State Estimation with Exogenous Information for Grids with Large Renewable Penetration

Basel Morsy *, Mahmoud AlSadat† and David Pozo‡
Center for Energy Science and Technology,
Skolkovo Institute of Science and Technology, Moscow, Russia
Email: *Basel.Morsy@skoltech.ru, †Mahmoud.Alsadat@skoltech.ru, ‡D.Pozo@skoltech.ru

***Abstract*—State estimation (SE) is one of the fundamental power system problems used in the determination of system state variables based on measurements. Current approaches for state estimation problems rely only on direct and pseudo-measurements of power system parameters, while some additional exogenous parameters can be measured and fit into the existing state estimation models, especially in renewable-rich grids where the power output of renewable power plants is dependent on weather conditions. In this work, we introduce a methodology for the inclusion of weather data as exogenous measured parameters into the state estimation problem. To test our proposed framework, a simulation environment was developed and validated in a 5-bus system. The introduced methodology was carried out 100 times to reach a statistically significant result and to investigate the model stability. The exogenous weather parameter measurement added to the model proved its capability of enhancing state estimation accuracy by up to 79%.**

***Keywords*—State estimation, renewable generation**

I. Introduction

A power system is fully determined by the network topology and system states (voltage magnitudes and phase angles) at each bus [1]. In order to obtain a full description of the power system, which is critical for power flow calculations, these parameters should be known with a high degree of accuracy. However, there is always a degree of uncertainty associated with these parameters arising from errors within measurement devices as well as bad data (inconsistent data) that may result from network topology changes or faults within measurements devices [2] [3]. State estimation is a power system analysis tool used in dealing with these uncertainties. It provides a method for approximating the true values of state variables as well as detecting and eliminating bad data. Current state estimation models rely on the use of power system measurements to perform state estimation. Of these models, the weighted least square model (WLS) is the most investigated and used one [4] [5]. It relays on minimizing the squared error between the measured parameters and their equivalent calculated values based on the estimated states, giving high weights to more accurate measurements and lower weights to less accurate ones [2]. Taking into consideration the increasing level of penetration of renewable energy sources, weather data can be a good candidate for enhancing the current state estimation models. This is motivated by the fact that these energy sources power output is highly dependant on weather parameters, which can be easily measured (e.g. wind speed, temperature and solar irradiance). For this exogenous measurements to fit into the current model, a relation between the measured parameter (wind speed in our work here) and the power system states must be obtained, which will be discussed in details in the following sections. While such concept may seem very promising, there exist a number of challenges that must be investigated before arriving at a stable model. In this work, a simulation environment was developed in *Julia* language, a 5-bus system was investigated at different operating points and SE was performed on the system with and without the addition of exogenous parameters to assess the newly proposed model and evaluate the enhancement based on arbitrarily defined quality metrics, which will be defined in details in the following sections. For convenience, vector and matrix variables are denoted in *bold* while scalar variables are not.

II. Methodology

A. Static-state estimation (SSE)

SE model used was static-state estimation based on WLS method. The model formulation does not take into considerations transients in the system governed by physical laws, instead it assumes a static steady state operation [1] [6]. This model will be used within our investigation hereby. The formulation goes as follows:
Assume there are some measurements $\boldsymbol{z}$ where $\boldsymbol{z}$ is a vector of all delivered measurements of the system. This follows that

$$\boldsymbol{z} = \boldsymbol{z}^{\boldsymbol{true}} + \boldsymbol{e} \tag{1}$$

where $\boldsymbol{z}^{\boldsymbol{true}}$ is a vector of the true values of these measurements and $\boldsymbol{e}$ is a vector of the corresponding error for each measurement. Let $\boldsymbol{h}(\boldsymbol{x})$ be a vector function, where each element is the governing physical law that relates the corresponding measurement in $\boldsymbol{z}$ to the system states (bus voltages and phase angles) defined by vector $\boldsymbol{x}$. Espousing the weighted least square approach (WLS) the SE problem can be written as follows [1] [5] [4]:

$$\min_{\boldsymbol{x}} J(\boldsymbol{x}) = \sum_{i=1}^{m} w_i (h_i(\boldsymbol{x}) - z_i)^2 = [\boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{z}]^T W [\boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{z}] \tag{2}$$

978-1-7281-5655-2/20/$31.00 ©2020 IEEE

Solution of such an optimization problem, which is non-linear and non-convex, can be achieved using Newton-Raphson method. Using first order optimality conditions, the iterative solution method could be easily concluded to be as follows:

$$\boldsymbol{G}(\boldsymbol{x^k}) = \boldsymbol{H^T}(\boldsymbol{x^k})\boldsymbol{W}\boldsymbol{H}(\boldsymbol{x^k}) \tag{3}$$

$$\boldsymbol{g}(\boldsymbol{x^k}) = \boldsymbol{H^T}(\boldsymbol{x^k})\boldsymbol{W}[\boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{z}] \tag{4}$$

$$\boldsymbol{x^{k+1}} = \boldsymbol{x^k} - \boldsymbol{G^{-1}}(\boldsymbol{x^k})\boldsymbol{g}(\boldsymbol{x^k}) \tag{5}$$

Elements of $\boldsymbol{h}(\boldsymbol{x})$ belong to a set of functions given as follows:

1) V_i : Voltage at bus i
2) P_i : Active power injection at bus i
3) Q_i : Reactive power injection at bus i
4) P_{ij} : Active power flow from bus i to bus j
5) Q_{ij} : Reactive power flow from bus i to bus j

From power flow equations derived in [5]:

$$P_i = V_i \sum_{j=1}^{n} V_j (G_{ij}\cos\delta_{ij} + B_{ij}\sin\delta_{ij}) \tag{6}$$

$$Q_i = V_i \sum_{j=1}^{n} V_j (G_{ij}\sin\delta_{ij} - B_{ij}\cos\delta_{ij}) \tag{7}$$

$$P_{ij} = V_i V_j (G_{ij}\cos\delta_{ij} + B_{ij}\sin\delta_{ij}) - V_i^2 G_{ij} \tag{8}$$

$$Q_{ij} = V_i V_j (G_{ij}\sin\delta_{ij} - B_{ij}\cos\delta_{ij}) - V_i^2 (B_{ij} - b_{s,ij}) \tag{9}$$

B. Assumptions

1) Steady state or quasi-steady state operation is investigated. Although a power system can never achieve a steady state due to the continuous changes in loads and generations, yet, it is reasonable to approximate the system as operating in steady state conditions for short periods of time [2].
2) Within this work, we assume no topological changes are taking place.
3) It is also assumed here that there is no bad data.
4) Renewable energy bus (Wind turbine in this work) is modelled as load (PQ) bus of which active power load is negative of the generation and reactive power load is *zero*.
5) Exogenous parameter used is wind speed measurement for wind turbine bus.
6) Wind speeds less than 3 m/s are set to *zero* and wind speeds greater than 12 m/s are set to 12 to resemble the pitching mechanism taking place with wind turbines.

C. Inclusion of exogenous parameters

In order to include exogenous parameters in the formulation of SE problem, it is necessary to find adequate equations that relate these exogenous parameters to power system states. In this work, we are concerned with the inclusion of wind speed measurements in SE. In [7], the physical law that relates wind speed to generated power from a wind turbine is given by:

$$P_w = \frac{1}{2}\rho S v_w^3 C_P \tag{10}$$

where P_w is wind turbine generated power, ρ is air density, S is the surface area swept by the wind turbine rotor, v_w is wind speed and C_P coefficient of power of the wind turbine. Let $\frac{1}{2}\rho S C_P$ be a constant K, then:

$$v_w^3 = \frac{P_w}{k} \tag{11}$$

Let the wind turbine bus be bus α connected to busses $j \in \Omega$, then the net injection in bus α has to be equal to wind power generated at it (since there are no loads assumed to be at bus α).

$$P_W = \sum_{j\in\Omega} P_{\alpha j} \tag{12}$$

From (11) and (12) it follows that:

$$v_w^3 = \frac{\sum_{j\in\Omega} P_{\alpha j}}{k} \tag{13}$$

Equation (13) was relaxed to be:

$$|v_w^3 - \frac{\sum_{j\in\Omega} P_{\alpha j}}{k}| \leq \epsilon \tag{14}$$

where ϵ is arbitrarily defined to be the summation of the variance of error in wind speed and line flows measurements. Getting back to SE formulation, the vector function $\boldsymbol{h}(\boldsymbol{x})$ now will have another element which is v_w, such that v_w has to satisfy the constraint equation (13), then a new measurement can be added to the measurements vector $\boldsymbol{z}(\boldsymbol{x})$ which is the *wind speed* measurement.

D. Simulation environment

In this work, we developed a simulation environment to investigate the effect of inclusion of exogenous parameters. The simulation environment can be viewed as a two-layer framework. The first layer resembles the physical system operation with 100% certainty of all information about it and satisfaction of all physical laws governing the power system. The second layer is a *representation model* of the physical system through erroneous measurements obtained about the system via measurement devices. A model of the physical power system has also to obey physical laws, thus state estimation is performed. Since we have certain information (obtained from layer 1) and uncertain estimated information (obtained from SE in layer 2), therefore, we can assess the performance of SE. Figure 1 shows a diagram of the simulation environment.

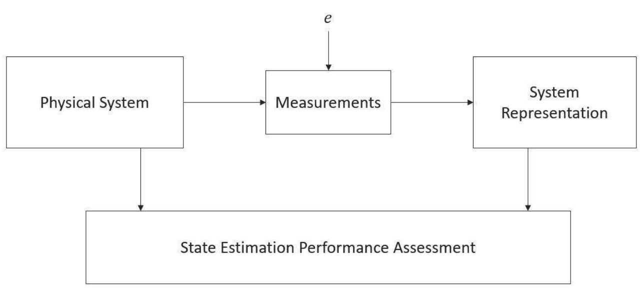

Fig. 1. Illustration of simulation environment

In fact, our simulation environment resembles what happens in real cases, except that we do not have certain information about the physical power system, we only have measurements, that is the reason why it is important to perform state estimation.
As mentioned before, the system being investigated here is a 5-bus system, its topology is given as shown in figure 2. Demands are located buses (2,3,4), wind generated power is located at bus 5, and bus 1 is a conventional generator bus. Constant k in equation (11) is calculated for our assumed wind turbine of 56 m diameter and 0.38 C_P to be 573.265, all values given are in SI units.

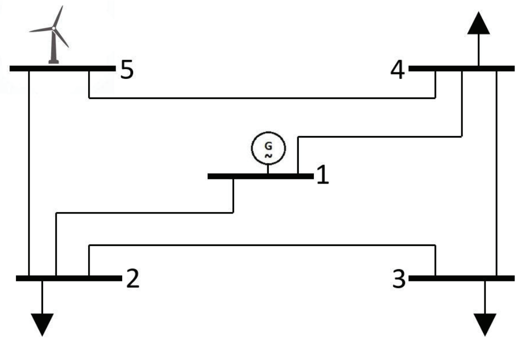

Fig. 2. Single-line diagram for the system investigated

Next we discuss the work flow of the simulation environment. The system shown in figure 2 is fed with values for demands at buses (2,3,4) and wind energy generation at bus (5) as a time-series resembling the system operation for 8 days on an hourly basis, which sums up to be 192 time instances, equivalently, 192 snap-shots of the system. For each snap-shot a set of operations are performed as follows:

1) Solve optimal power flow (OPF) and get full and true information about the system.
2) Generate measurements for the system measurable variables (bus voltages, real and reactive power injections at buses, real and reactive power flows at lines). Measurements generation is performed as per the following equation:

$$X_{measurement} = X_{true} + N(0, \sigma_X) \tag{15}$$

3) Perform state estimation with and without exogenous parameters and save results.

Then the estimated state variables at each time instance are used to assess both methods (SE with and without exogenous parameters) according to some certain quality metrics that will be discussed later. Figure 3 shows a flow chart for how the simulation environment work. The data used for demands and wind speed, hence wind power, were generated arbitrarily based on historical data. The ratios between load and demand where adjusted such that there is no excess wind generation at any time instance, because generator technical constraints does not allow for absorbing active power, and renewable energy generation curtailment was not implemented within the simulation environment.

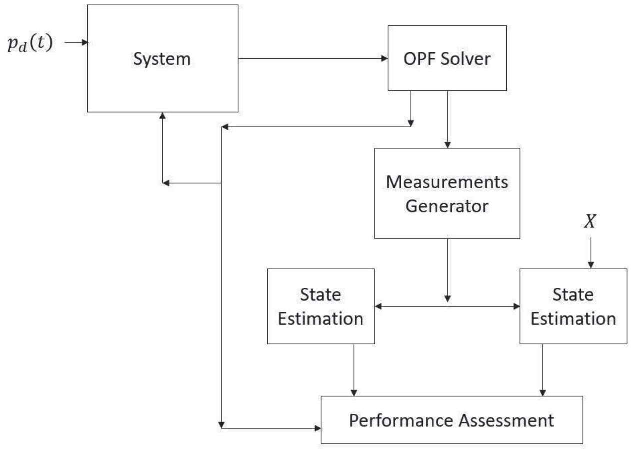

Fig. 3. Flow chart of the simulation environment

All the previous steps were repeated 100 times to reach a statistically significant result, that follows that every calculated quality metric is the average of quality metrics evaluated in each iteration.

E. Quality metrics

We have chosen the sum of mean square error of bus voltages ($SMSE_V$) and sum of mean square error of phase angles ($SMSE_\delta$) to construct an arbitrarily defined index which we call *error index* (E_i), the less it is, the closer the model to the real physical system operation. Following is a formal definition of the aforementioned metrics.
Let $\tilde{\boldsymbol{V}}^{\boldsymbol{t}}$ and $\tilde{\boldsymbol{\delta}}^{\boldsymbol{t}}$ be two vectors of estimated bus voltages and estimated phase angles at time instance t respectively, the dimensions of these vectors will be $size(\tilde{\boldsymbol{V}}^{\boldsymbol{t}}) = (5, 1)$. And let vectors $\boldsymbol{V}^{\boldsymbol{t}}$ and $\boldsymbol{\delta}^{\boldsymbol{t}}$ be bus voltages and phase angles obtained from solving *optimal power flow* for the system at time instance t. Let $t \in T$ and operator $\mathrm{M}(X, T)$ be defined as the mean of X over time period T. Then the metrics are defined as follows:

$$\boldsymbol{SE_V} = (\tilde{\boldsymbol{V}}^{\boldsymbol{t}} - \boldsymbol{V}^{\boldsymbol{t}})^2 \quad for \quad t \in T \quad dim = (5, T) \tag{16}$$

$$\boldsymbol{SE_\delta} = (\tilde{\boldsymbol{\delta}}^{\boldsymbol{t}} - \boldsymbol{V}\boldsymbol{\delta}^{\boldsymbol{t}})^2 \quad for \quad t \in T \quad dim = (5, T) \tag{17}$$

From equations (16) and (17) denoting squared errors array by SE and the dimension of each array of them by dim, we can define mean squared error vectors which defines the mean squared error of each state variable at each bus, thus MSE vectors have same dimensions as voltage and phase angle vectors:

$$\boldsymbol{MSE_V} = \mathrm{M}(\boldsymbol{SE_V}, T) \quad dim = (5, 1) \tag{18}$$

$$\boldsymbol{MSE_\delta} = \mathrm{M}(\boldsymbol{SE_\delta}, T) \quad dim = (5, 1) \tag{19}$$

From equations (18) and (19) we can then define the sum of mean squared error for bus voltages and phase angles (SMSE). This metric gives a measure of the total squared error for each variable and defined as follows:

$$SMSE_V = \sum_{i=1}^{5} MSE_{V,i} \tag{20}$$

$$SMSE_\delta = \sum_{i=1}^{5} MSE_{\delta,i} \tag{21}$$

From equations (20) and (21) the error index E_i will be defined as follows:

$$E_i = SMSE_V + SMSE_\delta \tag{22}$$

F. Input data fed into the system

The input demands and wind energy generation are given to the system as a time-series. Details on input data can be found in the online appendix [8].

An additional case was considered in which the wind energy generation data was assumed to be higher than the maximum possible capacity given the wind speed at this case. This actually resembles dishonest renewable generation (DRG). This case was assumed to be at the last two days of the period considered within our study. Power generation at wind energy bus is then given as follows in figure 4.

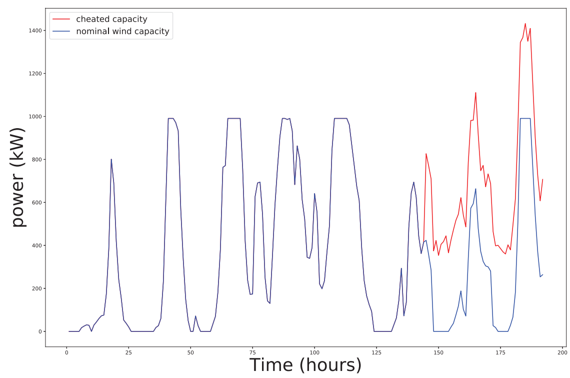

Fig. 4. Generation at wind turbine bus with DRG

III. Results and Discussion

The aforementioned input data was fed into the system, using the work flow discussed in the simulation environment description. The SE performance with and without inclusion of exogenous parameters was assessed subject to the different quality metrics discussed previously (mean squared error and error index). The following results were found.

A. No DRG case

As data was provided for the first 6 days with no DRG (144 hours, consequently 144 snap-shots of the system), all fundamental procedures were then performed 144 times, one time at each snap-shot of the system, including SE. The following two figures (5 and 6) show the value for mean squared error of bus voltages and phase angles respectively as per their definition discussed in the quality metrics section. It can be seen that for bus voltages, the MSE value for SE with exogenous parameters is lower than that of SE without exogenous parameters for all buses, while for the phase angles this is not the case for all buses.

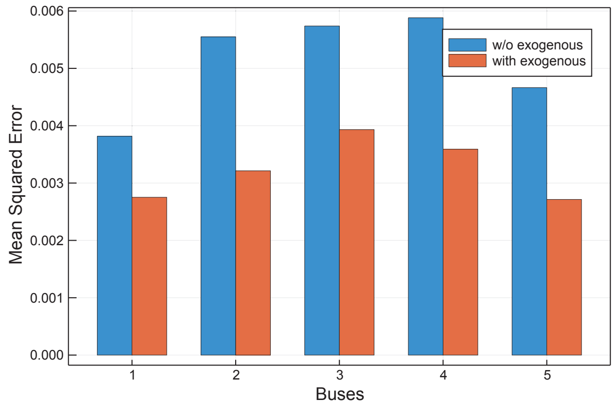

Fig. 5. MSE of bus voltages

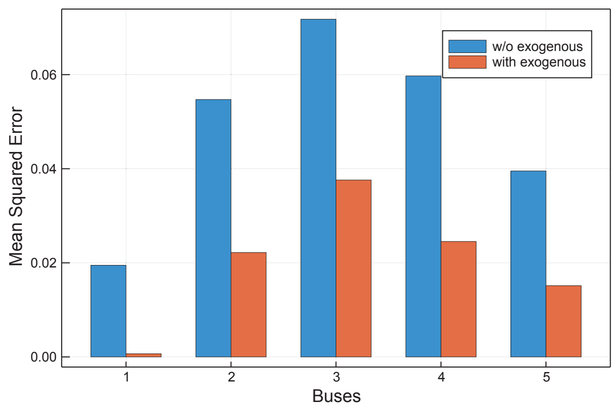

Fig. 6. MSE of phase angles

The main assessment parameter used to assess each of the SE models (with and without exogenous parameters) is the aforementioned *error index* (E_i). As shown in table I the value for error index for SE with exogenous parameters ($E_{i,X}$) is less than that for SE without exogenous parameters (E_i). The values are namely $E_{i,X} = 0.0546$ and $E_i = 0.2650$. This means a decrease in error index by 79% caused by inclusion of exogenous parameters.

B. In case of DRG

In case that wind turbine bus generates more active power than the available capacity concluded from wind speed, which can happen in case of DRG, the inclusion of exogenous parameters is expected to give worse results than not including them. It would also be the case when wind speed measurements are considered as bad data, both cases are equivalent. As shown in figures (7 and 8) the mean squared error for bus voltages and phase angles is maximum for SE with exogenous parameters inclusion.

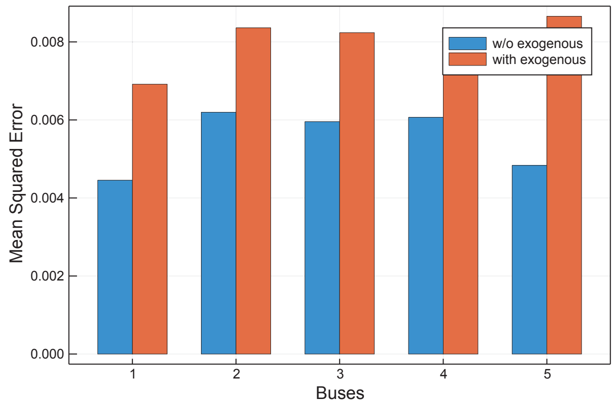

Fig. 7. MSE of bus voltages in case of DRG

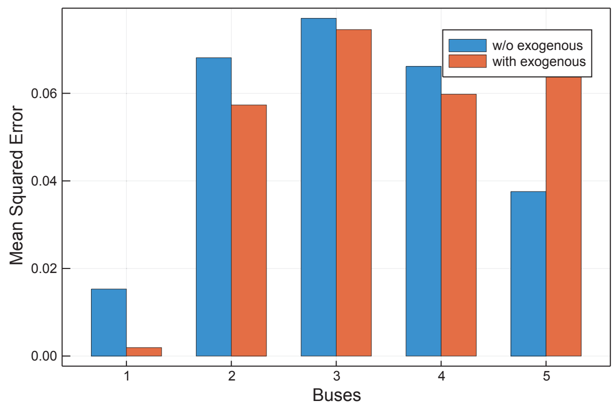

Fig. 8. MSE of phase angles in case of DRG

The overall assessment of both SE models is also done by evaluation of the *error index* for both of them. As shown in table I the error index for SE with exogenous parameters is much higher than that of SE without. $E_{i,X} = 0.1409$ while $E_i = 0.0001$, which indicates that the inclusion of exogenous parameters is not quiet useful when there is DRG.

To further elaborate on the overall assessment of the two models, figure 9 shows the results of the 100 experiments performed on the case study we have. As shown in the figure, the first two box plots shows the case of no dishonest renewable generation (DRG) and the inclusion of exogenous parameters enhanced the performance of SE indicated by a lower error index median value. The median value is chosen to be representative as 50% of the experiment population lies there. It is also noticed that the inter-quartile range of our introduced model is less than that of the conventional SE model, which indicates a higher stability. On the other hand, it can be seen that in case of DRG the value of the error index of our introduced model is higher than that of the conventional SE, and this is plausible as the DRG introduces inconsistency between the supposed renewable generation and the actual one we have. It can also be seen that the inter-quartile range is larger for SE with exogenous parameters which indicates less stability than conventional SE.

TABLE I
COMPARISON OF ERROR INDICES

	E_i	$E_{i,X}$
Without DRG	0.2650	0.0546
With DRG	0.0001	0.1409

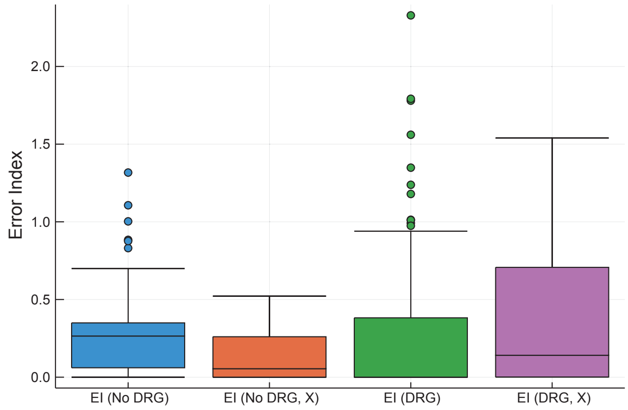

Fig. 9. Error index box plot

C. Confidence of absence of bad data

We consider a null-hypothesis that there is no bad data. We tested this hypothesis using the Chi-square test as done in [9] and [1]. As shown in figure 10 the period from hour 0 to hour 144, where there is no DRG, the two models are consistent in terms of the confidence level of the null hypothesis. However, in the last 48 hours, where there is DRG, we can see that the null hypothesis is totally rejected for the SE with exogenous parameters, which means that we are almost sure that there exists bad data during this period of time. It can be seen that the conventional SE could not detect this anomaly of DRG because no exogenous parameters were included, which means that the model does not distinguish between power generated from renewable and non-renewable sources (source blind).

IV. CONCLUSION

As discussed in the results section, adding exogenous measurements improved the accuracy of state estimation by about 79%. this accuracy improvement was during the first 6 days where there was no inconsistency in the measurements (no

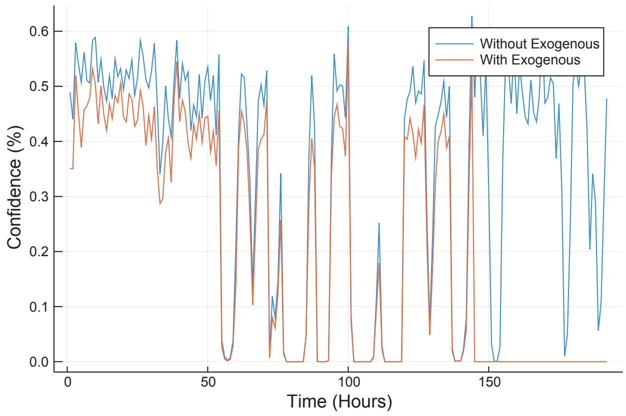

Fig. 10. Confidence percentage of null hypothesis

conventional generation is added to the wind generation). While during the last 2 days, where conventional generation is added (DRG case), the state estimation with exogenous measurements was much worse than the state estimation without exogenous parameters. In conclusion, by using our introduced model and confidence curve, DRG can be detected as well as state estimation can be enhanced in case of no DRG is present.

V. Future Work

Although the presented model showed an enhancement in the state estimation performance when tested on a 5-bus network, Testing on more complex networks needs to be performed to ensure the stability of the model. This includes networks of larger number of buses as well as different topologies.

Another important aspect to consider is the number of available measurements. In this work we considered that all the power system measurements are used in the state estimation. Where in reality this is not the case. Also, in order to further investigate the effect of the exogenous parameter on state estimation performance, we need to consider fewer number of power system measurements. We also need to consider the instantaneous ratio between renewable generation and conventional generation, as this would give an estimate of how critical and effective the inclusion of exogenous parameters is.

Machine learning models can also be considered in this problem. These models can help with regards to two main aspects. First is related to bad data detection and identification. Where a model could be optimized to detect inconsistent changes in measurement data. The second aspect is related to finding relations between relevant weather data and power system states of which a mathematical model is hard to obtain (e.g. temperature).

References

[1] A. G. E. Ali Abhur, *Power System State Estimation*. CRC Press, 2004.

[2] F. C. SCHWEPPE and J. WILDES, “Power system static-state estimation, part i: Exact model,” *IEEE Transactions on Power Apparatus and Systems*, 1970.

[3] O. A. N. Vempatis, et al, “Generalized state estimation,” *Proceedings of the 20th International Conference on Power Industry Computer Applications*, 1997.

[4] P. Zarco and A. G. Exposito, “Power system parameter estimation: A survey,” *IEEE TRANSACTIONS ON POWER SYSTEMS*, 2000.

[5] J. John J. Grainger, William D. Stevenson, *Power System Analysis*. McGraw-Hill, Inc., 1994.

[6] B. F. W. A. J. Wood, *Power generation, operation and control*. John Wiley and Sons, Ltd, 1996.

[7] T. Ackermann, *Wind Power in Power Systems*. John Wiley and Sons, Ltd, 2005.

[8] B. Morsy, M. AlSadat, and D. Pozo, “State estimation with exogenous information for grids with large renewable penetration,” 2020. [Online]. Available: https://sites.skoltech.ru/paco/?attachment%20id=496

[9] T. P. Vishnu, V. Viswan, and A. M. Vipin, “Power system state estimation and bad data analysis using weighted least square method,” in *2015 International Conference on Power, Instrumentation, Control and Computing (PICC)*, Dec 2015.