PHI UK | Local Health and Global Profits

Protecting people, place and equity

Building a synthetic population to assess the health impacts of local climate change policies in England

Andrea Serna-Castano and Eleonora Fichera

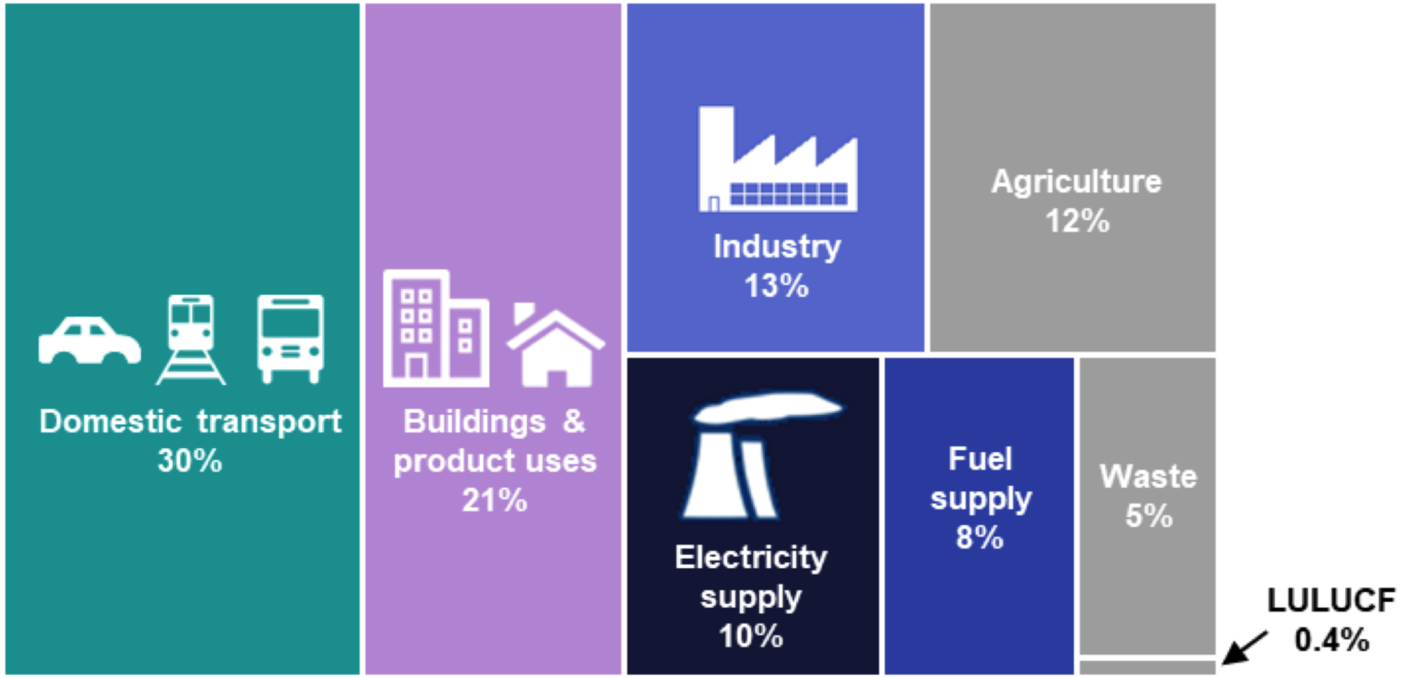
Context

- Links between environmental determinants and health
- approx. 24% of all deaths are attributable to preventable environmental risks (WHO, 2024)
- Equity dimension:

"*The top 10% wealthiest individuals contribute 6.5 times more to global warming than the average person*" (Callahan, 2025).

Context

Greenhouse gas emissions by sector UK, 2024



According to the Net Zero Strategy, 82% of emissions are within the scope of local authorities.

PHI UK | Local Health and Global Profits

Local Authority Roles and Challenges

In 49% LAs transport is the top emitter

Decision making power

Impact assessments

Long term effects

Structural barriers

Enabling Data-Driven Local Action with Synthetic Data

Why High-Resolution Population Data Matters

The Data Dilemma

Synthetic Data as a Solution

PHI UK | Local Health and Global Profits

Objective

To build a synthetic population that can serve as the basis to assess health impacts of local climate change policies

1. Determine what are the characteristics of a constraint variable that lead to a better-performing synthetic population
2. Evaluate the ability of different constraint sets to reproduce spatial patterns in health-relevant variables.
3. Compare the spatial alignment of health outcomes and environmental exposures.

Previous synthetic populations

Reference	Topic	Year	Area	Unit of analysis	Source of survey data	Variables included	Model used
Wu et al., 2022	Socioeconomic/ health	2018	Great Britain	Households individuals	Census, understanding society	sex, age, economic activity, highest educational qualifcation, marital status, ethnic group, composition of household, and housing tenure	Simulated annealing
Salat et al., 2023	Circadian activities	2020	England	Individuals	Census , Health Survey for England Time Use Survey, salary data, Trips to schools and retail, Google mobility reports, OpenStreetMap building footprint	Demographic, economic, social and health	Use of previous population SPENSER
Prédhumeau and Manley, 2023	Socioeconomic	2030	Canada	Households individuals	2016 census data and 2018 population projections	age, sex, income, education level, employment status, household size,	Synthetic reconstruction with Iterative Proportional Fitting
Ton et al., 2024	Socioeconomic/ health	2015	Global	Households individuals	Microdata from the Luxembourg Income Study (LIS) and Demographic and Health Survey (DHS)	age, education, gender, income/wealth, settlement type (urban/rural), household size, household type	Iterative proportional updating

Our Contribution

01 Constraint variable characteristics

02 Environmental exposure data

03 Focus domestic transport/ air pollution

04 Distributional health impacts

What is a synthetic population

- Synthetic populations have been around since the 1990s
- Artificially generated microdata
- Most methodologies generate microdata at the level of individuals from surveys by matching it to population-level constraints from the census
- We adopted combinatorial optimisation with simulated annealing following comparative studies (Ryan et al., 2009; Harland et al., 2012; Duran-Heras et al.,2018)

Ingredients of a synthetic population

- Geographic boundaries
- Data sources
- Constraint variables
- Method
- Outcome variables
- Validity measures

Geographic boundaries

- England
- Population over16 years old

PHI UK | Local Health and Global Profits

Data Sources

- 2021 Census for England
- Wave 14 of Understanding Society (USoc)
- Nitrogen dioxides (NO2) concentration data from Defra's Modelling of Ambient Air Quality (MAAQ)

Constraint variables

Topic	Variable	Categories
Demography	Sex	F, M
	Age	16–19, 20–24, 25–34, 35–49, 50–64, 65–74, 75–84, 85+
Ethnic group	Ethnic group	white, mixed, asian, black, other
Education	Highest level of qualification	level 4 and above, level 3, level 1 to 2, none
Labour market	Economic activity status	self-employed, in paid employment, unemployed, retired, other, looking after home, student, long term sick disabled
Health	General health	very good, good, fair, poor
Climate risk	Air pollution	Exposed, unexposed

Outline of synthetic populations

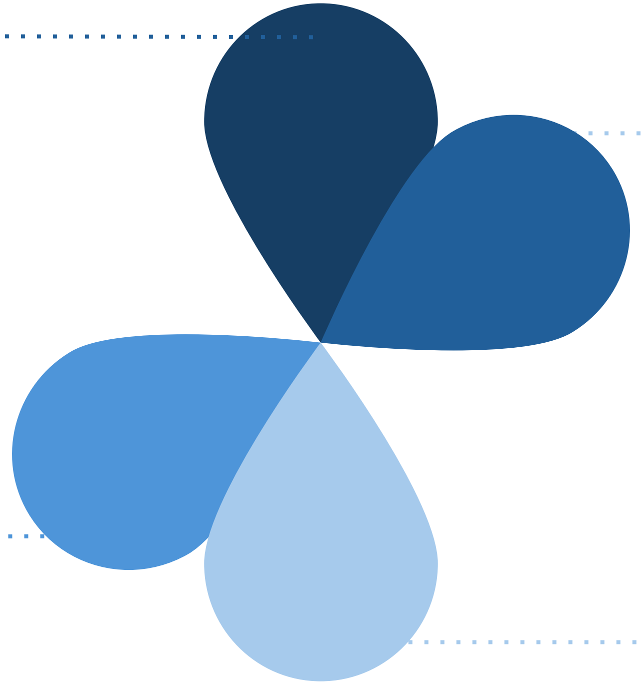
SP	Topic combination
SP1	Demography
SP2	Demography + Ethnic group
SP3	Demography + Ethnic group + Education
SP4	Demography + Ethnic group + Education + Labour market
SP5	Demography + Ethnic group + Education + Labour market + Health
SP6	Demography + Ethnic group + Education + Labour market + Health + Climate risk

PHI UK | Local Health and Global Profits

Method

Flexible Modelling Framework

(Harland et al., 2013)



Sampling with replacement

Individuals to be cloned multiple times, supporting flexibility in matching area-specific population sizes.

Goodness-of-Fit Metric

Overall Relative Sum of Squared Z-scores (Voas & Williamson, 2001)

Iterative Swapping Procedure

to improve the fit, retaining changes only if they enhance alignment with constraint distributions.

Outcome variables

Variable	Question label
hcond1	Diagnosed health conditions: Asthma
hcond11	Diagnosed health conditions: Chronic bronchitis
hcond13	Diagnosed health conditions: Cancer or malignancy
hcond14	Diagnosed health conditions: Diabetes
hcond21	Diagnosed health conditions: COPD (Chronic Obstructive Pulmonary Disease)
hcond27	Diagnosed health conditions: Lung cancer
hcond3	Diagnosed health conditions: Congestive heart failure
hcond4	Diagnosed health conditions: Coronary heart disease
hcond5	Diagnosed health conditions: Angina
hcond6	Diagnosed health conditions: Heart attack or myocardial infarction
hcond7	Diagnosed health conditions: Stroke

Validity measures

Internal validity:

Percentage of classification error (Duran-Heras et al., 2018)

$$\%CE_i^{M-dim} = \frac{\sum_{c_1=1\dots}^{C_1} \sum_{cM=1}^{CM} \left| o_{ic_1\dots c_M} - e_{ic_1\dots c_M} \right|}{2N_i} * 100$$

External validity:

Compare the value of the outcome variables which are the health conditions with respect external datasets e.g. the asthma prevalence at the LSOA level

Results :

Mean percentage of classification error

SP	Survey sample	Mean % CE Included variables	Mean % CE All variables
SP1	28086	6.05	16.45
SP2	27930	6.98	14.68
SP3	27056	5.53	12.25
SP4	26964	4.85	10.86
SP5	26487	5.44	10.09
SP6	26487	4.66	4.66

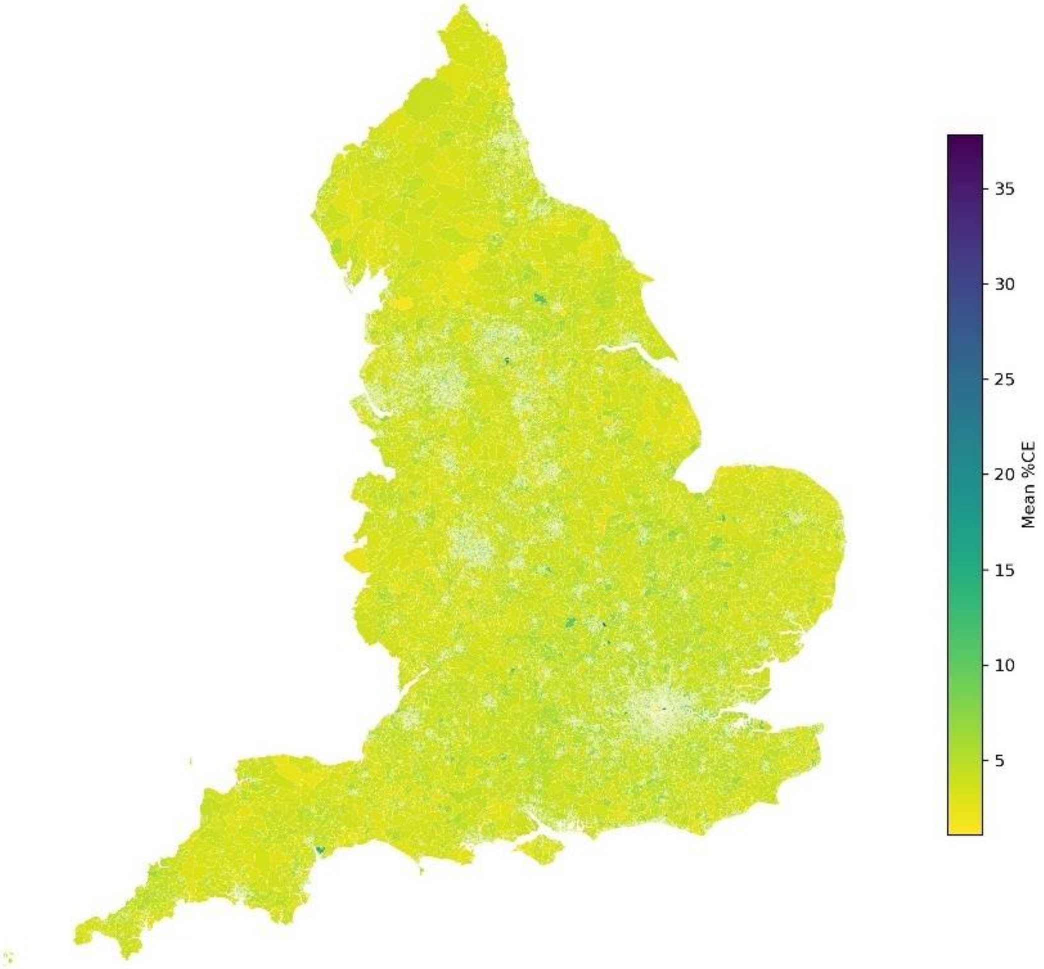
Results :

Mean percentage of classification error by variable

Variable	SP1	SP2	SP3	SP4	SP5	SP6	% Change
Sex	8.85	8.85	8.85	8.85	8.85	8.85	0%
Age	3.23	3.23	3.23	3.23	3.23	3.23	0%
Ethnic group	19.52	8.84	8.84	8.84	8.84	8.84	-54.71%
Highest level of qualification	18.91	19.04	1.18	1.18	1.18	1.18	-93.80%
Economic activity status	11.66	11.48	11.13	2.15	2.15	2.15	-80.68%
General health	13.29	13.05	14.52	13.81	8.35	8.35	-39.53%
Air pollution	39.67	38.30	37.95	37.97	38.04	0.00	-100%

Results :

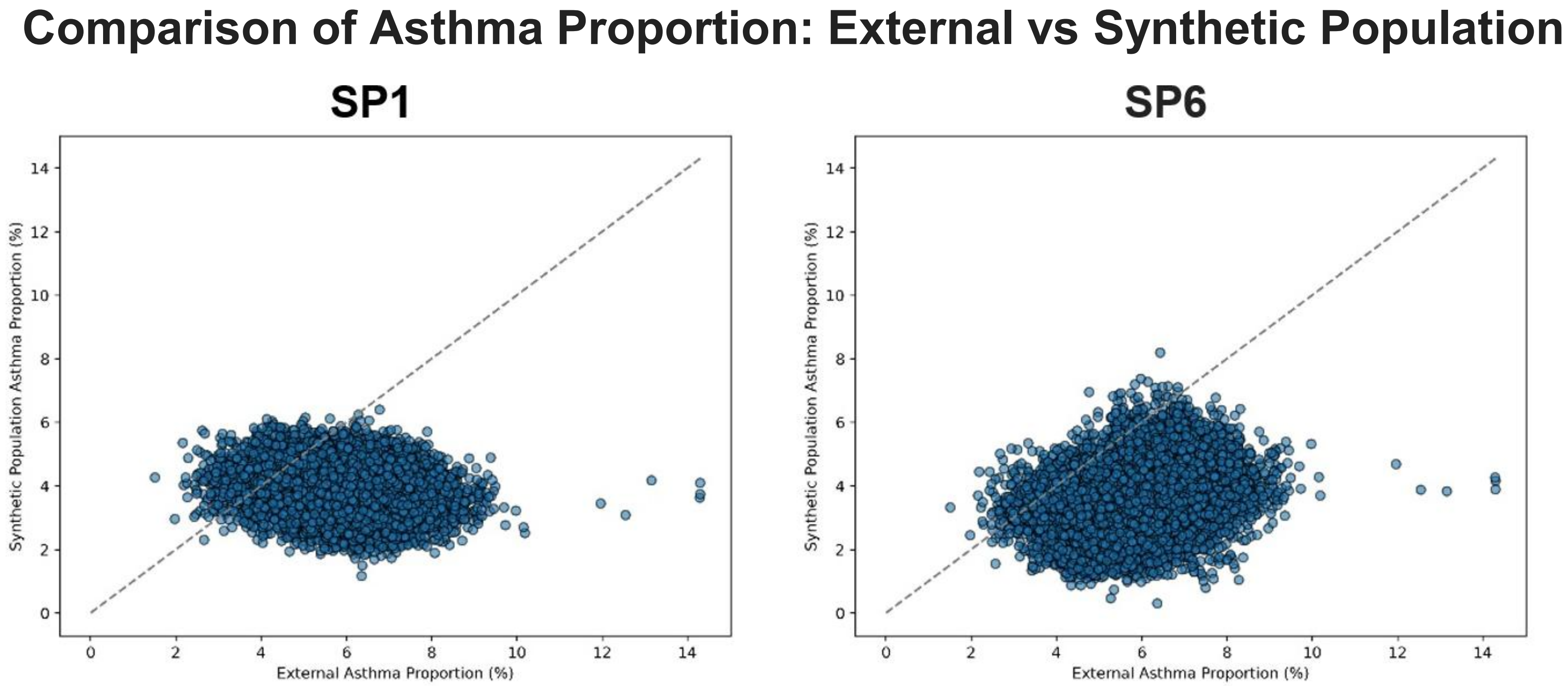
Map of mean %CE for SP6 by LSOA



PHI UK | Local Health and Global Profits

External validity

Rose et al. (2021)



Conclusion

- Gap in literature: constraint variable selection is often based on data availability
- **Internal validity**
  - improves with sequential constraint addition, greatest gains from education and economic activity. General health slightly reduced model fit.
  - Adding constraints did not reduce accuracy of earlier variables.
- **External Validity**
  - SP6 (most detailed model) improved fit across LSOAs, reducing high-error zones.
  - Asthma estimates saw only marginal external improvement Suggests need for better-aligned health constraints or richer health data.

Next steps

- Add more variables including environmental risks (e.g. flooding, heatwaves).
- Strengthen validation using linked datasets with joint health distributions (e.g. Health Episode Statistics).
- Integrate into dynamic microsimulation for long-term policy evaluation.
- Goal: support robust, equitable, and policy-relevant climate-health assessments.

Acknowledgements

This work was supported by UK Research and Innovation funding for "Local Health and Global Profits" (Grant no MR/Y030753/1) which is part of Population Health Improvement UK (PHI UK).

UK Research and Innovation

PHI UK | Local Health and Global Profits

Protecting people, place and equity

Thank you