## B.Selection of input variables

- Input variables have to be selected keeping these 3 aspects in mind:
  - Which exogenous Variables have significant correlation to wind generation
  - Which station's data is suitable for our forecasting model
  - How many lagged hours to be included in the wind generation time series
- Table 1 shows that wind generation has positive cross correlation to 2 exogenous variables namely wind direction and gust speed.
- Table 2 shows the cross correlation between different stations and wind generation.We set a benchmark of 0.60 and select stations which have a correlation (>0.60) i.e. Stations 2 , 3 , 5.

TABLE I
CROSS-CORRELATIONS BETWEEN AGGREGATED WIND GENERATION AND DIFFERENT OBSERVATIONS WITHIN THE WIND FARM

| | Wind direction | Barometric pressure | Temperature | Humidity | Gust speed |
|---|---|---|---|---|---|
| Aggregated WG | 0.14 | -0.02 | -0.02 | -0.07 | 0.62 |

TABLE II
CROSS-CORRELATIONS BETWEEN THE WIND SPEEDS IN THE WEATHER STATIONS AND THE AGGREGATED GENERATION OF THE WIND FARM

| Stations | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Cross-correlation With generation | 0.56 | 0.66 | 0.67 | 0.54 | 0.62 | 0.50 | 0.51 |

TABLE IV
LIST OF INPUT DATA OF THE BCD CLASSIFIER

| Input | Variable | Detail description |
|---|---|---|
| 1-6 | $G_I$ | Wind power generation series |
| 7-30 | $S_I$ | Wind speed series |
| 31-40 | Variables to determine prior probability | $G_0$: Average wind generation |
| | | $S_0$: Average wind speed |
| | | $D$: Forecasted wind direction |
| | | $H$: Average humidity |

- Table 4 is used for deciding the input variables for the BCD classifier.
- First 6 input variables - Wind power generation time series.
- Next 24 input variables - Wind speed time series.
- Last 10 input variables - Avg wind generation , Avg wind speed , Forecasted wind direction and average humidity which is used to determine the prior probability of the classifier.

- Table 3 is used for deciding the input variables for the SVR network.
- This table consists of hourly wind generation series of previous 6 hours as well forecasted and observed wind speeds of wind farm and weather stations 2 ,3 5 that were chosen.
- Here lag 1-6 indicates the observed wind generation for the previous 6 hours
- Similarly for wind speeds lag 0 indicated the forecasted wind speed at and lag 1-5 are for the observed wind speeds for previous 5 hrs

TABLE III
LIST OF INPUT DATA OF THE SVR NETWORK

| Input | Variable name | Lagged value (hours) |
|---|---|---|
| 1-6 | Hourly wind power generation ($G_I$) | 1,2,3,4,5,6 |
| 7-12 | Hourly wind speed 0 ($S_I$) | 0,1,2,3,4,5 |
| 13-18 | Hourly wind speed 2 | 0,1,2,3,4,5 |
| 19-24 | Hourly wind speed 3 | 0,1,2,3,4,5 |
| 25-30 | Hourly wind speed 5 | 0,1,2,3,4,5 |

The wind speed 0 means the speed within the wind farm, and accordingly number 2, 3, or 5 means weather station 2, 3, or 5.
The hour of the predication is assumed at 0, the lag 0 represents the target instant, and the 6 lagged hours means the values that were measured 6 h earlier than the hour of predication.

## C.The learning Algorithm - BCD classifier

- Method of clustering is used in the classifier using the principle of similarity of time series
- Step 1 of BCD classifier is to model the input time series using an autoregressive model. The matrix form of the model is : $x_j = X_j \beta_j + \varepsilon_j$ (eqn 4). Using given data and standard bayesian techniques , parameters are calculated.
- Step 2 of the BCD classifier is the forming of initial clusters which is done using a model based bayesian procedure.The matrix form of these clusters is $x_k = X_k \beta_k + \varepsilon_k$ (eqn 5). Here $x_k = \begin{pmatrix} x_{k1} \\ \vdots \\ x_{km_k} \end{pmatrix}, \quad X_k = \begin{pmatrix} X_{k1} \\ \vdots \\ X_{km_k} \end{pmatrix}.$

- Step 3 of the BCD Classifier is ranking these initial clusters using the posterior probability given by $P(M_C \mid x) \propto P(M_C) f(x \mid M_C)$ (eqn 6). To calculate the marginal likelihood we use $f(x \mid M_C) = \int f(x \mid \theta_C) f(\theta_C)\, d\theta_C$ (eqn 7) which is simplified to remove the integral and results in $f(x \mid M_c) = \frac{\Gamma(1)}{\Gamma(1+m)} \times \prod_{k=1}^{c} \frac{\Gamma(m_k/m + m_k)}{\Gamma(m_k/m)}$ (eqn 8).
$$\times \frac{(\mathrm{RSS}_k/2)^{(q-n_k)/2}\, \Gamma\,(n_k - q)/2}{(2\pi)^{(q-n_k)/2} \det(X_k^T X_k)^{(1/2)}}$$

- However these clusters grow exponentially and Step 4 of the BCD classifier is to reduce the no of clusters by merging them.There are two ways to do this :-
  - Agglomerative search strategy - Uses marginal likelihood to merge clusters
  - Heuristic strategy applied on agglomerative clustering - Uses a similarity measure namely euclidean distance between 2 times series $S_j = \{x_{i1}, \ldots, x_{in}\}$ and $S_j = \{x_{j1}, \ldots, x_{jn}\}$ given by $D(S_i, S_j) = \sqrt{\sum_{t=1}^{n} (x_{it} - x_{jt})^2}$ (eqn 9)