Feasibility study for a high-resolution Southeast Asian air pollution forecasting system

Final report

Distribution:	**CONFIDENTIAL**
Project Title:	Feasibility study for a high-resolution Southeast Asian air pollution forecasting system
Report prepared for:	UK Foreign, Commonwealth & Development Office (FCDO)
Report Status:	Draft
Report Reference:	CERC/FM1444/D6.4
CERC Job Number:	FM1444
Issue Version (Date):	**1 (23/02/2024)** **2 (05/03/2024)** **3 (25/03/2024)** **4 (18/06/2024)**
Authors:	Jenny Stocker, Christina Hood, Kate Johnson, Victoria Hamilton, Amy Stidworthy (CERC) Mohd Talib Latif, Liew Ju Neng, Maggie Ooi Chel Gee, Nur Nazmi Liyana Binti Mohd Napi (Universiti Kebangsaan Malaysia) Bang Quoc Ho, Do Thi Thuy Linh, Nguyen Thoai Tam, Tran Thi Hong Hien (Institute for Environment and Resources, Vietnam National University, Ho Chi Minh City)
Reviewers:	Jeremie Credo (FCDO)

Cambridge Environmental Research Consultants (CERC)
Web: cerc.co.uk. Email: enquiries@cerc.co.uk. Tel: +44 (0)1223 357773.
Post: CERC, 3 Kings Parade, Cambridge, CB2 1SJ, UK.

Contents

Acknowledgements

This study was supported by the South Asia Research Hub, Foreign, Commonwealth & Development Office (FCDO), Government of the United Kingdom. The views expressed in this report, however, do not necessarily reflect the official policies of the Government of the United Kingdom.

The project team would also like to acknowledge the following organisations within Malaysia and Vietnam:

- Malaysia:
  - Malaysian Department of Environment
  - Department of Health and Environment, Kuala Lumpur City Hall
  - Department of Civil Engineering and Urban Transport, Kuala Lumpur City Hall
  - Ministry of Transport Malaysia
- Vietnam:
  - Center for Environmental Monitoring, Division of Environmental Protection, Ho Chi Minh City Department of Natural Resources and Environment

Acknowledgements of additional data providers are given in Section 11.

1 Executive summary

Background

1.1 Air pollution is one of the greatest environmental risks to health. World Health Organisation guidance levels of air pollutant concentrations are exceeded in many Association of Southeast Asian Nations (ASEAN) member countries. There is an increased risk of certain health conditions where both temperatures and ambient air pollution levels are high, and high temperatures are exacerbated by climate change.

1.2 Providing air pollution forecasts in combination with actionable health advice contributes to improved health. Air pollution alerts can also discourage the public from generating additional air pollution from domestic and transport sources.

1.3 Project lead Cambridge Environmental Research Consultants have developed software systems that produce air pollution forecasts and are experienced in providing air quality forecasting services. Academic project partners from the Institute for Environment and Resources, Vietnam National University Ho Chi Minh City and the Department of Earth Sciences and Environment, Faculty of Science and Technology, Universiti Kebangsaan Malaysia have expertise in air pollutant measurements and emissions, as well as knowledge of the key factors that influence air quality in their countries.

1.4 This project addresses Sustainable Development Goal (SDG) 3 'Ensure healthy lives and promote well-being for all at all ages.'

Research aims

The project aimed to:

1.5 Demonstrate that pan-ASEAN street-scale air quality forecasts are achievable.

1.6 Stimulate interest in high-resolution air quality forecasting within the ASEAN stakeholder community.

1.7 Demonstrate that open data allows transferability and scalability.

1.8 Strengthen UK-ASEAN air quality modelling relationships.

1.9 Consider a route-to-market strategy for the deployment of street-scale air quality forecasting systems in ASEAN.

Methodology

Technical activities included:

1.10 Selection of two 200 km by 200 km domains: one domain covered Ho Chi Minh City in Vietnam, extending into the Svay Rieng Province in Cambodia; while the other included the Kuala Lumpur mega-urban area.

1.11 Collation and processing of model input data using a range of tools and custom scripts. Datasets included: mesoscale model meteorology, regional model pollutant concentration, and regional emissions. High-resolution air pollutant emissions inventories were generated using top-down and bottom-up approaches, with a focus on road traffic sources. For the Vietnamese domain, a 2 km resolution emissions inventory, developed by the Vietnamese partners prior to the current project, was integrated into the system inputs.

1.12 Modelling system configuration and execution to generate predictions of air pollutant concentration at measurement locations. Statistics generated during the model evaluation process included Air Pollutant Index values (APIs), which is the indexing system commonly used to report air quality levels to the public.

1.13 Generation of regional-to-street-scale air pollution maps.

1.14 Proposal for a system design that describes how this modelling could be taken forward as operational air quality forecasting systems throughout ASEAN.

Stakeholder engagement activities included:

1.15 Initial collation and continued updating of the identified stakeholders list.

1.16 Development of project dissemination materials, using the public-facing project name 'Demonstration of High-resolution Air Quality in South East Asia' (DHAQ-SEA).

1.17 A series of project and workshop notification emails to the stakeholder list.

1.18 Planning and execution of an online stakeholder engagement workshop.

1.19 Workshop follow up.

Results

Technical results:

1.20 Periods of poor air quality during 2022 were identified for both study domains.

1.21 The model evaluation process identified improvements to the model configuration including calibration of regional model concentration data inputs for one of the two domains.

1.22 High-resolution air pollution maps were generated at varying spatial scales, as both air pollutant concentrations and APIs. Animations[1] were generated demonstrating daily air pollutant variations.

Stakeholder engagement outcomes:

1.23 The final stakeholder list included representatives from over 60 different organisations. These were primarily from ASEAN countries, but the list also included non-ASEAN country representatives.

1.24 The workshop agenda covered not only project activities and results, but also motivation for the work in terms of impacts of air pollution on health, and an overview of the primary contributing factors to poor air quality.

1.25 Over 75 people attended the two-hour online stakeholder workshop, which took place on Thursday 16th May 2024.

1.26 To maximise project exposure, the stakeholder workshop presentation was circulated to all on the stakeholder list, including those who did not register or attend the workshop.

Impact & significance of project outcomes to targeted beneficiaries in the study sites

1.27 DHAQ-SEA addressed health priorities by demonstrating a system that can provide publicly accessible forecast air quality data. This information enables those vulnerable to air pollution to protect themselves.

[1] https://cerc.co.uk/environmental-research/assets/data/CERC_DHAQ-SEA_PM2.5_API_video.gif
https://cerc.co.uk/environmental-research/assets/data/CERC_DHAQ-SEA_PM2.5_Conc_video.gif

1.28 The project has developed a system that incorporates both regional and local components of air pollution, while the system domains provide examples of covering a single country and multiple countries. Air pollution is not confined by national borders; thus, regional cooperation is vital. By including neighbouring countries in the study domains, the DHAQ-SEA project acknowledges and addresses the transboundary nature of air pollution.

1.29 The stakeholder workshop emphasised the collaborative nature of addressing air pollution issues. The workshop brought together representatives from various ASEAN countries, facilitating an opportunity for them to share knowledge, experiences, and best practices.

1.30 The DHAQ-SEA approach aligns with the broader goals of ASEAN to promote sustainable development and environmental protection.

1.31 The collation of emission inventory data is the main challenge for air quality modelling in ASEAN. The project has highlighted that, although suitable datasets may exist, it can be challenging to obtain permission for access, and if access is granted, the time-scales involved with receiving the data can be prohibitive. The Vietnamese project partners have experience in local-scale emissions inventory development. A key recommendation of the project is to update the Regional Emission inventory in ASia (REAS) database, which was last released in 2020 with emissions data for 2015.

Scalability of the results at the ASEAN level

1.32 The air quality modelling system is suitable for applications over large domains throughout ASEAN. The system requires domain-specific inputs.

1.33 The project team used open data sources to generate system inputs where possible. These included open emissions, concentration, meteorology and road network data. However, the project results highlighted that some global datasets are inaccurate for ASEAN.

1.34 Air quality measurements from local organisations are required for system evaluation in terms of modelled pollutant concentrations and, in some cases, calibration of system inputs. Air quality measurements are usually available, but the data is not always reliable.

1.35 The project has highlighted that in order to resolve pollutant concentration variations of $PM_{2.5}$, which is the pollutant of highest concern in terms of health impacts in the region, some non-traffic sources must be resolved to local scales using suitable emissions proxies. Such local information relating to region-specific air pollution sources must be provided by local air quality experts.

2 Introduction

This report describes activities associated with the project to undertake a 'Feasibility study for a high-resolution Southeast Asian air pollution forecasting system'. The report has been prepared to fulfil 'Output 4' for the Research and Innovation for Development in ASEAN[2] (RIDA) project, referred to as deliverable D6.4 in the project plan.

During the project, a public-facing project name was adopted: 'Demonstration of High-resolution Air Quality forecasting in South East Asia', with the associated acronym DHAQ-SEA and email dhaqsea@ukm.edu.my.

This project involved both technical activities and stakeholder engagement.

A project overview is given in Section 3, providing project context (Section 3.1) and aims, objectives, and significance (Section 3.2). In addition, the robustness of project outcomes is discussed in Section 3.3, separately for technical items and stakeholder engagement. Section 4 begins with a discussion of the approaches used to quantify air pollutant concentrations (Section 4.1); the project method is then summarised (Section 4.2) to provide context for the technical results (Section 4.3). Section 5 summarises project stakeholder engagement; the activity timeline is presented (Section 5.1), followed by a description of stakeholder engagement outcomes (Section 5.2).

Following these project details, a summary of lessons learned and project impacts are presented (Section 6). The specific requirements of this final report are then given in relation to: a discussion of study site beneficiaries (Section 7.1); implications for the wider Southeast Asian region (Section 7.2); and system scalability (Section 7.3). A discussion of project limitations is given in Section 8, and potential next steps for the project are suggested in Section 9; a list of references is provided (Section 10).

Supplementary information is provided in the Appendices: Appendix A provides additional technical detail; Appendix B presents additional pollutant concentration evaluation and maps; and Appendix C gives additional stakeholder engagement information.

[2] Association of Southeast Asian Nations

3 Project overview

3.1 Context

According to the World Health Organisation (WHO), air pollution is one of the greatest environmental risks to health. By reducing air pollution levels, countries can reduce the burden of disease from stroke, heart disease, lung cancer, and respiratory conditions, including asthma. WHO estimates that air pollution in both cities and rural areas caused 4.2 million premature deaths worldwide in 2019. WHO guideline pollutant concentration levels [1] are exceeded in many member countries of the Association of Southeast Asian Nations (ASEAN). Annual average concentrations of toxic particulates[3] (PM_{10}, $PM_{2.5}$) and nitrogen dioxide (NO_2) were around 50, 25 [2] and 16 µg/m³ [3] for 2018/2019 in Ho Chi Minh City, Vietnam (population over 9 million), compared to WHO guidelines of 15, 5 and 10 µg/m³, respectively; corresponding concentrations for urban areas in Malaysia (2020) were 22, 13 and 16 µg/m³ [4]. There is an increased risk of certain health conditions where both temperatures and ambient air pollution levels are high [5]. ASEAN countries experience high temperatures, which are exacerbated by climate change, for example, recent record temperatures occurred in April 2023.

Providing air pollution forecasts in combination with actionable health advice contributes to **improved health**, thus the project addresses Sustainable Development Goal (SDG) 3 'Ensure healthy lives and promote well-being for all at all ages.' People who receive air pollution alerts can prepare their day ahead accordingly. Vulnerable people, in particular, can ensure they have any necessary medication at hand, avoid vigorous exercise, and, in extreme cases, stay indoors. An additional benefit of alerting the public to upcoming air pollution episodes is **pollution mitigation**. For example, the public can be discouraged from generating additional air pollution in their homes (e.g. wood burning stoves) and in their local area (e.g. avoiding the use of vehicles with combustion engines) [6].

Southeast Asian academics from the Institute for Environment and Resources, Vietnam National University Ho Chi Minh City (IER/VNU-HCM) and the Department of Earth Sciences and Environment, Faculty of Science and Technology, Universiti Kebangsaan Malaysia (UKM) have collaborated in an equitable partnership with Cambridge Environmental Research Consultants (CERC) to demonstrate the feasibility of regional-to-local scale air quality forecasting systems for domains encompassing large parts of Malaysia and Vietnam, and a small region of Cambodia. The system could be implemented in wider regions across ASEAN, as it relies primarily on open data sources.

Project aims and corresponding objectives are summarised in Table 3.1. All project objectives have been achieved except for two related to the route-to-market strategy, where only partial progress has been made. Details of the project team's plan for continued stakeholder engagement and commercialisation of the system are given in Section 9.1.

[3] For regulatory air quality purposes, toxic particulates are categorised by their diameter. Particles with diameters less than 10 and 2.5 micrometres (microns) are referred to as PM_{10} and $PM_{2.5}$ respectively.

Table 3.1 – Aims and objectives of the 'Feasibility study for a high-resolution Southeast Asian air pollution forecasting system' project; ✓ indicates that the objective has been completed; (✓) indicates that partial progress was made on the objective during the project.

Aim		Objective
Description	**Type**	
To stimulate interest in the project within the ASEAN stakeholder community	Project dissemination / Networking	• Collate a comprehensive list of ASEAN stakeholders, working with SE Asian project partners, and FCDO representatives. ✓ • Provide summary project information to ASEAN stakeholders. ✓ • At the end of the project, invite ASEAN stakeholders to an online workshop. The workshop will demonstrate the feasibility of the system for applications across the ASEAN region. ✓
To consider a route-to-market strategy	Commercial/ Networking	• Stimulate interest in the project from the ASEAN stakeholder community (refer to activities listed above). ✓ • Identify stakeholders interested in operational forecasting systems for their region. (✓) • Investigate possible funding avenues. (✓)
To strengthen UK-ASEAN air quality modelling relationships	Networking	• CERC are working with UKM, with whom they have collaborated on previous projects. ✓ • CERC are working with IER/VNU-HCM, who they have not worked with before. ✓
To demonstrate that pan-ASEAN street-scale air quality forecasts are achievable	Project scope	• Develop a system that accounts for the contribution of both local and regional emissions to local air quality. ✓ • Select study domains that include neighbouring countries. ✓
To develop high-resolution air pollutant emissions inventories for the study regions	Technical	• Assess the most suitable global emissions inventory datasets to use for downscaling, considering spatial resolution and year. ✓ • Increase resolution of road traffic open emissions datasets by applying top-down methods using road type & other parameters. ✓
To identify poor air quality periods in the selected study regions	Technical	• Collate air quality measurements from the study regions. ✓ • Analyse air quality data to identify air pollution episodes suitable for demonstration of the air quality modelling system. ✓
To generate accurate regional-to-street-scale air pollution maps for the study domains	Technical	• Review data processing steps to minimise errors. ✓ • Use the developed emissions inventory alongside other datasets as input to the coupled air pollutant modelling system; run system. ✓ • Compare model output to available measurement data; revise model inputs as required. ✓ • Present air pollution maps in a format that is accessible to the project stakeholders. ✓
To demonstrate that open data allows transferability and scalability	Project scope	Use open data sources where possible, for example: • Modelled datasets of global air pollutant concentrations and meteorology (from the Copernicus Atmosphere Monitoring Service, CAMS: reanalysis [7] and forecast [8]). ✓ • Global air pollutant emissions inventories (CAMS [9], EDGAR [10], REAS [11]). ✓ • Major road network data from OpenStreetMap [12], categorised by road type. ✓
To demonstrate that an operational air quality forecasting system can be developed	Technical	• Show examples of CERC's operational forecasts in other regions. ✓ • Document how an operational air quality forecasting service based on the demonstration modelling can be developed for any or all SE Asian countries (technical information in report, and summary information for project workshop). ✓

3.2 Significance

This project addresses health priorities by demonstrating a system that can provide publicly accessible forecast air quality data.

ASEAN governments have highlighted air quality as a health priority. Malaysia's Health White Paper [13] highlights the need for reform with four key pillars, of which pillar 2 focuses on 'preventative care and empowering people to maintain health' and 'data-driven public health practices'. Similarly, the Prime Minister of Vietnam approved a national plan for environmental quality management in 2021. This plan emphasises the need to improve air quality, key aspects of which include improved air monitoring and assessment of the impact of air quality on the health of citizens. Also in 2021, the Cambodian Ministry of Environment released the 'Clean Air Plan of Cambodia', which outlines monitoring and assessment of air quality, and highlights the impact of air pollution on health, crops and the climate.

Air quality forecasting systems allow those vulnerable to air pollution to protect themselves. In addition, archives of the pollution maps generated by forecasting systems can provide the necessary data to support governmental bodies in assessing current, and informing future, air quality improvement strategies.

3.3 Robustness of project outcomes

The project has produced robust technical and stakeholder outcomes.

Technical outcomes

Technical outputs from the proposed system have been demonstrated in the form of pollution maps for two distinct domains, both approximately 200 km × 200 km in extent. One domain covers Ho Chi Minh City in Vietnam, extending into the Svay Rieng Province in Cambodia; while the other includes the Kuala Lumpur mega-urban area in Malaysia, as shown in Figure 3.1. These pollution maps resolve the high pollution concentration gradients that occur in the vicinity of road sources and can be generated for a range of toxic pollutants; example maps are presented for key pollutants $PM_{2.5}$, PM_{10}, NO_2 and O_3. The robustness of these pollution maps is demonstrated through the 'model evaluation' procedure (described in Section 4.3), where modelled pollutant concentrations are compared with air quality measurements.

Extensive data processing was required to generate the inputs for the air quality modelling system. Processing scripts were written and tested to ensure that the modelling system data inputs were robust. The 'Multi-model Air Quality System' (MAQS, [14]) used to generate the air pollution maps is commercially available software developed by CERC according to strict coding standards. This system has been applied in other locations for a range of applications including country-wide baseline modelling, pollution mitigation scenarios and air quality forecasting [15].

Although out of scope for the current feasibility study, the project team have experience of, and expertise in, developing operational air quality forecasting systems [16]. Thus, given interest and funding from appropriate organisations, the modelling for these regions could be developed to form the basis for an air quality forecast and/or similar modelling approaches could be used for other ASEAN regions.

A summary of the technical results from the project is given in Section 4.

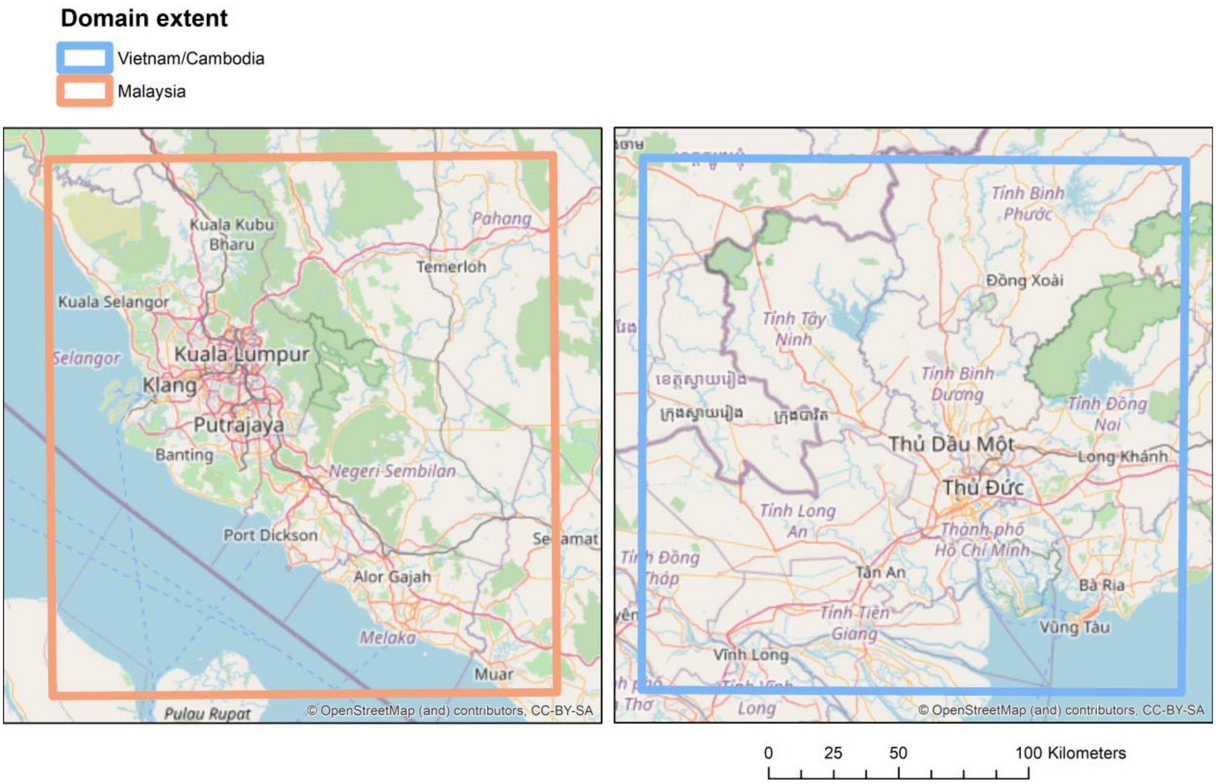


Figure 3.1 – Domain extents for air pollution maps, used to demonstrate the feasibility of developing a regional-to-local-scale air quality forecasting system

Stakeholder outcomes

The project team have engaged with a wide range of stakeholders, including local, regional and national government bodies, non-governmental organisations, and academic institutions. Furthermore, FCDO facilitated additional stakeholder engagement opportunities via the ASEAN Committee on Science, Technology, and Innovation (COSTI), and British High Commissions and Embassies.

An online workshop for key ASEAN project stakeholders was held on 16th May 2024. Workshop details are given in Section 5. The presentations included an overview of the project's technical methodology and outcomes, and an outline of the steps required to develop operational high-resolution systems for any ASEAN region. The post-project engagement resulting from the stakeholder workshop is vital to the commercial success of the project; further details of the commercial plan developed by the project team are given in Section 9.1.

4 Technical method and results

The project technical activities fulfil a number of the project aims including:

- ‘To identify poor air quality periods in the selected study regions.’
- ‘To develop high-resolution air pollutant emissions inventories for the study regions.’
- ‘To generate accurate regional-to-street-scale air pollution maps for the study domains.’

Identifying ‘poor air quality’ periods necessitates information on the relationships between air pollutant concentrations and population health; this is a broad and challenging topic. However, many countries have defined air pollution indexing systems which relate short-term (e.g. hourly, daily) pollution levels to health impacts. Details of these indexing systems for Vietnam and Malaysia are given in Section 4.1. This is followed by an overview of the technical method (Section 4.2) and results (Section 4.3). Further information is provided in Appendix A (method) and Appendix B (results).

4.1 Air pollution quantification

Within a forecasting system, modelled air pollutant concentration levels are often compared to air pollution indices; the daily Air Quality Index (AQI) system for Vietnam and Air Pollutant Index (API) system for Malaysia are summarised in Tables 4.1 and 4.2, respectively. In general, index values are calculated separately for each pollutant and averaging time, then the maximum index value for each day is reported as an overall indicator. In Vietnam, the AQI for O_3 is usually calculated from the daily maximum 8-hour average, however if this has a value higher than 400 $\mu g/m^3$, the AQI is calculated with the daily maximum 1-hour average. In Malaysia, where the daily maximum 1-hr SO_2 API values are greater than 200, indices are calculated using the 24-hour average SO_2 concentrations, and where the daily maximum 8-hour O_3 API values are greater than 300, indices are calculated using the daily maximum 1-hr O_3 concentrations. In some areas in Malaysia, both the 8-hour and 1-hour O_3 API values are calculated, and the maximum is taken.

Operational air quality forecasting systems send out air pollution alerts when the daily air quality is expected to lead to acute adverse health effects. The Vietnamese and Malaysian alert levels (second column of the tables below) have a lower range of pollutant concentrations that are ‘unhealthy for sensitive groups’, and higher thresholds for the general public. Health advice is usually sent out with an air pollution alert, allowing individuals to take action to reduce their exposure.

Table 4.1 – Daily Vietnam Air Quality Index (AQI) values and their corresponding air pollution levels; pollutant concentration ranges for each AQI level are shown; concentration units[4] are in µg/m³.

AQI values	Air pollution levels	O_3	O_3	CO	SO_2	NO_2	PM_{10}	$PM_{2.5}$
	Averaging Time	***1-hour***	***8-hour***	***1-hour***	***1-hour***	***1-hour***	***24-hour***	***24-hour***
<50	Good	0-160	0-100	0-10,000	0-125	0-100	0-50	0-25
51 - 100	Moderate	161-200	101-120	10,001-30,000	126-350	101-200	51-150	26-50
101 - 150	Unhealthy for sensitive groups	201-300	121-170	30,001-45,000	351-550	201-700	151-250	51-80
151 - 200	Unhealthy	301-400	171-210	45,001-60,000	551-800	701-1,200	251-350	81-150
201 - 300	Very unhealthy	401-800	211-400	60,001-90,000	801-1,600	1,201-2,350	351-420	151-250
>300	Hazardous	>800	-	>90,000	>1,600	>2,350	>420	>250

Table 4.2 – Daily Malaysia Air Pollutant Index (API) values and their corresponding air pollution levels; pollutant concentration ranges for each API value are shown; PM_{10} and $PM_{2.5}$ concentrations are in units of µg/m^3 and the gaseous pollutants are in ppm.

AQI values	Air pollution levels	O_3	O_3	CO	SO_2	NO_2	PM_{10}	$PM_{2.5}$
	Averaging Time	***1-hour***	***8-hour***	***1-hour***	***1-hour***	***1-hour***	***24-hour***	***24-hour***
<50	Good	-	0.000-0.059	0.0-4.4	0.000-0.035	0.000-0.053	0-54	0.0-12.0
51 - 100	Moderate	-	0.060-0.075	4.5-9.4	0.036-0.075	0.054-0.100	55-154	12.1-35.4
101 - 150	Unhealthy for sensitive groups	0.125-0.164	0.076-0.095	9.5-12.4	0.076-0.185	0.101-0.360	155-254	35.5-55.4
151 - 200	Unhealthy	0.165-0.204	0.096-0.115	12.5-15.4	0.186-0.304	0.361-0.649	255-354	55.5-150.4
201 - 300	Very unhealthy	0.205-0.404	0.116-0.374	15.5-30.4	-	0.650-1.249	355-424	150.5-250.4
>300	Hazardous	>0.404	-	>30.4	-	>1.249	>424	> 250.4

4.2 Method

The primary technical outputs from the project are example air pollution maps for the 200 km by 200 km Vietnamese/Cambodian and Malaysian domains. Since the recent year (2022) air pollution episodes identified by the Malaysian and Vietnamese project partners correspond to high $PM_{2.5}$ concentrations (more details are provided in Section A.1), this pollutant is the focus

[4] Note that the µg/m^3 units are defined in relation to 1 m^3 air at standard temperature (25 °C) and pressure (760 mmHg)

of the project. However, it is important to note that the particulate (PM_{10} and $PM_{2.5}$) concentrations are dominated by long-range pollutant transport and chemical processes rather than local emissions. Other health-related pollutants (e.g. NO_2, CO) are more strongly influenced by local pollution sources, including traffic.

Pollution maps corresponding to example days within these high pollution episodes have been generated, alongside maps of days with lower measured pollution for comparison. These maps:

- have been compared to available monitoring data to check accuracy; and
- show the variation of pollutant concentrations at high spatial resolution.

The technical method used to generate the regional-to-local air quality maps is outlined in Figure 4.1. Details of the steps are provided in Appendix A.

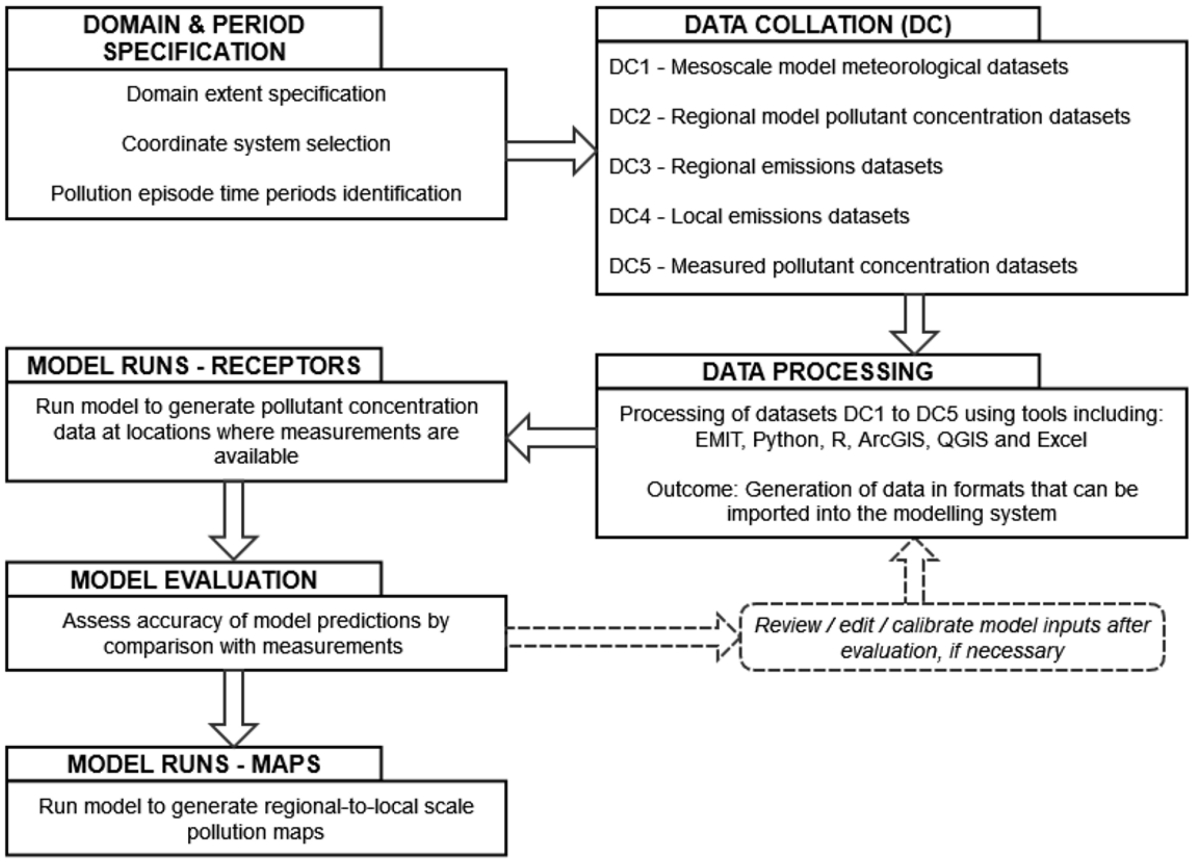


Figure 4.1 – Flow chart outlining the technical method for generating the demonstration regional-to-local-scale air pollution maps

4.3 Results

The modelling system has been run in two modes: firstly, with output at locations where measurements are available for a full month, used for model evaluation, and secondly, with gridded output throughout the domain for single-episode and non-episode days, used for creating concentration and pollution index map images. Run times for each domain are given in Table 4.3. Although both domains have the same spatial extent, the Malaysia domain has longer run times for both receptor and gridded output runs due to a larger number of explicitly modelled roads, with a particularly high spatial density of road sources within a few cells covering Kuala Lumpur. In contrast, the smaller road network in the Vietnam/Cambodia domain is more evenly spread across the domain, which allows an efficient distribution of computations. Run times could potentially be reduced in an operational system by further optimising the use of computational resources. Minimising run times is key to allowing operational forecasts to be issued in a timely manner. In addition to the coupled modelling system run time given in the table, additional time is required for pre-processing input data and post-processing outputs to generate map images.

Table 4.3 – Run time information (coupled modelling system only) for both modelling domains and run types. The receptor runs used for evaluation have a modelling period of one month and the gridded runs used for maps have a modelling period of one day.

Domain	Run time for run type, computer system	
	Receptors (evaluation), local computer	**Gridded (maps), virtual cloud computer**
Vietnam/Cambodia	16 minutes	3 hours 20 minutes
Malaysia	25 minutes	5 hours 50 minutes

Another consideration for operational forecasting systems is the size of output files, which affects requirements for archiving resources. For the Vietnam/Cambodia domain, single day modelling system output files are 4.6 GB in size, containing over 6 million output points. These allow map images to be generated at 20 m resolution. For the Malaysia domain, single day files contain around 4.5 million output points, leading to files that are 3.4 GB in size.

Evaluation

Concentrations generated by the modelling system at measurement site locations have been compared to the corresponding measurements. Information about the measurement sites is given in Appendix A. Some example evaluation graphs are shown in this section, while further information is given in Appendix B. Evaluation has been carried out with three levels of data processing: the raw measured and modelled hourly concentrations; the daily concentration metrics used in the air quality/pollution index calculations, e.g. maximum daily 1-hour average NO_2 concentration or 24-average $PM_{2.5}$; and the forecast index values associated with air quality categories, such as ‘good’ or ‘unhealthy’.

The simplest comparison between modelled and measured concentrations uses a scatter plot to compare period mean values. These are shown in Figure 4.2 for gaseous pollutants NO_2, NO_X and O_3, while Figure 4.3 shows equivalent plots for particulates (PM_{10} and $PM_{2.5}$), all for January 2022 in the Vietnamese/Cambodian domain. Two model configurations are shown in adjacent panels on these plots, one with both gridded and explicit emissions modelled for the full domain and one with finer-resolution gridded emissions (shared by project partners IER/VNU-HCM) for a subdomain covering Ho Chi Minh City. The modelled concentrations are of the same order of magnitude as the measured values for all of these pollutants. For O_3, $PM_{2.5}$ and PM_{10} the majority of measured values are predicted to within a factor of two by the model, however the model underestimates the variability of concentrations between measurement sites. There is more variation in the model predictions between sites with the IER 2 km emissions than the standard modelling approach, however there is a slight tendency for $PM_{2.5}$ to be underestimated in Ho Chi Minh City with these emissions.

Only data from the CEM national network of monitoring sites is available for NO_2, NO_X and PM_{10} in the Vietnamese/Cambodian domain. These sites are generally located in areas outside Ho Chi Minh City, with possible influences from local industrial activity. Within Ho Chi Minh City, for O_3 and $PM_{2.5}$ data are also available from the IER/VNU-HCM sensor network, while additional $PM_{2.5}$ data has been obtained from a sensor located on the US Consulate building, published via the US AirNow system [17].

Comparisons of model results with and without explicitly modelled roads for the Vietnam/Cambodia domain (shown in Appendix B) show some influence of the explicitly modelled roads on the concentrations of NO_2, NO_X and O_3 at a few monitoring sites, but no clear influence of explicitly modelled roads on concentrations of $PM_{2.5}$ or PM_{10}. This is due to both the relatively large distances between measurement sites in this domain and the nearest explicitly modelled road, and the smaller relative contribution of road emissions to total particulate concentrations compared to their contribution to gaseous pollutant concentrations.

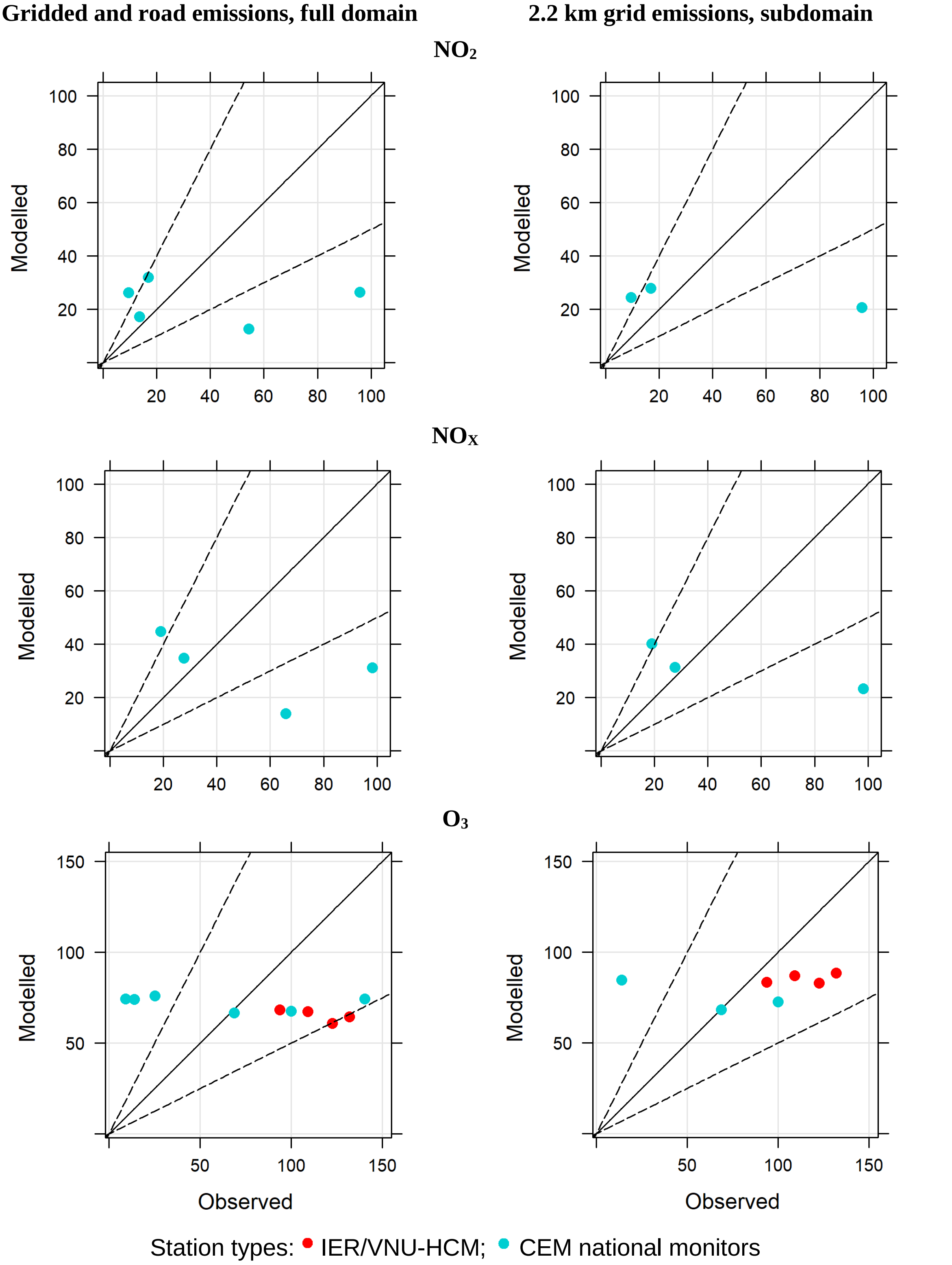


Figure 4.2 – Scatter plots comparing monthly mean NO_2, NO_X and O_3 concentrations (units $\mu g/m^3$) predicted by the modelling system to measured values, for the Vietnamese/Cambodian modelling domain and January 2022 modelling period. The left-hand panel for each pollutant shows modelled results with CAMS gridded and explicit road emissions for the full domain, while the right-hand panel uses the finer resolution IER gridded emissions in a subdomain. The solid line shows a 1:1 relationship between modelled and measured concentrations, while the dashed lines indicate factor of

two relationships.

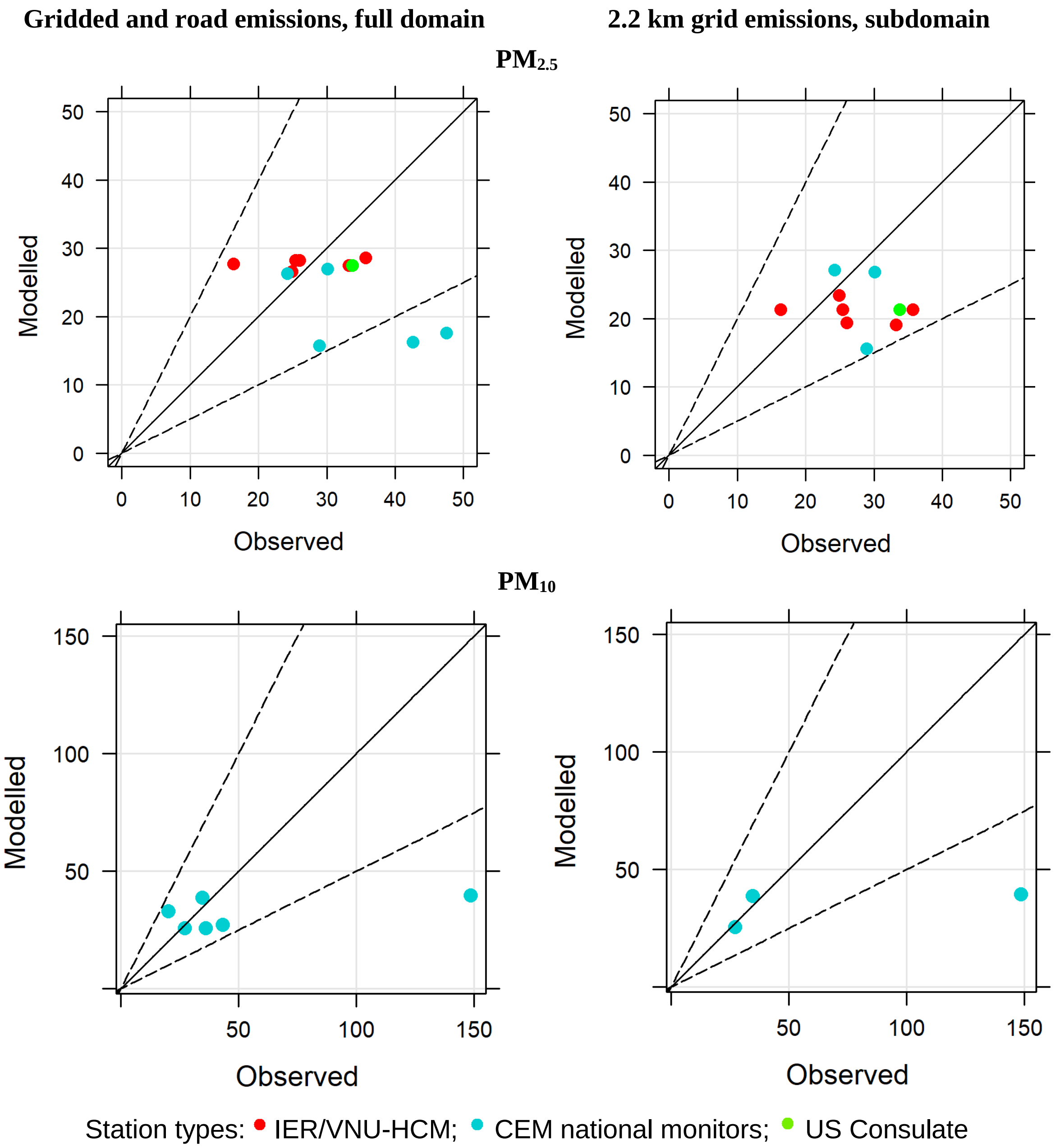


Figure 4.3 – Scatter plots comparing monthly mean $PM_{2.5}$ and PM_{10} concentrations (units µg/m^3) predicted by the modelling system to measured values, equivalent to Figure 4.2 for gaseous pollutants.

An initial evaluation of modelling results for the Malaysian domain in the 25th April – 25th May 2022 period showed substantial over-estimates of concentrations of most pollutants. This was traced to overpredictions by the CAMS global forecasts, likely due to overestimates of emission rates in the Kuala Lumpur mega-urban area and, to a lesser extent, the surrounding regions. Calibration factors to adjust the CAMS modelled concentrations were derived from the raw CAMS and observed monthly median values of the daily metrics relevant to API values for each pollutant. Calibration factor values are shown in Table 4.4. The calibration factor for SO_2 is notably low, suggesting that the CAMS emissions inventory has not yet captured recent regulatory changes that have reduced fuel sulphur levels, and hence SO_2 emissions and concentrations, in Malaysia. These calibration factors have also been applied for the month of January, to check their applicability in a season with distinct meteorology and emissions. Results shown in Appendix B suggest that the calibration factors derived from April – May are broadly appropriate for January. However, there is a tendency for the scaled model outputs to underestimate the lower observed concentrations and overestimate the higher observed concentrations, suggesting that either a spatially-varying calibration factor or an improvement of local emission data should be considered for future development of the system.

Table 4.4 – Calibration factors for CAMS modelled predictions in Malaysia, based on 25th April – 25th May 2022 modelling period

Pollutant	**O_3**	**NO_2**	**NO_X**	**SO_2**	**CO**	**PM_{10}**	**$PM_{2.5}$**
Calibration factor	0.48	0.57	0.57	0.32	0.78	0.54	0.52

Monthly average concentration scatter plots for the Malaysian domain, 25th April – 25th May 2022 period are shown in Figure 4.4 for gaseous pollutants and Figure 4.5 for particulate pollutants. The modelled concentrations include CAMS concentrations which have been adjusted using the calibration factors. For O_3, $PM_{2.5}$ and PM_{10}, monthly average concentrations are predicted within a factor of two for almost all sites. A few sites show increases in particulate concentrations due to the modelling of explicit road emissions. While most NO_2 and NO_X sites also show modelled concentrations within a factor of two of observations, it is clear that the model is predicting a larger concentration increment than observed between sites with lower and higher observed concentrations. The overestimate of observed concentrations within Kuala Lumpur by the raw CAMS modelled concentrations suggests that the gridded emissions within Kuala Lumpur are overestimated. This overestimate in gridded emissions leads to an overestimate in the top-down calculations of explicit road emissions and hence an excessive concentration increment from the explicitly modelled roads.

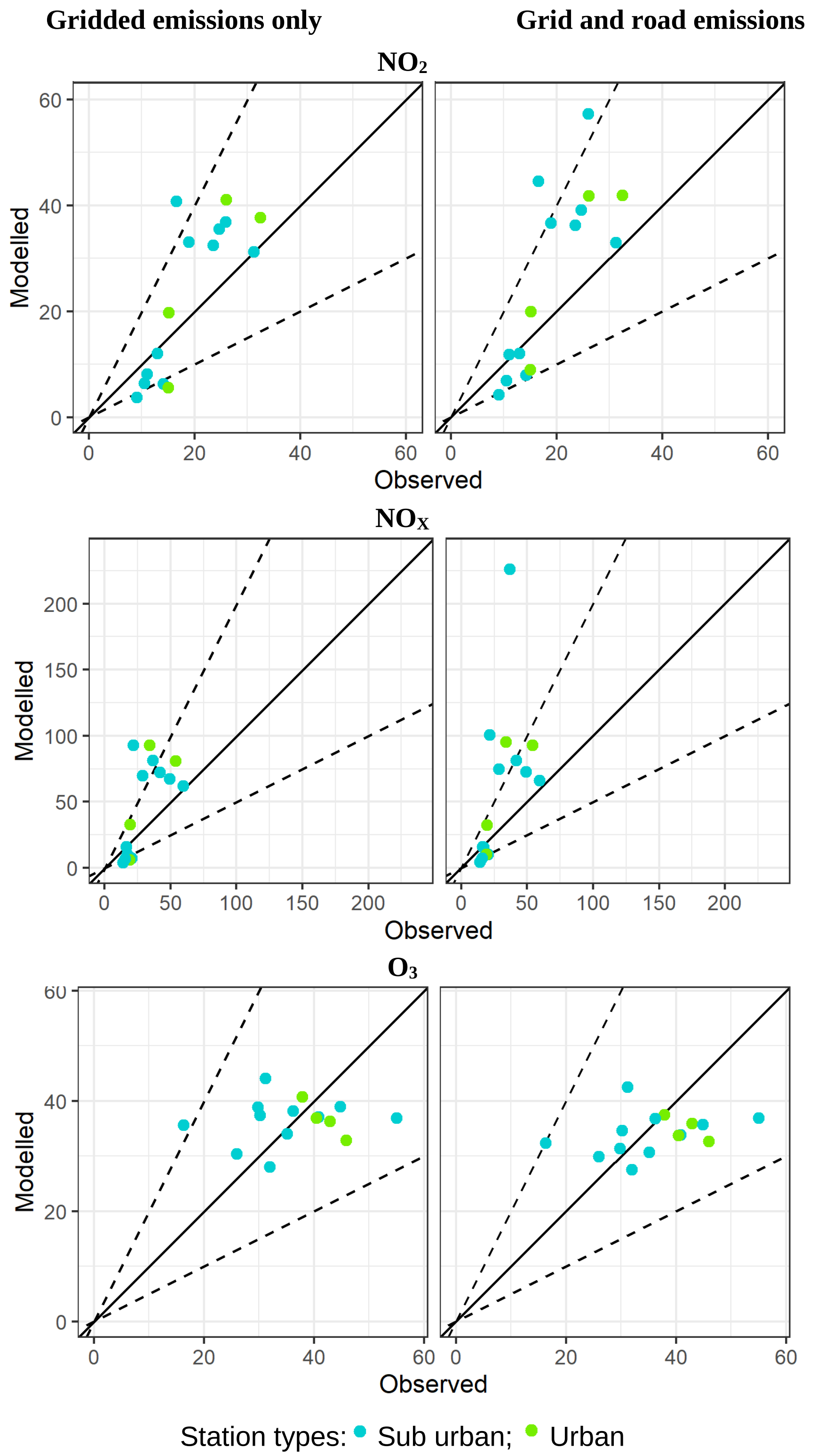


Figure 4.4 – Scatter plots comparing monthly mean NO_2, NO_X and O_3 concentrations (units μg/m^3) predicted by the modelling system to measured values, for the Malaysian modelling domain and 25th April – 25th May 2022 modelling period, including adjustment of CAMS concentrations by calibration factors. The left-hand panel for each pollutant shows modelled results with only gridded emissions, while the right-hand panel shows results with both gridded and explicit road emissions. The solid line shows a 1:1 relationship between modelled and measured concentrations, while the dashed lines indicate factor of two relationships.

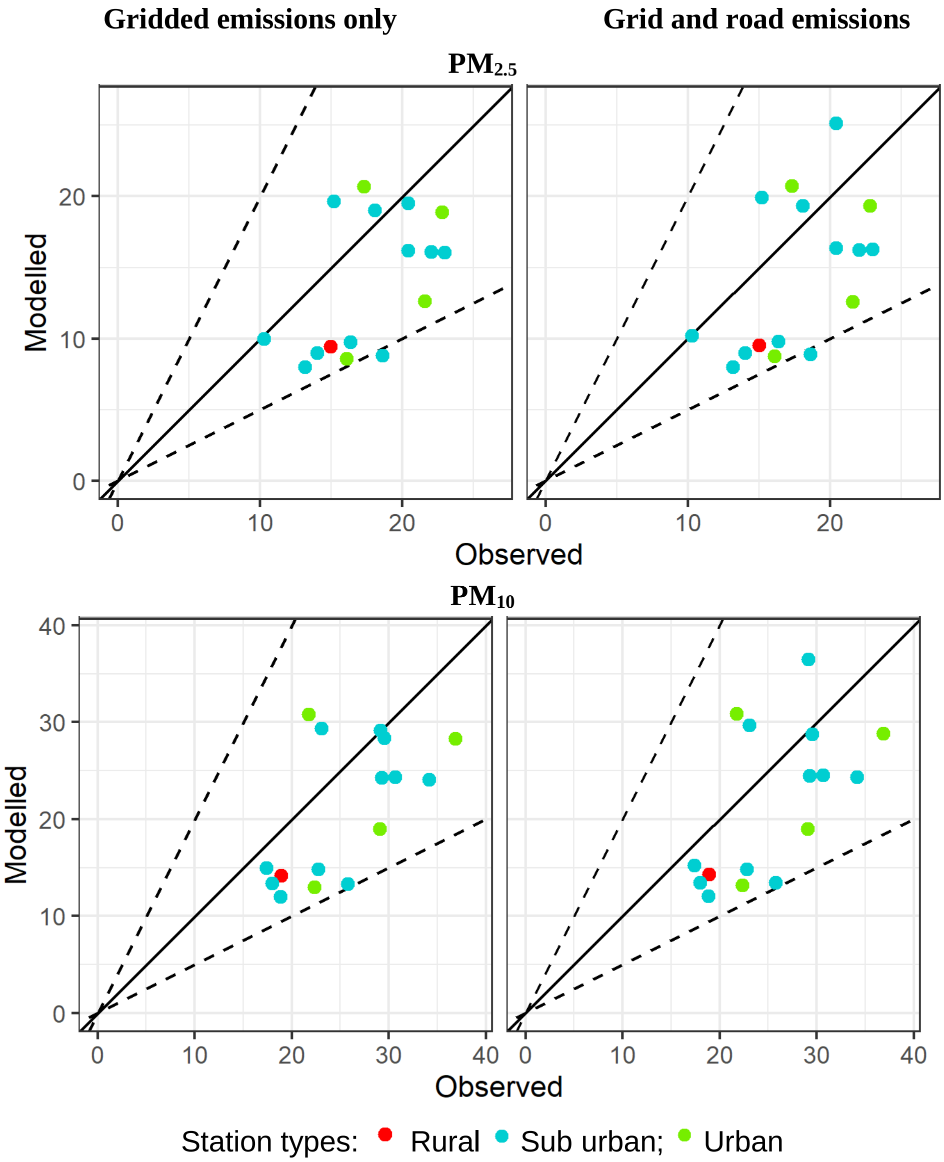


Figure 4.5 – Scatter plots comparing monthly mean $PM_{2.5}$ and PM_{10} concentrations (units μg/m^3) predicted by the modelling system to measured values, for the Malaysian modelling domain and 25th April – 25th May 2022 modelling period, including adjustment of CAMS concentrations by calibration factors. Equivalent to Figure 4.4 for gaseous pollutants.

The ability of the modelling system to capture the temporal variation of pollutant concentrations can be examined using time-series plots of the metrics which relate to the AQI values. Selected examples comparing model output (with explicit road influences) and measurements at individual sites in the Vietnamese/Cambodian domain are shown in Figure 4.6 for O_3 (8-hour rolling mean at IER/VNU-HCM Sensor 3) and Figure 4.7 for $PM_{2.5}$ (24-hour mean at site IER/VNU-HCM Sensor 1). The O_3 time series shows that the model is capturing most of the daily variation at this site, and is matching most of the days that reach the threshold concentration for the 'unhealthy for sensitive groups' AQI index level (121 μg/m^3). The $PM_{2.5}$ time series shows that the model is over-predicting measured concentrations for the first eight days of the month, but then captures most of the measured variation over the rest of the month, including days which reach the 'moderate' AQI index level (26 μg/m^3).

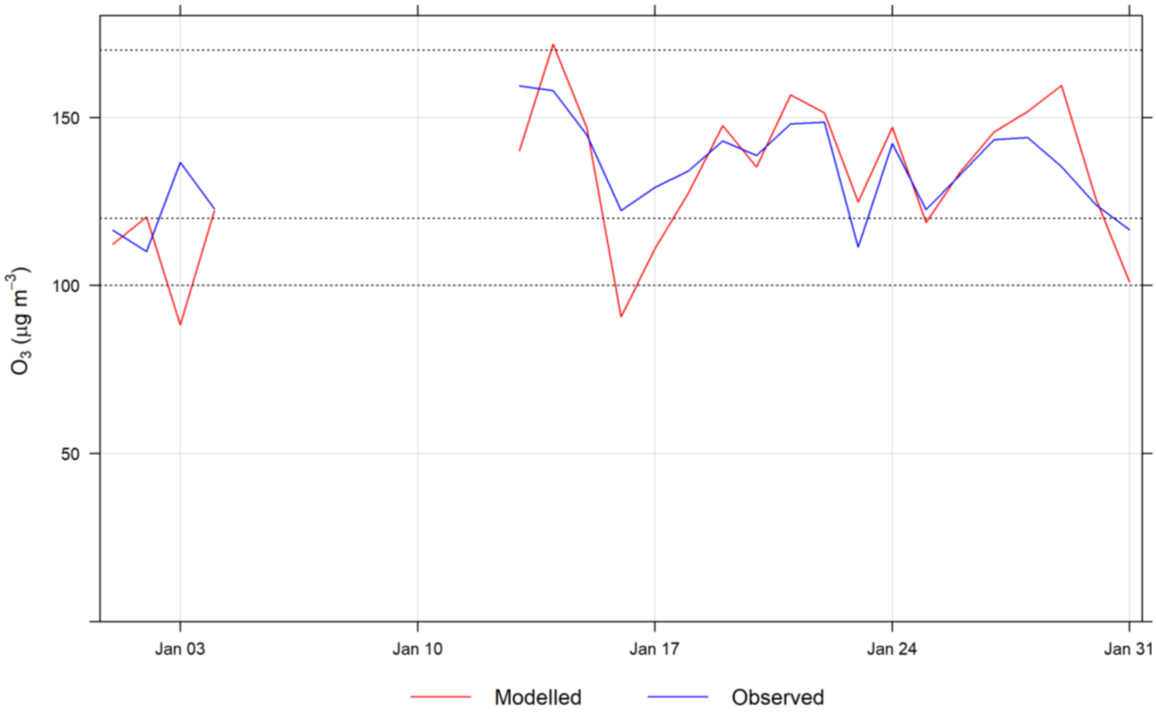


Figure 4.6 – Time series of modelled (red) and observed (blue) daily maximum rolling 8-hour average O_3 concentrations (units µg/m^3), for IER/VNU-HCM sensor 3 in Ho Chi Minh City, January 2022. Note that no measurement data is available from this sensor for the period 5th – 12th January. The dotted lines show the concentration thresholds associated with the 'Moderate', 'Unhealthy for sensitive groups' and 'Unhealthy' AQI levels.

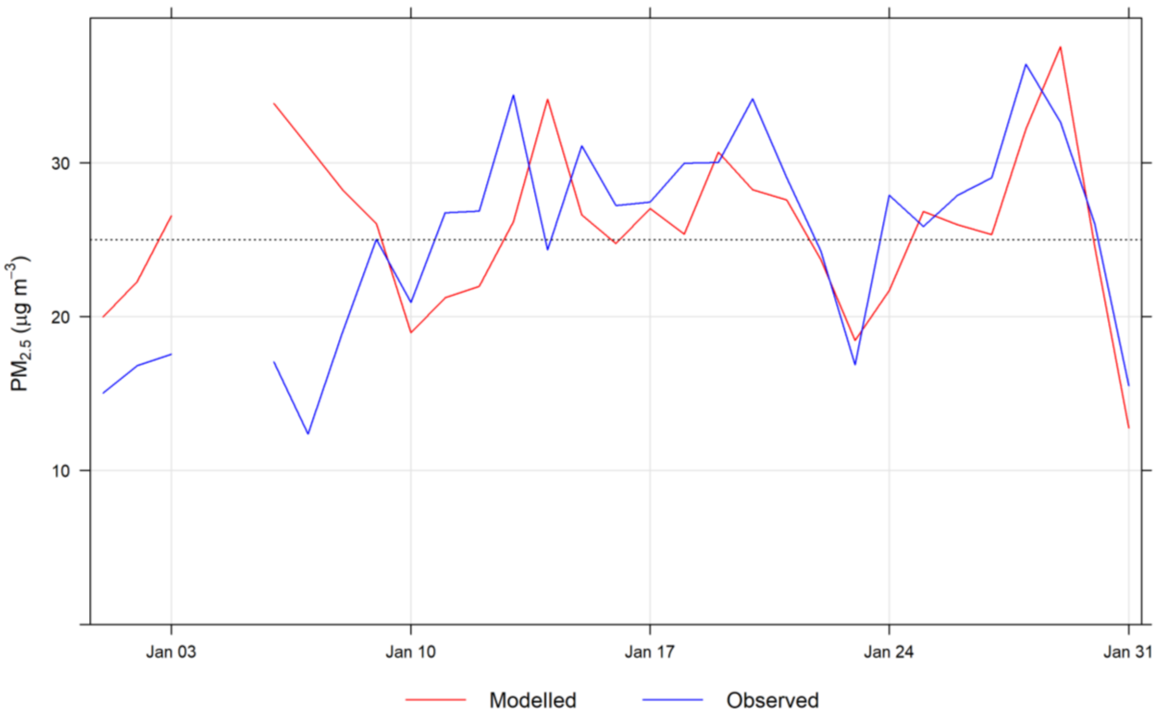


Figure 4.7 – Time series of modelled (red) and observed (blue) 24-hour average $PM_{2.5}$ concentrations (units µg/m^3), for IER/VNU-HCM sensor 1 in Ho Chi Minh City, January 2022. The dotted line shows the concentration threshold associated with the Air Quality Index 'Moderate' level.

Time-series of modelled and observed daily $PM_{2.5}$ concentrations at the CA20B monitor in the Malaysia domain for 25th April – 25th May 2022 are shown in Figure 4.8. During most of the modelling period, both the modelled and observed values are in the 'Moderate' API index level at this site. For 7th – 10th May, the model predicts reduced concentrations whereas the observations show concentrations remaining elevated. On 19th May, the observations reach the 'Unhealthy for sensitive groups' API level whereas the model prediction remains at 'Moderate'. Comparing similar plots for multiple sites in the Malaysia domain, the observations show much more variability in the day-to-day variation of concentrations than is predicted by the model, which gives a similar pattern across the modelling period at most sites. This suggests that local short-term emissions variations, not incorporated into the modelling, may be influencing the observed concentrations. Model performance could be improved by incorporating more local input data relating to emission source locations, activity and emission rates, and/or by assimilating observed concentrations in a near-real-time calibration process.

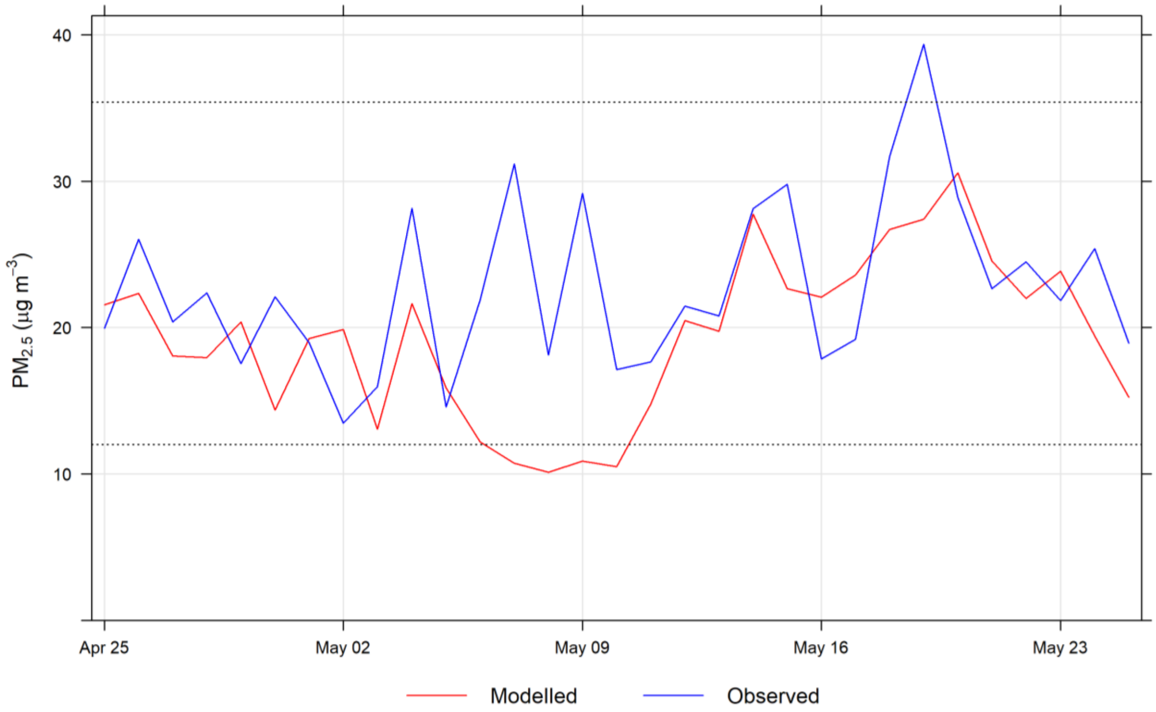


Figure 4.8 – Time series of modelled (red) and observed (blue) 24-hour average $PM_{2.5}$ concentrations (units µg/m^3), for monitor CA20B in Malaysia, 25th April – 25th May 2022. The dotted lines show the concentration thresholds associated with the Air Pollutant Index 'Moderate' and 'Unhealthy for sensitive groups' levels.

The comparison of modelled and measured air pollution levels in relation to the descriptive AQI/API bands (e.g. 'Good' or 'Unhealthy for sensitive groups') allows an assessment of the accuracy of forecasting these levels. Figure 4.9 shows the proportion of modelled O_3 air quality levels which are equal to, or within one level of, the corresponding measured value for ten sites in the Vietnamese/Cambodian domain for January 2022; the corresponding plot for $PM_{2.5}$ is shown in Figure 4.10. Both of these plots use the standard CAMS/EDGAR emissions inventory. There is some variability in the forecast accuracy between measurement sites and pollutants, with a greater proportion of modelled air quality level values within one level of measurements for $PM_{2.5}$, but a greater proportion of modelled levels equal to measurements for O_3. The two sites with predicted values only within one level of measurements for $PM_{2.5}$ both have daily measured concentrations systematically underestimated by the modelled forecast. They are both part of the CEM national network and located to the south-east of Ho Chi Minh City, outside the area of IER higher resolution emissions, where they may be affected by emissions from local industrial activities.

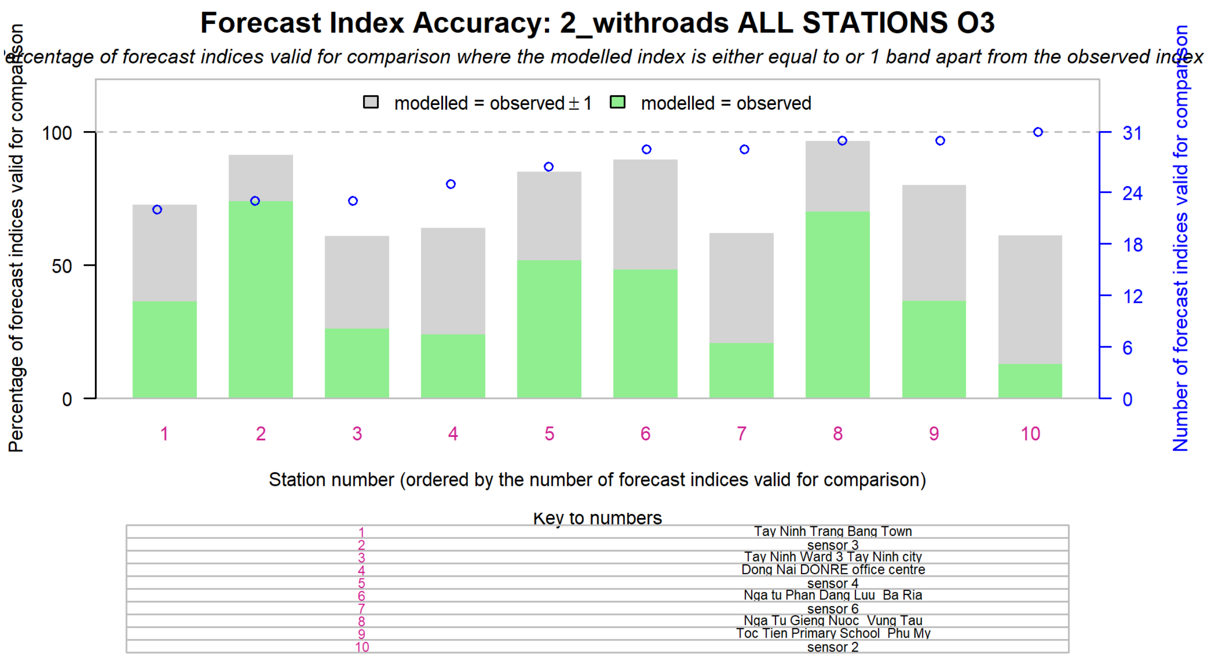


Figure 4.9 – Assessment of O_3 forecast accuracy based on the proportion of modelled forecast level (e.g. Good, Moderate) which are equal to the corresponding measured level (green bars) or within one level of the measured level (grey bars), for Vietnam/Cambodia domain January 2022. For example, if the observed level is Moderate, a modelled level of Moderate would be included in the green bar while a modelled level of Good or Unhealthy for Sensitive Groups would be included in the grey bar.

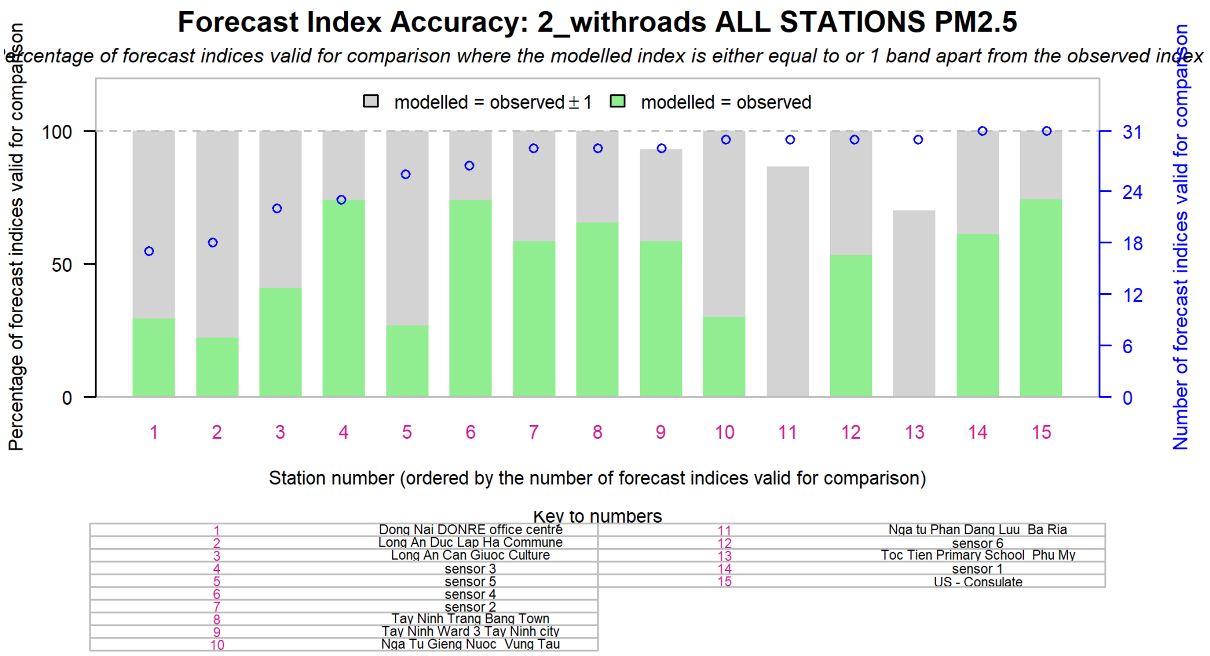


Figure 4.10 – Assessment of $PM_{2.5}$ forecast accuracy based on the proportion of modelled forecast

level (e.g. Good, Moderate) which are equal to the corresponding measured level (green bars) or within one level of the measured level (grey bars), for Vietnam/Cambodia domain January 2022.

The proportions of forecast AQI/API levels which match observed levels or are within one level of observed levels are summarised in Table 4.5 for both domains and all six index pollutants. Note that the values for Malaysia include calibration factors applied to the CAMS concentrations whereas the values for Vietnam/Cambodia do not include any calibration. The proportion of forecast levels which are the same as observed is higher for O_3 and $PM_{2.5}$ in Malaysia compared to Vietnam, which may indicate that the calibration approach is useful in improving forecast performance. However, the performance for $PM_{2.5}$ remains poorer than for O_3 in Malaysia, reflecting additional uncertainties in local emissions and measurements of this pollutant.

Table 4.5 – Proportion of forecast AQI/API levels which match the observed level or are ±1 level away from the observed for both domains and all index pollutants. Proportions calculated as an average over all available sites with at least 50% valid data.

Domain Modelling period	Pollutant	Number of sites	Forecast AQI/API level accuracy Proportion same level as observed	 Proportion ±1 level from observed
Vietnam/Cambodia January 2022	O_3	10	0.39	0.37
	NO_2	3	0.98	0.00
	$PM_{2.5}$	12	0.46	0.50
	PM_{10}	6	0.81	0.12
	CO	8	0.87	0.00
	SO_2	13	0.92	0.06
Malaysia 25th April – 25th May	O_3	15	0.91	0.08
	NO_2	15	0.90	0.10
	$PM_{2.5}$	16	0.65	0.35
	PM_{10}	16	0.99	0.01
	CO	15	1.00	0.00
	SO_2	15	0.95	0.05

Pollution maps

Pollution maps have been generated for the Vietnamese-Cambodian domain for 5th January 2022. Maps have been generated both for raw concentration values and for AQI levels, for the whole domain and a zoomed-in area covering Ho Chi Minh City. These areas are shown with background maps in Figure 4.11 for comparison, with areas of ‘industrial’ land use from the Open Street Map dataset [12] displayed for reference.

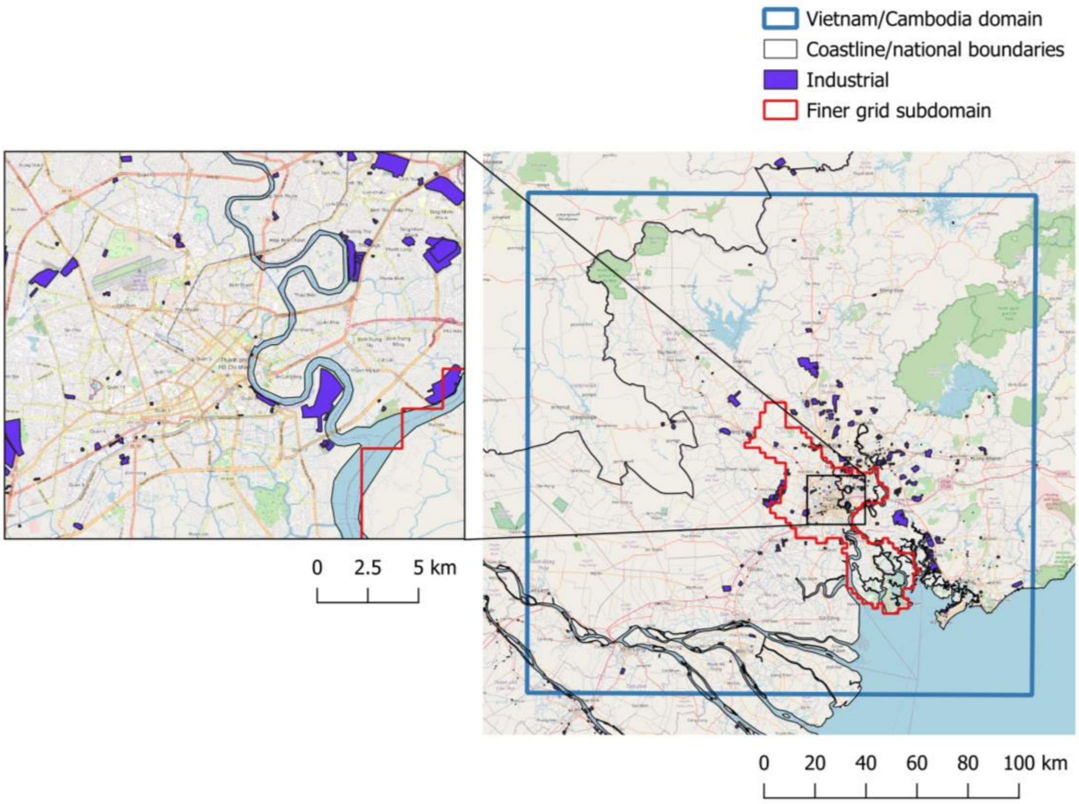


Figure 4.11 – Reference features for Vietnamese/Cambodian domain pollution maps, showing coastline and national borders (black lines), areas of 'industrial' land use from OSM and the finer resolution grid subdomain (red lines) (background data/maps copyright 2018 Geofabrik GmbH and OpenStreetMap Contributors.

Example pollution maps at 20 m spatial resolution for modelled maximum 1-hour average NO_2 concentrations on 5th January 2022 are shown in Figure 4.12, with the corresponding AQI levels shown in Figure 4.13. The concentration map shows a strong influence of the explicitly modelled road network, while the AQI map shows that levels are generally in the 'Good' level. Maps for maximum 8-hour rolling average O_3 concentration are included in Figure 4.14, with corresponding AQI levels in Figure 4.15. These show a wide area of 'Moderate' AQI level covering Ho Chi Minh City and further to the south-west, with reduced concentrations and AQI levels on the major road network, due to local chemical reactions between O_3 and NO_2. The equivalent maps for $PM_{2.5}$ are presented in Figure 4.16 for concentration and Figure 4.17 for AQI level. These show a much weaker influence of the explicitly modelled road network than NO_2 or O_3, with the overall spatial pattern of concentration dominated by a large area of high concentration to the north and west of Ho Chi Minh City centre, possibly associated with industrial activities. This leads to a widespread area of 'Moderate' AQI level predictions.

Additional example pollution maps for both domains are shown in Appendix B, including maps of the Vietnam/Cambodia domain showing PM_{10} on 5th January. The overall AQI value issued by an operational forecast would be calculated as the maximum of the values for individual pollutants. Aggregated AQI values would also be calculated for spatial areas such as local government districts in order to determine whether an alert should be issued to the relevant population. The spatial aggregation is typically based on the level predicted to be exceeded across 10% of the relevant area.

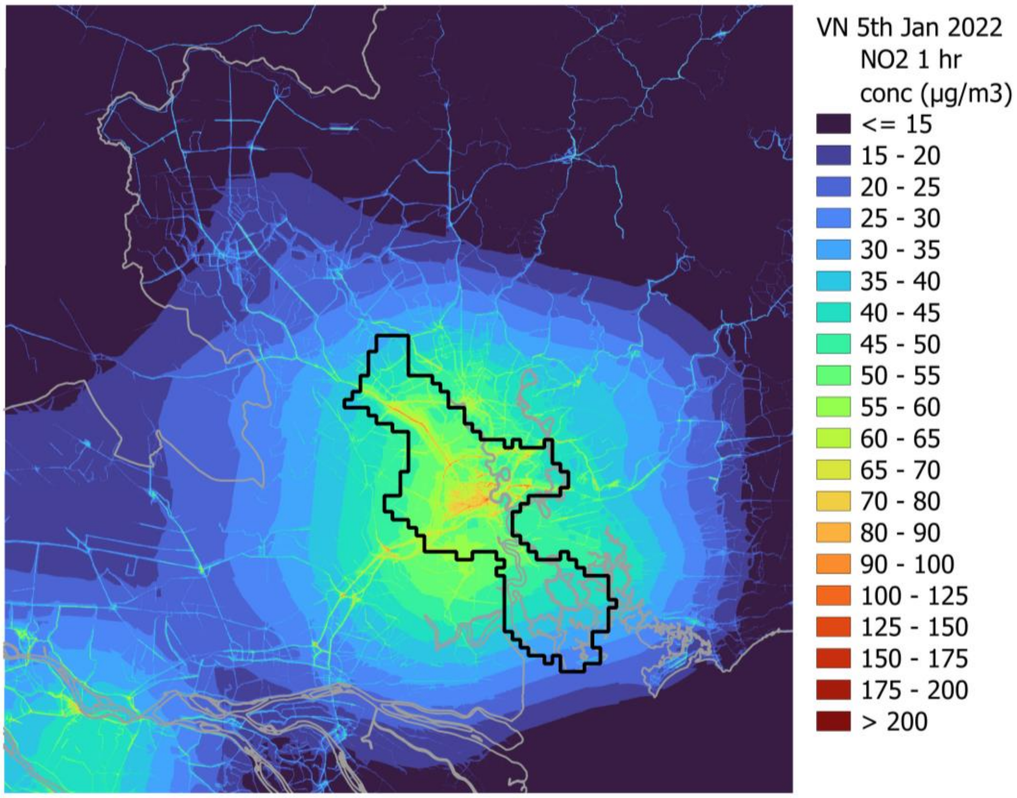


Figure 4.12 – Maximum 1-hour average NO_2 concentrations on 5th January from the MAQS coupled system for the full Vietnamese/Cambodian domain. The coastline and Vietnamese/Cambodian national border are shown as grey lines, while the extent of the subdomain with IER/VNU-HCM finer resolutions is outlined in black.

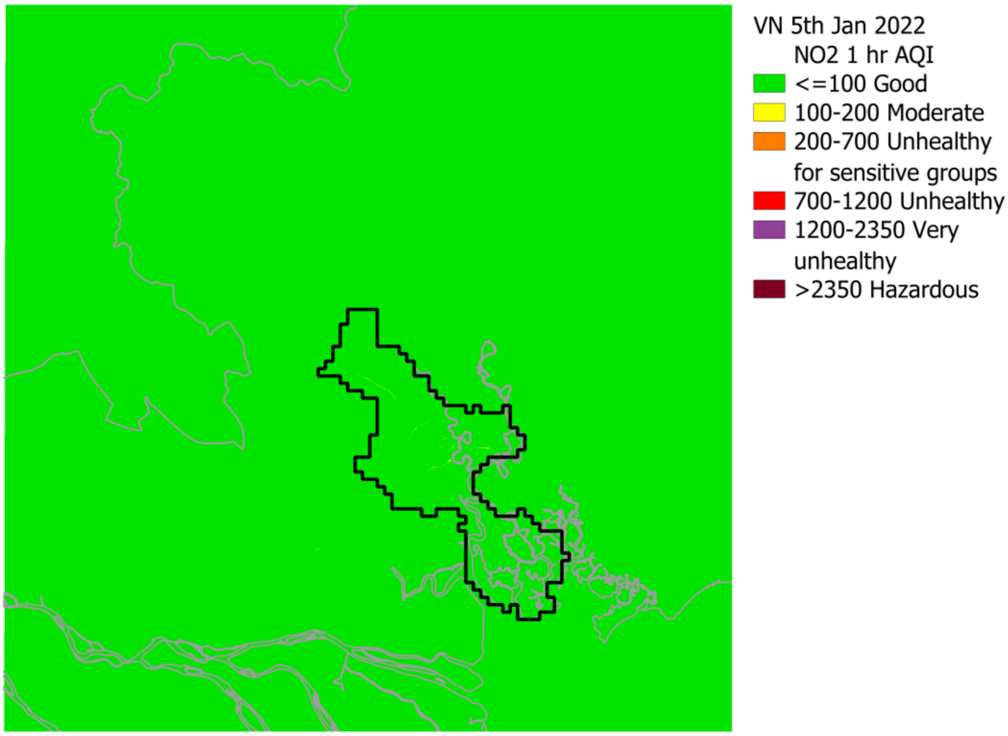


Figure 4.13 – NO_2 AQI levels on 5th January predicted by the MAQS coupled system, equivalent to

Figure 4.12 for concentration.

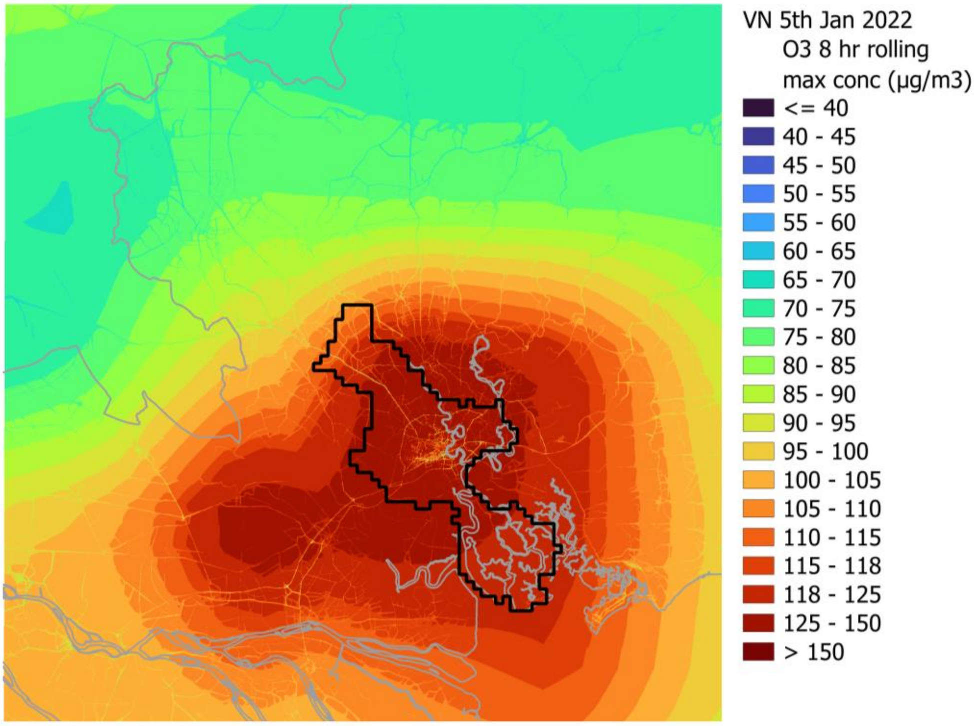


Figure 4.14 – Maximum 8-hour rolling average O_3 concentrations on 5th January from the MAQS coupled system for the full Vietnamese/Cambodian domain. The coastline and Vietnamese/Cambodian national border are shown as grey lines, while the extent of the subdomain with IER/VNU-HCM finer resolutions is outlined in black.

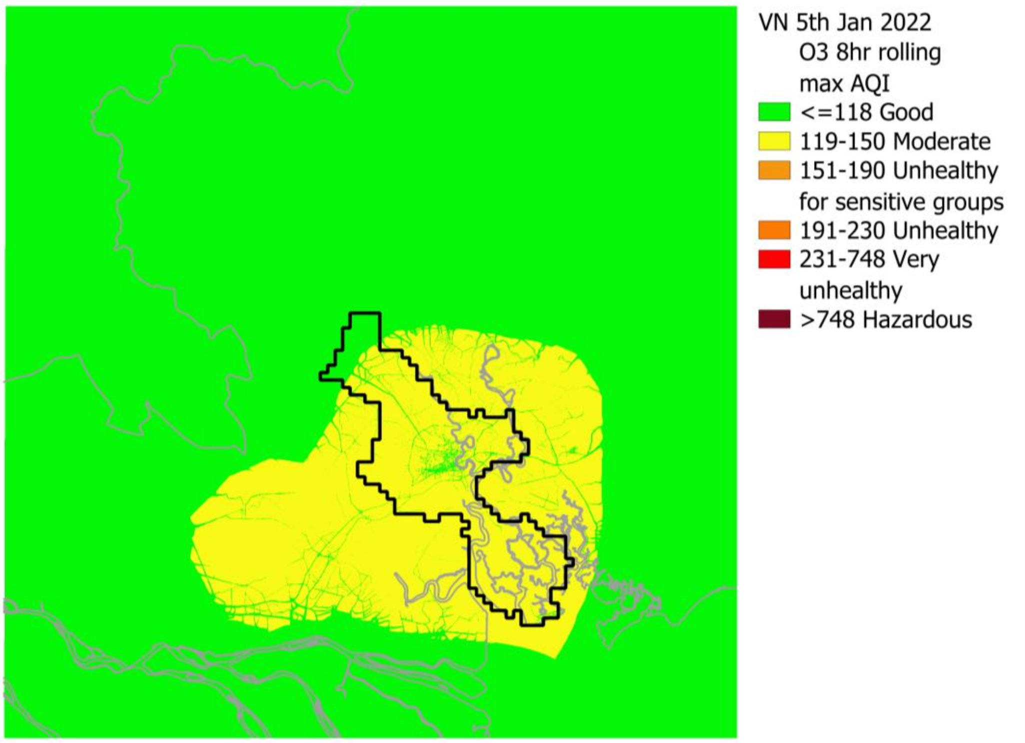


Figure 4.15 – O_3 AQI levels on 5th January predicted by the MAQS coupled system, equivalent to Figure 4.14 for concentration.

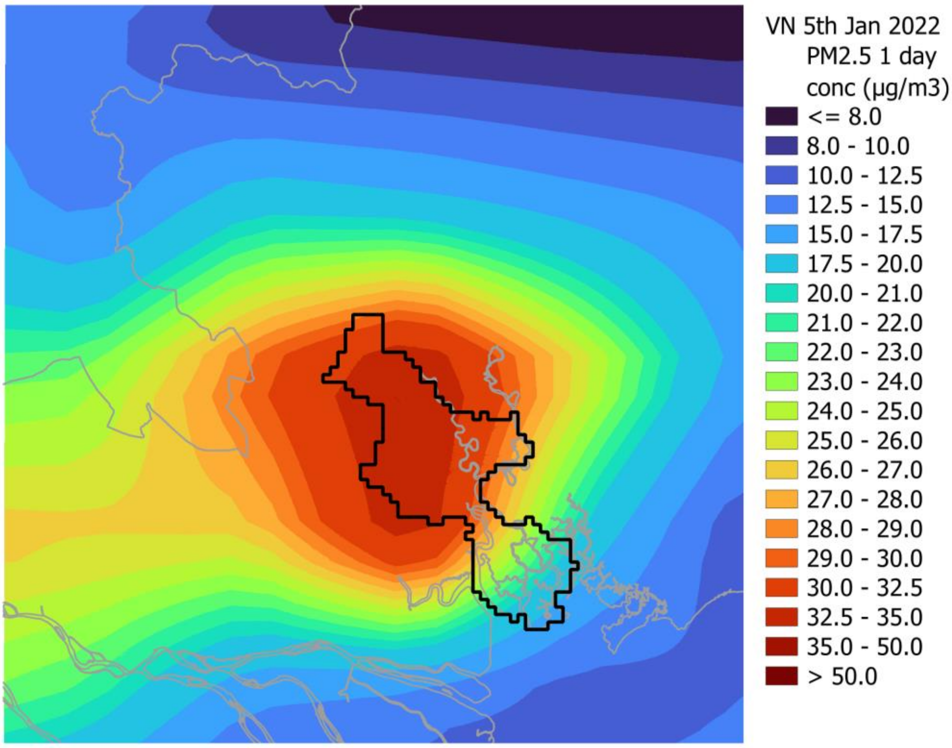


Figure 4.16 – 24-hour average $PM_{2.5}$ concentrations on 5th January from the MAQS coupled system for the full Vietnamese/Cambodian domain. The coastline and Vietnamese/Cambodian national border are shown as grey lines, while the extent of the subdomain with IER/VNU-HCM finer resolutions is outlined in black.

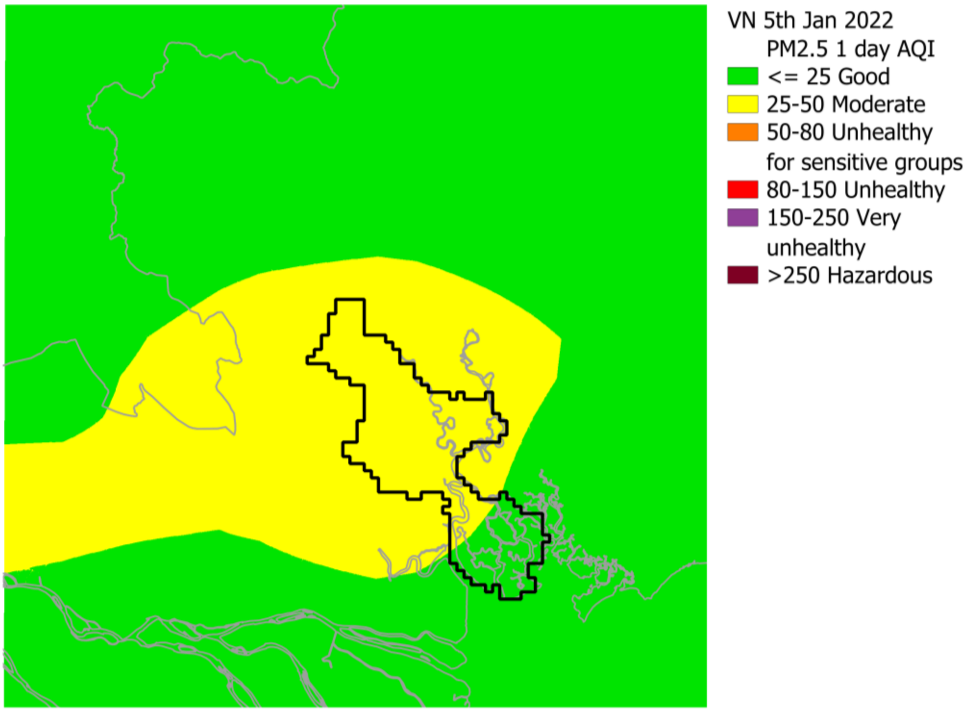


Figure 4.17 –$PM_{2.5}$ AQI levels on 5th January predicted by the MAQS coupled system, equivalent to Figure 4.16 for concentration.

The 2 km resolution gridded emissions shared by IER/VNU-HCM were used for modelling within a subdomain covering Ho Chi Minh City and surrounding provinces. Maps comparing model outputs from the coupled system, with low-resolution gridded and explicit road emissions, or high-resolution gridded emissions only, are compared in Figure 4.18 for NO_2 and Figure 4.19 for $PM_{2.5}$. The runs using the IER/VNU-HCM gridded emissions show a different spatial distribution of concentrations relating to non-road sources, particularly elevated NO_2 from port activities, and much lower concentrations of $PM_{2.5}$ within Ho Chi Minh City than runs using the standard emissions.

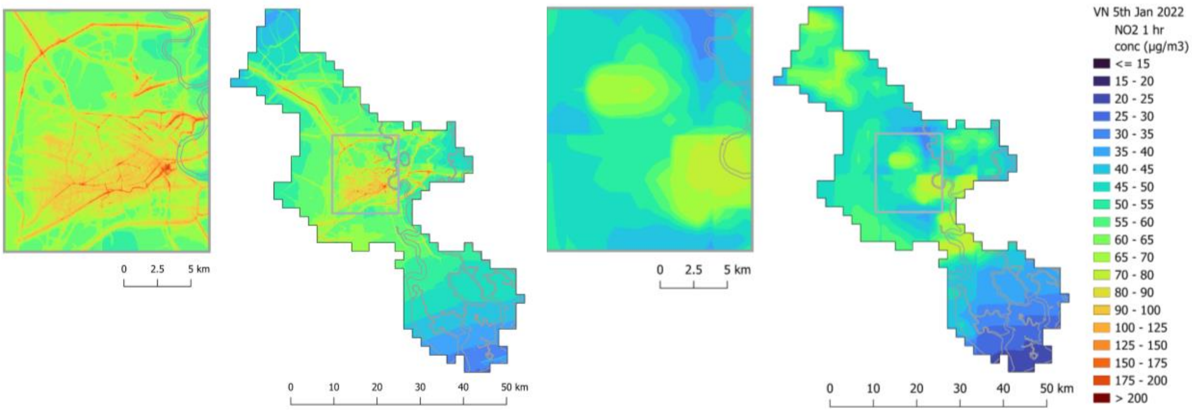


Figure 4.18 – Maximum 1-hour average NO_2 concentrations on 5th January from the MAQS coupled system for the subdomain around Ho Chi Minh City where IER/VNU-HCM emissions are available. The left-hand panels show model outputs with standard gridded and explicit road emissions, equivalent to Figure 4.12, while the right-hand panels show outputs with finer resolution gridded

emissions only.

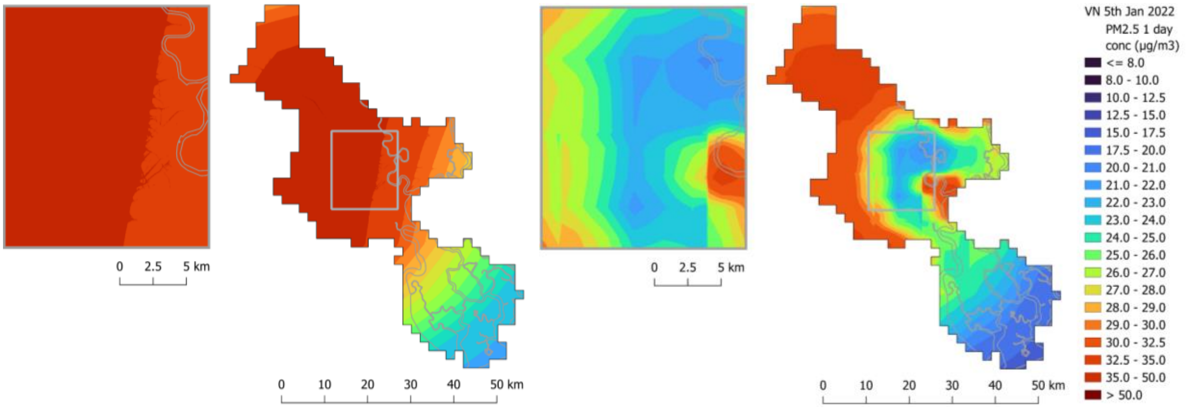


Figure 4.19 – 24-hour average $PM_{2.5}$ concentrations on 5th January from the MAQS coupled system for the subdomain around Ho Chi Minh City where IER/VNU-HCM emissions are available. The left-hand panels show model outputs with standard gridded and explicit road emissions, equivalent to Figure 4.16, while the right-hand panels show outputs with finer resolution gridded emissions only.

Example gridded forecast outputs for the Malaysia domain were generated for 13 days, from 10th – 22nd May 2022, in order to create animations showing the variation of concentrations/API levels between days. Sample NO_2 concentration and API maps for 14th May are presented in Figures 4.20 and 4.21. These show a substantial increment of concentrations within the Kuala Lumpur urban area and along the major road network, with 'Moderate' API level for the urban area and 'Unhealthy for sensitive groups' near major roads. Figures 4.22 and 4.23 show daily $PM_{2.5}$ API maps for 15th – 22nd May. These indicate 'Moderate' API levels within Kuala Lumpur on all days, with 'Unhealthy for sensitive groups' consistently along the road network and also in a wider area across the south of Kuala Lumpur on 20th May. The shape of the area of 'Moderate' API level associated with $PM_{2.5}$ varies between days due to varying emission rates and meteorological conditions. These plots demonstrate the coupled system capturing interactions between large-scale and local influences on air pollution levels.

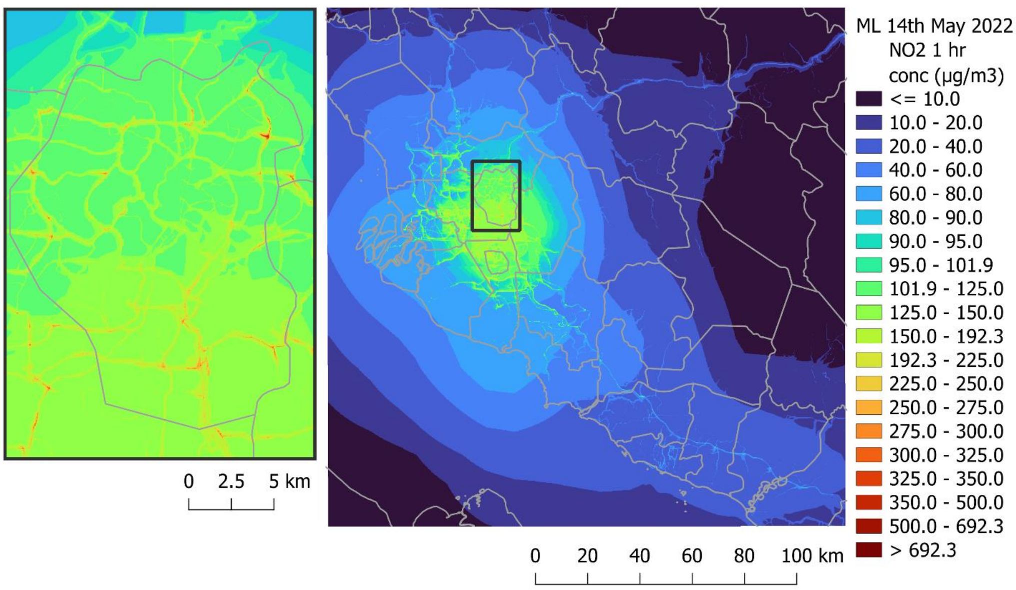


Figure 4.20 – Maximum 1-hour average NO_2 concentrations on 14th May from the MAQS coupled

system for the full Malaysian domain (right-hand panel) and a higher-resolution area over central Kuala Lumpur (left-hand panel). The coastline and internal province borders are shown as grey lines.

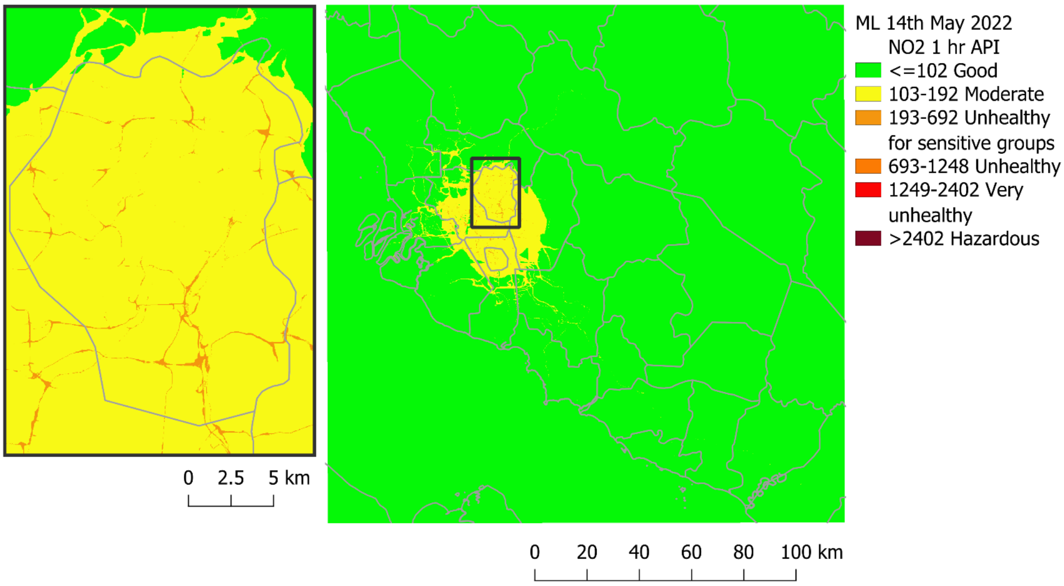


Figure 4.21 – API levels associated with maximum 1-hour average NO_2 concentrations on 14th May from the MAQS coupled system for the full Malaysian domain (right-hand panel) and a higher-resolution area over central Kuala Lumpur (left-hand panel). The coastline and internal province borders are shown as grey lines.

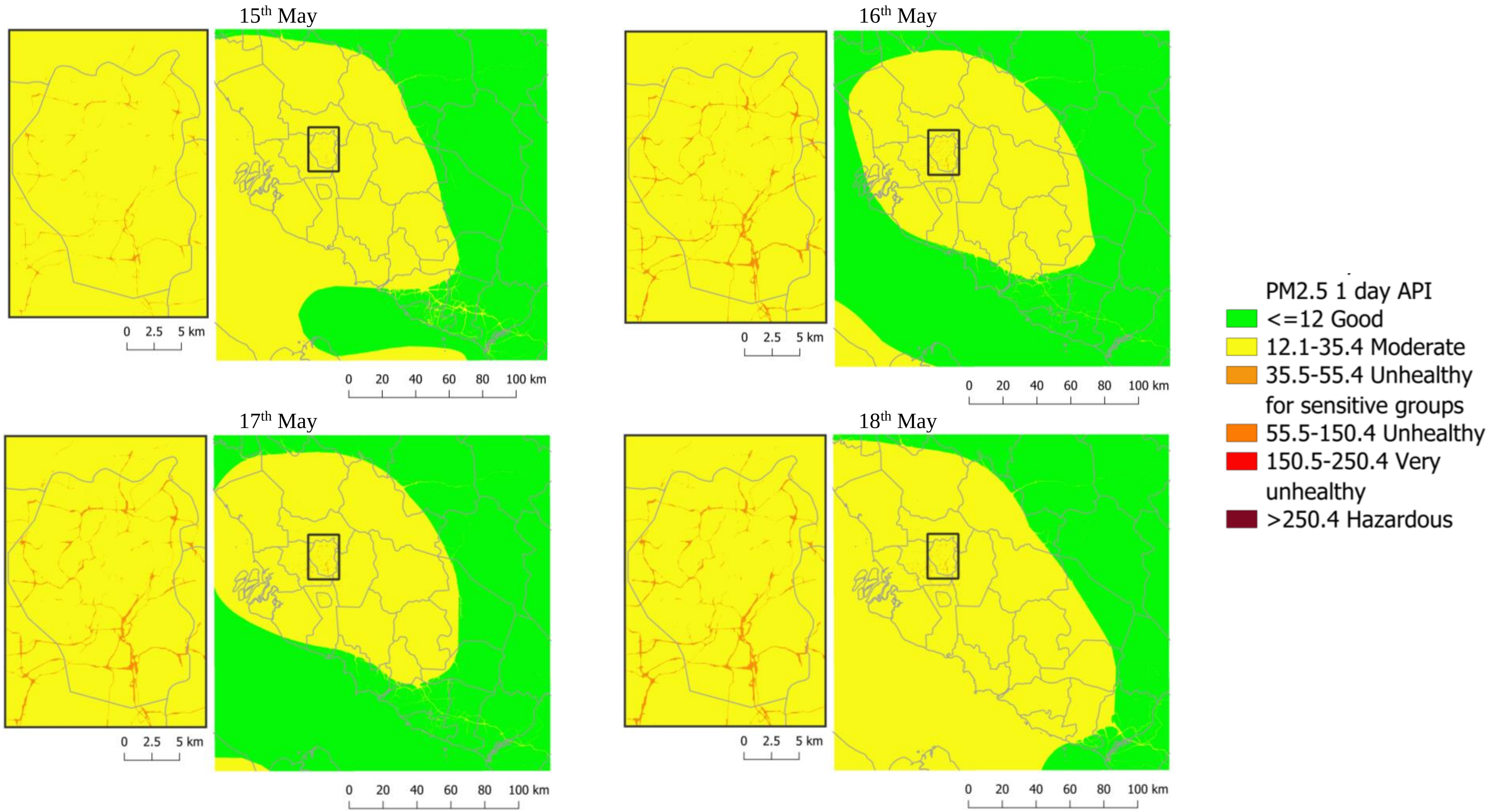


Figure 4.22 – API levels associated with 24-hour average $PM_{2.5}$ concentrations from 15th – 18th May from the MAQS coupled system for the full Malaysian domain (larger panels) and a higher-resolution area over central Kuala Lumpur (smaller panels with black outline). The coastline and internal province borders are shown as grey lines.

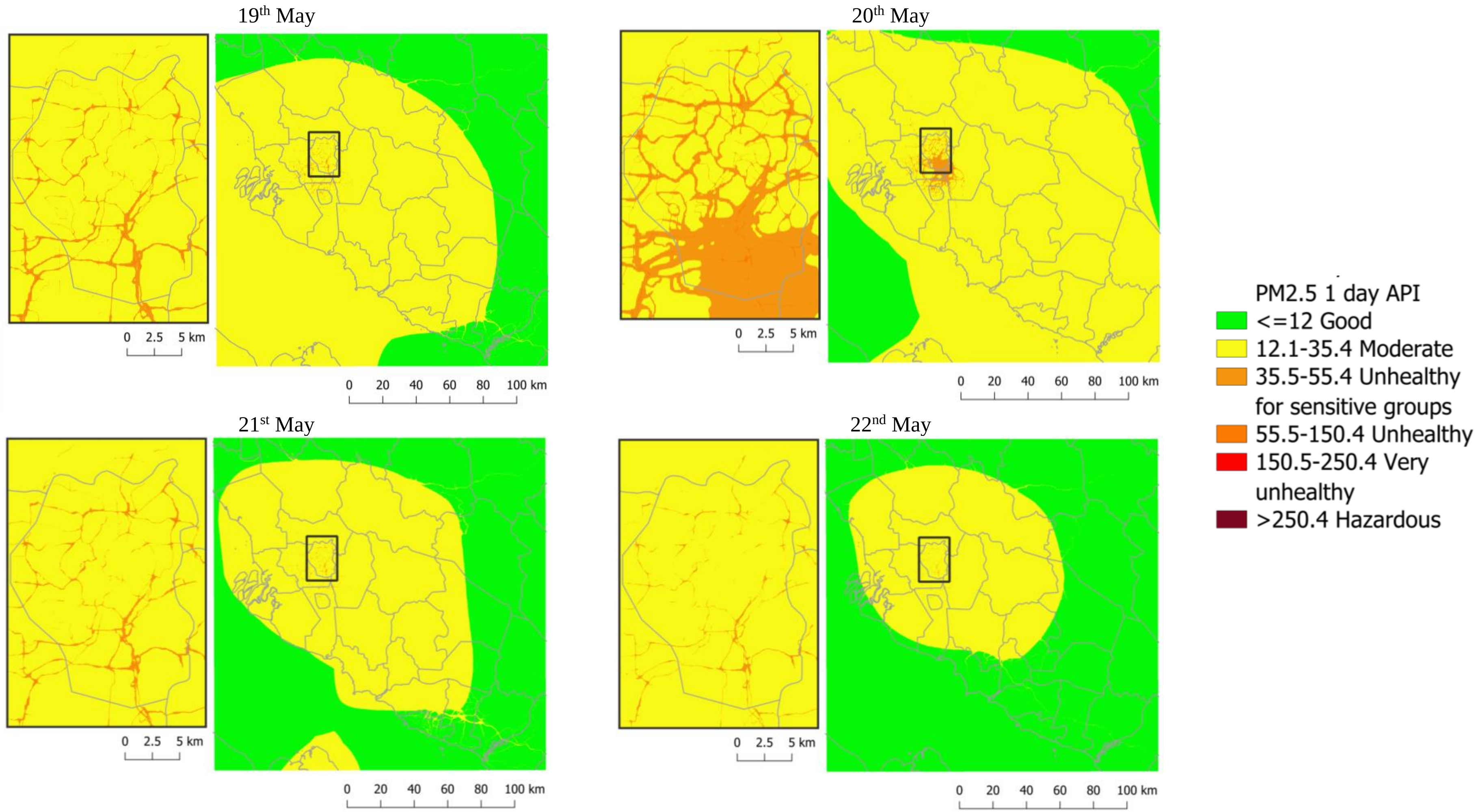


Figure 4.23 – API levels associated with 24-hour average $PM_{2.5}$ concentrations from 19th – 22nd May from the MAQS coupled system for the full Malaysian domain (larger panels) and a higher-resolution area over central Kuala Lumpur (smaller panels with black outline). The coastline and internal province borders are shown as grey lines.

5 Stakeholder engagement

The stakeholder engagement activities have been undertaken in line with the following project aim: 'To stimulate interest in the project within the ASEAN stakeholder community'. An online project stakeholder workshop took place on 16th May 2024. UKM led this activity, with support from CERC and IER/VNU-HCM.

An overview of the timeline for the preparation and execution of stakeholder activities is given in Section 5.1. The outcomes of the stakeholder engagement activities are described in Section 5.2. Further information is provided in Section 6.5 and Appendix C.

5.1 Activity timeline

The stakeholder activity timeline is summarised by the flowchart presented in Figure 5.1. The project team were aware that in order to disseminate project information to a wide range of stakeholders, engagement activities needed to start as early in the project as possible. The work began in January with the collation of a primary stakeholder list. This was followed by the creation of a range of materials for the dissemination of project information. Both of these items fed into the first project email, which was e-mailed to the primary set of stakeholders in March, in order to bring their attention to the project. Subsequently, the project team engaged additional stakeholders, where possible, for example via FCDO as well as through secondary contacts from the primary stakeholder list.

Workshop planning began with setting the workshop date and time, deciding what content should be presented, inviting external speakers and compiling a draft agenda. An understanding of how poor air quality impacts population health was considered to be helpful to motivate interest in air quality forecasting systems. For this reason, a UKM health researcher, Prof. Mazrura Sahani, was invited to speak at the workshop, to explain how short-term pollution exposure can lead to adverse health effects. Additionally, it was important for the workshop to be chaired by a representative from one of the ASEAN project teams. UKM identified Dr Nurfashareena Muhamad, Head of the Southeast Asia Disaster Prevention Research Initiative (SEADPRI-UKM), as a suitable workshop moderator.

Summary workshop information, including the draft agenda, was sent to all stakeholders in April. Workshop preparation continued throughout April and May, comprising of meeting logistics, organisation, technical content preparation, and further stakeholder emails. The project team and external speakers held a technical run-through three days before the event, in order to ensure the workshop ran smoothly. The two-hour online event took place on Thursday 16th May 2024, and ran to time.

Workshop follow-up activities have thus far involved the circulation of workshop materials presented at the meeting and a meeting between UKM and a Cambodian stakeholder.

5.2 Stakeholder engagement outcomes

This section provides further details of stakeholder engagement activities.

Project information dissemination materials

A project description is hosted on the CERC website (Figure C.1). This description includes details of the project focus, partners, funding and outputs. The text includes multiple embedded weblinks, which increase website traffic and improve search engine rankings.

The project team created a project description flyer (Figure C.2), providing project information in an accessible format for those who prefer not to click on website links from unknown sources.

The agenda for the project stakeholder workshop was designed to be visually pleasing, including an image of a street-scale pollution map, generated as part of the project (Figure C.3).

Stakeholder list

In order to ensure maximum participation in the workshop, the project team began compiling the stakeholder list at an early stage of the project. The primary stakeholder list included contact details for 22 individuals from government and academic organisations, representing the following ASEAN members: Philippines, Malaysia, Vietnam, Cambodia, Laos, Singapore, Indonesia, and Brunei Darussalam.

Following the circulation of the initial project information and subsequently the details of the online workshop, the stakeholder list expanded through workshop registration. Additionally, the project team's FCDO contact, Jeremie Credo, circulated project and workshop details to her contacts, which included specifically interested parties at the British Embassies and High Commissions in ASEAN, and the ASEAN COSTI.

Over 60 different organisations were included in the final project stakeholder list. Tables C.1 and C.2 summarise ASEAN and non- ASEAN project stakeholder organisations, respectively.

Online dissemination of project information

CERC posted a News Item describing the project[5]. This News Item will have been seen by CERC's followers on LinkedIn, Facebook and X. UKM also posted project information and publicised the workshop on their institute webpage[6] and on their institute Facebook page[7].

Both the project registration (via Google Forms) and the meeting link (Zoom) information was provided on the online project dissemination materials.

Communication with project stakeholders

A number of emails were sent to project stakeholders, specifically:

- The first project information email was circulated to those on the primary stakeholder list (20th March).
- A stakeholder workshop invitation email, including the Google Forms registration link,

[5] The News Item text is similar to overall project description and so has not been given explicitly here, but can be accessed via https://www.cerc.co.uk/about-us/news.php?newsitem=411#news411

[6] https://www.ukm.my/ipi/

[7] https://www.facebook.com/ipiukm and https://www.facebook.com/UKMAtmosphere

and draft agenda, was sent to all contacts (23rd April).

- A reminder email was sent to all stakeholders stating that those who registered for the workshop would receive the meeting link (13th May).
- The Zoom link and online calendar invitations were sent to those registered for the meeting (15th May).
- A slightly edited version of the workshop slide pack was circulated to all stakeholders, including those who did not attend the project workshop (27th May).

In addition to these direct communications, the project team's FCDO contact, Jeremie Credo, further disseminated project materials to her FCDO contacts, including the ASEAN COSTI members.

Following the initial stakeholder communication, FCDO science and technology representatives based at the British Embassy in Phnom Penh, Cambodia made contact and asked for further information about the project. UKM representatives hosted a follow-up meeting to answer their specific questions.

Communication with project stakeholders via online Stakeholder workshop

The workshop took place on Thursday 16th May between 15:00 and 17:00 Malaysian time, which corresponds to 14:00 and 16:00 Vietnamese time, and 08:00 and 10:00 British summer time. The information conveyed in the presentations was generally easy to understand. For example, when showcasing the project results, the team presented a visualisation of an air pollution episode developing over a number of days. The speakers also provided strong motivation for the development of air quality forecasting systems.

Over 75 people attended the workshop. The large attendance at the workshop by both ASEAN and non-ASEAN participants can be explained in part by the posting of both the registration and meeting links online. A number of people attended the workshop who had not registered. In addition, those people who registered via the Google Drive link appeared on the Zoom participants list with the host's name, rather than the participant name. Participants were encouraged to correct their names. At the end of the workshop there remained approximately 10 people with the host's name, so the project team were unable to make a record of these participants.

The audience asked some technical questions via the meeting chat and in the discussion at the end, specifically, relating to:

- Emissions inventories
- Biomass burning
- Air quality measurement networks
- The biggest challenges when developing an air pollution forecasting system
- System calibration using measurements

In addition, many attendees requested copies of the slides. The project team considered the workshop a success due to the high attendance and moderate engagement of the audience. The workshop slides, with some minor alterations, were circulated to all stakeholders after the workshop.

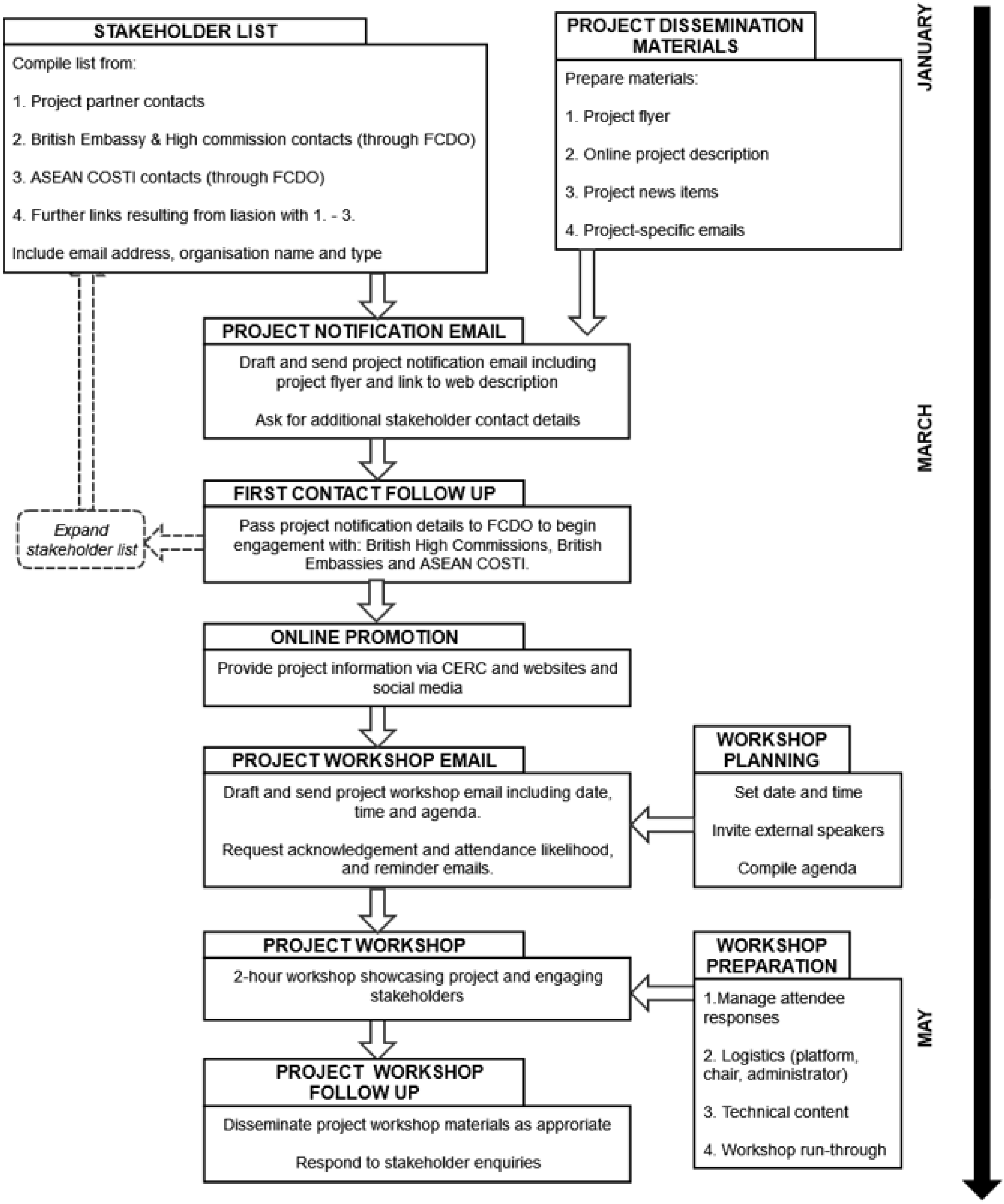


Figure 5.1 – Flow chart outlining the approach to stakeholder engagement, including project workshop.

6 Lessons learnt and project impacts

This section provides details of lessons learnt from the project, in addition to describing project impacts on the air quality and wider community in ASEAN.

6.1 Regional collaboration and study domains

A significant impact of the DHAQ-SEA project is the selection of study domains that include neighbouring countries, promoting a regional approach to air quality management. Air pollution is not confined by national borders; thus, regional cooperation is vital. By including neighbouring countries in the study domains, the DHAQ-SEA project acknowledges and addresses the transboundary nature of air pollution.

This regional approach is exemplified by the inclusion of cities such as Ho Chi Minh City in Vietnam and Kuala Lumpur in Malaysia within the study domains. This selection facilitates the examination of pollution patterns and the movement of pollutants across borders, including district borders within a country and international borders. The data and insights gained from these comprehensive studies are invaluable for developing coordinated policies and actions, both between districts within a single country and among ASEAN countries.

The stakeholder workshop, conducted as part of the DHAQ-SEA project, further emphasised this collaborative spirit. The workshop brought together representatives from various ASEAN countries, facilitating an opportunity for them to share knowledge, experiences, and best practices. Through these interactions, stakeholders could see first-hand how air quality modelling and forecasting could be applied to their regions, and fostered a collective effort to improve air quality across Southeast Asia. Dissemination of workshop materials to all stakeholders, including those who did not attend the event, ensures that as many people as possible are made aware of the project.

The DHAQ-SEA approach aligns with the broader goals of ASEAN to promote sustainable development and environmental protection. By focusing on both local and regional emissions and fostering regional cooperation, the project not only aided Vietnam and Malaysia in their air quality management efforts but also contributed to a more unified and effective regional strategy. This holistic approach ensures that future efforts to improve air quality are comprehensive, inclusive, and capable of addressing the shared challenges faced by ASEAN countries

6.2 Access to region-specific air pollutant emissions inventory data

The project has highlighted that good quality local air pollutant source data are needed to supplement global emissions inventories.

Malaysia

The collation of emission inventory data was found to be the main challenge for air quality modelling in Malaysia. Detailed information on emissions from industries, motor vehicles and open burning was lacking. The current Continuous Emissions Monitoring System [18] only records emissions data from power plants and palm oil mills, and access to this database is restricted. In terms of traffic, Kuala Lumpur City Hall's (DBKL) Integrated Transport

Information System (ITIS) was identified as a source of hourly varying traffic data, but because the data was not being used for research purposes, the process to obtain the data took a relatively long time and, consequently, could not be incorporated in the modelling within the constrained project timescale.

The main challenge when obtaining pollution source data in Malaysia is identifying the correct government agency and the specific department within the agency that can supply the necessary dataset. Then, the evaluation and assessment process must be passed from the department head to the unit head. This process is slow, and requires multiple follow-ups due to the vast number of services and additional duties provided by these departments. Once a specific person has been assigned to be in charge of a particular data request, there are additional challenges when transferring the data between departments.

The project team has learnt that it is necessary to allow for these data collation issues in future projects, and these issues are likely to apply to other ASEAN countries. Additional time to gain access to the required datasets must be assigned within the project timeline to ensure timely project completion. In addition, the procedure and experience has allowed the project team to identify shortcomings in the data management protocol at the agencies. The team can therefore provide feedback to the relevant agencies to encourage them to improve their data management procedures.

Vietnam

The Vietnamese group had quite recently generated an emissions inventory for an area including Ho Chi Minh City prior to this project, with a description of the work published in the literature [19]. This data was useful for the current project. However, there were difficulties accessing the underlying source data because the work was done by a researcher who had subsequently left the group's department.

This issue highlights the need to develop and support online platforms that can be used to hold data that may be of use for subsequent research and analysis. This applies to pollutant measurements, source activity data and source emissions data. In the UK and Europe there are many useful open data resources, for example the UK's National Atmospheric Emissions Inventory [20] and the European Environment Agency's database of air pollutant measurements [21]. However, whilst access to many datasets is open in Europe, some datasets remain inaccessible unless time-consuming negotiations with the relevant organisations are undertaken.

Pan-ASEAN

At the stakeholder workshop, the project team highlighted the opportunity for pan-ASEAN cooperation to generate a high-quality emissions inventory in order to improve the representation of air pollution in this region by global, regional and local models. This could involve a revision to the Regional Emission inventory in ASia (REAS) database [11], which was last released in 2020, with 2015 the last year of data available. For some source sectors, including traffic, the number of sources and their technologies may have changed significantly since 2015, so it is timely for this activity to be taken forward.

6.3 Coupled regional-to-local modelling approach

One of the primary impacts of the DHAQ-SEA project is the development of an air quality system that integrates both local and regional emissions. Air quality in Vietnam and Malaysia, particularly in highly populated urban areas such as Ho Chi Minh City and Kuala Lumpur, is influenced by a complex interplay of local sources such as traffic, industry, and domestic activities, as well as regional sources, including transboundary pollution. The DHAQ-SEA project has demonstrated the application of methodologies that quantify and model these diverse sources, offering a more accurate representation of air quality dynamics than previous modelling approaches.

By utilising open data sources, including emissions, concentration, meteorology, and road network datasets, the project has established a scalable and transferable modelling system. This system enables Vietnamese and Malaysian authorities to understand better the relative contributions of various emission sources, both within and outside the country's borders. The integration of high-resolution data, supported by local inputs where available, enhances the precision of these models, providing actionable insights for policymakers.

6.4 Global model (CAMS) performance

Prior to this project, CERC had extensive experience of using the CAMS air pollutant concentration forecast service within Europe, but less familiarity with the global equivalent. Working with the CAMS global forecast data product for domains in ASEAN has highlighted some limitations of the dataset. During the current project, the team identified a document [22] that summarised CAMS global forecasting system performance evaluation. Figure 6.1 presents a graphic from this report, comparing satellite measurements of NO_2 to the forecast, for approximately 6 years up to May 2022. Although CAMS model performance is improving, the system still significantly over-predicts NO_2 concentrations during the period shown.

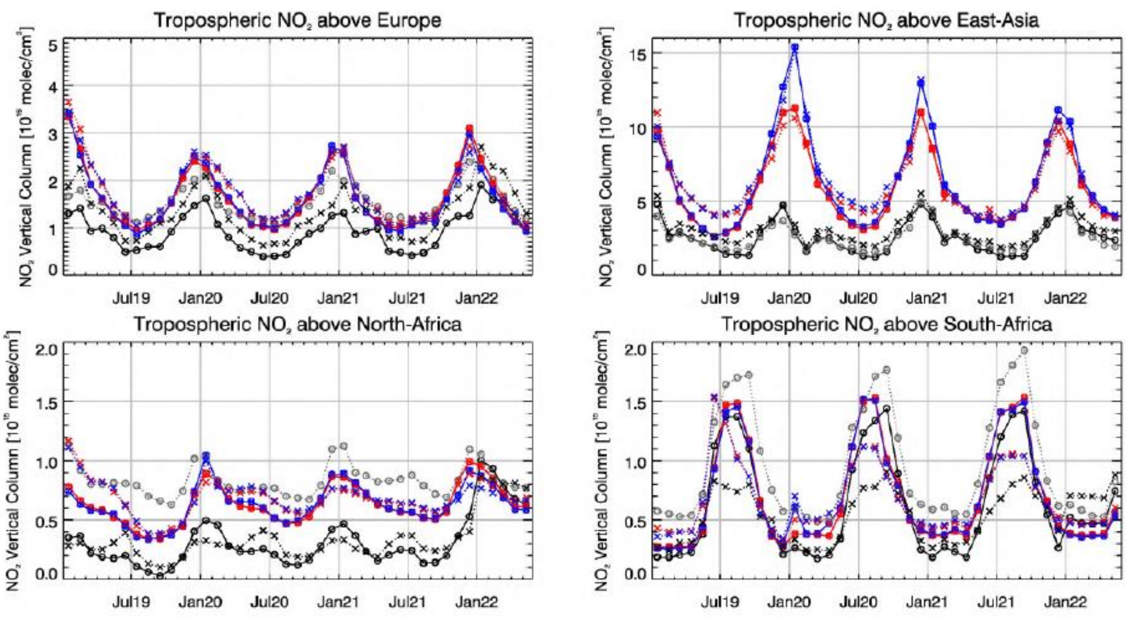


Figure S.3: Time series of average tropospheric NO_2 columns [10^{15} molec cm^{-2}] from (black and grey) satellite retrievals, (blue) control and (red) o-suite model results since Jan 2019 (see Annex 2 for definition of regions). The solid lines with circles show comparisons based on TROPOMI (in black the IUP-Bremen product and in grey the operational offline product), the dotted lines with crosses show comparisons for GOME-2C. The upper panels represent regions dominated by anthropogenic emissions, the panels in the middle represent those dominated by biomass burning.

Figure 6.1 – CAMS global forecast evaluation of NO_2 in comparison with satellite measurements for different global regions, taken from [22]

The overestimate of observed concentrations within Kuala Lumpur by the raw CAMS modelled concentrations suggests that the corresponding gridded emissions within Kuala Lumpur are overestimated. This overestimate in gridded emissions leads to an overestimate in the top-down calculations of explicit road emissions and hence an excessive concentration increment from the explicitly modelled roads in the coupled system.

The CAMS system is regularly updated. However, while improvements to the underlying model can be made by the CAMS community, they may not be able to update some of the system inputs without improved data from the global community e.g. emissions. An improved pan-ASEAN emissions inventory, such as an update to REAS, as described in the preceding section, could be incorporated into the global emissions inventories that drive CAMS, thus improving CAMS model performance in terms of air pollutant predictions across the ASEAN region.

6.5 Stakeholder engagement workshop

On 16th May 2024, the DHAQ-SEA Stakeholder Workshop showcased high-resolution air quality forecasting in Southeast Asia through a virtual event hosted on the Zoom platform from 15:00 to 17:00 Malaysian time (GMT +8). This workshop aimed to demonstrate the feasibility of developing a high-resolution air quality forecasting system for ASEAN regions, including Malaysia and Vietnam, including street-scale effects. Attendees included approximately 75 people from government and non-government agencies in Malaysia, Indonesia, Laos, the Philippines, Thailand, Myanmar, Vietnam, Singapore, and Brunei, as well as non-ASEAN countries such as Japan, Bangladesh, India and Pakistan.

The project team designed the workshop to ensure that the information presented was accessible to a wide audience. The team prioritised providing background information about the impacts of high air pollution, for example, in terms of poor health, to motivate the audience. Additionally, information on both local and regional contributors to air pollution were described, to help put the regional-to-local-scale air quality modelling approach into context. Figure 6.2 shows the sequence of presentations given at the workshop.

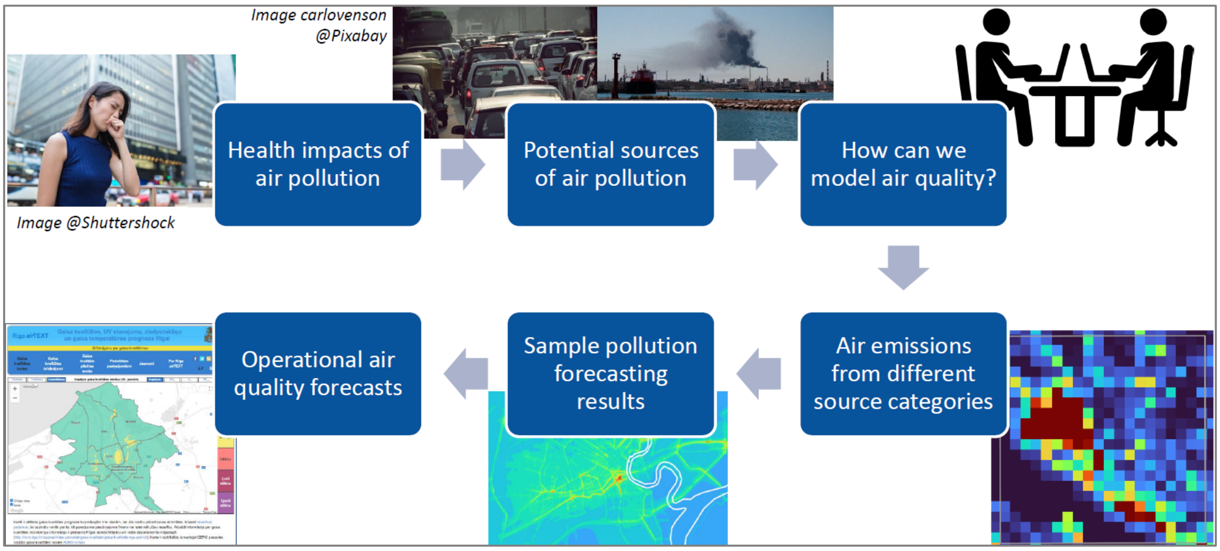


Figure 6.2 – DHAQ-SEA stakeholder workshop overview, describing the 'Why?' and 'How?' of air quality forecasting

7 Discussion

This section fulfils the specific requirements of this final project report in relation to: a discussion of study site beneficiaries (Section 7.1); implications for the wider Southeast Asian region (Section 7.2); and system scalability (Section 7.3).

7.1 Study site beneficiaries

This project focuses on 'Health', which is one of the RIDA priority areas; specifically, poor air quality impacts on people's health. It helps people to know when and where air pollution levels are likely to increase, as they can take appropriate action to minimise their exposure (as well as reduce their emissions). Thus, air pollution forecasts with associated actionable health advice contribute to improved health, addressing SDG 3 'Ensure healthy lives and promote well-being for all at all ages.' As highlighted in Figure 7.1, taken from a report by the Climate and Clean Air Coalition, air quality also has strong links to other SDGs, specifically: 7 'Ensure access to affordable, reliable, sustainable and modern energy for all'; 11 'Make cities and human settlements inclusive, safe, resilient and sustainable'; and 13 'Take urgent action to combat climate change and its impacts'.

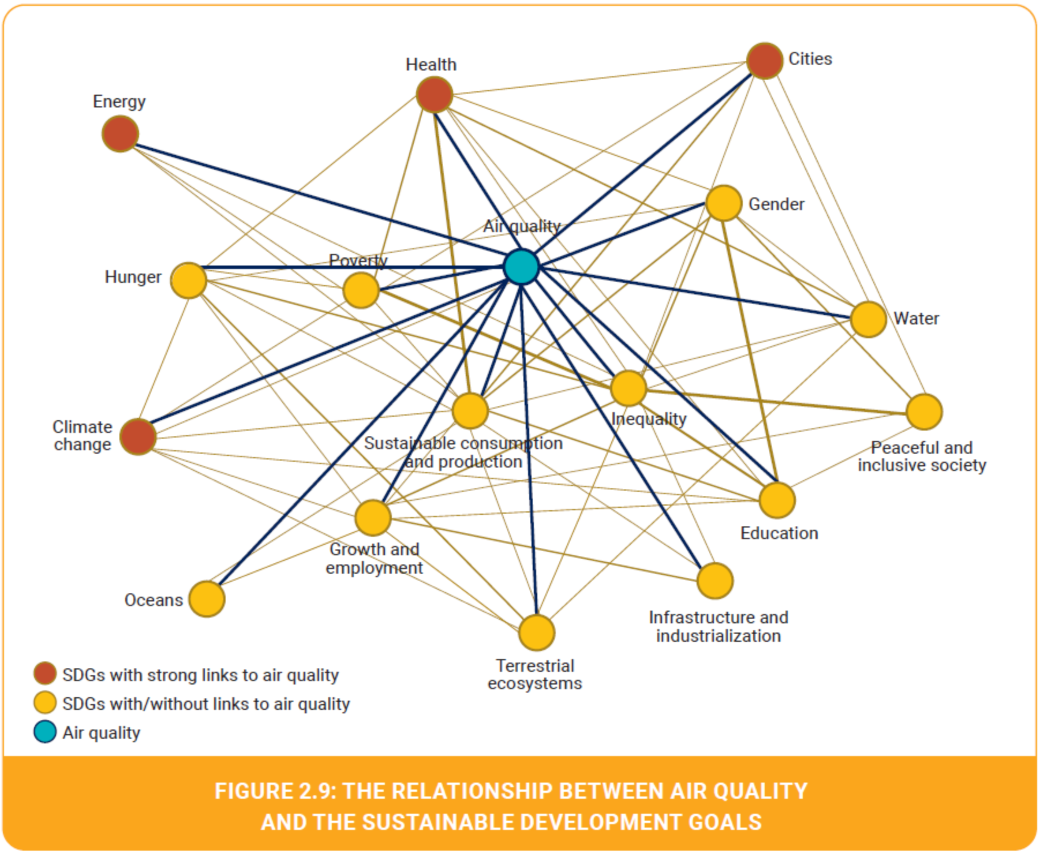


Figure 7.1 – Relationship between air quality and the sustainable development goals (SDGs) (taken from [24])

Pollution maps corresponding to air pollution episodes have been presented to the targeted beneficiaries (i.e. project stakeholders in ASEAN) during the workshop. The potential benefit of high-resolution air quality forecasting systems has been demonstrated in terms of providing the public with detailed information on the spatial and temporal distribution of air pollutants. The methodology for developing the pollution maps was explained at the workshop, including a description of how the information in the maps was evaluated in comparison to the air quality measurements recorded in the respective regions, thus providing context to the modelling work.

Details of how the modelling work can be extended to form part of an operational air quality forecasting system, with associated air pollution alerts, were described at the workshop by Amy Stidworthy from CERC in her 'Operational Air Quality Forecasts' presentation. Amy provided details of similar air quality forecasting services currently provided by CERC and also systems run by other organisations employing CERC software.

To the best of the project team's knowledge, air pollution modelling at street-scale resolution has not previously been undertaken for domains as large as those covered in this project in the Southeast Asian region. Therefore, the Vietnamese and Malaysian stakeholders may be interested in historical, as well as forecast, applications of the system. For example, historical air pollution maps can be used to perform detailed population exposure calculations, which can feed into economic calculations to quantify the costs of air pollution. Alternatively, forecast pollution map archives can provide data for this purpose, although forecast datasets are not usually as accurate as hindcasts because the latter include some measurement data assimilation.

7.2 Implications for the wider Southeast Asian region

All ASEAN countries may be subject to poor air quality as a consequence of transboundary pollution; local pollutant sources and meteorological conditions also influence air pollutant concentrations. ASEAN has an agreement on transboundary haze pollution [23]. Although the air quality in ASEAN is better than in some other countries in the wider Asia-Pacific region, many people are exposed to relatively high concentrations, particularly in the more densely populated urban areas; this has negative mortality and economic implications (shown in Figures 7.2 and 7.3 respectively, taken from [24]). It is clear from Figure 7.2 that air pollution from both household and external sources impacts mortality in ASEAN countries with relatively low GDP per capita, whereas in higher income ASEAN countries (Singapore, Malaysia and Brunei Darussalam), domestic solid fuel burning has been phased out in urban areas.

The level of local pollutant emissions can strongly influence air pollution levels. Figure 7.4 presents a review of local NO_X emissions levels in cities throughout ASEAN. The insets show emission rates for selected major cities: Kuala Lumpur (Malaysia), Ho Chi Minh City (Vietnam), Manila (Philippines) and Jakarta (Indonesia). It is clear that the emissions levels in the different cities are comparable. Therefore, by demonstrating that air quality modelling and forecasting is beneficial at the Vietnamese and Malaysian study sites, it follows that this system would be similarly useful in other areas in ASEAN,

During the stakeholder workshop, key representatives from other ASEAN member countries were shown how this type of modelling can be applied to their region.

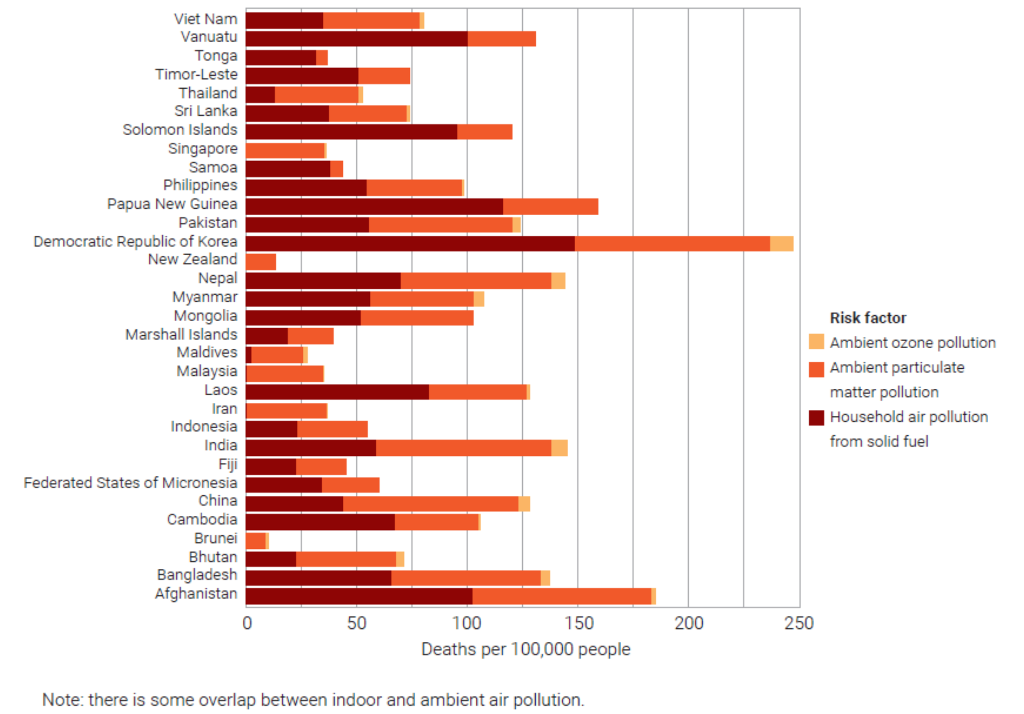


FIGURE 1.5: MORTALITY ASSOCIATED WITH AIR POLLUTION IN THE ASIA AND PACIFIC REGION, BY COUNTRY, 2016

Figure 7.2 – Mortality associated with air pollution in the Asia and Pacific region, by country, 2016 (taken from [24])

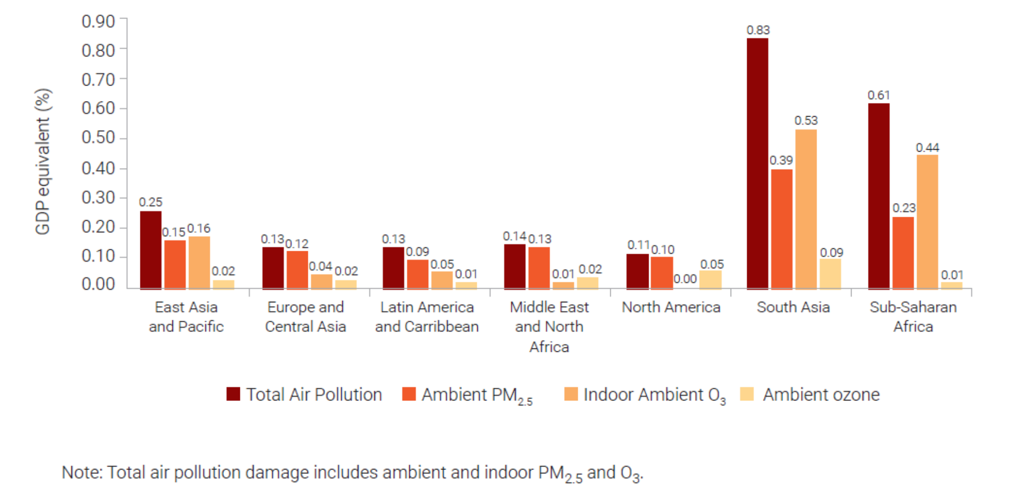


FIGURE 1.24: LABOUR OUTPUT LOST DUE TO AIR POLLUTION, BY REGION, 2013

Figure 7.3 – Labour output lost due to air pollution, by region, 2013 (taken from [24])

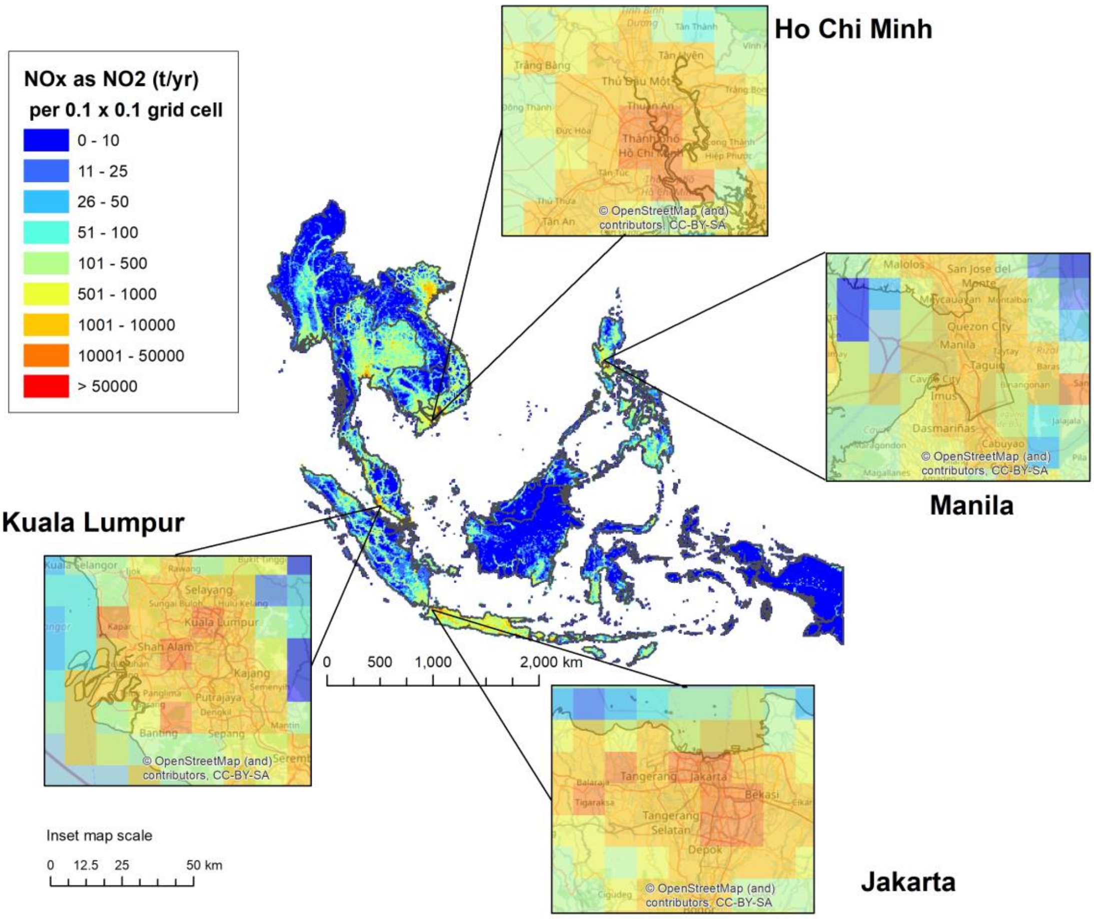


Figure 7.4 – A comparison of local emissions (NO_X, tonnes per year); data from the CAMS global emissions inventory [9]; grid cell size approximately 11 km × 11 km. Emissions from shipping lanes have been omitted from the image for clarity.

7.3 System scalability

One of the project aims was 'To demonstrate that open data allows transferability and scalability'. The project team have used open data sources where possible. These open emissions, concentration, meteorology and road network datasets are described in Appendix A. The use of open data allows transferability and scalability, as since these datasets are available across the entire ASEAN region, the modelling system could also be set up with equivalent inputs for other countries.

The accuracy of modelled concentrations relies heavily on the accuracy and resolution of model inputs. The available open emissions datasets are coarse resolution (approximately 11 km by 11 km). The project team have methods to increase the resolution of these emissions by using spatial proxies e.g. road network data for emissions from road traffic. However, local data should supplement these open data sources if available. For example, the Vietnamese project partners shared Ho Chi Minh City emissions data. If this modelling system were to be applied to other domains in the target countries, or elsewhere in ASEAN, local emissions may not be available. Less local model input data were available for the Malaysian domain compared to the Vietnamese domain, but the availability of good-quality reference measurements allowed calibration of model inputs. It is likely that a calibration component would need to be integrated within operational forecasting systems for countries where local emissions data are unavailable.

It is important to evaluate model performance using measured datasets. This project has highlighted that measurement data availability is sparse in some areas, and also that there is variability in terms of the quality of measurement data. Data quality relates in part to the high costs associated with maintaining reference standard measurement instruments; lower-cost measurement equipment also requires maintenance and calibration. Confidence in modelled concentrations is lower in areas where no measurements are available, and also for regions where there is a large uncertainty in measured values.

8 Limitations

The main limitations of this feasibility study result from difficulties with input data for the coupled modelling system.

The development of a local-to-regional air quality forecasting system requires access to forecast input regional-scale concentrations and meteorological data. Forecast concentration and meteorological data are available from the CAMS global service, but at relatively low spatial and temporal resolution (~40 km grid scale and 3-hourly data frequency for most variables). The quality of global-scale air quality datasets over ASEAN is reduced due to the limited public availability of input data in this region, such as routine concentration measurements or road traffic vehicle fleet composition and fuel characteristics.

The use of global CAMS and EDGAR emissions inventories at ~11 km resolution does not capture the finer scale variation of concentrations due to industrial, agricultural, domestic combustion or shipping activities. CAMS concentrations incorporate near-real-time estimates of biomass burning emissions from the GFAS inventory [25], however this data is only available at daily temporal and ~11 km spatial resolution. The emissions resolution for non-road transport emissions could potentially be improved by using additional spatial proxy datasets, such as industrial land use areas for industrial emissions or population density for domestic burning, to redistribute the gridded emissions to smaller geographical areas. Novel data sources, such as Automatic Identification System (AIS) data on shipping movements, GPS vehicle tracking or mobile phone data, may also be harnessed for generating higher-resolution emissions data.

The modelling carried out in this project has identified that some of the global emissions datasets may have additional inaccuracies in the ASEAN region due to slow updating. For example, the CAMS SO_2 emissions in the Malaysian domain resulted in very high modelled concentrations compared to measurements. This suggests that the emissions may not capture regulatory improvements in fuel sulphur content which have reduced emission rates in recent years. When top-down emissions calculations are used to derive source-specific emissions from gridded data, for example for explicit road emissions, inaccuracies in the gridded emissions lead to reduced accuracy of the explicit emissions and hence of the street-scale modelling.

Emissions inventories are generally published at annual or monthly resolution, with standardised profiles used to calculate temporal variations at daily and hourly resolution. Limited data availability in the ASEAN region leads to the use of profiles from European data sources which may not be appropriate, particularly for domestic combustion activities. The use of generalised temporal profiles also limits the accuracy of the forecasting system for episodes of increased air pollution related to short-term local emissions, e.g. construction activity or accidental fires.

Features of the urban built environment such as street canyons, flyovers or road tunnels can alter the dispersion of emissions from road traffic. The ADMS-Urban dispersion model can calculate the influence of these features but requires additional input data in order to identify the key characteristics of each emissions source. No public datasets representing these urban features specific to the modelling domains were identified during the project. The highly constrained project timescale did not allow the processing of global datasets such as Local

Climate Zone maps [26], nor the collection of local information e.g. through satellite imagery or site visits.

The evaluation of air quality modelling outputs relies on concentration measurements. Continuous concentration monitors require regular maintenance and calibration in order to produce reliable data. When different parts of a monitoring network are operated by multiple organisations, it is important to standardise maintenance and calibration procedures in order to avoid variations in data quality across the network. It is also important to identify monitoring site locations precisely, in particular the height of the air intake and the distance from the nearest major road.

9 Next steps

There are a number of possible ways forward following the completion of this feasibility study. Firstly, it is important to remain engaged with the project stakeholders, with the aim of identifying funding avenues for future work (Section 9.1). In terms of technical activities, the base air quality model configurations for the Vietnamese/Cambodian and Malaysian domains could be converted into operational air quality forecasting systems (Section 9.2). Additionally, baseline modelling and operational air quality forecasting systems could be set up for other parts of ASEAN (Section 9.3).

9.1 Continued stakeholder engagement and commercial plan

Consideration of a route-to-market strategy was one of the aims of the project. The project's commercial success in terms of follow-on funded activities is most likely when a large number of stakeholders are identified and engaged with the project. The team had a strong focus on stakeholder engagement throughout the project; lack of stakeholder engagement was considered one of the highest risks to project success.

The project stakeholder workshop provided the project team with an opportunity to showcase:

- The modelling system's capability to generate data that can be used to warn the public about air pollution episodes; and
- The project team's ability to deliver operational air pollution forecasting systems.

Previous sections in this report have described the successful delivery of the stakeholder workshop to a wide-reaching audience, including representatives from a diverse range of organisations from many ASEAN and non-ASEAN countries. The majority of attendees were from academia, but a number of government organisations also attended. The stakeholder workshop presentations concluded with 'Key outcomes' and 'Future potential' slides (Figure 9.1). These slides gave the key points relevant to the development of similar air quality modelling and forecasting systems, as well as providing team contact details. Workshop slides were circulated to all stakeholders, including those that were not able to attend the workshop. The project team believe that the relevant stakeholders now feel confident about investing in, or identifying funding for, such air quality forecasting systems, thus providing the project team with opportunities for further collaborative projects.

The route-to-market of ASEAN air quality forecasting systems requires strategic partnerships between commercial service providers (i.e. CERC), local organisations (e.g. academic partners with appropriate expertise), and government departments or other funding bodies. DHAQ-SEA is the third project in which CERC have collaborated with Professor Talib's team at UKM. For another project, CERC are delivering an operational air quality forecasting system for the Klang District in Selangor, Malaysia. For that project, a team within the University of Malaya are developing an app to disseminate air quality alert information. This team could potentially be brought together to provide a similar service as part of other future projects. CERC and IER/VNU-HCM have not collaborated prior to this project. During DHAQ-SEA, IER/VNU-HCM have demonstrated their emissions expertise. Generating sufficiently accurate emissions data is a key part of any follow-on ASEAN air quality modelling project; IER/VNU-HCM's emissions knowledge and skills will be highly beneficial for developing local emissions data for future projects.

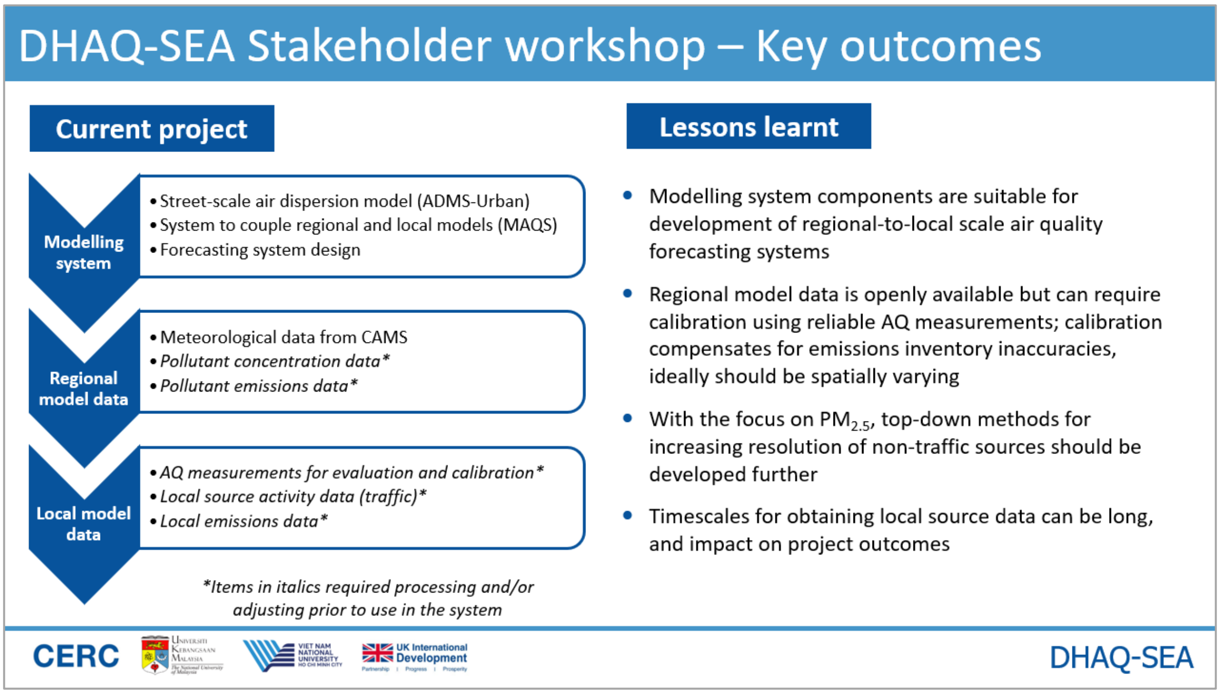


Figure 9.1 – 'Key outcomes' and 'Future potential' slides from the DHAQ-SEA stakeholder workshop

9.2 Regional-to-local-scale air quality forecasting system design

This feasibility project has demonstrated how air quality forecasting systems for the selected regions could be provided. The delivery of an operational system was out of scope for this project; however, the project team have extensive experience of developing dedicated operational air quality forecasting software and systems in other regions. Examples include airTEXT for London [27], Manchester [28], York (UK) [29] and Riga [30] (Latvia), as well as an air quality and temperature forecasting system for Kuala Lumpur [31] (Malaysia). Other organisations also licence CERC software systems to run their own air quality forecasting services, e.g. Barcelona [32] (Spain) have run such a system for over 10 years.

The Hong Kong University of Science and Technology (HKUST) run a regional-to-street-scale air quality forecasting system for Hong Kong (China) using CERC software. PRAISE-HK [33] is conceptually similar to the air quality forecasting system proposed here for ASEAN, but to date, CERC have not implemented this type of coupled forecasting system; CERC's current air quality forecasting services are for smaller domains, so do not make use of spatially varying forecast meteorological and concentration data.

The project team have considered the design of this type of regional-to-local air quality forecasting system. A system design diagram is presented in Figure 9.2. Key features of this system are:

- Use of regional-scale forecast meteorological and concentration data from CAMS[8];
- Efficient model domain scalability through the use of Cloud computing and spatial parallelisation of model runs; and
- Flexibility in terms of the generation of air quality outputs e.g. a variety of options for pollutants modelled, metrics calculated, and formats of pollution maps generated.

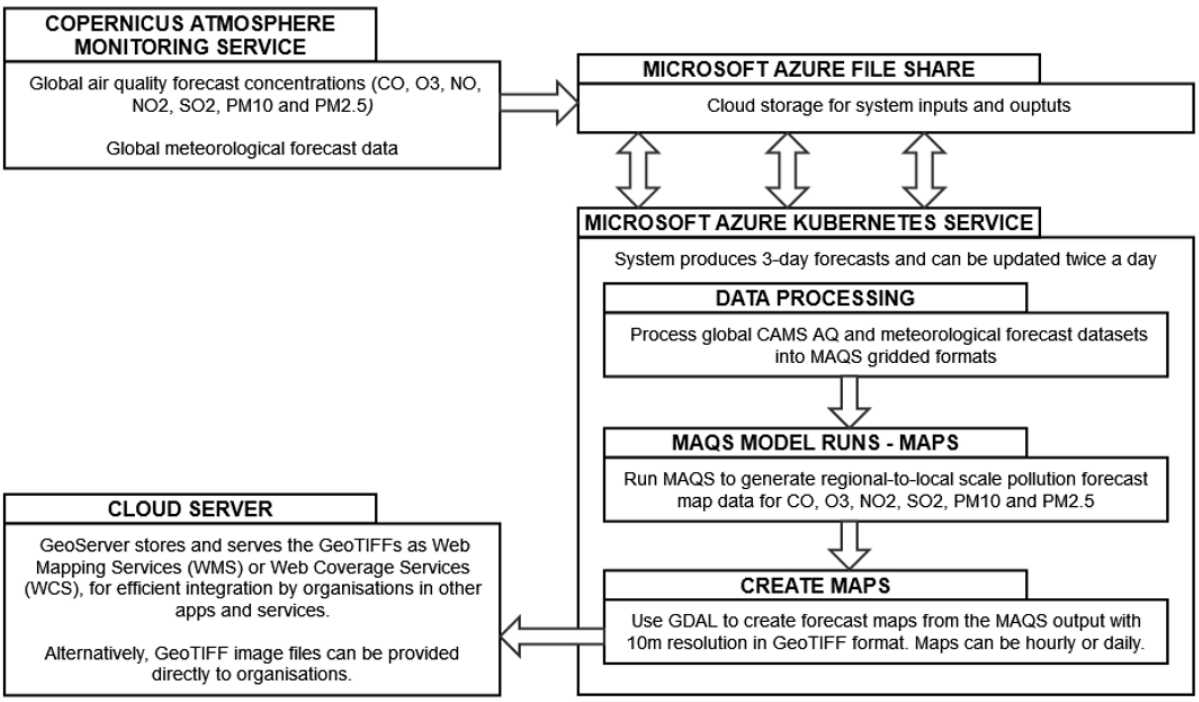


Figure 9.2 – Regional-to-local-scale forecasting system diagram

The best way to reach large proportions of the population is to provide air quality forecasts through mobile apps and/or social media; social media usage in some ASEAN countries is

---

[8] CERC have been partners in the CAMS user support / interaction strand for 15 years: https://atmosphere.copernicus.eu/user-interaction-activities

particularly high (for example an estimated 65.7% and 71.4% of the population have Facebook accounts in Vietnam and Malaysia respectively [34]). Behavioural experts have highlighted the importance of engaging positively with the public, and sharing air quality information in a way that provides actionable information [35]. The way in which this information should be shared will vary between countries and regions. As currently proposed, this system design includes the provision of air quality data via web APIs (Application Programming Interfaces), e.g. Web Mapping Services (WMS) and Web Coverage Services (WCS). Web APIs facilitate the retrieval of the required forecast datasets, allowing organisations to provide necessary information to the public in the most accessible way. CERC also have extensive experience in providing air quality alerting services and hosting websites, if this were required.

9.3 System extension to other ASEAN regions

The DHAQ-SEA project has demonstrated the application of the coupled regional-to-local-scale air quality modelling system (MAQS), which incorporates the ADMS-Urban street-scale air dispersion model. The system relies on both regional and local input data sources. For these system applications, openly available regional model data have been sourced from CAMS, specifically coarse-resolution meteorological data, pollutant concentration data, and pollutant emissions data. All these datasets require processing prior to input to the modelling system to a greater or lesser extent. Local air quality measurements are needed for evaluation and calibration. In addition, where available, local source activity data (e.g. traffic flow) and other emissions data can significantly enhance the accuracy of modelled concentrations.

A design for the extension of this modelling system to an operational forecast has been proposed. Air quality forecasting systems commonly use forecast meteorological and pollutant concentration data; CAMS makes these products publicly available. The emissions used within air quality forecasting systems are usually generated from historical data, although there can be opportunities associated with linking to real-time road traffic and/or shipping data.

The modelling system used for this project can be applied anywhere, if suitable input data are available. The CAMS regional datasets that have been used as input to the modelling undertaken for this project, and those required for an operational forecasting system, are openly available for the whole of ASEAN. The project team have, however, learnt that the accuracy of these global atmospheric products is not as high in Southeast Asia compared to other regions where the underlying emissions inventories are better quality e.g. in Europe. For ASEAN, calibration methods may have to be applied to system inputs to improve forecast accuracy. A simplified approach to calibrating regional transport system inputs has been applied for the Malaysian modelling domain presented here, which in general improves model outcomes. A more sophisticated approach would have to be implemented for an operational system, for instance, including spatially varying adjustments of regional concentrations, perhaps calculated using machine learning algorithms. All calibration methods require accurate local air quality measurement data; air quality measurements are also required for evaluating forecast outputs.

Local emissions data should be used to verify and enhance emissions derived from global estimates. For this project, the team have demonstrated this approach using the Ho Chi Minh City region emissions data shared by IER/VNU-HCM. Local knowledge is also invaluable, for example in terms of identifying the dominant emissions sources. A key recommendation from this project is for the ASEAN community to take steps towards updating their REAS emissions inventory.

This project has focused on $PM_{2.5}$ modelling, because this pollutant significantly impacts on health in ASEAN. The project team have learnt that to apply high-resolution modelling techniques for this pollutant, methods to increase resolution of non-transport pollution sources are required.

If ASEAN organisations were interested in going forward with developing similar regional-to-local-scale forecasting systems for their regions, the steps required are:

1. Define the spatial domain of interest and required forecast outputs
2. Assess the availability of local and global data for this domain
3. Process model inputs into suitable formats for use in the system
4. Configure the system for an historic year; perform evaluation to assess performance; develop calibration approaches, if required
5. Generate historic pollution maps
6. Extend to an operational forecasting system

10 References

1. Ambient (outdoor) air pollution, 2022. World Health Organisation factsheet. Accessed March 2024.
2. Hien, T.T., Chi, N.D.T., Nguyen, N.T., Takenaka, N. and Huy, D.H., 2019. Current status of fine particulate matter (PM2. 5) in Vietnam's most populous city, Ho Chi Minh City. *Aerosol and Air Quality Research*, 19(10), pp.2239-2251. Accessed March 2024.
3. Data provided in Private Communication with Vietnam National University, 2023
4. Environmental Quality Report 2020. Malaysian Ministry of Environment and Water. Accessed March 2024.
5. Sillmann, J., Aunan, K., Emberson, L., Büker, P., Van Oort, B., O'Neill, C., Otero, N., Pandey, D. and Brisebois, A., 2021. Combined impacts of climate and air pollution on human health and agricultural productivity. Environmental Research Letters, 16 (9), 093004.
6. What to do when there is an air pollution alert, 2019. World Health Organisation factsheet. Accessed March 2024.
7. CAMS Reanalysis. European Centre for Medium-Range Weather Forecasting. Accessed March 2024.
8. CAMS global atmospheric composition forecasts, 2021. Copernicus. Accessed March 2024.
9. CAMS global emission inventories, 2020. Copernicus. Accessed March 2024.
10. EDGAR – Emissions database for global atmospheric research. European Commission Joint Research Centre. Accessed March 2024.
11. Regional Emission Inventory in Asia (REAS), 2021. National Institute for Environmental Studies, Japan. Accessed March 2024.
12. OpenStreetMap for Malaysia and Vietnam. Accessed March 2024.
13. Health White Paper for Malaysia, 2023. Ministry of Health, Malaysia. Accessed March 2024.
14. MAQS – Multi-Model Air Quality System. CERC. Accessed March 2024.
15. Regional to local scale air quality modelling. CERC. Accessed March 2024.
16. Air quality forecasting. CERC. Accessed March 2024.
17. U.S. Embassy and Consulate AirNow Department of State air quality monitoring
18. Continuous Release Monitoring System, Department of Environment Ministry of Natural Resources And Environmental Sustainability. Accessed June 2024.
19. Ho, Q.B., Vu, H.N.K., Nguyen, T.T. and Nguyen, T.T.H., 2020. Traffic air emission inventory and measures to reduce air pollution in Ho Chi Minh City, Vietnam. *Journal of Urban Environment*, 38, pp.29-38. Accessed March 2024.
20. National Atmospheric Emissions Inventory, Department for Business, Energy & Industrial Strategy. Accessed June 2024.
21. European Environment Agency Air Quality download service. Accessed June 2024.

22. Validation report of the CAMS near-real-time global atmospheric composition service. Period March – May 2022. Ref: CAMS2_82_2022SC1_D82.1.1.3-MAM2022. Issued by: KNMI September 2022.
23. ASEAN Agreement on Transboundary Haze Pollution, 2002. ASEAN. Accessed March 2024.
24. Air Pollution in Asia and the Pacific - Science-based solutions 2019. Climate and Clean Air Coalition. Accessed March 2024.
25. Kaiser, J.W., Heil, A., Andreae, M.O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.J., Razinger, M., Schultz, M.G., Suttie, M. and Van Der Werf, G.R., 2012. Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. Biogeosciences, 9(1), pp.527-554.
26. Demuzere, M., Kittner, J., Martilli, A., Mills, G., Moede, C., Stewart, I.D., Van Vliet, J. and Bechtel, B., 2022. A global map of Local Climate Zones to support earth system modelling and urban scale environmental science. *Earth System Science Data Discussions*, *2022*, pp.1-57.
27. airTEXT Air quality, UV, pollen and temperature forecasts for Greater London and the South East, 2024. CERC. Accessed March 2024.
28. Clean Air Greater Manchester. Accessed March 2024.
29. City of York Council Air quality forecasts and alerts. Accessed June 2024.
30. Riga airTEXT, 2024. CERC. Accessed March 2024.
31. Disaster Resilient Cities - Forecasting Local Level Climate Extremes and Physical Hazards for Kuala Lumpur, 2024. Disaster Resilient Cities. Accessed March 2024.
32. Air Quality in Barcelona. Ajuntament de Barcelona. Accessed March 2024.
33. PRAISE-HK. Hong Kong University of Science and Technology. Accessed March 2024.
34. Facebook usage penetration statistics from Statistica.com for Vietnam and Malaysia. Accessed March 2024.
35. Riley, R., de Preux, L., Capella, P., Mejia, C., Kajikawa, Y. and de Nazelle, A. (2021). How do we effectively communicate air pollution to change public attitudes and behaviours? A review. Sustainability Science, pp.1-21.
36. Emissions of atmospheric Compounds and Compilation of Ancillary Data (ECCAD). Accessed January - March 2024.
37. Global Forecast System, GFS. Accessed March 2024.
38. Update to accessibility of ECMWF data via CAMS: IFS cycle 48r1. Accessed March 2024.
39. ECMWF Integrated Forecasting System, IFS. Accessed March 2024.
40. Emissions Inventory Toolkit EMIT. CERC. Accessed March 2024.
41. WorldPop population density data. Accessed March 2024.
42. Tran Thu Trang, Huynh Hai Van & Nguyen Thi Kim Oanh (2015) Traffic emission inventory for estimation of air quality and climate co-benefits of faster vehicle technology intrusion in Hanoi, Vietnam, Carbon Management, 6:3-4, 117-128, DOI:10.1080/17583004.2015.1093694

43. Huu, D.N.; Ngoc, V.N. Analysis Study of Current Transportation Status in Vietnam's Urban Traffic and the Transition to Electric Two-Wheelers Mobility. Sustainability 2021, 13, 5577. https://doi.org/10.3390/su13105577

44. Model Evaluation Toolkit. CERC. Accessed March 2024.

45. Malaysia Transportation Statistics 2022, 2023. Ministry of Transport Malaysia. Accessed March 2024.

11 Data acknowledgements

ECCAD (Emissions of atmospheric Compounds and Compilation of Ancillary Data, [36]) is the Global Emission InitiAtive (GEIA) data portal and is part of AERIS, the French data service for Atmosphere. ECCAD provides a large number of datasets at global and regional scales, at various spatial and temporal resolutions and time periods and has data download as interoperable NetCDF CF-compliant files.

For this project emissions data have been downloaded from the ECCAD site. Detailed metadata for each dataset is available on the ECCAD site, the references for the different products are given below.

- CAMS-GLOB-ANT: global dataset of anthropogenic emissions of gaseous pollutants, for a wide range of emissions sectors including road traffic, power generation, industrial activities, agriculture, domestic combustion, within-country (domestic) shipping and aviation.
    - Granier, C., S. Darras, H. Denier van der Gon, J. Doubalova, N. Elguindi, B. Galle, M. Gauss, M. Guevara, J.-P. Jalkanen, J. Kuenen, C. Liousse, B. Quack, D. Simpson, K. Sindelarova The Copernicus Atmosphere Monitoring Service global and regional emissions (April 2019 version) Report April 2019 version, doi:10.24380/d0bn-kx16, 2019
    - Soulie, A., C. Granier, S. Darras, N. Zilbermann, T. Doumbia, M. Guevara, J.-P. Jalkanen, S. Keita, C. Liousse, M. Crippa, D. Guizzardi, R. Hoesly, S. J. Smith Global Anthropogenic Emissions (CAMS-GLOB-ANT) for the Copernicus Atmosphere Monitoring Service Simulations of Air Quality Forecasts and Reanalyses Earth Syst. Sci. Data 2023
- CAMS-GLOB-SHIP: global dataset of emissions of gaseous pollutants from international shipping activities.
    - Granier, C., S. Darras, H. Denier van der Gon, J. Doubalova, N. Elguindi, B. Galle, M. Gauss, M. Guevara, J.-P. Jalkanen, J. Kuenen, C. Liousse, B. Quack, D. Simpson, K. Sindelarova The Copernicus Atmosphere Monitoring Service global and regional emissions (April 2019 version) Report April 2019 version, doi:10.24380/d0bn-kx16, 2019
    - J-P. Jalkanen, L. Johansson, and J. Kukkonen, A comprehensive inventory of ship traffic exhaust emissions in the European sea areas in 2011 Atmos. Chem. Physic. 71-84, https://doi.org/10.5194/acp-16-71-2016, 2016
    - Jalkanen, J.-P., Johansson, L., Kukkonen, J., Brink, A., Kalli, J., and Stipa, T. Extension of an assessment model of ship traffic exhaust emissions for particulate matter and carbon monoxide Atmos. Chem. Phys. 2641-2659, https://doi.org/10.5194/acp-12-2641-2012, 2012
    - Johansson, L., Jalkanen, J.-P., Kalli, J., and Kukkonen, J. The evolution of shipping emissions and the costs of regulation changes in the northern EU area Atmos. Chem. Phys. 11375–11389, https://doi.org/10.5194/acp-13-11375-2013, 2012
    - Johansson, L., Jalkanen, J.-P., and Kukkonen, J. Global assessment of shipping emissions in 2015 on a high spatial and temporal resolution Atmospheric Environment 167, 403-415, https://doi.org/10.1016/j.atmosenv.2017.08.042, 2017
- CAMS-GLOB-AIR: global dataset of gaseous pollutant emissions from international aviation.

  - Granier, C., S. Darras, H. Denier van der Gon, J. Doubalova, N. Elguindi, B. Galle, M. Gauss, M. Guevara, J.-P. Jalkanen, J. Kuenen, C. Liousse, B. Quack, D. Simpson, K. Sindelarova The Copernicus Atmosphere Monitoring Service global and regional emissions (April 2019 version) Report April 2019 version, doi:10.24380/d0bn-kx16, 2019
- CAMS-GLOB-BIO: global dataset of gaseous pollutant emissions from biogenic (predominantly vegetation) sources.
  - Sindelarova, K., Markova, J., Simpson, D., Huszar, P., Karlicky, J., Darras, S., and Granier, C. High-resolution biogenic global emission inventory for the time period 2000–2019 for air quality modelling Earth Syst. Sci. Data 14, 251-270, 2022
  - Granier, C., S. Darras, H. Denier van der Gon, J. Doubalova, N. Elguindi, B. Galle, M. Gauss, M. Guevara, J.-P. Jalkanen, J. Kuenen, C. Liousse, B. Quack, D. Simpson, K. Sindelarova The Copernicus Atmosphere Monitoring Service global and regional emissions (April 2019 version) Report April 2019 version, doi:10.24380/d0bn-kx16, 2019
- EDGAR v6: global dataset of pollutant emissions from anthropogenic activities in a wide range of sectors. In this project, EDGAR emissions have been used for particulate pollutants only.
  - Crippa, M., Oreggioni, G., Guizzardi, D., Muntean, M., Schaaf, E., Lo Vullo, E., Solazzo, E., Monforti-Ferrario, F., Olivier, J. and Vignati, E. Fossil CO2 and GHG emissions of all world countries: 2019 report Publications Office Luxembourg, 2019

Appendix A - Technical method details

Technical information about the modelling activities for the project is given in the following sections. Section A.1 introduces the two modelling domains, the associated coordinate systems, and the selection of pollution episodes. Descriptions of the datasets used in the project are given in Section A.2. These include the open datasets for forecast regional concentrations, forecast meteorology, gridded emissions, and road network geometry which were used as inputs to the modelling, along with details of the associated processing steps. Proprietary datasets shared by the project partners are summarised, followed by an assessment of available pollutant measurement data.

A.1 DOMAIN & PERIOD SPECIFICATION

Model domain and coordinate system definitions

The modelling domains used for Malaysia and Vietnam/Cambodia are shown in Figure 3.1. The standard national coordinate system for Vietnam is a Universal Transverse Mercator projection, zone 48 North, using the WGS84 ellipsoid datum. Processing using the GDAL spatial software library can access this coordinate system using the EPSG code 3405. The data files for the coupled system can specify this coordinate system using attributes which comply with the CF (climate forecast) file conventions.

The standard national coordinate system for Malaysia is an oblique Mercator projection, Hotine variant. This projection has EPSG code 3375. It is not possible to fully specify this coordinate system in the latest CF conventions. This means that caution is required when visualising data files in third party software, to ensure that the correct coordinate system is used.

Pollution episodes

The project partners from UKM and IER VNU-HCM analysed the available measurement data from their local domains in order to identify episodes of increased pollutant concentration. Figure A.1 shows examples of the time-series plots which were used to assess and present episodes of interest. The aim was to identify three episodes for each domain, ideally corresponding to different seasonal meteorological conditions and a range of pollutants. The pollutant of highest current concern in both Malaysia and Vietnam is $PM_{2.5}$. Some moderately elevated O_3 concentrations were also identified. For Vietnam, the chosen modelling periods were January, October and late December. January and December are both dry season periods, whereas October is transitional between wet and dry seasons. For Malaysia, the chosen modelling periods were January, March and late April – late May. These also cover a range of different seasonal meteorological conditions. Due to project timescale pressures, only the January period was modelled for the Vietnam/Cambodia domain while both January and late April – late May periods were modelled for the Malaysia domain.

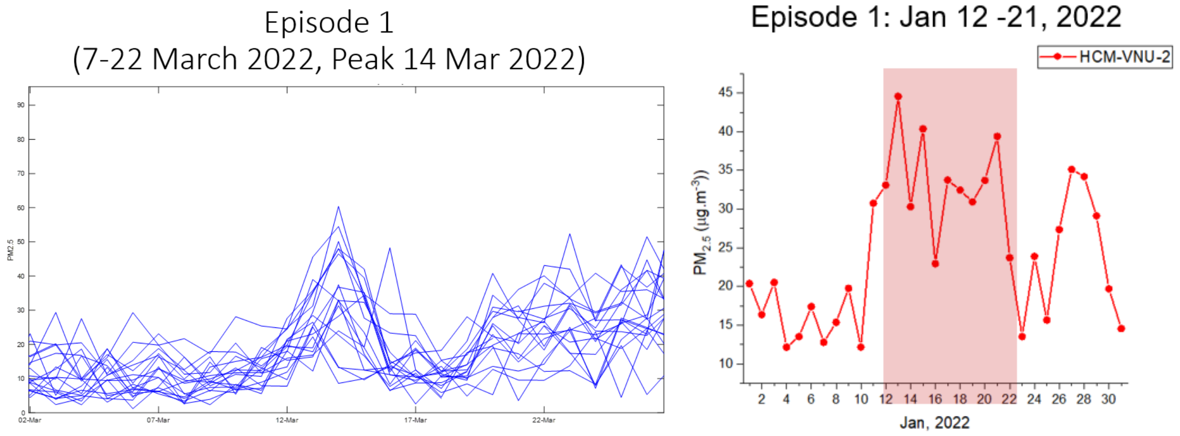


Figure A.1 – Time-series plots of measured data from Malaysia (left) and Vietnam (right), showing proposed episodes of increased $PM_{2.5}$ concentrations for modelling

A.2 DATA COLLATION

Mesoscale model meteorological datasets

The project originally planned to use regional gridded meteorological data from the GFS forecasting system [37]. CERC have previously worked with GFS meteorological data for a local model forecasting system. However, during the early stages of the project a new gridded meteorological data source became available from the global CAMS system [38]. This has the advantages of being consistent with the modelling system used for regional gridded concentrations, and having consistent file format with the gridded concentrations, allowing re-use of some processing code. Hence, CERC chose to use the CAMS meteorological data, which is derived from the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS) modelling system [39], in the modelling for this project. CERC developed Python processing scripts to automate the processing required which allowed this dataset to be used in the MAQS coupled system. Temporal interpolation was applied to the 3-hourly CAMS data to obtain hourly data for the coupled system forecast. The CAMS grid points were relabelled to a regular grid in the projected coordinate system and the data was written to the netCDF format required by the MAQS coupled system. The use of Python scripts for automated data processing facilitated re-use for different spatial regions and temporal periods, both within the current project and for potential future applications.

Regional model pollutant concentrations

Regional gridded concentration data have been obtained from the archive of the CAMS global forecast dataset [8]. The data have been downloaded at 3-hourly temporal frequency and 0.4° spatial resolution (around 40 km grid cell size). Python processing scripts have been developed to apply bilinear spatial and linear temporal interpolation to this data, before outputting in the netCDF format expected by the MAQS coupled system [14] at hourly temporal resolution and 11 km spatial resolution. An example of the processing of data for one hour in January 2022 for the Vietnamese domain is shown in Figure A.2. The processing was automated to facilitate repetition across the two domains and separate episode months within this project, with flexibility for application to other spatial regions and temporal periods if needed in the future.

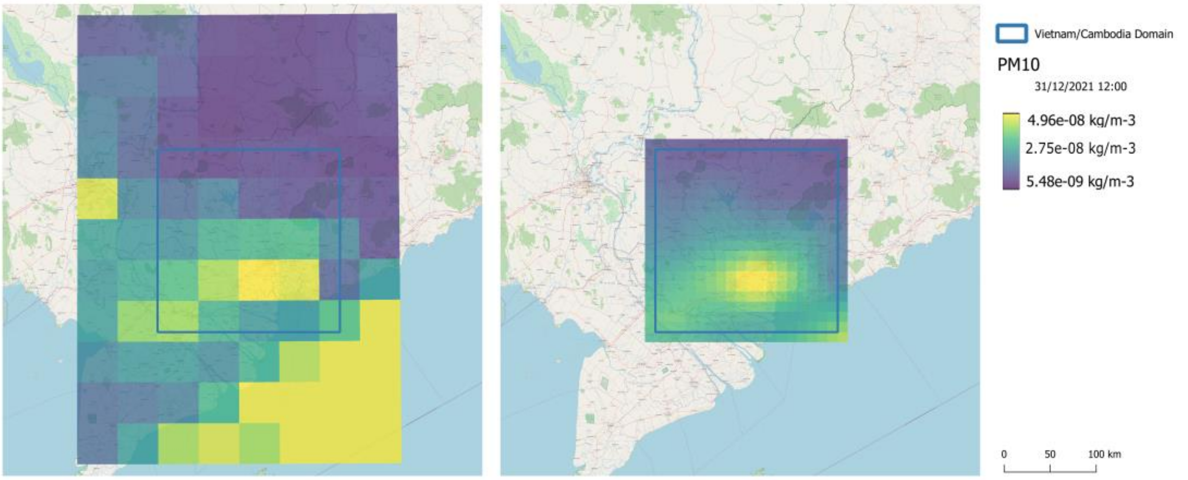


Figure A.2 – CAMS concentration (left) and interpolated concentration (right) for one hour for PM_{10} in the Vietnam/Cambodia domain

Regional emissions datasets

CERC reviewed the available global and regional gridded emissions inventories in order to identify suitable datasets for use in this project. Pollutant emissions can originate from human activities (anthropogenic), which include domestic (in-country) and international emissions, which are relevant to shipping and aircraft emissions. Some types of vegetation emit volatile organic compounds (VOCs) which can be important in long-range atmospheric chemistry, these are categorised as ‘biogenic’. The datasets which have been used in the demonstration forecasting system are summarised in Table A.1. Following communication with the ECCAD database provider, CERC have obtained updated CAMS global anthropogenic emissions for 2022 (version 6.2). The CAMS emissions data does not include anthropogenic total primary aerosol emissions (PM_{10} and $PM_{2.5}$) so these have been obtained from the EDGAR emissions inventory. The latest year available in the EDGAR inventory was 2018, at the time of this project. No reliable scaling factors have been identified to update these emissions to 2022, so the initial modelling assumption was that emissions of aerosol pollutants were unchanged between 2018 and 2022. The CAMS and EDGAR emissions are available as monthly totals; standard diurnal and weekly profiles by activity sector were applied to generate hourly emissions. For the Vietnam/Cambodia domain, specific time-varying profiles based on local traffic counts have been applied to the gridded traffic sector emissions.

Table A.1 – Summary of gridded global emissions data which were used in the coupled system. The anthropogenic emissions (CAMS/EDGAR) are available at 0.1°x0.1° spatial resolution and biogenic emissions at 0.25°x0.25°.

Emissions type	NO_X, CO, SO_2	VOC	PM_{10}, $PM_{2.5}$
Anthropogenic	ECCAD/CAMS glob-ant v6.2 2022	ECCAD/CAMS glob-ant v6.2 2022	EDGAR v6.1 2018
Biogenic	-	ECCAD/CAMS biogenic v3.1 2022	-

The gridded emissions were processed via CERC’s Emissions Inventory Toolkit (EMIT) Software [40]. Some data pre-processing was required to convert the raw CAMS and EDGAR data files into suitable formats for import into EMIT. Python processing scripts were developed to automate this pre-processing. Separate groups were set up in EMIT to contain each sector of gridded emissions data, for example domestic combustion or power generation, in order to assign appropriate vertical and time-varying profiles to each type of emission.

For the Vietnam/Cambodia domain, additional gridded emissions at higher resolution have been supplied by IER VNU-HCM for an area covering Ho Chi Minh City. Additional processing was used to convert these emissions for use in EMIT and the coupled system.

Top-down major road emissions

The coupled street-scale modelling system requires both gridded emissions and emissions for individual roads. The road network geometries for both forecasting domains have been obtained from Open Street Map (OSM). Several pre-processing steps have been carried out on the OSM road networks, initially by filtering for major roads, assumed to be those with an Open Street Map category definition of Motorway, Trunk, Primary, Secondary or Tertiary roads. The roads were then converted from multi to single part shapes, this ensured that each line feature was a single continuous line. The number of vertices on each road have been reduced without losing significant accuracy, by running a GIS tool which requires a tolerance input; a value of 3 m was used. Finally, a maximum length of 2 km was applied to roads, so any exceeding this length were defined as multiple roads. This step ensured that no explicit roads contained more than 50 vertices, which is the maximum permitted within the modelling system.

Initial testing of the Malaysia domain showed that there was a substantial system run time penalty due to the very high spatial density of explicitly modelled roads within the Kuala Lumpur mega-urban area. Hence, within this area the Tertiary category of roads were removed from explicit modelling. This reduced road numbers by 30% within nine problematic cells, which improved run times by around 50%. The final major road networks are shown in Figure A.3 for both domains.

Emissions for individual roads have been calculated by a top-down calculation based on correlation between road lengths of different road types and the total gridded road sector emissions for the relevant grid cell. Population density data from the WorldPop research programme [41], based in the School of Geography and Environmental Sciences at the University of Southampton, was also associated with each road and was incorporated in the correlation calculations. CERC have an existing method for carrying out these calculations, which was used for the Vietnam/Cambodia and Malaysia domains with minor adaptations.

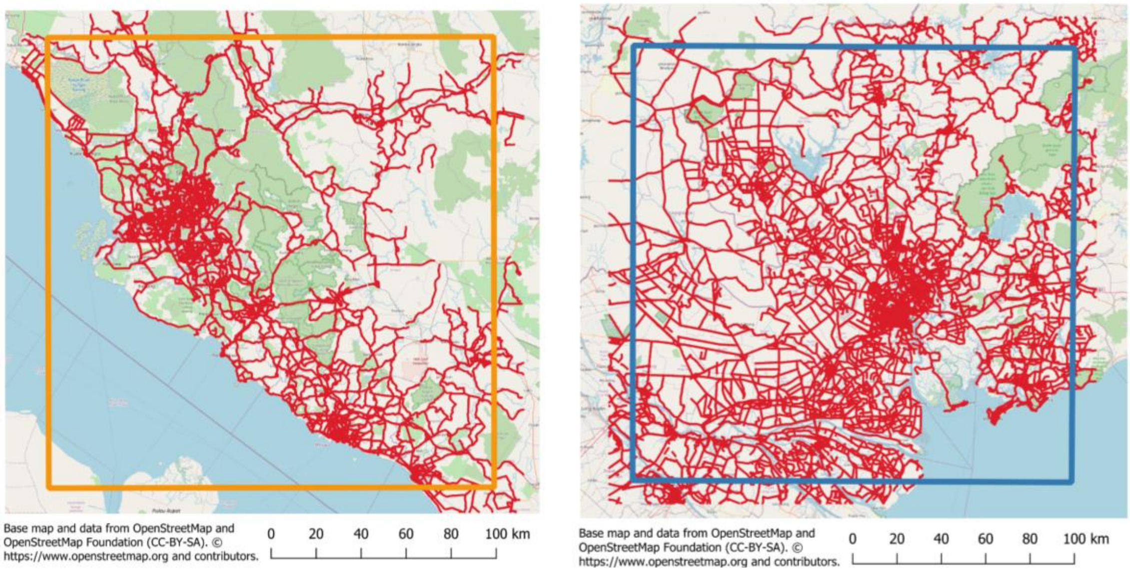


Figure A.3 – Major road networks for Malaysia (left) and Vietnam/Cambodia domains (right), processed from Open Street Map (OSM, [12]) data. There are 47687 explicit roads within the Malaysia domain and 37306 explicit roads within the Vietnam/Cambodia domain.

Local emissions datasets

In addition to the openly available global datasets described in the previous subsections, local project partners shared supplementary data specific to their region and/or country. The use of local data, where available, improves the quality of the modelling results.

- Data for Vietnam/Cambodia: IER VNU-HCM project partners shared the following datasets with CERC, which they generated or obtained through previous projects and which are not openly available.
    - Hourly gridded emissions at 2 km resolution were provided for a 90 km x 90 km area covering Ho Chi Minh City and neighbouring areas, as shown in Figure A.4 [19]. The emissions datasets were divided into four categories: area sources (near-surface anthropogenic emissions, e.g. domestic combustion and agriculture); biogenic sources (VOC emissions from vegetation); point sources (emissions from industrial production processes and fuel consumption in industrial activities); and transportation sources (roads including 'hot' emissions from the exhaust pipes of road vehicles, 'cold' emissions from engines starting up, evaporative emissions, non-exhaust emissions from brake, tyre and road wear, railways, local shipping and aviation). They include emissions for pollutants NO_X, VOC, SO_X, CO, PM_{10} and $PM_{2.5}$ which were modelled in the demonstration system, as well as CH_4 (methane) and TSP (total suspended particulate) which were not used.

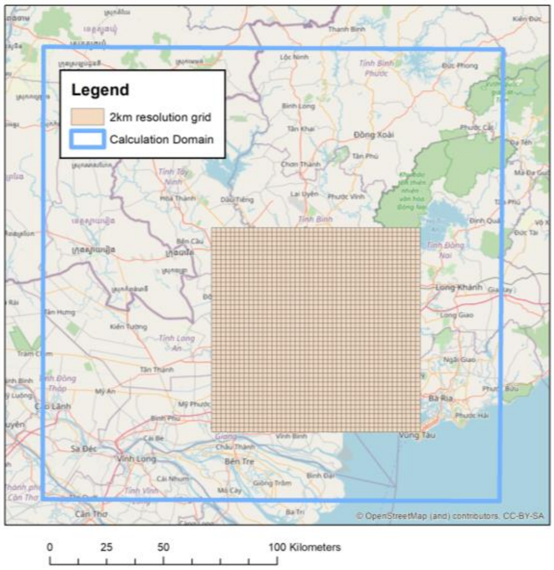


Figure A.4 – Location of 2 km resolution gridded emissions data supplied by project partner IER VNU-HCM within Vietnam/Cambodia domain.

- Traffic count data from 201 roads have been processed into typical hourly variations for five road types, with separate profiles for motorbikes, cars, buses, light duty vehicles (eg. vans weighing less than 3 tons) and heavy duty vehicles (eg. lorries). The road types include three urban types, highways and rural roads. These data have been processed further by CERC in order to calculate typical diurnal emissions variations for explicit roads, using estimates of the relative emission factors for each vehicle type based on descriptions of the Vietnam vehicle fleet from scientific literature ([42], [43]). OSM roads have been assigned to the road types according to the location (including population density and land use) and road category considerations given in Table A.2.
- Air quality measurement data from 20 sites in Vietnam were obtained and shared. These measurements are collected by at least four different organisations and use a variety of measurement techniques and reporting frequencies. Within the national monitoring CEM network, sites are managed at a provincial level before data are collated nationally. This appears to lead to discrepancies in data quality between sites in different provinces. The site locations, coloured by the relevant organisation, are shown in Figure A.5. IER VNU-HCM and CERC have collaborated on processing this data further into the correct format for CERC's Model Evaluation Toolkit [44]. Data from the four sites with discontinuous manual measurements have not been incorporated into the model evaluation process.

Table A.2 – Summary of application of time-varying emissions profiles, calculated from traffic count data shared by IER VNU-HCM, to OSM road categories and locations within the Vietnam/Cambodia such as Ho Chi Minh (HCM) City.

Traffic count road type	Source description	Location applied	OSM road categories
Urban 3	Main urban roads in HCM City	Within HCM	Motorway, Trunk and Primary
Urban 7	Suburban roads in HCM City	Within HCM City	Secondary and Tertiary
		Outside HCM City, population density > 10,000 people/km^2	Primary, Secondary and Tertiary
Urban 14	Industrial area roads	OSM industrial land use category	All except Motorway
Highway	Inter-city roads	Outside HCM City	Motorway and Trunk
Rural	Small rural roads	Population density < 10,000 people/km^2	Primary, Secondary and Tertiary

- Data for Malaysia: UKM project partners obtained the following datasets.
  - Hourly air quality measurement data from 18 routine measurement sites were obtained from the Malaysian Department of the Environment. UKM shared this data for 2022 with CERC. The site locations, coloured by site type, are shown in Figure A.5. CERC have worked with data from some of these sites in previous projects, modelling different years. The data format is straightforward to convert to the format expected by the Model Evaluation Toolkit.
  - Hourly traffic counts from selected roads in Kuala Lumpur from which typical diurnal variation of traffic counts were derived. This data is from the Integrated Transport Information System (ITIS) and is supplied by the Department of Health and Environment at Kuala Lumpur City Hall (Dewan Bandaraya Kuala Lumpur, DBKL). However, delays in UKM receiving this data from DBKL meant that an older version of this dataset, collated during a previous collaboration, was used for the current project.
  - Some typical traffic counts for named major roads in Malaysia are available in a published report [45]. UKM staff worked on geo-referencing this data onto the processed road network (shown in Figure A.3), in order to allow comparisons between bottom-up and top-down emissions calculations for these roads. However, delays in CERC receiving this dataset meant that it was not used within the constrained project timescale.

Measured pollutant concentration datasets

Air quality measurement data were not used for input to the demonstration forecasting system. However, measured concentrations were required in order to evaluate the accuracy of system outputs. Routine air quality measurement data have been collated for both domains by the local project partners. Data validity thresholds have been applied in the model evaluation, such that only data from measurement sites with at least 50% valid measurements within a modelling period were used for evaluation. Some concentration thresholds have also been used to identify invalid measurement data. For example, some sites in the Vietnam CEM network were reporting unrealistically low $PM_{2.5}$ concentrations, substantially below the typical global background concentration of around 5 μg/m^3. Hence a minimum valid $PM_{2.5}$ concentration threshold of 4 μg/m^3 was applied to the measurement data, with associated threshold of 6 μg/m^3 for PM_{10}.

The air quality measurement sites from which data have been provided by the project partners are shown on maps of each domain in Figure A.5. Note that not all of these sites record measurements for all pollutants of interest.

- In the Malaysia domain (left hand plot), most of the sites are in the more populated region near the coastline, with relatively few inland sites.
- In the Vietnam/Cambodia domain (right hand plot), the distribution of measurement sites is predominantly concentrated in and around Ho Chi Minh City, with fewer sites established in rural regions. Certain sites exhibited irregular or sporadic measurement patterns. Moreover, data collected from stations associated with the national monitoring network (denoted by the blue and pink dots) were designated solely for calibration purposes and for comparative analyses with model-derived outputs. No measurement sites were identified in the region of Cambodia included in the modelling domain.

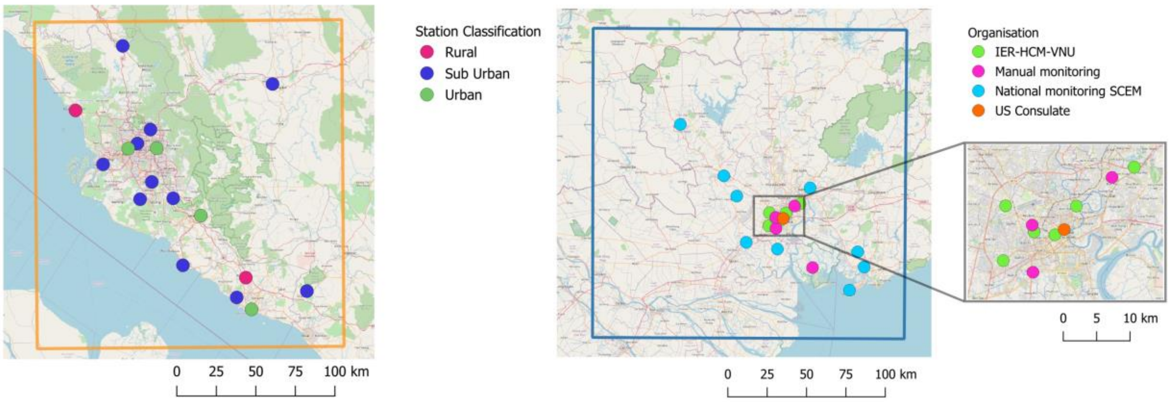


Figure A.5 – Spatial distribution of measurement site locations in Malaysia (left) and Vietnam/Cambodia domains, with an inset showing Ho Chi Minh City, (right). Site types in Malaysia and sites managed by different organisations in Vietnam are shown with different colours.

Appendix B - Supplementary model evaluation and pollution maps

Scatter plots comparing monthly average concentrations from model runs using gridded only or gridded and explicit road emissions for Vietnam/Cambodia are shown in Figure B.1 for NO_2, NO_X and O_3 and in Figure B.2 for PM_{10} and $PM_{2.5}$.

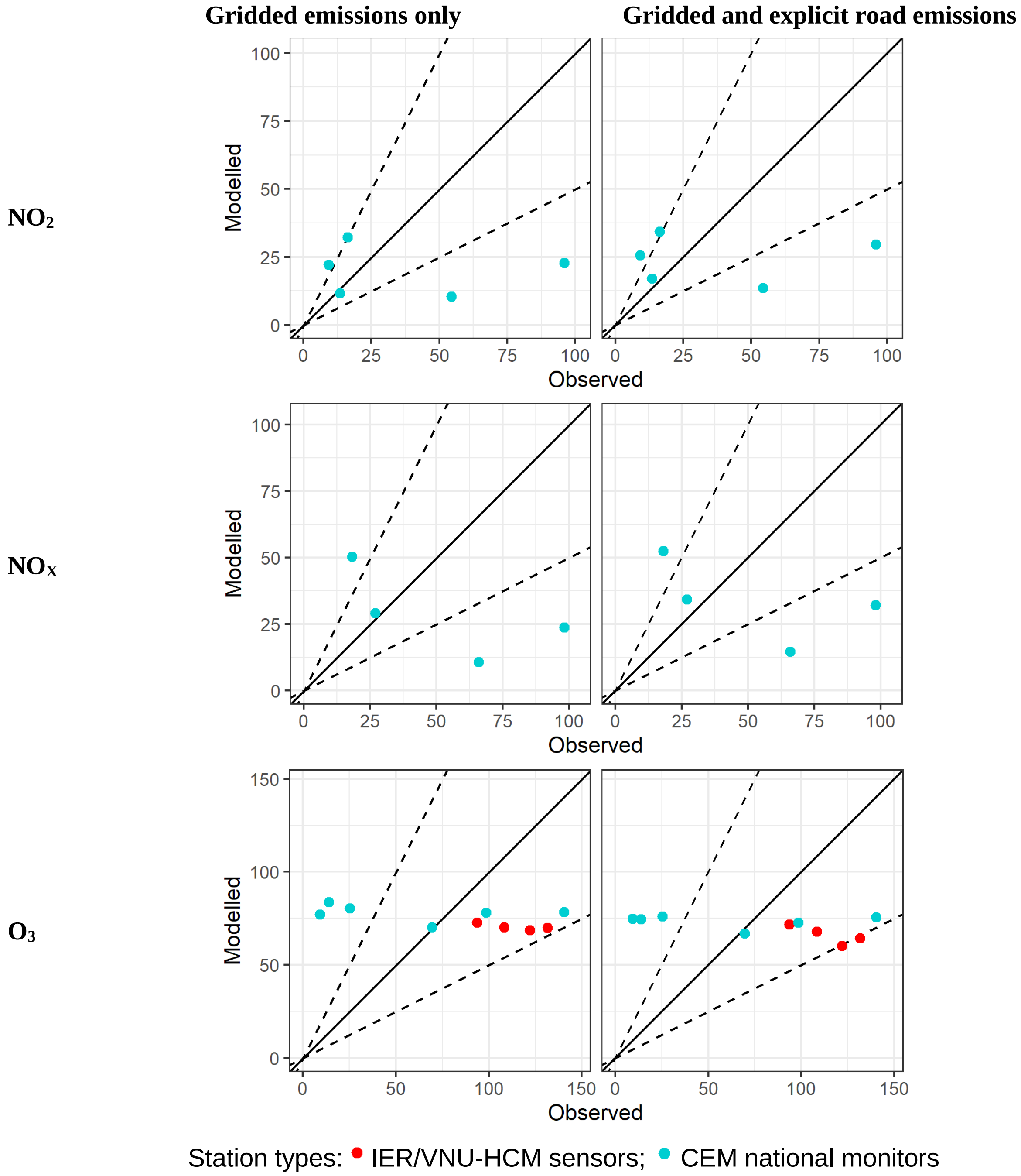


Figure B.1 – Scatter plots comparing monthly mean NO_2, NO_X and O_3 concentrations (units µg/m^3) predicted by the modelling system to measured values, for the Vietnamese/Cambodian modelling domain and January 2022 modelling period. The left-hand panel for each pollutant shows modelled results with only gridded emissions while the right-hand panel also includes explicit road emissions. The solid line shows a 1:1 relationship between modelled and measured concentrations, while the dashed lines indicate factor of two relationships.

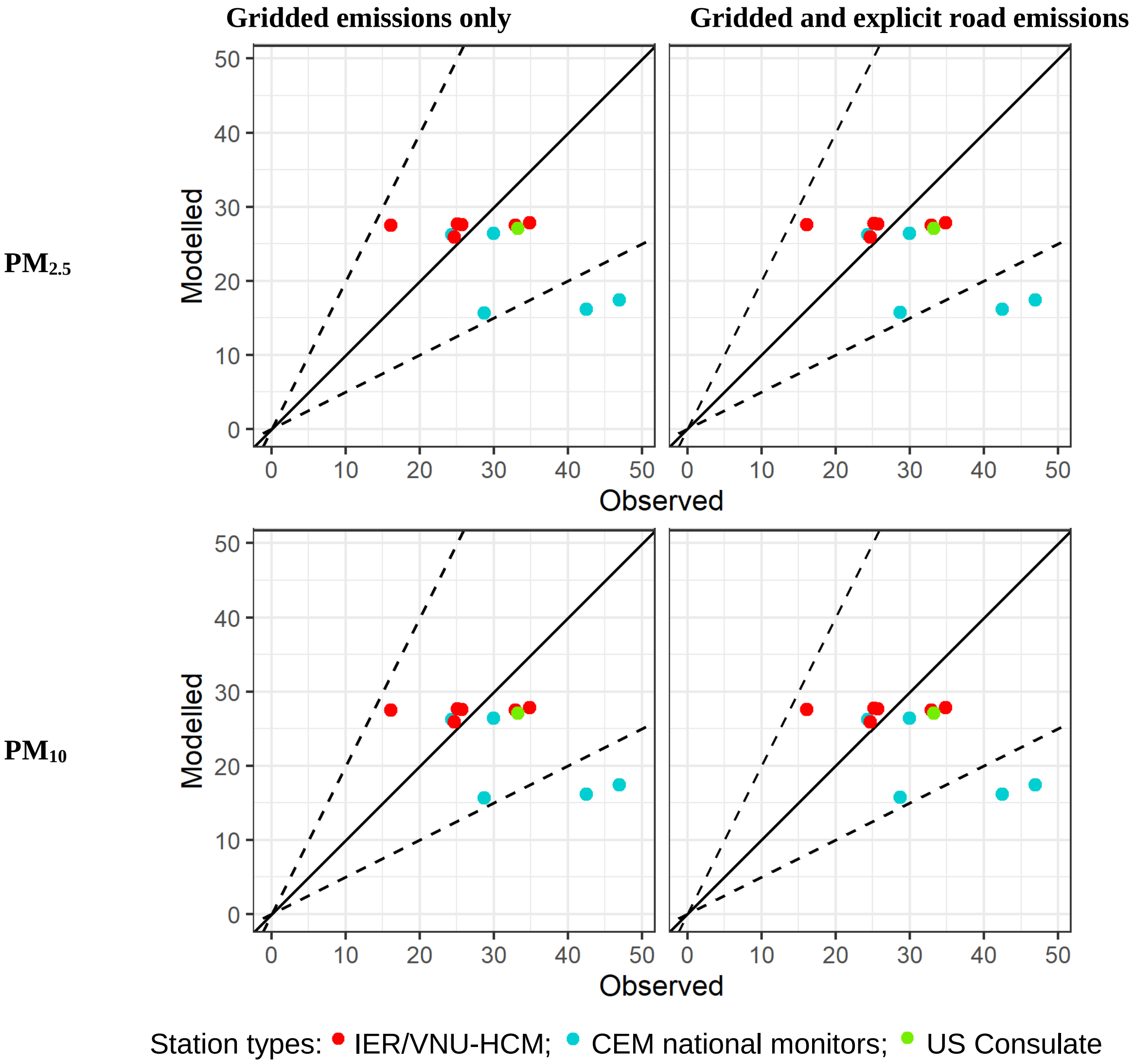


Figure B.2 – Scatter plots comparing monthly mean $PM_{2.5}$ and PM_{10} concentrations (units µg/m^3) predicted by the modelling system to measured values, for the Vietnam/Cambodia modelling domain and January 2022 modelling period. Equivalent to Figure B.1 for gaseous pollutants.

Figure B.3 shows monthly mean CO and SO_2 scatter plots for January 2022 in the Vietnam/Cambodia domain. Both of these show a single outlying measurement, with much higher concentrations than the remaining sites. Explanations for these outliers could be:

- For CO, where inspection of the temporal variation shows that the measurements are consistently very high but show a typical level of variation between days, the sensor may be very close to a combustion source, or there may be an issue with calibration or reported concentration units.
- For SO_2, inspection of the temporal variation shows two days with very high concentrations while other days are of typical magnitude, this may be related to a short-term local event, for example from an accidental industrial release.

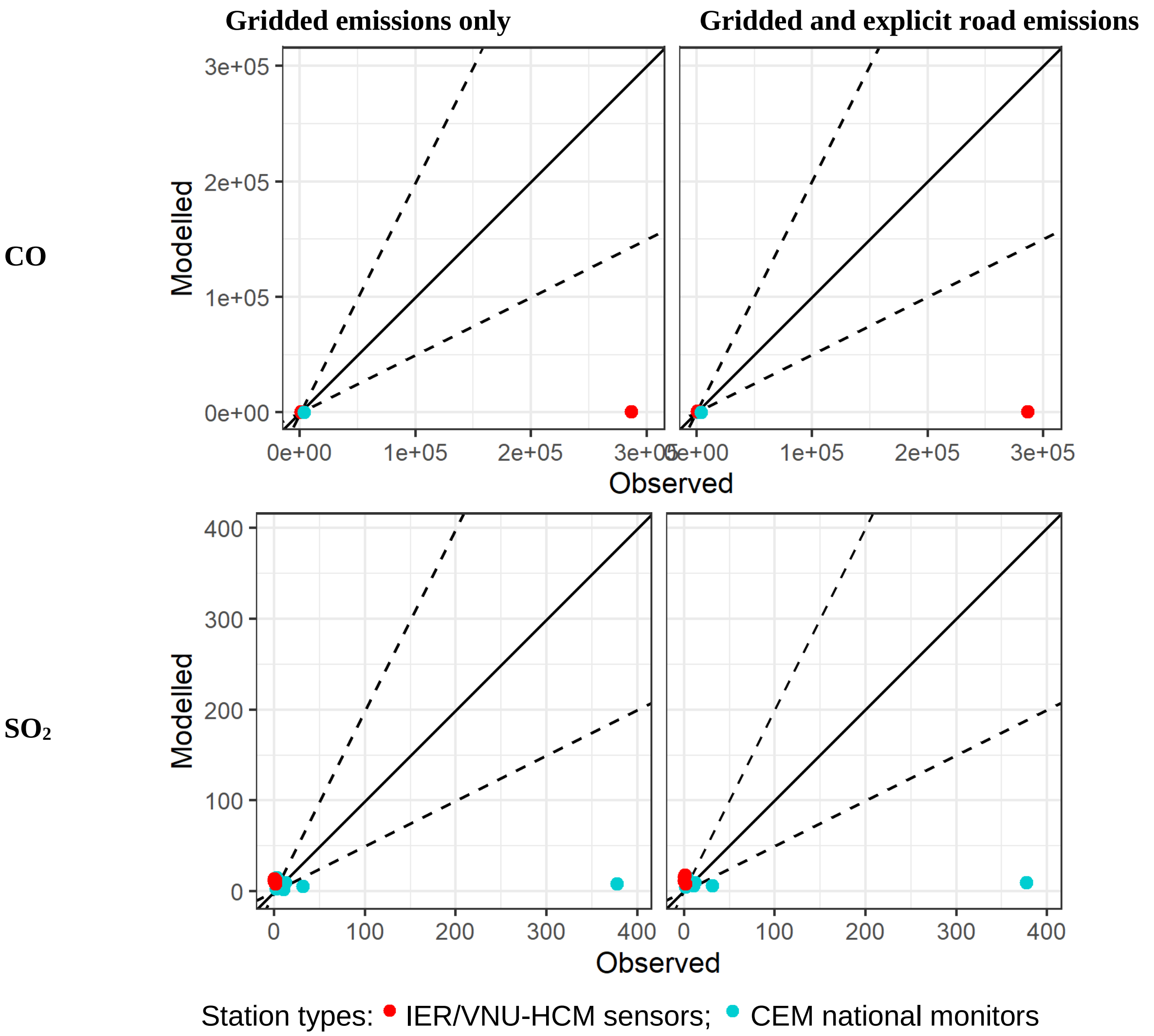


Figure B.3 – Scatter plots comparing monthly mean CO and SO_2 concentrations (units µg/m^3) predicted by the modelling system to measured values, for the Vietnamese/Cambodian modelling domain and January 2022 modelling period. The left-hand panel for each pollutant shows modelled results with only gridded emissions while the right-hand panel also includes explicit road emissions. The solid line shows a 1:1 relationship between modelled and measured concentrations, while the dashed lines indicate factor of two relationships.

Calibration factors for CAMS global forecast concentrations in the Malaysia domain were derived from modelled and observed data in the original 25th April – 25th May 2022 modelling period. The validity of these factors for a January 2022 period was tested using scatter plots without and with calibration, as shown in Figure B.4 for gaseous pollutants and Figure B.5 for particulate pollutants.

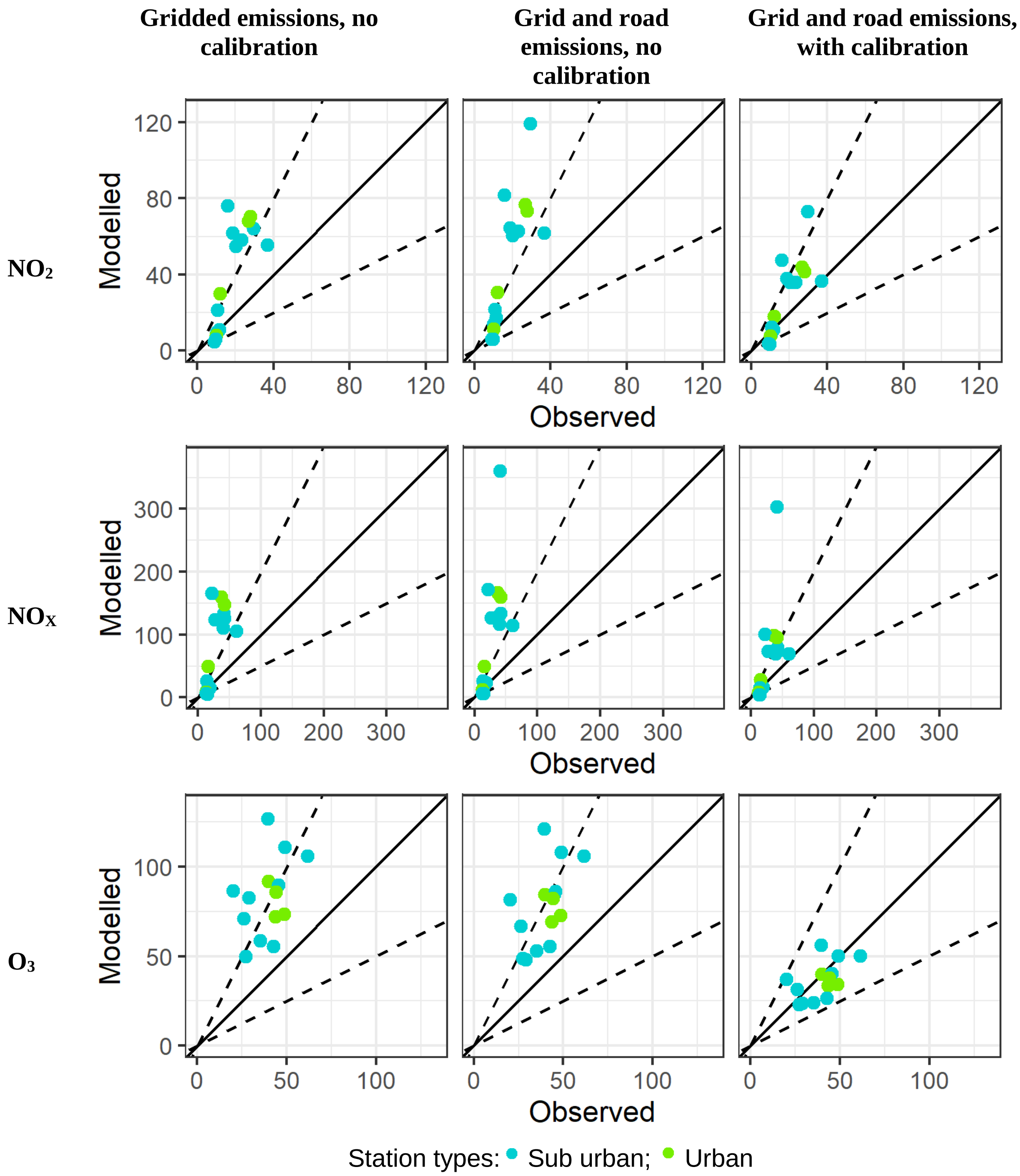


Figure B.4 – Scatter plots comparing monthly mean NO_2, NO_X and O_3 concentrations (units µg/m^3) predicted by the modelling system to measured values, for the Malaysian modelling domain and January 2022 modelling period. The left-hand panel for each pollutant shows modelled results with only gridded emissions and no calibration applied, the middle panel shows results with gridded and explicit road emissions while the right-hand panel includes gridded and explicit emissions as well as calibration factors. The solid line shows a 1:1 relationship between modelled and measured concentrations, while the dashed lines indicate factor of two relationships.

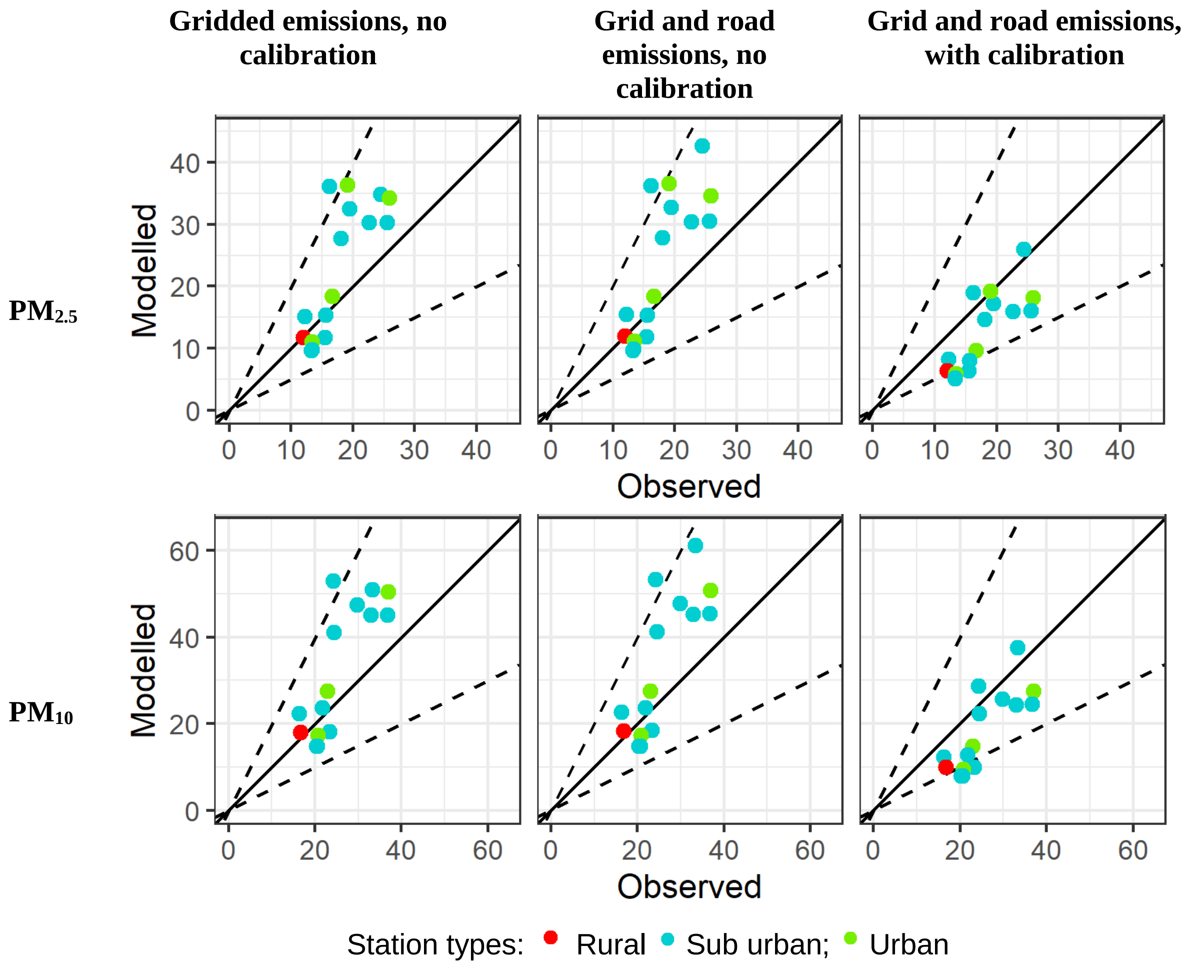


Figure B.5 – Scatter plots comparing monthly mean $PM_{2.5}$ and PM_{10} concentrations (units µg/m^3) predicted by the modelling system to measured values, for the Malaysian modelling domain and January 2022 modelling period. Equivalent to Figure B.4 for gaseous pollutants.

Time-series of modelled and observed daily maximum 1-hour average NO_2 concentrations during January 2022 are shown in Figure B.6 for the Vietnam CEM network Nga tu Phan Dang Luu – Ba Ria site. All NO_2 concentrations are within the 'Good' AQI level. Equivalent timeseries for daily average PM_{10} concentrations at the Nga tu Gieng Nuoc – Vung Tau site are shown in Figure B.7. This also has concentrations within the 'Good' AQI level.

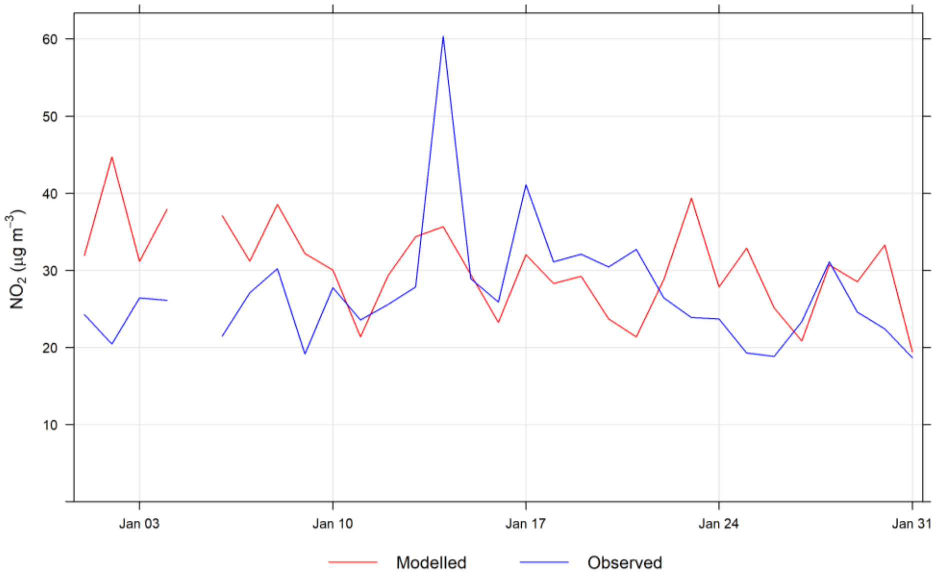


Figure B.6 – Time series of modelled (red) and observed (blue) maximum daily 1-hour average NO_2 concentrations (units μg/m^3), for the Vietnam CEM network Nga tu Phan Dang Luu – Ba Ria site, January 2022. All of the modelled and measured concentrations are within the 'Good' AQI level.

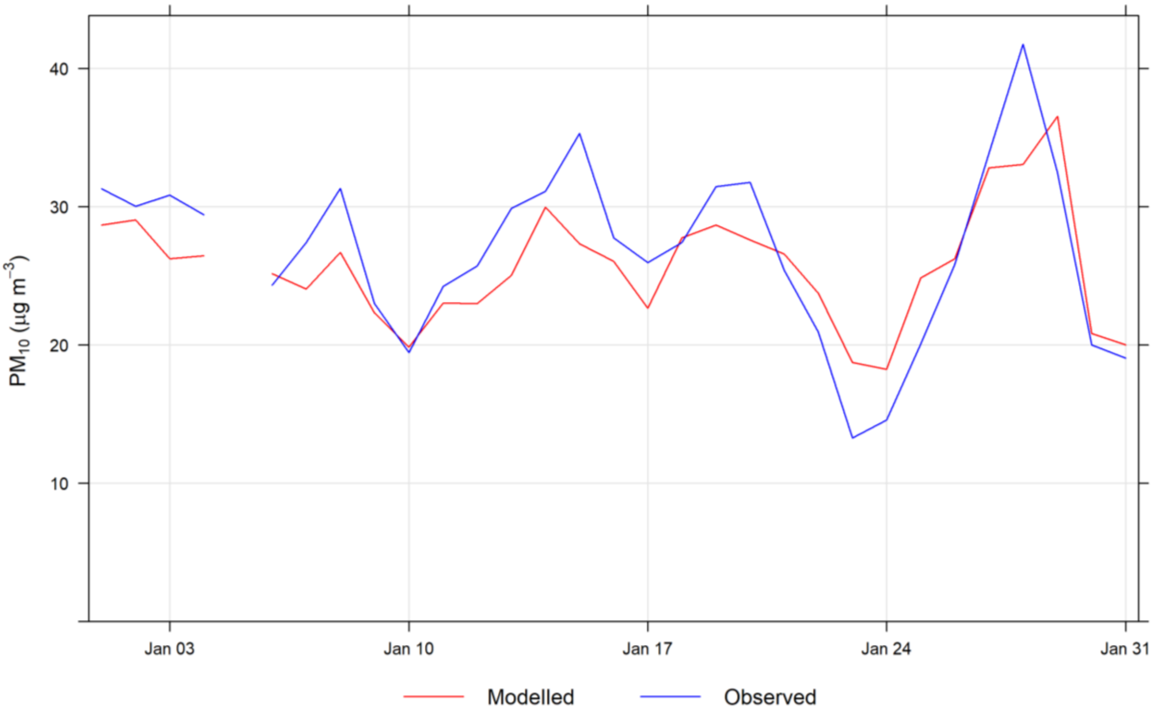
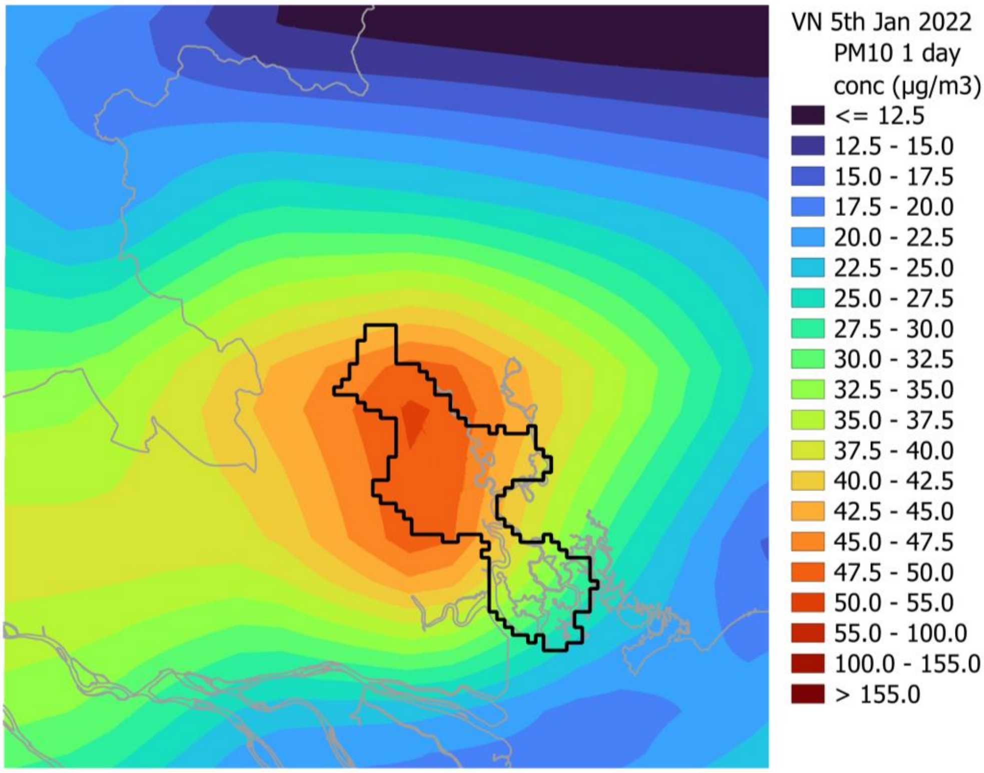


Figure B.7 – Time series of modelled (red) and observed (blue) for daily average PM_{10} concentrations (units μg/m^3), for the Vietnam CEM network Nga tu Gieng Nuoc – Vung Tau site, January 2022. All of the modelled and measured concentrations are within the 'Good' AQI level.



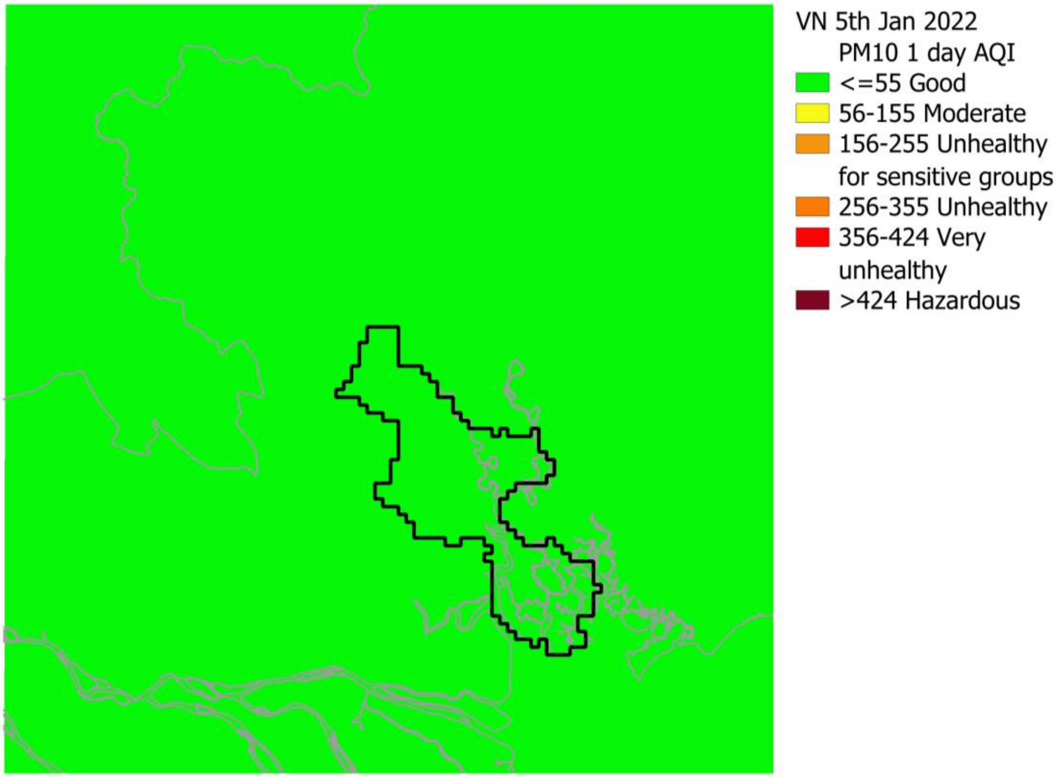


Figure B.8 –24-hour average PM_{10} concentrations (upper panel) and AQI values (lower panel) on 5th January from the MAQS coupled system for the full Vietnamese/Cambodian domain. The coastline and Vietnamese/Cambodian national border are shown as grey lines, while the extent of the subdomain with IER/VNU-HCM finer resolutions is outlined in black.

Maps of maximum daily 8-hour rolling average O_3 concentration and associated API values for the

Malaysian domain on 14th May are shown in Figure B.9. This shows a large area of elevated O_3 concentrations centred to the south-west of Kuala Lumpur, with API level ‘unhealthy for sensitive groups’ along the coast. The concentration plot shows reduced O_3 concentrations along the road network due to local chemical reactions with NO_2. Corresponding maps for $PM_{2.5}$ are shown in Figure B.10, with a wide area of ‘Moderate’ API around Kuala Lumpur and ‘Unhealthy for sensitive groups’ along the major road network. Figure B.11 shows equivalent maps of modelled PM_{10} concentrations and corresponding API values. There are concentration increments along the major road network but mostly ‘Good’ API values.

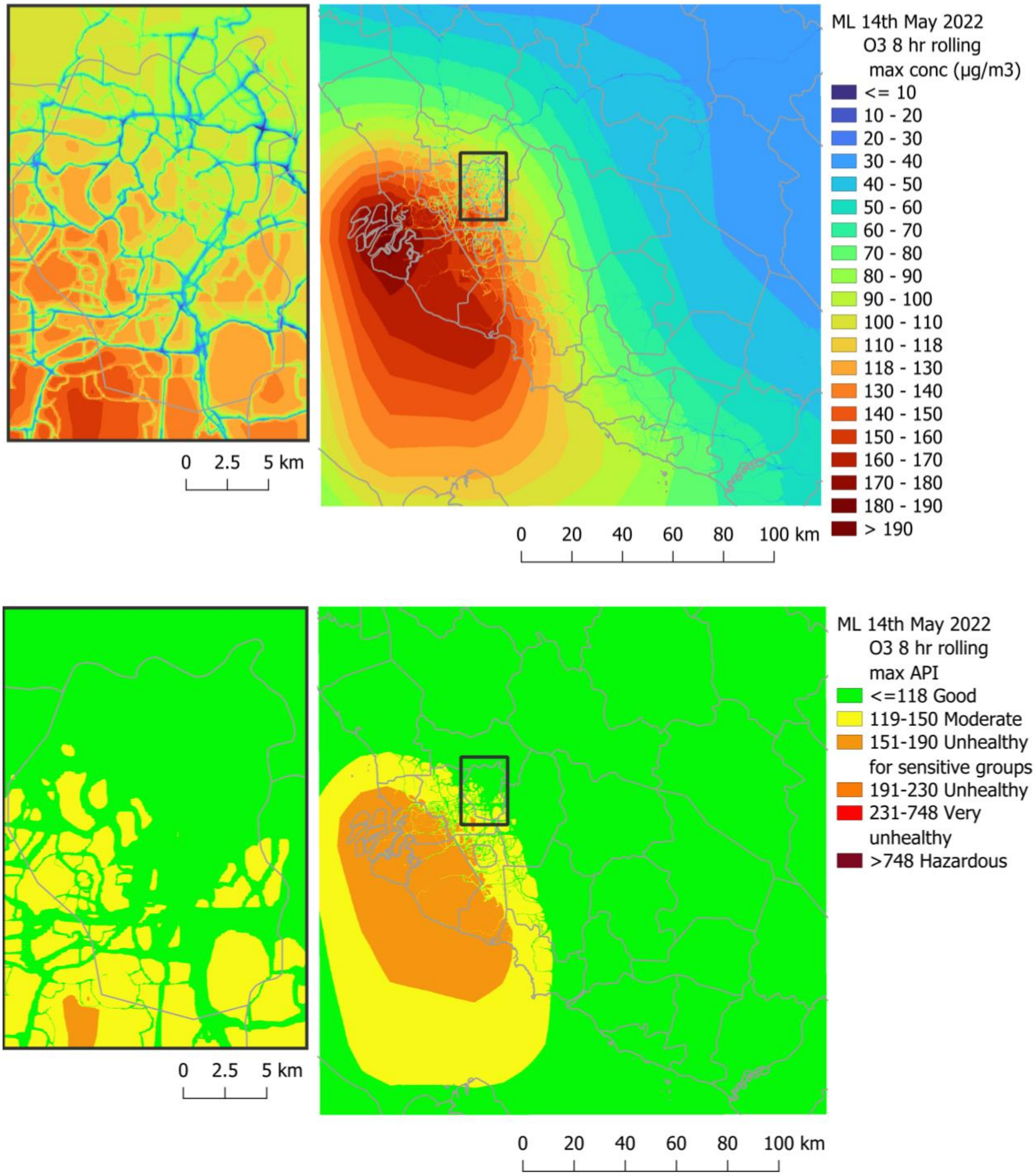


Figure B.9 –Maximum 8-hour rolling average O_3 concentrations (upper panel) and API values (lower panel) on 14th May from the MAQS coupled system for the full Malaysian domain (right-hand panels) with additional detail for central Kuala Lumpur (left-hand panels). The coastline and local authority boundaries are shown as grey lines.

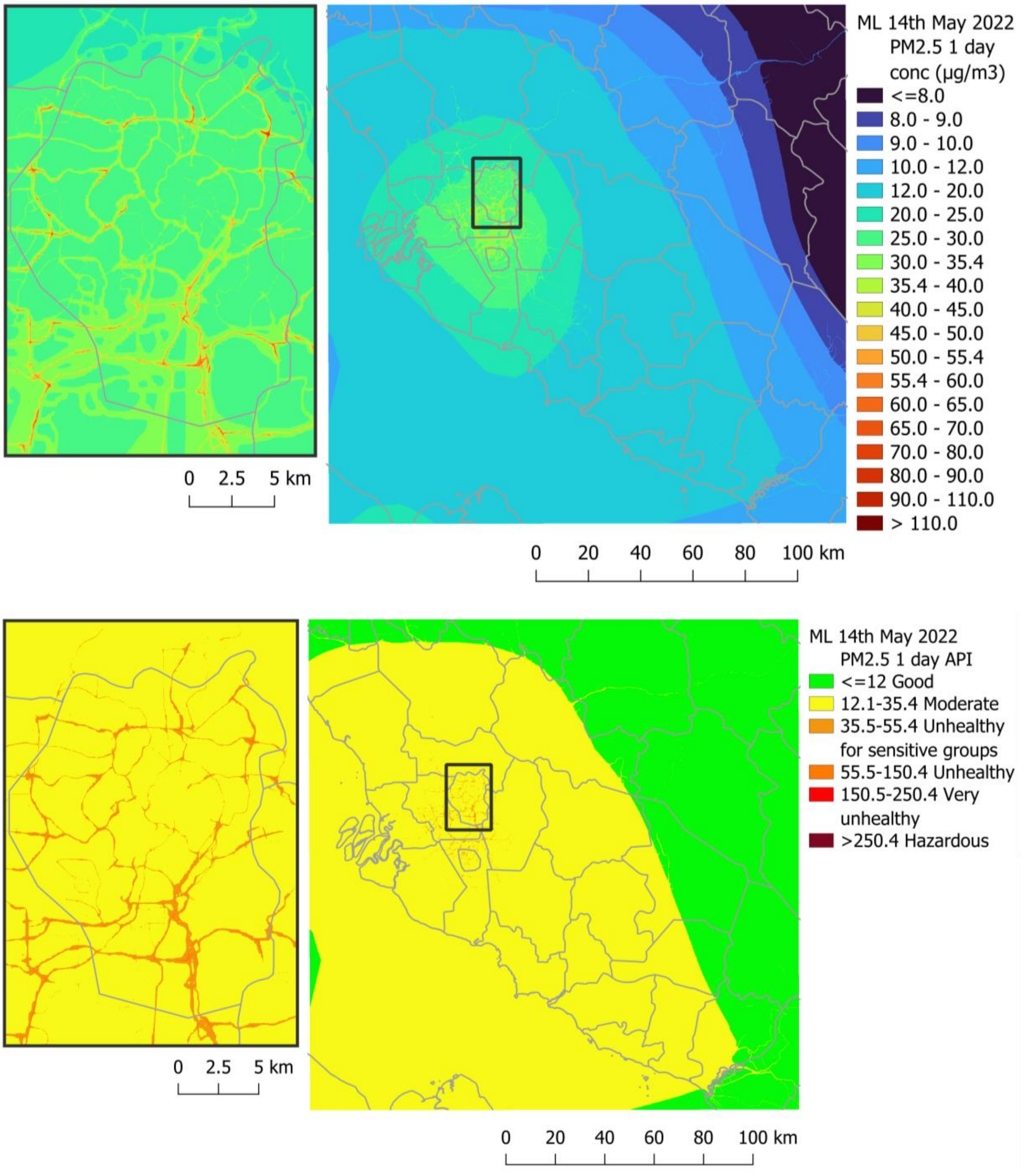


Figure B.10 –24-hour average $PM_{2.5}$ concentrations (upper panel) and API values (lower panel) on 14th May from the MAQS coupled system for the full Malaysian domain (right-hand panels) with additional detail for central Kuala Lumpur (left-hand panels). The coastline and local authority boundaries are shown as grey lines.

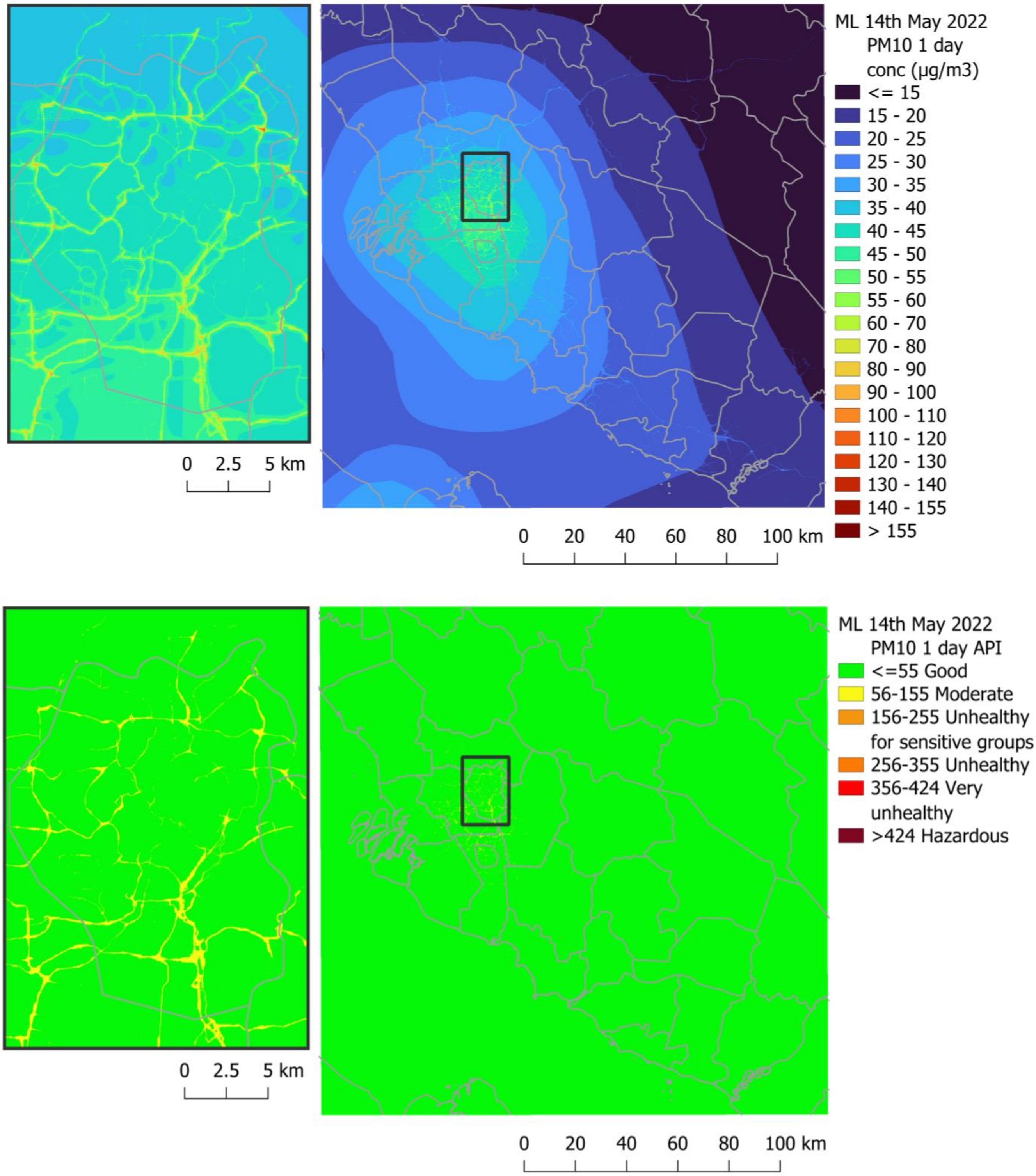


Figure B.11 –24-hour average PM_{10} concentrations (upper panel) and API values (lower panel) on 14th May from the MAQS coupled system for the full Malaysian domain (right-hand panels) with additional detail for central Kuala Lumpur (left-hand panels). The coastline and local authority boundaries are shown as grey lines.

Appendix C - Stakeholder engagement details

DHAQ-SEA - Demonstration of High-resolution Air Quality forecasting in South East Asia

DHAQ-SEA is a feasibility study that demonstrates the development of a high resolution (street-scale) air quality forecasting system for ASEAN regions. The MAQS coupled system is being used to ensure that both regional (using data from CAMS) and local (modelled using ADMS-Urban) pollution is accounted for. The system could be implemented throughout ASEAN, as it relies primarily on open data sources.

The project focus is on the priority area 'Health', addressing Sustainable Development Goal (SDG) 3 'Ensure healthy lives and promote well-being for all at all ages.' There is an increased risk of certain health conditions where ambient air pollution levels are high. ASEAN countries can be subject to poor air quality as a consequence of transboundary pollution as well as local pollutant sources. Alerting the public prior to air pollution episodes allows vulnerable and other individuals to take action to reduce the health impacts of poor air quality.

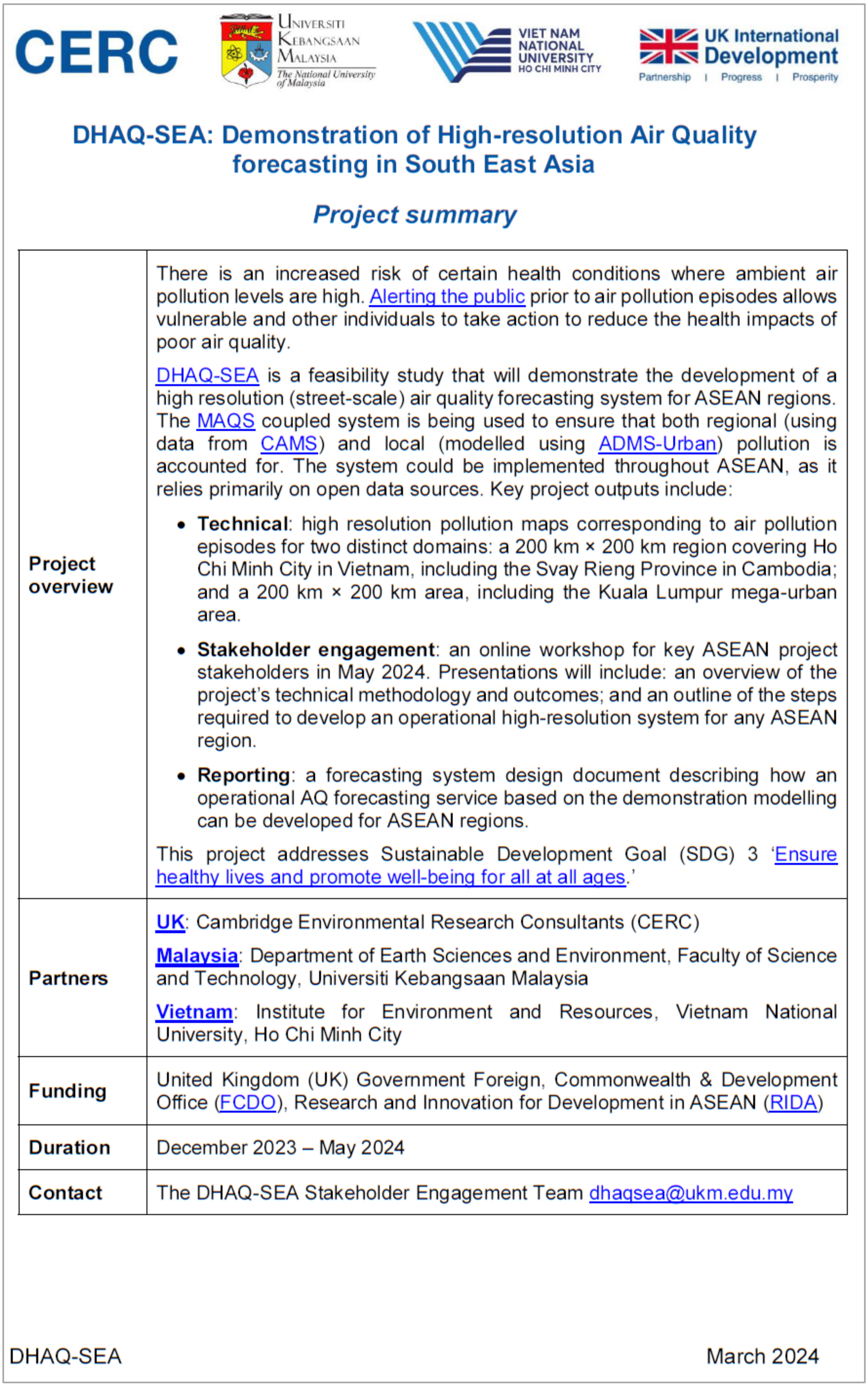
CERC are working on DHAQ-SEA with project partners from the Institute for Environment and Resources, Vietnam National University, Ho Chi Minh City and the Department of Earth Sciences and Environment, Faculty of Science and Technology, Universiti Kebangsaan Malaysia. The project has received Official Development Assistance (ODA) funding from the UK government Foreign, Commonwealth & Development Office (FCDO) under the Research and Innovation for Development in ASEAN (RIDA) programme.

Key DHAQ-SEA project outputs will include:

- **Technical**: high resolution pollution maps corresponding to air pollution episodes for two distinct domains: a 200 km × 200 km region covering Ho Chi Minh City in Vietnam, including the Svay Rieng Province in Cambodia; and a 200 km × 200 km area, including the Kuala Lumpur mega-urban area.
- **Stakeholder engagement**: An online workshop for key ASEAN stakeholders in May 2024. Presentations will include: an overview of the project's technical methodology and outcomes; and an outline of the steps required to develop an operational high-resolution system for any ASEAN region.
- **Reporting**: a forecasting system design document describing how an operational AQ forecasting service based on the demonstration modelling can be developed for ASEAN regions.

Figure C.1 – Online project description https://www.cerc.co.uk/environmental-research/regional-to-local.html#DHAQ-SEA

CERC — Universiti Kebangsaan Malaysia, The National University of Malaysia — Viet Nam National University Ho Chi Minh City — UK International Development: Partnership | Progress | Prosperity



CERC — Universiti Kebangsaan Malaysia, The National University of Malaysia — Viet Nam National University Ho Chi Minh City — UK International Development: Partnership | Progress | Prosperity

DHAQ-SEA: Demonstration of High-resolution Air Quality forecasting in South East Asia

Project summary

Project overview	There is an increased risk of certain health conditions where ambient air pollution levels are high. Alerting the public prior to air pollution episodes allows vulnerable and other individuals to take action to reduce the health impacts of poor air quality. DHAQ-SEA is a feasibility study that will demonstrate the development of a high resolution (street-scale) air quality forecasting system for ASEAN regions. The MAQS coupled system is being used to ensure that both regional (using data from CAMS) and local (modelled using ADMS-Urban) pollution is accounted for. The system could be implemented throughout ASEAN, as it relies primarily on open data sources. Key project outputs include: • **Technical**: high resolution pollution maps corresponding to air pollution episodes for two distinct domains: a 200 km × 200 km region covering Ho Chi Minh City in Vietnam, including the Svay Rieng Province in Cambodia; and a 200 km × 200 km area, including the Kuala Lumpur mega-urban area. • **Stakeholder engagement**: an online workshop for key ASEAN project stakeholders in May 2024. Presentations will include: an overview of the project's technical methodology and outcomes; and an outline of the steps required to develop an operational high-resolution system for any ASEAN region. • **Reporting**: a forecasting system design document describing how an operational AQ forecasting service based on the demonstration modelling can be developed for ASEAN regions. This project addresses Sustainable Development Goal (SDG) 3 'Ensure healthy lives and promote well-being for all at all ages.'
Partners	**UK**: Cambridge Environmental Research Consultants (CERC) **Malaysia**: Department of Earth Sciences and Environment, Faculty of Science and Technology, Universiti Kebangsaan Malaysia **Vietnam**: Institute for Environment and Resources, Vietnam National University, Ho Chi Minh City
Funding	United Kingdom (UK) Government Foreign, Commonwealth & Development Office (FCDO), Research and Innovation for Development in ASEAN (RIDA)
Duration	December 2023 – May 2024
Contact	The DHAQ-SEA Stakeholder Engagement Team dhaqsea@ukm.edu.my

DHAQ-SEA March 2024

Figure C.2 – Project flyer (.pdf format)

Figure C.3 – Workshop agenda providing details of the event date and time, speakers, platform, and introducing the workshop moderator.

Table C.1 – DHAQ-SEA stakeholder list: ASEAN

Country	Number of stakeholder organisations			Organisations
	Total	Registered for workshop	Attended workshop	
Philippines	3	1		Air Quality Management Section, Department of Environment and Natural Resources; Don Honorio Ventura State University; Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development (DOST-PCAARRD), FCDO
Thailand	5	1	4	Department of Meteorology and Hydrology; Mahidol University; Prince of Songkla University; Thai Meteorological Department; Chulalongkorn Medical School, Bangkok
Malaysia	22	3	17	Ahli Majlis, Subang Jaya City Council; Dewan Bandaraya Kuala Lumpur (DBKL); Department of the Environment; Institute for Medical Research; Malaysian Meteorological Department; JD Healthcare Holdings; Management and Science University; Maybank Malaysia; MSIG Insurance (Malaysia) Bhd; UKM; UKM Pakarunding; Universiti Malaysia Kelantan; Universiti Malaysia Pahang Al-Sultan Abdullah; Universiti Malaysia Perlis; Universiti Malaysia Sabah; Universiti Malaysia Terengganu; Universiti Pendidikan Sultan Idris Universiti Pertanian Malaysia (UPM); Universiti Putra Malaysia; Universiti Sains Malaysia; Universiti Sultan Zainal Abidin; Universiti Teknologi Malaysia; Xiamen University Malaysia
Myanmar	3	1	3	Department of Meteorology and Hydrology; University of Yangon; Environmental Conservation Department
Vietnam	6	2	2	Ho Chi Minh City Marine Port Authority; Vietnam National University – Ho Chi Minh City; Natural Resources and Environment Research Institute; Ho Chi Minh City Environmental Protection Department (HEPA); Ho Chi Minh City University of Technology; World Bank Vietnam
Cambodia	2			Royal University of Phnom Penh, Faculty of Science; FCDO
Indonesia	8	1	3	Bandung Institute of Technology; Environment Agency of DKI Jakarta; Institut Teknologi Sumatera; IPB University; Jakarta Environmental Agency; Meteorological, Climatological and Geophysical Agency; Indonesia National Nuclear Energy Agency (BATAN); FCDO
Singapore	1	1		National Environment Agency
Laos	1	1	1	National University of Laos, Dean of Environmental Faculty
Brunei	2		1	Sultan Sharif Ali Islamic University (UNISSA); Department of Environment, Parks and Recreation, Ministry of Development

Table C.2 – DHAQ-SEA stakeholder list: non-ASEAN

Country	Number of stakeholder organisations			Organisations
	Total	Registered for workshop	Attended workshop	
Bangladesh	2		1	SGS Bangladesh Ltd; Department of Environment, Bangladesh
Ethiopia	1		1	Environmental Protection Agency (EPA)
Japan	4	1	3	National Institute for Environmental Studies; Hokkaido University; The University of Tokyo; Kyoto University of Advanced Science
India	1		1	Banaras Hindu University Varanasi
Islamabad	1		1	International Islamic University
Mongolia	1		1	National University of Mongolia
Oman	1		1	University of Nizwa
Pakistan	1		1	Punjab Environmental Protection Agency
Peru	2		2	Municipalidad Metropolitana de Lima; Servicio Nacional de Meteorología e Hidrología del Perú (Senamhi)