

ORIGINAL ARTICLE

Discussion Paper



Quantifying the economic response to COVID-19 mitigations and death rates via forecasting purchasing managers' indices using generalised network autoregressive models with exogenous variables

Guy P. Nason | James L. Wei

Department of Maths, Imperial College,
London, UK

Correspondence

James L. Wei, Department of Maths,
Imperial College, 180 Queen's Gate,
London SW7 2AZ, UK.
Email: james.wei19@ic.ac.uk

Abstract

Knowledge of the current state of economies, how they respond to COVID-19 mitigations and indicators, and what the future might hold for them is important. We use recently developed generalised network autoregressive (GNAR) models, using trade-determined networks, to model and forecast the Purchasing Managers' Indices for a number of countries. We use networks that link countries where the links themselves, or their weights, are determined by the degree of export trade between the countries. We extend these models to include node-specific time series exogenous variables (GNARX models), using this to incorporate COVID-19 mitigation stringency indices and COVID-19 death rates into our analysis. The highly parsimonious GNAR models considerably outperform vector autoregressive models in terms of mean-squared forecasting error and our GNARX models themselves outperform

[Read before The Royal Statistical Society at the first meeting on 'Statistical Aspects of the Covid-19 Pandemic' held at the Society's 2021 annual conference in Manchester on Wednesday, 8 September 2021, the President, Professor Sylvia Richardson, in the Chair.] This discussion paper is linked to two other papers (<https://doi.org/10.1111/rssa.12874> and <https://doi.org/10.1111/rssb.12453>) published with discussion contributions and authors' replies in JRSSA 185:4.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* published by John Wiley & Sons Ltd on behalf of Royal Statistical Society.

GNAR ones. Further mixed frequency modelling predicts the extent to which that the UK economy will be affected by harsher, weaker or no interventions.

KEYWORDS

forecasting, joint medical and economic modelling, multivariate time series, network time series

1 | INTRODUCTION

The severe impact of COVID-19 on global economic conditions cannot be understated, with UK real gross domestic product (GDP) contracting by a record 20% in the second quarter (Q2) of 2020 alone. Timely estimates of economic activity through the current turbulent period and forecasts of the impact of government virus mitigation strategies will be crucial for evaluating the costs and benefits of different policy interventions.

This article introduces new methods to forecast Purchasing Managers' Indices (PMIs) during the COVID-19 pandemic and quantifies the potential impact of government virus mitigation strategies. PMIs are survey-based leading indicators of the direction of economic activity. We choose to model PMIs as they are used extensively for the nowcasting of slower-moving headline measures of economic activity, such as GDP (e.g. D'Agostino & Schnatz 2012). Our analysis builds on the class of generalised network autoregressive (GNAR) models (Knight et al., 2016, 2020), which model multivariate time series in a highly parsimonious fashion by associating the dependencies between constituent series with an underlying network structure. Hence, GNAR models can characterise a rich diversity of dependence structures, while incurring a lower risk of overfitting when compared to classical vector autoregression (VAR) models, translating into superior forecasting performance. Indeed, our work below demonstrates that GNAR models outperform VAR models by 33% based on mean-squared forecasting error for forecasting PMIs from historical PMI data.

A goal here is to model PMIs in terms of historical PMIs, PMIs from neighbouring countries via our network model and (i) stringency measures of government intervention policies enacted to mitigate the spread of COVID-19 and (ii) the monthly number of confirmed COVID-19 deaths per million. Incorporating information about past PMIs from neighbouring countries via a network structure improves forecasting as it is extra relevant information. Our network model uses countries' PMIs on the network nodes (vertices) and the strength of link between two countries (nodes) increases with increasing quantities of exports between the two. Figure 1 (below) shows an example of one of the networks we use. In addition, the incorporation of multiple exogenous time-dependent node-specific regressors is both new to GNAR modelling and practically essential for properly understanding the potential influence of (i) and (ii). Our new extension is named GNARX, where X indicates incorporation of fully time-dependent exogenous variables, which is in line with similar naming conventions, such as VAR to VARX.

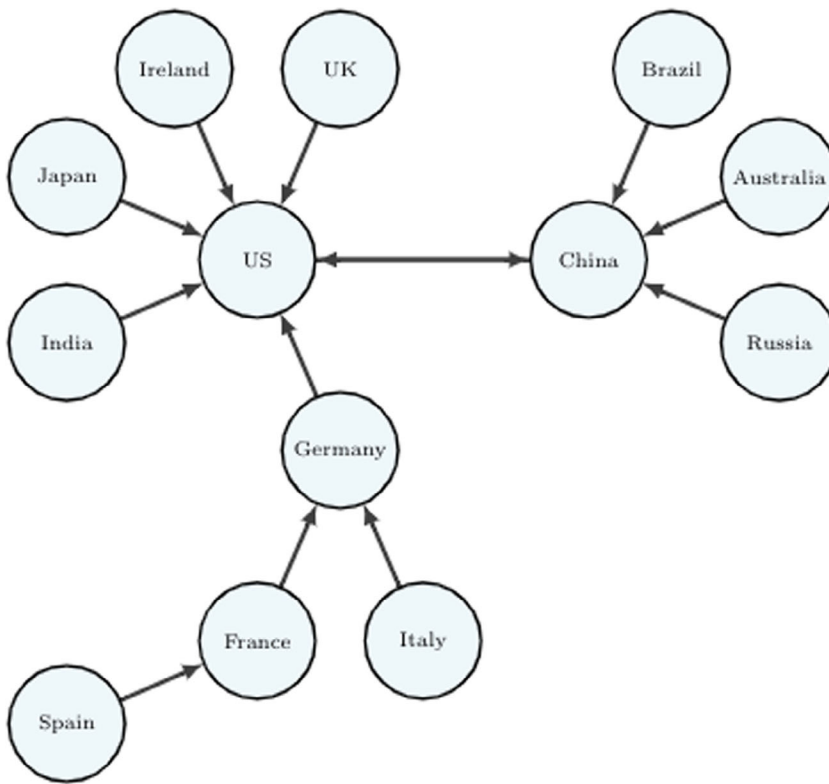


FIGURE 1 Nearest neighbour network based on 2019 export data. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

We structure our paper as follows. Section 2 introduces the new GNARX specification and characterises it as a specially constrained linear regression problem. Asymptotic properties of the corresponding estimators follow from this. Section 3 provides details of our multivariate time series and network data and discusses two experiments: the first compares out-of-sample forecasting performance of GNAR models (i.e. without using external regressors) with a VAR model, and the second evaluates the GNARX model estimates when the stringency indices and COVID-19 death rates are incorporated. Section 4 uses our new GNARX modelling ability to analyse UK economic conditions under three different scenarios pertaining to either tightening, easing or unchanged intervention policies over the next 6 months. Section 5 concludes with a brief discussion of the experimental results and avenues for future research.

2 | THE GNARX MATHEMATICAL MODEL AND NOTATION

Definition. Let our multivariate real-valued time series $\{Y_{i,t}\}$ be observations collected on variable $i \in \mathcal{K} = \{1, \dots, N\}$ at time $t = 1, \dots, T \in \mathbb{N}$, where $N \in \mathbb{N}$ is the number of variables. We sometimes write the multivariate series as the vector time series $\{\mathbf{Y}_t\} = (Y_{1,t}, \dots, Y_{N,t})^\top$, where $\top$ denotes transpose, for $t = 1, \dots, T$. Now let $\mathcal{G}$ be a graph consisting of a set of vertices/nodes, $\mathcal{K}$, and an edge set denoted by $\mathcal{E}$. Let $i \leftrightarrow j \in \mathcal{E}$ denote the presence of an undirected edge between nodes i and j and $i \rightsquigarrow j \in \mathcal{E}$ denote the existence of a directed edge from i to j , for $i, j \in \mathcal{K}$. For network time series a graph vertex $i \in \mathcal{K}$ corresponds to time series variable $\{Y_{i,t}\}$. Let $\{X_{h,i,t}\}$ be

the h -th real-valued stationary exogenous node-specific regressor series for node i at time t , for $h = 1, \dots, H \in \mathbb{N}$, the number of such regressors.

We introduce the GNARX $(p, \mathbf{s}, \mathbf{p}')$ model, where $(p, \mathbf{s}, \mathbf{p}') \in \mathbb{N} \times \mathbb{N}_0^p \times \mathbb{N}_0^H$, by

$$Y_{i,t} = \sum_{j=1}^p \left(\alpha_{ij} Y_{i,t-j} + \sum_{r=1}^{s_j} \beta_{j,r} \sum_{q \in \mathcal{N}^{(r)}(i)} \omega_{i,q} Y_{q,t-j} \right) + \sum_{h=1}^H \sum_{j'=0}^{p'_h} \lambda_{h,j'} X_{h,i,t-j'} + u_{i,t}, \quad (1)$$

where the parameters $\alpha_{ij}, \beta_{j,r}, \lambda_{h,j'} \in \mathbb{R}$ for all i, j, r, h, j', p is the autoregressive order of the model and also maximum order of neighbour time lags, p'_h (the h th element of $\mathbf{p}'$) denotes the maximum lag of the h th exogenous regressor involved in the model and s_j (the j th element of $\mathbf{s}$, which we sometimes write as $[s]$) is the maximum stage of neighbour dependence for time lag j . For example, $s_1 = 2$ implies that nodes are dependent on the first lags of its first and second stage neighbours in the network $\mathcal{G}$. The set $\mathcal{N}^{(r)}(i)$ is the r th stage neighbourhood set of node $i \in \mathcal{G}$ defined by

$$\mathcal{N}^{(r)}(i) = \mathcal{N} \{ \mathcal{N}^{(r-1)}(i) \} / [\{ \cup_{q=1}^{r-1} \mathcal{N}^{(q)}(i) \} \cup \{ i \}], \quad (2)$$

for $r = 2, 3, \dots$, $\mathcal{N}^{(1)} = \mathcal{N}$, where $\mathcal{N}(A) = \{ j \in \mathcal{K}/A : i \rightsquigarrow j; i \in A \}$. The $\omega_{i,q} \in [0, 1]$ in Equation (1) are network connection weights, typically set as the inverse of some prior notion of distance between nodes, as in Knight et al. (2020) using a similar approach to that used in the lifting scheme described by Jansen et al. (2009). Sometimes we consider the global- α model where $\alpha_{ij} = \alpha_j$, that is, use the same $\{\alpha_j\}_{j=1}^p$ sequence for each node. Consequently, the local- α model is where the α_{ij} sequence is different for each node i .

The GNARX model cannot be captured by the standard GNAR model of Knight et al. (2020) by incorporating the exogenous regressors into $\mathbf{Y}_t$, as p and p'_h may differ. The network VAR model proposed by Zhu et al. (2017) does incorporate dependence on node-specific covariates, but unlike model (1), they do not vary with time and have limitations on other parameters/quantities in the network. Full specifications for both of these models can be found in the supplementary material section 1.

GNARX models share the same advantages with respect to missing data as GNAR models, especially when compared to VAR-based approaches. This is because, for example, VAR models try to estimate parameters that link one specific variable to another. If, for example, one of those variables disappears for a significant block of time, then it can be difficult to estimate the parameter associated with any of its pairings. By contrast, in a GNARX model, a variable pair is just a nearest neighbour and every such pair in the data contributes to the 'nearest neighbour' parameter, one of the $\beta_{j,r}$. So, even if one node suffers from significant missingness, there will be usually many other pairs so that the respective parameter is still well-estimated. More details on this, particularly on how this works for r th-stage neighbours, can be found in section 2.6 of Knight et al. (2020).

GNARX Parsimony. The GNARX $(p, \mathbf{s}, \mathbf{p}')$ model with a single exogenous regressor ($H = 1$) contains a total of $M = Np + \sum_{j=1}^p s_j + p' + 1$ parameters. By comparison, the corresponding VAR model with exogenous variables (VARX) containing p lags of the modelled series $\{\mathbf{Y}_t\}$ and p' lags of the exogenous series $\{\mathbf{X}_t\} = \{(X_{1,t}, \dots, X_{N,t})^\top\}$, can be given in the reduced form by

$$\mathbf{Y}_t = \phi_1 \mathbf{Y}_{t-1} + \dots + \phi_p \mathbf{Y}_{t-p} + \Lambda_0 \mathbf{X}_t + \Lambda_1 \mathbf{X}_{t-1} + \dots + \Lambda_{p'} \mathbf{X}_{t-p'} + \mathbf{u}_t, \quad (3)$$

which contains $P = N^2(p + p' + 1)$ parameters. The parsimonious nature of GNARX compared to VARX is clear when parameter growth is $\mathcal{O}(N)$ for the former and $\mathcal{O}(N^2)$ for the latter. Indeed, it has been suggested by Bernanke et al. (2005) that VARs are not often used for macroeconomic datasets that contain more than six to eight time series. The parsimony of GNARX, combined with the flexibility to model dynamic networks representing multiple covariates, ensures our methodology is readily applicable to many applications involving the forecasting of multivariate time series.

GNARX Stationarity. Under the assumption of stationarity of $\{X_{h,i,t}\}$ for all h, i and parameter constraints

$$\sum_{j=1}^p \left(|\alpha_{i,j}| + \sum_{c=1}^C \sum_{r=1}^{s_j} |\beta_{j,r,c}| \right) < 1 \quad \forall i \in 1, \dots, N, \quad (4)$$

the GNARX process $\{\mathbf{Y}_t\}$ is stationary. We prove this in section 2 of the Supplementary Material. For simplicity, we will assume $H = 1$ for the remainder of this section only, replacing $X_{h,i,t}$ with $X_{i,t}$ and $\mathbf{p}'$ with p' . Extending the following results to multiple external node-specific regressors is straightforward.

Precise model definition and theoretical statistical properties. With the $H = 1$ assumption, the GNARX $(p, \mathbf{s}, p')$ model can be expressed as a VAR of the form (3), but where coefficient matrices are restricted so that $\phi_k = \text{diag}\{\alpha_{i,k}\} + \sum_{r=1}^{s_k} \beta_{k,r} W^{(r)}$, $[W^{(r)}]_{l,m} = w_{l,m} \mathbb{I}\{m \in \mathcal{N}^{(r)}(l)\}$, and $\Lambda_k = \lambda_k I_N$. In matrix form, the model can be written as $Y = BZ + U$, where $Y = [\mathbf{Y}_{p^*+1}, \dots, \mathbf{Y}_T]$, $p^* = \max(p, p')$, $B = [\phi_1, \dots, \phi_p, \Lambda_0, \Lambda_1, \dots, \Lambda_{p'}]$, $Z = [\mathbf{Z}_{p^*}, \dots, \mathbf{Z}_{T-1}]$, $\mathbf{Z}_t^\top = [\mathbf{Y}_t^\top, \dots, \mathbf{Y}_{t-p+1}^\top, \mathbf{X}_{t+1}^\top, \dots, \mathbf{X}_{t-p'+1}^\top]$, and $U = [\mathbf{u}_{p^*+1}, \dots, \mathbf{u}_T]$. Additional regressors can be incorporated by concatenating the corresponding coefficient and data matrices to B and $\mathbf{Z}_t$ respectively.

Following Lütkepohl (2005), the coefficient restrictions on B that result from the GNARX assumptions are $\text{vec}(B) = R\gamma$, where R is the $P \times M$ model matrix, defined in the supplementary material section 4, $\gamma = (\gamma_1^\top, \dots, \gamma_p^\top, \lambda_0, \dots, \lambda_{p'})^\top$, with GNAR parameters that we are interested in given by $\gamma_k = (\alpha_{1,k}, \dots, \alpha_{N,k}, \beta_{k,1}, \dots, \beta_{k,s_k})^\top$ and all elementary vectors are of length N . With the global alpha assumption, $\alpha_{i,k} = \alpha_k$ for all $i = 1, \dots, N$, we have $\gamma_k = (\alpha_k, \beta_{k,1}, \dots, \beta_{k,s_k})^\top$ and $A = L$.

Given the VAR matrix representation and linear coefficient restrictions from above, Lütkepohl (2005) proposes the feasible generalised least squares (FGLS) GNAR parameter estimator to be

$$\hat{\gamma}^{\text{FGLS}} = \left\{ R^\top \left(ZZ^\top \otimes \hat{\Sigma}_u^{-1} \right) R \right\}^{-1} R^\top \left(Z \otimes \hat{\Sigma}_u^{-1} \right) \text{vec}(Y), \quad (5)$$

where the maximum likelihood estimator of Σ_u is

$$\hat{\Sigma}_u = T^{-1} (Y - \hat{B}Z)(Y - \hat{B}Z)^\top, \quad (6)$$

and where $\hat{B}$ satisfies $\text{vec}(\hat{B}) = R\hat{\gamma}$ and $\hat{\gamma}$ is the ordinary least squares estimator of γ . In the case of missing observations, the corresponding residuals are set to zero. Under the conditions of stationarity of $\{\mathbf{Y}_t\}_{t=1, \dots, T}$ and $\{\mathbf{X}_t\}_{t=1, \dots, T}$, consistency and asymptotic normality are proved in section 2 of the supplementary material.

3 | PURCHASING MANAGERS' INDICES: THE DATA AND FORECASTING EXPERIMENTS

3.1 | Input data

3.1.1 | Purchasing managers' indices

IHS Markit Composite PMIs are computed by aggregating the views of senior purchasing executives from around 400 companies on a wide range of economic variables pertinent to the manufacturing and services sectors. Composite PMIs can take values between zero and 100, with a reading above 50 indicating economic expansion. All PMI series have been seasonally adjusted using the X-12 ARIMA method; full details of the survey methodology are available from <https://ihsmarkit.com/products/pmi.html>. Currently, monthly composite PMI data are available for 13 major economies, which are all included in our analyses. The longest series are from the United Kingdom, Italy and Germany, which start from Jan-98, and the shortest is from Australia, which starts from May-16.

PMI series are diffusion indices, which treat a fixed set of component economic indicators effectively as dummy variables that represent whether the trend in the component is positive or negative. For PMIs, these component series are the survey responses of individual purchasing managers. It is therefore reasonable to assume that diffusion indices are stationary, given the well-documented cyclical nature of macroeconomic time series, see Zarnowitz (1991). Moreover, Augmented Dickey–Fuller (ADF) unit root tests on the in-sample period between Jan-98 and Dec-17 result in p -values above 0.1 for all series, indicating stationarity for the time being, against the alternative hypothesis of first-order non-stationarity. Broadly, the series also do not reject the null of second-order stationarity according to the wavelet-based tests of Nason (2013, 2016).

3.1.2 | Network construction

GNAR and GNARX models require us to specify a network structure to associate with the PMI dataset. The time series for a particular country's PMI corresponds to a single vertex in the network—one vertex for each country. Rather than attempting to learn the network structure indirectly from the time series as in Leeming (2019) or Knight et al. (2020), we suggest that a construction based on the flow of exports between countries that provides a reasonable reflection of the association between component series. This rests on the assumption that a country's business confidence is partly dependent on the level of economic activity of its major export partners. Since (a) model orders are chosen by BIC optimisation, (b) network parameters are highly significant and (c) forecast performance is improved when compared to models that ignore network information, these all suggest that our choice of networks is reasonable. Evaluating performance by forecast accuracy is honest and exacting.

Our experiments evaluate two different networks:

1. **Fully connected trade network.** Each country is connected to every other country. Edge weight w_{ij} is set to be the export quantity between countries i and j . The weights are then normalised so that $\sum_j w_{ij} = 1$ for all i . In this case, $s_j \in \{0, 1\}$ for all j .
2. **Nearest neighbour trade network.** (Figure 1) Each country has a single outgoing edge to the country that it exports the most too.

Export data were obtained from the UN Comtrade International Trade Statistics Database (<https://comtrade.un.org/data/>), using the latest available observations.

3.1.3 | Exogenous variables

The GNARX model's advantage is that it permits the multivariate time series $\mathbf{Y}_t$ to depend on vertex-specific external time series regressors $\mathbf{X}_i$; two series in our examples. The first, $X_{1,i,t}$, are differenced country stringency indices (Hale et al., 2020) for country i , which aim to quantify the severity of COVID-19 mitigation policies. These are composite indices that take values between zero and 100, with higher values reflecting stricter measures. The stringency indices are mainly related to containment policies, such as the closing of schools, workplaces and public events. The full list of constituent indicators is available from <https://github.com/OxCGR/covid-policy-tracker/blob/master/documentation>. Values of the stringency indices before the COVID-19 outbreak are automatically set to zero, rather than treated as missing.

The second regressor, $X_{2,i,t}$, is the first difference of country i new COVID-19 death rates, which may impact economic conditions via channels beyond the previously discussed interventions. For example, consumers may extrapolate a rise in the confirmed number of COVID-19 deaths into the future, leading to consumer precautionary saving in anticipation of greater future lockdown measures. As higher death rates would likely also be associated with stricter mitigation policies, omitting death rates would likely lead to model errors being correlated with the stringency index regressor, resulting in inconsistent parameter estimates. Our experiments use the daily number of confirmed COVID-19 deaths per million in a given country, averaged for each month (Roser et al., 2020), available from <https://ourworldindata.org/covid-deaths>. The sample correlation between the differenced stringency indices and the differenced COVID-19 death rate of the same country is only 0.46, suggesting that multicollinearity is not a significant problem. The correlation only rises modestly to 0.57 for the non-differenced series.

Figure 2 shows the UK PMI and the two exogenous variables starting from January 2019, where all series have been standardised to have zero mean and unit variance so that they can be compared on the same axes. The most prominent feature of the UK PMI series is the large six and eight standard deviation declines during March and April 2020. This fall in purchasing manager confidence coincides with the first UK lockdown (rise in stringency index), but also reflects broader worsening of global macroeconomic conditions caused by the COVID-19 pandemic (e.g. International Monetary Fund, 2020).

3.2 | GNAR forecasting performance

Here, model order is selected to minimise the Bayesian information criterion (BIC) subject to an order limit of $p = 12$. For computational feasibility, we adopt a stagewise approach, where p is first selected to minimise BIC, followed by s_j for $j = 1, \dots, p$. This allows us to capture network effects lasting up to a year in duration. For the exogenous regressors in later experiments, we instead impose a lower order limit of three months, due to the lack of nonzero observations. In that case, p'_h for $h = 1, \dots, H$ are optimised at the same time as s_j . All PMI observations up to Dec-17 are treated as the in-sample period for model order selection and estimation of both GNAR and VAR models. We refer readers to the supplementary material section 7 for the results if model orders were instead selected to minimise forecasting errors in a subset of the in-sample period. Finally, we provide simulation evidence for the consistency of the model selection procedure using

stagewise BIC optimisation for GNARX in the supplementary material section 8, which shows similar performance to ‘global search’ BIC.

Unlike GNAR, VAR models do not easily handle missing values. Specifically, GNAR models deal with missingness by modifying connection weights (Knight et al., 2020, section 2.6) and because of this advantage our GNAR model can make use of the *entire* unbalanced panel starting from Jan-98. If Australia were included, then the in-sample period for the VAR model would contain only 20 months and, for this reason, we exclude Australia from this analysis.

Table 1 shows the composite PMI forecasting performance of four different stagewise-BIC selected GNAR models, a VAR model, a naive forecast (predicting next from current) and univariate AR models. The latter involve a separate AR model estimated for each country, with order determined by BIC optimisation. For all but three countries, AR(1) was selected. However, for fair comparison with the GNAR and VAR models where autoregressive order is the same for every series, AR(1) was used for all countries. It is worth noting that AR(1) is identical to GNAR(1,[0]) with local- α . We have used mean-squared forecast error (MSFE) in our article, which computes the empirical average of the squared differences between our point forecasts and the eventual out-turn. The lowest forecast error is attained by the local- α GNAR model with the fully connected network.

Furthermore, the MSFE of all four GNAR models are substantially lower than that of the VAR model, indicating overfitting in the latter. The local- α GNAR(1,[1]) model contains only 13 parameters, while the VAR(2) model contains $2N^2 = 288$ parameters. Our results indicate that when a simple model is appropriate, BIC optimisation for VAR may still lead to overfitting, but the same does not hold for GNAR models. Indeed, only three of the four GNAR models outperformed the naive forecast and one of the two local- α GNAR models outperformed the univariate AR models, reflecting the difficulty in forecasting economic indicators using only past values. On the other hand, the best GNAR model *does* improve the forecast performance and a few percent

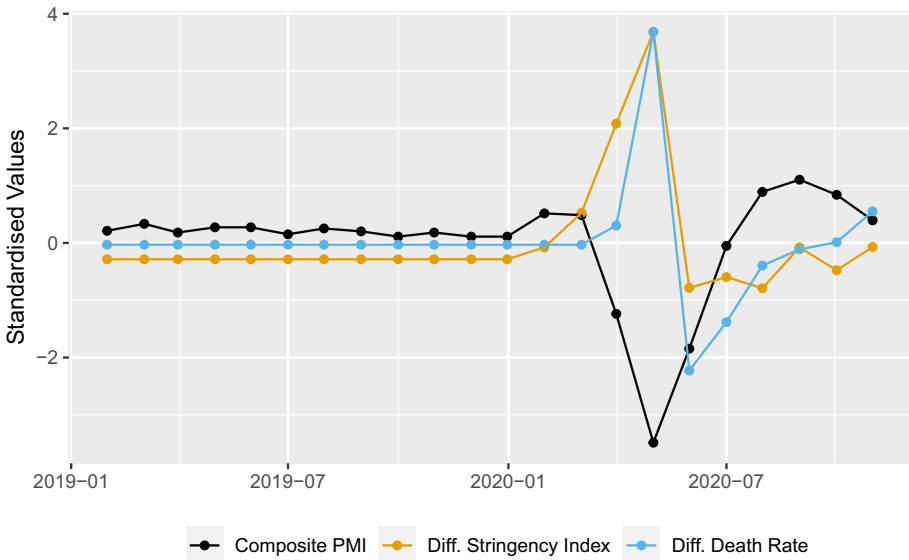


FIGURE 2 Plot of UK Composite Purchasing Managers’ Indices together with UK stringency index and death rate covariates, where all variables have been standardised to have zero mean and unit variance. [Colour figure can be viewed at wileyonlinelibrary.com]

improvement can translate into substantial economic value when one is dealing with GDP-like magnitudes, see Section 4.

Table 2 shows estimates for the local- α GNAR(1,[1]) model, with p -values computed using HC2 robust standard errors, which are heteroscedasticity-consistent estimators first proposed by MacKinnon and White (1985) that adjust residuals using the projection matrix leverage values. These were calculated with the **sandwich** package of Zeileis (2004). In this experiment, the use of robust standard errors produces more conservative p -values.

Figure 3 plots the UK out-of-sample one-step-ahead rolling forecasts for this best-performing model, along with those from the VAR(2) model (see supplementary material, section 6, for other countries). The VAR forecasts appear to underestimate the initial PMI decline at the start of the crisis period, but substantially overestimate the recovery. This leads to substantially higher MSFE during the COVID-19 crisis period starting from Feb-20, relative to GNAR.

Overall, BIC choice selects for higher autoregressive order in the global- α GNAR models compared to the local- α variant, which also results in higher out-of-sample MSFE. This suggests that the PMIs' autoregressive dynamics differs between countries, although the autoregressive parameters in Table 2 are similar. Moreover, BIC choice of the local- α models also favours exploiting the network structure implied by export flows, whether by using the fully connected or the nearest neighbour network: $\mathbf{s} \neq \mathbf{0}_p$ for both of the best-performing GNAR models.

TABLE 1 Out-of-sample composite Purchasing Managers' Indices (PMIs) forecast performance between Jan-18 and Oct-20 using GNAR, VAR and naive forecasting. MSFE = Out-of-sample mean-squared forecasting error; SE = standard error

Model	Model order	Parameters	MSFE (SE)
Global- α GNAR, fully connected net	2, [1, 1]	4	41.9 (8.2)
Global- α GNAR, nearest-neighbour net	2, [1, 1]	4	40.7 (7.4)
Local- α GNAR, fully connected net	1, [1]	13	38.4 (6.8)
Local- α GNAR, nearest-neighbour net	1, [1]	13	39.3 (7.1)
VAR	2	288	57.3 (9.3)
AR	1	12	39.1 (7.0)
Naive forecast	—	—	40.6 (7.2)

TABLE 2 Local- α GNAR(1,[1]) parameter estimates (to two decimal places) using the fully connected network. Parameter p -values were all $<0.001\%$

Parameter	Estimate	Parameter	Estimate	Parameter	Estimate
$\alpha_{USA,1}$	0.93	$\alpha_{Japan,1}$	0.87	$\alpha_{Germany,1}$	0.90
$\alpha_{UK,1}$	0.90	$\alpha_{Italy,1}$	0.89	$\alpha_{China,1}$	0.88
$\alpha_{Spain,1}$	0.92	$\alpha_{Ireland,1}$	0.92	$\alpha_{Brazil,1}$	0.88
$\alpha_{Russia,1}$	0.92	$\alpha_{India,1}$	0.91	$\alpha_{France,1}$	0.92
$\beta_{1,1}$	0.07				

3.3 | GNARX forecasting performance with external regressors

Given the shortness of the exogenous regressors, it is not feasible now to evaluate out-of-sample forecasting performance for the GNARX models, so we switch to in-sample forecast evaluation. For example, predicting ‘October’ based on exogenous variables that only existed since early Spring, that is, few observations, is not sensible. Here, we estimate model parameters on the whole series and then, given a time point, we use the estimates, but use observations only from previous time points. In future, with more data, we will be able to switch to out-of-sample forecasting performance evaluation.

We can now reintroduce the shorter Australian PMI series, using the same method for handling missing observations as in Section 3.2. On our relatively short time series with missing data, VARs with exogenous variables (VARX) as described in, for example, Tsay (2013), which contain $N^2(p + p' + 1)$ parameters, are infeasible. Recall that some VAR models may be estimated as separate linear regressions for each country. A simple VARX(1,0) model would involve the estimation of 13 coefficients corresponding to the stringency indices and 13 corresponding to the death rates in each country. However, our dataset contains fewer than 13 nonzero observations for each exogenous variable of each country, clearly leading to a parameter identification problem.

Table 3 shows in-sample mean-squared forecasting errors for our stagewise-BIC selected GNARX models using the fully connected network (results for the nearest neighbour network and no network were always worse and reported in the supplementary material, section 11). The inclusion of exogenous variables results in substantially lower BIC and an almost halving of the MSFE across all settings compared to GNAR without external regressors. The lowest BIC model is the global- α GNARX(5,[1,0,1,0,1],3,2), with estimates shown in Table 4 and p -values calculated as in Table 1.

Table 4 shows that the first and fifth positive autoregressive coefficients are significant. As before, current-period PMIs are also significantly positively dependent on previous-period PMIs of neighbours weighted by export volumes. On the other hand, the negative $\beta_{3,1}$ and $\beta_{5,1}$ parameter estimates are smaller in magnitude.

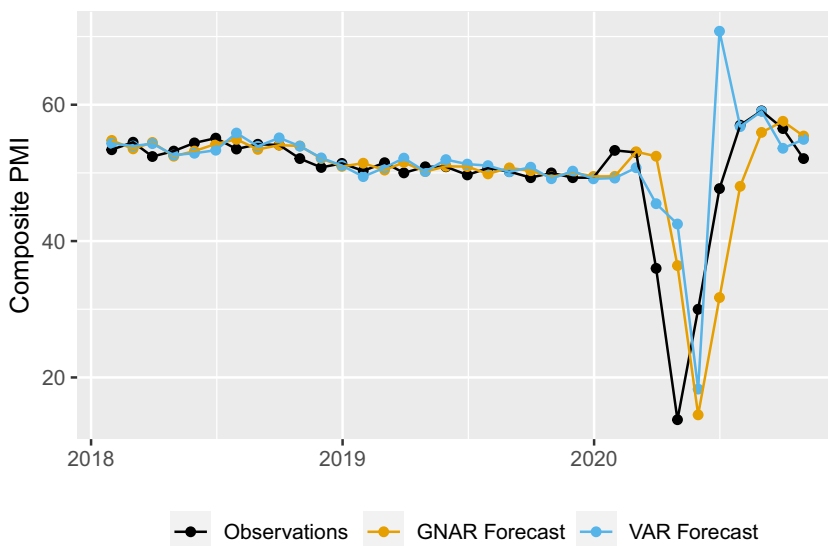


FIGURE 3 Out-of-sample one-step-ahead rolling composite UK Purchasing Managers’ Indices forecasts using the local- α GNAR(1,[1]) and the VAR(2) models. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

TABLE 3 In-sample composite Purchasing Managers' Indices (PMIs) forecast performance between Jan-98 and Oct-20 using GNARX models with the fully connected network. MSFE = In-sample mean-squared forecasting error

Model	Model order	BIC	MSFE (SE)
<i>Without external regressors</i>			
Global- α GNAR	4, [1, 1, 1, 0]	12653	7.1 (0.9)
Local- α GNAR	3, [1, 1, 1]	12806	6.7 (0.9)
<i>Stringency indices only</i>			
Global- α GNARX	5, [1, 0, 1, 0, 1], 3	11323	4.2 (0.3)
Local- α GNARX	2, [1, 1], 3	11331	4.0 (0.3)
<i>COVID-19 death rates only</i>			
Global- α GNARX	4, [1, 1, 1, 0], 3	12020	5.5 (0.8)
Local- α GNARX	2, [1, 1], 3	12212	5.5 (0.8)
<i>With both external regressors</i>			
Global- α GNARX	5, [1, 0, 1, 0, 1], 3, 2	11175	4.0 (0.3)
Local- α GNARX	2, [1, 1], 3, 1	11199	3.8 (0.3)

TABLE 4 Global- α GNARX(5,[1, 0, 1, 0, 1],3, 2) model parameter estimates using the fully connected network and both exogenous regressors. † means $p < 0.001$. Values significant at the 5% level in bold

Parameter	Estimate	<i>p</i> -value %	Parameter	Estimate	<i>p</i> -value %	Parameter	Estimate	<i>p</i> -value %
α_1	0.83	†	$\beta_{1,1}$	0.28	†	$\lambda_{1,2}$	0.15	0.8
α_2	0.00	97.5	$\beta_{3,1}$	−0.15	4.2	$\lambda_{1,3}$	0.10	3.3
α_3	0.03	59.2	$\beta_{5,1}$	−0.08	3.5	$\lambda_{2,0}$	−0.91	0.2
α_4	−0.05	34.0	$\lambda_{1,0}$	−0.31	†	$\lambda_{2,1}$	0.42	21.3
α_5	0.09	0.6	$\lambda_{1,1}$	−0.09	24.8	$\lambda_{2,2}$	0.41	11.3

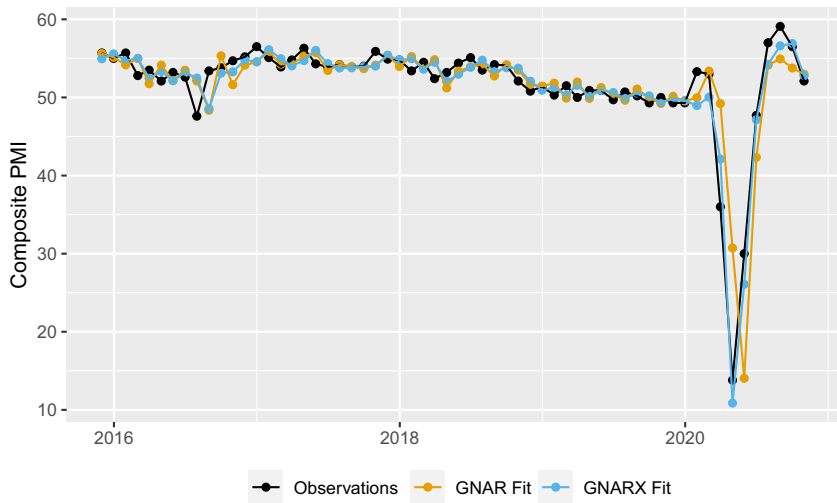


FIGURE 4 UK: Comparison of GNARX(5,[1, 0, 1, 0, 1]) and GNARX(5,[1,0,1,0, 1], 3, 2) fits. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

The zero-lag coefficient for COVID-19 intervention stringency indices is highly significant, matching our expectation of a strong coincident negative impact of containment policies on business confidence. A country’s COVID-19 death rate also has a significant negative contemporaneous impact on the PMI, indicating that the pandemic might be influencing business confidence through channels beyond policy interventions.

Figure 4 shows the global- α GNARX(5,[1,0,1,0,1],3,2) UK fits over the most recent 5-year period, to illustrate the quality of fit during the COVID-19 period (other countries’ charts are given in the supplementary material, section 10). The inclusion of stringency indices and COVID-19 death rates allows the GNARX model to anticipate both the initial composite PMI decline at the start of the crisis and the subsequent recovery, leading to a much lower mean-squared forecasting error than the corresponding GNAR model.

4 | UK SCENARIO ANALYSIS

We now use the fitted the global- α GNARX(5,[1,0,1,0,1],3,2) model from the previous section to produce 6-month forecasts of the UK composite PMI series under different assumptions on the evolution of the UK stringency index. For other examples of conditional forecasting on multivariate economic time series, we refer readers to the seminal paper by Waggoner and Zha (1999) and the more recent summary provided by Bańbura et al. (2015), which describe conditional forecasting techniques and applications for VAR models. We would note that these describe more difficult problem than ours, as we condition on exogenous variables. Figure 5 provides forecasts of the UK composite PMI under three different policy intervention scenarios for the 6-month period from Nov-20 to Apr-21, assuming no change in the stringency indices of other countries from Aug-20 or in COVID-19 death rates:

- 1. **Easing COVID-19 intervention measures.** The stringency index decreases linearly from the Oct-20 value of 67.9 to 0.0 between Nov-20 and Apr-21.

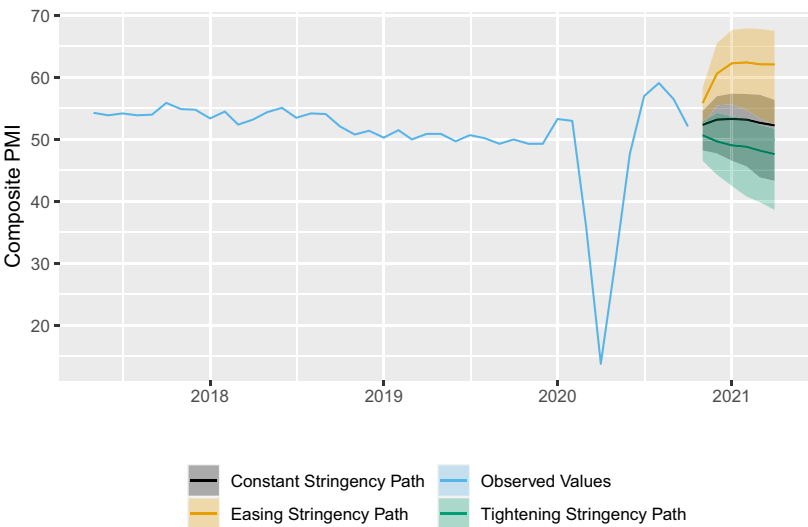


FIGURE 5 Forecasts of UK composite Purchasing Managers’ Indices under three different scenarios for policy interventions using the GNARX(5,[1,0,1,0,1],3,2) model. 95%-coverage prediction intervals estimated using 1000 bootstrap samples. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 5 Forecasts for UK quarterly real gross domestic product growth (%) based on Purchasing Managers’ Indices (PMIs) forecasts for the tightening, constant and easing intervention measure scenarios. 95% prediction intervals in brackets

Quarter	Tightening	Constant	Easing
2020 Q4	−0.1 (−4.1, 3.8)	−0.1 (−4.1, 3.8)	−0.1 (−4.1, 3.8)
2021 Q1	−1.2 (−5.8, 2.9)	0.3 (−4.1, 4.4)	3.3 (−1.3, 7.4)
2021 Q2	−1.6 (−7.1, 2.2)	−0.1 (−5.7, 3.2)	3.2 (−2.0, 7.1)

2. **No change in COVID-19 intervention measures.** The stringency index is constant at the Oct-20 value of 67.9 between Nov-20 and Apr-21.
3. **Tightening in COVID-19 intervention measures.** The stringency index increases linearly from the Oct-20 value of 67.9 to 100.0 between Nov-20 and Apr-21.

As might be expected, the GNARX model predicts significantly higher UK composite PMI under assumption of the UK stringency index falling to zero, holding death rates constant. Using bootstrap prediction intervals, computed using the algorithm introduced in the supplementary material section 12, the expected PMI under the easing scenario exceeds the 97.5th percentile of the PMI under the tightening and constant stringency scenarios after only 1 month. The Apr-21 prediction of PMI is 62.1 in the easing scenario (which would be a record high) and 47.6 in the tightening scenario. In all cases, the estimated sampling distribution of the PMI is negatively skewed.

We now estimate how these composite PMI paths would translate into headline quarter-on-quarter real GDP growth, using data sourced from <https://www.ons.gov.uk/economy/grossdomesticproductgdp/timeseries/ihyq/qna>. At the time of writing, 2020 Q3 is the latest available observation of UK GDP. Given that the GDP series is sampled quarterly, while the PMI series is sampled monthly, we assume a mixed-frequency linear relationship between the GDP growth and PMI series, using the MIDAS regression model (Ghysels et al., 2004). Analysis of both the unrestricted MIDAS model and the restricted MIDAS model, using exponential Almon lag polynomial weights, suggest that quarterly GDP is primarily dependent on the ‘2/3’ lag of composite PMI. E.g., 2020 Q3 GDP depends only on the Jul-20 PMI observation in our simplified model.

Table 5 shows conditional forecasts for 2020 Q4, 2021 Q1 and 2021 Q2 UK GDP growth using the PMI predictions produced by GNARX for each of the three COVID-19 intervention scenarios, assuming only the 2/3 lag of composite PMI is relevant. Our results suggest a 4.5% difference in 2021 Q1 GDP growth and a 4.8% difference in 2021 Q2 GDP growth between the easing and tightening scenarios. Table 5’s bootstrap prediction intervals take into account both the uncertainties in the GDP regression model and the underlying PMI forecasts.

Lending support to the strongly positive GDP growth path in our easing restriction scenario, the most recent (at the time of writing) Bank of England (2021) Monetary Policy Report forecasts a material recovery in UK GDP in the event of an easing in Covid-related restrictions. The authors suggest that the economic impact of such easing would be driven by a substantial increase in consumer spending, and a rise in business investment to a lesser extent as sales rise and uncertainty diminishes.

5 | DISCUSSION

We showed how to apply network time series models to multivariate PMI time series for many countries and how forecast performance could be improved by incorporating network

information. We introduced the GNARX model, which permits exogenous variables to be incorporated into GNAR models for the first time and improves forecasting results further still by using COVID-19 intervention stringency indices and COVID-19 death rates. Parameter estimates suggest a highly significant negative impact of harsher intervention measures and COVID-19 deaths on composite PMI and suggests that this effect can be transmitted to neighbours through the export channel.

GNAR(X) model order choice using BIC selects highly parsimonious models that result in excellent forecasting performance, considering the inherent difficulties in forecasting economic indicators. On the other hand, the BIC-chosen VAR model appears to overfit, leading to relatively high forecast errors. Future work might compare the GNAR forecasting performance to that of BVAR models (Giannone et al., 2015).

Our network was constructed by considering what looked 'sensible' rather than optimality with respect to forecasting performance and further development work is required. Some preliminary ideas have been suggested relating to network construction in different fields, see Leeming (2019); Knight et al. (2020) and references therein.

Our three different scenarios earlier all operated under the simplifying assumption that policy interventions would be modified independently of the COVID-19 death rate. Future work may relax this unrealistic assumption, and incorporate other variables to capture attempts by governments to mitigate the negative impact of containment measures through fiscal stimulus.

We have used mean-squared forecasting error throughout our work as it is a valid measure, well-understood and meaningful. There are, of course, many other forecast quality metrics that could be used. For example, the actual proportion of forecasts whose out-turn performed well with respect to (or within) previously established forecasting intervals, see Ehm et al. (2016) for more examples and discussion.

Network time series modelling is young and much remains unexplored. For GNARX models, an obvious extension might permit exogenous regressor coefficients to vary by node. Regarding the COVID-19 experiments discussed above, this would allow economies to react differently to a given stringency level of policy interventions. Intuitively, economies where a larger proportion of employees are able to work remotely should be less vulnerable to strict lockdown measures (Papanikolaou & Schmidt, 2020).

We are also conducting research where more than one network time series interacts. So, for example, one can already see suggestions in the media on how countries' lockdown policies influence each other. For example, the second UK lockdown on 5 November was made politically easier because of recent lockdowns in major European neighbours. Hence, the stringency index time series (and the COVID-19 deaths) could also be represented by network time series and there seems to be no reason why the PMI and stringency index networks would have to be the same. Finally, further extensions to network time series might be considered, such as time-varying parameters that can be used to create a locally stationary network time series process, or permitting edge distances to change and be modelled (e.g. as exports vary between countries over time).

ACKNOWLEDGEMENTS

This article was prepared under the auspices of the Imperial College COVID-19 Response Team (Economics Team). We are grateful to Dr Katharina Hauck, Deputy Director of the Abdul Latif Jameel Institute for Disease and Emergency Analytics (J-IDEA), School of Public Health, Imperial College, London for enthusiastic encouragement and helpful comments on an earlier version of this article.

REFERENCES

- Bañbura, M., Giannone, D. & Lenza, M. (2015) Conditional forecasts and scenario analysis with vector autoregressions for large cross-sections. *International Journal of Forecasting*, 31, 739–756.
- Bank of England. (2021) Monetary Policy Report. Available from: <https://www.bankofengland.co.uk/-/media/boe/files/monetary-policy-report/2021/february/monetary-policy-report-february-2021.pdf> [Accessed February 2021].
- Bernanke, B.S., Boivin, J. & Elias, P. (2005) Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics*, 120, 387–422.
- D'Agostino, A. & Schnatz, B. (2012) Survey-based nowcasting of US growth: a real-time forecast comparison over more than 40 years. Technical Report 1455, ECB, Frankfurt, Germany.
- Ehm, W., Gneiting, T., Jordan, A. & Krüger, F. (2016) Of quantiles and expectiles: consistent scoring functions, Choquet representations and forecast rankings. *Journal of the Royal Statistical Society: Series B*, 78, 505–562.
- Ghysels, E., Santa-Clara, P. & Valkanov, R. (2004) The MIDAS touch: mixed data sampling regression models. Technical Report qt9mf223rs, UCLA, Anderson GSM.
- Giannone, D., Lenza, M. & Primiceri, G.E. (2015) Prior selection for vector autoregressions. *Review of Economics and Statistics*, 97, 436–451.
- Hale, T., Webster, S., Petherick, A., Phillips, T. & Kira, B. (2020) Oxford COVID-19 Government Response Tracker. Available from: <https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker>
- International Monetary Fund. (2020) World economic outlook: a long and difficult ascent. October.
- Jansen, M., Nason, G.P. & Silverman, B.W. (2009) Multiscale methods for data on graphs and irregular multidimensional situations. *Journal of the Royal Statistical Society: Series B*, 71, 97–125.
- Knight, M.I., Nunes, M.A. & Nason, G.P. (2016) Modelling, detrending and decorrelation of network time series. *ArXiv e-prints*, 1603.03221.
- Knight, M., Leeming, K., Nason, G. & Nunes, M. (2020) Generalised network autoregressive processes and the GNAR package. *Journal of Statistical Software*, 96, 1–36.
- Leeming, K.A. (2019) New methods in time series analysis: univariate testing & network autoregression modelling. Ph.D. thesis, School of Maths, University Bristol, UK.
- Lütkepohl, H. (2005) *New introduction to multiple time series analysis*. Berlin: Springer Science & Business Media.
- MacKinnon, J.G. & White, H. (1985) Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29, 305–325.
- Nason, G.P. (2013) A test for second-order stationarity and approximate confidence intervals for localized autocovariances for locally stationary time series. *Journal of the Royal Statistical Society: Series B*, 75, 879–904.
- Nason, G.P. (2016) Locits: tests of stationarity and localized autocovariance, R package version 1.7.3.
- Papanikolaou, D. & Schmidt, L.D.W. (2020) Working remotely and the supply-side impact of Covid-19. Technical Report 27330, NBER.
- Roser, M., Ritchie, H., Ortiz-Ospina, E. & Hasell, J. (2020) Coronavirus pandemic (COVID-19). Available from: <https://ourworldindata.org/coronavirus>
- Tsay, R.S. (2013) *Multivariate time series analysis: with R and financial applications*. Hoboken, N.J.: Wiley.
- Waggoner, D.F. & Zha, T. (1999) Conditional forecasts in dynamic multivariate models. *The Review of Economics and Statistics*, 81, 639–651.
- Zarnowitz, V. (1991) What is a business cycle? Technical Report 3863, NBER.
- Zeileis, A. (2004) Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software*, 11, Issue 10.
- Zhu, X., Pan, R., Li, G., Liu, Y. & Wang, H. (2017) Network vector autoregression. *Annals of Statistics*, 45, 1096–1123.

How to cite this article: Nason, G.P. & Wei, J.L. (2022) Quantifying the economic response to COVID-19 mitigations and death rates via forecasting purchasing managers' indices using generalised network autoregressive models with exogenous variables. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1778–1792. Available from: <https://doi.org/10.1111/rssa.12875>

ORIGINAL ARTICLE

Discussion Paper



Small data, big time—A retrospect of the first weeks of COVID-19

Qingyuan Zhao 

Statistical Laboratory, University of
Cambridge, Cambridge, UK

Correspondence

Qingyuan Zhao, Statistical Laboratory,
University of Cambridge, Cambridge,
UK. Email: qyzhao@statslab.cam.ac.uk

Abstract

This article reviews some early investigations and research studies in the first weeks of the coronavirus disease 2019 (COVID-19) pandemic from a statistician's perspective. These investigations were based on very small datasets but were momentous in the initial global reactions to the pandemic. The article discusses the initial evidence of high infectiousness of COVID-19 and why that conclusion was not reached faster than in reality. Further reanalyses of some published COVID-19 studies show that the epidemic growth was dramatically underestimated by compartmental models, and the lack of fit could have been clearly identified by simple data visualization. Finally, some lessons for statisticians are discussed.

KEYWORDS

COVID-19, infectious disease modelling, model diagnostics, selection bias

1 | INTRODUCTION

Starting from a regional disease outbreak in Wuhan, China, the coronavirus disease 2019 (COVID-19) rapidly grew into a once-in-a-lifetime pandemic. As of the time of writing (December 2020), the pandemic has already taken at least 1.5 million lives around the world. For anyone living through the COVID-19 pandemic, it is rubbing salt into their wounds to repeat more statistics about the disruptions of lives.

[Read before The Royal Statistical Society at the Society's 2021 annual conference held in Manchester on Wednesday, 8 September 2021, the President, Professor Sylvia Richardson, in the Chair.] This discussion paper is linked to two other papers (<https://doi.org/10.1111/rssa.12875> and <https://doi.org/10.1111/rssb.12453>) published with discussion contributions and authors' replies in JRSSA 185:4.

Similar to many other new human diseases, there was an initial period of confusion. On 31 December 2019, the municipal Health Commission of Wuhan first reported a cluster of pneumonia cases of unknown aetiology. Because of the exponentially growing nature of a new epidemic outbreak, many researchers raced against time in the first weeks of 2020 to get a better understanding of COVID-19 with very limited data. For the same reason, the investigations in this period had a far-reaching impact and shaped the initial reactions to the pandemic around the world. This article aims to review some of these early investigations from a statistician's perspective.

It is important for the reader to keep in mind that this article is retrospective. Researchers always have limited data at the beginning of a new disease outbreak. This is especially the case for the COVID-19 pandemic: it is fair to say that nobody could have predicted all its rapid and dramatic developments. This article will try to be as objective as possible by presenting the information in a chronological order, although some information known by us now may not have been available to the investigators in the first weeks of the pandemic. Therefore, some degree of hindsight bias is perhaps unavoidable. The main goal of this article is to think about the lessons that can be learned from these investigations to help us respond to future crises. Another goal is to help statisticians to recognize their strengths and prompt others to appreciate the importance of statistics in the era of big data.

This retrospect is also inevitably personal because Wuhan is in fact my hometown. Right after the announcement of the initial cluster of pneumonia cases, I started to follow media reports and research studies. Like many other amateur epidemiologists, I tasked myself with analysing COVID-19 data and published a preprint in early February to warn that the spread of COVID-19 in Wuhan was much faster than initially thought (Zhao et al., 2020a). After several unsuccessful attempts to publish that article in a scientific journal, I was extremely distressed to watch how the pandemic was quickly developing. Because it is difficult to completely separate facts from speculations (I will try as best as I can), I ask the reader to treat this article as an opinion piece rather than an objective account of the history.

2 | CAN COVID-19 BE TRANSMITTED FROM HUMAN TO HUMAN?

For COVID-19, the most important question before 20 January 2020 was whether it could be transmitted from human to human and thus poses a major threat. To answer this question, the Chinese Center for Disease Control (CCDC) commissioned three groups of expert doctors and epidemiologists in early and mid-January. The answer to this question is extremely obvious in hindsight, but that was certainly not the case in early January. By reviewing the key events and evidence in this period, this section will try to understand why it was difficult to conclude that COVID-19 was highly infectious.

2.1 | The first warning and some background

Rumours of a novel coronavirus first appeared on Chinese social media on 30 December 2020. The most widespread message was initially posted by Dr. Li Wenliang, who tried to use a group chat to warn fellow doctors who went to the same medical school of a potential disease outbreak (Ni, 2020). Dr. Li wrote that there had been seven confirmed SARS (Severe Acute Respiratory Syndrome) cases in the Huanan seafood market and shared the report of a pathogen filtering test showing high confidence of SARS coronavirus, the pathogen responsible for SARS. The message

was quickly shared outside the original group. The same evening, two documents from the Wuhan Municipal Health Commission also started to circulate on the internet. Wuhan is the capital city of the Hubei province in China and home to more than ten million people. An official statement was released the next day by the Health Commission, who announced 27 cases of viral pneumonia, many with connections to the Huanan seafood market (Wuhan Health Commission, 2020a). The statement was also immediately picked up by a regional office of the World Health Organization (2020).

Before continuing the story, it may be helpful to go through some background information. Coronavirus is a group of RNA viruses that cause diseases in mammals and birds. Not all coronaviruses are lethal to humans; some strains just cause the common cold. Before the COVID-19 outbreak, two lethal coronaviruses—SARS and MERS (or more precisely, SARS-CoV and MERS-CoV, which are the official names of the viruses)—circulated in humans in this century. The 2002–2003 SARS epidemic infected at least 8000 people (mostly in China) and caused at least 774 deaths (World Health Organization, 2015). These figures seem minuscule compared to COVID-19, but SARS indeed had the potential of becoming a major pandemic before being contained in July 2003 after strict public health interventions. MERS first appeared in 2012 and has not led to any major outbreaks or sustained transmission due to its relatively low reproductive number. However, because of the frequent contact between humans and a natural host of MERS (camels), MERS keeps coming back almost every year. Both SARS-CoV and MERS-CoV are believed to have originated in bats, which also host many other SARS-like coronaviruses (Ge et al., 2013).

Importantly, it was quickly discovered that the pathogen of COVID-19 is a novel coronavirus that is now known as SARS-CoV-2. As the name suggests, the pathogen is not the original SARS virus but a novel SARS-like coronavirus. So technically speaking, the information provided by Dr. Li Wenliang was not entirely accurate and Dr. Li and several others were reprimanded by the local police for spreading misinformation. As the epidemic unfolded, the reprimand became immensely controversial and was eventually revoked after Dr. Li tragically died from COVID-19 in early February 2020.

To put the initial reactions to a potentially novel coronavirus in the right context, it is helpful to review the 2002–2003 SARS epidemic briefly. The SARS outbreak first began in the Guangdong province of China in November 2002, but the pathogen was not identified until mid-April 2003 (World Health Organization, 2003a). The Chinese government was criticized, both domestically and internationally, for its slow responses and delayed communications with the WHO (Crampton, 2003). In fact, the Minister of Health and Beijing's mayor were discharged in April 2003 after being accused of cover-ups (Jakes, 2003). The course of the epidemic then changed dramatically. After imposing some intensive public health measures (including school closure, travel quarantine and a purpose-built hospital in Beijing), the SARS epidemic was quickly contained by July.

Some of the initial investigations on COVID-19 were commissioned by the CCDC, which was established in January 2002 and quickly expanded after the SARS epidemic. However, media reports suggest that the CCDC usually only acts as a technical supervisor and lacks administrative authority over the provincial and municipal Health Commissions, which are financed locally (Badian Health News, 2020). In a landmark project after the SARS epidemic, the CCDC built a comprehensive information system to monitor infectious diseases. In a 2017 thesis that reviews the surveillance system for pneumonia of unknown aetiology (PUE), it is concluded that 'the system has played an important role in detection of cases in early stage of human infection with avian influenza outbreak ... , but the reporting rate and the positive rate of target pathogen in

PUE cases is very low’. Nonetheless, the director of CCDC, Dr. Gao Fu, reassured the public that ‘SARS-like viruses can be found at almost any time, but another SARS epidemic will not happen again’ in March 2019 (Ni, 2020), just months before the alarm was sounded again.

2.2 | The investigations of human-to-human transmissibility

After receiving the first warning, the CCDC immediately sent a group of experts to Wuhan, but the initial investigation did not reach a definitive conclusion. On 5 January 2020, the Wuhan Health Commission announced that the number of pneumonia of unknown aetiology cases had increased to 59, but there was ‘no clear evidence of human-to-human transmission’ (Wuhan Health Commission, 2020a).

A second group of experts then went to Wuhan on 8 January and reached, surprisingly, a more optimistic conclusion. A statement released on 11 January said that there were no new cases of pneumonia showing symptoms after 3 January and there were no infections among doctors and nurses (Wuhan Health Commission, 2020b). One clinician in the expert group, Dr. Wang Guangfa, told the media that the disease ‘can be prevented and contained’ (Liao & Li, 2020). In the meantime, the causative pathogen of the pneumonia was officially confirmed as a novel coronavirus on 9 January (Qu, 2020). The number of suspected cases kept growing rapidly in hospitals across Wuhan (Xiao, 2020; Yang, 2020c), but the official case count stagnated at 41 until 18 January (Table 1). This is perhaps why the Wuhan Health Commission maintained the optimistic tone in its statements during this period. Besides reiterating that ‘there is no clear evidence of human-to-human transmission’, a 14 January press release further made the following assessment: ‘Although we cannot exclude the possibility of limited human-to-human transmission, the risk of sustained transmission is low’ (Wuhan Health Commission, 2020c).

However, things took a dramatic turn around 17 January. Four new cases in Wuhan were confirmed that day, and more worryingly, the border controls in Thai and Japanese airports had just detected three additional cases who were travellers from Wuhan. By using an estimated number of international travellers from the Wuhan airport and Wuhan’s population, a group of researchers in Imperial College London estimated that a total of 1723 (95% confidence interval [CI]: 427–4471) coronavirus cases in Wuhan had onset of symptoms by 12 January, which is far more than the official number (Imai et al., 2020b). Their method was extremely simple, but the results were equally appalling. Moreover, it was discovered on 15 January that in an early family cluster of cases in Shenzhen, one family member had not been to Wuhan in the recent past (Southern Metropolis Daily, 2020). This proved that the pathogen could be transmitted between humans. A third expert group was then formed and sent to Wuhan on 18 January. Just 2 days later, it was confirmed for the first time to the public that the new coronavirus was indeed infectious and very likely to be highly infectious. And another 3 days later, the entire city of Wuhan entered an abrupt and

TABLE 1 Total number of confirmed cases in Wuhan in January 2020 by date of press release

Date	12-31	01-03	01-05	01-11	01-12	01-13	01-14	01-15	01-16	01-18	01-19
Cases	27	44	59	41	41	41	41	41	41	45	62
Date	01-20	01-21	01-23	01-24	01-25	01-26	01-27	01-28	01-29	01-30	01-31
Cases	198	258	425	495	572	618	698	1590	1905	2261	2639

Source: Press Releases from the Wuhan & Hubei Municipal Health Commissions.

unprecedented lockdown, which was repeated numerous times across the world over the next few months.

2.3 | Evidence of high infectiousness

After the lockdown on 23 January 2020, hospitals in Wuhan soon became overwhelmed with patients showing COVID-19 symptoms. The number of deaths also started to skyrocket. In the midst of the emotional outburst after Dr. Li Wenliang's tragic and emblematic death on 7 February, the initial investigations came under close scrutiny. Several interviews and articles were published by the Chinese media in the next few weeks and offered a lot of insights into the investigations. Unfortunately, these lessons did not appear to be widely recognized outside China in a quickly developing pandemic. Below I will try to bring us back in time and list some evidence of high infectiousness that was already available by 10 January. The reader should keep in mind that, even though the investigative journalism on these investigations appears to be of very high quality, all the interviews and media reports are still retrospective.

First of all, SARS-CoV-2 is not the first coronavirus that has circulated in the human population. Several human coronaviruses produce the generally mild symptoms of the common cold worldwide, and two quite lethal strains (SARS-CoV and MERS-CoV) as mentioned above had led to earlier outbreaks in this century. Being a novel coronavirus, it was quite likely that SARS-CoV-2 was also infectious to humans and poses a major threat to the public health.

Second, although the causative pathogen was officially confirmed as a novel coronavirus on 9 January 2020, there had already been many hints earlier. Weiyuan Gene, a company based in Guangzhou, received a patient sample from the Wuhan Central Hospital on 24 December 2020. Their metagenomic Next Generation Sequencing (mNGS) identified a SARS-like coronavirus in the sample and obtained a near-whole-genome sequence of the virus on 27 December Gao et al. (2020). The sequencing results were shared with the Institute of Pathogen Biology within the Chinese Academy of Medical Sciences. It was reported that a manager at Weiyuan Gene informed the hospital and disease control officials by telephone on the same day the sequencing results came out, and travelled to Wuhan on 29 December to discuss the results (Xiaoshangou (online id), 2020). According to another media report, at least nine samples were sent for testing by the end of December. Huada Gene, the industry leader in China, also received a sample and identified a SARS-like coronavirus on 30 December and informed the hospital about the results; the whole genome sequence and the testing reports were then shared with the CCDC and Wuhan Health Commission on 1 January (Gao et al., 2020). So it is fair to say that, before the official announcement, the investigators from the CCDC already knew that the causative pathogen was very likely a novel coronavirus.

Third, several hospitals in Wuhan were seeing rapidly increasing numbers of suspected cases in retrospective accounts. In an interview, Dr. Wang Guangfa, a respiratory disease clinician in the second group of experts who went to Wuhan, mentioned that one hospital in Wuhan had a 17% increase in pneumonia cases in December 2019. By 10 January, the 16 intensive care unit (ICU) beds at Zhongnan Hospital reserved for pneumonia patients were already filled up (Xiao, 2020). Jinyintan Hospital, a hospital specializing in treating infectious diseases which also became the first designated hospital for COVID-19 in Wuhan, had at least five hospital wards full of suspected cases, each with more than 40 patients; the emergency department at another major hospital was seeing 30–50 patients every day who required immediate hospitalization (Yang, 2020d).

In hindsight, there had been enough evidence by 10 January 2020 to conclude that the novel coronavirus was very likely to be highly infectious. So in the perfect scenario, the experts who went to Wuhan could have concluded that the disease outbreak posed a major threat and more stringent public health measures need to be imposed, which in reality only happened after 20 January.

2.4 | A faster decision?

So why was this conclusion not reached sooner? There are a multitude of reasons. The first is the cautious and conservative approach of the authorities towards infectious diseases. This is wise for most infectious diseases, but perhaps not for a completely novel virus with pandemic potential. On 3 January 2020, all laboratories not affiliated with the CCDC were asked by the National Health Commission on 3 January to destroy their samples and to not disclose their existing results to the public (Gao et al., 2020), due to the Chinese Law on the Prevention and Treatment of Infectious Diseases (Standing Committee of the National People's Congress, 2004). This means that evidence from the previous commercial tests, which had already identified a SARS-like coronavirus a week earlier, could not be relied on in the official investigations. The CCDC received their first patient sample from Wuhan on 2 January (Chinese Centre for Disease Control, 2020) and officially announced that the pathogen is a novel coronavirus on 9 January (Qu, 2020).

This conservative approach can also be seen from the case numbers in the press releases of the Wuhan and Hubei Health Commissions (Table 1). The Wuhan Health Commission made no official announcements between 6 January and 10 January, but changed the name of the disease from viral pneumonia of unknown aetiology to pneumonia due to a novel coronavirus on 11 January. Curiously, the case number reported on 11 January dropped from 59 (reported on 5 January) to 41; no explanation was specified, but a reasonable guess is that only the samples of 41 patients were tested positive for the new coronavirus. No new cases were announced before 17 January, because the local disease control had not received the approved test kits from the CCDC before 16 January. During that transition period, a rigorous and lengthy process (that took at least 3–5 days) was needed to confirm a coronavirus case (Xu, 2020), as demonstrated in Figure 1. The final diagnosis, which involves a comprehensive consideration of clinical, epidemiological and biological evidence, inevitably required a lot of time.

Another factor that delayed the decision is, as reported by several media outlets, strict case definition during the early outbreak. Around 3 January, the first CCDC investigators and local experts in Wuhan devised a preliminary guideline in a booklet with a green cover for the treatment of COVID-19 (this was still called viral pneumonia of unknown aetiology, or PUE, as the name COVID-19 was designated much later). This booklet followed a 2007 guideline by the Chinese Ministry of Health made after SARS and considered a patient to have a PUE if four clinical criteria were satisfied: fever ($> 38^{\circ}\text{C}$); imaging characteristics of pneumonia; normal or low white blood cells in the early course of disease, or decreased lymphocytes; illness had no improvement or worsened after at least 3 days of standard antibiotic treatment (Chinese Health Commission, 2007; Yang, 2020a). If the suspected case had been exposed to the Huanan seafood market, they only needed to satisfy three of the four criteria above.

What became quite controversial is another booklet with a white cover printed by the Wuhan Health Commission. This 'white booklet' included all the documents in the 'green booklet' and an additional document that makes epidemiological exposure a necessary condition. That is, PUE

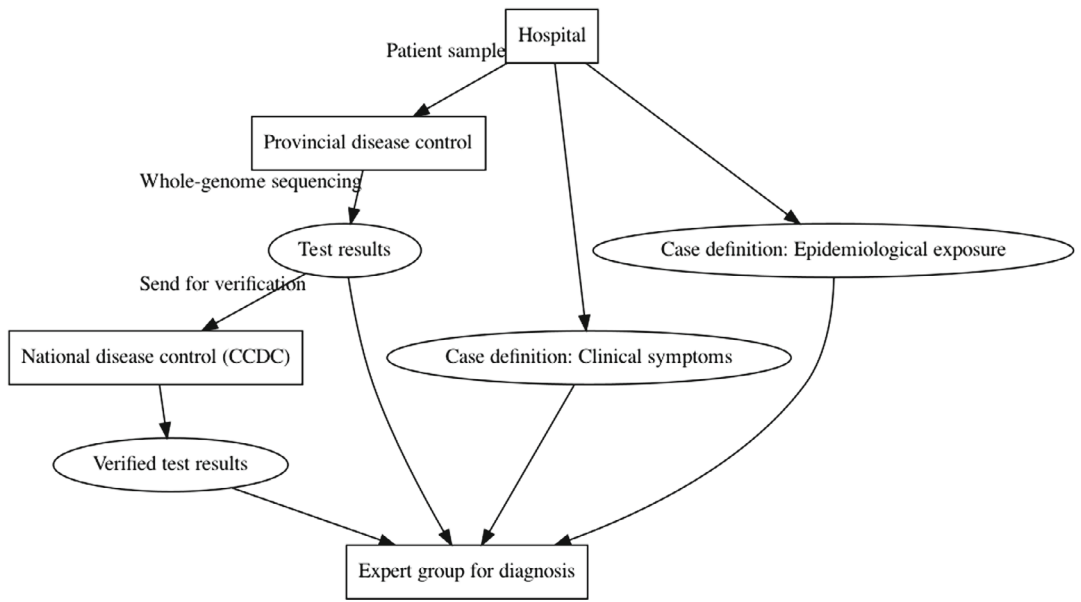


FIGURE 1 Case diagnostic process before 16 January 2020 (translated from a diagram in Xu (2020)).

cases not only needed to show the four clinical symptoms listed above but also had to have a direct epidemiological link to Huanan seafood market (Gao, 2020; Yang, 2020a). This is a critical mistake in hindsight: the additional requirement made it impossible to conclude that the coronavirus was also quickly circulating outside the Huanan seafood market. It remains unclear who edited the ‘white booklet’ or why the new inclusion criterion was added, but an anonymous investigator in the first CCDC expert group denied knowledge of the ‘white booklet’ and the additional document (Yang, 2020a). It is also reported that many doctors on the front line at the time believed that the epidemiological criterion was too strict (Xiao, 2020; Yang, 2020a).

A third reason is that the surveillance system designed by the CCDC to prevent another SARS-like epidemic (see Section 2.1) did not work as intended. According to the 2007 guidelines, once a pneumonia case satisfying the definition was found, hospitals were required to organize a consultation (by specialists in the hospital) within 12 hours. If no diagnosis could be reached, the hospital report it as a PUE case through the surveillance system. However, many doctors were unaware of this requirement or the surveillance system (Xin et al., 2020). Moreover, some hospitals never reported PUE cases because the patients did not satisfy the epidemiological exposure criterion discussed above (Wu, 2020). Only a few hospitals started to report their PUE cases through the surveillance system after 3 January. These direct reports were then visible to the municipal, provincial and national Health Commissions (Wei et al., 2020). However, the direct reports stopped around 10 January and only resumed in late January. According to media reports, officials at the Wuhan and Hubei Health Commissions asked the hospitals to be extremely cautious about reporting via the surveillance system (Yang, 2020b). From 4 January to 10 January, the requirement for reporting through the surveillance system went from ‘consultation within the hospital’ to ‘consultation with experts in the borough’, ‘centralized reporting by the city and province’, and finally ‘approval by the provincial Health Commission’ (Wei et al., 2020). So the surveillance system could not provide reliable evidence to infer the severity of the epidemic.

There also appears to have been a lack of statistical expertise and coordination in the first two expert groups. It is reported that the first investigation group (who went to Wuhan on 31 December 2019) included an expert from the CCDC specializing in pathogen identification and two doctors from hospitals in Beijing specializing in respiratory diseases (Li et al., 2020b). Although it is certainly important to determine the pathogen of the new disease and understand its clinical features, evaluating the infectiousness of the disease, another crucial question in early disease outbreaks, seemed to be beyond the group's expertise. The second group of experts included both clinicians and epidemiologists. However, I could not find any media reports suggesting that the epidemiologists tried to gather additional evidence beyond what was officially provided by the local disease control, before concluding that the outbreak 'can be prevented and contained' on 10 January 2020 (Liao & Li, 2020).

Intriguingly, the conclusion that the outbreak could be contained first appeared in an interview with Dr. Wang Guangfa, who is a respiratory disease clinician (Liao & Li, 2020). In a later interview, Dr. Wang acknowledged that he was not an epidemiologist and the decision was made by the group together (Qin, 2020). He also mentioned that the epidemiologists from the CCDC could not make a conclusion because there were not enough data to quantify the basic reproductive number R_0 (there were only 41 officially confirmed cases with two family clusters). I think this shows a confusion or miscommunication about what 'human-to-human transmissibility' means. If it simply means that the disease has the ability to be passed from one human, which is the literal meaning of the phrase in Chinese (and in English), calculating the R_0 is unnecessary and all we need is just one or a few instances like the family cluster in Shenzhen. If it means that R_0 is larger than 1 so a large outbreak will likely happen without interventions, then, due to the lengthy process of case confirmation during that period (Figure 1), the investigation should be focused on the growth of suspected cases or excessive cases of fever and pneumonia. In any case, the public messages sent out by the investigators on 10 January (Liao & Li, 2020) and 11 January (Wuhan Health Commission, 2020b) seem too optimistic and indefensible.

Finally, there appears to be a lack of considerations of risk in the initial investigations, given that SARS already had the potential of becoming a pandemic. In particular, the timing of Wuhan's COVID-19 outbreak was precarious: 10 January 2020 officially marks the beginning of the official travel season (*chunyun* in Chinese) for the Lunar New Year, which in 2020 falls on 25 January. Before Wuhan's eventual lockdown on 23 January, five million people left Wuhan (typically to go back home or for tourism) (Li et al., 2020b). If the new disease was indeed more contagious than SARS, there would be an extremely high risk of spreading the virus across China and to the world. Eventually, this became the main reason behind the unprecedented lockdown of Wuhan on 23 January, just 2 days before the Spring Festival (Sun, 2020). However, I could not find any articles or interviews discussing risk assessment and management in the initial investigations.

3 | HOW FAST WAS COVID-19 GROWING?

By late January 2020, it became abundantly clear that the novel coronavirus was infectious and a sizable number of people in Wuhan had already been infected. The next immediate question was: quantitatively speaking, how infectious was the novel coronavirus? How likely was it that the local outbreak would turn into a global pandemic?

Around the same time, I started to read COVID-19 studies and perform some statistical analyses. Prior to this, I had no experience with infectious disease modelling, but I knew a little bit about epidemiology from my research in causal inference. The first thing I noticed was that many

studies were trying to estimate the basic reproductive number, or R_0 , of COVID-19. The R_0 number is usually defined as the average number of infectees per infected person in the beginning of the outbreak. One can define R_0 more precisely by using a mathematical model for infectious diseases. Something I realized much later is that this also means the precise definition of R_0 is model dependent.

As some examples, the 2003 SARS epidemic had an estimated R_0 of 2–4 (World Health Organization, 2003b). MERS has an R_0 below 1; seasonal strains of influenza usually have an R_0 below 2 (World Health Organization, 2017). This number is important because if we view the epidemic as a birth process with mean number of offspring equal to R_0 , whether R_0 is larger than 1 critically decides the chance of a large outbreak. Moreover, the R_0 value also determines the threshold of ‘herd immunity’ in vaccine development. Besides R_0 , another important metric for the infectiousness of a disease is the initial doubling time, which better captures the urgency of the matter.

In this section, I will first briefly review two initial studies on the infectiousness of COVID-19. Because of how early they were published in prestigious medical journals, they attracted massive attention and greatly influenced the initial responses to COVID-19 across the world. Surprisingly, a careful reanalysis of the mathematical models in these studies shows that they do not fit the data in those studies well. I will then discuss alternative ways to analyse early outbreak data and how early studies of COVID-19 understated their uncertainty about R_0 .

When reading this section, the reader should keep in mind that, as we now know, there are a significant amount of asymptomatic COVID-19 infections and undetected cases. Under-ascertainment is an especially pronounced issue in the first weeks of the outbreak, and no statistical analysis can completely salvage poor data. Nonetheless, it is still important to apply the most appropriate methods to the available data and make correct interpretations of the results, which many initial COVID-19 studies failed to do.

3.1 | Initial studies and mistakes

On 29 January 2020, researchers at the CCDC published the first epidemic incidence curve (by symptom onset date) of COVID-19 on the *New England Journal of Medicine* (Li et al., 2020a). This paper analysed the first 425 confirmed cases in Wuhan and estimated that the initial doubling time was 7.4 days (95% CI 4.2–14), the R_0 was 2.2 (95% CI 1.4–3.9). These were the first peer-reviewed quantitative estimates of the infectiousness of COVID-19.

In the article, the authors described their statistical analysis as follows:

We estimated the epidemic growth rate by analysing data on the cases with illness onset between 10 December and 4 January, because we expected the proportion of infections identified would increase soon after the formal announcement of the outbreak in Wuhan on 31 December. We fitted a transmission model (formulated with the use of renewal equations) with zoonotic infections to onset dates that were not linked to the Huanan Seafood Wholesale Market, and we used this model to derive the epidemic growth rate, the epidemic doubling time and the basic reproductive number (R_0).

However, the article provided no particulars about the transmission model and only stated that the model was fitted using MATLAB.

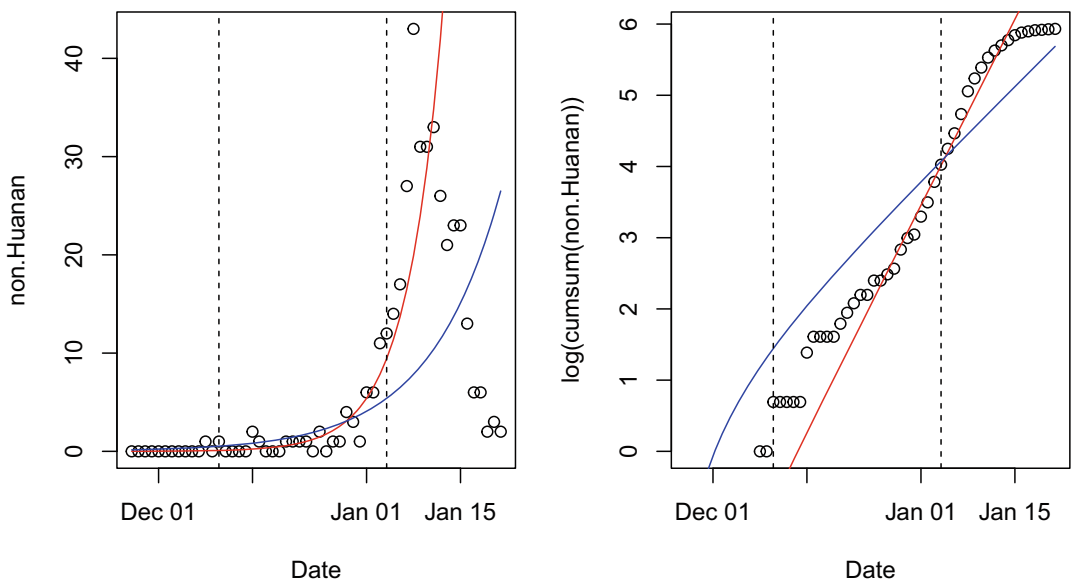


FIGURE 2 Initial epidemic curve (number of new symptom onsets per day) in Wuhan (Li et al., 2020a) and the fitted log-linear models. Poisson log-linear models were fitted using the incidences between 10 December and 4 January (dashed lines). The red curves correspond to an unrestricted fit; the blue curves correspond to the best fit assuming that the growth exponent correspond to a doubling time of 7.4 days. The left panel shows the new incidences with no exposure to the Huanan Seafood Market; the right panel shows the logarithm of the cumulative incidences.

Surprisingly, when fitting a simple Poisson log-linear model to the same epidemic incidence curve, I obtained a quite different estimate: the slope coefficient in the log-linear model, also known as the ‘growth exponent’ r , is estimated to be 0.187 (95% CI 0.135–0.245), which corresponds to a doubling time $\log(2)/r$ of 3.7 days (95% CI 2.8–5.1). Figure 2 shows that this simple log-linear model (red curve) provides a good fit with a residual deviance of 27.6 on 24 degrees of freedom, whereas the exponential growth curve corresponding to a doubling time of 7.4 days clearly does not fit the data. Beyond the lack of fit, the restriction to symptom onsets before 4 January was also questionable, as the observed incidences were higher than the predicted values from the log-linear model over the next few days.

If we model an epidemic outbreak as a birth process, the initial growth exponent r can be related to the basic reproductive number R_0 via the identity $R_0 = 1/M(-r)$, where $M(\cdot)$ is the moment generating function of the generation time (time between the infections of successive cases in a chain of transmission) (Wallinga & Lipsitch, 2006). Estimating the distribution of the generation time is a difficult if not daunting task because the infection times and transmission chains are rarely observed. In practice, it is common to estimate the distribution of serial intervals (time between the symptom onsets of successive cases) using some known transmission chains/clusters and use it to approximate the generation time distribution, but this method is not free from statistical issues (Bacallado et al., 2020). In Li et al. (2020a), the authors fitted a Gamma distribution to the serial intervals for just six pairs of presumed infector–infectees: 5, 9, 7, 7, 3 and 7. They used the estimated serial interval of SARS (mean 8.4 days, standard deviation 3.8 days) as an informative prior and estimated that the serial interval distribution of COVID-19 had a mean

of 7.5 days and standard deviation of 3.4 days. No uncertainty quantification was given to these numbers, but they were widely used in subsequent studies.

By using this distribution and the formula for R_0 , an initial doubling time of 7.4 days (95% CI 4.2–14) would correspond to a R_0 of 2.0 (95% CI 1.4–3.2), which is reasonably close to what the article reported. However, if we used the estimated growth exponent from our Poisson log-linear model, the corresponding R_0 would be 3.7 (95% CI 2.6–5.5). From a policy perspective, these numbers have very different interpretations given that the R_0 of SARS was estimated to be about 2–4.

Another influential study modelling the transmission of COVID-19 was published by the *Lancet* on 31 January 2020 (Wu et al., 2020). This study used 78 exported cases from Wuhan to areas outside mainland China to fit a susceptible-exposed-infectious-recovered (SEIR) model for the epidemic in Wuhan. To model case exportation, they used a non-homogeneous Poisson process with an estimated volume of outbound air travel from Wuhan. The basic reproductive number is a parameter in their SEIR model and was estimated to be 2.68 (95% CI 2.47–2.86); the initial doubling time was estimated to be 6.4 days (95% CI 5.8–7.1). A closer read of the article reveals that when fitting this model, the authors only used cases who first showed symptoms before or on 19 January 2020; this end date was chosen ‘to minimize the effect of lead time bias on case confirmation’. What the authors did not mention, however, is that this criterion left them with just 17 exported cases.

I fitted a simple Poisson log-linear model to the incidence curve of these 17 cases and the growth exponent was estimated to be 0.117 (95% CI 0.044–0.202), which corresponds to a doubling time of 5.9 days (95% CI 3.4–15.7). This corresponds to an R_0 of 2.58 (95% CI 1.44–4.96), if we follow the *Lancet* study and use the estimated serial interval of SARS (mean 8.4 days, standard deviation 3.8 days). The point estimates I obtained are reasonably close to the original numbers, but the original study seems to have massively understated the variability of the estimates.

An examination of the fitted and predicted values shows that this model substantially underestimates the number of exported cases in the few days after 19 January; see the blue curves in Figure 3. This shows that the under-ascertainment bias might not have been as large as the authors imagined. If the incidences between 20 January and 23 January are also used in model fitting, the initial doubling time would be estimated to be 3.9 days (95% CI 2.9–5.5) and the corresponding R_0 would be 4.16 (95% CI 2.77–6.60); see the red curves in Figure 3.

Finally, a key missing element in this analysis is Wuhan’s lockdown—virtually no civilians were allowed to leave Wuhan after 23 January 2020. This in fact has a nontrivial effect on the incidence curve. For example, consider two cases infected in Wuhan, one on 15 January and one on 22 January. It is much more likely for us to observe the first as an exported case than the second, because the second case had a much smaller chance of travelling outside Wuhan before the lockdown.

3.2 | Beyond modelling epidemic curves

Next I will explain the basic ideas that are needed to correct the selection bias described in the last paragraph. The illustration below is based on an early study (Zhao et al., 2020a) that I was involved in. This study appeared on medRxiv in early February 2020 and was one of the first to

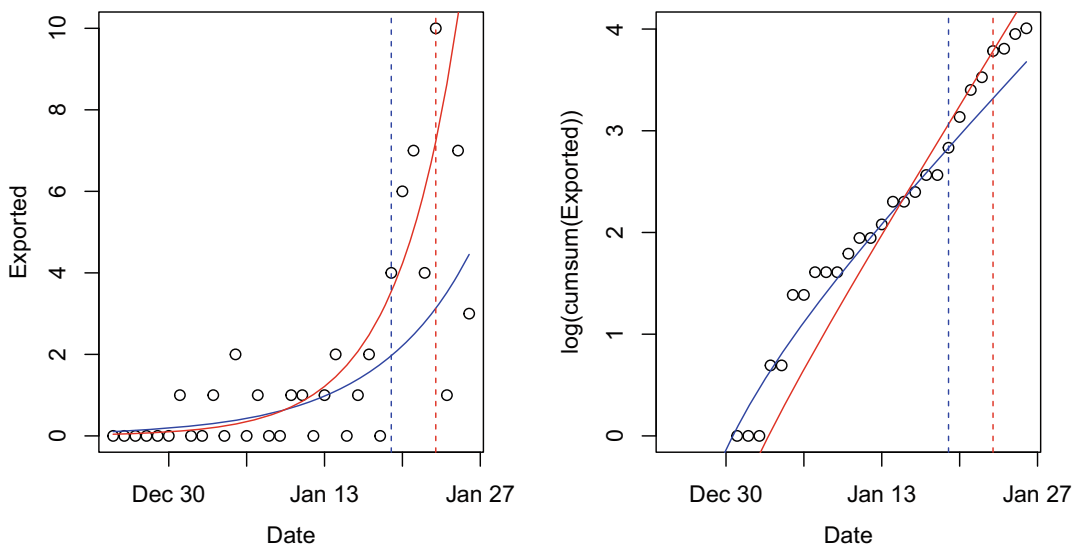


FIGURE 3 Initial epidemic curve (number of new symptom onsets per day) for Wuhan-exported cases (Wu et al., 2020) and the fitted log-linear models. The blue curves correspond to fit a Poisson log-linear model using the incidences between 25 December and 19 January (blue dashed lines), the choice in the original study. The red curves correspond to fitting the same model using data up to 23 January (red dashed lines). The left panel shows the new incidences with no exposure to the Huanan Seafood Market; the right panel shows the logarithm of the cumulative incidences.

suggest that the initial studies substantially underestimated the epidemic growth. Unfortunately, this study is never published.

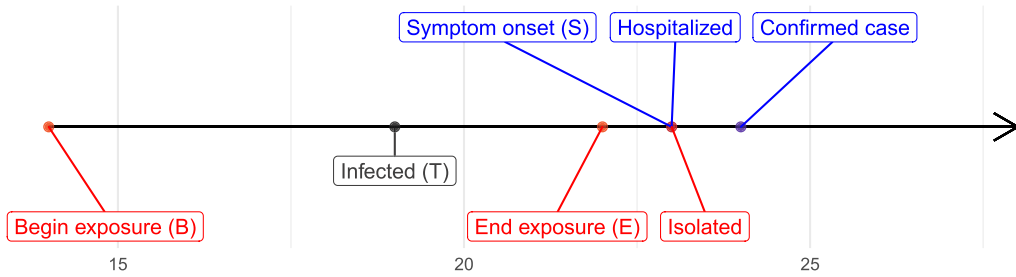
After Wuhan's lockdown, I started to collect data about the exported cases and immediately noticed that some countries and regions published detailed information about the confirmed cases. As an example, below is an excerpt of the Hong Kong government's press release on 24 January 2020 (translated from Chinese in Zhao et al. (2021)):

The other two cases are a married couple of residents of Wuhan, a 62-year-old female and a 63-year-old male, with good prior health conditions. Based on information provided by the patients, they took a high-speed train departing from Wuhan at 2:20 pm, 22 January, and arrived at the West Kowloon station around 8 pm. The female patient had a fever since yesterday with no respiratory symptoms. The male patient started to cough yesterday and had a fever today. They went to the emergency department at the Prince of Wales Hospital yesterday and were admitted to the hospital for treatment in isolation. Currently their health conditions are stable. Respiratory samples of the two patients were tested positive for the novel coronavirus.

The case description clearly contains richer information than the aggregated epidemic curves, although it might not be immediately clear how such information can be useful. In Zhao et al. (2020a, 2021) we identified three key events in the trajectory of each case:

1. The beginning of stay in Wuhan, B (the beginning of exposure to the disease);
2. The end of stay in Wuhan, E (the end of exposure to the disease);
3. The onset of symptoms, S .

HK-04



KR-01

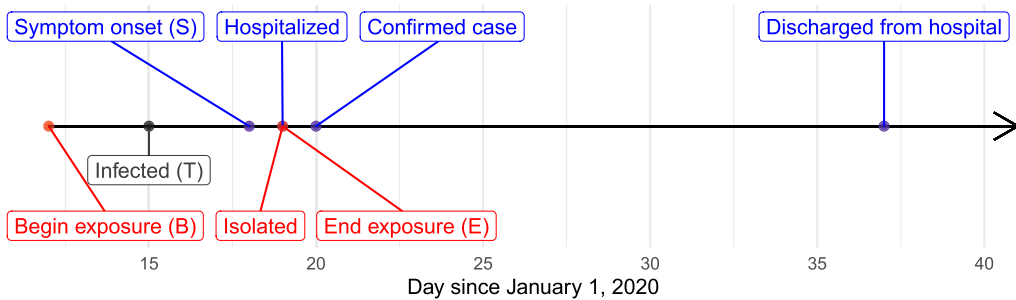


FIGURE 4 Timelines of two COVID-19 cases (Zhao et al., 2021). HK-04 is the fourth case in Hong Kong (62-year-old female) and KR-01 is the first case in South Korea (35-year-old female). The colour indicates type of event (blue: pathophysiological; red: epidemiological). [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

These events can help in inferring the latent time of infection, T . To understand the relationship between these events, we may model the data using marked point processes (see Figure 4 for two examples). Intuitively, we can use the case description to fill two timelines: a pathophysiological timeline that records the disease progression and recovery (blue in Figure 4), and an epidemiological timeline that records events related to disease transmission (red in Figure 4). Infection is both a pathophysiological and an epidemiology event.

For an exported case, the key event times satisfy the obvious constraint $B \leq T \leq E$. Moreover, $S - T \geq 0$ is the incubation period. Since S is related to the pathophysiological process and B , E are related to the epidemiological process, it may be reasonable to assume that S is conditionally independent of B , E given T (see Zhao et al. (2021) for more discussion on this assumption). This means that if the distribution of the incubation period is known, we can impute the latent infection time T from the symptom onset S by simulating $S - T$ from the incubation period distribution and rejecting the samples that do not satisfy $B \leq T = S - (S - T) \leq E$. We can then fit an appropriate growth model to the imputed T . This is the first basic idea in Zhao et al. (2020a).

Another key observation in Zhao et al. (2020a) is that, because of the travel quarantine of Wuhan, those potential travellers who were infected earlier would have a higher chance of being observed as exported cases. Therefore, instead of using a simple exponential growth model for the density of T

$$f_T(t) \propto e^{rt}, \quad (1)$$

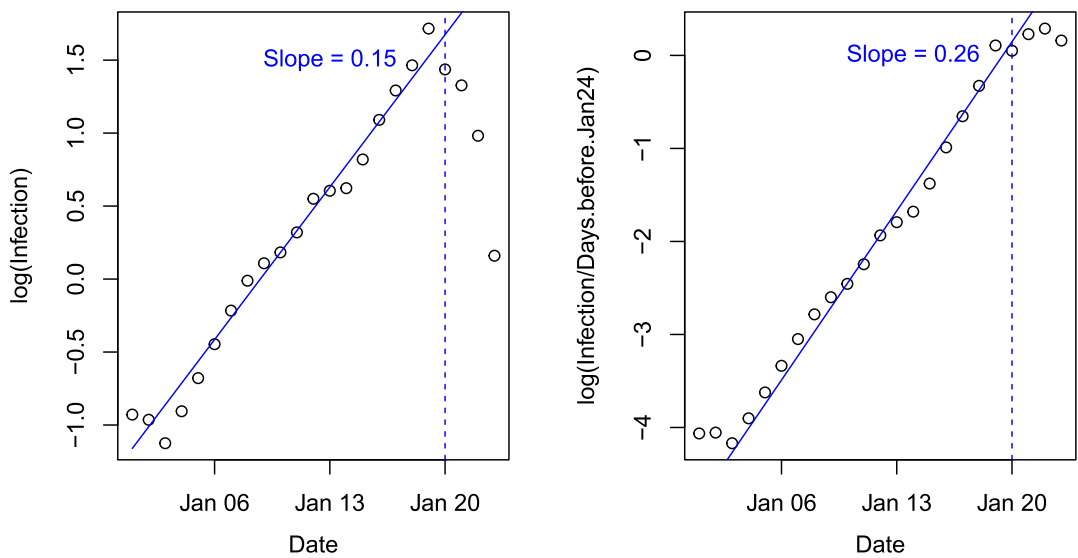


FIGURE 5 Fitting growth models to the imputed infection time. Data points correspond to the average imputed infection on each day after the transformations specified in the axis label. The left panel fits a simple exponential growth model to the imputed T ; The right panel fits the model (2) that is corrected for the observation bias. [Colour figure can be viewed at wileyonlinelibrary.com]

a better model is

$$f_T(t) \propto e^{rt} \cdot \max(L - t, 0), \quad (2)$$

where L is the time of travel quarantine (Zhao et al., 2020a). This model assumes that the probability of observing an exported case infected at time t is decreasing linearly in t until the lockdown; intuitively, this is reasonable if the international travels from Wuhan were stable.

Figure 5 shows a replication of the analysis in Zhao et al. (2020a) with 46 Wuhan-exported international cases collected in Zhao et al. (2021) the results are slightly different from the original study because the reanalysis used a slightly different definition of exported cases. The left panel fits the exponential growth model (1) to the expected infection counts (obtained from the imputation algorithm described above) between 1 January and 20 January. For the incubation period, I used a log-normal distribution with mean 5.2 and standard deviation 4.9, which is the initial estimate in Li et al. (2020a) and the choice in Zhao et al. (2020a). The estimated growth exponent r is 0.149, which corresponds to a doubling time of 4.6 days. However, this clearly overestimates the number of infections from 20 January to 23 January as shown in the figure. The right panel of Figure 5 fits the bias-corrected model (2) by including an offset. The new estimate of the growth exponent r is 0.260, which corresponds to a doubling time of 2.7 days.

Determining the sampling error of estimates is not trivial. To simplify this replication analysis, I did not use the Bayesian method described in the original analysis (Zhao et al., 2020a). Instead, I fitted quasi-Poisson models corresponding to (2) to 10,000 realizations of the transmission times and found that the 2.5% and 97.5% sample quantiles are 0.194 and 0.339. This is not a confidence or credible interval but gives us a sense of the variability of r .

A weakness of this approach is that the incubation period needs to be known. The better method is to jointly estimate the epidemic growth and the incubation period distribution using Wuhan-exported cases. See Zhao et al. (2021) for more detail, including a theoretical justification of model (2).

3.3 | Large uncertainty about R_0

Table 2 collects some early estimates of the initial growth exponent r and basic reproductive number R_0 of COVID-19 as well as the results of the replication analyses above. There is a significant discrepancy between these estimates and some are likely mistaken (Section 3.1). It is worth noting that the two studies that reported the highest R_0 (Sanche et al., 2020) analysed individual cases, while all the other studies used aggregated incidences.

Most of the epidemiological analyses of COVID-19 estimated the basic reproductive number R_0 instead of the growth exponent r . This is because R_0 can be expressed in terms of the parameters in standard compartmental epidemic models and is more relevant in forecasting the epidemic with various kinds of interventions. But from a statistical perspective, it is much easier to estimate r . Moreover, an initial doubling time of 2–3 days, which is what most countries in Europe and elsewhere experienced when the pandemic first hit (Pellis et al., 2021; Zhao, 2021) also highlights the urgency of the matter.

To determine R_0 , one needs to estimate not only r but also the distribution of the generation time. As mentioned above, the R_0 and r can be linked through the formula $R = 1/M(-r)$, where $M(\cdot)$ is the moment generating function of the generation time. In practice, it is common to approximate the distribution of generation time by the distribution of serial interval, but this approximation is not very accurate for COVID-19. As an example, the generation time is always positive by definition, but the serial interval can be negative if the infector is asymptomatic. The early estimates of COVID-19's serial interval are also very imprecise due to small sample sizes (the CCDC paper (Li et al., 2020a) only used six pairs of infector–infectees) and methodological issues (Bacallado et al., 2020).

It may come as a surprise to statisticians that all the early estimates of the R_0 of COVID-19 in Table 2 did not take into account the uncertainty about the generation time distribution. In fact, had that been considered, many confidence intervals would have become too wide to be useful. To illustrate this, consider the estimated generation time in Kremer et al. (2020) based on 91 cases in Singapore in February 2020. In this study, the generation time is modelled by the Gamma distribution and so can be parameterized by its mean and standard deviation. The left panel of Figure 6 shows 100 samples from the posterior distribution of the generation time and the density contours. This plot shows the following generation time distributions are not anomalous in the posterior: mean = 4 days, standard deviation = 2 days; mean = 6 days, standard deviation = 4 days. However, they can lead to very different estimates of R_0 , as shown in the right panel of Figure 6. In particular, by using the formula $R = 1/M(-r)$, a doubling time of 2.5 days corresponds to a $R_0 = 2.7$ under the first generation time distribution, and a $R_0 = 4.9$ under the second generation time distribution. These R_0 values have drastically different implications for policy decisions, not to mention that this illustration has not taken into account the uncertainty in r yet.

In conclusion, these initial analyses of COVID-19 not only underestimated its R_0 but also underestimated the uncertainty about R_0 . They had a far reaching impact on the early global responses to the developing pandemic.

TABLE 2 Early estimates of r and R_0 of COVID-19

First author	Published/ first posted (in 2020)	Data collected (in 2020)	Sample	Doubling time $\log(2)/r$	Reproductive number R_0
Imai et al. (2020a)	25 January	22 January	7 international exported cases		2.6 (1.5–3.5)
Read et al. (2021)	28 January	22 January	Case counts		3.1 (2.4–4.1)
Li et al. (2020a)	29 January	22 January	425 cases in Wuhan	7.4 (4.2–14)	2.2 (1.4–3.9)
Replication of Li et al. (2020a)				3.7 (2.8–5.1)	3.7 (2.6–5.5)
Zhao et al. (2020b)	30 January	24 January	Case counts		Varies with reporting rate: 2.2 (2.0–2.6) to 3.6 (2.9–4.4)
Wu et al. (2020)	31 January	25 January	78 international exported cases	6.4 (5.8–7.1)	2.7 (2.6–2.9)
Replication of Wu et al. (2020)			Only 17 cases meet the selection criterion	5.9 (3.4–15.7)	2.6 (1.4–5.0)
Zhao et al. (2020a)	9 February	5 February	46 international exported cases	2.9 (2.0–4.1)	5.7 (3.4–9.2)
Replication of Zhao et al. (2020a)				2.7 (2.0–3.6)	
Yang et al. (2020) (later withdrawn)	11 February	26 January	4021 cases in China		3.8 (3.5–4.1)
Sanche et al. (2020)	11 February	End of January	140 initial confirmed cases in other Chinese provinces	Method 1: 2.4 (1.9–3.3); Method 2: 2.3 (2.0–2.6)	Method 1: 6.3 (3.3–11.3); Method 2: 6.6 (4.0–10.5)

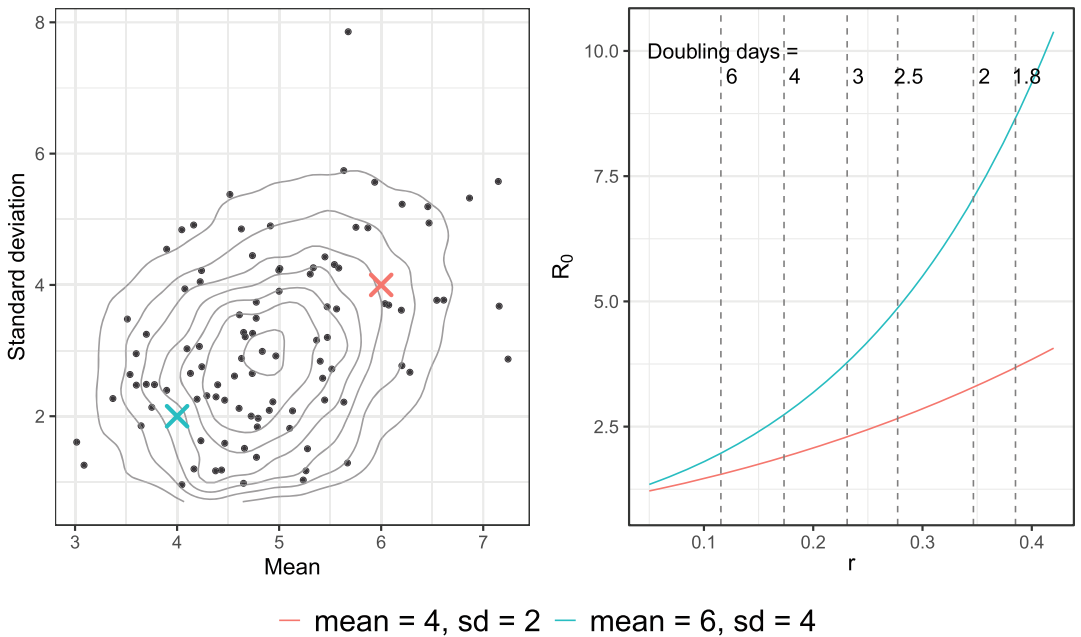


FIGURE 6 Estimated generation time and the implied basic reproductive number. The left panel shows 100 samples from the posterior distribution of the mean and standard deviation of the generation time in a replication analysis of Kremer et al. (2020). The right panel shows the implied R_0 for different values of r and two generation time distributions.

4 | DISCUSSION AND LESSONS LEARNED

This article has reviewed some of the most consequential outbreak analyses before COVID-19 turned into one of the worst pandemics in modern history. Reliable data on COVID-19 were scarce, so the initial analyses are understandably imperfect. Nonetheless, had the investigators and decision makers been better prepared, some of the mistakes could have been easily avoided.

In some instances, it is even possible to use common sense to see that the conclusions are likely erroneous. For example, as mentioned above the first CCDC article (Li et al., 2020a) estimated that the initial doubling time of COVID-19 in Wuhan was 7.4 days. This means that, if community transmission started in mid to late December 2019, as their data indicate, the epidemic would have gone through only six doubling cycles by the time that article was published (29 January 2020). Consequently, there should have been at most a few hundred cases (as $2^6 = 64$) in Wuhan by late January. However, the number of confirmed cases in Wuhan was already 1905 on 29 January, not to mention that several times more patients with the same clinical features were waiting to be tested. This kind of inconsistency creates confusion in an emerging crisis and can erode the public trust in science and policy.

For statisticians and anyone who engages in data-driven research, I think there are several lessons that can be learned from these mistakes.

Lesson 1: Small data analysis is crucial but not easy

In an emerging disease outbreak, decisions are inevitably made based on limited data. Inference for small samples is a hallmark of statistics and is routinely taught in introductory statistics

classes. However, small data analysis is not easy, as exemplified by the initial investigations of COVID-19. The common mistake in many of those investigations is that a general statistical method is applied to new datasets (small or big) without checking if the method is appropriate.

This lesson is of course not new, but its importance in all areas requiring quantitative analysis cannot be understated. In the era of data explosion, many new questions will be asked and will become possible to answer. It will be tempting to apply 'standard' statistical methods to answer these questions, but we must keep in mind that the new data may be collected in non-conventional ways and standard statistical models may be inappropriate.

Lesson 2: Practitioners are still unfamiliar with basic statistical concepts

Another rather discouraging lesson for statisticians is that many applied researchers and the public are still unfamiliar with some of the most basic principles in statistics. Almost no initial COVID-19 studies listed in Table 2 performed any model checking or published their code for reproducibility. Many studies and forecasting models relied on results from other studies, but they rarely considered uncertainty in the prior estimates. Almost all media coverage on the COVID-19 studies only reported point estimates and did not mention any form of uncertainty quantification. Finally, there were almost no serious considerations of selection bias in the initial COVID-19 studies.

Lesson 3: Better research appraisal is in urgent need

COVID-19 has presented a serious challenge to scientific journals and there were some encouraging moves. For example, many notable publishers, journals and institutions signed a statement on 31 January 2020 to make all publications related to COVID-19 immediately open access (Wellcome Trust, 2020). However, journals were quickly overwhelmed with submissions related to COVID-19 and the peer reviews were very unreliable. The high volume of research publications on COVID-19, often showing inconsistent results, further caused confusion. In the end, extra effort is often needed to appraise and synthesize the research studies.

Statisticians can contribute to improving this process in several ways. First, statisticians need to engage in applied research more actively and disseminate the basic statistical principles. Second, statisticians can take a bigger role in peer reviews and post-publication discussion of scientific studies. Many academic journals allow the manuscript type of 'letter to the editor' to discuss published articles and some journals also allow and moderate online comments. However, they are rarely used by or even known to statisticians. Finally, better methods for research synthesis are needed to take data quality and selection bias into account.

Lesson 4: Right data are often more important than right analysis

I would like to close this retrospect by going back further into history. In statistics education, it is common to assume that high-quality data are already available and are just 'waiting to be analysed'. But this is rarely the case in practice. The modern discipline of epidemiology is often traced

back to John Snow's inspiring investigation of the London cholera epidemic in 1854. He refuted the then popular miasma theory and identified that the source of the outbreak was a public water pump on Broad Street in London. From a statistical perspective, Snow's most innovative idea was to focus on death rates in districts served by two water companies that drew water from different stretches of the Thames River. The huge difference in death rates due to cholera provided the convincing evidence that germ-contaminated water was the cause of the epidemic (Winkelstein, 1995). By any standard, this is a remarkable achievement because Snow, unlike us, had no access to all the tools offered by modern microbiology and statistics.

Interestingly, a group of modern researchers recovered historical archives in Ferrara, Italy, which also suffered from another cholera epidemic in 1855 (Scapoli et al., 2002). They analysed the data in these archives using modern statistical methods and found that 'the better kept houses in the better parts of the town had less cholera morbidity and especially mortality', supporting a miasma theory. As pointed out by a discussant of the paper, this is not entirely surprising: the Ferrara data were recorded under the misguided miasmas theory and can never reveal the superiority of the competing germ theory as in Snow's analysis Vandenbroucke (2002).

The same thing can be said about COVID-19. In the initial investigations of human-to-human transmissibility, the main obstacle is a lengthy process of reporting and testing suspected cases (Section 2.2). Although a due process with convincing proofs is certainly needed to confirm transmissibility, more stringent public health interventions could have been imposed sooner if the right data were collected, as argued in Section 2.3. This teaches us another lesson in the era of data: statisticians should not be complacent with being the best analysts of 'big data'; instead, we need to step up and play a bigger role in study design and data collection.

AFTERWORD

The above article was initially written in December 2020. Back then, the United Kingdom, where I live, was heading to a third national lockdown. This article was then revised in June 2021 and prospects of the pandemic became a lot less grim due to successful vaccines. But unfortunately, the origins of COVID-19 continue to be politicized. When writing the article, it was difficult for me to not imagine a counterfactual scenario in which Wuhan was locked down a week earlier, even though I think the factual lockdown was already an incredibly resolute decision. Could COVID-19 be contained before turning into a pandemic? Could it buy more time for producing personal protective equipment and vaccine development? There are many more hypothetical questions one could ask, but the answers will never be known. Eventually, I reconciled myself through the history of infectious diseases and the lens of evolutionary biology. New kinds of viruses will eventually appear and circulate, and humanity will face greater challenges and need greater solidarity. What is important is not to find fault with others, but to learn from each others' failures.

ACKNOWLEDGEMENTS

Qingyuan Zhao was partially supported by an early career grant from the Isaac Newton Trust. The author is grateful to the Discussion Papers Editor, Dr. Adam Sykulski, and an anonymous referee for carefully reading a previous version of the manuscript and providing many helpful suggestions. Some of the references are originally in Chinese and the list can be found in the preprint of this article available from the author's personal website.

ORCID

Qingyuan Zhao  <https://orcid.org/0000-0001-9902-2768>

REFERENCES

- Bacallado, S., Zhao, Q. & Ju, N. (2020) Letter to the editor: generation interval for covid-19 based on symptom onset data. *Eurosurveillance*, 25(29). Available from: <https://doi.org/10.2807/1560-7917.es.2020.25.29.2001381>
- Badian Health News. (2020) Status does not match its name, things you do not know about the past of Chinese center for disease control, 28 February. Available from: <https://www.tmtpost.com/4259786.html> [Accessed 26th November 2020].
- Chinese Centre for Disease Control. (2020) On the front line, disease control warriors race against the new coronavirus, 1 February. Available from: http://www.chinacdc.cn/yw_9324/202002/t20200201_212137.html [Accessed 28th November 2020].
- Chinese Health Commission. (2007) National surveillance, investigation and management plan for pneumonia cases of unknown etiology. Available from: <http://www.nhc.gov.cn/bgt/pw10708/200708/4455f46a2f5e4908a8561c079ecbcf0e.shtml> [Accessed 28th November 2020].
- Crampton, T. (2003) W.H.O. criticizes China over handling of mystery disease. *The New York Times*, 7 April. Available from: <https://www.nytimes.com/2003/04/07/international/asia/who-criticizes-china-over-handling-of-mystery-disease.html> [Accessed 22nd October 2020].
- Gao, Y. (2020) Rashomon about contact with the Huanan seafood market: Wuhan Municipal Health Commission's 'double standard' is confusing. *Caixin*, 19 February. Available from: <https://www.caixin.com/2020-02-19/101517544.html> [Accessed 28th November 2020].
- Gao, Y., Peng, Y., Yang, R., Feng, Y. & Ma, D. (2020) Origin of the genetic sequencing of the novel coronavirus: when did the alarm first sound? *Caixin*, 26 February (the original post has been removed). Available from: <https://archive.li/YylMt> [Accessed 27th November 2020].
- Ge, X.-Y., Li, J.-L., Yang, X.-L., Chmura, A.A., Zhu, G., Epstein, J.H. et al. (2013) Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature*, 503(7477), 535–538. Available from: <https://doi.org/10.1038/nature12711>
- Imai, N., Cori, A., Dorigatti, I., Baguelin, M., Donnelly, C.A., Riley, S. et al. (2020a) Report 3—transmissibility of 2019-nCoV. Technical report, MRC Centre for Global Infectious Disease Analysis, 25 January.
- Imai, N., Dorigatti, I., Cori, A., Riley, S. & Ferguson, N.M. (2020b) Report 1—estimating the potential total number of novel Coronavirus (2019-nCoV) cases in Wuhan City, China. Technical report, MRC Centre for Global Infectious Disease Analysis, 17 January.
- Jakes, S. (2003) Beijing's SARS attack. *TIME*, 8 April. Available from: <http://content.time.com/time/magazine/article/0,9171,441615,00.html> [Accessed 22nd October 2020].
- Kremer, C., Ganyani, T., Chen, D., Torneri, A., Faes, C., Wallinga, J. et al. (2020) Authors' response: estimating the generation interval for covid-19 based on symptom onset data. *Eurosurveillance*, 25(29). Available from: <https://doi.org/10.2807/1560-7917.es.2020.25.29.2001269>
- Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y. et al. (2020a) Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *New England Journal of Medicine*, 382, 1199–1207.
- Li, X., Li, M., Peng, D. & Du, W. (2020b) Regret of Wuhan: how was the golden period to contain the epidemic missed? *Chinese News Weekly*, 10, February (the original web link has been removed). Available from: <https://archive.ph/SPwN3> [Accessed 2nd December 2020].
- Liao, J. & Li, C. (2020) Experts say Wuhan's unexplained viral pneumonia is preventable and controllable. *Xinhua News Agency*, 11 January. Available from: http://www.xinhuanet.com/local/2020-01/11/c_1125448549.htm [Accessed 26th November 2020].
- Ni, W. (2020) Gao Fu, director of the centers for disease control and prevention: tell everyone with confidence that China's vaccines are very good! *The Beijing News*, 4 March. Available from: <http://www.bjnews.com.cn/feature/2019/03/04/552608.html> [Accessed 26th November 2020].
- Pellis, L., Scarabel, F., Stage, H.B., Overton, C.E., Chappell, L.H.K., Fearon, E. et al. (2021) Challenges in control of COVID-19: short doubling time and long delay to effect of interventions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1829). Available from: <https://doi.org/10.1098/rstb.2020.0264>
- Qin, Z. (2020) After being discharged from hospital, Dr Wang Guangfa from Health Commission's expert team answers our 8 questions. *Chinese Youth Daily—Freezing Point*, 2 February. Available from: https://mp.weixin.qq.com/s/DWcRVz10zps27VlrmI_Khg [Accessed 2nd December 2020].

- Qu, T. (2020) Experts say it is a new type of coronavirus, and preliminary progress has been made in the etiological identification of the unexplained viral pneumonia epidemic in Wuhan. Xinhua News Agency, 9 January. Available from: http://www.xinhuanet.com/2020-01/09/c_1125438971.htm [Accessed 28th November 2020].
- Read, J.M., Bridgen, J.R.E., Cummings, D.A.T., Ho, A. & Jewell, C.P. (2021) Novel coronavirus 2019-nCoV (COVID-19): early estimation of epidemiological parameters and epidemic size estimates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1829). Available from: <https://doi.org/10.1098/rstb.2020.0265>
- Sanche, S., Lin, Y.T., Xu, C., Romero-Severson, E., Hengartner, N. & Ke, R. (2020) High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. *Emerging Infectious Diseases*, 26(7), 1470–1477. Available from: <https://doi.org/10.3201/eid2607.200282>
- Scapoli, C., Guidi, E., Angelini, L., Stefanati, A. & Gregorio, P. (2002) Sociomedical indicators in the cholera epidemic in Ferrara of 1855. *European Journal of Epidemiology*, 18(7), 617–622. Available from: <https://doi.org/10.1023/a:1024829328204>
- Southern Metropolis Daily. (2020) Details disclosure! the process of diagnosing Guangdong's first 'morbid' family is made public. 25 February. Available from: <https://wemp.app/posts/c483336b-1a74-49eb-a447-4321400a28c8> [Accessed 26th November 2020].
- Standing Committee of the National People's Congress. (2004) Law of the People's Republic of China on prevention and treatment of infectious diseases (2013 amendment). Available from: http://www.gov.cn/banshi/2005-08/01/content_19023.htm; English translation in Available from: <http://www.lawinfochina.com/display.aspx?id=14881&lib=law> [Accessed 28th November 2020].
- Sun, M. (2020) Soon before easing the lockdown, Dr Li Lanjuan for the first time discloses details of Wuhan's lockdown. Chinese Health Magazine, 26 March. Available from: https://mp.weixin.qq.com/s/1MQDiyfJt9YA9IK5b_i-A [Accessed 2nd December 2020].
- Vandenbroucke, J.P. (2002) Commentary: the 1855 cholera epidemic in Ferrara: lessons from old data reanalysed with modern means. *European Journal of Epidemiology*, 18(7), 595–598. Available from: <https://doi.org/10.1023/a:1024984102096>
- Wallinga, J. & Lipsitch, M. (2006) How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences*, 274(1609), 599–604. Available from: <https://doi.org/10.1098/rspb.2006.3754>
- Wei, F., Du, M. & Zhang, Y. (2020) Why did the direct network reporting system fail in the early stage of the Wuhan epidemic? *The Beijing News*, 14 March. [Accessed 1st December 2020].
- Wellcome Trust. (2020) Sharing research data and findings relevant to the novel coronavirus (covid-19) outbreak. Available from: <https://wellcome.org/press-release/sharing-research-data-and-findings-relevant-novel-coronavirus-ncov-outbreak> [Accessed 10th December 2020].
- Winkelstein, W. (1995) A new perspective on John snow's communicable disease theory. *American Journal of Epidemiology*, 142(Supplement 9), S3–S9. Available from: https://doi.org/10.1093/aje/142.supplement_9.s3
- World Health Organization. (2003a) Coronavirus never before seen in humans is the cause of SARS, 16 April. Available from: <https://www.who.int/mediacentre/news/releases/2003/pr31/en/> [Accessed 21st October 2020].
- World Health Organization. (2003b) Consensus document on the epidemiology of severe acute respiratory syndrome (SARS). Technical report, World Health Organization.
- World Health Organization. (2015) Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003. Available from: https://www.who.int/csr/sars/country/table2004_04_21/en/ [Accessed 21st October 2020].
- World Health Organization. (2017) WHO public health research agenda for influenza: minimizing the impact of pandemic, zoonotic, and seasonal epidemic influenza. Technical report, World Health Organization.
- World Health Organization. (2020) Timeline of WHO's response to COVID-19. Available from: <https://www.who.int/news/item/29-06-2020-covidtimeline> [Accessed 21st October 2020].
- Wu, J. (2020) Lost opportunities: neglected cases in small hospitals during the early coronavirus outbreak. *Badian Health News*, 11, February. Available from: <https://new.qq.com/rain/a/20200211A05AEA00> [Accessed 1st December 2020].
- Wu, J.T., Leung, K. & Leung, G.M. (2020) Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet*, 395(10225), 689–697. Available from: [https://doi.org/10.1016/s0140-6736\(20\)30260-9](https://doi.org/10.1016/s0140-6736(20)30260-9)

- Wuhan Health Commission. (2020a) Situation report on viral pneumonia of unknown etiology, 5 January. Available from: http://wjw.Wuhan.gov.cn/xwzx_28/gsgg/202004/t20200430_1199589.shtml [Accessed 26th November 2020].
- Wuhan Health Commission. (2020b) Experts explain the latest situation report on viral pneumonia of unknown etiology, 11 January. Available from: http://wjw.Wuhan.gov.cn/xwzx_28/gsgg/202004/t20200430_1199589.shtml [Accessed 26th November 2020].
- Wuhan Health Commission. (2020c) Q&A about the novel coronavirus pneumonia epidemic, 15 January. Available from: http://wjw.Wuhan.gov.cn/xwzx_28/gsgg/202004/t20200430_1199594.shtml [Accessed 26th November 2020].
- Xiao, H. (2020) Intensive care doctor describe: how do we rescue critically ill patients? Caixin, 5 February. Available from: <http://china.caixin.com/2020-02-05/101511802.html> [Accessed 26th November 2020].
- Xiaoshangou (online id). (2020) Record the experience of discovering the new coronavirus for the first time, 28 January (the original post has been removed). Available from: <https://project-gutenberg.github.io/nCovMemory-Web/post/a544e7b4ac5d5ccf01e8259a70b95044> [Accessed 27th November 2020].
- Xin, N., Wang, X., Sun, A. & Xin, Y. (2020) Why the direct reporting system for infectious diseases with an investment of rmb 730 million failed for 28 days? Caijing, 25 February. Available from: https://mp.weixin.qq.com/s/?__biz=MzA5NTM4Njk2Mw==&mid=2665968779&idx=1&sn=3fdb7a3736f53d81a139b4a15bb27a7&scene=21 [Accessed 2nd December 2020].
- Xu, B. (2020) The mystery of zero new cases in 12 days after 6th of January. Yicai, 28 January. Available from: <https://www.yicai.com/news/100485217.html> [Accessed 28th November 2020].
- Yang, H. (2020a) The white handbook and the green handbook: Changes in diagnostic criteria for the novel coronavirus pneumonia. Chinese Youth Daily–Freezing Point, 20 February Available from: <https://mp.weixin.qq.com/s/vysNta8IU2wbRBv-c3aS4Q> [Accessed 28th November 2020].
- Yang, H. (2020b) Why was the case report once suspended during the early outbreak in Wuhan? Chinese Youth Daily–Freezing Point, 5 March. Available from: <https://xw.qq.com/cmsid/20200305A003MV00> [Accessed 1st December 2020].
- Yang, H. (2020c) Doctor from the Wuhan Central Hospital: the infectious disease leaves us too little time to respond. Chinese Youth Daily–Freezing Point, 13 March. Available from: <https://terminus2049.github.io/archive/2020/03/13/bing-dian.html> [Accessed 26th November 2020].
- Yang, N. (2020d) ‘Reorganization’ Jinyintan: Secrets in the eye of the epidemic typhoon. Southern Weekly, 5 March. Available from: <https://www.infzm.com/contents/178385> [Accessed 28th November 2020].
- Yang, Y., Lu, Q.-B., Liu, M.-J., Wang, Y.-X., Zhang, A.-R., Jalali, N. et al. (2020) Epidemiological and clinical features of the 2019 novel coronavirus outbreak in China. *medRxiv*. Available from: <https://doi.org/10.1101/2020.02.10.20021675>.
- Zhao, Q., Chen, Y. & Small, D.S. (2020a) Analysis of the epidemic growth of the early 2019-nCoV outbreak using internationally confirmed cases. *medRxiv*. Available from: <https://www.medrxiv.org/content/early/2020/02/09/2020.02.06.20020941>.
- Zhao, S., Lin, Q., Ran, J., Musa, S.S., Yang, G., Wang, W. et al. (2020b) Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: a data-driven analysis in the early phase of the outbreak. *International Journal of Infectious Diseases*, 92, 214–217. Available from: <https://doi.org/10.1016/j.ijid.2020.01.050>.
- Zhao, Q., Ju, N., Bacallado, S. & Shah, R.D. (2021) BETS: the dangers of selection bias in early analyses of the coronavirus disease (COVID-19) pandemic. *Annals of Applied Statistics*, 15(1), 363–390.

How to cite this article: Zhao, Q. (2022) Small data, big time—A retrospect of the first weeks of COVID-19. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1793–1814. Available from: <https://doi.org/10.1111/rssa.12874>

Proposer of the vote of thanks and contribution to the Discussion of ‘Statistical Aspects of the Covid-19 Pandemic’

Ruth Studley

Office for National Statistics, London, UK

Correspondence: Ruth Studley, Office for National Statistics, London, UK.

Email: ruth.studley@ons.gov.uk

Having read these papers by such distinguished colleagues I believe the strength of this work is in the complement, what we can learn from the combination pulled together.

A key benefit of the proposal by Jiang et al. is to provide an intuitive prediction of the near future which could have a phenomenal impact on decision-making. There is considerable detail on the methodology and I recognise the difficulties of calculating changing growth rates and note this is further complicated as contextual factors continually change. Here, I would like to focus on the underlying assumptions of the data and any potential limitations this may have to influence decisions.

We do know a sizeable proportion of Covid-19 cases are asymptomatic, that is those who test positive who do not exhibit any symptoms or who do not necessarily report the presence of any symptoms. The inclusion of these cases would be critical to being able to more accurately predict the short term trends. I look forward to seeing an update of these results with more recent data reflecting later stages in the pandemic and the behaviours of the different variants.

Again, I am keen to understand the underlying data assumptions in Nason et al. A key tenet is the movement between countries and two options are proposed: nearest neighbour and equal connectivity. Going forward it will be interesting to understand the impact of the closed air and/or land borders during the pandemic on these options, and I look forward to seeing how this influences the results.

Taking this work forward a useful option will be the ability to provide more detailed regional information within a country where differences in non-pharmaceutical interventions may not be so different as across countries, and potentially have less of an impact on the findings.

I note this paper highlights that including additional exogenous information improves the forecasting of PMI. I applaud using innovative approaches such as including death rates when evidencing economic indicators. Going forward refining these to include age-related indicators and asymptomatic cases, as above, will be interesting to review.

Zhao provides a useful summary of the context for all this research. It struck me in some areas this paper is necessarily emotive and we do well to read this in that light. The empathy that is apparent within this paper is, I suspect, a key driver for so many involved in Covid-19 analysis to

add the value (analytical or otherwise) in these difficult times. Being able to convey this emotion in a statistical paper is impressive.

An interesting point is made of the Chinese Centre for Disease Control already knowing the causative pathogen was likely a coronavirus. Reflecting on this I questioned how often we have the findings within the data but struggle to determine signal from noise.

As already mentioned it is great this paper brings out the importance of understanding the asymptomatic cases and the impact they can have on the ability to calculate various metrics.

The point on the uncertainty of the estimates is well made and how much value can be gained from the estimation. New knowledge suggests these were underestimated. Understanding how to message findings with uncertainty is always a challenge and one we keep learning about it. Lessons can be learned for the future and Zhao shares his thoughts on how statisticians can better contribute to these important events.

I think there are some overarching themes from these very different papers, all considering different aspects of the pandemic. My takeaway points are:

1. We know as statisticians it is incumbent upon us to behave with integrity at all times. Being transparent with findings and underlying methods allows us and others to continually learn and improve. The publication of these papers is testament to this.
2. During times the pandemic, and in other times of crisis, innovation is more important than ever. Continually reviewing methods and considering new ways of working, whether that be to tweak existing approaches or doing something more imaginative, are all crucial activities. Sharing these new things with confidence and humility will simultaneously generate useful advice and stimulate others.
3. We need to be there. We are better statisticians if we are there 'in the moment'. Not only thinking explicitly about small data (being able to distinguish noise against signal using accurate data), and communicating effectively about findings and methods. But most importantly, by being there we can be involved in the design of future analyses and we will have a fundamentally better understanding of the complex context within which we conduct our work.

To summarise I have found these papers inspiring. Collectively they demonstrate how viewing life through different lenses brings together innovation. I am grateful for this opportunity to share my admiration to all those involved by proposing a sincere vote of thanks.

How to cite this article: Studley, R. (2022) Proposer of the vote of thanks and contribution to the Discussion of 'Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1815–1816. Available from: <https://doi.org/10.1111/rssa.12920>

Secunder of the vote of thanks and contribution to the ‘First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic’

Sheila M. Bird

MRC Biostatistics Unit, University of Cambridge, Cambridge, UK

Correspondence: Sheila M. Bird, MRC Biostatistics Unit, University of Cambridge, Cambridge, UK.

Email: sheila.bird@mrc-bsu.cam.ac.uk

Two of this evening’s papers are retrospective: modelling COVID-19 infection trajectories using a piecewise linear quantile trend model and the first weeks of COVID-19 in China – small data, big time. The third is economic. Intriguingly, in forecasting Purchasing Managers’ Indices (scored 0–100), Nason and Wei introduce COVID-19 mitigation stringency indices and monthly COVID-19 death-rates per million of population as time-series of exogenous variables in their Generalised Network Autoregressive model.

All three papers concern on how pandemic-related events evolve over time but, in infectious disease epidemics, the event-time which matters is occurrence-date. Prior to both Delta-variant and immunization, one-third of SARS-CoV-2 infections were never symptomatic. Since symptom-onset-date is not generally applicable, focus falls on swab-date.

Next steps after the swab is taken are: receipt by diagnostic laboratory, laboratory-analysis, swab-result reported-back to citizen, swab-positive-results reported-in centrally, daily count of newly-diagnosed SARS-CoV-2 infections reported to the media; additionally, report-date to the World Health Organization (WHO).

However, WHO-date, media-date and swab-date for new SARS-CoV-2 infections are seldom the same. When analysing COVID-19 infection trajectories internationally, the epidemiological-date that matters is swab-date: neither media-date nor WHO-report-date.

The piecewise linear quantile trend model by Jiang et al. can, of course, be implemented on a diversity of event-scales. But the scale closest to when the pandemic takes citizens out-of-circulation is **swab-date**. Others are confounded by a series of delays: country-specific, time-varying or both.

No10 press conferences during the first wave of COVID-mention deaths in England and Wales displayed a **date-scale** on which COVID-mention hospitalised deaths evolved. **Statisticians**, but *neither the media nor the general public*, immediately knew that report-date, **not death-date**, was the undisclosed date-scale. **Why?** Because the tell-tale dropping-off of reports for the most recent few days was not in evidence!

In late March 2020, the Department of Health and Social Care (DHSC) began to release publicly the death-dates for each day’s daily-reported COVID-mention deaths in hospitals in England. Hence, Bent Nielsen and I were able to estimate the reporting-delay distribution and put into

the public domain (via Science Media Centre) our nowcasting of hospitalised COVID-deaths in English regions and by age-group, see <http://users.ox.ac.uk/~nuff0078/Covid/index.htm>. Fellow scientists were grateful. Modellers' more sophisticated work, which informed SAGE, was not then routinely in the public domain but now is, for example: <https://www.mrc-bsu.cam.ac.uk/nowcasting/nowcasting-and-forecasting-7th-september-2021/>.

Progress in RSS's campaign to end the late registration of deaths in England, Wales and Northern Ireland has included: correct labelling of table-legends and time-axes by the Office for National Statistics (ONS) and provision by NHS Digital of registration-date as well as death-date when cohorts are flagged for mortality. As National Statistician, Sir Ian Diamond resolved to sort the issue but the pandemic supervened. Importantly, ONS reports COVID-mention deaths by occurrence-date, as well as by registration-date. For example, *England's first COVID-mention death occurred on 30 January 2020 but was not registered as such until September 2020 after this 84-year-old gentleman's family, who suspected COVID-19, asked for his autopsy to be reviewed.*

Jiang et al. do not model the 'COVID-19 infection trajectory' – rather the trajectory of report-dates for newly diagnosed SARS-CoV-2 cases. **Report-dates are confounded** by reporting-delays which vary internationally, suffer weekend and vacation perturbations, which contribute to heteroscedascity.

The authors' historical piecewise linear quantile trend model is useful in segmenting time-series into periods with relatively stable behaviour. But, inevitably, forecasts are generated from observations in the last segment where, in practice, variance may itself change: for example, due to variant-pressure on laboratories. Neatly, the authors propose 80% confidence intervals based on their forecasting ahead for the 10th, median and 90th quantiles.

There is much to like in Nason and Wei's quantification of national economic response to the stringency of COVID-19 mitigations (highly influential) and death-rates per 1 million of population (less so) via Purchasing Managers' Indices (PMIs) using Generalised Network Auto-Regressive models with exogenous variables (themselves time-series). Specifically, parsimony (order of N not N^2), dynamic networks (weighted by exports between countries), missingness managed (nearest neighbours with data 'represent' those without data), global versus local alpha (global better), experimentation enabled (e.g. fully connected trade network vs. nearest neighbour(s), say: country to which UK exports most).

But I am supposed to be critical. Hence, the public's and Purchasing Managers' response to stringency measures (e.g. lockdown) may be different second or third time around versus during the analysed first wave; available covariates are not necessarily the best – I note with interest that stringency measures, an early intervention, are more explanatory than registration-delayed mortality series. Vaccination or variant can disrupt the age-related patterns of progression from SARS-CoV-2 infection to death that applied during the first wave. Age-structure in network-countries matters: thus, COVID-mention deaths per million of population has different meaning in United Kingdom (ageing population) versus India or Africa. A week is a long time in politics: a month of fatalities is a very long time in this pandemic. Finally, PMIs were computed by aggregating the views of senior purchasing executives from around 400 companies: 400 in total or per network country; response-rate; how were executive selected and refreshed; and concordance between PMI-scores from trading-neighbours?

Finally, reporting delays in Hubei feature also in Zhao's account of 'Small Data, Big Time'. This paper was completed before infectious diseases physician, Sir Jeremy Farrar, published his inside story, **SPIKE, The virus vs the People** – in which he recounts how international scientists collaborated, and supported each other in the face of potential sanctions, in 'outing' quickly not only the sequence for SARS-CoV-2 which enabled polymerase chain reaction diagnostic tests to

be immediately developed but also critical early information about asymptomatic transmission of the virus.

Media – whether printed or social – can either be sampled representatively (with difficulty) or included selectively for illustrative purposes. **I think that Zhao has chosen the latter** but greater clarity on this aspect would be welcome.

My second remark relates back to **SPIKE**: early appreciation of a new infectious disease requires a diversity of experts: infectious disease clinicians, virologists, immunologists, veterinarians (if zoonosis is suspected), parasitologists in addition to infectious disease modellers and biostatistical expertise. **Statisticians do not work alone.** If we do, **we miss out on** anticipatory insights that other basic sciences bring to the table well before data emerge; and **they miss out on** what the theory of infectious diseases portends. Experts matter ... for their anticipation.

President, ladies and gentlemen, we have enjoyed a fascinating trio of read papers. I congratulate all tonight's co-authors and have great pleasure in seconding the vote of thanks.

The vote of thanks was passed by acclamation.

How to cite this article: Bird, S.M. (2022) Secunder of the vote of thanks and contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1817–1819. Available from: <https://doi.org/10.1111/rssa.12921>

DOI: 10.1111/rssa.12922

Tao Wang's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Tao Wang

Department of Economics, University of Victoria, Victoria, British Columbia, Canada

Correspondence: Tao Wang, Department of Economics, University of Victoria, Victoria, British Columbia, V8W 2Y2, Canada.

Email: taow@uvic.ca

I congratulate the authors for devising a novel segmentation algorithm that has enabled them to present a piecewise linear quantile trend model to analyse the trajectory of the COVID-19 daily new cases, which is resistant to outliers, exhibits heteroscedasticity, and effectively captures both point and interval forecasts. To broaden our understanding of the COVID-19 data and the piecewise linear quantile modelling, I would like to highlight some potential developments.

My main query concerns applicability in terms of COVID-19 data. In spite of the fact that quantile regression can be utilized to capture heteroscedasticity, in reality, individuals may

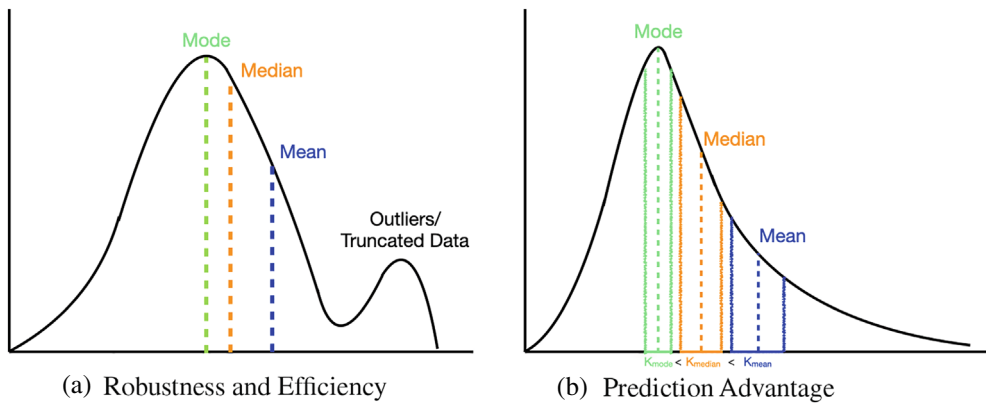


FIGURE 1 Features of modal estimation (a) Mode is robust to outliers and modal estimator can be more efficient than mean and median estimators; (b) Modal prediction tends to produce a shorter prediction interval compared to mean and median predictions, provided that the same coverage probability is given. [Colour figure can be viewed at wileyonlinelibrary.com]

be more concerned with the predicted mean value for COVID-19 data. The problem, however, is that the accuracy of prediction is largely dependent on the reliability of data, and among the scientific community, it is generally believed that the current official COVID-19 data are often noisy with outliers, biased, skewed and/or truncated (Linton et al., 2020). When using quantile regression to capture trends, efficiency may be sacrificed in order to achieve robustness. Due to this, among several potential developments, of particular interest is the extension to a robust mean regression context with the proposed piecewise linear quantile trend model to improve the estimation efficiency of its local least squares counterpart. For instance, one could follow Kai et al. (2010) to propose a composite piecewise linear quantile trend model when the error distribution is non-normal. The core principle is to take the average over all quantiles, that is, $\hat{\beta}_0 = \frac{1}{M} \sum_{m=1}^M \hat{\beta}_0(\tau_m)$. This kind of estimator is not only robust to aberrant observations, but can even be employed when the conditional variance is infinite.

Alternatively, the model can be extended to the mode-based or modal regression content (Ullah et al., 2022). It is well known that with symmetric data, the mode is identical to the mean. One can utilize mode to capture the mean estimate, which is highly resistant to extreme values or heavy-tailed distributions. Furthermore, the mode can be used for prediction and provides shorter prediction intervals compared to the mean or median. Ohta et al. (2019) suggested a new quantile regression-based conditional modal estimator. The procedure relies on the fundamental mechanism that the derivative of the conditional quantile function with respect to the quantile index is the reciprocal of the conditional density evaluated at the conditional quantile function. The conditional mode is then obtained by minimising the derivative of the conditional quantile function. It can be mathematically expressed by

$$\tau_x = \arg \min_{\tau \in (0,1)} s_x(\tau) = Q_x^{(1)}(\tau) := \frac{\partial Q_x(\tau)}{\partial \tau} = \frac{1}{f(Q_x(\tau)|x)},$$

where $\tau \in (0, 1)$ and $Q_x(\tau)$ represents the conditional τ -quantile of Y given X . Along these lines, the proposed piecewise linear quantile trend model may be readily extended to the corresponding

mode-based or modal regression (Figure 1). The estimate technique for the unknown number m_0 and location q of the change-points in the model should be followed straightforwardly.

REFERENCES

- Kai, B., Li, R. & Zou, H. (2010) Local CQR Smoothing: An Efficient and Safe Alternative to Local Polynomial Regression. *Journal of the Royal Statistical Society Series B*, 72(1), 49–69.
- Linton, N.M., Kobayashi, T., Yang, Y., Hayashi, K., Akhmetzhanov, A.R., Jung, S.M. et al. (2020) Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data. *Journal of Clinical Medicine*, 9(2), 538.
- Ohta, H., Kato, K. & Hara, S. (2019) Quantile Regression Approach to Conditional Mode Estimation. *Electronic Journal of Statistics*, 13, 3120–3160.
- Ullah, A., Wang, T. & Yao, W. (2022) Nonlinear modal regression for dependent data with application for predicting COVID-19. *Journal of the Royal Statistical Society Series A*, 185(3), 1424–1453.

How to cite this article: Wang, T. (2022) Tao Wang's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1819–1821. Available from: <https://doi.org/10.1111/rssa.12922>

DOI: 10.1111/rssa.12923

Xiaoping Shi, Yue Zhang and Yucheng Dong's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Xiaoping Shi¹ | Yue Zhang¹ | Yucheng Dong²

¹Department of Computer Science, Mathematics, Physics and Statistics, University of British Columbia, Kelowna, British Columbia, Canada

²Center for Network Big Data and Decision-Making, Business School, Sichuan University, Chengdu, China

Correspondence: Xiaoping Shi, Department of Computer Science, Mathematics, Physics and Statistics, University of British Columbia, Kelowna, BC, Canada.

Email: xiaoping.shi@ubc.ca

The authors Jiang et al. (2021) are to be congratulated on an important contribution, which is the multiple change point detection in the piecewise linear quantile trend model. There is an extension of the proposed model in (1) to the nonlinear quantile trend model like

$$Q_{\tau}(y_t) = f\left(t/n; \beta_1, \dots, \beta_q\right),$$

where f is a non-linear function of t/n with parameters $\beta_1, \dots, \beta_q$.

There are two reasons for estimating the change points in the COVID-19 trajectory using a non-linear model. Firstly, starting from 11 March 2020, when the WHO declared the outbreak

a pandemic, many countries have experienced multiple infection waves in an extended period. Nowadays, despite the increasing vaccination coverage, uncertainties like the spread of the delta variant are driving new waves worldwide, making the spread of COVID-19 an ongoing crisis. Chakraborty and Ghosh (2020) apply the ARIMA model and Wavelet-based model to explore the nonlinear characteristic of COVID-19 cases.

Secondly, the non-linear exponential growth model has shown numerous applications in the modelling of cumulative cases. For example, a recently proposed model by Shi et al. (2021), called clustering-segmented autoregressive sigmoid (CSARS), used change points to describe the COVID-19 trajectory as multiple S-shaped curves.

Then, a natural question comes: can we extend the GOALS algorithm to fit a more general non-linear case? Our answer is yes. We can replace the original two-dimensional quantile estimator of parameters with the q -dimensional one.

The following contributions were received in writing before the meeting.

REFERENCES

- Chakraborty, T. & Ghosh, I. (2020) Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: a data-driven analysis. *Chaos, Solitons and Fractals*, 135, 109850.
- Jiang, F., Zhao, Z. & Shao, X. (2021) Modeling the COVID-19 infection trajectory: a piecewise linear quantile trend model. Available from: https://rss.org.uk/RSS/media/File-library/Events/Discussion%20meetings/Jiang-et-al-_JRSSB_discussion-paper_16-Jul-21.pdf.
- Shi, X., Chen, M. & Dong, Y. (2021) Exploring the space-time pattern of log-transformed infectious count of COVID-19: a clustering-segmented autoregressive sigmoid model, <https://arxiv.org/abs/2102.13287>. <http://graph-clustering-system.com/>.

The following contributions were received in writing after the meeting.

How to cite this article: Shi, X., Zhang, Y. & Dong, Y. (2022) Xiaoping Shi, Yue Zhang and Yucheng Dong's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1821–1822. Available from: <https://doi.org/10.1111/rssa.12923>

Christine P. Chai's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Christine P. Chai

Microsoft Corporation, Redmond, Washington, USA

Correspondence: Christine P. Chai, Microsoft Corporation, Redmond, WA, USA.

Email: cpchai21@gmail.com

STATISTICAL ASPECTS OF THE COVID-19 PANDEMIC

Paper 1: 'Modeling the Covid-19 infection trajectory: a piecewise linear quantile regression'

Authors: F Jiang (Fudan University, Shanghai, China); Z Zhao, (University of Notre Dame, Indiana, USA); X Shao (University of Illinois Urbana-Champaign, USA)

Paper 2: 'Quantifying the economic response to Covid-19 mitigations and death rates via forecasting Purchasing Managers' Indices using Generalised Network Autoregressive models with exogenous variables'

Authors: G Nason & J Wei, Imperial College London, UK

Paper 3: 'Small Data, Big Time - A retrospect of the first weeks of Covid-19'

Author: Q Zhao, Statistical Laboratory, University of Cambridge, UK

Disclaimer: The opinions and views expressed here are those of the author and do not necessarily state or reflect those of Microsoft.

DISCUSSION

In the second paper by Nason and Wei (2021), I appreciate the use of network information in the forecasting models. I also like that the authors used out-of-sample forecasting wherever applicable. I have two questions:

1. What are the advantages and disadvantages of the fully connected trade network versus the nearest neighbour trade network?
2. When incorporating the network information, do we distinguish between leading indicators and lagging indicators?

The third paper by Zhao (2020) brings an interesting perspective regarding COVID-19 from a statistician. I would take this with a grain of salt because the article was written in retrospective, and I am also concerned about potential regulations from the Chinese government. Nevertheless, it is always better to be prepared for the pandemic. But we also need to be cautious in extreme measures because they may cause unnecessary economic loss.

I totally agree with the second lesson in Zhao (2020): "Practitioners are still unfamiliar with basic statistical concepts." Dr. Anthony Fauci mentioned statistical significance in his testimony to the US government in July 2020¹. It is appropriate to report experiment results,

¹<https://www.pbs.org/newshour/health/read-fauci-redfield-and-giroirs-full-witness-testimony-as-submitted-to-the-senate>

but over-emphasising $p < 0.05$ is not a good idea. Ronald Wasserstein and Nicole Lazar have given multiple talks on moving to a world beyond $p < 0.05$ (Wasserstein et al., 2019). The ASA (American Statistical Association) also published an official statement on the context of the p -values (Wasserstein & Lazar, 2016). There is still a long way to go in educating the practitioners on statistical communication.

Last but not least, I really like the conclusion in Zhao (2020): ‘What is important is not to find fault with others, but to learn from each others’ failures’. I hope we can all learn from the mistakes in fighting against the epidemic, so the same disaster would not repeat in the future.

REFERENCES

- Nason, G.P. & Wei, J.L. (2021) Quantifying the economic response to COVID-19 mitigations and death rates via forecasting purchasing managers’ indices using generalised network autoregressive models with exogenous variables. *arXiv preprint arXiv:2107.07605*.
- Wasserstein, R.L. & Lazar, N.A. (2016) The ASA statement on p -values: context, process, and purpose. *The American Statistician*, 70, 129–133.
- Wasserstein, R.L., Schirm, A.L. & Lazar, N.A. (2019) Moving to a world beyond “ $p < 0.05$ ”. *The American Statistician*, 73, 1–19.
- Zhao, Q. (2020) Small data, big time – a retrospect of the first weeks of COVID-19. *Journal of the Royal Statistical Society (Series A, Statistics in Society)*, 1–21.

How to cite this article: Chai, C.P. (2022) Christine P. Chai’s contribution to the ‘First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic’. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1823–1824. Available from: <https://doi.org/10.1111/rssa.12924>

DOI: 10.1111/rssa.12925

Anna L. Choi and Tze Leung Lai’s contribution to the ‘First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic’

Anna L. Choi

Center for Innovative Design, Stanford University, Stanford, California, USA

Correspondence: Anna L. Choi, Center for Innovative Design, Stanford University, Stanford, CA, USA.

Email: annahcchoi@gmail.com

Nason and Wei: The authors introduce a new GNAR model to quantify the economic response to COVID-19 mitigation strategies and economic indicators and to model and forecast the Purchasing Managers’ Indices (PMIs) of 11 countries. Here N represents network, and adding ‘network’ to vector autoregressive VAR models was introduced by Nason and Marina Knight, Kathryn Leeming and Mathew Nunes together with their GNAR

package. Motivated by applications to social networks, Hansheng Wang and Xuening Zhu (PKU) and their collaborators independently introduced their network vector autoregression (NVAR) in 2017. Nason and Wei compare in their supplementary material GNAR and NVAR models. They also extend GNAR to GNARX models, where X indicates a time-dependent exogenous node-specific regressor, defined in their section 2 and illustrated with PMIs data in sections 3.3 and 4. The paper provides a comprehensive development of GNAR and GNARX model that we find to be very useful for other spatiotemporal projects at Stanford's climate and environmental institutes in which we are engaged.

Qingyuan Zhao: The author, whose hometown was Wuhan, received his BSc in Mathematics from University of Science and Technology of China in 2011 and PhD in Statistics in 2016 from Stanford University. Sections 2 and 3.1 of this paper provide excellent reviews of the 'key events and evidence' in the period up to 31 January 2020, including SARS-CoV in 2002-3 and MERS-CoV in 2012; Figures 1–3 are especially noteworthy. In the afterword, he mentioned the difficulty of tracing back the origins of COVID-19, but the prospects of the pandemic has become 'a lot less grim due to the successful vaccines'. This is consistent with our research at Stanford's Center for Innovation in Global Health—developing treatments and response policies in different parts of the world, which is much more important and feasible than tracing the origins of COVID-19.

Jiang, Zhao and Shao: The authors propose a piecewise linear quantile trend model to analyse 'the trajectory of COVID-19 daily new cases (i.e., the infection curve)'. Let R_t denote the daily new cases at times t and $Y_t = \log(1 + R_t)$, they assume that $Q_\tau(Y_t) = \beta_{0t}(\tau) + \beta_{1t}(\tau)t/n$, $t = 1, \dots, n$, where n is the sample size. Although the assumption of piecewise linear trend for a pre-selected set of quantile levels seems overly restrictive, the data shown in Figure 1 show that the model fits the data quite well due to the (R_t, Y_t) transformation. Note that exogenous information, such as availability of vaccines mentioned by Qingyuan Zhao, is not used to improve efficiency of the estimate/forecast of daily infection curves.

How to cite this article: Choi, A.L. (2022) Anna L. Choi and Tze Leung Lai's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1824–1825. Available from: <https://doi.org/10.1111/rssa.12925>

Alex R. Cook, Kwok Pui Choi and Weng Kee Wong's contribution to the 'First Discussion Meeting on Statistical Aspects of the COVID-19 Pandemic'

Alex R. Cook¹ | Kwok Pui Choi² | Weng Kee Wong³

¹Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore

²Department of Statistics and Data Science, Faculty of Science, National University of Singapore, Singapore

³Department of Biostatistics, Fielding School of Public Health, University of California at Los Angeles, Los Angeles, California, USA

Correspondence: Weng Kee Wong, Department of Biostatistics, Fielding School of Public Health, University of California at Los Angeles, Los Angeles, CA, USA.

Email: wkwong@ucla.edu

Zhao's paper (Zhao, 2021) highlights the importance of statistical thinking in the early murk of an emerging outbreak, when uncertainty abounds not just on quantitative measures of spread, but even on the mechanisms of transmission. Not stressed in the paper, but equally important, are uncertainties on the prevalence and role of a-, pre- and pauci-symptomatic transmission, of the role of children, and the severity profile of the emergent disease. In retrospect, would better decisions have been made in the early weeks of 2020 had better data been available? As Zhao demonstrates, even 'small' data can be invaluable if their provenance—how were they identified, what are the inclusion criteria, at what time were the data curated—can be documented and richer information be captured—such as the travel history of the Wuhan couple in Hong Kong he reports. Collectively, we should in the aftermath of the pandemic advocate a standard codification of such case data to permit better analyses such as Zhao's (Zhao et al., 2021).

Beyond better evidence on the transmissibility and severity of the new disease, it strikes us that details of what policies are being implemented in different polities can and must be better shared so that the best decisions could be made. This is especially so for countries those are successful in mitigating the pandemic. We have been aghast at the failure of Western governments to learn from the numerous success stories in East and Southeast Asia (Legido-Quigley et al., 2020) in the early months of the pandemic—failure to control borders, failure to isolate all cases in a hospital (Dickens et al., 2020) or *fangcang* hospital (Fang et al., 2020), failure to establish effective contact tracing (Lewis, 2020). Most of all, failure to treat SARS-CoV-2 as a serious virus. The Oxford policy database (Hale et al., 2021) is an important step in this direction but missed the critical early period of the pandemic when its impact would have been greatest.

While better data ought to facilitate better analyses, we have to accept that—in the early days of a potential pandemic—a perfect analysis, in a month, is the enemy of a good analysis, in a day. Policy making cannot wait for a fully appropriate method to be developed, and in the post-pandemic era, our community should actively engage in developing statistical methods that can rapidly be deployed to overcome the issues raised in Zhao's summary of the nascent COVID-19 pandemic.

REFERENCES

- Dickens, B.L., Koo, J.R., Wilder-Smith, A. & Cook, A.R. (2020) Institutional, not home-based, isolation could contain the COVID-19 outbreak. *The Lancet*, 395, 1541–1542.
- Fang, D., Pan, S., Li, Z., Yuan, T., Jiang, B., Gan, D. et al. (2020) Large-scale public venues as medical emergency sites in disasters: lessons from COVID-19 and the use of Fangcang shelter hospitals in Wuhan, China. *BMJ Global Health*, 5, e002815. <https://doi.org/10.1136/bmjgh-2020-002815>
- Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T. et al. (2021) A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour*, 5, 529–538. <https://doi.org/10.1038/s41562-021-01079-8>
- Legido-Quigley, H., Asgari, N., Teo, Y.Y., Leung, G.M., Oshitani, H., Fukuda, K. et al. (2020) Are high-performing health systems resilient against the COVID-19 epidemic? *The Lancet*, 395, 848–850. [https://doi.org/10.1016/S0140-6736\(20\)30551-1](https://doi.org/10.1016/S0140-6736(20)30551-1)
- Lewis, D. (2020) Why many countries failed at COVID contact-tracing — but some got it right. *Nature*, 588, 384–387. <https://doi.org/10.1038/d41586-020-03518-4>
- Zhao, Q. (2021) Small data, big time—a retrospect of the first weeks of COVID-19. *Journal of the Royal Statistical Society Series A*.
- Zhao, Q., Ju, N., Bacallado, S. & Shah, R.D. (2021) BETS: The dangers of selection bias in early analyses of the coronavirus disease (COVID-19) pandemic. *The Annals of Applied Statistics*, 15, 363–390. <https://doi.org/10.1214/20-AOAS1401>

How to cite this article: Cook, A.R., Choi, K.P. & Wong, W.K. (2022) Alex R. Cook, Kwok Pui Choi and Weng Kee Wong's contribution to the 'First Discussion Meeting on Statistical Aspects of the COVID-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1826–1827. Available from: <https://doi.org/10.1111/rssa.12926>

DOI: 10.1111/rssa.12927

Peter J. Diggle's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Peter J. Diggle

CHICAS, Lancaster Medical School, Lancaster University, Lancaster, UK

Correspondence: Peter J. Diggle, CHICAS, Lancaster Medical School, Lancaster University, Lancaster, UK.

Email: p.diggle@lancaster.ac.uk

My comments relate to the interpretation of the parameters in equation (1) of Nason and Wei's very interesting paper.

Contribution to the discussion of the paper by Nason and Wei

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Author. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* published by John Wiley & Sons Ltd on behalf of Royal Statistical Society.

Firstly, 47 years ago the late Julian Besag presented a paper to the Society (Besag, 1974) that introduced what would now be called Markov random field models in which the joint distribution of a set of random variables, $Y_1, \dots, Y_n$ say, is defined indirectly by specifying the n full conditional distributions of each Y_i given all other Y_j . Besag's paper was set in the specific context of stochastic processes on a regular two-dimensional lattice. However, in seconding the vote of thanks Alan Hawkes made the prescient remark that this approach could be used for 'any multivariate distribution at all' (Hawkes, 1974), 6 years before Darroch et al. (1980) gave the the first explicit discussion of graphical models for multivariate data. Back then, there was considerable, and occasionally heated, debate about the relative merits of defining spatial autoregressive processes via Besag's conditional autoregressive models or simultaneous autoregressive models of the form

$$Y_i = \sum_{j \in \mathcal{N}(i)} \beta_j \omega_{ij} Y_j + u_i. \quad (1)$$

The parameters β_j in (1) are not regression parameters in the usual sense of relating to conditional expectations; specifically, it is not true that the full conditional distribution of Y_i depends only on $\{Y_j : j \in \mathcal{N}(i)\}$. For example, in the model $Y_i = \beta_1 Y_{i-1} + \beta_2 Y_{i+1} + u_i$ the full conditional distribution of Y_i depends on $Y_{i-2}, Y_{i-1}, Y_{i+1}$ and Y_{i+2} . Perhaps for this reason, as well as the convenient link between Markov random field models and MCMC algorithms, SAR models fell out of favour within the world of spatial statistics and, so far as am aware, are nowadays confined largely to the econometrics literature; see, for example, Pace and LeSage (2010).

Secondly, the positioning of the exogenous processes X_t in equation (1) seems to me counter-intuitive. I would have thought it more natural to define the model hierarchically, with (in a simple special case) a marginal model for X_t and a conditional model for $Y_t - X_t' \lambda$ given X_t , so that λ can be interpreted as a set of regression parameters.

REFERENCES

- Besag, J. (1974) Spatial interaction and the statistical analysis of lattice systems (with Discussion). *Journal of the Royal Statistical Society B*, 36, 192–225.
- Darroch, J.N., Lauritzen, S.L. & Speed, T.P. (1980) Markov fields and log linear interaction models for contingency tables. *Annals of Statistics*, 8, 522–539.
- Hawkes, A.G. (1974) Contribution to the discussion of Besag's paper. *Journal of the Royal Statistical Society B*, 36, 225–226.
- Pace, R.K. & LeSage, J. (2010) Spatial econometrics. In: Gelfand, A.E., Diggle, P.J., Fuentes, M. & Guttorp, P. (Eds.) *Handbook of spatial statistics*. Boca Raton: Chapman and Hall/CRC Press, pp. 427, 245–436, 260.

How to cite this article: Diggle, P.J. (2022) Peter J. Diggle's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1827–1828. Available from: <https://doi.org/10.1111/rssa.12927>

Alessio Farcomeni and Marco Geraci's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Alessio Farcomeni¹  | Marco Geraci²

¹University of Rome 'Tor Vergata', Rome, Italy

²Sapienza University of Rome, Rome, Italy

Correspondence: Alessio Farcomeni, University of Rome 'Tor Vergata', Via Columbia 2, 00133 Rome, Italy.

Email: alessio.farcomeni@uniroma2.it

We congratulate Feiyu Jiang, Zifeng Zhao and Xiaofeng Shao on a thorough and interesting work. Their change-point detection procedure for the piecewise linear quantile trend model clearly represents a competitive, innovative and flexible tool for several application areas. We would like to bring the authors' attention to an important issue concerning a potential disconnect between the proposed methods and the specific application discussed. More specifically, while their work is motivated by modelling COVID-19 log-counts $Y_t = \log(R_t + 1)$ (a discrete random variable), assumption 1 in section 3.1 requires the existence of a continuous density function that is bounded away from 0 (and ∞) at the quantile $Q_\epsilon(\tau)$. This assumption is crucial for the Bahadur representation and asymptotic behaviour of $\hat{\beta}(\tau)$ (Koenker, 2005). The good performance of GOALS and M-GOALS in the simulation studies reported by the authors revolves around normally distributed responses. Thus, it is natural to wonder whether the performance deteriorates when the response is discrete, especially when counts are small or irregularly spaced.

Quantile modelling of discrete distributions is notoriously troublesome because of the lack of consistency and asymptotic normality of L_1 estimators in the presence of ties (see, e.g., Jentsch & Leucht, 2016; Ma et al., 2011). In the present context, COVID-19 counts could be easily modelled using a jittering approach (Machado & Santos Silva, 2005), which would involve adding artificial random noise repeatedly to the counts R_t . Owing to some limitations of jittering and other existing approaches to discrete quantile regression, Geraci and Farcomeni (2021) have introduced mid-quantile regression, a conditional extension of Parzen's (1993) marginal mid-quantiles. Not only do mid-quantiles bridge modelling of discrete and continuous distributions, but they also handle different types of discreteness. The conditional mid-quantile estimator is versatile and well-behaved asymptotically (Geraci & Farcomeni, 2021).

We thank the authors for their stimulating work and hope that further contributions to the development of piecewise linear regression will see *conditional mid-quantiles* as a viable solution in change-point estimation problems with discrete responses, such as COVID-19 daily new cases. Modelling and forecast of the latter would also benefit from adjusting models for important confounders like daily number of swabs (e.g., see Alaimo Di Loro et al., 2021; Bartolucci & Farcomeni, 2021; Farcomeni et al., 2021; Girardi et al., 2021), as the authors seem to allude to in their final remarks.

ORCID

Alessio Farcomeni  <https://orcid.org/0000-0002-7104-5826>

REFERENCES

- Alaimo Di Loro, P., Divino, F., Farcomeni, A., Jona Lasinio, G., Lovison, G., Maruotti, A. et al. (2021) Nowcasting COVID-19 incidence indicators during the Italian first outbreak. *Statistics in Medicine*, 40(16), 3843–3864.
- Bartolucci, F. & Farcomeni, A. (2021) A spatio-temporal model based on discrete latent variables for the analysis of COVID-19 incidence. *Spatial Statistics*, 49, 100504.
- Farcomeni, A., Maruotti, A., Divino, F., Jona-Lasinio, G. & Lovison, G. (2021) An ensemble approach to short-term forecast of COVID-19 intensive care occupancy in Italian regions. *Biometrical Journal*, 63(3), 503–513.
- Geraci, M. & Farcomeni, A. (2021) Mid-quantile regression for discrete responses. *Statistical Methods in Medical Research*, 31(5), 821–838.
- Girardi, P., Greco, L. & Ventura, L. (2021) Misspecified modeling of subsequent waves during COVID-19 outbreak: A change-point growth model. *Biometrical Journal*, 64(3), 523–538.
- Jentsch, C. & Leucht, A. (2016) Bootstrapping sample quantiles of discrete data. *Annals of the Institute of Statistical Mathematics*, 68(3), 491–539.
- Koenker, R. (2005) *Quantile regression*. New York, NY: Cambridge University Press.
- Ma, Y., Genton, M.G. & Parzen, E. (2011) Asymptotic properties of sample quantiles of discrete distributions. *Annals of the Institute of Statistical Mathematics*, 63(2), 227–243.
- Machado, J.A.F. & Santos Silva, J.M.C. (2005) Quantiles for counts. *Journal of the American Statistical Association*, 100(472), 1226–1237.
- Parzen, E. (1993) Change PP plot and continuous sample quantile function. *Communications in Statistics - Theory and Methods*, 22(12), 3287–3304.

How to cite this article: Farcomeni, A. & Geraci, M. (2022) Alessio Farcomeni and Marco Geraci's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1829–1830. Available from: <https://doi.org/10.1111/rssa.12928>

Shuyi Ge, Oliver Linton and Shaoran Li's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Shuyi Ge | Oliver Linton | Shaoran Li

Faculty of Economics, University of Cambridge, Cambridge, UK

Correspondence: Oliver Linton, Faculty of Economics, University of Cambridge, Cambridge, UK.

Email: obl20@cam.ac.uk

This paper considers the modelling of network effect in multi-variate time series using generalised network auto-regressive (GNAR) models. The study of network spillover effects is an increasingly important question given we are living in a more and more inter-connected world and the authors have made a useful contribution.

In this paper, the network effect comes in a way that for one country i , its neighbours' dependent variables $y_{q,t-j}$ affect its own dependent variable $y_{i,t}$ only with some lag $j = 1, \dots, p$. Economists often work with 'structural VARs that allow for certain contemporaneous effects. For simplicity, consider $s_j = 1$ for all j (i.e. only the first stage neighbours matter). Then, an overarching network model that nests the spatial auto-regressive effect, the spatial-temporal (lead-lag) effect, the exogenous network effect, and the spatially correlated errors would be

$$Y_{i,t} = \sum_{j=1}^p (\alpha_{i,j} Y_{i,t-j} + \beta_j \sum_{q \in \mathcal{N}^{(i)}} \omega_{i,q} Y_{q,t-j}) + \rho_0 \sum_{q \in \mathcal{N}^{(i)}} \omega_{i,q} Y_{q,t} \\ + \sum_{h=1}^H \sum_{j'=0}^{p'_h} \lambda_{h,j'} X_{h,i,t-j'} + \sum_{h=1}^H \sum_{j'=0}^{p'_h} \rho_{1,h,j'} \sum_{q \in \mathcal{N}^{(i)}} \omega_{i,q} X_{h,q,t-j'} + u_{i,t},$$

where $u_{i,t} = \rho_2 \sum_{q \in \mathcal{N}^{(i)}} \omega_{i,q} u_{q,t}$. This model nests the GNAR model considered in this paper. In this paper's application of forecasting Purchasing Managers' Indices (PMIs), the contemporaneous network effect (the $\rho_0 \sum_{q \in \mathcal{N}^{(i)}} \omega_{i,q} Y_{q,t}$ term) is very likely to be important, and the ruling out of which a priori may not be justified. Following the notation of the paper, if $h = 1$, the structural model has a simple reduced form

$$\mathbf{Y}_t = \rho_0 \mathbf{W} \mathbf{Y}_t + \Phi_1^* \mathbf{Y}_{t-1} + \dots + \Phi_1^* \mathbf{Y}_{t-1} + \Lambda_0 \mathbf{X}_t + \dots, \Lambda_{p'} \mathbf{X}_{t-p'} + \mathbf{u}_t, \quad (1)$$

where $\Phi_k^* = \text{diag}\{\alpha_{i,k} + \beta_k W\}$, $\Lambda_k = \lambda_k I_N + \rho_{1,k} W$, and W is the $N \times N$ network matrix. This spatial auto-regressive model can be easily estimated using QMLE or GMM method.

The second comment we have regards the comparison with the VAR model. A VAR model with $N = 12$ is heavily parameterised, and maybe a fairer game would be to use some regularised VAR models such as Song and Bickel (2011).

The last comment is about the exogeneity of x_1 (the stringency indices). Can this variable really be taken as exogenous? We suspect that countries change their mitigation policies in response to PMIs. When the economy is doing badly, and the economic cost of harsh mitigations might be too high for governments to pursue. In this paper's GNAR model, everything on the right-hand side is weakly exogenous; maybe a generalisation of the model to allow for other endogenous variables (say a three-variable VARX with weakly exogenous network effect) is not too costly?

REFERENCE

Song, S. & Bickel, P. J. (2011) Large vector auto regressions, *arXiv*, 1106.3915.

How to cite this article: Ge, S., Linton, O. & Li, S. (2022) Shuyi Ge, Oliver Linton and Shaoran Li's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1831–1832. Available from: <https://doi.org/10.1111/rssa.12929>

DOI: 10.1111/rssa.12930

Kingsuk Jana, Lagnajita Basu and Kaushik Jana's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Kingsuk Jana | Lagnajita Basu | Kaushik Jana

School of Arts and Sciences, Ahmedabad University, Ahmedabad, Gujarat, India

Correspondence: Kaushik Jana, School of Arts and Sciences, Ahmedabad University, Ahmedabad, Gujarat, India.
Email: kaushik.jana@ahduni.edu.in

We congratulate the authors for their remarkable contributions towards addressing issues of modelling and predicting the COVID-19 infection trajectories. The authors proposed a novel segmentation algorithm for multiple change-point estimation at the multiple quantiles simultaneously. Also, a piece-wise linear quantile trend model is fitted on the basis of those estimated change-points at different quantiles simultaneously. The resulting model gives efficient prediction for the COVID-19 daily new infections.

In the context of high infection modelling, it would be interesting to see how this model can be improvised in the situation where contribution from the extreme quantiles are important (say

at probability level 0.95 or higher). It may be desirable to estimate the change-points which gives proper weight to the quantiles at higher levels. It could also be reasonable to assume that the slope of the quantile trend at the extreme quantiles are different from the quantiles at the bulk part of the distribution.

If the size of the dataset is not very large, then one may find it difficult to estimate the change-points or fitting quantile regression lines at higher quantiles. This problem can be circumvented by using parametric distribution for modelling the extreme observations. For example, one can use the observations over a certain threshold (extreme observations) using Generalised Pareto distribution (GPD). One needs to choose a suitable threshold and by using the fitted GPD one can estimate the quantiles at extreme tails.

Again, we congratulate the authors on such a thought-provoking paper.

How to cite this article: Jana, K., Basu, L. & Jana, K. (2022) Kingsuk Jana, Lagnajita Basu and Kaushik Jana's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1832–1833. Available from: <https://doi.org/10.1111/rssa.12930>

DOI: 10.1111/rssa.12931

Kuldeep Kumar's first contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Kuldeep Kumar

Centre for Data Analytics, Bond University, Robina, Queensland, Australia

Correspondence: Kuldeep Kumar, Centre for Data Analytics, Bond University, Robina QLD, Australia.

Email: kkumar@bond.edu.au

The Purchasing Managers' Index (PMI) is one of the indices which is quite often used to assess the business condition and to predict the direction of change in GDP. In view of the recent COVID19, which has affected the whole world, PMI could be an important indicator to assess the economic condition of a country. In this paper, the authors have tried to forecast PMI using Generalised Network Autoregressive Model with exogenous variables to quantify the economic response due to the COVID-19 mitigations and death rates which has drastically affected quite a few macroeconomic and microeconomic variables like employment, production, new order and supplier deliveries etc. The authors have taken COVID 19 mitigation stringency indices and COVID 19 death rates in the analysis. I personally think Age-Specific Death Rates, as well as infection rates in the demographic window, will be more appropriate than taking overall death rates. It is a bit more complicated as the health outcomes (e.g. the proportion of the population infected) is affected by the level of contact between individual households and the compliance of the households and

firms to public health policies. Health outcomes, in turn, will affect economic outcomes (e.g., gross domestic product) via the level of consumption by households, the government, or imports, and exports with other countries.

It will be interesting to compare the forecast generated by GNARX models with machine learning models like ANN.

ACKNOWLEDGEMENTS

Open access publishing facilitated by Bond University, as part of the Wiley - Bond University agreement via the Council of Australian University Librarians.

How to cite this article: Kumar, K. (2022) Kuldeep Kumar's first contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1833–1834. Available from: <https://doi.org/10.1111/rssa.12931>

DOI: 10.1111/rssa.12932

Kuldeep Kumar's second contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Kuldeep Kumar

Centre for Data Analytics, Bond University, Robina, Queensland, Australia

Correspondence: Kuldeep Kumar, Centre for Data Analytics, Bond University, Robina, QLD, Australia.

Email: kkumar@bond.edu.au

This is a very interesting paper by Dr. Zhao tracing the history of COVID-19 and the lessons we learned from the research done at the very early stage to combat future pandemics (if any, hopefully not!). It is clear from the paper that it was not possible at the early stage to predict the rapid development of this pandemic which engulfed the whole world. Even now, it still seems difficult to predict the impact of this pandemic and how long it will last. Quite often, to predict extraordinary events, we look at the historical analogy, and in the present context, it was the SARS outbreak. But as it is apparent from the paper, this historical analogy did not work for prediction.

Besides various lessons learned from the research studies in the early period of COVID-19, I would like to add a few more. One of the papers which received worldwide attention by Mehra et al. (2020) on the effectiveness of anti-malarial drug hydroxychloroquine published in *Lancet* in the early days and later retracted is worth mentioning in this context. In fact, *The Lancet* has apologised to readers after retracting a study that said the anti-malarial drug hydroxychloroquine

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Author. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* published by John Wiley & Sons Ltd on behalf of Royal Statistical Society.

did not help to curb COVID-19 and might cause death in patients. It may be mentioned that the study was based on big data with information coming from 671 hospitals around the world and the medical records of 96,000 patients. One of the reasons was that the company that provided data would not provide full access to the information for a third-party peer review. This may raise data integrity, and methodological issues as the authenticity of the primary data sources cannot be verified.

LESSON 5

Data-driven research should be performed in accordance with the highest ethical and professional guidelines and rely on data sources that adhere to our high standards. It may be mentioned that the Royal Statistical Society has published a detailed guideline for ethical data science jointly with the Institute and Faculty of Actuaries, and all data scientists should adhere to this.

LESSON 6

The peer-review process should be strengthened in the journals where data-driven research is published, and one of the reviewers should be a data scientist besides a statistician. The journals should specifically ask the reviewer if they have any concern about the research integrity or quality of data in the manuscript they are reviewing which has been introduced by The Lancet.

LESSON 7

Most of the models developed at the early stage, and even now, take only infected cases and not asymptomatic cases. These asymptomatic cases, which we can call 'dark data' have the potential of spreading the pandemic. In all the research, we should consider the possibility of dark data.

ACKNOWLEDGEMENTS

Open access publishing facilitated by Bond University, as part of the Wiley - Bond University agreement via the Council of Australian University Librarians.

REFERENCE

Mehra, M.R., Desai, S.S., Ruschitzka, F. & Patel, A.N. (2020) Retracted: Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *Lancet*, 2020. Available from: [https://doi.org/10.1016/S0140-6736\(20\)31180-6](https://doi.org/10.1016/S0140-6736(20)31180-6)

How to cite this article: Kumar, K. (2022) Kuldeep Kumar's second contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1834–1835. Available from: <https://doi.org/10.1111/rssa.12932>

Shaoran Li, Oliver Linton and Shuyi Ge's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Shaoran Li | Oliver Linton | Shuyi Ge

Faculty of Economics, University of Cambridge, Cambridge, UK

Correspondence: Oliver Linton, Faculty of Economics, University of Cambridge, Cambridge, UK.

Email: obl20@cam.ac.uk

The authors use modern structural change techniques to identify regime changes in a piecewise linear quantile trend regression as

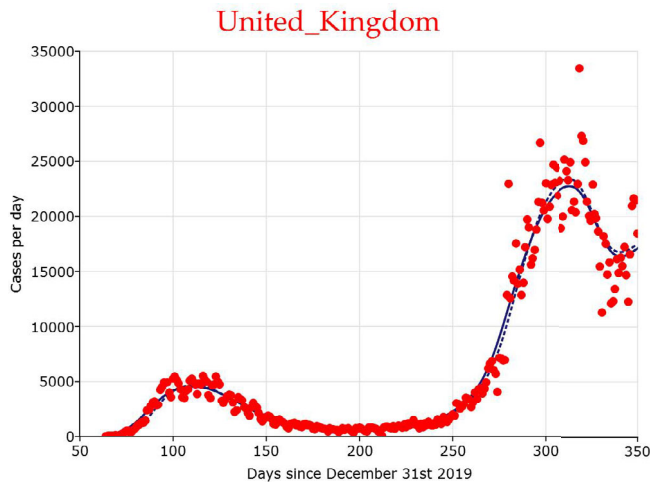
$$Q_{\tau}(Y_t) = \beta_{0t}(\tau) + \beta_{1t}(\tau)t/n, t = 1, \dots, n, \quad (1)$$

$$(\beta_{0t}(\tau), \beta_{1t}(\tau))^{\top} = \beta(\tau)^{(i)} = (\beta_0(\tau)^{(i)}, \beta_1(\tau)^{(i)})^{\top}, [nq_{i-1}] < t \leq [nq_i], i = 1, \dots, m_0 + 1, \quad (2)$$

where $Y_t = \log(R_t + 1)$, with R_t denoting the daily new cases on time t ; $\tau^M = \{\tau_1, \tau_2, \dots, \tau_M\} \subseteq (0, 1)$ is a given set of quantile levels; i indicates the i th segment of the curve.

They say that Li and Linton (2021) was not well adapted to the Covid modelling case since our paper only allowed a quadratic trend. This is not quite right, we discussed both global and local interpretations of our model: in the local interpretation we are approximating a smooth non-parametric trend by a local quadratic approximation, whereas in the global interpretation we imposed a polynomial structure globally, specifically a quadratic. Our paper was written in April 2020 with a view to forecasting and at that time a single peak structure was anticipated and a quadratic function seemed like the simplest way of implementing a single peaked structure, and using it for forecasting. With the benefit of hindsight, we could fit a multiple peaked polynomial model that would fit the full dataset better and provide forecasts consistent with that. The piecewise linear quantile fit is more of a competitor with the local non-parametric model and in this case, we think that smooth curves are more natural, since any kind of policy change designed to impede the progress of the pandemic would likely take effect slowly rather than abruptly, it is not the stock market after all. A further issue that arises is in going from the model specified in log space to the actual numbers and this introduces a bias for mean analysis but not so much for quantile analysis. Below we show an updated local quadratic applied to the log of new UK cases and transformed with a bias correction to the raw cases.

The change-point for each interval is detected by a proposed algorithm called 'GOALS' to facilitate the estimation of multiple change-point. There are two steps for GOALS to work, namely, the GLOBAl testing and the local scanning (GOALS). The main criteria for both steps is self-normalised (SN) statistics. The authors shows that the GOALS algorithm can consistently



[Colour figure can be viewed at wileyonlinelibrary.com]

estimate the change-point and therefore, can fit the infection trajectory better. The quantile regression within each segment can not only accommodate outliers but also deliver the confidence interval automatically. This method attains better in-sample fitting thanks to local interpretations. However, the prediction may be challenging when there are complicated pattern in the near future. To overcome this, the authors suggest to use local quadratic model for out-sample prediction, which is a popular way to depict the development of the COVID-19 pandemic as in Li and Linton (2021).

REFERENCE

Li, S. & Linton, O. (2021) When will the COVID-19 pandemic peak? *Journal of Econometrics*, 220(1), 130–157.

How to cite this article: Li, S., Linton, O. & Ge, S. (2022) Shaoran Li, Oliver Linton and Shuyi Ge's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1836–1837. Available from: <https://doi.org/10.1111/rssa.12933>

Jorge Mateu's first contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Jorge Mateu

Department of Mathematics, University Jaume I, Castellón, Spain

Correspondence: Jorge Mateu, Department of Mathematics, University Jaume I, E-12071 Castellón, Spain.

Email: mateu@uji.es

DISCUSSION

I congratulate the author of this paper as I recognize it is not easy to write such type of paper where there is not much statistics in it but lots of author's personal views of what happened in the initial weeks of the epidemic in early January 2020, and how the early reactions of science were scrutinised.

Overstating the fact that Statistics needs data, and data feeds statistical methods, when we do not have enough data, or the data itself is sort of poor, full of noise and mostly under-reported, then the use of statistical methods is over complicated and might be misleading. Thus let me reinforce that lesson 4 from the author is crucial, and without right data, strong statistical methods cannot perhaps help that much. Lack of data, or small data are facts linked with uncertainty. My personal view is that those many initial published scientific papers due to the urgency in publishing their research did not consider properly the inherent uncertainty, which was globally latent in those early months. I read a number of papers considering more mechanistic compartmental models, such as SIR or SEIR versions, that were fit to early data with no mention to uncertainty or to the under-reporting of cases. They seemed to provide good fits which at the end of the day were not that good and not good enough for predicting into the future. Clearly they did not consider handling such an important aspect. This brings us into statistical methods based on novel combinations between compartment models and temporal or spatio-temporal stochastic models. The author mentioned uncertainty about the reproductive number, and this can be handled using stochasticity in our modelling strategy. Thus, we can cope with having few data at the beginning of the pandemic, if we are prepared to have large residuals in our models, residuals that will be reduced as more clean data comes into play. Another related aspect is that this type of epidemiological disease is hard to understand at early stages, as it behaves differently to past apparently similar diseases. Science can only get close to the problem by assuming certain parameters that are known to be not valid after sometime. Again, the problem with the first pub-

lications was that they tried to show that the whole spread process was already known, which is a sort of fallacy.

How to cite this article: Mateu, J. (2022) Jorge Mateu's first contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1838–1839. Available from: <https://doi.org/10.1111/rssa.12934>

DOI: 10.1111/rssa.12935

Jorge Mateu's second contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Jorge Mateu

Department of Mathematics, University Jaume I, Castellón, Spain

Correspondence: Jorge Mateu, Department of Mathematics, University Jaume I, E-12071 Castellón, Spain.

Email: mateu@uji.es

DISCUSSION

The authors are to be congratulated on a valuable and thought-provoking contribution in the rapidly developing field of the covid-19 pandemic. In 1-year time hundreds of scientific contributions have appeared in the literature, and only time will say which are outstanding. This one is certainly scientific sound, and there is an underlying nice simplicity in the approach. There are several comments I would like to pose here.

The authors argue that their model is purely statistical in that they model the observed time series contrary to other existing mechanistic models, such as SIR. This is true but we add that a good combination of differential equations and stochastic methods can deal with the same sort of problem the authors treat in this paper. What I have in mind is a combination or mixture of generalised logistic differential equations with parameters (including the growth rate) being random variables. This approach is naturally stochastic, brings together growth rates and adapts to existing several waves by detecting the location of the change-points. In this context, using a Bayesian bootstrapping inferential framework we would be able to measure uncertainty in our estimation, and thus in model predictions. This is an issue that I have not found in this paper, in words, how to measure uncertainty in the parameters associated to the linear quantiles, and thus how to control how robust are the estimations under under-reporting or noise in the data. Working with covid-19 data, in particular, at the early stages of the pandemic brings lots of problems in relation to the accuracy of the reported cases.

Another aspect I would like to touch upon here is the spatial structure inherent to this type of data. The method here presented handles time series data with no dependence structure

amongst different regions over which we might have such time series. Of course, I am imaging countries but perhaps counties or health zones within a region for which we have time series of covid-19 data. Considering quantile trend models with spatial dependence structure reflected in the corresponding parameters would provide a wider scope within perhaps a more complex modelling framework.

A final quick comment has to do with implementing this method under the presence of covariates. These linear trends could be revealed by certain covariates, and I would like to see an extension of this model that adds this external information into the existing approach.

How to cite this article: Mateu, J. (2022) Jorge Mateu's second contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1839–1840. Available from: <https://doi.org/10.1111/rssa.12935>

DOI: 10.1111/rssa.12936

Garib Nath Singh's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Garib Nath Singh

Department of Mathematics & Computing, Indian Institute of Technology (ISM), Dhanbad, India

Correspondence: Garib Nath Singh, Department of Mathematics & Computing, Indian Institute of Technology (ISM), Dhanbad 826004, India.

Email: gnsingh@iitism.ac.in

I have read the manuscript entitled 'Small Data, Big Time—A retrospect of the first weeks of COVID-19'. In this manuscript, the author examines the early days of the COVID-19 pandemic, with the focus on data from China, where the virus/disease originated.

Section 2 is devoted to presenting a chronological summary of the events that transpired in the initial stages. Although the paper aims to review the early outbreak from a 'statistician's perspective', this section summarises the administrative decisions and policies during the early days of the outbreak, which are beyond the scope of statistical methodology, and serves as an extended Introduction/ Background.

Section 3 contains detailed discussions of drawbacks of methods used in various papers to analyse early outbreak data, and provides alternative ways to do the same. The author has focused mainly on estimation of the basic reproductive number R_0 , the growth rate r , and the doubling time. Relevant tables and graphs have been provided to support the results obtained.

In Section 3.1, the author has used the Poisson log-linear model to provide a better fit to the data than models proposed by other authors. However, it would be helpful to the readers, especially those from a non-statistical background, if the author mentions briefly the motivation and

reasons for using the specific model and the specific parameters. The model in Section 3.2 is discussed in details which include its derivation/motivation, drawbacks and application on real data.

Additionally, the following minor points require attention from the authors:

1. Section 2.1 mentions Dr. Li Wenliang warned the world of a possible pandemic and was reprimanded for spreading ‘misinformation’. It is relevant to note here that Dr. Li Wenliang worked at Wuhan Central Hospital, from where the first sample of COVID-19 was received and analysed (as mentioned in Section 2.3).
2. In the Afterword Section, the author notes that ‘Eventually, I reconciled myself through the history of infectious diseases and the lens of evolutionary biology’. It is worth mentioning here that medical science has greatly improved over the past decades. Many infectious diseases have been eradicated (e.g. small pox). Thus, historical occurrence of pandemics is not entirely comparable to today’s scenario.
3. Most of the references cited in the manuscript, especially those pertaining to the coverage of the early spread of the disease in China, are from websites or agencies based in China, which may not be completely neutral on the matter. It might be helpful to provide international, independent sources to back up the claims and provide an unbiased view of the scenario.

How to cite this article: Singh, G.N. (2022) Garib Nath Singh’s contribution to the ‘First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic’. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1840–1841. Available from: <https://doi.org/10.1111/rssa.12936>

DOI: 10.1111/rssa.12937

Wei Zhong, Chuang Wan and Changliang Zou’s contribution to the ‘First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic’

Wei Zhong¹ | Chuang Wan¹ | Changliang Zou²

¹Xiamen University, Xiamen, Fujian, China

²Nankai University, Tianjin, China

Correspondence: Wei Zhong, Xiamen University, Xiamen, Fujian, China.

Email: wzhong@xmu.edu.cn

We congratulate Prof. Jiang, Prof. Zhao and Prof. Shao for their timely and thought-provoking paper on modelling the COVID-19 infection trajectory using a piecewise linear quantile trend model. A novel segmentation algorithm based on a self-normalised (SN) test statistic is proposed to estimate multiple change points. The proposed model is interpretable and robust to outliers and heteroscedasticity.

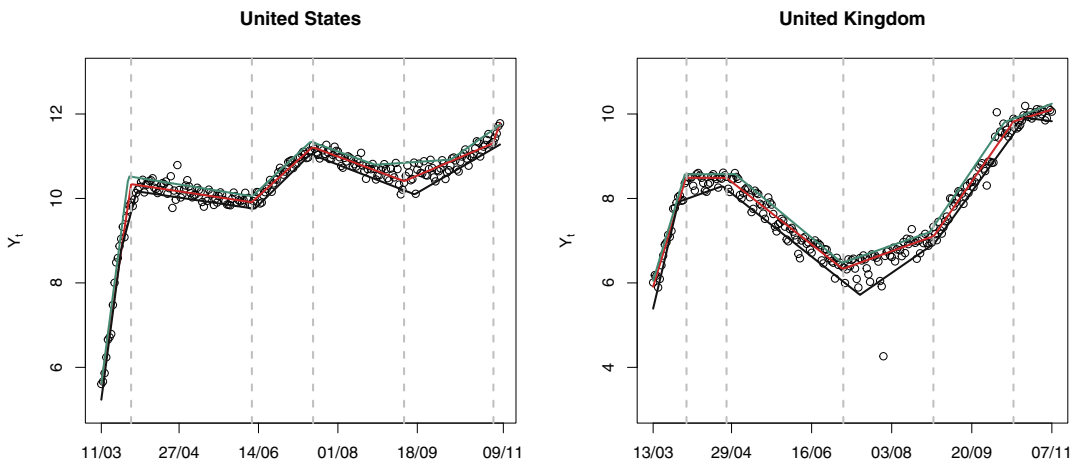


FIGURE D.1 Estimated infection curves of COVID-19 for United States and United Kingdom based on the Multi-Kink Quantile Regression model at different quantile levels, $\tau = 0.1$ (black), $\tau = 0.5$ (red) and $\tau = 0.9$ (green). Vertical lines (grey) indicate the estimated change/kink points at $\tau = 0.5$. [Colour figure can be viewed at wileyonlinelibrary.com]

Zhong et al. (2021) and Wan et al. (2021) developed an interpretable and robust Multi-Kink Quantile Regression (MKQR) for independent data and longitudinal data, respectively. The MKQR model can be considered as an alternative method to the piecewise linear quantile trend model. In longitudinal data, let Y_{ij} denote a response of interest, X_{ij} be a univariate threshold variable and $\mathbf{Z}_{ij}$ be a p -dimensional additional covariates vector at the j th observation for the i th subject, where $j = 1, 2, \dots, n_i$ and $i = 1, 2, \dots, n$. The MKQR model with an unknown number of change points is

$$Q_{\tau}(Y_{ij}) = \alpha_0 + \alpha_1 X_{ij} + \sum_{k=1}^{m_0} \beta_k (X_{ij} - \delta_k)_+ + \mathbf{Z}_{ij}^T \boldsymbol{\gamma}, \quad (\text{D.1})$$

where $(a)_+ = aI(a > 0)$ for any $a \in \mathbb{R}$, $(\delta_1, \dots, \delta_{m_0})^T$ is a vector of change/kink points and m_0 is unknown. By letting $X_{ij} = t/n$ for country i and ignoring the covariates $\mathbf{Z}_{ij}$, the MKQR model can be used to model the COVID-19 infection curves with multiple change points. The main difference of model (D.1) compared with model (1) in their paper is that the regression curve of model (D.1) is required to be continuous at change points. Hence, the MKQR model (D.1) is more parsimonious as it requires estimating $m_0 + 2$ unknown parameters while model (1) has $2m_0$ unknown parameters. As observed in Figure D.1 it may be reasonable to assume that the infection curves are continuous at change points. Figure D.1 depicts the estimated infection curves of COVID-19 for United States and United Kingdom based on the MKQR model (D.1) using the same data provided by the authors for Figure 1 in their paper. The further comparison between two kinds of models are worth further being investigated.

We believe that their paper will stimulate many further related researches. (1) How to accurately estimate the number of change points? Zou et al. (2020) developed a cross-validation-based procedure (COPSS) to consistently estimate the number of change points, which might be a solution. (2) How to make statistical inference for change points in the piecewise linear quantile trend model? Zhong et al. (2021) considered test-inversion confidence intervals for change points.

(3) How to control the false discovery rate of detected change points? (4) How to improve the short-term extrapolation forecast when a change point may exhibit in the follow-up short term?

REFERENCES

- Wan, C., Zhong, W., Zhang, W. & Zou, C. (2021) Multi-kink quantile regression for longitudinal data with application to progesterone data analysis. Available from: <https://arxiv.org/abs/2112.05045>.
- Zhong, W., Wan, C. & Zhang, W. (2021) Estimation and inference for multi-kink quantile regression. *Journal of Business & Economic Statistics*, 40, 1123–1139. <https://doi.org/10.1080/07350015.2021.1901720>
- Zou, C., Wang, G. & Li, R. (2020) Consistent selection of the number of change-points via sample-splitting. *The Annals of Statistics*, 48(1), 413–439.

The authors replied later, in writing, as follows:

How to cite this article: Zhong, W., Wan, C. & Zou, C. (2022) Wei Zhong, Chuang Wan and Changliang Zou's contribution to the 'First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1841–1843. Available from: <https://doi.org/10.1111/rssa.12937>

Nason and Wei's reply to the Discussion of 'The First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Guy P. Nason | James L. Wei

Department of Mathematics, Imperial College, London, UK

Correspondence: James L. Wei, Department of Mathematics, Imperial College, London, UK.

Email: james.wei19@ic.ac.uk

The authors would like to thank the proposer, seconder and all of the discussants for constructive, useful and incisive points.

Studley is correct that we experimented with two types of network: the fully connected one, where every country was connected to every other country, but weighted according to the degree of economic trade, and a nearest neighbour network, which linked countries to the single country that it exported the most to. *Studley* wonders what the influence of closed land or air borders might be. That is an interesting question and we do not have a detailed answer, but observe that relative inter-country trade was perhaps not so much affected—lorries still crossed borders and air freight still happened. We know of no country that was cut completely off. On the other hand, trade might well have been suppressed by such restrictions. One line of our research in network time series is performing the modelling we describe, but with dynamic networks. That is, the edges changing in some way over time. Computationally, it is something that can be done, but theoretically, it is a much harder problem to study.

Studley's idea of extending our work to regional trade is an excellent one, especially perhaps when considering a large country in that many other differences between regions will be less than differences between countries. However, less economic data is often generally available for very fine scale regions, so there will be a limit to what can be done.

Studley's idea of using age-related indicators and asymptotic case estimates as further exogenous variables is also excellent and one can think of many other variables that one *could* use, but one must be careful of association between such variables themselves. Both *Bird* and *Kumar* suggest that age-related variables might be useful. We can only agree, but the modelling complexity would increase, but it is something that could certainly be explored.

Bird also suggests some extensions along these lines too and the interesting suggestion that Purchasing Manger Index (PMI) response might be different at different times in the pandemic, after different waves perhaps. This is true, and not something we have investigated. However,

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* published by John Wiley & Sons Ltd on behalf of Royal Statistical Society.

could be modelled by a non-stationary version of what we propose, which exists in concept, but not yet, as far as we know, in theory and certainly not in accessible software. On the other hand, reducing the data into smaller windows (piecewise stationary) or weighted regions (locally stationary, say) might reduce the available PMI data, which is on a fairly coarse scale with respect to the pandemic data. So perhaps the quality of estimation would not be so good? *Bird* encapsulates this last point beautifully by ‘A week is a long time in politics: a month of fatalities is a very long time in this pandemic’. In response to *Bird* the precise details of how the PMI survey is carried out is the subject of a significant and well-carried out process, which has been around for many years, but even the best-designed surveys would not be able to completely eliminate systematic country differences.

On *Bird*’s point about occurrence time—it matters certainly for fine-scale epidemiological modelling, but because we are interested in the impact of COVID on economic conditions through the business confidence channel, our modelling involves using the data available to purchasing managers and the general public, that is, the headline reported death rates (which will not be the occurrence times perhaps, but ‘surrogates’ that people have access to).

We are grateful to *Choi and Leung Lai* for their comments on historical, independent and near-simultaneous development of the network autoregressive processes. We are very interested in their proposal to use such methods for their spatio-temporal projects at Stanford’s climate and environmental institutes.

Chai raises two interesting questions on our work. On the first: the relative performance of the nearest neighbour and fully connected networks very roughly tells us whether the surveyed purchasing managers take account of all countries’ economic conditions, or restrict their attention to only the main trading partners (we can think of this as a form of bounded rationality, where individuals are limited in time to make decisions). On the second: do we distinguish between leading and lagging indicators? Yes, in this model — PMIs are considered to be leading indicators for economic conditions, which we model on a network.

We would like to thank *Diggle* for his own prescient comments about the relationship of our work to Markov random field models and the SAR models. Hawkes’ prescient remark was one of many. *Diggle*’s suggestion about using a marginal model for the X_t and then a conditional model for $Y_t - X_t^T \lambda$ is certainly interesting and to be considered. We do not explicitly model the exogenous regressors and would be uncertain about using hierarchical modelling in this situation, but we stand to be educated.

Ge, Linton and Li’s perceptive comment on using contemporaneous effects is well-made for modelling. However, our key goal was forecasting and we are not sure how contemporaneous terms could be made to work well, if at all, in this situation. We do agree with *Ge, Linton and Li* that the vector autoregression (VAR) model for $N = 12$ is heavily parameterised and that is precisely why we are keen on the highly parsimonious generalised network autoregression (GNAR) models. It is this parsimony, which is key to the great success of GNAR models for forecasting (when it is an appropriate model, of course).

On *Ge, Linton and Li*’s mention regularisation as a viable alternative. Regularisation is an important class of approaches for when knows nothing much else about the system of study, for example, when one only has access to a multi-variate time series, say. The point about network time series is that very often one has access to a separate associated network (or one can be constructed from other covariates) that gives you **strong** information about how the variables might be connected. We believe we should use that information if possible and appropriate. Of course, it might turn out that the separate network information provides no information on the association between variables, but that can be simply determined from basic model fit assessment

techniques. However, networks often supply powerful information completely negating the need for mathematical techniques such as regularisation. Regularisation is not generally related to the underlying connectivity structure and/or is potentially negatively influenced by spurious or nonsense correlations (being based on the values of the process as they are).

We agree with *Ge, Linton and Li's* remarks on exogenous variables and their suggestion is certainly useful. We do think that countries, certainly the United Kingdom, used mitigations that are driven almost exclusively by (recent) historical COVID statistics, in spite of a wide variety of commentators strongly suggesting that other considerations should come into play, such as economic, mental health or non-COVID health issues. Latterly, we feel that economic considerations have now come to the forefront, certainly in England and less so in the devolved regions of Scotland and Wales. In truth, perhaps there are actually more likely to be weak feedback loops between several of these variables, which can then cause all kinds of effects that would be hard to disentangle and study.

We are grateful to and agree with *Kumar* in that it would be fascinating to compare our generalised network autoregression with exogenous variables (GNARX) forecasting results with those produce by machine learning techniques and indeed we have a project investigating that too.

Can we once again thank all respondents for their many positive and creative suggestions.

How to cite this article: Nason, G.P. & Wei, J.L. (2022) Nason and Wei's reply to the Discussion of 'The First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1844–1846. Available from: <https://doi.org/10.1111/rssa.12938>

DOI: 10.1111/rssa.12939

Zhao's reply to the Discussion of 'The First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Qingyuan Zhao 

Statistical Laboratory, University of Cambridge, Cambridge, Cambridgeshire, UK

Correspondence: Qingyuan Zhao, Statistical Laboratory, University of Cambridge, Cambridge, UK.

Email: qyzhao@statslab.cam.ac.uk

I would like to thank the Royal Statistical Society for organizing this special session for the statistical aspects of the COVID-19 pandemic, and *Mrs. Studley, Prof. Bird*, and many others for contributing their thoughts and discussion. I will first respond to comments pertinent to my article before discussing the broader points.

Bird asked me to offer clarity on how I sampled media articles in my writing. This is an obvious oversight on my part. I took advantage of two crowdsourcing projects hosted on GitHub that provided an archive of media articles (github.com, 2020b) and an organized timeline of the coronavirus outbreak in Wuhan (github.com, 2020a). Most of the articles were published by print media, digital media or government agencies, but all of them were accessed on the Internet. The

presentation of these references in my article was, of course, selective. But I read almost all the articles referenced in the crowdsourced timeline (github.com, 2020a) up till the end of January, 2020, and I tried to present stories from all sides as long as I found them credible. I should also take this opportunity to correct another omission in my acknowledgement: my sampling of the media was not possible without the journalists who wrote all the reports in unfavourable circumstances and the numerous anonymous volunteers who contributed to these crowdsourcing projects. They deserve all the credit.

Bird made another excellent insight that the event-time that matters in infectious disease epidemics is the occurrence date, or in the terminology I use, the infection time. This was touched briefly upon in section 3.2 and in much more depth in the references there. The same point was brought up to me by Prof Marc Lipsitch at the Harvard School of Public Health in late January 2020, when I was trying to collect and analyze some data about the outbreak in Wuhan. This point prompted me and my collaborators to think about what data could provide information about the infection time. A traditional approach in infectious disease modeling is to back-calculate the infection time using symptom-onset and incubation period distribution (Fraser, 2007). However, the method described there only does mean imputation and does not consider uncertainty in the imputation. This was addressed in the initial analysis made by me and coauthors (Zhao et al., 2020), but that was too unconventional for it to be published when the conclusion still mattered. We also brought up the same point in an online commentary on an influential article that estimated the curve of effective reproductive number (R_t) in Wuhan and associated it with public health interventions (Ju et al., 2020). However, in our personal communications it did not look like the authors of that article understood our suggestion. My observation is that the point raised by *Bird* seems to be a common struggle for infectious disease modelers who are familiar with dynamic systems (such as SIR-type models), while statisticians often have no difficulty appreciating the importance of infection time.

Relatedly, *Prof. Mateu* mentioned novel combinations between compartment models and temporal or spatio-temporal stochastic models for infectious diseases. In this era of data explosion, we can collect much more data than a incidence curve. Sections 2 and 3.1 of the article have clearly demonstrated why modeling the incidence curve is inadequate and potentially misleading for early outbreak analysis. I hope the next generation of infectious disease modeler will adopt this combination approach and make use of unconventional data. A first step would be standardising the case-specific data, as advocated by *Prof. Cook*, *Prof. Choi* and *Prof. Wong*. Nianqiao Ju and I manually curated a dataset of initial COVID-19 cases in some Chinese provinces and Asian countries (Zhao et al., 2021). During that painstaking process, we had to revise the data columns numerous times, as the health agencies publish their case reports in different ways and use different wordings. The dataset is documented in our paper and could serve as an initial template for the standard codification.

Prof. Singh suggested that because most of the references in my section 2 are from Chinese sources, they may not be completely neutral on the matter. I agree with this, but I think they are perhaps more critical of the government response and much more detailed than what can be found on a credible western media outlet. On this matter, the reader should be aware of at least two sources of implicit bias. First, almost all the authors, interviewees, and commentators are Chinese nationals or of Chinese descent, who may share a different ideology (in particular, towards containment of infectious diseases) than someone in a western society. Second, almost all the articles (especially those published in print media) were subject to media censorship; in fact, the original links to some of the referenced articles were deleted and I could only access them through Internet archives. *Singh* also asked me to briefly explain the Poisson log-linear

model used in section 3.1. The log-linear model assumes that the logarithm of the expected case count is linear in time, or in other words, the epidemic had an exponential growth. The model further assumes that the observed case count follows a Poisson distribution, which is the most common probability distribution for modelling count data.

I am delighted to see that most of the messages in the article appear to be well received by the discussants. Several discussants mentioned uncertainty due to asymptotic cases and under-reporting. *Prof. Kumar* called these ‘dark data’, which sound like a synonym for what statisticians call missing data or selection bias. We investigated this issue in another article (Zhao et al., 2021) and found that potential bias from asymptotic cases can be addressed by making a careful statistical analysis. *Cook et al.* also mentioned the role of children and severity profile, which should be important components of a good model for COVID-19 transmission if such data are available.

Turing to the broader lessons for statisticians, *Kumar* used the retraction of a highly controversial article about hydroxychloroquine to argue that the peer-review process and the ethics of data-driven research need to be strengthened. *Cook, Choi and Wong* proposed that we need to develop methods that can be rapidly developed in early outbreak analyses. *Studley* suggested that we need to ‘be there’ and ‘be involved in the design of future analyses’. Finally, *Bird* remarked that ‘early appreciation of a new infectious disease requires a diversity of experts’ and ‘statisticians do not work alone’. I think there is a common theme here: statisticians need to be more active and collaborative. But I would take one step further: I believe *statisticians need to be at the center of action* and not just be an advisor, because we understand the intricacies of data analysis, study design and evidence synthesis better than anyone else. If statisticians merely stay as advisors and do not get their hands dirty with study design and data analysis, the same mistakes we saw in COVID-19 analyses and decisions will inevitably happen again.

ORCID

Qingyuan Zhao  <https://orcid.org/0000-0001-9902-2768>

REFERENCES

- Fraser, C. (2007) Estimating individual and household reproduction numbers in an emerging epidemic. *PLoS One*, 2(8), e758. <https://doi.org/10.1371/journal.pone.0000758>
- github.com. (2020a) Epidemic situation and public opinion: timeline of new crown pneumonia. Available at: <https://github.com/lestweforget/COVID-19-Timeline> [Accessed 14th January 2022].
- github.com. (2020b) 2020 Memories of COVID-19: reports, non-fiction, and personal narratives. Available at: <https://github.com/2019ncovmemory/nCovMemory> [Accessed 14th January 2022].
- Ju, N., Bacallado, S., & Zhao, Q. (2020) Using symptom onsets instead of incidences results in time-lagged estimate of $R(t)$. Available at: <https://github.com/nianqiaoju/rt> [Accessed 14th January 2022].
- Zhao, Q., Chen, Y., & Small, D. S. (2020) Analysis of the epidemic growth of the early 2019-nCoV outbreak using internationally confirmed cases. *medRxiv*. <https://doi.org/10.1101/2020.02.06.20020941>. Available at: <https://www.medrxiv.org/content/early/2020/02/09/2020.02.06.20020941> [Accessed 14th January 2022].
- Zhao, Q., Ju, N., Bacallado, S. & Shah, R.D. (2021) BETS: The dangers of selection bias in early analyses of the coronavirus disease (COVID-19) pandemic. *Annals of Applied Statistics*, 15(1), 363–390.

How to cite this article: Zhao, Q. (2022) Zhao’s reply to the Discussion of ‘The First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic’. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1846–1848. Available from: <https://doi.org/10.1111/rssa.12939>

Jiang, Zhao and Shao's reply to the Discussion of 'The First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'

Feiyu Jiang^{ORCID} | Zifeng Zhao | Xiaofeng Shao^{ORCID}

Department of Statistics, University of Illinois at Urbana-Champaign, Urbana-Champaign, Illinois, USA

Correspondence: Xiaofeng Shao, Department of Statistics, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA.

Email: xshao@illinois.edu

1 | ALTERNATIVE MODELS FOR THE COVID-19 INFECTION TRAJECTORIES

Following Box's famous saying 'All models are wrong, but some are useful', we view our piecewise linear quantile trend model as an approximation to the true model (if there is one), and there are certainly many other legit approximations that may fit the data equally well or even better. A nice feature associated with our piecewise linear quantile trend model is its interpretability, that is, the slope parameter naturally corresponds to the growth rate of the infection trajectory. In addition, it provides accurate short-term forecasting in a relatively straightforward manner, as is demonstrated in the paper by extrapolating the last estimated segment. Our focus on the conditional quantile modelling is motivated by the well-known features of Quantile Regression (QR) (i.e. resistant to outliers, captures heteroscedasticity and provides both point and interval forecasts), as mentioned in Wang's discussion.

We thank Shi, Zhang and Dong for bringing some recent work on non-linear modelling of Covid-19 data to our attention. The model they described, that is, $Q_{\tau}(y_t) = f(t/n; \beta_1, \dots, \beta_q)$, where f is a nonlinear function of t/n , is a very broad one, and indeed our piecewise linear quantile trend model is just a member of this broad class. One could consider some other members, such as the piecewise quadratic quantile trend model, as implied in the comments of Li, Linton and Ge.

The GOALS algorithm can be applied to estimate change-points in the piecewise quadratic quantile model without much difficulty. For demonstration, Figure 1 provides the model fits to US and UK Covid-19 time series at log scale, which seem reasonable. The interpretation for

We are grateful to the discussants for their stimulating and insightful comments. For brevity, we apologise that we cannot address all the issues raised by the discussants. We have arranged our reply under four broad topic headings as follows.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](#) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* published by John Wiley & Sons Ltd on behalf of Royal Statistical Society.

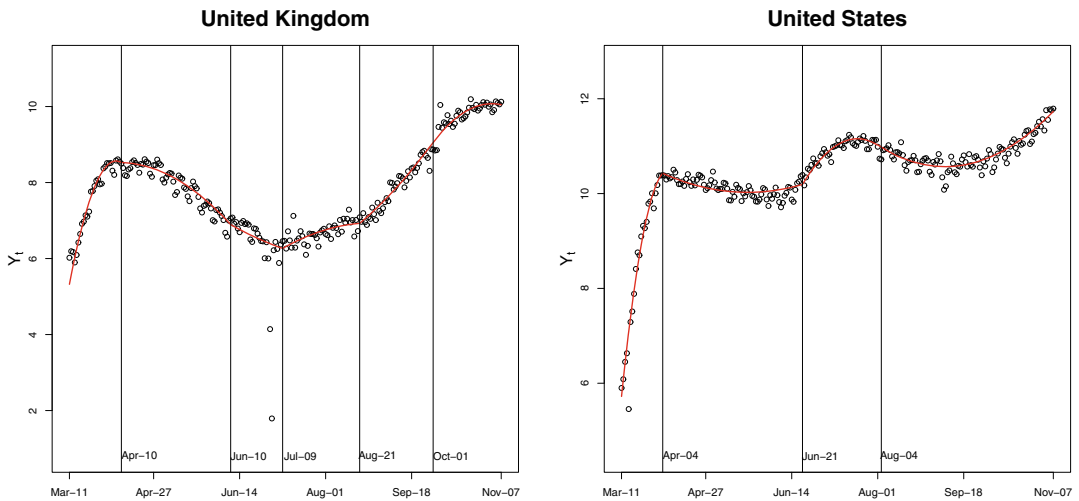


FIGURE 1 Estimated piecewise quadratic quantile trend models by M-GOALS for United Kingdom and United States at $\tau = 0.5$ (with imposed continuity constraint in the same spirit of (1)). [Colour figure can be viewed at wileyonlinelibrary.com]

piecewise linear model is simple and intuitive but this seems lost for local quadratic model. In general, we expect a fully non-parametric method, such as a local polynomial regression may fit the in-sample data better. However, it cannot detect change-points and thus may not be directly useful for downstream analysis such as evaluation of effectiveness of public health measures.

As Mateu points out, there are other mechanistic models such as SIR to deal with the same data and could involve similar change-point segmentation problem we encounter here. It would be interesting to extend the GOALS algorithm to estimate the change-points in the parameters of an model. The basic idea is to approximate the time-varying parameters in an SIR model by a piecewise constant function, which then requires change-point estimation, see Bai et al. (2020) for a related work in this direction.

2 | VARIANTS/ALTERNATIVES OF QUANTILE REGRESSION FOR THE COVID-19 DATA

As pointed out by Farcomeni and Geraci, our response $Y_t = \log(R_t + 1)$ is a discrete random variable, as R_t is an integer valued random variable. They mention the literature of quantile regression for discrete response and their recent work on mid-quantile regression in Geraci and Farcomeni (2019). Since the range of Covid-19 daily counts is relatively large for almost all countries, ignoring the discrete nature of the response variable and treating it as if it is continuous might be okay from the modelling and forecasting perspective. As a matter of fact, if we model the jittered response $Y_t = \log(R_t + 1 + U_t)$, where U_t are iid uniform(0, 1), we do not see any noticeable differences in the estimation of change-points. We also tried the mid-quantile regression using the R package `Qtools`, and only observed slight difference from our results. For discrete response with limited range and small number of support points, the mid-quantile regression might make a difference.

Jana, Basu and Jana mention that it is of practical interest to model and predict the extreme quantiles (at probability level 95% or higher), which we certainly agree. However, for the Covid-19 time series, the minimal segmentation length is of order 30–50 for many countries. Estimating an extreme quantile with such a short sample size is a difficult problem due to data sparsity at high quantiles and suffers from high variability, which may have a negative impact on the change-point estimation. As they point out, there could be ways to overcome data sparsity by using a parametric distribution such as the generalised Pareto distribution to model the observations above a threshold. It seems that such an approach may be difficult to implement as the choice of threshold in the change-point setting can be a thorny issue. It would be encouraging to see new methods yet to be developed for change-point analysis in extreme quantile regression settings such as Wang et al. (2012) and Wang and Li (2013).

Wang proposes the use of composite quantile regression (Kai et al., 2010; Zhao & Xiao, 2014; Zou & Yuan, 2008) to estimate the conditional mean in a robust way. We expect our GOALS algorithm can be used to estimate the change-points in conditional mean, and we call this method Mean-GOALS-CQ. Conceptually, we feel that Mean-GOALS-CQ and our M-GOALS share the same spirit, as both try to pool the information across several quantiles in estimating the change-point locations, and may deliver very similar results under the assumption that the change-point configurations across quantiles (and the mean) are identical. It would be interesting to confirm this point via numerical simulation.

The literature of modal regression, as pointed out by Wang, is new to us. We agree that change-point detection in modal regression framework would be interesting, but may be non-trivial. In particular, a combination of the estimator developed in Ota et al. (2019) and our GOALS algorithm seems difficult. The QR-based conditional modal estimator in Ota et al. (2019) requires the linear quantile regression at all quantiles $\tau \in (0, 1)$, which results in a complication in implementation due to the presence of short segments in Covid-19 time series, the involvement of recursive sub-sample estimator in the GOALS algorithm and the difficulty of choosing the necessary smoothing parameter in Ota et al. (2019).

We are grateful to Zhong, Wan and Zou for pointing out a related line of research: bent line quantile regression (for one change-point) or multi-kink quantile regression (for multiple change-points) (Li et al., 2011; Wan et al., 2021; Zhong et al., 2021), which have been applied to both independent data and longitudinal data. There are indeed some similarity between our piecewise linear quantile trend model and the multi-kink QR model, in the sense that the regression parameter is piecewise constant in both frameworks. Typically, in the bent line or multi-kink QR literature, the covariates are assumed to be random, and the piecewise constant regression parameter corresponds to a set of thresholds for a univariate threshold variable (i.e. in the state domain of the threshold variable), whereas in our model, since our covariate $(1, t/n)$ is fixed, it corresponds to a set of time-ordered change-points (i.e. in the time domain).

Nevertheless, as Zhong, Wang and Zou pointed out, the multi-kink QR model can be used to fit Covid-19 daily time series, although it requires the conditional quantile to be a continuous function of time t . By contrast, we do not assume the continuity of regression curve in our model and let the data speak for itself. Furthermore, we assume the change-points are identical across the three quantile levels (i.e. 10%, 50%, 90%) and estimate the change-points via the M-GOALS algorithm by pooling the information contained at these three quantile levels. It is worth pointing out that it is straightforward to extend our method to handle change-point estimation with continuity constraints. Specifically, for observations between $\{Y_t\}_{t=t_1}^{t_2}$ with a potential change-point k in between,

the following estimator imposes a continuity constraint at k of the estimated trends:

$$\begin{aligned} \left(\hat{\beta}_{0:t_1, t_2}^{(c)}, \hat{\beta}_{1:t_1, k}^{(c)}, \hat{\beta}_{1:k+1, t_2}^{(c)} \right) = \arg \min_{\alpha_0, \alpha_1, \alpha_2 \in \mathbb{R}} \sum_{t=t_1}^{t_2} \rho_{\tau} \left(Y_t - \alpha_0 - \alpha_1 \left[\left(\frac{t}{n} \right) 1(t \leq k) + \left(\frac{k}{n} \right) 1(t > k) \right] \right. \\ \left. - \alpha_2 \left(\frac{t-k}{n} \right) 1(t > k) \right). \end{aligned} \quad (1)$$

Naturally, the sub-sample SN statistic can be constructed by contrasting $\hat{\beta}_{1:t_1, k}^{(c)}$ and $\hat{\beta}_{1:k+1, t_2}^{(c)}$ instead of their unconstrained counterparts.

3 | COVID-19 DATA AND MODELLING ISSUES

Bird points out that our model is actually not fitted to 'COVID-19 infection trajectory'—rather the trajectory of report-dates for newly diagnosed SARS-CoV-2 cases. The infection date and reporting date are different and this difference can vary internationally, suffer weekend and vacation perturbations. This is an astute observation. There are also potential bias in the reported case due to under-reporting, as pointed out by Studley and Mateu.

Technically speaking, the infection date is unknown and should be before the reporting date. However, such information is not available for the data we analysed in our manuscript. Nevertheless, we feel the data analysis we provide using the piecewise linear trend model and the GOALS algorithm can still shed some light on the COVID-19 growth pattern in each country and the similarity between different countries, albeit with some systematic shift in time due to the reporting delay. It would be interesting to analyse a 'new' infection trajectory based on the infection date, if such data are available, and compare to what we did.

In response to Mateu's comments on incorporating spatial dependence into the model, our model and segmentation method is proposed for a univariate time series and its extension to modelling and estimating change-points in the trends of multiple time series is certainly interesting. For Covid-19 time series observed for neighbouring counties or states, it is quite likely there are non-trivial spatial dependence and cross-temporal dependence, and it is also possible they share similar trend pattern or change-points. How to capture these features/dependence using a multi-variate time series model allowing for change-points would be an interesting future research direction.

Several discussants (Choi and Lai, Mateu, Farcomeni and Geraci) mention the topic about incorporating covariates into the piecewise linear trend model. In terms of implementation, there is not much difficulty as we only need to change the fixed covariate $X_t = (1, t/n)^T$ in the regression model to incorporate other available covariates of interest, such as the vaccination rate, seasonality effect and timing of public health measures. The SN-based test statistic and the (multi-scanning) M-GOALS algorithm can be adjusted correspondingly in a straightforward manner.

However, the interpretation of such a model will be much different than the piecewise linear quantile trend model. The purpose of our study is mainly about analysing the trend change in the infection curve via the identification of potential change-points and our model has intuitive interpretation as the estimated slope in each segment naturally characterises the growth rate of the Covid-19 pandemic. In the downstream analysis, we then relate the detected change-points in trend to potential factors such as introduction of public health measures and emergence of

new variants. However, with the incorporation of more variables, such interpretation is lost. For example, if we include an indicator covariate for an exogenous event, such as the introduction of a new public health measure or the emergence of a new variant, the associated change in trend of the Covid-19 trajectory may be absorbed by the covariate, and thus cannot be detected as a change-point, which defies the purpose of our analysis.

On the other hand, if our focus is solely on the forecasting end, incorporating informative covariates should help improve the prediction accuracy, though the complication is that we would also need to forecast the covariate X_t reasonably well into the future, which may or may not be easy. In terms of theoretical justification, our current theory is established for fixed covariate and we expect the results should continue to hold for stationary random covariate. However, different framework and theory may need to be developed for handling non-stationary covariate.

4 | INFERENCE AND FORECASTING

It is obviously important to provide a measure of uncertainty for the estimated intercept and slope parameter in our piecewise linear quantile trend model, as pointed out by Mateu. Statistically speaking, it seems natural to ask whether the uncertainty associated with the change-point estimation has any impact on the SE of intercept and slope estimator in a particular segment asymptotically. How reliable is the standard error for the estimated parameter for a given segment post change-point estimation? Would a bootstrap method, such as Bayesian bootstrap, provide proper uncertainty measure in this change-point model? We shall leave these questions for future investigation.

Prompted by the comments of Zhong, Wan and Zou, we acknowledge the importance of inference for change-points, that is, constructing a confidence interval after we obtain the point estimator of the change-point location. Oka and Qu (2011) provide a solution to this issue for quantile regression with temporal independence. However, this is a more challenging topic in our setting due to unknown temporal dependence. It is also equally challenging to address potential forthcoming change-point in the short horizon with our segmentation-extrapolation forecast. Along the same line, Linton and Ge suggest that out-of-sample prediction should be performed using a local quadratic model. Actually, in our forecasting procedure, we apply a linear or quadratic QR model to the last segment and let the data determine which fits the last segment better. In our rolling window forecast, we have seen that for some periods (see figure S7 in the supplement), the quadratic model fits the last segment better and its corresponding out of sample forecast is more accurate.

5 | CONCLUSION

We started this paper at a relatively early stage of the pandemic and we hope it would stimulate more research on Covid-19 data from statisticians during this difficult time. The Royal Statistical Society has offered us a great opportunity to bring our work to other statisticians and data scientists, and we are pleased to see that there is general agreement on the importance of developing new models and methods for the analysis of Covid-19 data. We would like to thank Adam Sykulski and Judith Shorten for their efficient handling of the paper and smooth coordination of

our presentation at the RSS discussion paper meeting. Finally, We would like to express our great gratitude to all the discussants as well as participants at the discussion meeting for their interests, support and insightful comments.

ORCID

Feiyu Jiang  <https://orcid.org/0000-0001-9902-2768>

Xiaofeng Shao  <https://orcid.org/0000-0001-5822-8209>

REFERENCES

- Bai, Y., Safikhani, A. & Michailidis, G. (2020) Non-stationary spatio-temporal modeling of COVID-19 progression in the U.S. *medRxiv*. <https://doi.org/10.1101/2020.09.14.20194548>.
- Geraci, M. & Farcomeni, A. (2019) Mid-quantile regression for discrete responses. *arXiv:1907.01945*.
- Kai, B., Li, R. & Zou, H. (2010) Local composite quantile regression smoothing: an efficient and safe alternative to local polynomial regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72, 49–69.
- Li, C., Wei, Y., Chappell, R. & He, X. (2011) Bent line quantile regression with application to an allometric study of land mammals' speed and mass. *Biometrics*, 67, 242–249.
- Oka, T. & Qu, Z. (2011) Estimating structural changes in regression quantiles. *Journal of Econometrics*, 162, 248–267.
- Ota, H., Kato, K. & Hara, S. (2019) Quantile regression approach to conditional mode estimation. *Electronic Journal of Statistics*, 12, 3120–3160.
- Wan, C., Zhong, W., Zhang, W. & Zou, C. (2021) Multi-kink quantile regression for longitudinal data with application to the progesterone data analysis. *arXiv:2112.05045*.
- Wang, H.J. & Li, D. (2013) Estimation of extreme conditional quantiles through power transformation. *Journal of the American Statistical Association*, 108, 1062–1074.
- Wang, H.J., Li, D. & He, X. (2012) Estimation of high conditional quantiles for heavy-tailed distributions. *Journal of the American Statistical Association*, 107, 1453–1464.
- Zhao, Z. & Xiao, Z. (2014) Efficient regressions via optimally combining quantile information. *Econometric Theory*, 30, 1272–1314.
- Zhong, W., Wan, C. & Zhang, W. (2021) Estimation and inference for multi-kink quantile regression. *Journal of Business & Economic Statistics*, 40, 1123–1139.
- Zou, H. & Yuan, M. (2008) Composite quantile regression and the oracle model selection theory. *The Annals of Statistics*, 36, 1108–1126.

How to cite this article: Jiang, F., Zhao, Z. & Shao, X. (2022) Jiang, Zhao and Shao's reply to the Discussion of 'The First Discussion Meeting on Statistical Aspects of the Covid-19 Pandemic'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1849–1854. Available from: <https://doi.org/10.1111/rssa.12940>