

# **A new strategy of the Machine learning based data calibration method for boundary layer observations with a low-cost mini-radiosonde system**

**Shih-Hao Su<sup>1\*</sup>, Min-Lun Wu<sup>1</sup>**

<sup>1</sup>Chinese Culture University, <sup>2</sup>National Taiwan University

## **Abstract**

Even today, radiosondes remain the most direct method for measuring atmospheric vertical profiles. Generally, their high cost has limited sampling frequency in field campaigns. Low-cost mini-radiosondes now enable high-frequency, dense 3D observations, but maintaining consistent data quality across instruments is crucial. This study presents the strategies of the machine-learning (ML) based calibration method for the novel mini-radiosonde system, Storm Tracker (ST). To address temperature and moisture errors caused by solar radiation heating. We adopted a couple of ML modules as the kernel for data calibration. We used the co-launch radiosonde data from ST and Vaisala-RS41 between 2018 and 2022 as a training dataset to develop the ML model. The data calibration system can correct the temperature using the ML modules and then calibrate the moisture based on the corrected temperature.

The results show the linear calibration model exhibits good calibration capability for observing the planetary boundary layer (1000-700 hPa). On the other hand, the non-linear calibration model, with a more complex structure, maintains even better calibrating performance statistically. The data correction modules mentioned above are all based on a "point-to-point" data structure. However, for vertical atmospheric structure characteristics, such as atmospheric stability, it is likely more important. Based on this concept, we propose a new mini-radiosonde data correction strategy that can be more suitable for sequential data architectures. This approach integrates the sequential variation of radiosonde measurements into the correction model.

Key words : data calibration, machine learning, radiosondes