

# A prototype machine learning-based system for predicting thunderstorm wind gusts

Seung Hun Choi <sup>a</sup>, Sungsu Lee <sup>b</sup>, Junyeong Kim <sup>c</sup>, Chang Hee Won <sup>d</sup>

<sup>a</sup>Wind & Disaster Prevention Research Center, Chungbuk National University, Cheongju, Republic of Korea, [krcsh@hanmail.net](mailto:krcsh@hanmail.net)

<sup>b</sup>School of Civil Engineering, Chungbuk National University, Cheongju, Republic of Korea, [sungsulee@chungbuk.ac.kr](mailto:sungsulee@chungbuk.ac.kr)

<sup>c</sup>Cuber Solution Co., Ltd., Cheongju, Republic of Korea, [wind.jykim@gmail.com](mailto:wind.jykim@gmail.com)

<sup>d</sup>Department of Business Strategy, WeeGee Co., Ltd., Cheongju, Republic of Korea, [wonch1027@naver.com](mailto:wonch1027@naver.com)

## SUMMARY

This study presents a prototype system for predicting thunderstorm wind gusts using machine learning and wind observations. Wind gusts are defined using operational strong wind thresholds, and events accompanied by thunderstorm phenomena in South Korea are classified as thunderstorm wind gusts. One-minute maximum wind speed time series are analyzed, and moving averages over multi-hour windows are applied to extract precursor signals embedded in the temporal evolution of the wind field. These precursors are used to construct a training dataset and probability distributions that relate pre-event features to subsequent gust occurrence. In the operational phase, real-time wind data are processed with the same methodology to evaluate gust occurrence probability. In addition, an anemometer and field-deployable prototype system are developed to provide continuous measurements and to serve as a testbed for validating the proposed prediction framework.

**Keywords:** *thunderstorm wind gusts, machine learning, early warning, statistical methodology, moving average*

## 1. INSTRUCTION

A thunderstorm is a localized storm formed by the development of cumulonimbus clouds and is accompanied by various adverse weather phenomena, including strong wind gusts, severe turbulence, lightning, heavy rain, and hail (Ban, 2003). Among these, thunderstorm wind gusts, which occur and dissipate rapidly within a localized area while generating strong wind speeds instantaneously, are not only difficult to prepare for but also challenging to observe (Lee and Kim, 2013). These gusts can cause significant damage to buildings, infrastructure, and temporary structures, as well as disruptions to transportation systems and power networks.

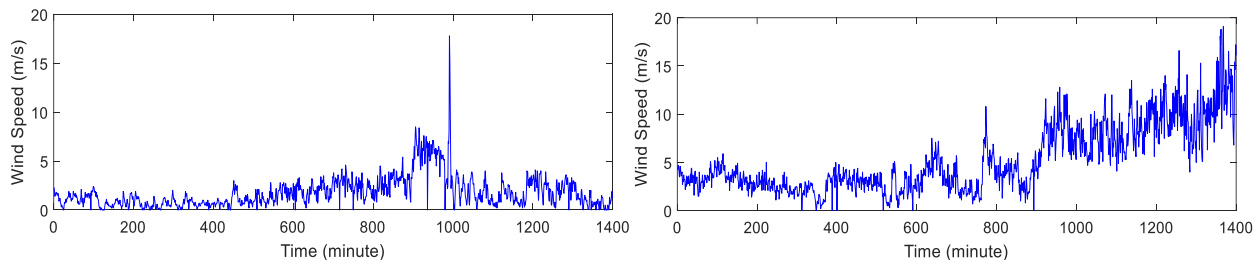
Despite their impact, thunderstorm wind gusts remain difficult to predict using conventional forecasting approaches that are primarily designed for larger-scale phenomena. The short-lived and small-scale nature of convective cells, along with limitations in observational coverage and model resolution, leads to substantial uncertainties in the timing and location of hazardous gusts. This study aims to present and evaluate a machine learning-based methodology for predicting thunderstorm wind gusts in advance, with the goal of contributing to a research-stage prototype system that can support early warning and risk management for localized wind hazards.

## 2. CLASSIFICATION FOR THUNDERSTORM WIND GUSTS

The Korea Meteorological Administration (KMA) issues strong wind advisories when meteorological phenomena are determined to pose a significant level of risk to life and property. In general, a strong wind advisory is issued when instantaneous wind speeds are forecasted to reach at least 20 m/s in inland regions or 25 m/s in mountainous regions. These operational thresholds are widely used as reference criteria for identifying hazardous wind conditions in South Korea. In this

study, wind gusts were defined as events in which the observed instantaneous wind speed exceeded 20 m/s at a given station. Among these, cases that coincided with thunderstorm phenomena—such as lightning, thunder, or hail reported in the synoptic observations—were classified as thunderstorm wind gusts, while the remaining cases were classified as non-thunderstorm wind gusts.

To analyze and compare the characteristics of thunderstorm and non-thunderstorm wind gusts, both conventional meteorological data and 1-minute maximum wind speed data provided by the KMA were utilized. The KMA conventional meteorological reports include information on the occurrence of thunderstorm phenomena at each observing station. In this study, days with recorded wind gusts were first identified using the wind observations, and the presence or absence of accompanying thunderstorm phenomena on those days was then determined from the meteorological reports. Subsequently, 1-minute maximum wind speed data were extracted for all stations where thunderstorm or non-thunderstorm wind gusts were observed across South Korea during the period from 2002 to 2009. These temporal-resolution data were used to construct time series for each gust event, enabling the examination of the temporal evolution of wind speed before, during, and after the occurrence of the gust. Figure 1(a) presents a representative example of the 1-minute average wind speed time series recorded on days when thunderstorm wind gusts occurred in South Korea, whereas Figure 1(b) shows a representative time series for non-thunderstorm wind gust events (Choi and Hidayat, 2002). Together, these figures highlight the differences in the temporal patterns associated with the two types of gusts.



(a) Thunderstorm wind gust (June 8, 2007, Yeongwol) (b) Non-thunderstorm wind gust (March 3, 2007, Sangju)  
Figure 1: The representative time series of 1-minute maximum wind speeds recorded on days with thunderstorm and non-thunderstorm wind gusts occurrences in South Korea

As shown in Figure 1(a), wind speed remains relatively low prior to the occurrence of the maximum wind speed and then undergoes a rapid increase and decay within a few minutes, clearly illustrating the short-lived and impulsive characteristics of thunderstorm wind gusts. This abrupt temporal behavior makes it difficult to visually distinguish any gradual buildup or precursor signal from high-frequency variability in the raw 1-minute time series. To extract underlying trends and identify potential precursors, the moving average method was utilized in this study (Kim *et al.*, 2014). By smoothing the high-frequency fluctuations while preserving the broader temporal evolution, the moving average approach provides a useful tool for examining how wind conditions evolve in the hours leading up to a wind gust event.

Figure 2 depicts the results of applying the moving average method to the time series in Figure 1(a), with calculations performed over averaging intervals of 3, 6, 9, and 12 hours. As the averaging window increases, short-term fluctuations are progressively filtered out, and a clearer background trend emerges. As shown in Figure 2, wind speeds exhibit a gradual but noticeable increase approximately 6 to 12 hours before the occurrence of the thunderstorm wind gust. This result

indicates that, despite the very short duration of the gust itself, information relevant to gust occurrence is embedded in the longer-term evolution of the wind field, providing a physical basis for incorporating pre-event time series features into a machine learning framework for thunderstorm wind gust prediction.

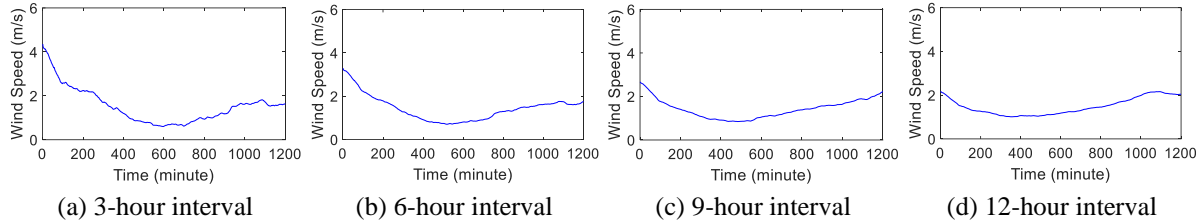


Figure 2: Results of moving average calculations for thunderstorm wind gust time series (when the time is 0, it represents the occurrence time of the thunderstorm wind gust)

### 3. DEVELOPING MACHINE LEARNING ALGORITHMS FOR THUNDERSTORM WIND GUSTS

This study seeks to develop a machine learning algorithm for the prediction of thunderstorm wind gusts, utilizing historical 1-minute maximum wind speed data from observed events. The approach assumes that precursor signals for gust occurrence are embedded in the temporal evolution of the moving-averaged wind speed. Based on this assumption, past observed wind speed data are first processed by calculating moving averages over selected time windows. The timing of the moving average rise point is then identified for each event, and their statistical characteristics are used to construct a probability distribution for the occurrence time and likelihood of thunderstorm wind gusts. These probability distributions, derived from historical data, constitute the machine learning dataset and provide the basis for evaluating the relationship between precursor features and subsequent gust occurrence.

The workflow of the proposed machine learning algorithm for thunderstorm wind gust prediction is outlined in Figure 3. In the operational phase, real-time wind speed is continuously measured and subjected to the same moving average calculation as applied to the historical data. When a moving average rise point is detected in the real-time series, its characteristics are compared with the learned probability distributions obtained from past events, and the occurrence time and probability of a thunderstorm wind gust are evaluated. Finally, the system checks whether the evaluated probability exceeds a predefined threshold and, if so, indicates the potential occurrence of a thunderstorm wind gust. Through this integrated process, the proposed framework links historical precursor patterns with real-time observations to provide a research-stage prototype for short-term prediction of thunderstorm wind gusts.

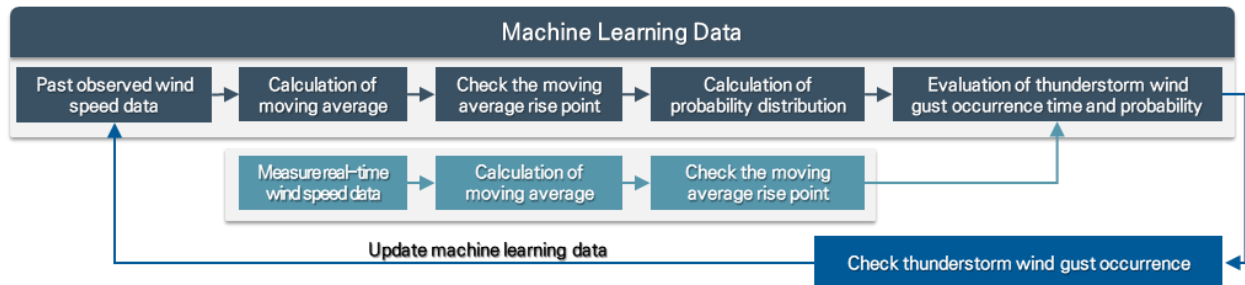


Figure 3: Workflow of the machine learning algorithm for predicting thunderstorm wind gusts

#### 4. DEVELOPMENT OF AN ANEMOMETER AND PROTOTYPE SYSTEM

As shown in Figure 4, this study also includes the development of an anemometer and prototype system for field deployment. The anemometer is designed to measure wind speed with high temporal resolution and improved durability under strong wind and precipitation conditions. Its sensor configuration, signal conditioning circuit, and data logging module are optimized to interface seamlessly with the proposed thunderstorm wind gust prediction framework. The prototype integrates real-time wind measurements, on-site data storage, and communication functions, enabling continuous monitoring at target locations and providing testbed conditions for validating the performance of the machine learning-based prediction system.



Figure 4: Anemometer and prototype system developed in this study (left: anemometer, right: prototype system)

#### 5. CONCLUSIONS

This study utilized machine learning to develop a predictive model for thunderstorm wind gusts. Wind gusts were defined as instantaneous wind speeds exceeding a specific threshold, and wind gust events observed on days with thunderstorm occurrences in South Korea were classified as thunderstorm wind gusts. The time series characteristics of 1-minute average wind speeds recorded on days with thunderstorm wind gust occurrences were analyzed. By applying the moving average method to these time series, precursors associated with thunderstorm wind gusts were identified. These findings were utilized to construct a training dataset, serving as the foundation for the development of a machine learning algorithm. In addition, this study includes the development of an anemometer and prototype system for field deployment, which serves as a testbed for validating the performance and practical applicability of the proposed thunderstorm wind gust prediction framework.

#### ACKNOWLEDGEMENTS

This research was supported by a grant (RS-2025-02633557) of Technology Development Program for Strengthening Resilience Against Urban Wind Disasters funded by Ministry of Interior and Safety (MOIS, Korea).

#### REFERENCES

- Ban, K.S., 2002. Advancements in Technology for Predicting Monsoon Characteristics and Adverse Weather Conditions Affecting Flight Safety. *Aviation Promotion*. 2 (30), 75-102.
- Choi, Edmund C.C and Hidayat, Ferry A, 2002. Gust Factors for Thunderstorm and Non-thunderstorm Winds. *Journal of Wind Engineering and industrial Aerodynamics*. 90(12). 1683-1696.
- Lee, S. and Kim, J.Y., 2013. Spatial and Temporal Analysis of Thunderstorm Wind Gust. *Journal of Korea Spatial Information Society*. 21 (4), 1-6.
- Lee, S., Kim, J.Y., Ham, H.J., and Kim, J., 2014. Stochastic Analysis of Prognosis of Thunderstorm Wind Gust. *Journal of the Wind Engineering Institute of Korea*. 18 (1), 21-27.