上海理工大學
UNIVERSITY OF SHANGHAI FOR SCIENCE AND TECHNOLOGY

# 机器学习辅助大气污染及暴露估算

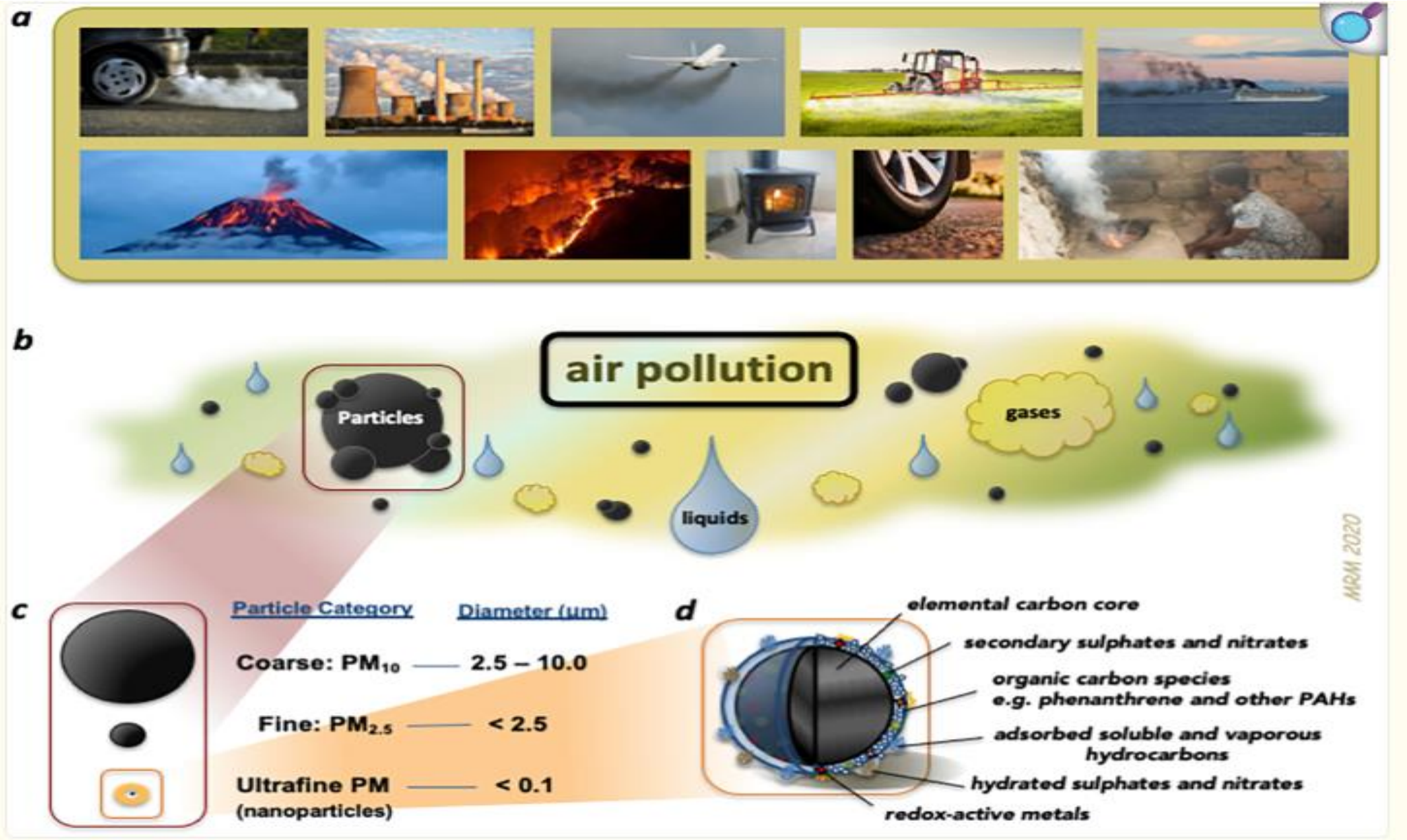





张宏亮

上海理工大学 环境与建筑学院

# 大气污染

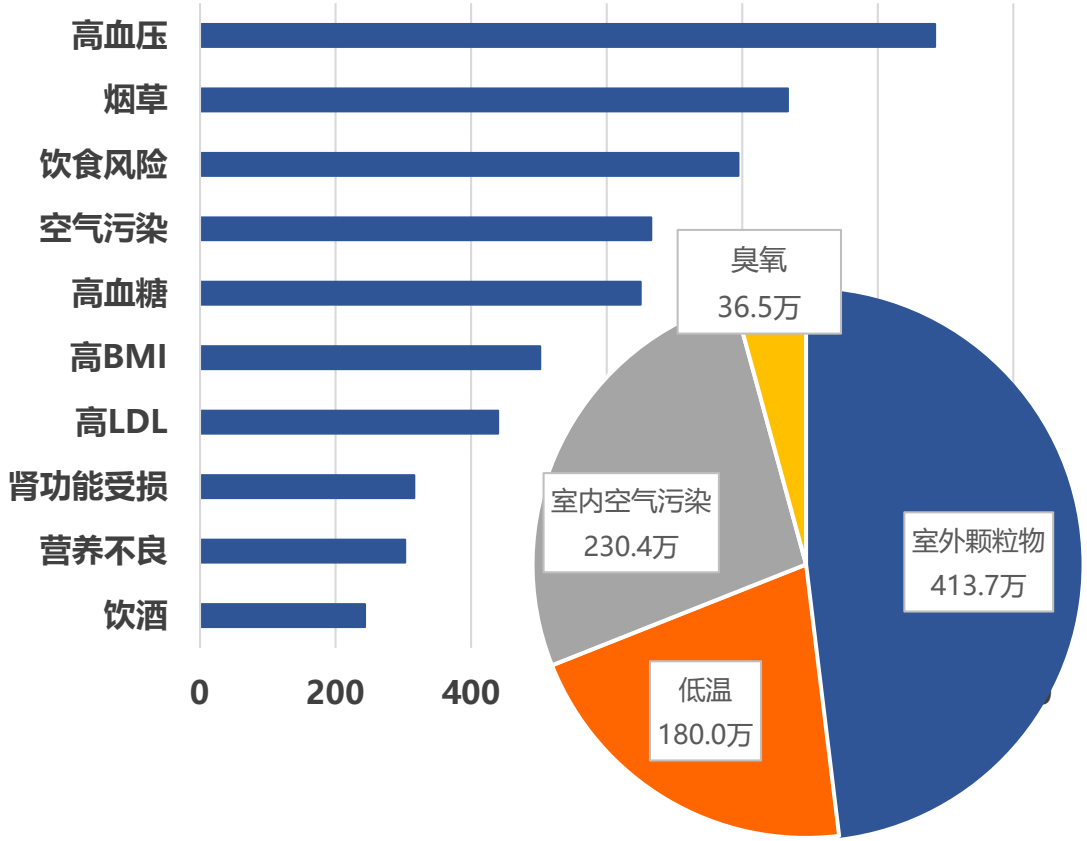



Miller， 2022

# 大气污染健康影响

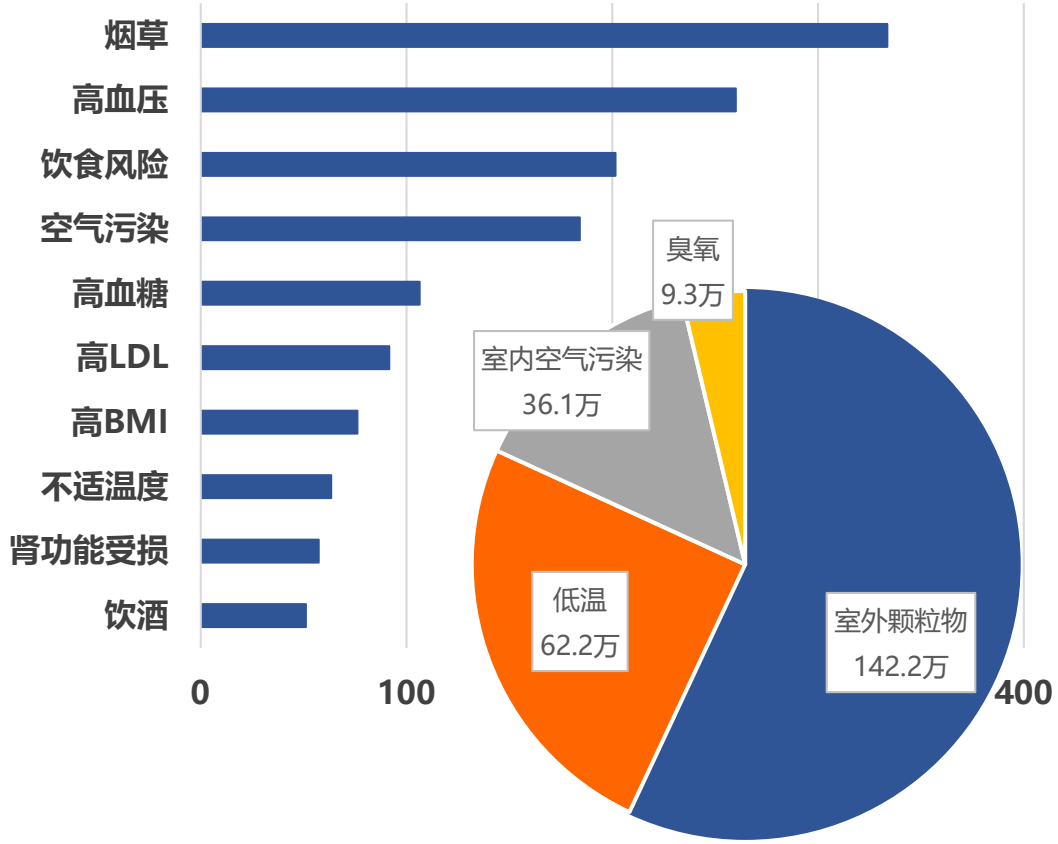
## 2019 全球前十致死危险因素

归因死亡人数（万）




## 2019 我国前十致死危险因素

归因死亡人数（万）




GBD, Lancet, 2020

# 暴露评估是研究关键

## 研究路径

- 空气污染监测与建模
  - （监测站 / 遥感 / 数值模式 / AI）
- 污染物暴露评估
  - （核心环节）
- 人群健康数据整合
  - （住院 / 队列 / 临床检测）
- 统计与因果分析
  - （时间序列 / C-R / 贝叶斯）
- 结果解读与政策建议
- （高风险人群 / 干预措施）

## 暴露评估

- 高精度：
  - 空间 <1 km，时间 ⩽1 小时
- 多污染物联合：
  - $PM_{2.5}$、$O_3$、$NO_2$、$SO_2$ 等
- 个体化暴露：
  - 结合活动轨迹与可穿戴设备
- 室内外整合：
  - 考虑生活/工作环境差异
- 减少误差：
  - 提升健康风险估计可靠性
- 人工智能 + 遥感 + 传感器革新暴露评估方法

# 暴露估算方法

| 方法 | 优点 | 局限性 |
| --- | --- | --- |
| 固定监测站 | 数据连续、长期记录，易获取 | 空间覆盖有限，不能反映个体差异 |
| 个体可穿戴监测 | 精细到个人，反映真实暴露 | 样本量小，成本高，设备依从性问题 |
| 卫星遥感反演 | 覆盖范围广，可跨区域比较 | 垂直分辨率不足，受云和气溶胶干扰 |
| 化学传输模型 | 多污染物同时模拟，可探究来源 | 依赖排放清单，计算量大，模型不确定性 |
| 人工智能 | 融合多源数据，精度可高，适应性强 | 需高质量训练数据，黑箱问题，外推性有限 |

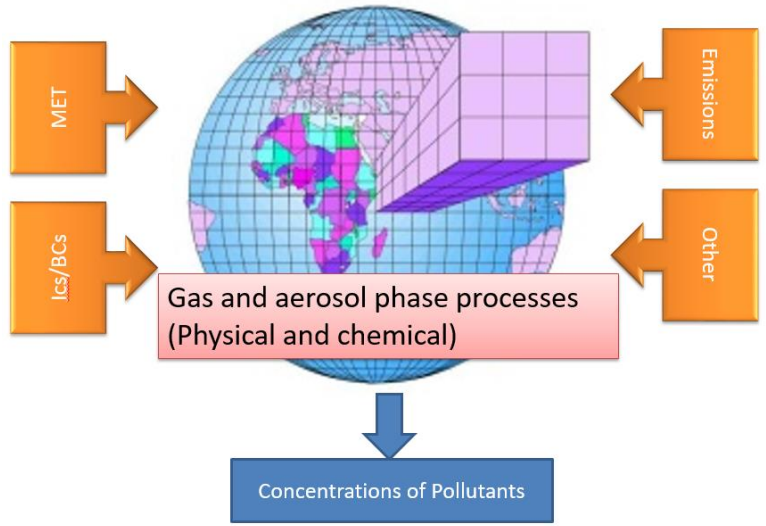
# 化学传输模型

$$\frac{\partial C}{\partial t} = -\frac{\partial (u_i C)}{\partial x_i} + \frac{\partial}{\partial x_i}\left(K_i \frac{\partial C}{\partial x_i}\right) + R + S - L \qquad i=1,2,3$$

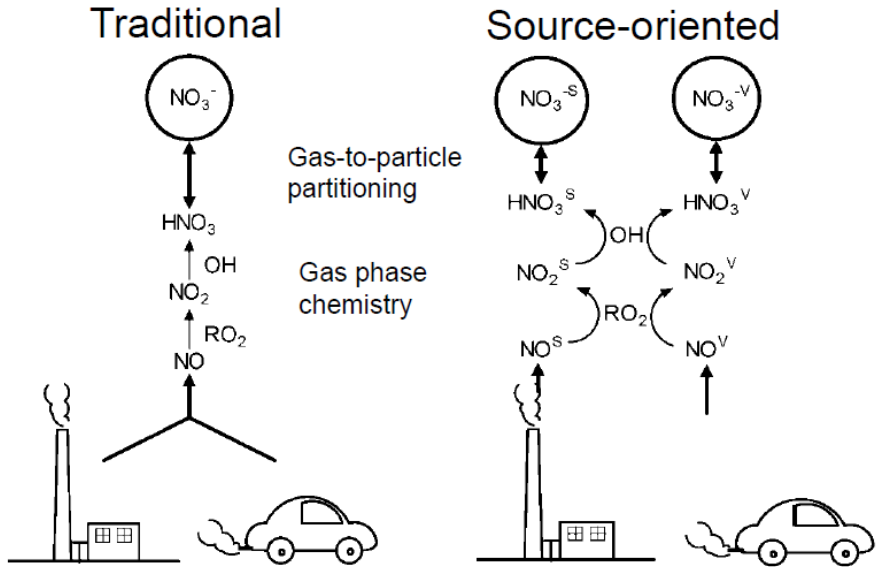
Emission

Advection　　Turbulent Diffusion　　Reaction　　Removal

## 传统空气质量模型



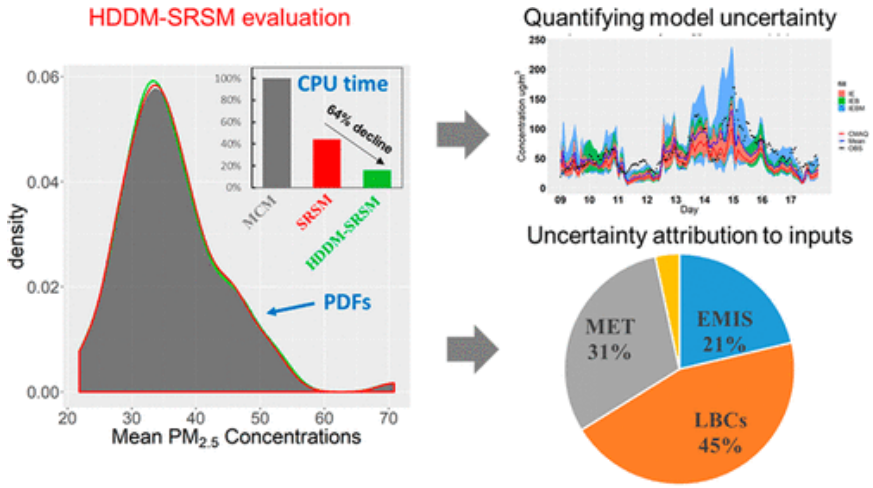

## 反应示踪物技术




不同来源对污染物的贡献：时空分布，一次污染物，二次污染物、区域贡献、行业贡献

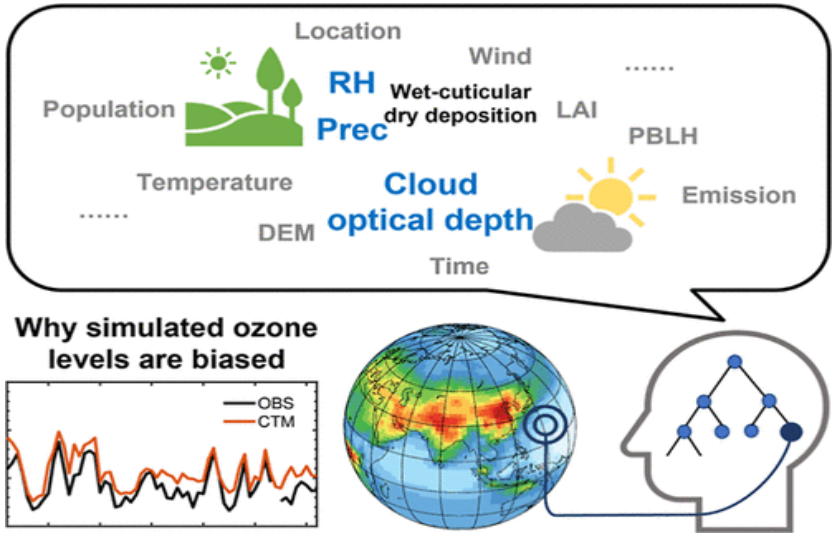
# 化学传输模型

- 模拟过程：不清楚的化学过程、简化的过程
- 模拟输入：气象场、排放清单、初始和边界条件
- 模拟结果：观测对比
- 模拟精度：全球-区域-城市
- 模拟需求：计算资源需求大



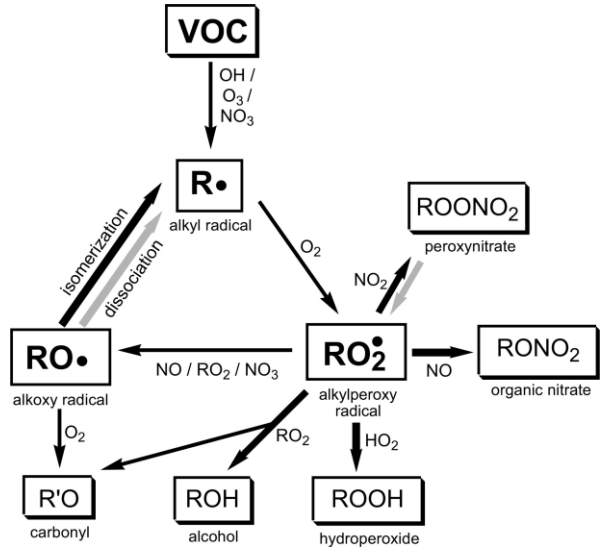

**气象，排放和横向边界条件对$PM_{2.5}$模拟偏差的贡献**

*Environ. Sci. Technol. 2019, 53, 6, 3110–3118*




**湿度，云光学厚度对$O_3$模拟偏差大**

*Environ. Sci. Technol. 2021, 55, 8, 4483–4493*




一般挥发性有机化合物大气氧化的简化机制

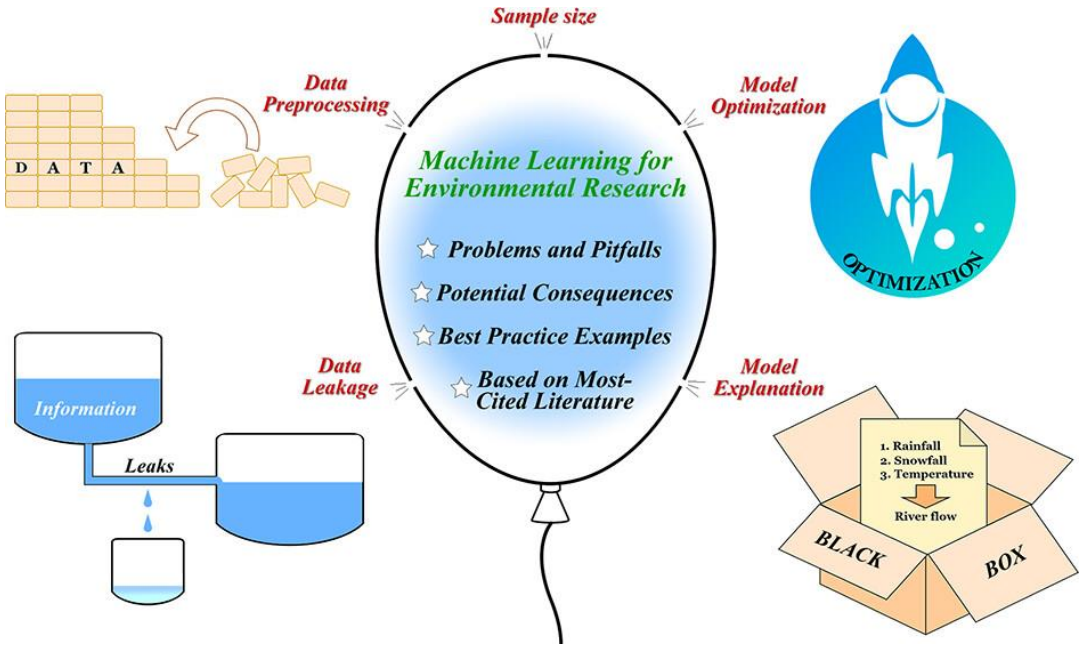
doi:10.1016/j.atmosenv.2008.01.003

所有模型都是不准确的，一些模型是有用的

# 人工智能、大数据、机器学习？

**AI4Science**

- 人工智能并不新鲜
- 大数据+人工智能时代
  - 模拟，卫星，地表观测，实验室测量数据大量增加（气象再分析数据量级达PB级以上）
  - 机器学习和深度学习算法快速发展
  - GPU等算力的大幅提升



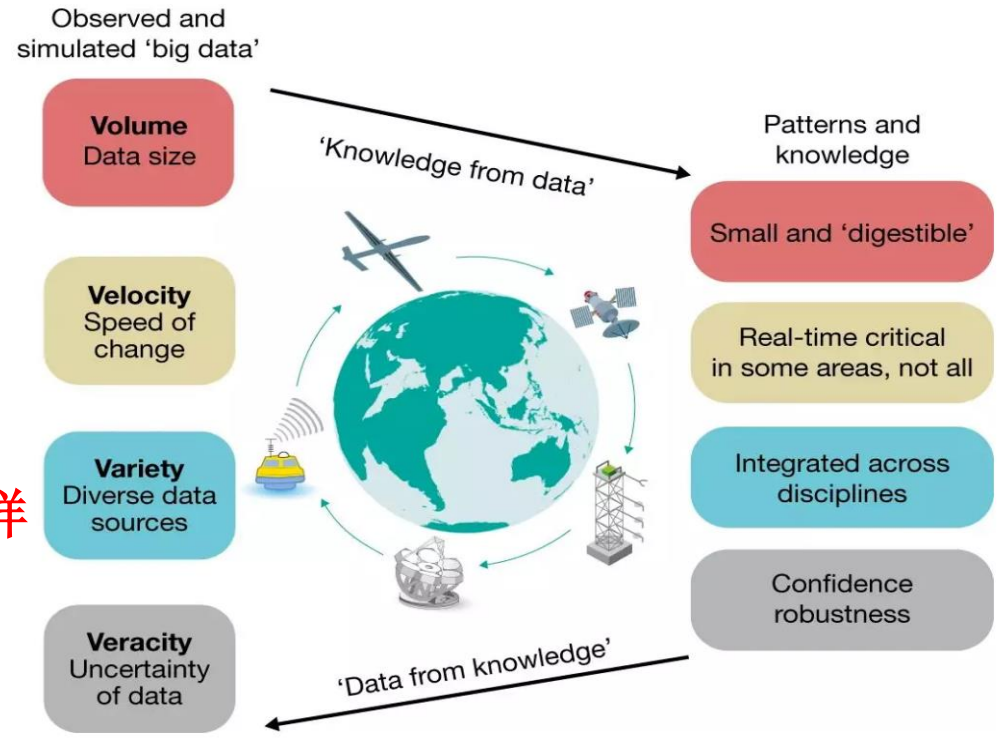

https://pubs.acs.org/doi/10.1021/acs.est.3c00026

体积大

更新快

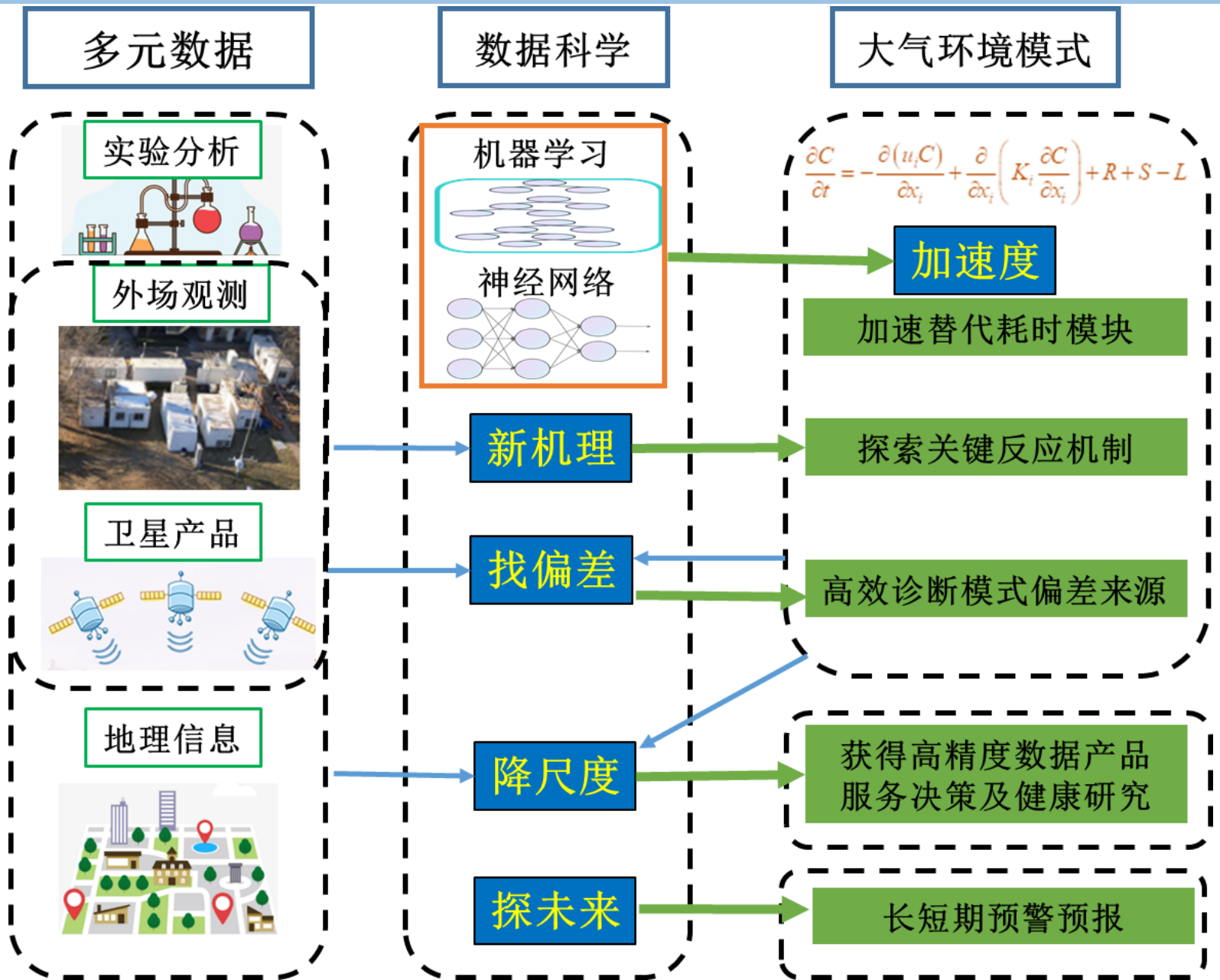
数据源多样

不确定性




Markus R, **Nature** 2019

# 人工智能辅助数值模拟


多元数据
数据科学
大气环境模式
实验分析
外场观测
卫星产品
地理信息
机器学习
神经网络
$\frac{\partial C}{\partial t}=-\frac{\partial(u_i C)}{\partial x_i}+\frac{\partial}{\partial x_i}\left(K_i\frac{\partial C}{\partial x_i}\right)+R+S-L$
加速度
加速替代耗时模块
新机理
探索关键反应机制
找偏差
高效诊断模式偏差来源
降尺度
获得高精度数据产品
服务决策及健康研究
探未来
长短期预警预报

# 系列尝试（王帅）

## 1. 多源数据集

- 地面观测数据（$PM_{2.5}$、$PM_{10}$、$PM_1$）
- 气象再分析数据（ERA5，ERA5-Land，NCEP-FNL）
- 卫星遥感数据（MAIAC AOD、NDVI、 Land use、DEM ）
- 大气再分析数据（MERRA2、LGHAP、TAP、CHAP）
- 污染物排放数据（ EDGAR / MEIC、 MEGAN、 FINN ）

## 2. 数值模式模拟

WRF → CMAQ

光化学机制: SAPRC-11
气溶胶机制: AERO6

初始条件、边界条件

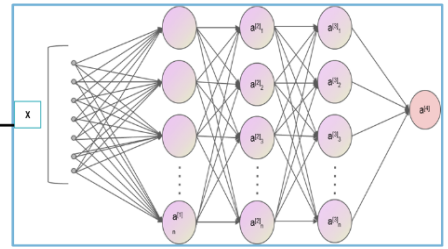
基础模拟 / 溯源模拟

数据预处理

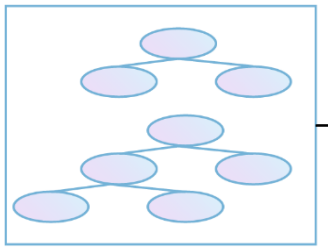
## 3. AI算法改进

- 时序特征提取：STID, Dlinear, Informer
- 图像特征提取：VGG, ResNet, Unet
- 特征融合网络：MLP, KAN, ResMLP

TRAIN | TEST



神经网络



树模型

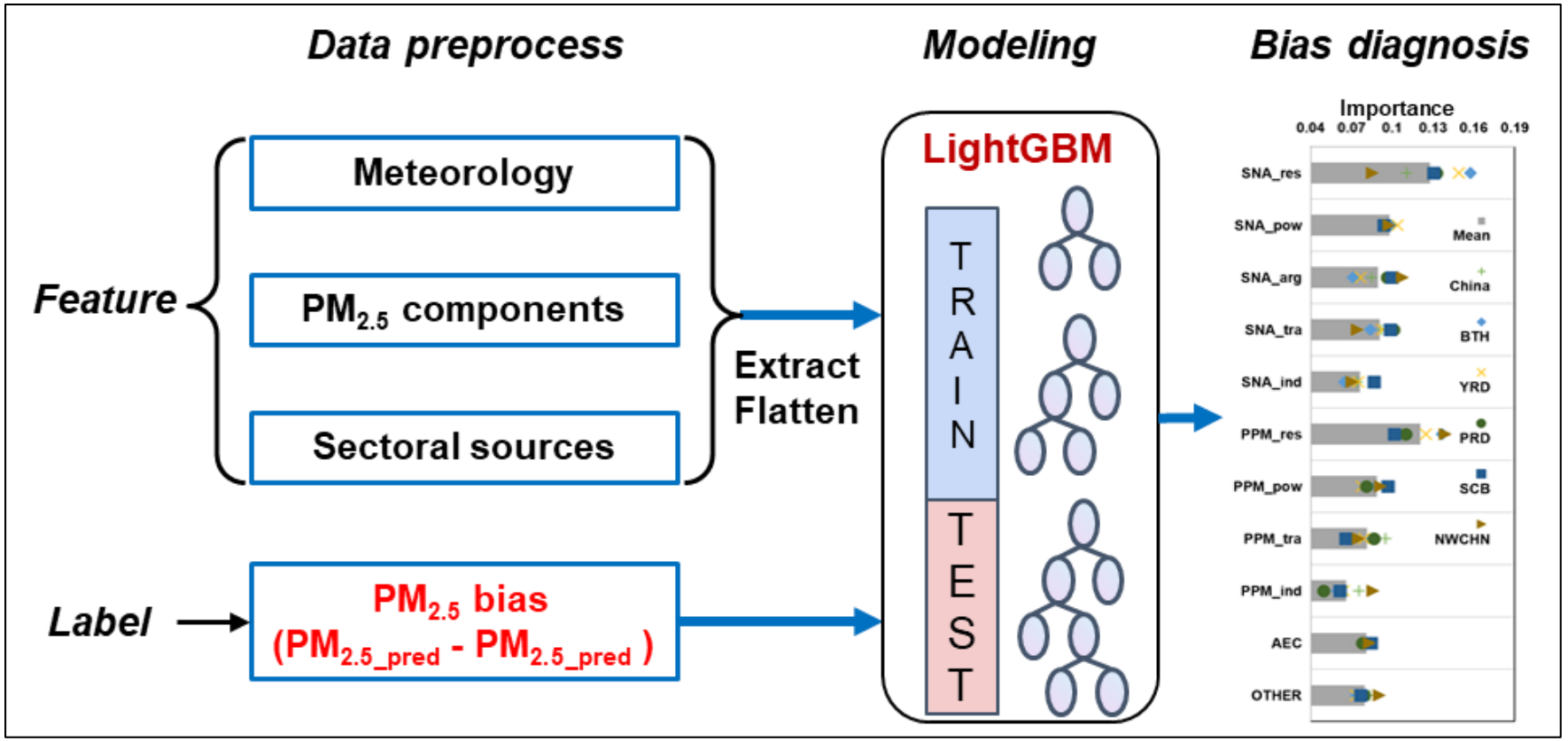
- 随机森林
- XGBoost
- LightGBM

AI模型推理

## 4. AI推理和应用

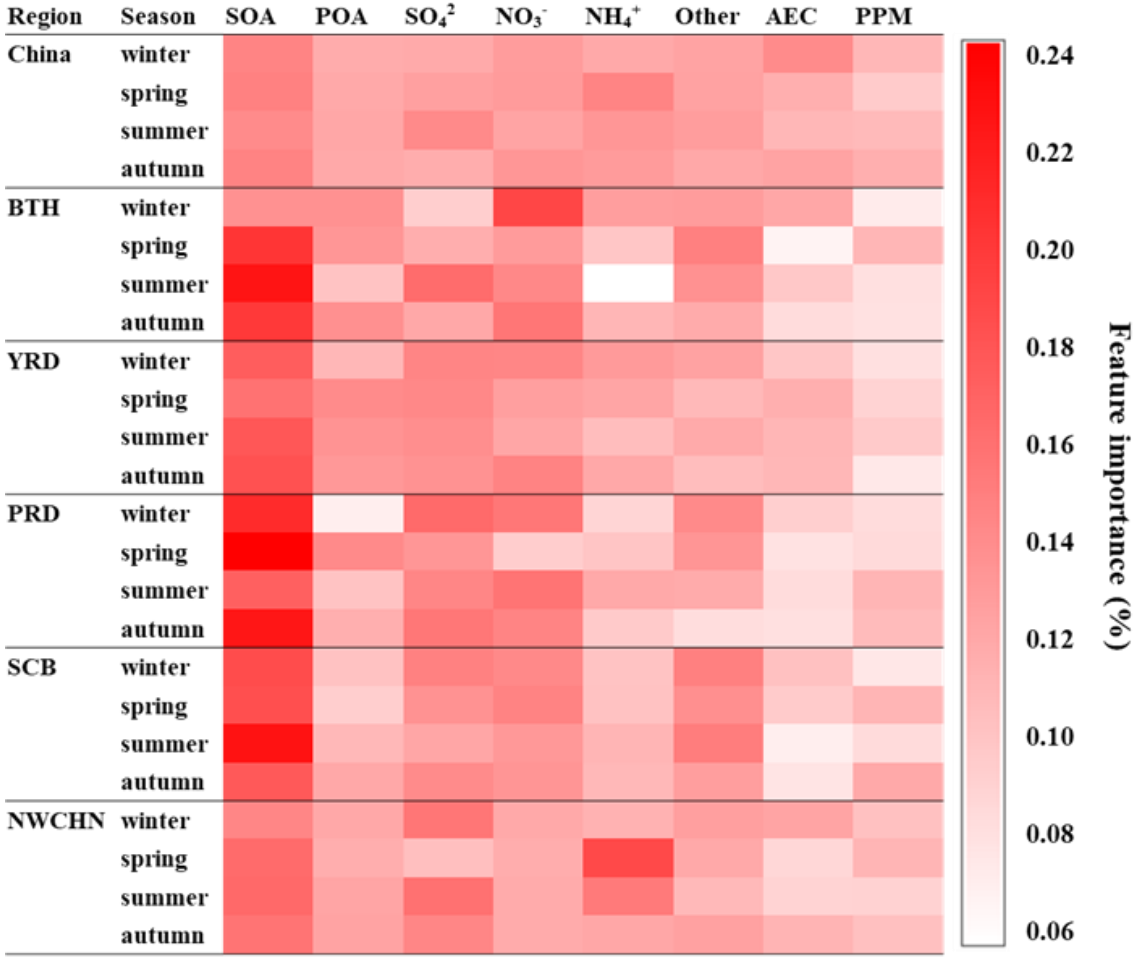
识别模拟偏差 → 建立估算模型 → 建立颗粒物数据集 → 亚洲$PM_{2.5}$健康影响 → 全球$PM_1$暴露不公平性

# 模拟偏差识别



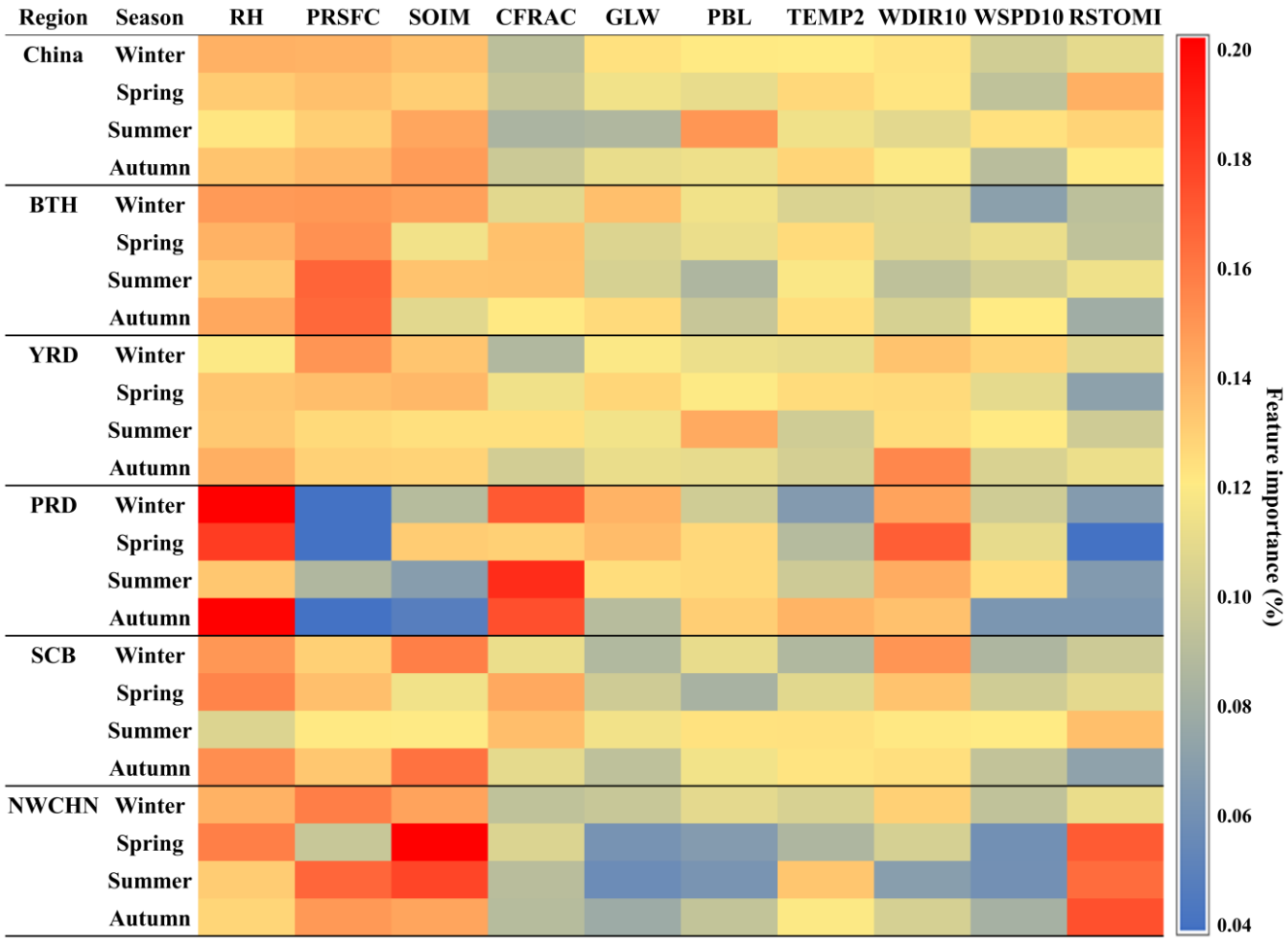

- 使用**lightGBM**模型（一种基于树的ML方法）来诊断CMAQ模拟地表$PM_{2.5}$浓度偏差
- **溯源式CMAQ**模型用来追踪$PM_{2.5}$的行业排放源
- 研究区域：中国；时间范围：2019年全年
- **排放清单：MEIC2019; 气象：WRF v4.1 (FNL)**

Wang et al., GMD 2023

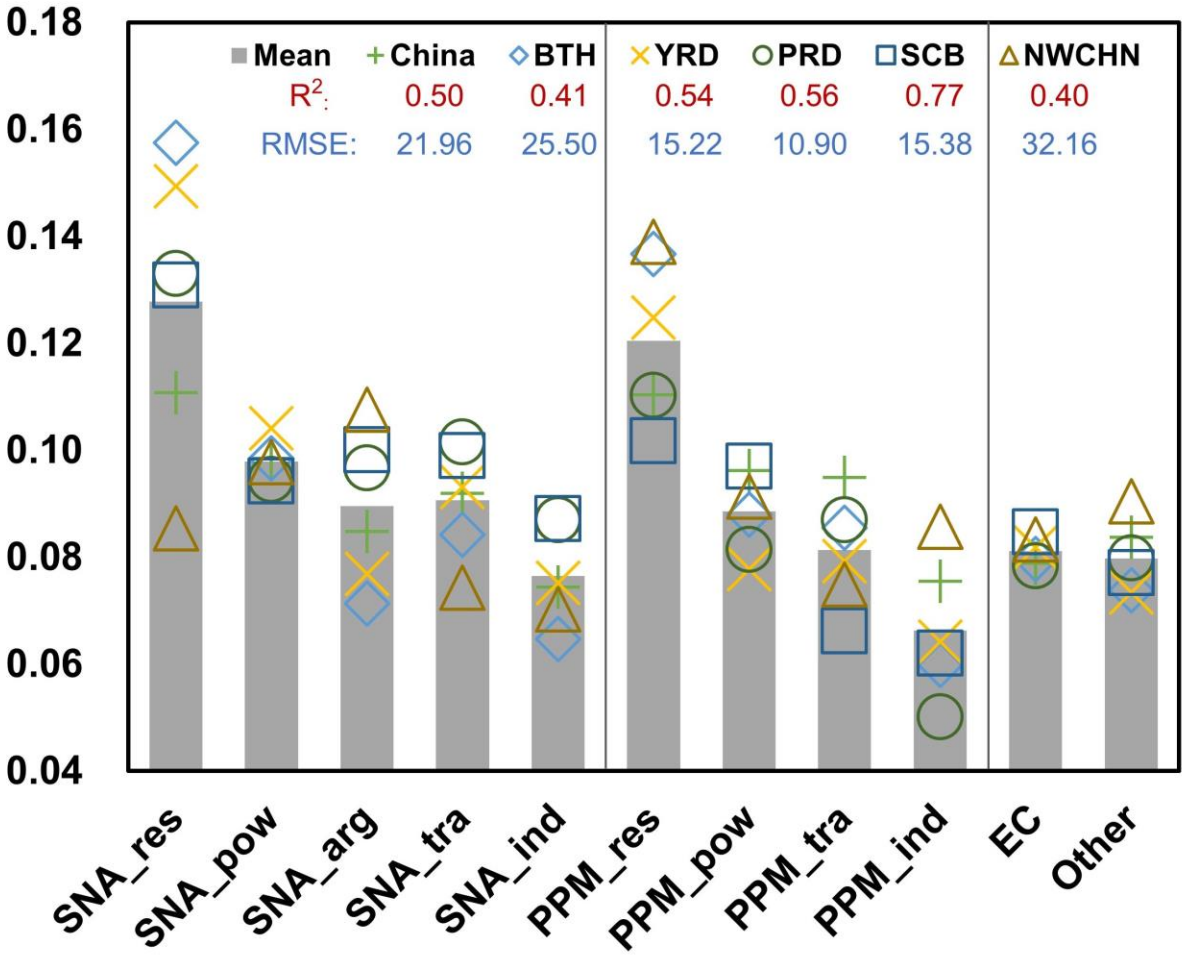



CMAQ模拟偏差的贡献（相对重要性%）

- 在$PM_{2.5}$成分中，**SOA对不同地区和季节的$PM_{2.5}$模拟偏差的贡献最大**




- 在BTH地区，表面压力和**相对湿度**对模拟偏差的贡献最大-------- **WRF模拟的不确定性（湿度 RMSE为20.38% ）**
- 在珠三角地区：相对湿度、**云量**、风速--------**气溶胶对气象的反馈作用**
- 在西北地区，**土壤表面湿度**和气孔导度--------**沙尘气溶胶**
- 干燥天气下，模拟低估更严重----------**干沉降方案不足**




各地区和季节的源排放对CMAQ模拟偏差的贡献（相对重要性%）

| Region | $PM_{2.5\_res}$ | $PM_{2.5\_pow}$ | $PM_{2.5\_tra}$ | $PM_{2.5\_arg}$ | $PM_{2.5\_ind}$ | EC | Other |
|---|---|---|---|---|---|---|---|
| BTH | **0.20** | 0.16 | 0.14 | 0.13 | 0.10 | 0.14 | 0.12 |
| China | **0.16** | 0.16 | 0.14 | 0.14 | 0.13 | 0.13 | 0.14 |
| FWP | **0.18** | 0.18 | 0.13 | 0.12 | 0.12 | 0.14 | 0.14 |
| NWCHN | **0.18** | 0.15 | 0.12 | 0.15 | 0.13 | 0.13 | 0.14 |
| PRD | **0.17** | 0.14 | 0.14 | 0.13 | 0.13 | 0.14 | 0.15 |
| SCB | **0.16** | 0.15 | 0.16 | 0.14 | 0.12 | 0.13 | 0.13 |
| YRD | **0.18** | 0.16 | 0.15 | 0.13 | 0.12 | 0.14 | 0.12 |

行业来源对CMAQ模拟偏差的贡献（%）

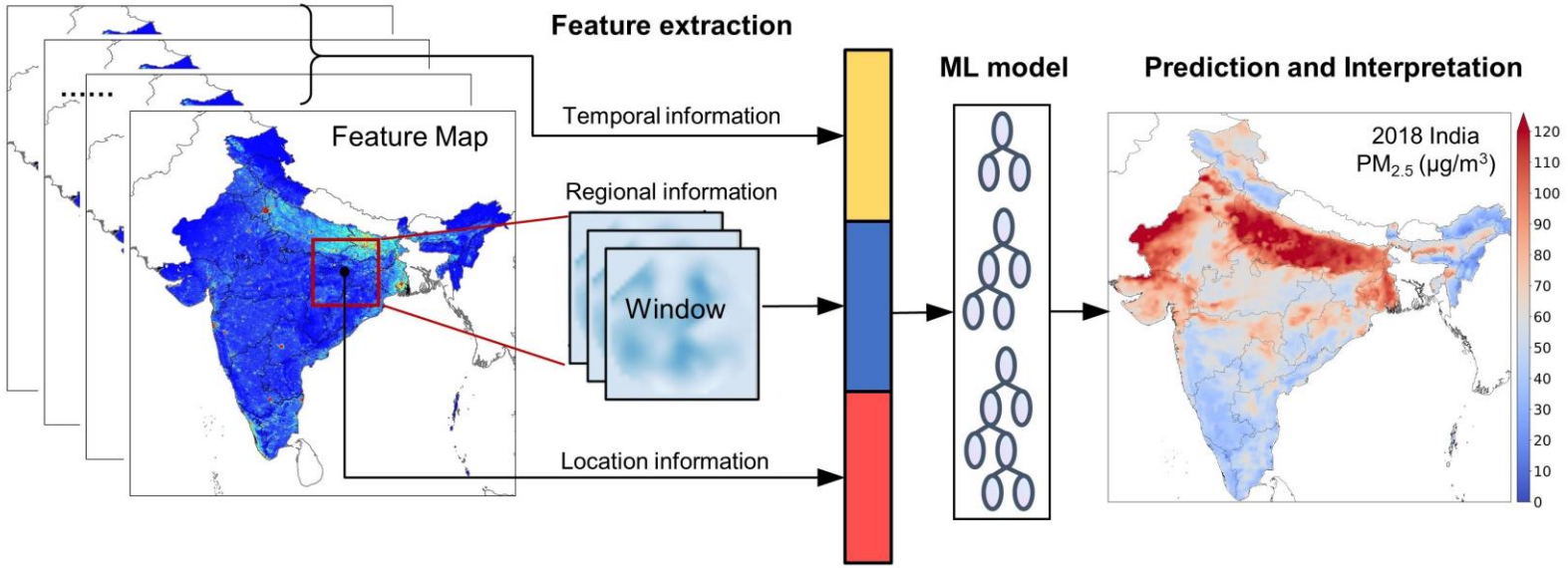
- 来自**居民部门**的PPM和SNA对$PM_{2.5}$模拟偏差的贡献最大。用不同行业来源的$PM_{2.5}$总浓度建立模型时，也得到了同样的结论。

Wang et al., GMD 2023

# 多数据源建模+改进

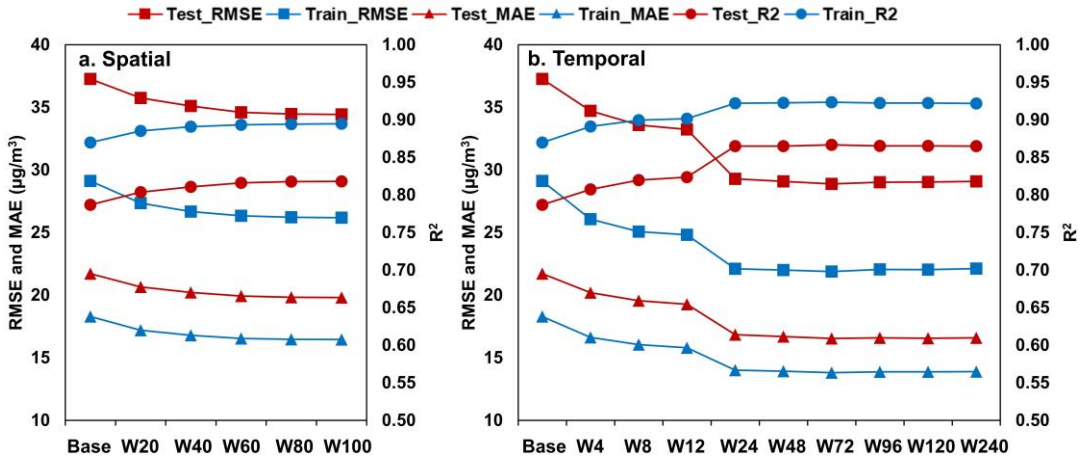
提取区域和时间特征，改进机器学习对印度小时颗粒物浓度建模

**关注高污染地区：印度**

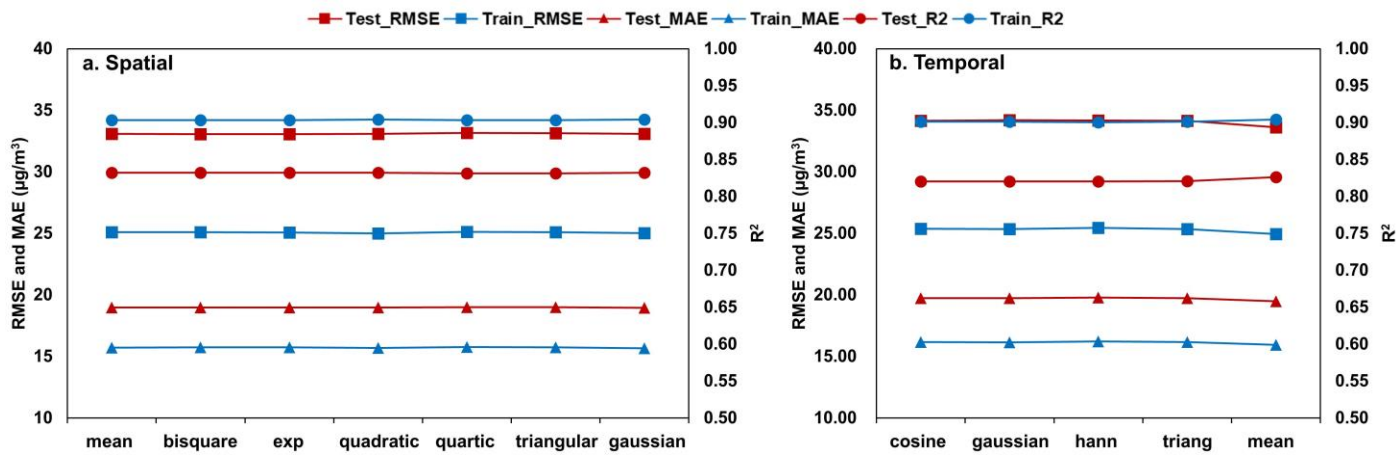



**空间特征提取：地理加权法**—距离越近权重越大

**时间特征提取：滚动加权法**—过去一段时间平均




时空加权窗口大小敏感性测试




时空加权核函数敏感性测试

In submission

# 建模改进——时空特征提取

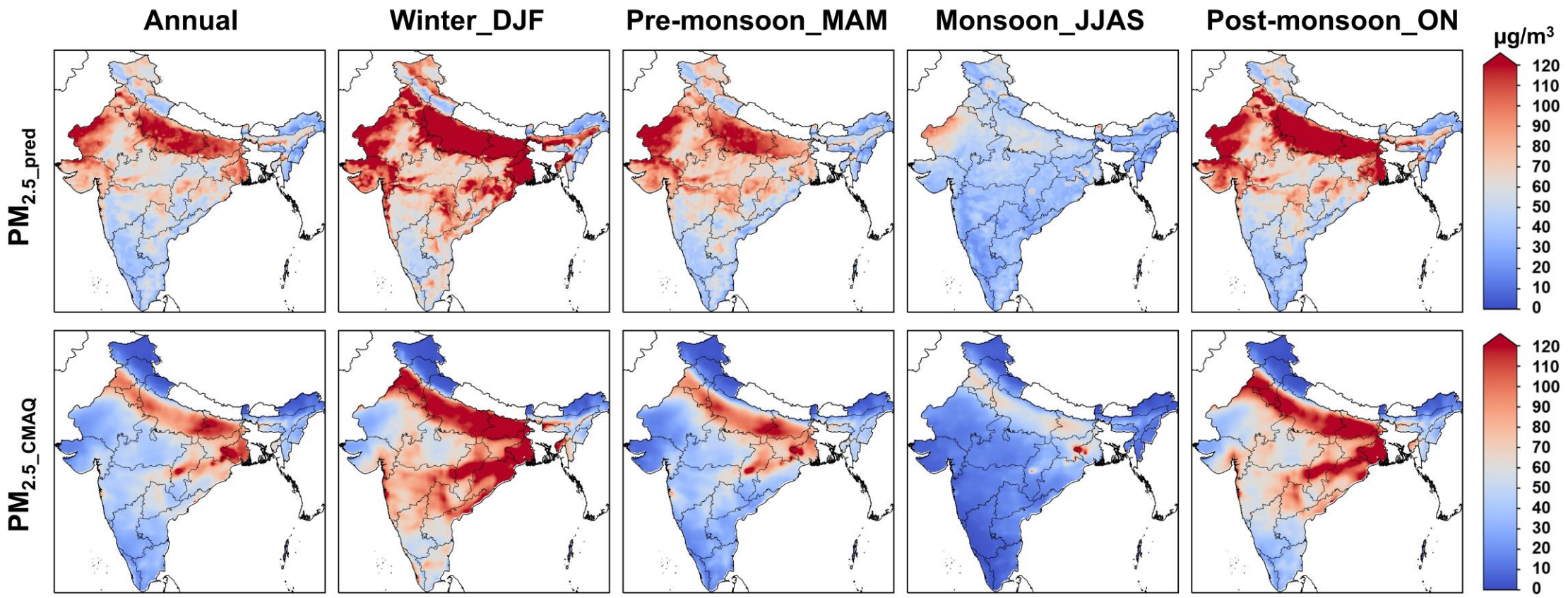
| Spec | Type | $R^2$ | | RMSE (μg/m$^3$) | | MAE (μg/m$^3$) | |
|---|---|---|---|---|---|---|---|
| | | Test | Train | Test | Train | Test | Train |
| $PM_{2.5}$ | Original | 0.79 | 0.87 | 37.27 | 29.13 | 21.72 | 18.31 |
| | Spatial | 0.80 | 0.89 | 35.73 | 27.35 | 20.65 | 17.20 |
| | Temporal | 0.82 | 0.90 | 34.09 | 25.75 | 19.85 | 16.40 |
| | Spatio-temporal | **0.83** | **0.90** | **33.53** | **24.88** | **19.45** | **15.92** |
| $PM_{10}$ | Original | 0.81 | 0.89 | 63.27 | 48.87 | 40.69 | 33.41 |
| | Spatial | 0.82 | 0.90 | 61.62 | 46.79 | 39.41 | 31.94 |
| | Temporal | 0.83 | 0.91 | 59.80 | 44.09 | 38.25 | 30.41 |
| | Spatio-temporal | **0.84** | **0.92** | **58.38** | **42.23** | **37.21** | **29.28** |

**时空特征的加入提高了模型表现**

$PM_{2.5}$估算的$R^2$, RMSE, and MAE 分别提高了0.04, -3.74 μg/m$^3$, and -2.27 μg/m$^3$

Wang et al. AE

# 建模改进——时空特征提取

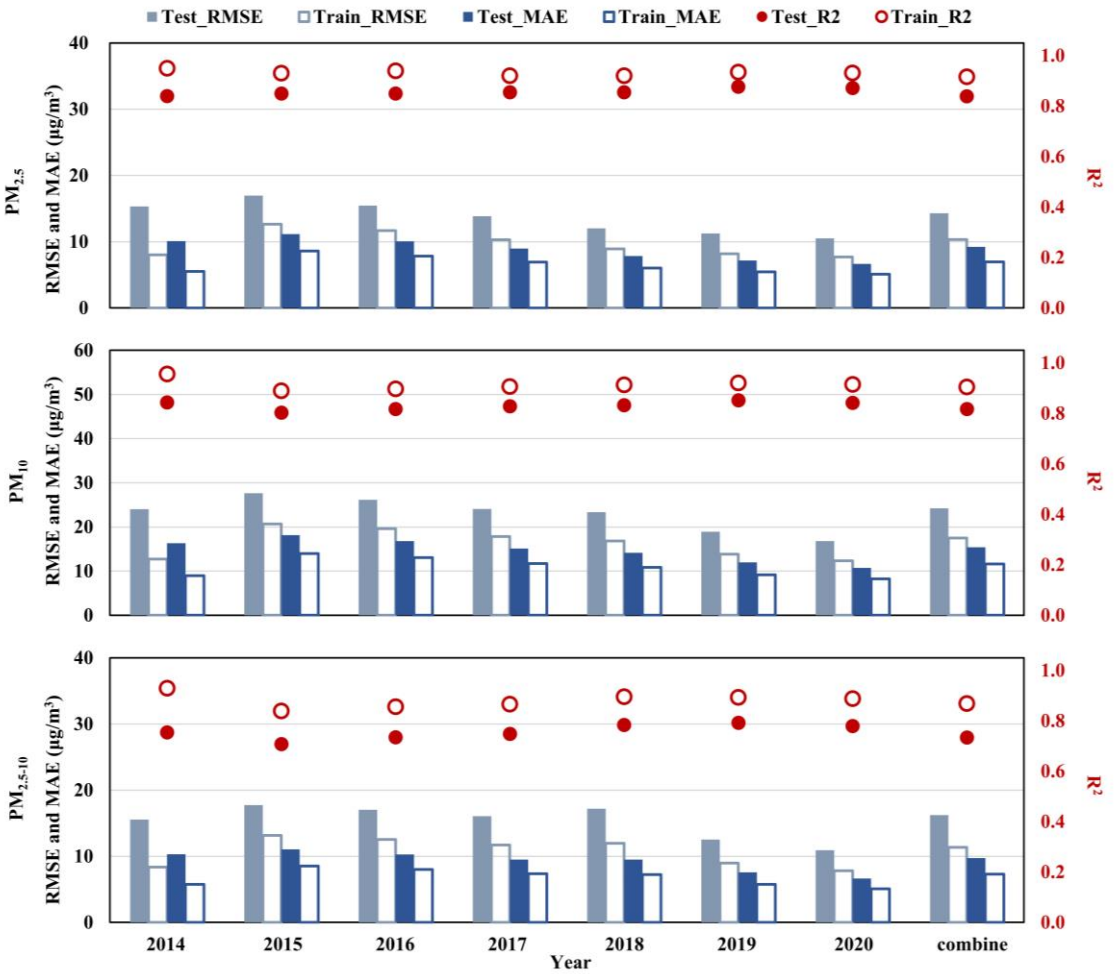



空间分布与CMAQ模拟结果相似，但**机器学习模型捕捉到了印度西部沙尘地区的高$PM_{2.5}$浓度**

In submission

# 长期数据集

**中国2000-2020年长期颗粒物浓度数据集：**利用气象再分析数据、卫星遥感产品、CMAQ模式结果、土地利用等数据，基于lightGBM模型，先填补AOD的缺失数据，然后对三种粒径段的颗粒物浓度进行估算（$PM_{2.5}$, $PM_{10}$, 和 $PM_{2.5-10}$）



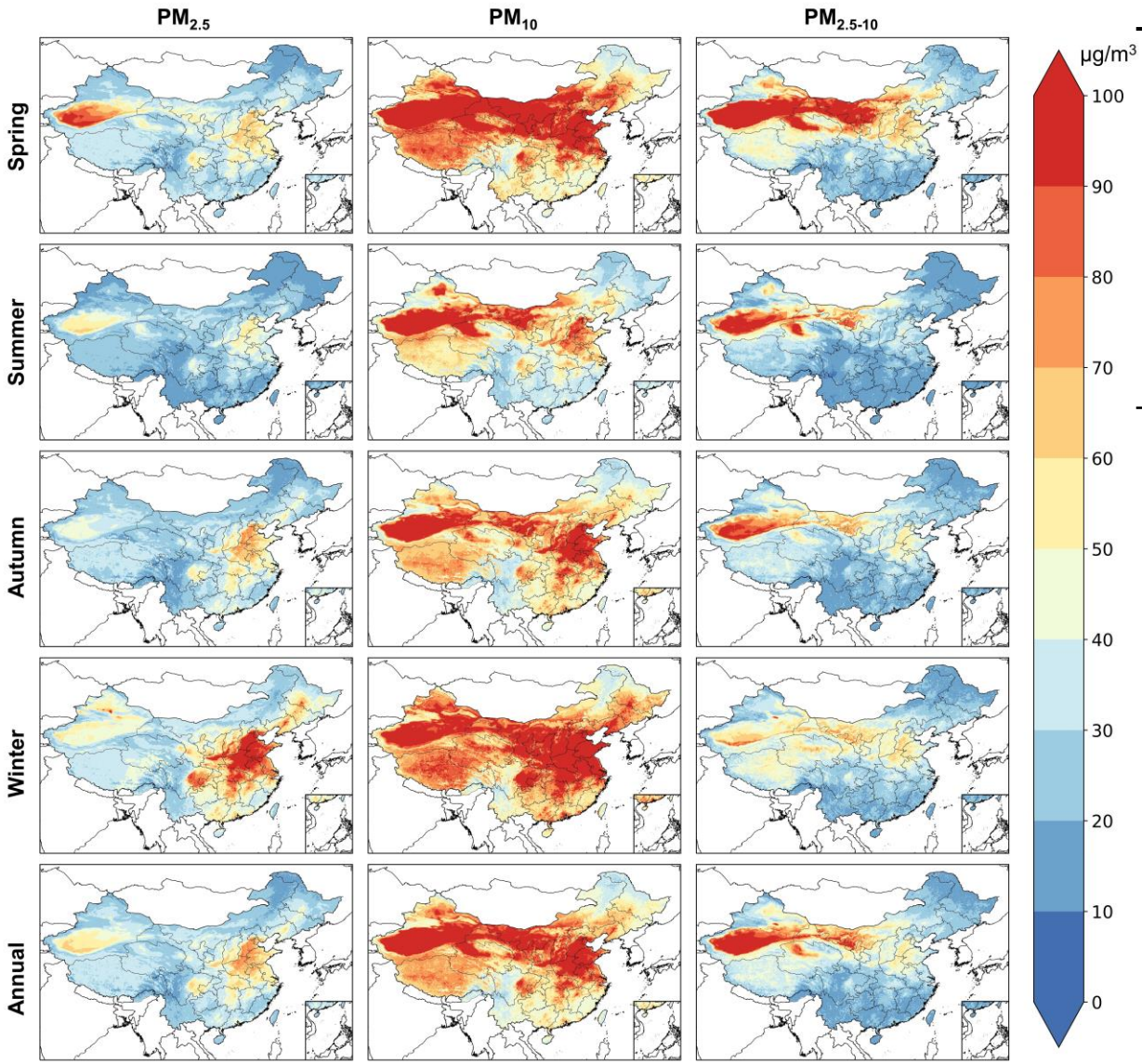

$PM_{2.5}$, $PM_{10}$, 和 $PM_{2.5-10}$ 2014-2020年模型训练和测试结果（out-of-sample）

不同统计模型之间训练和测试对比

| Model | Fit time | $R^2$ | | RMSE (μg/m³) | | MAE (μg/m³) | |
|---|---|---|---|---|---|---|---|
| | | Test | Train | Test | Train | Test | Train |
| MLR | 1.00 | 0.49 | 0.49 | 24.22 | 24.22 | 14.98 | 14.98 |
| PolyR | 37.16 | 0.64 | 0.64 | 20.31 | 20.16 | 12.59 | 12.52 |
| ERTs | 223.54 | 0.72 | 0.77 | 17.98 | 16.09 | 10.45 | 9.27 |
| RF | 1330.70 | 0.77 | 0.85 | 16.11 | 13.12 | 9.21 | 7.20 |
| XGB | 1175.64 | 0.83 | 1.00 | 13.91 | 0.00 | 8.00 | 0.00 |
| **LGB** | **26.24** | **0.83** | **0.88** | **13.85** | **11.51** | **8.70** | **7.83** |

- Out-of-sample验证结果：$PM_{2.5}$, $PM_{10}$, 和 $PM_{2.5-10}$ 模型准确，训练和测试误差的差异较小 (delta $R^2$: 0.07-0.12; delta RMSE: 4-6.7 μg/m³)。
- LightGBM和XGBoost具有相似的测试误差，但是前者过拟合情况更轻，并且速度更快；

# 长期数据集




| Species | Region | Spring | Summer | Autumn | Winter | Annual |
|---|---|---|---|---|---|---|
| $PM_{2.5}$ | China | 35.28 ± 3.17 | 25.07 ± 2.93 | 31.19 ± 3.89 | 43.95 ± 3.61 | 33.87 ± 7.68 |
| | BTH | 48.78 ± 5.40 | 39.70 ± 6.71 | 52.48 ± 7.42 | 77.30 ± 7.50 | 54.57 ± 15.52 |
| | YRD | 45.27 ± 4.78 | 32.22 ± 5.38 | 43.91 ± 6.54 | 67.78 ± 8.59 | 47.29 ± 14.43 |
| | PRD | 30.19 ± 3.26 | 19.90 ± 3.04 | 35.99 ± 6.49 | 46.64 ± 7.05 | 33.18 ± 11.03 |
| | SCB | 38.92 ± 7.16 | 27.83 ± 4.84 | 36.32 ± 7.11 | 61.43 ± 9.33 | 41.13 ± 14.4 |
| | NWCHN | 49.26 ± 8.37 | 33.05 ± 4.61 | 30.67 ± 3.46 | 37.68 ± 3.24 | 37.67 ± 8.90 |

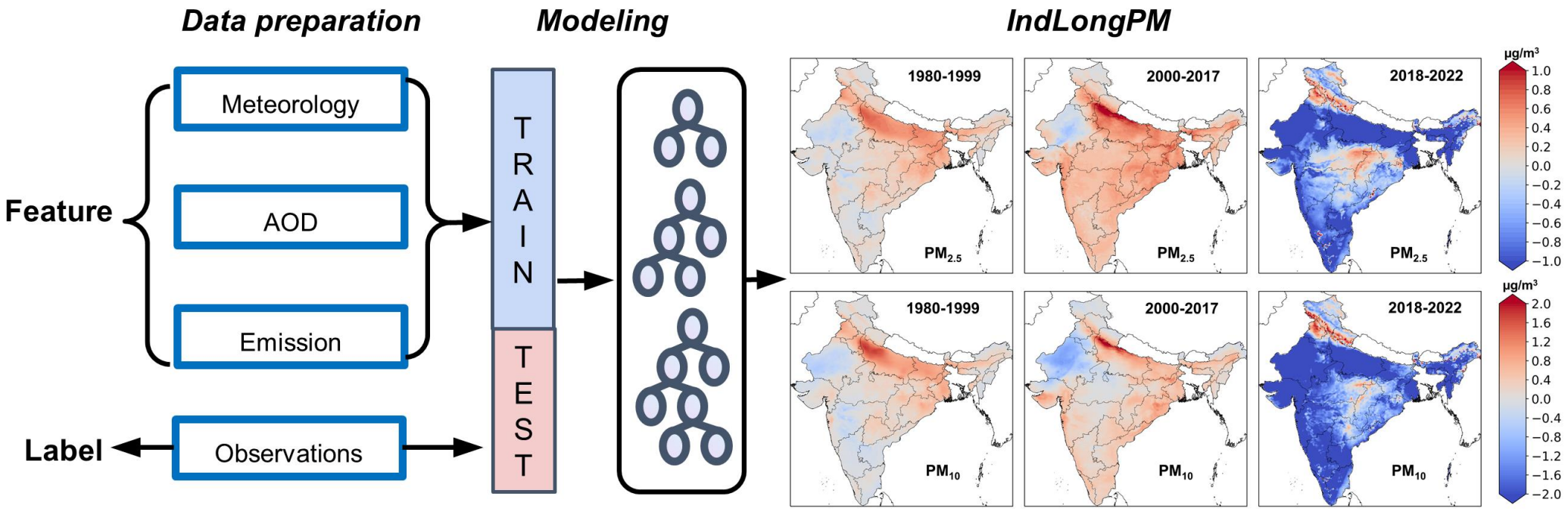
2000年至2020年中国三种粒径段颗粒物浓度季节均值，及其空间分布（10 km）：

- 春季西北大部分地区$PM_{10}$和$PM_{2.5-10}$浓度较高可能与频繁的沙尘事件有关；
- 冬季京津冀地区$PM_{2.5}$浓度较高，可能与污染物排放、不利的气象因素和独特的地形条件有关。

Wang et al., Chemosphere 2023

# 长期数据集

## 印度1980-2022年长期颗粒物浓度数据集：$PM_{2.5}$和 $PM_{10}$




- Machine learning model：**LightGBM**
- Meteorology: **ERA5-Land**; Emission: **MERRA2**
- Cross-validation: **out-of-sample CV, out-of-site CV, out-of-year CV**.
- **Train set: 2018-2022**; **test set: 202301-202306**.
- Mortality estimation: 2000-2019 (**GBD 2019**)

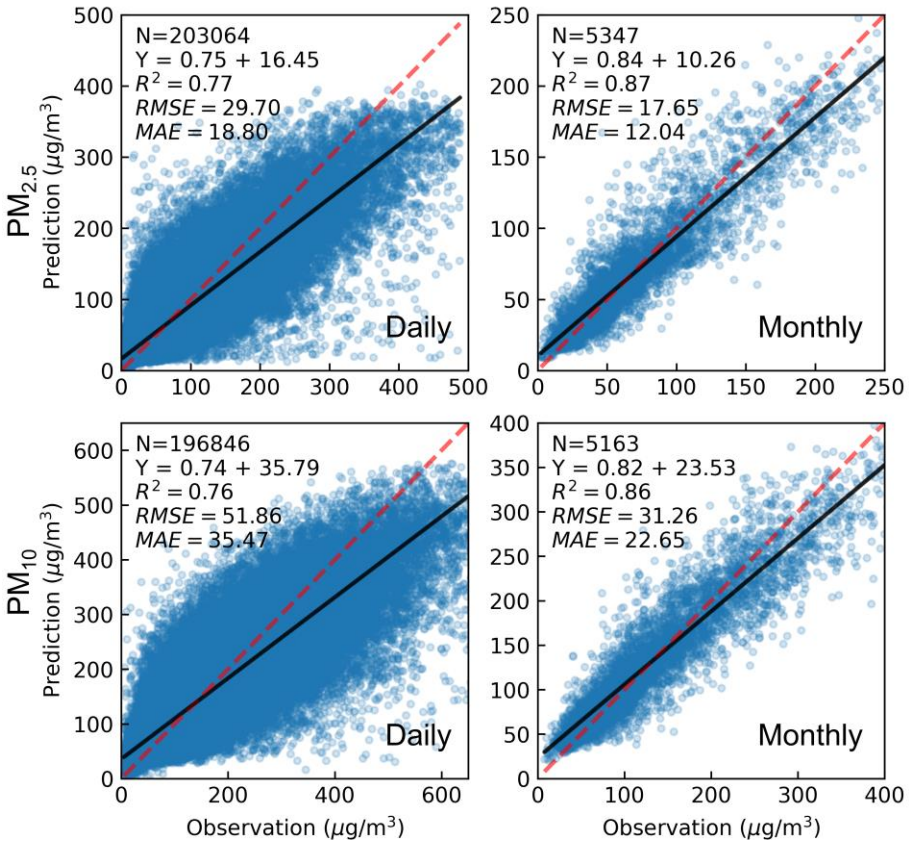
| Data description | IndLongPM dataset |
|---|---|
| Data type | Gridded |
| File format | NetCDF |
| Specie | $PM_{2.5}$, $PM_{10}$ |
| Spatial reference | WGS 84 |
| Horizontal resolution | 0.1° × 0.1° (≈ 10 km × 10 km) |
| Horizontal coverage | India, [60° E, 100° E], [5.0° N, 40.0° N] |
| Temporal coverage | 1980-2022 |

# 印度1980-2022年长期颗粒物浓度数据集：$PM_{2.5}$和 $PM_{10}$

Training and testing results of out-of-sample CV, out-of-site CV, and out-of-year CV for daily $PM_{2.5}$ and $PM_{10}$ (2018-2022). RSME and MAE unit: μg/m$^3$.

| Spec | Type | $R^2$ | | RMSE (μg/m$^3$) | | MAE (μg/m$^3$) | |
|---|---|---|---|---|---|---|---|
| | | Test | Train | Test | Train | Test | Train |
| $PM_{2.5}$ | out-of-sample | 0.77 | 0.79 | 29.57 | 28.51 | 18.76 | 18.25 |
| | out-of-site | 0.70 | 0.79 | 31.73 | 27.90 | 20.32 | 17.78 |
| | out-of-year | 0.66 | 0.79 | 35.35 | 27.61 | 21.54 | 17.61 |
| $PM_{10}$ | out-of-sample | 0.76 | 0.77 | 51.63 | 50.11 | 35.42 | 34.52 |
| | out-of-site | 0.65 | 0.77 | 57.37 | 49.42 | 39.92 | 33.94 |
| | out-of-year | 0.66 | 0.78 | 60.65 | 49.06 | 40.74 | 33.72 |

模型显示出良好的准确性，对于每日 $PM_{2.5}$ 和 $PM_{10}$，样本外 CV $R^2$ 分别为 0.77 和 0.76，RMSE 分别为 29.57 和 51.63 μg/m$^3$。

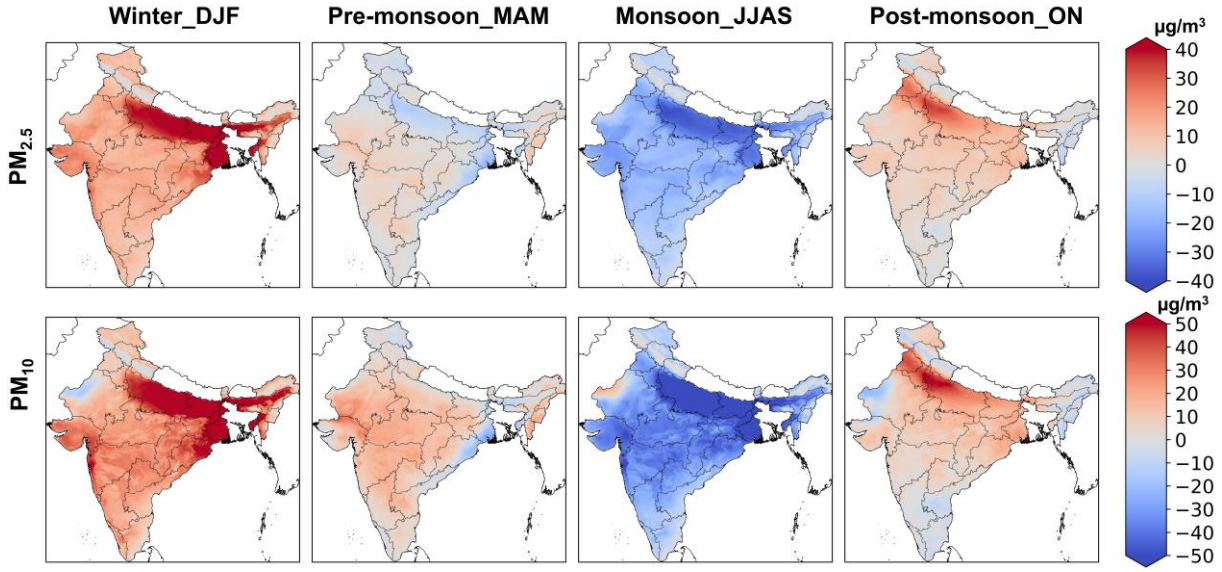



Comparison between observations and predictions of out-of-sample CV for daily and monthly $PM_{2.5}$ and $PM_{10}$. Dashed lines denote 1:1 line. Solid lines denote linear regression fitting. The sample numbers (N), linear regression function, $R^2$, RMSE, and MAE are also shown. Units of RMSE and MAE are μg/m$^3$.

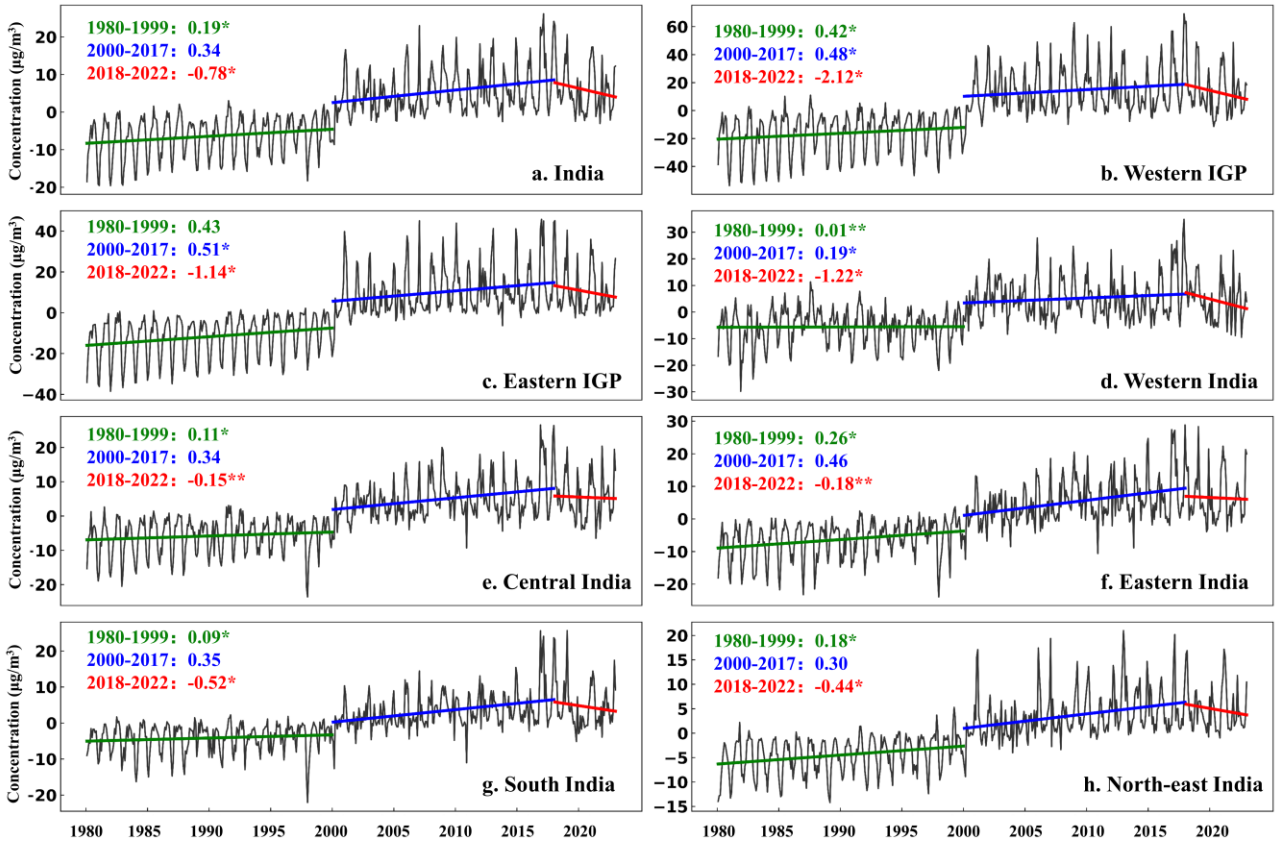
Wang et al., ESSD

# 长期数据集建立—印度

## 印度1980-2022年长期颗粒物浓度数据集：$PM_{2.5}$和 $PM_{10}$




Spatial patterns of seasonal $PM_{2.5}$ and $PM_{10}$ anomalies (the difference between seasonal mean and annual mean) in India during 1980-2022.



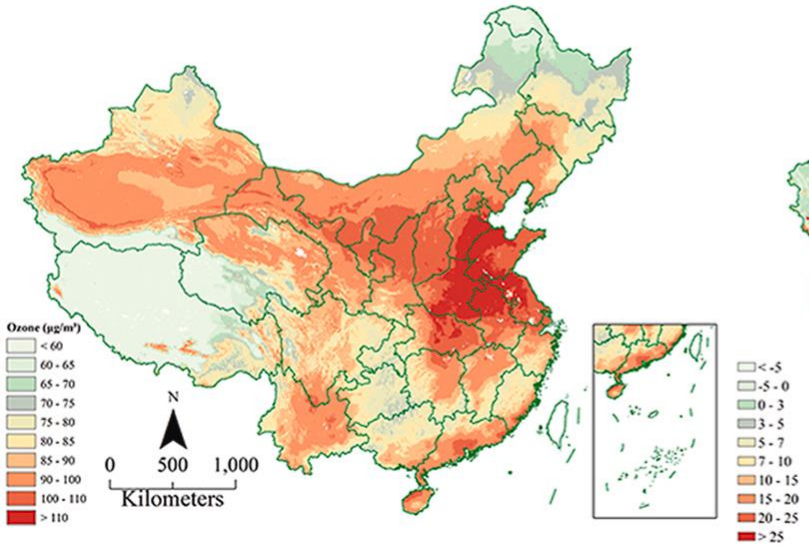

- 高浓度：
  冬季，印度恒河平原
- 低浓度：
  雨季，印度南部

- **1980-1999：$PM_{2.5}$浓度缓慢增加**
- **2000-2017：快速增加**
- **2018-2022：快速下降**

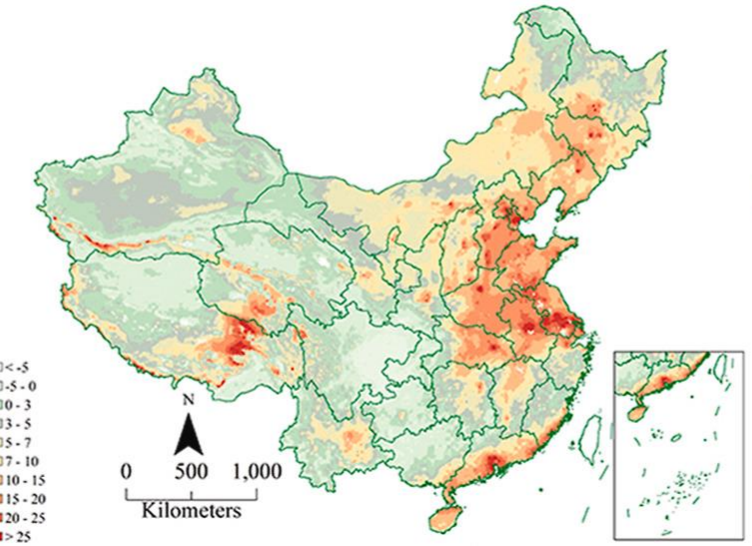
Wang et al., ESSD

# $O_3$和$NO_2$

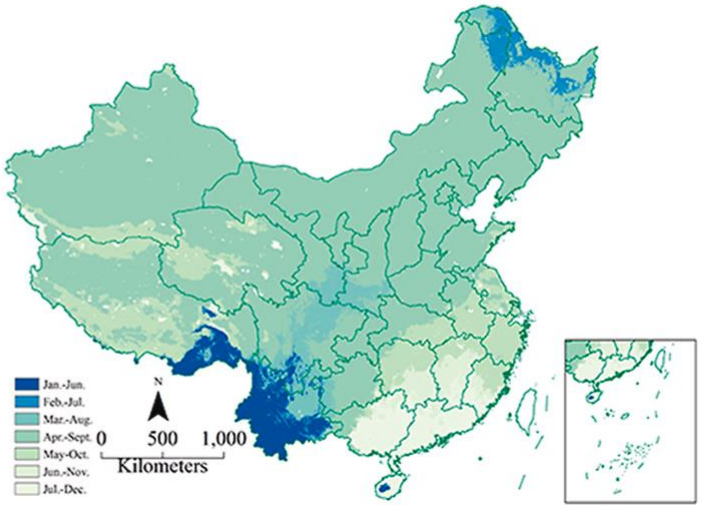
The annual mean MDA8 ozone concentrations in 2019



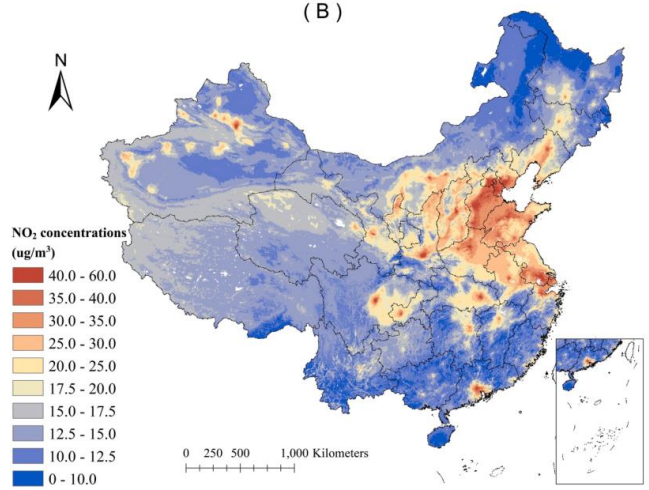

Change rates of annual mean MDA8 ozone concentrations from 2013 to 2019

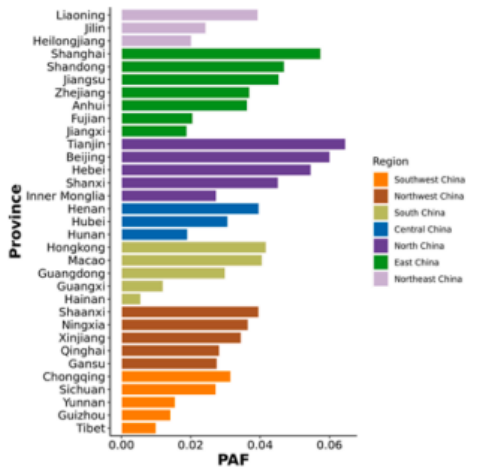



The spatial distribution of ozone peak season with different months in 2019

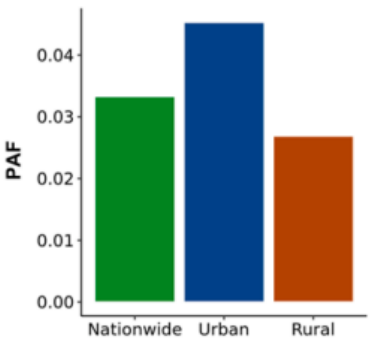



( B )

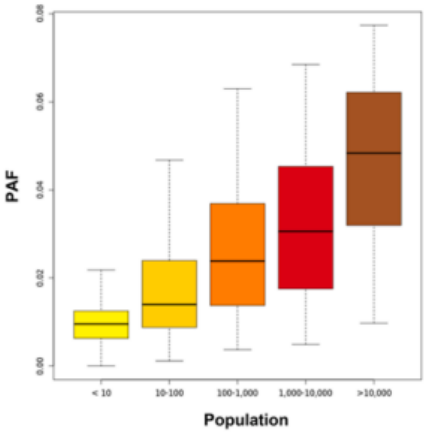



( A )




( B )




( C )

**Fig. 4.** Distribution of population attributable fraction for $NO_2$ long-term exposure in 2019. (A) Among provinces; (B) between urban and rural areas; and (C) in different population.

Meng et al., EP

Li et al., EI

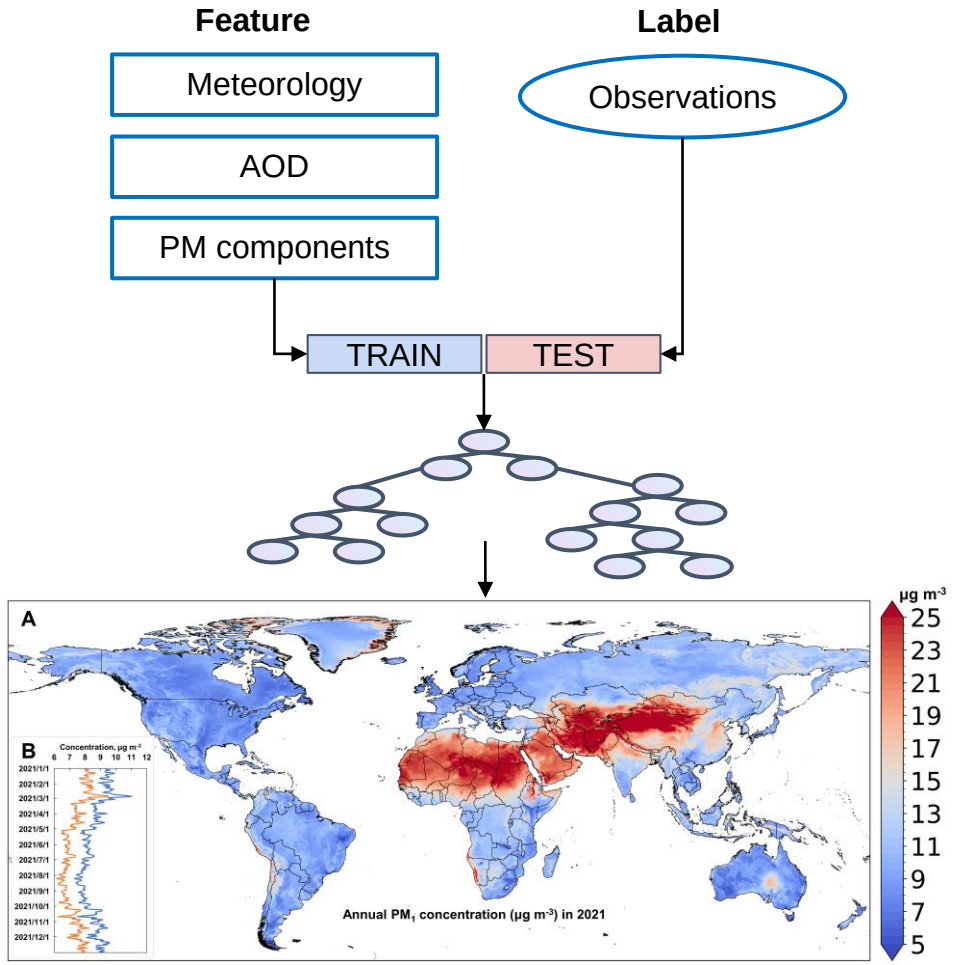
# $PM_1$全球分布

## 基于4D树模型的全球$PM_1$浓度估算

**Data sources** OpenAQ and ERA5+MERRA2

| Type | Variable | Description | Spatial Resolution |
|---|---|---|---|
| ERA5 | SSRD | Surface solar radiation | 0.1° × 0.1° |
| | EVAP | Evaporation | 0.1° × 0.1° |
| | TEMP2 | 2m air temperature | 0.1° × 0.1° |
| | DEWP2 | 2m dewpoint temperature | 0.1° × 0.1° |
| | SP | Surface pressure | 0.1° × 0.1° |
| | TPREC | Total precipitation | 0.1° × 0.1° |
| | UWIND10 | 10m u component of wind | 0.1° × 0.1° |
| | VWIND10 | 10m v component of wind | 0.1° × 0.1° |
| | BLH | Boundary layer height | 0.25° × 0.25° |
| | TCLOUD | Total cloud cover | 0.25° × 0.25° |
| MERRA2 | BCSMASS | Black Carbon | 0.5 ° × 0.625 ° |
| | OCSMASS | Organic Carbon | 0.5 ° × 0.625 ° |
| | DUSMASS25 | Dust– $PM_{2.5}$ | 0.5 ° × 0.625 ° |
| | DUSMASS | Dust | 0.5 ° × 0.625 ° |
| | SO2SSMASS | sulfur dioxide ($SO_2$) | 0.5 ° × 0.625 ° |
| | SO4SMASS | Sulfate | 0.5 ° × 0.625 ° |
| | TOTEXTTAU | Total Aerosol Extinction | 0.5 ° × 0.625 ° |

**Flow Chart**

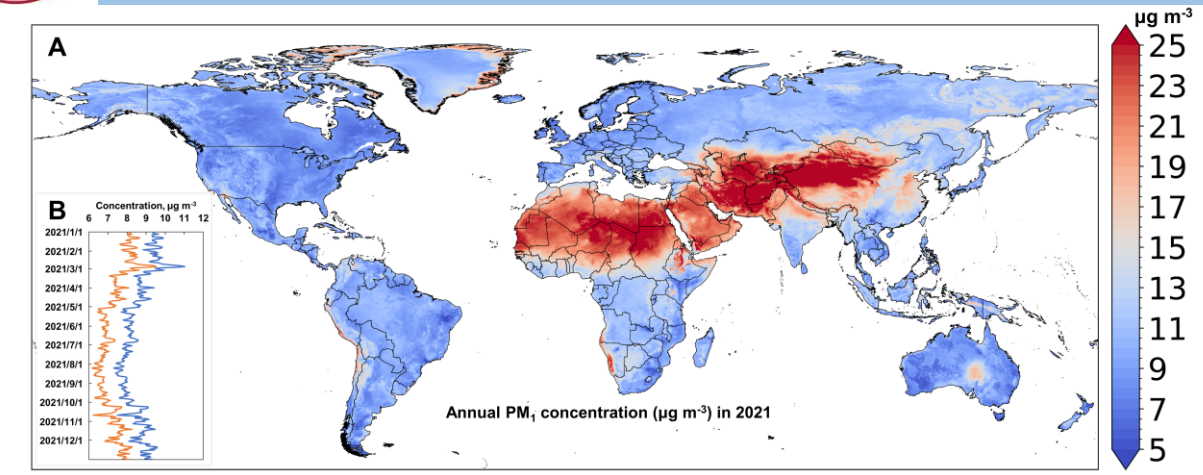


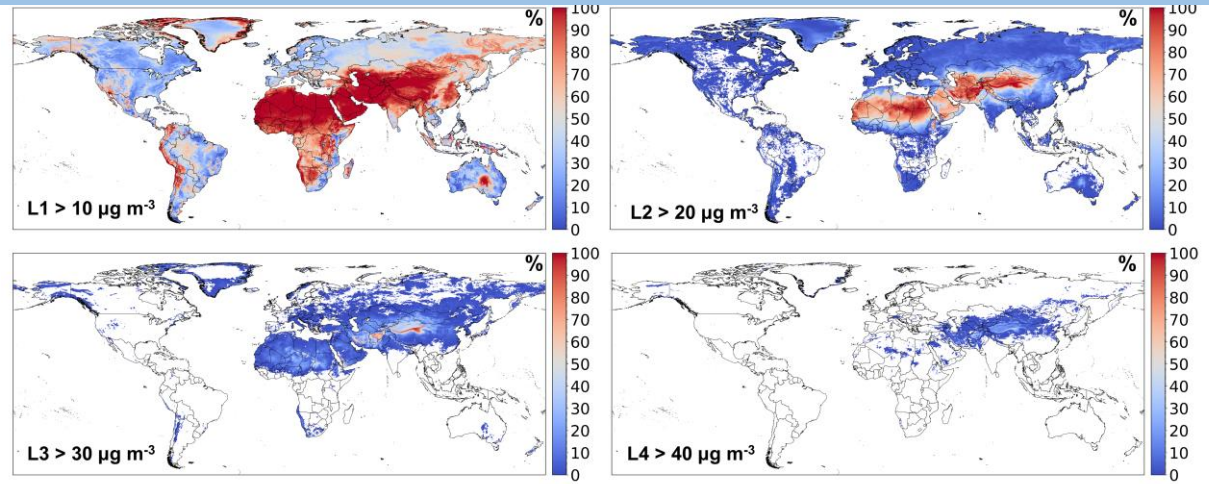

**Disparities in global $PM_1$ exposure:**

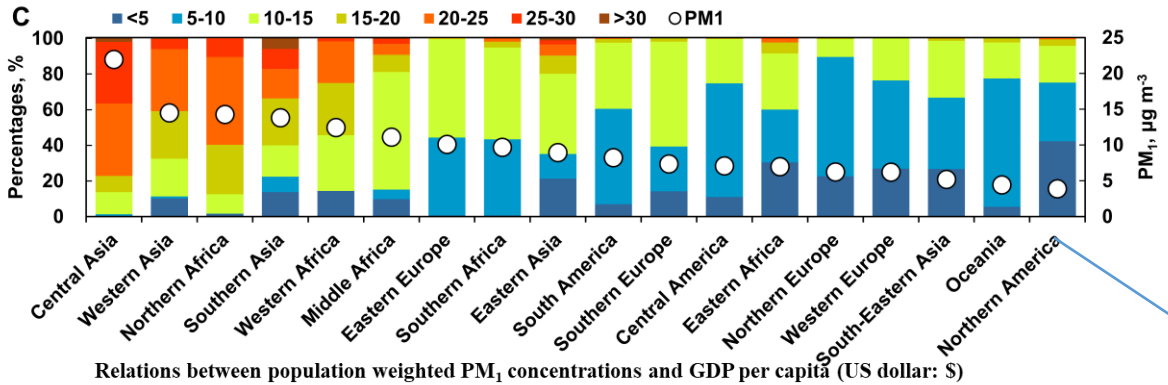
- 划分了四个等级的目标值：L1>=10 μg $m^{-3}$ , L2 =20 μg $m^{-3}$, L3=30 μg $m^{-3}$, L4=40 μg $m^{-3}$ 。
- 短期暴露风险的计算方法：计算一年中每日$PM_1$浓度超过标准值的天数比例；
- 长期暴露风险的计算方法：计算年$PM_1$浓度超过目标值的区域比例

*in preparation*

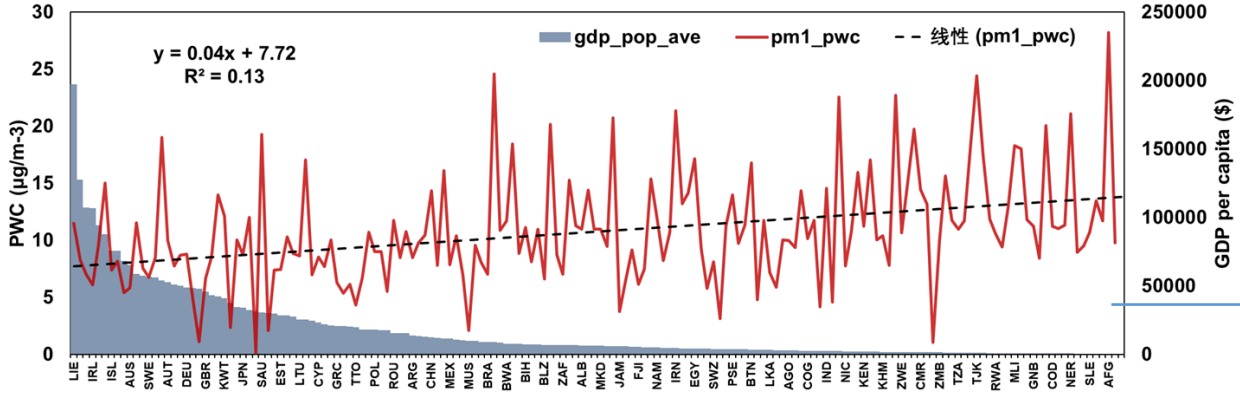
# $PM_1$全球分布

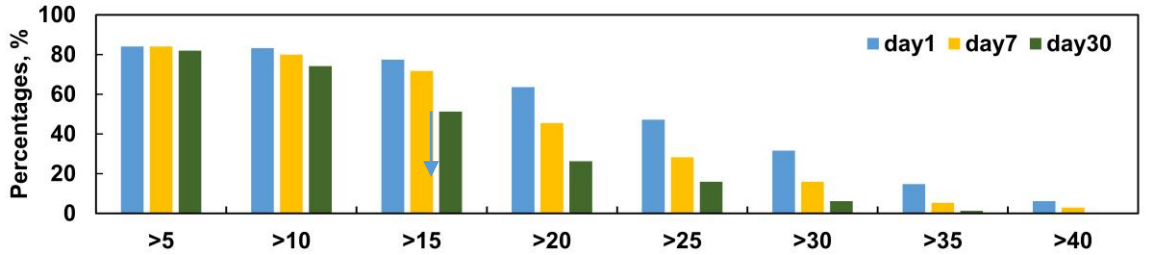



Relations between population weighted $PM_1$ concentrations and GDP per capita (US dollar: $)



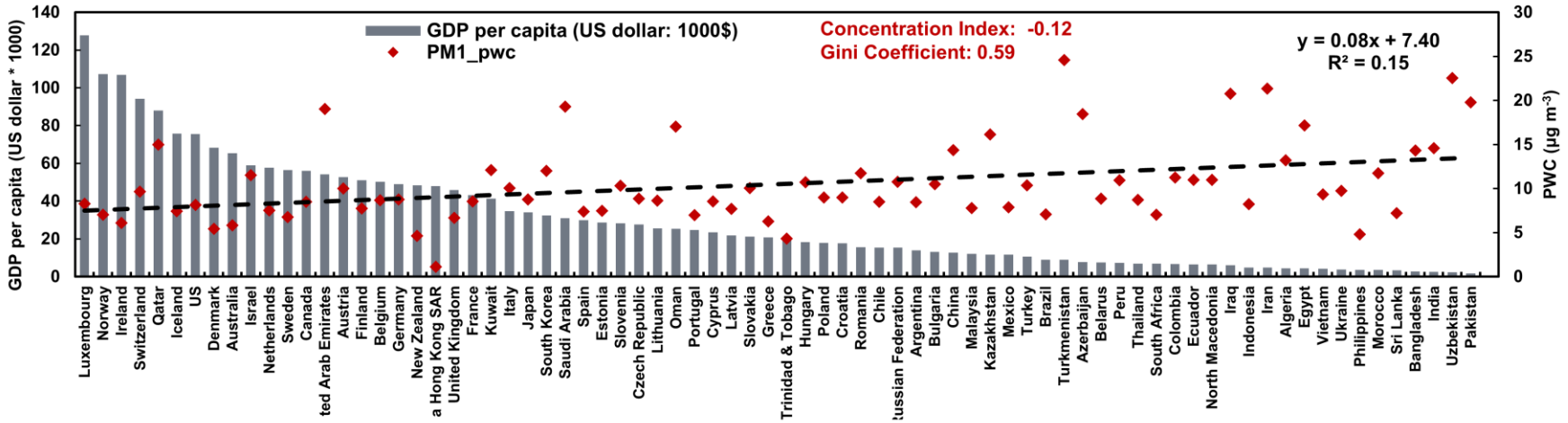

**短期暴露风险：**2021 年全球分别有 77%、71% 和 51%的人至少有一天、一周和一个月暴露在不健康的 $PM_1$（> 15 μg m-3）中。

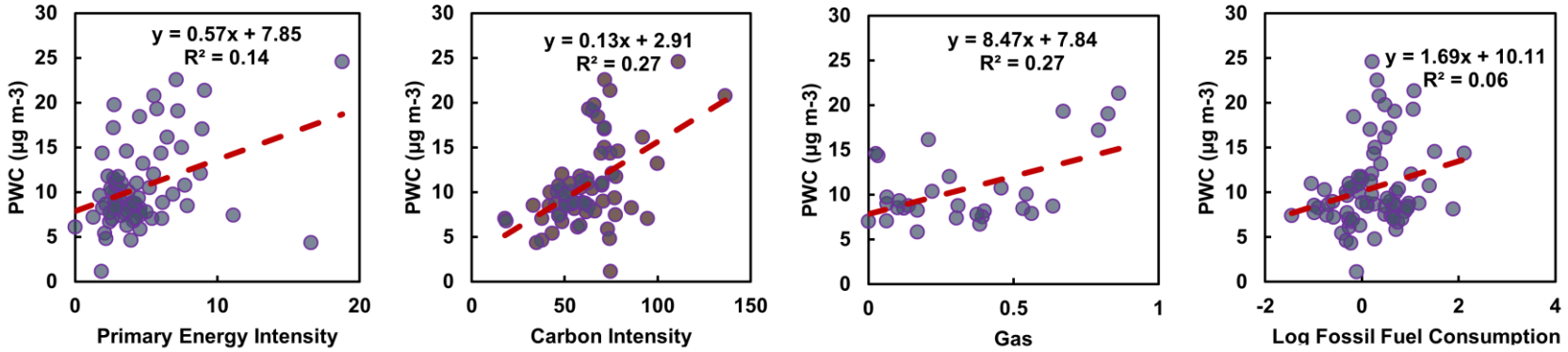
**长期暴露风险:** 2021 年中亚98%的地区，北非98%的地区，南亚79%的地区暴露在不健康的 $PM_1$（年均浓度> 10 μg m-3）中，而北美仅25%的地区，澳大利亚仅22%的地区，西欧仅23%的地区暴露在不健康的 PM1（年均浓度> 10 μg m-3）中。

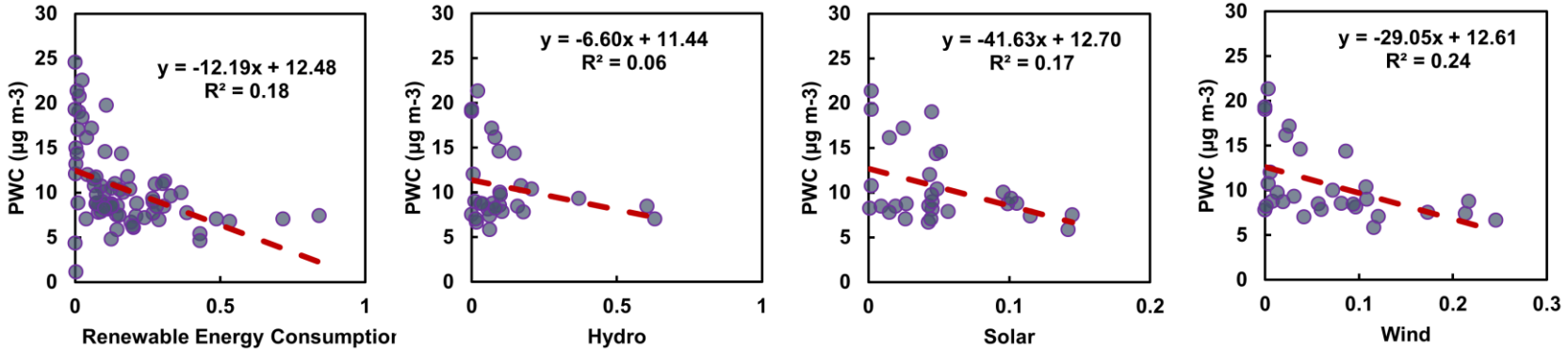
在全球范围内，经济水平与人口加权 $PM_1$ 浓度呈负相关；

经济发展水平低的国家和地区$PM_1$暴露水平更高

# $PM_1$ disparities




1. 在全球范围内，经济水平与人口加权 $PM_1$ 浓度呈负相关；基尼系数为 0.59，集中指数为-0.12；

2. 经济发展水平低的国家和地区 $PM_1$暴露水平更高；

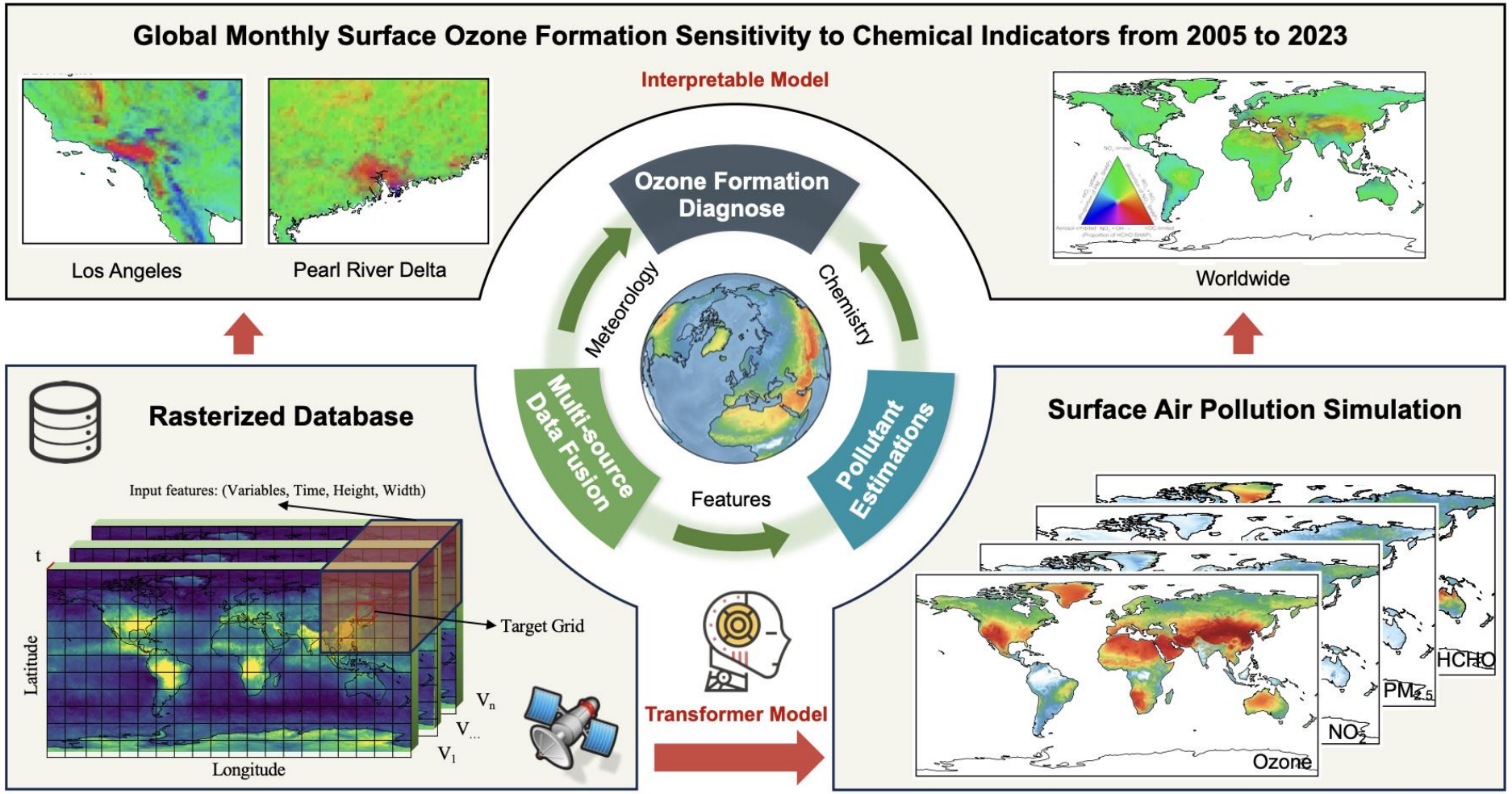
3. 与一次能源相关的指标，如碳强度、天然气发电比例和化石燃料消耗量与 PWC 呈正相；

4. 水力、太阳能和风能等清洁能源与 PWC 呈负相关。

*in preparation*

# 全球NOx-VOCs-$O_3$-$PM_{2.5}$

## 可解释机器学习诊断臭氧对VOCs (HCHO)、$NO_x$ ($NO_2$)和气溶胶($PM_{2.5}$)的敏感性




气象再分析

卫星遥感

人口和路网

植被指数

土地利用

10km分辨率

2005-2023年

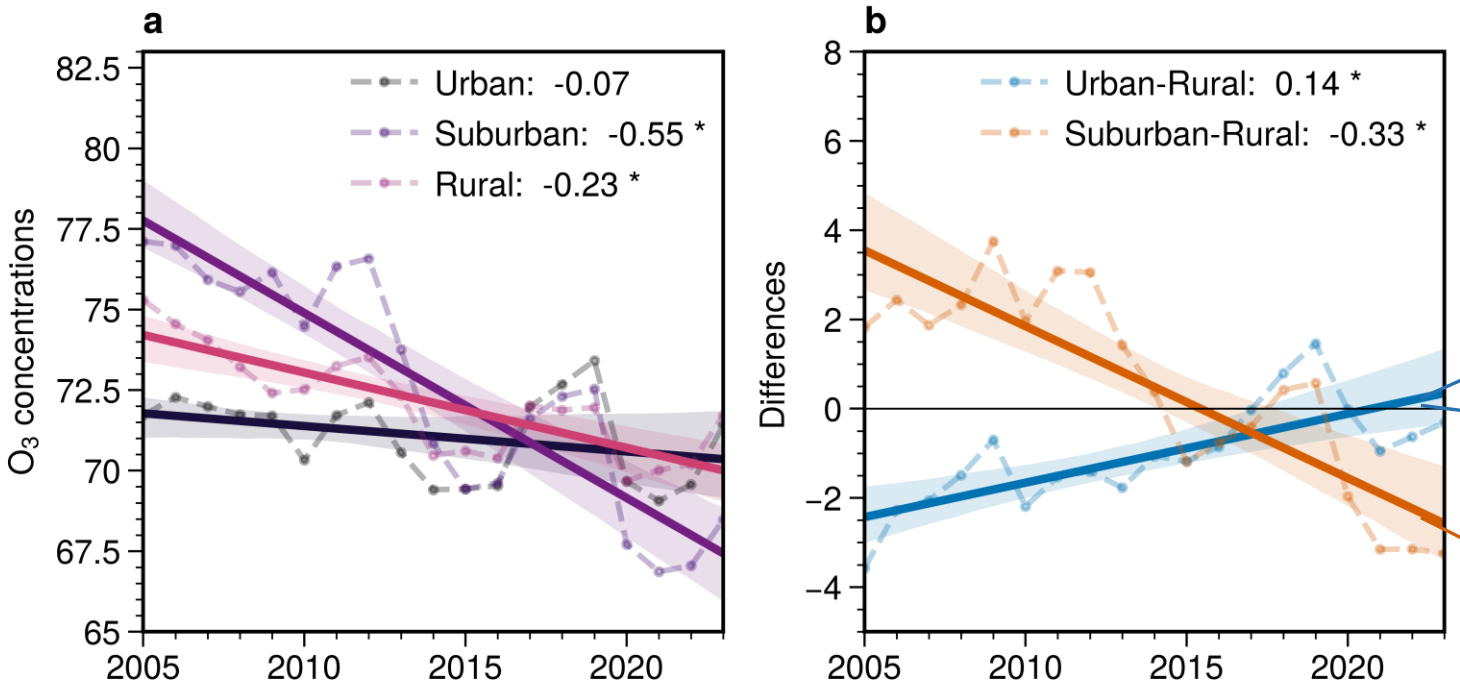
多源时空数据融合

多任务深度学习估算全球大气污染物浓度

*in preparation*

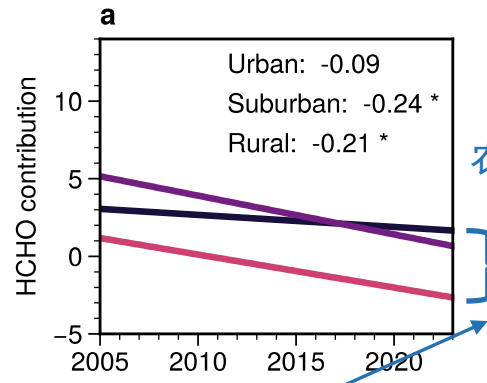
# 全球臭氧城乡差异的时空变化趋势

**由于农村地区臭氧的下降幅度更大，导致城乡差异不断缩小**



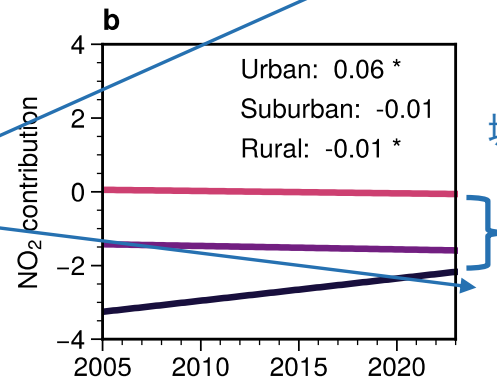

**2005年臭氧：郊区 > 农村 > 城市**

**2023年臭氧：农村 > 城市 > 郊区**



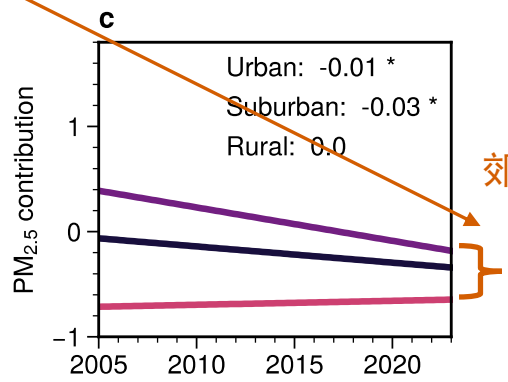

农村HCHO抑制效应增强

0.13 μg $m^{-3}$ $yr^{-1}$ *




城市NO2抑制效应减弱

0.07 μg $m^{-3}$ $yr^{-1}$ *

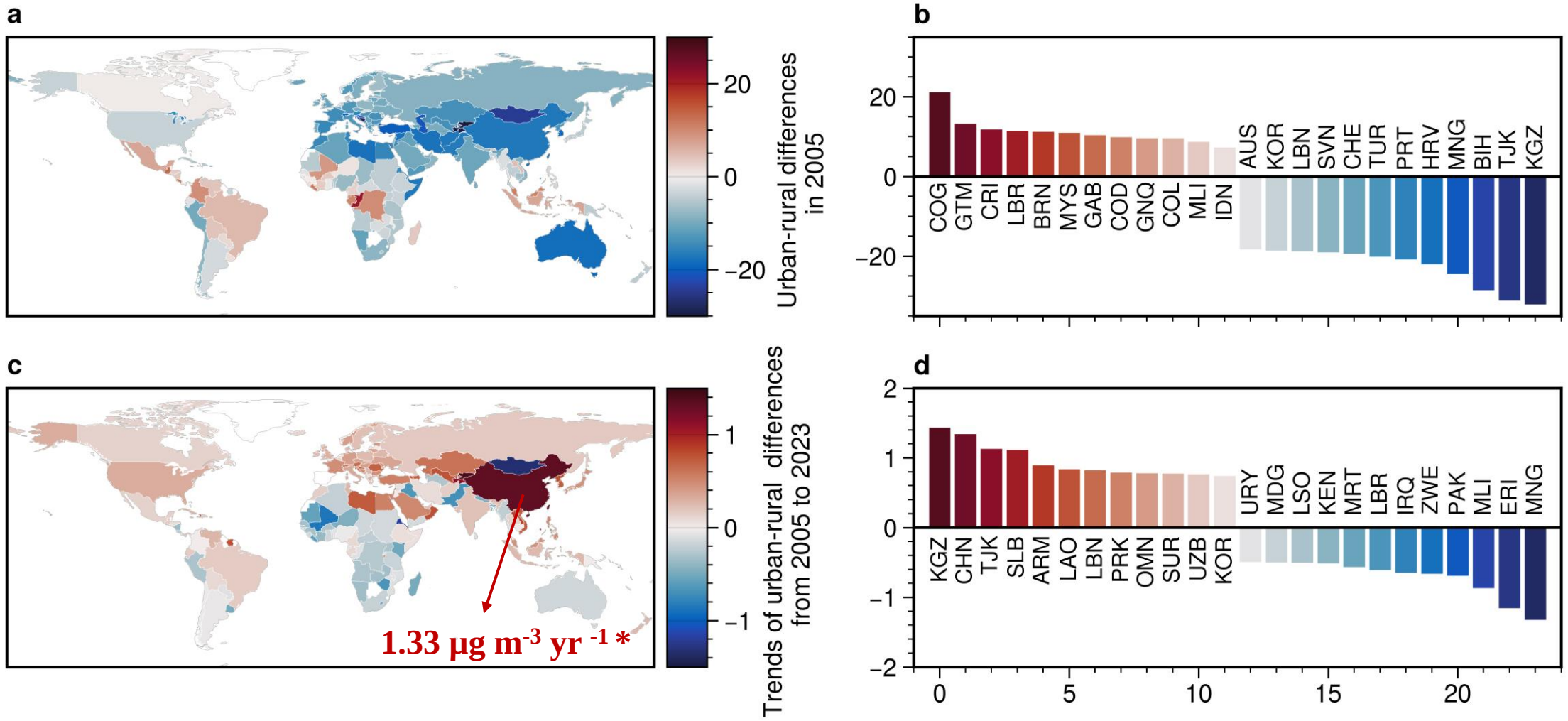



郊区颗粒物抑制效应增强

-0.03 μg $m^{-3}$ $yr^{-1}$ *

*in preparation*

# 国家尺度臭氧城乡差异的时空变化趋势



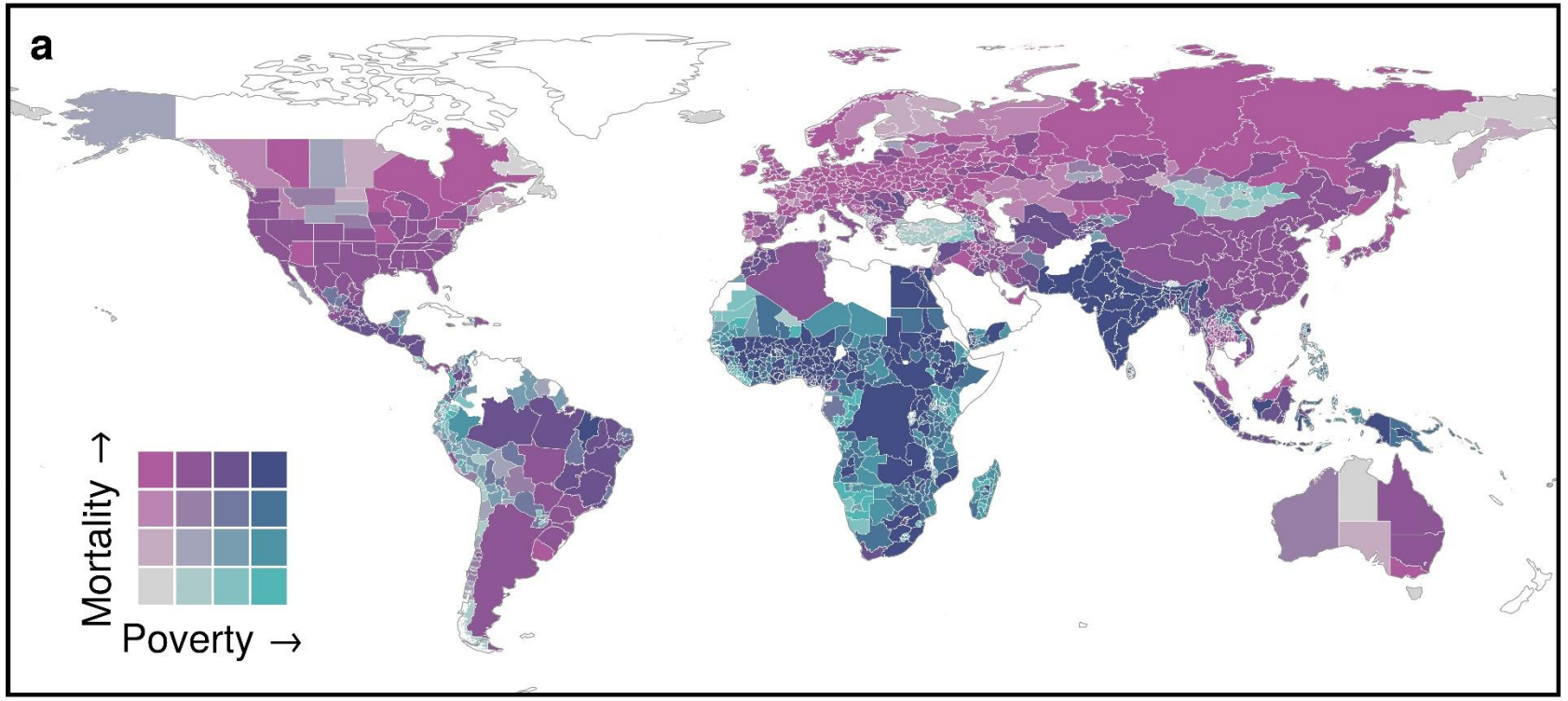

- **2005年，超过75%的国家城市臭氧低于农村**
- **2005-2023年，超过60%的国家臭氧城乡差异呈增加趋势**

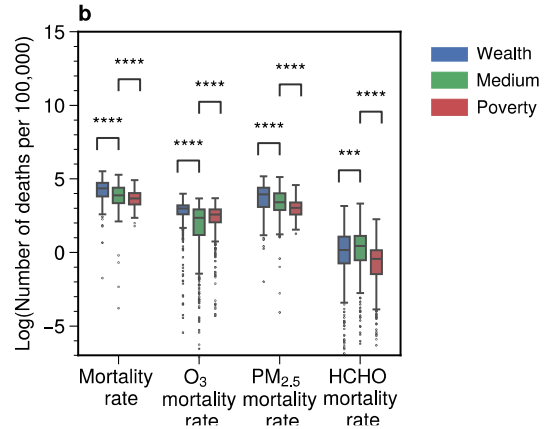
*in preparation*

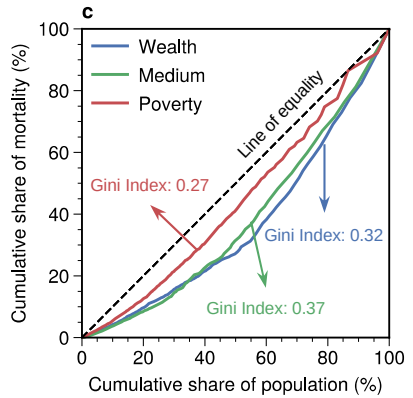
# 大气污染物过早死亡的社会经济不平等效应

**社会经济加剧了空气污染不平等，南亚和非洲低收入群体的二次污染物暴露死亡脆弱性更高**

**贫困率更低的富裕国家的归因于长期污染物暴露的总过早死亡率更高，且内部不平等性更高**




贫困和大气污染过早死亡的关联

Poverty: 2017年国际价格下，生活在消费或人均收入低于3.65美元贫困线家庭的人口占地区总人口的百分比

*in preparation*

# 全球健康负担分布呈现双重异质性

**高收入国家内部呈现逆社会经济梯度，即社会经济优势群体承受更显著的污染暴露疾病负担**

**欠发达地区结构性贫困持续强化着底层民众的健康负担与营养不良复合危机**

**红色：**

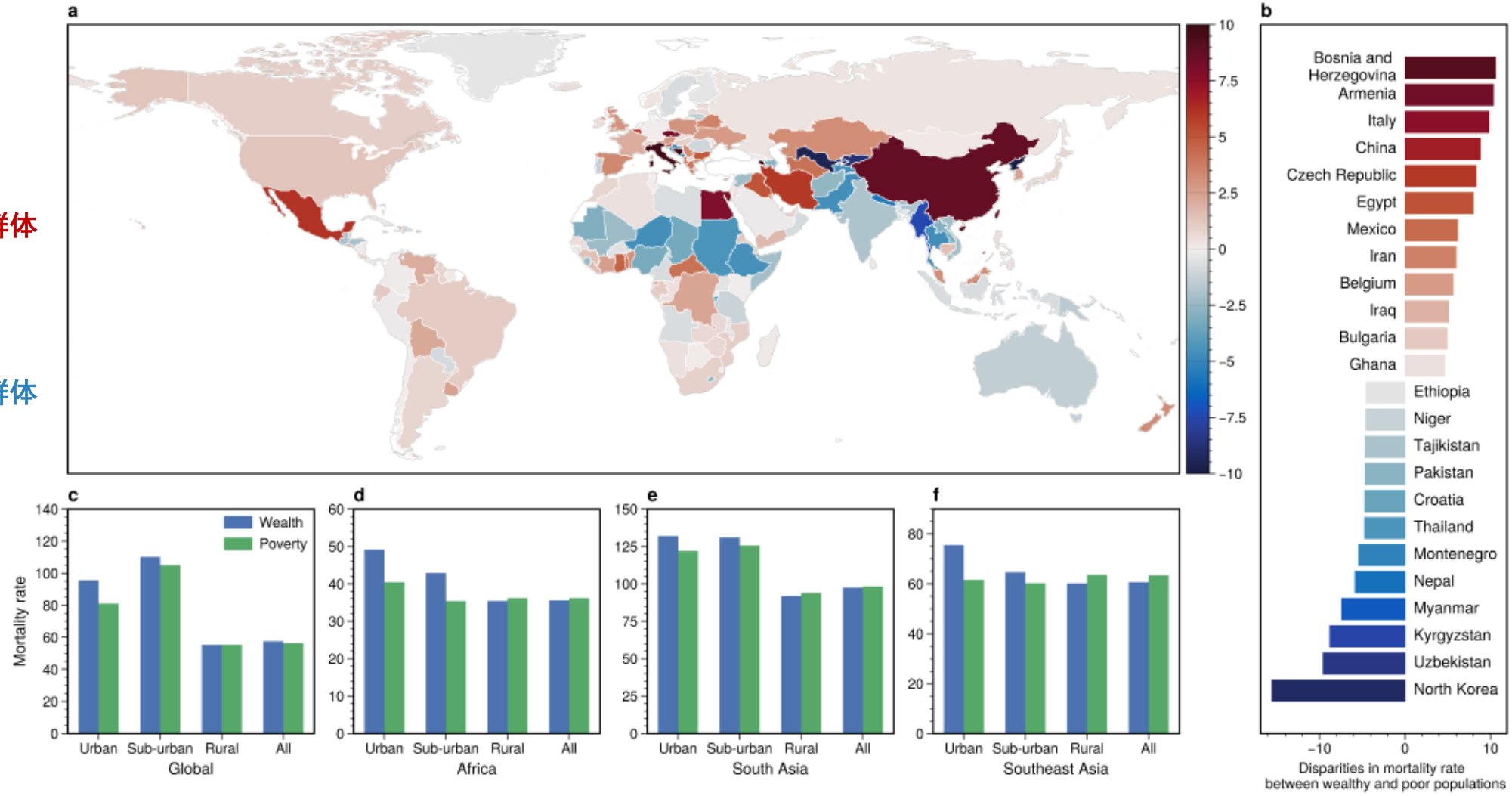
**富裕群体 > 贫困群体**

**(过早死亡率)**

**蓝色：**

**贫困群体 > 富裕群体**

**(过早死亡率)**



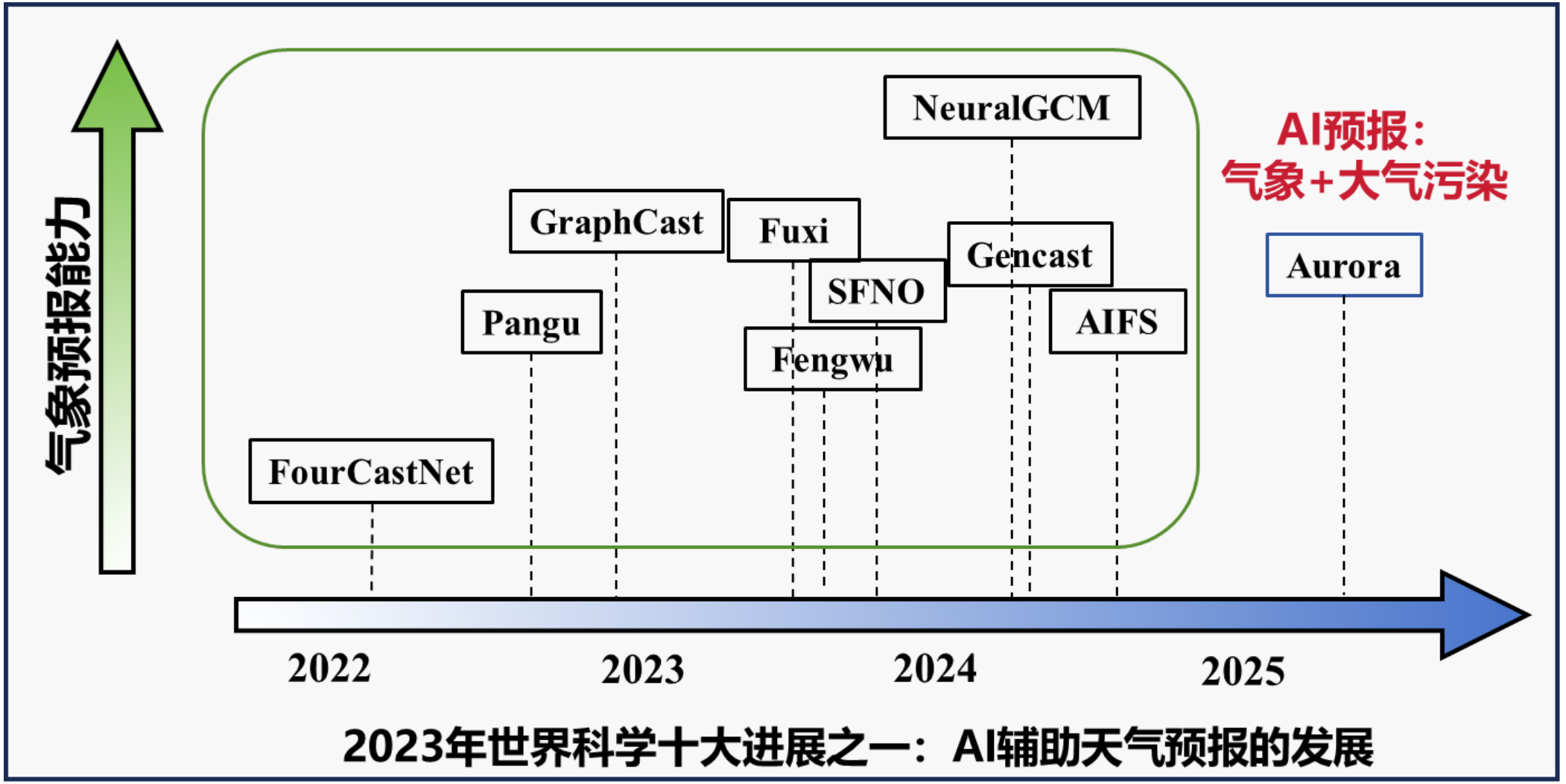

*in preparation*

# 更有价值的是什么？

- 人工智能辅助数值模拟
  - 更准确的污染物浓度
  - 不等于更准确的暴露
  - 室内室外、人的移动
- 人工智能替代数值模拟
- 数值模拟辅助人工智能

# 第二阶段：人工智能替代化学传输模型

**基于人工智能的气象预报获突破性进展，给大气污染智能预报带来机遇**



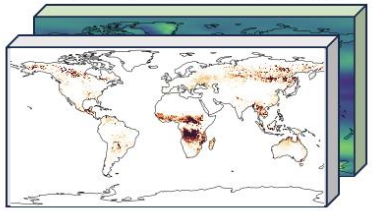

2023年世界科学十大进展之一：AI辅助天气预报的发展

# 深度学习模拟并归因排放-浓度响应关系

**Pre-training**

**Step1: Pre-training deep learning model on historical simulations of the climate-chemistry model**

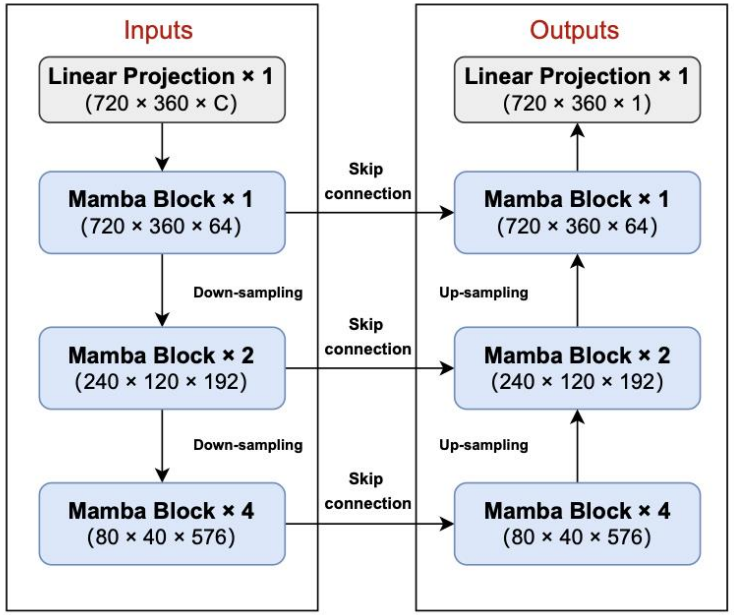
GFDL_ESM4 Meteorology



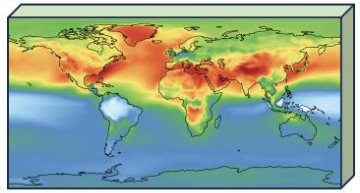
Input4MIPs Emissions

1850-2014

**Mamba-based Emission-Concentration Response Model**



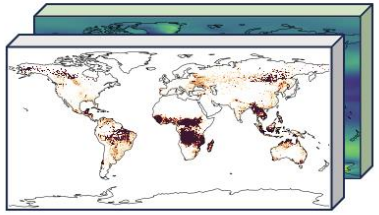

GFDL_ESM4 Surface Ozone



**Fine-tuning**

**Step2: Fine-tuning on reanalyze data**

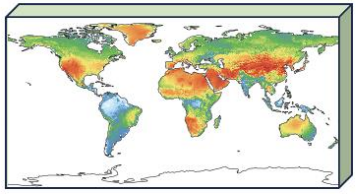
ERA5 Meteorology



CAMS Emissions

2005-2019

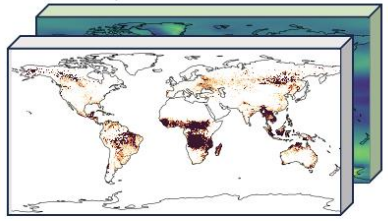
Observation + DL Surface Ozone



**Prediction and Attribution**

**Step3: Predicting of future ozone and revealing the contribution of biomass combustion**

Multiple CMIP6 Model Meteorology



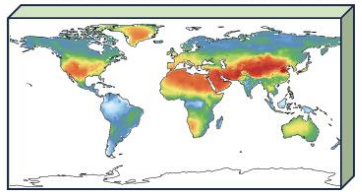
Input4MIPs Emissions

2020-2100

**Integrated Gradients Interpretable approach**

SSP119, SSP126, SSP245, SSP370, SSP585

Predicted Future Surface ozone



*in preparation*

# 深度学习预测未来碳中和路径下$PM_{2.5}$源贡献

模型训练

**MEIC排放清单**

$PM_{2.5}$、$PM_{10}$、$NH_3$、NOx、BC、OC、$SO_2$、CO、NMVOC

**WRF气象**

**气象因子：**风速、风向、降水、温度、湿度等

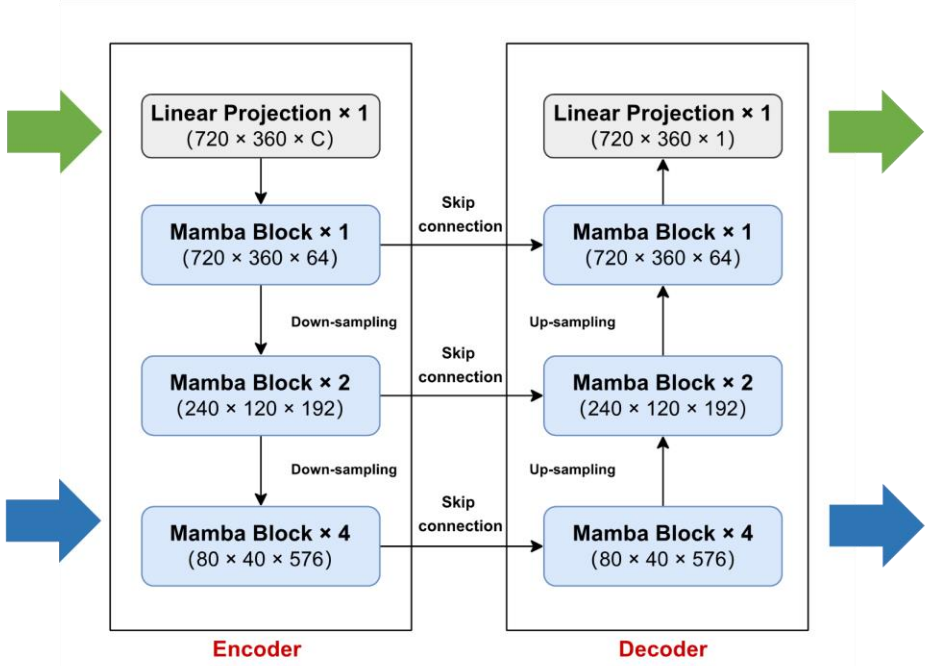
预测+分析

**DPEC排放清单**

**CMIP6未来气象**

SSP情景

## Mamba-Unet 排放-浓度响应模型

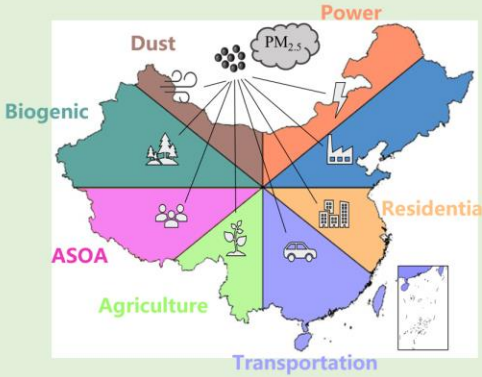



1、未来$PM_{2.5}$源贡献的时空变化

2、未来$PM_{2.5}$源贡献的健康效益

3、气候变化对$PM_{2.5}$源贡献的影响分析

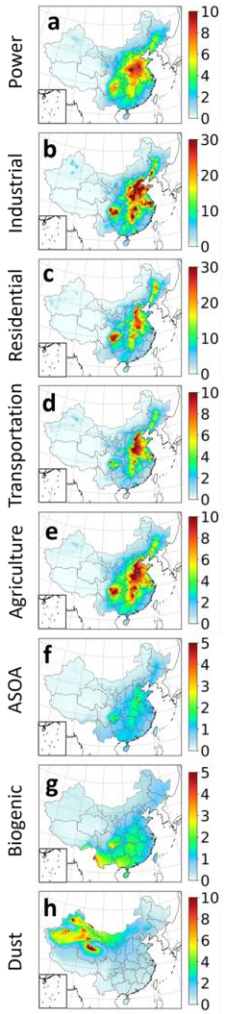
**CMAQ-$PM_{2.5}$源贡献**




**深度学习-$PM_{2.5}$源贡献**

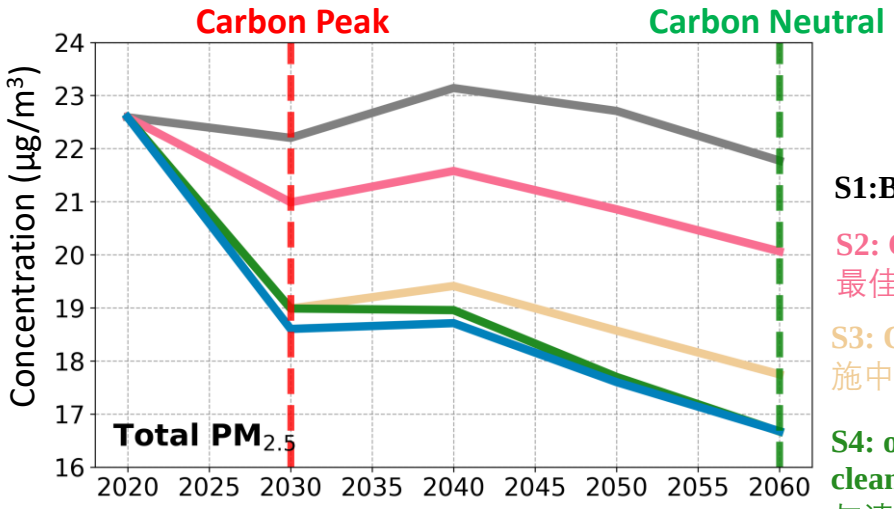
**来源：**工业、居民、交通、能源、农业、SOA…

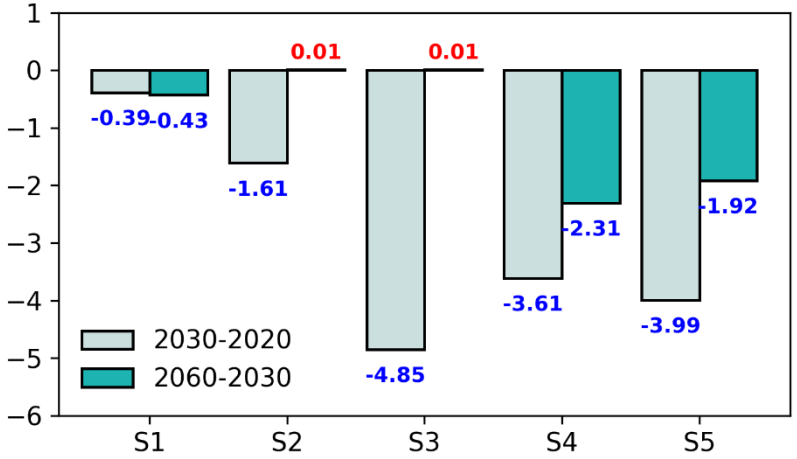
**多套排放情组合景**




*in preparation*

# $PM_{2.5}$源贡献的变化趋势



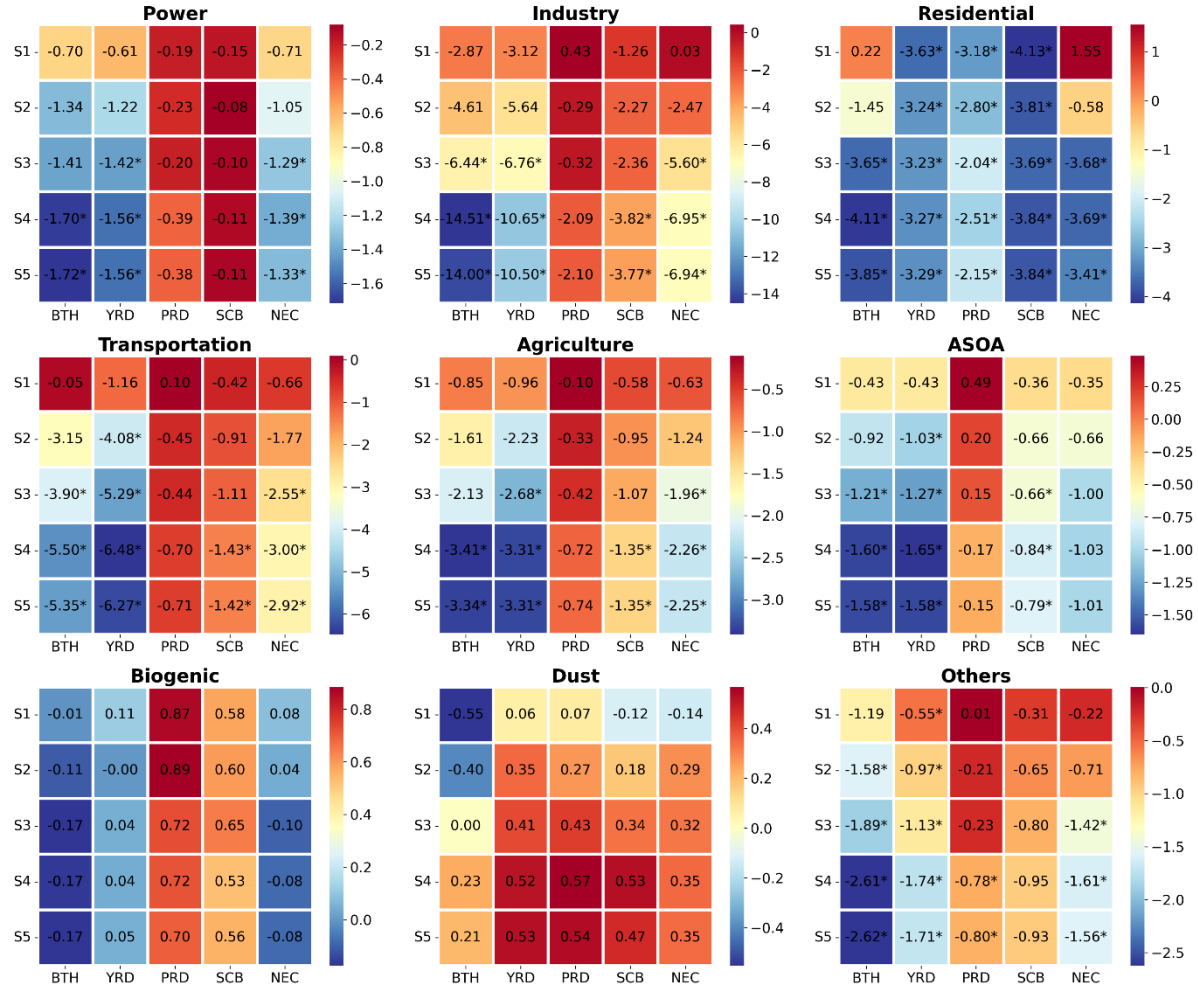

**S1:Baseline**

**S2: Clean air:** 逐步实施最佳污染控制

**S3: On-time clear air:** 实施中短期碳达峰减排政策

**S4: on-time peak-net zero-clean air:** 碳达峰、碳中和与清洁空气协同情景

**S5: early peak-net zero-clean air:** 强化了2030年前的碳减排政策



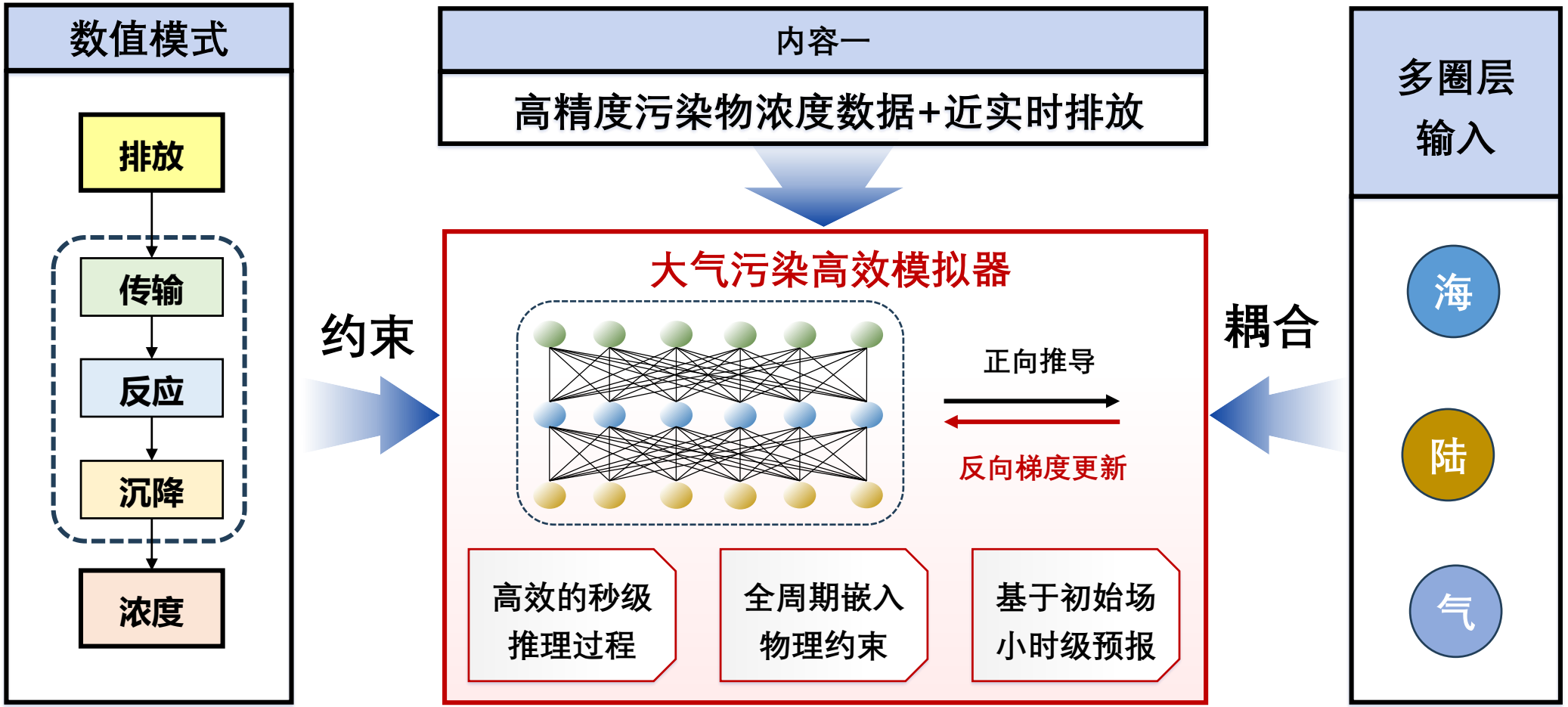

碳中和情景下全国$PM_{2.5}$的变化

碳中和情景下区域$PM_{2.5}$的变化趋势

*in preparation*

# 第三阶段：数值模拟辅助人工智能

## 深度学习驱动的高效精准数据模型



# 未来：数据驱动+大预言交互模型

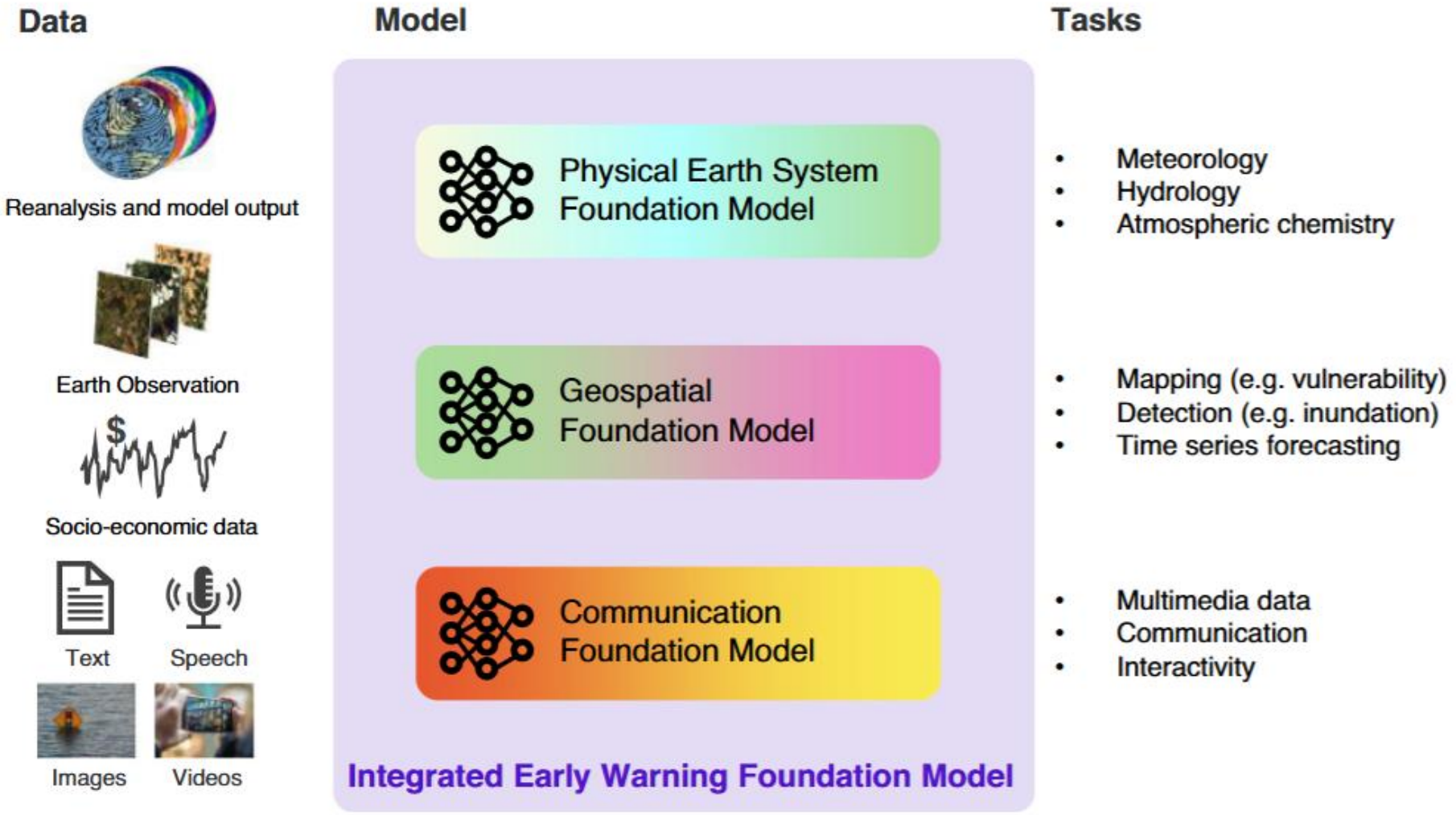



**Fig. 3 | Vision for an Integrated Early Warning Foundation Model.** Integration of currently developed foundation models into a modular Early Warning Foundation Model (center) allowing for ingestion of diverse data (left) and for addressing prediction and communication tasks (right).

Markus Reichstein, NC, 2024

KeAi CHINESE ROOTS GLOBAL IMPACT

# INTELLIGENT CLIMATE and ECO-ENVIRONMENT

Bridging AI, Climate Science and Environmental Health

**RENHE ZHANG**

Fudan University, China

• Committee co-chair

**DREW SHINDELL**

Duke University, United States

• Committee co-chair

**HONGLIANG ZHANG**

University of Shanghai for Science and Technology, China

• Editor-in-Chief


群聊：ICEE 作者群

- *First Issue Published in 2026*
- *Open Access Fee Waived*
- *Free editing service for invited articles*
- *Annual outstanding papers will be awarded certificates and bonuses*
- *Free promotion for published papers*

https://www.keaipublishing.com/en/journals/intelligent-climate-and-eco-environment/

该二维码7天内(9月4日前)有效，重新进入将更新

**2025**

# 谢谢！


**ACES Group**

School of Environment and Architecture
Shanghai University for Science and Technology
Email: zhanghl@usst.edu.cn

**ACES课题组**

上海理工大学
环境与建筑学院
zhanghl@usst.edu.cn