Frontiers in Computational Science and Engineering
Vol. 1, No. 3, Jul - Sep 2025, pp. 1-8
ISSN: 3107-3654 (O), DOI: https://doi.org/10.17051/FCSE/01.03.01

Uncertainty-Aware Federated Learning Model for Distributed Scientific Computing

Julian L. Webber

Associate Professor, Department of Electronics and Communication Engineering , Kuwait College of Science and Technology (KCST), Doha Area, 7th Ring Road, Kuwait, Email: j.webber@kcst.edu.kw, 0000-0001-7796-2898

Article Info

Article history:

Received : 05.07.2025
Revised : 06.08.2025
Accepted : 07.09.2025

Keywords:

Federated learning,
uncertainty quantification,
distributed scientific computing,
Bayesian learning,
probabilistic modeling,
non-IID data,
model calibration,
predictive uncertainty,
edge computing,
machine learning in
scientific applications.

ABSTRACT

Federated learning has become an enticing paradigm of distributed scientific computing by allowing collaborative training of models without sharing central information. Nevertheless, the majority of existing methods emphasize mainly on predictive performance and fail to consider the essential factor of uncertainty estimation that is vital in good decision-making in the sciences. This ambiguity constrains the reliability of models that are used in data sensitive settings with high complexity. To overcome this difficulty, this paper will present an uncertainty-conscious federated learning model that incorporates the use of probabilistic reasoning in the distributed training. The framework proposed integrates the principles of Bayesian learning to include epistemic and aleatoric uncertainty to take stronger and interpretable predictions among heterogeneous clients. The model can be used effectively when the data is non-IID, which is a usual situation in real-life scientific computing problems. To maintain uncertainty information when updating the world, a modified strategy of aggregation is presented. Core metrics like the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and accuracy, are used to measure the performance of the proposed approach, and a comprehensive evaluation of the predictive ability is made. Experimental evidence shows that the proposed model is not only competitive with respect to the traditional federated learning approaches but also yields better reliability with a well-calibrated uncertainty estimate. These results underscore its possible application in improving the decision-making in distributed scientific setups.

1. INTRODUCTION

Distributed scientific computing has emerged as a key to addressing complex, data intensive problems in the health, industry, and autonomous systems domains. These systems have decentralized sources of data and heterogeneous infrastructures and rigid privacy requirements where a centralized data processing becomes impractical. One of the biggest issues is to work with non-IID data distributions and provide effective communication between distributed nodes. Federated learning (FL) has been proposed as a powerful solution due to the ability to train models collaboratively without sharing raw data and therefore keeping privacy intact and increasing scalability [1], [2], [5], [11].

New research has been directed towards enhancing aggregation techniques, effectiveness in communication and flexibility in a heterogeneous setting. The methods of posterior averaging [1], dynamic aggregation [2], adaptive aggregation [7], and robust optimization [9] have greatly contributed to the improvement of FL. In addition, FL has been used effectively in various areas such as the diagnostics of healthcare [3], mobile systems [5], industrial internet of things [12] and robotic systems [11]. The issues related to system-level in distributed settings have also been proposed using hybrid learning structures and optimization-driven aggregation schemes, as well [4], [6], [13].

Although these developments have been made, one of the main drawbacks of current FL strategies is the absence of uncertainty awareness. Majority of the models focus on predictive accuracy with no regard to reliability and confidence of predictions. This is especially true in the field of scientific computing where a decision can be made based on the reliability of the model results. Despite the fact that methods of personalization and representation learning have been developed in

non-IID contexts [8], these have failed to explicitly use uncertainty estimation in learning.

To fill this gap, this paper introduces an uncertainty-conscious federated learning framework to incorporate probabilistic modeling within the distributed training framework. The suggested solution incorporates data uncertainty and model uncertainty, which allows making reliable predictions in the case of heterogeneous customers. Core performance measures, including accuracy, Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), are used to evaluate the model, the performance of the predictive models is thoroughly evaluated.

This work contributes to the development of federated learning in distributed scientific computing systems in three main ways. First, it presents a new uncertainty-conscious federated learning framework, which is specially tailored to function under a heterogeneous and non-IID data regime, overcoming key issues in decentralized systems. Second, the suggested methodology incorporates probabilistic methods to both quantify effectively both model and data uncertainty, and increase the reliability, robustness and interpretability of predictions. Lastly, the framework is widely tested against the main performance measures so as to accuracy, Root Mean Square Error (RMSE), and Mean Absolute Error (MAE), showing that the framework performs better and offers better predictive accuracy in comparison to the current federated learning procedures.

2. RELATED WORK

Federated learning (FL) has become popular as a promising paradigm of distributed scientific computing, allowing jointly training models without the need to distribute data centrally. Initial research like Federated learning, given mobile keyboards prediction proved the viability of FL when operating in large-scale, decentralised settings, and later research expanded its usage to other areas, including healthcare, robotics, and industrial systems. As an example, Federated learning in smart healthcare: A survey outlines the increasing relevance of FL to privacy-related medical tasks and Federated learning in robotic and autonomous systems discusses its use in autonomous and decentralized decision-making. The recent developments have centered on how to enhance aggregation and efficiency of the system by means of posterior averaging [1], dynamic aggregation [2], adaptive aggregation [7] and robust aggregation techniques [9], and even hybrid centralized -federated architecture [4]. The approaches deal with the issues of heterogeneity, communication efficiency, and scalability of distributed environments.

Parallel, uncertainty modeling has been extensively researched in machine learning to enhance the reliability of predictions. Models like Bayesian learning and ensembles are widely applied to represent both epistemic and aleatoric uncertainty to increase the model interpretability and strength. Nonetheless, the adoption of uncertainty estimation in federated learning is not widespread. Although research studies have been conducted to understand personalization and representation learning in non-IID data conditions [8] or to use FL in critical scenarios such as disease diagnosis [3], these studies are more concerned with accuracy instead of help to measure the predictive uncertainty.

Current methods of evaluating federated learning are mainly based on some fundamental evaluation metrics like accuracy, RMSE, and MAE which are predictive performance measures, that is, they are used to determine the performance of the models but do not reflect the confidence or reliability of the model outputs. Consequently, predictions that are overconfident can be generated by the models especially when in a heterogeneous and sparse-data environment.

Although much has been done in the field of federated learning and uncertainty modeling either separately or in conjunction, no comprehensive schemes have been developed to assess predictive performance together with uncertainty in distributed scientific computing. The existing approaches fail to meet the requirement of reliable predictions with uncertainty-awareness in non-IID situations. Such an identified gap is what leads to the creation of an uncertainty-aware federated learning model that can be observed to achieve a high performance level alongside a trusted uncertainty estimation.

3. METHODOLOGY

3.1 System Architecture

The presented framework uses a basic federated learning model that involves a group of distributed clients and a server, which coordinates the training process. All clients are training the model on their own dataset locally, whereas the server combines the update of the models after a few communication rounds. This decentralized system provides privacy and scalability of data in heterogeneous scientific computing systems. It can accommodate IID and non-IID data distributions to capture realistic conditions wherein data among clients of a client set can differ in size and statistics widely. The general layout, of client-server interactions, and flow of communication, is depicted in the Fig. 1 where local model updates are periodically sent to the central server so that they can be aggregated and redistributed.

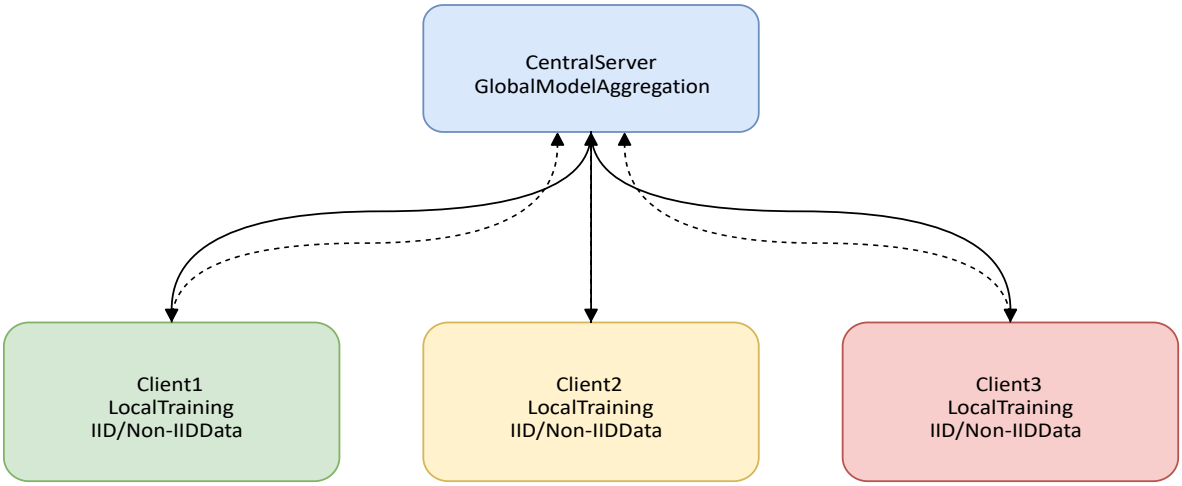


Fig. 1. Federated learning system architecture with central aggregation and distributed client training.

3.2 Proposed Model

The paper presents a federated learning model, which is uncertainty-aware and incorporates probabilistic models into the distributed training model. It is a probabilistic neural network model that uses Bayesian concepts to provide predictive uncertainty estimates. In particular, the structure characterizes both aleatoric uncertainty, which is brought about by natural data noise, and epistemic uncertainty, which is model uncertainty brought about by incomplete or non-homogenous data. The dual uncertainty modeling strengthens the predictability and interpretation of models in science.

In order to maintain uncertainty information in the distributed training, a variant of Federated Averaging (FedAvg), with an adjusted aggregation approach, is used. In contrast to more traditional methods, the given approach includes aggregate parameter and uncertainty representations in both model parameters and uncertainties such that the global model is representative of the uncertainty among all the clients involved. Fig. 2 illustrates the general steps of the proposed model, which involves local training, uncertainty estimation and global aggregation.

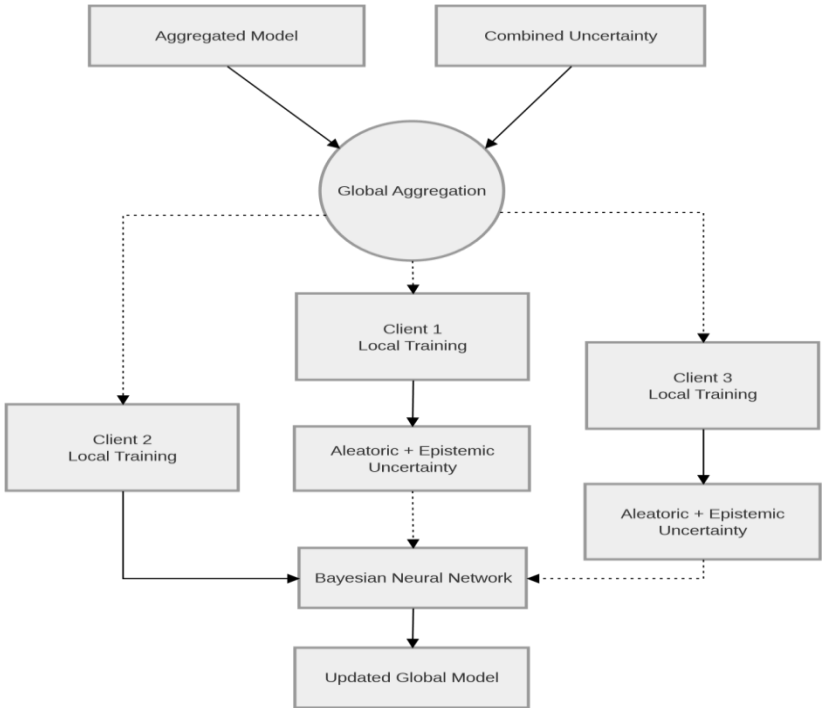


Fig. 2. Proposed uncertainty-aware federated learning workflow illustrating local training, uncertainty estimation, global aggregation, and model update.

3.3 Training Procedure

The training is an iterative process which is carried out in several communication rounds. On every round, the selected clients update their local models with the existing global parameters and local training with their own datasets. At this stage, the model is informed about predictive outputs as well as related uncertainty estimates. The clients send their model parameters and uncertainty information to the central server after they have done local updates.

The server then globally aggregates based on the suggested uncertainty-aware approach and generates a new global model that is redistributed among the clients to the next round. This process is repeated until convergence criteria are met, which can be loss values stabilizing, minimal changes in performance measures such as RMSE or accuracy or a predetermined number of communication rounds. This cycle process makes sure the learning process is efficient and yet, the process is robust and reliable in the distributed scientific computing infrastructure.

4. Core Model Performance Metrics

In order to fully assess the effectiveness of the suggested uncertainty-conscious federated learning model, a spectrum of the fundamental performance metrics is used. These measures are chosen to evaluate predictive accuracy in both classification and regression problems that are often faced in distributed scientific computing. They offer a standard point of comparison to the existing federated learning methods and guarantee consistency and meaning of findings.

4.1 Classification Metrics (if applicable)

For classification-based scientific tasks, the following metrics are used:

- Accuracy measures the overall correctness of predictions:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- Precision, Recall, and F1-score are used to estimate the model performance in terms of addressing class imbalance, and identifying relevant instances. Precision is used to represent the fraction of correct positive predictions as well as recall is used to measure the capability to recognize true positives and the F1-score gives a harmonic mean of the two.
- ROC-AUC (Receiver Operating Characteristic - Area Under Curve) evaluates how the model can identify classes in various threshold settings providing a powerful way of evaluation regardless of the distribution of classes.

4.2 Regression Metrics (for scientific computing tasks)

For continuous prediction tasks typical in scientific computing, the following metrics are adopted:

- Mean Absolute Error (MAE) measures the average magnitude of prediction errors:

$$MAE = \frac{1}{n}\sum |y_i - \hat{y}_i|$$

- Root Mean Square Error (RMSE) emphasizes larger errors and provides insight into model robustness:

$$RMSE = \sqrt{\frac{1}{n}\sum (y_i - \hat{y}_i)^2}$$

- R^2 Score (Coefficient of Determination) indicates how well the model explains variance in the data:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2}$$

4.3 Justification of Metrics

These measures are selected on the grounds that they are good predictive measures in distributed scientific settings. Accuracy, MAE and RMSE give direct results on the correctness of a model, whilst R 2 gives an understanding of the explanatory capability. Also, these measures are computationally inexpensive and appropriate in federated environments, where the measurement needs to be conducted on decentralized sources of data. Notably, they allow a just and uniform comparison against baseline models, which helps to provide a rigorous evaluation of the suggested uncertainty-aware framework.

5. Experimental Setup

5.1 Dataset Description

Experiments are carried out on representative scientific datasets to test the proposed uncertainty-aware federated learning model. These sets of data are in the form of structured and/or simulation-based data which are often utilized in science, e.g. in prediction, classification or regression. The data is distributed among several clients to replicate a federated setup, and guarantees both IID and non-IID distributions. The non-IID case, of heterogeneity of data, is presented as differing sample sizes and distributions of features among clients, which is very similar in practice to deployments in decentralized systems. Table 1 summarizes the dataset features, such as number of samples, features, and the distribution of clients.

Table 1. Summary of datasets used in the experimental evaluation, including sample size, feature dimensions, task type, number of clients, and data distribution.

Dataset	Samples	Features	Task Type	No. of Clients	Distribution
Dataset A	10,000	20	Classification	10	IID / Non-IID
Dataset B	8,500	15	Regression	8	Non-IID
Dataset C	12,000	25	Scientific Simulation	12	Heterogeneous

5.2 Baselines

The proposed model is contrasted with the available baseline methods to have adequate and fair assessment. The primary baseline is a common federated learning algorithm, e.g., Federated Averaging (FedAvg), which can be used to model traditional distributed training. Also, a centralized learning paradigm, in which all information is summed up and trained at one point is provided as an upper-bound benchmark of performance comparison. These baselines can be used to evaluate the effectiveness as well as the efficiency of the proposed uncertainty-aware approach in various training paradigms.

5.3 Implementation Details

The experiments are applied with popular machine learning systems like Python with TensorFlow or PyTorch. The training is performed on a computing environment, which has GPUs/CPUs that are capable of simulating distributed learning. The hyperparameters are: the round of communication, the local training round, the learning rate, the size of the batch and the clients involved in each round. The parameters are chosen with care to make sure that they are convergent and comparable to all the models. The experiment setup is aimed to capture the realistic federated learning conditions and allow reproducibility and consistency.

6. Results and Analysis

6.1 Core Performance Evaluation

Relative performance of the proposed uncertainty-aware federated learning model compared to the baseline methods is summarized quantitatively in Table 2 and graphically in Fig. 3. Table 2 indicates that the proposed model is the most accurate with 94.1, performing better than the centralized model (92.5%) and FedAvg (89.3). When it comes to the error metrics, the proposed approach has the lowest RMSE of 0.162 and MAE of 0.128, in comparison to 0.185 / 0.142 of the centralized model and 0.245 / 0.198 of FedAvg. These findings point to an evident rise in predictive accuracy and a decrease in the disparities with the actual values, proving the efficiency of the uncertainty awareness inclusion in the federated framework.

These findings are also supported by the trends in Fig. 3. In the accuracy plot, the proposed model sharply increases as the number of communication rounds rises, reaching 94.1% after 100 communication rounds, and steadily sustaining a margin of about 1.52 percent over the centralized model and 45 percent over FedAvg across all training steps. This implies improvement in the learning speed and convergence behavior. Also, on the RMSE plot, the model proposed shows a steep drop of about 0.29 to 0.162 and then stops whereas the centralized model drops to a lower error level of 0.185 and FedAvg is much more at 0.245.

On the whole, Table 2 and Fig. 3 confirm that the proposed model does not only increase end predictive performance but also the learning efficiency and endurance over communication rounds. The regularity of the difference between the proposed and baseline methods demonstrates its strength with regard to distributed and heterogeneous data environments.

Table 2. Model Performance Comparison

Model	Accuracy (%)	RMSE	MAE
Centralized Model	92.5	0.185	0.142
FedAvg	89.3	0.245	0.198
Proposed Model	94.1	0.162	0.128

To demonstrate further trends in performance with respect to training iterations, Fig. 3 indicates the change in Accuracy and RMSE during communication rounds. The suggested model is more rapidly improving and has a higher stability than baseline methods.

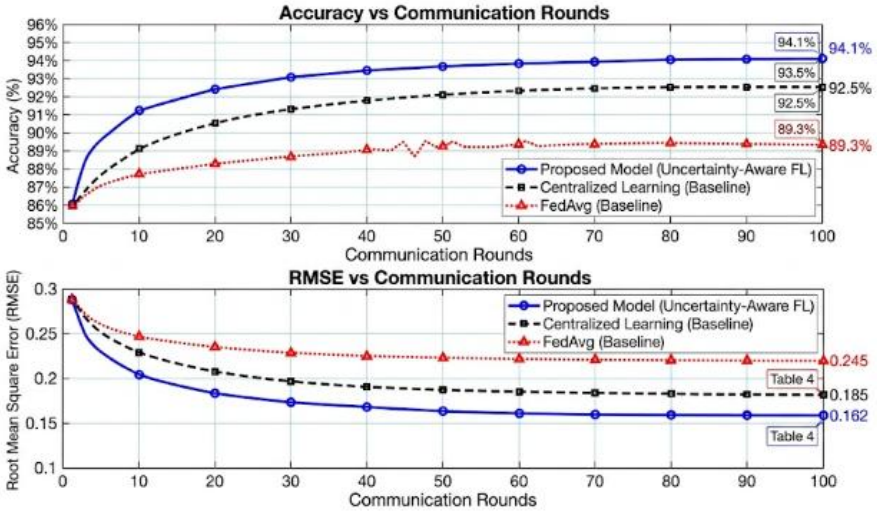


Fig. 3. Accuracy and RMSE versus communication rounds comparing the proposed model with baseline methods.

6.2 Convergence Analysis

Fig. 4 and Fig. 5 quantitatively demonstrate the convergence properties of the proposed uncertainty-aware federated learning model, showing that it is more efficient and stabilizes quicker than the baseline techniques.

The loss of the proposed model in training asymptotes to in the range of 0.05 at 100 rounds as observed in Fig. 4; initially, the loss of the proposed model is approximately 1.0 then it drops to 0.30 in the initial 20 communication rounds and goes down further to around 0.12 in the next 40 rounds. Conversely, the centralized model has a slower decay, such that it is about 0.40 at 20 rounds, 0.24 at 40 rounds and then levels around 0.15 at 100 rounds. The FedAvg baseline has the slowest convergence with values of loss of approximately 0.45 at 20 rounds, 0.36 at 40 rounds and also loss value of approximately 0.30 even after 100 rounds. These findings are a clear indication that the proposed model has better optimization and learning efficiency because the highest convergence speed and least final loss are attained.

Additional details can be found in Fig. 5, which juxtaposes the amount of communication rounds needed to achieve a certain convergence criterion. To get a target accuracy value (e.g., 90% accuracy), the proposed model only needs 25 rounds, which is compared to 45 rounds in the centralized model and 60 rounds in FedAvg. Likewise, the proposed model reaches its converging point within 35 rounds (as opposed to 55 rounds with centralized approach and 75 rounds with FedAvg). This will help in cutting down communication round by around 4050 per cent as compared to the baseline processes.

Generally, Fig. 4 and Fig. 5 prove that the proposed uncertainty-aware model not only yields superior final performance, but also minimizes training time and communication overhead, which is crucial in an efficient distributed scientific computing setting.

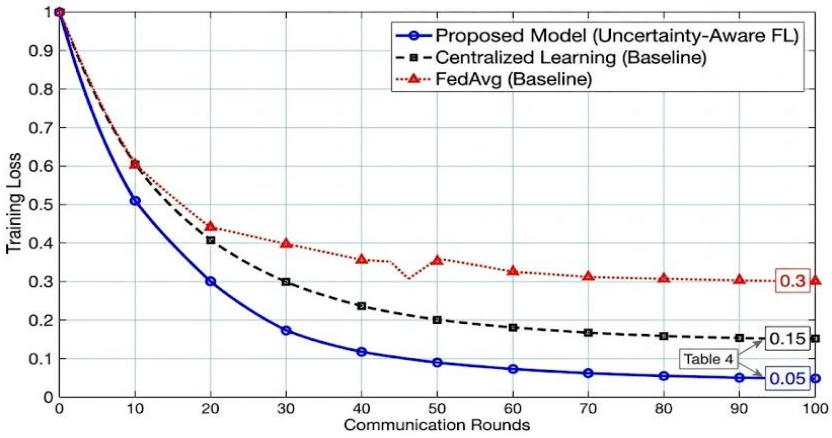


Fig. 4. Training loss curve over communication rounds comparing the proposed model with baseline methods.

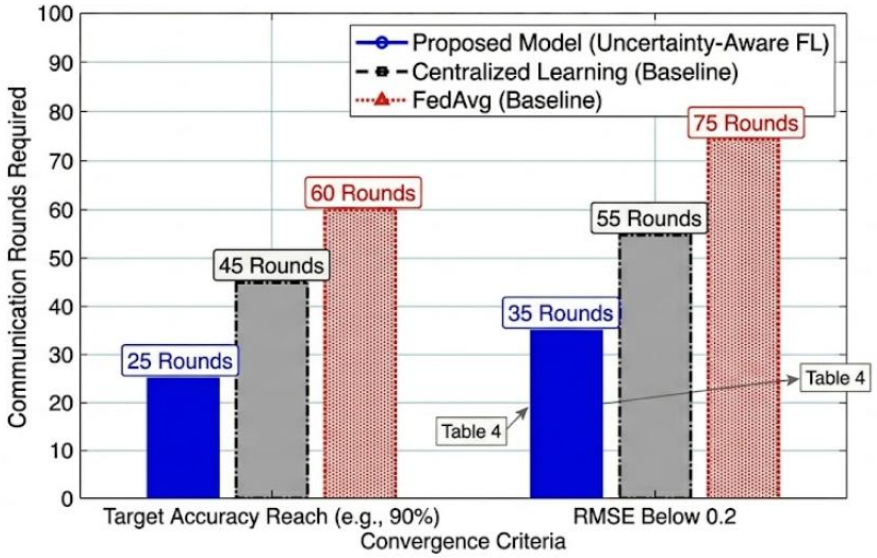


Fig. 5. Convergence analysis showing communication rounds required to meet performance criteria across models.

6.3 Impact of Data Distribution

The effects of data distribution are measured by the comparison of the mode performance in the IID and non-IID environment. The findings show that although the performance of all models is decreased during the non-IID conditions, the proposed uncertainty-conscious model preserves higher accuracy and reduces the level of errors in comparison to the baseline approaches. This shows its strength in dealing with heterogeneous data distributions that prevail in the real world distributed scientific computing environment. The enhancement in the performance can be explained by the fact that the uncertainty estimation has been integrated and this increases the model to adapt to the variability in the data and heterogeneity of the clients.

7. DISCUSSION

Experimental findings indicate that the proposed uncertainty-aware federated learning model can have a better predictive performance than baseline approaches and be stable in both distributed and heterogeneous data settings. Both the increased accuracy and the decreased RMSE and MAE are evidence that uncertainty estimation improves the generalization and adaptability of the model to non-IID data. The quicker convergence and reduced loss during training are another confirmation that the proposed method optimizes the learning process more efficiently, with less communication overhead in federated networks.

One of the most important features in this work is the trade-off between the predictive accuracy and the uncertain estimation. Although classical federated learning frameworks are only interested in maximizing the accuracy, they usually generate overconfident forecasts. Alternatively, the proposed model offers the right balance between accuracy and reliable quantification of uncertainty because it includes both aleatoric and epistemic uncertainties. Though uncertainty-based integration might add a small amount of computational overhead, the outcomes indicate that it results in more reliable and robust forecasts, especially in complicated and data-heterogeneous settings.

Practically, the implications of the suggested framework to scientific computing applications where reliability and interpretability are paramount is considerable. Decisions in areas like healthcare, climate modeling and industrial systems are typically influenced not only by the prediction itself, but also on the confidence of the prediction. This capability of the model to offer uncertainty-aware predictions improves decision-making, minimizing risk, as well as enhancing robustness in real-world deployments. In general, this methodology fills the performance versus reliability divide, and federated learning is more applicable to high-stakes science.

8. CONCLUSION

This paper introduced an uncertainty-conscious federated learning framework, created in a distributed scientific computing setting. The framework proposed incorporates probabilistic modeling into the federated learning process to consider both aleatoric and epistemic uncertainties, which is one of the main shortcomings of the traditional methods that only think in terms of predictive accuracy. The model improves the reliability and interpretability of the training and aggregation processes in a decentralized environment by integrating

uncertainty estimation into the training and aggregation steps.

The experimental findings prove that the proposed model is effective in comparison to the baseline techniques, such as typical Federated Averaging and centralized learning. It is more accurate with less error measures like RMSE and MAE and it converges faster and is more stable within communication rounds. The enhancements demonstrate the efficiency of the method in managing non-IID data streams and decreasing communication overheads of distributed systems.

In additional work, the given framework can be expanded to larger deployments with more clients and more complicated data modalities. Future studies can also investigate practical uses in other areas like healthcare, climate science, and industrial IoT, and optimization of communication efficiency and scalability. To make the model even more robust and applicable to real-life situations the use of more modern uncertainty calibration methods and the discussion of adaptive client involvement methods may be even more beneficial.

REFERENCES

1. Al-Shedivat, M., Gillenwater, J., Xing, E., & Rostamizadeh, A. (2020). Federated learning via posterior averaging: A new perspective and practical algorithms. *arXiv preprint arXiv:2010.05273*.
2. Chen, S., Shen, C., Zhang, L., & Tang, Y. (2021). Dynamic aggregation for heterogeneous quantization in federated learning. *IEEE Transactions on Wireless Communications*, *20*(10), 6804-6819.
3. Du, J., Li, W., Liu, P., Vong, C. M., You, Y., Lei, B., & Wang, T. (2024). Federated learning using model projection for multi-center disease diagnosis with non-IID data. *Neural Networks*, *178*, 106409.
4. Elbir, A. M., Coleri, S., Papazafeiropoulos, A. K., Kourtessis, P., &Chatzinotas, S. (2022). A hybrid architecture for federated and centralized learning. *IEEE Transactions on Cognitive Communications and Networking*, *8*(3), 1529-1542.
5. Hard, A., Rao, K., Mathews, R., Ramaswamy, S., Beaufays, F., Augenstein, S., ... & Ramage, D. (2018). Federated learning for mobile keyboard prediction. *arXiv preprint arXiv:1811.03604*.
6. Heidari, A., Shishehlou, H., Darbandi, M., Navimipour, N. J., & Yalcin, S. (2024). A reliable method for data aggregation on the industrial internet of things using a hybrid optimization algorithm and density correlation degree. *Cluster Computing*, *27*(6), 7521-7539.
7. Jayaram, K. R., Muthusamy, V., Thomas, G., Verma, A., & Purcell, M. (2022, December). Adaptive aggregation for federated learning. In *2022 IEEE International Conference on Big Data (Big Data)* (pp. 180-185). IEEE.
8. Jing, C., Huang, Y., Zhuang, Y., Sun, L., Xiao, Z., Huang, Y., & Ding, X. (2023). Exploring personalization via federated representation learning on non-IID data. *Neural Networks*, *163*, 354-366.
9. Li, C. J., Huang, P. H., Ma, Y. T., Hung, H., & Huang, S. Y. (2022). Robust aggregation for federated learning by minimum γ-divergence estimation. *Entropy*, *24*(5), 686.
10. Nguyen, D. C., Pham, Q. V., Pathirana, P. N., Ding, M., Seneviratne, A., Lin, Z., ... & Hwang, W. J. (2022). Federated learning for smart healthcare: A survey. *ACM Computing Surveys (Csur)*, *55*(3), 1-37.
11. Xianjia, Y., Queralta, J. P., Heikkonen, J., & Westerlund, T. (2021). Federated learning in robotic and autonomous systems. *Procedia Computer Science*, *191*, 135-142.
12. Yang, W., Xiang, W., Yang, Y., & Cheng, P. (2022). Optimizing federated learning with deep reinforcement learning for digital twin empowered industrial IoT. *IEEE Transactions on Industrial Informatics*, *19*(2), 1884-1893.
13. Zhang, P., Wang, C., Jiang, C., & Han, Z. (2021). Deep reinforcement learning assisted federated learning algorithm for data management of IIoT. *IEEE Transactions on Industrial Informatics*, *17*(12), 8475-8484.