A microscopic view of a coronavirus particle, showing its characteristic spherical shape and surface spikes, set against a light blue background.

NOVEL CORONAVIRUS, COVID-2019 IN CHINA

A DATA ANALYSIS TOOLBOX FROM SONG LAB  
U-M SCHOOL OF PUBLIC HEALTH

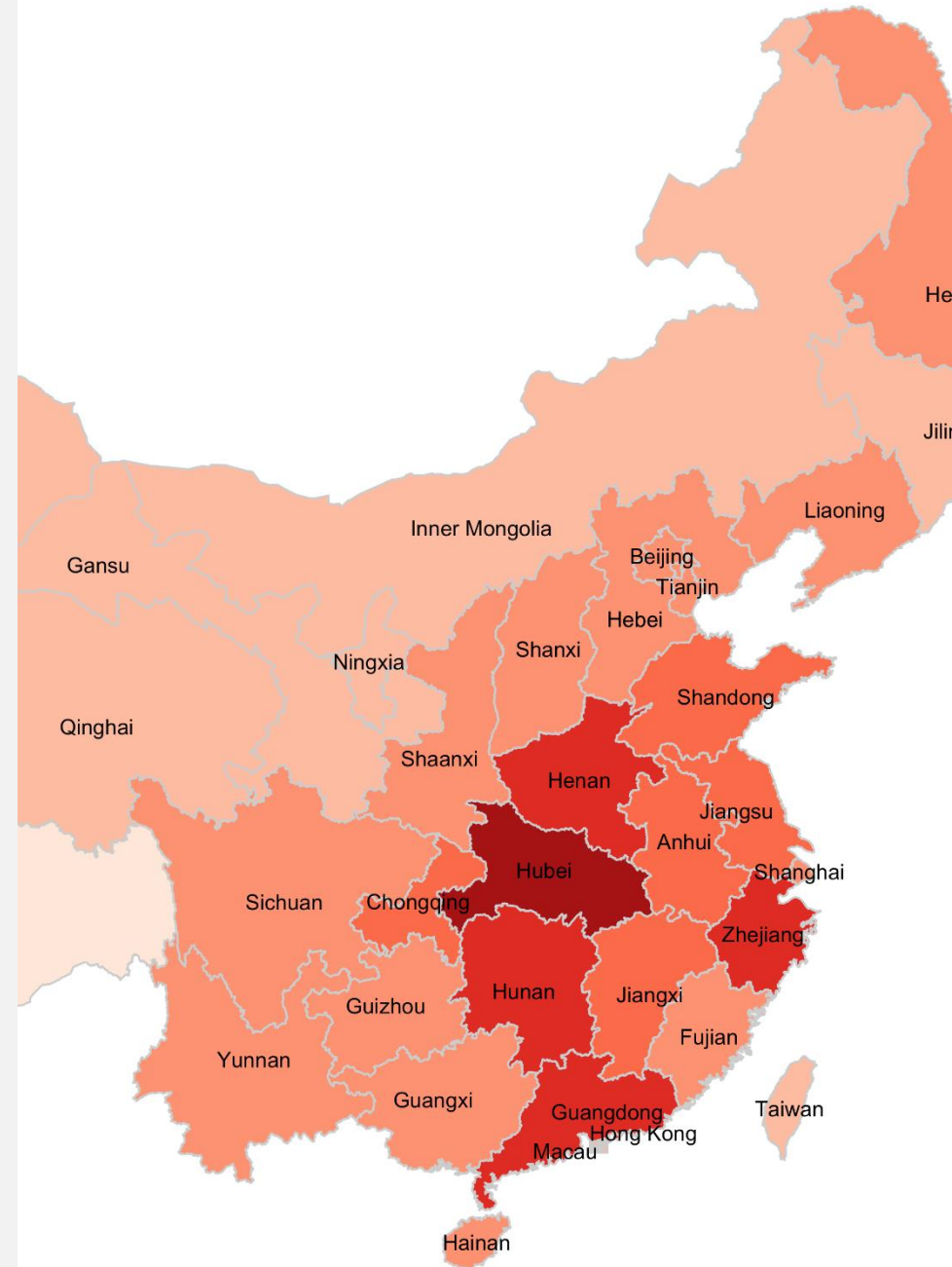
Lili Wang, Jie He, Yiwang Zhou, Bin Zhu, Fei Wang, Lu Tang

Marisa Eisenberg and Peter Song

*February 17, 2020*

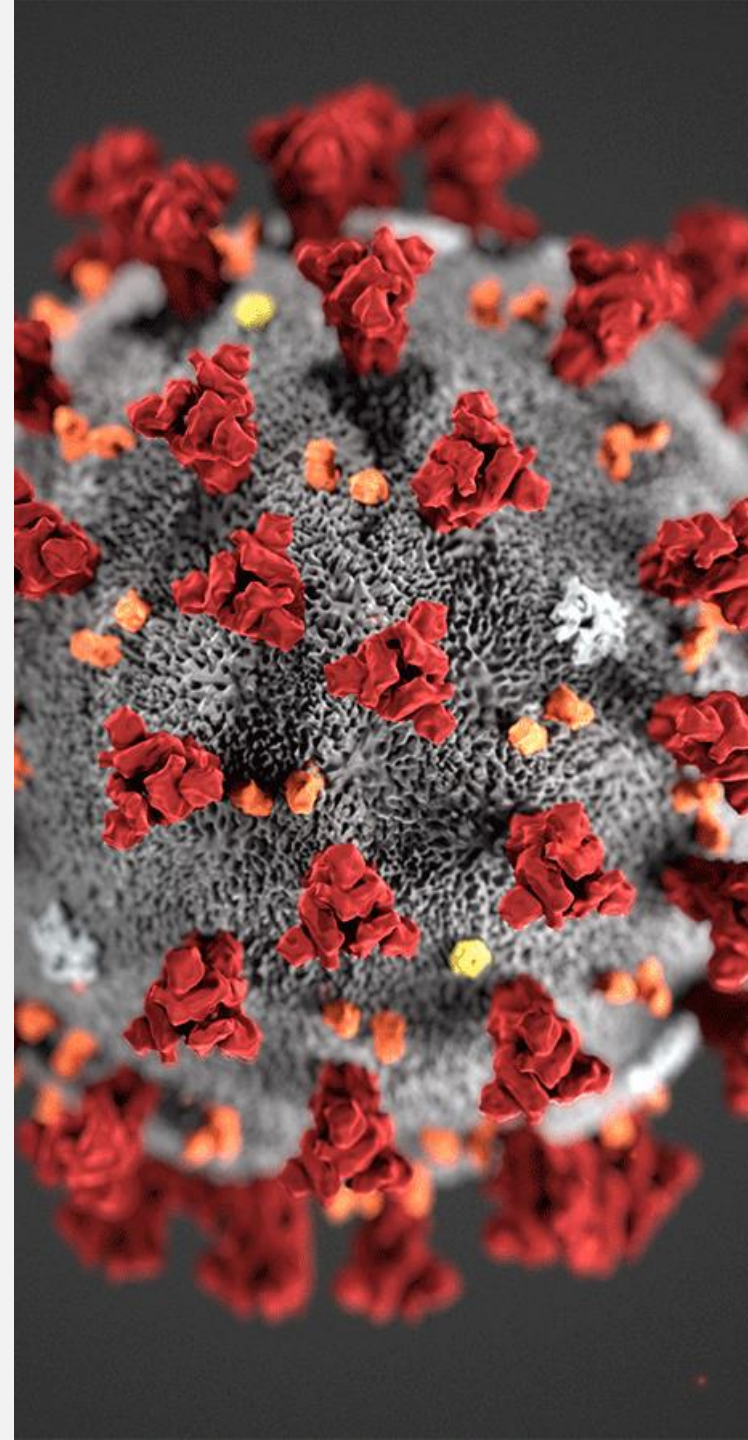
## COVID-2019 VIRUS

- The outbreak of novel coronavirus that began in Wuhan in December 2019 spreads quickly in the province of Hubei, and then China as well as many cities in the world.
- As of Feb 15th, there are a total of **68,595** infectious cases, **1,667** deaths, and **9,763** recovered cases.
- There are still **8,228** suspected cases.



## CURRENT KNOWLEDGE ABOUT COVID-2019

- An infectious disease caused by SARS-CoV-2, a virus closely related to the SARS virus.
- Spreading from person to person.
- High transmission rate, resulting in a large number of infected cases, especially in Hubei Province.
- The median incubation period is 3 days, which can be as long as 24 days.
- Highly likely infectious by asymptomatic person.



## QUARANTINE IS THE MOST EFFECTIVE

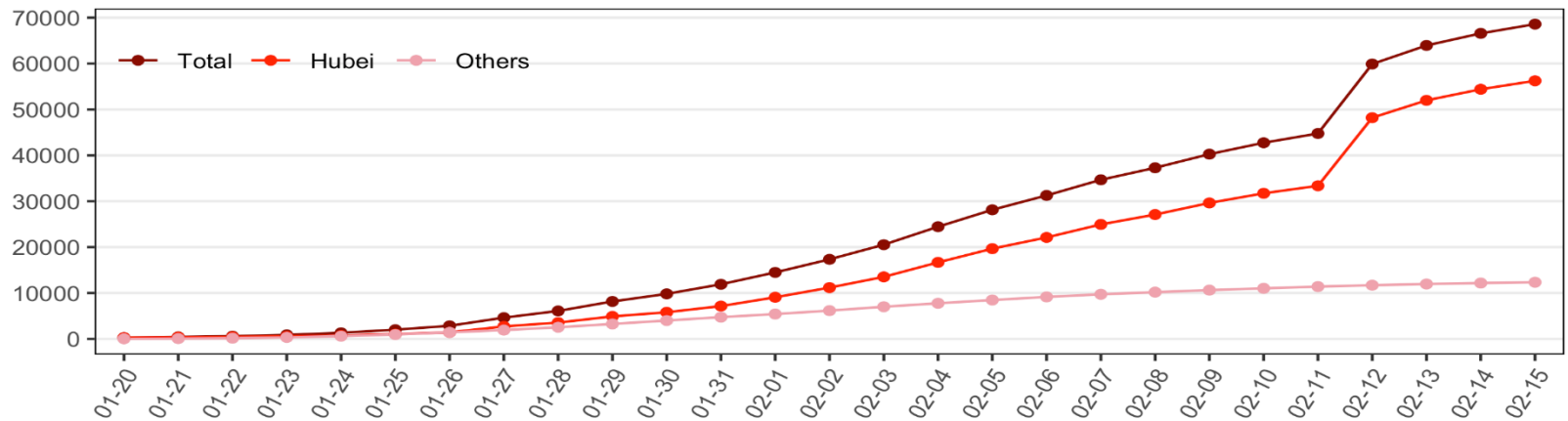
- Quarantine is the most effective way to stop spreading infectious disease, as proved again and again by all deadly infectious diseases in the human history (SARS, plague etc.).
- The Chinese government has started very strict in-home isolation protocols at all levels to control the transmission of COVID-2019.
- In the meanwhile, the government quickly increased the capacity of hospitals that took symptomatic patients to be quarantined in hospitals or as such.

## OUR OBJECTIVES

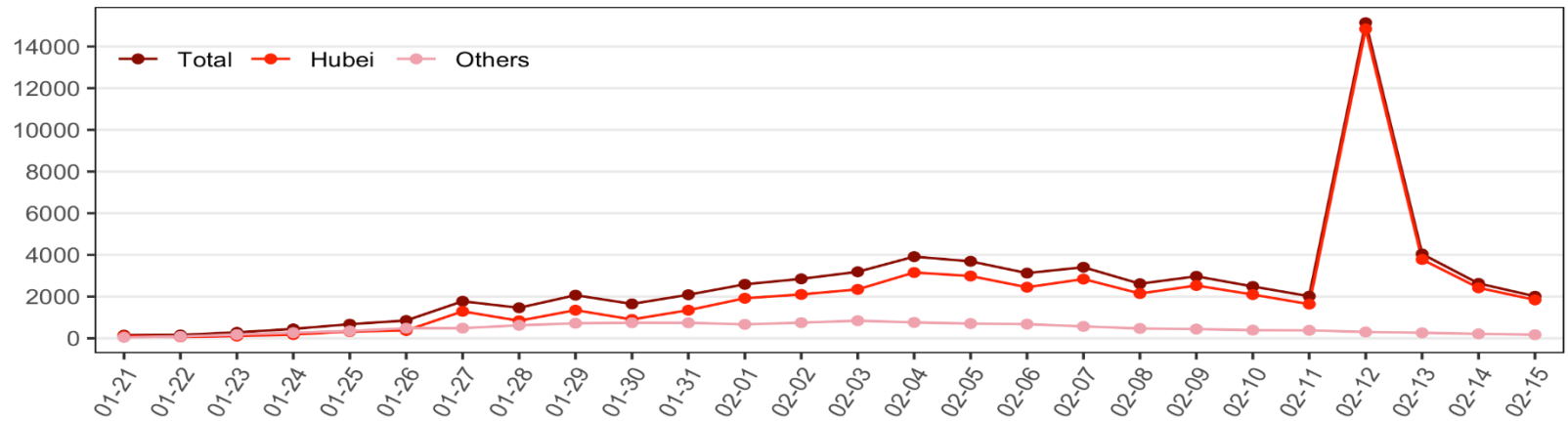
- **AIM 1:** Utilize publicly available data from China CDC to understand the trend of the coronavirus epidemic in Hubei Province and the other regions of China.
- **AIM 2:** Incorporate time-varying quarantine protocols in the modeling of infection dynamics using the classical SIR model (a system of differential equations) for infectious disease. **This dynamic infection system necessitates forecast the future trend of this epidemic.**
- **AIM 3:** Provide an R software package to CDCs for their own analyses of disease spreading patterns using their own data.
- **By no means in this project we attempt to predict turning points, but rather we provide an analytic toolbox to people who have good quality data.**

# AVAILABLE DATA: CUMULATIVE AND DAILY INFECTED CASES

**a** Cumulative number of infectious patients.



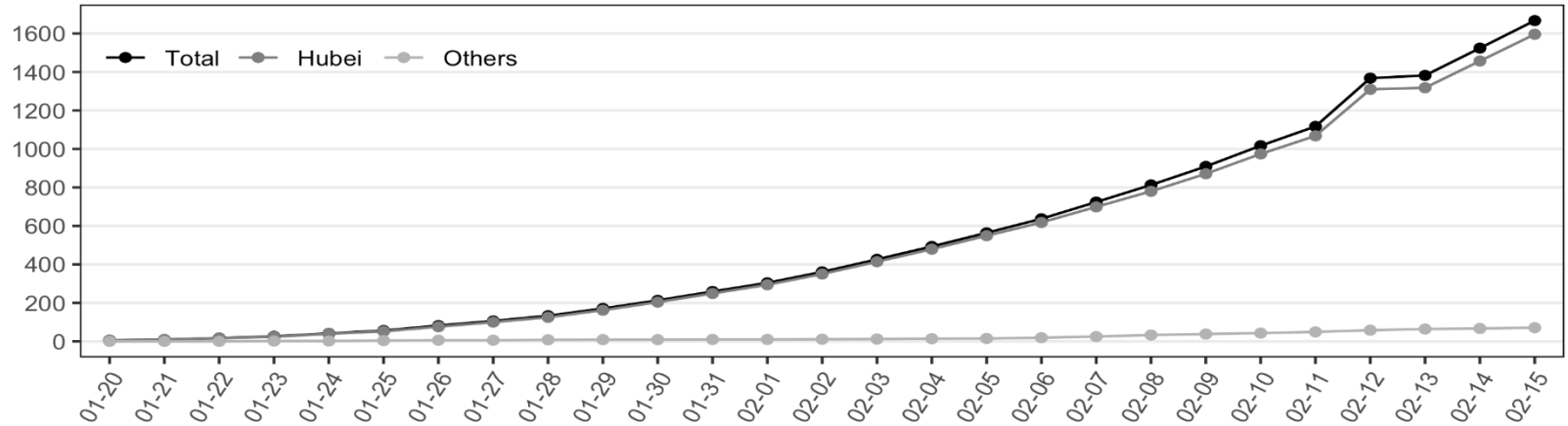
**b** Daily added number of infectious patients.



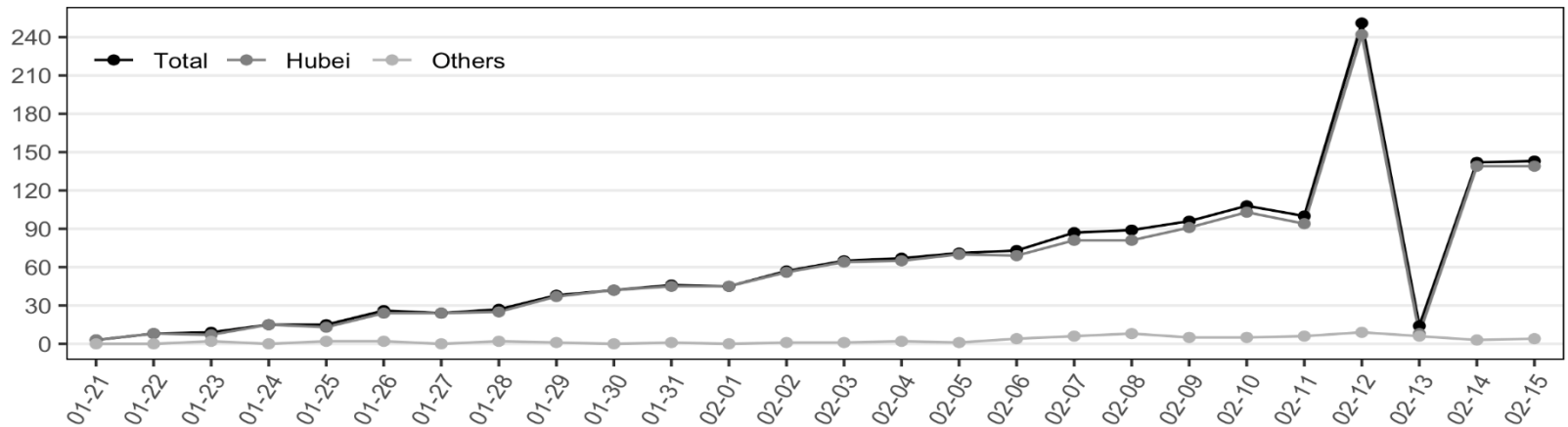


# AVAILABLE DATA: CUMULATIVE AND DAILY DEATHS

**a** Cumulative number of dead patients.

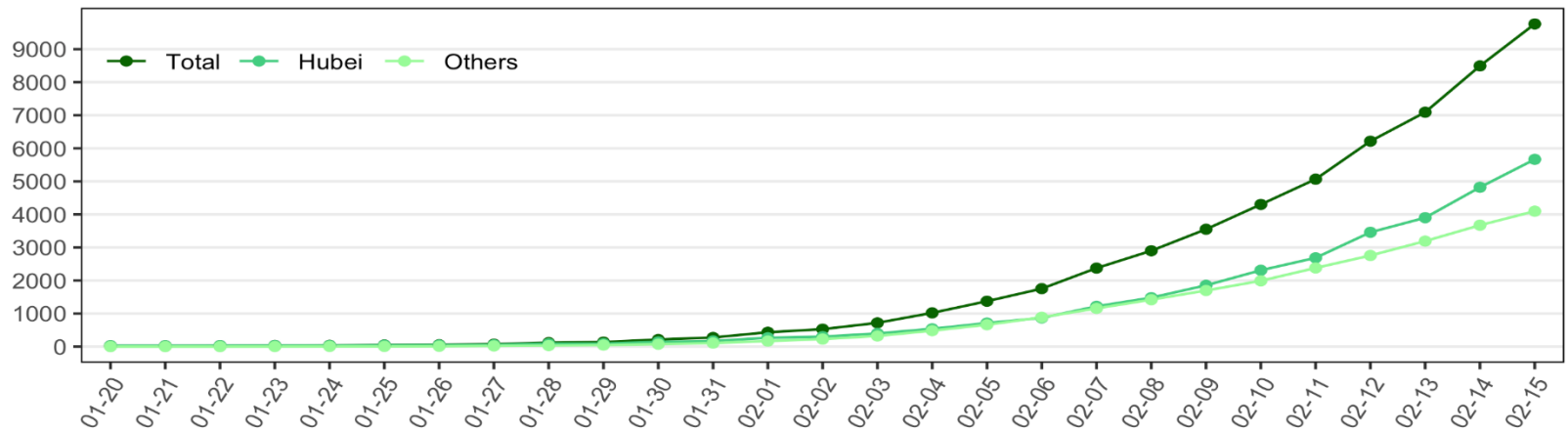


**b** Daily added number of dead patients.

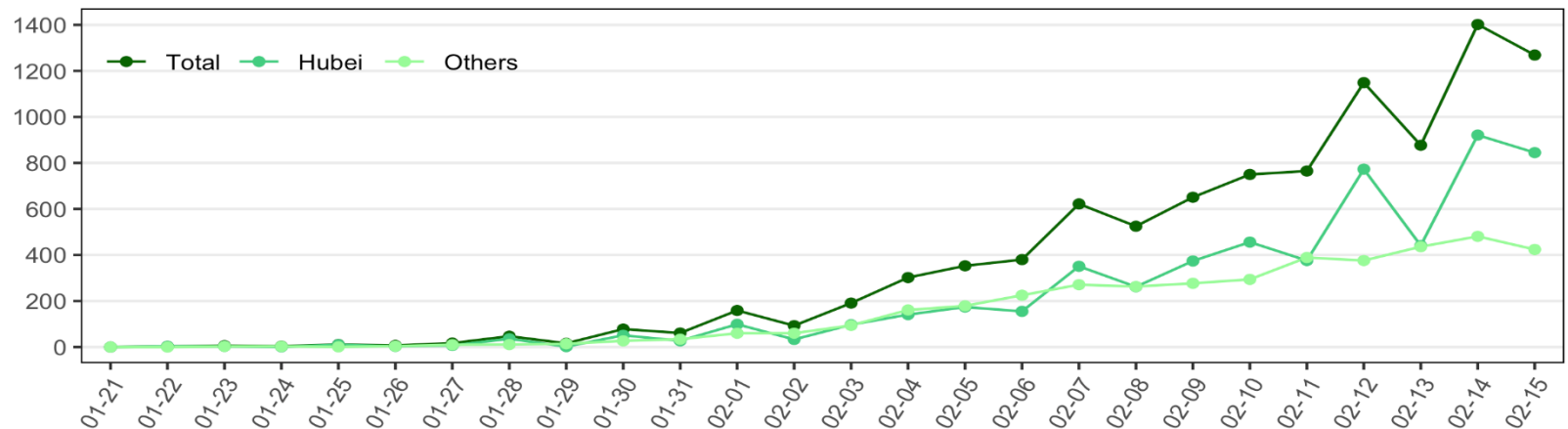


# AVAILABLE DATA: CUMULATIVE AND DAILY RECOVERED CASES

**a** Cumulative number of recovered patients.



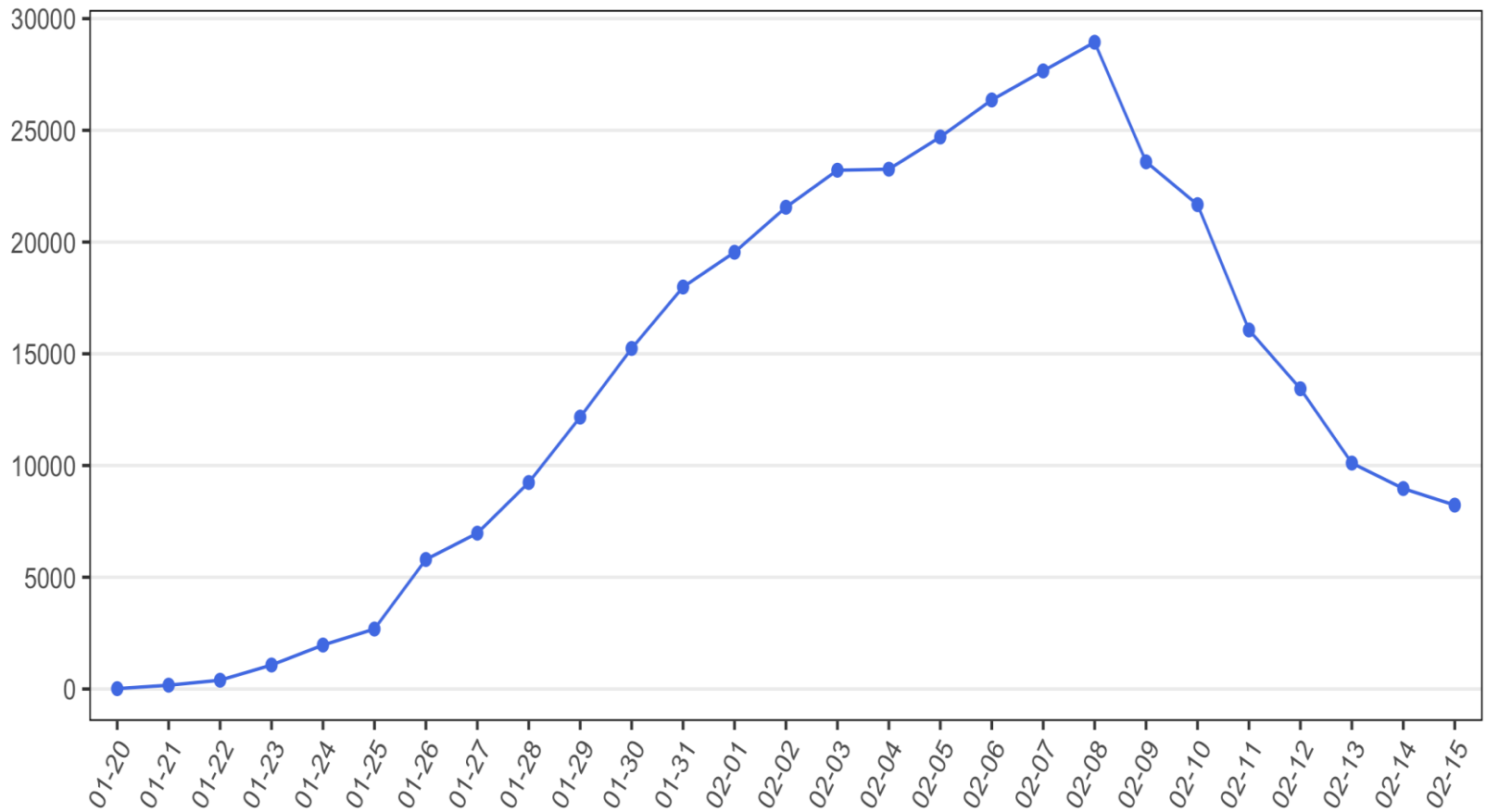
**b** Daily added number of recovered patients.





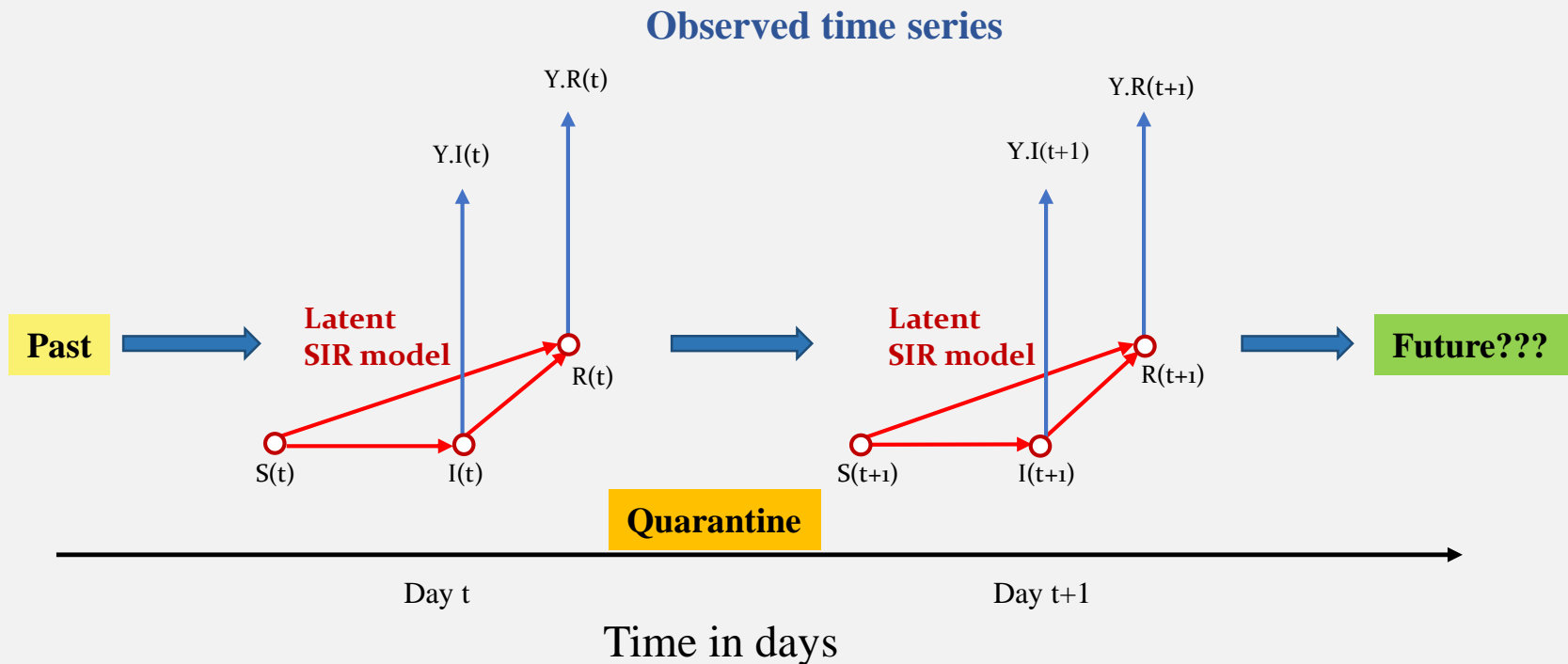
# AVAILABLE DATA: CUMULATIVE AND DAILY SUSPECTED CASES

Number of suspected patients.



# STATE-SPACE SIR MODEL

- Fitting the coronavirus data, including the respective time series of infected cases, recovered cases and deaths, with the dynamic SIR model (**a system of ordinary differential equations**) of infectious disease.



# THE STATE-SPACE SIR MODELING TOOLBOX

- Modeled the **processes of proportions** of susceptible, infected and removed (recovered and death) compartments
- Implemented estimation and inference by Markov Chain Monte Carlo (MCMC)
- Posteriors of model parameters (e.g. transmission and removal) and predicted proportions with their credible intervals
- **Two turning points: (i) When daily number of infected becomes smaller than previous ones; (ii) When daily number of removed cases is bigger than that of infected cases**
- An R package eSIR is available.