Physics-informed Machine Learning Method for Forecasting and Uncertainty Quantification of Partially Observed and Unobserved States in Power Grids

Ramakrishna Tipireddy
Pacific Northwest National Laboratory
Ramakrishna.Tipireddy@pnnl.gov

Alexandre Tartakovsky
Pacific Northwest National Laboratory
Alexandre.Tartakovsky@pnnl.gov

Abstract

We present a physics-informed Gaussian Process Regression (GPR) model to predict the phase angle, angular speed, and wind mechanical power from a limited number of measurements. In the traditional data-driven GPR method, the form of the Gaussian Process auto- and cross-covariance functions is assumed and its parameters are found from measurements. In the physics-informed GPR, we treat unknown variables (including wind speed and mechanical power) as a random process and compute the auto and cross-covariance functions from the resulting stochastic power grid equations. We demonstrate that the physics-informed GPR method is significantly more accurate than the standard data-driven one for immediate forecasting of generators' angular velocity and phase angle. We also show that the physics-informed GPR provides accurate predictions of the unobserved wind mechanical power, phase angle, or angular velocity when measurements from only one of these variables are available. The immediate forecast of observed variables and predictions of unobserved variables can be used for effectively managing power grids (electricity market clearing, regulation actions) and early detection of abnormal behavior and faults. The physics-based GPR forecast time horizon depends on the combination of input (wind power, load, etc.) correlation time and characteristic (relaxation) time of the power grid and can be extended to short and medium-range times.

1. Introduction

Real-time monitoring and immediate forecasting of the power grid dynamics is important for efficient operation of the power grid, including electricity market clearing and regulation actions [1]. Other applications that require forecasting include early detection of faults and instabilities and the effective operation of controllers. Despite the fact that modern power grids are heavily instrumented, directly measuring all power grid states remains impractical. Inherent high-frequency oscillations of the power grid dynamics, compounded by increasing penetration of renewable energy sources, make immediate forecasting very challenging.

Here, we present a "physics-informed" Gaussian Process Regression (GPR) method for monitoring (computing unobserved states) and immediate forecasting of power grid dynamics from measurements of (observed) states. We demonstrate that when partial (past) observations of variables are available, the proposed physics-informed GPR is significantly more accurate for forecasting than the standard "data-driven" GPR. When observations of the variables (states) of interest are not available, the physics-informed GPR still can provide equally good predictions based on measurements of other states, while the data-driven GPR becomes inapplicable.

Short-term forecasting in power grids has been addressed with various methods, including machine learning (support vector machines [2], Deep Neural Networks [3, 4, 5]) and statistical (e.g., the autoregressive integrated moving average method [6]) methods. GPR [7], another machine learning method, has been used for forecasting observed states with the model hyperparameters learned from measurements [8, 9]. In this work, we refer to this standard GRP method as the "data-driven" GPR. Challenges with the data-driven GPR include: 1) it requires a large amount of data collected at high frequency (relative to the process' correlation time), 2) the estimation of hyperparameters can be affected by the noise in measurements, and 3) it cannot predict unobserved states. Unlike some engineered and natural systems, power grid dynamics is well understood, and equations governing power grid dynamics have solid theoretical foundations. Using this characteristic of power grids, we propose a novel physics-informed GPR approach with the auto- and cross-covariance functions computed as a solution of stochastic power grid equations. We consider a grid with wind-powered generators. Because

HĭCSS

wind mechanical power is oscillating in time and deterministically unknown, it must be modeled as a correlated-in-time random process to properly describe the effect of wind fluctuations on the power grid dynamics and small signal and transient stability [10]. Random mechanical wind power transforms power grid equations into stochastic equations [11, 12]. To compute the covariance function, we solve stochastic power grid equation using the strong stability preserving Runge-Kutta (RK) scheme for different realizations of mechanical wind power [13, 12].

2. Problem description

To introduce the physics-informed GPR approach, we consider a power system with a single wind-powered generator and infinite bus, described by the swing equations [14, 11]:

$$\frac{d\theta}{dt} = \omega_B(\omega - \omega_S),$$
$$\frac{d\omega}{dt} = \frac{\omega_S}{2H}[P_m - P_e - D(\omega - \omega_S)], \qquad (1)$$

subject to the initial conditions $\theta(t = 0) = \theta_0$ and $\omega(t = 0) = \omega_0$. Here, $\theta \in [0, 2\pi)$ is the phase angle between the axis of the generator and the magnetic field, $\omega \in (-\infty, \infty)$ is the angular generator speed, H is the generator inertia, ω_B is the base speed, ω_s is the synchronization speed, and P_e is the electric power supplied by the generator to the infinite bus and is related to the phase angle θ as

$$P_e = P_{max} \sin\theta, \qquad (2)$$

where $P_{max} = \frac{EV}{X}$, E is the internal energy, V is the bus voltage, and X is the total reactance in the system. The mechanical power of wind,

$$P_m(t) = \overline{P_m} + P'_m(t), \qquad (3)$$

is modeled as a stochastic process, where $\overline{P_m}(t) > 0$ is a known deterministic function of time and P'_m is a zero mean Gaussian process with variance σ^2, correlation time λ, and the auto-covariance function

$$\overline{P'_m(t)P'_m(s)} = \sigma^2 \exp\left(-\frac{|t - s|}{\lambda}\right). \qquad (4)$$

To solve the governing equations, we model $P'_m(t)$ as an Ornstein-Uhlenbeck (O-U) process, satisfying the stochastic ordinary differential equation (ODE) [13]:

$$dP'_m = -\frac{1}{\lambda}P'_m dt + \sqrt{\frac{2}{\lambda}}\sigma dW, \qquad (5)$$

subject to the initial condition

$$P'_m(0) = P'_{m,0}, \qquad (6)$$

where W is the standard Wiener process.

The resulting equations are solved using a second-order strong stability preserving RK scheme [13] (described in Section 3). The initial condition (6) is obtained from the stationary Gaussian distribution with variance σ^2. The initial conditions for θ and ω are set to $\theta_0 = \arcsin\frac{\overline{P_m}}{P_{\max}}$ and $\omega_0 = \omega_S$.

3. Strong stability preserving Runge-Kutta scheme

To introduce the RK scheme, we define the vector $y = [\theta, \omega]^T$ and scalar $z = P'_m$ functions of time and rewrite the governing equations as

$$dy = f(y)dt + g(y)zdt,$$
$$dz = -azdt + bdW, \qquad (7)$$

where $f(y) = \begin{bmatrix} \omega_B(\omega - \omega_S)dt \\ \frac{\omega_S}{2H}[\overline{P_m} - P_{\max}\sin\theta - D(\omega - \omega_S)] \end{bmatrix}$, $g(y) = \begin{bmatrix} 0 \\ \frac{\omega_S}{2H}P'_m \end{bmatrix}$, $a = \frac{1}{\lambda}$, and $b = \sqrt{\frac{2}{\lambda}}\sigma$. The RK discretization of Eq (7) is [13]:

$$y_{k+1} = y_k + \frac{h}{2}\{(f + gz)_k + (f + gz)_{\bar{k}}\} + \frac{1}{\sqrt{12}}g_k bh^{3/2}\eta_k,$$

$$z_{k+1} = z_k + b\xi_k h^{1/2} + \frac{h}{2}a(z_k + z_{\bar{k}}) + \frac{1}{\sqrt{12}}abh^{3/2}\eta_k, \qquad (8)$$

where h is the time step, ξ_k and η_k are the realizations of independent standard Gaussian random variables ξ and η at time step k, $f_{\bar{k}} = f(y_{\bar{k}})$, $g_{\bar{k}} = g(y_{\bar{k}})$, $y_{\bar{k}} = y_k + (f + gz)_k h$, and $z_{\bar{k}} = z_k + b\xi_k h^{1/2} + az_k h$.

4. Gaussian Process Regression model

GPR gives an unbiased estimate of variable $x^f = [x_1^f, ..., x_{Nf}^f]^T$ ($x_i^f = x^f(t_i)$) at different times t_i and associated uncertainty given observations $x^o = [x_1^o, ..., x_{N_o}^o]^T$ ($x_i^o = x^o(\tilde{t}_i)$). Here, x^o and x^f can represent the same power grid states, where, for example, x^o is the observed ω, x^f is the forecasted ω, and $t_{N_o} < \tilde{t}_1$. Alternatively, x^o and x^f could be different states, e.g., x^o is observed ω, and x^f is predicted θ at the same times when ω is observed (i.e., $t_i = \tilde{t}_i$) or at different times.

Let us define $x = [x^o, x^f]^T$. Next, we assume that x is the Gaussian process with

$$\begin{bmatrix} x^o \\ x^f \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \bar{x}^o \\ \bar{x}^f \end{bmatrix}, \begin{bmatrix} C_{oo} & C_{of} \\ C_{fo} & C_{ff} \end{bmatrix} \right), \qquad (9)$$

where $\bar{x}^o$ and $\bar{x}^f$ are the so-called prior (or, unconditional) mean of x^o and x^f, respectively, and C_{oo}, C_{of}, C_{fo}, and C_{ff} are the respective prior covariances between x^o and x^o, x^o and x^f, x^f and x^o, and x^f and x^f. Then, GPR defines the conditional (or posterior) estimate of x^f given x^o as

$$\hat{x}^f(t) = C_{fo}C_{oo}^{-1}x^o \qquad (10)$$

and the conditional covariance as

$$\hat{C}_{ff} = C_{ff} - C_{fo}C_{oo}^{-1}C_{of}. \qquad (11)$$

The main challenge in GPR is to obtain prior statistics, i.e., prior mean and covariance functions. In the standard data-driven GPR method, prior statistics is found from measurements of x^o and x^f using the so-called marginal likelihood [8, 9]. Usually, a form of the covariance function is assumed, and the parameters (variance and correlation length) are found by minimizing the negative marginal likelihood of the Gaussian process. Such a purely data-driven approach is impossible to implement for reconstructing unobserved states from observed states without making assumptions about the correlation between unobserved and observed states.

Our proposed physics-informed GPR approach combines GPR and power grid equations. Specifically, we use Eqs (1) - (6) to compute the mean and auto- and cross-covariance functions in Eq (9). Let $\omega_k^{(n)}$, $\theta_k^{(n)}$, and $P_{m,k}^{'(n)}$ $k = 0, 1, ..., K$ be numerical solutions of Eqs (1) - (6) at time t_k with the initial condition $P_{m,0}^{'(n)}$ ($n = 0, 1, 2, \cdots, N$) drawn from the Gaussian distribution. Then, the mean of the state $x_k^{(n)}$ ($x = \omega, \theta, P_m'$) in in Eq (9) is computed as

$$\bar{x}(t_k) = \bar{x}_k = \frac{1}{N}\sum_{n=1}^{N} x_k^n, \quad k = 1, \cdots, K \qquad (12)$$

and the auto-covariance function in Eq (9) as

$$C_{xx}(t_j, t_k) = \frac{1}{N-1}\sum_{n=1}^{N}(x_j^n - \bar{x}_j)(x_k^n - \bar{x}_k),$$
$$j, k = 1, \cdots, K. \qquad (13)$$

The cross-covariance between different variables (e.g., between ω and θ) in in Eq (9) is computed as:

$$C_{\omega\theta}(t_j, t_k) = \frac{1}{N-1}\sum_{n=1}^{N}(\omega_j^n - \bar{\omega}_j)(\theta_k^n - \bar{\theta}_k),$$
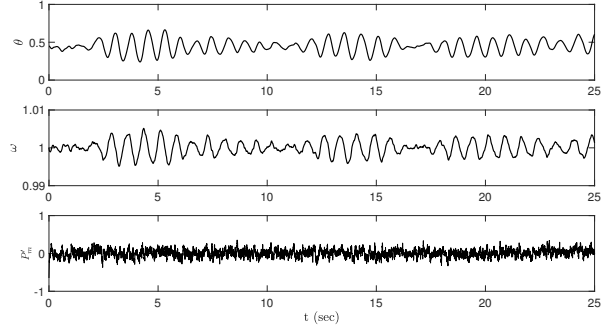$$j, k = 1, \cdots, K. \qquad (14)$$

5. Numerical results

Figure 1. One realization of $\theta(t)$, $\omega(t)$, and $P_m'(t)$.

The numerical solutions are obtained with $P_{\max} = 2.1$ p.u., $H = 5$ s, $D = 5$ p.u., $\langle P_m \rangle = 0.9$ p.u., $\omega_B = 120\pi$ rad/s, $\lambda = 0.026$ s, and $\omega_S = 1.0$. This parameters are taking from a recent study on small signal stability analysis in [11]. The solutions are computed for a total time of $T = 25$ s with a time step $h = 0.0025$ s. The initial conditions are assumed to be $\theta_0 = 0.45$, $\omega_0 = 1.0$. One thousand Monte Carlo realizations are performed with $P_{m,0}^{'(n)} = \xi^n$, ($n = 0, 1, ..., 999$), where ξ^n is a Gaussian random variable with zero mean and variance $\sigma = 0.1$.
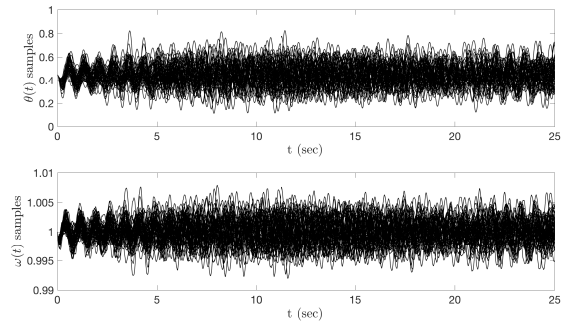

Figure 2. 50 realizations of $\theta(t)$ and $\omega(t)$.

Fig. 1 shows one realization of $\theta(t)$, $\omega(t)$, and $P_m'(t)$. It illustrates that θ and ω have higher-frequency oscillations on the scale of 1 s caused by the high-frequency oscillations in P_m' and lower-frequency

oscillations on the scale of 10 s that are proportional to the relaxation time of the system $\frac{H}{\omega_s D}$. In this work, we focus on modeling higher-frequency dynamics. Fig. 2 shows 50 realizations of $\theta(t)$ and $\omega(t)$. Mean and standard deviation of θ and ω approach asymptotic values after approximately 20 s, which are shown in Figs. 3 and 4.
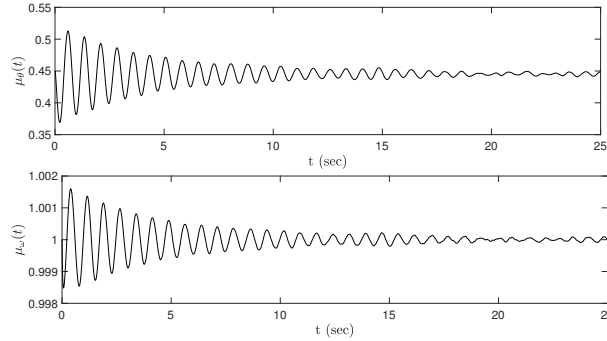

Figure 3. Mean of θ and ω.
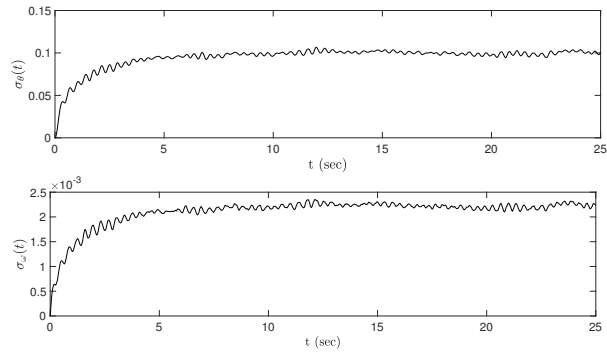
Figure 4. Standard deviation of θ and ω.

We consider three different cases for time series forecasting. In the first case, we assume that partial measurements of θ and ω up to a certain time are available and forecast of θ and ω beyond this time using conditional estimation of mean and standard deviation from the Gaussian process model. In the second case, we assume that partial measurements are available only for θ and predict θ, ω, and P'_m. In the third case, we predict θ, ω, and P'_m when only partial ω measurements are available.

5.1. Case 1: Forecast of θ and ω given past measurements of θ and ω

In this case, we assume that the $\theta(t)$ and $\omega(t)$ measurements are available until time $t = 8.3375$ s every 0.0025 s. Our objective is to forecast $\theta(t)$ and $\omega(t)$ for times greater than 8.3375 s. To test our approach, we compute prior statistics as described in Section 4 and
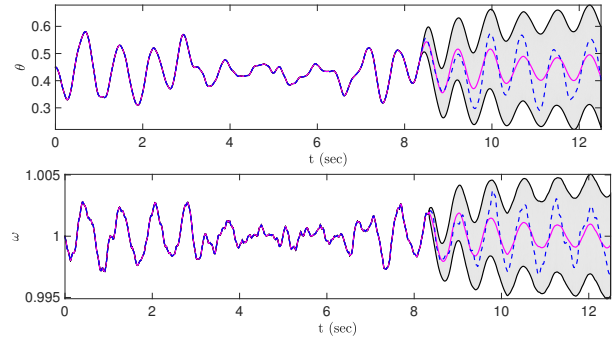
Figure 5. Forecast for θ and ω (magenta line) and the error bounds (gray area) corresponding to two standard deviations of θ and ω, respectively. The blue dashed line shows the ground truth (realization one).

use two randomly selected solutions of the governing equations as ground truth. Specifically, we treat the solution for $t < 8.3375$ s as observations of $\theta(t)$ and $\omega(t)$ and use these observations as input for the GRP model to predict $\theta(t)$ and $\omega(t)$ for $t > 8.3375$ s. The second part of the solutions is used to validate the GPR predictions.

For these two synthetic observation sets, Figs. 5 and 6 show the forecast (mean values) of θ and ω (magenta line) and uncertainty, represented by the gray area corresponding to two standard deviations of θ and ω. The blue dashed line shows the ground truth. In both cases, the physics-informed GPR method is able to provide an accurate deterministic forecast of θ and ω for the time horizon of at least 2 s or 50λ, i.e., the mean prediction of the GPR closely agrees with the ground truth for at least 20λ, after which the mean prediction quality deteriorates. More importantly, the "statistical forecast" is accurate for at least 4 s or 80λ, i.e., the ground truth lies within the predicted uncertainty envelope.

For comparison, we also forecast θ and ω using the standard data-driven GRP method with the covariance obtained from the marginal likelihood method (Fig. 7). The comparison of the two methods demonstrates that the the physics-informed GPR performs much better than the data-driven GPR. The data-driven GPR provides an accurate mean prediction for less than 5λ and statistical prediction for less than 20λ.

5.2. Case 2: Forecast θ, ω, and P'_m given measurements of θ

In this case, we assume that $\theta(t)$ measurements are available for $t < 8.3375$ s every 0.0025 s. Our goal is to predict $\theta(t)$ for $t > 8.3375$ and $\omega(t)$ and $P'_m(t)$ for the entire time interval $t \in [0, 12.5]$ s. As in Case 1, we
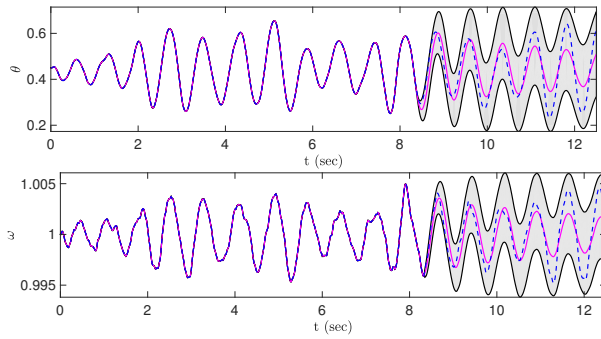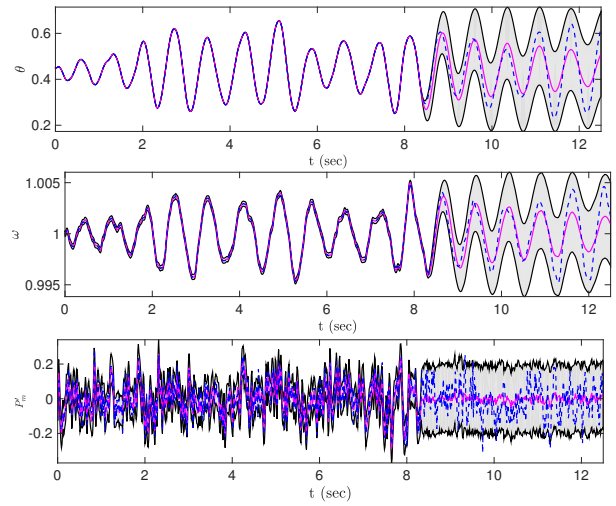
Figure 6. Forecast for θ and ω (magenta line) and the error bounds (gray area) corresponding to two standard deviations of θ and ω, respectively. The blue dashed line shows the ground truth (realization two).
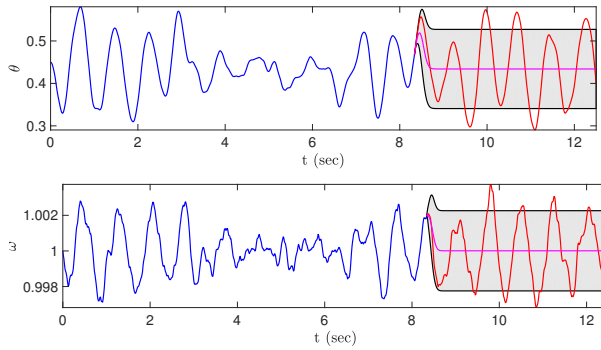
Figure 7. Forecast for θ and ω (magenta line) and the error bounds (gray area) corresponding to two standard deviations of θ and ω, respectively, computed using the standard data-driven GPR with covariance obtained from the marginal likelihood method. Red line shows the ground truth.

Figure 8. Forecast of θ, ω, and P'_m with no observations of θ for $t > 8.3375$.

very accurate for computing unobserved variables. This is important because some variables, including P_m, are difficult to measure directly.

5.3. Case 3: Forecast of θ, ω, and P'_m given measurements of ω

Figure 9. Forecast of θ, ω, and P'_m process with no observations of ω for $t > 8.3375$.
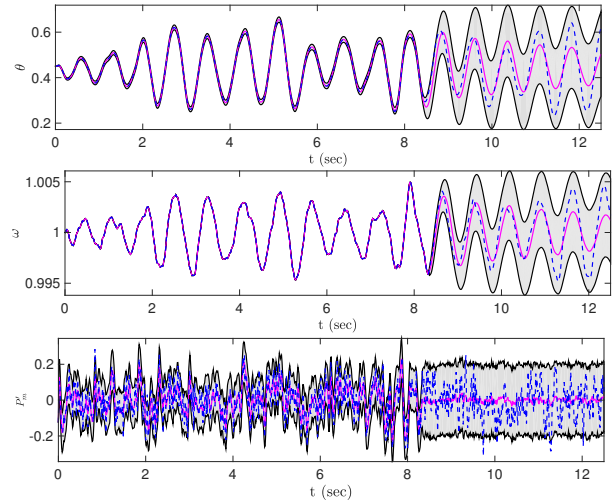
compute prior statistics by solving the stochastic power grid equations and use a randomly selected solution as observations of $\theta(t)$ for $t < 8.3375$ s and ground truth to validate the physics-informed GPR predictions of $\theta(t)$ (for $t > 8.3375$), $\omega(t)$, and $P'_m(t)$. Fig. 8 shows the predictions of θ, ω, and P'_m and the ground truth results. Notably, the forecast of ω is exactly the same as in Case 1 (Fig 6) because the ω observations are the same in these two cases. The prediction for $\theta(t)$ is excellent for $t < 8.3375$ when $\omega(t)$ measurements are available. This is evident by the mean prediction being very close to the ground truth and very small standard deviation (i.e., the narrow gray "uncertainty" area). For $t > 8.3375$, the agreement between mean prediction and ground truth is almost as good as in Case 1. Also, as in Case 1, the ground truth remains within the uncertainty envelope. Similar results are observed for P_m. These results demonstrate that the physics-informed GPR is

This case is similar to Case 2, except instead of $\theta(t)$ measurements, we have $\omega(t)$ measurements for $t < 8.3375$. Our objective is to predict $\omega(t)$ for $t > 8.3375$ and $\theta(t)$ and $P'_m(t)$ for the entire time interval $t \in [0, 12.5]$ s. Fig. 9 presents measurements of ω and predicted $\omega(t)$, $\theta(t)$, and $P'_m(t)$, as well as the ground truth solutions for these variables. As in Case
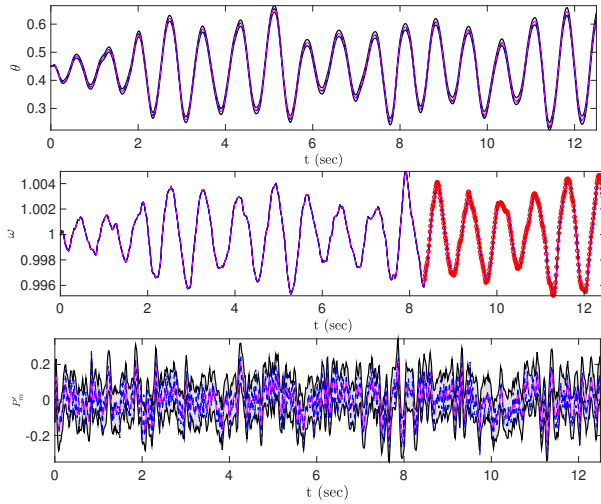
Figure 10. Forecast of θ, ω, and P'_m with 333 observations of ω for $t > 8.3375$.

2, physics-informed GPR provides excellent predictions of $\theta(t)$ and $P'_m(t)$ for $t < 8.3375$ and a similarly good prediction for $t > 8.3375$ s as in Case 1 where the $\theta(t)$ measurements for $t < 8.3375$ are available.

Next, we assume that 333 $\omega(t)$ measurements (shown by red circles in Fig. 10) are available for $t > 8.3375$. Fig. 10 shows that these measurements significantly improve prediction of the (unobserved) $\theta(t)$ and $P'_m(t)$ variables.

Fig. 11 depicts a comparison of the physics-informed GPR prediction of P'_m based on measurements of θ and ω. This comparison reveals that the predictions of $P'_m(t)$ are reasonably good in both cases and better when θ measurements are available instead of ω measurements. This shows that θ is correlated more strongly with P'_m than ω.
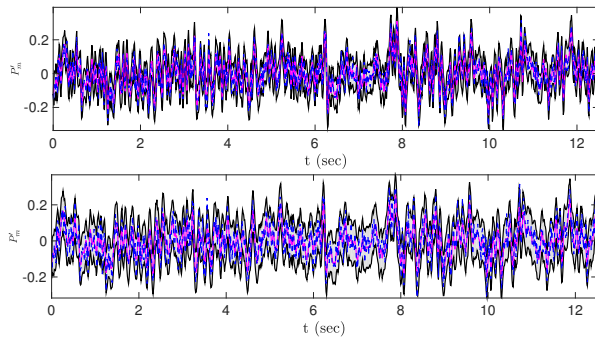
Figure 11. Forecast of P'_m process when observations of a) θ and b) ω are available.

6.   Discussion and conclusions

We have presented a physics-informed Gaussian Process Regression (GPR) model to predict the phase angle, angular speed, and wind mechanical power when past measurements of all or some of these states are available. In the traditional data-driven GPR method, a form of the covariance matrix of the Gaussian Process model is first assumed and then its parameters are found from measurements. In the physics-informed GPR, we treat unknown variables, including wind speed and mechanical power, as a random process and compute the covariance matrix of the GPR model from the resulting stochastic power grid equations. We have demonstrated that the physics-informed GPR method is significantly more accurate than the standard data-driven GPR method for forecasting generators' angular velocity and phase angle. We also have shown that the physics-informed GPR provides accurate predictions of the unobserved wind mechanical power, phase angle, or angular velocity when measurements of only one of these variables are available.

One advantage of GPR over neural-network-based methods is that its predictions come with uncertainty bounds, i.e., GPR predicts the mean value and standard deviation of quantities of interest. We have demonstrated that the physics-informed GPR accurately and reliably (with small uncertainty) predicts unobserved variables from observed ones for the entire time interval where observations of other (correlated) variables are available. For forecasting, the uncertainty in predicted values slowly increases with the time horizon. In the considered cases, the forecasted mean values were accurate in the time horizon equal to as many as 200 correlation times of the wind power fluctuations. For larger time horizons, we have noted the predicted values to be within two standard deviations from the ground truth.

In this study, we have modeled high-frequency power grid dynamics produced by high-frequency wind oscillations. The correlation length of the wind mechanical power in the simulated data was $\lambda = 0.026$ s, and the accurate mean values were obtained for up to 2 s time horizon. Such high-frequency oscillations have been shown to have a significant effect on small signal stability [10, 11, 12]. For economic load dispatch and planning, the considered wind oscillations are on the scale of minutes. Load dispatch and planning require forecast for the 6 hr time horizon and unit commitment, reserve requirement decisions, and maintenance scheduling need 24 hrs to one-week forecasts [1]. In our future work, we will investigate if the physics-informed GPR can provide an

accurate forecast on such time horizons by considering low-frequency wind and load fluctuations with the correlation time on the order of 2 - 30 min.

It is important to note that the proposed physics-informed GPR approach is "non-intrusive," i.e., it does not require any modifications of equations governing the power grid dynamics. Therefore, any existing power-grid code can be used to compute auto- and cross-covariance functions. Then, to apply the proposed approach to a larger power network with several generators and wind farms, the Gaussian process (9) must be extended by including auto-covariance functions for each generator and cross-covariance functions between different generators.

7. Acknowledgment

References

[1] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons," in *North American power symposium (NAPS), 2010*, pp. 1–8, IEEE, 2010.

[2] A. Fentis, L. Bahatti, M. Mestari, M. Tabaa, A. Jarrou, and B. Chouri, "Short-term pv power forecasting using support vector regression and local monitoring data," in *Renewable and Sustainable Energy Conference (IRSEC), 2016 International*, pp. 1092–1097, IEEE, 2016.

[3] J. Grant, M. Eltoukhy, and S. Asfour, "Short-term electrical peak demand forecasting in a large government building using artificial neural networks," *Energies*, vol. 7, no. 4, pp. 1935–1953, 2014.

[4] E. Yeung, S. Kundu, and N. Hodas, "Learning deep neural network representations for koopman operators of nonlinear dynamical systems," *arXiv preprint arXiv:1708.06850*, 2017.

[5] K. Thiyagarajan and R. S. Kumar, "Real time energy management and load forecasting in smart grid using compactrio," *Procedia Computer Science*, vol. 85, pp. 656–661, 2016.

[6] C. Nichiforov, I. Stamatescu, I. Făgărăşan, and G. Stamatescu, "Energy consumption forecasting using arima and neural network models," in *Electrical and Electronics Engineering (ISEEE), 2017 5th International Symposium on*, pp. 1–4, IEEE, 2017.

[7] C. E. Rasmussen, "Gaussian processes in machine learning," in *Advanced lectures on machine learning*, pp. 63–71, Springer, 2004.

[8] M. Blum and M. Riedmiller, "Electricity demand forecasting using gaussian processes," *Power*, vol. 10, p. 104, 2013.

[9] T. Hachino, H. Takata, S. Fukushima, and Y. Igarashi, "Short-term electric load forecasting using multiple gaussian process models," *World Academy of Science, Engineering and Technology, International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering*, vol. 8, no. 2, pp. 447–452, 2014.

[10] W. S. Rosenthal, Z. Huang, and A. M. Tartakovsky, "Ensemble kalman filter for dynamic state estimation of power grids stochastically driven by time-correlated mechanical input power," *IEEE Transactions on Power Systems*, vol. PP, no. 99, pp. 1–1, 2017.

[11] P. Wang, D. A. Barajas-Solano, E. Constantinescu, S. Abhyankar, D. Ghosh, B. F. Smith, Z. Huang, and A. M. Tartakovsky, "Probabilistic density function method for stochastic odes of power systems with uncertain power input," *SIAM/ASA Journal on Uncertainty Quantification*, vol. 3, no. 1, pp. 873–896, 2015.

[12] D. A. Barajas-Solano and A. M. Tartakovsky, "Probabilistic density function method for nonlinear dynamical systems driven by colored noise," *Phys. Rev. E*, vol. 93, p. 052121, May 2016.

[13] G. Milshtein and M. Tret'yakov, "Numerical solution of differential equations with colored noise," *Journal of Statistical Physics*, vol. 77, no. 3-4, pp. 691–715, 1994.

[14] P. M. Anderson and A. A. Fouad, "Power system control and stability," *IEEE Series on Power Engineering*, 2003.