3rd SEERC Conference Vienna 2021
Cooperation • Sustainability • Future

CIGRE SEERC South East European Regional Council

Paper no 1125

CIGRE South East European Regional Council Conference 2021 in Vienna, Austria

Transition to Future Power Systems

Experiences with Meteorological Models for Asset Scheduling in Local Energy Communities and Microgrids

[1]Michael H. SPIEGEL and [1,2]Thomas I. STRASSER
[1]AIT Austrian Institute of Technology, [2]Vienna University of Technology (TU Wien)
[1,2]Austria
michael.spiegel@ait.ac.at

SUMMARY

Local energy community networks and microgrids are built to meet current challenges in the power system such as reducing CO_2 emissions by increasing the share of renewables, increasing economic benefits, and enhancing power supply resiliency. One lever that can be pulled is to deploy controllers that presciently schedule the operation of all controllable assets in advance. Control algorithms may range from simple rules that include predicted load and infeed to complex optimization problems that minimize operation costs such that a resilient operation is guaranteed. Often several modelling assumptions are made to quantify volatile generation within the scheduling horizon and to optimally operate the assets. In case these assumptions are violated, the economic performance as well as the resiliency of the power systems can be seriously affected.

This work addresses the impact of common assumptions regarding meteorological inputs. It uses an exemplary meteorological dataset to assess the impact the assumptions have on the forecasted renewable energy generation. These forecasts are particularly needed to optimally schedule local energy community/microgrid assets and are thus vital for the system performance. Additionally, the paper showcases the impact of plant models on the eligibility of studied input assumptions and assesses the value of scheduling-time information such as numerical weather forecasts. Therefore, several models of wind speed and global horizontal irradiation are trained and evaluated on two independent sets of measurements. The accumulated daily wind speed, irradiation, as well as the daily capacity factor are calculated to estimate the quality of the input assumptions in predicting renewable infeed within a typical scheduling horizon.

One common modelling assumption is that a meteorological observable is independent from previous observations. It is demonstrated that the probability of days with high and low solar irradiation and wind speed is considerably underestimated on using the independence assumption. Consequently, probabilities of days with high and low renewable generation are systematically underestimated. It is shown that discrete Markov models that additionally consider one previous observation can ameliorate the goodness of fit and accurately estimate the distribution of infeed.

To quantify the effects of scheduling-time information on the accuracy, a re-forecasting dataset is used to estimate the distribution of meteorological observables given a forecasted value. Although it is demonstrated that scheduling-time forecasts can reduce underestimation of extreme values, their main benefit is found in reducing the uncertainty when predicting the most likely observation. Such a reduction may directly result in an improved economic performance in case reserves can be safely reduced.

By applying two exemplary plant models that estimate the infeed, the effects of the meteorological assumptions on the infeed distribution, and the prediction errors are demonstrated. It is shown that the plant model can influence target metrics such as the prediction error significantly. Hence, the plant characteristics need to be considered when assessing the eligibility of assumptions. By assessing common input models, their implications, and alternative formulations, it is believed that this work aids the design and validation of scheduling algorithms in the context of local energy communities and can help to improve the overall system performance.

KEYWORDS

Local Energy Community, Microgrid, Asset Scheduling, Meteorological Input, Discrete Markov Model, Forecasting Deviation

1 Motivation on Studying Scheduling Assumptions

Local Energy Communities (LEC) and microgrids are often realized as tightly controllable electrical networks within a limited geographical extent [1]. LECs mostly focus on economic aspects and community participation, while microgrids often include resiliency aspects such as fault mitigation. To increase economic benefits and resiliency in presence of highly volatile Renewable Energy Sources (RES), both concepts regularly utilize optimization techniques to schedule controllable generation, energy storage and controllable load in advance [1], [2]. Today, a vast amount of scheduling approaches having various unique propositions already exist. Many of them optimize day-ahead operation of controllable assets.

Since energy production of various renewables such as wind and Photovoltaics (PV) highly depends on environmental conditions, some modeling assumptions must be made to quantify problem production and to formulate the optimal scheduling. Proposed approaches include static interval representations [3], [4], probability distributions of meteorological observables [5], [6], as well as the availability of a forecast including its stochastic deviations [7].

Figure 1 illustrates different formulations of meteorological input assumptions. Although RES models can have a large impact on the system performance, including financial loss and stability issues, e.g., in case production is overestimated, these assumptions are hardly validated in the context of LECs and microgrids [1]. In [8], the assumption on normally distributed wind forecasting errors was disproved and a new approximation function is proposed. However, other assumptions such as the distribution of the forecasted values themselves are beyond the scope. This work addresses these assumptions by studying commonly used modeling assumptions on an exemplary meteorological dataset and aims at guiding the practical implementation of day-ahead scheduling.

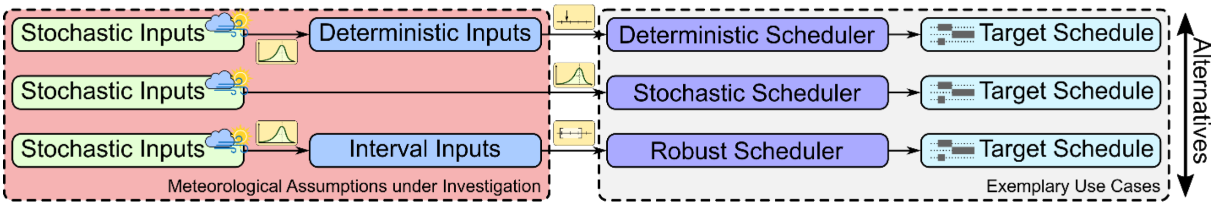

Figure 1: Overview of meteorological input assumptions in scheduling applications

2 Assessing the Quality of Meteorological Models

The performance of various models is evaluated based on meteorological observations that reflect the ground truth of the modeled quantity. To avoid smoothing effects, the case study is based on long-term measurements from a single measurement station instead of relying on aggregated data and satellite observations. Due to the high density of five available measurement stations within the region of Denver, Colorado, which may support future work, and the broadly available long-term datasets starting at 1989, the NREL measurement station [9] was selected. This dataset was divided into two periods, one seven-year training set that reflects past experiences, and one seven-year validation set that is used as ground truth. All models that do not solely operate on ad hoc information such as previous measurements were fitted on the training set without using the validation set for calibration.

Some approaches also assume the availability of a-priori knowledge delivered by a numerical weather forecast. For instance, [7] assumed that forecasting errors are independently, normally distributed. To assess the impact of numerical forecasts, NOAA's reforecast dataset [10] is used to fit a discrete model of measurement values given a certain forecast. It must be noted that due to the coarse temporal and spatial resolution of available data, results are taken as lower bound for forecasting accuracy.

Although other use-cases such as estimating threads exist, meteorological inputs are most commonly used to estimate the energy production of RES [1]. To first cover short-term interdependencies of measurement, without the need of relying on one specific plant model, modeled quantities have been accumulated over one day. This means that instead of directly studying single measurements, accumulated wind speed and solar irradiation values have been used as stochastic target observable.

In addition, two exemplary plant models are used to estimate the impacts of the models under test on the total energy production. The first model uses a power curve to map the wind speed to the power output of a wind turbine. For the sake of simplicity, the curve of an Enercon E-115/3000 turbine [11] was assumed without performing a detailed eligibility analysis. The second model estimates the PV output based on the Global Horizontal Irradiation (GHI), the orientation of the PV panels and the calculated position of the sun. All panels are oriented towards south having a tilt of 30°. The influence of the solar position on the energy output was simulated via the Python PVLib [12]. For both plant models, the daily Capacity Factor (CF), i.e., the ratio of daily generation to the theoretical infeed at constant nominal load, is evaluated under the discussed meteorological inputs. To calculate the CFs, a nominal PV array irradiation of 1 kW/m² is assumed. For the wind turbine, the nominal power of 3 MW was taken. Since the meteorological station does not measure the wind speed at hub height, the wind speed was scaled to 80 m by using the power law [13].

3 Observations on Assessing the Input Models

3.1 Questioning Independent Weather Observations

One commonly made input assumption is that observables such as the wind speed and the solar irradiation follow a certain probability distribution that is independent from previous samples [5], [6], [14]. However, original work [15] indicates that samples within a daily time frame are highly correlated, i.e., the observable of one hour highly depends on the measurement of previous hours. To study the effects of any correlation, two standard distributions of hourly independent samples were selected. According to [5], [6] and [15] which use Rayleigh distribution, a special case of Weibull distribution, Weibull distribution was selected to model the average wind speed of each hour independently. Similarly, the GHI was modeled by one Beta distribution per season and hour [5], [6], [16].

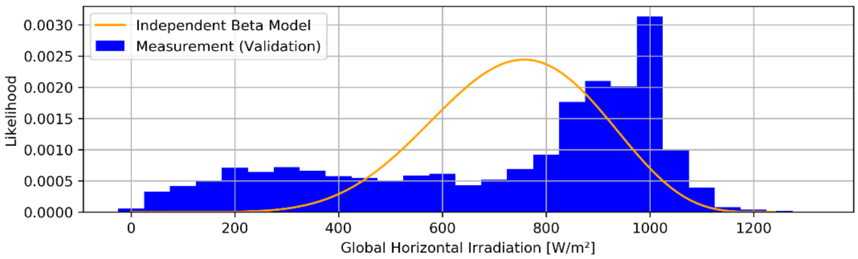

Figure 2: Fitted Beta-PDF and validation data likelihood of summer days at noon

Figure 2 shows the Beta Probability Density Function (PDF) for solar irradiation as well as the histogram of the validation dataset. One can observe that the shape of the Beta PDF does not always follow the weather data. To compensate these effects that are also partly covered in [6], and to focus on the independency assumption, additional temporally independent discrete probability functions were fitted for wind and solar irradiations, respectively.

Figure 3 shows the Cumulative Density Function (CDF) of accumulated daily GHI and wind speed, using the models as well as the reference distribution of the validation partition. Note that due to the integration over one day, the units are given in Wh/m² and km, respectively. To get a typical interval of one forecasted value per hour, all distributions were sampled hourly. For the validation dataset, a temporal resolution of ten minutes was kept to also cover the effects of down-sampling the temporal resolution. Due to the independency assumption, one can see that both models drastically underestimate the probability of having days with low and high irradiation situations. Similar effects are seen on modeling wind speed as observations that are independent from previous ones. As an effect, storage needs may be systematically underestimated.

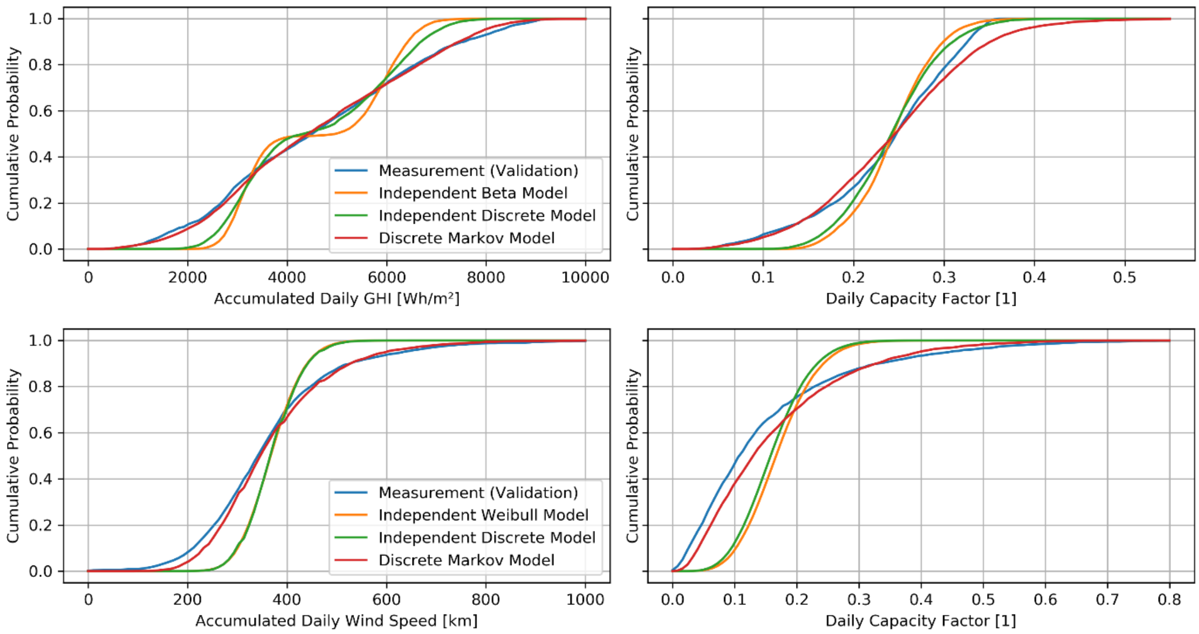

Figure 3: CDFs of daily accumulated observations (independent from scheduling-time information)

On applying the plant models, similar trends are visible. The probability of days with low infeed, denoted by the CF, are systematically underestimated. However, the considered orientation of the PV panels transforms the bimodal shape of the independent solar irradiation models to a CF distribution with one dominant mode. It is observed that the PV model reduced the systematic underestimation of days with high infeed. Specifically, the reference distribution that is computed from the GHI measurements shows a reduced likelihood of high infeed. For wind power plants, trends of underestimating the likelihood of high generation days are also visible on the exemplary CFs.

3.2 Modeling Temporal Dependencies

One of the main limitations of the proposed probability distributions in the context of day-ahead scheduling is the independence assumption. To directly study the effect of independence, the discrete wind and solar irradiation models were extended to discrete Markov models, which take the measurement from the previous hour into consideration. Since the GHI strongly depends on the current hour and season, the discrete Markov model for solar irradiation additionally depends on these inputs and returns the GHI probability given the season, the time, as well as the last observation.

Figure 3 includes the resulting CDF of accumulated wind speed and GHI samples. The Markov model closely follows the distribution of measured samples but does not include any information – such as numerical weather forecasts or past observations – that is only available at scheduling time. As for the independent GHI models, a considerable deviation from the reference distribution that cannot be directly witnessed in the input distribution is observed after applying the PV model. Contrary, the CFs for wind generation follow the expectations from the source CDFs well. Despite the deviations for the Markov GHI models, all Markov models that do not depend on scheduling-time information show fewer deviations to the validation CDF than their independent counterparts. The detailed figures are listed in Table 1 that contains the maximum absolute deviation of each CDF to the reference one. Hence, the table lists the maximum probability difference to the reference distribution. The higher one value is, the more the occurrence of some observation or CF ranges is under or overestimated.

Table 1: Maximum absolute CDF deviations to the validation distribution

Model	PV		Wind	
	GHI [1]	**CF [1]**	**Speed [1]**	**CF [1]**
Independent Beta/Weibull Model	0.1569	0.1290	0.2513	0.3793
Independent Discrete Model	0.1138	0.1055	0.2571	0.344
Discrete Markov Model	0.02735	0.09075	0.05509	0.09227
Independent Forecasting Model	0.0844	0.1247	0.1664	0.2309
Markov Forecasting Model	0.05529	0.1014	0.05215	0.06983

3.3 The Value of Numerical Weather Forecasts

The models given in Figure 3 only use information that is available at the time they were trained. However, when estimating the meteorological inputs, the current state and information from numerical weather forecasts may also be available. To estimate the impact of numerical weather forecasts, the discrete probability of the target observables given the current forecast for the time interval was trained and evaluated on the distinct validation dataset. For the wind speed, one distribution for all samples was fitted. For solar irradiation, one distribution per hour was trained. Results are given in Figure 4 that shows the resulting CDFs for both observables, accumulated solar irradiation and wind speed.

One may note that the accumulated observables (i.e., GHI and wind speed) for the independent forecasting models in Figure 4 follow the reference closer than the independent distributions given in Figure 3. For instance, the maximum absolute GHI CDF deviation for the independent discrete model, as given in Table 1, is eleven percent points while the independent forecasting distribution achieves a deviation of eight percent points. However, after applying the PV plant model, the observation cannot be made anymore.

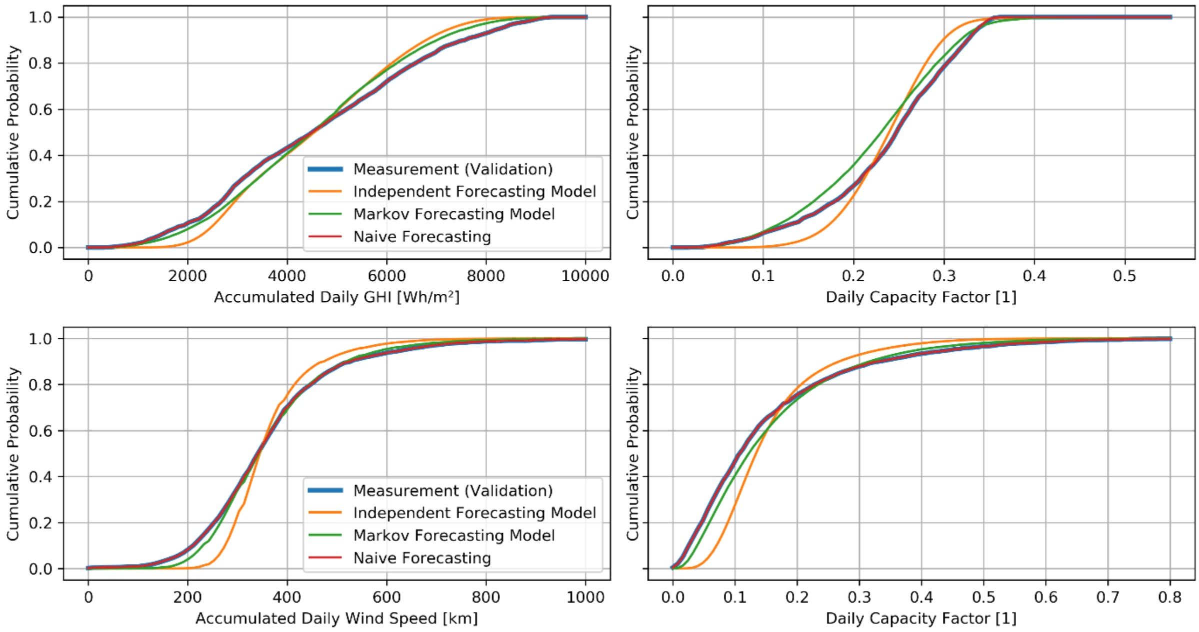

Figure 4: CDFs of Daily Accumulated Observations (Considering Scheduling-Time Information)

Since the forecasting error does not depend on previous errors, the probabilities of high and low wind days are still underestimated, in comparison to the discrete Markov models that are independent of numerical weather forecasts. To model temporal dependencies between forecasting deviations, the probability of an observable given the current forecast, the previous observation, and the current hour for solar irradiation was trained. One can see that the Markov forecasting models follow the reference distribution closer than all independent ones. In terms of the maximum absolute CDF deviation, only the discrete Markov model that does not include scheduling-time information showed better results. For wind speed and the wind CF, the Markov forecasting model outperforms all other models.

3.4 Deterministic Prediction Error

On the one hand, the goodness of fit of an input distribution is one important parameter for stochastic scheduling. On the other hand, deterministic scheduling approaches may only operate on the expected generation given a forecast. For all studied distributions, the expected value is taken as the predicted one and the deviations from the validation dataset are recorded. Additionally, naive forecasts, i.e., the observations from the previous day, are included for reference.

Naturally, the CDF of naive forecasting closely follows the reference CDF in Figure 4 because all samples are simply shifted by one day. Similarly, no bias is found for naive forecasting and the average deviation is asymptotically zero. Figure 5 illustrates the distribution of forecasting errors in estimating the accumulated daily observable. In the given box-plots, median values are marked with an orange line while mean values are pointed out by green triangles. The box itself marks the upper and lower quartile. Both whiskers show the range of the data excluding outliers beyond the 1.5-fold of the interquartile range. The bias for the wind, and GHI models that do not require scheduling-time information can be directly tracked down to a bias in the validation dataset of 11.7 Wh per day and -10.7 km per day, respectively. It is believed that the bias for the accumulated wind speed is caused by changes in the instrumentation, that are considered as a regular phenomenon in long-term operation.

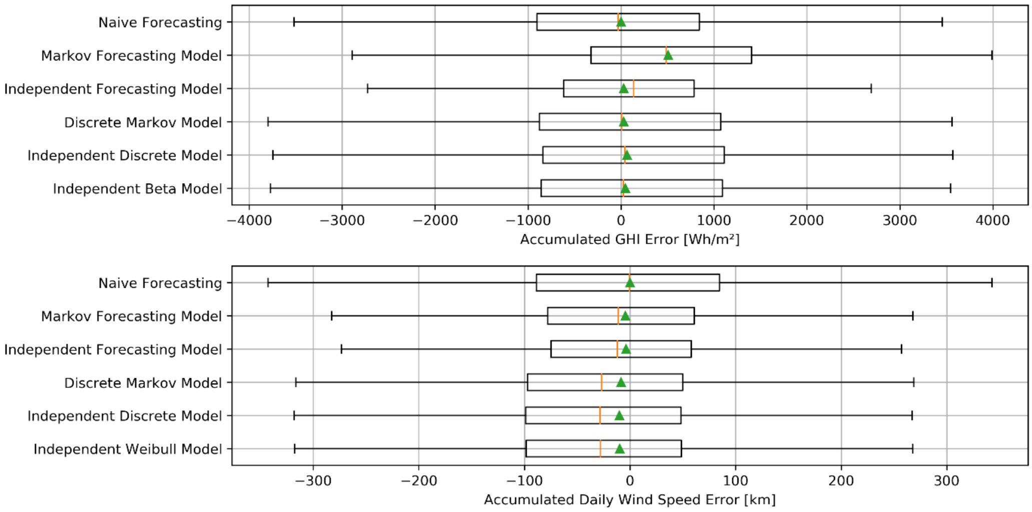

Figure 5: Error in Predicted Observables

For a further comparison, the absolute difference in daily CFs as illustrated in Figure 6 is taken. For both, the validation and the forecasting data, the same plant models were used. Using the Mann-Whitney U test with a significance level of 5%, it is witnessed that before applying the plant models, all distributions that utilize scheduling-time information show a significantly smaller absolute error than the ones that are independent of that information. After applying the plant models, the conclusion does not hold anymore. The independent forecasting still significantly outperforms all other distributions, but the Markov forecasting performs significantly worse than the independent beta model, the independent discrete model, and the discrete Markov model.

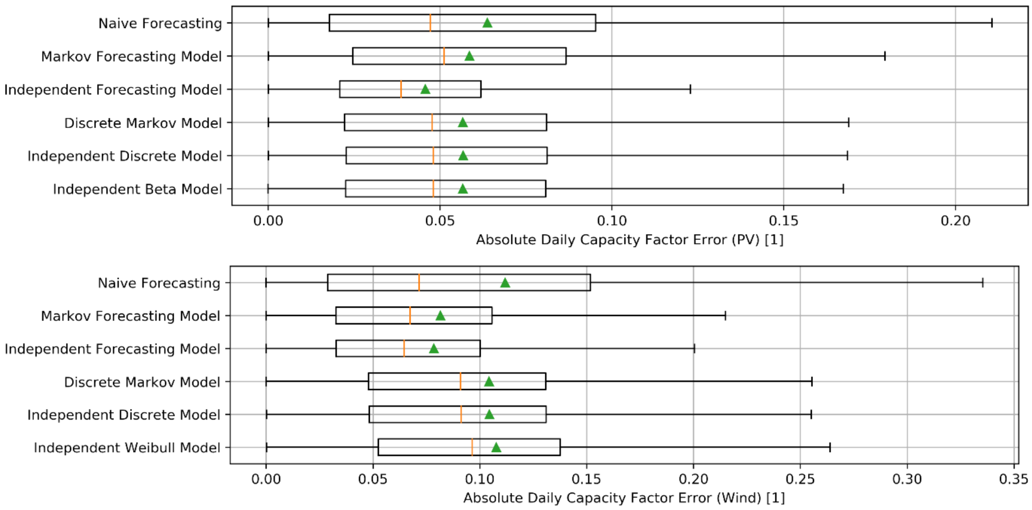

Figure 6: Absolute Error in Predicted Capacity Factors

4 Conclusion and Outlook

The paper indicates the importance of choosing adequate assumptions on meteorological input conditions for modeling day-ahead scheduling problems of geographically contained power systems. A case study demonstrates the effects of modeling RES power production as a series of independent samples. It is shown that the independence assumption leads to a categorical underestimation of days with particularly low and high RES production. Consequently, probabilistic and robust scheduling approaches, which rely on the stochastic assumptions, can suffer from a decreased performance.

The availability of numerical forecasts slightly relaxes the problem and provides effective means for reducing forecasting errors. However, due to the independence assumptions of forecasting errors, underestimation of low production still occurs. A simple, yet effective way of modeling timely interdependence is presented via discrete Markov models. Especially in regions with a high timely correlation of weather phenomena, the Markov model may improve the performance of microgrid and LEC scheduling algorithms.

Exemplary plant models were applied to the meteorological input observables to further assess the impact on the energy system. It is demonstrated that the plant potentially influences the performance measures of studied input models. Hence, decisions on suitable meteorological input models need to consider the target system as well.

Future work needs to study the effects of the model assumption on the outcome of various scheduling algorithms. It is still open to cover a more diverse set of measurement stations in different climate zones. Additionally, special correlations between neighboring sites may be studied to close the gap between geographically local installations and extended distribution systems. Finally, the discretized distributions used in the study may be replaced by suitable parametric continuous distributions to ease analytical analysis and reduce the need for extensive training data. Furthermore, alternative models such as hidden Markov models may be assessed as well.

BIBLIOGRAPHY

[1] S. Parhizi, H. Lotfi, A. Khodaei and S. Bahramirad, “State of the Art in Research on Microgrids: A Review,” *IEEE Access,* vol. 3, pp. 890-925, 2015.

[2] S. M. Nosratabadi, R.-A. Hooshmand and E. Gholipour, “A comprehensive review on microgrid and virtual power plant concepts employed for distributed energy resources scheduling in power systems,” *Renewable and Sustainable Energy Reviews,* vol. 67, pp. 341-363, 2017.

[3] A. Gholami, T. Shekari and S. Grijalva, “Proactive management of microgrids for resiliency enhancement: An adaptive robust approach,” *IEEE Transactions on Sustainable Energy,* 2017.

[4] A. Gholami, T. Shekari, F. Aminifar and M. Shahidehpour, “Microgrid Scheduling With Uncertainty: The Quest for Resilience,” *IEEE Transactions on Smart Grid,* vol. 7, pp. 2849-2858, 11 2016.

[5] F. S. Gazijahani, S. N. Ravadanegh and J. Salehi, "Stochastic multi-objective model for optimal energy exchange optimization of networked microgrids with presence of renewable generation under risk-based strategies," *ISA transactions,* vol. 73, pp. 100-111, 2018.

[6] A. Zakariazadeh, S. Jadid and P. Siano, "Smart microgrid energy and reserve scheduling with demand response using stochastic optimization," *International Journal of Electrical Power & Energy Systems,* vol. 63, pp. 523-533, 2014.

[7] H. Farzin, M. Fotuhi-Firuzabad and M. Moeini-Aghtaie, "A stochastic multi-objective framework for optimal scheduling of energy storage systems in microgrids," *IEEE Transactions on Smart Grid,* vol. 8, pp. 117-127, 2017.

[8] H. Bludszuweit, J. A. Domínguez-Navarro and A. Llombart, "Statistical analysis of wind power forecast error," *IEEE Transactions on Power Systems,* vol. 23, pp. 983-991, 2008.

[9] T. Stoffel and A. Andreas, "NREL Solar Radiation Research Laboratory (SRRL): Baseline Measurement System (BMS); golden, colorado (data)," 1981.

[10] T. M. Hamill, G. T. Bates, J. S. Whitaker, D. R. Murray, M. Fiorino, T. J. Galarneau Jr, Y. Zhu and W. Lapenta, "NOAA's second-generation global medium-range ensemble reforecast dataset," *Bulletin of the American Meteorological Society,* vol. 94, pp. 1553-1565, 2013.

[11] M. Petersen and C. Hofmann, "OEP - supply.wind_turbine_library," 4 2019. [Online]. Available: https://openenergy-platform.org/dataedit/view/supply/wind_turbine_library. [Accessed 24 February 2020].

[12] W. Holmgren, C. Hansen and M. Mikofski, "pvlib python: a python package for modeling solar energy systems," *Journal of Open Source Software,* vol. 3, p. 884, 2018.

[13] S. Emeis, Wind energy meteorology: atmospheric physics for wind power generation, Springer, 2018.

[14] S. Zolfaghari, G. H. Riahy and M. Abedi, "A new method to adequate assessment of wind farms' power output," *Energy Conversion and Management,* vol. 103, pp. 585-604, 2015.

[15] K. Conradsen, L. B. Nielsen and L. P. Prahm, "Review of Weibull statistics for estimation of wind speed distributions," *Journal of climate and Applied Meteorology,* vol. 23, pp. 1173-1183, 1984.

[16] Z. M. Salameh, B. S. Borowy and A. R. A. Amin, "Photovoltaic module-site matching based on the capacity factors," *IEEE transactions on Energy conversion,* vol. 10, pp. 326-332, 1995.