HKU Med LKS Faculty of Medicine School of Public Health 香港大學公共衛生學院

25

Weijia Xiong (PhD)
Primary Supervisor: Prof Tim K. Tsang

Nowcasting time-varying effective reproductive number in Hong Kong, 2020-2022

Weijia Xiong[1], Simon Cauchemez[2], Benjamin J. Cowling[1,3], Tim K. Tsang[1,3]

[1] WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China
[2] Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Université Paris Cité, UMR2000, CNRS, Paris, France.
[3] Laboratory of Data Discovery for Health Limited (D²4H), Hong Kong Science and Technology Park, New Territories, Hong Kong Special Administrative Region, China

Background

Real-time tracking of transmission of an infectious disease is critical to guide control policy. The time-varying effective reproductive number (Rt) has been widely used to quantify the transmissibility. However, it suffers lags of more than a week due to the latent period and reporting delays.

Objectives

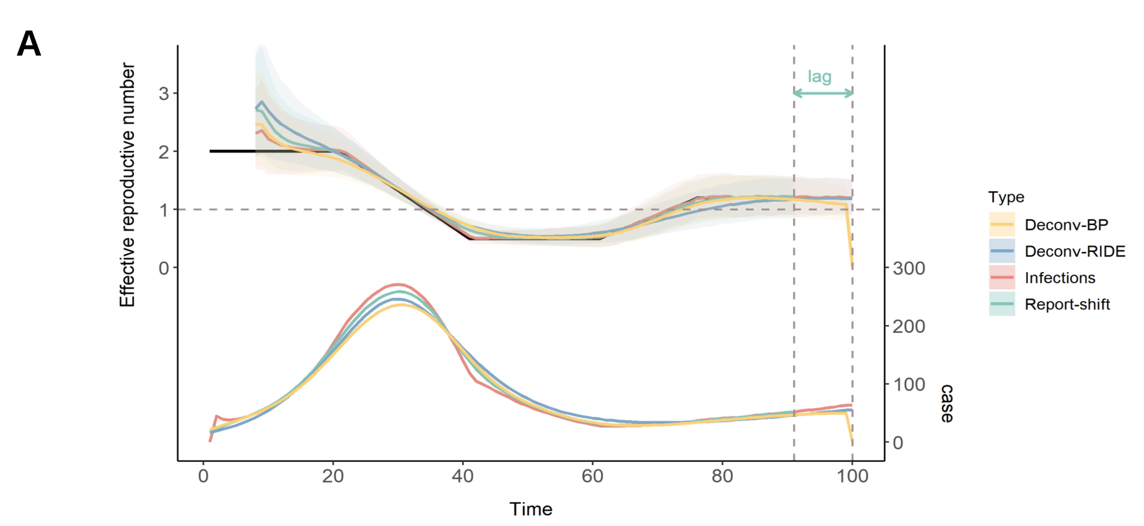
To overcome this, we develop a framework for nowcasting and forecasting R_t up to 7days ahead for COVID-19, incorporating previous covariates including epidemiological, meteorological and policy indices predictors.

Methods

We use the framework in Cori et al.[1] to estimate R_t based on the PCR-confirmed case number and apply a robust incidence deconvolution estimator (RIDE) in the deconvolution of time series of cases, with correction of censoring to estimate R_t in real-time(here after denoted as temporal R_t). Back-projection(BP) deconvolution approach is included for comparison.

We test models that integrate temporal R_t, from the previous 14 days to the current day, policy indices data, and meteorological data to perform nowcast and forecast the R_t up to 7 days ahead. Eight models are tested, including Autoregressive integrated moving average model (ARIMA), Generalized additive model (GAM), Gradient Boosting algorithm (GBM), Gaussian process regression model (GPR), Random Forest model (RF), Ridge model (RIDGE), Neural Network (NN) and support vector regression with linear kernels (SVR). For each model, we perform a forward stepwise selection method based on RMSE to choose predictors to improve the performance of our model. RIDE and BP constant models are included, wherein the forecast is set to the same constant value as the nowcast.

STEP 1： Deconvolution
Back-calculate the time series of infections from report case and estimated R_t based on Cori's method.



STEP 2：Correction
Incorporate other data sources to reduce the possible gap between R_t^k (estimated from deconvolution infection time series with data up to time k) and the R_t (estimated from data of whole study period)

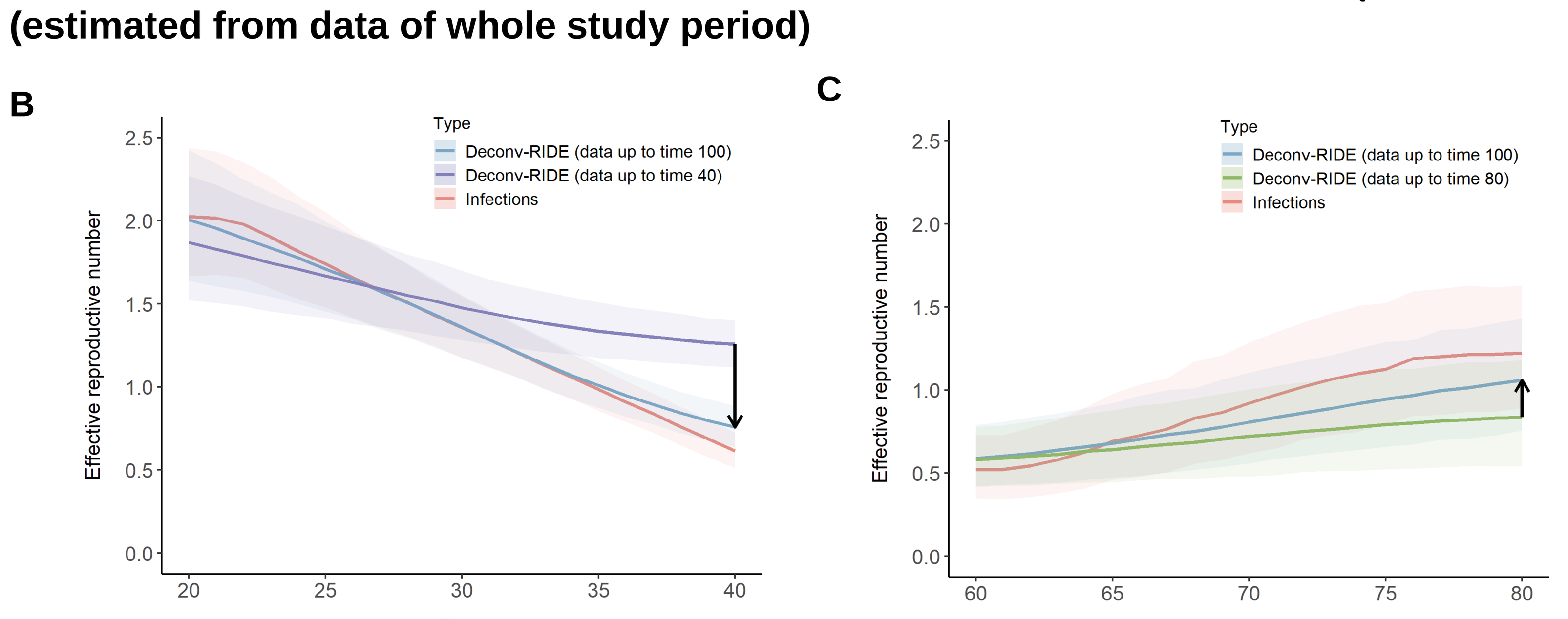


Figure1. Schematic of modeling approach.

Results

Our study applies this approach to estimate the R_t from December 2020 to December 2022 in Hong Kong. We analyze Omicron wave data from fifth wave (training) to the sixth wave (testing) and used the data from the fourth wave (ancestral strains) as another testing set to ensure the reliability and robustness(different circulating strains SARS-CoV-2).

Our results show that covariate ARIMA model performs best, could reduce the 7-day root mean square error by 29% for nowcast and 34% for 7-day forecast over the testing period, and sustain <10% mean absolute percentage error up to 7-day ahead forecast, in the sixth wave. In addition, the covariate ARIMA model demonstrates more precise trends when R_t is on the rise or declining towards. Moreover, the prediction performance of the ARIMA model is not sensitive to selected predictors.

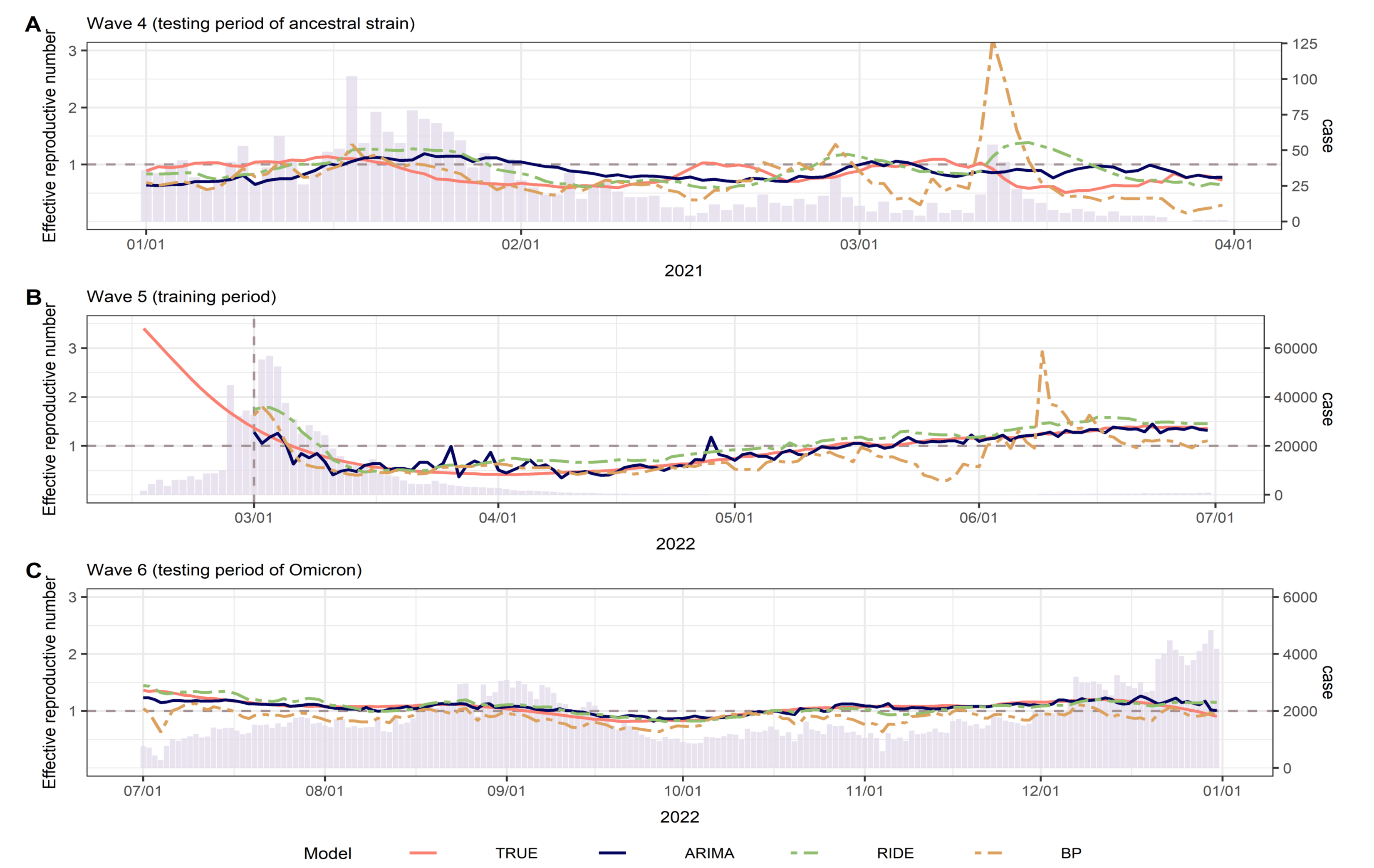


Figure2. Each panel shows the confirmed COVID-19 local cases and estimated R_t in Hong Kong during each epidemic wave from 2021 to 2022. The grey bars indicate daily numbers of COVID-19 cases confirmed by PCR for cases. The red lines indicate the R_t estimate at the end of outbreaks. The dashed green line and orange line represent the nowcast R_t estimated by RIDE, BP model respectively. The dark blue line indicates the nowcast R_t by the ARIMA models.

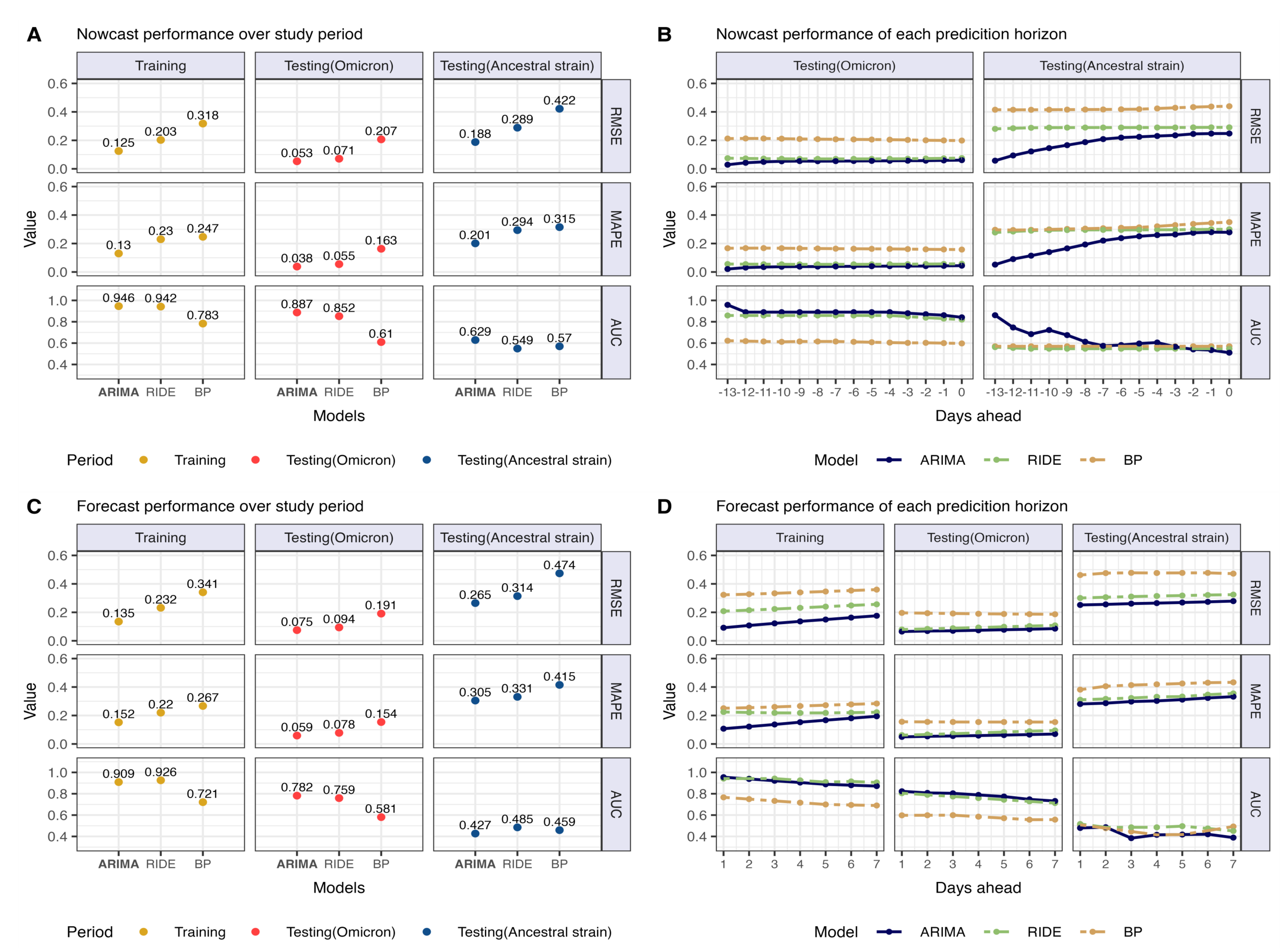


Figure3. Nowcast and forecast over study period. Panel A,C show the 7-day RMSE of each over the training period and testing period. Panel B, D show the RMSE of selected models at each prediction horizon.

Conclusion

In this study, we systematically evaluate the performance of different approaches to estimate R_t of SARS-COV-2 transmission in real-time, including deconvolution approaches to overcome the right censoring of cases, and statistical models to include temporal predictors to conduct nowcast and up to 7-day ahead forecast of R_t from 2020-2022 in Hong Kong. We find that using more sophisticated approaches for deconvolution and a simple ARIMA model could reduce the MAPE by 69% compared with the commonly used BP approach in the testing period. In conclusion, we demonstrate a comprehensive approach to improve real-time estimation and short-term forecasts of R_t, which could allow for timely and accurate tracking of epidemic transmission dynamics.

References

[1] Cori A, Ferguson NM, Fraser C, Cauchemez S. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. American Journal of Epidemiology 178, 1505-1512 (2013).

Acknowledgements

This project was supported by the Theme-based Research Scheme (Project No. T11-705/21-N) of the Research Grants Council of the Hong Kong SAR Government.