Generalized Curve-Fitting Propagation Models for Epidemics

L. Ardon[*], F. Buet-Golfouse[*†], A. Macrina[†‡] , G. Papadopoulos[*], G. W. Peters[§]

[*]J. P. Morgan Chase Bank
London E14 5JP, United Kingdom
[†]Department of Mathematics, University College London
London WC1E 6BT, United Kingdom
[‡]African Institute of Financial Markets & Risk Management
University of Cape Town
Rondebosch 7701, South Africa
[§]Department of Actuarial Mathematics & Statistics
Heriot-Watt University
Edinburgh EH14 4AS, United Kingdom

January 6, 2021

Abstract

We propose a novel approach for...

Keywords:

1 Introduction

- Discussion of physical (epidemiological, structural) versus statistical (reduced-form) models.

- Identify sub-class of models in which our proposed model fits or is closest related to, and summarize specific model features of sub-class in comparison to other types of modelling approaches.

- Description of merits and shortcomings of our modelling approach, e.g. parsimony, flexibility, efficiency, interpretability, accuracy, and ease of data feeding. Also, model robustness, inclusion of attributes such as mobility, and applicability for decision making.

- Concise presentation of "winning model" within the model class we develop.

- Short summary of data set, which will be discussed and used later in the paper.

2 Research project goals

- Data-driven probabilistic model for spatial-temporal propagation of epidemics, such that data can be fed with ease and model outputs are interpretable.

- Produce evidence for NPIs to be useful and accurate models when generalized curve-fitting stochastic propagation models are developed.

- Model robustness against a number of features, e.g. noise, sparsity, spatial structure, etc.

Corresponding author: a.macrina@ucl.ac.uk

3 Data

- Detailed description of the data set and precise source, so that reproducing and validating any results presented in the paper is possible.

- Discussion of any procedure to improve quality of used data set. Transparency is of paramount importance. Any bias must be discussed and treatment thereof explained. The goal is to eliminate any possibility to argue against the quality of results obtained by the model, which can be traced back to data quality or mishandling.

- Any other data, which is not available or has not been used (why?) and which may produce improved model performance.

- Comparison of used data for our model with other data sets used in other studies and models. For example, does our approach require less data while achieving comparable or even better model performance (e.g. predictability or stress test)?

4 Modelling approach

- Non-compartmental Gaussian process approach where GP-kernel is used to capture various data features.

- "Robustification" of "winning model", log-transform...

- Multi-output model apt for capturing spatial heterogeneity (e.g. US states, vs USA or GER Länder vs GER, UK counties vs UK)

5 Model calibration

- Gaussian processes and regression technique.

- Robustness: sparsity, noise, spatial structure, averaging, etc.

6 Results

- New class of reduced-form propagation models for the stochastic dynamics of epidemics.

- Robust against stress-testing

- "Winning model" that has been trained on a dense data set, applicable with high level of accuracy to case with sparse data (e.g. trained on US data and applied to epidemics in Turkey, Mexico, etc., South Africa?).

- No MCMC... ?

  ...

References

A Appendix