IGI Cropped for letter head

***CLEAN AIR IGI/IAS Pump Priming Funding 2022-23 – Round 1***

**\*\*Please complete and submit the form below by email to** [**clean.air@contacts.bham.ac.uk**](mailto:clean.air@contacts.bham.ac.uk) **by 5pm on Friday 20th January\*\***

**Lead Applicant Details**

| Name | Yuqing Dai |
| --- | --- |
| Position |  |
| School / College | School of Geography, Earth & Environmental Sciences |
| Email | yxd598@bham.ac.uk |
| Details of other staff involved (name, position, school / college) | Rob Mackenzie: Professor, GEES, Supervising the whole project  Jiahua Jiang: Assistant Professor, Mathematics, Supervising Machine learning model  Bowen Liu: Research Fellow, Economics, Data analysis |

**Project Overview**

| **Project title** | Air pollutant dispersion in street canyons based on machine learning |
| --- | --- |
| **Total value of funding sought** | £8,000 |
| **Proposed activity start date** | 1 Feb 2023 |
| **Proposed activity end date** | 30 June 2023 |

**Project Details**:

| Please provide a description of the project (using the headings: background, aim, activities) for which funding is sought *(max 1 side)* |
| --- |
| **Research background:** Air pollutants from intensive tail-pipe emissions can easily get trapped in urban canyons due to inadequate ventilation, leading to a great threat to public health and becoming an important concern. A better understanding of air pollutant behaviours and effective tools for evaluating air quality interventions are needed to support air quality management in such environmental “hotspots”.  **Aim:** The overarching goal of this project is to estimate and understand the causal impact of “canyon effects” on air quality in a regular canyon by the means of state-of-the-art machine learning (ML) techniques and economic modelling. This project consists of the following three objectives:  1). The first objective is to develop interdisciplinary methods to understand the causal impact of “street canyon” on air quality. In particular, we will focus on pollutants such as NOx and NO2, the former of which is often regarded as a passive scalar and the latter of which is engaged in NOx-O3 and VOC radical chemistry. Comparing NO2 and NOx can deliver insights into the influence of atmospheric processes on statistical simulations.  2). A fast next-generation machine learning (ML) model will be used to explain and understand the impact of environmental interventions or socio-economic activities on air quality in a regular street canyon. We will delve into the importance of different factors and proxies (i.e., model inputs), particularly, traffic volumes and compositions, on model performance. COV-19 lockdown data will be used as a case study.  3). What is the highest traffic volume that can be sustained in a standard canyon while still adhering to the air quality standard under “extreme” meteorological conditions (e.g., background winds perpendicular to the street axis)? We will identify the behavioural responses, and discover transport demand changes and “wiggle room” for traffic emissions in urban areas.  **Activities:**  1). Data collection: hourly air pollution data during 2011–2021 in London Marylebone Road (MR) and North Kensington (NK), observed and estimated (European Centre Reanalysis Version 5, ERA5) meteorological data (e.g., wind speed, wind direction, temperature, atmospheric pressure, relative humidity, surface net solar radiation, total cloud cover, boundary layer height), and traffic counts will be collected.  2). Model development: light gradient-boosting machine (LightGBM) is an efficient and high-performance machine learning approach invented by Microsoft in 2016, and it is particularly well-suited for tasks such as regression, classification, and ranking. It has risen to prominence within both academia and industry and received considerable success in many fields, such as bioinformatics, computer vision, natural language processing, finance, and recommender systems. Compared with recent renaissances in ML-based air pollution prediction methods, such as random forest and eXtreme Gradient Boosting, LightGBM results in more loss reduction and, in turn, higher accuracy while being faster and properly tuned, especially for large and complex datasets. In this work, a LightGBM regression model tailored for air pollutantconcentrations will be developed based on different meteorological factors and emission scenarios.  3). Model evaluation: Statistical indicators such as systematic/unsystematic root-mean-square error (RMSE) and index of agreement (IOA) will be applied to assess the efficacy of our proposed ML model. A comprehensive comparison between our model and a process-based multi-box model (Dai et al., 2021) for street canyon simulations will be provided.  4). Data analysis:   1. ML results will be compared against observed values during the COV-19 lockdown period. With the trained model, “business-as-usual” scenarios can be investigated, which will enable a better understanding of the impact of lockdown on air quality in street canyons. 2. The causal impact of street canyons on air quality will be discussed using economic models such as Augmented Synthetic Control Method (ASCM). Air quality data at North Kensington can be used as a “control group”. This will strengthen the knowledge of air pollutant behaviours in urban canyon environments. 3. “Counterfactual” scenarios will be built to evaluate the maximum traffic load under different meteorological conditions (e.g., prevailing winds). This can be beneficial for a plethora of intentions, such as formulating emergency plans and pinpointing potential pinch points or vulnerabilities in controlling air pollution in urban canyons. |

1. **Project costs:** *Please provide a breakdown of all project costs being sought*

*(please note funding for core funded staff will not be provided)*

| Cost (item and timescale – consider feasibility vs timescale). Include all costs- VAT, delivery etc as appropriate  Note: all funds must be spent by end of June 2023 | Amount (£) |
| --- | --- |
| DR student – 15 hours per week for 21 weeks  Workshops, meeting, and paper materials | £5,250 (£16.67 per h for 15 hrs per week)  £2,750 (estimated) |
| **TOTAL** | £8,000 |

1. **Project deliverables**

| *Please provide a list of project deliverables (e.g., datasets, publications, reports) and dates by which these will be available.* |
| --- |
| * Air quality and meteorological datasets will be collected. (estimated 15th Feb 2023). * The training of the LightGBM regression model will be finished. (estimated 30th March 2023). * A quantitative evaluation and analysis of air pollution will be undertaken, drawing upon the ML results. (estimated 30th April 2023). * One publication will be submitted in peer-reviewed journals of international standing (estimated 30th June 2023). * A research proposal for fellowship in relative subjects (see below). * The proposed research will be presented at least one conference and future engagement workshops (estimated before 30th June 2023). |

1. **Outcomes**

*The two main aims of the pump priming funding are to develop interdisciplinary funding proposals and/or to establish new strategic partnerships.*

| *Please provide details of the specific proposal (e.g., funder, funding scheme: note that there are often recurring funding schemes such as NERC Pushing the Frontiers, MRC Applied Global Health Research, Leverhulme or Wellcome Trust fellowships and standard grants) and/or development of a partnerships (e.g., progression to signature of MOU).* |
| --- |
| Air pollution in urban “hotspots” severely affects public health and has risen to the top of the policy agendas of local and regional officials. This pump-priming project aims to assist in addressing the need to tackle air quality issues in urban canyons and attain legal compliance with standard values. Our results will provide new insight into the knowledge of air pollutants behaviours in urban environments. By applying advanced interdisciplinary science, this project will help inform the development of effective control measures, such as emission reduction strategies (e.g., traffic flow management), to improve air quality.  Importantly, our vision is to establish a sustainable diverse interdisciplinary network, connecting researchers with different backgrounds and expertise to tackle complex environmental tasks. IGI funding will offer a strong foundation for the follow-on research collaboration and grant application. Specifically, air-quality alert systems (AQAS) cover more than 1.7 billion people and have significant impacts on social activities and economies, especially in developing countries. However, air quality improvements and health budgets due to AQAS have not been accurately assessed. Coupled with economic models (e.g., Augmented Synthetic Control Method and Synthetic Difference-in-Difference), the fast and powerful ML model developed in this work will be further enhanced and upgraded to evaluate the causal impact of AQAS on air quality and public health. Within this project, we will develop a practical solution and toolkit to environmental and health challenges.  Considering our team comprising experts from a diverse array of fields, including Environmental Science, Computer Science, Mathematics, Economics, a subsequent proposal based on the findings supported by IGI will be submitted to “Applied global health research” from Medical Research Council (MRC) and Foreign, Commonwealth and Development Office (FCDO). The topic and diversity of skills fit perfectly into the scope of the funding.Given the interdisciplinary expertise of our team, the proposal can also be submitted to “UKRI Economic and Social Research Council (ESRC) postdoctoral fellowships”, as well as “Leverhulme Early Career Fellowships”, where the PI can establish and extend hid research records after completion of PhD. Moreover, the team members can further deepen the interdisciplinary communications among environmental science, computer science and social science. By harnessing the power of LightGBM, our proposed machine learning model seeks to augment the capabilities of existing models in environmental science (Random Forest ML model) and social science (causal inference model). |

1. **Fit to Clean Air theme workstreams / priority areas** (Air Pollution & Health / Indoor Air Quality / Engagement & Impact / Clean Air Solutions / Net Zero & Climate)

| *Please outline fit to Clean Air theme workstreams / cross-cutting themes (50 words)* |
| --- |
| Our project shows a compelling fit to the objectives of the Clean Air theme’s priorities – Air Pollution & Health and to the theme of Clean Air Solutions. |

1. **Previous IGI Clean Air funding**

| *Please outline any previous IGI Clean Air funding and outcomes from this* |
| --- |
| The project is a proposal for growing scientific and environmental concerns, it has not yet been funded. |

1. **Confirmation:**

I understand all funds must be spent / invoiced by 30/6/23. I understand that (where relevant) all studies must have evidence of existing ethical approval and risk assessment undertaken before the start of the project. I will provide short reports (1 side) of project outcomes and planned next steps as requested, and a short presentation on my project at future Clean Air theme meetings.

PI Signature / electronic signature \_Yuqing Dai\_\_\_\_\_\_\_\_\_\_\_

Date \_12/01/2023\_\_\_\_\_\_\_\_\_\_