*Supplementary information*

# Tropospheric ozone trends and attributions over East and Southeast Asia in 1995-2019: An integrated assessment using statistical methods, machine learning models, and multiple chemical transport models

**Xiao Lu et al.**
**Correspondence to:** Xiao Lu (luxiao25@mail.sysu.edu.cn)

**Table S1.** Summary of meteorological variables being used in the multiple linear regression (MLR), the ridge regression (RR), and the random forest regression (RFR) models

| Names and units of variables | ERA5 variables | MERRA2 variables | Average period |
|---|---|---|---|
| temperature at 2 m (K) | t2m | T2M | 06:00 to 18:00 |
| Solar Radiation | ssrd | SWGDN | 06:00 to 18:00 |
| mean sea level pressure (Pa) | msl | SLP | 06:00 to 18:00 |
| relative humidity (%) | R | RH | 06:00 to 18:00 |
| boundary layer height (m) | blh | PBLH | 06:00 to 18:00 |
| zonal wind at 10m (m $s^{-1}$) | u10 | U10M | 06:00 to 18:00 |
| meridional wind at 10m (m $s^{-1}$) | v10 | V10M | 06:00 to 18:00 |
| total precipitation (mm) | tp | PRECTOT | 06:00 to 18:00[a] |
| zonal wind at 850hPa (m $s^{-1}$) | u | U | 06:00 to 18:00 |
| meridional wind at 850hPa (m $s^{-1}$) | v | V | 06:00 to 18:00 |
| vertical velocity at 850hPa (Pa $s^{-1}$) | w | OMEGA | 06:00 to 18:00 |

[a] Total amount in the period

**Table S2**. Configuration of physical parameterization schemes for the WRF model

| **Term** | **Option** |
|---|---|
| Microphysics scheme | Purdue Lin scheme |
| Longwave radiation scheme | RRTM scheme |
| Shortwave radiation scheme | Goddard shortwave |
| Surface-layer scheme | Revised MM5 Monin-Obukhov scheme |
| Land-surface scheme | Unified Noah Land Surface Model |
| Boundary-layer scheme | Yonsei University scheme |
| Cumulus scheme | Kain-Fritsch scheme |

**Table S3.** Mean and standard deviations in $R^2$ in the four city-clusters by different models.

| | | MLR | RR | RFR |
|---|---|---|---|---|
| BTH | ERA5 | 0.41 ± 0.07 | 0.39 ± 0.07 | 0.43 ± 0.08 |
| | MERRA-2 | 0.37 ± 0.05 | 0.35 ± 0.05 | 0.38 ± 0.07 |
| YRD | ERA5 | 0.45 ± 0.08 | 0.43 ± 0.08 | 0.55 ± 0.09 |
| | MERRA-2 | 0.38 ± 0.08 | 0.36 ± 0.08 | 0.47 ± 0.09 |
| PRD | ERA5 | 0.56 ± 0.04 | 0.54 ± 0.04 | 0.64 ± 0.05 |
| | MERRA-2 | 0.53 ± 0.06 | 0.50 ± 0.06 | 0.60 ± 0.06 |
| SCB | ERA5 | 0.29 ± 0.17 | 0.26 ± 0.18 | 0.25 ± 0.16 |
| | MERRA-2 | 0.46 ± 0.04 | 0.44 ± 0.04 | 0.47 ± 0.05 |

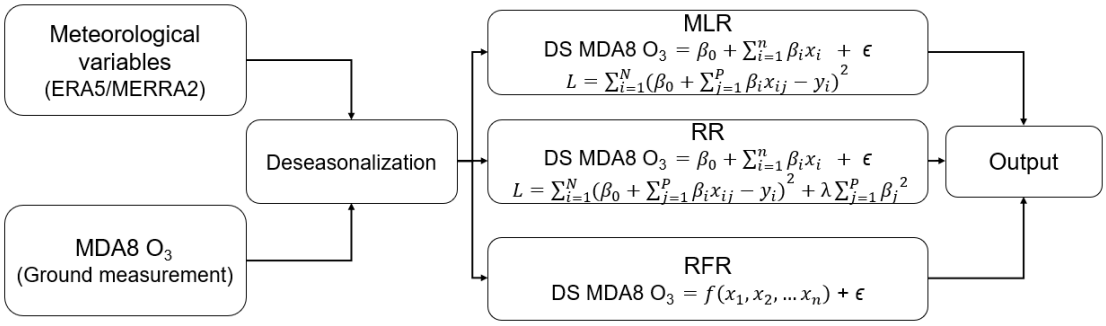


**Figure S1.** Overview of the statistical and machine learning models used in this study to attribute surface ozone trends.

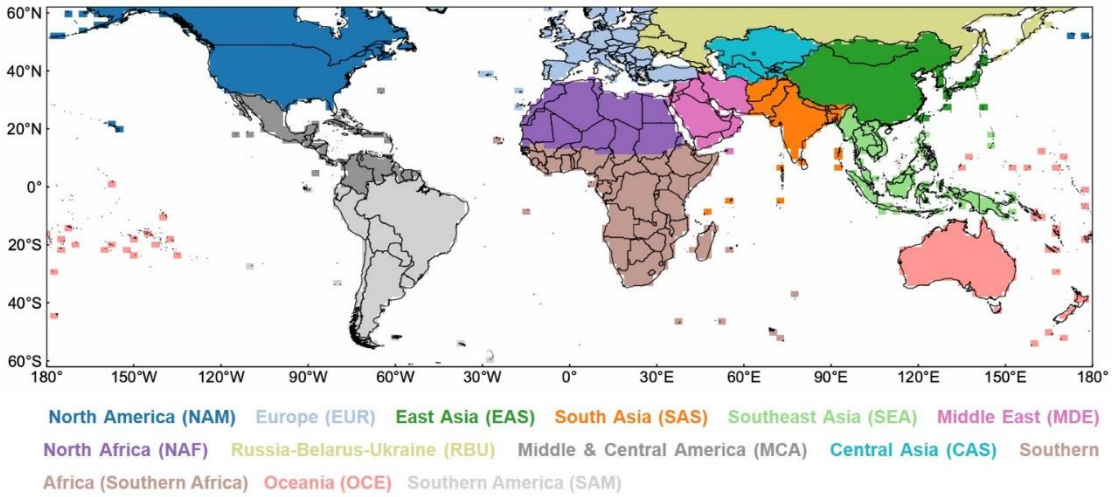


**Figure S2.** $NO_x$-tagged domain of the CAM4-chem simulation.

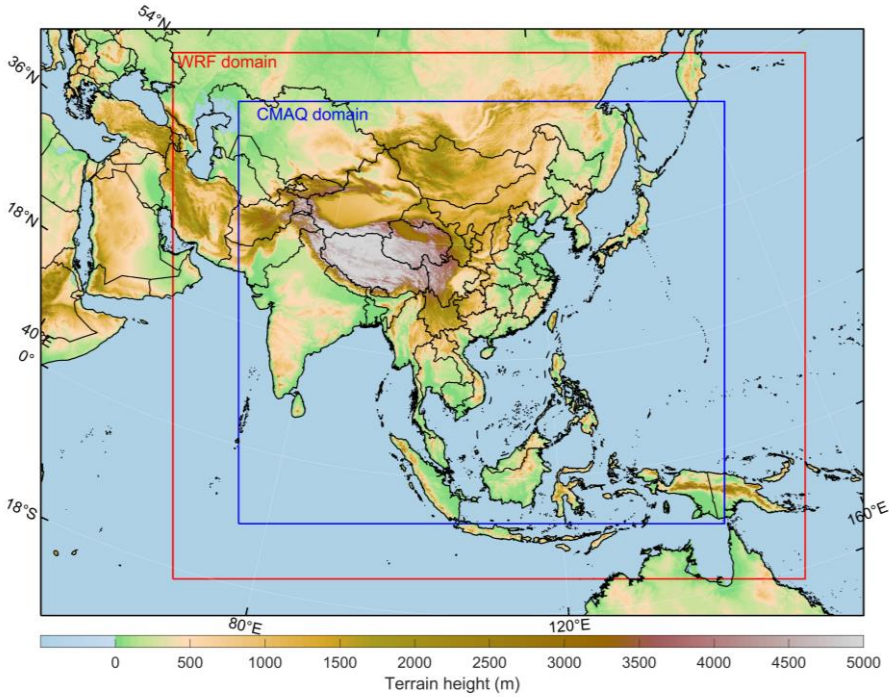


**Figure S3.** Domain of the WRF-CMAQ simulation.

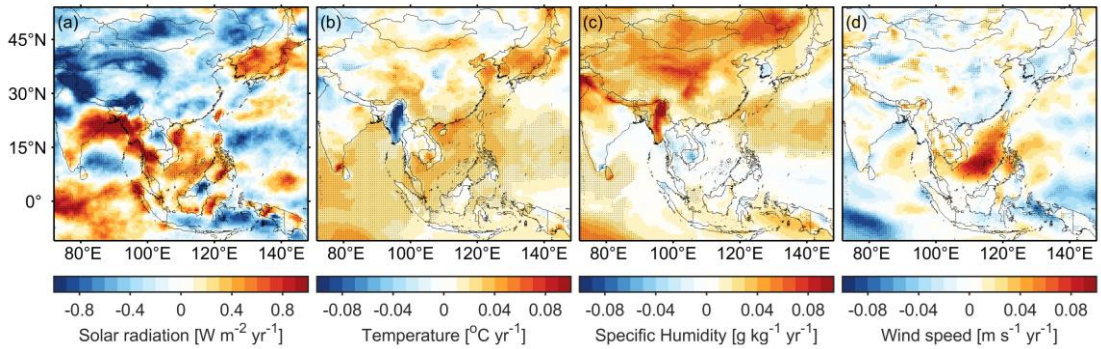


**Figure S4.** Spatial distributions of trends in meteorological variables over East and Southeast Asia, 1995-2019. Panels (a) to (d) displays trends in surface downward solar radiation, temperature at 2m height, specific humidity at 2m height and wind speed at 10m height, respectively. Meteorological parameters are derived from the MERRA2 re-analysis dataset. Black dots denoted linear trends with a p-value < 0.05.

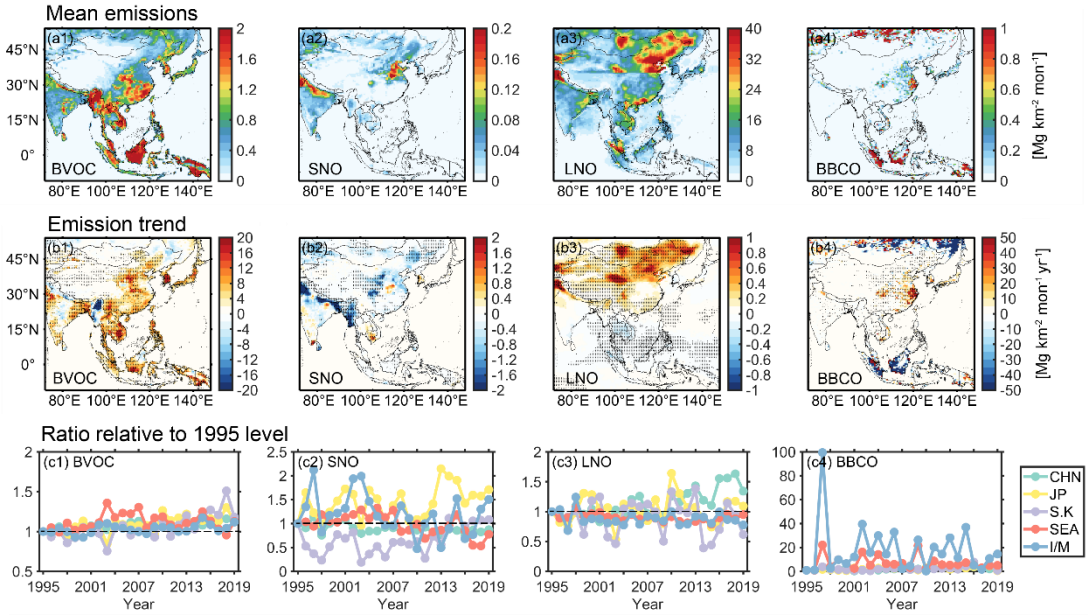


**Figure S5.** Same as Figure 2, but for BVOC, soil $NO_x$, and lightning NO emissions. Emissions are estimated from parameterization schemes implemented in the GEOS-Chem model, except for biomass burning emissions which are derived from the BB4CMIP6 and GFED inventory.

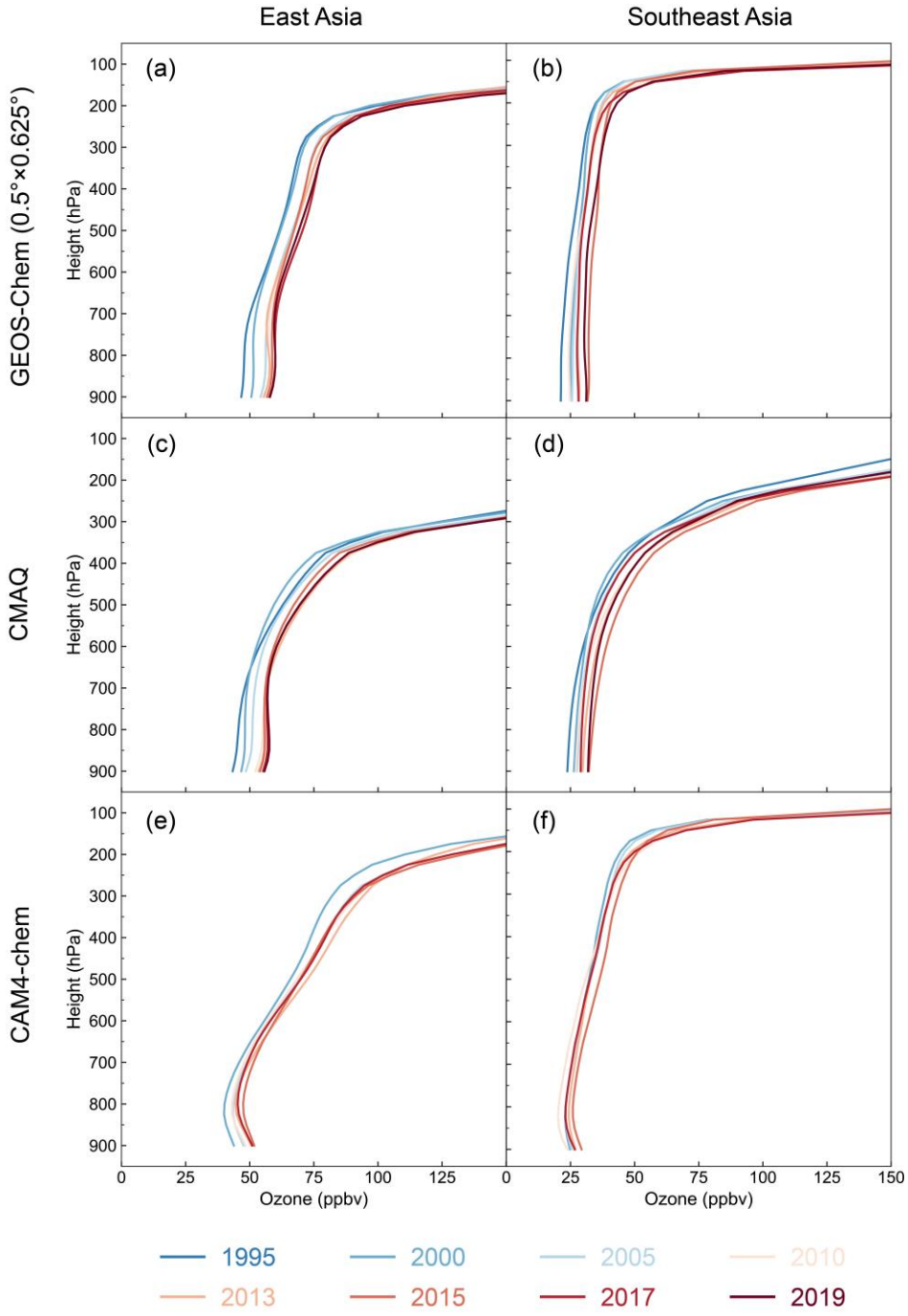


**Figure S6.** Comparison of GEOS-Chem, CMAQ, and CESM vertical ozone profiles averaged over East Asia and Southeast Asia in June, July, August in 1995-2019.

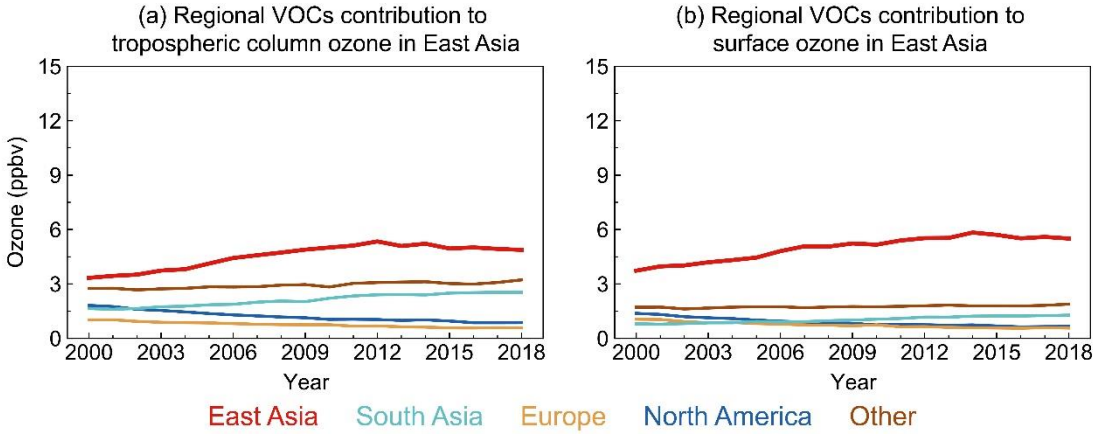


**Figure S7.** Same as Figure 10 but for RC-tagged results. Results for Southeast Asia is not available, because Southeast Asia is categorized as "Other" in RC-tagged simulation.

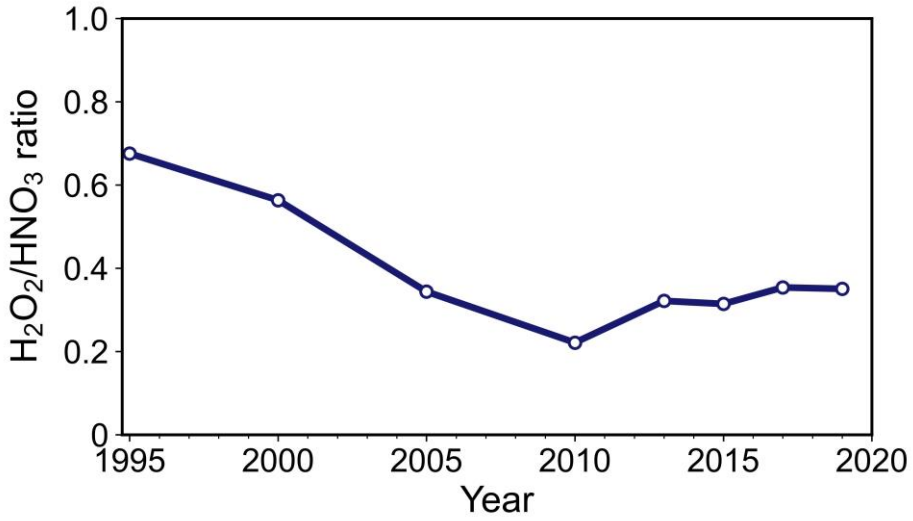


**Figure S8.** $H_2O_2/HNO_3$ time series in the North China Plain city cluster, China.

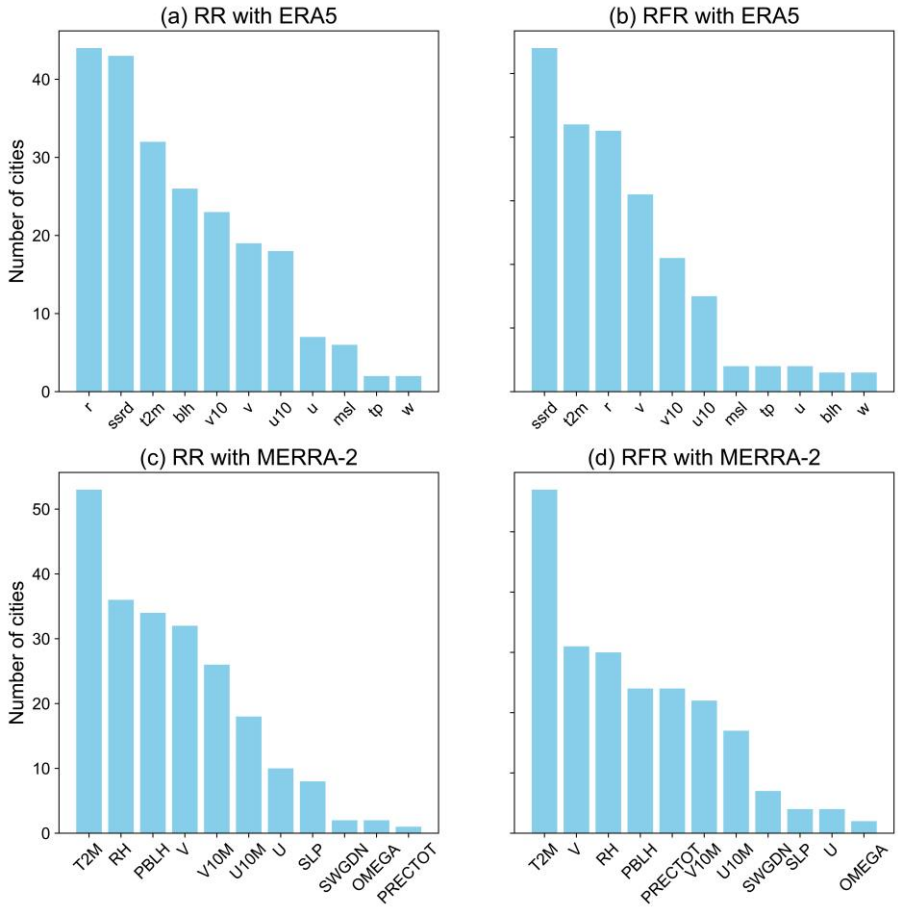


**Figure S9.** Number of cities with the meteorological parameters included as top-three most powerful predictors in the RR or RFR models.

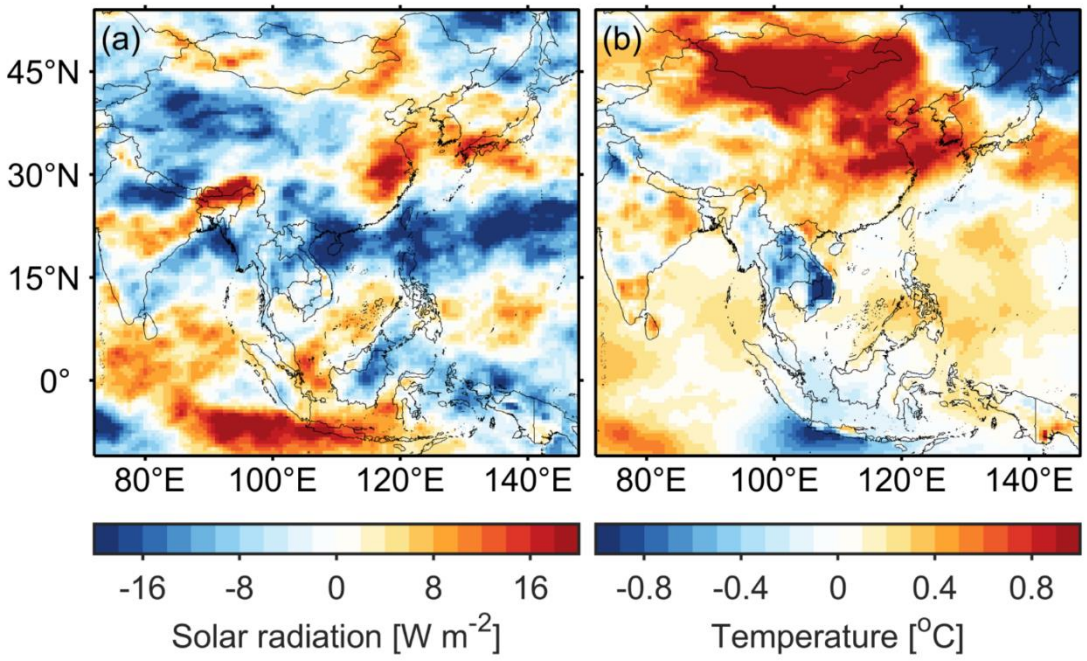


**Figure S10.** Difference in meteorological parameters in 2017-2019 and 2013-2015.