Authors' Response to Discussions of Efficient Bayesian Inference of Instantaneous Reproduction Numbers at Fine Spatial Scales, with an Application to Mapping and Nowcasting the Covid-19 Epidemic in British Local Authorities

Yee Whye Teh[(1)], Bryn Elesedy[(1)], Bobby He[(1)], Michael Hutchinson[(1)], Sheheryar Zaidi[(1)], Avishkar Bhoopchand[(2)], Ulrich Paquet[(2)], Nenad Tomasev[(2)], Jonathan Read[(3)], Peter J. Diggle[(3)]
[(1)]Dept Statistics, University of Oxford
[(2)]Dept Statistics, University of Oxford, seconded from DeepMind
[(3)]CHICAS, Lancaster Medical School, Lancaster University

We warmly thank Chris Jewell, Guy Nason, Sebastian Funk & Sam Abbott, and Gavin Gibson, for their time and effort in reading our paper and for their great insights and discussion points which we value highly. These can roughly be collected into two broad categories: evaluation and modelling.

1. Evaluation

We begin with points raised about method evaluation. Funk and Abbott brought up the challenges of fair comparisons of such complex methods. We agree. Besides specific modelling choices and evaluation objectives, there are also complexities associated with reproducible software, and availability and access to different data sources which can affect the quality of the outputs of different methods. We agree that this is best led by an independent arbitrator with commitment by the wider community of methods developers.

Funk and Abbott also raised the issue of evaluating different models using data simulated from a single model. Indeed our main aim in that section paper is not to compare across different methods, but to verify that the methods are able to recover the groundtruth reproduction numbers to within reasonable tolerance in simple settings, and to identify specific behaviours of different methods. With kind help from Funk and Abbott, we have updated the way we used their excellent EpiNow2 software and

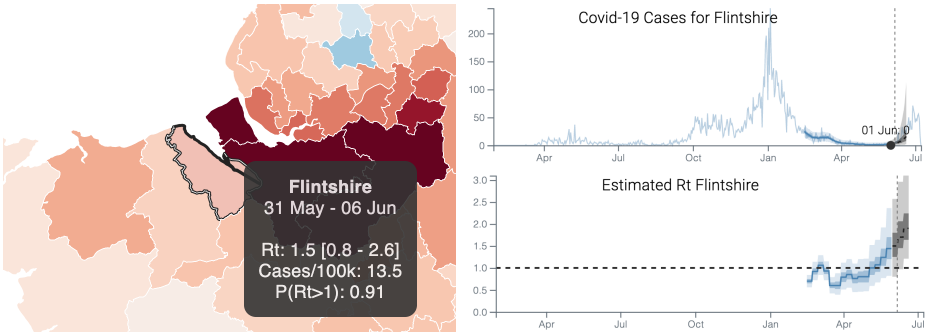

Fig. 1. EpiMap estimates for Flintshire on June 6, 2021.

updated the results in our paper accordingly. Gibson also suggested using data simulated from a more detailed model, which will be a good addition for a subsequent iteration of our project.

Nason brought up the important point of trustworthiness, given that estimates from different methods can differ so substantially from each other. The specific example used was that for Flintshire on June 6, 2021. Interestingly that was a great example of why we made the modelling choices that we made in the EpiMap model. The estimates are reproduced in Figure 1, and can be accessed at `https://localcovid.info/map.html?map=2021-06-08-bootstrap`. The estimated $R_{i,t}$ grew from a median of 1.4 during the last week modelled (May 24-30) to a forecasted median of 1.9 three weeks later (June 14-20), while cases went from 15 in the last week modelled to a median of 79 three weeks later. At that time Flintshire still had a very small number of cases and such growth would not have been predictable from the cases observed in Flintshire thus far. But it was right next to the hotspots of Wirral and Cheshire West and Chester, so the forecasted growth reflected good chance that the hotspots in the neighbouring LTLAs will spill over into Flintshire, which did happen. In fact the actual number of cases in Flintshire during the week of June 14-20 was 173, higher than the median forecasted but still within the 95% credible interval. This would not have been captured by a model that did not account for spatiotemporal patterns of epidemic growth.

2. Modelling

There are many excellent suggestions to consider different modelling variations. Gibson and Nason both suggested to consider non-pharmaceutical interventions (NPIs) as well as the effects of mobility within the model. If the data were readily available, we agree that these will likely improve the estimation and forecasting of the model, by incorporating them using the approach of Flaxman et al. (2020) in addition to the spatiotemporal Gaussian process.

Indeed the main limitation to further elaborating the model is the availability of high quality data that will help pin down the additional model complexity. As Gibson put it well "the richness of available data" needs to be well-matched to the "complexity of models". The same data limitation unfortunately applies to the nice suggestion of Jewell to model spatiotemporal variations in the generation interval distribution.

A related observation of Gibson is that of the role of mechanistic models and the interpretation of reproduction numbers, whether as an intrinsic quantity of interest or as a by-product of a detailed mechanistic model consisting of transmission processes, surveillance and control strategies. Our take on this is a pragmatic one: given the complexities of human social interactions and the lack of detailed data to fully pin down the underlying mechanisms of transmissions across the UK, it is best to summarise the complex mechanism using an interpretable quantity $R_{i,t}$ that can be fruitfully inferred from data.

The art of statistical modelling here finds an interesting parallel with the art of painting. How does the artist convey their thoughts and feelings accurately with a few pithy brush strokes, just as how the statistical modeller decide what to include in a model in view of available data and ultimate aim of inference? This is an important skill, and one that we are still honing and learning from the insightful comments of the discussants. Thank you!

References

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., Whittaker, C., Zhu, H., Berah, T., Eaton, J. W., et al. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, pages 1–8.