies conducted on the subject is given in Table 1. This research provides a comprehensive approach to improving the effectiveness of DRPs through HEMS by integrating thermal comfort, occupancy forecasting, and uncertainty analysis. Although each component has been studied individually in previous studies, the combined impact and potential benefits of integrating them have not been fully addressed. In thermal comfort, while existing multi-objective optimization studies have effectively balanced energy efficiency with occupant comfort, they often overlook the critical aspects of occupancy status and forecasting. For occupancy detection, despite the advantages of current technologies in real-time monitoring, issues such as privacy, accuracy, and high installation costs remain unaddressed, which our approach aims to resolve using non-intrusive learning algorithms. Regarding uncertainty analysis, traditional probabilistic models, while adept at quantifying uncertainties, face challenges with large data requirements, contrasting with non-probabilistic models that, despite their utility in data-limited scenarios, may lack detailed accuracy for complex systems. The problem in our study is formulated as a multi-objective model that aims to meet the requirements for reducing electricity demand, adjusting thermal comfort AC temperature setpoints, and minimizing consumer costs subject to the related constraints. To comprehensively assess the impact of applying DRPs and uncertainty on the results, the problem is simulated for four case studies where all possible combinations of applying and not applying DRPs and uncertainty are examined thereby addressing the limitations observed in existing research and highlighting the novel contributions of our integrated approach. Besides, the organizational flowchart of the simulation procedure in this study is shown in Fig. 1.

II. METHODOLOGY

This section outlines the techniques employed in this study.

A. OCCUPANCY FORECASTING

1) DATA COLLECTION AND PREPROCESSING

The Wi-Fi connection data, which serves as an indicator of occupancy levels, were collected from June 10th, 2019 to October 7th, 2019. This dataset represents the hourly history of Wi-Fi connections in a residential setting. Typically, the house is inhabited by three people, although the occupancy may vary over time. The corresponding electrical consumption data was obtained from the local utility provider in the same time frame. To ensure the integrity and reliability of these datasets, we conducted a thorough check for NaN (Not a Number) values to identify and address any missing or undefined data points. Additionally, we verified the length and consistency of both datasets to ensure alignment without discrepancies.

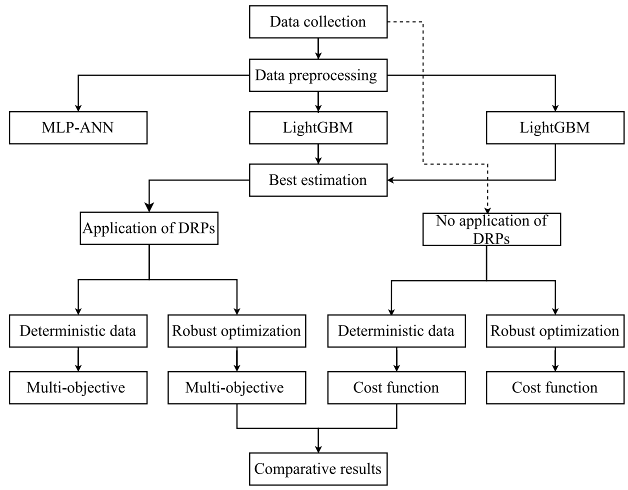


FIGURE 1. Organizational flowchart of the simulation procedure in this study.

To forecast occupancy levels based on learning algorithms, two sets of initial datasets are required, inputs (or features) and outputs (or responses). In this study, hourly electricity demand data was used as the input feature set, while Wi-Fi connection records served as the response dataset, indicating occupancy levels. Fig. 2 demonstrates the relationship between occupancy levels and electricity demand for a random day in September. It shows a clear correlation between occupancy levels and electricity demand. As occupancy increases, there is a corresponding rise in power usage, indicating a direct correlation. For instance, periods with three occupants consistently show higher electricity consumption compared to hours with fewer occupants. This pattern highlights the potential of using electricity demand as a predictive feature for accurately forecasting occupancy.

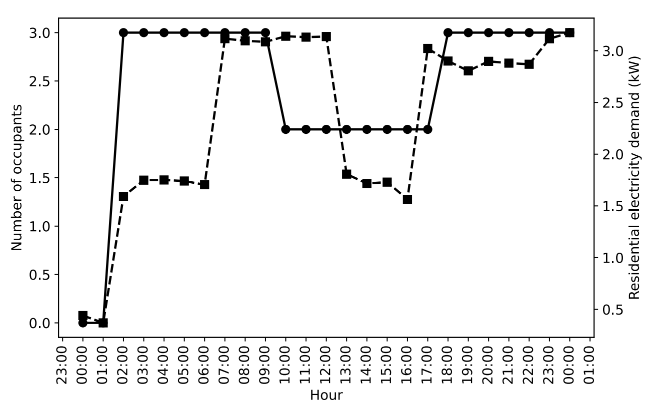


FIGURE 2. Hourly residential electricity demand and corresponding occupancy level number for a random day in September 2019.

TABLE 1. The summary of studies conducted on the subject.

Ref	Occupancy Consideration	Uncertainty Consideration	Objective functions	Method	Key Findings
[3]	×	×	Energy cost Carbon emissions Consumer comfort	Ant Colony Optimization	Simulation results demonstrate the effectiveness of the proposed framework in reducing energy costs, and carbon emissions, mitigating peak loads, and improving user comfort compared to existing frameworks.
[4]	×	×	Energy cost Thermal comfort Peak demand Electricity consumption	Mixed Integer Linear Programming	Achieve about 20% savings on electricity costs compared to baseline scenarios without HEMS. The HEMS allows for dynamic control of various household appliances and systems, including air conditioners, electric water heaters, and batteries.
[5]	×	×	Electricity cost Thermal comfort	Nondominated Sorting Genetic Algorithm	Approximately 25% reduction in peak load and 10% savings in daily electricity costs with the best demand response control strategy. Load flexibility of AC systems and household electrical appliances allows shifting of peak load to off-peak periods without significantly affecting occupant comfort.
[6]	×	×	Electricity consumption Energy cost Thermal comfort	Chebyshev Goal Programming	The framework effectively manages the power distribution system during extreme temperature events, balancing the load generation and minimizing operational costs. The solution considers the health and comfort of residents, especially vulnerable groups like children and the elderly, by conservatively using demand-responsive A/C units.
[7]	×	×	Thermal comfort Peak demand Electricity demand	Nondominated Sorting Genetic Algorithm	The model effectively balances users' satisfaction and the power grid company's profit. The model's application showed an increase in user satisfaction with electricity consumption as the load variation coefficient increased within a certain range, indicating efficient electricity task transfers under the dynamic pricing strategy.
[10]	✓	×	Economic revenue	Stackelberg Game Model	The study finds that consumers actively respond to price changes when they are at home, shifting their responsive flexible load to off-peak hours. This behavior varies among consumers, with each having different off-peak hours. The pricing scheme successfully flattened the aggregated power consumption
[11]	✓	×	Average accuracy	LSTM	LSTM models demonstrated over 95% accuracy in both controlled and uncontrolled environments in occupancy classification tasks. This algorithm is prone to false detections due to circuit noise.
[38]	×	✓	Energy cost Battery lifetime Thermal comfort Electricity demand	Chebyshev Goal Programming	The optimization model effectively ensures comfortable AC temperature set points in residential units. It is observed that considering uncertainties significantly alters the energy dispatch strategy and the values of the objective functions.
This study	✓	✓	Energy cost Thermal comfort Electricity demand	Multi-objective Genetic Algorithm	The study found that classification methods were ineffective for occupancy prediction due to imbalanced occupancy samples. The random forest regression model outperformed MLP-ANN and lightGBM. The comparative analysis revealed that participating in DRPs, even with uncertainties considered, can yield cost savings for consumers without compromising comfort.

The performance of the learning algorithms could be enhanced with data normalization techniques. Data normalization is a preprocessing technique to scale variables to a standard scale. In this study, Min-Max scaling was employed, a common practice that scales data to a fixed range, between 0 and 1 [28].

2) LEARNING ALGORITHMS

After scaling the data, they can be fed into the learning algorithms. Given the limited size of the dataset, there is a potential risk of overfitting, which could lead to poor results. To mitigate this, the study employs three advanced supervised learning algorithms, including random forest, light gradient boosting machine (LightGBM), and MLP-ANN to accurately forecast occupancy levels for effective use by HMES.

The random forest algorithm is an ensemble learning method that constructs multiple decision trees during the training phase [29]. Each of these trees is built using a random subset of electricity demand data and occupancy data. When making predictions, the algorithm aggregates the decisions of all these trees. Specifically, each tree independently predicts the occupancy levels based on electricity demand, and then these individual predictions are combined. The most frequent prediction from all trees is chosen as the final forecast for occupancy levels.

LightGBM is an advanced gradient-boosting learning method that constructs its trees sequentially, unlike random forests [30]. It creates a base tree to predict occupancy based on electricity demand. Then, it identifies areas where this prediction is inaccurate and constructs the next tree to address these errors specifically. Through this iterative refinement, LightGBM effectively captures and learns from the complex patterns and relationships in the data, resulting in highly accurate occupancy level predictions.

MLP-ANNs are a form of deep learning that consists of multiple layers of interconnected neurons [31]. Each layer in the network is trained using electricity demand data to predict occupancy levels. As the input data passes through these layers, each neuron processes the information by applying a weighted sum and a non-linear activation function. This allows the MLP-ANNs to identify and learn complex, non-linear patterns between electricity usage and occupancy levels. Besides, regularization techniques such as dropout are used to prevent overfitting [32]. The final output of the network is generated in the last layer, representing the estimated occupancy level based on the analyzed electricity demand data. It should be noted that using multiple algorithms can help identify the underlying patterns in the data and achieve better accuracy by comparing their results.

Supervised learning algorithms include regression and classification tasks that work with labeled datasets for prediction purposes. Regression is necessary in this study because many features with continuous values are lost when outputs are defined as discrete values in the classification process. This loss of information can reduce classification accuracy. Additionally, if the input data distribution is heavily imbalanced, the classification result may be non-uniform, with all outputs shifting to one class [33]. Further, to evaluate the performance of the learning algorithms in this study, root mean square error (RMSE), mean squared error (MSE), and mean absolute error (MAE) are used [34].

B. LOAD MODEL

For the purpose of this study, the internal load demand of a residential house has been categorized into four groups: AC, shiftable, non-shiftable, and miscellaneous loads [35]. Shiftable loads such as dishwashers, washers, and dryers can be operated at different time according to their patterns. Appliances like ovens, TVs, and personal computers are considered non-shiftable loads [36]. To forecast the current AC load, the methodology presented in [6] is employed, where a black-box reduced order model called auto-regressive with three exogenous inputs (outdoor temperature, occupancy, and AC temperature setpoints) and a previous cooling load is used to forecast the current AC loading given by [6]:

$$p_{i,j,h}^{AC} = \sum_{k\in K}\left[-\alpha_{i,j,k}\times p_{i,j,(h-k)}^{AC} + \sum_{m\in M}\beta_{i,j,k,m}\times Occ_{i,j,m,(h-k+1)}\right]$$
$$\forall_i \in N,\ \forall_j \in D,\ \forall_h \in H \tag{1}$$

Since this study is conducted just for one smart home, M, i, and j are set to 1. As a result, the load demand can be expressed as follows:

$$p_h^d = p_h^{shift} + p_h^{n\_shift} + p_h^{mis} + p_h^{AC}, \quad \forall_h \in H \tag{2}$$

C. ROBUST COUNTERPART OPTIMIZATION

The primary approach to dealing with uncertainty is to evaluate the worst-case scenario and conduct robust optimization based on that. However, using the worst-case scenario approach without considering uncertainty budgets may lead to overly conservative solutions. This is because uncertain data rarely hold their worst values at the same time. In optimization problems, uncertainty budgets are used to quantify the uncertainty in the data and determine what level of uncertainty is acceptable (not simply worst values).

1) ROBUST COUNTERPART

Consider the optimization problem given by:

$$O = \left\{c^T.x + d\right\} \tag{3}$$

Subjected to:

$$\boldsymbol{A}.x \leq b \tag{4}$$

If coefficients of c, d, $\boldsymbol{A}$, and b are uncertain and belong to uncertainty set U, Eq. (3) can be written as [37]:

$$O^{Ro} = \left\{\min_x \left\{c^T.x + d : \boldsymbol{A}.x \leq b\right\}\right\}_{(\boldsymbol{A},b,c,d)\in U} \tag{5}$$

The uncertainty set is affinely parameterized by a perturbation vector ς altering within a perturbation set Z. For instance, the parameter c with uncertainty can be written as [37]:

$$\tilde{c} = c^0 + \sum_{q\in Q}\varsigma_q \times \hat{c}_q : \varsigma \in Z \subset R^Q$$
$$Z = \left\{\varsigma \in R^L : |\varsigma| \leq 1,\right\} \tag{6}$$

Then the optimization problem assuming the worst-case scenario for all uncertain variables is as follows [37]:

$$O^{Ro} = \min_x \left\{\max_{(\boldsymbol{A},b,c,d)}\left[c^T.x + d\right], \boldsymbol{A}.x \leq b, (\boldsymbol{A}, b, c, d) \in U\right\} \tag{7}$$

2) UNCERTAINTY BUDGET BOX

While it is highly unlikely that all uncertain parameters coincide, by allocating budgets to uncertain parameters, the deviation from their nominal value can be controlled and limited. So, by taking budgets into account, the perturbation set Z can be rewritten as [37]:

$$Z = Z^0 + \sum_{l=1}^{L}\varsigma_l[Z^l] \tag{8}$$

To develop robust counterpart optimization in the general case, the robust counterparts of the objective functions and constraints must be obtained. If $\boldsymbol{A}$ and b are uncertain, they can be rewritten as follows using (6):

$$\boldsymbol{A} = \boldsymbol{A}^0 + \sum_{l=1}^{L} \varsigma_l [\boldsymbol{A}^l] \tag{9}$$

$$b = b^0 + \sum_{l=1}^{L} \varsigma_l [b^l] \tag{10}$$

Therefore, by using Eqs. (9) and (10), Eq. (4) is written as follows:

$$\left[\boldsymbol{A}^0\right]^T x + \sum_{l=1}^{L} \zeta_l \left[\boldsymbol{A}^l\right]^T x \leq \left[b^0\right]^T + \sum_{l=1}^{L} \zeta_l \left[b^l\right]^T \tag{11}$$

$$\sum_{l=1}^{L} \zeta_l \left(\left[\boldsymbol{A}^l\right]^T x - \left[b^l\right]\right) \leq \left[b^0\right] - \left[\boldsymbol{A}^0\right]^T x \tag{12}$$

The left-hand side of Eq. (12) can be written as follows [37]:

$$\sum_{l=1}^{L} \zeta_l \left(\left[\boldsymbol{A}^l\right]^T x - \left[b^l\right]\right) = \sum_{l=1}^{L} \left|\left[\boldsymbol{A}^l\right]^T x - \left[b^l\right]\right| \tag{13}$$

Accordingly, by integrating Eq. (13) into Eq. (12), Eq. (14) is obtained.

$$\left[\boldsymbol{A}^0\right]^T x + \sum_{l=1}^{L} \left|\left[\boldsymbol{A}^l\right]^T x - \left[b^l\right]^T\right| \leq \left[b^0\right] \tag{14}$$

Finally, Eq. (14) is linearized using axillary parameters Z and W. The uncertainty budget γ is applied to be less conservative as [37]:

$$\sum_{l=1}^{L} |Z_l| + \gamma . \max |W_l| + \left[\boldsymbol{A}^0\right]^T x \leq b^0 \tag{15}$$

where [37]:

$$Z_l + W_l = b^l - \left[\boldsymbol{A}^l\right]^T x, \quad l = 1, 2, \ldots, \ L \tag{16}$$

Generally, objective functions also have uncertain parameters and are modeled by auxiliary variables. Therefore, by using an auxiliary variable Γ^O, the robust counterpart of Eq. (3) can be written as follows [37]:

$$O^{RO} = \min \Gamma^O \tag{17}$$

If c is uncertain by introducing auxiliary parameters $Z', \ W'$ for uncertain parameter c, and considering budget γ', Eq. (17) is converted as follows:

$$\Gamma^O \geq c^T x + d + \gamma' . \max \left|W'_l\right| + \sum_{l=1}^{L} \left|Z'_l\right| \tag{18}$$

$$Z' + W' = \Delta O \tag{19}$$

Eq. (18) can be linearized, for instance, absolute value W can be linearized as follows:

$$\max |W| \leq T \tag{20}$$

$$-T \leq W \leq T \tag{21}$$

III. CASE STUDIES

To apply DRPs, it is assumed that the utility offers two options to consumers: first, the ability to adjust their indoor temperature by selecting desired AC temperature setpoints, and second, the option to shift selected appliance operations from peak hours to off-peak hours. Consumers who participate in DRPs receive financial rewards or face penalties. In this study, four cases are developed to comprehensively examine the impact of uncertainty in demand and occupancy data and the application of DRPs.

Case (a): This scenario operates the HEMS under standard conditions, where neither uncertainty in demand and occupancy data nor DRPs are considered. It serves as a baseline, illustrating the HEMS functionality in a controlled environment without external incentives or unpredictability.

Case (b): In this scenario, the HEMS accounts for uncertainties in forecasted occupancy levels and load demand enhancing its adaptability to unpredictable variations. However, this case does not involve the HEMS participating in DRPs, thereby focusing solely on the HEMS capabilities in the face of data uncertainties.

Case (c): Here, the HEMS actively participates in DRPs, allowing consumers to adjust AC temperature setpoints and shift appliance usage to off-peak hours. However, this scenario does not incorporate strategies for handling uncertainties in demand and occupancy data, thereby assessing the effectiveness of DRP participation under stable and deterministic conditions.

Case (d): This comprehensive scenario combines the application of DRPs with the consideration of uncertainties in occupancy and load demand. It represents an advanced implementation of the HEMS, showcasing how the system can optimize energy management by responding to both DRP incentives and data variability.

A. CASE (A)

As all data are assumed deterministic and there is no DRP contract between consumers and the utility, the consumers energy cost is as follows:

$$\mathrm{Cos}\, t = \sum_{h \in H} c_h^u p_h^d \tag{22}$$

B. CASE (B)

As this case examines the impact of uncertainties in data for demand and occupancy levels without any application of DRPs, based on the methodology described in the robust counterpart optimization with uncertainty budgets section, Eq. (22) can be written as an uncertain optimization problem given by:

$$O_{case\ (b)} = \min \left\{ \sum_{h \in H} c_h^u p_h^d \right\} \tag{23}$$

According to Eqs. (17)-(21), by applying auxiliary parameter Γ^1, Eq. (23) can be rewritten as follows:

$$O_{case\ (b)}^{RO} = \min \left\{ \Gamma^1 \right\} \tag{24}$$

where $O_{case\,(b)}^{RO}$ is robust counterpart of $O_{case\,(b)}$ subject to:

$$\Gamma^1 \geq \sum_{h \in H} c_h^u \, p_h^d \tag{25}$$

$$\forall_h \in H,\ Z_h^{p^d} = \left\{ \varsigma^{p^d} \in R^H : \left\| \varsigma^{p^d} \right\|_\infty \leq 1,\ \left\| \varsigma^{p^d} \right\|_1 \leq \gamma_h^{p^d} \right\} \tag{26}$$

$$\Gamma^1 \geq \sum_{h \in H} c_h^u \, p_h^d + \gamma_{h,case\ (b)}^{p^d} \cdot \max(\left| W_{h,case\ (b)}^{p^d} \right|) + \sum_{h \in H} \left| Z_{h,case\,(b)}^{p^d} \right| \tag{27}$$

$$W_h^{p^d} + Z_h^{p^d} = -\Delta p_h^d . c_h^u \tag{28}$$

Eq. (27) can be easily linearized by using Eqs. (20) and (21). Also, a power balance constraint is required to model the total internal load:

$$p_h^u = p_h^d, \quad \forall_h \in H \tag{29}$$

Since considering uncertainties increases the margin of safety, the power balance equality in Eq. (29) can be expressed as [37]:

$$p_h^u \geq p_h^d, \quad \forall_h \in H \tag{30}$$

Since consumer demand includes occupancy levels, Eq. (2) can be rewritten as follows:

$$p_h^d = p_h^{d'} + p_h^{AC} \tag{31}$$

$$p_h^{d'} = p_h^{shift} + p_h^{n-shift} + p_h^{mis} \tag{32}$$

By combining Eqs. (1) and (2) with Eqs. (31) and (32), the power balance constraint can be rewritten as:

$$p_h^u \geq \sum_{k \in K} \left[-\alpha_k \times p_{(h-k)}^{AC} + \beta_k \times Occ_{(h-k+1)} \right] + p_h^{d'},\ \forall_h \in H \tag{33}$$

According to Eqs. (14)-(16), the robust counterpart of Eq. (33) can be derived as:

$$\begin{aligned} p_h^u \geq & \sum_{k \in K} \left[-\alpha_k \times p_{(h-k)}^{AC} + \beta_k \times Occ_{(h-k+1)} \right] + p_h^{d'} \\ & + \gamma_{h,case\,(b)}^{p^{d'}} \cdot \max(\left| W_{h,case\,(b)}^{p^{d'}} \right|) + \left| Z_{h,case\,(b)}^{p^{d'}} \right| \\ & + \gamma_{h,case\,(b)}^{occ} \cdot \max(\left| W_{h,case\,(b)}^{occ} \right|) + \left| Z_{h,case\,(b)}^{occ} \right| \end{aligned} \tag{34}$$

$$\forall_h \in H : Z_{h,case\,(b)}^{p^{d'}} + W_{h,case\,(b)}^{p^{d'}} = \Delta p_h^{d'} \tag{35}$$

$$\forall_h \in H : Z_{h,case\,(b)}^{occ} + W_{h,case\,(b)}^{occ} = \sum_{k \in K} \beta_k \times \Delta Occ_{(h-k+1)} \tag{36}$$

Notice that Eq. (34) is linearized based on Eqs. (20) and (21).

C. CASE (C)

In this case, consumers participate in DRPs without considering uncertainties in data, and a multi-objective optimization is developed.

1) OBJECTIVE FUNCTIONS

The problem is formulated as a multi-objective optimization where (a) electricity load demand reduction, (b) adjustment in occupants thermal comfort AC temperature setpoints, and (c) minimizing consumer costs are achieved subject to the related constraints. Typically, during peak hours, the HEMS requires consumers to match their load demand with the expected load demand and minimize the difference between them at time h. This can be accomplished by adjusting the AC setpoint or shifting loads from peak hours to off-peak hours. Therefore, the load demand reduction objective function can be defined as [38]:

$$O_{case\ (c)}^1 = \min \sum_{h \in H} \left| p_h^d - p_h^{d,des} \right| \tag{37}$$

While HEMS effectively aims to reduce consumer demand, it might compromise occupants thermal comfort if it elevates the AC temperature setpoints beyond occupants desired thermal comfort levels. To minimize dissatisfaction regarding the AC temperature setpoints, the difference between the desired and AC temperature setpoint must be minimized. Accordingly, the occupant thermal comfort objective function is expressed as follows:

$$O_{case\ (c)}^2 = \min \sum_{h \in H} \left(T^{AC,set} - T^{AC,des} \right) . Occ_h \tag{38}$$

In addition, participation in DRPs should result in a reduction in customer costs. If consumers decrease their load demand relative to the utility desired load demand, they earn incentive fees. Otherwise, they will be monetarily penalized. This has been expressed as the consumer cost objective function given by:

$$O_{case\ (c)}^3 = \min \sum_{h \in H} c_h^u \, p_h^d + c_c \left(p_h^d - p_h^{d,des} \right) \tag{39}$$

2) CONSTRAINTS

If shiftable appliances are set to be switched on at time h, the following conditions must be met:

- They must have been off in the previous steps.
- They must remain on for the required time to complete their cycle.
- Once the cycle is completed, they must be switched off.

These constraints are formulated in Eqs. (40) and (41). Eqs. (42) and (43) state that appliances cannot start before the predetermined start time and must complete their processes before the predetermined end time. Moreover, if a demand shift is planned, the entire load will be shifted to the following time step [38]:

$$\forall_s \in S : \sum_{h \in H} u_{s,h} = N_s \tag{40}$$

$$\forall_s \in S,\ \forall_h \in H: u_{s,h+1} \geq \frac{u_{s,h}}{N_s}(N_s - \sum_{\tau=1}^{h} u_{s.\tau}) \quad (41)$$

$$\forall_s \in S: \sum_{h=1}^{H_s^{start}-1} u_{s,h} = 0 \quad (42)$$

$$\forall_s \in S: \sum_{h=H_s^{end}+1}^{H} u_{s,h} = 0 \quad (43)$$

Additionally, while optimizing, there is a risk of reaching uncomfortable temperatures for residents. To mitigate this, we set a desired AC temperature setpoint at 23.33°C for residences, ensuring it never exceeds 28.55°C as demanded by the utility [6]:

$$\forall_h \in H: \left|T_h^{AC,set} - T^{AC,des}\right| \leq 5.22 \quad (44)$$

Besides, the optimization process may consistently maintain AC temperature setpoints at their maximum value, as indicated by Eq. (44), which is not desired for occupants. To address this, we impose constraints on the total variations of AC temperature setpoints against desired AC setpoints as expressed by Eq. (45), where 19.44°C is chosen according to [6]:

$$\sum_{h=1}^{H} \left|T_h^{AC,set} - T^{AC,des}\right| \leq 19.44 \quad (45)$$

Finally, to prevent overcooling in summer, AC temperature setpoints should never be less than the desired temperature as expressed as follows [6]:

$$\forall_h \in H: T_h^{AC,set} \geq T^{AC,des} \quad (46)$$

D. CASE (D)

As described, the robust equivalent of the multi-objective function expressed in case (c) can be calculated. Since the second objective function is nonlinear, it can be linearized using an auxiliary variable denoted as y_h, as shown in Eqs. (20) and (21). In addition, the objective functions include uncertain variables. So, the auxiliary variable models their uncertainty [37]. Consequently, Eq. (37) can be rewritten as Eq. (47).

$$O_{case\ (d)}^{1,Ro} = \min \sum_{h \in H} y_h, \quad (47)$$

where,

$$\forall_h \in h: y_h \geq \left|p_h^d - p_h^{d,des}\right| \quad (48)$$

Since objective functions contain uncertain variables, auxiliary variables $\Gamma_{case\ (d)}^1$, $\Gamma_{case\ (d)}^2$, and $\Gamma_{case\ (d)}^3$ are used to obtain their robust counterpart.

$$O_{case\ (d)}^{1,Ro} = \min \Gamma_{case\ (d)}^1 \quad (49)$$

$$O_{case\ (d)}^{2,Ro} = \min \Gamma_{case\ (d)}^2 \quad (50)$$

$$O_{case\ (d)}^{3,Ro} = \min \Gamma_{case\ (d)}^3 \quad (51)$$

$$\Gamma_{case\ (d)}^1 \geq \sum_{h \in H} y_h \quad (52)$$

$$\Gamma_{case\ (d)}^2 \geq \sum_{h \in H} (T^{AC,set} - T^{AC,des}).Occ_h \quad (53)$$

$$\Gamma_{case\ (d)}^3 \geq \sum_{h \in H} (c_h^u p_h^d + c_c(p_h^d - p_h^{d,des})) \quad (54)$$

According to Eqs. (18) and (19), robust counterparts of Eqs. (52)-(54) can be derived as follows by including γ_h^{occ} and γ_h^d as occupancy level and demand budgets, respectively, and W_h^{occ}, Z_h^{occ}, W_h^d, $w_h^{d'}$, $Z_h^{d'}$, and Z_h^d as auxiliary variables.

$$\Gamma_{case\,(d)}^1 \geq \sum_{h \in H} y_h + \gamma_h^d.\max(\left|w_h^d\right|) + \sum_{h \in H} \left|Z_h^d\right| \quad (55)$$

$$\Gamma_{case\,(d)}^2 \geq \sum_{h \in H} (T^{AC,set} - T^{AC,des}).Occ_h + \gamma_h^{occ}.\max(\left|w_h^{occ}\right|) + \sum_{h \in H} \left|Z_h^{occ}\right| \quad (56)$$

$$\Gamma_{case\,(d)}^3 \geq \sum_{h \in H} (c_h^u p_h^d + c_c(p_h^d - p_h^{d,des})) + \gamma_h^d.\max(\left|w_h^{d'}\right|) + \sum_{h \in H} \left|Z_h^{d'}\right| \quad (57)$$

$$\forall_h \in H,\ w_h^{occ} + Z_h^{occ} = -d(Occ_h)\Delta(T^{AC,set} - T^{AC,des}) \quad (58)$$

$$\forall_h \in H,\ w_h^d + Z_h^d = -d(p_h^d) \quad (59)$$

$$\forall_h \in H,\ w_h^{d'} + Z_h^{d'} = -d(c_h^u p_h^d) \quad (60)$$

where perturbation set for each uncertain variable is as follows:

$$Z_h^d = \left\{\varsigma^{p^d} \in R^H : \left\|\varsigma^{p^d}\right\|_\infty \leq 1, \left\|\varsigma^{p^d}\right\|_1 \leq \gamma_h^{p^d}\right\} \quad (61)$$

$$Z_h^{occ} = \left\{\varsigma^{occ} \in R^H : \left\|\varsigma^{occ}\right\|_\infty \leq 1, \left\|\varsigma^{occ}\right\|_1 \leq \gamma_h^{occ}\right\} \quad (62)$$

It should be noted that all constraints related to this case are the same as cases (b) and (c).

Cases (c) and (d) are formulated as a multi-objective function where (a) electricity load demand reduction, (b) adjustment in occupants' thermal comfort AC temperature setpoints, and (c) minimizing consumer costs are achieved subject to the related constraints. A multi-objective function refers to a function that involves optimizing two or more conflicting objectives simultaneously [39]. The multi-objective genetic algorithm (GA) is used in this study, which has shown a great ability to solve multi-objective problems [40]. This solver was configured with carefully chosen parameters to balance and optimize these conflicting objectives effectively. Through this process, the GA algorithm iteratively evolved potential solutions. These solutions provide insights into the trade-offs between electricity demand, thermal comfort, and cost, aiding in the identification of strategies that optimally balance all three objectives.

IV. RESULTS AND DISCUSSION

All the simulations were performed on a system equipped with an Intel(R) Core (TM) i7-8700 CPU @ 3.20GHz and 16 GB of memory. Table 2 lists the execution times of the learning, simulations, and optimization processes for various stages of this study.

TABLE 2. Execution times for various simulation processes.

Model	Execution time (s)
Random forest	37.94
MLP-ANN	77
LightGBM	8.77
Case (a)	0.07
Case (b)	0.11
Case (c)	2.34
Case (d)	2.84

A. ASSUMPTION AND DATA COLLECTION

The following assumptions and approaches are considered in this study for numerical analysis.

- Washing machines and dishwashers were considered shiftable loads.
- It was assumed that indoor temperature equals ambient temperature.
- Forecasting and simulation were carried out for 12 hours (It is applicable for larger periods if needed).
- For all uncertain parameters, a box of uncertainty was defined as 0.9 to 1.1 times their nominal value [38].
- Each time slot was taken to be one hour without loss of generality, which means power demand in kW is equivalent to energy consumption in kWh.
- The financial fines and rewards are assumed to remain constant throughout all hours [38].
- It was assumed the house is demand responsive through HEMS.
- For forecasting and data preparation, Python and MATLAB were employed to solve the multi-objective problem.

B. OCCUPANCY FORECASTING RESULTS

The learning algorithms can provide non-intrusive, cost-effective, and accurate occupancy level information to able the HEMS to more effectively implements DRPs. However, in this study, the occupancy samples are heavily imbalanced. Specifically, the majority of the training samples fall under the category of resident presence, while a smaller proportion represents resident absence. In this case, the learning algorithms may face difficulties in accurately classifying the minority class. This is because it has not seen enough examples of the minority class to learn its characteristics effectively. This limitation becomes apparent as the precision, recall, and f1-score metrics for the resident absence class are zero. For instance, the parametric values for the classification process with MLP-ANN are given in Table 3, where the precision, recall, and f1-score parameters for the resident absence class are zero. This underscores the algorithm inability to correctly predict occupancy levels. Given these limitations, this study utilizes regression as a more suitable approach to address this challenge. The regression parameters for the random forest, MLP-ANN, and lightGBM algorithms are detailed in Table 3 and 4. Besides, Table 5 shows that MSE, RMSE, and MAE are 0.97, 0.99, and 0.66, respectively, for MLP-ANN regression. According to Table 5, using lightGBM improves MLP-ANN performance by 3.52%, 1.75%, and 5.47%, respectively while using the random forest algorithm improves them by 42.18%, 23.94%, and 40.6%. Therefore, random forest outperforms MLP-ANN regression and lightGBM. It is important to note that these results were obtained through multiple executions of the model, all conducted with fixed random state parameters. That means a specific seed was set for random number generation to ensure that the results of stochastic processes, like data shuffling or model initialization, are consistent in each execution.

TABLE 3. Parametric values for MLP-ANN regression and classification.

Regression parameter	Value	Classification parameter	Value
Solver	Adam	Solver	Adam
Alpha	0.0001	Alpha	0.0001
Batch_size	Auto	Batch_size	Auto
Learning_rate	0.001	Learning_rate	Constant
Power_t	0.5	Activation function	Relu
Max_iter	4000	Max_iter	4000
Shuffle	True	Shuffle	True
Random state	42	Epochs	60
Hyperparameter optimizer	Bayesian	Lost function	Binary_ crossentropy

TABLE 4. Parametric values for random forest and LightGBM regression.

Random forest parameter	Value	LightGBM parameter	Value
Criterion	Gini	Boosting type	Gbdt
Random-state	0	Random state	0
N_estimators	400	N_estimators	400
Min_impurity_decrease	0	Num-leaves	31
Max_depth	None	Max_depth	None
Bootstrap	True	Objective	Regression
Random state	0	Learning-rate	Callbacks
Max_features	Auto, Sqrt, Log2	Subsanple-for-bin	20000

TABLE 5. MSE, RMSE, and MAE for MLP-ANN, random forest, and lightGBM algorithms.

Evaluation metric	MSE	RMSE	MAE
Random forest	0.5630	0.750	0.3952
LightGBm	0.9392	0.9691	0.6290
MLP-ANN	0.9735	0.9864	0.6654

Forecasting results with the random forest algorithm are illustrated in Fig. 3. This graph shows a strong correlation between electricity demand and occupancy levels, indicating the effectiveness of using electricity demand as a predictive feature for accurately forecasting occupancy. Finally, the results derived from the random forest algorithm are employed by HEMS to optimize DRPs, ensuring consumer demand and cost reduction while preserving occupants thermal comfort.

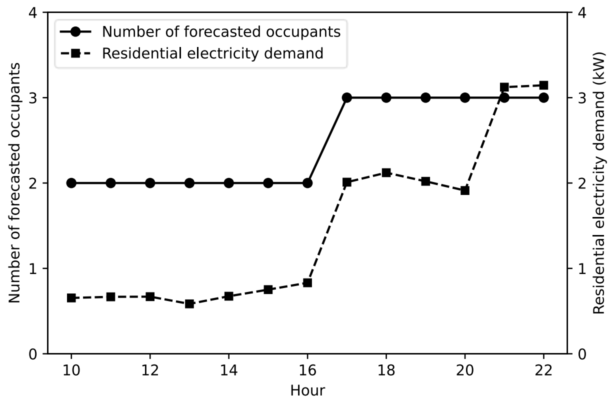


FIGURE 3. Forecasted occupants and residential electricity demand for 12 hours.

C. CASE STUDY RESULTS

1) CASE (A)

This case is considered as the benchmark without consideration of uncertainties and without applying DRPs. The total consumer cost for 12 hours is \$6.942.

2) CASE (B)

The robust counterpart optimization of case (a) is employed in case (b) incorporating uncertainty budgets. Each uncertain variable is associated with 13 distinct budgets, leading to a total of 169 possible states, considering uncertainties in demand and occupancy levels. Due to the large number of states, the analysis primarily focuses on scenarios where both uncertain parameters have identical budgets. Fig. 4 demonstrates the consumer total costs in this case. As illustrated in Fig. 4, when both uncertainty budgets are set to zero, the consumer total costs align with the costs observed in case (a). A budget of zero denotes the parameters operating at their nominal values. As the uncertainty budget increases, consumer costs increase. For instance, consumer total costs increase by 36 percent when uncertainty is considered at all hours. Additionally, robust counterparts become less conservative as budgets increase. When both uncertainty budgets transition from 0 to 1, robust counterpart optimization allocates these budgets to demand and occupancy during hours with the lowest electricity tariffs. However, uncertainty intensifies during hours with high electricity rates as the budget increases. when both uncertainty budgets transition from 11 to 12, the system experiences pronounced uncertainty. Numerically, by changing both budgets from 0 to 1 and 11 to 12 respectively, consumer costs rise by 0.08% and 5.37%. This behavior indicates that robust counterpart optimization is less conservative and can control uncertainty levels based on the system condition.

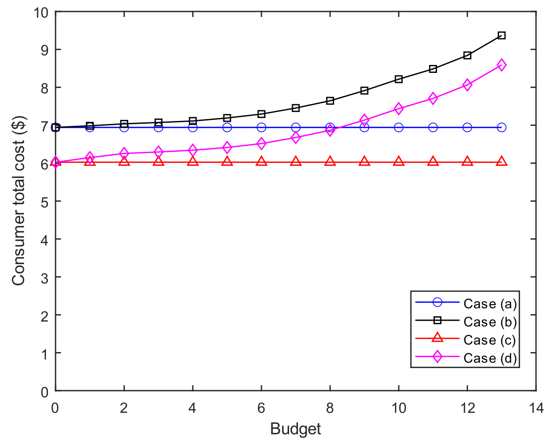


FIGURE 4. Consumer total costs in different cases. where the horizontal axis indicates occupancy and demand budgets, for example, number 3 represents that occupancy and demand budgets are both 3.

3) CASE (C)

In this case, consumers participate in DRPs without considering uncertainties. The multi-objective GA is employed to solve the optimization problem and its hyperparameters are detailed in Table 6. Besides, the multi-objective GA convergence curve is depicted in Fig. 5. Initially, the curve displays low values for the objectives, indicating that the exploration of multi-objective GA across a wide range of potential solutions. As multi-objective GA evolves and refines to produce more optimal solutions, these values progressively increase. The observed fluctuations before reaching stability highlight the algorithm dynamic process in balancing and effectively optimizing the interrelated objective functions of load demand reduction, occupant thermal comfort, and consumer cost. This trend underscores the efficacy and capability of multi-objective GA in simultaneously achieving an optimal solution across these interconnected objective functions. Besides, Table 7 shows the transferred demand for shiftable loads across different time intervals. According to Table 7, shiftable loads are transferred from the hot hours of the day, to the nighttime hours when cooling loads are less prevalent. Demand transfers arise from the multi-objective function, which seeks to minimize both consumer costs and the surplus demand discrepancy between the desired and actual consumer demand to avoid penalties. The AC temperature setpoints are shown in Fig. 6. The AC temperature setpoint is aligned with the desired temperature, set at 23.85°C for the majority of hours. This is driven by the occupant thermal comfort objective function, which aims to maintain the indoor temperature as close to the desired value as feasible. Furthermore, consumer total costs are \$6.025, which is improved by 13.2% compared to case (a). This improvement can be attributed to the load demand reduction objective function which aims to reduce consumer load demand by shifting their loads to lower-rate hours and increasing AC temperature setpoints. It should be noted that the occupant thermal comfort objective function makes the AC run at low temperatures (close to desired) during most hours, increasing costs and energy use.

TABLE 6. Hyperparameters of the multi-objective genetic algorithm.

Hyperparameters	values
Number of objective functions	3
Population size	500
Crossover rate	0.9
Mutation rate	0.1

TABLE 7. Transferred demand for shiftable loads in Case (c).

From hour	To hour	Transferred demand (kW)
13	22	0.5
20	22	0.5
17	21	1.5
18	22	1.5

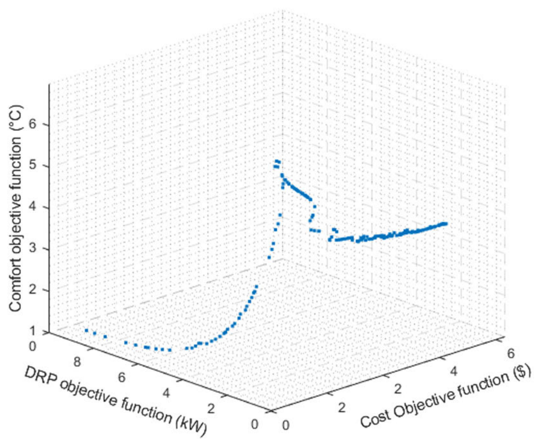


FIGURE 5. Multi-objective GA convergence curve in case (c).

This case underscores that participation in DRPs can yield cost savings for consumers without compromising comfort. However, ignoring uncertainties in input parameters can result in suboptimal or even impossible solutions. For instance, if actual demand surpasses predicted levels, consumers will face penalties resulting in an increase in their costs. The thermal comfort of occupants can also be compromised if fluctuations in occupancy levels are not accounted for. Accordingly, by integrating these uncertainties into the optimization problem, we can develop more resilient strategies that not only optimize energy consumption and costs but also ensure occupant thermal comfort.

4) CASE (D)

The concurrent effects of considering uncertainties and the application of DRPs are investigated in this case. Also, the convergence curve of the multi-objective GA is similar to Fig. 5. As demonstrated in Fig. 6, the AC temperature setpoints remain the same in uncertain and certain states for most hours, except for late hours of the night when electricity rates are lower, causing a decrease in AC setpoints due to uncertainty. This minor deviation is largely attributed to the occupancy level coefficient in the occupant thermal comfort objective function. Because any change in occupancy levels caused by uncertainty significantly increases the occupant thermal comfort objective function, and the robust counterpart optimization aims to maintain the AC temperature setpoints as close to desired value as possible. Moreover, the performance of shiftable loads exhibits sensitivity to uncertainty. According to Table 8, similar to case (c), washing machine operations transition from 17:00 and 18:00 to 21:00 and 22:00, respectively. While dishwasher operations continue at 13:00 and only change from 20:00 to 22:00. This change can be attributed to the hourly electric rates, which determine load transfer decisions. Given this study assumes deterministic hourly electric rates, the optimization problem reached the nearly identical schedules compare with case (c). In addition, uncertainty decreases the AC temperature setpoints at 21:00 to 22:00 (see Fig. 6), resulting in increased demand during these hours. Therefore, the optimization problem prevents appliances from being transferred to these hours to avoid extra costs. As illustrated in Fig. 4, consumer costs under case (c) align with those in case (d) when no budgets are applied. Moreover, even under worst-case conditions of case (d), where uncertainty exists in all hours, consumer costs improve by 9% compared to the worst point in case (b) due to the impact of DRPs.

TABLE 8. Transferred demand for shiftable loads in Case (d).

From hour	To hour	Transferred demand (kW)
20	22	0.5
17	21	1.5
18	22	1.5

Cumulatively, across all cases, using robust counterpart optimization with uncertain budgets allows HEMS to handle and measure uncertainty in a realistic manner. Instead of solely focusing on the worst-case scenarios, this approach considers varying levels of uncertainty that the HEMS might face. Fig. 4 illustrates all possible planning scenarios with various uncertainty budgets for all cases. Using the costs of case (a) as a benchmark, implementing DRPs and taking uncertainty into account for up to 8 out of 12 hours is cost-effective. For example, in case (d), when the data are uncertain for 8 hours, the consumer costs are $6.94 which is almost equal to the costs of case (a). However, in case (d) where uncertainty exists for more than 8 hours, the consumer

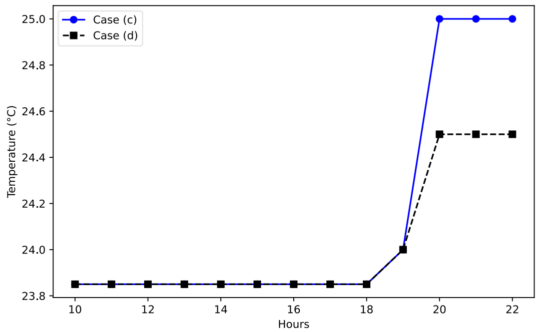


FIGURE 6. AC set-point temperature in cases (c) and (d).

costs rise and surpass the predefined benchmark indicating uncertainty offsets the savings from the DRPs. While this scenario might not seem cost-effective initially, accepting the higher costs could be justified for two key reasons: first, a more robust solution can increase system reliability and reduce the likelihood of unexpected failure. Secondly, it is less likely to encounter uncertainty spanning more than 8 out of 12 hours simultaneously.

V. CONCLUSION

This study addresses the significance of occupancy forecasting in the optimization of HEMS for effective participation in demand DRPs. While HEMS aims to adjust AC temperature setpoints the occupants thermal comfort should not be compromised. For HEMS to effectively leverage DRPs to minimize costs and load demand, while simultaneously preserving occupants thermal comfort, HEMS must be aware of occupancy levels. To address that, this work proposed non-intrusive, accurate, and cost-effective methods for predicting the occupancy level using advanced learning algorithms. The results demonstrated that the classification is ineffective because the occupancy samples are imbalanced. Accordingly, the regression task was performed by using advanced learning algorithms, such as random forest, MLP-ANN, and lightGBM. The results demonstrated the superiority of the random forest regression with MSE, RMSE, and MAE of 0.5630, 0.75038, and 0.39525, respectively. Additionally, the study emphasized the need to account for the inherent uncertainties tied to predicted occupancy and load demand when optimizing HEMS. Robust counterpart optimization along with uncertainty budgets were employed to pragmatically quantify uncertain parameters. This equipped the HEMS to reliably implement DRPs as well. The comparative analysis revealed that participation in DRPs, even with considering uncertainties, could yield cost savings for consumers without compromising comfort. The study showcased that while it is cost-effective to consider uncertainties up to 8 out of 12 hours, going beyond might lead to higher costs. However, these higher costs could be justified by increased system reliability and the relatively low likelihood of experiencing uncertainty for more than 8 hours. While this study provides a robust framework for HEMS optimization in the context of DRPs, future research may explore:

- Applying a smart city or multiple smart homes instead of a smart home.
- Formulating the problem from the investors or utility point of view.
- Considering uncertainties in other data such as market prices.
- Considering comfort in other sections, namely dishwasher and washing machine operation time, and finding the best function hours for these machines.
- Exploring the application of time series analysis techniques to forecast occupancy levels.

REFERENCES

[1] M. K. Metwaly and J. Teh, "Probabilistic peak demand matching by battery energy storage alongside dynamic thermal ratings and demand response for enhanced network reliability," *IEEE Access*, vol. 8, pp. 181547–181559, 2020.
[2] S. Hussain, C. Z. El-Bayeh, C. Lai, and U. Eicker, "Multi-level energy management systems toward a smarter grid: A review," *IEEE Access*, vol. 9, pp. 71994–72016, 2021.
[3] S. Ali, A. U. Rehman, Z. Wadud, I. Khan, S. Murawwat, G. Hafeez, F. R. Albogamy, S. Khan, and O. Samuel, "Demand response program for efficient demand-side management in smart grid considering renewable energy sources," *IEEE Access*, vol. 10, pp. 53832–53853, 2022.
[4] P. Munankarmi, H. Wu, A. Pratt, M. Lunacek, S. P. Balamurugan, and P. Spitsen, "Home energy management system for price-responsive operation of consumer technologies under an export rate," *IEEE Access*, vol. 10, pp. 50087–50099, 2022.
[5] Z. Chen, Y. Chen, R. He, J. Liu, M. Gao, and L. Zhang, "Multi-objective residential load scheduling approach for demand response in smart grid," *Sustain. Cities Soc.*, vol. 76, Jan. 2022, Art. no. 103530.
[6] M. Choobineh, A. Speake, M. Harris, P. C. Tabares-Velasco, and S. Mohagheghi, "End-user-aware community energy management in a distribution system exposed to extreme temperatures," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3753–3764, Jul. 2019.
[7] Q. Lu and Y. Zhang, "A multi-objective optimization model considering users' satisfaction and multi-type demand response in dynamic electricity price," *Energy*, vol. 240, Feb. 2022, Art. no. 122504.
[8] D. Zhang, H. Zhu, H. Zhang, H. H. Goh, H. Liu, and T. Wu, "Multi-objective optimization for smart integrated energy system considering demand responses and dynamic prices," *IEEE Trans. Smart Grid*, vol. 13, no. 2, pp. 1100–1112, Mar. 2022.
[9] K. Amasyali, Y. Chen, and M. Olama, "A data-driven, distributed game-theoretic transactional control approach for hierarchical demand response," *IEEE Access*, vol. 10, pp. 72279–72289, 2022.
[10] L. He, Y. Liu, and J. Zhang, "An occupancy-informed customized price design for consumers: A Stackelberg game approach," *IEEE Trans. Smart Grid*, vol. 13, no. 3, pp. 1988–1999, May 2022.
[11] Z. Chen, M. Wang, and Y. Wang, "Improving indoor occupancy detection accuracy of the SLEEPIR sensor using LSTM models," *IEEE Sensors J.*, vol. 23, no. 15, pp. 17794–17802, Aug. 2023.
[12] A. Naser, A. Lotfi, and J. Zhong, "Adaptive thermal sensor array placement for human segmentation and occupancy estimation," *IEEE Sensors J.*, vol. 21, no. 2, pp. 1993–2002, Jan. 2021.
[13] F. Nematkhah, F. Aminifar, M. Shahidehpour, and S. Mokhtari, "Evolution in computing paradigms for Internet of Things-enabled smart grid applications: Their contributions to power systems," *IEEE Syst. Man, Cybern. Mag.*, vol. 8, no. 3, pp. 8–20, Jul. 2022.
[14] J. Vanus, O. Majidzadeh Gorjani, P. Dvoracek, P. Bilik, and J. Koziorek, "Application of a new CO_2 prediction method within family house occupancy monitoring," *IEEE Access*, vol. 9, pp. 158760–158772, 2021.
[15] M. M. Ouf, M. H. Issa, A. Azzouz, and A.-M. Sadick, "Effectiveness of using WiFi technologies to detect and predict building occupancy," *Sustain. Buildings*, vol. 2, p. 7, Jul. 2017.
[16] B. Li, A. Tavakoli, and A. Heydarian, "Occupant privacy perception, awareness, and preferences in smart office environments," *Sci. Rep.*, vol. 13, no. 1, p. 4073, Mar. 2023.

[17] W. Zhang, Y. Wu, and J. K. Calautit, "A review on occupancy prediction through machine learning for enhancing energy efficiency, air quality and thermal comfort in the built environment," *Renew. Sustain. Energy Rev.*, vol. 167, Oct. 2022, Art. no. 112704.

[18] C. Liu, M. Li, Y. Yu, Z. Wu, H. Gong, and F. Cheng, "A review of multitemporal and multispatial scales photovoltaic forecasting methods," *IEEE Access*, vol. 10, pp. 35073–35093, 2022.

[19] C. Wan and W. Cui, "Machine learning-based probabilistic forecasting: A combined bootstrap and cumulant method," *IEEE Trans. Power Syst.*, vol. 39, no. 1, pp. 1370–1383, Jan. 2024.

[20] A. Belhadi, Y. Djenouri, G. Srivastava, D. Djenouri, J. C.-W. Lin, and G. Fortino, "Deep learning for pedestrian collective behavior analysis in smart cities: A model of group trajectory outlier detection," *Inf. Fusion*, vol. 65, pp. 13–20, Jan. 2021.

[21] R. Nematirad and A. Pahwa, "Solar radiation forecasting using artificial neural networks considering feature selection," in *Proc. IEEE Kansas Power Energy Conf. (KPEC)*, Apr. 2022, pp. 1–4.

[22] C. Sardianos, I. Varlamis, C. Chronis, G. Dimitrakopoulos, Y. Himeur, A. Alsalemi, F. Bensaali, and A. Amira, "A model for predicting room occupancy based on motion sensor data," in *Proc. IEEE Int. Conf. Informat., IoT, Enabling Technol. (ICIoT)*, Feb. 2020, pp. 394–399.

[23] D. H. Nguyen, "Residential energy consumer occupancy prediction based on support vector machine," *Sustainability*, vol. 13, no. 15, p. 8321, Jul. 2021.

[24] L. Wang, M. Han, X. Li, N. Zhang, and H. Cheng, "Review of classification methods on unbalanced data sets," *IEEE Access*, vol. 9, pp. 64606–64628, 2021.

[25] M. Imani and S. F. Ghoreishi, "Graph-based Bayesian optimization for large-scale objective-based experimental design," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 10, pp. 5913–5925, Oct. 2022.

[26] M. Eguaras-Martínez, M. Vidaurre-Arbizu, and C. Martín-Gómez, "Simulation and evaluation of building information modeling in a real pilot site," *Appl. Energy*, vol. 114, pp. 475–484, Feb. 2014.

[27] M. Jasinski, A. Najafi, O. Homaee, M. Kermani, G. Tsaousoglou, Z. Leonowicz, and T. Novak, "Operation and planning of energy hubs under uncertainty—A review of mathematical optimization approaches," *IEEE Access*, vol. 11, pp. 7208–7228, 2023.

[28] L. Munkhdalai, T. Munkhdalai, K. H. Park, H. G. Lee, M. Li, and K. H. Ryu, "Mixture of activation functions with extended min-max normalization for forex market prediction," *IEEE Access*, vol. 7, pp. 183680–183691, 2019.

[29] C. Zhan, Y. Zheng, H. Zhang, and Q. Wen, "Random-forest-bagging broad learning system with applications for COVID-19 pandemic," *IEEE Internet Things J.*, vol. 8, no. 21, pp. 15906–15918, Nov. 2021.

[30] X. Yao, X. Fu, and C. Zong, "Short-term load forecasting method based on feature preference strategy and LightGBM-XGboost," *IEEE Access*, vol. 10, pp. 75257–75268, 2022.

[31] S. Mohammadpouri, M. Sadeghnejad, H. Rezaei, R. Ghanbari, S. Tayebi, N. Mohammadzadeh, N. Mijani, A. Raeisi, S. Fathololoumi, and A. Biswas, "A generalized regression neural network model for accuracy improvement of global precipitation products: A climate zone-based local optimization," *Sustainability*, vol. 15, no. 11, p. 8740, May 2023.

[32] I. Salehin and D.-K. Kang, "A review on dropout regularization approaches for deep neural networks within the scholarly domain," *Electronics*, vol. 12, no. 14, p. 3106, Jul. 2023.

[33] D. Devi, S. K. Biswas, and B. Purkayastha, "A review on solution to class imbalance problem: Undersampling approaches," in *Proc. Int. Conf. Comput. Perform. Eval. (ComPE)*, Jul. 2020, pp. 626–631.

[34] J. Zhou, A. H. Gandomi, F. Chen, and A. Holzinger, "Evaluating the quality of machine learning explanations: A survey on methods and metrics," *Electronics*, vol. 10, no. 5, p. 593, Mar. 2021.

[35] A. A. Adamopoulou, A. M. Tryferidis, and D. K. Tzovaras, "A context-aware method for building occupancy prediction," *Energy Buildings*, vol. 110, pp. 229–244, Jan. 2016.

[36] R. Hendron and C. Engebrecht, *Building America House Simulation Protocols*. Washington, DC, USA: U.S. Dept. Energy, 2010, p. 79.

[37] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*, vol. 28. Princeton, NJ, USA: Princeton Univ. Press, 2009.

[38] A. Dindar, M. Choobineh, and S. Mohagheghi, "Socially-and-environmentally-aware power management in a residential neighborhood under heat wave events," in *Proc. IEEE Ind. Appl. Soc. Annu. Meeting*, Oct. 2020, pp. 1–8.

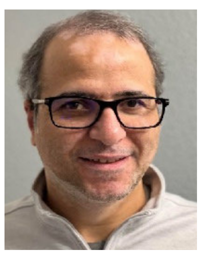
[39] M.-R. Chen, G.-Q. Zeng, and K.-D. Lu, "Constrained multi-objective population extremal optimization based economic-emission dispatch incorporating renewable energy resources," *Renew. Energy*, vol. 143, pp. 277–294, Dec. 2019.

[40] K.-D. Lu and Z.-G. Wu, "Multi-objective false data injection attacks of cyber–physical power systems," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 9, pp. 3924–3928, Sep. 2022.

REZA NEMATIRAD (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering from the Amirkabir University of Technology (Center of Excellence in Power Systems), Tehran, Iran, in 2018, and the M.Sc. degree in power system planning and management from the Amirkabir University of Technology, in 2021. He is currently pursuing the Ph.D. degree in electrical and computer engineering with Kansas State University. His research interests include optimization in power systems, state estimation, probabilistic and statistical analysis, renewable energy resources, learning algorithms, and time series analysis.

M. M. ARDEHALI received the Ph.D. degree from Iowa State University in collaboration with the U.S. Department of Energy, in 1995. He held a postdoctoral research position with Iowa State University in collaboration with the U.S. Department of Energy, in 1995. He is a registered professional engineer in several states. For over 20 years, he has served as a Faculty Member of the Power Systems Group, Department of Electrical Engineering, Amirkabir University of Technology (Tehran Polytechnic). He was the Principal Investigator with the Energy Systems Laboratory. His current research interests include the utilization of artificial intelligence for optimization, control, and fault diagnostics of energy systems.

AMIR KHORSANDI was born in Tehran, Iran, in 1986. He received the B.Sc. and M.Sc. degrees in electrical engineering from the Amirkabir University of Technology, Tehran, in 2008 and 2010, respectively, and the Ph.D. degree in electrical engineering from the Sharif University of Technology, Tehran, in 2015. He is currently an Assistant Professor with the Amirkabir University of Technology. His research interests include power quality, microgrid control, and applications of power electronics in power distribution systems.



ARASH MAHMOUDIAN (Student Member, IEEE) received the B.Sc. degree in electronics engineering from Tarbiat Moallem Sabzevar University (renamed to Hakim Sabzevari University, in 2011), Khorasan Razavi, Iran, in 2008. He is currently pursuing the master's degree in data science with Bellevue University, NB, USA. He is with Dish Wireless (O-RAN network). His research interests include the use of AI/ML in 5G network optimization, statistics, time series analysis, and vector databases.