Improving estimates of epidemiological quantities with wastewater data in Aotearoa New Zealand

Leighton M. Watson, Michael J. Plank, Bridget A. Armstrong, Joanne R. Chapman, Joanne Hewitt, Helen Morris, Alvaro Orsi, Michael Bunce, Christl A. Donnelly, **Nicholas Steyn***

*nicholas.steyn@univ.ox.ac.uk, github.com/nicsteyn2/NZWastewaterModelling

This project is a collaboration between the University of Oxford (UK), the University of Canterbury (NZ), and the Institute of Environmental Science and Research (NZ).

UNIVERSITY OF OXFORD

Motivation

- The reproduction number R_t is typically estimated using reported cases
- Tracking wastewater concentrations of viral material has been proposed as an alternative source of data
- Can we use both simultaneously?

The Institute of Environmental Sciences and Research New Zealand collected wastewater sampling data throughout the COVID-19 pandemic [1]. Figure 1 (below) shows reported cases and wastewater sampling data between Jan 2022 and Mar 2023.

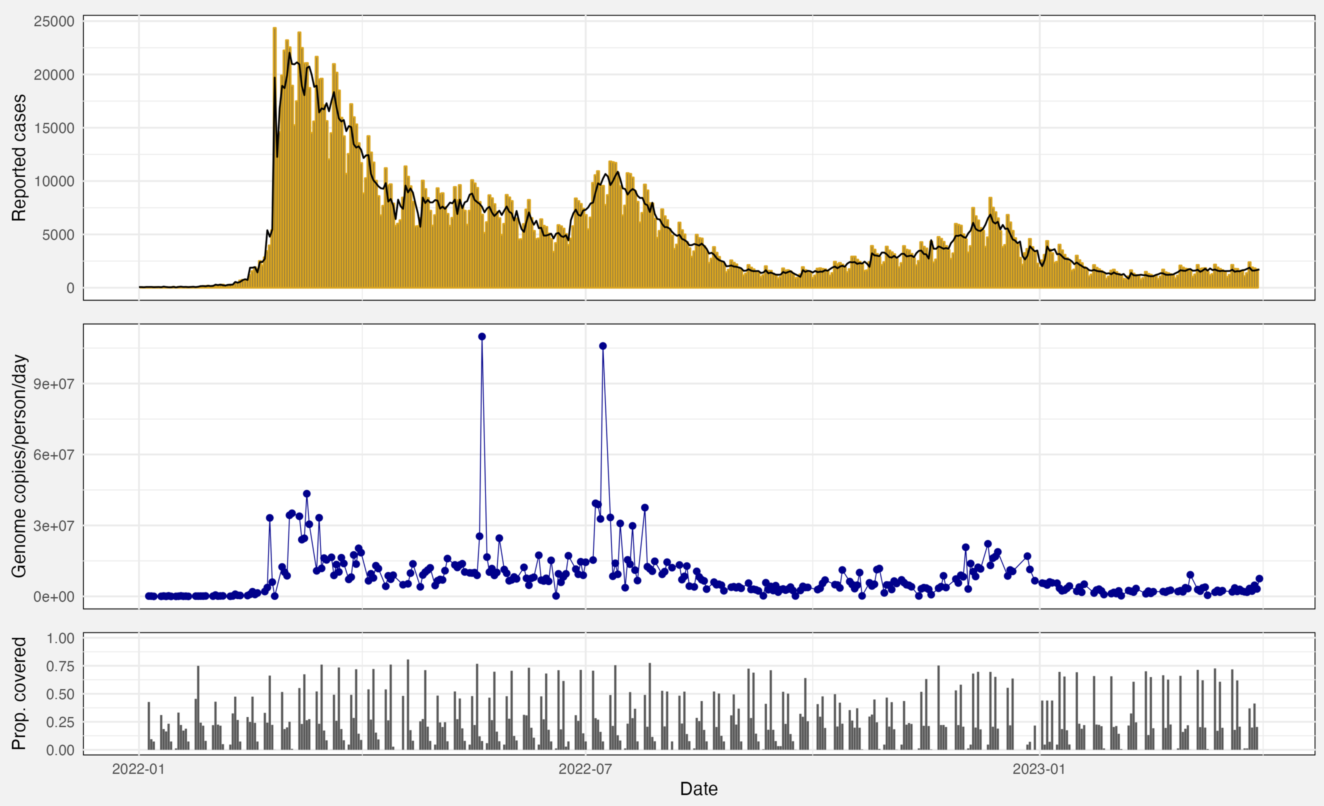


Figure 1: National data from Aotearoa New Zealand. Daily reported COVID-19 cases (orange bars), detrended cases (black line), SARS-CoV-2 concentration in wastewater (blue line and dots), and proportion of population covered by sampled wastewater catchments (gray bars).

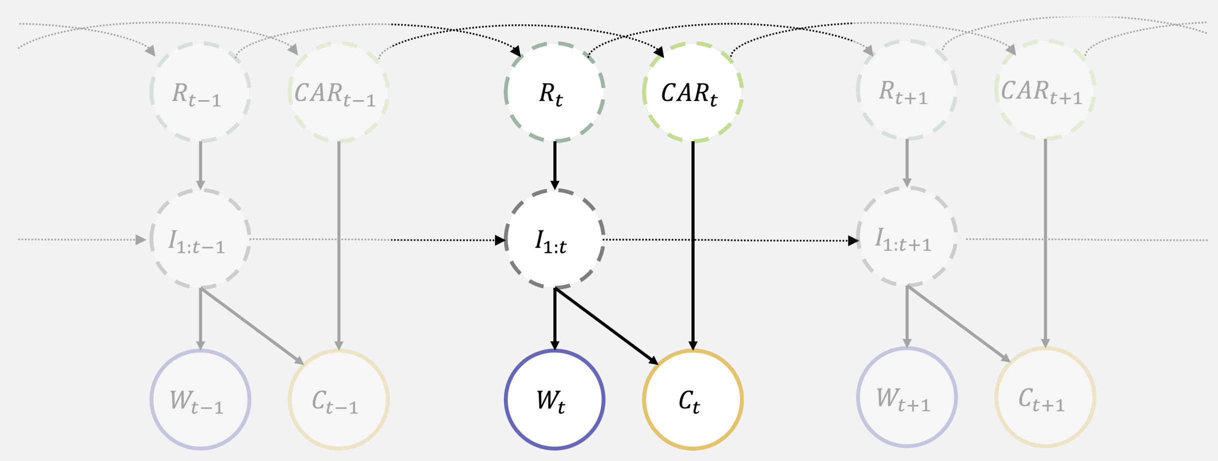
The model...

...is a hidden-state model that relates **observed data:**

- Reported cases C_t
- Wastewater samples W_t

to **unobserved states:**

- Reproduction number R_t
- Case ascertainment ratio CAR_t
- Infection incidence I_t



We let R_t and CAR_t vary smoothly over time, and let I_t follow a Poisson renewal model:

$$R_t \sim \text{Normal}(R_{t-1}, \sigma_R R_{t-1})$$

$$CAR_t \sim \text{Normal}(CAR_{t-1}, \sigma_{CAR})$$

$$I_t \sim \text{Poisson}(R_t \sum I_{t-u} g_u)$$

C_t is modelled using a negative binomial distribution and W_t using a Gamma distribution with expected values:

$$E[C_t] = CAR_t \sum I_{t-u} L_u$$

$$E[W_t] = \alpha \sum I_{t-u} \omega_u$$

g_u, L_u, and ω_u are the generation interval, case reporting delay, and wastewater shedding distributions respectively. α is the average total detectable viral material shed by an infected person.

References

[1] Joanne Hewitt et al. Sensitivity of wastewater-based epidemiology for detection of SARS-CoV-2 RNA in a low prevalence setting. *Water Research*, 211:118032, March 2022.

[2] Leighton M. Watson et al. Improving estimates of epidemiological quantities by combining reported cases with wastewater data: A statistical framework with applications to COVID-19 in Aotearoa New Zealand. *medRxiv preprint*, August 2023.

Results

We solve for R_t, CAR_t, and I_t using a particle filter with fixed lag resampling. The particle marginal Metropolis Hastings algorithm is used to fit model parameters. The final output is a collection of samples from the posterior distributions of our unobserved states at each time-step given our observed data $P(R_t, CAR_t, I_t | C_{1:T}, W_{1:T})$. These are presented in Figure 2:

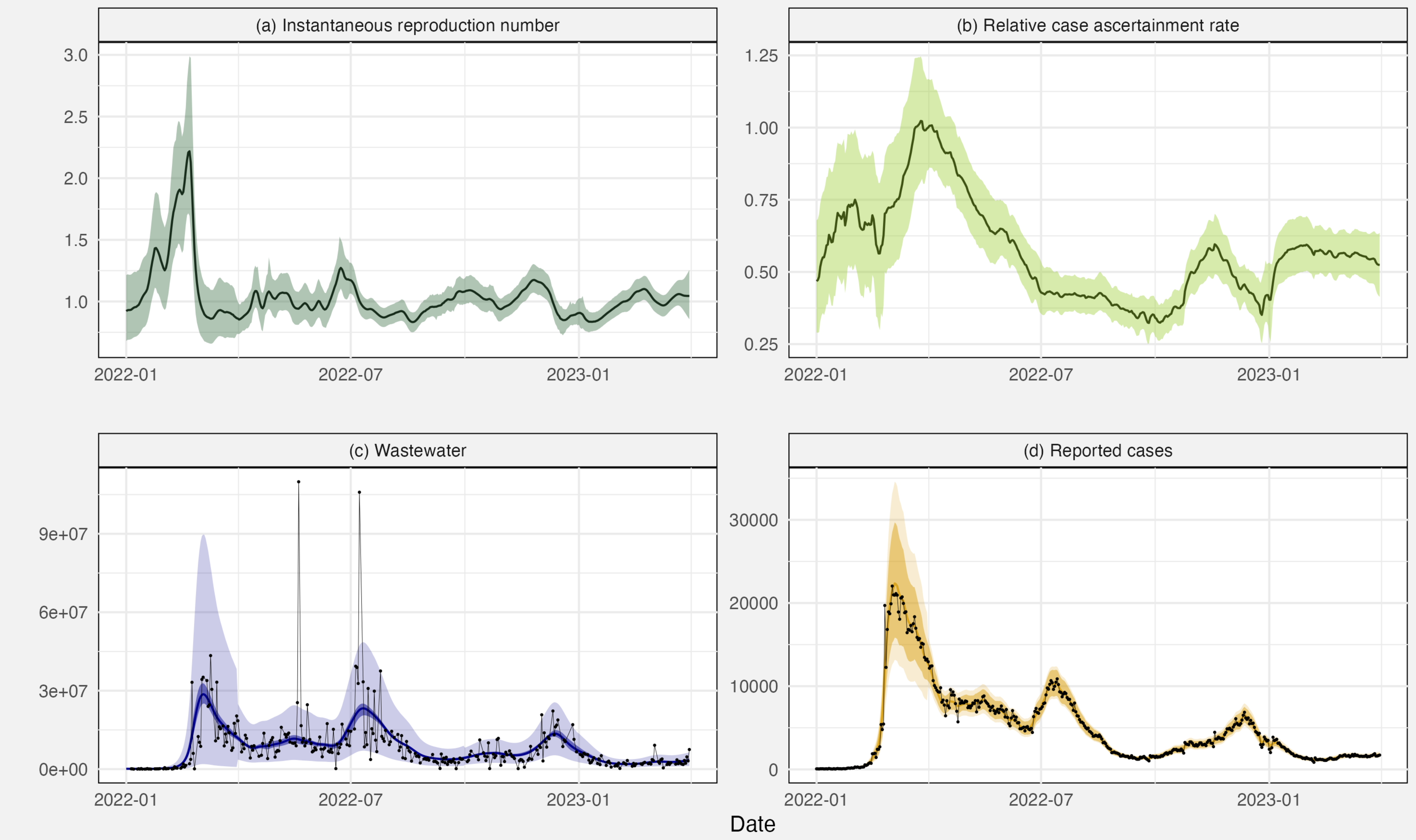


Figure 2: Estimated epidemiological quantities for SARS-CoV-2 in Aotearoa New Zealand. Shaded regions show 95% credible intervals. Lighter shading in panels (c) and (d) shows 95% predictive intervals (accounting for observation noise). Black dots and lines show observed data. CAR_t is reported relative to 1 April 2022, assuming α is constant.

Wastewater versus reported cases (work in progress)

- Wastewater sampling data are very noisy, so may we expect them to be less informative about R_t than reported cases.
- However, models based solely on reported cases assume that CAR_t is constant!
- What happens if we allow CAR_t to vary in a cases-only model?
- We can answer this by solving our model using only one source of data at a time.

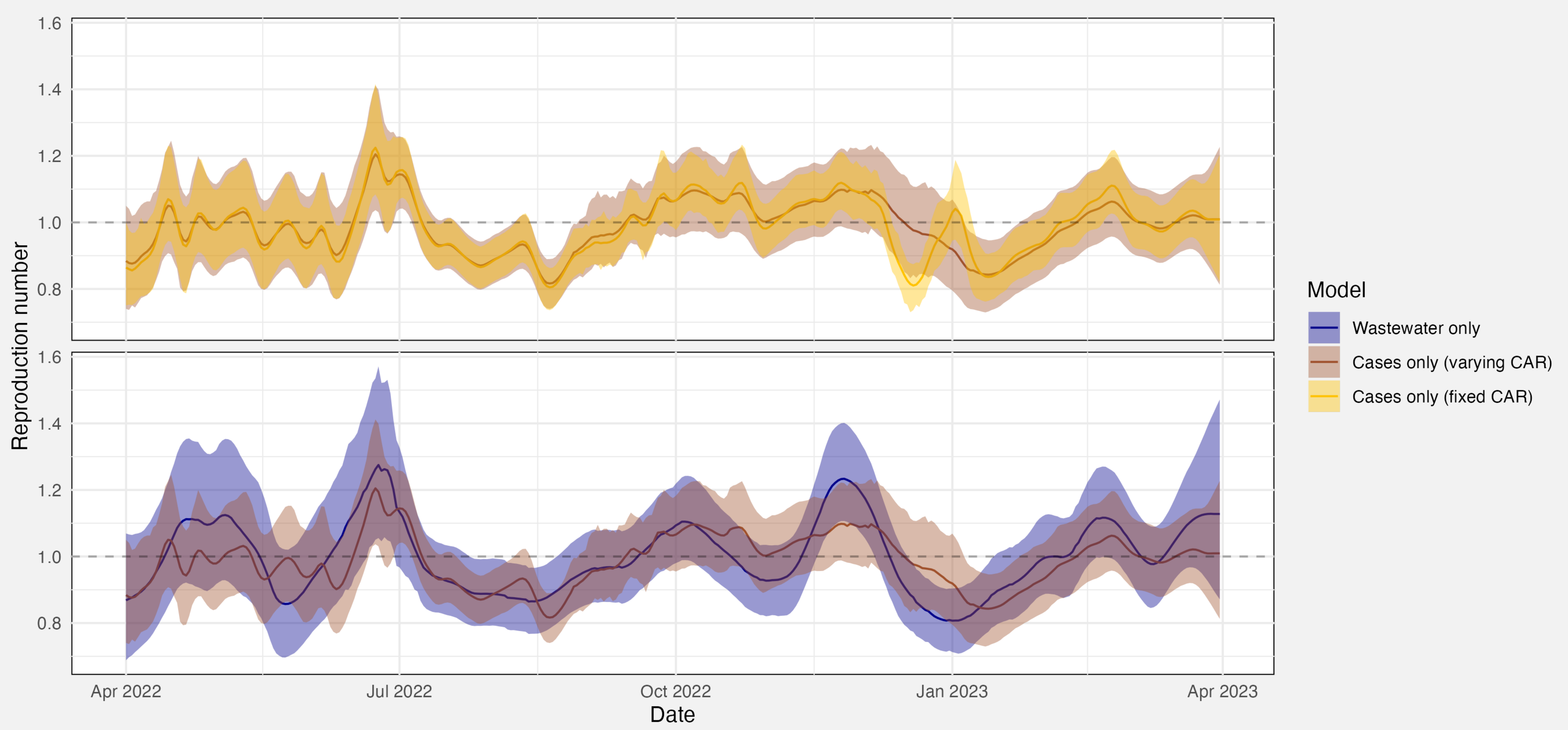


Figure 3: Estimates of R_t from wastewater data only, reported cases only with a variable CAR_t, and reported cases only with a fixed CAR_t. Shaded regions show 95% credible intervals. For presentation purposes, we show results between April 2022 and March 2023.

- Uncertainty in R_t (when fit to reported cases alone) is higher when we allow CAR_t to vary, than when we assume CAR_t is constant.
- In both scenarios, wastewater data appear to be less informative about R_t than reported cases. Particularly when we assume CAR_t is constant.
- This difference is less clear when CAR_t is allowed to vary.
- **As testing (or consistency in testing) decreases, wastewater could plausibly become the more informative source of data (with respect to R_t).**

In this work, we assume the **detectable viral material shedding rate** α is fixed. We have just kicked the can down the road!

Future work: can we use epidemic surveys, like REACT-1 in England, to quantify how α varies over time?