Flexibility prediction in Smart Grids
Making a case for Federated Learning

Čaušević, Selma; Snijders, Ron; Pingen, Geert; Pileggi, Paolo; Theelen, Mathilde; Warnier, Martijn; Brazier, Frances; Kok, Koen

DOI
10.1049/icp.2021.2196

Publication date
2021

Document Version
Final published version

Published in
IET Conference Proceedings

Flexibility prediction in Smart Grids: Making a case for Federated Learning

Selma Čaušević[1,2,,†] Ron Snijders[1,†], Geert Pingen[1], Paolo Pileggi[1], Mathilde Theelen[1], Martijn Warnier[2], Frances Brazier[2], Koen Kok[3]*

[1] *ICT Department, Netherlands Organisation for Applied Scientific Research (TNO), The Hague/Groningen, The Netherlands*
[2] *Systems Engineering Section, Delft University of Technology, Delft, The Netherlands*
[3] *Electrical Energy Systems, Eindhoven Univerisity of Technology, Eindhoven, The Netherlands*
† *These authors contributed equally to this article.*
* *selma.causevic@tno.nl*

Keywords: AGGREGATED DEMAND RESPONSE, LOAD FLEXIBILITY PREDICTION, FEDERATED LEARNING, SMART GRIDS

Abstract

High penetration of renewable energy sources brings both opportunities and challenges for Smart Grid operation. Due to their high contribution to energy consumption, aggregated load flexibility of small residential and service sector consumers has a potential to address the intermittency challenge of distributed generation. Predicting aggregated load flexibility of this consumer sector involves access to sensitive smart meter data, raising data collection and sharing concerns. Federated Learning, a decentralized machine learning technique that uses data distributed on user devices to construct an aggregated, global model, offers potential solutions to tackling this challenge. This paper explores the potential of using Federated Learning for flexibility prediction in Smart Grids through an analysis of its opportunities and implications for different stakeholders involved, as well as the challenges faced. The analysis shows that Federated Learning is a promising approach for building privacy-preserving energy portfolios of aggregated demand data.

1 Introduction

Power systems are large-scale, complex socio-technical systems that are continually in transition. Traditionally built as centralized systems, societal and technological changes drive modern power systems to become more decentralized. This transition is further supported by advancements in, and the incorporation of smart Information and Communication Technologies (ICT), resulting in Smart Grids. Such a decentralized energy system requires novel intelligent energy management techniques.

In an energy system with volatile, non-dispatchable production, these techniques rely on flexibility of energy consumers and prosumers to perform demand response (DR) and load shifting [1]. Flexibility refers to the extent to which these stakeholders can adapt their energy demand in response to changes. Reynders et al. [2] outline and discuss various definitions and quantification methods for flexible energy.

Advancements in smart home technologies empower small consumers and prosumers to become proactive members of the system and trade their flexibility in energy markets, which can render financial benefits [3]. Participation in demand response services and forming energy communities bring residential and service sector consumers at the core of the energy transition. Currently, DR programs are mainly focused on energy-intensive industrial and commercial consumers, as typical residential and service sector consumers are too small to individually participate in these programs [4]. However, due to their high contribution to electricity consumption, there is a large potential when considering their aggregated load flexibility [5]. Load flexibility aggregation for residential and service sector consumers comes with challenges regarding data collection and privacy preservation, as it involves collecting sensitive consumer data conform national and international directives on data collection and processing such as GDPR [6].

Another challenge lies in understanding demand heterogeneity of small residential and service sector consumers, as most existing models are based solely on residential load and do not include the service sector [7]. Therefore, service sector consumers' flexibility is not properly reflected in current models. To address these challenges, methods to predict aggregated load flexibility, while attaining data integrity and privacy in a decentralized energy system should be considered.

Federated Learning (FL), a decentralized Machine Learning (ML) technique [8], offers a solution to tackle these challenges. FL can enable different stakeholders to create global representative models without sharing raw data, addressing the challenges of data privacy [9]. In Smart Grids, FL can be used to create local energy flexibility models of different consumer and prosumer categories and to aggregate these into global models, without exchanging local data.

The main objective of this paper is to further explore the application of FL for load flexibility prediction (more specifically, demand response) of small-sized consumers in Smart

Grids through an analysis of its opportunities, challenges and implications for different stakeholders.

2 Load flexibility prediction using Federated Learning: A stakeholder analysis

This section gives an overview of FL and analyses the potential of using this technique for addressing the challenges faced by stakeholders involved in aggregating load flexibility of small residential and service sector consumers and prosumers.

2.1 An overview of Federated Learning

Capable of discovering complex patterns from raw data, ML techniques are used for load flexibility prediction in Smart Grids. However, due to privacy concerns and data ownership, collecting all required data becomes challenging [10]. Federated Learning, a class of ML, capable of learning a single model across distributed devices, offers a solution to this. In FL none of the original data samples needs to be shared among different devices, thus offering potential benefits in lowering bandwidth requirements, data ownership and preserving privacy of its users.

The term and general concept of FL has originally been introduced in 2016 by McMahan et. al. [8] in order to learn keyboard suggestions on mobile phones [11]. Since then, it has been applied in many other domains. For Deep Learning [12], Federated Stochastic Gradient Descent (FedSGD) [13] and Federated Averaging (FedAVG) [8] are well-known variants for FL. FL can also be used for other types of machine learning models, such as Support Vector Machines [14].

FL aims to provide a solution to the problem of learning a global model from distributed data sets that should remain local, on nodes (data owners) with sufficient computation power to fit local models. In the context of this paper, the nodes are consumers and prosumers. For these data owners, the main opportunity therefore lies in maintaining data privacy whilst enabling local contributions to a shared model. This means all individual nodes can reap the benefits of each others contributions to the global model and at the same time keep their personal data private. From a model owner perspective (e.g. an energy aggregator), computation is now partially offloaded to local nodes, which opens up opportunities for cost reduction.

Despite the applicability of FL in the energy domain, current research on this topic is limited [15, 16].

The following subsections discuss the potential implications of using FL by different stakeholders that participate in aggregated small residential and service sector demand response.

2.2 Consumers and prosumers

Aggregating load flexibility can enable small residential and service sector consumers and prosumers to jointly participate in energy markets through DR programs, which can bring them financial benefits. As the main concern lies in data privacy (as some consumers are reluctant to share private information with centralized ML models used by aggregators) [17], FL offers a potential solution to this, as private consumer and prosumer data can stay on their devices, without the need to be shared with other stakeholders. Therefore, using FL could potentially increase the willingness of small consumers and prosumers to participate in aggregated DR.

2.3 Aggregators

This aggregated flexibility can be used by aggregators to create energy portfolios, and trade them in energy markets [18, 19]. Depending on their business model, aggregators with different roles face a number of challenges. Information exchange is identified as one of these challenges, as different stakeholders need access to aggregator's data to enable accurate load forecasting [20]. The privacy-preserving property of FL offers a potential opportunity to address this challenge, as stakeholders involved do not have to share data with each other. Another challenge lies in data availability, as aggregators' access to data generated by smart energy devices is not guaranteed and can be incomplete, questioning data integrity [18]. To tackle this challenge, aggregators can use global flexibility models (potentially hosted on their own devices) generated for specific consumer and prosumer categories to get an estimation of offered flexibility.

2.4 Distribution System Operators

High penetration of renewable energy sources (RES) brings both opportunities and challenges for Distribution System Operators (DSOs) when it comes to grid planning and operation. Load flexibility prediction of small residential and service sector consumers' demand reponse is a promising approach to deal with variability and uncertainty of RES. To have an impact on network planning and RES integration, small consumer flexibility should be considered on an aggregated level [21]. From the perspective of DSOs, FL can offer a number of opportnunities to integrate small residental and service sector DR into grid operation planning, and better respond to uncertainties and non-dispatchability of RES. Global FL models that aggregate a large number of small residential and service sector consumers on geographically distant locations, can be used by DSOs to better understand the DR of this sector. This knowledge can be used to gain insight into local flexibility, when data is unavailable or insufficient, and help make better plans for future grid investments and planning.

2.5 Policy and decision makers

As mentioned before, another challenge lies in the lack of representative service sector models, which can lead to significant misestimations for RES integration [7]. FL can be used to construct global models of different consumer categories (e.g. residential, service sector) and subcategories (e.g. offices, supermarkets, schools) to better use their potential in demand response and RES integration. These models can support policy and decision makers to get better insight into future energy transition scenarios.

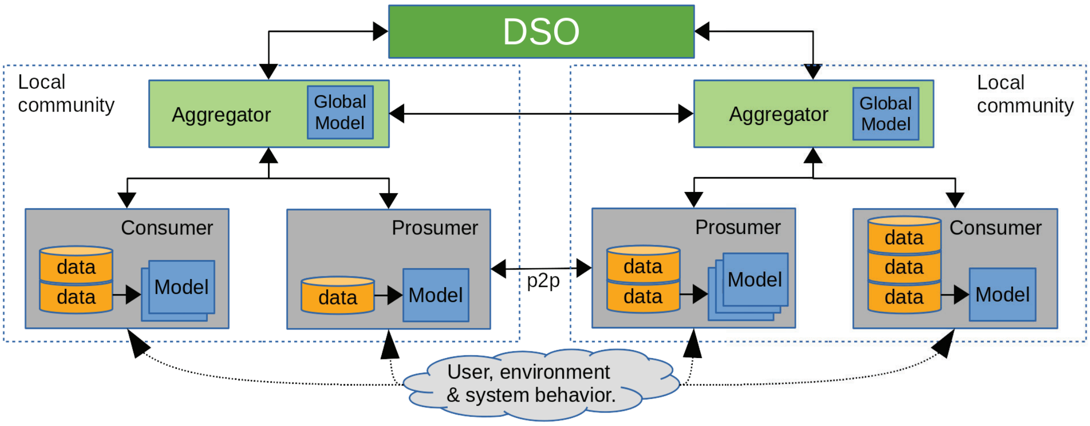

Fig. 1 An example application of using Federated Learning within a Smart Grid. A local community (e.g. an energy neutral village) typically consists of multiple consumers, producers and prosumers, which in turn consists of multiple devices that may require a different type of model. Data is typically distributed unevenly over multiple devices and influenced by different user behavior, system behavior and other environmental influences.

In this section, the potential of using FL for aggregated load flexibility prediction is discussed with respect to the stakeholders involved. However, despite its promising benefits, FL comes with challenges that need to be addressed. The next section outlines these challenges and gives guidelines and recommendations to address them.

3 Addressing the challenges of Federated Learning: Guideliness and recommendations

Fig. 1 provides an abstract example application of using FL within the energy domain and illustrates some of the challenges as discussed further in the following subsections. Most variants use a single central server to aggregate all changes to the global model from multiple clients. However, as the ecosystem grows in complexity and local communities become more independent, it is likely to assume that multiple of these central servers (or model aggregators) may exist for the purpose of improving scalability or to abstract the global model at a certain level. This might especially be relevant in the energy domain which consists of multiple (competing) stakeholders and communities that may heaviliy rely on self-organization.

3.1 Data Distribution

Many ML techniques often have the underlying assumption that training data is sampled Independent, and is Identically Distributed (IID) in order to obtain an unbiased estimate of the gradient that is being approximated [22, 23]. A major challenge in FL, therefore, is the use of non-IID, statistically heterogeneous, and vertically distributed data.

Solutions have been proposed such as, item reparametrization of existing models to help them converge in the heterogeneous scenario [24], sharing a small IID subset of the data globally, or using proxy data [22].

3.2 Communication Efficiency

Communication is a primary bottleneck of distributed computation, especially the case of FL that often requires many rounds of sharing updates from resource-constrained nodes. It is therefore key to balance the size of model updates, communication frequency, sparsity and model performance, which remains an open problem [23].

3.3 Privacy and Security

Providing that none of the data samples are shared among the devices, it seems intuitive to assume that this protects the privacy of the user. However, sensitive information can still become compromised during the updates of the model.

For instance, in the work of [25], a Generative Adversarial Network (GAN) is trained to generate the original samples as used to train the model of other users. Alternatively, the (aggregated) gradients, that are shared from users to a central server (and other users through updates of the global model), may compromise sensitive information about the private training data being used [26].

Apart from the original input, aggregated gradients can also be used to classify the presence of specific properties (not present in the input features) using a separate trained classifier [27]. Potential solutions to these attacks include the Double Blind Collaborative Learning algorithm [28], which uses random matrix sketching for the parameters on the central server side to obscure the information between model updates.

3.4 Model Bias, Fairness and Personalization

FL also poses challenges in reducing unwanted bias and increasing fairness in the model [29]. Global models might be biased towards specific users and may achieve poor accuracy on an individual level. For instance, devices with poor connectivity in rural areas might not be able to participate in the federated learning scheme as often. This increases the risk of biasing the global model towards devices within an urban area which may have a very different energy profile.

Potential solutions to these problems include Model Agnostic Meta Learning [30] or Agnostic Federated Learning [31]. The latter uses a mixture of client distributions to optimize the central model for any target distribution. An alternative is to use clustering in order to match the best global model to specific users, which is similar to what is done for energy demand prediction for electric vehicle networks [15].

The key challenge is to find a way to use generic information from the global model (e.g., by learning typical usage profiles over time) while adapting it to a particular situation (e.g., by including personal preferences and local differences in the environment). For this purpose, one could also consider using methods similar to Meta-Learning [32, 33] or Transfer Learning [34, 35] to adapt pretrained models to a particular situation. Or using local and global representations to account for the heterogeneity in the data [36].

4 Conclusion

Due to their high contribution to energy consumption, aggregating load flexibility of small residential and service sector consumers has a potential to address the intermittency challenge of distributed generation. However, predicting aggregated load flexibility of this consumer sector involves access to sensitive smart meter data, raising data collection and sharing concerns. With its privacy-preserving properties for data aggregation, FL offers a potential solution to tackle this challenge. The analysis discussed potential benefits for stakeholders involved, potentially resulting in higher consumer participation in demand response programs, and getting a better insight in residential and service sector flexibility. This paper shows that, given the need for privacy preservation, increased scalability and the shift towards decentralization, FL is a promising approach to support privacy-preserving data aggregation.

Using FL comes with challenges that need to be addressed. As using such models can potentially incentivise consumers to share their data, aggregators should start investigating methods to model the value of those contributions to a global model, and define billing and pricing models on top of that. Furthermore, as more consumers participate, security measures have to be researched, as there is a possibility for influx of potentially harmful data. Finally, as decentralized approaches for flexibility prediction become more prominent, the future work should also focus on how to share different models and adapt them for various applications.

5 Acknowledgements

This research has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 872613 (H2020 project BD4NRG "Big Data for Next Generation Energy"). The authors acknowledge Aliene van der Veen and Wilco Wijbrandi for discussions that improved the quality of the manuscript.

6 References

[1] M. H. Weck, J. van Hooff, and W. G. van Sark, "Review of barriers to the introduction of residential demand response: a case study in the Netherlands," *International Journal of Energy Research*, vol. 41, no. 6, pp. 790–816, 2017.

[2] G. Reynders, R. A. Lopes, A. Marszal-Pomianowska, D. Aelenei, J. Martins, and D. Saelens, "Energy flexible buildings: An evaluation of definitions and quantification methodologies applied to thermal storage," *Energy and Buildings*, vol. 166, pp. 372–390, 2018.

[3] R. Henriquez, G. Wenzel, D. E. Olivares, and M. Negrete-Pincetic, "Participation of demand response aggregators in electricity markets: Optimal portfolio management," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4861–4871, 2018.

[4] J. Iria, F. Soares, and M. Matos, "Optimal supply and demand bidding strategy for an aggregator of small prosumers," *Applied Energy*, vol. 213, no. September 2017, pp. 658–669, 2018. [Online]. Available: https://doi.org/10.1016/j.apenergy.2017.09.002

[5] A. Lucas, L. Jansen, N. Andreadou, E. Kotsakis, and M. Masera, "Load flexibility forecast for dr using non-intrusive load monitoring in the residential sector," *Energies*, vol. 12, no. 14, 2019. [Online]. Available: https://www.mdpi.com/1996-1073/12/14/2725

[6] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). European Commission. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj

[7] N. Voulis, M. Warnier, and F. M. Brazier, "Impact of service sector loads on renewable resource integration," *Applied Energy*, vol. 205, no. July, pp. 1311–1326, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.apenergy.2017.07.134

[8] H. B. McMahan, E. Moore, D. Ramage, S. Hampson *et al.*, "Communication-efficient learning of deep networks from decentralized data," *arXiv preprint arXiv:1602.05629*, 2016.

[9] Q. Yang, Y. Liu, Y. Cheng, Y. Kang, T. Chen, and H. Yu, "Federated Learning," *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 13, no. 3, pp. 1–207, 2020.

[10] R. Ahmadiahangar, T. Häring, A. Rosin, T. Korõtko, and J. Martins, "Residential Load Forecasting for Flexibility Prediction Using Machine Learning-Based Regression Model," *Proceedings - 2019 IEEE International Conference on Environment and Electrical Engineering and 2019 IEEE Industrial and Commercial Power Systems Europe, EEEIC/I and CPS Europe 2019*, pp. 10–13, 2019.

[11] B. McMahan and D. Ramage, “Federated learning: Collaborative machine learning without centralized training data,” *Google Research Blog*, vol. 3, 2017.

[12] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[13] J. Chen, X. Pan, R. Monga, S. Bengio, and R. Jozefowicz, “Revisiting distributed synchronous sgd,” *arXiv preprint arXiv:1604.00981*, 2016.

[14] T. S. Brisimi, R. Chen, T. Mela, A. Olshevsky, I. C. Paschalidis, and W. Shi, “Federated learning of predictive models from federated electronic health records,” *International journal of medical informatics*, vol. 112, pp. 59–67, 2018.

[15] Y. M. Saputra, D. T. Hoang, D. N. Nguyen, E. Dutkiewicz, M. D. Mueck, and S. Srikanteswara, “Energy demand prediction with federated learning for electric vehicle networks,” in *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2019, pp. 1–6.

[16] A. Taik and S. Cherkaoui, “Electrical Load Forecasting Using Edge Computing and Federated Learning,” *IEEE International Conference on Communications*, vol. 2020-June, 2020.

[17] “A review of residential demand response of smart grid,” *Renewable and Sustainable Energy Reviews*, vol. 59, pp. 166–178, 2016.

[18] C. Ziras, C. Heinrich, M. Pertl, and H. W. Bindner, “Experimental flexibility identification of aggregated residential thermal loads using behind-the-meter data,” *Applied Energy*, vol. 242, no. March, pp. 1407–1421, 2019. [Online]. Available: https://doi.org/10.1016/j.apenergy.2019.03.156

[19] J. Iria and F. Soares, “A cluster-based optimization approach to support the participation of an aggregator of a larger number of prosumers in the day-ahead energy market,” *Electric Power Systems Research*, vol. 168, no. July 2018, pp. 324–335, 2019. [Online]. Available: https://doi.org/10.1016/j.epsr.2018.11.022

[20] O. Okur, P. Heijnen, and Z. Lukszo, “Aggregator’s business models: Challenges faced by different roles,” *IEEE PES Innovative Smart Grid Technologies Conference Europe*, vol. 2020-October, pp. 484–488, 2020.

[21] Ö. Okur, P. Heijnen, and Z. Lukszo, “Aggregator’s business models in residential and service sectors: A review of operational and financial aspects,” *Renewable and Sustainable Energy Reviews*, vol. 139, no. March 2020, p. 110702, 2021. [Online]. Available: https://doi.org/10.1016/j.rser.2020.110702

[22] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, “Federated learning with non-iid data,” *arXiv preprint arXiv:1806.00582*, 2018.

[23] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, “Federated learning: Challenges, methods, and future directions,” *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 50–60, 2020.

[24] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, “Federated optimization in heterogeneous networks,” *arXiv preprint arXiv:1812.06127*, 2018.

[25] B. Hitaj, G. Ateniese, and F. Perez-Cruz, “Deep models under the gan: information leakage from collaborative deep learning,” in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, pp. 603–618.

[26] L. Zhu, Z. Liu, and S. Han, “Deep leakage from gradients,” in *Advances in Neural Information Processing Systems*, 2019, pp. 14 774–14 784.

[27] L. Melis, C. Song, E. De Cristofaro, and V. Shmatikov, “Exploiting unintended feature leakage in collaborative learning,” in *2019 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2019, pp. 691–706.

[28] S. Wang, “Matrix sketching for secure collaborative machine learning,” *arXiv preprint arXiv:1909.11201*, 2019.

[29] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings *et al.*, “Advances and open problems in federated learning,” *arXiv preprint arXiv:1912.04977*, 2019.

[30] Y. Jiang, J. Konečnỳ, K. Rush, and S. Kannan, “Improving federated learning personalization via model agnostic meta learning,” *arXiv preprint arXiv:1909.12488*, 2019.

[31] M. Mohri, G. Sivek, and A. T. Suresh, “Agnostic federated learning,” *arXiv preprint arXiv:1902.00146*, 2019.

[32] C. Finn, P. Abbeel, and S. Levine, “Model-agnostic meta-learning for fast adaptation of deep networks,” *arXiv preprint arXiv:1703.03400*, 2017.

[33] Q. Sun, Y. Liu, T.-S. Chua, and B. Schiele, “Meta-transfer learning for few-shot learning,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2019, pp. 403–412.

[34] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[35] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, “A survey on deep transfer learning,” in *International conference on artificial neural networks*. Springer, 2018, pp. 270–279.

[36] P. P. Liang, T. Liu, L. Ziyin, R. Salakhutdinov, and L.-P. Morency, “Think locally, act globally: Federated learning with local and global representations,” *arXiv preprint arXiv:2001.01523*, 2020.