Imperial College London

Hybrid AI and physical modelling for accurate and rapid environmental prediction and management

Fangxin Fang, Christopher Pain, Xiaofei Wu, Shengjuan Cai, Meiling Cheng, Boyan Cheng, Yanghua Wang, Jinxi Li and Jie Zheng

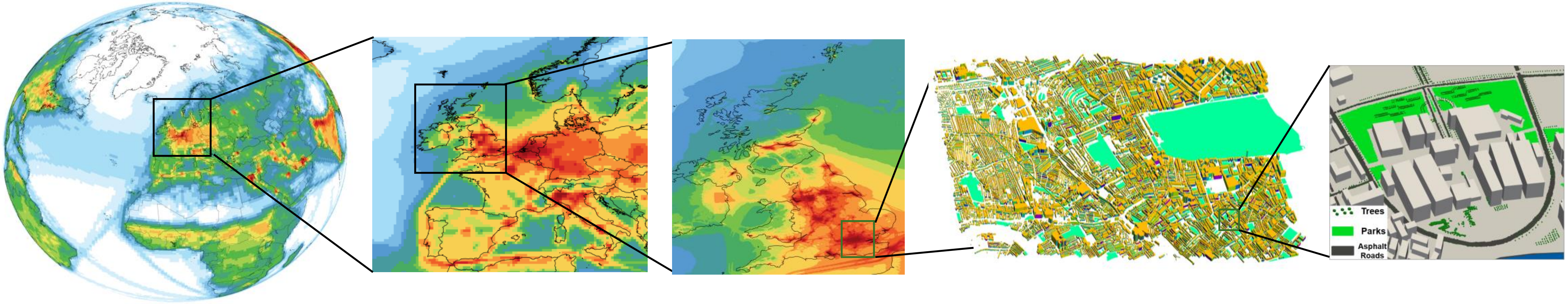
ECO-AI, Heriot-Watt University,12 March 2024

DEPARTMENT OF
EARTH SCIENCE AND ENGINEERING

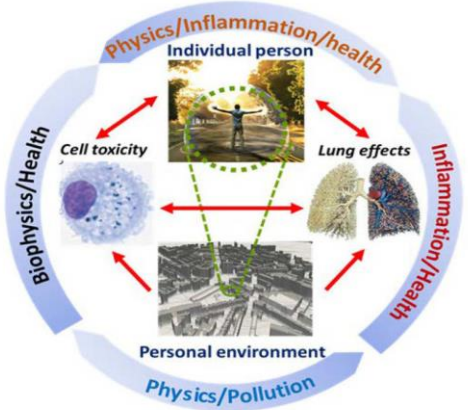
Digital tools for Urban Environment Management

Decarbonisation combating climate change– NetZero by 2050

Global scale | Regional scale | City scale | Neighbourhood scale | Building scale



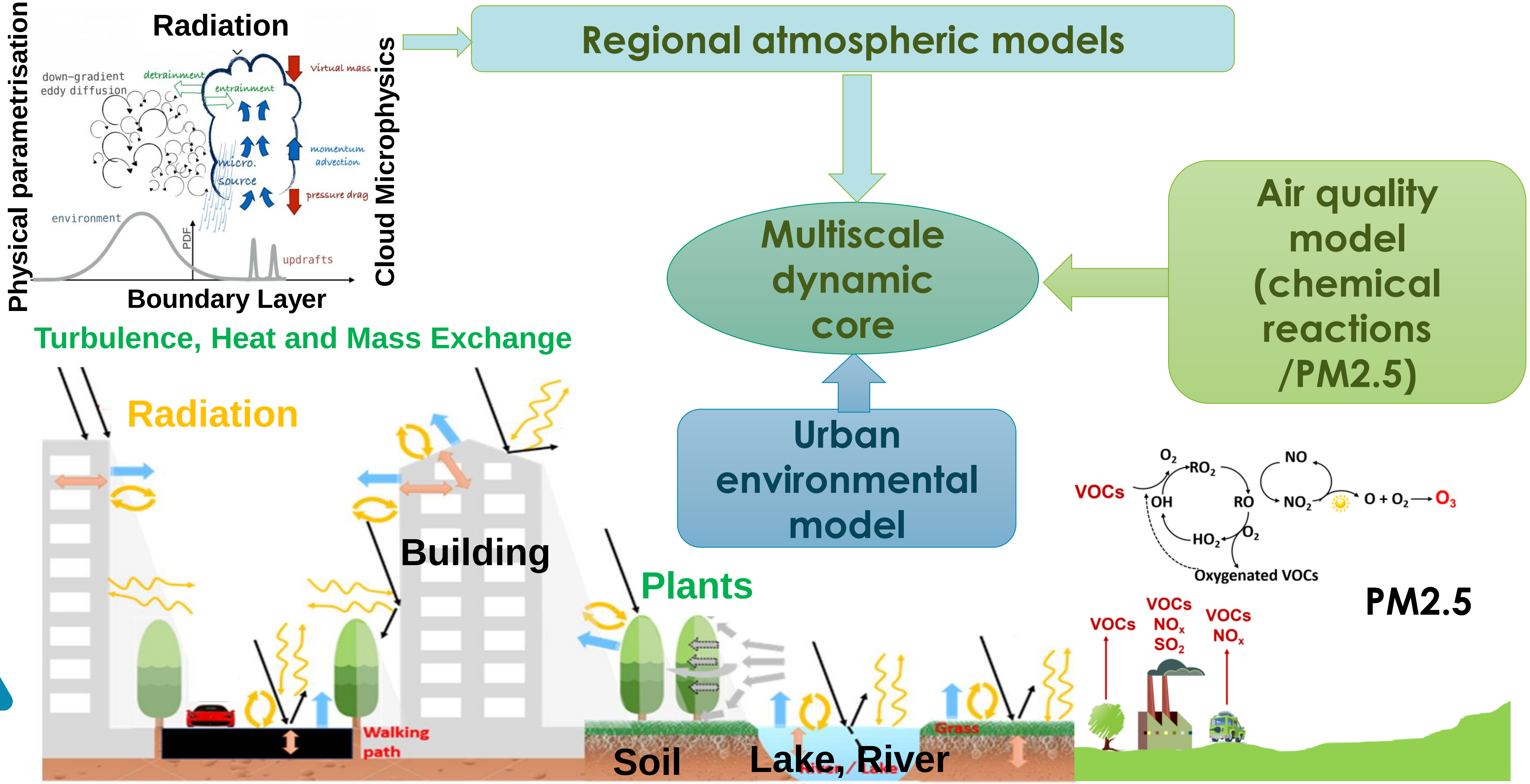
Understanding of relationship between health, economics, environment and climate change



Digital-twin operational tools for environment and energy management, which enable the urban population (as well as policy makers) to make both strategic and everyday decisions that help generate a zero pollution environment by 2050

Complex physical processes

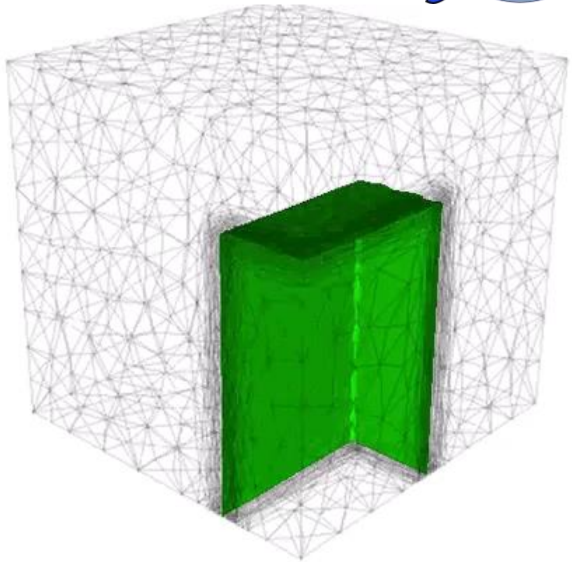
In Atmosphere and Urban Environments

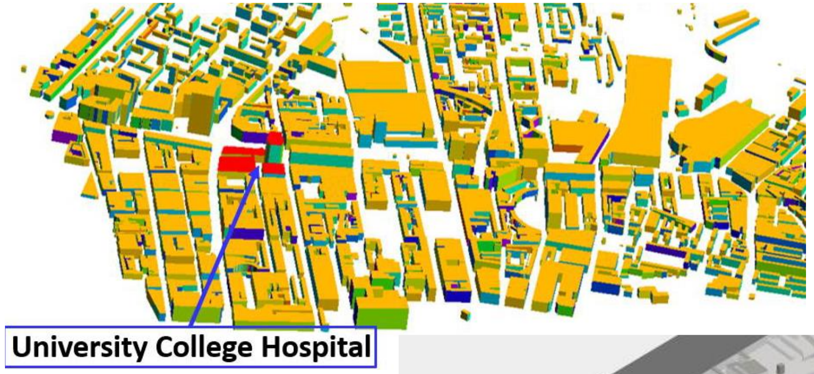


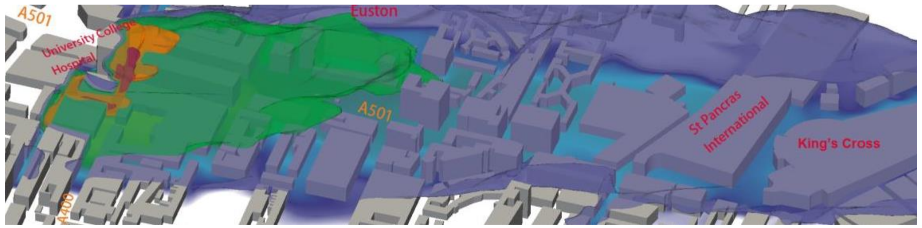
Physical Modelling Issues to be resolved	Hybrid-methodologies	Data science Issues to be resolved	Objectives Improved accuracy
Uncertainties and parameters in models	Machine learning, Back-propagation Adjoint uncertainty sensitivity, Goal-based approach	**Uncertainty in big datasets**	• **Uncertainty quantification** • **Identify pollutant sources**
Empirical subgrid models	Data driven model replaces empirical subgrid models in physical modelling, Data assimilation	**Data-driven modelling**	• **Uncertainty quantification** • **More accurate subgrid models**
Model error	Data assimilation, goal-based approach	**Error in datasets**	• **Reduce the misfit between modelling results & measurements**
	Introducing physical modelling to training	**Lacking dynamic knowledge**	• **Spatio-temporal data driven prediction**
CPU time	Rapid detailed machine learning models		

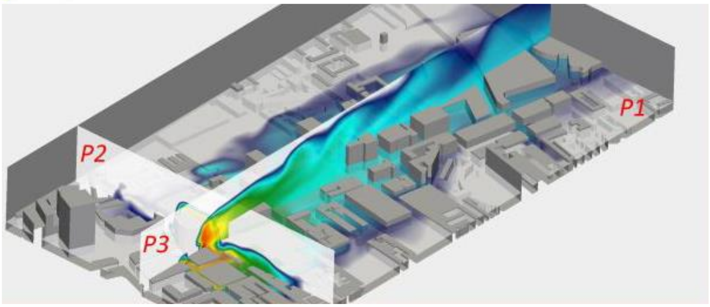
Introduction of an adaptive unstructured mesh fluid model – Fluidity

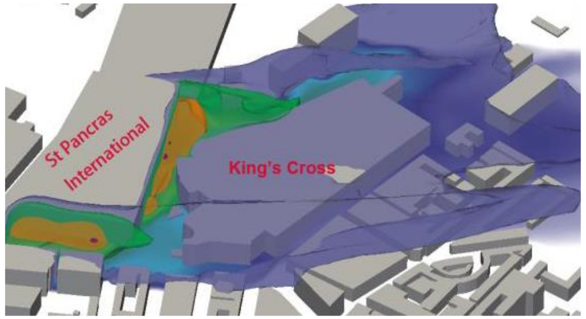
- Open Source Model Software for Multiphysics Problems
- Unstructured FEM Meshes
- Large Eddy Simulation (LES)
- Anisotropic Adaptive Mesh technology
- User-friendly GUI
- Python interface to calculate diagnostic fields, to set prescribed fields and user-defined boundary conditions

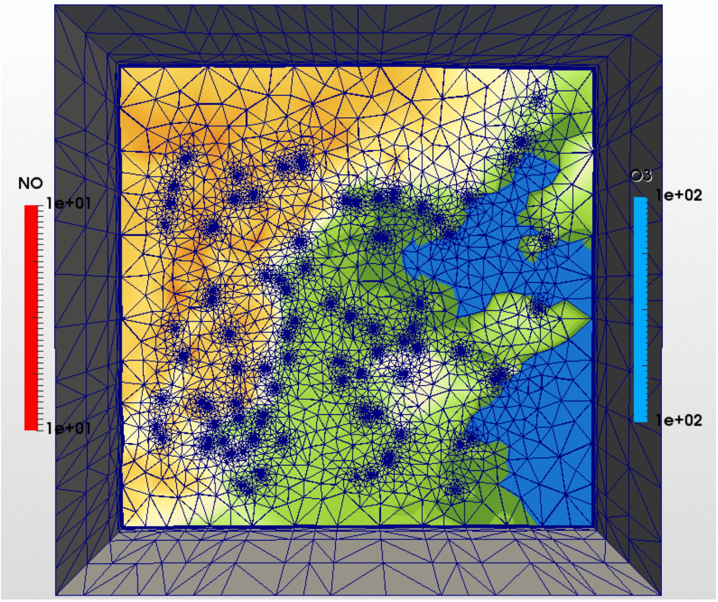












Physical Model
Completed Research

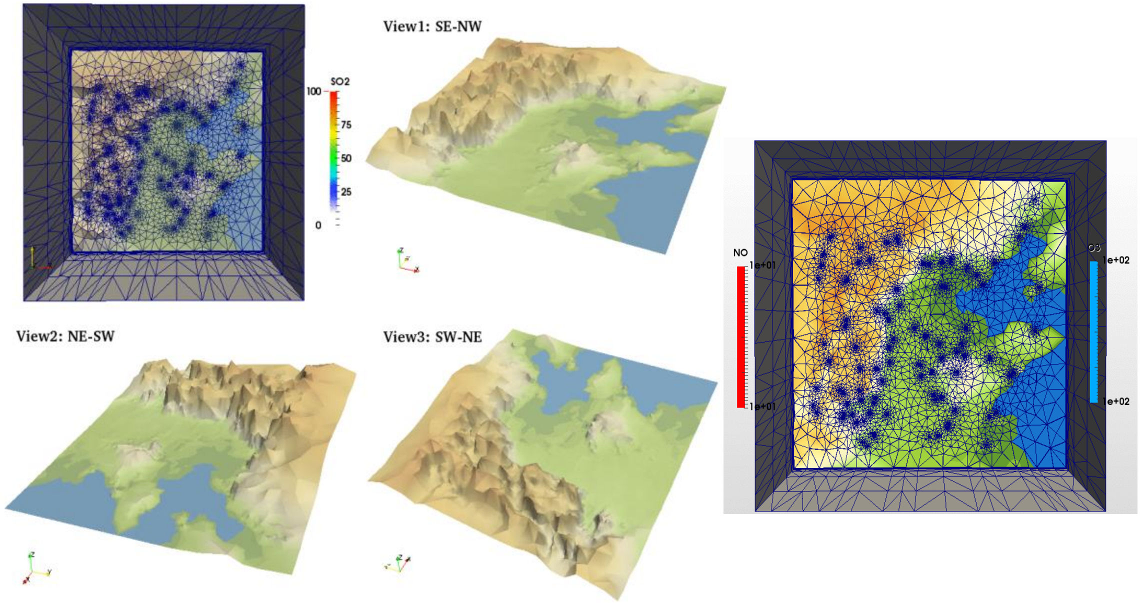
Imperial College London

Model

Regional Atmosphere

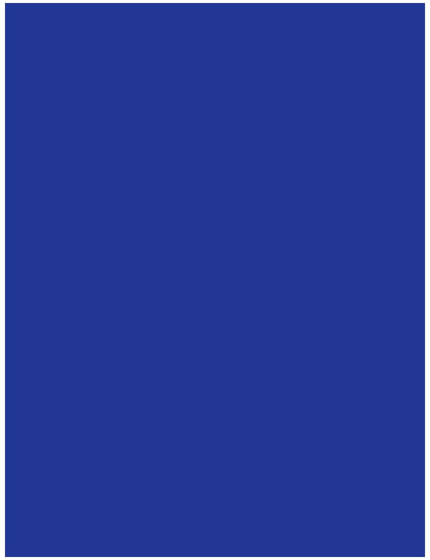
Air quality

Novelty: Provide a single united integrated model for resolving chemical and atmospheric processes over a wide range from meters up to kilo-meters. **Software:** Fluidity-Chem, Fluidity-Atmos

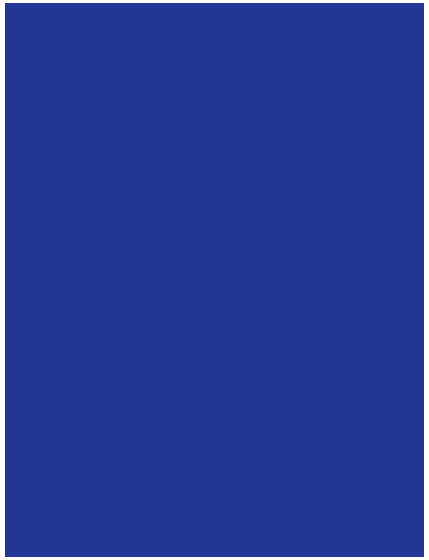


SO_2 released from over 100 power plants (left) and Chemical modelling, NO2 and O3 (right)

Cloud Water Mesh



Rainwater Structure



Physical Model
Completed Research

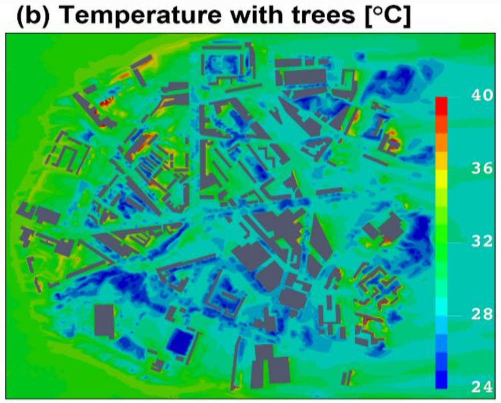
MAGIC
Envisaging a world with greener cities

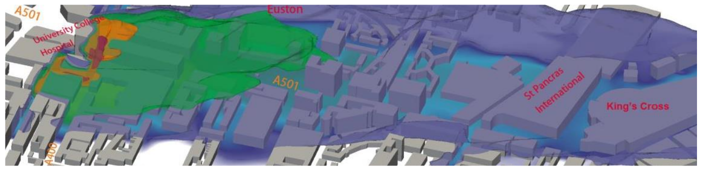
INHALE

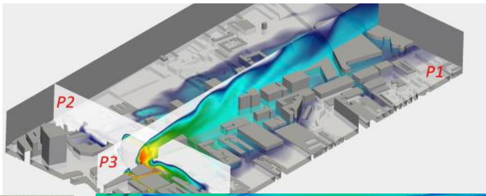
Model

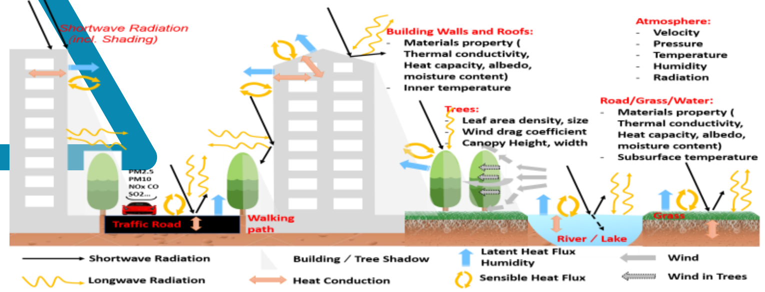
Urban environmental

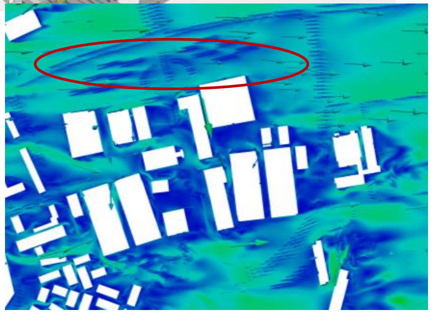
Novelty: Provide a high resolution spatial distribution of pollutants, temperatures humility by incorporating the impact of green-blue infrastructures, radiation and thermal dynamics.
Software: Fluidity-Urban, and the 3D urban generator.











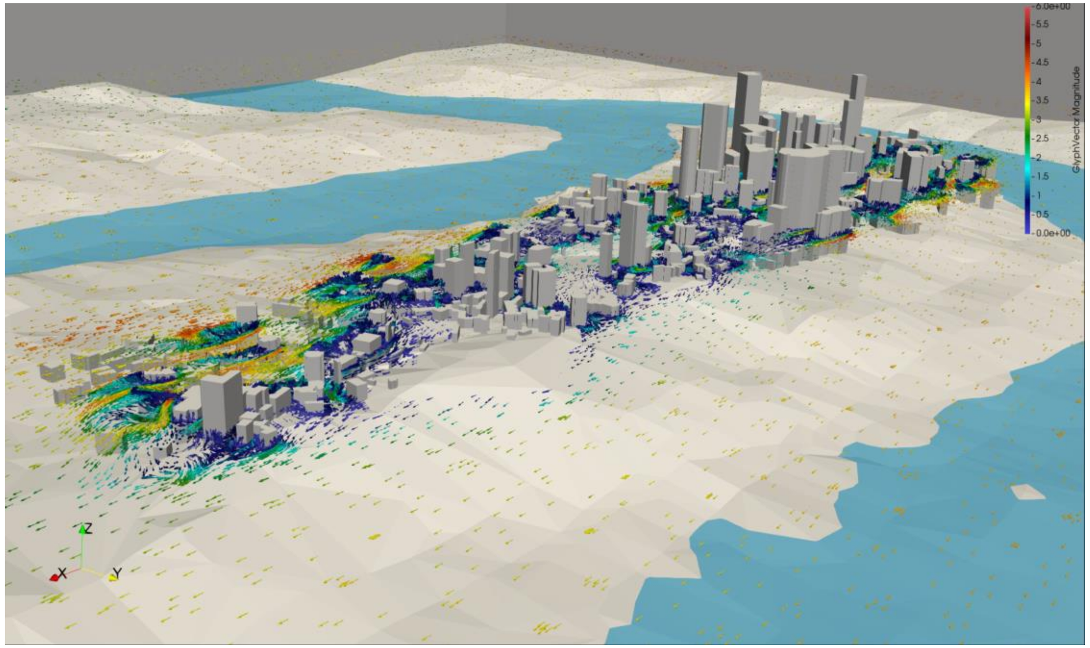
Physical Model
Completed Research

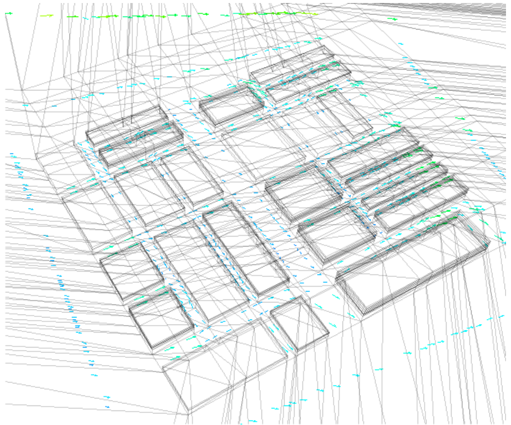
MAGIC
Envisaging a world with greener cities

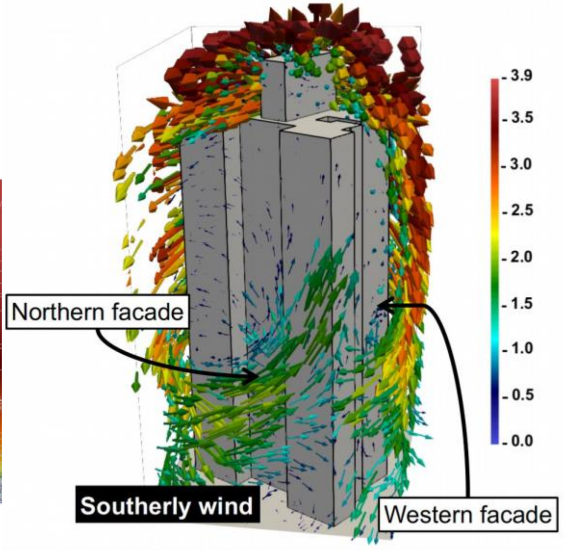
INHALE

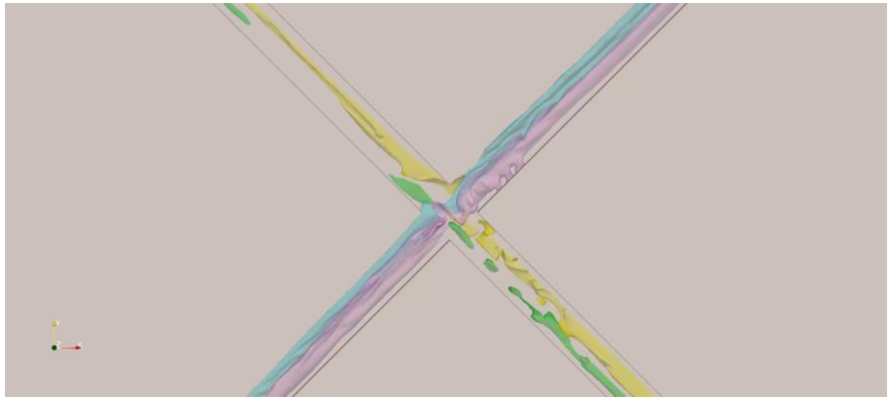
Model

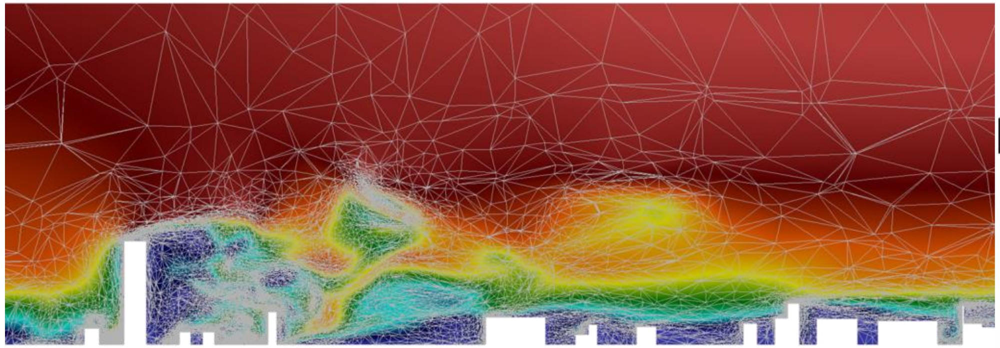
Urban environmental





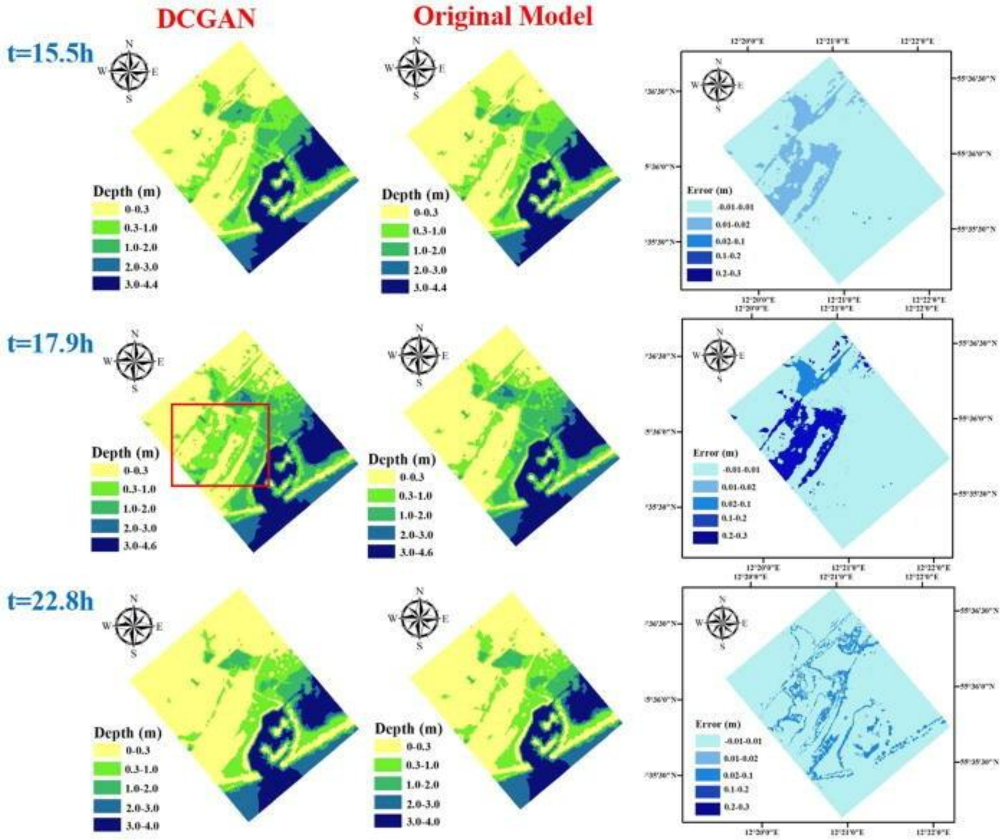


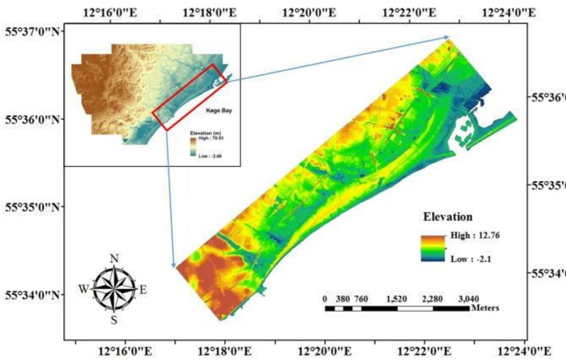




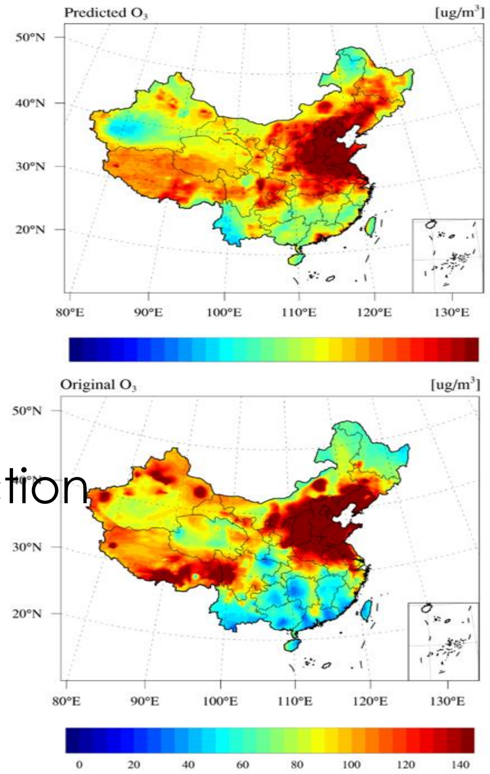
Spatial-temporal prediction using hybrid-machine learning and physical informed modelling & data assimilation (Dr. Cheng)

- Machine learning based rapid response tools for real-time operation prediction and uncertain analysis
- Sub-grid physical parameterization schemes in atmospheric modelling
- Real time air pollution forecasting at high spatial and temporal scales
- Machine learning-based coupling of multiple scale models from large (national/region), city to street sales
- Real-time operational tools for urban environment (traffic, green/blue, indoor/outdoor)





Real-time flooding prediction



Real-time ozone spatial distribution map in China

Machine Learning – Challenges

Spatiotemporal Forecasting via Machine Learning Models

- **High Dimensionality:** Spatiotemporal data is often high dimensional, which can make it difficult to train machine learning models that can capture the complex relationships between different variables.
- **Spatial and temporal correlations:** It is challenging to train models that can accurately capture both the spatial and temporal dynamics from data, specially from sparse monitoring measurements.
- **Complex topographic and meteorological conditions** resulting in highly variable spatial and temporal patterns of variables (e.g., PM2.5), making forecasting challenging at a high spatial resolution.
- **Predictive accuracy:** ML models lack of interpretability, thus it may lost its accuracy outside the range of data.

Machine Learning – Case Study

Imperial College London

Machine Learning (ML) & Reduce Order Modelling (ROM)

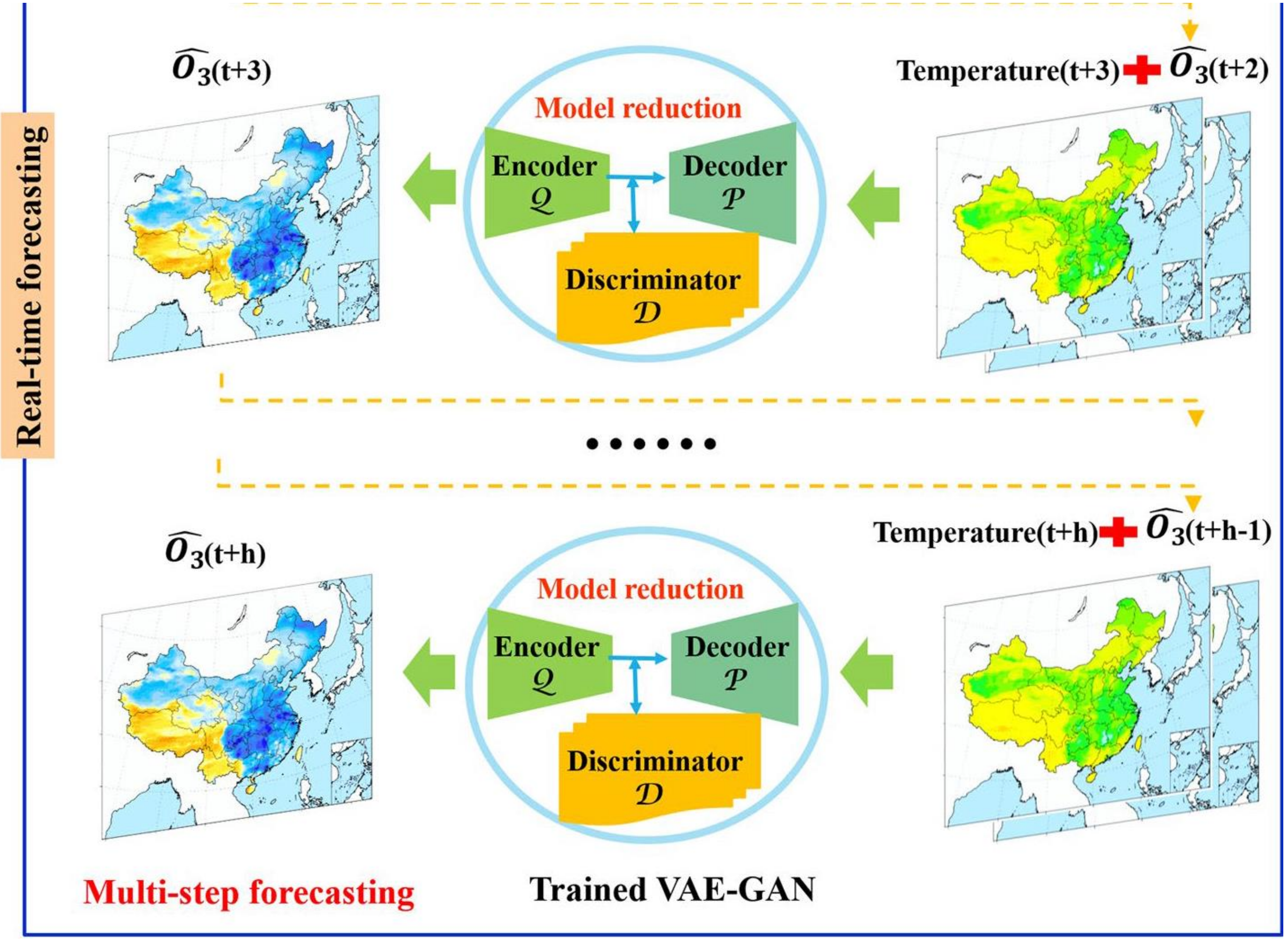
Spatio-temporal Hourly and Daily Ozone Forecasting in China

Method: Autoencoder and Generative Adversarial Networks

The reanalysis ozone datasets from 2013 to 2018 over China are used for processing different training and prediction scenarios

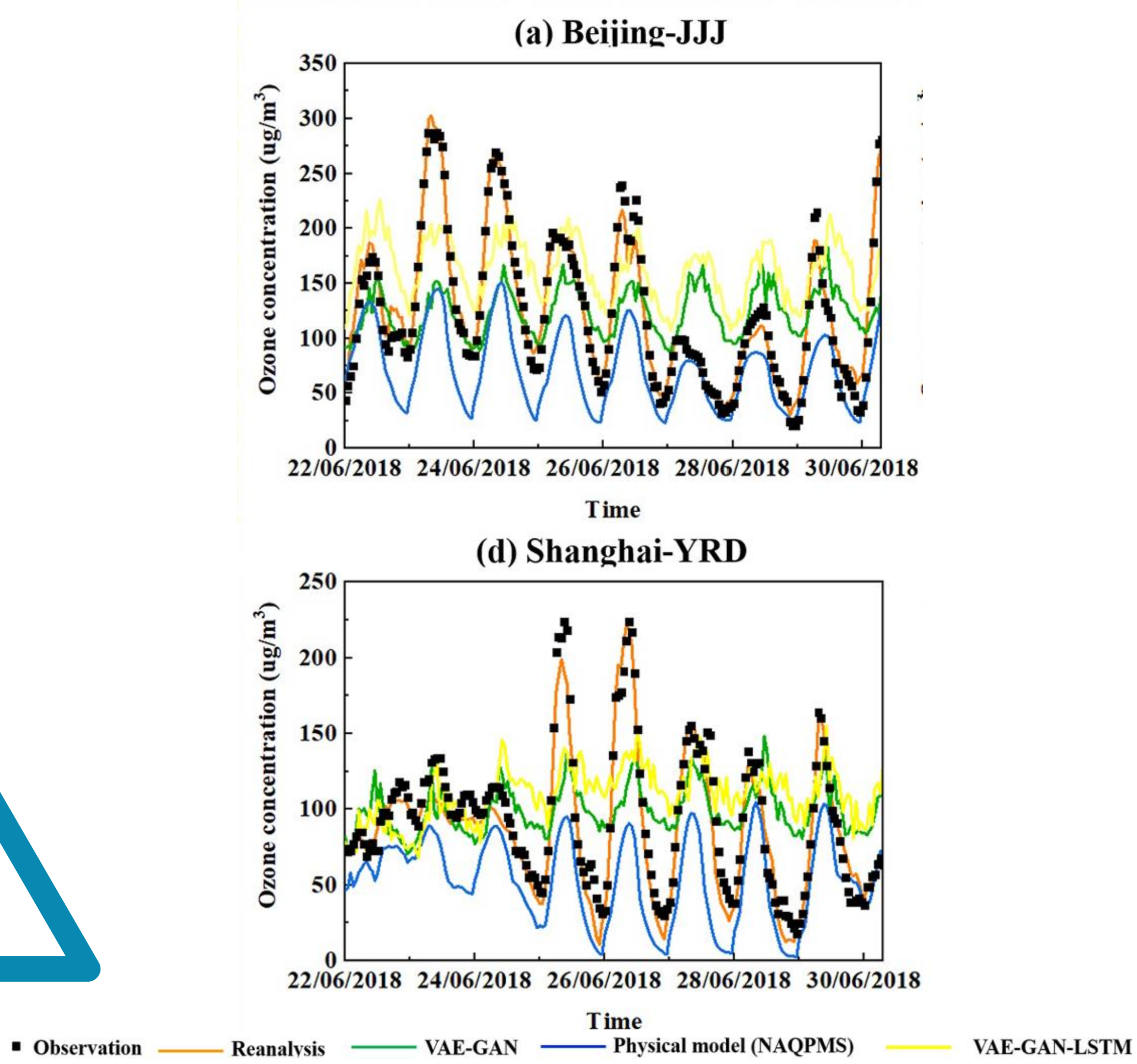
Inputs: Meteorological data (temperature, humility, wind speed etc) and the ozone concentration from the previous time levels

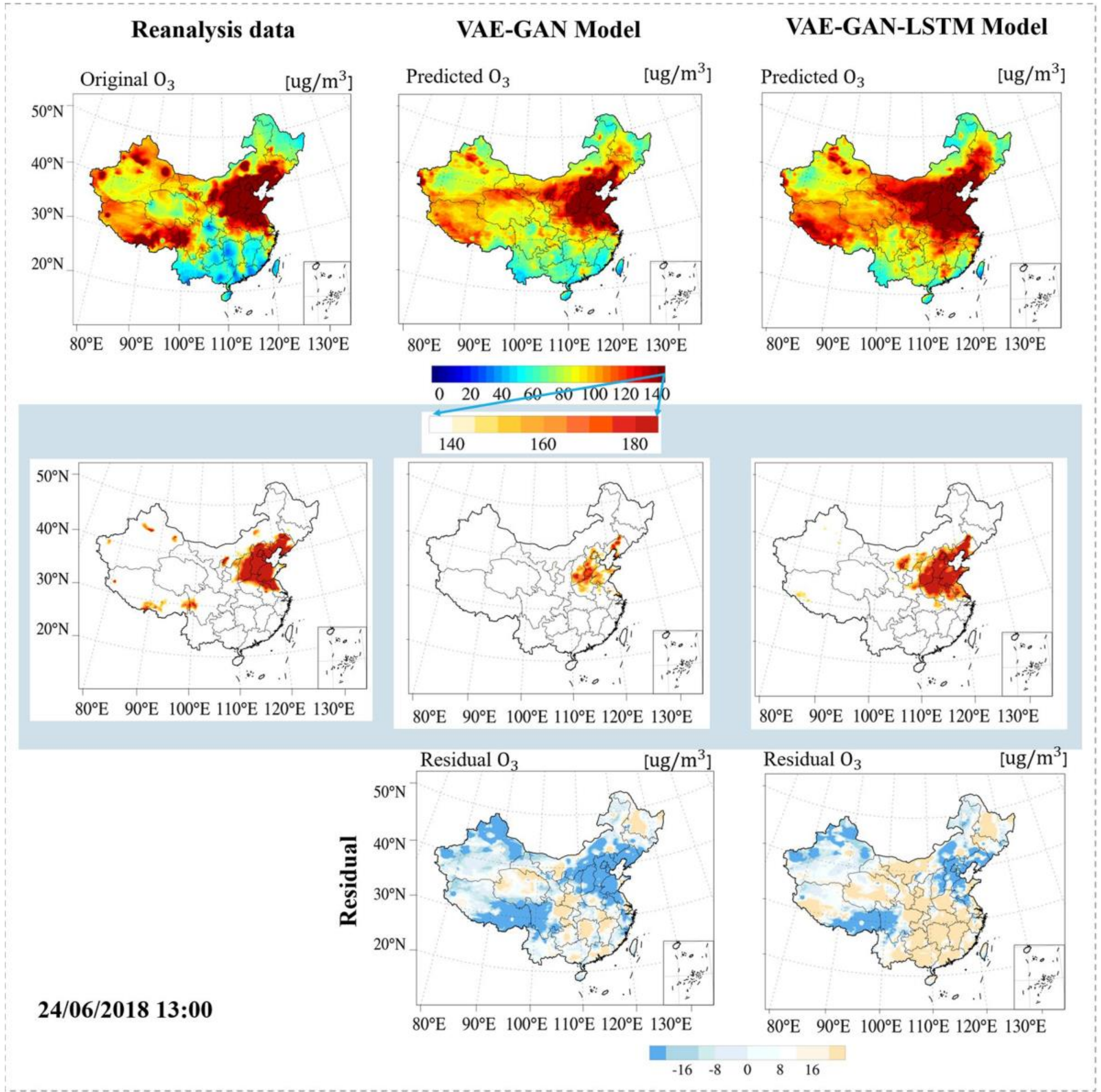
Output: Ozone at future time levels



Machine Learning – Case Study - Results

Machine Learning (ML)





Machine Learning – Case Study – Challenges

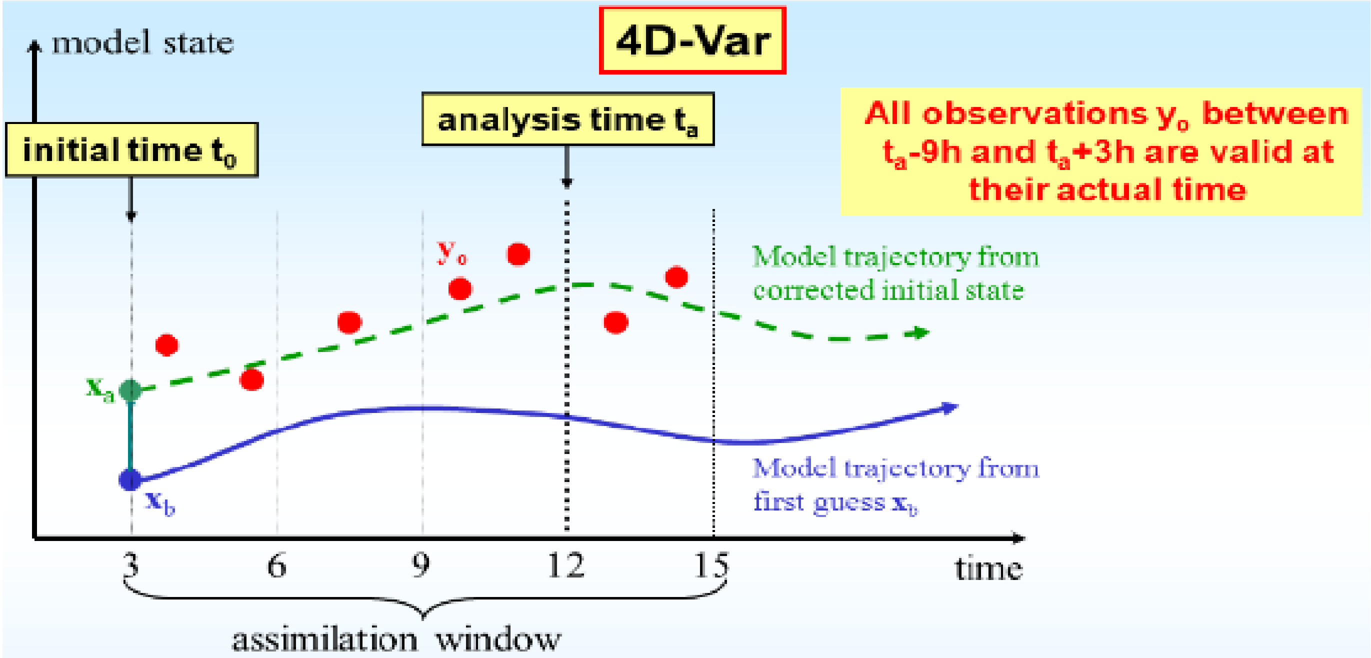
Biggest issues around long term forecasting

- It is beyond the range of the training data (beyond the training period)
- Complex nonlinear physical processes (uncertainties)
- Gradual accumulation of errors in long-term forecasting

To tackle these issues, data assimilation was introduced to ML-based model

Long term forecasting

Data Assimilation (DA)



- To improve the predictability of numerical models
- Uncertainty sensitivity analysis
- Optimisation of uncertainties in models
- Goal-based error measure and mesh adaptivity
- Design optimisation
- Adaptive observation (Optimisation of sensors locations)

Long term operational forecasting – Challenges
Traditional Data Assimilation (DA) model

- **Nonlinear Dynamics:** Spatiotemporal forecasting models often exhibit nonlinear dynamics, which can make it challenging to accurately assimilate data into the model.

- **Computationally expensive calculation**: Data assimilation for spatiotemporal forecasting can be computationally intensive, especially when dealing with high-dimensional data and complex models.

- **Sparse measurements:** Sparse observations used in the assimilation process can impact the accuracy of the assimilation process and difficult to capture the spatial and temporal dependences.

- **On-line data assimilation at a high spatial resolution** due to computationally expensive calculation

- **Model-Data Mismatch:** There can be differences between the model predictions and the observations used in the assimilation process, which can lead to inaccuracies in the assimilation process.

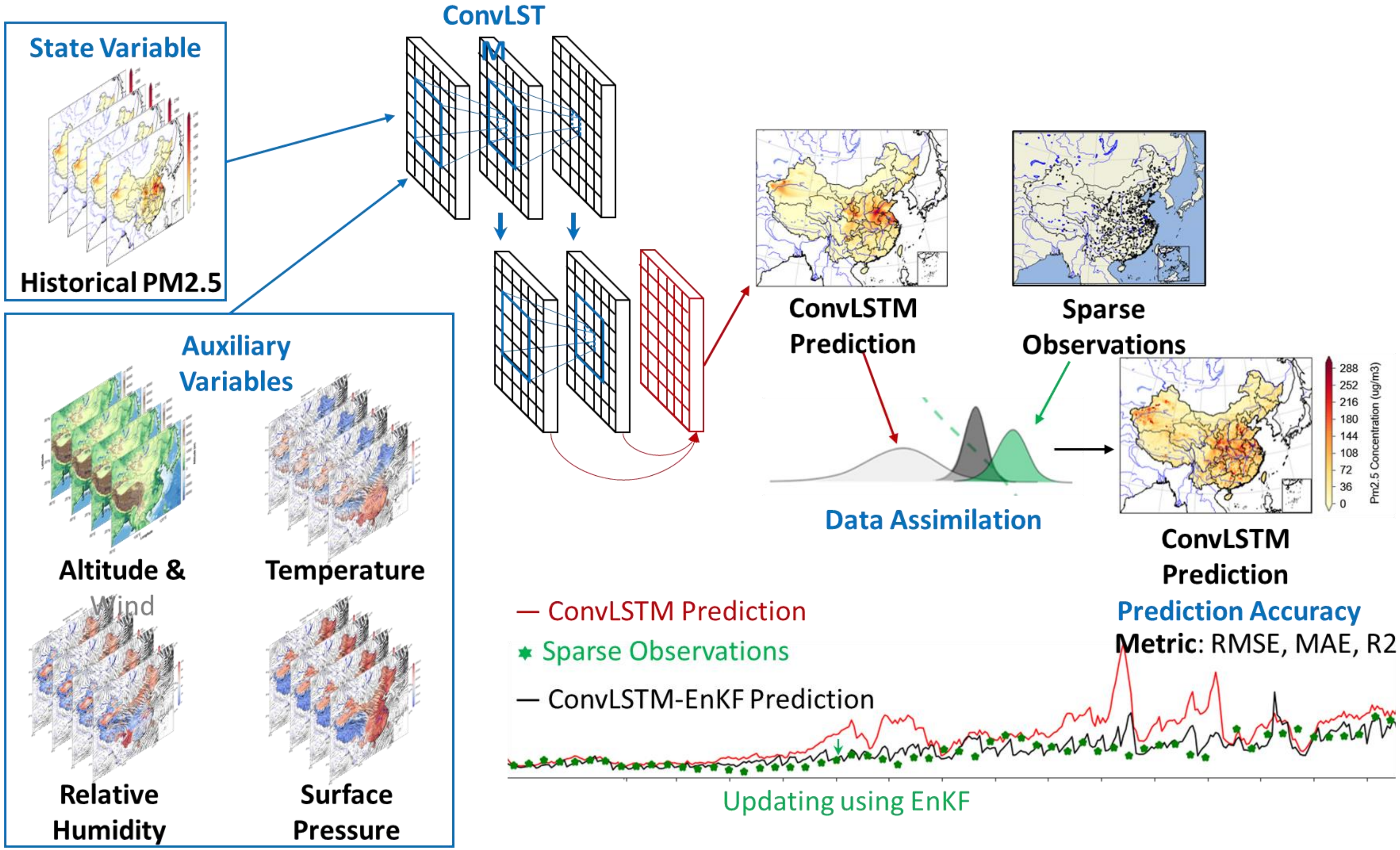
Long term forecasting – Improvements

Machine Learning (ML) & Data Assimilation (DA) model

- **ML-based long-term spatiotemporal forecasting** by updating initial conditions with incorporating data
- **Efficiency and accuracy of forecasting and data assimilation** by using ML methods
- **On-line data assimilation at a high spatial resolution** with sparse observations
- **Application** of the hybrid ML-DA in PM2.5 forecasting over China

Long term forecasting – Case Study (Ms Cai)

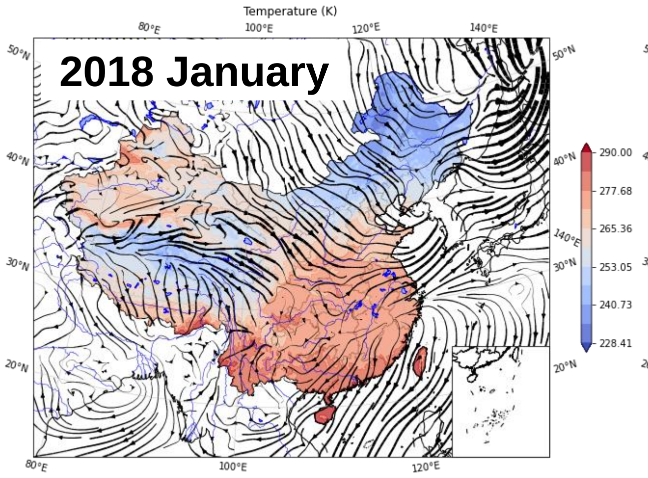
Machine Learning (ML) & Data Assimilation (DA) model [Modelling Process]



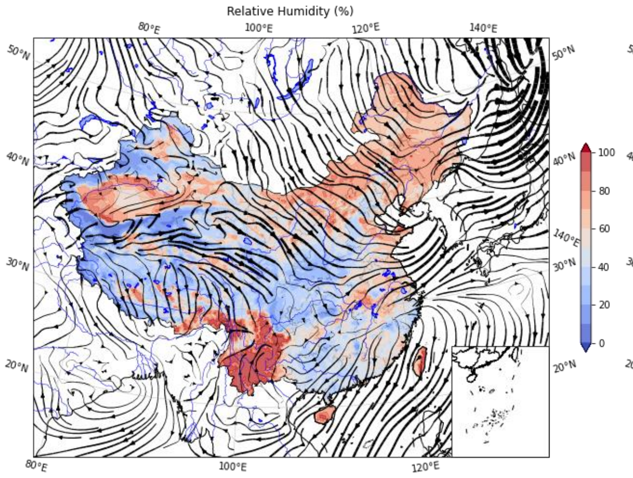
Long term forecasting – Case Study

ML & DA model – Reanalysis dataset

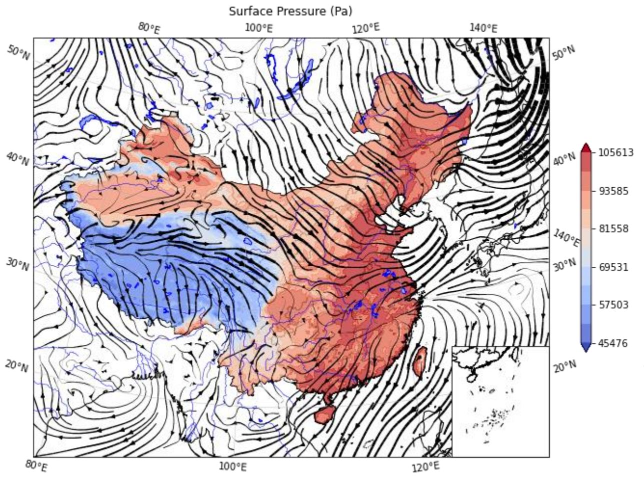
Temperature (K)



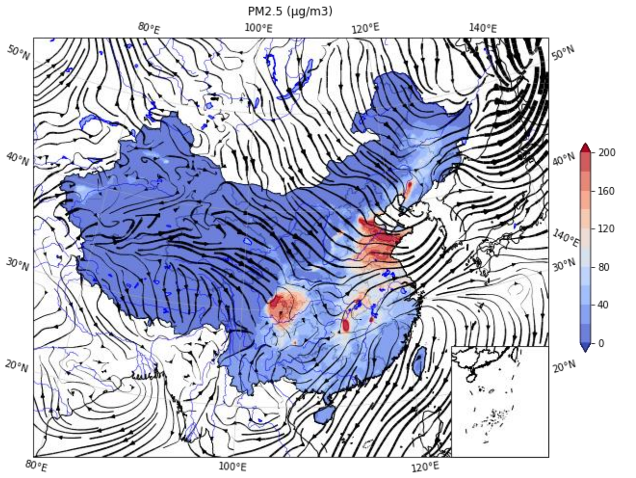
Relative Humidity (%)



Surface Pressure (Pa)



PM2.5 (ug/m3)



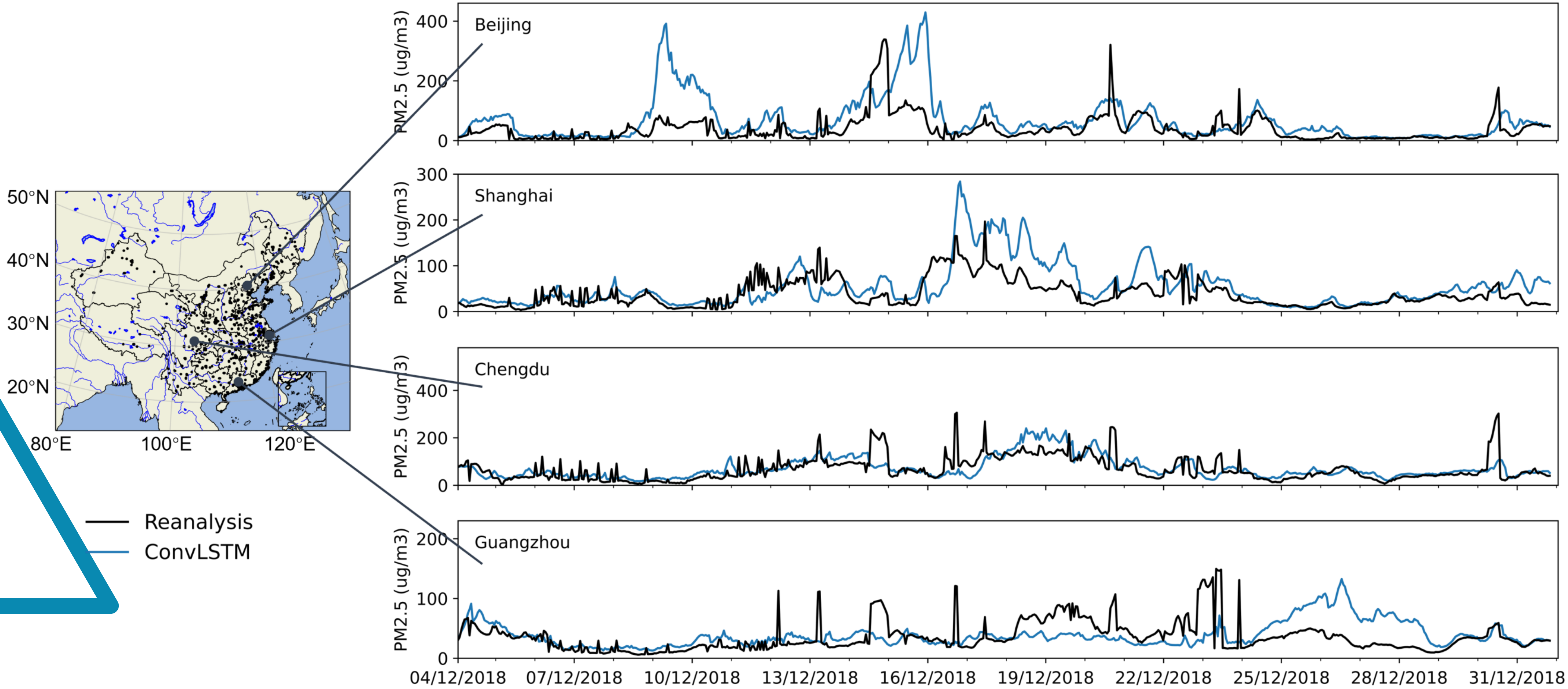
- Reanalysis data = EnKF (physical simulation, surface observation)
- Integrated physical models (WRF, NAQPMS) and observations
- High spatial resolution: 15km x 15km (339, 430)
- High temporal resolution: 1h (61344)

Training (90%) + validation (10%) :   2013-2018

Predicting: 2019

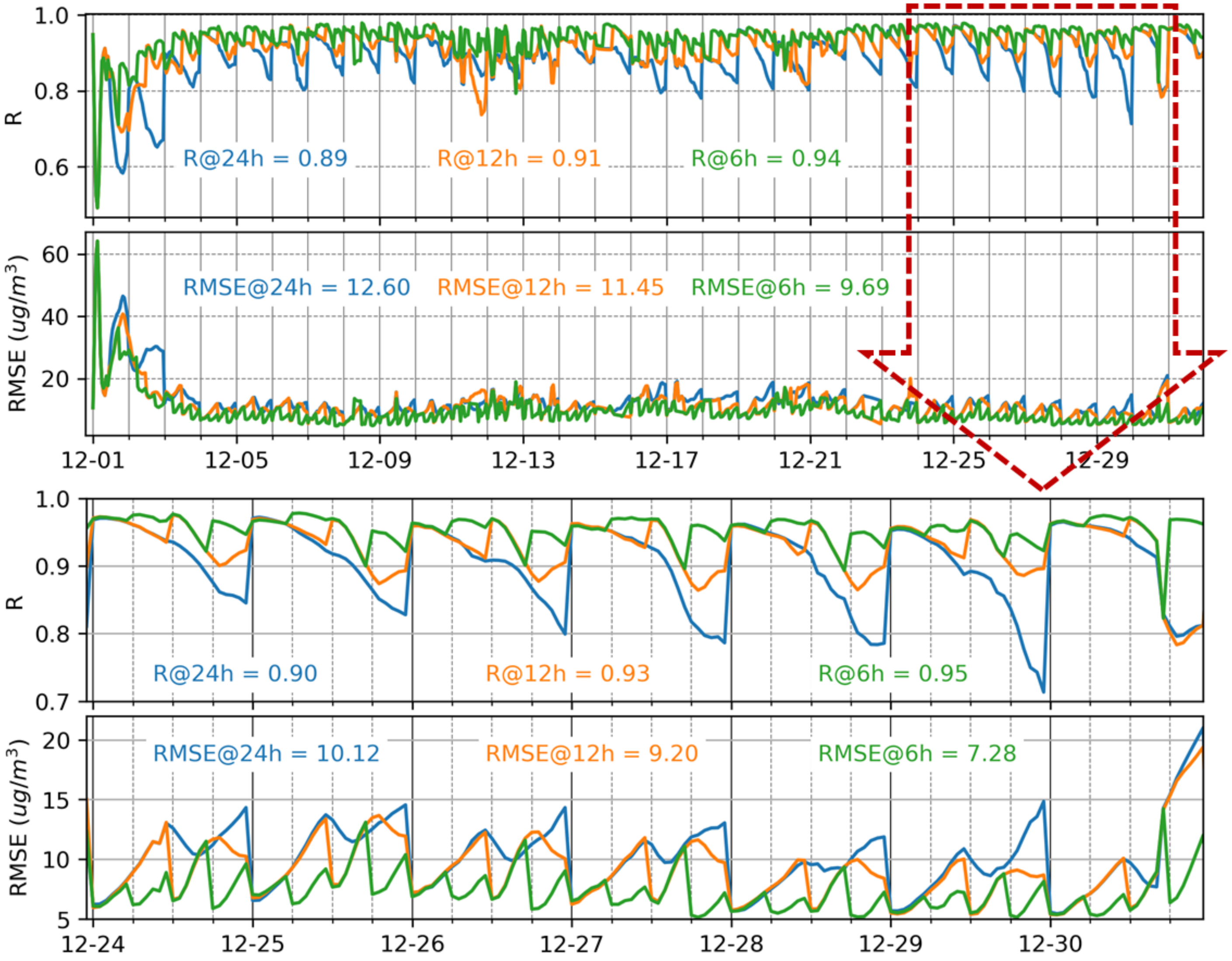
Long term forecasting – Case Study

ML & DA model – iterative multiple-hour forecasting (error accumulation – without DA)



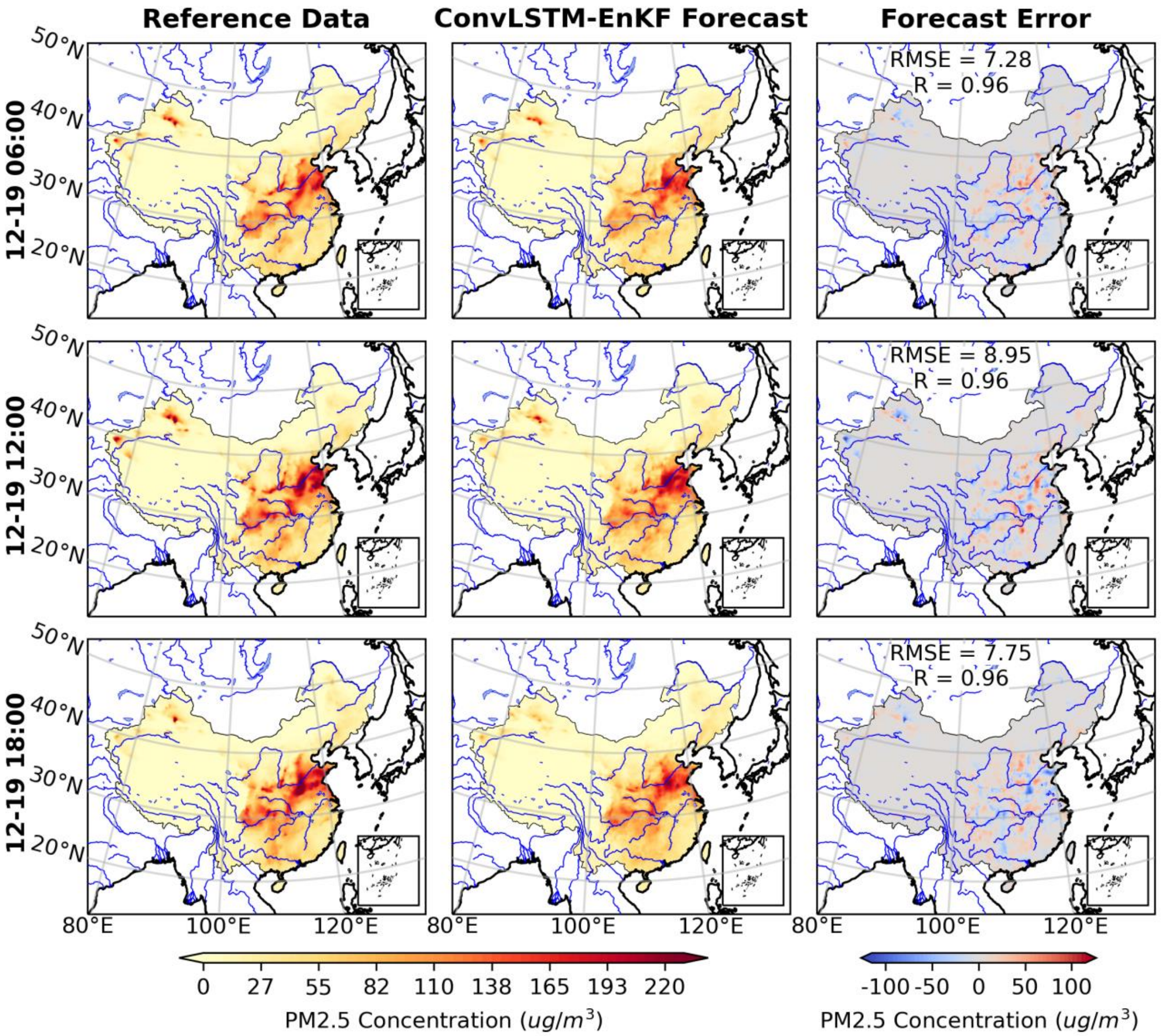
Long term forecasting – Case Study

ML & DA model – virtual spatial uniform observations



Long term forecasting – Case Study

ML & DA model – virtual spatial uniform observations

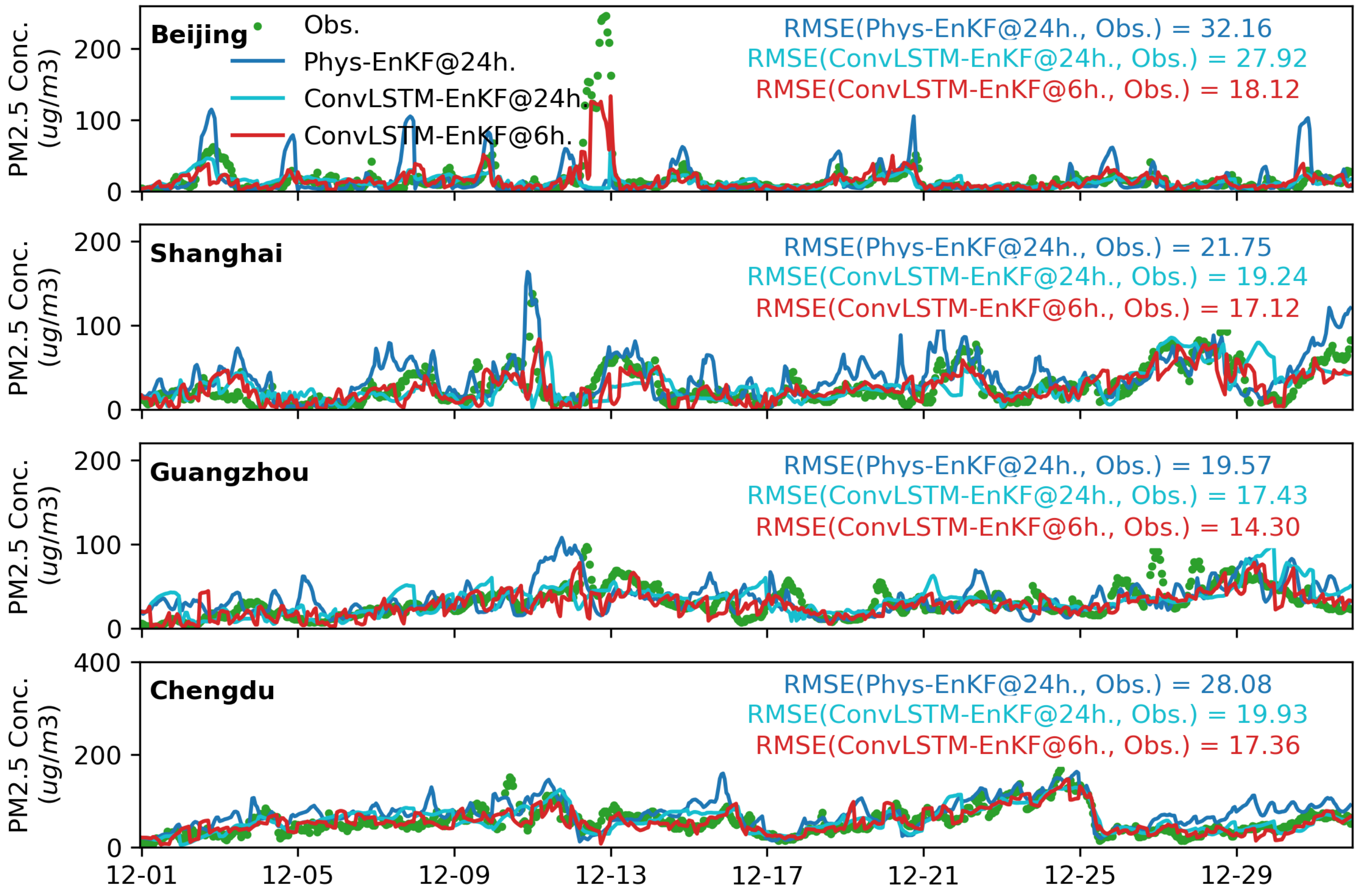


Ensemble size: 100

DA frequency: 6h

Long term forecasting – Case Study

Comparison between ML-DA and Physics-DA models



Long term forecasting – Case Study

ML & DA model – Efficiency

Methods	Ensemble size	CPU hours
NAQPMS-EnKF	50	166.67
ConvLSTM-EnKF	50	0.12

- CPU: Intel(R) Xeon(R) W-1290P CPU@3.70GHz GPU: NVIDIA RTX A4000
- **EnKF**: requires **large ensemble size** to represent the statistical distribution of the studied state variables (mean and variance)
- **Conventional DA system with physical models**:
  commonly use **~50** (computationally expensive, cannot afford large ensembles)
  mostly **offline** (monthly analysis data available)
- **ConvLSTM-EnKF**: enable large ensemble size, further improve DA accuracy

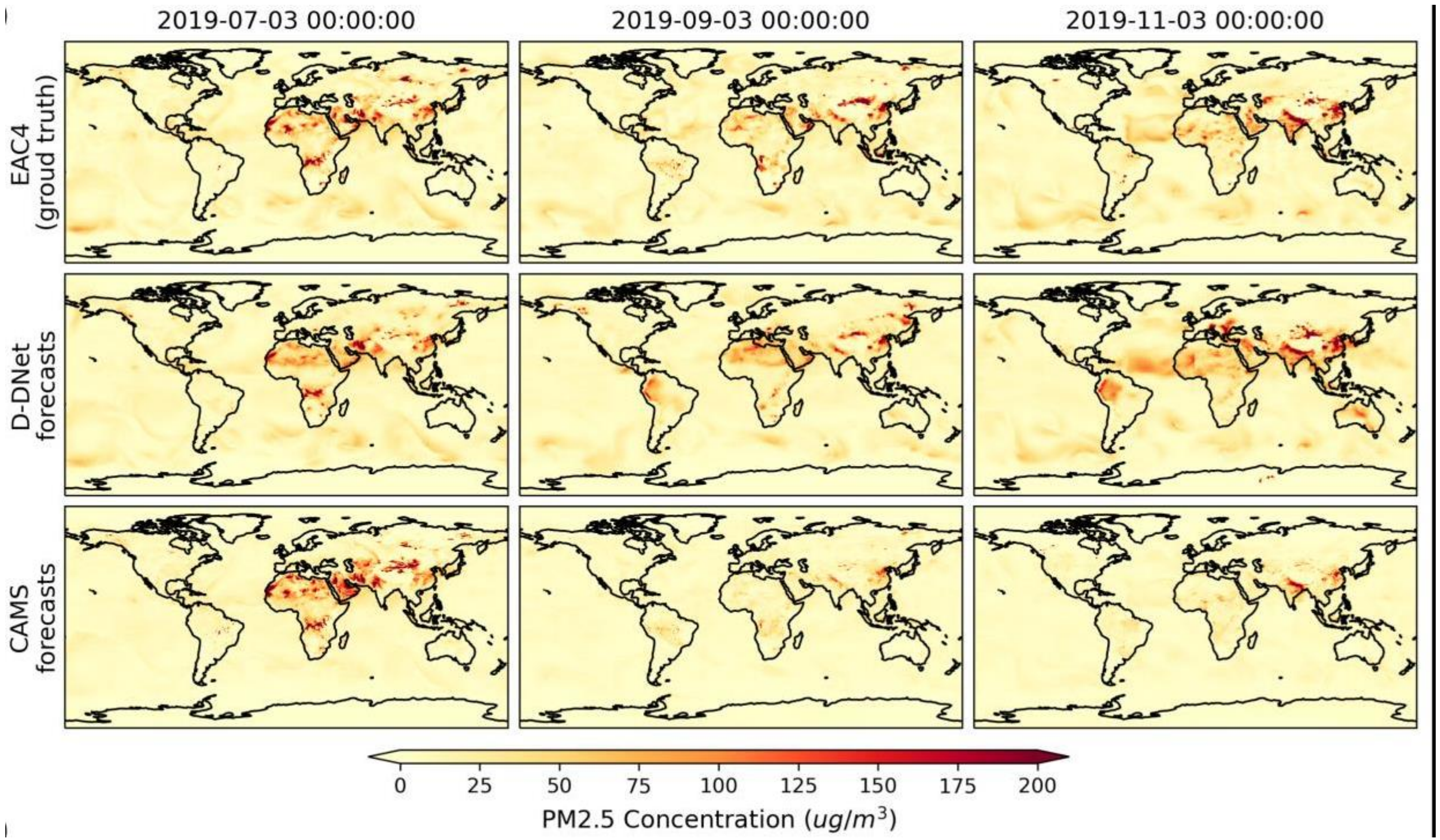
Long term forecasting – Case Study

Summary and impact of this work

- **Long-term forecasting**: Hourly spatiotemporal PM2.5 forecasting
  Existing modes: hourly forecasting for the whole China, up to 48 hours
  ML-DA: hourly forecasting up to one month plus.
- **Computational efficiency (CPU) online simulation (forecasting + DA)**:
  Physical modes: 166-hour CPU time for every hourly prediction
  ML-DA: 7 minutes for every hourly prediction
- **Impact:** pave the way for operational real-time prediction and management

Global PM2.5 forecasting–Case Study (Ms. Cai)

ML & DA model – sparse observations



ECMWF Atmospheric Composition Reanalysis

- Reanalysis data = Integrating surface observations and physical simulation - advanced Copernicus Atmosphere Monitoring Service (CAMS), operated by the European Centre for Medium-Range Weather Forecasts (ECMWF).
- High spatial resolution: 80km x 80km (60000 nodes)
- High temporal resolution: 3h
- Data assimilation frequency: 6h. Sparse data: 3258.

Training (90%) + validation (10%) : 2013-2018
Predicting: 2019

Digital tools for Urban Environment Management

EPSRC
Engineering and Physical Sciences Research Council

Aim to develop develop a hybrid AI-physics framework for optimal city design and management for decarbonisation

- Allow critical assessment of UK existing and drive new policy options on decarbonisation to achieve net zero by 2050
- Improve the existing regulations for decarbonization by providing valuable insights, optimising energy efficiency, and empowering decision-making processes with increased knowledge and awareness.

Existing Cities

London

Singapore

Ningbo

Future Smart Cities

Neom city (smart city) construction in Saudia Arabia – to demonstrate the AI approach

Government and Regulation

Department for Environment Food & Rural Affairs

Department of Health & Social Care

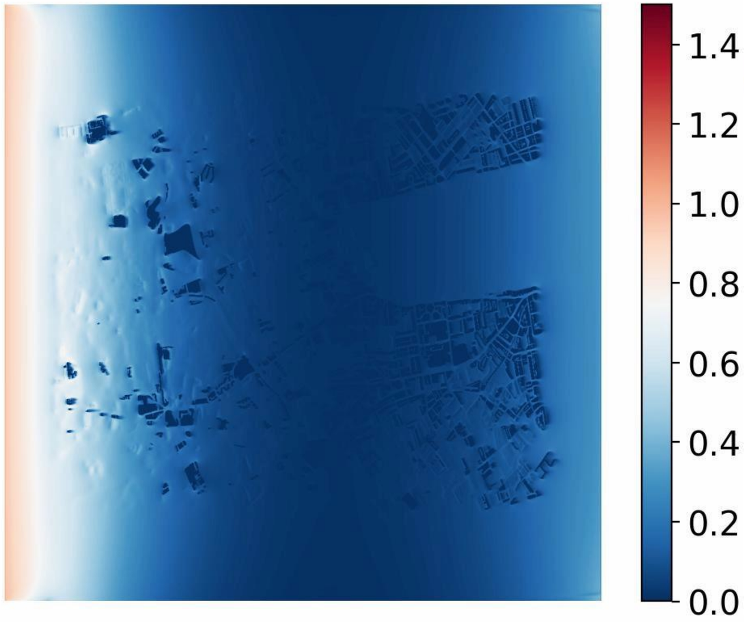
Department for Business, Energy & Industrial Strategy

Department for Transport

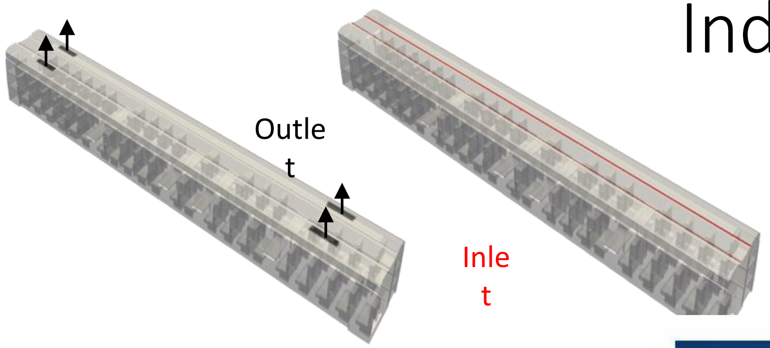
Environment Agency

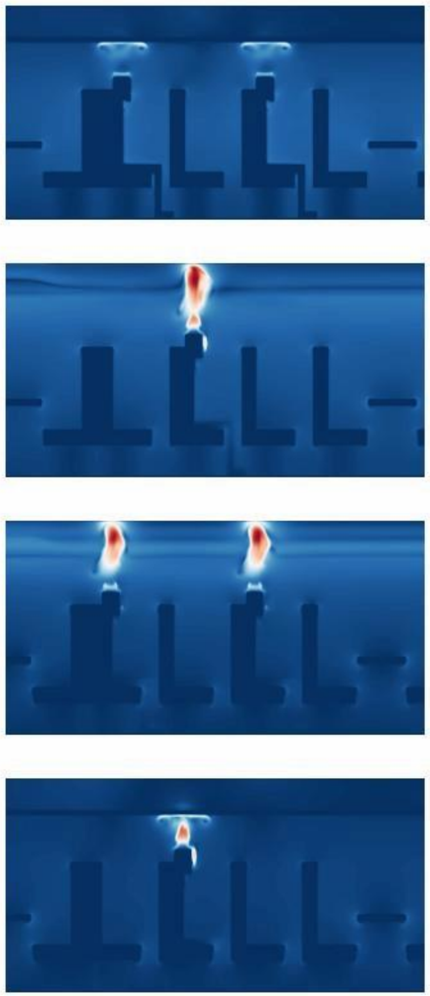
Rapid High-fidelity simulation of airflow in central London (Dr. B. Cheng)

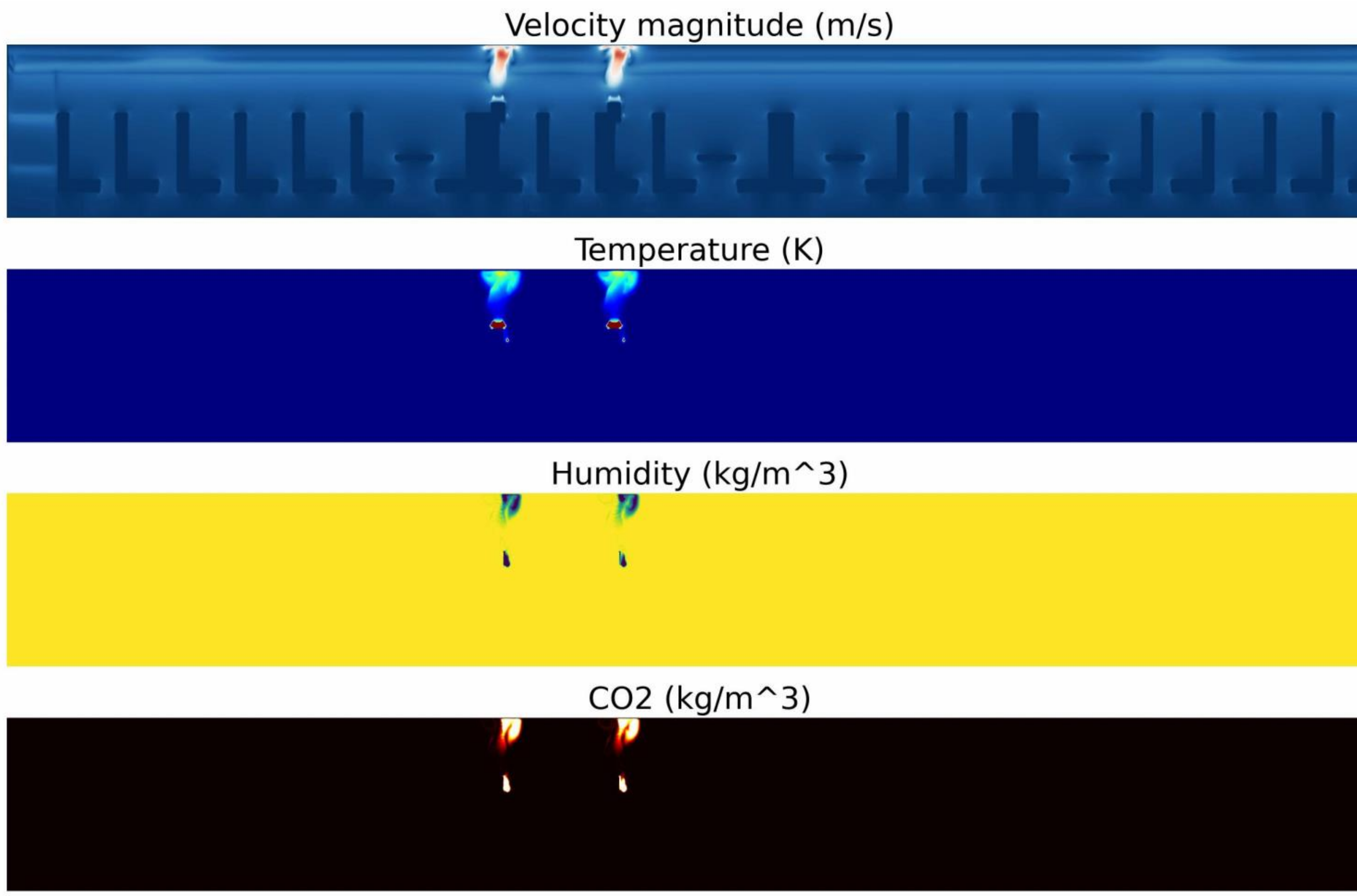
- 1024m x 1024m x 128m
- 134M structured element nodes (street level)
- One single GPU (NVIDIA RTX A5000)
- One-hour computational time → 5 hours



Indoor air quality modelling of train carriage (Dr. B. Cheng)



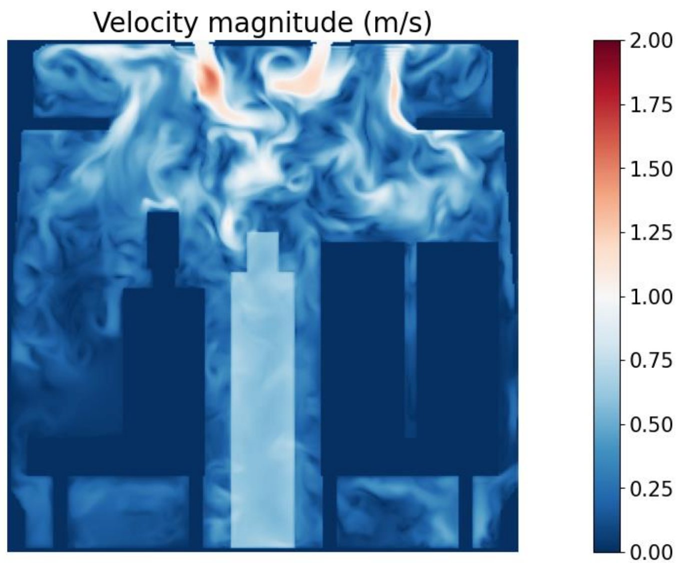


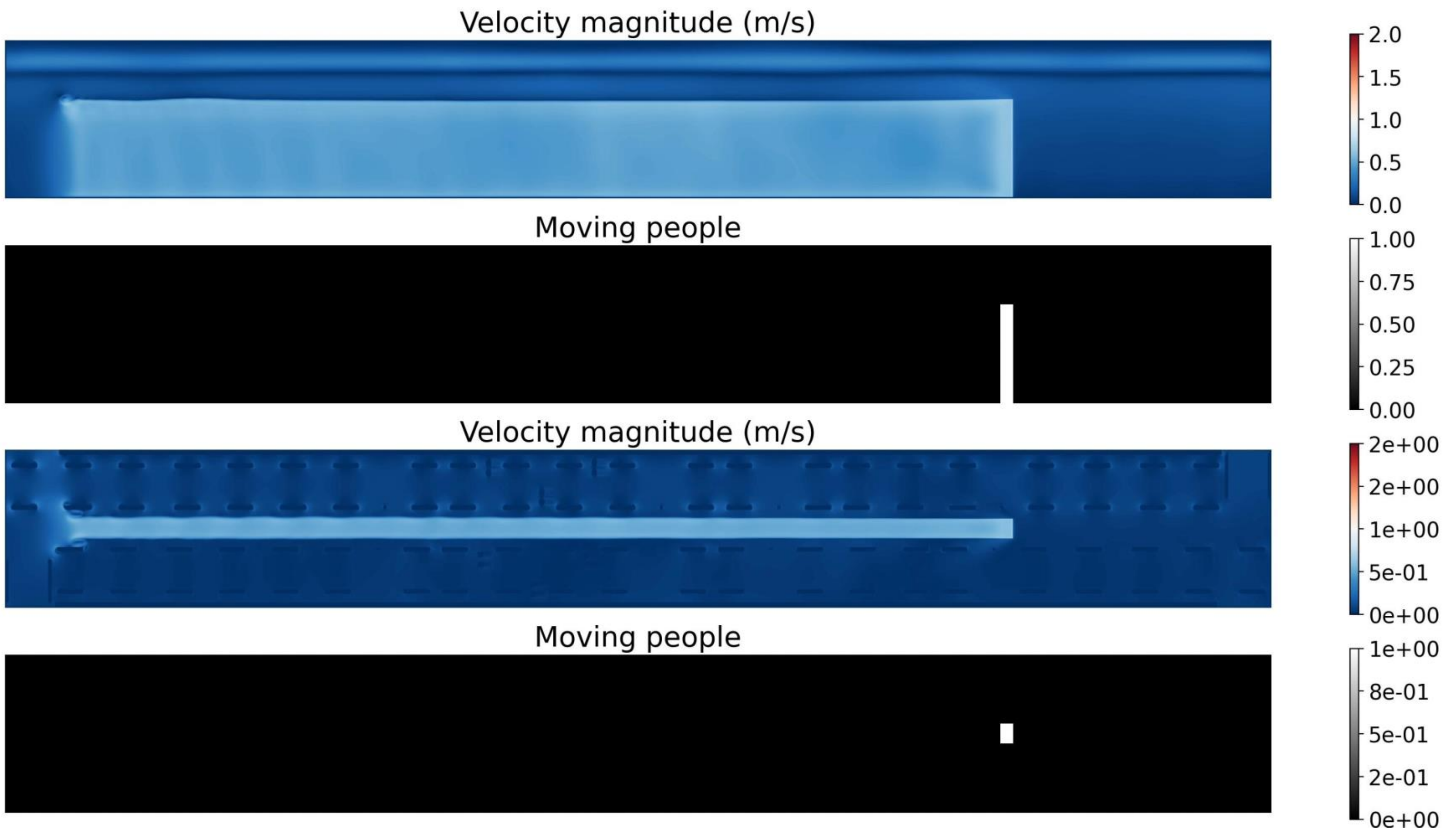


- 20.48m x 2.56m x 2.56m
- Physical modelling (thermal buoyancy, etc)
- People movement
- 134M structured element nodes
- One single GPU (NVIDIA RTX A5000)
- One-hour computational time → 2 days

People movement within the ventilated train carriage (Dr. B. Cheng)

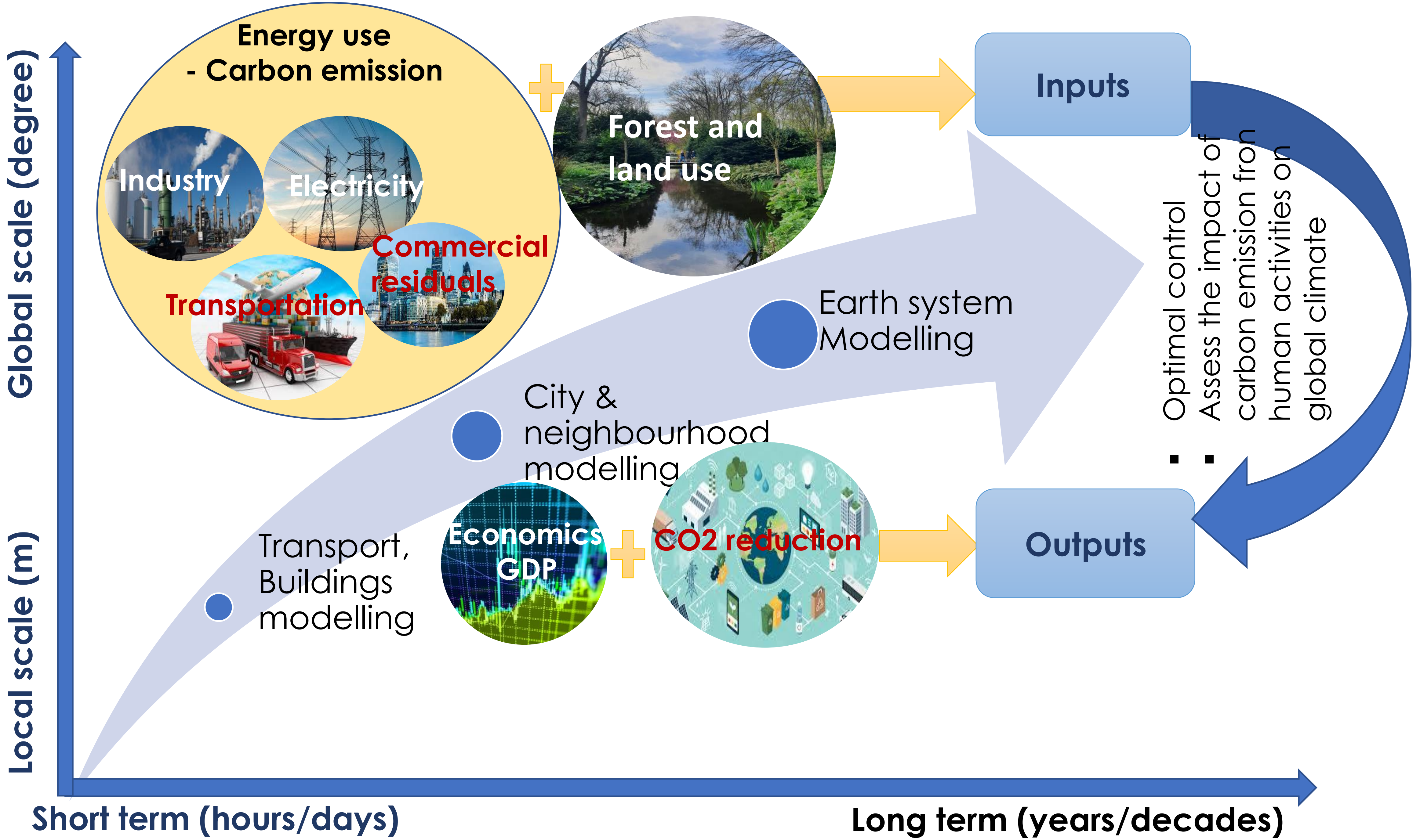
- Walking speed → 0.6 m/s
- Size → 1.6m (H) x 0.2m (L) x 0.3m (W)
- Moving pathway → backward and forward along the middle of the train carriage
- Breathing out air while walking





Digital tools for Urban Environment Management:
Questions to be addressed

- How do anthropogenic carbon emissions affect local urban and global climate change?
- Which optimal GI-BI, buildings, transportation, and sustainable city designs provide maximum mitigation of carbon emissions & climate change?
- What is the trade-off between carbon reduction, energy use and economics?
- How can detailed multi-scale models provide efficient and accurate prediction of carbon emissions and their impact on climate change?
- What are the feedbacks of the urban carbon contribution to global climate? (Assess tge improvement of global climate after carbon reduction via optimal management of infrastructures)



Energy use
- Carbon emission
Industry
Electricity
Commercial residuals
Transportation
Forest and land use
Inputs
Outputs
Earth system Modelling
City & neighbourhood modelling
Transport, Buildings modelling
Economics GDP
CO2 reduction
Optimal control
Assess the impact of carbon emission from human activities on global climate
Global scale (degree)
Local scale (m)
Short term (hours/days)
Long term (years/decades)

Physical image
“As Is”

Hybrid data generation approach

- Collecting data from sensors (e.g. drones, mobiles) and satellites;
- Physical modelling solutions

Hourly/daily physical nowcast/forecast

- Traffic emission spatial map
- Carbon/pollutant spatial map
- People map – linked to mobiles – people trace app
- Energy use/distribution map
- Extreme weather forecast (flooding, hurricane)

Digital Twin
(IoT)

Internet of things

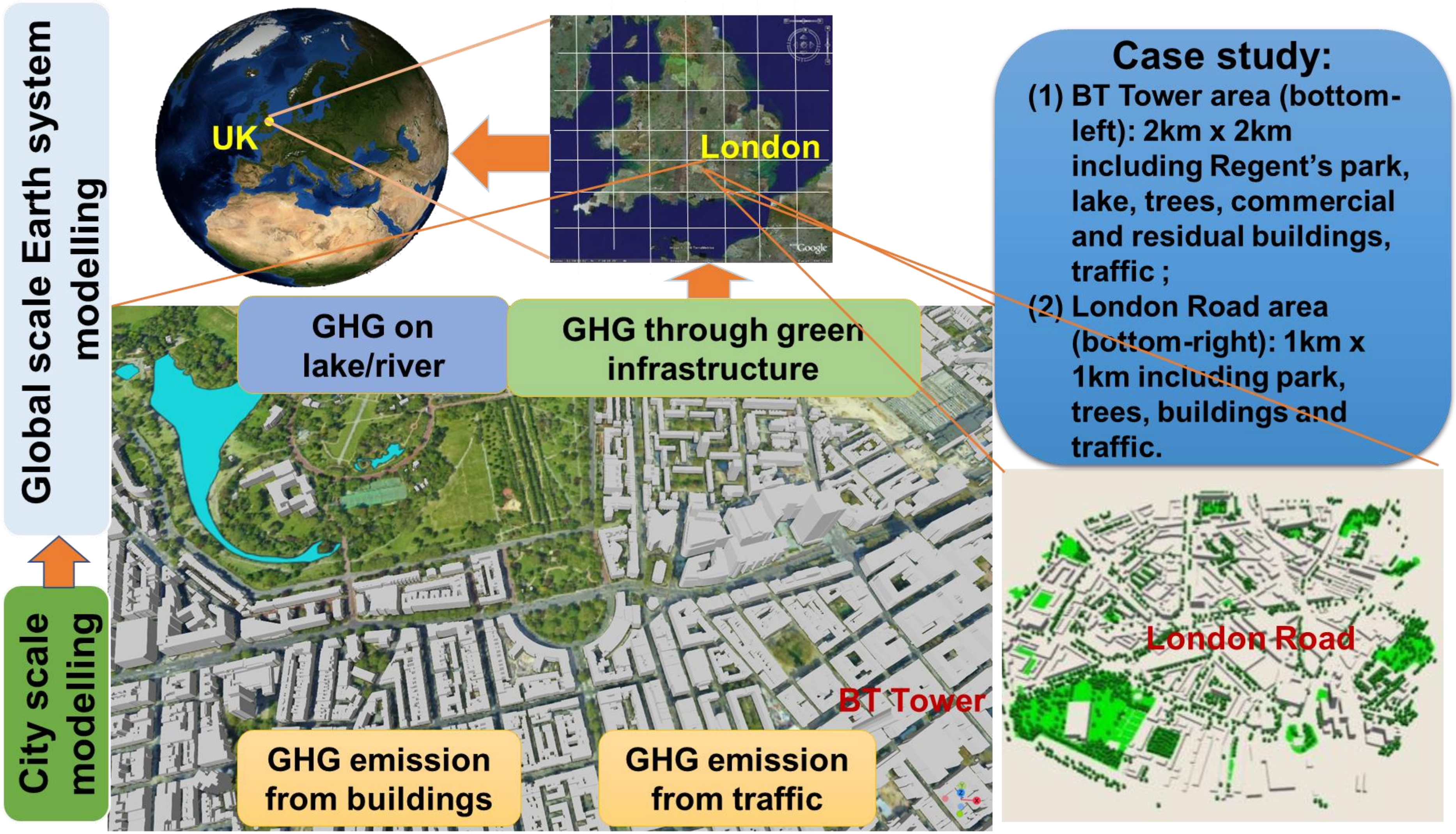
Virtual image
“To Be”

AI-enabling decision support system

- Autonomous carbon/pollutant monitoring and control
- Optimal traffic flow system
- Building environment control system (indoor and outdoor)
- Green and Blue infrastructures
- Efficient energy system
- Assessment of socio-economic & health impact

Digital tools for Urban Environment Management:

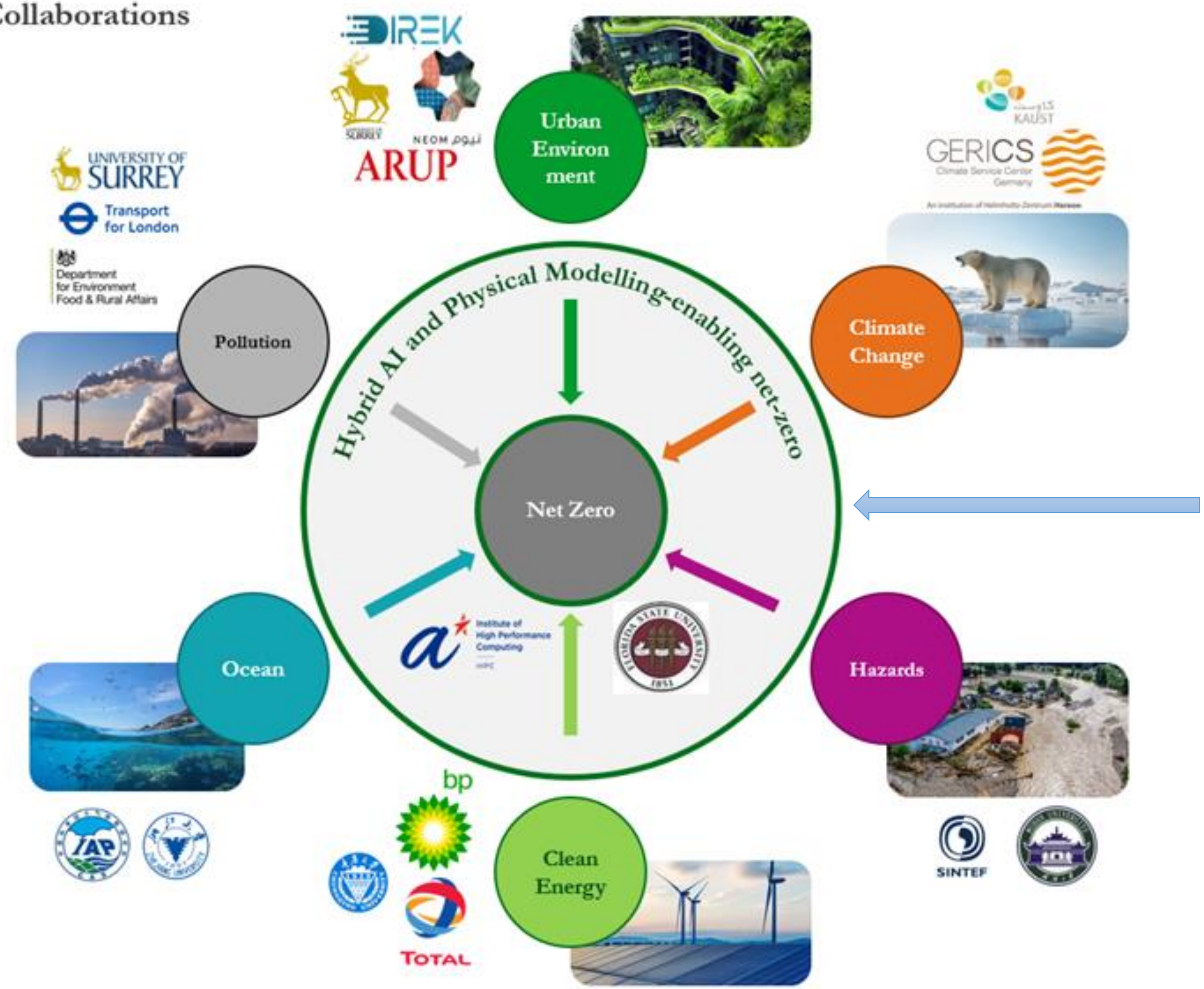
Integrated modelling from the neighbourhood, city to global scales showing the city GI-BI and human activities on local and global climate



Generally, the hybrid AI and multiscale physical modelling framework will facilitate a comprehensive evaluation of the existing conditions in the UK and enable the exploration of new policy options for decarbonization. These powerful tools possess the potential to significantly impact and improve current regulations related to decarbonization by providing valuable insights, optimizing energy efficiency, and empowering decision-making processes through increased knowledge and awareness.

- **Infrastructure Optimization**: Enable us to optimise existing infrastructures for energy efficiency and reduction of carbon emission and environment impact
- **Transportation planning**: Allow us to optimise the transportation routes and control the traffic flow, thus reduce the carbon/pollutant emission;
- **Urban planning and design**: Enable urban planners to visualize and plan for a sustainable city layout with green/blue spaces and efficient energy buidlings;
- **Efficient and resilience energy system**: Simulate energy consumption patterns in buildings and entire city systems. AI algorithms can then identify opportunities for energy efficiency and provide resilient energy plan in response to extreme climate;
- **Financial planning**: Provide cost-benefit analysis for different carbon reduction strategies, thus maximising its impact.
- **Real-time monitoring and data analysis**: Provide real-time data allowing us to monitor and measure the effectiveness of decarbonization efforts continuously;
- **Policy and regulation support**: Assess the impact of different policies and regulations on the city’s footprint, thus providing new regulations for decarbonisaton.

Collaborations



Internal collaboration

ESE;
Environmental research group;
Centre for environmental policy;
Civil and Environmental Engineering;
Physics Atmosphere;
Data Science Engineering;
I-X;
Grantham institute;
ICT

Thanks!

SCAN ME

Dr Fangxin Fang
Senior Research Fellow