IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 18 July 2025, accepted 4 August 2025, date of publication 15 August 2025, date of current version 25 August 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3599360

RESEARCH ARTICLE

Wind Turbine Condition Monitoring Based on Intra- and Inter-Farm Federated Learning

ALBIN GRATALOUP[1], STEFAN JONAS[1,2], AND ANGELA MEYER[1,3]
[1]School of Engineering and Computer Science, Bern University of Applied Sciences, 2501 Biel, Switzerland
[2]Faculty of Informatics, Università della Svizzera italiana, 6900 Lugano, Switzerland
[3]Department of Geoscience and Remote Sensing, Delft University of Technology, 2628 CD Delft, The Netherlands

Corresponding author: Stefan Jonas (stefan.jonas@bfh.ch)

This work was supported by the Swiss National Science Foundation under Grant 206342.

ABSTRACT As wind energy adoption is growing, ensuring the efficient operation and maintenance of wind turbines becomes essential for maximizing energy production and minimizing costs and downtime. Many AI applications in wind energy, such as in condition monitoring and power forecasting, may benefit from using operational data not only from individual wind turbines but from multiple turbines and multiple wind farms. Collaborative distributed AI which preserves data privacy holds a strong potential for these applications. Federated learning has emerged as a privacy-preserving distributed machine learning approach in this context. We explore federated learning in wind turbine condition monitoring, specifically for fault detection using normal behaviour models. We investigate various federated learning strategies, including collaboration across different wind farms and turbine models, as well as collaboration restricted to the same wind farm and turbine model. Our case study results indicate that federated learning across multiple wind turbines consistently outperforms models trained on a single turbine, especially when training data is scarce. Moreover, the amount of historical data necessary to train an effective model can be significantly reduced by employing a collaborative federated learning strategy. Finally, our findings show that extending the collaboration to multiple wind farms may result in inferior performance compared to restricting learning within a farm, specifically when faced with statistical heterogeneity and imbalanced datasets.

INDEX TERMS Condition monitoring, federated learning, industrial fleets, normal behavior model, privacy-preserving, wind energy, wind farm, wind turbine.

I. INTRODUCTION

A. PROBLEM FORMULATION

The deployment of wind turbines for renewable energy generation is witnessing exponential growth globally [1], [2], driven by the transition towards sustainable energy sources. Ensuring the efficient and reliable operation of wind turbines is critical to maximizing energy production and minimizing downtime and maintenance costs. Condition monitoring and anomaly detection play a pivotal role, offering insights into the health and performance of critical components. Deep learning methods, in particular, have risen as an efficient approach to anomaly detection [3], [4], [5], [6]. However, the demanding data prerequisites of deep learning models present a major challenge as they necessitate either an abundance of labeled data from faulty operation or, in our scenario, a large amount of fault-free data for training a *normal behaviour model* (NBM). A NBM operates by predicting target variables like component temperatures or power output that are crucial for assessing system health or performance. Anomalies are identified when the predicted target variable diverges significantly from the measured value of the target variable, such as detecting an abnormal spike in component temperatures compared to the system's normal operational values. Condition monitoring and anomaly detection are extensively studied fields within the area of wind turbine operations. In recent years, deep learning has emerged as a particularly promising approach for condition monitoring

The associate editor coordinating the review of this manuscript and approving it for publication was Long Xu.

tasks. Among the methodologies employed, NBMs have gained prominence. These models rely on the comparison between critical features measured in wind turbines and their corresponding predicted values, serving as indicator for assessing wind turbine health [7], [8], [9].

Training an effective NBM requires a substantial amount of data, which can be time-consuming or even impractical to obtain. A single wind turbine would require a significant amount of time to gather enough data to train a representative and accurate NBM. Scarcely available data may lead to an unrepresentative training set. For instance, measurements observed during a short two-week period in summer may lack certain wind speeds, power generation behaviour, or only cover a limited range of component temperatures. They would therefore not represent a sufficient range of possible normal WT operational states that are expected to occur at later stages, i.e., during test time. This lack of representative training data may cause a substantial drop in fault detection performance, as observed in [10]. For wind turbines, the fastest way to amass sufficient data is by collecting training data from multiple turbines. However, this approach raises significant data privacy concerns as manufacturers and operators are hesitant to share operational data due to strategic business interests [11]. Additionally, data sharing may introduce risks and complexities in complying with data privacy regulations, such as the EU General Data Protection Regulation (GDPR) [12]. To address this issue, we propose privacy-preserving collaborative learning methods to leverage training data collected from multiple wind turbines simultaneously. Federated Learning (FL) emerged as a promising paradigm to address these challenges [13]. FL enables collaborative decentralised model training across multiple wind turbines while preserving their data privacy. By exchanging only FL model parameters and not operational data, the sensitive operation data of each wind turbine remains local and inaccessible to others. This approach allows wind turbines to collaboratively train an effective NBM with less data, without compromising their privacy. Federated learning has gained traction across various domains [14]. It was also adopted in renewable energy sectors [13], notably in wind energy applications, for tasks such as wind power forecasting [15], [16], [17], [18], [19], to obtain significantly more accurate forecasts compared to local models. It also has shown promising results in fault detection applications, exhibiting improved performance over local training methodologies for tasks such as blade icing detection [20], [21], [22], [23], fault detection [24], [25] and condition monitoring with an NBM [10]. Despite these advancements, the application of FL for training NBMs for wind turbines remains largely unexplored.

B. RESEARCH CONTRIBUTIONS

Our study is the first to demonstrate collaborative federated learning strategies across different wind farms (inter-farm learning) for training NBMs. We demonstrate collaborative inter-farm learning for condition monitoring and fault detection, and compare it to collaborative learning within a single wind farm (intra-farm). We find that federated learning across multiple wind turbines consistently outperforms normal behaviour models trained on a single wind turbine, especially when training data is scarce. Further, we investigate the effects of statistical heterogeneity between different wind turbines and wind farms in collaborative machine learning model training. We show that extending collaboration to multiple wind farms may result in inferior performance compared to intra-farm learning if model training data are statistically heterogeneous or imbalanced. Finally, we present the first investigation of how FL affects the amount of training data needed to achieve accurate condition monitoring and its potential in accelerating model deployment. Our findings show that the amount of historical data necessary to train an accurate normal behaviour model can be significantly reduced by collaborative federated learning.

The objectives of our study are twofold: 1. We assess the effectiveness of collaborative federated learning strategies among wind turbines of multiple wind farms, comparing intra- and inter-farm collaborative federated learning (Fig. 1). We ask which learning method (local, intra-, or inter-farm) yields the best results to ensure the reliability of Federated Learning methods. 2. We quantify the time savings in collecting training data for a NBM through collaborative FL across multiple wind turbines compared to collecting training data from a single WT (referred to as "local" data). This investigates the concrete expected gain from using Federated Learning on newly installed farms. These findings should provide further insights into the added value of FL methods in this industrial application.

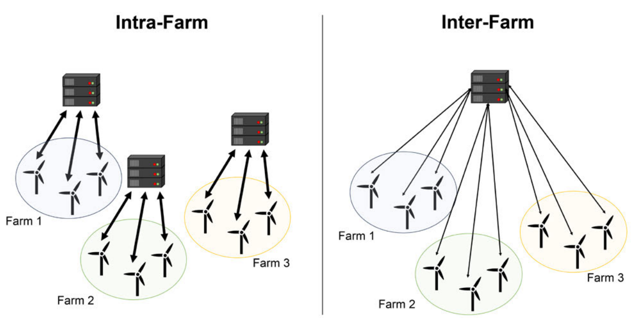

FIGURE 1. Left: Intra-farm learning, only the turbines from the same wind farms collaborate. Right: Collaborative inter-farm learning on the turbines of all wind farms.

II. FEDERATED LEARNING FOR CONDITION MONITORING OF WIND TURBINES

FL is a collaborative deep learning framework that involves distributed participants referred to as clients. In our scenario, each wind turbine acts as an individual client and aims to collaboratively train a machine learning model for condition monitoring. In FL, it is crucial that the clients never share their locally stored data in order to preserve data privacy. The iterative FL training process involves that clients train local models, such as a NBM, using only their private local dataset, and then transmit the model parameters to a central server where they are aggregated. This approach ensures the privacy of locally stored client data and can provide a viable solution to overcome data scarcity [10], [13]. FL with a central server involves the following iterative steps:

Algorithm 1 Centralized Federated Learning

Require: R: number of training rounds, central server, initialized server model
for $t = 1 \cdots R$ **do**
- N clients receive a global FL model from the server
- All clients independently perform training updates on this model using only their local datasets $\mathscr{D}_i$, $i = 1 \ldots N$;
- The clients send the parameters of their updated local models $\mathscr{M}_i$, $i = 1 \ldots N$, to the server
- The server aggregates the parameters of all models $\mathscr{M}_i$, to obtain the updated global FL model
end for

The most widely applied FL framework is the Federated Averaging (`FedAvg`) algorithm [26] in which the aggregation step consists of averaging the received model parameters

$$\omega^{t+1} = \sum_{i=1}^{N} \frac{n_i}{n} \omega_i^{t+1} \tag{1}$$

where ω^t and ω_i^t denote the global model parameters and the model parameters of client i, respectively, in training round t. n_i denotes the amount of data available to client i while n is the total amount of available data across all clients involved in the aggregation.

Although `FedAvg` has demonstrated empirical success and serves as a cornerstone in many FL algorithms, its effectiveness in real-world applications can be hindered by *statistical heterogeneity*, where the data distributions differ across the clients participating in the learning process. The clients' data may differ in their statistical properties and in size, for example, because of differences in feature distributions or in label distributions. In wind turbines, individual turbines may display differing mechanical characteristics and possibly even differing turbine models and supervisory control and data acquisition (SCADA) systems may be involved. Statistical heterogeneity poses a challenge for FL model training and convergence because the aggregated model must learn to generalize across the diverse datasets. The variations among clients result in differences in the statistical distributions of their local datasets, leading to non-identically distributed (non-iid) data distributions.

Fleets of industrial assets, such as wind turbines, can display significant statistical heterogeneity across clients. In such settings, global FL models tend to exhibit suboptimal performance [27], [28] compared to locally trained models. The latter may even achieve higher accuracy than their globally trained counterparts [10]. As a result, some clients may lack incentives to participate in training the global FL model ([10], [29], [30]). To address this challenge, Personalised FL (PFL) has been proposed to customize global FL models to individual clients. PFL retains the advantages of collaborative learning while tailoring the resulting global FL models to each client's specific local data. Various PFL approaches exist, including client clustering [31], [32], [33], [34], personalised model layers [35], meta-learning [36], and fine-tuning methods [37]. We refer to [38] and [39] for a comprehensive overview of customisation approaches, and to [13] for PFL applications in renewable energy contexts.

III. INTRA- AND INTER-FARM FEDERATED LEARNING OF NORMAL BEHAVIOUR MODELS

Training a NBM on data from a single WT requires a significant amount of data representative of the WT's normal operational behaviour, which may not always be available. For example, this is typically the case in newly installed wind farms or after component updates and replacements. The resulting lack of data to train a representative and accurate model is known as the *cold start problem* in computer science, e.g., [40]. We propose to exploit data gathered from multiple wind turbines to reduce the amount of time required for collecting data for training NBMs. We refer to the time savings as the *cold start speed up* because it is the speed up achieved by training a NBM from scratch through collaborative training rather than training on only local data. Due to privacy considerations, the data from individual turbines are kept confidential, so no data sharing with other wind turbines or servers is allowed. We employ FL for collaborative learning across different wind turbines and different wind farms. We assess the impact of having multiple turbines with different specifications in different wind farms on the efficacy of collaborative learning. Condition monitoring often relies on NBMs which simulate the normal operation behaviour of the monitored WT components under the current environmental and operation conditions. NBMs are trained on WT data from fault-free operation periods, and allow quantifying the deviations between the measured target variables and their expected values as simulated by the NBM.

A. WIND FARM DATA

We investigate FL for wind turbine condition monitoring using SCADA data of WTs from three different wind farms (Table 1).

The wind farms provide different WT models and site conditions, which can give rise to statistical heterogeneity of the WTs' SCADA variables. An illustrative example of the present statistical heterogeneity in our selected wind farms is shown in Fig. 3. The wind farms Penmanshiel and Kelmarsh exhibit similar configurations, sharing identical SCADA variables, whereas the EDP wind farm features a different SCADA system. We chose 10-minute averages of wind speed, ambient temperature, and wind direction as input features for the NBM, with gear-bearing temperature as the

TABLE 1. Description of the three wind farms used in this study.

Wind farm	Penmanshiel	Kelmarsh	EDP
Location	55.906°N, 2.262°W	52.402°N, 0.945°W	unknown
Turbine model	Senvion MM82	Senvion MM92	unknown
Time period	2016–2021	2016–2021	2016–2017
Number of turbines	14	6	4
Rated power	2.05 MW	2.05 MW	2.0 MW
Rotor diameter	82 m	92 m	unknown
Cut-out wind speed	25 m/s	24 m/s	25 m/s
SCADA data source	[41]	[42]	[43] & [44]

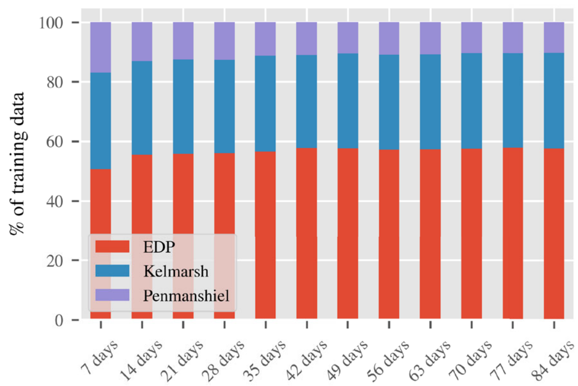

FIGURE 2. **Proportion of data available per wind farm when taking different time windows of data (1 to 12 weeks by 1-week intervals). Due to a larger share of missing data, SCADA data from the Penmanshiel farm represents only a comparably small amount of total data, whereas the EDP farm represents a share of roughly 50% to 60% across varying training set sizes.**

target variable to be predicted. Our selection follows one possible common set-up (e.g., [45]), while we further restrict ourselves to SCADA variables matching across our differing SCADA systems. The SCADA data were cleaned manually by removing curtailment periods and outliers (removing times when the turbine was not operational with wind for example). During the creation of input windows, we use linear interpolation to fill the gaps smaller than 1h and remove windows containing a larger gap. Wind speed and ambient temperature were normalised, while wind direction was cyclically encoded by a sine-cosine transformation. Details of our pre-processing are available in our provided implementation.

The SCADA datasets of the EDP, Kelmarsh and Penmanshiel wind farms contain significant fractions of missing values of 1%, 5%, and 30%, respectively. This leads to the data balance depicted in Fig. 2. The FL algorithm `FedAvg` applied in our study weighs the WTs' contribution to the training in accordance with the fraction of training data available from them (eq. 1). This data imbalance, depicted in Fig. 2, can therefore affect the learning process in intra- and inter-farm FL.

For our experiments in this case study, we randomly selected four turbines from each of the three wind farms

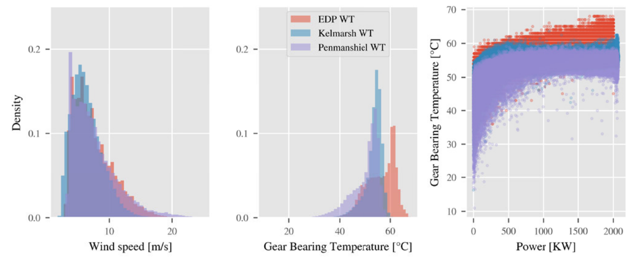

FIGURE 3. **Illustration of the statistical heterogeneity across the three considered wind farms. We used a randomly selected WT from each farm for comparison of the data distributions. Left: The distribution of the wind speed, an input feature of the NBM. Middle: Distribution of the target variable, the gear-bearing temperature. Right: A scatterplot of the power (input feature) in relation to the output feature (gear bearing temperature).**

to reduce data imbalance and computational cost for FL model training. SCADA data from the resulting 12 WTs in total was used to train a NBM with FL. From our dataset, we extracted 24-hour trailing windows as model input samples to ensure that the NBM can capture the temporal dynamics and operational conditions of the past 24 hours. That is, the task of the NBM is to predict the gear-bearing temperature at the end of the 24 hour window sample based on the respective historical data of the input features.

B. MODEL ARCHITECTURE AND TRAINING

The NBM trained in this study is an LSTM (Long Short-Term Memory) network, selected due to its suitability for time series data and its ability to capture temporal dependencies in the data, consisting of layers of LSTM units followed by fully connected layers. The NBM predicts the expected gear-bearing temperature at the end of the 24 hour input sample. The resulting LSTM model comprises two LSTM layers of sizes 16 and 64, respectively, and ReLU activation, followed by two fully connected layers of sizes 64 and 32 with ReLU activation. Hyperparameters of the LSTM network were optimized by searching for the best performance on a single randomly selected turbine from the Penmanshiel wind farm, whose data presented greater complexity than the turbines of other wind farms. We used random grid search and manually iteratively restricted the grid boundaries according to the insights from preliminary results (e.g. removing learning rates that are too low/high according to grid search) to get an increasingly improving search. The metric used for the selection was the MAE on the validation dataset and the search space was defined via standard parameters such as network depth, layer sizes, Adam optimizer parameters, learning rates and so on. The initial boundary of our search was taken excessively large and was successively refined to go from coarse to fine parameter tuning.

To assess the reduction of the training data accumulation time (cold start speed up), we train our NBM using increasing time intervals of training data. We will start by selecting a specific date as the start date. Then, for each start date, we will train multiple NBMs, each incorporating progressively more historical data by using different time ranges of training data. The time ranges commence from the selected start date and incrementally increase by one week, reaching a maximum of 12 weeks. This approach allows us to evaluate the impact of FL on reducing the time required to accumulate adequate data for model training. To account for seasonal variations and avoid bias towards any particular season, we select four different start dates spread evenly throughout the year: December 2016, March 2017, June 2017, and September 2017. For each of these start dates, we train models using all 12 different time ranges, enabling a comprehensive evaluation of the cold start speed up achieved by FL across different seasons. For each start date, a test set is given by the 4-week time window that follows the 12th week of training data as illustrated in Fig. 4.

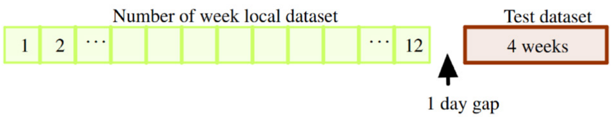

FIGURE 4. For each training time range from 1 to 12 weeks, we select the subsequent 4 weeks as the test set, beginning at the end of the 12-week training period. We introduce a one-day gap between the training and test sets to prevent data overlaps of the time windows. This is repeated for each starting date on the first of December 2016, March 2017, June 2017, and September 2017.

Each local dataset is split into 80% for training and 20% for validation. We note that in the case of only one or up to two weeks of data, this may result in a validation dataset that is not fully independent from the training dataset due to autocorrelation of environmental condition time series.

C. LEARNING STRATEGIES

For each training set described above (defined by a selected start date and time range), we trained normal behaviour models of gear bearing temperature according to three different learning strategies representing different types of collaboration:

- *Local learning*: Each wind turbine independently trains its own model using only its local data without any collaboration with other turbines.
- *Intra-farm learning*: Each wind farm utilizes `FedAvg` to train its own FL model, with no exchange of data or knowledge across different wind farms. Models are trained on similar clients, as WTs within a given wind farm typically exhibit similar and correlated SCADA data patterns. Employed in [10].
- *Inter-farm learning*: WTs of multiple farms participate in a single `FedAvg` learning process. The participating turbines involve different WT models, SCADA systems, and geographic locations, which results in significant data heterogeneity among the participating WT clients.

All models are trained by minimizing the mean squared error between the measured and the predicted gear bearing temperature using an Adam optimizer at a learning rate of 0.0008 and batch size of 256. Training is stopped once the validation loss (i.e., calculated either over the local WT's validation data for local training or as an average across all involved validation sets in FL) stops improving. The intra- and inter-farm learning strategies are illustrated in Fig. 1.

As outlined in Section II, globally trained models resulting from the `FedAvg` algorithm may in certain cases result in inferior performance compared to exclusively locally trained model. To this end, we additionally employ fine-tuning as a Personalized FL approach. For each FL strategy, we assessed the performance of the trained normal behaviour models with and without fine-tuning. Fine-tuning involves retraining the global model on each wind turbine's local training data after the FL training process. Fine-tuning is commonly used as a customization process in FL applications (e.g., [13]) and aims to retain knowledge embedded in the global model while further adjusting the model weights to local WT data, thereby possibly improving performance, as demonstrated in [10]. Moreover, it can be viewed as extra, post-training local updates, thereby not requiring any additional framework or coordination. We consider identical feature and label spaces. Other types of FL, such as vertical FL and federated transfer learning [46] are not considered in this study. We provide our implementation on GitHub.[1]

IV. RESULTS AND DISCUSSION

A. FL OUTPERFORMS LOCAL TRAINING IF TRAINING DATA ARE LIMITED

We compare the FL strategies by analyzing the quality of the NBMs trained for the gear-bearing temperatures of the twelve WTs from the three wind farms. We assess the average Mean Absolute Error (MAE) of the individual WTs when trained with the respective learning strategy and training data time range. This average is taken across all time ranges, start dates, and all twelve wind turbines. We will discuss both the impact of fine-tuning on the models' performances and compare the relative effectiveness of the different learning strategies. The results are given by the unweighted average of the MAE (predicted v. actual temperature) for each turbine (on the test set) using either the global model (non-fine-tuned results) or the specialized model for that turbine (fine-tuned models or local training).

As shown in Table 2, the overall best result is obtained by using intra-farm `FedAvg` with fine-tuning, followed by inter-farm learning with fine-tuning. Intra-farm `FedAvg` with fine-tuning demonstrates a significant improvement over local training, reducing the MAE by approximately 40%. While inter-farm `FedAvg` with fine-tuning also demonstrates improvement over local training, it falls slightly behind fine-tuned intra-farm learning.

a: FINE-TUNING

Across all strategies evaluated, fine-tuning always outperforms the performances of the global model. The results

[1] Code available at https://github.com/EnergyWeatherAI/FL-Wind-NBM

TABLE 2. Mean absolute errors (in °C) of the gear-bearing temperature NBMs trained with different learning strategies, averaged across all time ranges, start dates, and wind turbines.

	FedAvg		Local
Fine-tuned	Intra-farm	Inter-farm	
no	1.31	6.92	
yes	**0.87**	0.97	1.44

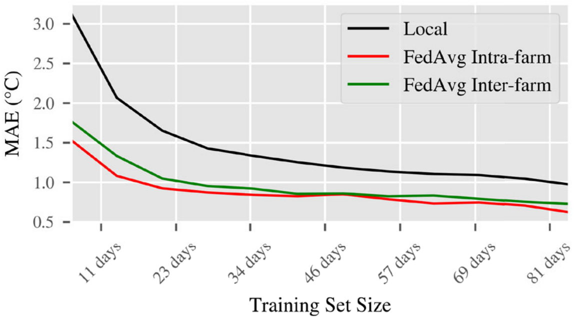

FIGURE 5. Mean absolute errors for fine-tuned `FedAvg` and local training. We show the error depending on the size of the training set (with evaluations in 1-week intervals). The losses are averaged across all start dates and wind turbines. The relative performance of the considered strategies remains consistent across varying training set sizes.

of Table 2 suggest that fine-tuning federated models combines both local and collaborative knowledge embedded in the global FL model (knowledge obtained from different turbines), even though non-fine-tuned models exhibit moderate to poor performances compared to local training. This highlights the importance of fine-tuning in consolidating collaborative learning gains and improving model performance.

b: LEARNING STRATEGIES

We obverse that after fine-tuning, the federated strategies sharply outperform local training. This is particularly so when little training data (less than a few months, as in this study) is available as shown in Fig. 5.

However, while intra-farm learning outperforms local training by approximately 9% without fine-tuning, inter-farm learning without fine-tuning performs very poorly (the MAE is about fivefold higher than local training according to Table 2). This discrepancy may be caused by the significant statistical heterogeneity between the wind farm SCADA variables.

The relative performance between local training and fine-tuned intra- and inter-farm `FedAvg` remains consistent across all time ranges considered, as shown in Fig. 5. FL consistently exhibits a strong improvement compared to local training, even with 12 weeks of training data.

c: SEASONALITY

Our above results were averaged over different start dates. We investigated whether the results depend on the time of year and found they remain largely consistent across different seasons, as shown in Fig. 6. In all seasons, intra- and interfarm FL enable more accurate NBMs than local training.

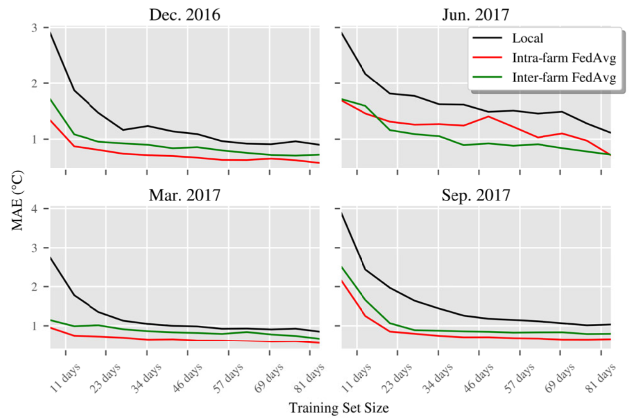

FIGURE 6. Mean absolute errors for each start date, with fine-tuned `FedAvg` and local training. These results are averaged across all turbines.

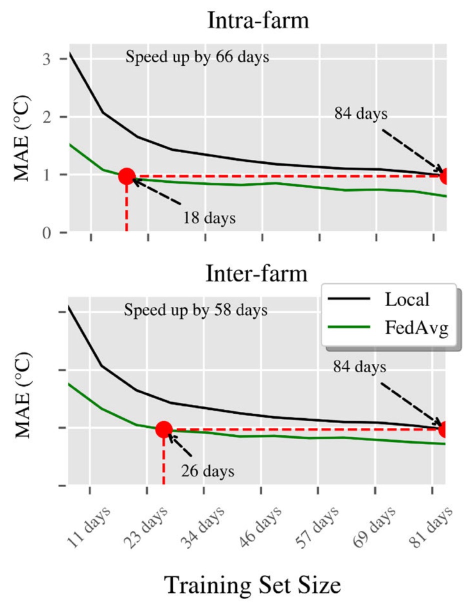

FIGURE 7. The mean absolute error for the intra-farm (top) and the inter-farm (bottom) setting. `FedAvg` with fine-tuning reduces the time needed to accumulate the amount of historical training data required to achieve performances comparable to local training by 66 days in intra-farm learning and by 58 days in inter-farm learning, when averaged across all WTs and start dates.

Fig. 6 shows that the evolution of the mean absolute errors for fine-tuned `FedAvg` and local training by starting date is also consistent across all seasons with the exception of the starting date in June 2017. A possible explanation for this behaviour is a seasonality-based data distribution shift between the training and test set for the Kelmarsh wind farm, which is discussed further in Section IV-C.

TABLE 3. **Mean absolute errors (in °C) of the gear-bearing temperature NBMs trained with different learning strategies from each wind farm, averaged across all time ranges, start dates, and wind turbines of the respective wind farm.**

		FedAvg		Local
	Fine-tuned	Intra-farm	Inter-farm	
EDP	no	1.74	2.33	
	yes	**1.35**	1.44	2.10
Kelmarsh	no	1.18	1.81	
	yes	0.87	**0.65**	1.23
Penmanshiel	no	1.02	16.65	
	yes	**0.40**	0.82	1.00

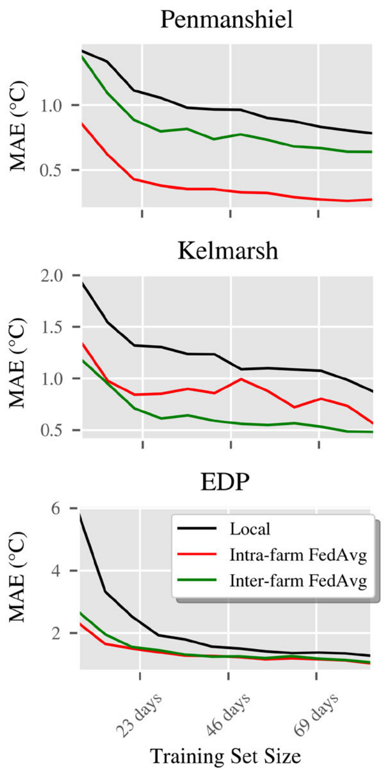

FIGURE 8. **Evolution of the mean absolute errors (in °C) for fine-tuned** `FedAvg` **and local training across varying training set sizes. These losses are the averaged results across all start dates and turbines within each wind farm. Fine-tuned federated learning consistently outperforms local training across wind farms.**

d: FL STRONGLY REDUCES THE TRAINING DATA COLLECTION TIME

The *cold start speed up* refers to the reduction of time needed to accumulate the amount of historical training data required to achieve performances comparable to local training when employing FL. As shown in Fig. 7, `FedAvg` with fine-tuning reduces the training data collection time by approximately eight to nine weeks out of the twelve weeks under consideration in intra-farm and inter-farm FL. For example in Fig. 7, to achieve the equivalent accuracy of the best performing gear bearing temperature NBM trained with local data, intra-farm `FedAvg` with fine-tuning

TABLE 4. **Mean absolute error (in °C) for each wind farm, start date and learning strategy, averaged across all time ranges, and wind turbines within each wind farm.**

		EDP			Kelmarsh			Penmanshiel		
		FedAvg		Local	FedAvg		Local	FedAvg		Local
	Fine-tuned	Intra-farm	Inter-farm		Intra-farm	Inter-farm		Intra-farm	Inter-farm	
2016-12-01	no	1.79	2.49		0.81	1.82		0.95	17.80	
	yes	**1.48**	1.53	2.24	**0.42**	0.44	0.72	**0.34**	0.77	0.93
2017-03-01	no	1.67	2.47		0.78	1.93		0.94	16.32	
	yes	**1.39**	1.51	2.23	**0.35**	0.40	0.59	**0.25**	0.68	0.83
2017-06-01	no	1.77	2.02		2.08	1.85		1.06	13.76	
	yes	**1.30**	1.32	1.88	1.89	**0.96**	2.08	**0.48**	0.85	1.09
2017-09-01	no	1.74	2.34		1.06	1.63		1.12	18.76	
	yes	**1.24**	1.38	2.05	0.83	**0.80**	1.53	**0.52**	0.97	1.16

requires only 18 days compared to the 84 days of training data required using local data only. This speed up allows for earlier deployment of accurate NBMs, enabling earlier fault detection and thereby reducing the risk of undetected incipient faults.

B. RESULTS BY WIND FARM

The statistical heterogeneity of the datasets of different clients can present a significant challenge to collaborative learning. Our results suggest that increasing the number of WTs involved in the training does not necessarily lead to improved collaborative learning outcomes, even after fine tuning. In particular, if the data distributions vary among the different wind turbines, the performance of collaborative learning across WTs of different wind farms (inter-farm learning) can be worse than that of collaborative learning within a given wind farm (intra-farm learning).

We also assessed the performance of FL for NBM training in the context of the individual wind farms. We averaged the accuracies of the NBMs across the four turbines involved in the FL training from each wind farm, as shown in Table 3.

`FedAvg` with fine-tuning consistently outperformed the other learning methods at each wind farm. Intra-farm FL with fine-tuning significantly surpasses the accuracy of inter-farm learning with fine-tuning at the Penmanshiel wind farm, as shown in Table 3 and Fig. 8. The comparatively poor performance of inter-farm FL is likely related to the low fraction of SCADA data from the Penmanshiel wind farm, as its small data contribution to the inter-farm FL training results in a small contribution to the inter-farm FL model. The non-fine-tuned inter-farm MAE is substantially higher (16.65 °C), indicating that the NBM primarily learns from the other two wind farms in the case of inter-farm learning. Conversely, the EDP wind farm exhibits minimal disparity between intra- and inter-farm learning. This is likely because a significant portion of the global model's influence stems from EDP's data, which accounts for roughly 50% of the total data across the various wind farms (as depicted in Fig. 2). Thus, the global model's performances on EDP's wind turbines are less affected by the heterogeneity introduced by other wind farms' data.

For each wind farm, the implementation of FL strategies results in a significant cold start speed up, with saved time

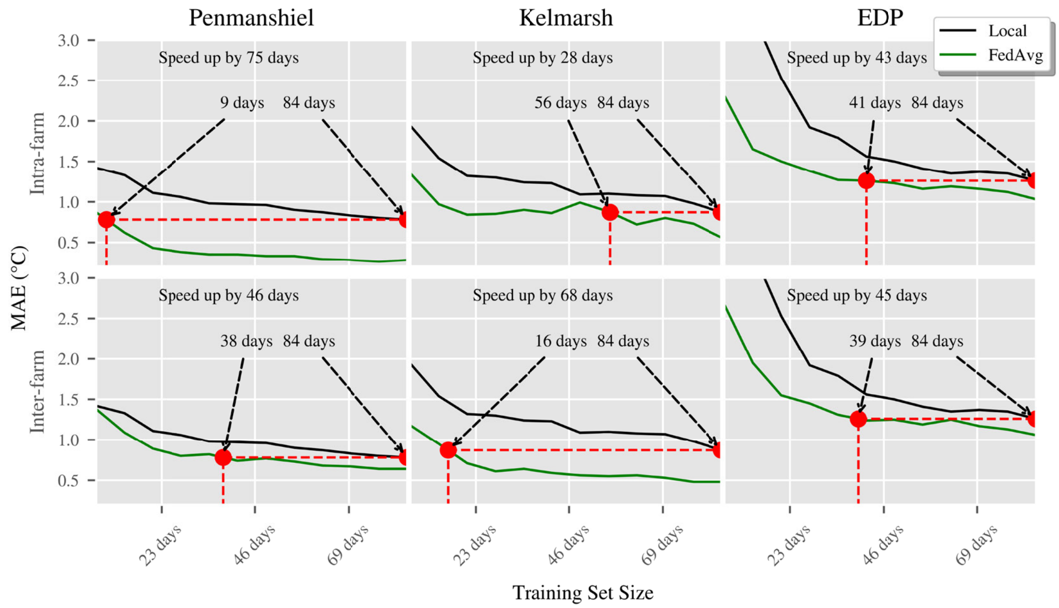

FIGURE 9. Cold start speed up for fine-tuned `FedAvg`. These MAEs are averaged across the different start dates and turbines within each wind farm. For each farm, `FedAvg` consistently provides a cold start speed by achieving the best considered performance of local training several days earlier.

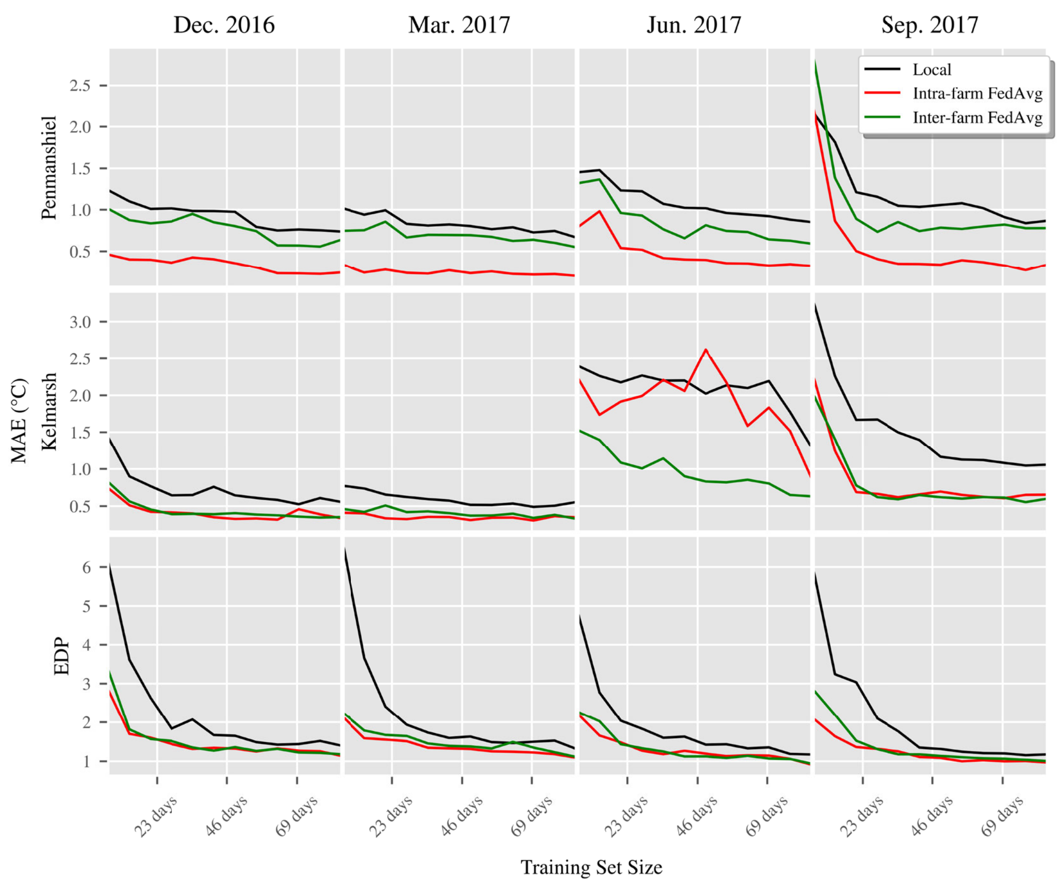

FIGURE 10. Mean absolute error for all four start dates and wind farms to illustrate the impact of seasonality on the results. These results are averaged across the turbines within a given wind farm.

ranging from four to more than ten weeks. This phenomenon is illustrated in Fig. 9, where the cold start speed up for fine-tuned `FedAvg` is depicted. The most substantial reduction in the time required to collect training data occurs with intra-farm learning with fine-tuning in the Penmanshiel wind farm.

We considered four different start dates, one in each season, to investigate how seasonality impacts the accuracy of NBMs trained with different FL strategies at each wind farm. This resulted in twelve combinations of start date and wind farm, as shown in Fig. 10 and Table 4.

Intra-farm learning with fine-tuning emerged as the best-performing strategy in most cases (10 out of 12), and inter-farm learning with fine-tuning in the remaining two cases which pertain to the Kelmarsh wind farm. In one of the two cases, the difference between intra- and inter-farm is insignificant, in the other case, we explain in more details what might have happened in Section IV-C. We also investigated how the different learning strategies perform for individual wind turbines and found that the previous results from Sections IV-A and IV-B are confirmed also at the WT level. Federated learning enables more accurate normal behaviour models than local training with limited data, and a reduction in the amount of time needed to collect the required model training data. Fine-tuning provided more accurate NBMs in all cases. Moreover, intra-farm FL tended to provide more accurate NBMs than inter-farm FL.

C. GROUND TRUTH EXAMPLES AND DISCUSSION

We examine the model prediction compared to the ground truth of the gear-bearing temperature for one selected turbine from each wind farm. We restrict ourselves here to models trained on three weeks of training data in December for Fig. 11 and June for Fig. 12, using the corresponding four-week test dataset for evaluation. This analysis provides insights into the performance of various learning methods, enabling a qualitative assessment of FL. Notably, predictions from local training (no collaboration) and fine-tuned `FedAvg` closely align with the ground truth. However, non-fine-tuned `FedAvg` exhibits inferior performance, especially in scenarios characterized by significant data heterogeneity among clients, such as inter-farm learning.

Fig. 11 illustrates the performance of the local, intra- and interfarm learning strategies on the respective test sets. A single WT has been picked randomly from each wind farm to this end. One important observation from Fig. 11 is the widening error of non-fine-tuned inter-farm `FedAvg` as the data proportion decreases. Specifically, the red curve representing non-fine-tuned inter-farm `FedAvg` for Penmanshiel (the wind farm with the smallest data proportion) shows a pronounced deviation from ground truth and is closer to the temperature ranges observed in the Kelmarsh and EDP wind farms. This disparity arises from the observable heterogeneity, with Penmanshiel exhibiting a temperature range of 25 °C to 35 °C, contrasting with the 40 °C to 60 °C range observed in Kelmarsh and EDP wind farms. `FedAvg` attributes a larger weight to clients with more data, resulting in Penmanshiel's turbine contributing significantly less than those of Kelmarsh and EDP wind farms (see Fig. 2).

Non-fine-tuned intra-farm learning outperforms non-fine-tuned inter-farm learning in this case study. This indicates that the different wind farms involved in the collaborative FL process end up competing to achieve different learning objectives rather than collaborate. However, this disparity is no longer visible after fine-tuning and opens the question of whether there is a retention of collaborative knowledge.

While these observations hold across most start dates and time ranges, a different behaviour emerges in situations where the local training fails to fit the test set, as shown in Fig. 12.

In Fig. 12, the overall observations remain consistent, except for the Kelmarsh wind farm during the testing period between September 11 and September 17. During this period, local training fails to effectively fit the ground truth, likely indicating data points lying outside the range of the model's training data, leading to poor generalization. Upon examining the data distribution of various features, we find no indications of anomalous behaviour during that period. Furthermore, we observe wind speed and ambient temperature value ranges slightly higher and lower, respectively, in the affected test set compared to the training set. Such variations are expected when comparing weather conditions between June and September. A distribution shift being a possible cause is further supported by the observation that during this period, inter-farm learning, both with and without fine-tuning, outperforms intra-farm learning (see also Fig. 6). This suggests that inter-farm learning may benefit from insights gained from other farms, enabling it to better adapt to locally unseen data. However, in practical scenarios, our primary concern is not whether a model trained on data from June will perform well on test data in September. Instead, our focus would lie on ensuring that the model performs effectively in the weeks that follow the training set. We could continuously retrain the model with new data as it becomes available, thereby mitigating any seasonality shift and increasing the likelihood of the model performing well on near-future data in a continuous learning setting.

D. OTHER ALGORITHMS

In addition to `FedAvg`, we implemented an alternative federated learning algorithm named `FedProx` [47]. `FedProx` follows a similar learning process as `FedAvg` but incorporates a regularization term in the loss function during local training, which measures the discrepancy between the current global model and the updated local model of the clients. Our preliminary results indicated that `FedProx` performs comparably but slightly worse than `FedAvg` for both inter-farm and intra-farm learning, for which we did

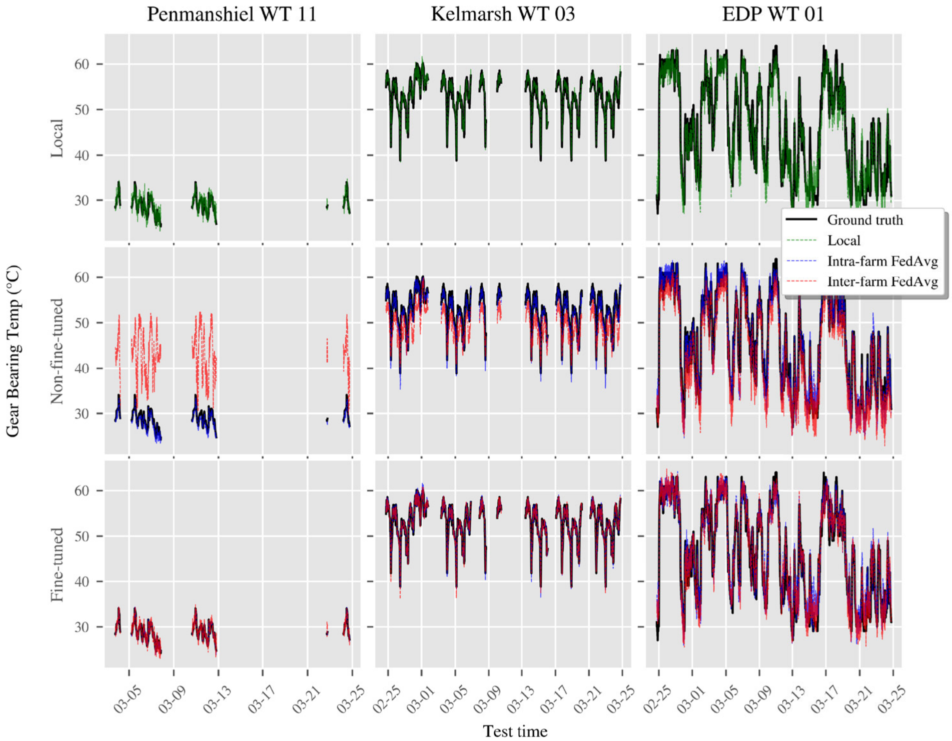

FIGURE 11. For each wind farm, we compare ground truth (in black) to the predicted gear bearing temperature (in °C) using local learning (first row), non-fine-tuned intra and inter-farm `FedAvg` (second row), and their fine-tuned version (third row). A single turbine has been picked from each wind farm. Results depicted for the models trained on three weeks of training data in December.

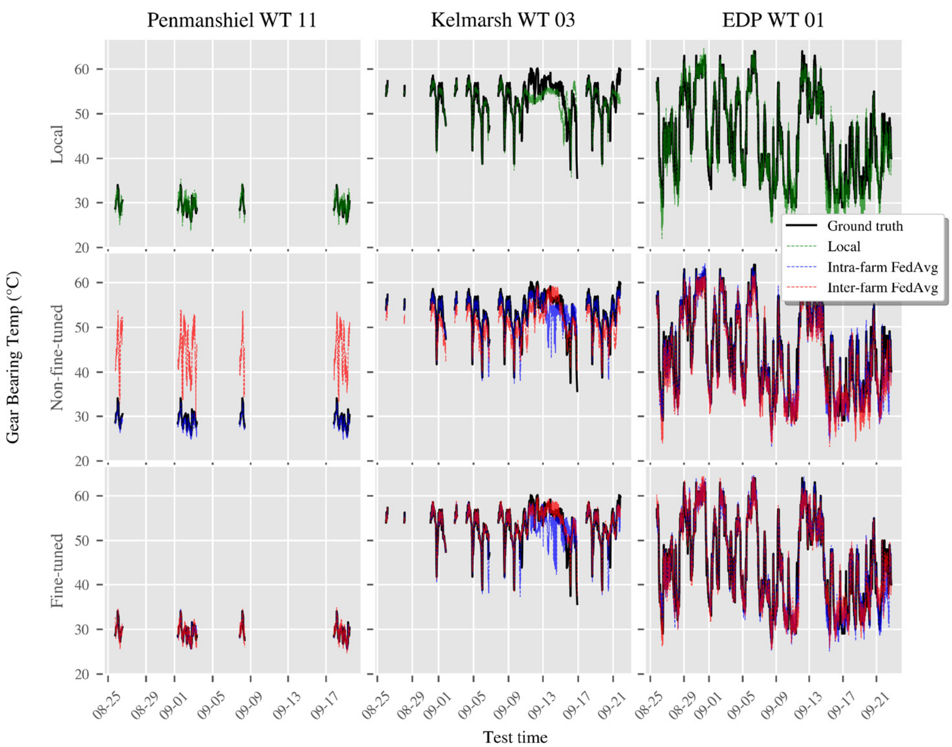

FIGURE 12. Actual versus predicted gear bearing temperatures using local learning (first row), intra and inter-farm `FedAvg` with and without fine-tuning (third and second, respectively). This visual evaluation is done on the test dataset and the models are trained on three weeks of data in June 2017.

not further consider it for further investigation. This may be attributed to the fact that `FedProx` slows down local training to retain more knowledge from the global FL model. Although `FedProx` was initially developed to handle statistical heterogeneity, it mainly addresses covariate shift, i.e., different distributions of the input variables. However, as visible in Fig. 3, our input variables, such as the wind speed, do not exhibit such distribution shifts. Instead, we are faced with a concept shift, where the task (input-output relationship) differs across the different wind farms. Such a shift cannot be resolved by `FedProx`.

V. CONCLUSION

Our study demonstrated the effectiveness of privacy-preserving collaborative learning of wind turbine normal behaviour models for condition monitoring. We demonstrated that federated learning enables a reduction of the training data accumulation time, allowing for an earlier detection of developing faults, compared to local training. We also found that having more collaborators is not necessarily better in collaborative learning. In the presence of high statistical heterogeneity, i.e., significant differences between the data distributions of the involved wind turbines, the performance of federated learning across different wind farms (*inter-farm* learning) is worse than that of federated learning within a given wind farm (*intra-farm* learning).

We assessed two distinct collaborative learning approaches: inter-farm learning, which involves collaboration across turbines from different wind farms, and intra-farm learning, which restricts collaboration to turbines in the same wind farm. Our analysis shows that high levels of statistical heterogeneity present significant challenges to collaborative learning. The accuracy of NBMs trained in intra-farm learning surpassed that of inter-farm learning in most situations, underscoring the adverse impacts of heterogeneity on collaborative learning. We demonstrated fine-tuning as a successful approach to address FL model training in view of significant statistical heterogeneity.

These findings highlight the trade-offs involved in deploying federated learning for industrial condition monitoring. While collaborative training can enable accelerated learning and improved generalization when data distributions are sufficiently aligned, the presence of heterogeneity introduces challenges that can degrade the model performances. These insights emphasize the importance of considering the heterogeneity for FL deployment.

This study is subject to several limitations. First, we focused our investigation on an application of a specific NBM. Further experiments with varying model constellations, e.g., different input and target variables, more wind farms, as well as more variations in the data imbalance and extent of heterogeneity among the farms are required to validate the generalizability and robustness of our results. Scaling FL to larger wind farms may also introduce additional challenges and considerations, for instance regarding efficiency and the practical deployment of FL. In the latter, there are furthermore numerous choices open to practitioners regarding data efficiency, possible privacy enhancements, and the structure of FL (e.g., centralized server or decentralized set-up), for which a suitable decision may depend on various factors. In our study, we considered a standard FL setting with a central server. In practice, this requires the set-up of a centralized server handling the coordination, organization, and communication overhead. Certain considerations might need to taken particularly when considering more complex models or scaling FL to larger fleets, as this would follow a rise in overhead and communication costs. We refer to [13] for a discussion of factors and challenges regarding the framework flexibility and choices, computational efficiency, communication overhead, and scalability of FL.

Lastly, our work is restricted to homogenous feature spaces and largely similar SCADA systems. An investigation into federated transfer learning [46] may prove insightful on how to adapt to changing feature spaces across wind farms.

There are several further directions of future research. Firstly, our model selection and hyperparameter tuning processes were conducted on a full dataset from a single turbine, potentially diverging from real-world conditions where historical data accumulation occurs incrementally and disregarding the contributions of the other turbines. Addressing this challenge entails the development of automated and adaptive model selection methods capable of accommodating evolving data volumes and complexities [48]. Moreover, integrating FL techniques for hyperparameter tuning [49] could enhance efficiency and scalability in FL settings. Furthermore, we highlight the potential of continuous learning strategies [50] to reduce communication costs and training time in FL by enabling incremental model updates instead of periodic full re-training. Finally, we considered identical feature and label spaces of all client wind turbines but did not consider other types of FL in this study, such as federated transfer learning [46]. They may be the subject of future research in wind energy applications.

In conclusion, our study demonstrated the potential and challenges of collaborative learning in wind turbine condition monitoring through FL. By advancing our understanding of effective collaboration strategies and addressing challenges such as statistical heterogeneity and model adaptation, we move closer to realizing the full potential of FL for enhancing the reliability and efficiency of wind farms and other renewable energy systems.

REFERENCES

[1] *Clean Energy Technology Observatory: Wind Energy in the European Union 2023 Status Report on Technology Development, Trends, Value Chains and Markets*, document KJ-NA-31-678-EN-N, 2023.

[2] GWEC. (2023). *Global Wind Report 2023*. [Online]. Available: https://gwec.net/wp-content/uploads/2023/03/GWR-2023_interactive.pdf

[3] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel, "Deep learning for anomaly detection: A review," *ACM Comput. Surv.*, vol. 54, pp. 1–38, Mar. 2021.

[4] J. L. C. Hoffmann, L. P. Horstmann, M. M. Lucena, G. M. de Araujo, A. A. Fröhlich, and M. H. N. Nishioka, "Anomaly detection on wind turbines based on a deep learning analysis of vibration signals," *Appl. Artif. Intell.*, vol. 35, no. 12, pp. 893–913, Oct. 2021.

[5] C. Zhang and T. Yang, "Anomaly detection for wind turbines using long short-term memory-based variational autoencoder Wasserstein generation adversarial network under semi-supervised training," *Energies*, vol. 16, no. 19, p. 7008, Oct. 2023.
[6] B. Altice, E. Nazario, M. Davis, M. Shekaramiz, T. K. Moon, and M. A. S. Masoum, "Anomaly detection on small wind turbine blades using deep learning algorithms," *Energies*, vol. 17, no. 5, p. 982, Feb. 2024.
[7] J. Maron, D. Anagnostos, B. Brodbeck, and A. Meyer, "Artificial intelligence-based condition monitoring and predictive maintenance framework for wind turbines," *J. Phys., Conf. Ser.*, vol. 2151, no. 1, Jan. 2022, Art. no. 012007.
[8] S. Jonas, D. Anagnostos, B. Brodbeck, and A. Meyer, "Vibration fault detection in wind turbines based on normal behaviour models without feature engineering," *Energies*, vol. 16, no. 4, p. 1760, Feb. 2023.
[9] P. Bangalore, S. Letzgus, D. Karlsson, and M. Patriksson, "An artificial neural network-based condition monitoring method for wind turbines, with application to the monitoring of the gearbox," *Wind Energy*, vol. 20, no. 8, pp. 1421–1438, Aug. 2017.
[10] L. Jenkel, S. Jonas, and A. Meyer, "Privacy-preserving fleet-wide learning of wind turbine conditions with federated learning," *Energies*, vol. 16, no. 17, p. 6377, Sep. 2023.
[11] A. Kusiak, "Renewables: Share data on wind energy," *Nature*, vol. 529, no. 7584, pp. 19–21, Jan. 2016.
[12] P. Voigt and A. Von dem Bussche, "The EU general data protection regulation (GDPR)," in *A Practical Guide* (GDPR), vol. 10. Cham, Switzerland: Springer, 2017, p. 5555.
[13] A. Grataloup, S. Jonas, and A. Meyer, "A review of federated learning in renewable energy applications: Potential, challenges, and future directions," *Energy AI*, vol. 17, Sep. 2024, Art. no. 100375.
[14] L. Li, Y. Fan, M. Tse, and K.-Y. Lin, "A review of applications in federated learning," *Comput. Ind. Eng.*, vol. 149, Nov. 2020, Art. no. 106854.
[15] Y. Li, R. Wang, Y. Li, M. Zhang, and C. Long, "Wind power forecasting considering data privacy protection: A federated deep reinforcement learning approach," *Appl. Energy*, vol. 329, Jan. 2023, Art. no. 120291.
[16] A. Ahmadi, M. Talaei, M. Sadipour, A. M. Amani, and M. Jalili, "Deep federated learning-based privacy-preserving wind power forecasting," *IEEE Access*, vol. 11, pp. 39521–39530, 2023.
[17] H. Moayyed, A. Moradzadeh, B. Mohammadi-Ivatloo, A. P. Aguiar, and R. Ghorbani, "A cyber-secure generalized supermodel for wind power forecasting based on deep federated learning and image processing," *Energy Convers. Manage.*, vol. 267, Sep. 2022, Art. no. 115852.
[18] Y. Wang, W. Zhang, Q. Guo, and Y. Wu, "A privacy-preserving wind speed prediction method based on federated deep learning," in *Proc. 4th Int. Conf. Power Energy Technol. (ICPET)*, Jul. 2022, pp. 638–643.
[19] L. Wang, W. Fan, G. Jiang, and P. Xie, "An efficient federated transfer learning framework for collaborative monitoring of wind turbines in IoE-enabled wind farms," *Energy*, vol. 284, Dec. 2023, Art. no. 128518.
[20] X. Cheng, F. Shi, Y. Liu, X. Liu, and L. Huang, "Wind turbine blade icing detection: A federated learning approach," *Energy*, vol. 254, Sep. 2022, Art. no. 124441.
[21] X. Cheng, F. Shi, Y. Liu, J. Zhou, X. Liu, and L. Huang, "A class-imbalanced heterogeneous federated learning model for detecting icing on wind turbine blades," *IEEE Trans. Ind. Informat.*, vol. 18, no. 12, pp. 8487–8497, Dec. 2022.
[22] D. Zhang, W. Tian, Y. Yin, X. Liu, X. Cheng, and F. Shi, "Human knowledge-based compressed federated learning model for wind turbine blade icing detection," in *Proc. Int. Conf. High Perform. Big Data Intell. Syst. (HDIS)*, Dec. 2022, pp. 277–281.
[23] X. Cheng, W. Tian, F. Shi, M. Zhao, S. Chen, and H. Wang, "A blockchain-empowered cluster-based federated learning model for blade icing estimation on IoT-enabled wind turbine," *IEEE Trans. Ind. Informat.*, vol. 18, no. 12, pp. 9184–9195, Dec. 2022.
[24] G. Jiang, W. Fan, W. Li, L. Wang, Q. He, P. Xie, and X. Li, "DeepFedWT: A federated deep learning framework for fault detection of wind turbines," *Measurement*, vol. 199, Aug. 2022, Art. no. 111529.
[25] W. Yang and G. Yu, "Federated multi-model transfer learning-based fault diagnosis with peer-to-peer network for wind turbine cluster," *Machines*, vol. 10, no. 11, p. 972, Oct. 2022.
[26] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. 20th Int. Conf. Artif. Intell. Statist.*, vol. 54, Apr. 2017, pp. 1273–1282.
[27] T.-M. H. Hsu, H. Qi, and M. Brown, "Measuring the effects of non-identical data distribution for federated visual classification," 2019, *arXiv:1909.06335*.
[28] Q. Li, Y. Diao, Q. Chen, and B. He, "Federated learning on non-IID data silos: An experimental study," in *Proc. IEEE 38th Int. Conf. Data Eng. (ICDE)*, May 2022, pp. 965–978.
[29] Y. Deng, M. M. Kamani, and M. Mahdavi, "Adaptive personalized federated learning," 2020, *arXiv:2003.13461*.
[30] T. Yu, E. Bagdasaryan, and V. Shmatikov, "Salvaging federated learning by local adaptation," 2020, *arXiv:2002.04758*.
[31] C. Briggs, Z. Fan, and P. Andras, "Federated learning with hierarchical clustering of local updates to improve training on non-IID data," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2020, pp. 1–9.
[32] G. Long, M. Xie, T. Shen, T. Zhou, X. Wang, and J. Jiang, "Multi-center federated learning: Clients clustering for better personalization," *World Wide Web*, vol. 26, no. 1, pp. 481–500, Jun. 2022.
[33] L. Tang, H. Xie, X. Wang, and Z. Bie, "Privacy-preserving knowledge sharing for few-shot building energy prediction: A federated learning approach," *Appl. Energy*, vol. 337, May 2023, Art. no. 120860.
[34] J. C. Bezdek, "Objective function clustering," in *Pattern Recognition With Fuzzy Objective Function Algorithms*. Boston, MA, USA: Springer, 1981, pp. 43–93.
[35] M. G. Arivazhagan, V. Aggarwal, A. K. Singh, and S. Choudhary, "Federated learning with personalization layers," 2019, *arXiv:1912.00818*.
[36] A. Fallah, A. Mokhtari, and A. Ozdaglar, "Personalized federated learning with theoretical guarantees: A model-agnostic meta-learning approach," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., 2020, pp. 3557–3568.
[37] L. Collins, H. Hassani, A. Mokhtari, and S. Shakkottai, "FedAvg with fine tuning: Local updates lead to representation learning," in *Proc. Adv. Neural Inf. Process. Syst.*, 2022, pp. 10572–10586.
[38] A. Z. Tan, H. Yu, L. Cui, and Q. Yang, "Towards personalized federated learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 12, pp. 9587–9603, Dec. 2023.
[39] V. Kulkarni, M. Kulkarni, and A. Pant, "Survey of personalization techniques for federated learning," in *Proc. 4th World Conf. Smart Trends Syst., Secur. Sustainability (WorldS4)*, Jul. 2020, pp. 794–797.
[40] E. Serral, P. Valderas, and V. Pelechano, "Improving the cold-start problem in user task automation by using models at runtime," in *Information Systems Development*, 2011, pp. 671–683.
[41] C. Plumley, "Penmanshiel wind farm data," *Dataset*, Feb. 2022, doi: 10.5281/zenodo.5946808.
[42] C. Plumley, "Kelmarsh wind farm data," *Dataset*, Feb. 2022, doi: 10.5281/zenodo.5841834.
[43] EDP. (2016). *Wind Turbine Scada Signals 2016*. Dataset. [Online]. Available: https://www.edp.com/en/innovation/open-data/wind-turbine-scada-signals-2016
[44] EDP. (2017). *Wind Turbine Scada Signals 2017*. Dataset. [Online]. Available: https://www.edp.com/en/innovation/open-data/wind-turbine-scada-signals-2017
[45] M. Schlechtingen, I. F. Santos, and S. Achiche, "Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 1: System description," *Appl. Soft Comput.*, vol. 13, no. 1, pp. 259–270, Jan. 2013.
[46] Q. Yang, Y. Liu, Y. Cheng, Y. Kang, T. Chen, and H. Yu, *Federated Transfer Learning*. Switzerland: Springer, 2020, pp. 83–93.
[47] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," in *Proc. Mach. Learn. Syst.*, 2020, pp. 429–450.
[48] B. Celik and J. Vanschoren, "Adaptation strategies for automated machine learning on evolving data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 9, pp. 3067–3078, Sep. 2021.
[49] M. Khodak, R. Tu, T. Li, L. Li, M.-F. Balcan, V. Smith, and A. Talwalkar, "Federated hyperparameter tuning: Challenges, baselines, and connections to weight-sharing," in *Proc. Adv. Neural Inf. Process. Syst.*, 2021, pp. 19184–19197. [Online]. Available: https://proceedings.neurips.cc/paper/2021/hash/a0205b87490c847182672e8d371e9948-Abstract.html

[50] L. Wang, X. Zhang, H. Su, and J. Zhu, "A comprehensive survey of continual learning: Theory, method and application," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 8, pp. 5362–5383, Aug. 2024.

ALBIN GRATALOUP received the B.S. and M.S. degrees in mathematics and application from the ENS of Lyon, and the Ph.D. degree in mathematics from the University of Montpellier, France. He is currently a Postdoctoral Researcher with Bern University of Applied Sciences, where his main focus lies in federated learning and personalization methods applied for renewable and sustainable energies, and medical (in particular mental health) applications.

STEFAN JONAS received the B.S. and M.S. degrees in computer science from the University of Bern, Switzerland. He is currently pursuing the Ph.D. degree in informatics with Bern University of Applied Sciences, Switzerland, and Università della Svizzera italiana, Switzerland. His primary research interests include applied machine learning, particularly anomaly detection and transfer learning with applications to data from industrial fleets.

ANGELA MEYER received the B.S. degree in physics from Imperial College London, the M.S. degree in mathematics from the University of Cambridge, U.K., and the Ph.D. degree in atmospheric physics from ETH Zürich, Switzerland. After postdoctoral research in the EU project DNICast on short-term solar irradiance forecasting, she was a Research Scientist and a Project Leader in the fields of applied machine learning and data-driven condition monitoring at the research and development centers of Hexagon AB and Siemens Smart Infrastructures. She is currently an Assistant Professor with BFH and TU Delft. Her research aims at developing intelligent decision support systems to increase the resilience and sustainability of industrial and energy systems with sensor-driven and machine learning approaches. Her research team is supported by the European Commission, the Swiss Innovation Agency Innosuisse, and the Swiss National Science Foundation.