MODELLING THE COVID-19 PANDEMIC REQUIRES A MODEL...
BUT ALSO DATA!

Marc Lavielle
Inria Saclay & Ecole Polytechnique

May 20th, 2020

Workshop: Modeling the propagation of Covid-19
https://coronavirus.jhu.edu/map.html
Time series summary (csse_covid_19_time_series)

This folder contains daily time series summary tables, including confirmed, deaths and recovered. All data is read in from the daily case report. The time series tables are subject to be updated if inaccuracies are identified in our historical data. The daily reports will not be adjusted in these instances to maintain a record of raw data.

Two time series tables are for the US confirmed cases and deaths, reported at the county level. They are named `time_series_covid19_confirmed_US.csv`, `time_series_covid19_deaths_US.csv`, respectively.
Cumulated number of confirmed cases
Cumulated number of deaths
• The objective is not to build a model... and try to “calibrate” it in order to fit the data as well as possible
• The objective is not to build a model... and try to “calibrate” it in order to fit the data as well as possible

• The objective is to develop a model
  - for the observed data, and validated by the data,
  - that provides good short-term predictions,
  - that is mechanistic and parsimonious,
  - that is implemented as an open-access interactive tool.
The epidemiological compartments

- $S_i$: susceptibles in site $i$
- $E_i$: exposed in $i$
- $P_i$: pre-symptomatic infectious in $i$
- $I_i$: symptomatic infectious in $i$
- $A_i$: asymptomatic/mildly symptomatic infectious in $i$
- $H_i$, $Q_i$: Hospitalized, Quarantined and isolated in $i$
- $D_i$, $R_i$: Deceased, Recovered in $i$

Marino Gatto talk
The epidemiological compartments

\[ \begin{align*}
S_i &: \text{susceptibles in site } i \\
E_i &: \text{exposed in } i \\
P_i &: \text{pre-symptomatic infectious in } i \\
I_i &: \text{symptomatic infectious in } i \\
A_i &: \text{asymptomatic/mildly symptomatic infectious in } i \\
H_i, Q_i &: \text{Hospitalized, Quarantined and isolated in } i \\
D_i, R_i &: \text{Deceased, Recovered in } i
\end{align*} \]

Here, site = country

Marino Gatto talk
The epidemiological compartments

- $S_i$: susceptibles in site $i$
- $E_i$: exposed in $i$
- $P_i$: pre-symptomatic infectious in $i$
- $I_i$: symptomatic infectious in $i$
- $A_i$: asymptomatic/mildly symptomatic infectious in $i$
- $H_i$, $Q_i$: Hospitalized, Quarantined and isolated in $i$
- $D_i$, $R_i$: Deceased, Recovered in $i$

Marino Gatto talk
The epidemiological compartments

- $S_i$: susceptibles in site $i$
- $E_i$: exposed in $i$
- $P_i$: pre-symptomatic infectious in $i$
- $I_i$: symptomatic infectious in $i$
- $A_i$: asymptomatic/mildly symptomatic infectious in $i$
- $H_i$, $Q_i$: Hospitalized, Quarantined and isolated in $i$
- $D_i$, $R_i$: Deceased, Recovered in $i$
The epidemiological compartments

- $S_i$: susceptibles in site $i$
- $H_i$: Hospitalized,
  Quarantined and isolated in $i$
- $D_i, R_i$: Deceased, Recovered in $i$
The epidemiological compartments

$S_i$: susceptibles in site $i$

$H_i$: Hospitalized,
Quarantined and isolated in $i$

$D_i, R_i$: Deceased, Recovered in $i$
The epidemiological compartments

\[ S_i : \text{susceptibles in site } i \]

\[ H_i : \text{Hospitalized,} \]
\[ \text{Quarantined and isolated in } i \]
\[ D_i, R_i : \text{Deceased, Recovered in } i \]
The epidemiological compartments
The epidemiological compartments

\[ \dot{S}(t) = -\beta \frac{S(t)}{N} I(t) \]

\[ \dot{I}(t) = \beta \frac{S(t)}{N} I(t) - \mu I(t) - \nu I(t) \]

\[ \dot{R}(t) = \mu I(t) \]

\[ \dot{H}(t) = \nu I(t) - \lambda H(t) \]

\[ \dot{D}(t) = \lambda H(t) \]
The epidemiological compartments

\[ \dot{S}(t) = -\beta \frac{S(t)}{N} I(t) \]

\[ \dot{I}(t) = \beta \frac{S(t)}{N} I(t) - \mu I(t) - \nu I(t) \]

\[ \dot{R}(t) = \mu I(t) \]

\[ \dot{H}(t) = \nu I(t) - \lambda H(t) \]

\[ \dot{D}(t) = \lambda H(t) \]

Approximation: \( S(t) = N \)
The epidemiological compartments

\[ \dot{I}(t) = \beta I(t) - \mu I(t) - \nu I(t) \]

\[ \dot{R}(t) = \mu I(t) \]

\[ \dot{H}(t) = \nu I(t) - \lambda H(t) \]

\[ \dot{D}(t) = \lambda H(t) \]
The epidemiological compartments

\[ \dot{I}(t) = \beta I(t) - \mu I(t) - \nu I(t) \]

\[ \dot{R}(t) = \mu I(t) \]

\[ \dot{H}(t) = \nu I(t) - \lambda H(t) \]

\[ \dot{D}(t) = \lambda H(t) \]

_R not needed for fitting the data_
The epidemiological compartments

\[ \dot{I}(t) = \beta I(t) - \mu I(t) - \nu I(t) \]

\[ \dot{H}(t) = \nu I(t) - \lambda H(t) \]

\[ \dot{D}(t) = \lambda H(t) \]
The transmission rate $\beta$ changes over time

$\dot{I}(t) = \beta I(t) - \mu I(t) - \nu I(t)$

$\dot{H}(t) = \nu I(t) - \lambda H(t)$

$\dot{D}(t) = \lambda H(t)$
The epidemiological compartments

\[
\begin{align*}
\dot{I}(t) &= \beta(t)I(t) - \mu I(t) - \nu I(t) \\
\dot{H}(t) &= \nu I(t) - \lambda H(t) \\
\dot{D}(t) &= \lambda H(t)
\end{align*}
\]
We will use a piecewise linear function for $\beta$.

The epidemiological compartments

\[
\begin{align*}
\dot{I}(t) &= \beta(t)I(t) - \mu I(t) - \nu I(t) \\
\dot{H}(t) &= \nu I(t) - \lambda H(t) \\
\dot{D}(t) &= \lambda H(t) \\
\beta(t) &= \beta_0 + a t + \sum_{k=1}^{K} h_k \cdot t \chi[t \geq \tau_k]
\end{align*}
\]
The observation model

Available data:

• \( w = (w_j, j = 1, 2, ...) \) where \( w_j \) is the cumulated number of confirmed cases

Then, \( w_j - w_{j-1} \) is the number of new confirmed cases on day \( j \)

• \( d = (d_j, j = 1, 2, ...) \) where \( d_j \) is the cumulated number of deaths

Then, \( d_j - d_{j-1} \) is the number of new deaths on day \( j \)
http://shiny.webpopix.org/covidix/app2/
http://shiny.webpopix.org/covidix/app2/
The observation model

The data:

- cumulated number of confirmed cases ($w_j$)

- cumulated number of deaths ($d_j$)

The epidemiological model:

\[
\begin{align*}
\dot{I}(t) &= \beta(t) I(t) - \mu I(t) - \nu I(t) \\
\dot{H}(t) &= \nu I(t) - \lambda H(t) \\
\dot{D}(t) &= \lambda H(t)
\end{align*}
\]
The observation model

The data:

• cumulated number of confirmed cases \( w_j \)

• cumulated number of deaths \( d_j \)

We assume that only a fraction \( \alpha \) of the infected people are confirmed.

The epidemiological model:

\[
\begin{align*}
\dot{I}(t) &= \beta(t)I(t) - \mu I(t) - \nu I(t) \\
\dot{H}(t) &= \nu I(t) - \lambda H(t) \\
\dot{D}(t) &= \lambda H(t)
\end{align*}
\]
The observation model

The data:

- cumulated number of confirmed cases \( w_j \)
- cumulated number of deaths \( d_j \)

We assume that only a fraction \( \alpha \) of the infected people are confirmed.

The epidemiological model:

\[
\begin{align*}
\dot{I}(t) &= \beta(t)I(t) - \mu I(t) - \nu I(t) \\
\dot{W}_c(t) &= \alpha \beta(t)I(t) \\
\dot{H}(t) &= \nu I(t) - \lambda H(t) \\
\dot{D}(t) &= \lambda H(t)
\end{align*}
\]
The observation model

The data:

• cumulated number of confirmed cases \( w_j \)
  \( w_j \) predicted by \( W_c(t_j) \)

• cumulated number of deaths \( d_j \)
  \( d_j \) predicted by \( D(t_j) \)

The epidemiological model:

\[
\begin{align*}
\dot{I}(t) &= \beta(t)I(t) - \mu I(t) - \nu I(t) \\
\dot{W}_c(t) &= \alpha \beta(t)I(t) \\
\dot{H}(t) &= \nu I(t) - \lambda H(t) \\
\dot{D}(t) &= \lambda H(t)
\end{align*}
\]
The observation model

A first statistical model for the daily counts:

\[ w_j - w_{j-1} = W(t_j) - W(t_{j-1}) + e_j \ ; \quad e_j \sim \mathcal{N}(0, \sigma_e^2) \]

\[ d_j - d_{j-1} = D(t_j) - D(t_{j-1}) + u_j \ ; \quad u_j \sim \mathcal{N}(0, \sigma_u^2) \]

Parameters of the model:

\[ \theta = (\alpha, \beta_0, a, h_1, \ldots, h_K, \tau_1, \ldots, \tau_K, \mu, \nu, \lambda, I_0, H_0, D_0, \sigma_e^2, \sigma_u^2) \]

\[ \dot{I}(t) = \beta(t) I(t) - \mu I(t) - \nu I(t) \]

\[ \dot{W}_e(t) = \alpha \beta(t) I(t) \]

\[ \dot{H}(t) = \nu I(t) - \lambda H(t) \]

\[ \dot{D}(t) = \lambda H(t) \]

\[ \beta(t) = \beta_0 + a t + \sum_{k=1}^{K} h_k t \times \mathbb{I}\{t \geq \tau_k\} \]
The observation model

A first statistical model for the daily counts:

\[ w_j - w_{j-1} = W(t_j) - W(t_{j-1}) + e_j \quad ; \quad e_j \sim \mathcal{N}(0, \sigma_e^2) \]
\[ d_j - d_{j-1} = D(t_j) - D(t_{j-1}) + u_j \quad ; \quad u_j \sim \mathcal{N}(0, \sigma_u^2) \]

Parameters of the model:

\[ \theta = (\alpha, \beta_0, a, h_1, \ldots, h_K, \tau_1, \ldots, \tau_K, \mu, \nu, \lambda, I_0, H_0, D_0, \sigma_e^2, \sigma_u^2) \]

\[ \theta \quad \text{obtained by Maximum Likelihood (ML) Estimation} \]

\[ K \quad \text{obtained by minimizing the Bayesian Information Criteria (BIC)} \]
Fitting the Italian data

Cumulated numbers

- Confirmed
- Deaths

March
April
May
Fitting the Italian data

Cumulated numbers

Transmission rate
Fitting the Italian data

Cumulated numbers

Daily numbers

Transmission rate
Fitting the Italian data

The residuals (daily numbers)
Fitting the Italian data

The residuals errors exhibit a weekly periodic component: the observation model should include this periodic component.
Fitting the Italian data

The residuals errors exhibit a weekly periodic component: the observation model should include this periodic component.
The residuals errors exhibit a weekly periodic component: the observation model should include this periodic component.

The magnitude of the errors increase with the prediction: the observation model should include a proportional error model.
The observation model

A second statistical model for the daily counts:

\begin{align*}
    w_j - w_{j-1} &= (W(t_j) - W(t_{j-1})) \left( 1 + A \cos \left( \frac{2\pi}{7} t_j + \phi \right) \right) (1 + e_j) \\
    e_j &\sim \mathcal{N}(0, \sigma_e^2) \\
    d_j - d_{j-1} &= (D(t_j) - D(t_{j-1})) \left( 1 + B \cos \left( \frac{2\pi}{7} t_j + \phi \right) \right) (1 + u_j) \\
    u_j &\sim \mathcal{N}(0, \sigma_u^2)
\end{align*}

\begin{align*}
    \dot{I}(t) &= \beta(t)I(t) - \mu I(t) - \nu I(t) \\
    \dot{W}_c(t) &= \alpha \beta(t)I(t) \\
    \dot{H}(t) &= \nu I(t) - \lambda H(t) \\
    \dot{D}(t) &= \lambda H(t) \\
    \beta(t) &= \beta_0 + a t + \sum_{k=1}^{K} h_k t \times \mathbb{I}\{t \geq \tau_k\}
\end{align*}
Fitting the Italian data

Cumulated numbers

Daily numbers

Transmission rate
Some fits with the periodic component
Some fits without the periodic component
About the basic reproduction number

\[ \dot{I}(t) = \beta(t)I(t) - \mu I(t) - \nu I(t) \]
\[ R_0 = \frac{\beta(t)}{\mu + \nu} \]

- Not so easy to understand (at least for me)
- Seems to depend on the model (and not only on the parameter values)

The difference \( \beta(t) - \mu + \nu \)

- (maybe) more informative than the ratio \( \beta(t)/(\mu + \nu) \)
- easy to interpret, as related to the “half-life”

\[ t_2 = \log(2)/(\beta - \mu + \nu) \]
\[ t_{1/2} = -\log(2)/(\beta - \mu + \nu) \]
<table>
<thead>
<tr>
<th>Country</th>
<th>Increase</th>
<th>Peak of daily deaths</th>
<th>Decrease</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t4 (↑)</td>
<td>Date</td>
<td>n</td>
</tr>
<tr>
<td>Denmark</td>
<td>12</td>
<td>4/3/2020</td>
<td>16</td>
</tr>
<tr>
<td>Portugal</td>
<td>19</td>
<td>4/11/2020</td>
<td>32</td>
</tr>
<tr>
<td>Switzerland</td>
<td>16</td>
<td>4/5/2020</td>
<td>58</td>
</tr>
<tr>
<td>Netherlands</td>
<td>13</td>
<td>4/4/2020</td>
<td>155</td>
</tr>
<tr>
<td>Germany</td>
<td>19</td>
<td>4/13/2020</td>
<td>226</td>
</tr>
<tr>
<td>Belgium</td>
<td>21</td>
<td>4/15/2020</td>
<td>268</td>
</tr>
<tr>
<td>France</td>
<td>15</td>
<td>4/5/2020</td>
<td>532</td>
</tr>
<tr>
<td>Italy</td>
<td>15</td>
<td>3/27/2020</td>
<td>835</td>
</tr>
<tr>
<td>Spain</td>
<td>16</td>
<td>4/3/2020</td>
<td>839</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>19</td>
<td>4/13/2020</td>
<td>927</td>
</tr>
</tbody>
</table>

Maximum daily number of deaths predicted by the model. For each country, t4 (↑) and t2 (↑) are, respectively, the number of days it took to multiply the number of deaths by 4 and 2; t1/2 (↓) and t1/4 (↓) are, respectively, the number of days it took to divide the number of deaths by 2 and 4.
A monitoring tool

This tool can be useful for detecting unexpected changes in the dynamics of the epidemics.

Portuguese data:
A monitoring tool

This tool can be useful for detecting unexpected changes in the dynamics of the epidemics.

Portuguese data:
A monitoring tool

This tool can be useful for detecting unexpected changes in the dynamics of the epidemics.

Portuguese data:
A monitoring tool

This tool can be useful for detecting unexpected changes in the dynamics of the epidemics.

Spanish data:
A monitoring tool

This tool can be useful for detecting unexpected changes in the dynamics of the epidemics.

Spanish data:
A monitoring tool

This tool can be useful for detecting unexpected changes in the dynamics of the epidemics.

Spanish data:
A monitoring tool

This tool can be useful for detecting unexpected changes in the dynamics of the epidemics.

... or to consider as expected what may seem unexpected
A monitoring tool

This tool can be useful for detecting unexpected changes in the dynamics of the epidemics.

... or to consider as expected what may seem unexpected.
A monitoring tool

This tool can be useful for detecting unexpected changes in the dynamics of the epidemics.

... or to consider as expected what may seem unexpected
A monitoring tool

This tool can be useful for detecting unexpected changes in the dynamics of the epidemics.

... or to consider as expected what may seem unexpected
A monitoring tool

This tool can be useful for detecting unexpected changes in the dynamics of the epidemics.

... or to consider as expected what may seem unexpected
Possible scenarios after the end of the lockdown

1) The transmission rate remains the same
Possible scenarios after the end of the lockdown

2) The transmission rate is multiplied by 1.5
Possible scenarios after the end of the lockdown

3) The transmission rate is multiplied by 2
Possible scenarios after the end of the lockdown

4) The lockdown ends 2 weeks later (May 25)
Possible scenarios before/after the lockdown

5) The lockdown starts one week before (March 10)
About the data

Some French data...
About the data

Ireland:

Casos nuevos diarios con coronavirus en España
About Bayesian methods

Bayesian approach is great!
... but it's often used in extremist ways:
About Bayesian methods

Bayesian approach is great!
... but it's often used in extremist ways:

• Some parameters are definitively fixed
  (looks like repressive Bayesian... $\beta=3$, no debate!)
About Bayesian methods

Bayesian approach is great!
... but it's often used in extremist ways:

• Some parameters are definitively fixed
  (looks like repressive Bayesian... \( \beta=3 \), no debate!)

• The other parameters are equipped with non informative priors
  (looks like Platonician Bayesian: I know that I know nothing... maybe \( \beta=3 \), but \( \beta=100\ 000 \)...)

Bayesian approach should be used for introducing prior information
(I think that \( \beta \) should not be far from 3... probably between 2.5 and 3.5)
Bayesian approach is great!
... but it's often used in extremist ways:

- Some parameters are definitively fixed
  *(looks like repressive Bayesian... $\beta=3$, no debate!)*

- The other parameters are equipped with non informative priors
  *(looks like Platonician Bayesian: I know that I know nothing... maybe $\beta=3$, but $\beta=100\,000...$)*

Bayesian approach should be used for introducing prior information
*(I think that $\beta$ should not be far from 3... probably between 2.5 and 3.5)*
Thank you!