

Article

Refining Open-Source Asset Management Tools: AI-Driven Innovations for Enhanced Reliability and Resilience of Power Systems

Gopal Lal Rajora ^{1,*}, Miguel A. Sanz-Bobi ^{1,*}, Lina Bertling Tjernberg ² and Pablo Calvo-Bascones ¹

¹ Institute for Research in Technology, Universidad Pontificia Comillas, 28015 Madrid, Spain

² Division of Electric Power and Energy Systems, KTH Royal Institute of Technology Stockholm, 114 28 Stockholm, Sweden

* Correspondence: glrajora@comillas.edu (G.L.R.); masanz@comillas.edu (M.A.S.-B.)

Abstract

Traditional methods of asset management in electric power systems rely upon fixed schedules and reactive measurements, leading to challenges in the transparent prioritization of maintenance under evolving operating conditions and incomplete data. In this paper, we introduce a new, fully integrated artificial intelligence (AI)-driven approach for enhancing the resilience and reliability of open-source asset management tools to support improved performance and decisions in electric power system operations. This methodology addresses and overcomes several significant challenges, including data heterogeneity, algorithmic limitations, and inflexible decision-making, through a three-module workflow. The data fidelity module provides a domain-aware pipeline for identifying structural (missing) values from explicit missingness using sophisticated imputation methods, including Multiple Imputation Chain Equations (MICE) and Generative Adversarial Network (GAN)-based hybrids. The characterization module employs seven complementary weighting strategies, including PCA, Autoencoder, GA-based optimization, SHAP, Decision-Tree Importance, and Entropy Weighting, to achieve objective feature weight assignment, thereby eliminating the need for subjective manual rules. The optimization module enhanced the action space through multi-objective optimization, balancing reliability maximization and cost minimization. A synthetic dataset of 100 power transformers was used to validate that the MICE achieved better imputation than other methods. The optimized weighting framework successfully categorizes Health Index values into five condition levels, while the multi-objective maintenance policy optimization generates decisions that align with real-world asset management practices. The proposed framework provides the Transmission and Distribution System Operators (TSOs/DSOs) with an adaptable, industry-oriented decision-support workflow system for enhancing reliability, optimizing maintenance expenses, and improving asset management policies for critical power infrastructure.

Keywords: asset health assessment; condition monitoring; power system asset management; predictive maintenance; multi-objective optimization; machine learning; data-driven insights



Academic Editors: Matheus Henrique Dal Molin Ribeiro, Géremi Gilson Dranka, Stefano Frizzo Stefanon and Gurmail Singh

Received: 12 December 2025

Revised: 5 January 2026

Accepted: 8 January 2026

Published: 11 January 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Electric power system asset management (AM) has been the subject of increasing interest due to aging infrastructure, growing demand, and the integration of renewable energy [1,2]. AM methods developed traditionally with static scheduling do not always

provide adequate consistency in prioritizing large groups of equipment under budget and labor limitations. In situations where there are gaps or heterogeneity in the data used for the analysis, adaptive and data-driven approaches can provide greater clarity, reproducibility, and defensibility in decision-making regarding maintenance. In this scenario, data-driven and adaptive strategies are essential for both predictive and preventive maintenance, which are not only cost-efficient but also desired by operations managers [3].

It is also important to recognize that early failures are due to a combination of causes other than sustained overloads, including harmonic-rich load conditions that accelerate thermal degradation, deterioration of insulation/moisture, switching/lightning surges, and/or other external events (e.g., vehicle impacts). Therefore, the proposed method will enable risk-based prioritization using available indicators/proxies, while remaining interpretable to allow engineers to assess whether the recommended priority aligns with likely failure mechanisms.

In recent times, the use of advanced technologies such as AI and ML has created new opportunities to enhance asset condition monitoring and inform decisions related to its maintenance. Many open-source toolboxes [4,5] and frameworks [6] have been developed for TSOs/DSOs during this transition. The authors of this study have also made a previous contribution (for improved continuity and deeper technical context, we recommend reviewing our earlier publication [7]) to this area as the developers of the ATTEST open-source toolbox (as part of the European Horizon project), which utilizes clustering and reinforcement learning (RL) to optimize maintenance strategies [7]. Table 1 presents the major methodological differences between ATTEST Toolbox [7] and the methodology used in this research for determining the comparative advantages of the proposed methodology as a solution to some of the problems common to other data-based asset management systems, such as those described in previous studies and field implementation experiences through improved data management processes, objective Health Index development, and transparent decision-making.

However, the real-world deployment of such AM tools, including the author's earlier work [7] and other state-of-the-art solutions reviewed here, has exposed certain systemic issues related to data handling and algorithmic robustness. While moving from theoretical models to real utility environments, researchers and operators often come across three general categories of constraints that limit the applicability of the currently available approaches:

- **Data Fidelity:** Utility data is rarely pristine [8]. It is characterized by extreme heterogeneity (mixing analog and digital records), explicit missingness (gaps in the recording), and implicit "structural zeros" (sensor errors that record zero instead of null). Standard imputation methods, such as mean/median replacement, used in many existing tools, are insufficient for this domain. They are inadequate to capture the nonlinear relationships between asset indicator variables (i.e., load vs. temperature), and therefore may introduce significant statistical bias into the health scores generated from those methods [9–11].
- **Algorithmic Objectivity:** Most current tools depend on a process for weighting features that is based on expert experience or other subjective methods. Although important, most of these manual, heuristic processes are static, i.e., do not adapt to the various statistical realities of the various data sets they are used to analyze. In addition, as noted with both commercial software and previous research studies, they are limited in their ability to capture non-linear relations and changes in degradation trends in aging fleets [9,10].
- **Strategic Flexibility:** A common limitation in data-driven AM tools is the rigidity of the decision-making policy. Many frameworks do not incorporate multi-objective

optimization, which is capable of balancing competing goals, such as minimizing cost versus maximizing reliability, across a full spectrum of preventive, corrective, and predictive actions [9,12].

Table 1. Comparison between the ATTEST Toolbox [7] and the present work.

Aspect	Previous Work: ATTEST Toolbox [7]	Present Work
Primary objective	Development and demonstration of an open-source asset management toolbox within the ATTEST project	Refinement and extension of open-source asset management tools to improve robustness, objectivity, and decision consistency
Motivation	Enable proactive maintenance planning using clustering and reinforcement learning	Address methodological limitations identified in prior research and implementation experience, including data heterogeneity, subjective weighting, and rigid decision policies
Data assumptions	Assumes preprocessed or sufficiently complete datasets; missing data handled implicitly or heuristically	Explicitly assumes incomplete, heterogeneous, and noisy utility data; introduces domain-aware missingness detection and validation-driven imputation
Treatment of missing data	Limited discussion; relies on basic preprocessing and user intervention	Dedicated data fidelity module distinguishing explicit missingness and structural zeros, with benchmarking of advanced imputation methods (MICE, GAN-based hybrids)
Feature weighting	Expert-driven and heuristic weights defined by users	Fully data-driven, multi-method weighting framework (Entropy, PCA, Autoencoder, GA, SHAP, Decision Tree) with clustering-based validation
Objectivity of Health Index (HI)	Partially subjective due to manual weight assignment	Objective and adaptive HI construction assessed using clustering-quality consistency metrics
Health state definition	Three discrete states (low, medium, high)	Five standardized condition levels with fixed thresholds over a normalized [0, 1] interval
Decision-making approach	Reinforcement learning (Q-learning) with discrete state space and hand-crafted rewards	Transition-free, interpretable multi-objective optimization balancing risk and cost
Maintenance policy generation	Policy learned via Q-matrix; dependent on reward tuning and state transitions	Direct optimization of maintenance actions using explicit risk–cost trade-offs without trajectory data
Interpretability	Moderate; learned policies can be difficult to explain to practitioners	High; explicit indicators, weights, thresholds, and penalties aligned with engineering reasoning
Practical deployment focus	Toolbox architecture and integration within the ATTEST research platform	Methodological redesign informed by prior implementation experience, emphasizing transparency and configurability under realistic data constraints
Validation data	Synthetic grid (600 transformers) and limited real utility dataset (92 transformers)	Physics-informed synthetic dataset with controlled missingness to enable objective benchmarking against known ground truth
Key contribution	Proof-of-concept open-source asset management toolbox	Second-generation, industry-oriented decision-support framework emphasizing robustness, transparency, and methodological rigor

The scientific value of this work is rooted in its ability to address the key challenges by offering advanced, rigorously developed, and domain-specific improvements to the AM process. This paper builds upon lessons learned from developing the ATTEST toolbox, as well as addressing the gaps in the broader literature, to develop an adaptable architecture

that provides a robust and industrially oriented decision-support system, by making the following contributions:

1. Implementing a domain-aware pipeline that distinguishes between explicit missingness and invalid structural zeros, using robust scaling and benchmarking advanced imputation techniques (e.g., MICE and GAN-based hybrids) to ensure data integrity [13].
2. Optimization of feature weight assignment: this approach can be utilized to incorporate various important features, such as a multi-method weighting framework, by integrating different factors, including entropy and genetic algorithms. This feature enhances the quality of data clustering and asset management [12].
3. Expanded action space: introducing a full spectrum of maintenance options optimized using a meta-heuristic and multi-objective optimization framework for sophisticated maintenance policy generation.
4. Case study on a new synthetic Power Transformers dataset: demonstrating the effectiveness of the improved methodology and rigorously benchmarking its performance.

By systematically addressing these universal data and algorithmic challenges, this work seeks to provide TSOs and DSOs with a more reliable and adaptable decision-support system. This improved structure is not merely a refinement of previous tools, but a necessary evolution to solve the common problems inherent in working with complex AM data, offering practical advantages such as greater reliability, lower maintenance expenses, and better AM policies.

2. Methodology

To address the complexity of power system data, the workflow is structured into three logical blocks: Module I (data imputation), Module II (characterization), and Module III (optimization).

- **Module I**, focuses on the reliability of the input data through ingestion, identification of valid versus invalid missing values, and application of sophisticated imputation for completion of a comprehensive dataset.
- **Module II** ensures that the assessment is objective. The “Total Indicator (Health Index)” for an asset is not based upon manual rules, but instead, it is derived from mathematical determination of the relative importance of each feature using multiple algorithms.
- **Module III** involves determining the optimal decision (i.e., prescription) of the maintenance action(s) that the system should take, as a function of the Health Index, and utilizing a meta-optimization process that balances the risk and cost of the prescribed actions.

Figure 1 illustrates the information flow and internal steps of this methodology.

2.1. Module I: Data Imputation for Power-System Assets

In the case of power systems, data incompleteness often occurs due to sensor failures, communication delays, manual input errors, or inadequate regular checks. Unlike in other domains, incomplete records cannot be easily scrapped, since all transformers, cables, circuit breakers, and substations are operationally vital for providing safe and reliable grid performance. As a result, a powerful and domain-specific imputation plan emerges as a fundamental requirement for the reliability of asset health assessment and maintenance judgments. Therefore, this module addresses the persistence issues of data analytics, i.e., the lack of data or its incompleteness.

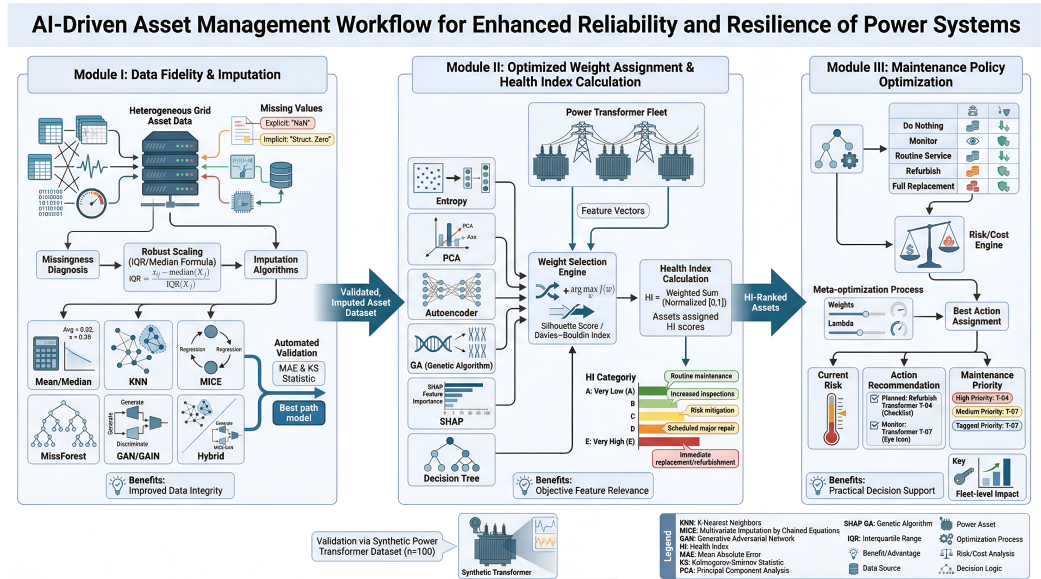


Figure 1. Information flow and internal steps of the proposed methodology.

Mean or median replacement, regression imputation, and expectation maximization (EM) are some of the most common traditional imputation methods used in studies of condition monitoring. These techniques are useful for stabilizing small-scale analyses; however, they also introduce statistical bias, underestimate variance, and fail to capture nonlinear relationships between correlated indicators (e.g., transformer temperature, dissolved gas, and fault current). These distortions can be transmitted downstream, resulting in unreliable health indices, misleading clusters, and suboptimal maintenance policies [14].

The proposed imputation framework differs from these classical methods. It proposes a domain-aware, adaptive, and validation-driven pipeline, which (i) differentiates between valid and invalid zeros, (ii) uses robust normalization as a way of reducing the impact of outliers, and (iii) compares various sophisticated imputation algorithms, automatically choosing the most appropriate one for a specific dataset. This will ensure that all the records are incorporated to the model without interfering with the data integrity and interpretation.

2.1.1. Mathematical Formulation

Let $X = [x_{ij}] \in \mathbb{R}^{n \times p}$ represent the condition-monitoring dataset containing n assets and p condition indicators. Corrupted or missing entries are indicated by a binary mask matrix $M = [m_{ij}]$, which is defined as

$$m_{ij} = \begin{cases} 1, & \text{if } x_{ij} \text{ is observed and valid,} \\ 0, & \text{if } x_{ij} \text{ is missing, invalid, or corrupted.} \end{cases} \quad (1)$$

The objective behind the imputation process is to estimate the missing values of X and form a full dataset \hat{X} , such that the imputed values preserve the statistical properties, relationships, and engineering interoperability of the observed values. The imputed matrix, \hat{X} , is formally calculated by minimizing the reconstruction error among the observed (non-missing) values, subject to preserving the empirical distributions and dependencies among condition indicators:

$$\hat{X} = \arg \min_{\tilde{X}} \mathbb{E} \left[\|(X - \tilde{X}) \odot M\|^2 \right] \quad (2)$$

where \tilde{X} represents a candidate imputed matrix, a possible reconstruction of the full dataset where all missing values have been filled in and \odot denotes feature-wise multiplication.

A diagnosis of the missingness must be conducted prior to imputation to determine explicit and implicit data missingness:

- **Explicit Missingness:** values explicitly recorded as NaN, null, or undefined due to missing field reports or communication failures.
- **Implicit Missingness:** structurally zero, abnormally low or abnormally high constants caused by an equipment failure or malfunctioning of loggers (e.g., energy throughput = 0 an active transformer).

Reassigning this invalid zero (missingness) is performed using domain-specific constraints, which assure consistency in engineering (e.g., H₂ criticality = 0 is valid, but energy = 0 when a transformer is operating is not).

All the numerical variables are converted to an interval scale by robust scaling [15], which introduces the median of each variable and then normalizes it by the interquartile range (IQR):

$$x'_{ij} = \frac{x_{ij} - \text{median}(X_{.j})}{\text{IQR}(X_{.j})} \quad (3)$$

where

x_{ij} = original value of feature in row i and column j ,

$X_{.j}$ = The set of all values rows of column j ,

x'_{ij} = scaled (robust-normalized) value for row i and column j .

The normalization reduces the effect of extreme or noisy measurements, which are frequently encountered in field-collected asset measurements, allowing distance-based and model-based imputers to work with more consistent feature levels. It also enhances the comparability, characterization and differentiation of values. The whole imputation process can be formulated as

$$\begin{array}{c} X' \xrightarrow{\text{Missingness Detection}} M \xrightarrow{\text{Robust Scaling}} X_{\text{norm}} \\ \xrightarrow{\text{Model-based Imputation}} \hat{X} \xrightarrow{\text{Validation}} \hat{X}_{\text{final}} \end{array} \quad (4)$$

where \hat{X}_{final} represents the verified and reconstructed dataset that will be used as the input in the clustering, weighting, and maintenance optimization modules in further modules.

2.1.2. Imputation Implicates and Rationale

The step estimates the missing entries in X using a number of consistency algorithms which depict a number of features of the data structure.

k -Nearest Neighbors (KNN)

The KNN implies imputation of missing values by the mean of the N -nearest neighbors and is computed with the help of the Euclidean Distance on robustly scaled features. It operates through local similarity and is effective in situations where there is local similarity among assets in terms of mode of operation.

Multiple Imputation Chain Equations (MICE)

MICE optimistically predicts every feature with a missing feature as a regression of all other features in a cyclic manner. The methodology reflects the existence of multivariate dependencies and the conditional relationships between the asset indicators [16].

MissForest

MissForest employs a random forest ensemble to repeatedly predict the missing values, enabling the model to capture both nonlinear interactions and mixed-type data, without making assumptions about a specific distribution.

GAN-Based Imputation (GAIN)

GAIN employs a generative-adversarial learning approach, where a generator predicts missing values and a discriminator determines the difference between the observed and imputed values. This approach generates real forces that are compatible with the joint feature distribution [17].

Hybrid Methods

Two hybrid models are adopted: (i) KNN-GAN, where the GAN is initialized by KNN imputations and includes locality, and (ii) MICE-GAN, which uses MICE results to include multivariate structure. These models combine deterministic stability with generative flexibility.

The selected algorithms include three families of instance generative GAN, and hybrid models are covered to ensure that the local, global and distributional attributes of asset data are well represented. This allows the automated validation module to select the most appropriate imputer for the specific dataset, based on its quantitative performance metrics.

2.1.3. Validation, Automation, and Final Selection

Once the imputed datasets have been produced by each model, the best-performing imputer is selected, and this process is automated based on the number and statistical consistency. This is achieved by hiding a set of known values to create the illusion of unobserved data, and then testing each algorithm to reconstruct the values. The performance is measured by the point-wise reconstruction error and the similarity between the imputed distribution of features and those observed, respectively, using the Mean Absolute Error (MAE) and the Kolmogorov–Smirnov (KS) statistic [18]. The smaller values of these metrics represent better imputations. In order to bring these criteria together, a composite score can be defined as

$$S = \alpha \widetilde{\text{MAE}} + (1-\alpha) \text{KS} \quad (5)$$

where $\widetilde{\text{MAE}}$ is the normalized mean absolute error of the above-mentioned methods and α is a weighting coefficient between accuracy and distributional similarity. The imputational algorithm, which has the lowest composite score S , automatically becomes the final algorithm to be applied to the present dataset. Where two methods are similar (within a 5% deviation), a hybrid model is considered more robust to data heterogeneity.

The final complete dataset X^{imp} will be the imputed dataset, which will be filled with the statistically consistent estimations of the missing or invalid values. This standardized input is the verified data that will be used by the following modules, e.g., clustering, weight assignment, and optimization of maintenance.

2.2. Module II: Optimized Weight Assignment and Health Index Categorization for Power-System Assets

A health indicator in AM usually brings together different pieces of information about an asset's life. This can include its physical condition (for example, results from DGA testing), how it has been operated over time (such as historical loading), and basic characteristics like its age. Because each of these factors matters differently, they are weighted given different levels of importance and combined into a single, clear metric or KPI that reflects the asset's overall health. To construct an effective Health Index (HI) and

facilitate the maintenance decision-making process in power system asset management, it is crucial to accurately identify the relative significance of condition indicators. The obtained feature weights used in our previous work [7] were mostly based on expert knowledge and heuristic scaling, which were interpretable but still subjective across datasets, and unresponsive to the nonlinear relationships that exist with real asset data. To address these shortcomings, this module proposes an entirely data-driven, adaptive, and validation-based weighting system that objectively estimates the relevance of features, premised on the statistical and structural properties of the data. The outcome is as follows: (i) the optimization of a feature-weight vector w^* , and (ii) standardized mapping of HI values into operationally meaningful condition categories.

2.2.1. Mathematical Formulation

Let $X \in \mathbb{R}^{n \times p}$ denote the fully preprocessed and normalized dataset obtained from Module 1, where n is the number of assets and p the number of condition indicators. The goal is to compute a non-negative, normalized weight vector:

$$w = [w_1, \dots, w_p], \quad w_j \geq 0, \quad \sum_{j=1}^p w_j = 1 \quad (6)$$

This is the contribution of each feature to the final Health Index.

Given a candidate weight vector w , the weighted dataset is computed as

$$X_w = X \odot w \quad (7)$$

where \odot denotes feature-wise multiplication. The optimal weight vector is selected by maximizing a clustering-quality objective:

$$w^* = \arg \max_w J(w) \quad (8)$$

where $J(w)$ is a composite measure derived from cluster separability and compactness and feature-weight vector w^* .

2.2.2. Multi-Method Weight Computation

To ensure robustness, interpretability, and generalizability across different utilities and asset types, seven complementary weighting strategies are employed. These methods capture feature relevance from statistical variability, latent structure, expert-informed comparisons, optimization search, and model-based interpretability.

Entropy Weighting

Entropy quantifies the information content of each feature [19]. The normalized probability distribution of feature j is

$$p_{ij} = \frac{x_{ij}}{\sum_i x_{ij} + \epsilon} \quad (9)$$

and its entropy is

$$E_j = -\frac{1}{\ln n} \sum_i p_{ij} \ln(p_{ij} + \epsilon) \quad (10)$$

The final normalized weight is

$$w_j^{(\text{ent})} = \frac{1 - E_j}{\sum_k (1 - E_k)} \quad (11)$$

PCA-Based Weighting

Principal Component Analysis (PCA) [20] captures variance in contributions. The weight of feature j is

$$w_j^{(\text{pca})} \propto \sum_{k=1}^K |\lambda_k l_{jk}| \quad (12)$$

where λ_k is the explained variance and l_{jk} is the loading.

Autoencoder-Based Weighting

A shallow autoencoder [21] is trained to minimize reconstruction error. Features with lower reconstruction error are deemed as more important:

$$w_j^{(\text{ae})} = \frac{1/(r_j + \epsilon)}{\sum_k 1/(r_k + \epsilon)}, \quad r_j = \frac{1}{n} \sum_i (x_{ij} - \hat{x}_{ij})^2 \quad (13)$$

GA-Based Optimization

A Genetic Algorithm [17] searches for weights that maximize clustering coherence:

$$w^{(\text{ga})} = \arg \max_w S(\text{KMeans}(X \odot w, k)) \quad (14)$$

where $S(\cdot)$ represents the Silhouette Score.

SHAP-Based Weighting

A Random Forest classifier is trained to predict cluster labels obtained from Module 1. The SHAP-based weight [22] is

$$w_j^{(\text{shap})} = \frac{1}{n} \sum_i |\phi_{ij}| \quad (15)$$

where ϕ_{ij} is the SHAP value.

Decision-Tree Importance

A CART [23] model is used to predict cluster labels. Feature importance-based weight is

$$w_j^{(\text{dt})} = \frac{I_j}{\sum_k I_k} \quad (16)$$

where I_j is the Gini-based importance.

2.2.3. Validation and Adaptive Hybrid Selection

For each method, clustering is performed on $X_w^{(m)} = X \odot w^{(m)}$ using the optimal number of k clusters from Module 1. Two Complementary validity indices are computed as follows:

Silhouette Score [24]

$$S = \frac{1}{n} \sum_i \frac{b_i - a_i}{\max(a_i, b_i)} \quad (17)$$

Davies–Bouldin Index [25]

$$D = \frac{1}{k} \sum_{i=1}^k \max_{j \neq i} \left(\frac{\sigma_i + \sigma_j}{d(c_i, c_j)} \right) \quad (18)$$

The unified evaluation metric is

$$J(w^{(m)}) = S(w^{(m)}) - D(w^{(m)}) \quad (19)$$

Let $w_{(1)}$ and $w_{(2)}$ be the highest-ranked methods by $J(w)$. The final weight vector is selected by

$$w^* = \begin{cases} \text{normalize}\left(\frac{w_{(1)} + w_{(2)}}{2}\right), & \text{if } \frac{|J(w_{(1)}) - J(w_{(2)})|}{J(w_{(1)})} < 0.05, \\ w_{(1)}, & \text{otherwise.} \end{cases} \quad (20)$$

This ensures robustness when two methods perform comparably and prevents overfitting to any single weighting paradigm. The 5% relative difference criterion is proposed as a statistical measure of robustness for selecting the weighting approach with the least amount of risk of poor performance due to incomplete or noisy data, and therefore does not represent an optimal threshold based on empirical testing. The 5% relative difference criterion is introduced as a robustness safeguard rather than as an empirically optimized threshold. When two weighting strategies yield objective values that differ by less than 5%, their performance is considered statistically and practically comparable given the uncertainty introduced by incomplete and heterogeneous data. In such cases, averaging the two best-performing weight vectors reduces sensitivity to noise and avoids overfitting the Health Index construction to marginal performance differences in a single method. Conversely, when the performance gap exceeds this threshold, the best-performing weighting strategy is selected directly to preserve discriminative power.

2.2.4. Health Index Categorization Using Fixed Thresholds

After determining the best weighting w^* , the HI of each asset is calculated as

$$HI_i = \sum_{j=1}^p w_j^* x'_{ij} \quad (21)$$

To facilitate the interpretation of total indicator (also known as HI) values across data sets and to allow consistency, The HI values are remapped to five condition categories with a fixed threshold over the normalized $[0, 1]$ interval, as shown in Table 2.

Compared to the 0.50–0.75 intervals used in the previous framework, the new 0.20-wide intervals are more realistic and smooth progressions of asset severity. This eliminates moderately stressed assets (e.g., $TI \approx 0.75$) from being overclassified to categories of very high. The fixed thresholds are independent of the datasets, which makes it consistent and prevents unwanted variance across the datasets.

Table 2. Condition ranking and recommended actions based on total indicator (TI).

Rank	TI Range	Level	Recommended Action
A	$TI \leq 0.20$	Very Low	Normal time-based maintenance
B	$0.20 < TI \leq 0.40$	Low	Increase detection and correct minor issues
C	$0.40 < TI \leq 0.60$	Moderate	Risk reduction by addressing key contributing factors
D	$0.60 < TI \leq 0.80$	High	Plan repair, rebuild, or major maintenance
E	$TI > 0.80$	Very High	Immediate replacement or refurbishment

The reason for using a 0.2-wide Health Index range is to maximize interpretability, stability, and progressive increases in health issue severity. Although previous frameworks had larger ranges, the smaller range provides an even more incremental increase in the number of condition states that can be represented by the Health Index and also reduces

the effect of small numerical differences in the total index on the categorization of the Health Index.

To avoid sudden changes in categorization due to small differences in imputed or normalized feature values, the framework discretizes the normalized Health Index into five equally sized bands, thereby avoiding artificial boundaries while maintaining sufficient resolution to allow for appropriate maintenance priority assignments. The fixed width of each band was chosen to be independent of the specific datasets being used, in order to maintain consistency and repeatability across multiple assets and operational environments. However, it could be adjusted if needed, based on utility-specific policies or risk tolerances.

The proposed Module II provides a stringent, data-driven framework that comprises seven complementary weighting algorithms, a clustering-based validation measure, and a hybrid selection rule to derive an optimal feature weight vector. The resulting HI is then categorized using standardized, domain-consistent thresholds, ensuring stability, interpretability, and operational usability. This improves upon earlier expert-driven weighting schemes and lays a robust foundation for clustering, prioritization, and reinforcement-learning-based maintenance optimization in Module III.

2.3. Module III: Data-Driven Maintenance Policy Optimization Using Risk–Cost Parameter Search

In practical power-system asset management, maintenance decisions must balance three competing factors: the physical condition of the asset (life assessment), the urgency of the necessary interventions (maintenance strategy), and the financial implications (economic impact). In the earlier framework, this decision-making step relied on a Q-learning algorithm operating on three condition states (low, medium, high). While effective, this approach assumed a small discrete state space, relied on manual rewards, and required transition dynamics that are difficult to estimate accurately in real industrial datasets. Moreover, the earlier policy did not introduce how life, maintenance, and economic risks should be weighted when evaluating long-term impact. These constraints reduced the flexibility of the model when operating under state transitions and long-horizon trajectories are not reliably available.

This module replaces the RL paradigm with a transparent, transition-free, data-driven meta-optimization framework that directly minimizes fleet-level risk–cost trade-offs. The resulting policy is fully interpretable, requires no historical trajectory data, respects hard engineering constraints, and consistently achieves zero under-treatment of critical assets and zero over-treatment of healthy ones. Table 3 presents an overview of the maintenance actions, including cost assumptions and their quantified impact on risk reduction, which form the basis for subsequent evaluation.

2.3.1. Mathematical Formulation

$x_i^{(L)}, x_i^{(M)}, x_i^{(E)} \in [0, 1]$ denote the normalized life assessment, maintenance strategy, and economic impact indicators for asset i , as obtained from Module II. Each indicator is mapped to one of five HI bands using fixed thresholds:

Band	Threshold	Risk Midpoint
VL	≤ 0.20	0.10
L	$0.20 < x \leq 0.40$	0.30
M	$0.40 < x \leq 0.60$	0.50
H	$0.60 < x \leq 0.80$	0.70
VH	> 0.80	0.90

The corresponding risk vector is $\mathbf{R}_i = (R_i^{(L)}, R_i^{(M)}, R_i^{(E)})$.

The objective is to determine risk weights $\mathbf{w} = (w_L, w_M, w_E)^\top$ satisfying $\sum_j w_j = 1$ and $w_j \geq 0$, and a risk-aversion parameter $\lambda > 0$. For each maintenance action a with normalized cost c_a and risk-reduction vector $\Delta_a = (\Delta_L, \Delta_M, \Delta_E)^\top$ (positive values indicate reduction), the expected future risk is

$$R_i^{\text{fut}}(a) = \mathbf{w}^\top \text{diag}(1 - \Delta_a) \mathbf{R}_i \quad (22)$$

The decision score for action a on asset i is defined as

$$S(i, a) = c_a + \lambda \cdot R_i^{\text{fut}}(a) \quad (23)$$

Given the overall health band $\mathcal{A}(s_i)$ of asset i , the optimal action is selected from the set of engineering-feasible actions s_i :

$$\pi^*(i) = \arg \min_{a \in \mathcal{A}(s_i)} S(i, a) \quad (24)$$

Table 3. Maintenance actions, costs, and risk-reduction effects (Module 3).

ID	Action	Cost	ΔR_L	ΔR_M	ΔR_E
0	Do Nothing	0.00	−0.05	−0.05	0.00
1	Routine Maintenance	0.20	0.15	0.20	0.00
2	Diagnostics	0.30	0.20	0.25	0.10
3	Minor Repair	0.40	0.40	0.30	0.15
4	Derating	0.25	0.20	0.10	0.40
5	Partial Rebuild	0.60	0.60	0.50	0.30
6	Full Replacement	1.00	0.90	0.90	0.50

Note: positive values indicate risk reduction; negative values in “Do Nothing” reflect expected natural deterioration.

2.3.2. Meta-Optimization of Policy Parameters

These hyper-parameters (\mathbf{w}, λ) are calculated by minimizing a composite fleet-level objective:

$$(\mathbf{w}^*, \lambda^*) = \arg \min_{\mathbf{w}, \lambda} [U + \alpha O + \beta C], \quad (25)$$

where

- U : fraction of high-risk (H or VH) assets receiving insufficiently aggressive actions (intensity < 3),
- O : fraction of low-risk (VL or L) assets receiving overly aggressive actions (intensity 4),
- C : fleet-average-normalized maintenance cost,
- $\alpha = 2.0$ and $\beta = 0.5$ are fixed penalty weights. These parameters are design choices introduced to shape the optimization behavior and ensure stable, interpretable maintenance decisions. They are not empirically calibrated and remain configurable to reflect different risk–cost trade-offs during deployment.

A random search is performed over the bounded four-dimensional parameter space. Due to the low dimensionality and smoothness of the objective, convergence is rapid and robust.

2.3.3. Final Decision Layer and Prioritization

Each asset is assigned a total current risk

$$R_i^{\text{total}} = \mathbf{w}^{*\top} \mathbf{R}_i \quad (26)$$

and a raw decision score equal to the cost of the selected action $\pi^*(i)$ plus $\lambda^* \times R_i^{\text{fut}}(\pi^*(i))$. This raw score is min–max normalized to the interval $[0, 1]$, and the final maintenance priority is assigned according to the following bands:

Normalized Score	Priority Level
≥ 0.80	Immediate
0.40–0.80	Planned
< 0.40	Monitor

2.3.4. Contributions and Practical Impact

Module III delivers a practical, deployable alternative to traditional RL approaches by eliminating the need for transition probabilities or hand-crafted rewards. The resulting policy is fully interpretable, automatically respects engineering constraints, and produces clear, ranked maintenance priorities suitable for direct integration into utility work-planning systems. The entire optimization can be re-executed instantly whenever new condition data becomes available, making the framework highly adaptive in operational environments.

The proposed framework has been developed with operational use in mind by transmission and distribution system operators (TSOs/DOS) utilizing historical and planning information that is readily available in most utility asset management systems. Inputs include asset age, install date, historical faults, maintenance record, network criticality, customer impact factors, and cost to replace or repair. The methodology does not require online monitoring or extensive failure mode definitions to function efficiently.

If there are no exact measurements (i.e., real-time loading of transformers in distribution networks), estimated or proxy indicators may be used based on assumptions made during planning phases, customer load profiles, or engineering judgment. These inputs are considered approximations of operating conditions, rather than actual measures. This approach is similar to how Utilities currently operate. Therefore, the framework can be utilized at various levels of data maturity among TSOs and DSOs.

The framework's output will be an asset prioritization based on relative risk factors, along with recommended maintenance actions, rather than explicit failure predictions. The intent of the prioritized results is to provide utilities with assistance in their asset management processes by pointing out assets that should receive greater attention in situations where resource constraints exist. As utilities acquire more historical data, they can utilize the modules of the framework individually to begin the process of adopting it, beginning with minimal amounts of historical data and eventually including additional indicators.

3. Case Study: Implementation of the Methodologies in Asset Management Tool

3.1. Data and Pre-Processing

The study utilizes a dataset of 100 transformers, which was created by combining publicly available data with physics-based synthetic models of the data. This is intended to create a dataset that is representative of all three asset management dimensions, life assessment (LA), maintenance strategy (MS), and economic impact (EI), and is also realistic, comprehensive, and internally consistent.

3.1.1. Data Generation and Sources

A synthetic dataset was created using data from three sources. The first one was the ETDataset [26], which offers high-resolution transformer telemetry, including load components (HUFL, MUFL, LUFL) and oil temperature. These measurements were taken to determine realistic annual energy output, load behavior, and thermal aging indicators. The

second source was a Colombian transformer asset dataset [27], containing real attributes such as rated power, number of customers, energy not supplied (EENS), failure/burn counts, and defecting exposure (DDT). The third source is the Spanish DSO, which has real values for transformer age, contracted power, key customers, historical faults, severity, and duration of defects.

As the three datasets did not perfectly align, they were harmonized through unit standardization, attribute renaming, and inference of missing temporal fields (installation and manufacturing year inferred from age). A stratified sampling approach was applied to ensure balanced representation across small-, medium-, and large-impact transformers. Load and temperature time-series from the ETDataset were first normalized and then scaled to each transformer's rated power to obtain realistic estimates of annual energy output and a corresponding 12-month load profile. Using these loader-derived characteristics together with age, defect severity, and customer information, we calculated the required condition, maintenance, and economic indicators. These included failure probability, Health Index, remaining useful life, MTBF, MTTR, cost of failure, and Risk Index. All indicators were computed using established engineering formulations based on thermal aging behavior, reliability models, and value-of-lost-load (VOLL)-based economic estimation. The result is a single file containing all required physical, operational, maintenance, and economic attributes. The dataset is complete and contains no missing values, so no further cleaning or preprocessing is required.

3.1.2. Dimensional Structure, Indicators, and Interpretation

The three main dimensions used in advanced asset management frameworks have been identified for analysis to determine how the variables relate to each other: life assessment includes aging, condition, and criticality; maintenance strategy indicates reliability behavior and maintenance requirements; and economic impact reflects customer exposure, energy value, and financial consequences of failure. Table 4 provides an overview of the variable indicators along with definitions of each indicator.

3.2. Data Imputation for Power-System Assets

To assess the performance of the proposed imputation method, the entire synthetic transformer dataset result from the dataset preparation step was used to gauge the imputation outcomes. The synthetic dataset has no missing data (there are no missing values), which also means that it is appropriate in terms of the objective evaluation of the adequacy of the imputation method. A total of 5% of synthetic data was randomly masked by missing values to obtain simulated field conditions in the real world, presenting sensor dropout, slow inspection, or incomplete logging, and thereafter, this artificially corrupted dataset has been subjected to the Module I pipeline.

The original dataset did not have any missing values, and, hence, the imputed measurements could be directly measured by evaluating the imputed values in direct comparison to the actual ones. The masked dataset was then passed through the data imputation techniques, including missingness detection, robust scaling, and all imputation models (mean, median, KNN, MICE, MissForest, GAN, and hybrid approaches). Table 5 summarizes the performance of all methods in terms of mean absolute error (MAE) and the Kolmogorov–Smirnov (KS) statistic. MICE achieved the lowest MAE (520.46), indicating highly accurate point-wise reconstruction, while MissForest obtained the best KS value (0.3889), showing strong distribution preservation. A combined metric (70% MAE + 30% KS) confirmed MICE as the overall best-performing model for transformer condition imputation.

Table 4. Dimensions, indicators, and interpretation.

Dimension	Indicator	Interpretation
Life Assessment (LA)	H1_age_years	Physical aging and insulation deterioration.
	H2_failure_probability	Estimated failure likelihood based on age, load stress, and defect severity.
	H3_criticality	Severity of network and customer impact in case of failure.
	Health_Index	Composite condition score combining aging, defects, and reliability indicators.
	Remaining_Useful_Life_years	Expected remaining operational lifetime under current conditions.
	Thermal_Aging_Factor	Hotspot-driven insulation aging acceleration derived from load profile.
Maintenance Strategy (MS)	MTBF_years	Mean time between failures; indicator of reliability performance.
	MTTR_hours	Estimated average time required to repair a failure event.
	Maintenance_Cost_per_Year_EUR	Annual expenditure for routine and corrective maintenance activities.
Economic Impact (EI)	Customers_affected	Number of end-users impacted by outages of this transformer.
	H4_Energy_kWh_per_year	Yearly delivered energy derived from ETDataset-based load scaling.
	EENS_kWh_case_of_failure	Expected unsupplied energy in the event of a failure.
	VOLL_EUR_per_kWh	Economic value of lost load depending on customer category.
	H2_Cost_of_failure_EUR	Total cost of failure computed from EENS, VOLL, and replacement cost.
	Annual_OPEX_EUR	Yearly operational expenditure, including maintenance and energy losses.
	Risk_Index	Composite risk metric combining failure probability, criticality, and economic consequence.

Table 5. MAE and KS statistics for all imputation methods (lower is better).

Method	MAE	KS
Mean Imputation	2257.05	0.8333
Median Imputation	2793.12	0.6111
KNN (k = 5)	1498.63	0.5000
MICE	520.46	0.4444
MissForest	1341.09	0.3889
GAN	1893.73	0.5612
Hybrid KNN–GAN	1592.84	0.5000
Hybrid MICE–GAN	1493.23	0.6167

Note: Values in bold indicate the lowest values among the compared methods.

Figures 2 illustrate the imputation quality for three representative indicators, failure probability, annual energy throughput, and thermal aging factor, by comparing the true values with predictions from MICE, alongside flat estimates from mean and median imputation. Across all indicators, MICE closely matches the true engineering behavior, whereas this behavior is completely lost under traditional mean/median filling. This confirms that the proposed Module I framework produces numerically precise and engineering-consistent imputations, providing a reliable foundation for clustering, weighting, and maintenance optimization in subsequent modules. The last imputed data serve as the validated base for further clustering, weighting, calculation of the health index, and optimization of maintenance.

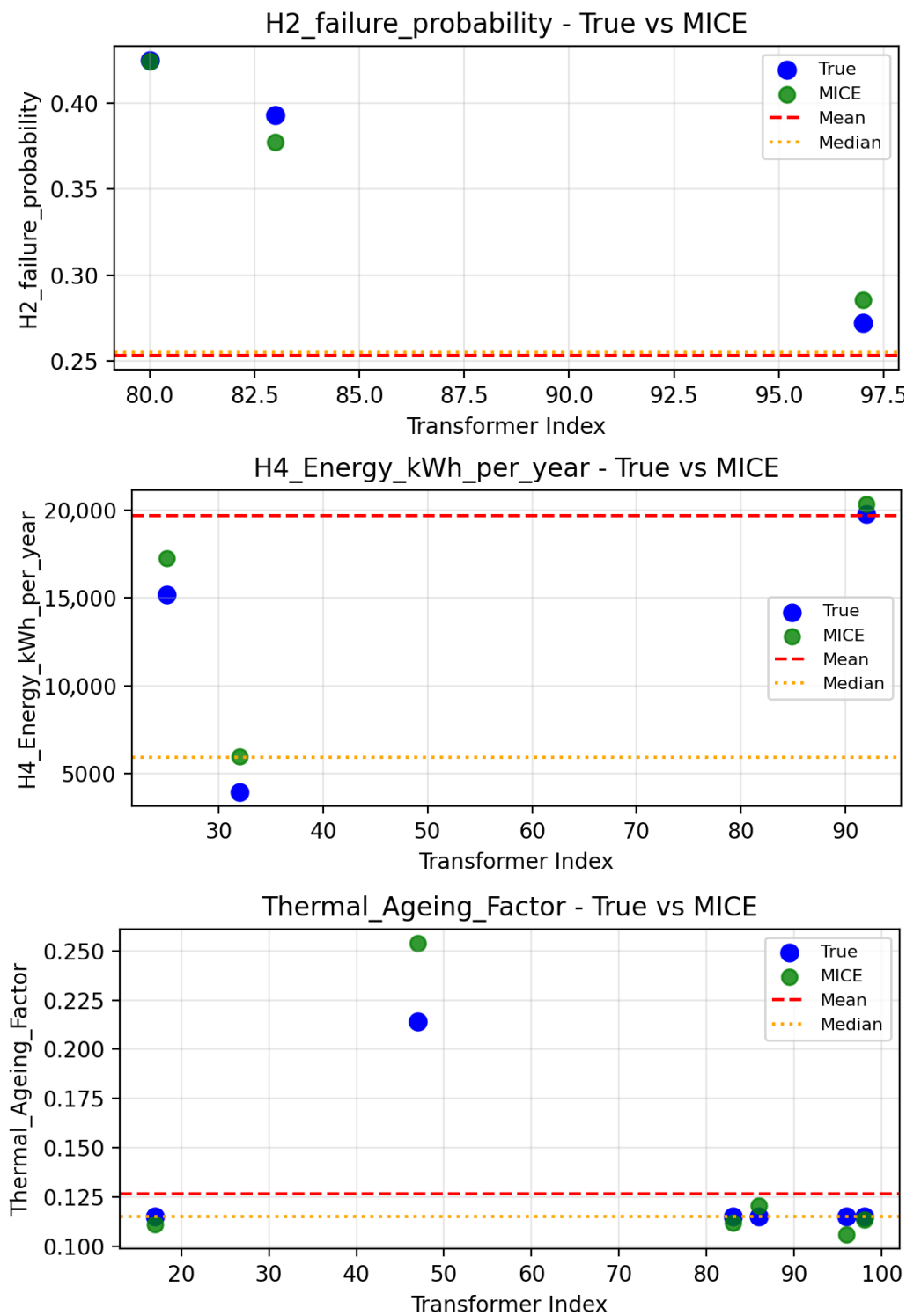


Figure 2. Comparison of ground-truth values and MICE-imputed values.

This case study confirms that the proposed imputation framework can accurately substitute missing data in transformer condition data, maintaining both numerical fidelity and distributional structure. By adding controlled missingness to a complete knowledge synthetic dataset, the evaluation separates the imputation error, not due to external noise or labeling uncertainty, but rather due to traditional heuristic imputation methods (e.g., mean or median replacement) that are common to most types of missing-data problems in general, which will always shift the properties of the original data and cause greater difficulty in interpreting the results of the further step in model development. Conversely,

a comparison of the six advanced imputation methods, based on automated comparisons made in this study, revealed that MICE was the most suitable method with respect to the minimum piecewise statistics errors and the natural statistical performance on this specific dataset. The results of these studies show that a data-driven, application-aware method selection strategy is more essential than a single, fixed imputation strategy that can be applied to all the datasets. The superior quality of the imputed data, which is validated by Module I, can now serve as a strong and sound foundation for the next modules, providing a high-quality and structurally consistent set of data for the downstream processes of clustering, weight assignment, and policy optimization.

3.3. Multi-Dimensional Weighting Results and Health-Based Recommendations

After completing the imputation, the optimized weighting framework was applied to the final transformer dataset, and the resulting weights were calculated for each of the three main factors: life assessment (LA), maintenance strategy (MS), and economic impact (EI). Each of the final weights is an average of the two top-performing methods for each factor, ensuring a robust feature weight.

Table 6 summarizes the indicators and their optimized weights for the three dimensions. In the LA dimension, the highest weight is given to age in years (0.51), followed by the thermal aging factor (0.18) and energy (0.21). The smallest weight, but still relevant, is the failure probability (0.11). For MS, the criticality (0.46) and the cost of maintenance (0.38) were ranked the first two weights in the optimization. Then, there was a less significant but still notable influence of MTBF years (0.07), Health Index (0.04), and remaining useful life years (0.05) in the optimization. In EI, the optimization process clearly determined that the most important weight is related to the cost of failure (0.45). Secondary weights were assigned to the annual OPEX (0.30) and to the customers affected (0.20). Finally, the EENS kWh during the time of failure (0.05) had the lowest weight.

These patterns are both statistically and operationally meaningful. A high weight indicates that an indicator is consistently informative for separating transformers into distinct health clusters; in practice, it means that small changes in that variable significantly affect the final HI, and, therefore, the maintenance recommendation. For example, the 0.51 weight assigned to age combined with the 0.18 weight assigned to thermal aging factor indicates that aging and thermal stress have the greatest influence on the degradation of all transformers across the entire fleet. It is common practice among electrical engineers to monitor both of these factors to determine whether a transformer is nearing the end of its useful life. Similarly, in the maintenance strategy (MS) dimension, the emphasis placed on criticality (0.46), maintenance cost (0.38) and mean time between failure (MTBF) (0.07), indicates that those transformers which are both critical to the operation of the electric system and costly to repair will be identified as being at a greater risk of failing, regardless of the MTBF of the transformer. This is similar to the manner in which asset managers identify and prioritize equipment that generates multiple trouble calls or affects important customer loads. In the economic impact (EI) dimension, the extremely high weight assigned to the cost of failure (0.65) indicates that transformers that could cause significant financial or social loss if they fail will have a substantially better HI score than those transformers with a lower potential loss of service. This is similar to how utilities would rationally wish to prioritize risk in a real-world environment. By contrast, earlier heuristic or expert-based weightings tended to distribute weights more evenly or based on intuition, which risked underestimating high-impact transformers or overemphasizing less informative indicators. The data-driven weights, therefore, not only maximize clustering quality, as defined in Module II, but also align closely with the way utilities would rationally want to prioritize risk. These optimized weights were used to calculate a total indicator (HI) for each

transformer and assign this indicator to a specific rank on a fixed, five-band condition rating scale (A–E) as specified by the method. The results for this data set of 100 transformers indicated that 44 ranked at A (very low), 40 ranked at B (low), 15 ranked at C (moderate), 1 ranked at D (high) and none ranked at the extreme E (very high), which is consistent with a synthetic but realistic fleet where only a few units are near critical condition. The distribution of ranks indicates that the framework is not overly conservative; however, it identifies most of the transformers as low risk, while a small but important subset is escalated for closer attention. Table 7 presents an example of 12 transformers, representing the entire spectrum from very low to high risk, and includes the three-dimensional scores, the final total indicator (HI), the categorical rank, and the recommended action for each transformer.

Table 6. Dimensions, indicators, and final optimized weights.

Dimension	Indicator	Description	Optimized Weight
Life Assessment (LA)	H1_Age_years	Physical aging	0.51
	H4_Energy	Annual delivered energy	0.21
	H2_failure_probability	Estimated failure likelihood	0.11
	Thermal_Aging_Factor	Hotspot thermal degradation	0.18
Maintenance Strategy (MS)	MTBF_years	Mean time between failures	0.07
	Maintenance_Cost_per_Year_EUR	Annual maintenance cost	0.38
	H3_criticality	Impact severity in case of failure	0.46
	Health_Index	Condition-related reliability score	0.04
	Remaining_Useful_Life_years	Expected remaining life	0.05
Economic Impact (EI)	H2_Cost_of_failure_EUR	Total cost of failure	0.65
	EENS_kWh_case_of_failure	Expected unsupplied energy	0.05
	Annual_OPEX_EUR	Yearly operating expenditure	0.20
	Customers_affected	Number of affected consumers	0.10

Table 7. Representative asset-level results across three dimensions and final Health Index.

Asset	LA	MS	EI	TI	Rank	Level	Recommended Action
T0090	0.00	0.07	0.00	0.03	A	Very Low	Normal time-based maintenance
T0078	0.01	0.06	0.00	0.03	A	Very Low	Normal time-based maintenance
T0036	0.08	0.08	0.00	0.05	A	Very Low	Normal time-based maintenance
T0015	0.13	0.09	0.02	0.08	A	Very Low	Normal time-based maintenance
T0024	0.50	0.10	0.01	0.20	B	Low	Increase detection and correct minor issues
T0016	0.12	0.50	0.03	0.22	B	Low	Increase detection and correct minor issues
T0091	0.41	0.70	0.05	0.39	B	Low	Increase detection and correct minor issues
T0012	0.16	0.93	0.30	0.47	C	Moderate	Address key contributing factors
T0048	0.37	0.71	0.41	0.50	C	Moderate	Address key contributing factors
T0013	0.48	0.82	0.36	0.55	C	Moderate	Address key contributing factors
T0070	0.46	0.81	0.47	0.58	C	Moderate	Address key contributing factors
T0047	0.60	0.70	0.98	0.76	D	High	Plan repair, rebuild, or major maintenance

This subset clearly shows the combination of the three weighted dimensions in the total indicator (HI) and the corresponding maintenance recommendations. The first four assets (T0090–T0015) have similar scores (all very low) across LA, MS, and EI; therefore, their TI values are all less than 0.10, and they are classified as Rank A. These transformers are considered low-risk units that require only routine, time-based maintenance. The next set of assets in rank B (T0024, T0016, T0091) illustrate different ways to reach an elevated, but still relatively moderate level of risk: T0024 has a relatively high life-aging contribution (LA \approx 0.50) relative to other risk categories, but low levels of maintenance

and economic stresses; T0016 exhibits relatively strong maintenance-related concerns ($MS \approx 0.50$); and T0091 represents an asset that has both elevated maintenance burdens (0.70) and elevated life assessments (0.41), making it appropriate to recommend that it be monitored more frequently. Assets in the C category (T0012, T0048, T0013, T0070) have a consistent combination of elevated scores across at least two dimensions (primarily elevated MS values > 0.70 in conjunction with notable economic exposure) with HI values ranging from 0.47 through 0.58; therefore, these units would be appropriately recommended to mitigate the primary causes of concern prior to the risk escalating further. Finally, asset T0047 is representative of a high-risk transformer; while it has moderate levels of LA and MS (in the range of 0.60–0.70), its economic impact ($EI = 0.98$) is extreme, thereby elevating its HI to 0.76 and classifying it in Rank D. This type of transformer would be given priority in real utility operations for planned repairs, rebuilds, or major maintenance as a result of the combination of its age, reliability concerns, and potential failure consequences.

Using the actual optimal weightings along with the three dimension model, this case study supports the notion that the proposed methodology does not only provide a mathematical consistency but is also practically applicable, the highest weighted indicators will be those on which utilities will have to focus their efforts, and the calculated total indicator (HI), as well as the provided recommendations, closely align with what an experienced asset manager would consider in making a decision.

3.4. Policy Optimization for Maintenance Decision Support Using Multi-Objective Parameter Search

The outputs from Module II, life assessment (LA), maintenance strategy (MS), economic impact (EI), and the overall total indicator (TI) represent the technical and economic condition of each transformer. Module III builds directly on these results. Instead of using RL, maintenance decisions are generated through a multi-objective parameter search, where each possible maintenance action is evaluated in terms of the risk it reduces and the value it provides. The model evaluates each transformer by producing three types of risk reduction components R_{life} (impact on aging and deterioration), R_{maint} (impact on reliability and maintenance burden), and R_{econ} (impact on economic exposure if failure occurs). These three types of risk reduction are then added together to determine R_{total} , an overall measure of the benefits associated with performing a given maintenance action at the current time.

Based on these scores, the algorithm selects the best action and assigns a priority level using a normalized decision score. This transforms the condition indicators from Module II into clear, asset-specific maintenance recommendations that account for both risk and cost.

Table 8 shows how the maintenance requirements of the fleet progress logically in terms of transformer health. For example, the most healthy transformers (T0090, T0078, T0036, and T0015), all have small amounts of LA, MS, and EI; therefore, since no action would greatly improve the already good status of these transformers, they will be given equal low scores for risk reduction by life, by maintenance and by economy ($R_{\text{life}} = R_{\text{maint}} = R_{\text{econ}} = 0.1$). Hence, their total combined score ($R_{\text{total}} = 0.10$), demonstrates that they should be assigned to “Do Nothing”, but as a priority to monitor.

When conditions worsen, the model will adjust its recommendations based on the new condition. Assets such as T0024 and T0016 have shown some early signs of aging or excessive maintenance burden, evidenced by larger R_{life} and/or R_{maint} values; while they remain in the low-priority groups, it is clear that they should receive increasing attention in upcoming maintenance cycles. The most significant changes can be observed in transformers such as T0091, T0012, and T0048. These transformers are experiencing greater reliability or economic concerns, as evidenced by significantly higher aggregated scores (R_{total} , 0.40–0.50); therefore, the model recommends routine maintenance at a planned

priority level. This indicates that proactive intervention will lead to a reduction in their long-term risk.

Table 8. Asset-level results from maintenance policy optimization.

Asset	R_{life}	R_{maint}	R_{recon}	R_{total}	Best Action	Decision Score	Priority
T0090	0.10	0.10	0.10	0.10	Do Nothing	0.00	Monitor
T0078	0.10	0.10	0.10	0.10	Do Nothing	0.00	Monitor
T0036	0.10	0.10	0.10	0.10	Do Nothing	0.00	Monitor
T0015	0.10	0.10	0.10	0.10	Do Nothing	0.00	Monitor
T0024	0.50	0.10	0.10	0.289	Do Nothing	0.362	Monitor
T0016	0.10	0.50	0.10	0.237	Do Nothing	0.263	Monitor
T0091	0.50	0.70	0.10	0.495	Do Nothing	0.757	Planned
T0012	0.10	0.90	0.30	0.411	Routine Maintenance	0.639	Planned
T0048	0.30	0.70	0.50	0.474	Routine Maintenance	0.754	Planned
T0013	0.50	0.90	0.30	0.600	Routine Maintenance	0.933	Immediate
T0070	0.50	0.90	0.50	0.637	Routine Maintenance	1.000	Immediate
T0047	0.50	0.70	0.90	0.642	Derating	0.987	Immediate

In contrast, transformers such as T0013, T0070, and T0047 demonstrate how the multi-objective approach can identify transformers for which delay is unacceptable. All three of these units exhibit significant degradation across multiple measures, specifically T0047, which carries an extremely high economic risk. Their R_{total} values (0.60–0.64) and high decision scores (0.93–1.00) indicate that additional maintenance provides significant benefit. As a result, T0013 and T0070 were placed into the immediate priority category with urgent maintenance recommendations, while T0047 was placed into derating, a strategy used to reduce loadings and decrease the risk of major failure. These decisions align closely with real-world asset management practice, where the most vulnerable assets receive attention first.

4. Scope and Limitations

The focus of this research is on developing a methodology and validating it; thus, to allow for an unbiased (objective) comparison among various imputation and weighting strategies, a synthetic data set was used to validate the framework. However, field testing on actual utility data sets with known and structured missing data will need to be accomplished in future research in order to assess the performance of the framework under operational conditions.

In contrast to the explicit modeling of individual failure mechanisms, the framework uses aggregate risk indicators, which are consistent with what is typically available in utility data. The choice of thresholds, weights, and parameters for determining action impact was made to ensure a stable and transparent operation of the framework, and is intended to be configurable by users in their operational environment. The framework has been designed to augment, rather than replace, existing engineering judgment and asset management processes.

5. Conclusions

This research has successfully designed and proven an all-encompassing framework utilizing AI and ML to enhance asset management in the electric power system, addressing the fundamental limitations of traditional approaches through systematic innovation across data processing, feature characterization, and optimization domains. The three-module architecture represents a significant advancement, providing a mathematically consistent

and practically applicable solution that bridges the gap between theoretical models and real utility environments.

Data imputation enables advanced treatment of missing data and structural zeroes via an enhanced imputation method, creating a robust foundation for reliable asset health assessment. Validations have demonstrated that ML-based approaches significantly outperform traditional approaches in terms of both numerical accuracy and distributional structure, which are critical requirements for accurate condition monitoring. This multi-method weighting framework represents a paradigm shift from manual subjective rule-based approaches to an objective methodically based approach to determine the relative importance of features. Through the combination of methods, the robustness and clarity of results are increased. By providing an objective measure, it also enables the production of asset health indices that truly represent the current state of assets, thereby avoiding bias in assessments through operator or subjective interpretation. This ultimately enhances the reliability and consistency of all maintenance decisions and overall asset evaluations across different times and operational conditions. The optimization module expands the action space, and a meta-heuristic multi-objective framework addresses the complexity of real-world maintenance decision-making by balancing competing objectives. The methodology has the ability to produce ranked maintenance priorities with those made by experienced asset managers' decision-making, while providing mathematical consistency, which represents a significant practical advancement.

Validation is provided using an extensive synthetic dataset of 100 transformers, generated from publicly available information combined with physics-based modeling. These results demonstrate the effectiveness of the methodology in realistic operating conditions. Results from case studies demonstrate that this methodology yields tangible improvements in the development of maintenance policies, risk assessments, and resource allocation.

As future studies will be focused on applying the framework to emerging assets (e.g., renewable energy infrastructures and battery storage), and in addition to this, to develop and apply real time data for dynamic optimizations as well as to investigate the possibilities of federated learning for the utilization of a collective industry knowledge base, at the same time protecting the confidentiality of the individual data. In addition, as part of the standardization process, future studies should also focus on developing standardized interfaces for connecting the proposed system to existing utility management systems, which would contribute to broader use and, thus, increased influence on current industry practices.

Ultimately, this work will establish a new benchmark for intelligent asset management within power systems, providing transmission and distribution system operators with the advanced tools necessary to address the numerous challenges associated with operating modern grids at high levels of reliability and cost-effectiveness. The open-source nature ensures broad accessibility and continued community-driven development, as well as the advancement of related R&D activities in power systems, and ultimately contributes to enhancing the overall resilience and reliability of electrical power infrastructures in our society.

Author Contributions: Conceptualization, G.L.R. and M.A.S.-B.; methodology, G.L.R., M.A.S.-B., and P.C.-B.; coding and machine learning, G.L.R.; validation, G.L.R. and P.C.-B.; formal analysis, G.L.R.; investigation, G.L.R., M.A.S.-B. and L.B.T.; writing—original draft preparation, G.L.R. and M.A.S.-B.; writing—review and editing, G.L.R., M.A.S.-B., L.B.T., and P.C.-B.; visualization, G.L.R., M.A.S.-B., L.B.T., and P.C.-B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The study is based on open-source datasets. Using these publicly available data as a foundation, synthetic data were generated to support the analysis presented in this work. No proprietary or confidential data was used.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wang, K.; Li, Y.; Wang, X.; Zhao, Z.; Yang, N.; Yu, S.; Wang, Y.; Huang, Z.; Yu, T. Full Life Cycle Management of Power System Integrated With Renewable Energy: Concepts, Developments and Perspectives. *Front. Energy Res.* **2021**, *9*, 680355. <https://doi.org/10.3389/fenrg.2021.680355>.
2. Lund, H. Renewable energy strategies for sustainable development. *Energy* **2007**, *32*, 912–919. <https://doi.org/10.1016/j.energy.2006.10.017>.
3. Moleda, M.; Małysiak-Mrozek, B.; Ding, W.; Sunderam, V.; Mrozek, D. From Corrective to Predictive Maintenance—A Review of Maintenance Approaches for the Power Industry. *Sensors* **2023**, *23*, 5970. <https://doi.org/10.3390/s23135970>.
4. Brown, T.; Hörsch, J.; Schlachtberger, D. PyPSA: Python for power system analysis. *arXiv* **2017**, arXiv:1707.09913.
5. Chassin, D.P.; Schneider, K.; Gerkenmeyer, C. GridLAB-D: An open-source power systems modeling and simulation environment. In Proceedings of the 2008 IEEE/PES Transmission and Distribution Conference and Exposition, Chicago, IL, USA, 21–24 April 2008; pp. 1–5.
6. Cui, Y.; Bangalore, P.; Bertling Tjernberg, L. A fault detection framework using recurrent neural networks for condition monitoring of wind turbines. *Wind Energy* **2021**, *24*, 1249–1262.
7. Rajora, G.L.; Sanz-Bobi, M.A.; Domingo, C.M.; Bertling Tjernberg, L. An Open-Source Tool-Box for Asset Management Based on the Asset Condition for the Power System. *IEEE Access* **2025**, *13*, 49174–49186. <https://doi.org/10.1109/ACCESS.2025.3551663>.
8. Zhang, Y.; Huang, T.; Bompard, E.F. Big data analytics in smart grids: a review. *Energy Inform.* **2018**, *1*, 8.
9. Rajora, G.L.; Sanz-Bobi, M.A.; Domingo, C. Application of Machine Learning Methods for Asset Management on Power Distribution Networks. *Emerg. Sci. J.* **2022**, *6*, 905–920. <https://doi.org/10.28991/esj-2022-06-04-017>.
10. Strielkowski, W.; Vlasov, A.; Selivanov, K.; Muraviev, K.; Shakhnov, V. Prospects and Challenges of the Machine Learning and Data-Driven Methods for the Predictive Analysis of Power Systems: A Review. *Energies* **2023**, *16*, 4025. <https://doi.org/10.3390/en16104025>.
11. Aminifar, F.; Abedini, M.; Amraee, T.; Jafarian, P.; Samimi, M.H.; Shahidehpour, M. A review of power system protection and asset management with machine learning techniques. *Energy Syst.* **2021**, *13*, 855–892. <https://doi.org/10.1007/s12667-021-00448-6>.
12. Alhamrouni, I.; Kahar, N.H.A.; Salem, M.; Swadi, M.; Zahroui, Y.; Kadhim, D.J.; Mohamed, F.A.; Nazari, M.A. A Comprehensive Review on the Role of Artificial Intelligence in Power System Stability, Control, and Protection: Insights and Future Directions. *Appl. Sci.* **2024**, *14*, 6214. <https://doi.org/10.3390/app14146214>.
13. Akhtar, S.; Adeel, M.; Iqbal, M.; Namoun, A.; Tufail, A.; Kim, K.H. Deep learning methods utilization in electric power systems. *Energy Rep.* **2023**, *10*, 2138–2151. <https://doi.org/10.1016/j.egy.2023.09.028>.
14. Little, R.J.; Rubin, D.B. *Statistical Analysis with Missing Data*; John Wiley & Sons: Hoboken, NJ, USA, 2019. <https://doi.org/10.1002/9781119482260>
15. Wongoutong, C. The impact of neglecting feature scaling in k-means clustering. *PLoS ONE* **2024**, *19*, e0310839. <https://doi.org/10.1371/journal.pone.0310839>.
16. Van Buuren, S.; Groothuis-Oudshoorn, K. mice: Multivariate imputation by chained equations in R. *J. Stat. Softw.* **2011**, *45*, 1–67.
17. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*. <https://doi.org/10.48550/arXiv.1406.2661>.
18. Massey, F.J., Jr. The Kolmogorov–Smirnov test for goodness of fit. *J. Am. Stat. Assoc.* **1951**, *46*, 68–78.
19. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.
20. Jolliffe, I. Principal component analysis. In *International Encyclopedia of Statistical Science*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 1094–1096. https://doi.org/10.1007/978-3-642-04898-2_455
21. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507.
22. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **2017**, *30*. <https://doi.org/10.48550/arXiv.1705.07874>
23. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.
24. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65.
25. Davies, D.L.; Bouldin, D.W. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *PAMI-1*, 224–227.

26. Zhou, H.; Zhang, Q.; Zhang, J.; Yuan, C. ETDataset: Electricity Transformer Temperature and Load Dataset. 2021. Available online: <https://github.com/zhouhaoyi/ETDataset> (accessed on 10 February 2025).
27. Bravo, D.; Alvarez, L.; Lozano, C. *Mendeley Data*, Version 4; Dataset of Distribution Transformers at Cauca Department: Colombia, Mexico, 2021. <https://doi.org/10.17632/yzyj46xpmy.4>

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.