Time Series Models in Epidemic Forecasting

Salvador Razatos

Institute for Computing in Research

July 28, 2025

Summary

Introduction

The goal of the research that I conducted was to determine if time series models are effective effective epidemic forecasters, and if using synthetic data generated from agent based models could potentially increase the accuracy of these models.

SIRS Model

Simple deterministic model, that works by integrating a set of differential equations. This will be used as a baseline.

$$\frac{dS}{dt} = \upsilon N - \beta \frac{SI}{N} + \xi R - \mu S \tag{1}$$

$$\frac{dI}{dt} = \beta \frac{SI}{N} - \gamma I - \mu I \tag{2}$$

$$\frac{dR}{dt} = \gamma I - \xi R - \mu R \tag{3}$$

Agent Based Models

Simulates Complex Behavior, as it calculates the movement of each individual agent, this will be used to generate data for the Times Series Model.

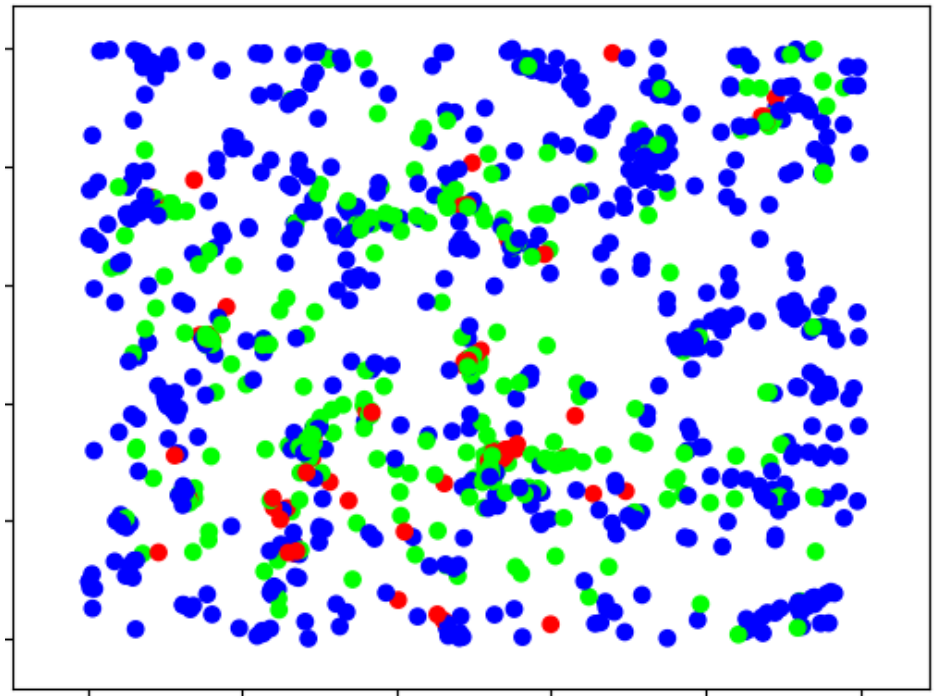

Figure 1: Agent Based Model Visualization

Time Series Models

Time series models are predictive neural networks that are trained on example data in an attempt to achieve an extremely high accuracy.

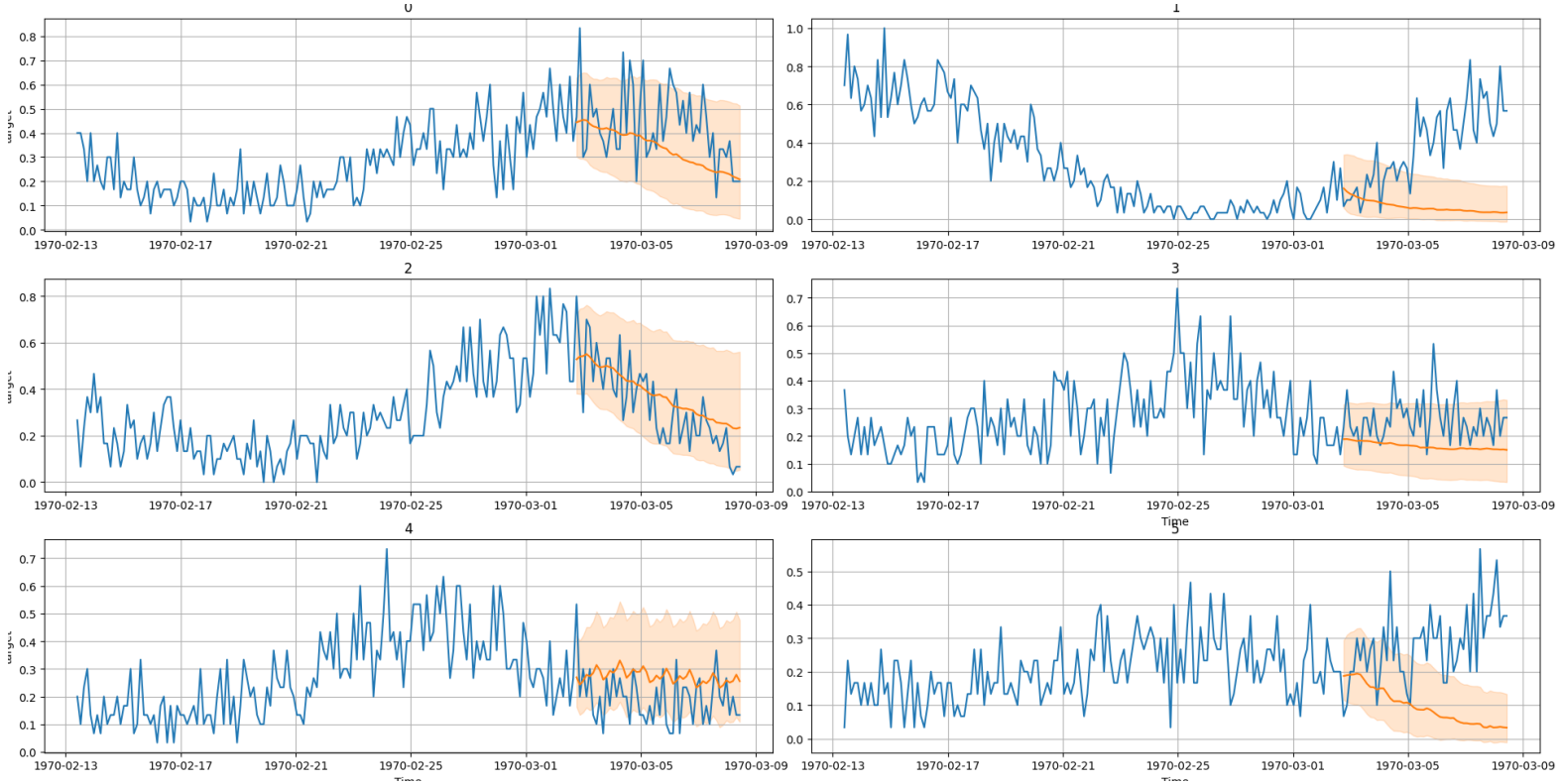

Figure 2: Zero Shot Model Predictions

Data Generation

Two hundred and fifty-five simulations were run on the agent-based model, with the data being split into three classes, each with 85 simulations. The difference between each of the classes was the clumping factor, which was set at 0.05, 0.25, and 0.5.

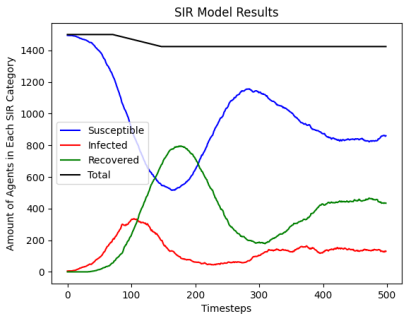

Figure 3: Clumping Factor of 0.05

Figure 4: Clumping Factor of 0.5

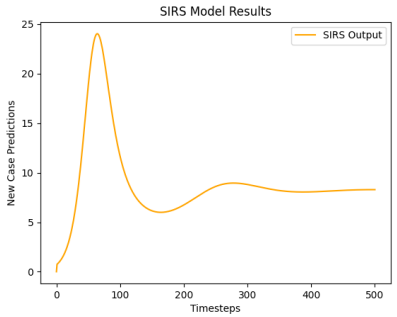

Figure 5: SIRS Output

Generating a Baseline with the SIRS Model

Overall, this model performed poorly with its predictions after the initial spike in cases, but occasionally aligned with some of the simulations

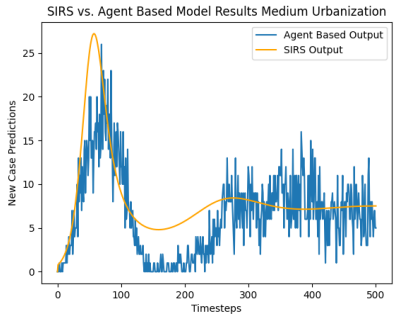

Figure 6: SIRS Loss Visual 1

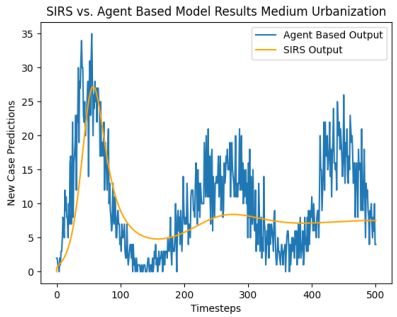

Figure 7: SIRS Loss Visual 2

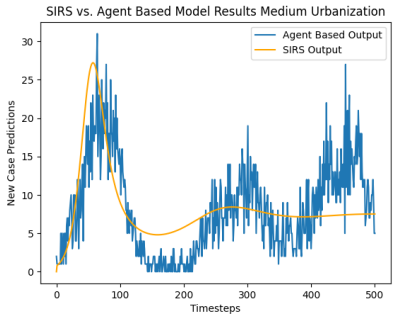

Figure 8: SIRS Loss Visual 3

Training the Time Series Model

I started by having the model make predictions before doing any training to get a baseline accuracy before beginning training. For this experiment, I trained four models, one in the low-urbanization data, one in the medium-urbanization data, one in the high-urbanization data, and one on all data.

Training Results, Part One

These are the results for the low urbanization training.

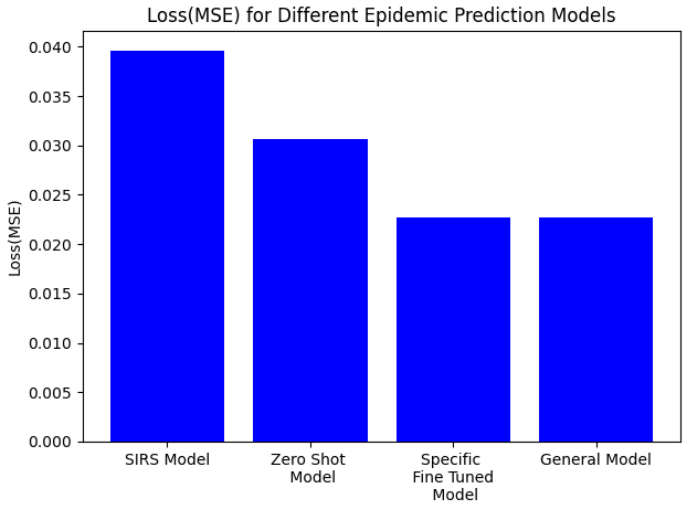

Figure 9: Low Urbanzation Loss for All Models

Figure 10: Chronos Low Urbanization Loss

Results Table

I compiled the results of the above bar graphs into a table, shown here.

Table 1: Compiled Results

Urbanization Level	SIRS	Zero Shot	Time Series	General
Low	0.0396	0.0306	0.0227	
Medium	0.064	0.025	0.0225	
High	40	65	153	