Analyzing the Characteristics of Cloud Condensation Nuclei (CCN) in Hebei, China Using Multi-Year Observation and Reanalysis Data

Hengqi Wang 1,*, Meng Zhang 1, Yiran Peng 1,*, and Jing Duan 2,*

1 Ministry of Education Key Laboratory for Earth System Modeling, Institute for Global Change Studies, Department of Earth System Science, Tsinghua University, Beijing 100084, China; wanghq19@mails.tsinghua.edu.cn (H.W.); meng-zha16@mails.tsinghua.edu.cn (M.Z.)

2 Key Laboratory for Cloud Physics of China Meteorological Administration (CMA), CMA Weather Modification Centre, Beijing 100081, China

* Correspondence: pyiran@mail.tsinghua.edu.cn (Y.P.); duanjing@cma.gov.cn (J.D.)

Abstract: The study of the characteristics and variations of cloud condensation nuclei (CCN) plays an important role in understanding the effects of aerosol–cloud interactions. This paper selected observation data in a city region of Shijiazhuang in North China from 2005 to 2007, along with the corresponding MERRA-2 and ERA5 data, to analyze the characteristics of CCN, sort the factors affecting the diurnal variation of CCN number concentration ($N_{CCN}$) according to their importance, and build the relationship between $N_{CCN}$ and supersaturation (SS) in the heavily polluted region. The results show that there was a bimodal distribution of a daily time series for $N_{CCN}$ in Shijiazhuang, China. By calculating the correlation between CCN and pollutants observed in winter 2007, we identified that the dominant factor for peaks of $N_{CCN}$ is SO$_2$ in the morning but NO$_2$ in the evening. We also ranked the factors affecting the diurnal variation of $N_{CCN}$ by using observation and reanalysis data and found that the concentration of pollutants is the greatest impact factor in summer, but the atmospheric stability is the greatest factor in winter. Finally, we determined the relationship between $N_{CCN}$ and SS according to the Twomey formula ($N_{CCN} = cSS^k$) and found there was a reasonable value range (i.e., 0.5~0.7) for the parameter $k$ in East and North China. Specifically, it is more reasonable for $k$ to be 0.5 in summer and 0.7 in winter.

Keywords: CCN; MERRA-2 and ERA5; long-term observation; heavily polluted region; Twomey formula; diurnal variation

1. Introduction

The multiphase system composed of solid and liquid particles suspended in the air is called aerosol. Changes in aerosol composition, particle size, and number concentration have important effects on the radiation budget, clouds, and precipitation in the earth system [1–4]. Aerosols can cool (and warm) the earth–atmosphere system by scattering (absorbing) solar radiation directly (i.e., aerosol–radiation interaction). In addition, a part of aerosols can also indirectly affect cloud radiative properties [2] and precipitation processes [5] by influencing the cloud droplet size/number of warm stratus clouds, namely aerosol–cloud interaction. These aerosols are called cloud condensation nuclei (CCN). CCN plays a significant role in aerosol and cloud physics as it links aerosol characteristics to the cloud formation process. Due to the complexity of aerosol–cloud interaction and cloud/rain processes, there are large uncertainties in estimating the effect of aerosol on climate using global climate models (GCM) [6–8]. To better evaluate climate model simulation on aerosol–cloud interaction and reduce the uncertainty, it is necessary to conduct a more in-depth study on the characteristics of CCN over large-scale and long-term spatial-temporal domains.
The number concentration of CCN (N\textsubscript{CCN}, cm\textsuperscript{-3}) has a large spatial variability. For example, in the case of 1% supersaturation (SS), the clean region usually has N\textsubscript{CCN} < 1000, the polluted region has $1000 \leq N\textsubscript{CCN} < 10,000$, and the heavily polluted region even shows $N\textsubscript{CCN} \geq 10,000$. It is relatively common to carry out studies in clean or polluted regions, such as New York (N\textsubscript{CCN} < 3000 at SS = 1 %, similarly hereinafter) [9], Helsinki (N\textsubscript{CCN} < 1500) [10], Shanghai (N\textsubscript{CCN} < 8500) [11], Azores (N\textsubscript{CCN} < 300) [12], and Seoul (N\textsubscript{CCN} < 10,000) [13]. Alternatively, few observations were carried out in heavily polluted regions, such as India (N\textsubscript{CCN} at SS = 0.6 % up to 25,000) or the North China Plain and Yangtze River Delta in China (13,000 $\leq N\textsubscript{CCN} < 30,000$) [14–19]. Fewer observations lead to a lack of thorough understanding of the characteristics of CCN in heavily polluted regions. Meanwhile, lacking long-term observational data is another reason for restricting the research of CCN characteristics in heavily polluted regions. Previous studies usually focused on a short-term variation of CCN due to the limitation of observational conditions. Most studies had an effective time of observation for less than one year [10–12,16–30], some for one season [16,18,19,21,26,30] and even for several days [17,24,25,28,29]. Although some studies have been carried out using long-term observations located at some representative areas [13,31–34], most of them are in clean or polluted areas where biomass burning is the main source of pollution [32–34], and the observations sampled on high-polluting areas affected by human activities are relatively rare [13,31]. Short-term data have limited representativeness and are not applicable for evaluation on large-scale simulations with GCM. Therefore, it is essential to carry out a long-term CCN observation for better characterizing CCN in heavily polluted regions.

N\textsubscript{CCN} has obvious diurnal changes, and usually shows a unimodal or bimodal distribution. Although many studies on CCN pointed out this pattern and gave some qualitative analyses, such as boundary layer height (BLH) [13,19,21,25,27–29], pollutant emissions [11,13,16–19,21,24–30], aerosol types [10,12,21,27], and photochemical processes [11,13,24,25,27], the quantitative analyses of influencing factors are rare. Taking BLH as an example, only Kim et al. [13] and Wang et al. [18] in the above studies used the micro-pulse lidar data for qualitative verification, and others just analyzed the theoretical feasibility. This is largely limited by the lack of observational data. At present, reanalysis data have been widely used in climate studies [35–38], and it is feasible to combine the reanalysis data with long-term CCN observation to analyze the influencing factors and their respective impacts on N\textsubscript{CCN} variations.

In addition, the parameterization of N\textsubscript{CCN} is also a focus of previous research. The Twomey formula is one of the most widely used parameterizations for determining N\textsubscript{CCN} through SS, namely, $N\textsubscript{CCN} = cSS^k$ [39]. Therefore, the determination of c and k in this parameterization scheme is an important objective as well. Some studies fitted parameters for a certain season by using single-year observation data. For example, Hudson [40] and Martins et al. [41] determined the parameters of autumn in America and Brazil, respectively. More generally, however, researchers could only obtain parameters for a short-term period (typically for days) because of the limitation of data, especially in China [18,24,25,27,29]. The lack of the parameterization scheme derived from long-term CCN data hinders the further application in GCM and large-scale studies.

Beginning in the 11th FYP (Five-Year Plan) in 2006–2010, the Chinese government set a goal of controlling air pollution and put efforts in reducing the sulfate emissions from coal combustion by utilizing new techniques of desulfurization. In September 2013, the “Air Pollution Prevention and Control Action Plan” was released and became the most influential policy for improving China’s environment in the subsequent decade. Correspondingly, the air quality in North China has been continuously improved since 2010, as shown from multiple surface measurements [42,43] or satellite datasets [44,45]. Therefore, we left out the CCN data after 2010 as our study focused on CCN in the heavily polluted period. Noting that 2006 was a year with severe pollution [46], the observation data from 2005 to 2007 were therefore selected for application in this study.
This paper selected multi-season measurements conducted in Shijiazhuang, a polluted megacity, and used the observations in summer and winter during 2005–2007 to analyze the characteristics of CCN. The concurrent data, which includes the Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2) and the ECMWF Reanalysis 5th Generation (ERA5), are also analyzed to identify the factors affecting the diurnal variation of \( N_{\text{CCN}} \), and build the relationship between \( N_{\text{CCN}} \) and SS in the heavily polluted region. This paper is organized as follows. The next section describes the data and methodology, followed by the results illustrated in Section 3. The conclusion and discussion are given in Section 4.

2. Data and Methodology
2.1. Station Data

Shijiazhuang (38°02′ N, 114°31′ E), the capital of Hebei, China, was selected as the studying site, which is a megacity in North China and located at one of the polluted regions in the world [47]. Here, we utilized the following observed data, including the concentrations of CCN, pollutants (PM\(_{10}\), NO\(_2\) and SO\(_2\)), and meteorological variables (temperature and wind speed), to study the characteristics of CCN in the heavily polluted region. The location of the study area and the observed periods are shown in Figure 1. CCN was observed at the Hebei Meteorological Service Station (38°02′ N, 114°32′ E) (OBS_S for short) from 2005 to 2007 (see all the CCN measurement periods in the caption of Table 1). The observation site of pollutants was located at the Environmental Monitoring Center (38°03′ N, 114°32′ E) (OBS_E) and was about 3 km away from the OBS_S, where the mass concentrations of PM\(_{10}\), NO\(_2\) and SO\(_2\) were measured from 28 February 2007–12 March 2007. To facilitate comparative analysis, we divided all sampling periods by season (panel d in Figure 1 or column “Time” in Table 1), where “winter” and “summer” are roughly defined as in the period of “December–January–February” and “June–July–August”, respectively, and allowing a half-month extension of the period for including as much data as possible. Meteorological data were obtained in the Meteorological Observation Station (38°02′ N, 114°25′ E) (OBS_M), approximately 9 km away from the OBS_S station, where surface temperature and wind speed at 10 m height were available during all the CCN measurement periods.

![Figure 1](image-url) 

**Figure 1.** The location of the study area and the observed periods in this study. (a) The location of the Beijing-Tianjin-Hebei region (red box) in the world map; (b) the location of Shijiazhuang in the Beijing-Tianjin-Hebei region and its topographic map. (c) Sampling stations (colored points, OBS_S: Hebei Meteorological Service Station; OBS_E: Environmental Monitoring Center; OBS_M: Meteorological Observation Station), as well as the extraction area (blue box) for obtaining the reanalysis data; (d) the observed periods include summer 2005, winter 2006, summer 2006, winter 2007, and summer 2007 (detailed time is listed in Table 1).
Table 1. The detailed information for variables used in this study.

| Variable | Description                               | Unit     | Data Source | Time                          |
|----------|-------------------------------------------|----------|-------------|-------------------------------|
| $M_{PM10}$ | Mass concentration of PM$_{10}$          | $\mu g$ m$^{-3}$ | OBS_E       | 28 February 2007–12 March 2007 (winter 2007) |
| $M_{SO2}$  | Mass concentration of SO$_2$              | $\mu g$ m$^{-3}$ | OBS_E       |                               |
| $M_{NO2}$  | Mass concentration of NO$_2$              | $\mu g$ m$^{-3}$ | OBS_E       |                               |
| $N_{CCN}$ | The number concentration of CCN           | cm$^{-3}$ | OBS_S       |                               |
| $M_{Dust}$ | Surface mass concentration of dust        | kg m$^{-3}$ | MERRA-2     | 24 June 2005–9 August 2005 (summer 2005), 1 January 2006–19 January 2006 (winter 2006), 21 May 2006–19 August 2006 (summer 2006), and 24 May 2007–30 June 2007 (summer 2007) |
| $C_{SO2}$  | Column mass concentration of SO$_2$       | kg m$^{-2}$ | MERRA-2     |                               |
| $C_{O3}$   | Total column concentration of ozone       | DU       | MERRA-2     |                               |
| $C_{BC}$   | Column mass concentration of black carbon | kg m$^{-2}$ | MERRA-2, OBS-M, ERA5 |                               |
| $T_{M}$    | Temperature                               | K        | MERRA-2, OBS_M, ERA5 |                               |
| $T_{O}$    | Temperature                               | K        | OBS-M, ERA5 |                               |
| $T_{E}$    | Temperature                               | K        | ERA5        |                               |
| $WS_{10M}$ | 10 m wind speed                           | m s$^{-1}$ | MERRA-2, OBS_M | 28 February 2007–12 March 2007 (winter 2007), and 24 May 2007–30 June 2007 (summer 2007) |
| $WS_{10O}$ | 10 m wind speed                           | m s$^{-1}$ | OBS-M, ERA5 |                               |
| $WS_{10E}$ | 10 m wind speed                           | m s$^{-1}$ | ERA5        |                               |
| $RH$       | Relative humidity                         | %        | ERA5        |                               |
| $WS$       | Surface wind speed                        | m s$^{-1}$ | MERRA-2     |                               |
| $BLH$      | Boundary layer height                     | m        | ERA5        |                               |
| $BLD$      | Boundary layer dissipation                | J m$^{-2}$ | ERA5        |                               |
| $CAPE$     | Convective available potential energy     | J kg$^{-1}$ | ERA5        |                               |
| $CIN$      | Convective inhibition                     | J kg$^{-1}$ | ERA5        |                               |

In the OBS_S site, a CCN Counter, model CCN-100 from Droplet Measurement Technologies (DMT; Boulder, CO, USA), was used to measure the $N_{CCN}$ at 4 different SS levels (0.1%, 0.3%, 0.5% and 1.0%) by utilizing the difference between the heat and moisture diffusion coefficients [48]. The theoretical error of CCN concentration observed by this CCN Counter is no more than 10% [49]. To ensure the accuracy of the CCN data, the CCN Counter was calibrated by the manufacturers every year before observations [14]. The CCN data have a temporal resolution of 1 s and were processed by hourly averaging before analysis. In the OBS_E and OBS_M sites, pollutants (e.g., the mass concentrations of PM$_{10}$ ($M_{PM10}$), NO$_2$ ($M_{NO2}$) and SO$_2$ ($M_{SO2}$)) and meteorological elements (the surface temperature ($T_{O}$) and 10 m wind speed ($WS_{10M}$)) data were provided, respectively. Both of them have a temporal resolution of 1 h. See Table 1 for more details.

2.2. Reanalysis Data

To analyze the factors affecting the long-term CCN variation more comprehensively, we needed corresponding climatological data to assist. Here, MERRA-2 and ERA5 data of Shijiazhuang region (38°00’ N–38°03’ N, 114°24’ E–114°33’ E) were extracted. MERRA-2 is the latest atmospheric reanalysis of the modern satellite era produced by NASA’s Global Modeling and Assimilation Office (GMAO) [50], which was often used to analyze aerosol characteristics and aerosol-climate interaction and relations [51–53]. MERRA-2 data have been proved to be reliable in representing the long-term temporal and spatial distribution of pollutants (i.e., PM$_{2.5}$, PM$_{10}$ sulfate, organic carbon, and black carbon) in China [54–56]. We downloaded the pollutants and meteorological data from Giovanni (https://giovanni.gsfc.nasa.gov/giovanni/, accessed on 21 January 2020), including the data of surface mass concentration of dust ($M_{Dust}$), the column mass concentration of SO$_2$ ($C_{SO2}$), the total column concentration of ozone ($C_{O3}$), the column mass concentration of black carbon ($C_{BC}$), the surface temperature ($T_{M}$), the 10-m wind speed ($WS_{10M}$), and the surface wind speed (WS). ERA5 performs decently in characterizing meteorological conditions (i.e., temperature, wind speed, and precipitation) for climate studies [57,58]. Therefore, ERA5 data (https://cds.climate.copernicus.eu/cdsapp#!/home, accessed on 8 January 2020) were used to analyze the effect of meteorological conditions on CCN variation in this study, we extracted the data of surface temperature ($T_{E}$), the relative
humidity (RH), the 10-m wind speed (WS10E), the boundary layer height (BLH), the boundary layer dissipation (BLD), the convective available potential energy (CAPE), and the convective inhibition (CIN). Both MERRA-2 and ERA5 data are processed as hourly data, and more information about them had shown in Table 1.

2.3. Methodology

We verified the consistency between MERRA-2 and ERA5 data with available observational data at the OBS-M site, e.g., WS10O for verifying 10 m wind speeds WS10M and WS10E, and T0 for verifying surface temperature TM and TE. The result showed that the Pearson correlation coefficient (r) values of WS10M and WS10E with WS10O are 0.677 and 0.591, respectively. For the temperature, the r values are 0.963 and 0.940, respectively. All results passed the consistency test with a 95% confidence level, indicating a reasonable consistency between the reanalysis data and the observations. Therefore, we used both MERRA-2 and ERA5 data to carry out analyses when the station data were not available.

To highlight the variation of each variable, we mainly used the normalized changes of a variable, as defined below:

\[ V = \frac{V_t - V_{\text{MIN}}}{V_{\text{MAX}} - V_{\text{MIN}}} \]  

(1)

where V (V can be any individual variable normalized in a specified period in Table 1, e.g., NCCN in winter 2007) means the time series of the normalized change of \( V_t \), which is the value of a variable at a certain observational time point t (e.g., measured NCCN at one day in winter 2007), and \( V_{\text{MAX}} \) and \( V_{\text{MIN}} \) represent the maximum and minimum values of the variable during the whole observational period (e.g., maximum and minimum values of NCCN in winter 2007). The time series of normalized data can highlight the relative variations of variables, reduce the complexity of graphics, and avoid the dimensional conversion, making our analysis process more concise.

We used the Pearson correlation coefficient (r) to characterize the correlation (linear correlation) between different variables [59]. It is worthy of noting that the use of the normalized method does not change the r values between different variables, ensuring the reliability of our correlation analysis.

3. Results

3.1. Diurnal Variation of NCCN

Considering the NCCN data with SS = 0.3 % had more valid data samples, we selected the NCCN under this supersaturation to analyze the characteristics of diurnal variation. The absolute variations of NCCN in five measurement periods are shown in Figure 2a. In addition, to highlight the characteristics of the diurnal variation, the normalized NCCN in summer and winter are plotted in Figure 2b,c. Overall, the diurnal variations of NCCN in different periods show bimodal distributions. In detail, there are some differences in NCCN daily series between winter and summer. On one hand, the intensities of the two peaks are not consistent. The peak in the morning is more prominent for summer, but the peak in the evening is stronger for winter. On the other hand, the times hitting peak points in different seasons are also different. In summer, the double peaks appear around 09:00 Local Standard Time (LST) in the morning and 22:00 LST in the evening, whereas the peaks appear around 08:00 LST in the morning and 18:00 LST in the evening in winter. The above results show that the diurnal variations of NCCN in winter and summer might be influenced by different factors. We analyze the possible influencing factors in more detail in Section 3.3.
peaks appear around 08:00 LST in the morning and 18:00 LST in the evening.

**3.2. Diurnal Variation of $N_{CCN}$ in Winter**

Pollutants in the atmosphere have an important effect on the aerosol ability for activating to form cloud droplets. Zhang et al. [60] pointed out that aerosol activity is linked not only to particle composition but also to mixing state, aging, and physical processes (such as coating and condensation). The impact of highly soluble pollutants, such as gas precursors such as $SO_2$ and $NO_2$, can affect the formation of hygroscopic matter, and influence particle hygroscopicity [61]. The more $SO_2$ and $NO_2$ coating in the aerosol, the more hygroscopic the particles are and the easier the particles to serve as CCN. Here, we obtained the hourly-averaged concentrations of pollutants ($NO_2$, $PM_{10}$, and $SO_2$) from the available station data in winter 2007 and made use of them to identify the possible reasons for the two peaks occurring in the diurnal variations of $N_{CCN}$. As shown in Figure 3, $N_{CCN}$ is highly consistent with the concentrations of pollutants, especially for $NO_2$ and $PM_{10}$. The $r$ values of $M_{NO_2}$ and $M_{PM_{10}}$ with $N_{CCN}$ are 0.789 and 0.771, and all pass the consistency test with a confidence of 99%. In terms of a.m. and p.m. periods, the positive relationship is apparent for $SO_2$ ($r = 0.797$) at a.m., and for $NO_2$ ($r = 0.933$) at p.m., suggesting that the dominant factor affecting the diurnal variation of $N_{CCN}$ is time-dependent. In the morning, $SO_2$ plays a critical role, but the role of $NO_2$ is more significant in the afternoon and becomes a dominant factor in the evening. $PM_{10}$ and $N_{CCN}$ always maintain a high positive correlation all day, which indicates that the total aerosol particle concentration plays a decisive role in the $N_{CCN}$ variation. Note that the data analyzed in this sub-section only reflect winter 2007. The pollutant of $SO_2$ is mainly emitted from the heating facilities in North China [62,63], and the frequently occurring temperature inversion in winter evenings may accumulate the pollutant to reach a peak in the morning. The pollutant of $NO_2$ is mainly from the traffic emission in the daytime [64], and likely accumulates to reach a peak in the evening.
3.3. Factors Affecting the Diurnal Variation of N_{CCN}

The diurnal variations of N_{CCN} are related to the factors affecting the aerosol activation, for instance, the physical and chemical characteristics of the aerosol, and the meteorological conditions. To quantify the effects of these factors, we plotted their variations and calculated the correlation r between N_{CCN} and the following variables: pollutants (C_{SO2}, C_{O3}, C_{BC}, and M_{Dust}), meteorological variables (T_{O}, WS10, WS and RH), and atmospheric stability variables (BLH, BLD, CAPE, and CIN), as shown in Figure 4 and Table 2. Here, pollutant and atmospheric stability data were sourced from the MERRA-2 and ERA5, respectively, and meteorological data were obtained from on-site observation and reanalysis data. The atmospheric stability variables are used to represent the environmental conditions for pollutants (emitted from the Earth’s surface) to diffuse away from the emission sources. BLH and BLD characterize the atmospheric state under tiny turbulence, and CAPE and CIN correspond to stronger atmospheric disturbances (i.e., convection). Note that CIN values are available only in summer. In general, pollutants will be diffused and the concentration near the source region will be diluted with high BLH, BLD, CAPE, and low CIN.

**Table 2.** The importance ranking of the factors affecting N_{CCN} diurnal variation according to the absolute value of correlation r for summer and winter during 2005–2007. Results that pass the 90% confidence test are indicated in bold. The superscripts of ***, **, and * present the correlations passing the consistency test with the significance levels of 99%, 95% and 90%, respectively.

| Order | Summer Variable | r     | p-Value | Winter Variable | r     | p-Value |
|-------|-----------------|-------|---------|-----------------|-------|---------|
| 1     | C_{SO2}         | 0.745 | ***     | BLH             | -0.714 | ***     |
| 2     | C_{BC}          | 0.694 | ***     | BLD             | -0.677 | ***     |
| 3     | WS              | -0.641 | ***     | C_{SO2}        | 0.607  | ***     |
| 4     | M_{Dust}        | 0.531 | ***     | WS10           | -0.530 | ***     |
| 5     | WS10           | -0.516 | ***     | WS              | 0.481  | **      |
| 6     | BLD             | -0.492 | **      | C_{O3}         | 0.455  | **      |
| 7     | CAPE            | -0.490 | **      | M_{Dust}       | 0.350  | *       |
| 8     | T_{O}           | -0.471 | **      | RH              | 0.329  | 0.1167  |
| 9     | RH              | 0.431 | **      | T_{O}           | -0.295 | 0.1612  |
| 10    | BLH             | -0.289 |       | CAPE           | -0.145 | 0.4990  |
| 11    | C_{O3}          | 0.178  |        | C_{BC}         | 0.049  | 0.8215  |
| 12    | CIN             | -0.095 |        | CIN            | –      | –       |
Figure 4. The hourly-averaged diurnal variations (normalized) of (a,d) pollutants (\(\text{SO}_2\), \(\text{CO}_3\), \(\text{BC}\) and \(\text{M}_{\text{Dust}}\)), (b,e) meteorological variables (\(T_0\), \(\text{WS}_{10}\), \(\text{WS}\) and \(\text{RH}\)), (c,f) atmospheric stability variables (BLH, BLD, CAPE and CIN), and \(\text{N}_{\text{CCN}}\) in summer (top panels) and winter (bottom panels) during 2005 and 2007. The \(r\) value represents the strength of correlation between the individual variable and \(\text{N}_{\text{CCN}}\) is shown in the bracket. The superscripts of ***, **, and * present the correlations passing the consistency test with the significance levels of 99\%, 95\% and 90\%, respectively.

The hourly-averaged time series of different variables and \(\text{N}_{\text{CCN}}\) in winter and summer are shown in Figure 4. In summer, the diurnal variation of pollutants and \(\text{N}_{\text{CCN}}\) has a positive correlation. Among them, \(\text{SO}_2\) has the best correlation, followed by BC and dust, and finally \(\text{O}_3\) (Figure 4a). In terms of meteorological variables, temperature and wind speed are negatively correlated while RH is positively correlated to \(\text{N}_{\text{CCN}}\). Wind speed (surface and 10 m wind speed) has the strongest correlation, followed by surface temperature and RH (Figure 4b). All atmospheric stability variables have a negative correlation with \(\text{N}_{\text{CCN}}\) in the diurnal variation, of which the strongest correlation is BLD, followed by CAPE, and finally BLH and CIN (Figure 4c). In winter, the positive and negative correlations between different variables and \(\text{N}_{\text{CCN}}\) in the diurnal variation are generally consistent with those in summer, but the strength of the correlation (i.e., \(r\)) is different. For the pollutants, the diurnal variation of \(\text{SO}_2\) still has the strongest impact on \(\text{N}_{\text{CCN}}\), followed by \(\text{O}_3\) and dust, and the least impacting pollutant is BC (Figure 4d). Wind speed is still the most important factor affecting \(\text{N}_{\text{CCN}}\) among meteorological variables (Figure 4e). However, the correlation between surface wind speed and \(\text{N}_{\text{CCN}}\) is negative in summer but positive in winter, which will be discussed later. There are strong negative correlation relationships between \(\text{N}_{\text{CCN}}\) and atmospheric stability variables, especially for BLH and BLD (Figure 4f).

As mentioned above, \(\text{WS}_{10}\) (wind speed at 10 m height) and \(\text{N}_{\text{CCN}}\) always show a negative correlation, but WS (surface wind speed) and \(\text{N}_{\text{CCN}}\) have different correlations in different seasons. We explain the phenomenon as follows. Wind speed is regarded as an important factor affecting the source and sink of aerosol and pollutants. \(\text{WS}_{10}\) is higher from the ground and is mostly determined by the synoptic meteorology or the large-scale air circulation. The larger \(\text{WS}_{10}\) usually indicates stronger horizontal air motions and favors the diffusion of pollutants, resulting in a lower \(\text{N}_{\text{CCN}}\). However, WS is strongly
affected by the surface properties such as vegetation types, soil grains, surface roughness, etc., and is most likely determined by the local topography, and small- to middle-scale weather situations. Therefore, the relationship between WS and N\textsubscript{CCN} is prone to change with seasons.

To analyze the main influencing factors of N\textsubscript{CCN} in winter and summer, we sorted the variables according to the absolute values of \(r\), as shown in Figure 5. In this chord diagram, the absolute values of \(r\) denote the correlations between the corresponding variables and N\textsubscript{CCN} connected by the belt. The wider belt indicates the stronger correlation between the individual variable and N\textsubscript{CCN}. For all the variables listed in Table 2, only those variables with correlations passing the test of 90% confidence level (i.e., \(p\)-value is less than 0.1) were selected to show in Figure 5. In summer, we see the variations in pollutants of C\textsubscript{SO2} and C\textsubscript{BC} have the greatest impact on that in N\textsubscript{CCN}, followed by meteorological conditions of wind speed (WS and WS\textsubscript{10}), and finally atmospheric stability of BLD and CAPE. In winter, the factor that has the greatest impact on N\textsubscript{CCN} variation is atmospheric stability of BLH and BLD, followed by the pollutant of C\textsubscript{SO2}, and the relatively minor factor is the meteorological condition of wind speed (WS\textsubscript{10} and WS).

Figure 5. A chord diagram to display the correlation relationships (represented with the absolute value of \(r\)) between pollutants (black arcs of the circumference; C\textsubscript{SO2}, C\textsubscript{O3}, C\textsubscript{BC} and M\textsubscript{Dust}), meteorological variables (dark grey arcs; T\textsubscript{O}, WS\textsubscript{10}, WS, and RH) and atmospheric stability variables (grey arcs; BLH, BLD and CAPE) and N\textsubscript{CCN} in summer (red arc) and winter (blue arc). The absolute values of \(r\) (numbers shown along the upper half circumference) denote the correlations between the corresponding variables and N\textsubscript{CCN} connected by the red/blue belt. The wider belt indicates the stronger correlation between the individual variable and N\textsubscript{CCN}.

Based on the above results, we summarize that high N\textsubscript{CCN} in Shijiazhuang is most likely related to high pollutant concentration and atmospheric stability, and low 10 m wind speed. In summer, the concentration of pollutants is relatively low, so the variations of pollutants have relatively strong impacts on N\textsubscript{CCN} variation. In winter, however, the concentration of the pollutant is relatively high, so it has less impact on the change of N\textsubscript{CCN}, but the variation of atmospheric stability plays a more important role.
3.4. \(N_{CCN}\)-SS Relationship

An overview of hourly-averaged CCN number concentration is shown in Figure 6. Overall, \(N_{CCN}\) keeps a high level, up to 14,026 cm\(^{-3}\) under SS = 1% in winter, and \(N_{CCN}\) is a little higher in winter than summer, especially at high SS. \(N_{CCN}\) has an increasing tendency with SS in both seasons in Shijiazhuang.

![Figure 6](image)

**Figure 6.** An overview of hourly–averaged \(N_{CCN}\). All the data are divided into five groups that present different observation periods. Each group contains four SS levels, namely 0.1%, 0.3%, 0.5% and 1.0%. “N” means the number of samples in different SS. On each box, the central mark (red line) indicates the median, and the bottom and top edges (blue line) of the box indicate the 25th and 75th percentiles, respectively. The outliers are plotted individually using the point symbol.

Many empirical formulas have been proposed to obtain \(N_{CCN}\) from the supersaturation (SS). A power-law equation proposed by Twomey [39] is most commonly used, namely \(N_{CCN} = cSS^k\). The parameters of \(c\) and \(k\) contain information concerning the particle number concentration and the chemical characteristics of aerosols. We fitted the Twomey formula and obtained the fitting parameters in different periods of this study, as shown in Figure 7. When fitting, we first took the logarithm of both sides of the Twomey formula, that is, \(\ln(N_{CCN}) = k\ln(SS) + \ln(c)\). Next, we used the QR method [65] to linearly fit the above equation (function `lm` in R language) using the observed data points to obtain \(k\) and \(\ln(c)\). The R-squared values were obtained when conducting linear fitting and high \(R^2\) denoted a better performance of the fitted curve in representing the observed \(N_{CCN}\)-SS relationship (Figure 7). The values of \(k\) in winter (1.08 and 0.69) are larger than those in summer (0.52, 0.49 and 0.58). Therefore, the fitting curves for winter are steeper than the ones for summer (Figure 7). The fitting parameter \(c\) can be regarded as a proxy to represent the particle number concentration. Obviously, the values are higher in winter (22,510 and 15,766) than those in summer (12,676, 14,116 and 13,428). The parameter \(c\) of Shijiazhuang maintaining a high level (\(c \geq 10,000\)) in both winter and summer shows the characteristic of a typical continental CCN (\(c \geq 2200, k < 1\)) [66]. However, we noticed that the fitting parameters in winter 2006 are significantly higher than the results in other years in North China [18,30]. This may be attributed to the highly frequent and severe pollution events in winter 2006 in Shijiazhuang. Qu et al. [46] analyzed the air pollution condition in Shijiazhuang based on the Air Quality Index (AQI). The results showed that there were 5 days with high-polluted events (AQI > 300) and 21 days with middle-polluted events (200 < AQI ≤ 300) from 2005 to 2009, and all of the high-polluted events and almost half of middle-polluted events occurred in winter and spring of 2006. Therefore, we think the observations in winter 2007 are more representative of the general air pollution condition in the winter of Shijiazhuang, and the fitting parameters in winter 2006 are more representative of a state of extremely heavy pollution. At the same time, the above results also remind us that it is necessary to consider using larger \(c\) and \(k\) values while using \(N_{CCN} = cSS^k\) in heavily polluted areas.
Figure 7. The relationship between $N_{CCN}$ and SS in (a) winter 2006, (b) winter 2007, (c) summer 2005, (d) summer 2006 and (e) summer 2007. The fitting formulas and the corresponding R-squared ($R^2$) values are marked on the upper left of each graph. Stars and the solid curve in each panel represent the observed data points and the fitted curve of $N_{CCN}$~SS relationship, with the SS ranging from 0 to 1%. The superscript, *** indicates that the fitting parameters pass the consistency test under the significance levels of 99%.

The comparison of parameters $c$ and $k$ derived from this study with parameters from different regions (continental or transitional between marine and continental) is shown in Table 3. We found that the value of $c$ in Shijiazhuang is about one order of magnitude higher than others, which may be related to the high pollution level in Shijiazhuang. Although there is an obvious difference between $c$ parameters in Shijiazhuang and other studies in China, the values of $k$ are similar, as shown in Table 3. Qi and Yao [30] studied the CCN characteristics at Qingdao, China in winter and found that the $k$ values were 0.51~0.72 at different weather conditions. Wang et al. [18] showed that the values of $k$ were 0.49~0.69 for different wind conditions in autumn of Tianjin, China. Based on the above results, we summarize the recommended values of parameters $c$ and $k$ to construct the fitting formula ($N_{CCN} = cSS^k$) when predicting $N_{CCN}$ in East and North China for a large-scale study. The parameter $c$ can be obtained from the average value of $N_{CCN}$ at SS = 1.0% if the observation data are available. Otherwise, $c$ within the range of 2000 to 15,000 is regarded as a general case, but $c > 15,000$ should be regarded as the extremely heavy pollution case. The recommended value of parameter $k$ is 0.5~0.7 (0.5 for summer and 0.7 for winter) in East and North China. During the extremely polluted situation, $k$ needs to be further increased, and it can be selected between 0.7 and 1.0.

Table 3. Comparison of parameters $c$ and $k$ from multiple studies. The type of CCN is determined according to Hobbs et al. [66].
4. Conclusions and Discussion

Studying CCN in the atmosphere is of great significance for understanding aerosol–cloud interactions. This study used multi-year data observed in Shijiazhuang from 2005 to 2007 and the reanalysis data (MERRA-2 and ERA5) to investigate the characteristics of CCN, sort the factors affecting the diurnal variation of \( N_{CCN} \) according to their importance, and build the relationship between \( N_{CCN} \) and SS in the heavily polluted region. Our comprehensive analysis based on CCN long-term observations combined with climate-scale reanalysis data can provide a reference for climatological CCN research analysis or model simulation.

The diurnal variations of \( N_{CCN} \) in different periods show a bimodal distribution, and the peak in the morning is more prominent for summer, but the evening peak is obvious for winter. By analyzing the observation data measured in winter 2007, we found the dominant factor affecting this pattern is the pollutant \( \text{SO}_2 \) in the morning but the pollutant \( \text{NO}_2 \) in the evening, and \( \text{PM}_{10} \) always maintains a high positive correlation with \( N_{CCN} \) all day. Additionally, we ranked the importance of factors affecting \( N_{CCN} \) diurnal variation by using observation and reanalysis data according to the absolute values of the Pearson correlation coefficient. The variation of pollutants (\( C_{\text{SO}_2} \) and \( C_{\text{BC}} \)) has the greatest impact on \( N_{CCN} \) in summer, but the variation of atmospheric stability (\( \text{BLH} \) and \( \text{BLD} \)) in winter. To study the relationship between \( N_{CCN} \) and SS, we fitted the Twomey formula (\( N_{CCN} = c_{SS}^k \)) and found the CCN has the typical continental characteristic in Shijiazhuang. According to this study and some previous research, we recommended a reasonable value range (0.5–0.7) of the parameter \( k \) for applying in East and North China. In detail, it is more reasonable for \( k \) to be 0.5 in summer and 0.7 in winter. When extreme pollution (\( c > 15,000 \)) occurs, the value of \( k \) can be further increased, recommending a value between 0.7 and 1.0.

Several points are worthy of mentioning for future studies. Although reanalysis data have great convenience in acquisition and application, these data still have great limitations in the characterization of small-scale regional pollution and meteorological conditions. Therefore, observations of pollutants and meteorological variables on a local scale should be carried out together with \( N_{CCN} \) observations in the future, so as to further improve the understanding of the complex relationship between CCN and pollutants and meteorological conditions. In addition, the \( N_{CCN} \) data used in this study are representative for 2005–2007 only. In recent years, China has achieved remarkable success in pollution control [68], and its impacts on CCN and aerosol–cloud interaction are of interest to study in the future.

Supplementary Materials: The data to produce the key plots (Figures 2–4 and 6) are available online at https://www.mdpi.com/article/10.3390/atmos13030468/s1.

Author Contributions: H.W. processed data, conducted analyses, and wrote the manuscript. M.Z. helped with the processing and analyses of data. Y.P. and J.D. led the work, instructed the analyses, and refined the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the National Important Project of the Ministry of Science and Technology in China (grant no. 2017YFC1301404), the S&T Development Fund of CAMS (2022KJ010), and the National Natural Science Foundation of China (grant nos. 42175096, 71690243, and 41675137).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data to produce the key plots Figures 2–4 and 6) have been uploaded as the Supplementary Materials.

Acknowledgments: The authors thank MERRA-2 and ERA5 teams for making the dataset publicly available. In addition, we would like to thank three anonymous reviewers for their constructive suggestions to improve the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.
29. Zhao, Y.; Niu, S.; Lv, J.; Xu, J.; Sang, J. Observational Analyses on Cloud Condensation Nuclei in Northwestern China in Summer of 2007. *Plateau Meteorol.* 2010, 29, 1043–1049.
30. Qi, J.; Yao, X. Observation Analysis of Cloud Condensation Nuclei over Qingdao in Winter. *Period. Ocean Univ.* 2019, 48, 10–19.
31. Che, H.C.; Zhang, X.Y.; Wang, Y.Q.; Zhang, L.; Shen, X.J.; Zhang, Y.M.; Ma, Q.L.; Sun, J.Y.; Zhang, Y.W.; Wang, T.T. Characterization and parameterization of aerosol cloud condensation nuclei activation under different pollution conditions. *Sci. Rep.* 2016, 6, 24497. [CrossRef]
32. Jurányi, Z.; Gysel, M.; Weingartner, E.; Bukowiecki, N.; Kammermann, L.; Baltensperger, U. A 17 month climatology of the cloud condensation nuclei number concentration at the high alpine site Jungfraujoch. *J. Geophys. Res. Atmos.* 2011, 116, D10204. [CrossRef]
33. Pöhler, M.L.; Pöhler, C.; Ditas, F.; Klimach, T.; De Angelis, I.H.; Araújo, A.; Brito, J.; Carbone, S.; Cheng, Y.; Chi, X.; et al. Long-term observations of cloud condensation nuclei in the Amazon rain forest—Part 1: Aerosol size distribution, hygroscopicity, and new model parameterizations for CCN formation. *Atmos. Chem. Phys.* 2016, 16, 15709–15740. [CrossRef]
34. Schmale, J.; Henning, S.; Decesari, S.; Henzing, B.; Keskinen, H.; Sellegri, K.; Ovdanevaite, J.; Pöhler, M.; Brito, J.; Bougiatioti, A.; et al. Long-term cloud condensation nuclei number concentration, particle number size distribution and chemical composition measurements at regionally representative observatories. *Atmos. Chem. Phys.* 2018, 18, 2853–2881. [CrossRef]
35. Frank, C.W.; Wahi, S.; Keller, J.D.; Pospichal, B.; Hense, A.; Crewell, S. Bias correction of a novel European reanalysis data set for solar energy applications. *Sol. Energy* 2018, 164, 12–24. [CrossRef]
36. Jiang, H.; Yang, Y.; Bai, Y.; Wang, H. Evaluation of the Total, Direct, and Diffuse Solar Radiations from the ERA5 Reanalysis Data in China. *IEEE Geosci. Remote Sens. Lett.* 2020, 17, 47–51. [CrossRef]
37. Mahto, S.S.; Mishra, V. Does ERA-5 Outperform Other Reanalysis Products for Hydrologic Applications in India? *J. Geophys. Res. Atmos.* 2019, 124, 9423–9441. [CrossRef]
38. Chen, S.; Gan, T.Y.; Tan, X.; Shao, D.; Zhu, J. Assessment of CFSR, ERA-Interim, JRA-55, MERRA-2, NCEP-2 reanalysis data for drought analysis over China. *Clim. Dyn.* 2019, 53, 737–757. [CrossRef]
39. Twomey, S. The nuclei of natural cloud formation part II: The supersaturation in natural clouds and the variation of cloud droplet concentration. *Geophys. Pur. Appl.* 1959, 43, 243–249. [CrossRef]
40. Hudson, J.G. Relationship between fog condensation nuclei and fog microstructure. *J. Atmos. Sci.* 1959, 16, 1854–1867. [CrossRef]
41. Martins, J.A.; Gonçalves, F.L.T.; Morales, C.A.; Fisch, G.F.; Pinheiro, F.G.M.; Leal, B.V., Jr.; Oliveira, C.J.; Silva, E.M.; Oliveira, J.C.P.; Costa, A.A.; et al. Cloud condensation nuclei from biomass burning during the Amazonian dry-to-wet transition season. *Meteorol. Atmos. Phys.* 2009, 104, 83–93. [CrossRef]
42. Huang, J.; Pan, X.; Guo, X.; Li, G. Health impact of China’s Air Pollution Prevention and Control Action Plan: An analysis of national air quality monitoring and mortality data. *Lancet Planet. Health* 2018, 2, e131–e323. [CrossRef]
43. Cai, S.; Wang, Y.; Zhao, B.; Wang, S.; Chang, X.; Hao, J. The impact of the “Air Pollution Prevention and Control Action Plan” on PM2.5 concentrations in Jing-Jin-Ji region during 2012–2020. *Sci. Total Environ.* 2017, 580, 197–209. [CrossRef] [PubMed]
44. Sogacheva, L.; De Leeuw, G.; Rodriguez, E.; Kolmonen, P.; Georgoulias, A.K.; Alexandri, G.; Kourtidis, K.; Proestakis, E.; Marinou, E.; Amiridis, V.; et al. Spatial and seasonal variations of aerosols over China from two decades of multi-satellite observations—Part 1: ATSR (1995–2011) and MODIS C6.1 (2000–2017). *Atmos. Chem. Phys.* 2018, 18, 11389–11407. [CrossRef]
45. Wei, J.; Peng, Y.; Mahmood, R.; Sun, L.; Guo, J. Intercomparison in spatial distributions and temporal trends derived from multi-source satellite aerosol products. *Atmos. Chem. Phys.* 2019, 19, 7183–7207. [CrossRef]
46. Qu, X.; Fu, G.; Jia, J.; Yang, G. Distribution characteristics of air quality and its relationship with meteorological factors from 2005 to 2009 in Shijiazhuang, Hebei province. *J. Meteorol. Environ.* 2011, 27, 29–32.
47. Xu, W.Y.; Zhao, C.S.; Ran, L.; Deng, Z.Z.; Liu, P.F.; Ma, N.; Lin, W.L.; Xu, X.B.; Yan, P.; He, X.; et al. Characteristics of pollutants and their correlation to meteorological conditions at a suburban site in the North China Plain. *Atmos. Chem. Phys.* 2011, 11, 4353–4369. [CrossRef]
48. Roberts, G.C.; Nenes, A. A continuous-flow streamwise thermal-gradient CCN chamber for atmospheric measurements. *Aerosol Sci. Technol.* 2005, 39, 206–221. [CrossRef]
49. Ward, D.S.; Cotton, W.R. Cold and transition season cloud condensation nuclei measurements in western Colorado. *Atmos. Chem. Phys.* 2011, 11, 4303–4317. [CrossRef]
50. Gelaro, R.; McCarty, W.; Suárez, M.J.; Todling, R.; Molod, A.; Takacs, L.; Randles, C.A.; Darmenov, A.; Bosilovich, M.G.; Reichle, R.; et al. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Clim.* 2017, 30, 5419–5454. [CrossRef]
51. Xing, Z.; Li, S.; Xiong, Y.; Du, K. Estimation of cross-boundary aerosol flux over the Edmonton- Calgary Corridor in Canada based on CALIPSO and MERRA-2 data during 2011–2017. *Atmos. Environ.* 2021, 246, 118084. [CrossRef]
52. Sun, E.; Che, H.; Xu, X.; Wang, Z.; Lu, C.; Gui, K.; Zhao, H.; Zheng, Y.; Wang, Y.; Wang, H.; et al. Variation in MERRA-2 aerosol optical depth over the Yangtze River Delta from 1980 to 2016. *Theor. Appl. Climatol.* 2019, 136, 363–375. [CrossRef]
53. McCoy, D.T.; Bender, F.A.-M.; Mohrmann, J.K.C.; Hartmann, D.L.; Wood, R.; Grosvenor, D.P. The global aerosol-cloud first indirect effect estimated using MODIS, MERRA, and AeroCom. *J. Geophys. Res. Atmos.* 2017, 122, 1779–1796. [CrossRef]
54. Zhao, Q.; Zhao, W.; Bi, J.; Ma, Z. Climatology and calibration of MERRA-2 PM2.5 components over China. *Atmos. Pollut. Res.* 2021, 12, 357–366. [CrossRef]
55. Ma, X.; Yan, P.; Zhao, T.; Jia, X.; Jiao, J.; Ma, Q.; Wu, D.; Shu, Z.; Sun, X.; Habtemicheal, B.A. Article evaluations of surface 
PM10 concentration and chemical compositions in MERRA-2 aerosol reanalysis over central and eastern China. Remote Sens. 2021, 
13, 1317. [CrossRef][PubMed]
56. Qin, W.; Zhang, Y.; Chen, J.; Yu, Q.; Cheng, S.; Li, W.; Liu, X.; Tian, H. Variation, sources and historical trend of black carbon 
in Beijing, China based on ground observation and MERRA-2 reanalysis data. Environ. Pollut. 2019, 245, 853–863. [CrossRef]
57. Wang, Q.; Li, W.; Xiao, C.; Ai, W. Evaluation of high-resolution crop model meteorological forcing datasets at regional scale: Air 
temperature and precipitation over major land areas of China. Atmosphere 2020, 11, 1011. [CrossRef][PubMed]
58. Jiang, Y.; Han, S.; Shi, C.; Gao, T.; Zhen, H.; Liu, X. Evaluation of HRCLDAS and ERA5 Datasets for Near-Surface Wind over 
Hainan Island and South China Sea. Atmosphere 2021, 12, 766. [CrossRef][PubMed]
59. Benesty, J.; Chen, J.; Huang, Y.; Speech, I.C.-N. Pearson Correlation Coefficient. In Noise Reduction in Speech Processing; Springer: 
Berlin/Heidelberg, Germany, 2009.
60. Zhang, F.; Wang, Y.; Peng, J.; Ren, J.; Collins, D.; Zhang, R.; Sun, Y.; Yang, X.; Li, Z. Uncertainty in Predicting CCN Activity of 
Aged and Primary Aerosols. J. Geophys. Res. Atmos. 2017, 122, 11723–11736. [CrossRef]
61. Wang, Y.; Zhang, F.; Li, Z.; Tan, H.; Xu, H.; Ren, J.; Zhao, J.; Du, W.; Sun, Y. Enhanced hydrophobicity and volatility of submicron 
aerosols under severe emission control conditions in Beijing. Atmos. Chem. Phys. 2017, 17, 5239–5251. [CrossRef]
62. Duan, X.; Jiang, Y.; Wang, B.; Zhao, X.; Shen, G.; Cao, S.; Huang, N.; Qian, Y.; Chen, Y.; Wang, L. Household fuel use for 
cooking and heating in China: Results from the first Chinese Environmental Exposure-Related Human Activity Patterns Survey 
(CEERHAPS). Appl. Energy 2014, 136, 692–703. [CrossRef]
63. Chai, L.; Ma, C.; Ni, J.Q. Performance evaluation of ground source heat pump system for greenhouse heating in northern China. 
Biosyst. Eng. 2012, 111, 107–117. [CrossRef]
64. Liu, H.; He, K.; Barth, M. Traffic and emission simulation in China based on statistical methodology. Atmos. Environ. 2011, 45, 
1154–1161. [CrossRef]
65. Kublanovskaya, V.N. On some algorithms for the solution of the complete eigenvalue problem. USSR Comput. Math. Math. Phys. 
1962, 1, 637–657. [CrossRef]
66. Hobbs, P.V.; Bowdle, D.A.; Radke, L.F. Particles in the lower troposphere over the high plains of the United States. Part II: Cloud 
condensation nuclei. J. Clim. Appl. Meteorol. 1985, 24, 1358–1369. [CrossRef][PubMed]
67. Miao, Q.; Zhang, Z.; Li, Y.; Qin, X.; Xu, B.; Yuan, Y.; Gao, Z. Measurement of cloud condensation nuclei (CCN) and CCN closure 
at Mt. Huang based on hygroscopic growth factors and aerosol number-size distribution. Atmos. Environ. 2015, 113, 127–134. 
[CrossRef][PubMed]
68. Dou, X.; Liao, C.; Wang, H.; Huang, Y.; Tu, Y.; Huang, X.; Peng, Y.; Zhu, B.; Tan, J.; Deng, Z.; et al. Estimates of daily ground-level 
NO2 concentrations in China based on Random Forest model integrated K-means. Adv. Appl. Energy 2021, 2, 100017. [CrossRef]