Probabilistic Planning of Minigrid with Renewables and Storage in Western Australia

Weixuan Gao
Dept. of Civil and Environ. Engr.
Stanford University
Stanford, CA 94305
gaow@stanford.edu

Dev Tayal
Visiting Scholar
Stanford University
Stanford, CA 94305
dev.tayal@horizonpower.com.au

Dimitry Gorinevsky
Dept. of Electrical Engineering
Stanford University
Palo Alto, CA 94306
gorin@stanford.edu

***Abstract*—This paper demonstrates analytical method for resource mix optimization. The case study is for an isolated large minigrid of Esperance in Western Australia. The method is based on discrete convolution of conditional probability distributions learned from data. It accounts for demand interdependence with wind and solar generation and for grid storage. The capacity value of these resources is established by computing Loss of Load Probabilities (LOLP) for combined probability distributions.**

***Index Terms*—Capacity value; variable generation resources; storage; reliability; planning; probabilistic analysis; analytical method.**

I. Introduction

Horizon Power, the Government-owned, vertically integrated, monopoly utility operating across regional Western Australia is hard pressed for effective planning of its 38 minigrids to include renewable energy and storage technologies. Maintaining reliability while optimizing the resource mix requires knowing capacity values of these variable resources. The state of the art is to use Monte Carlo analysis, see [1] for wind, [2] for solar, and [3] for storage capacity value.

Planning tools based on Monte Carlo simulation analysis are currently used by RTO/ISO organizations and large utilities but are poorly suited for high penetration of renewables and for smaller grids. The analysis is too labor and computation intensive for the size, number, and resource mix fragility of isolated micro- and minigrids.

This paper demonstrates analytical method devoid of these drawbacks. The method is based on discrete convolution approach, e.g., see [4]. It allows for fast and accurate reliability analysis of resource mix by combining data-driven models for each resource. The known approach convolves probability distributions of independent random variables; it has to be extended in several important ways. First, wind and solar generation as well as load depend on day hours and seasons. Second, even for a given hour and month, wind, solar, and load are not independent. (For example, see [5]). Finally, addressing energy storage is non-trivial.

This work was supported by Stanford Bits and Watts Initiative and Horizon Power.

Though a complete analytical method for resource mix has not been demonstrated earlier, there is prior work on many of the subproblems. Using a data-driven model for each hour of the year is discussed in [6]. For one given hour, it is more reasonable (though still inaccurate) to assume load and variable resources independent and convolve respective distributions. Such approach is used in operational forecasting, e.g., see [7] and [8].

Prior work on storage capacity value is much less extensive. The main approach used is simulation analysis assuming a fixed scenario of storage charge-discharge or optimizing the storage performance as part of simulation, see [9]. No analytical methods for storage capacity value seem to be available, apart from the earlier related work of the authors [10] that is extended in this paper.

The analytical method in this paper combines probabilities models for resources conditional on time regressors (such as hour and month) as described in [10]. These quantile models effectively define the load (or variable generation) duration curves for a given hour. Machine Learning methods used to obtain these models from limited historical data are described in [10]–[12].

The main contribution of this paper is demonstrating the analytical method for integrated resource planning in the Esperance minigrid. Importantly, this paper extends energy balance analysis in [10] to the computation of bounds for storage capacity value. In addition, several extensions of the modeling approach in [10] are demonstrated: combined model for simulation and historical data in Section III, intra-hour variation model in Section IV-A, and quantile bins model in Section III-D.

II. Esperance Minigrid

A. Problem description

The town of Esperance is located on the Southern Ocean coastline of Western Australia, some 720 kilometers (450 miles) from the nearest major city. It supplies electricity to just under $14,000$ people. Esperance is surrounded by national parks, where transmission lines to connect the town with the main grid would be an environmental issue even if they were not cost-prohibitive.

Considered against Horizon Power's total portfolio of isolated minigrids, Esperance has average total supply costs despite its large low-density 33 kV network. Rooftop solar has reached the maximum hosting capacity level for unmanaged installations set at 2.1 MW (an 11% penetration for Horizon Power's capacity).

Horizon Power has peak load of approximately 18.5MW, with a shoulder load of 12.5MW. Being on the Southern Ocean coast, Esperance has a winter peak with 4.7MW difference between winter and summer peak demand, which is atypical of Australia.

Esperance demand has two sources: approximately 75% is town load from residential and business customers and around 25% is local port load. Seven gas turbines in two power stations with total 38.5MW capacity currently provide around 85% of the supply. Existing 2.0MW Ten Mile Lagoon Wind Farm and 3.6MW Nine Mile Beach Wind Farm contribute around 5% and 10% of the generation supply respectively.

Esperance does not currently have any installed storage capacity, but it is considered. Interconnection to the bulk grid is not considered viable now.

III. Modeling

Figure 1 shows the block diagram of the resource mix analysis flow. Machine Learning approaches detailed in [10] are used to build probabilistic models for individual resources from historical data. The analytical method described here then integrates the component models for resource mix optimization. This section discusses the component models for the Esperance minigrid.

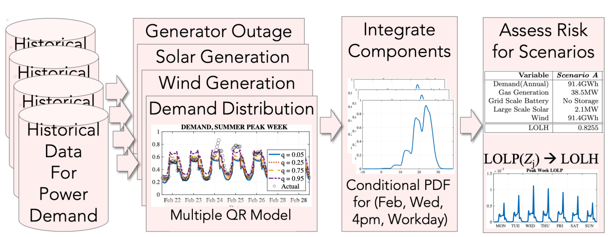

Fig. 1: Block Diagram of Modeling and Analysis Flow

A. Modeling Data

Horizon Power demand is represented by over 3 years of feeder data for the residential load sampled every 5 min. Port demand is modeled as 10% residential load and 90% block load profile with known total annual energy consumption. The block load profile approximates a typical industrial user and is shown in Figure 2 along with residential demand on a summer day in 2015.

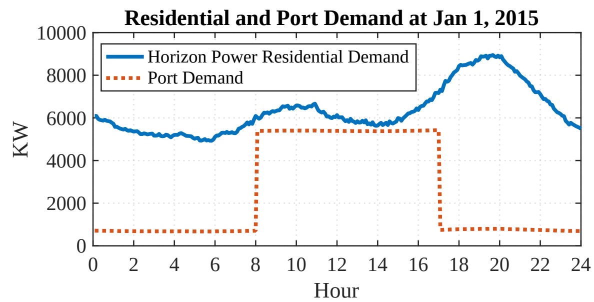

Fig. 2: Residential and Port Demand

Gas plant outages were modeled using NERC value of EFORD $= 11.02\%$ (i.e., the outages occur 11.02% of the time the gas plant should be running).

For wind farm and rooftop solar generation modeling, we used data from Renewables Ninja [13]. Assumed historical generation is based on standard values for efficiency factors and available wind and solar resource.

B. Conditional Probability Model

A distribution for a sampled random variable u can be represented by sampled values

$$p(u[k]) = p[k], \qquad (1)$$

where independent variable samples $u[k]$ are uniformly spaced on a grid covering support interval $[u_0, u_{max}]$.

We consider conditional distributions where regressor vector Z indicates time variables including month, weekday, hour, and holiday. Binary vectors Z include 24 indicators for hour, 7 for weekday, 12 for month, and 2 for holiday/workday. There are total of $24 \times 7 \times 12 \times 2 = 4032$ different regressor states.

Conditional probability model is described by a vector of samples p_i (1) for each state Z_i. Machine Learning methods described in [10]–[12] characterize a conditional distribution of $u|Z$ through sampled Cumulative Density Function (CDF)

$$\mathbf{P}(u \leq v_m|Z) = q_m, \quad v_m = v(Z, q_m), \qquad (2)$$

where quantile levels q_m are uniformly spaced on a grid covering interval $(0, 1)$. Quantile model $v(Z, q_m)$ in (2) can be estimated from data, see below. It can be converted to form (1) as discussed in [10].

C. Demand Model

Quantile Regression (QR) model is one special case of Quantile model. QR model assumes linear influence of each regressor in (2),

$$v(Z, q) = Z\beta(q) + \alpha(q). \tag{3}$$

where $q \in (0, 1)$ is quantile level. Machine Learning methods described in [10] allow building QR models. Such model is given by a set $\{q_i\}_{i=1}^m$ and corresponding QR parameters: vectors $\beta_i \in \mathbf{R}^n$ and scalars $\alpha_i \in \mathbf{R}$.

D. Solar Generation

Quantile Bins (QB) model is another special case of the Quantile model. QB models for each regressor state Z_i are independent. This requires more data then QR model, because there is less statistical averaging.

QB model was developed for Esperance solar generation. The only regressor components that of influence here are month and hour. Thus, there are $12 \times 24 = 288$ different regressors (bins) Z_i. The solar QB model is based on historical data from 2011 through 2015.

Figure 3 shows the predicted solar generation distribution and real points at Oct 21, 2015.

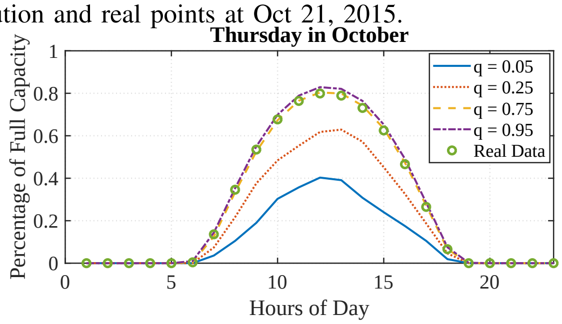

Fig. 3: Solar Generation and Model Quantiles

E. Generator Outage

For a sum of two independent random variables, probability distribution can be computed as discrete convolution of the two sampled probability density functions (PDFs). Assume u and v are two independent variables with sampled PDFs $p_u[\cdot]$ and $p_v[\cdot]$ of the form (1). Then

$$p_{u+v}[\cdot] = p_u[\cdot] * p_v[\cdot], \tag{4}$$

where $*$ denotes discrete convolution.

An example is Outage of multiple gas generators in Esperance minigrid. Outage of each generator is modeled as a sampled PDF based on EFORD parameter. The distribution of the total outage power can be computed as the convolution of the individual generator outage distributions, see [10]. The convolution method for independent variables is well known, see [4]. Its extension for dependent random variables is discussed next.

F. Combined Model of Demand, Wind, and Solar

Demand d, wind generation w, and solar generation s are three interdependent random variables. A nested model (conditional on time regressor vector Z) is built following [10]. The model uses s and w as predictors of d, and s as predictor of w,

$$w = \tilde{w} + \alpha \cdot s, \tag{5}$$

$$d = \tilde{d} + \beta \cdot s + \gamma \cdot w, \tag{6}$$

where $\tilde{w}$ is the prediction residual independent of s, and $\tilde{d}$ is the prediction residual independent of s and w. The coefficient impact factors α, β, γ are conditional on Z.

Multiple Quantile Model including parameters α, β, and γ in (5) and (6)) is built from historical data as described in [10]. The model parameters depend on regressor vector Z. Figure 4-6 shows 3-D plots of α, β, γ vs hour and month. For wind and solar generation, only hour and month components of Z matter.

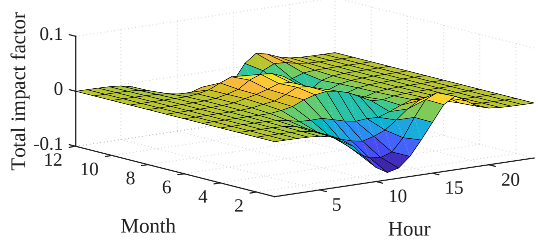

Fig. 4: Solar Impact on Wind, α(Month,Hour)

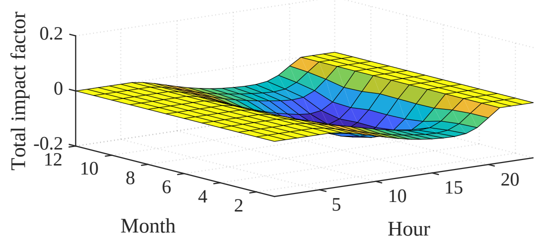

Fig. 5: Solar Impact on Demand, β(Month,Hour)

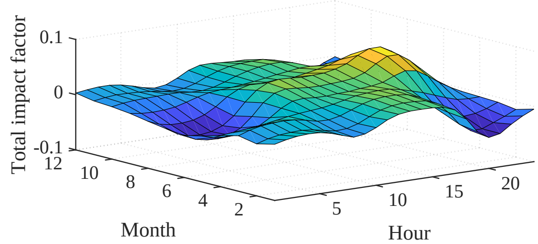

Fig. 6: Wind Impact on Demand, γ(Month,Hour)

Based on (5) and (6) and Subsection III-E, we can combine the PDFs after collecting the terms for three independent variables s, $\tilde{w}$, and $\tilde{d}$ in the sum $w+s-d$,

$$p_{w+s-d}[\cdot] = p_{(1+\alpha-\beta-\alpha\gamma)s}[\cdot] * p_{(1-\gamma)\tilde{w}}[\cdot] * p_{\tilde{d}}[\cdot]. \quad (7)$$

G. Model Validation

How do we know that the developed models, with all the assumptions made in the estimation, have good predictive power for the future years? The answer comes from validating the model on a separate test set.

The data from 2015 through 2016 were used for building dependent variable model (5), (6). The model was then validated with data from 2017. The Pearson's chi-squared goodness-of-fit test was computed for 11-bin empirical probability distribution. Table I shows the validation result, which confirms predictive power of the model.

TABLE I: Model Validation: Goodness-of-Fit Test

Variable	***Dimension***	95% ***Level***	χ^2 ***Statistics***
Demand	10	18.31	15.5810
Wind	10	18.31	5.6285
Solar	10	18.31	6.6206

IV. Power and Energy Reserve Margins

We used the developed probabilistic models to explore three scenarios for the Esperance grid planning. These scenarios are summarized in Table II. Today there are seven gas turbines with total dispatchable capacity of 38.5 MW. In Scenarios B and C, one of the turbines is replaced with variable wind or solar generation, respectively, supported by a 4-hour battery storage system.

TABLE II: Grid Planning Scenarios

Scenario	***Dispatchable***	***Storage***	***Solar***	***Wind***
Today (A)	38.5MW	-	2.1MW	5.6MW
Wind (B)	33.0MW	2MW	2.1MW	18.0MW
Solar (C)	33.0MW	2MW	11.0MW	5.6MW

A. Power Reserve Margin

We define reserve margin R as a random variable

$$R = G - O + S + W - D + V, \quad (8)$$

where G is the dispatchable generation capacity. Other terms are random variables modeled as follows

O is the outage power, see Subsection III-E.

W, S and D are wind generation, solar generation, and total demand, respectively. Subsection III-F describes the model for the sum $W+S-D$.

V is the intra-hour variation of demand from mean hourly value D. It is modeled using 5-minute data. Other variables are assumed to be constant within each hour. This paper extends the formulation in [10] by adding Quantile Bins model of V with month and hour indicators as regressors. Figure 7 shows the quantile model of the intra-hour deviation for Aug 30.

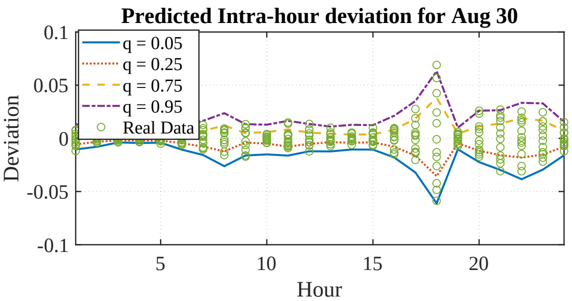

Fig. 7: Quantile Model for Intra-hour Variation

Distribution $R|Z$ in (8) is computed by combining conditional PDFs for independent variables O, $S+W-D$, and V through discrete convolution at each of the 4032=$24 \times 7 \times 12 \times 2$ regressor states Z_i.

B. LOLH Reliability Analysis:

Negative R in (8) means the generation cannot meet the demand. We computed Loss of Load Probability (LOLP) for a given state Z_j as

$$\text{LOLP}_j = \mathbf{P}(R \leq 0 \mid Z_j) \quad (9)$$

Numerical value of LOLP$_j$ can be obtained from the conditional PDF for $R|Z_j$ computed as a convolution of the component PDFs.

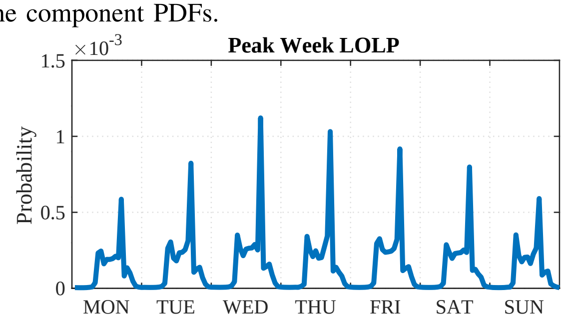

Fig. 8: LOLP$_j$ in the peak week

Figure 8 shows the LOLP for peak week. The expected Loss of Load Hours (LOLH) is (see [10] for detail),

$$\text{LOLH} = 365 \times 24 \times \sum_{i=1}^{N} \text{LOLP}(Z_j) \cdot \mathbf{P}(Z_j) \quad (10)$$

NERC 1-in-10 requirement for capacity adequacy is equivalent to LOLH≤ 2.4.

C. Energy Reserve Margin

The LOLH analysis of the previous subsection can be extended to the scenarios with grid-scale energy storage (battery). Batteries can store energy when excess power is available and generate power when reserve margin is depleted. The battery contribution can be analyzed through *energy margin* introduced in [10]. If we consider a reserve margin (without a battery) as a random variable, the battery can only avert loss of load if the sum of negative reserve margins in n consecutive hours is less than battery energy capacity B_n for any n. The energy margin is defined as

$$R_n = G_n - O_n + S_n + W_n - D_n, \quad (11)$$

where R_n is computed as the convolution of R in (8) with a rectangular filter window of n ones. (Note that $V_n = 0$, since it has zero average for any given hour). To model R_n, the already described methods are used with the component signals in (8) after rectangular window (moving sum) filtering. For a given window width n, $\text{LOLP}_n = \mathbf{P}(R_n \leq -B_n)$. The energy balancing risk can be evaluated through $\text{LOLH} = \max_n\{\text{LOLH}_n\}$. More details can be found in [10]. Since $R_n + B_n \geq 0$ is a necessary condition of energy balancing, the energy margin method yields the lower bound of actual LOLH.

An upper bound of the LOLH can be obtained by assuming a specific operational profile of the battery, which modifies the demand model used for the LOLH estimation in Subsection IV-A. Figure 9 provides battery operation profile assumed for the Esperance minigrid. The battery is fully charged every night from midnight-6am, discharges at a certain rate during daytime, and runs empty at the end of the day.

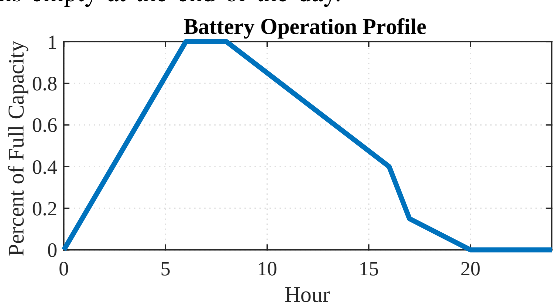

Fig. 9: Net Generation Distribution

In the evaluated scenarios, the upper and lower bounds are not too far apart. This means the presented analysis of energy balance is reasonable and does not overestimate LOLH by much. The result shows how much improvement might be achievable by optimizing the battery operation further compared to the profile in Figure 9.

V. Analysis Results and Conclusions

Table III shows the results of the LOLH analysis for three scenarios in Table II. Since Scenarios B and C include grid-scale battery storage, the upper and lower bounds of LOLH are shown. In Scenario B, replacing one 4.5MW gas turbine out of seven, requires at least 12.4MW of wind nameplate capacity and 2MW of 4-hour storage to keep the same LOLH reliability. In Scenario C, the same result is achieved through 8.9MW of extra solar nameplate capacity and 2MW of storage.

TABLE III: LOLH for Analyzed Scenarios

Scenario	**LOLH**
A	0.66
B	$(0.66, 1.25)$
C	$(0.65, 1.38)$

With an increasing share of renewables in the generation mix, fuel costs and dispatch requirements from the gas plant can be greatly reduced. Full economic analysis is beyond the scope of this paper. Yet one can note that demonstrated reliability analysis is useful for negotiation of long term power purchase agreements, especially as the costs of renewables and storage continue to fall.

References

[1] A. Keane, M. Milligan, et al, "Capacity value of solar power," *IEEE Tran. Power Systems*, vol. 26, no. 2, pp. 564–572, 2011.

[2] C. Dent, R. Sioshansi, J. Reinhart, et al., "Capacity value of solar power," *PMAPS*, Beijing, China, October 2016,.

[3] I. Konstantelos and G. Strbac, "Capacity value of energy storage in distribution networks," *J. of Energy Storage*, vol. 18, 2018, pp. 389–401.

[4] R. Billington and R.N. Allan, *Reliability Evaluation of Power Systems*, Plenum Press, 1984.

[5] N. Krakauer and D. Cohan, "Interannual variability and seasonal predictability of wind and solar resources," *Resources*, vol. 6, no. 3, 2017.

[6] B.A. Frew, W. J. Cole, Y. Sun, et al., "8760-based method for representing variable generation capacity value in capacity expansion models," *NREL Tech Report*, Golden, CO, 2017.

[7] M. A. Matos and R. J. Bessa, "Setting the operating reserve using probabilistic wind power forecasts," *IEEE Tran. on Power Systems*, vol. 26, no. 2, pp. 594–603, 2011.

[8] L. Hirth and I. Ziegenhagen, "Balancing power and variable renewables: Three links," *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 1035–1051, 2015.

[9] R. Sioshansi, S. H. Madaeni, P. Denholm, "A dynamic programming approach to estimate the capacity value of energy storage," *IEEE Tran. Power Systems*, vol. 29, no. 1, pp. 395–403, 2014.

[10] W. Gao and D. Gorinevsky, "Probabilistic balancing of grid with renewables and storage," *PMAPS*, Boise, ID, , 2018.

[11] S. Shenoy and D. Gorinevsky, "Data-driven stochastic pricing and application to electricity market," *IEEE J. of Selected Topics in Signal Processing*, vol. 10, no. 6, pp. 1029–1039, 2016.

[12] S. Shenoy and D. Gorinevsky, "Stochastic optimization of power market forecast using non-parametric regression models," *IEEE PES GM*, Denver, CO, July 2015.

[13] *Renewables Ninja.* Available: https://www.renewables.ninja.