

Atmos. Chem. Phys. Discuss., referee comment RC2  
<https://doi.org/10.5194/acp-2021-196-RC2>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## **Comment on acp-2021-196**

Anonymous Referee #1

---

Referee comment on "Forecasting and identifying the meteorological and hydrological conditions favoring the occurrence of severe hazes in Beijing and Shanghai using deep learning" by Chien Wang, Atmos. Chem. Phys. Discuss.,  
<https://doi.org/10.5194/acp-2021-196-RC2>, 2021

---

## **Interactive comment on "Forecasting and Identifying the Meteorological and Hydrological Conditions Favoring the Occurrence of Severe Hazes in Beijing and Shanghai using Deep Learning" by C. Wang**

### **General comments**

This paper presents a Deep Learning based research of severe air pollution events in Southeast China, namely "haze events". The technological backbone of this research is a deep Convolutional Neural Network (CNN) that is commonly used in advanced image processing algorithms and applications but recently introduced to the climate and weather prediction communities in the private sector. The author well justifies the model choice and clearly elaborates on its technical structure. The model performance on the two selected regions of interest (Beijing and Shanghai) shows impressive accuracy when predicting severe haze events, as defined by the author. An additional classification analysis extended from that same model could also reveal the significance of each meteorological feature for severe haze event occurrence, contributing to the optimization of any future CNN analyses for similar purposes and raising new insights for further physical discussion. This paper introduces new techniques to the field of weather, climate, and environmental prediction, leveraging the growingly available computing power. Furthermore, this work joins the cutting edge of spatiotemporal physical models, recently shifting from deterministic into statistic methods. While most of the relevant research in the private sector is already implementing similar weather and climate prediction approaches, I see great promise and importance in introducing such work to our academic community.

Therefore, I recommend this paper to be published in *Atmospheric Chemistry and Physics*, with reservation to the author's response to my comments below, hoping that my notes could help the author improve his paper.

### **Specific comments**

- The selected threshold for defining a severe haze event in the 2-class training is set to days whose surface visibility decreases below the 25<sup>th</sup>. While the value of that percentile varies in time and space, I suggest elaborating more on selecting that exact percentile threshold.
- Please elaborate on the input data and its possible effects on the results. How many ground stations are used for the analysis in each city? What are the potential limitations that are resulted from the ERA5 spatial resolution of 0.25 degrees?
- Please include, possibly as a supplemental, some technical details regarding the CNN analysis. What were the data and computational volumes and costs? What kind of computation platform was used, how long each training session take? What were optimization and approximation procedures implemented? Such information could assist future researchers when planning their analyses and also provide the scientific community with a technological benchmark for comparison with future projects.

### **Technical corrections**

- Please keep consistency in number representation along the manuscript (e.g. "11,376" in P. 4 Line 132 vs. "14975" in P. 11 Line 338).
- Please keep consistency in technical terminology along the manuscript (e.g., "class-1" vs. "class 1", etc.).
- Please specify the unit following physical quantities (e.g. "...heights at 500 (Z500) and 850 (Z850) hPa" in P. 6 Line 188).
- Please follow a consistent terminology for classes 0 and 1 in the 2-class analysis (e.g., "non-haze events" and "severe haze events").
- 2 Line 38: Please change "event" to "events".
- 2 Line 39: Please change "has" to "has".
- 4 Line 119: Please change "Introduction" to "introduction".
- 5 Line 158: Please add punctuation marks where necessary.

- 5 Line 176: I suggest replacing or dropping the words "longitude-latitude" that are already self-embedded in surface map objects.
- 6 Line 214: Please change "metrics" to "metric".
- 6 Lines 219-220: Please clarify that sentence.
- 7 Line 225: Please correct a typo "class0weight"
- 8 Line 267: Please modify to "there are many hyper-parameters in HazeNet that need ...".
- 12 Line 398: Please change "soli" to "soil".
- 8 – caption (Lines 405-408): Please specify the explicit variable descriptions for better readability.