Federated-Transfer Learning for Scalable Condition-based Monitoring of Nuclear Power Plant Components

Koushik A. Manjunatha[a], Vivek Agarwal[a], and Harry Palas[b]
[a] Idaho National Laboratory, Idaho Falls, ID, USA
[b] Public Service Enterprise Group, Nuclear LLC, Hancocks Bridge, NJ, USA

Abstract: Condition-based monitoring (CBM) techniques are widely being used for maintenance activities in nuclear power plants (NPPs). As faults are rare events, it is highly unlikely that all potential fault modes are captured for a single component. In addition, fault signatures extracted from a single component cannot be robust enough to handle unseen fault patterns. On the other hand, privacy, security, legal, and commercial concerns restrict data-sharing across different plant systems/different systems within one plant. This research presents federated-transfer learning (FTL) to scale machine learning (ML) models for CBM across a component or plant system by combining federated learning (FL) and transfer learning (TL) approaches. FL enables a centralized server to develop an aggregated global CBM model, while the training data are safely and privately distributed on the devices of plant systems, and TL enables application of the developed aggregated model to different but related systems within the same plant site, or to the same system at different plant sites. FTL was demonstrated using circulating water system data for two plant sites—one two-unit plant and one single-unit plant—to predict the health condition of a circulating water pump. The FTL framework was verified using a multi-kernel adaptive support vector machine and an artificial neural network. The results show significant improvement in prediction performance, reducing overfitting issues and data heterogeneity.

1. INTRODUCTION

Over the years, the domestic nuclear power plant (NPP) fleet has relied on labor-intensive and time-consuming preventive maintenance (PM) programs, thus driving up operation and maintenance (O&M) costs to achieve high-capacity factors [1]. As part of the PM strategy, plant systems, structures, and components (SSCs) undergo manual, labor-intensive periodic maintenance checks—such as inspection, testing, calibration, replacement, and refurbishment—irrespective of their condition. Predictive maintenance (PdM) strategies recommend that action be taken 'as required' based on the health condition of the SSCs.

A well-constructed risk-informed PdM approach for an identified plant asset will take advantage of advancements in data analytics, machine learning (ML), artificial intelligence (AI), physics-informed modeling, and visualization. PdM strategies utilize plant asset current and historic data to develop diagnostic and prognostic models. Diagnostic models provide the current state of health of the plant assets. If the diagnosis indicates a potential incipient fault, the prognostic model predicts the time to failure or the remaining useful life, thus allowing plant personnel to develop a maintenance plan accordingly.

Domestic NPPs are currently focusing on transitioning from a PM strategy to a PdM strategy to achieve long-term economic sustainability in the current competitive energy market [2]. One of the challenges with this transition is the need to develop a scalable PdM strategy that is deployable across plant assets and the NPP fleet. In collaboration with Public Service Enterprise Group (PSEG) Nuclear, LLC, Idaho National Laboratory (INL) has developed a risk-informed PdM strategy, as observed in Figure *1*, and a scalability framework, as shown in Figure *2* [2]. As part of this framework, scalability is defined as expanding capabilities of a target entity to meet current and future application-specific requirements. 'Entity' in this context is each element of the suggested framework. Expanding each element of the framework is outside the scope of this paper, which focuses on the 'methodology' entity.

Figure 1. Research and development for achieving risk-informed PdM strategy.

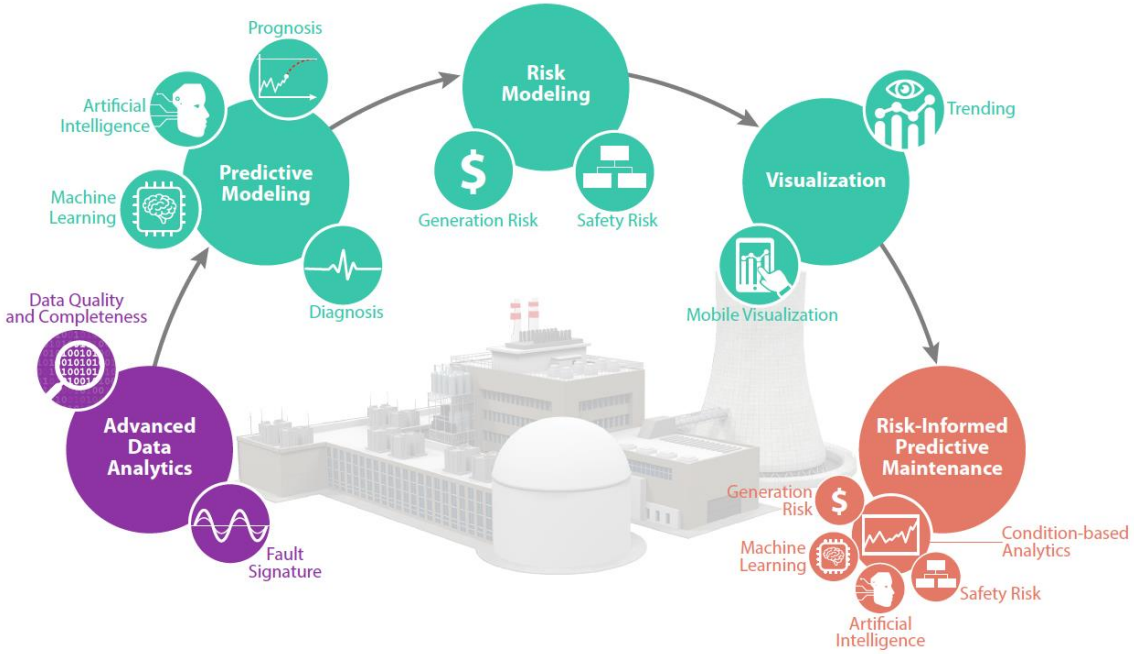
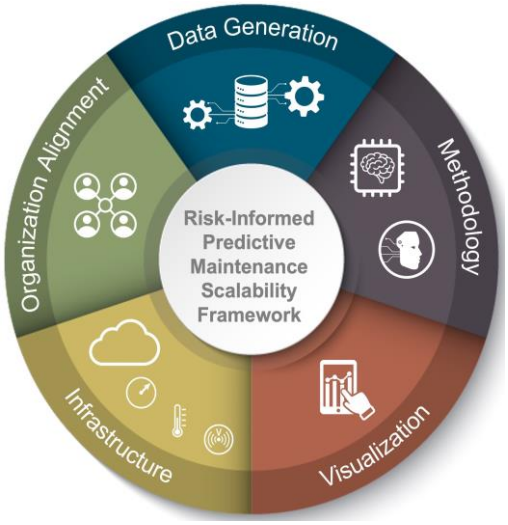


Figure 2. Risk-informed PdM scalability framework.



One of the biggest challenges in building scalable methodologies is the dynamic and heterogeneous nature of the data. The data collected on a particular asset can vary significantly in types and resolutions across plants. A particular signal may not always be available due to instrumentation errors or sensor failure. Finally, plants might add a new sensor modality to enhance situation awareness, resulting in new information. These data dynamics lead to a problem of identifying a 'global' ML model capable of representing the state of the plant asset. This warrants some sort of active learning [4] that constantly updates the ML model based on data dynamics. While active learning could be a potential solution, it requires regular retraining of the ML model, which is undesirable because of the effort required to train the models and the impact on model explainability.

To address dynamics in data, this paper presents the application of a federated-transfer learning (FTL) approach on circulating water system (CWS) data obtained from the PSEG-owned plant sites. The coupling of federated learning (FL) with transfer learning (TL) provides a foundation that aligns with the goals of the scalable risk-informed PdM strategy framework by: (1) collaborating and building a

robust ML model addressing privacy, security, legal, and commercial concerns that restrict data-sharing across the fleet; (2) eliminating the need to create a centralized data repository (i.e., data stays at the source); (3) eliminating the cost associated with building and maintaining a centralized data repository; (4) minimizing the technical challenges, such as cyber-scenarios, latency, throughput, and bandwidth optimization (for wireless data transmission); (5) enabling the transfer of a 'globally' trained and validated ML model to similar plant assets across the NPP fleet; and (6) adapting the transferred ML model with minimal retraining for a different plant asset within the same plant site. A summary of related work on the application of FL towards a PdM strategy is shown in Table 1.

Table 1. Summary of the Related Work on the Application of FL towards PdM.

Algorithm	Ref	Problem Area	Contribution
Federated Multi-Task Learning	[5]	Minimal computational complexity and low communication overhead.	Distributed PdM.
SplitPred	[6]	Collaborative PdM.	Privacy-aware resource sharing.
FL+Blockchain	[7]	PdM on cross-company data.	Collaborative learning with blockchain.
FL	[8]	Equipment condition monitoring in manufacturing environment.	Intelligent prediction and monitoring engine.

2. CIRCULATING WATER SYSTEM

To develop initial scalable methods and models, the CWS at two NPPs was selected as the identified plant asset. The CWS is an important non-safety-related system. As the heat sink for the main steam turbine and associated auxiliaries, the CWSs at Plant Site 1 and Plant Site 2 are designed to maximize steam power cycle efficiency [2]. A CWS consists of the following major equipment [2]:

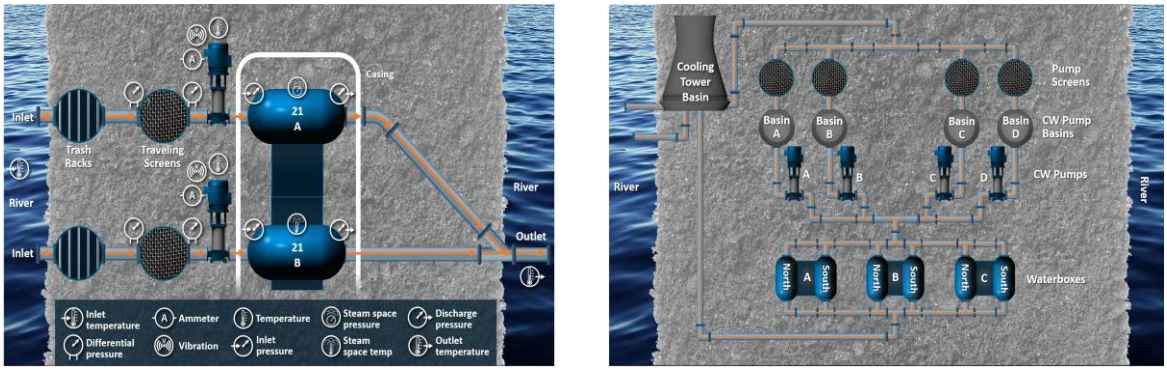
- vertical, motor-driven circulating pumps (e.g., 'circulators'), each with an associated fixed trash rack and traveling screen at the pump intake to filter out debris and marine life
- main condenser (tube side only)
- condenser waterbox air removal system
- circulating water sampling system
- screen wash system
- necessary piping, valves, and instrumentation/controls to support system operation.

The NPP at Plant Site 1 comprises a two-unit pressurized water reactor that features six circulators at each unit. A schematic representation of the main condensers for Plant Site 1 Unit 2 is shown in Figure 3 (left). Each pair of waterboxes is named using the following convention: Unit #, Condenser #A, Unit #, and Condenser #B. The NPP at Plant Site 2 consists of a single-unit boiling water reactor with four circulators. A schematic representation of Plant Site 2 CWS is shown in Figure 3 (right). Several distinct differences can be seen between the two plants. These include: (1) the water supply to the Plant Site 2 CWS comes from a cooling tower water basin, not directly from the river; (2) the Plant Site 2 CWS does not have traveling screens, but each circulator has a single-pump screen to prevent debris transmission to the waterboxes; and (3) the Plant Site 2 CWS has four circulators feeding six waterboxes via a common header, unlike the Plant Site 1 CWS, in which each waterbox had its own circulator.

A general functional description of the Plant Site 1 CWS, component integration, and design basis are found in [2]. The description of the Plant Site 2 CWS is similar, with minor differences in the integration due to the previously highlighted changes in the design basis. The CWS equipment that most impact the unit's gross load output are the circular water pumps (CWPs) and their motors. The number of CWPs operating together impacts the generated gross load. A derate is a percentage decrease in gross

load, which is due to unavailability of plant assets supporting power generation. A trip is a reactor shutdown in which one or more plant assets are unavailable, thereby leaving the plant unable to maintain safe reactor operation. An outage occurs when reactor power is at zero for an extended period (e.g., usually less than a month) to address scheduled fuel cycle maintenance. During this time, all plant assets are non-operational, leading to an observable pattern in the plant gross load, which is labeled as operational, derate, trip, and outage.

Figure 3. Schematic representation of the CWS at Plant Site 1 Unit 2 CWP combination 21A and 21B (left) and Plant Site 2 (right).



2.1 Plant Site 1

The Plant Site 1 Unit 1 and Unit 2 CWS process data are collected once per minute and stored in the plant's OSIsoft process information system. Due to file size restrictions, the project team received CWS process data on an hourly frequency for both units from 2009 to 2020. Continuous CWP motor current data for both Units 1 and 2 are available only from September 2017 onward. Figure 4 (top) shows a sampling of CWS process data for a Plant Site 2 unit.

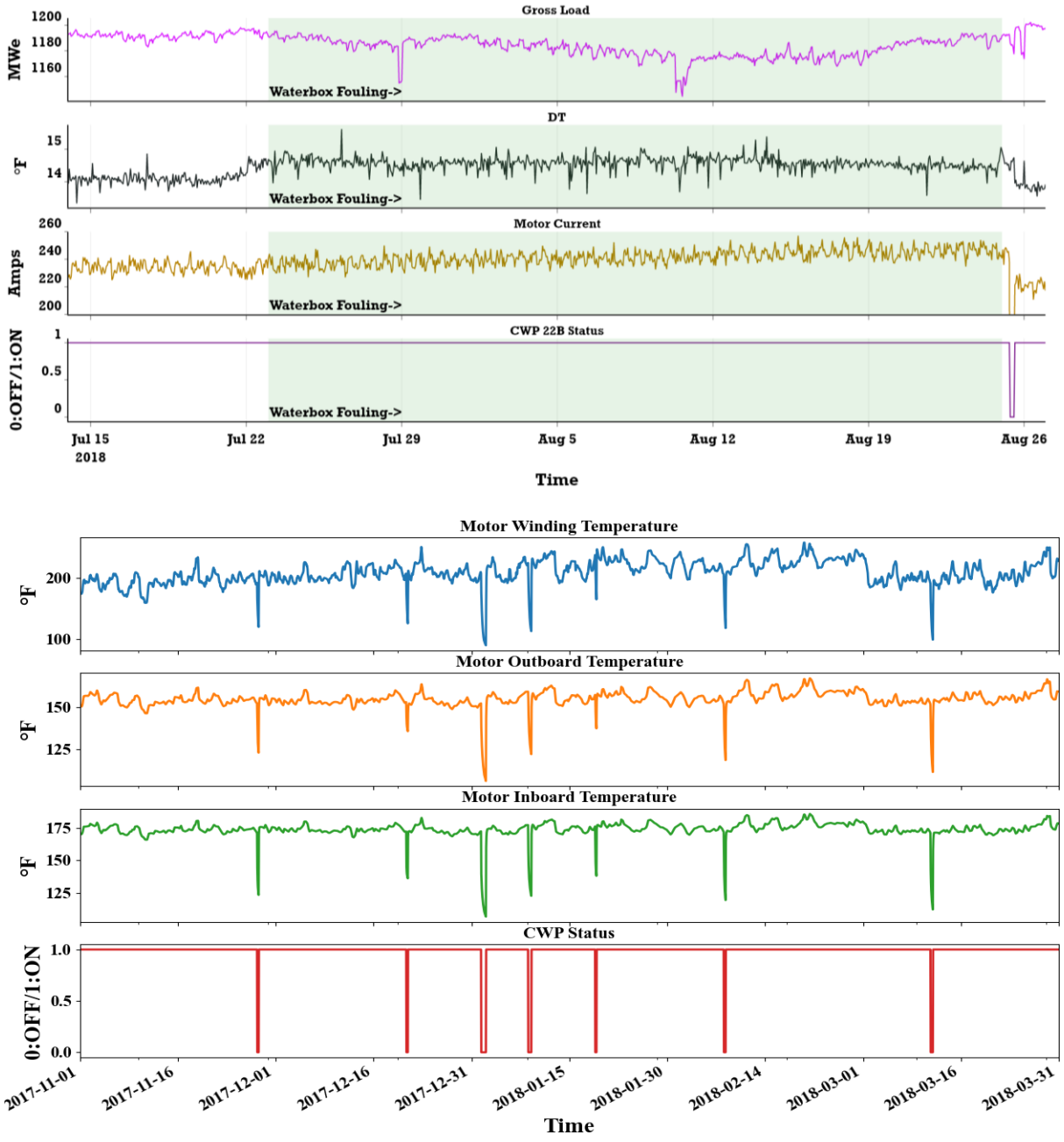
2.2 Plant Site 2

The Plant Site 2 CWS data consists of hourly measurements spanning from 1 January 2010 to 18 May 2021. Overall plant status data include gross load in megawatts and a 15-minute average of the ambient outside temperature. Each of the four CWP measurements include the basin level, discharge pressure (DP), motor winding temperature, motor axial position, outbound and thrust bearing temperatures, and discharge valve position. The pump run status (on/off) is also recorded. Figure 4 (bottom) shows a sampling of CWS process data from the Plant Site 2. Table 2 details CWP specific measurement types observed at both Plant Site 1 and Plant Site 2.

Table 2. CWP specific measurement types observed at Plant Site 1 and Plant Site 2 NPPs.

Measurements	Plant Site 1	Plant Site 2
Timestamp	Y	Y
CWP Status	Y	Y
Gross Load	Y	Y
Differential Temperature (DT)	Y	Y
Motor Current	Y	×
Motor Temperatures (stator, outboard, inboard)	Y	Y
Motor Vibration (axial, outboard, inboard)	Y	×

Figure 4. An example of changes to the CWS process data both before and after waterbox fouling. (a) Plant Site 1 Unit 2's CWP 22B (Top) and (b) Plant Site 2's waterbox A (Bottom).

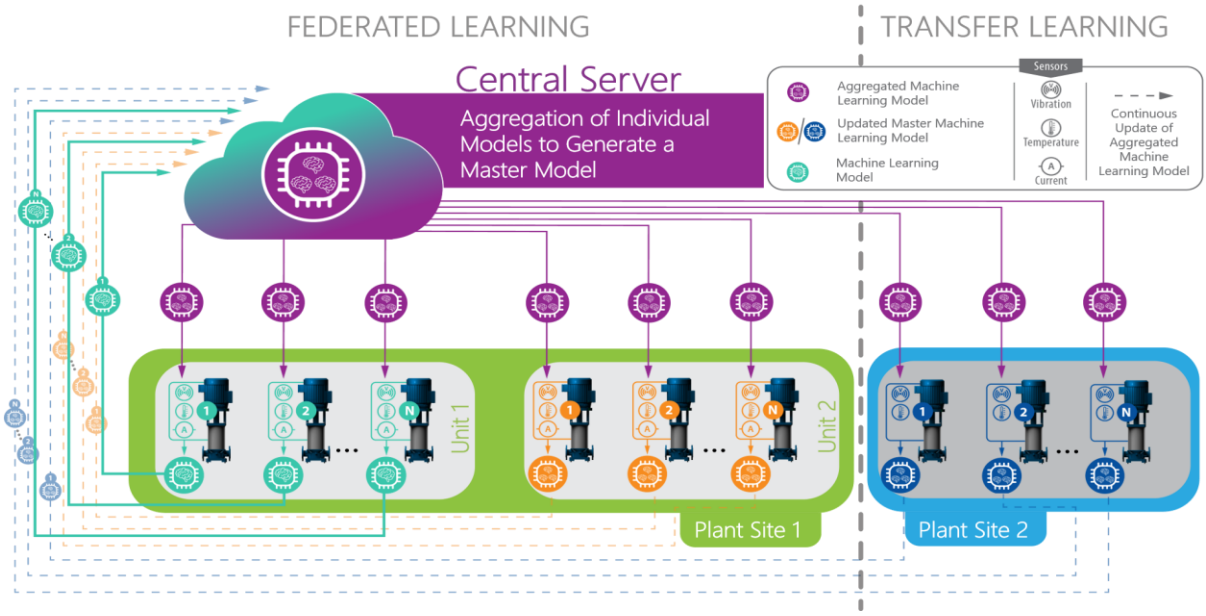


3. FEDERATED-TRANSFER LEARNING

FTL [3] is a combination of FL and TL. In FTL, datasets differ in the feature space. This applies to datasets collected from NPPs of different but similar nature. Due to differences in the nature of operation and monitoring, such plants may share only a small overlap in the feature space. FL is a collaborative learning where many clients collaboratively train a model under the orchestration of a central controller without exchanging the user's original data. FL enables focused data collection and data minimization by reducing systematic privacy risks and costs resulting from traditional, centralized ML. The FL process is typically driven by a model engineer developing an AI model for a particular application. TL builds an effective model for the target domain while leveraging knowledge from the other (source) domains. The main advantage of TL is that the training time is reduced significantly, while a very small amount of training data or no training data are required to leverage pre-trained models.

Using the FTL approach, this work focuses on: (1) developing an individual component level model using component-specific available data sources; (2) consolidating the knowledge gained from individual component models for a given plant asset into a master model; (3) using the master model for diagnostic and prognostic estimations of the entire system; and (4) applying (i.e., transferring) the master model for diagnostic and prognostic estimations of similar plant systems either at the same plant site or a different NPP entirely. The schematic representation of FL and TL is shown in Figure 5. The FL learning is demonstrated on Plant Site 1, while the TL framework is demonstrated on Plant Site 2. The FTL framework is demonstrated on neural networks (NNs) [9] and support vector machines (SVMs) [10].

Figure 5. FTL framework for condition assessment of vertical motor-driven pumps of the CWS across two plant sites.



4. NUMERICAL RESULTS

4.1 Plant Data for FTL Modeling

This section discusses two examples of waterbox fouling, one each from Plant Sites 1 and 2, which will be used in the development of FTL for condition-based monitoring (CBM). The two examples highlight the similarities and differences in fault signatures for waterbox fouling and are a perfect example of the need for FTL in predictive modeling. The primary issue noted with the Plant Site 1 CWS is fouling of the waterboxes by grass and debris. Fouling of the waterboxes typically occurs due to the accumulation of grass/debris in the waterbox, thus resulting in condenser tube blockage and reduced circulator water flow. This is a unique and frequent issue at Plant Site 1 where the CWP intake comes directly from the river, resulting in a significant quantity of grass/debris. The grassing season typically occurs between February 1 and May 31. Grassing often emerges from the river during high-wind conditions associated with storms. During these periods, the motor current can oscillate with river level changes. Operations monitors the waterbox motor current and inlet pressure, and schedules waterbox cleanings based on deviations in motor current and inlet pressure when compared against the historical baseline data. Waterbox fouling is typically identified via motor current increase (also, though far less frequently, motor current decrease), inlet pressure increase, waterbox differential temperature (DT) increase, and condenser thermal performance loss.

Figure 4 (top) shows an instance of waterbox fouling diagnosed in Plant Site 1 Unit 2's CWP 22B. An upward drift in DT and motor current was identified on 23 July 2018. Consequently, the gross load began to dip. Note that in Figure 4 (top), the CWP 22B motor current increased from 231 to 245 amps,

and the DT increased from 14°F to 16°F, with the gross load not trending as expected. The motor current and DT decreased to 220 amps and 14°F, respectively, following the waterbox cleaning on 25 August 2018, resulting in a 30–40 MWe improvement in gross load. The waterbox fault and approximate date of the shutdown were found by searching the work order database and narrative log information.

In the case of Plant Site 2, waterbox fouling is not a major fault, yet still of interest. The cause of waterbox fouling for Plant Site 2 is once again debris (limited grassing) in the water circulated in and out of the cooling tower basin. Figure 4 (bottom) shows an instance of waterbox fouling in Plant Site 2 waterbox A. Under normal operating conditions with no faults, the DP across the Plant Site 2 CWPs are, on average, 40–41 pounds per square inch gauge (PSIG). Note that in Figure 4 (bottom), on 23 December 2017, CWP A's DP began trending upward and spiked above 43 PSIG. Following the DP trend, the DT across the north and south ends of waterbox A also trended upward in the same amount of time. These slow, steady increases in differential pressure and DT trends are indicative of waterbox fouling. Following a waterbox cleaning around 23 January 2018, the DP reduced to near 41 PSIG and the DT stabilized as well.

These two examples show that different fault features can indicate the same fault. Developing a comprehensive fault signature for each fault mode is key to achieving scalable, accurate predictive models. For other CWP fault signatures, see reference [2].

4.2 Feature Extraction

To develop the FTL-based predictive model, features were extracted based on the identified fault signatures:

1) Plant Site 1: From the CWS-associated plant operational data, the following features are extracted for each motor and pump (M&P) set:

- DT is calculated as the difference between the outlet water temperature associated with the M&P set and the inlet river temperature.
- The measured motor inboard (MIB) temperature, motor outboard (MOB) temperature, and motor stator temperature.
- The M&P run-hours from one replacement to the next are considered in calculating the motor age (MAge) and pump age (PAge), as determined from historical CWS M&P replacement/refurbishment dates.
- To consider the seasonal effects on the data, the week of the year (WoY) is calculated for every timestamp and is used as a feature.

Thus, a total of seven features are extracted from the CWS plant operational data for each M&P set. Detailed information on the feature extraction from the plant operational data—as well as from the vibration data—can be found in reference [2]. For model development, plant operational data after 2016 were considered because Plant Site 1 adopted a new six-year CWP replacement PM at that time. Since 2016, each unit of the Plant Site 1 NPP has periodically replaced their CWPs as per the updated PM strategy. Based on the replacement date of each CWP, the age of the M&P set is estimated. If any faults in the M&P are identified after their replacement, the data corresponding to that fault and time period is labeled as unhealthy; otherwise, it is labeled as healthy.

2) Plant Site 2: From the Plant Site 2 CWS-associated plant operational data, the DT, MIB temperature, MOB temperature, motor stator temperature, and WoY features are extracted for each M&P set. As there were no historical CWS M&P replacement/refurbishment dates available for Plant Site 2, the MAge and PAge are not calculated. There are other faults represented in the CWP data prior to 2016 that are not discussed here for clarity; however, the approach is extendable to those faults as well.

4.3 FL-based CWP Motor Health Prediction

FL was demonstrated using an SVM, which classifies whether a CWP is in a healthy or unhealthy state, on the Plant Site 1 data with each local model being developed for a pair of CWPs connected to a common waterbox, as shown in Figure *3* (left). Since there are three waterboxes for each Plant Site 1 Unit, this gives six local models that will be combined into a master model via the FL approach, as observed in Figure 5. The samples were grouped based on the CWP combinations and split into training and test samples in accordance with an 80:20 ratio, as shown in Table 3.

Table 3. Data Split into Training and Test Sets Among Groups for CWP Condition Prediction.

Data Group	Training Samples	Test Samples
Group 1 (CWP 11A and CWP 11B)	6174	974
Group 2 (CWP 12A and CWP 12B)	8303	1309
Group 3 (CWP 13A and CWP 13B)	4366	822
Group 4 (CWP 21A and CWP 21B)	1720	288
Group 5 (CWP 22A and CWP 22B)	2356	476
Group 6 (CWP 23A and CWP 23B)	1496	358

The results of the SVM-based individual learning and FL on Plant Site 1 are given in Table 4. From Table 4, it is seen that individual models from each group achieved a performance of close to 100% in most of the SVM models. This is a clear indication of overfitting in individual models, with the models being unable to predict other datasets or unseen data with the same accuracy. In addition, for some models, the accuracy of the test samples is higher than that of the training samples, since the test data were sometimes easier for the model to predict than the training data. After applying FL-based model aggregation (by averaging the kernel matrix from all the local models) and retraining each individual model, the accuracy levels came down for most of the models, but the performances remained at acceptable levels. FL aggregation over several iterations can further improve overfitting, while maintaining acceptable performance of the diagnostic model. This exemplifies that FL-based model aggregation enables aggregation of diagnostic models from the component level to the plant level. In addition, the fact that the models are trained with limited datasets also impacts the performance of FL. It is anticipated that FL performance will improve with larger training datasets.

As a parallel method to FL via SVM, we explored the feasibility of using artificial NNs to perform FL on the Site 1 data. As with the SVM approach, we trained six different local models, one for each CWP combination at Site 1. For consistency, the same full dataset and training/testing splits used for the SVMs were used in the NN model. In NN model, an aggregated model is generated by averaging the weight and bias parameters from all the six local models.

The high training accuracies of the individual models for NNs suggest that several of the current individual models are overfit to the training data. Future iterations should focus on further regularization techniques. This overfitting is at least partly reflected in the low-test accuracy displayed by many of the models. (Note that an extremely high individual test accuracy can indicate a high level of model bias and poor generalizability of the model to new observations.) Further work must be done to build more robust individual models that generalize more readily to previously unseen data. However, as noted above, a sufficiently high accuracy is obtainable while retaining a uniform (and thus simpler) architecture for all individual models.

For the FL approach, we see much stronger test set performance on the CWPs that previously showed low accuracy in the individual phase. The added information afforded by examining all pump data en masse provided clear advantages to the federated-model building process.

Table 4. Individual Training and FL Performance on Plant Site 1 Data Using SVM and NN.

CWP Combination	SVM				NN			
	Individual Learning		FL		Individual Learning		FL	
	Train	Test	Train	Test	Train	Test	Train	Test
11A and 11B	97.11%	95.58%	98.32%	93.73%	100%	88.30%	97.96%	96.30%
12A and 12B	100%	99.92%	100%	99.92%	93.75%	97.33%	98.71%	96.02%
13A and 13B	100%	98.90%	100%	98.90%	100%	97.57%	100%	94.28%
21A and 21B	100%	99.30%	100%	99.30%	100%	75%	100%	96.88%
22A and 22B	100%	98.94%	98.64%	91.80%	100%	75%	100%	98.118%
23A and 23B	100%	98.32%	99.26%	99.16%	100%	82.12%	98.54%	99.72%

4.4 TL-based CWP Motor Health Prediction

TL aims to train a model on data from one domain in Plant Site 1 and adapt that model to another domain in Plant Site 2 by partial or full retraining. For the demonstration of TL on the Plant Site 2 data, binary classification was considered with the healthy and unhealthy class labels being determined based on the plant operational data. For the unhealthy state, waterbox fouling fault data were extracted; data prior to the occurrence of waterbox fouling were considered healthy. Healthy and unhealthy samples extracted from CWPs A and C are shown in Table 5. For the extracted samples, the master model from FL is used—both with and without retraining for comparison—to predict the condition of the CWP using the Plant Site 2 data. Note that for TL without retraining, the extracted data are not split into training and test data; instead, all the data are considered test data and the FL model will be used to predict the labels on the entire dataset.

Table 5. Plant Site 2 Data for TL.

CWP	Date	Number of Samples	
		Healthy	Unhealthy
A	November 1, 2017 - March 1, 2018	1916	966
C	April 1, 2017 - April 1, 2018	6502	2558

For SVM-based TL, the overall performance on both CWP A and CWP C data is around 80%. The approach involves using all the samples from CWP A and CWP C as test data to classify health using the master model from the FL framework. The performance dictates that the SVM parameters must be further optimized to improve prediction accuracy. Typically, in TL, a small set of sample data is used to retrain the transferred model to fine-tune the model parameters for the new environment in Plant Site 2. For example, only 10–20% of the total number of samples will be used to retrain the model and optimize the parameters of the SVM for the Plant Site 2 data. After retraining with 20% of the data, the performance of CWP A did not improve, whereas the performance of CWP C significantly improved—to higher than 95% accuracy. The performance of CWP A with TL indicates there were insufficient samples for building the ML model.

For comparison with TL, individual models were also trained on the Plant Site 2 data, with an 80:20 split between the training and test data. Individual model performance—particularly for CWP A—clearly shows the same overfitting trend as seen in the FL case. More samples for training are required to generalize the model and avoid overfitting.

The same overall Plant Site 2 dataset from SVM was used as the starting point for TL with NNs. The classification target of healthy or unhealthy is the same as was described previously. Unlike with the SVM approach, the NN approach to TL amounts to more than merely applying the raw FL model from Plant Site 1 to the new data. Furthermore, no model is fit solely to the Plant Site 2 data. This affords us

the benefit of truly transferring knowledge from one domain to another. After examining the initial results from the FL model developed above, we then further trained the FL-associated NN by using some of the Plant Site 2 data. We used a standard 80:20 train-test split of the Plant Site 2 data. No further training was done on the FL model for this step; it was "raw" and was employed as though the Plant Site 2 data were merely yet unseen Plant Site 1 data: the overall average accuracy was just above 75%. Wanting to take advantage of the benefits of TL, we then used the FL results presented here as a weight initialization (initial model) to further train the NN only for the five epochs for both CWP A and CWP C. The results given in Table 6 for CWP C are the result of this TL training step. For CWP A, we obtain around 88% accuracy with the five training epochs. Training for a further 15 epochs yielded the training and test accuracies given in Table 6.

Table 6. TL Performance on Hope Creek Using SVM and NN from FL.

CWP Combination	**SVM**				**NN**			
	Individual learning		**TL without retrain**		**TL without retrain**		**TL with retrain**	
	Train	Test	Train	Test	Train	Test	Train	Test
CWP A	99.9%	66.1%	-	80.74%	-	80.43%	93.93%	94.01%
CWP C	99.92%	93.1%	-	79.92%	-	76.90%	97.38%	97.35%

4. CONCLUSION

The FL using the Plant Site 1 data and the TL using the Plant Site 2 data were verified using SVM and NN. For SVM, the features were grouped based on the type of measurement and trained with separate kernel functions. While the performance of SVM for FL is satisfying, the performance of TL can be further improved by adopting a partial retraining approach that optimizes transferred model parameters without going through comprehensive training. The performance of NN is comparable on FL with SVM. But TL for NN performed better than SVM when TL is performed with retraining. The performance can be improved using more training samples. This demonstrated the significance of FTL approach and avoids building of exclusive predictive models for each NPP and each system.

Acknowledgments

This research was made possible through funding from the U.S. Department of Energy (DOE)'s Light Water Reactor Sustainability program under the contract DE-AC07-05ID14517. We are grateful to Alison Hahn of DOE and Bruce P. Hallbert and Craig A. Primer at Idaho National Laboratory (INL) for championing this effort.

5. REFERENCES

[1] S. W. Foon and M. Terziovski, "The impact of operations and maintenance practices on power plant," *Journal of Manufacturing Technology Management*, 2014.

[2] V. Agarwal, K. A. Manjunatha, J. A. Smith, A. V. Gribok, V. Yadav, H. Palas, M. Yarlett, N. Goss, S. Yurkovich, B. Diggans, N. J. Lybeck, M. Pennington, and N. Zwiryk, "Machine Learning and Economic Models to Enable Risk-Informed Condition Based Maintenance of a Nuclear Plant Asset," Idaho National Laboratory, Idaho Falls, ID, USA, 2021.

[3] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated Machine Learning: Concept and Applications," *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(2), 1-19, 2019.

[4] B. Settles, "Active Learning Literature Survey," 2009.

[5] K. J. Tilsted, “Federated Multi-Task Learning on Acoustic Signals for Predictive Maintenance,” 2021.

[6] S. Bharti and A. McGibney, “Privacy-Aware Resource Sharing in CrossDevice Federated Model Training for Collaborative Predictive Maintenance,” IEEE Access, 2020.

[7] M. Mohr, C. Becker, R. Möller, and M. Richter, Towards Collaborative Predictive Maintenance Leveraging Private Cross-Company Data, INFORMATIK, 2020.

[8] J. Zhou, X. Li, A. J. Andernroomer, Y. S. Wong, and G. S. Hong, “Intelligent Prediction Monitoring System for Predictive Maintenance in Manufacturing,” in *31st Annual Conference of IEEE Industrial Electronics Society*, 2005.

[9] A. K. Jain, J. Mao, and K. Mohiuddin, “Artificial Neural Networks: A Tutorial,” *Computer*, 29(3), 31-44, 1996.

[10] C. Cortes and V. Vapnik, “Support-Vector Networks,” in *Machine Leaming*, Springer, 273-297, 1995.