RELIABLITY MODELING AND SIMULATION OF COMPOSITE POWER SYSTEMS WITH RENEWABLE ENERGY RESOURCES AND STORAGE

A Dissertation

by

HAG-KWEN KIM

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Chanan Singh
Committee Members,	Alex Sprintson
	Jim Ji
	Lewis Ntaimo
Head of Department,	Chanan Singh

August 2013

Major Subject: Electrical Engineering

ABSTRACT

This research proposes an efficient reliability modeling and simulation methodology in power systems to include photovoltaic units, wind farms and storage. Energy losses by wake effect in a wind farm are incorporated. Using the wake model, wind shade, shear effect and wind direction are also reflected. For solar modules with titled surface, more accurate hourly photovoltaic power in a specific location is calculated with the physical specifications. There exists a certain level of correlation between renewable energy and load. This work uses clustering algorithms to consider those correlated variables. Different approaches are presented and applied to the composite power system, and compared with different scenarios using reliability analysis and simulation. To verify the results, reliability indices are compared with those from original data.

As the penetration of renewables increases, the reliability issues will become more important because of the intermittent and non-dispatchable nature of these sources of power. Storage can provide the ability to regulate these fluctuations. The use of storage is investigated in this research.

To determine the operating states and transition times of all turbines, Monte Carlo is used for system simulation in the thesis. A conventional power system from IEEE Reliability Test Systems is used with transmission line capacity, and wind and solar data are from National Climatic Data Center and National Renewal Energy

Laboratory. The results show that the proposed technique is effective and efficient in practical applications for reliability analysis.

ACKNOWLEDGEMENTS

I would like to express sincere appreciation to my committee chair, Dr. Chanan Singh, and committee members, Dr. Alex Sprintson, Dr. Jim Ji, and Dr. Lewis Ntaimo, for their guidance and encouragement during the course of this research work.

Also, I express my indebtedness to my parents and relatives, friends, and colleagues, for their constant encouragement and moral support.

Finally, financial assistance at Texas A&M University through the Graduate Scholarship is gratefully acknowledged. The work in this dissertation was partly supported from a research project from Power Systems Engineering Research Center.

NOMENCLATURE

CAES	Compressed Air Energy Storage
ENSI	Energy Not Supplied Interruption
ET	Equation of Time
EENS	Expected Energy Not Supplied
FGFCM	Fast Global Fuzzy C Means
FGKM	Fast Global K Means
FCM	Fuzzy C Means
FPCM	Fuzzy Probabilistic C Means
GFCM	Global Fuzzy C Means
GKM	Global K Means
KM	K Means
LOLE	Loss of Load Expectation
LSTM	Local Standard Time Meridians
NREL	National Renewable Energy Laboratory
NCDC	National Climatic Data Center
PV	Photovoltaic
PCM	Probabilistic C Means
PHES	Pumped Hydro Energy Storage
RTS	Reliability Test Systems
RHS	Right Hand Side

WAAAP	Wind Atlas Analysis and Application Program

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Renewable energy is becoming more pervasive as it is a sustainable resource and there is no carbon dioxide emission. For these reasons, many countries have set targets to include a significant share of wind and solar power into their energy portfolios. However, wind or solar power is unstable because of its intermittent and fluctuating characteristics. So a detailed reliability analysis and estimation of power systems is critical for the growing penetration of such resources. This research has focused on several problems to make this analysis closer to the reality.

The first problem is development of more accurate models of the wind farms. A model based on birth and death process was proposed in [1]. It only considers transition behaviors between adjacent states. This model would be accurate if the wind was modeled on a continuous basis. However, the wind speed data is collected and used on basis of intervals like 10 minutes to an hour. During these periods transitions are also possible to other states. This research proposes a more accurate model as shown in Figure 1 called an exact transition rate model [2]-[3] which demonstrates state changes over time with transition rates. The number indicates a state, and arrows indicate possible transitions between states. Transition behaviors are determined by transition rates from a state to another state. This approach considers all transition behaviors between wind speed or direction states from an actual wind data in wind farm. The probability of each wind state can be derived by using the transition rates. However this model cannot incorporate the correlation between load and wind speed, even though it

uses all transitions between the various wind speed or direction states. So this method is appropriate for modeling of conventional generating turbines or wind turbines if there were no correlation between the renewable energy and the load but not for the wind or solar power when such a correlation does exist. As efficient data mining methods, clustering approaches [4]-[5] for considering correlation are presented and compared in this work to solve this problem. Several different clustering methods are introduced and applied to an example power system with wind farm and photovoltaic (PV) system. This research shows an efficient methodology for the modeling and simulation of the system using clustering which is able to keep the correlation between renewable energy and load, reducing the original data size.

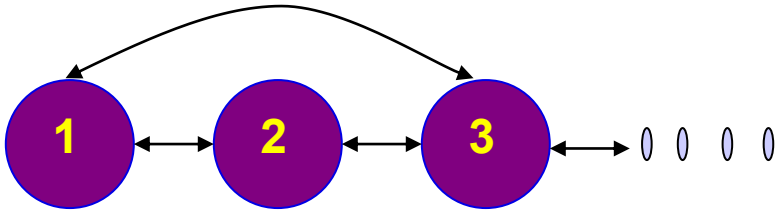

Figure 1. Transition Rate Matrix Method

Another problem dealt in this work is the more accurate modeling of energy production from wind and solar farms. Because of the energy conservation principle, the wind speed (energy) entering a turbine is higher than that leaving it, since turbines generate electricity from the entering wind. During this process, in the area behind a

turbine, leeside, turbulent flow occurs. This power loss is called wake effect [6]-[8]. Because of this effect, downstream wind turbines generate less power, because of lower wind speed. In general, for a free standing wind turbine, there is no wake effect by other wind turbines. However, when a turbine is located in a relatively close spacing with others, wake effect can have a significant influence on the calculations. As the effect gets stronger, loadability of wind farms decreases. Therefore, it is essential to consider the wake effect and to examine its impact on the wind farm reliability and economics for more accurate and reliable power calculation. PV array consists of a number of PV modules connected in series or parallel. Its power depends on a radiation from solar energy which is changed by the absorption, diffuseness, or reflection of clouds and atmosphere dusts. Physical characteristics of a module determine actual PV power over time. For more accurate calculation of solar power generation, this work considers inclination of solar modules and cloud effect. The third problem considered here is that of the role of storage. Integration of storage into the system improves loadability of the system by providing extra energy at an appropriate time as well storing energy when there is surplus. A methodology for optimal storage deployment is also proposed and studied in this work.

Monte Carlo Simulation [9]-[10] is used for system simulation. This dissertation uses both random sampling and the next event form of sequential method and the results are compared. The wake effect is incorporated with each method and reliability indices [11]-[14] such as Loss of Load Expectation (LOLE), Expected Energy Not Supplied (EENS), and Energy Not Supplied Interruption (ENSI) are used for quantitative

reliability analysis. Wind data is taken from the National Renewable Energy Laboratory (NREL) [15] and National Climatic Data Center (NCDC) [16], and load data are from IEEE Reliability Test Systems (RTS) [17].

2. SIMULATION AND ESTIMATION OF RELIABILITY IN A WIND FARM CONSIDERING WAKE EFFECT*

This section describes the improvements made in the models of wind farms. One of these is the inclusion of wake effect models and the second is the development of more accurate state transition model.

2.1 Wake Effect Models

Wind energy is becoming more pervasive as it is a sustainable resource and there is no carbon dioxide emission. For these reasons, many countries have set targets to include significant share of wind and solar power into their energy portfolios. However, wind power is unstable because of its intermittent and fluctuating characteristics. So a detailed reliability analysis and estimation of the impact of wind power systems is critical for the growing penetration of such resources.

Because of the energy conservation principle, the wind speed (energy) entering a turbine is higher than that leaving it, since turbines generate electricity from the entering wind. During this process, in the area behind a turbine, leeside, turbulent flow occurs. This power loss is called wake effect [6]-[8]. Because of this effect, downstream wind turbines generate less power, because of lower wind speed. In general, for a free

* Reprinted with permission from "Simulation and Estimation of Reliability in a Wind Farm Considering the Wake Effect" by H. Kim, C. Singh and A. Sprintson, April 2012, *IEEE Transactions Sustainable Energy,* vol. 3, no. 2, pp. 274-282, Copyright [2012] by IEEE.

standing wind turbine, there is no wake effect by other wind turbines. However, when a turbine is located in a relatively close spacing with others, wake effect can have a significant influence on the calculations. As the effect gets stronger, loadability of wind farms decreases. Therefore, it is essential to consider the wake effect and to examine its impact on the overall wind power system reliability and economics for more accurate and reliable wind power calculation. This dissertation shows that wake effect does make an impact on the reliability of a wind farm and quantifies this impact.

There are many types of wake models [18]-[20] for the wind speed modeling and analysis. The generally accepted models, N.O. Jensen model [21]-[22], Eddy Viscosity (J.F. Ainslie) model [23], and G.C. Larsen model [24] are adopted, and modified in this work. Jensen and Larsen model are based on the kinematic model employing momentum equation. Eddy model comes from the field model using flow field on a wind farm. From original wake models, this dissertation develops newly modified wake models by wind shade and wind shear effect.

Wake effect relies on wind direction as well as wind speed. So by varying wind direction, upstream and downstream turbines are updated over time so that waked speed of downstream turbines is calculated using proposed wake models. So wind direction also plays an essential role on power loss by wake effect in a wind farm generation as well as wind speed data. So every hour wind speed and direction determine waked speed in a wind farm.

2.1.1 N.O. Jensen Model

The N.O. Jensen model was first developed by N.O. Jensen in 1984. It is a simple model with linearly expanding wake effect. Figure 2 shows the schematic for the model description and waked speed from natural undisturbed wind is generated using (1).

$$v_w = v_f\{1-(1-\sqrt{1-C_t})(\frac{d}{d_x})^2\} \quad (1)$$

where x [m] is an axial distance between turbines, v_f [m/s] is undisturbed wind speed, C_t is thrust coefficient, v_w [m/s] is the waked speed, $d_x = d + 2kx$, and k is a wake decreasing constant. It is assumed to be 0.075 for onshore in this dissertation. Because of the combined wake effect by adjacent turbines, Equation (1) is modified to (2).

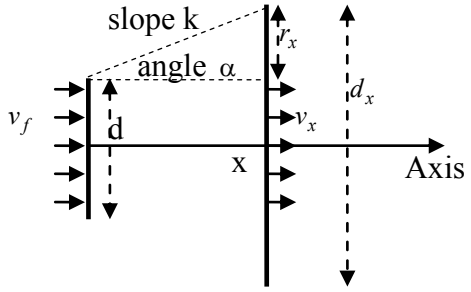

Figure 2. Schematic of N.O. Jensen model

$$v_w = v_f\{1-\sum_i^n(1-\sqrt{1-C_t})(\frac{d}{d_{xi}})^2\} \quad (2)$$

where n is the number of upstream turbines for a downstream turbine, and d_{xi} [m] is the wake region by upstream turbine i.

Wind speed is also affected by adjoining natural terrains or artificial constructions, which is called the wind shade effect [22]. In less than about 1 [km] of the atmosphere layer, wind is becoming stronger, as the height goes up. This is the wind shear effect [25]. By taking into account for each effect, (2) is modified to (3) and (4).

$$v_w = v_f\{1-\sum_i^n (1-\sqrt{1-C_t})(\frac{d}{d_{xi}})^2(\frac{A_{si}}{A})\} \quad (3)$$

$$v_w = v_f\{1-\sum_i^n (1-\sqrt{1-C_t})(\frac{d}{d_{xi}})^2(\frac{A_{si}}{A})\}\frac{\ln(h/h_o)}{\ln(h_{ref}/h_o)} \quad (4)$$

where $A\,[m^2]$ is the rotor disc area and $A_{si}\,[m^2]$ is the shades area by upstream turbine i, depending on wind direction. And h_o [m] is a roughness length [26], h_{ref} [m] is the reference height, and h[m] is the height of the downstream turbine. For a wind farm with turbines of different height, (4) is useful to apply.

2.1.2 Eddy Viscosity Model

Eddy Viscosity model was first constructed by J.F. Ainslie in 1988, solving the equivalent thin shear layer approximated from Navier-Strokes equation. It is assumed that the wake behavior is axis symmetric and stationary so that the model has two dimensional description of field. The waked speed is given in (5.1).

$$v_w = v_f[1 - B\exp\{-3.56(\frac{r}{b})^2\}] \tag{5.1}$$

$$B = C_t - 0.05 - (16C_t - 0.5)(TI/10) \tag{5.2}$$

$$b = \sqrt{3.56C_t/\{4B(2-B)\}} \tag{5.3}$$

where r[m] is a radial distance and TI is turbulence intensity.

Similarly, by considering cumulative wake effect, wind shade and shear effect, waked speed in a wind farm is derived by (6).

$$v_w = v_f[1 - \sum_i^n B\exp\{-3.56(\frac{r_i}{b})^2\}(\frac{A_{si}}{A})]\frac{\ln(h/h_o)}{\ln(h_{ref}/h_o)} \tag{6}$$

where r_i[m] is a radial distance from upstream turbine i.

2.1.3 G.C. Larsen Model

G. C. Larsen model was first developed by G.C. Larsen in 1988. It is based on Prandtl's turbulent boundary layer equation. In this dissertation, the first order approximation solution is described. Following equation shows the mathematical expression for waked speed calculation.

$$v_w = v_f[1 - \frac{1}{9}\{C_t A(x + x_o)^{-2}\}^{\frac{1}{3}}\{r^{1.5}(3c_1{}^2 C_t A x)^{-0.5} - (\frac{35}{2\pi})^{0.3}(3c_1{}^2)^{-0.2}\}^2]. \tag{7.1}$$

$$x_o[\mathrm{m}] = \frac{9.5d}{(\frac{2R_{9.5}}{D_{eff}})^3 - 1} \tag{7.2}$$

$$c_1 = (0.5D_{eff})^{2.5}(\frac{105}{2\pi})^{-0.5}(C_t A x_o)^{-\frac{5}{6}} \tag{7.3}$$

$$D_{eff}[\mathrm{m}] = d\sqrt{\frac{1+\sqrt{1-C_t}}{2\sqrt{1-C_t}}} \tag{7.4}$$

$$R_{9.5}[\mathrm{m}] = 0.5\{R_n + \min(h, R_n) \tag{7.5}$$

$$R_n[\mathrm{m}] = \max\{1.08d, 1.08d + 21.7d(TI - 0.05)\} \tag{7.6}$$

Similarly, cumulative wake effect, wind shade and shear effect are incorporated as shown in (8).

$$v_w = v_f[1 - \sum_i^n \frac{1}{9}\{C_t A(x_i + x_o)^{-2}\}^{\frac{1}{3}}\{r_i^{1.5}(3c_1{}^2 C_t A x_i)^{-0.5} - (\frac{35}{2\pi})^{0.3}(3c_1{}^2)^{-0.2}\}^2 \frac{A_{si}}{A}]\frac{\ln(h/h_o)}{\ln(h_{ref}/h_o)} \tag{8}$$

where x_i [m] is an axial distance from upstream turbine i.

2.2 Wind Turbine Output Power

In this work, power generated by wind energy conversion systems is derived using general power curve [27] of a wind turbine, given by (9).

$$
\begin{aligned}
P_{out} &= 0, && \text{for} \quad 0 \le V \le V_{ci} \\
&= (A + BV + CV^2)P_r, && \text{for} \quad V_{ci} \le V \le V_r \\
&= P_r, && \text{for} \quad V_r \le V \le V_{co} \\
&= 0, && \text{for} \quad V \ge V_{co}
\end{aligned} \tag{9}
$$

where parameters A, B, and C are given by following equations (10.1), (10.2), and (10.3). They depend on cut in speed and rated speed, respectively.

$$A = \frac{1}{(V_{ci} - V_r)^2}[V_{ci}(V_{ci} + V_r) - 4(V_{ci}V_r)(\frac{V_{ci} + V_r}{2V_r})^3] \tag{10.1}$$

$$B = \frac{1}{(V_{ci} - V_r)^2}[4(V_{ci} + V_r)(\frac{V_{ci} + V_r}{2V_r})^3 - (3V_{ci} + V_r)] \tag{10.2}$$

$$C = \frac{1}{(V_{ci} - V_r)^2}[2 - 4(\frac{V_{ci} + V_r}{2V_r})^3] \tag{10.3}$$

where P_{out} is power output generated, V_{ci} is cut in speed, V_r is rated speed, V_{co} is cut out speed, and P_r is rated power.

At cut in speed, turbines begin generating power and then power increases nonlinearly with the speed. From rated speed to cut out speed, turbine keeps generating rated power, and above cut out speed, turbine is shut down for the equipment safety.

2.3 Transition Rate Matrix Approach

Wind farm model consists of wind turbine model and wind speed model, as shown in Figure 3. In conventional power systems, two or three state models are used to represent a generation unit [28], representing fully available, forced outage and forced

derated conditions. However, in wind turbine model, for each wind speed level, power generated is assigned to the fully available or forced outage state of a turbine. Figure 3 shows the state space of a wind turbine for each wind speed state. The circles in the left box represent wind speed states, and the squares in the right box indicate wind turbine states. The arrows between the circles represent transitions between different wind speed states and the arrows between the squares are failure and repair transitions of turbines with specific rates. State Up is properly working state, and state Down is out of service state. The lines between the circle and the squares indicate that for a given wind speed, a turbine could be either up or down. The transition rates between the up and down states could be represented as a function of the wind state as the failure rate may depend on the wind speed. It can be seen that the possible capacity states is two times the number of wind speed states. The aggregate state of wind farm is updated with sampling so that the corresponding power is determined on simulation time.

For wind speed model, reference [1] suggests birth and death Markov chain. This is based on a stochastic process [29] where each state of a system transits as a function of time. In the Markov chain, each state moves to the next neighboring state through birth and death process. If the sampling time is small i.e., is close to zero, wind speed can be considered to change smoothly over time. In this situation, the model suggested by [1] would represent the physical reality correctly. However, in practice sampling is done at intervals like 10 minutes. In such a situation speed cannot be assumed to transit smoothly and transitions to remote states can occur more frequently. So if the birth and death model is used for a wind data sampled at finite intervals as is the case in practice,

some transitions between wind states can be lost. Then, resultant steady state probabilities will not be correct. This dissertation introduces a transition rate matrix method [2]-[3] by which all possible transition rates between states from original wind data can be captured. The resultant probability will be very close to each state's fraction of total operation time. From the given wind data, transition rates between any set of states are calculated using (11).

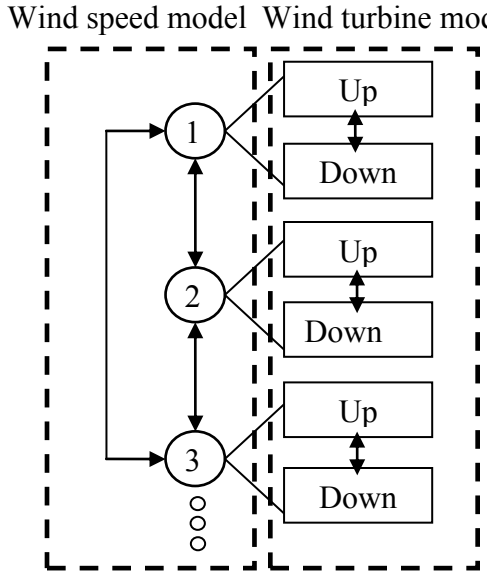

Figure 3. State Space of a Wind Turbine

$$\lambda_{ij} = \frac{F_{ij}}{P_i} = \frac{\frac{N_{ij}}{T}}{\frac{T_i}{T}} = \frac{N_{ij}}{T_i} \tag{11}$$

where λ_{ij} [#/s] is transition rate from state i to j, F_{ij} [#/s] is frequency from state i to j, P_i is time in state i as fraction of total time, N_{ij} [#] is number of transition from state i to j, T [s] is total time, and T_i [s] is duration time of state i.

Using calculated transition rates, steady state probability of each state is derived by the following algorithm [3].

1. Find transition rate matrix, A.

   where element $a_{ij} = \lambda_{ij}$ for $i \neq j$

   $$a_{ii} = -\sum_{j} \lambda_{ij} \quad \text{for } i = j$$

2. Transpose matrix A to A^t.
3. In A^t, replace the elements of a randomly chosen row k by one and call this B.
4. Find state probabilities using P=inv (B)*C

where P is steady state probability matrix, C is column vector such that kth element is one and the others are all zero, and inv (B) is the inverse matrix of matrix B.

In the model described, wind speed states are defined and then corresponding power for a turbine is determined depending on its status of up or down. It should be noted that the transition rate matrix developed in this approach using equation 10 represents average transition rates and the matrix when solved will provide correct average probabilities over the period of study. This approach is useful when load and wind can be assumed to vary independently of each other.

2.4 Case Studies and Results

To illustrate application of the methodology discussed in this dissertation, a simple wind farm system shown in Figure 4 is used. There are 16 identical wind turbines. Symbol o indicates the placement of a wind turbine which is located at the center of hub. The layout consists of 6d by 6d square spacing for wind turbine where d is the rotor diameter of a turbine. In practice, spacing between turbines is represented in terms of rotor diameter. One conventional unit is also added to the wind farm. The power generated by this unit is characterized by failure and repair rates regardless of wind speed. Its capacity is 40 [MW] which is the same as the capacity of the wind farm. Table 1 shows the total system capacity and peak load. To reduce fluctuations of wind power, a conventional generating unit is added into the proposed wind farm. The load data comes from IEEE RTS [17].

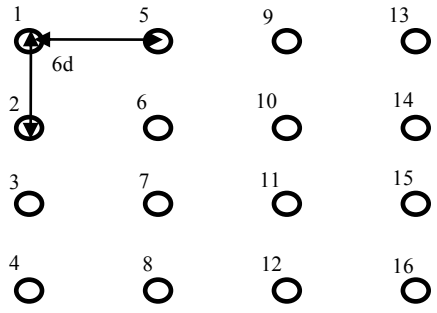

Figure 4. Layout of a Wind Farm

Table 1. System Capacity and Peak Load for Wind Farm

Generation [kW]		Peak load [kW]
Wind farm	**Conventional unit**	
2500 *16=40000	40000	30000

Tables 2 and Table 3 show wind data and wind turbine data respectively. Wind speed and direction data comes from Western Wind Resources, National Renewable Energy Laboratory (NREL) [15]. The location is Texas station number 1. The data starts from 01/01/2006 12:00 am and goes up to 12/31/2006 11:50 pm. In the extreme weather (above 19m/s), failure rates increase as described in Table 4. This is based on the information from many studies that have shown that as the wind speed goes up, failure rates also increase [30]-[31]. The given wind speed data varies with sampling time of 10 [min].

Table 2. Wind Speed Data

Wind speed data	
Peak wind[m/s]	33.01
Minimum wind[m/s]	0.27
Mean wind[m/s]	7.79
Standard deviation[m/s]	3.35
Sampling time[min]	10
Number of samples	52560

Table 3. Wind Turbine Data

Wind turbine data	
Cut in speed[m/s]	6
Rated speed[m/s]	11
Cut out speed[m/s]	19
Rated power[kW]	2500
Rotor diameter[m]	80
Hub height[m]	70

Table 4. Transition Rates of a Wind Turbine for Different Speed

	Conventional unit	**Wind turbine**	
Weather		**Normal speed**	**Extreme speed(>19m/s)**
Failure rate[#/yr]	6	6	36
Repair rate[#/yr]	130	130	36

It should be noted that there can exist periods of calm during which wind is continuously below 5 [m/s]. From the given data, it is observed that calm periods of 5 hours or above occur on 36 days. If transition rate matrix approach is used, then using (11), the calm period would increase the probability of wind speed states 1 and 2 and decrease the outward frequency from these states. Thus the average probability of the

states will be accounted for correctly but using this approach the correlation of load with the calm period can not be captured using this approach.

Wind direction is shown by wind rose with sampling time of 10 [min]. The wake effect on wind the farm depends on the changing wind direction over time. By embedding the wake effect with varying wind direction, resultant overall power of wind farm is computed. Suppose that wind blows from the west at one time. Then, from the given layout of this wind turbine farm in Figure 4, turbines 1, 2, 3, and 4 have no waking effect. For turbine 9, its upstream turbines are, 1, 2, 3, 4, 5, 6, 7, and 8. However, for upstream turbines, 2, 3, 4, 6, 7, and 8, the shade area is zero. So, turbine 9 is actually influenced by waking upstream turbines 1 and 5, according to equation 5. By virtue of layout and wind direction, each turbine in the following groups has the same wake effect: {5, 6, 7, and 8}, {9, 10, 11, and 12} and {13, 14, 15, and 16}.

Wind speed states are first defined as shown in Table 5. Then, frequency and duration of each state is also calculated. Using the general power curve, we can calculate power output for each state. Power generated gets bigger as the wind speed increases above the cut in speed and over the rated speed, power is maintained. Above the cut out speed, there is no power generated.

Tables 6 and Table 7 describe differences between the birth and death Markov chain and the exact transition rate method. For example, for state 1, transitions to state 3, 4, 5, 6, 7, and 9 are neglected in the birth and death model. Basically state probabilities by the birth and death model do not have all transitions, since transitions between neighboring wind speed levels only are considered. On the other hand, probabilities

using exact transition rates cover all the transitions. So by using the exact transition rates, more accurate results are obtained.

Table 5. Wind Speed States for Wind Farm

State	Range[m/s]	Freq[#/yr]	Dur[yr/#]	Power[kW]
1	0-4	407	0.00030	0
2	4-5	921	0.00009	0
3	5-6	1128	0.00009	0
4	6-7	1210	0.000093	215.26
5	7-8	1253	0.000094	668.94
6	8-9	1236	0.000094	1153.5
7	9-10	1119	0.000087	1668.94
8	10-11	920	0.000080	2215.26
9	11-12	719	0.000081	2500
10	12-13	543	0.000071	2500
11	13-14	379	0.000065	2500
12	14-15	252	0.000054	2500
13	15-16	166	0.000050	2500
14	16-17	123	0.000053	2500
15	17-18	74	0.000070	2500
16	18-19	48	0.000064	2500
17	19-20	32	0.000039	0
18	20-21	23	0.000031	0
19	21-34	13	0.000109	0

Table 6. Transitions from State 1 for Different Approaches

Birth and death Markov chain		**Exact transition rate method**	
Transitions	**Transition rates[#/y]**	**Transitions**	**Transition rates[#/y]**
1-2	2925	1-2	2925
		1-3	208
		1-4	56
		1-5	24
		1-6	32
		1-7	8
		1-9	8

In the wake model, the thrust coefficient which is a function of wind speed is needed. In general, it is provided by the manufacturer of wind turbines. In this work, it is calculated using the Wind Atlas Analysis and Application Program (WAAAP) [22], one of the wind analysis tools for wind turbines or wind farms. Thrust coefficient curve is shown in Figure 5. At the cut in speed, it has a steep rise, and then it decreases as wind speed goes up. This means that the wake effect is relatively weak at high wind speed. A west wind is considered for examining the wake effect.

To find the changes of wind states for different wake models in details, let us observe a case of south wind direction. When the wind blows from the south, upstream turbines are {4,8,12,16} and downstream turbines are {3,7,11,15},{2,6,10,14}, and {1,5,9,13} from Figure 4. Figure 6, Figure 7 and Figure 8 show the state changes of

downstream turbines for three different models. Arrows indicates the change of the state. Red color shows only one state drop, blue dotted ones show two state drops. Wind speed drop by wake effect relatively decreases, as the wind speed grows up. For example, for turbines {1,5,9,13} in Figure 6, states from 4 to 11 drop by two states. However, states from 12 to 16 falls only one state. And for high speed beyond state 17, there is no change by wake effect. Turbines {1,5,9,13} show more wake loss than {3,7,11,15} regardless of wake models. This is because that for turbines {1,5,9,13} more wake effects are influenced by around upstream turbines as the combined wake effect increases, shown by Figure 4.

For turbines {3,7,11,15}, waked speed from different models have a similar value. However, for turbines {1,5,9,13}, the state changes from Jensen and Eddy model have a similar value, which is a little bit different from Larsen model. Basically, in the original Larsen model (7.1), cumulative wake effect by neighboring several upstream turbines and wind shear effect are neglected. So it is appropriate for situations like neither cumulate wake effect nor wind shear effect.

As these effects are included on the original Larsen model, waked speed from modified model (8) shows different value from that by two other wake models. So it means that for a wind farm with cumulative wake effect or wind shear, modified Larsen model has some deficiencies as a wake model.

Table 7. Probability of Each State Using Different Approaches

Range [m/s] for each state		Fraction Probability	Birth and death model Probability	Transition rate matrix Probability
1	0-4	0.12477	0.15555	0.12482
2	4-5	0.08622	0.10243	0.08621
3	5-6	0.10664	0.11825	0.10660
4	6-7	0.11305	0.11846	0.11303
5	7-8	0.11860	0.11801	0.11860
6	8-9	0.11672	0.11021	0.11670
7	9-10	0.09754	0.08725	0.09750
8	10-11	0.07372	0.06334	0.07369
9	11-12	0.05878	0.04900	0.05877
10	12-13	0.03858	0.03142	0.03859
11	13-14	0.02486	0.01985	0.02488
12	14-15	0.01377	0.01026	0.01378
13	15-16	0.00840	0.00580	0.00842
14	16-17	0.00658	0.00421	0.00660
15	17-18	0.00519	0.00303	0.00520
16	18-19	0.00308	0.00150	0.00309
17	19-20	0.00127	0.00062	0.00127
18	20-21	0.00072	0.00029	0.00072
19	21-34	0.00142	0.00045	0.00143

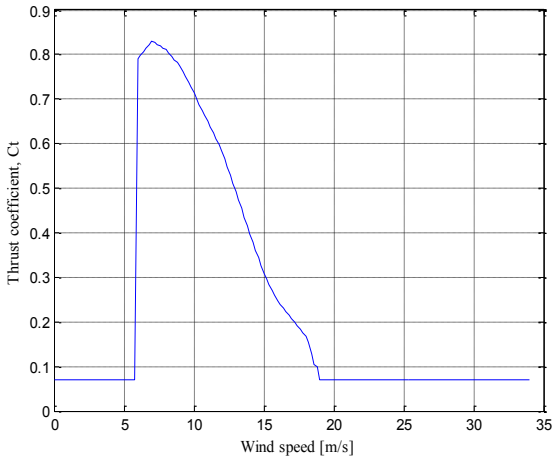

Figure 5. Thrust Coefficient Curve

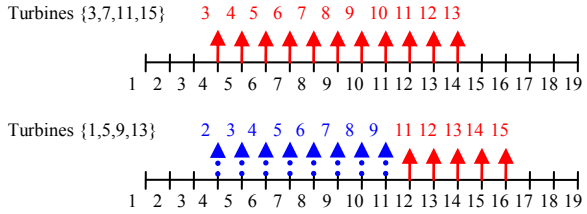

Figure 6. State Changes of Turbines Using Jensen Model

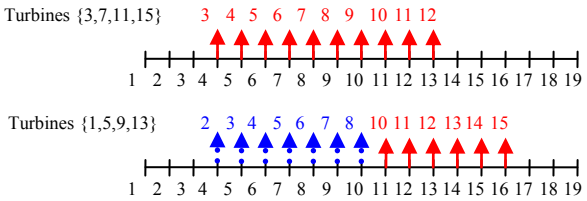

Figure 7. State Changes of Turbines Using Eddy Model

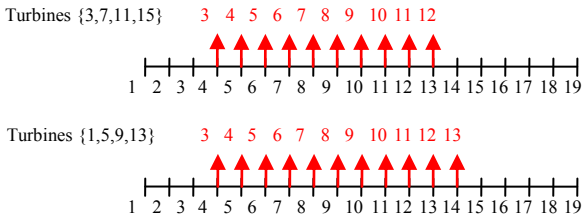

Figure 8. State Changes of Turbines Using Larsen Model

It is assumed that the hub height of turbines 1, 5, 9, and 13 is changed from 70 [m] to 150 [m]. It is assumed that roughness length [26] is 0.03 [m]. This value is used for a typical landscape which is the open agricultural area without fences and very scattered buildings. Then using model (4), wind state changes are shown in Figure 9. Similarly, there is wake effect on the whole. Overall, wind states drop less than Figure 6, because of wind shear effect. Here is an interesting point. States 14, 15, and 16 in Figure 9 goes up even with the wake effect, although they go down from Figure 6 or Figure 7. This is because wind shear effect removes energy losses by wake effect at high speed. In practice, to build up turbines with 150 [m] hub height is uneconomical because of construction costs. However, equation (4) is used effectively for wind farm with changing ground heights or with different hub heights of turbines.

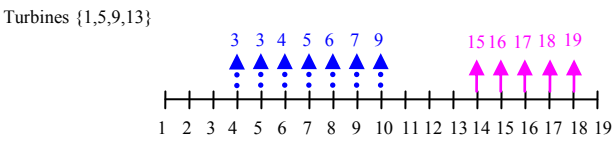

Figure 9. State Changes of Turbines Using Jensen Model with Shear Effect

For Monte Carlo simulation, random sampling and next event are used as non-sequential and sequential approach, respectively. Table 8 shows reliability indices with wake and without wake using the birth and death model. As wake model, Jensen model using (3) is used to calculate waked speed. Using changing wind direction data, wake effect is applied for the proposed wind farm. The results from random sampling are almost the same as those from next event. As would be expected, when wake effect is included, reliability level drops.

Table 8. Reliability Indices Using Birth and Death Markov Chain

Random sampling			**Next event**	
Indices	**Without wake**	**With wake**	**Without wake**	**With wake**
LOLE[h/y]	146.82	163.90	150.98	162.28
EENS[kWh-yr]	278.81	362.53	291.82	350.30

Table 9 is based on the transition rate method using Jensen model. The exact transition rate approach is considered more accurate, since all transition rates are included. LOLE and EENS are lower when the full transition rate matrix is used as all wind speed transitions are properly accounted for.

Table 9. Reliability Indices Using Transition Rate Matrix Method

Non-sequential			Sequential	
Indices	**Without wake**	**Jensen**	**Without wake**	**Jensen**
LOLE[h/y]	130.79	178.30	132.17	180.48
EENS[kWh-yr]	255.18	296.82	261.01	301.60

Figure 10 shows the correlation between load data and wind speed. These data are based on average value during a day. The original load data are scaled down by dividing 3000 in order to present them with wind speed at a picture. It is observed that power generated from wind speed is very small during the peak load time. When wake effect is included in the wind farm, the reliability level intends to drop as shown by Table 10. It is observed that reliability index LOLE from Jensen and Eddy model has the similar value, while LOLE from Larsen model is relatively smaller than those from others. It appears that Larsen model does not explain the combined wake effects well. LOLE from transition rate approach and that from original data are different, since transition rate approach cannot take the correlation between wind speed and load over time, even though it considers all transitions between different wind speed states. Using transition rate approach, Figure 11 compares EENS by different wake models with pattern similar to LOLE. Table 11 shows the computation time to get EENS for different wake effect scenarios. As wake effect is incorporated on the wind power system, it takes a longer time to simulate. This is because that waked speed calculation by varying wind direction is added on the simulation. And calculation time for three different wake models is similar.

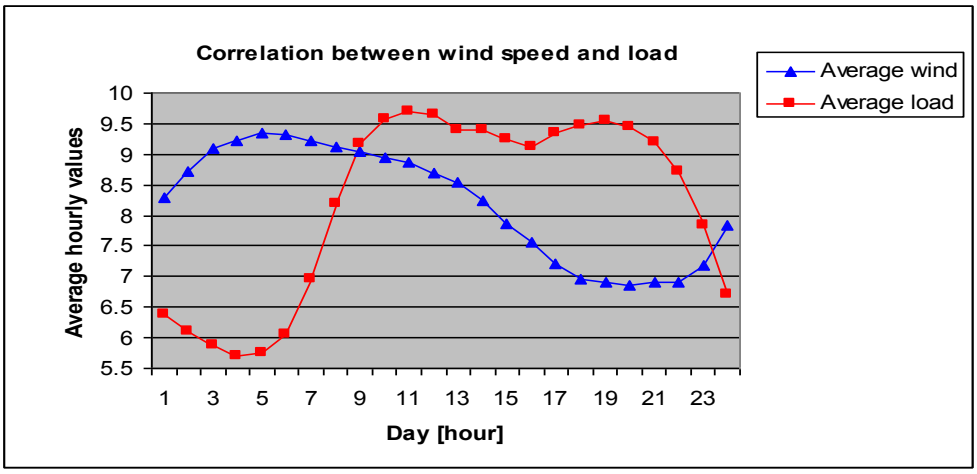

Figure 10. Correlation between Load and Wind Speed

Table 10. LOLE Comparison by Different Wake Models

Index	Approaches	Without wake	Jensen	Eddy	Larsen
LOLE [h/y]	Exact transition	130.79	180.48	172.22	151.67
	Original data	216.59	269.61	274.08	229.38

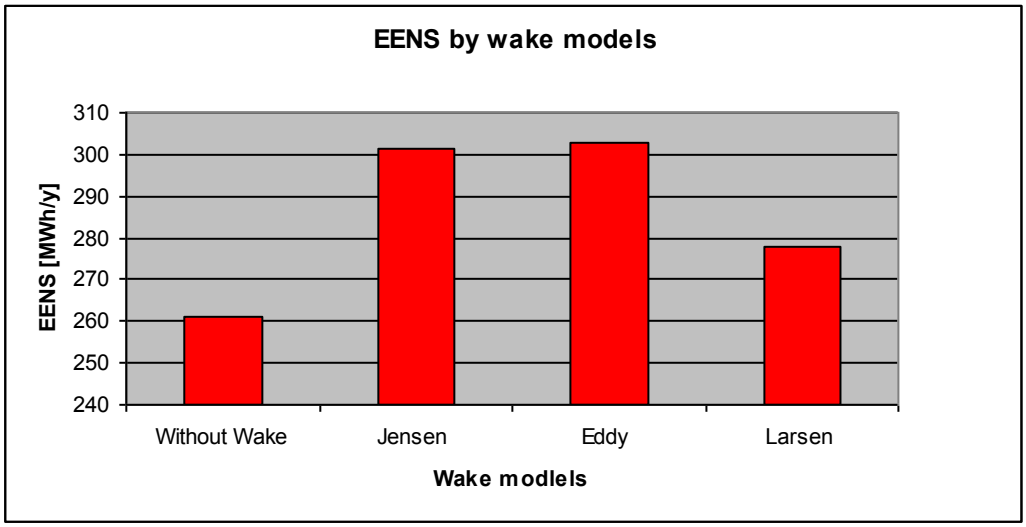

Figure 11. EENS Comparison by Wake Models

Table 11. Running Time Comparison by Different Wake Models

Models	**Without wake**	**Jensen**	**Eddy**	**Larsen**
Computation time [min]	4.03	11.46	11.37	11.57

3. CLUSTERING APPROACH FOR RELIABILITY SIMULATION IN A WIND FARM*

3.1 Clustering Techniques

Clustering is an effective approach for data mining as it groups original data into several clusters. It maintains data characteristics but reduces data size. From the clustering perspective, the approach can be classified into two categories, partitional clustering and hierarchical clustering [32]-[33]. The objective of partitional clustering is to partition the original data into the specific cluster sizes with a criterion function. On the other hand, hierarchical clustering generates clusters as a hierarchical tree. In this dissertation, partitional clustering approach is presented in details and discussed. As correlated variables, the observation data consists of wind, solar and load data. Figure 12 shows the clustering concept. Here N is the total number of data, and M is the dimension of the data - in this case M is two. K is the number of optimal clusters. The clustering approach is used to demonstrate the correlation between renewable energy and load. For partitional clustering, eight different clustering methods are presented and compared. These methods require the number of clusters as an input value. To find the optimal number of clusters, validity measurement [34] is used as shown by (12.1). Here C_i is the

* Reprinted with permission from "Comparison of Clustering Approaches for Reliability Simulation of a Wind Farm" by H. Kim and C. Singh, Oct. 2012, *Power System Technology (POWERCON) IEEE International Conference,* pp. 1-6, Copyright [2012] by IEEE.

center of cluster i , I_a is the distance between a point and a cluster center, and I_b is the distance between different clusters. So the desirable and compact clustering size is achieved by finding the minimum value of the validity.

Figure 13 shows the procedure to find the optimal clustering size. Before we start clustering process, we need to determine the optimal cluster size as input of clustering algorithm. The optimal cluster size is taken from validity measurement for best clustering results.

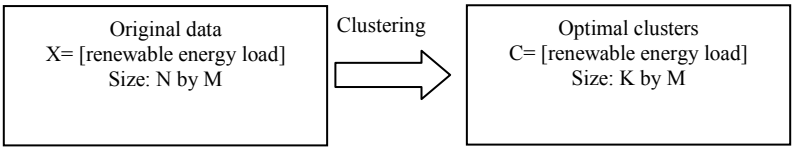

Figure 12. Description of Clustering

$$\text{Validity} = \frac{I_a}{I_b} \tag{12.1}$$

$$I_a = \frac{1}{N}\sum_{i=1}^{k}\sum_{x \in C_i} \| x - C_i \|^2 \tag{12.2}$$

$$I_b = \min(\| C_i - C_j \|)^2, \tag{12.3}$$

$$i \in \{1,2,...,k-1\}, j \in \{i+1,...,k\}, k \geq 2$$

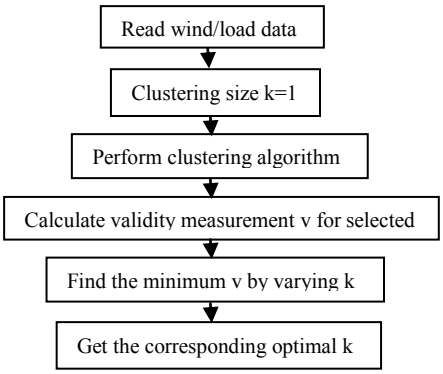

Figure 13. Procedure for Validity Measurement

3.1.1 K Means (KM)

K Means (KM) [4], [32] is simple, has fast simulation time and described in Figure 14. However, it depends on the initial choice of clusters, which can cause the local optimum. It means that different initial clusters can produce different optimal clusters. Here d_{is} is the distance between point i and cluster s, and convergence coefficient η is assumed to be 0.00001 in the work.

3.1.2 Fuzzy C Means (FCM)

Fuzzy C Means (FCM) [35] is fuzzy, which means that one point may be in several clusters. So it returns not only the optimal clusters but also the membership distribution, shown by Figure 15. The cluster with the highest membership is chosen to one of a point. Similar to KM, this method is sensitive to the initial choice of clusters.

Here m is a user-defined parameter, which is set to be two in this work. Sometimes, d_{ir} or d_{is} can be zero. For these singularity issues, membership is plugged into zero.

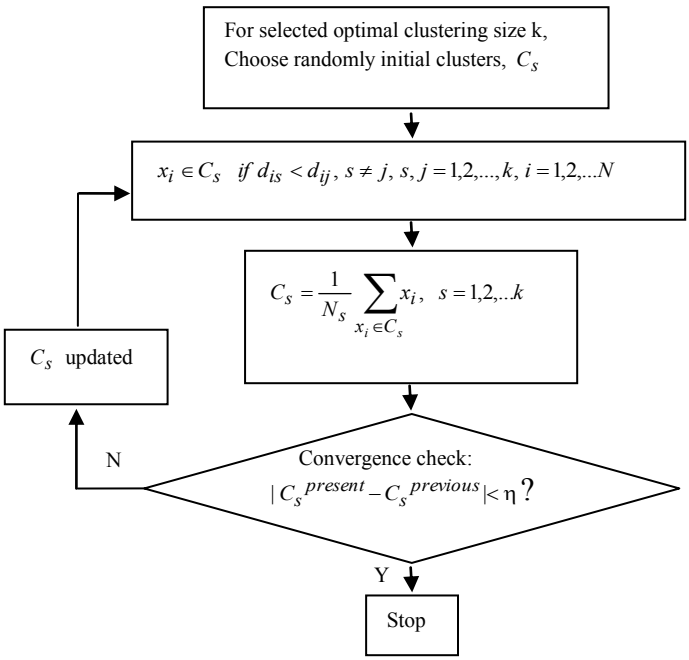

Figure 14. KM Clustering

3.1.3 Probabilistic C Means (PCM)

Probabilistic C Means (PCM) [36] is appropriate for the noisy data. In Figure 15, instead of U_{ir}, T_{ir} is evaluated, as described by (13.1). Contrary to FCM, there is no singularity problem in this approach, even if d_{ir} becomes zero.

$$T_{ir} = [1 + (\frac{d_{ir}^{\ 2}}{\eta_r})^{1/(m-1)}]^{-1} \tag{13.1}$$

$$\eta_r = \frac{\sum_{i=1}^{N} U_{ir}^{\ m} d_{ir}^{\ 2}}{\sum_{i=1}^{N} U_{ir}^{\ m}} \tag{13.2}$$

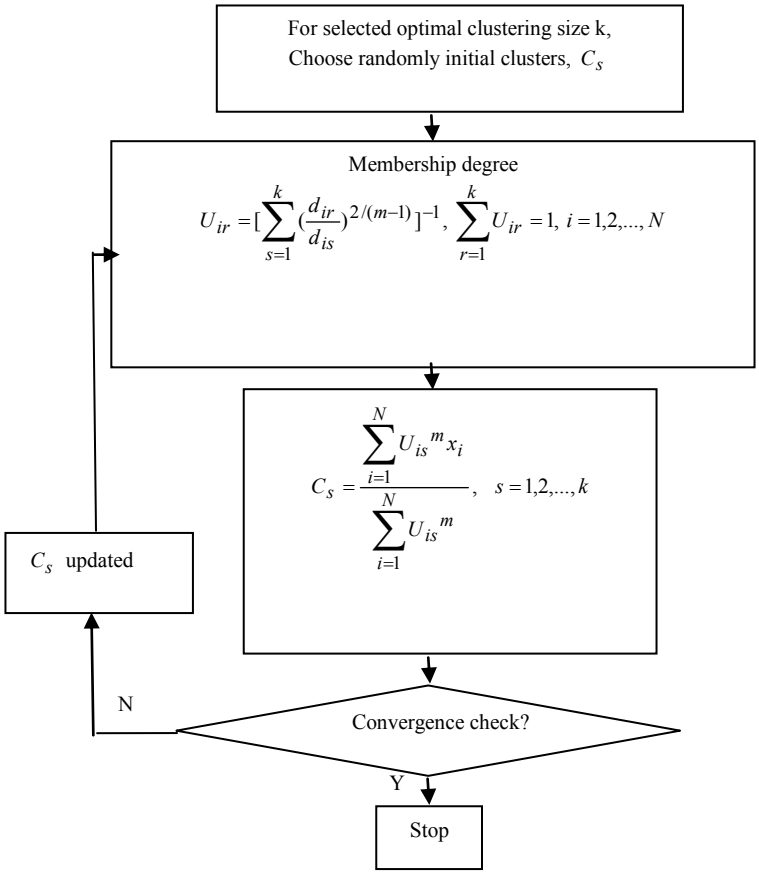

Figure 15. FCM Clustering

3.1.4 Fuzzy Probabilistic C Means (FPCM)

Fuzzy Probabilistic C Means (FPCM) [36] combines FCM and PCM. In Figure 15, clusters are calculated using (14.1). Here U_{is} is the same as in FCM. Parameters a and b are set to be zero in this work.

$$C_s = \frac{\sum_{i=1}^{N} F_{is}^{\ m} x_i}{\sum_{i=1}^{N} F_{is}^{\ m}}, \quad s = 1,2,...,k \tag{14.1}$$

$$F_{is} = aU_{is}^{\ m} + bB_{is}^{\ m} \tag{14.2}$$

$$B_{is} = [\sum_{g=1}^{N} (\frac{d_{is}}{d_{gr}})^{2/(m-1)}]^{-1}, \ \sum_{i=1}^{N} B_{is} = 1, \ s = 1,2,...,k \tag{14.3}$$

3.1.5 Global K Means (GKM)

Global K Means (GKM) [37] is independent of the initial choice of clusters as shown in Figure 16.

3.1.6 Fast Global K Means (FGKM)

The big issue in GKM is that it takes a long time to cluster data. The initial kth cluster is determined only after calculation of the cluster error for all points x_i. On the other hand, FGKM [37] calculates b_i to find the best point x_i^*, described in (15.1), instead of computing the cluster error from Figure 16. The calculation of b_i does not

require kth cluster value, and only uses ($C_1, C_2, ..., C_{k-1}$). A point x_i^* with the maximum b_i is selected as the initial kth cluster. If the term $d_{k-1}{}^j - p_{ij}{}^2$ is negative for j, it is plugged into zero. Here p_{ij} is the distance between point i and point j.

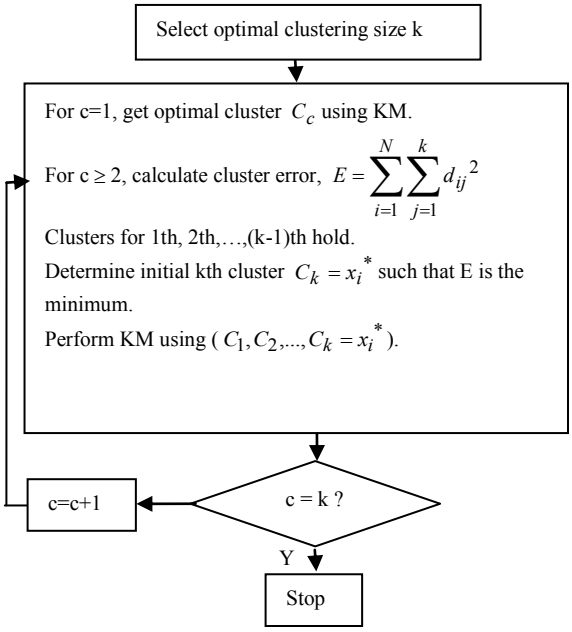

Figure 16. GKM Clustering

$$b_i = \sum_{j=1}^{N} (d_{k-1}{}^j - p_{ij}{}^2),\ i = 1,2,...,N \tag{15.1}$$

$$d_{k-1}{}^j = \min(d_{ir}{}^2),\ r = 1,2,...,k-1 \tag{15.2}$$

3.1.7 Global Fuzzy C Means (GFCM)

For FCM, the global approach can be applied as well [38]. In Figure 16, cluster error E_m is used to determine the initial kth cluster, shown by (16). To calculate E_m, all membership degrees for clusters, s=1,2,…k are necessary. A point x_i^* with the minimum E_m becomes the initial kth cluster. And then, FCM is carried out with clusters $(C_1, C_2, ..., C_k = x_i^*)$.

$$E_m = \sum_{i=1}^{N} \sum_{s=1}^{k} U_{is}{}^{m} d_{is}{}^{2} \tag{16}$$

3.1.8 Fast Global Fuzzy C Means (FGFCM)

To reduce the running time in GFCM, the fast approach is applied as well. FGFCM [38] computes J_m to find the best initial candidate of x_i, shown by (17). A point x_i^* with the minimum J_m is chosen as the initial kth cluster. And then, FCM is similarly operated.

$$J_m = \sum_{i=1}^{N} [\sum_{s=1}^{k} d_{is}{}^{2(1-m)}]^{1-m} \tag{17}$$

3.2 Comparison between Different Clustering Approaches

This section discusses about the best choice from the proposed clustering techniques. For the different clustering methods, the corresponding number of iterations is developed, and compared, shown by Table 12. Here N is the data size, k is the number of clusters, and η is a convergence coefficient. In KM, one term Nk is the calculation of the distance between points and clusters, and another one Nk is iterations for calculation of new clusters. In FCM, the term Nk^2 is the iterations for calculation of membership degree, $2Nk$ is iterations for the calculation of new clusters. In PCM, a new term Nk is added to FCM case which is related to T_{ir} in (13.1). In FPCM, new term N^2k is added to FCM case which calculates B_{is} in (14.3). In GKM, term $(Nk + Nk)\eta$ is changed to $(Nk + N)\eta$, compared with KM. This is because the global approach calculates only kth cluster. And the other one k is the calculation of cluster error E in Figure 16. In FGKM, term $N^2(k-1)$ represents iterations for the calculation of b_i in (15.1). In GFCM, $(Nk^2 + 2N)\eta$ is substituted instead of $(Nk^2 + 2Nk)\eta$, since the global approach computes only kth one cluster as well. The other Nk is for cluster error E_m in (16). Finally, FGFCM has term N^2k because of J_m in (17).

From the viewpoint of simulation time in Table 10, KM is very efficient. However it is sensitive to the initial selection of clusters, because of local optima. FCM still has local optimum, although it provides the additional membership degree information. PCM is not appropriate, since it is appropriate for noisy data. Global

approach gets over the sensitivity of initial choice of clusters, and fast global method reduces running time of global approach, holding the global optimum. Therefore, FGFCM is the efficient clustering approach, considering a compromise between simulation speed and clustering accuracy.

Table 12. Iteration for Clustering

Iterations	
KM	FCM
$(Nk + Nk)\eta$	$(Nk^2 + 2Nk)\eta$
PCM	FPCM
$(Nk^2 + 2Nk + Nk)\eta$	$(Nk^2 + N^2k + 2Nk)\eta$
GKM	FGKM
$Nk[(Nk + N)\eta + k]$	$k[N^2(k-1) + (Nk + N)\eta]$
GFCM	FGFCM
$Nk[(Nk^2 + 2N)\eta + Nk]$	$k[(N^2k + (Nk^2 + 2N)\eta]$

3.3 Generation from a Wind Farm

In practice, a wind farm has wake effect [7] between wind turbines. There are many models [18]-[19] to describe the wake effect in a wind farm. In this dissertation, N. O. Jensen model is used for a wake model, since it is simple and shows good performance for the modeling and simulation of waked speed. Figure 17 shows the basic schematic of the model. From the figure, wake speed at distance x is calculated using (18).

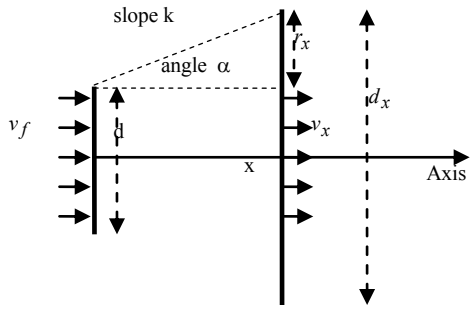

Figure 17. Proposed N. O. Jensen Model for Clustering

$$v_w = v_f\{1-\sum_i^n(1-\sqrt{1-C_t})(\frac{d}{d_{xi}})^2\} \tag{18}$$

where v_f [m/s] is free wind speed, C_t is thrust coefficient, v_w [m/s] is the waked speed, d is the diameter of an upstream turbine, $d_x = d + 2kx$, and k is a wake decreasing constant which is set to be 0.075 for onshore in this dissertation [35]. By considering the combined wake effect and shade effect in [22], [25],

$$v_w = v_f\{1-\sum_i^n(1-\sqrt{1-C_t})(\frac{d}{d_{xi}})^2(\frac{A_{si}}{A})\} \tag{19}$$

where n is the number of upstream turbines, and d_{xi} [m] is the wake region by upstream turbine i where A [m^2] is the rotor area and A_{si} [m^2] is the shade area by upstream turbine i. Wake effect depends on the wind direction. As the wind direction changes, the upstream and downstream turbines are updated and shade area between them is also changed. As a result, waked speed is calculated to include these effects.

A wind farm consists of a number of wind turbines which generate power depending on wind speed. A general wind power curve [27] is adopted in this work. To incorporate the wake effect, the power of each turbine is determined by waked speed and not by free wind speed with varying wind direction. The total power of a wind farm is calculated by the sum of the power of all turbines and by their failure/ repair characteristics.

3.4 Case Studies and Results

Monte Carlo [2], [10] is adopted for system simulation procedure. Wind data used is from National Renewable Energy Laboratory (NREL) [15] and load data is from IEEE Reliability Test System (RTS) [17]. Final reliability indices are calculated by operating inner product of indices from each cluster seed and their probabilities. The reliability index from each cluster is taken by convergence criterion to use coefficient of variation [13], [37]. Equation (18) shows computation of final reliability index, Loss of Load Expectation (LOLE). Here LOLE is the final index we desire to calculate, C_i is the probability for optimal cluster i, $LOLE_i$ is sub index calculated by cluster i, and n is the clustering size. The optimal cluster seeds represent characteristics of the original data, and their probabilities are distribution of the original data. Other indices, like Expected Energy Not Supplied (EENS) and Energy Not Supplied Interruption (ENSI) can be computed in this way. Based on the proposed different clustering algorithms, the corresponding programming is developed and simulated using Matlab. Reliability

indices, LOLE and EENS [12] are computed and compared with different clustering approaches. To see the validity of results, reliability indices using original data are also calculated and compared.

$$\mathrm{LOLE} = \sum_{i=1}^{n} C_i \mathrm{LOLE}_i \tag{20}$$

The layout of a wind farm is illustrated in Figure 18. It has 16 identical wind turbines whose distance is six times the diameter of turbines. To regulate the fluctuating characteristics of wind generation, one conventional generating unit is included in the wind farm whose capacity is the same as that of the wind farm, described in Table 13. Wind turbine information and wind data are gleaned from National Renewable Energy Laboratory (NREL) [15], and load data from IEEE Reliability Test System (RTS) [17], shown by Table 14 and Table 15. From the original wind data, wind states are identified, and the corresponding frequencies, durations, power and probabilities are calculated in Table 16. Power output of each state is computed using power curve [27]. Figure 19 shows an example of iterations required for KM, GKM, and FGKM using Table 12. KM is the fastest. In GKM the iterations exponentially increase for simulation as the number of clusters goes up. Fast global approach overcomes the drawback of global approach by shortenings the running time extremely.

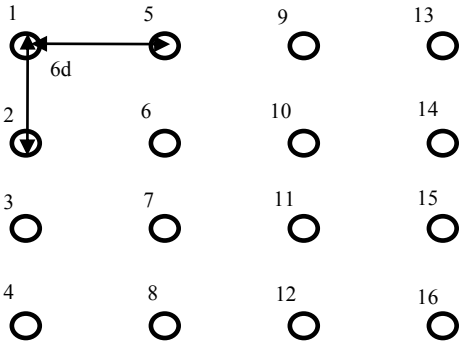

Figure 18. Proposed Wind Farm for Clustering

Table 13. System Capacity and Peak Load for Clustering

Capacity [kW]		**Annual peak load [kW]**
Wind farm	**Conventional unit**	
100 *16=1600	1600	1000

Table 14. Failure and Repair Rates of Turbines

Units	**Wind turbine**		**Conventional unit**
Weather	**Normal speed**	**Extreme speed(>19m/s)**	
Failure rate[#/yr]	6	36	6
Repair rate[#/yr]	130	36	130

Table 15. Wind Speed and Turbine Data for Clustering

Peak wind speed [m/s]	29.07
Minimum wind speed [m/s]	0.29
Mean wind speed [m/s]	8.10
Cut in speed [m/s]	6
Rated speed [m/s]	11
Cut out speed [m/s]	19
Rated power [kW]	100
Rotor diameter [m]	80
Hub height [m]	70

For the proposed wind farm, the optimal cluster size and the corresponding validity measurement for each clustering approach are shown in Table 17. Total data size of 8736 in one year is sharply decreased to 7 - 11.

Table 18 compares the running time of each clustering method for selected optimal cluster size. Figure 20 shows the distribution of data and clusters for FGKM. There are seven distinct clusters, and each black circle symbol indicates cluster center of the points.

Table 16. Wind Speed States for Clustering

State	Range [m/s]	Freq [#/yr]	Dur [h/#]	Power [kW]	Prob
1	0-4	603	2.596	0	0.179
2	4-5	1136	0.588	0	0.077
3	5-6	1269	0.580	0	0.084
4	6-7	1342	0.601	8.61	0.092
5	7-8	1393	0.629	26.76	0.100
6	8-9	1271	0.628	46.14	0.092
7	9-10	1135	0.627	66.76	0.082
8	10-11	1047	0.582	88.61	0.070
9	11-12	870	0.573	100	0.057
10	12-13	670	0.551	100	0.042
11	13-14	514	0.497	100	0.029
12	14-15	424	0.511	100	0.025
13	15-16	360	0.480	100	0.020
14	16-17	267	0.427	100	0.013
15	17-18	190	0.458	100	0.010
16	18-19	154	0.428	100	0.008
17	19-20	118	0.474	0	0.006
18	20-21	82	0.435	0	0.004
19	21-34	38	2.344	0	0.010

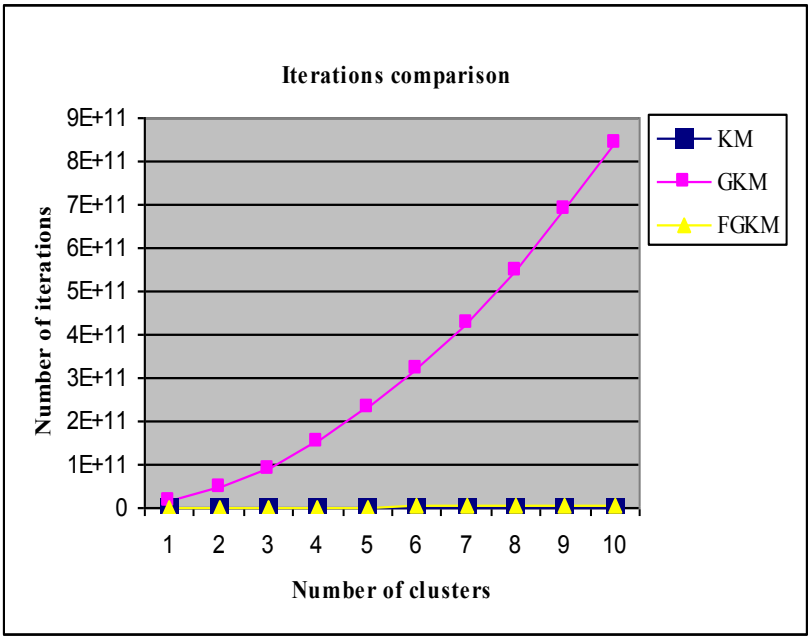

Figure 19. Example of Iterations for Different Cluster Size

Table 17. Optimal Clustering Size

Approach	**KM**	**FCM**	**PCM**	**FPCM**
Optimal k	10	11	8	8
Validity	0.125	0.126	0.110	0.250
Approach	**GKM**	**FGKM**	**GFCM**	**FGFCM**
Optimal k	7	7	7	7
Validity	0.122	0.113	0.183	0.179

Table 18. Running Time of Clustering

Approach	KM	FCM	PCM	FPCM
Running time [min]	0.016	2.1	3.5	151
Approach	**GKM**	**FGKM**	**GFCM**	**FGFCM**
Running time [min]	100.2	76.2	454	84

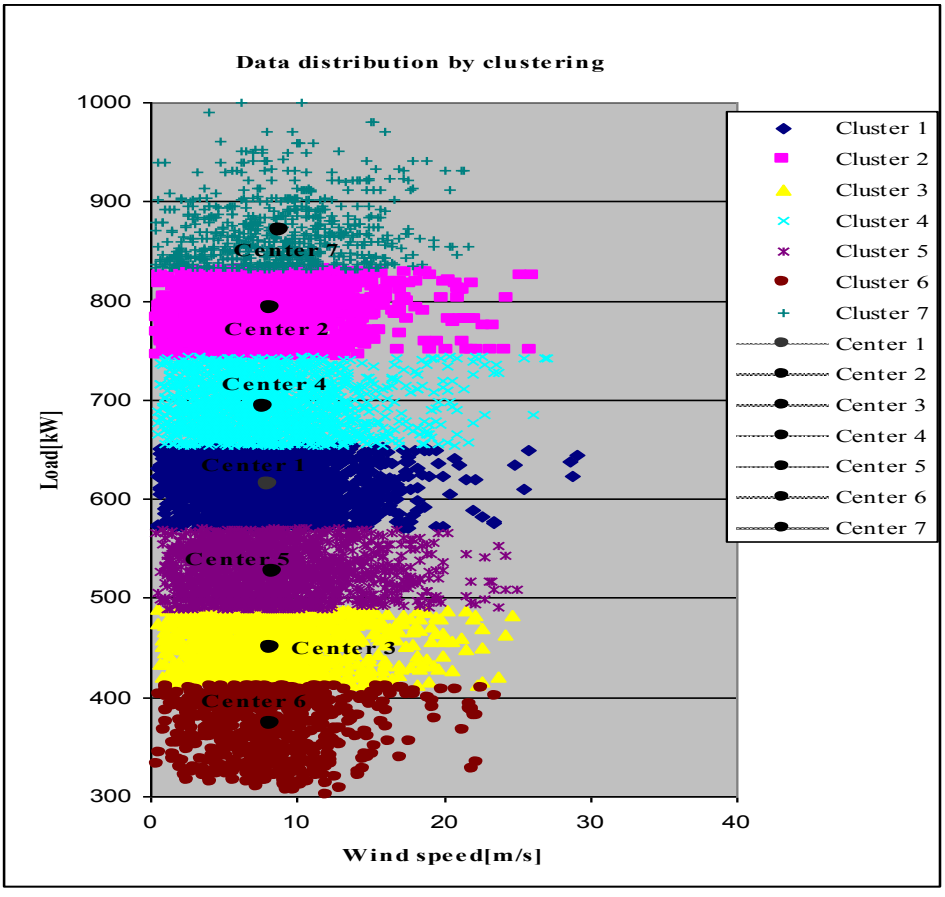

Figure 20. Example of Data Distribution Using FGKM

Using these results, LOLE is calculated and compared in Table 19. KM, FCM, and FPCM have similar value, which is different from the original case. This means these approaches converge to the local optimum by the sensitivity of initial choice of clusters. And PCM is also different since wind and load data are not noisy.

Global and fast global methods are closer to the original case, although they take longer time to simulate. These methods optimally add one cluster in an incremental way, instead of choosing the initial k clusters. So it's so time consuming. EENS is also calculated and compared in Figure 21. For non global approaches, resultant EENS shows fluctuations, depending on initial randomly selected clusters. However, for global approaches, the results are very close to those from original data approach.

Table 19. Reliability Index Comparison by Different Clustering Approaches

Index	**Original data**	**Clustering Approach**			
		KM	**FCM**	**PCM**	**FPCM**
LOLE[h/y]	259.22	328.78	332.38	370.05	335.72
		GKM	**FGKM**	**GFCM**	**FGFCM**
		263.81	269.18	268.12	270.68

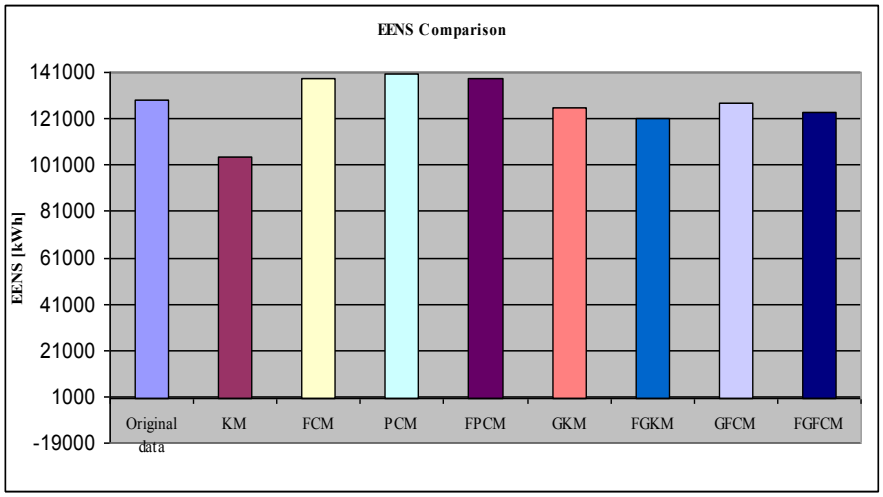

Figure 21. EENS by Different Clustering Approaches

4. RELIABILITY EVALUATION IN COMPOSITE POWER SYSTEMS WITH PHOTOVOLTAICS AND WIND USING MULTI-DIMENSIONAL CLUSTERING

4.1 Modeling for Photovoltaic System Generation

Photovoltaic (PV) array consists of a number of PV modules connected in series or parallel. Its power depends on a radiation from solar energy which is changed by the absorption, diffusion, or reflection of clouds and atmospheric dusts. And physical characteristics of a module determine actual PV power over time.

4.1.1 Hourly Clearness Index Modeling

The hourly ratio of the irradiance on a horizontal plane to the extraterrestrial solar radiation is called hourly clearness index. So this demonstrates cloud effect in the atmosphere. This dissertation uses reliable probability density function [39]-[40] of the clearness index, k_t for randomness of cloud distribution, shown by (21.1). Here k_{tu} is a maximum limit, and k_{tm} is a mean value of k_t. So fixing those two parameters, we can specify the cloud distribution anywhere over time. From (21.1), cumulative probability function is derived to (22.1) by the integration of time.

`

$$p(k_t) = \frac{Ce^{\lambda k_t}(k_{tu} - k_t)}{k_t} \tag{21.1}$$

$$C = \frac{\lambda^2 k_{tu}}{e^{\lambda k_{tu}} - 1 - \lambda k_{tu}} \tag{21.2}$$

$$\lambda = \frac{2r - 17.519e^{-1.3118r} - 1062e^{-5.0426r}}{k_{tu}} \tag{21.3}$$

$$r = \frac{k_{tu}}{k_{tu} - k_{tm}} \tag{21.4}$$

$$P(k_t) = C\frac{e^{\lambda k_t}(1 - r_1 k_t) - 1}{k_{tu}\lambda r_1} \tag{22.1}$$

$$r_1 = \frac{\lambda}{1 + \lambda k_{tu}} \tag{22.2}$$

To find k_t, we can use Monte Carlo simulation [9]-[10] in (21.1). Here Z is a uniformly distributed random variable ranging from zero to one. Some papers [41]-[42] use Lambert W function to solve the equation. In this dissertation, Newton Raphson method [43] is applied as an iterative solution, illustrated by (23)-(25.3). In (24), only right side term depends on k_t. By the approximation of Taylor series expression [43], it is iteratively calculated until a convergence criterion is satisfied. Jacobian J is one dimensional, and we choose an initial k_t as k_{tm} to reduce the number of iterations.

$$P(k_t) = Z \tag{23}$$

$$\frac{Zk_{tu}\lambda r_1}{C} + 1 = e^{\lambda k_t}(1 - r_1 k_t) \tag{24}$$

$$k_t\big|_{k_t = k_t^{i+1}} = k_t + J^{-1}\left[\frac{Zk_{tu}\lambda r_1}{C} + 1 - \{e^{\lambda k_t}(1 - r_1 k_t)\}\right]\Bigg|_{k_t = k_t^{i}} \tag{25.1}$$

$$J = \frac{\partial\{e^{\lambda k_t}(1 - r_1 k_t)\}}{\partial k_t}\Bigg|_{k_t = k_t^{i}} \tag{25.2}$$

$$J^{-1} = \frac{1}{e^{\lambda k_t}(\lambda - \lambda r_1 k_t - r_1)}\Bigg|_{k_t = k_t^{i}} \tag{25.3}$$

4.1.2 Calculation of Solar Radiation with Inclination

It would be incorrect if we directly use statistical radiation data which are based on the flat horizontal surface, since a PV module has a specific titled angle from the surface for the best insolation. An hourly radiation with inclined angle β [degree] [44] is shown in (26.1). Here R_b is a ratio of direct radiation on the surface to that on the flat surface, ϕ [degree] is latitude of a location, δ [degree] is solar declination, and n is day number. For instance, n=1 on January 1, and n=365 on December 31. ω [degree] is hourly angle, T [h] is local clock time, T_c [min] is time correction factor, ET [min] is equation of time, and LSTM [degree] is local standard time meridians to define time zone. This dissertation uses 90 degrees as central time zone. k_d represents diffuse fraction which is modeled by the piecewise-linear approximation, and ρ is reflectivity.

$$I_\beta = I_t\{R_b + (\frac{1+\cos\beta}{2} - R_b)k_d + \rho\frac{1-\cos\beta}{2}\} \quad (26.1)$$

$$R_b = \frac{\cos(\phi-\beta)\cos\delta\cos\omega + \sin(\phi-\beta)\sin\delta}{\cos\phi\cos\delta\cos\omega + \sin\phi\sin\delta} \quad (26.2)$$

$$k_d = 1 - 0.249k_t (k_t \le 0.35),\ 1.557 - 1.84k_t (0.35 < k_t < 0.75),$$
$$0.177 (k_t \ge 0.75) \quad (26.3)$$

$$\delta = 23.45\sin B \quad (26.4)$$

$$\omega = 15\{12 - (T + \frac{T_c}{60}\} \quad (26.5)$$

$$T_c = 4(LSTM - longitude) + ET \quad (26.6)$$

$$ET = 9.87\sin(2B) - 7.53\cos B - 1.5\sin B \quad (26.7)$$

$$B = \frac{360(n-81)}{365} \quad (26.8)$$

$$I_t = 1367\{1 + 0.033\cos(\frac{360n}{365})\}(\cos\phi\cos\delta\cos\omega$$
$$+ \sin\phi\sin\delta)k_t \quad (26.9)$$

For the high insolation, we need to find β that makes Sun's beam perpendicular to a PV module [44]. Figure 22 illustrates this fact. From the figures, we can determine β in (27)-(28). The tilted angle changes day by day, since solar declination depends on the number of the day.

$$\beta = 90^o - \beta_N \quad (27)$$

$$\beta_N = 90^o - \phi + \delta \quad (28)$$

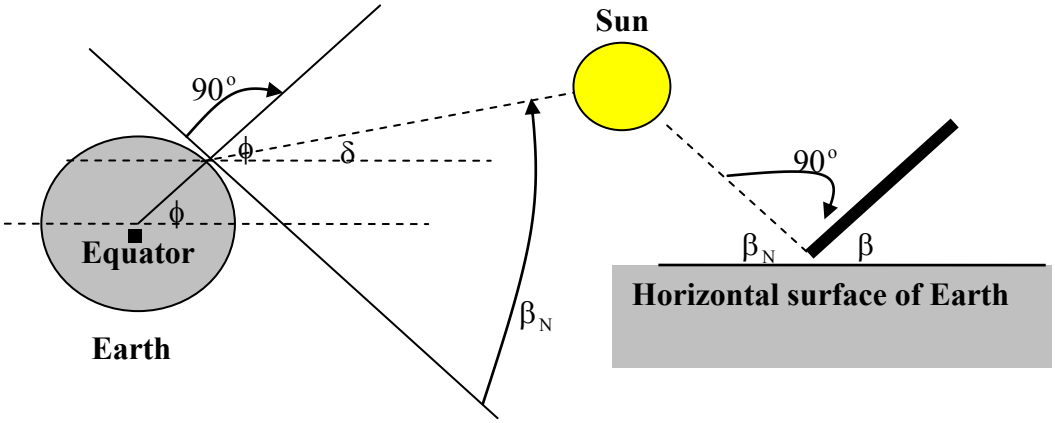

Figure 22. Tilted Angle in a PV Module

4.1.3 PV Array Output Power

A PV module produces power by a product of PV current and PV voltage across the load and these are determined by ambient temperature and solar radiation with inclination, shown by (40)-(44), respectively. Here $T_C\,[C^o]$ is a cell temperature, $N_{OT}\,[C^o]$ is nominal operating cell temperature, $I_{SC}\,[A]$ is short circuit current, $k_i\,[mA/C^o]$ is current temperature coefficient, $v_{OC}\,[V]$ is open circuit voltage, $k_v\,[mV/C^o]$ is voltage temperature coefficient, and $T_A\,[C^o]$ is ambient temperature. As $I_\beta\,[W]$ increases, T_C also does. As a result, PV current increases, while PV voltage decreases. Finally, PV array power is determined by (43). Here FF is fill factor, $V_{mpp}\,[V]$ is voltage at maximum power point, $I_{mpp}\,[A]$ is current at maximum power point, and N is the number of PV modules. Table 20 shows physical specification of PV Modules.

Table 20. Physical Specification of PV Module

Specification of PV modules
Nominal operating cell temperature
Current temperature coefficient
Voltage temperature coefficient
Short circuit current
Open circuit voltage
Voltage at maximum power point
Current at maximum power point

4.2 Generation from a Wind Farm

In practice, a wind farm has wake effect [7] between wind turbines. There are many models [18]-[19] to describe the wake effect in a wind farm. In this dissertation, N. O. Jensen model is used for a wake model, since it is simple and shows good performance for the modeling and simulation of waked speed. Figure 23 shows the basic schematic of the model. From the figure, wake speed at distance x is calculated using (29).

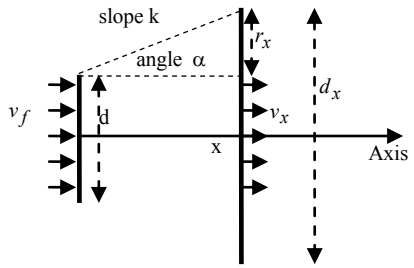

Figure 23. Proposed N. O. Jensen model for PV Systems

$$v_w = v_f\{1 - \sum_i^n (1 - \sqrt{1 - C_t})(\frac{d}{d_{xi}})^2\} \quad (29)$$

where v_f [m/s] is free wind speed, C_t is thrust coefficient, v_w [m/s] is the waked speed, d is the diameter of an upstream turbine, $d_x = d + 2kx$, and k is a wake decreasing constant which is set to be 0.075 for onshore in this dissertation [35]. By considering the combined wake effect and shade effect in [45],

$$v_w = v_f\{1 - \sum_i^n (1 - \sqrt{1 - C_t})(\frac{d}{d_{xi}})^2 (\frac{A_{si}}{A})\} \quad (30)$$

where n is the number of upstream turbines, and d_{xi} [m] is the wake region by upstream turbine i where A [m^2] is the rotor area and A_{si} [m^2] is the shade area by upstream turbine i. Wake effect depends on the wind direction. As the wind direction changes, the upstream and downstream turbines are updated and shade area between them is also changed. As a result, waked speed is calculated to include these effects.

A wind farm consists of a number of wind turbines which generate power depending on wind speed. A general wind power curve [27] is adopted in this research. To incorporate the wake effect, the power of each turbine is determined by waked speed and not by free wind speed with varying wind direction. The total power of a wind farm is calculated by the sum of the power of all turbines and by their failure/ repair characteristics.

4.3 Clustering of Data

It is generally known that there is a correlation between power produced by photovoltaic and wind and load, that is, they have a specific pattern with time, even though their generation has random behavior. To deal with the correlation, this work uses a clustering approach. Clustering is a grouping process to reduce original data size while keeping characteristics of the data. As one of partitional clustering methods, Fast Global Fuzzy C Means (FGFCM) [38] is applied to the power system with PV units and wind farms, shown by Figure 24. Here, J_m is described in (31). Using this algorithm, a programming code to simulate is developed with Matlab. It finds the optimal clusters and membership degree of each point by adding initial cluster with the minimum objective function step by step. To determine the optimal cluster size, validity measurement [34] is used in this dissertation.

$$J_m = \sum_{i=1}^{N} [\sum_{s=1}^{k} d_{is}^{\,2(1-m)}]^{1-m} \tag{31}$$

4.4 Dimensional Clustering with Renewable Energy

If we consider only PV array in composite power system, the dimensions of clustering are two, PV generation and load, since there is no correlation between conventional generating units and load. The solar power and load from IEEE RTS are

the inputs of clustering process. As a result, system reliability indices are calculated by product and sum of reliability index of each cluster and its probability. Figure 25 shows the two dimensional clustering for PV system and load. Here N is the original data size, and K is the optimal clustering size.

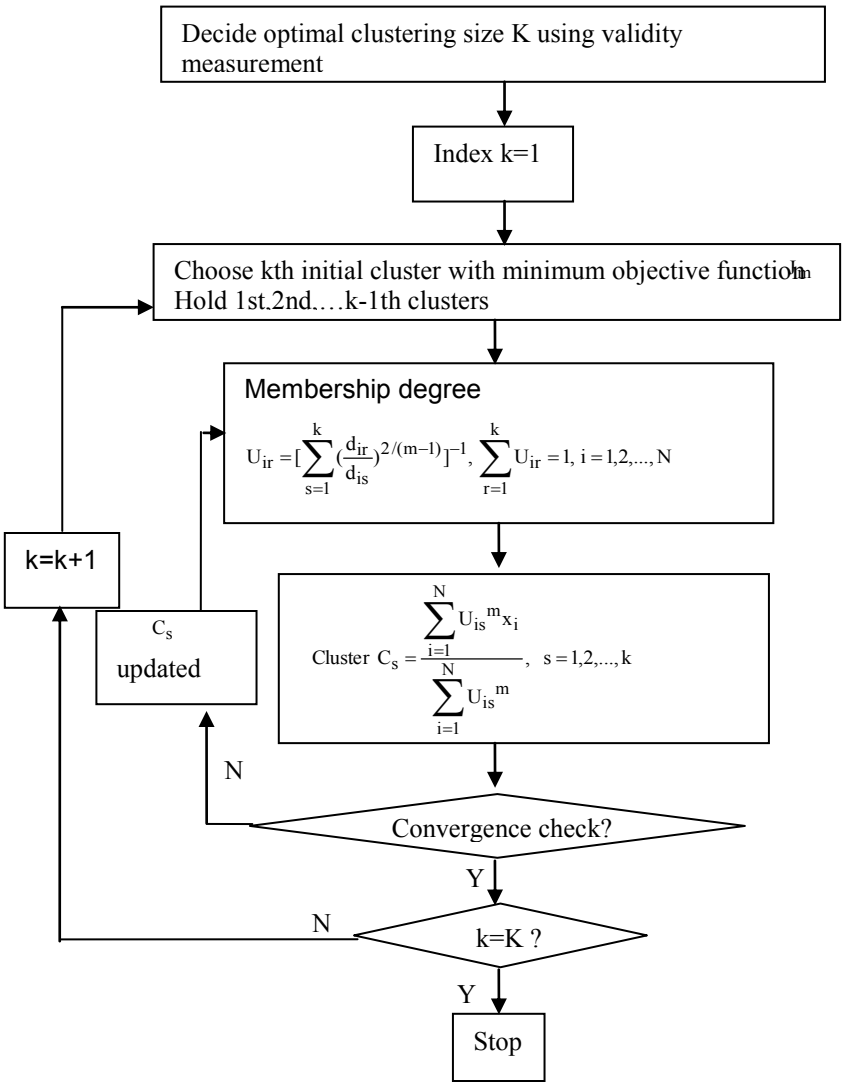

Figure 24. FGFCM Algorithm

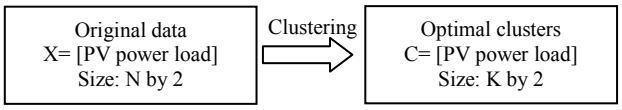

Figure 25. Description of Two Dimensional Clustering

Let us consider the combination of wind farms and composite power system except solar units. This case is three dimensions, wind speed, direction, and load, since wind power depends on wind direction by wake effect as well as wind speed. Wind speed and load are scalars, however, wind direction data are represented by degree or cardinal direction. If we directly use the data for calculation of distance between points and cluster seeds, the clustering results turn out to be incorrect. For example, suppose there are two directions, 0 degree and 360 degree. If we directly use those values to calculate the distance between them, the distance is 360 and this is wrong. The distance should be just zero. This the problem to calculation of distance between points and clusters in wind direction data.

To solve this problem, this research uses a unit circle with a radius of one to calculate the distance. We can put wind direction data on the unit circle. The distance is the straight line connecting between two points. From the original wind data, a degree begins increasing clockwise from vertical axis. However, in mathematical rectangular coordinates, it starts increasing counter clockwise from horizontal axis. From given wind direction degree, we can transform into rectangular coordinates, illustrated in Table 21. The distance is the biggest when the difference between degrees is 180 degree. Using this approach, it is possible to do correct clustering for wind direction data. Here in the

points C and D, the first column is hourly wind speed, the second column is hourly wind direction, and the third column is hourly load. Clustered wind speed and direction are used to calculate more accurate power of wind farm lost by wake effect. Figure 26 shows three dimensional clustering for wind farm and load.

Table 21. Transformation to Rectangular Coordinates

Degree direction	**Rectangular coordinates**
A	B=[cos(-A+90), sin(-A+90)]
Distance between point C(x1,y1,z1) and D(x2,y2,z2)	
$\sqrt{(x1-x2)^2+\sqrt{\{\cos(-y1+90)-\cos(-y2+90)\}^2+\{\sin(-y1+90)-\sin(-y2+90)\}^2}+(z1-z2)^2}$	

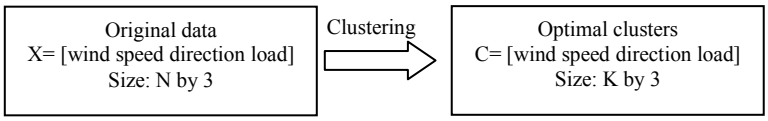

Figure 26. Description of Three Dimensional Clustering

Finally, let us consider a case of all combination, PV arrays and wind farms in IEEE RTS which consists of solar power, wind speed, direction and load for clustering, illustrated by Figure 27. PV units and wind farms are installed in some buses of the composite power system. Instead of taking original solar and wind data, clusters and

their probabilities are used to evaluate system reliability so that the results are compared with those by actual data.

Table 22 compares memory size required to store data for different cases. Here N is the number of original data, and K is the optimal cluster size. For clustering approach, one more dimension in column is added to a memory, because of probability information of each cluster.

For example, the coefficient of two dimensional clustering is three, since it includes solar power, load and probability. Clustering approach is so efficient for reliability evaluation of power systems, noting that K is much less than N.

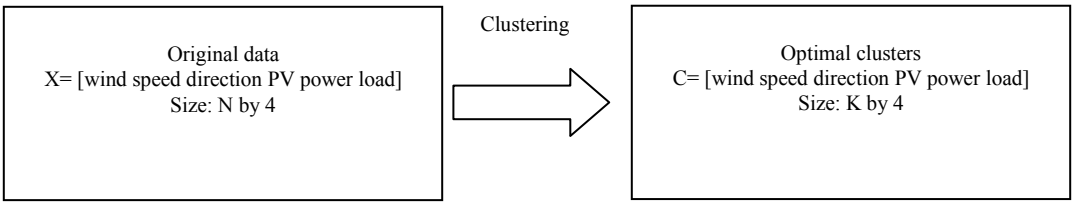

Figure 27. Description of Four Dimensional Clustering

Table 22. Comparison of Memory Size

Cases	**Approaches**	
	Actual data	**Clustering**
Two dimensions	2N	3K
Three dimensions	3N	4K
Four dimensions	4N	5K

4.5 System Optimization

As non-sequential Monte Carlo, random sampling method [9]-[10] is used for the transition of turbines' operation with clustering. In composite system level calculations, all transmission constraints are considered for reliability analysis and DC power flow is embedded in the formulation of minimum curtailment of load formulation, shown in (32.1)-(32.5). In this formulation, N is the number of buses, C_k is load curtailments at bus k, g_k is generation at bus k, f_{kj} is real power flow between bus k and j, d_k is load at bus k, g_{lower} is lower limit of generation, g_{upper} is upper limit of generation, f_{lower} is lower limit of power flow, and f_{upper} is upper limit of power flow. Simplex method uses the reduced costs of the system problem to get the final optimal solution with iterations. At the beginning, it is required to choose the initial basic feasible solution once the system problem is converted into the standard form. However, we cannot choose the initial basic variable for basic feasible solution in some constraints because of reverse (negative) power flow limits in transmission lines. As an alternative method, artificial variables are added to the problem. There are generally two approaches for using artificial variables [46]; two phase method, and big M method - two phase method is used in this work. Two phase method has two phases to optimize a problem. At the phase one level, its objective function is the sum of all artificial variables. If the optimal value is not zero, it does not have any feasible solutions, since artificial variables are added to the original problem. Otherwise, it goes to next level, phase two. If some

artificial variables are in basic variable set, they are replaced by other non-basic variables and the simplex process iterates using the reduced costs to find the final optimal solution to the original problem. Programming code is developed for the algorithm of two phase method using a computer tool Matlab.

$$\text{Objective function} = \min \sum_{k=1}^{N} C_k \tag{32.1}$$

Constraints

$$C_k + g_k - \sum_j f_{kj} = d_k \tag{32.2}$$

$$0 \le C_k \le d_k \tag{32.3}$$

$$g_{lower} \le g_k \le g_{upper} \tag{32.4}$$

$$f_{lower} \le f_{kj} \le f_{upper} \tag{32.5}$$

To reduce the running time, we can use sensitivity analysis [46]-[47] for linear programming. It determines whether current optimization process should be taken directly from previous optimization information, or carried out again. Figure 28 shows the sensitivity analysis procedure. If the inverse of coefficient matrix by basic variables in the power flow equations, B^{-1} times current Right Hand Side (RHS) vector, b is bigger than zero, we can directly get current load curtailments by product of coefficient vector of objective function by basic variables, C_B and B^{-1}, and b. Otherwise, the optimization should be newly started again, since previous basic matrix no longer holds feasibility. Loss of Load Expectation (LOLE) is used as one of reliability indices. It indicates mean system failure hour during simulation period, one year. The results for different dimensional cases are compared with those from actual data. The system

simulation to calculate LOLE from actual data is based on next event method, one of sequential Monte Carlo [9]-[10].

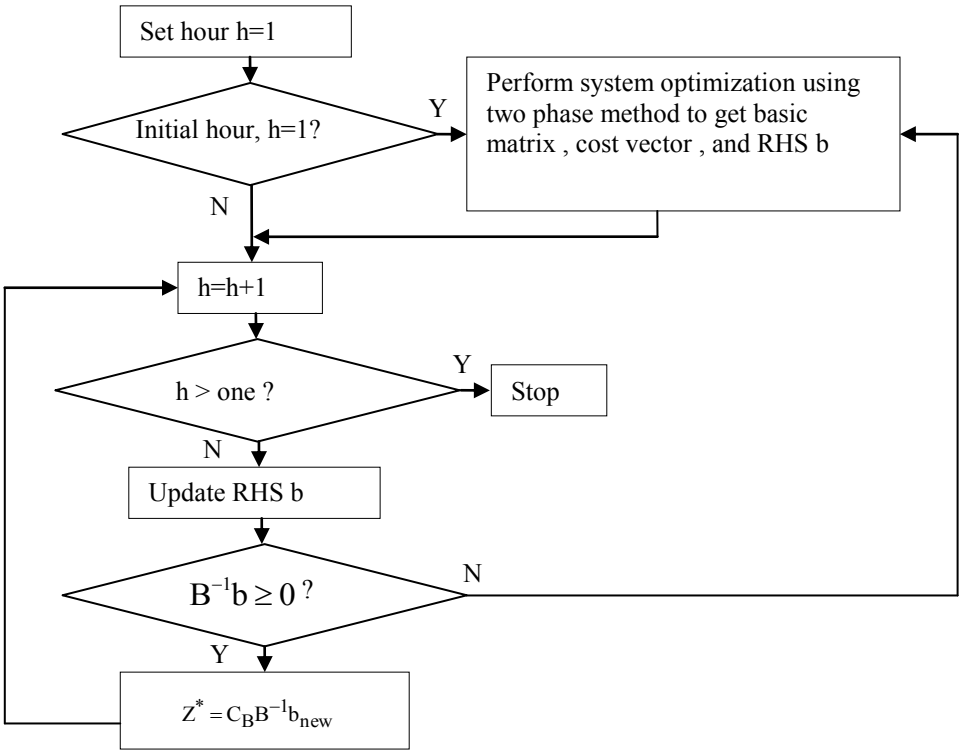

Figure 28. Sensitivity Analysis Procedure

4.6 Case Studies and Results

Figure 29 illustrates a layout of the proposed power system. Energy from conventional generating units and wind farm and PV array is supplied to the load through transmission lines of IEEE RTS.

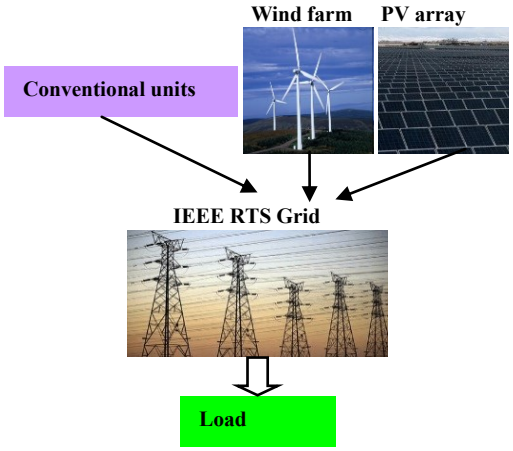

Figure 29. Layout of Proposed Power System

Table 23 shows system capacity and annual peak load. By the product of PV current and PV voltage across the load in a cell, solar power is generated. So no failure/repair process of in PV modules is considered although there may be failure/repair process for some components like diode, or resistance. Five identical PV arrays are deployed on some buses in the IEEE RTS.

For hourly clearness index modeling, maximum limit, k_{tu} and mean value, k_{tm} are set to be 0.864 and 0.4, respectively. We assume that convergence coefficient is 0.000001 for iteration of Newton Raphson method to calculate clearness index. And reflectivity is 0.26 for calculation of solar radiation in (26.1), which is used for green grass surface. Table 24 describes PV module's operating specification. From the proposed model and the specification for PV array generation, Figure 30 shows average hourly PV power. It has a peak value at noon. The latitude and longitude of the PV array located are (30.595^0, 96.366^0).

Table 23. System Capacity and Peak Load for PV Systems

PV array	**Peak Watt of a module [W]**	**# of modules**	**Locations of bus**	**Capacity [MW]**
	264	90000	15-19	118.8
Wind farm	**Capacity of a turbine [MW]**	**# of turbines**	**Locations of bus**	**Capacity [MW]**
	5	16	3,20,24	240
IEEE RTS	**Peak load [MW]**			
	3405			
	Total conventional generation [MW]			
	3405			

Table 24. Proposed Specification of PV Module

Open circuit voltage [V]	24
Short circuit current [A]	13
Voltage at maximum power [V]	23
Current at maximum power [A]	12
Voltage temperature coefficient [mV/ C^o]	60
Current temperature coefficient [mA/ C^o]	1.5
Nominal cell operating temperature [C^o]	42

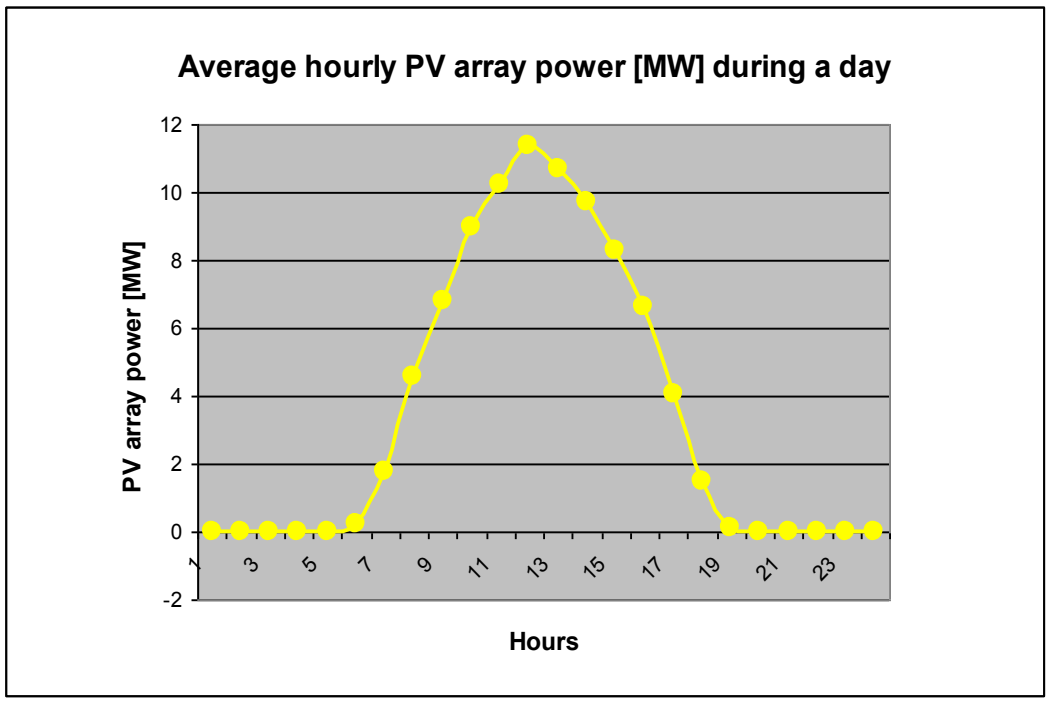

Figure 30. Average Hourly PV Array Power

Figure 31 shows layout of a wind farm proposed in this work. Three identical wind farms are installed at buses of IEEE RTS. The transition rate information of turbines is described in Table 25. Table 26 and Table 27 show wind speed state and direction state, respectively.

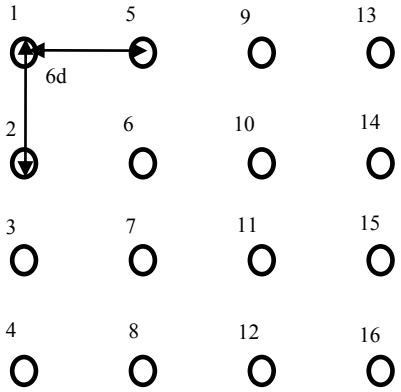

Figure 31. Proposed Wind Farm for PV Systems

Table 25. Transition Rates of Turbines

Units	**Conventional unit**	**Wind turbine**	
Weather		**Normal speed**	**Extreme speed (>19m/s)**
Failure rate[#/yr]	6	6	36
Repair rate[#/yr]	130	130	36

For the proposed power system, 124 slack variables and 3 artificial variables are initialized for optimization in the two phase method. Equation (33.1)-(33.6) shows original problem for bus number one in IEEE RTS. Here A_1 is the upper limit of generation at bus one. It changes every hour depending failure/repair behavior of turbines. The upper limit of transmission line connected with bus one is 175 [MW].

Table 26. Wind Speed States for PV Systems

State	Range [m/s]	Prob.	Freq [#/yr]	Dur [h/#]	Power [MW]
1	0-4	0.1856	546	2.9689	0
2	4-5	0	0	0	0
3	5-6	0.0833	562	1.2954	0
4	6-7	0.0814	581	1.2238	0.43
5	7-8	0.0777	560	1.2125	1.33
6	8-9	0.0767	556	1.2050	2.30
7	9-10	0.0725	523	1.2103	3.33
8	10-11	0.0672	491	1.1955	4.43
9	11-12	0.0619	467	1.1585	5
10	12-13	0	0	0	5
11	13-14	0.0590	426	1.2089	5
12	14-15	0.0474	357	1.1597	5
13	15-16	0.0458	341	1.1730	5
14	16-17	0.0314	224	1.2232	5
15	17-18	0.0311	235	1.1574	5
16	18-19	0.0224	171	1.1462	5
17	19-20	0	0	0	0
18	20-21	0.0160	120	1.1667	0
19	21-34	0.0406	171	2.0760	0

Table 27. Wind Direction States for PV Systems

States	Direction	Prob.	Freq [#/yr]	Dur [h/#]
1	N	0.1956	637	2.6829
2	NNE	0.0291	170	1.4941
3	NE	0.0211	120	1.5333
4	ENE	0.0222	125	1.5520
5	E	0.0309	159	1.6981
6	ESE	0.0394	204	1.6863
7	SE	0.0689	359	1.6769
8	SSE	0.1419	649	1.9106
9	S	0.2603	763	2.9803
10	SSW	0.0496	286	1.5140
11	SW	0.0239	139	1.5036
12	WSW	0.0117	73	1.3973
13	W	0.0117	78	1.3077
14	WNW	0.0143	83	1.5060
15	NW	0.0294	169	1.5207
16	NNW	0.0500	231	1.8918

To convert into standard form, slack variables are added on the problem in (34.1)-(34.6). Here S_i indicates slack variable for i=1, 2, 3, 4, and 5. And variables are

non-negative. To have the initial basic matrix with non-negativity of RHS, an artificial variable should be added in (34.1) because of negative power flow. So (34.1) is replaced by (35) where Z_1 is artificial variable. S_i and Z_1 are used for the initial basic variables to make up basic matrix in phase one. After the simplex process, if Z_1 is zero, it goes to next phase two so that the final solution to the original problem is derived. One can do the numbering manually as below.

$$C_1 + g_1 - f_{12} - f_{13} - f_{15} = d_1 \quad (33.1)$$
$$0 \le C_1 \le d_1 \quad (33.2)$$
$$0 \le g_1 \le A_1 \quad (33.3)$$
$$0 \le f_{12} \le 175 \quad (33.4)$$
$$.0 \le f_{13} \le 175 \quad (33.5)$$
$$0 \le f_{15} \le 175 \quad (33.6)$$

$$C_1 + g_1 - f_{12} - f_{13} - f_{15} = d_1 \quad (34.1)$$
$$C_1 + S_1 = d_1 \quad (34.2)$$
$$g_1 + S_2 = A_1 \quad (34.3)$$
$$f_{12} + S_3 = 175 \quad (34.4)$$
$$f_{13} + S_4 = 175 \quad (34.5)$$
$$f_{15} + S_5 = 175 \quad (34.6)$$

$$C_1 + g_1 - f_{12} - f_{13} - f_{15} + Z_1 = d_1 \quad (35)$$

Clustering results are presented in Table 28, Table 29, and Table 30, respectively. LOLE of each cluster tends to decreases as wind farms and PV units are incorporated on the power system.

Table 28. Clustering in Two Dimensions

Cluster #	**Clusters=[PV power,load]=[MW,MW]**	**Prob.**	**LOLE[h/y] of each cluster**
1	[3.5,2091.5]	0.1844	7
2	[1.5,1629.1]	0.1571	3.9
3	[5.6,2718.5]	0.1394	106.7
4	[0,1328]	0.1220	3.6
5	[4.6,2381.3]	0.1883	45.9
6	[5.7,2973.1]	0.0687	105
7	[2.5,1841]	0.1401	3.2

Table 29. Clustering in Three Dimensions

Cluster #	**Clusters=[wind speed,direction,load]=[m/s,state,MW]**	**Prob.**	**LOLE[h/y] of each cluster**
1	[9.1,8,2091.5]	0.1840	5.7
2	[7.4,8,1628.3]	0.1566	3.6
3	[10.8,9,2718.6]	0.1394	103
4	[6,7,1327.9]	0.1221	3.3
5	[10,8,2380.7]	0.1884	40.6
6	[10.2,9,2973.1]	0.0687	103.9
7	[8.9,8,1840.7]	0.1408	3

Table 30. Clustering in Four Dimensions

Cluster #	**Clusters=[wind speed,direction,PV power,load] =[m/s,state,MW,MW]**	**Prob.**	**LOLE[h/y] of each cluster**
1	[9.1,8,3.5,2091.5]	0.1838	5.4
2	[7.4,8,1.2,1628.3]	0.1567	2.57
3	[10.8,9,5.6,2718.6]	0.1399	98.7
4	[6,7,0.2,1327.9]	0.1221	2.17
5	[10,9,8,5.1,2380.1]	0.1881	36.2
6	[10.2,9,6,2973.1]	0.0687	99.98
7	[8.9,8,2.4,1840.7]	0.1407	2.07

Table 31 compares LOLE for different cases using the original data as well as clustering method. The results from original data are based on the sequential simulation using next event method.

LOLE using proposed clustering approach with random sampling is calculated by using cluster information shown in Table 28, Table 29, and Table 30. The actual data size, 8736 is grouped into the optimal cluster size, 7. The correlation between renewable energy and load is well preserved by clustering, observing that LOLE is very similar each other.

Table 31. LOLE [h/y] for Different Cases

Cases	LOLE [h/y]	
	Original data	Clustering
Without renewable energy	36.50	36.11
IEEE RTS+PV arrays	34.11	33.52
IEEE RTS+Wind farms	32.01	31.58
IEEE RTS+PV arrays+Wind farms	30.14	29.43

5. DEPLOYMENT OF OPTIMAL STORAGE BUSES IN COMPOSITE POWER SYSTEMS WITH WIND FARMS

5.1 Modeling for Wind Farm Generation

Conventional units' ability to generate power depends on their failure and repair status. Unlike these units, the ability of wind turbines in a wind farm to generate power depends on the wind speed as well as their operational status. This is why a wind farm is represented by a product of wind speed model and wind turbine model. The output of each wind turbine is determined by combining these two models. This research uses the general wind power curve [27] whose input information is a cut in speed, rated speed, cut out speed, and rated power to determine wind power. Then overall power of wind farm, which is equal to the sum of the total contribution of all wind turbines, is supplied to the load at a given time.

For wind speed models, two approaches are presented in the dissertation and applied to the system; transition rate method [2]-[3] and clustering method [32]-[33]. From original wind speed data, wind states are identified so that the frequency, duration and probability of each state can be calculated.

5.1.1 Transition Rate Matrix Approach

This approach is based on the transition rates among wind states. Transition rate

is the ratio of number of state changes and duration of the stay in the state before transition. To consider wind speed changes, all transition rates are extracted from the original wind speed. Using these rates, the frequency, duration, and probability of all wind states can be derived.

5.1.2 Clustering Method

Load and wind speed may have a pattern of variation relative to each other as both values have some relation to the time of the day, season, or weather. In other terms, load and wind speed may be correlated variables. Exact transition method does not capture this correlation as the transition rate matrix contains average rates over the period of study and these are assumed constant. Clustering approach is proposed for such cases. Actual load and wind speed data as functions of time are collected. Each pair of wind speed and the corresponding load constitute one data point. Then using clustering algorithm, all given data points are grouped into several clusters using the nearest cluster seeds sorting based on the Euclidean distance. Clustering is generally categorized into partitional methods or hierarchical methods [32]-[33]. In this work, partitional clustering is used and its various approaches compared. It partitions original data into the specific data size holding data characteristics. K-Means (KM) [4], [33] is a simple and fast method for this purpose. Once the clustering size is determined, it iterates to find the optimal clusters with the closest distance between clusters and points, starting from the initially selected clusters. Fuzzy C-Means (FCM) [35] provides

additional membership probability of clusters for each point to show fuzziness. These two methods basically depend on the initially randomly selected clusters. So the final clusters can be different depending on the initial selection. To solve this issue, global approach can be applied for clustering like Global K-Means (GKM) [37] and Global Fuzzy C-Means (GFCM) [37]. For the faster simulation procedure, Fast Global K-Means (FGKM) [38] and Fast Global Fuzzy C-Means (FGFCM) [38] are also examined in his research. These global or global fast approaches find the optimal clusters by adding a cluster step by step, instead of starting initial guess with preselected clustering size. They are independent of the initial guess so that it is possible to make more accurate and reliable clustering from original data.

As an input to clustering, cluster size should be determined using validity measurement [34]. Data consist of two dimensional observations; wind speed and load, and the size of the data is simulation period, one year in this research. The number of iterations for different clustering approaches is developed and compared in this work, shown by Table 32. Here n is the data size, k is the clustering size, and c is the number required to satisfy convergence. As the size of data or cluster becomes bigger, the simulation time of the global approach exponentially increases. And n is much bigger than k in general composite power systems so that fast global approaches like FGKM and FGFCM are efficient for simulation.

Table 32. Iterations of Clustering

Number of iterations for different clustering approaches	
KM	FCM
$2nkc$	$nk^2c + 2nkc$
GKM	GFCM
$n^2k^2c + n^2kc + nk^2$	$n^2k^3c + 2n^2kc + n^2k^2$
FGKM	FGFCM
$n^2k(k-1) + nk^2c + nkc$	$n^2k^2 + nk^3c + 2nkc$

For wind turbine model, Monte Carlo simulation [9]-[10] is used in this work. Each turbine is assumed to have two generating states; fully available and out of service. From failure/repair rates of turbines, probabilities of two operating states can be calculated. As one of sequential methods, next event approach [48] is applied for the system simulation. Probability distribution function for transition duration time of each turbine is assumed to be exponential. And then the operating state of a turbine and its transition time demonstrates failure and repair behaviors of the turbine.

5.2 Wake Effect Models

In wind farm, there is wake effect which is turbulent air flow in the area of leeside of turbines. This causes energy loss in power productivity. As one of wake

models, N. O. Jensen model [21]-[24] is used in the dissertation. From undisturbed wind speed, waked speed is calculated, as shown by Figure 32 and (36). To consider cumulative wake [45] and wind shade effect [45], modified Jensen model is applied to the system in (37). Here v_f [m/s] is free wind speed, C_t is thrust coefficient [22], d [m] is the diameter of turbines, d_x [m] $= d + 2kx$, k is wake decreasing coefficient, set to be 0.075 for onshore, and v_w [m/s] is waked speed from distance x [m]. And N is number of upstream turbines by wake effect, d_{xi} [m] $= d + 2kx_i$, x_i [m] is a distance between upstream turbine i and its downstream turbine. A_{si} [m^2] is shade area of the downstream turbine by upstream turbine i, and A [m^2] is the rotor disc are of the downstream turbine. Wake effect also depends on wind direction, since upstream and downstream turbines vary with different direction. New wind state by waked speed becomes input of general wind power curve [27] so that wind power by wake effect is finally generated considering the operating states of turbines.

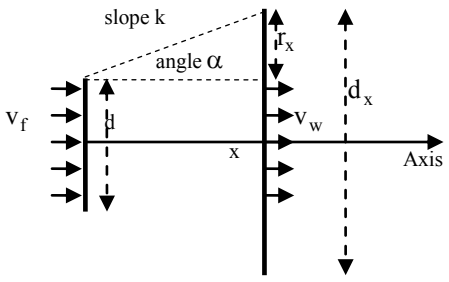

Figure 32. Proposed N. O. Jensen Model for Storage

$$v_w = v_f\{1-(1-\sqrt{1-C_t})(\frac{d}{d_x})^2\} \quad (36)$$

$$v_w = v_f\{1-\sum_i^n(1-\sqrt{1-C_t})(\frac{d}{d_{xi}})^2(\frac{A_{si}}{A})\} \quad (37)$$

5.3 System Optimization

In composite system level calculations, all transmission constraints are considered for reliability analysis and DC power flow is embedded in the formulation of minimum curtailment of load formulation, shown in (3). In this formulation, N is the number of buses, C_k is load curtailments at bus k, g_k is generation at bus k, f_{kj} is real power flow between bus k and j, d_k is load at bus k, g_{lower} is lower limit of generation, g_{upper} is upper limit of generation, f_{lower} is lower limit of power flow, and f_{upper} is upper limit of power flow. Simplex method uses the reduced costs of the system problem to get the final optimal solution with iterations. At the beginning, it is required to choose the initial basic feasible solution once the system problem is converted into the standard form.

In general, the initial basic feasible solution may be unavailable from the original problem, since there are some constraints with reverse real power flows. As an alternative method, artificial variables are added to the problem. There are generally two approaches for using artificial variables [46]; two phase method, and big M method - two phase method is used in this research. Two phase method has two phases to

optimize a problem. At the phase one level, its objective function is the sum of all artificial variables. If the optimal value is not zero, it does not have any feasible solutions, since artificial variables are added to the original problem. Otherwise, it goes to next level, phase two. If some artificial variables are in basic variable set, they are replaced by other non-basic variables and the simplex process iterates using the reduced costs to find the final optimal solution to the original problem. Programming code is developed for the algorithm of two phase method using a computer tool Matlab.

$$\text{Objective function} = \min \sum_{k=1}^{N} C_k \tag{38.1}$$

Constraints

$$C_k + g_k - \sum_j f_{kj} = d_k \tag{38.2}$$

$$0 \le C_k \le d_k \tag{38.3}$$

$$g_{lower} \le g_k \le g_{upper} \tag{38.4}$$

$$f_{lower} \le f_{kj} \le f_{upper} \tag{38.5}$$

The Right Hand Side (RHS) of the power flow problem consists of load, available generation, and power flow capacity of the transmission, as shown in (3). So it changes over time, since load and generation vary every hour. It can be also changed by storage deployment, since the upper bound of generation vector with storage increases. So if we run optimization process every time, it is very time consuming. Sensitivity analysis [46]-[47] is used to calculate the final optimal solution to sum of load curtailments in the power system. By using the basic matrix B (coefficient matrix of

constraints by basic variables), cost vector C_B (coefficient vector of objective function by basic variables), and newly changed RHS b_{new}, it can be determined whether the optimization should be restarted or directly calculated by sensitivity analysis. If C_B times inverse of B is greater or equal to zero, it means that previous basic matrix can hold the problem for feasible availability. So the final optimal solution is directly taken using $Z^* = C_B B^{-1} b_{new}$. Otherwise, optimization should be restarted, since basic matrix is changed by b_{new}.

5.4 Storage Techniques

Renewable energy resources like wind power or solar power have fluctuating characteristics, since wind speed or solar radiation are based on random behaviors. So using renewable energy alone, it is hard to satisfy varying load. To mitigate this problem, energy storage can be added to the IEEE RTS. In general, significant amount of electric energy cannot be stored itself. So it is required to convert into other types of energy like kinetic, potential, or chemical energy and so on. Table 33 shows general storage types [49]-[52].

Figure 33, 34, 35 and 36 illustrate energy storage principles, respectively. Pumped Hydro Energy Storage (PHES) pumps water up using motor in off peak load, and generates power to use turbine in peak load. It requires expensive capital costs and suitable topography. Compressed Air Energy Storage (CAES) compresses air in

underground tank during off peak load, and generates power using heat from expanded air through recuperator. It needs lower construction costs than PHES. Flywheels spins flywheel in vacuum vessel in off peak load, and provides energy using mechanical energy. And batteries are used as storage by charging/discharging process of chemical energy. Figure 36 shows an example of battery complex in Northern Chile with 20MW capacity Lithium ion.

Table 33. Energy Storage Techniques

Types	Usage level	Storage	Capacity	Efficiency
Pumped Hydro Energy Storage (PHES)	Composite	Pumping water	1 [GW]	70 [%]
Compressed Air Energy Storage (CAES)		Compressing air	100-300 [MW]	80 [%]
Battery Complex	Distribution	Chemical process	20-50 [MW]	90-95 [%]
Flywheels		Spinning flywheel	25-30 [kW]	85-90 [%]

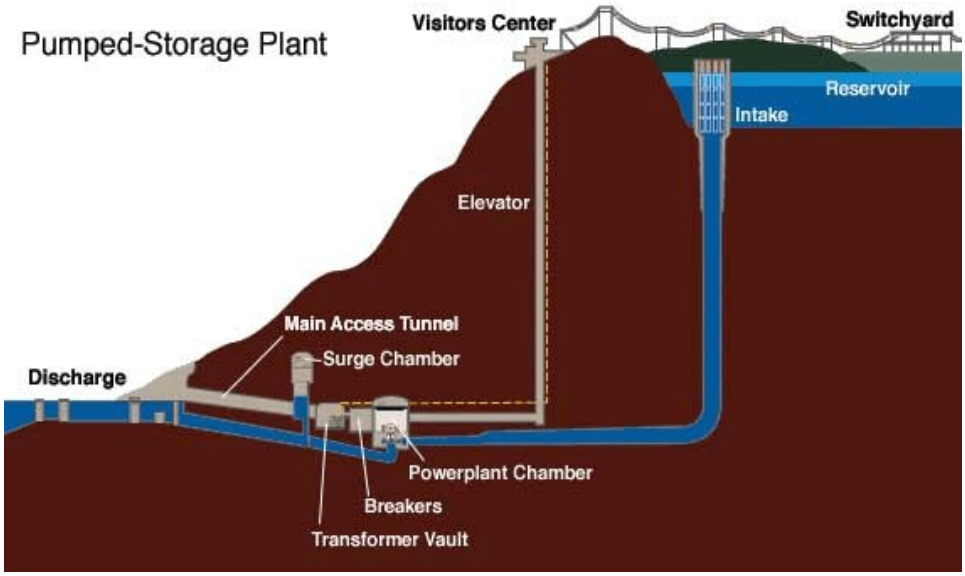

Figure 33. Layout of PHES

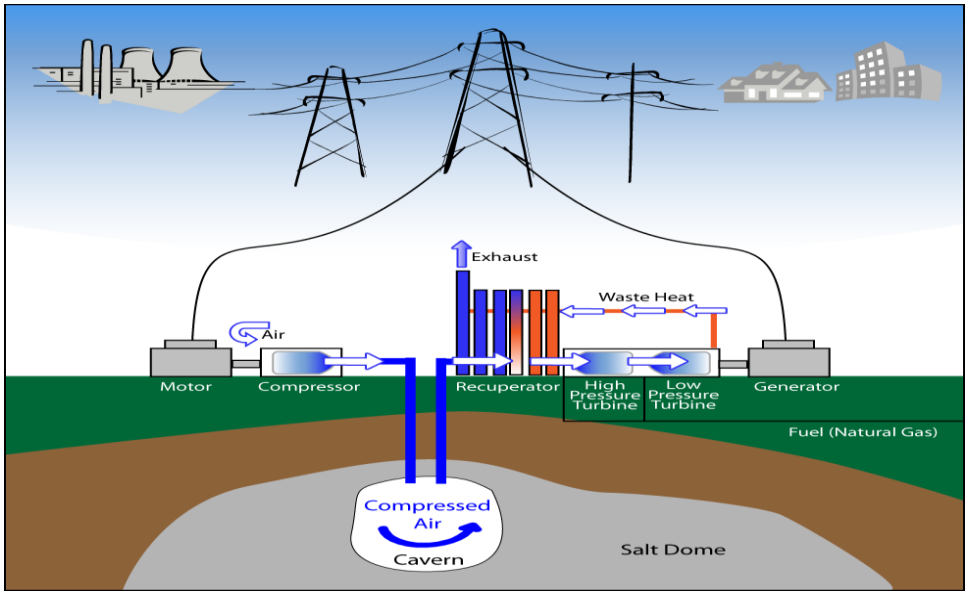

Figure 34. Layout of CAES

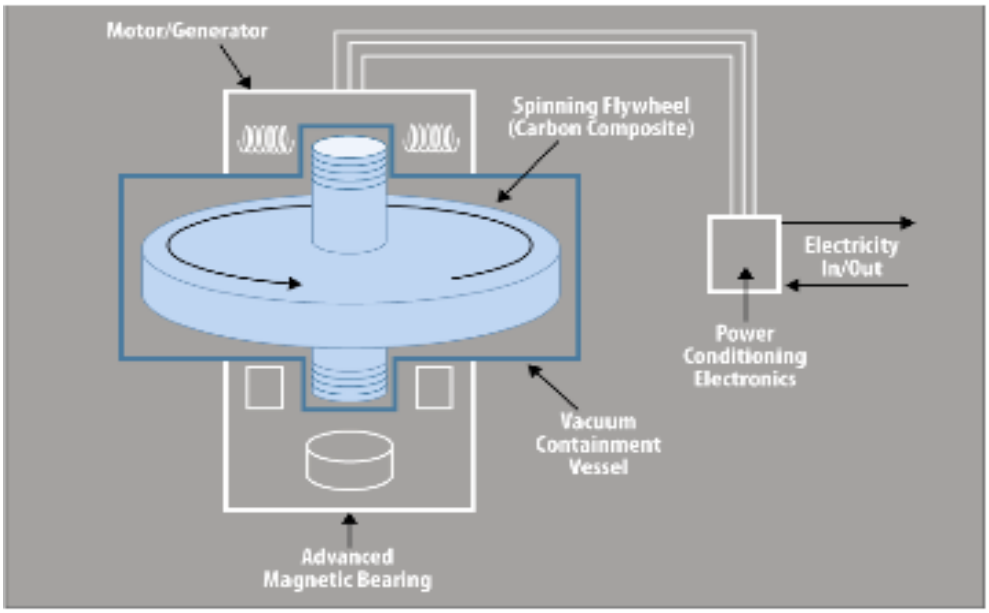

Figure 35. Layout of Flywheel

Figure 36. Example of Battery Complex in Northern Chile

5.5 Integration of Storage with Wind Farms

Energy storage technologies are incorporated on some buses in IEEE RTS. For buses without storage, conventional linear programming is used in (38.1)-(38.5). For buses with storage, the upper limit of the generation constraint is changed by creating storage vector with charge/discharge rates, illustrated in (39). Here x_k is storage vector at bus k, which is calculated using energy balance vector, x_{bk} and charge/discharge rates of storage. x_{bk} is taken from the difference between generation and load at bus k. Figure 37 shows the process of calculation of storage vector at bus k, x_k every time. Here i is the sequence of time, and max_cap is the maximum capacity of storage. Positive x_{bk} means energy storing mode, and negative x_{bk} means energy generating mode.

$$g_{lower} \le g_k \le g_{upper} + x_k \tag{39}$$

5.6 Optimal Storage Deployment

One issue is to determine the buses where to place the storage. The placement of storage becomes important because of the transmission constraints. If there were no transmission constraints, then storage could be placed anywhere.

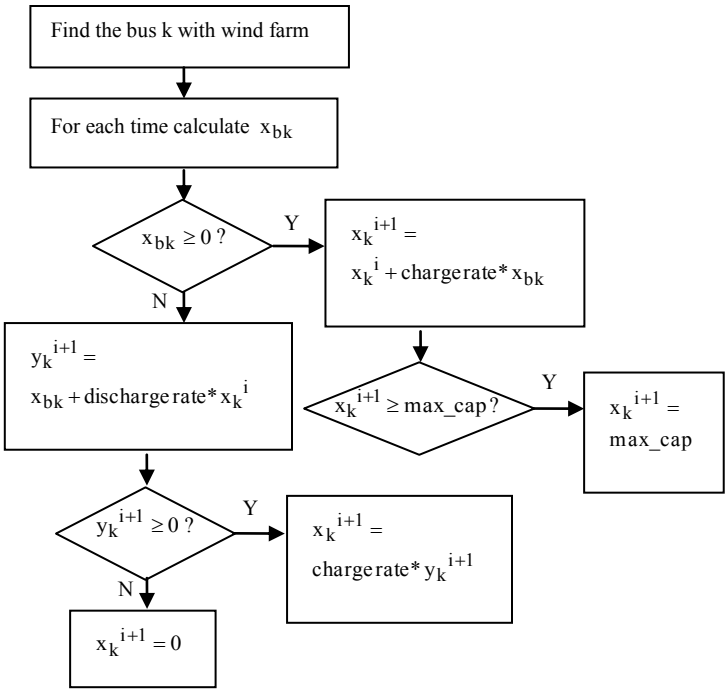

Figure 37. Flowchart of Calculation of Storage Vector

Storage bus deployment cases can be a dimensionally complex problem depending on the power system size. For IEEE RTS [17], there are 24 buses. If we choose 3 buses as storage buses with wind farm, the number of possible locations is 3 combinations out of 24, which is 2024. To find the candidates for optimal storage buses, this dissertation introduces an approach based on expected capacity [3].

The upper limit of generation vector with storage consists of maximum available generation vector and storage vector. Available generation depends on the failure/repair process of turbines. Storage vector is determined by the energy balance vector and charge/discharge rates. So once charge/discharge rates are fixed, the upper limit of generation vector with storage is determined by expected capacity of a bus,

shown in (40). Here N is the number of available generation states at a bus, C_i is the capacity of state i, P_i is the probability of state i. One possible simple approach seems to be to use a bus with higher expected capacity as a candidate of the optimal storage buses with wind farm. The idea is that if the capacity is high then at low load periods, the excess capacity could be used for charging. Then one could perform simulations on a selected number of candidates to make the final choice.

$$\text{Expected capapcity} = \sum_{i=1}^{N} C_i P_i \tag{40}$$

5.7 Case Studies and Results

Figure 38 shows the schematic of the proposed system from viewpoint of a bus. IEEE RTS has 24 buses and 32 conventional generating units having total capacity of 3405 [MW]. The annual peak load is 2850 [MW]. Three buses of the system are assumed to have wind farm and storage. Generation system of IEEE RTS is placed for a swing bus or PV buses. Load is connected in PQ buses. Grid represents the transmission network of the system represented by the bus admittance matrix. Figure 39 shows the layout of wind farm which has 16 identical wind turbines with square by square structure. Here d [m] indicates the diameter of the turbine. Three wind farms are assumed to be installed at different buses. The capacity of each wind farm is 80 [MW], having each wind turbine 5 [MW]. Wind data is from National Renewable Energy

Laboratory (NREL) [10]. The number of site area is 10. The wind speed and wind turbine data are shown in Table 34.

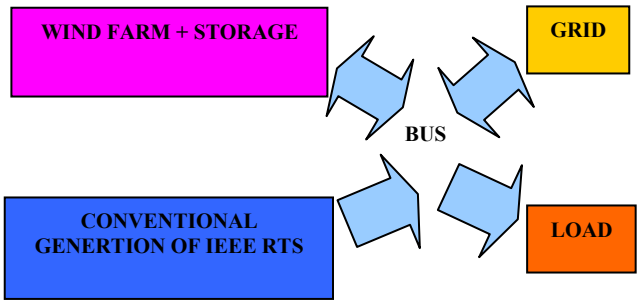

Figure 38. System Configuration from Viewpoint of a Bus

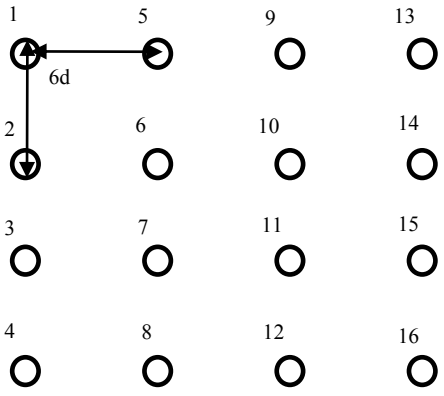

Figure 39. Proposed Wind Farm for Storage

Table 34. Wind Speed and Turbine Data for Storage

Wind speed data	
Peak wind speed [m/s]	27.57
Mean wind speed [m/s]	8.20
Standard deviation of wind speed [m/s]	3.19
Wind turbine data	
Cut in speed [m/s]	6
Rated speed [m/s]	11
Cut out speed [m/s]	19
Rated power [MW]	5
Rotor diameter [m]	80
Hub height [m]	70

Table 35 shows failure/repair rates of conventional units and wind turbines. Transition rate of wind turbines depend on wind speed [30]-[31]. In the proposed transition rate method for wind speed modeling, all transition rates among wind speed states are required to be calculated. For example, Transition rates between state 12 and other state are shown in Table 36 from the original wind speed data. From original wind data, wind speed states are identified in Table 37.Clustering approach for wind speed model is also applied to the system. Table 38 represents the cluster seed and probability using FGFCM. First column of each cluster is mean wind speed, and second one is mean

load. From validity measurement, optimal cluster size is chosen to be seven.

Table 35. Transition Rates of Operating Turbines

Units	**Conventional unit**	**Wind turbine**	
Weather		**Normal speed**	**Extreme speed(>19m/s)**
Failure rate[#/yr]	6	6	36
Repair rate[#/yr]	130	130	36

Table 36. Transition Rate of Wind State 12

Wind speed states		**Transition rates [#/h]**
From	**To**	
12	7	0.0080
	9	0.0399
	10	0.1200
	11	0.2560
	13	0.1520
	14	0.0160
	15	0.0160

Table 37. Identification of Wind Speed States

State	Range [m/s]	Prob.	Freq [#/yr]	Dur [h/#]	Power [MW]
1	0-4	0.0872	321	2.37	0
2	4-5	0.0751	731	0.89	0
3	5-6	0.0996	997	0.87	0
4	6-7	0.1098	1138	0.84	0.43
5	7-8	0.1162	1215	0.83	1.33
6	8-9	0.1199	1202	0.87	2.30
7	9-10	0.1088	1108	0.85	3.33
8	10-11	0.0939	977	0.84	4.43
9	11-12	0.0730	805	0.79	5
10	12-13	0.0487	562	0.75	5
11	13-14	0.0328	347	0.82	5
12	14-15	0.0138	200	0.60	5
13	15-16	0.0082	125	0.57	5
14	16-17	0.0050	89	0.49	5
15	17-18	0.0034	60	0.50	5
16	18-19	0.0017	39	0.37	5
17	19-20	0.0009	20	0.39	0
18	20-21	0.0006	18	0.30	0
19	21-34	0.0011	10	0.99	0

Table 38. Clustering Information Using FGFCM

Clusters	Seeds ([m/s],[MW])	Probabilities
1	(8.21,1750.6)	0.1835
2	(9.02,1362.7)	0.1563
3	(7.26,2274.1)	0.1402
4	(8.73,1111.2)	0.1220
5	(7.91,1990.4)	0.1884
6	(7.87,2488.2)	0.0687
7	(8.23,1540.5)	0.1409

Table 39 shows the expected capacity of IEEE RTS. Probability of each capacity state is calculated using transition matrix approach [2]-[3]. From the Table 39, if we select three buses with storage, there exist two choices for optimal storage deployment; (23, 13, 18) or (23, 13, 21). Using sensitivity analysis to compare reliability indices, the final optimal storage buses can be determined. From that Table, it is also possible to choose three more storage buses as candidates of optimal deployment. For example, if we select five storage buses, there is one choice; (23, 13, 18, 21, 22) by expected capacity order.

Table 40 compares LOLE [h/y] by different wind speed models using two phase method and sensitivity analysis. Wind farms are assumed to be installed at bus 3, 17, and 24. Clustering method is more accurate than exact transition, since it deals with correlation between load and wind speed.

Table 39. Expected Capacity of Bus

Bus	Expected capacity [MW]
1	184.9639
2	184.9639
7	288
13	561.4500
15	207.6000
16	148.8000
18	352
21	352
22	297.0000
23	619.4052

Especially, global or fast global approach of clustering is much closer to using exact original wind data approach by making sure that it is convergent to the global optimum. Instead, it takes longer time than traditional clustering, KM or FCM in Table 41. The running time is the duration of clustering process for the optimal cluster size which is determined by validity measurement. Fast global approach accelerates the simulation speed. Table 42 shows difference between without and with sensitivity analysis using clustering FGFCM. LOLE [h/y] is almost the same. With sensitivity, however, the number of optimizations required during simulation period, one year significantly decreases, 12 times in this case. As the proposed wake model is incorporated on the system, reliability level drops in Table 43. As peak load increases,

LOLE [h/y] also goes up.

Table 40. LOLE [h/y] by Different Wind Speed Model

LOLE[h/y]	**Exact transition**		**original**			
	26.399		22.90			
	Clustering					
	KM	FCM	GKM	FGKM	GFCM	FGFCM
	25.37	25.91	22.55	23.21	23.00	22.74

Table 41. Running Time of Different Clustering Approaches

Time [min]	**Clustering**					
	KM	**FCM**	**GKM**	**FGKM**	**GFCM**	**FGFCM**
	0.012	1.21	73.2	59.53	391.8	65.4

Table 42. LOLE [h/y] without and with Sensitivity Analysis

LOLE[h/y]/number of optimizations	**FGFCM**	
	Without sensitivity	**With sensitivity**
	22.6421/8736	22.7418/723.474

Table 43. LOLE [h/y] without and with Wake Effect

Peak load[MW]	Without wake	With wake
2750	11.3741	14.3104
2850	22.7418	25.6343
2950	38.1095	41.0188

To determine the optimal storage buses, Table 44 compares LOLE [h/y] without and with storage for candidates of optimal bus using FGFCM with wake effect. As can be seen, for cases without storage, LOLE [h/y] is almost the same regardless of the location of wind farms. And we know that the final optimal storage buses with wind farms should be bus 23, 13, and 18 by observing the changes of LOLE [h/y]. Table 45 shows LOLE [h/y] for selected optimal storage buses, 23, 13, and 18 by different peak load. For a case with peak load 2850[MW], reliability indices are compared by different storage performances in Figure 40 and Figure 41. As charge/discharge rates and capacity of storage increases, LOLE [h/y] tends to decreases. Finally, EENS [MWh/y] becomes greater for higher peak load, shown by Figure 42.

Table 44. LOLE [h/y] without and with Storage

Cases	Without storage		With storage	
Bus location of wind farms	23,13,18	23,13,21	23,13,18	23,13,21
LOLE[h/y]	25.6535	25.6543	21.8774	22.3725

Table 45. LOLE [h/y] of Optimal Storage

LOLE [h/y]	
Peak load[MW]	**Optimal storage**
2750	12.3003
2850	21.8774
2950	36.7522

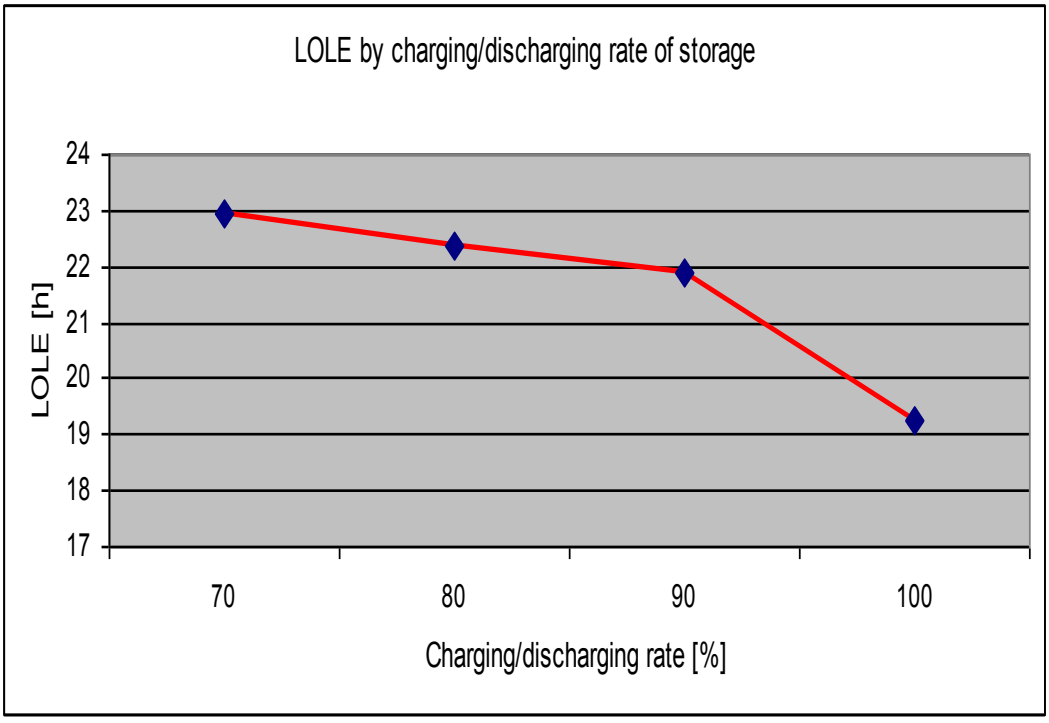

Figure 40. LOLE [h/y] by Different Charge/Discharge Rates

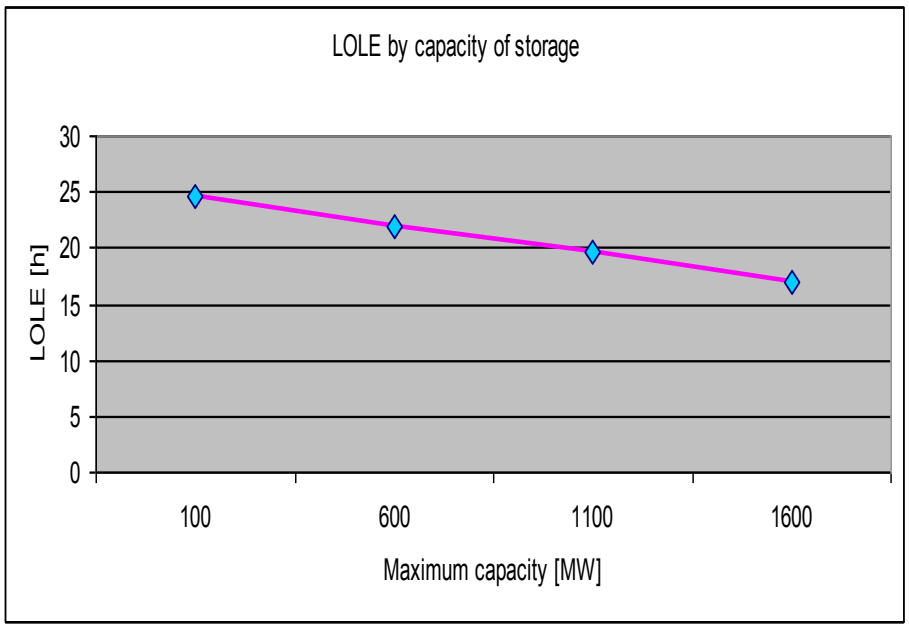

Figure 41. LOLE [h/y] by Different Maximum Capacity

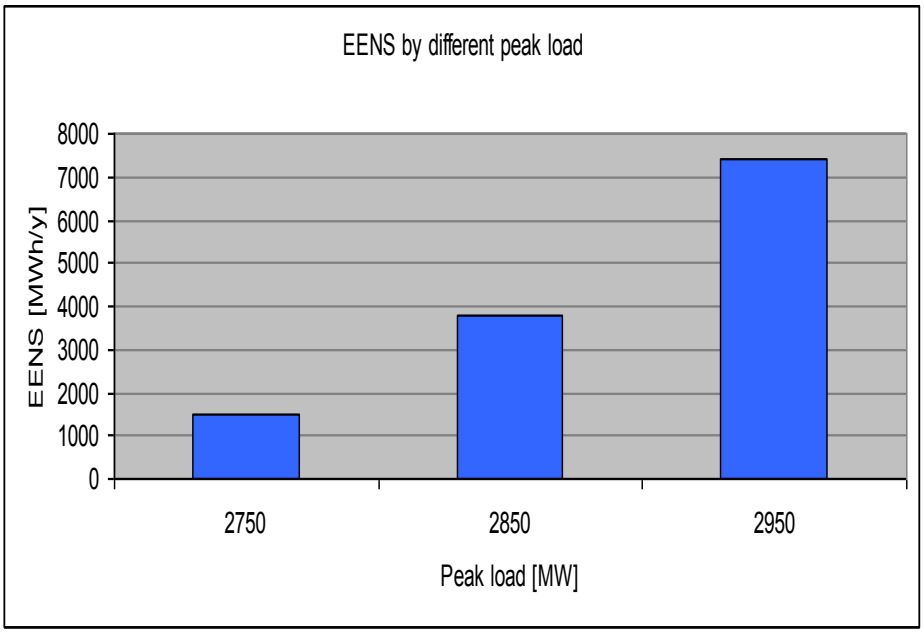

Figure 42. EENS [MWh/y] by Different Peak Load

6. CONCLUSIONS

The proposed power system is modeled by a combination of conventional generation, wind farms, PV systems, and energy storage. From some original wake models, newly developed wake models are proposed for the wind farm to generate more accurate waked speed by considering cumulative wake effect, wind shade and shear effect in a wind farm. Improved N.O. Jensen model or Eddy Viscosity works well as a wake model. However, modified Larsen model shows some deficiencies as a wake model for a wind farm with cumulative wake effect or wind shear. As wake effect is incorporated into the system, reliability level drops due to energy losses. This means that not including wake effect can over estimate reliability.

Turbines are modeled on two generating states; fully available and out of service. For wind speed modeling, birth and death process and transition rate matrix approach are discussed. In birth and death Markov chain, each state only moves to the next neighboring state. If the sampling time is small which is close to zero, wind speed can be considered to change smoothly over time. In this situation, the model suggested by [1] would represent the physical reality correctly. However, in practice sampling is done at intervals like 10 minutes. In such a situation speed cannot be assumed to transit smoothly and transitions to remote states can occur more frequently. So if the birth and death model is used for a wind data sampled at finite intervals as is the case in practice, some transitions between wind states can be lost. This dissertation introduces a transition rate matrix approach [2]-[3] by which all possible transition rates between states from

original wind data can be captured. This approach is useful when load and wind can be assumed to vary independently of each other. However, in practice, there is correlation between them. To deal with it efficiently, clustering approach is applied to power systems with renewable energy. Some different clustering algorithms are presented and compared. As renewable energy is incorporated into the composite power system, different dimensional clustering approaches are demonstrated in details. For the failure/repair process of turbines, Monte Carlo simulation including random sampling and next event method is applied for the proposed power system.

The flow model embedded in the linear program is DC power flow. To ensure that an initial basic feasible solution is available, artificial variables are added to the original constraints. As one approach to use artificial variables, two phase method is applied to the system to get reliability indices. Optimization process is needed for every hour of simulation. Using sensitivity analysis, we can reduce simulation running time. To regulate the fluctuation of renewable energy, energy storage is integrated into the proposed power system. Optimal storage bus using sensitivity analysis and clustering method are chosen and compared. Simulation methodology to select the optimal storage buses is developed and applied to the system, using LOLE [h/y], EENS [MWh/y], or ENSI [MWh/#] to figure out the system reliability explicitly. The upper limit of generation vector at the bus is updated by the storage vector. As storage is added to the system, it is observed that the reliability is improved.

REFERENCES

[1] F.C. Sayas, and R.N. Allan, "Generation availability assessment of wind farms," *in Proc. IEEE Gener. Transm. Distrib.* Vol. 143, no. 5, pp. 507-518, Sep. 1996.

[2] C. Singh and R. Billinton, *System Reliability Modelling and Evaluation*, 1st ed. London, U.K.: Hutchinson Educational, Jun. 1977.

[3] "*Electrical Power System Reliability*," Dr Singh's Homepage, [Online]. Available: http://www.ece.tamu.edu/People/bios/singh/, May 2013.

[4] C. Singh, and Y. Kim, "An efficient technique for reliability analysis of power systems including time dependent sources," *IEEE Trans. Power Syst*., vol. 3, no. 3, pp. 1090-1096, Aug. 1988.

[5] "*Cubic Clustering Criterion (CCC)*," SAS technical report, [Online]. Available: http://support.sas.com/documentation/onlinedoc/v82/techreport_a108.pdf, May 2013.

[6] F. Koch, M. Gresch, F. Shewarega, I. Erlich, and U. Bachmann, "Consideration of wind farm wake effect in power system dynamic simulation," *in Proc. IEEE Power Tech. Conf*., vol. 37, no. 1, pp. 1-7, Jun. 2005.

[7] Vaughn Nelson, *Wind Energy: Renewable Energy and the Environment*, 1st ed. Boca Raton, FL: CRC Press, Apr. 2009.

[8] "*The Prediction of the Energy Production of a Wind Farm*," Wind Energy the Facts. [Online]. Available: http://www.wind-energy-the-facts.org/, May 2013.

[9] C. Singh and J. Mitra, "Monte Carlo simulation and intelligent search methods", *IEEE Tutorial on Electric Delivery System Reliability Evaluation (editor: J. Mitra)*, pp 23–38. IEEE, Piscataway, NJ: Jul. 2005.

[10] R. Billinton and X. Tang, "Selected considerations in utilizing Monte Carlo simulation in quantitative reliability evaluation of composite power systems," *Elect. Power Syst. Res.*, vol. 69, no. 2–3, pp. 205–211, May 2004.

[11] M. T. Schilling, J. C. Stacchini de Souza, and M. B. Do Coutto Filho, "Power system probabilistic reliability assessment-current procedures in Brazil," *IEEE Trans. Power Syst*., vol. 23, no. 3, pp. 868–876, Aug. 2008.

[12] R. Billinton, and R. N. Allan, *Reliability Evaluation of Power Systems*, 2nd ed. New York: Plenum, Jan. 1996.

[13] J. Wen, Y. Zheng, and F. Donghan, "A review on reliability assessment for wind power," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 9, pp.2485-2494, Dec. 2009.

[14] W. Yichun, "Research on reliability evaluation and expansion planning of power system including wind farms," PhD Dissertation in Wind Energy from the Hefei University of Technology, Jul. 2006.

[15] "*Western Wind Resource Dataset*," National Renewable Energy Laboratory (NREL). [Online]. Available: http://wind.nrel.gov/Web_nrel/, May 2013.

[16] "*Local Climatological Data in National Climatic Data Center (NCDC)*" [Online]. Available: http://www.ncdc.noaa.gov/most-popular-data, Jan. 2013.

[17] Power Systems Engineering Committee, "The IEEE reliability test system," *IEEE Trans. Power App. Syst*., vol. PAS-14, no. 3, pp.1010–1020, Aug. 1999.

[18] Douwe J. Renkema, "Validation of wind turbine wake models," Master Thesis in Wind Generation from Delft University of Technology, Jun. 2007.

[19] David Ryan VanLuvanee, "Investigation of observed and modeled wake effects at horns rev using Windpro," Master Thesis in Wind Energy from the Technical University of Denmark, Aug. 2006.

[20] H. Kim, C. Singh, and A. Sprintson, "Comparison of wake models for reliability analysis", *The 12th International Conference on Probabilistic Methods Applied to Power Systems*, Istanbul, Turkey, Jun. 2012.

[21] X. Zhang and W. Wang, "Wind farm and wake effect modeling for simulation of a studied power system," *IEEE Power Systems Conference and Exposition (PSCE)*, vol. 3, no. 2, pp. 1-6, Aug. 2010.

[22] "*Wind Farm Wake Effect Model in WAsP8*," The Wind Atlas Analysis and Application Program. [Online]. Available: http://www.risoe.dk/vea/storpark/presentations/WAsP8%20Wake-effect%20model.pdf, Dec. 2012.

[23] J.F. Ainslie, "Calculating the flowfield in the wake of wind turbines," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 27, no. 1-3, pp. 213-224, Jan. 1988.

[24] G.C. Larsen, "A simple wake calculation procedure," *RISO National Laboratory*, DK-4000 Roskilde, Denmark, Dec. 1988.

[25] S. Mathew, *Wind Energy: Fundamentals, Resource Analysis and Economics*, 1st ed. New York: Springer, Jun. 2006.

[26] "*Surface Roughness Length*", The Meteorological Resource Center. [Online]. Available: http://www.webmet.com/met_monitoring/663.html, Jan. 2013.

[27] P. Giorsetto, and K. F. Utsurogi, "Development of a new procedure for reliability modeling of wind turbine generators," *IEEE Trans. Power App. Syst.*, vol. PAS-102, no. 1, pp. 134-143, Jan 1983.

[28] C. Singh, and A. Lago-Gonzalez, "Reliability modeling of generation systems including unconventional energy sources," *IEEE Trans. Power App. Syst.*, vol. PAS-104, no. 5, pp. 1049-1056, May 1985.

[29] H. Ascher and H. Feingold, *Repairable Systems Reliability Modeling, Inference, Misconceptions, and Their Causes*, 1st ed. New York: Marcel Dekker, Mar. 1984.

[30] J.A. Carter, B. Johnson, R.W. Sherwin, Rade-Makers, and L.W.M.M. "Failure modes and effects analysis for the AOC 15/50 wind turbine," *Proc. of the AWEA Wind Power Conf.*, pp. 175-182, Jul. 1993.

[31] D. Quarton, "Technical and economic effects of connecting increasing numbers of large wind turbines to weak public utility networks," *Garrad Hassan and Partners LTD*, New York, Dec. 1992.

[32] R. Xu and D. C. Wunsch, *Clustering*, IEEE Press Series on Computational Intelligence, John Wiley & Sons Inc., Feb. 2009.

[33] H. Kim and C. Singh, "Comparison of clustering approaches for reliability simulation of a wind farm," *Power System Technology (POWERCON) 2012 IEEE International Conference*, Oct. 2012.

[34] Siddheswar Ray and Rose H. Turi, "Determination of number of clusters in K-means clustering and application in colour image segmentation," *Int. Conf. on Advances in Pattern Recognition and Digital Techniques (ICAPRDT),* ISBN: 81-7319-347-9, pp 137-143, Dec. 1999.

[35] D. Zigras and K. Moennich, "Farm efficiencies in large wind farms," *Deutsches Windenergie-Institut (DEWI)*, New York, Nov. 2006.

[36] N. R. Pal, K. Pal, J. M. Keller, and J. C. Bezdek, "A possibilistic fuzzy c-means clustering algorithm*," IEEE Trans. Fuzzy Syst*., vol. 13, no. 4, pp. 517-530, Aug. 2005.

[37] A. Likas, M. Vlassis, and J. Verbeek, "The global k-means clustering algorithm," *Pattern Recognition*, vol. 36, pp. 451–461, Mar. 2003.

[38] W. Wang, Y. Zhang, Y. Li, and X. Zhang, "The global fuzzy c-means clustering algorithm," in *Proc. of 6th World Congress on Intelligent Control and Automation Conf.*, pp. 3604-3607, Jul. 2006.

[39] K. G. T. Hollands and R. G. Huget. "A probability density function for the clearness index with applications," *Solar Energy*, vol. 30, pp. 195-209, Jun. 1983.

[40] R. M. Moharil and P. S. Kulkarni. "Reliability analysis of solar photovoltaic system using hourly mean solar radiation data," *Solar Energy*, vol. 84, pp. 691-702, Apr. 2010.

[41] H. Wang and X. Bai, “Adequacy assessment of generating systems incorporating wind, PV and energy storage,” *IEEE Innovative Smart Grid Technologies*, vol. 3, no.4, pp. 1-6, May 2012.

[42] Y. M. Atwa, “Distribution system planning and reliability assessment under high DG penetration,” PhD Dissertation in Power Generation from the University of Waterloo, Apr. 2010.

[43] H. Yang, F. S. Wen, and L. Wang, “Newton-Raphson on power flow algorithm and Broyden Method in the distribution system,” *Power and Energy Conf., IEEE 2nd International*, pp. 1613-1618, Dec. 2008.

[44] R. A. Messenger and J. Ventre, *Photovoltaic Systems Engineering*, 2nd ed. CRC Press, Feb. 2000.

[45] H. Kim, C. Singh, and A. Sprintson, "Simulation and estimation of reliability in a wind farm considering the wake effect," *IEEE Trans. Sustainable Energy*., vol. 3, no. 2, pp. 274-282. Apr. 2012.

[46] MokhtarS. Bazaraa, John J. Jarvis, and Hanif D. Sherali, *Linear Programming and Network Flows*, 4th ed. Wiley, New York, Dec. 2009.

[47] “*Practical Optimization: A Gentle Introduction*” [Online]. Available: http://www.sce.carleton.ca/faculty/chinneck/po.html, Dec. 2012.

[48] H. Kim, and C. Singh, “Reliability modeling and simulation in power systems with aging characteristics,” *IEEE Trans. Power Syst*., vol. 25, no. 1, pp. 21-28, Feb. 2010.

[49] A Report by the Electricity Advisory Committee, "Bottling electricity: storage as a strategic tool for managing variability and capacity concerns in the modern grid," *Office of Electricity Delivery and Energy Reliability*, vol. 53, no. 1, pp. 1051-1058, Dec. 2008.

[50] P. Taylor, "Companies race to develop utility-scale power storage," *New York Times*, vol. 1, no. 2, pp. 1-4, Sep. 2009.

[51] D. Rastler, "New demand for energy storage: electric perspectives," *Electric Perspectives*, vol. 33, no. 5, pp. 30–47, Sep. 2008.

[52] C. Vartanian, "The coming convergence, renewables, smart grid and storage," *IEEE Energy 2030*, vol. 99, no.6, Nov. 2008.