Supplementary Information

Indian annual ambient air quality standard is achievable by completely mitigating emissions from household sources

Sourangsu Chowdhury[1,2], *Sagnik Dey[1,3], Sarath Guttikunda[4], Ajay Pillarisetti[2], *Kirk R. Smith[2,5], Larry Di Girolamo[3]

[1]Centre for Atmospheric Sciences, Indian Institute of Technology Delhi, New Delhi, India.

[2]School of Public Health, University of California, Berkeley, California, USA

[3]Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

[4]Urban Emissions, New Delhi, India.

[5]Collaborative Clean Air Policy Centre, Delhi, India

The supplementary information contains descriptions of the preparation of the emission inventory for cooking, lighting, and space and water heating. It also contains four figures.

www.pnas.org/cgi/doi/10.1073/pnas.1900888116

Preparation of emission inventory

The improvements made in the cooking and lighting emissions estimation methodology are as follows:

- All the emissions were calculated at grid level. Grids are designated at 0.25° resolution (~25 km) and spread over a working domain covering the Indian subcontinent as shown in Figure S1.
- Each of the grids is mapped to the districts, listed in the Census-India (2011). This also allows for fractional mapping. For example, if a grid is covering multiple districts, then the area of the overlap is taken into consideration for appropriate mapping of the grid.
- The total number of households in each district is calculated based on the total population and the household size for each district.
- Each of the districts and grids are further broken down into urban and rural areas – based on the urban build-up retrieved from MODIS Land Cover Type Yearly L3 (MCD12Q1) obtained at 500 m resolution for the year 2013 (https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mcd12q1).
- The urban and rural split within the district, linked to the gridded population, prevented overestimation of emissions over the urban grids and underestimations of the emissions over the rural grids, especially for the cooking sector.
- Each of the districts is mapped to the data fields from Census-India (2011) – HH10 for the HEC for cooking and heating and HH7 for lighting. The data is further segregated at the district level into urban and rural cooking and inside and outside cooking.
- The HH10 database allows analysis of nine fuel categories (including electric) – this was utilized for cooking and heating categories. The assumption here is that the fuel used for cooking is also utilized for heating. The electric share of households is listed as zero emissions. The survey records only one fuel per household, but it is likely not true that a given household will use only one fuel category. In the calculations, the share of the households from the survey is used as a proxy since a mix of the fuels are likely being used by all the households in a district. For fuel consumption, corrections were introduced, especially for LPG consumption, based on the level of penetration of new connections at the state level(obtained from MoPNG, Govt. of India) and matching the overall LPG sales every

year. This allowed us to make adjustments to the overall Census data with respect to more recent and relevant data for year 2015.

- Fuels included in the cooking and heating emissions estimation are crop residues, wood, coal, cow dung, kerosene, coal and charcoal, LPG, biogas, and others (a small share of unknown fuel types).
- The average energy consumed by households for cooking is calculated based on the NSSO (2012) survey database, which lists the amount of food varieties cooked at the state level. Within the state, this is assumed constant for all districts. The average energy consumption is 1.07 and 1.16 GJ per capita per year for rural and urban settings, respectively. The lowest averages are observed in the Northeastern states(1).
- The HH7 database from the Census allows analysis for four fuel categories (including solar and electric) for lighting. The solar and electric share of households are listed as zero emissions. Most rural lighting needs are met via kerosene.
- The water and space heating emissions estimations are linked to a dynamic meteorological database, which allows for daytime and nighttime temperature profiles at the grid level and the population database available as age groups. This was used to improve the emissions in two ways:
    1. In the past, the global inventories assumed that the southern states do not use space or water heating, based on the monthly average temperatures. It is our assumption that water heating is a year-long process, especially for certain age groups – children under 15 years and elderly over 55 years.
    2. A temperature trigger was set to estimate space and water heating for all age groups. However, for the children and the elderly, this trigger was nullified.
- The dynamic temperature profiles and correction based on age groups prevented underestimation of emissions in Southern states (for all months) and in Northern and Northeastern states (during the summer months). For example, the Western Ghats in the Southern states are known to be cold and in need of space heating even in the summer months, which would be assumed zero if a state average temperature is assumed for the region.
- The cooking and lighting emissions are considered constant through the year. The water heating and space heating emissions vary by hour and day.

Atmospheric dispersion modelling was conducted to study the movement of emissions on a regional scale, the formation of the secondary sulfate particulates (part of $PM_{2.5}$), and their contribution to the health impacts. The Comprehensive Air Quality Model with Extensions (CAMx) version 6.2, an Eulerian photochemical dispersion model, was utilized for dispersion modelling. Meteorological data (3D wind, temperature, pressure, relative humidity, and precipitation fields) was derived from the National Center for Environmental Prediction (NCEP) global reanalysis database and processed through WRF meteorological model (v6.2) at 1 hour temporal resolution. The model was simulated for the entire year of 2015 with the emissions discussed in the preceding section. The model was simulated at 0.25º×0.25º resolution over coordinates covering the entire Indian landmass (7-39ºN and 67-99ºE). Biogenic emissions were obtained from EDGAR Global inventory (2), and gas phase chemistry obtained from the SAPRC99 (3) mechanism was utilized. The initial and boundary conditions were obtained from MOZART-4 offline model (4). Further details of the model simulation and model validation with satellite and in-situ data are provided elsewhere (3, 5, 6). Two sets of simulations were performed: (a) total $PM_{2.5}$ was modelled with all emissions from all anthropogenic and natural sources like transport, industries, household, power plants, brick kilns and agriculture (1, 7, 8) (b) $PM_{2.5}$ was modelled with all emissions minus household emissions. The difference between (a) and (b) estimates the contribution of household sources to ambient $PM_{2.5.}$ Secondary organic aerosols were not considered while modelling the total $PM_{2.5}$; this assumption is not expected to have a significant impact on the contribution of households to ambient $PM_{2.5}$ concentrations (9).

Figures S2 (a) and (b) depict the emissions from cooking and lighting at grid level over India. Figure S3 and S4 depict the emission of $PM_{2.5}$ at monthly scale over India due to space heating and water heating.

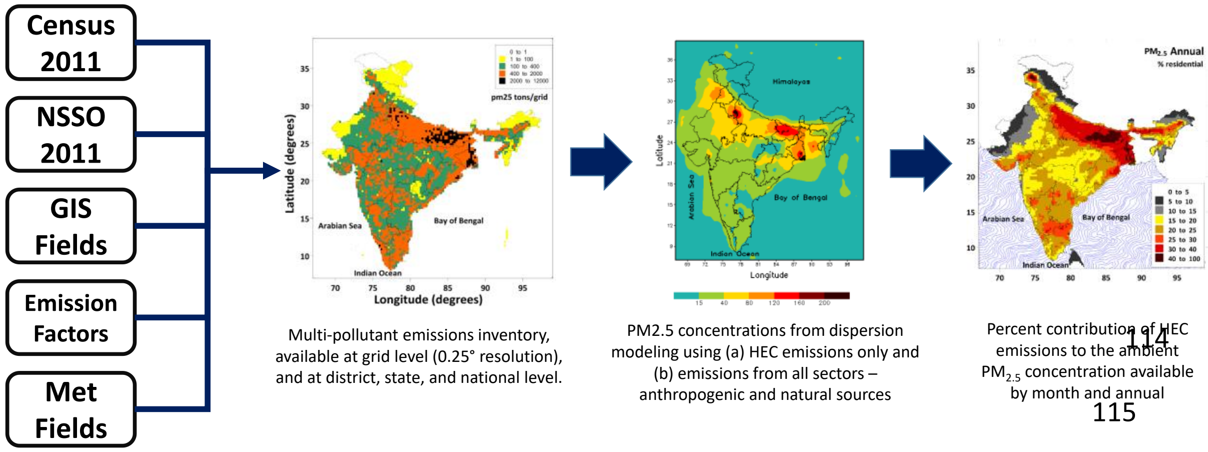

Figure S1. Depiction of the methodology to estimate the contribution of household emissions towards ambient air pollution.

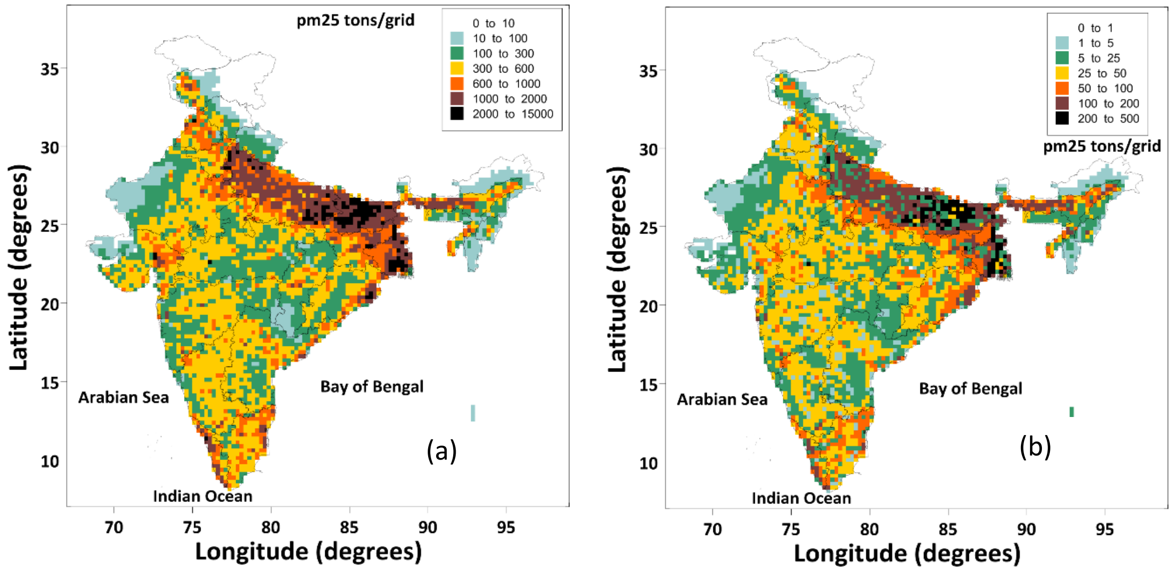

Figure S2. Total annual $PM_{2.5}$ emissions from cooking (a) and lighting (b) in India.

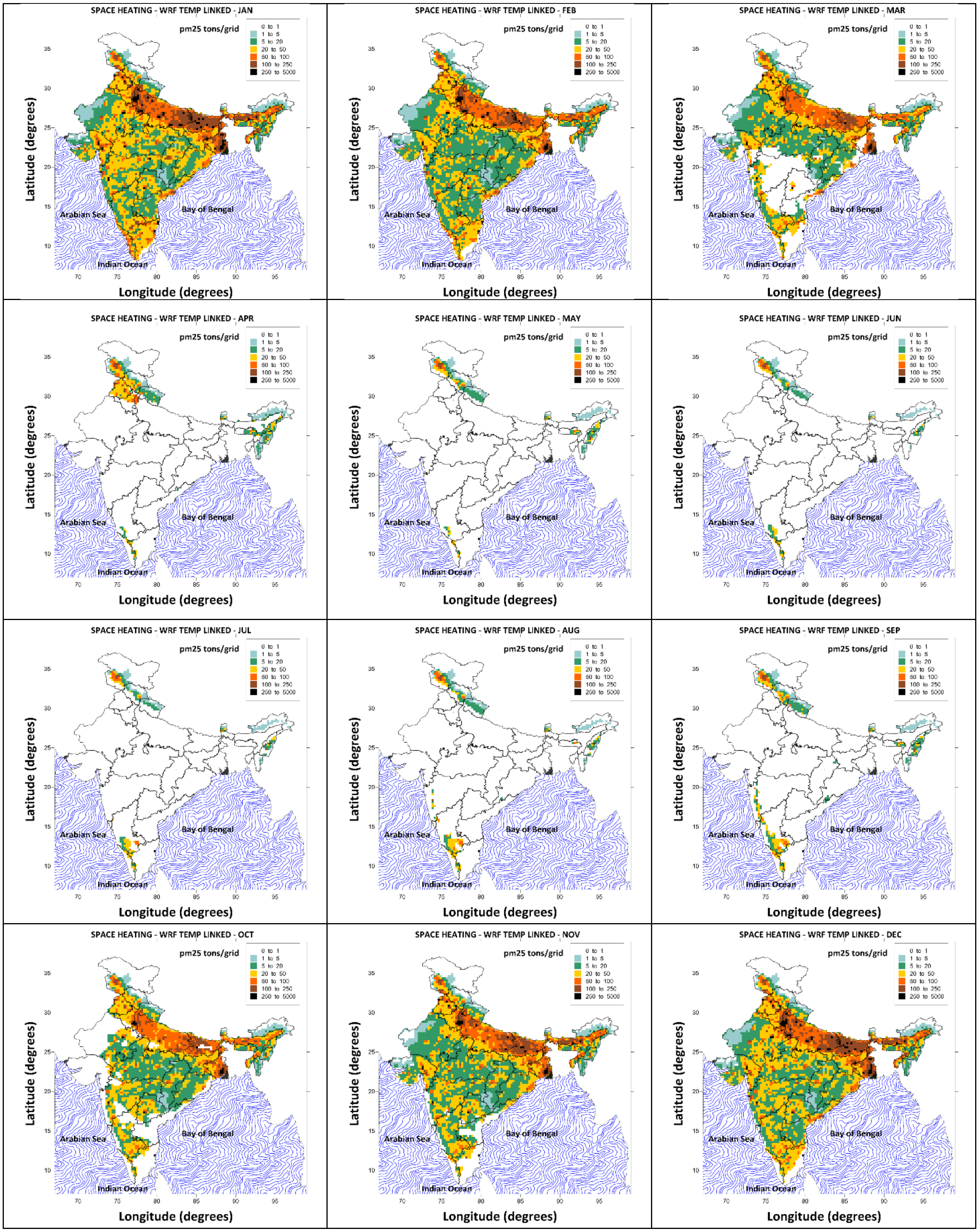

Figure S3. Monthly total $PM_{2.5}$ emissions from energy consumption for space heating in India.

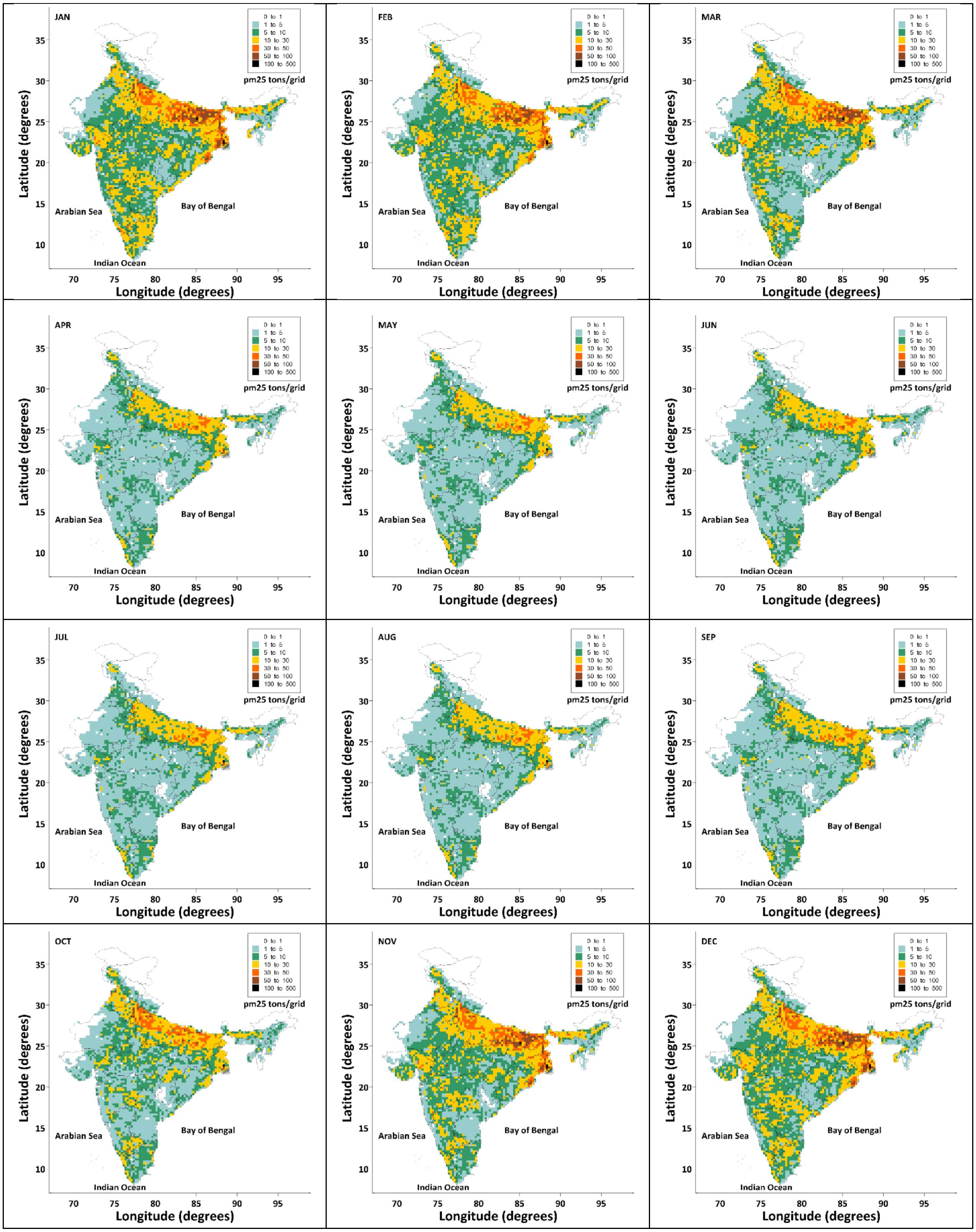

Figure S4. Monthly total $PM_{2.5}$ emissions from energy consumption for water heating in India.

Estimation of premature mortality burden

Premature mortality (M) may be estimated using the following equation:

$$\sum_{i=1,k=1}^{N,N1} M_{i,k,} = \sum_{i=1,k=1}^{N,N1,N2} y_{i,k} \times \frac{RR_{i,k}-1}{RR_{i,k}} \times P_i$$

Where $M_{i,,k}$ is the premature mortality in a particular district *i* for a disease *k*. $M_{i,j}$ is estimated as a function of baseline mortality $y_{i,k}$. Baseline mortality is adjusted as a function of gross development product (GDP) following our earlier study(10). Relative risks ($RR_{i,k}$) are estimated with integrated exposure-response (IER) functions(11). The adult population (P_i) for a district *i* above 25 years is obtained from the Census of India, 2011. We estimate the averted premature mortality as the change between premature mortality for a given scenario relative to that for the baseline year of 2015. There are large uncertainties in both the IERs and in the input data required for estimation of premature mortality. As such, we quantified the % health benefits (in terms of % averted mortality) of our devised mitigation measures by using the central value of premature mortality estimates.

Attribution Method

We used the fractional reduction in ambient $PM_{2.5}$ exposure to scale the averted premature mortality(12) (attribution method, in Figure 5, main paper). The difference in premature mortality estimated using the IER and the attribution method (Figure 5) can be attributed to the non-linear shape of the IER curves in these exposure ranges.

Supplementary References

1. Pandey A, Sadavarte P, Rao AB, Venkataraman C (2014) Trends in multi-pollutant emissions from a technology-linked inventory for India: II. Residential, agricultural and informal industry sectors. *Atmos Environ* 99:341–352.

2. Olivier J, et al. (1996) *Description of EDGAR Version 2.0: A set of global emission inventories of greenhouse gases and ozone-depleting substances for all anthropogenic and most natural sources on a per country basis and on 1 degree x 1 degree grid* Available at: https://www.rivm.nl/bibliotheek/rapporten/771060002.html.

3. Sarkar M, Venkataraman C, Guttikunda S, Sadavarte P (2016) Indian emissions of technology-linked NMVOCs with chemical speciation: An evaluation of the SAPRC99 mechanism with WRF-CAMx simulations. *Atmos Environ* 134:70–83.

4. Emmons LK, et al. (2010) Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4). *Geosci Model Dev* 3(1):43–67.

5. Guttikunda SK, Nishadh KA, Jawahar P (2019) Air pollution knowledge assessments (APnA) for 20 Indian cities. *Urban Clim* 27(August 2018):124–141.

6. MAPS Working Group G (2018) Burden of Disease Attributable to Major Air Pollution Sources in India. (January).

7. Pandey A, Sadavarte P, Rao AB, Venkataraman C (2014) Trends in multi-pollutant emissions from a technology-linked inventory for India: I. Industry and transport sectors. *Atmos Environ* 99:353–364.

8. Pandey A, Venkataraman C (2014) Estimating emissions from the Indian transport sector with on-road fleet composition and traffic volume. *Atmos Environ* 98:123–133.

9. Rooney B, et al. (2018) Impacts of Household Sources on Air Pollution at Village and Regional Scales in India Brigitte. *Atmos Chem Phys Discuss* (December):1–51.

10. Chowdhury S, Dey S (2016) Cause-specific premature death from ambient PM<inf>2.5</inf>exposure in India: Estimate adjusted for baseline mortality. *Environ Int* 91. doi:10.1016/j.envint.2016.03.004.

11. Burnett RT, et al. (2014) An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environ Health Perspect* 122(4):397–403.

12. Conibear L, Butt EW, Knote C, Arnold SR, Spracklen D V. (2018) Residential energy use emissions dominate health impacts from exposure to ambient particulate matter in India. *Nat Commun* 9(1):1–9.