air

MDPI

Article

An Evaluation of Ground-Level Concentrations of Aerosols and Criteria Pollutants Using the CAMS Reanalysis Dataset over the Himawari-8 Observational Area, Including China, Indonesia, and Australia (2016–2023)

Miles Sowden

Sigma Theta, Scarborough 6019, Australia; miles.sowden@sigmatheta.biz

Abstract: This study assesses the performance of the Copernicus Atmosphere Monitoring Service (CAMS) reanalysis dataset in estimating ground-level concentrations (GLCs) of aerosols and criteria pollutants across the Himawari-8 observational area, covering China, Indonesia, and Australia, from 2016 to 2023. Ground-based monitoring networks in these regions are limited in scope, making it necessary to rely on satellite-derived aerosol optical depth (AOD) as a proxy for GLCs. While AOD offers broad coverage, it presents challenges, particularly in capturing surface-level pollution accurately during episodic events. CAMS, which integrates satellite data with atmospheric models, is evaluated here to determine its effectiveness in addressing these issues. The study employs square root transformation to normalize pollutant concentration data and calculates monthly–hourly long-term averages to isolate pollution anomalies. Geographically weighted regression (GWR) and Jacobian matrix (dY/dX) methods are applied to assess the spatial variability of pollutant concentrations and their relationship with meteorological factors. Results show that while CAMS captures large-scale pollution episodes, such as the 2019/2020 Australian wildfires, discrepancies in representing GLCs are apparent, especially when vertical aerosol stratification occurs during short-term pollution events. The study emphasizes the need for integrating CAMS data with higher-resolution satellite observations, like Himawari-8, to improve the accuracy of real-time air quality monitoring. The findings highlight important implications for public health interventions and environmental policy-making, particularly in regions with insufficient ground-based data.

Keywords: CAMS reanalysis; ground-level concentrations (GLCs); geographically weighted regression (GWR); Jacobian matrix; data normalization; spatiotemporal analysis; aerosol

Citation: Sowden, M. An Evaluation of Ground-Level Concentrations of Aerosols and Criteria Pollutants Using the CAMS Reanalysis Dataset over the Himawari-8 Observational Area, Including China, Indonesia, and Australia (2016–2023). *Air* **2024**, *2*, 419–438. https://doi.org/10.3390/air2040024

Academic Editors: Kamila Widziewicz-Rzońca, Patrycja Rogula-Kopiec, Jan Stefan Bihałowicz and Ling Tim Wong

Received: 17 August 2024
Revised: 24 November 2024
Accepted: 28 November 2024
Published: 5 December 2024

1. Introduction

Ground-level concentrations (GLCs) of air pollutants are critical for human health [1], yet accurate monitoring remains challenging due to sparse surface measurements and limitations in modeling [2]. Globally, air pollution is a significant issue, impacting public health, climate change, and economies worldwide. Effective monitoring and management of GLCs are essential for mitigating these impacts. This study evaluates the Copernicus Atmosphere Monitoring Service (CAMS) reanalysis dataset [3] over an eight-year period for the Himawari-8 Observational Area [4], which includes China, Indonesia, and Australia. These regions experience significant pollution from both natural and anthropogenic sources [1,5], making accurate GLC monitoring essential for public health and policy-making. Current methods for GLC monitoring, including surface measurements [6] and traditional modeling [7,8], often carry significant uncertainties due to the limited availability of ground-based measurements, particularly in remote areas. The CAMS reanalysis dataset, which integrates multiple sources of data, may provide a comprehensive resource for determining pollution dynamics in regions where traditional methods are insufficient [9].

This study aims to assess the accuracy and resolution of CAMS in capturing annual, seasonal, and episodic pollution trends. Key results include statistical summaries, compositional analyses, vertical profile analysis of GLC to aerosol optical depth (AOD) ratios, and identification of significant pollution events and trends. This study considered a recent review advocating for increased use of satellite remote sensing for monitoring air pollution and sustainable development goals [10], and it examined the effectiveness of geographically weighted regression (GWR) and Jacobian matrix (dY/dX) methods to determine the accuracy of pollutant concentration determinations as a function of aerosol speciation and their correlations with meteorological factors [11].

1.1. Context and Importance of Air Pollution Monitoring and Modeling

Effective air quality monitoring and modeling are crucial for protecting public health and formulating environmental policies. Understanding ground-level concentrations (GLCs) of criteria air pollutants, such as particulate matter (PM_1, $PM_{2.5}$, and PM_{10}) and gases (ozone, nitrogen oxides, and sulfur dioxide), which directly impact human health in urban areas through higher pollution levels [1], is critical in reducing the adverse impact of air pollution. However, traditional ground-based monitoring systems cover limited areas, lack real-time data reporting, and often fail to capture short-term episodic pollution events, such as wildfires, volcanic eruptions, or seasonal agricultural burning [2]. These systems also face logistical and financial challenges in maintaining widespread sensor networks, particularly in remote or underdeveloped regions [12].

The Copernicus Atmosphere Monitoring Service (CAMS) reanalysis dataset integrates satellite remote sensing, ground-based measurements, and atmospheric modeling with the aim of creating a detailed and consistent global dataset [9] that provides comprehensive air quality information, tracking air quality across local and regional scales. CAMS data have been claimed to be able to monitor air quality trends and detect transient pollution spikes, and may provide data essential for timely public health responses [3].

The Himawari-8 satellite, which has been operational since mid-2015, provides ten-minute near-real-time monitoring of the observational area at a 2 km resolution (at nadir) [4]. These data offer the ability to provide near-instantaneous updates on events, like wildfires, dust storms, and volcanic eruptions, with unparalleled temporal resolution [5]. However, Himawari data are not as yet utilized in CAMS reanalysis.

The study period spans eight years, from 2016 to 2023, during which significant global events, like the 2019/2020 Australian bushfires and the COVID-19 pandemic, occurred. These occurrences provide scenarios to observe natural and anthropogenic impacts on air quality, which is essential for evaluating CAMS efficacy in varied environmental conditions. By addressing the challenges of sparse surface measurements and modeling uncertainties, this study aims to enhance understanding of air quality dynamics and support targeted public health interventions and robust environmental policies through advocating for greater use of the CAMS data [7].

1.2. Challenges in Modern Air Quality Monitoring and Determining GLCs

As discussed, modern air quality monitoring systems struggle to provide comprehensive, real-time data. The limited number of publicly available ground-based stations leads to data gaps in the coverage and underreporting of air quality issues. These fixed monitoring stations focus on common pollutants, like sulfur dioxide, nitrogen oxides, and particulate matter, while overlooking harmful substances, like volatile organic compounds, and ultrafine particles. These overlooked pollutants can impact health, and their omission creates gaps in public health data and an inability to adapt to rapid changes in air quality due to weather, traffic, and industrial activities. In contrast, continuous, real-time monitoring is challenging due to the high costs of expanding and maintaining a dense network of monitoring stations.

Air pollution modeling has evolved with recent approaches, such as land use regression (LUR) [13], which uses the homogeneous assumption Y = f(X), and geographically

weighted regression (GWR), which operates on a per-point basis $Y_i = f_i(X_i)$ where f_i is a function of the temporal correlation of Y_i with X_i [14]. These methods predict pollution levels but often focus on long-term averages [15], overlooking short-term incidents crucial for public health, i.e., they are useful for determining chronic but not acute events. They fail to incorporate dynamic variables, such as the atmospheric pressure changes that drive wind patterns and air turbulence which drive pollutant dispersion. Regressors, like terrain height [16], wind direction through north [17], or categorical data [18], can misrepresent pollution drivers due to their static nature and co-correlation effects (e.g., terrain height does not influence air pollution levels as PM_{10} is not a function of height, but terrain does affect the siting of emission sources. These models may also ignore source–receptor contributions, assuming a homogeneous concentration [19] proportional to wind speed and ignoring emission strength [20] when atmospheric dynamics are overlooked. This study has examined the application of the Jacobian matrix, which evaluates the sensitivity of model outputs to various input parameters [21,22], with the aim of refining pollution models' predictive capabilities and identify which factors affect air quality.

In addition to monitoring and statistical approaches, dispersion models (including the output of CAMS) rely on input data that can be inaccurate due to changes in weather, traffic, and industrial activity. Remote sensing techniques, particularly those using satellite data, offer broader coverage but struggle to accurately assess ground-level pollutants due to their columnar nature. Aerosol optical depth (AOD) measurements from remote sensing estimate total columnar particulate matter concentrations but may face difficulties due to variability in aerosol types and sizes, atmospheric conditions, and cloud cover uncertainties [23]. Furthermore, AOD data requires complex algorithms to infer ground-level concentrations, often leading to uncertainties.

1.3. CAMS Reanalysis: Rationale and Limitations

The Copernicus Atmosphere Monitoring Service (CAMS) reanalysis integrates satellite and ground-based observational data with numerical models, creating a statistically consistent atmospheric dataset that uses 4D-Var data assimilation to enhance model accuracy [3]. Studies have used ERA5 reanalysis data to predict tropical cyclones with high accuracy, demonstrating reanalysis data's utility in severe weather forecasting [24]. CAMS provides meteorological variables, such as wind components, temperature, mean sea level pressure, and relative humidity. Studies have confirmed the CAMS data's reliability for environmental AI and stochastic modeling [25].

CAMS offers atmospheric composition data crucial for air quality and climate studies, including gases, like carbon monoxide, isoprene, formaldehyde, nitrogen dioxide, nitrogen monoxide, ozone, and sulfur dioxide, and aerosols, such as dust, sea salt, hydrophilic and hydrophobic black carbon, and organic matter [3]. Meteorological data in CAMS are more reliable for model validation compared to atmospheric chemistry data due to their easier measurability, greater number of measurement sites, and global data assimilation. These meteorological data support methodology testing and predictive capability enhancement [26,27]. However, the accuracy of CAMS outputs depends heavily on the quality and consistency of the input data, including the specific satellite products used [3,28], as discussed in the previous section.

1.4. Specific Problems in Environmental Modeling

In addition to the lack of surface monitoring and the dependency of models on input data, there are further methodological challenges which need to be addressed to ensure accurate environmental modeling and analysis. Data have to be normalized, multifactorial regression techniques must be enhanced, and methods to evaluate large spatiotemporal datasets need development.

Normalizing atmospheric concentration data, which may vary by four to five orders of magnitude across a domain, is crucial for accurately modeling pollutant dispersion across diverse geographical scales. Many studies simply attempt to determine annual

averages ignoring short term chronic events. Sources and environmental conditions vary widely in location and magnitude, necessitating standardization to manage the exponential decay that occurs in pollutant dispersal—as described by Gaussian dispersion models. Normalization mitigates biases and enhances the comparability of datasets from different sources and times. Without normalization, variations in data scales can obscure actual environmental patterns and trends. Employing normalization ensures that conclusions reflect actual atmospheric behaviors rather than fluctuations above a mean [29].

Gaussian dispersion models describe the spread of pollutants in the atmosphere as a bell-shaped curve, proportional to the exponential of the square of the distance from the source [30–32]. The square root transformation addresses the skewed distribution of concentrations typically observed downwind from emission sources. Unlike logarithmic transformations, which can create errors by attempting to process log (0), the square root transformation maintains zero values—sqrt (0) remains zero. This adjustment scales dispersion to a more linear scale, reducing the range and impact of extreme values and simplifying computational analysis. This method furthermore maintains accuracy in datasets with zero values and avoids distortions in environmental data analysis by methods that add plus one (eliminating log (0)) or mentally computing the exponential of a number.

Large spatiotemporal datasets, including CAMS, require accurate validation. Metrics, like root mean square error (RMSE) and mean absolute error (MAE) are typically used to quantify prediction errors [33]. However, some metrics, like RMSE and MAE, are point-based, and do not capture spatiotemporal interdependencies within large datasets. Advanced techniques, such as spatiotemporal cross-validation, highlight the importance of assessing model robustness [34]. The accuracy of land use regression (LUR) models, which typically assume homogeneity (i.e., pollution concentrations are a global function of parameters), may be improved by considering localized anomalies, like specific source–receptor pairs of high-concentration pollution plumes and high-resolution spatiotemporal data (i.e., dY/dt and dX/dt), rather than simply dY/dX to capture atmospheric dynamics [35].

Visualization tools, like Taylor plots, compare model predictions with actual observations, providing insights into model performance through visual representations of correlation, RMSE, and standard deviation. However, Taylor plots are point-based statistics and do not address the dY/dX variability. Proper understanding of the source of variability is critical in facilitating targeted improvements in environmental modeling and ensuring the realistic representation of observed data [36].

The sensitivity of the predicted to observed parameters is given by the Jacobian matrix (dY/dX) which evaluates the sensitivity of model outputs to various input parameters and identifies which factors may affect air quality and can assist in fine-tuning models to improve predictions [37]. The Jacobian matrix, which is the crux of land use regression models assumes homogeneity; however, geographically weighted regression (GWR), which assesses spatial variability in pollutant concentrations using the localized Jacobian matrix and represents how pollutants disperse and accumulate, does not assume homogeneity [38].

1.5. Research Question and Hypothesis

This research evaluates the capability of CAMS reanalysis data to capture temporal and seasonal air quality trends and major pollution events. It has investigated the accuracy of CAMS in documenting incidents, like the 2019 Australian bushfires, pollution episodes in cities, like Beijing, and emissions reductions during the COVID-19 lockdowns. The hypothesis is that CAMS, through its integration of satellite data, ground-based observations, and advanced modeling, may provide a reliable framework for air quality monitoring. The aim is to evaluate the validity and precision of CAMS model estimates in the absence of empirical verification data, i.e., to determine whether the results are consistent with known events and what the data indicate regarding input sensitivities.

This was achieved by employing and assessing the localized Jacobian matrix (dY/dX), i.e., GWR, to assess the apparent spatial variability in pollutant concentrations and their correlations to meteorological factors. Additionally, vertical profiles of ground-level con-

centration (GLC) to aerosol optical depth (AOD) ratios were analyzed to understand the relationship between ground-level pollution and columnar aerosol measurements. Spatiotemporal pollution episodes with significant short-term impacts on air quality were investigated to evaluate CAMS' response to sudden atmospheric changes caused by extreme weather and large-scale human activities.

2. Materials and Methods

2.1. Study Area and Timeframe

The study focuses on the observational area of the Himawari-8 satellite (Figure 1a), which provides high-spatiotemporal-resolution data. The area was cropped to reduce non-habitable areas (sea, space, and earth curvature), resulting in a dataset approximately 45% the size of the original Himawari-8 dataset (Figure 1b). The selected region extends from 85° to 160° east longitude and from −45° to 50° north latitude, covering Australia, Indonesia, and significant portions of China, each with unique environmental characteristics impacting air quality.

Australia features diverse climates from deserts to forests, affecting local and regional air quality. Topography influences pollutant dispersion, especially from bushfires and mining. Coastal areas, particularly in the southeast, experience marine aerosol influxes that interact with urban pollution. Indonesia, which is an archipelago of over 17,000 islands, experiences volcanic activity and biogenic burning which generate significant amounts of particulate matter, and the country's coastal nature impacts dispersion patterns. China has experienced rapid industrialization and urbanization which have caused severe pollution in cities, like Beijing and Shanghai. Large mountain ranges influence atmospheric circulation, affecting pollutant transport. Seasonal winds and dust storms from northern deserts impact regional air quality.

The dataset covers the eight-year period from 2016 to 2023 and includes major environmental events, like the 2019 Southeast Asia haze, the 2019/2020 Australian bushfires, Melbourne's 2016 pollen events, and industrial reductions during the 2020 COVID-19 lockdowns. These events offer insights into aerosol composition variability and its impact on regional and global climate patterns. The extensive selection of geographical extent and timeframe ensures a detailed analysis of pollution sources and environmental conditions in China, Indonesia, and Australia.

2.2. CAMS Reanalysis Data

This study uses CAMS reanalysis data, integrating extensive observational data with advanced numerical models for a consistent dataset crucial for global air quality and climate studies. The data, accessed on 26 April 2024, include three-hour timesteps saved into eight yearly netCDF files, covering a comprehensive range of atmospheric composition and meteorological variables. The dataset included only single-level data (at the ground level of the 60 available atmospheric pressure levels) for surface and total columnar parameters to determine GLCs and total column statistics. Key variables include wind components, temperature, mean sea level pressure (MSLP), relative humidity, carbon monoxide, nitrogen oxides, sulfur dioxide, and various particulates.

Variables were chosen based on their relevance to the study's goals, which were to examine the temporal and spatial trends in respirable air pollution and to assess responses to significant pollution events. Ground-level air quality was prioritized, excluding vertical profile data and secondary pollutants, to focus on aerosols and criteria pollutants which may provide insights into regional air quality dynamics.

While the study relies on CAMS reanalysis data, supplementary ground-based data from monitoring stations and AERONET were considered but not used. AERONET's 17 Australian sites provide insufficient AOD coverage, and the focus was on GLCs, not columnar AOD. Australian surface data are too sparse, and while China's extensive monitoring network could verify grid variability, data access issues prevented its use. Furthermore, the mismatch between point-based ground data and the CAMS grid size further compli-

cates direct calibration. The study acknowledges these constraints and highlights the need for improved data availability to enhance verification and air quality modeling.

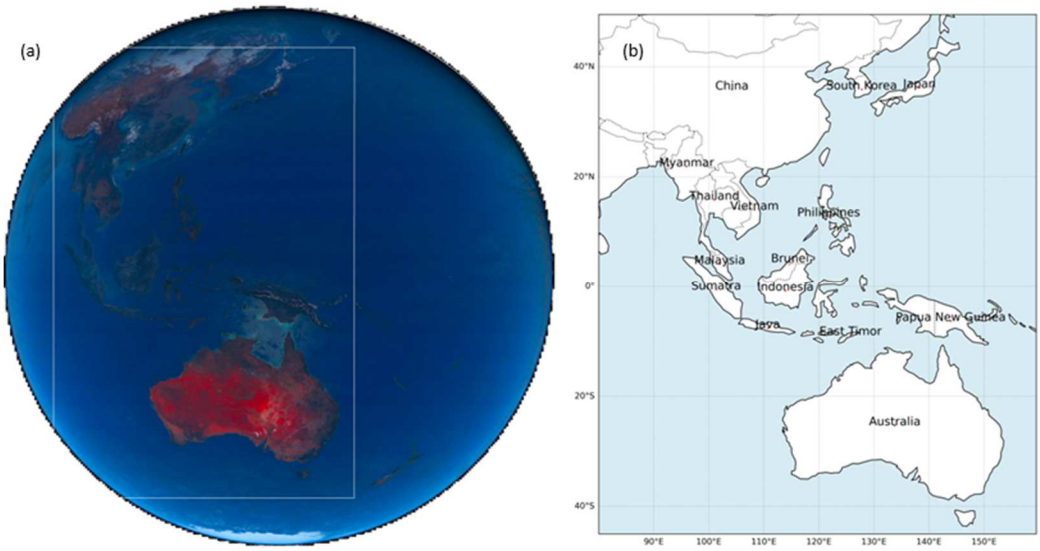

Figure 1. (**a**) Himawari-8 observational area (pseudo true color) and study domain (white-box cropped at (x1 = 500, x2 = 3500; y1 = 500, y2 = 5000 to exclude outer edges), and (**b**) approximate geographic extent of the study area from 85° E to 160° E and 45° S to 50° N displayed as equal area projection.

2.3. Detailed Methodological Approach

This study processed and analyzed atmospheric data from the CAMS Reanalysis Dataset using BASH and Python scripts as provided in the Supplementary Files. Key steps included data normalization, statistical summarization, compositional analysis, anomaly detection, and visualization. Data were processed annually from 2016 to 2023. Ground-level and total column data were downloaded into annual files from the CAMS servers. CDO (climate data operator) commands were used to adjust units and standardize atmospheric parameters to common units [39].

Normalization was performed using the square root transformation to mitigate skewness and handle zero values, thus avoiding computational challenges. Visual examination shows a balanced representation across the domain of concentration levels, highlighting significant pollution areas. Monthly–hourly means were calculated for each month and hour combination and subtracted from the normalized values to reduce background variation and focus on air quality changes.

In-depth statistical analyses were conducted on the consolidated datasets, including metrics, such as maximum, mean, minimum, and standard deviation over the study period. Specifically, we examined the maximum hourly concentration, which is the highest pollutant level recorded in any single hour, the maximum daily concentration, which is the highest average concentration over a 24 h period, and the year mean, which represents the average concentration of pollutants across the entire year. Additionally, the standard deviation ("std dev") was calculated to measure the variability or dispersion of pollutant concentrations from the mean, indicating how much the concentrations fluctuate over time. This was achieved using CDO operations, like timmax, timmean, timmin, and timstd, with data divided into smaller segments (monthly, daily, hourly) using splitmon, splitday, and splithour commands. Monthly datasets examined variations and anomalies and identified significant deviations that indicated unusual pollution events or changes in atmospheric conditions. Composition spatiotemporal analysis identified short-term, seasonal, and long-term pollution trends.

The dataset included concentrations of sulfur dioxide (SO_2), nitrogen oxides (NO_x), carbon monoxide (CO), ozone (O_3), and speciated aerosols, like sulfate, nitrate, ammonium, black carbon, organic carbon, sea salt, and mineral dust. This provided a comprehensive view of pollutant distribution and concentrations over time. Mean sea level pressure (MSLP) and specific humidity were evaluated for their impact on air pollution dynamics. MSLP influences air mass movement and stability, while specific humidity impacts cloud formation and precipitation, crucial for pollutant dispersion and removal. Carbon monoxide (CO) and ground-level ozone (O_3) concentrations were analyzed, focusing on urban and industrial areas. Nitrogen dioxide (NO_2) and sulfur dioxide (SO_2) were studied for their anthropogenic sources and impacts on health and the environment. Nitric oxide (NO) and isoprene were analyzed to understand their roles in atmospheric photochemical reactions. Particulate matter ($PM_{2.5}$ and PM_{10}) concentrations were examined across Asia and Australia to understand their sources and impacts. Aerosol speciation considered black carbon (BC) and organic carbon (OC) aerosols to identify pollution sources, particularly from industrial activities and wildfires. The distributions of sulfate (SU) and sea salt (SS) aerosols were examined to understand their sources and effects.

The relationship between ground-level concentrations (GLCs) and aerosol optical depth (AOD) was analyzed to highlight the limitations of using satellite data for ground-level air quality. These included specific humidity, sea salt, dust aerosol, organic matter, black carbon, and sulfate aerosol. The goal was to assess how the vertical ratio of these components is influenced by environmental factors and to evaluate the feasibility of computing GLC data from AOD columnar data.

Anomalies in NO_2, SO_2, PM_{10}, and organic carbon (OC) were analyzed over an eight-year period to identify significant changes and events affecting air quality. Isopleth maps were used to highlight deviations from typical conditions, providing insights into the impacts of events, like the COVID-19 pandemic and wildfires, on air quality. Maximum hourly concentrations per day for various pollutants were analyzed using time-series plots. The analysis utilized the square root transformation and anomaly removal based on monthly–hourly averages to focus on specific pollution incidents rather than long-term averages. This aimed to identify and analyze acute pollution episodes in major cities across the region.

Geographically weighted regression (GWR) approaches using the Jacobian matrix (dY/dX) and temporal correlation were employed to illustrate the spatial variability between air pollutants and meteorological factors across the study domain. GWR operates on a by-pixel basis, making it ideal for capturing local variations in data. This method determines the Jacobian matrix for selected species-factors, highlighting the complexity of air pollution dynamics and the need for localized air quality strategies. The analysis focused on key meteorological variables and GLCs, including specific humidity (q), mean sea level pressure (msl), temperature in Kelvin (1/T), and wind components (U10m and V10m). Pollutants, such as organic matter (OM, aermr07), black carbon (BC, aermr09), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and particulate matter (PM_1, $PM_{2.5}$, PM_{10}), were included to assess their responses to environmental changes. The spatial and temporal changes in these parameters, their correlations, and the evolution of the Jacobian matrix (dY/dX) over space and time were examined.

3. Results

Given the extensive study domain—covering approximately one-eighth of the globe—along with the eight-year timeframe and the inclusion of multiple pollutants and atmospheric variables, it is impractical to present high-resolution figures within this article. Readers are, therefore, directed to the Supplementary Information, which includes individual images of the compound plots as well as additional plots that could not be included here due to space constraints.

Figure 2 presents the normalized maximum hourly concentrations for NO_2, SO_2, PM_{10}, and OM through isopleth maps. The original, squared, and log-transformed values are

depicted for each pollutant. The normalization process highlights distinct spatial patterns and concentration gradients, which are crucial for identifying high-risk areas. The square root transformation effectively maintains a coherent plume and simplifies the identification of pollution hotspots. This approach ensures that significant pollution incidents are clearly visible, making it easier to focus on areas that require targeted interventions.

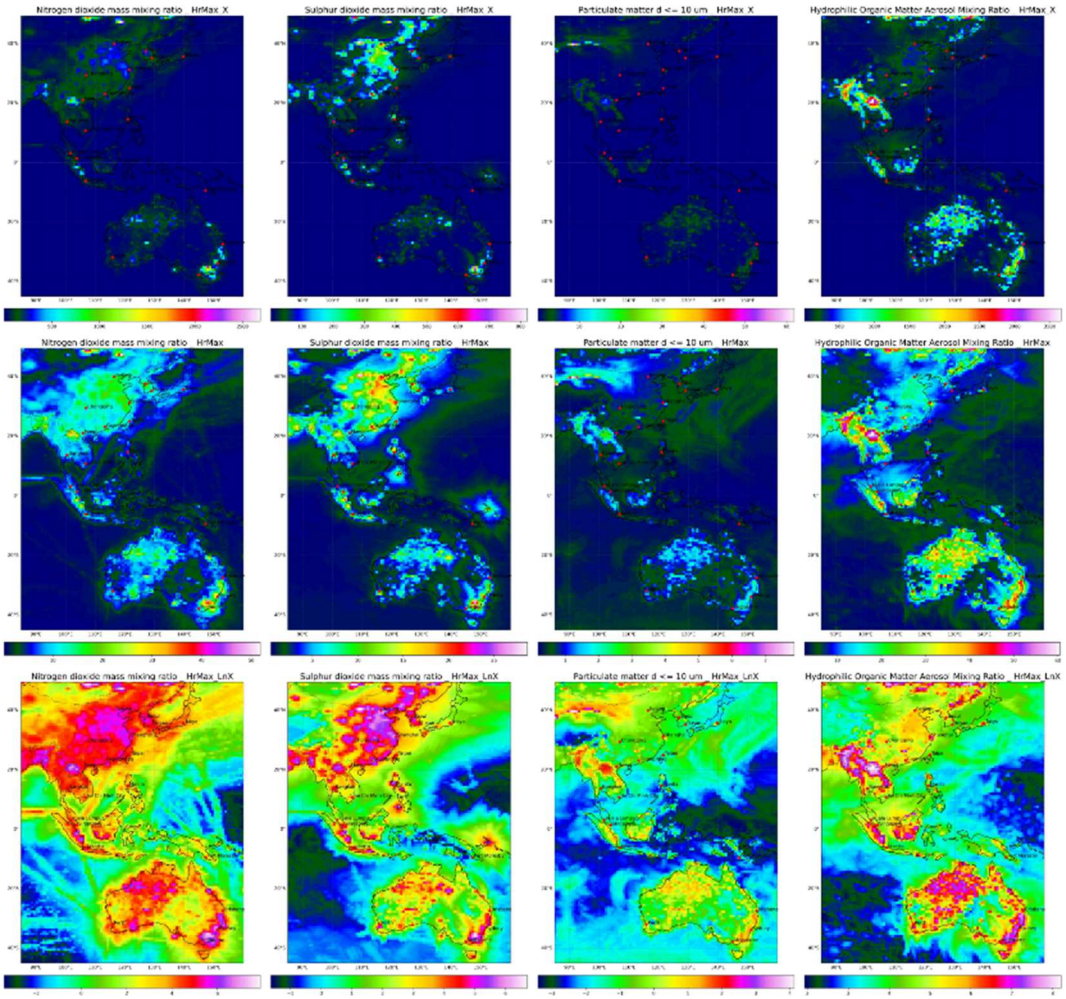

Figure 2. Data normalization (depicting maximum hour). Columns left to right are NO_2, SO_2, PM_{10}, and OM, respectively. Rows top to bottom are X, square root (X), and Log_{10} (X), respectively.

Using climate data operators (CDOs), statistical summaries, including minimum and maximum hourly, maximum daily mean, and annual average and variance for various factors were derived as depicted in Table 1. These environmental parameters were grouped into meteorological parameters, criteria pollutants, particulate matter, aerosol speciation, and other pollutants. The information was depicted spatially in Figures 3–7 showing the maximum hourly, maximum daily mean, annual averages, variance, and average monthly (January, April, July, October) isopleths. Due to space constraints, only one month per season (January, April, July, October) is displayed in the figures. Figure 3 depicts this for the U and V wind components, mean sea level pressure, and specific humidity, and revealed significant spatial and seasonal variations. For example, stronger winds over Australia compared to China indicated differences in pollutant dispersion, while seasonal patterns showed how variations in wind strength and direction impacted pollutant transport.

Figure 4 depicts isopleths of CO, O_3, NO_2, and SO_2 exhibiting distinct spatial and temporal patterns. CO and NO_2 levels were highest in urban and industrial areas, particularly in China. Seasonal variations highlighted the influence of heating emissions and

atmospheric conditions on pollutant levels. For example, higher levels of NO_2 during winter months were associated with increased heating activities and stable atmospheric conditions that trap pollutants close to the ground.

The isopleths of CO, O_3, NO_2, and SO_2 in Figure 4 reveal significant regional differences and seasonal variations. China shows the highest pollutant levels, driven by industrial and urban activities. Seasonal peaks are evident for different pollutants, emphasizing the need for targeted air quality management strategies. Indonesia and Australia have lower pollutant levels, with specific events, like regional fires, influencing local air quality.

Figure 5 depicts isopleths of particulate matter (PM_1, $PM_{2.5}$, PM_{10}) and dust aerosols. Visual analysis reveals that non-industrial sources, such as wind-blown dust and fires, significantly contribute to high PM concentrations in Australia and northern China. Seasonal variations indicate consistent high concentrations in interior Australia, with transient dust events from fires and storms impacting short-term readings. In contrast, long-term averages highlight persistent industrial emissions, particularly in China, where industrial areas show consistently higher concentrations.

Table 1. Data statistics across the study region grouped according to parameter type.

Parameter	CAMS	Min Hourly	Max Hourly	Max Daily	Annual Average	Std Dev
Meteorological parameters						
2 m dewpoint temperature (°C)	d2m	−1.36	24.49	23.51	14.54	1.96
Mean sea level pressure (kPa)	msl	98.93	103.08	102.87	101.33	0.71
Specific humidity (g/kg)	q	2.07	4.39	4.28	3.31	0.62
Temperature (°C)	t	7.13	30.55	28.22	19.85	1.85
2 m temperature (°C)	t2m	7.25	31.02	28.34	20.10	1.87
Total column water vapor (kg/m^2)	tcwv	9.60	65.39	61.05	31.86	3.09
10 m U wind component (m/s)	u10	−13.02	14.25	11.00	−0.54	1.85
10 m V wind component (m/s)	v10	−13.25	13.53	10.48	0.48	1.81
Criteria Pollutants (μg/kg PM $\mu g/m^3$)						
Carbon monoxide mass mixing ratio	co	38.05	1373.0	754.93	110.44	1.89
Ozone mass mixing ratio (full chemistry scheme)	go3	8.84	142.44	105.97	45.80	1.14
Nitrogen dioxide mass mixing ratio	no2	0.00	18.65	6.77	0.69	0.30
Particulate matter d $\leq$ 1 μm	pm1	0.00	116.92	60.84	4.99	29.05
Particulate matter d $\leq$ 2.5 μm	pm2p5	0.00	198.35	105.16	11.24	41.80
Particulate matter d $\leq$ 10 μm	pm10	0.00	315.69	166.03	17.86	53.80
Sulfur dioxide mass mixing ratio	so2	0.02	8.47	4.34	0.63	0.25
Other Pollutants (μg/kg)						
Ethane	c2h6	0.15	8.73	4.55	0.47	0.16
Propane	c3h8	0.02	4.50	2.24	0.14	0.12
Isoprene	c5h8	0.04	36.63	13.52	1.47	0.42
Hydrogen peroxide	h2o2	0.00	6.10	3.64	0.67	0.30
Formaldehyde	hcho	0.14	10.67	5.17	0.78	0.22
Nitric acid	hno3	0.00	8.19	3.79	0.21	0.23
Nitrogen monoxide mass mixing ratio	no	0.00	12.77	2.58	0.05	0.16
Hydroxyl radical	oh	0.00	0.00	0.00	0.00	0.00
Peroxyacetyl nitrate	pan	0.01	9.51	4.80	0.41	0.26
GLC of AOD of constituents (μg/kg)						
Sea salt aerosol (0.03–0.5 μm) mixing ratio	aermr01	0.00	3.02	1.82	0.17	0.20

Table 1. *Cont.*

Parameter	CAMS	Min Hourly	Max Hourly	Max Daily	Annual Average	Std Dev
Sea salt aerosol (0.5–5 μm) mixing ratio	aermr02	0.00	253.32	152.47	14.51	1.79
Sea salt aerosol (5–20 μm) mixing ratio	aermr03	0.00	207.73	120.80	7.11	1.40
Dust aerosol (0.03–0.55 μm) mixing ratio	aermr04	0.00	18.22	10.63	0.29	0.45
Dust aerosol (0.55–0.9 μm) mixing ratio	aermr05	0.00	38.59	21.52	0.55	0.64
Dust aerosol (0.9–20 μm) mixing ratio	aermr06	0.00	80.23	35.48	0.70	0.71
Hydrophilic organic matter aerosol mixing ratio	aermr07	0.00	82.96	44.53	2.90	0.80
Hydrophobic organic matter aerosol mixing ratio	aermr08	0.00	41.80	17.78	0.84	0.43
Hydrophilic black carbon aerosol mixing ratio	aermr09	0.00	4.79	2.53	0.11	0.18
Hydrophobic black carbon aerosol mixing ratio	aermr10	0.00	5.44	2.19	0.04	0.12
Sulfate aerosol mixing ratio	aermr11	0.00	15.48	9.54	1.16	0.44
Total Column (mg/m^2)						
GEMS total column ozone	gtco3	4880.98	7738.72	7520.36	6029.06	2.68
Total column ethane	tc_c2h6	1.42	14.35	11.19	2.79	0.24
Total column propane	tc_c3h8	0.12	6.28	4.50	0.48	0.19
Total column isoprene	tc_c5h8	0.02	16.10	9.37	0.64	0.33
Total column methane	tc_ch4	9415.60	9966.03	9944.68	9713.68	0.39
Total column hydrogen peroxide	tc_h2o2	0.27	34.59	28.48	6.79	0.70
Total column nitric acid	tc_hno3	0.36	26.86	18.99	2.53	0.44
Total column nitrogen monoxide	tc_no	0.00	3.29	0.83	0.19	0.40
Total column hydroxyl radical	tc_oh	0.00	0.01	0.00	0.00	0.02
Total column peroxyacetyl nitrate	tc_pan	1.11	28.84	20.65	4.97	0.43
Total column carbon monoxide	tcco	405.9	2678.6	2161.8	726.7	2.7
Total column formaldehyde	tchcho	0.61	11.89	8.04	2.07	0.27
Total column nitrogen dioxide	tcno2	0.29	10.12	5.09	1.40	0.29
Total column sulfur dioxide	tcso2	0.10	9.56	6.37	0.93	0.28
AOD						
Black carbon aerosol optical depth at 550 nm	bcaod550	0.000	0.141	0.096	0.006	0.077
Dust aerosol optical depth at 550 nm	duaod550	0.000	0.440	0.279	0.016	0.130
Organic matter aerosol optical depth at 550 nm	omaod550	0.002	0.845	0.606	0.062	0.218
Sea salt aerosol optical depth at 550 nm	ssaod550	0.000	0.410	0.252	0.034	0.156
Sulfate aerosol optical depth at 550 nm	suaod550	0.001	0.500	0.318	0.051	0.180
Total aerosol optical depth at 469 nm	aod469	0.006	1.772	1.239	0.197	0.344
Total aerosol optical depth at 670 nm	aod670	0.004	1.261	0.857	0.139	0.291
Total aerosol optical depth at 865 nm	aod865	0.003	1.031	0.683	0.109	0.262
Total aerosol optical depth at 1240 nm	aod1240	0.002	0.847	0.542	0.079	0.231

The isopleths of black carbon (BC), organic carbon (OC), sulfate, and sea salt aerosols in Figure 6 revealed the influence of both anthropogenic and natural sources. Higher BC and OC concentrations were observed in industrial regions, while sea salt aerosols were prominent in coastal areas. Other pollutants, including NO, isoprene, formaldehyde, and ethane, showed significant spatial and temporal variability, as in Figure 7. Higher concentrations were observed in urban and industrial regions, influenced by both anthropogenic and biogenic sources. Figure 8 illustrates the ratio of ground-level concentration (GLC) to aerosol optical depth (AOD) for various aerosols, including specific humidity, sea spray, dust aerosol, organic material, black carbon, and sulfate aerosol. This ratio reflects how

different aerosol sources influence the vertical distribution of pollutants. For instance, aerosols from wildfires, which generate significant heat, often result in a lower ratio at the surface, indicating higher concentrations aloft. In contrast, dust storms tend to have a more uniform vertical profile, leading to a stronger correlation between GLC and AOD. Understanding these variations is essential for accurately assessing air quality and the environmental impact of different aerosol sources. This ratio is not spatially constant, and varies by four orders of magnitude across parameters, indicating that calibration must be location- and parameter-specific, thus supporting the use of geographically weighted regression (GWR) over land use regression (LUR). The ratio reaches a maximum correlating with specific events, such as fires and dust storms, as evidenced by the different maxima across the panels. It is the temporal variation of source strengths that dictates this ratio, causing it to change both spatially and temporally as the plume dissipates.

Anomalies from monthly–hourly means for NO_2, SO_2, $PM_{2.5}$, and PM_{10} over an eight-year period (2016–2023) (Figure 9) highlighted significant changes and events affecting air quality. The analysis revealed the impact of events, such as the COVID-19 pandemic (2020 column five) and wildfires (2019 column four), on pollutant levels and lower 2023 (column eight) levels.

The timeseries plots in Figure 10 of the maximum hourly concentration per day for different pollutants revealed distinct pollution patterns in cities across Australia, China, and Southeast Asia. The 2019/2020 Australian bushfires and the COVID-19 lockdown's impact on pollution levels were evident in the data. These plots illustrated how specific events lead to sharp increases or decreases in pollutant concentrations, providing valuable insights for air quality management.

Geographically weighted regression (GWR) analysis illustrated the spatial variability in the correlation in Figure 11 between air pollutants and meteorological factors. The analysis revealed significant relationships in the Jacobian matrix (dY/dX) in Figure 12 between wind components, specific humidity, and pollutant concentrations. This highlighted the need for localized air quality management strategies that consider the unique environmental conditions of each region. By understanding these spatial relationships, policymakers can develop more effective interventions to reduce pollution and protect public health.

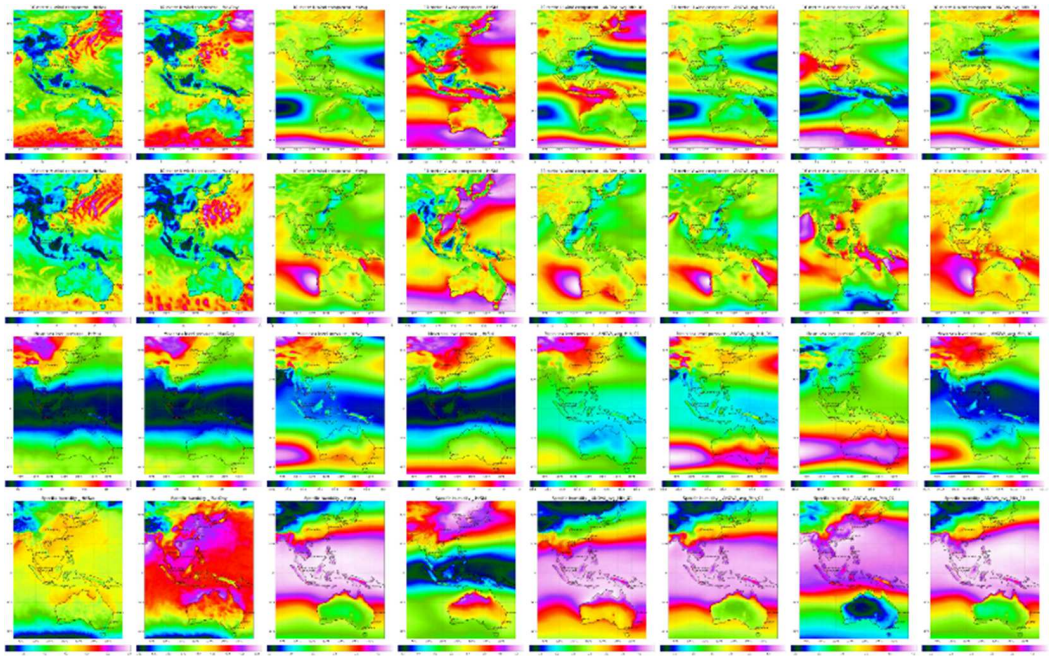

Figure 3. Spatiotemporal variation by row of U, V, MSLP, and specific humidity. Columns are max hourly, max daily, annual average, Std Dev, and mean: January, April, July, and October, respectively.

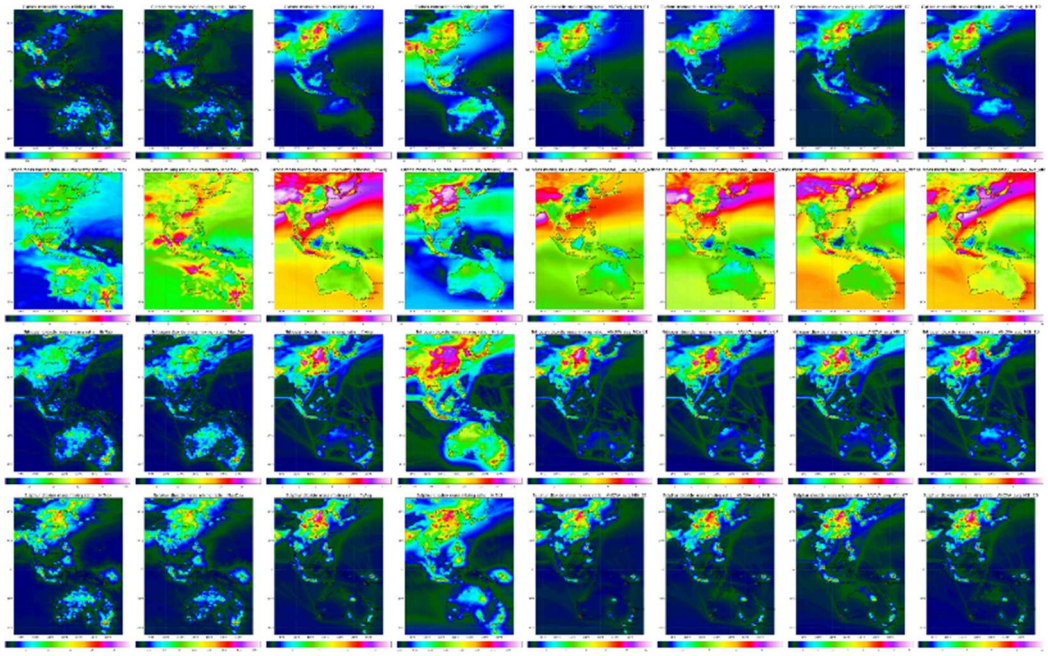

Figure 4. Isopleths for CO, O_3, NO_2, and SO_2 (by row, respectively). Columns are max hourly, max daily, annual average, Std Dev, and mean: January, April, July, and October, respectively.

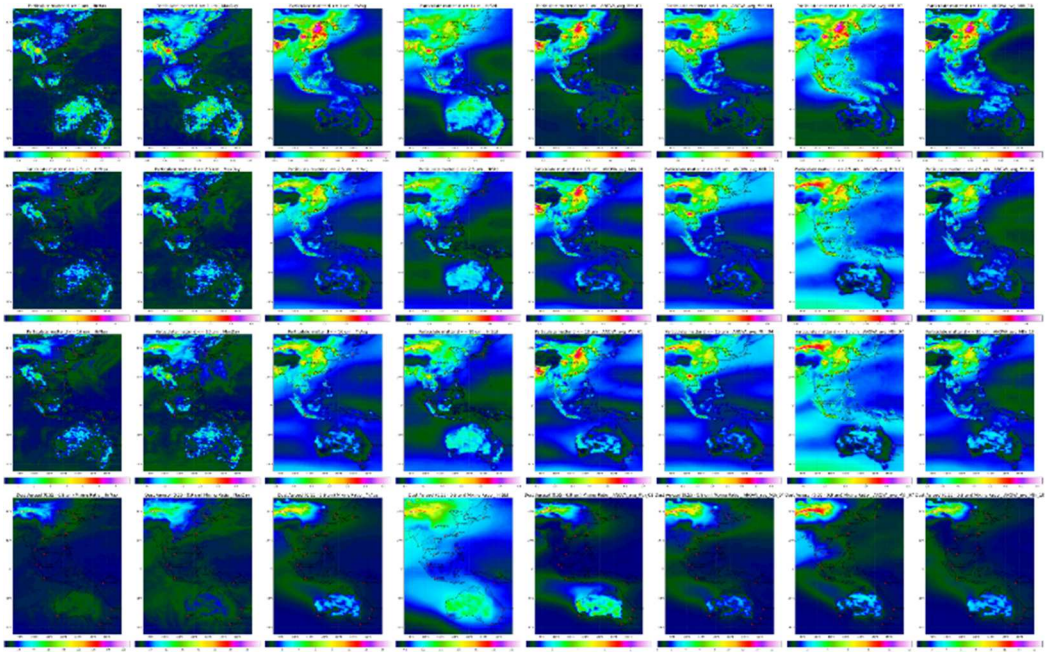

Figure 5. Isopleths of PM_1, $PM_{2.5}$, PM_{10} and, dust $aerosol_{2.5}$ (by row, respectively). Columns are max hourly, max daily, annual average, Std Dev, and mean: January, April, July, and October, respectively.

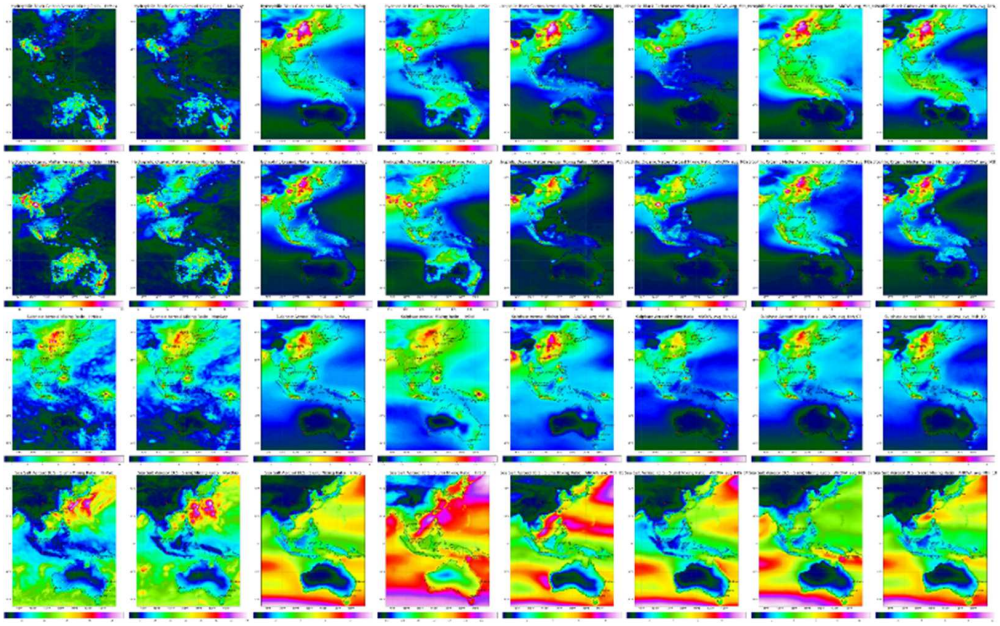

Figure 6. Isopleths of BC, OC, SU, and SS aerosols (by row, respectively). Columns are max hourly, max daily, annual average, Std Dev, and mean: January, April, July, and October, respectively.

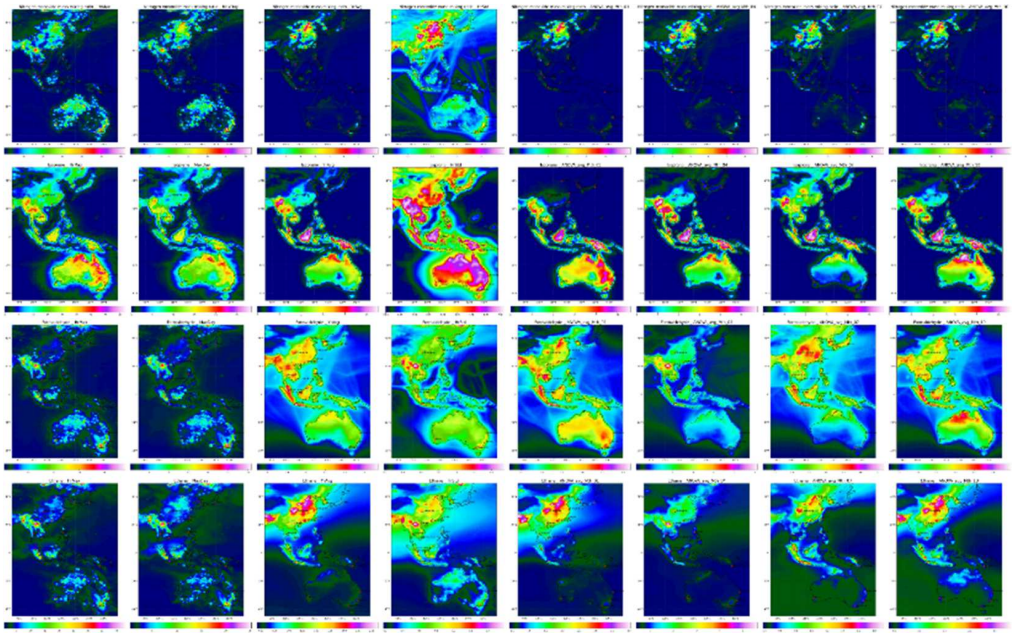

Figure 7. Isopleths of NO, isoprene, formaldehyde, and ethane (by row, respectively). Columns are max hourly, max daily, annual average, Std Dev, and mean: January, April, July, and October, respectively.

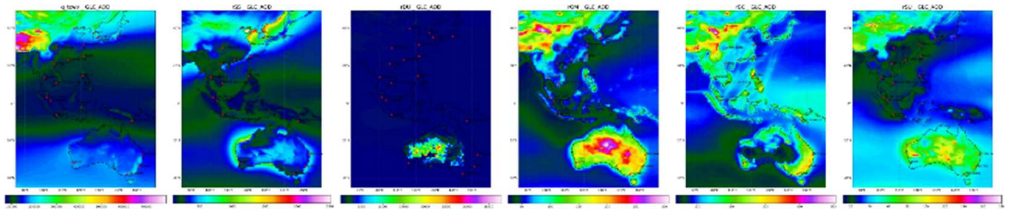

Figure 8. Vertical ratio GLC/AOD of specific humidity (g/m^2), sea spray, dust aerosol, organic material, black carbon, and sulfate aerosol (unitless 0 to 1×10^6).

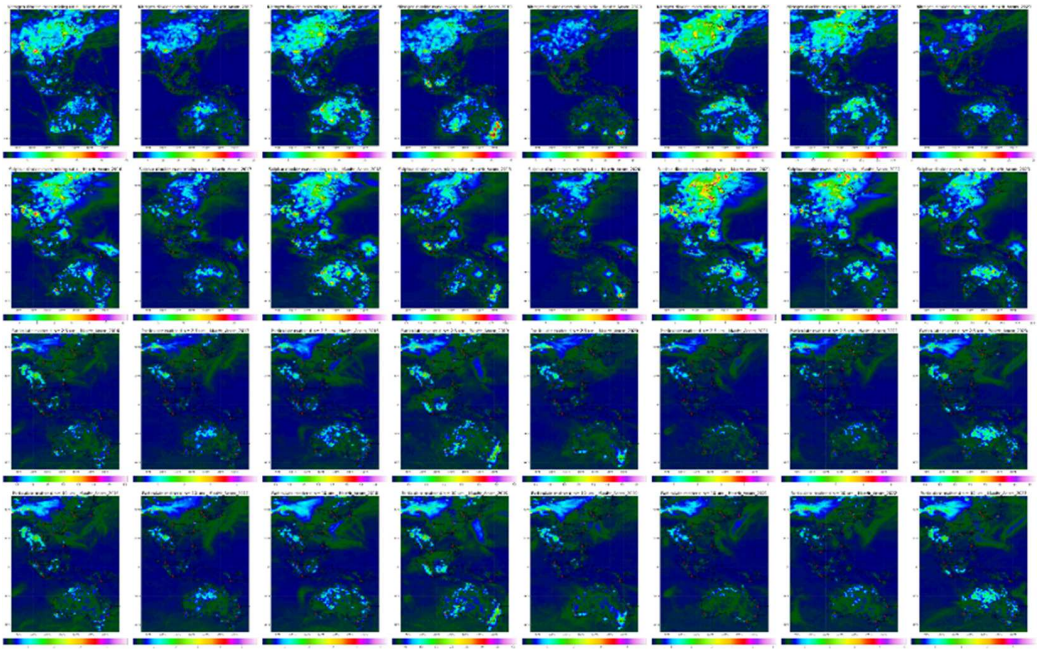

Figure 9. Annual analysis of anomalies for 2016 to 2023 (columns) for the criteria pollutants NO_2, SO_2, $PM_{2.5}$, and PM_{10}.

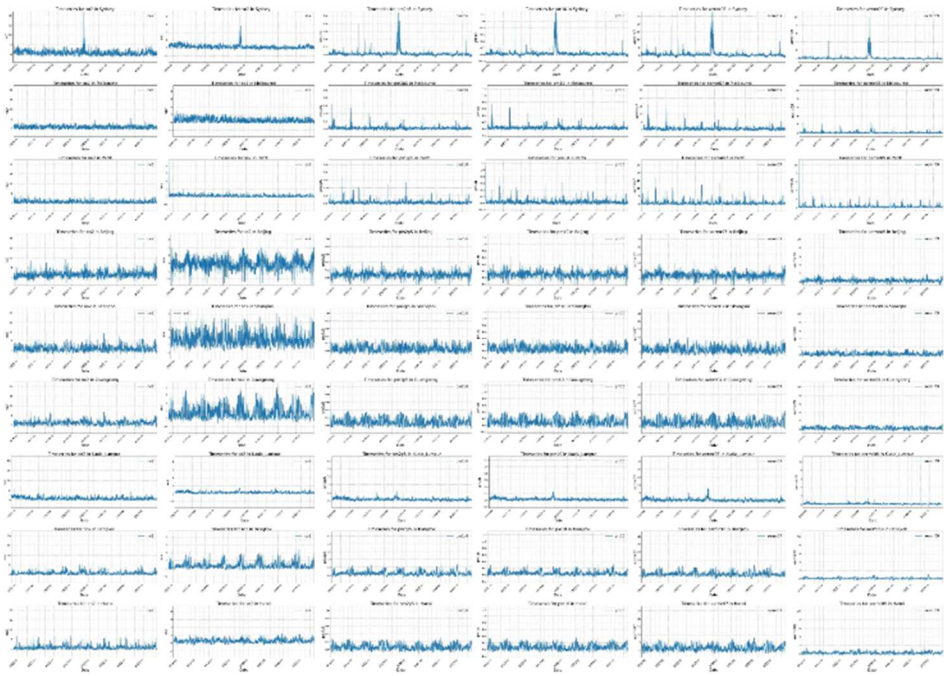

Figure 10. Timeseries of maximum hourly concentration per day for the pollutants (columns) NO_2, SO_2, $PM_{2.5}$, PM_{10}, OC, and BC, respectively, and for the cities (rows) of Sydney, Melbourne, Perth, Beijing, Shanghai, Guangdong, Kuala Lumpur, Bangkok, and Hanoi, respectively, over the eight-year period.

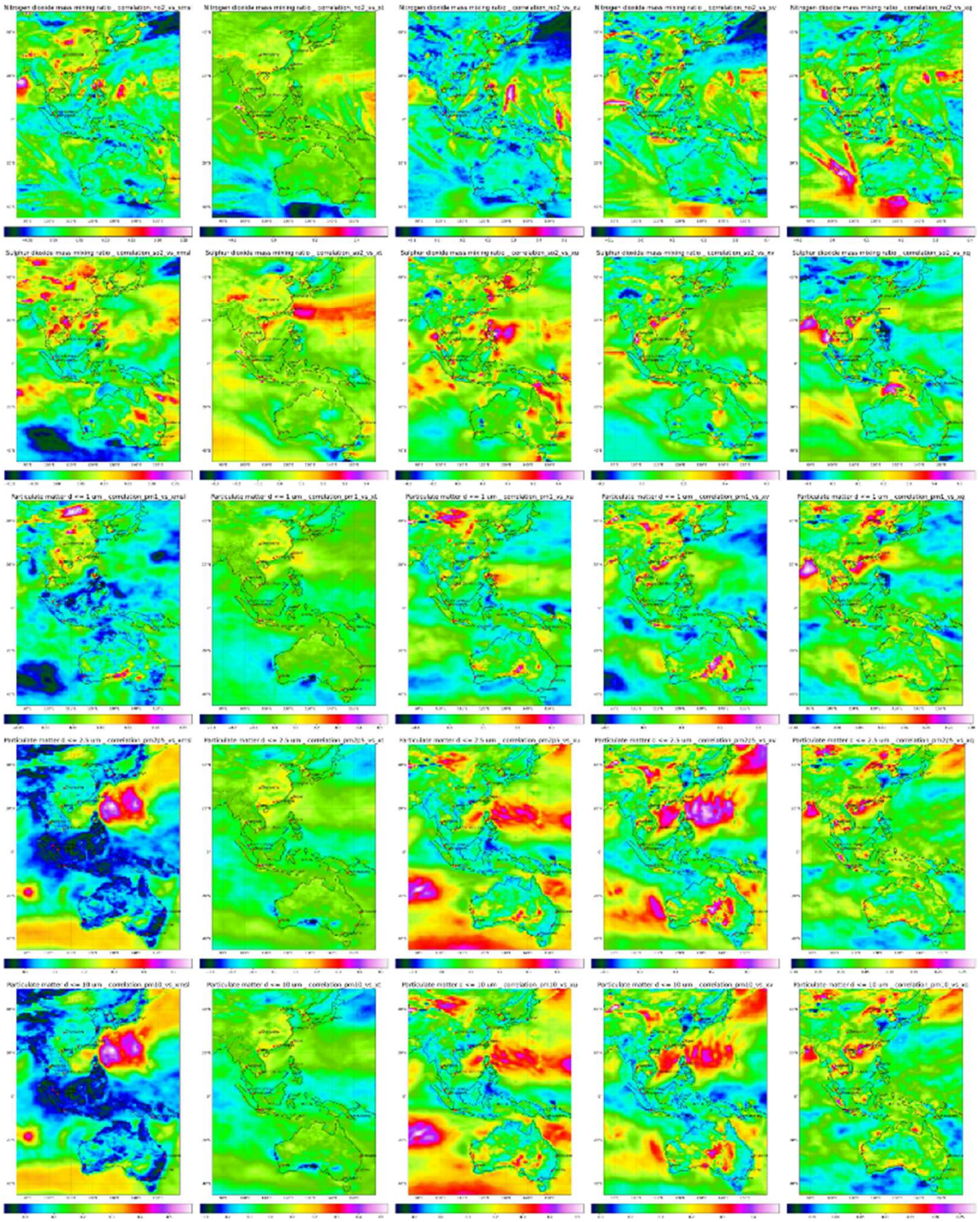

Figure 11. Correlation matrix analysis for the meteorological parameters (columns = X) of msl, 1/T (Kelvin), U10, V10, q and the GLC (rows = Y) of NO_2, SO_2, PM_1, $PM_{2.5}$, and PM_{10} over the eight-year period.

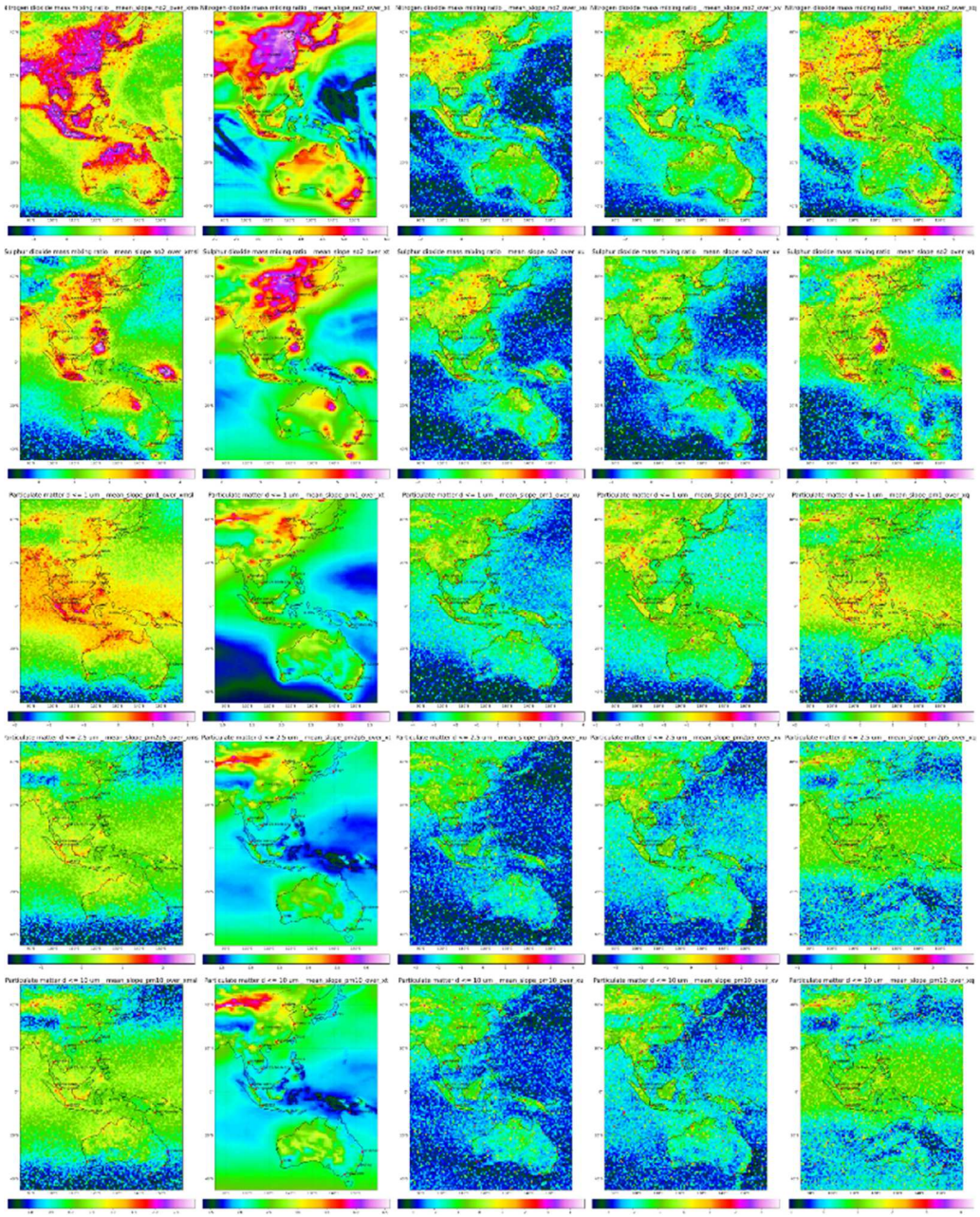

Figure 12. Jacobian matrix (LOG dY/dX) analysis for the meteorological parameters (columns = X) of msl, 1/T (Kelvin), U10, V10, q, respectively, and the GLC (rows = Y) of NO_2, SO_2, PM_1, $PM_{2.5}$, and PM_{10} over the eight-year period, respectively.

4. Discussion

Evaluating CAMS data offers insights into the strengths and limitations of such datasets for supporting air quality monitoring and policy decisions. Importantly, in the absence of dense monitoring data in certain regions, this study validates that CAMS data aligns well with anticipated patterns, with certain constraints. Key findings related to normalization techniques, interactions between atmospheric dynamics and pollutant distribution, and broader environmental implications are examined, along with a discussion of limitations and directions for future research.

Normalization techniques, particularly the square root transformation, were effective in reducing skewness due to low mean values and helped to identify critical pollution incidents, which are often significantly above mean levels. By preserving zero values and

providing a more intuitive interpretation compared to exponential transformations, the square root approach facilitated a more accurate visualization of pollution events.

Our study identified complex interactions between atmospheric dynamics—such as wind components, humidity levels, and pollutant dispersion—across different regions. For instance, analysis of U and V wind components highlighted the impact of seasonal winds, such as the winter monsoon in China, on pollutant distribution, emphasizing the importance of these factors for air quality modeling. The analysis of mean sea level pressure (MSLP) and specific humidity underscored their roles in pollution dispersion, with high-pressure systems typically leading to pollutant accumulation and low-pressure systems enhancing dispersion. Specific humidity also plays a role in cloud formation, which affects photochemical reactions and pollutant dispersion, demonstrating the need to integrate these variables for accurate predictions in air quality models.

The evaluation of CO and ground-level O_3 concentrations provided insights into urban pollution, particularly in China and Southeast Asia, where high CO levels are indicative of significant vehicular and industrial emissions, precursors for ozone formation. Mapping NO_2 and SO_2 concentrations highlighted combustion-related hotspots, suggesting that regions with high values may benefit from stricter emission standards and cleaner industrial technologies.

Analysis of $PM_{2.5}$ and PM_{10} highlighted distinct geographical patterns, reflecting both industrial activity and natural events. High particulate matter levels in Asia were primarily driven by industrial and urban sources, whereas natural sources, like the 2019/2020 Australian bushfires, contributed significantly to particulate pollution in Australia. These findings underscore the necessity of integrating dynamic meteorological data into air quality models to accurately capture the interactions between environmental conditions and pollutant behaviors.

The study's analysis of aerosol components, such as black carbon (BC), organic carbon (OC), sulfates, and sea salt aerosols, revealed the influence of both anthropogenic and natural sources on air quality. Source–receptor relationships varied by region, illustrating the non-uniform nature of air pollution. The use of geographically weighted regression (GWR) demonstrated the importance of accounting for spatial variability, contrasting with the limitations of traditional land use regression (LUR), which assumes homogeneity and may misinterpret the factors driving pollutant dispersion.

Vertical profile analysis provided a deeper understanding of how specific atmospheric components, like specific humidity and dust aerosols, vary both spatially and temporally. These observed variations reinforce the need to incorporate multiple atmospheric variables to accurately assess air quality dynamics.

Examining anomalies over eight years revealed correlations with major regional events, such as reductions in pollutant levels during the COVID-19 pandemic and increases from natural events, like wildfires. For example, pollutant reductions in 2020, except in regions affected by the Australian wildfires, highlight the impact of reduced human activity and natural pollution sources. These findings suggest that anomaly analysis can provide critical insights into the environmental impacts of disruptive events.

A notable contribution of this study is the application of the Jacobian matrix to evaluate the sensitivity of pollutants to atmospheric factors, offering insights into source–receptor influences. This approach, unlike LUR, which assumes uniform conditions, highlighted the location-specific and often non-linear relationships between emission sources and pollutant concentrations. The per-pixel approach of GWR, as shown in Figure 11, captured spatial variability more effectively than LUR, while Figure 12 demonstrates the temporal and spatial variability of the Jacobian matrix, reinforcing the need for comprehensive datasets when developing global air quality methodologies.

The study emphasizes the necessity of location-specific air quality modeling to inform public health interventions and policy decisions. Improved air quality monitoring has broader implications for public health, economic factors, and climate action. Accurate monitoring can help reduce respiratory and cardiovascular illnesses, lower healthcare costs,

and enhance workforce productivity. Although the CAMS data resolution is sufficient for large-scale analysis, it may not capture fine-scale pollution dynamics essential for local policy. Integrating higher-resolution data sources, like Himawari-8, could provide more localized insights. Future research should investigate the feasibility of such integrations and further assess the applicability of advanced normalization techniques across varied datasets to adapt to different pollutants and atmospheric conditions.

5. Conclusions

This study underscores the significant potential of the Copernicus Atmosphere Monitoring Service (CAMS) reanalysis for monitoring ground-level concentrations (GLCs) of aerosols and criteria pollutants across varied regions, including China, Indonesia, and Australia. The accuracy of CAMS data is fundamentally dependent on the quality of input data, which varies widely by region. For instance, China's dense monitoring network may contribute to more reliable CAMS data, while Australia's sparse network could introduce data uncertainties. Despite these challenges, CAMS remains a valuable resource for assessing broad air quality trends in regions that otherwise lack sufficient data.

The evaluation indicates that normalization techniques, such as the square root transformation, enhance data clarity by highlighting extreme pollution events, potentially enabling more targeted air quality interventions. Additionally, geographically weighted regression (GWR) has proven effective for capturing spatial variability in pollutant concentrations, illustrating how meteorological factors—including wind direction, humidity, and pressure—can drive pollutant dispersion across different areas. This method, in accounting for local conditions, provides a more balanced approach to modeling air quality than traditional models, such as land use regression (LUR).

While CAMS has limitations in detecting localized pollution events due to its reliance on dispersion models and existing ground-based data, integrating CAMS with satellite observations, like Himawari-8, could address these limitations. The high spatial and temporal resolution of satellite data may enhance CAMS's capability to detect episodic pollution events and improve monitoring accuracy in areas where ground data are sparse.

In conclusion, although CAMS's accuracy is regionally dependent due to differences in data density, it remains a critical tool for monitoring regional air quality. Future efforts should focus on combining CAMS data with high-resolution satellite observations to strengthen local monitoring capacity and capture pollution events more effectively. This integrated approach has the potential to provide more robust data for public health and environmental policy decisions, ultimately leading to more efficient and responsive air quality management.

Supplementary Materials: The following supporting information can be downloaded at https://github.com/MilesSowden/CAMS_Him8, accessed on 26 April 2024, all high resolution images of each individual subplot; Cams.sh (the main processing code); PlotStats.sh (the main plotting code); plot_netcdf.py (the python script per subplot); plot_timeseries.py (plotting timeseries); and gws.sh (the GWR code).

Funding: This research received no funding and was conducted independently of all funding or affiliation by the author.

Data Availability Statement: The data supporting the findings of this study are available upon request from the author due to their large file size. Additionally, the original data can be accessed from the Copernicus Atmosphere Monitoring Service (CAMS). For more information and to access the data, please visit https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-reanalysis-eac4?tab=form, accessed on 26 April 2024.

Conflicts of Interest: Author Miles Sowden was employed by the company Sigma Theta. The Author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Cresswell, I.; Janke, T.; Johnston, E. *Australia State of the Environment*; Commonwealth Department of Agriculture, Water, and the Environment: Canberra, Australia, 2021. [CrossRef]
2. Sowden, M.; Blake, D. Using infrared geostationary remote sensing to determine particulate matter ground-level composition and concentration. *Air Qual. Atmos. Health* **2021**, *17*, 1183–1192. [CrossRef] [PubMed]
3. Inness, A.; Ades, M.; Agustí-Panareda, A.; Barré, J.; Benedictow, A.; Blechschmidt, A.-M.; Dominguez, J.J.; Engelen, R.; Eskes, H.; Flemming, J.; et al. The CAMS reanalysis of atmospheric composition. *Atmos. Chem. Phys.* **2019**, *19*, 3515–3556. [CrossRef]
4. JMA. Meteorological Satellite Center, Himawari Real-Time Image. 2024. Available online: https://www.data.jma.go.jp/mscweb/data/himawari/sat_img.php?area=fd (accessed on 26 April 2024).
5. CNEMC. China National Environmental Monitoring Centre. 2024. Available online: https://www.cnemc.cn/ (accessed on 26 April 2024).
6. Marvin, M.R.; Palmer, P.I.; Yao, F.; Latif, M.T.; Khan, F. Uncertainties from biomass burning aerosols in air quality models obscure public health impacts in Southeast Asia. *Atmos. Chem. Phys.* **2024**, *24*, 3699–3715. [CrossRef]
7. Cheng, M.; Fang, F.; Navon, I.M.; Zheng, J.; Zhu, J.; Pain, C. Assessing uncertainty and heterogeneity in machine learning-based spatiotemporal ozone prediction in Beijing-Tianjin- Hebei region in China. *Sci. Total. Environ.* **2023**, *881*, 163146. [CrossRef] [PubMed]
8. Gokhale, S.; Patil, R. Uncertainty in modelling PM_{10} and $PM_{2.5}$ at a non-signalized traffic roundabout. *Atmos. Pollut. Res.* **2010**, *1*, 59–70. [CrossRef]
9. ECMWF. CAMS Reanalysis Data Documentation. 2024. Available online: https://confluence.ecmwf.int/display/CKB/CAMS:+Reanalysis+data+documentation (accessed on 28 August 2023).
10. Stratoulias, D.; Nuthammachot, N.; Dejchanchaiwong, R.; Tekasakul, P.; Carmichael, G.R. Recent developments in satellite remote sensing for air pollution surveillance in support of Sustainable Development Goals. *Remote Sens.* **2024**, *16*, 2932. [CrossRef]
11. Sun, H.; Wang, D.; Han, W.; Yang, Y. Quantifying the Impact of Aerosols on Geostationary Satellite Infrared Radiance Simulations: A Study with Himawari-8 AHI. *Remote Sens.* **2024**, *16*, 2226. [CrossRef]
12. Pinder, R.W.; Klopp, J.M.; Kleiman, G.; Hagler, G.S.; Awe, Y.; Terry, S. Opportunities and challenges for filling the air quality data gap in low- and middle-income countries. *Atmos. Environ.* **2019**, *215*, 116794. [CrossRef]
13. Azmi, W.N.F.W.; Pillai, T.R.; Latif, M.T.; Koshy, S.; Shaharudin, R. Application of land use regression model to assess outdoor air pollution exposure: A review. *Environ. Adv.* **2023**, *11*, 100353. [CrossRef]
14. Zhou, Q.; Wang, C.; Fang, S. Application of geographically weighted regression (GWR) in the analysis of the cause of haze pollution in China. *Atmos. Pollut. Res.* **2019**, *10*, 835–846. [CrossRef]
15. Peng, B.; Xie, B.; Wang, W.; Wu, L. Enhancing Seasonal PM2.5 Estimations in China through Terrain–Wind–Rained Index (TWRI): A Geographically Weighted Regression Approach. *Remote Sens.* **2024**, *16*, 2145. [CrossRef]
16. Zhang, Y.; Cheng, H.; Huang, D.; Fu, C. High Temporal Resolution Land Use Regression Models with POI Characteristics of the $PM_{2.5}$ Distribution in Beijing, China. *Int. J. Environ. Res. Public Health* **2021**, *18*, 6143. [CrossRef] [PubMed]
17. Patton, A.P.; Zamore, W.; Naumova, E.N.; Levy, J.I.; Brugge, D.; Durant, J.L. Transferability and Generalizability of Regression Models of Ultrafine Particles in Urban Neighborhoods in the Boston Area. *Environ. Sci. Technol.* **2015**, *49*, 6051–6060. [CrossRef]
18. Ghassoun, Y.; Löwner, M.-O.; Weber, S. Wind direction related parameters improve the performance of a land use regression model for ultrafine particles. *Atmos. Pollut. Res.* **2019**, *10*, 1180–1189. [CrossRef]
19. Mölter, A.; Lindley, S. Developing land use regression models for environmental science research using the XLUR tool—More than a one-trick pony. *Environ. Model. Softw.* **2021**, *143*, 105108. [CrossRef]
20. Li, Z.; Ho, K.-F.; Chuang, H.-C.; Yim, S.H.L. Development and intercity transferability of land-use regression models for predicting ambient PM10, PM2.5, NO2 and O3 concentrations in northern Taiwan. *Atmos. Chem. Phys.* **2021**, *21*, 5063–5078. [CrossRef]
21. Nesser, H.; Jacob, D.J.; Maasakkers, J.D.; Scarpelli, T.R.; Sulprizio, M.P.; Zhang, Y.; Rycroft, C.H. Reduced-cost construction of Jacobian matrices for high-resolution inversions of satellite observations of atmospheric composition. *Atmos. Chem. Phys.* **2021**, *14*, 5521–5534. [CrossRef]
22. Jacobson, M.Z.; Lu, R.; Turco, R.P.; Toon, O.B. Development and application of a new air pollution modeling system-part I: Gas-phase simulations. *Atmos. Environ.* **1996**, *30*, 1939–1963. [CrossRef]
23. Petrenko, M.; Kahn, R.; Chin, M.; Soja, A.; Kucsera, T. The use of satellite-measured aerosol optical depth to constrain biomass burning emissions source strength in the global model GOCART. *J. Geophys. Res. Atmos.* **2012**, *117*, D18. [CrossRef]
24. Kumar, S.; Biswas, K.; Pandey, A.K. Forecasting formation of a Tropical Cyclone Using Reanalysis Data. *arXiv* **2022**, arXiv:2212.06149. [CrossRef]
25. Chen, X.; Bannister, R.; Shaddick, G.; Zidek, J.V. On the Stochasticity of Reanalysis Outputs of 4D-Var. *arXiv* **2023**, arXiv:2304.03648.
26. Baklanov, A.; Brunner, D.; Carmichael, G.; Flemming, J.; Freitas, S.; Gauss, M.; Hov, Ø.; Mathur, R.; Schlünzen, K.H.; Seigneur, C.; et al. Key Issues for Seamless Integrated Chemistry–Meteorology Modeling. *Bull. Am. Meteorol. Soc.* **2018**, *98*, 2285–2292. [CrossRef] [PubMed]
27. Bocquet, M.; Elbern, H.; Eskes, H.; Hirtl, M.; Žabkar, R.; Carmichael, G.R.; Flemming, J.; Inness, A.; Pagowski, M.; Camaño, J.L.P.; et al. Data assimilation in atmospheric chemistry models: Current status and future prospects for coupled chemistry meteorology models. *Atmos. Chem. Phys.* **2015**, *15*, 5325–5358. [CrossRef]

28. Agustí-Panareda, A.; Barré, J.; Massart, S.; Inness, A.; Aben, I.; Ades, M.; Baier, B.C.; Balsamo, G.; Borsdorff, T.; Bousserez, N.; et al. Technical note: The CAMS greenhouse gas reanalysis from 2003 to 2020. *Atmos. Chem. Phys.* **2023**, *23*, 3829–3859. [CrossRef]
29. Peterson, R.A. Finding Optimal Normalizing Transformations via best Normalize. *R J.* **2021**, *13*, 310–329. [CrossRef]
30. A&WMA. *Theories, Methodologies, Computational Techniques, and Available Databases and Software*; A&WMA: Philadelphia, PA, USA, 2010; Volume IV.
31. Leelőssy, Á.; Molnár, F.; Izsák, F.; Havasi, Á.; Lagzi, I.; Mészáros, R. Dispersion modeling of air pollutants in the atmosphere: A review. *Cent. Eur. J. Geosci.* **2014**, *6*, 257–278. [CrossRef]
32. IASTATE. Dispersion for Point Sources. CE 424. 2011. Available online: https://www.engineering.iastate.edu/~leeuwen/CE%20524/Notes/Dispersion_Handout.pdf (accessed on 16 June 2024).
33. Gleckler, P.J.; Taylor, K.E.; Doutriaux, C. Performance metrics for climate models. *J. Geophys. Res. Atmos.* **2008**, *113*, 8972. [CrossRef]
34. Colleges, I.R.; Cenita, J.A.; Asuncion, P.R.; Victoriano, J. Performance Evaluation of Regression Models in Predicting the Cost of Medical Insurance. *Int. J. Comput. Sci. Res.* **2023**, *7*, 2052–2065. [CrossRef]
35. Hoek, G.; Beelen, R.; de Hoogh, K.; Vienneau, D.; Gulliver, J.; Fischer, P.; Briggs, D. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos. Environ.* **2008**, *42*, 7561–7578. [CrossRef]
36. Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **2001**, *106*, 7183–7192. [CrossRef]
37. Eslahchi, M.; Dehghan, M.; Asl, S.A. The general Jacobi matrix method for solving some nonlinear ordinary differential equations. *Appl. Math. Model.* **2012**, *36*, 3387–3398. [CrossRef]
38. Fotheringham, A.S.; Charlton, M.E.; Brunsdon, C. Geographically Weighted Regression: A Natural Evolution of the Expansion Method for Spatial Data Analysis. *Environ. Plan A Econ. Space* **1998**, *30*, 1905–1927. [CrossRef]
39. Schulzweida, U. CDO User Guide (2.3.0). 2023. Available online: https://code.mpimet.mpg.de/projects/cdo/wiki (accessed on 16 June 2024). [CrossRef]