arXiv:2207.10781v1 [stat.ML] 21 Jul 2022

Data-Driven Stochastic AC-OPF using Gaussian Processes

Mile Mitrovic , Aleksandr Lukashevich , Petr Vorobev , Vladimir Terzija , Semen Budenny , Yury Maximov , Deepjoyti Deka

Abstract

In recent years, electricity generation has been responsible for more than a quarter of the greenhouse gas emissions in the US. Integrating a significant amount of renewables into a power grid is probably the most accessible way to reduce carbon emissions from power grids and slow down climate change. Unfortunately, the most accessible renewable power sources, such as wind and solar, are highly fluctuating and thus bring a lot of uncertainty to power grid operations and challenge existing optimization and control policies. The chance-constrained alternating current (AC) optimal power flow (OPF) framework finds the minimum cost generation dispatch maintaining the power grid operations within security limits with a prescribed probability. Unfortunately, the AC-OPF problem's chance-constrained extension is non-convex, computationally challenging, and requires knowledge of system parameters and additional assumptions on the behavior of renewable distribution. Known linear and convex approximations to the above problems, though tractable, are too conservative for operational practice and do not consider uncertainty in system parameters. This paper presents an alternative data-driven approach based on Gaussian process (GP) regression to close this gap. The GP approach learns a simple yet non-convex data-driven approximation to the AC power flow equations that can incorporate uncertainty inputs. The latter is then used to determine the solution of CC-OPF efficiently, by accounting for both input and parameter uncertainty. The practical efficiency of the proposed approach using different approximations for GP-uncertainty propagation is illustrated over numerous IEEE test cases.

Preprint submitted to IJEPES *July 29, 2022*

Keywords: optimal power flow, chance constraints, Gaussian processes.
PACS: 0000, 1111
2000 MSC: 0000, 1111

1. Introduction

Further development of new technologies and the growing need for renewable energy sources (RES) bring significant benefits in reducing the cost of energy production, and in protecting the environment. Despite these benefits, the increase in RES generation such as wind and solar lead to significant challenges to the Transmission System Operators (TSOs) [1]. Instability of RES generation is probably the most important of the above challenges as it substantially compromises grid security and may lead to a power blackout. To this end, grid optimization and control algorithms that account for generation uncertainty are needed.

This paper focuses on the Chance-Constrained Optimal Power Flow (CC-OPF) [2, 3, 4, 5, 6] problem that minimizes energy generation cost and provides optimal power dispatch in response to changing demand and operational conditions of a grid. Further, the operational grid conditions such as voltage or transmission capacity limits are satisfied with a high probability thus guaranteeing grid immunity against a sudden change in RES generation. The chance-constrained (CC) OPF remains a hard grid optimization problem both in terms of solution accuracy and computational complexity. Monte Carlo scenario-based CC-OPF [7, 8], with nonlinear AC-PF equations, does not computationally scale for large systems or a large number of scenarios [9]. Consequently, advanced sampling policies (importance sampling, active sampling) have been recommended to reduce the number of samples necessary [10, 11, 9, 12, 13]. In contrast, analytical approach to CC-OPF states the chance constraints using distributional information of the uncertainty and is the focus of this work. A substantial number of papers is devoted to designing convex/linear approximations to the chance-constraints problem [14, 15, 16, 17, 18]; however, most of these formulations are conservative and hence are unable to address a large change in power generation. Another line of research is focused on the nonlinear Polynomial chaos expansion for modeling uncertainties in power demand and generation [19, 20]. Although this approach often gives acceptable accuracy, its solubility lacks good computational performance. Furthermore, most prior work requires knowledge

of the power flow model parameters that may not be accurate due to insufficient and irregular equipment calibration. Moving forward to the realm of micro-grids and low-voltage grids, such parametric uncertainties will be major due to reduced real-time metering.

This paper proposes a flexible alternative to the traditional approaches for solving the CC-OPF problem that balances the computational complexity and accuracy of the solution. In particular, having empirical samples of RES and load fluctuations, we propose to learn a Gaussian process-based approximation to the AC power flow balance equations and use them inside CC-OPF. Gaussian process (GP) is a well-known model-free machine learning algorithm that can fit well any smooth nonlinear function [21] and yields a suitable complexity-accuracy tradeoff [22, 23]. GP has a competitive advantage against traditional approximation algorithms that are focused on fitting the power balance equations around the current state and, therefore, fail to capture non-linearity from large fluctuations. In prior work, GP and neural network-based power-flow models have been implemented in deterministic OPF formulations [24, 25, 26, 27]; however, such models do not consider CC-OPF with input uncertainty. Neural network-based stochastic AC-OPF has been proposed in [28], but considers scenario-based chance constraints. In contrast, our GP-based formulation exploits uncertainty propagation in Gaussian Processes [29]. It includes analytical CC-OPF constraints that account for both parametric, i.e., line impedances, and input, i.e., RES, loads, uncertainties.

Our case studies over various IEEE test cases demonstrate the better or comparable performance of Gaussian processes based stochastic AC-OPF over the scenario and model-based CC-OPF approaches, despite not having access to network parameters.

The paper is organized as follows. Section 2 introduces the problem, summarizes related works and paper notation. Section 3 outlines Gaussian process background. The algorithm for solving the chance constrained AC-OPF problem is introduced in Section 4. The results of the case study are given in Section 6. Sections 6.6 and 7 contain a critical discussion of the approach, its possible extensions and applications respectively.

2. Problem Setup and Notation

This section covers the notation used throughout the paper and introduces the setup of the chance-constrained AC optimal power flow problem

considered in this paper.

α	participation factor	$g(\cdot) = 0$	AC power flow equations
w	random independent variable (fluctuation)	Ω	total active power imbalance
n_u	number of controllable generators	n_L	number of loads
n_R	number of renewables	γ	power ratio
$p(\cdot), E[\cdot]$	probability, expectation	$\texttt{Prob}(\cdot)$	chance (probability) constraint
ε	acceptable violation probability	ϱ	grid loss percentage
$\mathcal{B}$	set of buses	ψ	uniform distribution
$\mathcal{E}$	set of lines	x	vector of input features
G_{bk}/B_{bk}	real/imaginary part of admittance of line bk	$\mathcal{N}(\mu, \Sigma)$	Gaussian distribution
p/q	active/reactive power injections	$diag(\cdot)$	diagonal matrix
v/θ	voltage magnitudes/phase angels	$tr(\cdot)$	trace of a square matrix
s	apparent power flow	k	squared exponential kernel
$\mathcal{B}_c$	set of controllable generators	K	Gram matrix
$c2, c1, c0$	cost coefficients	I	identity matrix
d	vector of random independent variables	$m(\cdot)$	prior mean
u	vector of controllable variables	r	quantile function of $\mathcal{N}(0, 1)$
y	vector of output dependent variables	N	number of training samples
η/ν	log-normal distributions of renewables		

Table 1: Paper notation

2.1. Notation

For the given matrix X, X_{ij} denotes element ij of a matrix X, and $[x]_i$, $[x]_{:,j}$ its i-th row and j-th column, respectively. With $[x]_{*,j}$, we denote $*$ element of j-th column vector. Further, we will use $\cdot_*$ subscript to denote unseen test data. For the given vector x, $diag(x)$ is used to refer to a diagonal matrix with entries of vector x. The probability density of x is denotes as $p(x)$, $p(x|y)$ representing the probability of x for given y. $\frac{\partial \cdot (x_*)}{\partial x_*}$ is partial derivative of function $\cdot$ evaluated at x_*. The power grid related notation is summarized in Table 1.

2.2. AC-OPF

Let $\mathcal{G} = (\mathcal{B}, \mathcal{E})$ be a transmission power grid given by its buses $b \in \mathcal{B}$ and lines $e_{bk} \in \mathcal{E}$[1]. The alternating current power flow (PF) equations are then

$$p_{bk} = v_b v_k \left(G_{bk} \cos\theta_{bk} + B_{bk} \sin\theta_{bk}\right) \tag{1a}$$

$$q_{bk} = v_b v_k \left(G_{bk} \sin\theta_{bk} - B_{bk} \cos\theta_{bk}\right) \tag{1b}$$

[1]We also use $b \sim k$ to represent that b, k share an edge.

where p_{bk} and q_{bk} are the active and reactive power flows from bus b to bus k along e_{bk} line; $\theta_{bk} := \theta_b - \theta_k$ is a phase angle difference between adjacent buses k and b, $k \sim b$; G_{bk} and B_{bk} are the the real and imaginary parts of the admittance on the line between b and k. The system is balanced when the power flows leaving each bus is equal to the network injection at that bus:

$$p_b = v_b^2 G_{bb} + \sum_{k \sim b} p_{bk} \tag{2a}$$

$$q_b = -v_b^2 B_{bb} + \sum_{k \sim b} q_{bk}, \tag{2b}$$

where the first term represents nodal shunt elements; p_b and q_b are active and reactive power injections at bus b. To simplify the notation we refer the PF Eqs. (1),(2) as $g(p, q, v, \theta) = 0$.

The AC-OPF problem minimizes the cost of generation while satisfying the load and system constraints, as formulated below.

$$\min_{p_g} f(p_g) \tag{3a}$$

$$\text{s.t. } g(p, q, v, \theta) = 0, \tag{3b}$$

$$p_b^{\min} \le p_b \le p_b^{\max}, \quad \forall b \in \mathcal{B}_c \tag{3c}$$

$$q_b^{\min} \le q_b \le q_b^{\max}, \quad \forall b \in \mathcal{B}_c \tag{3d}$$

$$v_b^{\min} \le v_b \le v_b^{\max}, \quad \forall j \in \mathcal{B} \tag{3e}$$

$$s_{bk}(w) \le s_{bk}^{\max}, \quad \forall b \sim k \tag{3f}$$

where p_g is the controllable power generations. The voltage at generators is fixed at their prescribed values. $\mathcal{B}_c$ is a set of controllable buses, and $s_{bk} := \sqrt{p_{bk}^2 + q_{bk}^2}$ is a value of apparent power on the line e_{bk} that is limited from above by $s_{bk}^{\max}$ to prevent extreme currents and line overheating. Problem (3) can be equally stated as:

$$\min_{u} f(u) \tag{4a}$$

$$\text{s.t. } g(y, u, d) = 0 \tag{4b}$$

$$h^{min} \le h(y, u, d) \le h^{max} \tag{4c}$$

where control variable $u = p_g$ is bounded in a range $u \in [u^{\min}; u^{\max}]$; d is a vector of input variables (active power on loads p_l, q_l and RES p_{rs}, q_{rs}). The reactive power at loads q_l and RES q_{rs} are implicitly included under a

constant power factor $cos(\phi)$ assumption, i.e., $q_l = \gamma p_l$, $q_{rs} = \gamma p_{rs}$, where power ratio $\gamma = \frac{\sin(\phi)}{\cos(\phi)}$.

Let y be a vector of (dependent) output variables given by the solution of (1),(2); and h is a set of constraints given by Ineq. (3c)–(3f). Usually, the dependent constrained variables include the voltage magnitude v at PQ (load) and slack busses, reactive power q at PV (generator) buses, and apparent power flows s on the lines.

2.3. Chance Constrained AC Optimal Power Flow

Deviations in loads and renewables are caused due to forecast errors, external fluctuations, or intra-day electricity trading. Therefore, we propose that both load $d_l = [\bar{p}_l, \bar{q}_l]$ and RES $d_{rs} = [\bar{p}_{rs}, \bar{q}_{rs}]$ as uncertain power injections. We assume that active power uncertainties are modeled as the sum of the forecast set points p_d^0 and a fluctuation $\omega \sim \mathcal{W}^2$ as,

$$p_d(\omega) = p_d^0 + \omega,$$

where ω is an independent random vector with zero mean and known standard deviation σ_ω. The reactive power fluctuations are modeled with constant power factor as mentioned before as, $q_d(\omega) = \gamma p_d(\omega) = \gamma p_d^0 + \gamma\omega$. Thus, the ratio of active and reactive power injections at each node remains unchanged during fluctuations.

Designing power dispatch $p_g(\omega)$ that is absolutely immune against any uncertainty realization leads to an infeasible or a very conservative solution with limited impact in practice. The chance-constrained formulation, which admits a small probability for each of the inequality constraints not being feasible, allows deriving practically methods having rigorous theoretical guarantees [30]. The chance-constrained AC optimal power flow problem is given

[2] W is the uncertainty distribution.

as follows for $\omega \sim W$:

$$\min_{u} \mathbb{E}_{\omega} f(u(\omega)) \tag{5a}$$

$$\text{s.t.} \quad g(y, u, d) = 0 \tag{5b}$$

$$\begin{cases} \texttt{Prob}_{\omega \sim W}(h(y,u,d)) \leq h^{\max}) \geq 1 - \varepsilon_y \\ \texttt{Prob}_{\omega \sim W}(h(y,u,d)) \geq h^{\min}) \geq 1 - \varepsilon_y \\ \texttt{Prob}_{\omega \sim W}(u(\omega) \leq u^{\max}) \geq 1 - \varepsilon_u \\ \texttt{Prob}_{\omega \sim W}(u(\omega) \geq u^{\min}) \geq 1 - \varepsilon_u \end{cases} \tag{5c}$$

where the dependence on ω for g and h is suppressed. Also, ε_y and ε_u correspond to probabilities of the constraint violation that are taken according to the measure W of uncertainty in power demands and RES generation.

2.4. Automated Generation Control

In the rest of the paper, we maintain the power balance and AC OPF security constraints by changing the active power generation only. The Automatic Generation Control (AGC) [31], which is widely accepted in modern power systems, stands for compensating the power injections mismatch $\Omega = \sum_{b=1}^{n} \omega_b$ by proportionally changing selected power generations in set $\mathcal{B}_c$ according to the vector of non-negative participation factors $\{\alpha_b\}$

$$u_b(\omega) = u_b + \alpha_b \Omega, \quad b \in \mathcal{B}_c, \quad \sum_{b \in \mathcal{B}_c} \alpha_b = 1, \tag{6}$$

The parameters $\{\alpha_b\}$ can be optimized along with the generator set points p_g^0 to define the corrective control efficiency.

2.5. Data Driven AC Power Flow Optimization Pipeline

In this paper, we use the following pipeline for solving the chance constrained AC optimal power flow problem using a data-driven approach.

In the first stage, we collect fluctuations of power demand and RES power generation and dependent variables using historical data or a simulator. We use this data and the Gaussian process regression in the second stage to provide a data-driven probabilistic approximation for the dependent variables in the power balance equations, Eqs. (1),(2). The approximation is substantially more tractable than the PF equations and leads to a simpler optimization problem. Figure 1 illustrates the above for a synthetic example.

In the third stage, we compute the probability for constraint violation (termed *uncertainty margins*) that are given by Eqs. (5c). This probability has an explicit expression for Gaussian process-based modeling with input uncertainty and enables the CC-OPF problem to be cast as a deterministic optimization. Solving the latter gives us the minimal cost of power dispatch, satisfying the probabilistic security constraints.

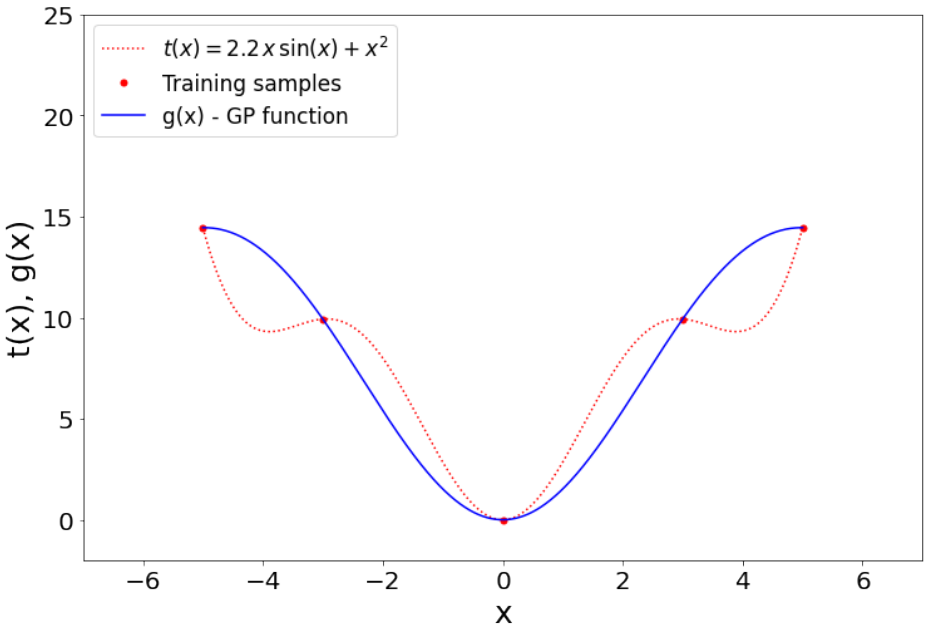


Figure 1: A synthetic example illustrates the power of the Gaussian process regression. Function $t(x) = 2.2x\sin x + x^2$ stands for the initial equation and $g(x)$ is its data-driven approximation via Gaussian process with squared exponential kernel. We refer to the approximation as a GP function.

In the next section, we introduce the data-driven Gaussian process regression framework used throughout the paper in our CC-OPF formulation.

3. Gaussian Processes

3.1. Gaussian Process Regression

Gaussian process regression (GPR) is a data-driven non-parametric regression approach that balances approximation accuracy and complexity. An attractive property of GPR is that it provides second-order statistics for the

predicted variables. We employ Gaussian process regression to infer the power flow function $g(y, u, d)$ from collected input and output measurement samples $[(x_i, y_i), i = 1, ..., N]$. To reiterate, each input data sample $x_i \in \mathbb{R}^{n_x}$ includes the vector of controllable active power generators $u = p_g$, and active power injections $p_d^\top = [p_l^\top \; p_{rs}^\top]$. Thus, $x_i^\top = [u_i^\top, p_{d_i}^\top]$. The output samples $y_i \in \mathbb{R}^{n_y}$ corresponds to all dependent variables involved in constraint (4), and $y_i^\top = [v_i^\top, q_i^\top, s_i^\top]$. For the quasi-stationary power system, the output samples can be considered as outputs from the nonlinear AC-PF equation as:

$$y_i = g\left(x_i, q_{l,i}, q_{rs,i}\right) + \zeta_i, \quad i = 1, \ldots, N \tag{7}$$

where N is total number of samples and $\zeta_i \sim \mathcal{N}(0, \Sigma_\zeta)$ is the measurement Gaussian noise/modeling error with covariance Σ_ζ.

Here we want to note that the reactive power at q_l (loads) and g_{rs} (RES) are not considered as inputs to the GP model as they are given by a fixed power factor ratio from the active power. By excluding redundant features, we reduce computational complexity without compromising the model performance.

For the given input data matrix $X = [x_i^\top]_{1 \le i \le N} \in \mathbb{R}^{N \times n_x}$, each output dimension $[y]_{:,a} \in \mathbb{R}^N$ is learnt independently. A GP [32] with prior mean function $m_a(\cdot)$ and chosen kernel $k_a(\cdot, \cdot)$ in each output dimension $a \in \{1, ..., n_y\}$, models the output $[y]_{:,a}$ as normally distributed data:

$$[y]_{:,a} \sim \mathcal{N}(m_a(X), K_a + \sigma_{\zeta,a}^2 I)$$

where $\sigma_{\zeta,a}^2$ is the covariance of measurement noise ζ in (7); I is identity matrix; $m_a(X) = [m_a(x_1), ..., m_a(x_N)]^\top$, and $[K_a]_{ij} = k_a(x_i, x_j)$ is the Gram matrix of the sampled data points. In this paper, we consider a zero prior mean $m(X) = 0$ and the squared exponential (SE) kernel with Automatic Relevance Determination (ARD):

$$k_a(x_i, x_j) = \sigma_{f,a}^2 \exp\left(-\frac{1}{2}(x_i - x_j)^\top \Lambda_a^{-1}(x_i - x_j)\right)$$

in which $\sigma_{f,a}^2$ is the signal variance and $\Lambda_a = diag[l_{a,1}^2, l_{a,2}^2, ..., l_{a,n_x}^2]$ is a positive diagonal length scale matrix.

3.1.1. Hyperparameters learning

The parameters of GP to be optimized are the aforementioned hyperparameters $\Theta = [\Lambda_a, \sigma_{f,a}^2, \sigma_{\zeta,a}^2]$ of the covariance function K_a and process noise. Note that Λ_a is a diagonal matrix. We employ the Maximum Likelihood Estimate (MLE) with the log marginal likelihood to optimize hyperparameters as follows:

$$\log\ p([y]_{:,a}|X,\Theta) = \frac{1}{2}[y]_{:,a}^T(K_a + \sigma_{\zeta,a}^2 I)[y]_{:,a} - \frac{1}{2}\log|K_a + \sigma_{\zeta,a}^2 I| - \frac{N}{2}\log 2\pi \tag{8}$$

The negative log marginal likelihood (NLL) is applied as the objective function to minimize the function value subject to the hyperparameters. The nonlinear optimization solver was employed since the NLL is a non-convex function. This paper uses Sequential Least Squares Programming (SLSQP) solver to find the optimal solution.

3.1.2. Prediction with Deterministic Input

For a given unseen sample data x_*, the predictive distribution of $[y]_{*,a}$ is provided as a marginal distribution with zero mean:

$$[y]_{*,a} \sim \mathcal{N}(0, K_{**})$$

Thus, a joint distribution of the training data and unseen point x_* is given by:

$$p([y]_{*,a}, [y]_{:,a}) \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}K_a + \sigma_{\zeta,a}^2 I & k_*\\ k_*^\top & K_{**}\end{bmatrix}\right)$$

where $k_* = k_a(x_i, x_*)$ is the vector of covariance between unseen point and training samples and $K_{**} = k_a(x_*, x_*) = 1$ is scalar. Since the joint distribution of the variables being Gaussian, the resulting distribution conditioned on the training data $p([y]_{*,a}|[y]_{:,a}) = \mathcal{N}(\mu_a(x_*), \sigma_a^2(x_*))$ is also Gaussian [32] with:

$$\mu_a(x_*) = k_*^\top(K_a + \sigma_{\zeta,a}^2 I)^{-1}[y]_{:,a} \tag{9a}$$

$$\sigma_a^2(x_*) = K_{**} - k_*^\top(K_a + \sigma_{\zeta,a}^2 I)^{-1}k_* \tag{9b}$$

Generally, the predictive distribution of multivariate output dimension y_* can be given by the equations:

$$\mu(x_*) = [\mu_1(x_*), \mu_2(x_*), ..., \mu_{n_y}(x_*)]^\top$$
$$\Sigma(x_*) = diag([\sigma_1^2(x_*), \sigma_2^2(x_*), ..., \sigma_{n_y}^2(x_*)])$$

To minimize the computational effort in inverting $(K_a + \sigma_{n,a}^2 I)$ during prediction, we pre-compute it by applying the Cholesky decomposition.

3.2. Uncertainty Propagation

So far, we have assumed that the input to the GP is deterministic and the output is Gaussian with errors due to measurement/modeling errors. For CC-OPF, we need to consider uncertainty in input forecasts propagating through the learned GP model. Following standard convention in prior work, we model all uncertain inputs have a Gaussian distribution $x_* = \mathcal{N}(\mu_{x_*}, \Sigma_{x_*})$. The distribution of the predicted uncertain output for a particular dimension a is obtained by integrating over the input distribution as follows:

$$p(y_*|\mu_{x_*}, \Sigma_{x_*}) = \int p(y_*|x_*)\, p(x_*|\mu_{x_*}, \Sigma_{x_*})\, dx_*$$

This posterior distribution cannot be computed analytically for general Kernel functions since the Gaussian input is mapped through a nonlinear function. To overcome this issue, we approximate the predictive distribution by a Gaussian and use different methods to derive its statistics, thus propagating the uncertainty. In the following, we gave a brief overview of Taylor Approximation (TA) and Exact Moment Matching (EM) that we consider in our setting. For details we refer the reader to [33, 34].

3.2.1. Taylor Approximation

We approximate output mean $\mu(x_*)$ by its first-order Taylor expansion and output variance $\sigma^2(x_*)$ by its 1st and 2nd order Taylor expansion around μ_{x_*}. With the first-order Taylor expansion, the mean prediction at a random input x_* does not provide any correction over the zero-order, as in (9a):

$$\mu_{TA}(x_*) = k_*^\top(\mu_{x_*})(K + \sigma_n^2 I)^{-1} y \tag{10}$$

For output variance, we consider the case with the first order Taylor approximation (TA1) as follows:

$$\sigma_{TA1}^2(x_*) = \sigma^2(\mu_{x_*}) + \frac{\partial \mu_{TA}(\mu_{x_*})}{\partial \mu_{x_*}}^\top \Sigma_{x_*} \frac{\partial \mu_{TA}(\mu_{x_*})}{\partial \mu_{x_*}}$$

as well as the second order Taylor approximation (TA2):

$$\sigma^2_{TA2}(x_*) = \sigma^2_{TA1}(x_*) + \frac{1}{2} tr\left(\frac{\partial^2 \sigma^2(\mu_{x_*})}{\partial \mu_{x_*} \partial \mu_{x_*}^\top} \Sigma_{x_*}\right)$$

where $\sigma^2(\mu_{x_*})$ corresponds to (9b).

3.2.2. Exact Moment Matching

According to [34], for a zero prior mean function and SE covariance kernel with ARD, the predicted mean and variance could be analytically computed by the law of iterated expectations (Fubini's theorem). Thus, the predicted mean is given by:

$$\mu_{EM}(x_*) = k_*^\top(\mu_{x_*}, \Sigma_{x_*})(K + \sigma_n^2 I)^{-1} y \tag{11}$$

where the kernel $k_*^\top$ is simultaneously a function of input mean and covariance:

$$k_*^\top(\mu_{x_*}, \Sigma_{x_*}) = \sigma_f^2 |\Sigma_* \Lambda^{-1} + I|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x - \mu_*)^\top (\Sigma_* + \Lambda)^{-1}(x - \mu_*)\right\}$$

The variance of the predictive distribution is obtained as follows:

$$\sigma^2_{EM}(x_*) = \sigma_f^2 - tr\left((K + \sigma_n^2 I)^{-1} Q\right) + \beta^\top Q \beta - \mu_{x_*}^2$$

in which $\beta = (K + \sigma_n^2 I)^{-1} y$ and elements of $Q \in \mathbb{R}^{N \times N}$ are computed as:

$$[Q]_{ij} = \frac{k_*(x_i, \mu_{x_*}) k_*(x_j, \mu_{x_*})}{|2\Sigma_{x_*} \Lambda^{-1} + I|^{\frac{1}{2}}} \cdot \exp\left((z_{ij} - \mu_{x_*})^\top \left(\Sigma_{x_*} + \frac{1}{2}\Lambda\right)^{-1} \Sigma_{x_*} \Lambda^{-1} (z_{ij} - \mu_{x_*})\right)$$

with $z_{ij} = \frac{1}{2}(x_i + x_j)$.

4. Algorithm

4.1. GP CC-OPF Formulation

The model given by Eq. (5) is reformulated as a nonlinear programming (NLP) problem with learned GPs in the constraints, such that an optimization solver can handle it.

Following [14], we reformulate the objective function as:

$$E_{\omega}[f(\cdot)] = \sum_{k \in n_u} \left\{ c_{2k}(u_k^2 + tr(\Sigma_w)\alpha_k^2) + c_{1k}u_k + c_{0k} \right\} \tag{12}$$

where $\{c_{2k}, c_{1k}, c_{0k}\}_{k=1}^{n_u} \geq 0$ are scalar cost coefficients and α is the participation factor. Note that α (participation factor) is not a standard (fixed) value, but a variables to be optimized with constraints $\sum_{k=1}^{n_u} \alpha_k = 1$ and $\alpha_k \geq 0$. Accordingly, the objective function is a convex quadratic function of u and α.

Reformulating chance constraints to the tractable form implies reformulating both output and decision input constraints. According to [29], a GP individual output probabilities, could be expressed in terms of the mean μ_y and variance σ_y^2:

$$\begin{cases} \texttt{Prob}(\mu_{y_z} + \sqrt{\sigma_{y_z}^2} \leq y_z^{\max}) \leq 1 - \varepsilon_{y_z} \\ \texttt{Prob}(\mu_{y_z} - \sqrt{\sigma_{y_z}^2} \geq y_z^{\min}) \leq 1 - \varepsilon_{y_z} \end{cases} \tag{13}$$

that is equivalent to:

$$y_z^{\min} + r_{y_z}\sqrt{\sigma_{y_z}^2} \leq \mu_{y_z} \leq y_z^{\max} - r_{y_z}\sqrt{\sigma_{y_z}^2} \tag{14}$$

where r_{y_z} represents the quantile function $\Phi^{-1}(1-\varepsilon_{y_z})$ of the standard normal distribution and $z = 1, ..., n_y$.

Similarly, the decision input constraints are [14]:

$$\begin{cases} \texttt{Prob}(u_k + \alpha_k\, tr(\Sigma_w) \leq u_k^{max}) \leq 1 - \varepsilon_{u_k} \\ \texttt{Prob}(u_k - \alpha_k\, tr(\Sigma_w) \geq u_k^{min}) \leq 1 - \varepsilon_{u_k} \end{cases}$$

As in (13), this formulation is equivalent to:

$$u_k^{min} + r_{u_k}\alpha_k\, tr(\Sigma_w) \leq u_k \leq u_k^{max} - r_{u_k}\alpha_k\, tr(\Sigma_w) \tag{15}$$

where $r_{u_k} = \Phi^{-1}(1 - \varepsilon_{u_k})$ and $k = 1, ..., n_u$.

By taking into account of the formulations in (12), (14), and (15), the

following approximate GP CC-OPF is derived:

$$\min_{u,\alpha,\mu_y,\sigma_y} \sum_{k\in n_u} \{c_{2k}(u_k^2 + tr(\Sigma_w)\alpha_k^2) + c_{1k}u_k + c_{0k}\} \tag{16a}$$

$$\text{s.t.} \sum_{k\in n_u} \alpha_k = 1,\ \alpha_k \geq 0 \tag{16b}$$

$$\sum_{k\in n_u} u_k = \sum_{i\in n_L} p_{l_i} - \sum_{j\in n_R} p_{rs_j} \tag{16c}$$

$$\mu_y = \mu(x_*) \tag{16d}$$

$$\sigma_y^2 = \sigma^2(x_*) \tag{16e}$$

$$y_z^{min} + \lambda_{y_z} \leq \mu_{y_z} \leq y_z^{max} - \lambda_{y_z} \tag{16f}$$

$$u_k^{min} + \lambda_{u_k} \leq u_k \leq u_k^{max} - \lambda_{u_k} \tag{16g}$$

where $\lambda_{y_z} = r_{y_z}\sqrt{\sigma_{y_z}^2}$ for $z \in n_y$ and $\lambda_{u_k} = r_{u_k}\alpha_k\sqrt{tr(\Sigma_w)}$ are uncertainty margins. By $\sqrt{tr(\Sigma_w)}$ the standard deviation of the total active power imbalance Ω is represented. Equation (16c) describes the balance equation in which decision variables of controllable generators have to be satisfied in the power system keeping the balance between generation and consumption, while (16d) and (16e) represent GP output mean and variance formulation. Thus, we replace the standard AC power flow equation with a data-driven method using these equations.

The proposed OPF optimization method with GP-based model results in a non-convex optimization problem which is often hard to solve. The second-order derivative information of all quantities is available for the chosen twice differentiable SE kernel and zero prior means. Thus, the problem can be solved using sequential quadratic programming or the nonlinear interior-point method. We employed a primal-dual interior-point algorithm with filter line-search. The available IPOPT [35] solver is used in non-linear optimization framework CasADi [36]. To solve this optimization problem, we assume simultaneous approaches that simultaneously treat u, α, μ_y, and σ_y^2 as optimization variables.

4.2. Input Covariance

As previously mentioned in Section 3.2, we assume that unseen input points x_* for the GP to be normally distributed with specified vector mean $\mu_{x_*} \in \mathbb{R}^{n_x}$ and covariance matrix $\Sigma_{x_*} \in \mathbb{R}^{n_x \times n_x}$. Mean values correspond to

optimized control variables and forecast uncertain parameters as:

$$\mu_{x_*} = [u^\top, p_l^\top, p_{rs}^\top]^\top$$

where $u = [p_g^1, p_g^2, ..., p_g^{n_u}]^\top$, $p_l = [p_l^1, p_l^2, ..., p_l^{n_L}]^\top$ and $p_{rs} = [p_{rs}^1, p_{rs}^2, ..., p_{rs}^{n_R}]^\top$. We need to correctly define the input covariance matrix Σ_{x_*} as u is correlated with uncertainty w in p_l and p_{rs} due to the linear feedback, see Eq. (6). This matrix is:

$$\Sigma_{x_*} = \begin{bmatrix} \Sigma_u & \Sigma_{uw} \\ \Sigma_{uw}^\top & \Sigma_w \end{bmatrix} \tag{17}$$

so that $\Sigma_u \in \mathbb{R}^{n_u \text{x} n_u}$, $\Sigma_w \in \mathbb{R}^{n_d \text{x} n_d}$ and $\Sigma_{uw} \in \mathbb{R}^{n_u \text{x} n_d}$ are covariance sub-matrices.

The covariance sub-matrix Σ_w represents uncertain variations in loads and renewable sources. We consider this matrix as diagonal:

$$\Sigma_w = \begin{bmatrix} \sigma_{l_1}^2 & ... & 0 & 0 & ... & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & ... & \sigma_{l_L}^2 & 0 & ... & 0 \\ 0 & ... & 0 & \sigma_{rs_1}^2 & ... & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & ... & 0 & 0 & ... & \sigma_{rs_R}^2 \end{bmatrix}$$

and assume that fluctuations between stochastic power injections are independent, consistent with prior work [14].

Sub-matrix Σ_u refers to controllable generator deviations caused by loads and renewable sources fluctuations keeping the balance of the system. Since controllable variables u are dependent on forecasted uncertain injections p_l and p_{rs}, the deviations of u are also dependent on fluctuations of w as follows:

$$\Sigma_u = \begin{bmatrix} \alpha_1^2 tr(\Sigma_w) & ... & \alpha_1 tr(\Sigma_w)\alpha_{n_u} \\ ... & ... & ... \\ \alpha_{n_u} tr(\Sigma_w)\alpha_1 & ... & \alpha_{n_u}^2 tr(\Sigma_w) \end{bmatrix}$$

The covariance elements between dependent decision variables and stochastic parameters are expressed in sub-matrix Σ_{uw} as:

$$\Sigma_{uw} = \begin{bmatrix} \alpha_1\sigma_{l_1}^2 & ... & \alpha_1\sigma_{l_{n_L}}^2 & \alpha_1\sigma_{rs_1}^2 & ... & \alpha_1\sigma_{rs_{n_R}}^2 \\ ... & ... & ... & ... & ... & ... \\ \alpha_{n_u}\sigma_{l_1}^2 & ... & \alpha_{n_u}\sigma_{l_{n_L}}^2 & \alpha_{n_u}\sigma_{rs_1}^2 & ... & \alpha_{n_u}\sigma_{rs_{n_R}}^2 \end{bmatrix}$$

5. Training Dataset Simulation

As our GP-based data-driven CC-OPF is model-free and uses estimated relation between input and output variables, we next discuss the creation of a synthetic dataset to learn the model and for our validation. In practice, historical data or a power grid simulator can be used to generate such data for learning.

Motivated by [37], we generated the synthetic dataset using three different parts of injection sampling, first active and reactive power loads, then renewable sources, and finally active power (controllable) generation.

5.0.1. Sampling active power loads and renewable sources

To simulate real active powers for i-th load and j-th renewable source, we define stochasticity as:

$$p_l^i = \eta^i p_{l_{ref}}^i,\ \forall 1 \leq i \leq n_L,\ p_{rs}^j = \nu^j p_{rs_{ref}}^j,\ \forall 1 \leq j \leq n_R$$

where $p_{l_{ref}}^i$ and $p_{rs_{ref}}^j$ are fixed reference values of n_L loads and n_R renewable sources; η^i and ν^j are a random variables that follow a required distribution. Further, we will consider log-normal distribution. As in [37], we separate random variables η^i and ν^j into two components:

$$\eta^i = \eta_{corr}\eta_{uncorr}^i,\ \forall 1 \leq i \leq n_L \tag{18a}$$

$$\nu^j = \nu_{corr}\nu_{uncorr}^j,\ \forall 1 \leq j \leq n_R \tag{18b}$$

where η_{corr} and ν_{corr} denotes spatio-temporal correlations, while η_{uncorr}^j and ν_{uncorr}^j are local variations. We propose that both variables have *log-normal* distribution. We consider correlated variables η_{corr} at all loads and ν_{corr} at all RES to take the same values with mean and standard deviation, $(\mu(\eta_{corr}) = -1,\ \mu(\nu_{corr}) = 0.2)$ and $(\sigma(\eta_{corr}) = 0.1,\ \sigma(\nu_{corr}) = 0.4)$. In contrast, uncorrelated variables are different for each load and renewable source, distributed with means $(\mu(\eta_{corr}) = 1,\ \mu(\nu_{corr}) = 1)$ and standard deviations $(\sigma(\eta_{corr}) = 0.05, \sigma(\nu_{corr}) = 0.3)$.

5.0.2. Sampling reactive power loads and renewable sources

Reactive powers are sampled as a fixed power ratio from the active injections as follows:

$$q_l^i = \gamma^i p_l^i,\ \forall 1 \leq i \leq n_L,\quad q_{rs}^j = \gamma^j p_{rs}^j,\ \forall 1 \leq j \leq n_R$$

The power ratio is derived form the reference values $\gamma^i = \frac{q_{l_{ref}}^i}{p_{l_{ref}}^i},\ \gamma^j = \frac{q_{rs_{ref}}^j}{q_{rs_{ref}}^j}$.

5.0.3. Sampling active power generation

Since the voltages of the generation at PV nodes are fixed to their nominal values, we only considered the sampling of the active power generation p_g. To sample generation correctly, we have to ensure that the quasi-stationary power system equation has power balance:

$$\sum_{k=1}^{n_u} p_g^k = \sum_{i=1}^{n_L} p_l^i - \sum_{j=1}^{n_R} p_{rs}^j + losses \tag{20}$$

Consequently, the total production is approximately some percentage ϱ higher than the total demand, covering losses in the power system. In our case, this percentage ϱ is calculated as the ratio of the referent values between the total production and the total load for the various systems[3]. Following [37], we first formulate the part in which each production sample has to target the total demand increased by losses percentage as:

$$\breve{p}_g^k = \psi^k \cdot \varrho \cdot \frac{\sum_{i=1}^{n_L} p_l^i}{\sum_{k=1}^{n_u} p_{g_{ref}}^k} \cdot p_{g_{ref}}^k$$

where $\psi^k \sim \mathcal{U}[0.8, 1.2]$ is the uniform distribution to ensure no deterministic behavior for the productions.

We applied the second step to ensure that the relation $\sum_{k=1}^{n_u} p_g^k + \sum_{j=1}^{n_R} p_{rs}^j = \varrho \cdot \sum_{i=1}^{n_L} p_l^i$ holds perfectly, as:

$$p_g^k = \breve{p}_g^k \cdot \varrho \cdot \frac{\sum_{i=1}^{n_L} p_l^i}{\sum_{k=1}^{n_u} \breve{p}_g^k}$$

6. Empirical Evaluation

The results presented in this paper are obtained on an Intel Core i7-5500U CPU @ 2.40GHz and 8GB of RAM. To simulate the experiment, we use our python gp-ccopf[4] framework developed for this article and pandapower [38] package to validate results.

[3]For tested IEEE 9 bus system ϱ=1.0139 that means the total losses are 1.39% of the total output. Similarly, for IEEE 39 bus system ϱ=1.0086

[4]`https://github.com/mile888/gp_cc-opf`

6.1. Case Study

We use the 9 and 39-bus IEEE Test Systems to evaluate the performance and scalability of the proposed GP CC-OPF method. Both systems are provided with pandapower [38].

IEEE 9-bus Test Systems consists of 3 generators and 3 loads on the high-voltage level of $V_n = 345kV$. We introduce two renewable source generators at buses 3 and 5 with the total forecast power output of 80 MW, which is approximately 25.4% of the total active power demand. We assume that the power ratio is 0.3. All loads and renewable sources are uncertain.

IEEE 39-bus Test Systems consists of 10 generators and 21 loads with $V_n = 345kV$. Renewable source generators are placed at buses 1, 11, 14, 21, 23, and 28 with the total forecast power output of 1260 MW, which is approximately 20.2% of the total active power demand. The assumed power ratio is 0.3. Also, we consider that all loads and renewable sources are uncertain.

The forecast errors are modeled as zero mean, multivariate Gaussian random variables with a standard deviation corresponding to 15% of the forecasted loads ($\sigma_l = 0.15\, p_l$) and 30% of the forecasted renewable sources ($\sigma_r = 0.3\, p_r$). The acceptable violation probabilities ε are set to $\varepsilon_u = 0.1\%$ and $\varepsilon_y = 2.5\%$. Since generation constraints are physically impossible to violate, we assume a very small percentage of violations (ε_u), as proposed in Swissgrid (the swiss system operator) [39] for their reserve procurement process. In contrast, output constraints are soft constraints where some violations can be tolerated (ε_y) if the magnitude and duration are not too large or removed through additional control actions such as generation redispatch [5].

6.2. Evaluation Procedure and Metrics

To evaluate the prediction quality of the *GP regression*, optimized using training data, we use the root mean squared error (RMSE) metric for each output variable a as:

$$RMSE_a = \sqrt{\frac{1}{N_*}\sum_{i=1}^{N_*}([y]_{i,a} - [\hat{y}]_{i,a})^2}$$

where y and $\hat{y}$ are the actual (AC-PF) and predicted (GP) mean values for testing data, and N_* is a number of test samples. The average evaluation

score of all output variables is calculated as follows:

$$RMSE = \frac{\sum_{a=1}^{n_y} RMSE_a}{n_y} \tag{21}$$

Solving the data-driven CC-OPF with GP-based constraints for different test data, we compute the mean of output variables $\hat{y}$ (Eq. (16d)) and report the RMSE against the variables computed by AC-PF. To analyze chance constraints feasibility, we compare the distribution of output variables, computed analytically in GP CC-OPF, with the true empirical spread computed via AC-PF with Monte-Carlo samples of uncertainty ω around the mean inputs (loads, RES, and optimized generator set-points).

To measure the performance of the solution, we also compute the computation time, generation cost, Eq. (16a)), and the empirical spread of output variables for (a) model-based scenario CC-OPF [8, 9], (b) AC-OPF on individual test samples (i.e., full recourse), and (c) AC-OPF for base case (mean loads and RES). Note that AC-OPF on the base case doesn't optimize for input uncertainty and presents the worst-case violations. On the other hand, the full-recourse solves AC-OPF for each sample separately and is the benchmark for feasibility as it is not restricted to linear generator feedback.

6.3. Models Performance

Both GP models for IEEE 9 and 39-bus systems are trained on randomly sampled data, generated as discussed in Section 5. The GP model of the IEEE 9-bus system consists of 8 inputs features and 15 outputs, while the IEEE 39 bus has 37 inputs, 74 outputs. In the case of the IEEE 9 bus case, we applied 75 random samples for training and 25 for validation. For the IEEE 39 bus system, training is conducted on 200 samples, and 65 samples are used for validation. The average RMSE metric in Eq. (21) for IEEE 9 bus and IEEE 39 bus are RMSE = $7.72e^{-5}$ $p.u.$ and RMSE = $8.91e^{-4}$ $p.u.$, respectively. The small errors suggest that the models for both systems have good generalization for deterministic inputs and that the hyperparameters are well optimized.

Next, we use the GP CC-OPF method with three approximation methods: first-order Taylor approximation (TA1), second-order Taylor approximation (TA2), and Exact moment matching (EM). The iterative solution interior-point algorithm is applied to find a feasible solution in k iterations. In our simulations, the algorithm converged within a relatively small number of iterations for the convergence tolerance of $\epsilon_{tol} = 10^{-5}$, as shown in

Table 2. Table 2 also presents each approximation method's RMSE for GP-output against AC-PF based outputs, for same inputs. The RMSE results are the same for TA1 and TA2 algorithms since both algorithms have the same mean function (10). In contrast, the EM mean function depends on input covariance; in our case, the RMSE is better. However, the CPU time of optimization is greater for EM, with a sharper increase for the larger system.

Table 2: GP CC-OPF method results for IEEE 9 and IEEE 39 bus systems

System	parameters	TA1	TA2	EM
IEEE 9	RMSE [p.u.]	$6.35 \cdot 10^{-3}$	$6.35 \cdot 10^{-3}$	$5.95 \cdot 10^{-3}$
	No. iteration	16	14	16
	CPU time [s]	0.35	4.97	6.60
IEEE 39	RMSE [p.u.]	$1.37 \cdot 10^{-2}$	$1.37 \cdot 10^{-2}$	$1.35 \cdot 10^{-2}$
	No. iteration	18	18	20
	CPU time [s]	19.38	717.6	1221.2

6.4. Chance-constraint feasibility and uncertainty margins

We plot the spread of centered CC-OPF output variables (voltages at PQ buses and reactive powers at controllable generators) for the IEEE 9-bus system in Fig. 2. The three standard deviations (3 STD) intervals for uncertainty margins are analytically derived using the estimated GP-variance, see Eq. (16e). The different approximations considered for uncertainty propagation are marked as blue (TA1), red (TA2), and black (EM). The empirical spread for the same output variables is derived using both GP-function and AC-PF equation with 5000 Monte Carlo (MC) samples of the input uncertainty ω. Brown and yellow bars represent the upper and lower margins of the MC-based GP model, while dark and light green bars are for upper and lower margins of the MC-based AC-PF equation. Due to the asymmetrical MC-based distributions owing to nonlinear AC-PF, the upper and lower uncertainty margins, computed as 99.73% and 0.27% probability, respectively (3 STD), are depicted separately.

Similar plots for nodal voltage, reactive injections, and apparent powers on lines in the 39 bus systems are presented in Fig. 3. Comparing the analytical with the Monte-Carlo based uncertainty margins, it is clear that they are close for most outputs, with a few exceptions where the analytical spread *overestimate* the MC ones.

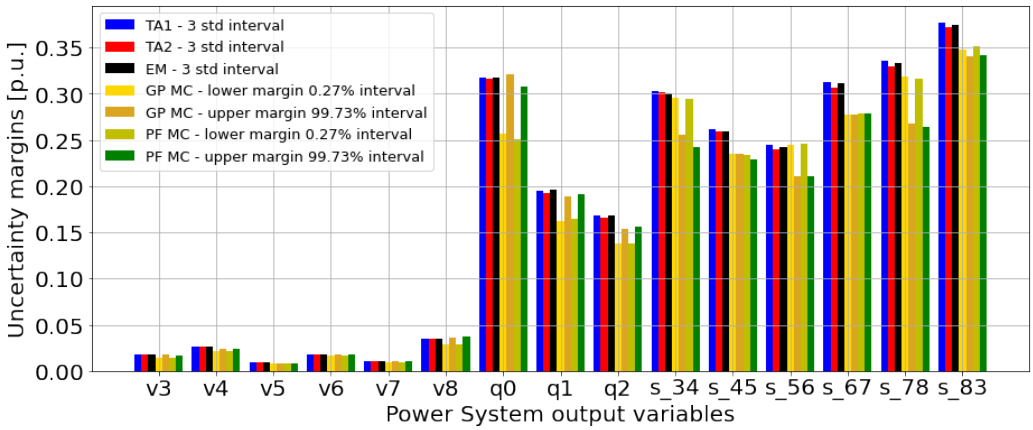


Figure 2: Output uncertainty spread of IEEE 9.

Note that as analytical margins satisfy the chance constraints, see Eq. (16f) and Eq. (16g), the smaller MC empirical violations will also be within the chance-constrained limits, and hence feasible. Moreover, from Fig. 2 and Fig. 3, MC margins for GP model and AC-PF equation are comparable, which implies that the GP model can learn the nonlinear AC-PF accurately. For the IEEE 39 bus system, we do observe small deviations between upper and/or lower margins for some variables, but this can be overcome by increasing the number of training data samples.

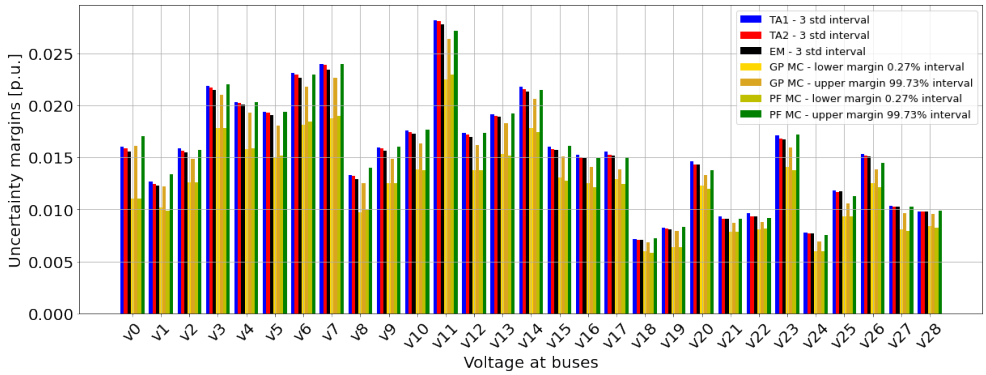
6.5. Comparison with Model-based Scenario CC-OPF

Here, we compare cost, computational time, and CC-feasibility of GP CC-OPF method (with TA1 approximation), with model-based scenario CC-OPF [8, 9], with a varying number of scenarios. The empirical constraint violation for determined set-points is computed using 1000 test MC samples. To benchmark, we also compare against two model-based setups:

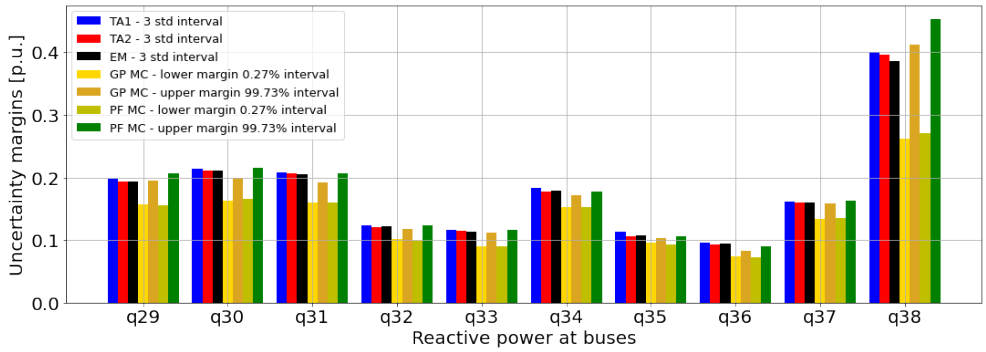
(A) AC-OPF will full recourse, i.e., separate AC-OPF for each MC sample,

(B) and AC-OPF for the base case, i.e., AC-OPF for the mean input.

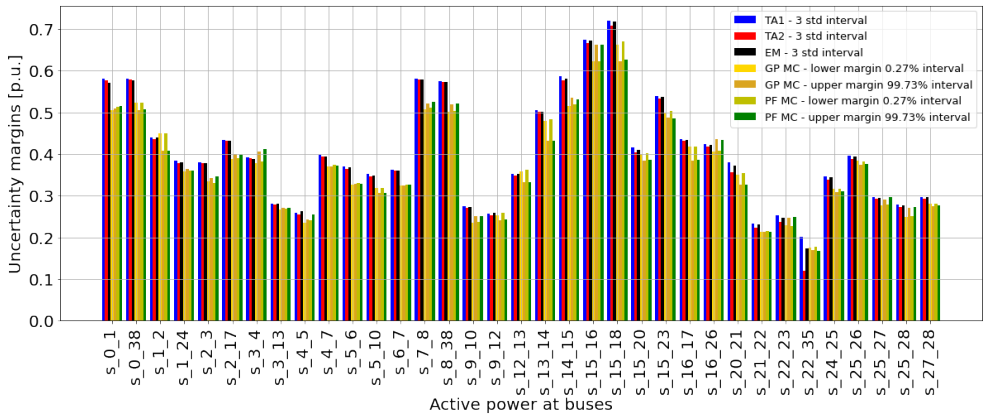
Note that though unrealistic, (A) considers generalized nonlinear feedback, while (B) doesn't optimize the participation factor α and uses equal generator participation for test MC samples. Hence they represent, respectively, model-based solutions with the best and worst feasible solutions.



(a) Voltage output distributions



(b) Reactive power output distributions

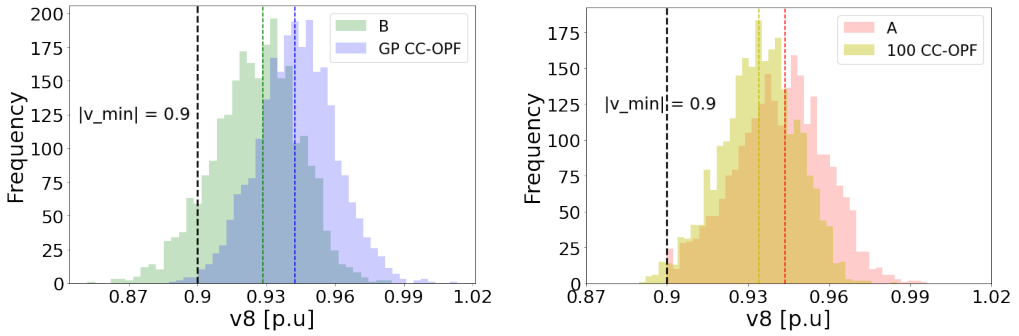


(c) Apparent power output distributions

Figure 3: Output uncertainty spread of IEEE 39.

In Table 3, we can observe that GP CC-OPF solution produces the desired reliability level of $\varepsilon_y = 2.5\%$, with a cost between that of (A) and (B) for both test cases. Note that (A) solves separate AC-OPF for each MC sample and hence has no constraint violation but higher cost. On the other hand, (B) 's set point is optimized by OPF for the mean loads and RES. Thus, it has a lower cost but large violations for test uncertainty. Their computational time is that of a single AC-OPF. For scenario CC-OPF, increasing the number of scenarios increases cost and reduces constraint violations for either test case. Although the number of scenarios necessary to get desired CC-feasibility is greater than 100 for IEEE 39, and greater than which leads to a significantly higher computational time for scenario CC-OPF over GP CC-OPF where CC are implemented analytically.

To study the empirical spread of output variables, we illustrate the voltage under uncertainty at bus 8 of IEEE9 in Fig. 4, and voltage at bus 25 of IEEE39 in Fig. 5. As one can observe, voltage values for GP CC-OPF for the IEEE9 (Fig. 4a) are centered further from the minimal voltage than the corresponding voltage values for 100 scenario CC-OPF (Fig. 4b). The latter leads to reduced violation of GP CC-OPF compared to scenario CC-OPF for the important lower voltage limit. Moreover, the latter's spread is more skewed to the constraint border. Note that (A) with full recourse has no violation while (B) for base-case has a higher violation. Similarly, for voltage at bus 25 of IEEE39, the GP CC-OPF voltage spread (Fig. 5a) is further from the important upper limit compared to the solution of scenario CC-OPF (Fig. 5b).



(a) Spread of voltages for GP CC-OPF and (B) AC-OPF for base case.

(b) Spread of voltages for scenario CC-OPF and (A) AC-OPF with full recourse.

Figure 4: Case IEEE-9. Empirical Spread of voltages at Bus 8 under uncertainty.

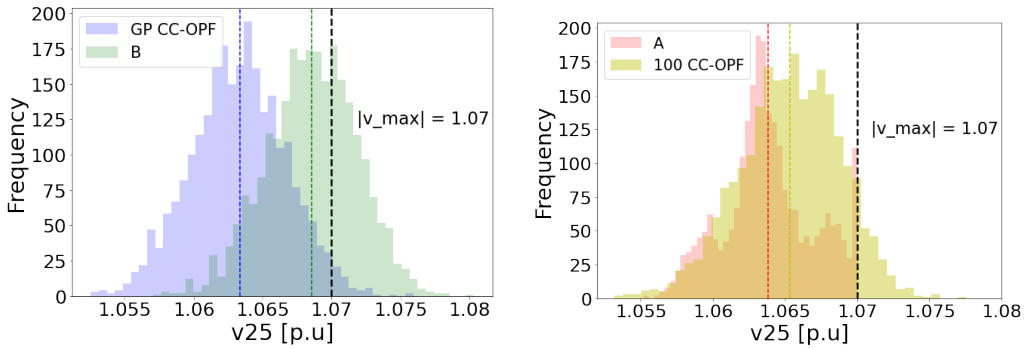
System	IEEE 39		
Parameters	Cost [$]	Infeas. Prob. [%]	Time (s)
A (full recourse)	$7.869 \cdot 10^6$	-	0.93
B (base case)	$7.533 \cdot 10^6$	35.36	0.93
GP CC-OPF	$7.752 \cdot 10^6$	2.36	19.38
50 CC-OPF	$7.642 \cdot 10^6$	14.20	96.4
100 CC-OPF	$7.696 \cdot 10^6$	8.24	184.7
200 CC-OPF	$7.813 \cdot 10^6$	0.16	505.0

System	IEEE 9		
Parameters	Cost [$]	Prob. [%]	Time (s)
A (full recourse)	$4.056 \cdot 10^3$	-	0.86
B (base-case)	$3.467 \cdot 10^3$	9.76	0.86
GP CC-OPF	$4.039 \cdot 10^3$	0.44	0.35
20 CC-OPF	$3.481 \cdot 10^3$	7.52	4.4
50 CC-OPF	$3.836 \cdot 10^3$	4.92	12.1
100 CC-OPF	$3.985 \cdot 10^3$	1.24	22.8

Table 3: Cost function values and probability of violation of a constraint at a solution

6.6. Discussion

Comparing the results of different approximation methods within GP CC-OPF, we can conclude that all three methods (TA1, TA2, and EM) approximate uncertainty margins very well with small probability violations for MC test samples for both systems (Figs. 2, 3). Table 2 shows that with the system scaling, the accuracy (RMSE) of the mean performance decreases. It can be explained by the fact that for the IEEE 39 bus system, we considered 200 training data samples randomly sampled in the space. More samples and active learning techniques can help improve the accuracy. In general, the problem of GP is scalability, where the computational complexity of training and prediction increases with the number of samples and input/output dimensions. Due to the inclusion of the second derivative term, TA2 and EM are more complicated than TA1 as confirmed by the CPU time results in Table 2. Computational complexity problems can be overcome by using sparse GP and dimensionality reduction techniques [29], which we plan to explore in the future. Generally, from the obtained results that the TA1 method gives similar performance accuracy as TA2 and EM, but with less



(a) Spread of voltages for GP CC-OPF and (B) AC-OPF for the base case.

(b) The spread of voltages for scenario CC-OPF and (A) AC-OPF with full recourse.

Figure 5: Case IEEE-39. Empirical Spread of voltages at Bus 25 under uncertainty.

complexity.

7. Conclusion

In this research, we have proposed a data-driven and model-free stochastic OPF method by replacing the standard AC power flow equation and security constraints with the Gaussian process regression model. We show that a relatively small number of samples spread correctly in the space can generalize the GP function to model the AC-PF equations and furthermore enables accurate propagation of uncertainty from input to output.

We show that applying the first Taylor Approximation method in GP CC-OPF outperforms other proposed approximation methods and two widely used sample-based approaches in our experiments. Although the accuracy of the data-driven CC-OPF is notable, it depends on the propagation space of training data. In future work, we will attempt to overcome this issue by proposing a new hybrid GP CC-OPF method that combines linear and non-linear approximation methods. Also, our further research will be aimed at improving the scalability of the proposed approach using sparse GP formulations.

References

[1] Z. Liu, L. v. d. Sluis, W. Winter, Challenges, experiences and possible solutions in transmission system operation with large wind integration,

OSTI (2012).

[2] B. Stott, O. Alsaç, Optimal power flow: Basic requirements for real-life problems and their solutions, in: SEPOPE XII Symposium, Rio de Janeiro, Brazil, Vol. 11, 2012, pp. 1–10.

[3] A. Venzke, L. Halilbašić, A. Barré, L. Roald, S. Chatzivasileiadis, Chance-constrained ac optimal power flow integrating hvdc lines and controllability, International Journal of Electrical Power & Energy Systems 116 (2020) 105522.

[4] L. Roald, F. Oldewurtel, T. Krause, G. Andersson, Analytical reformulation of security constrained optimal power flow with probabilistic constraints, in: 2013 IEEE Grenoble Conference, IEEE, 2013, pp. 1–6.

[5] L. Roald, G. Andersson, Chance-constrained AC optimal power flow: Reformulations and efficient algorithms, IEEE Transactions on Power Systems 33 (3) (2017) 2906–2918.

[6] N. Viafora, S. Delikaraoglou, P. Pinson, J. Holbøll, Chance-constrained optimal power flow with non-parametric probability distributions of dynamic line ratings, International Journal of Electrical Power & Energy Systems 114 (2020) 105389.

[7] G. C. Calafiore, M. C. Campi, The scenario approach to robust control design, IEEE Transactions on Automatic Control 51 (5) (2006) 742–753.

[8] M. Vrakopoulou, K. Margellos, J. Lygeros, G. Andersson, A probabilistic framework for reserve scheduling and security assessment of systems with high wind power penetration, IEEE Transactions on Power Systems 28 (4) (2013) 3885–3896.

[9] I. Mezghani, S. Misra, D. Deka, Stochastic ac optimal power flow: A data-driven approach, Electric Power Systems Research 189 (2020) 106567.

[10] F. Capitanescu, S. Fliscounakis, P. Panciatici, L. Wehenkel, Cautious operation planning under uncertainties, IEEE Transactions on Power Systems 27 (4) (2012) 1859–1869.

[11] A. B. Owen, Y. Maximov, M. Chertkov, Importance sampling the union of rare events with an application to power systems analysis, Electronic Journal of Statistics 13 (1) (2019) 231–254.

[12] A. Lukashevich, V. Gorchakov, P. Vorobev, D. Deka, Y. Maximov, Importance sampling approach to chance-constrained dc optimal power flow, arXiv preprint arXiv:2111.11729 (2021).

[13] A. Lukashevich, Y. Maximov, Power grid reliability estimation via adaptive importance sampling, IEEE Control Systems Letters 6 (2021) 1010–1015.

[14] D. Bienstock, M. Chertkov, S. Harnett, Chance-constrained optimal power flow: Risk-aware network control under uncertainty, Siam Review 56 (3) (2014) 461–495.

[15] M. Lubin, Y. Dvorkin, L. Roald, Chance constraints for improving the security of ac optimal power flow, IEEE Transactions on Power Systems 34 (3) (2019) 1908–1917.

[16] M. Lubin, Y. Dvorkin, S. Backhaus, A robust approach to chance constrained optimal power flow with renewable generation, IEEE Transactions on Power Systems 31 (5) (2015) 3840–3849.

[17] L. Roald, S. Misra, T. Krause, G. Andersson, Corrective control to handle forecast uncertainty: A chance constrained optimal power flow, IEEE Transactions on Power Systems 32 (2) (2016) 1626–1637.

[18] X. Du, X. Lin, Z. Peng, S. Peng, J. Tang, W. Li, Chance-constrained optimal power flow based on a linearized network model, International Journal of Electrical Power & Energy Systems 130 (2021) 106890.

[19] T. Mühlpfordt, L. Roald, V. Hagenmeyer, T. Faulwasser, S. Misra, Chance-constrained AC optimal power flow: A polynomial chaos approach, IEEE Transactions on Power Systems 34 (6) (2019).

[20] T. Mühlpfordt, T. Faulwasser, V. Hagenmeyer, Solving stochastic ac power flow via polynomial chaos expansion, in: 2016 IEEE Conference on Control Applications (CCA), IEEE, 2016, pp. 70–76.

[21] R. M. Dudley, Sample functions of the gaussian process, in: Selected Works of RM Dudley, Springer, 2010, pp. 187–224.

[22] E. Schulz, M. Speekenbrink, A. Krause, A tutorial on gaussian process regression: Modelling, exploring, and exploiting functions, Journal of Mathematical Psychology 85 (2018) 1–16.

[23] H. Liu, Y.-S. Ong, X. Shen, J. Cai, When gaussian process meets big data: A review of scalable gps, IEEE transactions on neural networks and learning systems 31 (11) (2020) 4405–4423.

[24] P. Pareek, H. D. Nguyen, Gaussian process learning-based probabilistic optimal power flow, IEEE Transactions on Power Systems 36 (1) (2020) 541–544.

[25] A. S. Zamzam, K. Baker, Learning optimal solutions for extremely fast ac optimal power flow, in: 2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), IEEE, 2020, pp. 1–6.

[26] X. Pan, T. Zhao, M. Chen, S. Zhang, Deepopf: A deep neural network approach for security-constrained dc optimal power flow, IEEE Transactions on Power Systems 36 (3) (2020) 1725–1735.

[27] M. K. Singh, V. Kekatos, G. B. Giannakis, Learning to solve the AC-OPF using sensitivity-informed deep neural networks, IEEE Transactions on Power Systems (2021).

[28] S. Gupta, S. Misra, D. Deka, V. Kekatos, Dnn-based policies for stochastic ac opf, arXiv preprint arXiv:2112.02441 (2021).

[29] L. Hewing, J. Kabzan, M. N. Zeilinger, Cautious model predictive control using gaussian process regression, IEEE Transactions on Control Systems Technology 28 (6) (2019) 2736–2743.

[30] B. Liu, B. Liu, Theory and practice of uncertain programming, Vol. 239, Springer, 2009.

[31] D. Apostolopoulou, P. W. Sauer, A. D. Domínguez-García, Automatic generation control and its implementation in real time, in: 2014 47th Hawaii International Conference on System Sciences, IEEE, 2014, pp. 2444–2452.

[32] C. E. Rasmussen, Gaussian processes in machine learning, in: Summer school on machine learning, Springer, 2003, pp. 63–71.

[33] A. Girard, C. E. Rasmussen, J. Quinonero-Candela, R. Murray-Smith, Gaussian process priors with uncertain inputs application to multiple-step ahead time series forecasting, in: Advances in neural information processing systems, Vol. 15, 2003, pp. 1–10.

[34] M. P. Deisenroth, Efficient reinforcement learning using Gaussian processes, Vol. 9, KIT Scientific Publishing, 2010.

[35] A. Wächter, L. T. Biegler, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, Mathematical programming 106 (1) (2006) 25–57.

[36] J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, M. Diehl, Casadi: a software framework for nonlinear optimization and optimal control, Mathematical Programming Computation 11 (1) (2019) 1–36.

[37] B. Donnot, Deep learning methods for predicting flows in power grids: novel architectures and algorithms, Ph.D. thesis, Université Paris Saclay (COmUE) (2019).

[38] L. Thurner, A. Scheidler, F. Schäfer, J.-H. Menke, J. Dollichon, F. Meier, S. Meinecke, M. Braun, pandapower—an open-source python tool for convenient modeling, analysis, and optimization of electric power systems, IEEE Transactions on Power Systems 33 (6) (2018) 6510–6521.

[39] F. Abbaspourtorbati, M. Zima, The swiss reserve market: Stochastic programming in practice, IEEE Transactions on Power Systems 31 (2) (2015) 1188–1194.