

Data Science and Analytics for Air Pollution



Dr. Congbo Song
University of Birmingham
27 June 2022

Journey of Congbo



A pandemic post-doc career

China—> UK (2019.8-)



Natural
Environment
Research Council

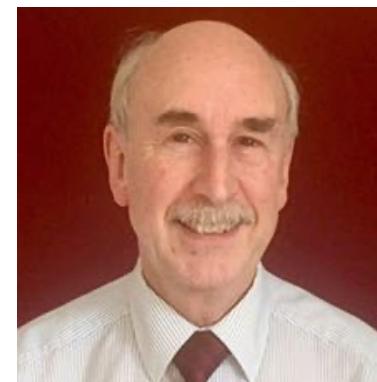
Research fellow for a NERC-funded project SEANA at



£1,245,811
Jan 19 - Jun 26

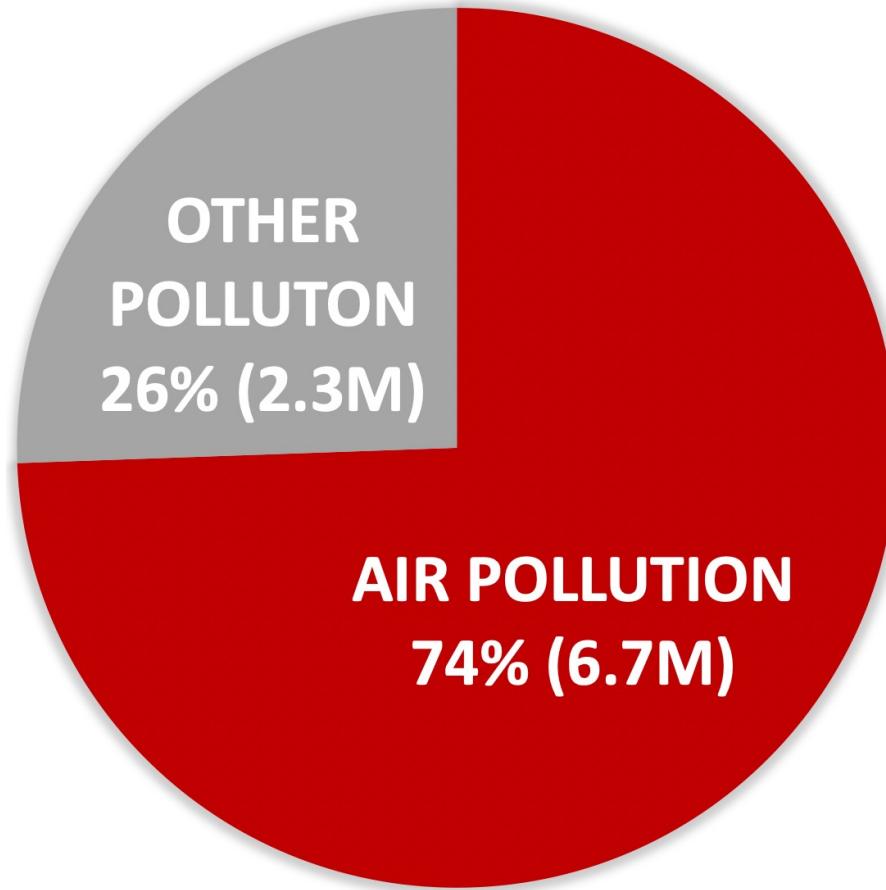
SEANA -Shipping Emissions in the Arctic and North Atlantic atmosphere

NERC award to [University of Birmingham](#) and Zongbo Shi



~3 years

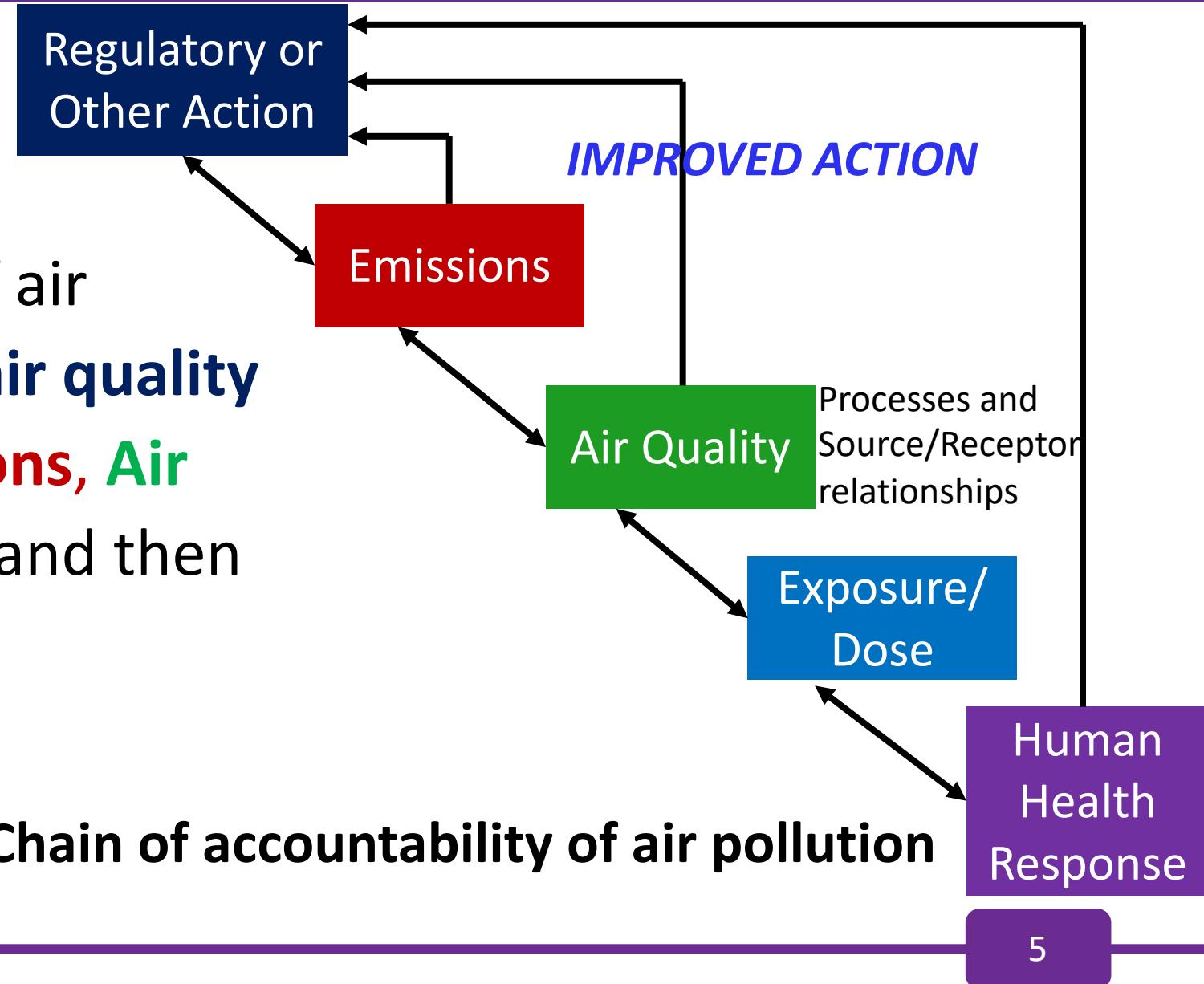
Air Pollution and Public Health



- **Majority** of pollution-related premature deaths is from air pollution.
- The **largest** environmental risk factor.
- UK, **~ 40K** premature deaths/yr could be attributed to air pollution.

Clean Air Strategy

Chain of accountability of air pollution describes how **air quality policy** can impact **Emissions**, **Air Quality**, **Exposure/Dose**, and then **Human Health**.



Clean Air Strategy

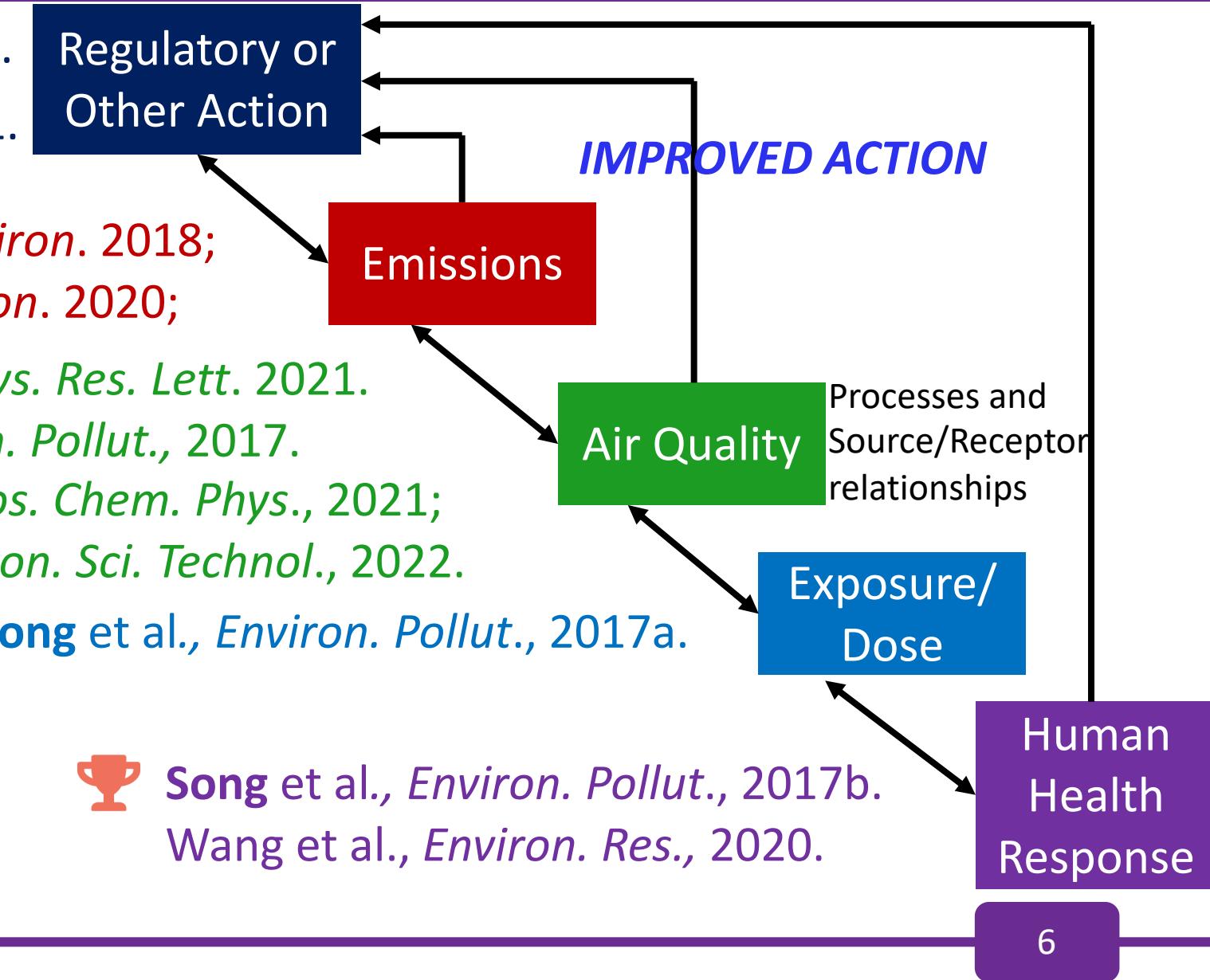
🏆 Shi and Song et al., *Sci. Adv.*, 2021.
Dai et al., *Geophys. Res. Lett.* 2021.

Song et al., *Sci. Total Environ.* 2018;
Song et al., *Atmos. Environ.* 2020;

🏆 Liu et al., *Geophys. Res. Lett.* 2021.
🏆 He et al., *Environ. Pollut.*, 2017.
Song et al., *Atmos. Chem. Phys.*, 2021;
Song et al., *Environ. Sci. Technol.*, 2022.

🏆 Song et al., *Environ. Pollut.*, 2017a.

🏆 Song et al., *Environ. Pollut.*, 2017b.
Wang et al., *Environ. Res.*, 2020.



Data Science Techniques for Air Pollution Studies

Data Science Techniques in analyzing sources and processes of air pollution

Positive matrix factorization (PMF) of VOCs and PNC.

Song et al., *Atmos. Environ.* 2019;
Dai et al., *Sci. Total Environ.* 2021.

Receptor modelling

Explainable machine learning technique.

Hou et al., *Environ. Sci. Technol. Lett.*, 2022.

Drivers of haze

Two-step SA for size-resolved aerosol.

Song et al., *Environ. Sci. Technol.*, 2022.

SA for wide-range PN

Multivariate Statistical Models and Machine Learning Techniques

Chemical mass balance (CMB) and Size-resolved chemical mass balance (SR-CMB).

Dai et al., *Atmos. Environ.* 2019.
Zhang et al. *Environ. Pollut.*, 2020.
Receptor modelling and improvement

k-means clustering of PVSD. Song et al., *Atmos. Chem. Phys.*, 2021.

Sources of coarse particles

A **coupling** receptor model and explainable machine learning technique.

Song et al., *Environ. Sci. Technol.*, 2022.
Drivers of source-specific aerosols

Data Science Techniques for Air Pollution Studies

Direction: Real-time Sources and Drivers of Particle Number Size Distribution?



pubs.acs.org/est

Article

¹ Understanding Sources and Drivers of Size-Resolved Aerosol in the High Arctic Islands of Svalbard Using a Receptor Model Coupled with Machine Learning

⁴ Congbo Song,* Silvia Becagli, David C. S. Beddows, James Brean, Jo Browse, Qili Dai, Manuel Dall’Osto, ⁵ Valerio Ferracci, Roy M. Harrison, Neil Harris, Weijun Li, Anna Jones, Amélie Kirchgässner, ⁶ Agung Ghani Kramawijaya, Alexander Kurganskiy, Angelo Lupi, Mauro Mazzola, Mirko Severi, ⁷ Rita Traversi, and Zongbo Shi*

Cite This: <https://doi.org/10.1021/acs.est.1c07796>

Read Online

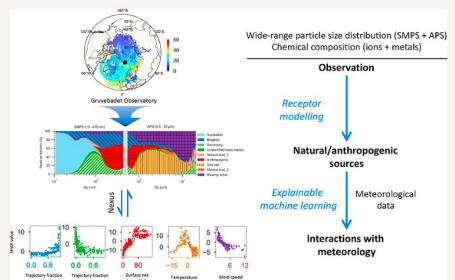
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

⁸ ABSTRACT: Atmospheric aerosols are important drivers of Arctic climate change through aerosol–cloud–climate interactions. However, large uncertainties remain on the sources and processes controlling particle numbers in both fine and coarse modes. Here, we applied a receptor model and an explainable machine learning technique to understand the sources and drivers of particle numbers from 10 nm to 20 μm in Svalbard. Nucleation, biogenic, secondary, anthropogenic, mineral dust, sea salt and blowing snow aerosols and their major environmental drivers were identified. Our results show that the monthly variations in particles are highly size/source dependent and regulated by meteorology. Secondary and nucleation aerosols are the largest contributors to potential cloud condensation nuclei (CCN, particle number with a diameter larger



<http://pubs.acs.org/journal/aescq>

ML for rapid SA

Article

Novel Application of Machine Learning Techniques for Rapid Source Apportionment of Aerosol Mass Spectrometer Datasets

Paritosh Pande, Manish Shrivastava,* John E. Shilling, Alla Zelenyuk, Qi Zhang, Qi Chen, Nga Lee Ng, Yue Zhang, Masayuki Takeuchi, Theodora Nah, Quazi Z. Rasool, Yuwei Zhang, Bin Zhao, and Ying Liu

Cite This: <https://doi.org/10.1021/acsearthspacechem.1c00344>

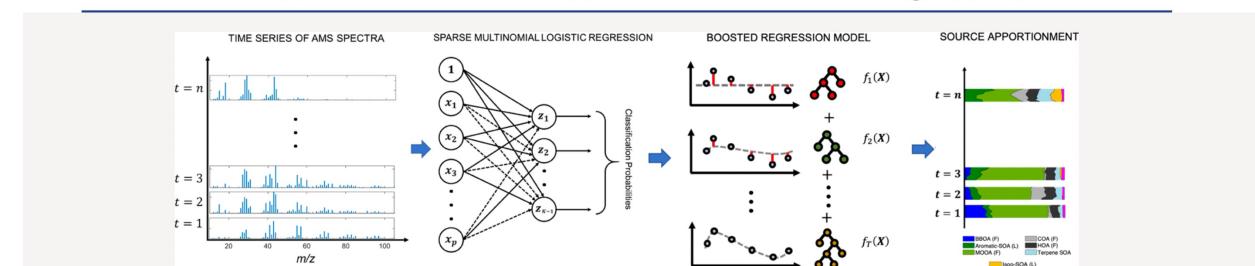
Read Online

ACCESS |

Metrics & More

Article Recommendations

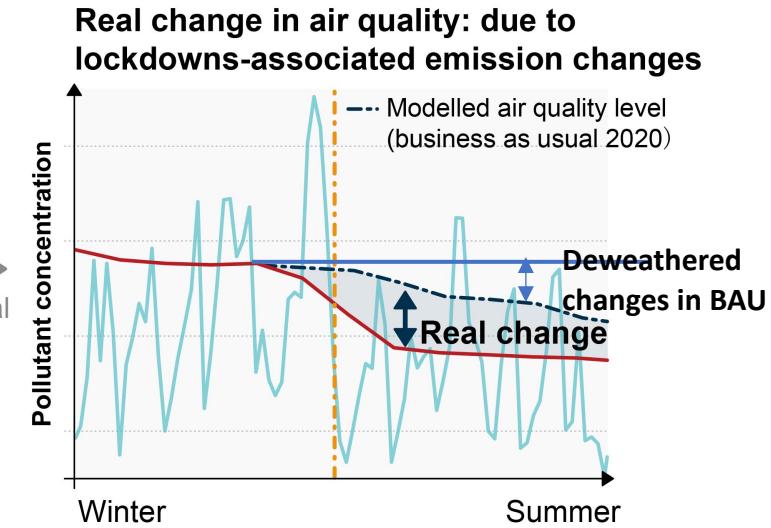
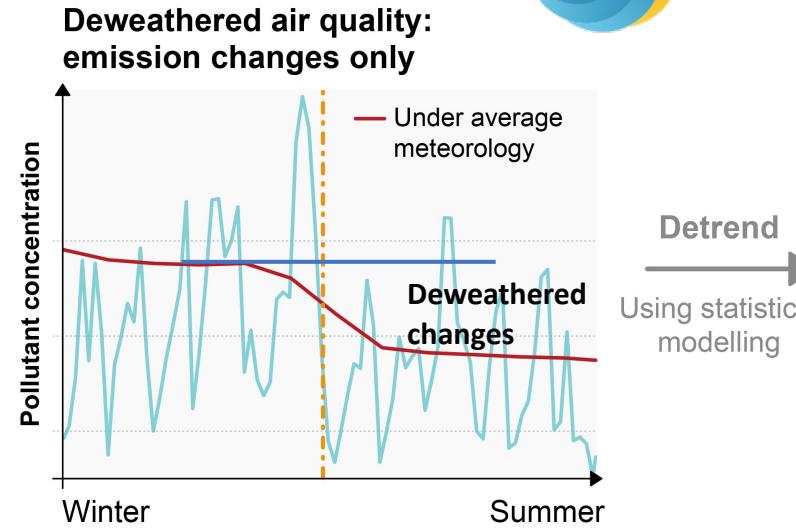
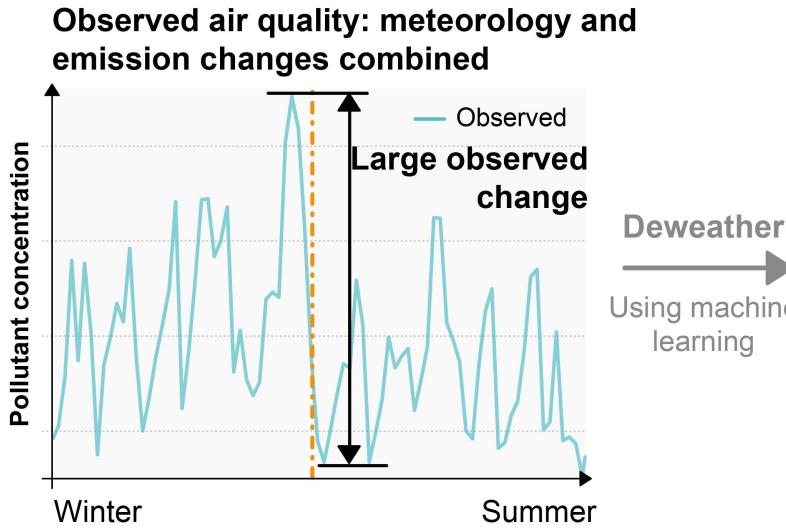
Supporting Information



Causality of air quality management



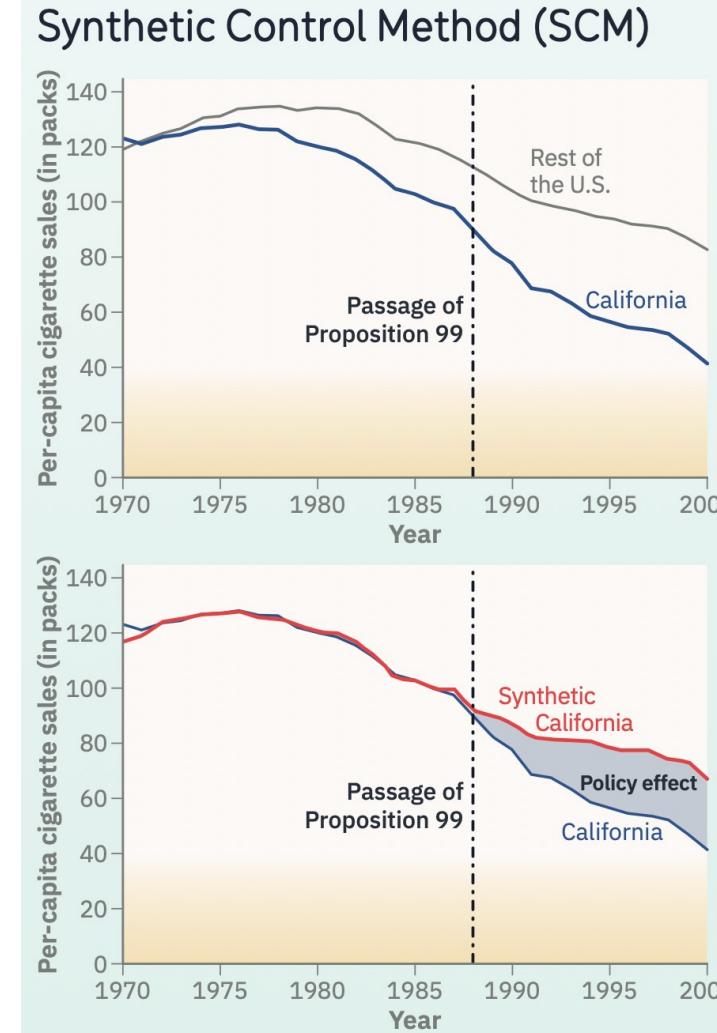
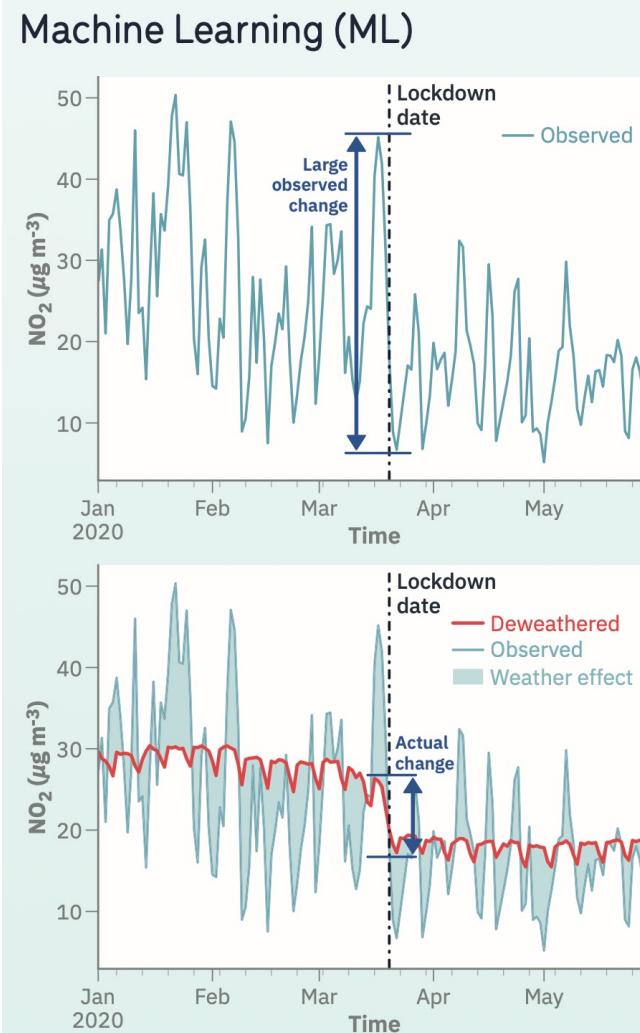
Shi and Song et al., Sci. Adv., 2021



- “Deweathering” plus “Detrending”
- **Corrected** numerous errors in published papers about impacts of COVID-19 lockdowns on air quality. ——Comments from a reviewer.
- Received much **attention** from researchers (97 citations) and media (Picked up by 51 news outlets, Blogged by 7, Tweeted by 129; Altmetrics: 550).

Causality of air quality management

Detrending → Synthetic Control Method (SCM)



Song et al., PNAS (under submission), 2022

Propose a coupling weather normalization and augmented synthetic control approach, which has a potential to be a **Golden-Standard** approach for Air Quality Intervention Studies and effective of **policies**.

A successful cross-disciplinary study with **Economists**.

Near/Longer Term Vision

Near term

- Support development of digital twins and urban observatory.
- Integrate observatory resources in Manchester (MUO) and Birmingham (BAQS and AQST) to understand causes and changes in air pollution in the age of net-zero emissions;
- Funding to support my research.

Longer term

- A leadership in understanding sources and processes of air pollution with the aid of data-driven models;
- A leadership in air quality intervention studies;
- Cross-disciplinary studies with public health, economics, policy, etc.

Future Funding Sources from UK, US and China



Natural
Environment
Research Council



Engineering and
Physical Sciences
Research Council



Medical
Research
Council



Economic
and Social
Research Council

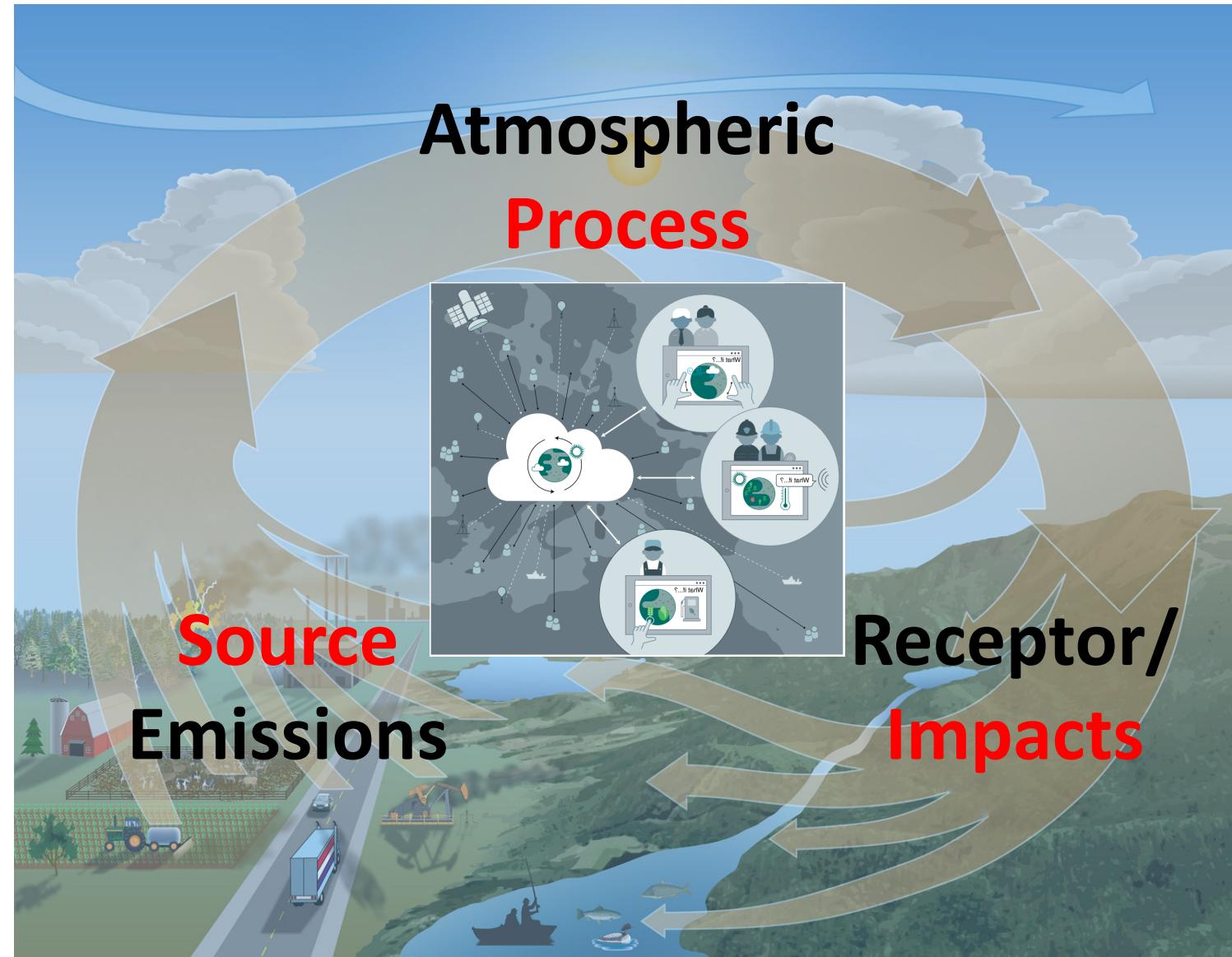


Health Effect
Institution



National Natural
Science Foundation
of China

Extensive Experience in Data-driven Models



- Extensive experience in applying data science techniques in understanding sources, processes and impacts of air pollution.
- Support development of urban observatory and digital twins.

Developing supersites and urban observatory

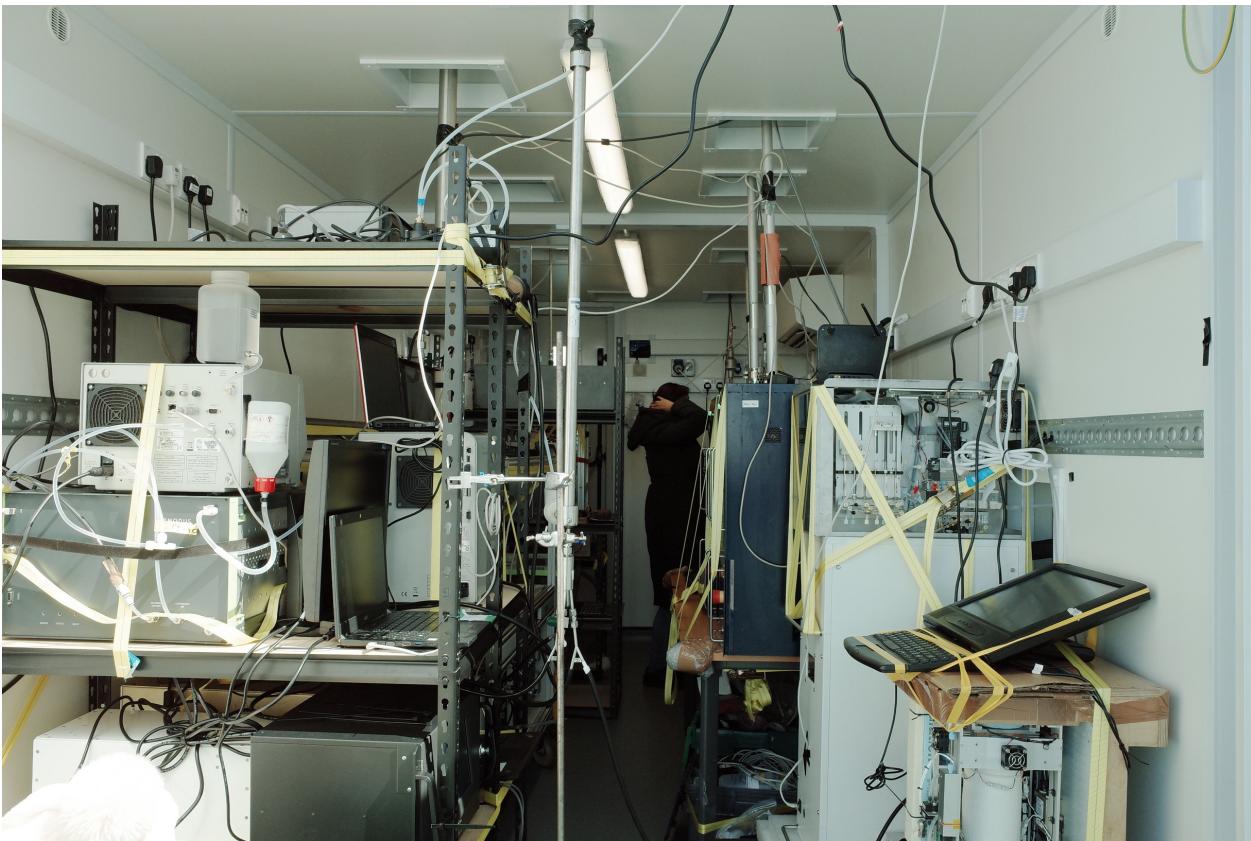


Actively involved in WM-AIR, OSCA projects to support measurements
@ Birmingham Air Quality Supersite (BAQS)



Improve measurement/data capability
@ Manchester Urban Observatory (MUO)

Leadership in large field campaigns



- Successfully **coordinated** a SEANA research cruise.
- Prepare, Install, Setup and Maintain a **large instrument suite** in the mobile air quality supersite (AQST) on ship.



UK Air Quality Supersite
Triplet (AQST)

Global Collaborators

Cross-disciplinary Collaborations with Colleagues in the UK, Europe, US & China



SINCE 2020,
22 PUBLICATIONS

NCAS's Mission & Vision

Three problems to tackle in the next five years.

*"Our five-year strategy sets out how we will unlock the answers to the **biggest environmental issues** of our time. Our research, skills and infrastructure will bring benefits to the UK, such as **clean air solutions**, resilience to **climate change**, and early-warning systems for high-risk weather."*

Air Pollution

Prof. Huge Coe

Prof. David Toppings

Prof. Gordon McFiggans

Dr. James Allen

Dr. James Lee

...

Chain of accountability

Prof. David Toppings

Dr. James Lee

Dr. James Allen

...

Net-zero emissions and lower uncertainties in Aerosol-Climate interaction

Prof. David Toppings

Prof. David Schultz

Prof. Ann Webb

Prof. Mike Burton

...



UNIVERSITY OF
BIRMINGHAM



CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



Thanks!

SEANA – Shipping Emissions in the Arctic and North Atlantic atmosphere

Air Quality Supersite Triplets (**UK-AQST**)

COP-AQ: UK-China collaboration to optimise net-zero policy options for air quality and health

West Midlands Air Quality Improvement Programme (**WM-AIR**)

Integrated Research Observation System for Clean Air (**OSCA**)

Impact of COVID-19 lockdown on air quality in selected megacities

Winter heating and air pollution in northern China



**Natural
Environment
Research Council**



UNIVERSITY OF
BIRMINGHAM

THE INSTITUTE FOR
GLOBAL INNOVATION