

Article

Global Air Quality: An Inter-Disciplinary Approach to Exposure Assessment for Burden of Disease Analyses

Gavin Shaddick^{1,*}, James M. Salter¹, Vincent-Henri Peuch² , Giulia Ruggeri³, Matthew L. Thomas¹, Pierpaolo Mudu³, Oksana Tarasova⁴, Alexander Baklanov⁴  and Sophie Gumy³

¹ Joint Centre for Excellence in Environmental Intelligence, University of Exeter & Met Office, Exeter EX4 4PY, UK; J.M.Salter@exeter.ac.uk (J.M.S.); m.l.thomas@exeter.ac.uk (M.L.T.)

² European Centre for Medium-Range Weather Forecasts, Reading RG2 9AX, UK; Vincent-Henri.Peuch@ecmwf.int

³ World Health Organization, CH-1211 Geneva, Switzerland; ruggerig@who.int (G.R.); mudup@who.int (P.M.); bonjourso@who.int (S.G.)

⁴ World Meteorological Organization, CH-1211 Geneva, Switzerland; otarasova@wmo.int (O.T.); abaklanov@wmo.int (A.B.)

* Correspondence: g.shaddick@exeter.ac.uk

Abstract: Global assessments of air quality and health require comprehensive estimates of the exposures to air pollution that are experienced by populations in every country. However, there are many countries in which measurements from ground-based monitoring are sparse or non-existent, with quality-control and representativeness providing additional challenges. While ground-based monitoring provides a far from complete picture of global air quality, there are other sources of information that provide comprehensive coverage across the globe. The World Health Organization developed the Data Integration Model for Air Quality (DIMAQ) to combine information from ground measurements with that from other sources, such as atmospheric chemical transport models and estimates from remote sensing satellites in order to produce the information that is required for health burden assessment and the calculation of air pollution-related Sustainable Development Goals indicators. Here, we show an example of the use of DIMAQ with the Copernicus Atmosphere Monitoring Service Re-Analysis (CAMSR) of atmospheric composition, which represents the best practices in meteorology and climate monitoring that were developed under the World Meteorological Organization's Global Atmosphere Watch programme. Estimates of PM_{2.5} from CAMSR are integrated within the DIMAQ framework in order to produce high-resolution estimates of air pollution exposure that can be aggregated in a coherent fashion to produce country-level assessments of exposures.

Keywords: air pollution; fine particulate matter; PM_{2.5}; health effects; global burden of disease; data integration; DIMAQ



Citation: Shaddick, G.; Salter, J.M.; Peuch, V.-H.; Ruggeri, G.; Thomas, M.L.; Mudu, P.; Tarasova, O.; Baklanov, A.; Gumy, S. Global Air Quality: An Inter-Disciplinary Approach to Exposure Assessment for Burden of Disease Analyses. *Atmosphere* **2021**, *12*, 48. <https://doi.org/10.3390/atmos12010048>

Received: 1 December 2020

Accepted: 24 December 2020

Published: 31 December 2020

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Air pollution represents the largest environmental risk to public health worldwide. The World Health Organization (WHO) estimates that 4.2 million premature deaths every year can be attributed to fine particulate ambient air pollution (PM_{2.5}), and that over 90% of people worldwide are exposed to harmful levels of PM_{2.5} [1]. The quality of the air that we breathe varies greatly across the globe, with populations in many low- and middle-income countries suffering from the highest exposures, with some experiencing levels of PM_{2.5} that are over five times the WHO Air Quality Guidelines [2]. Two of the Sustainable Development Goals (SDG) reflect the importance of air quality for our health and wellbeing, for which the WHO is the custodial agency: (i) SDG 11, which includes air pollution exposure in urban settings; and, (ii) SDG 3, which includes mortality due to air pollution (ambient and household).

There are a variety of methods for producing estimates of levels of air pollution in the context of its health effects. Short-term forecasting supports communication and timely public health response [3], especially for the most vulnerable, whereas the assessment of longer-term exposures is required for the estimation of burden of disease. The latter is of crucial importance, as, while reducing exposure during air pollution episodes will reduce the risk of acute harm, the greatest health benefit is likely to be achieved with long term reductions that reduce the risk of chronic effects [1].

Producing estimates (of air pollution exposures) for the burden of disease assessments and the information that is required to calculate SDG indicators requires the following information:

1. Population distributions of exposures for each country—high-resolution (over space) estimates of pollution concentrations matched to population estimates.
2. The relationship between increased exposures and risks to health. For a given level of air pollution, an estimate of the increase in risk (above that associated with underlying country-level mortality rates) is required for the causes of disease that are being considered.
3. Country-level mortality rates. The number of deaths, by cause, gender, and age, together with corresponding populations at risk.

WHO has been producing estimates of exposure and burden of disease from ambient air particulate matter since the early 2000s, first at the regional level [4], which led to national guidance for performing burden of disease assessment [5,6] Since these early attempts, the methods for assessing global exposure to particulate matter and related health impacts have improved drastically, which has allowed for country-level estimates of burden to be reported. These methods have developed from simple linear regression to models allowing spatial variability, including geographical weighted regression and hierarchical (random effects) models, and they have incorporated information from remote sensing satellites, chemical transport models, population estimates, and land-use to estimate levels of PM_{2.5} [7–12].

Assessing global air pollution and health requires an inter-disciplinary approach and, since 2014, the WHO has been leading the Global Platform on Air Quality and Health [13] which convenes experts across a wide range of scientific disciplines, including air quality modelling and monitoring, satellite remote sensing, statistics, epidemiology, health effects assessment, public health, from research institutions, space research agencies, national bodies, and other UN agencies that are involved in air pollution and health activities. To date, there have been three expert consultations for the Global Platform, in 2014, 2015, and 2017, with the aim of assisting countries and cities to measure air pollution and reduce exposures, its associated burden of disease, and effectively tackle the sources of health-damaging air pollution.

The establishment of the Task Force on Data Integration, which led the subsequent development of the Data Integration Model for Air Quality (DIMAQ), was an important output from the first Global Platform meeting. This collaborative effort was driven by the need to produce estimates of population-exposures (to air pollution) for the burden of disease calculations, and latterly for SDG indicators. While many countries have a history of monitoring air pollution (e.g., in Western Europe and North America) and it is increasing in many parts of the world (e.g., in China and India), there are still many countries and regions, notably in Africa, South America, and parts of South Eastern Asia, in which monitoring information is sparse, and possibly non-existent [14]. Even in countries with extensive monitoring, measurements tend to be focused on larger towns and cities, and there may be large proportions of the population, particularly in rural areas, for which localized information is generally scarce [9,12]. This is a particular problem with PM_{2.5}, due to its high variability. A lack of consistency in measurement operations (including height of air intake, quality control, and maintenance) is also an issue when considering raw observations. DIMAQ was developed in order to combine information from ground-based measurements with that from other sources, such as chemical transport

models and estimates from remote sensing satellites that are based on aerosol optical depth (AOD) [7,9,11,12] in order to produce estimates of exposures to air pollution with comprehensive coverage across the globe. However, estimates from products, such as these are often available on relatively low-resolution grids and they may not incorporate the fine-scale variation in pollution in relation to a population that is required for health burden calculations. They may also not align with ground measurements (where available).

DIMAQ was used in order to provide the latest assessments of the global burden of disease from the WHO [1], the Institute of Health Metric Evaluation's (IHME) 'Global Burden of Disease' project [15–17] and the Health Effects Institute's 'State of Global Air' [18]. It provides the information that is used in order to quantify the number of global annual deaths that are attributable to ambient air pollution (and also on a country- and region-level basis) and the proportion of people that are exposed to levels of pollution above WHO guidelines (currently estimated to be 4.2 million and 91%, respectively) [1].

In this paper, we report the findings of an action from the most recent Global Platform meeting, which was convened by WHO in collaboration with World Meteorological Organization (WMO) to review progress, identify gaps, and propose the next steps and further development needed for ensuring high quality estimates of human exposure to air pollution and the related burden of disease. The action was to investigate whether data products from the WMO Global Atmosphere Watch (GAW) could be integrated within the DIMAQ framework, complementing the information from satellite-based estimates that combine aerosol optical depth retrievals with information from the GEOS-Chem chemical transport model [11] that have been used previously [19] in order to provide additional information for the assessment of the health burden that is associated with air pollution. GAW partners produce a number of products integrating information from atmospheric chemical-transport models, remote sensing, and ground monitoring data using data assimilation [20]. Specifically, here, we consider the Copernicus Atmosphere Monitoring Service Re-Analysis of atmospheric composition (CAMSR) that provides estimates of $PM_{2.5}$ at 0.7° spatial resolution that is based on combining information from an integrated atmospheric composition and meteorological model with AOD (from remote sensing satellites) while using data assimilation [21], resulting in gridded estimates that have comprehensive coverage and are consistent over time. Importantly, CAMSR provides vertical information regarding the distribution of aerosol and, while retaining the information that is available from satellite AOD information through data assimilation, benefits from the modelling of the aerosol sources, of their transport, and their eventual removal from the atmosphere. If DIMAQ can be used to calibrate the information on annual total $PM_{2.5}$ from CAMSR to produce estimates that are consistent with those using other inputs, then it opens the possibility of including sub-year temporal variation and the ability to apportion the components of particulate matter to specific emission sources or sectors; information that is available in the rich set of outputs of CAMSR (CAMSR includes a split of total $PM_{2.5}$ into five components: dust, organic matter, black carbon, sea salt, and sulphates).

Here, we present a case study that shows how CAMSR can be incorporated within the DIMAQ framework in order to produce high-resolution estimates of annual air pollution exposures (for total $PM_{2.5}$) that are similar to previously reported estimates and it can be aggregated to produce country-level estimates of population exposures, together with associated measures of uncertainties. The remainder of this paper is organised, as follows: Section 2 provides details of DIMAQ and its role in producing country, regional, and global estimates of air quality and population-exposures for health burden assessments; Section 3 presents case study of implementing DIMAQ while using the CAMSR data and Section 4 provides a concluding discussion.

2. Materials and Methods

The Data Integration Model for Air Quality (DIMAQ) is based on the idea of spatio-temporal statistical calibration, which is also known as measurement-modelling fusion: if the relationship between ground measurements and sources of data that have compre-

ensive coverage can be modelled, then ground level concentrations can be predicted in locations where PM_{2.5} data from ground monitors are not available. The relationship between ground measurements and gridded data is then used to predict ground level PM_{2.5} concentrations that are based on the information in the gridded data and on spatial information. There are challenges associated with this: (i) data at multiple scales of measurement in both space and time need to be ‘joined together’ in a coherent fashion; (ii) each of the different data sources represent fundamentally different quantities; (iii) the relationships between the ground-based measurements and gridded products may vary over space and time; and, (iv) each of the inputs will have different error structures and uncertainties, and these may vary over both space and time. DIMAQ addressed each of these challenges, providing a major step forward in the ability to produce high-resolution estimates of air pollution with comprehensive coverage across the globe. Notably, for the first time, there was a framework in which to propagate the uncertainty in exposure estimates through to the overall estimates of the burden of disease, and the ability to calculate probabilities of exceedance, e.g., the probability that the levels of pollution in any area exceeded the WHO air quality guidelines.

Subsequent to its initial release in 2016, DIMAQ has been further developed [2,22] providing a series of improvements in the availability of high-quality estimates of global air quality and the impacts on health. The international need to track changes and progress towards the SDGs and be able to couple the evidence for effective interventions through global, regional, and local indicators of air pollution trends drove these developments. In addition, there was a need for exposure estimates that reflected finer scale variation in air quality, particularly in urban areas, which is now addressed by allowing for the gridded products to be downscaled [23,24] to increase the spatial resolution. Additionally, in the original formulation of DIMAQ, the coefficients of the calibration equations were allowed to vary on a country-by-country basis, with yearly outputs being treated independently. The formulation of DIMAQ was enhanced to allow the coefficients of the calibration equations to also vary smoothly over time (and space) in order to produce the level of information that is required for the SDGs. This allows for trends to be modelled, making full use of the multi-year data now available in the WHO database of ambient air quality [14]. The inclusion of multi-year data also required an update to computational methods that are used to calculate the exposure estimates and associated measures of uncertainties.

The following provides a brief introduction to DIMAQ. In the interests of clarity, the first part of the description focuses on the spatial aspects of the model, followed by the introduction of the temporal aspects. DIMAQ contains both the random and fixed effects:

$$Y_s = \tilde{\beta}_{0s} + \sum_{p \in P} \beta_p X_{pA_{ps}} + \sum_{q \in Q} \tilde{\beta}_{qs} X_{qA_{qs}} + \epsilon_s \quad (1)$$

where Y_s are ground measurements at point locations, s , with two groups of covariates, P and Q , containing fixed effects (across space) and spatially-varying random effects, respectively. Here, A_{ps} denotes the grid cell that contains point location s . The residual error terms are at the level of the point locations, $\epsilon_s \sim^{iid} N(0, \sigma_\epsilon^2)$.

The spatially-varying coefficients, $\tilde{\beta}_{0s}$ and $\tilde{\beta}_{qs}$, $q \in Q$, comprise a set of both fixed and random effects

$$\tilde{\beta}_{0s} = \beta_0 + \beta_{0s}^R + \beta_{0s} \quad (2)$$

$$\tilde{\beta}_{qs} = \beta_q + \beta_{qs}^R + \beta_{qs} \quad (3)$$

Here, β_0 and β_q are fixed effects that represent the overall mean value of the coefficients, with β_{0s}^R and β_{qs}^R being regional random effects. The latter allow for the estimation of the calibration coefficients in areas with limited ground monitoring to ‘borrow’ information from the calibration coefficients for other countries in the surrounding region [19]; however, it should be noted that the accurate estimation of the coefficients will be reliant on the accuracy, and the representiveness of the available ground measurements. The β_{0s} and β_{qs} represent further spatial adjustments, allowing for the intercept and calibration coefficients

to vary continuously over space. These spatially-continuous random effects allow for the spatial variation within grid cell to be modelled and they are assumed to follow a stationary, isotropic, zero-mean Gaussian random field with a Matern covariance function [25].

Predictions can be made at any given resolution, although it is noted that the density of ground monitoring information within a localised area will determine the extent of any fine scale variation.

Incorporating time (year) into Equations (3) and (4) gives:

$$\tilde{\beta}_{0st} = \beta_{0t} + \beta_{0st}^R + \beta_{0st} \quad (4)$$

$$\tilde{\beta}_{qst} = \beta_{qt} + \beta_{qst}^R + \beta_{qst} \quad (5)$$

where the random effects can now vary over both time and space, with the temporal variation in $\beta_{.st}$ being assigned a first-order autoregressive process, e.g., $\beta_{.st} \sim N(\beta_{\rho.s,t-1}, \sigma_{\beta}^2)$.

DIMAQ uses ‘approximate’ Bayesian inference based on integrated nested Laplace approximations [26,27]. Originally, the estimation (of the model parameters) and prediction (of air pollution on a high-resolution grid) were performed simultaneously. As there were significant computational challenges that are associated with this, DIMAQ was revised to adopt a new method: (joint) samples of all the model’s parameters are taken and predictions are calculated as a linear combination of the (sampled) parameters and the model inputs. This process is repeated through taking multiple samples, which results in the construction of posterior distributions (of estimated PM_{2.5}) for each cell in the grid. In addition to being computationally efficient, this provides a coherent method for aggregation and for the first time, country-level, multi-year summaries of air quality, together with associated measures of uncertainties, can be directly produced in the form that is required for the SDGs (e.g., country-level average exposures for 11.6.2) and the burden of disease assessments (e.g., population-distributions of exposure).

3. Results

In this section, we present the results of a case study: producing country-level indicators of air quality while using DIMAQ with the Copernicus Atmosphere Monitoring Service a re-analysis of atmospheric composition. The WMO GAW produces a number of products regarding atmospheric constituents and greenhouse gases while using data assimilation techniques, which allow for the integration of information from chemical-transport models, ground-based, and satellite remote sensing, as well as in-situ monitoring data at the surface and onboard aircraft and ships. These estimates are produced at hourly time steps that can be averaged over longer periods with the spatial resolution for global models being currently of the order of 50 km. Of these, the Copernicus Atmosphere Monitoring Service (CAMS) provides estimates of PM_{2.5}, and other pollutants, at high temporal resolution (3 h) and (in some regions) up to 10 km × 10 km spatial resolution. CAMS has recently released a re-analysis of atmospheric composition that is based on combining information from an integrated atmospheric composition and meteorological model with observations of AOD while using data assimilation, as described in Section 1 [21]. This reanalysis (CAMSR) provides gridded estimates of PM_{2.5} at 0.7° resolution that have comprehensive coverage and that are consistent over the period from 2003 to present (the time period that is covered by CAMSR is continually being extended).

However, estimates from products, such as CAMSR, may not immediately be representative of the exposures to air pollution that drive adverse health effects. There is a need to align spatially-resolved estimates of concentrations with population distributions at a higher resolution than is available and ensure close alignment with ground measurements in locations where they are available in order to produce the population-weighted measures of exposures required to estimate health effects. Figure 1 shows comparisons between the estimates from CAMSR and ground measurements for 2016 and it shows clear biases, particularly for Eastern and South-Eastern Asia, where (gridded) estimates from CAMSR are markedly higher than ground measurements. The R² (and root mean squared error,

RMSE) for each of the regions are 0.42 ($55.65 \mu\text{g m}^{-3}$) for Eastern/South-Eastern Asia; 0.18 ($5.59 \mu\text{g m}^{-3}$) for Europe/Northern America; 0.03 ($3.71 \mu\text{g m}^{-3}$) for Australia/New Zealand; 0.42 ($65.85 \mu\text{g m}^{-3}$) for Central/Southern Asia; 0.00 ($46.26 \mu\text{g m}^{-3}$) for Latin America/Caribbean; 0.34 ($21.89 \mu\text{g m}^{-3}$) for Northern Africa/Western Asia; and, 0.94 ($6.58 \mu\text{g m}^{-3}$) for in Sub-Saharan Africa. It has to be noted that some of these biases may be at least partly related to representativeness issues when comparing values that are valid for surface areas of about 5000 km^2 (0.7°) and point observations, which may be representative of only few tens of km^2 . The corresponding plots of the estimates from DIMAQ used in the WHO assessment of the global burden of disease attributable to ambient air pollution [1] can be seen in Figure 2, where much closer alignment with ground measurements can be seen. With the estimates from DIMAQ, the R^2 (and RMSE) for each of the regions are 0.91 ($5.45 \mu\text{g m}^{-3}$) for Eastern/South-Eastern Asia; 0.90 ($1.88 \mu\text{g m}^{-3}$) for Europe/Northern America; 0.85 ($0.53 \mu\text{g m}^{-3}$) for Australia/New Zealand; 0.81 ($21.92 \mu\text{g m}^{-3}$) for Central/Southern Asia; 0.89 ($4.27 \mu\text{g m}^{-3}$) for Latin America/Caribbean; 0.88 ($7.85 \mu\text{g m}^{-3}$) for Northern Africa/Western Asia; and, 0.96 ($4.91 \mu\text{g m}^{-3}$) for in Sub-Saharan Africa.

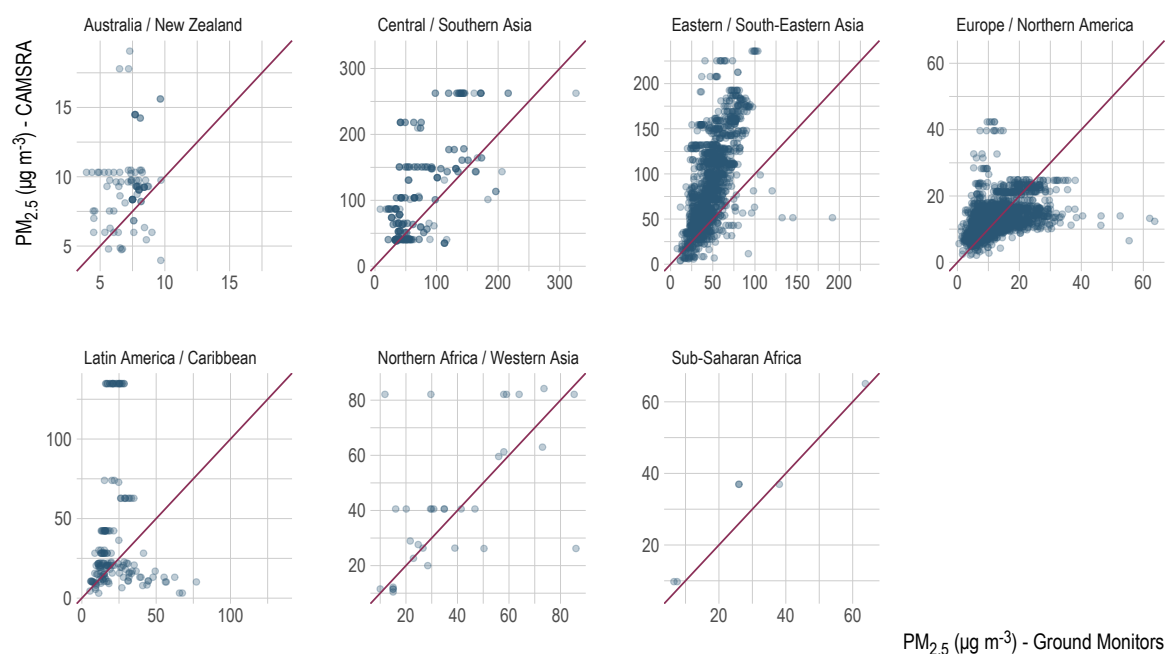


Figure 1. Comparison between the annual average of $\text{PM}_{2.5}$ estimated by the Copernicus Atmosphere Monitoring Service multi-annual global reanalysis of atmospheric composition and ground measurements, by Sustainable Development Goals (SDG) region, for 2016. Red lines indicate a one-to-one relationship.

If the estimates from CAMSRA were used in their ‘raw’ form, then the biases (e.g., the overestimates that were seen in Eastern/Southern Asia) would lead to biases in estimates in summaries of exposures and thus subsequent biases in the assessment of disease burden. Equally, if the summaries are based on ground measurements alone, then a lack of representativeness in monitoring networks and a lack of quality control would induce biases, for example, in the number of days that contribute to the calculation of annual means, whether the exact location of monitors are known or whether the monitoring location is of a prescribed type, i.e., background, rural, etc. Figure 3 shows the population-weighted annual average concentrations by country for 2016 that were calculated using CAMSRA compared to the equivalent metric from SDG indicator 11.6.2 (calculated using DIMAQ, see [2] for further details) for the same year. It should be noted that, to aid comparison, the values for the SDG indicator are those for both urban and rural areas (the reported indicator 11.6.2 is for cities/urban areas). Clear differences are seen for countries, notably in Eastern and South-Eastern Asia, with, for example, the estimate for China being markedly higher at $92 \mu\text{g m}^{-3}$ when compared to $46 \mu\text{g m}^{-3}$. This particular case may be due to

a lag between the very strong emissions reduction efforts that were seen in China since 2012, and updates to global emissions inventories that are used in the model that provides a-priori inputs for the reanalysis.

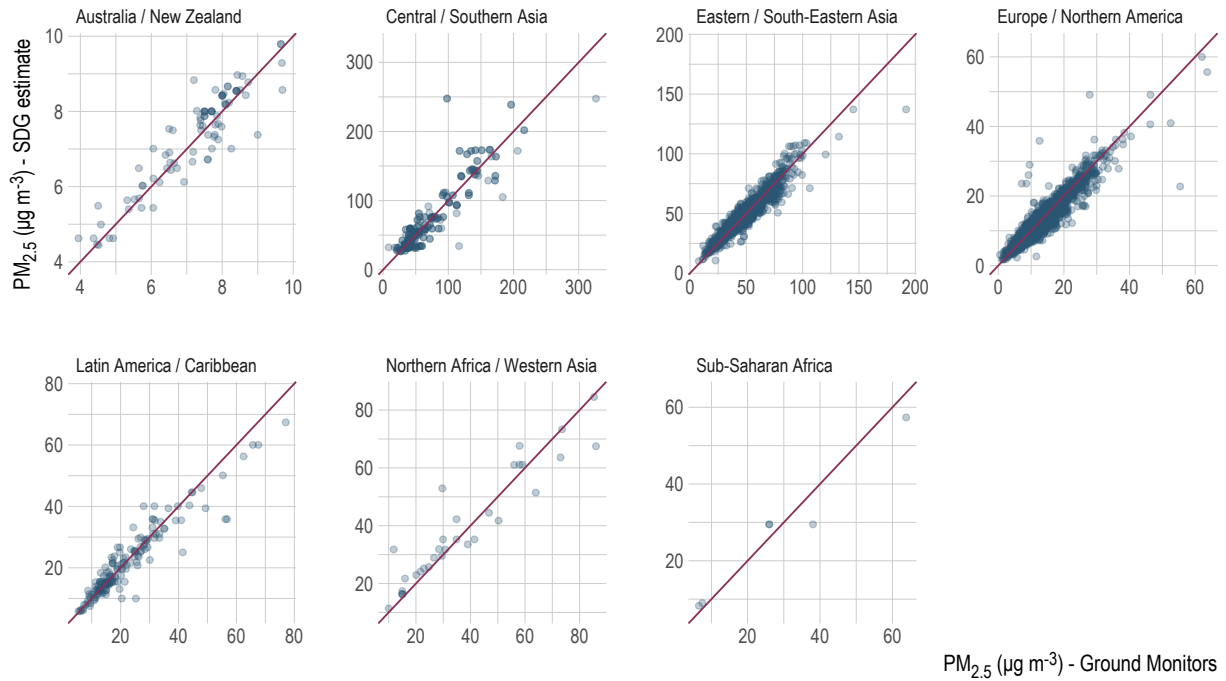


Figure 2. Comparison between the estimates of annual average of $PM_{2.5}$ from Data Integration Model for Air Quality (DIMAQ) and ground measurements, by SDG region, for 2016. Red lines indicate a one-to-one relationship.



Figure 3. Comparison between the country-level population-weighted annual average of $PM_{2.5}$ estimated by the Copernicus Atmosphere Monitoring Service multi-annual global reanalysis of atmospheric composition and SDG indicator 11.6.2 calculated using DIMAQ, by SDG region, for 2016. Grey line indicates a one-to-one relationship.

However, by including the CAMSRA estimates within the DIMAQ framework, assigning them spatio-temporally varying random effects (as in Equations (4) and (5)), the gridded estimates can be calibrated in order to reduce the biases that are seen in Figure 2. The result can be seen in Figure 4, which shows a marked reduction in the differences that were seen between the estimates and ground measurements (where they are available). The biases observed in Figure 2 have largely been corrected with close alignment (between ground measurements and CAMSRA estimates) now seen in all of the regions. Notably, in Eastern/South-Eastern Asia, the R^2 (and RMSE) improves from 0.42 to 0.90 (55.65 to $6.44 \mu\text{g m}^{-3}$) and in Europe/Northern America from 0.18 to 0.89 (5.59 to $1.89 \mu\text{g m}^{-3}$). Changes in the values of R^2 for other SDG regions were 0.03 to 0.79 (RMSE 3.71 to $0.72 \mu\text{g m}^{-3}$) for Australia/New Zealand; 0.42 to 0.82 (65.85 to $20.40 \mu\text{g m}^{-3}$) in Central/Southern Asia; 0.00 to 0.86 (46.26 to $4.80 \mu\text{g m}^{-3}$) in Latin America/Caribbean; 0.34 to 0.88 (21.89 to $7.67 \mu\text{g m}^{-3}$) in Northern Africa/Western Asia; and, 0.94 to 0.95 (6.58 to $5.25 \mu\text{g m}^{-3}$) in Sub-Saharan Africa. The RMSEs improve for each region, with the global figure reducing from $33.52 \mu\text{g m}^{-3}$ to $5.80 \mu\text{g m}^{-3}$, the latter comparing with a global RMSE of $5.77 \mu\text{g m}^{-3}$ when using DIMAQ.

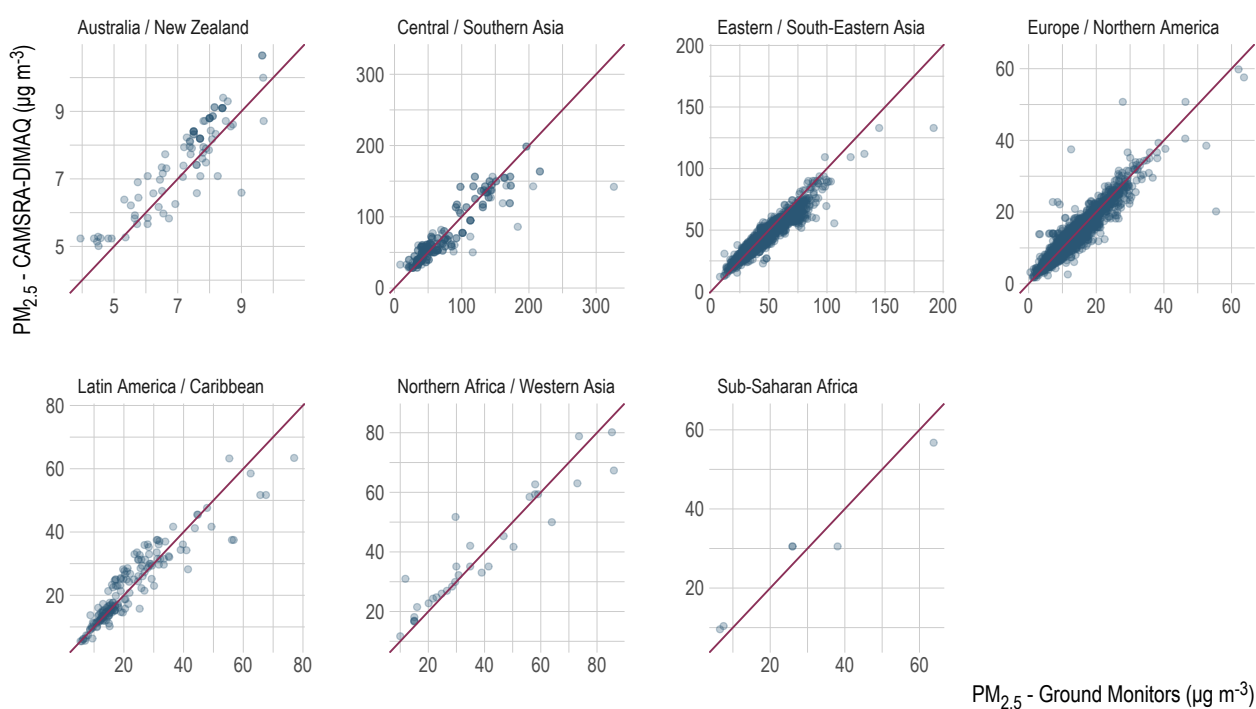


Figure 4. Comparison between the annual average of $\text{PM}_{2.5}$ estimated by DIMAQ while using the Copernicus Atmosphere Monitoring Service multi-annual global reanalysis of atmospheric composition and ground measurements, by SDG region, for 2016. Red lines indicate a one-to-one relationship.

Figures 5 and 6 show the maps of the estimates of $\text{PM}_{2.5}$ from DIMAQ while using CAMSRA, and the uncertainty that is associated with those estimates, for the Eastern/South-Eastern Asia and Central/Southern Asia SDG regions. Figure 5 shows the median values (of the posterior distributions) for each grid cell and Figure 6 presents the coefficients of variation (mean/standard deviation). A clear pattern in uncertainty can be seen with, as might be expected, the higher uncertainty seen in areas where there is sparse ground monitoring information.

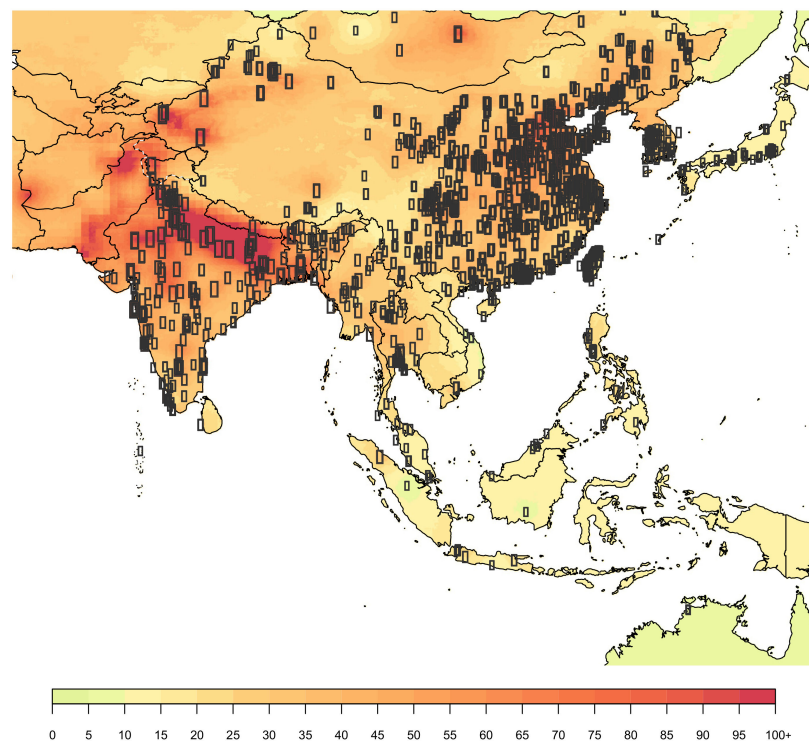


Figure 5. Estimated annual average concentrations of PM_{2.5} ($\mu\text{g m}^{-3}$) in the Eastern/South-Eastern Asia and Central/Southern Asia SDG regions for 2016. Estimates are the medians of the posterior distributions from DIMAQ while using the Copernicus Atmosphere Service multi-annual global reanalysis of atmospheric composition.

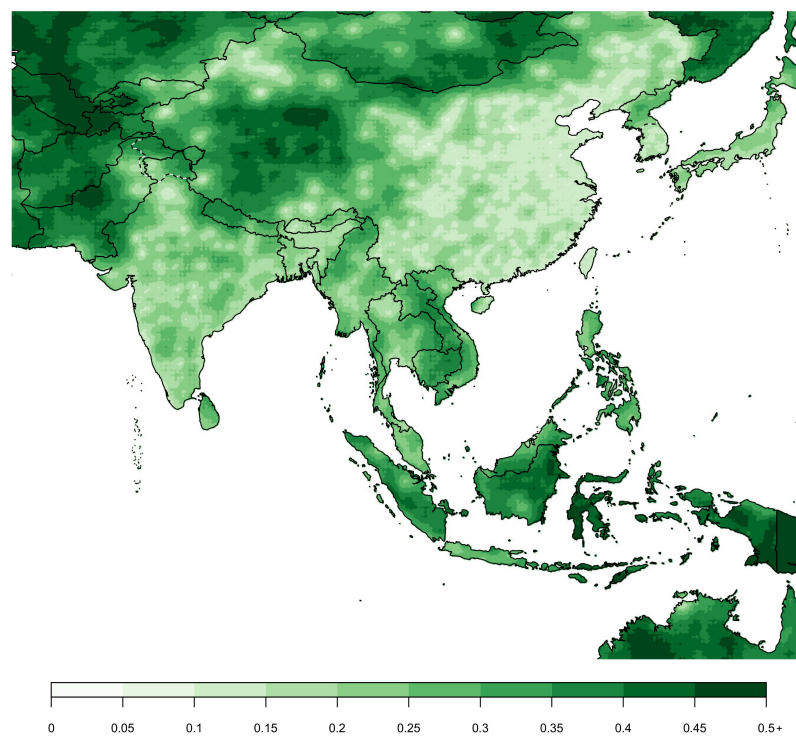


Figure 6. Uncertainty associated with estimated annual average concentrations of PM_{2.5} in the Eastern/South-Eastern Asia and Central/Southern Asia SDG regions for 2016. Uncertainty is represented by the coefficient of variation (mean/standard deviation).

4. Discussion

Long-term policies for reducing air pollution have been shown to be effective and they have been implemented in many countries, notably in Europe and the United States. However, even in countries with the cleanest air, there are large numbers of people exposed to harmful levels of air pollution. The SDGs, which aim to achieve a better and more sustainable future for all, include indicators that are specifically related to air pollution: “By 2030, reduce the adverse per capita environmental impact of cities, including by paying special attention to air quality and municipal and other waste management”; and SDG 3.9: “By 2030, substantially reduce the number of deaths and illnesses from hazardous chemicals and air, water, and soil pollution and contamination”. The ability to quantify progress towards the SDGs and assess the health effects of air pollution requires comprehensive estimates of the exposures that are experienced by populations in every country. but there are many countries and regions in which ground monitoring networks are sparse, or non-existent, and, even in countries with extensive monitoring, there may be large proportions of the population, particularly in rural areas, for which localised information is poor. Whilst other sources of information on air quality are available, such as that from remote sensing satellites and chemical transport models, none of these approaches is perfect if taken in isolation. Ground monitoring suffers from inhomogeneous, and often scarce, distribution of monitoring networks that may lead to issues of representativeness and there may be a lack of quality assurance information; information from satellites from passive sensors monitor vertically integrated amounts of aerosol and they lack information on the vertical distribution, which is essential in deriving surface particulate matter concentrations; and, chemical transport models are imperfect, due to insufficient knowledge, quality of input data (e.g., emissions), and spatial resolution.

The Data Integration Model for Air Quality (DIMAQ) framework, which was developed as part of the actions from the first WHO Global Platform for Air Quality meeting, held in Geneva in 2014, provides a framework that allows for multiple sources of data to be used in estimating exposures to air pollution across the globe. DIMAQ has been applied by the WHO and the Institute of Health Metric Evaluation’s Global Burden of Disease (GBD) project in order to estimate the global burden of disease associated with air pollution [1,15–17]. In this paper, we have successfully incorporated estimates of PM_{2.5} from the CAMS re-analysis of atmospheric composition, which combines information from an integrated atmospheric composition and meteorological model with observations of AOD while using data assimilation, within the DIMAQ framework. Estimates from CAMSRA, are available globally at 0.7° resolution and they are continually and operationally updated. They are calibrated (with ground measurements) and downscaled in order to align with population estimates, which allows for the distributions of population-level exposures required for burden of disease calculations to be produced. Calculating the global burden of disease using results from implementing DIMAQ while using CAMSRA resulted in a very similar estimate of premature deaths (4.2 m per year for 2016) to that published in the WHO’s 2018 assessment of the global burden of disease attributable to air pollution [1].

An important feature of models, such as that used to produce CAMSRA is the ability to apportion the components of particulate matter to specific emission sources or sectors. CAMSRA includes a split of total PM_{2.5} into five components: dust, organic matter, black carbon, sea salt, and sulphates. Using DIMAQ with products, such as this, offers great potential for future assessments of the health impacts of individual sources of pollution and for informing potential mitigation policies by identifying the contributions from anthropogenic and non-anthropogenic sources.

Author Contributions: Conceptualization, G.S., J.M.S., V.-H.P., P.M., O.T., A.B. and S.G.; methodology, G.S., J.M.S., V.-H.P., G.R., M.L.T. and S.G.; software, J.M.S., G.R. and M.L.T.; validation, G.S., J.M.S., G.R. and M.L.T.; formal analysis, G.S., J.M.S. and G.R.; data curation, J.M.S. and G.R.; writing—original draft preparation, G.S., V.-H.P., P.M., G.R., P.M. and S.G.; writing—G.S., J.M.S., V.-H.P., P.M.,

O.T., A.B., S.G.; visualization, J.M.S. and G.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

- World Health Organization Ambient Air Pollution: A Global Assessment of Exposure and Burden of Disease. Available online: <https://www.who.int/phe/publications/air-pollution-global-assessment/en/> (accessed on 1 November 2020).
- Shaddick, G.; Thomas, M.L.; Mudu, P.; Ruggeri, G.; Gumy, S. Half the world's population are exposed to increasing air pollution. *npj Clim. Atmos. Sci.* **2020**, *3*, 1–5. [[CrossRef](#)]
- Kumar, R.; Peuch, V.-H.; Crawford, J.H.; Brasseur, G.P. The Need for a Globally Coordinated System to Observe and Forecast Regional Air Quality. *Nature* **2018**, *561*, 27–29. [[CrossRef](#)] [[PubMed](#)]
- Cohen, A.; Anderson, H.R.; Ostro, B.B.K.; Krzyzanowski, M.; Künzli, N.; Gustschmidt, K. Urban Air Pollution. In *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*; Ezzati, M., Lopez, A.D., Rodgers, A., Muraay, C.J.L., Eds.; World Health Organization: Geneva, Switzerland, 2004; pp. 1353–1433.
- Desai, M.A.; Mehta, S.; Smith, K.R. *Indoor Smoke from Solid Fuels: Assessing the Environmental Burden of Disease at National and Local Levels*; WHO Environmental Burden of Disease Series, No. 4; World Health Organization: Geneva, Switzerland, 2004.
- Ostro, B. *Outdoor Air Pollution: Assessing the Environmental Burden of Disease at National and Local Levels*; WHO Environmental Burden of Disease Series, No. 5; World Health Organization: Geneva, Switzerland, 2004.
- Van Donkelaar, A.; Martin, R.V.; Brauer, M.; Kahn, R.; Levy, R.; Verduzco, C.; Villeneuve, P.J. Global Estimates of Ambient Fine Particulate Matter Concentrations from Satellite-Based Aerosol Optical Depth: Development and Application. *Environ. Heal. Perspect.* **2010**, *118*, 847–855. [[CrossRef](#)] [[PubMed](#)]
- Murray, C.J.L.; Vos, T.; Lozano, R.; Naghavi, M.; Flaxman, A.D.; Michaud, C.; Ezzati, M.; Shibuya, K.; Salomon, J.A.; Abdalla, S.; et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **2012**, *380*, 2197–2223. [[CrossRef](#)]
- Brauer, M.; Freedman, G.; Frostad, J.; Van Donkelaar, A.; Martin, R.V.; Dentener, F.; Van Dingenen, R.; Estep, K.; Amini, H.; Apte, J.S.; et al. Ambient Air Pollution Exposure Estimation for the Global Burden of Disease 2013. *Environ. Sci. Technol.* **2016**, *50*, 79–88. [[CrossRef](#)] [[PubMed](#)]
- Burnett, R.T.; Pope, C.A., III; Ezzati, M.; Olives, C.; Lim, S.S.; Mehta, S.; Shin, H.H.; Singh, G.; Hubbell, B.; Brauer, M.; et al. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environ. Health Perspect.* **2014**, *122*, 397–403. [[CrossRef](#)] [[PubMed](#)]
- Van Donkelaar, A.; Martin, R.V.; Brauer, M.; Boys, B.L. Use of Satellite Observations for Long-Term Exposure Assessment of Global Concentrations of Fine Particulate Matter. *Environ. Heal. Perspect.* **2015**, *123*, 135–143. [[CrossRef](#)] [[PubMed](#)]
- Shaddick, G.; Thomas, M.L.; Amini, H.; Broday, D.M.; Cohen, A.; Frostad, J.; Green, A.; Gumy, S.; Liu, Y.; Martin, R.V.; et al. Data Integration for the Assessment of Population Exposure to Ambient Air Pollution for Global Burden of Disease Assessment. *Environ. Sci. Technol.* **2018**, *52*, 9069–9078. [[CrossRef](#)] [[PubMed](#)]
- World Health Organization. Global Platform on Air Quality and Health. Available online: www.who.int/airpollution/global-platform (accessed on 1 November 2020).
- Available online: https://www.who.int/airpollution/data/AAP_database_summary_results_2018_final2.pdf (accessed on 1 November 2020).
- Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. *Lancet* **2016**, *388*, 1659–1724. [[CrossRef](#)]
- Cohen, A.J.; Brauer, M.; Burnett, R.; Anderson, H.R.; Frostad, J.; Estep, K.; Balakrishnan, K.; Brunekreef, B.; Dandona, L.; Dandona, R.; et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: An analysis of data from the Global Burden of Diseases Study 2015. *Lancet* **2017**, *389*, 1907–1918. [[CrossRef](#)]
- Murray, C.J.; Aravkin, A.Y.; Zheng, P.; Abbafati, C.; Abbas, K.M.; Abbasi-Kangevari, M.; Abd-Allah, F.; Abdalalim, A.; Abdollahi, M.; Abdollahpour, I.; et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **2020**, *396*, 1223–1249. [[CrossRef](#)]
- Health Effects Institute State of Global Air. Available online: <https://web.archive.org/web/20201002080950/https://www.stateofglobalair.org/> (accessed on 1 November 2020).
- Shaddick, G.; Thomas, M.L.; Green, A.; Brauer, M.; Van Donkelaar, A.; Burnett, R.; Chang, H.H.; Cohen, A.; Van Dingenen, R.; Dora, C.; et al. Data integration model for air quality: A hierarchical approach to the global estimation of exposures to ambient air pollution. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2017**, *67*, 231–253. [[CrossRef](#)]
- World Meteorological Organization. WMO Global Atmosphere Watch (GAW) Implementation Plan: 2016–2023. Available online: https://library.wmo.int/doc_num.php?explnum_id=3395 (accessed on 1 November 2020).
- Inness, A.; Ades, M.; Agustí-Panareda, A.; Barré, J.; Benedictow, A.; Blechschmidt, A.-M.; Dominguez, J.J.; EngeleniD, R.J.; Eskes, H.; Flemming, J.; et al. The CAMS reanalysis of atmospheric composition. *Atmos. Chem. Phys.* **2019**, *19*, 3515–3556. [[CrossRef](#)]

22. GBD 2017 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **2018**, *392*, 1923–1994.
23. Chang, H.H. Data assimilation for environmental pollution fields. In *Handbook of Spatial Epidemiology*; Lawson, A.B., Banerjee, S., Haining, R.P., Ugarte, M.D., Eds.; CRC Press: Boca Raton, FL, USA, 2016; pp. 289–302.
24. de Hoogh, K.; Gulliver, J.; van Donkelaar, A.; Martin, R.V.; Marshall, J.D.; Bechle, M.J.; Cesaroni, G.; Pradas, M.C.; Dedele, A.; Eeftens, M.; et al. Development of West-European PM_{2.5} and NO₂ land use regression models incorporating satellite-derived and chemical transport modelling data. *Environ. Res.* **2016**, *151*, 1–10. [[CrossRef](#)] [[PubMed](#)]
25. Cressie, N. *Statistics for Spatial Data*; John Wiley & Sons, Inc: Hoboken, NJ, USA, 2015.
26. Rue, H.; Martino, S.; Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2009**, *71*, 319–392. [[CrossRef](#)]
27. Lindgren, F.; Rue, H.; Lindström, J. An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2011**, *73*, 423–498. [[CrossRef](#)]