Domain-Constrained Machine Learning for Decarbonization of Thermal Power Plants to support Climate Action

Waqar Muhammad Ashraf[1,2,3+], Vivek Dua[1], and Ramit Debnath[2,4]

[1]University College London, London, WC1E 7JE, UK
[2]University of Cambridge, Cambridge, CB2 1PZ, UK
[3]The Alan Turing Institute, London, NW1 2DB, UK
[4]Caltech, Pasadena, 91125, US
[+]Corresponding author: waqar.ashraf.21@ucl.ac.uk

Abstract

Industrial decarbonization of global thermal power plants is critical for achieving climate action goals from the energy sector—the largest contributor of greenhouse gas emissions. First-principle models, built on certain assumptions, may not accurately predict the power generation dynamics of ageing large-scale industrial power systems. Whereas pure data-centric machine learning (ML) algorithms lack the interpretability and domain-inconsistency aspects in their predictions. The development of data-metrics-guided trained ML models and / or carrying out data-driven domain-constrained optimization analytics have the promise and potential to transform how we design, operate, and control industrial thermal power plants. In this perspective, we identify some data-metrics and multivariate constraints to quantify and / or represent the data-driven domain of thermal power plants. We also highlight that the domain-constrained optimization framework optimizes the component or system design of thermal power plants and can reduce process-related emissions. The domain-constrained optimization framework promotes artificial intelligence (AI) adoption in industrial environment and can lever up the improvement in energy efficiency, optimizing the fuel consumption, and reducing the emissions throughout the life cycle of thermal power plants. The direct reduction in process-related emissions achieved through AI-powered domain-consistent analytics contributes to the decarbonization of thermal power plants and support climate action.

1 Introduction

With the advent of a new data-driven industrial age, the global digital transformation of our society, also known as digitalization, is rapidly shaping our existing anthropogenic systems and processes. However, we have yet to fully understand the impact of this digitalization, raising open-ended questions about its potential to cause more harm to the environment, economy, and society than its perceived benefits [1, 2]. Big data, machine learning (ML), and artificial intelligence (AI) algorithms drive the digital transformation at the system level, attempting to meaningfully interpret the vast amount of data generated from our economic activities. This is often referred to as the building blocks of the 4^{th} Industrial Revolution or Industry 4.0 [3, 4].

There is a direct link between economic activities and greenhouse gas (GHG) emissions, driven by the production of goods and services in the energy, chemical and manufacturing industries for societal consumption [5, 6]. Rapidly increasing GHG emissions are one of the main causes of climate change and are potentially emitted from thermal power plants [7]; it is urgent that we reduce these emissions through various climate actions such as mitigation, adaptation, and carbon dioxide removal [8]. Therefore, thermal power plants are crucial leverage points for reducing the huge volumes of emissions that are discharged to the environment. Key enabling technologies such as the electrification of industrial processes, carbon capture, storage, and utilization, biomass-derived biofuels, hydrogen and ammonia as energy vectors, and cofiring of biomass with fossil fuels hold the promise of decarbonizing the power sector [9]. However, manufacturing and supply chain issues, techno-economic limitations of renewable-source-driven power systems, and the lack of harmonized global efforts to adopt these technologies across sectors hinder industrial decarbonization at scale [10]. The social tipping elements of low-carbon and renewable energy technologies pose a significant threat to the economic vulnerability of low-income countries, which are at the forefront of the fight against climate change [11]. The techno-socio-political challenges of installing decarbonization technologies across sectors are not only highly complex and heterogeneous in nature [12], but also require global efforts to develop just and sustainable transition pathways for the decarbonization of energy, industry, and the sectors difficult to decarbonize.

In this spectrum of issues surrounding "just" decarbonization of power sector, the International Energy Agency (IEA) advocates for flexible system design and improving the operational efficiencies of existing energy and industrial power assets [13]. The improved energy efficiency provides the same service with the reduction in consumption of fossil fuels that are converted into lower emissions discharge to the environment. Flexible design with improved demand-side management, improved operation, and real-time control are the key application domains where AI can significantly contribute to increasing the efficiency bottlenecks of thermal power plants, and can help make informed decisions toward climate action [3, 14]. Research shows that industrial decarbonization using AI on a large scale depends on the development of cyber-physical systems (CPS) [4]. CPS is an integration and intelligent orchestration of computation and physical processes in which they operate on different spatial and temporal scales, exhibit multiple and distinct behavioral modalities, and interact with each other in ways that change with context [15]. Implementing a full-scale and robust CPS is technologically challenging and demands innovation in AI algorithm design that is interpretable and explainable to humans. This is especially true when CPS is created for industrial processes, where the distribution of data becomes highly non-linear and complex, and empirical models do not explain the operation of this high-dimensional system [16]. At the same time, we argue that these systems must also be accessible and understandable to policymakers who will ultimately design appropriate decarbonization policies at the local, regional, and international level [17].

AI serves as a lever to drive the higher performance of industrial thermal power systems through optimal consumption of resources to provide the same energy services, reduction in the intensity of greenhouse gases, and management of demand and supply of the power grids through deep predictive capabilities [18]. A recent research study finds that AI can reduce 2.4 giga tonnes of GHG while adding 5.2 trillion USD to the global economy relative to the baseline projections until 2030 [19]. Considering the enormous potential of AI towards the environmental sustainability, it becomes imperative to understand the interpretable performance of the AI-based algorithms. Theoretically, we rely on the following definitions of AI interpretability, proposed by Doshi-Velez and Kim: *"the ability to explain or to present in understandable terms to a human"* [20] and Miller *"the degree to which a human can understand the cause of a decision"* [21].

There are three major frameworks for evaluating AI interpretability: application-grounded, human-grounded, and functionally-grounded evaluations [20, 22]. Application-based evaluation checks how the results of the interpretation process affect the human, domain expert, or end-user in terms of a specific and well-defined task or application. In human-grounded evaluations, any human end-user can serve as a tester instead of a domain expert. The objective is not to assess an interpretation's suitability for a particular application but to assess its quality in a broader context and gauge the accuracy of capturing general concepts. Functionally grounded evaluation does not require any experiments that involve humans, but instead uses formal, well-defined mathematical definitions of interpretability to evaluate the quality of an interpretability method. For example, once a class of models passes some interpretability criteria through human- or application-based experiments, we can use mathematical definitions to further rank the interpretability models' quality [20].

Functionally grounded evaluation approaches can be broadly categorized into ante hoc and post hoc techniques and are the strong foundations of eXplainable AI (XAI). In the ante-hoc method, ML models are simple and transparent by design and can provide insights into how the model makes the prediction. Examples include linear models, generalized additive models, and decision trees, as described in [23–26]. Since many real-life applications have complex and non-linear relationships, simple ML algorithms may not model the system with reasonable accuracy, leading to incomplete and inaccurate evaluation of interpretability performance. The ante-hoc approach includes embedding the trained ML models, which are complex and not transparent, into the interpretability evaluation environment, e.g., partial dependency, SHapley Additive exPlanations (SHAP), permutation feature importance, Local Interpretable Model-agnostic Explanations (LIME), etc. The ante-hoc approach is commonly used by the research community for a range of applications [27–31]. The potential challenge with the ante-hoc approach is building another model on top of the trained ML model that propagates error in the interpretability assessment. Moreover, the fundamental architecture and algorithm of ML models such as neural networks remain non-transparent, requiring research toward a design of interpretable ML models for improved interpretability evaluation.

Here, we expand upon the theoretical framework of functionally grounded AI interpretability, leveraging data metrics to train the ML models for thermal power plants. The operation dynamics of the large-scale power plants are well captured in the data collected from the years of operation. The stored dataset also serves as the domain knowledge of the power plant since different modes of power generation and the effect of variable interactions and couplings are available in the data to understand the operating characteristics and response of the power plant. The variable dependencies and relationships can be quantified through different statistical metrics. These metrics can be embedded into the algorithm design of ML models, which contribute to the domain-consistent predictions of the models [32] as well as accurate interpretability performance [33, 34]. Furthermore, data-driven multivariate techniques, which simultaneously handle the variable dependencies, can also be deployed to construct data-driven constraints on the collected data of the power plants and are embedded in the domain-constrained optimization or control problems. The domain-constrained data-driven analytics can significantly improve the efficacy of the optimal solutions for the design, operation and control applications [35–37], which can also be applied to thermal power plants for performance gains and techno-economic and environmental benefits. AI-powered energy efficient power generation reduces the process-related emissions from the industrial thermal power plants and contributes to industrial decarbonization to support climate action [18, 38].

We structure this perspective as follows: we expand on the evolutionary paradigm of industrial data and how it relates to industrial decarbonization's objectives for thermal power plants in section 2 . Next, we present the existing challenges and opportunities of data-centric industrial decarbonization in section 3. After scoping the challenges and opportunities, we provide insights into fostering innovation in ML algorithm design through embedding the statistical data-metrics and data-driven multivariate constraints and aligning it with the objectives of industrial decarbonization's goal for thermal power systems in section 4. Finally in section 5, we present a way forward for supporting climate action and industrial decarbonization via data-driven domain-constrained ML model-based design, operation, and control of industrial thermal power plants.

2 Historical shift in human-learning paradigm to data-centricism

The human capacity to understand the physical world has evolved through empirical, theoretical, and computational science approaches, leading us now into the era of data-driven science. Initially, empirical sciences attempted to explain an event without theoretical insights. The empirical sciences eventually evolved into the realms of theoretical sciences that lasted until around 1950, and the discoveries of universal laws enabled us to understand and explain the physical phenomena in nature. The theoretical sciences era gave way to the computational sciences approach, which combined both empirical and theoretical methods and persisted until roughly 2010. Since then, better computers and technology have created a new way of learning called "data-driven science". This way of learning focuses on getting knowledge from the data we collect at the points where different fields of knowledge meet [39]. We argue that this has the potential to fundamentally transform how industrial systems—particularly thermal power plants—are designed, operated, and controlled, enabling techno-economic and environmental benefits that support decarbonization and climate action.

2.1 Data driving the evolution of analytics for thermal power plants

Traditionally, experimentally validated mathematical models are created to explain the characteristics of engineering systems [40, 41]. These mathematical models provide mechanistic insights into the operation of the system and comply with collective human understanding and physics of the system . Historically, the industrial operations of thermal power plants are being maintained and controlled through first-principle models [42, 43]. However, their accuracy and adaptability deteriorate for aging thermal power plants because many complex processes are difficult to model to predict the behavior of aging power plant [44]. Furthermore, first-principle modeling of the operation of large-scale thermal power plant is quite challenging considering the large input space and nonlinear interactions between the systems. This is further emphasized conceptually in Figure 1(a) that as we observe increasingly larger data observations, the mathematical models become more and more complex to interpret the data occurrence. As data occurrence and distribution become highly nonlinear and complex, the mathematical models fail to explain the data generation process of high-dimensional data. Applying this concept to the scale of industrial operations of thermal power plants where the hyper-dimensional control space and data volumes are quite large, creating accurate mathematical models for the dynamic operation of large thermal power plants is quite challenging [16]. Moreover, the simulation and optimization analyses based on large mathematical models become computationally prohibitive, which further limits the applicability of the mathematical models for real-time optimization of large-scale thermal thermal power plants.

To address these challenges, we turn to the data-driven learning paradigm, which seeks to take advantage of the data to build customized models of the system under consideration. Data-driven models, including ML, require lower computational resources for modeling and optimization analyses compared to traditional mathematical models [45]. Moreover, ML algorithms are known to capture nonlinear dynamics and uncover underlying relationships out of the hyper-dimensional variable space [34, 46–49]. In recent years, ML has evolved from "shallow" application domains, such as recognizing cats on the internet to more complex constructs such as graphs and symbolic representations, in-variances, and positional embeddings, which represent deeper levels of scientific abstraction. Currently, classical ML methods are making significant impacts in biological and medical fields [50–52], despite the fact that these fields often have poorly defined entities and governing laws to interpret the behavior of the system.

In contrast, the physical sciences have rigorously defined entities and relationships based on conservation laws, in-variances, and causal relationships [53]. Applying ML models without taking these constraints into account can result in inaccurate results if model training violates physical relationships. Therefore, pure data-centric approaches face criticism for their lack of transparency and explainability, which limits their adoption in complex and safety-critical applications [54]. One powerful ML approach that has recently emerged is symbolic regression. Methods such as SISSO [55], SinDy [56], and PySR [57] derive symbolic expressions from observational data for applications such as predicting chemical reactivity, relating crystallographic structures to atomic radii, and defining electron transfer directions between elements. However, the interpretability of these expressions remains challenging for medium- to large-scale problems, often missing the opportunity to address the model interpretability problem. Recently, Kolmogorov-Arnold Networks (KAN) algorithm is reported, which provides interpretable expression(s) to explain the model-based predictions [58]. However, the applicability of the KAN model for noisy datasets and higher-dimensional problems, which are typical of industrial thermal power systems, remains unexplored to investigate its computational performance and predictive accuracy compared to standard ML models.

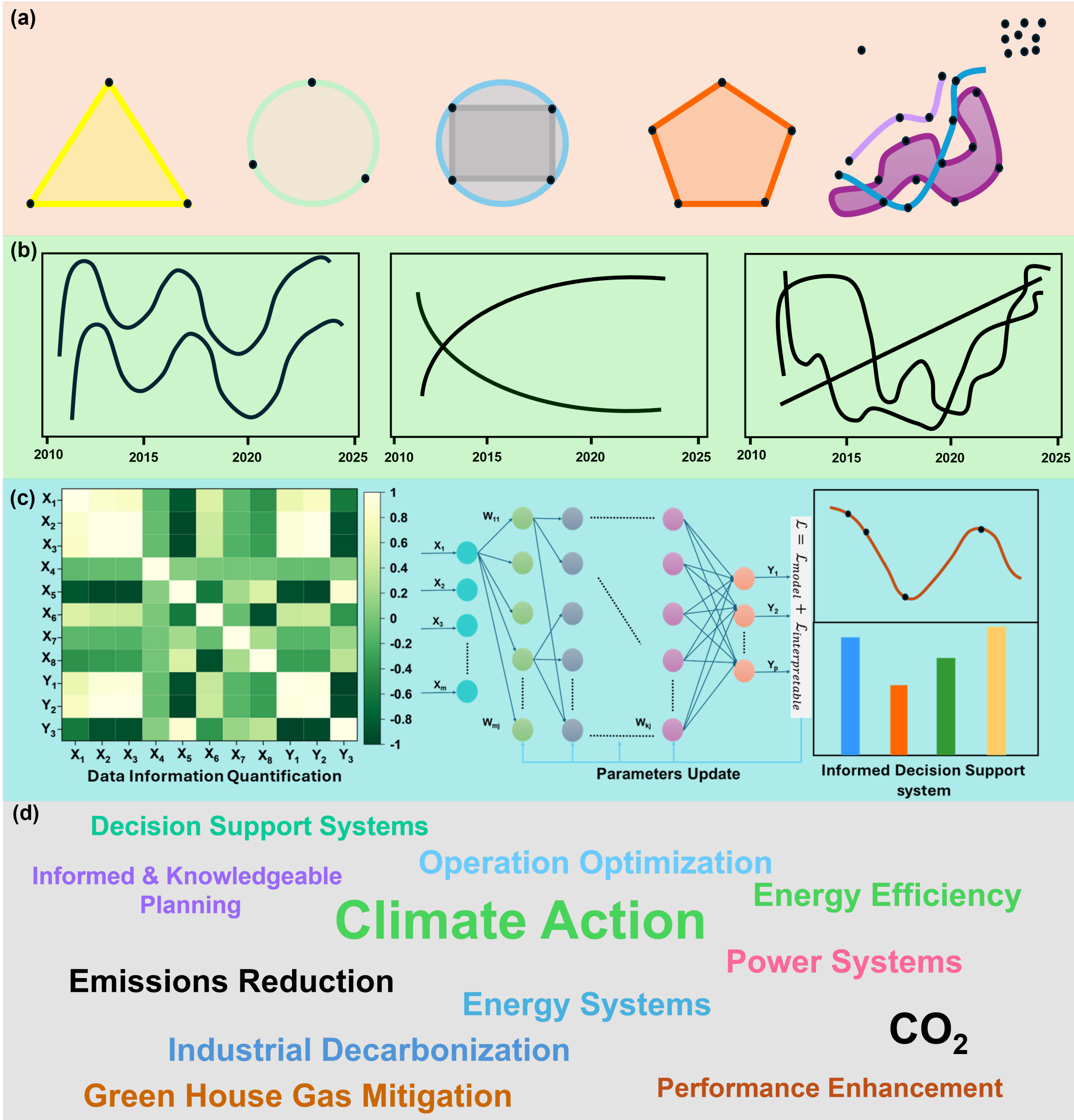

Figure 1: How data direct to update the model for explaining an event. (a) Observing small number of data observations results in fitting different low-dimensional mathematical models. As the size of randomly observed data increases, traditional mathematical models fail to explain the data-generation process. (b) Observing the long-term variables interactions for the time-delayed, monotonic, and highly nonlinear relationships typically characterized for the industrial complexes. (c) The data-based information is extracted by the relevant statistical metrics to get an insight of the behavior of the industrial systems. The loss function of the neural network comprises of two loss terms, (i) typical modeling-based loss, and (ii) loss occurred due to the interpretation metric embedded in the loss function. (d) The data-metric(s) guided trained neural network and domain-constrained analytics can contribute to decarbonization of thermal power plants and support climate action by energy efficiency gains and reducing the emissions discharge.

3 Challenges and opportunities for data-driven models to support decarbonization of thermal power plants

Physics-informed Machie Learning (PIML) is a subdomain of ML that integrates physical laws into the architecture of the ML algorithm to train a model tailored for the application domain [59]. PIML-driven models can extrapolate the function profile and may require less data for training [60]. PIML shows remarkable promise for applications like modelling spring-mass systems, thermodynamic property estimation, and predicting turbulent fluid flow behaviour, where well-established and universally accepted physical laws are available. However, the development of mathematical models built on physical laws requires highly specific domain expertise and extensive parameter estimation of the model [61]. The multiobjective optimization analysis carried out with large mathematical models becomes computationally prohibitive for strategic-level performance enhancement for component-to-system-level operation excellence of chemical and process industries [62]. Moreover, as industrial systems undergo aging and wear and tear, the validation of the developed mathematical models becomes even more challenging [63], leading to inaccurate state estimation of industrial systems. For these potential issues of mathematical and PIML models, data-driven modelling, including ML, is an alternative approach that can train and/or update existing models flexibly. Data-driven modelling requires less computational power for model training, update and calibration, holding great promise for its utility for industrial-scale problems [64].

Although the potential of data-driven models for real-life problems is promising, we must address significant challenges in training a generalized and deployable ML model to support decarbonization of thermal power plants. Importantly, it is not just about collecting data and feeding it into an ML algorithm, as the adage "garbage in, garbage out" reminds us of for data-driven modeling [65]. The data obtained from the experiments often contain observational biases and confounders, which limit their interpretability. Similarly, sensor-based measurements can have intermittent, freezing, and drift errors, potentially introducing noise into the recorded observations.

Another challenge is to define the system boundaries and the scope of the problem for ML-based studies in terms of the selection of variables. During training, a deep learning model minimizes its loss by absorbing spurious correlations, including confounding factors and selection biases that adversely affect the generalization capability of the model. These connections, which have nothing to do with real causes, can change between training and application domains, which makes it difficult to do well with out-of-distribution data [66].

Some suggestions exist in the literature for the selection of the ML technique considering the nature of data distribution, the number of variables, and suitability of the algorithm for the quantitative and qualitative nature of the problem [67–69], and the selection of the ML technique for data-driven modeling remains an open question. Despite the selection and subsequent training of the ML model for predictive tasks, the interpretability assessment of commonly used modeling algorithms such as neural networks, random forests and gradient-boosted trees becomes less transparent. The lack of understanding of the predictive mechanisms of the parameter-intensive ML architectures hinders their deployment in industrial environment where human experts want to know the interpretability of ML models that is consistent with their domain knowledge [26]. The reduced transparency and/or explainability of complex ML models has historically resulted in their low adoption rate for the performance enhancement and energy-efficient process control of fossil-powered power systems, chemical, and process industries, which are the major drivers of climate change [18]. The interpretable techniques should be used on top of the ML-based workflow for the industrial decarbonization-related decision-making regarding operation management, balancing the trade-off between internal energy consumption and emissions discharge rate, and optimizing the techno-economic performance of industrial systems.

Another important challenge is to assess the uncertainty bounded with the point predictions made by the ML model. The point prediction envelopes the uncertainty that sources back to the modeling inaccuracy and the hyper-dimensional parameter space associated with the model structure [70]. The Bayesian and ensemble methods lack the validity issues of the constructed prediction intervals [71] while the transductive conformal prediction approach is computationally expensive to draw the valid prediction intervals, requiring the model to be retrained for the new input conditions [72].

Explicit prediction based on models, formally called XAI, is a hot research topic among research communities. The XAI approach seeks to investigate the transparency, justification, information, and reliability of model-based predictions [73]. The XAI techniques are categorized into feature attribution, instance-based, graph convolution-based, self-explaining and uncertainty estimation [74]. Counterfactual, attention mechanisms, human-readable rule extraction, adversarial explanation, transfer learning, integrating human feedback, etc., are some of the XAI categories used for explaining model-based predictions. These techniques, either ante-hoc or post-hoc, suffer owing to their working principles for explaining the model-based predictions for hyper-dimensional parameter space embedded in the ML models.

In [75], authors have provided guidelines to assess the scope of model-based explanation methods when designing new model explanations. Researchers are also exploring active learning, based on Gaussian processes, to enhance the interpretability of the model. However, these methods face significant challenges when assuming a Gaussian error distribution, which limits their applicability to complex real-world datasets. In-variances in model explanations for the input conditions demonstrate the ineffectiveness of the chosen XAI method for explanation tasks. The existing XAI methods can be independent of models and data-generating processes, and the XAI methods failing the tests designed on cascading randomization and independent randomization of model parameters are inadequate for the explanation tasks that are sensitive to data (detecting outliers) or model (explaining relationship between inputs and outputs).

Considering the applicability and scope of the XAI methods described earlier, it is apparent that a single explanation method is not an ideal technique for all types of ML models. Data-centric, interpretable modeling can improve the model-based explanation performance through innovation in data-driven constraints and their embeddings in the algorithmic design, holding a promise on providing domain-consistent prediction and explanations [32, 34]. The data-infused interpretability in the ML models and explained by the XAI methods can satisfy the regulatory requirements, debug the model, and reveal bias and inaccurate effects learned by the model that lead to informed decision-making for decarbonization of thermal power plants.

Table 1 summarizes the potential challenges associated with the development of data-driven workflows and provides mitigation strategies to implement ML-driven solutions for decarbonization of thermal power plants. Industrial systems are equipped with supervisory information systems and data banks that record historical operation data. Industrial units also rely on manual data compilation and data logging that should be digitalized to improve data curation, storage, and access to data-driven analytics. Even with access to data, data cleaning and preprocessing are the major steps that require statistical and domain expertise to prepare the final set of "cleaned" data to learn the underlying relationships and discover the data patterns. The right selection of the ML-based modeling algorithm trained on data-driven constraints, evaluating the modeling performance through relevant metrics, uncertainty quantification in the model-based predictions, and deciphering the model predictions via ante hoc or post hoc approaches are the major milestones in training a generalized ML model.

A very next and logical step after model development is to estimate the optimal set points / operating conditions on which component or system of a thermal power plant should be operated. Many research studies reported in the literature underscore the importance of ML-built optimization frameworks that carry out extensive data mining and leverage the rigor of mathematics and statistics to estimate optimal solutions for optimisation problems subjected to hierarchical levels of uncertainties surrounding industrial decarbonization (parameters, processes, systems, enterprises, clusters, planetary boundaries) [76–79]. However, as the ML models are put into operation, the models need to be updated regularly since the production data may undergo distribution shift and model performance deteriorates. The frequency of model adaptation may require human intervention or can be performed autonomously as long as the threshold on distribution shift is crossed. Moreover, ML-based optimization analysis should be carried out with domain-constrained framework that provides domain-consistent solutions [80]. However, the right choice of data-driven constraints formation and their satisfaction during solving an optimization problem remains challenging. Another challenge of ML operations in the industrial environment is the vulnerability of models to cyber-attacks, data theft, and system security, which requires the addition of a security layer for the smooth and uninterrupted operation of industrial systems towards industrial decarbonization.

Despite the spectrum of challenges bound with data-driven analysis workflow, the ML models trained under extensive data pre-processing, feature engineering, and rigorous hyperparameter tuning have demonstrated practical utility for applications such as power systems [81–86] and industrial cooling and heating systems [87, 88]. The predictive accuracy of the ML models is improved under adaptive training as the data generation pattern and process change in real-life applications. A well-predictive ML model enables the development of informed operating strategies and decision support systems, thereby contributing to energy efficiency gains in thermal power plants. As industrial systems become increasingly digital, data-driven science is used to improve the design, operation, and control of thermal power plants. Data-driven science can estimate robust and energy-efficient operating conditions for existing thermal power plants' operation to reduce emissions discharge without significant retrofit solutions [81, 89] that help industry act on climate change in the face of economic and societal challenges.

Table 1: Challenges that exist for the data-driven analytics integrating ML models and optimisation techniques for the performance enhancement of industrial systems.

Scope	Challenges	Mitigation
Data Curation & Access	• Manual data collection from the on-site systems • Temporal inconsistencies of data-recording for multi-level operation • Inadequate data recording and storage facilities • Human error and bias in the measurement • Data loss • No access to data storage or partial data retrieval at improper data sampling rate	• Digitalization of the data recording • Developing strategies to reduce the data sampling frequencies through advanced metering and measurement systems • Implementing secure data storage systems and maintaining the data-logs of the industrial systems • Integrating advanced supervisory information system for data extraction at the desirable sampling rate
Data processing	• Data distribution is skewed and imbalanced • Sensor-measurements induced errors • Feature selection for the data-driven modeling • Data scaling for the modeling tasks	• Deploying techniques to sample the data from big-clusters for balancing the data-distribution across the operating ranges • Applying outliers' removal techniques and visualize the data to identify the potential outliers [84] • Consulting with domain-experts and applying statistical techniques for feature selection and dimensionality reduction

Continued on next page

Table 1 – continued from previous page

Scope	Challenges	Mitigation
ML model development	• Selection of the ML model corresponding to the operating level (component, system, strategic) of the industrial systems • Rigorous model evaluation through robust metrics on the unseen data (data sample from the industrial operation) • Estimating the parametric uncertainty and investigating its impact on the model-based prediction and explanation	• Neural network-based algorithms for the quantitative nature of the data and classification-based algorithms for the qualitative nature of the dataset [90] • Apply prediction interval estimation techniques (conformal prediction, direct interval estimation) for the uncertainty quantification [70, 91] and XAI methods for model-based explanation
Model-based optimisation under uncertainty	• The non-convex optimisation problem integrating the ML models poses computational issues to estimate a solution • The solver(s) may converge to a feasible solution that may not align well with the domain-knowledge • Introducing uncertainty and its evaluation for model-based optimisation is challenging and computationally expensive	• Introducing data-information-based constraints in the optimisation problem that guide the solver(s) to estimate an efficient solution • Development of scalable and computationally effective tools for ML model-based optimisation under uncertainty for large-scale industrial data problems

Continued on next page

Table 1 – continued from previous page

Scope	Challenges	Mitigation
Model calibration and security	• Consistent model-calibration requires experts for in-house analysis • Cyber-attacks and model security • Self model adaptation and tuning	• Robust algorithm design for the self-model calibration with human-in-the-loop to monitor the model deterioration and accuracy • Secure model deployment environment against cyber attacks

4 Algorithm design innovation aligned with industrial decarbonization

The need for innovations in algorithm design to train reliable models for thermal power plants is clear. Information and communication technologies record vast amounts of real-time data from industrial operations. These data sets possess information on the dynamics of the system accumulated over years of industrial system operation, as shown graphically in Figure1(b). In particular, the data sets capture the interactions and non-linear dynamics at the intersection of interdisciplinary knowledge domains such as chemical and / or biochemical processes, material science, and electromechanical characterization that essentially represent the "physics" of the'specific' system [40]. This provides an opportunity to incorporate system physics into algorithmic design, similar to how the physics-informed neural network (PINN) in particular and the physics-informed machine learning (PIML) in general update the model parameters during the learning phase. The predictions made by PINN or PIML are consistent with those of the physical laws incorporated into the loss function (L) or embedded as constraints in the algorithm [59]. Thus, there is a clear need to use existing or innovative novel statistical terms that accurately quantify the strength of the feature associations established in the big data of thermal power plants to make better interpretable models.

The commonly used statistical metrics for quantifying the association between the two variables considering their nature (continuous, categorical, mixed) are mentioned in Table 2. Pearson's correlation coefficient (r) and Spearman's Rank Correlation Coefficient (ρ) quantify linear and monotonic relationships, respectively, between two continuous variables. Generally, the variables involved with the control of power generation processes are continuous in nature and therefore r or ρ can be used for the development of customized algorithms, similar to Data Information Integrated Neural Network (DINN) as reported in [34]. A Pearson correlation-based deviation constraint on the actual and model-based predicted response is developed and integrated in the loss function (L) of DINN for the parameter tuning. Similarly, the association between categorical or mixed nature of variables can be quantified, for instance, by chi-square (χ^2) and Point-biserial Correlation Coefficient (r_{pb}) and later these metrics can be used for the development of customized algorithms for modeling industrial-scale problems within the context of decision making for industrial decarbonization. More details about the nature of variables and the properties of the metrics are described in Table 2.

Methodologically, there is an opportunity to leverage the 'physics' that exists in the data-driven feature-relationships in the form of the statistical data metric, quantifying the association between the variables. Similar to PINN, ML models can integrate relevant statistical metrics into L, or as a constraint in the algorithm or as a stopping condition that contributes to updating the parameter with data information during model training. For example, the customised L integrating r and the parameter update by gradient descent with momentum algorithm for DINN are given as [34]:

$$L = a_1 \cdot \frac{\sum(D - Z)^2}{N} + a_2 \cdot \frac{\sum_{i=1}^{m}(r_{X_i|D} - r_{X_i|Z})^2}{N} \tag{1}$$

$$W_1^{\text{new}} = W_1 - \eta V_{W_1} \tag{2}$$

$$V_{W_1} = \beta V_{W_1} + (1 - \beta)\frac{\partial L}{\partial W_1} \tag{3}$$

$$\frac{\partial L}{\partial W_1} = \frac{\partial L}{\partial Z}\frac{\partial Z}{\partial p_2}\frac{\partial p_2}{\partial y_1}\frac{\partial y_1}{\partial p_1}\frac{\partial p_1}{\partial W_1} \tag{4}$$

Here, D and Z are the true and simulated response of the model, N is the total data observations, m is the total number of input variables, $r_{X_i|D}$ and $r_{X_i|Z}$ are the Pearson correlation coefficients calculated with respect to the input variables (X_i) in D and Z, respectively. a_1 and a_2 are the weighting factors; W_1 is the matrix containing the weight connections from input to hidden layer of the DINN algorithm, V_{W_1} is the velocity matrix defined on W_1; η and β are the learning rate and momentum parameters, respectively; and $\frac{\partial L}{\partial W_1}$ is the partial derivative of L with respect to W_1 calculated by equation. (4) and contributes to the weight update (W_1^{new}). The rest of the model parameters (W_2, b_1, b_2) are also updated by the customized L during iterative model training. The graphical representation of the information flow and processing in DINN is presented in Figure 2.

It is important to note here that the custom L defined for DINN updates the parameters through chain rule and error backpropagation; the custom L or the embedding of constraints driven by statistical metrics in any ML algorithm architecture can contribute to update the model parameters through optimization solvers. The incorporation of data patterns quantified through statistical metrics in ML algorithms, either in loss function or as constraints, may contribute to constructing effective functional mapping and developing true relationships between the variables (input-input and input-output). The accurate function approximation through consistent variable associations in the trained ML model provides reliable model-based explanations when coupled with XAI techniques including SHAP, LIME or Monte Carlo simulations, which also rely on functional mappings that exist in the trained ML model. In [34], DINN-based explanations obtained through Monte Carlo simulations have been found to be better domain-compliant than those of the standard artificial neural network to predict the operating characteristics of industrial-scale gas turbine system. Integrating data information via statistical metrics or constraints, which represents the operating behavior of the industrial system, into ML algorithms appears to be a flexible and intuitive approach that holds a promise to improve the interpretable performance of ML algorithms that can facilitate decision making for industrial decarbonization of thermal power plants.

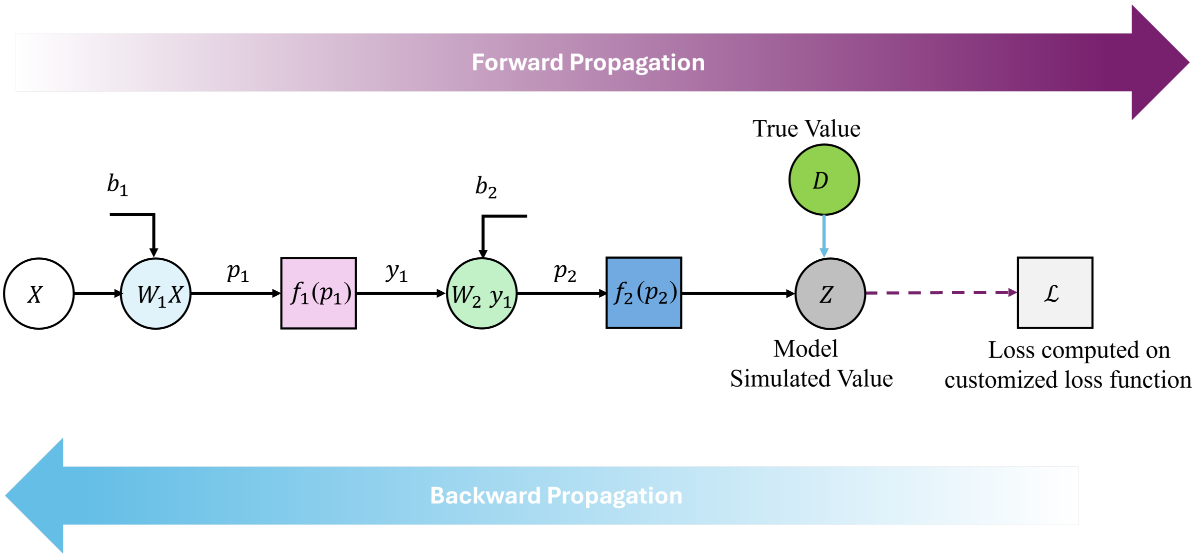

Figure 2: The information flow in the architecture of data information integrated neural network (DINN) algorithm. The parameters are updated iteratively on the customized loss function. (Figure adapted from [34]).

Table 2: Leveraging the association of the variables (continuous, categorical, mixed) for the development of interpretable ML model.

Nature of Variables	Metrics	Function
Continuous	Pearson Correlation Coefficient (r): $$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$ x and y are two variables with the mean of $\bar{x}$ and $\bar{y}$ respectively	• Measures the linear relationship between the two variables • Ranges from -1 to +1
	Spearman's Rank Correlation Coefficient (ρ): $$\rho = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)}$$ d is the deviation between the ranks of two variables for each data pair; n is the total number of data points	• Measures the strength and direction of the monotonic relationship between two ranked (ordinal) variables • Ranges from -1 to 1

Continued on next page

Table 2 – continued from previous page

Nature of Variables	Metrics	Function
	Coefficient of determination (R^2): $R^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$	• Represents the proportion of the variance in the dependent variable that is predictable from the independent variable • Ranges from 0 to 1
Categorical	Chi-Squared (χ^2): $\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$ O and E represent the observed and expected frequencies to observe the values of the categorical variable	• Provides a p-value indicating the significance of the association
	Cramér's V: $\sqrt{\frac{\chi^2}{n(k-1)}}$ n is the number of observations and k is the smaller of the number of levels of the two variables	• Measures the strength of association between two categorical variables • Ranges from 0 to 1
	Goodman and Kruskal's lambda (λ): $\lambda = \frac{E_1 - E_2}{E_1}$ • E_1 is the overall non-modal frequency • E_2 is the sum of the non-modal frequencies for each value of the independent variable	• Measures the strength of association between two categorical variables, indicating the proportion of error reduced in predicting one variable using the other • Ranges from 0 to 1

Continued on next page

https://doi.org/10.33774/coe-2024-3r5v0-v2 **ORCID:** https://orcid.org/0000-0003-0727-5683 Content not peer-reviewed by Cambridge University Press.

Table 2 – continued from previous page

Nature of Variables	Metrics	Function
Mixed	Point-biserial Correlation Coefficient (r_{pb}): $$r_{pb} = \frac{\bar{X}_1 - \bar{X}_0}{s_X}\sqrt{\frac{n_1 n_0}{n(n-1)}}$$ where, • $\bar{X}_1$ is the mean of the continuous variable for the group where the binary variable is 1 • $\bar{X}_0$ is the mean of the continuous variable for the group where the binary variable is 0 • s_X is the standard deviation of the continuous variable • n_1 is the number of observations where the binary variable is 1 • n_0 is the number of observations where the binary variable is 0 • n is the total number of observations	• Measures the strength and direction of association between a continuous variable and a binary variable • Ranges from -1 to 1
	Eta-squared (η^2): $$\eta^2 = \frac{SS_{effect}}{SS_{total}}$$ where, • SS_{effect} is the sum of squares for the effect • SS_{total} is the total sum of squares for all effects, errors and interactions	• Measures the proportion of variance in a continuous dependent variable that is associated with one or more categorical independent variables • Ranges from 0 to 1

5 Way forward: ML for decarbonization of thermal power plants

In the previous sections, we describe the potential opportunities as well as challenges in training domain-guided ML models and implementing the ML-based solutions in the industrial environment of thermal power plants. We also discuss that the domain of thermal power plants can be quantified by statistical metrics. The domain-constrained, machine learning–based optimization framework represents a promising paradigm in data-driven science, with the potential to revolutionize the design, operation, and control of thermal power plants, enhancing energy efficiency and reducing emissions to support decarbonization [18] (refer to Figure 3). We discuss the potential of domaine-constrained ML analytics in the design, operation and control phases of thermal power plants in the following.

5.1 ML for Component / System Design

The optimal design of components and systems is critical to maintain the energy-efficient and economic operation of industrial thermal power plants. Historically, the design of industrial components like pumps, compressors, bearings, impellers etc., is made by rigorous computer simulations, which are computationally expensive as well as a slow process. However, this design paradigm has ushered in rich design datasets, which are valuable and can accelerate the component design through a data-constrained approach. The design engineer can deploy the available design data as constraints for the training of the ML model and / or the hybrid modeling approach [92, 93]. The data-constrained driven trained models predict design conditions that not only reflect the design geometry available in the historical datasets, but can also optimize the design features of industrial components for the given design specification of components [94, 95].

The data-driven constraints can sample the design conditions for the desired specification of the new component from the joint distributions of the design features made on the historical data of the design conditions [96, 97]. The data-constraints-informed ML model can be embedded in the design-optimisation problem [98], and the optimal design specifications of the new component can be iteratively sampled from the joint distributions of the design features such that the desirable specification of the component has been estimated [99]. The joint distribution of the design features can be established through various sampling techniques, including copula sampling, Markov chain Monte Carlo, Gaussian mixture modelling, generative models, etc. These sampling techniques capture the nonlinear and complex joint interaction of the design features, which are useful to optimize the design specification of the new component [100]. Similarly, rule-informed ML models and rules-compliant generative AI models including large language models are also revolutionizing the design of industrial processes and systems through the automatic generation of process drawings and diagrams [101] and extraction of the design specification from the system design [102]. The optimal design of components and systems, determined by data-constrained ML algorithms and optimization frameworks, can offer higher energy efficiency and optimal utilization of resources that improve the techno-economic performance. This reduces the material waste and emissions intensity from the optimally designed components and / or systems of the thermal power plants to support industrial decarbonization, as highlighted on Figure 3.

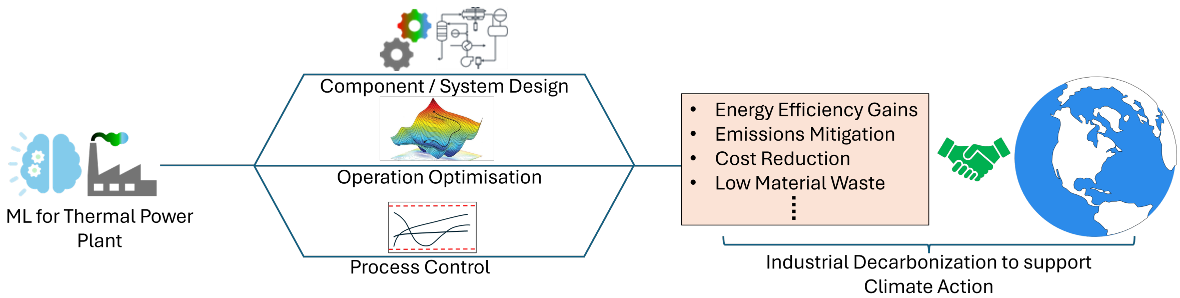

Figure 3: Role of domain-constrained ML for component/system design, operation optimisation and process control of thermal power plants to support decarbonization and climate action.

5.2 ML for Operation Optimization

Recently, ML is being used for the system-level modeling and optimization of industrial systems, including process industries [103, 104], power plants [105, 106], energy grids [26][107], batteries [108] and transportation systems [109].The operation of the thermal power plants can be analyzed with the view of minimization of resource consumption and multi-objective optimization of the techno-economic and environmental performance. Leveraging the excellent modeling capacity of ML models and embedding the data-driven constraints, discovered and/or constructed from the industrial datasets, significantly improves the solution quality and its practical implementation towards the operation excellence of thermal power plants [80, 110].

Table 3 depicts some data-driven multivariate constraints that can be built on the data and are embedded in the optimization problem. Principal component analysis (PCA) determines a number of principal components that explain the maximum variability of the data. PCA-based constraint limits the variation of the optimal solution in the orthogonal direction of PCA-constructed subspace $(I - P_k P_k^T)$ up to τ, and contributes to preserving the correlation structure in the data. Similarly, Mahalanobis distance-based constraints allow the deviation of the optimal solution up to τ from the mean of the data distribution. Both PCA and Mahalabonis distance-based constraints guide the optimization solvers to preserve the linear correlation structure of the input variables; the nonlinear interactions are not well captured by these two constraints.

On the other hand, the empiric cumulative distribution function (ECDF)-based constraint minimizes the local deviation between the pair of variables up to τ for estimating the solution. A major advantage of ECDF-based constraint is that it is non-parametric and is established on the empiric data distribution, unlike the Mahalanobis distance which assumes Gaussian distribution of the data. A key limitation of the pairwise ECDF-based constraints is the scalability for higher input dimensions and conflicts between the overlapping pairs. Partial least squares (PLS) method explains the variation in the target variable with respect to variation in the input variables. Hotelling T^2 computes the distance of the PLS scores from the mean of latent PLS space and can be incorporated in the objective function of the optimization problem as a soft penalty approach. The Hotelling T^2 guides the optimization solvers to estimate the solutions that remain within the operability limits of the system. PLS with Hotelling T^2 captures the linear correlation in the variables of the power generation process of industrial thermal power plants.

It is important to mention here that the data-driven multivariate constraints are not developed as "hard" constraints, but a tolerance (τ) is provided for the optimization solver to explore nonlinear interactions between the trained ML model and the constraints. Tuning of τ remains a challenge for the estimation of effective solutions. The multivariate constraints, as mentioned in Table 3, can be embedded in the multi-objective optimization problem, formulated for optimizing the key performance variables of thermal power plants. The formulated multi-objective optimization problem can aim at minimizing the fuel consumption rate for the set value of power generation given that energy efficiency of the process has been maximized. The data-driven constraints guardrail the optimization solver against the violation of process dynamics and contribute towards the estimation of domain-consistent optimal solutions. The data-driven domain-constrained optimization analytics for thermal power plants are rarely reported in the literature. However, the domain-compliant AI solutions obtained from domain-constrained optimization problems increase the energy efficiency of thermal power plants via optimal fuel consumption rate that can significantly reduce the cost of operation of thermal power plants [83, 111]. The reduction in power generation process-related emissions through AI-based energy efficiency gains supports the decarbonization of thermal power plants [18] (refer to Figure 3), and is emerging an a practical approach to strengthen the climate action from the power sector.

Table 3: Data-driven constraints for embedding multivariate relationships in optimization problems.

Constraint Name	**Mathematical Expression**	**Advantages and Limitations**
Principal Component Analysis	$\left\|(I - P_k P_k^T)(x-\mu)\right\|_2^2 \leq \tau$ x is input, μ is mean, P_k contains top-k eigenvectors, τ is tolerance	+ Captures dominant linear structure + Reduces dimensionality – Ignores nonlinear dependencies – Discards low-variance directions that may be important
Mahalanobis Distance	$\sqrt{(x-\mu)^T\Sigma^{-1}(x-\mu)} \leq \tau$ μ is mean vector, Σ is covariance matrix	+ Preserves full multivariate Gaussian-like shape + Simple to compute when Σ^{-1} exists – Assumes Gaussian distribution – Computationally expensive in higher dimensions
Empirical Cumulative Distribution Function (ECDF)	$\lvert \mathrm{ECDF}_i(x_i) - \mathrm{ECDF}_j(x_j) \rvert \leq \tau$ $F_i^{\min} \leq \mathrm{ECDF}_i(x_i) \leq F_i^{\max}$	+ Fully non-parametric + Captures complex local patterns – Non-differentiable (unless smoothed) – Hard to scale in high dimensions for pairwise control
		Continued on next page

Table 3 – continued from previous page

Constraint Name	Mathematical Expression	Advantages and Limitations
Partial Least Squares with Hotelling's T^2	$T^2 = t^T \Lambda^{-1} t \leq \tau, \quad t = W^T x$ t is latent score, W is weight matrix, Λ is score variance	+ Captures input-output latent structure + Enforces process-operating bounds – Limited to linear latent relations – Requires response variable for training

5.3 ML for Control Applications

Data-driven control [112] and learning from the data [113] are promising domains in control systems that leverage data-driven models for the control of the systems. The problem-specific data metrics (mentioned in Table 2) and data-driven constraints (refer to Table 3) can approximate the performance of the existing controllers being used in the power plants [114]. Later, the approximated controllers can estimate the control actions that are domain-consistent and enhance the system stability and control performance [80, 115]. The promising utility of ML-based control is to forecast the optimal control trajectories for the dynamic operation of thermal power plants that the control system should follow under human supervision to realize the optimal operation control of thermal power plants[116]. The domain-constrained ML model-based control systems can estimate energy-efficient control actions to limit emissions discharge to allowable limits during the operation of power systems [117, 118], energy grids [119, 120], carbon capture technologies [121], and battery systems [122], thus reducing emission volumes to support decarbonization and climate action from the power sector (shown on Figure 3).

5.4 Adopting AI in industrial environment

Machine learning algorithms are capable of uncovering complex patterns within high-dimensional, large-scale datasets of thermal power plants. Engineers and industrial managers can harness this capability to enable energy-efficient and cost-effective design, operation, and control of industrial systems—especially thermal power plants. Embedding the data-driven domain in the ML algorithm design and optimization problems to effectively model and analyze the system provides domain-compliant optimal solutions. This aspect of AI-based analytics is important for safety-critical industrial operation of thermal power plants. The validity of industrial domain ensured during data-driven analytics enhances the confidence and trust of industrial professionals, and contributes to AI adoption in industrial environment [123].

Furthermore, data-led ML model development and its deployment in domain-constrained control structures lay the foundation for the construction of digital twins that should be robust, resilient and safe for power system applications. Domain-constrained, ML-driven analytics enable timely adaptation in response to adverse events, supporting informed decision-making for the technical, economic, and environmental management of thermal processes and systems, advancing the decarbonization of thermal power plants.

Concluding Remarks

As we begin to live in the fourth learning paradigm-data-driven sciences, it becomes crucial to extend the reliable application of machine intelligence for the modeling and optimization of industrial power systems to support industrial decarbonization. ML-based models offer a flexible approach to modeling the large-scale operation of thermal power plants, as opposed to first-principle modeling, and the data-driven models can be adapted flexibly to predict the state of thermal power plants.

Industrial thermal power plants store a massive amount of data on years of operation, and the representative association between variables is almost established. Thus, the feature association as computed on the data metrics represents the domain quantification of the system, which can be introduced in the ML algorithms, similar to how PINN works. The integration of the data metrics into the loss function or as constraints in the ML algorithms can update the model parameters well, and the trained model offers improved interpretability performance [34]. As a result, the trained ML model can be deployed in a data-driven domain-constrained optimization problem for the optimal design, operation, and control of thermal power plants. This enhances the performance of thermal power plants from the dimensions of energy efficiency improvement, emissions reduction, optimal resource utilization to developing robust decision support system and effective operation management to contribute to industrial decarbonization and support climate action.

On the one hand, AI drives the higher performance of thermal power plants, it also contributes to the reduction of GHGs (2.4 Gt) and the addition of value in the economy (USD 5.2 Trillion by 2030) [19]. Therefore, it is anticipated that ML-based industrial analytics backed by domain expertise and exploiting the rich industrial data can accelerate the AI adoption in thermal power plants in particular and in industrial settings in general that enable a sustainable future against the climate crisis.

Acknowledgement

WMA acknowledges to have received the funding (CMMS-PHD-2021-006) from The Punjab Education Endowment Fund (PEEF) to pursue PhD at University College London. WMA acknowledges that this work was supported by The Alan Turing Institute's Enrichment Scheme/Turing Studentship Scheme. RD acknowledges the support of the Bill and Melinda Gates Foundation [OPP1144], Cambridge Humanities Research Grants (2023) and UKRI Responsible AI IO0008 Grant Ref: EP/Y009800/1.

Author contributions

W.M.A. initiated the project and convened the team. W.M.A., V.D., and R.D. conceptualized the study. W.M.A wrote the first draft of the manuscript with inputs from R.D. and V.D. All authors read, revised, and approved the manuscript.

Competing Interests

The authors declare that they have no competing interests to disclose in carrying out this study.

References

1. Creutzig, F. *et al.* Digitalization and the Anthropocene. *Annual review of environment and resources* **47,** 479–509 (2022).

2. Debnath, R. *et al.* Harnessing human and machine intelligence for planetary-level climate action. *npj Climate Action* **2,** 20 (2023).

3. Coyle, D. & Weller, A. "Explaining" machine learning reveals policy challenges. *Science* **368,** 1433–1434 (2020).

4. Inderwildi, O. & Kraft, M. *Intelligent Decarbonisation: Can Artificial Intelligence and Cyber-Physical Systems Help Achieve Climate Mitigation Targets?* (Springer Nature, 2022).

5. Chen, Z., Wang, C. & Bai, F. Greenhouse gas emissions and global real economic activities. *Finance Research Letters* **64,** 105404 (2024).

6. Bhatti, U. A. *et al.* Global production patterns: Understanding the relationship between greenhouse gas emissions, agriculture greening and climate variability. *Environmental Research* **245,** 118049 (2024).

7. Jafari, M., Botterud, A. & Sakti, A. Decarbonizing power systems: A critical review of the role of energy storage. *Renewable and Sustainable Energy Reviews* **158,** 112077 (2022).

8. Buck, H. J. *et al.* Adaptation and carbon removal. *One Earth* **3,** 425–435 (2020).

9. Victoria, M. *et al.* Early decarbonisation of the European energy system pays off. *Nature communications* **11,** 6223 (2020).

10. Upham, P., Sovacool, B. & Ghosh, B. Just transitions for industrial decarbonisation: A framework for innovation, participation, and justice. *Renewable and Sustainable Energy Reviews* **167,** 112699 (2022).

11. Sovacool, B. K., Del Rio, D. F. & Zhang, W. The political economy of net-zero transitions: Policy drivers, barriers, and justice benefits to decarbonization in eight carbon-neutral countries. *Journal of Environmental Management* **347,** 119154 (2023).

12. Chen, P. *et al.* The heterogeneous role of energy policies in the energy transition of Asia–Pacific emerging economies. *Nature Energy* **7,** 588–596 (2022).

13. IEA. *An energy sector roadmap to carbon neutrality in China* `https://www.iea.org/reports/an-energy-sector-roadmap-to-carbon-neutrality-in-china` (2021).

14. Rolnick, D. *et al.* Tackling climate change with machine learning. *ACM Computing Surveys (CSUR)* **55,** 1–96 (2022).

15. Lee, E. A. *Cyber physical systems: Design challenges* in *2008 11th IEEE international symposium on object and component-oriented real-time distributed computing (ISORC)* (2008), 363–369.

16. Bhutani, N., Rangaiah, G. & Ray, A. First-principles, data-based, and hybrid modeling and optimization of an industrial hydrocracking unit. *Industrial & engineering chemistry research* **45,** 7807–7816 (2006).

17. Inderwildi, O., Zhang, C. & Kraft, M. in *Intelligent Decarbonisation: Can Artificial Intelligence and Cyber-Physical Systems Help Achieve Climate Mitigation Targets?* 17–28 (Springer, 2022).

18. IEA. Energy and AI. *IEA, Paris https://www.iea.org/reports/energy-and-ai, Licence: CC BY 4.0* (2025).

19. Joppa, L. & Herweijer, C. *How AI can enable a Sustainable Future* Technical Report (Microsoft, PwC UK, 2019). `https://news.microsoft.com/wp-content/uploads/prod/sites/53/2019/04/PwC-Executive-Summary.pdf`.

20. Doshi-Velez, F. & Kim, B. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608* (2017).

21. Miller, T. Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence* **267,** 1–38 (2019).

22. Linardatos, P., Papastefanopoulos, V. & Kotsiantis, S. Explainable ai: A review of machine learning interpretability methods. *Entropy* **23,** 18 (2020).

23. Liu, X. *et al.* The medical algorithmic audit. *The Lancet Digital Health* **4,** e384–e397 (2022).

24. Murdoch, W. J. *et al.* Definitions, methods, and applications in interpretable machine learning. *Proceedings of the National Academy of Sciences* **116,** 22071–22080 (2019).

25. Lisboa, P. J. *et al.* The coming of age of interpretable and explainable machine learning models. *Neurocomputing* **535,** 25–39 (2023).

26. Cremer, J. L., Konstantelos, I. & Strbac, G. From optimization-based machine learning to interpretable security rules for operation. *IEEE Transactions on Power Systems* **34,** 3826–3836 (2019).

27. Kumar, R. & Singh, A. K. Chemical hardness-driven interpretable machine learning approach for rapid search of photocatalysts. *npj Computational Materials* **7,** 197 (2021).

28. Huang, J. *et al.* Optimisation led energy-efficient arsenite and arsenate adsorption on various materials with machine learning. *Water Research,* 122815 (2024).

29. Ashraf, W. M. & Dua, V. Partial derivative-based dynamic sensitivity analysis expression for non-linear auto regressive with exogenous (NARX) modelcase studies on distillation columns and model's interpretation investigation. *Chemical Engineering Journal Advances* **18,** 100605 (2024).

30. Roh, J. *et al.* Interpretable machine learning framework for catalyst performance prediction and validation with dry reforming of methane. *Applied Catalysis B: Environmental* **343,** 123454 (2024).

31. Zhao, S. *et al.* Interpretable machine learning for predicting and evaluating hydrogen production via supercritical water gasification of biomass. *Journal of Cleaner Production* **316,** 128244 (2021).

32. Xu, H. *et al.* Biogeochemistry-informed neural network (binn) for improving accuracy of model prediction and scientific understanding of soil organic carbon. *arXiv preprint arXiv:2502.00672* (2025).

33. Dash, T. *et al.* A review of some techniques for inclusion of domain-knowledge into deep neural networks. *Scientific Reports* **12,** 1040 (2022).

34. Ashraf, W. M. & Dua, V. Data Information integrated Neural Network (DINN) algorithm for modelling and interpretation performance analysis for energy systems. *Energy and AI* **16,** 100363 (2024).

35. Khanzadeh, S. *et al.* An exploratory study on domain knowledge infusion in deep learning for automated threat defense. *International Journal of Information Security* **24,** 71 (2025).

36. Cheng, J., Luo, X. & Jin, Z. Integrating domain knowledge into transformer for short-term wind power forecasting. *Energy* **312,** 133511 (2024).

37. Heyrani Nobari, A., Regenwetter, L. & Ahmed, F. *Towards Domain-Adaptive, Resolution-Free 3D Topology Optimization With Neural Implicit Fields* in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* **88360** (2024), V03AT03A012.

38. Barrett, J. *et al.* Energy demand reduction options for meeting national zero-emission targets in the United Kingdom. *Nature Energy* **7,** 726–735 (2022).

39. Karagiorgi, G. *et al.* Machine learning in the search for new fundamental physics. *Nature Reviews Physics* **4,** 399–412 (2022).

40. Pantelides, C. C. & Renfro, J. G. The online use of first-principles models in process operations: Review, current status and future needs. *Computers & Chemical Engineering* **51,** 136–148 (2013).

41. Strejc, V. Mathematical Modelling of Large Scale Systems (General Approach). *IFAC Proceedings Volumes* **20,** 1–9 (1987).

42. Sreepradha, C., Panda, R. C. & Bhuvaneswari, N. S. Mathematical model for integrated coal fired thermal boiler using physical laws. *Energy* **118,** 985–998 (2017).

43. Pena, R., Masada, G. & Flake, R. A Boiler-Turbine Mathematical Model for Power Plant Operation Studies. *IFAC Proceedings Volumes* **20,** 681–686 (1987).

44. Bystritskaya, E., Pomerantsev, A. L. & Rodionova, O. Y. Prediction of the aging of polymer materials. *Chemometrics and intelligent laboratory systems* **47,** 175–178 (1999).

45. Bishnu, S. K., Alnouri, S. Y. & Al-Mohannadi, D. M. Computational applications using data driven modeling in process Systems: A review. *Digital Chemical Engineering* **8,** 100111 (2023).

46. Uddin, G. M. *et al.* Artificial intelligence-based Monte-Carlo numerical simulation of aerodynamics of tire grooves using computational fluid dynamics. *Ai Edam* **33,** 302–316 (2019).

47. Tariq, R., Abatal, M. & Bassam, A. Computational intelligence for empirical modeling and optimization of methylene blue adsorption phenomena using available local zeolites and clay of Morocco. *Journal of Cleaner Production* **370,** 133517 (2022).

48. Tariq, R. *et al.* Deep learning artificial intelligence framework for sustainable desiccant air conditioning system: Optimization towards reduction in water footprints. *International Communications in Heat and Mass Transfer* **140,** 106538 (2023).

49. Ashraf, W. M. *et al.* Artificial intelligence based operational strategy development and implementation for vibration reduction of a supercritical steam turbine shaft bearing. *Alexandria Engineering Journal* **61,** 1864–1880 (2022).

50. Libbrecht, M. W. & Noble, W. S. Machine learning applications in genetics and genomics. *Nature Reviews Genetics* **16,** 321–332 (2015).

51. Angermueller, C. *et al.* Deep learning for computational biology. *Molecular systems biology* **12,** 878 (2016).

52. Shehab, M. *et al.* Machine learning in medical applications: A review of state-of-the-art methods. *Computers in Biology and Medicine* **145,** 105458 (2022).

53. Kalinin, S. V. *et al.* Physics is the new data. *arXiv preprint arXiv:2204.05095* (2022).

54. Haibe-Kains, B. *et al.* Transparency and reproducibility in artificial intelligence. *Nature* **586,** E14–E16 (2020).

55. Ouyang, R. *et al.* SISSO: A compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates. *Physical Review Materials* **2,** 083802 (2018).

56. Kaptanoglu, A. A. *et al.* Pysindy: A comprehensive python package for robust sparse system identification. *arXiv preprint arXiv:2111.08481* (2021).

57. Cranmer, M. *et al.* Discovering symbolic models from deep learning with inductive biases. *Advances in neural information processing systems* **33,** 17429–17442 (2020).

58. Liu, Z. *et al.* Kan: Kolmogorov-arnold networks. *arXiv preprint arXiv:2404.19756* (2024).

59. Karniadakis, G. E. *et al.* Physics-informed machine learning. *Nature Reviews Physics* **3,** 422–440 (2021).

60. Kashinath, K. *et al.* Physics-informed machine learning: case studies for weather and climate modelling. *Philosophical Transactions of the Royal Society A* **379,** 20200093 (2021).

61. Mohamed, O. *et al.* Mathematical modelling for coal fired supercritical power plants and model parameter identification using genetic algorithms. *Electrical Engineering and Applied Computing,* 1–13 (2011).

62. Papadrakakis, M. *et al.* Large scale structural optimization: computational methods and optimization algorithms. *Archives of computational Methods in Engineering* **8,** 239–301 (2001).

63. Pfenninger, S., Hawkes, A. & Keirstead, J. Energy systems modeling for twenty-first century energy challenges. *Renewable and sustainable energy reviews* **33,** 74–86 (2014).

64. Larrañaga, P. *et al.* *Industrial applications of machine learning* (CRC press, 2018).

65. Geiger, R. S. *et al.* *Garbage in, garbage out? Do machine learning application papers in social computing report where human-labeled training data comes from?* in *Proceedings of the 2020 conference on fairness, accountability, and transparency* (2020), 325–336.

66. Nagarajan, V., Andreassen, A. & Neyshabur, B. Understanding the failure modes of out-of-distribution generalization. *arXiv preprint arXiv:2010.15775* (2020).

67. Raschka, S. Model evaluation, model selection, and algorithm selection in machine learning. *arXiv preprint arXiv:1811.12808* (2018).

68. Doan, T. & Kalita, J. *Selecting machine learning algorithms using regression models* in *2015 IEEE International Conference on Data Mining Workshop (ICDMW)* (2015), 1498–1505.

69. Mahesh, B. Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).[Internet]* **9,** 381–386 (2020).

70. Shafer, G. & Vovk, V. A tutorial on conformal prediction. *Journal of Machine Learning Research* **9** (2008).

71. Dewolf, N., Baets, B. D. & Waegeman, W. Valid prediction intervals for regression problems. *Artificial Intelligence Review* **56,** 577–613 (2023).

72. Saunders, C., Gammerman, A. & Vovk, V. Transduction with confidence and credibility (1999).

73. Haque, A. B., Islam, A. N. & Mikalef, P. Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. *Technological Forecasting and Social Change* **186,** 122120 (2023).

74. Islam, M. U. *et al.* The past, present, and prospective future of xai: A comprehensive review. *Explainable Artificial Intelligence for Cyber Security: Next Generation Artificial Intelligence,* 1–29 (2022).

75. Adebayo, J. *et al.* Sanity checks for saliency maps. *Advances in neural information processing systems* **31** (2018).

76. Zheng, X. *et al.* A multi-scale time-series dataset with benchmark for machine learning in decarbonized energy grids. *Scientific Data* **9,** 359 (2022).

77. Peng, J. *et al.* Human-and machine-centred designs of molecules and materials for sustainability and decarbonization. *Nature Reviews Materials* **7,** 991–1009 (2022).

78. Li, P.-H. *et al.* Revealing effective regional decarbonisation measures to limit global temperature increase in uncertain transition scenarios with machine learning techniques. *Climatic Change* **176,** 80 (2023).

79. Guo, J. *et al.* A merged continental planetary boundary layer height dataset based on high-resolution radiosonde measurements, ERA5 reanalysis, and GLDAS. *Earth System Science Data* **16,** 1–14 (2024).

80. Ashraf, W. M. *et al.* Domain-Informed Operation Excellence of Gas Turbine System with Machine Learning. *arXiv preprint arXiv:2507.08697* (2025).

81. Ashraf, W. M. & Dua, V. Artificial intelligence driven smart operation of large industrial complexes supporting the net-zero goal: Coal power plants. *Digital Chemical Engineering* **8,** 100119 (2023).

82. Ashraf, W. M. *et al.* Optimization of a 660 MW e supercritical power plant performance—a case of industry 4.0 in the data-driven operational management part 1. Thermal efficiency. *Energies* **13,** 5592 (2020).

83. Ashraf, W. M. *et al.* Strategic-level performance enhancement of a 660 MWe supercritical power plant and emissions reduction by AI approach. *Energy Conversion and Management* **250,** 114913 (2021).

84. Ashraf, W. M. *et al.* Construction of operational data-driven power curve of a generator by industry 4.0 data analytics. *Energies* **14,** 1227 (2021).

85. Ashraf, W. M. & Dua, V. Machine learning based modelling and optimization of post-combustion carbon capture process using MEA supporting carbon neutrality. *Digital Chemical Engineering* **8,** 100115 (2023).

86. Ashraf, W. M. *et al.* Artificial intelligence enabled efficient power generation and emissions reduction underpinning net-zero goal from the coal-based power plants. *Energy Conversion and Management* **268,** 116025 (2022).

87. Krzywanski, J. *et al.* AutoML-based predictive framework for predictive analysis in adsorption cooling and desalination systems. *Energy Science & Engineering* **12,** 1969–1986 (2024).

88. Chou, J.-S. & Bui, D.-K. Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. *Energy and Buildings* **82,** 437–446 (2014).

89. Ashraf, W. M. & Dua, V. Driving towards net-zero from the energy sector: Leveraging machine intelligence for robust optimization of coal and combined cycle gas power stations. *Energy Conversion and Management* **314,** 118645 (2024).

90. Ashraf, W. M. *et al.* Artificial intelligence modeling-based optimization of an industrial-scale steam turbine for moving toward net-zero in the energy sector. *ACS omega* **8,** 21709–21725 (2023).

91. Ashraf, W. M. & Dua, V. Storage of weights and retrieval method (SWARM) approach for neural networks hybridized with conformal prediction to construct the prediction intervals for energy system applications. *International Journal of Data Science and Analytics,* 1–15 (2024).

92. Vogel, G., Balhorn, L. S. & Schweidtmann, A. M. Learning from flowsheets: A generative transformer model for autocompletion of flowsheets. *Computers & Chemical Engineering* **171,** 108162 (2023).

93. Schweidtmann, A. M. Generative artificial intelligence in chemical engineering. *Nature Chemical Engineering* **1,** 193–193 (2024).

94. Song, B., Zhou, R. & Ahmed, F. Multi-modal machine learning in engineering design: A review and future directions. *Journal of Computing and Information Science in Engineering* **24,** 010801 (2024).

95. Samuel, K. M. & Ahmed, F. Continual Learning Strategies for 3D Engineering Regression Problems: A Benchmarking Study. *arXiv preprint arXiv:2504.12503* (2025).

96. Regenwetter, L., Nobari, A. H. & Ahmed, F. Deep generative models in engineering design: A review. *Journal of Mechanical Design* **144,** 071704 (2022).

97. Ahmed, F. *et al.* Design by Data: Cultivating Datasets for Engineering Design. *Journal of Mechanical Design* **147,** 040301 (2025).

98. Ghorbani, M. *et al.* An active machine learning approach for optimal design of magnesium alloys using Bayesian optimisation. *Scientific Reports* **14,** 8299 (2024).

99. Regenwetter, L., Abu Obaideh, Y. & Ahmed, F. Multi-objective counterfactuals for design: A model-agnostic counterfactual search method for multi-modal design modifications. *Journal of Mechanical Design* **147,** 021401 (2025).

100. Regenwetter, L. *et al.* Constraining generative models for engineering design with negative data. *Transactions on Machine Learning Research* (2024).

101. Mann, V. *et al.* eSFILES: Intelligent process flowsheet synthesis using process knowledge, symbolic AI, and machine learning. *Computers & Chemical Engineering* **181,** 108505 (2024).

102. Khan, M. T. *et al.* From Drawings to Decisions: A Hybrid Vision-Language Framework for Parsing 2D Engineering Drawings into Structured Manufacturing Knowledge. *arXiv preprint arXiv:2506.17374* (2025).

103. Jablonka, K. M. *et al.* Machine learning for industrial processes: Forecasting amine emissions from a carbon capture plant. *Science Advances* **9,** eadc9576 (2023).

104. Min, Q. *et al.* Machine learning based digital twin framework for production optimization in petrochemical industry. *International Journal of Information Management* **49,** 502–519 (2019).

105. Ibrahim, M. S., Dong, W. & Yang, Q. Machine learning driven smart electric power systems: Current trends and new perspectives. *Applied Energy* **272,** 115237 (2020).

106. Deon, B. *et al.* Digital twin and machine learning for decision support in thermal power plant with combustion engines. *Knowledge-Based Systems* **253,** 109578 (2022).

107. Kruse, J., Schäfer, B. & Witthaut, D. Revealing drivers and risks for power grid frequency stability with explainable AI. *Patterns* **2** (2021).

108. Wu, B. *et al.* Battery digital twins: Perspectives on the fusion of models, data and artificial intelligence for smart battery management systems. *Energy and AI* **1,** 100016 (2020).

109. Boukerche, A. & Wang, J. Machine learning-based traffic prediction models for intelligent transportation systems. *Computer Networks* **181,** 107530 (2020).

110. Ashraf, W. M., Dua, V. & Debnath, R. Domain Consistent Industrial Decarbonisation of Global Coal Power Plants. *arXiv preprint arXiv:2503.03571* (2025).

111. Rangel, N., Li, H. & Aristidou, P. An optimisation tool for minimising fuel consumption, costs and emissions from Diesel-PV-Battery hybrid microgrids. *Applied energy* **335,** 120748 (2023).

112. Hou, Z.-S. & Wang, Z. From model-based control to data-driven control: Survey, classification and perspective. *Information Sciences* **235,** 3–35 (2013).

113. Hewing, L. *et al.* Learning-based model predictive control: Toward safe learning in control. *Annual Review of Control, Robotics, and Autonomous Systems* **3,** 269–296 (2020).

114. Balhorn, L. S., Degens, K. & Schweidtmann, A. M. Graph-to-SFILES: Control structure prediction from process topologies using generative artificial intelligence. *Computers & Chemical Engineering,* 109121 (2025).

115. Yu, N. *et al.* Data-driven control, optimization, and decision-making in active power distribution networks. *Applied Energy* **397,** 126253 (2025).

116. Chen, C. *et al.* Enhancing the load cycling rate of subcritical coal-fired power plants: A novel control strategy based on data-driven feedwater active regulation. *Energy* **312,** 133627 (2024).

117. Sahoo, S., Wang, H. & Blaabjerg, F. *On the explainability of black box data-driven controllers for power electronic converters* in *2021 IEEE Energy Conversion Congress and Exposition (ECCE)* (2021), 1366–1372.

118. Yan, Z. & Xu, Y. Data-driven load frequency control for stochastic power systems: A deep reinforcement learning method with continuous action search. *IEEE Transactions on Power Systems* **34,** 1653–1656 (2018).

119. Zhang, K. *et al.* Explainable AI in deep reinforcement learning models for power system emergency control. *IEEE Transactions on Computational Social Systems* **9,** 419–427 (2021).

120. Zhang, K., Xu, P. & Zhang, J. *Explainable AI in deep reinforcement learning models: A shap method applied in power system emergency control* in *2020 IEEE 4th conference on energy internet and energy system integration (EI2)* (2020), 711–716.

121. Shekhar, A. R., Moar, R. R. & Singh, S. A hybrid mechanistic machine learning approach to model industrial network dynamics for sustainable design of emerging carbon capture and utilization technologies. *Sustainable Energy & Fuels* **7,** 5129–5146 (2023).

122. Bui, V.-H., Hussain, A. & Kim, H.-M. Double deep Q-learning-based distributed operation of battery energy storage system considering uncertainties. *IEEE Transactions on Smart Grid* **11,** 457–469 (2019).

123. Afroogh, S. *et al.* Trust in AI: progress, challenges, and future directions. *Humanities and Social Sciences Communications* **11,** 1–30 (2024).