The 2020 World Congress on
Advances in Civil, Environmental, & Materials Research (ACEM20)
25-28, August, 2020, GECE, Seoul, Korea

Estimation of wind profile through the machine learning of remote and wind data

*Dong Hyeok Lee[1)] and Thomas Kang[2)]

[1), 2)] Department of Architecture & Architectural Engineering, Seoul National University, Seoul, Korea
[2)] tkang@snu.ac.kr

ABSTRACT

Wind profile is commonly defined using the vertical distribution of mean wind speed. This is one of the important indicators for the design of wind power generators and high-rise buildings. It is difficult to predict wind profiles numerically as a result of a variety of effects from complex variables, such as surrounding terrain and spatial distribution of neighboring buildings. Moreover, data in only a fewer places are being measured directly due to physical limitations. Thus, this study attempts propose data-based wind profile prediction method through machine learning of remote and wind data. Estimated wind profiles are verified using the collected remote and wind speed data.

1. INTRODUCTION

Wind profile is the distribution of wind speed from the surface in the vertical direction. Wind profile is used as the base for the design wind speed when designing loads for wind power generators and high-rise buildings. Until now, theoretical and simple experiments on the vertical distribution were conducted, and several models of the vertical distribution were presented. KBC 2016 uses exponential distribution as shown in Fig. 1 This parametric distribution works well in flat areas, whereas in urban areas uncertainty increases due to the influence of various variables. As shown in Fig. 2, it has been confirmed that there are significant differences between the observed data and the exponential distribution suggested by KBC 2016. This study deals with the improvement through non-parametric models and simple machine learning methods. Non-parametric models learn flexible models directly from the data. The non-parametric model has a problem of overfitting when there is little data, but it is expected that many recent observed data will solve this problem.

[1)] Graduate Student
[2)] Professor

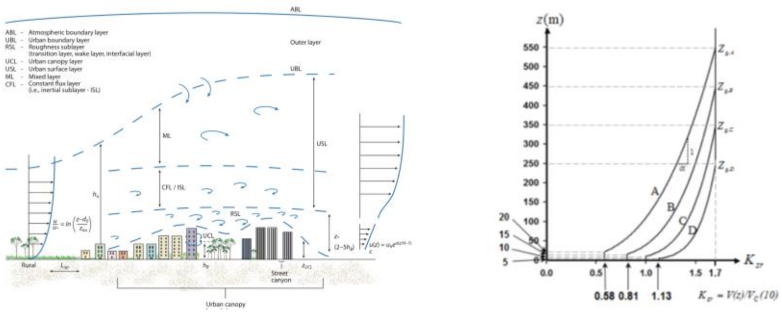

Fig. 1 (Left) Various zones of the urban boundary layer (Oke 1998)
(Right) Exponential wind profile (KBC 2016)

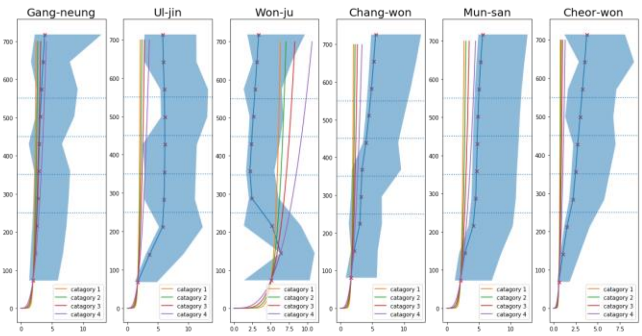

Fig. 2 Comparison between field-measured and calculated wind profiles based on KBC 2016

2. EXPERIMENT

In this experiment, distributions estimated by two methods are compared, including KBC 2016 method and proposed method. The vertical distribution data were used for one year in 2015 in Cheorwon, South Korea. The exposure coefficient is set to B by referring to KBC 2016.

2.1 Problem Definition.

Many variables affect the actual wind vertical distribution. In this experiment, variables were set by reference to KBC 2016. Fig. 3 shows this as a form of Bayesian network. However, θ_{loc} and θ_{exp} are difficult to be quantified. Thus, this experiment solved the problem by using data from a single region (middle of Fig. 3). In addition, the KBC 2016 method requires a standard height and wind speed. Thus, these were assumed to be conditional variables for wind speeds at the lowest height (right of Fig. 3).

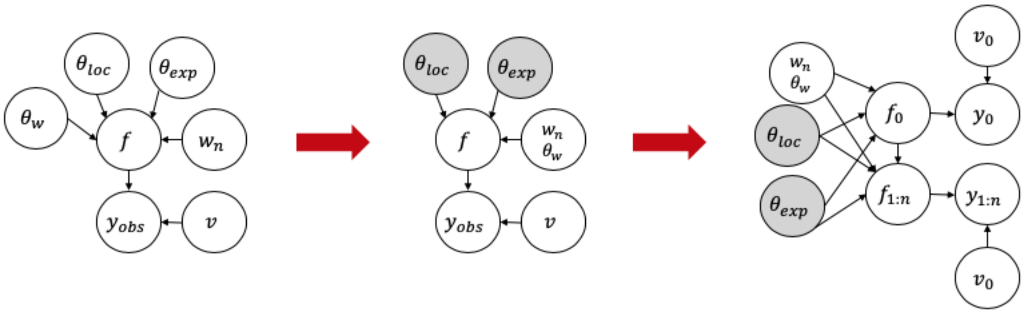

Fig 3. Bayesian network assumed in experiment,
θ_{loc} : factor of location, θ_{exp} : factor of exposure, θ_w : factor of weather
w_n : noise or unknown something, v: noise of observation

2.2 Error function.

The error function is as follows. The *L2 loss* of observed data and estimated values was used.

$$E(f_{obs}, g) = E_{f_{obs}}[\|f_{obs} - g(f_{obs})\|_2] \ , \ f_{obs} = \{(x, y)_{1:N}\}$$

In this experiment, Gaussian process (GP) was used as a non-parametric model. The vertical distributions estimated by KBC method and GP method are as follows, respectively.

$$g_{GP}(f) = \{m \mid f_{GP} \sim GP(m, K; f, \theta(f_0))\}, \ g_{KBC}(f) = \left\{ y_0 \left(\frac{x_i}{x_0}\right)^{\alpha}_{i=0:N} \mid f, \alpha \right\}$$

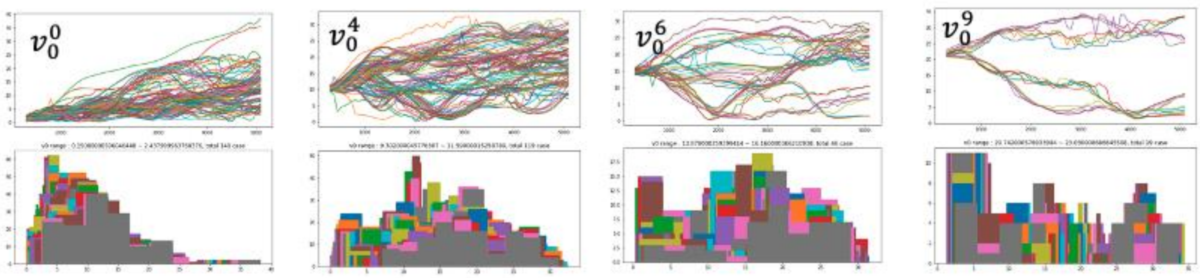

Fig. 4 Sample data according to the standard height wind speed (top) and the distribution of wind speed by height (bottom)

2.2 Preprocessing and obstacle

Data was preprocessed based on the wind speed at the reference standard height. The result is shown in Fig. 4. In the results, the higher the wind speed, the more specific patterns of the data were identified. At high wind speeds, the pattern was so unusual. Three assumptions are proposed in regard to this. The first is that the assumption applied in the experiment is incorrect. In the experiment, conditional estimation was made at the reference standard height wind speed for the comparison with KBC method. The second is a lack of samples. In this experiment, only one year-data was used. Thus, it is estimated that relatively high wind speed data are insufficient and that uncertainty has increased. Finally, it is assumed that the vertical wind speed distribution has a complex distribution. In this case, it is unreasonable to use GP method.

In the experiment, the second and third assumptions were adopted, respectively. Particularly, in order to use GP in the third assumption, the distribution of data was changed to Gaussian through the Flow model.

2.3 Experimental results

The results of the experiment are shown in Table 1. Based on the standard height, error values and variance obtained by bootstraps are obtained. The difference between the two methods according to the standard wind speed was confirmed. Overall, the proposed method shows better performance, especially in areas with lower wind speeds. However, in the case of variance, much better results are found at higher wind speeds.

Table. 1 Experimental results

	$E(\boldsymbol{f}_{\boldsymbol{obs}}, g_{KCI})$	$v_{voot}[E(\boldsymbol{f}_{\boldsymbol{obs}}, g_{KCI})]$	$E(\boldsymbol{f}_{\boldsymbol{obs}}, g_{GP})$	$v_{voot}[E(\boldsymbol{f}_{\boldsymbol{obs}}, g_{GP})]$
$0.15 \leq v_0 < 2.438$	52.337	31.208	**29.808**	18.968
$2.438 \leq v_0 < 4.726$	62.513	42.296	**42.464**	23.787
$4.726 \leq v_0 < 7.014$	42.845	30.453	**37,288**	22.006
$7.014 \leq v_0 < 9.302$	38.257	18.773	35.217	19.369
$9.302 \leq v_0 < 11.59$	49.316	17.405	47.728	16.325
$11.59 \leq v_0 < 13.878$	51.837	20.694	45.642	18.283
$13.878 \leq v_0 < 16.166$	62.375	35.607	51.227	26.988
$16.166 \leq v_0 < 18.454$	82.884	47.121	60.152	**24.201**
$18.454 \leq v_0 < 20.742$	83.984	50.736	58.726	**26.223**
$20.742 \leq v_0 < 23.03$	109.333	73.6	81.089	**14.912**

A simple additional experiment was conducted on the third assumption. With Flow model, the distribution of y-space was converted to z-space following multi-normal distribution. GP was conducted in z-space (Fig. 5). The RealNVP model was used as a Flow model. Fig. 6 shows the sample vertical distribution. As shown in the results, it has been confirmed that samples are quite similar to observed data.

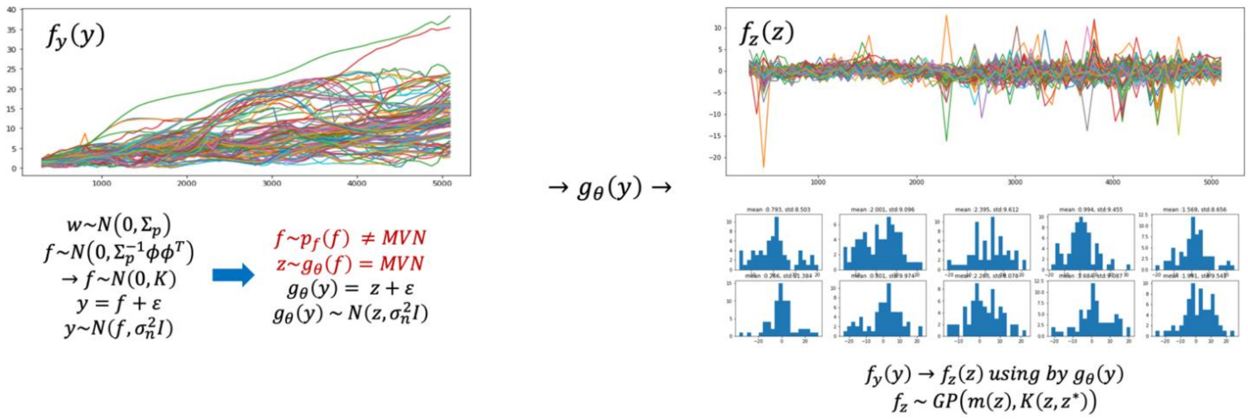

Fig. 5 Transformation of distribution using autoregressive model

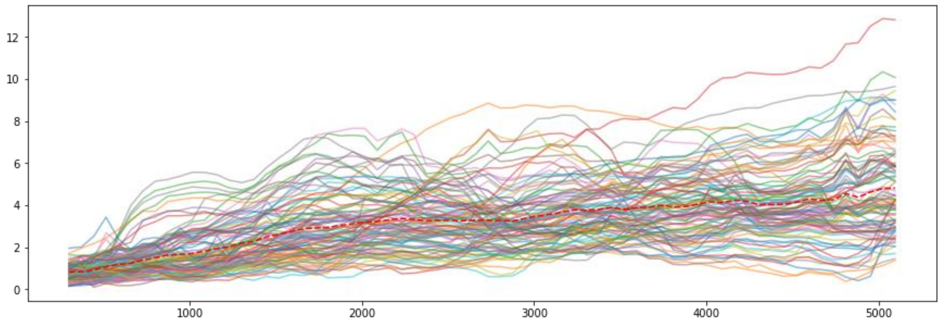

Fig. 6 Sampling from z-space

3. CONCLUSIONS

This study has shown that non-parametric models perform better compared to parametric models. It is expected that this method will greatly improve wind speed estimation and data conversion in the other two regions. Some limitations and future research possibilities could be identified. First of all, many assumptions were made in this experiment. Because of this, wind speed estimation is possible only in the areas where data are measured. If these variables can be quantified and measured, it is expected that reliable vertical wind speeds will be estimated where there is no direct observed data.

REFERENCES

Jung, S.-H., Kim, B.-J. and Ha, Y.-C. (2014), “Revision of Basic Wind Speed Map of KBC-2009.”, *Journal of the Architectural Institute of Korea*, **30**, p.37-47. (in Korean)

Fernando, H. J. S. (2010), “Fluid Dynamics of Urban Atmospheres in Complex Terrain”, Annual Review of Fluid Mechanics, vol. 42, p.365-389

Dinh, L., Sohl-Dickstein, J. and Bengio, S. “Density estimation using Real NVP.” arXiv: 1605.08803, 2017