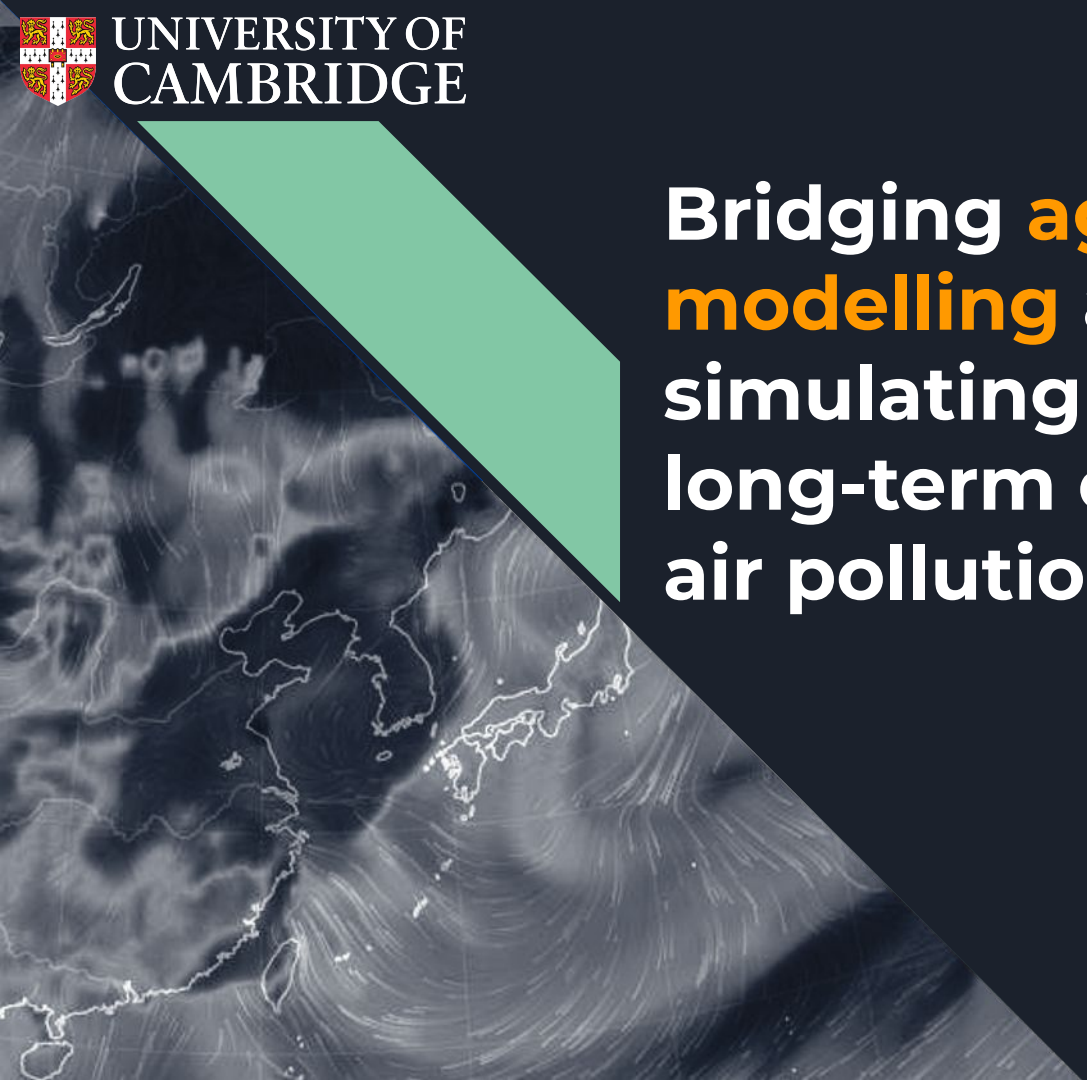


UNIVERSITY OF CAMBRIDGE

Bridging agent-based modelling and R with nlrx: simulating pedestrian's long-term exposure to air pollution

Hyesop Shin

Ph.D Candidate

Dept. of Geography and Fitzwilliam College

University of Cambridge

@hyesop

TOC

Files, Codes and Tutorials:
http://tiny.cc/nlrx

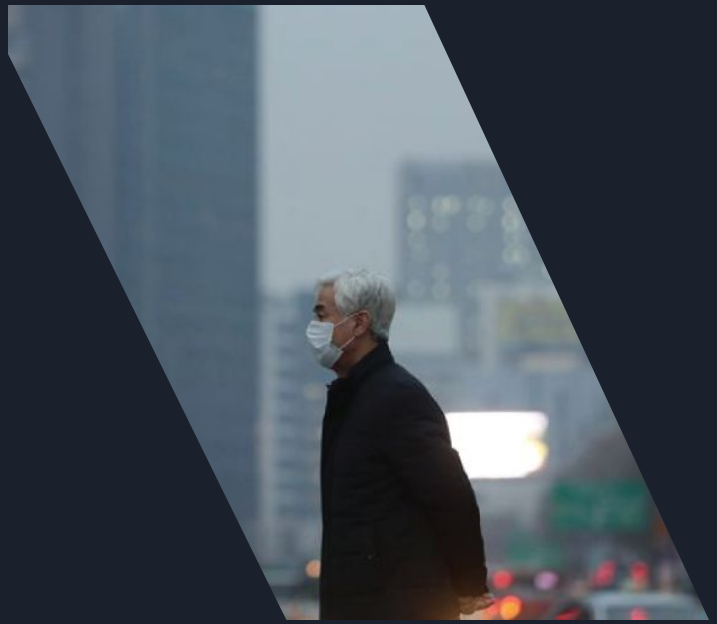
Project Objective

Air Pollution in South Korea (March, 2019)

Songdo, Incheon (Mar.5th, 2019)

Central Seoul (Mar.6th, 2019)

Project objective



- This study aims to estimate pedestrian's exposure to acute air pollution in Seoul districts using agent-based simulation
  - How does socioeconomic group potentially affect health outcomes?
  - How could health levels change under different pollution scenarios?

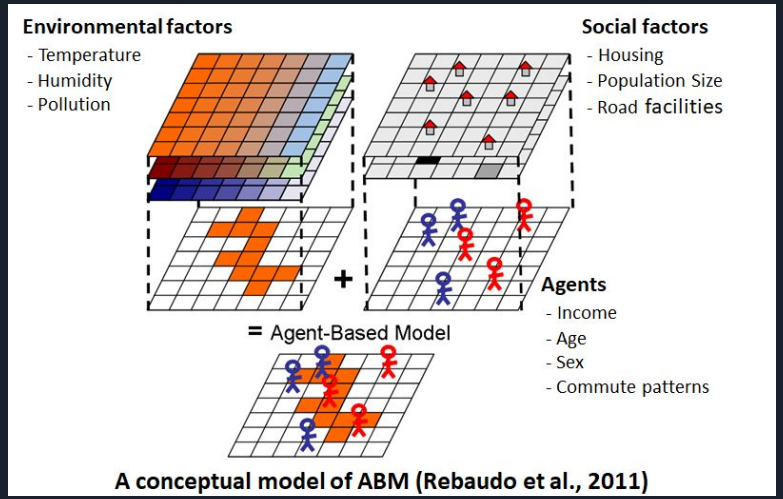
Agent-based modelling

Agent-based modelling (ABM) is a computational method that focuses on individual's movements and interactions that can affect the system structure

e.g. Social media, epidemics, decision-making

Components

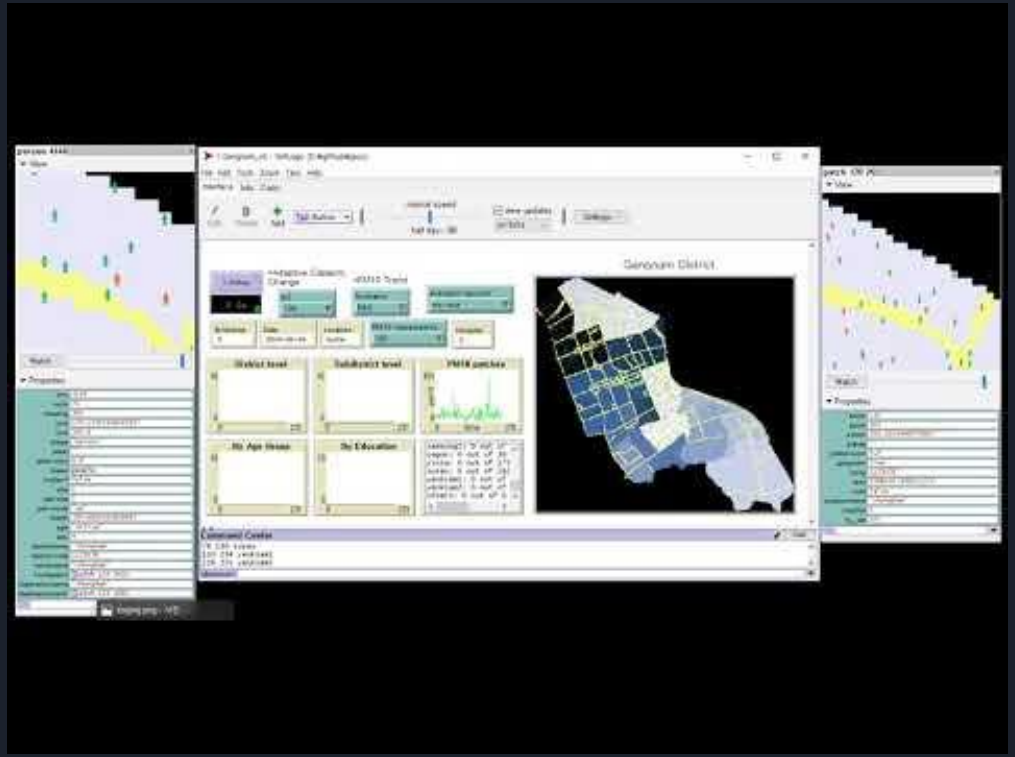
- Agents = *turtles*
- Environments = *patches*



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

ABM setup

How ABM works: NetLogo example



Agent-based modelling: Settings

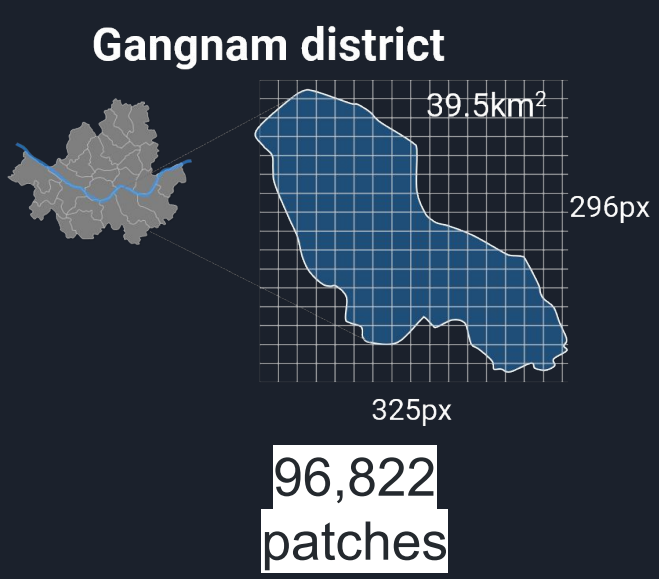
Study area + demographic info

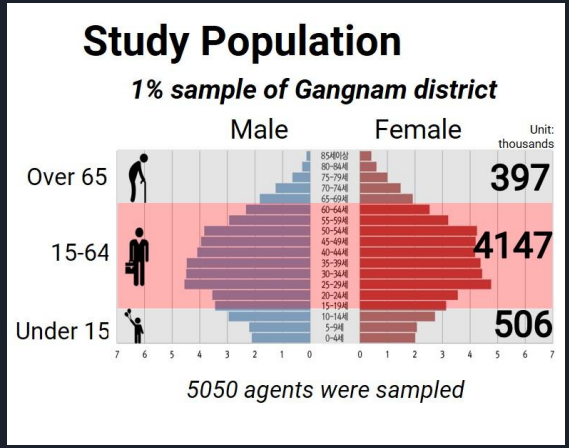
Agent setup and create destination

Measure health loss and recovery

Export file as export.csv

Import files to R





UNIVERSITY OF CAMBRIDGE

Agent-based modelling: Settings

Study area + demographic info

Agent setup and create destination

Measure health loss and recovery

Export file as export.csv

Import files to R

- The population of Gangnam is allocated into three groups: Under 15, 15-64, Over 65
- A day is splitted into two time sequences: Work hours (09-19 hrs), Home (20-08 hrs)
- Agents have no previous exposure experienced
- Agents aged 15-64 follow OD matrix while restricting other groups' movement range close to their origin

Origin \ Destination	ap	chungdam	daechi1	daechi2	daechi4	dogok1	dogok2	gaepo1	gaepo2	gaepo4	ilwon1	ilwon2	ilwon	nonhyun1	nonhyun2	samsung1	samsung2	segok	sinsa	suseo	yeoksam1	yeoksam2	others	Σ
ap	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	53.8	100
chungdam	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
daechi1	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
daechi2	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
daechi4	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
dogok1	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
dogok2	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
gaepo1	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
gaepo2	0.5	0.8	1	3.2	2	3.1	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
gaepo4	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
ilwon1	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
ilwon2	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	3.1	1.6	0.6	1	1.4	1.9	1.3	36.9	100
ilwon	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
nonhyun1	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
nonhyun2	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
samsung1	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
samsung2	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
yeoksam1	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
yeoksam2	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

Agent-based modelling: Settings

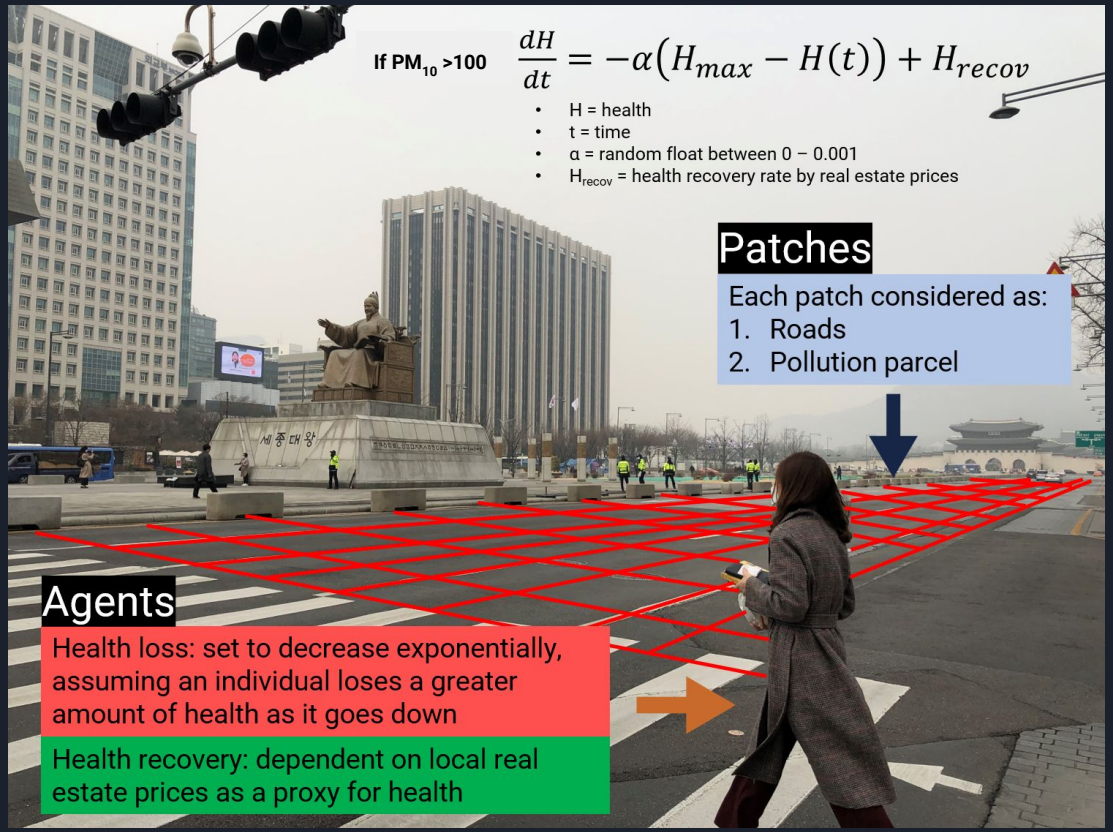
Study area + demographic info

Agent setup and create destination

Measure health loss and recovery

Export file as export.csv

Import files to R



UNIVERSITY OF CAMBRIDGE

Agent-based modelling: Settings

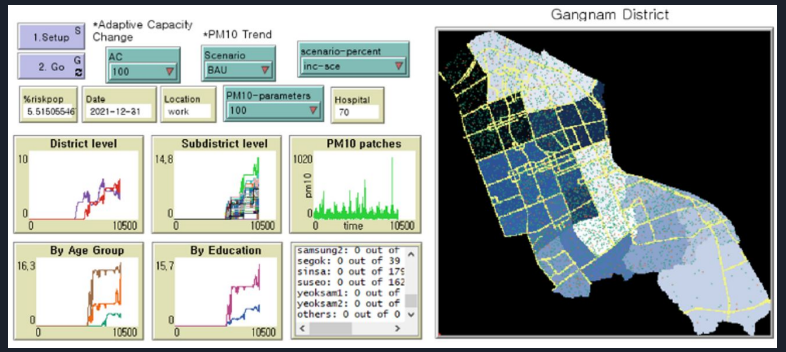
Study area + demographic info

Agent setup and create destination

Measure health loss and recovery

Export file as export.csv

Import files to R



1.2 hours
for a single run

50 iterations

Low Quality Images

Exported as (uncleaned) .csv

UNIVERSITY OF CAMBRIDGE

You will end up doing this...

Solving current problems from NetLogo

01 How can R reduce human intervention during analysis?

02 How can R improve the quality of figures?

03 How can R increase the simulation's clock speed?

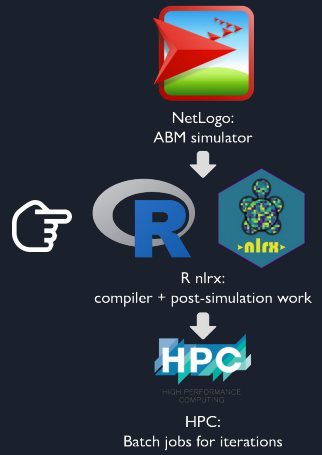
ABM with nlrx:

a tool to setup and execute NetLogo simulations from R

Netlogo + R + XML

UNIVERSITY OF CAMBRIDGE

Stage 1: Install packages nlrx



Java setup

```
Sys.setenv(JAVA_HOME= 'XXXX') #Varies by OS
```

Load pkgs

```
library(nlrx)
library(tidyverse)
library(rcartocolor)
library(ggthemes)
```

Assign path

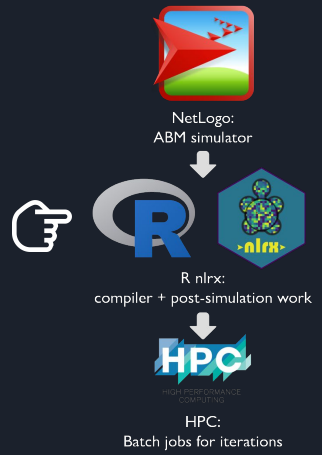
```
nlpath <- file.path("/home/hs621/NetLogo 6.0.4")
modelpath <- file.path(path,"Gangnam.nlogo")
outpath <- file.path("/home/hs621/out")
```

Create an nl object

```
nl <- nl(nlversion = "6.0.4",
         nlpath = nlpath,
         modelpath = modelpath,
         jvmmem = 1024)
```


Codes and tutorials: http://tiny.cc/nlrx

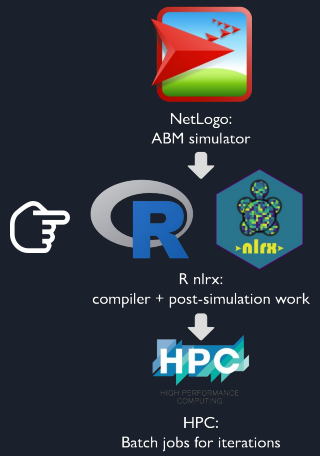
Stage 2: Attach an experiment (1/2)




```
nl@experiment <- experiment(expname = "seoul",
                            outpath = outpath,
                            repetition = 1,
                            tickmetrics = "true",
                            idsetup = "setup",
                            idgo = "go",
                            runtime = 8764,
                            evalticks=seq(1,8764, by = 100),
```


Codes and tutorials: http://tiny.cc/nlrx

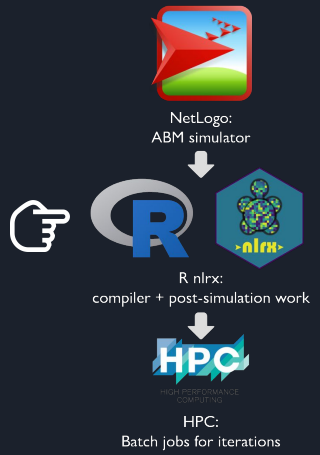
Stage 3: Attach an experiment (2/2)




```
constants = list("PM10-parameters" = 100,
                 "Scenario" = "\"BAU\"",
                 "scenario-percent" =
                 "\"inc-sce\""),
variables = list('AC' = list(values=c(100,150,200))),
metrics.turtles =  list("people" = c("xcor", "ycor",
                   "homename", "destinationName", "age",
                   "health")
metrics.patches =  list("patch" = c("pxcor", "pycor",
"pcolor"))
```


Codes and tutorials: http://tiny.cc/nlrx

Stage 4: Attach a simulation design and run model



Iteration sampling

```
nl@simdesign <- simdesign_distinct(nl=nl,
                                   nseeds=1)
```

Run

```
init <- Sys.time()
results <- run_nl_all(nl = nl)
Sys.time() - init
```

Add results to nl list

```
setsim(nl, "simoutput") <- results

write_simoutput(nl)
```


Codes and tutorials: http://tiny.cc/nlrx

Stage 5: Submit batch jobs to reduce time

NetLogo:
ABM simulator

R nlrx:
compiler + post-simulation work

HPC:
Batch jobs for iterations

hs621@login-e-16:~/github/nlrx

File Edit View Search Terminal Help

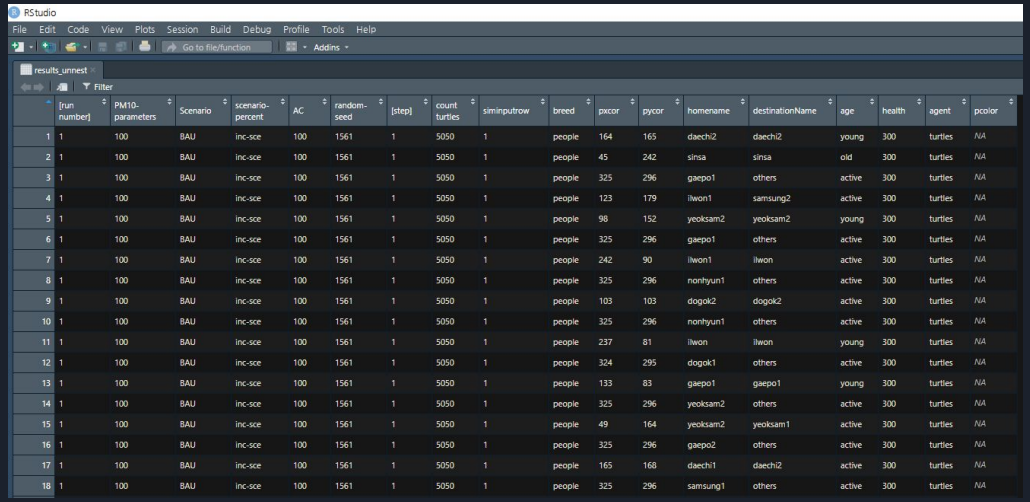
GNU nano 2.3.1 File: slurm_submit.peta4-skylake_nlrx

```
#!/bin/bash
#!
#! Example SLURM job script for Peta4-Skylake (Skylake CPUs, OPA)
#! Last updated: Mon 13 Nov 12:25:17 GMT 2017
#!

#!#############################################################
#!#### Modify the options in this section as appropriate ######
#!#############################################################

#! sbatch directives begin here ###############################
#! Name of the job:
#SBATCH -J nlrx
#! Which project should be charged:
#SBATCH -A BITHELL-SL3-CPU
#! How many whole nodes should be allocated?
#SBATCH --nodes=1
#! How many (MPI) tasks will there be in total? (<= nodes*32)
#! The skylake/skylake-himem nodes have 32 CPUs (cores) each.
#SBATCH --mem=99999
#! How much wallclock time will be required?
#SBATCH --time=12:00:00
#! What types of email messages do you wish to receive?
#SBATCH --mail-type=END
#! Uncomment this to prevent the job from being requeued (e.g. if
#! interrupted by node failure or system downtime):
##SBATCH --no-requeue

#! For 6GB per CPU, set "-p skylake"; for 12GB per CPU, set "-p skylake-himem":
#SBATCH -p skylake-himem
```

Post-simulation

Result structure: nested tibble

⇒ unnested result

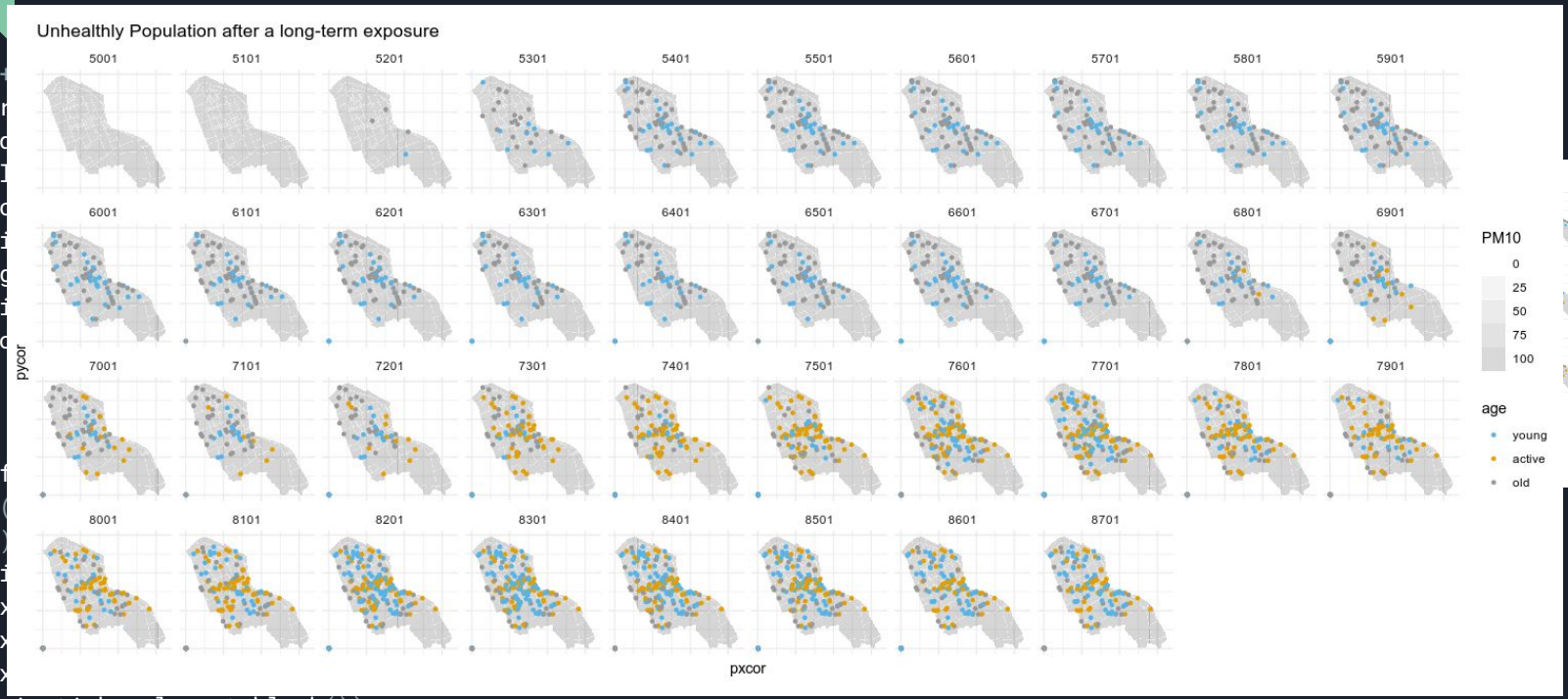


	[run number]	PM10-parameters	Scenario	scenario-percent	AC	random-seed	[step]	count turtles	siminputrow	breed	pxcor	pycor	homename	destinationName	age	health	agent	pcolor
1	1	100	BAU	inc-sce	100	1561	1	5050	1	people	164	165	daechi2	daechi2	young	300	turtles	NA
2	1	100	BAU	inc-sce	100	1561	1	5050	1	people	45	242	sinsa	sinsa	old	300	turtles	NA
3	1	100	BAU	inc-sce	100	1561	1	5050	1	people	325	296	gaepo1	others	active	300	turtles	NA
4	1	100	BAU	inc-sce	100	1561	1	5050	1	people	123	179	ilwon1	samsung2	active	300	turtles	NA
5	1	100	BAU	inc-sce	100	1561	1	5050	1	people	98	152	yeoksam2	yeoksam2	young	300	turtles	NA
6	1	100	BAU	inc-sce	100	1561	1	5050	1	people	325	296	gaepo1	others	active	300	turtles	NA
7	1	100	BAU	inc-sce	100	1561	1	5050	1	people	242	90	ilwon1	ilwon	active	300	turtles	NA
8	1	100	BAU	inc-sce	100	1561	1	5050	1	people	325	296	nonhyun1	others	active	300	turtles	NA
9	1	100	BAU	inc-sce	100	1561	1	5050	1	people	103	103	dogok2	dogok2	active	300	turtles	NA
10	1	100	BAU	inc-sce	100	1561	1	5050	1	people	325	296	nonhyun1	others	active	300	turtles	NA
11	1	100	BAU	inc-sce	100	1561	1	5050	1	people	237	81	ilwon	ilwon	young	300	turtles	NA
12	1	100	BAU	inc-sce	100	1561	1	5050	1	people	324	295	dogok1	others	active	300	turtles	NA
13	1	100	BAU	inc-sce	100	1561	1	5050	1	people	133	83	gaepo1	gaepo1	young	300	turtles	NA
14	1	100	BAU	inc-sce	100	1561	1	5050	1	people	325	296	yeoksam2	others	active	300	turtles	NA
15	1	100	BAU	inc-sce	100	1561	1	5050	1	people	49	164	yeoksam2	yeoksam1	active	300	turtles	NA
16	1	100	BAU	inc-sce	100	1561	1	5050	1	people	325	296	gaepo2	others	active	300	turtles	NA
17	1	100	BAU	inc-sce	100	1561	1	5050	1	people	165	168	daechi1	daechi2	active	300	turtles	NA
18	1	100	BAU	inc-sce	100	1561	1	5050	1	people	325	296	samsung1	others	active	300	turtles	NA

General info: Iteration, Scenario, Random seed, Step,
Agent info: Count turtles, Breed, X, Y, home, destination, age, health
Patch info: X, Y, patch colour

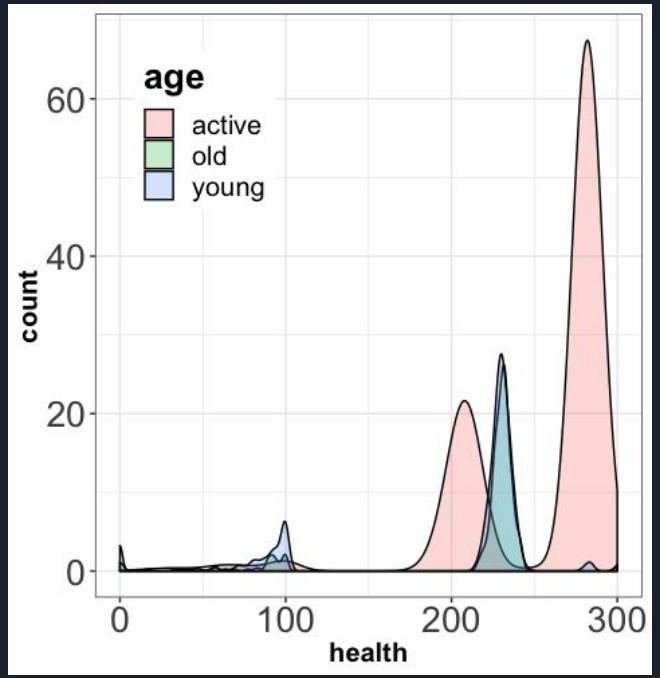
Mapping unhealthy population with `ggplot2`

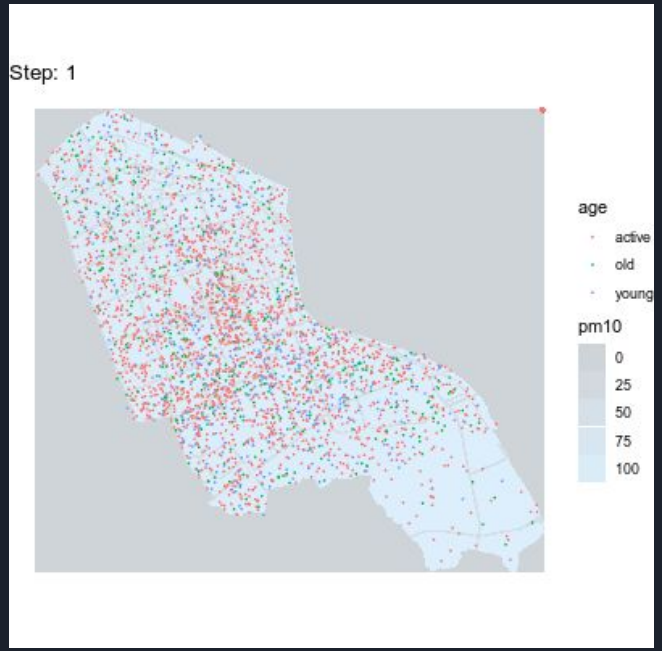
```
ggplot() +
  facet_wr
  coord_ec
  geom_til
fill=pcolo
  geom_poi
color = ag
  scale_fi
  scale_co



  guides(f
  ggtitle(
exposure")
  theme_mi
  theme(ax
        ax
        ax
        axis.ticks=element_blank())
```



Codes and tutorials: http://tiny.cc/nlrx

Density plot with `ggplot2` & Animations with `gganmiate`

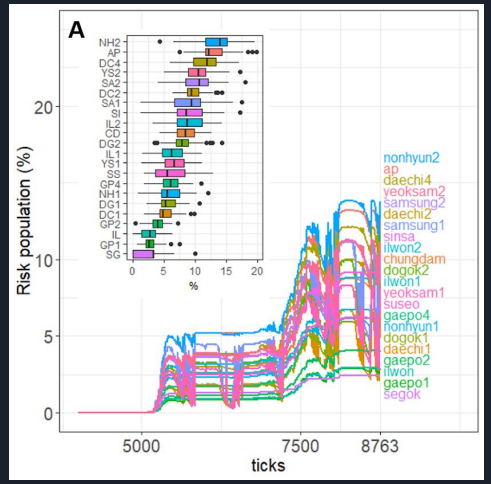




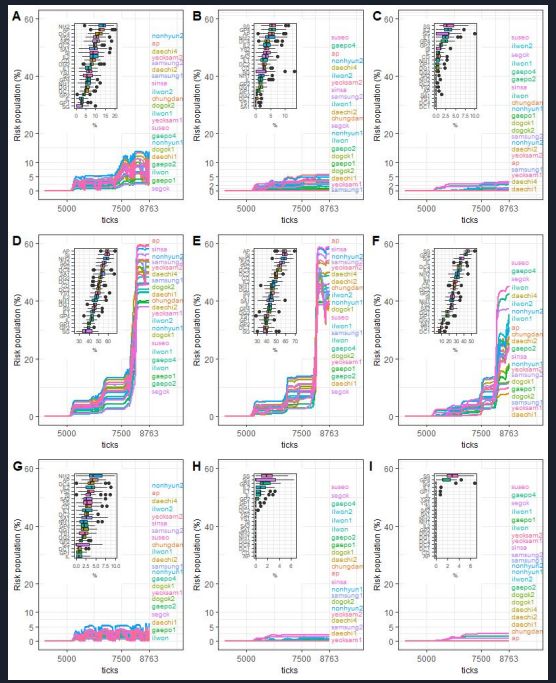
Codes and tutorials: https://tiny.cc/nlrx

`gridextra` & `directlabels` for HQ images

`direct.label()`



`grid.arrange()`



UNIVERSITY OF CAMBRIDGE

Codes and tutorials: http://tiny.cc/nlrx

Summary and Contribution

Summary

- Contents
  - Disparities in health outcomes are likely to depend on demographic status
  - When the vulnerable group (the old and young) is exposed over a long period, road proximity causes additional health degradation
- R & nlrx
  - Workload: 60% on NetLogo, 30% on R, 10% HPC
  - nlrx works as a compiler
  - nlrx results in a tibble format that can be plotted in a variety of figures for different purposes with `ggplot, gganimate`

Contributions from nlrx

- Time saving
  - Fast iterative process
  - The simulation ends with only a paragraph of codes
- Solves *fat finger* issues
  - No need to delete readme text from the NetLogo output
- Convenience
  - Doesn't need *rJava* installation *(Really helps when running HPC)*
  - Silent machines
  - Code categorical variables

For more information...

Paper

JASSS

An Agent-Based Assessment of Health Vulnerability to Long-Term Particulate Exposure in Seoul Districts

Hyesop Shin[1] and Mike Bithell[1]

[1] *Department of Geography, University of Cambridge, Downing Place CB2 3EN, United Kingdom*
Correspondence should be addressed to *hs621@cam.ac.uk*

Journal of Artificial Societies and Social Simulation 22(1) 2, 2019
Doi: 10.18564/jasss.3940 Url: http://jasss.soc.surrey.ac.uk/22/1/2.html

Received: 22-01-2018 Accepted: 11-01-2019 Published: 31-01-2019

Tutorial https://tiny.cc/nlrx

Thank you!

 @hyesop

 hs621@cam.ac.uk

 @mrsensible

Agent-based modelling: advantages

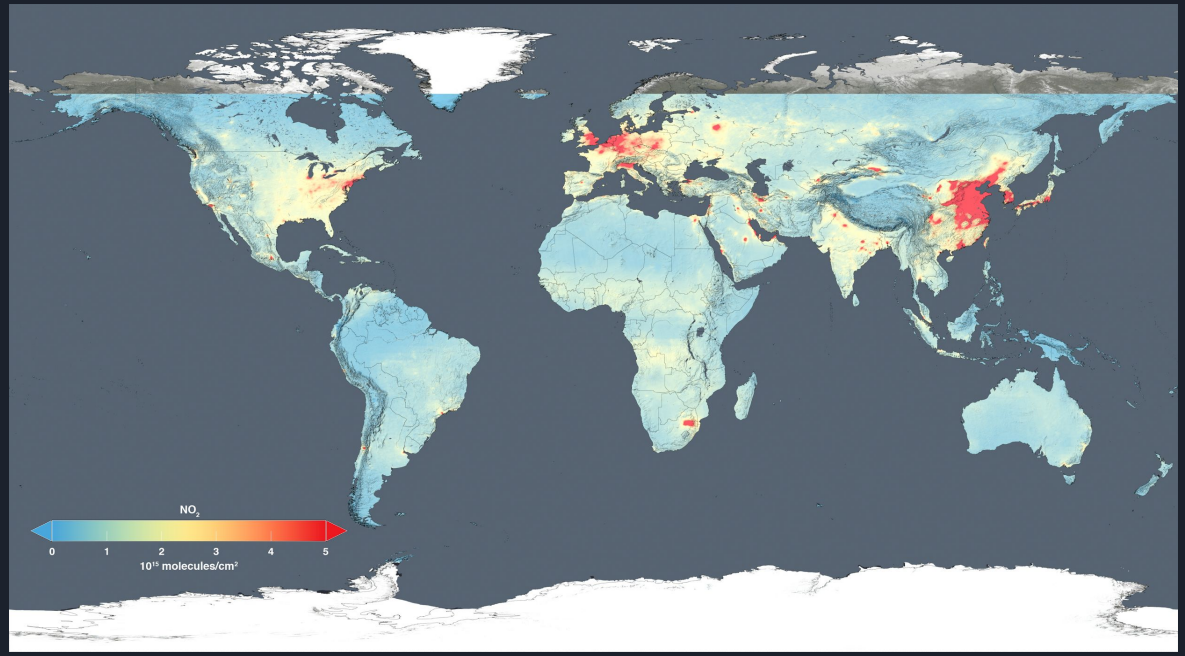
- Is able to simulate human-environment interaction
- Can analyse adverse health impact by social groups
- Follows trajectories of individuals over time
- Envisages effects from possible scenarios ("What if…?")
- NetLogo is the most widely used software in the ABM world

Air Pollution in South Korea (March, 2019)

Songdo, Incheon (Mar.5th, 2019)

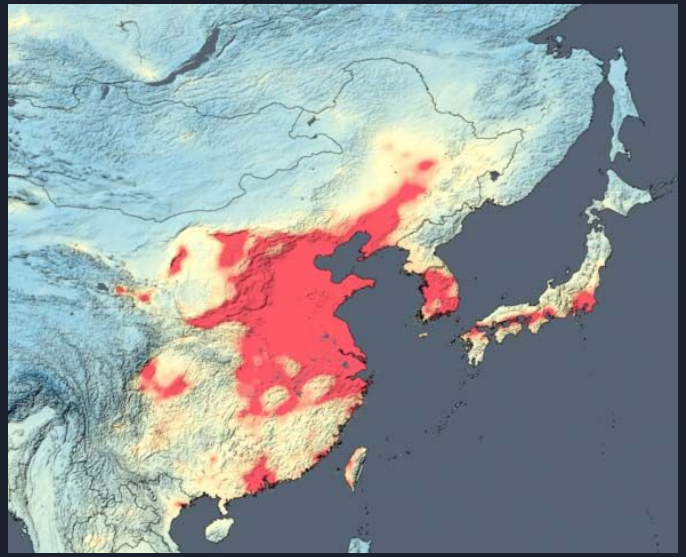
Central Seoul (Mar.6th, 2019)

Urbanisation and Air Pollution



NO_2 trend in a decade (2005-2015), Satellite imagery taken by NASA, 2015

Urbanisation and Air Pollution



NO_2 change in East Asian countries

Traffic in Seoul CBD

Health threat may differ by demographic factors

- Children are inherently more susceptible to air pollution as their lung function and immunological systems are still developing (Pearce et al., 2006)
- Higher risks due to the PM_{10} exposure were observed for elderly individuals - COPD, stroke, etc (Halonen et al., 2016; Wang et al., 2016)

More attention should be given to how travel behaviours differ by social groups (e.g. age), and how health loss are manifested in each group after a long-term pollution exposure

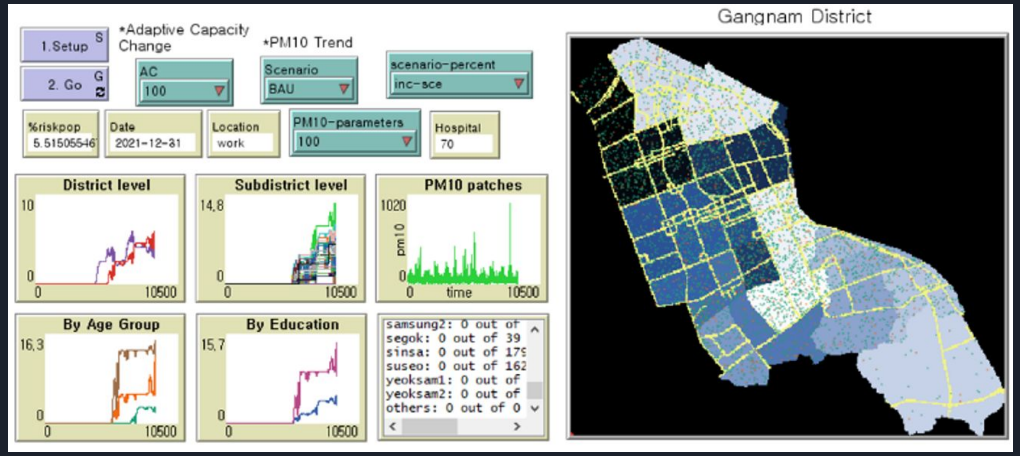
Stage 1: Install and finish coding in NetLogo

NetLogo:
ABM simulator

R nlrx:
compiler + post-simulation work

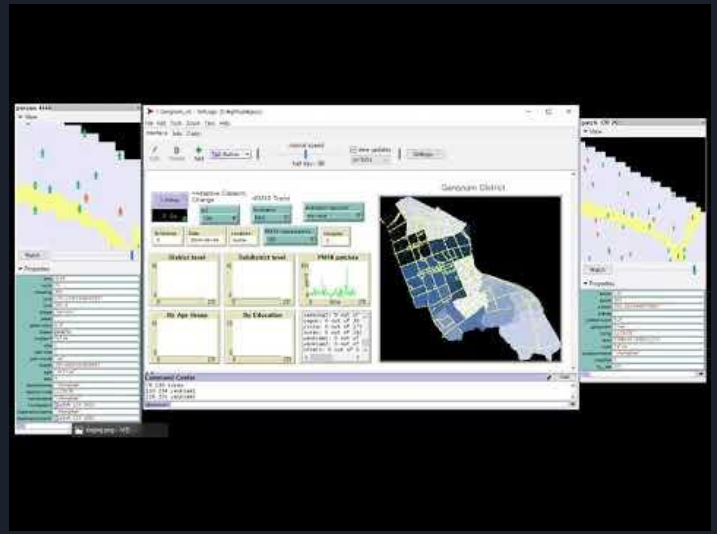
HPC:
Batch jobs for iterations

- Install NetLogo >= 5.3, NetLogo 6 is preferred
- Java required



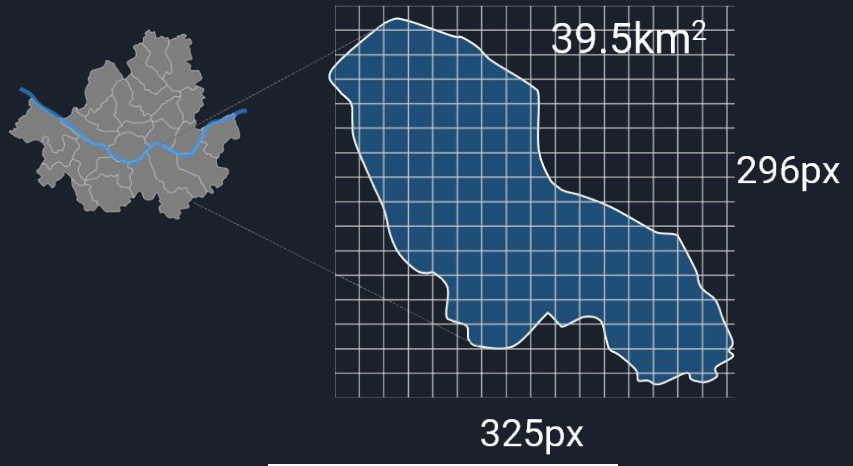
Agent-based modelling: advantages

- Is able to simulate human-environment interaction
- Can analyse adverse health impact by social groups
- Follows trajectories of individuals over time
- Envisages effects from possible scenarios ("What if…?")
- NetLogo is the most widely used software in the ABM world



NetLogo Example

Gangnam district



96822 patches