COVID-19: Data-Driven Mean-Field-Type Game Perspective

COVID-19: Data-Driven MFTG Modelling

Hamidou Tembine
Learning & Game Theory Laboratory
Center on Stability, Instability and Turbulence
New York University Abu Dhabi
tembine@nyu.edu

Abstract

In this article, a class of mean-field-type games with discrete-continuous state spaces is considered. We establish Bellman systems which provide sufficiency conditions for mean-field-type equilibria in state-and-mean-field-type feedback form. We then derive unnormalized master adjoint systems (MASS). The methodology is shown to be flexible enough to capture multi-class interaction in epidemic propagation in which multiple authorities are risk-aware atomic decision-makers and individuals are risk-aware non-atomic decision-makers. Based on MASS, we present a data-driven modelling and analytics for mitigating Coronavirus Disease 2019 (COVID-19). The model integrates untested cases, age-structure, decision-making, gender, pre-existing health conditions, location, testing capacity, hospital capacity, mobility map on local areas, in-city, inter-cities, and international. It shown that the data-driven model can capture most of the reported data on COVID-19 on confirmed cases, deaths, recovered, number of testing and number of active cases in 66+ countries. The model also reports non-Gaussianity and non-exponential properties in 15+ countries.

Contents

1 Introduction **3**

2 Basic models **10**
2.1 Data-Driven SIR 12
2.2 Basic SIR with location, age and economic wealth 14
2.3 Control of SIRD with spatial and age structure 15
2.4 A basic SIRD using Mean-Field Game 16
2.5 A basic SIRD using Mean-Field-Type Game 17

3 Data-Driven MFTG **17**
3.1 Authorities, Firms and People 17
3.1.1 Individual state 20
3.1.2 Global state 21
3.1.3 Individual state transitions 22
3.1.4 Mobility per area/city/country 27
3.1.5 Data-driven mean-field trajectories 29
3.1.6 Beyond pairwise interaction 30
3.2 Risk-aware MFTG 31
3.2.1 Mean-field-type equilibrium 32

4 Model outputs and Data visualization **35**
4.1 Data used and constraints 35
4.2 Visualization 37

5 Conclusion **38**

1 Introduction

COronaVIrus Disease 2019 (COVID-19) is one of the severe acute respiratory syndrome-related coronavirus. COVID-19 is a disease caused by the virus SARS-CoV-2. SARS-CoV-2 is infecting people across several countries within few months and causing deaths, economic losses and food insecurity. As COVID-19 spreads around the globe, authorities are making decisions about local migration, confinement, isolation, quarantine and heath care system capabilities. Individuals, depending on the country, culture, social-economic context, age and season, are also making decisions concerning COVID-19. In this paper we propose an integrated mathematical model and analysis strategies for combatting the propagation of COVID-19 and for examining the economic consequences. For each infected person, one often defines the number of secondary infections that are generated by that patient. The probability distribution of that latter random variables is referred to as offspring distribution. The offspring distribution depends on its infectiousness which can be biological, environmental, behavioral or combination of these. The offspring distribution has been estimated from empirical data for outbreaks of several infectious diseases such as tuberculosis, cholera, plaque, smallpox, severe acute respiratory syndromes (SARS), MERS and Ebola. There are two important sectors that authorities are considering to mitigate COVID-19. The public health sector which tries to understand the dynamics of COVID-19 (via epidemiological models and efforts) and take measure to flatten the curve of active COVID-19 patients (mitigation and suppression). The economic policy sector has been used to alleviate negative consequences implied by COVID-19 and by the restrictive policy measures. The economic policies are also being used to prevent the amplification and the multiplication of the crisis. Unfortunately, these two sectors and approaches have not been coordinated enough. The effectiveness of the public health measures may depend on the perception of people in practice and their expectations because each individual also make its own decisions. Therefore both decisions, from economic policies by authorities and from country-specific individual preferences, should be integrated all together.

Literature review on SARS-CoV-2

There is an important ongoing research on COVID-19. Below we review some closely related works. For SARS-CoV-2 (and hence COVID-19) the offspring

distribution is not known and the dynamic COVID-19 data set which is an information that is updated on a daily basis or so, meaning it changes over time as new information becomes available.

- Effect of mobility on the propagation COVID-19: The movement and migration of people between locations add an additional difficulty in estimating the infectivity as each infected person may be moved, flown to another city, state, country or continent. Some estimations have been proposed by using contact-tracing or mobile tracking of the path of the each infected person and all its possible contacts using mobile networks and recognition. However, these estimations on COVID-19 remain limited at a global level [1–4]. The work [5] focuses on a Global Epidemic and Mobility model that integrates sociodemographic and population mobility data set in a spatially structured stochastic disease approach to simulate the spread of epidemics at the worldwide scale. In contrast to [5], we integrate a data-driven optimization setup for authorities and for individuals as well.

- Age and gender structure on COVID-19 death rate: Because the COVID-19 reported deaths depend on ages and gender, the role of demography on COVID-19 spread and fatality is examined in [6]. The work in [7] explains the use of SARS-CoV-2 RT-PCR testing outcomes in Italy. The work in [8] provides some early estimates of clinical severity of COVID-19 by age.

- Epidemiological data set is needed during epidemics to best monitor and anticipate spread of infection. The work in [9] provides geo-positioned records of case of of COVID-19 outbreak. COVID-19-related deaths are not clearly defined in the international reports available so far, and differences in definitions of what is or is not a COVID-19-related death might explain some of the variations in the reported data.

Several basic compartmental models have been suggested in the literature to simplify the complexity of mathematical modelling of infectious diseases. These include

- SIR (susceptible-infected-removed),

- SEIR, that is extended SIR with exposed (E) agents

Feature	SIR	SEIR	SEIRD	SEIR-spatial	SEIR -age structured	MFTG
locality	no	no	no	no	no	yes
city	no	no	no	no	no	yes
sector per city	no	no	no	no	no	yes
position	no	no	no	yes	no	yes
infection status	3 states	4	5	4	4	17
age	no	no	no	no	yes	yes
gender	no	no	no	no	no	yes
height	no	no	no	no	no	yes
health	no	no	no	no	no	yes
family size	no	no	no	no	no	yes
economic status	no	no	no	no	no	yes
testing strategy	no	no	no	no	no	yes
decision-making	no	no	no	no	no	yes
untested positive	no	no	no	no	no	yes
data-driven	no	no	no	no	no	yes

Table 1: Key features of the data-driven MFTG model.

- SEIR with quarantine (Q)
- SEIR with age-structure
- SEIR with spatial location
- SEIRD that is R is divided into recovered (R) or deceased (D)
- SEIRD with data fitting via parameters.

These models need to be adapted to include important and key features of COVID-19 such as context-awareness, decision-making, co-morbidities per area, mobility, age, gender, healthcare status, testing capabilities to capture the evolving COVID-19 data set [10–14].

Table 1 displays some of the key differences and key features of the data-driven MFTG model.

Data-driven mean-field-type game theory

The term 'mean-field' has been referred to a physics concept that attempts to describe the effect of an infinite number of particles on the motion of a

single particle. Researchers began to apply the concept to social sciences in the early 1960s to study how an infinite number of factors affect individual decisions. However, the key ingredient in a game-theoretic context is the influence of the distribution of states and or control actions into the payoffs of the decision-makers. There is no need to have a large population of decision-makers. A mean-field-type game (MFTG) is a game in which the (instantaneous) payoffs and/or the state dynamics coefficient functions involve not only the state and actions profiles but also the distributions of state-action process (or its marginal distributions).

In a mean-field-type game, the quantities-of-interest (such as instantaneous payoff, cost, performance functionals, state kernel/coefficients) depend not only on the type-state-actions of the decision-makers but also on the distribution of them. This allows us to consider higher order (non-linear) dependence in the distribution of state and or actions on the expected utility/cost functions. It also allows us to capture risk-awareness in the performance metric whenever one or more decision-makers are facing uncertainties. We refer the reader to [15–25] for more details on MFTG.

We build upon the MFTG model a data-driven methodology. The data-driven mean-field-type game theory (MFTG) model of the pandemic COVID-19 presented here is built upon Markov decision process model of epidemics. Despite the simplicity of this model, important features are extracted that can assist potential decisions on the strategy to combat the outbreak, essentially configuring a scenario ranging from short-term suppression to long-term mitigation depending on the achieved reduction in the contact number, exposed number, testing number, pre-existing health conditions, age, gender, location, mobility pattern and spatial contact rate. A coopetition and dedicated effort from individuals, governments, decision-makers and stakeholders will help in reducing the propagation but with a global high cost despite the increased rate of recovered people. Basically, data-driven MFTG is given as follows.

- Time is continuous
- Data is reported at discrete-time instants.
- Decision-makers: individuals, firms and authorities
- Control Actions: Individuals' actions include contact/meeting rates and a movement path to go outside for food, pharmacies or other ba-

sic needs. Firms' action include the production essential, moderate-essential and less-essential goods and total working hours. An authority at specific location controls the testing actions and migration rates within that area, between different areas of a city or between cities or the country level.

- Objective: an individual aims to reduce its risk of being infected and to limit its economic consequences. Each firm aims to maximize its revenue while keeping its employees from infections. Each authority aims to reduce the number of deaths under cost-benefit considerations (see Figure 1).

Contributions

The proposed data-driven MFTG model has the following key features

- context-awareness: captures some areas/city/country-specific features
- integrated decision-making and risk-aware cost-benefit analysis for authorities and individuals.
- the mobility map/pattern per areas/cities/countries is endogenous to the model
- the country-specific age-structure is endogenous to the model
- the pre-existing health conditions per city/country is endogenous to the model
- an estimation of the probability of being exposed (coughing, sneezing, surface contact)
- the hospital capacity is endogenous to the model. It is an aggregated data from number of hospital beds, intensive care facilities, ventilators, nurses, doctors, point-of-care etc.
- the testing capacity per country is endogenous to the model (the number of testing per week, number of testing facilities per areas/city/country)
- the infection status of the untested individuals is endogenous to the model

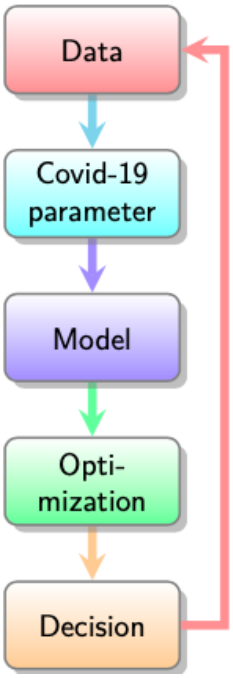

Figure 1: Data-Driven MFTG Model: Data, Covid-19 parameter, Model, Optimization, Decision and back to the data measurement. The verification-validation procedure used in this article.

The goal of the article is to examine the data-driven MFTG modelling aspect of COVID-19. From the mean-field-type game theory perspective, one examines carefully each individual process at the microscopic level. The individual characteristics are age, gender, location, and infection status. As each infected individual may affect its evolving neighbors and path, its create an evolving mean-field term at the local, regional, national and international level. Thus, a macroscopic description of the proportion of infected per location will emerge as a global mean-field term. In addition, risk-awareness at the individual level as well as the global level is proposed. As several coun-

tries may not have the same season, culture or socio-economic context, we introduce temperature field, seasonal variation and local specificities across the countries and couple them with the decisions on the duration of the soft/medium/hard confinement. The age-structured dynamics is introduced to capture the dependence on age and on the prior medical conditions of the COVID-19 patient. For the pandemic COVID-19, many local hospitals cannot sustain a massive influx of infection and this will create a congestion and saturation. Saturation and the local capacities of the healthcare system are introduced into the model and de-congestion and curve flattening strategies are analyzed. We also examine the impact of interventions, control and risk-aware mitigation policies from the authorities. Both early control, late control and no control scenarios are examined from a network perspective using a risk-aware mean-field-type game theory. At the individual level we formulate a risk-aware decision-making problem that the probability of being infected and its possible consequences. We build upon prior optimization-based multi-agent models developed in [26–30]. The number of deaths in the infected population is estimated in two ways. The number of deaths at hospitals and the number of deaths who are not tested. While there is a reported data set for the number of deaths at hospitals that directly and indirectly related to COVID-19. Currently, there is no clear data set for the number of untested deaths. We use a model to estimate it.

The COVID-19 pandemic has already an enormous impact on economic, finance, agriculture, energy, and geopolitics. The decisions made by the authorities such as confinement, lockdown, isolation, curfew measures, may have severe side effects and consequences on other sectors such as employment, supply chain, agriculture and food security. Therefore a cost-benefit analytics need to be conducted for each policy per area/city/country.

The risk-aware mean-field-type game techniques presented here can be used to better understand the propagation of virus across geographical areas. As an integrated model, the impact of control strategies and switching between different locations in the propagation of the virus include

- international connectivity via flight, boat or ground transportation,
- inter-cities connectivity via flights, trains, buses, cars, motorcycles,
- within city connectivity with pedestrian mobility to pharmacies, local markets, open street air markets, supermarkets, workplaces, offices,

and isolation/confinement strategies and testing strategy based on the spreading characteristics. For finite networks, the mean-field term is stochastic and characterizes the evolution of the total number of infected/recovered/tested/untested agents per location. By introducing control strategies at both atomic and non-atomic levels, the number of deaths can be reduced. However, the control is costly and need to be carefully planned (Figure 11). Table 2 summarizes the heterogenous data sources.

Data	Sources
International connectivity per country	Flightradar24 [31]
Age structure per country	[32]
COVID-19 data set	[10–13, 33]
Local mobility	Google Community Mobility Reports [34]
Country-specific hospital capacity data	World Bank Group [35]
Geographical map	OpenStreet map
Economic policy	IMF

Table 2: Heterogenous data sources.

Structure

The rest of the article is organized as follows. In Section 2 we present some basic models. The data-driven MFTG setup is presented in Section 3. Section 4 presents the COVID-19 data set used. Section 5 concludes the article.

Notation

Table 3 summarizes the notations used the proposed model.

2 Basic models

SIR We start with a basic system of differential equations to estimate the propagation in an homogenous population with size $n = n(t_0) >> 1$ at time t_0 and three compartments (Susceptible "S", Infectious "I", Removed "R") There are three different variables (S, I and R) representing the status of each

Symbols	Meaning
l	locality
$l = (l_1, l_2, l_3)$	country, city, area
$T(t, l)$	temperature field
η_1	migration between countries
η_2	migration between cities
η_3	local mobility between areas of a city
z	age variable
$\rho(l, z, t)$	initial age distribution
$\bar{t}_l(t)$	testing capacity per week
δ_l	testing strategy at l
$c = (c_1, c_2, c_3, c_4)$	pre-existing health condition, gender, height, family size, economic wealth
$x = (x_1, x_2, h)$	position of an individual agent
v	velocity of an individual agent
s	individual epidemiological states
n_l	subpopulation size at locality l
u	meeting/activity strategy
m	mean-field
f	Drift
β	interaction kernel
ϵ	physical distancing of 2 m
e	economic wealth

Table 3: Some notations used the proposed model.

compartment. β represents the disease transmission rate and γ describes the "removed" rate.

$$\frac{dS}{dt} = -\lambda \cdot S$$

$$\frac{dI}{dt} = \lambda \cdot S - \gamma \cdot I$$

$$\frac{dR}{dt} = \gamma \cdot I,$$

$$\lambda = \beta \cdot \frac{I}{n}$$

An important parameter, called basic reproduction number, is given by:

$$R_0 = \frac{\beta}{\gamma}$$

2.1 Data-Driven SIR

We would like to have an idea of $I(t_0), \beta, \gamma, R_0$. We aim to estimate them from the dynamic data set that reported (publicly) by authorities.

- Given the data sets
  - infected cases $\hat{I}(t_0),\ \hat{I}(t_0+1),\ \ldots, \hat{I}(t_0+k-1)$,
  - Removed cases $\hat{R}(t_0),\ \hat{R}(t_0+1),\ \ldots, \hat{R}(t_0+k-1)$,

  propose an estimate of $\beta_{data} = \hat{\beta}$ (for β) and $\gamma_{data} = \hat{\gamma}$ (for γ)
- Compute an estimate for R_0

We estimate the rate parameters by solving a dynamic optimization problem.

- Given the data sets, solve the basic optimization problem

$$(\hat{\beta}, \hat{\gamma}, \hat{S}_0, \hat{I}_0)$$

$$\in \arg \min_{(\beta,\gamma,S(t_0),I(t_0))\in Spec}$$

$$\int_{t_0}^{t_0+k} |I(t) - \hat{I}(t)|^2 + |R(t) - \hat{R}(t)|^2 \ dt,$$

- This provides $(\beta_{data,k} = \hat{\beta}, \gamma_{data,k} = \hat{\gamma})$
- $$R_{0,data,k} = \frac{\beta_{data,k}}{\gamma_{data,k}}$$

where *Spec* means the specification parameters and $k \geq 1$.

This leads to a Data-Driven SIR:

$$
\begin{aligned}
&S(t_0) = S_0, I(t_0) = I_0 \neq 0, R(t_0) = 0, \\
&\tfrac{dS}{dt} = -\hat{\beta} \cdot \tfrac{I}{N} \cdot S \\
&\tfrac{dI}{dt} = \hat{\beta} \cdot \tfrac{I}{N} \cdot S - \hat{\gamma} \cdot I \\
&\tfrac{dR}{dt} = \hat{\gamma} \cdot I
\end{aligned}
$$

From the latter system one can obtain the trajectories of (S, I, R) over time. Then, several questions arise

- What if the data set changes regularly?
- Weekly estimates?
- Monthly estimates?
- Long-term estimates?

It is important to notice that for some countries, the data set provided need to be structured per week, two weeks, month time windows or a proper moving average to learn a meaningful trend of the data.

Data-Driven SIR

- Pros
  - we have a data-driven model of SIR
  - it is a simple model
  - it is compartment-based
  - it is based on the data set $\{\hat{I}, \hat{R}\}$ during the time window $[t_0, t_0 + k - 1]$
  - it is deterministic
- Contras:
  - N is fixed and global
  - not localized
  - it is homogenous
  - Does the shape of I, R capture the observed data $\hat{I}, \hat{R}$?

Next, we relax the restrictive assumptions made in the previous model and include context-awareness into the data-driven model.

2.2 Basic SIR with location, age and economic wealth

We introduce the SIRD (susceptible-infected-recovered-deceased) dynamics with spatial location x over a map $\mathcal{D}_l$ at locality l, socio-economic status e and age-structure z.

$$\begin{array}{ll} \text{Susceptible:} & S(x,z,e,t) \\ \text{Infected:} & I(x,z,e,t) \\ \text{Recovered:} & R(x,z,e,t), \\ \text{Deceased:} & D(x,z,e,t) \end{array}$$

$$\begin{array}{l} \text{Domain: } (x,z,e,t) \in \mathbb{R}^3 \times \mathbb{R} \times \mathbb{R} \times (t_0,t_1), \\ S_t(x,z,e,t) = \\ -uS(x,z,e,t) \int \beta((x,z,e),(x',z',e'))I(dx',dz',de',t), \\ I_t(x,z,e,t) = \\ uS(x,z,e,t) \int \beta((x,z,e),(x',z',e'))I(dx',dz',de',t) \\ -\gamma I(x,z,e,t) - \delta(x,z,e)I(x,z,e,t), \\ R_t(x,z,e,t) = \gamma I(x,z,e,t), \\ D_t(x,z,e,t) = \delta(x,z,e)I(x,z,e,t), \\ \text{ Initial condition} \\ S(dx,dz,de,t_0) = S_0(dx,dz,de), \\ I(dx,dz,de,t_0) = I_0(dx,dz,de), \\ R(dx,dz,de,t_0) = 0 = D(dx,dz,de,t_0), \\ x \in \mathcal{D}_l \end{array} \tag{1}$$

where u is the meeting rate, γ is the recovery rate, δ is the death rate. These are nonnegative real-valued.

From (1) we observe that the function

$$t \mapsto \frac{d}{dt} \int (S+I+R+D)(dx,dz,de,t),$$

is identically zero. We set $\int (S+I+R+D)(dx,dz,de,t_0) = n_0$, at time $t = t_0$. The condition for I_t to be negative yields $uS(\int \beta I) < (\gamma+\delta)I$. This means that basic reproduction number R_0 depends not only on $\frac{u}{\gamma+\delta}$ but also the kernel β which captures physical distancing and economic wealth interaction in the society. We choose $I(x,z,e,t_0)$ to be higher for very high e to capture the initial infection from (international) travelers and people who are in touch with international businesses. Then the contact rate kernel β is inversely proportional with e. The spread is transferred to areas with

low income, the dynamics exhibits a higher infection of COVID-19 in some places with low income per family size. However the spread transfer from higher income population to a population with lower income may be slower in some countries due less interaction β with infectious people.

2.3 Control of SIRD with spatial and age structure

We introduce the control of the SIRD (susceptible-infected-recovered-deceased) dynamics with spatial location x, socio-economic status e and age-structure z. Let $t_0 < t_1$ and $(x, z, e, t) \in \mathbb{R}^3 \times \mathbb{R} \times \mathbb{R} \times (t_0, t_1)$.

$$
\begin{aligned}
&S_t + S_z - \epsilon(s,z,e)\Delta_x S = -\tilde{d}(s,z,e)S - \mathrm{div}_x(vS)\\
&-uS(x,z,e,t)\int \beta((x,z,e),(x',z',e'))I(dx',dz',t),\\
&I_t + I_z - \epsilon(i,z,e)\Delta_x I = -\tilde{d}(i,z,e)I - \mathrm{div}_x(vI)\\
&+uS(x,z,e,t)\int \beta((x,z,e),(x',z',e'))I(dx',dz',de',t)\\
&-\delta(x,z,e)I(x,z,e,t) - \gamma I(x,z,e,t),\\
&R_t + R_z - \epsilon(r,z,e)\Delta_x R\\
&= -\tilde{d}(r,z,e)R + \gamma I(x,z,e,t) - \mathrm{div}_x(vR),\\
&D_t = \delta(x,z,e)I(x,z,e,t),\\
&\\
&\text{Initial condition}\\
&S(dx,dz,de,t_0) = S_0(dx,dz,de),\\
&I(dx,dz,de,t_0) = I_0(dx,dz,de),\\
&R(dx,dz,de,t_0) = 0 = D(dx,dz,de,t_0),\\
&\\
&\text{New borns (age 0):}\\
&I(x,0,e,t) = R(x,0,e,t) = 0,\\
&S(x,0,e,t) =\\
&\int \tilde{b}(x,z,e,s,t)S(x,dz,e,t)\\
&+\int \tilde{b}(x,z,e,r,t)R(x,dz,e,t)\\
&+\int \tilde{b}(x,z,e,i,t)I(x,dz,e,t),
\end{aligned}
$$

where the number of infectious people inside the ball $\mathcal{B}_\epsilon(x)$, within two meters radius ϵ is given by

$$n_{I(t)}(\mathcal{B}_\epsilon(x)) = \int_{x' \in \mathcal{B}_\epsilon(x)} \int_{(z',e')} I(dx', dz', de', t),$$

u is a control functional, $\epsilon(.)$ is a diffusion coefficient, Δ_x denotes the Laplacian operator, div_x denotes the divergence operator, $\beta((x,z),(x',z'))$ is the

interaction kernel by age and spatial position of the agents, $\tilde{b}$ is the fertility rate of the population of Susceptible (S), Infected (I), recovered (R), $\tilde{d}(s,z,e)$ is the death rate by age and infection status, v is the velocity (three-dimensional), and e is the economic wealth.

Then the optimal control of SIRD becomes a Mean-Field-Type Control Problem, and is given by

$$\begin{aligned}
&\sup_{(u,v)} \int S(dx,dz,de,t_1) - D(dx,dz,de,t_1) \\
&- \int_{t_0}^{t_1} dt \int h_0(n_I(\mathcal{B}_\epsilon(x)))\|v\|_2^2(x,z,e,s,t)S(dx,dz,t) \\
&- \int_{t_0}^{t_1} dt \int \|v\|_2^2(x,z,e,s,t)S(dx,dz,de,t) \\
&- \int_{t_0}^{t_1} dt \int \|v\|^2(x,z,e,r,t)R(dx,dz,de,t) \\
&- \int_{t_0}^{t_1} dt \int \|v\|_2^2(x,z,e,i,t)I(dx,dz,de,t) \\
&+ \int_{t_0}^{t_1} dt \int [\log(u) - u^2](x,z,e,s,t)S(dx,dz,de,t) \\
&+ \int_{t_0}^{t_1} dt \int [\log(u) - u^2](x,z,e,r,t)R(dx,dz,de,t), \\
&\text{subject to} \\
&\mu = (S,I,R,D) = (S^{u,v,\mu_0}, I^{u,v,\mu_0}, R^{u,v,\mu_0}, D^{u,v,\mu_0}), \\
&\mu(.,t_0) = \mu_0,
\end{aligned} \tag{2}$$

2.4 A basic SIRD using Mean-Field Game

We formulate a Mean-Field Game Problem as follows

$$\begin{aligned}
&\sup_{(u,v)} \int S(dx,dz,de,t_1) - D(dx,dz,de,t_1) \\
&- \int_{t_0}^{t_1} dt \int h_0(n_I^*(\mathcal{B}_\epsilon(x)))\|v\|^2(x,z,e,s,t)S(dx,dz,de,t) \\
&- \int_{t_0}^{t_1} dt \int [\|v\|^2(x,z,e,s,t)S(dx,dz,de,t) \\
&- \int_{t_0}^{t_1} dt \int \|v\|^2(x,z,e,r,t)R(dx,dz,de,t) \\
&- \int_{t_0}^{t_1} dt \int \|v\|^2(x,z,e,i,t)I(dx,dz,de,t) \\
&+ \int_{t_0}^{t_1} dt \int [\log(u) - u^2](x,z,e,s,t)S(dx,dz,de,t) \\
&+ \int_{t_0}^{t_1} dt \int [\log(u) - u^2](x,z,e,r,t)R(dx,dz,de,t), \\
&\text{subject to} \\
&\mu = (S,I,R,D) = (S^{u,v,\mu_0}, I^{u,v,\mu_0}, R^{u,v,\mu_0}, D^{u,v,\mu_0}), \\
&\mu(.,t_0) = \mu_0, \\
&m = (S^*, I^*, R^*, D^*) \\
&= (S^{u^*,v^*,m_0}, I^{u^*,v^*,m_0}, R^{u^*,v^*,m_0}, D^{u^*,v^*,m_0}), \\
&n_{I^*(t)}(\mathcal{B}_\epsilon(x)) = \int_{x' \in \mathcal{B}_\epsilon(x)} \int_{(z',e')} I^*(dx',dz',de',t)
\end{aligned} \tag{3}$$

Note that Problem (3) is linear w.t.r the measure μ (but non-linear in

m).

2.5 A basic SIRD using Mean-Field-Type Game

We now introduce a basic Mean-Field-Type Game (MFTG) as follows

$$\begin{array}{l}
\sup_{(u,v)} \int S(dx,dz,de,t_1) - D(dx,dz,de,t_1) \\
- \int_{t_0}^{t_1} [\epsilon_1 \mathrm{var}_\mu(v) + \epsilon_2 \mathrm{var}_\mu(u)] dt \\
- \int_{t_0}^{t_1} dt \int h_0(n_I^*(\mathcal{B}_\epsilon(x))) \|v\|^2(x,z,e,s,t) S(dx,dz,de,t) \\
- \int_{t_0}^{t_1} dt \int \|v\|^2(x,z,e,s,t) S(dx,dz,de,t) \\
- \int_{t_0}^{t_1} dt \int \|v\|^2(x,z,r,t) R(dx,dz,de,t) \\
- \int_{t_0}^{t_1} dt \int \|v\|^2(x,z,e,i,t) I(dx,dz,de,t) \\
+ \int_{t_0}^{t_1} dt \int [\log(u) - u^2](x,z,e,s,t) S(dx,dz,de,t) \\
+ \int_{t_0}^{t_1} dt \int [\log(u) - u^2](x,z,e,r,t) R(dx,dz,de,t), \\
\text{subject to} \\
\mu = (S,I,R,D) = (S^{u,v,\mu_0}, I^{u,v,\mu_0}, R^{u,v,\mu_0}, D^{u,v,\mu_0}), \\
\mu(.,t_0) = \mu_0, \\
m = (S^*, I^*, R^*, D^*) = \\
(S^{u^*,v^*,m_0}, I^{u^*,v^*,m_0}, R^{u^*,v^*,m_0}, D^{u^*,v^*,m_0}), \\
n_{I^*(t)}(\mathcal{B}_\epsilon(x)) = \int_{x' \in \mathcal{B}_\epsilon(x)} \int_{(z',e')} I^*(dx',dz',de',t)
\end{array} \tag{4}$$

Note that Problem (4) is nonlinear w.t.r the measure μ (and non-linear in m).

3 Data-Driven MFTG

There are $K \geq 1$ countries distributed over a three-dimensional geographical area. A projected map is given in Figure 2. The set of countries is represented by a set of locations $\{1, \ldots, K\}$ with geographical coordinates (latitude, longitude) on the map.

3.1 Authorities, Firms and People

- Authorities: In each country/location, there is an authority who can decide on the testing policy (when, how many), migration, confinement policies of the country and international connections via flights, boats and ground transportations. The authority in l, has a testing capacity,

and the migration rate $\eta_{ll'}, l' \in \mathcal{N}_l$, where $\mathcal{N}_l$ is the set of connecting areas/cities/countries from l. The intra-city mobility in city l_3 is represented by the mobility rate η_3, the inter-city mobility from/to cities l_2 is represented by the mobility rate η_2, and the international connectivity from/to country l_1 is represented by the mobility rate η_1.

The country specific seasons and temperature field $T(t, l)$ are given. The temperature field will be taken into consideration when making a decision on the duration of a lockdown. In each country in $\{1, \ldots, K\}$ there is a subpopulation of people, called agents. The distribution of age of all agents in country l is given by $\rho(l, dz, t)$ where the (chronological) age z ranges from 0 to 124 years. The function ρ is fitted from real data set making it heterogenous from one country to another. The age distribution data per country is available in [32]. An illustrative location-specific age distribution (unnormalized) function is displayed in Figure 4. Figure 3 illustrates a sample (unnormalized) death rate per age-structure. A spectrum of gender and height are also given.

The authorities play the role of atomic decision-makers as their decisions can influence significantly the internal and external migration from that country, and the subpopulation at that location. The authority of country l aims to reduce infections, deaths and economic losses. The authorities strategic objectives are the following:

- reduce the number of deaths
- Interrupt agent-to-agent transmission including reducing secondary infections among close contacts and preventing further spread at a larger-scale;
- Identify, isolate and care for patients early (testing-tracking-tracing);
- Realizing a testing strategy at a given time window. The testing capacity $\bar{t}_l(t)$ can be improved by the authority in l over time t.
- Minimize social and economic impact through multisectoral partnerships between countries, geographical locations and airports hubs.

- Firms: Production goods are classified as essential, moderate-essential and less-essential. Each firm decides its total working hours and aims to reduce the number of infected employees.

- Agents per country: Each agent has an age z and a position located in one of the countries. An agent of any age can be infected by a virus (SARS-CoV-2 for the Coronavirus disease 2019). An agent with located in area l_3 of city l_2 of province state/country l_1, age z,pre-existing health condition (c), gender(g), height (h), family size (fs), average revenue per family. We rename (c, g, h, fs) of the agent to be $c = (c_1, c_2, c_3, c_4)$. and has an infection status described as follows.

  - For untested individuals, there are six states. Non-infected agents are susceptible (lac-S). Non-infected agents including recovered agents can be re-infected by another mutated version (if any) of the virus. An agent can be exposed (lac-E), for example, by an infected agent who is coughing, sneezing, from a surface contact or from a directional air conditioning (if any). Some of the exposed agents become passive (lac-P) which means dormant (asymptomatic infectious or mild symptom) and some others become active (lac-A). One absorbing state: deceased (lac-D). Some of the passive or active agents become recovered (lac-R).
  - Some of exposed people will be tested. There are ten additional states plus one specific state actioned by healthcare authorities to identify, isolate the potential exposed people for example by means of phone calls or contact-tracing and tracking. These states are Testing (lac-T), lac-testing negative, lac-testing positive, lac-testing unclassified/unknown, lac-contact tracing of the last two-weeks for each COVID-19 tested positive, lac-isolation at home, at hotel, at a point-of-care, lac-hospitalized (severely active case), lac-recovered, lac-deceased

  The 17 infection status are given by

  - lac-susceptible, lac-exposed
  - lac untested: lac-passive-untested, lac-active-untested, lac-recovered-untested, lac-deceased-untested,
  - lac-testing, lac-testing negative, lac-testing positive, lac-testing unclassified/unknown,
  - lac-contact-tracing-tracking: for each COVID-19 tested positive agent, a contact-tracing-tracking action is taken by the authori-

ties to reach out and identify people who may have been potentially exposed by getting in the close contact of the COVID-19 positive agent in the last two weeks. This is not an individual but a collective action taken by the authority. A person who has been identified from the list of contract-tracing person may be considered as potentially exposed and a testing decision will be made after evaluation.

  - lac-isolation at home, at hotel, at a point-of-care
  - lac-hospitalized (severely active case), lac-recovered, lac-deceased

Figure 6 illustrates the evolution of the 17 states over time.

3.1.1 Individual state

The internal state of individual p is $(l_p(t), x_p(t), s_p(t), z_p(t), c_p(t))$ where

- $l_p(t) = (l_{1,p}(t), l_{2,p}(t), l_{3,p}(t))$ is an index representing latitude-longitude of the location of person p at time t, $l_{1,p}$ for the country, $l_{2,p}$ for the city/province, $l_{3,p}$ for the area of the city. $l_p(t) \in \mathcal{L}$ describes a discrete set of cities map of the countries in $\{1, \ldots, K\}$.

- $x_p(t) = (x_{1,p}(t), x_{2,p}(t), h_p(t))$ represents the position of agent p at time t

- $s_p(t) \in \{S, E, P, A, R, D, s_7, \ldots, s_{17}\}$ is the infection status of person p at time t,

- $z_p(t) \in [0, 124y]$ is the age of person p at time t, up to 124 years. The number y is a conversion from one year into a proper time units (miliseconds)

- $c_p(t) = (c_{1,p}(t), c_{2,p}(t), c_{3,p}(t), c_{4,p}(t))$ with $c_{1,p}(t) \in C_1$, which is the set of all subset of $C_1' = \{phc_1, phc_2, phc_3, \ldots\}$ a finite list of pre-existing health conditions (phc) that are considered as relevant by health authorities. Notice that C_1 contains the empty set $\emptyset$. If person p has no prior health issue then $c_{1,p}(t)$ is set the empty set $c_{1,p}(t) = \emptyset$. The set C_1 includes also some combinations of diabetes (type 2), severe asthma, pulmonary hypertension, cardiovascular disease, high blood pressure, cerebrovascular disease, chronic obstructive pulmonary disease, chronic

kidney disease, among the others. Older people and people with pre-existing medical conditions are assumed to be more vulnerable to becoming severely ill with the virus.

$c_{2,p} \in C_2$ represents the gender of person p. C_2 is a finite set representing the spectrum of gender, and $c_{3,p}(t) = h_p(t) \in C_3$ is the height of person p at time t. C_3 represents the spectrum of height and C_4 represents the set of family size or number of roommates at residential areas.

The set of states of an agent in location $l \in \mathcal{L}$ is $\mathcal{X} \times C \times \text{Status}$ where the infection status of the agent can be $\{S, E, P, A, R, D, s_7, \ldots, s_{17}\}$ and $C = C_1 \times C_2 \times C_3 \times C_4$.

3.1.2 Global state

The population size is a large but finite number n. The global state of the system is the measure $(M(l, dx, s, dz, c, t),\ l \in \mathcal{L}, s \in \{s_1, \ldots, s_{17}\}, z \in [0, 124y], c \in \mathcal{C}$ where

$$\begin{aligned} &M(l, \mathcal{V}_x, s, \mathcal{V}_z, c, t) = \\ &\textstyle\sum_p \mathbb{I}_{\{(l_p(t),,s_p(t),c_p(t))=(l,s,c)\}} \mathbb{I}_{x_p(t)\in\mathcal{V}_x, z_p(t)\in\mathcal{V}_z}, \end{aligned} \tag{5}$$

$\mathcal{V}_x$ is a Lebesgue measurable set of positions at l and $\mathcal{V}_z$ is a measurable set of ages at l. The intra-state status of the system in location l, age z and characteristics (c_1, c_2, c_3, c_4) at time t is given by

$$\begin{aligned} &M(l, dx, dz, c, t) \\ &= (S, E, P, A, R, D, s_7, \ldots, s_{17})(l, dx, dz, c, t), \end{aligned} \tag{6}$$

representing a measure of the total number of people located at l for each of the 17 status. Let $n(l, z, c, t)$ be the population size at location l with age z and characteristics (c_1, c_2, c_3, c_4) at time t. Let $n_l(t) = \sum_c \int n(l, dz, c, t)$ be the population size at location l at time t. The occupancy measure in area l_3 of city l_2 of country l_1 is

$$M^n(l, \mathcal{V}(x), z, c, t) = \frac{1}{n_l V} \int_{x\in\mathcal{V}} M(l, x, z, c, t),$$

where V of the volume of the neighborhood $\mathcal{V}$.

3.1.3 Individual state transitions

We describe below the transitions between infection status.

- Susceptible agents can be exposed to the virus in different ways. (i) become exposed from a surface contact that has been infected recently. (ii) exposed from a coughing or sneezing of an infectious person (with or without symptoms) with rate λ_{se}

- Exposed agents: Recent observations suggest that for the COVID-19, the risk of exposure can be affected by surface contact, coughing and sneezing (with a face mask vs without a face mask). We propose to examine the mathematics of coughing and sneezing in order to estimate the range and the risk of exposure of a person. A simple dynamics of coughing and sneezing is proposed in [36]. The trajectory of the puff is governed by the evolution of its momentum I and buoyancy B. The buoyancy force is vertical, causing the cloud to follow a curvilinear trajectory denoted by (w, θ, h) or equivalently (x, y, h)

$$\begin{cases} V_t = 3\eta\alpha^3 w^2 w_t - v_p N_t, \\ \rho_{c,t} = \frac{1}{V}(v_p N_t(\rho_c - \rho_p) + 3\eta\alpha^3 w^2 w_t(\rho_a - \rho_c)), \\ B_t = g\rho_a V_t - g(V\rho_c)_t, \\ I_{y,t} = B, \\ w_t = (I_y^2 + \frac{I_0^2 \cos^2\theta_0}{V\rho_c})^{\frac{1}{2}}, \\ N_t = -\frac{3NU_s}{2\alpha w} \end{cases} \tag{7}$$

where V volume of the cloud, $\eta = \alpha = 0.1$ constant, s distance from the source, v_p speed, N is the number of particles within the cloud, with particles concentration $\frac{N(t)}{V(t)}$ at time t, ρ_c (density of the cloud) , ρ_p ρ_a (density of the ambient air), B cloud buoyancy, g is the acceleration due to gravity, $I = (I_x, I_y)$, momentum, I_0 initial momentum, θ_0 initial angle. $U_w = w_t$ is the speed of the suspended particles with a time scale of $\frac{h}{U_w}$

The quantities-of-interest include the fallout time, position, the horizontal range of deposition. These quantities affect the exposure to the range and possibly the transmission of the disease by emitted pathogens. We numerically compute the parameters $\beta, \lambda_{ep,l}$ by looking at the distance over which the cough/sneeze travelled around 0.7 m,

An exposed person

- may become active with a probability $\lambda_{ea}(l)$
- may become passive via two ways. (i) λ_{Sm} is the probability of getting infected by a surface of contact or in close contact with an infected agent. The exposed agent can decide to receive or not an item or a visit to a workplace, grocery store, pharmacy, $\{o, \bar{o}\}$. The contact probability is $u_l(o|e, z, c)$. In practice, the meeting or contact probability are estimated through the path trajectory of exposed people and the use of physical distancing to an exposed person or an exposed surface. (ii) $\lambda_{ep,l} P_l^n(t)$ models the probability of encounter a passive agent where $P_l^n(t) = \int_{x',z'} \sum_c [P^n + t_+^n(l, dx', dz', c, t)]$. An example of interaction between an exposed agent and an active or passive infectious agent is given by

$$\begin{aligned} & m(l, x, s, z, c, t) \int_{x',z'} \sum_{c'} \beta((x', z'), (x, z)) \\ & \sum_{s' \in \{p, a, t_+, \text{tpc}, \text{th}_1, \text{th}_2, \text{thos}\}} m(l', dx', s', dz', c', t)], \end{aligned} \tag{9}$$

where β is the spatial exposure and cross-age kernel between positions x and x' as in (7), and $s \in \{\text{lac-S}, \text{lac-exposed}\}$. For infection at homes and residential areas, we take the average family size per area/city/country. This is made endogenous to the model by setting c_4 to the 'family' size (or number of roommates) of the individual.

- Untested Passive agents: In practice, the confirmed cases (asymptomatic and symptomatic) are done through a two-step process via testing for COVID-19 which can be positive, negative or underdetermined with some probabilities $(\lambda_{po}, \lambda_{ne})$. Since testing is still limited in many countries, we estimate the untested passive agents from a tracking of exposed agents obtained by phone calls, surveys, mobile tracking, contact-tracing, among the others. A large number of the population has not been tested yet. A person in country l is untested with probability $1 - \delta_l$. where δ_l is the testing strategy adopted by the authority in l.
  - may become recovered with probability λ_{pr}.
  - may become susceptible with probability λ_{ps}.
  - may become active with a rate $\lambda_{pa,1}$.
  - may opportunistically encounter with another infectious agent or has comorbidities, and both become active. The rate is $\lambda_{pa,2}(z, c)(P_l^n(t) - \frac{1}{n_l})$. The passive agent can decide to be in physical contact or not with the other infected agent, so there are two possible actions: $\{m, \bar{m}\}$. The meeting probability is $u_l(m|p, z, c)$.
  - may cause death with age.
  - newborns with passive status with age 0. This will be considered as a very small number as in the (publicly) reported data set.
- Untested Active agents:
  - may become susceptible with probability $\lambda_{as}(l, z, c, \rho)$ and depends on age.
  - may become passive with probability $\lambda_{ap}(z, \rho, l)\frac{P_l^n(t)}{h_l + P_l^n(t)}$. This will be considered as a very small number as in the (publicly) reported data set.

- may cause death with age. The death rate $\tilde{d}(z, c, \rho)$, due to the virus, increases with the age z of the agent. The distribution of age in location l solves

$$\rho_t + \tilde{v}\rho_z = -\tilde{d}(z, c, \rho, l)\rho(l, z, t)$$

  where $\tilde{v}$ is the aging speed and $\tilde{d}$ is the death rate. The initial distribution of age in location l is $\rho(l, dz, 0)$. The birth rate is $\rho(l, 0, t) = \int fert(l, z, \rho)\rho(l, dz, t)$ where fert(l, z, ρ) is the fertility rate at age z in country l. Older people and people with pre-existing medical conditions such as severe asthma and pulmonary hypertension, high blood pressure, diabetes, are assumed to be more vulnerable to becoming severely ill with the virus.
- newborns with active status with age 0. The newborn part of the population contributes as $\rho(l, 0, t)$.

- Untested Recovered agents: Depending on the type of virus mutation and immunization of the agent, a recovered agent can be re-infected by a mutated version of the virus and hence,
  - may become susceptible with $\lambda_{rs}(z, c)$ but it is assumed to be small as suggested by the publicly reported data set.
  - may become passive with $\lambda_{rp}(z, c)$ but it is assumed to be small as suggested by the publicly reported data set.

- Untested Death: this is an absorbing state. No transition to any other state once in. The death rate depends not only on the age of the patient but also on the temperature field or season at location l. It also depends the capacity of the Healthcare system at location l. The constraint is $\int \sum_{c'} \text{hospitalized}(l, dx', dz', c', t) \leq \bar{c}_l(t)$, where $\bar{c}_l(t)$ is the capacity of the healthcare system at location l at time t. When the capacity $\bar{c}_l(t)$ is exceeded, the death rate $\tilde{d}$ increases drastically. Authorities anticipate the saturation and find alternatives such as installations of equipments at stadiums and new hospitals. An aggregated hospital capacity data set $\bar{c}_l$ is available from World Health Organization and World Bank Group [35]. Depending on the country, there is also a certain rate λ_{co} of undetected COVID-19 patients (not reported at hospitals or from their homes) that lead to death. Figure 3 illustrates a data-driven death rate by age.

- lac-testing: During a period of asymptomatic infection, the testing can reveal infection depending on the outcome. The testing result can lead to different outcomes and the isolation can be at home, hotel or point-of-care as displayed in Figure 5. Testing occurs in country l with probability δ_l.The result of the testing is often classified as positive, negative or underdetermined (unknown). Quarantine at home, hotel or point-of-care policy is case-dependent. It can be dependent on whether a case is unknown, known positive, known negative, or recovered. Testing therefore makes possible the identification and quarantine of infected individuals and release of non-infected individuals. Each country has its own testing capacity $\bar{t}_l$ and testing rule δ_l (deciding by the authority at l) and the constraint is captured by

$$\sum_{c'} \int_{x',z'} \text{testing}(l, dx', dz', c', t) \leq \bar{t}_l(t).$$

- The testing state moves to one of the three states: lac-testing negative, lac-testing positive, lac-testing unclassified/unknown.

- lac-contact tracing of the last two-weeks for each COVID-19 tested positive agent.

- lac-isolation at home, at hotel, at a point-of-care:

  - may become hospitalized with probabilities

$$\begin{array}{l} \lambda_{\text{th}_1,\text{thos}}(z,c) u_1 \bar{u}_1 (\bar{\text{th}}^n_{1,l}(t) - \frac{1}{n_l}), \\ \lambda_{\text{th}_2,\text{thos}}(z,c) u_2 \bar{u}_2 (\bar{\text{th}}^n_{2,l}(t) - \frac{1}{n_l}), \\ \lambda_{\text{tpc},\text{thos}}(z,c) u_3 \bar{u}_3 (\bar{\text{tpc}}^n_l(t) - \frac{1}{n_l}), \end{array} \tag{10}$$

    respectively.

  - may become recovered. The rates are $\lambda_{\text{th}_1,\text{tr}}(z,c)$, $\lambda_{\text{th}_2,\text{tr}}(z,c)$, $\lambda_{\text{tpc},\text{tr}}(z,c)$,

  - There may be also a transition from home to one of the point-of-care with probability $\lambda_{\text{th}_1,\text{tpc}}(z,c)$.

- lac-hospitalized (severely active case),

  - may become recovered with rate $\lambda_{\text{thos},\text{tr}}(z,c)(\text{thos}^n_l(t) - \frac{1}{n_l})$.

- may become dead with rate $\lambda_{\text{thos,td}}(z,c) + (\text{thos}_l - \bar{c}_l)_+$.

- lac-recovered
  - may become susceptible with rate $\lambda_{\text{tr},s}(z,c)$ but it is assumed to be very small.
  - may become exposed with rate $\lambda_{\text{tr},e}(z,c)$ but it is assumed to be very small.

- lac-deceased: this is an absorbing state. There is no transition to any other state once in.

When $\delta_l = 0$, the model (without testing at locality l) reduces to six basic individual states. Figure 5 represents some of the intra-dynamics for both tested and untested individual at a given location wit pre-existing health condition c, gender, height, family size, age z, and untested options. Connection with SIR, SEIR, SIRD, SEIRD can be found in Figures 7, 8, 9 and 10.

3.1.4 Mobility per area/city/country

- The mobility pattern of the agents through ground transportation within the same street/area of the city, between areas of the cities, between cities via ground and air transportation and international connectivity are given by a mobility map or graph with rate functions $\eta_{ll'}(T, C, c, \lambda, m, \rho, u, v, \delta, n)$ between locations l and l'. We decompose η into three layers: η_1 connecting countries, η_2 for connecting cities, and η_3 for connecting areas of the same city.

The mobility pattern by area per city can be formulated as follows. Let $l = (l_1, l_2, l_3)$ be a location of a specific area l_3 of the city l_2 of region/country l_1 and $x = (x_1, x_2, h) \in \mathbb{R}^3$ a geographical location of an individual agent inside the area l_3. Let $m(l, x, s, z, c, t)$ be a measure of the agents satisfying the Kolmogorov equation (18 in Table **??**). The limiting measure m can be

written as

$$\begin{cases} \text{time-variation} \\ = \text{advection-jump-diffusion} \\ \text{local spread of the disease: coughing and sneezing} \\ \text{local spread of the disease via surface contact} \\ \text{intra-city mobility in } l_3 \\ \text{inter-city mobility from/to } l_2 \\ \text{international connectivity from/to } l_1 \\ \text{population aging and death} \\ \text{number of hospitalized patients vs hospital capacity in } l_3 \\ \text{number of tests vs testing capacity in } l_3 \\ \text{initial population} \\ \text{new borns with age 0,} \\ \text{city architecture constraint} \\ \text{map constraint} \\ \text{Allowed/feasible mobility areas/exits} \end{cases} \tag{11}$$

The evolution of fraction of agents per state is given by

- The number of agents for which there is a transition in one time slot is always less than 2.
- Payoff contribution from switching : $\{0, \frac{1}{n_l}, -\frac{1}{n_l}, \frac{2}{n_l}, -\frac{2}{n_l}\}$.
- The intensity of interaction is in the order of $1/n_l$.
- The drift (the expected change of M^n in one time-step given the current state) of the system is: $\frac{1}{n_l}[D_{\text{intra}} + D_{\text{inter}}]$, with D_{intra} given in (**??**) and

$$D_{inter} := \begin{pmatrix} \sum_{l'} s_{l'}\eta_{l'l} - s_l \sum_{l'} \eta_{ll'} - vs_{z,l} - \tilde{d}s_l \\ \sum_{l'} e_{l'}\eta_{l'l} - e_l \sum_{l'} \eta_{ll'} - ve_{z,l} - \tilde{d}e_l \\ \sum_{l'} p_{l'}\eta_{l'l} - p_l \sum_{l'} \eta_{ll'} - vp_{z,l} - \tilde{d}p_l \\ \sum_{l'} a_{l'}\eta_{l'l} - a_l \sum_{l'} \eta_{ll'} - va_{z,l} - \tilde{d}a_l \\ \sum_{l'} r_{l'}\eta_{l'l} - r_l \sum_{l'} \eta_{ll'} - vr_{z,l} - \tilde{d}r_l \\ 0 \\ \sum_{l'} s_{l'j}\eta_{l'l} - s_{lj} \sum_{l'} \eta_{ll'} - vs_{lj,z} - \tilde{d}s_{lj} \\ 0 \end{pmatrix},$$

where $j \in \{7, \ldots, 16\}$, and m is the actual population mean-field profile as a function of (l, x, s, z, c, t) given in (18) where yq denotes a

quarantined agent with status y. Agents with non-empty pre-existing conditions $c \neq \emptyset$ are handled separately from the above dynamics with a certain quarantine rate at the early stage. Here $\tilde{b}_{00} = 0$ for $z > 0$ and The new borns are counted in the calculations of m from the functional $\tilde{b}$ with age $z = 0$. The transition rate $\tilde{q}(x, s, z, c; (l, x', s', z', c', t, m))$ is obtained from the functions λ following the diagram in (**??**).

Proposition 1. *As $n_l \longrightarrow +\infty$ the occupancy measure $M^n(t)$ converges in probability to the solution of the (controlled) differential equation given by*

$$\dot{m}(l, x, s, z, c, t) = f(l, x, s, z, c, t, \lambda, u, v, \eta, \delta, m),$$

starting from $m(., t_0) = m_0(.)$.

The convergence is in distribution. It is, in general, weaker than the convergence in probability. However, here, the limiting object m is a deterministic (parametrized) object, and hence, the convergence is also in probability. The proof follows from [27, 28] by using a martingale approach with an infinitesimal generator driven by the drift

$$f := \lim_{n \to +\infty} [D_{\text{inter}} + D_{\text{intra}}],$$

where the limit is taken in (**??**).

3.1.5 Data-driven mean-field trajectories

We build upon the works on mean-field-type filters developed in [37–39] to incorporate the economic data and the pandemic data together in the model. We tune the parameters of the model such that the solution of the dynamical system becomes as close as possible to the dynamically changing data set. We rewrite the model with the parameters λ to be tuned from the reported data set using the mean-field dynamics in Proposition 1.

- The geographical coordinates are taken from the world map data set.
- The sequences of reported data set $\hat{y}_l(t_k)$, $k \in \{0, 1, \ldots, \}$, $t_k \leq t$ is given. The reported data set is observed at discrete time instants $t_0 < t_1 < t_2 < \ldots$. t_k captures the time instant at which the k-th data point is officially reported by the corresponding state and healthcare

authorities at location l. The reported data set is a sequence of aggregated data set: $\hat{y}_l(s, t_k)$ where $s \in \{\text{tr,td,testing}, t_+\}$. The conditions of untested agents are not in the reported data set. The reported data set is assumed to be noisy and inaccurate.

- The reported data set is used to construct a filtered trajectory minimizing the global error within $[t_0, t_1]$ by choosing the tensor λ and h_l. Then the resulting dynamics is used to forecast the mean-field within an uncertainty region. The sequence of optimization problems up to time t_k, is given by

$$\inf_{\lambda,h,m_0} \sum_s \int_{t_0}^{t_k} \| \int_x \int_z \sum_c m(l, dx, s, dz, c, t) - \hat{m}_l(s, t)\|_2^2 dt,$$

  where $\hat{m}_l(t) = \hat{m}_l(t_{i-1}) + \frac{t-t_{i-1}}{(t_i-t_{i-1})} \mathbb{I}_{(t_{i-1},t_i)}(t)(\hat{y}_l(t_i) - \hat{m}_l(t_{i-1}))$ for $t \in (t_{i-1}, t_i)$. As a new data point $\hat{m}_l(t_{k+1})$ is observed, we need to re-update the estimation over the integral $[0, t_{k+1}]$ by solving the new minimization problem $\inf_{\lambda,h} \int_0^{t_{k+1}} \|m(t) - \hat{m}(t)\|_2^2 dt$, and so on. The continuation of the mean-field starts from $m(t_k)$ at time t_k.

The above optimization provides $\lambda_{\text{data}}, h_{\text{data}}$ that are specific to COVID-19 country-specific spread. Moreover, λ_{data} and h_{data} are context-aware as they depend on country-specific, age, gender, testing and pre-existing health conditions. The data-driven dynamics yields

$$\dot{m}(l, x, s, z, c, t) = f(l, x, s, z, c, t, \lambda_{\text{data}}, u, v, \eta, \delta, m),$$

with $m(., t_0) = m_0(.)$. Instead of $\hat{m}$ above, the estimation is replaced by the distribution of $m_{|\{\hat{y}(t_k),\ t_k \leq t\}}$ leading to a mean-field-type filtering problem [37–40]. The diagram of Figure 1 represents the main steps of the algorithm of the data-driven MFTG model.

3.1.6 Beyond pairwise interaction

We have detailed above the pairwise interaction case with a kernel β in (9). A 3-wise, 4-wise,... j-wise interaction case, depending on the local markets, supermarkets, pharmacies, family members at home, and the number of in-

fected people at the neighborhood $\mathcal{V}(x)$ can be formulated as

$$\begin{aligned}
&\text{infectious} = \{p, a, t_+, \text{tpc}, \text{th}_1, \text{th}_2, \text{tpc}, \text{thos}\},\\
&m(l, x, s, z, c, t) \Big[\textstyle\int_{x' \in \mathcal{V}(x)} \beta_1(x', x, t) \times\\
&\textstyle\int \sum_{c'} \sum_{s' \in \text{infectious}} m(l', dx', s', dz', c', t)\\
&+ \textstyle\sum_{j \geq 2} \lambda_j(l, x) \int_{(x'_1, \ldots x'_j) \in \mathcal{V}^j(x)}\\
&\beta_j(x'_1 \ldots x'_j, x; \eta) \Big(1 - \textstyle\prod_{\kappa=1}^{j} (1 - m_{++}(l, dx'_\kappa, t))\Big)\Big],\\
&m_{++}(l, x'_\kappa, t) = \textstyle\sum_{c'} \sum_{s \in \text{infectious}} \int_{z'} m(l, x'_\kappa, s, dz', c', t)
\end{aligned}$$

where $\lambda_j(l, x)$ denotes the probability of having j agents around x.

3.2 Risk-aware MFTG

Introduce $\bar{u} = \sum_k u_k \mu_k$ and $\text{var}_\mu(u) = \sum_k (u_k - \bar{u})^2 \mu_k$. The benefit-cost functional is a risk-aware function as it depends explicitly on the variance. Each individual agent chooses u and solves the following best response problem given (m, η)

$$\begin{aligned}
&\sup_{u,v} \mathbb{E}[(s + r)(t_1) +\\
&\textstyle\int_{t_0}^{t_1} ((s + r) - q\text{var}_\mu(u) - \bar{q}\bar{u}^2 - g_0(u, v, \mu, \eta, m))dt]\\
&\text{subject to}\\
&\dot{\mu}(l, x, s, z, c, t) = f(l, x, s, z, c, t, \lambda, u, v, \eta, \delta, m),\\
&\mu(l, dx, s, dz, c, t_0) = m_0(l, dx, s, dz, c),\\
&\dot{m}(l, x, s, z, c, t) = f(l, x, s, z, c, t, \lambda, u^*, v^*, \eta, \delta, m),\\
&m(l, dx, s, dz, c, t_0) = m_0(l, dx, s, dz, c),\\
&\textstyle\int m(l, dx, \text{hospitalized}, dz, t) \leq \bar{c}_l(t),\\
&\textstyle\int m(l, dx, \text{testing}, dz, t) \leq \bar{t}_l(t),\\
&x \in \mathcal{D},\\
&v_{|\partial\mathcal{D}} = 0,\\
&v_{Exits} = k(l, x, m),
\end{aligned} \tag{12}$$

with $\partial\mathcal{D}$ being the border of the map $\mathcal{D}$,

$$\begin{aligned}
&g_0(u, \mu, \eta, \delta, m)\\
&= \|v\|_2^2 [h_1(\textstyle\int_{x' \in \mathcal{V}_x} (t_+ + p)(dx'))\\
&+ h_2(x, \mathbb{I}_{\int_{x' \in \mathcal{V}_x} \int_z \sum_c (t_+ + p)(l, dx', dz, c, t) > 1})]\\
&+ \text{var}(v) - b_1(v),
\end{aligned}$$

and b_1 is the benefit of the agent, h_1, h_2 are positive and non-decreasing functions (in the second variable for h_2.)

The l-th authority chooses a finite dimensional vector η and solves the following best-response problem given m and given the migration restrictions imposed by the other authorities.

$$\begin{array}{l}
\inf_{\eta_l,\delta_l}(d_l + a_l + t_{+,l} + \mathrm{th}_{1,l} + \mathrm{th}_{2,l} + \mathrm{tpc}_l + \mathrm{thos}_l + td_l)(t_1) \\
+ \int_{t_0}^{t_1}((d_l + a_l + \mathrm{thos}_l + \mathrm{td}_l)dt \\
+ \int_{t_0}^{t_1} q_l \mathrm{var}_m(\eta_l) + \bar{q}_l \bar{\eta}_l^2 + g_l(\eta, \delta, m))dt \\
\text{subject to} \\
\dot{m}(l, x, s, z, c, t) = f(l, x, s, z, c, t, \lambda, u, v, \eta, \delta, m), \\
m(l, dx, s, dz, c, t_0) = m_0(l, dx, s, dz, c), \\
\int m(l, dx, \text{hospitalized}, dz, t) \leq \bar{c}_l(t), \\
\int m(l, dx, \text{testing}, dz, t) \leq \bar{t}_l(t), \\
x \in \mathcal{D}_l, \\
v_{|\partial\mathcal{D}} = 0, \\
v_{\text{Exits}} = k(l, x, m),
\end{array} \tag{13}$$

where $\eta_l = (\eta_{l'l})_{l'}$, $\bar{\eta}_l = \sum_{l'} m_{l'}\eta_{l'l}$ and $\mathrm{var}_m(\eta_l) = \sum_{l'}(\eta_{l'l} - \bar{\eta}_l)^2 m_{l'}$ and $g_l(\eta, \delta, m)$ is the economic cost-benefit at l.

3.2.1 Mean-field-type equilibrium

The MFTG problem is then obtained by combining the two interdependent problems (13) and (12). A solution $(u^*, v^*, \eta^*, \delta^*)$ to this MFTG problem is a mean-field-type Nash equilibrium.

Proposition 2. *If there is a solution to the following Hamilton-Jacobi-Bellman system in the space of measure $\mathcal{P}_2(\mathbb{R}^2 \times \mathcal{S} \times [0, 124y] \times C)$ given by*

$$\begin{array}{l}
W: \ [t_0, t_1] \times \mathcal{P}_2(\mathbb{R}^2 \times \mathcal{S} \times [0, 124y] \times C) \\
\times \mathcal{P}_2(\mathbb{R}^2 \times \mathcal{S} \times [0, 124y] \times C) \to \mathbb{R}, \\
V_l : [t_0, t_1] \times \mathcal{P}_2(\mathbb{R}^2 \times \mathcal{S} \times [0, 124y] \times C) \to \mathbb{R},
\end{array} \tag{14}$$

where $\mathcal{P}_2$ is the set of square (in space and age) integrable (positive) measures

and

$$\begin{aligned}
&W_t + H_0 = 0,\\
&W(t_1, \mu, m) = (s + r + tr)(t_1)\\
&V_{l,t} + H_l = 0,\\
&V_l(t_1, m) =\\
&(d_l + a_l + t_{+,l} + th_{1,l} + th_{2,l} + tpc_l + thos_l + td_l)(t_1)\\
&l \in \mathcal{L},
\end{aligned} \tag{15}$$

where the Hamiltonians are given by

$$\begin{aligned}
&H_0 = \sup_{u,v}[b_0 + \langle f, W_\mu\rangle] + \langle f(., u^*, v^*), W_m\rangle,\\
&H_l = \inf_{\eta_l,\delta_l}[b_l + \langle f, V_{l,\mu}\rangle],\\
&b_0 = (s + r) - qvar_\mu(u) - \bar{q}\bar{u}^2 - g_0(u, v, \mu, \eta, m),\\
&b_l = d_l + a_l + thos_l + td_l + q_l var_m(\eta_l) + \bar{q}_l\bar{\eta}_l^2 + g_l(\eta, \delta, m),\\
&l \in \mathcal{L}
\end{aligned} \tag{16}$$

Then, $(u^*, v^*, \eta^*, \delta^*)$ is a mean-field-type Nash equilibrium.

The proof of Proposition 2 follows similar steps as in [20]. From the formulation of the minimization problem, the number of infected people and the number deaths are significantly reduced at the mean-field equilibrium. However, the cost-benefit analysis suggests that it depends on the functionals g_0, g_l, h_1, h_2.

The master adjoint system (MASS) is given by

$$\begin{aligned}
&U_{0,t}(l, x, s, z, c, t, \mu, m) + H_{0,\mu} = 0,\\
&U_0(l, x, s', z, c, t_1, \mu, m) = \mathbb{I}_{s' \in \{s,r,\mathrm{tr}\}}\\
&U_{l,t} + H_l = 0,\\
&U_l(t_1, m) =\\
&(d_l + a_l + t_{+,l} + \mathrm{th}_{1,l} + \mathrm{th}_{2,l} + \mathrm{tpc}_l + \mathrm{thos}_l + \mathrm{td}_l)(t_1)\\
&l \in \mathcal{L}
\end{aligned} \tag{17}$$

The MASS system provides the equilibrium strategies but not necessarily the equilibrium cost.

Other relevant payoffs

- Authorities The authority in locality l decides on
  - Migration rules (lockdown, confinement, isolation, curfew) $(\eta_{1,l}, \eta_{2,l}, \eta_{3,l})$, testing strategy δ_l

- Budget allocation and incentives ls_l, tr_l

Multiple objectives for authority in locality l :

- reduce the number of deaths, infected,
- reduce economic losses
- maximize the number of recovered-deaths

$$\begin{cases} \text{Optimize long-term objectives subject to} \\ \text{constraints} \\ \text{dynamics} \end{cases}$$

- Consumption goods firms Firms (producing essential, moderate-essential, less-essential goods) in locality l decide on

  - Production, Total hours worked $A_1 = \int a_1 m$
  - Budget constraint

  Multiple objectives for j :

  - Reduce the number of infected employees ,
  - Maximize profit

$$\begin{cases} \text{Optimize long-term objectives subject to} \\ \text{constraints} \\ \text{dynamics} \end{cases}$$

- Each individual's risk-awareness An individual in locality l decides on

  - Meeting rate a_2, local movement x, v, a_3, consumption a_1

  Multiple objectives for an individual

  - reduce the risk of being infected,

– reduce the risk of exposing the others (if co-opetitive),

– reduce economic losses

$$\begin{cases} \text{Optimize long-term objectives subject to} \\ \text{constraints, dynamics} \\ \text{(locality, position, infection status, age, pre-existing health} \\ \text{ gender, height, family size, income per family size)} \end{cases}$$

for susceptible: economic: $\log(\frac{w\ a_1+ls_g}{1+tr_g})+b_1(v,spot)-\frac{\hat{\lambda}_1}{2}a_1^2-\hat{\lambda}_2\ var_\mu(a_1)$
pandemic: $s-q\ var_\mu(a_2)-\bar{q}\bar{a}_2^2-h_0(n_{infectious}(\mathcal{B}_\epsilon(x)))\|v\|^2-\hat{\lambda}_3\ var_\mu(v)$

4 Model outputs and Data visualization

4.1 Data used and constraints

We present some illustrative examples with $K = 195$ countries and more detailed analysis for 66+ countries. The COVID-19 data set is taken from [10–13, 33]. The age-structure data per country is provided in [32]. For the international connectivity graph, η_1 we have used Flightradar24 [31], which is a global flight tracking service that provides information about flights from 1200 airlines and 4000 airports around the world. The data set does not capture all international connections but it provides a sample. As the initial propagation of COVID-19 started from flights we have modulated the flight data set to capture it. After a certain waiting time, the soft/partial/hard flight restriction decision has been made between some countries. Those countries we made $\eta_{1,l'l}$ to be zero or a very small number (0.001) if travels from l' to l are restricted. For the geographical mobility and co-location map of people in country l_1, we have used OpenStreetMap before and after confinement decisions (if any) per country. The Google Community Mobility Reports can be found in [34].

The country-specific hospital capacity data set is obtained from World Bank Group [35] and is used to capture

$$\int \sum_{c'} \text{hospitalized}(l, dx', dz', c', t) \leq \bar{c}_l(t).$$

However, the data set is limited to the country level and it is not per area/city/county. The country-specific testing capacity data set is not available but the number of tests per country is provided in [12]. The set $\mathcal{D} = \cup_{l \in \mathcal{L}} \mathcal{D}_l$ where $\mathcal{D}_l$ is a constrained non-empty set. Given two countries l_1 and l_1', the dynamics $m(l, x, s, z, c, t)$ and $m(l', x, s, z, c, t)$ will be related to each other via the switching rates η_1. Table 2 summarizes the heterogenous data sources. Figure 12 represents the proportion of infected people

$$\sum_{l,c} \int \sum_{s \in \{p, a, t_+, \mathrm{th}_1, \mathrm{tpc}, \mathrm{th}_2, \mathrm{thos}\}} m(l, dx, s, dz, c, t),$$

at the 195 countries as of March 30, 2020. Figure 13 illustrates the status of the fifteen most (reported) infected states as of April 1, 2020. Figure 14 samples G7 countries. Detailed testing results are displayed in Figure 15 for the cities of New York and Los Angeles.

We make the following observations form the model.

- the model is flexible enough to capture some local context-awareness including local mobility and age structure per country. The provided model can be filtered to track the data set provided in [12] even for some of the countries with few number of cases.

- The connectivity of the local graph η_3 plays a key role in the propagation of the virus across different areas of the city

- The connectivity of the local graph η_2 plays a key role in the propagation of the virus across different cities

- The connectivity of the local graph η_1 plays a key role in the propagation of the virus across states/countries.

- The current case fatality rate (data-CFR) per country data set is the ratio between the number of (reported) confirmed cases and the total number of (reported) deaths from COVID-19 outbreak data set [33] for that country. We observe that there are some significant differences in terms of data-CFR per country even for whose who have made similar decisions at similar infection time. Some of the differences comes from the implementation in practice and the response of the people to the measures. The proposed MFTG model captures the data-CFR per country as illustrated in 66+ countries.

4.2 Visualization

The optimization algorithm is implemented in C++ using [41]. The plots on the geographical-map (see Figure 12) are done using MATLAB. We combine several layers: the transportation layer (air and ground), economics layer, population mobility map after the first infection of COVID-19 in that city/country, and the epidemic data-driven model. We display below the outputs (tested positive, tested-recovered, tested-dead, tested-active) of the epidemic model in 66+ countries. We choose the constraints $\bar{c}_l, \bar{t}_l$ to be high enough so that the interior solution can be found. This means extending the capacity by means of new resources. The COVID-19 initial death rate is chosen as in Figure 3 for the countries where the death rate due to COVID-19 is unknown. We estimate the death rate from the model using the reported data [12]. We choose $k = 0.7m(.)$ for the numerics at the exits on the grid over OpenStreet map. The rate λ_{ap} is set 10^{-7} or zero depending on the country/locality. The other λ functions are obtained from the minimization with the smallest norm allowed to derive λ_{data}. As new data comes, λ_{data} can be recalculated. The newborns are taken from the last year growth scale. From λ_{data} we plot the behavior of the dynamics (18) and integrates it out. The initial t_0 is set to January 22, 2020 for most countries.

We observe that the proposed model captures the behavior in most countries. The data-driven MFTG model reveals that the number active COVID-19 positive patients in countries such as Venezuela, Germany, Italy, Malaysia, Senegal, Zambia, Togo have several local peaks and oscillating behaviors. Some of them cannot be fitted with a Gaussian or bi-Gaussian. It cannot be fitted with a single exponential. By changing h_l from 0.1 to 0.9 we observe a limit cycle of the dynamics. Figures 89 and 90 represents a COVID-19 sample of Non-Gaussianity and non-exponential. Also, Non-Gausianity of the reported active cases and hospitalized cases are observed in 15+ countries. It includes Albania, Azerbaijan, Bahamas, Bosnia, Costa Rica, Djibouti, Georgia, Irak, Iran, Jordan, Lebanon, Lituana, Malta, San Marino, Uzbekistan, etc. Tables 4, 5, 6 summarize lists of figures and illustrations.

Limitations of the proposed model

The proposed data-driven modelling has several interesting features including epidemiological aspects (hospitalization, tested negative, untested active,

untested deceased, untested active) as well as economics aspects (average revenue per family size, income status, on-site worked hours during COVID-19) and mobility aspects (pedestrian within a sector, within a city, between cities, provinces, countries by ground/air/sea transportation, and movement to hotspots, local markets, grocery stores, pharmacies, etc). However, the model still need extension and integration of food security and supply chain which are relevant aspects in many countries.

Acknowledgment

The author would like to thank the six webinars' audience at UCLA IPAM, NYU, Next Einstein Forum, CNRIA 2020, CDC meeting on COVID-19, and IEEE St-Maurice section for their valuable inputs. The author would like to thank Boualem Djehiche, Nader Masmoudi, Julian Barreiro-Gomez, and Salah Eddine Choutri for interesting discussions on the topic of data-driven modelling. We gratefully acknowledge support from U.S. Air Force Office of Scientific Research under grant number FA9550- 17-1-0259 on Mean-Field-Type Game Theory and the Center on Stability, Instability and Turbulence.

5 Conclusion

Understanding the transmission characteristics of infectious diseases in communities, counties, provinces, regions, states and countries can lead to better approaches to decreasing the transmission of these diseases. In this paper, a context-aware data-driven MFTG model. The proposed epidemiological and economic model can track and captures the country-specific COVID-19 data in 66+ countries. The context-aware data-driven MFTG model can be used in visualizing, comparing, planning, implementing, evaluating, and optimizing various detection, testing, prevention and control programs.

Data-Driven MFTG modeling can contribute to the design and analysis of the pandemic COVID-19 surveys, suggest crucial data that should be collected, identify trends, make short-term forecasts, and quantify the uncertainties in short-term forecasting. In addition to predictive safety measures, COVID-19 opens also other challenges such as

- Supply chain: Identifying ways to optimize supply chain among many different industries, so the impact of the COVID-19 can be minimized.

- Testing, Tracing and Tracking: Finding ways to better test, trace and track the infection propagation including those asymptomatic infected, best practices in how to test, trace and track in both rural and urban areas as early as possible.

- Community: Finding solutions to prevent social and economic disruption in all communities. Proposing innovative de-confinement strategies in how to create a smooth transition from physical distancing to guaranteeing minimum services and quality-preserving activities and back to regular activities.

Number of technical questions remain unanswered. This include (i) impact of the propagation in local economy (including informal economy), (ii) the duration of the limit cycle of active cases, (iii) the stabilization of the system by means of controls and co-opetitive decision-making, (iv) impact of COVID-19 on food insecurity. We leave these questions for future research.

References

[1] M. U. Kraemer, C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D. M. Pigott, L. du Plessis, N. R. Faria, R. Li, W. P. Hanage *et al.*, "The effect of human mobility and control measures on the covid-19 epidemic in china," *Science*, 2020.

[2] Y. Fu, Y. Cheng, and Y. Wu, "Understanding sars-cov-2-mediated inflammatory responses: from mechanisms to potential therapeutic tools," *Virologica Sinica*, pp. 1–6, 2020.

[3] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, J. Y. Wong *et al.*, "Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia," *New England Journal of Medicine*, 2020.

[4] C. Chen and B. Zhao, "Makeshift hospitals for covid-19 patients: where health-care workers and patients need sufficient ventilation for more protection," *Journal of Hospital Infection*, 2020.

[5] D. Balcan, B. Gonçalves, H. Hu, J. Ramasco, V. Colizza, and A. Vespignani, "Modeling the spatial spread of infectious diseases: the global

epidemic and mobility computational model," *Journal of computational science*, vol. 1, no. 3, pp. 132–145, 08 2010.

[6] J. B. Dowd, L. Andriano, D. M. Brazel, V. Rotondi, P. Block, X. Ding, Y. Liu, and M. C. Mills, "Demographic science aids in understanding the spread and fatality rates of covid-19," *Proceedings of the National Academy of Sciences*, 2020.

[7] G. Onder, G. Rezza, and S. Brusaferro, "Case-fatality rate and characteristics of patients dying in relation to covid-19 in italy," *Jama*, 2020.

[8] J. T. Wu, K. Leung, M. Bushman, N. Kishore, R. Niehus, P. M. de Salazar, B. J. Cowling, M. Lipsitch, and G. M. Leung, "Estimating clinical severity of covid-19 from the transmission dynamics in wuhan, china," *Nature Medicine*, pp. 1–5, 2020.

[9] B. Xu, B. Gutierrez, S. Mekaru, K. Sewalk, L. Goodwin, A. Loskill, E. L. Cohn, Y. Hswen, S. C. Hill, M. M. Cobo *et al.*, "Epidemiological data from the covid-19 outbreak, real-time case information," *Scientific data*, vol. 7, no. 1, pp. 1–6, 2020.

[10] CSSE, *Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) https://github.com/CSSEGISandData/COVID-19*, 2020.

[11] WHO, *World Health Organization (WHO), https://www.who.int/*, 2020.

[12] WorldoMeters, *https://www.worldometers.info/coronavirus/*, 2020.

[13] COVID, *COVID Tracking Project, https://covidtracking.com/data.*, 2020.

[14] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, J. Y. Wong *et al.*, "Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia," *New England Journal of Medicine*, 2020.

[15] A. K. Cissé and H. Tembine, "Cooperative mean-field type games," in *Proceedings of the 19th World Congress The International Federation of Automatic Control*, Cape Town, South Africa, 2014, pp. 8995–9000.

[16] H. Tembine, "Risk-sensitive mean-field-type games with l^p-norm drifts," *Automatica*, vol. 59, pp. 224–237, 2015.

[17] ——, "Uncertainty quantification in mean-field-type teams and games," in *Proceedings of the IEEE Control Conference on Decision and Control (CDC)*, Osaka, Japan, 2015, pp. 4418–4423.

[18] ——, "Mean-field-type games," *AIMS Mathematics*, vol. 2, no. 4, pp. 706–735, 2017.

[19] B. Djehiche, A. Tcheukam, and H. Tembine, "Mean-field-type games in engineering," *AIMS Electronics and Electrical Engineering*, vol. 1, no. 4, pp. 18–73, 2017.

[20] A. Bensoussan, B. Djehiche, H. Tembine, and S. C. P. Yam, "Mean-field-type games with jump and regime switching," *Dynamic Games and Applications*, pp. 1–39, 2019.

[21] A. S. Tcheukam and H. Tembine, "On the distributed mean-variance paradigm," *13th International Multi-Conference on Systems, Signals & Devices. Conference on Systems, Automation & Control - Leipzig, Germany*, pp. pp. 604–609, 2016.

[22] T. Duncan and H. Tembine, "Linear-quadratic mean-field-type games: A direct method," *Games*, vol. 9, no. 7, 2018.

[23] J. Barreiro-Gomez, T. E. Duncan, and H. Tembine, "Linear-quadratic mean-field-type games: Jump-diffusion process with regime switching," *IEEE Transactions on Automatic Control*, vol. 64, no. 10, pp. 4329–4336, 2019.

[24] J. Barreiro-Gomez, T. E. Duncan, B. Pasik-Duncan, and H. Tembine, "Semi-explicit solutions to some non-linear non-quadratic mean-field-type games: A direct method," *IEEE Transactions on Automatic Control*, pp. 1–14, 2020, dOI 10.1109/TAC.2019.2946337.

[25] J. Barreiro-Gomez, T. E. Duncan, and H. Tembine, "Discrete-time linear-quadratic mean-field-type repeated games: Perfect, incomplete, and imperfect information," *Automatica*, vol. 112, p. 108647, 2020.

[26] H. Tembine, P. Vilanova, and M. Debbah, "Noisy mean field game model for malware propagation in opportunistic networks," *Game Theory for Networks - 2nd International ICST Conference, GAMENETS, Shanghai, China, Revised Selected Papers, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Springer*, vol. 75, pp. 459–474, 2011.

[27] H. Tembine, "Population games in large-scale networks: Time delays, mean field dynamics and applications," *LAP Lambert Academic Publishing*, 2010.

[28] H. Tembine, P. Vilanova, and M. Debbah, "Noisy mean-field stochastic games with network applications," *Report, Ecole superieure delectricite, Paris*, 2010.

[29] H. Tembine, R. Tempone, and P. Vilanova, "Mean field interaction in biochemical reaction networks," *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 991–997, 2011.

[30] A. Tcheukam and H. Tembine, "Spatial mean-field games for combatting corruption propagation," *2016 Chinese Control and Decision Conference (CCDC)*, pp. 820–825, 2016.

[31] FlightRadar24, "National and international flight connections," *https://www.flightradar24.com/*, 2020.

[32] Data, "Age structure data per country," *https://ourworldindata.org/age-structure*.

[33] A. Zlojutro, D. Rey, and L. Gardner, "Optimizing border control policies for global outbreak mitigation," *Scientific Reports 9:2216. DOI https://doi.org/10.1038/s41598-019-38665-w*, 2019.

[34] Google, "Covid-19 community mobility reports," 2020.

[35] WorldBank, "Hospital capacity data," *https://data.worldbank.org/indicator/SH.MED.BEDS.ZS*, 2020.

[36] L. Bourouiba, E. Dehandschoewercker, and J. W. Bush, "Violent expiratory events: on coughing and sneezing," *Journal of Fluid Mechanics*, vol. 745, pp. 537–563, 2014.

[37] J. Gao and H. Tembine, "Distributed mean-field-type filters for traffic networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 507–521, 2019. [Online]. Available: https://doi.org/10.1109/TITS.2018.2816811

[38] ——, "Distributed mean-field-type filter for vehicle tracking," *2017 American Control Conference, ACC 2017, Seattle, WA, USA, May 24-26, 2017*, pp. 4454–4459, 2017. [Online]. Available: https://doi.org/10.23919/ACC.2017.7963641

[39] ——, "Correlative mean-field filter for sequential and spatial data processing," *IEEE EUROCON 2017 -17th International Conference on Smart Technologies, Ohrid, Macedonia, July 6-8, 2017*, pp. 243–248, 2017. [Online]. Available: https://doi.org/10.1109/EUROCON.2017.8011113

[40] K. J. Law, H. Tembine, and R. Tempone, "Deterministic mean-field ensemble kalman filtering," *SIAM Journal on Scientific Computing*, vol. 38, no. 3, pp. A1251–A1279, 2016.

[41] B. Djehiche, A. Tcheukam, and H. Tembine, "A mean-field game of evacuation in multi-level building," *IEEE Transactions on Automatic Control*, vol. 62, no. 10, pp. 5154–5169, 2017.

$m = (m(l,x,s,z,c,t))_{l,s,c},$	
$m_t = -\mathrm{div}_x(vm) - div_v(u_3 m) + \frac{1}{2}\mathrm{trace}[(\sigma^*\sigma m)_{xx}]$	movement and diffusion with
$+\int_\Theta [m(l,x-\gamma,s,z,c,t) - m(l,x,s,z,c,t) + \langle m_x,\gamma\rangle]\nu(d\theta)$	jump
$+\int_{\mathcal{B}_v(x)} m(l,dx',s,z,t)q(x,s,z;(l,x',s,z,t,m,v))$	
$-m\int_{\mathcal{B}_v(x)} q(dx',s,z;(l,x,s,z,t,m,v))$	
$+\sum_{s'\neq s,c'\neq c}\int m(l,dx',s',dz',c',t)\tilde{q}(x,s,z,c;(l,x',s',z',c',t,m))$	
$-m(l,x,s,z,c,t)\sum_{s'\neq s,c'\neq c}\int \tilde{q}(dx',s',dz',c';(l,x,s,z,c,t,m))$	local spread of the disease: c
	local spread of the disease vi
$+\sum_{l'\neq l_3}\int m(l',dx',s',z,c,t)\eta_3(l,x,s,z,c;(l',x',s',z,c,t,m))$	
$-m(l,x,s,z,c,t)\sum_{l'\neq l_3}\int \eta_3(l',dx',s',z,c;(l,x,s,z,c,t,m))$	intra-city mobility in l_3
$+\sum_{l'\neq l_2}\int m(l',dx',s',z,c,t)\eta_2(l,x,s,z,c;(l',x',s',z,c,t,m))$	
$-m(l,x,s,z,c,t)\sum_{l'\neq l_2}\int \eta_2(l',dx',s',z,c;(l,x,s,z,c,t,m))$	inter-city mobility from/to l_2
$+\sum_{l'\neq l_1}\int m(l',dx',s',z,c,t)\eta_1(l,x,s,z,c;(l',x',s',z,c,t,m))$	
$-m(l,x,s,z,c,t)\sum_{l'\neq l}\int \eta_1(l',dx',s',z,c;(l,x,s,z,c,t,m))$	international connectivity fro
$-\tilde{v}m_z - \tilde{d}(l,z,c,m)m$	aging and death
$\int m(l,dx,\mathrm{hospitalized},dz,t) \leq \bar{c}_l(t),$	
$\int m(l,dx,\mathrm{testing},dz,t) \leq \bar{t}_l(t),$	
$m(l,x,s,z,c,0) = m_0(l,x,s,z,c)$	initial population
$m(l,x,s,0,c,t) = \int \tilde{b}(l,x,s,z,c,t,m)m(l,x,s,dz,c,t),$	
$x \in \mathcal{D}_l,$	City map constraint
$v_{\mid\partial\mathcal{D}_l} = 0,$	map constraint
$v_{Exits} = k(l,x,m),$	Allowed/feasible mobility are
$s' \in \{s, sq, p, pq, a, aq, r, d\},$	infection status,
$s \in \{s, e, p, r, d, t_-, t_u, t_+, \mathrm{home}, pc, \mathrm{hotel}, tt, \mathrm{hospitalized}, tr, td\},$	infection status,

Illustration	Meaning
Figure 1	Flow diagram
Figure 2	first wave of spread
Figure 3	death rate distribution by age
Figure 4	sample age distribution
Figure 6	evolution of the epidemiological ODE with 17 states
Figure 7	connection with SIR
Figure 8	connection with SEIR
Figure 9	connection with SIRD
Figure 10	connection with SEIRD
Figure 11	pandemic vs economic consequences of COVID-19
Figure 12	spread worldwide
Figure 13	propagation of the virus in 15 most reported active cases as of April 11, 2020
Figure 14	G7 countries
Figure 15	New York and Los Angeles testing results: data vs model
Figure 16	Brazil: Sequential data-driven model
Figure 17	USA: Sequential data-driven model

Table 4: List 1 of figures and illustrations

Illustration	Meaning
Figure 64	Argentina : data vs model
Figure 65	Bolivia: data vs model
Figure 66	Brazil : data vs model
Figure 67	Chile : data vs model
Figure 68	Colombia : data vs model
Figure 69	Ecuador : data vs model
Figure 70	: data vs model
Figure 71	Paraguay : data vs model
Figure 72	Uruguay : data vs model
Figure 73	Venezuela : data vs model
Figure 74	Cuba : data vs model
Figure 75	Guatemala : data vs model
Figure 76	Haiti : data vs model
Figure 77	Mexico : data vs model
Figure 78	Bangladesh : data vs model
Figure 79	India : data vs model
Figure 80	Indonesia : data vs model
Figure 81	Korea : data vs model
Figure 82	Malaysia : data vs model
Figure 83	Pakistan : data vs model
Figure 84	UAE : data vs model
Figure 85	Germany : data vs model
Figure 86	Italy : data vs model
Figure 87	Russia : data vs model
Figure 88	Spain : data vs model
Figure 89	Non-Gaussianity and non-exponential
Figure 90	Not a single exponential

Table 5: List 2 of figures and illustrations. data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

Illustration	Meaning
Figure 18	Covid-19 samples in selected countries in Africa
Figure 19	Selected COMESA countries
Figure 20	Selected ECOWAS countries
Figure 21	Selected SADC countries
Figure 22	Tunisia: Sequential data-driven model
Figure 23	Algeria: data vs model
Figure 24	Angola: data vs model
Figure 25	Botswana: data vs model
Figure 26	Benin: data vs model
Figure 27	Burkina Faso : data vs model
Figure 28	Burundi : data vs model
Figure 29	Cameroon : data vs model
Figure 30	Chad : data vs model
Figure 31	Djibouti : data vs model
Figure 32	Egypt : data vs model
Figure 33	Eritrea : data vs model
Figure 34	Eswatini : data vs model

Table 6: List 3 of figures and illustrations. data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

Illustration	Meaning
Figure 35	Ethiopi : data vs model
Figure 36	Gabon : data vs model
Figure 37	Ghana : data vs model
Figure 38	Guinea : data vs model
Figure 39	Guinea Bissau : data vs model
Figure 40	Kenya : data vs model
Figure 41	Liberia : data vs model
Figure 42	Libya : data vs model
Figure 43	Madagascar : data vs model
Figure 44	Malawi : data vs model
Figure 45	Mali : data vs model
Figure 46	Mauritania : data vs model
Figure 47	Mauritius : data vs model
Figure 48	Morocco : data vs model
Figure 49	Mozambique : data vs model
Figure 50	Namibia : data vs model
Figure 51	Niger : data vs model
Figure 52	Nigeria : data vs model
Figure 53	Rwanda : data vs model
Figure 54	Senegal : data vs model
Figure 55	Seychelles : data vs model
Figure 56	Somalia : data vs model
Figure 57	South Africa : data vs model
Figure 58	Sudan : data vs model
Figure 59	Togo : data vs model
Figure 60	Tunisia : data vs model
Figure 61	Uganda : data vs model
Figure 62	Zambia : data vs model
Figure 63	Zimbabwe : data vs model

Table 7: List 4 of figures and illustrations. data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

Figure 2: First flow of infection of COVID-19 from Asia to Europe to America and then to Africa.

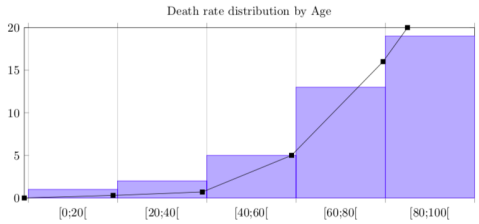

Figure 3: Death rate distribution by age.

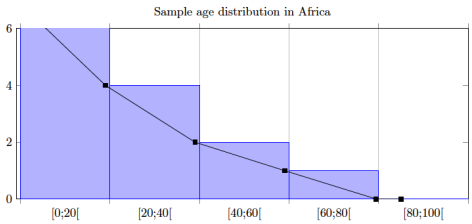

Figure 4: A sample age distribution (unnormalized) for a specific sector.

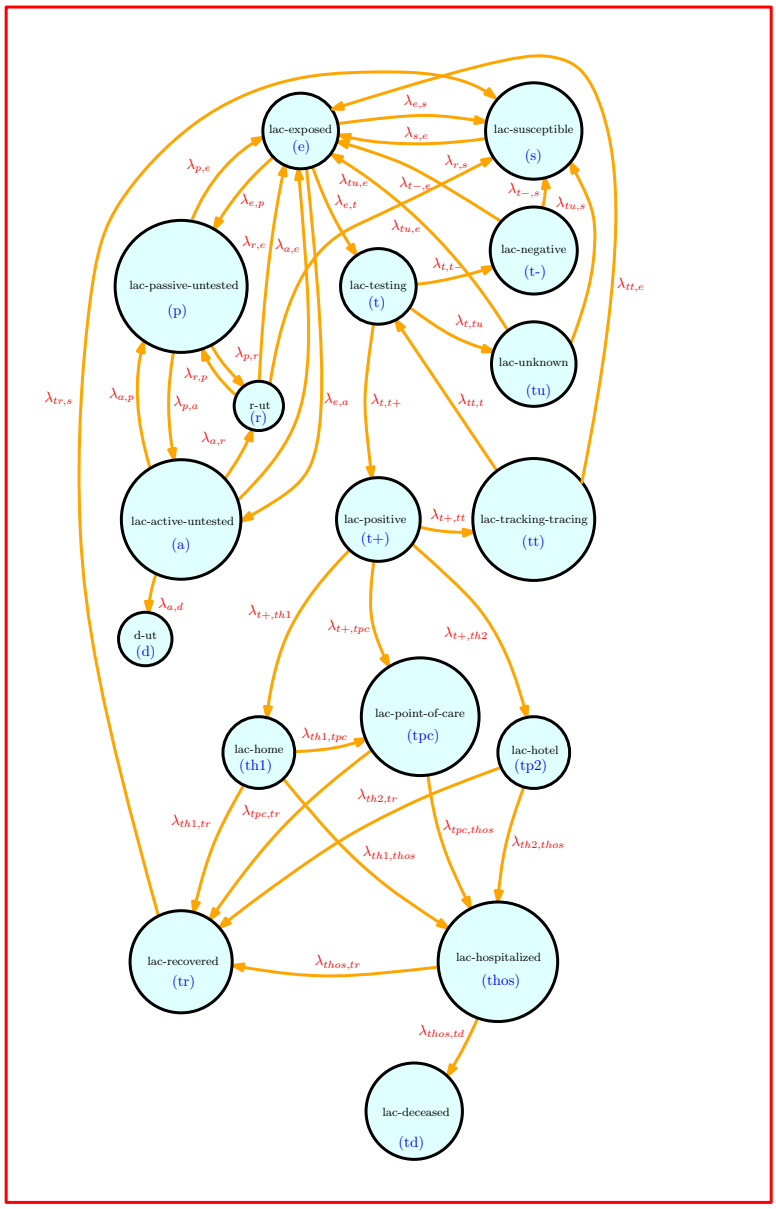

Figure 5: Intra-dynamics at a given location with untested and tested options. Home, hotel, point-of-care isolation. l.a.c=location(l)-age (a)-pre-existing condition (c), gender, height, family size, economic status

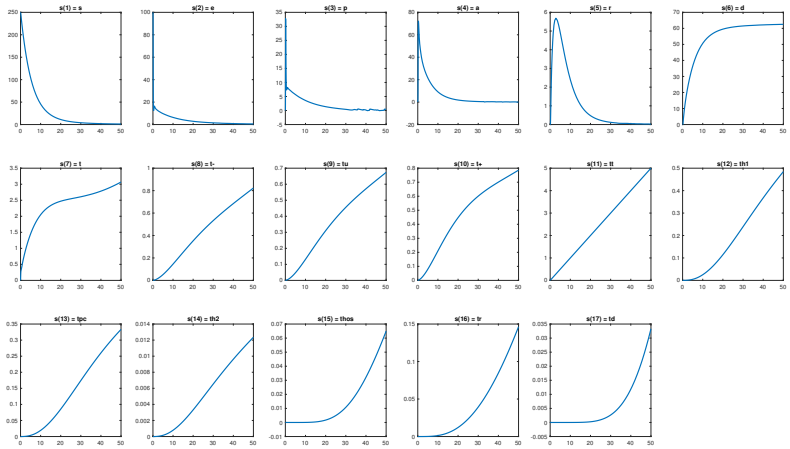

Figure 6: Evolution of the ODE with 17 infection states.

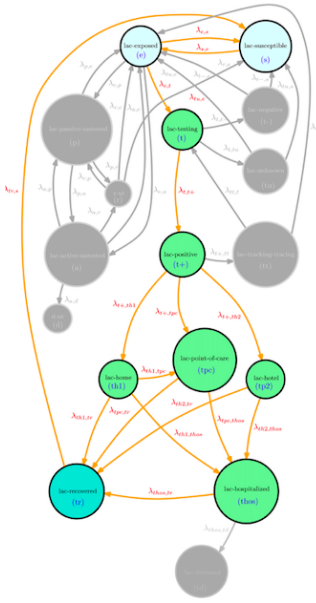

Figure 7: Connection of the dynamics of SIR

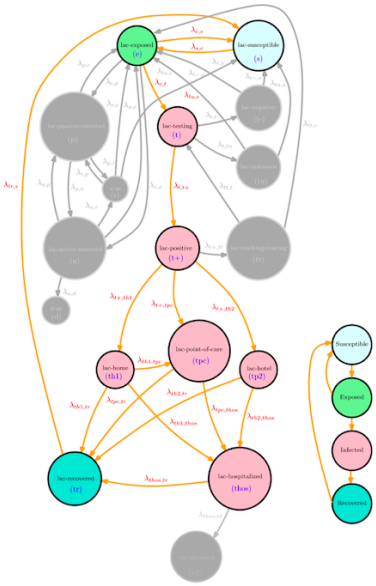

Figure 8: Connection of the dynamics of SEIR

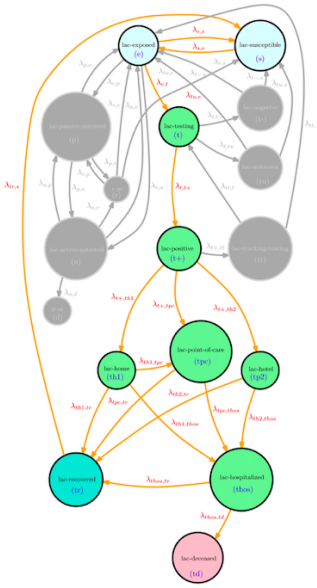

Figure 9: Connection of the dynamics of SIRD

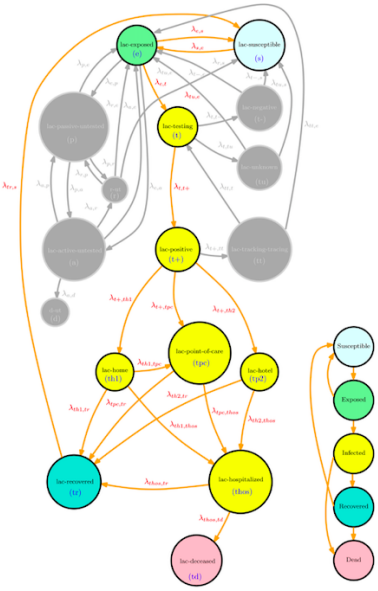

Figure 10: Connection of the dynamics of SEIRD

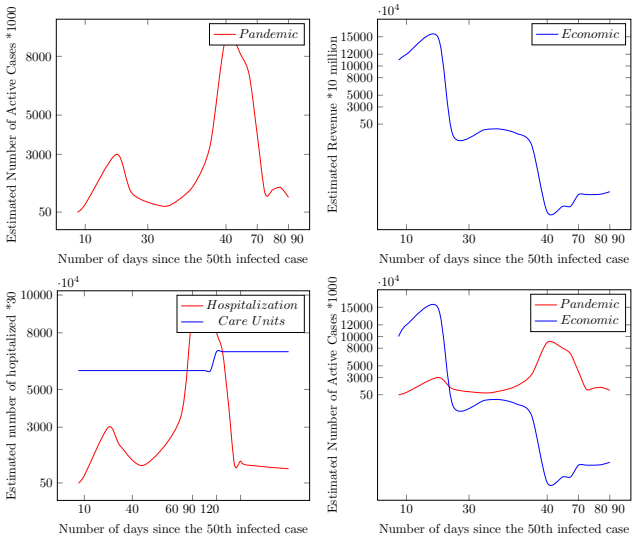

Figure 11: Pandemic vs Economic Consequences of COVID-19.

Figure 12: World map: proportion of infected as of March 30, 2020

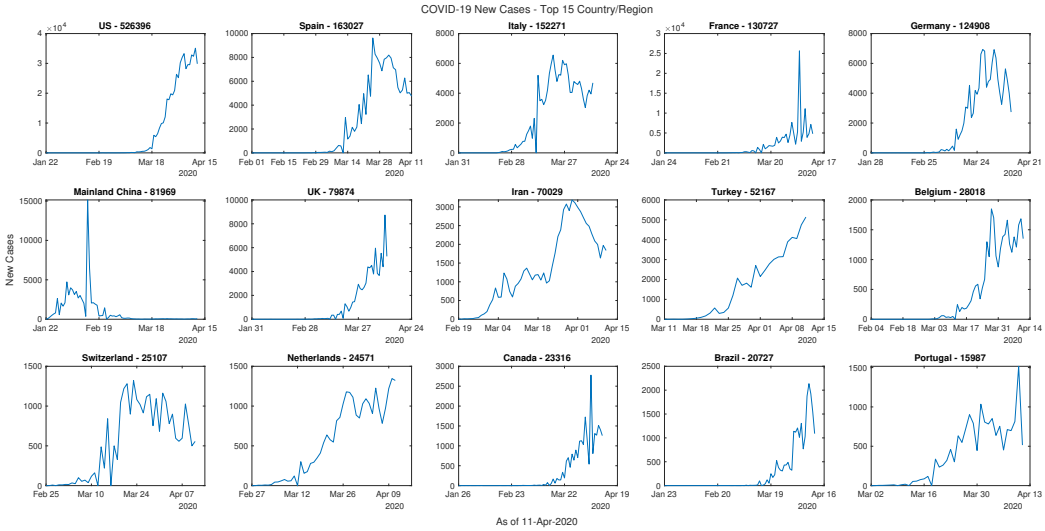

Figure 13: Propagation of the virus in 15 most reported active cases as of April 11, 2020.

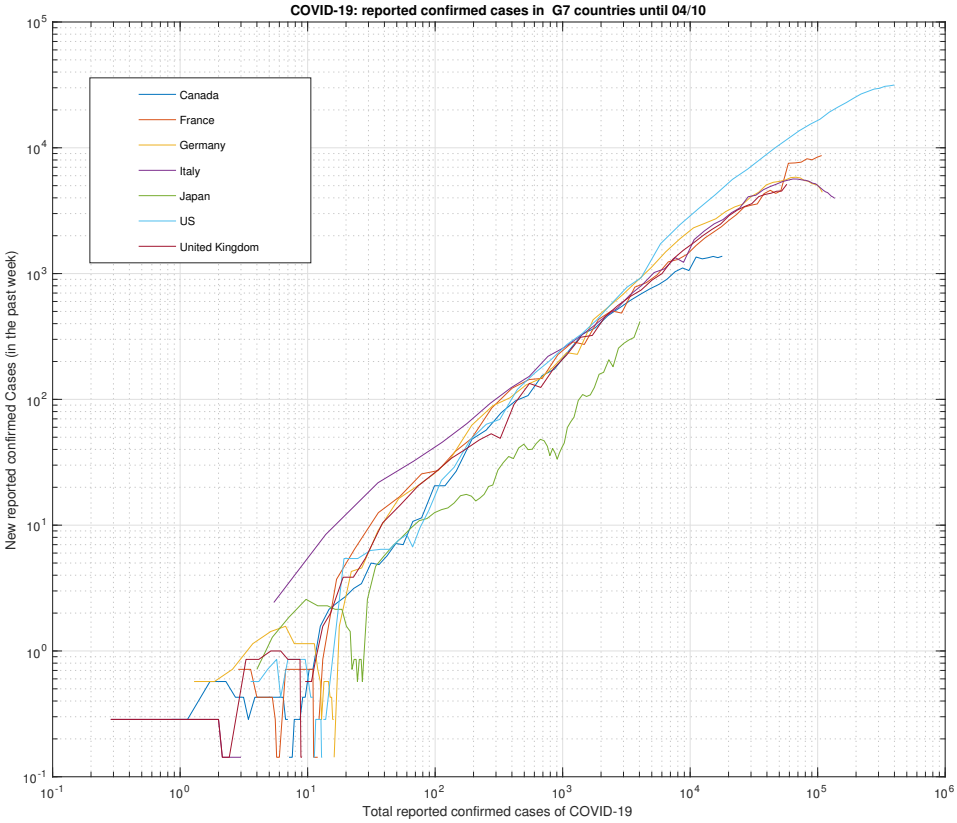

Figure 14: COVID-19 samples in G7 countries

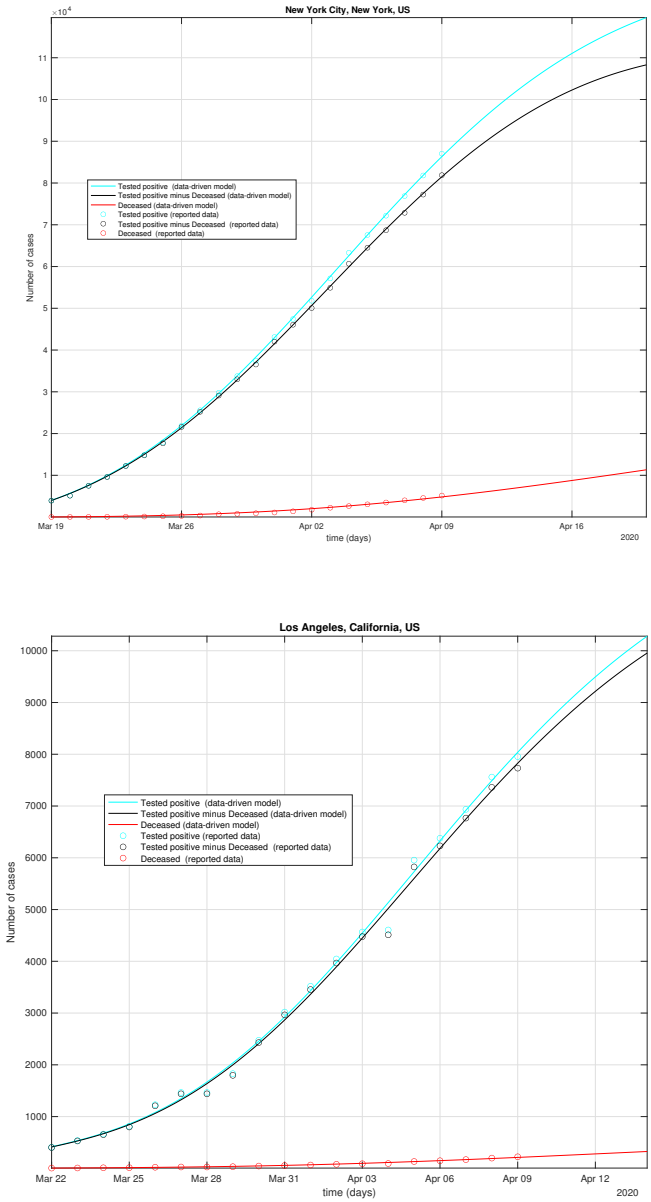

Figure 15: New York and Los Angeles testing results: data vs model.

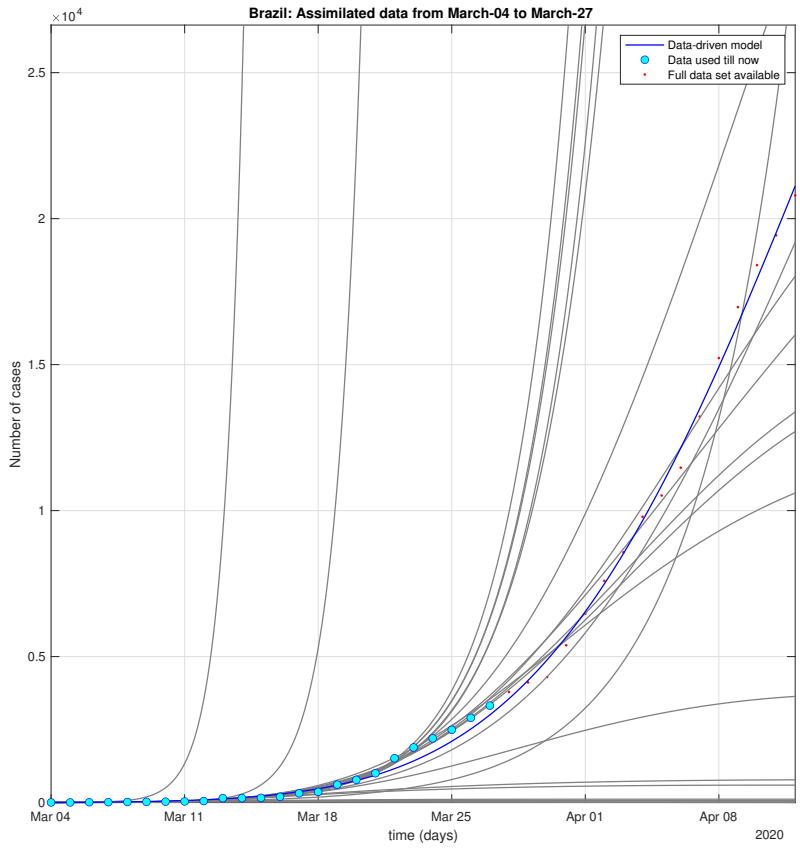

Figure 16: Brazil: Sequential data-driven model.

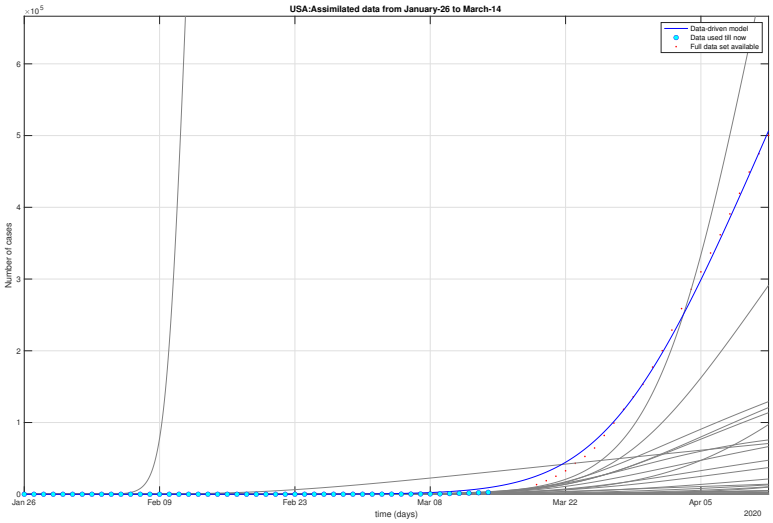

Figure 17: USA: Sequential data-driven model.

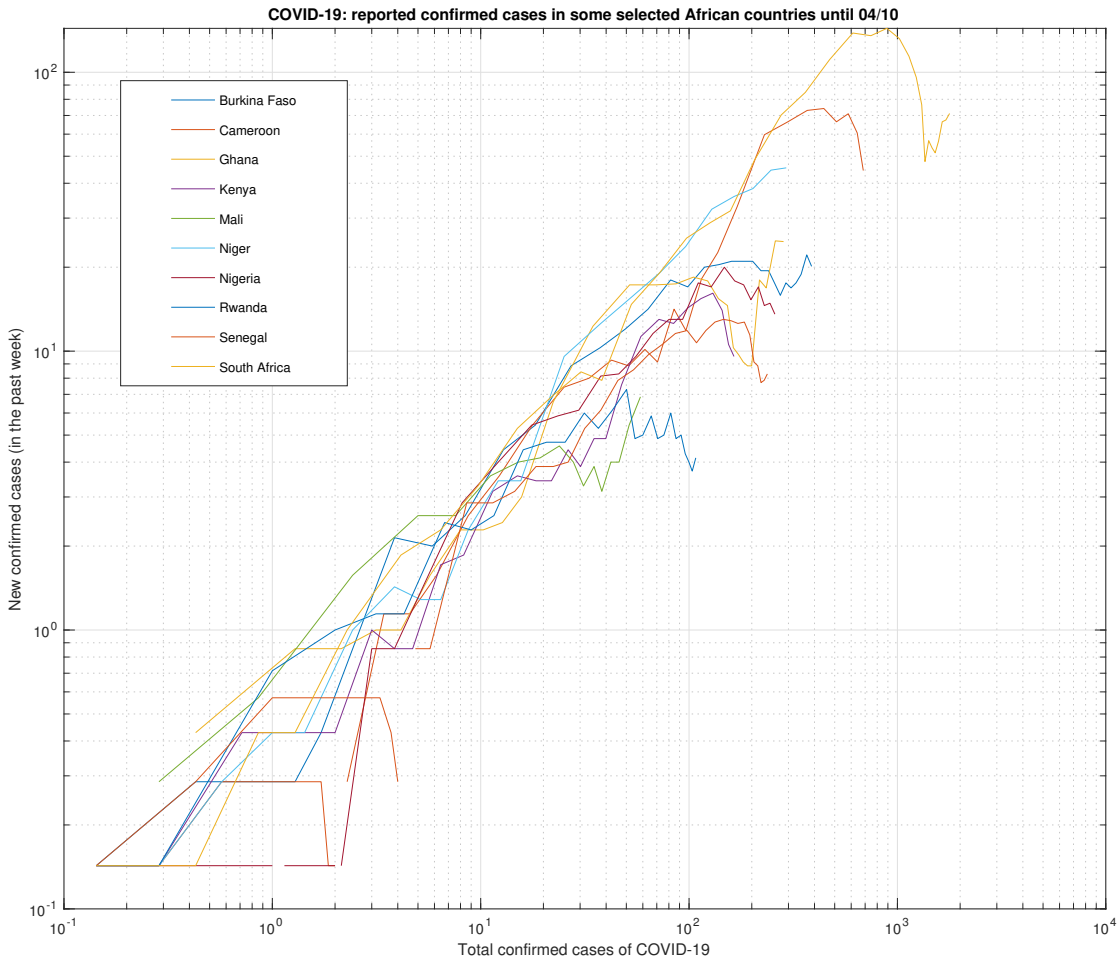

Figure 18: Covid-19 samples in selected countries in Africa

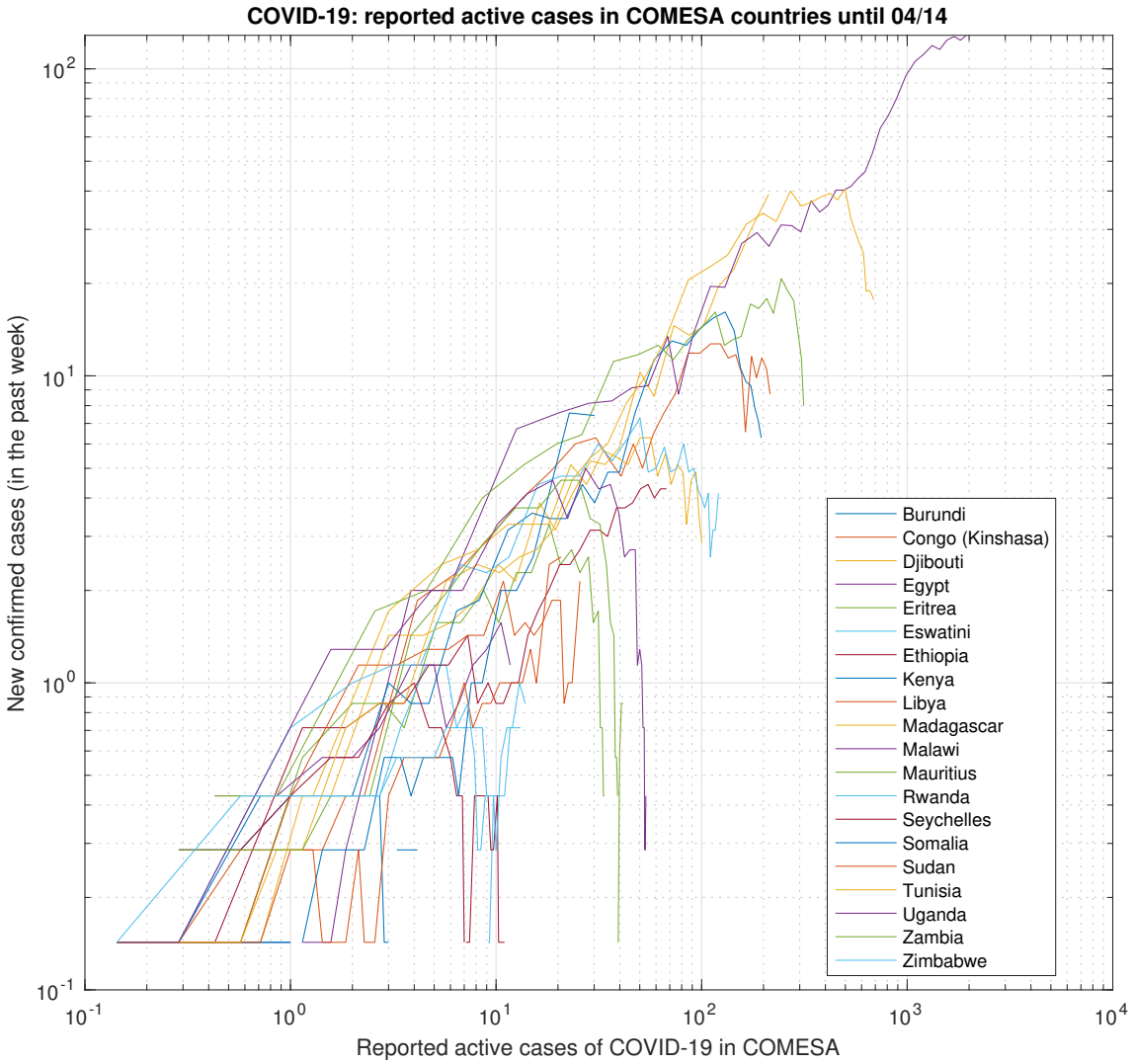

Figure 19: Selected COMESA countries

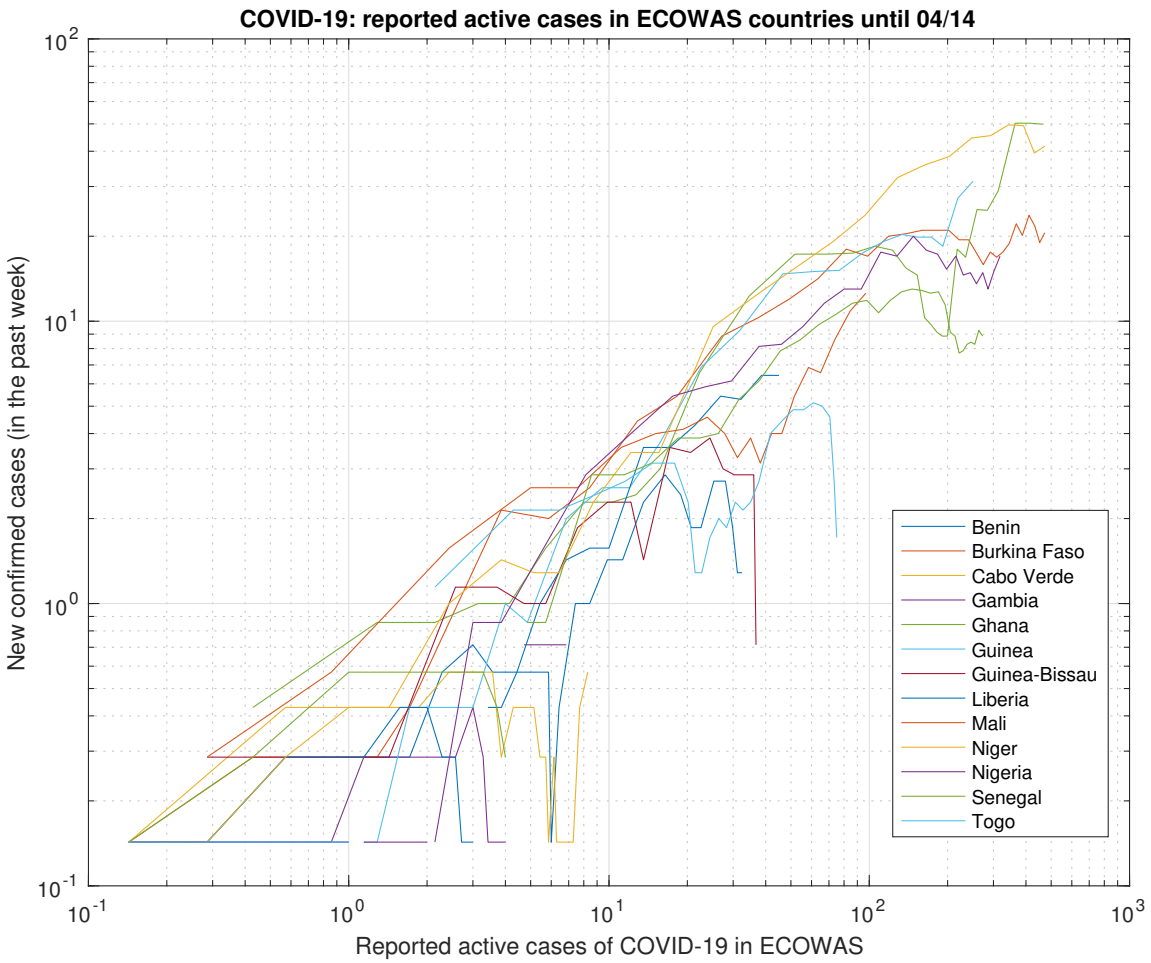

Figure 20: Selected ECOWAS countries

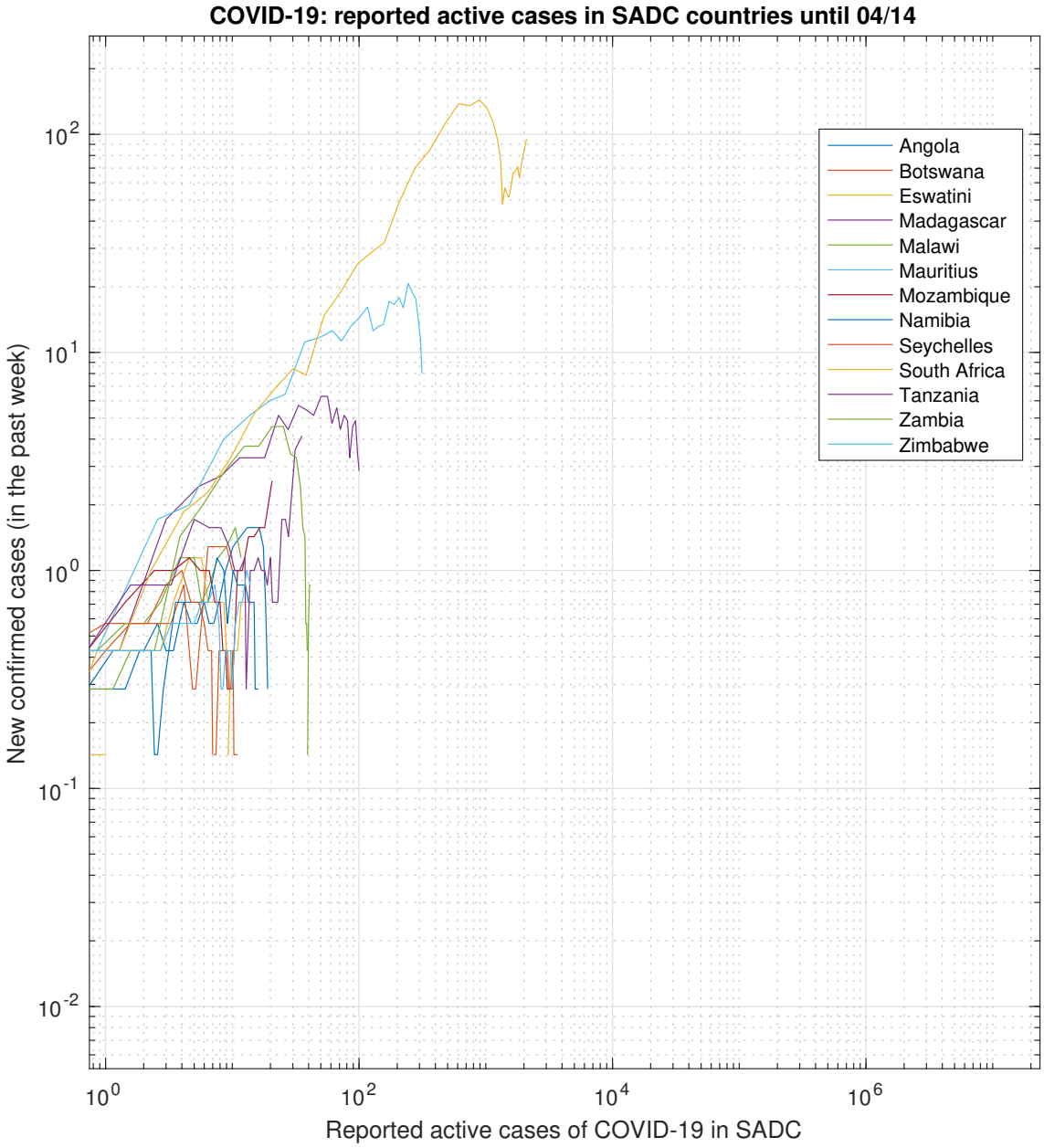

Figure 21: Selected SADC countries

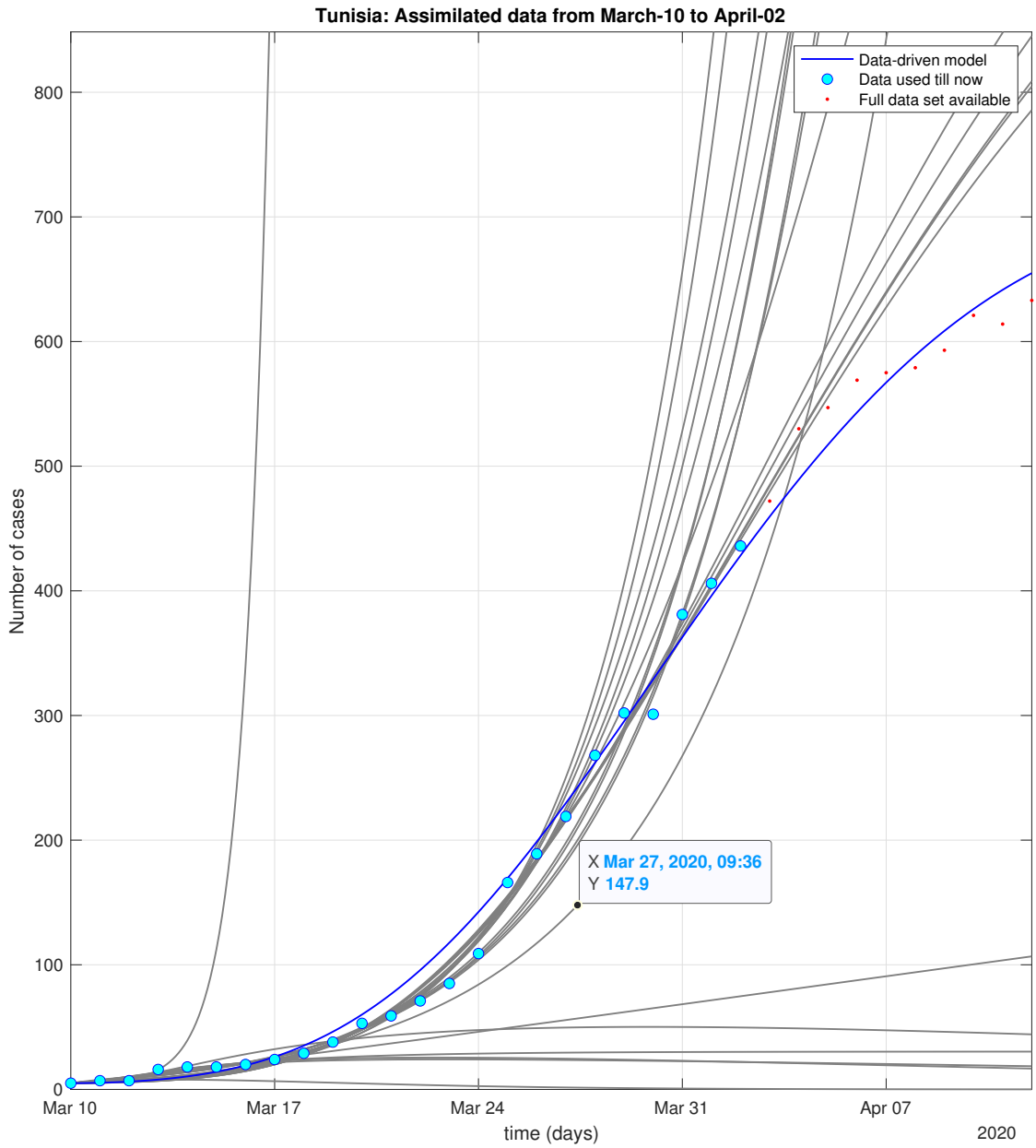

Figure 22: Tunisia: Sequential data-driven model

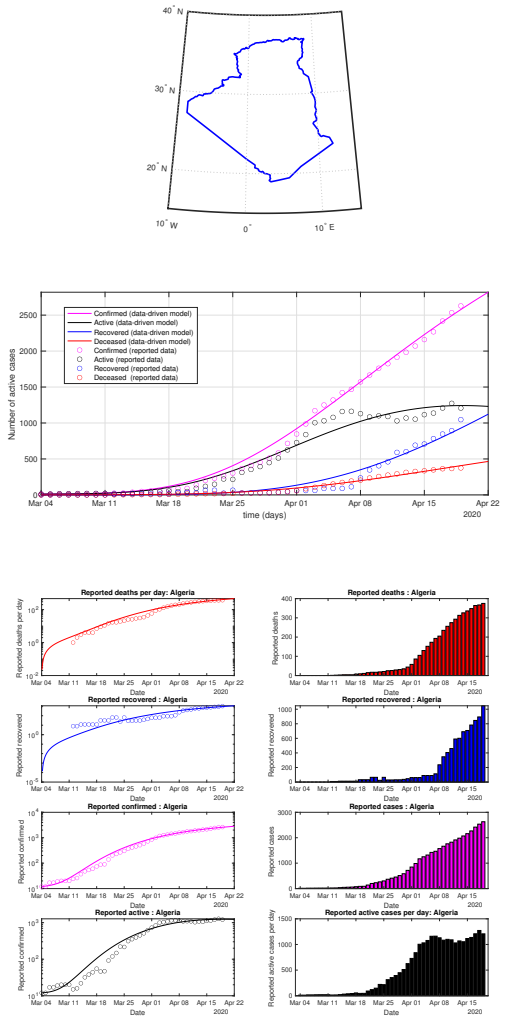

Figure 23: Algeria: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

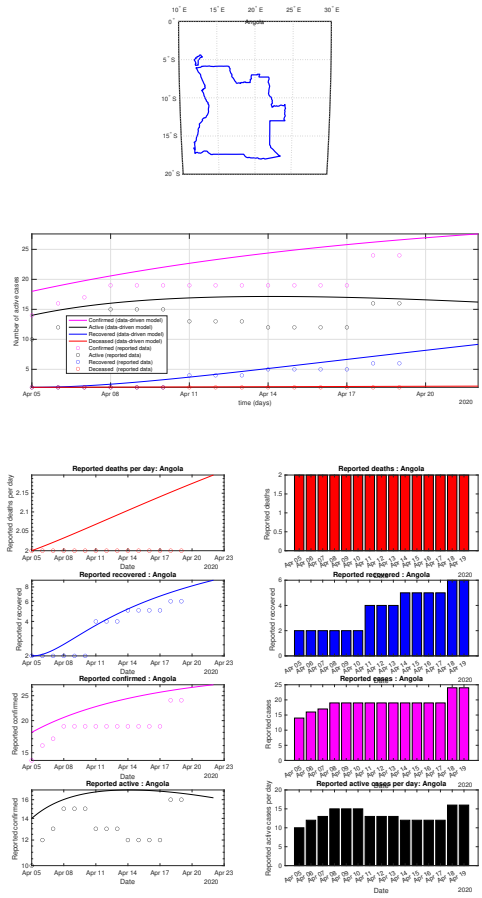

Figure 24: Angola: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

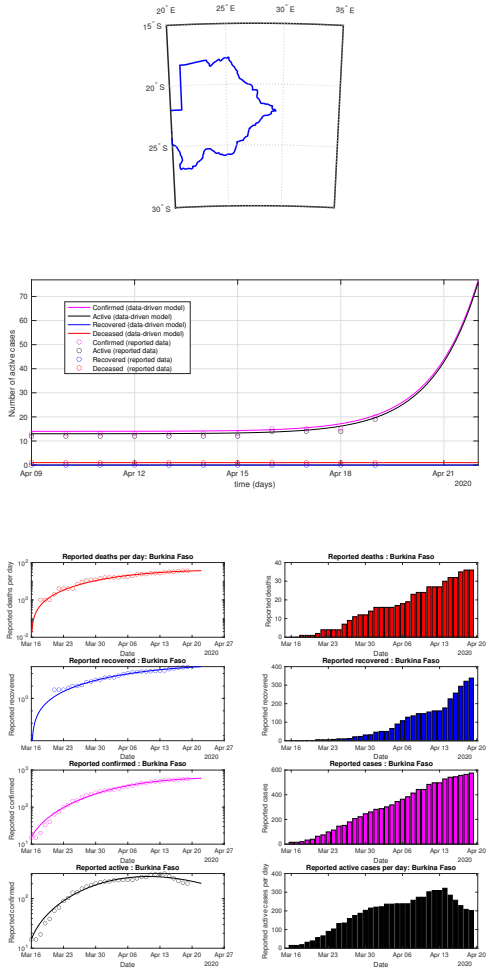

Figure 25: Botswana: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

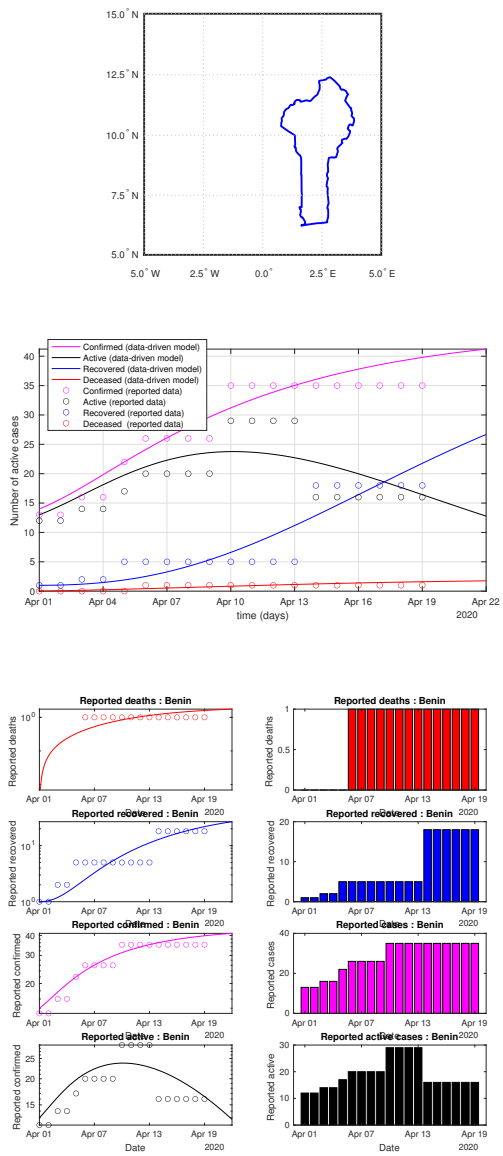

Figure 26: Benin: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

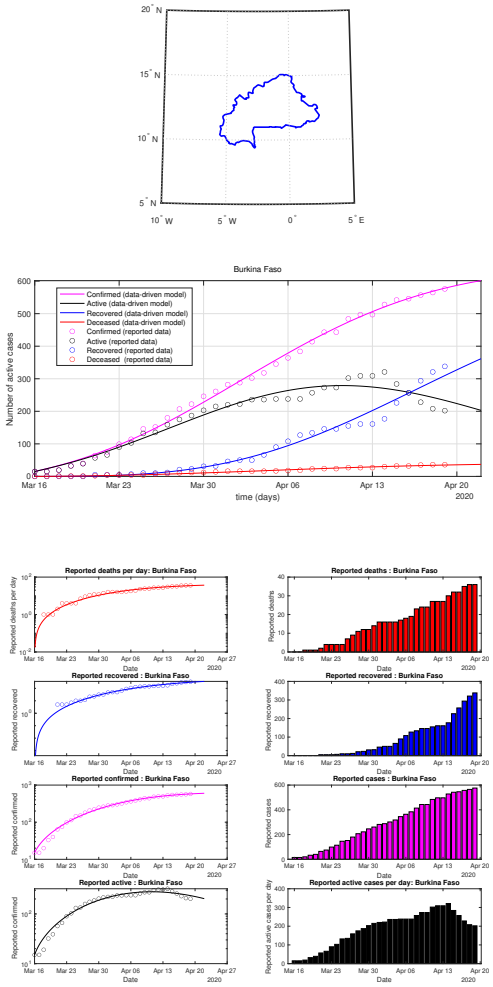

Figure 27: Burkina Faso: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

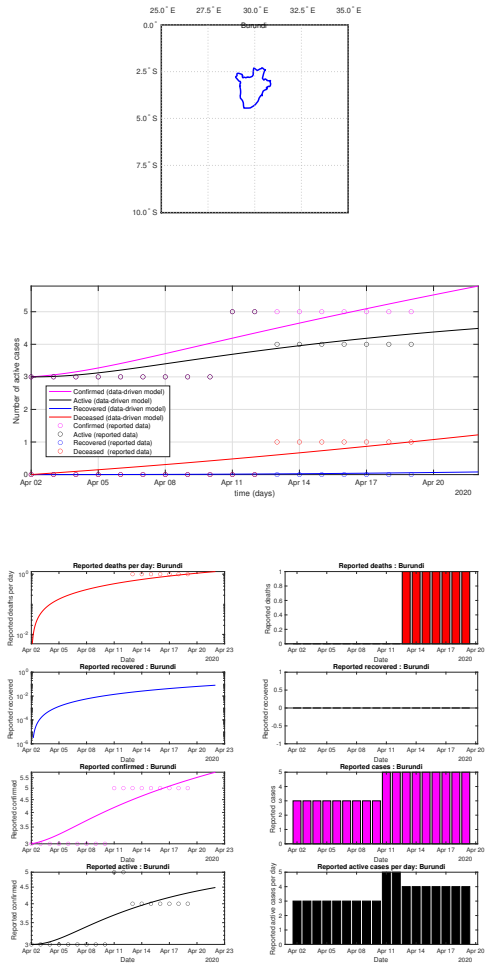

Figure 28: Burundi: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

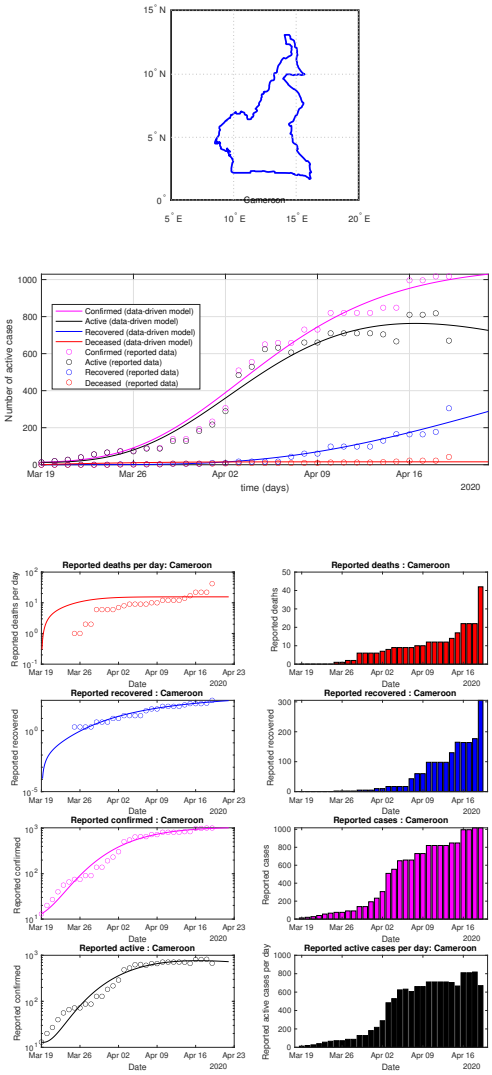

Figure 29: Cameroon: data vs model

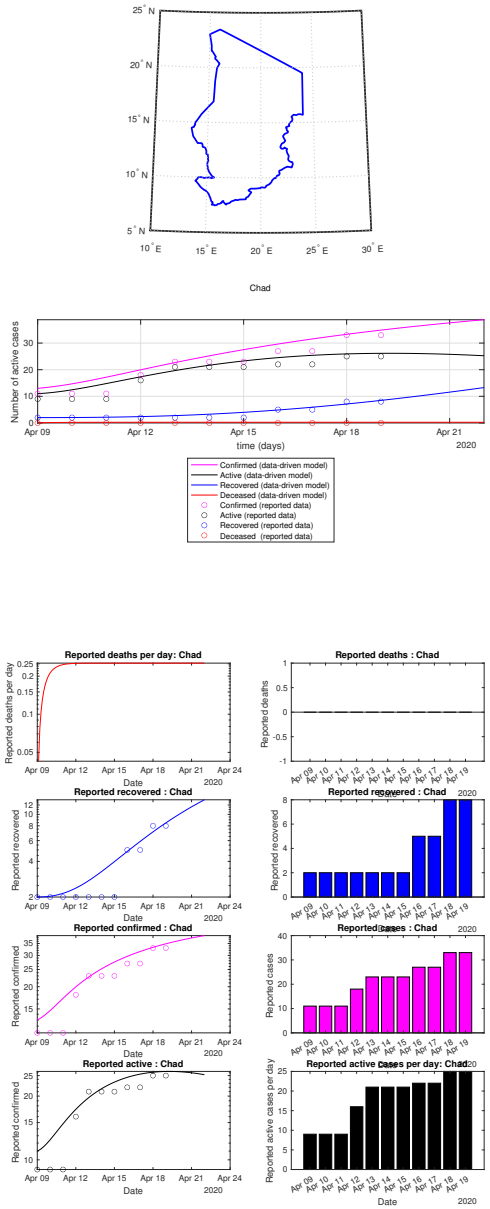

Figure 30: Chad: data vs model

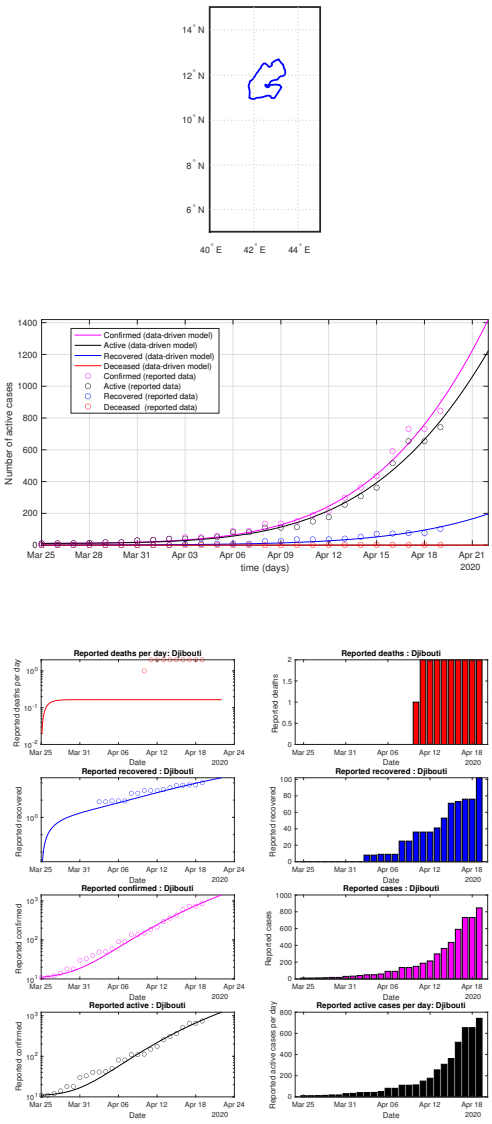

Figure 31: Djibouti: data vs model

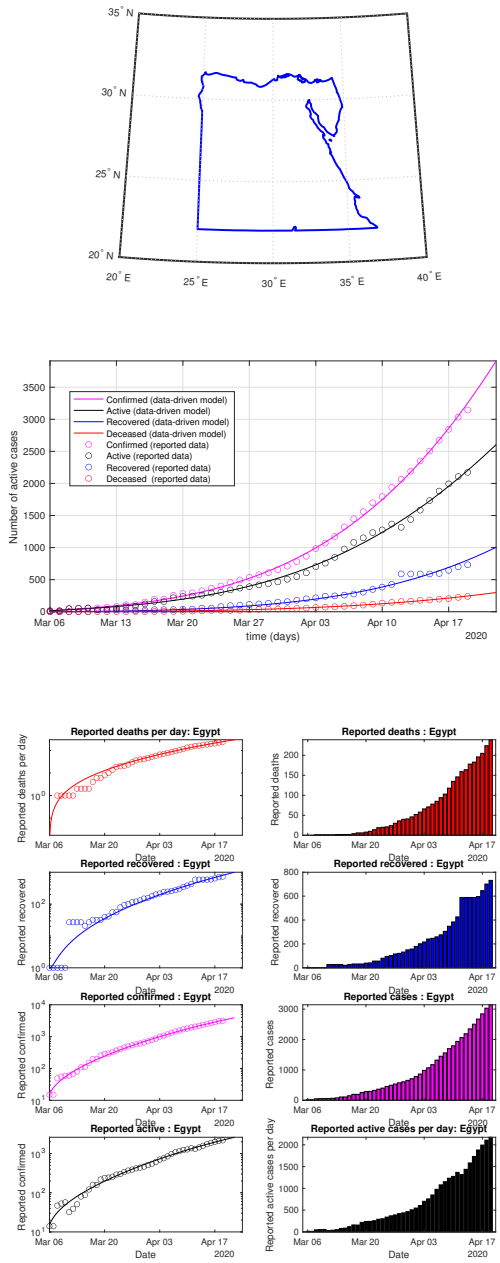

Figure 32: Egypt: data vs model

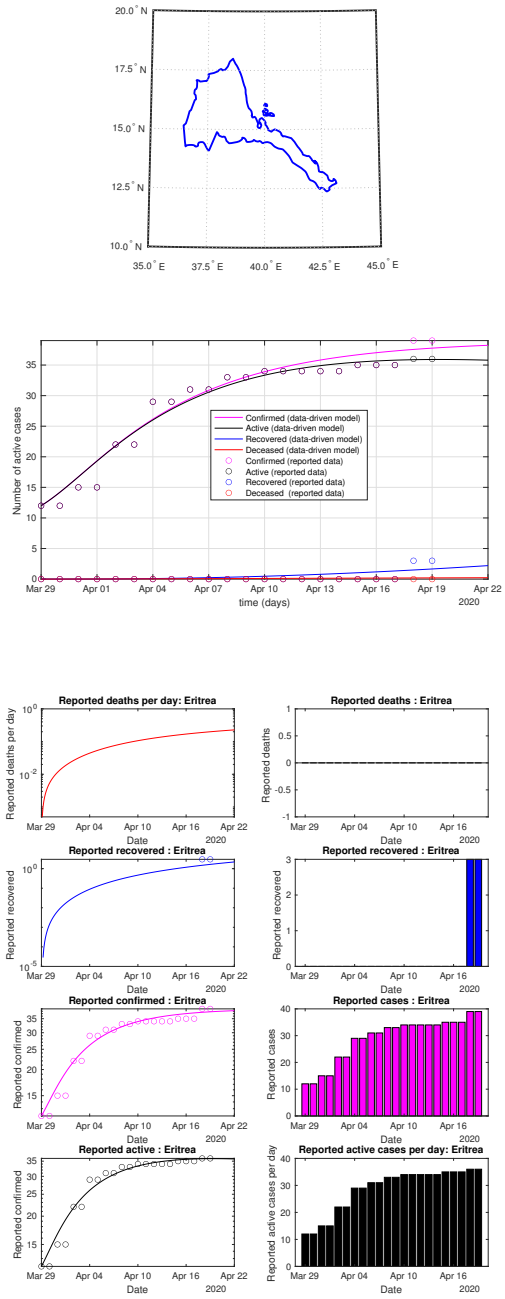

Figure 33: Eritrea: data vs model

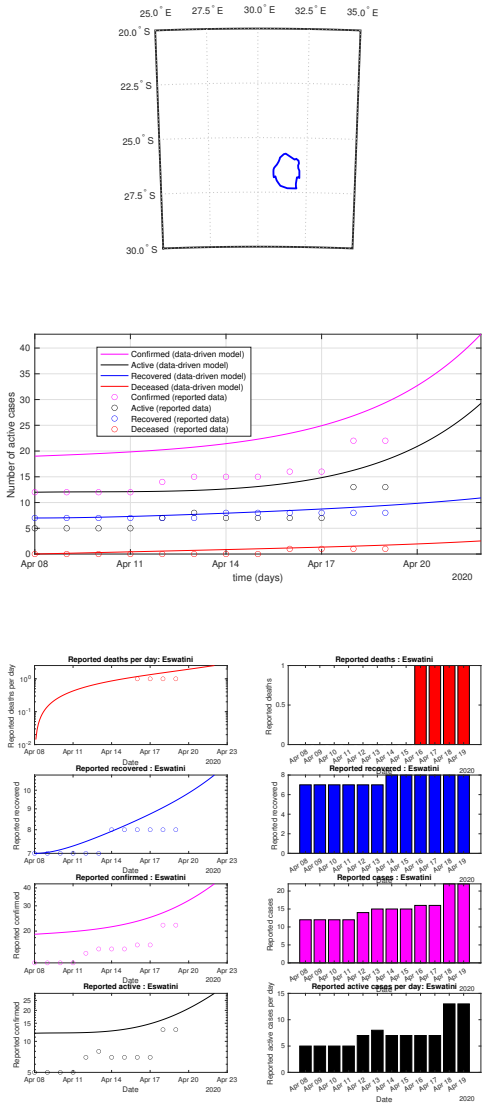

Figure 34: Eswatini: data vs model

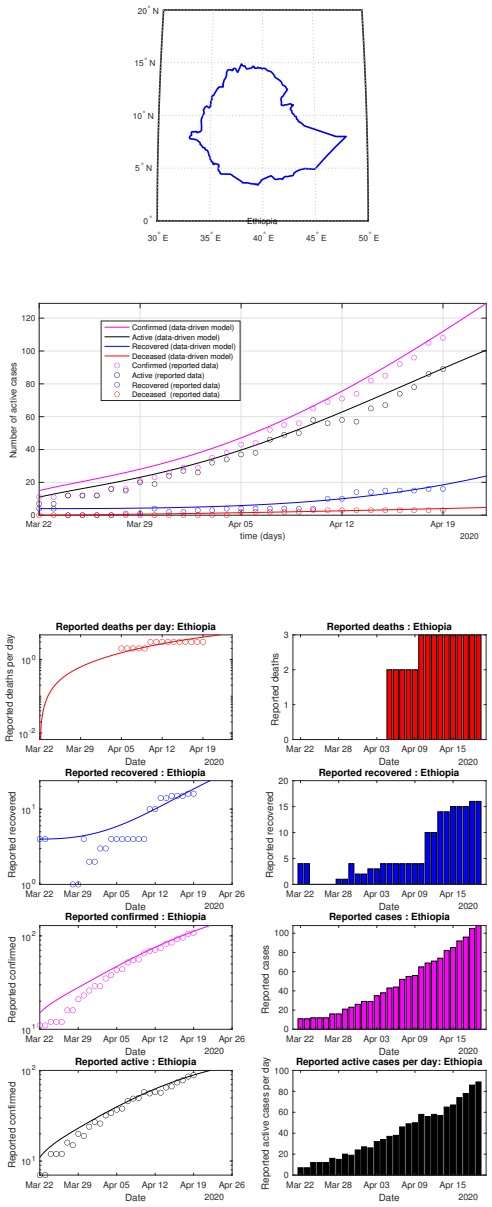

Figure 35: Ethiopia: data vs model

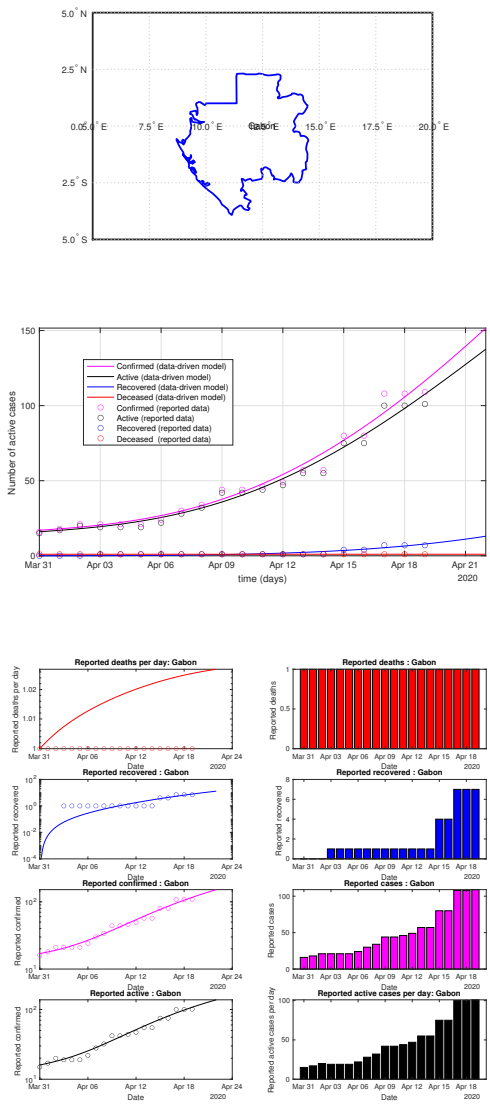

Figure 36: Gabon: data vs model

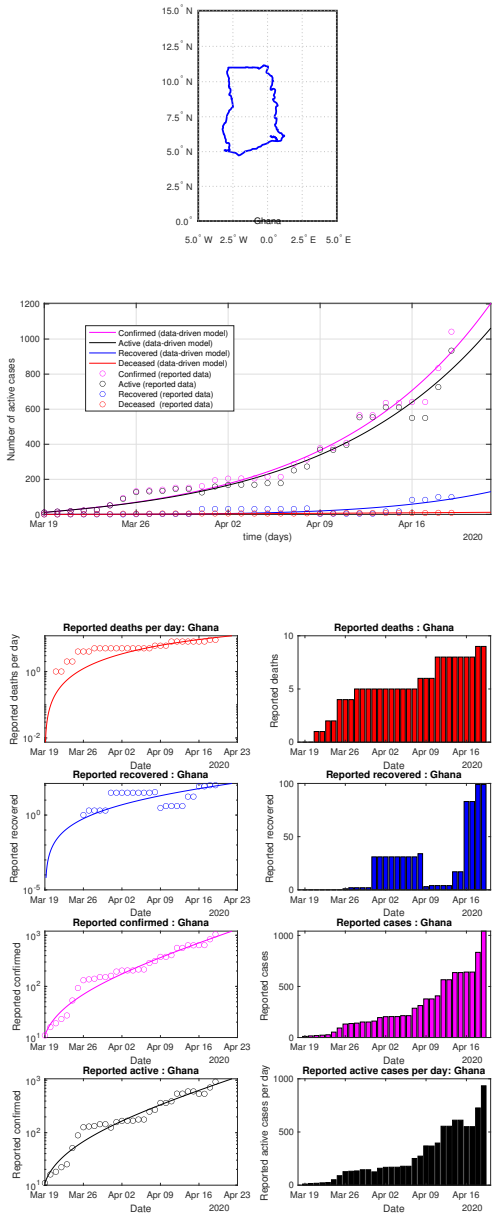

Figure 37: Ghana: data vs model

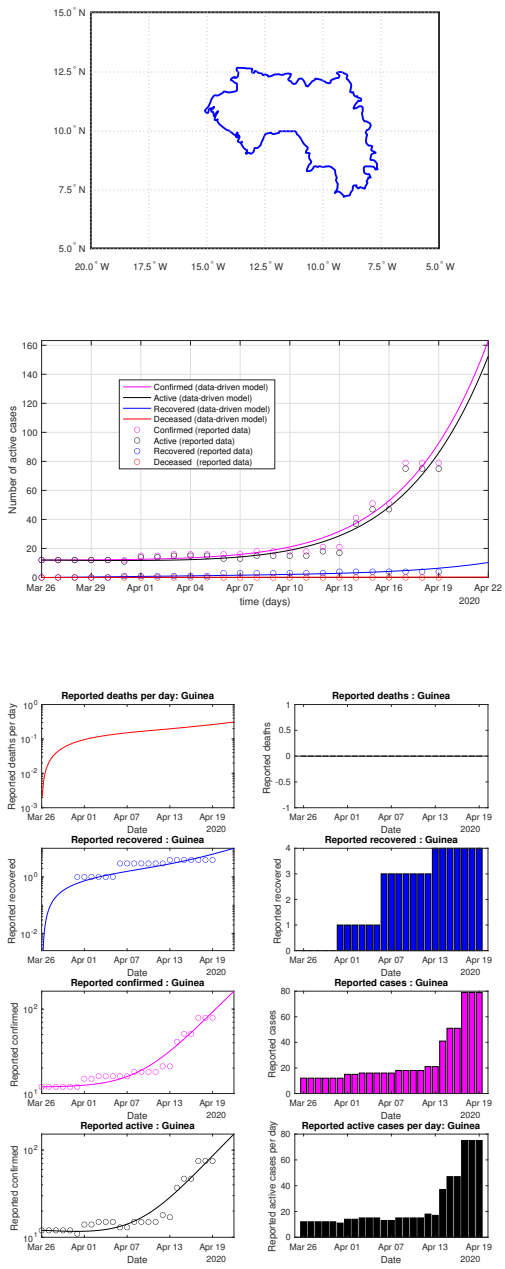

Figure 38: Guinea: data vs model

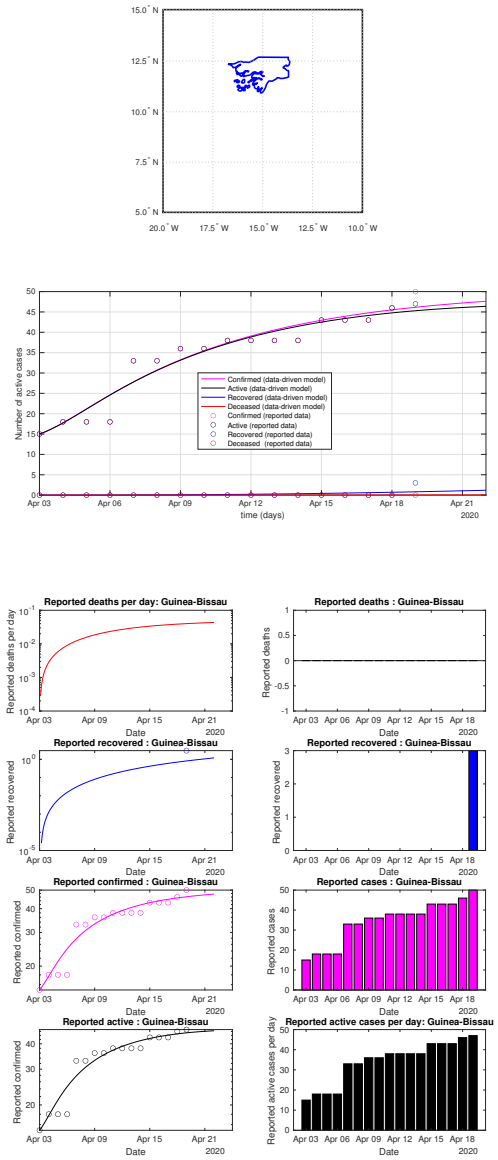

Figure 39: Guinea-Bissau: data vs model

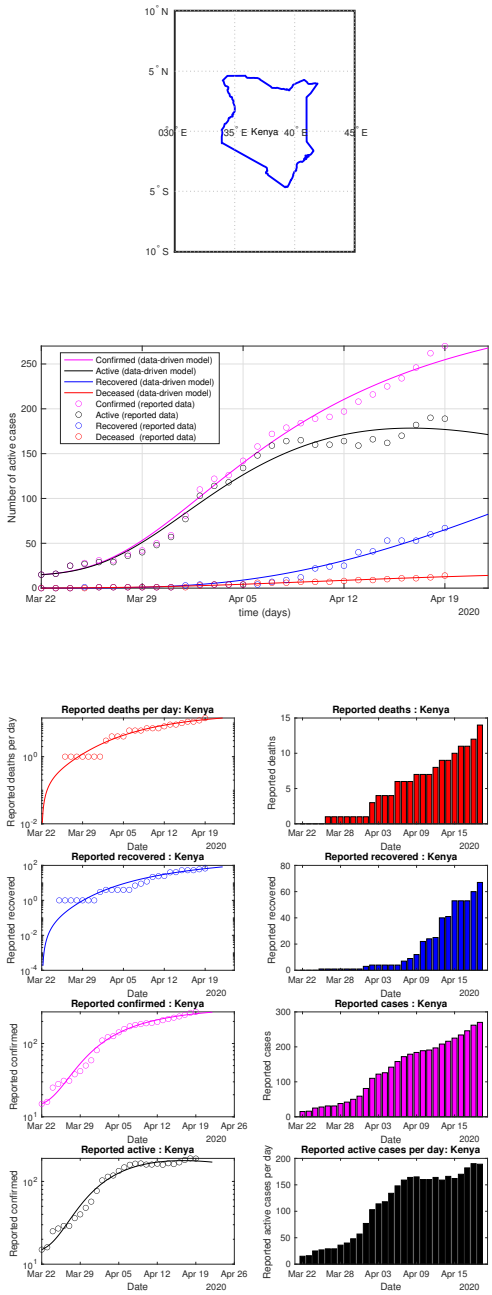

Figure 40: Kenya: data vs model

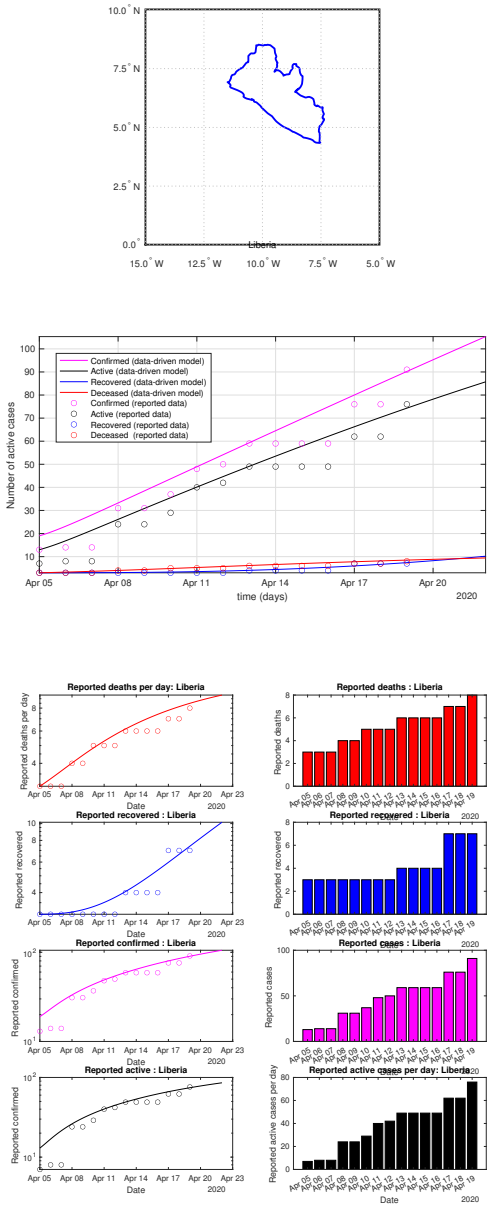

Figure 41: Liberia: data vs model

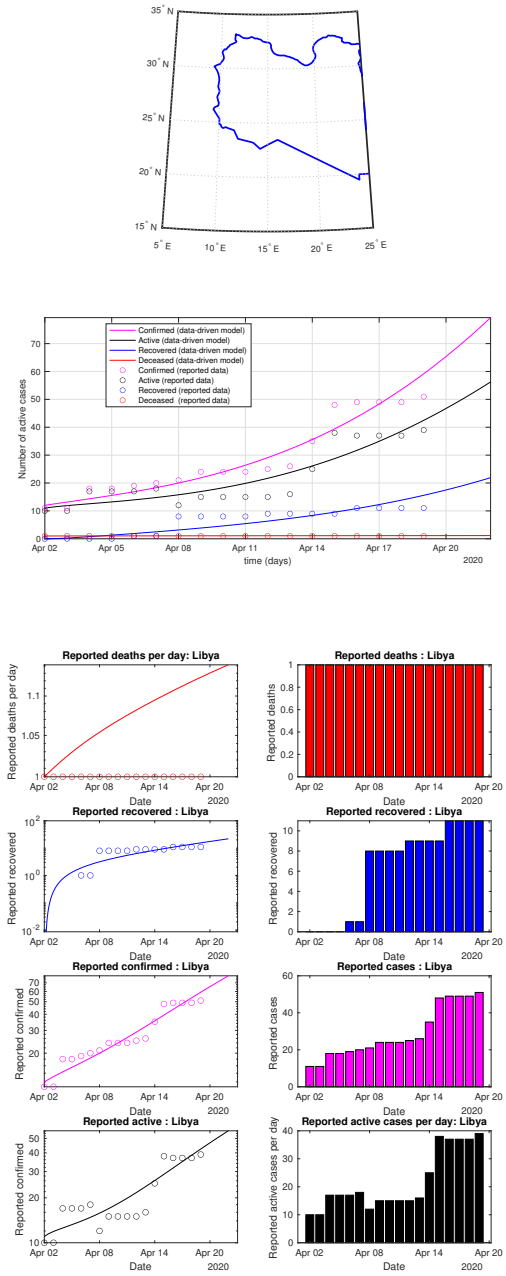

Figure 42: Libya: data vs model

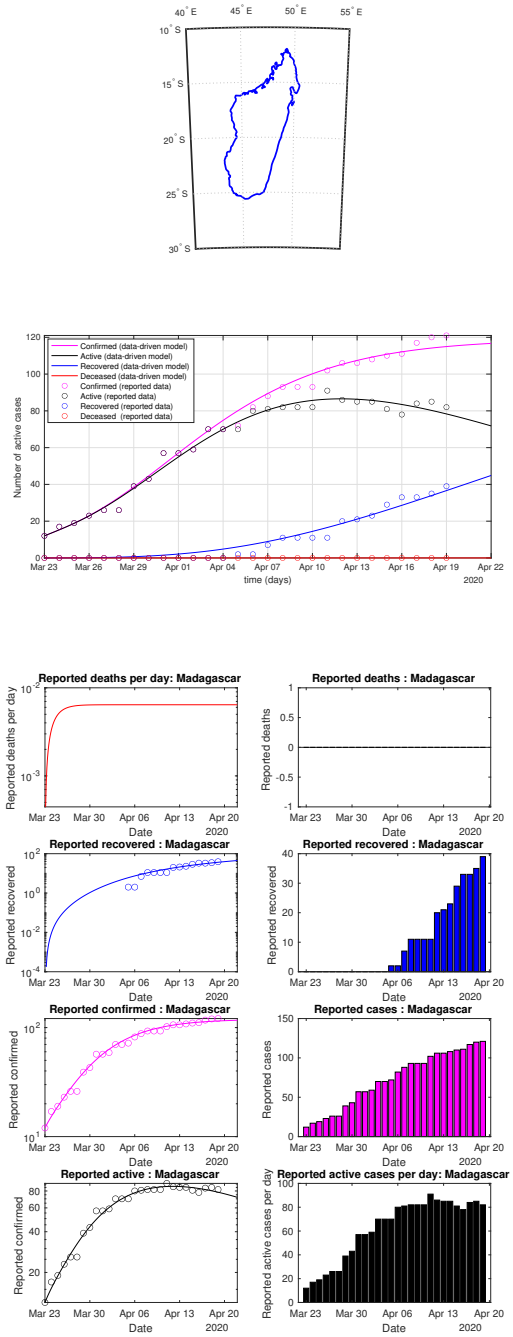

Figure 43: Madagascar: data vs model

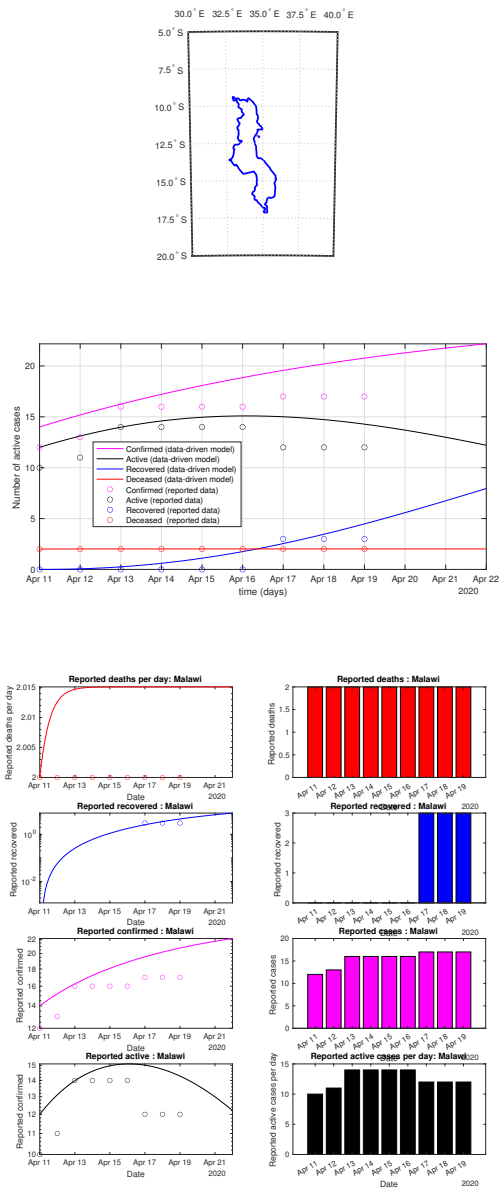

Figure 44: Malawi: data vs model

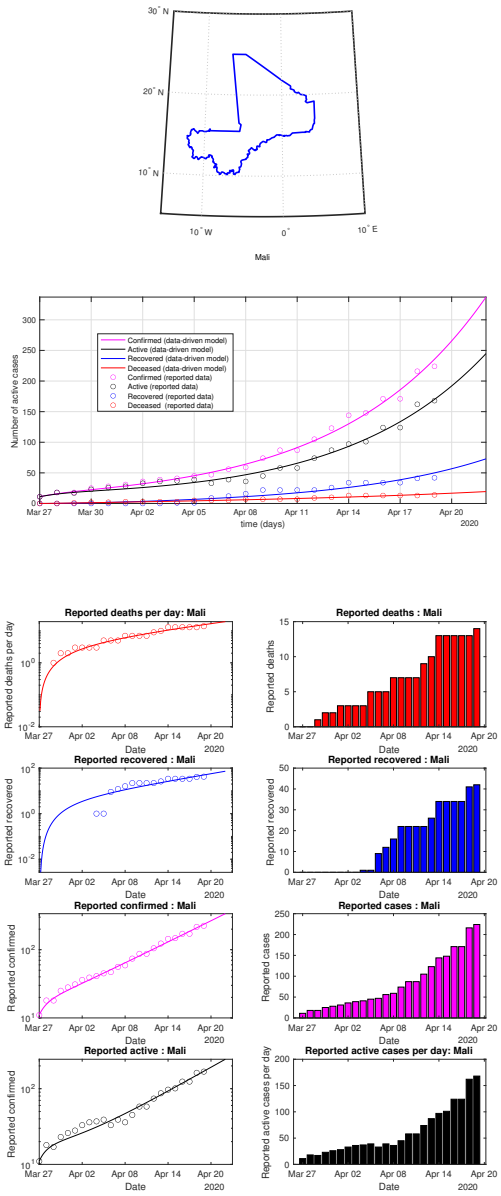

Figure 45: Mali: data vs model

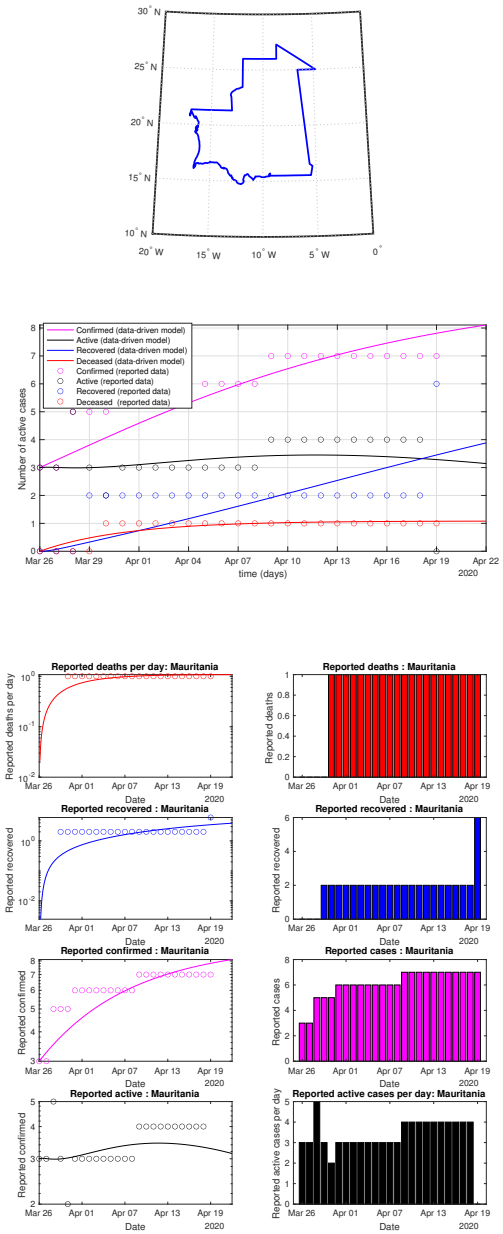

Figure 46: Mauritania: data vs model

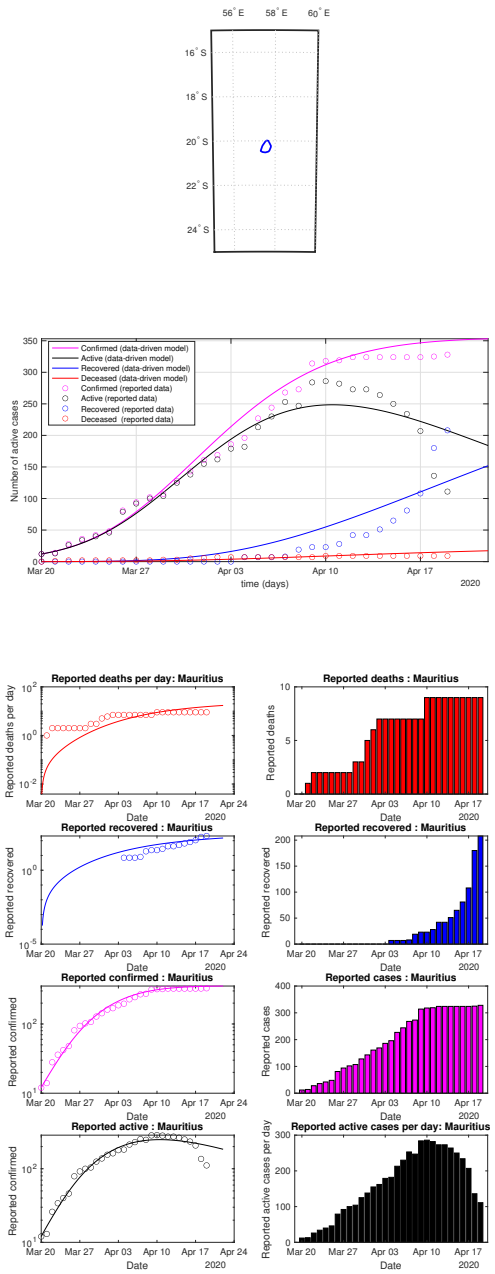

Figure 47: Mauritius: data vs model

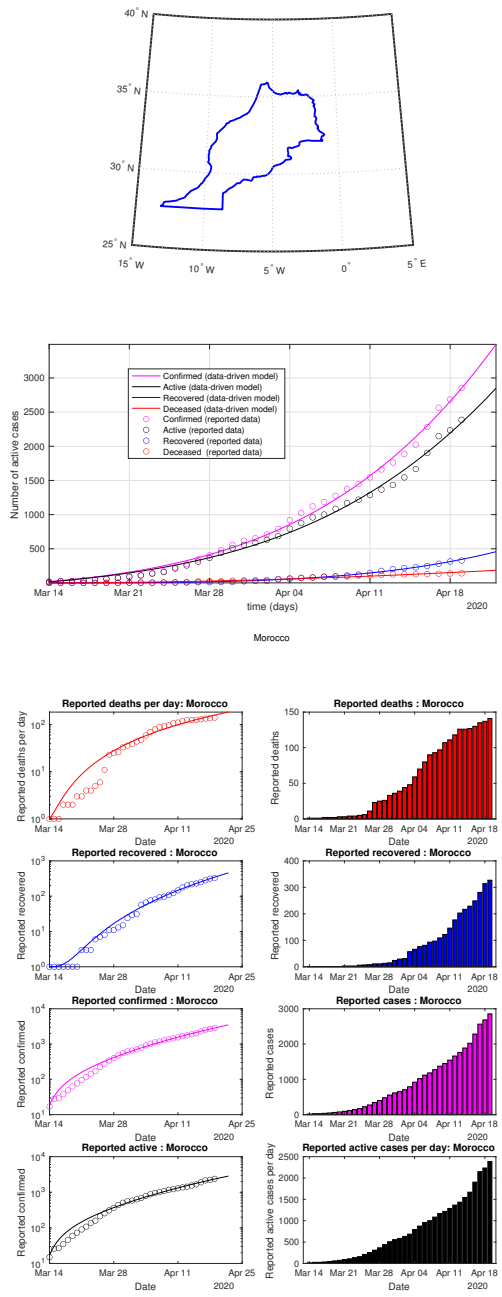

Figure 48: Morocco: data vs model

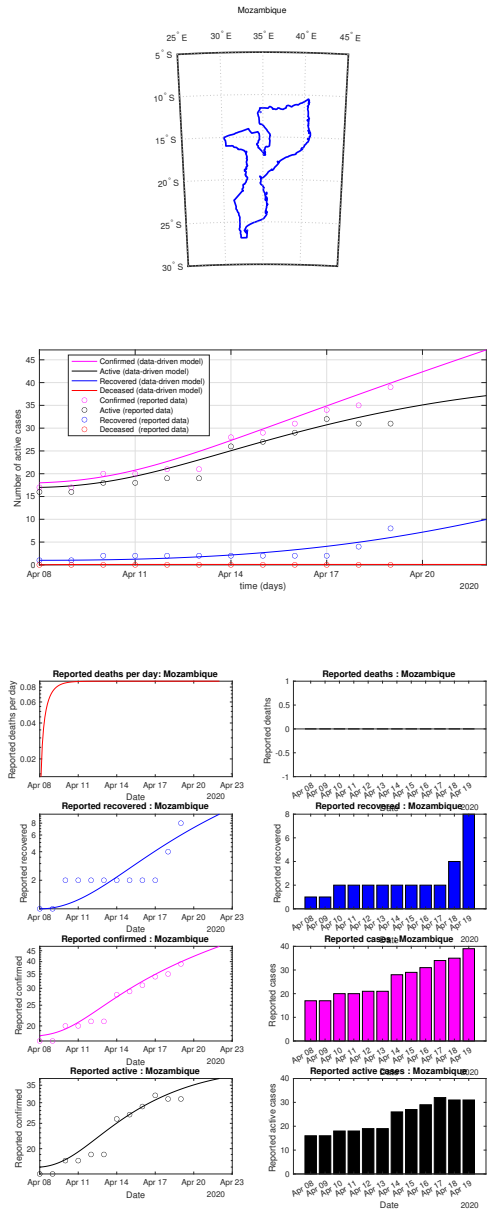

Figure 49: Mozambique: data vs model

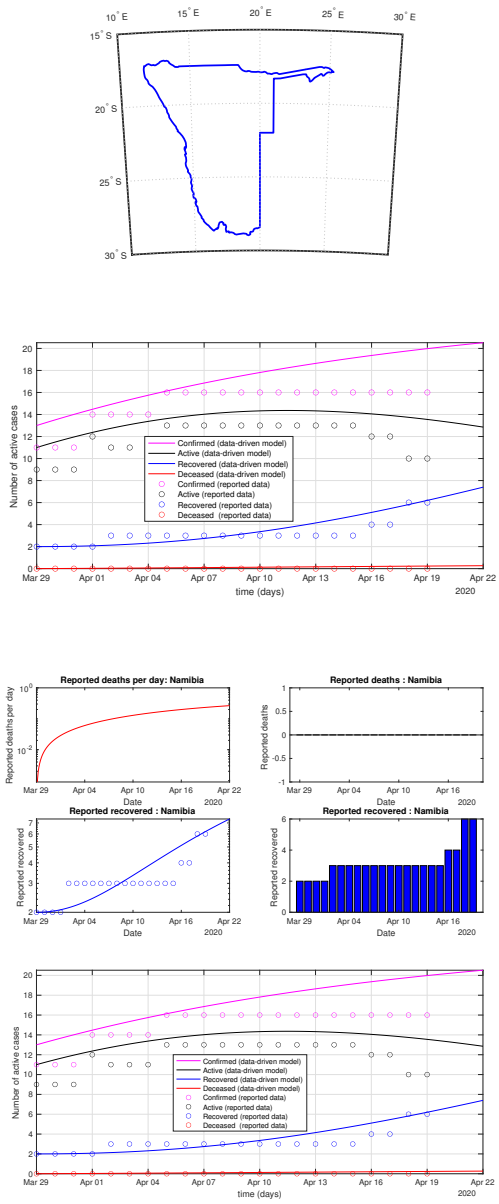

Figure 50: Namibia: data vs model

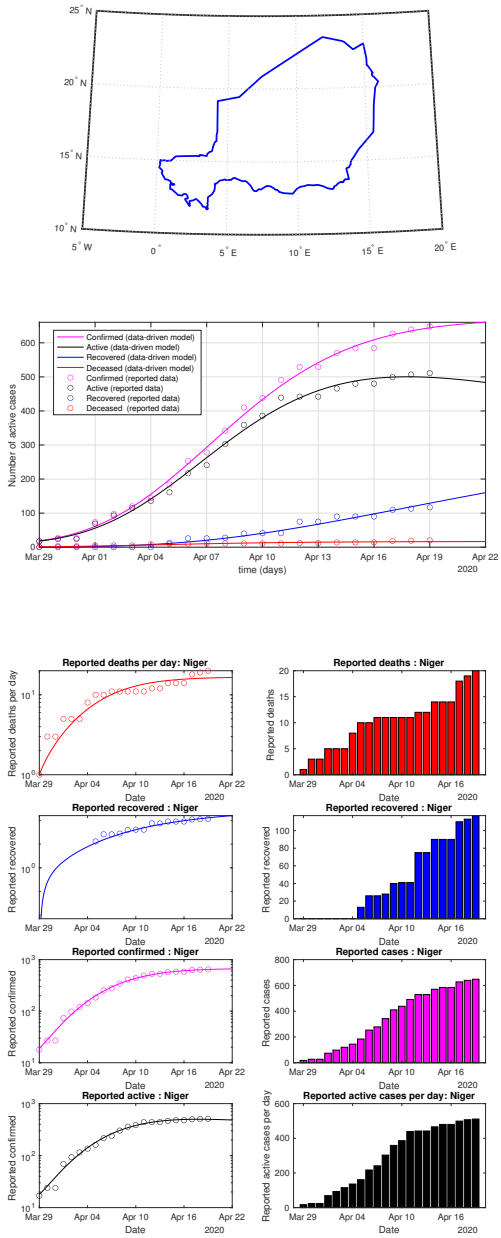

Figure 51: Niger: data vs model

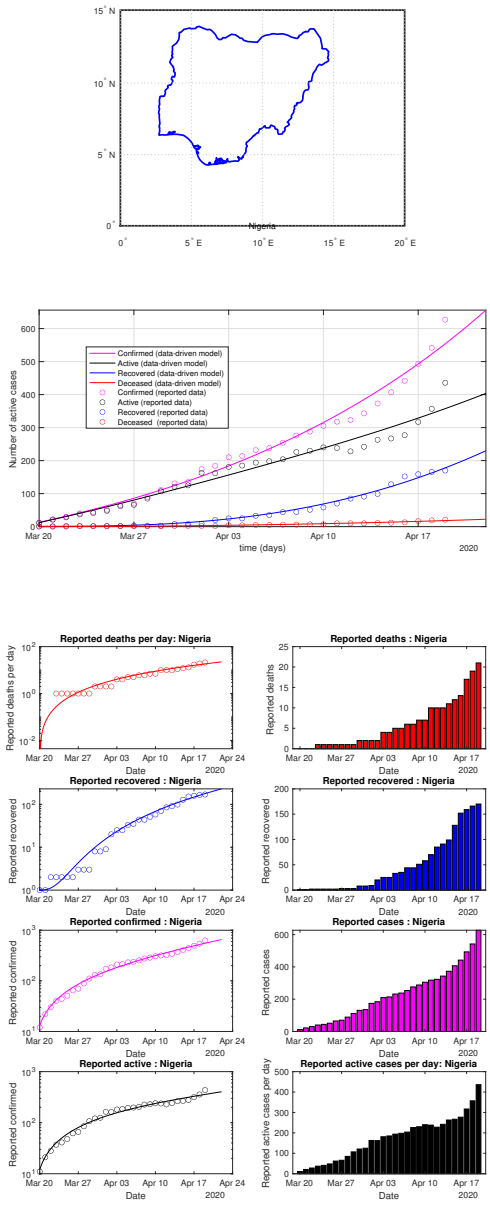

Figure 52: Nigeria: data vs model

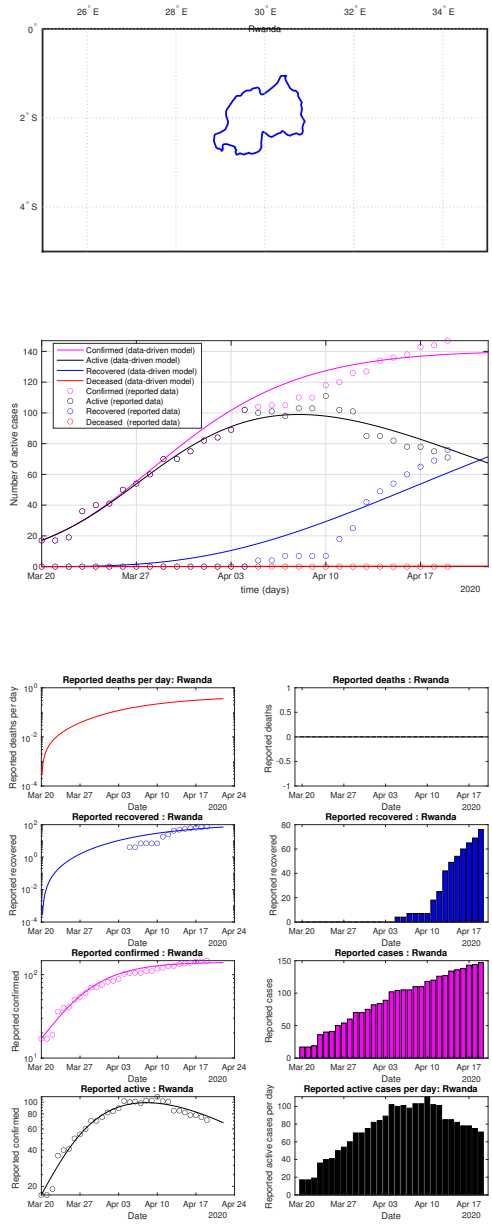

Figure 53: Rwanda: data vs model

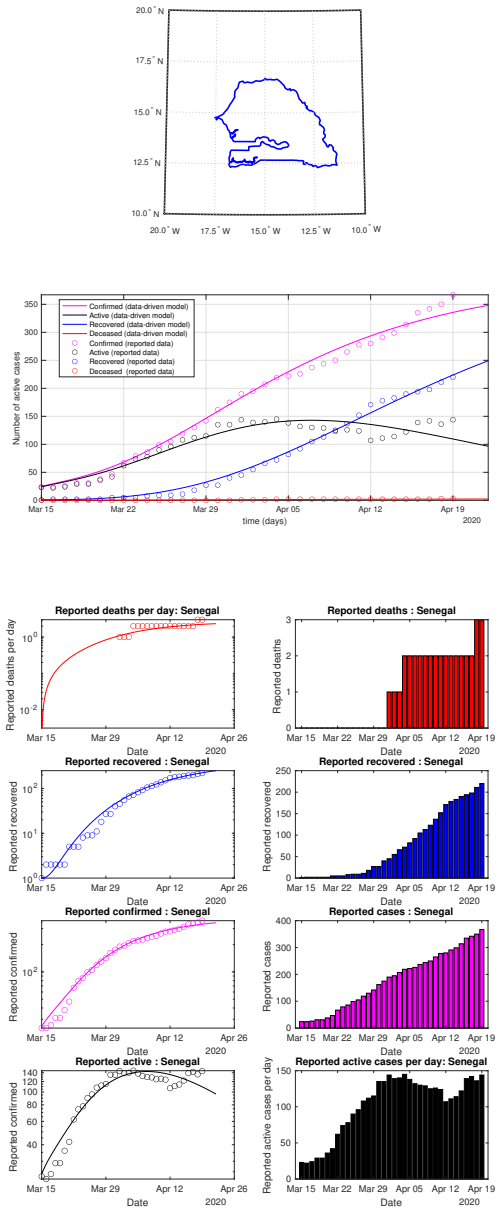

Figure 54: Senegal: data vs model

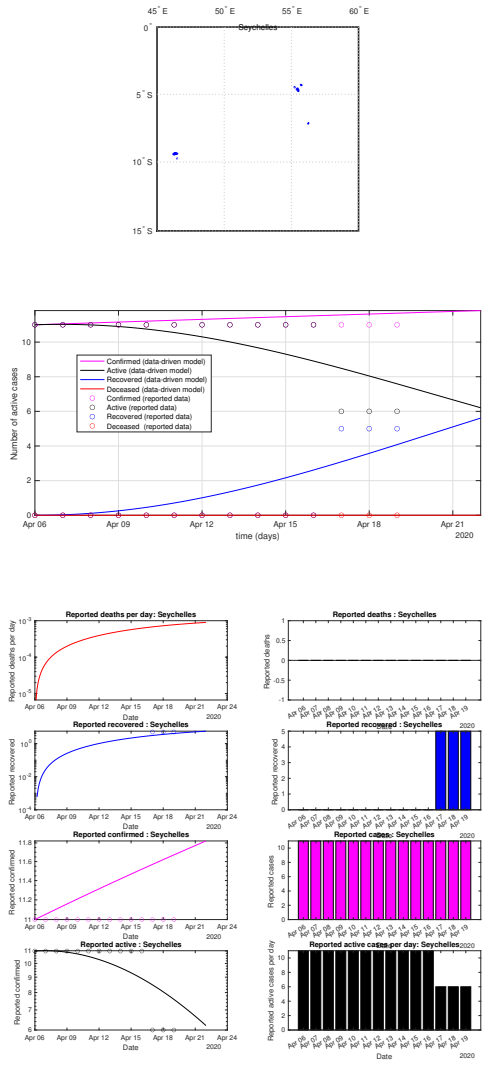

Figure 55: Seychelles: data vs model

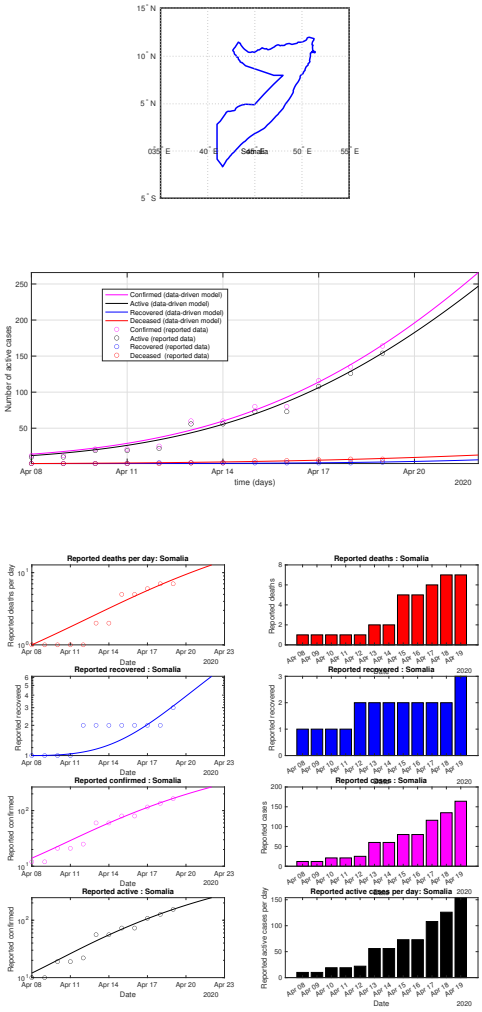

Figure 56: Somalia: data vs model

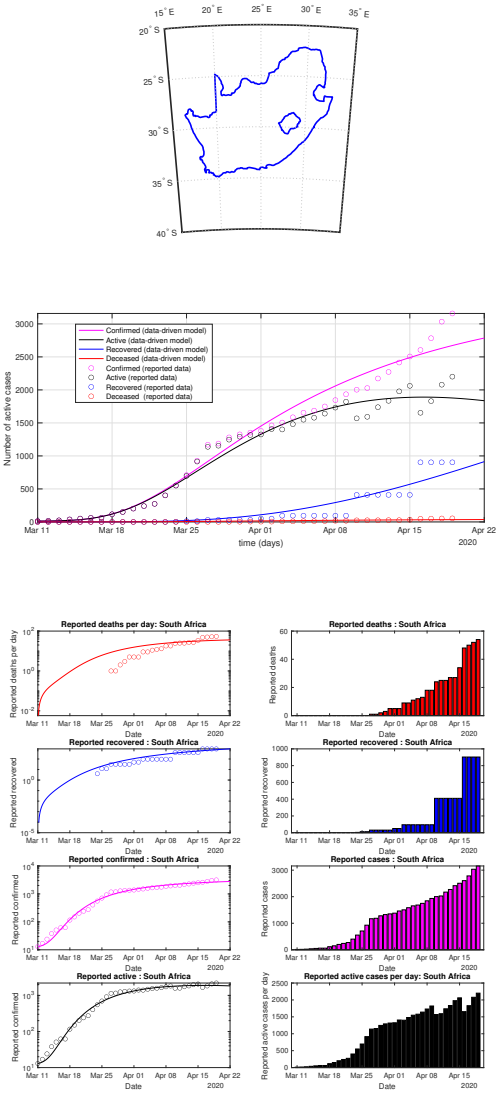

Figure 57: SouthAfrica: data vs model

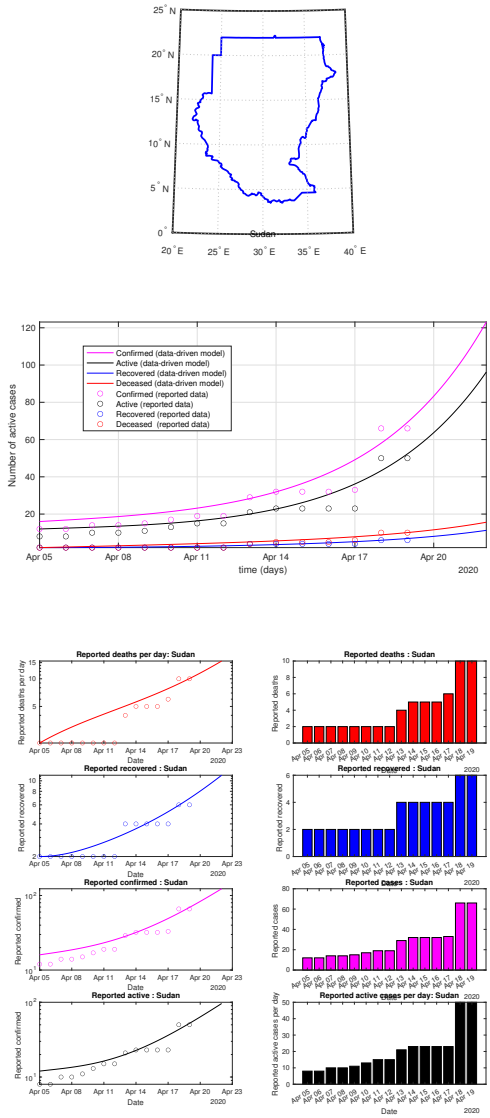

Figure 58: Sudan: data vs model

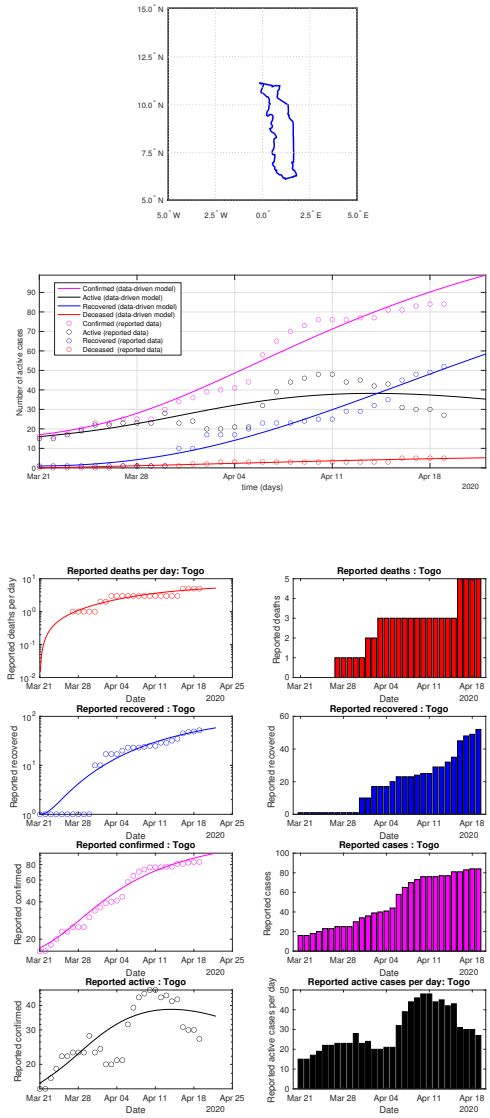

Figure 59: Togo: data vs model

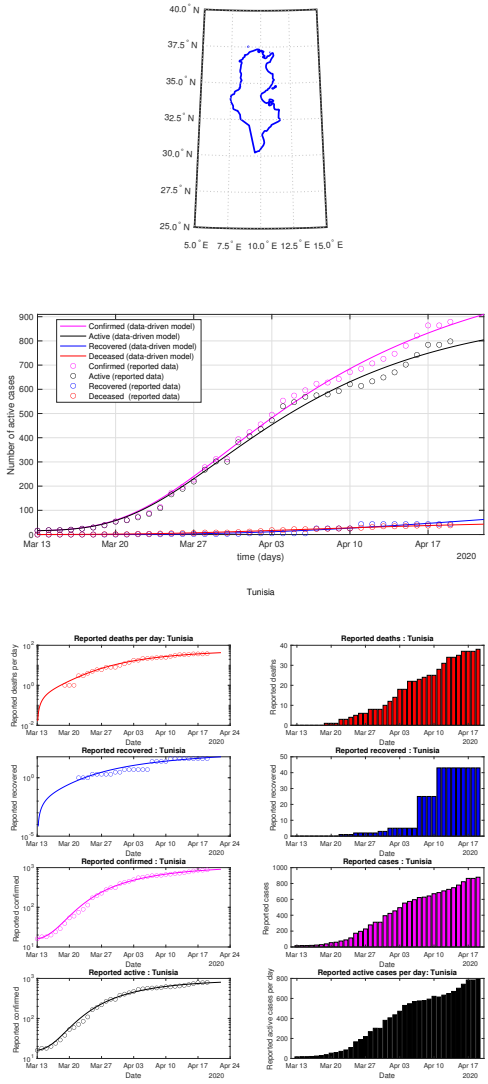

Figure 60: Tunisia: data vs model

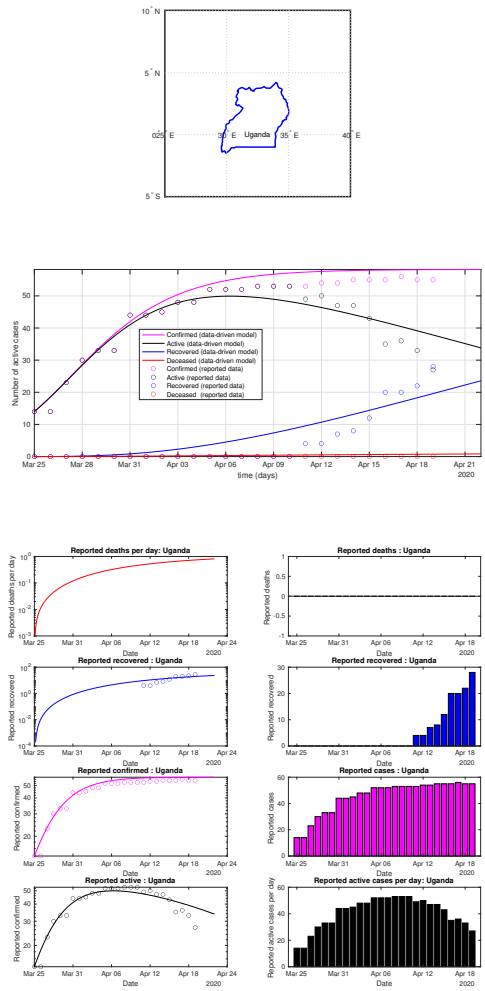

Figure 61: Uganda: data vs model

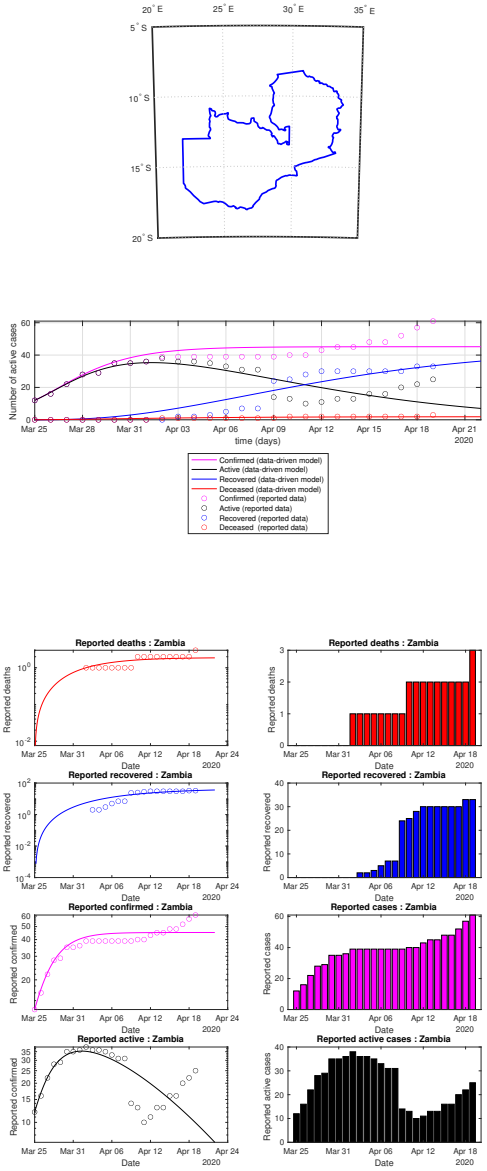

Figure 62: Zambia: data vs model

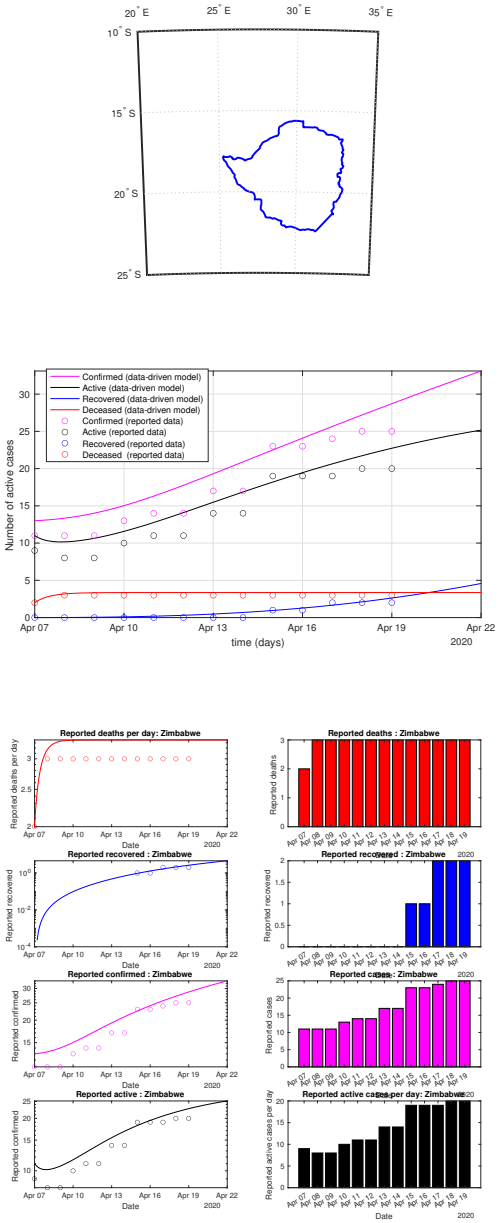

Figure 63: Zimbabwe: data vs model

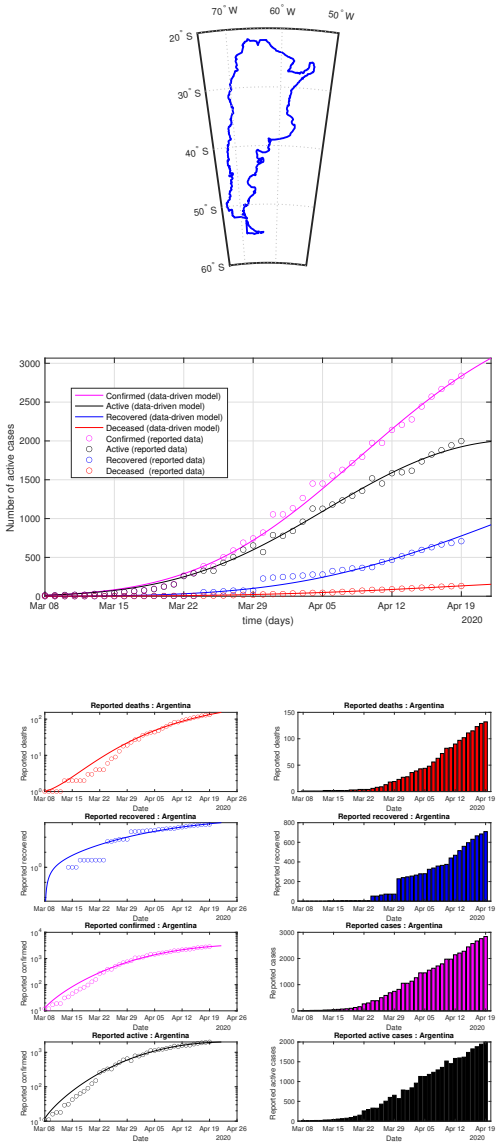

Figure 64: Argentina: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

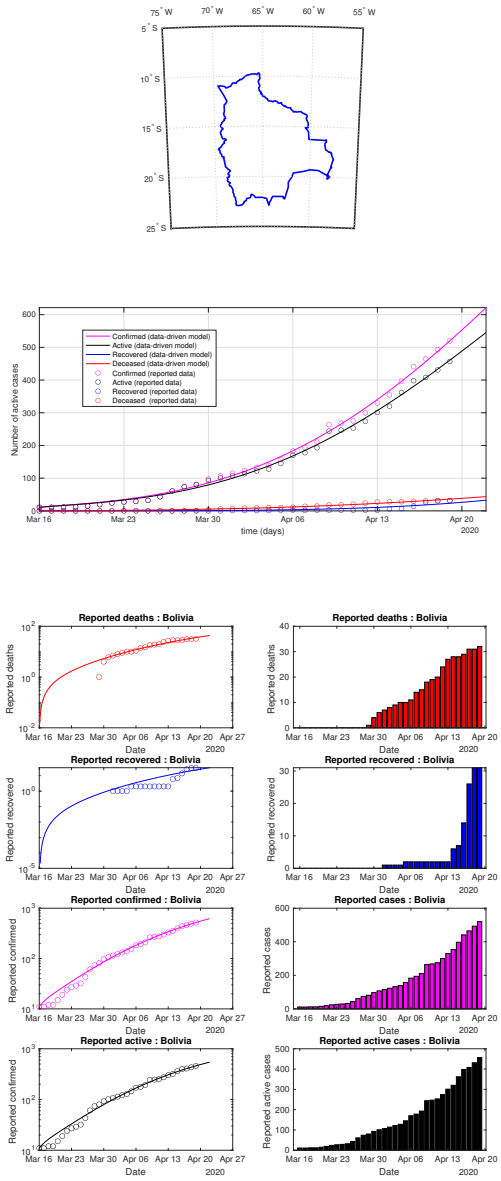

Figure 65: Bolivia: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

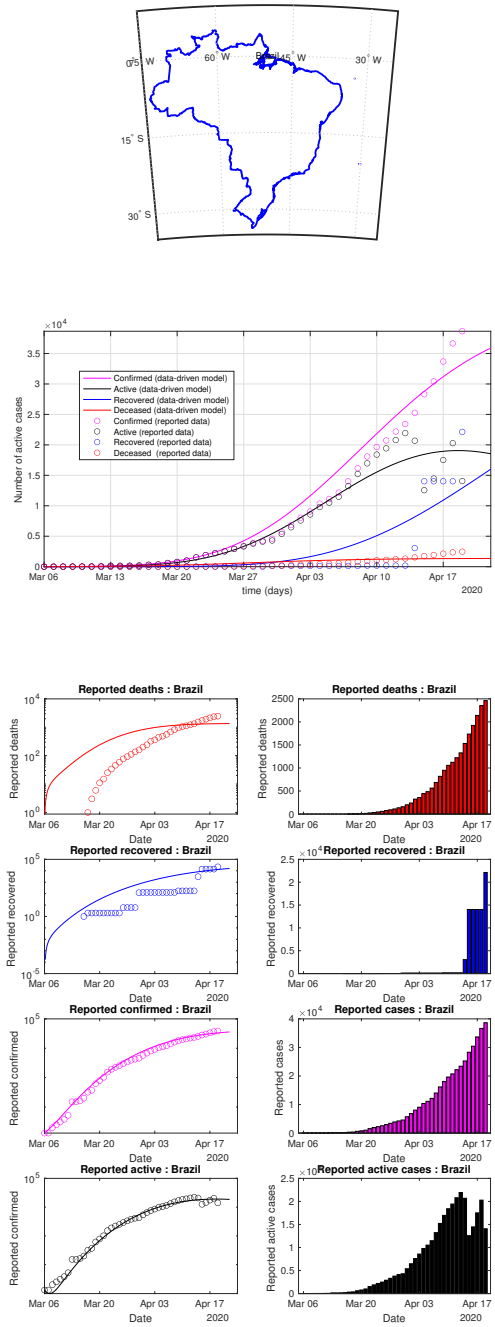

Figure 66: Brazil: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

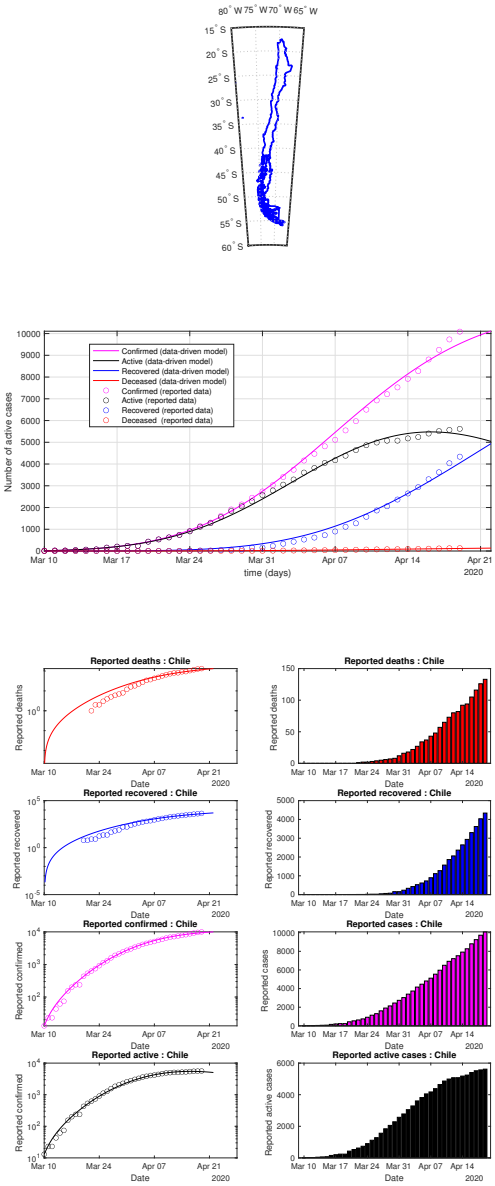

Figure 67: Chile: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

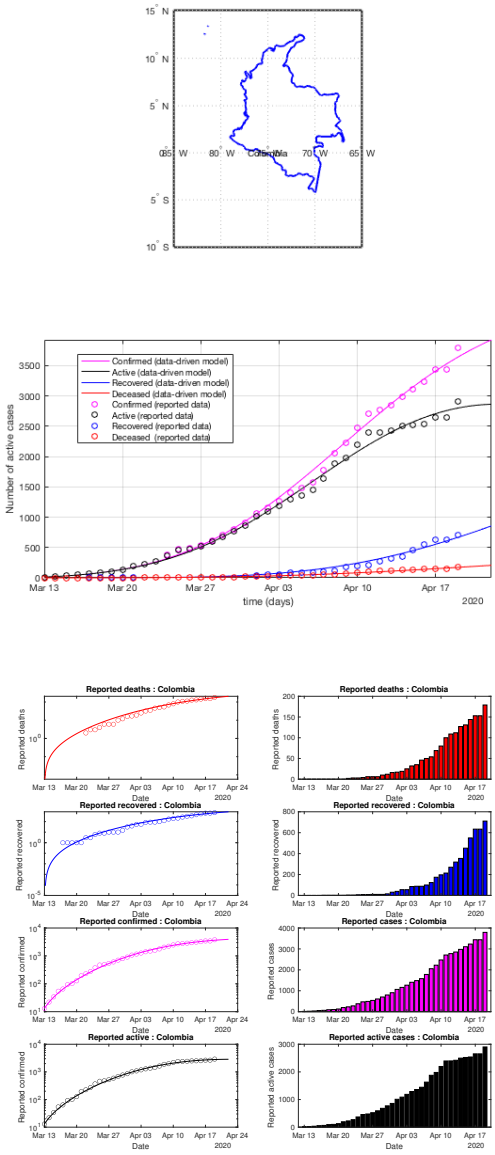

Figure 68: Colombia: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

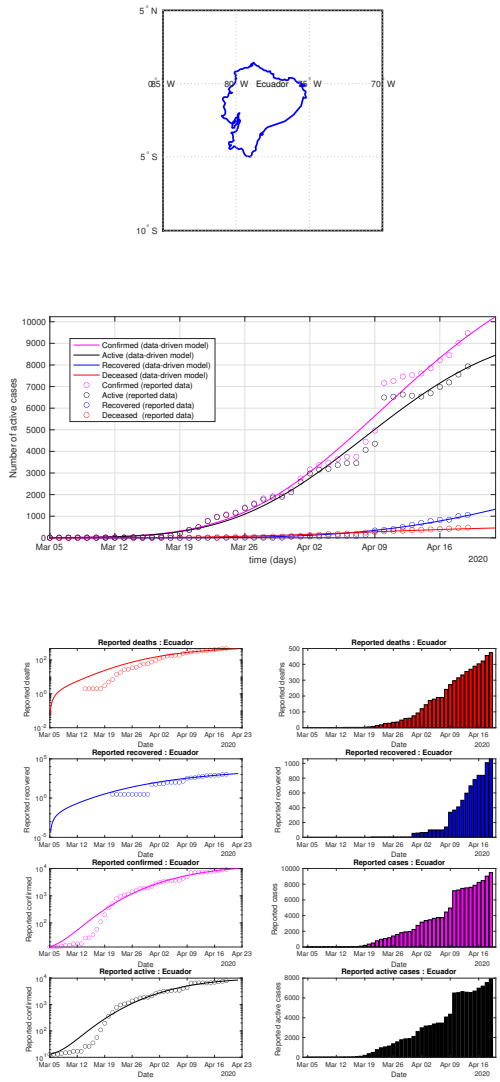

Figure 69: Ecuador: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

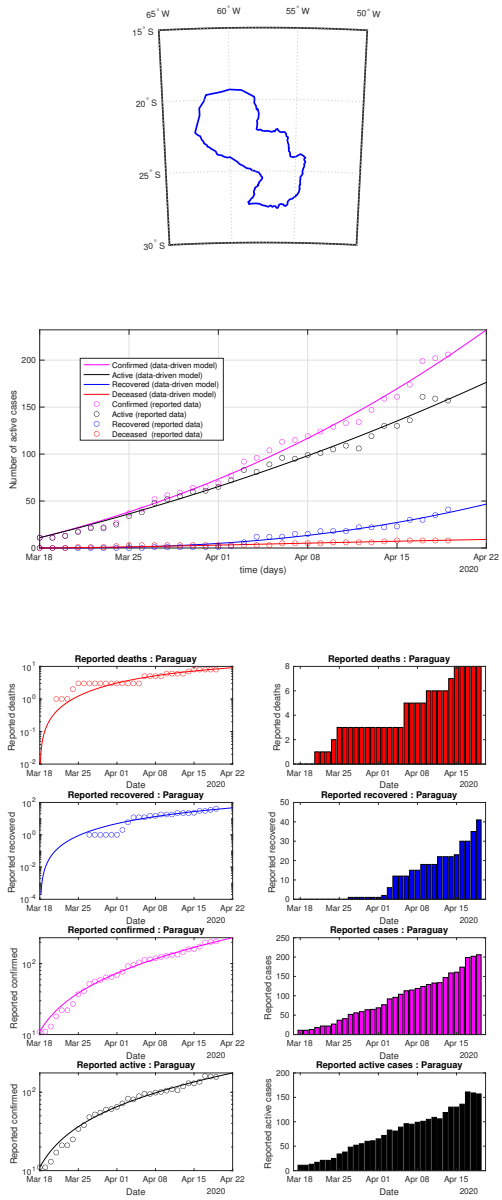

Figure 70: Paraguay: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

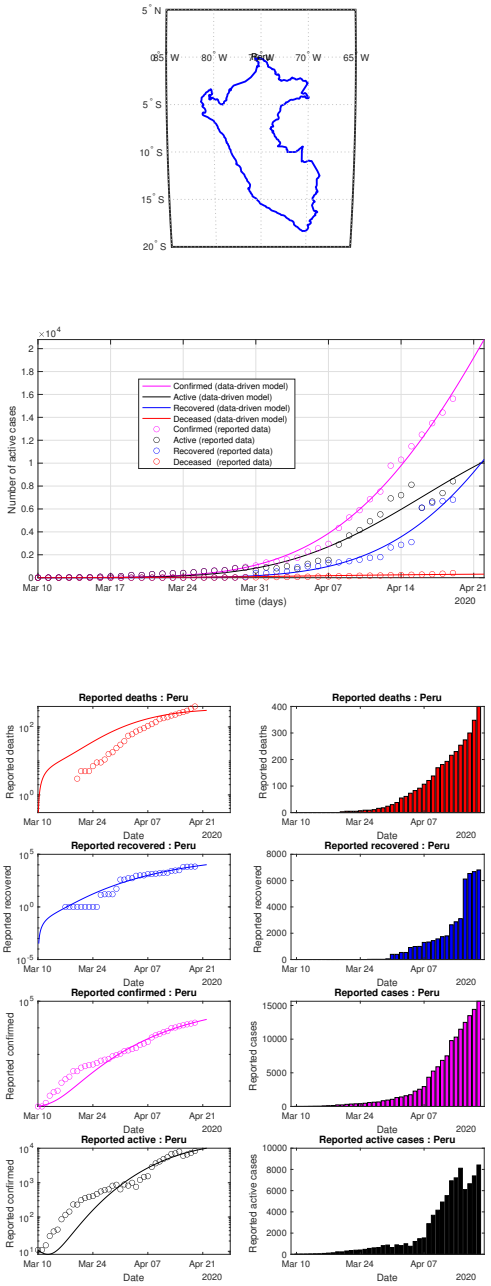

Figure 71: Peru: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

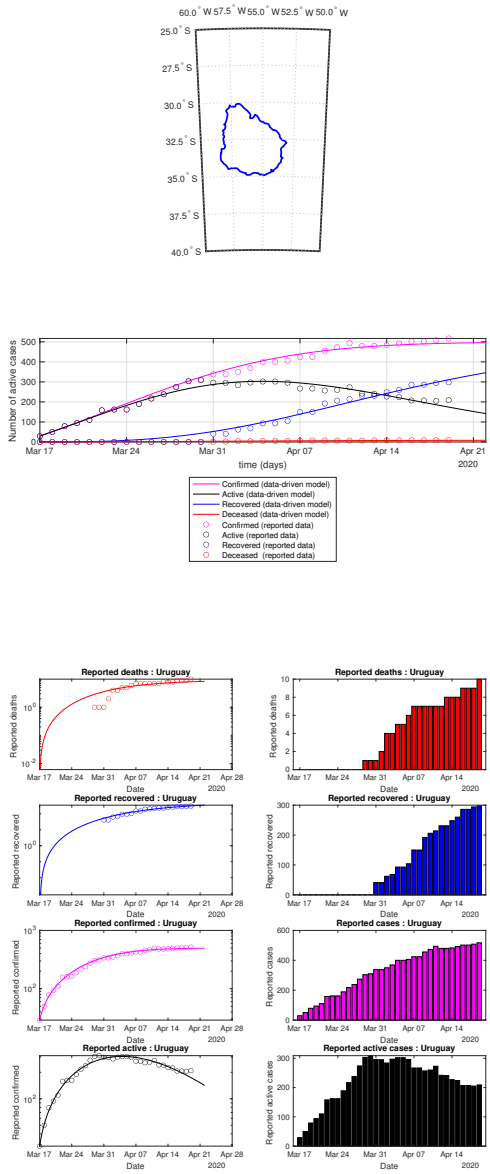

Figure 72: Uruguay: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

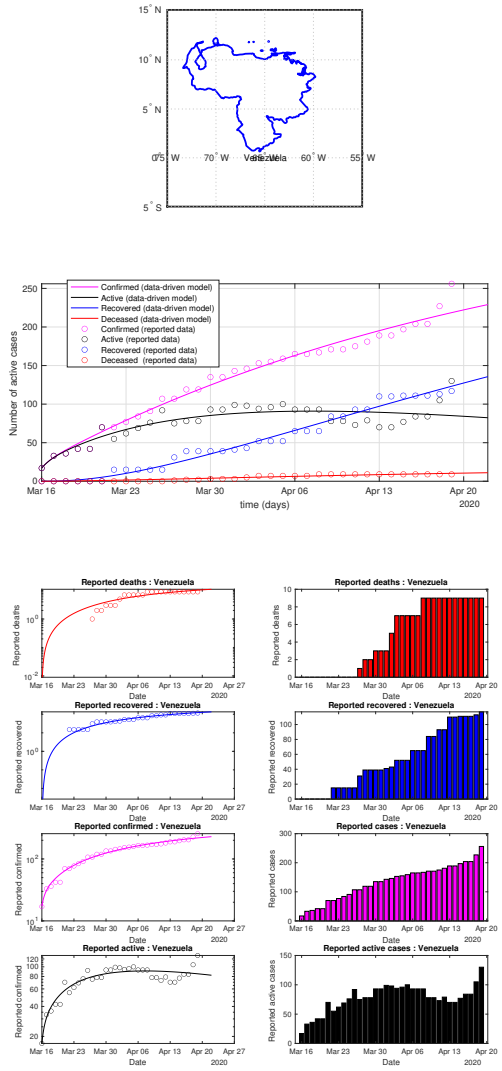

Figure 73: Venezuela: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

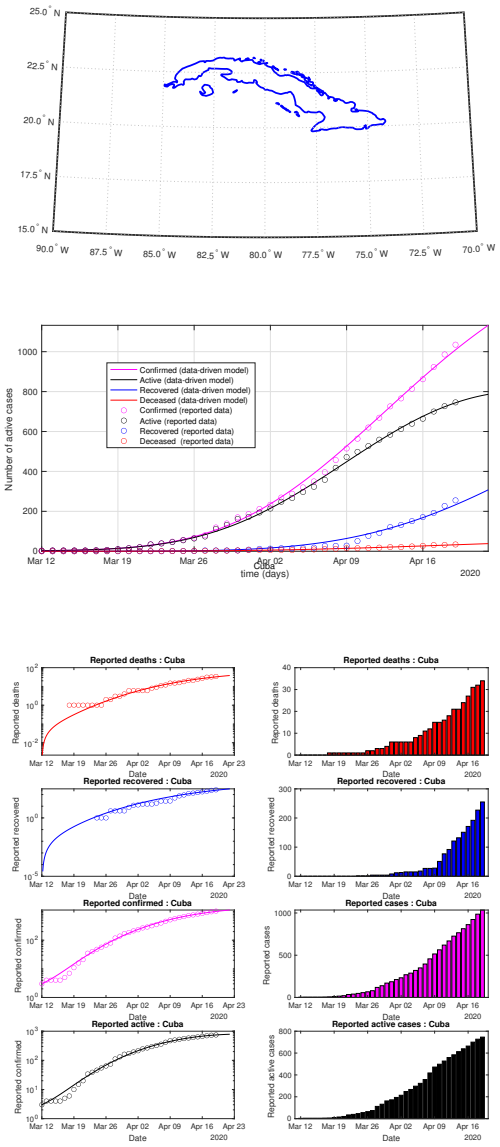

Figure 74: Cuba: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

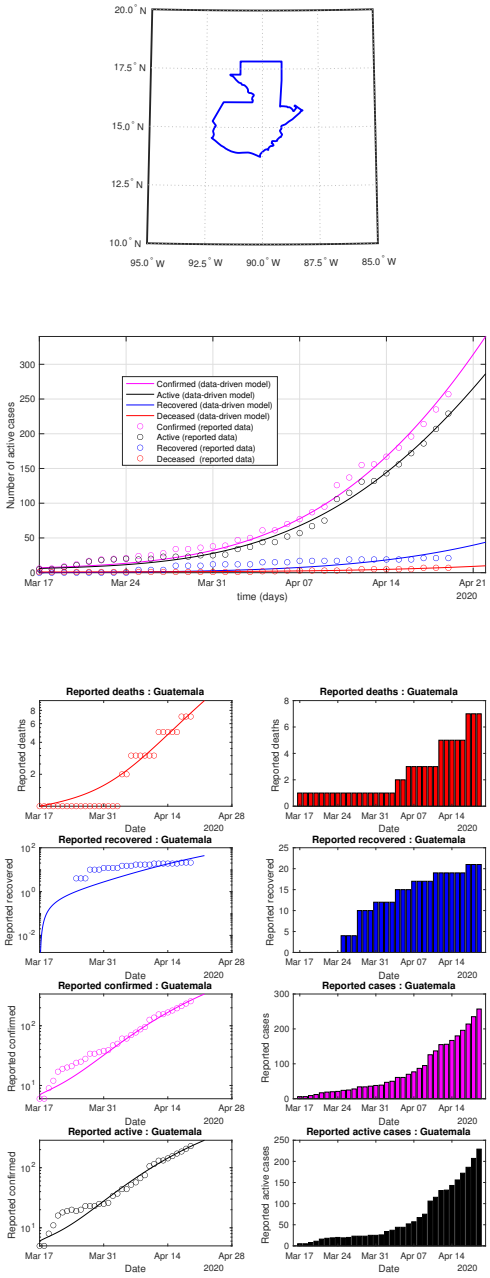

Figure 75: Guatemala: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

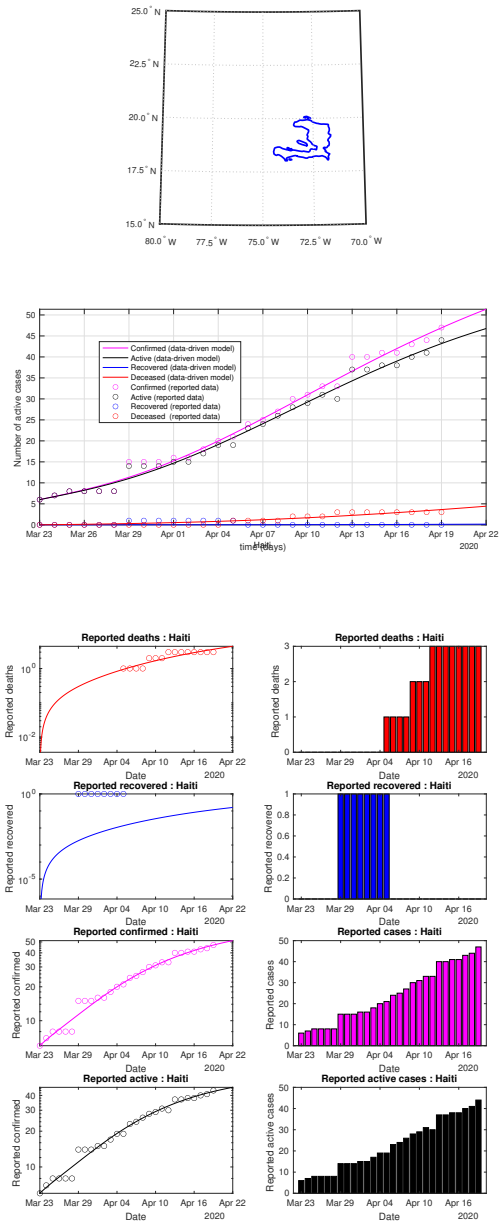

Figure 76: Haiti: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

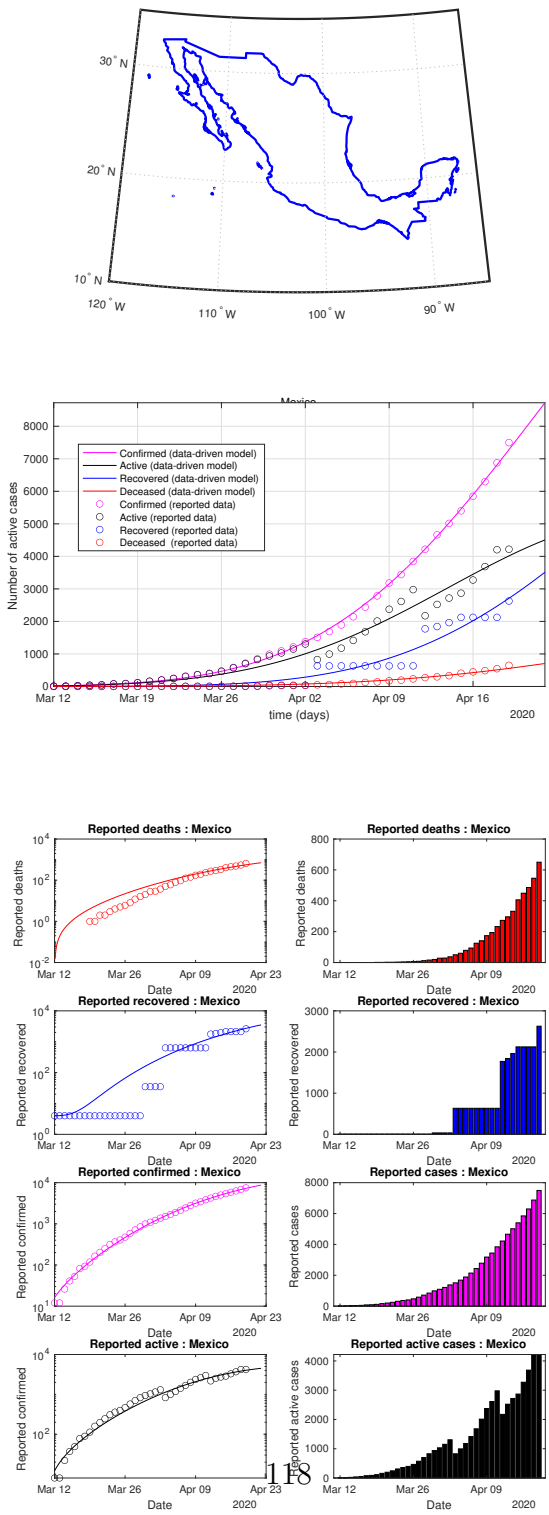

Figure 77: Mexico: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases over time. Predictive analytics for the next couple days

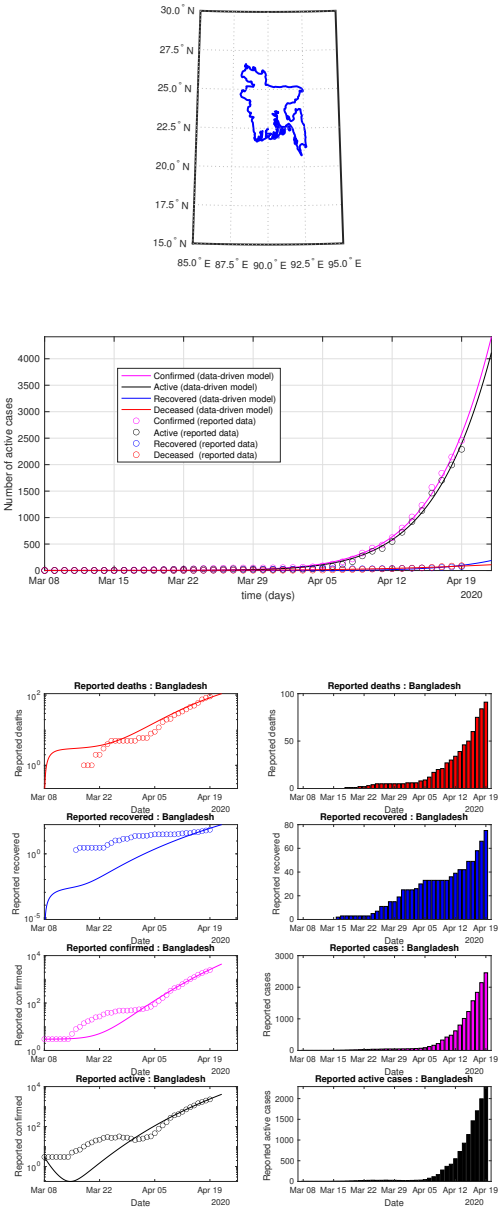

Figure 78: Bangladesh: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

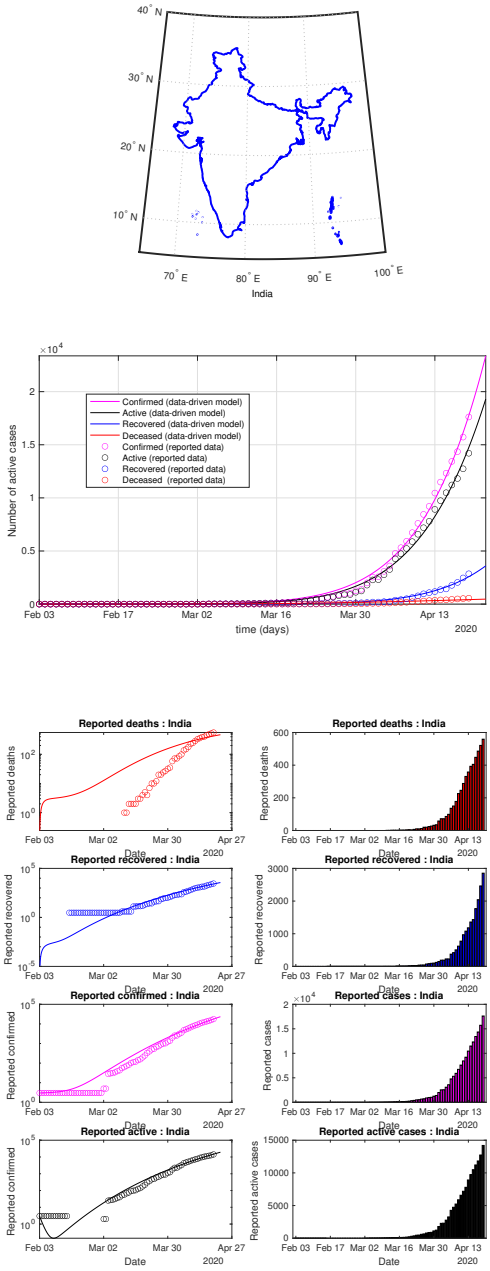

Figure 79: India: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

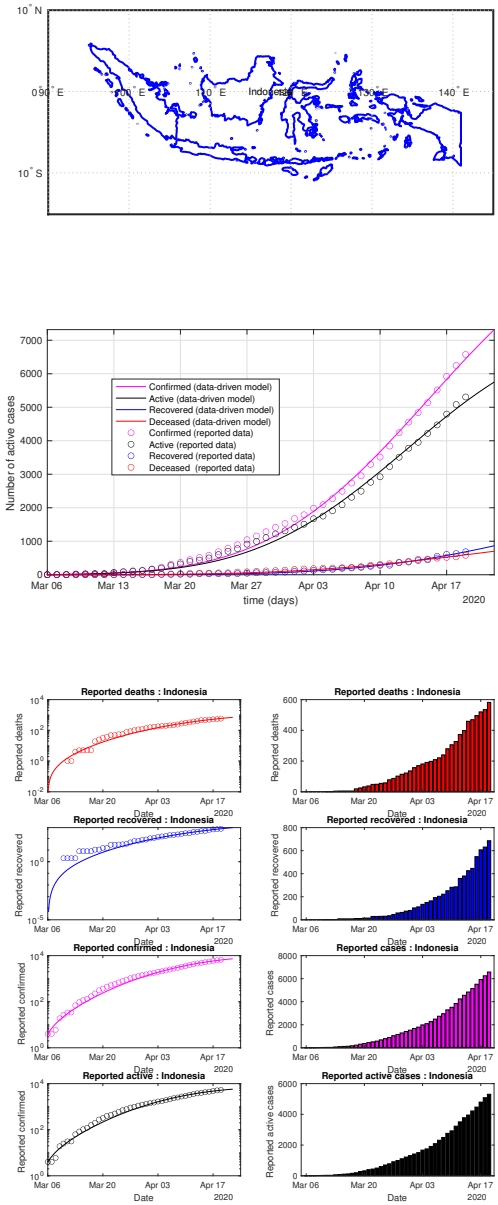

Figure 80: Indonesia: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

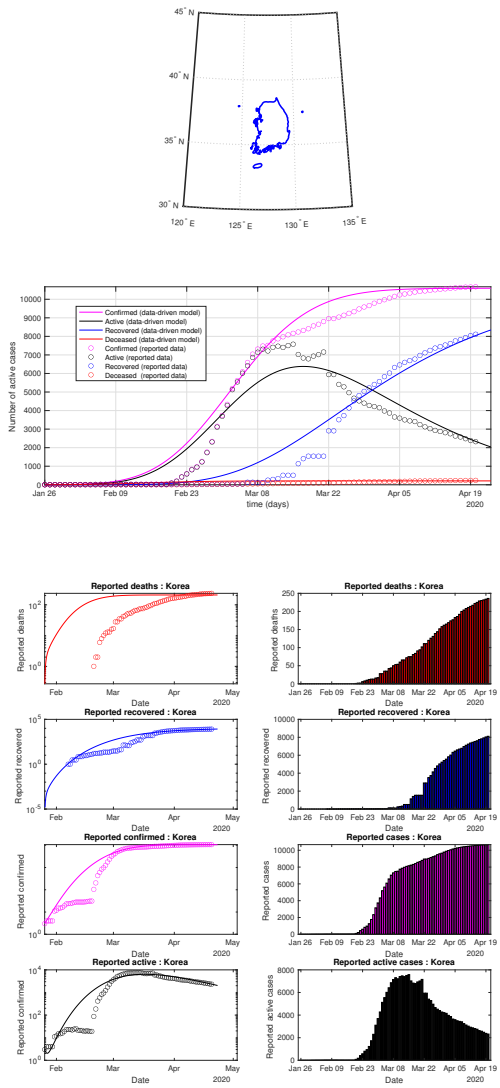

Figure 81: South Korea: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

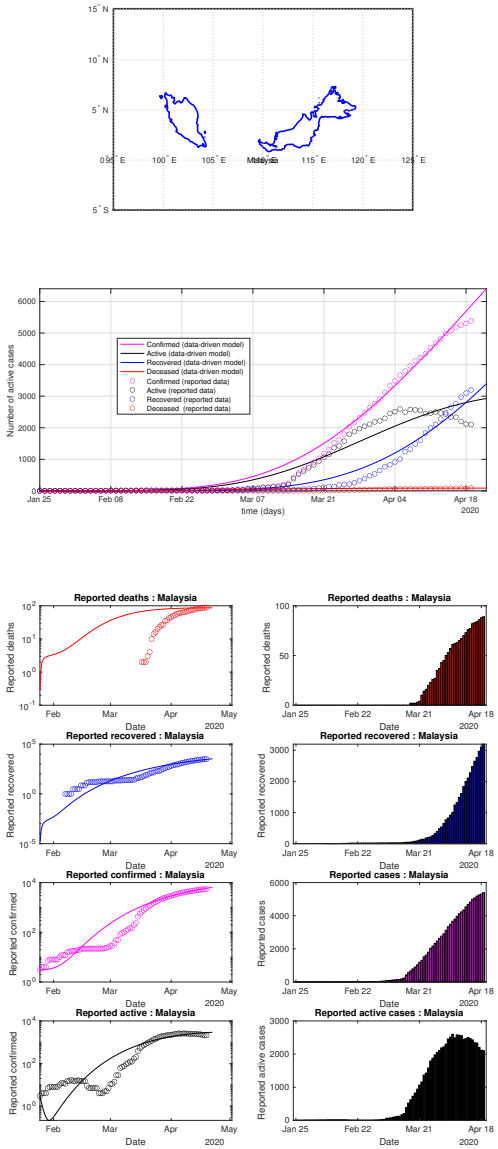

Figure 82: Malaysia: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

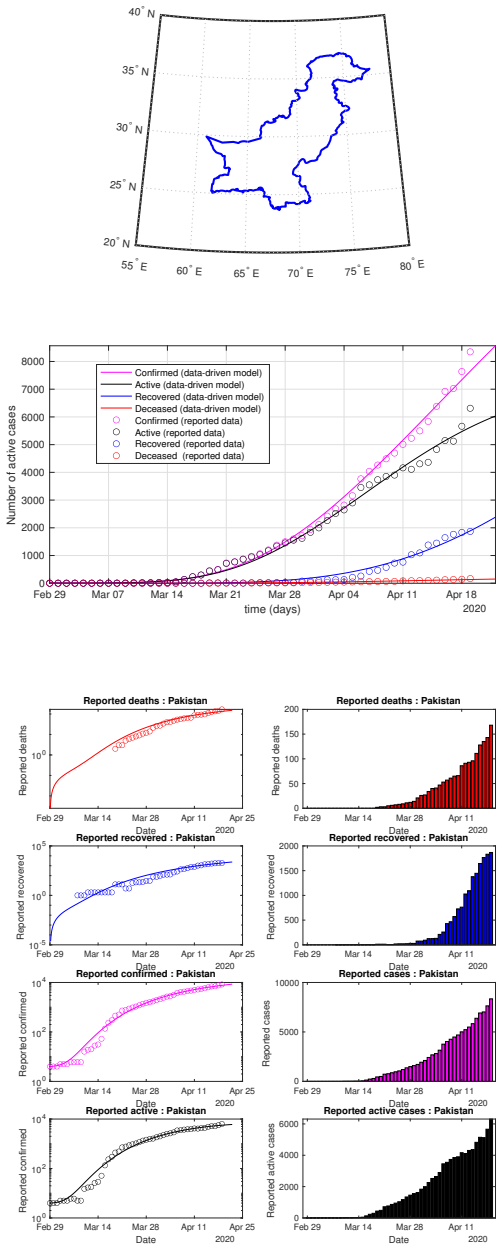

Figure 83: Pakistan: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

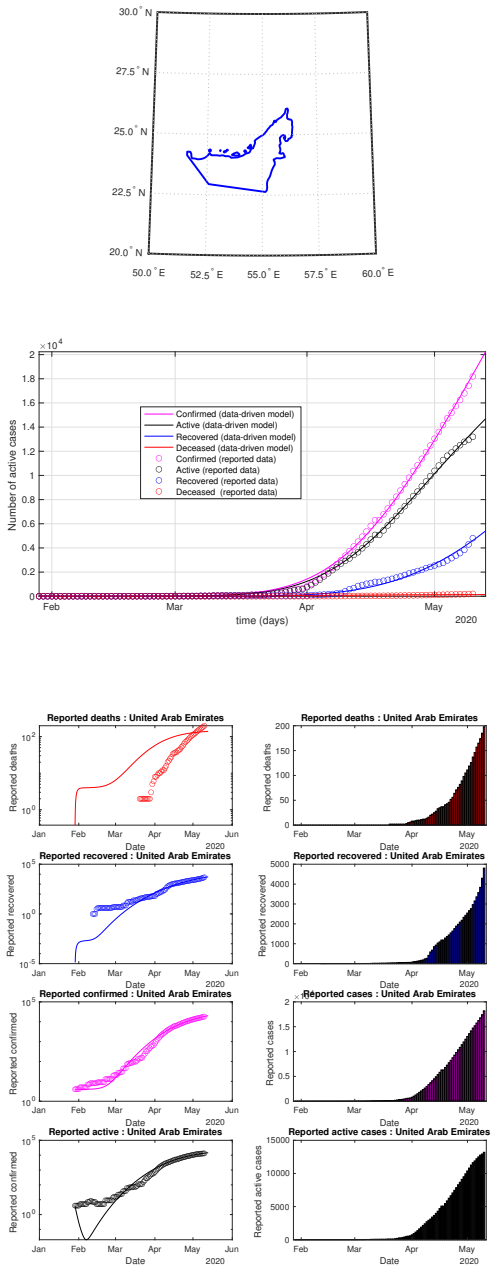

Figure 84: UAE: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

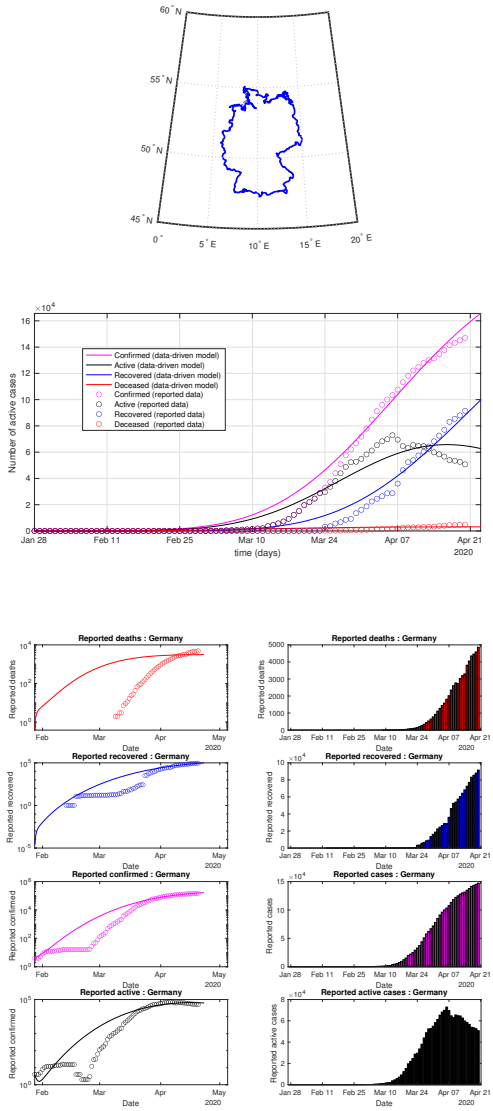

Figure 85: Germany: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

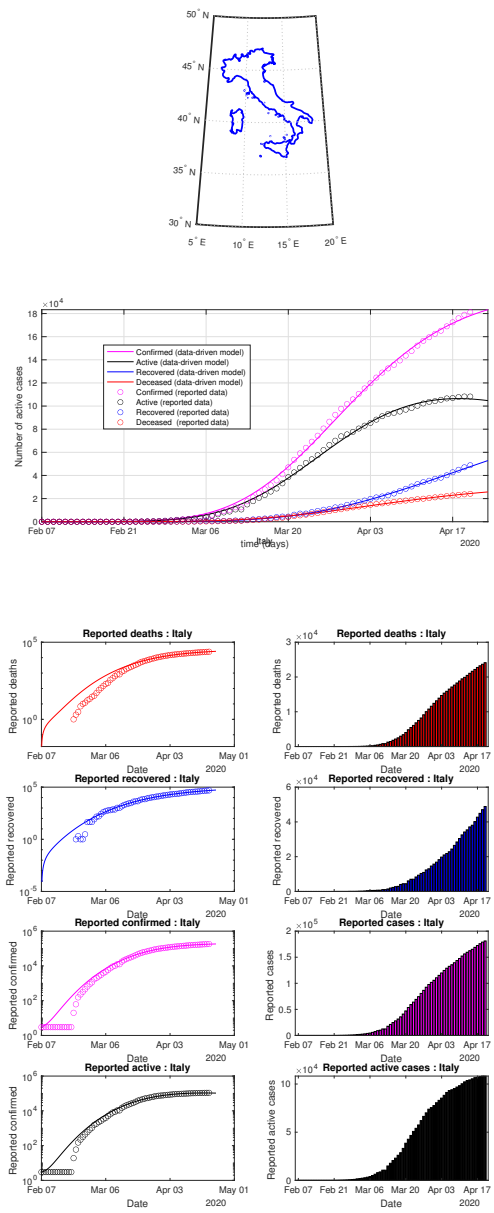

Figure 86: Italy: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

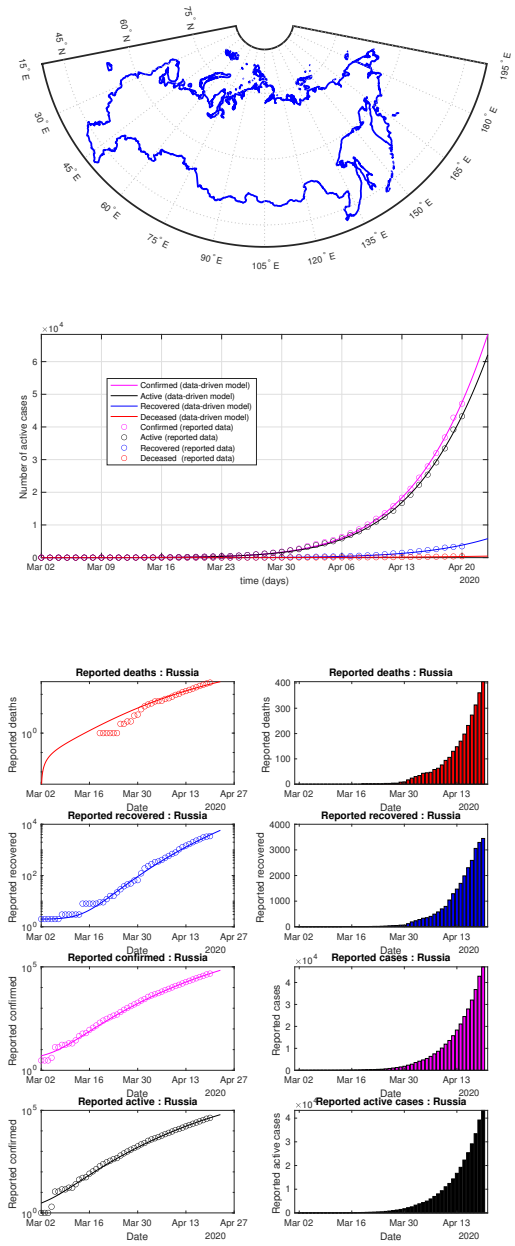

Figure 87: Russia: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

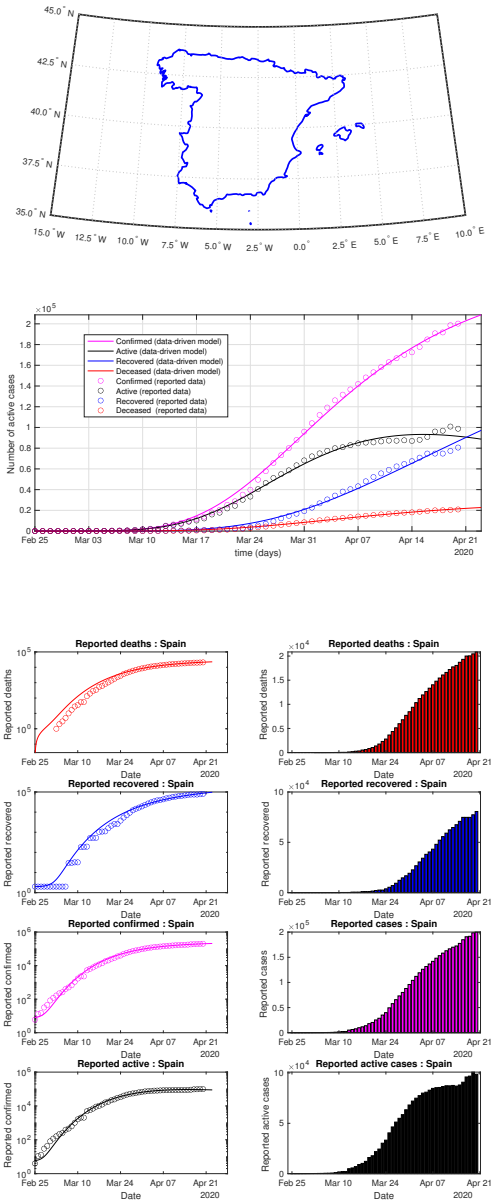

Figure 88: Spain: data of active, recovered, deceased (in circle or bar) vs model of active, recovered, deceased over time. Tracking the number of active cases. Predictive analytics for the next couple days

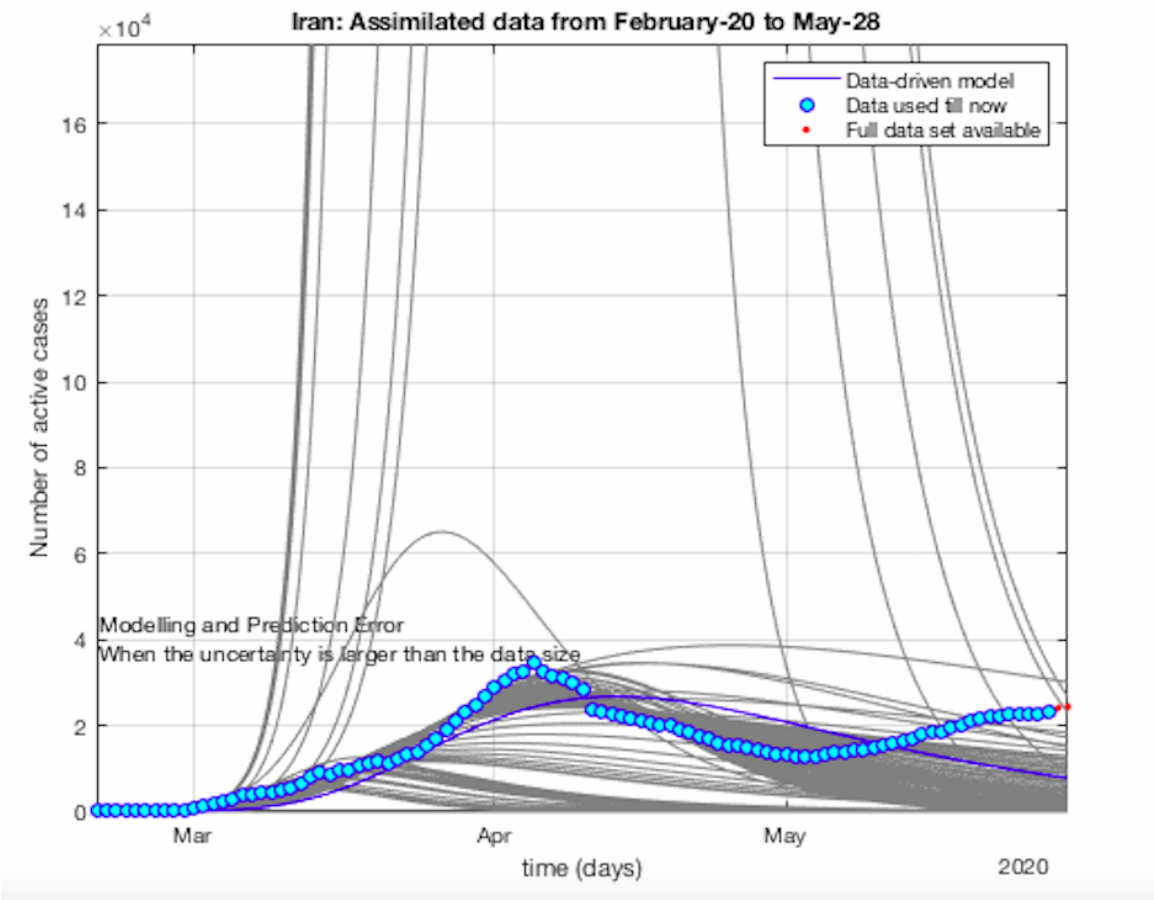

Figure 89: COVID-19 sample of Non-Gaussianity and non-exponential. The plots are obtained by using the data set sequentially till the latest data point. A **Non-Gausianity** of the reported active cases and hospitalized cases is observed. The shape of the sequential data-driven curve is clearly not a single exponential.

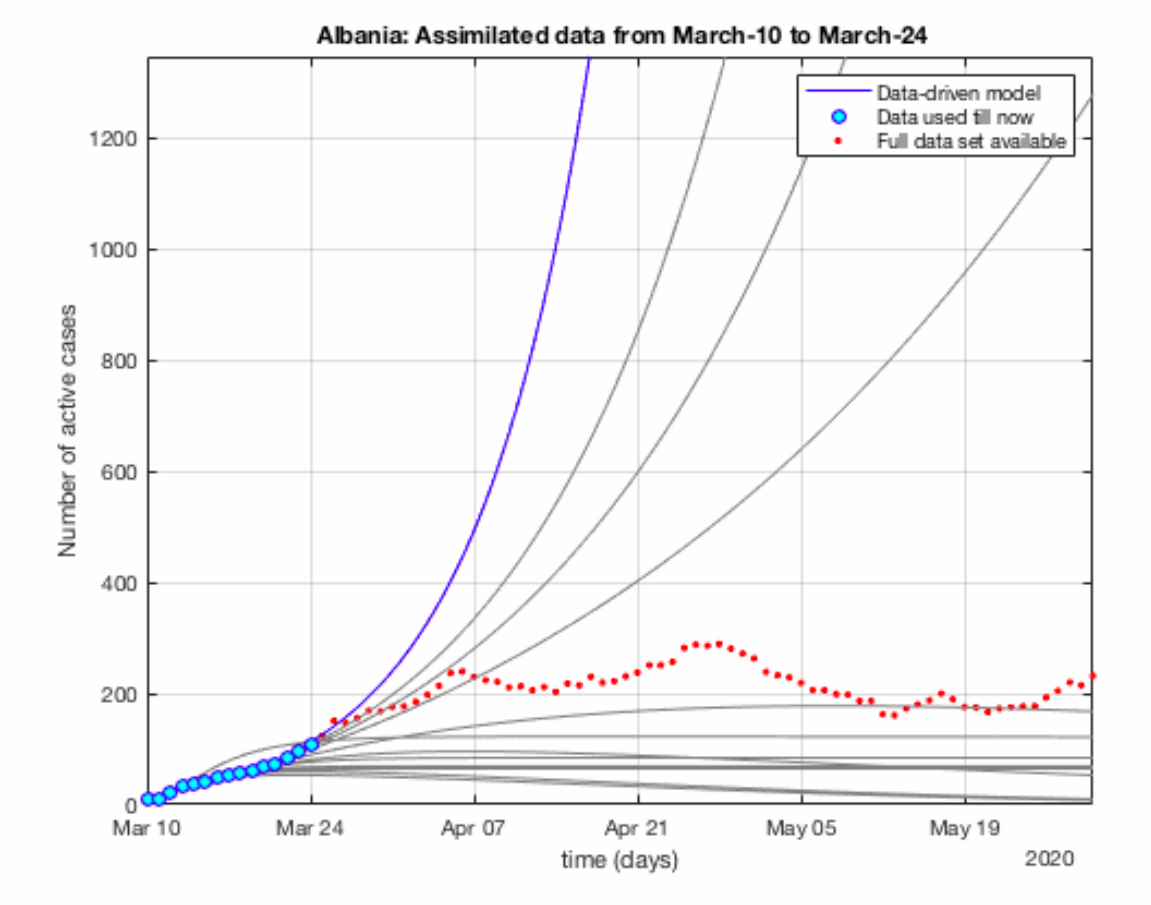

Figure 90: Another COVID-19 sample of Non-Gaussianity and non-exponential. The plots are obtained by using the data set sequentially till the latest data point. A **Non-Gausianity** of the reported active cases and hospitalized cases is observed. The shape of the sequential data-driven curve is clearly not a single exponential.