Data Integration Model for Exposures (DIMEX-UK)

Dr. Matthew Thomas
Lecturer in Environmental Intelligence

JCEEI - UoM Joint Meeting

27th July 2022

With thanks to...

University of Exeter

- Gavin Shaddick
- Matthew Thomas
- Mike Diessner
- Stefan Siegert

Technical University of Denmark

- Karyn Morrisey

University of British Columbia

- Jim Zidek

University of Manchester

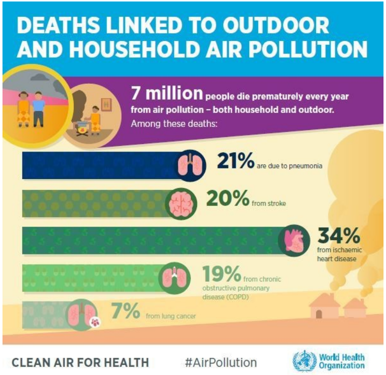
- Dave Topping
- Tom Bannan
- Hugh Coe
- James Evans

Alan Turing Institute

- Ruth Bowyer
- Fernando Benitez-Paez

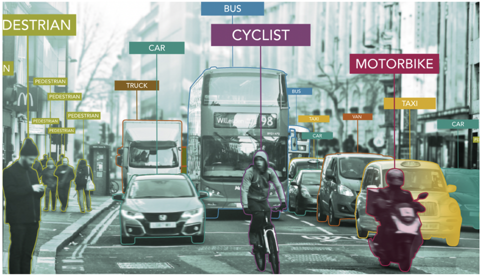
Introduction

- Air pollution poses significant threats to human health.
- In the UK, poor air quality is the largest environmental risk to public health.
- Majority of the research related to the health effects of air pollution has been at a population level
  - measured or modelled concentrations of ambient pollution
  - matched to residential addresses.
- This does not necessarily reflect individuals exposure to air pollution.
- People move through series of micro-environments throughout the day (work, home, school, outdoor, car, . . .)



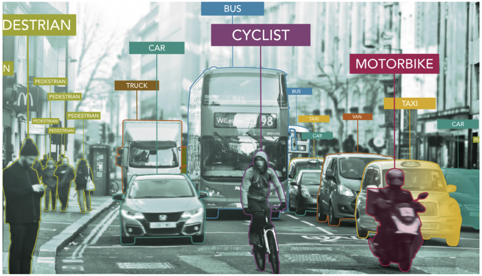
Introduction

- This kind of exposure model can be obtained through *direct* measurements using exposure monitors.
  - costly
  - time-consuming
  - sample sizes are (often) small
- Alternatively, *indirect* methods have been used such as agent-based modelling (or micro-simulation)
- Simulate/ human activities and combines them with modelled concentrations of air pollution in each micro-environment.
- Previous models
  - APEX
  - pcNEM
  - SHEDS



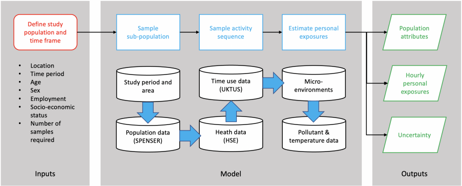
Aims

- Create an integrated modelling framework that estimates personal exposures to air pollution
- Simulate daily exposure of different population groups using agent-based modelling
- Understand differences between personal exposures and ambient concentrations.
- Investigate sub-population differences in exposures.



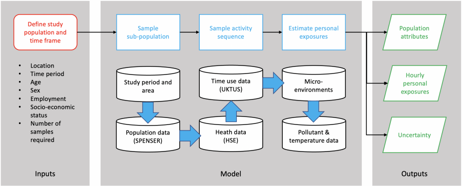
Data Integration Model for Exposures

- The Data Integration Model for Exposures (DIMEX) builds upon the framework developed for pCNEM
- Incorporates new modelling techniques and exploits the increased availability of granular information.
  - activity patterns across time.
  - different micro-environments.
  - demographic information.
- DIMEX generates a sequence of pollutant concentrations to which a randomly selected individual is exposed over time.



Data Integration Model for Exposures

- Stochastic process that follows a randomly selected individual in their activities over the period of the simulation.
- The individual is thought of as visiting one microenvironment after another as their activities change through time.
- Concentrations of $PM_{2.5}$ are estimated for each of the corresponding micro-environments and matched to each sampled individual.
- Result is hourly estimates of exposure to $PM_{2.5}$, which can be aggregated to yield multiple quantities of interest



DIMEX: Underlying Population

- Individuals are sampled from an underlying *synthetic* population from the area of interest.
- SPENSER combines census data with small scale surveys and datasets to create a geo-referenced synthetic population forecast at a high resolution.
- Supplemented with include data from the United Kingdom Time Use Survey and the Health Survey of England based on demographic information.
- Each individual are assigned to a Middle Layer Super Output Area (MSOA).

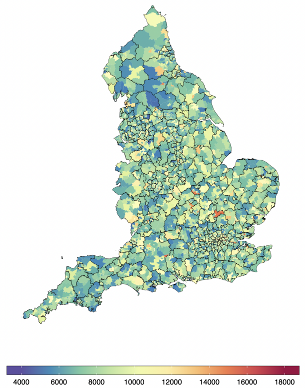


Figure: Estimated 2020 population counts by MSOA in England from SPENSER. Colours represent the level of population. and the black lines represent the local authority boundaries in England. In total, there is an estimated 56 million synthetic individuals in England.

DIMEX: Activity Sampler

- UK Time Use Survey (UKTUS) is a nationally-representative survey that provides information on how people aged eight years and over in the UK spend their time.
- Contains diaries consisting of sequences of activities and the locations that they take place between 4am to 4am in 10-minute intervals.
- Locations of each activity from the activity diaries were grouped into four micro-environments: Home, Indoor-not-home, Outdoor and Transport.

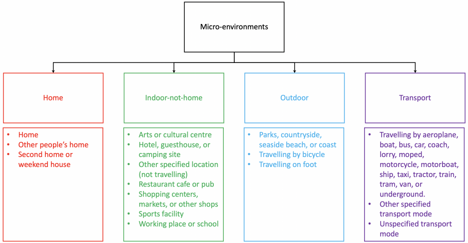


Figure: Grouping of the micro-environments.

DIMEX: Activity Sampler

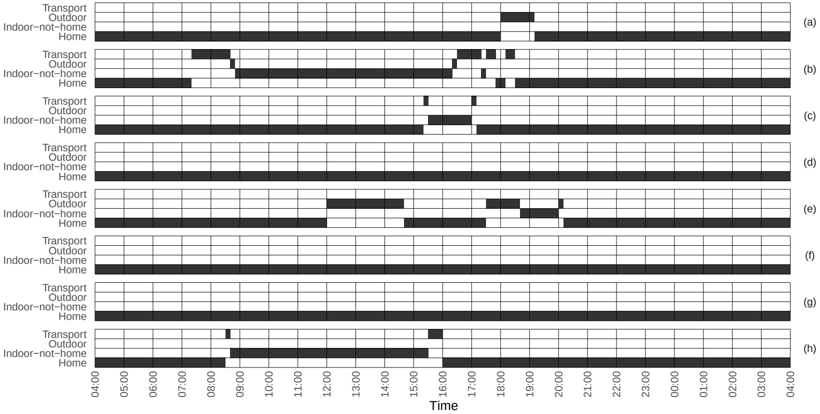


Figure: Example activity sequences in 10 minute intervals; (a) a full-time working female, during the weekend, (b) a full-time working female, during a weekday, (c) a retired male, during the weekend, (d) a retired male, during a weekday, (e) an unemployed male, during the weekend, (f) an unemployed male, during a weekday, (g) a school-child, during the weekend, (h) a school-child, during a weekday at term-time. Activities have been grouped by Home, Indoor-not-home, Transport and Outdoor.

DIMEX: Activity Sampler

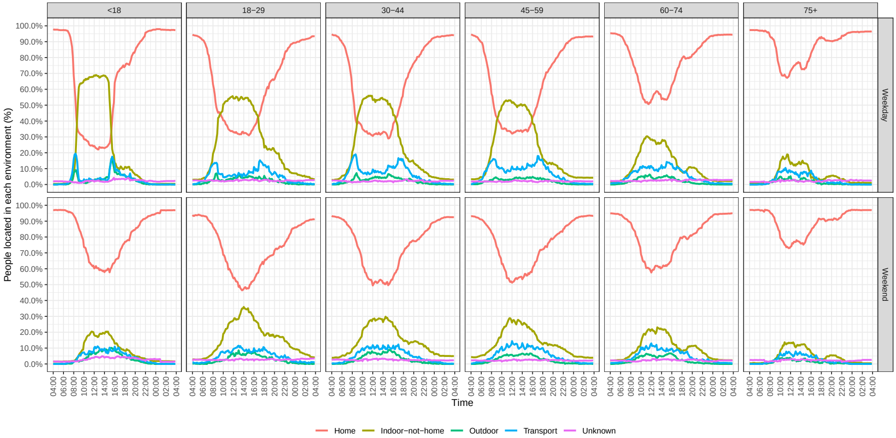


Figure: Proportion of surveyed individuals at Home, Indoor-not-home, Outdoor, Transport and Unknown locations by age group and day type from time activity diaries extracted from the UKTUS.

DIMEX: Activity Sampler

- People spend the majority of their time indoors (mean of 21.6 hours a day; 90%), the outputs from DIMEX-UK demonstrates the importance of indoor, specifically residential (17.9 hours a day; 75%),
- On a particular day, activity sequences of a sampled individual from SPENSER are generated by concatenating randomly selected UKTUS diaries.

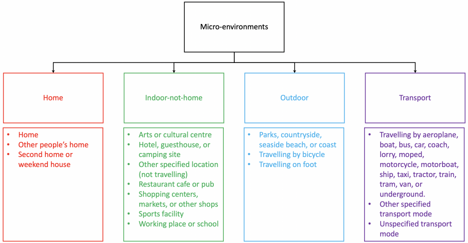


Figure: Grouping of the micro-environments.

DIMEX: Exposure Estimation

- Each of the individuals activity sequence needs to be matched to the corresponding concentrations of air pollutions they are exposed to in each micro-environment.
- Concentrations of $PM_{2.5}$ at home are modelled as a function of the the ambient outdoor concentrations and non-ambient sources of air pollution

$$local = f(ambient, source).$$

- Mass balance equations.

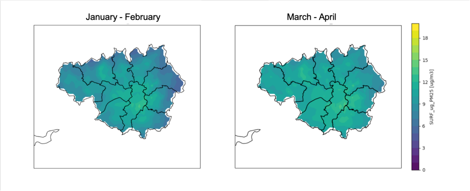


Figure: Predictions of average $PM_{2.5}$ in Manchester from EMEP for (Left) Jan-Feb and (Right) Mar-Apr.

DIMEX: Exposure Estimation

- Concentrations of $PM_{2.5}$ in indoor-not-home and transport are modelled as a function of the the ambient outdoor concentrations

$$local = f(ambient).$$

- Ambient air pollution concentrations are taken from the European Monitoring and Evaluation Programme for Transboundary Long-Range Transported Air Pollutants (EMEP) model.

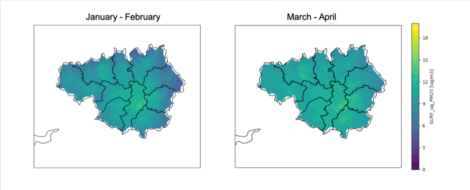
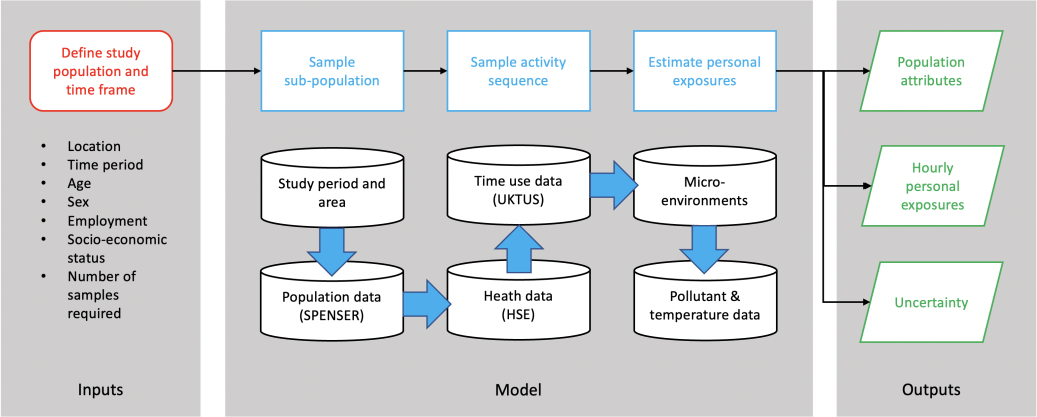


Figure: Predictions of average $PM_{2.5}$ in Manchester from EMEP for (Left) Jan-Feb and (Right) Mar-Apr.

DIMEX: Exposure estimation



Case study

- DIMEX was used to estimate personal exposures for simulated individuals in Greater Manchester
  - 2.8 million individuals
  - 10 metropolitan boroughs
  - 364 MSOAs
- Using modelled outputs from EMEP for January - March 2021.
- Sampled 100 synthetic individuals from each MSOA.

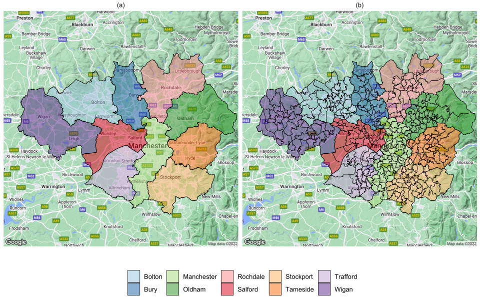


Figure: Study region of Greater Manchester split by (a) metropolitan boroughs and (b) MSOAs.

Case study:

- Difference between estimated personal exposures and ambient concentrations.
- Importance of indoor micro-environments as an exposure pathway for both sexes and all ages.
- Difference is particularly marked for the youngest and oldest age groups.
- Reflects the increased proportion of time spent at home by these population groups

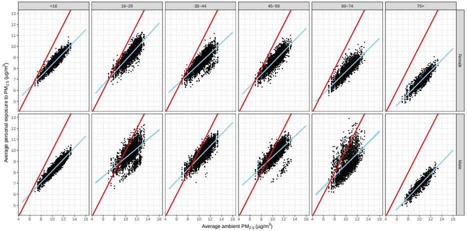


Figure: Comparison of estimated personal exposures to $PM_{2.5}$ with modelled ambient concentrations (from EMEP), see text for details). Results are presented by age group and gender. The red lines denote a 1:1 relationship (the idenity line), with the light blue lines showing the results of fitting least-squared regression to each age-gender combination.

Case study: Time Series of Exposures

- Time series plots of estimated daily personal exposures and ambient concentrations.
- Personal exposures to $PM_{2.5}$ are generally lower than the ambient concentrations.
- Potentially up to $15\mu g/m^3$ at peak days.

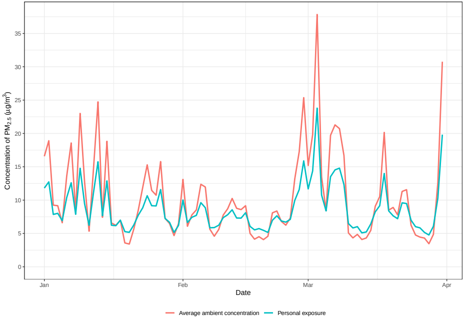


Figure: Comparison of daily estimated personal exposures to $PM_{2.5}$ using modelled ambient concentrations (from EMEP). The dotted red line denotes a difference of zero between the two series of data.

Case study: Time Series of Exposures

- Time series plots of estimated daily personal exposures and ambient concentrations.
- Personal exposures to $PM_{2.5}$ are generally lower than the ambient concentrations.
- Potentially up to $15\mu g/m^3$ at peak days.

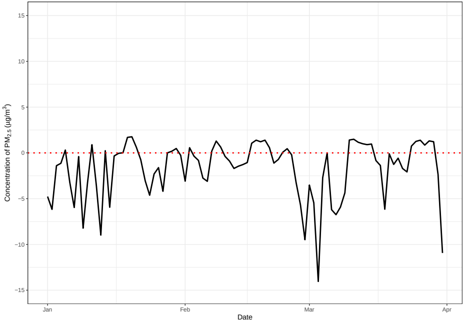


Figure: Differences between daily estimated personal exposures to $PM_{2.5}$ using modelled ambient concentrations (from EMEP). The dotted red line denotes a difference of zero between the two series of data.

CASE STUDY: SPATIAL DISTRIBUTION

- Spatial distribution of estimated personal exposures and ambient concentrations.
- Personal exposures to $PM_{2.5}$ are
  - generally lower than the ambient concentrations.
  - higher in urban areas.
- Largest differences between the personal exposures and the ambient concentrations in these urban areas.

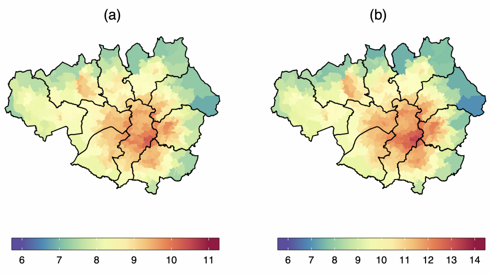


Figure: Maps of (a) estimated average personal exposures to $PM_{2.5}$, (b) modelled ambient concentrations (from EMEP).

Case study: Spatial Distribution

- Spatial distribution of estimated personal exposures and ambient concentrations.
- Personal exposures to $PM_{2.5}$ are
  - generally lower than the ambient concentrations.
  - higher in urban areas.
- Largest differences between the personal exposures and the ambient concentrations in these urban areas.

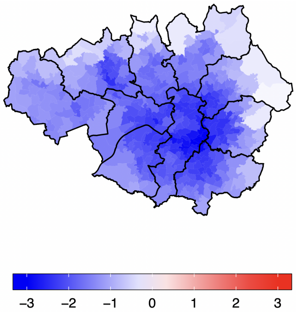


Figure: Map of the differences between daily estimated personal exposures to $PM_{2.5}$ with modelled ambient concentrations (from EMEP).

Case study: Uncertainity

- Able to produce uncertainty estimates.
- Distributions of personal exposure estimates.
- Variability in exposure across the different population groups.
- Some individuals being exposed to levels of $PM_{2.5}$ substantially above the mean/median values that would commonly be used as summary measures.

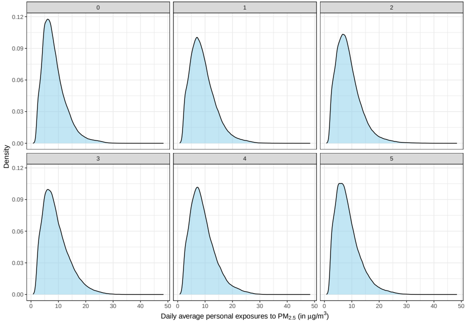


Figure: Distribution of daily estimated personal exposures to $PM_{2.5}$ using modelled ambient concentrations (from EMEP) from sampled individuals within the Manchester study area, by NSSEC socio-economic classification. 0 = Not applicable/Economically inactive, 1 = Higher managerial, administrative and professional occupations, 2 = Intermediate occupations, 3 = Small employers and own account workers, 4 = Lower supervisory and technical occupations, 5 = Semi-routine and routine occupations.

Case study: Impacts of air quality strategies

- Able to examine the impact of proposed air quality strategies and interventions.
- DIMEX run with ambient concentrations across the study area and time reduced to the recently updated WHO air quality guideline of $5\mu g/m^3$.
- Note that in some areas this might actually result in a small increase in concentrations.
- Following these new guidelines would result in a mean (estimated) reduction in personal exposures between 2.7 and $3.1\mu g/m^3$, with a median reduction between 1.8 and $2.0\mu g/m^3$.

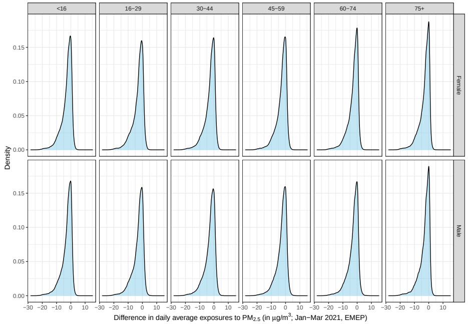


Figure: Distribution of differences between estimated personal exposures using modelled ambient concentrations (from EMEP) and with ambient concentrations set to the WHO air quality guideline of 5 $\mu g/m^3$, by age group and gender.

SUMMARY

- Open-source integrated model framework that allows
  - the estimation of personal exposures to air pollution
  - variations in exposures between different populations to be quantified
  - exposure patterns over space and time to be assessed.
- Understanding differences between the exposures experienced by individuals is crucial for informing policy decisions
  - health
  - reducing inequalities,
  - developing interventions
  - tracking progress to targets and compliance to standards.

Potential for the future

- Lots of potential for future work
  - Activity sampler
  - Micro-environments
  - Movements
  - Ambient concentrations
  - Exposure estimation
  - Computation
- Many, many potential impactful uses of the methodology.
- In conclusion, I think we have enough here to set up a digital twin.

Any Questions?