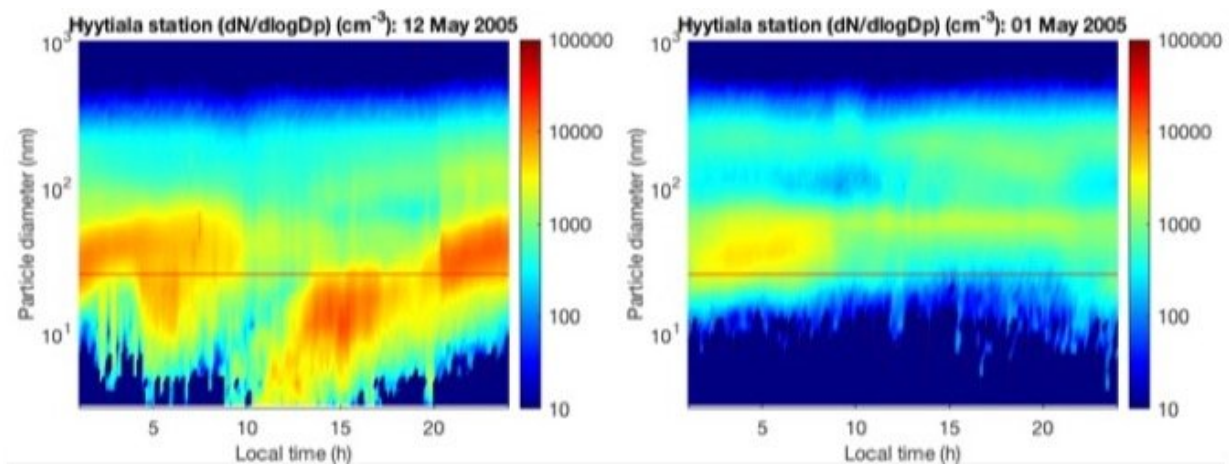# Getting the most out of atmospheric data analysis

October 29 2018

Atmospheric new particle formation non-event (a) and event (b) days at Hyytiälä forest, Finland, in May 2005. A non-event day is assumed when the day is clear of all traces of particle formation whilst an event day occurs when there is a growing new mode of aerosol particles in the nucleation size range (3-25 nm) prevailing over several hours. Credit: Kanazawa University

New particle formation in the atmosphere provides the nucleation centres required for the formation of clouds, making it an important process for understanding climate. Efforts to investigate the complex balance of chemistry and physics that leads to new particle formation have resulted in the acquisition of very large data sets. A team of researchers based at a number of centers, including Kanazawa University, has developed a mutual information approach to interpreting

atmospheric data collected over an 18-year period at the SMEAR II station in Hyytiälä, Finland. Their findings were published in *Atmospheric Chemistry and Physics.*

The study, an extension of previous data mining work that was carried out on measurements acquired at the same station, uses an additional 10 years of data, which is expected to enhance the reliability of the results and accuracy of the conclusions. In contrast to the previous studies that used data mining methods based on clustering and classification, the new [approach](#) looked at the mutual information between observed new particle formation events and a variety of measured variables. The method was shown to be powerful and computationally light, highlighting its potential as a useful tool.

"The conclusions reached using our analysis were found to agree with previous findings; however, the analysis was achieved without supervision and did not require a deep understanding of the physics," study coauthor Adam Foster says. "Previous studies have required field, lab, and theory work, so being able to get to where we have through data analysis is a very positive step for future studies."

The work showed that new particle formation was strongly correlated with water content and sulfuric acid concentration, as well as other factors such as temperature, relative humidity, condensation sink (how quickly molecules and small particles condense onto existing particles), and radiation.

"We hope that the method will provide a robust first option for analyzing [atmospheric data](#) sets," says Foster. "We aim to extend our approach to [data sets](#) from other SMEAR stations, as well as to use it to investigate different phenomena and their influences."

These findings are expected to be implemented widely in the field of

atmospheric science. It is hoped that the effect of other variables such as volatile organic compounds and aerosol particles below 3 nm can be assessed using the reported approach.

More information: Martha A. Zaidan et al, Exploring non-linear associations between atmospheric new-particle formation and ambient variables: a mutual information approach, *Atmospheric Chemistry and Physics* (2018). DOI: 10.5194/acp-18-12699-2018

Provided by Kanazawa University