A Type-2 Fuzzy-Based Feature-Driven Rule Reduction Approach for Demand Forecasting within an Energy Smart Grid System

Mahmoud Alfayan, Hani Hagras
School of Computer Science and Electronic Engineering
University of Essex
Colchester, United Kingdom

***Abstract*— Accurate short-term energy demand forecasting is essential for the stability and efficiency of modern energy smart grids, especially with the growing integration of renewable energy sources, which introduce significant uncertainty. Traditional machine learning models, such as neural networks and deep learning, are "black boxes" that deliver accurate predictions but lack transparency, limiting operator trust and actionable insights. To help users better understand the prediction model and increase their confidence in its outputs, it is necessary to employ eXplainable Artificial Intelligence (XAI), whose models and predicted results can be easily understood, analyzed, and improved by business users. In this paper, we present an Interval Type-2 Fuzzy Logic System (IT2 FLS)-based XAI system, designed to manage noise, non-stationarity, and uncertainty in hourly demand data. We utilize SHapley Additive Explanations (SHAP) and a Random Forest (RF) to enhance model interpretability and prediction reliability by reducing the number of inputs and rules in the IT2 FLS. In our approach, SHAP is used to rank features and combine their importance with RF-based significance for optimal feature selection and rule pruning in the IT2 FLS. This method reduces the root-mean-squared error (RMSE) on testing data by 55%, from 5.94 kilowatt-hours (with a full rule base of 2,781 rules) to 2.74 kilowatt-hours, using only three inputs and 16 rules—a 250-fold reduction in the number of rules. In comparison, the Type-1 FLS showed lower performance with an RMSE of 5.376, while the neural network achieved a lower RMSE of 1.410 but lacked interpretability. The proposed IT2 FLS-based XAI system thus offers a practical balance between accuracy and explainability, enabling real-time smart grid applications. This approach will be further applied within a smart grid to facilitate adaptive decision-making and optimization.**

Keywords— Type-2 fuzzy logic, smart grid, explainable AI

I. Introduction

Thermal power stations utilizing fossil fuels, including coal and oil, account for approximately 80% of global electricity generation, although this proportion is decreasing due to the shift towards renewable energy sources [1]. Conventional power systems necessitate a continuous balance between supply and demand; disruptions may compromise critical infrastructure and result in extended outages [2]. Insufficient real-time monitoring constrains operators' capacity to effectively address faults and variations in demand, underscoring the necessity for more adaptable and robust energy systems. Smart grids tackle these challenges through the integration of digital technologies, facilitating automation, real-time monitoring, and enhanced efficiency, while also supporting the integration of renewable energy and improving reliability [3]. The system integrates components including energy storage, power electronic interfaces, and advanced metering and communication systems, as depicted in Fig. 1 [4], facilitating real-time data acquisition and swift system response.

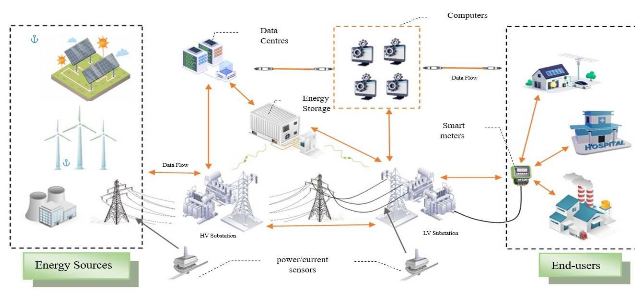

Fig. 1. Smart grid components. ([4])

The global renewable energy market is projected to grow substantially, necessitating around 3,700 GW of additional capacity [5]. The increase in growth heightens system complexity, thereby diminishing the efficacy of traditional control methods, which may lead to challenges in managing energy distribution and reliability in the face of fluctuating renewable energy sources. Artificial intelligence (AI), especially machine learning, facilitates data-driven decision-making, enhancing efficiency and lowering costs [6]. Many AI models function as “black boxes,” which restricts interpretability and trust, thereby rendering explainability crucial for their adoption in energy systems [7]. Explainable AI (XAI) enhances transparency, while fuzzy logic systems provide benefits owing to their rule-based and interpretable framework. Interval type-2 fuzzy logic systems (IT2FLSs) effectively address uncertainty in dynamic smart grid environments [8].

This paper presents an XAI framework that combines SHapley Additive Explanations (SHAP) with a Random Forest (RF) model to improve feature selection, rule pruning, and prediction reliability in IT2FLSs. SHAP assesses the significance of features, whereas RF identifies pertinent features. The IT2FLS employs a streamlined rule base to address noise and uncertainty in hourly demand data, facilitating precise predictions and enhancing real-time decision-making in smart grid applications. This paper is structured as follows: Section II outlines the proposed method, Section III analyzes the experimental results, and Section IV concludes the paper while suggesting directions for future research.

II. The Proposed Type-2 Fuzzy-Based Feature-Driven Rule Reduction Approach for Demand FORECASTING WITHIN Energy Smart Grid Systems

Our proposed framework uses feature-driven input selection to identify the most important factors for predicting next-hour kWh consumption. This simplifies fuzzy rules, improves storage efficiency, and improves prediction accuracy. Additionally, the method enhances operational decision-making and builds user trust by providing a small set of concise rules that are easy to analyze and interpret. Fig. 2 shows an overview of the proposed system, which will be explained in detail in the following subsections.

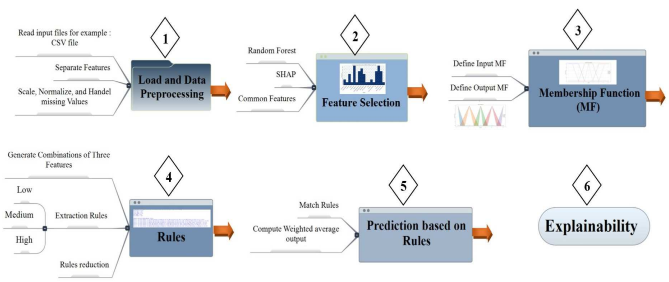

Fig. 2. An Overview of the Proposed System.

A. Load and Preprocessing Data

The data preprocessing steps are as follows:

1) Data Loading: Consumption data is loaded using Pandas DataFrame in Python, which enables data manipulation, missing value handling, and structuring for analysis.

2) Data Transformation: After reading the CSV file via pd.read_csv, additional columns such as minimum, maximum, and total kWh consumption, along with the target kWh, are added.

3) Datetime Conversion: The timestamp column is transferred to a datetime object to ensure accurate modeling of times and dates for visualization and time series forecasting.

4) Feature Engineering: Additional features, such as weekdays and months, are generated to assist in identifying periodic trends, consumption patterns, and variations affecting energy usage.

5) Data Cleaning and Scaling: Filling missing values, applying one-hot encoding to categorical variables, and scaling numerical data help create a complete and consistent dataset with detailed temporal and contextual features. This forms the basis for feature selection and Type-2 fuzzy modeling.

B. Feature Selection

Feature selection is essential for reducing model complexity, enhancing computational efficiency, and improving interpretability in smart grid demand forecasting. The dataset contains load measurements and contextual factors, but not all variables equally impact prediction accuracy [9]. Choosing the most relevant features ensures the model focuses on inputs that strongly influence load behavior, which is critical for accurate and reliable smart grid operation [10]. The proposed framework employs a feature-driven selection approach that integrates multiple techniques, as explained in the following subsections.

1) Random Forest Regressor Model

Breiman introduced Random Forest (RF) in 2001 to tackle classification and regression tasks by combining multiple learning algorithms [11]. The RF regressor combines an ensemble of decision trees, each trained on a randomly selected subset of features from a bootstrap sample of the original dataset. The final prediction is made by averaging the output of all trees, as outlined in:

$$\hat{y}_{RF} = \frac{1}{B}\sum_{B}^{1} h_b(X) \tag{1}$$

Here $h_b(X)$ represents the output from the $b$the decision tree, and B Is the total number of trees in the ensemble. Equation (2) shows that an accurate forest requires trees to be both individually precise and minimally correlated [11].

$$PE^*(forest) \le \bar{\rho} \cdot PE^*(tree) \tag{2}$$

Where:

- $\bar{\rho}$ = average correlation between trees
- PE^* = prediction error

One significant benefit of RF is that it can provide feature importance rankings. According to [12], the Mean Decrease Impurity (MDI) of a feature X_m Is defined as:

$$Imp(X_m) = \sum_{k=0}^{p-1} \frac{1}{C_p^k} \frac{1}{p-k} \sum_{B\in P_k(V^{-m})} I\,(X_m; Y \mid B) \tag{3}$$

where $P_k(V \setminus \{X_m\})$ is the set of subsets of size k of the other features, and $I(X_m; Y \mid B)$ is the conditional mutual information of X_m and the output Y given B. The sum of all feature importances equals the mutual information between all input variables and the output:

$$\sum_{m=1}^{p} Imp\,(X_m) = I(X_1, \ldots, X_p; Y) \tag{4}$$

This decomposition captures all possible interactions, providing a rigorous theoretical justification for using MDI to rank features, while RF remains robust, accurate, and interpretable for variable selection.

2) SHapley Additive exPlanation Feature Importance

SHapley Additive exPlanation (SHAP) values provide a principled framework to interpret machine learning models by quantifying the contribution of each feature to a specific prediction. According to cooperative game theory, each feature receives a “credit” for its marginal contribution to the model output. SHAP values explain model behaviour by examining all feature combinations and capturing direct impacts and feature interactions. For a single instance x, the Shapley value ϕ_j of the feature j is computed as [13]:

$$\phi_j(f, x) = \sum_{S\subseteq F\setminus\{j\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} \left[f_{S\cup\{j\}}(x_{S\cup\{j\}}) - f_S(x_S)\right] \tag{5}$$

Where:

- F is the set of all features
- S is a subset not containing j
- $f_S(x_S)]$ is the model prediction using only the features in subset S.

Classic Shapley value estimation methods, including Shapley regression values, Shapley sampling values, and Quantitative Input Influence (QII), operationalize this principle by either retraining models on all subsets or applying Monte Carlo approximations, while satisfying three key properties: local accuracy (the sum of feature contributions equals the model output [13]:

$$f(x) = \phi_0 + \sum_{i=1}^{M} \phi_i \quad (6)$$

Missingness (features absent from the input receive zero contribution) $x_i = 0 \implies \phi_i = 0$, and consistency (a feature's contribution cannot decrease if its marginal effect increases). These properties uniquely determine the additive feature attribution method, providing a strong theoretical foundation for SHAP values. In practice, for ensemble tree models such as XGBoost, the Tree SHAP algorithm computes Shapley values efficiently across all trees [13]:

$$\phi_j = \frac{1}{N} \sum_{t=1}^{T} \left| \phi_j{}^{(t)} \right| \quad (7)$$

Where $\phi_j{}^{(t)}$ is the contribution from tree t, capturing feature interactions with reduced computational complexity [13]. To assess global feature importance across the dataset, the absolute Shapley values are averaged over all N samples:

$$SHAP_{importance}(j) = \frac{1}{N} \sum_{i=1}^{N} \left| \phi_j(i) \right| \quad (8)$$

This provides a ranking of features by their overall impact on model predictions.

3) *Common Features Between Random Forest and Shapley Additive ExPlanation*

Random Forest (RF) and SHapley Additive Explanations (SHAP) provide rankings of feature relevance; however, these rankings might be different when features are linked or when interactions are complicated [14].

Let the top-k features selected by RF be:

$$F_{RF} = \{f_1, f_2, \dots, f_k\} \quad (9)$$

and the top-l features selected by SHAP are:

$$F_{SHAP} = \{f_1, f_2, \dots, f_l\} \quad (10)$$

The intersection of these sets obtains the common feature subset:

$$F_{Common} = F_{RF} \cap F_{SHAP} \quad (11)$$

Selecting F_{Common} yields a simple and robust feature set with the most relevant variables from both techniques. This subset simplifies Type-2 fuzzy modeling, increases computing efficiency, and preserves predictive accuracy. This hybrid methodology reduces feature ranking inconsistencies, improving the interpretability of smart grid demand forecasting and its resilience [14], [15].

C. *Membership Function Generation*

Fuzzy systems rely on mapping crisp numerical inputs into linguistic labels (e.g., *low*, *medium*, *high*) through membership functions (MFs) [8].

Each input vector at time T is defined as:

$$x = (x_1, x_2, x_3 \dots \dots, x_{18})^T \quad (12)$$

where x_i corresponds to a feature such as facility load, gas consumption, heating load, or average consumption [8].

An interval, Type-2 Fuzzy Set $\tilde{A}$ is formally defined as:

$$\tilde{A} = \int_{x \in X} \int_{u \in J_x} 1/(u, x) \quad J_x \subseteq [0,1] \quad (13)$$

For IT2F sets, we only specify upper and lower MFs $(\bar{\mu}_{\tilde{A}}, \underline{\mu}_{\tilde{A}})$:

$$\bar{\mu}_{\tilde{A}} \leq \mu_{\tilde{A}} \leq \underline{\mu}_{\tilde{A}} \quad (14)$$

The region between $(\underline{\mu}_{\tilde{A}}, \bar{\mu}_{\tilde{A}})$ is the Footprint of Uncertainty (FOU):

$$FOU(\tilde{A}) = \cup_{x \in X} [\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x)] \quad (15)$$

The secondary MF can be expressed in terms of upper and lower MF as:

$$\mu_{\tilde{A}}(x) = \int_{u \in J_x [\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x)]} 1/u \quad (16)$$

To generate interval trapezoidal type-2 MFs for each input (x), the first step is to define the $[x_{min}, x_{max}]$ for each x. Then divide equally into low, medium, and high intervals. Mathematically, the lower $\underline{\mu}_{\tilde{A}}(x)$ and upper $\bar{\mu}_{\tilde{A}}(x)$ MFs for trapezoidal will be defined as:

$$\mu_{\tilde{A}}(x) = \begin{cases} 0, & x \leq a \text{ or } x \geq d \\ \frac{x-a}{b-a}, & a < x \leq b \\ 1, & b < x \leq c \\ \frac{d-x}{d-c}, & c < x < d \end{cases} \quad (17)$$

Fig. 3 represents the upper (a_U, b_U, c_U, d_U) and lower (a_L, b_L, c_L, d_L) correspond to the critical points of the trapezoidal MFs.

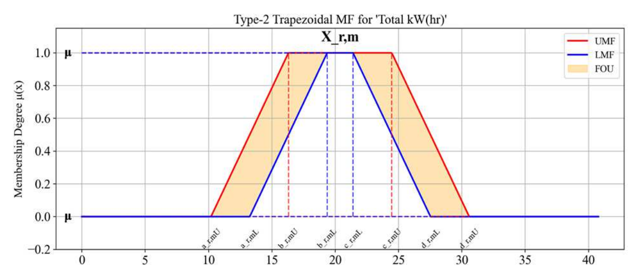

Fig. 3. Trapezoidal MFs

D. *Rules extraction and Pruning*

Fuzzy rule-based systems are fundamental to Explainable Artificial Intelligence (XAI), as they combine predictive accuracy with human interpretability [10]. In the proposed IT2 FLS, the rule base is constructed from IF–THEN rules extracted from the input features. Each input x_i is mapped to a linguistic term represented by an Interval Type-2 Fuzzy Set (IT2 FS), which captures uncertainty through its Footprint of Uncertainty (FOU) [8], [16].

The relationship between the output $y = (y_1 \dots, y_k)$ and the input $x = (x_1 \dots, x_n)$ It is expressed as in [19]:

IF x_1 is $\widetilde{A}_1^l$ … and x_n is $\widetilde{A}_n^l$ THEN y_1 is $\widetilde{B}_1^l$ … and y_k is $\widetilde{B}_k^l$ (18)

The dataset uses fixed input-output $(x^{(t)}; y^{(t)})$ pairs, $(t = 1, \dots N)$, compute the upper $\overline{\mu}_{\tilde{A}}(x)$ end lower $\underline{\mu}_{\tilde{A}}(x)$ membership values for each fuzzy set and input variable.
To find the centre of gravity $(\mu_{\tilde{A}_s^q}^{cg})$ at $(x_s^{(t)})$:

$$\mu_c^{cg}\left(x_s^{(t)}\right) = \frac{1}{2}\left(\underline{\mu}_{\widetilde{A}_s^q}\left(x_s^{(t)}\right) + \overline{\mu}_{\widetilde{A}_s^q}\left(x_s^{(t)}\right)\right) \tag{19}$$

For each input, type-2 fuzzy sets are used to define it, enabling many possible rules. Nonetheless, only rules with dominant regions that contain data points are generated. Initially, a rule is created for each input-output pair by selecting the fuzzy set with the highest membership value [16]. This is not the final rule, as further refinement and weighting can be calculated as follows:

$$wi^{(t)} = \prod_{s=1}^{n} \mu_{\tilde{A}_s^q}^{cg}(x_s(t)) \tag{20}$$

The completion of the Fuzzy Rule Extraction process, including detailed methodologies for generating and refining rules using type-2 fuzzy sets, is thoroughly discussed in [16].

A straightforward, effective method for pruning the rules, regardless of the shape of the MFs. The Takagi-Sugeno-Kang (TSK) fuzzy regression model with R rules are the basic idea is to identify rules that are similar enough, combine them, and then re-tune all remaining rules [17]. Start from the rule R_0 , iterative pruning (for $t = 2, \dots, T$) Compute normalized firing $\bar{f}_r$ levels for each r:

$$\bar{f}_r(x) = \frac{f_r(x)}{\sum_{k=1}^{R} f_k(x)}, \quad \mathrm{r} = 1, \dots, \mathrm{R}, \mathrm{n} = 1, \dots, \mathrm{N} \tag{21}$$

Compute average normalized firing strength per rule:

$$\bar{f}_r(x) = \frac{1}{N}\sum_{n=1}^{N} \bar{f}_r(x_n) \tag{22}$$

Remove rules if their normalized firing strength is too low [17]:

$$\bar{f}_r < \gamma \cdot median\{\bar{f}_k\}_{k=1}^{R} \tag{23}$$

Update the rule base and denote the remaining rules as R, and construct $S \in R^{R \times R}$. The following equation represents the similarity of the normalized firing levels between rules i, and j:

$$s(i,j) = \frac{\min(\bar{f}_i, \bar{f}_j)}{\max(\bar{f}_i, \bar{f}_j)}, \ i \neq j \tag{24}$$

A large value of the ratio represents a more substantial overlap between rules. While a weighted average is employed to update the consequent rule i (c_i, c_j) and weights rule j $is(w_i, w_j)$ [18]. Then merge similar rules, while $\max S > \theta$ update the rule i consequent as:

$$c_i \leftarrow \frac{w_i c_i + w_j c_j}{w_i + w_j}, w_i \leftarrow w_i + w_j \tag{25}$$

- Remove rule j and update S.
- Set R←R−1

E. *Prediction Based on Rules (Inference)*

The proposed framework uses a zero-order TSK fuzzy model, where each rule outputs a constant value c_r [18]. For a new input x, the average firing strengths are computed to find the output interval as:

$$y_r = \frac{\sum_{r=1}^{R} \bar{f}_r \, c_r}{\sum_{r=1}^{R} \bar{f}_r}, \ y_l = \frac{\sum_{r=1}^{R} \underline{f_r} \, c_r}{\sum_{r=1}^{R} \underline{f_r}} \tag{26}$$

The single crisp output is the midpoint of the interval:

$$y^{(x)} = \frac{y_l + y_r}{2} \tag{27}$$

F. *Explainability*

Explainability enables operators to understand how predictions are derived, supports root cause analysis, provides actionable insights, and facilitates timely decision-making [19]. Equations. 28 (center of the firing interval). Here, f_c measures of how strongly the merged rule is activated:

$$f_c = \frac{f_l + f_r}{2} \tag{28}$$

Each kWh prediction depends on the set of active rules R, with contributions weighted by their firing strength and centroid, and the total firing strength is:

$$S = \sum_{r \in R} f_c \tag{29}$$

For each rule r with firing strength f_c, its weight w (contribution to the final output) is computed by normalizing against the total firing strength S:

$$w = \frac{f_c \cdot C_r}{S} \tag{30}$$

Each rule is associated with a label reflecting a feature or feature value [4]. Contributions for a label l within a feature column c are aggregated as:

$$Contribution_{c,l} = \sum_{\substack{r \in R \\ label\ of\ r = l}} w \tag{31}$$

The maximum contribution is scaled to 100% and assigned to a fixed length, while all other contributions are scaled in proportion to this reference.

$$Contribution\% = \ v_c / v_{max} \times 100 \tag{32}$$

Where:v_c is the contribution of a feature and v_{max} is max (v_c) across all features.

III. EXPERIMENTS AND RESULTS

We employed a real-world dataset from residential neighborhoods, capturing residential and commercial energy usage over a one-year period [20]. The input vector consists of 26 features covering temporal (Time, Weekday, Month), environmental (Temperature, Dew Point, Humidity, Wind Speed, Wind Gust, Pressure, Precipitation), and operational parameters (Facility load, Gas:Facility, Heating, Cooling, HVAC components, Interior/Exterior lights, Water heater, Interior equipment). Historical consumption metrics such as Total kWh, Max_Consumption, and Average Consumption were also included. The output variable, Target kWh, represents the predicted energy consumption for the next hour.

The feature selection pipeline mentioned earlier was used to improve interpretability and model efficiency. Random Forest (RF) ranked features based on their direct effect on prediction accuracy, while SHAP values calculated with an XGBoost model identified features that contributed to explainable predictions. The intersection of features selected by both methods ensured strong rule generation (Fig. 4).

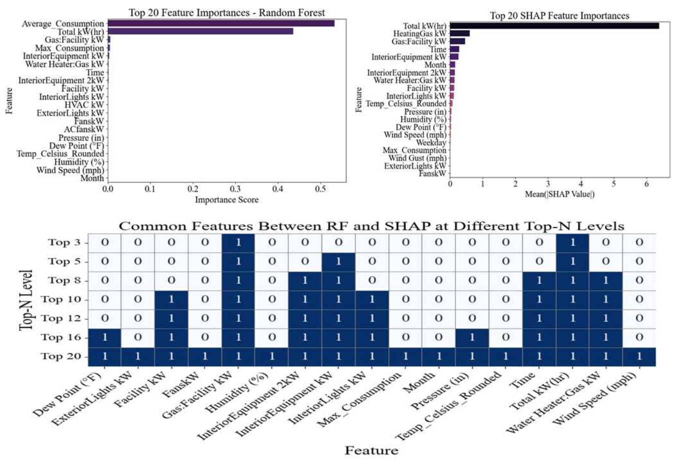

Fig. 4. Top 20 feature importances – RF – SHAP–common features

The prediction engine models uncertainty in real-world energy consumption using an Interval Type-2 Fuzzy Logic System (IT2 FLS). Each fuzzy rule has three antecedents mapped to input attributes (Low, Medium, High), with outputs described linguistically from "Very Low" to "Very High." Rules are derived from the top-ranked features via feature selection. Each input pattern activates rules, and zero-order TSK inference computes the prediction as a weighted average of outputs based on firing strengths. The model was tested with the top N feature subsets (N = 3, 6, 8), recording rule counts and predictive accuracy. Table I summarizes RMSE and rule counts.

TABLE I. RMSE VS. NUMBER OF RULES (COMMON RF-SHAP FEATURES)

Top N Features (SHAP)	Top N Features (RF)	Top N Features (Common)	Common - Testing RMSE	No. Rules
5	5	3	2.74	16
8	8	6	3.088 ▲	370
10	10	8	3.185 ▲	616

We compared the IT2 FLS with a Type-1 FLS and a black-box neural network. All three models were trained in 80% of the data and tested on the remaining 20%, with metrics reported on the test set only. For both fuzzy systems, rule bases were built from the three common features identified by RF and SHAP, ensuring a compact and consistent feature set. Using this approach, the IT2 FLS achieved a training RMSE of 2.696 and a test RMSE of 2.74 with just 16 rules—a 250-fold reduction compared to the 2,781 rules reported in [4]. The Type-1 FLS performed worse (test RMSE 5.376), while the neural network produced a lower test RMSE (1.410) but lacked interpretability. Fig. 5 shows actual versus predicted consumption for 50 test samples, illustrating that the IT2 FLS closely tracks real values and provides transparent, rule-based predictions, unlike the opaque neural network or less accurate Type-1 FLS.

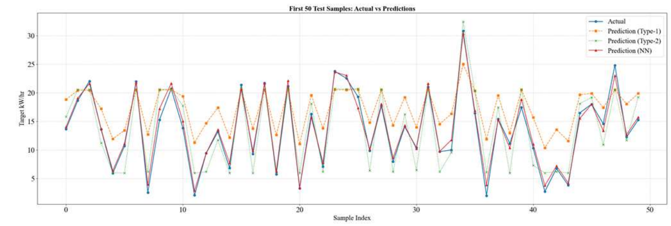

Fig. 5. Interval Type-2 FLS, and Type-1 FLS for 50 testing data samples.

To demonstrate the interpretability and diagnostic advantages of the IT2 FLS, we analyze two representative cases: Scenario 1, which shows unexpectedly low usage, and Scenario 2, which indicates unexpectedly medium consumption. These scenarios illustrate how feature contributions clarify unexpected results and how expert knowledge can enhance predictive accuracy and inform operational planning.

Scenario 1 – Unexpected Low Consumption: On Sunday, September, at 6:00 AM, with an ambient temperature of 11 °C, the IT2 FLS predicted energy consumption of 13.2 kWh (Low). The primary contributors were Low Interior_Equipment consumption, Low Gas Facility, and Medium Total_kWh (previous hour) (Fig. 6). Despite the previous hour's medium facility load, minimal appliance and gas usage led to lower overall consumption. This illustrates that early-morning loads are mainly driven by appliance and gas activity, helping operators understand the situation. An expert rule such as, "If previous hour Total_kWh is Medium but Interior_Equipment__kW and Gas_Facility trend Low, maintain Low prediction unless occupancy increases," could further improve forecasting accuracy. This scenario demonstrates the IT2 FLS's ability to clarify unexpected predictions and support managers in distinguishing between real efficiency gains and potential underutilization.

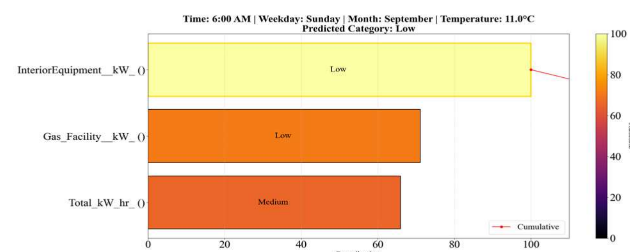

Fig. 6. Contributions of input features to low consumption prediction

Scenario 2: On Monday, November, at 6:00 PM, with an ambience temperature of 3 °C, the IT2 FLS predicted 19.7 kWh (Medium category). Feature contributions indicate that Interior Equipment__kW High is the main factor, followed by Medium consumption (Fig. 7). This result is unexpected since cold, peak-hour conditions typically lead to heating-dominated (Gas) loads; however, internal appliances and equipment mainly influence consumption. The IT2 FLS clarifies this pattern by quantifying per-feature contributions, enabling operators to investigate appliance schedules, server activity, or EV charging as potential causes. An expert rule such as "If InteriorEquipment__kW is High during peak hours, bias the prediction toward

Medium/High and flag for operational review” may be added to improve robustness.

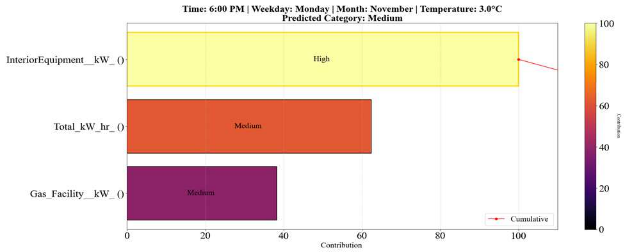

Fig. 7. Contributions of fuzzy rules to medium consumption prediction

This case demonstrates the model’s ability to identify subtle, actionable demand drivers, thereby aiding targeted load management efforts.

The scenarios illustrate the practical advantages of explainable predictions in operational decision-making. Accurate low forecasts facilitate efficient generation and storage, whereas moderate predictions assist in load balancing and the implementation of demand-response strategies, ultimately leading to improved energy management and cost savings for utilities and consumers. The findings indicate that IT2 FLS is appropriate for real-time smart grid operations, effectively merging accuracy, interpretability, and computational efficiency while delivering actionable insights through the integration of human expertise and data-driven principles.

IV. Conclusions and Future Work

This paper introduces a comprehensive framework for short-term demand forecasting in smart grids. It incorporates feature-driven rule selection, SHAP, and Random Forest for input prioritization, alongside Interval Type-2 Fuzzy Logic Systems (IT2 FLS) to address uncertainty and dynamic conditions. The proposed method effectively minimizes rule complexity by utilizing only 16 rules based on the top three selected features, resulting in a test RMSE of 2.74 kWh, in contrast to 5.376 for Type-1 FLS and 1.410 for a neural network. The system offers interpretability and computational efficiency by associating linguistic categories with feature contributions and utilizing zero-order TSK inference. This lets operators understand what affects predictions, find outliers, and make smart operational decisions in real time, all while keeping a balance between accuracy and explainability. The framework's effectiveness is illustrated through two representative scenarios that give concrete insights into energy management, including the identification of low and medium-consumption patterns and an understanding of their contributing factors. The examples illustrate the approach's facilitation of proactive decision-making, anomaly detection, and energy optimization within complex smart grid environments.

Future research will focus on integrating the IT2 FLS-based XAI model within a virtual smart grid framework using real weather and energy data. The system will connect to a real-time optimization platform to manage load balancing, demand response, and distributed energy resources (DERs). Continuous feedback will refine forecasts and explanations, enabling timely, data-driven operational adjustments and strategic planning for energy operators.

References

[1] Energy Inst., *Statistical Review of World Energy 2023*, London, UK, 2023. [Online]. Available: https://www.energyinst.org/statistical-review

[2] R.A.C. van der Veen and R.A. Hakvoort, "The electricity balancing market: Exploring the design challenge," *Utilities Policy*, vol. 43, pt. B, pp. 186–194, Dec. 2016. [Online]. Available: https://doi.org/10.1016/j.jup.2016.10.008

[3] M.S. Hossain, N.A. Madlool, N.A. Rahim, J. Selvaraj, A.K. Pandey, and A.F. Khan, "Role of smart grid in renewable energy: An overview," *Renew. Sustain. Energy Rev.*, vol. 60, pp. 1168–1184, Sep. 2016, doi: 10.1016/j.rser.2015.09.098.

[4] M. Alfayan and H. Hagras, "Towards an Explainable Artificial Intelligence approach for smart grid systems," *Discover AI*, vol. 5, art. no. 40, Apr. 2025.

[5] Int. Energy Agency, "Corrigendum: Renewables," IEA, Jan. 17, 2024. [Online]. Available: https://www.iea.org/reports/renewables-2023

[6] F. Doctor, H. Hagras, “Neuro type-2 fuzzy based method for decision making”. US Patent 8515884.

[7] V. Callaghan, M. Colley, H. Hagras, J. Chin, F. Doctor, G. Clarke, “Programming iSpaces: A tale of two paradigms”, “*Intelligent Spaces: The Application of Pervasive ICT”, (Eds: A. Steventon, S. Wright),* Springer-Verlag, Chapter 24, pp.389- 421, December 2005.

[8] J.M. Mendel and R.I.B. John, "Type-2 fuzzy sets made simple," *IEEE Trans. Fuzzy Syst.*, vol. 10, no. 2, pp. 117–127, Apr. 2002.

[9] C. Li, Z.Y. Dong, F. Luo, Y. Chen, Y. Xu, and K.P. Wong, "Feature selection for short-term load forecasting using data mining," *IET Gener., Transm., Distrib.*, vol. 11, no. 4, pp. 1031–1038, Mar. 2017, doi: 10.1049/iet-gtd.2016.1432.

[10] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, Mar. 2003.

[11] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[12] G. Louppe, L. Wehenkel, A. Sutera, and P. Geurts, "Understanding variable importances in forests of randomized trees," in *Proc. 26th Int. Conf. Neural Inf. Process. Syst. (NeurIPS)*, Lake Tahoe, NV, USA, 2013, pp. 431–439.

[13] S.M. Lundberg, *SHAP (SHapley Additive exPlanations) Documentation*. [Online]. Available: https://shap.readthedocs.io/en/latest/. [Accessed: 21 Jul. 2025]

[14] E. Ali and T.A. Zarma, "Energy demand forecasting for hybrid microgrid systems using machine learning models," *Proc. Eng. Technol. Innov.*, vol. 29, pp. 68–83, 2025.

[15] X. Feng and J. Wieland, "Integration of SHAP and Random Forest for feature importance analysis in precipitation prediction," *Environ. Sci. Technol.*, vol. 55, no. 2, pp. 1234–1245, 2021.

[16] K. Almohammadi, H. Hagras, B. Yao, A. Alzahrani, D.Alghazzawi, G. Aldabbagh,”A Type-2 Fuzzy Logic Recommendation System for Adaptive Teaching” Soft Computing, Vol.21, No.4. pp.965-967, 2017.

[17] D. Wu, "MBGD-RDA training and rule pruning for concise TSK fuzzy regression models," *IEEE Trans. Fuzzy Syst.*, early access, Oct. 08, 2021, doi: 10.1109/TFUZZ.2021.3118760.

[18] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, no. 1, pp. 116–132, Jan.–Feb. 1985.

[19] J. Senoner, S. Schallmoser, B. Kratzwald, S. Feuerriegel, and T. Netland, "Explainable AI improves task performance in human-AI collaboration," *arXiv*, Jun. 2024. [Online]. Available: https://arxiv.org/abs/2406.08271. [Accessed: 3 Sep. 2025]

[20] C. Goncalves, R. Barreto, P. Faria, L. Gomes, and Z. Vale, "Dataset of an energy community's consumption and generation with appliance allocation for one year (1.0)," *Zenodo*, 2022. [Online]. Available: https://doi.org/10.5281/zenodo.6778401