



**HAL**  
open science

# ML-CI: Machine Learning Confidence Intervals for Covid-19 forecasts

Alice Lacan, Isabelle Guyon

► **To cite this version:**

Alice Lacan, Isabelle Guyon. ML-CI: Machine Learning Confidence Intervals for Covid-19 forecasts. BayLearn - Machine Learning Symposium 2021, Oct 2021, San Francisco, United States. hal-03501101

**HAL Id: hal-03501101**

**<https://hal.science/hal-03501101>**

Submitted on 23 Dec 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ML-CI: Machine Learning Confidence Intervals for Covid-19 forecasts

**Alice Lacan**  
LISN/Inria/CNRS, U. Paris-Saclay  
alice.b.lacan@gmail.com

**Isabelle Guyon**  
ChaLearn, San Francisco, California  
LISN/Inria/CNRS, U. Paris-Saclay  
guyon@chalearn.org

## Abstract

Epidemic forecasting has always been challenging and the recent Covid-19 outbreaks emphasizes it. We introduce a novel approach to address the problem of evaluating confidence intervals (CI) of time series prediction forecasts for compartmental models, using machine learning. We evaluate our approach using real data of the Covid pandemic on 27 countries. Compartmental models were trained taking into account non pharmaceutical governmental measures. A Random Forest regressor was trained, using various engineered features, to predict the forecasting error for various horizons on synthetic data, then applied to estimate CI on real data forecasts. Our method outperforms baselines using forecast likelihood as metric.

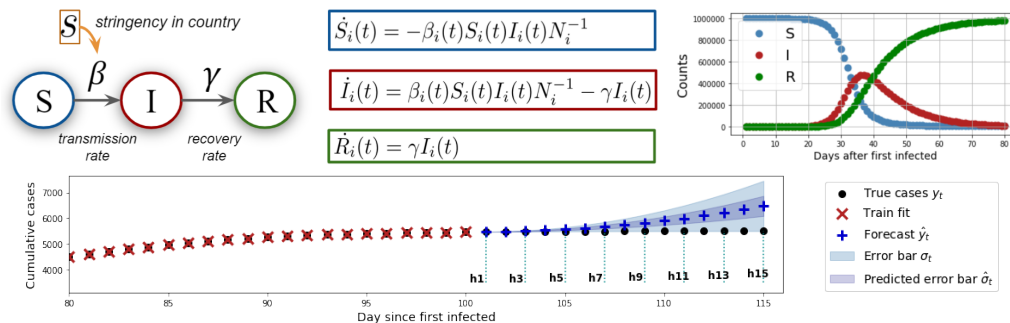


Figure 1: [Top] Compartmental Susceptible-Infected-Recovered model (SIR), with dynamic transmission rate  $\beta$  oppositely related to the stringency of government Non-Pharmaceutical Interventions (NPI), and a fixed recovery rate  $\gamma = 07$ . [Bottom] Country cumulative cases (sum of S and R compartments). Model fit (red crosses) and 15 day forecasts (blue crosses). A ML model is trained as estimator  $\hat{\sigma}_t$  of errors  $\sigma_t$  on synthetic data at each horizon  $t$ ; then applied to real data.

Since January 2020, policy-making has been steered by Covid-19 cases and deaths forecasts. Forecasting methods encompass epidemiological compartmental models [1][2] and deep learning models [3][4][5] among others. Poorly advised policies can greatly impact populations with human and economical loss [6]. Due to the high degree of uncertainty of predictions, point-wise forecasts are less relevant than interval forecasts. But, despite the vast amount of literature on Covid-19 forecasting, few studies have provided confidence intervals (CI) [7]. In addition, the few papers mentioning CI, rarely document the actual methods used for interval forecasting. Bayesian approaches including Gaussian processes [8], have gained in popularity, but make unverifiable assumptions [9][10][11][12]. In this paper, we present a global method (Figure 1) mapping the decision-making process to the disease evolution with a Confidence Intervals estimated with a Machine Learning method (ML-CI).

Table 1: Negative log-likelihood for the 15-day horizon interval forecast for 27 countries. Best methods are highlighted in bold (smaller values are better).

|                     |                |               |           |           |           |             |           |           |             |
|---------------------|----------------|---------------|-----------|-----------|-----------|-------------|-----------|-----------|-------------|
|                     | United Kingdom | United States | Poland    | Portugal  | Belgium   | Bulgaria    | Canada    | Denmark   | Finland     |
| <b>ML-CI (OURS)</b> | <b>13</b>      | <b>15</b>     | <b>17</b> | <b>14</b> | <b>11</b> | <b>113</b>  | <b>12</b> | <b>10</b> | <b>9</b>    |
| Naive               | 14             | 18            | 128       | 18        | <b>11</b> | 385         | 18        | 11        | 11          |
| Bootstrap           | <b>13</b>      | 18            | 284       | 633       | 56        | 565         | 14        | 77        | 35          |
|                     | France         | Germany       | Russia    | India     | Israel    | Italy       | Japan     | Norway    | Netherlands |
| <b>ML-CI (OURS)</b> | <b>13</b>      | <b>12</b>     | <b>16</b> | <b>12</b> | <b>10</b> | <b>13</b>   | <b>9</b>  | <b>11</b> | <b>11</b>   |
| Naive               | <b>13</b>      | <b>12</b>     | 121       | 128       | <b>10</b> | 21          | 13        | <b>9</b>  | 12          |
| Bootstrap           | 16             | 16            | 20        | 17        | 54        | 26          | 53        | <b>9</b>  | 184         |
|                     | Luxembourg     | South Africa  | Greece    | Australia | Austria   | South Korea | Spain     | Sweden    | Taiwan      |
| <b>ML-CI (OURS)</b> | <b>11</b>      | <b>12</b>     | 9         | 9         | 10        | 10          | 13        | 11        | 9           |
| Naive               | <b>14</b>      | 63            | 9         | 7         | <b>9</b>  | <b>7</b>    | <b>11</b> | <b>10</b> | <b>7</b>    |
| Bootstrap           | <b>10</b>      | 18            | <b>8</b>  | 7         | 10        | 7           | 15        | 11        | 284         |

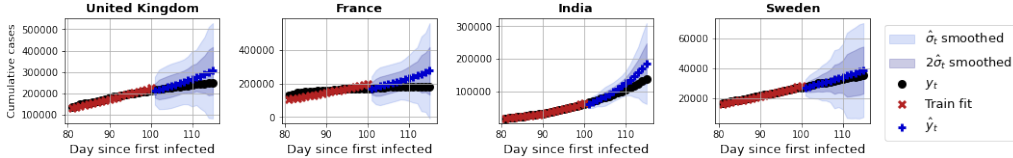


Figure 2: Examples of error bar prediction with our method ML-CI.

We focus on the widely used compartmental Susceptible-Infected-Recovered model (SIR) [1], from which we forecast Covid-19 cumulative cases (Figure 1). We limit our study to this simple model because more complex models have shown only minor (or null) performance advantage [7], and comparing predictive models is not the object of this paper. SIR models are trained on 27 countries from the OxCGRT dataset [13], which provides cumulative cases and government Non-Pharmaceutical Interventions (NPIs), from January 2020 to December 2020. We fit as few parameters as possible: The recovery rate  $\gamma$  is fixed as  $1/N_r$ , with  $N_r$  the average number of days to recover [2]. The transmission rate  $\beta$  is obtained through an affine transformation of the Oxford index summarizing all NPIs. The two parameters of the affine transformation are fit, on a 100 day window, using a dual annealing optimization, essentially a simulated annealing approach with a local search ([Documentation](#)).

The novel contribution of this paper is a method for estimating confidence intervals, based on training a Random Forest (RF) regressor, inspired by recent work [12]. To train the regressor, we created a simulated dataset of 2000 virtual countries, using the SIR model described above, displaying a wide range of NPI scenarios. The inputs to the RF regressor include a variety of features derived from past data, such as past errors, past cumulative forecasts, past daily case forecasts, and past stringency. We train one regressor per horizon  $h$ , for a 15 day forecast, using as target value the absolute difference between prediction and actual number of cases. Subsequently, the output of the RF regressor  $h$  is used to estimate the forecast error at horizon  $h$ , for predictions made on real data by a SIR model. We use a simple heuristic to lightly smooth the error bar. Our code is available on [Github](#).

We compare our method (ML-CI) with two baseline methods for computing error bars: (1) a naive method [14] assuming a simple diffusion model, advocating to use  $s_{t-15:t}\sqrt{h}$ , where  $s_{t-15:t}$  is the average RMSE over the past 15 days, and  $h$  the prediction horizon; (2) a bootstrap approach using as prediction error the standard deviation of predictions of 4 SIR models trained on data subsamples. Table 1 shows negative log-likelihood performance for a Gaussian error model, at 15-day horizon forecast. Figure 2 provides examples of 15-day forecasts for a few countries (results on all 27 countries are in Appendix).

Our preliminary results are encouraging, indicating that ML regression is a viable alternative method to more conventional techniques, for predicting confidence intervals in forecasting applications. Robust methods for epidemiological time series interval forecasting being in great need, this line of research to obtain reliable confidence intervals seems promising. Our future work includes conducting systematic experiments on more countries and with a time sliding window, and compare ML-CI with more baseline methods on a variety of datasets. Improvements may be obtained by experimenting with other machine learning regressors, better feature engineering, and calibrating the output of the regressor. Together with improvements in our forecasting model (e.g. taking into account mobility data and people compliance to NPIs) we hope to advance the state of the art.

## Acknowledgments and Disclosure of Funding

The authors wish to acknowledge the members of the Transatlantic team<sup>1</sup> with whom we were part of the XPRIZE-Cognisant Pandemic Response challenge: Prasanna Balaprakash, Xavier Baro, Martin Cepeda, Romain Engele, Sergio Escalera, Julio Jacques, Rafael Munoz, Alma Paris, Sebastien Treguer, and Wei-Wei Tu, for stimulating discussions that prompted this work. This work was supported by the ANR (Agence Nationale de la Recherche, National Agency for Research) under AI chair of excellence HUMANIA, grant number ANR-19-CHIA-00222 and INRIA.

## References

- [1] Kermack William Ogilvy and McKendrick A. G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A* 115700–721, 1927. <http://doi.org/10.1098/rspa.1927.0118>.
- [2] Yi-Cheng Chen, Ping-En Lu, Cheng-Shang Chang, and Tzu-Hsuan Liu. A time-dependent sir model for covid-19 with undetectable infected persons. *IEEE Transactions on Network Science and Engineering*, 7(4):3279–3294, 2020.
- [3] Zifeng Yang, Zhiqi Zeng, Ke Wang, and Sook-San Wong et al. Modified seir and ai prediction of the epidemics trend of covid-19 in china under public health interventions. *Journal of Thoracic Disease*, 12(3), 2020.
- [4] Duxin Chen, Yifan Yang, Yifan Zhang, and Wenwu Yu. Prediction of covid-19 spread by sliding mseir observer. *Sci. China Inf. Sci.* 63, 222203, 12(3), 2020.
- [5] Abdelhafid Zeroual, Fouzi Harrou, Abdelkader Dairi, and Ying Sun. Deep learning methods for forecasting covid-19 time-series data: A comparative study. *Chaos, Solitons and Fractals*, 140:110121, 2020.
- [6] John P.A. Ioannidis, Sally Cripps, and Martin A. Tanner. Forecasting for covid-19 has failed. *International Journal of Forecasting*, 2020.
- [7] Janyce Eunice Gnanvi, Kolawolé Valère Salako, Gaëtan Brezesky Kotanmi, and Romain Glèlè Kakaï. On the reliability of predictions on covid-19 dynamics: A systematic and critical review of modelling techniques. *Infectious Disease Modelling*, 6:258–272, 2021.
- [8] Risto Miikkulainen, Olivier Francon, Elliot Meyerson, Xin Qiu, Darren Sargent, Elisa Canzani, and Babak Hodjat. From prediction to prescription: Evolutionary optimization of non-pharmaceutical interventions in the covid-19 pandemic. *IEEE Transactions on Evolutionary Computation*, 2021.
- [9] Mario Castro, Saúl Ares, José A. Cuesta, and Susanna Manrubia. The turning point and end of an expanding epidemic cannot be precisely forecast. *Proceedings of the National Academy of Sciences*, 117(42):26190–26196, 2020.
- [10] Alex Arenas, Wesley Cota, Jesús Gómez-Gardeñes, Sergio Gómez, Clara Granell, Joan T. Matamalas, David Soriano-Paños, and Benjamin Steinegger. Modeling the spatiotemporal epidemic spreading of covid-19 and the impact of mobility and social distancing interventions. *Phys. Rev. X*, 10:041055, Dec 2020.
- [11] H. Nishiura, N. M. Linton, and A. R. Akhmetzhanov. Serial interval of novel coronavirus (covid-19) infections. *Int.J. Infect. Dis.* 93, P284, 2020.
- [12] Moloud Abdar, Farhad Pourpanah, Sadiq Hussain, Dana Rezazadegan, Li Liu, Mohammad Ghavamzadeh, Paul Fieguth, Xiaochun Cao, Abbas Khosravi, U. Rajendra Acharya, Vladimir Makarenkov, and Saeid Nahavandi. A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *CoRR*, abs/2011.06225, 2020.

---

<sup>1</sup>Transatlantic team github repository: <https://github.com/transatlantic-team>.

- [13] Anna Petherick, Beatriz Kira, Emily Cameron-Blake, Helen Tatlow, Laura Hallas, Thomas Hale, and Toby Phillips et al. Variation in government responses to covid-19. *Blavatnik School working paper*, 2020.
- [14] R.J. Hyndman and G. Athanasopoulos. *Forecasting: principles and practice*. OTexts: Melbourne, Australia, 2018.

## Appendix

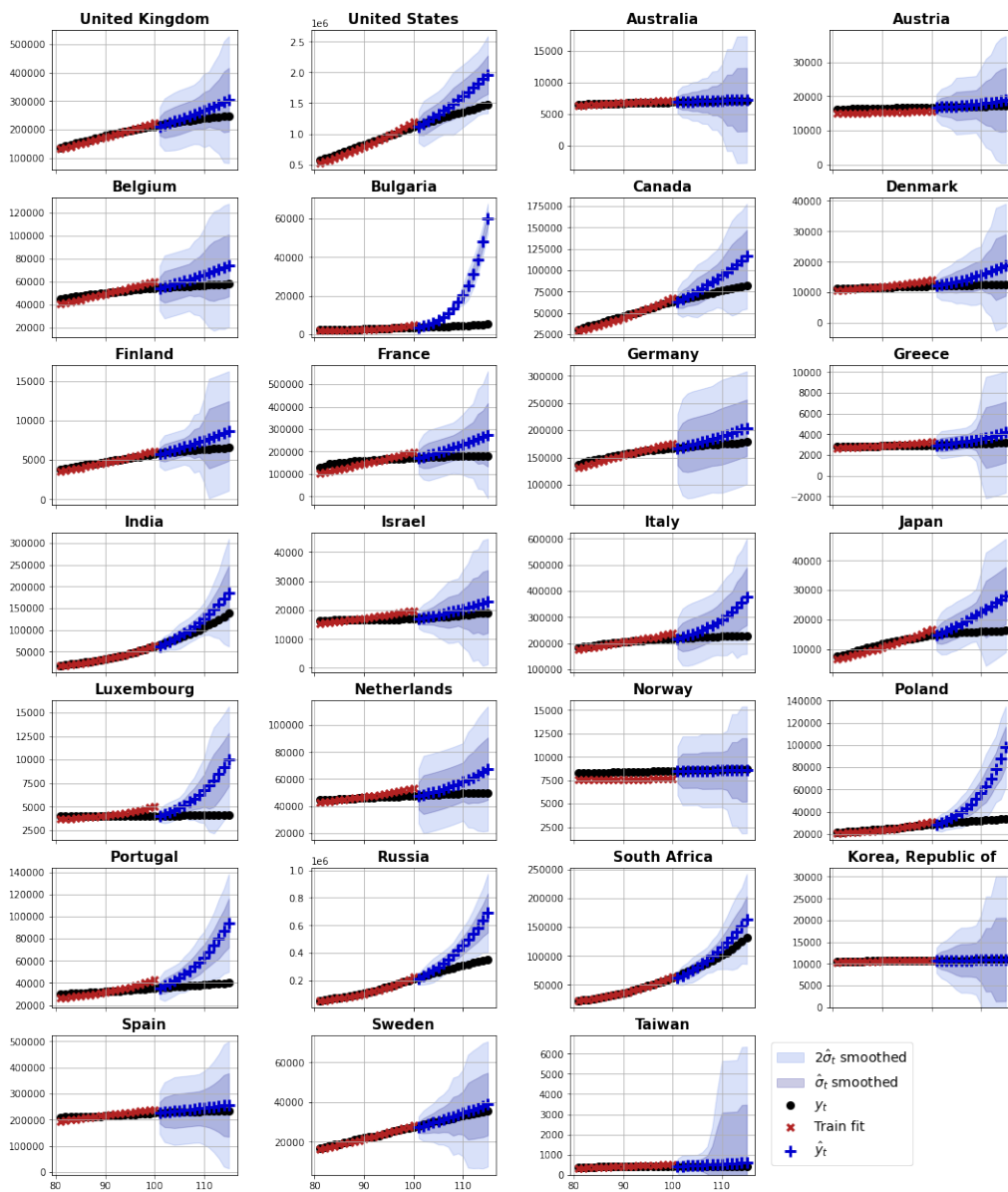


Figure 3: Error bar prediction with our method ML-CI for 27 countries.