Journal of Information & Systems Management

DLINE JOURNALS

ISSN: 2230 – 8776

JISM 2026: 16 (4)
https://doi.org/10.6025/jism/2026/16/4/190-212

Sensitivity Analysis and Model-Based Interpretability of PWR Fuel-Assembly Neutronics Simulation Parameters: A 39-Variable Statistical and Predictive Assessment

Darya-Anna Kaybiyaynen
Kazan National Research Technological University
Kazan, Russia
dariya.anna@gmail.com

ABSTRACT

The integration of artificial intelligence and machine learning (AI/ML) in nuclear engineering offers significant potential for reactor-state prediction and digital twins; however, the "black-box" nature of data driven models limits their adoption in safety-critical, regulatory environments. This study addresses the critical gap in systematic sensitivity analysis and model-based interpretability for high-fidelity Pressurized Water Reactor (PWR) fuel-assembly neutronics simulations. We present a comprehensive 39-variable statistical and predictive assessment framework applied to a dataset of 24,360 simulation records, targeting the infinite multiplication factor (K∞) and the pin power peaking factor (PPPF). By integrating classical statistical methods with modern explainable AI (XAI) techniques including Multiple Linear Regression, HistGradientBoosting, Permutation Importance, SHapley Additive exPlanations (SHAP), and Partial Dependence Plots (PDP) we evaluate how material and geometric parameters propagate to key neutronics responses. Results reveal distinct predictive structures: is predominantly linear (), whereas PPPF exhibits nonlinear characteristics best captured by ensemble methods (). Furthermore, a strong inverse correlation () and contrasting sensitivity directions between the two responses are identified, alongside response-specific predictive drivers. This research establishes a transparent, multi-tiered interpretability pipeline that bridges high fidelity neutronics simulation with physics informed insight, providing a foundational framework for multi objective reactor optimization and large scale uncertainty quantification.

Keywords: Pressurized Water Reactor (PWR), Neutronics Simulation, Sensitivity Analysis, Explainable AI (XAI), Surrogate Modeling; SHAP, Pin Power Peaking Factor (PPPF), Multi-objective Optimization

Received: 18 February 2026, Revised 5 June 2026, Accepted 21 June 2026

Copyright: DLINE

1. Introduction

Artificial intelligence (AI) and machine learning (ML) have been increasingly applied across nuclear engineering for reactor state prediction, transient diagnosis, high dimensional physics modeling, condition monitor

-ing control, and digital twins. Early neural-network applications demonstrated the feasibility of identifying nuclear power plant transients directly from plant variables [1], while subsequent work developed dynamic neural network ensembles capable of jointly detecting, classifying, and predicting reactor transients. More recent studies have expanded these capabilities using deep learning; for example, transfer learning has been investigated to improve nuclear power plant fault diagnosis under changing operating conditions [2], and comparative evaluations of deep neural architectures have demonstrated the potential of convolutional–recurrent models for equipment fault diagnosis under limited and noisy data [3]. These developments complement nuclear-specific studies that combine simulation data with deep learning and uncertainty quantification [4] analyze reactor simulation outputs and uncertainty using data driven methods [5], forecast loss of coolant-accident transients [6], and approximate high-dimensional full core BWR neutronics [7].

2. Background

Beyond prediction and diagnosis, AI is increasingly being integrated with physics based models. Gong et al. combined reduced order modeling and machine learning to construct a physics informed digital twin capable of rapidly reconstructing high dimensional reactor neutron-flux and power fields [7], while Song et al. developed an online machine learning calibration strategy to improve the accuracy of nuclear power plant digital twins while maintaining real time capability [8]. AI applications have also progressed toward autonomous decision making, including deep reinforcement learning for microreactor load-following control [10] and multistep criticality search and power shaping [11]. Variational digital twins provide another means of combining physical models and observations for inference [12], while symbolic learning has been explored to improve interpretability of AI models for advanced energy systems [13]. The breadth of this research is reflected in recent surveys by Jendoubi and Asad, who reviewed AI applications spanning nuclear plant monitoring, diagnostics, prediction, and operation [14], and by Ejigu et al., who examined the integration of AI and big-data methods for nuclear power plant control [15]. The comprehensive discussion of this literature demonstrates substantial progress in using AI to predict, reconstruct, diagnose, and control nuclear systems. However, most existing approaches learn a direct input output mapping, reduced-order representation, diagnostic model, or control policy from the available data [16].

The rapid evolution in machine learning has catalyzed its widespread adoption across diverse domains over the past decade. Data-driven and/or physics-informed ML methodologies have offered enhanced solutions with superior efficiency for computationally intensive, complex problems, presenting significant appeal for reactor design [17]. Numerical optimization has been widely used to solve design engineering problems in recent years. However, due to the large number of numerical simulations required, the optimization process is often computationally expensive. In order to accelerate the optimization process, strategies based on surrogate models are frequently employed.

Reactor physics is the study of neutron properties, focusing on the use of models to examine the interactions between neutrons and materials in nuclear reactors. Artificial intelligence has made significant contributions to reactor physics, such as in operational simulations, safety design, real-time monitoring, core management, and maintenance [18]. Zhang presents a comprehensive review of AI approaches in reactor physics, especially considering the category of Machine Learning (ML, which we also refer to as AI/ML), with the aim of describing the application scenarios, frontier topics, unsolved challenges, and future research directions [18].

Traditional light water reactor simulations are usually either high fidelity, requiring hundreds of node-hours, or low fidelity, requiring only seconds to run on a common workstation. It is desirable to combine the positive aspects of both of these simulation types while minimizing their associated negative costs [19]. High-fidelity multiphysics simulation is essential for resolving the strongly coupled phenomena governing nuclear reactor behavior, but coupling detailed physics models across multiple spatial and temporal scales remains computationally demanding [20, 21]. Computational fluid dynamics (CFD), for example, provides detailed descriptions of coolant transport, mixing, and flow heterogeneity that are difficult to obtain from lower-order thermal-hydraulic models, but at substantially greater computational cost [22, 23]. Consequently, generating sufficiently long and diverse high-fidelity CFD datasets for direct exploration of large inverse, optimization, or

uncertainty spaces can become computationally prohibitive.

Predicting fuel assembly bow in pressurized water reactors requires solving tightly coupled fluid-structure interaction problems, whose direct simulations can be computationally prohibitive, making large-scale uncertainty quantification very challenging. Frameworks that couple Gaussian process surrogate models representing distinct physical solvers have been introduced to enable rigorous uncertainty quantification in coupled multiphysics systems [24].

Amidst the growing global emphasis on nuclear safety, the integrity of nuclear reactor systems has garnered increased attention. The rapid development of artificial intelligence technology has provided immense opportunities to enhance the safety and economy of nuclear energy. However, data-driven deep learning techniques often lack interpretability, which hinders their applicability in the nuclear energy sector. To address this problem, hybrid data-driven and knowledge-driven artificial intelligence models based on physics-informed neural networks have been proposed to accurately compute the neutron flux distribution inside a nuclear reactor core [25, 26].

Despite these advances, a critical gap remains in the systematic sensitivity analysis and model-based interpretability of the numerous input parameters that govern pressurized water reactor (PWR) fuel-assembly neutronics simulations. Understanding how individual and combined variations in material, geometric, and nuclear data parameters propagate to key outputs, such as eigenvalue, power distribution, and reaction rates, is essential for robust design, uncertainty quantification, and regulatory confidence. The present work addresses this need through a comprehensive 39-variable statistical and predictive assessment that integrates sensitivity analysis with interpretable machine learning models, thereby bridging high-fidelity neutronics simulation with transparent, physics-informed insight.

2.1 Research Statement

Despite the rapid integration of artificial intelligence and machine learning (AI/ML) in nuclear engineering for reactor state prediction and digital twins, a critical gap remains in the systematic sensitivity analysis and model based interpretability of high fidelity Pressurized Water Reactor (PWR) fuel assembly neutronics simulations. Data driven deep learning techniques often operate as “black boxes,” lacking the interpretability required for regulatory confidence, robust design, and uncertainty quantification. This study addresses this challenge by presenting a comprehensive 39-variable statistical and predictive assessment framework. By bridging high fidelity neutronics simulation data with transparent, physics informed interpretability techniques, this research evaluates how material, geometric, and nuclear data parameters propagate to key outputs provisionally identified as the infinite multiplication factor ($K\infty$) and the pin power peaking factor (PPPF). The study reveals distinct predictive structures and contrasting sensitivity directions for different neutronics responses, establishing a foundational, interpretable framework for multi-objective neutronics optimization and large scale uncertainty quantification.

2.2 Methodology

The methodology employs a multi-tiered analytical framework that integrates classical statistical methods with modern machine learning interpretability techniques to evaluate the input response relationships:

• Statistical Association Analysis: Pearson and Spearman correlations are utilized to identify marginal linear and monotonic relationships. Partial correlation analysis is applied to isolate conditional linear associations by controlling for the remaining 38 input variables, mitigating the effects of shared statistical information.

• Surrogate Modeling & Regression: Multiple Linear Regression (MLR) serves as the baseline to capture adjusted linear effects, accompanied by HC3 heteroskedasticity-robust 95% confidence intervals to assess coefficient stability. Ensemble methods (Extra Trees and HistGradientBoosting) are deployed to capture potential nonlinearities and interaction effects.

• Feature Importance & Attribution:

o *Permutation Importance* assesses predictive reliance by measuring the degradation in test-set R^2 upon random feature shuffling.

o *SHAP (SHapley Additive exPlanations)* provides an exact feature-attribution framework for the linear models, quantifying the magnitude of individual feature contributions relative to training-set means.

• Response Surface Visualisation: Partial Dependence Plots (PDP) and Individual Conditional Expectation (ICE) profiles are generated to examine average predicted response behaviours and individual prediction heterogeneity, allowing for the identification of nonlinear thresholds and interaction effects within the surrogate models.

2.3 Design or Architecture

The architecture of the study is designed as a sequential analytical pipeline that transitions from data exploration to predictive modeling, culminating in interpretable sensitivity mapping:

1. Data Ingestion & Partitioning Architecture: The dataset is structurally divided into a large scale training cohort for model fitting and an independent test cohort for unbiased generalization assessment. The feature space is mapped across 39 provisional input variables (enrichment related parameters) and 2 response variables (Y_1, Y_2).

Multi-Model Evaluation Pipeline: A comparative architecture evaluates linear versus non linear surrogate models independently for each response variable. This design determines the optimal modeling strategy per response (e.g., MLR for Y_1 vs. HistGradientBoosting for Y_2), acknowledging that a single modeling strategy is not universally appropriate for all neutronics responses.

2. Interpretability & Sensitivity Mapping Layer: A dedicated analytical layer applies model-agnostic (Permutation) and model-specific (SHAP) interpretability algorithms to rank feature importance. This is coupled with a visualization layer (PDP/ICE) to map the functional form of the surrogate responses across the 39-dimensional input space.

3. Multi-Objective Optimization Framework Design: Based on the empirical discovery of a strong negative correlation ($r \approx -0.75$) and opposing sensitivity directions between the two responses, the architecture proposes a conceptual multi objective optimization design. This framework is structured to guide future Pareto-based reactor physics optimizations, balancing multiplication characteristics against pin-level power distribution.

2.4 Experimental Setup

The experimental setup is configured to rigorously evaluate the predictive adequacy, generalization, and sensitivity of neutronics simulation parameters:

Dataset Configuration:

o *Training Set:* 24,000 independent simulation records used for model fitting and SHAP baseline calculations.

o *Test Set:* 360 independent records reserved strictly for out-of-sample validation and permutation testing.

o *Variables:* 41 numerical dimensions per observation (39 inputs X_1- X_{39} / E_1 - E_{39} 2 responses).

Model Training & Benchmarking:

o *Linear Regression:* Fitted on the training set; evaluated on the test set yielding an of 0.9872 () and 0.6058 ().

o *Ensemble Models:* Extra Trees and HistGradientBoosting trained and tested to benchmark against the linear

baseline, revealing that HistGradientBoosting outperforms MLR for ().

Interpretability Hyperparameters & Grid Setup:

o *Permutation Importance:* Executed with 20 random shuffling repetitions on the independent test set using the scoring metric.

o *PDP/ICE Grid:* A 20-point evaluation grid spanning the 5th to 95th percentile range of each input variable, generating 25 individual ICE profiles per feature to assess response heterogeneity and conditional behavior.

Evaluation Metrics: Predictive performance is quantified using the Coefficient of Determination (), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). Statistical significance of regression coefficients is assessed via HC3 heteroskedasticity-robust 95% confidence intervals to ensure reliability against variance anomalies.

3. Data structure and analytical scope

The investigation considers a PWR fuel assembly neutronics simulation dataset comprising 24,000 training records and 360 independent test records. Each observation contains 41 numerical variables. In the absence of an explicit variable schema, the first two columns are provisionally designated as response variables (Y1 and Y2), while the remaining 39 columns are considered input variables (X1–X39).

The analytical framework integrates descriptive statistics, Pearson and Spearman correlation, partial correlation, standardized multiple linear regression, permutation importance, SHAP-based model interpretation, and partial dependence and individual conditional expectation analysis.

The combination of these methods provides complementary perspectives on the relationships between the input parameter space and the two simulation responses. Statistical correlation identifies marginal relationships, partial correlation examines conditional associations, regression quantifies adjusted linear effects, and machine-learning interpretability techniques examine predictive dependence.

The principal methodological strength is the use of multiple analytical approaches rather than relying on a single feature-ranking method. This makes it possible to examine whether important input variables are consistently identified across statistical and predictive analyses.

Data qualification: The positional labels are provisional. The physical meanings, units, and identities of the variables must be established before engineering-specific conclusions can be drawn.

4. Results

4.1 Model performance and predictive adequacy

Table 1. Evaluation of the multiple linear regression surrogate models

Metric	Y1	Y2
Training R^2	0.98679	0.57277
Test R^2	0.98716	0.60583
Test RMSE	0.00220	0.07739
Test MAE	0.00155	0.06078

Table 1 presents the predictive performance of the multiple linear regression surrogate models developed independently for the two response variables. The evaluation uses the training dataset for model fitting and an independent test dataset for assessing generalization.

The Y1 model demonstrates excellent predictive accuracy, achieving a training R^2 of 0.98679 and an independent test R^2 of 0.98716. This indicates that approximately 98.72% of the observed variation in Y1 is explained by the fitted linear relationship on the test dataset. The low RMSE of 0.00220 and MAE of 0.00155 further demonstrate the model's strong predictive agreement with the observed values.

In comparison, the Y2 model achieves a training R^2 of 0.57277 and test R^2 of 0.60583. Its test RMSE and MAE are 0.07739 and 0.06078, respectively. These results indicate a substantially weaker linear relationship between the 39 input variables and Y2.

The difference in predictive performance suggests that the response represented by Y1 is more adequately described by a linear surrogate than Y2. The remaining unexplained variation in Y2 may be associated with nonlinear response characteristics, interactions between input variables, or stochastic effects in the simulation process. These explanations remain hypotheses until specifically tested.

Research implication: The substantially different model performance suggests that a single modeling strategy may not be equally appropriate for both neutronics responses. While linear regression provides a reliable approximation for Y1, more flexible nonlinear models may be required to characterize Y2.

A further consideration is that the test set contains only 360 records. Therefore, repeated validation or bootstrap-based uncertainty estimation would strengthen the reliability of the reported performance metrics.

4.2 Identification of influential enrichment variables

4.21 Infinite multiplication factor (k_{∞})

Feature	Pearson correlation	Standardized coefficient
E27	0.2005	0.1917
E10	0.1999	0.1835
E22	0.1983	0.1964
E13	0.1960	0.1810
E16	0.1939	0.1939

Table 2. Most influential enrichment variables associated with the infinite multiplication factor

Table 2 presents the five leading enrichment variables identified through Pearson correlation and standardized multiple regression coefficients for the infinite multiplication factor, provisionally associated with Y1.

All five variables exhibit positive Pearson correlation coefficients, ranging from 0.1939 to 0.2005. Similarly, their standardized regression coefficients are positive and range from 0.1810 to 0.1964.

Among the reported variables, E27 exhibits the highest marginal Pearson correlation with the response ($r=0.2005$), suggesting that it has the strongest individual linear association among the five selected

variables. E22 has the largest standardized regression coefficient (0.1964), indicating the strongest adjusted linear association among the listed variables when the remaining inputs are controlled.

The similarity between the correlation and regression results indicates that these variables retain positive associations with the response after accounting for the other predictors.

The findings suggest that multiple enrichment-related inputs contribute to the variation in the infinite multiplication factor rather than a single parameter dominating the relationship.

The positive associations indicate that increases in the selected variables tend to accompany increases in the modeled \(k_{\infty}\) response within the observed dataset. However, the magnitude of these relationships is moderate rather than exceptionally strong.

Importantly, the five variables should be interpreted as leading statistical indicators rather than independently established physical control parameters. Their engineering significance depends on their actual definitions, units, and roles in the simulation configuration.

The analysis also reports that all 39 enrichment variables have positive standardized regression coefficients. This indicates a consistent positive direction across the fitted linear model, although the statistical significance and uncertainty of every individual coefficient must be assessed using the complete confidence-interval table.

4.3 Pin power peaking factor (PPPF)

Table 3. Leading enrichment variables associated with the pin power peaking factor

Feature	Pearson correlation	Standardized coefficient
E18	-0.1720	-0.1746
E17	-0.1711	-0.1686
E10	-0.1624	-0.1534
E13	-0.1578	-0.1460
E11	-0.1539	-0.1596

Table 3 identifies the five leading enrichment variables associated with the pin power peaking factor (PPPF). Unlike the positive associations observed for the infinite multiplication factor, all five variables demonstrate negative Pearson correlation and standardized regression coefficients.

E18 exhibits the strongest negative marginal association (\(r=-0.1720\)), followed by E17 (\(r=-0.1711\)). E18 also exhibits the largest absolute standardized regression coefficient (-0.1746), indicating the strongest adjusted linear association among the listed variables.

The negative coefficients suggest that higher values of these variables are associated with lower predicted PPPF values within the observed parameter space.

The analysis further indicates that all 39 enrichment variables have negative standardized regression coefficients for PPPF.

The contrasting coefficient directions between the two responses represent an important finding. The

enrichment-related variables exhibit positive associations with the infinite multiplication factor but negative associations with the pin power peaking factor.

This suggests that the two neutronics responses exhibit different statistical relationships with the input parameter space. Consequently, optimizing one response independently may not necessarily produce an equivalent improvement in the other.

From a multi-objective neutronics analysis perspective, this observation motivates further investigation of the trade-offs between multiplication behavior and power distribution.

Nevertheless, these results do not establish that increasing any individual enrichment parameter will physically reduce PPPF under all operating conditions. Such conclusions require verified parameter definitions and controlled simulation experiments.

4.4 Comparative evaluation of regression and ensemble models

Table 4. Comparative predictive performance of regression and ensemble learning models

Target	Model	Test R^2
k_{∞}	Linear Regression	0.9872
k_{∞}	Extra Trees	0.5921
k_{∞}	HistGradientBoosting	0.8301
PPPF	Linear Regression	0.6058
PPPF	Extra Trees	0.4519
PPPF	HistGradientBoosting	0.6854

Table 4 compares the predictive performance of three modeling approaches: Linear Regression, Extra Trees, and Histogram-based Gradient Boosting (HistGradientBoosting). The comparison is conducted independently for the two neutronics response variables using the test-set coefficient of determination.

For k_{∞}, Linear Regression achieves the highest predictive performance, with an R^2 of 0.9872, followed by HistGradientBoosting at 0.8301 and Extra Trees at 0.5921. This indicates that the linear model provides the most accurate representation of the observed response among the three evaluated methods.

For PPPF, HistGradientBoosting achieves the highest R^2 of 0.6854, outperforming Linear Regression (0.6058) and Extra Trees (0.4519). This result suggests that the gradient-boosting approach captures predictive patterns in PPPF that are not fully represented by the linear regression model.

The different rankings across the two responses highlight the importance of response-specific model selection. Linear Regression performs exceptionally well for k_{∞}, whereas HistGradientBoosting provides a better approximation for PPPF.

The comparative results suggest that the relationship between the enrichment-related input variables and k_{∞} is predominantly represented by a stable linear structure within the sampled parameter space. In contrast, PPPF appears to benefit from a more flexible modeling approach.

However, the superiority of HistGradientBoosting for PPPF does not, by itself, prove the presence of nonlinear

physical interactions. Further analysis using interaction effects, nonlinear response profiles, and physics-informed sampling is required to establish the underlying response structure.

Important consistency check: The linear regression values in Table 4 agree with the test R^2 values in Table 1. The comparison should therefore be presented as a complementary model evaluation rather than an independent experiment.

4.5 Descriptive statistical analysis of the 39 input variables

Table 5. Descriptive statistical characteristics of the simulation responses and input variables

Table 5 presents the descriptive statistical characteristics of the two response variables and all 39 input variables. The reported statistics include sample count, mean, standard deviation, minimum, first quartile, median, third quartile, and maximum.

4.51 Statistical distribution of the response variables

The first response variable, Y1, has a mean of approximately 1.3251 and a standard deviation of 0.0178. Its observed values range from 1.24636 to 1.38558.

The relatively small standard deviation indicates that Y1 exhibits comparatively limited variation around its mean. The difference between its minimum and maximum values nevertheless indicates that the response is not constant across the simulation records.

The second response variable, Y2, exhibits a mean of approximately 1.8836 and a standard deviation of 0.1200. Its values range from approximately 1.5275 to 2.4730.

Compared with Y1, Y2 demonstrates substantially greater absolute dispersion. This difference is consistent with the weaker predictive performance of the linear surrogate model for Y2.

The interquartile distribution also indicates that Y2 varies more substantially across the simulation records than Y1.

4.52 Distribution of the input variables

The 39 input variables exhibit relatively similar descriptive statistical characteristics.

Their mean values are concentrated approximately between 2.833 and 2.876, while their standard deviations range approximately from 1.229 to 1.248.

The minimum values are close to 0.7, and the maximum values approach 5.0 across the input variables. Their first, second, and third quartiles are also relatively similar.

This indicates that the simulation dataset contains a broadly comparable numerical range across the input dimensions.

The similarity in marginal distributions is useful because it reduces the likelihood that apparent differences in feature importance are simply caused by substantial differences in numerical scales. Nevertheless, standardized coefficients and normalized importance measures remain necessary for meaningful comparisons.

The descriptive statistics indicate that the input variables occupy broadly similar numerical ranges, whereas the two response variables demonstrate different levels of dispersion.

This provides an important statistical foundation for the subsequent sensitivity analysis. In particular, the limited dispersion of Y1 is consistent with its low prediction error, whereas the greater variability of Y2 is accompanied by comparatively higher prediction errors.

Recommended table correction: Table 5 is currently titled Variable Sensitivity Analysis, but its contents are descriptive statistics. A more accurate title is Descriptive Statistical Analysis of Simulation Responses and Input Variables. Sensitivity rankings should be reported separately.

4.56 Pearson, Spearman, and partial-correlation analysis

The correlation analysis examines the statistical relationships between the two responses and the 39 input variables.

4.56.1 Pearson correlation

Pearson correlation measures the strength and direction of linear association between two variables.

The two response variables exhibit a Pearson correlation coefficient of:

$$ r_{Y1,Y2}=-0.7596 $$

This indicates a strong negative linear association between the two responses. In other words, higher values of one response tend to accompany lower values of the other response within the supplied simulation dataset.

The largest absolute Pearson correlation among the 39 input variables is approximately 0.0223. This indicates that the input variables exhibit very weak pairwise linear correlations in the training dataset.

Interpretation: The weak intercorrelation suggests that the inputs contain relatively limited linear redundancy. This is beneficial for interpreting individual regression coefficients because strong linear multicollinearity is less apparent.

However, low Pearson correlations do not rule out nonlinear dependencies or higher-order relationships.

4.56.2 Spearman correlation

Spearman correlation evaluates monotonic relationships using ranked observations and is less dependent on the assumption of linearity.

The relationship between the two responses is:

$$ \rho_{Y1,Y2}=-0.7496 $$

The similarity between the Pearson coefficient (-0.7596) and Spearman coefficient (-0.7496) indicates that the negative association is not limited to a purely linear relationship.

The consistency in magnitude suggests that the inverse relationship between the responses is broadly monotonic over the observed dataset.

Interpretation: The agreement between Pearson and Spearman results strengthens the evidence for a stable inverse statistical relationship between the two response variables.

It also suggests that the relationship is not driven exclusively by a small number of extreme observations, although formal robustness testing would be required to establish this.

4.6 Partial correlation analysis

Partial correlation measures the relationship between an input variable and a response after controlling for the remaining 38 input variables.

Unlike ordinary Pearson correlation, partial correlation attempts to isolate the unique linear association between an individual predictor and the response.

The partial correlation analysis is particularly relevant because it helps determine whether an input-response relationship remains after accounting for the other variables.

Interpretation: If an input exhibits similar marginal and partial correlations, its association with the response is relatively stable after controlling for the remaining predictors. Substantial differences between these two measures may indicate shared statistical information or suppression effects.

Since the input variables demonstrate very weak pairwise Pearson correlations, relatively similar marginal and partial associations may be expected. Nevertheless, this must be verified from the individual entries in the complete partial-correlation matrix.

The partial correlations should not be interpreted as causal effects. They represent conditional statistical associations under a linear adjustment framework.

4.7 Standardized regression coefficients and confidence intervals

The standardized multiple regression analysis provides an estimate of the direction and magnitude of the relationship between each input variable and the corresponding response while controlling for all other inputs. The standardized coefficient can be represented as:

$\beta_j^* = \beta_j \frac{\sigma_{X_j}}{\sigma_Y}$ where:

- β_j^* represents the standardized regression coefficient.
- β_j represents the original regression coefficient.
- σ_{X_j} represents the standard deviation of the input.
- σ_Y represents the standard deviation of the response. The coefficients are accompanied by HC3 heteroskedasticity robust 95% confidence intervals

The positive standardized coefficients reported for all 39 enrichment variables in the \(k_{\infty}\) analysis indicate positive adjusted linear associations.

Conversely, the negative coefficients reported for PPPF indicate inverse adjusted linear associations.

The confidence intervals provide additional information about the stability and uncertainty of these estimates.

- Coefficients with confidence intervals entirely above zero indicate positive conditional associations supported by the fitted model.
- Coefficients with confidence intervals entirely below zero indicate negative conditional associations.
- Coefficients with confidence intervals spanning zero provide insufficient evidence of a stable nonzero linear conditional association at the conventional 5% significance level.

The standardized regression results complement the correlation analysis by distinguishing marginal associations from adjusted relationships.

The consistency of coefficient directions across the enrichment variables is particularly noteworthy. It suggests that the two responses exhibit contrasting linear sensitivities across the modeled input space.

However, the regression coefficients represent conditional linear effects within the supplied dataset. They should not be treated as physical sensitivity coefficients without confirmation of the simulation parameter definitions and experimental design.

4.8 Permutation importance and SHAP-based interpretation

4.8.1. Permutation importance

Permutation importance evaluates how much a fitted model's predictive performance decreases when the values of an individual input variable are randomly shuffled.

In this study, importance is evaluated using the independent test dataset and R^2 scoring, with 20 permutation repetitions.

A larger positive decrease in R^2 indicates greater dependence of the model's predictive performance on the corresponding input.

The analysis identifies different leading variables for the two responses.

For Y1, the leading permutation-importance variables include:

- X23
- X8
- X22
- X19
- X16

For Y2, the leading variables include:

- X17
- X11
- X18
- X29
- X13

The difference between the two rankings indicates that the two response variables depend on different subsets of input variables within their respective fitted models.

For Y1, the prominent role of X23 and X8 indicates that these variables contribute substantially to the predictive performance of the linear surrogate.

For Y2, X17 and X11 emerge as particularly important predictors.

This response-specific pattern reinforces the observation that the two neutronics responses exhibit different statistical and predictive structures.

Permutation importance can produce negative values when shuffling a variable improves predictive performance. Such values should be retained because they may indicate sampling variability, model instability, or limited predictive contribution.

4.8.2 SHAP importance analysis

SHAP analysis provides a feature-attribution framework for explaining individual model predictions.

For the fitted linear regression models, SHAP contributions are calculated analytically relative to the training-set feature means: $\phi_j = \beta_j \ (X_j _ E[X_j])$where:

- ϕ_j is the contribution of feature j.
- β_j is its fitted regression coefficient.
- X_j is its observed value.
- $E[X_j]$ is its training set mean.

The mean absolute SHAP value is used to summarize the magnitude of each feature's contribution across the 360 independent test observations.

The leading SHAP variables are reported as follows.

For Y1:

1. X23
2. X9
3. X10
4. X8
5. X19

For Y2:

1. X18
2. X17
3. X11
4. X10
5. X20

The SHAP results provide a complementary perspective to permutation importance.

For Y1, X23 appears as the leading variable in both methods, indicating relatively consistent evidence of predictive importance. X8 and X19 also appear in both rankings.

For Y2, X17, X18, and X11 are identified as important by both permutation importance and SHAP analysis. This agreement strengthens their status as candidate predictive drivers.

However, differences between the rankings are also meaningful. For example, X10 appears among the leading SHAP variables but not among the five leading permutation-importance variables for Y1.

This difference arises because the two techniques quantify different properties:

- Permutation importance evaluates the loss in predictive performance after disrupting a feature.
- SHAP measures the magnitude of the feature's contribution to individual predictions.

Consequently, agreement between the methods provides stronger evidence of model dependence, whereas disagreement may reveal differences in attribution, predictive redundancy, or sampling sensitivity.

Important methodological qualification: These SHAP values are exact for the fitted linear surrogate models. They do not represent SHAP explanations of the Extra Trees or HistGradientBoosting models.

4.9 Partial dependence and individual conditional expectation analysis

The partial dependence plot (PDP) and individual conditional expectation (ICE) analyses examine how the fitted model responds when individual input variables are varied.

The analysis uses a 20-point evaluation grid spanning the 5th–95th percentile range of each input variable. For each input, 25 individual ICE profiles are generated.

9.1 Mathematical interpretation

The partial dependence function for input X_j can be expressed as: $PD_j(x_j) = \frac{1}{n}\sum_{i=1}^{n} \hat{f}\left(x_j, \boldsymbol{x}_{i,-j}\right)$ where:

- X_j represents the selected value of the input variable.
- $X_{i\text{-}j}$ represents the remaining input values for observation i.
- $\hat{f}$ represents the fitted predictive model.
- η represents the number of observations used for averaging.

ICE profiles provide corresponding individual predictions rather than averaging across all observations.

4.91 Interpretation of PDP and ICE patterns

The PDP curves provide an overall representation of how predicted responses change as individual inputs vary.

The ICE profiles reveal whether the response patterns are consistent across individual simulation records.

Three patterns are particularly important when interpreting these plots:

1. Approximately linear profiles: Indicate that the fitted model predicts a relatively consistent linear response over the evaluated input range.

2. Curved profiles: Suggest nonlinear response behavior within the fitted model.

3. Diverging or heterogeneous ICE profiles: Indicate that individual observations exhibit different response

patterns, potentially suggesting interactions or differences in conditional response behavior.

When PDP curves are approximately flat, the corresponding variable contributes relatively little to changes in the model's average prediction over the evaluated range.

It is important to distinguish the appearance of a PDP curve from the physical response of the actual neutronics system. These plots describe the fitted surrogate model, not independently simulated physical outcomes.

4.92 Figure 2: PDP and ICE profiles for Y1

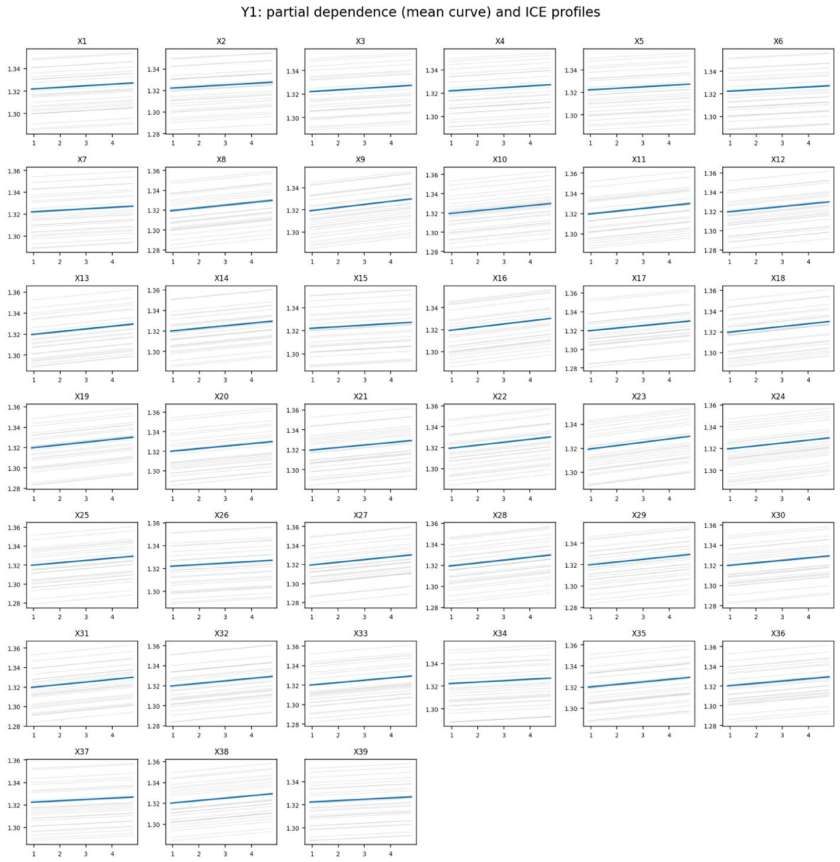

Figure 2 - All 39 Input Variables

Figure 2. Partial dependence and individual conditional expectation profiles for the 39 input variables with respect to Y1.

Figure 2 presents the PDP and ICE profiles for all 39 input variables associated with Y1. The figure provides a comprehensive visualization of the predicted response behavior across the observed input parameter space.

The PDP curves represent the average predicted Y1 response as each individual input is varied, while the ICE profiles illustrate the corresponding predictions for individual test observations.

The figure complements the sensitivity rankings by revealing the functional form of the model response rather than simply reporting the relative importance of individual predictors.

For variables identified as highly influential through permutation and SHAP analysis, the PDP/ICE profiles provide additional insight into the direction and consistency of their contributions.

In particular, the prominence of X23 and X8 in the predictive importance analysis makes their response profiles especially relevant for examining the behavior of the Y1 surrogate.

Figure 2 provides a visual basis for identifying potentially important enrichment related parameters and examining whether their predicted effects remain consistent across individual simulation records.

4.9.4 Figure 4: PDP and ICE profiles for Y2

FIGURE 4 — ALL 39 INPUT VARIABLES

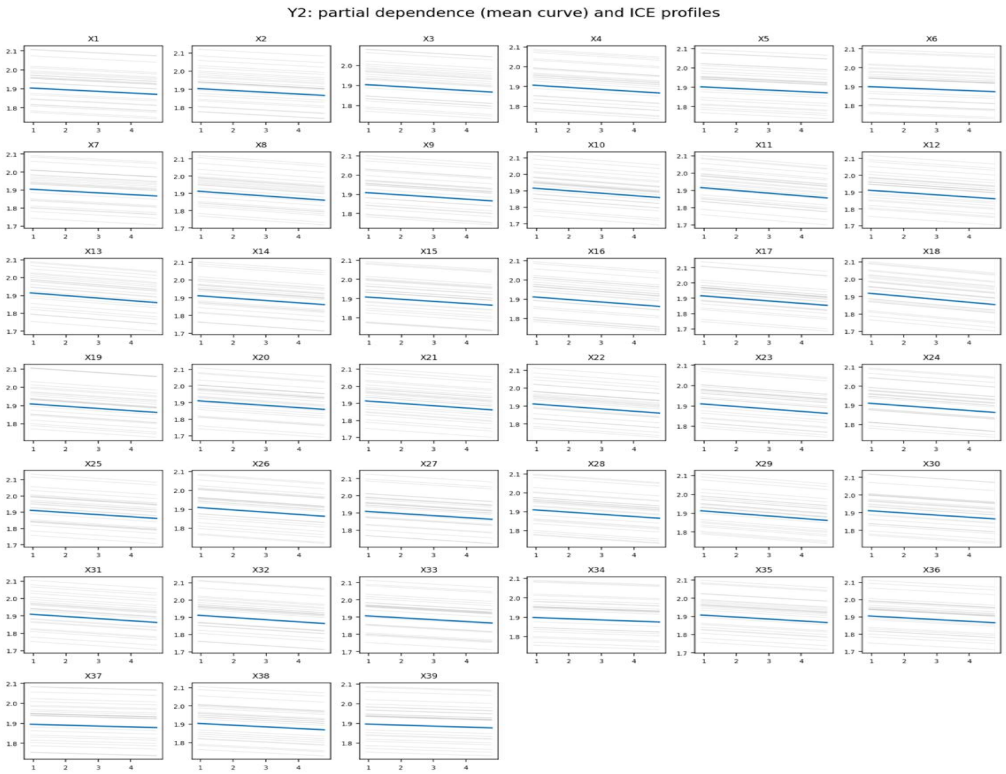

Figure 4. Partial dependence and individual conditional expectation profiles for the 39 input variables with respect to Y2

Figure 4 illustrates the PDP and ICE profiles for the 39 input variables associated with Y2.

Compared with Y1, Y2 exhibits lower predictive accuracy under the linear regression model. Therefore, the PDP/ICE analysis is particularly useful for investigating the behavior of its fitted response approximation.

The leading variables identified through permutation importance and SHAP analysis, particularly X17, X18, and X11, provide important candidates for examining response variation.

The ICE profiles allow the investigation of whether the predicted response changes similarly across individual observations or exhibits heterogeneous patterns.

A systematic difference between the individual ICE profiles and their average PDP curves may indicate that the average response does not adequately represent every observation.

Research implication: The figure provides a useful interpretability framework for identifying candidate variables that may require more detailed nonlinear and interaction based investigation.

Important qualification: The report does not provide numerical PDP slopes or detailed descriptions of individual curve shapes. Therefore, claims about specific nonlinear thresholds, turning points, or interaction effects should only be included after inspecting the actual plotted curves.

4.10 Sensitivity ranking and comparative feature importance

4.10.1 Figure 1: Sensitivity ranking for Y1

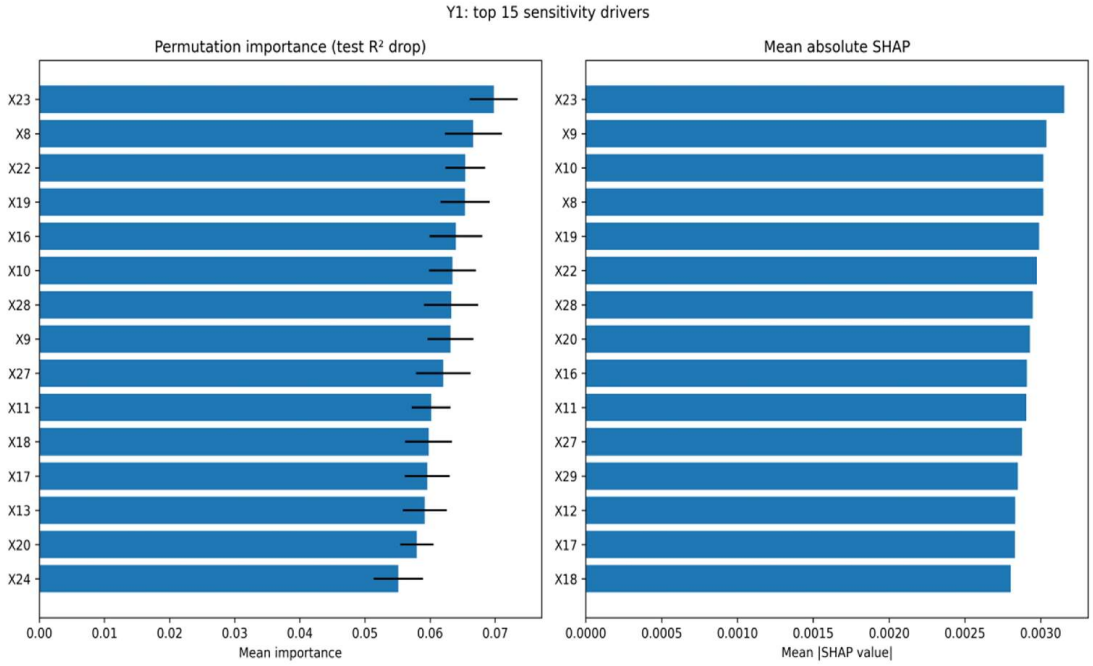

Figure 1. Comparison of test set permutation importance and mean absolute SHAP importance for Y1

Figure 1 compares the importance of the 39 input variables using two independent model-interpretation approaches.

Permutation importance evaluates the extent to which predictive performance deteriorates when a feature is randomly shuffled, whereas SHAP importance measures the average magnitude of the feature's contribution to model predictions.

The comparison identifies X23 as a particularly important variable because it appears at the top of both importance rankings.

X8 and X19 also appear among the leading variables in both methods. This agreement suggests that these variables contribute consistently to the fitted model's predictions.

However, X9 and X10 are prominent in the SHAP ranking, illustrating that feature attribution and predictive reliance do not necessarily produce identical rankings.

The convergence of the two importance measures provides stronger evidence for identifying candidate predictive drivers of Y1.

X23, X8, and X19 should therefore be prioritized for further investigation using physical parameter mapping and controlled simulation experiments.

The results also demonstrate the importance of using multiple interpretability methods when examining high-dimensional simulation datasets.

4.10.2 Figure 3: Sensitivity ranking for Y2

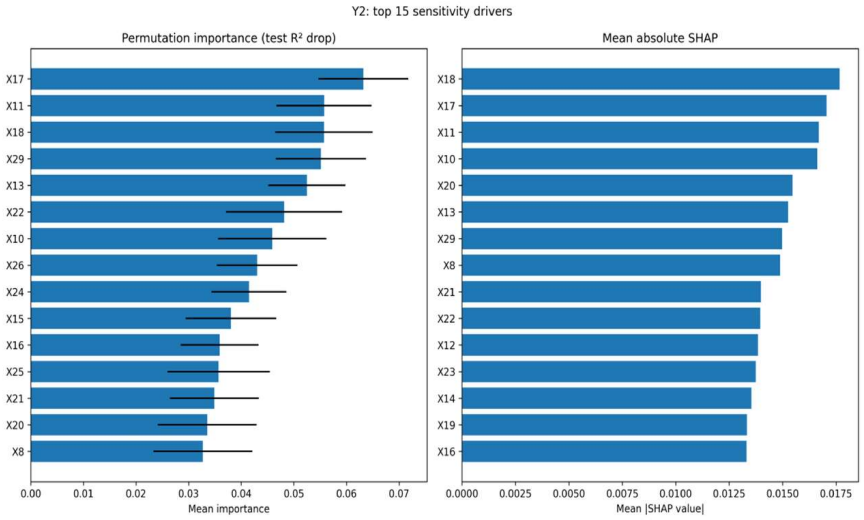

Figure 3. Comparison of test-set permutation importance and mean absolute SHAP importance for Y2

Figure 3 presents the comparative importance of the 39 input variables for predicting Y2.

The permutation and SHAP rankings identify X17, X18, and X11 as consistently important variables.

The identification of these variables through both approaches suggests that they make substantial contributions to the predictive behavior of the Y2 model.

The presence of X10 and X20 among the leading SHAP variables, despite their absence from the leading permutation-importance list, indicates that their average contribution to individual predictions does not necessarily translate into an equally large reduction in overall predictive performance when permuted.

The results indicate that the predictive structure of Y2 differs from that of Y1.

Whereas X23 is prominent for Y1, X17 and X18 are more prominent for Y2.

This distinction is relevant for multi-response neutronics modeling because it suggests that different simulation parameters may need to be prioritized depending on the response being evaluated.

However, feature importance alone cannot establish whether modifying a particular parameter would improve the physical performance or safety characteristics of a fuel assembly.

11. Integrated discussion of analytical findings

The combined results provide a multi-dimensional assessment of the statistical relationships and predictive importance of the 39 input variables associated with two neutronics responses.

Several important findings emerge from the analysis.

4.11.1 Different predictive structures across neutronics responses

The first major finding concerns the difference in predictive adequacy between Y1 and Y2.

The linear regression model achieves a test R^2 of 0.9872 for Y1, demonstrating that the response is very well represented by a linear combination of the supplied inputs within the sampled dataset.

In contrast, the test R^2 for Y2 is 0.6058, indicating a substantially weaker linear approximation.

The comparative model evaluation further demonstrates that HistGradientBoosting performs better than Linear Regression for Y2, achieving an R^2 of 0.6854.

These results suggest that the predictive structure of Y2 deserves additional investigation using nonlinear modeling approaches.

4.11.2 Contrasting relationships between the responses

The second important finding concerns the opposing directions of the standardized enrichment-related coefficients.

All 39 enrichment variables are reported to have positive standardized coefficients for \(k_{\infty}\), whereas the corresponding coefficients for PPPF are negative.

This opposing direction suggests that the two neutronics responses exhibit contrasting statistical relationships with the enrichment-related parameter space.

The Pearson and Spearman correlations between the two responses are also strongly negative, at “0.7596 and “0.7496, respectively.

Together, these observations indicate that the two responses should be considered jointly when developing a comprehensive fuel-assembly optimization framework.

4.11.3 Identification of response-specific predictive drivers

The permutation and SHAP analyses identify different leading input variables for the two responses.

For Y1, X23, X8, and X19 exhibit relatively consistent importance across the two methods.

For Y2, X17, X18, and X11 demonstrate similar cross-method prominence.

This response-specific ranking suggests that a single universal feature-importance ranking may not adequately represent the predictive structure of both neutronics responses.

Consequently, sensitivity analysis should be conducted separately for each response before developing multi-objective optimization strategies.

4.11.4 Complementary value of statistical and model-based interpretability

An important contribution of this work is the integration of marginal correlation, partial correlation, standardized regression, permutation importance, SHAP, and PDP/ICE.

Each method addresses a different analytical question:

Analytical method	**Main contribution**
Pearson correlation	Identifies linear relationships
Spearman correlation	Identifies monotonic relationships

Partial correlation	Identifies conditional linear associations
Standardized regression	Quantifies adjusted linear effects
Confidence intervals	Evaluates coefficient uncertainty
Permutation importance	Measures predictive reliance
SHAP	Explains feature contributions
PDP	Describes average predicted response profiles
ICE	Reveals individual response heterogeneity

The agreement between methods increases confidence in identifying candidate influential variables.

Conversely, disagreement between rankings provides useful information about differences in feature attribution, predictive dependence, and model behavior.

4.11.5 Implications for multi-objective neutronics optimization

The contrasting response relationships provide a foundation for developing a multi-objective optimization framework.

In particular, the simultaneous consideration of k_{∞} and PPPF may help investigate trade-offs between multiplication characteristics and pin-level power distribution.

A possible future optimization framework could be formulated as:

$$\begin{aligned} \text{Objective 1:} \quad & \max f_1(\mathbf X)=k_{\infty}\\ \text{Objective 2:} \quad & \min f_2(\mathbf X)=PPPF \end{aligned}$$

subject to verified engineering constraints on the input parameters and the relevant reactor-physics and operational requirements.

The resulting optimization problem could be investigated using Pareto-based multi-objective algorithms.

However, the present statistical results are insufficient to define physical feasibility constraints or recommend operational parameter changes. Those require validated physical models and engineering specifications.

5. Limitations

The present investigation provides a comprehensive exploratory analysis, but several limitations must be addressed before the findings can support definitive engineering conclusions.

1. Missing variable definitions

The supplied dataset does not contain verified physical descriptions or units for the 39 input variables and two responses. The provisional X1–X39 and Y1–Y2 labels must be replaced with their actual simulation parameter names.

2. Limited independent test sample

The independent test dataset contains 360 records. Although this provides an independent evaluation, repeated

validation and bootstrap confidence intervals are recommended to quantify uncertainty in predictive metrics and feature rankings.

3. Dependence on linear surrogate models
The exact SHAP values and regression coefficients explain the fitted linear models. They should not be interpreted as complete representations of the underlying physical neutronics system.

4. Nonlinear response characteristics
The lower predictive performance for Y2 indicates that further nonlinear modeling is warranted. Models such as Random Forest, Extra Trees, XGBoost, and gradient boosting should be evaluated using consistent validation procedures.

5. Interaction effects
The current analysis focuses primarily on individual feature contributions. Interactions between enrichment-related variables may influence neutronics responses and should be investigated through explicit interaction analysis.

6. Physical sensitivity versus predictive importance
Permutation importance and SHAP measure model-dependent predictive contributions. They do not quantify the proportion of physical output variance attributable to individual input parameters.

For a stronger physics-oriented study, Morris screening and Sobol global sensitivity analysis should be considered, provided that the simulation sampling design supports their assumptions.

7. Simulation data quality and distribution shift
The distributions of the training and test datasets should be compared to establish whether the independent test records represent the same parameter space as the training observations. Residual analysis, outlier investigation, and simulation convergence checks are also recommended.

8. Variable naming consistency
The report uses E-based identifiers in Tables 2 and 3 and X-based identifiers in the remaining analysis. The correspondence between these naming systems must be verified before publication.

6. Conclusion

This study presents an integrated statistical and model-based sensitivity analysis of 39 input variables associated with two PWR fuel-assembly neutronics responses.

The findings demonstrate substantial differences in the predictive characteristics of the two responses. The linear regression surrogate achieves excellent predictive performance for Y1, with a test R^2 of 0.9872, while the corresponding model for Y2 achieves a more moderate R^2 of 0.6058.

The comparative model evaluation further identifies HistGradientBoosting as a more effective predictive approach for Y2, achieving a test R^2 of 0.6854.

The correlation analysis reveals a strong negative association between the two responses, while the regression results indicate contrasting directions of association across the enrichment-related inputs.

The permutation and SHAP analyses identify response-specific predictive drivers, with X23, X8, and X19 emerging as important variables for Y1 and X17, X18, and X11 for Y2.

Furthermore, the PDP and ICE analyses provide an interpretability framework for examining predicted response profiles and individual prediction heterogeneity across the 39-dimensional input space.

Overall, the study establishes a comprehensive exploratory framework for examining input response relationships in PWR fuel assembly neutronics simulation datasets. Its principal contribution is the integration of statistical association analysis, regression based sensitivity assessment, predictive feature importance, and model interpretability.

Future work should prioritize verified physical parameter mapping, nonlinear interaction analysis, repeated validation, and physics informed global sensitivity methods. These extensions would strengthen the physical interpretability and generalizability of the findings and provide a more reliable foundation for multi objective neutronics optimization.

References

[1] Santosh, V., Vinod, G., Saraf, R. K., Ghosh, A. K., Kushwaha, H. S. (2007). Application of artificial neural networks to nuclear power plant transient diagnosis. *Reliability Engineering System Safety, 92*(10), 1468–1472. https://doi.org/10.1016/j.ress.2006.10.009.

[2] Yao, Y., Ge, D., Yu, J., Xie, M. (2022). Model-based deep transfer learning method to fault detection and diagnosis in nuclear power plants. *Frontiers in Energy Research, 10*, 823395. https://doi.org/10.3389/fenrg.2022.823395.

[3] Qian, G., Liu, J. (2022). A comparative study of deep learning-based fault diagnosis methods for rotating machines in nuclear power plants. *Annals of Nuclear Energy, 178*, 109334. https://doi.org/10.1016/j.anucene.2022.109334.

[4] Radaideh, M. I., Kozlowski, T. (2019). Combining simulations and data with deep learning and uncertainty quantification for advanced energy modelling. *International Journal of Energy Research, 43*(14), 7866–7890.

[5] Radaideh, M. I., Kozlowski, T. (2020). Analyzing nuclear reactor simulation data and uncertainty with the group method of data handling. *Nuclear Engineering and Technology, 52*(2), 287–295.

[6] Radaideh, M. I., Pigg, C., Kozlowski, T., Deng, Y., Qu, A. (2020). Neural-based time series forecasting of loss of coolant accidents in nuclear power plants. *Expert Systems with Applications, 160*, 113699.

[7] Saleem, R. A., Radaideh, M. I., Kozlowski, T. (2020). Application of deep neural networks for high-dimensional large BWR core neutronics. *Nuclear Engineering and Technology, 52*(12), 2709–2716.

[8] Gong, H., Cheng, S., Chen, Z., Li, Q. (2022). Data-enabled physics-informed machine learning for reduced-order modeling digital twin: Application to nuclear reactor physics. *Nuclear Science and Engineering, 196*(6), 668–693. https://doi.org/10.1080/00295639.2021.2014752.

[9] Song, H., Song, M., Liu, X. (2022). Online autonomous calibration of digital twins using machine learning with application to nuclear power plants. *Applied Energy, 326*, 119995. https://doi.org/10.1016/j.apenergy.2022.119995.

[10] Tunkle, L., Abdulraheem, K., Lin, L., Radaideh, M. I. (2025). Nuclear microreactor transient and load-following control with deep reinforcement learning. *Energy Conversion and Management: X, 27*, 101090.

[11] Radaideh, M. I., Tunkle, L., Price, D., Abdulraheem, K., Lin, L., Elias, M. (2025). Multistep criticality search and power shaping in nuclear microreactors with deep reinforcement learning. *Nuclear Science and Engineering*, 1–13.

[12] Burnett, L. A., Nabila, U. M., Radaideh, M. I. (2026). Variational digital twins. *Energy and AI*, 100756.

[13] Panczyk, N. R., Erdem, O. F., Radaideh, M. I. (2025). Opening the AI black-box: Symbolic regression with Kolmogorov-Arnold networks for advanced energy applications. *Energy and AI*, 100595.

[14] Jendoubi, C., Asad, A. (2024). A survey of artificial intelligence applications in nuclear power plants. *IoT, 5*(4), 666–691. https://doi.org/10.3390/iot5040030.

[15] Ejigu, D. A., Tuo, Y., Liu, X. (2024). Application of artificial intelligence technologies and big data computing for nuclear power plants control: A review. *Frontiers in Nuclear Engineering, 3*, 1355630. https://doi.org/10.3389/fnuen.2024.1355630.

[16] Kim, H., Persico, A., Burnett, L., Forget, B., Radaideh, M. I. (n.d.). *Surrogate-guided generative modeling with physics verification for expensive multiphysics inverse problems*.

[17] Zhu, S., Chen, S., Zhou, J., Zou, C., Chen, J., Cai, X. (2025). *Status of machine learning technology applications in reactor design*.

[18] Zhang, R. Z., Zhu, S. F., Wang, K., et al. (2026). Artificial intelligence in reactor physics: Current status and future prospects. *Nuclear Science and Techniques, 37*, 100. https://doi.org/10.1007/s41365-026-01928.

[19] Shriver, F., Gentry, C., Watson, J. (2021). Prediction of neutronics parameters within a two-dimensional reflective PWR assembly using deep learning. *Nuclear Science and Engineering, 195*(6), 626–647.

[20] Permann, C. J., Gaston, D. R., Andrš, D., Carlsen, R. W., Kong, F., Lindsay, A. D., Miller, J. M., Peterson, J. W., Slaughter, A. E., Stogner, R. H., et al. (2020). MOOSE: Enabling massively parallel multiphysics simulation. *SoftwareX, 11*, 100430.

[21] Price, D., Roskoff, N., Radaideh, M. I., Kochunas, B. (2024). Multiphysics modeling of heat pipe microreactor with critical control drum position search. *Nuclear Science and Engineering*, 1–20.

[22] Brunton, S. L., Noack, B. R., Koumoutsakos, P. (2020). Machine learning for fluid mechanics. *Annual Review of Fluid Mechanics, 52*(1), 477–508.

[23] Price, D., Roskoff, N., Radaideh, M. I., Kochunas, B. (2023). Thermal modeling of an EVinci™-like heat pipe microreactor using OpenFOAM. *Nuclear Engineering and Design, 415*, 112709.

[24] Abboud, A., de Lambert, S., Garnier, J., Leturcq, B. (2026). Uncertainty quantification in coupled multiphysics systems via Gaussian process surrogates: Application to fuel assembly bow. *International Journal for Uncertainty Quantification, 16*(4).

[25] Zeng, F. L., Zhang, X. L., Zhao, P. C., Liu, Z. J. (2026). Application research of a hybrid data-and knowledge-driven artificial intelligence scientific computing model in neutron diffusion calculation for nuclear reactors. *Nuclear Science and Techniques, 37*(2), 24.

[26] Appriou, T., Rullière, D., Gaudrie, D. (2024). Bayesian optimization in high-dimension via a combination of Kriging sub-models. *MASCOT-NUM 2023*, April 3–6, Le Croisic, France.