Accessible infectious disease forecasting

Group 135 : Linh Dinh, Vandana Rana, Naveen Vipin, Abrar Manzar, Jessica Manzione, Erik Magnusson

CSE 6242 Group 135

MOTIVATION / INTRODUCTION

Problem:

Quantitative dynamic models are useful in forecasting the spread of infectious diseases and allow planning and management of often scarce medical resources. However, widespread use of these models has the following challenges:

- **Technical know-how of the users, particularly in third world countries**
- **Easy availability of models for use**
- **Restrictions on shareability of data outside their location**

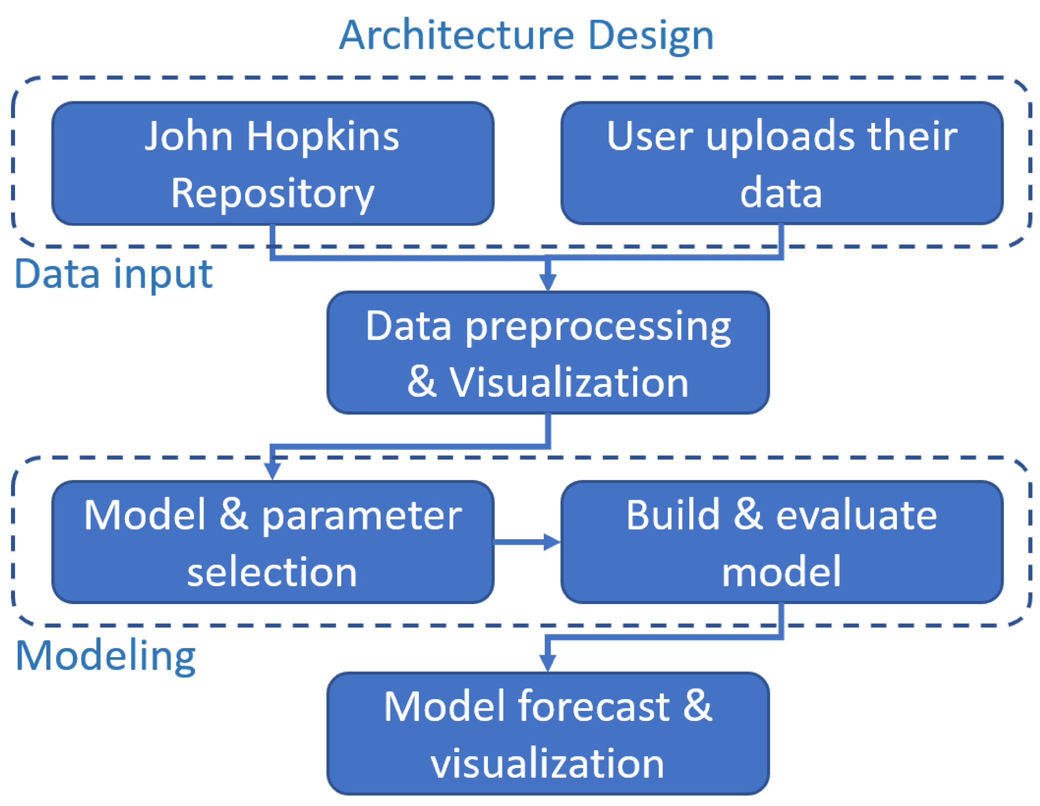
Solution:

Easily accessible forecasting models to improve awareness and allow more informed decisions by the users including **public health professionals, policy makers, scientific community and even general public**

APPROACH

Project Objectives:

1. Create an **easily accessible platform for disease modeling** allowing optional user data upload
2. Implement, **test and compare the selected models** with published results
3. Test, **identify limitations and recommend guidelines** for model usage



Modeling:

Most common options include **Deterministic models** which can be further grouped into:

- **Compartmental models** (modeling platforms exist e.g. COMO).
- **Phenomenological Growth models** (to our knowledge modeling platform with user data upload options do not exist)

Phenomenological Growth models are the main focus of this project. They use only a few parameters and approximate to standard compartmental models. These do not incorporate transmission mechanism of the disease but can be easily extended to almost all infectious diseases, making them more versatile. The current application includes following options for this family of models:

Exponential growth model	$\frac{dC}{dt} = rC_t$
Generalized growth model	$\frac{dC}{dt} = rC_t^p$
Logistic growth model	$\frac{dC}{dt} = rC_t(1 - \frac{C_t}{K})$
Generalized logistic model	$\frac{dC}{dt} = rC_t^p(1 - \frac{C_t}{K})$
Generalized Richards model	$\frac{dC}{dt} = rC_t^p(1 - (\frac{C_t}{K})^a)$

r = Intrinsic growth rate, p = Deceleration of growth
a = Scaling parameter K = Maximum cumulative case incidence
C_t = Cumulative incidence at time t

Implementation:

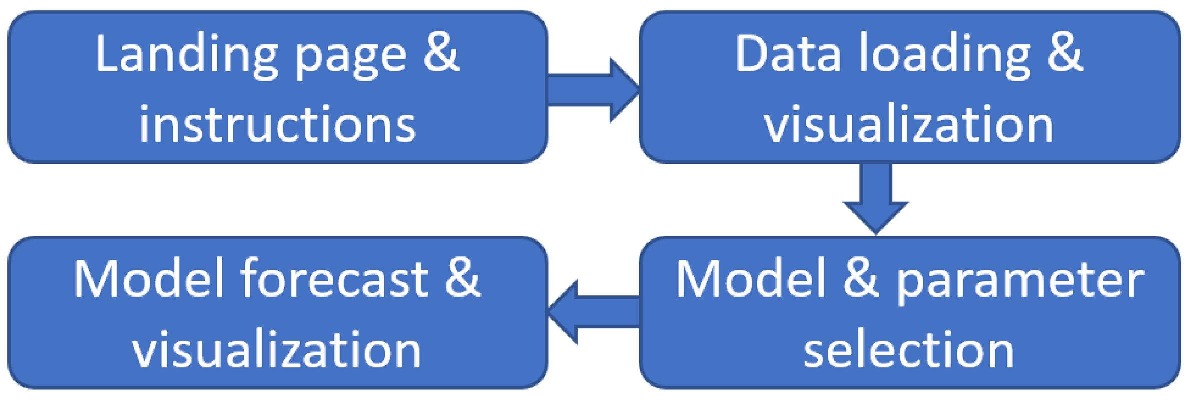
- Use part of the data for **training** model and estimating parameters, and rest for **testing** the model.
- **Model parameter estimate** by solving differential equations & model fitting
- **Parameter uncertainty estimation** using an error structure e.g. Poisson

DATA

- Current COVID-19 data for countries, downloaded from Our World in Data **via Johns Hopkins University (CSSE)**. Updated daily - 196 countries with over 126,000 rows (one row per location and date)
- Alternatively, **users uploads their own data via a csv, tsv, or txt file**

APPLICATION DESIGN

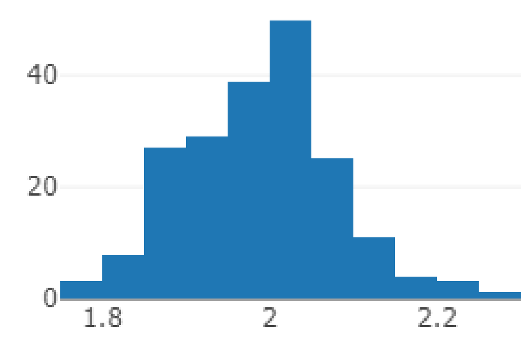
R-Shiny package used for application development. It follows a multi-tab design to guide the users between steps of workflow as shown here:



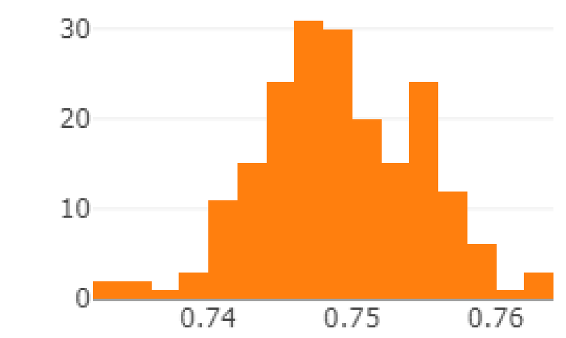
EXPERIMENTS / RESULTS

Model Performance:

After building each model, parameter estimates, and model performance can be evaluated using **MAE** & **RMSE**



r=1.99 (95%CI: 1.84, 2.17)



p=0.749 (95%CI: 0.738, 0.759)

Benchmark testing:

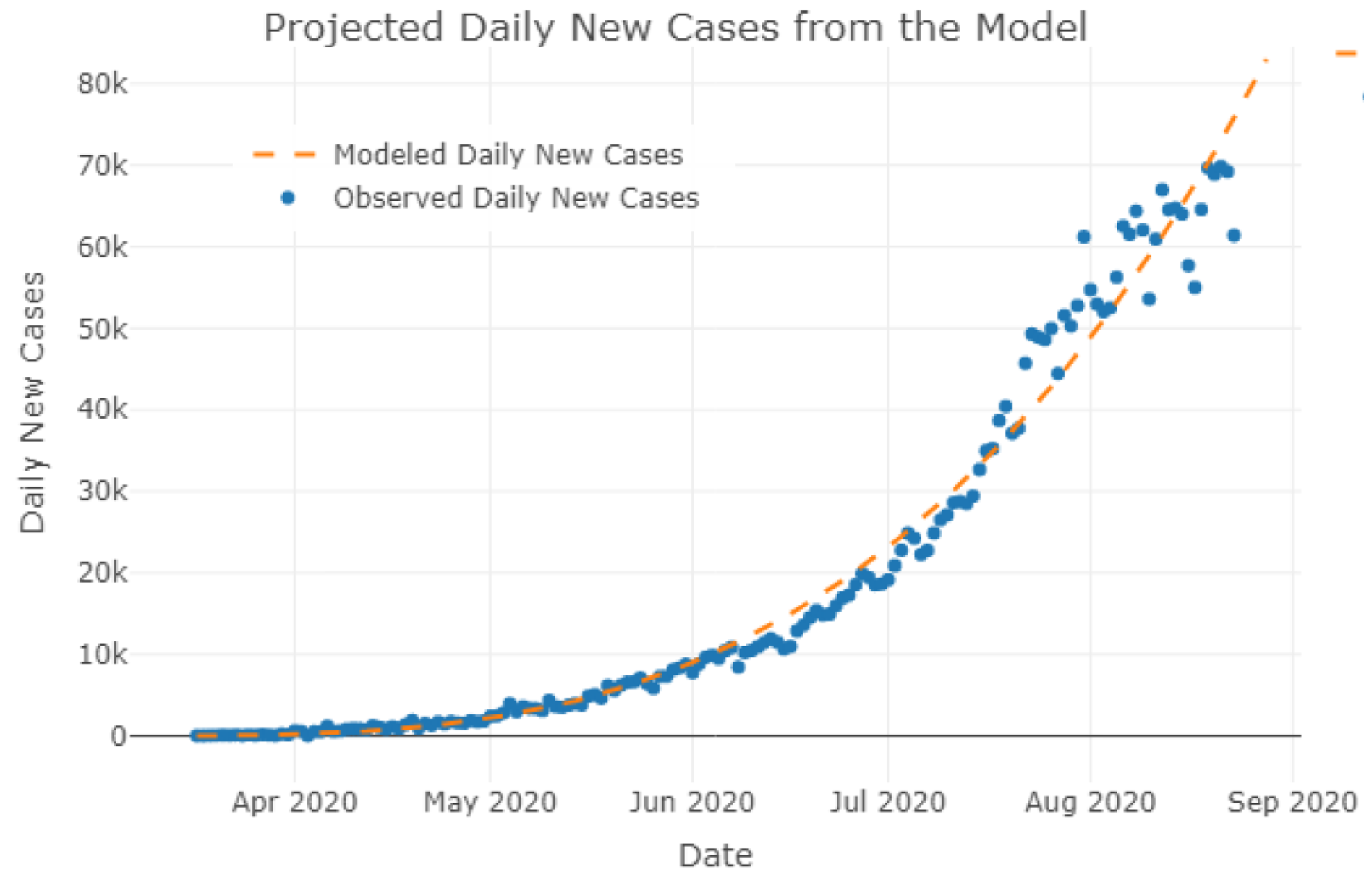
Comparison of the implemented models with published literature has been conducted using select cases as shown in the table

Disease	Location	Time period	Model
COVID-19	India	17 Mar-23 Aug 2020	GGM
COVID-19	Columbia	30 Mar-23 Aug 2020	GGM
COVID-19	Italy	25 Feb-15 Apr 2020	GRM
COVID-19	Italy	25 Feb-15 May 2020	GRM
COVID-19	Italy	25 Feb-31 May 2020	GRM
COVID-19	Italy	25 Feb-10 Jun 2020	GRM

*GGM: Generalized growth model
GRM: Generalized Richards model

Testing Results:

Model performance and comparisons have been satisfactory as **shown here for the cases in India using GGM**. Some observed differences are likely due to input data being not exactly the same.



Model Selection & Usage Guidelines:

- These models have limitations for multi-wave epidemics like COVID-19. **Stage-wise multi-modal approach likely to work better**
- **Generalized & Exponential growth models** for early stage of the wave
- **Logistic growth models** when the outbreak is slowing or passed its peak
- Above mentioned models are special case of Generalized Richards model which allows slower decline rate past the peak

FUTURE ENHANCEMENTS

- **Auto-detection of stages** of outbreaks, making more precise now-casting
- **Better estimation of confidence / prediction intervals** for large outbreaks as in COVID-19
- A **user feedback survey** to improve application useability and features
- **Implement sub-epidemics model** for multi-wave infections