9th Asia-Pacific Conference on Wind Engineering
Auckland, New Zealand
3-7 December 2017

Gust forecasting by using numerical weather prediction and on-site measurement

A. Yamaguchi[1] and T. Ishihara[1]

[1]Department of Civil Engineering, School of Engineering,
The University of Tokyo, Tokyo 113-8656, JAPAN

Abstract

A gust forecasting based on ARX model which uses numerical weather prediction and on-site measurement as inputs was proposed and the model parameters were estimated by using non-parametric regression with forgetting factors. The prediction accuracy of the dynamically adaptive model was improved compared to the conventional static MOS (Model Output Statistics) model. It was also shown that the prediction accuracy of maximum gust improves by utilizing the numerical weather prediction with higher horizontal resolution. The predictability of the gust with the maximum wind speed larger than 15m/s was evaluated by using ROC (Receiver Operating Characteristic) curve and AUC (Area Under the Curve) and was improved by the proposed method.

Introduction

Forecast of gust event is important for the maintenance of civil structures. It is also important for the transportation infrastructures which required to be closed when strong gust event is happening in near future.

Shimamura and Matsunuma [1] developed a Karman Filter based model for the forecast of wind speed up to 15 minutes ahead based on the measured wind speed. Hopmann et al.,[2] proposed a method to perform forecast up to 2 minutes ahead by using linear trend model. However, both methods are based on only measurement data and forecast horizon is limited. For longer forecasting, the use of NWP to account for the change in weather condition is needed.

On the other hand, in the field of wind power forecasting, extensive researches have been carried out [3]. For example, Nielsen et al. [4] proposed to use an ARX model the parameters of which is estimated by using non-parametric regression. They performed wind power forecast up to 24hours ahead in Denmark and showed its applicabiliy. However, in wind power application, the interest is an averaged power and not strong gust, which means there do not exist a method to forecast the gust wind speed. In addition, the applicablity of this model to Japan, where the terrain is very complex, has not been investigated.

Thus, in this study, a strong gust forecasting model is developed by using ARX model. The method is applied to the gust wind speed forecasting to investigate the applicability of this model to the forecasting of strong wind event. The effect of numerical weather prediction data on the forecast accuracy is also investigated. The predictability of the strong gust event is investigated quantitatively by using ROC curve and AUC.

Forecast Model

The forecast system developed in this study uses numerical weather prediction data and onsite wind speed measurement data as an input, and output the maximum gust wind speed up to 24 hours ahead. The outline of the system is shown in figure 1.

This system consists of three sub-models, i.e., mean wind forecast mode, peak factor estimation model and fluctuating wind speed forecast model. Each model forecasts or estimates its target value based on the input data. In addition, the forecast accuracy is considered and the upper limit of the forecast is also estimated. In this section, the input data used in this study are described and each sub models are described.

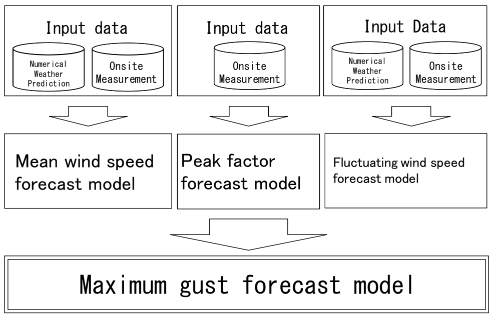

Figure 1 The outline of the proposed model

Table 1 summarises the numerical weather prediction data used in this study. Two numerical weather prediction data with different resolution.

Table 1 Summary of the Numerical Weather Predicion

Model	GPV-MSM	GPV-GSM(Japan region)
Forecast horizon	39 hours	84 hours
Available time	3 hours after the initial	3 hours after the initial time
Temporal resolution	1hour (surface), 3hours(pressure level)	
Initial time (JSR)	3:00, 9:00, 15:00, 21:00 (JST)	3:00, 9:00, 15:00, 21:00 (JST)
Variables	Surface pressure, Altitude (of isobars), horizontal wind vertical wind, temperature, humidity, precipitation, cloud	
Horizontal resolution	60	
Surface horizontal resolution	0.05 degree in NS×0.0625 degree in EW	0.2 degree in North-South×0.25 degree in East West
Domain	20.4N～47.6N, 120E～150E	20N～50N, 120E～150E

Forecast Schedule

The forecast schedule when GSM is used as an input, is shown in figure 2. The delivery of the on-line NWP data is around 3 hours after their initial time. Thus, in this study, it is assumed that NWP data can be used for the on-line forecast six hours after their initial time.

For example, when the occurrence of the strong gust event is forecasted at six o'clock in the morning, then, the initial time of the available NWP data for on-line forecast is 21 o'clock on the previous day. This NWP data contains 39 hours of forecast value, which means that forecasting can be performed until 12 o'clock of the following day.

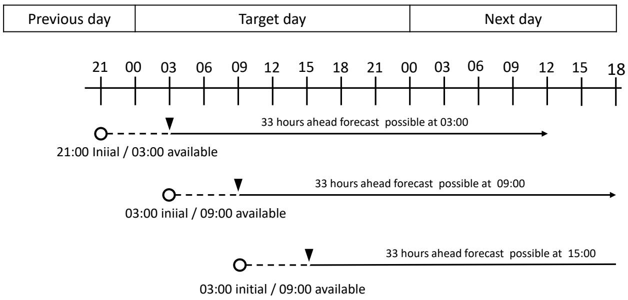

Figure 2 Forecast Schedule

Mean Wind Speed Forecast Model

Numerical weather prediction data has limited resolution and detailed elevation and surface roughness are not taken into account. When the atmosphere is neutral, the change in the wind speed and direction can be assumed to be the function of only wind direction. In his study, the local wind speed which is affected by local terrain is assumed to be a function of wind speed and direction in NWP, and the forecasted local wind speed $\left|\bar{\mathbf{u}}_{t+k|t}^{\text{local}}\right|$ forecasted at time t targeted at after k steps can be written as

$$\left|\bar{\mathbf{u}}_{t+k|t}^{\text{local}}\right| = f\left(\left|\bar{\mathbf{u}}_{t+k|t}^{\text{nwp}}\right|, \theta_{t+k|t}^{\text{nwp}}\right) \tag{1}$$

Where, $\left|\bar{\mathbf{u}}_{t+k|t}^{\text{nwp}}\right|$ and $\theta_{t+k|t}^{\text{nwp}}$ are the wind speed and direction of NWP, which can be calculated from east-west component and north-south component as following.

$$\left|\mathbf{u}_{t+k|t}^{\text{nwp}}\right| = \sqrt{\left(v_{t+k|t}^{\text{nwp}}\right)^2 + \left(u_{t+k|t}^{\text{nwp}}\right)^2} \tag{2a}$$

$$\theta_{t+k|t}^{\text{nwp}} = \text{atan2}\left(v_{t+k|t}^{\text{nwp}}, u_{t+k|t}^{\text{nwp}}\right) \tag{2b}$$

$f\left(\left|\bar{\mathbf{u}}_{t+k|t}^{\text{nwp}}\right|, \theta_{t+k|t}^{\text{nwp}}\right)$ is the smooth function to consider the effect of local terrain and estimated by using nonparametric regression, which is described later.

The wind speed obtained as explained above does not contain the bias but the phase error of the forecast, which comes from the NWP is still contained. In his study, in order to reduce the phase error mainly for short forecast horizon, the forecasted wind speed is collected as follows.

$$\begin{aligned}\left|\bar{\mathbf{u}}_{t+k|t}^{\text{pred}}\right| &= a\left(k, \theta_{t+k|t}^{\text{nwp}}\right)\left|\bar{\mathbf{u}}_{t}^{\text{meas}}\right| \\ &+ b\left(k, \theta_{t+k|t}^{\text{nwp}}\right)\left|\bar{\mathbf{u}}_{t+k|t}^{\text{local}}\right|\end{aligned} \tag{3}$$

Where, $\left|\bar{\mathbf{u}}_{t}^{\text{meas}}\right|$ is the measured wind speed at time t. a and b are the smooth function which is estimated by using non parametric regression with forgetting factor.

Fluctuating Wind Speed Forecast Model

The fluctuating wind speed can be modelled similarly. Followings are the equation for the forecast of fluctuating wind speed.

$$\sigma_{t+k|t}^{\text{local}} = f_{\sigma}\left(|\mathbf{u}|_{t+k|t}^{\text{nwp}}, \theta_{t+k|t}^{\text{nwp}}\right) \tag{4}$$

$$\sigma_{t+k|t}^{\text{pred}} = a_{\sigma}\left(k, \theta_{t+k|t}^{\text{nwp}}\right)\sigma_{t}^{\text{meas}} + b_{\sigma}\left(k, \theta_{t+k|t}^{\text{nwp}}\right)\sigma_{t+k|t}^{\text{local}} \tag{5}$$

Maximum Gust Forecast Model

Based on the mean wind speed $\left|\bar{\mathbf{u}}_{t+k|t}^{\text{pred}}\right|$ and fluctuating wind speed $\sigma_{t+k|t}^{\text{pred}}$, the maximum gust forecast at time t targeted at k step ahead can be written as follows.

$$u_{t+k|t}^{\max} = \left|\bar{\mathbf{u}}_{t+k|t}^{\text{pred}}\right| + p_t \sigma_{t+k|t}^{\text{pred}} \tag{6}$$

Where, p_t is the peak factor which assumed not to depend on wind speed, wind direction nor forecast horizon and estimated so that the error shown in equation (7) is minimized.

$$\epsilon = p_t - \frac{u_t^{\text{max,meas}} - \left|\bar{\mathbf{u}}_t^{\text{meas}}\right|}{\sigma_t^{\text{meas}}} \tag{7}$$

Maximum Gust Upper Limit Estimation Model

The maximum gust obtained by the equation above forecasts the average value of maximum wind speed because equation (6) shows the average relationship between maximum gust and mean wind speed. In actual strong wind event, maximum gust which exceeds this value can be observed. In this study, the root mean square error of the maximum gust is modelled as shown in equation (8) and some portion of this error is added to the maximum gust to estimate the upper limit of maximum gust as shown in equation (9).

$$\epsilon_t\left(k, u_{t+k|t}^{max,pred}\right) = \sqrt{\left(u_t^{\text{max,meas}} - u_{t+k|t}^{\text{max,pred}}\right)^2} \tag{8}$$

$$u_{t+k|t}^{\text{max,upper}} = u_{t+k|t}^{\text{max,pred}} + \gamma\epsilon_t\left(k, u_{t+k|t}^{max,pred}\right) \tag{9}$$

Here, γ is a model parameter, the value of which will be discussed later.

Non Parametric Regression with Forgetting Factor

The model parameters and the functions which appear in the previous section can be estimated by using non-parametric regression with forgetting factor. In this method, any model which can be expressed as equation(10) is the target.

$$y_s = \mathbf{z}_s \mathbf{\Phi}^T(\mathbf{q}_s) + \epsilon_s \tag{10}$$

Where, y_s is an objective variable and $\mathbf{z}_s = (z_{s(1)} \ \ z_{s(2)} \ \ \cdots \ \ z_{s(M)})^T$ and $\mathbf{q}_s = (q_{s(1)} \ \ q_{s(2)} \ \ \cdots \ \ q_{s(N)})$ are the explanatory variables,

$\boldsymbol{\phi}^T(\mathbf{q}_s) = (\phi_{(1)}(\mathbf{q}) \quad \phi_{(2)}(\mathbf{q}) \quad \cdots \quad \phi_{(M)}(\mathbf{q}))$ are the functions that we want to identify. All the models proposed in this study can be shown in this form. For example, equation (1) can be expressed as equation (11) and $M = 1, N = 2$.

$$\phi_{(1)}(\mathbf{q}) = f\left(q_{(1)}, q_{(2)}\right) \\ \begin{cases} q_{(1)} = u_{t+k|t} \\ q_{(2)} = \theta_{t+k|t} \end{cases} \\ z_{(1)} = 1 \tag{11}$$

In the case of equation (3), $M = 2, N = 2$ and can be expressed as equation (12).

$$\begin{cases} \phi_{(1)}(\mathbf{q}) = a\left(q_{(1)}, q_{(2)}\right) \\ \phi_{(2)}(\mathbf{q}) = b\left(q_{(1)}, q_{(2)}\right) \end{cases} \\ \begin{cases} q_{(1)} = k \\ q_{(2)} = \theta_{t+k|t} \end{cases} \\ \begin{cases} z_{(1)} = u_{t+k|t}^{\text{local}} \\ z_{(2)} = u_{t+k|t}^{\text{meas}} \end{cases} \tag{12}$$

In non-parametric regression, the pair of the objective variables and explanatory variables are used to estimate the function $\boldsymbol{\phi}_{(\mathbf{q})}$ near the grid point $\mathbf{q}_p = (q_{p(1)} \quad q_{p(2)} \quad \cdots \quad q_{p(N)})^T$ locally. The estimation is done by minimizing the equation (13).

$$\epsilon = \sum_{s=1}^{t} \lambda^{t-s} w\left(\mathbf{q}_s, \mathbf{q}_p\right) \left(y_s - \mathbf{z}_s^T \hat{\boldsymbol{\phi}}_{p,t}(\mathbf{q})\right)^2 \tag{13}$$

Where, y_s, and z_s, are the measurement data obtained in the past. $\lambda^{t-s} w\left(\mathbf{q}_s, \mathbf{q}_p\right)$ is the weight used to evaluate the error and can be divided into two parts. λ is a parameter with the range of $0 < \lambda \leq 1$ and called forgetting factor. t is the current time stamp and s is the time when the measurement data were obtained, which means λ^{t-s} shows smaller value for older measurement data. By using this forgetting factor, it is realized to forget the older data and relative weight of the recent data is large. The larger value of λ means the use of longer period of data, and smaller value of λ means the use of shorter period. The effective number of data used for the estimation can be calculated by equation (14).

$$N_{eff} = \frac{\lambda}{1 - \lambda} \tag{14}$$

For example, when $\lambda = 1$, the number of data used becomes infinity, which is equivalent to use all the past data. Typically, in the field of wind power forecasting, 0.999 is used as a forgetting factor.

$w\left(\mathbf{q}_s, \mathbf{q}_p\right)$ shows larger weight when the distance between the point $\mathbf{q}_p$ where the function is estimated and the point $\mathbf{q}_s$ where we have training data. By using this weight, it is realized that the weight of the training data located closer to the point where the function is estimated, is larger. Multi-dimensional weight function $w(\mathbf{q_s}, \mathbf{q_i})$ can be expressed as shown in equation (15) as the product of the one dimensional weight.

$$w\left(\mathbf{q}_s, \mathbf{q}_p\right) = \prod_{j=1}^{N} W\left(\frac{\left|q_{s(j)} - q_{p(j)}\right|}{\hbar_j}\right) \tag{15}$$

Where, $\hbar_j$ is a parameter called band width which determines the smoothness of the estimated functions. The one dimensional weight function $W(x)$ is defined as equation (16).

$$W(x) = \begin{cases} (1 - x^3)^3 & 0 \leq x < 1 \\ 0 & 1 \leq x \end{cases} \tag{16}$$

Forecast Results

Figure 3 shows the time history of measured and predicted maximum wind speed for a day when strong gust event occurred. When adaptive model proposed in this study is used, the forecast accuracy is improved compared to the conventional static MOS model. Especially on the 31 of January, 2014, when the measured and predicted maximum wind speed differs considerably, forecast values differ considerably just after the beginning of the forecast. It is also pointed out that when MSM is used as an input as NWP, the prediction accuracy is better.

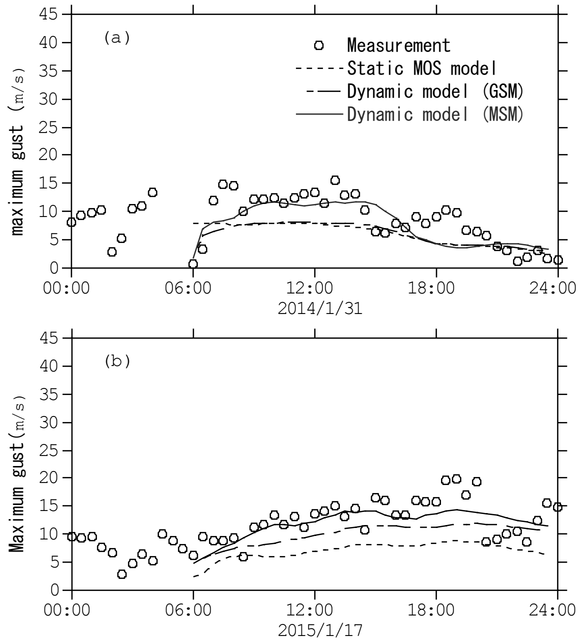

Figure 3. Measured and Forecasted maximum gust (a)31st of January, 2014; (b) 17th of January, 2015

Predictability of the gust event

For the predictability of gust event, the maximum wind speed upper limit estimation model is important. As maximum gust forecast model estimates the expected value of the maximum gust, the maximum gust upper limit estimate model is needed for the prediction of whether strong wind event occurs or not. When larger value of the model parameter γ in the maximum gust upper limit forecast model is used, the number of strong wind event forecasted becomes large. On the other hand, when smaller value of γ is used, less number of strong wind event is forecasted. Here, strong wind speed event is defined as number of the day when the maximum gust exceeds 15m/s.

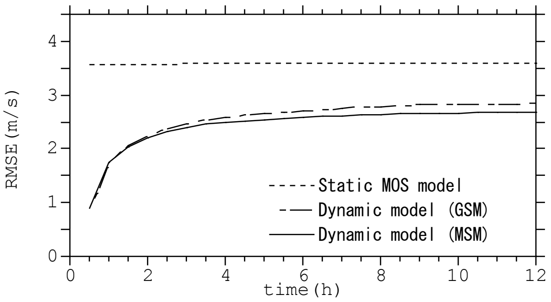

Figure 4 Root mean square error of the conventional MOS model and proposed model

ROC curve is widely used to evaluate the forecast model with such a parameter. ROC curve takes false positive rate on the x axis and true positive rate on the x axis and investigate the sensitivity of model parameter on the true positive rate and false positive rate. The contingency table of the strong wind event is shown in table 2. True positive is the case when strong wind event is forecasted and happened and the number of true positive case is denoted by a. False positive, true negative and false negative are defined according to table 2 and the number of each case are denoted by b, c and d, respectively. False positive rate is defined as $b/(b+d)$ and true positive rate is defined as $a/(a+c)$. If the forecast model is perfect, the true positive rate is 1 and false positive rate is 0, which means the ROC curve passes through the left top corner of the ROC curve. In a random forecast model, it is expected that the true positive rate is equal to the false positive rate, which means the ROC curve is the straight line connecting the bottom left corner and top right corner. Actual forecast model is located between these two curves, and when the curve passes closer to the top left corner, the model is better. The area under the curve (AUC) of ROC is a one of the index of the evaluation of the model. The larger value of AUC shows better model.

Figure 5 shows the ROC curve of the conventional static MOS mode, proposed dynamic model which uses GPV-GSM as an input NWP and proposed dynamic model which uses GPV-MSM as an input NWP. In any model, true positive rate can be larger by adopting larger value of γ, but false positive rate increases. Proposed model with an input of GP-GSM shows the curve which passes most close to the top left corner, implying that this is the best model. Table 3 shows the AUC value for each model, and it is shown that the AUC increases to 0.941 for dynamic model with GPV-MSM input from 0.842 for static MOS model.

Table 2 Contingency table

		Measurement	
		Yes	No
Forecast	Yes	True positive (a)	False positive (b)
	No	True negative (c)	False negative (d)

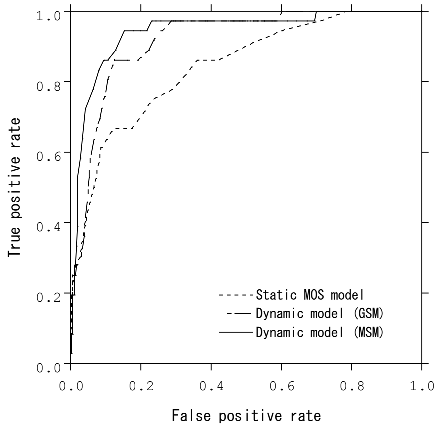

Figure 5 The ROC curve of conventional and proposed model

Table 3 The AUC of conventional and proposed model

Model	AUC
Static model (MOS)	0.842
Dynamic model (GPV-GSM)	0.919
Dynamic model (GPV-MSM)	0.941

Conclusions

The prediction accuracy by using dynamically adaptive model was improved compared to the conventional static MOS (Model Output Statistics) model. It was also shown that the prediction accuracy of maximum gust improves by utilizing the numerical weather prediction with higher horizontal resolution. The predictability of the gust with the maximum wind speed larger than 15m/s was evaluated by using ROC (Receiver Operating Characteristic) curve and AUC (Area Under the Curve) and was improved by the proposed method.

Acknowledgments

This research was carried out as a part of the collaboration of East Japan Railway Company (JR East) and the University of Tokyo (UTokyo). Dr. Yayoi Misu and Mr. Yosuke Nagumo helped up on collecting the online measurement data. .

References

[1] M. Shimamura and M. Matsunume : Development and implementation of Strong wind alart system, *JR EAST Technical Review*, No.13, pp. 36-43, 2005. (In Japanese)

[2] U. Hoppmann, S. Koenig, T. Tielkes and G. Matschke: A short-term strong wind prediction model for railway application: design and verification, *J. Wind Eng. Indust. Aerodyn.*, vol. 90, pp. 1127-1134, 2002.

[3] G. Giebel: The State-Of-The-Art in Short-Term Prediction of Wind Power – A Literature Overview, Project ANEMOS, 2003.

[4] H. A. Nielsen, T. S. Nielsen, A.K. Joensen, H. Madesen and J. Holst: *International Journal of Adaptive Control and Signal Processing*, Vol. 14, No. 8, pp. 813-828, 2000.