

Energy Consumption Probabilistic Prediction of HVAC Systems in Public Buildings Based on Deep Learning Fusion Model

Siyu Jiang

*State Key Laboratory of Internet of Things for Smart City
University of Macau
Macao, China*

sy.jiang@connect.um.edu.mo

Hongxun Hui*

*State Key Laboratory of Internet of Things for Smart City
University of Macau
Macao, China*

hongxunhui@um.edu.mo

Abstract—The energy consumption prediction of heating ventilating and air-conditioning (HVAC) systems in public buildings is essential for promoting energy efficiency. However, HVAC energy consumption often fluctuates significantly due to weather variations and occupancy uncertainties within public buildings. To address this issue, this paper introduces a probabilistic prediction model based on the deep learning fusion model to quantify the energy consumption ranges with specific confidence intervals. First, the impact of temporal and environmental characteristics on HVAC energy consumption are analyzed to select relevant features. Second, we combine Long Short-Term Memory with Conformalized Quantile Regression model to obtain prediction intervals of energy consumption. Finally, to improve the model's generalization performance, an ensemble learning method is introduced to adapt varying time-series lengths through homogeneous model enhancements. Based on realistic data from an office building in the University of Macau, case studies validate the superior accuracy and generalization of the proposed model.

Index Terms—HVAC energy consumption, probabilistic prediction, long short-term memory, conformalized quantile regression, ensemble learning

I. INTRODUCTION

Since the start of the global energy crisis, countries representing over 70% of the world's energy consumption have strengthened energy efficiency policies. Buildings, in terms of their operations, account for close to one-third of the global energy consumption [1]. The proliferation of heating ventilating and air-conditioning (HVAC) systems stands as the primary driver of the increasing electricity consumption in buildings. Globally, HVAC services constitute approximately 40% of the total energy usage in buildings [2]. In certain subtropical regions, such as Singapore and Macao, HVAC systems account for over 50% of the electricity consumption in building complexes [3]. By 2050, it is predicted that cooling

This work is funded in part by the Guangdong Basic and Applied Basic Research Foundation, China (Grant No. 2024A1515010141), in part by the Science and Technology Development Fund, Macau SAR (File no. 001/2024/SKL), and in part by the Start-up Research Grant of University of Macau (File no. SRG2023-00063-IOTSC). And the corresponding author is Hongxun Hui.

demand will drive an additional global electricity usage of 2800 TWh [2].

As a consequence of climate change, heat waves are expected to become more frequent and intense, further amplifying the cooling demand. The escalating demand for electricity attributed to HVAC systems within buildings presents a significant challenge to the power system operation. Therefore, accurate predictions of HVAC energy consumption facilitate the development of dispatch plans and enable flexible regulation of HVAC resources for power systems. However, HVAC energy consumption often fluctuates significantly due to weather variations and occupancy uncertainties. Consequently, researchers have been dedicated to predicting HVAC energy consumption.

In recent years, attention has predominantly focused on deterministic point predictions, with data-driven models assuming primacy. These models, encompassing machine learning and deep learning techniques, have advantages in elucidating intricate nonlinear relationships across heterogeneous data. Notable models include Support Vector Regression (SVR) [4], Convolutional Neural Networks (CNN) [5], and Recurrent Neural Networks (RNN) [6]. Nonetheless, deterministic point prediction of HVAC energy consumption is difficult to provide comprehensive knowledge due to the inherent uncertainties of environment features [7]. In contrast, probabilistic prediction of HVAC energy consumption offers crucial insights into quantifying uncertainties through prediction intervals (PI), quantile regression, or probability distributions.

Probabilistic energy consumption prediction models can be categorized into parametric and non-parametric estimation techniques. Parametric estimation assumes specific distribution for prediction errors. However, HVAC energy consumption displays significant variability across scenarios, which are unable to be accurately modeled using a prior specific distribution [8]. Conversely, non-parametric estimation formulates a customized distribution based on the intrinsic characteristics of sample data, which is particularly robust for irregular distributions. Common non-parametric models include Quantile

Regression (QR) [9], Kernel Density Estimation (KDE) [10], Gaussian Process Regression, among others [11]. Nonetheless, purely statistical models may fail to capture interrelationships among heterogeneous data.

To address the challenges mentioned above, current research efforts have focused on developing fusion models that combine statistical models with data-driven models. These models aim to effectively leverage heterogeneous data for prediction and quantify uncertainties of HVAC energy consumption. Afrasiabi *et al.* [12] propose a CNN-RNN model integrated into a mixture density network for predicting the probability density functions of residential loads. Cheng *et al.* [10] introduce a CNN architecture with squeeze-and-excitation modules to capture load characteristics and employ piecewise KDE for residential load prediction based on micrometeorological data and customer consumption patterns. Wang *et al.* [13] devise a probabilistic model utilizing point predicting residuals by modeling conditional prediction residuals. Wen *et al.* [14] propose a doubly residual stacking strategy based on neural basis expansion analysis and Conformalized Quantile Regression (CQR) to minimize coverage for load probabilistic prediction. However, given the significant temporal and occupancy flow induced variations in HVAC energy consumption within public buildings, most aforementioned models do not fully account for local variability in energy consumption data.

Based on the aforementioned analysis, this paper acknowledges the challenge of probabilistic energy consumption prediction in public buildings due to the variability of HVAC energy consumption over time. To address this issue, this paper proposes a Long Short-Term Memory Ensemble Conformalized Quantile Regression model (LECQR). The contributions are described as follows:

1) We analyze the relevant features of HVAC energy consumption consisting of ambient features and temporal features, in which primary features are selected based on linear and non-linear methods.

2) Long Short-Term Memory network and Conformalized Quantile Regression model are combined to both establish temporal dependencies from HVAC energy consumption and improve energy consumption prediction accuracy by introducing residual truncation.

3) LECQR fusion model is developed by introducing the ensemble learning method to enhance the model adaptability to heteroscedastic energy consumption data.

The remainder of this paper is organized as follows. Section II presents the probabilistic predicting framework of public buildings HVAC energy consumption. Section III formulates feature selection methods. Section IV illustrates LECQR model in details. Case studies are presented in Section V. The conclusion is presented in Section VI.

II. PROBABILISTIC PREDICTION FRAMEWORK

Fig. 1 illustrates the framework of the fusion model. Recognizing the limitations of traditional single model in capturing the complex variations of HVAC energy consumption, this paper proposes a deep learning fusion model LECQR.

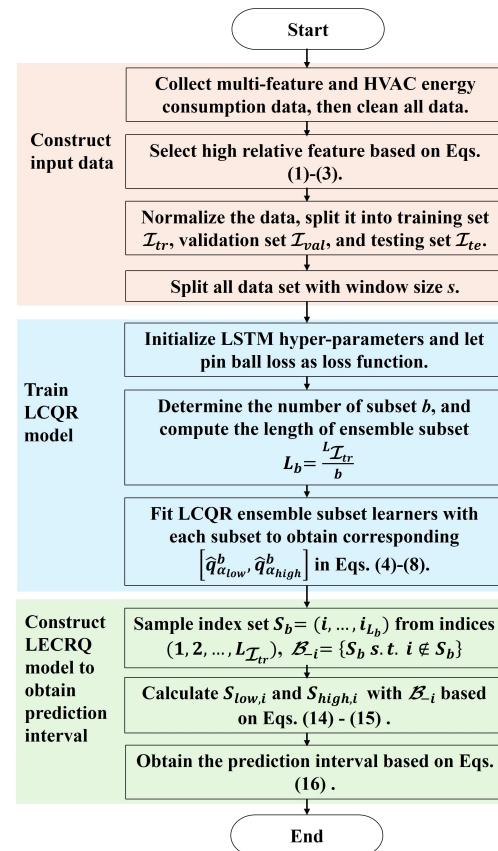


Fig. 1. The probabilistic prediction framework of HVAC systems energy consumption.

Specifically, this model consists of the Long Short-Term Memory (LSTM) prediction module, CQR probabilistic prediction module, and an ensemble learning module.

First, the process commences by constructing the input data series. Ambient features, temporal features, and HVAC energy consumption are constructed and collected. After addressing outliers and missing values, primary features are selected using the methods outlined in Eqs. (1)-(3). Subsequently, these selected features are merged with HVAC energy consumption to form the pre-normalized dataset. Then the dataset is normalized, and partitioned into subsets \mathcal{I}_{tr} , \mathcal{I}_{val} , \mathcal{I}_{te} with a predetermined sliding window size s .

Then, each subset fits subset LCQR model. This step entails the specification of initial LSTM parameters, including the learning rate, number of hidden layer units, loss function etc. Notably, the pinball loss function is adopted here. Determination of the subset count b is followed by the calculation of the data length for each subset L_b . Subsequently, the LCQR model is fitted to each subset to obtain upper quantile and lower quantile, as detailed in Eqs. (4)-(10).

Finally, an ensemble method is introduced to derive LECQR PI. Sample set is extracted to form S_b , and the complement of S_b is constructed using leave-one-out cross-validation B_{-i} . B_{-i} is utilized as input to compute conformity scores $S_{low,i}$ and $S_{high,i}$. After aggregating each subset conformity score, we can obtain the output PI, following the steps outlined in

Eqs. (11)-(16).

III. HVAC ENERGY CONSUMPTION RELATED FEATURES IN PUBLIC BUILDINGS

Analyzing the correlation between constructed features and HVAC energy consumption holds significance. This section illustrates methods for computing correlation coefficients between distinct constructed features and HVAC energy consumption.

HVAC energy consumption is subject to the influence of diverse features, primarily encompassing ambient features and temporal features. Ambient features include ambient temperature, humidity, wind speed, wind direction, occupancy flow, etc. Temporal features are derived from parameters such as day of the week, event features, and hour of the day. In detail, event features categorize weekdays as 1 and holidays as 0.

The Pearson correlation coefficient [14] is expressed as follows:

$$r_p = \frac{\sum_{k=1}^n (H_{i,k} - \bar{H}_i)(H_{j,k} - \bar{H}_j)}{\sqrt{\sum_{k=1}^n (H_{i,k} - \bar{H}_i)^2} \sqrt{\sum_{k=1}^n (H_{j,k} - \bar{H}_j)^2}}, \quad (1)$$

where $i, j = 1, 2, \dots, 19$. Parameters H_i and H_j represent the sample sizes of different features, respectively. Specifically, $H_1 - H_{19}$ respectively denote ambient temperature, daily maximum temperature, daily minimum temperature, apparent temperature, relative humidity, dew point temperature, total daily rainfall, 10-minute average wind speed, gust speed, wind direction, wind direction in degrees, hour of the day, day of the week, day of the month, week number, event feature, lighting and socket energy consumption, occupancy flow, and HVAC energy consumption, respectively. The variable n signifies the number of samples for each feature, while \bar{H}_i and \bar{H}_j denote the mean value of H_i and H_j , respectively.

Although the Pearson correlation coefficient is a commonly used method for measuring the strength of relationships between variables, it is sensitive to outliers. To comprehensively evaluate the variables' correlation, this paper also employs the Spearman rank correlation coefficient. Unlike Pearson correlation coefficient, Spearman does not require data to follow a linear relationship or be normally distributed. The expressions are as follows:

$$r_s = 1 - \frac{6 \sum_{k=1}^n dist_k^2}{n(n^2 - 1)}, \quad (2)$$

$$dist_k = rank(H_{i,k}) - rank(H_{j,k}), \quad (3)$$

where $dist_k$ denotes the level difference between different features.

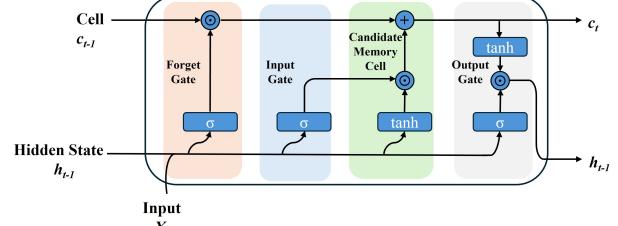


Fig. 2. The structure diagram of LSTM.

IV. LONG SHORT-TERM MEMORY ENSEMBLE CONFORMALIZED QUANTILE REGRESSION

A. Long Short-Term Memory

Long Short-Term Memory (LSTM) is a new architecture for RNN, which is specially designed to overcome the vanishing or exploding gradients [15]. Internal structure of LSTM is based on a set of connected units. To understand the function of LSTM memory units, Fig. 2 depicts the structure of an LSTM unit.

LSTM is equipped with a cell state feature for storing information, which learns temporal information across various time scales. The forget gate controls retention of historical temporal feature information in c_{t-1} . The input gate identifies historical features essential for predicting HVAC energy consumption at the current time step t . And the output gate governs dissemination of the current hidden state h_t . The computational process of LSTM unfolds as described:

$$\begin{bmatrix} i_t \\ f_t \\ o_t \end{bmatrix} = \sigma(X_t \begin{bmatrix} w_{xi} \\ w_{xf} \\ w_{xo} \end{bmatrix} + h_{t-1} \begin{bmatrix} w_{hi} \\ w_{hf} \\ w_{ho} \end{bmatrix} + \begin{bmatrix} b_i \\ b_f \\ b_o \end{bmatrix}), \quad (4)$$

where X_t represents the input information at time step t ; i_t , f_t and o_t denote the states of the input gate, forget gate, and output gate at time step t , respectively; the activation function at the fully connected layer is sigmoid function σ ; h_{t-1} represents the hidden state from the previous time step $t-1$; w denotes the weights of different gates; b represents the bias terms of different gates.

The calculation of candidate memory is similar to the three gates' calculation steps in Eq. (4), while it utilizes the tangent function. At time step t , the equations are as follows:

$$\tilde{c}_t = \tanh(X_t w_{xc} + h_{t-1} w_{hc} + b_c), \quad (5)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t, \quad (6)$$

$$h_t = o_t \odot \tanh(c_t), \quad (7)$$

where \tilde{c}_t , c_t and h_t denote the candidate cell memory, cell state and hidden state at step t , respectively.

B. Conformalized Quantile Regression

Conformalized Quantile Regression (CQR) is a probabilistic prediction model based on statistical theory that combines the advantages of conformal prediction (CP) and QR [16]. It is

capable of constructing effective PI for heteroscedastic data without assuming the data distribution. Moreover, it can adapt to local variations in the data. Hence, it is suitable for HVAC energy consumption probabilistic prediction with complex data distribution.

Similar to CP, CQR assumes that samples are exchangeable and splits data into training set \mathcal{I}_{tr} and validation set \mathcal{I}_{val} . Initially, on \mathcal{I}_{tr} , CQR utilizes QR to fit the conditional quantile function:

$$[\hat{q}_{\alpha_{low}}, \hat{q}_{\alpha_{high}}] = f(X_i, Y_i), i \in \mathcal{I}_{tr}, \quad (8)$$

$$\alpha_{low} = \frac{\alpha}{2}, \quad (9)$$

$$\alpha_{high} = 1 - \frac{\alpha}{2}, \quad (10)$$

where α denotes the confidence level; α_{low} and α_{high} represent the lower and upper quantiles, respectively; f represents LCQR model; X_i represents input features and Y_i represents observed values; $\hat{q}_{\alpha_{low}}$ signifies the quantile lower bound regression function; $\hat{q}_{\alpha_{high}}$ denotes the quantile regression upper bound regression function.

Subsequently, conformity scores are utilized to conformalized the PI:

$$S_i = \max \{ \hat{q}_{\alpha_{low}}(X_i) - Y_i, Y_i - \hat{q}_{\alpha_{high}}(X_i) \}, i \in \mathcal{I}_{val}. \quad (11)$$

Then the CQR PI is as follows:

$$C_\alpha(X_{i+1}) = [\hat{q}_{\alpha_{low}}(X_{i+1}) - Q_{1-\alpha}(S, \mathcal{I}_{val}), \hat{q}_{\alpha_{high}} + Q_{1-\alpha}(S, \mathcal{I}_{val})], \quad (12)$$

$$Q_{1-\alpha}(S, \mathcal{I}_{val}) = (1 - \alpha)(1 + \frac{1}{|\mathcal{I}_{val}|}). \quad (13)$$

C. Long Short-Term Memory Ensemble Conformalized Quantile Regression Prediction Intervals

Ensemble learning is a widely adopted method to enhance model performance by combining multiple distinct subset learners.

Long Short-Term Memory Ensemble Conformalized Quantile Regression (LECQR) establishes b disjoint subsets, each of length L_b , where every subset trains a distinct learner. When \mathcal{B}_{-i} is applied for training, the subset learner can generate out-of-sample residuals. Within each subset, multiple overlapping input-output sequences are extracted to serve as batches in subset learner training process. The aggregated PI obtained through ensemble subset learners are conformalized using asymmetric conformity scores, defined as:

$$S_{low,i} = \hat{q}_{\alpha_{low}}(X_i) - Y_i, \quad (14)$$

$$S_{high,i} = Y_i - \hat{q}_{\alpha_{high}}(X_i). \quad (15)$$

The LECQR PI is constructed by LECQR using the estimated quantile functions and absolute conformity scores along with asymmetric conformity scores. And the LECQR PI is described as follows:

$$\hat{C}_{\alpha,\phi}(X_t) = [\hat{q}_{\alpha_{low}}(X_t) - Q_{1-\alpha}(\varepsilon_{low}), \hat{q}_{\alpha_{high}}(X_t) + Q_{1-\alpha}(\varepsilon_{high})], \quad (16)$$

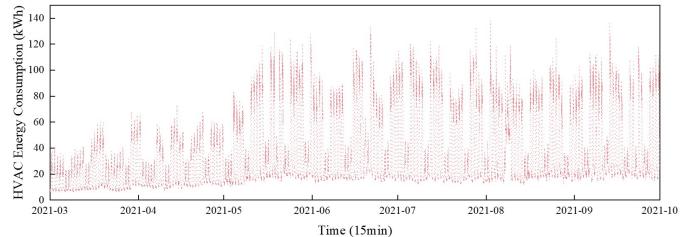


Fig. 3. HVAC energy consumption from March to September 2021.

where ϕ is the aggregation function; ε denotes the set of residuals, with $\varepsilon_{low} = \{S_{low,i}\}_{t-T}^{i=t-1}$ and $\varepsilon_{high} = \{S_{high,i}\}_{t-T}^{i=t-1}$.

V. CASE STUDIES

A. Experiment Settings

1) *Data Description:* The paper conducts case studies using realistic HVAC energy consumption, occupancy flow, and numerical weather forecast data obtained from a specific office building at the University of Macau. The data spans from March to September 2021, with a resolution of 15 minutes. HVAC energy consumption is shown in the Fig. 3.

The data is partitioned as follows: the training set encompasses data from March through August and is denoted by \mathcal{I}_{tr} ; the validation set is labeled as \mathcal{I}_{val} , comprising the initial half of September, from the 1st to the 15th; the test set includes the latter half of September, from the 16th to the 30th, and is labeled as \mathcal{I}_{te} . Furthermore, the training set is divided into three subsets, each indicated by the subset count b , which corresponds to the following periods: March to April, May to June, and July to August.

2) *Evaluation Metrics:* In order to evaluate the model comprehensively, this subsection introduces a variety of metrics. Prediction Intervals Coverage Probability (PICP) calculates the effectiveness of PI $\hat{C}_{\alpha,\phi}(X_t)$ by determining the proportion of predicted values \hat{Y}_i falling within PI.

$$\text{PICP} = \frac{1}{n} \sum_{i=1}^n k_i, \quad (17)$$

$$k_t = \begin{cases} 1, & \hat{Y}_i \in \hat{C}_{\alpha,\phi}(X_i) \\ 0, & \hat{Y}_i \notin \hat{C}_{\alpha,\phi}(X_i) \end{cases}. \quad (18)$$

Prediction Intervals Normalized Average Width (PINAW) offers insights into the precision of PI by measuring average width relative to the range of observed values.

$$\text{PINAW} = \frac{1}{n(Y_{\max} - Y_{\min})} \sum_{i=1}^n \hat{C}_{\alpha,\phi}(X_i). \quad (19)$$

To summarize the quality of the PI briefly, a modified version of the coverage width-based criterion CWC is adopted, which penalizes the values out of coverage [16].

$$\text{CWC} = (1 - \text{PINAW}) e^{-\lambda(\text{PICP} - (1 - \alpha))^2}, \quad (20)$$

where let λ be equal to 30.

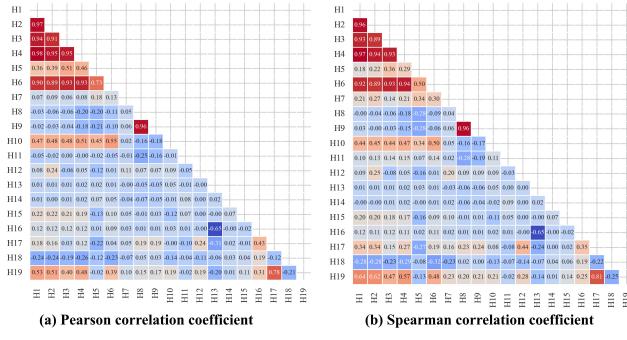


Fig. 4. Correlation analysis results.

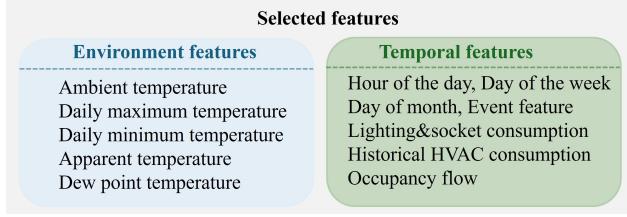


Fig. 5. Selected input features for HVAC energy consumption prediction.

Mean Absolute Error (MAE) measures prediction accuracy by computing the average magnitude of errors between predicted and observed values.

$$\text{MAE} = \frac{\sum_{i=1}^n |\hat{Y}_i - Y_i|}{n}. \quad (21)$$

Root Mean Squared Error (RMSE) quantifies the overall magnitude of residual errors by calculating quadratic mean of the differences between the observed values and predicted values.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\hat{Y}_i - Y_i)^2}{n}}. \quad (22)$$

Coefficient of Determination (R^2) assesses the proportion of variance in the dependent variable explained by the independent variables.

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{Y}_i - Y_i)^2}{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}. \quad (23)$$

3) Training Environment: The proposed model related hyper-parameters are listed in TABLE I. Notably, an early stopping mechanism is incorporated during LSTM training to prevent overfitting of the model. All the experiments are conducted on a desktop with Intel(R) Core(TM) i7-12700 CPU and NVIDIA GeForce RTX 3060TI GPU (64GB RAM) on a Windows 11 platform.

B. Correlation Analysis Results

A correlation analysis between HVAC energy consumption and constructed features is conducted. To provide a more intu-

TABLE I
HYPER-PARAMETERS SETTING DURING THE INITIAL TRAINING PROCESS

Hyper-parameter	Value
Learning rate	0.002
Batch size	24
Number of hidden layers	2
Units in each hidden layer	[256,128]
Optimizer	Adam

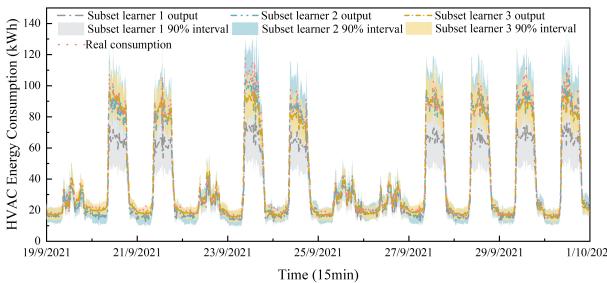


Fig. 6. The prediction results of different subset learners at a 90% confidence interval.

itive visualization of the analysis results, a heatmap displaying the correlation coefficients is depicted in Fig. 4.

Based on Fig. 4, it is evident that the primary correlation feature is lighting and socket energy consumption. This is followed by correlations with ambient temperature, daily maximum temperature, daily minimum temperature, apparent temperature, and dew point temperature. However, correlations with features such as relative humidity, total daily rainfall, 10-minute average wind speed, gust speed, wind direction, and week number are relatively weak, with their impact on HVAC energy consumption being negligible. Therefore, ambient temperature, daily maximum temperature, daily minimum temperature, apparent temperature, dew point temperature, hour of the day, day of the week, day of the month, event feature, lighting, socket energy consumption, history of HVAC energy consumption, and occupancy flow are selected as the primary influencing features for HVAC energy consumption, as the Fig. 5 shows.

C. Probabilistic Prediction Results

Fig. 6 illustrates the prediction results on the test set at a 90% confidence interval by three subsets-trained subset

TABLE II
EVALUATION METRICS RESULTS OF DIFFERENT MODELS

Model	PICP	PINAW	CWC	MAE	RMSE	R ²
Subset learner 1	68.33%	0.18	0.20	11.06	15.22	0.75
Subset learner 2	77.30%	0.14	0.53	3.75	5.28	0.97
Subset learner 3	90.60%	0.14	0.86	3.30	5.21	0.97
LCQR	88.72%	0.16	0.80	4.77	7.06	0.95
LEQCR	91.11%	0.17	0.84	4.51	6.23	0.96

TABLE III
EVALUATION METRICS RESULTS AT DIFFERENT INTERVALS

Confidence Interval	PICP	PINAW	CWC	MAE	RMSE	R ²
95% interval	96.71%	0.23	0.68	4.70	6.84	0.95
90% interval	91.18%	0.17	0.84	4.51	6.23	0.96
80% interval	82.01%	0.19	0.67	5.70	8.49	0.92
70% interval	70.03%	0.09	0.34	5.04	7.62	0.94
50% interval	51.22%	0.09	0.29	4.27	6.11	0.96

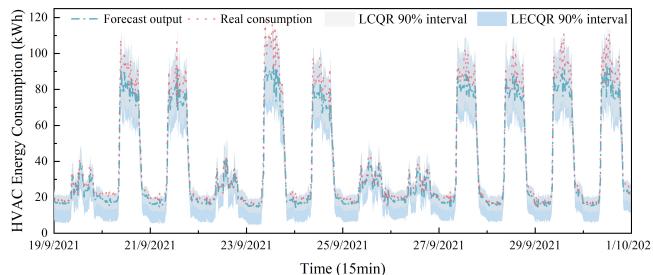


Fig. 7. The prediction results of LCQR and LECQR at a 90% confidence interval.

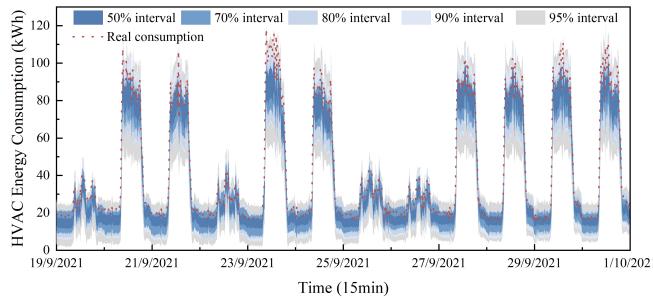


Fig. 8. The prediction results of LECQR at different confidence intervals.

learners, respectively. It is obvious that the mean output value of subset learner 1 exhibits a significant deviation from the realistic consumption. This discrepancy can be attributed to the limited information conveyed by the HVAC energy consumption during March and April when electricity demand does not reach peak. In contrast, subset learners 2 and 3 demonstrate higher precision. Additionally, only subset learner 3 achieves the required 90% confidence interval.

Furthermore, to validate the effectiveness of the proposed fusion model, a comparison with the LCQR model is conducted at a 90% confidence interval. As depicted in Fig. 7, the coverage interval of LECQR is notably broader. TABLE II presents various performance metrics for different subset learners, LCQR, and LECQR. Compared with LCQR, LECQR leads to about 11.84% reduction in RMSE, 5.15% enhancement in CWC.

Finally, probabilistic predictions of HVAC energy consumption are carried out at different confidence intervals. As depicted in Fig. 8 and detailed in TABLE III, an increase in the confidence interval results in a widening of PI.

VI. CONCLUSION

This paper proposes a probabilistic energy consumption prediction model for public buildings HVAC energy consumption based on LECQR. First, various features are constructed, which have high relevance to HVAC energy consumption. Then the LCQR model are utilized to fit every subset learner. Finally, by introducing an ensemble method into LCQR, this paper obtains the LECQR model to predict PI. The case study demonstrates that the proposed model achieves PICP of 91.11% and CWC of 0.84 at a 90% confidence interval. In comparison with LCQR, the proposed model shows a CWC improvement of 5.15%.

REFERENCES

- [1] Efficient grid-interactive buildings. Technical report, IEA, 2023.
- [2] More efficient and flexible buildings are key to clean energy transitions. Technical report, IEA, 2024.
- [3] World energy outlook 2023. Technical report, IEA, 2023.
- [4] Elena Mocanu, Phuong H Nguyen, Madeleine Gibescu, and Wil L Kling. Comparison of machine learning methods for estimating energy consumption in buildings. In *Probabilistic Methods Applied to Power Systems (PMAPS)*, pages 1–6. IEEE, 2014.
- [5] M Jayashankara, Priyansh Shah, Anshul Sharma, Prasenjit Chanak, and Sanjay Kumar Singh. A novel approach for short-term energy forecasting in smart buildings. *IEEE Sensors Journal*, 23(5):5307–5314, 2023.
- [6] Yidong Liu, Leibo Liu, Fabrizio Lombardi, and Jie Han. An energy-efficient and noise-tolerant recurrent neural network using stochastic computing. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 27(9):2213–2221, 2019.
- [7] Chirag Deb, Fan Zhang, Junjing Yang, Siew Eang Lee, and Kwok Wei Shah. A review on time series forecasting techniques for building energy consumption. *Renewable and Sustainable Energy Reviews*, 74:902–924, 2017.
- [8] Adrian Chong, Yaonan Gu, and Hongyuan Jia. Calibrating building energy simulation models: A review of the basics to guide future work. *Energy and Buildings*, 253:111533, 2021.
- [9] Yaoyao He, Yang Qin, Shuo Wang, Xu Wang, and Chao Wang. Electricity consumption probability density forecasting method based on lasso-quantile regression neural network. *Applied Energy*, 233:565–575, 2019.
- [10] Dunjian Xie, Hongxun Hui, Yi Ding, and Zhenzhi Lin. Operating reserve capacity evaluation of aggregated heterogeneous tcls with price signals. *Applied Energy*, 216:338–347, 2018.
- [11] Dennis W Van der Meer, Mahmoud Shepero, Andreas Svensson, Joakim Widén, and Joakim Munkhammar. Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using gaussian processes. *Applied Energy*, 213:195–207, 2018.
- [12] Mousa Afrasiabi, Mohammad Mohammadi, Mohammad Rastegar, Lina Stankovic, Shahabedin Afrasiabi, and Mohammad Khazaei. Deep-based conditional probability density function forecasting of residential loads. *IEEE Transactions on Smart Grid*, 11(4):3646–3657, 2020.
- [13] Yi Wang, Qixin Chen, Ning Zhang, and Yishen Wang. Conditional residual modeling for probabilistic load forecasting. *IEEE Transactions on Power Systems*, 33(6):7327–7330, 2018.
- [14] Honglin Wen, Jie Gu, Jinghuan Ma, Lyuzerui Yuan, and Zhijian Jin. Probabilistic load forecasting via neural basis expansion model based prediction intervals. *IEEE Transactions on Smart Grid*, 12(4):3648–3660, 2021.
- [15] Yong Yu, Xiaosheng Si, Changhua Hu, and Jianxun Zhang. A review of recurrent neural networks: Lstm cells and network architectures. *Neural Computation*, 31(7):1235–1270, 2019.
- [16] Yaniv Romano, Evan Patterson, and Emmanuel J. Candès. Conformalized quantile regression. *Neural Information Processing Systems*, 2019.