Online learning and forecasting in the prediction horizon

Principle of OL method

In this subsection, the forecasting method used in the MPC is presented. The forecasting methods and further explanations regarding the method used in this work can be found in [1, 2]. However, a brief introduction and overview of the method is given in this section. Firstly, the introduction to regression is given. Secondly, the reason behind using the regression and online learning (OL) method is given. Lastly, the procedure of OL method is presented in detail.

The linear regression method (LRM) is used to forecast by creating the relationship between the response and explanatory variables for any given time. As an example, explanatory variable such as wind speed helps to predict the response variable such as wind power. By considering continues variables, such as wind speed and wind power, and a set of observations in the past of response and explanatory variables, the relationship between wind speed and wind power can be written through LRM. LRM introduces regression parameters which values can be found by using least square (LS) estimation. By creating relationship between those two sets of variables and obtaining the regression parameters, the previous time steps are ignored, and linear relationship can be found for each time instant between the wind speed and wind power. Therefore, obtained relationship can be used to forecast in future. However, linear regression method used to forecast the future leads to three concern that should be taken into account. Firstly, in order to forecast in the future, the wind speed for the future is acquired. Thus, wind speed must be forecasted in order to further obtain future forecasted wind power values. By using a forecasting values to forecast the future, the forecasting error is increased. Secondly, uncertain parameters can vary with time. For example, seasonal effects can be seen in wind power. Seasonal effects can be taken into account through the estimation of model parameters on the sliding window. In other words, the LS estimation is performed for a season or a specified window size. For each time step, the regression parameters are recalculated. That leads to the third concern which is increased computational burden.

Two linear regression models can be differentiated. First model considers simple mapping of explanatory variable to response variable or in other words mapping of wind speed to wind power as explained above. The second model is autoregressive model. Autoregressive model is called OL method, as the OL method is based on recursivity, i.e. the new update is based on the relationship between the last and the new data point [2]. Due to highlighted concerns, like computational burden and forecast error, the OL method is implemented in this study based to the research done in [2]. By using OL method, the computational burden is not increasing and the explanatory variable such as wind speed forecasts are not needed as it will be shown below.

The model of OL method is as follows. Autoregressive model is similar to simple mapping model, with exception that explanatory variable is a response variable in the previous or present time. The autoregressive model of order p is shown in Eq. (1) and compact formulation is given in Eq. (2). β are the model parameters and ε is the noise accounting for deviation between the observed measurement and modelled relationship. y_t is the response variable. For example, y_t can be wind power output at time *t*. $\boldsymbol{x}_t = [1\ y_{t-1} \ldots y_{t-p}]^T$ is the explanatory variable; $\boldsymbol{\beta}_t = [\beta_{t,0}\ \ \beta_{t,1} \ldots\ \ \beta_{t,p}]^T$ is the model parameter. Hence, the recusivity of the autoregressive model means that the model parameters, β_t, for p=1 are estimated at time *t-1* and the data before time step *t-1* can be deleted. The recursivity of the autoregressive model is controlled by the order of the model.

$$y_t = \beta_{t,0} + \beta_{t,1} y_{t-1} + \beta_{t,2} y_{t-2} + \ldots + \beta_{t,p} y_{t-p} + \varepsilon_t \quad ,\forall t \tag{1}$$

$$y_t = \boldsymbol{\beta}_t^{\mathrm{T}} \mathbf{x}_t + \varepsilon_t \quad ,\forall t \tag{2}$$

To estimate model parameters at time *t*, β_t, of the autoregressive model, the regressive least square (RLS) method is applied as presented in Eq. (3). λ is the forgetting factor with a value less than 1 and close to 1. Usually, λ ranges between 0.95 to 0.99 [2]. An update for $\hat{\beta}_t$ can be recursively obtained through the Newton-Rapson method as shown in Eq. (4). Following Eq. (3), derivatives are acquired to obtain the update scheme for RLS. There will be four steps to follow. Firstly, assuming that the current step is t-1 and the future step is t, $S_t(\hat{\beta}_{t-1})$ is rewritten in Eq. (5). Secondly, the first derivative is applied on both sides of Eq. (5) resulting in Eq. (6). The second derivative on Eq. (6), results in Eq. (7). In the third step, in order to simplify the notation, Eqs. (8)-(9) are introduced by denoting similar parts from Eqs. (6)-(7) into R_t and R_{t-1}. Once the mentioned simplified notation is available, the left and right hand side of Eq. (7) can be replaced by R_t and R_{t-1}, and the recursive equation for R_t as shown in Eq. (10) is obtained. The forth steps is to obtain the updating equations for R and $\hat{\beta}$.

The update equation for R is shown in Eq. (10) and the update for $\hat{\beta}$ can be obtained in similar manner. Recursive $\hat{\beta}_t$ can be obtained by replacing denominator in Eq. (4) with R_t from Eq. (8). The update $\hat{\beta}_t$ is shown in Eq. (11). The current error at time t should not be ignored and, hence, the forgetting factor is set to one. This leads to the update for β_t as shown in Eq. (12).

$$\hat{\beta}_t = \arg\min_{\beta} S_t\beta_t = \arg\min_{\beta} \frac{1}{2}\sum_{i<t} \lambda^{t-i}(y_i - \mathbf{\beta}_t^T \mathbf{x}_i)^2 \tag{3}$$

$$\hat{\beta}_t = \hat{\beta}_{t-1} - \frac{\nabla S_t(\hat{\beta}_{t-1})}{\nabla^2 S_t(\hat{\beta}_{t-1})} \tag{4}$$

$$S_t(\hat{\beta}_{t-1}) = \frac{1}{2}\lambda \sum_{i<t-1} \lambda^{t-i-1}(y_i - \hat{\beta}_{t-2}^T x_i)^2 + \frac{1}{2}(y_t - \hat{\beta}_{t-1}^T x_t)^2 = \lambda S_{t-1}(\hat{\beta}_{t-2}) + \frac{1}{2}(y_t - \hat{\beta}_{t-1}^T x_t)^2 \tag{5}$$

$$\nabla S_t(\hat{\beta}_{t-1}) = \lambda \nabla S_{t-1}(\hat{\beta}_{t-2}) - x_t(y_t - \hat{\beta}_{t-1}^T x_t) \tag{6}$$

$$\nabla^2 S_t(\hat{\beta}_{t-1}) = \lambda \nabla^2 S_{t-1}(\hat{\beta}_{t-2}) + x_t x_t^T) \tag{7}$$

$$R_t = \nabla^2 S_t(\hat{\beta}_{t-1}) \tag{8}$$

$$R_{t-1} = \nabla^2 S_{t-1}(\hat{\beta}_{t-2}) \tag{9}$$

$$R_t = \lambda R_{t-1} + x_t x_t^T \tag{10}$$

$$\hat{\beta}_t = \hat{\beta}_{t-1} - \frac{\nabla S_t(\hat{\beta}_{t-1})}{\nabla^2 S_t(\hat{\beta}_{t-1})} = \hat{\beta}_{t-1} - R_t^{-1}\nabla S_t(\hat{\beta}_{t-1}) = \hat{\beta}_{t-1} - R_t^{-1}\sum_{i<t}\lambda^{t-i}(-x_i)(y_i - \hat{\beta}_{t-1}^T x_i) \tag{11}$$

$$\hat{\beta}_t = \hat{\beta}_{t-1} + R_t^{-1} x_t (y_t - \hat{\beta}_{t-2}^T x_t) \tag{12}$$

However, one final step of deciding the initial parameters remains and it is explained later on. The scheme to recursively estimate the parameters $\hat{\beta}_t$ by applying RLS is presented in Fig. 1. Firstly, historical data is obtained. Secondly, order p for Eq. (2) is decided. In most cases, order p equals to two is enough [2]. In this study, updated forecasts are obtained for a period of 24 hours. The time resolution of the data is 15 min and the forecasted data is updated every 15 minutes. Thirdly, the forgetting factor, λ, equals to 0.99 and $R_{t-1}$and $\beta_{t-1}$are set to zeros initially.

In step 1, the first two values of time series in the historical data are left outside as the order p equals two and those first two values are used as input in the first loop in step 1. After step 1 is completed and initial parameters are obtained, the learning step is no longer needed. Step 2 can be applied providing the future forecasted values.

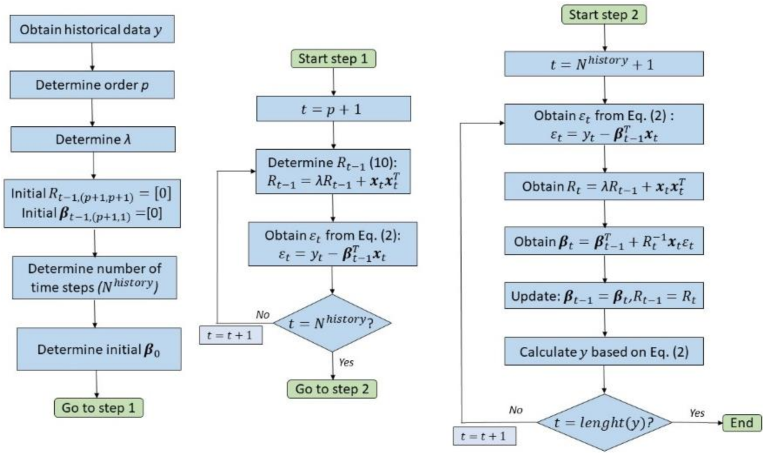

Fig. 1. Scheme for updating of the parameters and obtaining forecast values

Application of OL method

There are two horizons to be considered, i.e. scheduling and prediction horizon. An overview of the OL method is summarized in Fig. 2 when applied to the MPC based RT scheduling. The learning process is related to step 1 in Fig. 1. The learning process takes 96 time steps for wind power. Once the learning process is completed, step 2 from Fig. 1 is initiated. Moreover, since the scheduling horizon is 24 hours, 97 time steps are used in the historical data, whereas 1 day before the selected day and the first time step of the selected day in order to forecast one step ahead. The OL method can predict the information of the next step by obtaining the actual value of the uncertain parameter in the current step and the previous step as seen in the third step of the scheme in Fig. 2. Moreover, the OL method forecasts one step ahead based on the last two updated actual values of the uncertain parameter. As the OL method forecasts one step ahead, the early part of the prediction horizon of RT scheduling is two time steps equal to 30 minutes.

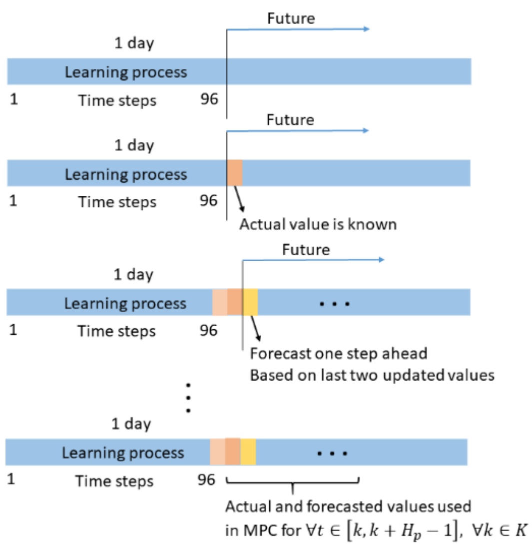

Fig. 2. Process of OL method in MPC based RT scheduling

Assume that a prediction horizon has H_p time steps in the MPC based RT scheduling. At the first time step of the prediction horizon, the actual measured values of the uncertainty sources are utilized, including wind power, electric demand, gas demand and heat demand. At the second time step, the forecast values of each uncertainty source obtained by OL method are utilized. For the rest H_p-2 time steps, the predicted values are generated based on forecast error distribution. It is assumed that the forecast errors of each uncertainty source follow the normal distribution with zero mean and a certain standard deviation for the rest H_p-2 time steps. The standard deviation of each uncertainty source is calculated based on the average standard deviation of the previous day. Then the forecasted random errors sampled from the normal distribution are added to the predicted value obtained by OL method.

Initialization of the parameters, validation of online learning method and forecasting parameters

For the purposes of further explaining the OL method, wind power data is used. As mentioned earlier, final step of deciding the initial parameters, $\boldsymbol{\beta_0}$ and R_{t-1}, remains, and it is thoroughly explained below. Selecting reasonable parameters for the forecasting model plays an important role in the OL method.

Initialization of the parameters is based on the number of historical values in the time series and the amount of data for calculating initial $R_{t-1}$and β_{t-1} must be defined. That is specified in the initial step where the number of time steps should be determined in Fig. 1. In order to obtain the new initial values for $R_{t-1}$and β_{t-1} based on historical data, entire scheme in Fig. 1 must be processed. Meaning, worthy values $R_{t-1}$and β_{t-1} are found after completing step 1 and step 2 in Fig. 1.To obtain β_{t-1}, i.e. $\boldsymbol{\beta_0}$ in $N^{history}$ is defined randomly and the initial $\boldsymbol{\beta_0}$ is based on saturation of all values. Parameter $\boldsymbol{\beta_0}$ can be defined by applying 5 historical values and forecasting the rest of the values in the historical data. That would mean $N^{history}$ equals 5 only for the purpose of obtaining initial $\boldsymbol{\beta_0}$. Meaning, the step 1 and step 2 can be performed and once the entire process is finished and all the steps have been used, the initial $\boldsymbol{\beta_0}$ can be found based on saturation of the all values. The parameter $\boldsymbol{\beta_0}$ is shown in Fig. 3. It can be seen that β_0, β_1 and β_2 can be initialized at 0, 1.75 and -0.75 respectively.

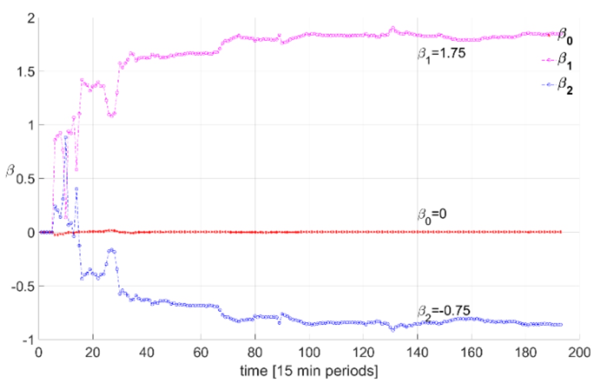

Fig. 3. Results for estimation of initial β parameter based on historical values of wind power output

Once $\boldsymbol{\beta_0}$ is obtained, the initialization for R_{t-1} is performed. In order to obtain R_{t-1}, the root-mean-square-error (RMSE) is calculated between the predicted and measured values in historical time series for a different values of $N^{history}$. The process in Fig. 1 is completed for each value of $N^{history}$. The RMSE results are shown in Fig. 4. It is shown that RMSE values decrease to zero after a number of time steps used. The smallest RMSE decides the $N^{history}$ value and consequently the initial R_{t-1}. For that reason, one day before the chosen day is used to obtain the initial parameters. The final value of $N^{history}$, also called learning process time, equals 96 time steps for the wind power output.

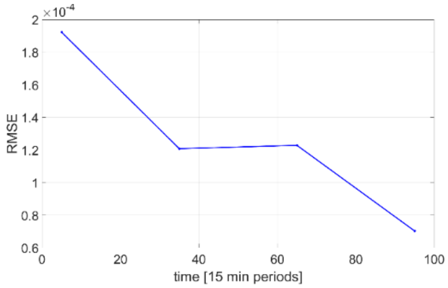

Fig. 4. RMSE between updated forecasted and measured actual wind power output

The parameters of each uncertainty source differs, as shown in Table 1. In Table 1, the electric demand and wind power data with a time resolution of 15 minutes are available online. Due to the high variability of wind power, the standard deviation of forecast error for wind power is higher than that of the electric demand. The gas and heat demand data can be found for hourly values. Hence, the predictions are based on hourly time resolution. Additionally, the actual gas and heat demand have a small deviation from the forecasted values, leading to a small standard deviation. To conform with the time resolution of 15 minutes, the gas and heat demand data are extended to 96 time steps, with four equal values in each hour.

Table 1. Parameters for uncertainties used in OL method and forecasting

Uncertainty	Learning process duration	Time resolution	β_0, β_1 and β_2			Prediction	Standard deviation
Wind power	1 day	15 min	0	1.75	-0.75	15 min ahead	0.0027
Electric demand	1 day	15 min	0	1.75	-0.75	15 min ahead	0.0012
Gas demand	2 days	1 hour	0	1.5	-0.5	1 h ahead	0.00057961
Heat demand	2 days	1 hour	0	1.5	-0.5	1 h ahead	0.00057961

References

[1] Juan M. Morales, Antonio J. Conejo, Henrik Madsen, Pierre Pinson, Marco Zugno, Integrating Renewables in Electricity Markets, New York: Springer, 2014.

[2] DTU CEE Summer school, "Data-Driven Analytics and Optimization for Energy Systems," 2019. [Online]. Available: https://energy-markets-school.dk/summer-school-2019/. [Accessed 02 03 2020].