CMIP6 Multi-model Data Assimilation and Urban-rural Downscaling on Ambient Ozone

A Focus on Data-Driven Methodological Innovations

Professor Alexander Archibald (PI)
Dr Haitong Sun (Presenter)

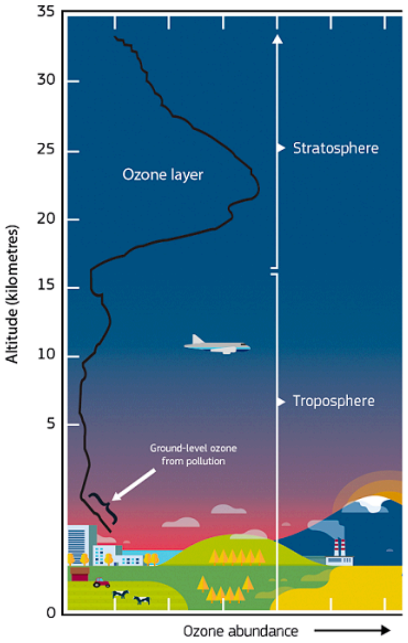
Centre for Atmospheric Science, Yusuf Hamied Department of Chemistry, University of Cambridge

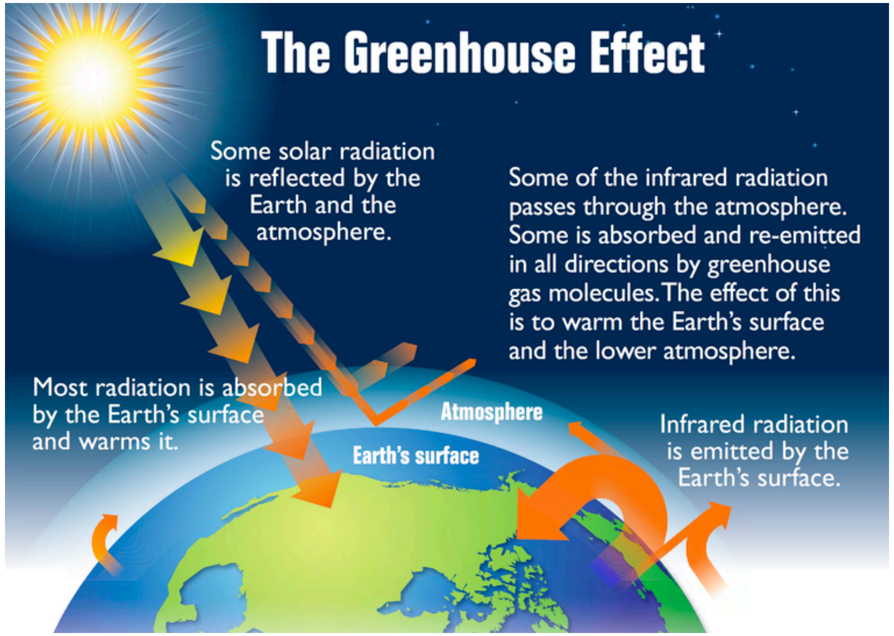
A Focus on O_3: Unique Properties

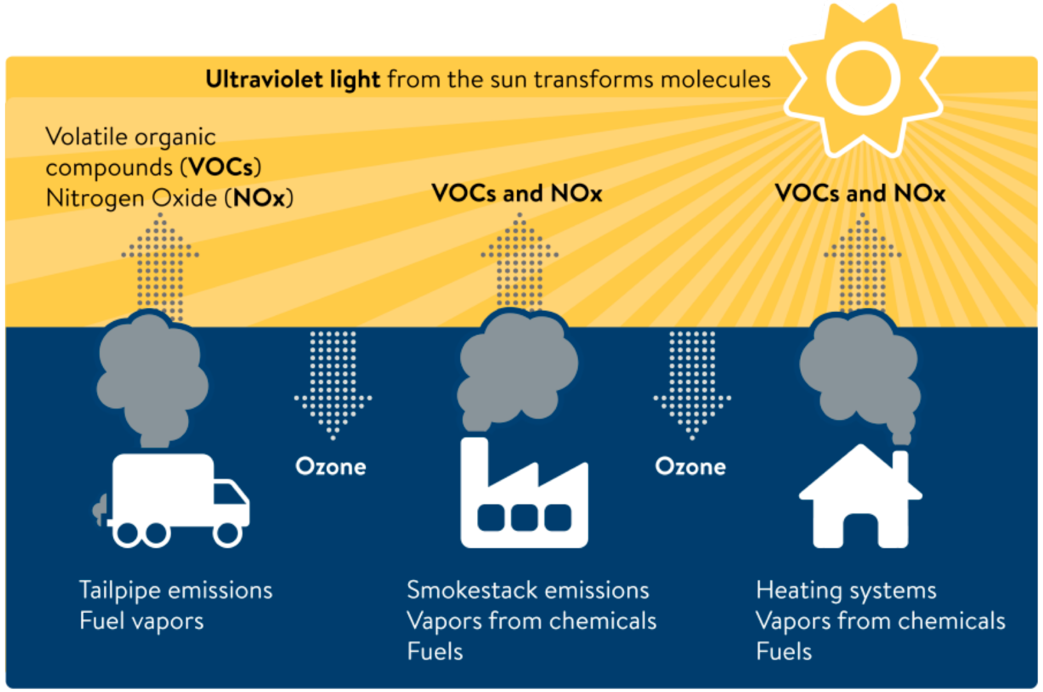
Ozone vertical profile: higher concentration in higher layers

Greenhouse effect: modifying the radiative forcing and atmospheric dynamics

O_3-NO_X-VOC system: non-linear response to precursors







Coupled Earth System Models: Chemistry Climate Models

Fig. 1 **What an Earth System Model (ESM) simulates.**

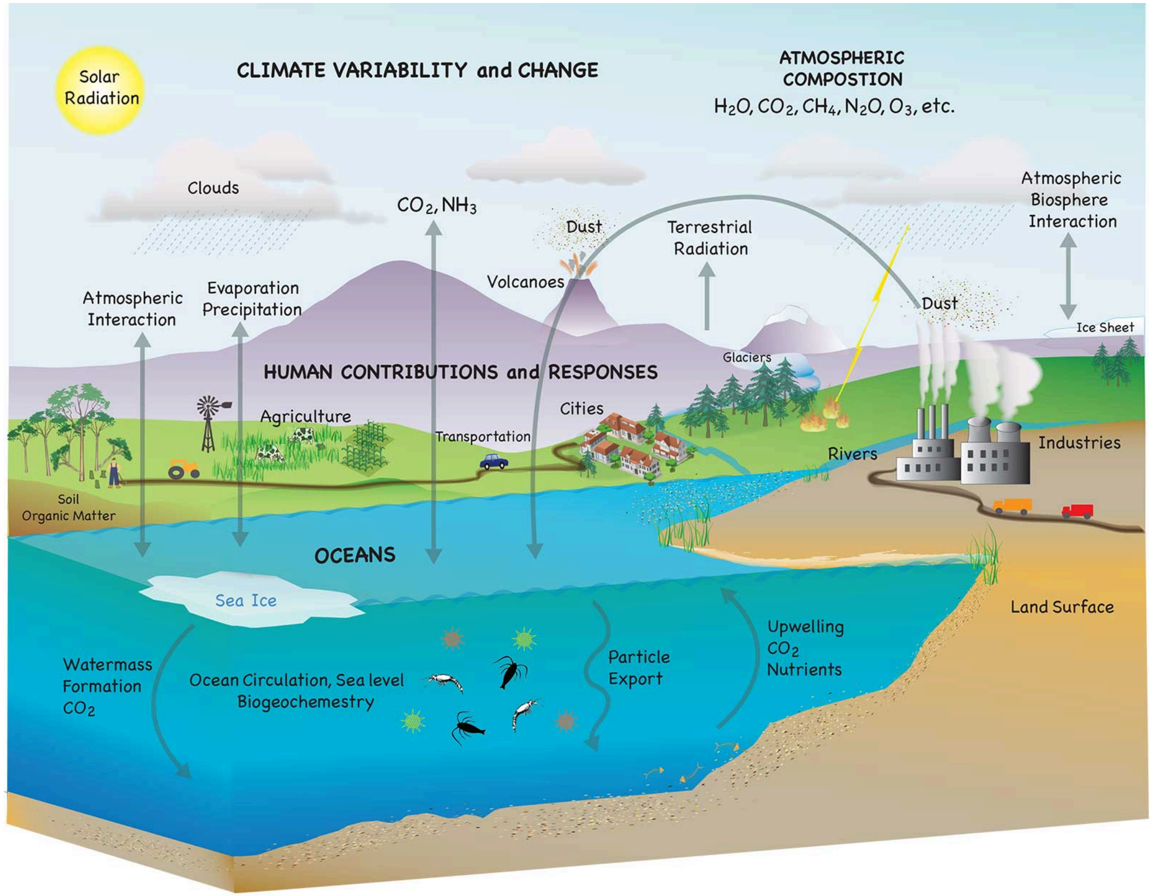
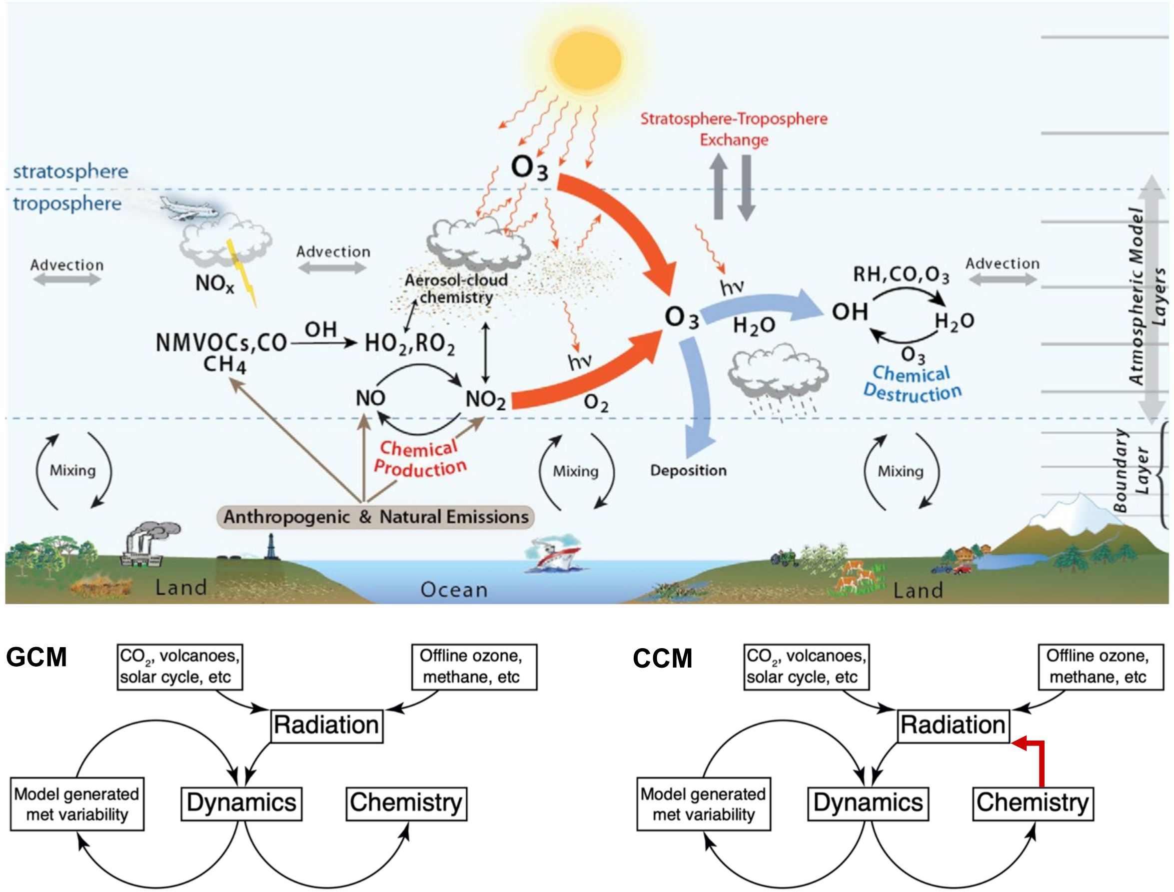


Fig. 2 **How a Chemistry Climate Model (CCM) understand the nature.**



Dunne, J. P., Horowitz, L. W., Adcroft, A. J., et al. The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics. *J Adv Model Earth Syst* **2020,** *12*, (11), e2019MS002015.
Young, P. J., Naik, V., Fiore, A. M., et al. Tropospheric Ozone Assessment Report: Assessment of global-scale model performance for global and regional ozone distributions, variability, and trends. *Elementa-Sci Anthrop* **2018,** *6*, (10), 1-49.

CMIP6 AerChemMIP ESMs

Fig. 3 Surface O_3 simulation by 8+1 CMIP6 models.

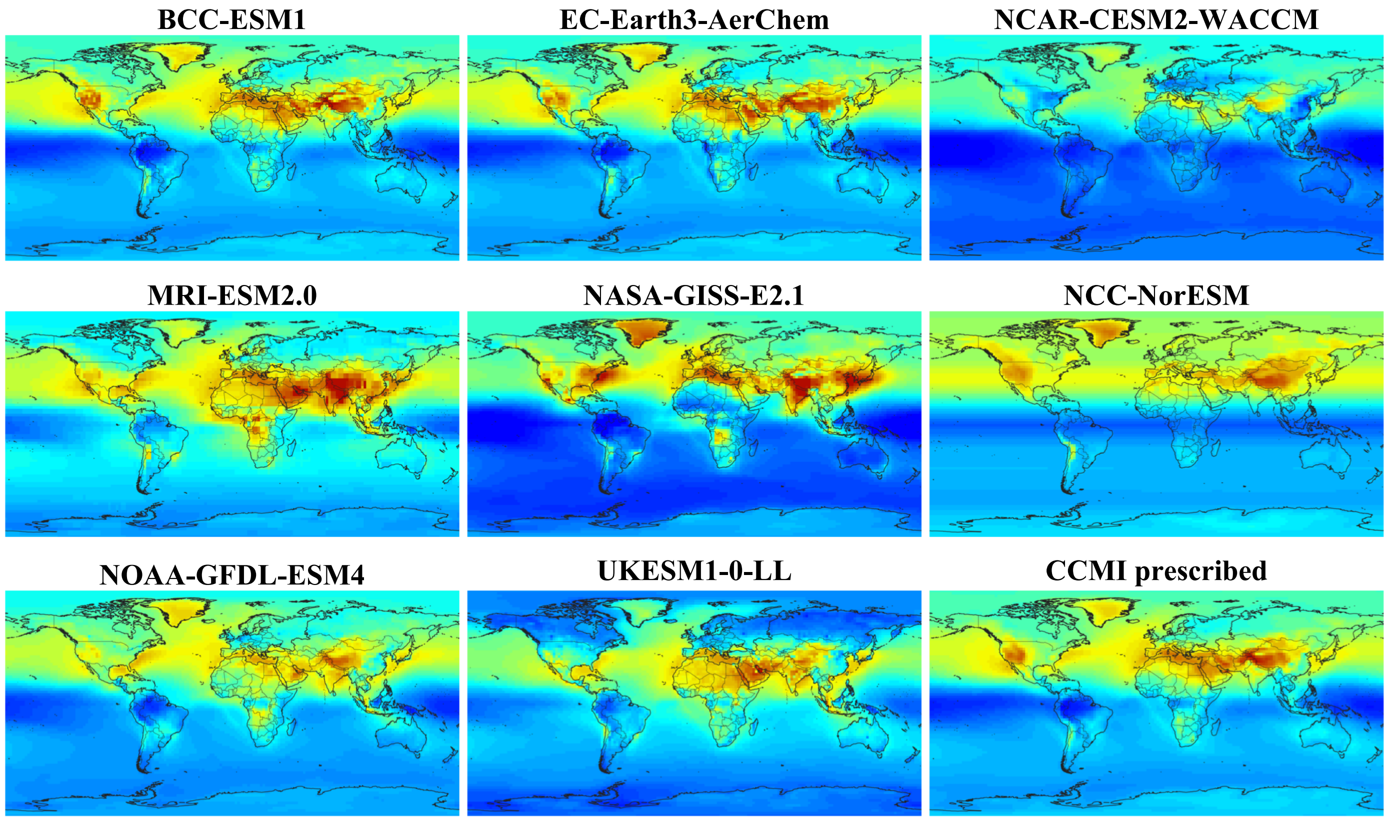
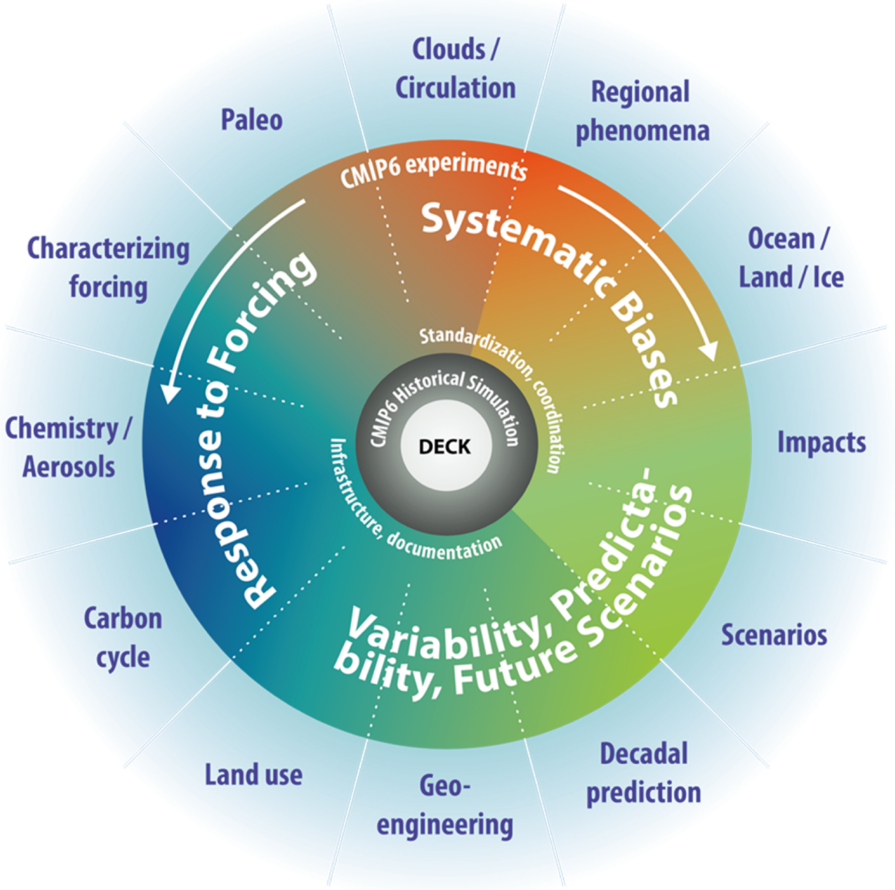


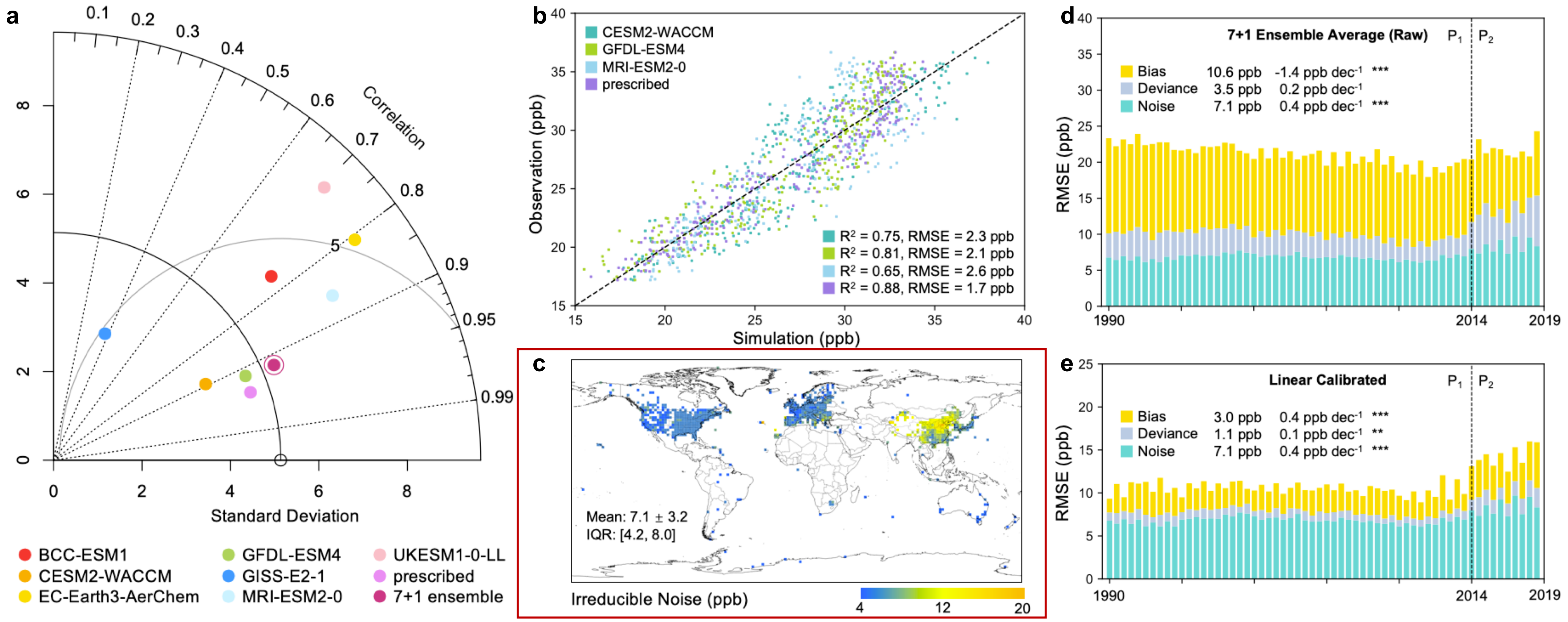
Fig. 4 CMIP6: The Coupled Model Intercomparison Project.



Sun, Z. and Archibald, A. T. Multi-stage ensemble-learning-based model fusion for surface ozone simulations: A focus on CMIP6 models. *Environmental Science and Ecotechnology* **2021,** *8*, 100124.
Eyring, V., Bony, S., Meehl, G. A., et al. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organisation. *Geosci Model Dev* **2016,** *9*, (5), 1937-1958.

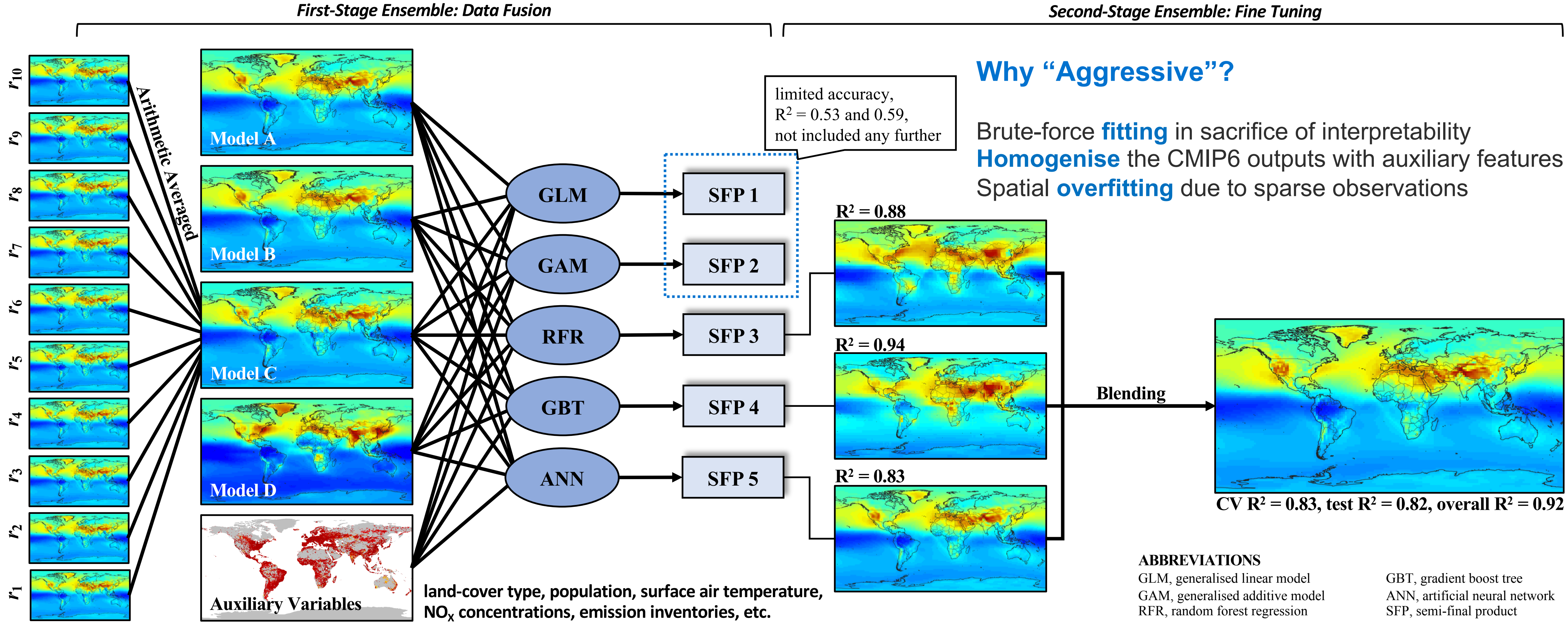
Observations and Model Performance Evaluation

Fig. 5 Observation-supervised model performance evaluation. a | Space-aggregated overall performance evaluations for individual models. b | Linear calibration potential assessment for model ensemble average. c | Irreducible square root noises by error decomposition. 30-year temporal tendencies of bias, deviation, and root noise by 7+1 model ensemble before (**d**) and after linear calibration (**e**) are assessed.



Sun, H. Z., Oliver, W., Staniaszek, Z., et al. Surface ozone simulation ensemble enhances model diagnosis assisted with machine learning. *Preprint*.

Classical Ensembled Learning: “Aggressive” Approach



Why “Aggressive”?

Brute-force **fitting** in sacrifice of interpretability
Homogenise the CMIP6 outputs with auxiliary features
Spatial **overfitting** due to sparse observations

ABBREVIATIONS

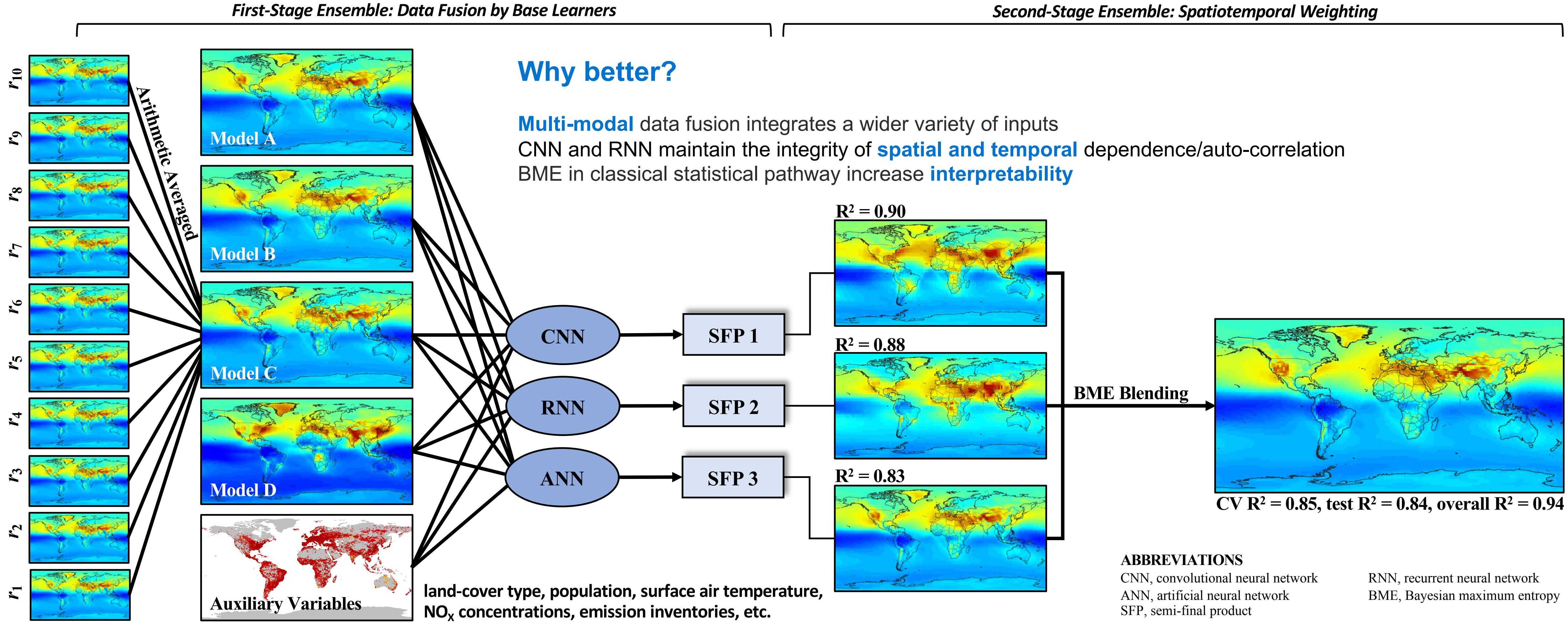
GLM, generalised linear model
GAM, generalised additive model
RFR, random forest regression
GBT, gradient boost tree
ANN, artificial neural network
SFP, semi-final product

Lyu, B., Hu, Y., Zhang, W., et al. Fusion Method Combining Ground-Level Observations with Chemical Transport Model Predictions Using an Ensemble Deep Learning Framework: Application in China to Estimate Spatiotemporally-Resolved $PM_{2.5}$ Exposure Fields in 2014-2017. *Environ Sci Technol* **2019,** *53*, (13), 7306-7315.

Optimisation: M^3Fusion with Bayesian Maximum Entropy

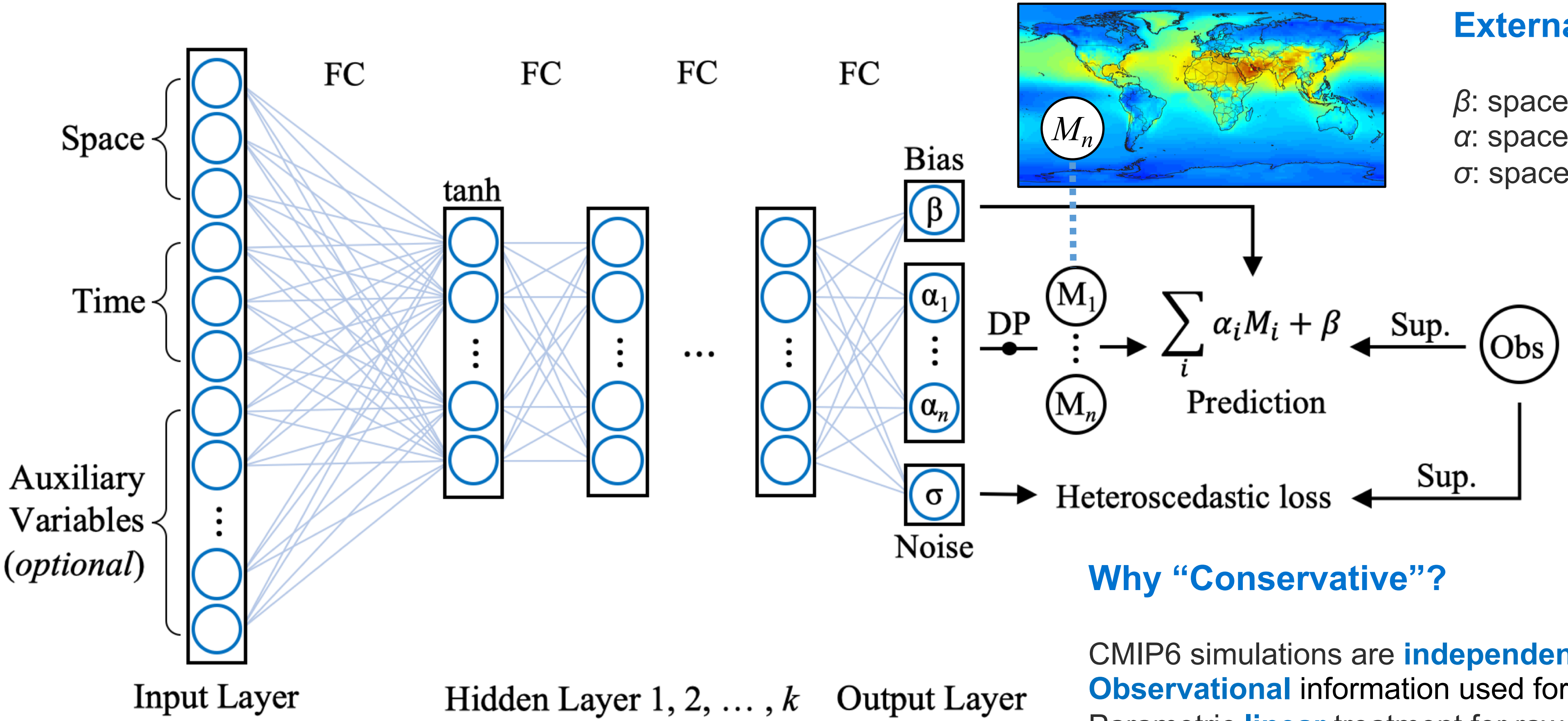
Why better?

Multi-modal data fusion integrates a wider variety of inputs
CNN and RNN maintain the integrity of **spatial and temporal** dependence/auto-correlation
BME in classical statistical pathway increase **interpretability**



Chang, K. L., Cooper, O. R., West, J. J., et al. A new method (M^3Fusion v1) for combining observations and multiple model output for an improved estimate of the global surface ozone distribution. *Geosci Model Dev* **2019**, *12*, (3), 955-978.

Space-time Bayesian Neural Network Ensembler: "Conservative" Approach



External Parameters

β: space-time variant bias corrector
α: space-time variant weight
σ: space-time variant stochastic noise

Why "Conservative"?

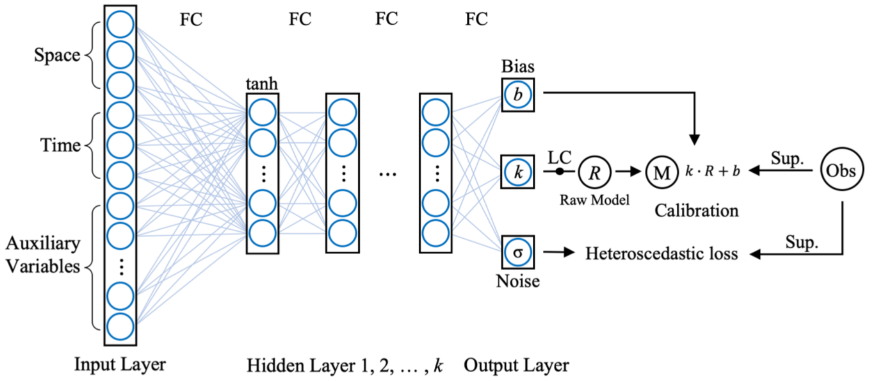
CMIP6 simulations are **independent** from neural network layers
Observational information used for global restriction
Parametric **linear** treatment for raw simulations in essence

Sun, Z. and Archibald, A. Multi-stage Ensemble-learning-based Model Fusion for Surface Ozone Simulations: A Focus on CMIP6 Models. *Environmental Science & Ecotechnology* **2021,** *8,* 100124.

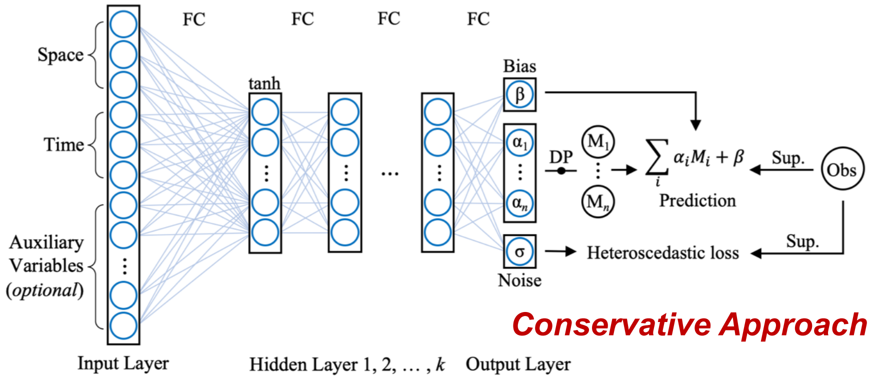
Space-time Bayesian Neural Network Ensembler: Weighting

Fig. 6 Evaluation and presentation of multi-model fused land surface O_3. a) Taylor diagram of performance evaluation for 7+1 model fusion by multiple approaches. b) Spatiotemporal averaged weights for individual models by 15 algorithms of fusion. Algorithms affected by hyperparameters are tuned to control overfitting. **c) 30-year averaged spatial weights for 7+1 individual models by space-time Bayesian neural network ensembler.** The overall spatiotemporal average weights for each model are inserted in each subplot.

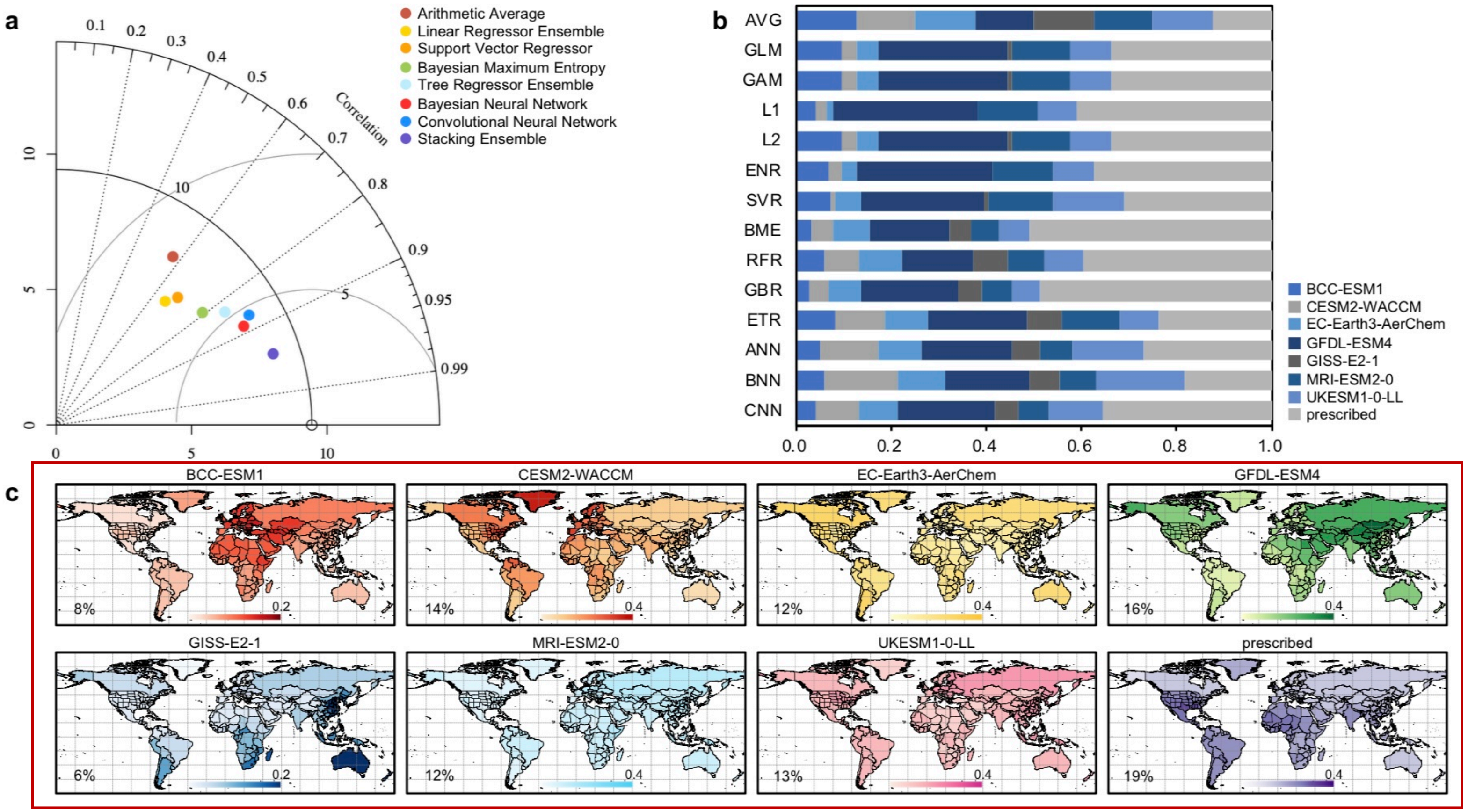
Step 1: Space-time-variant linear calibration



Step 2: Space-time-variant weighting



Conservative Approach

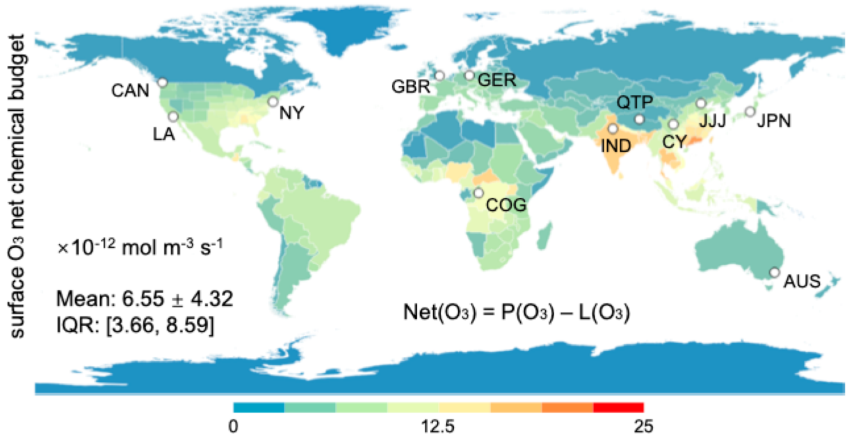


UNIVERSITY OF CAMBRIDGE

Sengupta, U. and Amos, M. et al. Ensembling geophysical models with Bayesian neural networks. *Advances in Neural Information Processing Systems*, **2020**, *33*, 1205-1217.
Sun, Z. and Archibald, A. T. Multi-stage ensemble-learning-based model fusion for surface ozone simulations: A focus on CMIP6 models. *Environmental Science and Ecotechnology* **2021**, *8*, 100124.
Sun, Z., Archibald, A. T., Wild, O. et al. Surface ozone simulation ensemble can strengthen model diagnosis. *arXiv* **2022**.

Earth System Model O_3-NO_X-VOC Diagnosis

12 Representative Regions



Precursor Sensitivities

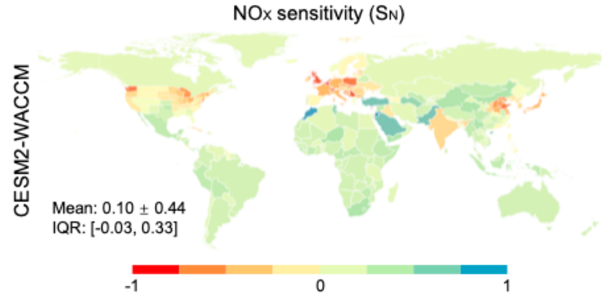
CESM2-WACCM

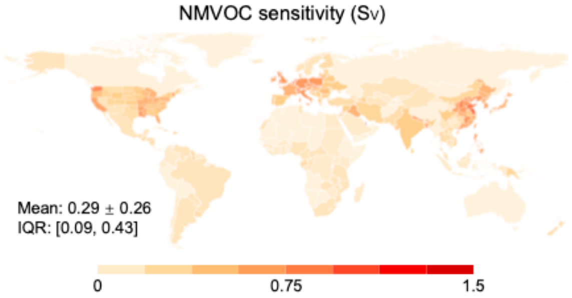
	LA	GBR	GER	JPN	JJJ	NY	IND	CY	CAN	AUS	QTP	COG	
$E(NO_x)$	-10.8	-10.5	-10.3	-10.4	-10.2	-10.7	-10.5	-10.9	-11.2	-11.6	-11.7	-11.6	*lg* kg m^{-2} s^{-1}
E(NMVOC)	-12.2	-12.0	-11.7	-11.6	-11.4	-11.7	-11.4	-11.4	-11.3	-10.9	-10.6	-10.8	*lg* kg m^{-2} s^{-1}
$Net(O_3)$	-11.0	-11.2	-11.2	-11.1	-11.0	-11.0	-10.8	-11.2	-11.5	-11.7	-12.1	-11.8	*lg* mol m^{-3} s^{-1}
S_N	-0.38	-0.48	-0.42	-0.43	-0.30	-0.28	-0.13	0.17	0.07	0.10	0.16	0.32	*lg* mol kg^{-1} m^{-1}
S_V	0.36	0.53	0.51	0.55	0.43	0.45	0.27	0.23	0.06	0.05	0.07	0.10	*lg* mol kg^{-1} m^{-1}
S_N/S_V	-1.06	-0.92	-0.82	-0.78	-0.69	-0.61	-0.46	0.74	1.07	1.90	2.42	3.20	unitless
regime	V_t	V_t	V_t	V_t	V_t	V_t	V_t	V	T	N	N	N	

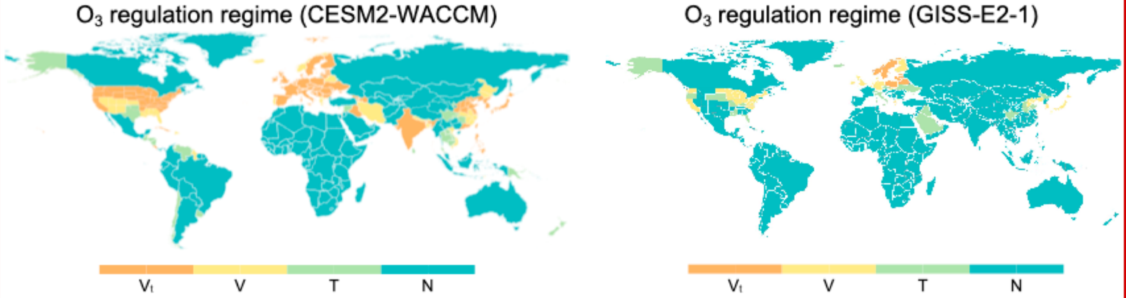
GISS-E2-1

	LA	GBR	GER	JPN	JJJ	NY	IND	CY	CAN	AUS	QTP	COG	
S_N	0.35	0.13	0.58	0.52	0.37	0.32	0.89	0.81	0.43	0.89	0.45	0.81	*lg* mol kg^{-1} m^{-1}
S_V	0.68	0.96	0.72	0.76	0.52	0.64	0.68	0.79	0.19	0.25	0.22	0.32	*lg* mol kg^{-1} m^{-1}
S_N/S_V	0.51	0.14	0.81	0.69	0.71	0.51	1.31	1.03	2.23	3.58	2.04	2.51	unitless
regime	V	V	T	V	V	V	N	T	N	N	N	N	

Geographical Distribution







Sun, H. Z., Oliver, W., Staniaszek, Z., et al. Surface ozone simulation ensemble enhances model diagnosis assisted with machine learning. *Preprint*.

Rural-Urban Disparity of Observed Ambient O_3

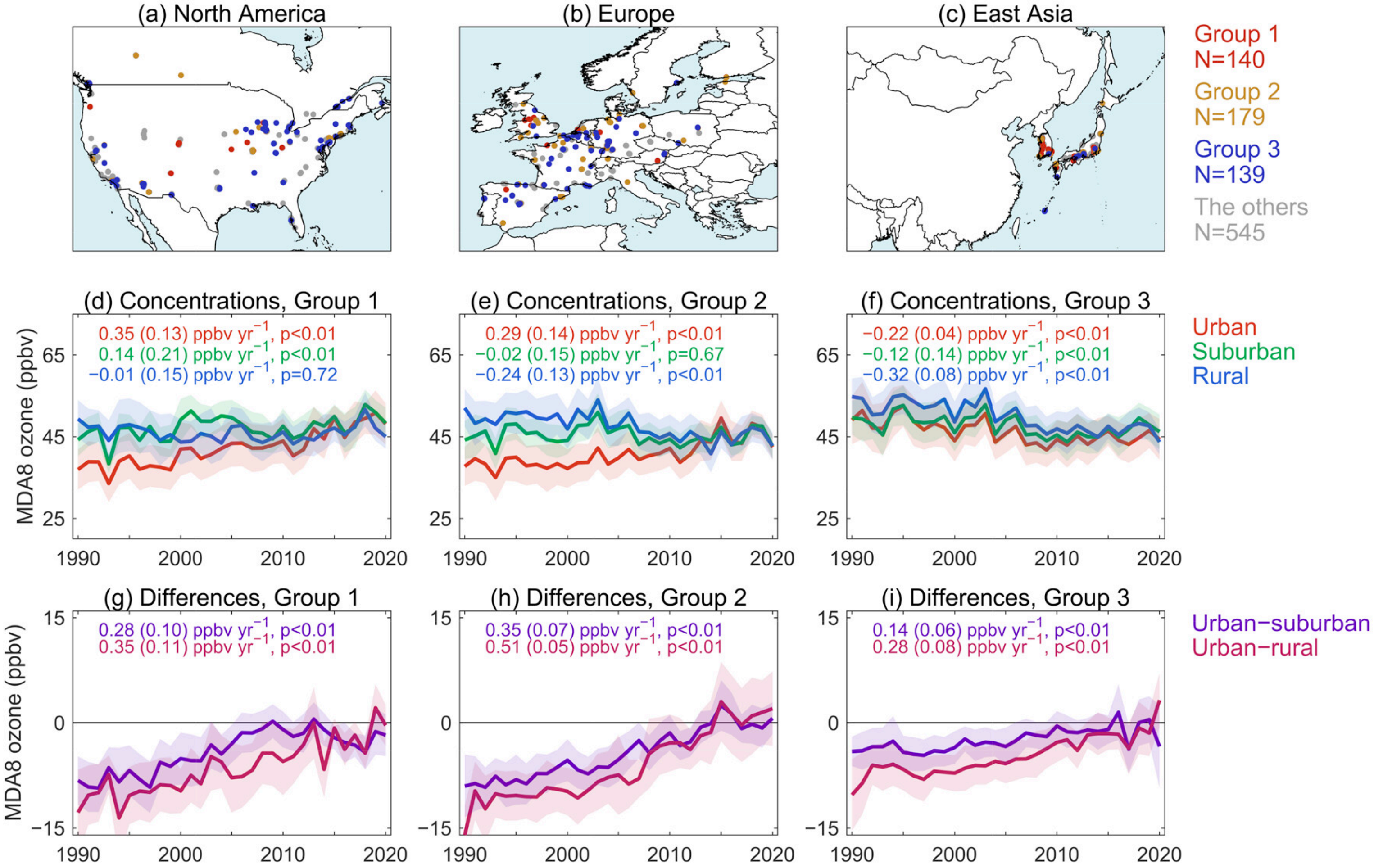
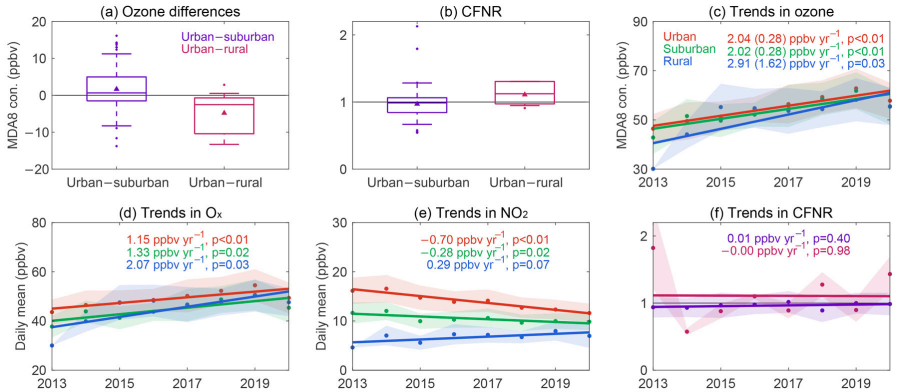


Fig. 7 Urban vs nonurban MDA8 ozone in the three groups of site pairs over 1990–2020.

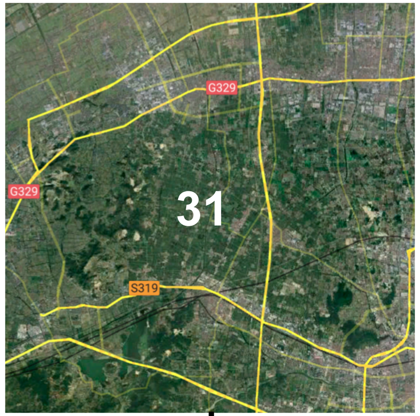
(a–c) Spatial distribution of the paired ozone observation sites in each group. (d–f) Time series of MDA8 ozone concentrations over urban, suburban, and rural areas in the three groups in North America, Europe, South Korea, and Japan. (g–i) The same as (d–f), but for urban vs suburban and urban vs rural MDA8 ozone differences. 'N' shows the number of urban sites in each group. The shading areas in (d–i) indicate $\pm 50\%$ the standard deviation for the site pairs in each group. The numbers in (d–i) show the trends with the contributions of changes in climate enclosed in the brackets.



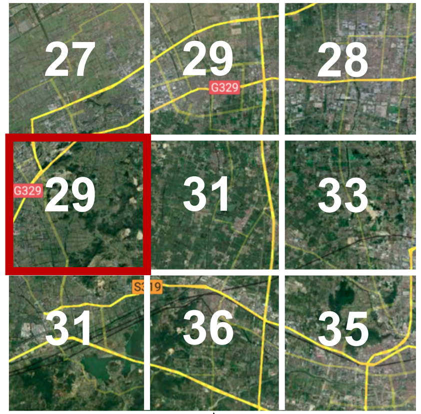
Han, H., Zhang, L., Liu, Z., et al. Narrowing Differences in Urban and Nonurban Surface Ozone in the Northern Hemisphere Over 1990–2020. *Environ Sci Tech Lett* **2023,** *10,* (5), 410-417.

Urban-Rural Downscaling Schematic Concept

Raw coarse grid



Downscaled fine grid



Theoretical Minimum Cell

Conventional Downscaler

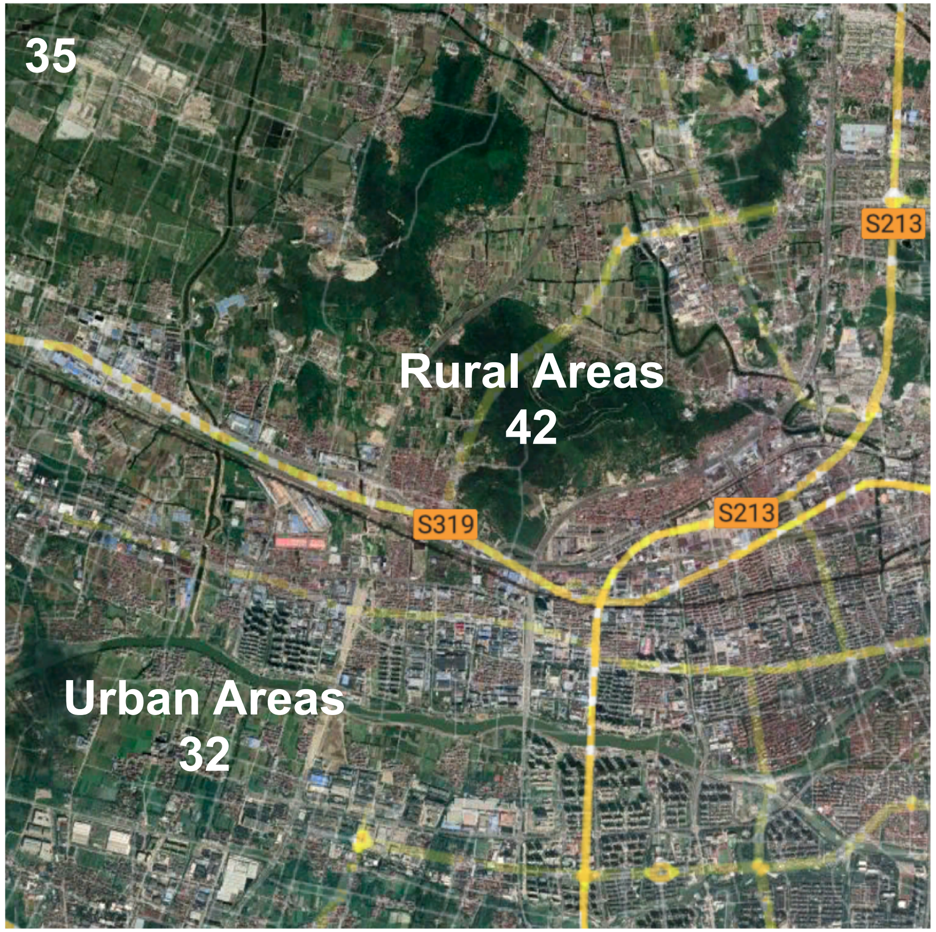
Stacked Downscaler

42 *A fictitious rural cell*

32 *A fictitious urban cell*

$$C_{s,t}^{\text{urban}} = C_{s,t}^{\text{fusion}} \cdot k_{s,t}^{\text{urban}} + \sigma_{s,t}^{\text{urban}}$$

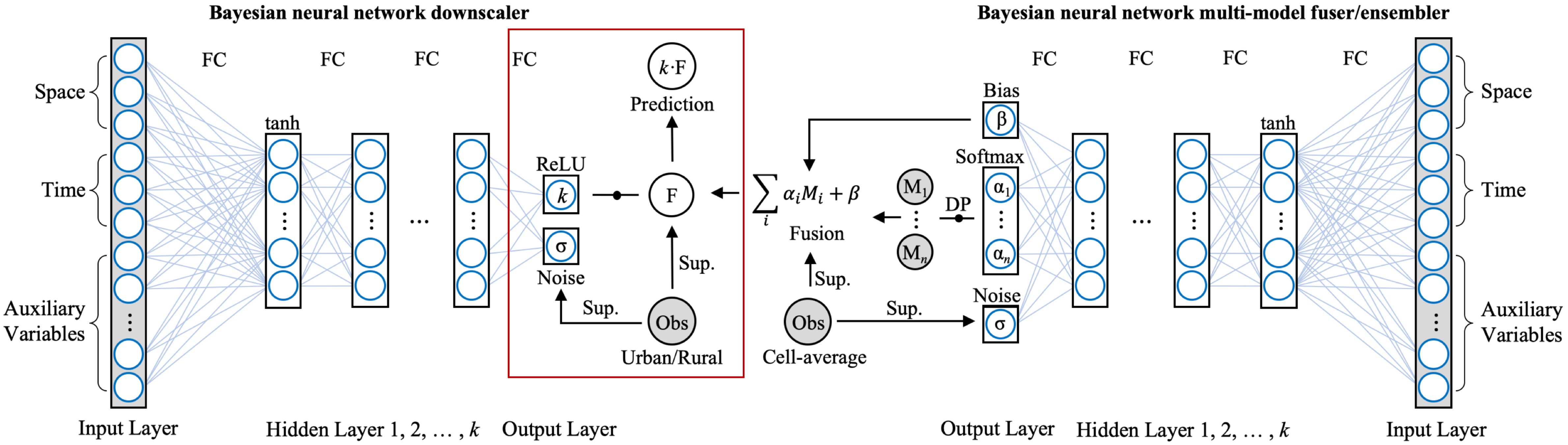
$$C_{s,t}^{\text{rural}} = C_{s,t}^{\text{fusion}} \cdot k_{s,t}^{\text{rural}} + \sigma_{s,t}^{\text{rural}}$$



Sun, H., Shin, Y. M., Archibald, A. T., et al. Spatial Resolved Surface Ozone with Urban and Rural Differentiation during 1990-2019: A Space-Time Bayesian Neural Network Downscaler. *Environ Sci Technol* **2021**.

High-resolution Downscaling Model: Urban-Rural Differentiation

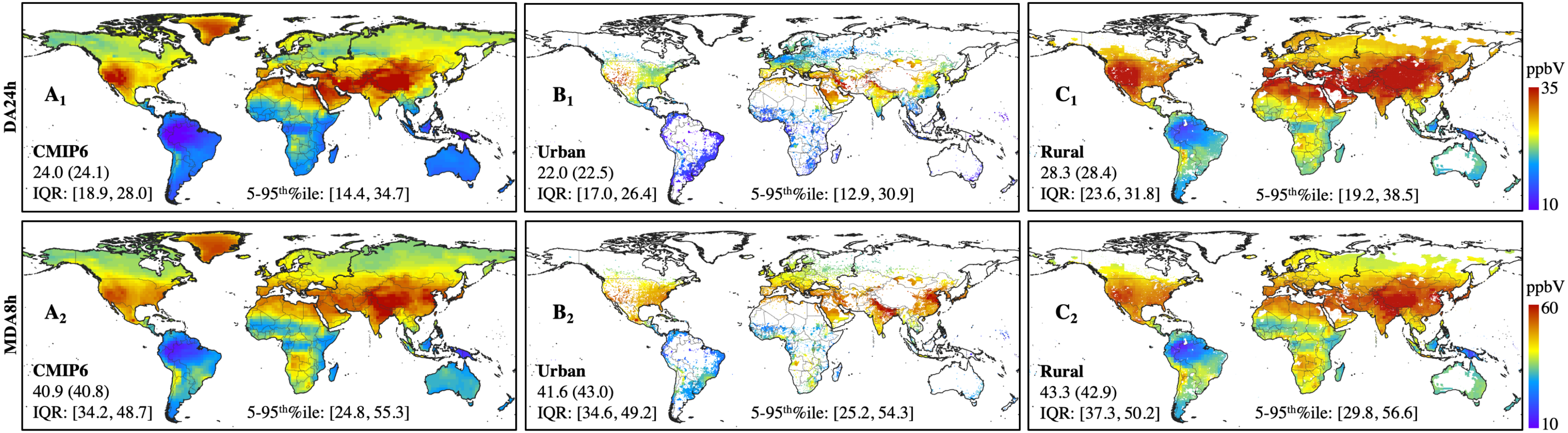
Fig. 8 Schematic diagram of space-time Bayesian neural network multi-model fuser and downscaler. The shaded elements refer to the external datasets not affected by neural network; the rectangle circumscribed elements indicate the input, processing and output variates inside the neural network; and non-rectangle circumscribed elements represent the final products.



Sun, H., Shin, Y. M., Archibald, A. T., et al. Spatial Resolved Surface Ozone with Urban and Rural Differentiation during 1990-2019: A Space-Time Bayesian Neural Network Downscaler. *Environ Sci Technol* **2021**.

Ambient O_3 Downscaled Products

Fig. 9 30-year average urban and rural surface ozone concentrations by space-time Bayesian neural network prediction during 1990-2019. Both DA24h and MDA8h metrics are predicted with spatial resolution as 0.125-degree (10 *km*) and temporal interval as per month. Global range statistics of the 30-year averaged surface ozone concentrations are inserted in each panel as arithmetic mean and median (in brackets), inter-quartile range (IQR), and 5^{th} to 95^{th} percentile ($5\text{-}95^{th}$ %ile).



Multiple database inter-comparison programme led by **Professor Jason West (UNC Chapel Hill)**, see more information at:
https://igacproject.org/human-health-impacts-ozone-focus-working-group

Sun, H., Shin, Y. M., Archibald, A. T., et al. Spatial Resolved Surface Ozone with Urban and Rural Differentiation during 1990-2019: A Space-Time Bayesian Neural Network Downscaler. *Environ Sci Technol* **2021**.

Ambient O_3 Exposure Associated Global Mortality Burden

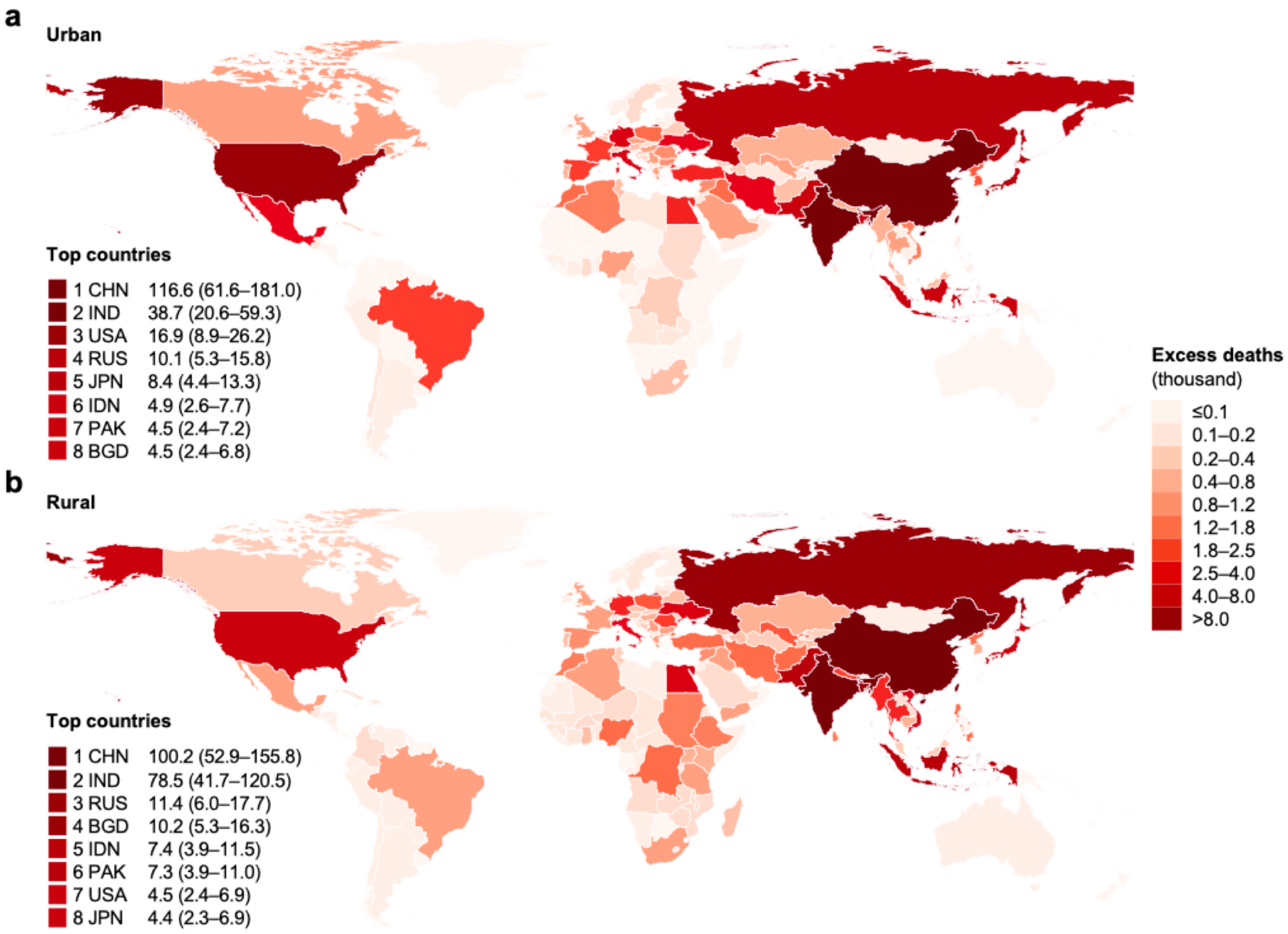


Fig. 10 Global mapping of ambient O_3 exposure associated excess cardiovascular mortalities in 2019 by urban and rural populations.

Estimated cases of O_3 exposure-associated cardiovascular premature deaths differentiating **(a)** urban and **(b)** rural populations are presented for 204 global countries and/or regions for geographical mapping. Exposure-response relationship for cardiovascular mortality is estimated by random-effects meta-analysis pooling 14 unique-cohort-based epidemiological studies identified from the most up-to-date systematic review.

UNIVERSITY OF CAMBRIDGE

Sun, H.Z.*, Daalen, K.R., Guillas, S., Giorio, C., Di, Q., Kan, H., Guo, Y., Archibald. A.T. Global cardiovascular mortality associated with ambient ozone exposure and rural-urban disparities. Under 3rd round revision.

Regional O_3 Isopleths of Chinese Representative Cities

Fig. 11 O_3 isopleths for six representative cities. The black cross marks the $O_{3,\ pop}$ under 100% of NO_X and VOC emissions, i.e., the current positions (2017) on the isopleths. The black dashed line is where the sensitivity to NO_X emission changes is zero (above which, increased NO_X emissions decrease O_3), and the solid black line is where the VOC and NO_X emission sensitivities are the same.

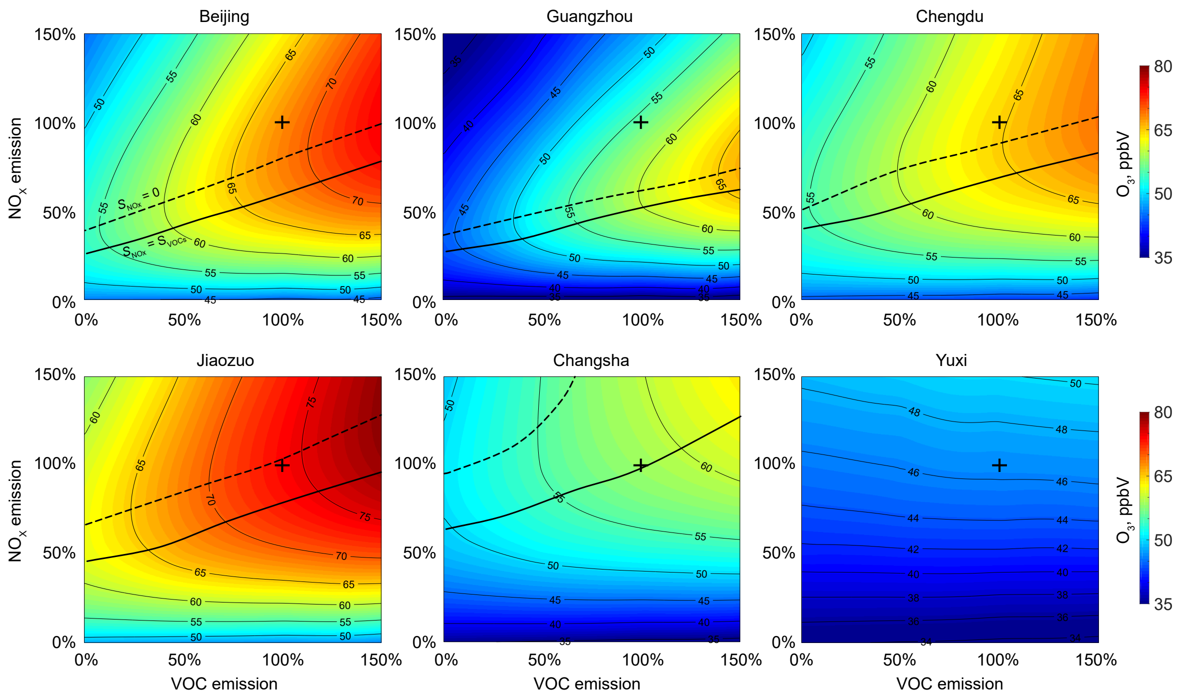
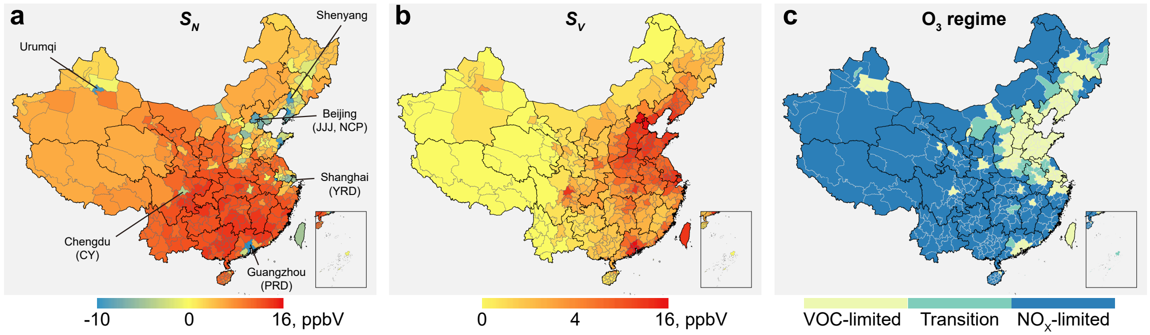
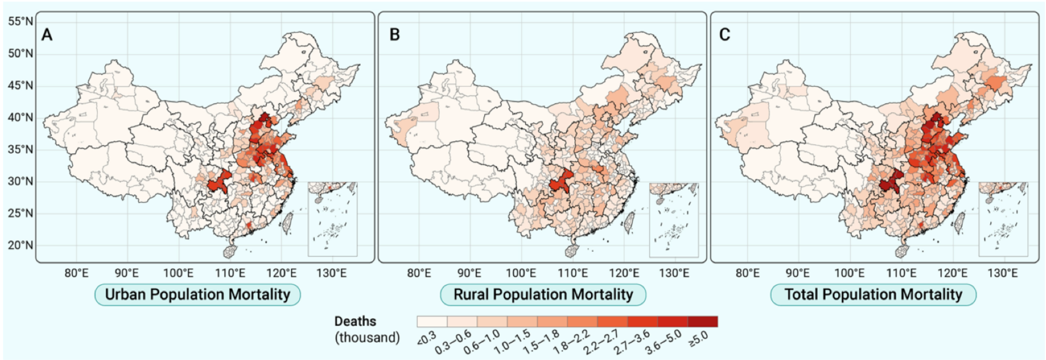


Fig. 12 The spatial patterns of the city-level S_N (a), S_V (b), and O_3 regimes (c). S_N and S_V are the modeled sensitivities of O_3 concentrations to anthropogenic NO_X and VOC emissions, respectively, and are represented by the O_3 concentration change due to a 100% change in NO_X and VOC emissions, respectively. The O_3 regime for each city is determined by the ratio of S_N to S_V (NO_X-limited if $S_N/S_V > 1.2$, transitioning if $0.8 < S_N/S_V \leq 1.2$, VOC-limited otherwise). Several representative cities are marked with the regions where the cities are located: Jing-Jin-Ji (JJJ), Yangtze River Delta (YRD), Pearl River Delta (PRD), and Cheng-Yu (CY).





Relevant Publications

Sun, Z.* and **Archibald, A. T.*** Multi-stage ensemble-learning-based model fusion for surface ozone simulations: A focus on CMIP6 models. *Environ Sci Ecotechnol* **2021**, *8*, 100124.

Sun, H.*, Shin, Y. M., Xia, M., Ke, S., Wan, M., Yuan, L., Guo, Y., Archibald, A. T.* Spatial Resolved Surface Ozone with Urban and Rural Differentiation during 1990-2019: A Space-Time Bayesian Neural Network Downscaler. *Environ Sci Technol* **2022,** *56*, (11), 7337-7349.

Shen, H.⊥*, Sun, Z.⊥, Chen, Y., Russell, A. G., Hu, Y., Odman, M. T., Qian, Y., Archibald, A. T., Tao, S. Novel Method for Ozone Isopleth Construction and Diagnosis for the Ozone Control Strategy of Chinese Cities. *Environ Sci Technol* **2021,** *55*, (23), 15625-15636.

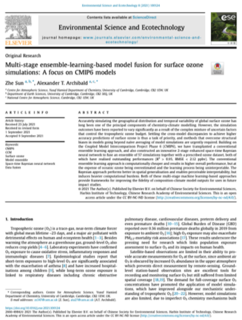
Sun, H. Z., Yu, P., Lan, C., Wan, M. W. L., Hickman, S., Murulitharan, J., Shen, H., Yuan, L., Guo, Y.*, Archibald, A. T.* Cohort-based long-term ozone exposure-associated mortality risks with adjusted metrics: A systematic review and meta-analysis. *The Innovation* **2022,** *3*, (3), 100246.

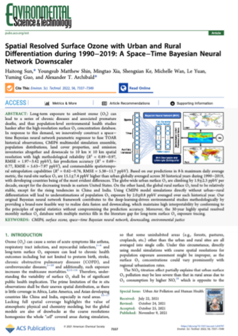
Sun, H. Z., Zhao, J., Liu, X., Qiu, M., Shen, H., Guillas, S., Giorio, C., Staniaszek, Z., Yu, P., Wan, M. W. L., Chim, M. M., Daalen, K. R., Li, Y., Liu, Z., Xia, M., Ke, S., Zhao, H., Liu, H.*, Guo, Y.*, Archibald, A. T.* Antagonism between ambient ozone increasing and urbanization-oriented population migration on Chinese cardiopulmonary mortality. *The Innovation* **2023**, Accepted.

Haitong Z. Sun, PhD

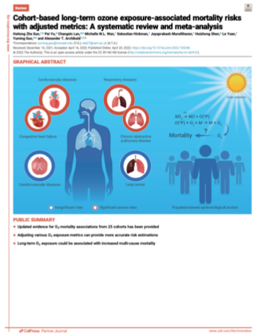
Research Fellow, National Centre for Atmospheric Science, UK
Guest Assistant Professor, Zhejiang University, China
Incoming Assistant Professor in 2024

Email: csuen27@gmail.com
Twitter: csuen27


Multi-stage ensemble-learning-based model fusion for surface ozone simulations: A focus on CMIP6 models


Spatial Resolved Surface Ozone with Urban and Rural Differentiation during 1990−2019: A Space−Time Bayesian Neural Network Downscaler


Novel Method for Ozone Isopleth Construction and Diagnosis for the Ozone Control Strategy of Chinese Cities


Cohort-based long-term ozone exposure-associated mortality risks with adjusted metrics: A systematic review and meta-analysis

Coming soon

CMIP6 Multi-model Data Assimilation and Urban-rural Downscaling on Ambient Ozone

A Focus on Data-Driven Methodological Innovations

Professor Alexander Archibald (PI): ata27@cam.ac.uk
Dr Haitong Sun (Presenter): zs347@cam.ac.uk

Centre for Atmospheric Science, Yusuf Hamied Department of Chemistry, University of Cambridge