A Digital Twin Approach to System-Level Fault Detection and Diagnosis for Improved Equipment Health Monitoring

Tat Nghia Nguyen[1*], Roberto Ponciroli[1], Paul Bruck[2], Thomas C. Esselman[2], Joseph A. Rigatti[3], Richard B. Vilim[1]

[1]Nuclear Science and Engineering Division, Argonne National Laboratory, Lemont, IL-60439.

[2]LPI Inc., Amesbury, MA-01913

[3]Dominion Energy, Richmond, VA 23219

Abstract

Automating the task of fault detection and diagnosis is crucial in the effort to reduce the operation and maintenance cost in the nuclear industry. This paper describes a physics-based approach for system-level diagnosis in thermal-hydraulic systems in nuclear power plants. The inclusion of physics information allows for the creation of virtual sensors, which provide improved fault diagnosis capability. The physics information also serves to better constrain diagnostic solutions to the physical domain. As a demonstration, various test cases for fault diagnosis in a high-pressure feedwater system were considered. The use of virtual sensors allows constructing performance models for two first-point feedwater heaters which would not have been possible otherwise due to the limited sensor set. Real-time plant data provided by a utility partner were used to assess the diagnostic approach. The detection of an abnormal event immediate after a plant startup pointed to faulty behaviors in the two first-point feedwater heaters. This double-blind fault diagnosis was subsequently confirmed by the plant operator. In addition, several simulated sensor fault events demonstrated the capability of our algorithms in detecting and discriminating sensor faults.

Keywords: Feedwater system; Real-time data; Model-based diagnosis; Fault detection; Digital twin.

[*] Corresponding author

Email address: nguyent@anl.gov

1. Introduction

The economic competition with lower cost electric generation alternatives is one of the main challenges facing the nuclear industry. A significant factor affecting the profitability is the high operations and maintenance (O&M) cost. Efforts by the nuclear utilities to address this challenge include the use of automated monitoring and fault detection capabilities in optimizing the O&M activities. This paper describes our physics-based system-level approach to fault detection and diagnosis for such applications in nuclear power plants. The proposed framework has the capability to detect and diagnose both equipment faults and instrument faults and can be adapted for various thermal-hydraulic (TH) systems in the plants.

The monitoring capability of most commercial nuclear power plants that entered service in the early 1980's was designed with sensor sets that provided the most basic required functionality. Compared to the monitoring capability being envisioned today for advanced reactor concepts, the sensor set in those plants is relatively limited. Furthermore, the available sensor data, obtained from past startup and normal operation events, are often not rich enough to support the training of data-driven algorithms. Purely data-driven approaches in current industry practice are susceptible to increased false detection rates under changes of operating conditions [1, 2, 3] and suffer from the inherent characteristic of the black-box treatment whose solutions are often considered shallow, brittle, and opaque [4, 5, 6, 7].

Motivated by the rapid development of machine learning (ML) techniques, various efforts have been made in recent studies to develop advanced data-driven diagnostic methods [8]. Wang et al. introduced a hybrid framework, combining the least squares support vector machine and Gaussian process regression methods, to detect and assess the severity of plant-level faults [9]. Li and Lin proposed to integrate different classifications methods in ensemble learning models to improve performance in detecting plant-level faults [10]. He et al. introduced an adaptive convolutional neural network architecture to solve the difficulties in constructing neural network models for fault diagnosis applications [11]. Wang et al. presented a deep transfer learning method to improve the performance of convolutional neural network methods when applied to different plant operating power levels [12]. Comparisons of five typical data-driven fault diagnosis methods for nuclear power plant applications were investigated in [13]. These studies offered different perspectives on mitigating the limitations of conventional data-driven methods for nuclear engineering applications.

An alternative solution to overcome the short-comings of data-driven methods consists in combining physics with machine learning methodologies in the so-called physics-informed or hybrid learning methods [14]. In particular, the knowledge of the physics of a system can be used develop dedicated models for data preprocessing (physics-based preprocessing), to construct ML networks (physics-based network architecture), or to formulate constraint equations for ML model regularization (physics-based regularization) [14]. Recent work on the application of this approach to fault detection and diagnosis can be found in [15, 16]. Nonetheless, there are several challenges that needs to be addressed to fully exploit hybrid models [14]. For applications with complex engineering systems, the need for an accurate system representation is often hindered by the unavailability of the design parameters of the components and the complexity of the underlying process. In particular, for the target cost reduction applications of the current work, an approach requiring the inputs from subject matter experts to provide design parameters would be economically counterproductive [17, 18].

This paper describes a physics-based system level approach to the problem of detecting and diagnosing slow performance degradations of components and sensors in nuclear power plants. The paper is mainly focused on solving the difficulties in constructing physics-based models for complex multi-component systems with reduced sensor sets without relying on component-specific design parameters. The available physics-based information in a system is fully utilized as the supplement to the limited sensor set to constrain the diagnostic solutions to the physical domain. In previous work [19], we developed an approach to construct physics-based components models to facilitate a model-based diagnostic framework and demonstrated the reasoning processes for fault diagnosis at the component level. The performance of the diagnostic framework at component level was further assessed for a feedwater pump system using real-time plant data [20]. Here we show how that diagnostic framework can be adapted to the system level for generic nuclear engineering systems.

At the component level, representation of the underlying physics, including the mass, momentum, and energy conservation equations, are used to formulate the parametric form of the component models. Each of the component models may contain several model parameters left to be determined by fitting against training data. It follows that the model training process for each model has a minimum sensor set requirement. To allow the construction of diagnostic models for maximum coverage, we introduce the concept of virtual sensors to be used in place of missing sensors for model training. Virtual sensors are obtained analytically for unmeasured process variables by solving the conservation laws and related constitutive equations available at the system level. Besides, the use of virtual sensors allowed us to formulate the system level approach such that the reasoning framework for fault diagnosis developed in [19] can be directly applied. The focus of this paper is to show how diagnostic models can be constructed and model residuals can be generated at the system level in that formulation.

This work is part of the effort at Argonne National Laboratory to develop digital twins with the capabilities to support both the continued operation of the current nuclear fleet and the design of advanced reactors. For this context, a digital twin, as defined in [21], is the use of a digital model of a physical system with automated real-time information exchange between the two. In our proposed framework, the available sensor set is fully utilized for the construction of a physics-based digital representation of the system, which in turn is used to monitor the health status of both the system equipment and the sensor set. The digital twin serves as a parallel representation that allows for comparison with actual plant operation (and potentially faulted operation) to diagnose faults. The resulting diagnostic capabilities can be applied for online fault detection and diagnosis to assist O&M activities [2] as well as for the inverse problem of determining optimal placement for new sensors in a system [1].

A significant feature of the proposed framework is the capability to produce explainable diagnostic results (*explainable agency* [22]). When presented with a fault diagnosis, a plant operator needs to understand and verify it before taking further corrective actions [23]. It is therefore crucial that the fault diagnoses are explainable, trustworthy, and actionable. In our approach, the explainability and trustworthiness are achieved from using physics-based models and virtual sensors to facilitate tracking the cause-effect relations between possible faults and fault symptoms throughout the system [2]. Details on the methodology of the proposed framework are discussed in Section 2. The high-pressure feedwater system and the results for various test cases using real-time plant data are described in Section 3. A summary of the proposed framework and discussion of the results is finally provided (Section 4).

2. Methodology

In previous work [19, 20], we developed and demonstrated a model-based diagnostic framework with physics-based models focusing on the component level of thermal hydraulic components. In this section, we will describe how the model-based framework can be generalized and applied to multi-component systems. To apply the model-based diagnostic framework, the first step consists in constructing a model that can describe the system behavior during fault-free operation [24, 18].

Our strategy to construct a physics-based model for a thermal hydraulic system is to decompose the system into building-blocks whose operations can be described by first principles. Thermal hydraulic systems in nuclear power plants are made of standard components, e.g., valves, pumps, heat exchangers. Each component type is designed to perform a specific function of mass, momentum or energy transfer, and their performance in fault-free conditions can be described by the underlying physics. One can formulate physics-based models for each generic component type from first principles. Each of such generic models may contain one or several parameters left to be determined by fitting against experimental data. The process of model fitting, also referred to as model calibration, requires certain operating variables to be measured for each component which imposes a sensor set requirement. In an ideal scenario with an abundant sensor set, one can construct diagnostic models and treat the system components independently. In current practice, the available sensor set is often limited, and it is not possible to construct a diagnostic model for every standalone component. Our approach to resolve this difficulty is by (1) relying on conservation laws and the relations between the components in a system to solve for missing sensors and (2) construct diagnostic models for each combination of components to the degree allowed by the sensor set.

2.1. Conservation Laws and Virtual Sensors

The operation of a T-H system is described by conservation laws. Consider a one-dimensional incompressible flow through with a single inlet/outlet control volume, the integral equations for the conservation of mass, momentum and energy are respectively given by [25]:

$$\frac{dm}{dt} = w_{\text{in}} - w_{\text{out}} \tag{1}$$

$$\left(\frac{l}{A}\right)_T \frac{dw}{dt} = P_{\text{in}} - P_{\text{out}} + \rho g (z_{\text{in}} - z_{\text{out}}) + \frac{w^2}{2\rho}\left(\frac{1}{A_{\text{in}}^2} - \frac{1}{A_{\text{out}}^2}\right) - \Delta P_{\text{loss}} \tag{2}$$

$$\frac{d}{dt} H = w_{\text{in}} h_{\text{in}} - w_{\text{out}} h_{\text{out}} + Q_{\text{eng}} \tag{3}$$

where the indices 'in' and 'out' denote the inlet and outlet location; m is the total fluid mass enclosed in the control volume; w denotes the mass flow rate; P denotes pressure; A denotes cross-sectional area; $(l/A)_T$ is the equivalent inertia length for the component defined by its geometric dimensions; ρ is the fluid density; g is the gravitational acceleration constant; z denotes the relative elevation at each location of the flow; ΔP_{loss} denotes the total pressure loss; H denotes the total energy enclosed in the control volume; h denotes the specific enthalpy and Q_{eng} is the combined energy source/sink term.

Under quasi-static conditions, the time derivative terms can be neglected, and the left sides of Eqns. (1)-(3) can be set to zero to obtain quasi-static conservation equations. We can use these quasi-static conservation equations as constraints between the process variables at different locations in the system to solve for the missing sensors.

By utilizing such balance equations and other constitutive equations across the systems, one can solve for certain process variables that are not measured by physical sensors. We will refer to this concept, i.e., an unmeasured variable obtained analytically to be used in the model calibration process and subsequently in fault diagnosis, as a *virtual sensor*.

Definition 1. A *virtual sensor* is the analytical solution of a process variable that can be used in place of a missing sensor for model construction and calibration.

Process variables in T-H systems include flowrate, pressure and temperature or enthalpy. When the system is fault-free, there is no loss of mass and external heat loss can be considered negligible. In that case, collections of system level mass and heat balance equations can be used to solve for virtual sensors of flowrate and enthalpy type. On the other hands, since the pressure loss term cannot be neglected, and is flowrate-dependent, it is not possible to generate virtual pressure sensors in the same manner from the momentum equations.

For example, the mass balance equation for a block of a system with an arbitrary number of inlet and outlet points is given by:

$$\sum w_{\mathrm{in}} = \sum w_{\mathrm{out}} \tag{4}$$

This equation remains valid when there is no fault among the flowrate sensors and no leakage in any component in the block. Similarly, the heat balance equation for a system block with no heat-exchanging component is given by:

$$\sum w_{in} \cdot h_{in} = \sum w_{out} \cdot h_{out} \tag{5}$$

The validity conditions for heat balance equations in general include no sensor faults, no leakage for the involved component and no faults that can cause significant external heat loss, if applicable.

If a virtual sensor can be determined, its value is valid if and only if all the involved balance equations are valid, i.e., the involved sensors and components must be free of the related faults. Such conditions can be assumed in the model construction process during which the system can be considered fault-free. Virtual sensors are computed firstly for model calibration purposes. In the subsequent use of the virtual sensors for fault diagnosis, one must keep track of the relevant faults that can invalidate the analytical value of each virtual sensor.

2.2. Model Construction and Residual Generation

After all solvable virtual sensors have been identified for a system, the combined set of physical sensors and virtual sensors can be used for model construction and calibration. Each diagnostic model has a specific sensor set requirement. To optimize the use of the sensor set, we then proceed to construct diagnostic models to the extent allowed by the available sensors. It is often the case that even with the addition of virtual sensors, it is not possible to construct models for every separate component. In that case, one must resort to extending the model set to include models from a combination of adjacent components.

In the current framework, diagnostic models include models for a specific component type, e.g., centrifugal pump, single-phase counterflow heat exchanger, as well as model for a set of generic components. In previous work, we demonstrated the use of the diagnostic models for a single-phase heat exchanger [19] and a compound system of fixed-speed induction motor and centrifugal pump [20], which can serve as example of standalone component models. Basic diagnostic models for typical TH components can be found, for example, in [26]. In the rest of this Section, we provide examples of multi-component models that can be applied for any combination of adjacent components given that certain conditions are met.

Mass Balance Model

Mass balance models are created for the purpose of detecting leakage and therefore should only involve physical flowrate sensors. Based on the mass conservation given by Eqn. (1) for quasi-static equation, an aggregate mass balance model can be constructed for any block of components if a physical flowrate sensor is available at every inlet and outlet point. The aggregate mass model is expressed by:

$$\sum w_{\text{in}} - \sum w_{\text{out}} = 0 \tag{6}$$

where each inlet w_{in} or outlet w_{out} flowrate must be available by a physical sensor.

Pressure Loss Models

By using Eqn. (2) under quasi-static equation, we can construct a pressure loss model for any combination of components with fixed geometry. We limit our consideration to a block of *isolated* components between single inlet and outlet points. More specifically, for a block of adjacent components under single-phase flow between a single inlet and outlet, the overall pressure loss can be expressed as a parametric model of the flowrate as:

$$P_{\text{in}} - P_{\text{out}} = f(w) \tag{7}$$

where P_{in} and P_{out} are the inlet and outlet pressure readings and w is a flowrate reading at either the inlet and outlet; $f(w)$ is the parametric form of the model. As an example, one can use a quadratic form for $f(w)$, in that case, we have the following parametric form of the pressure loss model:

$$P_{\text{in}} - P_{\text{out}} = \theta_0 + \theta_1 w + \theta_2 w^2 \tag{8}$$

Here θ_0, θ_1, and θ_2 are parameters to be determined by the model calibration process. As discussed, there is no virtual sensor for pressure variables thus both P_{in} and P_{out} must be provided by physical sensors. On the other hand, the required flowrate reading may be obtained from a virtual sensor.

After the available models have been calibrated, one can then generate model residuals for fault diagnosis. Each model may provide one or multiple analytical redundancy relations, each of which can be used to generate one residual [24, 19]. For example, the residuals from the mass balance model in Eqn. (6) and the pressure loss model in Eqn. (8) are given, respectively, by:

$$r_{\text{mass}} = \sum w_{\text{in}} - \sum w_{\text{out}} \tag{9}$$

$$r_{\text{press}} = \left(P_{\text{in}} - P_{\text{out}}\right) - \left(\theta_0 + \theta_1\ w + \theta_2\ w^2\right) \tag{10}$$

Recall that for the current diagnostic framework, each residual serves as a possible fault symptom for the reasoning process. To optimize the resolution of the diagnostic results, we are interested in maximizing the number of independent residuals. For a given model, additional residuals can be generated by using different (independent) sensor set, if available. In the calculation of a model residual, instead of using a real sensor reading, one can use the value obtained analytically for the underlying variable, thus generate a new, independent, residual. We generalize the use of such analytical variable by the concept of *type-II virtual sensors*.

Definition 2. A *type-II virtual sensor* is the analytical solution of a process variable to be used as the alternative of a physical sensor or a virtual sensor in the calculation model residuals.

Here the term “type-II” is used to differentiate this concept from the (type-I) virtual sensor concept in the last section. In general, type-II virtual sensors can be obtained by utilizing the predictions of available diagnostic models. More details on these concepts can be found in [18]. As with type-I virtual sensors, when using a type-II virtual sensor for residual generation in fault diagnosis, one must keep track of the relevant faults that can invalidate its value.

2.3. Fault Diagnosis at the System Level

In this work, an extension of the same reasoning framework as developed and demonstrated for component level diagnosis [19, 18] is presented. In particular, non-zero residuals are treated as fault symptoms and various reasoning methods, including deterministic and probabilistic reasoning, can be used to obtain diagnoses from a set of symptoms. The basis for the reasoning process is the causality relations between different faults and the fault symptoms. For that, one needs keep track of the dependency structure of each residual, i.e., the relevant faults that can cause a residual to become non-zero. For deterministic reasoning, the list of relevant faults is needed to construct a conflict relation from each non-zero residual. For probabilistic reasoning, the structure of each residual is needed to construct a Bayesian network for the system.

The dependency structure of a generic model residual in the system-level approach is summarized by the flowchart in Figure 1. To calculate each model residual, one or multiple virtual sensors of each type, whose validity depends on other system components and sensors, may be needed. A fault can cause a model residual to deviate either by directly affecting the model or by invalidating one of the involved virtual sensors. The list of relevant faults for each residual may include not just faults from the components and sensors directly involved in the underlying model but also the faults that can affect the calculation of the involved virtual sensors.

The use of virtual sensors thus allows us to keep track of the cause-effect relations between possible faults and fault symptoms across the system, from which reasoning rules for fault diagnosis can be formulated. More specifically, each residual can represent a fault symptom. The causality relations between possible faults and each residual are governed by the dependency structure of the residual. Using backward chaining inference, fault diagnoses can be inferred from the set of observed symptoms. Diagnostic results can be presented along with the observed symptoms and can be explained by the forward causality relations. The explainability of the diagnostic result is a significant feature of this approach.

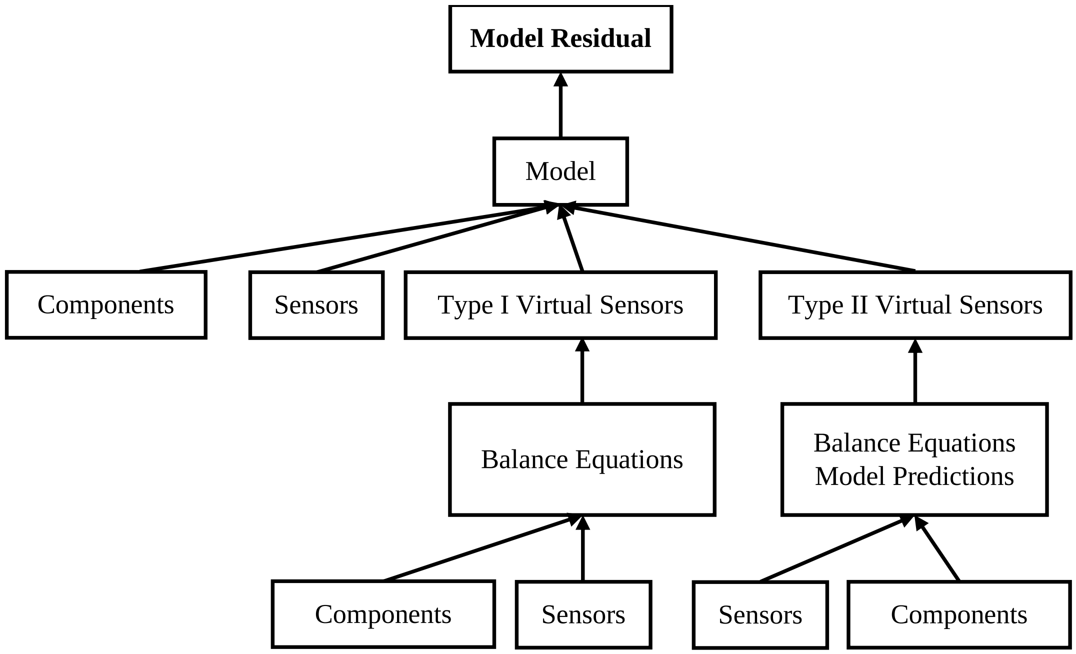

Figure 1. Dependency structure of a model residual in system level diagnosis.

As demonstrated in previous work [19], given a list of model residuals for a component, a Bayesian network can be built to represent the dependencies between each residual and its relevant faults, from which inferences can be made to perform component-level diagnostics. In this paper, we have formulated the system level approach in a way the same diagnostic framework can be applied. Specifically, through the concepts of virtual sensors, the physics-based knowledge of the system is similarly condensed to be represented by a list of residuals and their dependencies on component and sensor faults. Although in this case each residual may involve different components and sensors across the system, the same procedure of constructing the probabilistic network and performing diagnostics can be applied directly.

Each type-I virtual sensor is obtained from solving a set of balance equations involving certain physical sensors. Analytically, the uncertainty of such virtual sensor value can be evaluated from propagating the uncertainty of the individual physical sensors. Similarly, the uncertainty of a type-II virtual sensor is the combined effect of the measurement uncertainty of the physical sensors and the prediction uncertainty of the models involved. Each virtual sensor therefore has a higher "measurement" uncertainty compared to a typical physical sensor of the same variable type. As a result, model residuals involving virtual sensors have higher uncertainties which lead to higher detection delays given the same detection threshold.

As illustrated, when considering the dependencies of each model residual, the involved virtual sensors represent analytical expressions connecting the residual to other physical sensor measurements (Figure 1). The uncertainty of each residual can be computed directly using a training data set without the middle step of evaluating the uncertainty of individual virtual sensors. Ultimately, only the uncertainties of the model residuals are needed to perform fault detection and diagnosis. The demonstration of the uncertainty treatment and the probabilistic reasoning approach for fault diagnosis at the system level fall outside the scope of the current paper. In Section 3 we will show how virtual sensors can be obtained and used for model construction and residual generation at the system level. The fault diagnosis step is demonstrated by a deterministic reasoning approach.

3. Demonstration – A High Pressure Feedwater System

For a demonstration, let us consider the high-pressure feedwater system in a typical PWR plant. The corresponding P&ID is shown in Figure 2. The system in consideration is part of the condensate and feedwater system that is responsible for the supply of pre-heated feedwater to the steam generators. Feedwater from the condenser is heated through multiple stages of feedwater heaters [27]. We consider the last two heating stages closest to the inlets of the steam generators known as the first-point and second-point stages. Each stage consists of two feedwater heaters in parallel piping lines. The system as shown also include three feed pumps and three drain pumps.

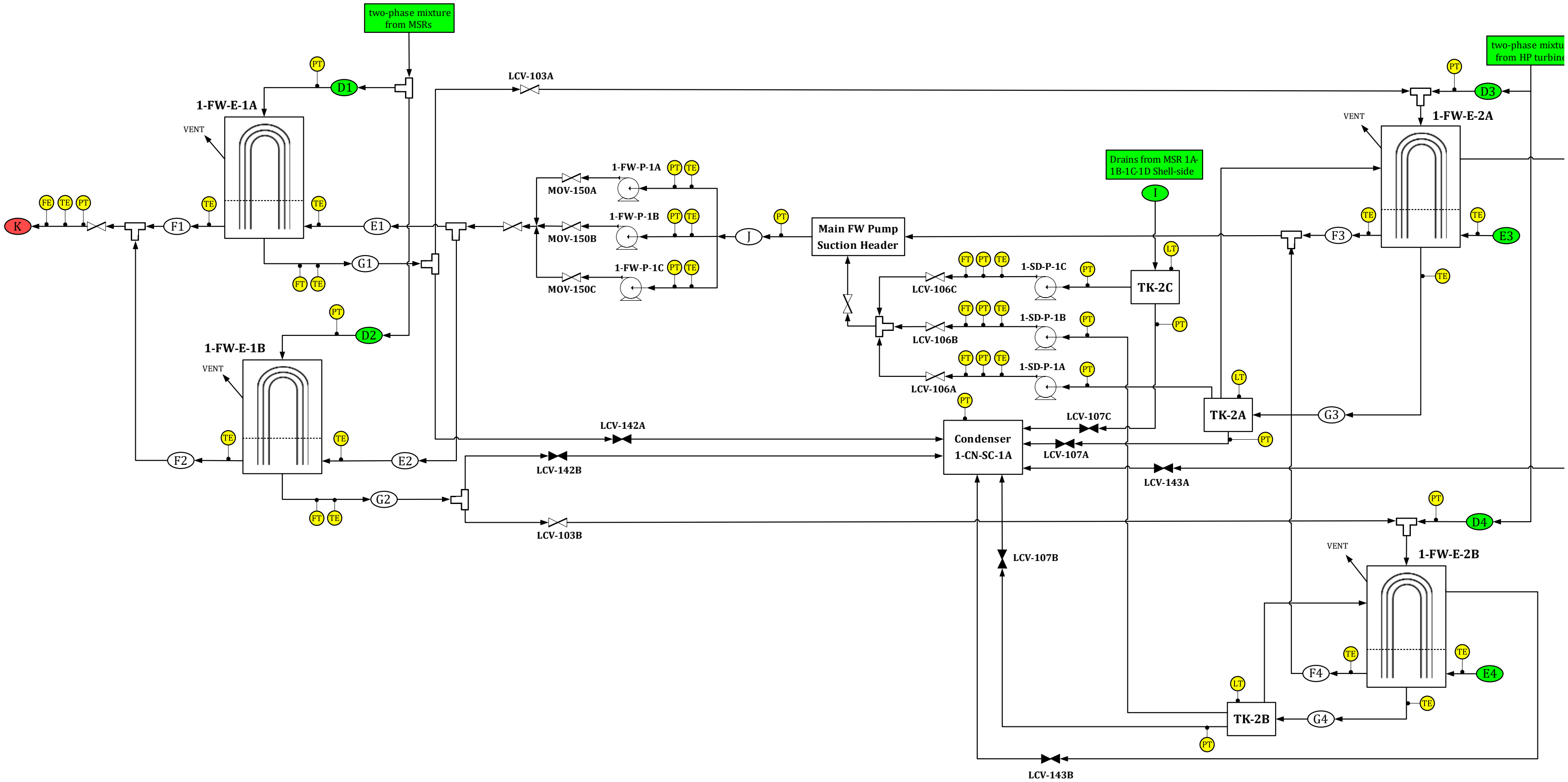

Figure 2. P&ID of the high-pressure feedwater system.

From the top of the P&ID, two-phase mixtures from the high-pressure turbine and moisture separator reheaters flow into the shells of the first and second-point heaters. The first-point heaters drain to the shells of the second-point heaters. On the right of the P&ID, feedwater from earlier heating stages flows through the two second-point heaters to the suction header of the feed pumps. Drains from the two-point heater shells are collected by the high-pressure heater drain receivers. Drains from the moisture separators (not shown in Figure 2) are collected by a third drain receiver. The three drain pumps pump condensate from the drain receivers to the suction header of the feed pumps. During normal operation, only two of the three feed pumps operate to pump feedwater through the two first-point heaters to a discharge header to supply the steam generators. In emergency situations, excessive drains from the first-point heaters and the drain receivers are collected by a condenser. All the valves to the condenser are otherwise closed off during normal operation. The yellow tags in Figure 2 indicate the sensors whose measurements are recorded by the plant data acquisition system.

3.1. Component Models and Sensor Set Requirements

The main components in the system are the four FWHs, three feed pumps and three drain pumps. Each of the pumps is powered by an electric motor operated in fixed-speed mode. Possible diagnostic models for such pump-motor sets were discussed in [20].

For each FWH, we are interested in monitoring its heat transfer performance. In particular, we will focus on the heat transfer coefficient of the condensing zone which is responsible for most of the heat transfer. The two first-point FWHs are of the two-zone vertical channel-down design. A simplified schematic of a vertical feedwater heater is provided in Figure 3.

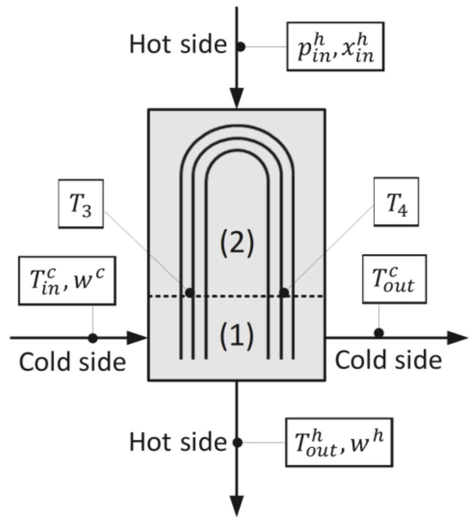

Figure 3. Simplified schematic of a vertical feedwater heater.

The log-mean temperature difference (LMTD) of the condensing zone is given by:

$$\mathrm{LMTD}_{cond.} = \frac{T_4 - T_3}{\ln \dfrac{T_s - T_3}{T_s - T_4}} \tag{11}$$

where T_3, T_4, respectively, are the feedwater temperature on the tube side entering and exiting the condensing zone; and T_s is the saturation temperature on the shell side of the condensing zone. Since the FWH does not have a de-superheating zone, one can approximately use the FWH tube side outlet temperature as T_4. T_3 and the heat transfer rate in the condensing zone can be calculated from the heat balance equations. One can then compute the overall heat transfer coefficient in condensing zone, defined as the ratio between the heat transfer rate and the LMTD:

$$UA_{cond.} = \frac{\dot{Q}_{cond.}}{\mathrm{LMTD}_{cond.}} \tag{12}$$

Such estimate of the heat transfer coefficient is typically used to assess the performance of the condensing zone for FWHs [28, 29]. It depends on the heat transfer processes on both sides of the FWH, which in turn depend on the operating conditions. To monitor the performance of the FWH, one must develop a model for $UA_{cond.}$ as a function of relevant operating variables.

With regard to the condensing process on vertical surfaces on the shell side, the most relevant variables are the feedwater mass flowrate w^c, drain flowrate w^h, steam quality x_{in}^h, and steam inlet pressure p_{in}^h:

$$UA_{cond.} = f(w^c, w^h, x_{in}^h, p_{in}^h) \tag{13}$$

Detailed development of the exact form for $UA_{cond.}$ is out of the scope of the current paper which we will omit here. The construction of such diagnostic model for a standalone FWH requires at least six sensor measurements, i.e., feedwater flowrate, feedwater inlet temperature, feedwater outlet temperature, drain flowrate, drain temperature, and steam inlet pressure.

3.2. Virtual Sensors and Available Diagnostic Models

Let us consider the sensor set as shown in the P&ID in Figure 2. The available sensor set for each first-point FWH does not include the required feedwater flowrate sensor. However, by utilizing the available sensors and balance equations in the system, a (type-I) virtual sensor can be obtained for the missing flowrate measurement. The resulting virtual sensors will then allow us to construct a heat transfer model for each first-point FWH. The two virtual flowrate sensors can also be used to monitor the flow ratio between the two parallel feedwater flows.

For the second-point FWHs, both the feedwater flowrate and drain flowrate are missing and cannot be obtained from solving balance equations. It is therefore not possible to construct any diagnostic models for the second-point FWHs.

For each drain pump, a model can be constructed for the pump head as a function of the flowrate. As the pump is powered by a motor in fix-speed service, the shaft-speed-dependency of the pump head can be neglected. For each feed pump, the discharge pressure and flowrate sensors are missing and thus, no diagnostic models can be constructed.

The set of available sensors along with the additional virtual sensors that can be generated for this system is shown by the updated P&ID in Figure 4. A list of the available models as described above is summarized in Table 1.

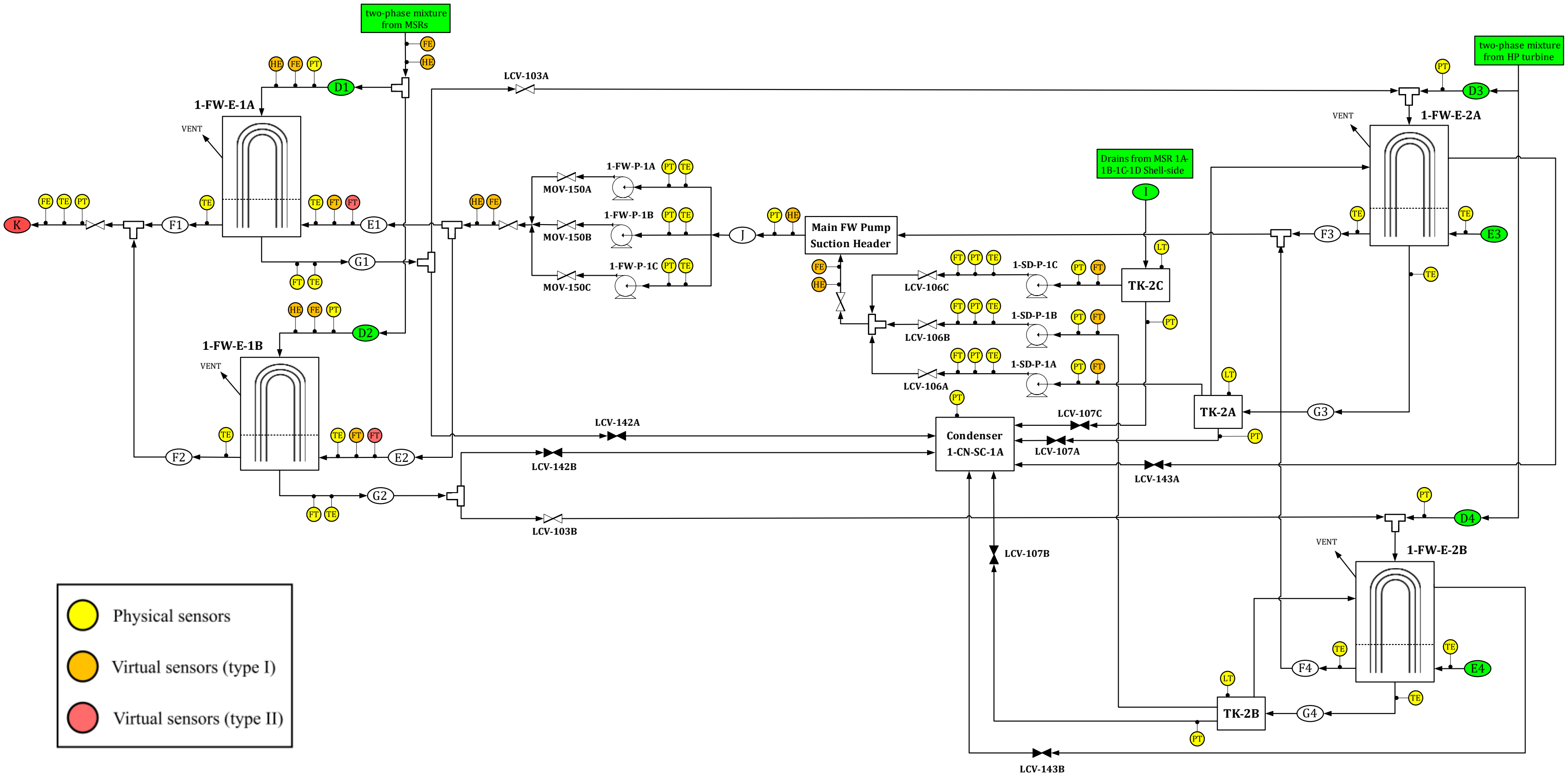

Figure 4. P&ID of the high-pressure feedwater system with the available sensor set along with additional virtual sensors

Table 1. Available diagnostic models for the high-pressure feedwater system

ID	***Name***	***Model Type***	***Components***	***Relevant Fault Types***
1	FR-1	Flow ratio	FWHs 1A and 1B	Tube leakage, Blockage
2	FWH-1A	HX performance	FWH 1A	Leakage, Fouling
3	FWH-1B	HX performance	FWH 1B	Leakage, Fouling
4	SDP-1A	Pump head	Drain pump 1A	Pump fault
5	SDP-1B	Pump head	Drain pump 1B	Pump fault
6	SDP-1C	Pump head	Drain pump 1C	Pump fault

We now turn to the generation of model residuals and related type-II virtual sensors. The two (type-I) virtual sensors for the feedwater flowrates, obtained previously, can be used in the flow ratio model FR-1 to generate a residual. From each FWH model, FWH-1A or FWH-1B, we can compute a residual to monitor the heat transfer performance. The calculation of each FWH model residual involves the use of a virtual sensor for the required feedwater flowrate. As described, the two required feedwater flowrates can be provided by the two type-I virtual sensors. Additionally, we can use the prediction of the flow ratio model FR-1 to compute the values, i.e., type-II virtual sensors, for the two feedwater flowrates, independently of the two type-I virtual sensors.

Similarly, we can compute a residual from the pump head model, SDP-1A, SDP-1B, and SDP-1C, for the drain pumps. Additionally, we can make use of the parity relation between the inlet temperatures of the two FWHs 1A and 1B. The parity relation arises from the energy balance as the two feedwater flows were from the same flow splitter. A list of available residuals is provided in Table 2.

Table 2. Model residuals for the high-pressure feedwater system.

Residual	***Model***	***Sensors***
r_2	FR-1	Two type-I virtual flowrate sensors
r_3	SDP-1A	Inlet pressure, outlet pressure, flowrate sensors
r_4	SDP-1B	Inlet pressure, outlet pressure, flowrate sensors
r_5	SDP-1C	Inlet pressure, outlet pressure, flowrate sensors
r_6	FWH-1A	Five physical sensors, 1 type-I virtual sensor
r_7	FWH-1B	Five physical sensors, 1 type-I virtual sensor
r_8	FWH-1A	Five physical sensors, 1 type-II virtual sensor
r_9	FWH-1B	Five physical sensors, 1 type-II virtual sensor

r_{10}	Parity	Two temperature sensors

Next, we turn to the description of fault signatures as constructed from model residuals. A list of possible faults across the two FWHs and their sensors is given in Table 3. Table 4 in turn presents the list of fault signatures for single-fault events that appear in Table 3. All the faults in Table 3 can be detected but a diagnosis may contain multiple possibilities. Only the faults with a unique signature in Table 4 can be uniquely identified. For example, Fault ID 15 (sensor E2.T for FWH 1B feedwater inlet temperature) gives rise to five non-zero residuals r_2, r_6, r_7, r_9 and r_{10}, a signature that is unique among the 19 different Fault IDs. The Deterministic II reasoning process [19, 18] can then identify E2.T as the unique minimal diagnosis.

On the other hand, Fault IDs 16, 17 and 18 (respectively for FWH 1B outlet feedwater temperature, drain temperature and drain flowrate) all have the same signature. In that case, the diagnosis for this residual signature has three fault possibilities. It should be noted that the fault signature table is only constructed here for illustration. Deterministic diagnoses can be obtained by backward chaining reasoning without explicitly constructing fault signature tables.

Table 3. List of faults for the two first-point FWHs and relevant sensors

Fault ID	***Comp. Name***	***Comp. Type***	***Fault***
1	FWH 1A	FWH	Fouling
2	FWH 1A	FWH	Tube leak
3	FWH 1A	FWH	Shell leak
4	FWH 1A	FWH	Tube block
5	FWH 1B	FWH	Fouling
6	FWH 1B	FWH	Tube leak
7	FWH 1B	FWH	Shell leak
8	FWH 1B	FWH	Tube block
9	K.w	Flow sensor	Sensor fault
10	E1.T	Temp. sensor	Sensor fault
11	F1.T	Temp. sensor	Sensor fault
12	G1.w	Flow sensor	Sensor fault
13	G1.T	Temp. sensor	Sensor fault
14	D1.p	Press. sensor	Sensor fault
15	E2.T	Temp. sensor	Sensor fault
16	F2.T	Temp. sensor	Sensor fault
17	G2.w	Flow sensor	Sensor fault
18	G2.T	Temp. sensor	Sensor fault
19	D2.p	Press. sensor	Sensor fault

Table 4. Fault signatures for single-fault events in the two first-point FWHs and relevant sensors

Fault ID	*Residuals*					
	r_2	r_6	r_7	r_8	r_9	r_{10}
1	0	1	0	1	0	0
2	1	1	1	1	1	0
3	0	1	1	1	0	0
4	1	0	1	0	1	0
5	0	0	1	0	1	0
6	1	1	1	1	1	0
7	0	1	0	1	1	0
8	1	0	1	0	1	0
9	0	1	1	1	1	0
10	1	1	1	1	0	1
11	1	1	1	1	0	0
12	1	1	1	1	0	0
13	1	1	1	1	0	0
14	0	1	0	1	0	0
15	1	1	1	0	1	1
16	1	1	1	0	1	0
17	1	1	1	0	1	0
18	1	1	1	0	1	0
19	0	1	1	0	1	0

For each drain pump, we only have one model residual as shown in Table 5. Either a pump fault or a fault in one of the sensors can cause the residual to be non-zero. Therefore, we can detect when a fault occurs but cannot discriminate between the possibilities of pump fault and sensor faults.

Table 5. Fault signature for each drain pump

Residuals	*Component Faults*	*Sensor Faults*		
	Pump	Q	P_{in}	P_{out}
$r_{\Delta P}$	1	1	1	1

3.3. Model Calibration Using Plant Data

To qualify the performance of the component models, we used a data set recorded by the real-time data acquisition system in a PWR plant, provided by a utility collaborator. Sensor data during a plant startup were used for the calibration of the models. For illustration, a plot of the plant power during the startup is shown in Figure 5.

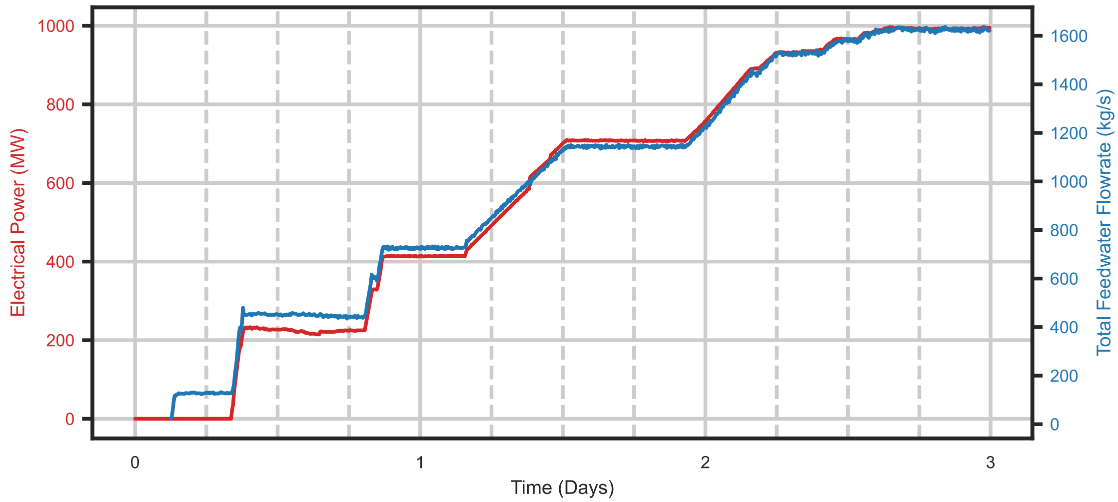

Figure 5. Plant electrical power and total feedwater flowrate during startup.

The data in the window between 40% and 100% power were selected for the training set. For illustration, the result for the fitted model of the condensing zone heat transfer coefficient for FWH 1A is compared in Figure 6 against the training set. The timestamps are days from the beginning of the startup.

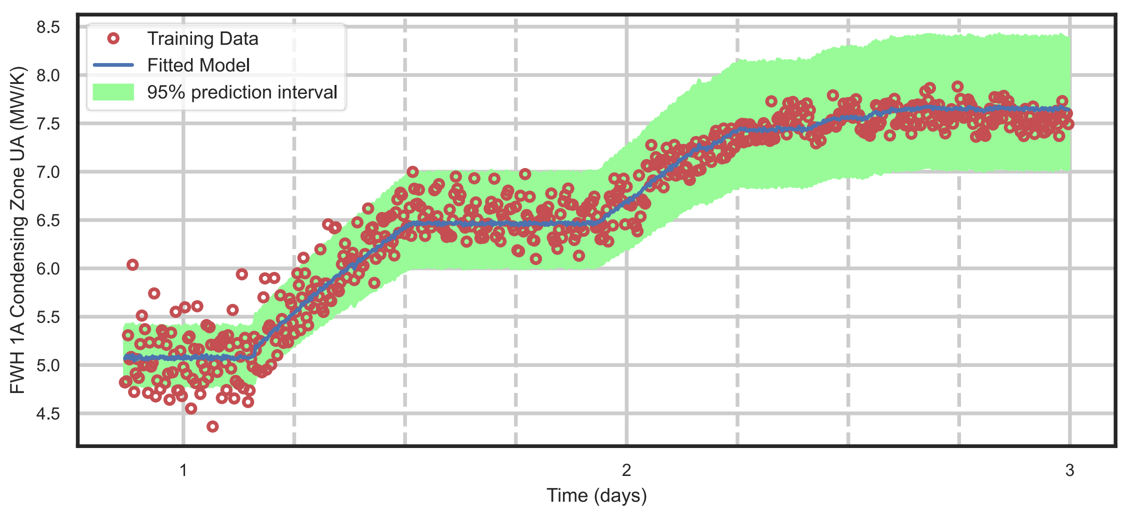

Figure 6. Fitted model for FWH 1A condensing zone UA against training data.

3.4. Diagnostic Results

In this section, we consider various fault scenarios to assess the performance of the proposed diagnostic approach.

3.4.1 Component Faults

Firstly, let us consider a time period immediately after the startup, that is the time after the plant nominally reached full power. Plots of the plant power and the main operating variables of the two first-point FWHs are shown in Figure 7 and Figure 8. For a better visualization of the changing trends, 60-minute moving average values were used for the plots. All timestamps are shown in the unit of days since the beginning of the startup. During this period, one sees in the figures that the plant power further increased by approximately 3%. The operating conditions of the two first-point FWHs remained relatively the same, except for the steam inlet pressures, which slightly increased.

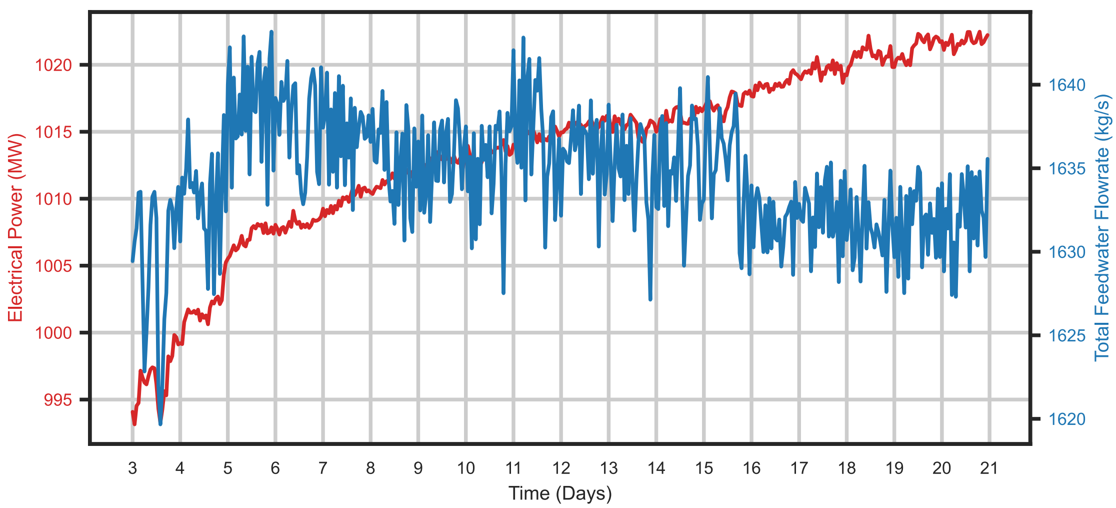

Figure 7. Plant power and feedwater flowrate during the period after startup.

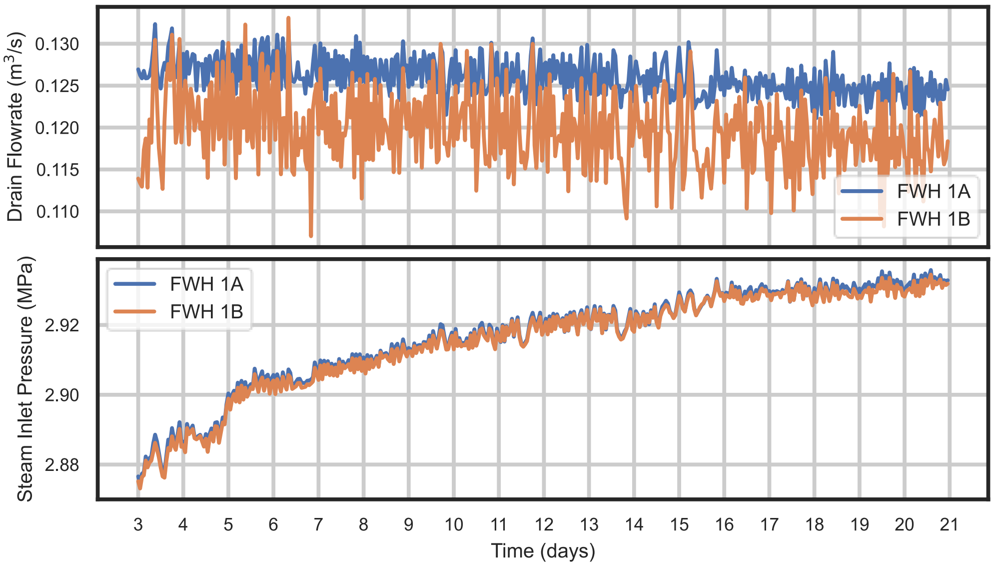

Figure 8. Steam inlet pressures and drain flowrates of the two first-point FWHs after startup.

Diagnostic model prediction for the FWH 1A condensing zone heat transfer coefficient is shown in Figure 9 against measured data. A significant discrepancy between the predicted and measured value can be observed. This gap leads to a non-zero residual computed from the model. Similarly, the residuals computed for the FWH 1B condensing zone UA can be observed to be non-zero, regardless of the virtual sensor used for feedwater flowrate reading.

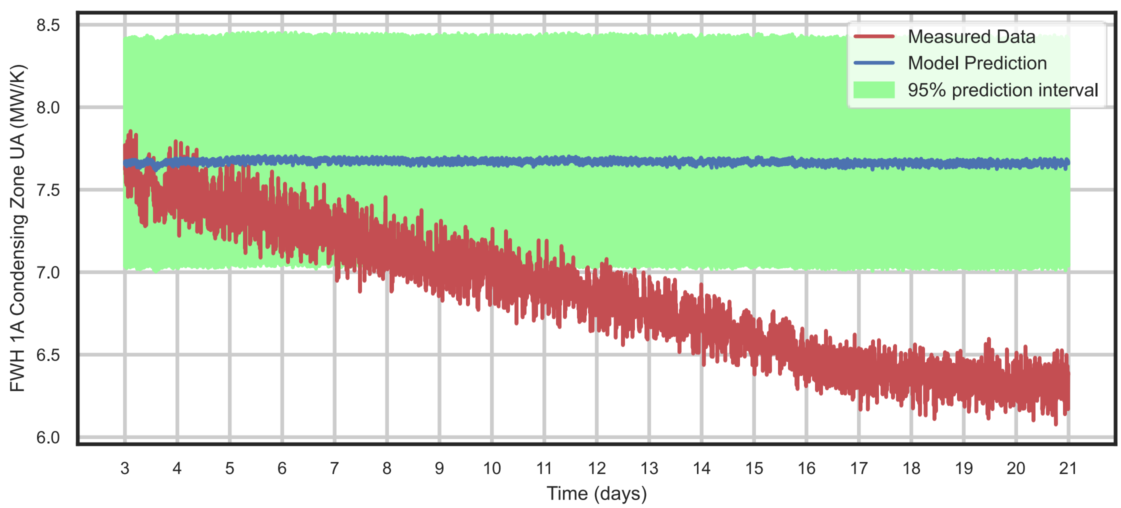

Figure 9. Model predictions and measured data for the FWH 1A condensing zone UA after startup.

Overall, the non-zero residuals in this case are r_6, r_7, r_8 and r_9, computed from the heat transfer models of FWH 1A and 1B. We omitted results of statistical tests for the residuals here as they are not relevant to the discussion. These non-zero residuals are symptoms of a decrease in the condensing zone heat transfer coefficients of both first-point FWHs, which cannot be accounted for by the slight changes of operating conditions. Using a deterministic reasoning method for fault diagnosis [19, 18], the possible causes based on the residual signatures are:

1. Fouling in both FWHs 1A and 1B
2. Sensor fault for the total feedwater flowrate reading.

To explain these results, note that from the way the diagnostic models were formulated for the FWHs, in the first possibility, a *fouling* fault implies any fault or internal changes in the FWH that can affect the heat transfer performance. In the second possibility, a total flowrate sensor fault would invalidate the virtual sensors for the feedwater flowrates of both FWHs, without affecting the flowrate ratio, thus would result in the four non-zero residuals as observed. This second possibility also includes any instrument anomalies that may have affected the total feedwater flowrate reading.

An alternate way to analyze these abnormal performance changes, which serves as an independent check on the above, is to consider the three performance indicators commonly used for FWHs, namely the feedwater temperature rise (FTR), drain cooler approach (DCA), and terminal temperature difference (TTD) [12]. Results for these indicators, computed directly from sensor data, for the two first-point FWHs are shown in Figure 10. Given that the operating conditions of the two FWHs are basically unchanged during this period, the increasing trends of the TTDs indicate a decreasing heat transfer performance in the condensing zone of both FWHs. The constant trend of DCAs for both FWHs indicate that the drain cooling zones were not affected. The decreasing FTRs simply indicate reduced heat transfer rates, which is not inconsistent with the findings above obtained by residual signature analysis.

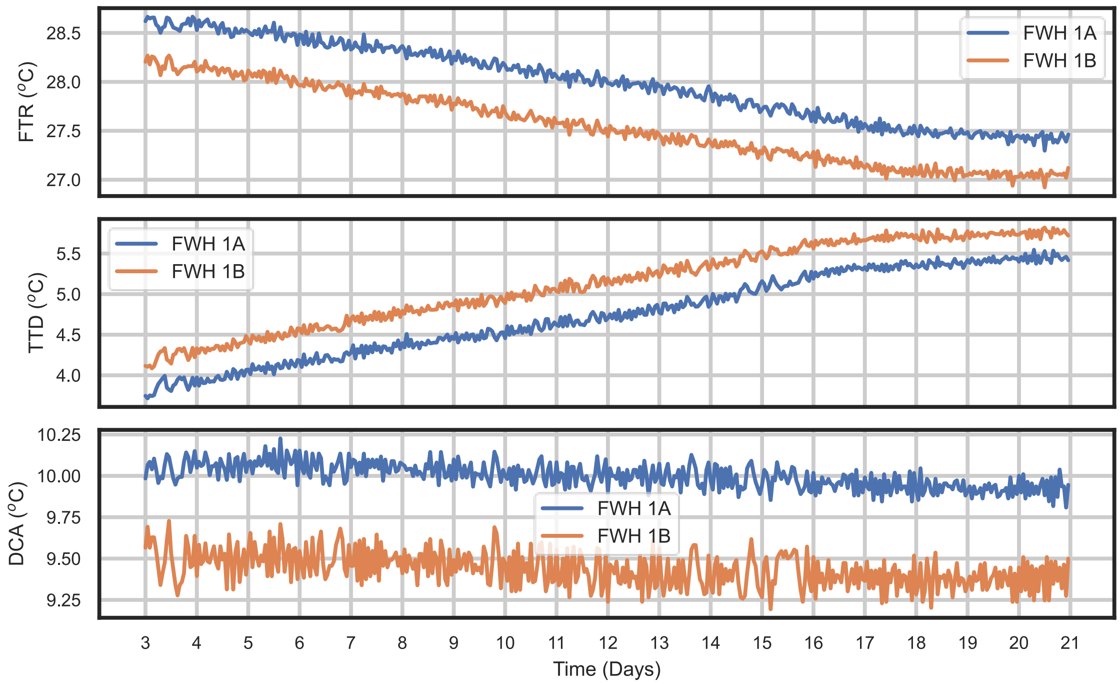

Figure 10. Performance indicators of the two first-point FWHs after startup.

Though the available sensor set did not allow us to differentiate between the two possibilities, we can speculate as to the more likely of the two. First, it is unlikely that the change in feedwater flowrate sensor error needed to account for the observations would occur over the relatively short period of 17 days and then stabilize. Second, there is a chemical treatment procedure carried out prior to shutdown to rejuvenate heat transfer surfaces by dissolving magnetite deposits. It is possible a new equilibrium of deposits in the condensing zone is established following startup, which would alter the condensing heat transfer coefficient.

In any case, our diagnostic analysis did result in a double-blind finding that that indicates abnormal behaviors of the two first point FWHs after startup. This was subsequently confirmed by the plant operator. While the exact cause is not immediately clear, further study is warranted to assess if there are possibly lost megawatts that can be recovered.

3.4.2 Simulated Sensor Faults

To simulate sensor faults, we used the sensor data collected from a post-startup period as the base data set and injected a bias into one sensor reading in each of the scenarios described below. The model residuals were realigned to the mean values at the start of the period, which would allow us to zero-out the effects of the issue described in Section 3.4.1.

FWH 1A Feedwater Inlet Temperature (E1.T)

A bias of $+3^{o}C$ was added to the reading values of the temperature sensor at the feedwater inlet of FWH 1A to simulate a temperature sensor fault. A plot of the resulting signal is shown in Figure 11in comparison with the corresponding sensor for FWH 1B.

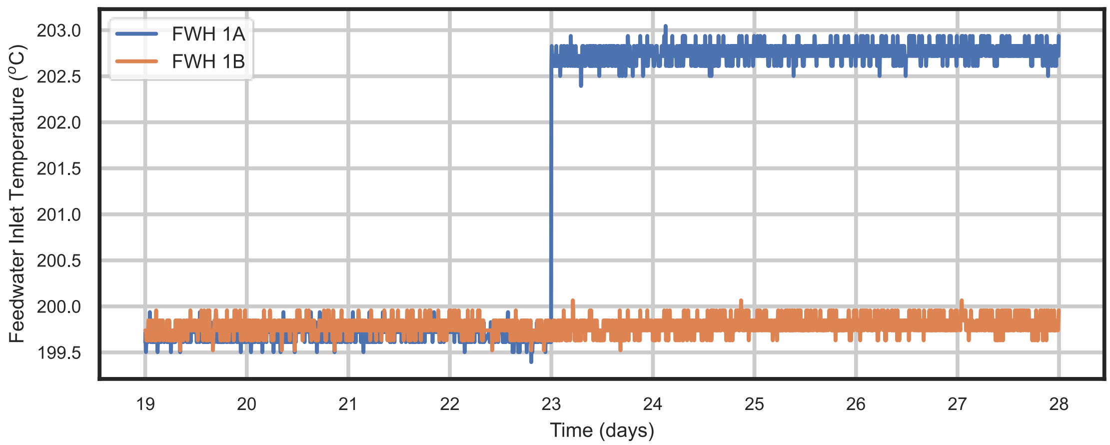

Figure 11. FWH 1A feed water inlet temperature with an injected bias.

To detect non-zero residuals, one can use a statistical change detection tool, e.g., the generalized likelihood ratio (GLR) test [30, 19]. Using the GLR test, a decision function is computed for each residual and a change is detected when the decision function exceeds a predefined detection threshold. A maximum number of 400 past data points was used in the calculations of the GLR decision functions. The results for this case are plotted in Figure 12. We can observe five non-zero residuals $(r_2, r_6, r_7, r_8, r_{10})$ with negligible false positive rates, and the fault can be detected and identified uniquely. This result can be verified manually using Table 4.

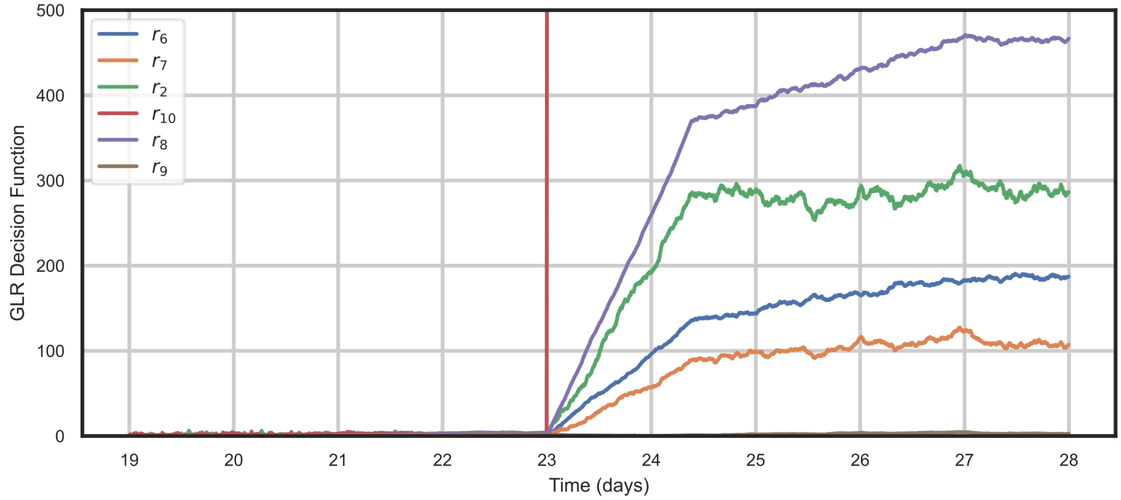

Figure 12. GLR decision functions for the detection of non-zero residuals in the E1.T sensor fault scenario.

FWH 1A Feedwater Outlet Temperature (F1.T)

Similarly, a bias of $+3^oC$ was added to the reading values of the feedwater temperature sensor at outlet of FWH 1A to simulate a temperature sensor fault. A plot of the resulting signal is shown in Figure 2 in comparison with the corresponding sensor for FWH 1B. Results for the GLR decision functions for the detection of non-zero residuals are shown in Figure 14

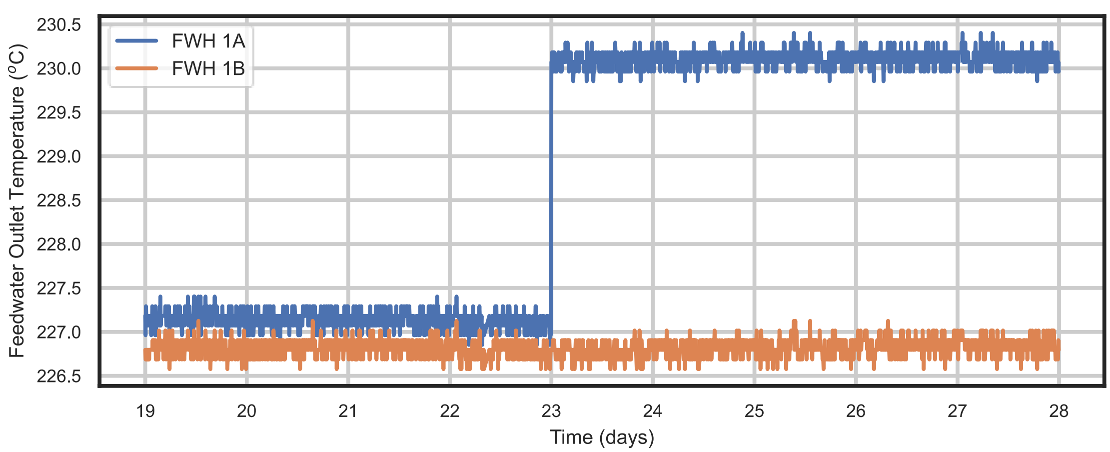

Figure 13. FWH 1A feed water outlet temperature with an injected bias.

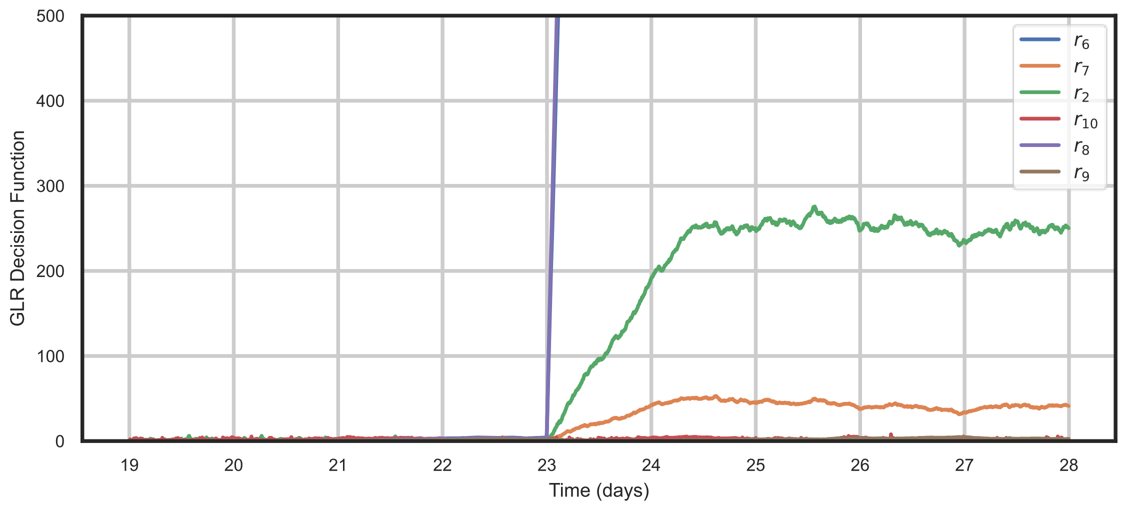

Figure 14. GLR decision functions for the detection of non-zero residuals in the E1.T sensor fault scenario (note that the lines for r_6 and r_8 are overlapping).

In this case, we can observe four non-zero residuals (r_2, r_6, r_7, r_8). The diagnostic result includes three possibilities of sensor faults: F1.T, G1.w or G1.T. Unlike the previous scenario, which had the help of a parity relation between the two feedwater inlet temperature sensors, in this case the available sensor set does not allow us to differentiate between three possible faults.

4. Discussion and Conclusion

In this paper, we presented a physics-based approach to fault detection and diagnosis at the system level for O&M cost reduction applications in commercial nuclear power plants. Physics-based information of the system is fully utilized as a supplement to sensor data to overcome sensor set limitations and serve to constrain diagnostic predictions to those that are physically consistent. At the component level, the underlying physics laws are employed to formulate the parametric component performance models. At the system level, available conservation laws and related constitutive equations are utilized to solve for unmeasured process variables. Such analytically solved variables, generalized by the concept of virtual sensors, which allow us to construct additional diagnostic models that are otherwise not possible due to the sensor set limitations.

We considered multiple test cases with a high-pressure feedwater system of a PWR plant. The main components of this system include four feedwater heaters, three feed pumps and three drain pumps. Construction of standalone diagnostic models for the feedwater heaters was not possible due to the limited sensor set available for the system. This difficulty was resolved by the use of *virtual sensors* obtained from solving the system balance equations. These virtual sensors allowed us to construct heat transfer models for the two first-point feedwater heaters. Additional, pump head models were constructed for the three drain pumps. Real-time plant data, provided by our collaborating utility partner, were used for both model training and diagnosis of various test cases of component faults and sensor faults. The diagnosis of an abnormal event immediate after a plant startup pointed to faulty behaviors in the two first-point feedwater heaters – a result confirmed by the plant operator. Results for various simulated sensor fault events also demonstrated the capability of our algorithms in detecting and isolating sensor faults.

Acknowledgment:

This work was supported by the U.S. Department of Energy, Office of Nuclear Energy. The submitted manuscript was created by UChicago Argonne, LLC, operator of Argonne National Laboratory. Argonne, a DOE Office of Science laboratory, is operated under contract DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan at http://energy.gov/downloads/doe-public-access-plan.

References

[1] R. Vilim and T. N. Nguyen, "Physics-based Automated Reasoning for Health Monitoring: Sensor Set Selection," in *12th Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies*, Virtual Meeting, June 14-17, 2021.

[2] R. Vilim, T. N. Nguyen and R. Ponciroli, "Explainable and Trustworthy Diagnostics Achievable through Process-Based Automated Reasoning," in *12th Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies*, Virtual Meeting, June 14-17, 2021.

[3] EPRI, "Requirements for On-Line Monitoring in Nuclear Power Plants," Palo Alto, CA, 2008.

[4] A. Plebe and G. Grasso, "The Unbearable Shallow Understanding of Deep Learning," *Minds and Machines,* vol. 29, pp. 515-553, 2019.

[5] G. Marcus and E. Davis, Rebooting AI: Building Artificial Intelligence We Can Trust, Knopf Doubleday Publishing Group, 2019.

[6] V. Venkatasubramanian, R. Rengaswamy, K. Yin and S. N. Kavuri, "A Review of Process Fault Detection and Diagnosis Part I: Quantitative Model-Based Methods," *Computers and Chemical Engineering,* vol. 27, pp. 293-311, 2003.

[7] Z. Gao, C. Cecati and S. X. Ding, "A Survey of Fault Diagnosis and Fault-Tolerant Techniques—Part I: Fault Diagnosis With Model-Based and Signal-Based Approaches," *IEEE Transactions on Industrial Electronics,* vol. 62, no. 6, pp. 3757-3767, 2015.

[8] H. Wang, M.-j. Peng, A. Ayodeji, H. Xia, X.-k. Wang and Z.-k. Li, "Advanced fault diagnosis method for nuclear power plant based on convolutional gated recurrent network and enhanced particle swarm optimization," *Annals of Nuclear Energy,* vol. 15, p. 107934, 2021.

[9] Z. Wang, H. Xia, B. Peng, B. Yang, S. Zhu, J. Zhang and M. Annor-Nyarko, "A multi-stage hybrid fault diagnosis approach for operating conditions of nuclear power plant," *Annals of Nuclear Energy,* vol. 153, p. 108015, 2021.

[10] J. Li and M. Lin, "Ensemble learning with diversified base models for fault diagnosis in nuclear power plants," *Annals of Nuclear Energy,* vol. 158, p. 108265, 2021.

[11] C. He, D. Ge, M. Yang, N. Yong, J. Wang and J. Yu, "A data-driven adaptive fault diagnosis methodology for nuclear power systems based on NSGAII-CNN," *Annals of Nuclear Energy,* vol. 159, p. 108326, 2021.

[12] Z. Wang, H. Xia, J. Zhang, M. Annor-Nyarko, S. Zhu, Y. Jiang and W. Yin, "A deep transfer learning method for system-level fault diagnosis of nuclear power plants under different power levels," *Annals of Nuclear Energy,* vol. 166, p. 108771, 2022.

[13] J. Li and M. Lin, "Research on robustness of five typical data-driven fault diagnosis models for nuclear power plants," *Annals of Nuclear Energy,* vol. 165, p. 108639, 2022.

[14] R. Rai and C. K. Sahu, "Driven by Data or Derived Through Physics? A Review of Hybrid Physics Guided Machine Learning Techniques With Cyber-Physical System (CPS) Focus," *IEEE Access,* vol. 8, pp. 71050-71073, 2020.

[15] S. Shen, H. Lu, M. Sadoughi, C. Hu, V. Nemani, A. Thelen, K. Webster, M. Darr, J. Sidon and S. Kenny, "A physics-informed deep learning approach for bearing fault detection," *Engineering Applications of Artificial Intelligence,* vol. 103, p. 104295, 2021.

[16] M. Russell and P. Wang, "Physics-informed deep learning for signal compression and reconstruction of big data in industrial condition monitoring," *Mechanical Systems and Signal Processing,* vol. 168, p. 108709, 2022.

[17] R. Vilim, A. Grelle, R. Boring, K. Thomas and R. L. T. Ulrich, "Computerized Operator Support System and Human Performance in the Control Room," in *10th International Topical Meeting on Nuclear Plant Instrumentation, Control and Human Machine Interface Technologies,* San Francisco, CA, 2017.

[18] T. N. Nguyen, "Model-Based Diagnostic Frameworks for Fault Detection and System Monitoring in Nuclear Engineering Systems (Doctoral dissertation)," University of Michigan, Ann Arbor, MI, 2020.

[19] T. N. Nguyen, T. Downar and R. Vilim, "A probabilistic model-based diagnostic framework for nuclear engineering systems," *Annals of Nuclear Energy,* vol. 149, p. 107767, 2020.

[20] T. N. Nguyen, R. Ponciroli, T. Kibler, M. Anderson, M. J. Strasser and R. B. Vilim, "A Physics-Based Parametric Regression Approach for Feedwater Pump System Diagnosis," *Annals of Nuclear Energy,* vol. 166, p. 108692, 2022.

[21] B. Kochunas and X. Huan, "Digital Twin Concepts with Uncertainty for Nuclear Power Applications," *Energies,* vol. 14, p. 4235, 2021.

[22] P. Langley, B. Meadows, M. Sridharan and D. Choi, "Explainable Agency for Intelligent Autonomous Systems," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence,* San Francisco, CA, 2017.

[23] H. E. Garcia, S. E. Aumeier, A. Y. Al-Rashdan and B. L. Rolston, "Secure embedded intelligence in nuclear systems: Framework and Methods," *Annals of Nuclear Energy,* vol. 140, p. 107261, 2020.

[24] J. d. Kleer and J. Kurien, "Fundamentals of Model-Based Diagnosis," *IFAC Proceedings Volumes,* vol. 36, no. 5, pp. 25-36, 2003.

[25] N. E. Todreas and M. S. Kazimi, Nuclear Systems Volume I: Thermal Hydraulic Fundamentals, London: CRC Press, 2011.

[26] R. Isermann, Fault-Diagnosis Applications - Model-based condition monitoring: Actuators, drives, machinery, plants, sensors, and fault-tolerant Systems, Berlin: Springer, 2011.

[27] Dominion Energy Virginia, "Revision 54 to Updated Final Safety Analysis Report, Chapter 10, Steam and Power Conversion System," U.S. NRC, 2018.

[28] British Electricity International, Turbines, Generators and Associated Plant, London: Pergamon, 1991.

[29] Heat Exchange Institute (HEI), "Standards for Closed Feedwater Heaters 9th Edition," 2015.

[30] M. Basseville and I. V. Nikiforov, Detection of Abrupt Changes - Theory and Application, Prentice Hall, In., 1993.