ENVIRONMENTAL RESEARCH
COMMUNICATIONS



LETTER • OPEN ACCESS

# Statistical bias correction for CESM-simulated $PM_{2.5}$

To cite this article: Qi Ran *et al* 2023 *Environ. Res. Commun.* **5** 101001

Environmental Research Communications

LETTER

# Statistical bias correction for CESM-simulated $PM_{2.5}$

Qi Ran[1,2,3], John Moore[4], Tianyun Dong[1,3], Shao-Yi Lee[5] and Wenjie Dong[1,3,*]

[1] School of Atmospheric Sciences, Key Laboratory of Tropical Atmosphere-Ocean System, Ministry of Education, Sun Yat-sen University, Zhuhai, 519082, People's Republic of China
[2] Meteorological Administration Key Open Laboratory of Transforming Climate Resources to Economy, Chongqing, 401147, People's Republic of China
[3] Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, 519000, People's Republic of China
[4] Arctic Centre, University of Lapland, Rovaniemi, 96101, Finland
[5] Disaster Prevention Research Institute, Kyoto University, Uji, 6110011, Japan
[*] Author to whom any correspondence should be addressed.

**E-mail:** dongwj3@mail.sysu.edu.cn

**Keywords:** $PM_{2.5}$ correction, multiple linear regression, global climate models

Supplementary material for this article is available online

**OPEN ACCESS**

**RECEIVED**
9 May 2023

**REVISED**
8 June 2023

**ACCEPTED FOR PUBLICATION**
12 September 2023

**PUBLISHED**
17 October 2023



## Abstract

Global climate models are good tools for simulating transnational and interregional transport of pollutants such as $PM_{2.5}$, which is of growing interest and importance, for example in human health and socio-economic development studies. However, reliable estimates of $PM_{2.5}$ are very challenging for such relatively coarse and simplified models, and even state of the art models fare poorly in matching satellite observations in many highly polluted, and some almost pristine environments. This work describes a novel bias correction method based on multiple linear regression (MLR) modelling. The target data we aim for is global satellite-based data and the $PM_{2.5}$ precursors simulated by the Community Earth System Model Version 1.2.2. The statistical method greatly reduced the simulation biases of $PM_{2.5}$ worldwide compared with satellite-derived $PM_{2.5}$, especially in highly-polluted regions, such as northern China, the Indo-Gangetic plains, the Democratic Republic of Congo and northwestern Brazil. Root-mean-square differences (RMSD) between continental-averaged observations and simulations are reduced from 75% to 9%. The ensemble RMSD for 13 countries exemplified here is reduced from 116% to 3%. One virtue of the MLR method is that details of the classification of internal mixed modes of each aerosol and their spatial differences are not required. The MLR coefficients are designed to be highly aerosol- and country-dependent, so they provide new perspectives of relative importance of each aerosol to local $PM_{2.5}$ and offer clues on observational and simulation biases. The bias-correction method is easily applied for air pollutants simulated by global climate models due to its low computational cost.

## 1. Introduction

Regional climate models (RCMs), such as the weather research and forecasting model (WRF), are commonly used to simulate fine particulate matter and assess local air quality. RCMs have higher spatial resolution than global climate models (GCMs) and tailored representations of the regional or urban contexts (Bai *et al* 2021, Bran and Srivastava 2017, He *et al* 2018, Yang *et al* 2020). However, RCMs are not suitable for the simulation of long-range transport of air pollutants over larger regions, such as a hemispheric domain. In these cases, GCMs are more suitable, and the accurate simulation of fine particulate matter in GCMs becomes a concern.

The Community Earth System Model (CESM, Hurrell *et al* 2013) is a state-of-the-art GCM, providing computer simulations of the Earth's past, present, and future climate states. The CESM has been used in a number of studies to simulate $PM_{2.5}$. Recent studies such as Banks *et al* (2022) and Xu and Lamarque (2018) have simulated aerosol concentrations using CESM with the Modal Aerosol Module, using a combination of the fine

modes (Aitken and Accumulation modes) as a proxy for surface $PM_{2.5}$ concentrations. Sulfate, black carbon, primary organics, and secondary organics were included, but the fine components of sea salt and dust aerosols were not considered. Compared with the satellite-derived $PM_{2.5}$ (van Donkelaar 2018), the modeled $PM_{2.5}$ concentrations in Banks *et al* (2022) were underestimated by more than 10 $\mu g\ m^{-3}$ over northern India, southern China, as well as most Southeast Asian and European countries, while concentrations were overestimated by more than 10 $\mu g\ m^{-3}$ over central South America. Possible factors responsible for simulation biases are uncertainties in emission inventories, meteorological input, and over-simplified chemistry or aerosol processes (Hur *et al* 2021, Gao *et al* 2011, Tilmes *et al* 2015, Liu 2012). Improving $PM_{2.5}$ simulations with more accurate emission inventories or better process modelling are ongoing long-term efforts. Meanwhile, carefully chosen bias-correction methods are a rapid and convenient way for adjusting GCM model output for use in downstream impact modelling.

Some bias correction methods have been effectively applied to reduce the $PM_{2.5}$ bias in air quality models, such as the simple running mean average, Kalman-Filtering (Djalalova *et al* 2010), the analog ensemble method (Huang *et al* 2017), the Bayesian model (McMillan *et al* 2009) and a cluster-analysis-based synoptic weather pattern (WP) classification (Cheng *et al* 2021). A recent machine learning-based correction method found great improvements in correcting the $PM_{2.5}$ biases over China (Liu and Xing 2022). The general applicability of all these methods on global domains has not yet been tested.

In this work, we use correlation analysis and a multiple linear regression (MLR) model to correct the bias between CESM-simulated $PM_{2.5}$ and the satellite-based data (van Donkelaar *et al* 2021). The details of the bias-correction method and its global utility are discussed here. Section 2 describes the data needed and methodology used, section 3 shows the results at grid point level across the globe, and section 4 discusses limitations and potential of the approach.

## 2. Method

### 2.1. Model

The Community Earth System Model version 1.2.2 was used to simulate global annual mean $PM_{2.5}$ concentrations. The 'B_2000_CAM5_CN' component set was used, including Community Atmosphere Model 5 (CAM5) (Neale *et al* 2010), Parallel Ocean Program version 2 (POP2), Community Land Surface Model (CLM) version 4.0, the Los Alamos sea ice model (CICE) version 4, all coupled together using the CESM coupler CPL7. The horizontal resolution of CAM5 was $0.9° \times 1.25°$ latitude-longitude. The Modal Aerosol Module with three modes (MAM3) was used to simulate the aerosol size distribution and the mixing between aerosol components, and the complex aerosol processes and physical, chemical and optical properties of aerosols are treated in a physically-based mode (Liu 2012). Black carbon (BC), primary organic matter (POM), sulfate aerosol, secondary organic aerosol (SOA), dust and sea salt aerosols were simulated in three internally mixed modes by MAM3, namely Aitken Mode (with dry diameter between 0.015–0.053 $\mu m$), Accumulation Mode (0.058–0.27 $\mu m$) and Coarse Mode (0.80–3.65 $\mu m$). Input emissions were based on Aerosol Comparisons between Observations and Models (Textor *et al* 2006).

To simulate the modern $PM_{2.5}$ concentrations worldwide, the default 'B_2000_CAM5_CN' experiment with greenhouse gas (GHG) levels and aerosol emissions of the year 2000 was adjusted to that of the year 2020. The atmospheric GHG concentrations were changed to that of the year 2020 from NOAA (available at https://gml.noaa.gov/ccgg/trends/). The aerosol emissions were updated to the year 2020 values from the Representative Concentration Pathway 6.0 (RCP6.0) scenario experiment in CESM. The experiment was run for 300 years to allow the simulation to reach the climate equilibrium state of 2020, and the last 100 years were used for the following analysis.

### 2.2. $PM_{2.5}$ concentrations

Since there is no straightforward output of $PM_{2.5}$ concentrations in CESM 1.2.2, we considered annual mean aerosol concentrations with diameters smaller than 2.5 $\mu m$ (Aitken, Accumulation and part of Coarse mode) in MAM3 as a proxy for simulated surface $PM_{2.5}$. Integrating the Coarse mode log-normal spectrum up to 2.5 microns suggests that around 36% of total coarse mode can be considered $PM_{2.5}$ (Qin 2003). All of the simulated black carbon (BC) and primary organic matter (POM) were $PM_{2.5}$, since they only have the Accumulation mode. All of the SOA was $PM_{2.5}$, since it only has Aitken and Accumulation modes. Sulfate and sea salt (SS) aerosols have all three modes, while dust aerosol has Accumulation and Coarse modes. Accordingly, 36% of the coarse mode of sulfate, sea salt and dust were taken as $PM_{2.5}$. Equation (1) summarises the calculation described above. This method is similar to Banks *et al* (2022) and Xu and Lamarque (2018), but additionally includes the contribution to $PM_{2.5}$ by sea salt, dust aerosols and the Coarse mode of sulfate.

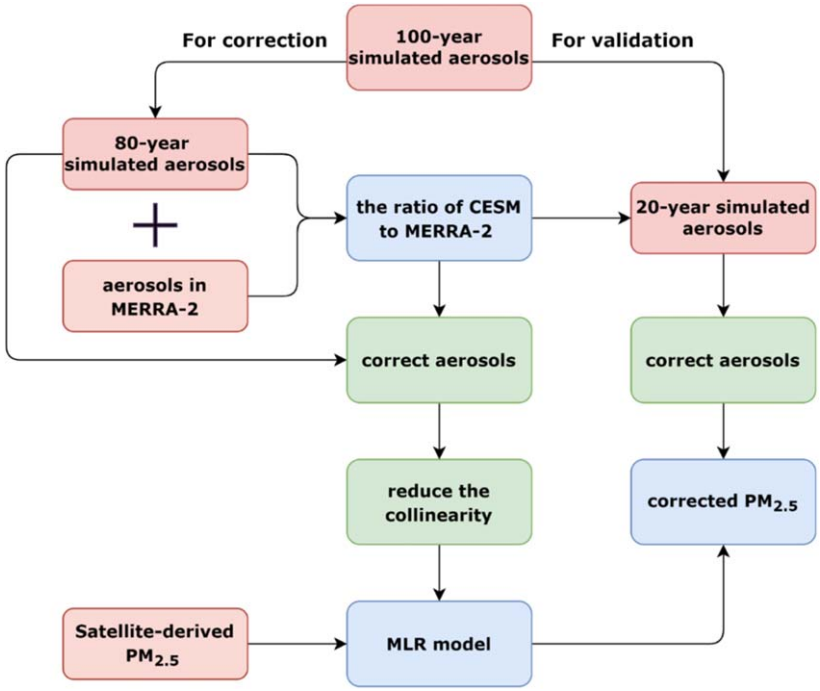


**Figure 1.** Flowchart of bias correction process of CESM-simulated $PM_{2.5}$. The process is divided into two parts: (i) establishing bias-correction model using 80-year simulated aerosols, MERRA-2 reanalysis aerosols over 2000–2019 (GMAO 2015) and the satellite-derived $PM_{2.5}$ over 2000–2019 (van Donkelaar *et al* 2021); (ii) validating the bias-correction method using 20-year simulated aerosols, the ratio of 80-year mean simulated aerosols to MERRA-2 and the MLR model. The pink boxes are the input data for bias-correction, the green boxes show the intermediate processing steps, and the blue boxes are the output results.

$$\begin{aligned} Simulated \quad PM_{2.5} &= BC_{m1} + POM_{m1} + SOA_{m1} + SOA_{m2} + Sulfate_{m1} + Sulfate_{m2} \\ &+0.36 * Sulfate_{m3} + +SS_{m1} + SS_{m2} + 0.36 * SS_{m3} + dust_{m1} + 0.36 * dust_{m3} \end{aligned} \quad (1)$$

In equation (1) above, the subscripts $m1$, $m2$and$m3$ refer to Accumulation, Aitken and Coarse modes.

The 2000–2019 multi-year mean aerosol concentrations from the Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2, GMAO 2015) were used as the reference for the first step of bias correction. This data was of $0.5^\circ \times 0.625^\circ$ horizontal resolution. After bias-correction of each aerosol type by MERRA-2, a satellite-derived $PM_{2.5}$ with $0.1^\circ \times 0.1^\circ$ horizontal resolution (van Donkelaar *et al* 2021) over the same 2000–2019 period was used for a secondary correction for total $PM_{2.5}$ and model validation. This satellite-derived dataset combines satellite observations, chemical transport modeling and ground-based monitoring.

### 2.3. Correction process

We selected thirteen countries worldwide (Europe: Germany, Poland, Italy, Asia: China, India, Indonesia, Oceania: Australia, Africa: Nigeria, Egypt, Democratic Republic of Congo, The Americas: Mexico, Brazil, Peru) with relatively large $PM_{2.5}$ biases compared with the observations (see figure S1), to test the effectiveness of the correction method on a global scale. In all the following discussion we limit ourselves to just these countries as they span a range of behaviors representative of the global dataset and avoid the need to discuss every individual country. The results for 6 continents are also shown in table S1. A flowchart of the multiple-step correction process is given in figure 1.

The bias correction process is divided into two parts, obtaining the correction parameters and validating the correction model. We randomly divided the CESM-simulated 100-year annual mean aerosol concentrations into 5 groups, taking the first 4 groups of 80 years to establish the parameters of the bias correction model, and the last group of 20 years was used to validate the correction effect. The random selection was performed five times, which we used for uncertainty in the bias correction method.

The 2000–2019 multi-year average aerosol concentrations from the MERRA-2 reanalysis dataset were used to calibrate the CESM-simulated sulfate, BC, POM, dust and sea salt aerosols (only considering aerosols less than 2.5 $\mu$m in diameter, and SOA is not included in MERRA-2). The $0.5^\circ \times 0.625^\circ$ horizontal resolution of MERRA-2 was bilinearly interpolated to the $0.9^\circ \times 1.25^\circ$ resolution for comparison with CESM outputs (figure S2). In general, the spatial distribution of simulated aerosol concentrations matched well with that of MERRA-2, except for the distribution biases of sulfate and sea salt in the Democratic Republic of Congo, and sulfate in Italy. However, the magnitudes of sulfate, BC and POM concentrations were underestimated by CESM

in most regions, particularly in northern China, northern India and southern Nigeria compared with MERRA-2 (figures S2.2 and S2.3). In addition, POM concentrations over Brazil were also underestimated (figure S2.4). The most plausible reason is that the emission inventory used by MAM3 did not capture the high levels of anthropogenic emissions in regions where sulfate, BC and POM were severely underestimated. As for dust aerosols, significant overestimation occurred in northwestern China, Egypt and northern Nigeria (figures S2.2 and S2.3), probably due to the overestimation of the effects of deserts. Since the ratio of the simulated 80-year mean aerosol concentrations to the 2000–2019 mean aerosol concentrations of MERRA-2 was fairly uniform worldwide (figure S2), we corrected the simulated 80-year annual mean aerosols by multiplying them by the ratio at every grid point.

The MERRA-calibrated aerosols (sulfate, BC, POM, dust and sea salt) and SOA need to be added together using reasonable weights to be a useful proxy for $PM_{2.5}$. In MAM3, sulfate is partially in the form of $NH_4HSO_4$, so that fraction of sulfate is already prescribed. So, a multiplication factor for sulfate aerosols is needed since the species tracer in CESM outputs is the sulfate ion. Nitrate aerosol is not simulated in the MAM output we have, but it can be important on regional scales, especially in East Asia (Li *et al* 2015), where it is expected to be increasingly important in the future due to reductions in sulfur dioxide and increases in nitrogen-oxides emissions. Dust aerosols span a broad size range from 0.058 $\mu$m to 3.65 $\mu$m (dry particle diameter), and the use of geometric diameter rather than aerodynamic diameter in MAM3 may overestimate the contribution of dust aerosols to the total $PM_{2.5}$ (Yang *et al* 2022). The proportions of aerosols in different size ranges will be affected by physical and chemical reactions, such as coagulation, and also temperature, humidity and ultraviolet radiation, which will not be perfectly simulated by CESM. Therefore, reasonable weights for each aerosol are necessary to calibrate the contribution of each aerosol to the total $PM_{2.5}$ concentrations.

We apply multiple linear regression (MLR) to create statistically reasonable weights for each aerosol. The target data is the satellite-derived $PM_{2.5}$ (van Donkelaar *et al* 2021) and the MERRA-calibrated 80-year aerosols are potential explanatory variables. Many aerosols have common sources and are well-correlated, for example sulfate, BC and POM in China and India. So, to reduce collinearity and avoid over-fitting the MLR model, we did a correlation analysis of MERRA-calibrated aerosols before the MLR process (Sheet 1 in Supplementary Excel). Aerosol components that are correlated better than 0.7 (and with $p < 0.01$) over a specific country were combined into a single explanatory variable (see Supplementary Excel) for the MLR model for that country. In practice, the sources of most dust and sea salt aerosols are different from sulfate, BC, POM and SOA, and so we only considered combinations of sulfate, BC, POM and SOA, keeping dust and sea salt aerosols as independent variables. We also considered the specific spatial distributions of each aerosol (figure 2) to test or correct the combination of different aerosols. In our 13-country example dataset we will discuss exceptions to these general rules, as here, in the case of Peru where POM and SOA are highly correlated ($R > 0.9$), and although the correlation coefficients between BC and POM or SOA are 0.67 or 0.51 ($p < 0.01$), and hence smaller than the criterion of $R > 0.7$, the spatial distributions of BC, POM and SOA are quite similar to each other (figure 2). So, in this case, we combined BC, POM and SOA into a single component for Peru.

The regression method we use (Matlab: lsqnonneg.m, Lawson & Hanson 1974) is an optimization code rather than a plain linear regression since we specify only positive coefficients for the explanatory variables for physical consistency. The 100-year CESM-simulated aerosols come from a climate equilibrium experiment for the year 2020, which means that a grid point actually has one value for a specific aerosol as the rest of the values are random scattered around that value. So, country-specific MLR models were built using 2000–2019 average satellite-derived $PM_{2.5}$ (van Donkelaar *et al* 2021) as the dependent variable, and the MERRA-calibrated 80-year average aerosols as explanatory variables, after reducing them where possible to account for collinearity as described above (equation (2)). The number (N) of the whole grid points in a country (table S2) is the number of data points in the MLR. According to the One in Ten Rule (Harrell *et al* 1984, Harrell *et al* 1996, Peduzzi *et al* 1996), N should be greater than 10 times the number of explanatory variables (minimum is 3 here) to ensure the reliability and predictive power of the regression model. The constant term $\epsilon$ for the MLR formula was ignored and the slope coefficients $\beta$ were required to be larger than zero, and these were taken as the weights for each aerosol. Then, the bias-corrected $PM_{2.5}$ concentrations can be represented by the weighted sum of the six aerosols (equation (3)). We also estimated the contributions of the MERRA-calibrated aerosols to the total bias-corrected $PM_{2.5}$ concentrations for a specific country as equation (4) (Wang *et al* 2022).

$$y_{i(r)} = \sum_{k=1}^{m(r)} x_{k,i(r)} \times \beta_{k,r} + \epsilon \qquad i = 1, 2, \cdots, N(r) \tag{2}$$

$$PM_{2.5,i(r)} = \sum_{g=1}^{6} Aerosol_{g,i(r)} \times w_{g,r} \tag{3}$$

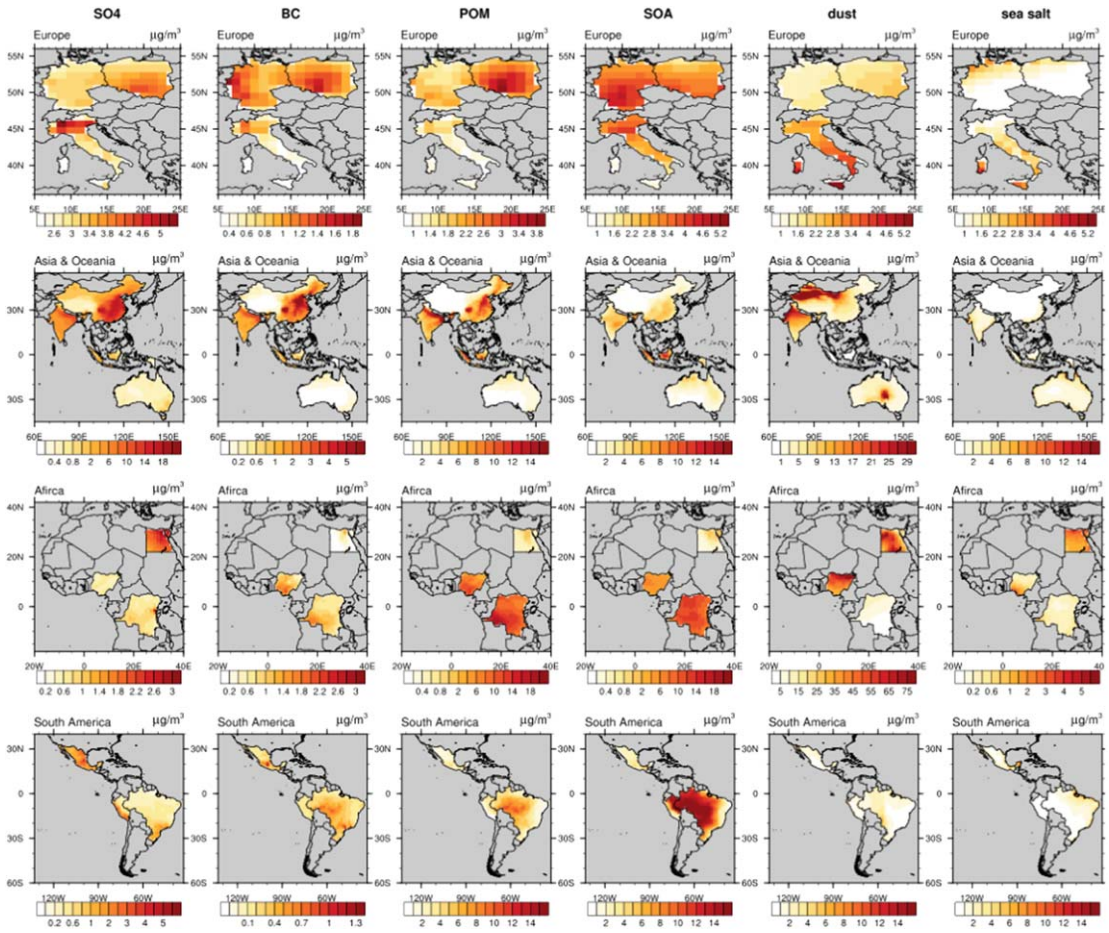


**Figure 2.** The average of MERRA-calibrated 80-year annual mean surface aerosol concentrations for different countries worldwide.

$$C_{g,r} = \frac{w_{g,r} \times \overline{Aerosol_{g,r}}}{\sum_{g=1}^{6} w_{g,r} \times \overline{Aerosol_{g,r}}} \tag{4}$$

Here, $y_{i(r)}$ is the satellite-derived $PM_{2.5}$ concentration of grid cell $i$ located in country $r$. N($r$) is the total number of grid points of country $r$. $x_{k,i(r)}$ is the concentration of component $k$ combined by the MERRA-calibrated aerosols. $m(r)$ is the number of components in country $r$. $\beta_{k,r}$ is the slope coefficient for component $k$ in country $r$. $\epsilon$ is the residuals. $PM_{2.5,i(r)}$ is the corrected $PM_{2.5}$ concentration of grid cell $i$ located in country $r$. $Aerosol_{g,i(r)}$ represents the concentration of Aerosol $g$ corrected by MERRA-2 data. $w_{g,r}$ is the regression coefficient for aerosol $g$ in country $r$. $\overline{Aerosol_{g,r}}$ is the regional mean MERRA-calibrated aerosol $g$ for country $r$, and $C_{g,r}$ is its contribution to regional mean corrected $PM_{2.5}$ for this country.

To assess the correction effect of $PM_{2.5}$ worldwide, we used the same method for the 13 representative countries to correct $PM_{2.5}$ in the six continents (Europe, Asia, Oceania, Africa, North America and South America) but based on continent-specific MLR coefficients ($r$ in equations (2) and (3) indicating a specific continent in this case).

For validating the correction method, we used the grid point ratios of the simulated 80-year aerosols to MERRA-2 to correct the remaining 20-year aerosols, and then applied the regression coefficients to obtain the corrected $PM_{2.5}$ (equation (3)). The corrected $PM_{2.5}$ concentrations in the validation data sets are compared with satellite observations for the 13 representative countries in figure 3, and for the whole globe in figure S3.

## 3. Results

The bias-corrected $PM_{2.5}$ in figures 3 and S3 show similar spatial distributions as the satellite-derived $PM_{2.5}$, and reproduced well the high concentrations in northern China, the Indo-Gangetic plains, Democratic Republic of Congo and northwestern Brazil, which were severely underestimated before correction (figure S1). However, the corrected $PM_{2.5}$ concentrations show a significant underestimation in eastern Nigeria and an overestimation in western Nigeria. This is a consequence of the different relations between east and west for the variables in the

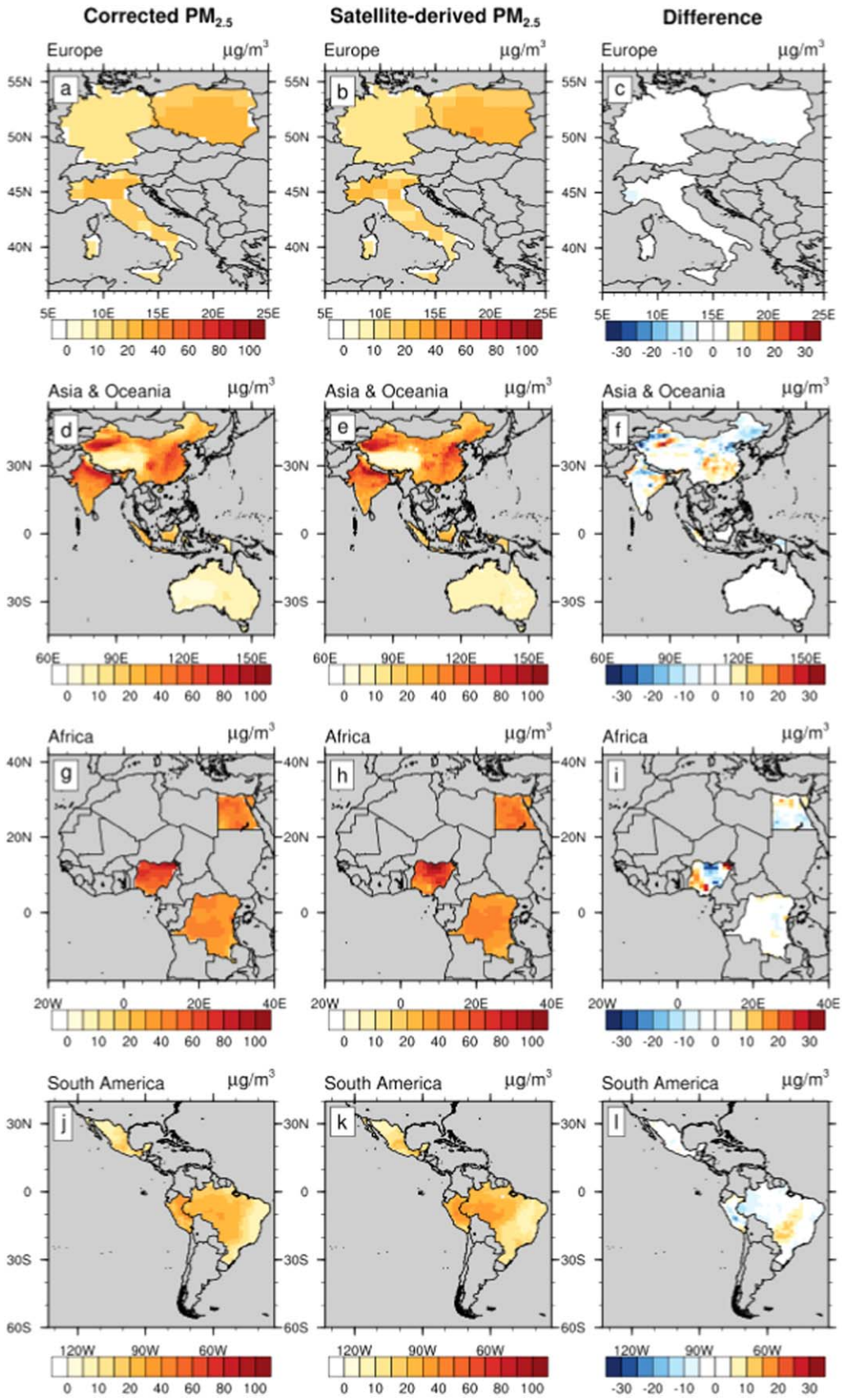


**Figure 3.** Model verification. Left column: bias-corrected 20-year mean surface $PM_{2.5}$ concentrations (the average of five validation data set selections) with horizontal resolution 0.9° × 1.25°, Middle column: 20-year mean satellite-derived surface $PM_{2.5}$ (van Donkelaar *et al* 2021) with horizontal resolution 0.1° × 0.1° over 2000–2019. Right column: differences (simulated $PM_{2.5}$ minus observations). Each row shows the representative countries in Europe, Asia, Africa and the Americas.

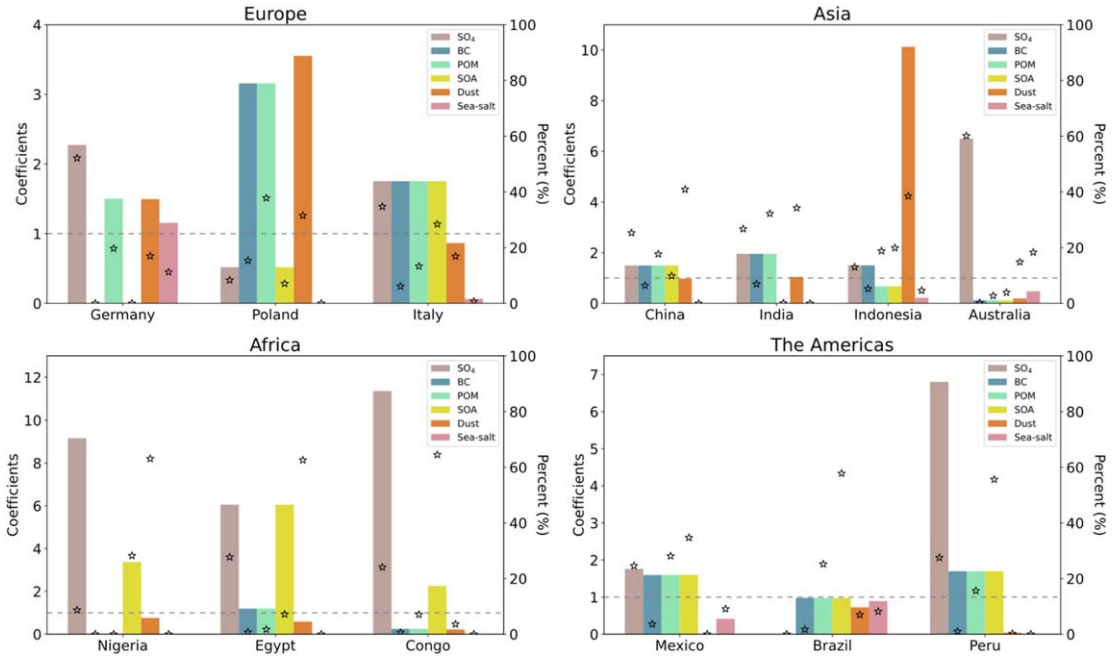


**Figure 4.** The MLR coefficients of each aerosol for the representative countries in Europe, Asia, Africa and the Americas. The pentagrams on the bars indicate the contribution of each aerosol to total $PM_{2.5}$ concentration for a specific country. A value of unity (dashed line) indicates that the MERRA-calibrated 80-year aerosols are the same as the MLR model. Higher coefficients indicate the MLR contribution of the aerosol to total $PM_{2.5}$ concentration is underestimated in this country. The absence of bars indicates that this aerosol was not important for the spatial distributions of $PM_{2.5}$ in this country. The uncertainty in the MLR coefficients deduced from the variations between the 5 random calibration sets is smaller than 0.01.

MLR regression. This can be related to the different pattern of surface winds in the country (figure S4b). Despite this, the correction method significantly reduces the differences between the regional mean $PM_{2.5}$ concentrations simulated by CESM and the observed one (van Donkelaar *et al* 2021). The bias-corrected $PM_{2.5}$ concentrations in most of the 13 representative countries are within 5% of observations (table S1), and the 13-country ensemble RMSD (equation S1) of simulated regional mean $PM_{2.5}$ from observations at country scale is reduced from 115.8% to 3.3%. The 6-continent RMSD of bias-corrected mean $PM_{2.5}$ based on continent-specific MLR models from the observations is 9.7% compared with 74.7% with no bias correction. This decreases to 8.6% if we use the 13 representative countries bias-corrected $PM_{2.5}$ within the continental models (table S1). So, while country-specific MLR models are more accurate than continental-wide MLR models, differences from observations are still lowered by a factor of 7.

The MLR coefficients of each aerosol for the 13 countries are shown in figure 4 and provide information on the contribution of aerosols to total local $PM_{2.5}$. In the case of Indonesia, for example, the contribution of MERRA-calibrated sulfate, BC and dust aerosols to the total $PM_{2.5}$ were underestimated, while that of POM, SOA and sea salt aerosols were overestimated. The large coefficient of dust aerosols means that the importance of dust aerosols to Indonesia's actual $PM_{2.5}$ distributions was highly underestimated. The source of the dust in this case may be attributed to transport from the deserts of Australia towards Indonesia (figure S4a), or seasonal biomass burning due to frequent fires occurring in degraded forests and peatlands (Lestari and Mauliadi 2009, Siregar *et al* 2022). The large overestimations of sea salt (with a coefficient of 0.2) in Indonesia, may be related to the simulation biases of sea surface temperatures and sea ice, which can affect the distribution of global sea salt aerosols (Jiang *et al* 2021).

We also estimated the relative importance of different aerosols to total $PM_{2.5}$ in specific countries. For Germany, Italy and Australia, the contributions of sulfate are larger than that of any other aerosols, and the largest contributions of SOA can be seen in the Democratic Republic of Congo and the three American countries. In Nigeria and Egypt, dust aerosol contributed largest to local $PM_{2.5}$. Comparing the spatial distributions between each aerosol and the bias-corrected $PM_{2.5}$ (figures 2 and 3), we found that the spatial distributions of sulfate, SOA or dust aerosol dominated the local $PM_{2.5}$ distributions for the above countries. However, for China, India and Indonesia, although dust aerosols made the largest overall contribution to total $PM_{2.5}$, the country spatial distribution of $PM_{2.5}$ was more similar to that of sulfate, BC, POM or SOA than to dust aerosols, hence we may expect that the $PM_{2.5}$ variance explained by these variables would be greater than by dust.

## 4. Discussion and conclusion

The long-range transport of air pollutants over different countries or even continents is important in the evaluation of future human health and socio-economic development pathways. Global climate models are a potentially excellent tool to simulate this process and potential impacts (Ran *et al* 2023). However, global climate models such as CESM are not good at simulating aerosol and $PM_{2.5}$ concentrations due to their relatively coarse resolutions and oversimplified representation of chemical processes compared with regional climate models. Thus, correcting the simulation biases in global climate models is a legitimate concern. In this work, we introduced and evaluated a multi-step correction method based on correlation analysis and multiple linear regression modelling for $PM_{2.5}$ concentrations simulated by an earth system model. Our results show that the correction method can considerably improve the simulations of the spatial distributions and magnitudes of $PM_{2.5}$ concentrations worldwide, reducing the root-mean-square differences of simulated $PM_{2.5}$ at continental scales by a factor of 7.

Randomizing the calibration and validation datasets shows that the coefficients derived from the model are robust. However, coefficients can be changed significantly by the selection of aerosol combinations for the MLR process. This is a consequence of the over-fitting caused by collinearity in the fitting variables. Additionally, manual checking of the spatial pattern within countries is useful to avoid effects such as different climate factors driving aerosol species across a country (e.g. as we illustrated for Nigeria). But even in these cases, the MLR bias correction still offers improvements over the climate model simulations. One benefit of using a simple tool like the MLR model to produce coefficients for each aerosol is that the spatially complex and variable classification of internal mixed modes for each aerosol are handled implicitly. These aerosol modes are greatly affected by physical and chemical reactions amongst themselves, as well as the atmospheric background conditions. Furthermore, these conditions will also change temporally as states make efforts to reduce $PM_{2.5}$ concentrations in future, decarbonize their economies and alter their energy production balance.

Here we take the number of whole grid points in a country (or continent) as the number of data points in each MLR model. Most of the 13 representative countries have $N > 10$ times the number of independent variables (the $x$ in equation (2)) in the MLR. For some small countries such as Kuwait, Lebanon and Palestine, this country-based method has very small N, and may not provide reliable information about local emission features.

We choose countries rather than regions having the minimum N grid points to build the MLR model because countries are the most important administrative areas where the emissions are always related to local policies and economy. The MLR coefficients are highly aerosol- and country-dependent, showing the relative bias in aerosol simulation and modifying the understanding of the importance of species contributions to total $PM_{2.5}$ concentrations in an individual country. We discussed in section 3 how underestimated dust aerosols in Indonesia contributed the most (39%) to local $PM_{2.5}$ and could come from seasonal biomass burning or long-range transport from Australian deserts (figure S4a). Unexpectedly, results for Poland also suggest a relatively large contribution of dust aerosols (32%) to overall $PM_{2.5}$, which is comparable to that of POM (37%). There are significant local sources of dust in Poland especially in heavily industrialized Silesia and southern Poland, and advected dust from the Ukrainian steppe and Saharan sources (Milinevsky *et al* 2020) that varies greatly on seasonal scales and from year to year (Milinevsky *et al* 2020, Szczepanik *et al* 2021). This combination of sources in Poland differentiates it from countries in western and southern Europe (Di Iorio *et al* 2009, Israelevich *et al* 2012), where dust is relatively less important, e.g., Germany and Italy in figure 4.

The MLR derived $PM_{2.5}$ in this work may be useful for various end-users of climate model output, and may be improved by local insights and observations to reduce bias related to aerosol sources, transport and local climate interactions. The small computational costs and time required to use the method means that adjustments can easily be made to accommodate better knowledge and provide improved air pollutant output from global climate models at regional and higher resolution.

## Acknowledgments

This study was supported by the National Natural Science Foundation of China (U21A6001 and 42075044), CMA Key Open Laboratory of Transforming Climate Resources to Economy (2023002K) and the Second Tibetan Plateau Scientific Expedition and Research Program (2019QZKK0103).

## Data availability statement

CESM-simulated aerosol concentrations used in this paper are available online (https://doi.org/10.5281/zenodo.8382318). Other processed data can be found in the supplementary materials.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## ORCID iDs

Qi Ran https://orcid.org/0000-0002-9051-2928

## References

Bai Y-Q, Wang Y, Kong S-F, Zhao T-L, Zhi X-F, Zheng H and Xiong J 2021 Modelling the effect of local and regional emissions on PM2.5 concentrations in wuhan, China during the COVID-19 lockdown *Adv. Clim. Chang. Res.* **12** 871–80

Banks A, Kooperman G J and Xu Y Y 2022 Meteorological influences on anthropogenic PM2.5 in future climates: species level analysis in the community earth system model v2 *Earths Future* **10** 20

Bran S H and Srivastava R 2017 Investigation of PM2.5 mass concentration over India using a regional climate model *Environ. Pollut.* **224** 484–93

Cheng F-Y, Feng C-Y, Yang Z-M, Hsu C-H, Chan K-W, Lee C-Y and Chang S-C 2021 Evaluation of real-time PM2.5 forecasts with the WRF-CMAQ modeling system and weather-pattern-dependent bias-adjusted PM2.5 forecasts in Taiwan *Atmos. Environ.* **244** 117909

Di Iorio T, di Sarra A, Sferlazzo D M, Cacciani M, Meloni D, Monteleone F and Fiocco G 2009 Seasonal evolution of the tropospheric aerosol vertical profile in the central Mediterranean and role of desert dust *J. Geophys. Res.* **114** D02201

Djalalova I, Wilczak J, McKeen S, Grell G, Peckham S, Pagowski M and Mathur R 2010 Ensemble and bias-correction techniques for air quality model forecasts of surface O3 and PM2.5 during the TEXAQS-II experiment of 2006 *Atmos. Environ.* **44** 455–67

Gao Y, Liu X, Zhao C and Zhang M 2011 Emission controls versus meteorological conditions in determining aerosol concentrations in Beijing during the 2008 Olympic Games *Atmos. Chem. Phys.* **11** 12437–51

Global Modeling and Assimilation Office (GMAO) 2015 MERRA-2 tavgM_2d_aer_Nx: 2d, Monthly Mean, Time-averaged, Single-level, Assimilation, Aerosol Diagnostics V5.12.4 [Dataset]. Goddard Earth Sciences Data and Information Services Center (GES DISC), Greenbelt, MD, USA (https://doi.org/10.5067/FH9A0MLJPC7N)

Harrell F E Jr, Lee K L, Califf R M, Pryor D B and Rosati R A 1984 Regression modelling strategies for improved prognostic prediction *Stat. Med.* **3** 143–52

Harrell Jr F E, Lee K L and Mark D B 1996 Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors *Stat. Med.* **15** 361–87

He H, Liang X-Z and Wuebbles D J 2018 Effects of emissions change, climate change and long-range transport on regional modeling of future U.S. particulate matter pollution and speciation *Atmos. Environ.* **179** 166–76

Huang J, McQueen J, Wilczak J, Djalalova I, Stajner I, Shafran P and Delle Monache L 2017 Improving NOAA NAQFC PM2.5 Predictions with a bias correction approach *Weather Forecast.* **32** 407–21

Hur S K, Ho C H, Kim J, Oh H R and Koo Y S 2021 Systematic bias of WRF-CMAQ PM10 simulations for Seoul, Korea *Atmos. Environ.* **244** 117904

Hurrell J W, Holland M, Gent P, Ghan S, Kay J, Kushner P and Marshall S 2013 The community earth system model: a framework for collaborative research *Bull. Am. Meteorol. Soc.* **94** 1339–60

Israelevich P, Ganor E, Alpert P, Kishcha P and Stupp A 2012 Predominant transport paths of saharan dust over the mediterranean sea to europe *J. Geophys. Res.* **117** D02205

Jiang B, Xie Z, Lam P K S, He P, Yue F, Wang L and Wu X 2021 Spatial and temporal distribution of sea salt aerosol mass concentrations in the marine boundary layer from the arctic to the antarctic *JGR Atmos.* **126** e2020JD033892

Lestari P and Mauliadi Y D 2009 Source apportionment of particulate matter at urban mixed site in Indonesia using PMF *Atmos. Environ.* **43** 1760–70

Li J, Wang W-C, Liao H and Chang W 2015 Past and future direct radiative forcing of nitrate aerosol in East Asia *Theor. Appl. Climatol.* **121** 445–58

Liu J and Xing J 2022 Identifying contributors to PM2.5 simulation biases of chemical transport model using fully connected neural networks *J. Adv. Model. Earth Syst.* **15** e2021MS002898

Liu X 2012 Toward a minimal representation of aerosols in climate models: description and evaluation in the community atmosphere model CAM5 *Geosci. Model Dev.* **5** 709–39

McMillan N J, Holland D M, Morara M and Feng J 2009 Combining numerical model output and particulate data using Bayesian space–time modeling *Environmetrics* **21** 48–65

Milinevsky G, Miatselskaya N, Grytsai A, Danylevsky V, Bril A, Chaikovsky A and Serozhkin Y 2020 Atmospheric aerosol distribution in 2016–2017 over the eastern european region based on the GEOS-Chem model *Atmosphere* **11** 722

Peduzzi P, Concato J, Kemper E, Holford T R and Feinstein A R 1996 A simulation study of the number of events per variable in logistic regression analysis *J. Clin. Epidemiol.* **49** 1373–9

Qin Y 2003 *Fundamentals of Atmospheric Chemistry* (Beijing, China: China Meteorological Press)

Ran Q, Lee S-Y, Zheng D, Chen H, Yang S, Moore J C and Dong W 2023 Potential health and economic impacts of shifting manufacturing from China to Indonesia or India *Sci. Total Environ.* **855** 158634

Siregar S, Idiawati N, Lestari P, Berekute A K, Pan W-C and Yu K-P 2022 Chemical composition, source appointment and health risk of PM2.5 and PM2.5-10 during forest and peatland fires in riau, Indonesia *Aerosol Air Qual. Res.* **22** 220015

Szczepanik D M, Stachlewska I S, Tetoni E and Althausen D 2021 Properties of saharan dust versus local urban dust—a case study *Earth Space Sci.* **8** e2021EA001816

Textor C, Schulz M, Guibert S, Kinne S, Balkanski Y, Bauer S and Tie X 2006 Analysis and quantification of the diversities of aerosol life cycles within AeroCom *Atmos. Chem. Phys.* **6** 1777–813

Tilmes S, Lamarque J F, Emmons L K, Kinnison D E, Ma P L, Liu X and Val Martin M 2015 Description and evaluation of tropospheric chemistry and aerosols in the community earth system model (CESM1.2) *Geosci. Model Dev.* **8** 1395–426

van Donkelaar A *et al* 2018 Global annual PM2.5 grids from MODIS, MISR and SeaWiFS aerosol optical depth (AOD) with GWR, 1998–2016. NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4ZK5DQS

van Donkelaar A, Hammer M S, Bindle L, Brauer M, Brook J R, Garay M J and Martin R V 2021 Monthly global estimates of fine particulate matter and their uncertainty *Environ. Sci. Technol.* **55** 15287–300

Wang J, Moore J C and Zhao L 2022 Changes in apparent temperature around the Beijing-Tianjin megalopolis under greenhouse gas and stratospheric aerosol injection scenarios *Earth Syst. Dynam. Discuss.* **2022** 1–24

Xu Y and Lamarque J-F 2018 Isolating the meteorological impact of 21st century GHG warming on the removal and atmospheric loading of anthropogenic fine particulate matter pollution at global scale *Earth's Future* **6** 428–40

Yang A, Tan Q, Rajapakshe C, Chin M and Yu H 2022 Global premature mortality by dust and pollution PM2.5 estimated from aerosol reanalysis of the modern-era retrospective analysis for research and applications, version 2 *Front. Environ. Sci.* **10** 975755

Yang J, Kang S, Ji Z, Tripathee L, Yin X and Yang R 2020 Investigation of variations, causes and component distributions of PM2.5 mass in China using a coupled regional climate-chemistry model *Atmos. Pollut. Res.* **11** 319–31