28th May 2021

Air pollution and road exposure variables for ALSWH participants (by Luke Knibbs, UQ)

Overview and usage

Several air pollution exposure estimates have been linked to the geocoded residential addresses of ALSWH participants from 1996 to 2017. Estimates have been done for each person and each year in this period, regardless of whether a survey was performed.

Please note, there are no costs payable for non-commercial research-related use of these estimates. It is expected that any academic publications (journal article, conference paper, book chapter, etc.) that use these air pollution and/or roads data offer the data contributor (A/Prof Luke Knibbs, l.knibbs@uq.edu.au) the opportunity to be a co-author. Apart from recognition of the work of the contributor, this is also a way to ensure the data are applied and interpreted in the most valid way, which benefits users and saves time. The contributor may choose not to take up the option of co-authorship, but can still offer useful content knowledge and feedback.

Applying to use the data

1. In the first instance, collaborators wishing to use these data must contact Luke Knibbs (l.knibbs@uq.edu.au) to discuss their research proposal.
2. Lodge your Expression of Interest http://www.alswh.org.au/how-to-access-the-data/alswh-data

Acknowledging the data in research output

In cases where A/Prof Luke Knibbs is **not** a co-author on the paper, he must be acknowledged with the statement: "The authors thank Associate Professor Luke Knibbs of the School of Public Health, The University of Queensland, for permission to use the linked Air Pollution data".

Variables

The following air pollution and road variables have been linked to the geocoded addresses of ALSWH participants from 1996 to 2017.

- Annual mean nitrogen dioxide (NO_2) for each year from 1996 to 2017 (inclusive). These variables are named 'PRED_NO2_year'. **NO2 estimates are in units of parts per billion (ppb)**. The papers describing the development and validation of the NO_2 models used to estimate air pollution are:

  Knibbs, LD, Coorey, CP, Bechle, MJ, Marshall, JD, Hewson, MG, Jalaludin, B, Morgan, GG, Barnett, AG. Long-term nitrogen dioxide exposure assessment using back-extrapolation of satellite-based land-use regression models for Australia. *Environmental Research* 2018;163:16-25.

  Knibbs, LD, Coorey, CP, Bechle, MJ, Cowie, CT, Dirgawati, M, Heyworth, JS, Marks, GB, Marshall, JD, Morawska, L, Pereira, G, Hewson, MG. Independent validation of national satellite-based land-use regression models for nitrogen dioxide using passive samplers. *Environmental Science & Technology*, 2016;50:12331-12338.

Knibbs, LD , Hewson, MG, Bechle, MJ, Marshall, JD, Barnett, AG. A national satellite-based land-use regression model for air pollution exposure assessment in Australia. *Environmental Research* 2014; 135:204-211.

- Annual mean fine particulate matter ($PM_{2.5}$) for each year from 1998 to 2016 (inclusive). Estimates prior to 1998 and after 2016 are not available for $PM_{2.5}$. Predictions from two different models (SATWOOD [referred to as 'SAT' in paper below], SATnoWOOD [referred to as 'SAT-W' in paper below]) are included ('PRED_PM25_SATWOOD_year' and 'PRED_PM25_SATnoWOOD_year'). **$PM_{2.5}$ estimates are in units of micrograms per cubic metre (µg/m^3)**. The paper describing the development and validation of the $PM_{2.5}$ models used to estimate air pollution is:

  Knibbs, LD, van Donkelaar A, Martin, RV, Bechle, MJ, Brauer, M, Cohen, DD, Cowie, CT, Dirgawati, M, Guo, Y, Hanigan, IC, Johnston, FH, Marks, GB, Marshall, JD, Pereira, G, Jalaludin, B, Morgan, GG, Barnett, AG. Satellite-based land-use regression for continental-scale long-term ambient $PM_{2.5}$ exposure assessment in Australia. *Environmental Science & Technology*, 2018;52:12445-12455.

  The optimum $PM_{2.5}$ model for a given application will vary and users should contact Luke Knibbs (l.knibbs@uq.edu.au) to discuss their requirements. $PM_{2.5}$ levels do not generally have substantial year-to-year variability, and if necessary, users can try to fill in data for 1996, 1997 and 2017 using their preferred imputation method.

- The total length of roads in circular buffers with radii of 100, 200 and 500 m. Estimates are presented for all road types ('SUM_ALLROADS_buffer') and major roads only ('SUM_MAJROADS_buffer'), based on standard Australian road classifications. **The road length estimates are in units of kilometres (km)**.

- The straight line distance from each address to the nearest road. Estimates are presented for all road types ('ROAD_DIST_ALL') and major roads only (ROAD_DIST_MAJOR), based on standard Australian road classifications. **The road distance estimates are in units of metres (m).** No attempt has been made to remove extreme values (e.g. living on an island far from the nearest road).

- **IMPORTANT ADDITIONAL NOTES**:
  - A missing value for a pollution variable generally means one or more of the variables used to estimate it were not available at that location.
  - The accuracy of the geocoded address ('accuracy' field) should be considered when using the estimates, as this can have a substantial impact on the validity of pollutant and road estimates. Contact Luke Knibbs to discuss this issue.
  - Some people move house, sometimes many times – not considering this can yield misleading results. Contact Luke Knibbs to discuss.