

Preface



Cite this article: Panovska-Griffiths J, Waites W, Ackland GJ. 2022 Technical challenges of modelling real-life epidemics and examples of overcoming these. *Phil. Trans. R. Soc. A* **380**: 20220179.
<https://doi.org/10.1098/rsta.2022.0179>

Received: 3 July 2022

Accepted: 5 July 2022

One contribution of 18 to a theme issue 'Technical challenges of modelling real-life epidemics and examples of overcoming these'.

Subject Areas:

computer modelling and simulation,
mathematical modelling

Keywords:

epidemics, mathematical modelling, COVID-19

Author for correspondence:

J. Panovska-Griffiths

e-mail: jasmina.panovska-griffiths@ndph.ox.ac.uk


Technical challenges of modelling real-life epidemics and examples of overcoming these

J. Panovska-Griffiths^{1,2}, W. Waites³ and G. J. Ackland⁴

¹The Big Data Institute and the Pandemic Sciences Institute, Nuffield Department of Medicine, and ²The Queen's College, University of Oxford, Oxford, UK

³Department of Computer and Information Sciences, University of Strathclyde, Glasgow G1 1XH, UK

⁴Institute of Condensed Matter and Complex Systems, School of Physics and Astronomy, University of Edinburgh, Edinburgh EH9 3FD, UK

 JPG, 0000-0002-7720-1121; WW, 0000-0002-7759-6805; GJA, 0000-0002-1205-7675

The coronavirus disease 2019 (COVID-19) pandemic has highlighted the importance of mathematical modelling in informing and advising policy decision-making. Effective practice of mathematical modelling has challenges. These can be around the technical modelling framework and how different techniques are combined, the appropriate use of mathematical formalisms or computational languages to accurately capture the intended mechanism or process being studied, in transparency and robustness of models and numerical code, in simulating the appropriate scenarios via explicitly identifying underlying assumptions about the process in nature and simplifying approximations to facilitate modelling, in correctly quantifying the uncertainty of the model parameters and projections, in taking into account the variable quality of data sources, and applying established software engineering practices to avoid duplication of effort and ensure reproducibility of numerical results. Via a collection of 16 technical papers, this special issue aims to address some of these challenges alongside showcasing the usefulness of modelling as applied in this pandemic.

© 2022 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License <http://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, provided the original author and source are credited.

Since the onset of the COVID-19 pandemic, in a number of countries, including the UK, mathematical and statistical modelling has been at the forefront of informing the status of the epidemic and providing evidence around possible future scenarios to decision makers. Modelling gives a flexible theoretical framework which, parameterized with and calibrated to data, allows exploration of different transmission scenarios and evaluation of different interventions. Although mathematical modelling of infectious diseases has a long history, this pandemic highlighted its applicability and importance for real-time decision-making. The pandemic has also driven the development and application of novel mathematical and computational methods, some of which are presented in this special issue. Undertaking real-time modelling is challenging as it requires a trade-off between fast and less robust modelling and delayed but more robust modelling. The former can be appropriate when data is scarce and we know very little, like at the onset of the pandemic, or when new variants of the virus occur. The latter can be appropriate when we are faced with an already developed epidemic and changes in response may be necessary; for example when faced with a fragmented epidemic where the classic measures such as national reproduction number R values may not be appropriate or when developing and rolling out a vaccination strategy.

During the current pandemic, most of the modelling groups involved in advising the UK Government have been undertaking responsive and real-time modelling, often balancing the need to give quick analysis outcomes with assuring accuracy and uncertainty quantification. While undertaking and publishing timely modelling results is important to advice on the current pandemic, as part of the lessons learned from this pandemic and in preparation for the next, it is important that modellers invest in testing the rigorousness in deriving outcomes. Real-time modelling and the vast amount of models developed in the last 24 months have highlighted gaps in the existing technical frameworks that ought to be addressed in preparedness for possible future pandemics.

Much of the work in this edition was supported by the Royal Society Rapid Assistance in Modelling the Pandemic [1] initiative which mobilized modelling expertise from a diverse range of disciplines to support the pandemic modelling community already working on Coronavirus (COVID-19). Details of the initiative, its origin, purpose, aims and outcomes are described in a separate article within this edition [2].

This special issue showcases a combination of articles that address overarching technical issues identified in the course of the current epidemic that would be relevant when modelling future ones. Through a series of articles, we generate a library of tools that promote open and transparent science and would be relevant as pandemic outbreak analysis tools in future.

Although each of the articles within the collections makes a fruitful singular contribution to scientific modelling, there are themes that link them. Figure 1 illustrates one possible way of classifying the articles in this issue according to subject matter. Most articles are concerned with more than one of these concepts and the edges show this. For example, the article by Marion *et al.* [3] presents an improved version of the approximate Bayesian computation algorithm and shows how this can be used to fit to data. The article by Moore *et al.* [4] on the use of scoring rules to refine forecasts sits in a similar conceptual space.

Agent-based models (ABMs) figure prominently in this issue with [5–8] all using them to explore different questions. All of these involve fitting to real data, of course, and all consider the effect of different variants of SARS-CoV-2. Sanz-Leon *et al.* [7] investigate vaccine hesitancy and the δ variant in Australia. Panovska-Griffiths *et al.* [5] and Hinch *et al.* [6] are focused on the English COVID-19 epidemics and both estimate the transmissibility of variants and evaluate interventions, with the former combining the ABM with a statistical regression and making a policy contribution with the results, while the latter is taking a geospatial approach.

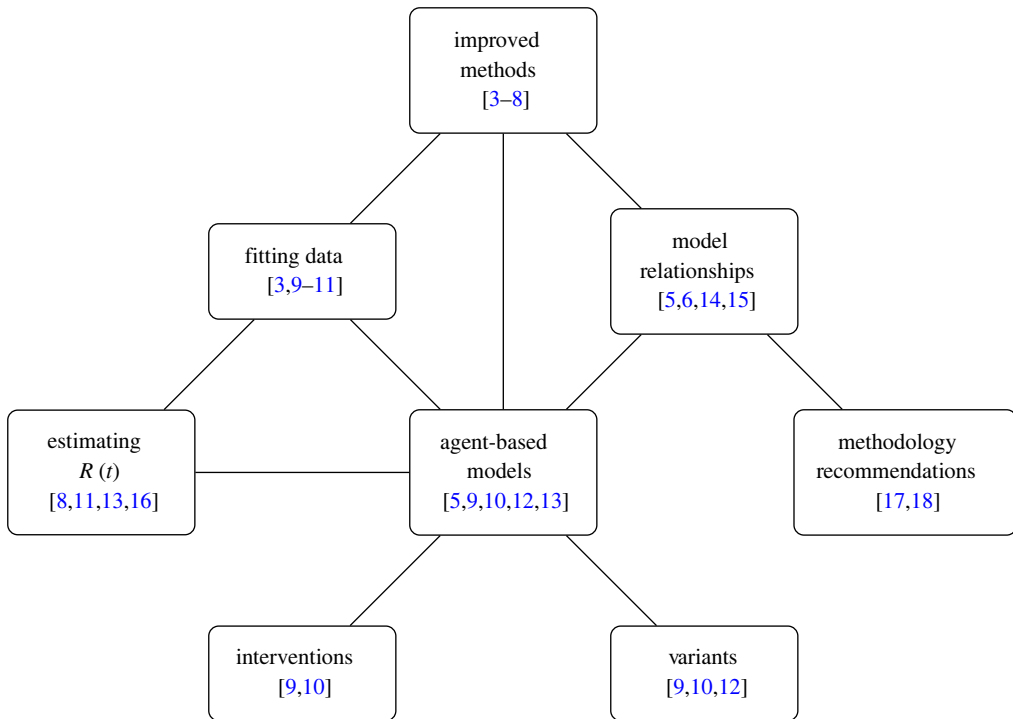


Figure 1. Conceptual map of the articles in this issue. Each article fits into at least one of these subject areas. Where an article treats more than one subject, an edge is drawn between them.

A measure that has been important for policy decision-making and public understanding of the pandemic is the basic reproduction number, R_0 , and its time-varying analogue, effective reproduction number, $R(t)$. Ackland *et al.* [9] show how to estimate this function in the context of UK data and explains why it behaves as it does. Li *et al.* [8] use an ABM to estimate $R(t)$ and demonstrate improved statistical methods to do this in the face of noisy data, while Creswell *et al.* [10] use classic modelling techniques to explore the role of imported COVID-19 cases on $R(t)$. Finally, Madzvamuse *et al.* [11] present analytical results showing how to reformulate the standard differential equation epidemic model so that it is expressed in terms of observed information rather than the unobserved actual state of the population.

Complementing the foregoing, which are concerned with particular models, are several papers that consider relationships between models and borrow methods from other aspects of mathematics to answer epidemiological questions. Fairbanks *et al.* [12] show how the mathematical property of compositionality can be used to construct complex compartmental models out of simpler ones in a semi-automated, modular, and type-safe way. Waites *et al.* [13] also demonstrate how the compositional and modular properties of rule-based models—used widely in physics—can be used to show how a simple immune response model can produce the heterogeneity that we observe at the population level. Probert *et al.* [14] also used rule-based methods—this time borrowed from vote-processing analogy—to explore the uncertainty of combining models' outcomes. Vernon *et al.* [15] show how simpler 'emulators' models can simulate more complex ABMs to assist in parameter estimation and sensitivity analysis. Swallow *et al.* [16] borrow methodology from statistics and use principal component analysis to derive key epidemic metrics across the four UK nations.

Finally, two papers take a high-level view and make recommendations about modelling in general. Mitchell *et al.* [17] emphasize the importance of provenance, describing a pipeline for managing data and promoting a practice of transparency and reproducibility in scientific

workflows that should be adopted in all modelling endeavours. Chen *et al.* [18] tackle visualization and make recommendations for the challenges surrounding communication of results.

In summary, this special issue addresses a number of questions around modelling that have not been considered in the rapid responsive modelling of the ongoing epidemics. Via a collection of papers that include multiscale, cross-discipline mathematical and statistical modelling, and focus on technical aspects of modelling rather than solely applications, while being timely and relevant to the current pandemic, this special issue aims to withstand time and be relevant for possible future pandemics also.

Data accessibility. This article has no additional data.

Authors' contributions. J.P.-G.: conceptualization, formal analysis, project administration, supervision, writing—original draft; W.W.: conceptualization, formal analysis, visualization, writing—review and editing; G.J.A.: conceptualization, formal analysis, project administration, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. This theme issue was put together by the Guest Editor team under supervision from the journal's Editorial staff, following the Royal Society's ethical codes and best-practice guidelines. The Guest Editor team invited contributions and handled the review process. Individual Guest Editors were not involved in assessing papers where they had a personal, professional, or financial conflict of interest with the authors, or the research described. Independent reviewers assessed all papers. Invitation to contribute did not guarantee inclusion.

Funding. No funding has been received for this article.

References

1. Royal Society. 2021 Rapid Assistance in Modelling the Pandemic (RAMP) initiative.
2. Ackland GJ, Panovska-Griffiths J, Waites W, Cates ME. 2022 The Royal Society RAMP modelling initiative. *Phil. Trans. R. Soc. A* **380**, 20210316. (doi:10.1098/rsta.2021.0316)
3. Pooley CM, Doeschl-Wilson AB, Marion G. 2022 Estimation of age-stratified contact rates during the COVID-19 pandemic using a novel inference algorithm. *Phil. Trans. R. Soc. A* **380**, 20210298. (doi:10.1098/rsta.2021.0298)
4. Moore RE, Rosato C, Maskell S. 2022 Refining epidemiological forecasts with simple scoring rules. *Phil. Trans. R. Soc. A* **380**, 20210305. (doi:10.1098/rsta.2021.0305)
5. Panovska-Griffiths J *et al.* 2022 Statistical and agent-based modelling of the transmissibility of different SARS-CoV-2 variants in England and impact of different interventions. *Phil. Trans. R. Soc. A* **380**, 20210315. (doi:10.1098/rsta.2021.0315)
6. Hinch R *et al.* 2022 Estimating SARS-CoV-2 variant fitness and the impact of interventions in England using statistical and geo-spatial agent-based models. *Phil. Trans. R. Soc. A* **380**, 20210304. (doi:10.1098/rsta.2021.0304)
7. Sanz-Leon P *et al.* 2022 Modelling herd immunity requirements in Queensland: impact of vaccination effectiveness, hesitancy and variants of SARS-CoV-2. *Phil. Trans. R. Soc. A* **380**, 20210311. (doi:10.1098/rsta.2021.0311)
8. Li W, Bulekova K, Gregor B, White LF, Kolaczyk ED. 2022 Estimation of local time-varying reproduction numbers in noisy surveillance data. *Phil. Trans. R. Soc. A* **380**, 20210303. (doi:10.1098/rsta.2021.0303)
9. Ackland GJ, Ackland JA, Antonioletti M, Wallace DJ. 2022 Fitting the reproduction number from UK coronavirus case data and why it is close to 1. *Phil. Trans. R. Soc. A* **380**, 20210301. (doi:10.1098/rsta.2021.0301)
10. Creswell R *et al.* 2022 Heterogeneity in the onwards transmission risk between local and imported cases affects practical estimates of the time-dependent reproduction number. *Phil. Trans. R. Soc. A* **380**, 20210308. (doi:10.1098/rsta.2021.0308)
11. Campillo-Funollet E, Wragg H, Van Yperen J, Duong D-L, Madzvamuse A. 2022 Reformulating the susceptible–infectious–removed model in terms of the number of detected cases: well-posedness of the observational model. *Phil. Trans. R. Soc. A* **380**, 20210306. (doi:10.1098/rsta.2021.0306)

12. Libkind S, Baas A, Halter M, Patterson E, Fairbanks JP. 2022 An algebraic framework for structured epidemic modelling. *Phil. Trans. R. Soc. A* **380**, 20210309. (doi:10.1098/rsta.2021.0309)
13. Waites W *et al.* 2022 Compositional modelling of immune response and virus transmission dynamics. *Phil. Trans. R. Soc. A* **380**, 20210307. (doi:10.1098/rsta.2021.0307)
14. Probert WJM, Nicol S, Ferrari MJ, Li S-L, Shea K, Tildesley MJ, Runge MC. 2022 Vote-processing rules for combining control recommendations from multiple models. *Phil. Trans. R. Soc. A* **380**, 20210314. (doi:10.1098/rsta.2021.0314)
15. Vernon I *et al.* 2022 Bayesian emulation and history matching of JUNE. *Phil. Trans. R. Soc. A* **380**, 20220039. (doi:10.1098/rsta.2022.0039)
16. Swallow B, Xiang W, Panovska-Griffiths J. 2022 Tracking the national and regional COVID-19 epidemic status in the UK using weighted principal component analysis. *Phil. Trans. R. Soc. A* **380**, 20210302. (doi:10.1098/rsta.2021.0302)
17. Natalie Mitchell S *et al.* 2022 FAIR data pipeline: provenance-driven data management for traceable scientific workflows. *Phil. Trans. R. Soc. A* **380**, 20210300. (doi:10.1098/rsta.2021.0300)
18. Dykes J *et al.* 2022 Visualization for epidemiological modelling: challenges, solutions, reflections and recommendations. *Phil. Trans. R. Soc. A* **380**, 20210299. (doi:10.1098/rsta.2021.0299)