A new method for combining observations and multiple models for an improved estimate of the global surface ozone distribution

K.-L. Chang*(NRC postdoc at NOAA/ESRL), O.R. Cooper (NOAA/ESRL & CIRES), J.J. West, M. Serre (UNC, Chapel Hill), M.G. Schultz (IEK-8), M. Lin (NOAA/GFDL), V. Marecal (Météo-France), M. Deushi (MRI), K. Sudo (Nagoya U), J. Liu, and C.A. Keller (NASA/GMAO)
*kai-lan.chang@noaa.gov

Background

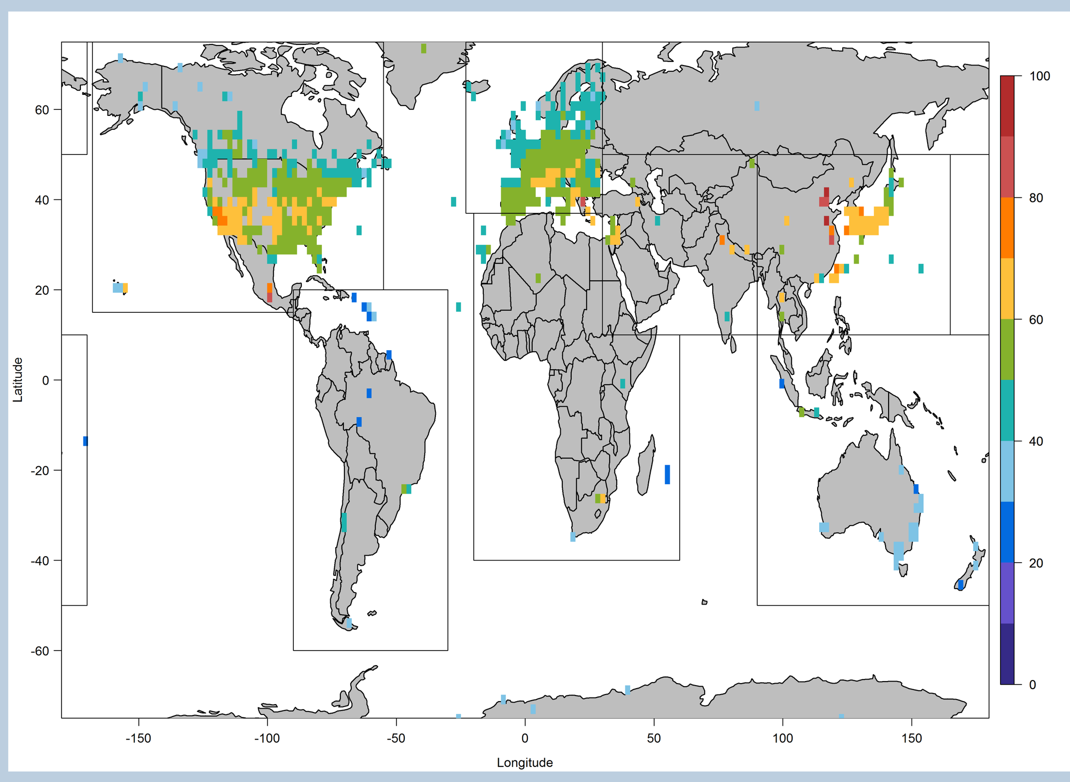
- Tropospheric ozone is a pollutant detrimental to human health and has been associated with a range of adverse cardiovascular and respiratory health effects due to short-term and long-term exposure.
- Assessing the human health impacts of ozone on the global scale requires accurate exposure estimates at any given inhabited location.

Aim: to combine the output from many models in a way that takes advantage of the strengths of each model and minimizes the weaknesses.

Sources

- *Tropospheric Ozone Assessment Report (TOAR) database*: The TOAR has produced the world's largest database of surface ozone metrics based on hourly observations at over 9000 sites around the globe (Schultz et al., 2017). Spatial coverage is high in North America, Europe, South Korea and Japan, but much lower across the rest of the world with very low data availability across Africa, the Middle East, Russia and India.
- *Atmospheric chemistry models*: Output from 5 models (CHASER, GEOSCCM, GFDL-AM3, MOCAGE and MRI-ESM1r1) was made available through the phase 1 of the Chemistry-Climate Model Initiative (CCMI). The high resolution GEOS-5 nature run with chemistry (G5NR-Chem) was provided by the NASA GMAO.

Annual maximum of the 6-month running mean of the maximum daily 8-hour average (DMA8) was calculated at every monitoring site and every model grid. This metric was selected because it aligns with the ozone metric used by Turner et al. (2016) to quantify the impact of long-term ozone exposure on human mortality.



Reference

[1] Chang et al. (2017) *Regional trend analysis of surface ozone observations from monitoring networks in eastern North America, Europe and East Asia.* Elem Sci Anth
[2] Schultz et al. (2017) *Tropospheric Ozone Assessment Report: Database and metrics data of global surface ozone observations.* Elem Sci Anth
[3] Turner et al. (2016) *Long-term ozone exposure and mortality in a large prospective study.* Am. J. Respir. Crit. Care Med

Data-model synthesis

We proposed the following procedure to combine model outputs and observations for data integration:

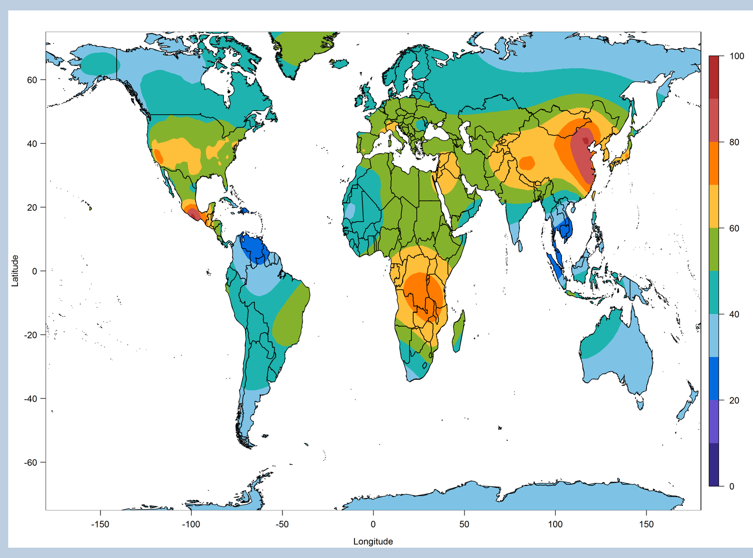
1. *Registering irregularly located monitoring observations & regular cell model output*: We use a statistical kriging technique to interpolate ozone observations from irregularly located monitoring sites onto a 0.125° × 0.125° grid, using the SPDE (stochastic partial differential equation) approach.
2. *Matching spatially interpolated ozone and model output by regions*: We divide the global land surface into 8 regions, roughly matching the continental outlines or major population regions.
3. *Weighting model output against spatially interpolated ozone*: Next regress the observations on multi-model output by a constrained least square approach within each of the eight regions. Let s_g be the grid cell at resolution 0.125° × 0.125°, and $\hat{y}(s_g)$ be the interpolated observations. The optimization equation is

$$\text{minimize} \sum_{s_g \in Region} \left(\hat{y}(s_g) - \alpha - \sum_{k=1}^{6} \beta_k \text{Model}_k(s_g)\right)^2, \text{subject to} \sum_{k=1}^{6} \beta_k = 1 \text{ and } \beta_k \geq 0.$$

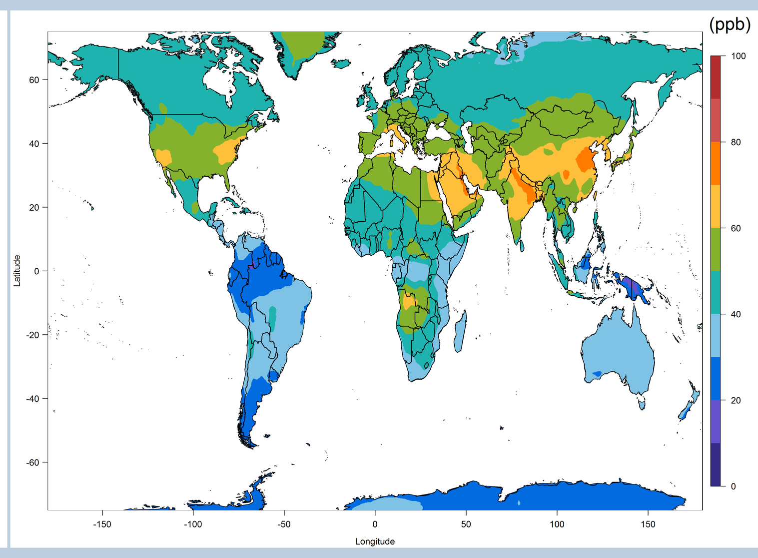
4. *Correcting multi-model bias*: We cannot correct the model without actual observations, thus we only correct the output grid where there is at least one station within a 2° radial distance of the station: We then end up using

$$\begin{cases} \hat{y}(s_g), & \text{if } s_g \text{within a 2 degree radial distance of the station;} \\ \alpha + \sum_{k=1}^{6} \beta_k \eta_k(s_g), & \text{otherwise,} \end{cases}$$

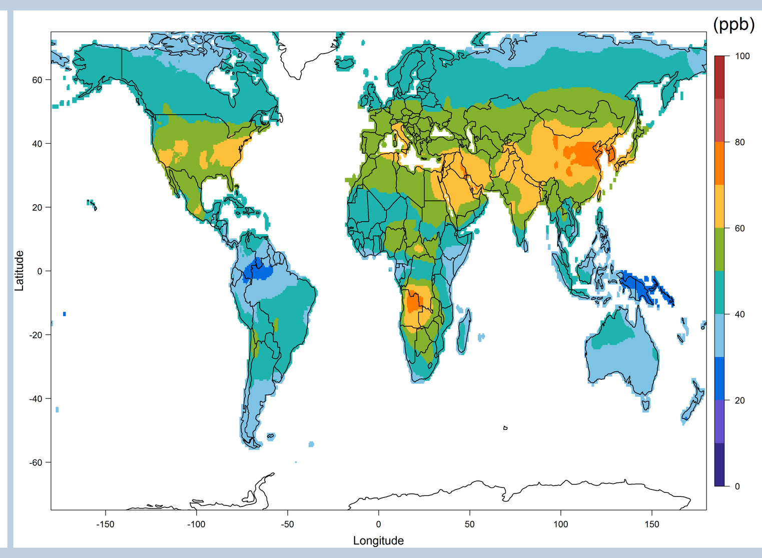
to generate our high resolution global surface ozone estimate.



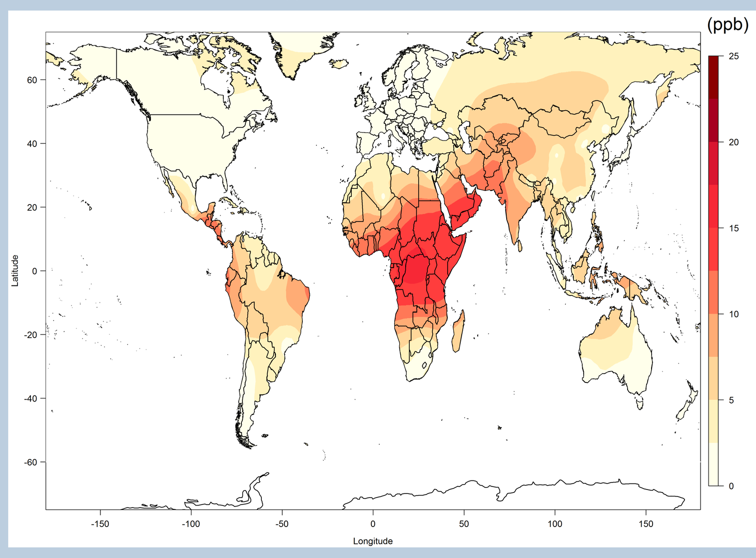
Interpolated mean



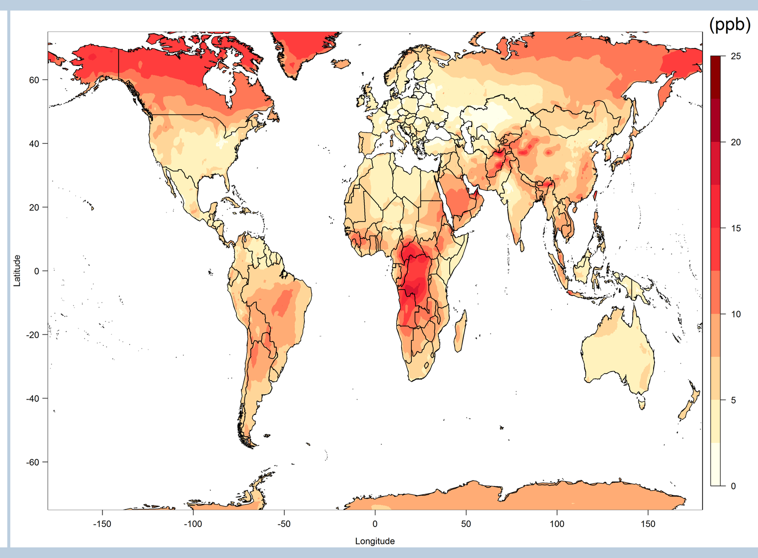
Multi-model average



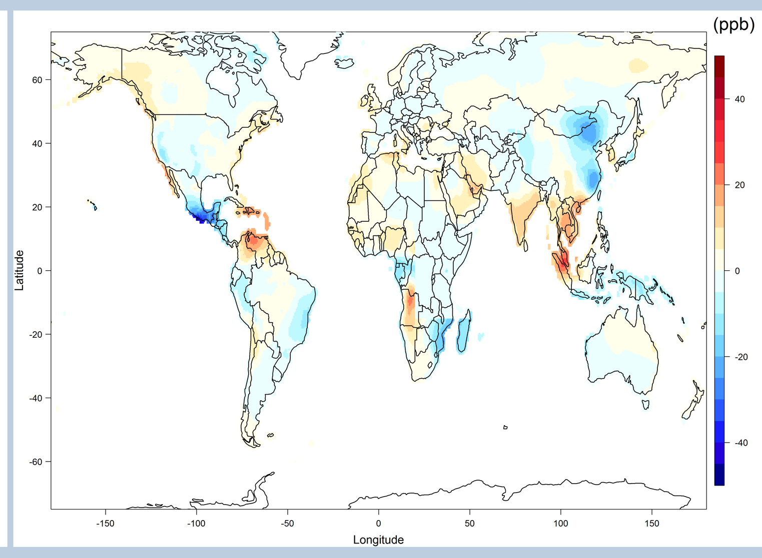
Multi-model composite



Interpolated SE

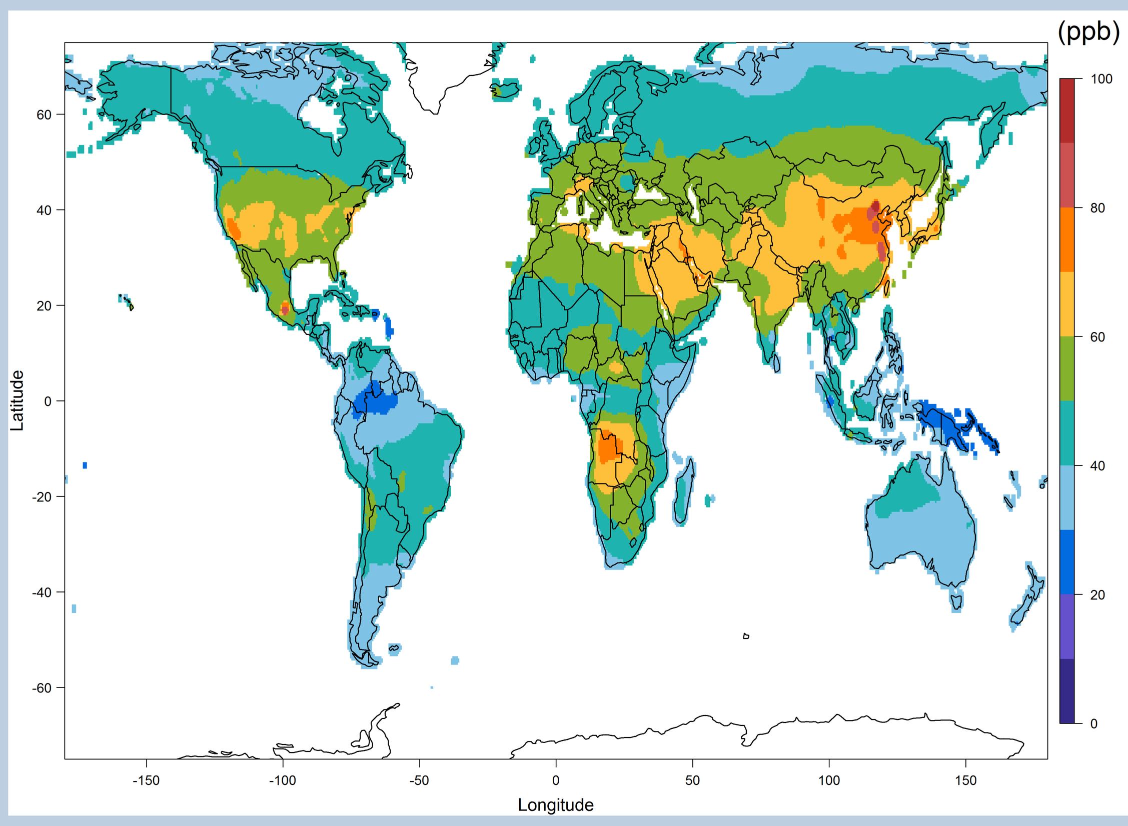


Multi-model SD

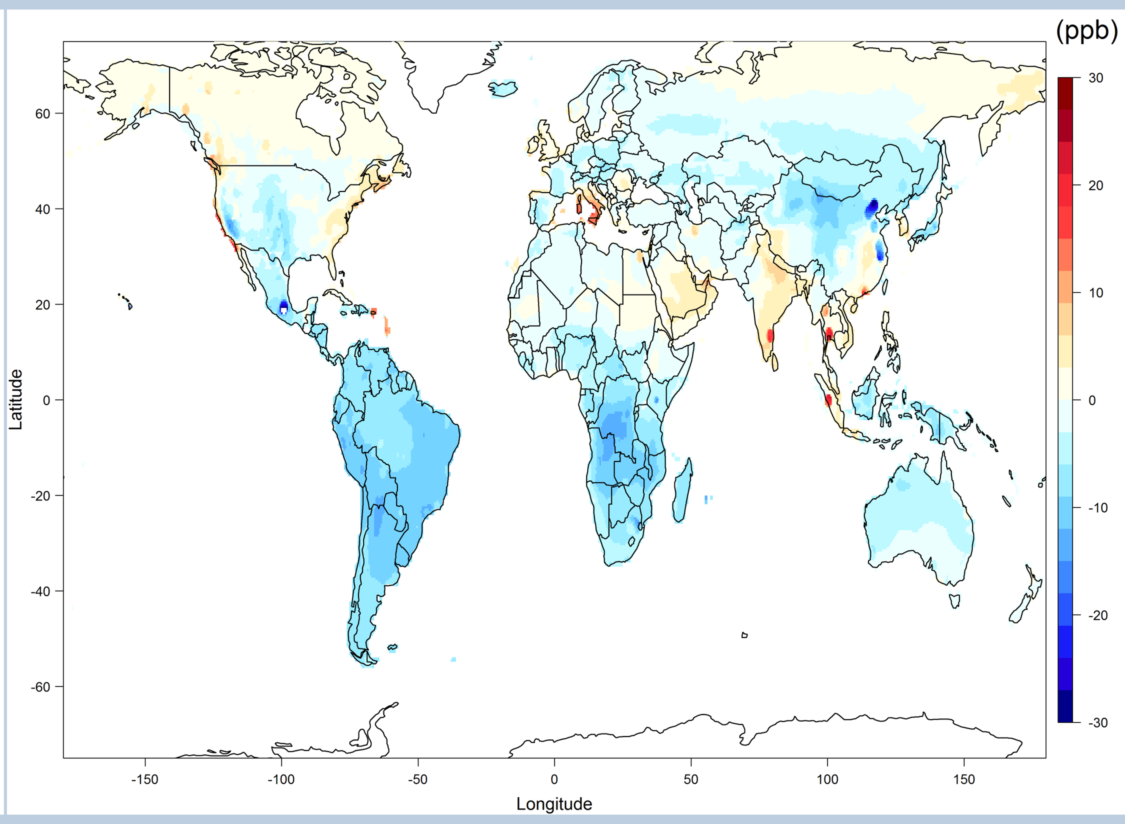


Multi-model bias

Model correction through observations



2° degree correction surface



Multi-model average minus the fused surface o3

RMSE against TOAR observations.

	MMA	Composite	Fusion
E Asia	11.43	8.70	5.29
Europe	7.33	5.61	4.50
N America	7.75	5.68	3.27
Overall*	8.48	6.76	4.35