A System Identification Approach for Building Forecast Models of Green Energy Systems

ADRIAN DĂNILĂ, LUMINIŢA POPA
Department of Electrical Engineering and Computer Science, Chair of Automatics
Transylvania University of Brașov
Brașov, M. Viteazu Street, no. 5, 500185
ROMANIA
adrian.danila@unitbv.ro, mluminita2001@yahoo.com, http://www.unitbv.ro

Abstract: - The paper focuses on the forecasting models of green energy power systems and loads. Since electric energy cannot be stored efficiently in large quantities, accurate tracking of the energy generation in conjunction with the consumer requirements can lead to the optimization of energy costs in the context of the liberalization of the electric energy market and the increased usage of the alternative energy resources. System identification approach can provide useful techniques to predict both green energy generation as well as hourly or daily loads such as the appropriate decision regarding the energy balance within the system are taken.

Key-Words: - system identification, forecasting model, wind turbine, building.

1 Introduction
System identification comprises activities whose purpose is to produce a mathematical model of systems based on excitation and response measurements. Such models are useful for analyzing system properties, performing simulation, prediction, filtering, state estimation, monitoring and fault diagnosis, as well as control. System identification is characterized by strongly interdisciplinary nature that draws from systems theory, signal processing, optimization, and statistics. System identification comprises a collection of activities that include experiment design, model structure selection, parameter estimation, and model validation.

A promising direction in system identification is the short-term load forecasting i.e. the prediction of power-system loads over an interval ranging from less than one hour to one week. Since electricity cannot be stored efficiently in large quantities, accurate tracking of the load by the system generation at all times is a basic requirement in the operation of power systems. Load forecasts are increasingly important not only from the production side but also from a financial perspective due to the liberalization of the electric energy markets worldwide.

From the resources perspective, classical fuels commonly used to produce electricity such as coal and petrol are going to be less attractive in the near future because of their ecological impact. Instead, the green energy such as the wind energy, the solar energy, and so are becoming more and more attractive in producing electrical energy. Because the green energy production is closely related to the weather conditions system identification is especially useful to provide forecasts both for the load side and for the generator side.

Studies available in the literature [1], [2] have used long time series containing hourly load values taken from a given consumer or from multiple substations to build a local pattern over a given time period. This pattern is then used to build a model for load forecasting. In [1] various system identification techniques for modeling and forecasting are discussed and compared such as the non-parametric models with seasonal components, the Box-Jenkins seasonal ARMA models or the application of the seasonally varying parameters in an auto regression. Regarding the weather forecast, a large scale model has been developed based on the Global Numerical Weather Prediction with a Climate Model. Another approach, presented in [3] is based on the Monte Carlo techniques for data assimilation and the ensemble Kalman Filter for combined state and parameter estimation.

In this work, the authors have studied the dynamical models of both a green energy supplier and a factory that uses the green energy to provide the proper working conditions within its building. The green energy supplier has a wind turbine farm and an extended photovoltaic array with the proper energy converter to adapt voltage level to the industrial grid. To fulfill this analysis the authors have used the dedicated software environment TRNSYS and...
the MatLAB System Identification toolbox. The analysis below can provide useful information both to the optimal design and operation of the green energy supplier. It is also important from the consumer side providing information on the seasonal profile of the energy necessities.

2 System Identification Approach for Building Forecasting Models

In System Identification, we start from large amounts of measured data both from the input to the system and from its output. Subsequently, a model structure has to be adopted. Various structured elements can be incorporated into the model formulation, including prior knowledge or information about the problem to hand. Model's parameters can then be estimated based on the loss function's minimization. Finally, the results must be verified to decide if the selected model is a valid representation of the underlying system.

2.1 Building a Forecasting Model of a Wind Turbine Farm

Wind turbines extract kinetic energy from the wind and convert it into electrical energy, which can be then fed directly into a utility grid. Wind turbines can be either variable or fixed speed. Fixed speed turbines are usually configured to connect directly to the utility grid, so the electricity they produce must have the same frequency as the grid. Variable speed turbines use power-processing equipment to convert variable-frequency electricity to the fixed electrical grid frequency. Energy is transferred from the rotating blades, or rotor, to the generator through a drive train that may consist of a single shaft or multiple shafts connected through one or more gear boxes. Modern wind turbines use variable speed structure to maximize the aerodynamic efficiency over a wide range of wind speeds. Additionally, actuators for the yaw drive, the blade-pitch drives and for the electrical generator are commonly used. This complex structure of the turbine leads to a complex representation of the dynamic model of the turbine. Even so, the electric power transferred to the grid over a given time period depends on the local wind speed profile. Therefore the local weather profile has to be taken into account when building the dynamic model of the system.

In this approach [4], a set of explanatory variables such as: (1) an autoregressive set of lagged values of the wind speed and temperature-related variables (2) day and hour of the day information are included into the regression vector of the model

\[ p(k) = f(\phi(k)) + e(k) \]  (1)

where \( f \) is an unknown function, \( p(k) \) is the electric power provided at the hour \( k \), \( e(k) \) is a white noise process, and \( \phi(k) \in \mathbb{R}^n \) is the regression vector

\[ \phi^T(k) = [-y(k-1), \ldots, -y(k-n_a), u(k), \ldots, u(k-n_b)]. \]  (2)

The prediction at time \( k \) is given by

\[ \hat{y}_k = y_k - e_k. \]  (3)

The unknown function \( f \) is identified as a function of a parameter vector \( \theta \):

\[ \theta = [a_1 \ldots a_{n_a}, b_1 \ldots b_{n_b}]. \]  (4)
An estimate $\hat{\theta}$ is obtained as the solution of an optimization problem. If $f$ is parameterized as a linear function
\[
f[f(k)] = \theta^T \cdot \phi(k) + b,
\]
then the unknown function $f$ is an ARX model, and the estimates are obtained as the solution of a least-square optimization problem. Linear models assume a linear effect from the input variable to the output variable. When the correlation between the input and output of the underlying system is not linear then the function $f$ is parameterized as a nonlinear function. In this case a NARX model can be taken into consideration. Furthermore the NARX model can be extended to the case where the residuals $e_k$ follow an autoregressive process of a given order leading to an AR-NARX structure [1].

In this work the forecasting model of the wind turbine is build based on the least squares method and a linear ARX structure, the least square method and a nonlinear ARMA structure, the prediction error methods and a generalized parameter model. The least square estimate minimizes the sum of the squared equation errors and is given by the expression
\[
\hat{\theta} = \left[ \frac{1}{N} \sum_{k=1}^{N} \phi(k) \cdot \phi^T(k) \right]^{-1} \left[ \frac{1}{N} \sum_{k=1}^{N} \phi(k) \cdot y(k) \right].
\]

The estimate is consistent if $E\left[\phi(k) \cdot \phi^T(k)\right]$ is nonsingular and $E\left[\phi(k) \cdot v(k)\right] = 0$ i.e. the disturbance $v(k)$ to the underlying process must be uncorrelated ($E[\bullet]$ is the expectation operator). Within the identification to track the dynamics of the system the least squares loss function to be minimized has been modified according to
\[
V_k(\theta) = \sum_{s=1}^{k} \lambda^{k-s} \cdot e^2(s),
\]
where $\lambda$ is the forgetting factor. The prediction error methods can be applied to the general linear parametric model
\[
y(k) = G^{-1} \cdot \theta \cdot u(k) + H^{-1} \cdot \theta \cdot e(k),
\]
The model's parameter vector $\theta$ is determined such that the prediction error $e(k,\theta) = y(k) - \hat{y}(k|k-1,\theta)$ is small. Within the relation (10) the expression $\hat{y}(k|k-1,\theta)$ denotes the mean-square optimal prediction of $y(k)$.

### 2.2 The Model of the Load

The load of the green energy provider within this analysis is a factory building. The electrical energy provided by the supplier is used to provide optimal working conditions within the building. The model of the building has to take into account the specific parameters of the construction such as the geometry, the materials of the walls, the windows orientation with respect to the solar incident radiation, the local weather conditions and so.

In purpose to accomplish this task, the TRNSYS simulation environment has been used. The specific weather local profile was simulated based on the Meteонorm data record. The temperature inside the building was estimated for a given time interval. These data were then used into a MatLAB application to build a dynamic model for the building.

### 2.3 Model Validation

Model validation provides several techniques to decide which the model is that picks out the best within the chosen structure. The following validation techniques commonly used are: the comparison between the simulated and the measured output, the usage of validation criteria such as the AIC or the FPE test criteria, the pole-zero cancellation method and the residual analysis method. Within this work the residual analysis method was used to perform the validation of the model.

The residual analysis consists in assessing the quality of the model by checking the properties of the residuals regression. If the regression is correctly specified, all correlations, in time and across input variables, are captured by the model. The residuals of a correctly specified model are expected to show a small degree of autocorrelation, as well as a little correlation with any of the input series. Significant autocorrelation suggests that there is a part of the response components that originates from the past input that has not been properly picked up by the model.
3 Implementation and Results

3.1 The Dynamic Model of the Wind Turbine

The mathematical model of the wind energy conversion system consists of a model for the wind turbine that calculates the power output based on a power versus wind speed characteristic provided on a look-up table. The impact of air density changes and wind speed increases with height is also modeled. In addition to the wind turbine, a model for the power conditioning unit is also implemented. The model is based on empirical efficiency curves for electrical inverters AC/DC and DC/AC. The wind turbine model parameters are given in Table 1.

<table>
<thead>
<tr>
<th>Denomination</th>
<th>Units</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of turbines</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>Hub height</td>
<td>m</td>
<td>46</td>
</tr>
<tr>
<td>Maximum Power per Unit</td>
<td>kW</td>
<td>15</td>
</tr>
<tr>
<td>Barometric Pressure</td>
<td>Pa</td>
<td>101325</td>
</tr>
<tr>
<td>Rated voltage</td>
<td>V</td>
<td>230</td>
</tr>
</tbody>
</table>

The implementation of an ARX model to the input/output measured data is presented in Figure 3. The residuals analysis of the results is depicted in figure 4. The correlations between the residuals of the output of the system denote a maximum of 0.23, which denotes low correlation. However, the cross correlation between the input and the output denotes the existence within the model of an unmatched influence of the lagged values of the regressors that provides insight for further refinement of the model. The implementation of the ARMA model is presented in Figure 5. The residuals analysis is presented in Figure 6. As seen from the plot, the response of the model lags the plot of the measurements. Also the model doesn’t track the measurements especially its high rated modifications. These results are sustained by the residuals analysis presented in Figure 6. The correlations between the regressors of the output are less that 0.18, which means the ARMA model matches the system better than the ARX

Figure 3. Comparison between the measured output power of the wind turbine – thick line and the estimated output power, ARX model – thin line.

Figure 4. Residuals analysis for the ARX model of the wind turbine.

Figure 5. Comparison between the measured output power of the wind turbine – thick line and the estimated output power, ARMA model – thin line.
model. However, the cross correlations between the input and the output denotes that there still is an influence of the lagged regressors that haven’t been picked-up within the model, thus the model could be improved.

To take into account the influence of the lagged regressors a forgetting factor has to be inserted into the model. In Figure 7 the response of a recursive RARX model with a forgetting factor of 0.985 is presented.

As seen from the plot the RARX model tracks the measurements even at high rated modifications.

3.2 The Dynamic Model of the Factory Building

The dynamic model of the factory building is based on the mathematical model provided by the TRNSYS software environment. The model comprises 3 thermal zones, based on the physical characteristics of the walls and windows, and takes into account the influence of the location and weather conditions. The output of the model is the requested thermal power to provide proper conditions within the building.

The geometry of the building has been set-up within a dedicated file.

The building seasonal requested power is presented in Table 2.

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Time [h]</th>
<th>Q-tot [W]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>400</td>
<td>3446</td>
</tr>
<tr>
<td>2</td>
<td>800</td>
<td>3776</td>
</tr>
<tr>
<td>3</td>
<td>1200</td>
<td>2694</td>
</tr>
</tbody>
</table>

Table 2

![Figure 6](image1.png)

Figure 6. Residuals analysis for the ARMA model of the wind turbine.

![Figure 7](image2.png)

Figure 7. Comparison between the measured output and the response of the RARX model of the wind turbine – above. The per-unit error of the model - below.
The summarized seasonal data shows that, for the conditions under study, within the summer period there is no need of thermal energy for the building. This is important to know if an accurate prediction for the power consumption is requested. The magnitude of the requested power during the cold days is about 3.5 kW, instead, the wind turbine can provide an average of about 1.5 kW – for the weather conditions under study. Therefore there is about 2.0 kW to obtain from the grid. It is also interesting to observe that there are situations when the amount of electric power provided by the wind turbine can exceed 45 kW. This is much more than the local necessities and can be directed to the grid to partially compensate the requested electric power. Based on these data, an RARX dynamic model was built for the factory building under study. The response and the per-unit errors of the model are presented in Figure 8. As seen from the plot, the RARX model tracks very well the dynamics of the process. The mean of the per-unit errors is 0.0198 and the variance is 0.0355.

4 Conclusions

This work illustrates the application of the systems identification techniques to the problem of forecasting the green energy provided by a wind turbine in conjunction with a local consumer i.e. a factory building to supply proper climate conditions inside the building. The study above proves that the dynamics of complex systems can be tracked via the system identification modeling. Through the validation techniques the authors proved that the model structure under study that tracks the dynamics of the system is the recursive RARX model both for the energy supplier and for the load. Further research on a more dedicated structure model, a much more detailed underlying system model and accurate input data might lead to improvements and extensions of these results.

References: