Optimal Storage Sizing using Two-Stage Stochastic Optimization for Intra-Hourly Dispatch

Kyri Baker, *Student Member, IEEE*, Gabriela Hug, *Member, IEEE*, and Xin Li, *Senior Member, IEEE*.

***Abstract*—With the increasing penetration of renewable energy sources into the electric power grid, a heightened amount of attention is being given to the topic of energy storage, a popular solution to account for the variability of these sources. Energy storage systems (ESS) can also be beneficial for load-levelling and peak-shaving, as well as reducing the ramping of generators. However, the optimal energy and power ratings for these devices is not immediately obvious. In this paper, the energy capacity and power rating of the ESS is optimized using two-stage stochastic optimization. In order to capture the wind and load variations in the different days throughout the year, it is advantageous to use a large number of scenarios. Optimizing generator outputs and storage decisions at the intra-hour level with a high number of scenarios will result in a very large optimization problem, and thus scenario reduction is employed.**

A relationship between the variance of the system price for each scenario and the optimal storage size determined for that scenario is shown. The correlation between these parameters allows for a natural clustering of similar scenarios. Scenario reduction is performed by exploiting this relationship in conjunction with centroid-linkage clustering, and stochastic optimization with the reduced number of scenarios is used to determine the optimal ESS size.

I. Introduction

MULTIPLE benefits can be achieved by the use of energy storage systems (ESS) in the power grid. The variability from intermittent energy sources can be balanced with the use of storage; daily peaks of high electricity usage can be shifted to off-peak hours, and the ramping and capacity requirements for conventional generators may be reduced.

Depending on the application, certain storage technologies may be more appropriate for certain purposes. The performance of each of these technologies differ by their charge/discharge rate and maximum energy capacity. Storage technologies include, but are not limited to: pumped hydro, compressed air, flywheels, double-layer or super/ultra capacitors, and batteries (lead-acid, lithium-ion, sodium/sulfur) [1]. In this paper, the focus is on intra-hour generation dispatch to balance out fluctuations in the net load, i.e., demand minus wind generation, of the system. In the considered problem formulation, the storage device is characterized by a maximum energy capacity, maximum power rating and a roundtrip efficiency.

Much of the literature regarding the problem of optimal ESS sizing is based on determining an optimal operating strategy for the storage device [2]- [5], but stochastic optimization techniques are not employed in these papers. Stochastic optimization is used in papers [6]- [8], which focus on an hourly dispatch scale. In [6], wind and load correlation probabilities are used for scenario generation. Scenarios with similar net load shapes and levels are grouped together in [7], and fuzzy clustering is used for scenario generation. In [8], uncertainties in wind and electricity prices are taken into account and a sample average approximation method is used to reduce the dimensionality of the scenarios.

The objective in our paper is to optimally size storage while minimizing generation costs and maximizing the use of renewable energy fluctuations on an intra-hourly scale. We show that there is a relationship between optimal storage size and variance in system price for this application, allowing scenarios which are similar with respect to storage needs to be clustered together and represented by a single scenario. Hence, the clustering operates on the similarities in optimal storage size rather than on similarities on the inputs of the scenarios. This reduced set of scenarios, each weighted with an appropriate probability, is then taken into account in the two-stage stochastic optimization problem, significantly reducing the overall problem size. Hence, two-stage stochastic optimization becomes feasible even for a large number of considered scenarios.

The outline of this paper is as follows: Section II gives the model of the system components and the formulation of the objective function and the constraints. Section III provides insights into the correlation between storage sizing and variance in marginal cost and how this is used to design the clusters to enable scenario reduction. In Section IV, simulation results are shown for a 10-generator system. Finally, in Section V, conclusions and a discussion are given.

II. Modeling

In this section, the optimization formulation and ESS model are provided, along with the notation used for the variables that appear in the paper.

1) Storage Model: The model for the ESS used in this paper is the following:

$$\begin{aligned} E(t+\Delta t) = & \quad E(t) + \eta_c \Delta t P_{in}(t) - \frac{\Delta t}{\eta_d} P_{out}(t), && (1)\\ 0 \leq & \quad E(t+\Delta t) \quad \leq E_{ss}, && (2)\\ 0 \leq & \quad P_{in}(t) \quad \leq P_{ss}, && (3)\\ 0 \leq & \quad P_{out}(t) \quad \leq P_{ss}, && (4)\end{aligned}$$

where $E(t)$ is the energy level in the storage at time instant t. The model incorporates separate variables for charging

The authors are with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, 15213 USA. E-mail: kyri@cmu.edu, ghug@ece.cmu.edu, xinli@ece.cmu.edu.

978-1-4799-5904-4/14/$31.00 ©2014 IEEE

and discharging power, P_{in} and P_{out}, as well as separate constants for the charging and discharging efficiencies, η_c and η_d. Variables E_{ss} and P_{ss} correspond to the energy capacity and the power rating of the storage device. The variable Δt is the time between control decisions. Since the focus in this paper is on intra-hourly economic dispatch, Δt will be set to 10 minutes in the simulations.

2) Cost Function and Constraints: Each scenario corresponds to a 24-hour period of net load, i.e., load minus wind generation. It is assumed that the generators have quadratic cost curves defined by cost parameters a_i, b_i, and c_i, upper and lower limits P_{Gi}^{min} and P_{Gi}^{max}, and ramping limitations R_{Gi}. The economic dispatch optimization problem to be solved for one scenario if storage size and charging and discharging limits are given is as follows:

$$\min_{P_{Gi}(t)} \sum_{t=1}^{N_T} \left(\sum_{i=1}^{N_G} a_i P_{Gi}^2(t) + b_i P_{Gi}(t) + c_i \right) \quad (5)$$

$$\begin{aligned} \text{s.t.} \quad & P_{Gi}^{min} \leq P_{Gi}(t) \leq P_{Gi}^{max}, & (6)\\ & |P_{Gi}(t+\Delta t) - P_{Gi}(t)| \leq R_{Gi}, & (7)\\ & \sum_{i=1}^{N_G} P_{Gi}(t) - P_L(t) + P_W(t) & \\ & \qquad + P_{out}(t) - P_{in}(t) = 0, & (8) \end{aligned}$$

$$\begin{aligned} & 0 \leq P_{out}(t) \leq P_{ss}, & (9)\\ & 0 \leq P_{in}(t) \leq P_{ss}, & (10)\\ & 0 \leq E(t) \leq E_{ss}, & (11)\\ & E(N_T) = E_0, & (12)\\ & E(t+\Delta t) = E(t) + \eta_c \Delta t P_{in}(t) - \frac{\Delta t}{\eta_d} P_{out}(t), & (13) \end{aligned}$$

with $t = 1, ..., N_T$ for all constraints and $i = 1, ..., N_G$, where N_T is the number of steps in the optimization horizon and N_G is the number of generators in the system. The generation output for generator i at time step t is given by $P_{Gi}(t)$, total wind generation by $P_W(t)$ and total load by $P_L(t)$. The initial energy level in the storage device is set to E_0. As described in Section III-A, P_{ss}, E_{ss} and E_0 all become variables in the two-stage stochastic optimization problem and problem formulation (5) - (13) corresponds to the second stage problem.

III. Methods

This section discusses the methods used in the proposed approach to determine the optimal size of a storage device.

A. Two-Stage Stochastic Optimization

Stochastic optimization is a technique which minimizes the total cost over a chosen number of scenarios while accounting for uncertainties in the problem. In two-stage stochastic optimization [9], there are first stage variables, common to all scenarios, and second stage variables, specific to each scenario and dependent on the first-stage variables. A set of scenarios are generated and a probability is assigned to each of these scenarios. The stochastic optimization problem is then formulated to find the optimal solution for all variables while taking into account the probabilities for each of the scenarios.

With regards to the storage sizing problem, the first-stage variables are the storage parameters E_{ss}, P_{ss}, and E_0, and the second-stage variables are the generation values, charging/discharging rate of the storage, and the energy level of the storage. This concept is illustrated in Figure 1.

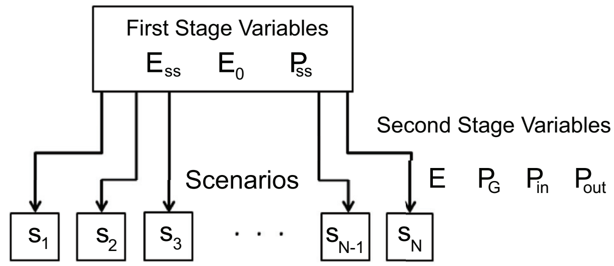

Fig. 1. Two-stage stochastic optimization concept for the storage sizing problem

The overall problem formulation for the two-stage stochastic problem is given by:

$$\min_{P_{Gi}(s,t)} \sum_{s=1}^{N_S} \left(w_s \cdot T_L \cdot \sum_{t=1}^{N_T} \left(\sum_{i=1}^{N_G} a_i P_{Gi}^2(s,t) + b_i P_{Gi}(s,t) + c_i \right) \right) + dE_{ss} + eP_{ss} \quad (14)$$

$$\begin{aligned} \text{s.t.} \quad & P_{Gi}^{min} \leq P_{Gi}(s,t) \leq P_{Gi}^{max}, & (15)\\ & |P_{Gi}(s,t+\Delta t) - P_{Gi}(s,t)| \leq R_{Gi}, & (16)\\ & \sum_{i=1}^{N_G} P_{Gi}(s,t) - P_L(s,t) + P_W(s,t) & \\ & \qquad + P_{out}(s,t) - P_{in}(s,t) = 0, & (17)\\ & 0 \leq P_{out}(s,t) \leq P_{ss}, & (18)\\ & 0 \leq P_{in}(s,t) \leq P_{ss}, & (19)\\ & 0 \leq E(s,t) \leq E_{ss}, & (20)\\ & E_0 = E(s, N_T), & (21)\\ & E(s,t+\Delta t) = E(s,t) + \eta_c \Delta t P_{in}(s,t) - \frac{\Delta t}{\eta_d} P_{out}(s,t). & (22) \end{aligned}$$

Hence, constraints of this optimization problem are equivalent to those given in (6)-(13), but now with distinct variables P_{Gi}, P_{out}, P_{in}, and E as well as values P_L and P_W for each scenario $s = 1...N_s$. Variables E_{ss}, P_{ss}, and E_0 are not dependent on t or s. These variables are common to all scenarios; their optimal values are calculated while taking into account all of the considered scenarios simultaneously. The constant values w_s correspond to the probability of occurrence of scenario s. The constants d and e correspond to the cost parameters of the storage device with respect to the capacity and charging speed, respectively, and T_L is the expected lifetime of the storage in number of days.

It is advantageous to determine what factors directly impact the optimal solution for the storage sizing problem, so individual scenarios which result in a similar optimal storage size may be grouped together and a new representative scenario for that cluster is chosen and weighted accordingly. It is obvious that the more scenarios that are considered in the problem, the more accurate the frequency of certain cases of wind and load in the system will be represented. Thus, it is desirable to utilize as many scenarios as possible. However, the number of variables and constraints increases tremendously with increasing number of scenarios rendering stochastic optimization computationally very intensive.

B. Relationship Between Optimal Capacity and Variance in System Price

The price of electricity is determined by the marginal cost of generation, i.e. the cost to generate one additional unit of power in the system. As the Lagrange Multiplier of the power balance equation (8) corresponds to the sensitivity of the objective function, in this case the overall generation cost, with respect to a change in this equation, i.e. a change in load, the incremental cost is equal to the value of this Lagrange Multiplier. In the following, we refer to this as the system price and denote it by λ_p.

Due to the fact that load and infeeds from renewable generation vary significantly throughout the course of the day, the system price also varies over the day. The variance of the system price, measured over the time period of one day, is defined as:

$$var(\lambda_p) = \frac{1}{N_T}\sum_{t=1}^{N_T}(\lambda_p(t) - mean(\lambda_p))^2. \tag{23}$$

For illustration purposes, we show the correlation between the system price variance and the optimal storage size for a small system with three conventional generators and two wind generators. First, the economic dispatch problem without storage is solved for a range of different scenarios. This corresponds to optimizing for (5) and including constraints (6) and (7) and the power balance equation (8) but without the charging/discharging from the storage for each of these scenarios. The resulting variance in system price for each of the scenarios is stored. Next, the optimization problem with storage as a variable, i.e. (14)-(22) is solved for each scenario separately. Hence, only one single scenario is taken into account in the optimization problem each time and the optimal storage size and charging/discharging rate are determined as if this is the only occurring scenario. The resulting variances in system price and the optimal storage sizes are plotted in Figure 2.

It is clear from the figure that the optimal capacity of the storage is positively correlated with the variance in system price for that scenario. That is, the bigger the variance in price over that scenario, the more beneficial it is to use storage. This can be attributed to a great extent to the quadratic cost functions of the generators. Changes in the level of generation

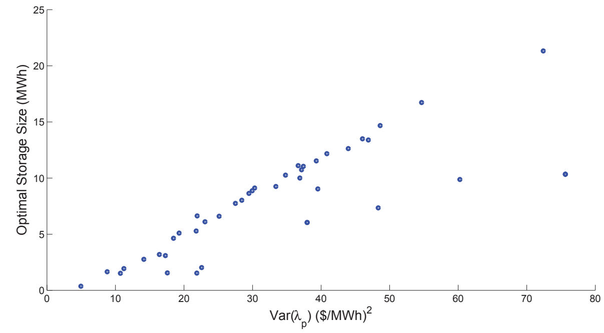

Fig. 2. Correlation between marginal price and optimal storage size for system with three conventional generators and two wind generators.

and therefore ramping are implicitly penalized because of this quadratic cost. Storage helps to alleviate the ramping of these generators, thus lowering the overall cost of generation. In the presence of an increased penetration of intermittent sources such as wind, the required ramping increases thereby increasing the value of storage. However, no direct correlation was found between the optimal storage capacity and variance in wind or load. The variance in system price was found to be the strongest indicator with respect to optimal storage size for the considered problem formulation with quadratic cost functions.

This dependency has important implications on how to cluster scenarios. To show how this is different from clustering scenarios based on similarities in net load, we show the optimal storage sizes for five scenarios and their respective net load curves in Figures 3 and 4 for a system of 8 conventional generators and two wind generators. The two scenarios with the closest optimal capacity/variance in system price are scenarios 2 and 3. It is interesting to note that the net load for scenarios 2 and 3 have a very large difference in magnitude. Scenario generation by methods which assume that scenarios with similar net load values produce similar optimal storage capacities may therefore lead to suboptimal storage sizes for the considered problem formulation and cost function. Consequently, we propose to use the correlation between system price and optimal storage size as a means to cluster scenarios.

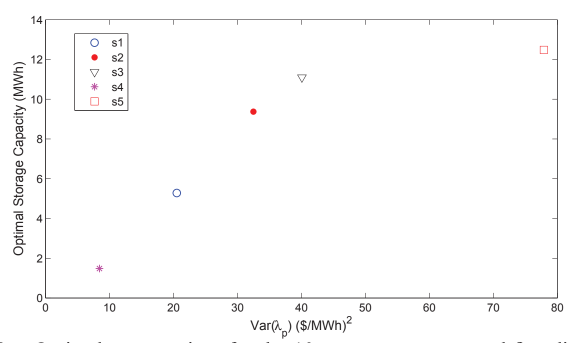

Fig. 3. Optimal storage sizes for the 10 generator system and five different scenarios.

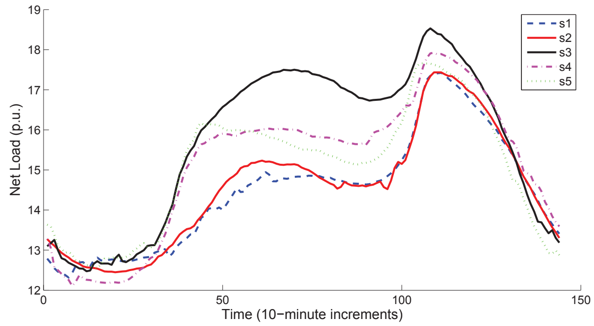

Fig. 4. Net load of five scenarios in the 10 generator system.

C. Scenario Reduction Using Centroid-Linkage Clustering

As the number of scenarios considered in the optimization increases, the more accurately the distribution of possible realizations of net load are represented. However, this also increases the problem size to unmanageable levels, especially on a 10-minute dispatch scale. Scenario reduction techniques have been employed for the energy storage sizing problem, e.g. in [6] and [7]. However, as described earlier, these techniques focus on load/wind correlations and net load analysis to determine similarity between scenarios.

Here, scenario reduction is performed by utilizing the discovered relationship between the optimal storage capacity and variance in system price. We use a hierarchical centroid-linkage clustering method [10] to form clusters of similar scenarios. In this clustering technique, each scenario is first considered to be a separate cluster, and clusters are subsequently combined into larger clusters until the desired number of clusters is obtained. Hierarchical clustering is chosen over other conventional clustering methods because other methods may group outlier clusters with other clusters instead of keeping them distinct, which is desirable in our application. At each iteration of the process, centroids, which correspond to the mean of all data points in their respective cluster, are calculated. The centroid c_i for cluster i is therefore defined as:

$$c_i = \frac{1}{N_i} \sum_{k \in \Omega_i} x_k, \tag{24}$$

where Ω_i includes the set of points $\mathbf{x_i} = [var(\lambda_p);\ E_{ss}]$ included in cluster i and N_i is the number of points in cluster i. Next, the Euclidean distances between all possible cluster pairs (i, j) where $i \neq j$, are determined and compared. The pair that minimizes $\|c_i - c_j\|_2$ is combined into a new cluster m where the data points $\mathbf{x_m} = \mathbf{x_i} \cup \mathbf{x_j}$. This process is repeated until the desired number of clusters is achieved. Next, a representative scenario is chosen for each cluster. This scenario $\hat{\mathbf{x}}_\mathbf{i}$, for each cluster i, is chosen to be the one that is closest to the mean of that cluster; i.e.,

$$\hat{x}_i = \underset{x_k \in \Omega_i}{\operatorname{argmin}} \ \|x_k - c_i\|_2 \,. \tag{25}$$

For the considered application, each scenario corresponds to a specific realization of load and wind generation for one day. Each of these scenarios results in one data point in the correlation between storage size and variance in marginal cost. The clustering technique is then used to cluster this two dimensional data into a pre-defined number of clusters.

As an example, 150 scenarios were run on the 8-generator, 2-wind plant system and these scenarios were grouped into 10 clusters. In Figure 5, the result of the clustering is shown. The number of scenarios in each cluster determines the probability of the resulting representative scenario of that cluster, as shown in Figure 6.

The scattering is related to the fact that different sets of generators reach their limits for different levels of net load, and therefore a different generator is setting the system price. I.e., in one case with a large amount of wind, a coal plant that was usually producing at capacity for most of the other scenarios was not at capacity. However, even with the scattering multiple linear trends can be observed.

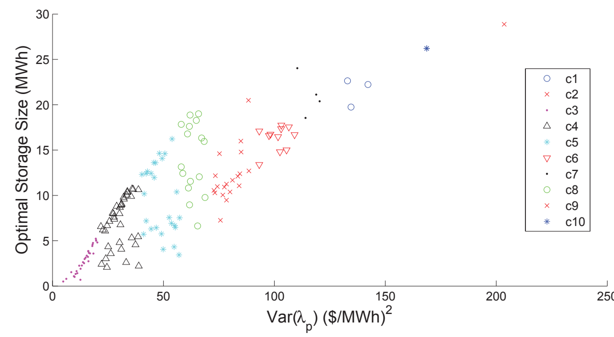

Fig. 5. Correlation and clustering for the 10 generator system, 150 scenarios and 10 clusters.

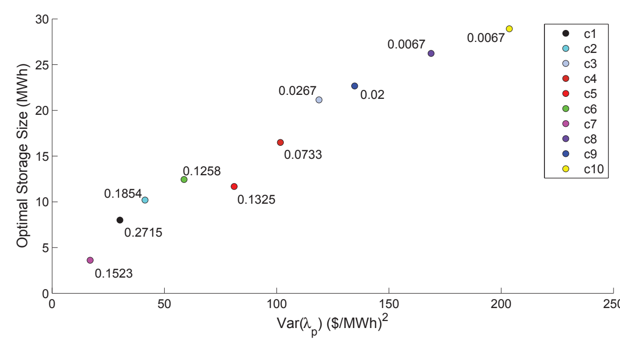

Fig. 6. Representative clusters weighted by probabilities for the 10 generator system.

D. Overall Proposed Approach

An overview over the proposed approach to determine the optimal sizing of the storage device is given in Figure 7. First, the economic dispatch problem in (5) - (13) is solved without any storage device in the system for every single scenario separately in order to determine the variance in marginal generation cost before the deployment of storage. Then, the problem in (14) - (22) is solved separately for each scenario with storage as a variable, determining the optimal E_{ss}, P_{ss}, and E_0 for that scenario. Next, the scenarios are clustered into groups and representative scenarios are chosen for each cluster. This reduced set of scenarios is used in the stochastic problem formulation (14) - (22) to determine the overall optimal storage size.

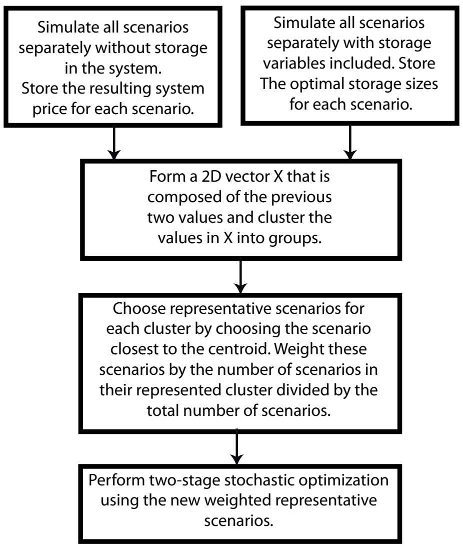

Fig. 7. Flowchart of the overall algorithm.

IV. Simulation Results

In this section, we first give an overview over the simulation setup and then discuss the simulation results.

A. Simulation Setup

Simulations were performed on a power system with eight conventional generators and two wind power plants. The 10-minute demand data for 150 days was taken from ISO New England [11] and the data for the wind outputs was taken from the Eastern Wind Integration Transmission Study (EWITS) [12]. Results are given for various levels of wind energy penetration. The chosen cost function data and capacities for the generators are given in Table I. The storage technology used in these simulations has a roundtrip efficiency of 95% and capital costs of the energy capacity and power converter size as $\$10,000/kWh$ and $\$500/kW$, respectively. Reasonable parameters were taken from the list of costs and efficiencies for various storage technologies described in [1]. The storage is assumed to be operating without degradation for the assumed time period T_L and generation costs are minimized over a period of 20 years; i.e., $T_L = 20 \cdot 365$.

B. Simulation Results

Taking into account the full set of 150 scenarios each corresponding to a day's worth of data already results in over $237,600$ variables for the considered formulation. Hence, clustering of scenarios is indispensable and we use the proposed clustering technique to reduce the number of scenarios to 70, 50, 30 and 10 clusters and compare the results to the "full" case of 150 scenarios. The optimal solution for the storage size with the above parameters and a 20% penetration of

TABLE I
Generator parameters.

Generator	a ($/MW/MWh)	b ($/MWh)	c ($/h)	Capacity
Gas	0.76	15	370	250 MW
Coal	0.0079	18	772	330 MW
Nuclear	0.00059	5	240	350 MW
Gas	0.76	13	370	250 MW
Coal	0.0133	18	440	340 MW
Nuclear	0.00059	5.2	240	350 MW
Coal	0.014	18	772	330 MW
Coal	0.0078	17.7	440	330 MW

TABLE II
Simulation results for varying cluster sizes.

No. Clusters	Optimal E_{ss}	P_{ss}	E_0	Time
10	7.1749	1.4822	1.0914	133s
30	8.281	1.726	0.2313	159s
50	8.258	1.837	0.3347	239s
70	8.1601	1.8373	0.3884	353s
100	8.1254	1.8024	0.4133	646s
150	8.0685	1.7911	0.4377	1342s

wind energy for each number of clusters is given in Table II. The computations were performed using the IBM ILOG CPLEX Optimizer [13] through MATLAB 2012a on an Intel i7 processor with 32 GB of RAM. The main purpose of showing the computation time is not to give an indication of how fast the solution can be computed in absolute values but to provide a way to demonstrate the effectiveness in reducing the problem size by the means of scenario reduction. As seen from the results, as the number of clusters decreases, the deviation from the presumably optimal solution given by the 150 scenarios case increases. However, even with a reduction to 30 clusters, the solution is very close to that optimal solution and it only takes a fraction of the time to compute the solution.

The value of using stochastic optimization can be analyzed in comparison with other methods. In Table III, the stochastic solutions for various numbers of clusters are compared with using a simple weighted average of representative scenarios. The "weighted average" of clusters as shown in the third column of the table refers to the average of optimal E_{ss} sizes from the representative scenarios weighted by their probability as calculated from scenario reduction. In the case of 150 scenarios, the "weighted average" does not include representative scenarios, but rather refers to the average of the optimal solution of the original 150 scenarios, each with equal probability. It can be seen that performing stochastic optimization with a reduced set of scenarios results in a storage size which is significantly closer to the solution of the overall stochastic optimization than if just averaging is used.

TABLE III
Optimal E_{ss} size using various cluster sizes and techniques.

No. Clusters	Stochastic Solution	Weighted Average
10	7.1749	11.02
30	8.281	9.869
50	8.258	9.5817
70	8.1601	9.9234
100	8.1254	9.806
150	8.0685	13.35

The optimal amount of storage in the system can be analyzed for various levels of wind penetration. In Figure 8, the optimal amount of storage capacity for 100 scenarios is shown for varying levels of wind penetration. Here, wind energy penetration level is defined as the percentage of demand in terms of energy that is supplied by wind energy on average over all considered scenarios. Table IV lists the results of stochastic optimization with increasing number of clusters for 0%, 10%, and 20% of wind energy penetration. As seen from the figure and table, the optimal amount of storage increases with the level of penetration. This can be attributed to the fact that a higher penetration of wind results in more variation in the net load, making it more beneficial to deploy a larger storage device.

TABLE IV

OPTIMAL E_{ss} SIZE FOR DIFFERENT LEVELS OF WIND PENETRATION.

No. Clusters	0% Wind	10% Wind	20% Wind
10	3.285	4.675	7.818
30	3.049	4.650	7.942
50	3.103	4.656	7.612
70	3.068	4.635	7.287
100	3.054	4.636	7.546

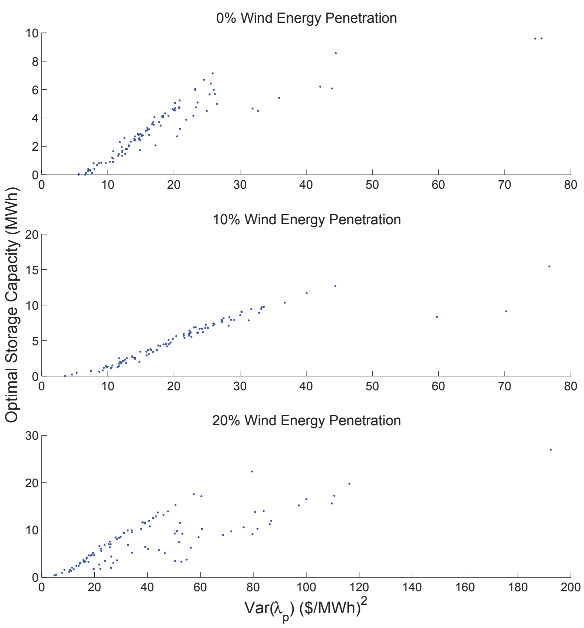

Fig. 8. Optimal storage capacity for various levels of wind penetration.

V. CONCLUSION

In this paper we presented a method for determining the optimal storage size in a power system focusing on a 10 minute economic dispatch. A correlation between optimal storage capacity and variance in the system price is found for the case in which generation costs are modeled as quadratic functions. Based on this relationship, a scenario reduction method was derived which accurately captures the similarities in storage needs across scenarios. It uses centroid linkage clustering to group scenarios with similar needs in storage into a reduced number of clusters. This set of clusters is then taken into account in the two-stage stochastic optimization problem significantly reducing the overall problem size. The problem structure was defined in a way which allows for an integration of various storage technologies on different time scales.

Future work will not only address the question of how much storage, but also where the storage should be placed. By incorporating line constraints, the value of storage increase as it allows for the alleviation of congestions. For that purpose, the relationship between the optimal storage location and variance in locational marginal price at particular buses will be examined.

ACKNOWLEDGEMENT

This work was supported in part by the Grid Technologies Collaborative, funded by the National Energy Technology Laboratory (NETL), and the SYSU-CMU Joint Institute of Engineering. The authors are very grateful for this financial assistance that made this project possible.

REFERENCES

[1] S. Schoenung. Energy Storage Systems Cost Update: A Study for the DOE Energy Storage Systems Program. Technical report, Sandia National Laboratories, 2011.

[2] S. Chakraborty, T. Senjyu, H. Toyama, A.Y. Saber and T. Funabashi, "Determination methodology for optimising the energy storage size for power system," IET Generation, Transmission and Distribution, vol. 3, pp. 987-999, Aug 2009.

[3] Oudalov, A.; Cherkaoui, R.; Beguin, A.; "Sizing and Optimal Operation of Battery Energy Storage System for Peak Shaving Application," IEEE PowerTech, Lausanne, 1-5 July 2007.

[4] P. Harsha and M. Dahleh, "Optimal sizing of energy storage for efficient integration of renewable energy, in *Proc. 50th IEEE Conf. Decision and Control and Eur. Control Conf.*, Dec. 2011.

[5] Brekken, T.K.A.; Yokochi, A.; Von Jouanne, A.; Yen, Z.Z.; Hapke, H.M.; Halamay, D.A., "Optimal Energy Storage Sizing and Control for Wind Power Applications," *IEEE Transactions on Sustainable Energy*, vol.2, no.1, pp. 69-77, Jan. 2011.

[6] Chad Abbey and Gza Jos, "A Stochastic Optimization Approach to Rating of Energy Storage Systems in Wind-Diesel Isolated Grids", *IEEE Trans. Power Syst.*, vol. 24, No. 1, pp. 418-426, Feb. 2009.

[7] P. D. Brown, J. A. P. Lopes, and M. A. Matos, "Optimization of pumped storage capacity in an isolated power system with large renewable penetration," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 523-531, May 2008.

[8] Zhen Shu; Jirutitijaroen, P., "Optimal sizing of energy storage system for wind power plants," IEEE Power and Energy Society General Meeting, July 2012.

[9] J. Birge, F. Louveaux, *Introduction to stochastic programming*, Springer-Verlag, New York, 1997.

[10] Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 1998; 95: 14863-14868.

[11] ISO New England Historical Five-Minute Data. http://www.iso-ne.com/markets/5min_data/index.html. Last accessed Aug. 8, 2013.

[12] Eastern Wind Dataset. http://www.nrel.gov/electricity/transmission/eastern_wind_methodology.html. Last accessed July 10, 2013.

[13] "IBM ILOG CPLEX Optimization Studio." [Online]. Available: http://www-01.ibm.com/software/integration/optimization/