Global Sensitivity Analysis of Large Distribution System With PVs Using Deep Gaussian Process

Ketian Ye, Junbo Zhao, *Senior Member, IEEE*, Fei Ding, Rui Yang, Xiao Chen, and George W. Dobbins

***Abstract*—Global sensitivity analysis (GSA) of voltage to uncertain power injection variations plays an important role for appropriate Volt-VAR optimization. This paper proposes a data-driven GSA method for large-scale distribution systems with a large number of uncertain sources. Specifically, the deep Gaussian process is used to identify the mapping relationship between uncertain power injections and voltages. This allows resorting to the analysis of variance framework to calculate the Sobol indices for GSA. Unlike the existing polynomial chaos expansion and Gaussian process-based approaches, our proposed method has much better scalability. Test results on the EPRI 1747-node K1 circuit with different numbers of uncertain sources with various uncertain levels and different PV distributions demonstrate that the proposed method can achieve accurate GSA.**

***Index Terms*—Distribution system analysis, deep Gaussian process, PVs, global sensitivity analysis, Sobol indices.**

I. Introduction

THE increased penetration of uncertain distributed energy resources (DERs), especially PVs, brings challenges for Volt-VAR optimization. The Jacobian matrix-based local sensitivity analysis (LSA) of voltage to real and reactive power injections has been widely used to guide the voltage control [1], [2]. However, LSA could not exhibit the overall impacts of the uncertain sources on voltage variations. By contrast, global sensitivity analysis (GSA) allows us to quantify how voltage variations are correlated with variation of each uncertain input and their joint contributions. To this end, the polynomial chaos expansion (PCE)-based GSA is developed for distribution systems [3] and for microgrids [4]. However, PCE is model-based and has the curse of dimensionality issue when the number of uncertain inputs is large, which is typically the case for large-scale distribution systems. To mitigate the modeling challenge, data-driven method for GSA can also be developed via kriging, i.e., Gaussian process (GP) [5]. But GP still has the curse of dimensionality issue. This letter develops a scalable data-driven GSA framework for large-scale distribution systems. Instead of using the single layer GP, we advocate the deep Gaussian process (DGP) by constructing the GP model hierarchically [7]. The DGP is used to identify the surrogate model that retains the relationships between uncertain power injections from uncertain sources and voltage variations. We use the Sobol indices (SI) in assessing the GSA.

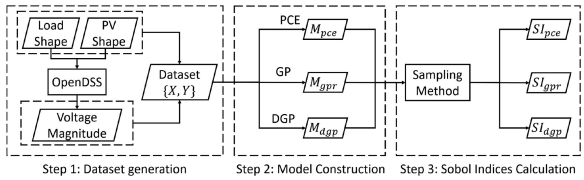

Fig. 1. The overall framework of GSA.

II. Proposed Data-Driven GSA Framework

The proposed framework is shown in Fig. 1 and it contains three main procedures, namely the data collection, surrogate modeling and SI calculation. The surrogate model can be achieved offline by PCE, or Kriging, or DGP. After that, SI is computed online using Monte-Carlo (MC) sampling method.

A. Voltage Sensitivity Analysis

Let $\boldsymbol{y} = \mathcal{M}(\boldsymbol{x})$ the model with d-dimensional random input vector $\boldsymbol{x} = [x_1, \ldots, x_d]^T$ and model response $\boldsymbol{y}$. For three-phase distribution system voltage sensitivity analysis, the uncertain model inputs are PV injections and loads while the outputs are typically bus voltage magnitudes V. Sensitivity analysis aims to quantify how the model response is affected by each input or their combinations. LSA studies the impacts of small perturbations on the model output by estimating the partial derivative with respect to the inputs at a specific point: $(\partial y/\partial x_k)_{x=x^*}$. This approach is called local since the perturbations occur in the neighborhood of the nominal values. By contrast, GSA reveals the global relationship between input and output variations by providing a quantitative importance ranking with respect to each input and the combination thereof.

B. Sobol Indices

Assume the input variable $\boldsymbol{x}$ is with support $D_x = [0, 1]^d$. Using the analysis of variance framework [3], the function $\mathcal{M}(\boldsymbol{x})$ can be decomposed as:

Manuscript received January 9, 2021; revised March 10, 2021 and April 10, 2021; accepted May 17, 2021. Date of publication May 27, 2021; date of current version August 19, 2021. This work was supported by the U.S. Department of Energy Solar Energy Technology Office and National Science Foundation under Grant ECCS-1917308. Paper no. PESL-00008-2021. *(Corresponding author: Junbo Zhao.)*

Ketian Ye and Junbo Zhao are with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS 39762 USA (e-mail: ky291@msstate.edu; junbo@ece.msstate.edu).

Fei Ding and Rui Yang are with National Renewable Energy Laboratory, Golden, CO 80401 USA (e-mail: fei.ding@nrel.gov; rui.yang@nrel.gov).

Xiao Chen is with Lawrence Livermore National Laboratory, Livermore, CA 94550 USA (e-mail: chen73@llnl.gov).

George W. Dobbins is with the Southern Company, Avenue Southwest, Reform, AL 35481 USA (e-mail: gwdobbin@southernco.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPWRS.2021.3084455.

Digital Object Identifier 10.1109/TPWRS.2021.3084455

$$\mathcal{M}(x_1,\ldots,x_d)=\mathcal{M}_0+\sum_{i=1}^{d}\mathcal{M}_i(x_i)+\sum_{1\le i<j\le d}\mathcal{M}_{ij}(x_i,x_j)+\cdots+\mathcal{M}_{1,\ldots,d}(x_1,\ldots,x_d) \quad (1)$$

with the condition: $\int_0^1 \mathcal{M}_{i_1,\ldots,i_s}(x_{i_1},\ldots,x_{i_s})\,dx_i=0$.

Using the variance operator over D_x, the expression of variance decomposition is:

$$var(Y)=V=\sum_{i=1}^{d}V_i+\sum_{i<j}^{d}V_{ij}+\cdots+V_{1,\ldots,d} \quad (2)$$

Based on that, the SI is defined as: $S_I=V_I/V$, where $I\subset\{1,\ldots,d\}$. SI measures the contribution of each input variable to the output variability. The first-order Sobol index represents the impact of each individual uncertain source on the variance of the output. The higher order SI quantifies the effect of the interaction of grouped inputs. To consider the overall importance of each input, the total SI is $S_i^T=\sum_{\{i_1,\ldots,i_s\}\supset i}S_{i_1\ldots i_s}$. A commonly used approach to estimate SI is based on MC simulations, i.e.,

$$\begin{cases}\widehat{V}_0=\frac{1}{N}\sum_{n=1}^{N}\mathcal{M}(\boldsymbol{x^{(n)}})\\ \widehat{V}=\frac{1}{N}\sum_{n=1}^{N}\mathcal{M}^2(\boldsymbol{x^{(n)}})-\widehat{V}_0^2\\ \widehat{V}_i=\frac{1}{N}\sum_{n=1}^{N}\mathcal{M}(x_i^{(n)},\boldsymbol{x}_{\sim i}^{(n)})\mathcal{M}_{kr}(x_i^{(n)},\boldsymbol{x'}_{\sim i}^{(n)})-\widehat{V}_0^2\end{cases} \quad (3)$$

where $\boldsymbol{x}_{\sim i}^{(n)}$ denotes the collocation points with ith input variable excluded; $\boldsymbol{x'}$ is another sample of uncertain space that is independent with $\boldsymbol{x}$. To improve the computational efficiency of calculating (3), the PCE and GP-based surrogate models $\mathcal{M}_{pce}$ and $\mathcal{M}_{krg}$ can be developed. But they have curse of dimensionality issue when the number of uncertain inputs is large.

C. Deep Gaussian Process-Based GSA

GP has strong theoretical justifications in surrogate modeling. It is data-driven and does not need the physical model and probability distributions of uncertain inputs. However, the standard GP has high computational burden with a large number of uncertain inputs [6]. In this letter, we develop the DGP-based GSA to address that.

GP infers a real-valued function $f:\boldsymbol{X}\to\boldsymbol{y}$ with a set of N pairs $\boldsymbol{X}\in\mathcal{R}^{N\times D}$ and $\boldsymbol{y}\in\mathcal{R}^{N\times 1}$. Its regression form is [7]:

$$\boldsymbol{y}=f(\boldsymbol{X})+\boldsymbol{\epsilon} \quad (4)$$

where f is drawn from GP defined by the mean function and the covariance function, i.e., $f\sim\mathcal{GP}(m(\boldsymbol{X}),k(\boldsymbol{X},\boldsymbol{X}))$, and $\boldsymbol{\epsilon}\sim\mathcal{N}(\boldsymbol{0},\sigma_\epsilon^2\boldsymbol{I})$. The inference of GP is computationally expensive with $\mathcal{O}(N^3)$ complexity. To this end, variational compression approach is utilized to capture the information of the original data. A set of M inducing points $\boldsymbol{Z}=\{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_M\}$ jointly distributed as a GP with the original data is used for compression, yielding the following joint density

$$p(\boldsymbol{y},\boldsymbol{f},\boldsymbol{u})=p(\boldsymbol{y}|\boldsymbol{f})p(\boldsymbol{f}|\boldsymbol{u})p(\boldsymbol{u}) \quad (5)$$

where $\boldsymbol{f}$ and $\boldsymbol{u}$ are the function values $f(\boldsymbol{X})$ and $f(Z)$, respectively. In (5), the conditional $p(\boldsymbol{f}|\boldsymbol{u})$ can be expressed as $\mathcal{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$, where $\boldsymbol{\mu}=K_{\boldsymbol{XZ}}K_{\boldsymbol{ZZ}}^{-1}\boldsymbol{u}$, $\boldsymbol{\Sigma}=K_{\boldsymbol{XX}}-K_{\boldsymbol{XZ}}K_{\boldsymbol{ZZ}}^{-1}K_{\boldsymbol{XZ}}^T$ with K_{XZ} denotes the covariance matrix with terms $[K_{XZ}]_{ij}=k(X_i,Z_j)$. By minimizing the Kullback-Leibler divergence between the variational and the true posteriors, the joint posterior is approximated with the variational posterior $q(\boldsymbol{f},\boldsymbol{u})=p(\boldsymbol{f}|\boldsymbol{u})q(\boldsymbol{u})$, where $q(\boldsymbol{u})=\mathcal{N}(\boldsymbol{u}|\boldsymbol{m},\boldsymbol{S})$, allowing for the closed-form marginal

$$q(\boldsymbol{f}|\boldsymbol{m},\boldsymbol{S})=\int p(\boldsymbol{f}|\boldsymbol{u})q(\boldsymbol{u})\,d\boldsymbol{u}=\mathcal{N}(\boldsymbol{f}|\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{\Sigma}}) \quad (6)$$

Similarly, the mean and covariance functions are expressed as: $\tilde{\boldsymbol{\mu}}=K_{\boldsymbol{XZ}}K_{\boldsymbol{ZZ}}^{-1}\boldsymbol{m}$, $\tilde{\boldsymbol{\Sigma}}=K_{\boldsymbol{XX}}-K_{\boldsymbol{XZ}}K_{\boldsymbol{ZZ}}^{-1}(K_{\boldsymbol{ZZ}}-\boldsymbol{S})K_{\boldsymbol{ZZ}}^{-1}K_{\boldsymbol{XZ}}^T$.

By stacking GP into a composite multilayer model, DGP can be obtained [6]:

$$\boldsymbol{y}=f_L(f_{L-1}(\cdots f_1(\boldsymbol{X})))+\boldsymbol{\epsilon} \quad (7)$$

where f_l denotes the l-th layer function with inducing inputs $\boldsymbol{Z}_l$ and corresponding outputs $\boldsymbol{u}_l$. Analogously, the joint density of multilayer Gaussian process is:

$$p(\boldsymbol{y},\{\boldsymbol{f}_l\}_{l=1}^L,\{\boldsymbol{u}_l\}_{l=1}^L)=p(\boldsymbol{y}|\boldsymbol{f}_L)\prod_{l=1}^{L}p(\boldsymbol{f}_l|\boldsymbol{u}_l)p(\boldsymbol{u_l}) \quad (8)$$

The complexity and nonlinearity introduced during GP stacking bring difficulties for handling correlations within and between layers. Doubly stochastic variational inference is used to achieve accurate approxiamtions while maintaining the conditional structure. The posterior of layer outputs is similar to single-layer GP, i.e.,

$$q(\{\boldsymbol{f}_l\}_{l=1}^L)=\prod_{l=1}^{L}\mathcal{N}(\boldsymbol{f}_l|\tilde{\boldsymbol{\mu}}_l,\tilde{\boldsymbol{\Sigma}}_l) \quad (9)$$

where $\tilde{\boldsymbol{\mu}}$ and $\tilde{\boldsymbol{\Sigma}}$ denotes the mean and covariance functions of the layer inputs. The variational parameters are determined by optimizing the lower bound on the marginal likelihood:

$$\mathcal{L}=\mathbf{E}_{q(\boldsymbol{f}_L)}[\log p(\boldsymbol{y}|\boldsymbol{f}_L)]-\sum_{l=1}^{L}KL[q(\boldsymbol{u}_l)||p(\boldsymbol{u}_l)], \quad (10)$$

where the expectation is estimated by MC sampling based on the variational posterior. The DGP model $\mathcal{M}_{dgp}$ can be obtained once the parameters are determined.

Implementation of DGP-based GSA: i) with given historical PV, load and voltage measurements, the DGP is used to build the surrogate model $\mathcal{M}_{dgp}$ via (4)-(10); ii) In (3), the SI can be computed by replacing $\mathcal{M}$ in (3) with $\mathcal{M}_{dgp}$.

Remark: GSA aims to assess the importance of all uncertain sources in a global view, which provides valuable information about how each uncertain PV affects the desired output and to what degree the influence is. GSA can characterize the joint impacts of PV injections on the model outputs via the global voltage to real and reactive power sensitivity indices. It can also benefit other applications, such as importance ranking of uncertain sources [1], overall model sensitivity estimation [4], network clustering [8], distribution network management [9], etc. This work focuses on the impacts of utility scale PVs and the aggregated PVs, i.e., aggregations of many low voltage side PVs on voltage sensitivity. Under these scenarios, the uncertain PV generation data are available for the utilities. For the low-voltage distribution systems, the rooftop PV generations are typically not available to utilities. In this case, it is challenging to specifically assess the impacts of rooftop PVs on the voltage sensitivity. However, the smart meters that show the net-loads are available to utilities. Our method can treat the net-loads as uncertain sources and assess the impacts of them on voltage sensitivity. Note

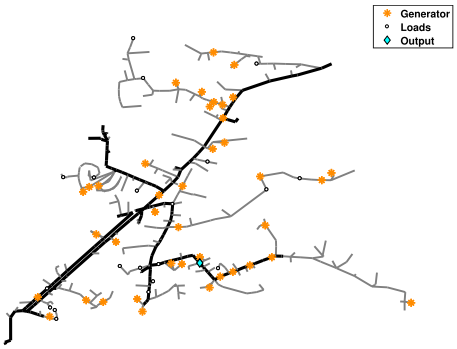

Fig. 2. Case 1: K1 system with 60 uncertain sources.

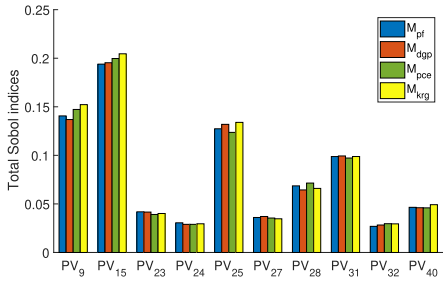

Fig. 3. Case 1: total SIs for 10 most significant inputs of each method.

that we can also aggregate several customers with rooftop PVs and assess their impacts on voltage sensitivity.

III. Numerical Results

The proposed method is tested on the EPRI K1 feeder, which contains 1747 nodes and 320 loads [10]. The uncertain inputs include flexible loads and PVs while the model responses are node voltage magnitudes. Three surrogate models are examined and compared with the benchmark obtained from MC simulations using the OpenDSS software. The mean absolute percentage error is used to quantify the accuracies of surrogate model prediction and total SI, i.e., e_M and e_{SI}^{ave}. Since most SIs are close to zero, only the 10 significant nodes with large total SIs are shown due to space limitation. The average total SI error e_{SI}^{ave} is also provided. All simulations are conducted in MATLAB on a PC with with 2.60 GHz Intel Core i7-6700HQ and 6 GB Nvidia GTX 1060. The AC power flow model $\mathcal{M}_{pf}$ in OpenDSS is viewed as the original model, on which the corresponding benchmark SIs are calculated via MC simulations. Three surrogate models, $\{\mathcal{M}_{pce}, \mathcal{M}_{krg}, \mathcal{M}_{dgp}\}$, are constructed using PCE, Kriging, DGP, respectively.

Two cases are investigated: 1) Case 1: 60 uncertain sources, i.e., 20 loads and 40 PVs as shown in Fig. 2 and 2) Case 2: 100 uncertain sources, i.e., 40 loads and 60 PVs as illustrated in Fig. 4. The settings of loads and PVs are following [10]. For PCE, an adaptive sparse scheme is utilized for high-dimensional inputs [11]. The degree of PCE is $n=2:4$ and the number of used historical measurements is 2000. For Kriging with Gaussian kernel, 500 historical data are sufficient for these two cases. More data leads to limited improvement on accuracy but affects the computational efficiency. The parameters of DGP are similar to those of Kriging except that DGP has 2 layers and 150 inducing points.

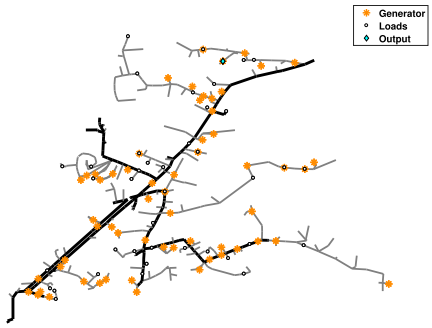

Fig. 4. Case 2: K1 system with 100 uncertain sources.

TABLE I
Tests With Different Kernel Functions

Kernel	Accuracy	
	$e_M(\times 10^{-2}\%)$	$e_{SI}^{ave}(\%)$
Gaussian	0.23	2.84
Exponential	0.34	3.04
Matern 3-2	0.23	2.93
Matern 5-2	0.23	2.88

TABLE II
Comparison Results of Different Methods on K1 System With 60 Uncertain Sources

Model	Accuracy		CPU time (s)
	$e_M(\times 10^{-2}\%)$	$e_{SI}^{ave}(\%)$	
$\mathcal{M}_{pf}$	—	—	6147.38
$\mathcal{M}_{pce}$	0.46	4.33	182.19
$\mathcal{M}_{krg}$	0.49	3.72	76.99
$\mathcal{M}_{dgp}$	0.23	2.84	71.43

In the literature, the most widely used kernel functions are Gaussian, exponential, and Matern. The hyperparameters of the kernels are calculated using the maximum likelihood estimator. To justify the choice of Gaussian kernel function in this letter, we have tested the performance by using different kernel functions and the results are shown in Table I. In general, the differences between these kernel functions are minor but the Gaussian kernel performs best.

The test results for Case 1 are shown in Fig. 3 and Table II. It can be found that with 60 uncertain inputs, all three methods can achieve good surrogate model prediction accuracy as well as SIs. Kringing and the proposed DGP-based methods have much higher computational efficiency than the MC and PCE-based methods. DGP-based method slightly outperforms the Kringing in terms of calculating accuracy of SIs. By contrast, the results for Case 2 with uncertain inputs increased from 60 to 100 are demonstrated in Fig. 5 and Table III. It can be observed that the performances of the PCE- and Kringing-based approaches have been significantly affected. Their average total SI estimation errors are 19.32% and 16.36%, respectively. The proposed method only has 4.11% estimation error while being much more computationally efficient than all other methods. This case clearly demonstrates the curse of dimensionality issues of the PCE and Kringing-based approaches. Our proposed method can deal with that and thus more suitable for practical large-scale distribution system GSA with a high penetration level of uncertain sources.

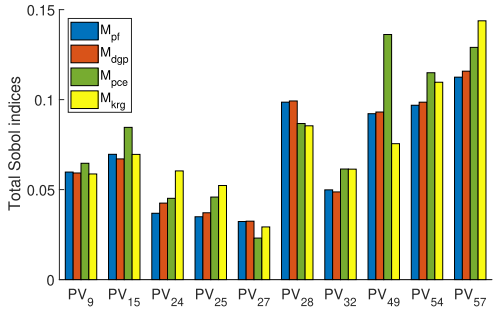

Fig. 5. Case 2: total SIs for 10 most significant inputs of each method.

TABLE III
COMPARISON RESULTS OF DIFFERENT METHODS ON K1 SYSTEM WITH 100 UNCERTAIN SOURCES

Model	Accuracy		CPU time (s)
	$e_M(\times 10^{-2}\%)$	$e_{SI}^{ave}(\%)$	
$\mathcal{M}_{pf}$	—	—	11361.61
$\mathcal{M}_{pce}$	3.26	19.32	509.20
$\mathcal{M}_{krg}$	2.41	16.36	382.14
$\mathcal{M}_{dgp}$	0.31	4.11	188.73

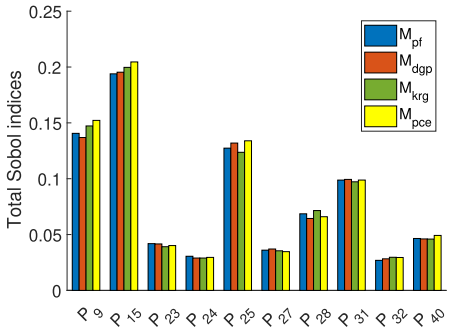

Fig. 6. Total SIs of 10 most significant inputs with Beta distribution of PVs.

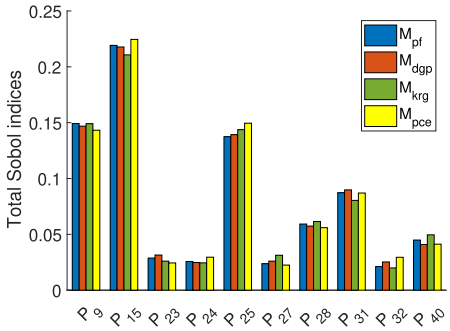

Fig. 7. Total SIs of 10 most significant inputs with Gaussian distribution of PVs.

To investigate the impacts of different distributions of PVs on SIs, the Beta, Gaussian, and uniform distributions for PVs are tested. The results are shown in Figs. 6–8 and it can be found the proposed method is able to achieve similar outcomes with the benchmark Monte Carlo method, demonstrating its high accuracy. From the theoretical point of review, with the accumulation of data, especially the adoption of deep layers, the data distribution will tend to be Gaussian according to the central limit theorem. This justifies why

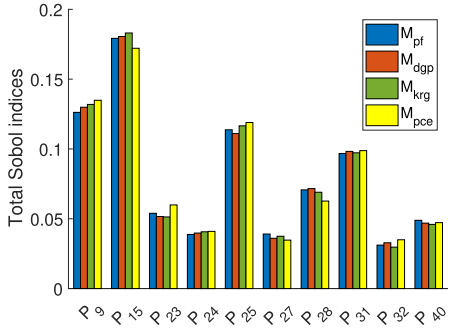

Fig. 8. Total SIs of 10 most significant inputs with Uniform distribution of PVs.

the performance of DGP is only slightly affected by non-Gaussian distributions of uncertain inputs.

IV. CONCLUSION

This paper proposes a GSA framework for large-scale three-phase distribution systems with uncertain flexible loads and PVs. The framework consists of surrogate modeling and Monte Carlo-based Sobol indices calculation. In particular, the data-driven DGP is developed to deal with the curse of dimensionality issues faced by PCE and Kringing-based approaches. Test results on the large-scale EPRI K1 system with different number of uncertainty sources, different level of uncertain inputs, and different PV distributions demonstrate that our proposed DGP-based GSA approach can significantly outperform other approaches.

REFERENCES

[1] I. Dzafic, R. Jabr, E. Halilovic, and B. Pal, "A sensitivity approach to model local voltage controllers in distribution networks," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1419–1428, May 2014.
[2] Z. Zhang, L. Ochoa, and G. Valverde, "A novel voltage sensitivity approach for the decentralized control of DG plants," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1566–1576, Mar. 2018.
[3] F. Ni, M. Nijhuis, P. H. Nguyen, and J. F. G. Cobben, "Variance-based global sensitivity analysis for power systems," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1670–1682, Mar. 2018.
[4] X. Xu *et al.*, "Maximum loadability of islanded microgrids with renewable energy generation," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 4696–4705, Sep. 2019.
[5] K. Ye, J. B. Zhao, C. Huang, N. Duan, Y. Zhang, and T. Field, "A data-driven global sensitivity analysis framework for three-phase distribution system with PVs," *IEEE Trans. Power Syst.*, to be published, doi: 10.1109/TPWRS.2021.3069009.
[6] H. Salimbeni and M. Deisenroth, "Doubly stochastic variational inference for deep Gaussian processes," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 4588–4599.
[7] A. Damianou and N. Lawrence, "Deep Gaussian processes," in *Proc. Artif. Intell. Statist.*, Stellenbosch, South Africa, 2013, pp. 207–215.
[8] H. Ruan, H. Gao, Y. Liu, L. Wang, and J. Liu, "Distributed voltage control in active distribution network considering renewable energy: A novel network partitioning method," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4220–4231, Nov. 2020.
[9] S. Nowak, Y. C. Chen, and L. Wang, "Measurement-based optimal DER dispatch with a recursively estimated sensitivity model," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4792–4802, Nov. 2020.
[10] F. Ding, Y. Zhang, J. Simpson, A. Bernstein, and S. Vadari, "Optimal energy dispatch of distributed PVs for the next generation of distribution management systems," *IEEE Open Access J. Power Energy*, vol. 7, pp. 287–295, 2020, doi: 10.1109/OAJPE.2020.3009684.
[11] X. Sun *et al.*, "Probabilistic load flow calculation based on sparse polynomial chaos expansion," *IET Gener. Transmiss. Distrib.*, vol. 12, no. 11, pp. 2735–2744, Jun. 2018.