A Multi-Objective Multi-Constraint Explainable AI-based Approach for Smart Energy Grid Systems

A Multi-Objective Multi-Constraint Explainable AI-based Approach for Smart Energy Grid Systems

Mahmoud Alfayan, Hani Hagras
School of Computer Science and Electronic Engineering
University of Essex
Colchester, United Kingdom
malfayan@essex.ac.uk , hani@essex.ac.uk

Abstract

As the global shift toward sustainable energy accelerates, modern electrical grids are increasingly incorporating renewable sources like wind and solar PV. Moving from traditional centralized power to decentralized smart grids demands better optimization and forecasting methods to ensure reliability and efficiency. The inclusion of renewables, energy storage solutions, and bidirectional energy exchanges boosts the flexibility of smart grids and supports sustainable energy practices. Successfully planning and managing hybrid renewable energy systems (HRES) ***requires accurate and clear forecasting methods in conjunction with robust optimization strategies.***

In this paper, we present a two-stage Multi-Objective Multi- Constraint Evolutionary computing based on a Genetic Algorithm and Non-dominated Sorting Genetic Algorithm III (GA-NSGA-III) optimization, which enables the procurement of renewable energy sources while concurrently minimizing the Levelized Cost of Energy (LCOE) and the Expected Energy Not Supplied (ENS). We have performed various experiments with real-world data in which, for cost-optimal configurations, NSGA-III decreases LCOE by 26.1% and worst-year ENS by 58.7% relative to NSGA-II. At balanced design, the ENS decreases by 72.7%, and the LCOE falls by 16.3%. This proposed system promotes the utilization of renewable energy sources while also enhancing the efficiency of storage. To enhance the design of hybrid renewable energy systems under uncertain conditions, the proposed method integrates multi-objective optimization, absence-aware reliability modeling, and interpretable load forecasting. These results support more consistent operation and simplify decision-making across multi-year planning horizons.

Load forecasting is provided using an Explainable Artificial Intelligence (XAI)-based Interval Type-2 Fuzzy Logic System (IT2FLS). Compared to Type-1 FLS, IT2FLS decreases RMSE by 48.6% and mean absolute error by 50.2% and gains R^2 from 0.625 to 0.901. These findings show that advanced XAI based multi-objective optimization improves system reliability and resilience.

1. Introduction

The global transition to sustainable energy systems has expedited the incorporation of renewable energy sources, including solar photovoltaic (PV) and wind power, into contemporary electrical grids [1]. Conventional power systems depended on centralized generation, consistent load demands, and one-way power flows [1],[2]. The growing integration of distributed renewable energy sources has converted traditional grids into intricate, decentralized systems that necessitate sophisticated monitoring, forecasting, and optimization approaches [3]. Smart grids improve conventional networks through the integration of digital communication technologies, advanced control systems, and distributed energy resources, resulting in enhanced efficiency, reliability, and longevity (see Figure.1) [2],[4]. They facilitate bidirectional energy flows between producers and consumers, promoting the integration of renewable energy, deployment of energy storage, and dynamic demand management, thereby assisting in the balance of supply and demand while reducing environmental impacts [2],[3].

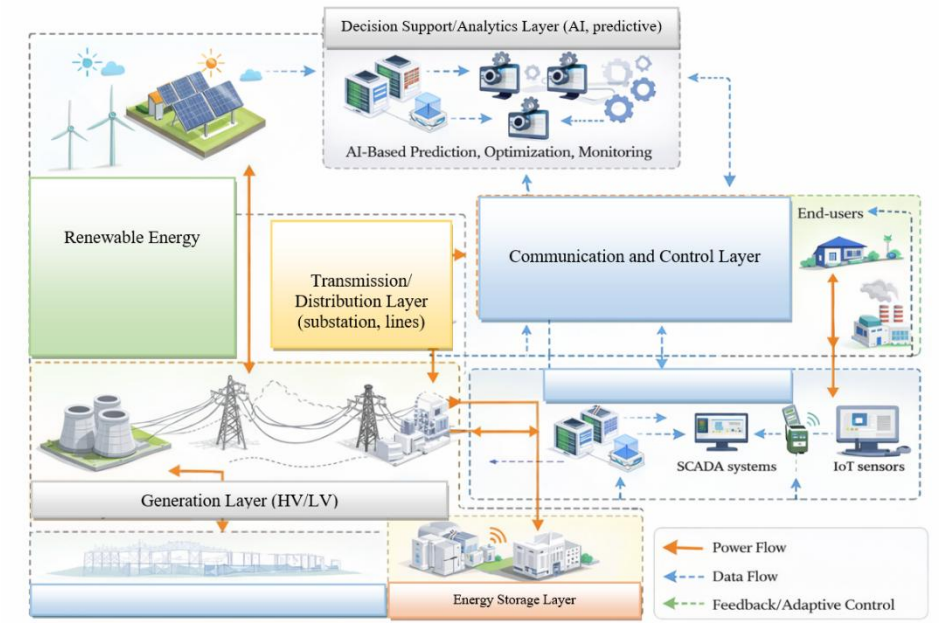

Fig. 1. Smart Grid System

High The adoption of renewable energy at high levels, while beneficial for reducing emissions and enhancing sustainability, presents significant planning and operational challenges for modern power systems [5]. Renewable energy sources, particularly photovoltaic (PV) and wind power, exhibit inherent variability and are contingent upon weather conditions, resulting in fluctuations that can affect grid stability and reliability [6],[7]. An illustrative example is the duck curve observed in California, characterized by substantial midday solar generation that significantly reduces net system demand (total demand minus renewable generation). This phenomenon leads to a distinct dip in demand during midday and a pronounced increase in the evening as solar output declines and electricity demand escalates [5], [7]. The rapid increase in demand places additional pressure on dispatchable generation

assets and system managers. Wider trends, including electrification, urban growth, and the increasing prevalence of distributed loads—such as electric vehicles and intelligent devices, contribute to the unpredictability of electricity consumption patterns [6], [7]. Thus, the design of renewable-based power systems requires management of variability and the determination of appropriate generation asset capacity. System planners must evaluate the quantity and capacity of photovoltaic panels and wind turbines required to adequately meet load demands across various operating scenarios. Figure 2, based on CAISO data reported by the U.S. Energy Information Administration, illustrates daily net-load variations caused by increasing solar generation, including the midday reduction in net load and the steep evening ramping requirement [5].

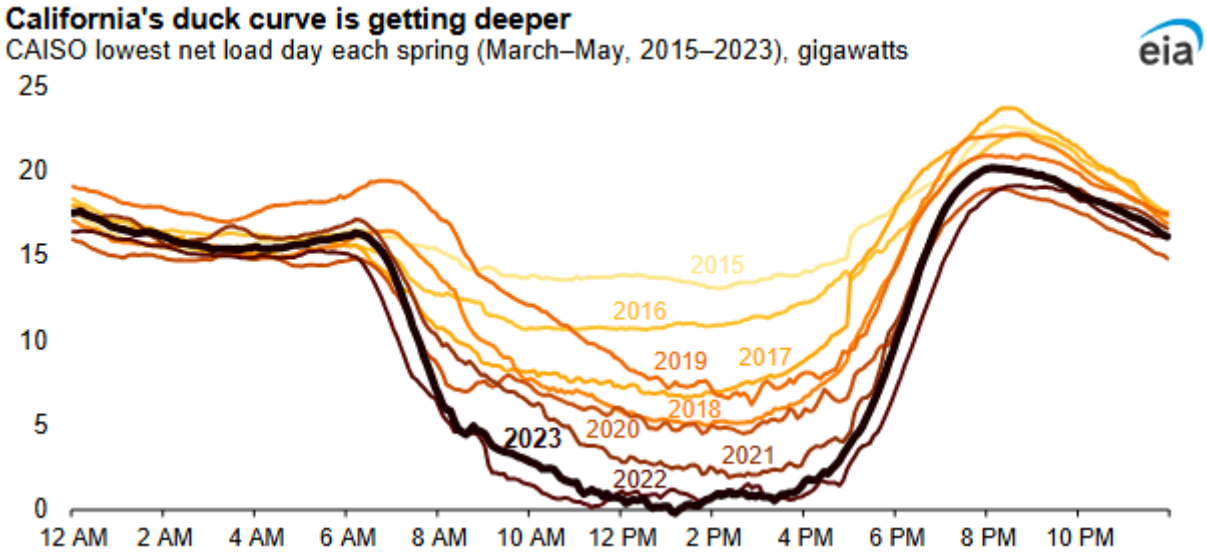

Fig 2. Deepening the California duck curve associated with increasing solar generation. The figure shows the reduction in midday net load and the steep evening ramping requirement caused by higher solar integration. Data source: CAISO, adapted from EIA [5].

Renewable generation is increasingly integrated with energy storage technologies to balance supply and demand and improve system flexibility [6],[7]. Establishing the optimal storage capacity remains a significant design challenge. Storage systems must be appropriately sized to accommodate surplus renewable energy during periods of high generation, such as midday solar peaks, and to discharge this energy during periods of low renewable production or heightened demand, such as in the evening. Excessive storage capacity can significantly increase capital expenditure and reduce the economic viability of the system. Consequently, system planning must concurrently determine the optimal configuration of renewable generation, including the quantity and capacity of photovoltaic arrays and wind turbines, as well as the appropriate battery storage capacity required to satisfy the load and support charging and discharging cycles. Additionally, practical factors such as battery degradation, constraints on charging and discharging rates, depth-of-discharge limitations, and various operational restrictions contribute to the complexity of sizing and managing storage systems [7],[8].

Planning renewable-dominant smart grids entails balancing conflicting objectives, including the minimization of the levelized cost of electricity (LCOE), reduction of energy not served (ENS), limitation of renewable curtailment, and maximization of renewable utilization [8],[9]. Multi-objective optimization methods are essential for determining configurations that equilibrate technical, economic, and environmental considerations [8]. Accurate and reliable load forecasting data is essential, as inaccuracies can lead to inadequate storage and increased Energy Not Supply (ENS) or excessive capacity and elevated costs [9].

Recent advancements in artificial intelligence (AI) and data-driven methods have demonstrated their utility in modeling complex, uncertain load behaviors, enabling predictive models that capture nonlinear dependencies between historical load and weather data [10-12]. However, many AI approaches operate as "black-box" models, offering limited interpretability regarding the reasoning behind predictions [10,11].

To overcome limitations in energy system planning, Explainable Artificial Intelligence (XAI) techniques have garnered significant attention. Interval Type-2 Fuzzy Logic Systems (IT2FLS) offer a balance between predictive accuracy and interpretability [11]. Incorporating uncertainty into membership functions allows IT2FLS to more effectively capture variations in load behavior and environmental conditions compared to conventional Type-1 fuzzy systems [13]. Their rule-based framework delineates explicit connections between inputs and outputs, rendering them especially appropriate for smart grid applications [13]. The incorporation of IT2FLS-based load forecasts into multi-objective optimization frameworks allows designers of hybrid renewable energy systems (HRES) to enhance decision-making processes. Forecasts incorporating uncertainty and offering interpretability enable optimization algorithms to identify configurations of renewable generation and storage that effectively balance system cost, reliability, renewable curtailment, and environmental impact [14],[15].

The design of HRES in smart grids can be framed as a multi-objective optimization problem, involving competing objectives including the minimization of the levelized cost of electricity (LCOE), reduction of energy not supplied (ENS), limitation of renewable curtailment, and maximization of renewable penetration. Optimization methods should determine solutions that balance these objectives while adhering to operational and environmental constraints [16].

This paper is organized as follows: Section 2 presents a brief overview of optimization techniques in smart grids and renewable energy systems. Section 3 presents a brief overview of Explainable AI and Type-2 Fuzzy Logic Systems. Section 4 presents the proposed type-2 fuzzy framework for multi-objective Optimization of smart grid applications. Section 5 presents the experiments and results. Section 6 presents the conclusions and future work.

2. Overview of Optimization Approaches in Smart Grid and Renewable Energy

The design and planning of hybrid renewable energy systems (HRES) and smart grids involve various technological, economic, and environmental objectives. Traditional deterministic optimization methods frequently demonstrate less-than-ideal performance due to nonlinear interactions, random generation patterns, and complex operational constraints within the renewable energy sector.

Previous research predominantly employed linear programming (LP) [16], mixed-integer linear programming (MILP) [17], and dynamic programming (DP) [18]. These techniques produce accurate outcomes for linear or discretized systems and demonstrate computational efficiency; however, they are inadequate in the presence of uncertainty or nonlinearity, requiring substantial simplification [19], [20].

2.1 Genetic Algorithms (GA) in Smart Grid (SG)

Genetic Algorithms (GAs) provide an adaptable alternative to traditional methods by simulating the processes of natural evolution and selection [21]. Instead of immediately identifying an exact solution, genetic algorithms generate a collection of potential solutions and iteratively refine them across generations. Each candidate, termed a "chromosome," signifies options such as the number of solar panels, the size of wind turbines, and the capacity of battery storage. The solutions are evaluated using objective functions that denote system performance indicators, including overall cost, reliability, and the proportion of energy derived from renewable sources. The optimization problem can be expressed mathematically as [19], [21]:

$$\min F(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_n(\mathbf{x})], \quad \text{s.t.} \quad g_i(\mathbf{x}) \leq 0,\ i = 1,2, \dots, m \tag{1}$$

Where:

- $F(\mathbf{x}) = [f_1(\mathbf{x}), \dots, f_n(\mathbf{x})]$ denotes multiple objective functions to be minimized.
- $g_i(\mathbf{x}) \leq 0$ are inequality constraints.
- $\mathbf{x}$ represent the vector of decision variables.

Genetic algorithms operate through three primary processes for solution generation: selection identifies the most effective solutions; crossover combines elements from two solutions to create new variations; and mutation introduces small random changes to investigate alternative possibilities. This evolutionary approach enables genetic algorithms to efficiently explore nonlinear and discrete search spaces, rendering them ideal for optimizing the characteristics of renewable energy systems. Traditional genetic algorithms are designed primarily for single-objective problems. The integration of multiple objectives into a weighted sum can guide solutions toward specific goals while concealing the complete range of trade-offs.[20].

2.2 Multi-Objective Evolutionary Optimization: Non-dominated Sorting GA (NSGA-II)

In complex renewable energy contexts, advanced optimization methods are often necessary to determine several near-optimal solutions that balance conflicting objectives. This domain often utilizes Multi-Objective Evolutionary Algorithms (MOEAs). NSGA-II is commonly employed in multi-objective optimization, especially within the fields of renewable energy and smart grids [19]. NSGA-II aimed to overcome the shortcomings of its predecessor, NSGA-I, which demonstrated significant processing requirements, an absence of elitism, and dependence on a single parameter. This method produces an initial array of alternative solutions across multiple generations, promoting convergence to the Pareto-optimal front while preserving solution diversity. The NSGA II method focuses on two main operators:

- Fast Non-Dominated Sorting involves assigning a rank to each solution according to Pareto dominance, thereby categorizing candidates into distinct non-dominated fronts. Solutions in the first front are the best, while the second front is only surpassed by the first front, and this pattern continues. This ranking establishes a hierarchy of solution quality.
- Crowding Distance (CD) is an operator that assesses the density of solutions around a candidate to maintain diversity within each front. A greater crowding distance signifies that a solution is in a less congested area of the objective space, thereby facilitating uniform representation of the Pareto front. The CD is calculated as [19]:

$$d_j(k) = \sum_{i=1}^{n} \frac{f_i(k-1) - f_i(k+1)}{f_i^{max} - f_i^{min}} \tag{2}$$

- Let d represent the distance, where n denotes the number of objectives, and $f_i(k)$ indicates the value of the i-th objective for the k-th neighboring solution. Fig. 3 depicts the methodology for coding and calculating crowding distance.

```
l=|Γ| // number of solutions in Γ (non-dominated set)

For each i∈Γ
   Γ[i]distance=0 // initialize the distance of solution i

For each objective m
   Γ=sort(Γ,m) // sort the non-dominated set based on the value of each objective function
   Γ[0]distance=Γ[l]distance=0 // boundary points are always selected
For i = 2 to (l-1) // for all other points
   Γ[i]distance=Γ[i]distance+((Γ[i+1].m-Γ[i-1].m)/(fmax-fmin))
```

Fig 3. The algorithmic process of the crowding distance criterion [19]

NSGA-II merges non-dominated sorting with crowding distance to achieve convergence toward the Pareto-optimal front while preserving solution diversity. This approach is particularly effective for designing multi-objective energy systems, as it is essential to reconcile cost, reliability, and renewable

integration [20], [22]. The increase in the number of optimisation objectives limit the effectiveness of the NSGA-II where as objectives increase, several solutions become non-dominated, diminishing selection pressure and complicating the differentiation of candidate solutions [19]. As a result, the crowding distance mechanism fails to maintain variety in high-dimensional objective spaces when the population size is limited or the objectives are strongly connected [19]. The Non-dominated Sorting Genetic Algorithm III employs a reference-point-based selection technique to maintain variety and enhance efficiency in many-objective optimization [23].

2.3 Multi-Objective Evolutionary Optimization: (NSGA-III)

The limitations of NSGA-II prompted the development of NSGA-III [23]. In scenarios involving four or more objectives, NSGA-III enhances NSGA-II by integrating a unique set of reference points within the objective space, which facilitates a diverse array of solutions and supports the selection process. The reference points established by NSGA-III are uniformly distributed throughout the objective space. The use of reference points ensures a consistent distribution of Pareto solutions generated during the selection phase across the goal space. Enhancing systems that balance conflicting objectives, including cost, reliability, and renewable energy use, can improve performance and facilitate better decision-making in complex areas such as resource allocation and energy management. Optimal outcomes require a balance among these objectives [23]. NSGA-III consists of the following primary steps:

1. Each point in the normalized objective space serves as a reference point. The total number of reference points H for M goals divided into p parts is calculated using the following [20]:

$$H = \binom{M + p - 1}{p} \tag{3}$$

2. Population Merging and Sorting involve the integrating Parent (P_t) and offspring (Q_t) populations are merged into a combined set, R_t, which is subsequently organized non-dominated fronts.

3. Selection through Reference Points: Individuals from the final non-dominated front are allocated to reference points. Adaptive normalization transforms the ideal point $\mathbf{z} = (z_{\min 1}, z_{\min 2}, \dots, z_{\min M})$ to zero, thereby normalizing the objectives [23]:

$$f_i'(x) = f_i(x) - z_{\min i}, i = 1, \dots, M \tag{4}$$

$$f_{n_i}(x) = \frac{f_i'(x)}{a_i - z_{\min i}}, \sum_{i=1}^{M} f_{n_i}(x) = 1 \tag{5}$$

where a_i is the intercept of the hyperplane that connects the extreme points of the objectives.

4. Individual Retention: When the archive surpasses the population size K, reference point correlation and selection facilitate a uniform and representative distribution of solutions across the Pareto front [23].

NSGA-III enables a structured examination of trade-offs between multiple objectives, allowing designers to grasp how different combinations of PV arrays, wind turbines, and energy storage influence system performance under varying generation and demand conditions. It can pinpoint setups that uphold supply-demand equilibrium during times of diminished solar radiation or limited wind availability. Combining NSGA-III with Interval Type-2 Fuzzy Logic Systems (IT2FLS) enhances decision-making clarity, guaranteeing that chosen system configurations are both technically feasible and practically applicable. Fig 4 depicts the Pareto front in a two-objective context, emphasizing optimal trade-offs, non-dominated solutions, and dominated results. NSGA-III solutions demonstrate improved coverage and diversity compared to NSGA-II, providing planners with a broader array of possible HRES designs.

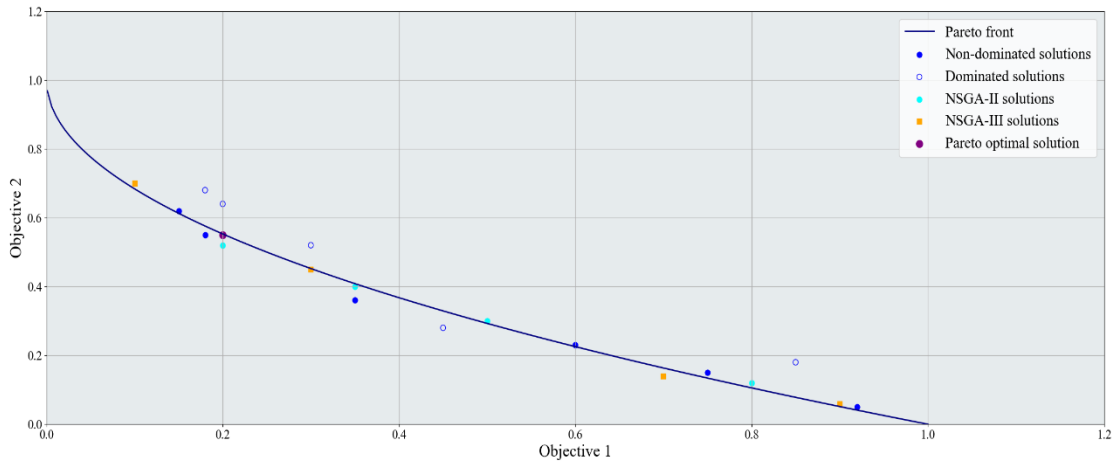

Fig. 4. An Example Pareto-optimal solutions of NSGA-II and NSGA-III evolutionary multi-objective optimization methods.

The development of smart grids and hybrid renewable energy systems has transitioned from traditional linear methodologies to sophisticated evolutionary strategies. Genetic Algorithms provide flexible solutions for nonlinear systems. NSGA-II identifies trade-offs for two or three objectives, whereas NSGA-III addresses many objective problems with improved diversity and coverage. These algorithms, especially when integrated with forecasting systems like IT2FLS, enable planners to make informed, resilient, and sustainable energy planning decisions.

3. Overview of Explainable AI (XAI) Using Type-2 Fuzzy Logic System.

Fluctuations in energy consumption and shifting climatic conditions hinder the accuracy of demand forecasting for renewable energy systems. The lack of transparency in traditional machine learning

models complicates their understanding and implementation, despite their high predictive accuracy. Explainable Artificial Intelligence (XAI) aims to develop models and outcomes that could be easily understood, analyzed, and augmented by the relevant business stakeholders [11].

Fuzzy logic systems are recognized as understandable methodologies due to their utilization of IF-THEN constructs. Human operators can readily comprehend these requirements. Interval Type-2 Fuzzy Logic Systems (IT2FLS) enhance Type-1 fuzzy systems by handling uncertainty via the Footprint of Uncertainty (FOU), thereby increasing their resilience to input noise and variations in energy consumption [24]. Fuzzy rule-based systems provide a natural framework for human-understandable AI, enhancing trust and decision-making in smart grid operations [24], [25].

The IT2FLS generates interpretable load predictions while handling uncertainty through a sequential procedure (Fig. 5). An interval Type-2 fuzzy set (IT2FS) is formally defined as [24]:

$$\tilde{A} = \int_{x\in X} \int_{u\in J_x} \mu_{\tilde{A}}(x,u)/(u,x), J_x \subseteq [0,1] \tag{6}$$

where x is the primary variable in the universe of discourse X, u is the secondary membership grade, and $J_x$defines the interval of possible primary membership values. The union of all $J_x$intervals form the Footprint of Uncertainty (FOU), capturing input variability, measurement noise, and imprecise expert knowledge [13], [24].

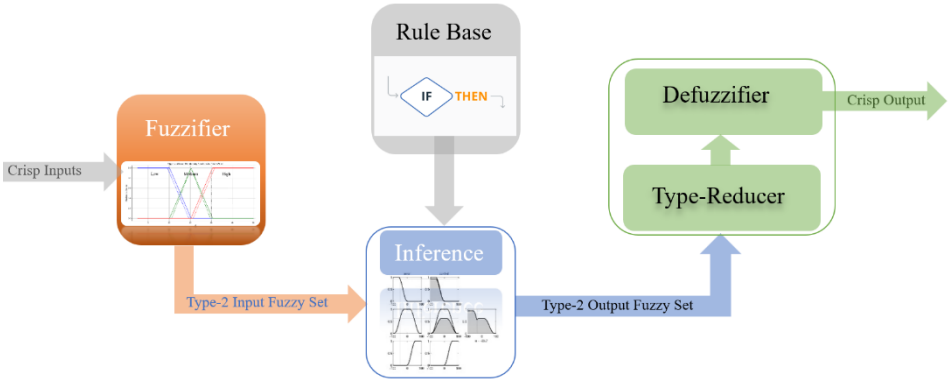

Fig. 5. Type-2 Fuzzy Logic System

The IT2FLS procedure begins with fuzzification, mapping crisp inputs into interval Type-2 fuzzy sets using predefined membership functions. These fuzzy inputs are evaluated by a rule base, which contains Lmulti-input multi-output (MIMO) linguistic IF–THEN rules [13], [24], [25]:

$$R_l: \text{IF } x_1 \text{ is } \tilde{A}_{1l} And \cdots And\ x_n \text{ is } \tilde{A}_{nl} \text{ THEN } y_1 \text{ is } \tilde{B}_{1l} And \cdots and\ y_k \text{ is } \tilde{B}_{kl}, l = 1,2,\dots,L \tag{7}$$

where $\tilde{A}_{il}$and $\tilde{B}_{jl}$are interval Type-2 fuzzy sets associated with the i-th input and j-th output. The inference engine computes the activation strength of each rule using a t-norm operator (minimum or product), producing interval Type-2 fuzzy outputs.

Since the outputs of an IT2FLS are themselves fuzzy sets, they cannot be directly defuzzified. A type-reduction step is performed to convert the interval Type-2 outputs into an interval Type-1 fuzzy set, which is then defuzzified to produce a crisp predicted load [13], [25]. To illustrate the procedure, consider a crisp hourly load of 16 kWh within a universe of discourse $[0, 40]$ kWh (see Fig. 6. and Table 1.) Using trapezoidal membership functions for the linguistic terms Low, Medium, and High, the fuzzification step produces the interval memberships values and their averages shown in Table 1.

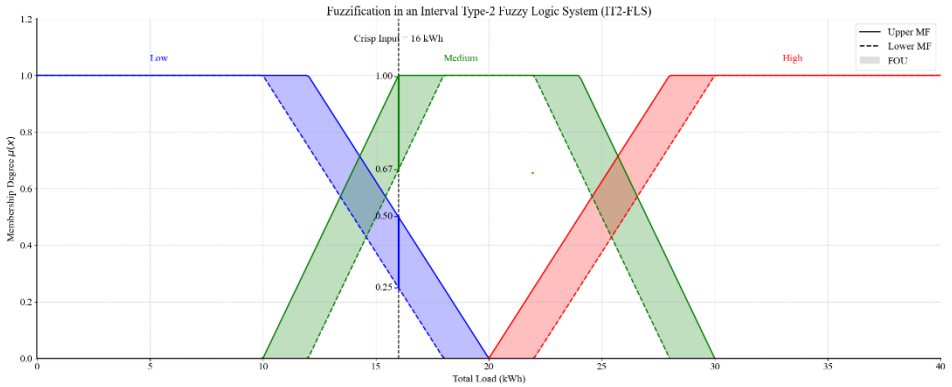

Fig. 6. Fuzzification of a Crisp Load Input Utilizing Interval Type-2 Trapezoidal MFs with a FOU

Linguistic Term	Membership Interval	Average (Lower + Upper)/2
Low	[0.25, 0.40]	0.325
Medium	[0.67, 0.93]	0.80
High	[0, 0]	0

Table 1. Growth Parameters for Scenario Generation

The inference engine subsequently utilizes these interval values to assess the pertinent rules, resulting in interval Type-2 outputs. The width of the type-reduced interval indicates the transmitted uncertainty [11], and the crisp load prediction is achieved through defuzzification and type reduction.

The forecasting of energy demand in systems primarily focused on renewable sources could gain from the IT2FLS's capacity to handle uncertainty within a clear and interpretable rule-based system.

As a result, the subsequent multi-objective optimization phase is guided by the load profiles produced by the IT2FLS-based forecasting module. Subsequently, we will examine the specifics of the system architecture that identifies the optimal renewable generation and energy storage capabilities by combining IT2FLS predictions with the NSGA-III optimization method.

4. The Proposed Type-2 Fuzzy Framework for Multi-Objective Optimization of Smart Grid Applications.

In the context of smart grids, we propose a framework for the stochastic development of hybrid renewable energy systems (HRES). By combining interpretable Interval Type-2 fuzzy load forecasting

with knowledge-driven multi-objective optimization (GA + NSGA-III), the approach balances cost, reliability, and renewable energy utilization, while incorporating multi-year hourly load profiles. Integrating transparent forecasts with evolutionary optimization enables informed decision-making and robust planning for renewable generation and storage (see Fig. 7). The following subsection provides a detailed examination of each framework component and its role in designing efficient HRES for smart grids.

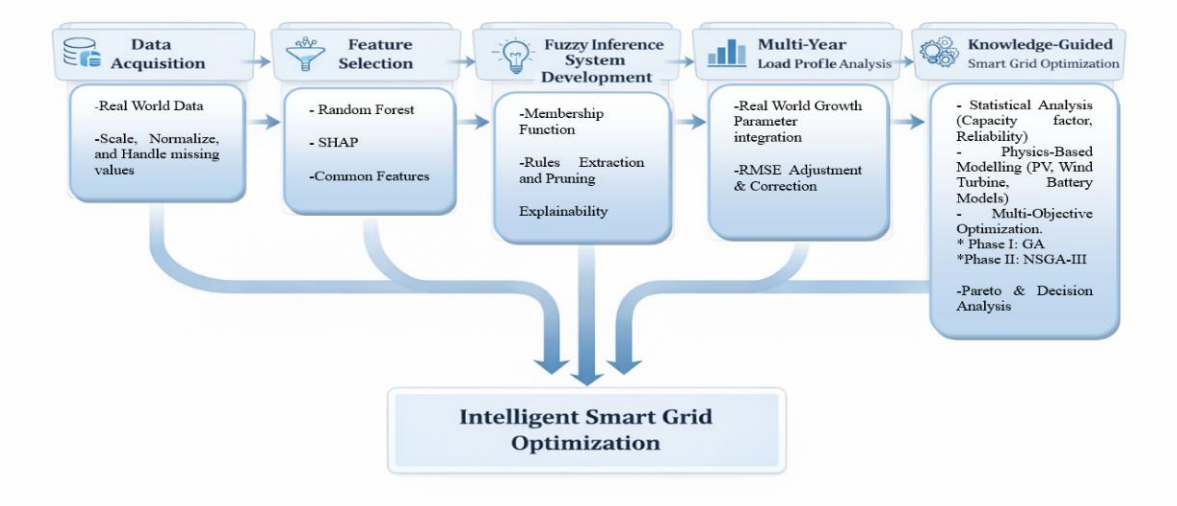

Fig. 7. An Overview of the Proposed System

4.1 Data Preparation and Preprocessing

The dataset is first imported into Python as a Pandas Data-Frame, which helps organize and modify the data. After loading the CSV file, key metrics such as minimum, maximum, total, and target kWh are identified. The timestamp column is converted to a datetime format to ensure precise time series analysis [11]. To analyze temporal usage patterns, additional variables like hour, day of the week, and month are created. Missing values are handled through statistical imputation, and categorical variables are encoded when necessary. The dataset is further enriched by adding important weather variables like temperature, humidity, wind speed, and air pressure to improve model performance. These preprocessing steps result in a well-structured and consistent dataset, ready for feature selection and Interval Type-2 Fuzzy Logic modeling [11].

4.2 Hybrid Importance Ranking for Feature Selection

A hybrid approach to feature selection is applied to decrease model complexity and enhance interpretability and prediction accuracy. The proposed method merges Random Forest feature importance with SHAP (SHapley Additive exPlanations) analysis to effectively identify the key input variables.

1) The Random Forest Regressor, including B decision trees, is used to evaluate the relative importance of each input attribute. The ensemble forecast of the model is obtained by averaging the outputs of the various decision trees [26].

$$\hat{y}_{RF} = \frac{1}{B}\sum_{b=1}^{B} h_b\,(X) \tag{8}$$

where $h_b(X)$ is the output of the b-th tree. The generalization error is constrained by $PE^*_{forest} \le \bar{\rho} \cdot PE^*_{tree}$, with $\bar{\rho}$ denotes the mean correlation among trees [27].

2) SHAP-Based Feature Importance: SHAP values are calculated to measure the contribution of each feature to the prediction. For feature j, the Shapley value ϕ_j is defined as follows [28]:

$$\phi_j(f,x) = \sum_{S \subseteq F\setminus\{j\}} \frac{|\,S\,|\,!\,(|\,F\,| - |\,S\,| - 1)!}{|\,F\,|\,!}\left[f_{S\cup\{j\}}(x_{S\cup\{j\}}) - f_S(x_S)\right] \tag{9}$$

where F signifies the comprehensive feature set and S specifies a subset of features that omits feature j. The global importance of each characteristic is ascertained by computing the average of the absolute SHAP values over all observations in the dataset.

3) Common feature set: Let F_{RF} and F_{SHAP} represent the sets of highest-ranked features nominated by Random Forest and SHAP, respectively. The final feature subset is defined as:

$$F_{common} = F_{RF} \cap F_{SHAP} \tag{10}$$

This consensus-driven approach bolsters robustness by preserving just those traits consistently identified as significant by both methods, thus reducing duplication and enhancing model dependability.

4.3 Interval Type-2 Fuzzy Logic Modelling

To explicitly incorporate uncertainty in load behavior and environmental factors, each input variable is modeled as an Interval Type-2 Fuzzy Set, as defined in (6). The primary membership interval for each input is maintained as [13],[24]:

$$J_x = \left[\underline{\mu}(x), \bar{\mu}(x)\right] \tag{11}$$

representing the range of possible membership grades for the input x.

The Footprint of Uncertainty (FOU) captures the uncertainty between the upper and lower membership functions, and is defined as [13], [24]:

$$\text{FOU}(\tilde{A}) = \bigcup_{x \in X} [\underline{\mu}(x), \bar{\mu}(x)] \tag{12}$$

4.3.1 Membership Function Construction

To construct the interval trapezoidal Type-2 membership functions (MFs) for each input variable x, the universe of discourse is first defined using the minimum and maximum values $[x_{\min}, x_{\max}]$. The range is divided into three linguistic regions: Low, Medium, and High, each indicating a specific load level. Trapezoidal membership functions are selected due to their computational efficiency and ease of interpretation [13].

For each fuzzy set, the trapezoidal breakpoints are labelled as a, b, c, d for the upper MF (a_U, b_U, c_U, d_U) and lower MF (a_L, b_L, c_L, d_L). The primary division points are calculated from the input range as:

$$c = \frac{a + e}{2} \tag{13}$$

$$b = \frac{a + c}{2} \tag{14}$$

$$d = \frac{c + e}{2} \tag{15}$$

where $a = x_{\min}$and $e = x_{\max}$. These points define the upper membership value primary points for the Medium fuzzy set and the Low and High fuzzy sets upper membership values points are formulated accordingly as shown in Fig.6.

A Footprint of Uncertainty (FoU) is introduced to adjust the lower MF boundaries (typically 5% of the range). The lower membership values for the Medium fuzzy parameters (the Low and High fuzzy sets low membership values points are formulated accordingly as shown in Fig.6) are then computed as follows:

$$\text{Lower Medium:} \quad a_L = a + \text{FoU}, \quad b_L = b + FOU, c_L = c - FOU, d_L = d - \text{FoU} \tag{16}$$

For each linguistic term, trapezoidal membership functions are defined as [13],[24]:

$$\mu_A(x) = \begin{cases} 0, & x \le a \text{ or } x \ge d \\ \dfrac{x - a}{b - a}, & a < x \le b \\ 1, & b < x \le c \\ \dfrac{d - x}{d - c}, & c < x < d \end{cases} \tag{17}$$

Through this construction, the upper membership function (UMF) $\bar{\mu}_{\tilde{A}}(x) = (a_U, b_U, c_U, d_U)$ define the outer boundary, while the lower membership function (LMF) $\underline{\mu}_{\tilde{A}}(x)$is parameterized as

(a_L, b_L, c_L, d_L). The structure is entirely nested, creating a clearly defined Footprint of Uncertainty (FoU) for each fuzzy set. This facilitates seamless transitions among Low, Medium, and High sets, effectively accounting for uncertainty in the membership values.

4.3.2 Rule Extraction and Pruning

The fuzzy inference system functions through interval Type-2 IF–THEN rules, as formalized in (7). Each rule associates input linguistic terms, represented as IT2FSs, with their corresponding output IT2FSs. The firing strength of each rule is calculated by aggregating the membership degrees of all antecedent inputs through a t-norm operator, commonly the minimum or product [13],[24],[25].

For a dataset containing Ninput-output pairs $(x^{(t)}; y^{(t)})$, the membership of the input $x_s(t)$ in an interval type-2 fuzzy set (IT2 FS) $\tilde{A}_s^q$ is represented as an interval [13],[24]:

$$\mu_c^{cg}\left(x_s^{(t)}\right) = \frac{1}{2}\left(\underline{\mu}_{\tilde{A}_s^q}\left(x_s^{(t)}\right) + \overline{\mu}_{\tilde{A}_s^q}\left(x_s^{(t)}\right) \right) \tag{18}$$

which delineates its Footprint of Uncertainty (FOU). To simplify rule development while preserving uncertainty information, this interval is often defuzzified to a single center-of-gravity (CoG) membership value, providing a precise and actionable measure for subsequent inference.

Each input is characterized by a type-2 fuzzy set, enabling a potentially extensive array of rules. Nevertheless, only rules pertaining to areas with real data points are produced, which limits the applicability of the fuzzy set approach to only those regions where data is available. A rule is established for each input-output pair by choosing the fuzzy set with the greatest membership value [13], [24]:

$$wi^{(t)} = \prod_{s=1}^{n} \mu_{\tilde{A}_s^q}^{cg}\left(x_s(t)\right) \tag{19}$$

This process produces an initial collection of N candidate rules, each associated with a particular data point.

To enhance interpretability and mitigate overfitting, rules with minimal contribution are eliminated. The normalized firing strength of rule r for sample n is expressed as follows [29]:

$$\bar{f}_r\,(x) = \frac{f_r(x)}{\sum_{k=1}^{R} f_k(x)}, \quad \mathrm{r} = 1, \ldots, \mathrm{R}, \mathrm{n} = 1, \ldots, \mathrm{N} \tag{20}$$

The average normalized firing strength across all samples is defined as [29]:

$$\bar{f}_r\,(x) = \frac{1}{N}\sum_{n=1}^{N} \bar{f}_r\,(x_n) \tag{21}$$

This quantifies the rule's overall significance. Rules with $\bar{f}_r < \gamma \cdot median\{\bar{f}_k\}_{k=1}^{R}$ are removed. The threshold γ is generally established within the range of 0.05 to 0.15, achieving a balance between rule base compactness and fidelity [29], [30].

Let the rule base consist of R rules. A similarity matrix $S \in \mathbb{R}^{R\times R}$ It is constructed to quantify the similarity between rules. The similarity between rules $i$and $j$is calculated using their normalized firing levels as [29]:

$$s(i,j) = \frac{\min\,(\bar{f}_i, \bar{f}_j)}{\max\,(\bar{f}_i, \bar{f}_j)}, i \neq j \tag{22}$$

A higher ratio indicates a greater degree of overlap between the two rules. Let (c_i, c_j)denote the consequents of rules $i$and j, and (w_i, w_j)represent their corresponding weights [29]. Similar rules are merged using a weighted averaging approach. If $\max\,(S) > \theta$, the consequent and weight of rule $i$are updated as [29],[30]:

$$c_i \leftarrow \frac{w_i c_i + w_j c_j}{w_i + w_j}, w_i \leftarrow w_i + w_j \tag{23}$$

Rule $j$is then removed from the rule base, the similarity matrix $S$is updated accordingly, and the number of rules is reduced as $R \leftarrow R - 1$.

4.3.3 Inference

A zero-order Takagi–Sugeno–Kang (TSK) inference mechanism is adopted [29]. For a given input x, the final predicted load is calculated as a weighted average of rule outputs [30]:

$$y_r = \frac{\sum_{r=1}^{R} \bar{f}_r\ c_r}{\sum_{r=1}^{R} \bar{f}_r}, \qquad y_l = \frac{\sum_{r=1}^{R} \underline{f_r}\ c_r}{\sum_{r=1}^{R} \underline{f_r}} \tag{24}$$

where R denotes the number of active rules. The final crisp prediction is obtained as the midpoint of the resulting interval [29]:

$$\hat{y}(x) = \frac{y_l + y_r}{2} \tag{25}$$

4.3.4 Explainability

Explainability improves transparency by helping operators understand how predictions are made. It aids in root cause analysis, offers actionable insights, and facilitates timely decisions [11]. The core of its firing interval indicates when a merged rule is activated. The firing strength is calculated as:

$$f_c = \frac{f_l + f_r}{2} \tag{26}$$

Each kWh prediction is derived from a set of active rules R. The sum of the rules' firing strengths contributing to the given prediction can be written as follows:

$$S = \sum_{r \in R} f_c \tag{27}$$

The contribution of each rule r is determined by normalizing its firing strength with respect to S, weighted by its centroid C_r:

$$w = \frac{f_c \cdot C_r}{S} \tag{28}$$

Rules are associated with labels corresponding to features or feature values [11]. The contribution of the label l within the feature column c is computed as:

$$\text{Contribution}_{(c,l)} = \sum_{\substack{r \in R \\ \text{label of } r = l}} w \tag{29}$$

For interpretability, contributions are normalized by scaling the maximum value to 100%:

$$\text{Contribution\%} = \frac{v_c}{v_{\max}} \times 100 \tag{30}$$

where $v_c$is the feature contribution and $v_{\max}$is the maximum contribution across all features.

4.4 Multi-Year Load Construction

After the development of the Interval Type-2 Fuzzy Logic System (IT2FLS), a systematic two-step extension method is employed to construct multi-year load profiles. The IT2FLS baseline output is modified by incorporating:

1) Hour-specific RMSE values: derived from historical performance data. This maintains the integrity of peak demand and daily variations, preventing smoothing effects that could underestimate true storage needs.

2) Long-term growth integration involves modeling demand fluctuations over several years through annual compound growth factors, which consider demographic, economic, and electrification trends. The data is sourced from reputable international entities to ensure the accuracy of the projections in relation to the scenario and parameters articulated in Table 2.

Parameter	Annual Rate	Source	Uncertainty Range
Population Growth	+1.2%	UN World Population Prospects 2022 [31]	±0.3%
EV Adoption	+15%	IEA Global EV Outlook 2023 [32]	±5%
Heat Pump Growth	+8%	National Energy Agencies	±2%
Industry Growth	+2.5%	World Bank GDP Projections [33]	±1%
Electrification	+3%	IEA Electrification Reports	±1%
Efficiency Gain	-1%	National Efficiency Standards	±0.5%
Urbanization Boost	+0.5%	UN Urbanization Statistics	±0.2%

Table 2. Growth Parameters for Scenario Generation

This method generates load trajectories characterized by hourly resolution, structural consistency, and realistic long-term growth, rendering them appropriate for physics-based modeling and multi-objective optimization.

4.5 Knowledge-Guided Smart Grid Optimization

In this paper, we present an evolutionary multi-objective optimization framework that combines physics-based modeling with multi-year statistical forecasts to produce reliable, economical, and physically consistent system designs. Fig. 8 illustrates an architecture composed of five interconnected layers.

The validated technology catalogs are incorporated into the data layer, alongside long-term forecasts of electricity consumption and the availability of renewable resources. The statistics layer performs quantitative assessments of system capacity and reliability metrics across diverse operational scenarios. The physics layer represents the physical behavior and interactions among the components of a hybrid renewable energy system. The text delineates the mechanisms through

which photovoltaic panels generate electricity, wind turbines convert kinetic energy from wind into electrical power, and battery systems store and supply energy. The optimization layer employs the GA–NSGA-III algorithm to determine optimal trade-offs among multiple design objectives. The analysis layer evaluates the identified Pareto-optimal solutions to assist decision-makers in selecting the most appropriate system configuration while complying with all operational and technical requirements. The following sections present an overview of each layer.

Fig. 8. Hybrid Renewable Energy System Design Code Framework

4.5.1 Data Layer

The Data layer provides essential inputs for system modeling and optimization. This dataset comprises classified technology inventories, information on renewable energy resources, and projections of load demand profiles. The technical catalogs provide detailed specifications for key

system components, including photovoltaic (PV) modules, wind turbines, and energy storage systems. The specifications encompass module capacity, conversion efficiency, temperature coefficients, turbine power ratings, hub heights, and manufacturer-provided performance curves.

The information is derived from reputable producers and validated industry records to ensure accuracy and practical applicability. The Data layer provides the foundational framework necessary for subsequent statistical analysis, physics-based modeling, and optimization techniques by delivering reliable technical and operational parameters derived from precise representations of hybrid renewable energy system components.

4.5.2 Statistics Layer

The variability of multi-year renewable resources significantly impacts energy storage needs and system reliability. Analyzing wind and solar data over multiple years reveals that low-generation periods last longer than single-year studies indicate, sometimes spanning several days or entire seasons. This results in higher storage requirements [34]. Long-duration storage, beyond daily (diurnal) timescales, is crucial for maintaining reliability in systems with high renewable energy penetration [35]. Historical data show many hours annually when combined wind and solar output cannot meet demand without sufficient storage [35], [36]. We define the following metrics for reliability driven by absence to quantify these effects explicitly:

- Joint Absence Hours (JAH) refer to the duration during which wind speed is outside acceptable limits, indicating the concurrent unavailability of both resources [36].
- Longest Energy Gap (LEG): The maximum number of consecutive hours of joint absence [36].
- Worst-Gap Energy (WGE) refers to the total load energy consumed during the longest gap period.
- Battery sizing requires a minimum storage capacity sufficient to address 95% of absence gaps, taking into account both the worst-case energy requirements and the longest duration of gaps [34], [36].

This multi-year strategy effectively captures extreme renewable scarcity events, offering reliable guidance for storage sizing and system design.

4.5.3 Physics Layer

- *Photovoltaic (PV) Generation Model:* The DC output power of the photovoltaic generator is estimated using the method proposed by Osterwald, which scales the rated peak power according to incident irradiance and module temperature. This approach uses irradiance and

operating temperature as inputs and applies a linear temperature correction to account for deviations from standard test conditions (STC) [37]:

$$P_m = P_{m,\text{ref}} \frac{G}{G_{\text{ref}}} \left[1 + \gamma \left(T_m - T_{m,\text{ref}}\right)\right] \tag{31}$$

where $P_m$is the instantaneous DC output power of the module, $P_{m,\text{ref}}$ is the rated peak power at STC,$G$is the global irradiance on the PV plane, $G_{\text{ref}} = 1000/m^2$ is the reference irradiance at STC, $T_m$is the module temperature, $T_{m,\text{ref}} = 25°\text{C}$is the module temperature at STC, and $\gamma$is the temperature coefficient of power (typically expressed in %/°C). The total DC output of the PV array is obtained by aggregating the contributions of all modules [37], [38]:

$$P_{\text{dc}} = N_{pv} \cdot P_m \tag{32}$$

where $N_{pv}$is the number of PV modules in the array. The AC power delivered to the grid is then derived by applying the inverter efficiency [38]:

$$P_{ac} = \eta_{inv} \cdot P_{\text{dc}} \tag{33}$$

where η_{inv}denotes the inverter efficiency, $T$is the operating temperature, and $T_{stc}$is the temperature at standard test conditions.

- *Wind Generation Model:* The power output of a wind turbine is modeled using a standard **piecewise power curve**, consistent with conventional wind turbine aerodynamic theory [39], [40]. The turbine produces no power below the **cut-in speed** $v_{\text{cut-in}}$and above the **cut-out speed** $v_{\text{cut-out}}$, increases cubically between cut-in and rated speed v_{rated}, and maintains rated power between rated and cut-out speeds [39], [41]:

$$P_w(v) = \begin{cases} 0, & v < v_{\text{cut-in}} \text{ or } v > v_{\text{cut-out}} \\ P_{\text{rated}} \left(\dfrac{v - v_{\text{cut-in}}}{v_{\text{rated}} - v_{\text{cut-in}}} \right)^3, & v_{\text{cut-in}} \le v < v_{\text{rated}} \\ P_{\text{rated}}, & v_{\text{rated}} \le v \le v_{\text{cut-out}} \end{cases} \tag{34}$$

where:

- $v$is the wind speed at hub height,
- $v_{\text{cut-in}} = 3$ m/s, $v_{\text{rated}} = 12$ m/s, $v_{\text{cut-out}} = 25$ m/s,
- $P_{\text{rated}}$is the turbine rated power.

The expected energy output of a wind turbine at a site can be expressed using the capacity factor (CF) formulation [41]:

$$\mathrm{CF} = \frac{P_{\mathrm{WTe}}}{P_r} = \frac{1}{P_r}\int_0^{\infty} P_w\,(v) f(v)\,dv \tag{35}$$

where $f(v)$is the probability density function of wind speeds at the site (commonly Weibull), and $P_r$is the turbine rated power.

- *Capacity Factor Constraints:* To ensure physically realistic system configurations, capacity factor (CF) constraints are imposed:
  - *Wind turbines:* $0.15 \leq \mathrm{CF} \leq 0.45$, reflecting local wind regimes, turbine characteristics, and operational limitations [39], [41], [42].
  - *Photovoltaic (PV) systems:* $0.10 \leq \mathrm{CF} \leq 0.25$, accounting for site irradiation, maintenance practices, system degradation, and technology limitations [37].

  These bounds ensure that modeled energy yields remain consistent with real-world system performance.

- *Microgrid-Scale Turbine Scaling*

  To maintain compatibility with microgrid-scale applications, turbines with rated power exceeding the system limit (100 kW) are scaled linearly while preserving relative capacities [37], [38]:

$$\text{scale factor} = \frac{100}{\max\,(P_{\text{rated}})}, P_{\text{rated,scaled}} = P_{\text{rated}} \cdot \text{scale\_factor} \tag{36}$$

  This preserves turbine characteristics while ensuring feasibility in small-scale systems.

4.5.4 Optimization Layer: Framework for Identifying Optimal Energy System Configurations.

Designing hybrid renewable energy systems (HRES) entails interrelated decisions, including the determination of the number of photovoltaic panels (*N_PV*), wind turbines (*N_WT*), battery capacity (*C_bat*), and relevant power electronics. These decisions simultaneously influence economic costs, reliability, and sustainability [22], [23]. This paper presents a two-phase evolutionary optimization framework (see Figure. 9) to effectively explore the design space while maintaining operational feasibility. Initially, a Genetic Algorithm (GA) produces a varied collection of feasible configurations. Subsequently, the Non-Dominated Sorting Genetic Algorithm III (NSGA-III) is employed to determine Pareto-optimal solutions across multiple objectives.

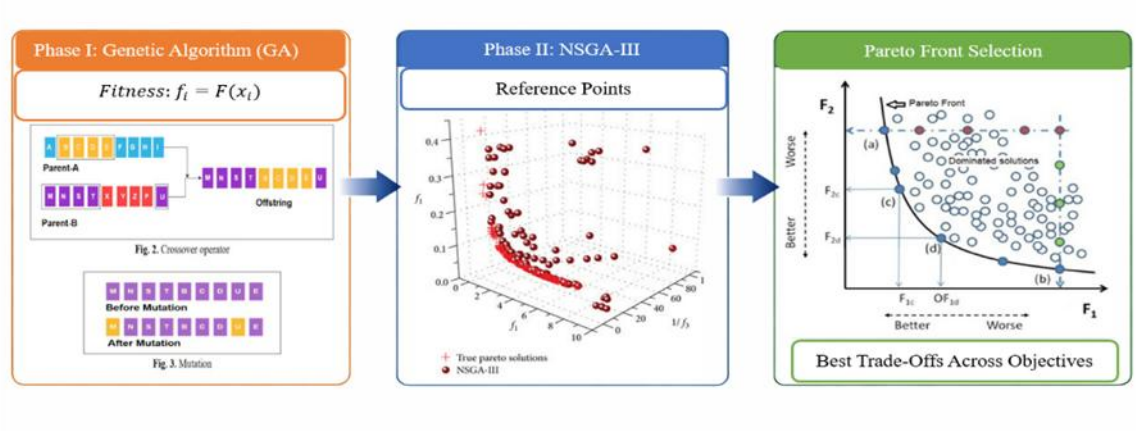

Fig. 9. Two-phase evolutionary optimization framework for the smart grid (SG) system, that combines GA and NSGA-III.

- ***Phase I: Initialization of the Genetic Algorithm (GA)***

The GA is employed to explore the feasible solution space efficiently. Each candidate solution, or chromosome, encodes a hybrid system configuration. The population evolves through selection, crossover, and mutation, allowing a balance between exploitation of good solutions and exploration of new possibilities [21]. Tournament selection reduces the risk of premature convergence and ensures robust exploration of high-quality configurations [43]

$$x_{parent} = arg\,\min_{x_j \in T} \; F(x_j) \tag{37}$$

where $T \subseteq$ population is a randomly chosen subset and $F(x)$is the fitness function. This method ensures that fitter solutions have a higher probability of producing offspring while maintaining diversity.

Simulated Binary Crossover (SBX) is commonly employed in continuous-variable problems due to its ability to maintain parental characteristics while facilitating the exploration of intermediate solutions [44]. SBX generates offspring from parent pairs (p_1, p_2) [44]:

$$\begin{aligned} c_{1,i} &= 0.5\left[(p_{1,i} + p_{2,i}) - \beta(p_{2,i} - p_{1,i})\right] \\ c_{2,i} &= 0.5\left[(p_{1,i} + p_{2,i}) + \beta(p_{2,i} - p_{1,i})\right] \end{aligned} \tag{38}$$

where βfollows a probability distribution controlled by the distribution index η_c. For example, If Parent 1 has a PV capacity of 20 kW and Parent 2 has a capacity of 30 kW, the resultant generation could show a PV capacity of 24 kW, reflecting features of both parent entities.

Polynomial Mutation perturbs individual genes $x_i$with probability p_m [19]:

$$x_i' = \begin{cases} x_i + \delta(x_i^{\max} - x_i^{\min}), & \text{with probability } p_m \\ x_i, & \text{otherwise} \end{cases} \quad (39)$$

where $\delta$is sampled from a polynomial distribution. This operator introduces controlled variability to prevent the population from stagnating, thereby ensuring the search explores less obvious but potentially optimal regions [19]. For instance, a battery capacity of 100 kWh may increase to 120 kWh, enabling the consideration of larger storage alternatives.

- ***Phase II: NSGA-III Multi-Objective Optimization***

In Phase II, NSGA-III is used to optimize multiple competing objectives simultaneously [23]. The multi-objective problem is formulated as in (1) [23]:

For Clarification:

- x: vector of decision variables [N_{PV}, N_{WT}, C_{bat}, etc.]
- f = Each f denotes an objective within the optimization problem. For instance, $f_1$may represent the Levelized Cost of Energy (LCOE), f_2 may denote the Expected Energy Not Supplied (EENS), f_3 may indicate the negative Renewable Fraction (−RF), and f_4 may refer to the Curtailment.

Subject to:

$$g_j(x) \leq 0, h_k(x) = 0, x_i^{\min} \leq x_i \leq x_i^{\max} \quad (40)$$

Let $g_j(x)$ represent inequality constraints such as battery state of charge, $h_k(x)$denote equality constraints (example: load), *and* x_i^{min}, x_i^{max}*: min/max values for design variables.* This formulation integrates cost, reliability, and sustainability within operational constraints, offering a systematic framework for decision-making.

Objective Functions

The objectives encompass essential aspects of system performance, including operational efficiency, environmental sustainability, energy production costs, and supply reliability. In developing a system to address actual energy needs, it is critical to consider all relevant factors. Enhancing one objective often degrades another; therefore, the goals are incompatible. Increasing renewable energy may raise system costs, while lowering costs may compromise system reliability, which is essential for energy stability. Thus, no single solution can optimize all objectives.

Pareto-based multi-objective optimization addresses this issue by evaluating each objective simultaneously rather than in isolation. This approach yields multiple Pareto-optimal solutions, each illustrating a unique trade-off among cost, dependability, and renewable energy utilization.

These solutions provide flexibility in selecting the optimal system configuration, informed by design priorities such as cost efficiency, reliability, and sustainability goals.

1. **Levelized Cost of Energy (LCOE)** captures the total cost per unit of delivered energy, essential for cost-effectiveness evaluation [45]:

$$\text{LCOE} = \frac{\sum_{t=1}^{T} \frac{I_t + M_t + F_t}{(1+r)^t}}{\sum_{t=1}^{T} \frac{E_t}{(1+r)^t}} \tag{41}$$

Here, I_t denotes the investment cost in the year t, M_t represents the operation and maintenance cost in a year t, and $F_t$is the fuel cost for that year. The energy delivered in a year t is given by E_t. The parameter $r$is the discount rate, and Trepresents the system lifetime. The objective is to minimize the LCOE, which measures how cost-effective a system is. For example, installing a battery increases the original outlay. Over the system's lifetime, it may reduce the total LCOE while also reducing energy waste. The LCOE is determined by the amount invested and the quantity of energy produced. These modifications allow the system to better leverage renewable energy sources. This modification might lead to a reduced LCOE. The data demonstrates that the system's energy use is efficient while being fairly priced. The LCOE varies depending on system design and usage. LCOE is an important indicator since it indicates how much a system really costs.

2. **Expected Energy Not Supplied (EENS)** quantifies reliability by measuring unserved demand, highlighting trade-offs between cost and system dependability [46].

$$\text{EENS} = \sum_{t=1}^{T} max\ (0, P_{\text{load}}(t) - P_{\text{supplied}}(x,t)) \tag{42}$$

P_{load} represents the power demand at time t, while $P_{\text{supplied}}(x,t)$ denotes the energy delivered by a hybrid renewable energy system (HRES) with configuration x at that same moment. $JAH(t)$indicates the hours when both solar and wind resources are simultaneously unavailable. The Expected Energy Not Supplied (EENS) measures the total energy demand that is unmet over a specified period. The objective is to minimize EENS, as lower values indicate a more reliable system. In periods of minimal or no wind and solar generation, the system relies on stored energy. A failure of the battery to meet demand leads to a supply shortfall, which increases Expected Energy Not Supplied (EENS) and reduces system reliability. From a systemic perspective, EENS is notably influenced by the temporal mismatch between energy production and consumption. During periods of reduced renewable availability, as indicated

by $JAH(t)$ insufficient storage capacity or an ineffective dispatch strategy result in heightened unmet demand. Reducing Expected Energy Not Supplied (EENS) typically requires increasing battery capacity or enlarging generating facilities. However, this may conflict with economic objectives, such as reducing LCOE, leading to a trade-off between reliability and cost.

3. **Renewable Fraction (RF)** indicates the proportion of energy met by renewables. Maximizing RF supports sustainability goals while minimizing fossil fuel dependency [47].

$$\mathrm{RF} = \frac{\sum_{t=1}^{T} P_{\mathrm{RES}}\,(x,t)}{\sum_{t=1}^{T} P_{\mathrm{load}}\,(t)} \tag{43}$$

$P_{\mathrm{RES}}(x,t)$ represents the renewable energy supplied by the system at time t for a specific configuration x, while $P_{\mathrm{load}}(t)$denotes the total energy demand at that time. The Renewable Fraction (RF) denotes the percentage of overall energy demand satisfied by renewable energy sources. If wind and solar power fulfill 75% of the total energy demand, the RF is 75%. The aim is to maximize RF, as elevated values signify increased dependence on renewable energy sources. From a system design perspective, enhancing RF necessitates either greater renewable generation capacity or more efficient utilization of energy storage. Achieving RF values exceeding 80% generally necessitates considerable photovoltaic and wind capacity, in addition to adequate battery storage to address generation variability.

Increasing RF typically results in elevated system costs and greater energy curtailment. This illustrates a triadic trade-off among RF, LCOE, and curtailment, indicating that enhancements in one objective may adversely impact the others.

4. **Curtailment** measures potential energy losses due to storage or generation limits, helping assess how efficiently renewable resources are utilized [48].

$$Curt = \sum_{t=1}^{T} max\,(0, P_{RES}(x,t) - P_{load}(t) - P_{bat}(x,t)) \tag{44}$$

for a certain system configuration *x*, the amount of energy that the battery has absorbed or stored at time *t*. Curtailment denotes the quantity of renewable energy that is unusable, unstageable, or undeliverable in relation to demand. The aim is to reduce curtailment, as elevated levels of curtailment signify inefficient utilization of available renewable resources.

For instance, during a sunny weekend with low electricity demand, a solar array can produce 10 kWh of energy. If this energy cannot be utilized immediately or stored in the battery, it is deemed curtailed. Curtailment exhibits significant sensitivity to system sizing. Excessive photovoltaic capacity without adequate storage results in surplus energy during peak generation, thereby heightening curtailment rates. Enhancing battery capacity may mitigate

curtailment through the storage of excess energy; however, it concurrently elevates system costs.

Minimizing curtailment necessitates a careful equilibrium among generation capacity, storage size, and demand profile.

Constraints in Renewable Energy Systems and Their Handling in NSGA-III

When designing Hybrid Renewable Energy Systems (HRES), the proposed solutions must be operationally and physically feasible while also adhering to policy and sustainability requirements. To ensure feasibility, various constraints are established and integrated into the NSGA-III algorithm via the concept of total constraint violation $CV(x)$.

1. Reliability Constraint – Restricting Expected Energy Not Supplied (EENS)

The system must provide electricity consistently, and the Expected Energy Not Supplied (EENS) measures the energy deficit when demand surpasses supply [46]. The reliability constraint is articulated as follows:

$$g_1(x) = \frac{EENS(x)}{\sum_{t=1}^{T} P_{load}(t)} - \epsilon \leq 0 \tag{45}$$

where $P_{\text{load}}(t)$ represents the electricity demand at time t, $P_{\text{supplied}}(x, t)$ denotes the total energy supplied by the system for the configuration x, and ϵ indicates the maximum allowable fraction of unmet demand, which is typically set between (5–10%). This constraint guarantees that the system achieves an acceptable level of reliability; any solution that exceeds ϵ is deemed infeasible. Battery and Storage Adequacy Constraint

2. Battery / Storage Adequacy Constraint

To mitigate the intermittency of renewable energy sources, the system requires adequate battery capacity to accommodate the longest continuous energy deficit [49].

$$g_2(x) = \max_t \sum_{\tau=0}^{G_{longest}} (P_{load}(t+\tau) - P_{supplied}(x, t+\tau)) \cdot \Delta t - C_{battery} \leq 0 \tag{46}$$

where G_{longest} represent the duration of the longest deficit period, $C_{\text{battery}}$is the usable battery capacity, and $\Delta t$is the time step. This constraint eliminates infeasible designs in which the battery fails to support the load during extended periods of renewable energy shortages.

3. **Renewable Fraction Constraint**

A minimum renewable energy share is mandated to promote sustainability and decrease reliance on fossil fuels [49].

$$g_3(x) = f_{\min} - RF(x) \leq 0 \tag{47}$$

In this context $f_{\min}$ represents the minimum required renewable fraction, which ranges from 50–80%. This constraint guarantees adherence to environmental and policy objectives.

4. **Total Constraint Violation in NSGA-III**

All constraints are aggregated into a single measure of constraint violation [19], [23]:

$$CV(x) = \sum_{j=1}^{3} \max(0, g_j(x)) \tag{48}$$

If $CV(x) = 0$, the solution is feasible; otherwise, it is infeasible. NSGA-III handles' constraints using **constraint dominance**:

- Feasible solutions dominate infeasible ones.
- Among infeasible solutions, smaller $CV(x)$is preferred.

Pareto Dominance: Solution $x_a$dominates $x_b$if:

$$x_a \prec x_b \iff f_m(x_a) \leq f_m(x_b), \forall m \text{ and } \exists m: f_m(x_a) < f_m(x_b) \tag{49}$$

Pareto dominance ensures that no objective can be improved without worsening another, guiding the selection of balanced HRES configurations.

5. **NSGA-III Workflow with Constraint Handling**

The integration of constraints within NSGA-III is illustrated in Fig. 10, providing a clear and intuitive understanding of the overall workflow. The figure illustrates the sequential process of objective evaluation and constraint handling using $g_1(x)$, $g_2(x)$, and $g_3(x)$, computation of $CV(x)$, and the subsequent ranking and diversity preservation mechanisms. Each candidate HRES configuration is first evaluated in terms of objective functions, followed by constraint evaluation using Equations (45)– (47). The total constraint violation $CV(x)$ (Equation (48)) is then used to classify solutions as feasible or infeasible. Feasible solutions are ranked using Pareto dominance (Equation (49)), while infeasible solutions are penalized according to their degree of violation.

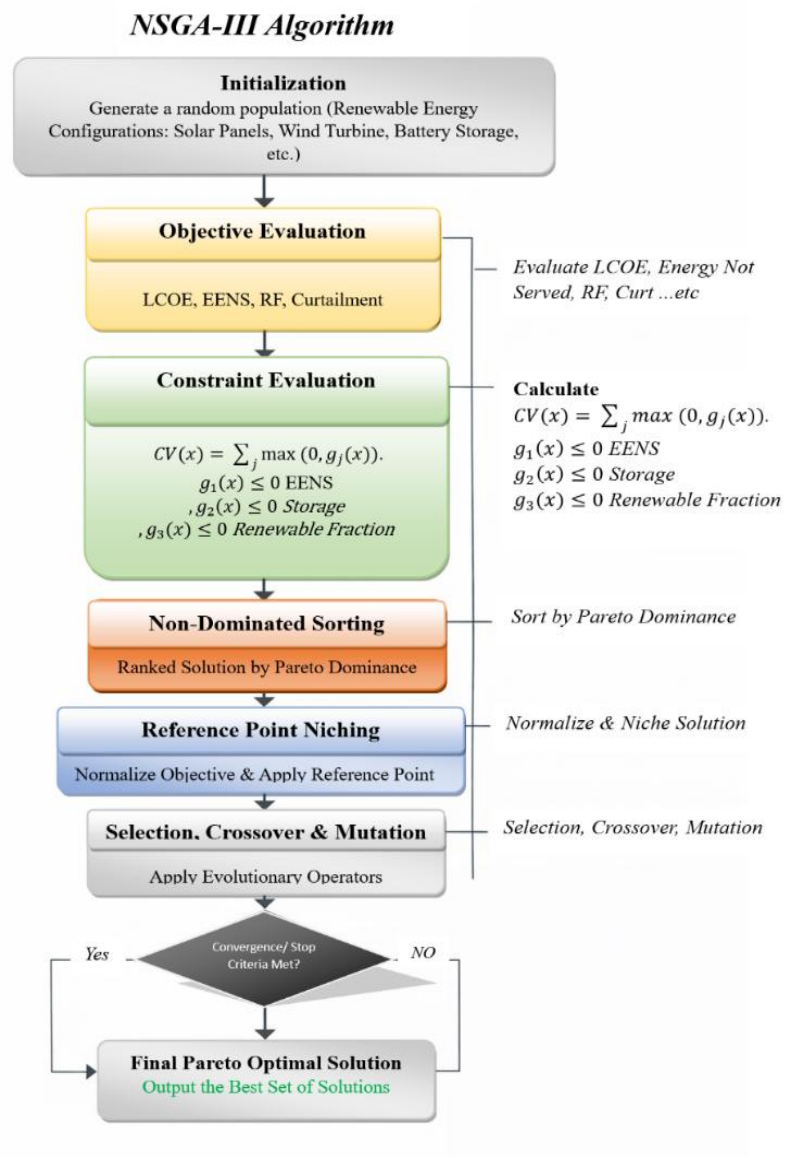

Fig. 10. NSGA-III-based optimization workflow for HRES planning, including objective and constraint evaluation, non-dominated sorting, reference-point niching, genetic operations, and final Pareto-optimal solution selection.

Normalization and Reference Directions maintain diversity and evenly distribute solutions along the Pareto front, enhancing interpretability [23]:

$$\hat{f}_i(x) = \frac{f_i(x) - z_i^{\text{ideal}}}{z_i^{\text{nadir}} - z_i^{\text{ideal}}} \tag{50}$$

Where z_i^{ideal}, z_i^{nadir} represents the best/worst observed values for the objective i.

$$d_\perp(x, r_h) = \| \hat{F}(x) - \frac{\hat{F}(x) \cdot r_h}{\| r_h \|^2} r_h \| \tag{51}$$

r_h is an established reference vector in the objective space, and $d_\perp$indicates the perpendicular distance of a solution from its reference line. In NSGA-III, this measure helps preserve diversity

among solutions. For instance, it confirms that the Pareto front includes both a low-LCOE/high-curtailment configuration and a high-RF/low-curtailment configuration, maintaining a variety of trade-offs.

4.5.5 Analysis Layer

The final layer analyzes the non-dominated solution set derived from NSGA-III within the objective space. After optimization, visualization tools like parallel coordinate plots or radar charts are used to analyze and interpret the trade-offs within the non-dominated set. This layer allows stakeholders to choose configurations that align with their priorities, balancing cost, reliability, and the use of renewable energy. In contexts with constrained renewable resources, resilience assessments analyze the system's capacity to endure unfavorable conditions over prolonged durations. This analysis does not arbitrarily identify a single optimal design; instead, it emphasizes Pareto efficiency, technological feasibility, environmental sustainability, and operational robustness. The final validation indicates that the proposed hybrid configurations are cost-effective, technically robust, and capable of sustained operation over several years [22, 23].

5. Experiments and Results

We have conducted various experiments based over one year for collected 8,761 hourly energy usage samples from both residential and business units [50]. The input vector consists of many features included time, weekday, month, environmental parameters (temperature, dew point, humidity, wind speed, wind gust, pressure, and precipitation), operational indicators (such as facility load, gas, heating, cooling, HVAC components, interior and exterior lights, water heaters, and interior equipment), and historical consumption metrics (total kWh, max_consumption, and average_consumption). Target kWh predicts electricity use over the next hour. A two-pass cross-validation approach was used: in each pass, 70% of the data was used for training and 30% for testing. All reported results (Table 4) correspond to the testing sets, ensuring unbiased evaluation.

To reduce dimensionality and improve interpretability, feature selection was conducted using Random Forest (RF) importance and SHAP analyses. The results, summarized in Table 3, show the number of top features identified by each method, the overlap (common features), and the resulting IT2 FLS rule complexity. The most influential features consistently identified were InteriorEquipment_kW, Total_kW_hr, and Gas_Facility_kW, which were used in subsequent modeling. Three linguistic fuzzy sets (Low, Medium, and High) were employed to represent each input variable in the Type-1 and IT2 fuzzy logic systems. The resulting fuzzy rule bases produced 16, 370, and 616 rules for 3, 6, and 8 common selected features, respectively.

Top N Features	Top N Features (SHAP)	Top N Features (RF)	Top N Features (Common)	Number of Rules
5	5	5	3	16
8	8	8	6	370
10	10	10	8	616

Table 3. Feature Selection and IT2 FLS Rule Complexity

The selected features were then used to evaluate three prediction models: Type-1 FLS (T1FLS), IT2 FLS, and Neural Network (NN). To provide an impartial comparison, all three forecasting models were trained and evaluated using the identical input features: InteriorEquipment_kW, Total_kW_hr, and Gas_Facility_kW. The neural network was built as a feedforward multilayer perceptron while tuning the hyper-parameters using grid search to achieve the best performance over training data. The tuned hyper-parameters included two hidden layers containing 64 and 32 neurons, respectively (tuned to achieve best performance). Both hidden layers employed ReLU activation, whereas the output layer used a linear activation function for hourly load predictions. To reduce overfitting, a dropout rate of 0.2 was used after the first hidden layer. The model uses the Adam optimizer, a mean squared error loss function, a batch size of 32, and a maximum training time of 50 epochs. Early stopping was applied by monitoring validation loss with a patience value of 10, and the best weights were restored. The input and target variables were normalized before training. Table 4 presents the performance metrics on the testing sets, showing that IT2 FLS outperformed T1FLS and NN in terms of generalization. NN provided competitive accuracy but exhibited increased fold variability. The top 3 features proved to be the most important predictors in hourly demand estimation, enhancing interpretability and decision relevance. With only 16 rules, the IT2 FLS achieved a testing RMSE of 2.717, representing a reduction in rule complexity of 250 times compared to the 2,781 rules required in [11].

Model	MAE	RMSE	R^2
Type-1 FLS	4.599	5.292	0.625
IT2 FLS	2.288	2.717	0.901
Neural Network	1.226	1.639	0.964

Table 4. Model Performance

To demonstrate the explainability of the IT2 FLS model, we show below examples of the extracted rules showing comprehensible correlations between operational factors and anticipated demand:

- *Rule 1: IF Water_Heater_Gas__kW_ is **Medium** AND Gas_Facility__kW_ is **Low** AND Electricity:Facility is **High** THEN Target_kW_hr is **Low***
- *Rule 2: IF Water_Heater_Gas is **Low** AND Gas_Facility is **Medium** AND Time is **High**, THEN Target_kW_hr is **Medium**.*
- *Rule 3: IF InteriorEquipment__kW_ is **Low** AND Gas_Facility__kW_ is **High** AND InteriorLights__kW_ is **Medium**, THEN Target_kW_hr_ is **High.***

The rules show that subsystem linkages significantly influence hourly load forecasts. Rule-based energy forecasting helps operators grasp how operational variables affect load predictions. Rule 1 suggests that low gas use in water heaters and minimal gas facility demand can keep energy goals low even when electrical facility use is high. Rule 2 indicates that moderate gas facility operation coupled with low water heater gas use leads to medium peak energy demand. Rule 3 explains that energy use might rise despite limited interior equipment, substantial gas facility activity, and moderate lighting. These rules highlight that gas-related factors have a strong impact on building energy behavior. The IT2 fuzzy inference system enhances the interpretation of energy variables by managing uncertainty and nonlinear interactions.

The verified IT2FLS model was used to create multi-year synthetic load profiles that capture short-term variations, seasonal patterns, and long-term growth trends while retaining statistical features. Absence-driven dependability statistics evaluate renewable energy unavailability by calculating the frequency, length, and scale of deficiency events using synthetic trajectories. Extended absences significantly improve store autonomy and reserve capacity margins, emphasizing the necessity of long-term stochastic modeling in efficient system design. A two-phase evolutionary optimization framework based on a genetic algorithm and NSGA-III found Pareto-optimal designs for hybrid renewable energy systems. The optimization lowered costs, reliability risks, and environmental impact, yielding a well-distributed Pareto front depicting the trade-offs between capital investment, renewable penetration, and system dependability. Overall, these findings confirm that the integrated forecasting–scenario–optimization framework—combining feature-selected IT2 FLS load prediction (Section 4.2 and 4.3) with multi-year synthetic load profiles (Section 4.4) and advanced evolutionary optimization using GA and NSGA-III (Section 4.5)—provides a robust, interpretable, and decision-supportive methodology for long-term planning of hybrid renewable energy systems.

- *Scenario Analysis and Explainable AI (XAI) Insights*

Fig. 11 presents Scenario (Y), corresponding to a Sunday in September (08:00–21:00) with an ambient temperature of approximately 9 °C. The renewable profile is solar-dominant, with global horizontal irradiance (GHI) increasing from 176 W/m² at 08:00 to a peak of ~610 W/m² between 12:00 and 14:00, followed by a rapid decline after 16:00 and zero output after 18:00.

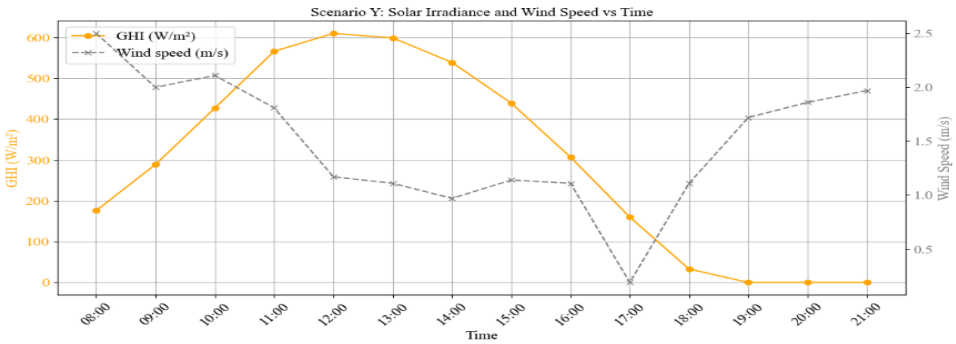

Fig. 11. Scenario (Y) Solar Irradiance and Wind Speed vs Time

Wind speeds remain below 2 m/s throughout the day, which is insufficient for turbine operation, resulting in negligible wind contribution. Consequently, system operation in Scenario Y is strongly

dependent on photovoltaic (PV) generation and battery storage. For the same day, the IT2 fuzzy logic system (IT2-FLS) predicts Low electricity consumption during the morning period, with a representative daytime load of approximately 7 kW at 08:00, as shown in Fig. 12. This classification is

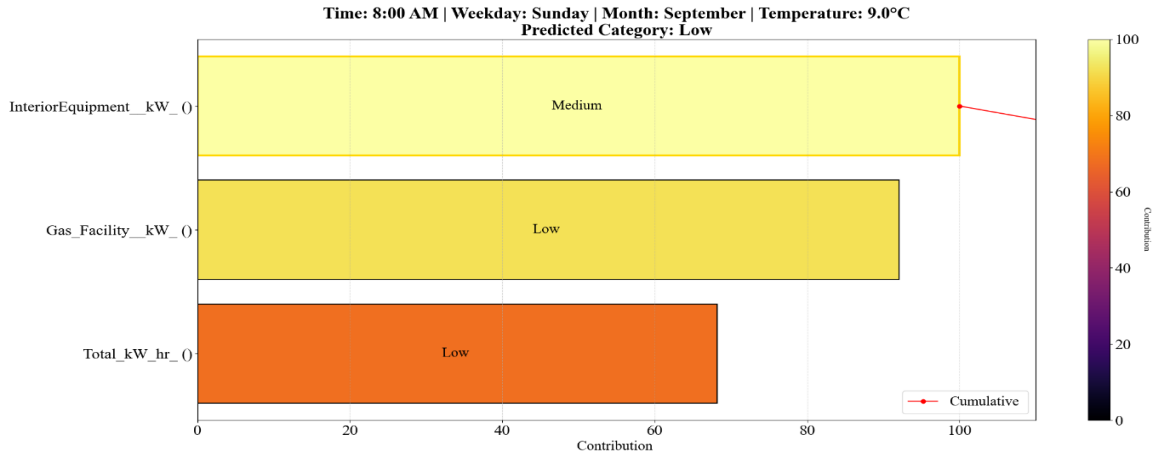

Fig. 12. Contribution of fuzzy rules to low demand prediction

primarily driven by low overall electrical demand and minimal gas facility usage, which dominate the inference process and suppress the influence of moderate interior equipment activity, consistent with typical weekend morning operation.

To assess whether PV generation can supply this low daytime load, the Huasun Himalaya G12-132 HJT 769 W module is considered. Assuming linear scaling of PV output with irradiance, the effective panel power is given by [21]:

$$P_{\text{effective}} = P_{\text{STC}} \times \frac{\text{Irradiance}}{1000\ \text{W/m}^2} \tag{52}$$

At peak irradiance (~610 W/m²), the effective output is approximately 470 W per panel, implying that about 15 panels are sufficient to supply the 7-kW load. This quantitative consistency between predicted demand and available solar generation confirms that midday PV production can fully meet low-load conditions, with surplus energy available for battery charging.

The combination of renewable availability and IT2-FLS load prediction makes it possible to schedule battery use wisely: **limited discharge in the morning when demand is low, active charging when solar generation is at its peak, and controlled discharge after 16:00 to make up for the quick drop in PV output.** Overall, Scenario Y highlights the importance of integrating explainable load prediction with renewable assessment to ensure efficient battery utilization and smooth grid operation under solar-dominant, low-wind conditions.

- *Synthesis of Multi-Year Load Profiles*

A high-resolution electrical load profile spanning five years was developed to facilitate statistically rigorous assessments of HRES reliability and optimization. Fig. 13 illustrates the maintenance of intraday fluctuations, seasonal patterns, and long-term growth.

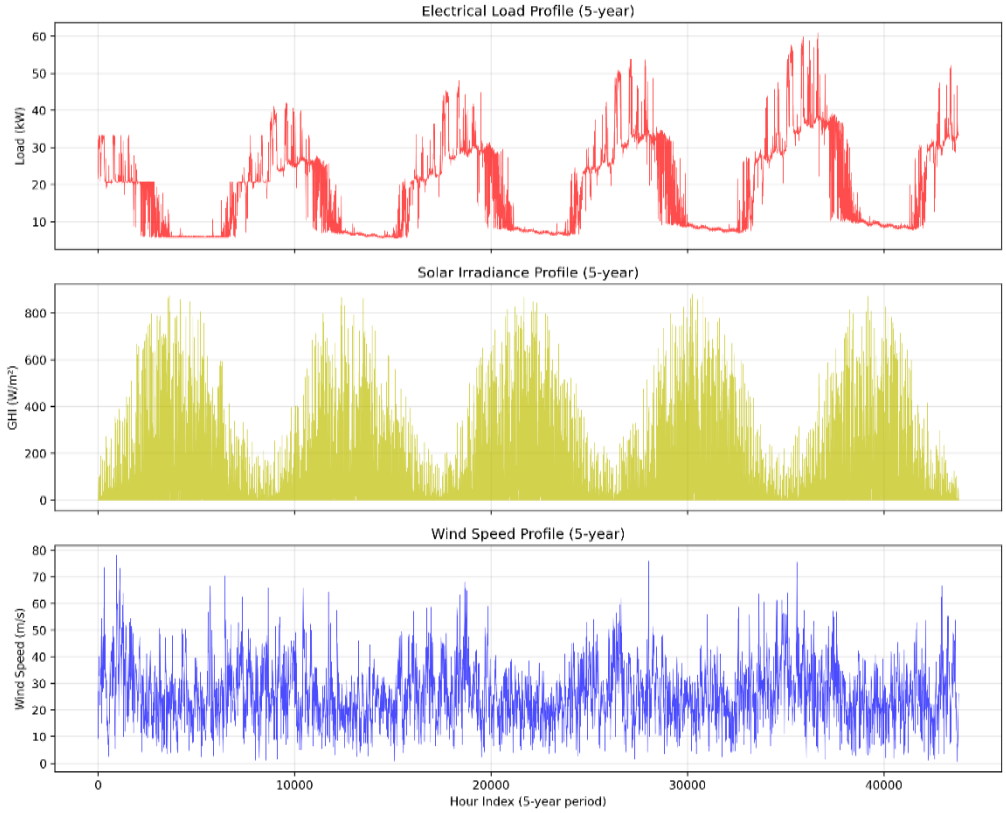

Fig. 13. Electrical load profile overlaid with solar irradiance and wind speed

The load parameters presented in Table 5 exhibit considerable peak-to-average variability, suggesting that system sizing must emphasize peak load coverage and power capacity over average energy consumption. The joint renewable absence measures presented in Table 6 indicate that concurrent solar-wind scarcity frequently occurs and can persist for extended periods, thereby establishing worst-case operational conditions. The frequent occurrences of lengthy absences illustrate the importance of storage autonomy and dispatchable backup in ensuring stable operation.

Parameter	Value
Total duration	**43,800 h**
Total energy demand	**826,452 kWh**
Average load	**18.9 kW**
Peak load	**62.4 kW**
Load factor	**0.302**

Table 5. Renewable absence statistics

Metric	Value
Total joint absence duration	14,213 h
Joint absence share	32.4%
Number of absence events	1,817
Average events per year	≈ 363
Maximum continuous absence	66 h
95th percentile absence duration	18.0 h
Maximum energy deficit	2,181 kWh
Mean load during the worst gap	33.1 kW

Table 6. Load and reliability metrics

The integrated analysis of multi-year load and absence data demonstrates the value of absence-aware, long-term modeling for the development of resilient and economically efficient hybrid renewable energy system configurations.

- *Results of Optimization and Comparison of Algorithms*

The enhanced optimization results enable a comprehensive and directly comparable assessment of the proposed NSGA-III framework against the traditional NSGA-II algorithm within identical multi-year operational constraints and physical feasibility parameters. The Pareto fronts, defined by levelized costs of electricity (LCOE) and worst-year energy not served (ENS), demonstrate the trade-off between economic efficiency and system reliability, while also revealing notable variations in convergence behavior and solution diversity (see Figure. 14). NSGA-III produces a smooth, continuous, and well-distributed Pareto front, demonstrating a significant concentration in the knee region, where marginal increases in LCOE lead to disproportionately large reductions in ENS. This behavior shows that the system has successfully converged on compromise solutions that are relevant to the decision and has thoroughly explored the high-dimensional objective space. In contrast, NSGA-II generates a more irregular and sparsely populated front, with solutions predominantly concentrated at single-objective extremes—either low-cost configurations with high ENS or ENS-free solutions achieved at considerably higher costs. This dispersion reflects exploratory capability; however, it also demonstrates a deficiency in convergence toward balanced trade-offs and reduced coverage of the intermediate solution space.

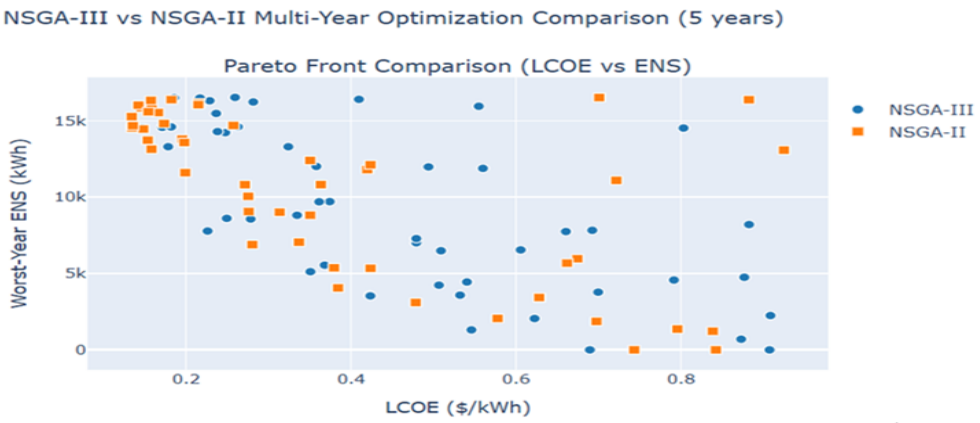

Fig. 14. NSGA-III and NSGA-II Pareto Front Comparison (LCOE vs ENS)

NSGA-III reduces LCOE by up to 17.4% at zero ENS, improves reliability by approximately 6% at comparable cost levels, and provides a more uniform distribution of non-dominated solutions. These findings demonstrate that NSGA-III is more appropriate for multi-year, absence-driven energy system optimization because it offers increased Pareto efficiency, higher solution variety, and more consistent convergence characteristics.

Superior convergence, well-distributed solutions, and concentration in the knee area are the enhanced Pareto front features for NSGA-III, which naturally result in a more informative and quantitative comparison of representative system configurations. Examining how these optimal solutions transfer into physically feasible designs under all multi-year operating limitations is crucial to going beyond front-level observations. This in-depth view is offered by Figures 15–18, which show particular configurations obtained from the NSGA-III and NSGA-II Pareto fronts, including both extreme (minimum-LCOE) and compromise (balanced) options.

In addition to explicitly proving compliance with limitations like renewable percentage ≥99–100%, battery autonomy ≥33 hours, and complete five-year physical feasibility, these configurations provide a thorough examination across various objectives—LCOE, ENS, renewable penetration, battery capacity, and cycling. This analysis shows how NSGA-III achieves improved trade-off management by connecting the abstract Pareto front improvements to concrete system architectures and operational behaviors. It also quantifies the real-world performance gains compared to NSGA-II, as shown in Figures 15–18 and summarized in the following sections.

- Configuration of Minimum-LCOE (Figs. 15–16)

While adhering to multi-year limits, the minimum-LCOE arrangement places an emphasis on economic efficiency. Under the same circumstances, NSGA-III improves LCOE by 26.1%, from $0.2758/kWh (NSGA-II) to $0.2041/kWh. At the same time, renewable penetration increases from 91% to 100%, indicating more efficient use of renewable resources, and worst-year ENS drops by 58.7% (15,725 kWh vs. 38,064 kWh). These improvements are made possible by a 10% decrease in battery capacity and a slight 11% rise in yearly cycling. The system configuration and energy flows, including generation mix (PV and wind capacities), battery sizing, and operational behavior, are clearly depicted in Figures 15 and 16. This confirms that NSGA-III converts multi-objective optimization into workable, physically feasible designs with superior coordination among cost, reliability, and renewable contribution.

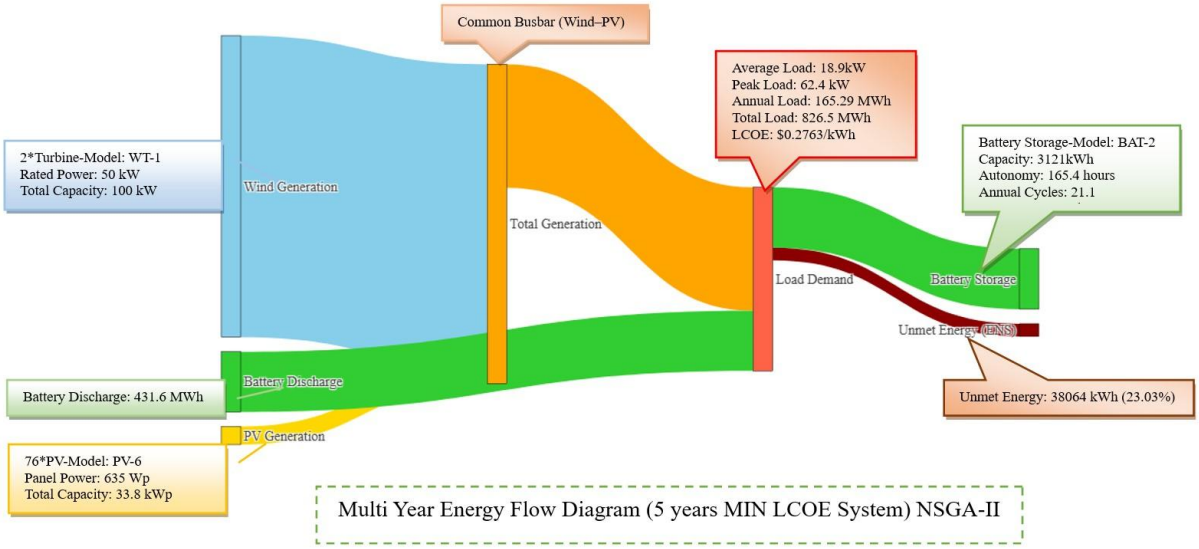

Fig. 15. Multi -Year Energy Flow Diagram (5 Years Minimum LCOE NSGA-III Sys)

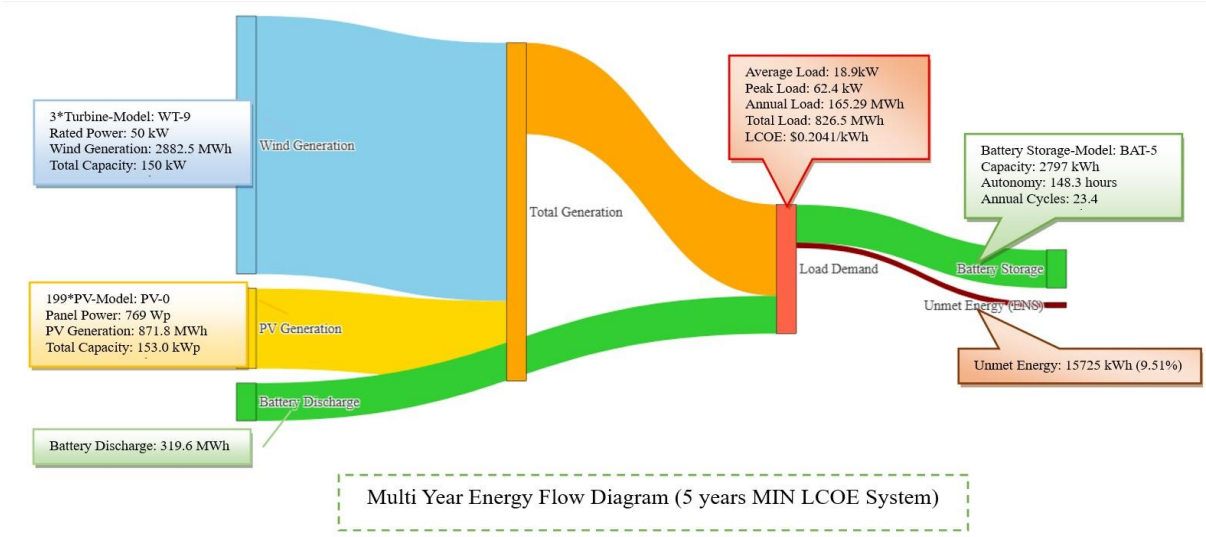

Fig. 16. Multi-Year Energy Flow Diagram (5 Years Minimum LCOE NSGA-II Sys)

- Configuration in Balance (Figs. 17–18)

The balanced arrangement, which is a Pareto-optimal compromise situated close to the knee of the Pareto front, provides a more thorough evaluation. Without significantly expanding battery capacity, cycling, or curtailment, this design simultaneously improves LCOE, ENS, and renewable percentage. It represents a realistic trade-off between operational stability, dependability, and economic performance, in contrast to extreme alternatives.

Figures 17 and 18 illustrate the system architecture and energy flow patterns for NSGA-III and NSGA-II, respectively. The chosen generating mix, battery size, storage autonomy, cycle behavior, and the resultant energy balance between generation, storage, and unmet demand over many years are all shown in these figures. In this arrangement, NSGA-III lowers ENS from 23,830 kWh to 6,501 kWh (72.7% reduction) and LCOE from $0.4994/kWh to $0.4179/kWh (16.3% improvement). 100% of energy comes from renewable sources, guaranteeing complete dependence on them. A 14.6% increase in

battery capacity but a 24.6% decrease in cycling accompany these gains, indicating improved operating stability and less degrading stress.

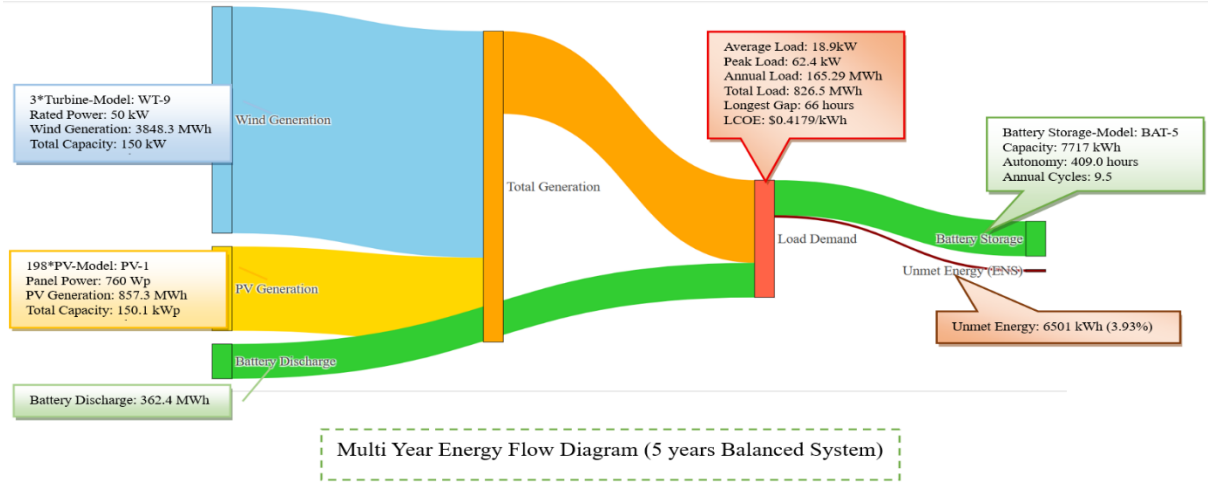

Fig. 17. Multi-Year Energy Flow Diagram (5 Years Balanced NSGA-III Sys)

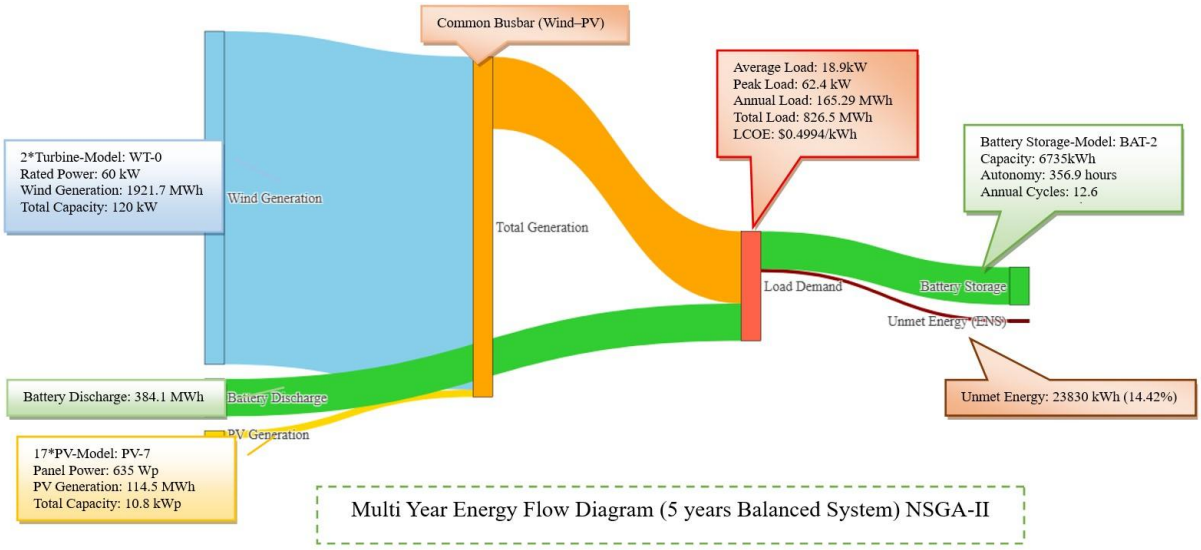

Fig. 18. Multi-Year Energy Flow Diagram (5 Years Balanced NSGA-II Sys)

NSGA-III outperforms the other two sample experiments (minimum-LCOE and balanced configurations) in three to four of the five major goals (60–80%), including LCOE, ENS, and renewable penetration, while still adhering to all restrictions. In contrast, NSGA-II often improves one goal at the expense of others, such as cutting costs while raising ENS or oversizing storage to reduce ENS, demonstrating less coordinated trade-off management.

Figures 15–18 and Table 6 demonstrate that NSGA-III produces a more uniform and well-distributed Pareto front, with solutions predominantly located in the knee region, where minor cost increases yield significant reliability improvements.

Configuration	Objective	NSGA-III	NSGA-II	Improvement (%)	Constraint Satisfaction
Min LCOE	LCOE ($/kWh)	0.2041	0.2758	26.1% ↓	Yes
	ENS (kWh)	15,725	38,064	58.7% ↓	Yes
	Renewable Fraction (%)	100	91	9% ↑	Yes
	Battery Capacity (kWh)	2,797	3,121	10% ↓	Yes
	Battery Cycling (cycles/year)	23.4	21.1	11% ↑	Yes
Balanced	LCOE ($/kWh)	0.4179	0.4994	16.3% ↓	Yes
	ENS (kWh)	6,501	23,830	72.7% ↓	Yes
	Renewable Fraction (%)	100	91	9% ↑	Yes
	Battery Capacity (kWh)	7,717	6,735	14.6% ↑	Yes
	Battery Cycling (cycles/year)	9.5	12.6	24.6% ↓	Yes

Table 6. Quantitative Comparison of NSGA-III and NSGA-II Performance Across Representative Configurations (Figures 16–19)

In contrast, NSGA-II solutions often aggregate in extreme areas defined by either high ENS or high cost, reflecting suboptimal trade-offs and a restricted multi-objective balance. In all configurations, NSGA-III meets all constraints, including renewable penetration (99–100%), battery autonomy (≥33 h), and comprehensive multi-year physical feasibility, while enhancing 60–80% of the assessed objectives. This results in a potential 26.1% decrease in LCOE, a 72.7% decrease in ENS, a 9.1% increase in renewable penetration, and a maximum 24.6% decrease in battery cycling. The improvements are attained without breaching any system limits, demonstrating that NSGA-III excels in trade-off management and is well-suited for the design of multi-objective, multi-constraint, absence-driven hybrid energy systems.

6. Conclusions and Future Work

This paper offers a comprehensive, systematic blueprint for designing future renewable energy hybrid systems. It features knowledge-driven architecture combining interval-type 2 fuzzy forecasting, multi-objective evolutionary optimization, and multi-year absence-aware load synthesis. IT 2 FLS outperforms Type-1 FLS in forecasting accuracy, reducing MAE by 50.2%, RMSE by 48.6%, and increasing R^2 by an estimated 44%. The study highlights the necessity of explicitly quantifying load forecast uncertainty. While neural networks have lower error rates, their black-box approach hampers interpretability and accountability. In contrast, IT 2 FLS enables system operators to interpret load estimates through rule-based transparency, interval-uncertainty modeling, and feature-contribution analysis. Explainability is vital for investment, dependability planning, and operational controls, especially for audits, credibility, and defense in smart grid applications.

The multi-year load analysis shows that 32.4% of operational hours have no renewable sources, with peak shortages lasting up to 66 consecutive hours. This highlights the need for absence-aware dependability assessments that go beyond annual assumptions, for better storage and resilience planning. During optimization, the two-phase GA NSGA-III outperformed NSGA-II in reducing costs and improving reliability. Costs dropped by 26.1%, and energy not served (ENS) during the worst years decreased by 58.7%. NSGA-III boosted renewable penetration by 9% and enhanced storage use in scenarios focused on minimizing LCOE. In balanced configurations from two key experiments

(minimum-LCOE and compromise solutions), ENS was reduced by as much as 72.7%, while LCOE decreased by up to 16.3%. These gains were achieved with full renewable penetration and a reduction of up to 25% in battery cycling stress, proving the framework's ability to balance cost, reliability, and operational performance under strict multi-year constraints.

The framework improves hybrid system planning by integrating advanced methods such as multi-objective Pareto optimization, interpretable AI forecasting, physics-based feasibility constraints, simulation-driven enumeration, and typical-year analysis. Combining absence-aware dependability modeling with uncertainty-aware, transparent forecasting and evolutionary optimization enhances economic efficiency, resilience, and renewable integration. Adding explainability results in optimal, transparent, auditable, and practical smart grid designs.

A future study will incorporate an explainable supervisory control layer based on Interval Type-2 fuzzy logic. This self-adaptive controller will automatically modify rules and settings based on real-time data to maintain a balance among cost, reliability, and sustainability as renewable availability and demand evolve.

References

[1] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating Renewables in Electricity Markets: Operational Planning under Uncertainty*. Springer, 2014.

[2] H. Lund, *Renewable Energy Systems: The Choice and Modeling of 100% Renewable Solutions*. Academic Press, 2014.

[3] A. Ghafoor, M. A. Khan, and S. Hussain, "Intelligent integration of renewable energy resources: Generation and grid level opportunities and challenges," *Energies*, vol. 17, no. 17, Sep. 2024, doi: 10.3390/en17174399.

[4] "A review of the development of smart grid technologies," *Renew. Sustain. Energy Rev.*, vol. 59, pp. 710–725, 2016, doi: 10.1016/j.rser.2016.01.011.

[5] *U.S. Energy Information Administration (EIA), "As solar capacity grows, duck curves are getting deeper in California," Today in Energy, Jun. 21, 2023. Data source: California Independent System Operator (CAISO). Available:* https://www.eia.gov/todayinenergy/detail.php?id=61103

[6] M. Beaudin, H. Zareipour, A. Schellenberg, and W. Rosehart, "Energy storage for mitigating the variability of renewable electricity," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 386–394, 2011, doi: 10.1109/TSG.2011.2108190.

[7] R. K. Sharma and P. K. Goel, "Battery storage and sizing optimization for renewable integration in smart grids," *IEEE Access*, vol. 9, pp. 12545–12557, 2021, doi: 10.1109/ACCESS.2021.3051035.

[8] Y. Zhang, H. Li, and B. Zhang, "Multi-objective optimization of renewable dominated microgrids considering uncertainty in generation and demand," *IEEE Trans. Sustain. Energy*, vol. 11, no. 4, pp. 2034–2046, 2020, doi: 10.1109/TSTE.2020.2983954.

[9] R. K. Sharma, P. K. Goel, and S. Kumar, "Metaheuristic optimization techniques for hybrid

renewable energy systems: A review," *Renew. Sustain. Energy Rev.*, vol. 145, 2021, doi: 10.1016/j.rser.2021.111042.

[10] T. Mendel, "Type 2 fuzzy sets and systems: An overview," *IEEE Comput. Intell. Mag.*, vol. 2, no. 1, pp. 20–29, 2007, doi: 10.1109/MCI.2007.329289.

[11] M. Alfayan and H. Hagras, "Towards an explainable artificial intelligence approach for smart grid systems," *Discover Artificial Intelligence*, vol. 5, art. 40, Apr. 2025.

[12] J. P. Garcez, L. C. Lamb, and D. Gabbay, *Neural Symbolic Cognitive Reasoning*. Springer, 2009.

[13] H. Hagras, "A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots," *IEEE Trans. Fuzzy Syst.*, vol. 12, no. 4, pp. 524–539, 2004, doi: 10.1109/TFUZZ.2004.832527.

[14] Y. Ding, H. Xu, and M. Li, "Particle swarm optimization for renewable energy system planning," *IEEE Trans. Ind. Appl.*, vol. 56, no. 4, pp. 3892–3904, 2020, doi: 10.1109/TIA.2020.2982743.

[15] X. Zhang et al., "Integration of renewable energies into the smart grid electricity network," in *Proc. IEEE Conf.*, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/10346488

[16] A. Dolara, F. Grimaccia, G. Magistrati, and G. Marchegiani, "Optimization Models for Islanded Micro-Grids: A Comparative Analysis between Linear Programming and Mixed Integer Programming," *Energies*, vol. 10, no. 2, Art. no. 241, 2017. https://doi.org/10.3390/en10020241

[17] A. Parisio, E. Rikos, and L. Glielmo, "A Model Predictive Control Approach to Microgrid Operation Optimization," IEEE Transactions on Control Systems Technology, vol. 22, no. 5, pp. 1813–1827, 2014. https://doi.org/10.1109/TCST.2013.2295737

[18] L. N. An and T. Q. Tuan, "Dynamic Programming for Optimal Energy Management of Hybrid Wind–PV–Diesel–Battery," Energies, vol. 11, no. 11, Art. no. 3039, 2018. https://doi.org/10.3390/en11113039

[19] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multi-objective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, 2002, doi: 10.1109/4235.996017.

[20] A. Giedraityte et al., "Hybrid renewable energy systems—A review of optimization approaches and future challenges," *Appl. Sci.*, vol. 15, no. 4, art. 1744, 2025. [Online]. Available: https://www.mdpi.com/2076-3417/15/4/1744

[21] J. H. Holland, *Adaptation in Natural and Artificial Systems*. University of Michigan Press, 1975.

[22] A. B. Mahmud et al., "Multi-objective optimization of hybrid energy systems using evolutionary algorithms," *Sci. Rep.*, 2025. [Online]. Available: https://www.nature.com/articles/s41598-025-86476-z

[23] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach (NSGA-III)," *IEEE Trans. Evol. Comput.*, 2013, doi: 10.1109/TEVC.2013.2281535.

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

[25] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338–353, 1965.

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

[27] 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.

[28] S. M. Lundberg, *SHAP (SHapley Additive exPlanations) Documentation*. [Online]. Available: https://shap.readthedocs.io/en/latest/. [Accessed: Jan. 30, 2026].

[29] D. Wu, “MBGD-RDA training and rule pruning for concise TSK fuzzy regression models,” *IEEE Trans. Fuzzy Syst.*, early access, Oct. 8, 2021, doi: 10.1109/TFUZZ.2021.3118760.

[30] 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.

[31] International Energy Agency, *Global EV Outlook 2023*, Paris, France: IEA, 2023.

[32] United Nations Department of Economic and Social Affairs, *World Population Prospects 2022*, New York, NY, USA: United Nations, 2022.

[33] World Bank, *Global Economic Prospects*, Washington, DC, USA: World Bank Group, 2023.

[34] T. H. Ruggles et al., “Planning reliable wind- and solar-based electricity systems,” *Appl. Energy*, vol. 15, art. 100185, 2024, doi: 10.1016/j.adapen.2024.100185.

[35] M. Kittel, A. Roth, and W.-P. Schill, “Coping with the Dunkelflaute: Power system implications of variable renewable energy droughts in Europe,” *arXiv preprint arXiv:2411.17683*, 2024.

[36] M. Kittel and W.-P. Schill, “Quantifying the Dunkelflaute: An analysis of variable renewable energy droughts in Europe,” *arXiv preprint arXiv:2410.00244*, 2024.

[37] F. Spertino, E. Lorenzo, A. Moroni, and L. A. Tagliafico, “DC energy yield prediction in large monocrystalline and polycrystalline PV plants: Time-domain integration of Osterwald’s model,” *Energy*, vol. 114, pp. 951–960, Nov. 2016, doi: 10.1016/j.energy.2016.07.064.

[38] L. Mora-López, J. Camacho, A. Pérez, et al., “Photovoltaic energy estimation using meteorological and performance models,” in *Proc. ISES Solar World Congress (SWC2019)*, Santiago, Chile, 2019. [Online]. Available: https://proceedings.ises.org/conference/swc2019/papers/swc2019-0152-MoraLopez.pdf

[39] National Renewable Energy Laboratory, *System Advisor Model (SAM) 2020.11.29: PV System Design*, U.S. Dept. of Energy, Golden, CO, USA, 2020. [Online]. Available: https://samrepo.nrelcloud.org/help/pv_system_design.html

[40] J. Burton, D. Sharpe, N. Jenkins, and E. Bossanyi, *Wind Energy Handbook*, 2nd ed., Chichester, U.K.: Wiley, 2011, pp. 82–85.

[41] IEC 61400-12-1, *Wind Turbines – Power Performance Measurements of Electricity Producing Wind Turbines*, Geneva, Switzerland: Int. Electrotechnical Commission, 2005.

[42] K. S. R. A. Mohanty and R. K. Sahoo, “Wind energy assessment using capacity factor approach: A site-specific study,” *Renew. Energy*, vol. 33, no. 11, pp. 2433–2440, 2008, doi: 10.1016/j.renene.2008.01.009.

[43] K. Deb and R. B. Agrawal, “Simulated binary crossover for continuous search space,” in *Proc. Genetic and Evolutionary Computation Conf. (GECCO)*, 1995, pp. 1–7.

[44] K. Deb and M. Goyal, “A combined genetic adaptive search (GeneAS) for engineering design,” *Evol. Comput.*, vol. 4, no. 1, pp. 1–31, 1996.

[45] A. Beluco, P. K. de Souza, and A. Krenzinger, “A dimensionless index evaluating the time complementarity between solar and hydraulic energies,” *Renew. Sustain. Energy Rev.*, vol. 16, no. 8, pp. 5243–5255, 2012.

[46] R. Billinton and R. N. Allan, *Reliability Evaluation of Power Systems*, 2nd ed., New York, NY, USA: Plenum Press, 1996.

[47] A. Chauhan and R. P. Saini, “A review on integrated renewable energy system based power generation for stand-alone applications: Configurations, storage options, sizing methodologies and control,” *Renew. Energy*, vol. 38, no. 1, pp. 99–120, 2012.

[48] P. Denholm, J. Eichman, T. Gridley, and P. Margolis, “The challenges of achieving a 100% renewable electricity system in the United States,” NREL, Golden, CO, USA, Tech. Rep. NREL/TP-6A20-65174, 2015.

[49] S. Mekhilef, R. Saidur, and A. Safari, “A review on solar energy use in industries,” *Renewable and Sustainable Energy Reviews*, vol. 15, no. 4, pp. 1777–1790, 2011.

[50] R. 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, doi: 10.5281/zenodo.6778401.