UNIVERSITY OF CAMBRIDGE

SIMULATING PEDESTRIAN'S LONG-TERM EXPOSURE TO AIR POLLUTION

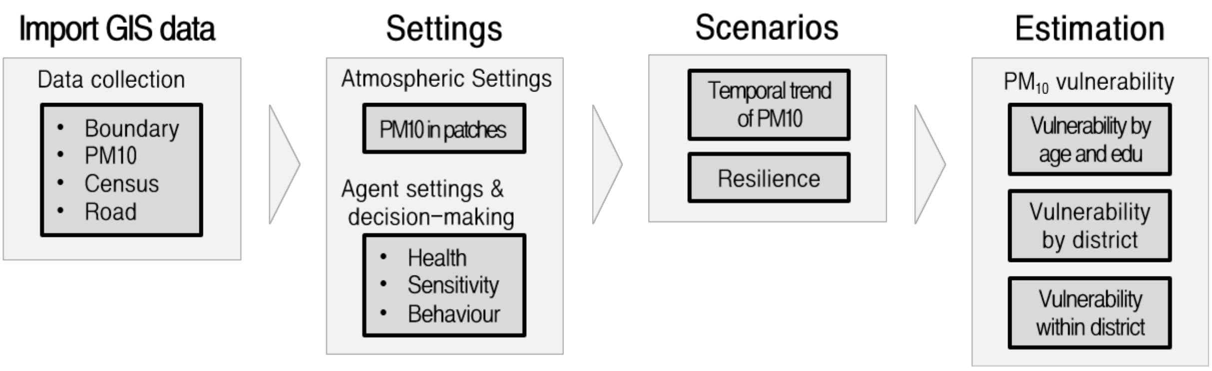
First Developed: February 2017

Latest Edit: October 2018

Developed by Hyesop Shin and reviewed by Mike Bithell

1. Model Overview

Our model was implemented in NetLogo 6.0.4, each district being run as a separate self-contained model: this means that some agents with home locations within a district but work destinations in another district had to be excluded from consideration. To simulate our model, we used 'RNetLogo' in R 3.4.4. implement 9 scenarios for 100 iterations each.



1.1. Purpose

The model objective is to understand the cumulative effects on population vulnerability due to its exposure to PM_{10} by different age and educational groups in both Gangnam (wealthy) and Gwanak (deprived) districts.

1.2. Entities, State Variables and Scales

There are 3 types of agents in the model based on age. Agent's coloured in sky blue age under 15, mustard yellow between 15 and 64, and grey over 65. An agent's birth and death are not considered. An agent's education is either given as 'well-educated' or 'less-educated' based on the education statistics by sub-district level in the 2010 census.

We used a 1% population sample of agents (individuals) in each district to generate a simple synthetic population in Gwanak district and Gangnam district. The idea was made for the sake of speed in this initial investigation. For example, the total population of *Boramae* (one of the sub-district in Gwanak) was 24,351 persons of which 2,219 (9%) were under 15 (excluding infants under 5), 19,520 (80%) were between 15 and 64, and 2,612 (11%) were 65 or higher. In the model, each group were converted to 22, 195, and 26 people respectively. Hence, the total population of Gangnam was converted to 5050 agents, and Gwanak to 4915 agents.

Table 1 2010 Census in Gwanak sub-districts (excluding people under 5 years old assuming they cannot mobilise themselves)

Name	Population	Name	Population	Name	Population
Boramae	24,351	Jowon	15,630	Samseong	27,070
Cheongnim	15,350	Jungang	13,484	Seowon	23,635
Cheongryong	30,851	Miseong	31,765	Seowon	22,797
Daehak	24,775	Nakseongdae	20,444	Shillim	18,428
Euncheon	33,095	Namheon	16,627	Shinsa	23,664
Haengwoon	27,709	Nangok	28,496	Shinsa	23,664
Inheon	26,805	Nanhyang	16,004	Shinwon	18,418
				Sunghyeon	32,156

Table 2 2010 Census in Gangnam sub-districts (excluding age under 5, it also has the same name called Shinsa but the Chinese characters are different)

Name	Population	Name	Population	Name	Population
Apgujeong	24,630	Gaepo1	20,998	Nonhyun2	21,276
Cheongdam	26,567	Gaepo2	33,137	Samseong1	13,927
Daechi1	23,677	Gaepo4	21,757	Samseong2	28,430
Daechi2	39,999	Ilwon1	17,136	Segok	4,023
Daechi4	20,615	Ilwon2	17,814	Shinsa	17,792
Dogok1	19,311	Ilwonbon	20,977	Suseo	17,792
Dogok2	29,319	Nonhyun1	23,983	Yeoksam1	16,169
				Yeoksam2	32,534

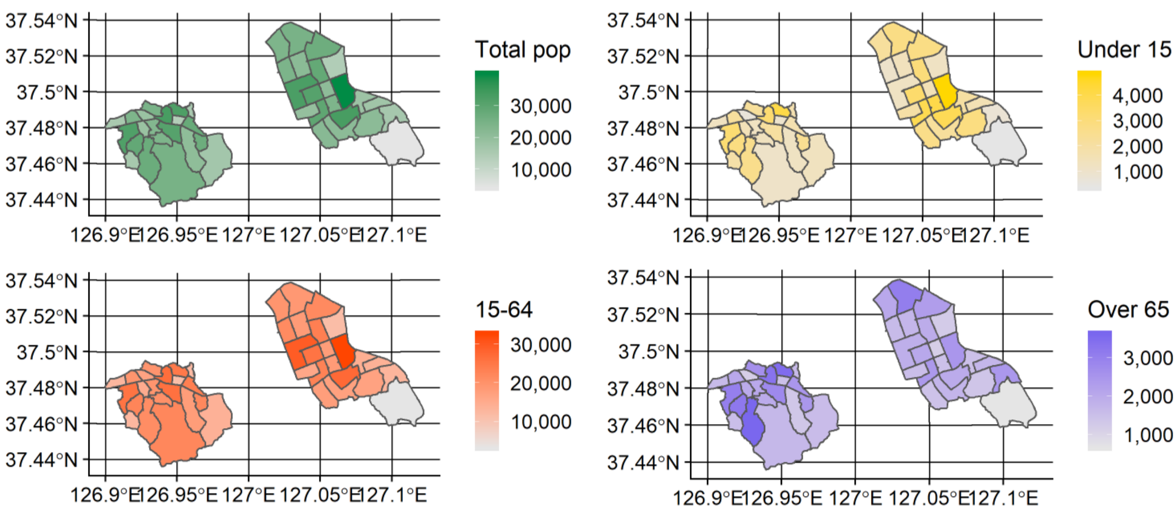


Figure 1 Population distribution in Gwanak (Left) and Gangnam (Right) by total population and age groups.

Each agent has a list of attributes that include the name of home location, coordinates of home location, name of destination, coordinates of destination, age group (i.e. young, active, old), education level, and health. Every agent will move to their given destination location during working hours and come back home again on alternate model ticks (see more details in Section 2.2).

We assumed that the agents have not experienced any previous exposure, as we lack either health records or exposure histories for the population. Our simulation is thus only able to track the likely rate at which exposure effects accumulate over time in a given district relative to others. Depending on the way this simulated history works out, we hope to then be able to estimate the accumulated risk that an unexposed population would have developed over the period of the simulation. Each agent is assigned a nominal "health" level, which is an integer with initial value 300. Depending on their socioeconomic status, they will lose health when exposed to a PM_{10} of over a threshold near 100µg/m^3. We choose this as it is a Korean hourly ambient air quality standard (https://www.airkorea.or.kr/eng/information/airQualityStandards), although lack of data on how this relates to disease means that the relation to actual health impact is not entirely clear. We assume that, consistent with the idea of an air quality standard, the adverse effects on health only begin to operate when pollution is above this threshold. More detail on this is given below.

1.3. Process overview and scheduling

Figure 3 shows the conceptual outline of the model. During the setup process, every agent is assigned a fixed home name (sub-district) and home patch as well as their destination name and patch. Destination names and patches differ by age group. Agents between *age 15 and 65*, also known as the economically active population, will move to their destination patches according to the fraction moving to a given destination in the origin-destination (OD) matrix. Those who commute to other districts are allocated to dummy patches outside the district during working hours – these agents are not included in the overall statistics, as we do not have data for them during the day.

Agents *aged under 15* will move to a random patch within radius 3, while those *aged over 65* will move to a random patch within radius 1: this is intended to represent a more restricted range of movement for this fraction of the population. As we do not currently model the traffic flow, we simplify movement by translating the agents to their destination patch during the day (1 tick) and back to the home patch at night (next tick). Every agent will start without any medical history, and with full health: all are susceptible to local levels of PM_{10} within their current patch. An agent will lose health when exposed to a patch exceeding 100µg/m^3 of PM_{10}. Agents in the youngest or oldest age groups, and those with lower educational levels receive and additional penalty. Recovery can occur according to local property values up to a maximum adaptive capacity (see details in Section 2.4).

As the agents continue to lose health from high PM_{10}, we assume that their health status can change suddenly when they cross a given threshold. Those with health less than 2/3 of the starting value change status to 'highly-exposed', and those with health less than 1/3 of the initial value are labelled at risk'. The model terminates either if all agents turn to 'at risk' status or if it reaches the end of year 12 i.e. tick 8764.

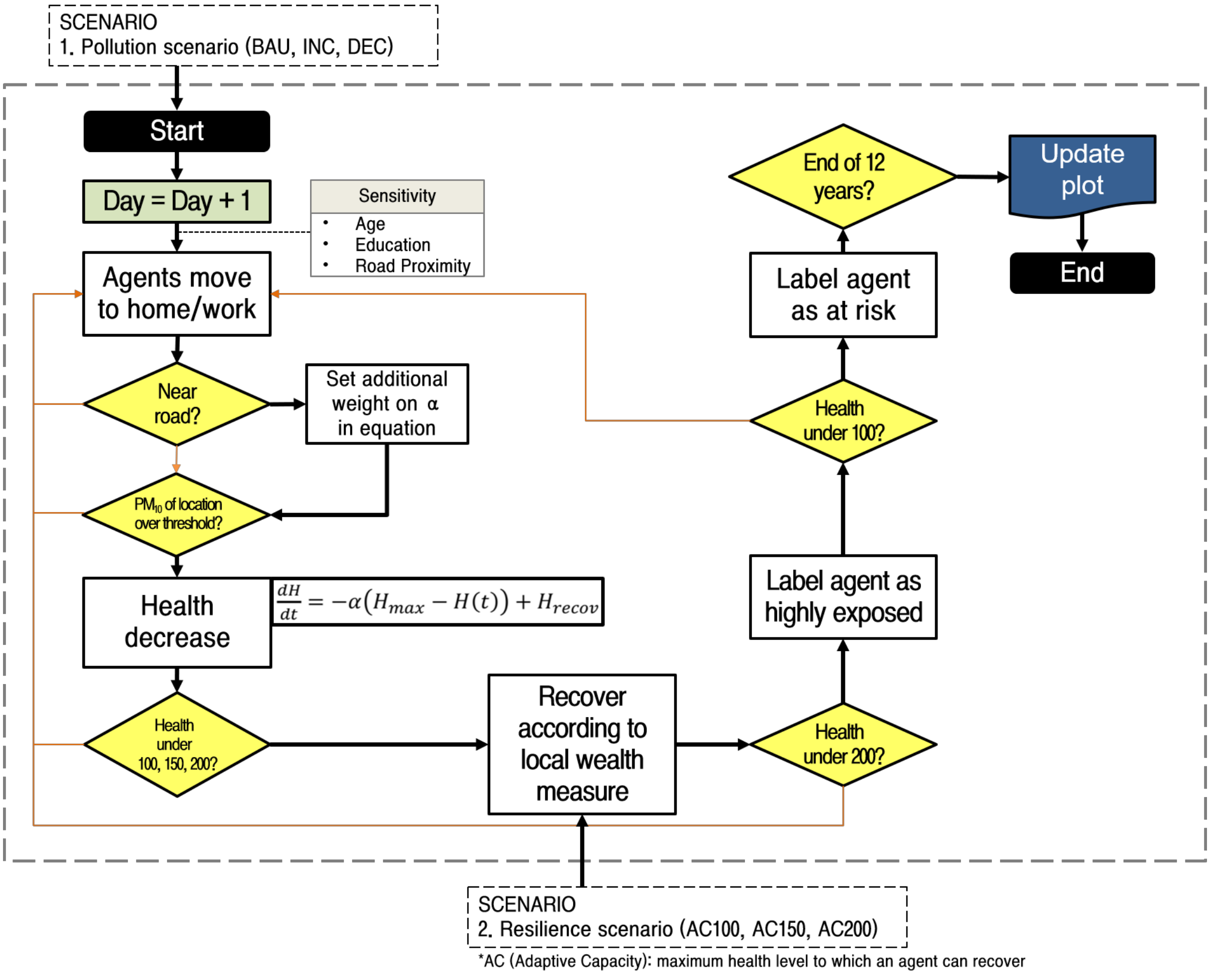


Figure 2 Implementation algorithm for each period

2. Design Concepts

2.1. Theoretical and Empirical Background

The cognitive behaviour of pedestrians in this model stems from the Urban-suite Pollution model introduced in the NetLogo library (Felsen & Wilensky 2007). The general idea of the Urban-suite pollution model was to examine the competition between the predators (pollution) and prey (people's reproduction) in an enclosed landscape. This model gave an insight that an agent who is close to a harmful pollution patch (no matter what the pollutant type is!) inhales the pollutant and the potential body status gets harmed.

There are two hypotheses in our model. First, the agents start to lose health when the PM_{10} exceeds 100µg/m^3. The number was designated by the "Unhealthy" standard provided by the Korean Ministry of Environment (KME). Second, individual behaviour and socioeconomic status can affect and be affected by pollution exposure. For instance, residents near polluted areas are more likely to be exposed, as well as to be socially deprived (David & Don 2012; Kan & Chen 2004; Kan et al. 2008; O'Neill et al. 2003)

Even though the universal pollution standard exists, people's exposure to pollution differs to which geographical space they tread on as well as their physical status e.g. age, education, immunity. Also, the metaphor of uncertainty is applied to the agent's random movement. From the complexity point of view, this model allows us to better understand the rise of vulnerable population in a more realistic way.

2.2. Pedestrian's decision-making process

The agents in the model have little by way of reasoning ability. Behaviour is entirely driven by a simple schedule.

Agents, in common, follow the hypotheses mentioned as follows:

- An agent's birth, death, and ageing are not considered
- Agents have minimal cognitive representation, but understand their area boundaries
- 1 tick is equivalent to half a day i.e. working and home hours
- Every agent starts with a health status of 300, but this drops when an agent is exposed to pollution
- Agents commute to the same location until the simulation ends
- For visualisation purposes, if the health status of an agent drops below 200, the agent colour turns to purple, when it drops below 100, the colour then turns to red
- If an agent's health reaches 0, they will be sent to hospital for remedy
- All agents stop if the system reaches 8764 ticks (equivalent to 12 years), or the 'at risk' population reaches 100% of the total population

Given the lack of address information, every agent was randomly distributed in their home location based on the census. Agents aged 15-64 commute within a sub-district, but also can move to different sub-districts derived from the Origin-Destination matrix. Movement is random, within these constraints, but with an additional tendency to randomly select patches with roads within permitted sub-districts as a way of approximating transportation in the absence of a full traffic model. By contrast, movement range of the young and elderly are restricted close to their origin. This was based on the idea that most children go to schools and do after-school activities near home, and the elderly travel to distant places less frequently. We did not consider agents who commute to different districts outside the study area, or the mode of transport for trips (so there is not accounting for differences between pedestrians, car travel, or buses, or for the fact that pollution levels may be lower within buildings). The effect of roads proves to be significant in determining relative outcomes between groups.

2.3. Interaction

Since the diseases resulting from pollution exposure are non-communicable, the model does not consider any transmission effects between individuals, but considers the effects resulting from continuous interactions between individual spatial trajectories, constrained by daily activity patterns, with the measured spatial macro-scale atmospheric pollution distributions.

2.4. Heterogeneity

All agents are heterogeneous in their personal profile: home district and location, destinations, and socioeconomic status i.e. age and education. For this study, we attempted to classify agent's PM_{10} exposure by including a sensitivity factor. This can work as a modifier of the adverse factors degrading the agent's health and use the land price as a recovery factor.

Sensitivity is composed of two socioeconomic attributes of an agent, which control the degree of health loss. One determinant is age, which works as a proxy of physical health variance, assuming that very young or very old population members are more likely to suffer when exposed to pollutants. The other is educational level, designed to take into account possible lack of pollution awareness amongst those with lower educational depth. By contrast, land price is a proxy of health recovery based on the idea that medical facilities (e.g. hospitals, clinics, pharmacies) have greater chances to be located in areas with higher property values, or that higher values may mean greater ability to avoid polluted air, as mentioned above. The model does not consider any transmission effects between individuals or environments because the diseases resulting from pollution exposure are non-communicable, although in practice compromised pulmonary systems may make agents with high exposure more likely to suffer ill effects from viral or bacterial diseases.

2.5. Stochasticity

In this model, we set a stochastic allocation of PM_{10} for each patch because each district had only one station to capture the pollution level (see details in section 3.4.1). This gives a spatial variation of pollution across the heterogeneous space during the simulation as well as in each iteration. This stochasticity affects the agent's exposure level regardless of their movement distance. In other words, agents aged under 15 and over 65 who are given a randomness within small movement range (from home to a shop nearby), will have a high pollution exposure. Finally, assuming each person has different susceptibility to pollution, agents have a small randomness to their health loss when they are exposed to over 100ug/m^3 of PM_{10}.

2.6. Observation

The model environment was derived from a GIS dataset of each district. For simplicity, we excluded building and traffic information for the present. The spatial extent of Gangnam and Gwanak are around 40km^2 and 30km^2 respectively, with a 30m by 30m spatial resolution. We used two-time steps (home hours, working hours) per day, and simulated 12 years, of which the earlier 6 years of PM_{10} was from the observational dataset. For the second 6 years we re-used the first six years again but modified according to create scenarios for future projection.

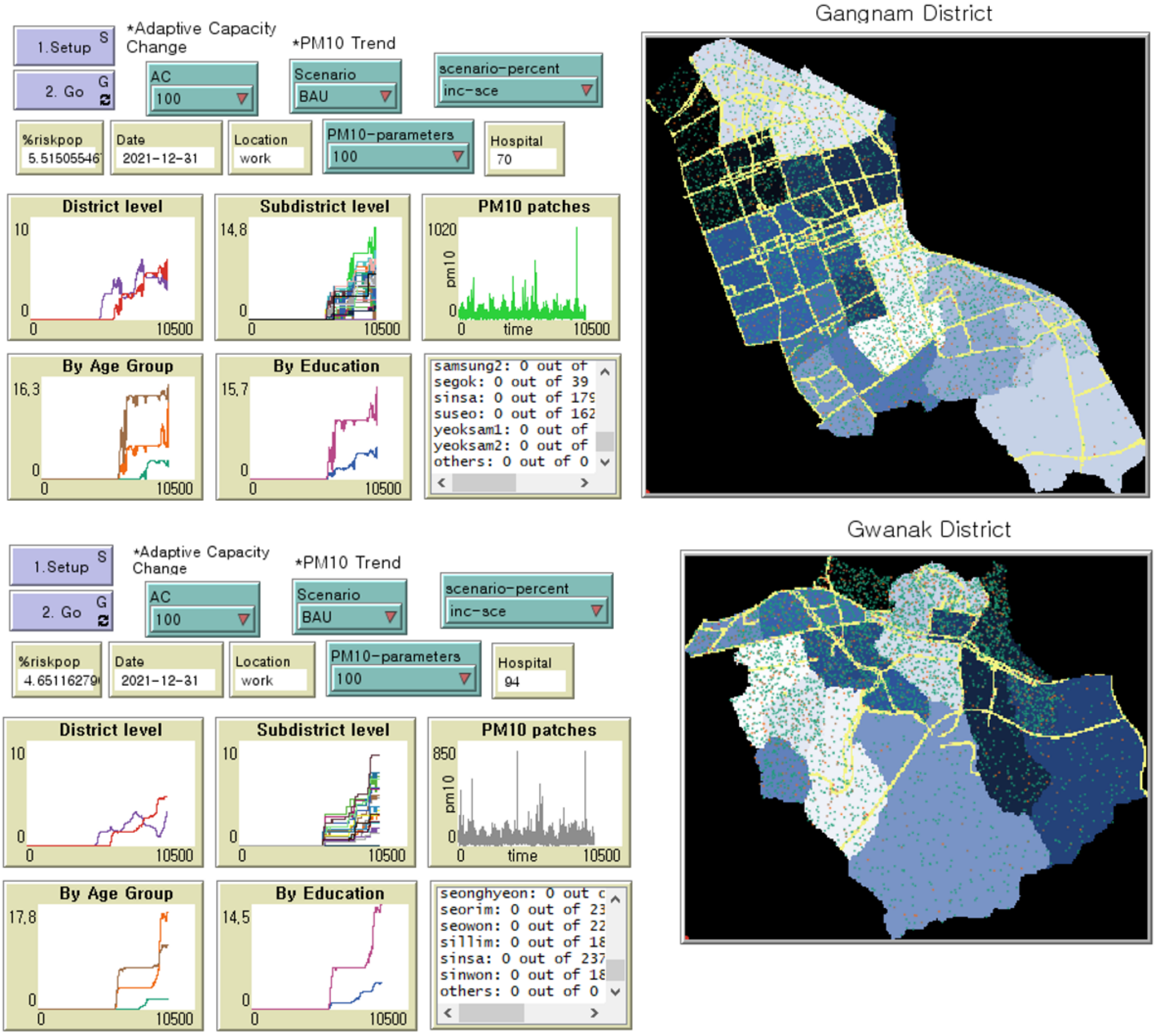


Figure 3 Implementation in Netlogo 6.0.4. for Gangnam (top) and Gwanak (bottom)

3. Detail

3.1. Implementation Details

The model was coded in NetLogo 6.0.4 (Wilensky, 1999), as illustrated in Figure 3. Agents move around the domain and are exposed to local levels of pollution in each tick. The agents' response to the exposure is broken down into a sensitivity, which determines their decrease in health on exposure, and a resilience, which allows their health to recover up to a maximum adaptive capacity.

3.2. Initialisation

During the setup process, the model is busily allocating agent's profile as well as pollution and land price levels. Figure 3 illustrates the initial state of the *in silico* world, which visualises the administrative distribution, roads, and agents coloured in their age groups. Agent's profile and PM_{10} patches differ in each setup, but the land price remains equivalent.

We attempted to investigate various ways to quantify an individual's health, however could not secure any data for a population-wide study. As an alternative, we used some arbitrary values that similarly mimics an individual's health loss. Firstly, each agent will have a health of 300 in the beginning of the simulation (assuming agents have no previous exposure experienced). Secondly, the agents will lose health based on an arbitrary factor if they are exposed to over 100ug/m3. The factors depend on an agent's socioeconomic status and road proximity.

3.3. Input

Table 1 provides an overview of the variables used in this initial model. We collected station data, individual personal attributes, road layouts, land prices and an origin-destination matrix for daily travel.

Table 1 List of Variables

Components	Variable	Value	References
Pollution	PM_{10}	Numeric	National Institute of Environmental Research
Sensitivity	Age	Numeric	Korean Statistical Integrated System (KOSIS)
	Education	Dichotomous	
	Road proximity	Dichotomous	Intelligent Transportation Systems (ITS)
	Official land price	Grade	National Spatial Data Integration (NSDI)
Movement	Origin-Destination matrix	Numeric	Korea Transport Database (KTDB)

We chose PM_{10} as a representative pollutant not only because of the data availability for a long period, but also of the growing emphasis of PM_{10} impacts as 1st class carcinogen that can cause pulmonary disease, cardiovascular problems, and stroke (IARC 2013; Kan et al. 2008; Lee, Hwang, & Kim 2014; Loomis et al. 2013).

Hourly PM_{10} was collected from Gangnam and Gwanak stations between 1st of January 2010 and 31st of December 2015, then grouped by home hours (assumed to be 20:00 - 08:00) and working hours (09:00 - 19:00). To account for the gaps in the time series (1129 out of 52584 hours (2.15%) in Gangnam and 785 hours (1.49%) in Gwanak), we inferred the missing values using a Kalman algorithm for each season from a "ImputeTS" package in R (https://cran.r-project.org/web/packages/imputeTS/index.html).

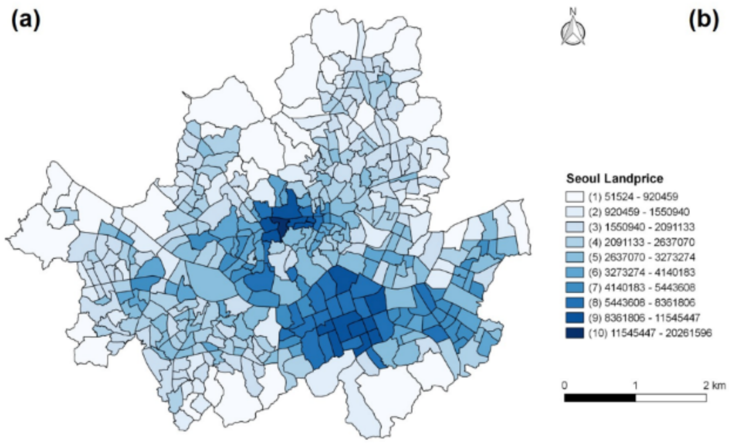
At this stage of the model development we do not track agent movement along the road system. However, we include a measure of road proximity as a means of demonstrating the negative impact of enhanced emissions near roadsides, irrespective of social categories. The data was retrieved from the Intelligent Transport Systems (ITS, http://its.go.kr/) institution run by the Korean Ministry of Land and Transport.

For this preliminary research, age and education were the characteristics that featured in pollution sensitivity. Age was included because different age groups presumably have distinctive differences in their physical health status and sensitivity to pollution. The raw data was retrieved from the 2010 Census in the Korean Statistical Office (http://kosis.kr/), which was grouped in every five years (e.g. 0-4, 5-9, 10-14…). We aggregated the groups into three: 0-15, 15-64, and 65+. Education level was included as representing the awareness of potential harm from pollution. The original education data from the 2010 census was provided in 8 categories, but for simplicity, was aggregated into two groups: above and below middle school graduation.

Official land price was selected as an indicator of ability to recover from pollution impacts. Coffee et al. (2013) attempted to show a relationship between residential properties with high socioeconomic status and health as part of a study of housing properties and social wellbeing. The reason of using residential property was not only because it was immovable and location specific capital, but also for its expression of price in the economy: those more able to afford higher priced housing are also perhaps better able to access health care, or to adapt their life-stlyes to compensate for high pollution levels. Here, we used official land prices as a proxy for regional socioeconomic status. 2015 land price were retrieved from the National Spatial Data Integration website (NSDI, http://www.nsdi.go.kr/).

The raw data were given in tables of prices by census block units. Using the Natural Breaks method, we aggregated block units to sub-district level, then categorised the price distribution into 10 sub-divisions (see Figure 2(a)). Figure 2(b) indicates hypothetical health changes determined by land price. It is assumed that agent's nominal health level can recover up to a maximum adaptive capacity level at a rate which is higher for those living in areas with higher property values.

(a)



(b)

Class	Land Price (in Korean Won)	Health
1	< 920,459	+ 0.01
2	920,459 - 1,550,940	+ 0.01
3	1,550,940 - 2,091,133	+ 0.02
4	2,091,133 - 2,637,070	+ 0.04
5	2,637,070 - 3,273,274	+ 0.05
6	3,273,274 - 4,140,183	+ 0.05
7	4,140,183 - 5,443,608	+ 0.06
8	5,443,608 - 8,361,806	+ 0.07
9	8,361,806 - 11,545,447	+ 0.12
10	11,545,447 - 20,261,596	+ 0.15

Figure 2 2015 official land price map of Seoul in sub-district scale (a), and table of hypothetical health changes determined by land price in Korean currency (£1 ≈ ₩1,500) (b)

Origin-destination matrices for Gangnam and Gwanak were downloaded from the Korean Transportation Database (https://www.ktdb.go.kr) and converted to fractions so that we can allocate home locations and associated work locations to a sub-sample of the total population (see tables below).

UNIVERSITY OF CAMBRIDGE

Gangnam O/D Matrix

Origin	Destination: Apgujeong	Cheongdam	Daechi 1(il)	Daechi 2(i)	Daechi 4(sa)	Dogok 1(il)	Dogok 2(i)	Gaepo 1(il)	Gaepo 2(i)	Gaepo 4(sa)	Irwon 1(il)	Irwon 2(i)	Irwonbon	Nonhyeon 1(il)	Nonhyeon 2(i)	Samseong 1(il)	Samseong 2(i)	Segok	Sinsa	Suseo	Yeoksam 1(il)	Yeoksam 2(i)	Other	Total
Apgujeong	12.4	7	0.3	1.6	0.8	0.4	2.7	0.1	0.2	0.2	0.1	0.4	0.7	2.2	1.7	1.4	2.6	0.4	7	0.4	2	1.6	53.9	100
Cheongdam	6.4	19.9	1.4	1.3	0.8	0.8	0.6	0.1	0.7	0.1	0.2	1.3	0.2	4.6	3	5.9	8.5	0.1	2	0.2	1.9	1.2	38.8	100
Daechi 1(il)	0.7	1.1	8.7	12.5	7.8	2.4	2.9	1	2.2	0.6	0.5	2.2	0.9	0.5	0.5	3.9	2.6	0.3	0.9	0.5	2.3	2.7	42.3	100
Daechi 2(i)	0.9	1.3	6.8	16.6	8.3	2.7	2.6	0.6	2.3	4.1	0.9	4.4	0.9	0.4	0.4	2.4	3.2	0.2	0.6	0.5	4.2	3.2	32.5	100
Daechi 4(sa)	1.2	2.6	4.7	11.4	10.6	2.4	2.6	0.8	1.2	1.7	0.7	2.5	0.7	0.8	1.2	2.3	4.6	0.1	1.5	1	4.2	2.2	39	100
Dogok 1(il)	1	1	1.7	3.1	2.5	13.6	14.5	1.5	0.7	1.5	0.2	1.6	0.5	0.9	0.4	1.2	2.7	0.2	0.8	0.8	7	3.9	38.7	100
Dogok 2(i)	3.6	0.7	1.7	3.3	2.6	11.9	11.9	2.4	4.7	2.1	0.4	2.3	0.9	0.9	0.8	1.4	1.8	0	1.1	0.4	2.7	2.8	39.6	100
Gaepo 1(il)	0.6	0.8	1.1	3.4	2	3.2	8.9	10.7	7.6	7.3	1.4	4.1	1.4	0.7	0.4	1.6	4	0	1.5	0.7	1.4	1.3	35.9	100
Gaepo 2(i)	0.5	0.8	1	3.2	2	1.1	5.7	4.5	14.2	5.8	6	5.5	6.5	0.5	0.7	2.6	3.2	0.9	0.2	2.8	1.6	1	29.7	100
Gaepo 4(sa)	0.3	0.5	0.9	6.5	3.1	3.1	6.5	6.4	10	12.3	1.5	4.2	0.6	4	0.4	0.4	2.7	1	0.9	0.5	1.3	1.9	31	100
Irwon 1(il)	0.6	0.4	0.4	2.7	1.4	0.6	1.6	0.9	10.4	1.3	9	7	12	0.6	0.2	5.1	4.5	1.1	0.3	4	1.2	0.4	34.3	100
Irwon 2(i)	0.6	0.7	3.1	8.3	3	2	2.9	3.2	6.9	5.3	5.4	5.7	6.1	0.7	0.3	1.1	1.6	0.6	1	1.4	1.9	1.3	36.9	100
Irwonbon	1.3	0.3	0.8	2.1	1.8	0.6	1.3	0.5	5	0.6	6.3	3.5	11.8	0.6	0.3	10.2	1.6	0.7	0.5	4.2	2.4	0.5	43.1	100
Nonhyeon 1(il)	2	3.9	0.2	0.4	0.5	0.9	0.3	0.2	0.1	1.3	0.1	0.3	0.4	8.4	6.5	1.6	2.3	0.1	2.9	0.1	4.5	1.6	61.4	100
Nonhyeon 2(i)	1.8	2.3	0.2	0.2	0.4	0.7	0.4	0.1	0.3	0.2	0.1	0.1	0.2	6.6	8.7	1.4	1.6	0.2	2.1	0.2	4.8	1.6	65.8	100
Samseong 1(il)	1.6	2.6	1.4	1.4	0.9	1.2	0.5	0.2	0.9	0.2	1.3	0.5	2.1	1.4	1	11.9	9.1	0.3	0.8	0.2	3.9	1	55.6	100
Samseong 2(i)	1	2.9	0.8	1.4	2	0.8	0.5	0.5	1	0.4	0.3	0.4	0.6	1.3	1.2	8.8	10.6	0	0.8	0.2	5	1.6	57.9	100
Segok	1	0.6	0.6	0.4	0.5	0.2	0.2	0.1	1.2	0.5	0.9	0.7	0.9	0.3	0.5	0.8	0.2	52.1	0.9	11.7	1.9	0.2	23.6	100
Sinsa	5.6	1.5	0.3	0.2	0.8	0.2	0.7	0.2	0.1	0.1	0.1	0.3	0.2	2.5	2.2	1.4	1.1	0.2	9.6	0.1	3.9	1.1	67.6	100
Suseo	1.4	0.6	0.4	1.2	2.1	1.7	0.8	0.6	5.1	0.7	3.8	1.4	7.7	0.3	0.3	0.8	2.8	8.9	0.3	14.8	1.7	0.2	42.4	100
Yeoksam 1(il)	0.6	0.6	0.4	0.8	1.1	1	0.7	0.1	0.3	0.3	0.1	0.3	0.5	1.5	1.5	3.2	3.5	0.2	2.3	0.1	11.4	5	64.5	100
Yeoksam 2(i)	2.1	0.8	1	2.1	1.5	3.4	2	0.3	0.6	0.6	0.1	0.8	0.5	1.4	1.3	1.5	3	0.1	2.6	0.2	13.3	9.4	51.4	100

UNIVERSITY OF CAMBRIDGE

Gwanak O/D Matrix

		Destination																						
		Boramae	Cheong nim	Cheong nyong	Daehak	Eun cheon	Haengun	Inheon	Jowon	Jungang	Miseong	Nak seongda e	Nam hyeon	Nangok	Nan hyang	Samseo ng	Seonghy eon	Serim	Sewon	Sillim	Sinsa	Sinwon	Other	Total
Origin	Boramae	24.7	0.8	1.2	1.3	3.6	1.2	0.8	1.1	2.8	0.3	0.9	0.3	0.4	0.3	0.5	1	0.8	1.4	5.3	1	1.2	49.1	100
	Cheongnim	1.1	15	1.2	4.3	1.3	2.6	0.2	0	1.9	0.2	0.3	0.2	0.1	0.2	0.7	2.4	4	3.6	11.5	0.5	1	47.7	100
	Cheongnyong	2.3	2.3	34.8	2.7	4.1	1.7	1.1	0.6	2.2	0.4	4.7	0.2	0.6	0.8	0.5	4.2	1.9	1.1	3.7	0.1	0.6	29.4	100
	Daehak	1.8	3.7	1.5	37.7	0.9	3.1	0.3	0.2	0.3	0.4	0.6	0.3	0.6	0.5	3.4	1.2	1.8	3.5	3.3	0.5	0.4	34	100
	Euncheon	4.7	2	5.6	2.4	22.7	5	1.3	0.3	1.5	0.7	1.4	0.1	0.3	0.7	2.2	8.5	0.9	0.6	5.9	0.6	0.7	31.9	100
	Haengun	1	2.2	1.6	4.6	2.1	15	1.9	0.3	0.8	0.4	2.8	0.4	0.1	0.7	0.9	1.9	1.1	1	2.7	0.2	0.1	58.2	100
	Inheon	1.3	0.2	3.3	1.9	1.5	2.9	47.6	0	0.9	0.5	10.7	1.5	0.3	0.6	0	2.4	0.4	0.1	1.5	0.4	1.1	20.9	100
	Jowon	1.1	0.1	0.9	0.6	0.4	0.2	0.1	32.4	0.4	10	0	0	3.5	0.8	0.6	0	0.1	0.3	2.5	1.4	0.2	44.4	100
	Jungang	1.7	3.8	7.1	1.8	2.9	2.3	2.3	0.7	28.3	1	1.6	0.5	0.3	4.2	0	4.3	0.3	0.6	4.4	0.3	1.6	30	100
	Miseong	0.6	0.2	0.3	0.7	0.4	0.5	0.3	3.8	3	42.9	0.2	0.1	3.8	0.8	0.4	0.2	0.4	1.7	19.2	2.6	0.5	17.4	100
	Nagseongdae	2.8	0.3	14.1	2.6	1.5	2.7	12.3	0	1.1	0.3	24.9	0.2	0.6	0.2	0.6	1.9	1.1	1.5	3.7	0.8	0.1	26.7	100
	Namhyeon	0.6	0.4	0.7	0.2	0.5	2	2.6	0	1.4	0.1	0.2	48.8	0	1.4	0	0.9	0.5	0	0.1	0	0.1	39.5	100
	Nangok	0.8	0.2	1	1.6	0.6	0.3	0.3	2.3	0.2	5.2	0.7	0	43.6	2.2	1.2	0.4	0.6	2.3	15.4	0.5	0.2	20.4	100
	Nanhyang	1.1	1.2	2	3.2	0.7	1.3	0.5	1.5	2.5	1.3	0.1	1.1	3.1	34.8	1.7	0.4	0.3	0.5	15.2	1	0.1	26.4	100
	Samseong	1.6	1.4	1.2	5.7	3.4	0.8	0	1.2	0	0.6	0.7	0	0.6	1.3	37.1	1	1.3	1.4	22.7	0.9	0.2	16.9	100
	Seonghyeon	2.3	12.5	6.1	3.3	4.4	4.5	1.6	0	1.6	0.2	0.9	0.2	0.1	0.3	0.3	22.1	1.6	2.2	2.8	0.6	0.3	32.1	100
	Serim	1.5	8.7	5.2	2.7	0.9	1.2	0.2	0.2	0.1	0.4	0.8	0	0.8	0.2	0.9	2	20.5	3.7	14.6	0	1.5	33.9	100
	Sewon	2.9	5.5	1.8	2.5	0.3	1.3	0.2	1.2	0.1	0.4	0.3	0	0.5	0.1	0.9	2.6	1.1	28.1	5.4	0.4	2.8	41.6	100
	Sillim	2.3	6.1	1.8	1.9	1.6	1.8	0.3	0.8	0.5	4.9	0.7	0.2	5.2	3.3	4	1.1	3	2.5	9.5	1.2	0.7	46.6	100
	Sinsa	1.7	1.3	0.5	0.7	0.7	0.5	0.6	2.1	0.2	4.1	0.6	0	1.1	1.6	1.4	0.4	0	0.4	2.6	43.3	1.5	34.7	100
	Sinwon	2.8	0.7	1.7	2	2.8	1.2	1.7	0.1	2.3	0.8	0.1	0	0.5	0.1	0.3	1.5	1.6	6.9	2.4	1.1	34.2	35.2	100

3.4. Sub-models

3.4.1: Atmospheric Sub-model

Most pollutants have the which vary quite strongly across location and time. Given that each district has only one background station, estimating personal exposure based on uniform air quality might overlook a significant amount of variance of pollution across space and time, and can under- or overestimate personal exposure level, particularly when the study region is larger than a census block level (Dias & Tchepel 2018). Hence, rather than distributing a uniform value homogeneously to a region, we allowed each patch to randomly select one value from the daily pollution field measured at the local background station. For example, on the first day of January 2010, each patch will randomly select one of the PM_{10} values between 09:00 and 19:00. This allows for the likely spatial variability between patches to be preserved, rather than all patches varying in exactly the same way as the local station.

(a) Randomly choose one of the value in the first row

No	Date	Work	h1	h2	h3	h4	h5	h6	h7	h8	h9	h10	h11	h12	h13
1	2010-01-01	home	30	35	36	$\boxed{31}$	33	36	28	38	38	52	43	43	39
2	2010-01-01	work	37	33	31	41	29	41	36	36	37	35	37		
3	2010-01-02	home	39	37	38	40	43	39	41	40	99	85	83	77	70
4	2010-01-02	work	46	44	47	42	39	42	47	55	68	75	102		
5	2010-01-03	home	64	62	59	56	54	52	47	40	46	59	61	64	65
6	2010-01-03	work	49	36	33	41	45	39	46	37	44	50	43		
7	2010-01-04	home	57	48	54	51	44	37	35	36	34	62	54	44	45
8	2010-01-04	work	25	26	28	24	25	27	31	30	33	27	26		
9	2010-01-05	home	46	56	53	60	61	62	62	59	40	33	31	31	37
10	2010-01-05	work	65	63	56	58	41	46	68	64	47	29	41		

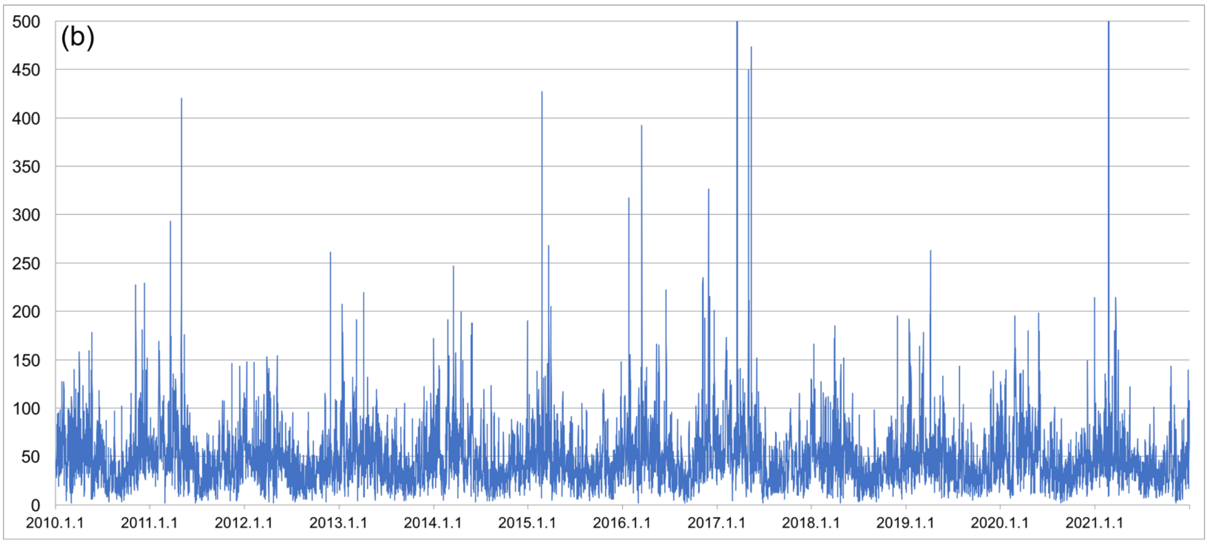


Figure 4 (a) stochastic process in selecting PM10 for each patch per row, and (b) a temporal trend of PM10 from a random location.

We also applied an additional weight factor of pollution to road exposure, from the evidence that adverse effects on health are more likely due to road proximity. According to Harrison, Jones, & Lawrence (2004), the 3 sample sites of London and 1 site in Leeds indicated an average mass increments of 11.5μg/m^3 of PM_{10} and 8.0μg/m^3 increment of $PM_{2.5}$ between roadside (daily mean 34.7μg/m^3 PM_{10}) and urban background areas (daily mean 23.2μg/m^3). Based on the averaged PM_{10} levels in 2010-2015, we discovered that the urban roadside stations in Gangnam and Gwanak showed a 140% and 138% higher concentration compared to the background stations. Thus, we applied these levels to the road patches in each district.

In the DEC scenario, negative pollution levels can be shown close to the end of the simulation.

However, the fact that the pollution drops below zero does not change or affect the outcomes of our research, since this model only detects the threshold levels of 100μg/m^3.

3.4.2. Measurement of exposure

In accordance with the PM_{10} concentration criteria provided by the Korean Ministry of Environment, the air quality is considered as *Good* for PM_{10} in the range 0-30μg/m^3, *Moderate* for 31-80μg/m^3, *Unhealthy* for 81-150μg/m^3, and *Very Unhealthy* for over 150μg/m^3. As mentioned above. Since the diseases resulting from pollution exposure are generally non-communicable, at least to first order, the model does not consider any transmission effects between individuals, but considers the effects resulting from continuous interactions between individual spatial trajectories constrained by daily activity patterns, with the simulated spatial micro-scale atmospheric pollution distributions.

It was assumed that, compared to a healthy person, a person suffering from disease symptoms will lose a steadily greater amount of health the more they are exposed to pollution. For example, an agent with health 50 will lose health more rapidly than the agent with health 150 when they are equally exposed to over 100μg/m^{3}of PM_{10}. Accordingly, we set the rate of change of an individual's health status caused by PM_{10} exposure to vary linearly with health:

$$If\ PM10\ \geq 100 \qquad \frac{dH}{dT} = -\propto \left(H_{max} - H(t)\right) + H_{recov}$$

Equation 1 Exposure measurement for each agent

where H_{max} denotes an agent's health status at the beginning, and H(t) (strictly less than H_{max}). Thus, in the absence of any recovery and with constant α and PM_{10} always above the threshold, agent health values would decrease exponentially away from their initial value H(0). H(t) is the current health value. The factor α sets the rate of change per unit time when health impact applies. This factor is chosen from a random uniform distribution between zero and a maximum on each tick, to allow for the fact that even within a patch, since these are 30m across, individual exposure levels will be very different. To some extent this mimics the fact that people may move in and out of buildings, for example. For vulnerable populations we apply the first term in eq. 1 again (so for example, an agent in the youngest age group has double the probability of experiencing an effect). H_{recov} is a health recovery rate that varies by the real estate price of the agent's home location, as in Figure 2.

Overall, this gives us a threshold model where impacts begin to accumulate whenever air quality is marked as unhealthy. However, there are random variations applied both to the above agent selection and to the health impact to allow for the fact that exposure will not be uniform across agents even when they are more or less spatially co-located (e.g. they may be inside or outside a building, or inside a vehicle rather than walking at the roadside).

Reference

DAVID and DON. (2012). An integrated agent-based framework for assessing air pollution impacts. *Journal of Environmental Protection*, *2012*.

DIAS and TCHEPEL. (2018). Spatial and Temporal Dynamics in Air Pollution Exposure Assessment. *International Journal of Environmental Research and Public Health*, *15*(3), 558.

FELSEN and WILENSKY. (2007). NetLogo Urban Suite - Pollution model. Illinois: Center for

Connected Learning and Computer-Based Modeling, Northwestern University, Evanston. Retrieved from http://ccl.northwestern.edu/netlogo/models/UrbanSuite-Pollution

HARRISON JONES and LAWRENCE. (2004). Major component composition of PM10and PM2.5from roadside and urban background sites. *Atmospheric Environment*, *38*(27), 4531–4538. https://doi.org/10.1016/j.atmosenv.2004.05.022

KAN and CHEN. (2004). Particulate air pollution in urban areas of Shanghai, China: Health-based economic assessment. *Science of the Total Environment*, *322*(1–3), 71–79. https://doi.org/10.1016/j.scitotenv.2003.09.010

KAN LONDON CHEN ZHANG SONG ZHAO … CHEN. (2008). Season, sex, age, and education as modifiers of the effects of outdoor air pollution on daily mortality in Shanghai, China: The Public Health and Air Pollution in Asia (PAPA) Study. *Environ Health Perspect*, *116*(9), 1183–1188.

O'NEILL JERRETT KAWACHI LEVY COHEN GOUVEIA … SCHWARTZ. (2003). Health, wealth, and air pollution: advancing theory and methods. *Environmental Health Perspectives*, *111*(16), 1861.