London School of Hygiene and Tropical Medicine

Improving Health Worldwide

Big data in environmental epidemiology

Arturo de la Cruz Libardi

Environment and Health Modelling (EHM) Lab
London School of Hygiene and Tropical Medicine

2024-03-12

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

intended learning outcomes

by the end of this session, you will be able to:

1. Critically define big data

2. Describe some implications and applications of big data in public health and epidemiology

3. Evaluate sources of big health and environmental data

4. Explain why and how environmental and health data are linked

table of contents

London Hybrid Exposure Model: Improving Human Exposure Estimates to NO2 and PM2.5 in an Urban Setting [4]

residential vs modelled exposure

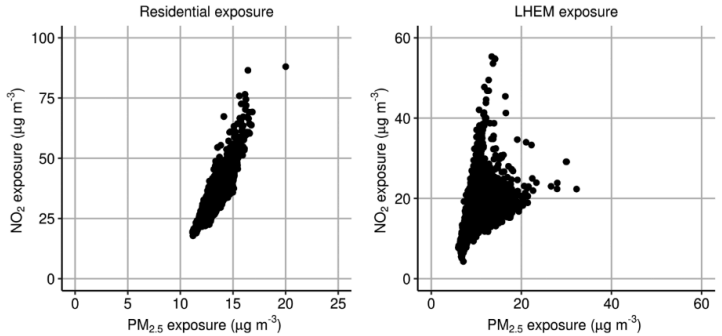


Figure 4. Scatter plots of NO_2 and $PM_{2.5}$ exposure outdoors at the residential address (left) (R = 0.90) and using the LHEM (right).

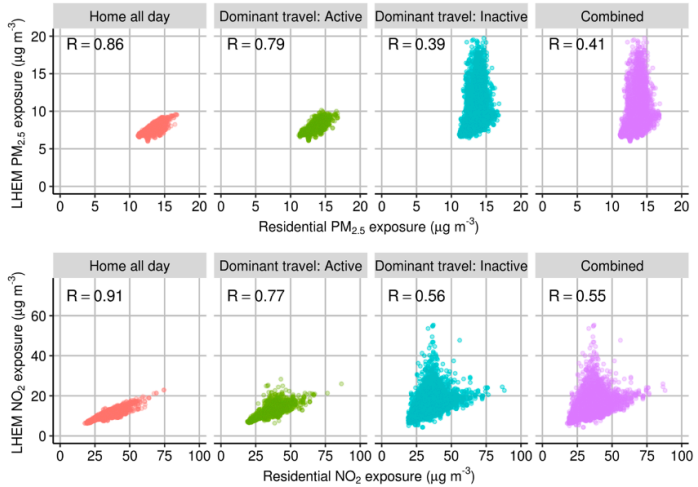
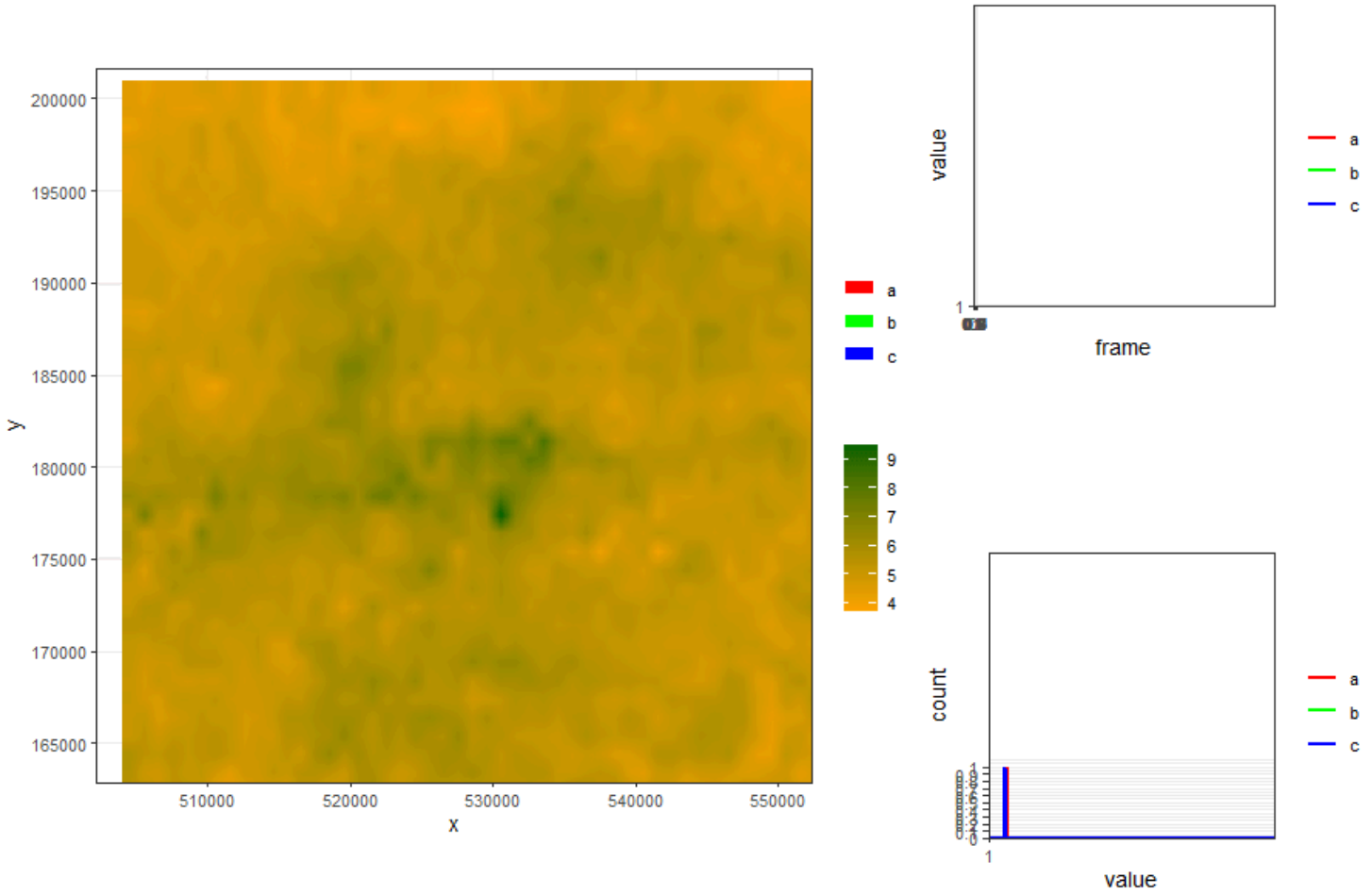


Figure 3. Scatter plots of NO_2 (bottom) and $PM_{2.5}$ (top) LHEM exposure versus exposure at the residential address - demonstrating the relative strength of the relationship between those who undertake active travel (cycle and walk), those that stay at home, and those who undertake inactive travel (car, motorcycle, bus, train, and tube).

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

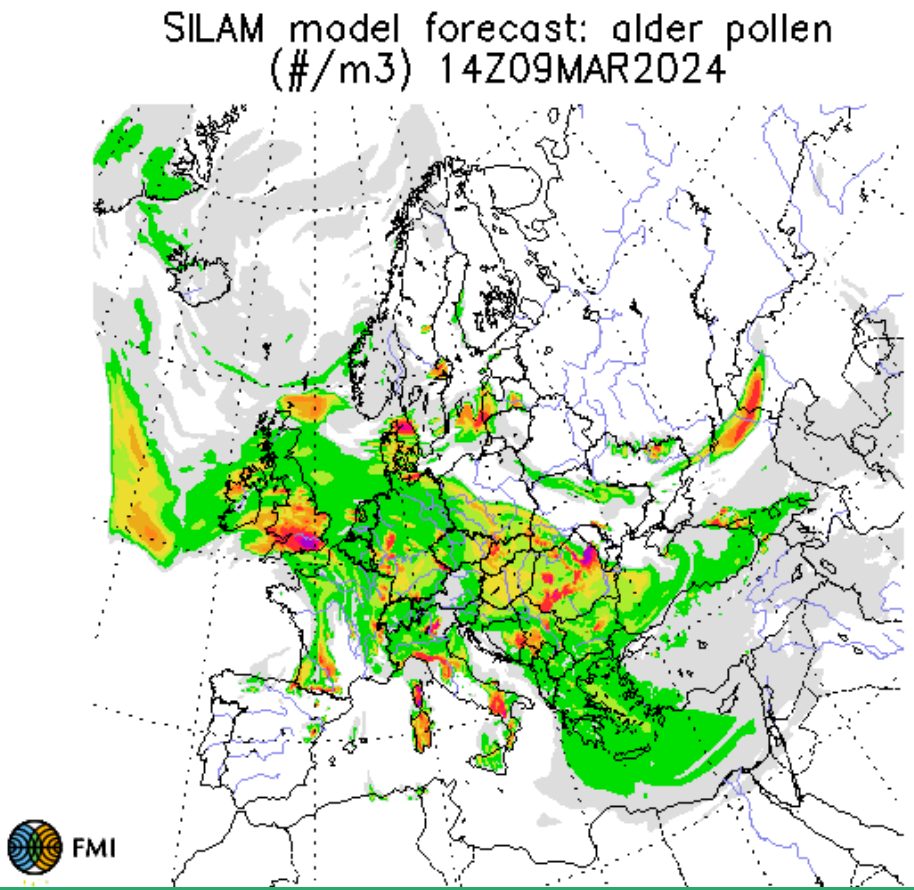

200000
195000
190000
185000
180000
175000
170000
165000
y
510000
520000
530000
540000
550000
x
a
b
c
9
8
7
6
5
4
value
frame
a
b
c
count
value
a
b
c

SILAM v.5.7

System for Integrated modeLling of Atmospheric coMposition

FINNISH METEOROLOGICAL INSTITUTE (http://en.ilmatieteenlaitos.fi/)

Region: europe



Control panel

Pollen taxon:

Alder | Birch | Grass

Olive | Ragweed | Mugwort

Pollen Index | Allergy Risk

Region:

Europe

Hour: -58

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

OpenSAFELY query (ehrQL) reliability testing using generative artificial intelligence!

we have learned to

1. Critically define big data as big data processes

2. Describe some implications and applications of big data in public health and epidemiology

- **classical (measurement error, confounding) challenges**
- **new (comprehensive health data, real-time action) opportunities**

3. Evaluate sources of big health and environmental data

- **health genetic data, EHRs, wearable sensors**
- **environment reanalyses, satellite, ground sensors**

4. Explain why and how environmental and health data are linked

- **depends on nature of health data and resolution of environmental data**

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

end of part 1

references

[1] T. R. Gadekallu, Q. Pham, T. Huynh-The, et al. *Federated Learning for Big Data: A Survey on Opportunities, Applications, and Future Directions*. En. 2021.

[2] C. Tonne, X. Basagaa, B. Chaix, et al. "New frontiers for environmental epidemiology in a changing world". In: *Environment International* 104 (Jul. 2017), pp. 155-162. ISSN: 0160-4120. DOI: 10.1016/j.envint.2017.04.003. URL: https://www.sciencedirect.com/science/article/pii/S0160412017301459 (visited on 02/15/2024).

[3] J. Vanoli, M. N. Mistry, A. De La Cruz Libardi, et al. "Reconstructing individual-level exposures in cohort analyses of environmental risks: an example with the UK Biobank". En. In: *Journal of Exposure Science & Environmental Epidemiology* (Jan. 2024). ISSN: 1559-0631, 1559-064X. DOI: 10.1038/s41370-023-00635-w. URL: https://www.nature.com/articles/s41370-023-00635-w (visited on 03/10/2024).

[4] J. D. Smith, C. Mitsakou, N. Kitwiroon, et al. "London Hybrid Exposure Model: Improving Human Exposure Estimates to NO2 and PM2.5 in an Urban Setting". En. In: *Environmental Science & Technology* 50.21 (Nov. 2016), pp. 11760-11768. ISSN: 0013-936X, 1520-5851. DOI: 10.1021/acs.est.6b01817. URL: https://pubs.acs.org/doi/10.1021/acs.est.6b01817 (visited on 02/02/2023).

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

references

[5] R. Schneider, A. Vicedo-Cabrera, F. Sera, et al. "A Satellite-Based Spatio-Temporal Machine Learning Model to Reconstruct Daily PM2.5 Concentrations across Great Britain". En. In: *Remote Sensing* 12.22 (Nov. 2020), p. 3803. ISSN: 2072-4292. DOI: 10.3390/rs12223803. URL: https://www.mdpi.com/2072-4292/12/22/3803 (visited on 02/03/2022).

Presentation made with xaringan in RStudio.

suggestions?

- DASH 26th March opening event
- hundreds of hours of free and open resources
- a lot of local and global circumstances to improve