Smart Operation for AC Distribution Infrastructure Involving Hybrid Renewable Energy Sources

Ahmed A. Mohamed, Mohamed A. Elshaer and Osama A. Mohammed
Energy Systems Research Laboratory, Department of Electrical and Computer Engineering, Florida International University, Miami, Florida, USA, (e-mail: mohammed@fiu.edu).

Abstract: This paper presents an effective algorithm for optimizing the distribution system operation in a smart grid from cost and system stability points of view. This algorithm is mainly dependent on forecasted data of the power available from different renewable energy sources as well as the load demand. Hence, full attention was paid to the forecasting process. A non-linear regression technique was applied to build accurate forecasting models for different sources and load conditions. These models help in monitoring and predicting the total power generation and demand online. The main objective of the optimization process is to control the power shared from different sources such that they satisfy the load demand with the least cost while giving higher priority to renewable energy sources. Moreover, batteries are controlled in such a way that they are allowed to discharge only when there is no very big load predicted within a future period so that they become available to act as a buffer for a predicted large load increase which may affect the stability of the system and hence reduce voltage dips. A fuzzy controller was utilized to determine how many hours before a predicted occurrence of load the battery system should be turned off and what percentage of power should be taken out of the batteries and the grid to solve this issue. Different case studies were investigated to verify the validity of the proposed algorithm and define the system performance under several conditions.

Keywords: Optimization, energy storage, commitment, fuzzy controller, smart grid, renewable energy.

1. INTRODUCTION

Challenges related to existing distribution systems such as demand response, distributed energy source integration and distributed energy storage impose an imperative necessity to leap forward towards smart operation of distribution systems. Researchers have worked on various ways to make a futuristic power system able to easily integrate renewable energy sources. Atwa et al. investigated different renewable energy sources available and worked on finding the optimal mix to minimize the system energy losses. They achieved a significant reduction in the annual energy losses for the proposed scenarios. Liserre et al. investigated the integration of different renewable energy sources into the smart grid. Battery and energy storage systems generally are involved in such research. Smith et al. discussed the state-of-the-art status of renewable energy sources and energy storage systems in rural areas. They introduced methods to deal with non-dispatchable renewable sources.

A system that is mainly dependent on renewable energy sources should have another backup supply to cover any deficiency in the power demand during peak loads. Battery system can then be used for this purpose. However, power which can be drawn from the batteries depends on its state of charge (SOC). In this paper, we are presenting a unit commitment problem that may face a grid-connected distribution system which depends mainly on renewable energy sources to supply its loads.

The main contributions of this paper can be summarized as follows:

1. Mathematical models for PV, wind and load demand were obtained based on previous actual data from available databases.
2. A fuzzy controller was designed to control the amount of power that should be taken out of each of the battery system and the grid in case of power deficiency to cover the load demand, while maintaining the high voltage stability of the system
3. The unit commitment problem of a prototype system containing PV, wind, battery system and loads was investigated in the case of system grid connection.

2. SYSTEM AND PROBLEM DESCRIPTION

In order to examine the proposed commitment scheme, a prototype system was used. A single line diagram of the system under study is given in Fig. 1. This system is an AC distribution system which depends mainly on renewable energy sources to supply its local loads. However, the system includes a backup battery system that can support load deficiencies. In addition, the system is connected to the main grid, which also can supply the load in case the power available from different renewable energy sources is insufficient. However, in this paper we try to make the system under study as self satisfied as possible. This means that we are minimizing the power drawn from the grid. The maximum peak load is assumed as 300 kW. The PV system has a capacity of 100 kW. Whereas, the wind system has a capacity of 150 kW.

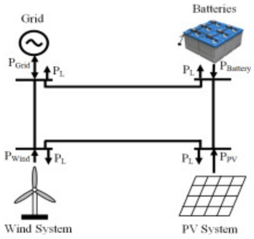


Fig. 1. The prototype system under study.

3. DATA FORECASTING

3.1 Data Collection

In order to solve the commitment problem involving renewable energy sources and coordinate the sources in an economic way, information about the total generation available out of renewable energy sources as well as the load demand should be known in advance. Hence, we count on real data forecasting of PV and wind output power as well as the demand. The data forecasting process has been based on PV data collected over 15 years on an hourly basis for a unit in the state of Texas, wind data collected over four years period on a 10 minutes basis and load data over four years period on an hourly basis for the same region. A non-linear regression modelling technique has been employed to mathematically model the output power of each of the renewable sources and the load demand. Different model evaluation indices have been used to validate the mathematical models obtained.

3.2 Non-Linear Regression Modelling

The non-linear regression model used here has the ability to cope with the non-linearity of the data and form an accurate model. It is based on the idea of transformation of the data using a pre-defined set of non-linear functions in order to achieve linearity (Wahab et al.). The non-linear model given in and designated as Y_{nlm} has the following form:

$$Y_{nlm} = b_0 + y_1 + y_2 + \cdots + y_m = b_0 + \sum_{i=1}^{m} y_i\,, i = 1, 2, \dots, m. \tag{1}$$

$$y_i = a_{i1}b_{i1}f_1(x_i) + a_{i2}b_{i2}f_2(x_i) + \cdots + a_{ij}b_{ij}f_j(x_i) + \cdots + a_{ik}b_{ik}f_k(x_i) \tag{2}$$

where:

K	is the total number of non-linear functions.
m	is the total number of variables included
y_i, $i=1, 2,\dots, m$.	is a non-linear model for each variable and it is the summation of all terms resulting from transforming input x_i through a pre-selected set of non-linear functions.
a_{ij}, b_{ij}	are the constants that need to be determined, $j=1, 2, \dots, k$.
$f_1, f_2, \dots, f_k$	are pre-selected set of non-linear functions which will be used for transformation of inputs. The set of non-linear functions may contain x^{α}, $1/x$, e^{-x} and $\ln(x)$.
x_i	is the numerical values for a given input to be used for deducing the model.

A detailed description of the way to calculate the non-linear regression model parameters is given in Appendix A (Wahab et al.).

3.3 Model Evaluation Indices

Different model evaluation indices have been implemented to measure the accuracy of the proposed mathematical models. They are the mean absolute percentage error (MAPE) calculated by (3) and the coefficient of determination R^2 computed by (4):

$$MAPE = \frac{\{\sum|\vec{d}-\vec{y}|\times 100/|\vec{y}|\}}{n} \tag{3}$$

$$R_d^2 = \frac{\sum(\vec{d}-\vec{y})^2}{\sum(\vec{y}-y_{avg})^2} \tag{4}$$

Where, $\vec{d}$ and $\vec{y}$ are the vectors of real and predicted data, respectively. The value of R_d^2 for a model is ranging from 0 to 1 and it implies that R_d^2 of the sample variation is attributable to or explained by one or more of the variables as long as it approaches unity. The better regression fits the data the closer the value of R_d^2 is to one.

3.4 Mathematical Modelling Results

Mathematical models for PV and wind systems output power in addition to the load demand have been deduced. These mathematical models are given by (5)-(7), respectively.

$$P_{PV} = -962.8 + 5915H^{0.4} + 797.6D^{0.8} + 114.5H^{0.7}D^{0.8} \tag{5}$$

$$P_{Wind} = 30.34 - 106H^{0.4} + 17.11D^{0.7} + 3.22H^{0.4}D^{0.8} + 10.87H^{0.4}D^{0.7} \tag{6}$$

$$P_{Load} = 120 + 320H^{0.8} + 84D^{0.4} + 3.22H^{0.7}(\tfrac{100}{D}) \tag{7}$$

Where, H and D are the hour and month, respectively.

The PV mathematical model was trained using the sets of data of the fourteen previous years. However, the model was tested using the data of the most recent year, which were not included during the training process. Figure 2 shows the mathematical modelling results of the PV data versus the actual data. We notice that the model results are successfully tracking the actual ones during the whole year. The *MAPE* of this model is 4.65, which is a reasonable value taking into consideration that we are minimizing the inputs to the model (variables of the non-linear functions) to only time bases. However, if we were to take other inputs related to environmental variations corresponding to sun radiations, we would definitely obtain a more detailed model as these inputs are much more correlated to the output power of the PV than just time. Moreover, the value of R_d^2 is 0.951, which means that the transformed inputs used are representative of the output power of the PV system. Here, we try to count on only the time to predict the output power. Non-linear regression is helpful in this case as it transforms sets of inputs into other forms that are more correlated to the desired output.

The Wind data were categorized into two groups; data of the first two years available were used as training data and data of the most recent year was used as testing data. Figure 3 shows the mathematical model results versus the actual data of the wind. We can also notice that the model is successfully representing the actual data. The *MAPE* was 6.1%. Such a small value proves the accuracy of the model. Moreover, the value of R_d^2 was 0.941, which is again acceptable.

The load data of four consequent years were used to model the load duration curve and they have been categorized as follows, data of three years as training data whereas data of another year as testing data. Figure 4 shows the results of the obtained load demand model. Actual and modelling data are close to each other, which validates the model obtained. Moreover, values of the *MAPE* and R_d^2 are 6.45 % and 0.934, respectively. The value of *MAPE* is relatively small. Whereas, the value of R_d^2 is close to one. These two facts support our conclusion that the mathematical model is well representing the actual data.

4. THE UNIT COMMITMENT PROBLEM

The main objective of the unit commitment problem solved in this paper is to minimize the power drawn from the grid, keep the battery's SOC above 60% to be ready as a buffer for sudden large loads and to use the energy stored in the batteries to shift peaks and, consequently, save money.

Hence, intuitively we commit both the PV and the wind systems to supply all the power available in them. This means that both sources are working at the maximum power point tracking level.

In this kind of systems we generally have two different scenarios. Firstly, if the power available from renewable energy sources exceeds the load demand, the power is injected back to the grid or used to charge the batteries. Secondly, if the load demand is larger than the power available from renewable energy sources, we have power deficiency as given by (8).

$$P_d = P_{load} - (P_{PV} + P_{Wind}) \tag{8}$$

Where, P_d is the power deficiency.

Generally, we have two different sources to supply this deficiency in power. That is either by using the power stored in the battery system or using the grid. In this case, as we previously stated the objective here is to make the system as self-dependent as possible. Hence, the priority is given to the batteries to supply the deficiencies. However, if it is predicted to have a big peak load within the coming few hours the priority is given to have the batteries ready with a relatively high state of charge (SOC) by the time of occurrence of that peak load. The purpose of this is to maintain high voltage stability of the system while minimizing the cost.

Moreover, a special care has been given to whether it is a peak or an off-peak hour as the cost of energy is different in both cases. Managing the power corresponding differently corresponding to peak and off-peak hours reduces the total annual cost. The commitment problem is run continuously. This means that the futuristic load and total supply powers are predicted and based on these values in addition to the time at which the coming peak load is taking place and the current SOC of the batteries, the percentage of power that will be taken from each of the grid and the batteries will be decided. The mathematical models derived are used to forecast the peak load and the hour of its occurrence as well as the renewable energy power. In addition, in one of the cases they will be used to calculate the energy that will be required during the coming peak hours by integrating the area under the power curve. Moreover, a fuzzy system is used to solve a part of this commitment problem as fuzzy systems have the ability to solve this type of complicated problems.

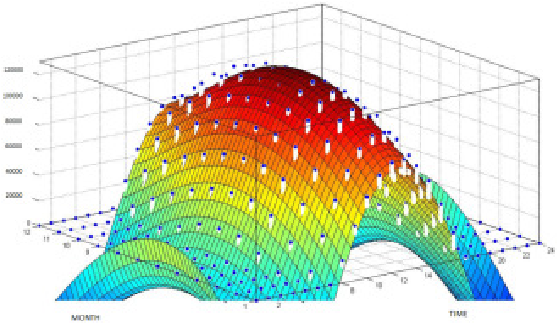


Fig. 2. PV actual versus modelling data for one year

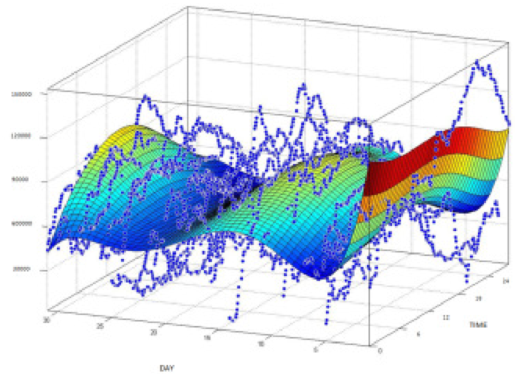


Fig. 3. Wind actual versus modelling data for one month

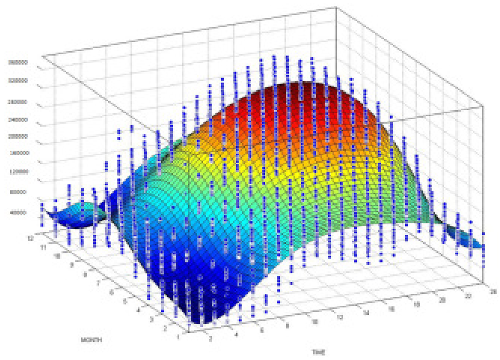


Fig. 4. Load demand actual versus modelling data for one year

The mathematical models derived were used to predict the peak load and generation available at the time it occurs. At peak load, the partial derivative of the curve with respect to hours tends to go to zero. Hence the hour at which the peak load will take place at a given day can be calculated as in (9)

$$\left.\frac{\partial P_{load}}{\partial H}\right]_{D=D1} = 0 \tag{9}$$

Where, D_1 is the day in which we are calculating. Solution of (9) yields the hour H_1, which is the hour at which the coming peak load is taking place. Substituting in (5), (6) and (7) with the value of H_1, we get values of the load demand, PV and wind output power. These values are P_{load1}, P_{PV1} and P_{wind1}, respectively. The energy during the coming peak hours, used in the energy management algorithm proposed in this paper, is predicted as follows:

$$E = \int_{H_{min}}^{H_{max}} P_{load} . dH \qquad (10)$$

Where, H_{min} and H_{max} are the starting and end hour of the coming peak period.

Customers can save an average of 6-7% annually over the basic plan by shifting some energy use to off-peak hours as shown in Fig. 5. This has been taken into consideration in order to have the economic operation of the system.

Fig. 5. Peak and Off-peak hours

The available sources are, PV and Wind (P_{PV}+P_{Wind}), The battery during the discharging mode (P_{batt}), Utility Grid (P_u). The loads are, Normal loads (P_{load}) and the battery during the charging mode (P_{batt})

The proposed algorithm has three inputs; the difference between the renewable power (P_{PV}+P_{Wind}) and the load demand, if the current time is within peak hours or not, SOC of the battery, predicted renewable power at the hour of the coming week and the hour at which it occurs.

There are two possible cases:

Case 1: surplus in power $P_{PV} + P_{Wind} - P_{load} \geq 0$

If H lies within the Off-peak hours, $P_{surplus}$ goes 100% to charge the battery system and the rest is injected to the grid

If H lies within the peak hours, the power follows this proposed pattern as shown in Fig. 6.

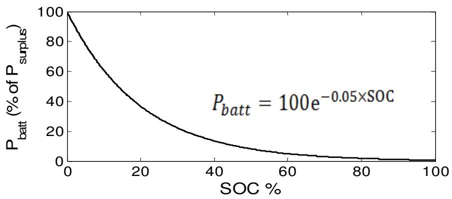


Fig. 6. Battery power as a function of its SOC

Case 2: deficiency in power $P_{PV} + P_{Wind} - P_{load} < 0$

If the SOC is less than 60%, we disconnect it to charge it when there is a surplus in power. Whereas, if the SOC is greater than 60%, the algorithm goes in with the following conditions:

If H lies within the peak hours, 100% of the power stored in the battery system is supplied to the loads to help covering P_{def}, the rest is drawn from the grid

If H lies within the Off-peak hours, P_{def} is covered partially by the battery system according to a fuzzy system that has been proposed.

A block diagram of the complete energy management algorithm is as shown in Fig. 7.

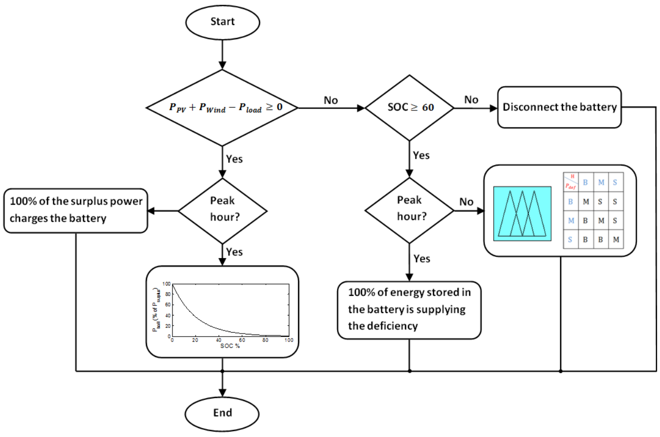


Fig. 7. A block diagram of the energy management algorithm applied to the smart battery charger proposed in this paper

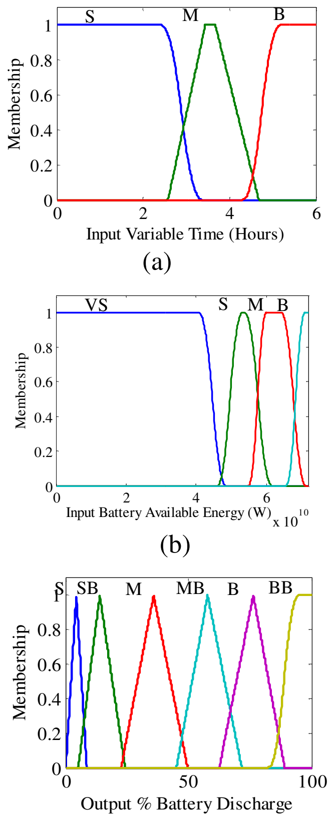
5. FUZZY SYSTEM

Fuzzy control is a powerful method which can be applied to different systems. It is based on the experience of the user about the system behaviour rather than modelling the system under control mathematically such as linear control theory. This makes fuzzy control a powerful technique especially with non-linear systems in which it is difficult to derive an accurate approximated mathematical model of the system and expect its behaviour. Fuzzy control is a rule-based control technique that is approached by linguistic fuzzy rules, which describe the output desired out of the system under different operating conditions. Fuzzy rules are in the form of if-then rules that the proficient user should design such that they cover all the conditions the system is expected to go through.

In this model fuzzy is used only when the instantaneous demand load is higher than the instantaneous available source from renewable and the system is not in the peak hour zone. At this state the battery will be operated at the discharge mode. Hence, Fuzzy determines the amount of power to be extracted out of the battery while taking in consideration the time left for the peak hour, the SOC of the battery and energy needed during the peak hour.

Designing a fuzzy logic controller is achieved through three basic steps; fuzzification, inference mechanism and defuzzification. The Sugeno type fuzzy system was used in this paper.

In fuzzification, the time left for the peak hour and the current SOC of the batteries are the inputs to the control system which are mapped into a certain linguistic values. The output of the fuzzy is a percentage that determines the percentage of load to be satisfied by the batteries. Three fuzzy memberships have been used in this paper; these are shown in Fig. 8. Each membership is divided into fuzzy subsets. Hence, the input battery available energy has four subsets; Very small (VS), small (S), medium (M) and big (B). For the other input, current time, three fuzzy subsets are used; small (S), medium (M) and big (B). The output is best represented by six fuzzy subsets; small (S), small big (SB), medium (M), medium big (MB), big (B) and big-big (BB). The B membership function is the one to the most right and other membership functions appear to the left in the same order mentioned at the beginning of this paragraph. These membership functions are used to map the input variable into fuzzy set. Operation of the membership functions on the input variable yields the extent to which that variable is a member of a particular rule. The process of converting control variables into linguistics rules is called fuzzification. However, in inference Engine and Rule base step, the output of fuzzy controller is managed through putting certain linguistic rules. These control rules are constructed based on given conditions (inputs) such that the fuzzy controller decides the proper control action.



(a)

(b)

(c)

Fig. 8. Membership functions of input variables of the fuzzy controller.

Finally, in defuzzification, as the output of the fuzzy controller is in the form of fuzzy set, it has to be transformed from linguistic form into a number that can be used to control the system. The following rules have been used,

Table 1: fuzzy rules

$E_{available}$ \ Hour	B	M	S
B	BB	B	MB
M	MB	M	SB
S	SB	S	S
VS	S	S	S

Many defuzzification methods like weighted average (wt_{aver}) or weighted summation (wt_{sum}) methods have been proposed. In this paper, we utilized the wt_{aver} method.

6. Results and Discussion

A simulated environment based on the forecasted data was built to prove the validity of the proposed method for estimating the amount of power to be supplied by the battery array each day. A dynamic operation of power flow in a one day period is shown in Fig. 9. The amount of power supplied by the battery array is controlled every day of the year and it is a function of the PV power forecasted data and the wind power forecasted data. The amount of power generated by wind and PV are added and subtracted from the load demand at every time instant. There are four possibilities that can be encountered:

Load demand is higher than the available renewable energy $P_{load} \geq P_{PV}+P_{Wind}$ not during peak period

Load demand is higher than available renewable energy $P_{load} \geq P_{PV}+P_{Wind}$ during peak hour

Renewable energy is higher than load demand during peak hour $P_{PV}+P_{Wind} \geq P_{load}$

Renewable energy is higher than load demand not during peak hour $P_{PV}+P_{Wind} \geq P_{load}$

Fig. 9 is divided into 6 sections to illustrate the operation of the battery at the different 4 possibilities stated above. Section 1, section 3 and section 5 of the figure represent the same state where there is a deficiency in renewable power not during the peak hour. Hence, fuzzy is used to determine the percentage of power the battery will share with the grid in order to minimize the power taken from the grid and at the same time make sure the battery has enough energy for the coming peak period. To illustrate the use of forecasting, the SOC in section 3 and section 5 have a smaller slope than in section 1, since it is near the peak hour, the battery will discharge slower. Section 2 represents the state where renewable sources are satisfying loads and having excess energy that can either be sold to the grid or used to charge the battery. In this case, since this period does not fall in the peak hour, it would be more beneficial to utilize the excess energy from renewable to charge the battery array. When the battery is charged to 100%, the extra energy available is sold to the grid. During peak hour the battery has to satisfy a big portion of the load or sell its energy to the grid. Section 6 represents the peak hour. It can be observed from the SOC of the battery in that section, the slope is large and the battery is used as the main source to satisfy the loads during this period. During peak hour the battery SOC is reduced from 97% to 64%. Therefore, the battery was successful in satisfying a big

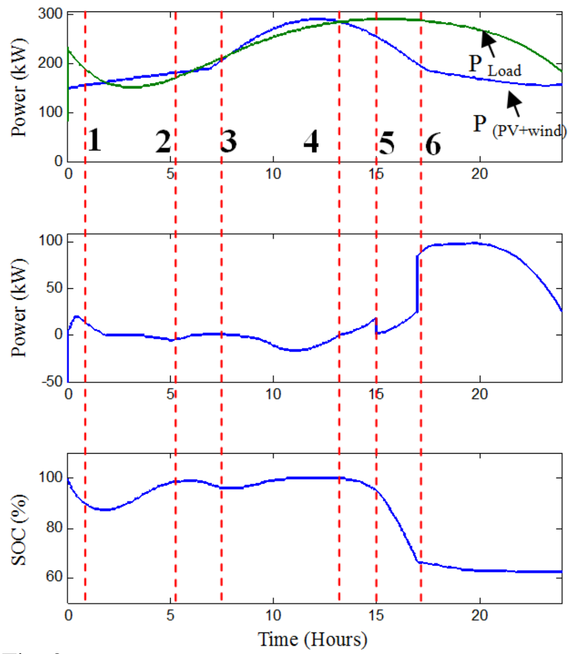


Fig. 9. Dynamic operation of power flow in a one day period

portion of the load during the peak hour.

7. CONCLUSIONS

A unit commitment scheme for an AC distribution system involving sustainable energy sources has been proposed. The commitment scheme aims at minimizing the cost of the power served to the loads and depending mainly on renewable energy sources. In case of deficiency in power, a fuzzy system has been proposed to control the sharing of the power between the grid and the battery system. A prototype system has been simulated to validate the applicability of the proposed scheme. Results show accurate performance of the commitment scheme proposed.

REFERENCES

Atwa, Y.M.; El-Saadany, E.F.; Salama, M.M.A.; Seethapathy, R. (2010). Optimal Renewable Resources Mix for Distribution System Energy Loss Minimization, in *IEEE Trans. On Power Systems*, 360-370

Liserre, M.; Sauter, T.; Hung, J.Y. (2010). Future Energy Systems: Integrating Renewable Energy Sources into the Smart Power Grid Through Industrial Electronics, in *IEEE Industrial Electronics Magazine,* 18-37

Smith, S.C.; Sen, P.K.; Kroposki, B.; Malmedal, K. (2010). Renewable energy and energy storage systems in rural electrical power systems: Issues, challenges and application guidelines, in Rural Electric Power Conference (REPC)

Wahab, M, Hamada, M and Mohamed, A. "Artificial Neural Network and Non-linear Models for Prediction of Transformer Oil Residual Operating Time," in *Electr. Power Sys.* Res., doi: 10.1016/j.epsr.2010.08.010

Appendix A. NON-LINEAR REGRESSION MODELLING

The set of non-linear functions selected for transforming input vectors contains $x_i^{\mu_1}, x_i^{\mu_2}, x_i^{\mu_3}, \frac{\beta}{x_i}, \ln(x_i)$. These functions are useful to be used for transforming input vectors to achieve linearity between them and the output vector. The constants $\mu_1, \mu_2, \mu_3, \beta$ are determined so as to maximize the correlation index r given by (A-1), which represents the linear relationship between any two vectors,

$$r = \frac{\Sigma\{(T-T_{avg})\times(U-U_{avg})\}}{\Sigma(T-T_{avg})^2\times\Sigma(U-U_{avg})^2} \tag{A-1}$$

Where,

T, U are the two vectors representing $f_j(x_i), f_j(x_i)$ for $j \neq j, i \neq i$, simultaneously, and

T_{avg}, U_{avg} are the average values of *T, U*.

The trial and error method was used for such determination; the constants were set to a small value (0.01) and varied in a wide range until the maximum correlation between the resulting transformed input vector and the output is obtained. Values of the constants $\mu_1, \mu_2, \mu_3, \beta$ are 0.4, 0.8, 0.7 and 100, respectively. Hence, The set of non-linear functions selected to model each input vector were $\{x_i^{0.4}, x_i^{0.8}, x_i^{0.7}, {}^{100}/_{x_i}, ln(x_i)\}$.

Therefore, the non-linear model for each variable (y_i) for modelling the power is as given by (A-2):

$$y_i = a_{i1}b_{i1}\,x_i^{0.4} + a_{i2}b_{i2}\,x_i^{0.8} + a_{i3}b_{i3}\,x_i^{0.7} + a_{i4}b_{i4}\,{}^{100}/_{x_i} + a_{i5}b_{i5}\ln(x_i) \tag{A-2}$$

Determination of constants $\{a_{i1}, a_{i2}, \dots, a_{ik}\}$

The constants $\{a_{i1}, a_{i2}, \dots, a_{ik}\}$ have just two possible values 0 or 1, so they control the presence of the transformed vectors in the model. These constants are determined through correlation analysis through two subsequent steps,

Firstly, the correlation index r between $f_j(x_i), y$ is determined by (A-1), where y is the real output values (experimental data which are required to be evaluated by the model). a_l will equal 1 if the absolute value of the correlation index is greater than pre-specified index r_1 (r_1 has been taken as 0.5 for the non-linear model) and will equal 0 if the correlation index is less than r_1.

Through this step only the functions having substantial effect on the output y is retained for further processing.

The correlation index r between T and U is determined. If the absolute value of the cross correlation is smaller than a pre-specified value r_2 (r_2 has been taken as 0.8 for the non-linear model) both terms are retained otherwise only the term with the greater correlation with respect to y is retained and the other is eliminated to avoid information overlap.

Up to this step, the non-linear basic matrix X is updated to a number of d_1 basic columns. The reduced matrix is given the designation X_r and any linear combination of any of the vectors included in the reduced matrix forms a basis for the non-linear model.

Determination of constants $\{b_{i1}, b_{i2}, \dots, b_{ik}\}$

If X_{r_1} is one of the combinations of the reduced basic matrix X_r having dimension d_2 so the non-linear multivariable model has the following form:

$$Y = b^T X_{r_1} = b_0 + y_1 + y_1 + \dots + y_{d_2} \tag{A-3}$$

The unknown vector of constant coefficients is given by:

$$\vec{b} = (b_0, b_1, b_2, \dots, b_{d_2})^T \tag{A-4}$$

$$\vec{b} = ({X_{r_1}}^T Q X_{r_1})^{-1}({X_{r_1}}^T Q)\vec{y} \tag{A-5}$$

Where $\vec{y} = (y_1, y_2, \dots, y_n)^T$ is the vector of real outputs and Q is the following weight matrix:

$$Q = \begin{pmatrix} w^{n-1} & 0 & . & 0 & 0 \\ 0 & w^{n-2} & . & . & 0 \\ . & . & . & . & . \\ 0 & . & . & . & . \\ 0 & 0 & . & . & w^0 \end{pmatrix} \tag{A-6}$$

and w is a weighting factor. w has been varied from 0.8 to 1 in step of 0.01 for the non-linear model.

The final model has the lowest value of *MAPE.*