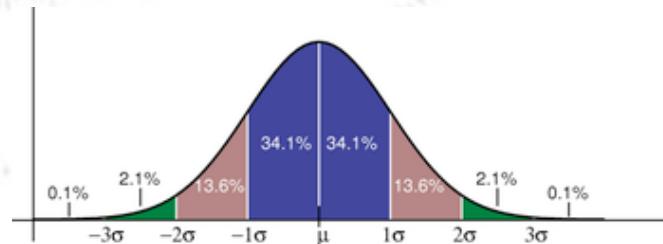


Modelling Covid-19

Agent Based Network & effects of inhomogeneities



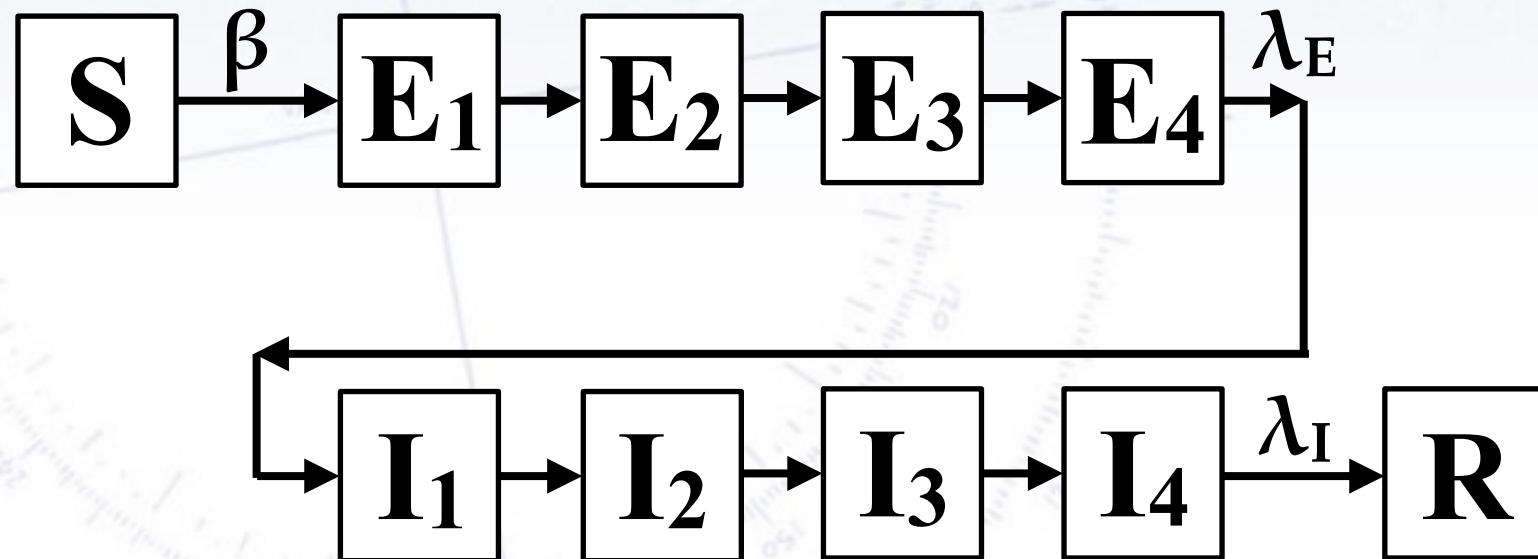
Mathias Heltberg (Paris), Christian Michelsen (NBI), Troels C. Petersen (NBI)



Modified SEIR model (for reference)

Modified SEIR model

In order to be able to check our model against a reference, and subsequently quantify the size of effects introduced into the agent based network model, we build a modified SEIR model (mSEIR) as follows:

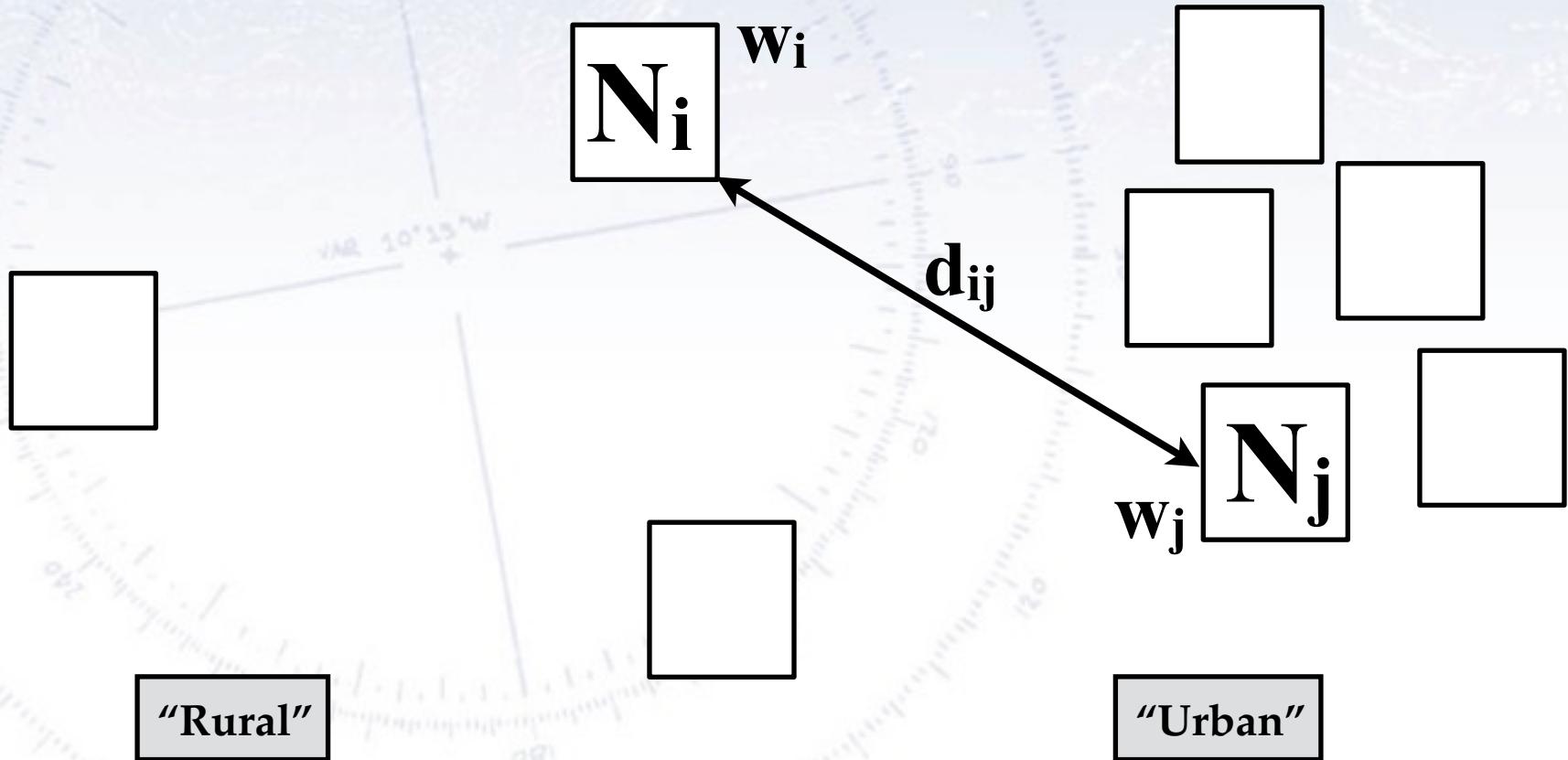


The choice of four E and I states is arbitrary and can be changed. Model serves as **reference**, to see **effects of spatial and infection rate inhomogeneties**.

Agent Based Network model (for exploration of effects)

Agent Based Network

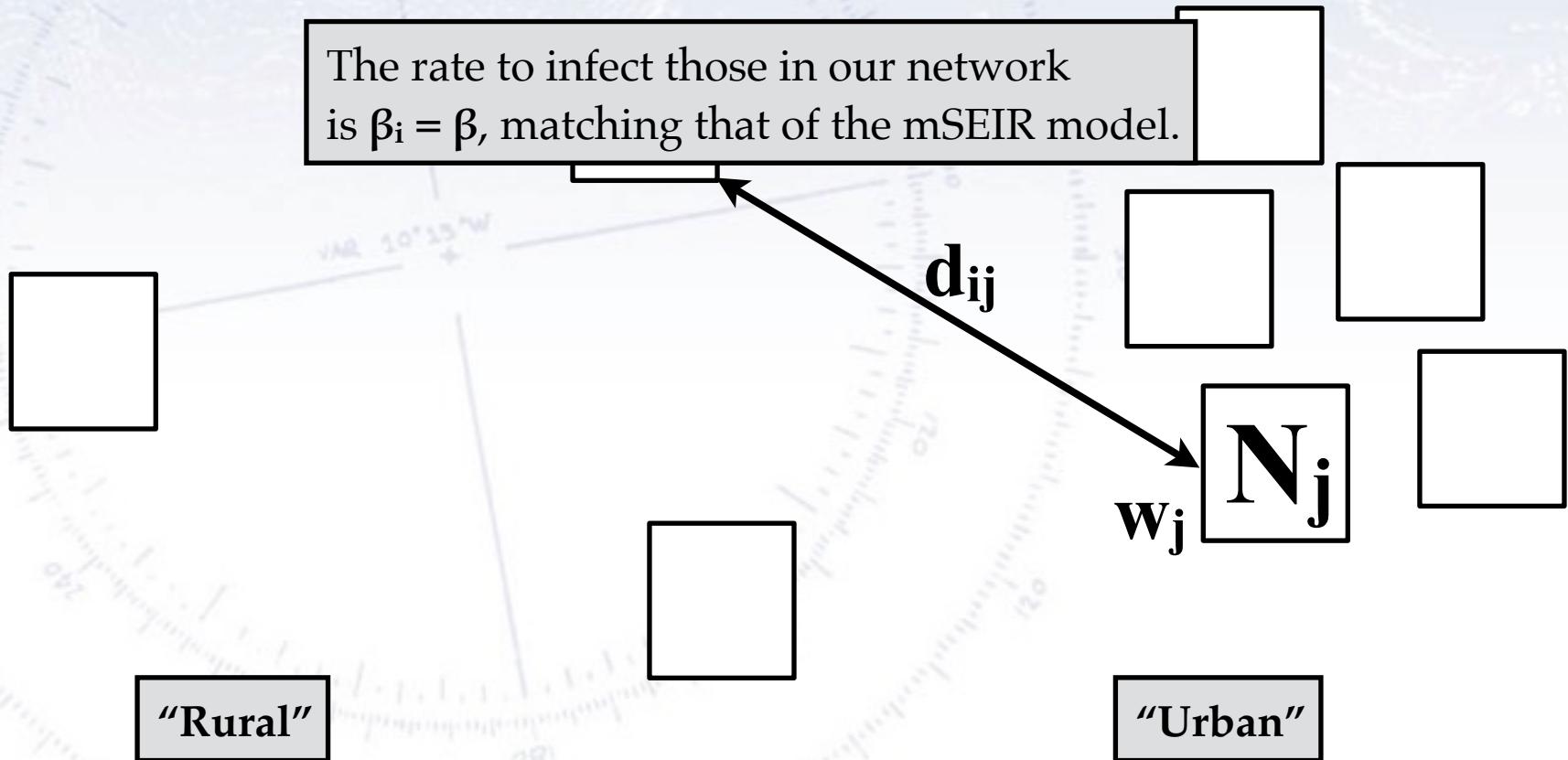
For an Agent Based Network (ABN) model, we choose N_{total} nodes each with a weight w_i and choose an average number of connections, μ .



We choose two random nodes (i,j) with probability proportional to w_i and w_j , and connect over distance d_{ij} with probability proportional to $p(d_{ij}) = \exp(-\rho d_{ij})$

Agent Based Network

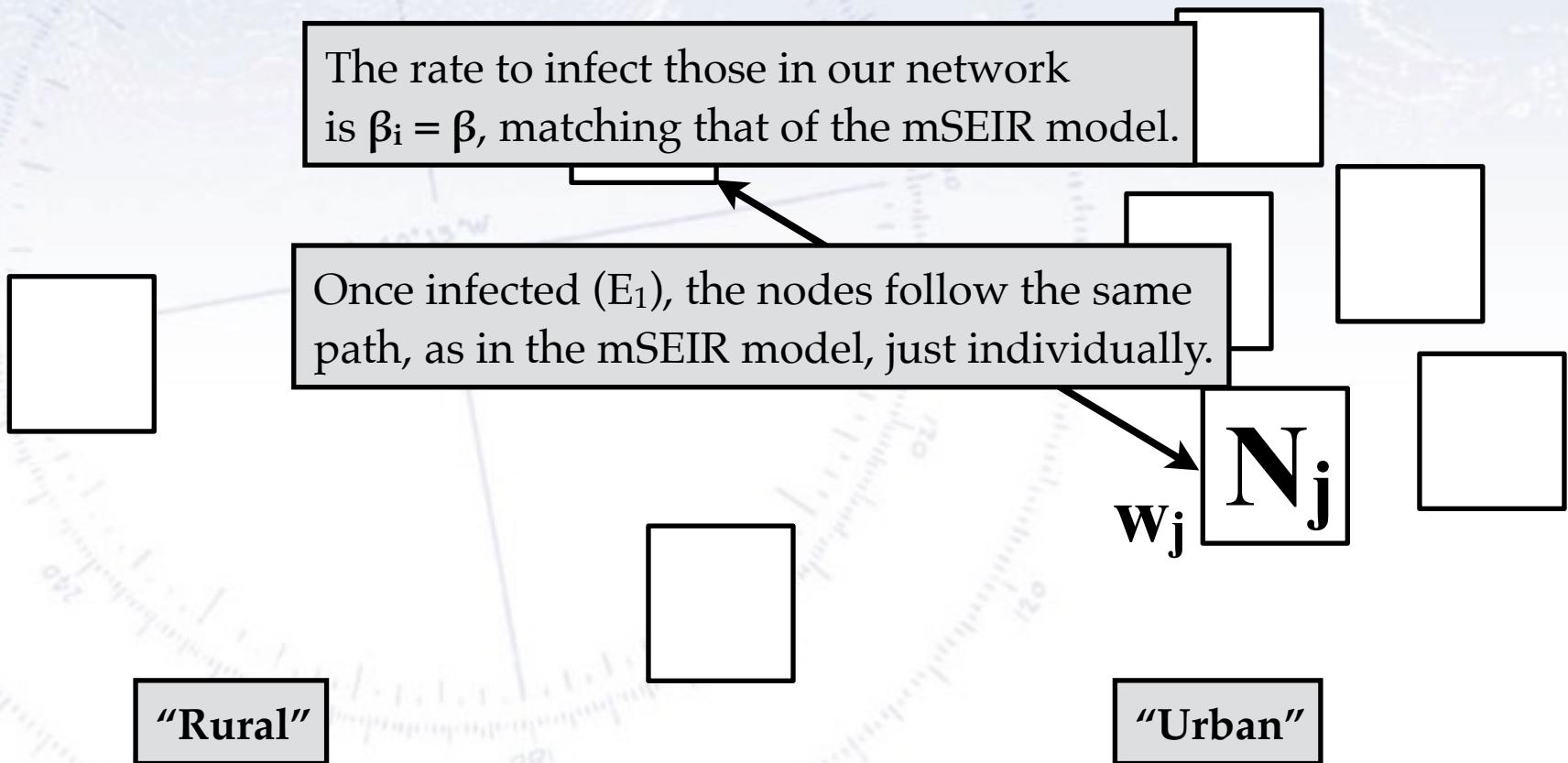
For an Agent Based Network (ABN) model, we choose N_{total} nodes each with a weight w_i and choose an average number of connections, μ .



We choose two random nodes (i,j) with probability proportional to w_i and w_j , and connect over distance d_{ij} with probability proportional to $p(d_{ij}) = \exp(-\rho d_{ij})$

Agent Based Network

For an Agent Based Network (ABN) model, we choose N_{total} nodes each with a weight w_i and choose an average number of connections, μ .



We choose two random nodes (i,j) with probability proportional to w_i and w_j , and connect over distance d_{ij} with probability proportional to $p(d_{ij}) = \exp(-\rho d_{ij})$

Agent Based Network

For an Agent Based Network (ABN) model, we choose N_{total} nodes each with a weight w_i and choose an average number of connections, μ .

The rate to infect those in our network
is $\beta_i = \beta$, matching that of the mSEIR model.

Once infected (E_1), the nodes follow the same path, as in the mSEIR model, just individually.

To begin with, we choose:
all weights the same ($w_i = 1$), i.e. connection rate homogeneity,
all distance measure the same ($\rho = 0$), i.e. spatial homogeneity.

“Rural”

“Urban”

We choose two random nodes (i,j) with probability proportional to w_i and w_j ,
and connect over distance d_{ij} with probability proportional to $p(d_{ij}) = \exp(-\rho d_{ij})$

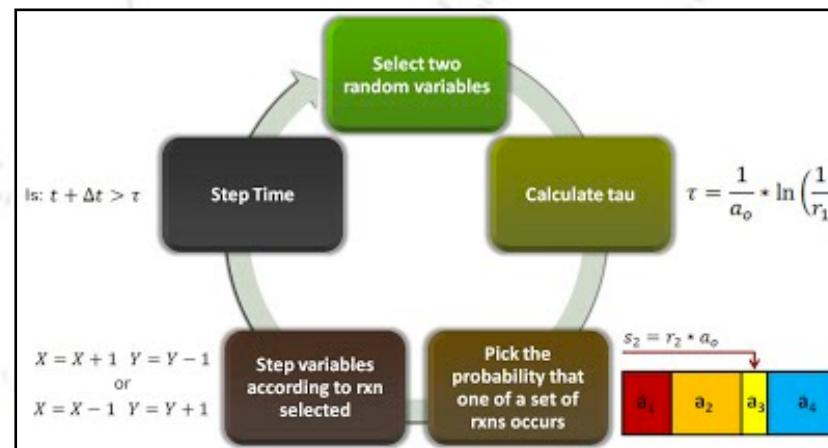
Implementation

The ABN model is implemented as a Doob-Gillespie model, i.e. calculating the time to the next “event”.

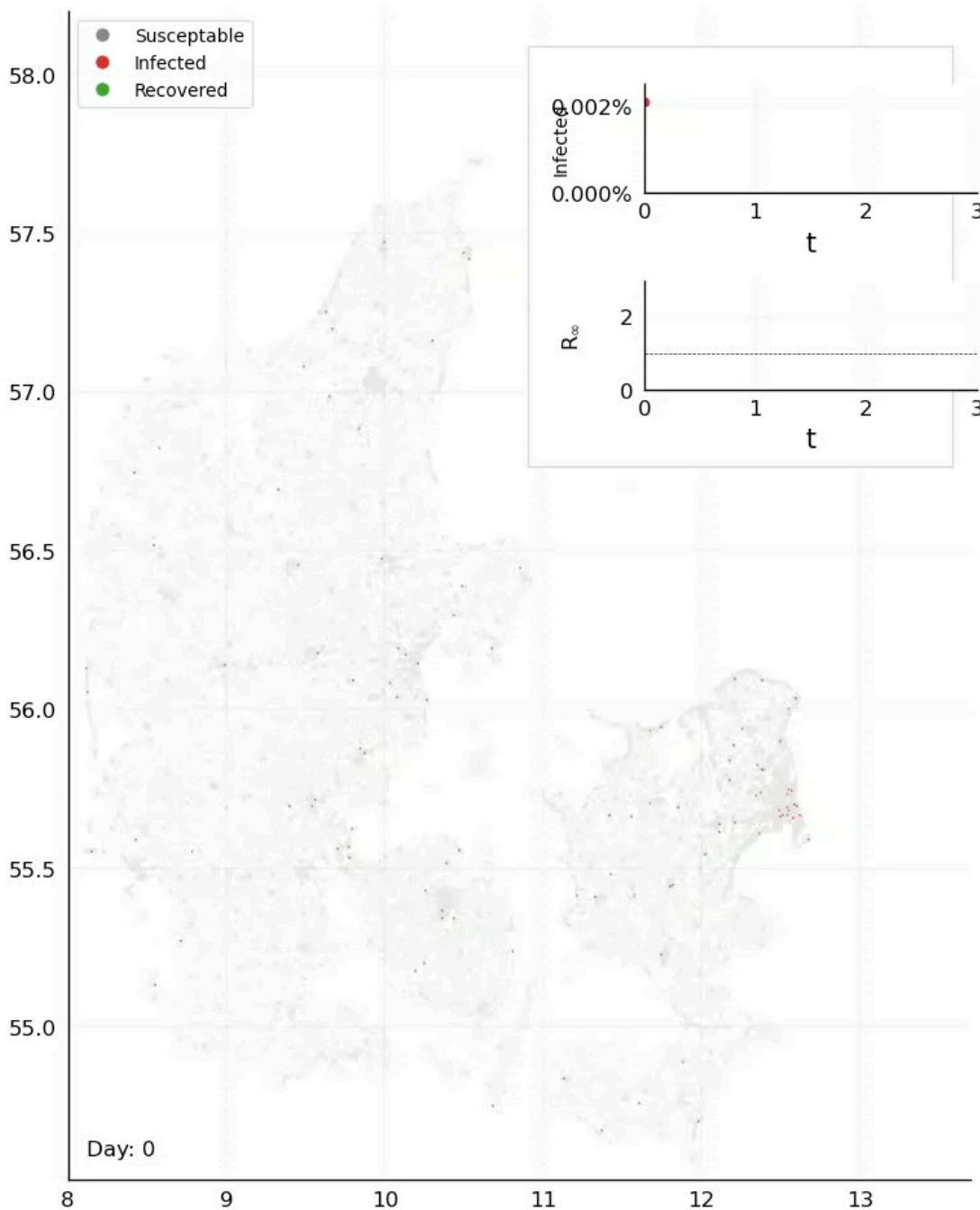
We have managed to run the model for the full Danish population, which takes 25 minutes.

We use housing sales GPS coordinates 2007-2019 [from Boligsiden] as a proxy for where people live, and distribute the nodes according to a 2D kernel fit to data.

We run several different parameter settings in parallel, and typically repeat the models simulation 100 times (but “only” for a population of 500.000).



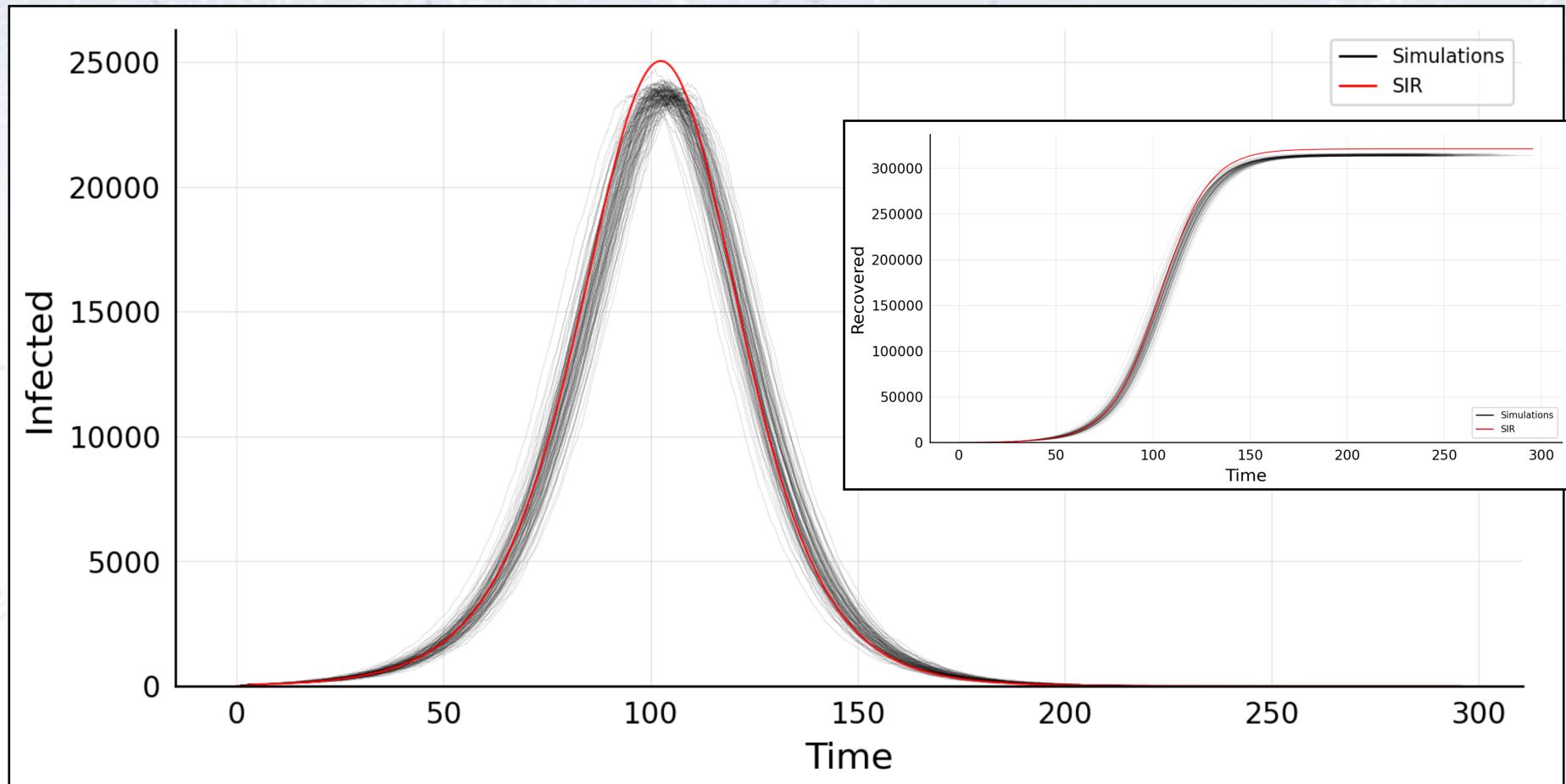
$N=5M$, $N_{\text{init}}=100$, $\mu=20.0$, $\sigma_\mu=0.0$, $\rho=0.0$, $\beta=0.01$, $\sigma_\beta=0.0$, $\lambda_E=1.0$, $\lambda_I=1.0$, $\epsilon_\rho=0.01$, $f_{02}=0.0$, connect algo = 1



Comparison of mSEIR and basic ABN models

Infection time distribution

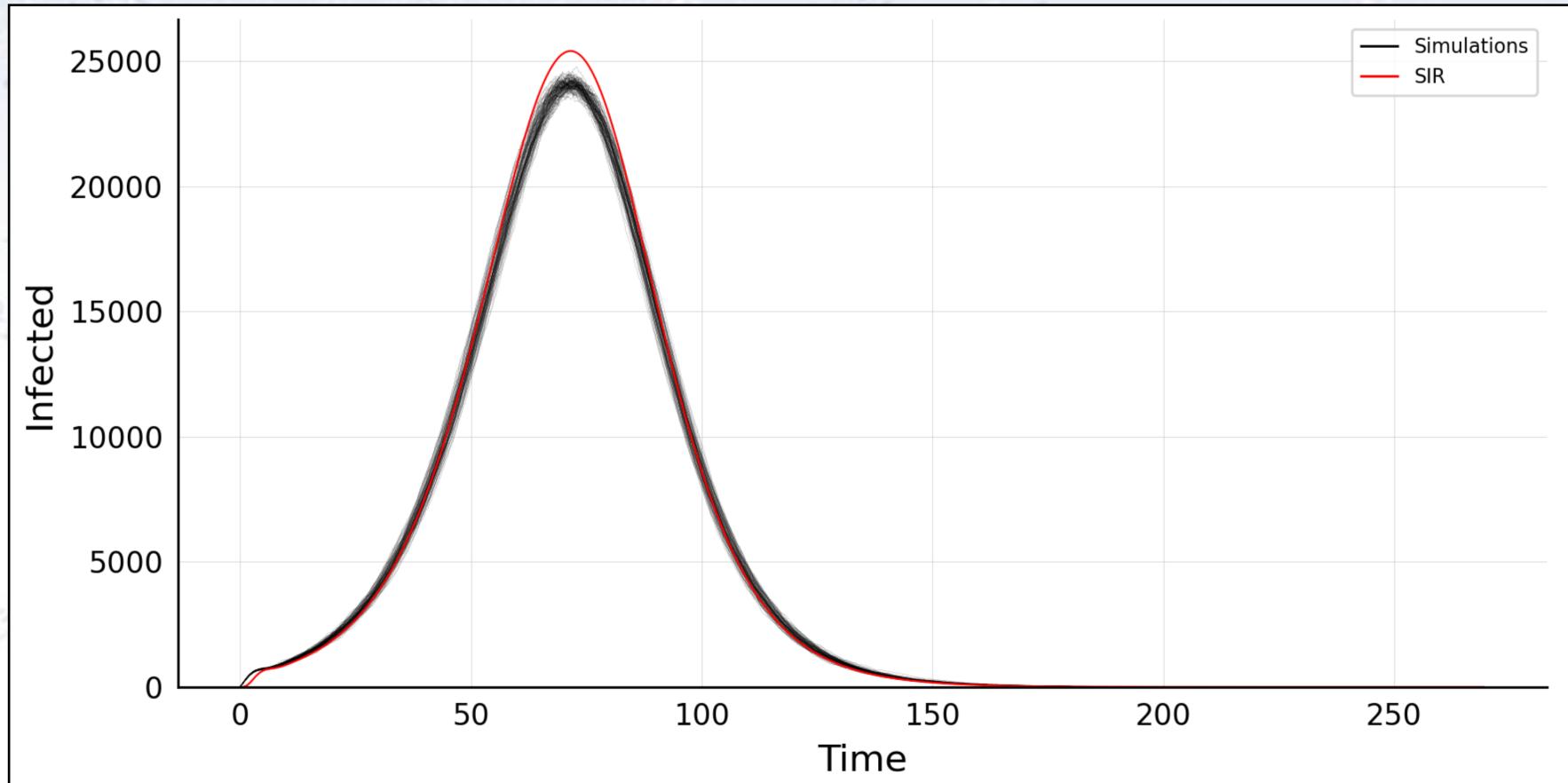
As can be seen below, the mSEIR and basic ABN models correspond very well.



The small time delay we believe are due to the continuity of the mSEIR model.
The peak falls 5% below while the integral is the same within 1.5%.

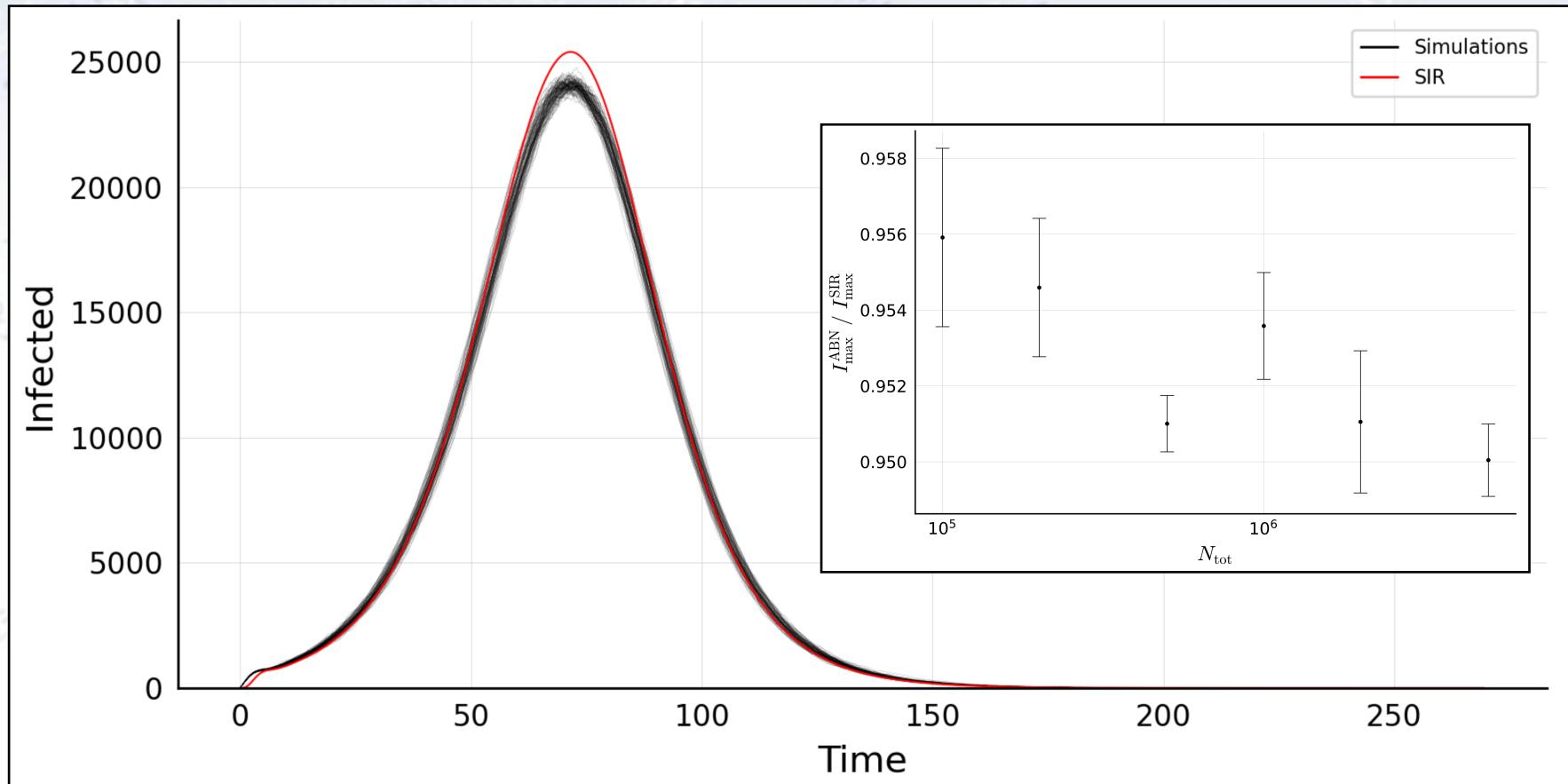
Fluctuations in ABN model

The fluctuations are a result of the number of initially infected ($N=100$ default). However, if we start with 1000 infected, they diminish as they should.



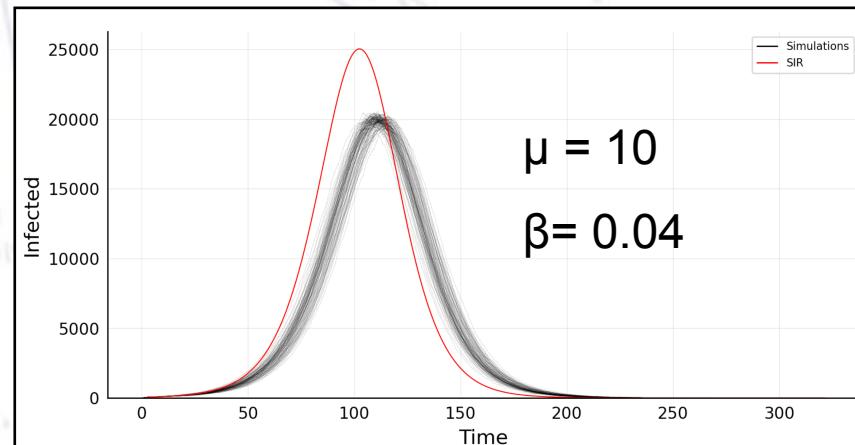
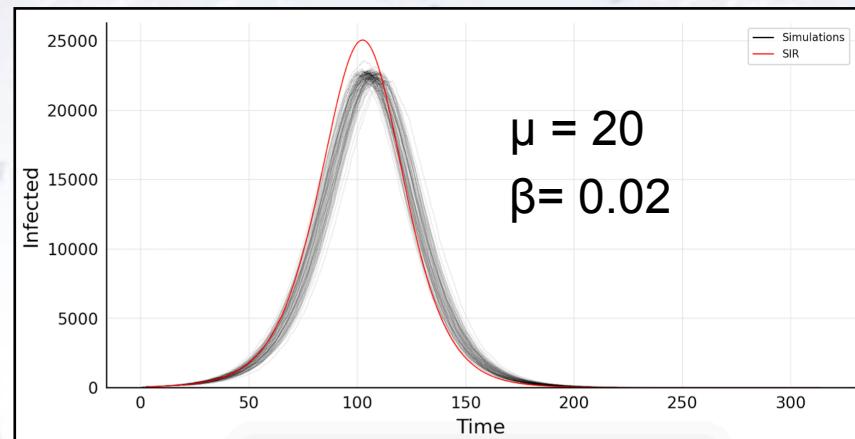
Fluctuations in ABN model

The fluctuations are a result of the number of initially infected (N=100 default). However, if we start with 1000 infected, they diminish as they should.



Fluctuations in ABN model

As we keep the theoretical R_0 constant, but decreases the mean connections, we get a reduction in the number of infected in the ABN model



Introducing inhomogeneties

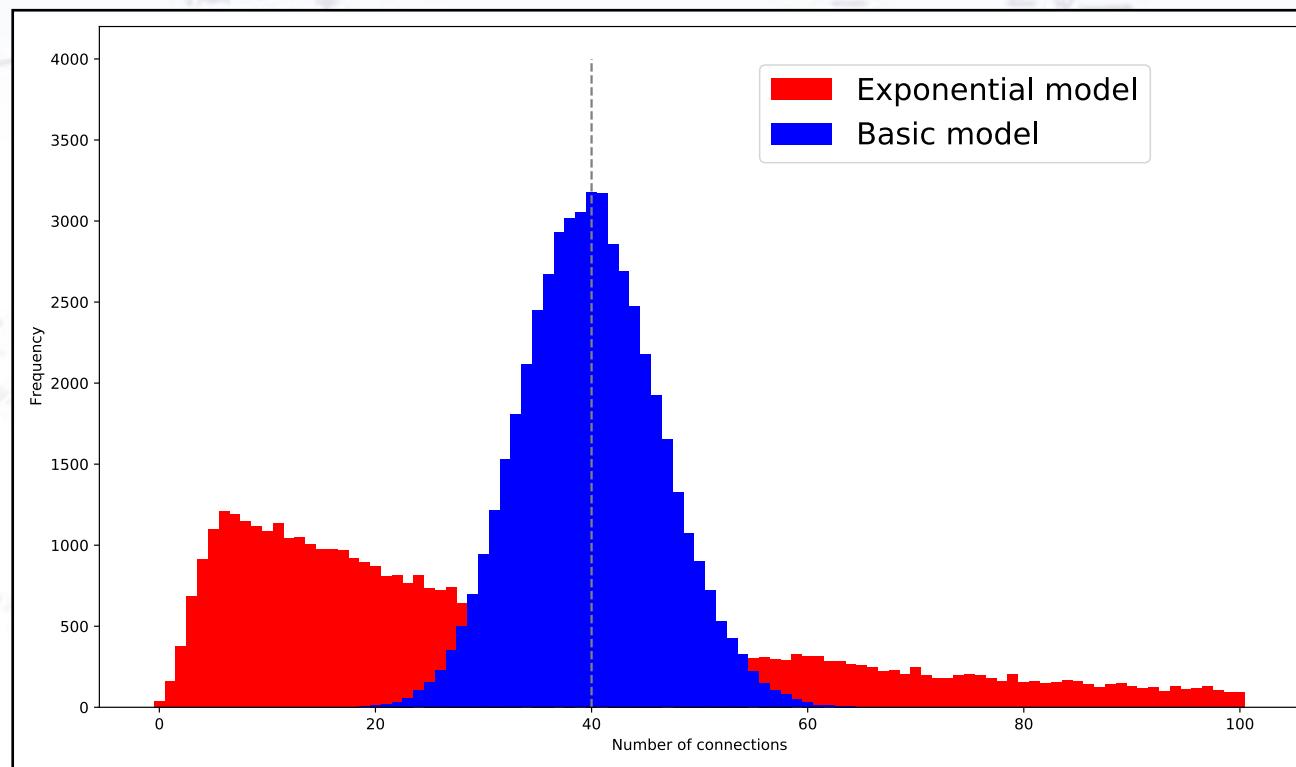
Exploring effects of connection and infection inhomogeneities

Connection inhomogeneity

First we introduce **connection inhomogeneity**:

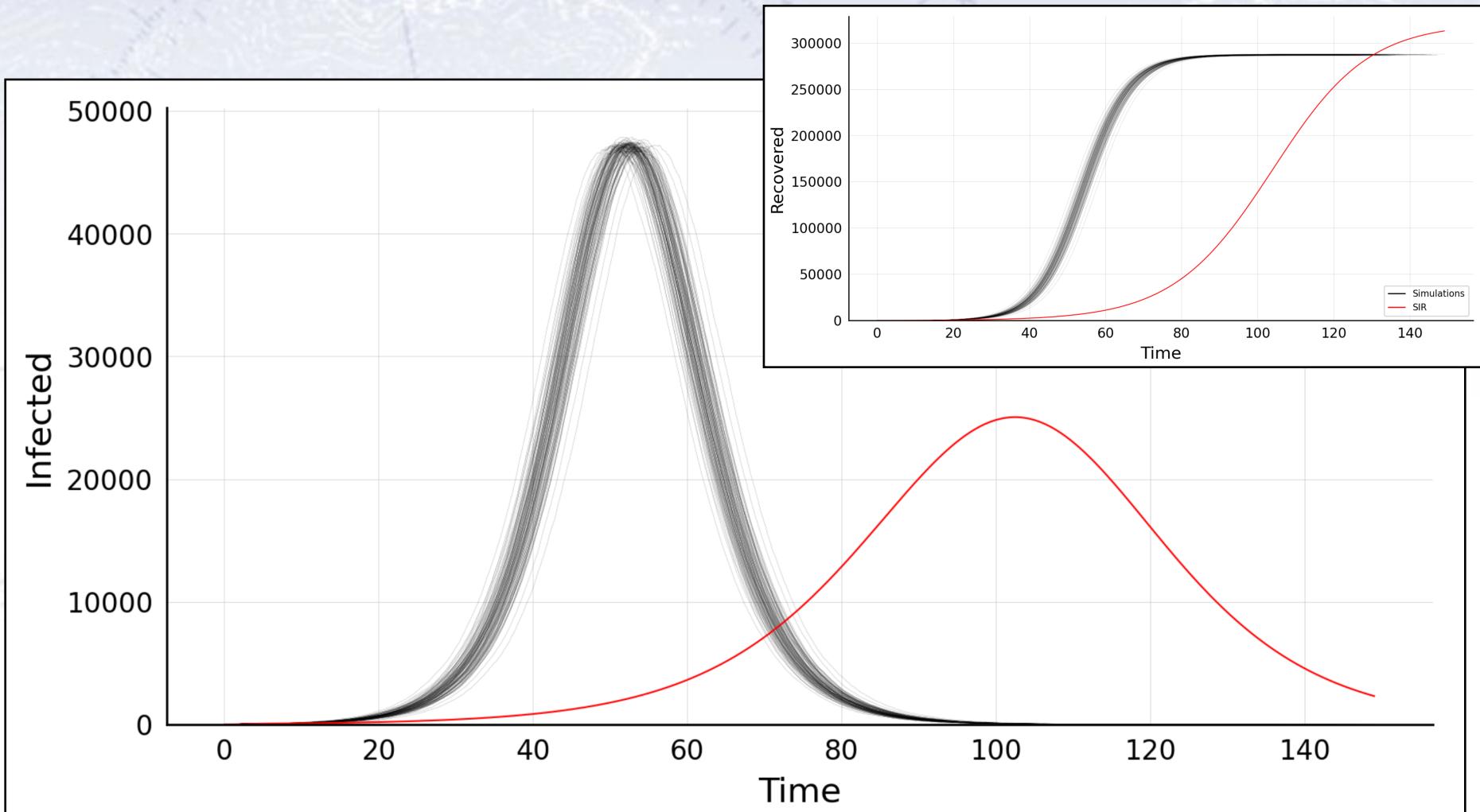
This is done by assigning non-constant weights to each node. Now, there are some nodes, that are more popular / attractive than others.

We have chosen to use shifted exponentially distributed weights, inspired by mobile phone data from the sensible-DTU project.



Connection inhomogeneity

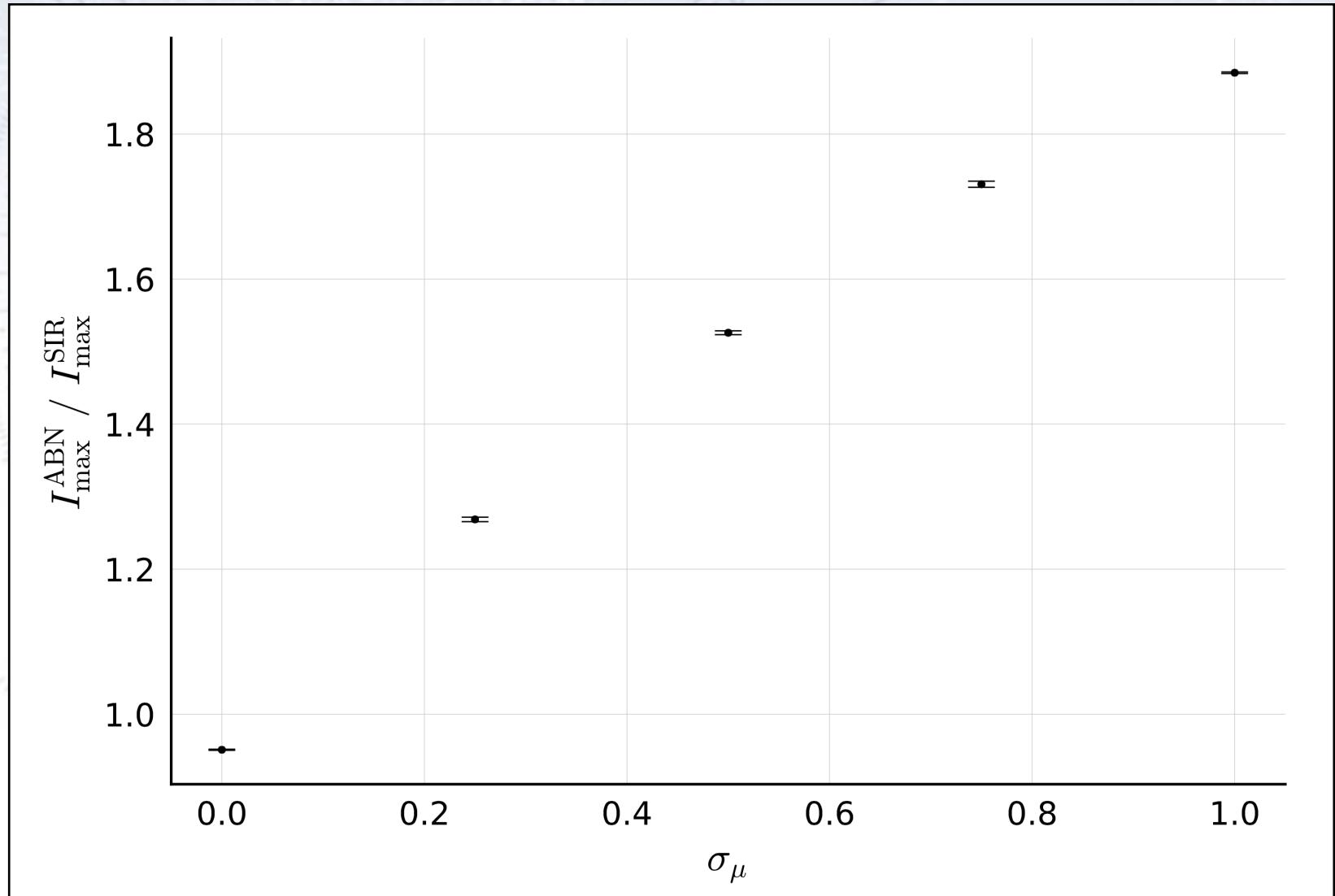
The result of choosing an inhomogeneous number of connections is as follows:



The effect is VERY clear: Super-spreaders ($N_{\text{conn}} \gg \mu$) rule the dynamics.

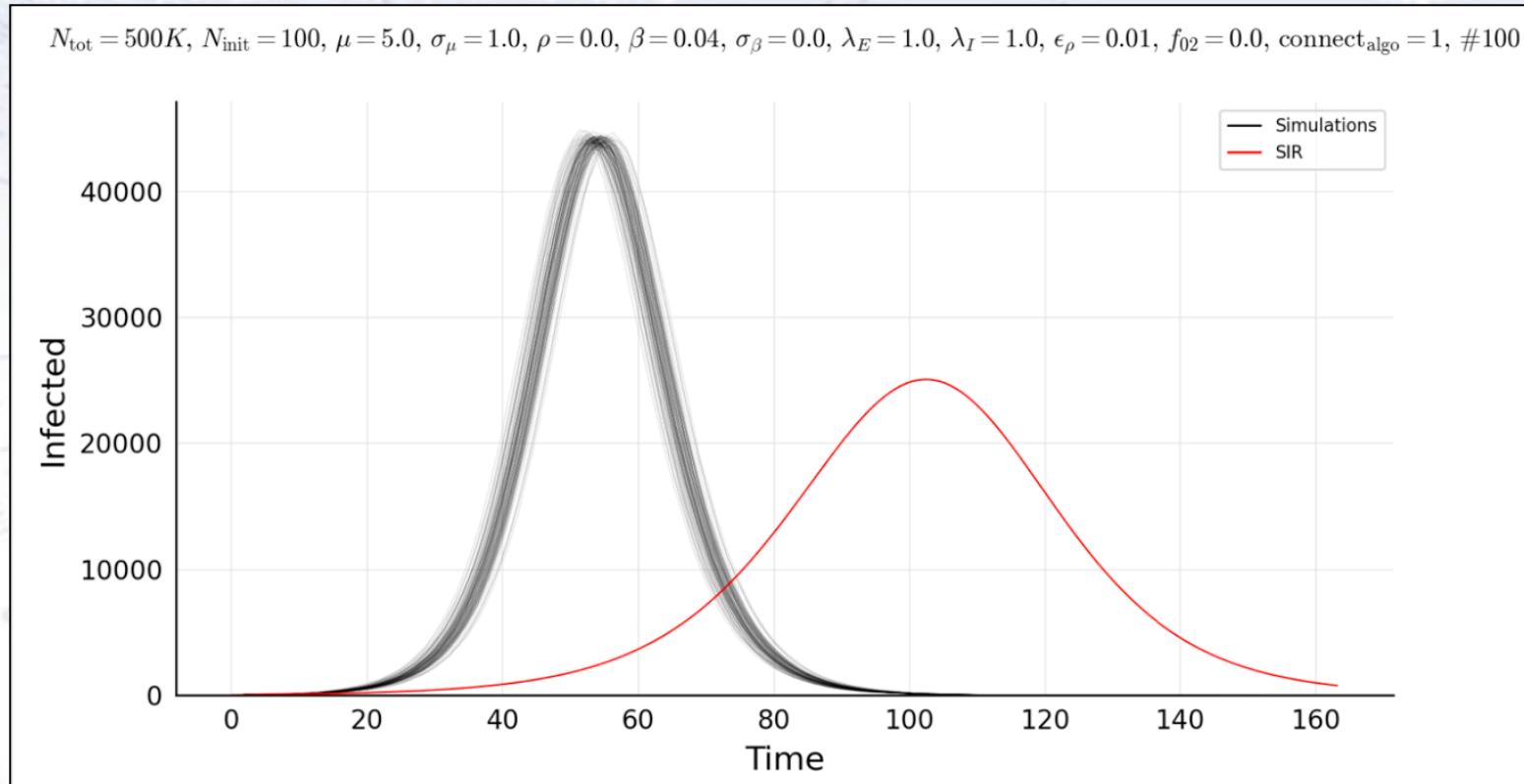
Connection inhomogeneity

Just to see the effect, we interpolate between the two models, and get:



Connection inhomogeneity

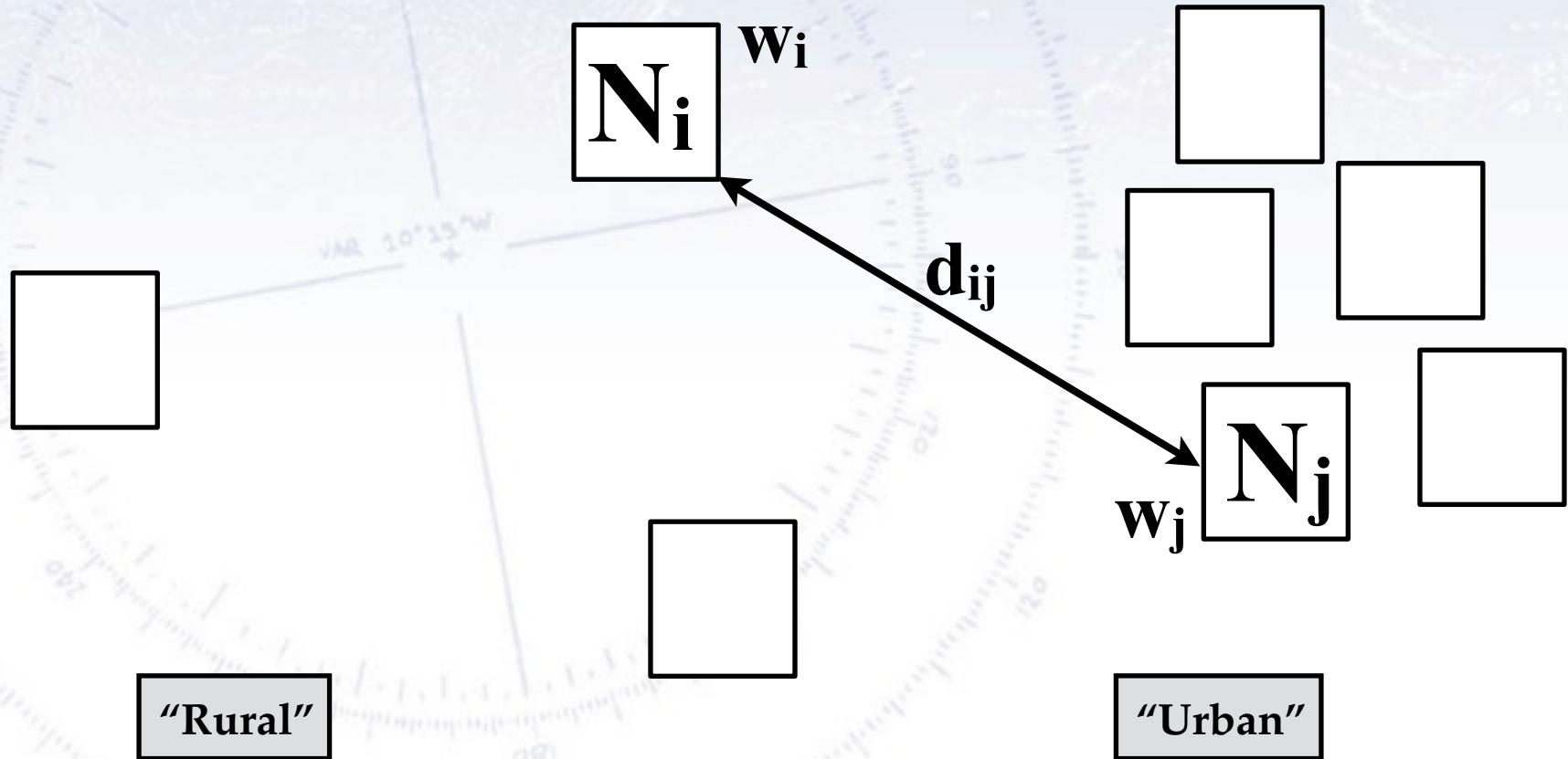
Just to see the effect, we interpolate between the two models, and get:



Infection inhomogeneity

In addition, we introduce infection inhomogeneity:

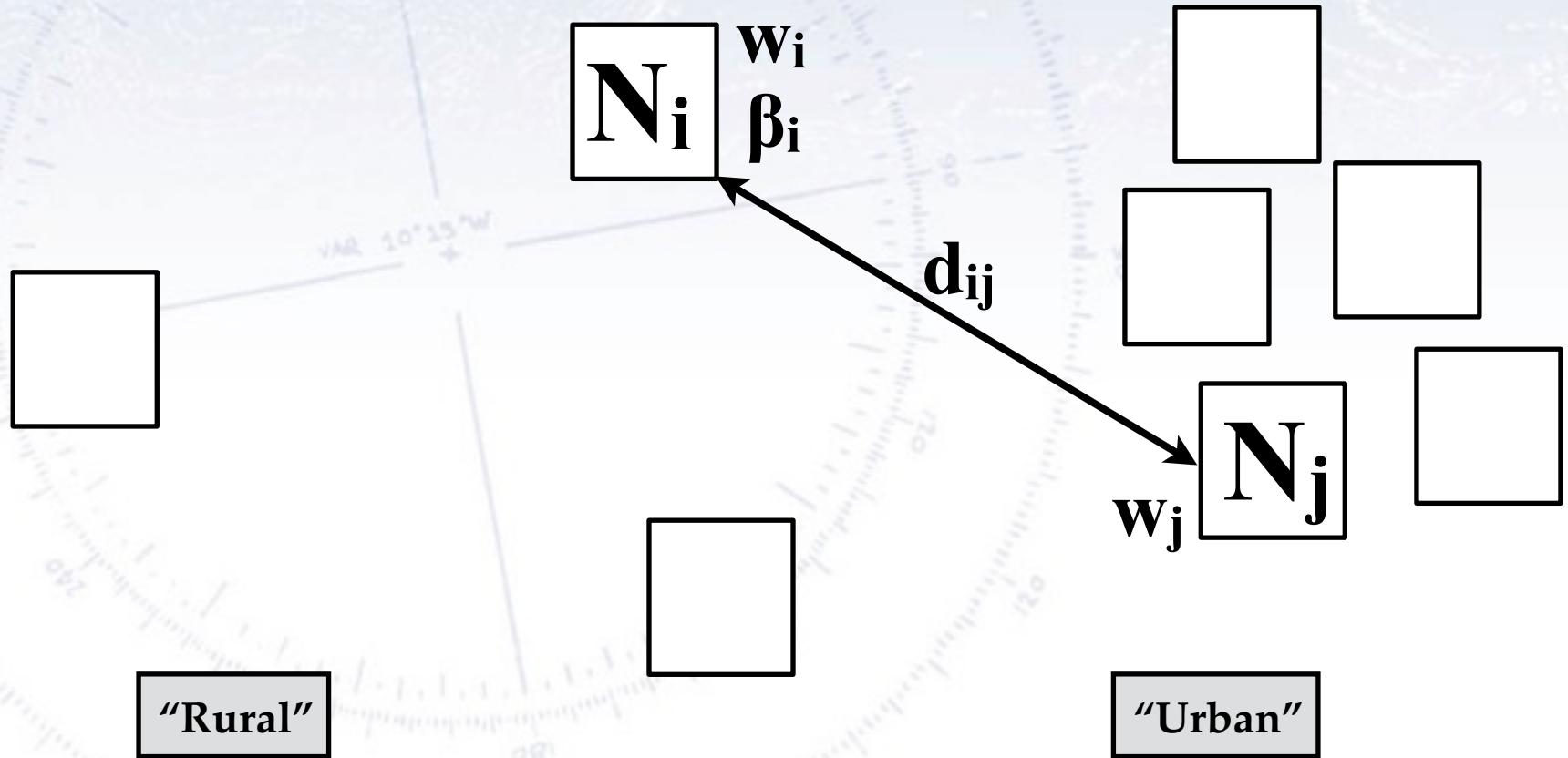
This is done by assigning different infection rates, β_i for the infector:



Infection inhomogeneity

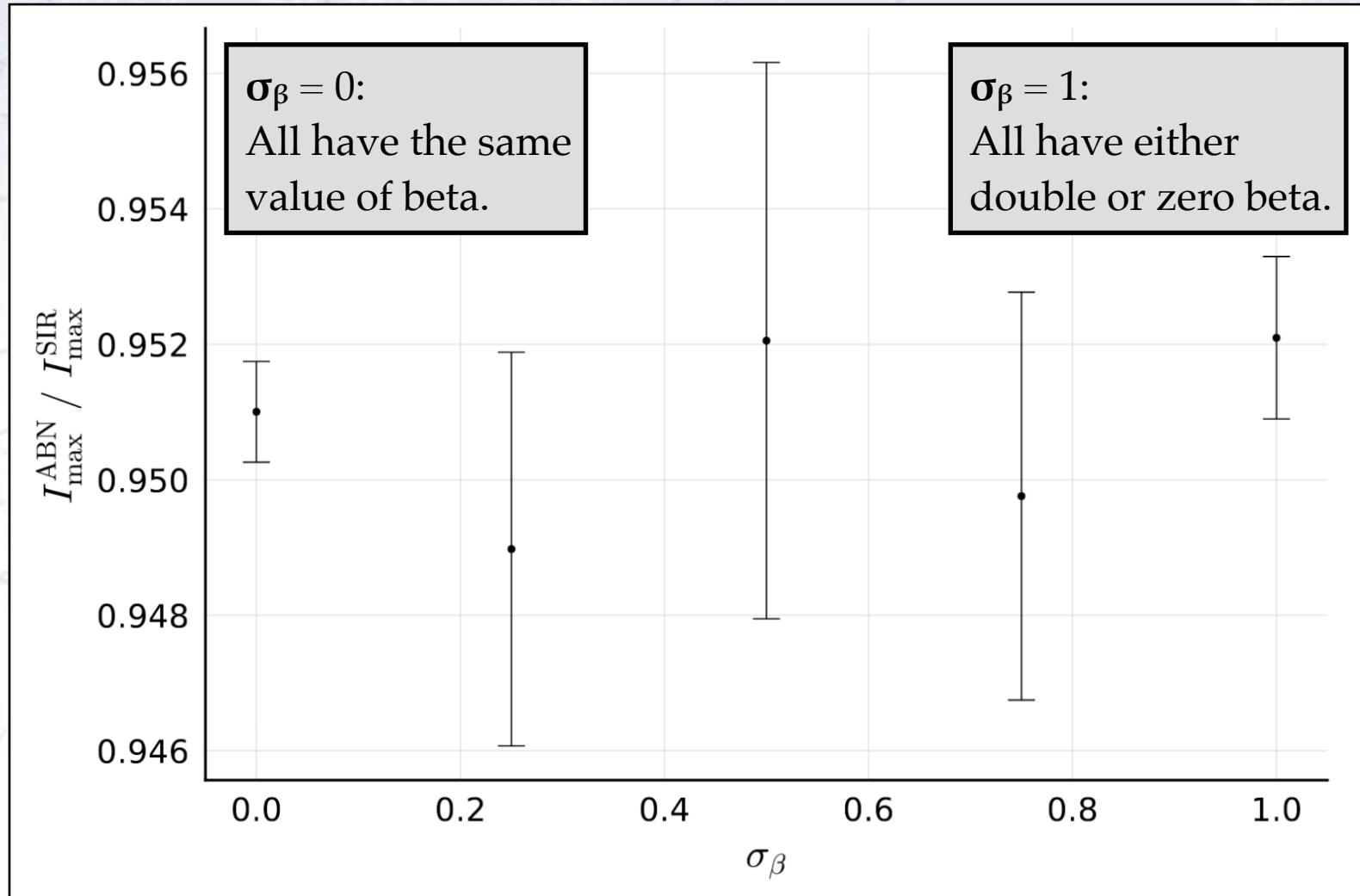
In addition, we introduce infection inhomogeneity:

This is done by assigning different infection rates, β_i for the infector:



Infection inhomogeneity

To our mild surprise, there is no effect of this! “Super shedders” play no role!



Exploring effects of spatial inhomogeneities

Spatial inhomogeneity

Secondly we introduce spatial inhomogeneity:

This is done by choosing a value of ρ , which controls the probability of a connection with distance: $p(d_{ij}) = \exp(-\rho d_{ij})$

A value of $\rho = 10$ (0.001 km^{-1}), means that the probability of connecting falls exponentially with typical distance $1/\rho = 100 \text{ km}$.

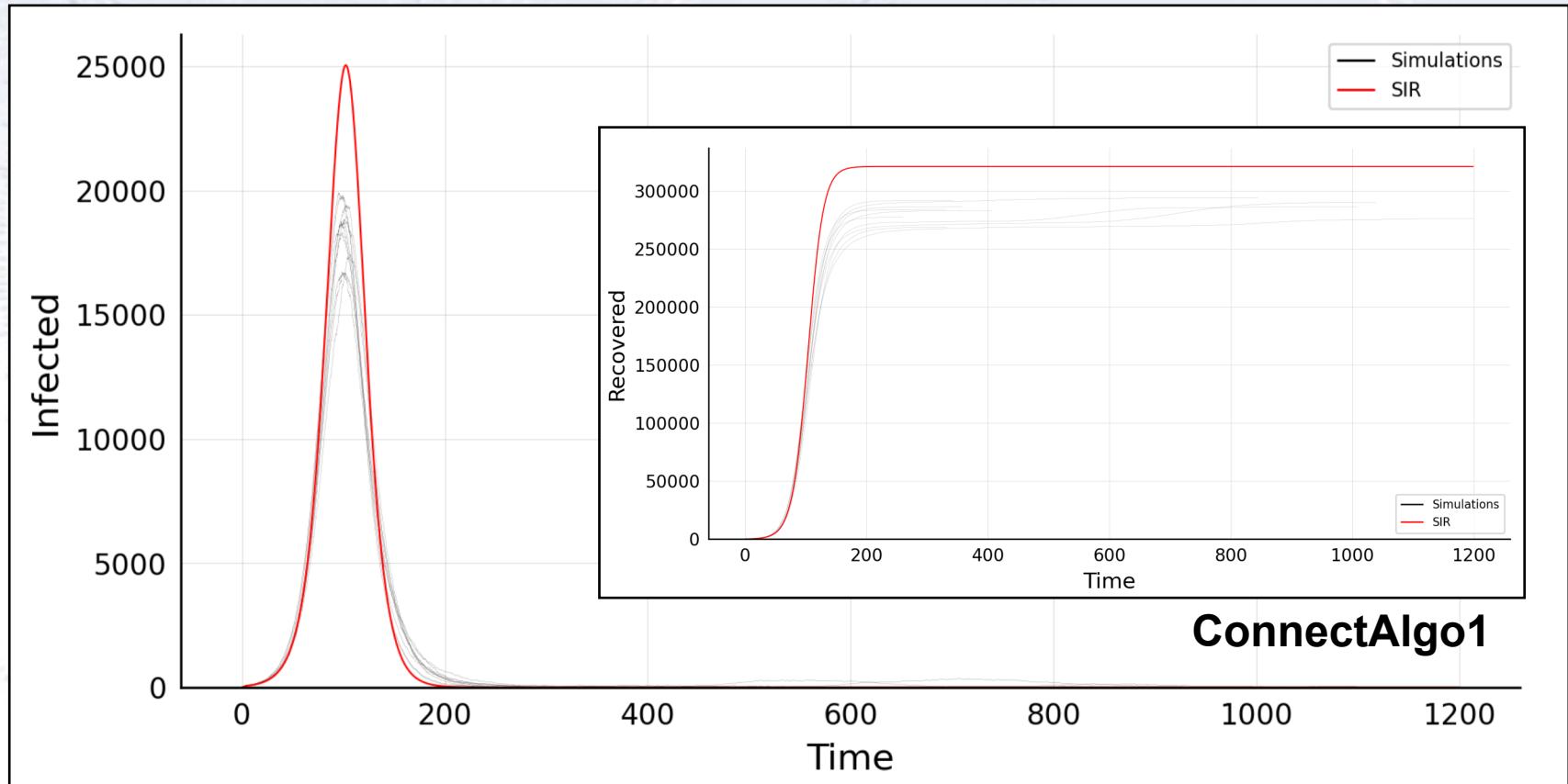
However, 1% of connections are independent of distance (reflecting that most know “someone”, who lives “far away”).

We have two models, if two random nodes are not connected:

1. Only exchange second node: All nodes “similarly” connected
2. Consider two new nodes: Urbans get much more connected

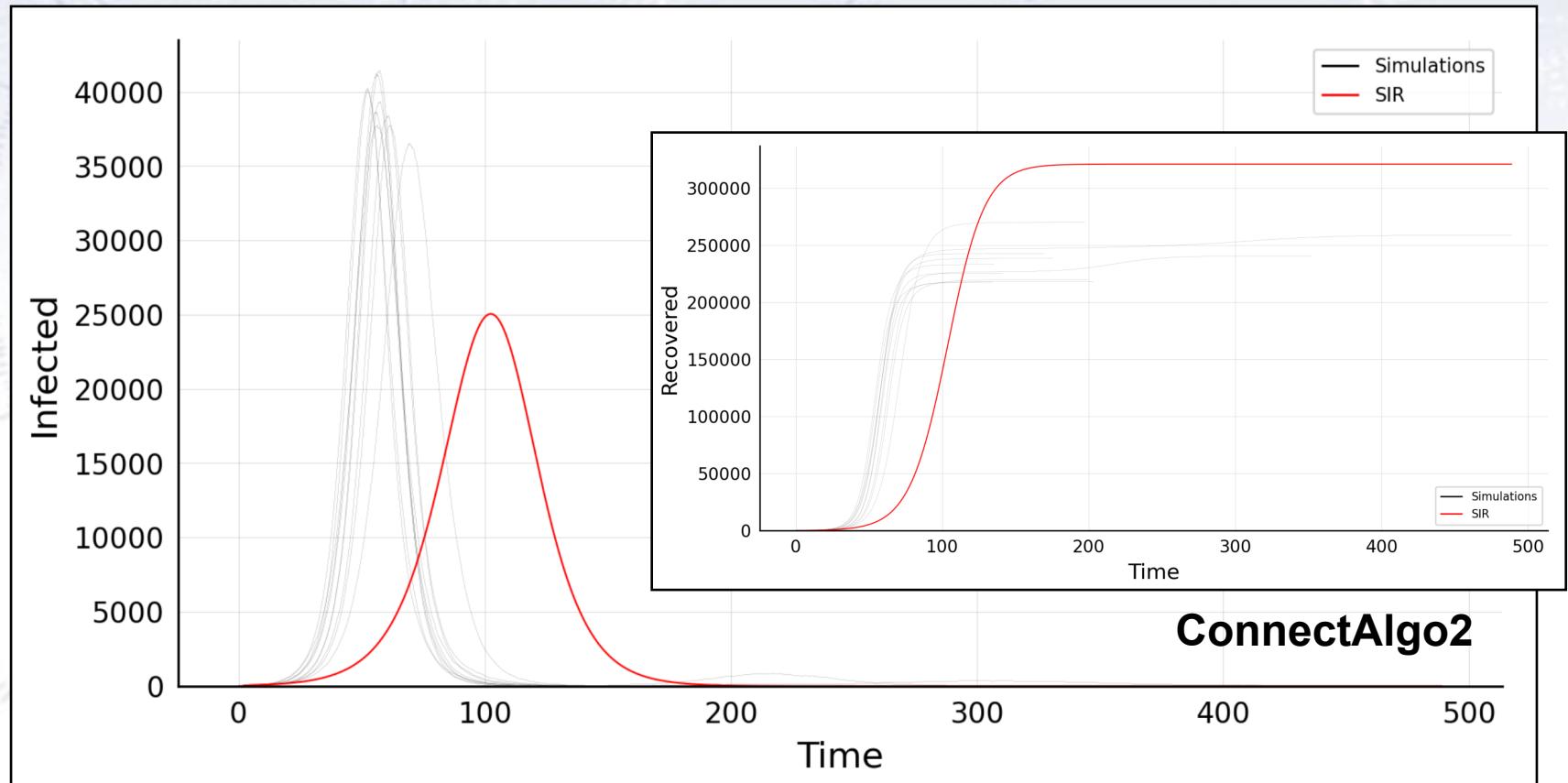
Spatial inhomogeneities

The result of choosing $\rho = 300$ (thus 3 km typical connection distance) with ConnectAlgo1 is:



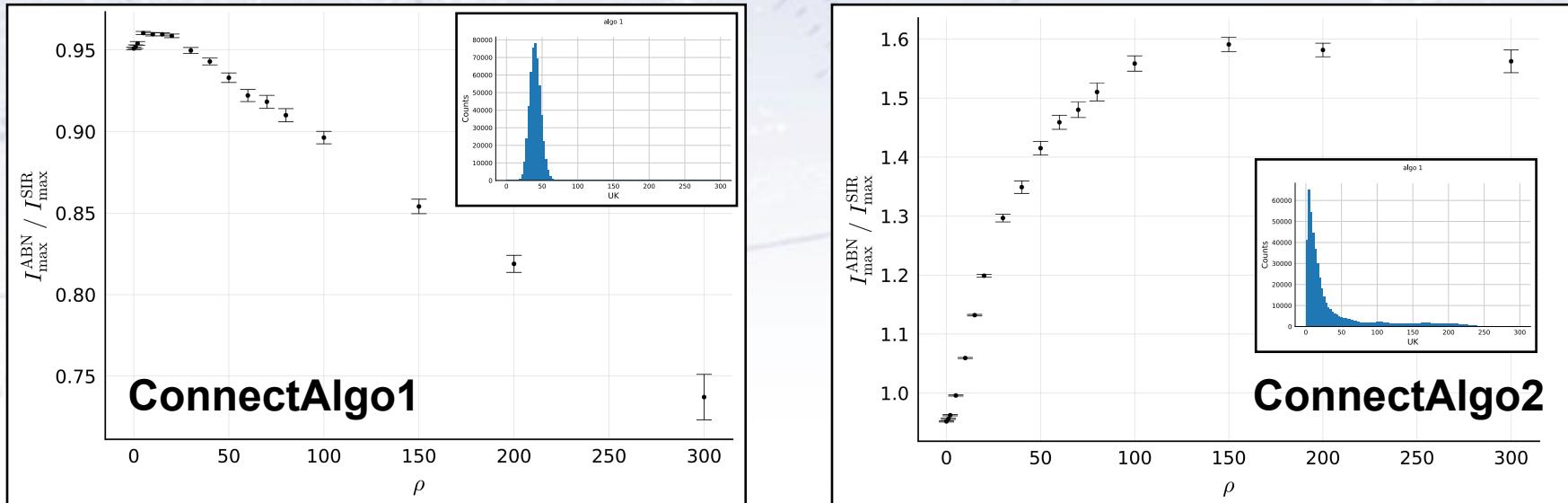
Spatial inhomogeneities

The result of choosing $\rho = 300$ (thus 3 km typical connection distance) with ConnectAlgo2 is:



Spatial inhomogeneities

The result as a function of ρ and connection algorithm is interesting:



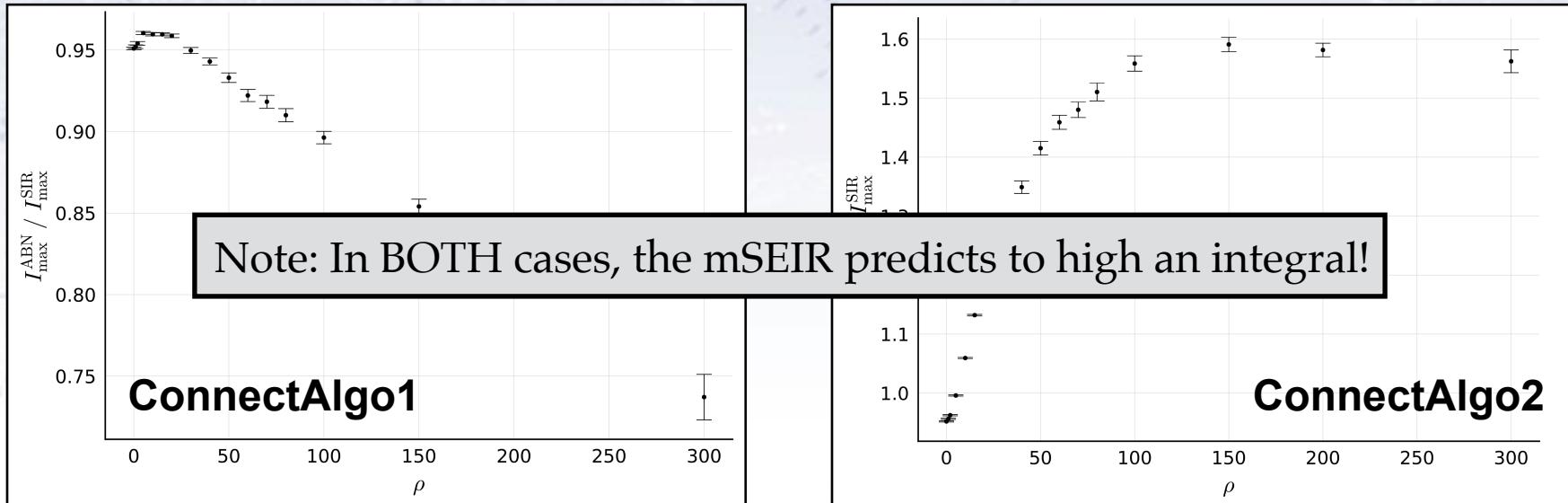
If we insist on connecting rural and urban population “similarly” (ConnectAlgo1), then we see that the ABN predicts a lower peak than the mSIR.

If we simply connect people according to distance (ConnectAlgo2), we see a much larger effect in the opposite direction!

We think that ConnectAlgo2 is more realistic, but not entirely realistic. However, this algorithm “mixes” the effects of distance and number of connections.

Spatial inhomogeneities

The result as a function of ρ and connection algorithm is interesting:



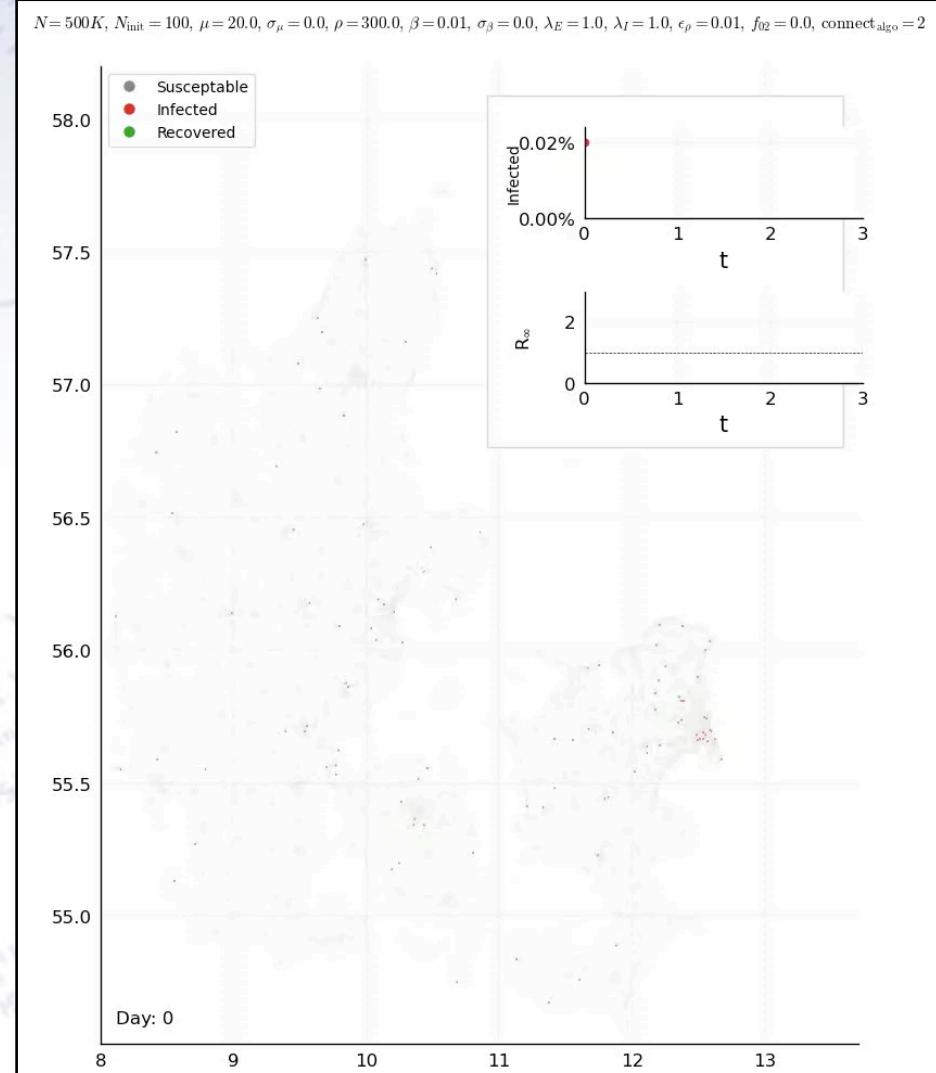
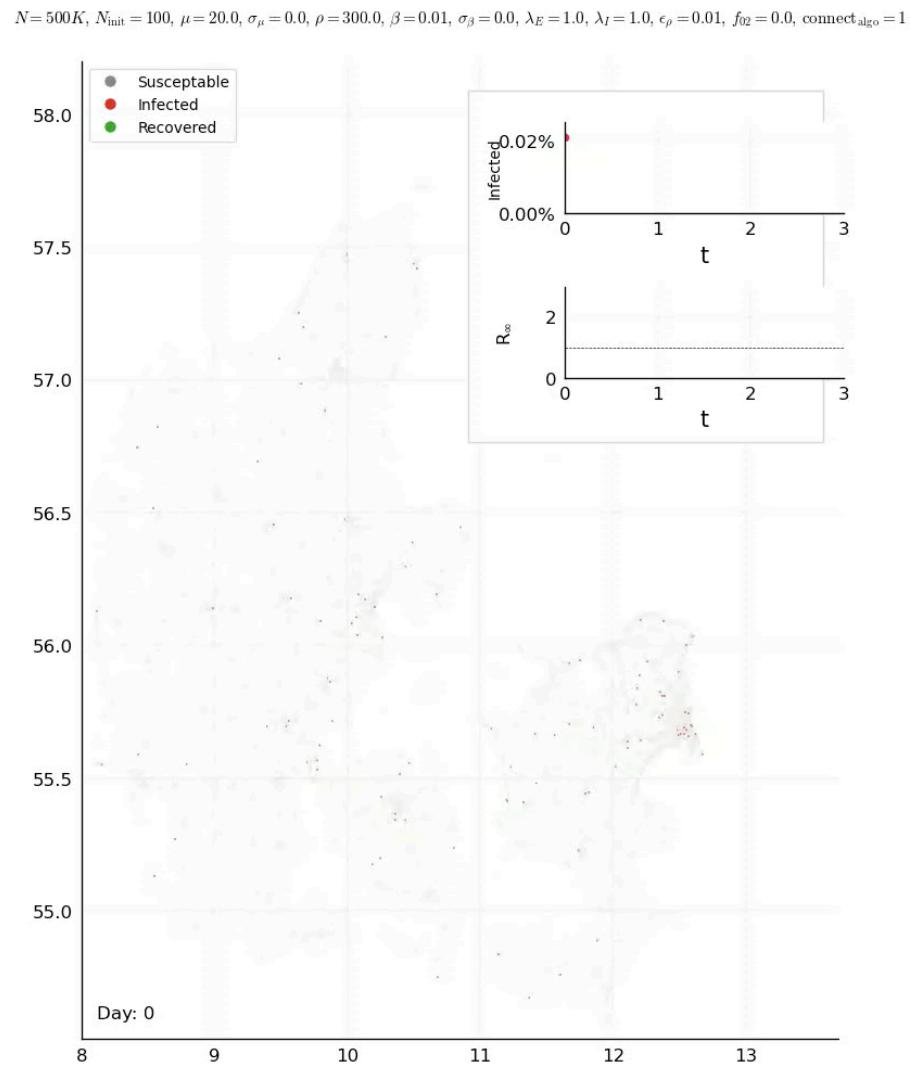
If we insist on connecting rural and urban population “similarly” (ConnectAlgo1), then we see that the ABN predicts a lower peak than the mSEIR.

If we simply connect people according to distance (ConnectAlgo2), we see a much larger effect in the opposite direction!

We think that ConnectAlgo2 is more realistic, but not entirely realistic. However, this algorithm “mixes” the effects of distance and number of connections.

Spatial inhomogeneity visually

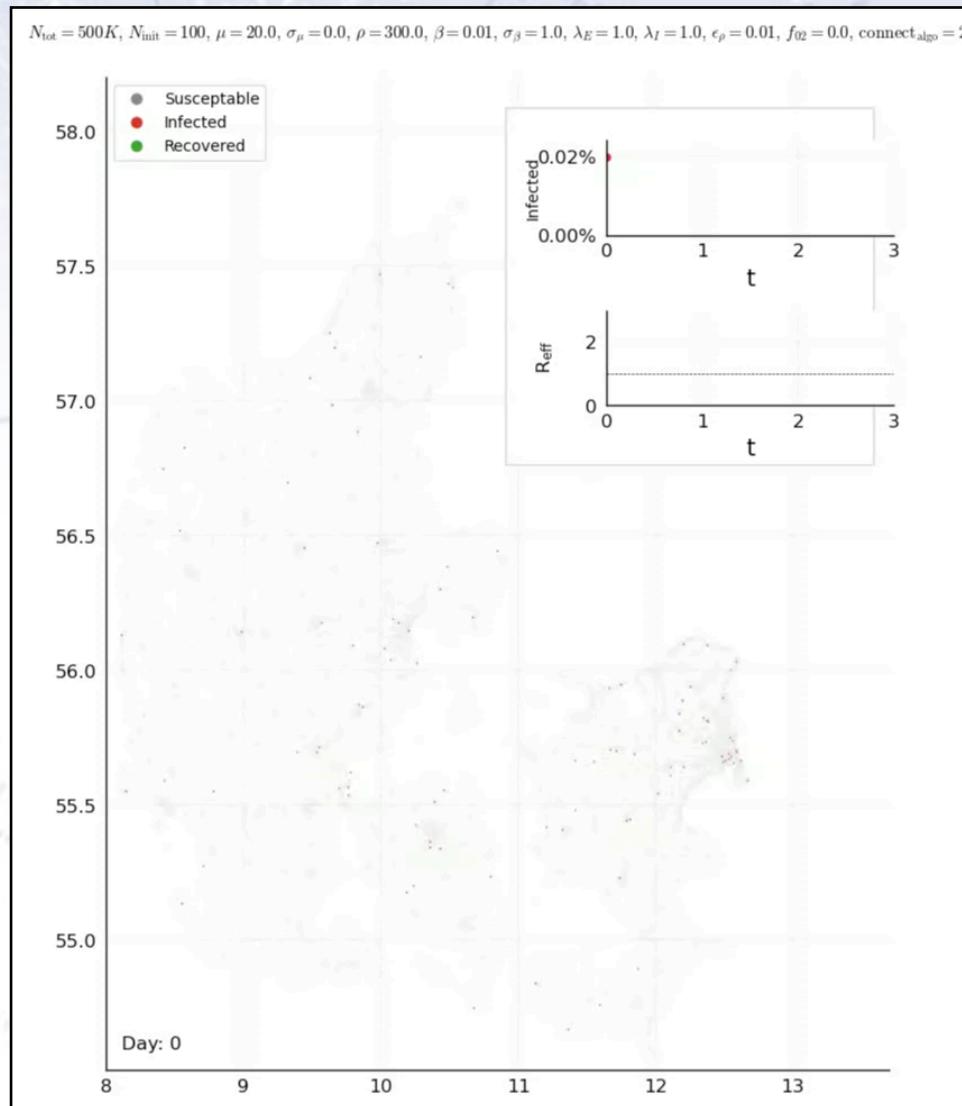
The spatial effects are best seen visually, as shown below:



Note: In ConnectAlgo2, there is no 1% connections distributed with no distance dependency. However, this should have little effect.

Spatial inhomogeneity visually

While the rates do not affect the overall outcome they do affect locally

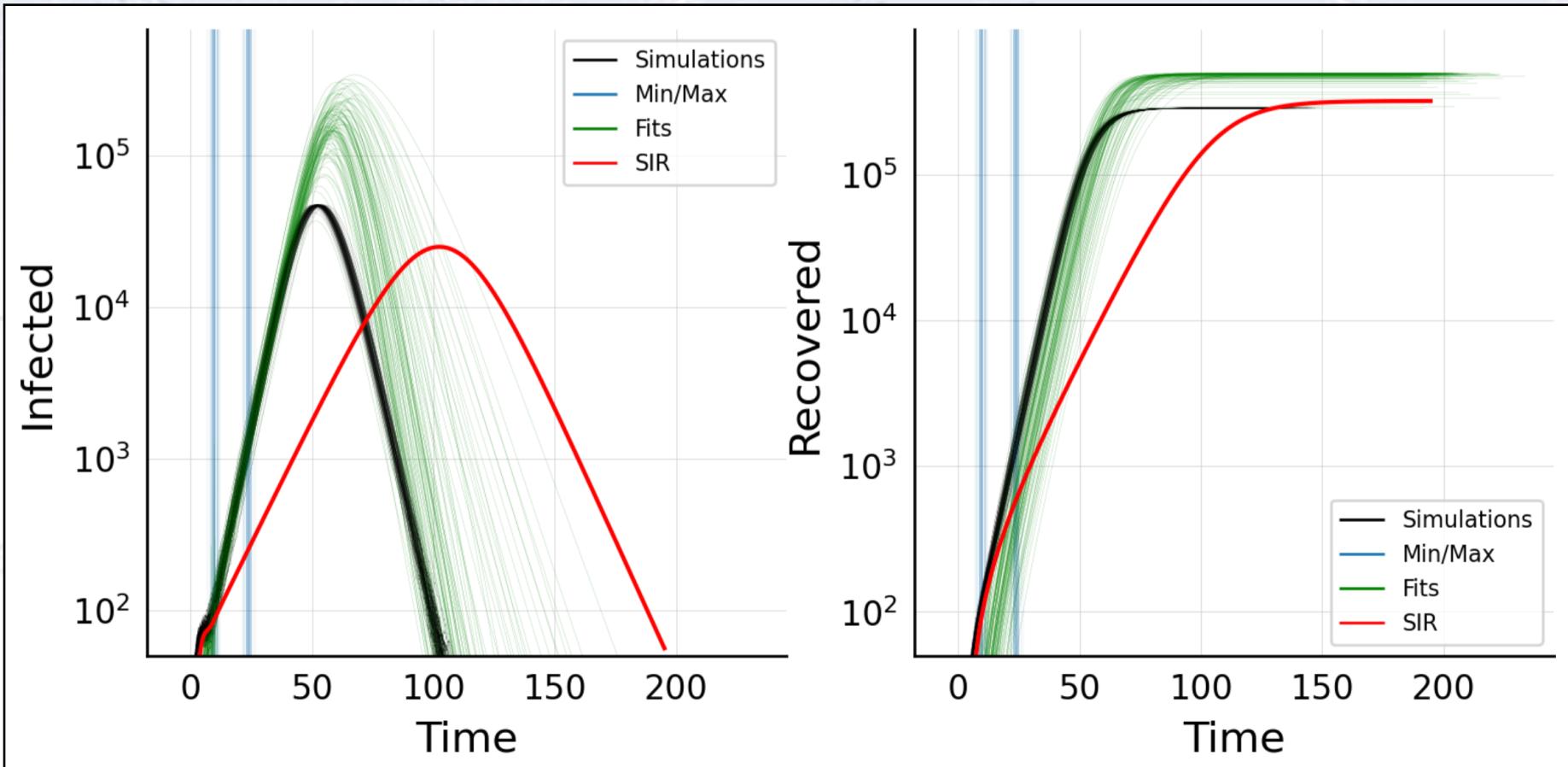


Note: In ConnectAlgo2, there is no 1% connections distributed with no distance dependency. However, this should have little effect.

Early prediction error

Early prediction power

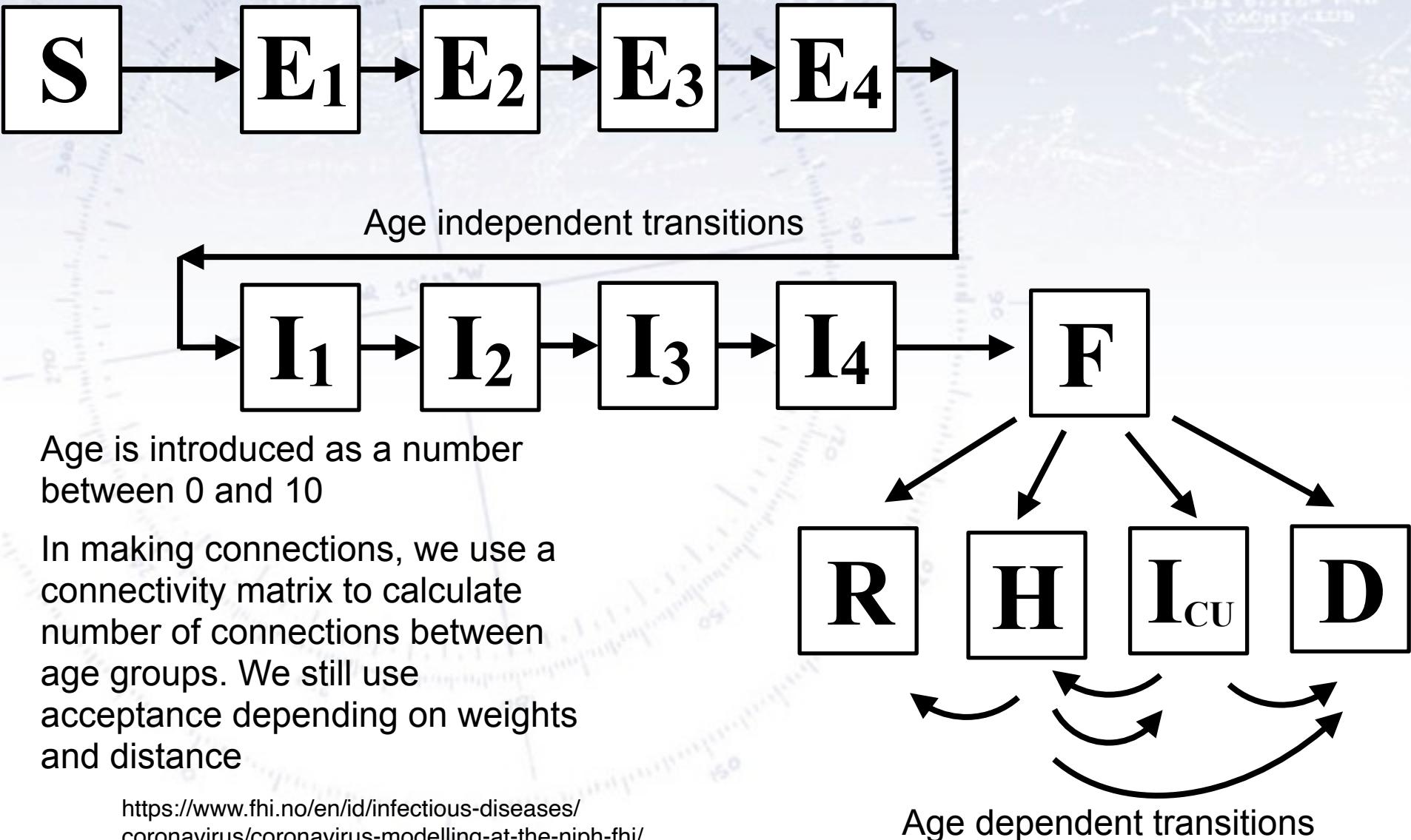
Given early data (before peak), how well does the mSEIR model predict peak size (as simulated in an ABN), assuming we have perfect infection data?



We have tried to answer this, by conducting fits based on data from the blue interval. Here, dark gray is the ABN model, while green are the mSEIR fits.

Practical implementations - not presented here

Age & Hospitalisation data



Fit with ABN model?

A core element in the practical implementation in the model is to include actual data of hospitalised patients.

To overcome this we estimate the effects of inhomogeneities, and include this as a multiplication function to the initial curve.

Another way to do this, is to run the ABN model multiple times to fit the data. To estimate the effects of political actions this is the way to proceed.

How to include actions?

Another core element in the practical implementation is the ability to predict the effect of political decisions.

With the current setup, we can at any time remove links between nodes.

This means that if we consider a specific type of job, we can estimate the effect of this by taking the number of nodes resembling the number of persons working in this section, and removing a given number of connections.

We also have the possibility to insert divide people in boxes, family, work and others. The given infection rates can then be grouped as well.

Questions from our side

What resources could be available for this? Computer memory might be a limiting factor.

What is the best way to move now?

What do you see as a timeline?

Conclusions

We have so far managed to produce a working Agent Based Network model, which can potentially incorporate a large number of effects and their interplay. **It is fast, reasonably realistic, and quite general.**

Research results:

Investigating the effects of inhomogeneties, we have found that the distribution of number of connections plays the most important role, followed by spatial effects. Generally, the mSIER model gives too high prediction / result compared to the ABN, in some cases by a factor 2-3. Shedders essentially play no role. Seemingly, there is basis to believe that herd immunity requires fewer infected.

Practical results:

The ABN model provides a way to include essentially all known effects. This provides a tool for investigating the impact of (political) actions, provided accurate input data.

The animations give a further tool for investigating effects, and provide a way to communicate with.

Bonus Slides

Estimate of infected

Danish estimate (so far) is based on $12/1071 = 1.12 \pm 0.32$ (68%) infected persons.
This yields a **30% statistical uncertainty!**

Consider the numbers from Stockholm Region:

Testing 1200/week for 8 weeks, the result was 7.3% (May 3rd in week 18, but estimate is then for week 15). Given about 9600 tested with 701 positives, this yields 7.30 ± 0.28 (68%) infected persons, thus 4% uncertainty!

With the number of deaths in Stockholm region at 1,855 people (by 19th of May), compared to the Danish number around 530, this would indicate:

$$(7.30 \pm 0.28) \times (530 / 1855) = 2.09 \pm 0.08$$

This calculation does not include test error rates nor precise timing corrections. But my guess is that the real number is higher than 1.1%

Swedish study also shows age variations: 0-19: 4.7%, 20-64: 6.7%, 65-70: 2.7%.

Norway: 0.58% - 0.73%... how many did they test and when?

Philosophy of modelling

Overall philosophy:

Given “sparse” data, we wanted to build a **simple** model with few assumptions.

On model parameters:

We wanted to extract model values from **fitting** available (and unbiased) data.

Measuring development:

Fitting data to a model, we hope to get **the best sensitivity to development**.

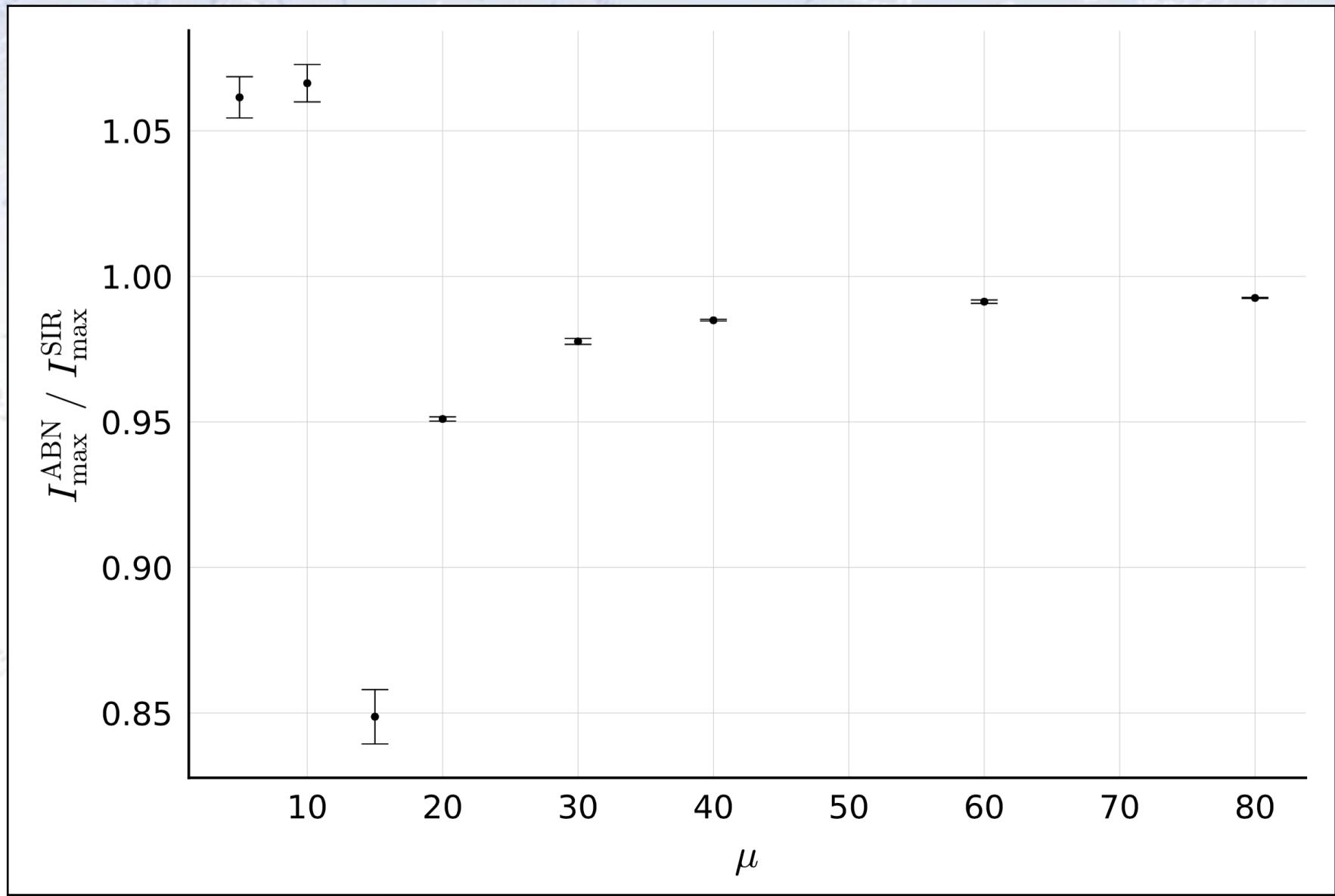
Predictive power:

Provide predictions of number of hospitalisations, intensive care, respirators, etc. both overall, but also for each region in Denmark, 1-2 weeks into the future.

Overall aim:

Contribute best we can with epidemic insights for planning/decision taking.

Peak ratio vs. N connections



Background and warranty

Limited warranty: We are NOT virology experts!

However, we are very accustomed to analysing complex data, and providing insights and estimates with uncertainties. This is a first (very preliminary) presentation of our attempt to model the Danish Corona epidemic.



Troels C. Petersen
Associate Prof. (NBI)

Worked with statistics,
simulation, modelling,
fits, etc. for many years
(in particle physics).



Christian Michelsen
Ph.D. student (NBI)

Worked with statistics,
simulation, and data
analysis. Very capable
with programming.



Mathias Heltberg
PostDoc (Paris/NBI)

Worked with statistics,
simulation, networks,
biocomplexity, etc.

Improved model

This should contain “the full Monty”, which we in principle have ready! But...
...it requires better input on probabilities and distributions, than we have right now!

