Appendix:

In this study, we tried several possible ways of grouping the data into two categories for each megacity, and various regression models in terms of a formula set containing two stepwise regression models were derived. The models with best simulation results were regarded as the optimal empirical regression models (while the others were not shown). The grouping methods of optimal empirical regression models in each city and their reasons are summarized below.

Relative humidity in Beijing shows a large variation compared to other three cities (Fig. a). Given that the relationship between particle scattering capability and relative humidity is nonlinear, the impacts of relative humidity on visibility in high or low relative humidity environments are different (Malm & Day, 2001, Wang et al., 2010). Therefore, we defined the cases with relative humidity > 51% (mean relative humidity in Beijing) as “High RH” cases; otherwise they were classified as “Low RH” cases.

In Shanghai, hazy days with heavy aerosol pollution always occurred on days with low wind speed (Fu et al., 2008). In this study we found that the impact of wind speed variation on visibility change was different on days when wind speed was greater or less than 3 m/s. In particular, when wind speed was < 3 m/s, the correlations between wind speed and visibility (R^2=0.140) and the slope of linear regression (2.45 km per m/s) were both stronger than those when wind speed was > 3 m/s (R^2=0.057, 1.78 km per m/s respectively) (Fig. b). Therefore, “High WS” and “Low WS” cases with wind speed greater and less than 3 m/s respectively were grouped.

In Guangzhou, negative regression slope (-0.451) between temperature and visibility was found when temperature was < 23°C (mean temperature in Guangzhou). In contrast, the slope was positive (0.770) when temperature was > 23°C (Fig. c). This result suggested that the dispersive effect of high temperature on visibility improvement was observed only on days with temperature > 23°C. As a result, we grouped the dataset into two groups (i.e., “High TEMP” cases with temperature higher than 23°C and “Low TEMP” cases otherwise).

Furthermore, Tsai (2005) found that the parameter of ln[PM_{10}] could better predict the impacts of PM_{10} on visibility than [PM_{10}] (concentration of PM_{10}). In this study, we also found that correlation between [PM_{10}] and visibility in Beijing was –0.647, while that between ln[PM_{10}] and visibility could be raised to 0.740. Such a phenomenon was also observed in the data of Shanghai and Guangzhou, where the correlations between ln[PM_{10}] and visibility (0.688 and 0.559 respectively) were slightly larger than that between [PM_{10}] and visibility (–0.657 and –0.532 respectively). As a result, we used ln[PM_{10}] rather than [PM_{10}] as an independent variable in the optimal empirical regression model for Beijing, Shanghai and Guangzhou.

In Chengdu, the original model tended to underestimate visibility when both PM_{10} concentration and relative humidity were high (Fig. d), which was probably because of the nonlinear growth of particle scatting capability in high relative humidity environment (Malm & Day, 2001). We thus defined “episode day” as a day with relative humidity > 75% (mean relative humidity in Chengdu) and PM_{10} concentration > 100 μg/m^3 (annual average of NAAQS II), then separated the data into two groups (i.e., “episode day” and “non-episode day”).

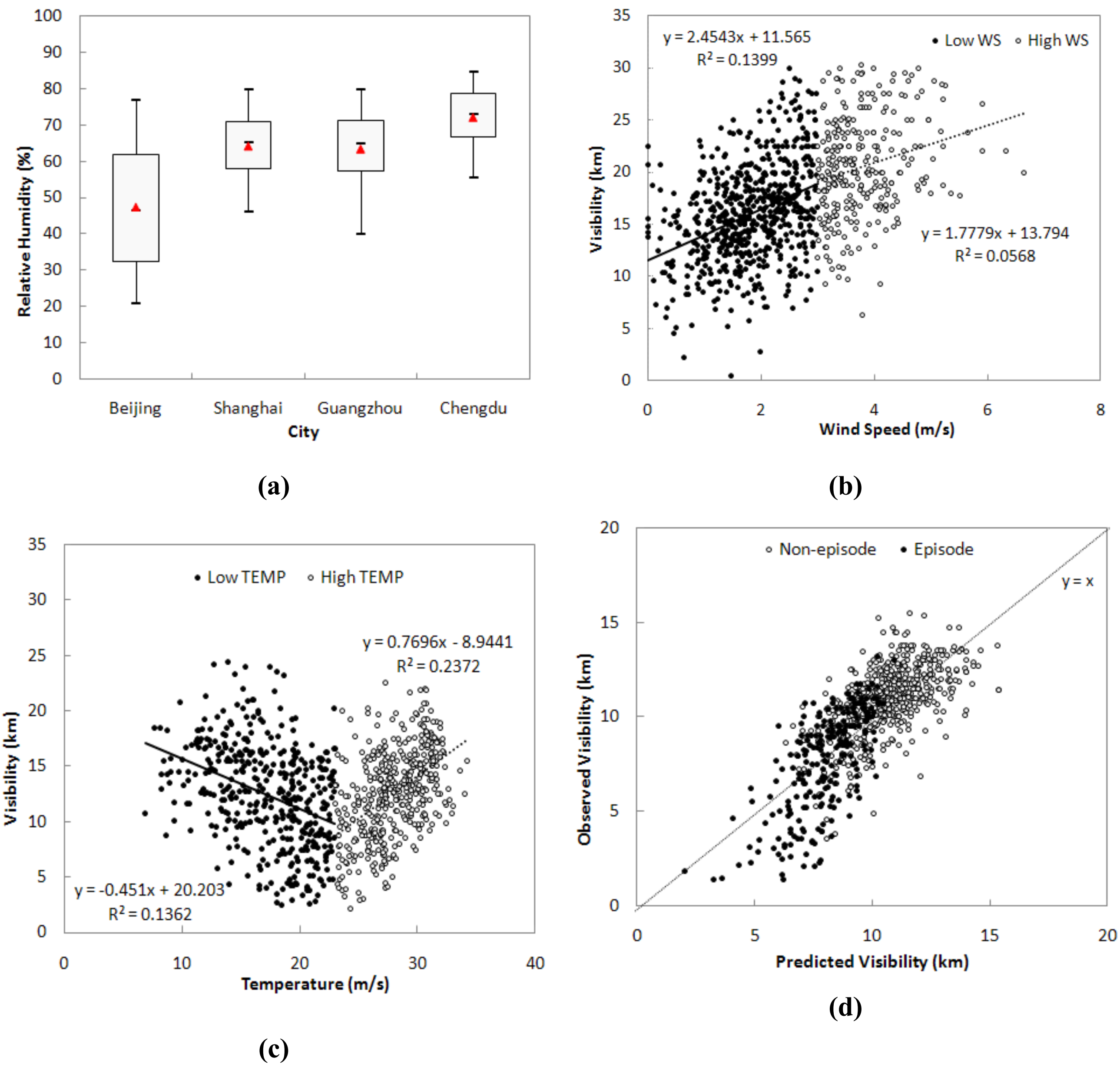

Fig. (a) Different percentiles (box & whisker) and average (triangle) of relative humidity in Beijing, Shanghai, Guangzhou and Chengdu; (b) Scatter plot of visibility and wind speed in Shanghai; (c) Scatter plot of visibility and temperature in Guangzhou; (d) Scatter plot of observed and predicted visibility in Chengdu.