Calibrating Agent-based Exposure Simulation to Clinical Records

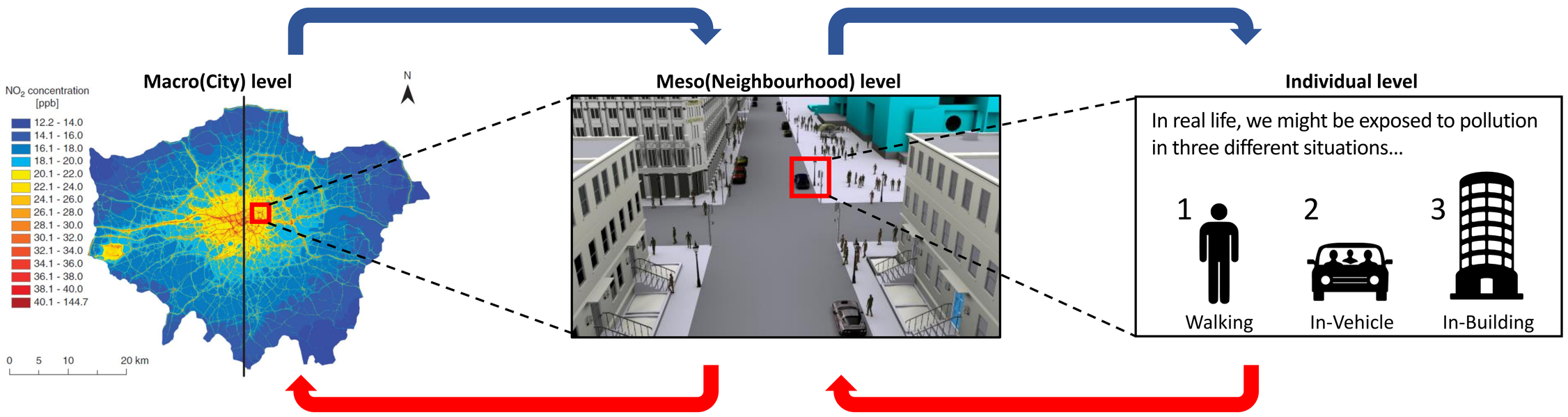
Hyesop Shin, Mike Bithell
Department of Geography, University of Cambridge

UNIVERSITY OF CAMBRIDGE

I. Presentation Scope

Modelling Health Vulnerability of Heterogeneous individuals across Seoul city

- How can we better understand the link between clinical data, air pollution and different individual behavioural patterns?



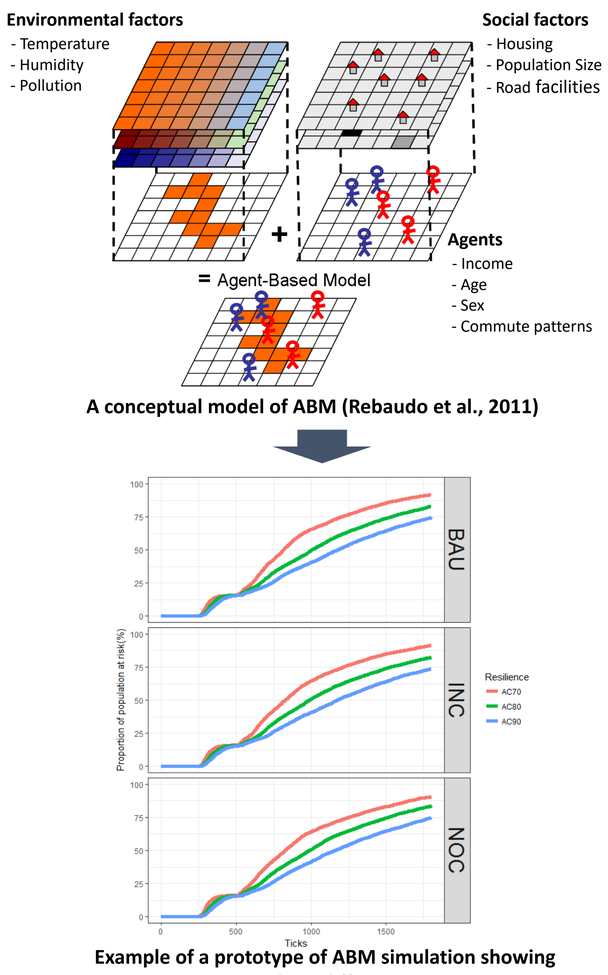
II. Background

- As megacities have increased in size and population density, steadily larger numbers of residents and commuters have been exposed to high levels of ambient air pollution, including particulates, and gaseous toxins such as NOx.
- Resulting health problems vary by pollution intensity levels, individual exposure patterns, and individual vulnerability, which may vary by age, social class, residency and other socio-economic factors.
- Resulting clinical cases may be produced by acute short term exposure, longer term low level insult or a mixture of the two, with varying patterns depending on pollutant type and time since exposure.

» **We need a spatio-temporal simulation to simulate exposure levels of heterogeneous individuals moving across heterogeneous space.**

- Our project uses *agent-based modelling (ABM),* a generic (or bottom-up) approach for simulating actions and interactions between agents, as a tool to test individual's behavioural patterns to gain a better estimation of exposure to atmospheric pollutants.
- However, there is an underlying concern in regard to uncertainty whilst performing synthetic population modelling and analysis, especially for city scale modelling
- To minimise the ranges of uncertainty, it is essential to compare actual records for calibration.
- Few validation methods are introduced in published studies (Klügel, 2008)
- Here, we present our conceptual framework of exposure simulation, and then compare different types hospital records that could calibrate our simulation results.

III. Agent-based modelling (ABM)



A conceptual model of ABM (Rebaudo et al., 2011)

Example of a prototype of ABM simulation showing exposure under 3 different scenarios

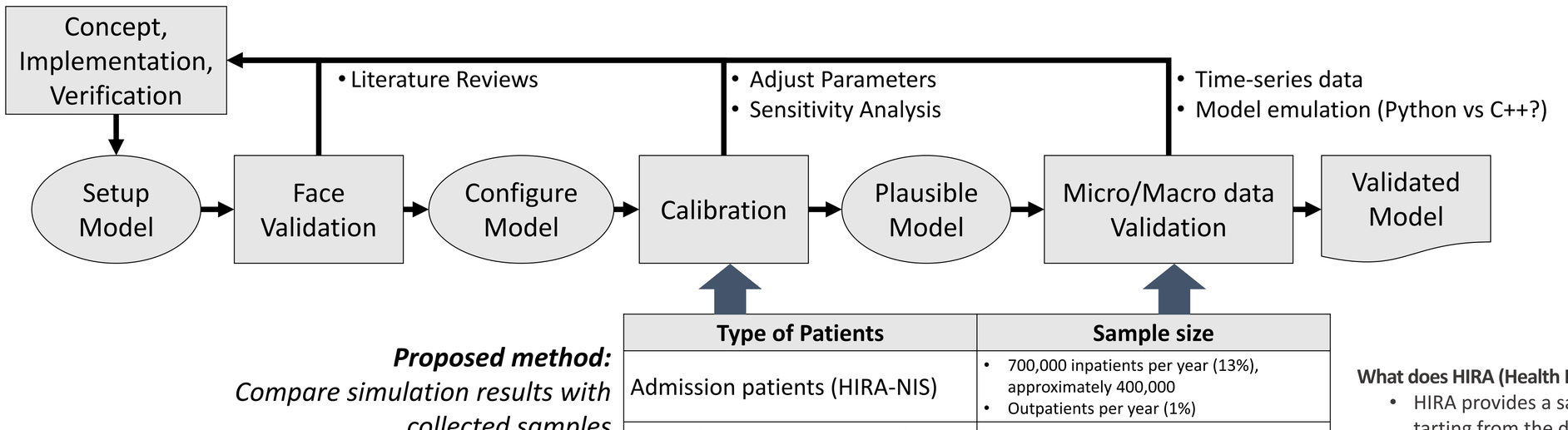
- ABM is a generic approach for simulating actions and interactions between agents to view their effects on the entire system (Gilbert, 2008).
- The artificial society is composed of autonomous, decision making entities called agents
- Individual movement resulting from daily activity allow for both simulated emissions from traffic flow and exposure to pollutants to be *jointly* estimated.
- Short term (Days to multiple months) and long term (30+ years) exposure can be de-convolved from physiological presentation of symptoms
- The social factors that exacerbate poor health can be examined independently of macro-scale pollution patterns.
- Massive artificial societies with millions of individuals can now be simulated – resulting data streams are large and complex

» **Big models can *generate* big data!**

- **Uncertainty**: ABM explain seek to explain societal phenomena as a result of individual human behaviour, but uncertain initial and boundary conditions, poorly known parameters (especially for translation of exposure into disease) and stochastic effects make this difficult.

➢ **Calibration and validation is needed to make the model realistic**

IV. Validation Framework



Proposed method:
Compare simulation results with collected samples

Type of Patients	Sample size
Admission patients (HIRA-NIS)	• 700,000 inpatients per year (13%), approximately 400,000 • Outpatients per year (1%)
Overall patients (HIRA-NPS)	• 1.4 million patients overall per year (3%)
Elderly patients (HIRA-APS)	• Approximately 1 million patients over the age of 65 per year (20%)
Infant patients (HIRA-PPS)	• Approximately 1.1 million patients under the age of 20 per year (10%)

HIRA-NIS: National Inpatient Sample
HIRA-NPS: National Patient Sample
HIRA-APS: Adult Patient Sample
HIRA-PPS: Pediatric Patient Sample

What does HIRA (Health Insurance Review and Assessment Service) do for Big Data?

- HIRA provides a sampled data of patients who have been treated for 1 year s tarting from the date of treatment

Possible enquiries

- Can we use machine learning to calibrate the model results?
- What other information or data about medical practices and experience are there that can help inform model calibration and validation?

V. Conclusion

- The project simulates pollution exposure using a large-scale artificial society
- Dealing with uncertainty is a major problem, but big data along with sophisticated models may allow for better understanding of what leads to clinical cases
- Models may suggest behavioural or policy changes that could lead to better prevention

References

- Bert, F.E. et al., 2014. Lessons from a comprehensive validation of an agent based-model: The experience of the Pampas Model of Argentinean agricultural systems. Ecological Modelling, 273, pp.284–298.
- Crooks, A., Castle, C. & Batty, M., 2008. Key challenges in agent-based modelling for geo-spatial simulation. Computers, Environment and Urban Systems, 32(6), pp.417–430.
- Gilbert, N., 2008. Using Agent-Based Models in Social Science Research. Agent-Based Models, pp.30–46.
- Klügl, F., 2008. A validation methodology for agent-based simulations. Proceedings of the 2008 ACM symposium on Applied, pp.39–43.
- Rebaudo, F. et al., 2011. Agent-based modeling of human-induced spread of invasive species in agricultural landscapes: insights from the potato moth in Ecuador. Journal of Artificial Societies and Social Simulation, 14(3), p.7.