Article

Emulation of a Chemical Transport Model to Assess Air Quality under Future Emission Scenarios for the Southwest of Western Australia

Stephen Vander Hoorn [1],*, **Jill S. Johnson [2]**, **Kevin Murray [1]**, **Robin Smit [3]**, **Jane Heyworth [1]**, **Sean Lam [4]**
and **Martin Cope [5]**

[1] School of Population and Global Health, The University of Western Australia, Crawley, WA 6009, Australia
[2] School of Mathematics and Statistics, University of Sheffield, Sheffield S3 7RH, UK
[3] Transport Energy/Emission Research, Brisbane, QLD 4068, Australia
[4] Department of Water and Environmental Regulation, Joondalup, WA 6027, Australia
[5] CSIRO Oceans and Atmosphere, Aspendale, VIC 3195, Australia
* Correspondence: stephen.vanderhoorn@research.uwa.edu.au

Abstract: Simulation outputs from chemical transport models (CTMs) are essential to plan effective air quality policies. A key strength of these models is their ability to separate out source-specific components which facilitate the simulation of the potential impact of policy on future air quality. However, configuring and running these models is complex and computationally intensive, making the evaluation of multiple scenarios less accessible to many researchers and policy experts. The aim of this work is to present how Gaussian process emulation can provide a top-down approach to interrogating and interpreting the outputs from CTMs at minimal computational cost. A case study is presented (based on fine particle sources in the southwest of Western Australia) to illustrate how an emulator can be constructed to simultaneously evaluate changes in emissions from on-road transport and electricity sectors. This study demonstrates how emulation provides a flexible way of exploring local impacts of electric vehicles and wider regional effects of emissions from electricity generation. The potential for emulators to be applied to other settings involving air quality research is discussed.

1. Introduction

Dispersion and chemical transport models are essential tools for estimating exposure to ambient air pollution in epidemiologic analyses and health impact assessments (HIA) [1]. Dispersion models simulate the transport, dispersion and deposition of pollutant emissions and can provide predictions at different temporal and spatial scales [2]. Chemical transport models (CTM) incorporate atmospheric chemistry and aerosol dynamics in addition to simulating the dispersion and deposition of pollutant emissions [3–5]. A key strength of these models is their ability to separate out source-specific components enabling the exploration of the impact of emission reduction policy on future air quality [6,7]. Examples include reducing on-road emissions through the promotion of public transportation, switching to cleaner low (or zero) emission fuels, or changes to the electricity sector to support the uptake of electric vehicles. CTMs are complex and require significant expertise to operate. Therefore, policy practitioners look to air quality modelling experts to provide model simulations, and advice regarding uncertainties, when using modelling to assess the costs and benefits of an intervention that can reduce outdoor air pollution.

Substantial gains in computational efficiency have been made in recent years, however, the complexity of CTMs can still be a limitation when applied to research questions that require many model runs [8]. This is especially the case for HIA that can involve multiple intervention scenarios, thus making the use of CTMs for such applications more problematic. Recent research in this area has led to the development of surrogate models that

aim to approximate a CTM under simplified assumptions, e.g., annual time scales and/or aggregated areas. One approach is the use of reduced-complexity climate models (RCMs) to predict health impacts from point source emission sources such as coal fired power plants [9]. Another example is the TM5-FASST model, which is based on a region-to-region source-receptor matrix and is developed from TM5 global CTM simulations [10]. However, the results may be less reliable than those based on full form modelling as they are less able to account for atmospheric chemistry and transport [8]. Furthermore, neither of these approaches have focused on a specific urban area let alone an assessment of intra-urban level HIAs. One approach which can reduce the computational challenge of many source sensitivity runs is coded into the Community Multiscale Air Quality System (CMAQ), which has separate downloadable versions of the CMAQ model—CMAQ-DDM-3D and ISAM which can be used to carry sensitivity analysis and calculate source attribution in a single simulation [11–13]. These methods can also be used to run scenario analysis but are nevertheless still complex and computationally intensive to run.

An alternative approach to this problem is the use of Gaussian process emulation (GPE) [14]. One early application of this approach has been the quantification of model parameter uncertainty and sensitivity to the complex interactions between aerosols and clouds [15,16]. Emulation has also been used more recently as a flexible and computationally efficient proxy of a CTM for predicting both shorter term (monthly average concentrations) and annual exposure to fine particulate matter ($PM_{2.5}$) [17–19]. In those applications, emulators are applied to study impacts of the changes to emissions aggregated by sector on air pollution across broad regional areas. Use of emulation to approximate high resolution pollution maps at an intra-urban scale has also been demonstrated [20]. However, few applications at this local spatial scale have been carried out, e.g., at the scale of neighbourhood, local government area, or administrative area within a city.

This study aims to demonstrate how GPE can assist policy makers to examine effects of city-wide air quality interventions. The example of electric vehicles (EVs) is used here as it applies to the entire transport sector thereby potentially improving air quality for urban-based populations. The case study is set in the southwest of Western Australia, and it considers changes to annual average $PM_{2.5}$ under future scenarios. Emissions from the transport and electricity sectors are varied simultaneously across a range of plausible levels and an emulator is built by analysing the outputs from multiple CTM runs. Emulators are then used to illustrate their capacity to assess the local impacts of a city-wide transition to EVs which in our case study represents an area of approximately six thousand and four hundred square kilometres. Wider regional impacts from increased electricity generation are then studied by coupling the electricity sector to the transport sector and comparing to scenarios involving renewable energy sources. The potential for emulation to simulate a CTM output and thereby make this method of scenario analysis more accessible to less technical audiences is discussed.

2. Materials and Methods

We briefly describe the scientific model used to generate experimental data needed to build and test the emulator. The case study is then introduced, and details of the emulation approach are provided. In Australia, detailed spatial assessments are limited to regions with contemporary state-of-the-art inventories [21]. Therefore, we also present the results of the CCAM-CTM applied to the case study region (Supplementary Materials, Section S1) as further background and to facilitate comparison of model performance in other applications.

2.1. The CCAM-CTM

The Commonwealth Scientific and Industrial Research Organisation (CSIRO) have developed a CTM, which is used by CSIRO and research partners to model the transport, dispersion, deposition, and chemical transformation of airborne pollutants. A summary of the science modules is given by the World Meteorological Organization [22] and can also be compared with other commonly used CTMs. In this study, the approach involves the

use of the Conformal Cubic Atmosphere Model (CCAM) for simulating regional meteorology, and the CSIRO CTM (thus CCAM-CTM) [23]. The CCAM provides a prediction of meteorological fields including temperature, wind velocity and direction, water vapour mixing ratio (including clouds), radiation and turbulence [24]. The CTM is configured with the Carbon Bond V chemical transformation mechanism for gas phase species, the Global Model of Aerosol Processes for aerosol dynamics [25], the Volatility Basis Set for treating semi-volatile organic aerosol [26], and ISORROPIA [27] (a thermodynamic equilibrium model) for treatment of secondary inorganic aerosols. Natural and anthropogenic aerosols (primary and secondary) are included in the simulations. Natural sources include sea salt aerosols, wind-blown dust, primary aerosols from ambient first, and secondary organic aerosols formed by oxidation of organic compounds emitted from vegetation and ambient fires [28–31]. Major sources of anthropogenic emissions include industry, commercial, motor vehicle, and domestic wood heaters.

The CTM has been applied to many applications involving air quality in an Australian context, but its main use is to run short-term forecasts (here coupled to the Australian Bureau of Meteorology weather forecasts) dispersion of smoke and particulates from bushfire and prescribed burning [32]. The CCAM-CTM configuration has been used as part of HIA analysis to study air quality impacts both historically and for the future [33,34]. When applied in this way, the CCAM-CTM is used in "scenario mode" by modifying emissions and to assess resulting pollutant concentrations under a counterfactual scenario.

The CCAM-CTM can be conceptualised as a complex function involving parameters X to reflect the many science processes mentioned above, and it provides model output $Y = f(X)$. However, Y is unknown until the simulation is run for a given time period across the selected geographical area(s) of interest. We define Y with respect to a base case as the absolute difference between the base case and scenario of interest, i.e.,

$$\Delta Y = Y^{x(n)} - Y^{bc}$$

where $x(n)$ is the selected scenario and 'bc' refers to the base case. We run the CCAM-CTM down to a 1 km resolution providing an estimate of ΔY in any defined grid cell within the model domain.

2.2. Case Study Region

In this study, potential impacts of EVs on air quality are investigated for the greater metropolitan region of Perth in Western Australia (GMR Perth), an area with approximately 2 million people [35]. A map of the study area is provided below in Figure 1 based on published population data from the ABS [36]. This shows how population density varies across GMR Perth (panel A) and surrounding areas to the south (panel B). Areas outside GMR Perth were included to study potential co-effects related to changes in electricity generation for the wider region. The GMR Perth and surrounds provide a relevant case study for other cities across Australia, and more broadly for other countries, where fossil fuel power plants are still a major source of electricity supply.
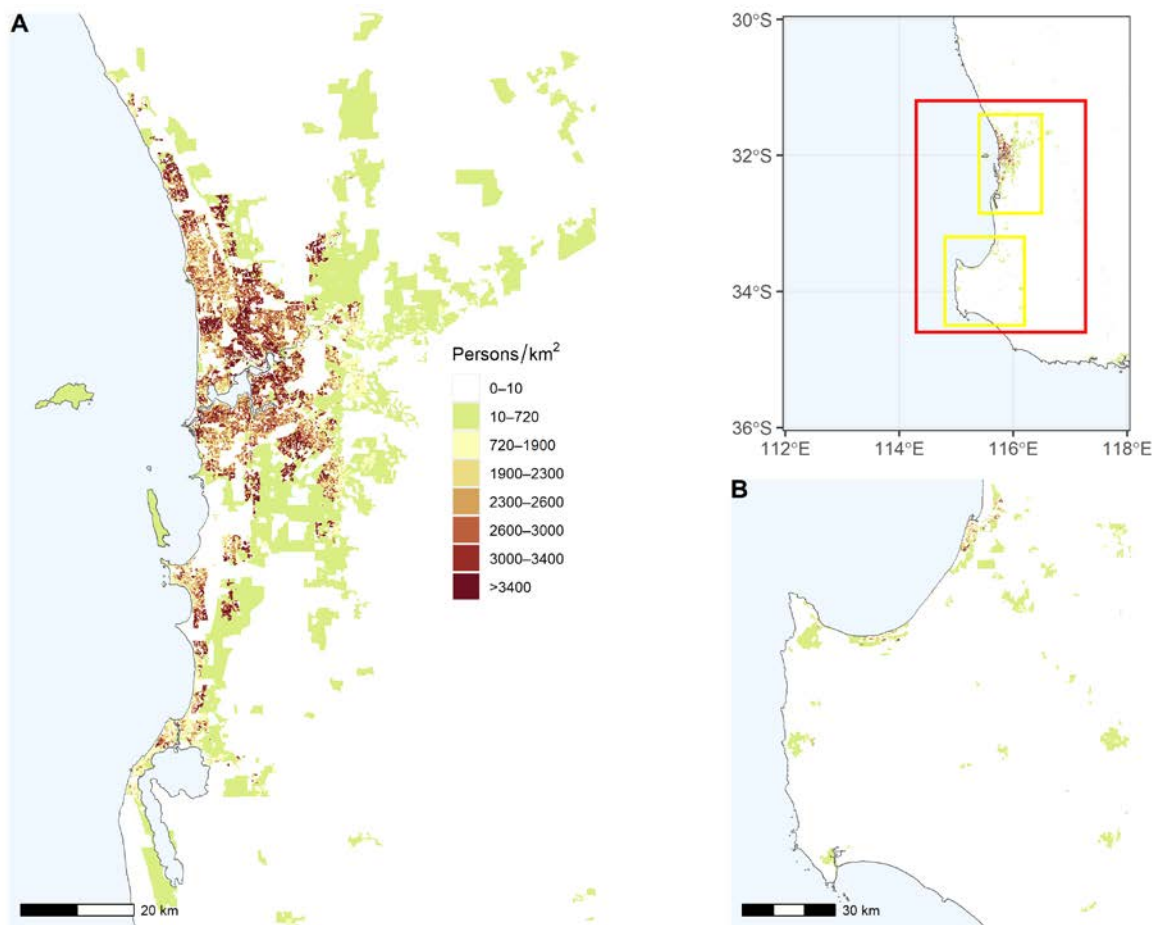
Figure 1. Study area and population. Map on top right shows the study region within the red rectangle and two sub-regions enclosed within yellow rectangles. Map on left shows population density for the city of Perth region (panel (**A**)) and map on lower right shows the population density for the southwest of the study area (**B**).

2.3. Emissions Data

Anthropogenic emissions for GMR Perth were obtained from the Department of Water and Environmental Regulation (DWER) [37] for the period July 2011 to June 2012. Total annual emissions data for on-road vehicle exhaust and electricity generation are presented in the appendix (Supplementary Materials, Tables S1 and S3). Details of emissions for all other sources of emissions for the study area are also provided in the appendix (Supplementary Materials Section S1.1). We refer to these data as the baseline emissions inventory.

The supply of electricity for the study region supports approximately 1 million households and is generated by the Southwest Inter-connected System (SWIS) of Western Australia. The locations of major power stations by energy type were obtained from the national electricity transmission lines and power stations database [38] as shown in Figure 2. The SWIS extends from Kalbarri in the north, to Kalgoorlie in the east and Albany in the south. Coal fired power stations for the region are located approximately 200 km south of GMR Perth. Conversely, gas turbines are spread across the SWIS network with several of the larger gas turbines located near main urban areas.
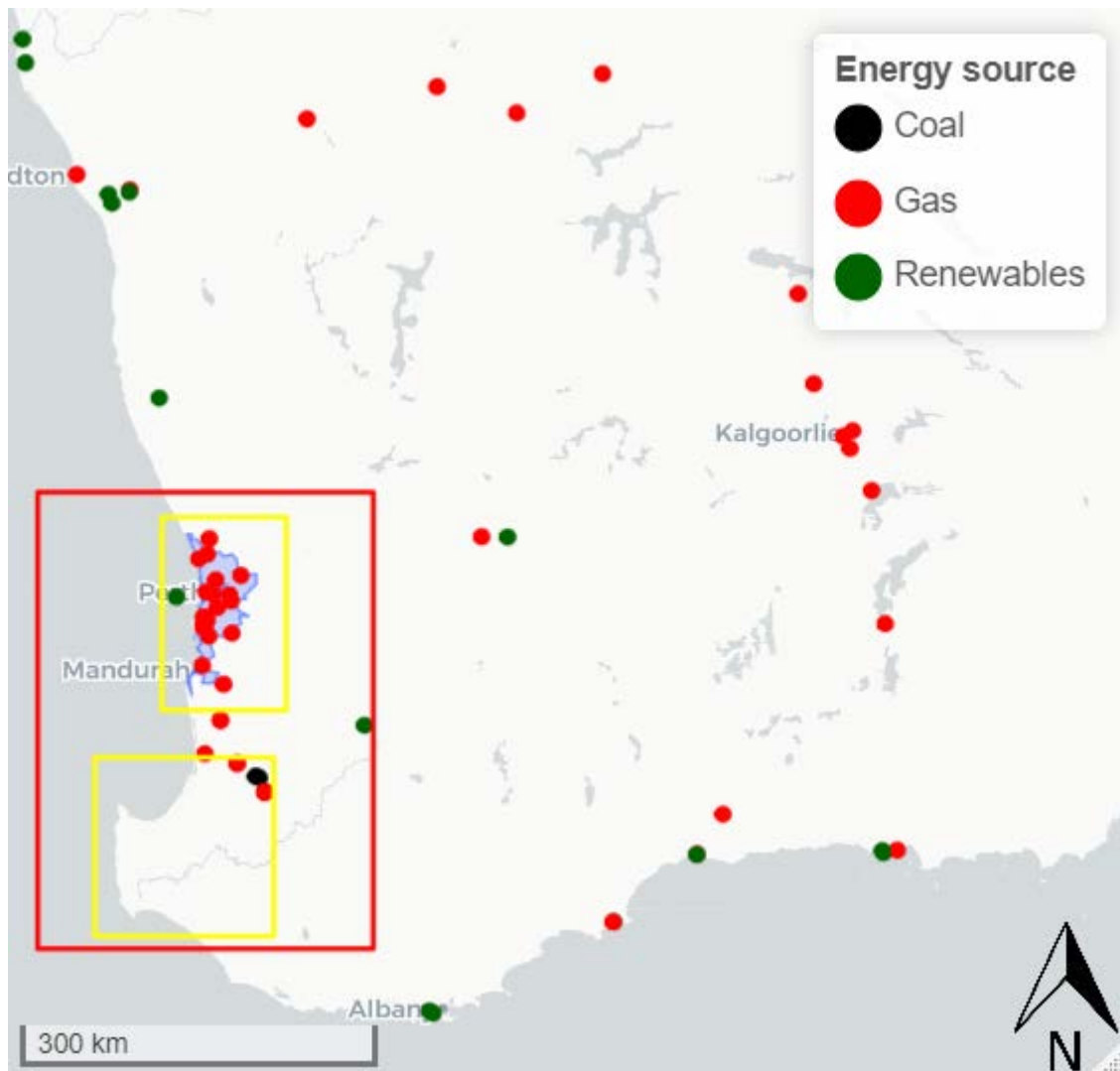
Figure 2. SWIS Network and location of power stations. Type of energy source shown as colour coded circles and the GMR Perth is shown in blue. Study region is enclosed within the red rectangle and two sub-regions (GMR Perth and region to southwest of Perth) enclosed within yellow rectangles.

2.4. Exposure Assessment

Annual average $PM_{2.5}$ (ug/m^3) were modelled using CCAM-CTM and results from the two inner domains were used to derive exposure for the study population. That is, data from the 1 km domain (for GMR Perth) and the 3 km domain for areas to the south. Gridded annual average $PM_{2.5}$ concentrations were first extracted to census mesh blocks, which are the smallest geographic areas defined by the Australian Bureau of Statistics (ABS) [39]. Figure 3 presents baseline annual mean $PM_{2.5}$ concentrations ($PM_{2.5}^{bc}$) for GMR Perth simulated by the CCAM-CTM from the baseline emissions inventory.
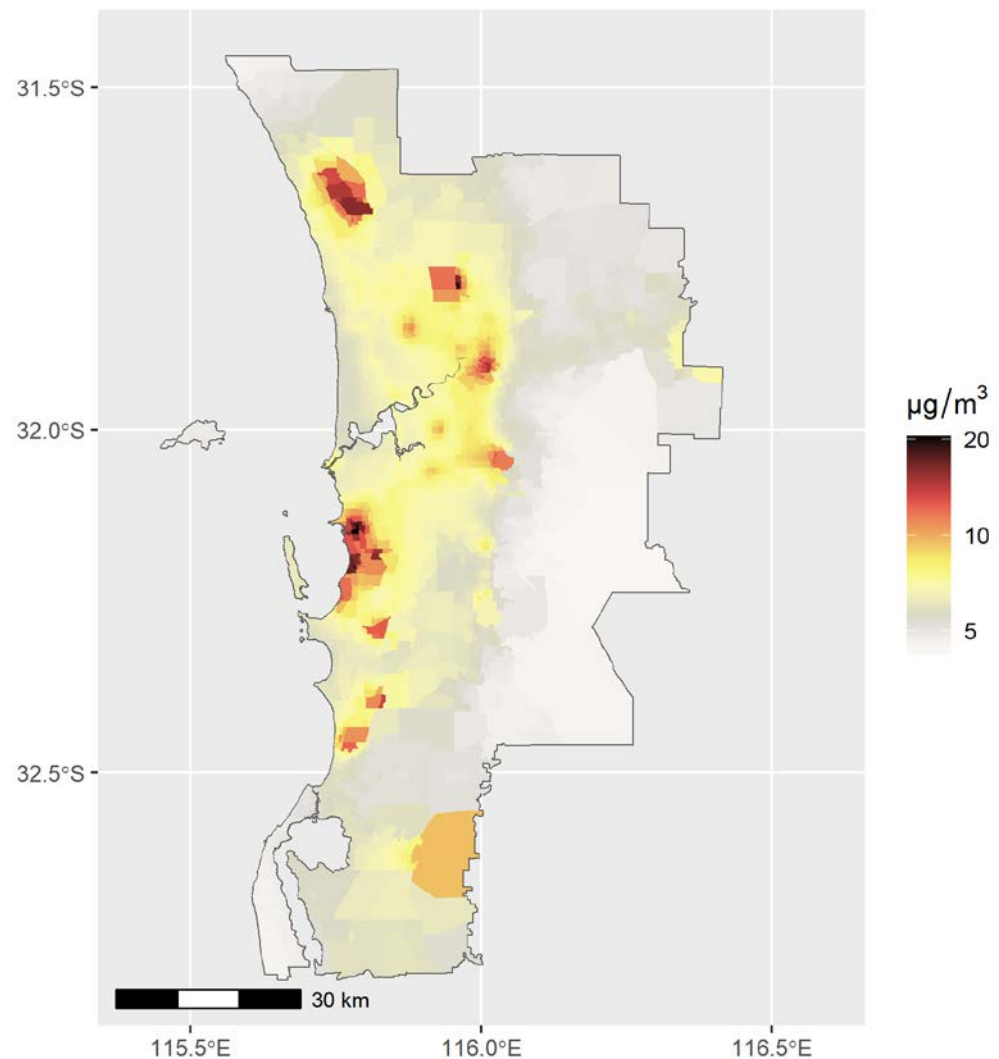
Figure 3. Baseline annual mean $PM_{2.5}$ ($PM_{2.5}^{bc}$) for GMR Perth generated by the CCAM-CTM from the baseline emissions inventory and extracted to ABS mesh blocks.

Estimates of the incremental change in annual average $PM_{2.5}$ concentration ($\Delta PM_{2.5}$) were produced by subtracting $PM_{2.5}^{bc}$ from any subsequent CTM simulation based on perturbations to emissions from the two sectors of interest. Thus, $\Delta PM_{2.5}$ can be interpreted as the change in $PM_{2.5}$ expected for a given change in on-road vehicle emissions and/or change in emissions produced by electricity generation.

For our case study, we base emulators on a spatial area unit that facilitates exploring intra-urban gradients but with sufficiently large populations to warrant HIA. We calculated mean $\Delta PM_{2.5}$ at a statistical area level 4 (SA4) for the GMR Perth and statistical area level 3 (SA3) for surrounding regions. SA3s are included to model net changes in $PM_{2.5}$ due to perturbations in emissions from electricity generation only. Thus, we have a total of six SA4s that make up GMR Perth and three selected SA3s in the surrounding region. We run the emulator in two modes. Firstly, without population weighting applied by equally weighting all model grids falling into each area. Secondly, by extracting to mesh blocks and population-weighting across mesh blocks. We chose to present the unweighted mean as an initial assessment, because we want to assess how well the emulator performs in terms of predicting the CCAM-CTM $\Delta PM_{2.5}$ across the study area without the added complexity of spatially weighting by population density. Population weighting is applied to assess how emulation would be used for policy development. Further details of this process and what each study area represents is provided in the appendix (Supplementary Materials Section S2.1).

2.5. Emulation

As noted in the introduction, emulation can be used to encapsulate many computationally demanding simulations of a complex model into an output that can be readily interrogated by non-modelling specialists. This approach uses scenario analysis to estimate the output of interest Y from the CCAM-CTM for a set of model inputs $\mathbf{X} = \{X_1, X_2, \ldots, X_d\}$, defined over a d-dimensional parameter space. For scenario S, we refer to a set of emission parameter inputs x and are interested in knowing the CCAM-CTM output $Y = f(x)$ but do not know this until the model has been run. An emulator estimates the CCAM-CTM simulation, thus providing $\hat{Y} = \hat{f}(x)$ where $\hat{f}$ is derived using Gaussian Process Emulation (GPE) based on a set of simulation runs (training data). Training data for an emulator are obtained by running different emission data input combinations of x through the CCAM-CTM to obtain the corresponding model outputs y.

The GPE model employs a Bayesian technique in which the prior probability distribution of f is conditioned on the training data to produce a posterior probability distribution [15,40]. The prior distribution is the Gaussian process with mean $m(x) = \mathrm{h}(x)^T \beta$ where $\mathrm{h}(\cdot)$ is a known function of x with unknown coefficients β. We set $\mathrm{h}(\cdot)$ here as the simple linear regression function with coefficients calculated based on the training data. The prior covariance function for any pair of parameter settings x and x' is defined by:

$$cov(x, x') = \sigma^2 c(x, x')$$

where

$$\mathrm{c}(x, x') = \exp\left\{ -\sum_{i=1}^{d} \left(\frac{(x - x')}{\delta_i} \right)^2 \right\}$$

is the Gaussian correlation function that depends on the distance between pairs of model input combinations. The smoothness of the model response to each emission parameter i, described by δ_i, is calculated from the training data and included in the model fitting using an uninformative prior. The resulting emulator $\hat{f}$ is a conditional probability distribution of Gaussian form that characterises the CCAM-CTM output with respect to the training data, having a posterior mean function given by:

$$m^*(x) = \mathrm{h}(x)^T \beta + \mathrm{t}(x)^T A^{-1}(\mathrm{y} - \mathrm{H}\hat{\beta})$$

and a posterior covariance matrix:

$$cov^*(x, x') = \sigma^2 [\mathrm{c}(x, x') - t(x)^T A^{-1} t(x') + (h(x) \\ - t(x)^T A^{-1} H)(H^T A^{-1} H)^{-1}(h(x') - t(x')^T A^{-1} H)^T]$$

where

$$y^T = (f(x_1), \ldots, f(x_n))$$

$$H^T = (h(x_1), \ldots, h(x_n))$$

$$A = \begin{pmatrix} 1 & c(x_1, x_2) & \ldots & c(x_1, x_n) \\ c(x_2, x_1) & 1 & & \vdots \\ \vdots & & \ddots & \\ c(x_n, x_1) & \ldots & & 1 \end{pmatrix}$$

$$t(x)^T = (c(x, x_1), \ldots, c(x, x_n))$$

$$\hat{\beta} = (H^T A^{-1} H)^{-1} H^T A^{-1} y$$

and

$$\hat{\sigma}^2 = \frac{y^T \left(A^{-1} - A^{-1} H (H^T A^{-1} H)^{-1} H^T A^{-1} \right) y}{n - q - 2}$$

where n is the number of training runs and q is the number of model parameters plus 1.

O'Hagan et al. [14] present a full derivation of these equations and Johnson et al. [15] provide more discussion relating to the underlying parameters involved when fitting GPE models.

For our case study, a total of eight emission parameters were selected, as summarised in Table 1. Oxides of nitrogen tailpipe emissions (NOx_EMS) and tailpipe $PM_{2.5}$ emissions (PM25_EMS) are included in the emulator (parameters X1 and X3) as they are both well-established contributors to $PM_{2.5}$ concentrations from on-road vehicles. Volatile organic compounds from on-road vehicles (VOC_EMS) is included (parameter X2) as it is well known that VOCs react with other pollutants to form secondary $PM_{2.5}$ [41]. The impacts of carbon monoxide (CO), as a primary pollutant have been known for a long time [42]. However, given that vehicles emit large quantities of CO emissions (CO_EMS), we decided to include as parameter X4 to investigate whether it may have an indirect influence on secondary aerosol production via a contribution to the urban photochemical reactivity. Although less dominant (on a mass emission basis) compared to the pollutants already discussed, sulphur dioxide (SO2_EMS) and ammonia emissions (NH3_EMS) were also included to investigate their role for this case study (parameters X5 and X6). Emissions for each pollutant were varied in a relative manner uniformly across the entire spatial domain. For example, NOX_EMS was varied in the range zero to 1.5 times the base inventory.

Table 1. Model parameters and their distribution ranges for this study.

Parameter No.	Abbreviated Name	Description	Range *
X1	$PM_{2.5}$	Particulate matter tailpipe emissions	0 to 150%
X2	NO_x	Oxides of nitrogen tailpipe emissions	0 to 150%
X3	VOC	Volatile organic compound emissions	0 to 150%
X4	CO	Carbon monoxide tailpipe emissions	0 to 150%
X5	SO_2	Sulphur dioxide tailpipe emissions	0 to 150%
X6	NH3	Ammonia emissions tailpipe emissions	0 to 150%
X7	GAS	Emissions from gas fire power stations	0 to 150%
X8	COAL	Emissions from coal fire power stations	0 to 150%

* Range over which parameter was varied as a percentage of baseline emissions.

Emissions from both gas and coal fired power stations (GAS_EMS; COAL_EMS;) were also varied and included in the emulator (parameters X7 and X8) in order to investigate the impact of electricity generation on $PM_{2.5}$ concentrations across the broader study area. In the case of GAS_EMS, emissions were varied across the SWIS relative to the magnitude of emissions from each turbine, i.e., as a percentage of the baseline emissions prescribed for the modelled inventory year.

We varied the emission parameters using a Latin hypercube sampling approach [43] to produce a series permuted emissions datasets according to the two selected emission sectors of interest (on-road vehicles and electricity generation). The CCAM-CTM was then applied to each emission configuration to obtain a series of gridded estimates of $PM_{2.5}$ and $PM_{2.5}^{bc}$ was subtracted to give $\Delta PM_{2.5}$. For each area of interest, these training data were used to fit separate emulators with and without population weighting as described in Section 2.4. Thus, for any given area, $\Delta PM_{2.5}$ is bounded below/above by the CCAM-CTM predictions with all emission parameters set to their minimum/maximum values. In our case study, the lower bound effectively represents an estimate of $PM_{2.5}$ attributed to the two selected emission sectors.

For each area of interest, these training data were used to fit separate emulators with and without population weighting as described in Section 2.4.

Guidance on how many training runs is needed to achieve a reasonable fit for the emulator is provided by Loeppky et al. [44] whereby n = 10 × d (where d is equal to number of parameters included in the emulator model) is generally recommended. For our case

study, we found that 32 runs of the CCAM-CTM achieved good emulator accuracy and we therefore decided not to continue with further simulations.

For validation, a second set of 16 CCAM-CTM runs were generated using the maximum Latin hypercube experiment design algorithm to ensure good coverage of the parameter space (validation data). The validation data effectively contain output from the CCAM-CTM in order to test how well an emulator predicts CCAM-CTM output for the 16 emission input combinations, which have purposely been selected to differ from those used for training the emulators. The validity of each fitted emulator model was assessed using methods outlined in Bastos and O'Hagan [45]. This entailed evaluating how well the emulator predicted the validation data using a series of scatter plots and comparing emulator mean predictions versus the actual CCAM-CTM output. Uncertainty around emulator predictions is calculated by using a defined critical level which is typically 5% thus providing 95% confidence intervals, however, this can be defined by the user. Validation includes calculating how many of the CCAM-CTM model output lie outside the 95% confidence bounds, and by comparing to the critical level. A listing of the emission scenarios used to train and validate the emulator is presented in the appendix (Supplementary Materials, Section S2.2).

Each emulator defines a fitted surface incorporating uncertainty, with respect to the CCAM-CTM outputs, across the eight-parameter uncertainty space. An attractive property of the GPE approach is that each fitted surface will pass through the training data points with no uncertainty whilst the emulator interpolates $\Delta PM_{2.5}$ from the CCAM-CTM to provide an estimate together with a measure of uncertainty for this prediction. We also present results for secondary inorganic aerosols (SIA) as an additional pollutant, in order to explore how the same emission parameters can be used to predict changes to components of $PM_{2.5}$.

A variance-based sensitivity analysis was carried out to determine how each emission parameter contributed to the overall variance in CCAM-CTM output. This method uses the emulator to explore how CCAM-CTM outputs vary with respect to individual contributions (main effects) of each parameter to $\Delta PM_{2.5}$ and also to quantify possible interactions between them. A detailed overview of this method has been provided by Saltelli et al. [46] and Oakley and O'Hagan [47], with the latter being one of the first articles to suggest using an emulator in place of a complex computer model for the purpose of sensitivity analysis.

All statistical analysis, developing, and validation of the emulator was carried out using R 4.2.1. [48] Each emulator was fit using the km () function from the DiceKriging package [49]. The extended-FAST (Fourier Amplitude Sensitivity Test) approach [50] was used for sensitivity analyses and implemented using the methods available from the R package sensitivity [51].

3. Results

3.1. Assessment of the Emulator

Figure 4 shows results from the emulator validation using the unweighted annual average $\Delta PM_{2.5}$ for GMR Perth. Emulator predictions are plotted against CTM output for each of the 16 validation runs and shown as a red circle. Uncertainty in the emulation with respect to the CTM is calculated as a 95% uncertainty interval for the emulator prediction and is presented as error bars. The largest decrease from baseline in $PM_{2.5}$ predicted by the CTM was 0.2 ug/m³ corresponding to validation run #15 whereby emission parameters were generally set lower than baseline emission levels. In this case, the emulator uncertainty is less than 2% of the CTM prediction. Figure 4 demonstrates that the emulator performs very well given the narrow uncertainty intervals and that the true CCAM-CTM output lies within these 95% uncertainty intervals for 15 of the 16 validation runs. Results of the fitted emulator applied to population weighted annual average $\Delta PM_{2.5}$ similarly showed accurate predictions of the CCAM-CTM with uncertainty intervals for all but one of the predictions containing the true CCAM-CTM output (Supplementary Materials Figure S13, Section S3.1). Weighting exposure by population yielded higher mean concentrations,

as would be expected given highest concentrations of on-road vehicle related $PM_{2.5}$ are generally in areas nearer to roads and corresponding higher population density.
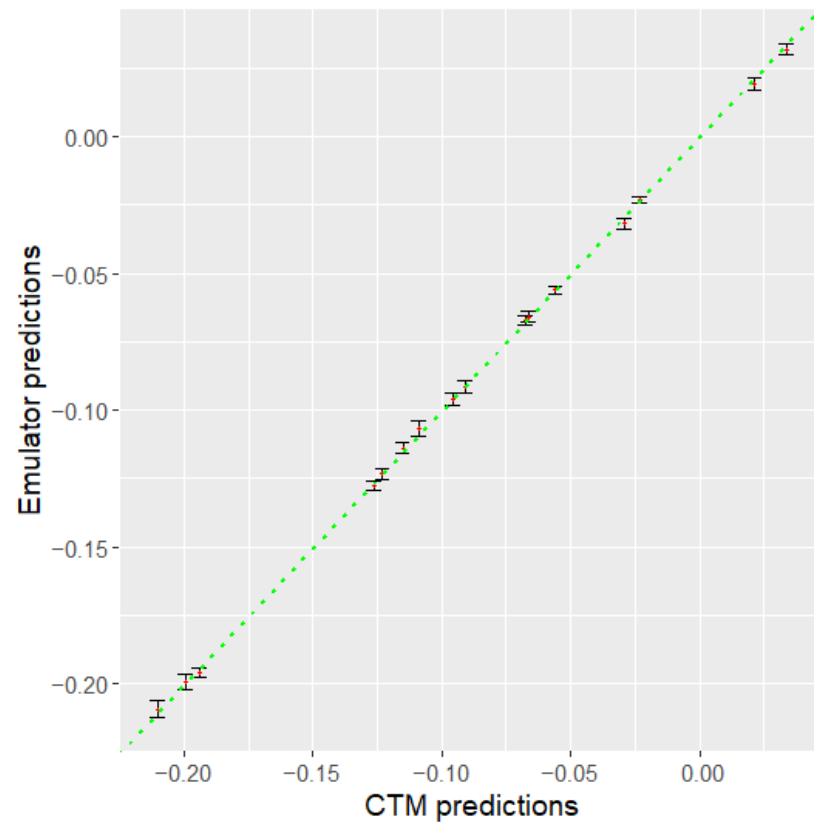

Figure 4. Emulator predictions of $\Delta PM_{2.5}$ ($\mu g/m^3$) in Greater Perth versus the CTM. The CCAM-CTM output (*x*-axis) is plotted against the corresponding emulator predictions (*y*-axis) for the validation dataset, with 95% confidence bounds on the emulator predictions. The green line is a line of equality.

Similar assessments were made of the emulators developed for different spatial scales including SA4s within GMR Perth, SA3s for surrounding areas, and for selected mesh block areas. These results (not shown) indicate that the emulators provide a very accurate estimate of the CCAM-CTM across a range of pollutant outcomes including nitrogen dioxide, ozone, secondary organic aerosols, and secondary inorganic aerosols. and at differing spatial scales.

3.2. Associations between Parameters and Output

The relationships between mean annual $\Delta PM2.5$ and each parameter for GMR Perth are plotted in Figure 5 based on equally weighting all grids from the inner CTM domain that fall inside the study area. The plot shows how much the emulator predicts the CTM to vary for a given change in each parameter, while holding the other seven parameters fixed at their baseline value. This effectively allows us to quantify the partial differential of $\Delta PM_{2.5}$ for a given change in a selected emission parameter. Note that the emulator also allows us to investigate whether joint combinations of changes to the parameters affect the output differently, e.g., by fixing other parameters at different values. However, results (not shown) of such analyses were very similar to that presented here indicating that associations were additive. There is a clear positive linear association between $PM_{2.5}$ tailpipe emissions from on-road vehicles (PM25_EMS) and $\Delta PM_{2.5}$ with an estimated annual decrease from baseline of 0.12 ug/m³ if emissions from this source were reduced to zero. There is also a clear linear association between emissions from coal-fired power stations (COAL_EMS) with an estimated decrease of 0.06 ug/m³ if emissions from this source were reduced to zero. Ammonia emissions from on-road vehicles (NH3_EMS) has a relatively moderate

association with an estimated decrease of 0.04 ug/m^3 if emissions from this source were reduced to zero. There were small associations with nitrogen oxide emissions from on-road vehicles and emissions from gas turbines. The other three parameters (VOC_EMS, CO_EMS, and SO2_EMS) show little/no association with $\Delta PM_{2.5}$.
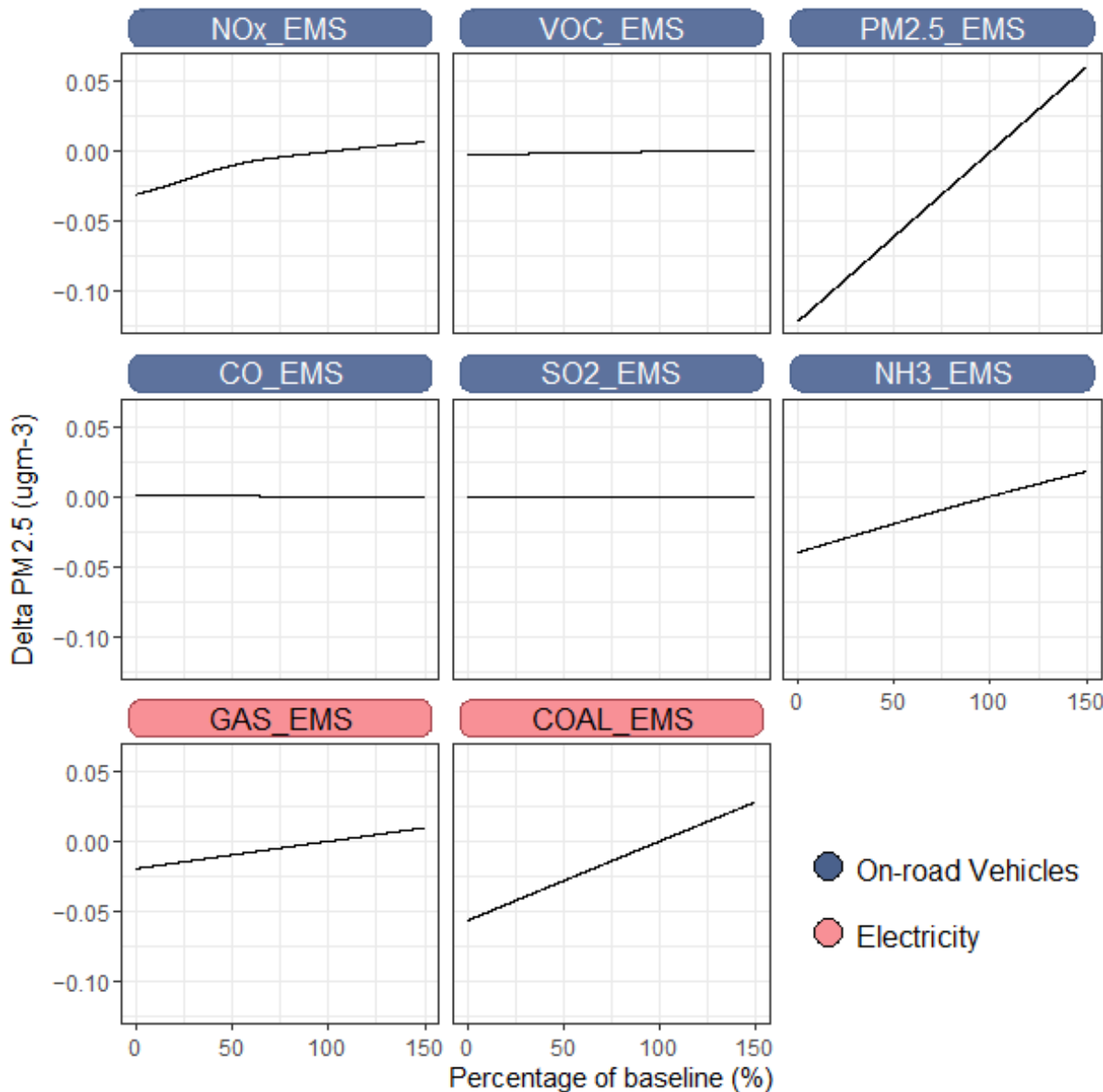

Figure 5. Relationship between mean annual average delta $PM_{2.5}$ concentration and each of the 8 parameters for GMR Perth. Each plot is based on all other parameters held fixed at their baseline level.

There is considerable spatial variation in associations between the eight parameters and predicted $\Delta PM_{2.5}$. To illustrate this, three mesh block areas were selected where we would expect to see contrasting results for the two sectors studied (Supplementary Materials Section S3.2). The association for PM25_EMS is steepest for the mesh block where on-road vehicle $PM_{2.5}$ emissions are highest, i.e., central Perth. Conversely, GAS_EMS was the strongest predictor for the selected mesh block located in an area with heavy industry. For the third selected mesh block, located near the northern fringe of GMR Perth, associations of a comparable magnitude were identified for NOx_EMS, PM$_{2.5}$_EMS, NH3_EMS, GAS_EMS, and COAL_EMS.

3.3. Sensitivity Analyses

For GMR Perth overall, variance-based sensitivity analyses showed that 72% of the variance in $\Delta PM_{2.5}$, across the emission parameters studied, is due to PM_EMS. Similarly, 17% of the variance is due to COAL_EMS, 7% due to NH3_EMS, 3% from NOx_EMS, and a further 2% due to GAS_EMS. Interaction effects between the eight parameters summed to less than 1%. To investigate the extent to which variance contributions varied spatially, sensitivity analyses were carried out based on the SA4 emulators. Figure 6 presents the main effects of the eight parameters for each of the six SA4 regions. COAL_EMS explains 85% of the variance in $\Delta PM_{2.5}$ for the SA4 located to the south of GMR Perth (Mandurah) where on-road traffic is lowest. Similarly, the proportion of variance explained by PM_EMS was lower for SA4 regions with less on-road traffic.
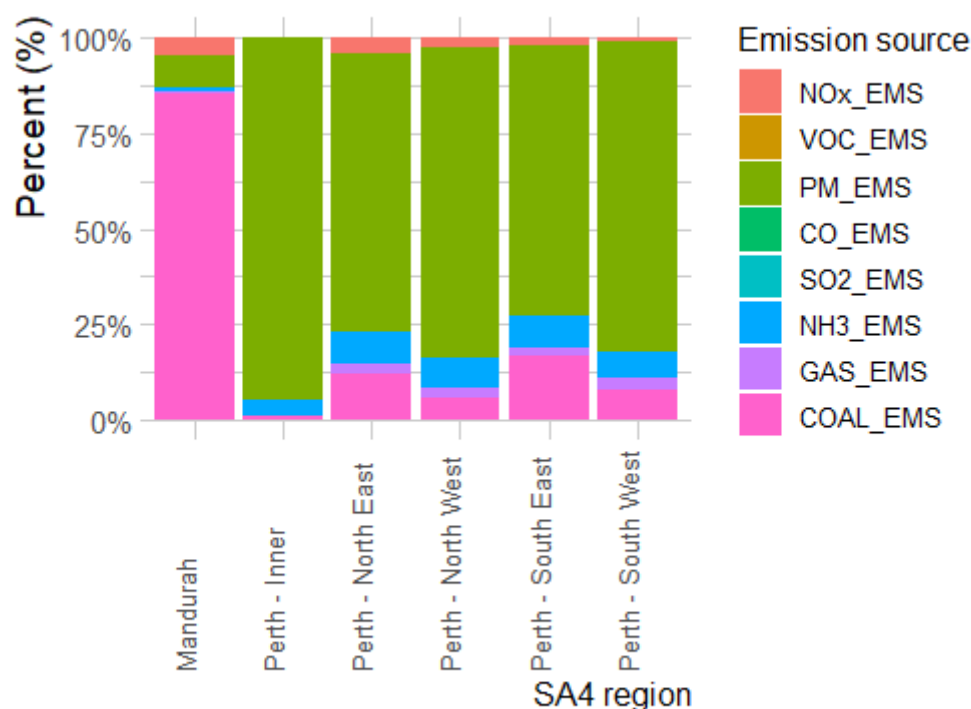

Figure 6. Percentage contribution to $\Delta PM_{2.5}$ variance by SA4 region within GMR Perth.

Applying the same steps as above for ΔSIA demonstrated that emulation also predicted the CCAM-CTM well with small emulator uncertainty (Supplementary Materials Section S3.3). Results of a sensitivity analyses based on SA4 emulators constructed for ΔSIA are also presented in the appendix. These analyses indicated that the largest contribution to variation in ΔSIA across GMR Perth is COAL_EMS (75%) followed by NOx_EMS (20%). The remaining variance is shared across $PM_{2.5}$_EMS, NH3_EMS, and GAS_EMS with around 1.5% of the variance explained by interactions between emission parameters. Although COAL_EMS is the main contributor to ΔSIA overall, differences by SA4 were noted whereby NOx_EMS and PM_EMS explained as much as 40% of the variation.

3.4. Scenario Analyses

A series of hypothetical emission scenarios are presented to demonstrate how emulation can be used to carry out scenario analysis for policy development. Specifically, we evaluate potential reductions in $PM_{2.5}$ for GMR Perth in future scenarios involving a transition to EVs and the use of renewables for electricity generation. For our case study, we employed SA4 emulators based on population weighted exposure data to assess impacts on areas within GMR Perth. We then use SA3 emulators to assess co-effects on Perth surrounds associated with changes to emissions from the electricity generation.

In order to quantify the emissions for each scenario, data from the DWER inventory (refer Table S1) were used to define expected reduction for each of the six major on-road emission precursors resulting from a city-wide transition to EVs. A business as usual (BAU) was defined as the expected increase in emissions from the SWIS to meet increased demand in electricity assuming the same energy mix as reported in 2012 (refer Table S2). That is, the expected increase in emissions produced by the SWIS, compared to baseline, assuming on-road vehicles in GMR Perth are replaced with EVs. Electricity consumption by EVs in the SWIS is projected to be approximately 14 TWh by 2050 [52] which equates to 50% more emissions generated by coal and gas turbines under the BAU. We then considered what the impact would be if electricity generation were transitioned to 100% renewable energy by setting X7 and X8 to zero. Thus, each scenario was defined as a specific emission data input combination which was then offered to each emulator yielding predictions of $\Delta PM_{2.5}$. This gave us four scenarios to compare as follows:

S1.   All on-road diesel vehicles switched to electric under BAU electricity generation
S2.   All on-road petrol vehicles to electric under BAU electricity generation
S3.   All on-road vehicles switched to electric under BAU electricity generation
S4.   All on-road vehicles switched to electric combined with 100% renewable electricity

Results from this set of scenario analyses are presented in Table 2. Under BAU for SWIS (Scenario S3), the emulation approach predicts the largest impact on $PM_{2.5}$ would be for the population living in Perth-Inner which would be 0.69 ug/m^3 lower than the baseline. The dominant contribution from replacing diesel vehicles is expected given that PM_EMS is the strongest predictor of $\Delta PM_{2.5}$ especially for the Perth—Inner region. Impacts for the other five areas within GMR Perth are predicted to be smaller in magnitude. For example, under scenario S1, predicted impacts from replacing diesel vehicles for other SA4s ranges from 0.42 ug/m^3 lower $PM_{2.5}$ for Perth—South East down to 0.09 ug/m^3 lower $PM_{2.5}$ for Mandurah. The net impact on GMR Perth under scenario S3 is estimated to be 0.44 ug/m^3 lower PM2.5. When combined with a transition to 100% renewable energy (scenario S4), this reduces further to 0.56 ug/m^3 lower PM2.5. Under the BAU for SWIS, the predicted increase in $PM_{2.5}$ is 0.06 ug/m^3 higher for regions to the south of GMR Perth with larger effects predicted for the Bunbury region. Conversely, SA3 emulators predict that a switch to 100% renewables for electricity generation would result in 0.12 ug/m^3 lower $PM_{2.5}$ for populations living in these regions.

Table 2. Estimated change in $PM_{2.5}$ concentration * by area for selected scenarios.

Scenario	S1 [ad]	S2 [bd]	S3 [cd]	S4 [e]
Geographical area [f]				
SA4 areas, population size (000 s)				
Perth—Inner (170)	−0.60	−0.06	−0.69	−0.81
Perth—North East (251)	−0.37	−0.04	−0.42	−0.53
Perth—North West (535)	−0.39	−0.05	−0.45	−0.57
Perth—South East (488)	−0.42	−0.05	−0.48	−0.60
Perth—South West (403)	−0.31	−0.02	−0.35	−0.49
Mandurah (97)	−0.09	0.01	−0.08	−0.23
GMR combined (1944)	−0.38	−0.04	−0.44	−0.56
SA3 areas, population size (000 s)				
Bunbury (102)	0.02	0.06	0.08	−0.16
Manjimup (23)	0.01	0.02	0.03	−0.06
Augusta—MR—Busselton (51)	0.01	0.02	0.03	−0.06
Surrounding areas combined (176)	0.01	0.05	0.06	−0.12

* Population weighted mean annual change in $PM_{2.5}$ concentration ($\mu g/m^3$). [a] tailpipe emissions from on-road diesel vehicles within Greater Perth reduced to zero. [b] Tailpipe emissions from on-road petrol vehicles within Greater Perth reduced to zero. [c] Tailpipe emissions from all on-road vehicles within Greater Perth reduced to zero. This includes vehicles with LPG fuel. [d] BAU = business-as-usual: emission levels from the SWIS increase in line with that expected to meet the increased demand in electricity assuming the same energy mix as reported in 2012. [e] Tailpipe emissions from on-road petrol vehicles within Greater Perth reduced to zero combined with electricity generation from 100% renewables. [f] Geographical areas are Statistical Area Level 4 (SA4) for GMR Perth and Statistical Area Level 4 (SA3) for areas considered in surrounding areas to the South.

4. Discussion

This study describes the use of statistical emulation for assessing impacts on air quality from interventions that target emissions from selected sectors. Our case study illustrates how emulation can be used to assess impacts on within-city pollutant concentrations from changes to emissions from on-road vehicles and the electricity sector. We show that emulators can accurately predict model output from the CCAM-CTM and quantify the relative contribution of emission precursors on $PM_{2.5}$ concentrations. The main contributions that this research offers are twofold. Firstly, it shows how emulation can be used to study exposure to air pollution at an intra-urban scale. Secondly, the case study highlights how emulators can be used separately from the more complex air quality model making scenario analysis more accessible to a wider audience.

An important advantage of complex air quality models is their ability to separate source contributions and quantify the impact of changes to emissions on pollutant concentration levels. The direct way to achieve this is to modify emissions according to a scenario of interest, re-run the model, and evaluate the results. However, running the complex model for many scenarios will be limited by computational and time resources available, particularly if full-year simulations are required. Furthermore, simulation outputs from models such as the CCAM-CTM or CMAQ are used by health professionals and policy specialists, but it is not straight forward to run a CTM. Making emulators available alongside simulation outputs would provide the policy analyst with an alternative approach without the need to interact directly with the complex model.

Air quality in Australia's urban centres is generally good relative to other countries [53] yet scientific evidence has linked air pollution to health effects in populations that have lower concentrations of exposure [54–56]. There is ongoing debate in Australia around the concept of "no safe limit" and whether current ambient air quality standards will be further tightened [57]. This makes assessing the effectiveness of policies, such as increasing the adoption of EVs, important for cities like Perth. To give further context to our case study, a mean $\Delta PM_{2.5}$ of -0.5 ug/m^3 is consistent with what source apportionment studies conducted in other Australian jurisdictions have reported for the same emission sources considered here [21]. The scenario analysis presented in Section 3.4 is only one of many possible permutations needed to properly evaluate such policies. In practice, a more in-depth analysis may reveal how impacts on air quality can depend on many factors including location, vehicle type, fuel type, and emission source used to meet increased demand in electricity. Specific examples include further examining different sub-sectors of the on-road transport industry such as heavy truck, buses, sport utility vehicles, etc. We illustrate to what extent switching on-road vehicles in GMR Perth to EVs may result in co-effects for populations living in surrounding areas. This could be further explored by employing emulators at a more local area level than administrative areas (SA4/SA3) as presented here. Insights into many such questions can be easily investigated directly with emulators by setting emission parameters to the specified level(s) of interest.

The concept of substituting the complex air quality model with an alternative surrogate model to enable more rapid scenario analysis is not new. This research area has led to the development of surrogate models such as AERIS [58] and TM5-FASST [10] both of which have been shown to reliably predict the more complex model. We chose a statistical emulation approach as it provides a direct way to study associations between a set of input parameters and model outputs. Our case study highlights this by presenting associations of the selected emission parameters with $PM_{2.5}$ concentration and how they varied spatially across different areas within GMR Perth. Of the two sectors considered here, on-road vehicles and electricity generation, $PM_{2.5}$ emissions from vehicle exhaust was the dominant predictor of change in $PM_{2.5}$ concentrations. This was expected as PM emissions are still relatively high despite recent improvements to vehicle combustion engines [59]. Furthermore, secondary organic aerosols tend to be the dominant source of secondary particulate matter formation in Australian settings [60]. We included VOC emissions from vehicles as a parameter as we were not sure what to expect in terms of

its role in formation of secondary particulates in an urban setting. Furthermore, we also wanted to present a complete set of the major emission precursors. In future iterations of this research, we plan to examine reactive VOC components more closely and consider other natural sources of VOC.

Statistical emulation provides a way to understand outputs from complex processes more generally [61]. The variance-based sensitivity analysis utilises the emulator to quantify the contribution of multiple factors to variability in model output. This approach allows us to assess the sensitivity of a CCAM-CTM output to individual input parameters and can identify potential nonlinear effects and interactions. While associations were generally found to be linear and close to additive in our case study, this may not be the case in other applications. One such area of research relates to studying the role of NH_3 and secondary inorganic aerosol (SIA) formation. In combination with NOx and $PM_{2.5}$, NH_3 is important due to its role as a precursor of ammonium particulates [62], however, the chemistry between these precursor pollutants is highly nonlinear [63]. Recently, it has been suggested that vehicular emissions of NH_3 may have an important role in particle formation within an urban environment despite accounting for only a small fraction of NH_3 emissions compared to other sources such as agriculture [64]. This topic may become even more relevant in the future given ammonia is being considered as a versatile energy vector of hydrogen production [65].

Depending on the research question, the use of emulation as applied in our case study can be redesigned in other ways. For example, in the context of transitioning to electric vehicles, a more policy relevant question might be to think more directly in terms of vehicle fuel efficiency rather than total emission levels. In this way, an emulator might be constructed to reflect characteristics of the vehicle itself, (e.g., engine size, age and fuel type) or we could define a different emulator for each vehicle type, and then include total Vehicle Kilometres Travelled (VKT) in the emulator model as a scaling factor. Another related research question relevant to on-road vehicle emissions relates to non-exhaust particulates. This is an area that is attracting more attention, particularly given that vehicle non-exhaust emissions may play a more dominant role on local air pollution in the future [66]. An emulator could be developed to investigate emissions that result from brake wear, tyre wear, and resuspended road dust. Exposure to non-exhaust emissions could be studied in this way by including parameters in the emulator such as vehicle weight, tyre characteristics, and road surface condition. Finally, we chose to work with annual average $PM_{2.5}$ concentrations for our case study as the focus of HIAs is often chronic health effects of long-term exposure. Future work could investigate how emulation can be applied to incorporate other time scales (seasonal, monthly, daily) to study how contribution of emission sources depends on time and/or to more generally study short-term health effects of an intervention.

There is an increasing awareness from local communities around the health impacts of air pollution [67]. At the same time, researchers are now collecting more data at a finer scale using advanced monitoring technologies [68]. Thus, there is a growing need for models that can not only utilise such data, but which can also be used alongside to assist with understanding their outputs. We plan to further investigate the use of emulation in other applications involving air quality related HIAs including exposure to air pollution from wood heater smoke. Emulators could provide a flexible way to address policy questions at a local community level such as evaluating the cost effectiveness of more compliant and/or low emission wood heaters. Similarly, emulation may be useful to complement chemical transport models as a way of examining the health impacts on vulnerable populations of smoke and other pollutants from prescribed burns.

5. Conclusions

The results of this research show that statistical emulation can be used to study interventions on air quality at an intra-urban scale. Our case study demonstrates that GPE performs very well in terms of reproducing simulation output from the CCAM-CTM

when applied to a typical Australian urban setting. Emulators can potentially supplement complex simulation models in other applications by facilitating fast and reliable assessments of interventions that improve air quality. If put into practical use, this approach could substantially reduce the human resources and costs involved in formalising air pollution control strategies. Further research is warranted that aims to apply emulation science as a pathway to looking at policy.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/atmos13122009/s1. CCAM-CTM setup and verification; Additional information on how the emulator is developed; Supplementary results. Table S1: Summary of tailpipe emissions data for on-road vehicles. Table S2: Summary of emission sources used for the CTM. Table S3: Summary of SWIS electricity generation and emissions. Table S4: CCAM-CTM domain definitions. Table S5: CCAM-CTM training runs used to fit emulators. Table S6: CCAM-CTM validation runs used to validate emulators. Figure S1: Spatial summary of on-road vehicle tailpipe $PM_{2.5}$ emissions. Figure S2: CTM computational domains. Figure S3: Scatter plots of observed and modelled temperature. Figure S4: Rose plots of modelled wind speed and wind direction. Figure S5: Rose plots of observed wind speed and wind direction. Figure S6: Probability density plots of observed and modelled wind direction. Figure S7: Diurnal plots by monitor site of modelled (green) versus observed (red) $PM_{2.5}$. Figure S8: Hourly time series plots of observed and modelled 1-h $PM_{2.5}$. Figure S9: Scatter plots of observed and modelled daily $PM_{2.5}$ concentrations by season. Figure S10: Bugle plots of mean fractional bias for CCAM-CTM modelling. Figure S11: Bugle plots of mean fractional error for CCAM-CTM modelling. Figure S12: ASGS spatial boundaries within the study area. Figure S13: Emulator predictions of $\Delta PM_{2.5}$ ($\mu g/m^3$) in GMR Perth versus the CTM. Figure S14: Associations with $PM_{2.5}$ for selected mesh block area A located in central Perth. Figure S15: Associations with $PM_{2.5}$ for selected mesh block area B located in the Kwinana industrial zone. Figure S16: Associations with $PM_{2.5}$ for selected mesh block area C located near the northern border of North-West Perth. Figure S17: Emulator predictions of ΔSIA ($\mu g/m^3$) in Greater Perth versus the CTM. References [69–75] are cited in the Supplementary Materials.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Scripts and selected data to reproduce our emulation results are available in an online repository: Emulation of a CTM for Western Australia, (2022), GitHub repository, https://github.com/stevevdhoorn/emulation-perth, accessed on 2 February 2020. Requests to access to the complete emissions data inventory and air quality monitor data used to carry out and validate chemical transport model simulations must be made directly to the West Australian Department of Water and Environmental Regulation (https://dwer.wa.gov.au/, accessed on 2 February 2020). Emissions data on emissions from the southwest interconnected system are freely available from the online National Pollutant Inventory (https://www.dcceew.gov.au/environment/protection/npi, accessed on 2 February 2020).

Conflicts of Interest: All authors declare that they have no conflict of interest.

References

1. Hoek, G. Methods for Assessing Long-Term Exposures to Outdoor Air Pollutants. *Curr. Environ. Health Rep.* **2017**, *4*, 450–462. [CrossRef]
2. Carruthers, D.; Holroyd, R.; Hunt, J.; Weng, W.; Robins, A.; Apsley, D.; Thompson, D.; Smith, F. UK-ADMS: A new approach to modelling dispersion in the earth's atmospheric boundary layer. *J. Wind. Eng. Ind. Aerodyn.* **1994**, *52*, 139–153. [CrossRef]
3. Appel, K.W.; Napelenok, S.L.; Foley, K.M.; Pye, H.O.T.; Hogrefe, C.; Luecken, D.J.; Bash, J.O.; Roselle, S.J.; Pleim, J.E.; Foroutan, H.; et al. Description and evaluation of the Community Multiscale Air Quality (CMAQ) modeling system version 5.1. *Geosci. Model Dev.* **2017**, *10*, 1703. [CrossRef]
4. Morris, R.E.; Yarwood, G.; Wagner, A. Recent Advances in CAMx Air Quality Modelling. In *Air Pollution Modelling and Simulation*; Sportisse, B., Ed.; Springer: Berlin/Heidelberg, Germany, 2002; pp. 79–88. [CrossRef]
5. Matthias, V.; Arndt, J.A.; Aulinger, A.; Bieser, J.; Van Der Gon, H.D.; Kranenburg, R.; Kuenen, J.; Neumann, D.; Pouliot, G.; Quante, M. Modeling emissions for three-dimensional atmospheric chemistry transport models. *J. Air Waste Manag. Assoc.* **2018**, *68*, 763–800. [CrossRef]
6. Mann, M.; Bertok, I.; Borken-Kleefeld, J.; Cofala, J.; Heyes, C.; Höglund-Isaksson, L.; Klimont, Z.; Nguyen, B.; Posch, M.; Rafaj, P.; et al. Cost-effective control of air quality and greenhouse gases in Europe: Modeling and policy applications. *Environ. Model. Softw.* **2011**, *26*, 1489–1501. [CrossRef]
7. Askariyeh, M.H.; Khreis, H.; Vallamsundar, S. Air pollution monitoring and modeling. *Traffic-Relat. Air Pollut.* **2020**, *5*, 111–135. [CrossRef]
8. Henneman, L.R.F.; Dedoussi, I.C.; Casey, J.A.; Choirat, C.; Barrett, S.R.H.; Zigler, C.M. Comparisons of simple and complex methods for quantifying exposure to individual point source air pollution emissions. *J. Expo. Sci. Environ. Epidemiol.* **2020**, *31*, 654–663. [CrossRef]
9. Baker, K.R.; Amend, M.; Penn, S.; Bankert, J.; Simon, H.; Chan, E.; Fann, N.; Zawacki, M.; Davidson, K.; Roman, H. A database for evaluating the InMAP, APEEP, and EASIUR reduced complexity air-quality modeling tools. *Data Brief* **2019**, *28*, 104886. [CrossRef]
10. Van Dingenen, R.; Dentener, F.; Crippa, M.; Leitao, J.; Marmer, E.; Rao, S.; Solazzo, E.; Valentini, L. TM5-FASST: A global atmospheric source-receptor model for rapid impact analysis of emission changes on air quality and short-lived climate pollutants. *Atmos. Chem. Phys.* **2018**, *18*, 16173–16211. [CrossRef]
11. Napelenok, S.; Cohan, D.; Odman, M.; Tonse, S. Extension and evaluation of sensitivity analysis capabilities in a photochemical model. *Environ. Model. Softw.* **2008**, *23*, 994–999. [CrossRef]
12. Daniel, C.S.; Napelenok, S.L. Air Quality Response Modeling for Decision Support. *Atmosphere* **2011**, *2*, 407–425.
13. Simon, H.; Valin, L.C.; Baker, K.R.; Henderson, B.H.; Crawford, J.H.; Pusede, S.E.; Kelly, J.T.; Foley, K.M.; Chris Owen, R.; Cohen, R.C.; et al. Characterizing CO and NOy Sources and Relative Ambient Ratios in the Baltimore Area Using Ambient Measurements and Source Attribution Modeling. *J. Geophys. Res. Atmos.* **2018**, *123*, 3304–3320. [CrossRef]
14. O'Hagan, A. Bayesian analysis of computer code outputs: A tutorial. *Reliab. Eng. Syst. Saf.* **2006**, *91*, 1290–1300. [CrossRef]
15. Johnson, J.S.; Cui, Z.; Lee, L.A.; Gosling, J.P.; Blyth, A.M.; Carslaw, K.S. Evaluating uncertainty in convective cloud microphysics using statistical emulation. *J. Adv. Model. Earth Syst.* **2015**, *7*, 162–187. [CrossRef]
16. Lee, L.A.; Carslaw, K.S.; Pringle, K.J.; Mann, G.W.; Spracklen, D.V. Emulation of a complex global aerosol model to quantify sensitivity to uncertain parameters. *Atmos. Chem. Phys.* **2011**, *11*, 12253–12273. [CrossRef]
17. Conibear, L.; Reddington, C.L.; Silver, B.J.; Chen, Y.; Knote, C.; Arnold, S.R.; Spracklen, D.V. Sensitivity of Air Pollution Exposure and Disease Burden to Emission Changes in China Using Machine Learning Emulation. *GeoHealth* **2022**, *6*, e2021GH000570. [CrossRef]
18. Conibear, L.; Reddington, C.L.; Silver, B.J.; Chen, Y.; Knote, C.; Arnold, S.R.; Spracklen, D.V. Statistical Emulation of Winter Ambient Fine Particulate Matter Concentrations from Emission Changes in China. *GeoHealth* **2021**, *5*, e2021GH000391. [CrossRef]
19. Conibear, L.; Reddington, C.L.; Silver, B.J.; Arnold, S.R.; Turnock, S.T.; Klimont, Z.; Spracklen, D.V. The contribution of emission sources to the future air pollution disease burden in China. *Environ. Res. Lett.* **2022**, *17*, 064027. [CrossRef]
20. Mallet, V.; Tilloy, A.; Poulet, D.; Girard, S.; Brocheton, F. Meta-modeling of ADMS-Urban by dimension reduction and emulation. *Atmos. Environ.* **2018**, *184*, 37–46. [CrossRef]
21. Broome, R.A.; Powell, J.; Cope, M.E.; Morgan, G. The mortality effect of PM2.5 sources in the Greater Metropolitan Region of Sydney, Australia. *Environ. Int.* **2020**, *137*. [CrossRef]
22. *Training Materials and Best Practices for Chemical Weather/Air Quality Forecasting*; WMO: Geneva, Switzerland, 2020.
23. Cope, M.; Keywood, M.D.; Emmerson, K.; Galbally, I.E.; Boast, K.; Chambers, S.D.; Cheng, M.; Crumeyrolle, S.; Dunne, E.; Fedele, R. Sydney Particle Study-Stage-II, Study Undertaken by the Centre for Australian Weather and Climate Research (CAWCR) on behalf of the New South Wales Office of Environment and Heritage. 2014. Available online: https://www.environment.nsw.gov.au/-/media/OEH/Corporate-Site/Documents/Air/sydney-particle-study-2010-13.pdf#page=3&zoom=auto,-19,367 (accessed on 13 November 2022).
24. Mcgregor, J.L.; Dix, M.R. The CSIRO Conformal-Cubic Atmospheric GCM. In Proceedings of the IUTAM Symposium on Advances in Mathematical Modelling of Atmosphere and Ocean Dynamics, Limerick, Ireland, 2–7 July 2000; Springer: Dordrecht, The Netherlands, 2001; pp. 197–202.

25. Mann, G.W.; Carslaw, K.S.; Spracklen, D.V.; Ridley, D.A.; Manktelow, P.T.; Chipperfield, M.P.; Pickering, S.J.; Johnson, C.E. Description and evaluation of GLOMAP-mode: A modal global aerosol microphysics model for the UKCA composition-climate model. *Geosci. Model Dev.* **2010**, *3*, 519–551. [CrossRef]

26. Donahue, N.M.; Robinson, A.L.; Stanier, C.O.; Pandis, S.N. Coupled partitioning, dilution, and chemical aging of semivolatile organics. *Environ. Sci. Technol.* **2006**, *40*, 2635–2643. [CrossRef] [PubMed]

27. Nenes, A.; Pandis, S.; Pilinis, C. ISORROPIA: A New Thermodynamic Equilibrium Model for Multiphase Multicomponent Inorganic Aerosols. *Aquat. Geochem.* **1998**, *4*, 123–152. [CrossRef]

28. Meyer, C.P.; Luhar, A.K.; Mitchell, R.M. Biomass burning emissions over northern Australia constrained by aerosol measurements: I—Modelling the distribution of hourly emissions. *Atmos. Environ.* **2008**, *42*, 1629–1646. [CrossRef]

29. Gong, S.L. A parameterization of sea-salt aerosol source function for sub- and super-micron particles. *Glob. Biogeochem. Cycles* **2003**, *17*, 1097. [CrossRef]

30. Lu, H.; Shao, Y. A new model for dust emission by saltation bombardment. *J. Geophys. Res. Atmos.* **1999**, *104*, 16827–16842. [CrossRef]

31. Cope, M.E.; Hess, G.D.; Lee, S.; Tory, K.; Azzi, M.; Carras, J.; Lilley, W.; Manins, P.C.; Nelson, P.; Ng, L.; et al. The Australian Air Quality Forecasting System. Part I: Project Description and Early Outcomes. *J. Appl. Meteorol.* **2004**, *43*, 649–662. [CrossRef]

32. Commonwealth Scientific and Industrial Research Organisation. National AQFx Prototype System. 2022. Available online: https://research.csiro.au/aqfx/ (accessed on 12 November 2022).

33. Horsley, J.A.; Broome, R.A.; Johnston, F.H.; Cope, M.; Morgan, G.G. Health burden associated with fire smoke in Sydney, 2001–2013. *Med. J. Aust.* **2018**, *208*, 309–310. [CrossRef]

34. Broome, R.A.; Fann, N.; Cristina, T.J.N.; Fulcher, C.; Duc, H.; Morgan, G.G. The health benefits of reducing air pollution in Sydney, Australia. *Environ. Res.* **2015**, *143*, 19–25. [CrossRef]

35. Australian Bureau of Statistics. Location: Australian Census. 2021. Available online: https://www.abs.gov.au/statistics/people/people-and-communities/location-census/latest-release (accessed on 30 August 2022).

36. Australian Bureau of Statistics. 2074.0—Census of Population and Housing: Mesh Block Counts, Australia, 2016. 2017. Available online: https://www.abs.gov.au/ausstats/abs@.nsf/mf/2074.0 (accessed on 12 March 2022).

37. Department of Water and Environmental Regulation. Perth Air Emissions Inventory 2011–12. Available online: https://www.der.wa.gov.au/our-work/programs/460-perth-air-emissions-study-2011-2012 (accessed on 2 February 2020).

38. Jiao, W.; Hagler, G.; Williams, R.; Sharpe, R.; Brown, R.; Garver, D.; Judge, R.; Caudill, M.; Rickard, J.; Davis, M.; et al. Community Air Sensor Network (CAIRSENSE) project: Evaluation of low-cost sensor performance in a suburban environment in the southeastern United States. *Atmos. Meas. Technol.* **2016**, *9*, 5281–5292. [CrossRef]

39. Australian Bureau of Statistics. *Australian Statistical Geography Standard (ASGS): Volume 1—Main Structure and Greater Capital City Statistical Areas, July 2016*; Australian Bureau of Statistics: Canberra, ACT, Australia, 2016.

40. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press Ltd.: Cambridge, MA, USA, 2006.

41. Ylisirniö, A.; Buchholz, A.; Mohr, C.; Li, Z.; Barreira, L.; Lambe, A.; Faiola, C.; Kari, E.; Yli-Juuti, T.; Nizkorodov, S.A.; et al. Composition and volatility of secondary organic aerosol (SOA) formed from oxidation of real tree emissions compared to simplified volatile organic compound (VOC) systems. *Atmos. Chem. Phys.* **2020**, *20*, 5629–5644. [CrossRef]

42. Jaffe, L.S. Ambient Carbon Monoxide And Its Fate in the Atmosphere. *J. Air Pollut. Control Assoc.* **1968**, *18*, 534–540. [CrossRef] [PubMed]

43. Morris, M.D.; Mitchell, T.J. Exploratory designs for computational experiments. *J. Stat. Plan. Inference* **1995**, *43*, 381–402. [CrossRef]

44. Loeppky, J.L.; Sacks, J.; Welch, W.J. Choosing the Sample Size of a Computer Experiment: A Practical Guide. *Technometrics* **2009**, *51*, 366–376. [CrossRef]

45. Bastos, L.S.; O'Hagan, A. Diagnostics for Gaussian Process Emulators. *Technometrics* **2009**, *51*, 425–438. [CrossRef]

46. Salteli, A.; Chan, K.; Scott, E. *Sensitivity Analysis*; John Wiley: New York, NY, USA, 2000.

47. Oakley, J.E.; O'Hagan, A. Probabilistic sensitivity analysis of complex models: A Bayesian approach. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2004**, *66*, 751–769. [CrossRef]

48. R Core Team. *R: A Language and Environment for Statistical Computing*; R Core Team: Vienna, Austria, 2021.

49. Roustant, O.; Ginsbourger, D.; Deville, Y. DiceKriging, DiceOptim: Two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization. *J. Stat. Softw.* **2012**, *51*, 1–55. [CrossRef]

50. Saltelli, A.; Tarantola, S.; Chan, K.P.S. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics* **1999**, *41*, 39–56. [CrossRef]

51. Bertrand Iooss, A.; Da Veiga, S.; Janon, A.; Pujol, G. Package "Sensitivity" Title Global Sensitivity Analysis of Model Outputs. 2021. Available online: https://cran.r-project.org/web/packages/sensitivity/sensitivity.pdf (accessed on 13 November 2022).

52. Graham, P.; Havas, L. *Electric Vehicle Projections 2021*; Commonwealth Scientific and Industrial Research Organisation: Melbourne, VIC, Australia, 2021. [CrossRef]

53. Paton-Walsh, C.; Rayner, P.; Simmons, J.; Fiddes, S.L.; Schofield, R.; Bridgman, H.; Beaupark, S.; Broome, R.; Chambers, S.D.; Chang, L.T.-C.; et al. A Clean Air Plan for Sydney: An Overview of the Special Issue on Air Quality in New South Wales. *Atmosphere* **2019**, *10*, 774. [CrossRef]

54. Barnett, A.G.; Williams, G.M.; Schwartz, J.; Best, T.L.; Neller, A.H.; Petroeschevsky, A.L.; Simpson, R.W. The effects of air pollution on hospitalizations for cardiovascular disease in elderly people in Australian and New Zealand cities. *Environ. Health Perspect.* **2006**, *114*, 1018–1023. [CrossRef]

55. Hanigan, I.C.; Rolfe, M.I.; Knibbs, L.D.; Salimi, F.; Cowie, C.T.; Heyworth, J.; Marks, G.B.; Guo, Y.; Cope, M.; Bauman, A. All-cause mortality and long-term exposure to low level air pollution in the '45 and up study' cohort, Sydney, Australia, 2006–2015. *Environ. Int.* **2019**, *126*, 762–770. [CrossRef]

56. Dirgawati, M.; Hinwood, A.; Nedkoff, L.; Hankey, G.; Yeap, B.B.; Flicker, L.; Nieuwenhuijsen, M.; Brunekreef, B.; Heyworth, J. Long-term Exposure to Low Air Pollutant Concentrations and the Relationship with All-Cause Mortality and Stroke in Older. *Men's Epidemiol.* **2019**, *30*, S82–S89. [CrossRef] [PubMed]

57. Zosky, G.R.; Hoorn, S.V.; Abramson, M.J.; Dwyer, S.; Green, D.; Heyworth, J.; Jalaludin, B.B.; McCrindle-Fuchs, J.; Tham, R.; Marks, G.B. Principles for setting air quality guidelines to protect human health in Australia. *Med. J. Aust.* **2021**, *214*, 254–256.e1. [CrossRef] [PubMed]

58. Vedrenne, M.; Borge, R.; Lumbreras, J.; Rodríguez, M.E. Advancements in the design and validation of an air pollution integrated assessment model for Spain. Environ. *Model. Softw.* **2014**, *57*, 177–191. [CrossRef]

59. Winkler, S.L.; Anderson, J.E.; Garza, L.; Ruona, W.C.; Vogt, R.; Wallington, T.J. Vehicle criteria pollutant (PM, NO$_x$, CO, HCs) emissions: How low should we go? *NPJ Clim. Atmos. Sci.* **2018**, *1*, 26. [CrossRef]

60. Keywood, M.; Guyes, H.; Selleck, P.; Gillett, R. Quantification of secondary organic aerosol in an Australian urban location. *Environ. Chem.* **2011**, *8*, 115–126. [CrossRef]

61. Lee, L.A.; Carslaw, K.S.; Pringle, K.J.; Mann, G.W. Mapping the uncertainty in global CCN using emulation. *Atmos. Meas. Technol.* **2012**, *12*, 9739–9751. [CrossRef]

62. Sokan-Adeaga, M.A.; Ree, A.G.; Ayodeji, S.A.M.; Deborah, S.A.E.; Ejike, O.M. Secondary inorganic aerosols: Impacts on the global climate system and human health. *Biodivers. Int. J.* **2019**, *3*, 249–259. [CrossRef]

63. Liu, Z.; Zhou, M.; Chen, Y.; Chen, D.; Pan, Y.; Song, T.; Ji, D.; Chen, Q.; Zhang, L. The nonlinear response of fine particulate matter pollution to ammonia emission reductions in North China. *Environ. Res. Lett.* **2021**, *16*, 034014. [CrossRef]

64. Cao, H.; Henze, D.K.; Cady-Pereira, K.; McDonald, B.C.; Harkins, C.; Sun, K.; Bowman, K.W.; Fu, T.M.; Nawaz, M.O. COVID-19 Lockdowns Afford the First Satellite-Based Confirmation That Vehicles Are an Under-recognized Source of Urban NH3Pollution in Los Angeles. *Environ. Sci. Technol. Lett.* **2022**, *9*, 3–9. [CrossRef]

65. Ghavam, S.; Vahdati, M.; Wilson, I.A.G.; Styring, P. Sustainable Ammonia Production Processes. *Front. Energy Res.* **2021**, *9*, 34. [CrossRef]

66. Vanherle, K.; Lopez-Aparicio, S.; Grythe, H.; Lükewille, A.; Unterstaller, A.; Mayeres, I. Transport Non-Exhaust PM-Emissions. An Overview of Emission Estimates, Relevance, Trends and Policies—Eionet Portal. In *ETC/ATNI Report 5/2020*; European Environment Agency: Copenhagen, Denmark, 2021.

67. Noël, C.; Vanroelen, C.; Gadeyne, S. Qualitative research about public health risk perceptions on ambient air pollution. A review study. *SSM Popul. Health* **2021**, *15*, 100879. [CrossRef] [PubMed]

68. Larkin, A.; Hystad, P. Towards Personal Exposures: How Technology Is Changing Air Pollution and Health Research. *Curr. Environ. Health Rep.* **2017**, *4*, 463. [CrossRef] [PubMed]

69. NSW EPA. Air Emissions Inventory for the Greater Metropolitan Region in NSW. Available online: https://www.epa.nsw.gov.au/your-environment/air/air-emissions-inventory (accessed on 20 March 2021).

70. Goldsworthy, L.; Goldsworthy, B. Modelling of ship engine exhaust emissions in ports and extensive coastal waters based on terrestrial AIS data—An Australian case study. *Environ Model Software.* **2015**, *63*, 45–60. [CrossRef]

71. Department of the Environment and Energy. National Pollutant Inventory 2019. Available online: http://npi.gov.au/ (accessed on 2 February 2020).

72. Australian Standard AS/NZS. *AS/NZS 3580.1.1:2007 Methods for Sampling and Analysis of Ambient Air Part 1.1: Guide to Siting Air Monitoring Equipment*; Standards Australia: Sydney, NSW, Australia; Standards New Zealand: Wellington, New Zealand, 2007.

73. Simon, H.; Baker, K.R.; Phillips, S. Compilation and interpretation of photochemical model performance statistics published between 2006 and 2012. *Atmos Environ.* **2012**, *61*, 124–139. [CrossRef]

74. Thunis, P.; Georgieva, E.; Galmarini, S. *A Procedure for Air Quality Models Benchmarking*; Joint Research Centre: Ispra, Italy, 2011.

75. Boylan, J.W.; Russell, A.G. PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models. *Atmos Environ.* **2006**, *40*, 4946–4959. [CrossRef]