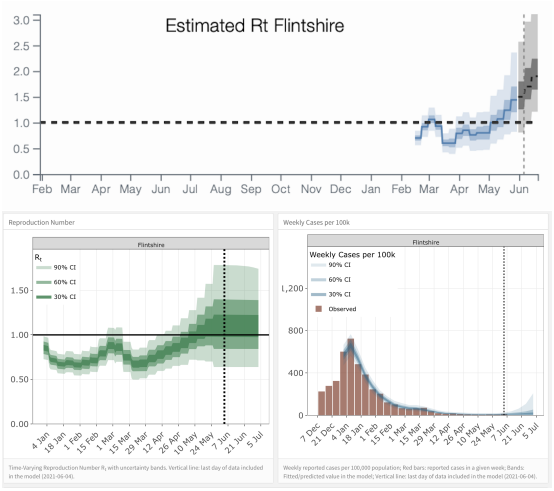

Figure 1: Flint estimate R_t from Teh *et al.* paper (top), and Mishra *et al.* (bottom left) for the Flint local authority for 6th June 2021 (right hand end)

Discussion of the papers by Teh et al. *and Mishra* et al.

Guy Nason, Department of Mathematics, Imperial College, London

I would like to congratulate both sets of authors for fascinating and worthy articles. They are clearly the result of hard and rapid work that has contributed to the public good during these terrible times. Both articles refer to websites that exhibit the outcomes of their contributions; presenting their work graphically and dynamically. The websites should be taken together in conjunction with the research articles I discuss below.

The top panel (from Teh *et al.*) shows that the forward projections of R_t are approaching the value of 2, and yet, the bottom panel (from Mishra *et al.*) shows their 90% CI is well below 2. Of course, with different models, there is no reason for the estimates of R_t to be the same or even close, but it is surprising the extent to which the models appear to give quite different answers. If one were a public health official, deciding on whether a local lockdown is necessary, what could they conclude from these plots? To be fair, in many other local authority areas, the estimates are closer. However, there is a serious statistical point here. How much modelling uncertainty has been incorporated into these models and, particularly, their credible intervals?

The articles stem from the common root of Flaxman *et al.* (2020) but in different ways. Teh *et al.* "share strength across localities and time" by introducing dependencies between reproduction numbers across neighbouring localities and model transmission across localities by using a spatial metapopulation model informed by 2011 UK Census commuter flow data. Mishra *et al.* incorporate case counts as well as deaths, incorporate prevalence survey data (e.g. REACT and ONS) and introduce time variation to concepts such as the infection ascertainment rate. These are all worthwhile ideas.

However, neither model non-pharmaceutical interventions (NPIs), but this is freely admitted by both. Of course, the effect of the NPIs will eventually feed through, but when and how? It was also difficult to see how pharma interventions (e.g. vaccines) were modelled, if at all? Both NPIs and PIs will have strongly influence the epidemic. It was also surprising that not more recent mobility information, or other variables, such as stringency indices, were incorporated, given the ease of access and similar granularity of such data. Overall, it seems odd that articles that identify as Bayesian do not make use of some of the strongest prior information available.

Both articles explained their modelling choices well. I am not personally a big fan of the"weekly jump" R_t choice. I can see why it has been employed and there is possibly an issue of whether the data directly tell you much about the 'finer than weekly' resolution. I think it does and smoother representations for R_t might have done a better job. For example, some kind of Bayesian nonparametric regression model using Fourier, splines or (even better)wavelets (for which Bayesian nonparametrics makes a lot of sense) and, indeed, the parametrisation might have been simpler.

Both models undertake model evaluation. Mishra *et al.* limits itself to median case projections versus observed weekly case numbers in Fig 5, highlighting poor performance in Dec 20 and Jan 21. Their conclusion is that there is a 'reasonable correspondence', but it seems to me that Sep 20 is systematically underestimating and Oct 20 seems skew. Teh *et al.* are to be congratulated for a more extensive evaluation and shows their model to be working reasonably well compared to a range of other methods (but not Mishra *et al.* as far as I can see) apart from after NPIs.

Although both articles are about local R_t estimation, they both use local, regional and national case counts of model quality and invite us to judge them on this too. Before the meeting, I wondered whether and how simpler statistical methods might compete in terms of case projections and suggested this to the authors. The response was along the lines of simpler methods could not possibly compete because they do not have the same significant amount of prior information incorporated. However, I thought it worth a try. Figure 5 in Teh *et al.* show case count estimates and projections for Scotland. Figure 2 reproduces part of this figure and uses automatic ARIMA fitting techniques (Hyndman and Khandakar, 2008) to compute an appropriate forecast and (frequentist) prediction intervals. It is interesting that the median 'prediction' from Teh *et al.* and the ARIMA point forecast are very similar. The 95% intervals for Teh *et al.* are (i) roughly the same as ARIMA at the midpoint and (ii) considerably larger the endpoint, near the beginning of May. The eventual true outcomes are also plotted on the figure and although both lower intervals seem to be doing a good job, to my mind, the upper ARIMA interval is better than the one from Teh *et al.* Of course, this is one small selected example.

The real message here is that more genuine evaluations would be welcome to develop a mature understanding of the efficacy of both methods, including comparing with each other, other similar methods and, definitely, simpler well-understood methods. For example, if I was director of public health in Scotland,

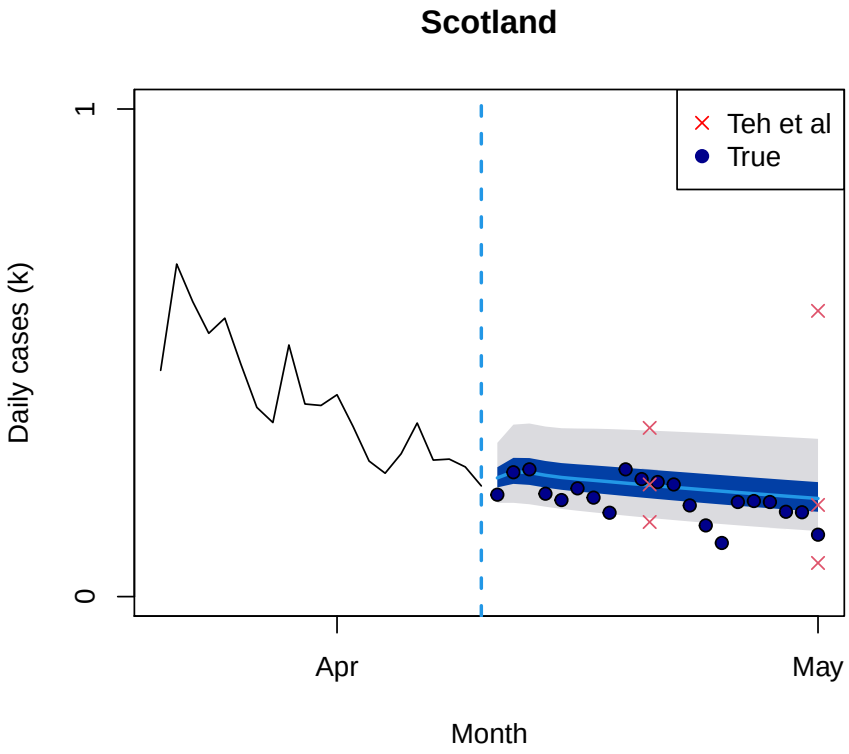

Figure 2: Scotland daily cases in 2021 (black). Blue dashed line: last date included in model (data available on 10th April). Red crosses: median projection and lower/upper 95% credible intervals (manually copied from Teh *et al.* figure). Blue solid line: mean `auto.arima()` forecast [Hyndman and Khandakar, 2008] and 50% and 95% prediction intervals (dark blue and grey regions). Blue dots: true outcome.

I would certainly be interested in R_t, but also case projections.

Overall though, both articles are impressive and extensive pieces of work and I once again extend my congratulations to both sets of authors.

Disclaimer: I belong to the same institution as some of the Mishra *et al.* authors and have recently had a minor authorship role on another recently accepted paper. I did not contribute to the review of Mishra *et al.*

References:

Hyndman R. J. and Khandakar, Y. (2008). "Automatic time series forecasting: the forecast package for R.", *Journal of Statistical Software*, **26**, (3), 1-22. `https://www.jstatsoft.org/article/view/v027i03`.