Again, I truly appreciate the constructive suggestions from the reviewers. The following are the point-to-point responses to the reviewers' comments (marked with Italic font).

Reviewer #2

Specific comments:
1. Page 2, abstract: The abstract reads like a short summary of the introduction. I suggest the author includes some key findings (e.g., the performance of the model and/or the identified haze favorable environment in Beijing and Shanghai) in the abstract.

The reviewer's suggest has been accepted. The revised Abstract should read more specific on the findings of this study:
"Severe haze or low visibility event caused by abundant atmospheric aerosols has become a serious environmental issue in many countries. A framework based on deep convolutional neural networks containing more than 20 million parameters, namely HazeNet, has been developed to forecast the occurrence of such events in two Asian megacities: Beijing and Shanghai. Trained using time sequential regional maps of up to 16 meteorological and hydrological variables alongside surface visibility data over the past 41 years, the machine has achieved a good overall performance in identifying the haze versus non-haze events and thus their respectively favorite meteorological and hydrological conditions, with a validation accuracy of 80% in both Beijing and Shanghai cases, exceeding the frequency of non-haze events or no-skill forecasting accuracy, and a F1 score specifically for haze events nearly 0.5. Its performance is clearly better during months with high haze frequency, that is all months except dusty April and May in Beijing and from late autumn through entire winter in Shanghai. Certain valuable knowledge has also obtained from the training such as the sensitivity of the machine's performance to the spatial scale of feature patterns that could benefit future applications using meteorological and hydrological data. Furthermore, an unsupervised cluster analysis using features with a greatly reduced dimensionality produced by the trained HazeNet has, arguably for the first time, successfully categorized typical regional meteorological-hydrological regimes alongside local quantities respectively associated with haze and non-haze events in the two targeted cities, providing substantial insights to advance our understandings of this environmental extreme. Interesting similarities in associated weather and hydrological regimes between haze and false alarm clusters, or differences between haze and missing forecasting clusters have also been revealed, implying that factors such as energy consumption variations, long-range aerosol transport, and beyond could also influence the occurrence of hazes, even under unfavorite weather conditions".
.
2. Page 4-7, Section 2: I still feel that the structure of this section is quite complex for readers to follow. I think it is better to add subsection titles and rearrange paragraphs a little bit. The subsections can be defined as follows: 2.1 network architecture (including content of lines 128-154), 2.2 kernel size optimization, 2.3 Data (lines 155-191), and 2.4 Training methodology (lines 192-214). The subsection bullet points (e.g., 2.1) may not be necessary but the author could at least have titles bolded like what you already had for kernel size optimization.

This is a good point. Three subsections have been created, respectively with titles of: 2.1 Network architecture; 2.2 Training data and methodology; and 2.3 Kernel size optimization.

3. Page 7, line 232: "…. see next section and Method)…." I am not sure about which "Method" you refer to here.

This perhaps is a leftover from an older version. The words of "see next section and Method" have been removed.

4. Page 9, line 301: The author has a subtitle for this section for reducing input features, but there is no subtitle for the prior paragraphs. This is quite confusing for readers to understand the number of input features you used in the results presented in Figures 4 and 5. I suggest the author add bold subtitles for the prior paragraphs: "Model performance using 16 input features".

The reviewer's point is well taken. In the captions of Fig. 4 and 5, the number of features used in training were indicated in the previous version. To avoid further confusion, an additional note of "with 16 features" has been added in Fig. 4 caption on original Line 257 for (Right Top) panel. In addition, a sentence was added in the end of the opening paragraph (original Line 252): "Also note that, unless otherwise indicated, results shown in this Section are obtained using 16 features". Therefore, the only exception of the above is in Fig. 4 (right Bottom), where a note of "16 and 9 features" already existed (so did the legend notes of the figure). Lastly, "(Fig. 4)" in the Line 321 (original) has been revised to "(Fig. 4, Right Bottom)".

5. Page 10, lines 321-322: It seems like the following cluster analysis uses the model results with nine input features. Maybe the author can rewrite or add another sentence clearly stating that the subsequent cluster analysis is conducted using the model outputs with nine input features.

Yes, indeed a sentence is needed here. It has been added in the original Line 340-341, "…the trained HazeNet for Beijing and Shanghai using 9 instead of 16 features, benefited from the effort of reducing the number of input features as described in the end of last Section, have been used here to…".

6. Page 12, lines 412-416: It will be helpful if the author could provide more analyses, discussions, or insights on why HazeNet misses the FN cases, especially the cases in cluster 1 since they are major cases (Table S1). It seems to me that the key differences between cluster 1 in FN cases and TP cases are shown in U10, V10, DTCV, and SW1. Do the patterns of cluster 1 (or other clusters) in FN cases represent weather patterns unfavorable for haze events?

The original discussions from Line 409 to 416 have been revised as: "On the other hand, among three FN clusters (also associated with haze events but missed in prediction), only the first cluster (the major cluster of FN) displays certain similarity to TP clusters across various features. Even for this cluster, the characters of the airmasses distantly surrounding Beijing differ substantially from those of TP clusters, as seen from the patterns of temperature (DT2M, T2SM), wind particularly V10, and column water (DTCV) that reflect a much weaker weather system on the west. The patterns of BLH, SW1, and SW2 also differ from those of TP, indicating a different near site boundary layer and hydrological condition. Such differences appear to be even more evidently in the two other (minor) clusters, *e.g.*, the size and strength of high relative

humidity center covering Beijing are even different. This result suggests a possible reason for why HazeNet missed these targets, that is haze might occur under unfavorable weather and hydrological conditions owing to, *e.g*., certain energy consumption scenarios".

7. Page 14, lines 438-442: In my previous comment #13 I asked how the trend or seasonality would affect the clustered features. I agree that the analysis could be for future work, but I suggest the author includes several sentences mentioning this point and potential influences of trend or seasonality on the clustered results.

The original sentence of "…to appreciate the conventional regional and local meteorological and hydrological patterns associated with various events" has been revised to "…to appreciate the conventional regional and local meteorological and hydrological patterns, and to detect thus to implement additional analysis, if necessary, on the possible impact of seasonality or trend associated with various events".