



A Coupled Data Assimilation Method for Visibility and Relative Humidity CAOP v1.0 module within the WRFDA v4.3 framework: Improving both PM_{2.5} and Humidity Forecasts

Lina Gao¹², Junli Jin¹², Zhiqian Liu³, Wei You⁴, Xiang Xie^{1,5}, Peng Yan¹

¹Meteorological Observation Center, China Meteorological Administration, Beijing, 100081, China

²State Key Laboratory of Environment Characteristics and Effects for Near-space, Beijing, 100081, China

³National Meteorological Information Centre, China Meteorological Administration, Beijing, 100081, China

⁴College of Meteorology and Oceanography, National University of Defense Technology, Changsha, 410073, China

⁵Aletai Meteorological Bureau, Aletai, 836500, China

Correspondence to: Junli Jin (jinjl@cma.gov.cn) and Zhiqian Liu (liuzhq@cma.gov.cn)

Abstract. High precision prediction of near-surface PM_{2.5} mass concentration is critical for effective air quality management and mitigating health risks. While Data assimilation (DA) of the aerosol mass concentration or the aerosol optical properties (AOP) observation optimizes the initial conditions of an atmospheric chemistry prediction model and improves aerosol prediction skill, the potential of assimilating surface visibility remains constrained by the overlooked regulatory effect of relative humidity (RH) on visibility variance. To address this, we developed a coupled DA method of AOP and RH (CAOP v1.0) within the WRFDA (version 4.3) framework. Four parallel experiments were conducted across mainland China in December 2020: a control experiment and three DA experiments. Results demonstrate that assimilating visibility alone (VisOnly) reduced RMSE by 21% for visibility and 9% for PM_{2.5}, validating the CAOP module. Furthermore, DA of both visibility and RH -- whether performed simultaneously but independently in the VisRH_Indp experiment or via the coupled scheme in the VisRH_Copl experiment -- effectively corrected the PM_{2.5} analysis bias originating from background RH deviations. The coupled scheme (VisRH_Copl) achieved superior performance, lowering PM_{2.5} RMSE by 17% compared to 16% in the independent run (VisRH_Indp). Crucially, the CAOP scheme more efficiently leveraged visibility data to constrain the RH field, improving the RH forecast correlation coefficient by 25% (versus 23% in the VisRH_Indp). This study confirms that visibility observations



31 substantially enhanced RH analysis and demonstrated that the coupled DA method outperforms
32 independent DA approaches for joint PM_{2.5} and RH forecasting.

33 **1. Introduction**

34 Aerosols are closely linked to human health and the environment. Particulate matter with a diameter
35 $< 2.5 \mu\text{m}$ (PM_{2.5}) is one of the most important air pollutants, which is capable of being inhaled into the
36 lungs, thereby degrading air quality, decreasing visibility and damaging human health and safety (Cohen
37 et al., 2017). Severe air pollution was caused by rapid urbanization and industrialization during the first
38 decade of the 21st century in China. To address this issue, a series of clean air projects and action plans
39 were executed by the Chinese government (Jing et al., 2025). The weighted annual mean PM_{2.5}
40 concentration declined from $61.8 \mu\text{g}/\text{m}^3$ in 2013 to $42.0 \mu\text{g}/\text{m}^3$ in 2019 (Zhang et al., 2019).
41 However, the declining rate of PM_{2.5} has slowed down, from $-4.5 \mu\text{g}/\text{m}^3/\text{yr}$ before 2017 to -2.3
42 $\mu\text{g}/\text{m}^3/\text{yr}$ after 2018 (Geng et al., 2024). Nevertheless, aerosol pollution episodes remain relatively
43 prevalent during the autumn and winter seasons (Gao et al., 2025b). Therefore, accurate prediction of
44 near surface PM_{2.5} concentration and in-depth exploration of its spatial and temporal variation are crucial
45 in guiding refined management of air pollution and control, as well as safeguarding human health and
46 safety.

47 With the rapid development of ground and satellite observation networks, an increasing amount of
48 observations are being utilized to optimize the chemical initial conditions, thereby upgrading the
49 numerical model's aerosol prediction skills (Sogacheva et al., 2020; Wang et al., 2022; Wang et al., 2022;
50 Wang et al., 2026). However, traditional ground-based air quality monitoring stations are unevenly
51 distributed, resulting in a lack of observation data in sparsely populated areas. Furthermore, aerosol
52 products retrieved from satellite remote sensing are constrained by the visible spectrum, leading to data
53 gaps during nighttime, in areas with persistent cloud cover or with high pollution levels. Therefore, there
54 is an urgent need for more high-precision ground-based observations to fill these gaps. Atmospheric
55 visibility is defined as the maximum distance at which objects can be distinctly perceived (Hyslop, 2009).
56 It is a key metric that characterizes the transparency and cleanliness of the atmosphere. Its variation is
57 primarily governed by dry aerosol mass concentration, particle size distribution, and relative humidity
58 (RH), collectively reflecting the direct pollution level of aerosols and their hygroscopic growth and



59 extinction effects under specific environmental conditions (Zhao et al., 2022). In recent years, with the
60 improvement of atmospheric environmental monitoring networks and the development of numerical
61 forecast techniques, visibility data assimilation (DA) or retrieval has become an important approach for
62 optimizing the initial aerosol fields and enhancing the accuracy of fog-haze forecasts (Kim et al., 2020;
63 Zhang et al., 2024). There are several mainstream directions for the assimilation of the aerosol optical
64 properties (AOP) observation, i.e. visibility. Previous studies have shown that incorporating the visibility
65 observation into DA systems can effectively constrain aerosol mass concentration in models and improve
66 the forecast performance of visibility and aerosol mass concentration (Mei et al., 2025; Gao et al., 2025a).
67 However, existing visibility DA schemes exhibit significant limitations in their theoretical foundations
68 and technical pathways, restricting their effectiveness in complex high-humidity or high-pollution
69 scenarios.

70 The first approach incorporates the aerosol hygroscopic growth scheme into the DA module,
71 trusting the RH from the model's background field. It assimilates the AOP observation (typically
72 excluding high-humidity conditions) to constrain the dry aerosol mass concentration. Liu et al. (2011)
73 assimilated aerosol optical depth (AOD) within the National Centers for Environmental Prediction
74 Gridpoint Statistical Interpolation 3DVAR DA system. The Community Radiative Transfer Model
75 (CRTM) was extended to calculate AOD using aerosol profiles and RH as input in the forward operator.
76 Water uptake effects of sulfate, sea salts, and hydrophilic OC and BC were included, but there was no
77 explicit expression of RH in the Jacobian term (Chin et al., 2002). Using a similar method, Bao et al.
78 (2019) assessed the impact of Chinese FY-3/MERSI AOD DA on air quality forecasts during sand and
79 dust events. Due to the oversimplification of aerosol chemical species and hygroscopic growth schemes,
80 there was bias between the AOP analysis and the observation. Gao et al. (2025a) assimilated the visibility
81 observation via the WRFDA-Chem (Weather Research & Forecasting Data Assimilation system coupled
82 with Chemistry version 4.3) model with the AOP DA module. The AOP DA module incorporated a more
83 comprehensive range of aerosol chemical species and a more detailed aerosol hygroscopic process. The
84 results showed that the root mean square error (RMSE) of $PM_{2.5}$ during winter in China was significantly
85 improved. However, it also yielded an obvious $PM_{2.5}$ bias after visibility-only DA, which may be related
86 to uncertainties of RH simulation or the hygroscopic process. There is inevitably a large uncertainty in
87 RH simulation under complex weather conditions, such as dust storms and fogs. The DA system may



88 mistakenly attribute and compensate for the forward simulation error originally caused by RH error by
89 adjusting aerosol mass concentration. This may introduce analysis bias and reduce assimilation efficiency.

90 The second approach attempts to circumvent the complexity of humidity influence by converting
91 AOP observations (e.g., visibility, AOD, aerosol extinction coefficient (AExtC)) into "dry aerosols"
92 through empirical relationships, then directly assimilating the dry aerosol to correct the aerosol
93 concentration field. Xiang et al. (2021) obtained the vertical distribution of $PM_{2.5}$ in the boundary layer
94 by combining the near-ground AExtC with in-situ $PM_{2.5}$ surface observations using the empirical formula
95 fitting method. Both the correlation coefficient (R) and RMSE for aerosol mass concentration were
96 significantly improved. This study assumes that within the same air mass or during the same pollution
97 episode, the aerosol chemical-physical-optical properties are relatively uniform in the vertical direction.
98 Thus, the mass extinction efficiency per unit mass does not vary significantly. Uncertainty is introduced
99 in situations where the vertical structure of polluted air is complex or there is a large RH gradient. Shuang
100 et al. (2025) converted visibility to AExtC based on the Koschmieder formula under dry conditions and
101 Rosenfeld et al. (2007) under high RH conditions. It turned out that visibility assimilation can
102 significantly improve the model's simulation and prediction of aerosol mass concentration, especially in
103 inland areas where satellite AOD data are missing. Their study demonstrated the value of visibility
104 observations as a necessary supplementary data source for aerosol forecasts. However, the applicability
105 of this formula was not verified for different geographical environments, aerosol types and climatic zones.
106 Besides, the aerosol hygroscopicity process was highly simplified, without consideration of either aerosol
107 chemical species or size distribution.

108 Although these two strategies have their respective advantages, their fundamental common flaw
109 lies in treating the adjustment of the aerosol concentration field and the humidity field as two independent
110 or sequential processes, failing to achieve their joint optimization. Haywood et al. (2008) and Clark et al.
111 (2008) developed a coupled visibility DA module, in which aerosol mixing ratio and RH were optimized
112 simultaneously. The decrease of visibility was caused by dry AExtC and aerosol hygroscopic growth,
113 and the coupled DA method contributed significantly to low visibility forecasts. However, due to extreme
114 simplified aerosol chemical species (one species only) and aerosol hygroscopic scheme (sulfate only),
115 the performance of the model was limited. For example, the AExtC was overestimated by 1.5-2.0 times
116 when RH was greater than 80%. Kim et al. (2020) conducted automatic visibility observation DA using



117 the same framework as Haywood and Clark, including localization of murky aerosol parameters and
118 optimization of the cycling DA strategy. Their work significantly enhanced the fog and low visibility
119 forecast skills. This lightweight and efficient model was suitable for low visibility and pollution scenarios,
120 but it performed poorly in a clean atmosphere, failing to meet the demand for high-precision aerosol and
121 air quality prediction.

122 Numerous observational studies indicate a close physical relationship between visibility and
123 relative humidity (Jin et al., 2022). Under high RH conditions (above 80%), RH and aerosol hygroscopic
124 properties are important factors affecting visibility. On one hand, the hygroscopic growth effect of aerosol
125 particle is significant under a high RH environment. Aerosol particle size, mass, and extinction efficiency
126 can increase nonlinearly with minute changes in humidity, making visibility extremely sensitive to
127 humidity variations (Chen et al., 2014). On the other hand, the observation error of the humidity-sensitive
128 capacitance sensor is larger in higher RH ranges ($RH > 80\%$) than that in low RH ranges ($RH < 80\%$) (Liu
129 et al., 2022). Under high RH conditions, a tiny variation of RH can largely influence visibility variation
130 (Chen et al., 2012). An increasing RH of 1% at $RH < 90\%$ results in a change of AExtC less than 10%,
131 whereas an increasing 1% at $RH > 95\%$ results in a change about 20%. Therefore, visibility could be a
132 more precise measurement of the impact of RH as uncertainties of RH measurement are usually large
133 under high RH conditions. This indicates that through visibility DA, it is expected to improve the analysis
134 and forecasts of RH under high RH conditions. In other words, during high-humidity, low-visibility
135 events, visibility observations effectively "encapsulate" precious information about both aerosol loads
136 and humidity conditions, which existing methods cannot fully and jointly integrate into the analysis field.

137 Based on the above understanding, this study aims to break through the limitations of existing
138 visibility assimilation schemes. The core innovation involves developing and validating a novel coupled
139 DA operator for visibility and RH. The outcomes of this research are expected to provide new technical
140 tools for high-precision aerosol and fog-haze analysis and forecast, promoting the deeper application of
141 DA technology in the atmospheric environment field. The structure of this paper is as follows: Section 2
142 will detail the numerical model, physical basis and operator construction of the coupled DA method,
143 together with the observational data, and experimental design. Section 3 will present the improvement
144 effects of coupled DA on the forecast field and will thoroughly demonstrate the superiority of the method
145 through comparative experiments. Section 4 will summarize the full text and discuss future research



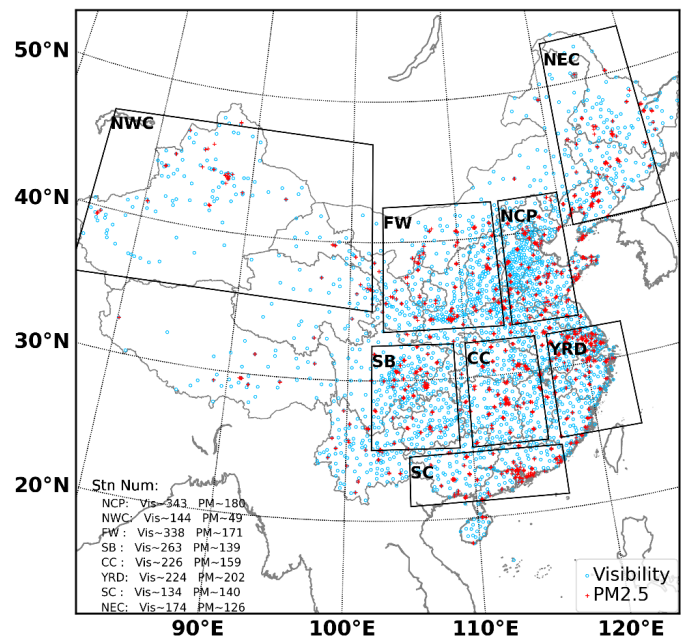
146 directions.

147 **2. Model, observations, and data assimilation methodology**

148 The WRF-Chem model (version 4.0) was used to predict aerosol chemical species in this study, and
149 the WRFDA-Chem three-dimensional (3D) variational model was used to perform RH and/or visibility
150 data assimilation (Gao et al., 2025a).

151 **2.1 The WRF-Chem model and emissions**

152 The WRF-Chem was used to simulate the aerosol physical and chemical processes. The physical
153 and chemical schemes are shown in Table 1. The simulation region covered most of China with a
154 horizontal resolution of 40.5 km (Fig. 1), and there were 57 vertical levels extending from the surface to
155 50 hPa.



156
157 **Figure 1: Computational domain.** Blue circle denotes the location of the visibility observation. And red plus
158 denotes the location of the PM_{2.5} observation. Eight black rectangles are as follows: North East China (NEC),
159 North China Plain (NCP), Fen Wei (FW), North West China (NWC), Yangtze River Delta (YRD), Central
160 China (CC), Sichuan Basin (SB), and South China (SC). The number of visibility and PM_{2.5} stations at 8



161 regions are listed in the lower left corner.

162 The multi-resolution Emission Inventory for China (MEIC) with a horizontal resolution of $0.25^\circ \times$
163 0.25° for base year 2016 was adopted as the anthropogenic emission input (Li et al., 2014). The temporal
164 mismatch could introduce systematic biases in the absolute magnitude of aerosol concentration
165 simulations. This would primarily affect the bias of the baseline run but would not necessarily invalidate
166 the relative differences between the assimilation experiments. As we focus on aerosol forecast
167 improvement by DA rather than absolute aerosol forecast itself, the base year differences of the
168 anthropogenic emissions were neglected. Biogenic, sea salt, and dust emissions were calculated online.

169 **Table 1.** Model configuration of WRF-Chem model

Scheme	Configuration
Microphysics	WSM 6-class graupel scheme (Hong et al., 2006)
Short/Long wave radiation	RRTMG (Iacono et al., 2008)
Planetary boundary layer	YSU scheme (Hong et al., 2006)
Land surface	Noah scheme
Aerosol	MOSAIC 4 BIN scheme
Gas-phase chemistry	CBM-Z
Meteorological initial and boundary conditions	6 hourly GFS forecasts with spatial resolution $0.5^\circ \times 0.5^\circ$
Grid nudging	No
Initial condition for chemical species	The forecast results from previous prediction or WRFDA-chem analysis
Boundary condition for chemical species	Averages of mid-latitude aircrafts profiles
Biogenic emissions	Online Gunther Scheme
Seasalt emissions	Online Gocart seasalt emission
Dust emission	Online Gocart dust emission with MOSAIC aerosol scheme (Zeng et al., 2020)
Aerosol radiation feedback	On

170 2.2 Coupled DA of AOP and RH (CAOP) method

171 There were 32 control variables in the WRFDA-Chem AOP observation operator, including black
172 carbon BC, organic carbon OC, sulfate SO_4 , nitrate NO_3 , ammonium NH_4 , sodium, chloride, and other
173 inorganic compounds OIN in 4 size bins ranging from 0.039-0.1, 0.1-1.0, 1.0-2.5 and 2.5-10 μm (Gao et
174 al., 2025a).

175 Based on this module, RH is incorporated as one of the control variables within the AOP framework.
176 We abbreviate this newly coupled AOP module as CAOP (version 1.0) hereafter. RH and aerosol



chemical species are connected mainly through aerosol liquid water content (ALWC), and ALWC contributed by inorganic species is determined by the following Eq. (1) (Petters and Kreidenweis, 2007).

$$ALWC_{inorganic} = V_i \kappa_i \frac{a_w}{1-a_w} \rho_w \quad (1)$$

V_i represents the volume concentration of each chemical component. κ_i is the hygroscopic parameter of each chemical component. a_w is the water activity, which is assumed to be equivalent to RH for simplicity. ρ_w is the density of water. And the κ_i values for NO_3 , SO_4 , NH_4 are the weighted averages of the mole fractions of the following chemical species, including NH_4NO_3 , $(NH_4)_2SO_4$, NH_4HSO_4 and H_2SO_4 . These chemical components are reconstructed through thermodynamic equilibrium of model chemical species, including NO_3 , SO_4 , NH_4 . The volume concentration of these chemical components was calculated based on the ion-pairing scheme using the following equations (Gysel et al., 2007; Jin et al., 2020). The hygroscopic parameter for NH_4NO_3 , $(NH_4)_2SO_4$, NH_4HSO_4 , H_2SO_4 , and NaCl are 0.58, 0.48, 0.56, 1.19 and 1.16 (Petters and Kreidenweis, 2007). The κ_i for OIN and BC are 0.14 and 1×10^{-6} , respectively (Saide et al., 2020).

$$\left\{ \begin{array}{l} n_{NH_4NO_3} = n_{NO_3^-} \\ n_{NH_4HSO_4} = \min(2n_{SO_4^{2-}} - n_{NH_4^+} + n_{NO_3^-}, n_{NH_4^+} - n_{NO_3^-}) \\ n_{(NH_4)_2SO_4} = \max(n_{NH_4^+} - n_{NO_3^-} - n_{SO_4^{2-}}, 0) \\ n_{H_2SO_4} = \max(n_{SO_4^{2-}} - n_{NH_4^+} + n_{NO_3^-}, 0) \\ n_{H_2SO_4} = \max(n_{SO_4^{2-}} - n_{NH_4^+} + n_{NO_3^-}, 0) \\ n_{HNO_3} = 0 \end{array} \right. \quad (2)$$

In eq. (2), n represents the number of moles, and the “max” and “min” denote the maximum and minimum values, respectively. The aerosol chemical species Cl and Na in the model are treated as a single species, NaCl, for simplicity. The mass concentrations of Cl and Na are 1-2 orders of magnitude lower than those of other inorganic water-soluble components over mainland China in the model, resulting in a negligible uncertainty in the ALWC calculation. Thus, the ion-pairing scheme for the NaCl is given by Eq. (3) for simplicity.

$$n_{NaCl} = (n_{Cl} + n_{Cl})/2. \quad (3)$$

The ALWC contributed by the OC is also determined by Eq. (1) (Jin et al., 2020), where the hygroscopic parameter κ for the OC is assumed to be 0.1 in this study (Wu et al., 2020; Nguyen et al., 2016). The maximum value of RH is set to be 95% in this study to ensure the calculated ALWC is within a reasonable range. This truncation was implemented for two main reasons. First, under near-saturated conditions (RH > 95%), the hygroscopic growth factor becomes asymptotically sensitive to minute



203 fluctuations in both RH measurements and temperature. Minor instrumental uncertainties would
204 propagate into disproportionately large errors in ALWC estimates. Second, the RH regime approaching
205 and exceeding 95% is physically associated with the activation of cloud/fog droplets, where the
206 fundamental assumptions of sub-saturated hygroscopic growth parameterizations are no longer strictly
207 valid.

208 Under the WRFDA-Chem AOP framework, we have developed the non-linear, tangent-linear, and
209 adjoint modules for the aforementioned process, which explicitly links ALWC to aerosol chemical
210 species. Accordingly, the tangent-linear and adjoint modules of this relationship can be directly expressed.
211 Whether to use the CAOP function can be set via the namelist. The lower RH limit for activating the
212 constraint through visibility DA is set to 80% in this study, which can also be set via the namelist.
213 Similarly, the lower visibility limit for activating RH constraint through visibility DA is set to be 1 km.
214 The limiting conditions of $RH > 80\%$ and visibility $< 1\text{ km}$ here can ensure that visibility data during foggy
215 weather, rather than sand/dust weather, is used to constrain and optimize RH.

216 Comparisons are made between the newly developed ALWC scheme and the original ALWC
217 scheme (ZSR method) from the WRF-Chem model (Barnard et al., 2010). Aerosol extinction coefficients
218 are calculated using both the CAOP method and the ZSR method every six hours during December 2020
219 over China (Fig. 2). The two methods yield highly consistent results, with the bias nearly 0 km^{-1} , the root
220 mean square error (RMSE) 0.05 km^{-1} , and the correlation coefficient (R) 0.995. These statistical results
221 imply that the new method enables precise calculation of ALWC across varying RH conditions,
222 particularly for aerosols with an aerosol extinction coefficient below 3 km^{-1} . The purpose of this
223 comparison is primarily diagnostic, confirming that the ALWC calculation scheme in AOP is correctly
224 implemented. The practical advantage of CAOP over original ZSR method is that its formulation is fully
225 differentiable, enabling the construction of tangent-linear and adjoint operators required by variational
226 DA, and it avoids iterative thermodynamic equilibrium calculations, thereby improving computational
227 efficiency.

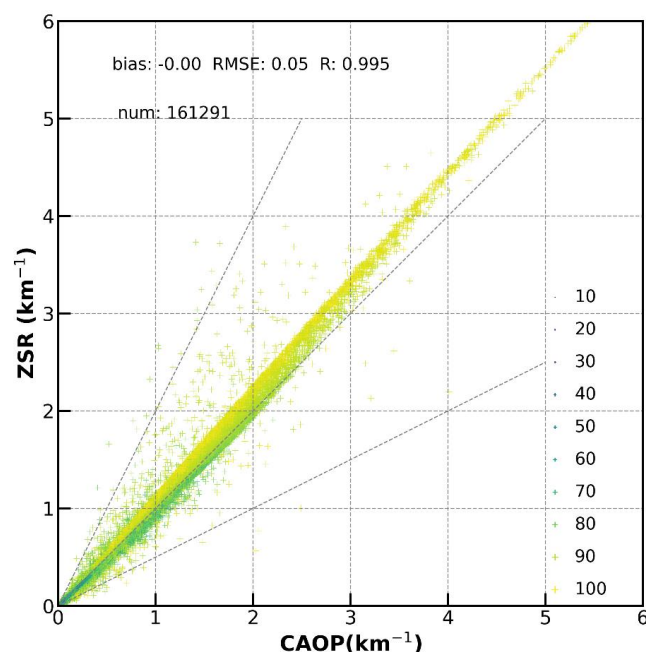


Figure 2: AExtC diagnostic by the CAOP method and the ZSR method in the WRF-Chem model six hourly (00, 06, 12, 18 UTC) from 1-31 December 2020 at 2424 synoptic stations over China.

2.3 Observations

Visibility from forward scattering visibility meter at 2424 synoptic stations in China from 1-31 December 2020 were collected and used in the DA and verification process (blue circles in Fig. 1) (Zhang et al., 2020). The visibility network has relatively uniform coverage in the eastern and central regions but sparser density in the western arid and high-elevation areas.

Visibility observations with precipitation > 0.0 mm were eliminated, as precipitation is not within the focus of this study. Furthermore, visibility observations < 500 m were eliminated. These low visibility events are mostly sandstorms or heavy fog, which are also not the focus of this study. Finally, visibility data were converted to AExtC before being used in DA by the WRFDA-Chem CAOP module according to the following Formula (4) (Koschmieder, 1924).

$$\text{AExtC} = -\frac{\ln \varepsilon}{\text{visibility}} \quad (4)$$

In Formula (4), visibility was calculated assuming an exponential scattering law. ε is the liminal

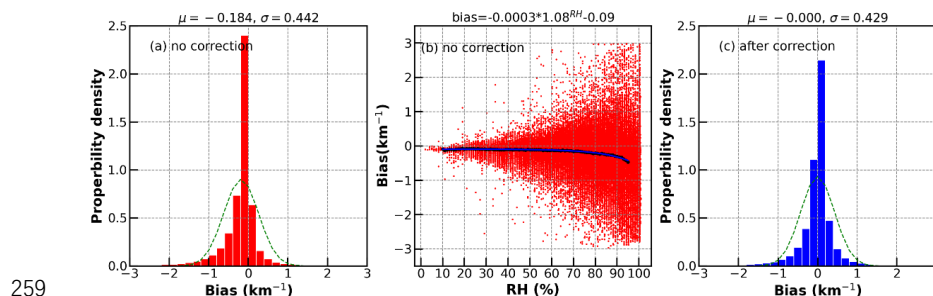


243 contrast within the range 0.02-0.05. ϵ value of 0.05 was adopted for the visibility observation via
244 visibility meter. The unit of AExtC was km^{-1} , with the visibility in unit of km.

245 The bias (model simulation minus observation) between visibility meter AExtC and simulated AExtC
246 was corrected to ensure that unbiased data is assimilated in DA (Gao et al., 2025a). The probability
247 density of the AExtC bias was checked at 2424 synoptic stations from 1-31 December 2020, as shown in
248 Fig. 3(a). The WRF-Chem model control run underestimated the AExtC by a mean value of -0.184 km^{-1} ,
249 which may be due to uncertainty in either the RH simulation or the AOP diagnostic module. The AExtC
250 bias increased with the RH observation (red dots in Fig. 3(b)) and the mean negative AExtC bias also
251 increased with RH (black dots in Fig. 3(b)). Thus, an exponential bias fitting against RH was conducted
252 (blue line in Fig. 3(b)) as follows:

$$253 \text{ bias} = -0.0003 \cdot 1.08^{RH} - 0.09 \quad (5)$$

254 The AExtC used in DA was the sum of the bias and the original/uncorrected AExtC. The corrected
255 AExtC is shown in Fig. 3(c) with the mean bias decreased remarkably from -0.184 km^{-1} to almost 0.0
256 km^{-1} . The corrected AExtC (hereinafter referred to as AExtC for simplicity) was used in DA and for
257 results verification. It's worth clarifying that the bias correction in Eq. (5) is a static calibration of the
258 forward operator's systematic error, not an iterative optimization driven by the assimilation feedback.



259 **Figure 3: The histogram of the AExtC bias before correction (a), the scatter plot of AExtC bias (red dots),**
260 **mean bias (black dots) and fitting line (blue line) against RH observations (in the RH range 10%-95%, b),**
261 **and the histogram of the corrected AExtC bias (c) at 2424 synoptic stations over China from 1-31 December**
262 **2020. μ is the mean bias, and σ is the standard square error.**
263

264 $\text{PM}_{2.5}$ observations at approximately 1400 national urban monitoring stations across China from 1-
265 31 December 2020 were also collected (red pluses in Fig. 1) to verify the visibility DA. The $\text{PM}_{2.5}$
266 network exhibits a pronounced east-west gradient in density. In comparison, the visibility network covers
267 a broader area with a more spatially uniform distribution. The eastern regions (NCP, YRD, FW, CC, SB)



are densely sampled, with 200-340 visibility stations and 140-200 PM_{2.5} stations, supporting robust regional statistics. The lower station density in western regions, such as NWC (144 visibility and 49 PM_{2.5} stations), does not necessarily imply poor regional representativeness. Western China is characterized by relatively homogenous land surface types, sparse population and substantially lower anthropogenic emissions compared to the eastern industrial regions, resulting in a weaker spatial heterogeneity of aerosol concentrations. Consequently, a smaller number of stations may still adequately capture the regional mean state and temporal variability.

2.4 Background error and observation error

The Gen-BE v2.0 was used to calculate the background error covariance for the 32 aerosol chemical composition control variables together with the meteorological control variables u , v , T , RH , and P_s via the background error option CV7 (Sun et al. 2020). The National Meteorological Center (NMC) method was applied to estimate the background error covariance (BE) over the computational domain where data assimilation (DA) was performed based on one-month WRF-Chem forecasts for January 2020 (31 days in total). Both the relative and absolute magnitudes of the background error (BE) standard deviations for the 32 aerosol chemical species control variables are generally consistent with previous studies, as detailed in Gao et al. (2025a).

The observation error for AExtC was set to 0.025 km^{-1} for simplicity. The maximum deviation allowed between observation and model background for AExtC DA was set to 3 km^{-1} . The observation error for RH was the default value for synoptic observation, which is 10%, and the maximum RH deviation allowed for DA was set to 30%.

2.5 Experiments and verifications

Four parallel experiments were conducted, including the control run and three DA runs. No observation was assimilated in the control run (CTR hereafter). The AExtC was assimilated in the visibility- only experiment (VisOnly hereafter). Only aerosol chemical species were optimized in VisOnly. The AExtC and RH were assimilated simultaneously but independently in the VisRH_Indp experiment (VisRH_Indp hereafter). Both RH and aerosol chemical species were optimized in the VisRH_Indp, with the optimization of the RH solely from the synoptic observation. In the VisRH_Copl experiment (VisRH_Copl hereafter), the AExtC and RH were assimilated using the newly developed



296 CAOP. The optimization of RH in VisRH_Copl relied on both the RH and the visibility.

297 **Table 2.** Settings of the four experiments.

Experiment	Visibility DA	RH DA	DA time (UTC)	Description
CTR	NO	No	-	Control run
VisOnly	Yes	No	00,06,12,18	AExtC DA run
VisRH_Indp	Yes	Yes	00,06,12,18	Independent AExtC and RH DA run
VisRH_Copl	Yes	Yes	00,06,12,18	Coupled AExtC and RH DA run

298 In the CTR, the WRF-Chem model conducted a 24-hour forecast at 00:00 UTC every day from 1-
 299 31 December 2020. To ensure the comparability with other three experiments, meteorological initial
 300 conditions (ICs) for the 1st day were from the GFS dataset at 00:00, and the meteorological ICs for the
 301 remaining 30 days were from the previous WRF-Chem forecast. The boundary conditions (BCs) were
 302 from the GFS dataset every day, which were the same with other three DA experiments. The chemical
 303 ICs were from the previous WRF-Chem forecasts.

304 DA were conducted every 6 hours (i.e., 00:00, 06:00, 12:00, and 18:00 UTC) from 1-31 December
 305 2020 in the other three DA experiments. A 24-hour forecast was conducted at 00:00 UTC each day. The
 306 meteorological ICs were the DA analysis at 00:00, when the RH field could be updated by the observation.
 307 The chemical ICs were also the DA analysis at 00:00.

308 Bias, RMSE and R between the model predictions and observations were calculated to evaluate the
 309 contribution of AExtC and/or RH DA on AExtC, PM_{2.5} and RH forecasts. To address the statistical
 310 significance of the difference between these experiments, we estimated the minimum detectable
 311 difference (MDD) at the 95% confidence level ($\alpha = 0.05$), accounting for the paired-comparison design
 312 and the high error correlation expected between the two assimilation methods (Jolliffe and Stephenson,
 313 2012). A difference is considered statistically significant when its absolute value exceeds the
 314 corresponding MDD.

315 **3. Results**

316 The coupled AExtC and RH DA was applied to improve the accuracy of both the RH and chemical
 317 ICs of the WRF-Chem model. Three other parallel runs were also conducted to further illustrate the
 318 performance of CAOP. Twenty-four-hour forecasts were performed at each 00:00 from 1-31 December
 319 2020. The forecast improvements in the AExtC, PM_{2.5}, and RH are evaluated in the following sections.



3.1 Forecast improvement of AExtC

The temporal and station mean bias, RMSE, and R between the model predicted and visibility meter observed AExtC with respect to the 24-hour forecast range are shown in Fig. 4(a1-a3). The mean bias, RMSE, and R for AExtC over China and in 8 regions are shown in Fig. 5(a) and Table 3. The CTR underestimated the AExtC with a mean bias of -0.02 km^{-1} over China, with the mean RMSE and R of 0.46 km^{-1} and 0.19, respectively (Fig. 5(a)). The relatively poor performance of CTR experiment (lower R or higher RMSE) may be attributed to the daily cold-start strategy without DA of meteorological variables, emission uncertainties from the 4-year temporal mismatch between the MEIC 2016 base year and the 2020 simulation period, or the absence of secondary organic formation in the current model configuration.

DA of the visibility meter AExtC observation only (the VisOnly) significantly improved the AExtC underestimation, with a mean bias of 0.03 km^{-1} (as shown in Fig. 4(a)). Absolute values and the time evolution of AExtC also were significantly modified, with the mean RMSE and R improved by 21% and 173%, respectively. The positive performance of the VisOnly is consistent with the previous research results of Gao et al. (2025a) and Mei et al. (2025), in which the ZSR method was used in the ALWC calculation. This further validates the effectiveness of the calculation method for the ALWC in the CAOP module developed in this study.

DA of both the AExtC and the RH also significantly improved the AExtC underestimation, with the mean bias of 0.031 km^{-1} and 0.033 km^{-1} in the VisRH_Copl and VisRH_Indp, respectively. Both experiments achieved improvements in RMSE and R, and the performances of the two experiments were comparable, with reductions of RMSE of 16% for both. Compared to Clark et al.'s (2008) reported <10% RMSE reduction over Eastern Europe (30 Jan-6 Feb 2006), our CAOP method achieves more substantial improvements. The R improvement in the VisRH_Copl and VisRH_Indp were 149% and 148%, respectively. It is worth noting that the performance of these two experiments was not as outstanding as that of the VisOnly, which will be detailed in the following Sect. 3.3.

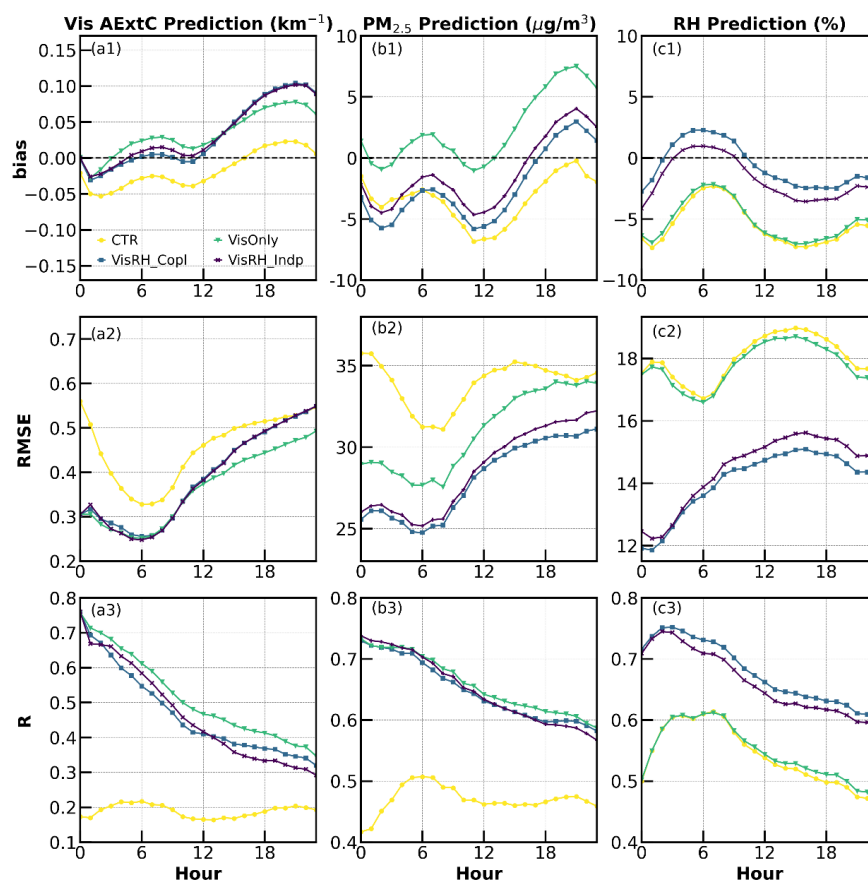
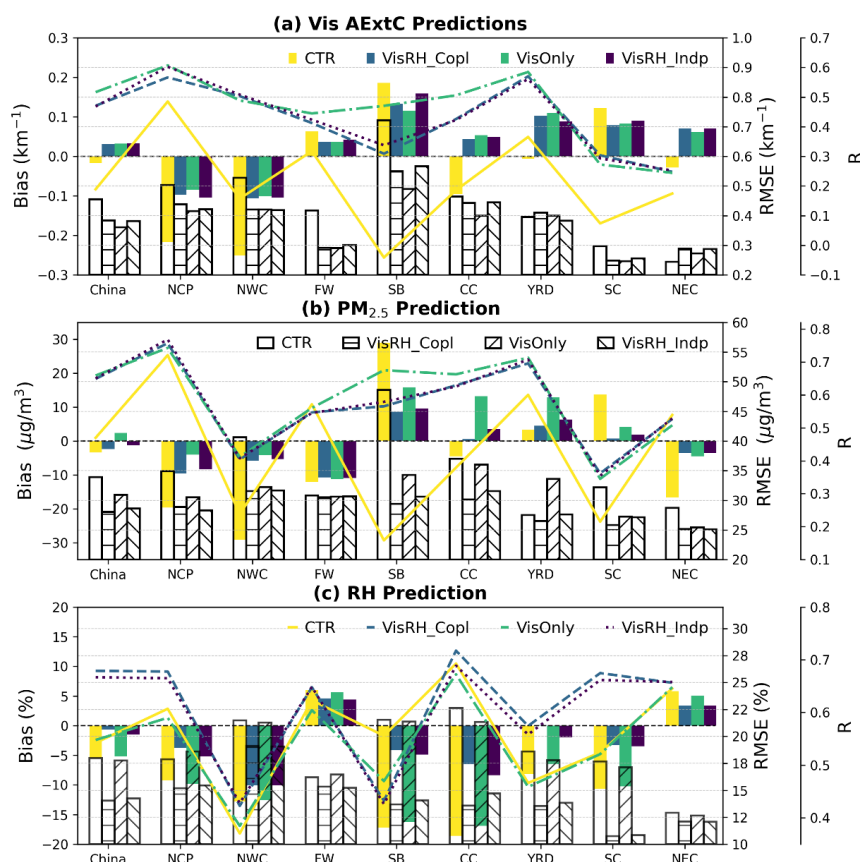


Figure 4: Mean bias, RMSE and R for the AExtC (in unit km^{-1} , a1-a3), the $\text{PM}_{2.5}$ (in unit $\mu\text{g}/\text{m}^3$, b1-b3) and the RH (in unit %, c1-c3) predictions as a function of the 24-hour forecast range for the four experiments (CTR in yellow, VisOnly in green, VisRH_Indp in purple, and VisRH_Copl in blue). The statistical results were verified against the hourly observation over China from 1-31 December 2020. The forecasts with absolute AExtC bias $> 3 \text{ km}^{-1}$ were not used in the verification process to avoid the uncertainty of emissions and or cloud process. The forecasts with absolute $\text{PM}_{2.5}$ bias $> 100 \mu\text{g}/\text{m}^3$ were not used in the verification process. And the forecasts with RH bias $> 40\%$ were not used in the verification process.



354

355 **Figure 5:** Station and temporal mean bias (colored bars), RMSE (hatch bars) and R(lines) values of the hourly
356 AExtC (a), $\text{PM}_{2.5}$ (b) and RH (c) forecasts from the four experiments (CTR in yellow, VisRH_Copl in blue,
357 VisOnly in green and VisRH_Indp in purple) against ground observation over China, and in 8 regions from
358 1-31 December 2020.

359 **Table 3.** Statistic metrics of the hourly AExtC, $\text{PM}_{2.5}$ and RH forecasts from the four experiments against
360 ground observations over China and in 8 regions from 1-31 December 2020.

Region	CTR			VisRH_Copl			VisOnly			VisRH_Indp		
	Bias	RMSE	R	Bias	RMSE	R	Bias	RMSE	R	Bias	RMSE	R
AExtC (km^{-1})												
China	-0.02	0.46	0.19	0.03	0.38	0.47	0.03	0.36	0.52	0.03	0.38	0.47
NCP	-0.22	0.50	0.48	-0.10	0.44	0.57	-0.08	0.41	0.61	-0.1	0.42	0.60
NWC	-0.25	0.53	0.16	-0.11	0.42	0.50	-0.1	0.42	0.49	-0.1	0.42	0.51
FW	0.06	0.42	0.32	0.04	0.29	0.41	0.04	0.29	0.45	0.04	0.30	0.42
SB	0.19	0.72	-0.04	0.13	0.55	0.31	0.12	0.49	0.47	0.16	0.57	0.34



CC	-0.1	0.46	0.19	0.04	0.44	0.43	0.05	0.4	0.51	0.05	0.44	0.42
YRD	-0.01	0.40	0.37	0.10	0.41	0.57	0.11	0.4	0.58	0.09	0.38	0.56
SC	0.12	0.30	0.07	0.08	0.25	0.30	0.08	0.25	0.27	0.09	0.26	0.29
NEC	-0.03	0.24	0.17	0.07	0.29	0.25	0.06	0.27	0.24	0.07	0.29	0.25
PM _{2.5} (µg/m ³)												
China	-3	34	0.47	-2	28	0.65	2	31	0.66	-1	29	0.65
NCP	-20	35	0.72	-10	29	0.76	-4	30	0.74	-8	28	0.77
NWC	-29	41	0.24	-6	32	0.40	-4	32	0.41	-5	32	0.41
FW	-12	31	0.57	-11	30	0.55	-11	31*	0.56	-11	31*	0.55
SB	29	49	0.16	9	29	0.56	16	34	0.68	10	31	0.58
CC	-4	37	0.38	0	30	0.63	13	36	0.66	4	32	0.62
YRD	3	28	0.6	4	26	0.70	13	34	0.71	6	28*	0.71
SC	14	32	0.21	1	26	0.36	4	27	0.34	2	27	0.35
NEC	-17	29	0.54	-4	25	0.53	-4	25	0.51	-4	25	0.53
RH (%)												
China	-6	18	0.54	-1	14	0.68	-5	18	0.55	-2	14	0.67
NCP	-9	18	0.61	-4	15	0.68	-10	19	0.59	-5	16	0.67
NWC	-13	22	0.37	-10	19	0.42	-12	21	0.39	-10	19	0.43
FW	6	16	0.62	5	15	0.65	6	16	0.60	4	15	0.65
SB	-17	22	0.56	-4	14	0.43	-16	21*	0.47	-5	14	0.43
CC	-19	23	0.69	-6	14	0.72	-17	21	0.67	-8	15	0.69
YRD	-8	19	0.47	0	14	0.57	-6	18	0.46	-2	14	0.56
SC	-11	18	0.52	-3	11	0.68	-10	17*	0.52	-4	11	0.66
NEC	6	13	0.65	3	12	0.66	5	13*	0.65	3	12	0.66

Note: * denotes improvements from DA experiments relative to CTR that are not statistically significant at $\alpha=0.05$ (paired MDD, assuming correlation coefficient 0.85 of the pair)

From the perspective of the regional distribution, the CTR significantly underestimated AExtC in regions like NCP, CC, and NWC, and seriously overestimated AExtC in FW and the southwestern part of SB (Fig. 6(a)). In the VisOnly, both the underestimation and the overestimation of the AExtC were generally improved, except for the overestimation in the southwestern part of SB (Fig. 6 (c)). The performance of the VisRH_Copl and VisRH_Indp (Fig. 6(b) and (d)) was comparable to that of the VisOnly. However, the difference lied in the fact that the overestimation of AExtC remained quite significant across the southern SB after assimilating both AExtC and RH. Deviations in either aerosol chemical species or RH may lead to deviation in visibility analysis. To illustrate the reasons behind this phenomenon, forecast improvements of aerosol chemical species (presented by PM_{2.5} mass concentration) and RH were also included in the statistics in the following sections.

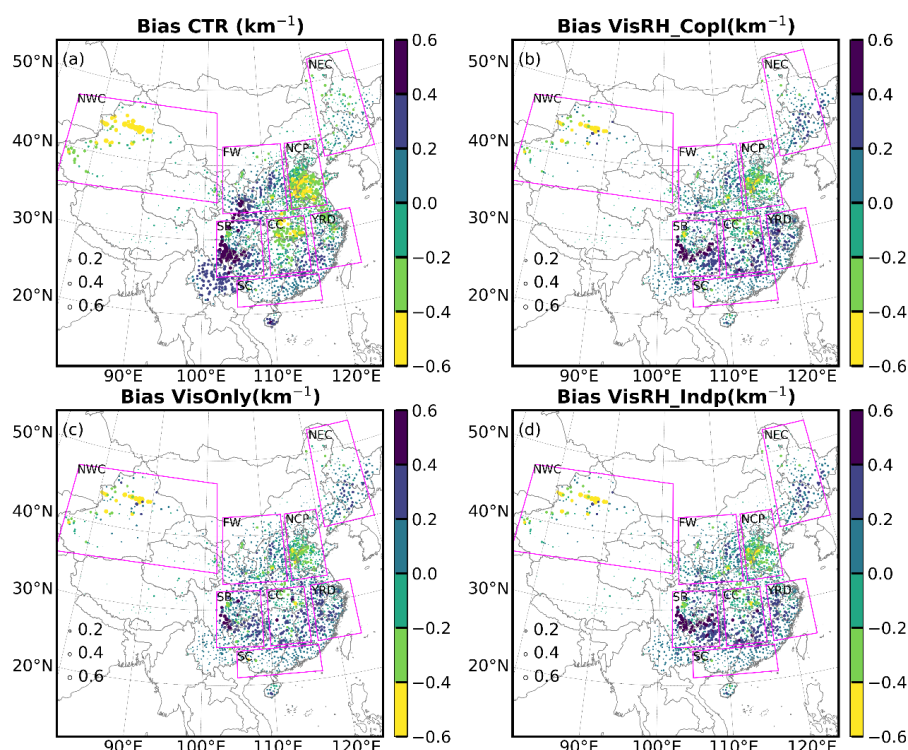


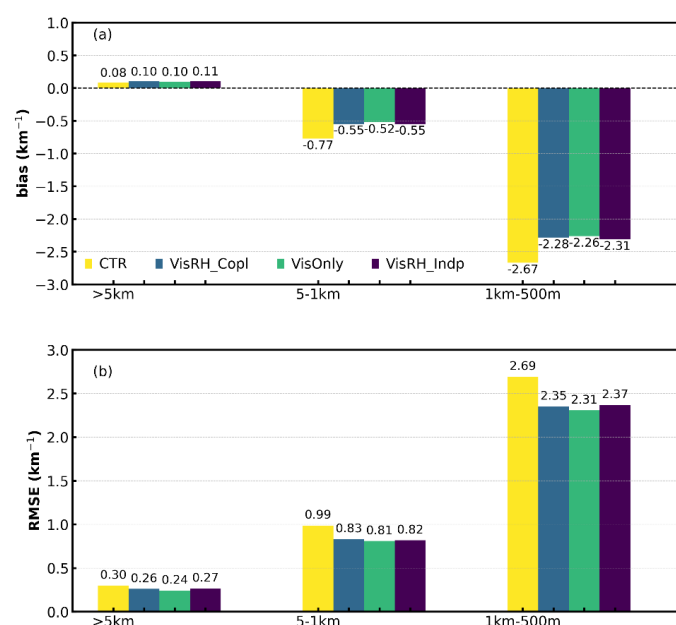
Figure 6: Temporal mean bias of the AExtC predictions over China in the CTR (a), VisRH_Copl (b), VisOnly (c) and VisRH_Indp (d) experiments from 1-31 December 2020. Yellow and red colors denote AExtC overestimations and green and blue colors denote AExtC underestimations.

Figure 7 shows the bias, RMSE, and R of AExtC predictions stratified by three visibility categories: >5 km, 1-5 km and 1 km-500 m. Compared with the no-DA baseline (CTR, yellow bars), the CAOP method (VisRH_Copl, blue bars) yields consistent improvements in the AExtC forecasts across all visibility regimes. For visibilities >5 km, the RMSE decreased from 0.30 km^{-1} to 0.26 km^{-1} (a 13% reduction), partially addressing the systematic underestimation of AExtC noted by Haywood et al. (2008). For the 1-5 km range, the RMSE decreased from 0.99 km^{-1} to 0.83 km^{-1} , representing a 16% reduction. For the lowest visibility category (1 km–500 m), the RMSE decreases from 2.69 km^{-1} to 2.35 km^{-1} (a 13% reduction), with the bias improving from -2.7 km^{-1} to -2.3 km^{-1} .

Clark et al. (2008) reported a RMSE reduction of less than 10% for visibility assimilation over Eastern Europe during 30 January-6 February 2006, and noted that the DA impact was most significant for visibility below 1 km, whereas emissions and advection dominated at slightly higher visibilities (1-5 km). In contrast, our CAOP method achieved a 20% RMSE reduction in the 1-5 km range,



389 demonstrating that the comprehensive chemical speciation and κ -based hygroscopic growth scheme
390 extend the benefit of visibility DA beyond the low-visibility regime identified by Clark et al. We attribute
391 this broader effectiveness to the improved representation of ALWC under different RH conditions, which
392 is critical for the 1-5 km visibility range where RH-dependent processes also play a dominant role.



393
394 **Figure 7: Bias (a) and RMSE (b) of AExtC predictions stratified by visibility categories for the CTR (in**
395 **yellow), VisRH_Copl (in blue), VisOnly (in green) and VisRH_Indp (in purple) experiments. Results are**
396 **stratified into three categories: >5 km, 1-5 km, and 1 km -500 m.**

397 3.2 Forecast improvement of $\text{PM}_{2.5}$

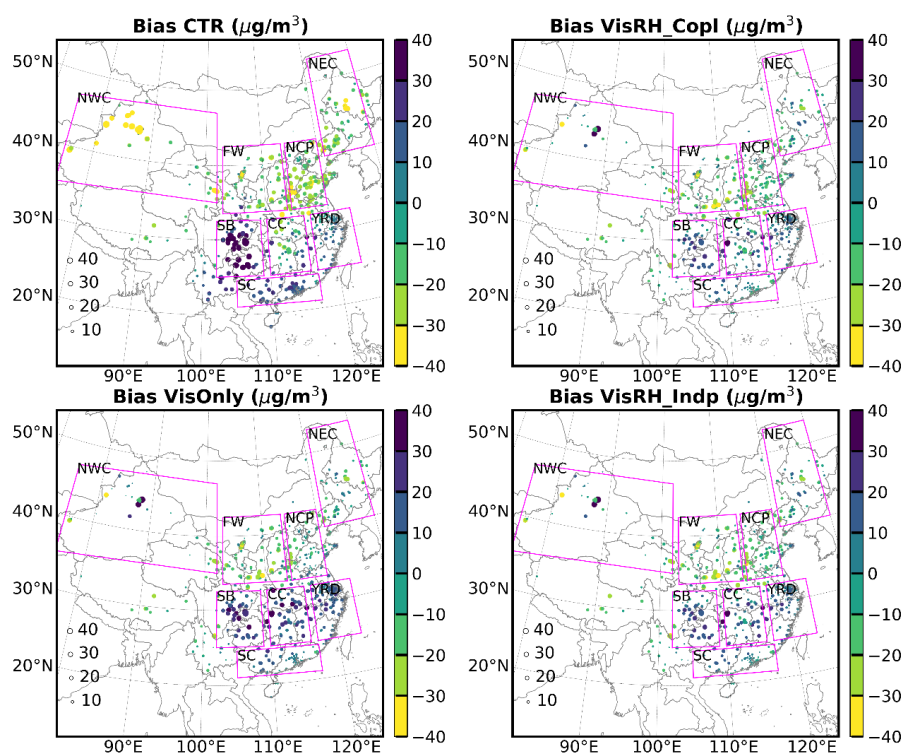
398 The temporal and station mean bias, RMSE, and R for the model predicted and observed $\text{PM}_{2.5}$ are
399 shown in Fig. 4 (b1-b3) and Fig. 5(b) and Table 3. The CTR underestimated $\text{PM}_{2.5}$ with a mean value of
400 $-3.4 \mu\text{g}/\text{m}^3$ over China. This is consistent with the overall underestimation of AExtC. Other regions,
401 such as NCP, NWC, SB, CC, SC, and NEC, also exhibited the phenomenon where the signs of mean bias
402 for $\text{PM}_{2.5}$ and for AExtC were the same (yellow bars in Fig. 5 (a, b), and Fig. 8). Therefore, in these
403 regions, the $\text{PM}_{2.5}$ biases were significantly improved through AExtC DA (in the VisOnly, VisRH_Indp,
404 and VisRH_Copl). For example, the VisRH_Indp obtained the best $\text{PM}_{2.5}$ forecast bias, which was -1.3



405 $\mu\text{g}/\text{m}^3$ over China. The mean RMSE for $\text{PM}_{2.5}$ over China was $33.9 \mu\text{g}/\text{m}^3$ in the CTR (Fig. 5). The
406 improvements in RMSE in the VisOnly, VisRH_Indp, and VisRH_Copl were 9%, 16%, and 17%. The
407 RMSE reduction achieved by VisRH_Copl relative to VisRH_Indp exceeded the MDD, indicating the
408 improvement is statistically significant at the 95% confidence level. As can be seen from Fig. 5 (b), the
409 VisRH_Copl exhibited the most significant improvement in RMSE, both over China and in most regions
410 with the exception of NCP. Moreover, the improvement was significant for the 24-hour forecast range.
411 This demonstrates the superiority of the coupled DA of AExtC and RH in predicting aerosol mass
412 concentration. The mean R for $\text{PM}_{2.5}$ over China was 0.47 in the CTR experiment. The VisOnly improved
413 the $\text{PM}_{2.5}$ trend by 40%. In terms of R improvement, the performances of the VisRH_Indp and
414 VisRH_Copl were slightly inferior to the VisOnly, both achieving improvements of 38%.
415 It can be found that CRT substantially overestimated $\text{PM}_{2.5}$ and AExtC in the SB region, particularly in
416 the southwest. This can be primarily attributed to the 40.5 km horizontal resolution being insufficient to
417 resolve the complex terrain of the basin, which smooths topographic barriers and distorts valley wind
418 circulations critical for pollutant dispersion (Hong et al., 2006). Additionally, the temporal mismatch
419 between the 2016 MEIC emission inventory and the 2020 simulation period is particularly pronounced
420 in the SB region, where $\text{PM}_{2.5}$ -related emissions declined more significantly ($\sim 34\%$ reduction in $\text{PM}_{2.5}$
421 concentration from 2016 to 2020; Department of Ecology and Environment of Sichuan Province, 2017;
422 2021) than the national average ($\sim 18\%$ target), leading to systematic overestimation of actual emission
423 levels (Zheng et al., 2018; 2021). In addition, it is noteworthy that the $\text{PM}_{2.5}$ RMSE improvement by
424 VisRH_Copl was less pronounced in NCP and NEC compared to other regions (Fig. 5b). We attribute
425 this to region-specific factors. For NCP, the monthly mean $\text{PM}_{2.5}$ concentration in December 2020 over
426 the Beijing-Tianjin-Hebei and surrounding areas (2+26 cities), which largely overlap with the NCP
427 region in this study, was $76 \mu\text{g}/\text{m}^3$ (MEE, 2021). Under such high-concentration conditions, the ALWC
428 calculation based on κ -Köhler theory may encounter increased uncertainty due to (1) the dilute-solution
429 assumption inherent in the single-parameter representation, which may not fully capture non-ideal
430 solution effects at elevated solute concentrations, and (2) the heightened sensitivity of the $\frac{a_w}{1-a_w}$ (Zhao
431 et al., 2019). For NEC, the relatively lower $\text{PM}_{2.5}$ concentration and different aerosol composition
432 (dominated by primary coal combustion for heating) meant that ALWC contributed less to total extinction,
433 reducing the marginal benefit of coupled RH assimilation on $\text{PM}_{2.5}$ forecasts. Furthermore, CTR



434 exhibited a positive RH bias in NEC (overestimation) (Fig. 5(c)), unlike most other regions where RH
435 was underestimated. Despite these regional differences, VisRH_Copl still achieved the best or near-best
436 performance for both PM_{2.5} across all regions, confirming the robustness of the coupled approach.



437
438 **Figure 8: Temporal mean bias of the PM_{2.5} predictions over China in the CTR (a), VisRH_Copl (b), VisOnly**
439 **(c) and VisRH_Indp (d) experiments from 1-31 December 2020. Blue and purple colors denote PM_{2.5}**
440 **overestimations and green and yellow colors denote PM_{2.5} underestimations.**

441 3.3 Forecast improvement of RH

442 In addition to AExtC, the RH was also assimilated in the VisRH_Copl and VisRH_Indp experiments.
443 To further reveal the impact of the CAOP operator on RH analysis, temporal and station mean bias,
444 RMSE, and R for RH are shown in Fig. 4(c1-c3) and Fig. 5(c). It should be noted that the VisOnly did
445 not assimilate the RH observations. The slight improvement in the RH forecast may be caused by the
446 PM_{2.5} improvement in the VisOnly via the aerosol-radiation interaction during the forecast process. Thus,
447 the following analysis is targeted at the CTR, VisRH_Indp, and VisRH_Copl.

448

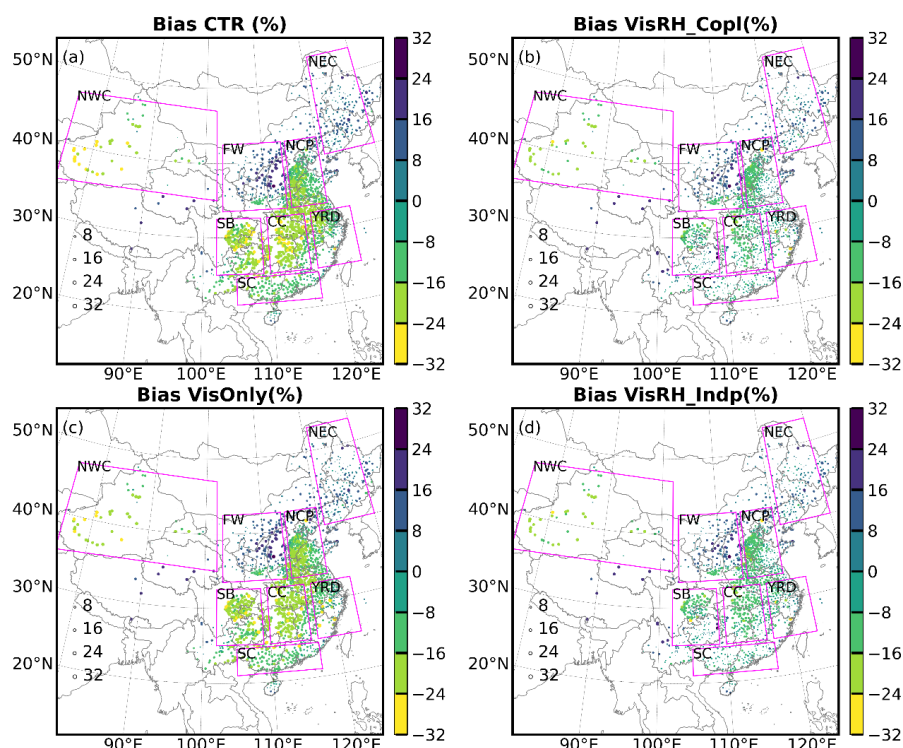


Figure 9: Temporal mean bias of the RH predictions over China in the CTR (a), VisRH_Copl (b), VisOnly (c) and VisRH_Indp (d) experiments from 1-31 December 2020. Blue and purple colors denote RH overestimations and yellow and green colors denote RH underestimations.

The CTR underestimated RH by a mean value of -5.5% over China (Table 3). Furthermore, RH was underestimated in most parts of China, including NCP, NWC, YRD, CC, SB, and SC, while it was overestimated in NEC and FW (Fig. 9). The systematic RH underestimation (-5.5% averaged over China) may partly be attributed to the YSU planetary boundary layer (PBL) scheme. As a nonlocal closure scheme, YSU enhances entrainment at the PBL top and strengthens vertical mixing (Hong et al., 2006), particularly under stable winter conditions (Zhang et al., 2012), which can mix drier free-tropospheric air downward and consequently lower near-surface RH. The positive RH bias in NEC may reflect the combined effect of initial snow field errors and the limitations of the single-layer snow parameterization in the Noah LSM (Chen et al., 1997; Zhang et al., 2021). For the Fenwei Plain (FW) region, the positive RH bias may be attributed to the combined effect of unresolved basin topography and aerosol radiative feedback. The FW region is characterized by a complex basin-valley topography that cannot be adequately resolved at 40.5 km horizontal resolution, leading to stagnant wind conditions and moisture



465 accumulation in the model. Additionally, the positive $PM_{2.5}$ bias in FW may enhance the aerosol radiative
466 cooling effect, reducing surface temperature and saturation vapor pressure, which further elevates RH.

467 The RH bias was reduced to -1.5% through the DA of the RH in the VisRH_Indp, in which the RH
468 improvement was solely attributed to the RH observations. Through further coupled DA of the AExtC
469 and RH (the VisRH_Copl), the RH bias was further improved to -0.7% and this improvement was
470 significant in the 24-hour forecast range (Fig. 4 (c1)). In eight regions across the country, the VisRH_Copl
471 also achieved the best bias performance, regardless of whether the RH was overestimated or
472 underestimated in the CTR (colored bars in Fig. 5 (c)).

473 The mean RMSE for RH was 18.0% over China in the CTR. The RMSE improvements achieved by
474 the VisRH_Indp and VisRH_Copl were 21% and 22%, respectively, with the positive effects lasting for
475 the 24-hour forecast range. From the perspective of improving RMSE, the VisRH_Copl performed the
476 best over China and in the eight regions.

477 The mean R for RH was 0.54 in the CTR over China. The VisRH_Indp and VisRH_Copl achieved R
478 improvements of 23% and 25%, respectively. The R improvement (2%) and RMSE reduction (1%) of
479 VisRH_Copl relative to VisRH_Indp were both statistically significant. The VisRH_Copl demonstrated
480 better RH prediction skill than the VisRH_Indp, further verifying the contribution of the visibility
481 observations towards the RH forecast. As stated by Clark et al. (2008), visibility observations indeed
482 incorporate information about humidity and make a positive contribution to RH analysis and prediction.

483 In Sect. 3.1, a seemingly contradictory phenomenon was mentioned, namely, the AExtC forecast
484 performance obtained through DA of AExtC and RH was inferior to that obtained through DA of only
485 AExtC. With the evaluation of the forecast performance of $PM_{2.5}$ and RH, an explanation can be provided
486 here. From Fig. 5 (b) and (c), taking the SB region as an example, better positive AExtC bias and RMSE
487 were obtained in the VisOnly than those in either the VisRH_Indp or the VisRH_Copl. However, the
488 VisOnly yielded larger positive $PM_{2.5}$ bias because it had a higher negative RH bias than the VisRH_Indp
489 and VisRH_Copl did. The positive $PM_{2.5}$ bias and negative RH bias canceled each other out, resulting in
490 what appeared to be a better AExtC forecast. This further underscores the necessity of joint assimilation
491 of visibility and RH data to avoid the occurrence of uncoordinated $PM_{2.5}$ and meteorological fields.
492 Furthermore, the VisRH_Copl achieved better $PM_{2.5}$ and RH forecasts than the VisRH_Indp over China,
493 and in regions like SB, CC, and YRD. By coupling DA of the AExtC and RH, the aerosol and



494 meteorological analysis fields can better reflect the actual atmospheric condition. This further verified
495 the effectiveness of the CAOP module developed in this study.

496 **4. Conclusions and Discussions**

497 Within a unified 3D variation data assimilation framework (WRFDA-Chem AOP module), we
498 developed a coupled DA method for AExtC and RH (CAOP). It simultaneously optimized the aerosol
499 mass concentration field and the near-surface humidity field at the same time by accurately describing
500 the physical relationships between visibility, aerosol chemical species, and RH. The visibility observation
501 was converted into AExtC, and the model-observation bias for AExtC was corrected before being used
502 in DA or forecast result verification.

503 Visibility meter AExtC and RH were assimilated using the newly developed CAOP method, and
504 the analysis field obtained was used for WRF-Chem prediction during wintertime over China
505 (VisRH_Copl). To evaluate the effectiveness and superiority of the CAOP method, this study designed
506 systematic numerical experiments. First, a basic control run (CTR) with no data assimilation was
507 performed. Second, a comparative experiment (VisOnly) was designed to examine the impact of
508 simultaneously DA of AExtC and RH on PM_{2.5} forecast. Third, to disentangle the effect of the coupled
509 method, we performed the VisRH_Indp, which assimilated AExtC and RH separately while maintaining
510 temporal concurrency, thereby assessing the unique contribution of visibility to RH analysis.

511 The CTR underestimated the AExtC with a mean bias of -0.02 km⁻¹ over China. Regions like NCP,
512 CC, and the northwest part of NWC showed obvious underestimation. The VisOnly demonstrated the
513 largest gains in AExtC, with a 21% RMSE reduction and a remarkable 173% increase in R. Although the
514 VisRH_Indp and VisRH_Copl also enhanced AExtC forecasts considerably, they remained slightly less
515 effective than the VisOnly.

516 The CTR underestimated PM_{2.5} nationwide with a mean value of -3.4 μg/m³. Regions like NEC,
517 NCP, and the northwestern part of NWC showed obvious underestimation. The VisRH_Copl had the
518 most significant RMSE reduction of 17%, with the performance of the VisRH_Indp slightly worse (by
519 16%) and that of VisOnly the worst (9%). In terms of R improvement, the three DA experiments were
520 not significantly different, with the improvement ranging from 38%-40%.

521 The CTR underestimated RH by -5.5% across China, with RH notably underestimated over key



522 regions including NCP, NWC, YRD, CC, and SB. The VisRH_Copl achieved the best performance,
523 yielding the largest RMSE reduction and R improvement for RH, thereby highlighting the value of
524 visibility observations towards the RH forecast. We obtained physically coherent $PM_{2.5}$ and
525 meteorological analysis through the coupled DA of the AExtC and RH, thereby validating the capability
526 of the CAOP module developed in this study.

527 This study conducted DA experiments based on data from December 2020 (wintertime), primarily
528 motivated by the frequent occurrence of heavy pollution episodes during this period and the ALWC effect
529 was pronounced. Nevertheless, we fully recognize that a single-month wintertime experiment is
530 insufficient to fully characterize the complete seasonal picture of aerosol hygroscopic growth.
531 Pronounced differences in aerosol chemical composition, RH, and mixing states across seasons, such as
532 enhanced secondary formation under hot and humid summer conditions, and elevated insoluble mineral
533 dust fractions during dust storm episodes, may systematically modulate the relative contribution of
534 ALWC to AExtC (Zhang et al., 2020; Wang et al., 2021). The CAOP module based on κ – Köhler
535 theory established in this study is not intrinsically constrained by seasonality. Thus, this approach can be
536 readily extended to other temporal windows. We recommend future investigations undertake inter-
537 seasonal comparative DA experiments to encompass the full meteorological spectrum.

538 In addition, correlations between the RH and 32 aerosol chemical species were excluded from the
539 BE statistic here. Incorporating these covariances represents a promising pathway for future model
540 enhancement, considering that RH plays an important role in liquid-phase chemistry. The CAOP
541 method's primary advantage that can propagate RH analysis increments to aerosol species through the
542 BE cross-covariance can not be fully realized when the cross-covariance is set to be zero. Under the
543 current implementation, the CAOP operates primarily through the joint observation operator, which
544 simultaneously constrains both RH and AExtC in the cost function, rather than through the multivariate
545 BE structure. This explains why VisRH_Copl and VisRH_Indp yield similar performance in most metrics.
546 Besides, the use of a 2016 emission inventory for 2020 simulations introduces uncertainty in the absolute
547 aerosol loading and relative proportion of aerosol chemical species. This may potentially reduce the
548 relative apparent impact of DA. Future work should incorporate updated multi-year emission inventories
549 to improve the representation of aerosol predictions.



550 **Code and data availability**

551 The original WRFDA and WRF-Chem codes are open sources and can be found at
552 https://www2.mmm.ucar.edu/wrf/users/download/get_source.html. WRF-Chem v4.0 (Grell et al., 2005)
553 and WRFDA-Chem v4.3 (Sun et al., 2020) are used in this study. The CAOP v1.0 module developed in
554 this study is available on Zenodo at <https://doi.org/10.5281/zenodo.22887371> (Gao, 2026a). The ground
555 synoptic observations (including RH data, visibility) is available at China Integrated Meteorological
556 Information Service System (<https://data.cma.cn/>; Zhao et al., 2018). The PM_{2.5} observations are from the
557 China National Environmental Monitoring Center (CNEMC) (<https://www.cnemc.cn/>; Gao et al., 2020).
558 For reproducibility, the processed visibility, PM_{2.5} and RH observations in ascii format and the namelist
559 configuration file used in DA experiments are available on Zenodo at
560 <https://doi.org/10.5281/zenodo.22886008> (Gao, 2026b).

561 **Author contributions**

562 LG and ZL conducted the development of the DA system. JJ, WY, XX and PY designed the research;
563 LG performed experiments and results verification. JJ provided the synoptic and PM observations, and
564 LG wrote the paper with contributions of all authors.

565 **Competing interests**

566 The contact author has declared that none of the authors has any competing interests.

567 **Acknowledgement**

568 This study was supported by the Jing-Jin Ji Regional Integrated Environmental Improvement-
569 National Science and Technology Major Project (2026ZD1212900) and the National Natural Science
570 Foundation of China under grant 42430612.

571 **Financial support**

572 This research has been supported by the Jing-Jin Ji Regional Integrated Environmental
573 Improvement-National Science and Technology Major Project (grant no. 2026ZD1212900) and the



574 National Natural Science Foundation of China (grant no. 42430612).

575 **References**

- 576 Bao, Y., Zhu, L., Guan, Q., Guan, Y., Lu, Q., Petropoulos, G.P., et al.: Assessing the impact of Chinese
577 FY-3/MERSI AOD data assimilation on air quality forecasts: Sand dust events in northeast China, *Atmos.*
578 *Environ*, 205, 78-89, <https://doi.org/10.1016/j.atmosenv.2019.02.026>, 2019.
- 579 Barnard, J.C., Fast, J.D., Paredes-Miranda, G., Arnott, W.P., Laskin, A.: Technical Note: Evaluation of
580 the WRF-Chem “Aerosol chemical to Aerosol Optical Properties” module using data from the
581 MILAGRO campaign, *Atmos. Chem. Phys.* 10, 7325-7340, <https://doi.org/10.5194/acp-10-7325-2010>,
582 2010.
- 583 Clark, P.A., Harcourt, S.A., Macpherson, B., Mathison, C.T., Cusack, S., Naylor, M.: Prediction of
584 visibility and aerosol within the operational Met Office Unified Model. I: Model formulation and
585 variation assimilation, *Q. J. R. Meteorol. Soc.*, 134, 1801-1816, <https://doi.org/10.1002/qj.318>, 2008.
- 586 Cohen, A.J., Brauer, M., Burnett, R., Anderson, H.R., Frostad, J., Estep, K., et al.: Estimates and 25-year
587 trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the
588 Global Burden of Diseases Study 2015, *Lancet*, 389, 1907–1918, [https://doi.org/10.1016/S0140-](https://doi.org/10.1016/S0140-6736(17)30505-6)
589 [6736\(17\)30505-6](https://doi.org/10.1016/S0140-6736(17)30505-6), 2017.
- 590 Chen, F., Mitchell, K., Schaake, J., Xue, Y., Pan, H.-L., Koren, V., et al.: Modeling of land surface
591 evaporation by four schemes and comparison with FIFE observations, *J. Geophys. Res. Atmos.*, 101(D3),
592 7251-7268, <https://doi.org/10.1029/95JD02165>, 1996.
- 593 Chen, J., Zhao, C.S., Ma, N., Liu, P.F., Göbel, T., et al.: A parameterization of low visibilities for hazy
594 days in the North China Plain, *Atmos. Chem. Phys.*, 12, 4935-950, [https://doi.org/10.5194/acp-12-4935-](https://doi.org/10.5194/acp-12-4935-2012)
595 [2012](https://doi.org/10.5194/acp-12-4935-2012), 2012.
- 596 Chin, M., Ginoux, P., Kinne, S., Torres, O., Holben, B. N., Duncan, B. N., et al.: Tropospheric aerosol
597 optical thickness from the GOCART model and comparisons with satellite and Sun photometer
598 measurements, *J. Atmos. Sci.*, 59, 461-483, [https://doi.org/10.1175/1520-](https://doi.org/10.1175/1520-0469(2002)059<0461:TAOTFT>2.0.CO;2)
599 [0469\(2002\)059<0461:TAOTFT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2002)059<0461:TAOTFT>2.0.CO;2), 2002.
- 600 Gao, L., Ren, L., Liu, Z., Sun, W., Jin, J., You, W., et al.: Improving PM_{2.5} forecasts with three-
601 dimensional variation data assimilation of visibility observations in China, *Atmos. Environ.*, 356, 121296,
602 <https://doi.org/10.1016/j.atmosenv.2025.121296>, 2025a.
- 603 Gao, L., Yue, X., Meng, X., Du, L., Lei, Y., Tian, C., et al.: Comparison of ozone and PM_{2.5}
604 concentrations over urban, suburban, and background sites in China [dataset]. *Adv. Atmos. Sci.*, 37(12),
605 1297–1309, <https://doi.org/10.1007/s00376-020-0054-2>, 2020.
- 606 Gao, L.: CAOP v1.0 within WRFDA v4.3 [Computer software]. Zenodo.
607 <https://doi.org/10.5281/zenodo.22887371>, 2026a.
- 608 Gao, L.: Processed visibility, PM, and RH observations and configuration file for CAOP v1.0 [Dataset].
609 Zenodo. <https://doi.org/10.5281/zenodo.22886008>, 2026b.
- 610 Gao, X., Wang, X., Li, F., Jiang, W., Zhe, M., Sun, J., et al.: PM_{2.5} concentration 7-day prediction in the
611 Beijing-Tianjin-Hebei region using a novel stacking framework, *Sci. Rep.*15, 20731,
612 <https://doi.org/10.1038/s41598-025-07719-7>, 2025b.
- 613 Geng, G., Liu, Y., Liu, Y., Liu, S., Cheng, J., Yan, L., et al.: Efficacy of China's clean air actions to tackle
614 PM_{2.5} pollution between 2013 and 2020, *Nat. Geosci.*, 17, 987-994, <https://doi.org/10.1038/s41561-024->



- 615 01540-z, 2024.
- 616 Grell, G.A., Peckham, S.E., Schmitz, R., McKeen, S.A., Frost, G., Skamarock, W.C., et al.: Fully coupled
617 “online”_chemistry with the WRF model [Software]. *Atmospheric Environment*, 39(37), 6957-6975,
618 <https://doi.org/10.1016/j.atmosenv.2005.04.027>, 2005.
- 619 Hong, S.-Y., and Lim, J. O. J.: The WRF Single-Moment 6-Class Microphysics Scheme (WSM6), *J.*
620 *Korean Meteorol. Soc.*, 42(2), 129-151, <https://doi.org/10.1175/MWR-D-14-00005.1>, 2006.
- 621 Hyslop, N.P.: Impaired visibility: the air pollution people see, *Atmos. Environ.*, 43, 182-195,
622 <https://doi.org/10.1016/j.atmosenv.2008.09.067>, 2009.
- 623 Haywood, J., Bush, M., Abel, S., Claxton, B., Coe, H., Crosier, J., et al.: Prediction of visibility and
624 aerosol within the operational Met Office Unified Model. II: validation of model performance using
625 observational data. *Q. J. R. Meteorol. Soc.* 134, 1817-1832, <https://doi.org/10.1002/qj.318>, 2008.
- 626 Iacono, M.J., Delamere, J.S., Mlawer, E.J., Shepard, M.W., Clough, S.A., and Collins, W.D.: Radiative
627 forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models. *J. Geophys.*
628 *Res. Atmos.* 113, D13103, <https://doi.org/10.1029/2008JD009944>, 2008.
- 629 Jolliffe, I.T., and Stephenson, D.B. (Eds.): *Forecast Verification: A Practitioner's Guide in Atmospheric*
630 *Science* (2nd ed.). Wiley, 292 pp, ISBN: 978-0-470-66071-3, 2012.
- 631 Jing, D., He, Y., Yin, Z., Huang, K., Liu, F., and Yi, F.: Evolution of tropospheric aerosols over central
632 China during 2010–2024 as observed by lidar, *Atmos. Chem. Phys.*, 25, 17047-17067,
633 <https://doi.org/10.5194/acp-25-17047-2025>, 2025.
- 634 Jin, X., Li, Z., Wu, T., Wang, Y., Su, T., Ren, R., et al.: Differentiating the contributions of particle
635 concentration, humidity, and hygroscopicity to aerosol light scattering at three sites in China, *J. Geophys.*
636 *Res. Atmos.*, 127(24), D036891, <https://doi.org/10.1029/2022JD036891>, 2022.
- 637 Jin, X., Wang, Y., Li, Z., Zhang, F., Xu, W., Sun, Y., et al.: Significant contribution of organics to aerosol
638 liquid water content in Beijing, China, *Atmos. Chem. Phys.*, 20, 901-914, <https://doi.org/10.5194/acp-20-901-2020>, 2020.
- 640 Kim, M., Lee, K., Lee, Y. H.: Visibility data assimilation and prediction using an observations network
641 in South Korea, *Pure Appl. Geophys.* 177, 1125-1141, <https://doi.org/10.1007/s00024-019-02288-z>,
642 2020.
- 643 Koschmieder, H., Theorie der horizontalen sichtweite. *Beitr Phys.d.freien Atm.*, 12,
644 https://doi.org/10.1007/978-3-663-04661-5_2, 1924.
- 645 Li, M., Zhang, Q., Streets, D., He, Y., Cheng, L.K., Emmons, H., et al.: Mapping Asia anthropogenic
646 emissions of non-methane volatile organic compounds to multiple chemical mechanisms. *Atmos. Chem.*
647 *Phys.* 14, 5617–5638, <https://doi.org/10.5194/acp-14-5617-2014>, 2014.
- 648 Liu, J., Xie, Y., Chen, Y.: Analysis on the difference of relative humidity between automatic station and
649 manual station in Hunan Province, *Climate Change Research Letters*, 11(4), 455-459. DOI:
650 10.12677/ccrl.2022.114047, DOI: 10.12677/CCRL.2022.114047, 2022. (In Chinese)
- 651 Liu, Z., Liu, Q., Lin, H., Schwartz, C.S., Lee, Y., Wang, T.: Three-dimensional variational assimilation
652 of MODIS aerosol optical depth: implementation and application to a dust storm over East Asia. *J.*
653 *Geophys. Res. Atmos.* 116, D23206, <https://doi.org/10.1029/2011JD016159>, 2011.
- 654 MEE (Ministry of Ecology and Environment of China): Report on the national surface water and ambient
655 air quality conditions for December 2020 and January–December 2020, Last access:
656 https://www.mee.gov.cn/xxgk/xxgk15/202101/t20210115_817499_wh.html, 2021. (in
657 Chinese)
- 658 Mei, S., You, W., Zhong, W., Zang, Z., Guo, J., Cheng, X., and Gao, L.: Improving aerosol forecasting



659 through visibility data assimilation in WRF-Chem: Methodology and evaluation, *Atmos. Res.*, 327,
660 108383, <https://doi.org/10.1016/j.atmosres.2025.108383>, 2025.

661 Nguyen, T.K.V, Zhang, Q., Jimenez, J.L., Pike, M., Carlton, A.G.: Liquid water: ubiquitous contributor
662 to aerosol mass, *Environ. Sci. Technol. Lett.*, 3, 257-263, DOI: 10.1021/acs.estlett.6b00167, 2016.

663 Rosenfeld, D., Dai, J., Yu, X., Yao, Z., Xu, X., Yang, X., Du, C.: Inverse relations between amounts of
664 air pollution and orographic precipitation, *Science*, 315, 1396-1398, DOI: 10.1126/science.113794, 2007.

665 Saide, P. E., Gao, M., Lu, Z., Goldberg, D. L., Streets, D. G., Woo, J.-H., et al.: Understanding and
666 improving model representation of aerosol optical properties for a Chinese haze event measured during
667 KORUS-AQ, *Atmos. Chem. Phys.*, 20, 6455-6478, <https://doi.org/10.5194/acp-20-6455-2020>, 2020.

668 Department of Ecology and Environment of Sichuan Province: 2016 Sichuan Provincial Environmental
669 Conditions Bulletin, retrieved from <https://sthjt.sc.gov.cn>, 2017.

670 Department of Ecology and Environment of Sichuan Province, 2020 Sichuan Provincial Environmental
671 Conditions Bulletin, retrieved from <https://sthjt.sc.gov.cn>, 2021.

672 Sogacheva, L., Popp, T., Sayer, A.M., Dubovik, O., Garay, M.J., Heckel, A. et al.: Merging regional and
673 global aerosol optical depth records from major available satellite products, *Atmos. Chem. Phys.*, 20,
674 2031-2056, <https://doi.org/10.5194/acp-20-2031-2020>, 2020.

675 Sun, W., Liu, Z., Chen, D., Zhao, P., and Chen, M.: Development and application of the WRFDA-chem
676 three-dimensional variational (3DVAR) system: Aiming to improve air quality forecasting and diagnose
677 model deficiencies [Software]. *Atmospheric Chemistry and Physics*, 20(15), 9311-9329,
678 <https://doi.org/10.5194/acp-20-9311-2020>, 2020.

679 Wang, D., You, W., Zang, Z., Pan, X., Hu, Y., and Liang, Y.: A three-dimensional variational data
680 assimilation system for aerosol optical properties based on WRF-Chem V4.0: design, development, and
681 application of assimilating Himawari-8 aerosol observations, *Geosci. Model Dev.*, 15, 1821-1840,
682 <https://doi.org/10.5194/gmd-15-1821-2022>, 2022.

683 Wang, H., Yang, T., Wang, Z., Li, J., Chai, W., Tang, G., et al.: An aerosol vertical data assimilation
684 system (NAQPMS-PDAF v1.0): development and application, *Geosci. Model Dev.*, 15(9), 3555-3585,
685 <https://doi.org/10.5194/gmd-15-3555-2022>, 2022.

686 Wang, H., Martin, C., Barré, J., Li, R., Weygandt, S., Huang, J., Tang, Y., et al.: PM_{2.5} assimilation
687 within JEDI for NOAA's regional Air Quality Model (AQMV7): application to the September 2020
688 Western US wild fires, *Geosci. Model Dev.*, 19(6), 2479-2495, [https://doi.org/10.5194/gmd-19-2479-](https://doi.org/10.5194/gmd-19-2479-2026)
689 2026, 2026.

690 Wu, T., Li, Z., Chen, J., Wang, Y., Wu, H., Jin, X., et al.: Hygroscopicity of different types of aerosol
691 particles: case studies using multi-instrument data in megacity Beijing, China, *Remote Sens.*, 12, 785,
692 <https://doi.org/10.3390/rs12050785>, 2020.

693 Xiang, Y., Zhang, T., Ma, C., Lv, L., Liu, J., et al.: Lidar vertical observation network and data assimilation
694 reveal key process driving the 3-D dynamic evolution of PM_{2.5} concentrations over the North China Plain,
695 *Atmos. Chem. Phys.*, 21, 7023-7037, <https://doi.org/10.5194/acp-21-7023-2021>, 2021.

696 Zeng, Y., Wang, M., Zhao, C., Chen, S., Liu, Z., Huang, X., et al.: WRF-Chemv3.9 simulations of the
697 East Asian dust storm in May 2017: modeling sensitivities to dust emission and dry deposition schemes.
698 *Geosci. Model Dev.*, 13, 2125-2147, <https://doi.org/10.5194/gmd-13-2125-2020>, 2020.

699 Zhao, C., Yu, Y., Kuang, Y., Tao, J., Zhao, G.: Recent Progress of Aerosol Light-scattering Enhancement
700 Factor Studies in China. *Adv. Atmos. Sci.*, 36(9): 1015-1026, [https://doi.org/10.1007/s00376-019-8248-](https://doi.org/10.1007/s00376-019-8248-1)
701 1, 2019.

702 Zhao, F., He, C., Zhang, X., and Zhang, Z.: Construction of China integrated meteorological information



703 sharing system [Dataset]. *Advances in Meteorological Science and Technology*, 8(1), 171-180, 2018.
704 <https://data.cma.cn/>
705 Zhao, P., Ge, S., Su, J., Ding, J., Kuang, Y.: Relative Humidity Dependence of Hygroscopicity Parameter
706 of Ambient Aerosols, *J. Geophysical Research: Atmospheres*, 127(8), e2021JD035647,
707 <https://doi.org/10.1029/2021JD035647>, 2022.
708 Zhang, B., Liu, S., Liu, H., Ma, Y.: The effect of MYJ and YSU schemes on the simulation of boundary
709 layer meteorological factors of WRF, *Chinese Journal of Geophysics*, 55(07), 2239-2248, doi:
710 10.6038/j.issn.0001-5733.2012.07.010, 2012. (in Chinese)
711 Zhang, H., Zhang, F., Che, T., Yan, W., Ye, M.: Investigating the Ability of Multiple Reanalysis Datasets
712 to Simulate Snow Depth Variability over Mainland China from 1981 to 2018, *J. Climate*, **34**, 9957-9972,
713 <https://doi.org/10.1175/JCLI-D-20-0804.s1>, 2021.
714 Zhang, Q., Zheng Y, Tong D, Shao M, Wang S, Zhang Y, et al.: Drivers of improved PM_{2.5} air quality
715 in China from 2013 to 2017, *Proc. Natl. Acad. Sci. U.S.A.*, 22: 24463-24469,
716 <https://doi.org/10.1073/pnas.1907956116>, 2019.
717 Zhang, S., Wu, J., Fan, W., Yang, Q., Zhao, D.: Review of aerosol optical depth retrieval using visibility
718 data. *Earth Sci. Rev.* 200, 102986, <https://doi.org/10.1016/j.earscirev.2019.102986>, 2020.
719 Zhang, X., Gui, K., Zeng, Z., Fei, Y., Li, L., Zheng, Y., et al.: Mapping the seamless hourly surface
720 visibility in China: a real-time retrieval framework using a machine-learning-based stacked ensemble
721 model, *npj Climate and atmospheric Science*, 7, 68, <https://doi.org/10.1038/s41612-024-00617-1>, 2024.
722 Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., et al.: Trends in China's anthropogenic emissions
723 since 2010 as the consequence of clean air actions, *Atmos. Chem. Phys.*, 18, 14095-14111,
724 <https://doi.org/10.5194/acp-18-14095-2018>, 2018.
725 Zheng, B., Zhang, Q., Geng, G., Chen, C., Shi, Q., Cui, M., et al.: Changes in China's anthropogenic
726 emissions and air quality during the COVID-19 pandemic in 2020, *Earth Syst. Sci. Data*, 13, 2895-2907,
727 <https://doi.org/10.5194/essd-13-2895-2021>, 2021.