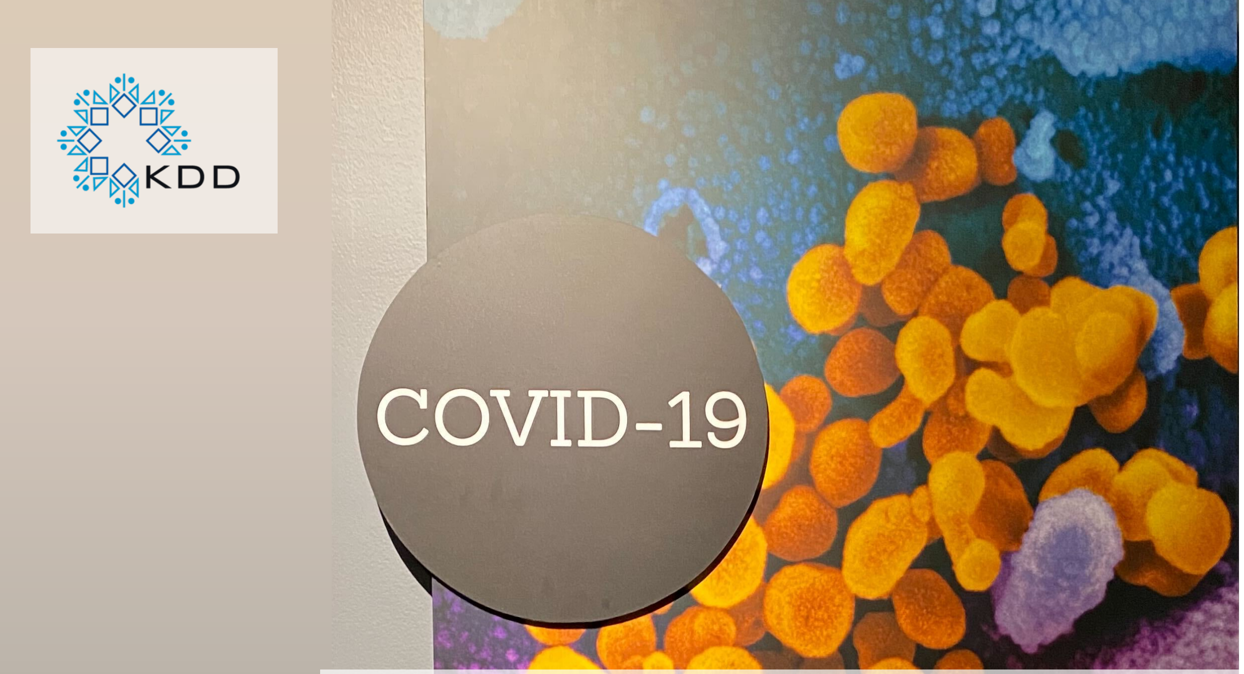



# Fast Mining and Forecasting of Co-evolving Epidemiological Data Streams

Tasuku Kimura, Yasuko Matsubara, Koki Kawabata, Yasushi Sakurai

SANKEN, Osaka University

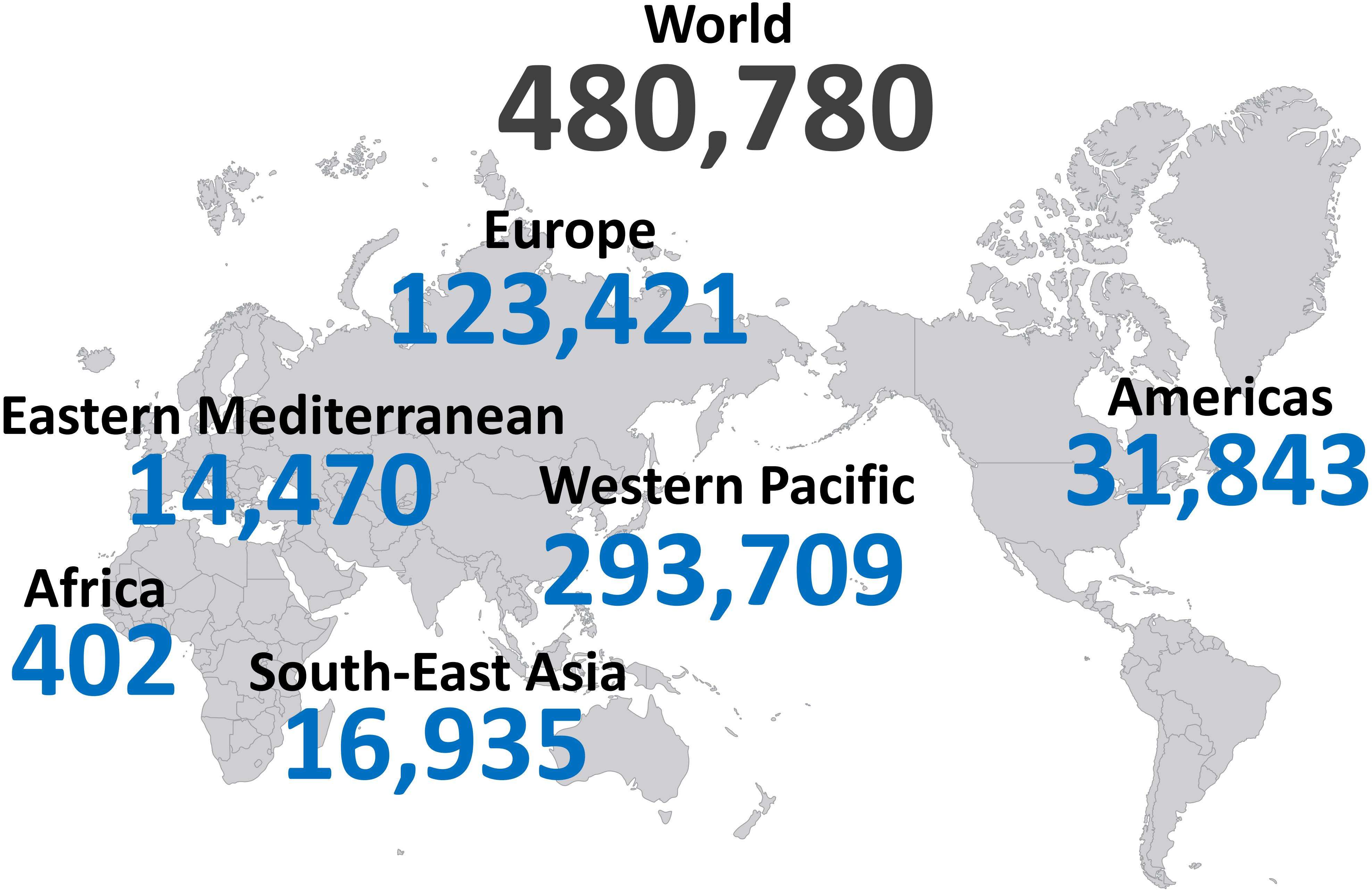
Sakurai Lab.
Data Mining Group @ Osaka University

# Outline

- **Motivation**
- Modeling power of EpiCast
- Problem definition
- Proposed model
- Algorithms
- Experiments
- Conclusions

# 480,780


All data is as of July 14, 2022.




All data is as of July 14, 2022.

# Motivation

**Given**: Epidemiological data stream $\mathcal{X}$

e.g., COVID-19 active cases/deaths/recovered

# Motivation

**Given**: Epidemiological data stream $\mathcal{X}$

e.g., COVID-19 active cases/deaths/recovered

# Motivation

**Given**: Epidemiological data stream $\mathcal{X}$

e.g., COVID-19 active cases/deaths/recovered

# Motivation

**Given**: Epidemiological data stream $\mathcal{X}$

e.g., COVID-19 active cases/deaths/recovered


How can we find **important patterns and rules**?

How can we forecast **future outbreaks and pandemics**?

# Motivation

**Given**: Epidemiological data stream $\mathcal{X}$

e.g., COVID-19 active cases/deaths/recovered


How can we find **important patterns and rules**?

How can we forecast **future outbreaks and pandemics**?

How can we monitor **the changing situations of epidemics**?

# Motivation

**Given**: Epidemiological data stream

e.g., COVID-19 active cases/deaths/recovered

**Challenges**:

- **Find** global and local-level representative patterns/regimes
- **Report** $l_s$- steps-ahead outbreaks, continuously, in a streaming fashion

# Motivation

**Given**: Epidemiological data stream

e.g., COVID-19 active cases/deaths/recovered

**Challenges**:

- **Find** global and local-level representative patterns/regimes
- **Report** $l_s$- steps-ahead outbreaks, continuously, in a streaming fashion

**Requirements**: handle non-stationary/spatially co-evolving time-series

- **Monitor** the changing situations of epidemics
- **Estimate/update** current dynamical patterns

# Motivation

**Given**: Epidemiological data stream

e.g., COVID-19 active cases/deaths/recovered

**Challenges**:

- **Find** global and local-level representative patterns/regimes
- **Report** $l_s$- steps-ahead outbreaks, continuously, in a streaming fashion

**Requirements**: handle non-stationary/spatially co-evolving time-series

- **Monitor** the changing situations of epidemics
- **Estimate/update** current dynamical patterns

**EpiCast: forecasting method for epidemiological data stream**

# Outline

~~Motivation~~

▶ **Modeling power of EpiCast**

Problem definition

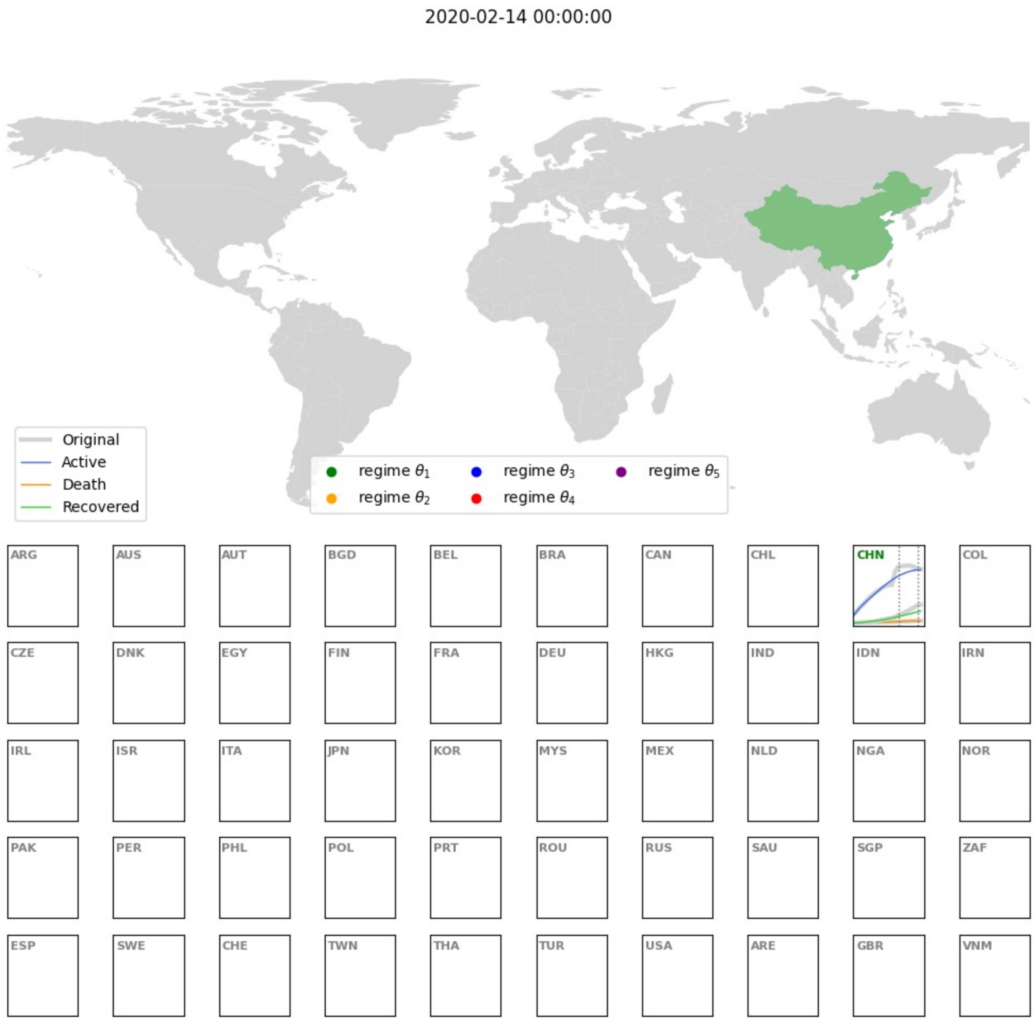
Proposed model

Algorithms

Experiments

Conclusions

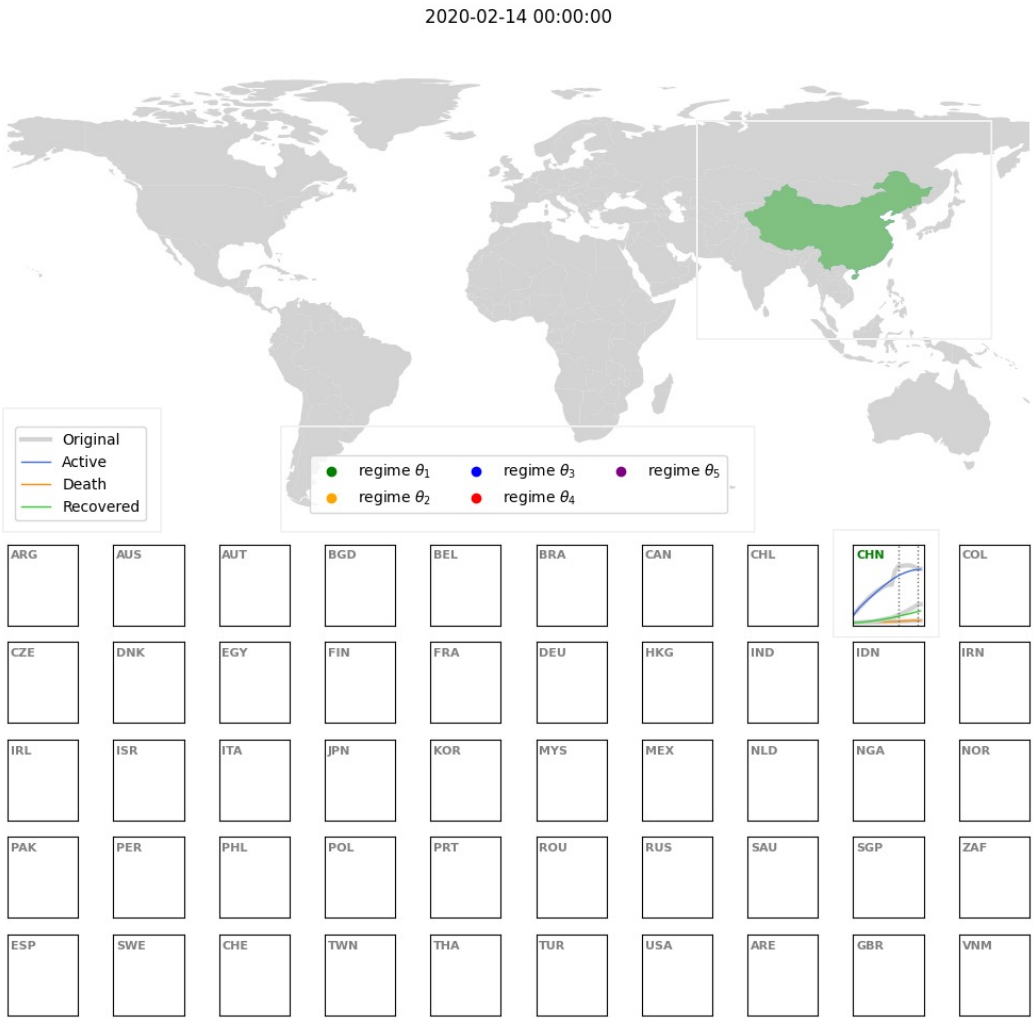
# Modeling power of EpiCast



# Modeling power of EpiCast

## Dataset

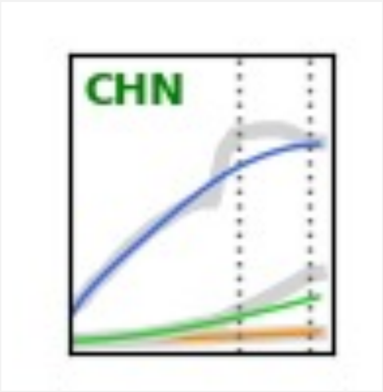
– 3-dimensional streams

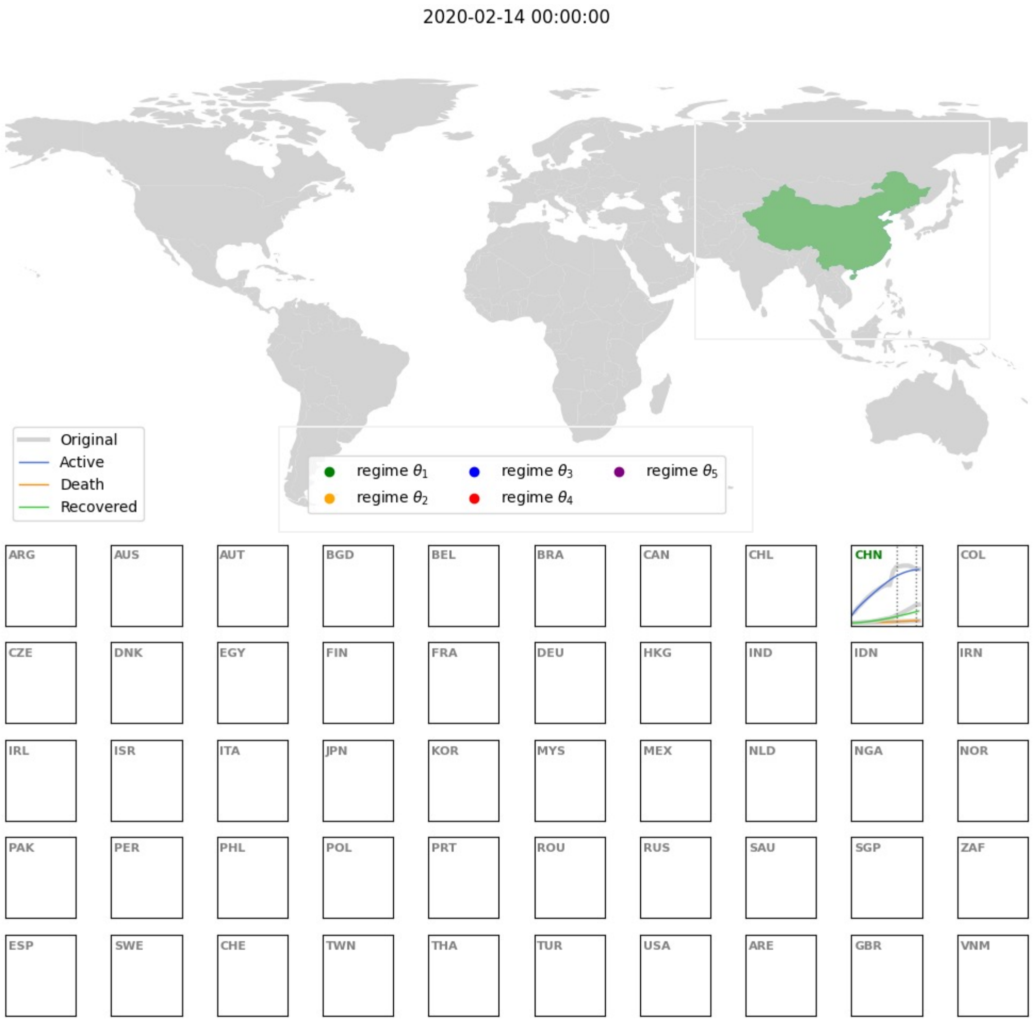


# Modeling power of EpiCast

## Dataset

– 3-dimensional streams

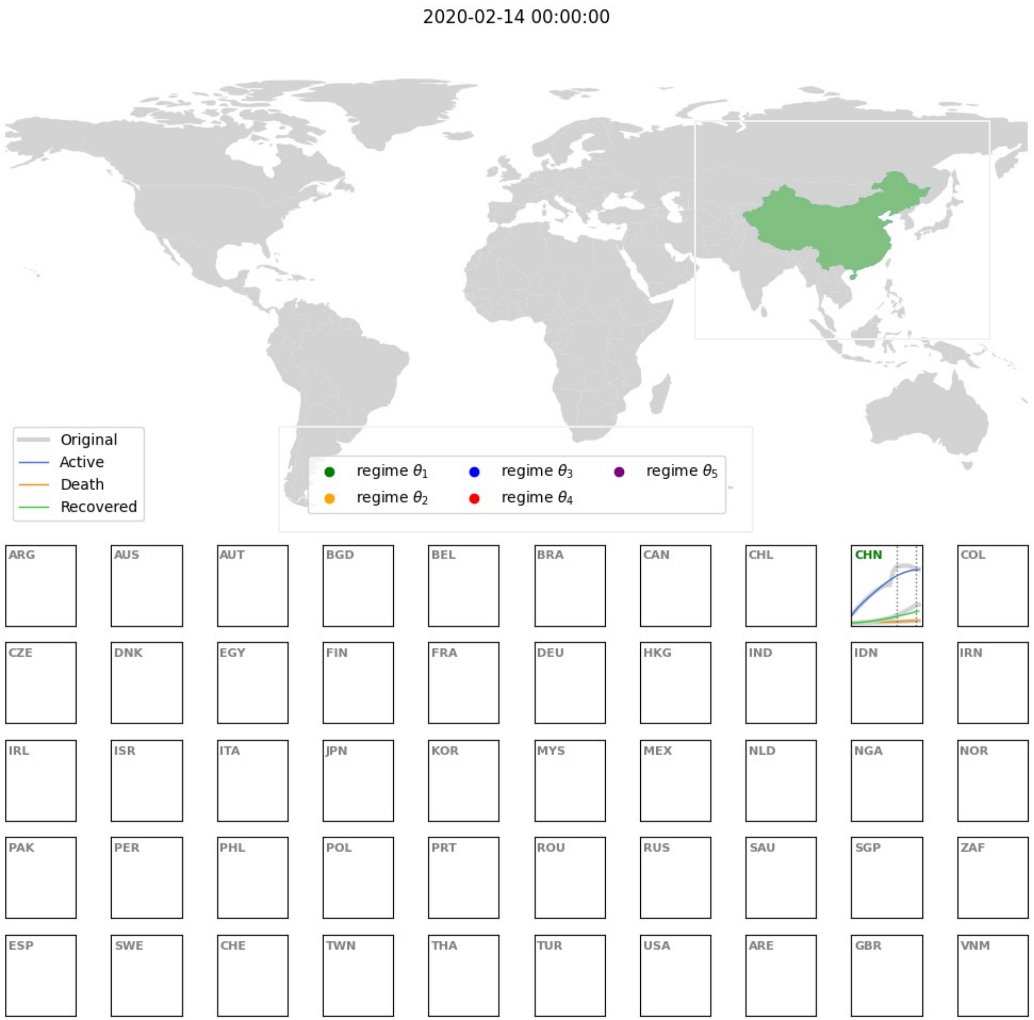



# Modeling power of EpiCast

## Dataset

– 3-dimensional streams






**Current window**

**7-days-ahead forecasted values**

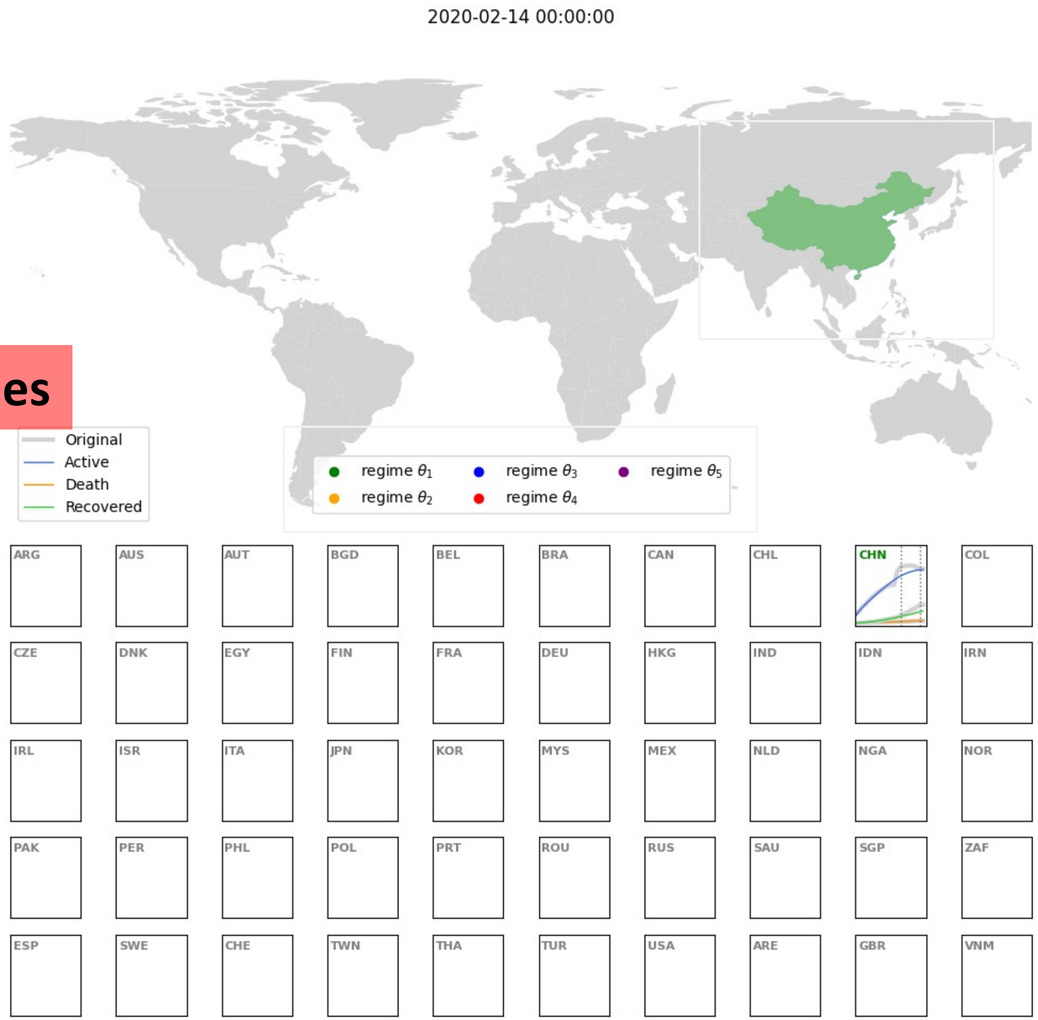
# Modeling power of EpiCast

## Dataset

– 3-dimensional streams


**Current window**

**7-days-ahead forecasted values**



# Modeling power of EpiCast

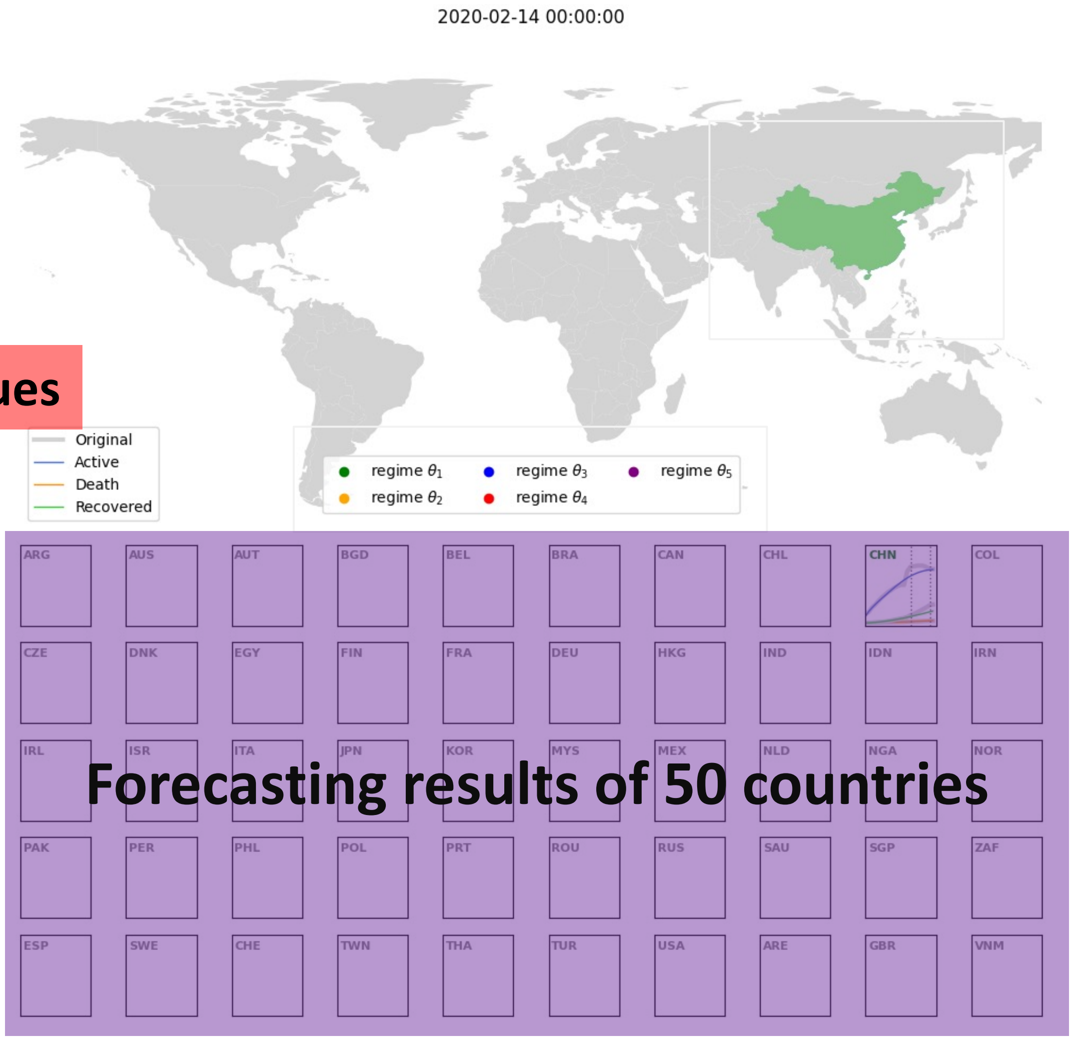
## Dataset

- 3-dimensional streams


Current window

7-days-ahead forecasted values

- 50 countries
- 600 days on a daily basis

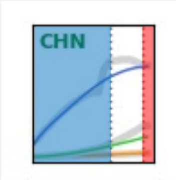



Forecasting results of 50 countries

# Modeling power of EpiCast

## Dataset

– 3-dimensional streams

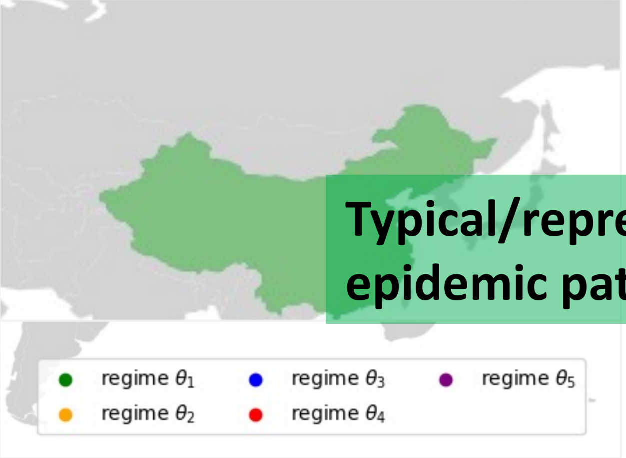




Current window

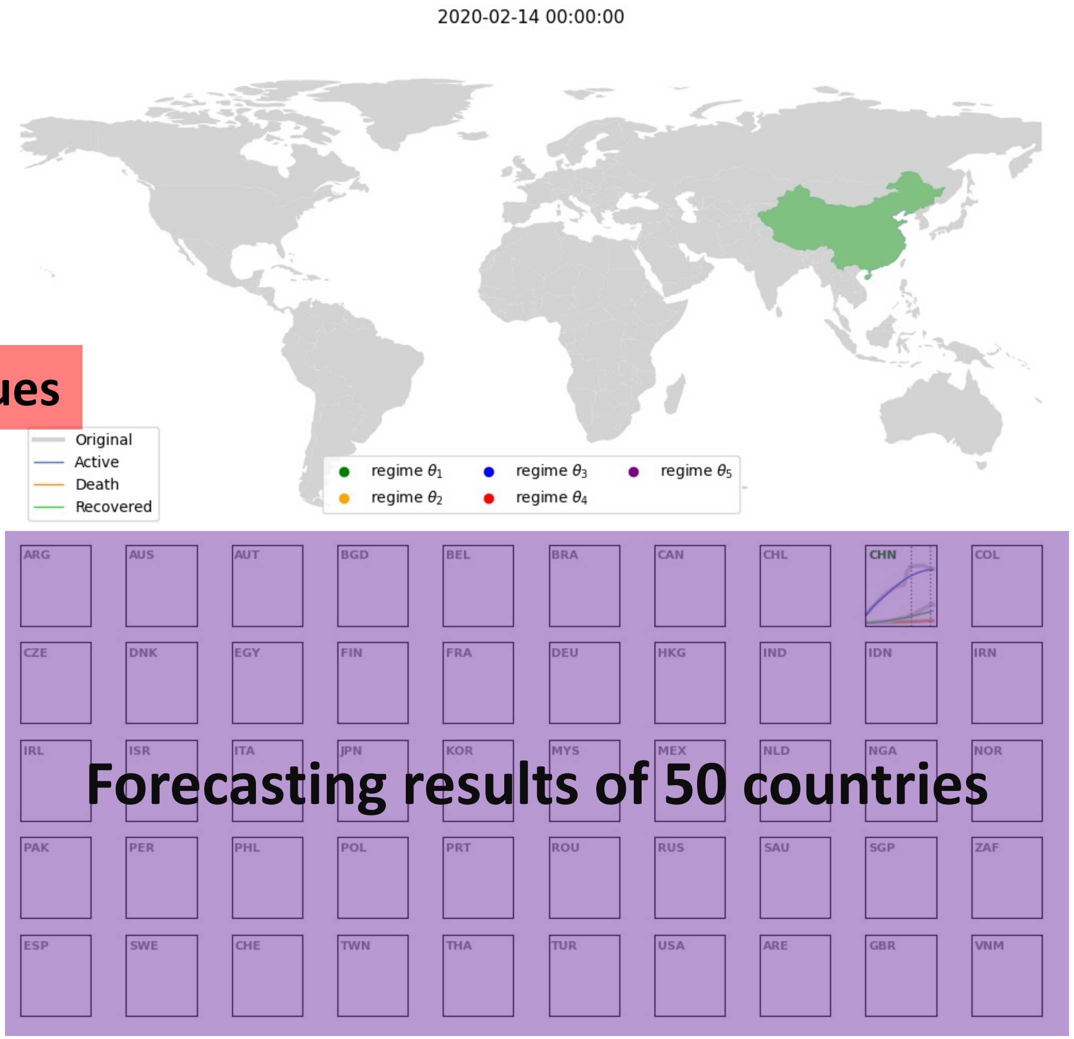
7-days-ahead forecasted values

– 50 countries

– 600 days on a daily basis




Typical/representative epidemic patterns (i.e., $\theta_1, \theta_2$)

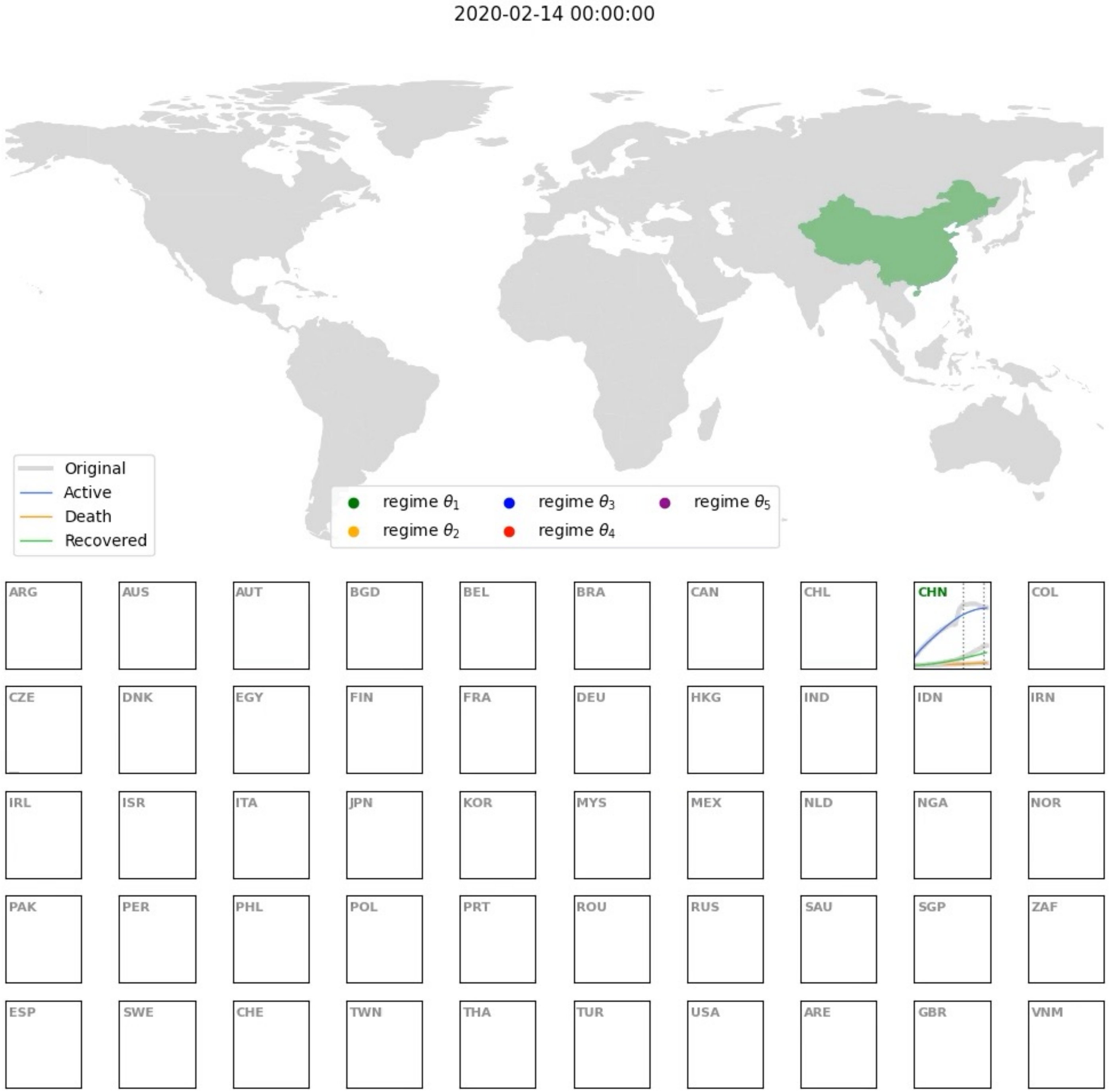



Forecasting results of 50 countries

# Modeling power of EpiCast

## Dataset

- 3-dimensional streams
- 50 countries
- 600 days on a daily basis



# Outline

~~Motivation~~

~~Modeling power of EpiCast~~

▶ **Problem definition**

Proposed model

Algorithms

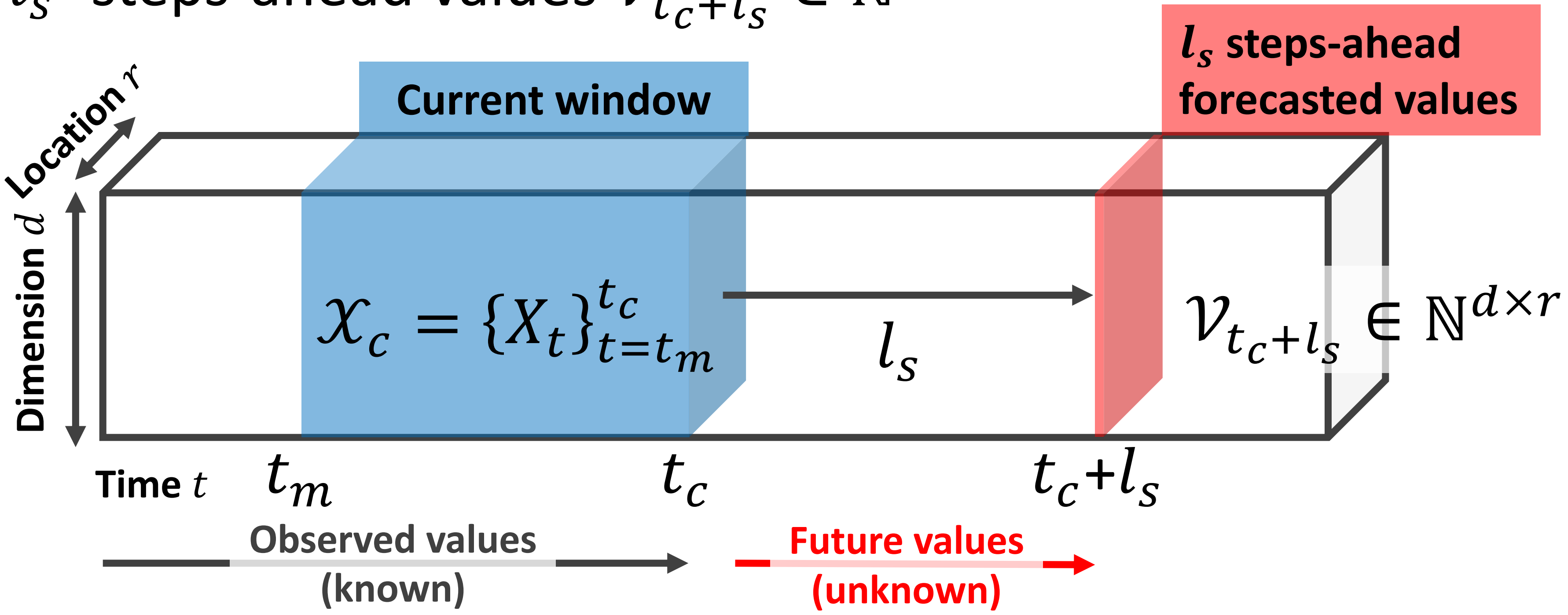
Experiments

Conclusions

# Problem definition

**Given**: Epidemiological data stream $\mathcal{X} \in \mathbb{N}^{d \times r \times t_c}$

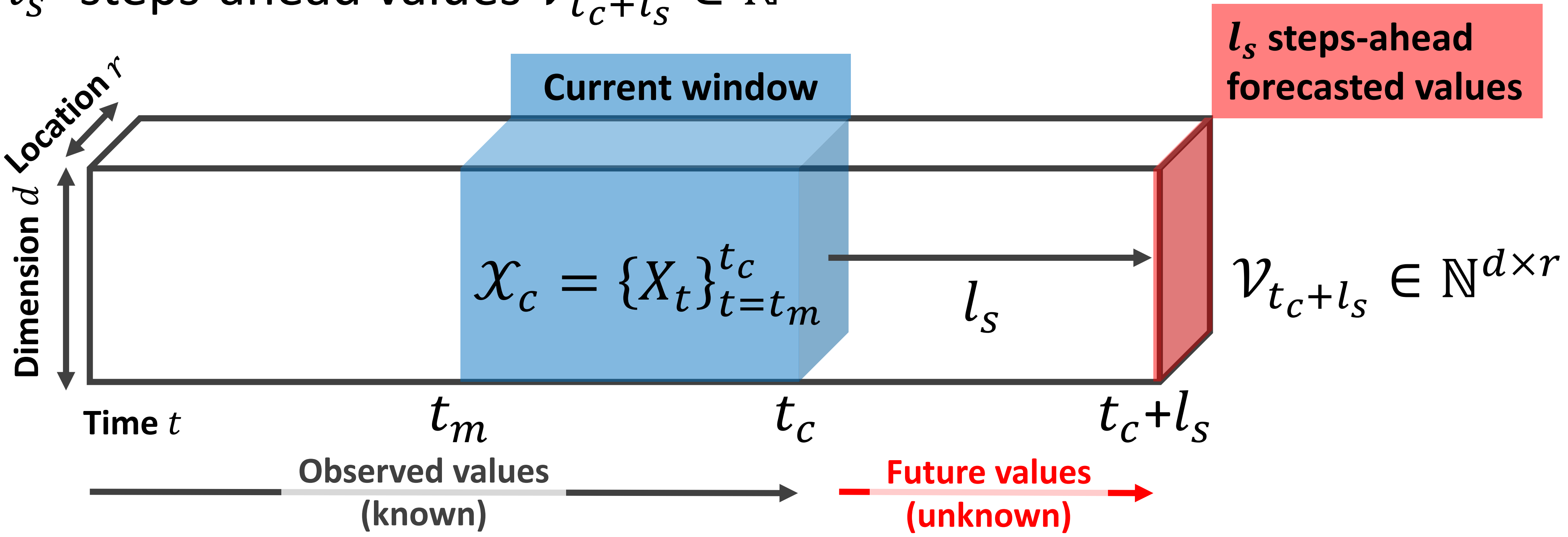
**Report**: $l_s$- steps-ahead values $\mathcal{V}_{t_c+l_s} \in \mathbb{N}^{d \times r}$



# Problem definition

**Given**: Epidemiological data stream $\mathcal{X} \in \mathbb{N}^{d \times r \times t_c}$

**Report**: $l_s$- steps-ahead values $\mathcal{V}_{t_c+l_s} \in \mathbb{N}^{d \times r}$



# Outline

~~Motivation~~

~~Modeling power of EpiCast~~

~~Problem definition~~

▶ **Proposed model**

Algorithms

Experiments

Conclusions

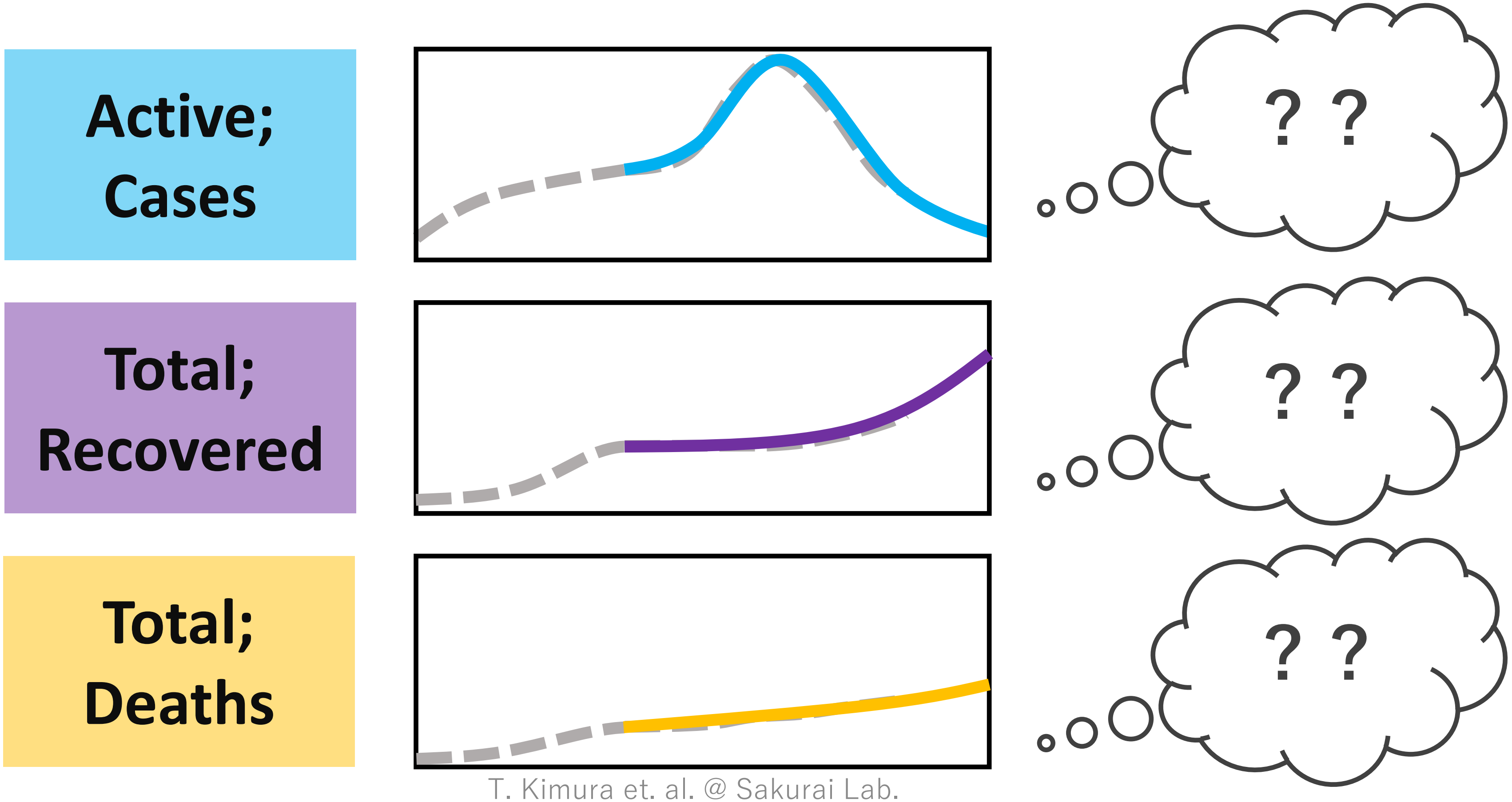
# Proposed model

**Q1**. How can we model complex, non-linear dynamics in epidemic streams?

**Q2**. How can we efficiently and effectively estimate
such non-linear models in multiple locations?

# Proposed model

**Q1**. How can we model complex, non-linear dynamics in epidemic streams?

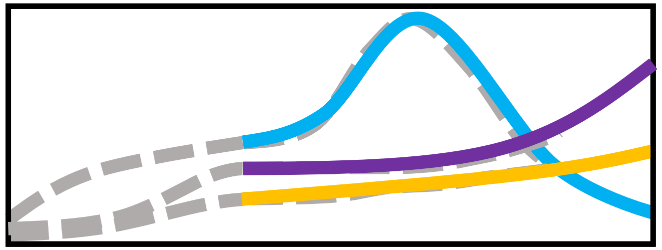


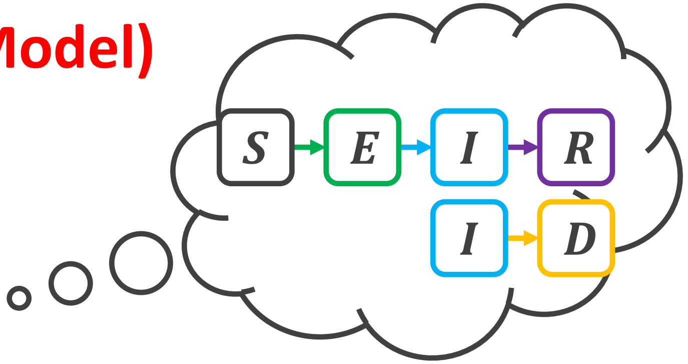
# Proposed model

**Q1**. How can we model complex, non-linear dynamics in epidemic streams?

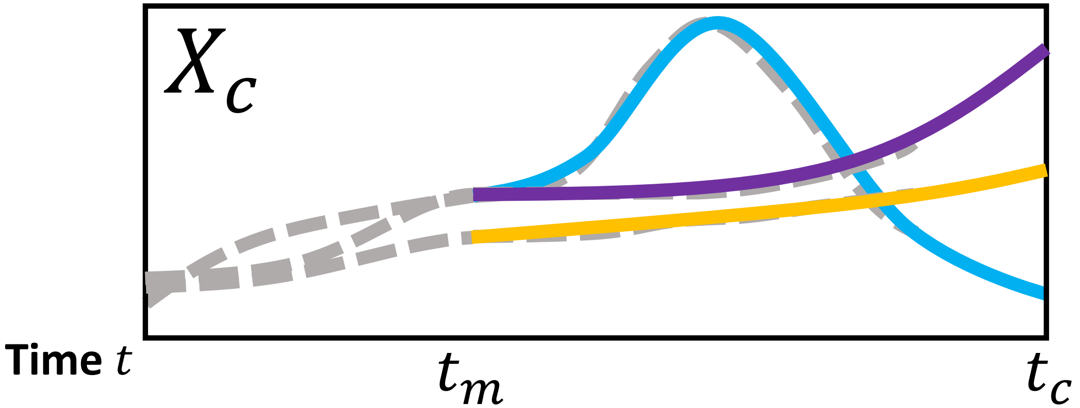
- **P1**. **Gray-box model – domain knowledge SEIRD (Model)**

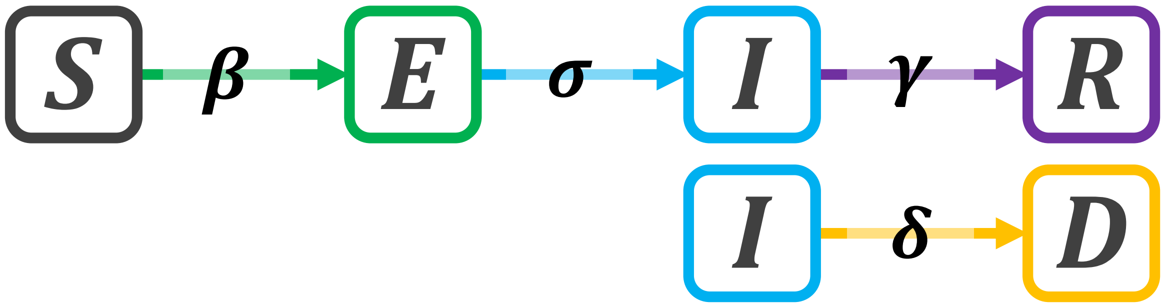
**A simple model of epidemics**





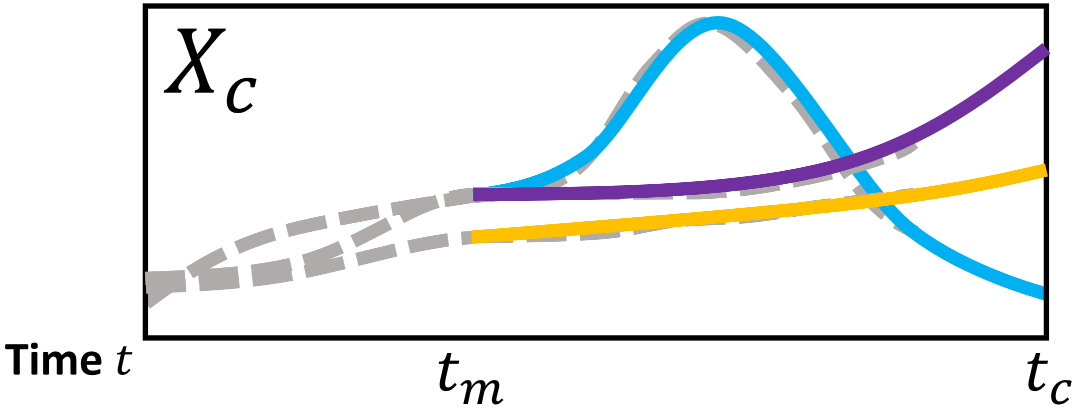
# P1. Grey-box model – domain knowledge SEIRD (Model)

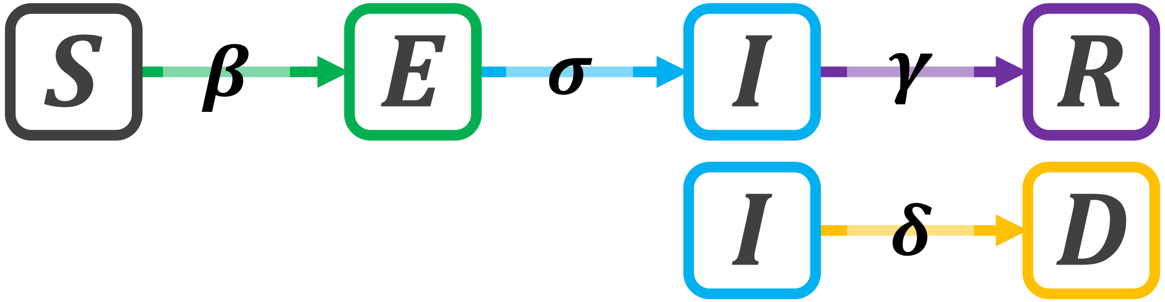



**Regime:** $\theta = \{N, \beta, \sigma, \gamma, \delta, E_0, I_0, R_0, D_0\}$ based on SEIRD model

# P1. Grey-box model – domain knowledge SEIRD (Model)



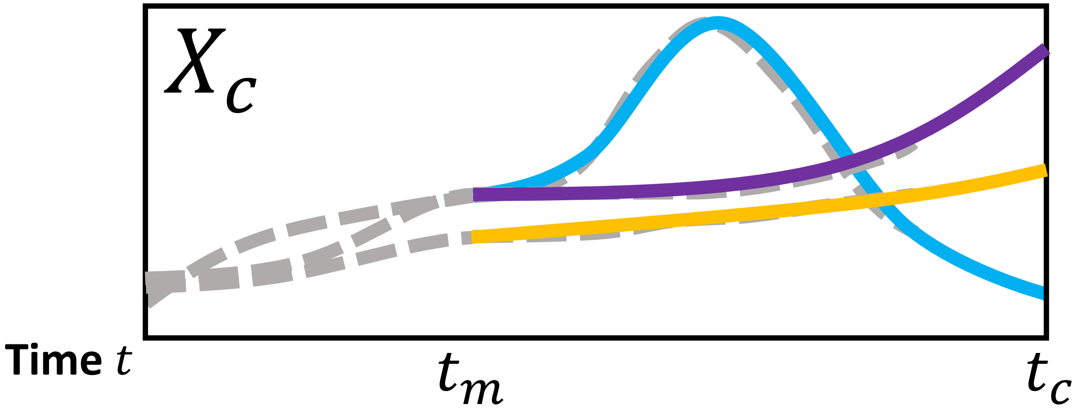



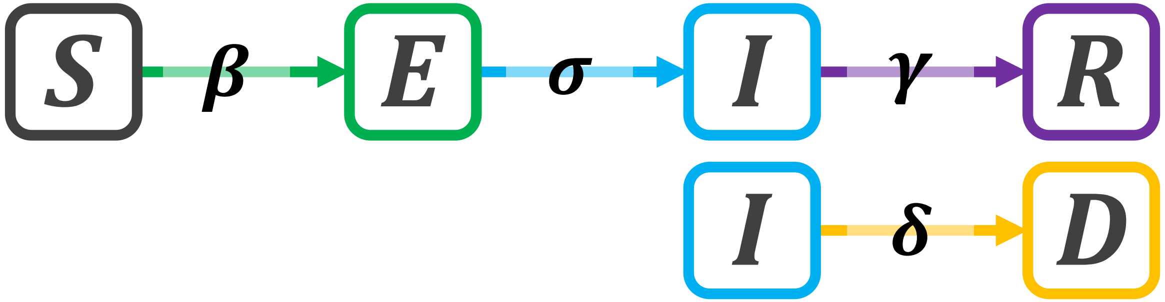

**Regime:** $\theta = \{N, \beta, \sigma, \gamma, \delta, E_0, I_0, R_0, D_0\}$ based on SEIRD model

$$\underbrace{\frac{dS}{dt} = -\beta S(t)I(t),}_{\text{Susceptible}} \quad \underbrace{\frac{dE}{dt} = \beta S(t)I(t) - \sigma E(t),}_{\text{Exposed}}$$

$$\underbrace{\frac{dI}{dt} = \sigma E(t) - \gamma I(t) - \delta I(t),}_{\text{Infected}} \quad \underbrace{\frac{dR}{dt} = \gamma I(t),}_{\text{Recovered}} \quad \underbrace{\frac{dD}{dt} = \delta I(t)}_{\text{Dead}}$$

# P1. Grey-box model
# – domain knowledge SEIRD (Model)




**Regime:** $\theta = \{N, \beta, \sigma, \gamma, \delta, E_0, I_0, R_0, D_0\}$ based on SEIRD model

$$\frac{dS}{dt} = -\beta S(t)I(t), \quad \frac{dE}{dt} = \beta S(t)I(t) - \sigma E(t),$$

Susceptible

Exposed

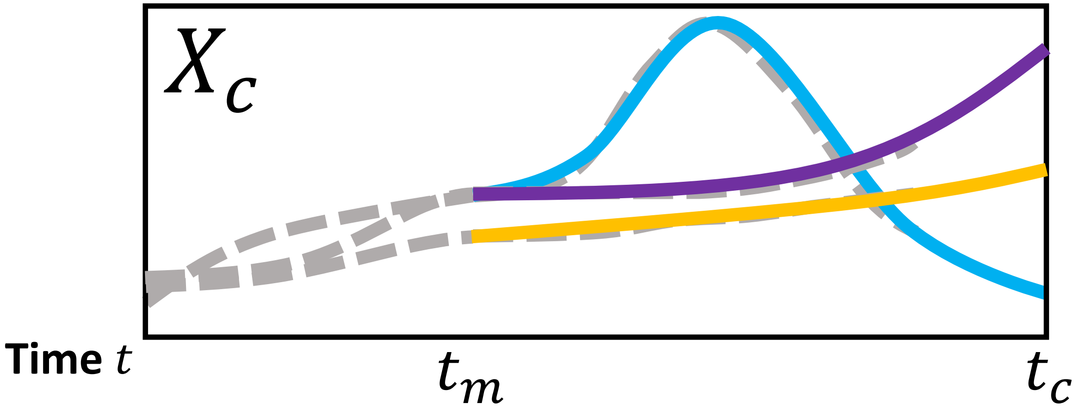
$$\frac{dI}{dt} = \sigma E(t) - \gamma I(t) - \delta I(t), \quad \frac{dR}{dt} = \gamma I(t), \quad \frac{dD}{dt} = \delta I(t)$$

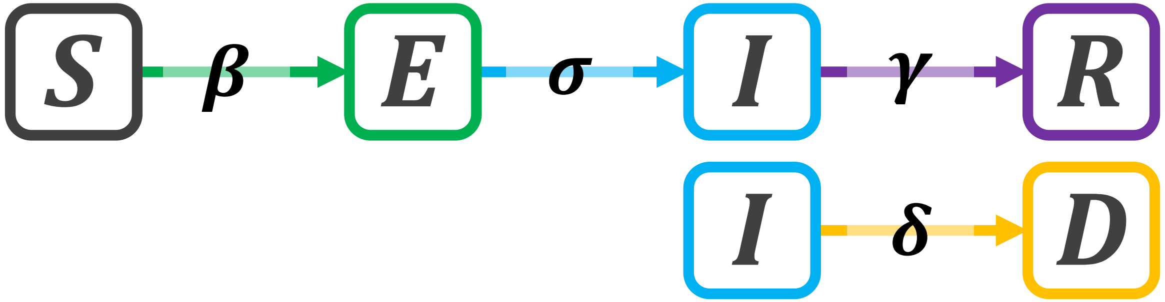
Infected

Recovered

Dead

# P1. Grey-box model – domain knowledge SEIRD (Model)

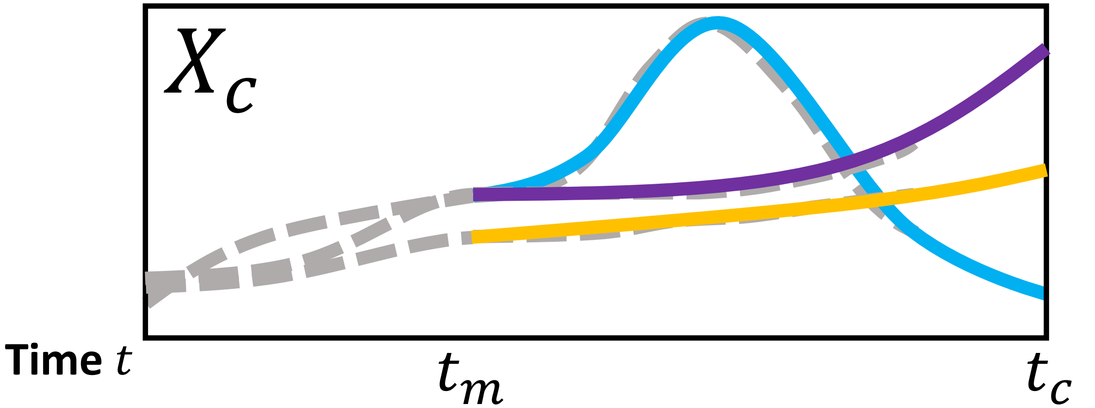


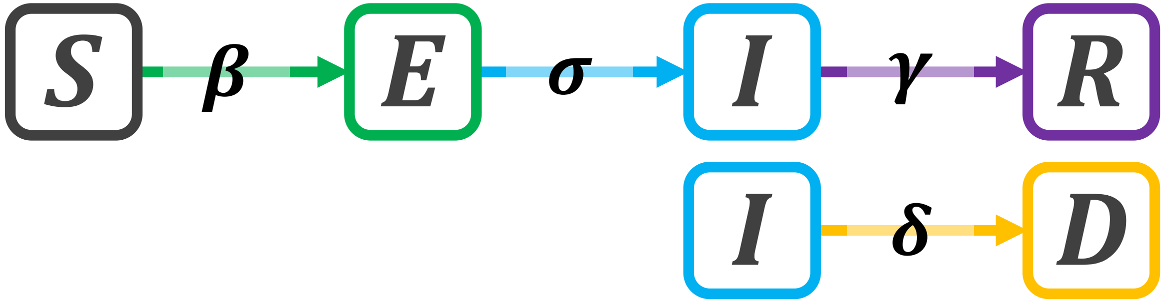




**Regime:** $\theta = \{N, \beta, \sigma, \gamma, \delta, E_0, I_0, R_0, D_0\}$ based on SEIRD model

$$\underbrace{\frac{dS}{dt} = -\beta S(t)I(t),}_{\text{Susceptible}} \quad \underbrace{\frac{dE}{dt} = \beta S(t)I(t) - \sigma E(t),}_{\text{Exposed}}$$

# P1. Grey-box model
# – domain knowledge SEIRD (Model)



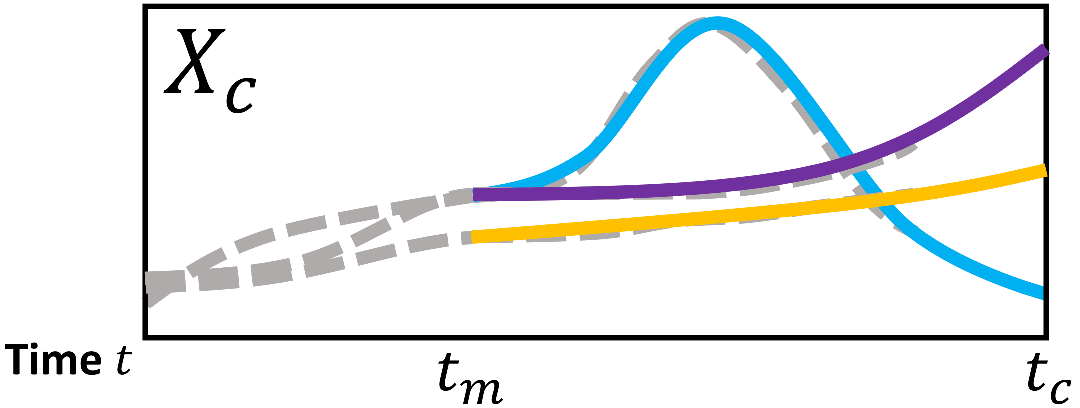



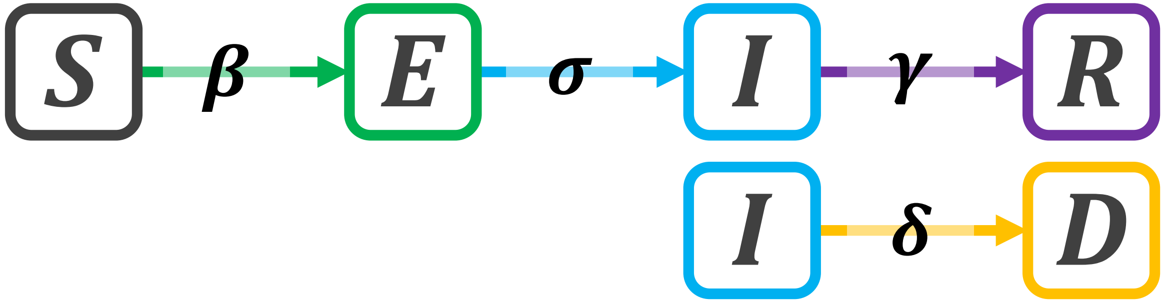

**Regime:** $\theta = \{N, \beta, \sigma, \gamma, \delta, E_0, I_0, R_0, D_0\}$ based on SEIRD model

$$\underbrace{\frac{dS}{dt} = -\beta S(t)I(t),}_{\text{Susceptible}} \quad \underbrace{\frac{dE}{dt} = \beta S(t)I(t) - \sigma E(t),}_{\text{Exposed}}$$

$$\underbrace{\frac{dI}{dt} = \sigma E(t) - \gamma I(t) - \delta I(t),}_{\text{Infected}}$$

# P1. Grey-box model – domain knowledge SEIRD (Model)




**Regime:** $\theta = \{N, \beta, \sigma, \gamma, \delta, E_0, I_0, R_0, D_0\}$ based on SEIRD model

$$\frac{dS}{dt} = -\beta S(t)I(t), \quad \frac{dE}{dt} = \beta S(t)I(t) - \sigma E(t),$$

Susceptible — Exposed

$$\frac{dI}{dt} = \sigma E(t) - \gamma I(t) - \delta I(t), \quad \frac{dR}{dt} = \gamma I(t), \quad \frac{dD}{dt} = \delta I(t)$$

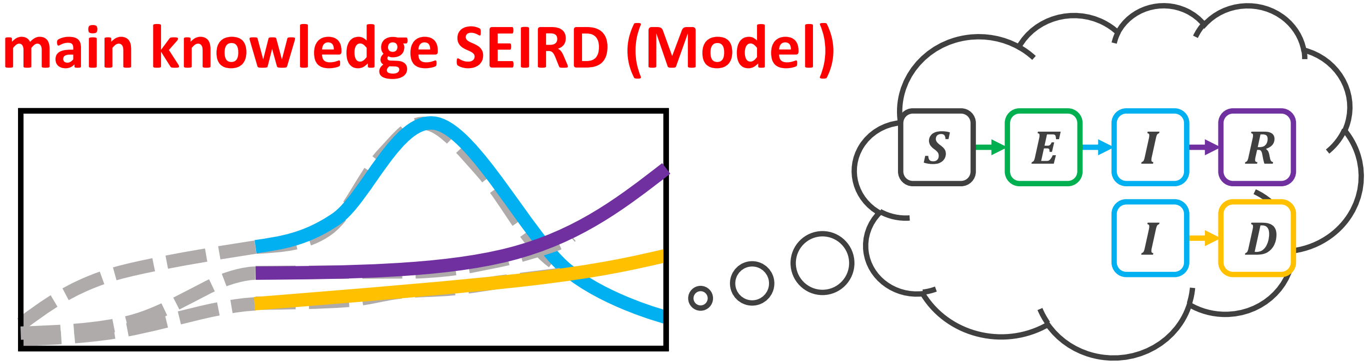
Infected — Recovered — Dead

# Proposed model

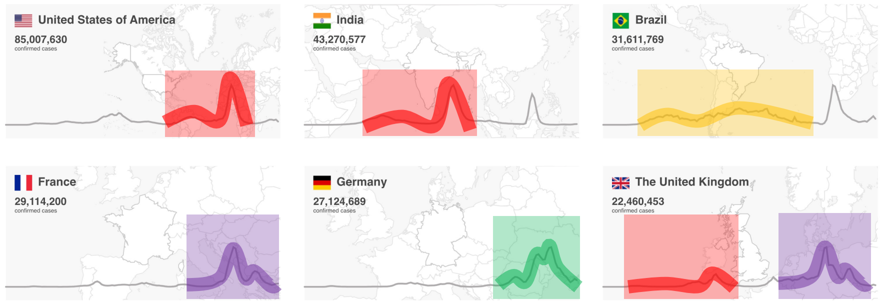
**Q1**. How can we model complex, non-linear dynamics in epidemic streams?

– **P1**. **Gray-box model – domain knowledge SEIRD (Model)**

**A simple model of epidemics**




**Q2**. How can we efficiently and effectively estimate such non-linear models in multiple locations?

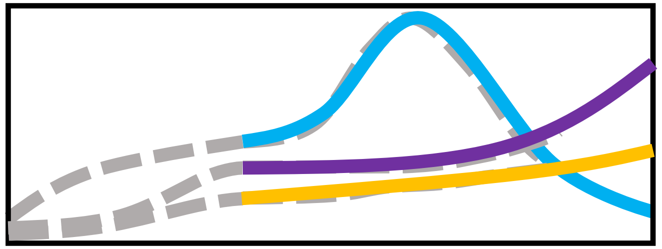


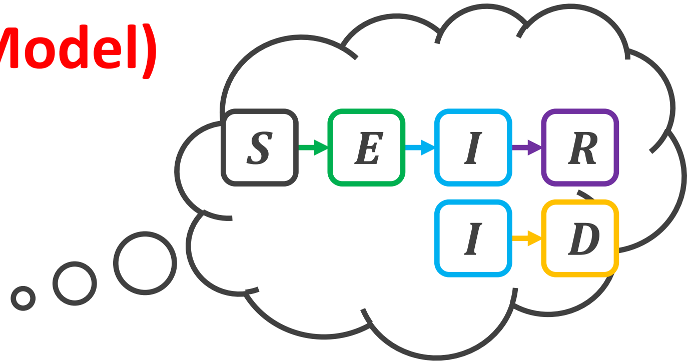
# Proposed model

**Q1**. How can we model complex, non-linear dynamics in epidemic streams?

– **P1**. **Gray-box model – domain knowledge SEIRD (Model)**

**A simple model of epidemics**

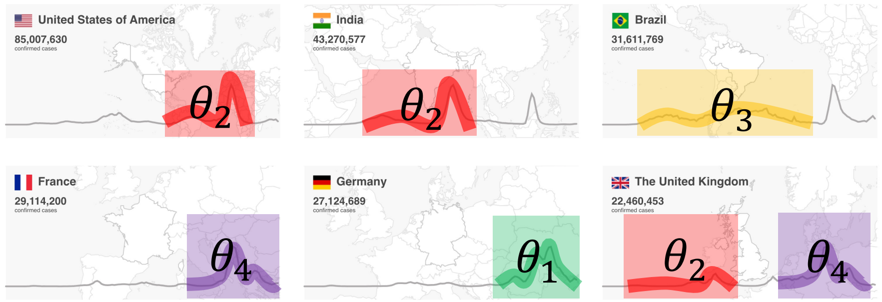




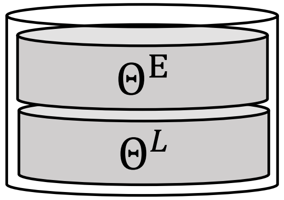

**Q2**. How can we efficiently and effectively estimate such non-linear models in multiple locations?

– **P2**. **Share global parameters and adjust local parameters**

**EpiCast parameters**




**EpiModel DB $\Theta$**

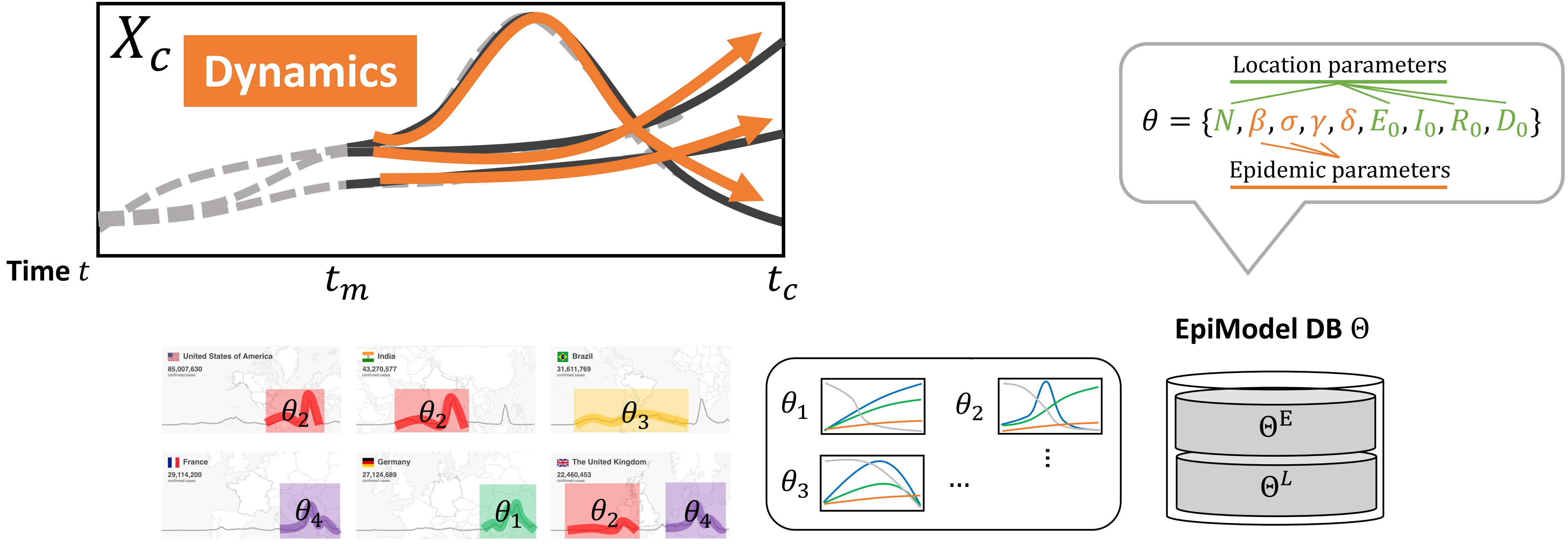



Location parameters

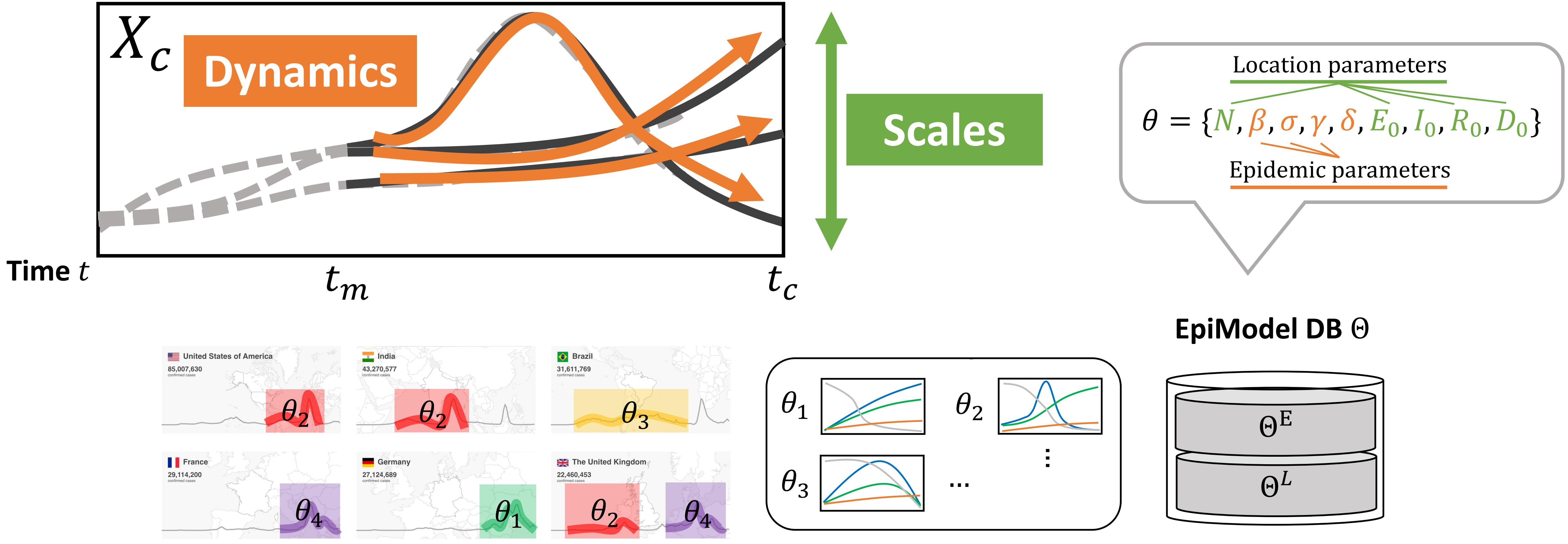
$$\theta = \{N, \beta, \sigma, \gamma, \delta, E_0, I_0, R_0, D_0\}$$

Epidemic parameters

# P2. Share global parameters and adjust local parameters



# P2. Share global parameters and adjust local parameters



# Outline

# Streaming algorithm: EpiCast

**1. Given:**

Current window: $\mathcal{X}[t_m : t_c]$
EpiModel DB: $\{\Theta^E, \Theta^L\}$

**2-1. Find: (EpiFinder)**

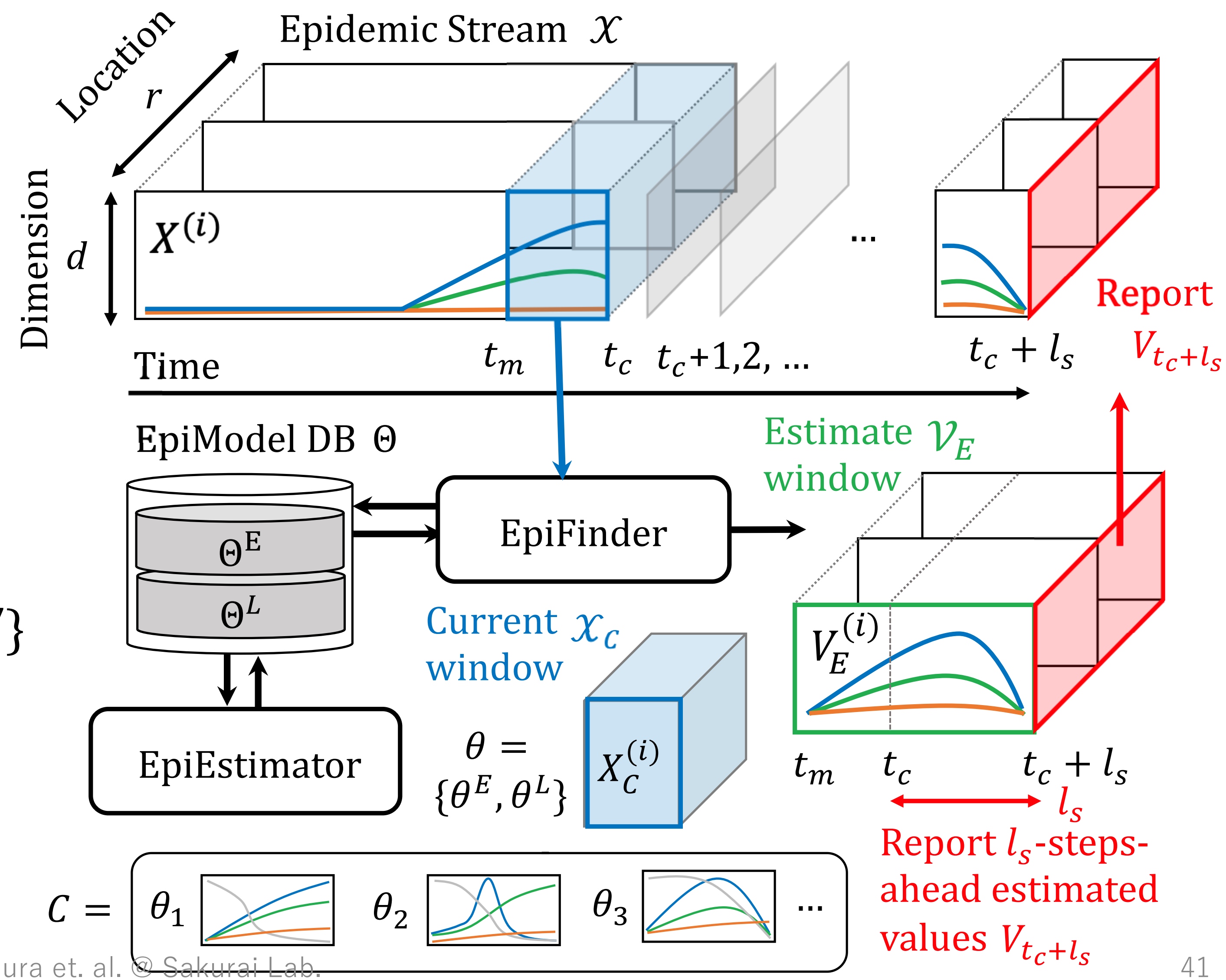
Model parameter set: $\{\theta^E, \theta^L\}$

**2-2. Estimate: (EpiEstimator)**

New model parameter set: $\{\theta^{E\prime}, \theta^{L\prime}\}$

**3. Report:**

$l_s$- steps-ahead values: $\mathcal{V}[t_m + l_s]$



# Streaming algorithm: EpiCast

**1. Given:**

Current window: $\mathcal{X}[t_m : t_c]$
EpiModel DB: $\{\Theta^E, \Theta^L\}$

**2-1. Find: (EpiFinder)**

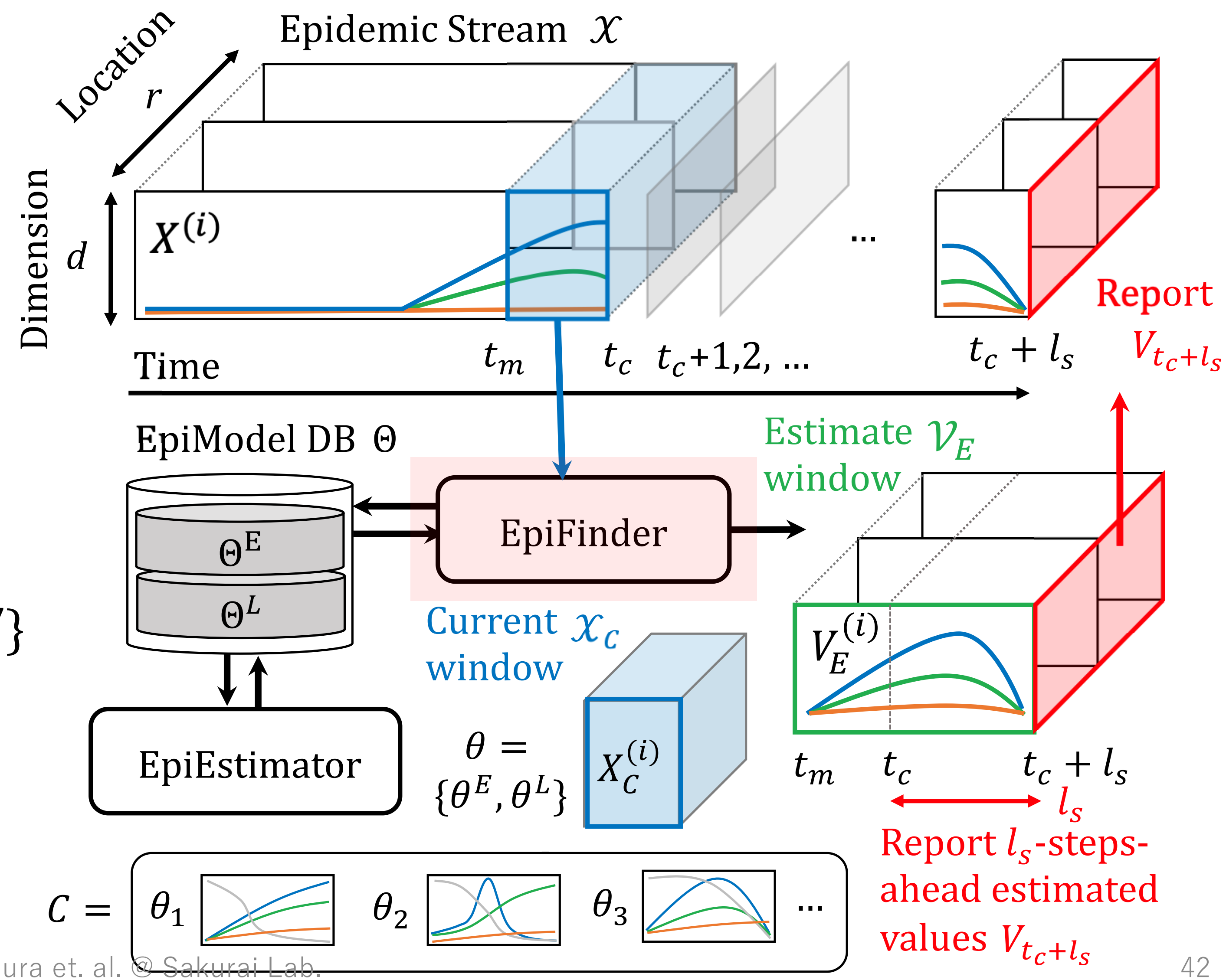
Model parameter set: $\{\theta^E, \theta^L\}$

**2-2. Estimate: (EpiEstimator)**

New model parameter set: $\{\theta^{E\prime}, \theta^{L\prime}\}$

**3. Report:**

$l_s$- steps-ahead values:$\mathcal{V}[t_m + l_s]$



# Streaming algorithm: EpiCast

**1. Given:**

Current window: $\mathcal{X}[t_m : t_c]$
EpiModel DB: $\{\Theta^E, \Theta^L\}$

**2-1. Find: (EpiFinder)**

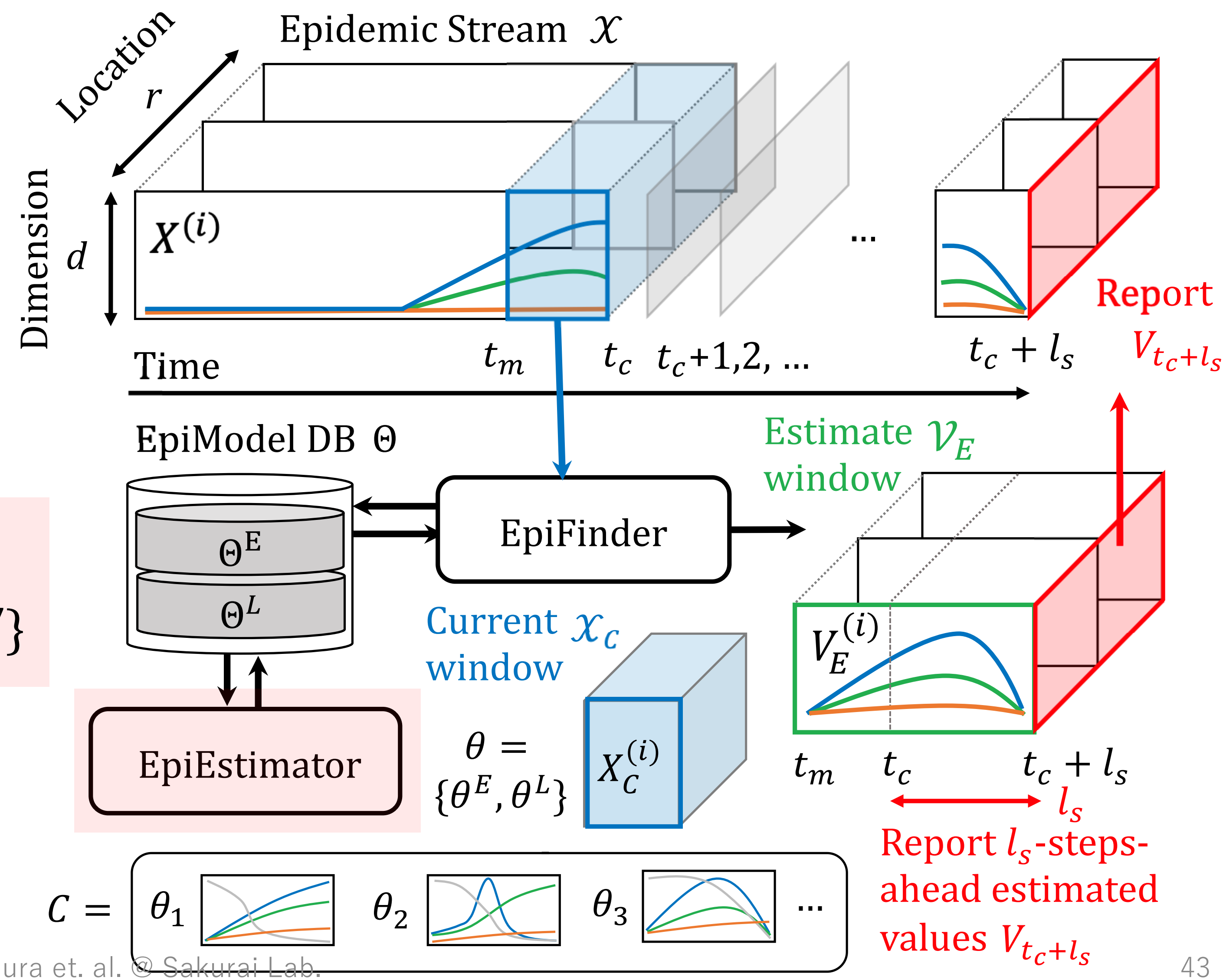
Model parameter set: $\{\theta^E, \theta^L\}$

**2-2. Estimate: (EpiEstimator)**

New model parameter set: $\{\theta^{E\prime}, \theta^{L\prime}\}$

**3. Report:**

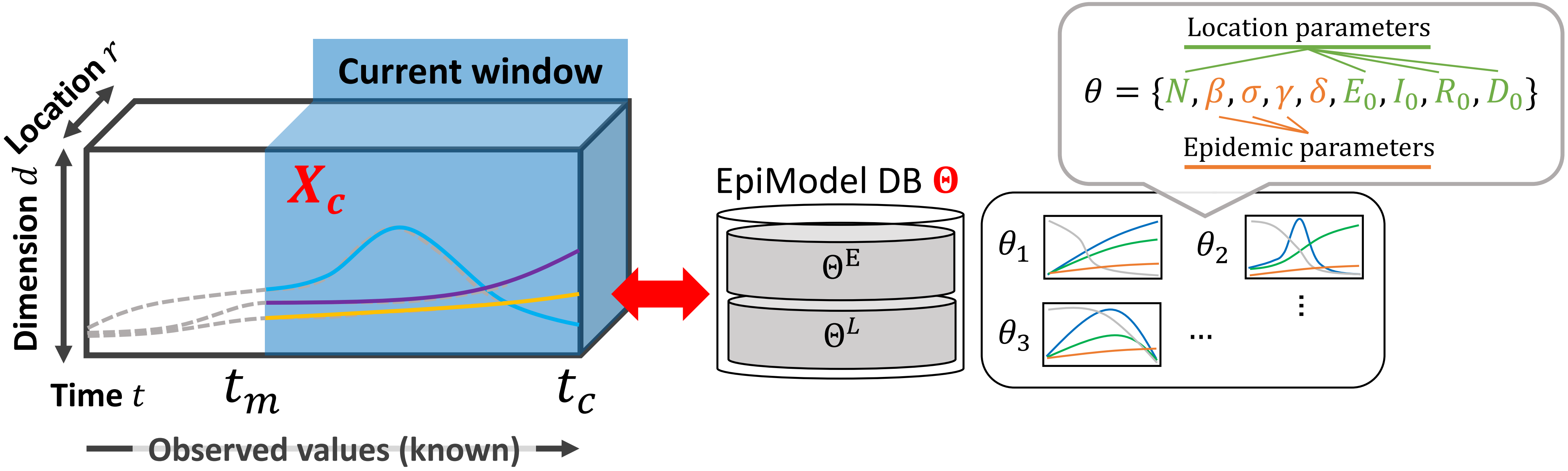
$l_s$- steps-ahead values:$\mathcal{V}[t_m + l_s]$



# Streaming algorithm: EpiCast

## EpiFinder

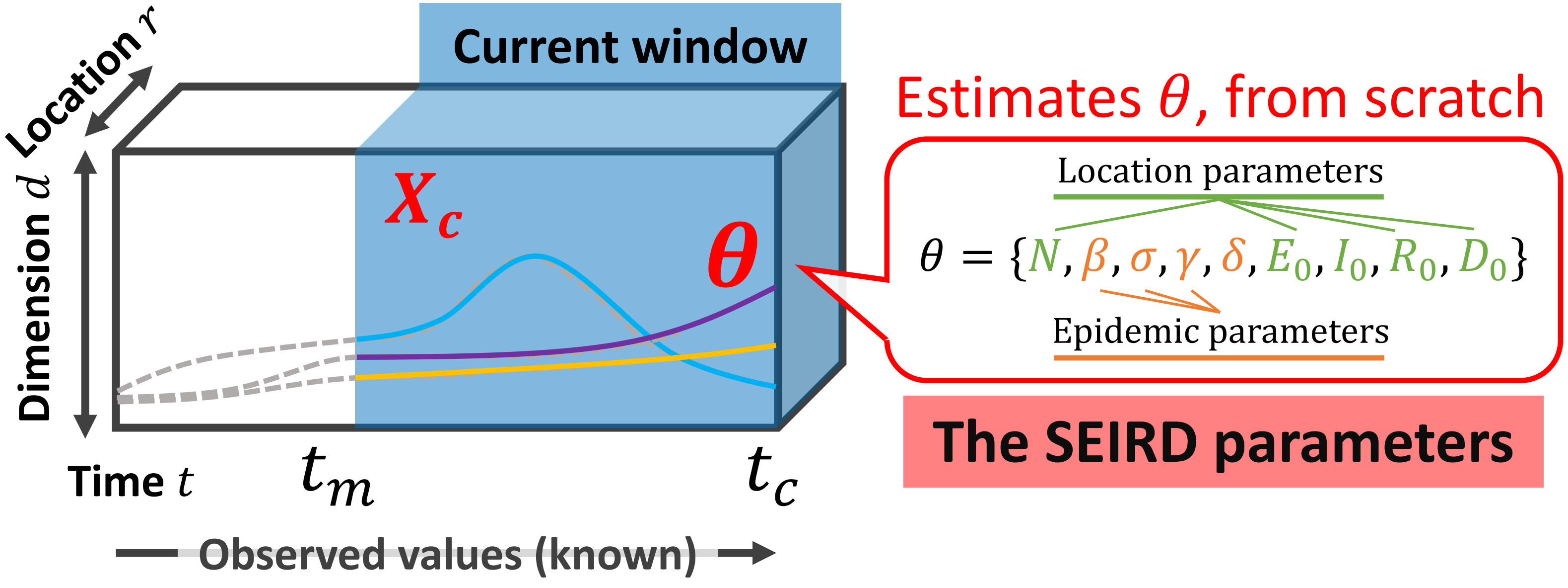
**Searches** for the best parameters $\boldsymbol{\theta} = \{\theta^E, \theta^L\}$ in the current parameter set $\boldsymbol{\Theta} = \{\Theta^E, \Theta^L\}$



# Streaming algorithm: EpiCast

## EpiEstimator

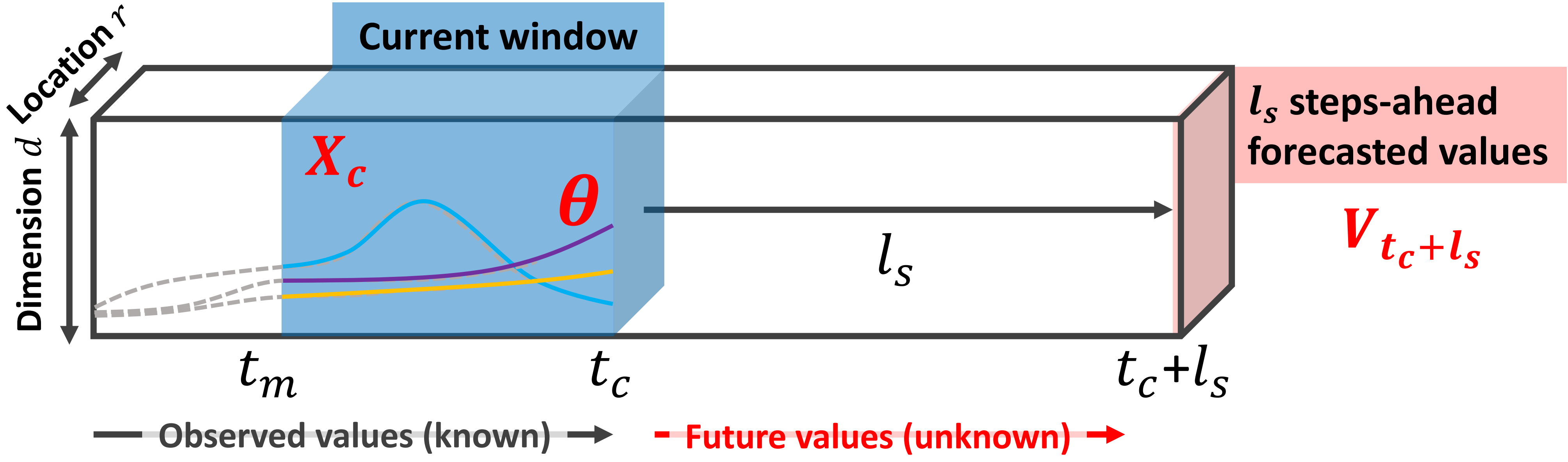
**Estimates** a new model parameters $\boldsymbol{\theta} = \{\theta^E, \theta^L\}$ from $\boldsymbol{X_C} = X[t_m : t_c]$



# Streaming algorithm: EpiCast

## EpiCast

Generates $l_s$- steps-ahead values $\boldsymbol{V}_{\boldsymbol{t_c+l_s}} = V[t_c + l_s]$ from an optimal model parameter set $\boldsymbol{\theta} = \{\theta^E, \theta^L\}$



# Outline

~~Motivation~~

~~Modeling power of EpiCast~~

~~Problem definition~~

~~Proposed model~~

~~Algorithms~~

▶ **Experiments**

Conclusions

# Experiments

## Q1. Effectiveness

How well does it capture co-evolving epidemic patterns?

## Q2. Accuracy

How accurately does it forecast future outbreaks?

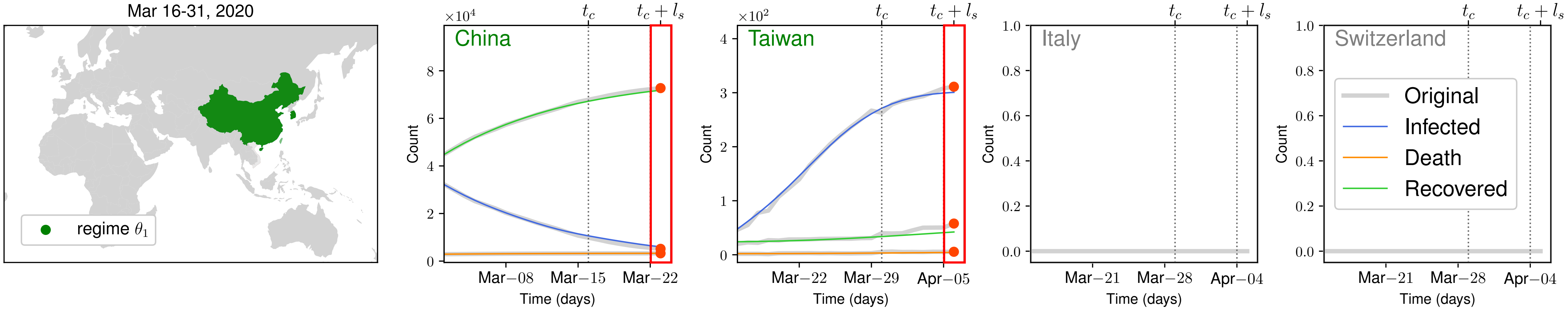
## Q3. Scalability

How does it scale in terms of computational time?

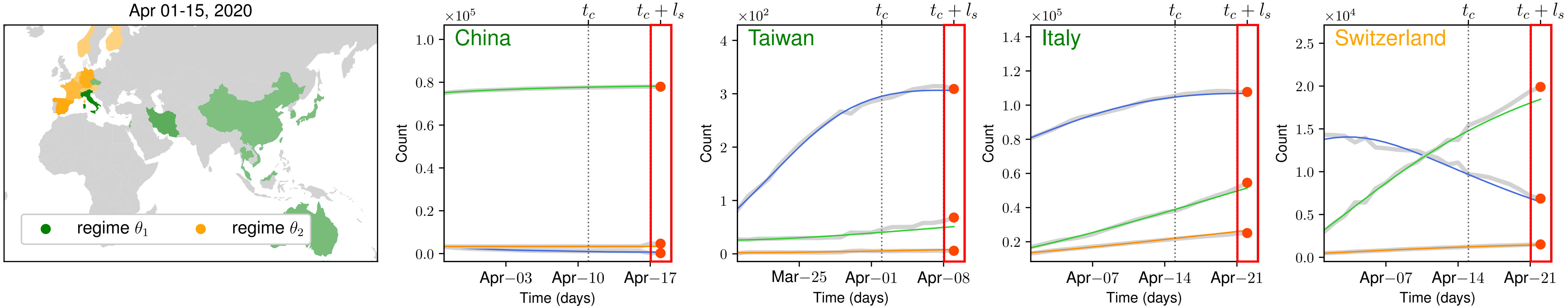
Competitors:

- SIRD
- ARIMA
- GRU
- TCN (CVPR‘17)
- EpiDeep (KDD‘19)

# Q1. Effectiveness

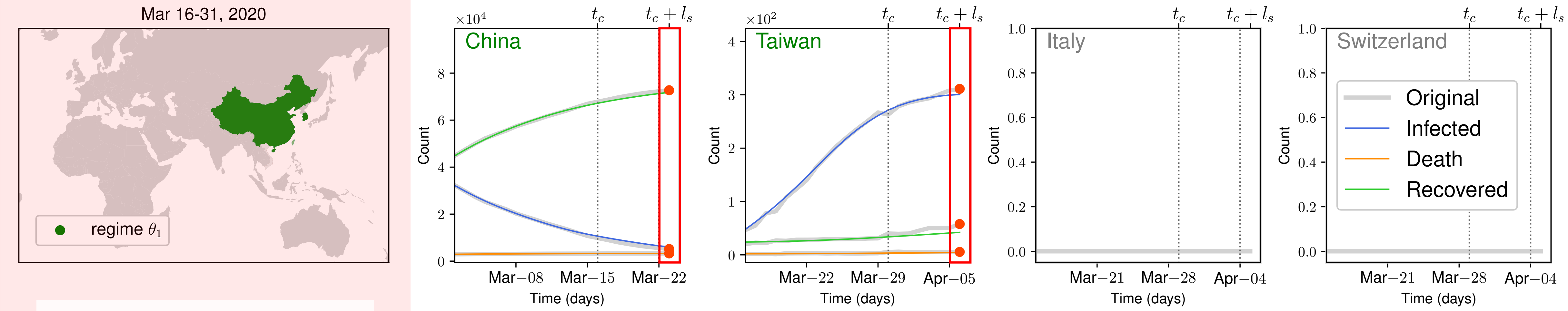



(a) Snapshots of EpiCast for the initial spread of COVID-19 in Asia



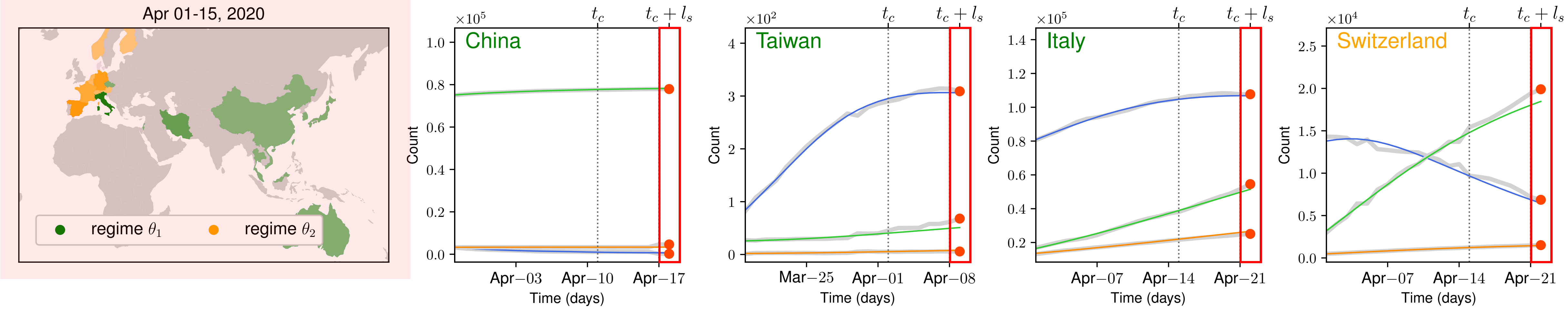

(b) Snapshots of EpiCast for the spreading process of COVID-19 from Asia to Europe

# Q1. Effectiveness



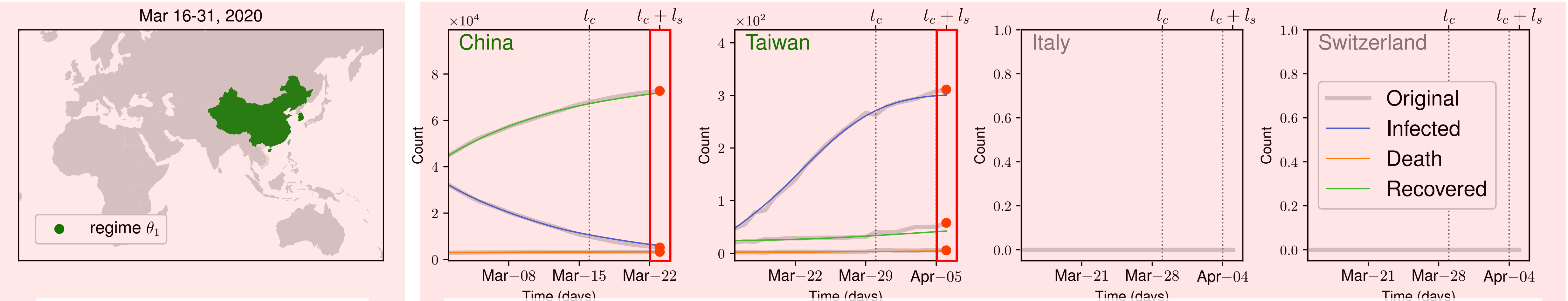

**EpiCast-Map**

(a) Snapshots of EpiCast for the initial spread of COVID-19 in Asia




(b) Snapshots of EpiCast for the spreading process of COVID-19 from Asia to Europe

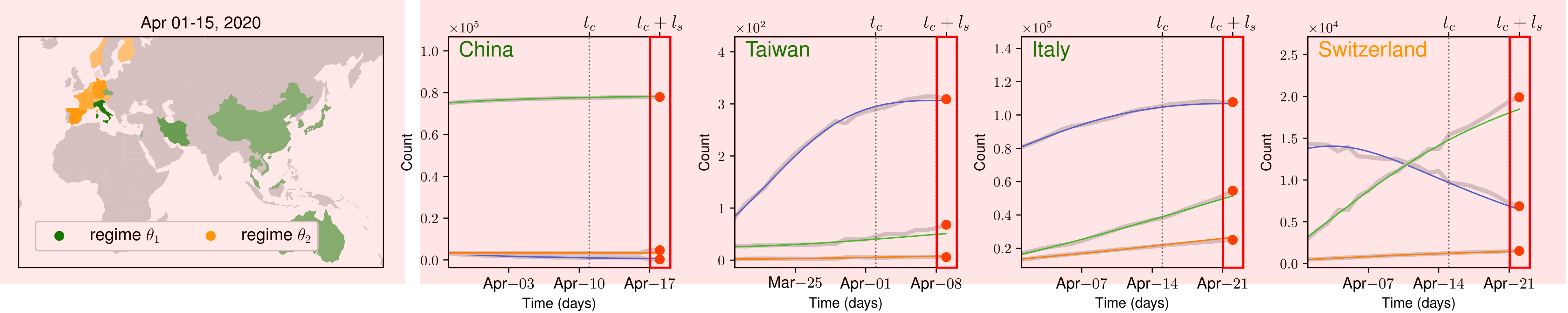
# Q1. Effectiveness




EpiCast-Map

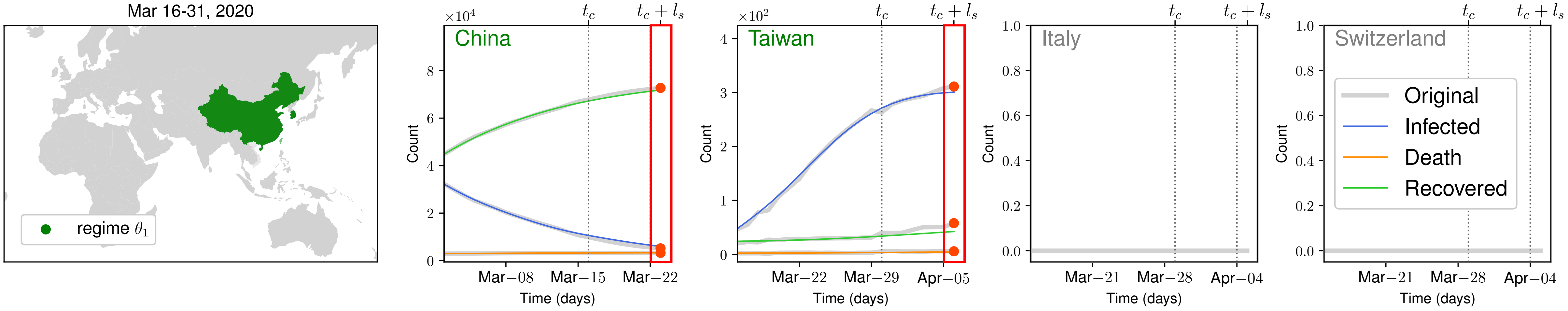
(a) Snapsho

Forecasting results of 7-days-ahead values

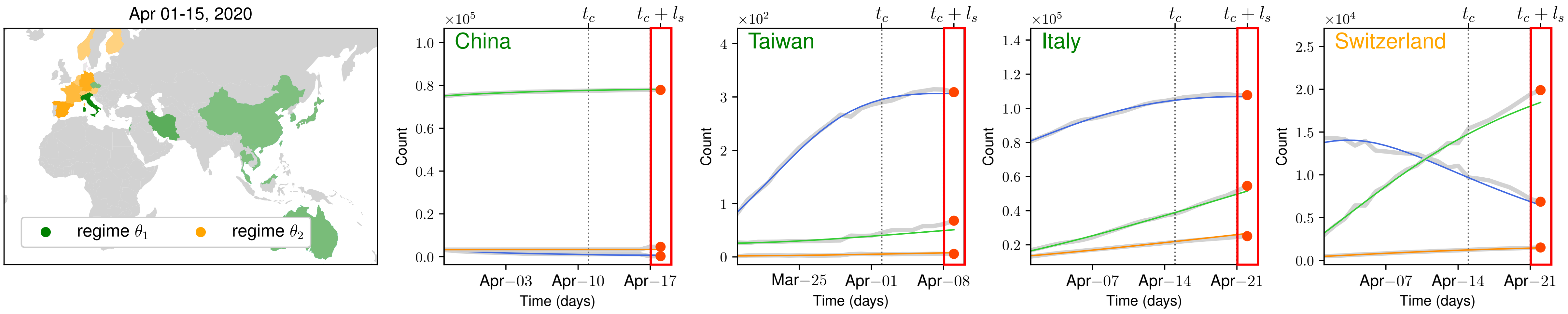



(b) Snapshots of EpiCast for the spreading process of COVID-19 from Asia to Europe

# Q1. Effectiveness



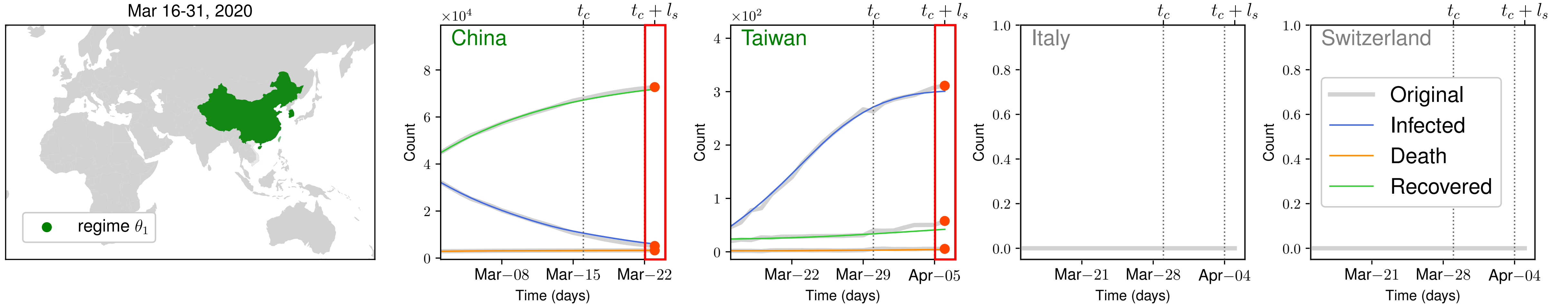

(a) Snapshots of EpiCast for the initial spread of COVID-19 in Asia



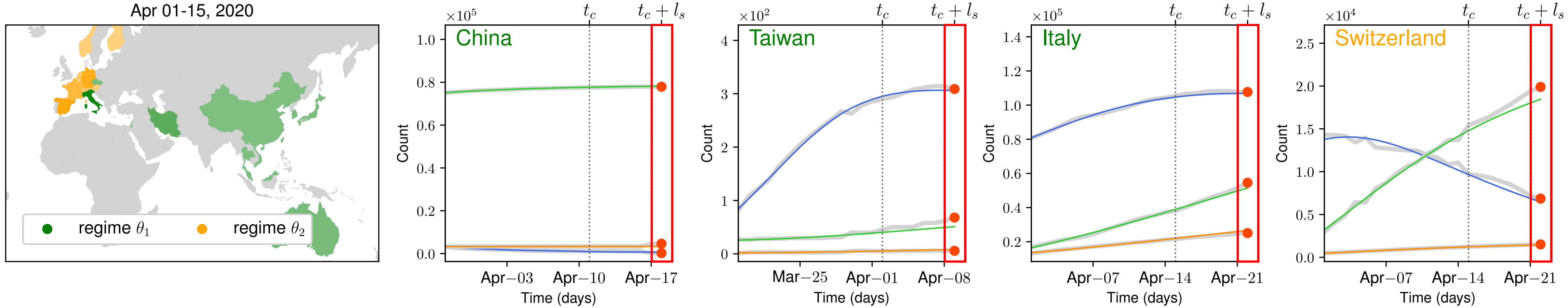

(b) Snapshots of EpiCast for the spreading process of COVID-19 from Asia to Europe

# Q1. Effectiveness

## The initial spread of COVID-19

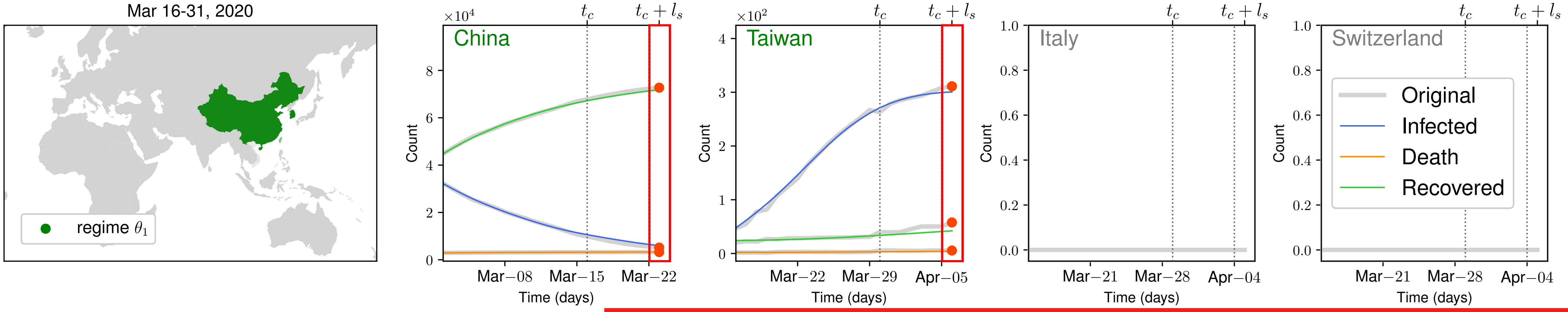



(a) Snapshots of EpiCast for the initial spread of COVID-19 in Asia



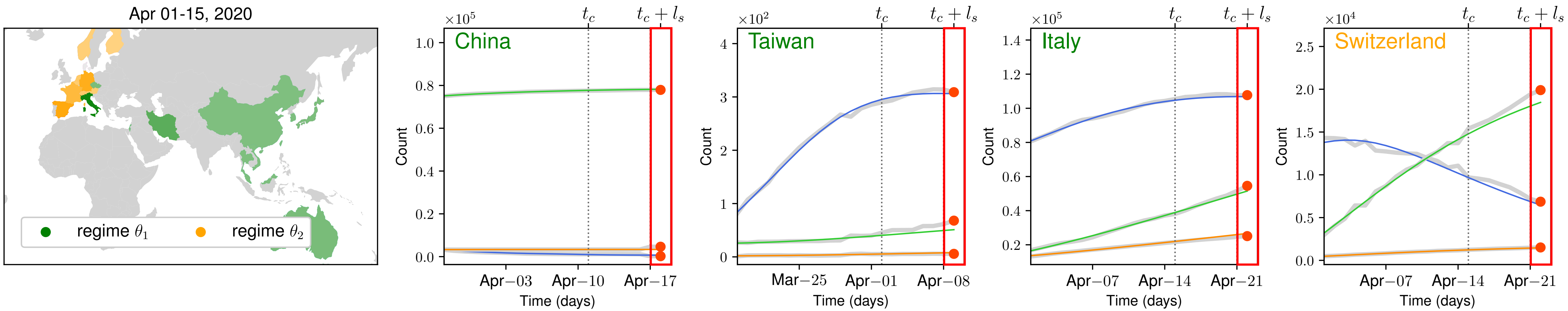

(b) Snapshots of EpiCast for the spreading process of COVID-19 from Asia to Europe

# Q1. Effectiveness



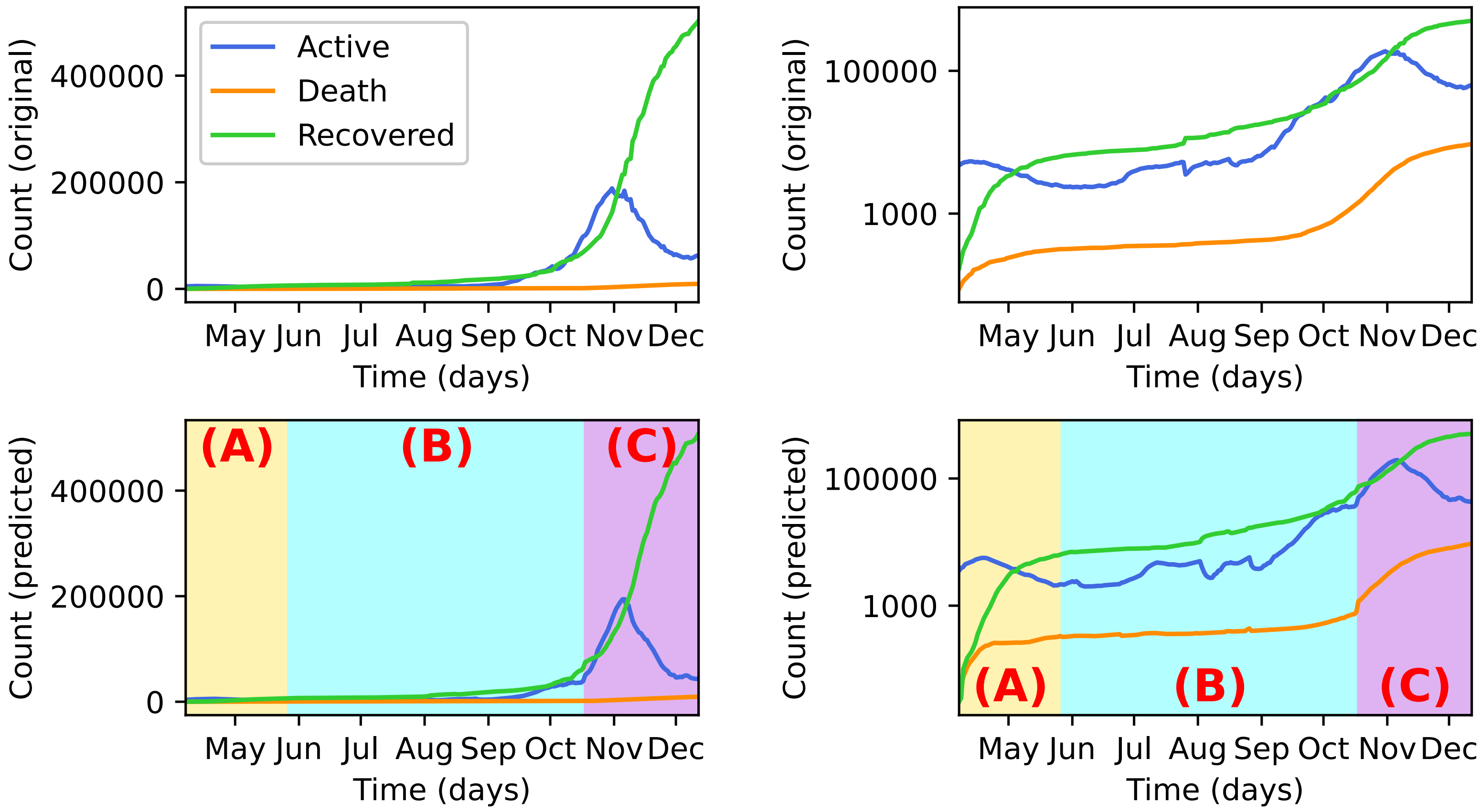

(a) Snapshots of EpiCast for the initial spread of COVID-19 in Asia

**The infection spread from Asia to Europe**



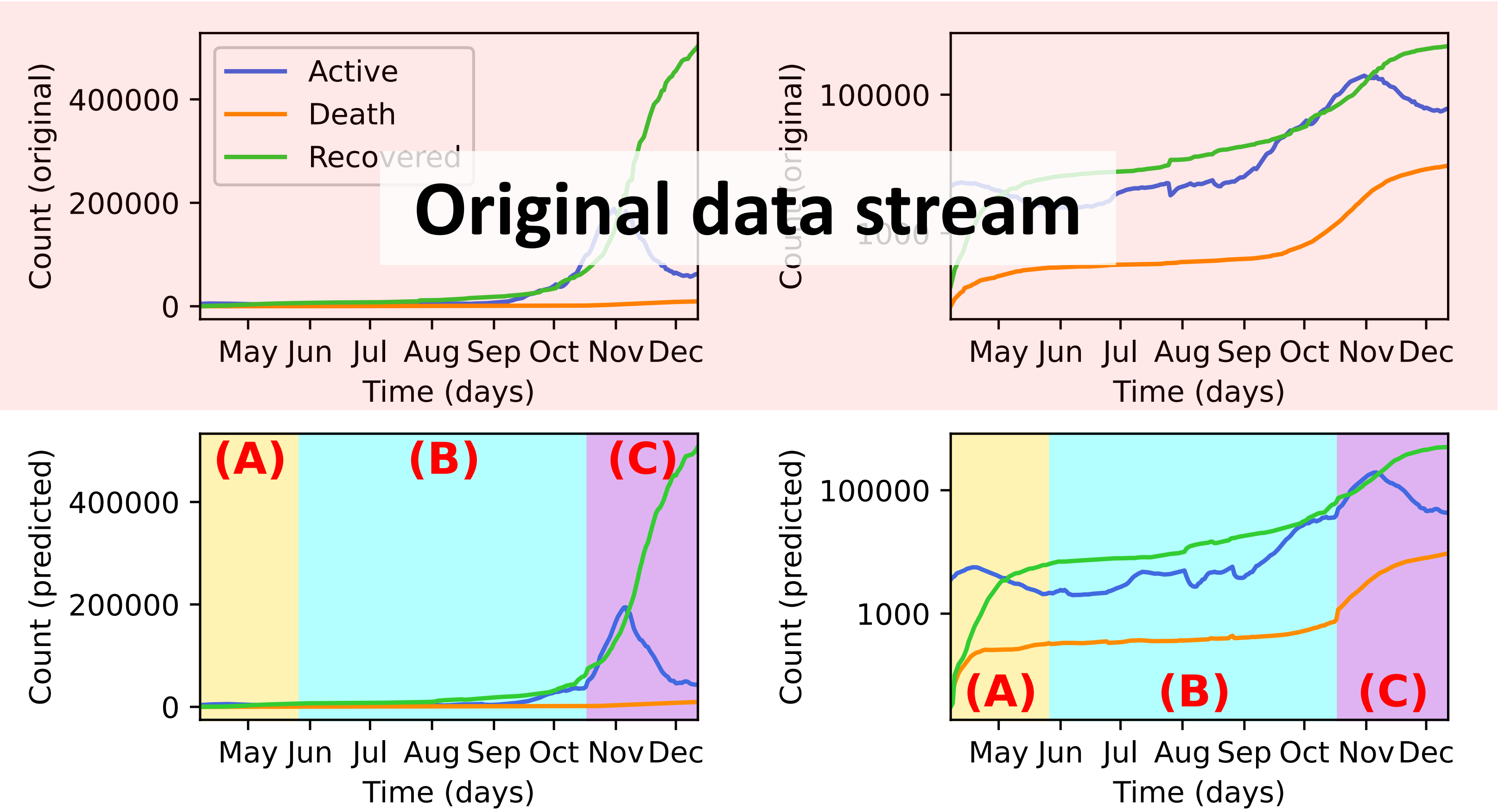

(b) Snapshots of EpiCast for the spreading process of COVID-19 from Asia to Europe

# Q1. Effectiveness



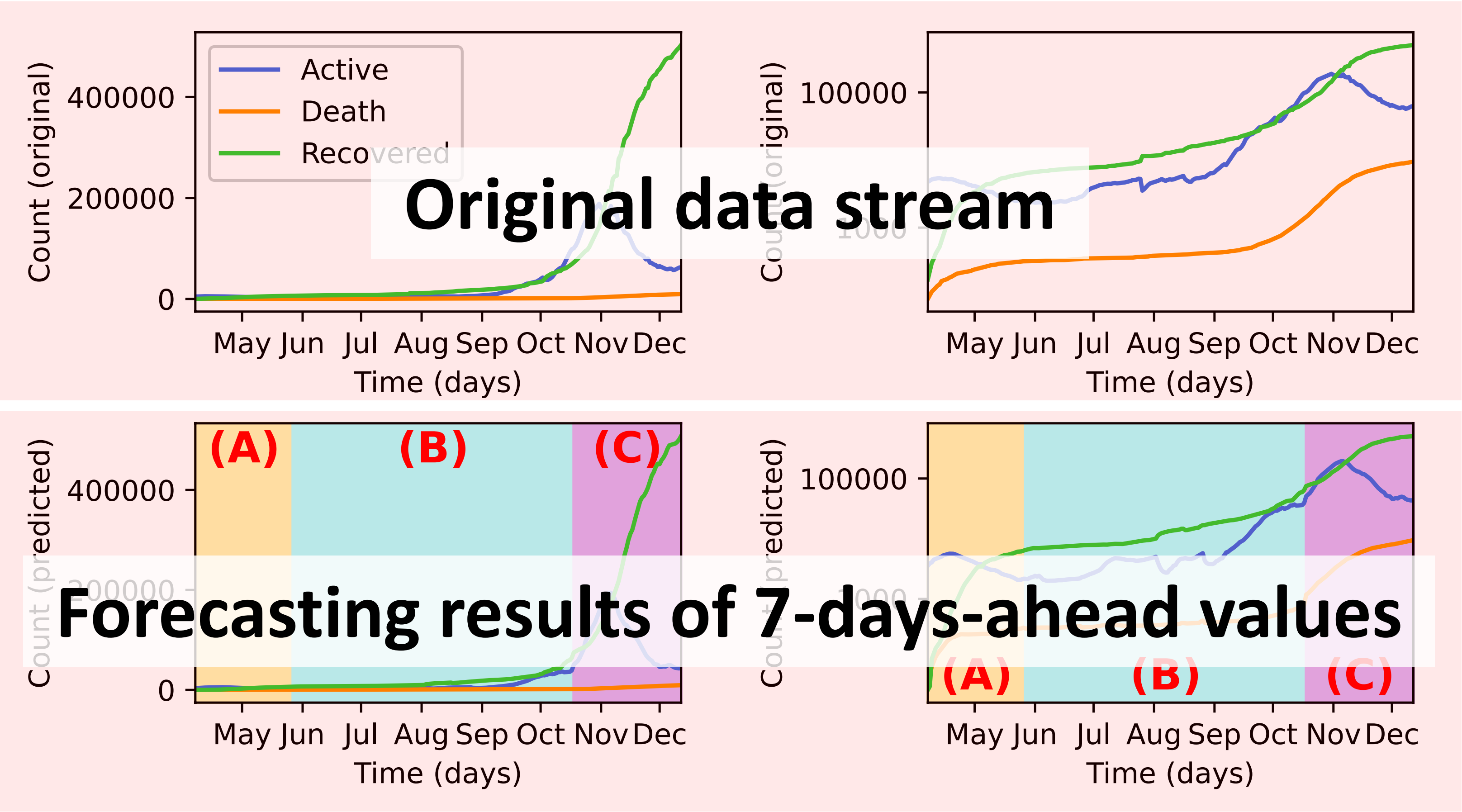

**linear-linear scale** **linear-log scale**

# Q1. Effectiveness



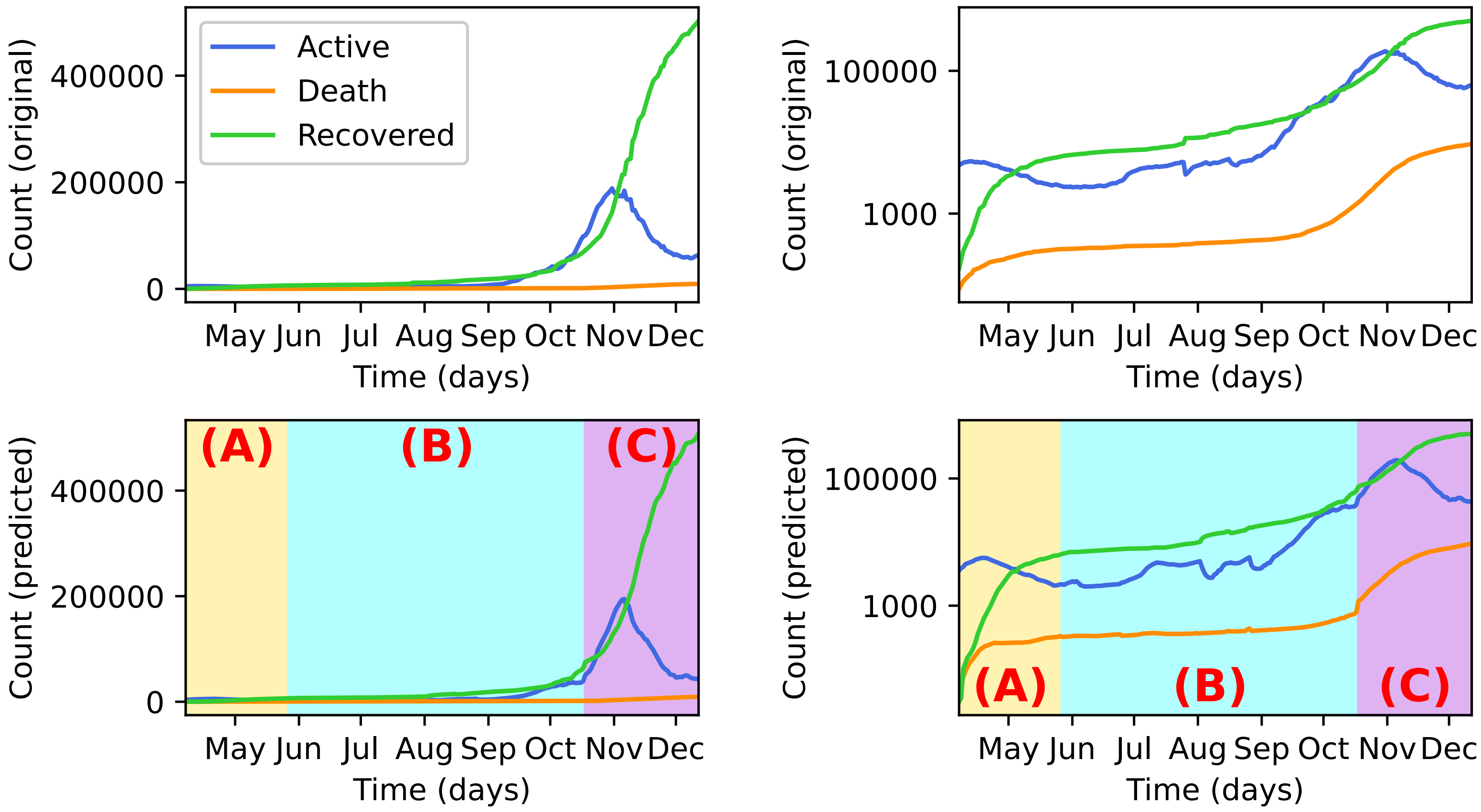

**linear-linear scale** **linear-log scale**

# Q1. Effectiveness




**linear-linear scale** **linear-log scale**

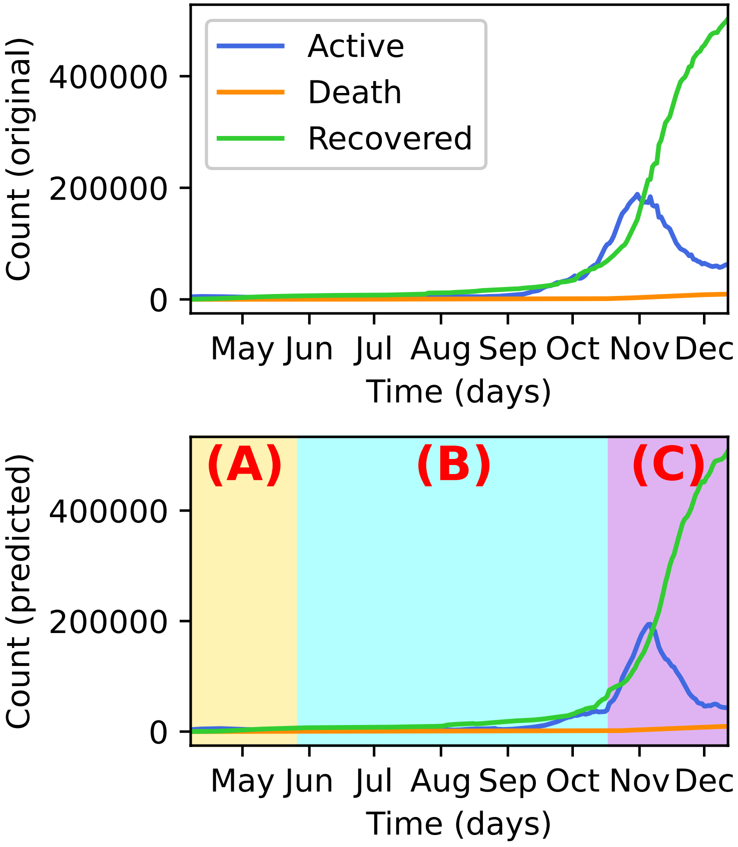
# Q1. Effectiveness



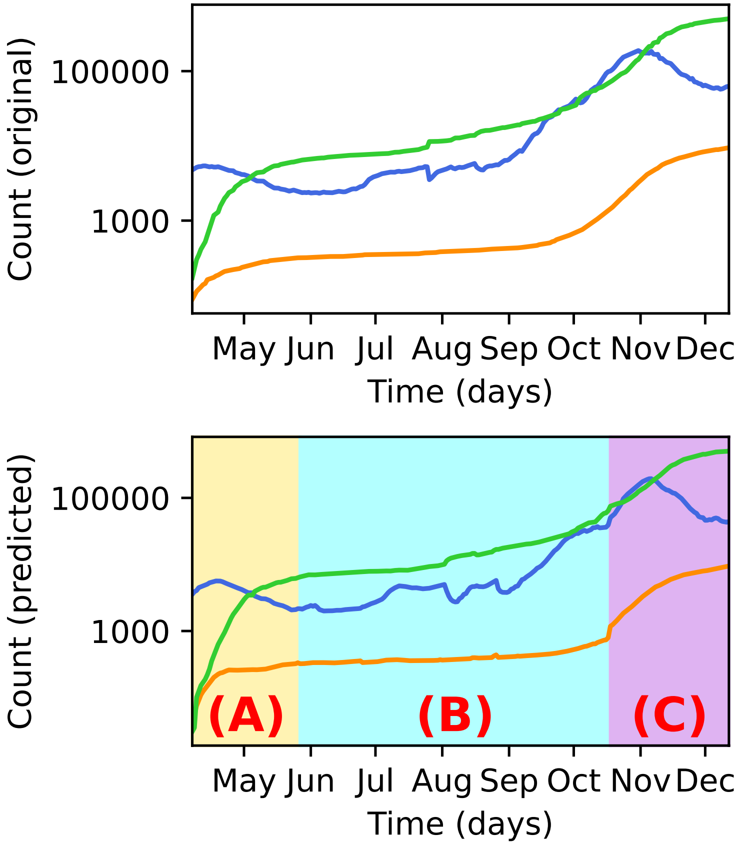

**linear-linear scale**

**linear-log scale**

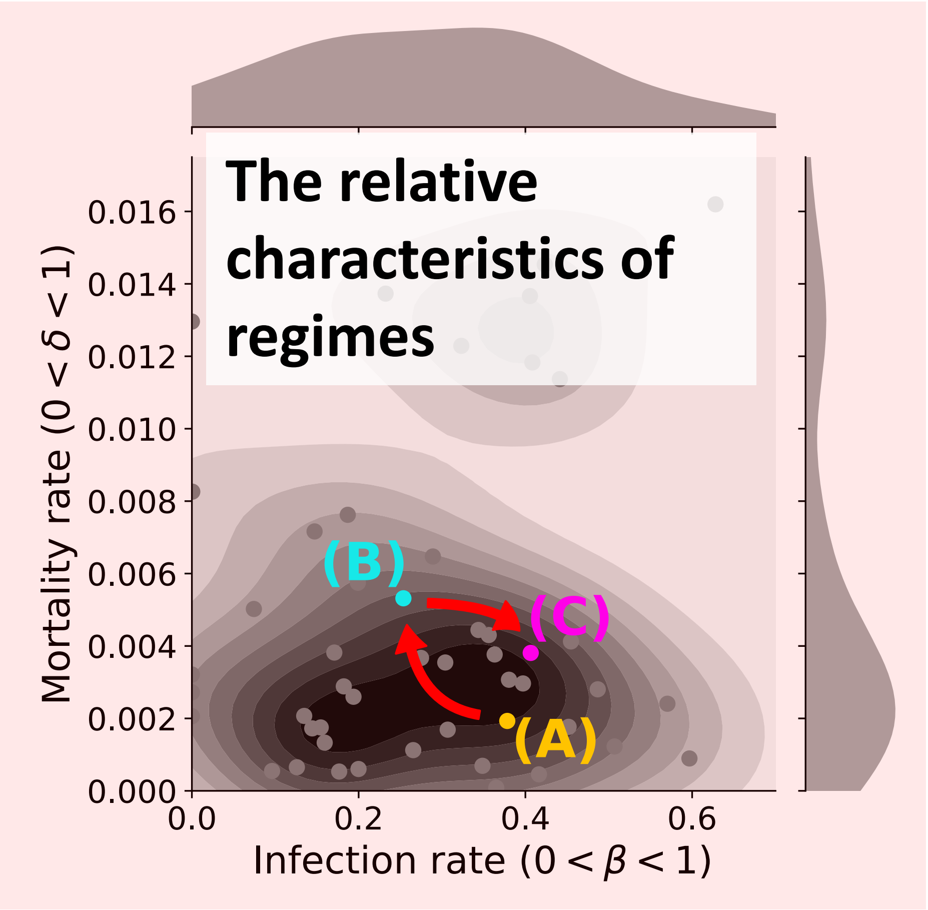
# Q1. Effectiveness



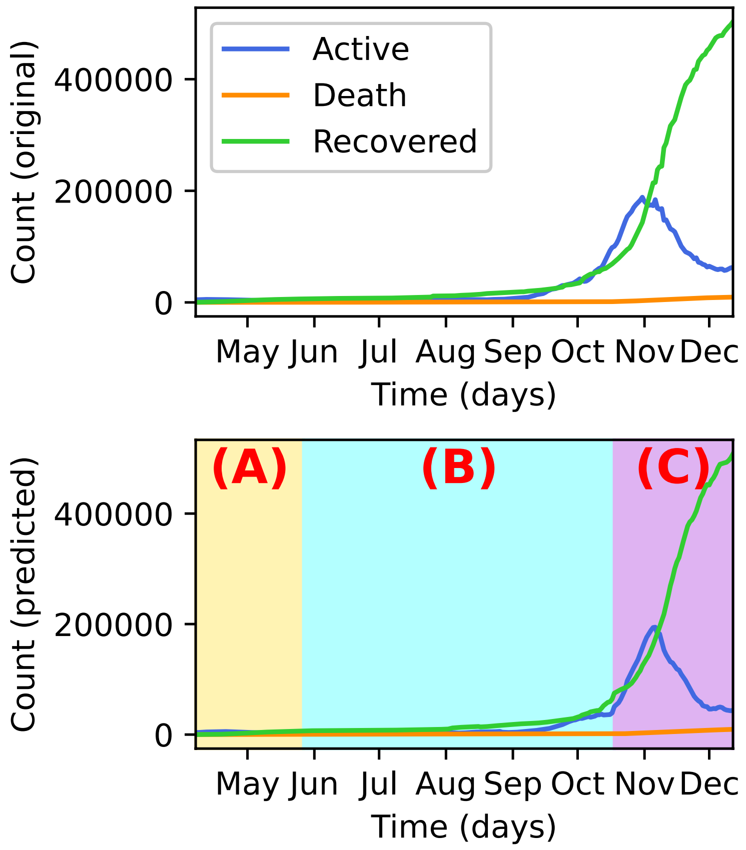

linear-linear scale

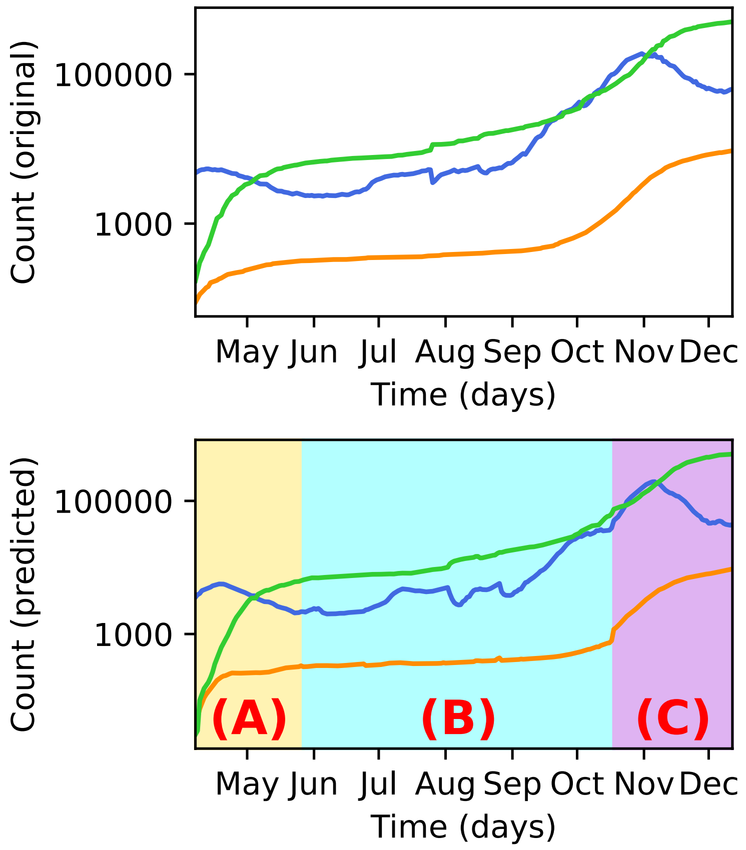



linear-log scale



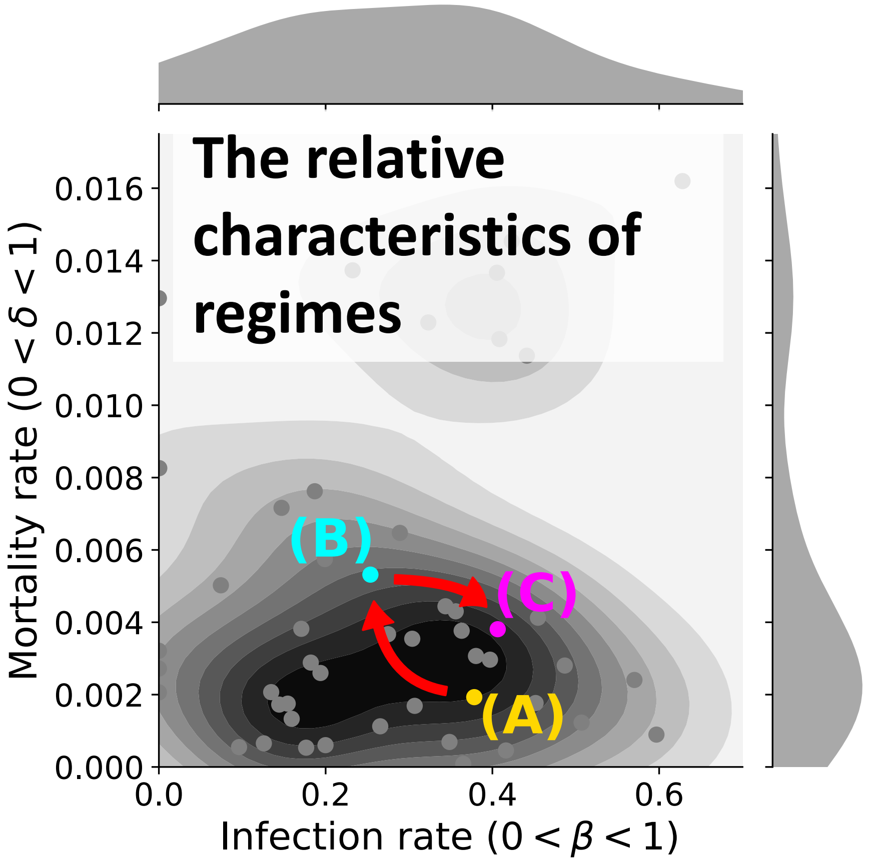
# Q1. Effectiveness



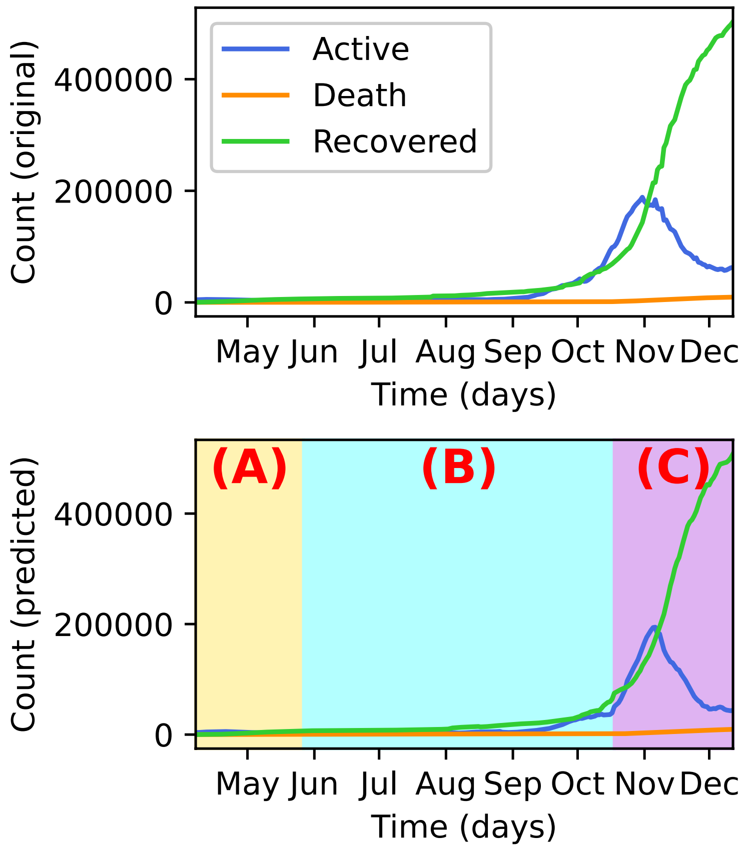

**linear-linear scale**

**linear-log scale**

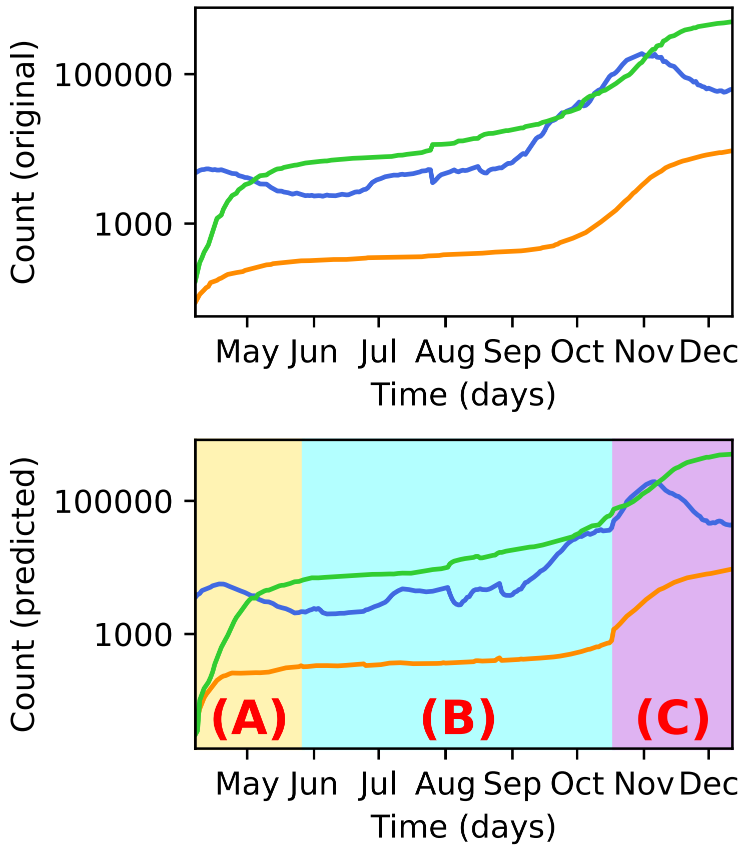


# Q1. Effectiveness

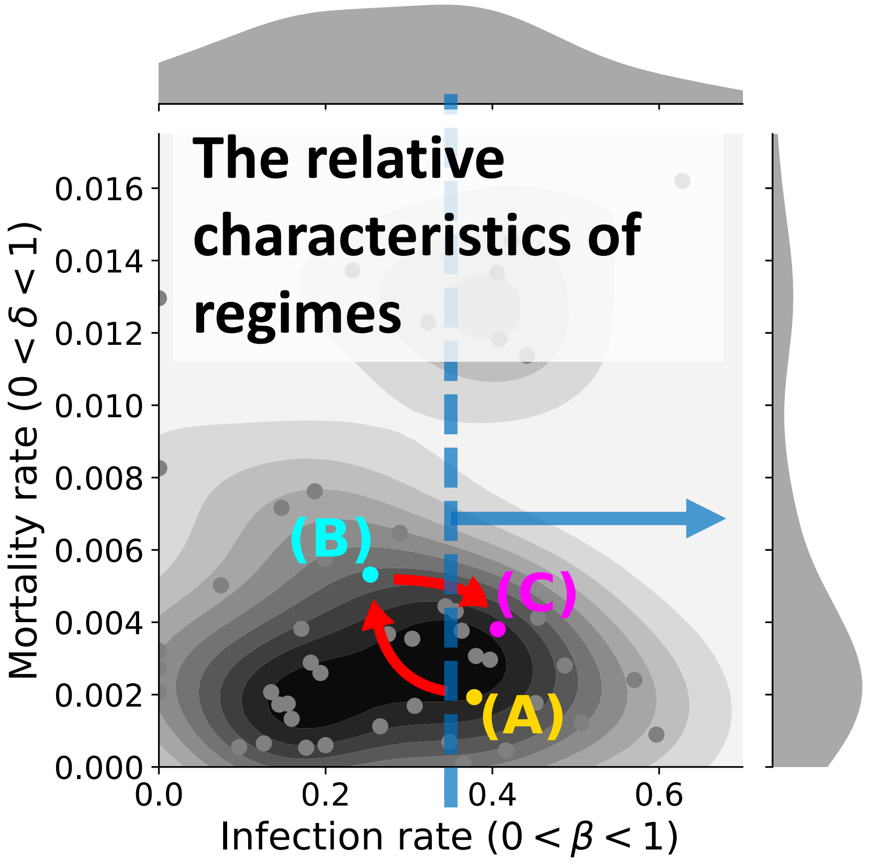
**Regimes A and B have high infection rates $\beta$**




**linear-linear scale**



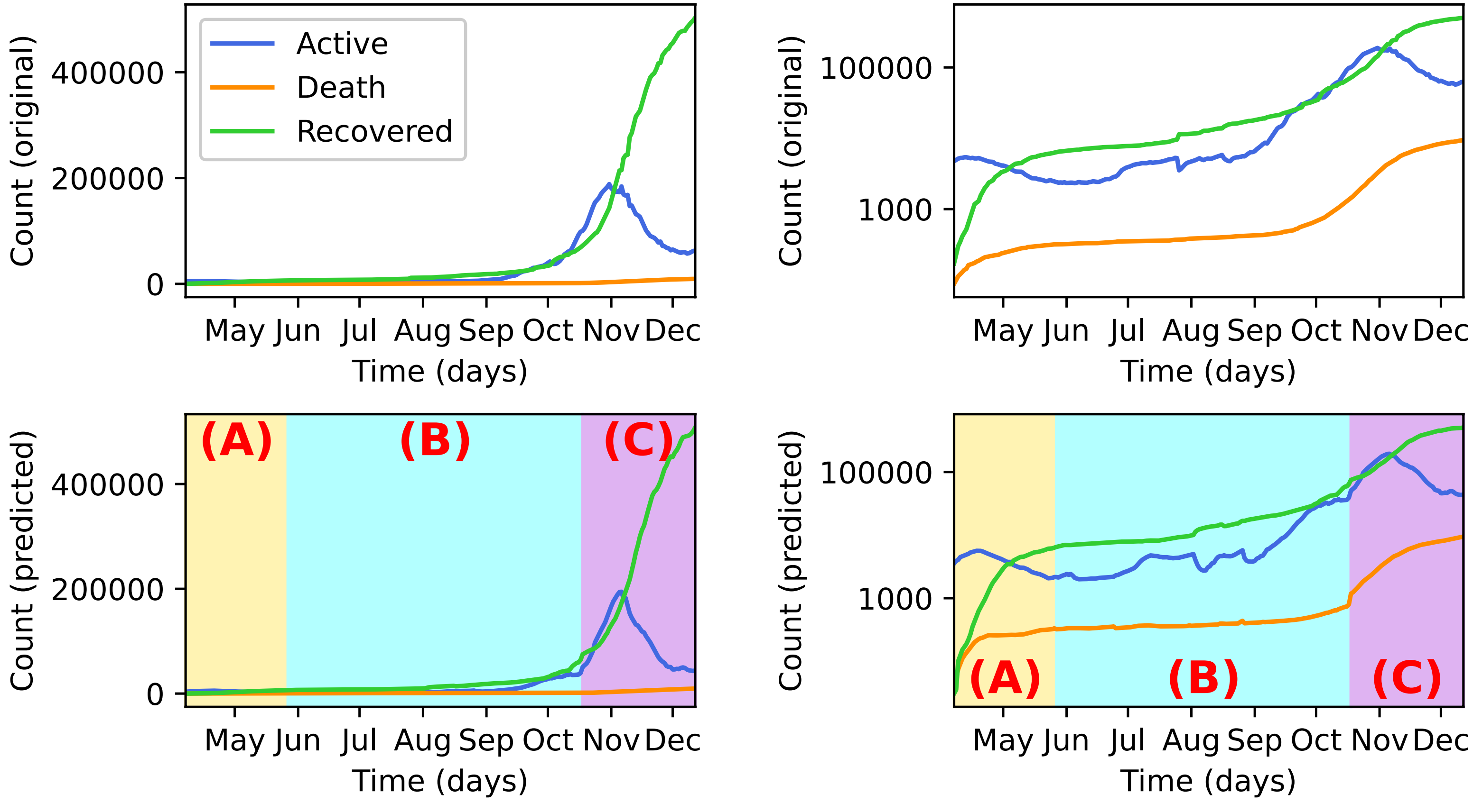

**linear-log scale**



# Q1. Effectiveness

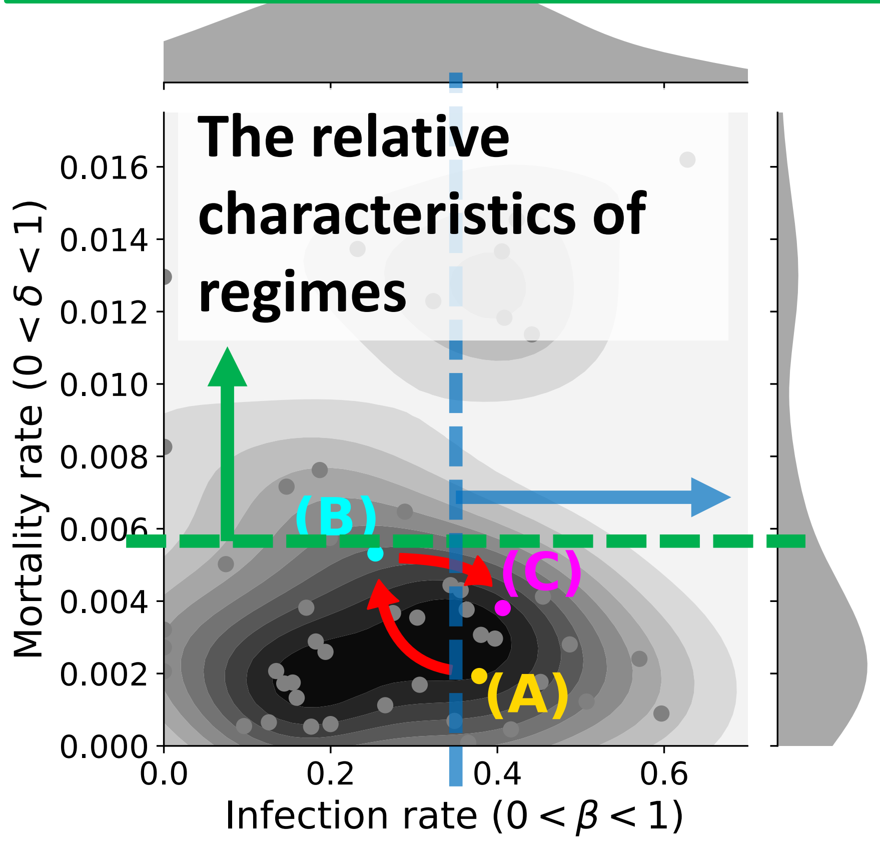
**Regimes A and B have high infection rates $\beta$**

**Regime B has high mortality rates $\gamma$**




**linear-linear scale**

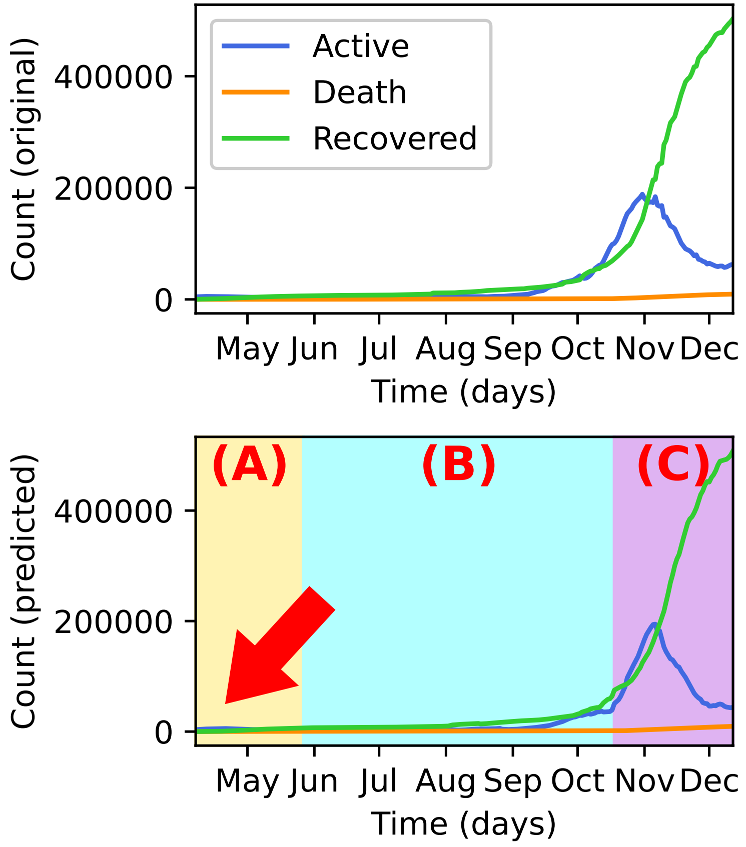
**linear-log scale**



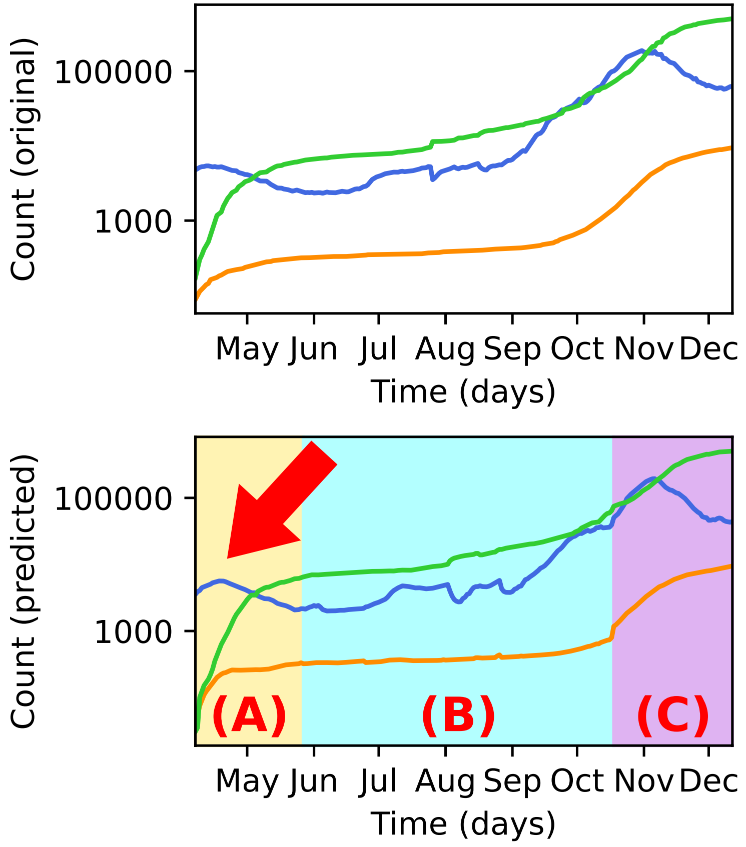
# Q1. Effectiveness

**Regimes A and B have high infection rates $\beta$**

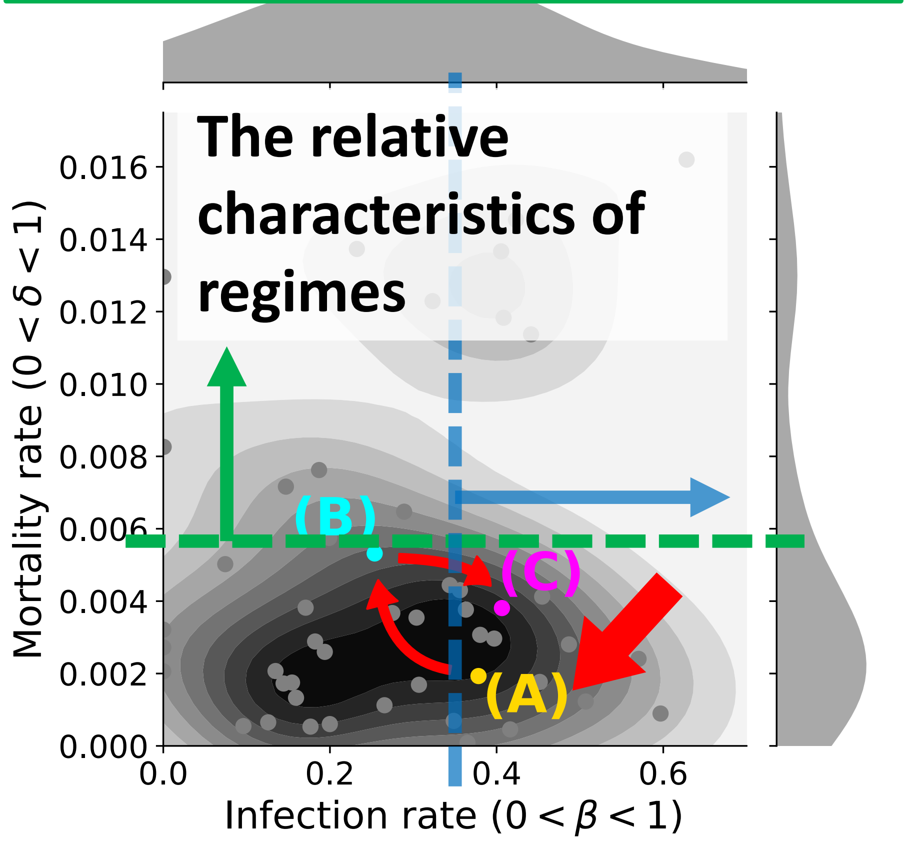
**Regime B has high mortality rates $\gamma$**




**linear-linear scale**



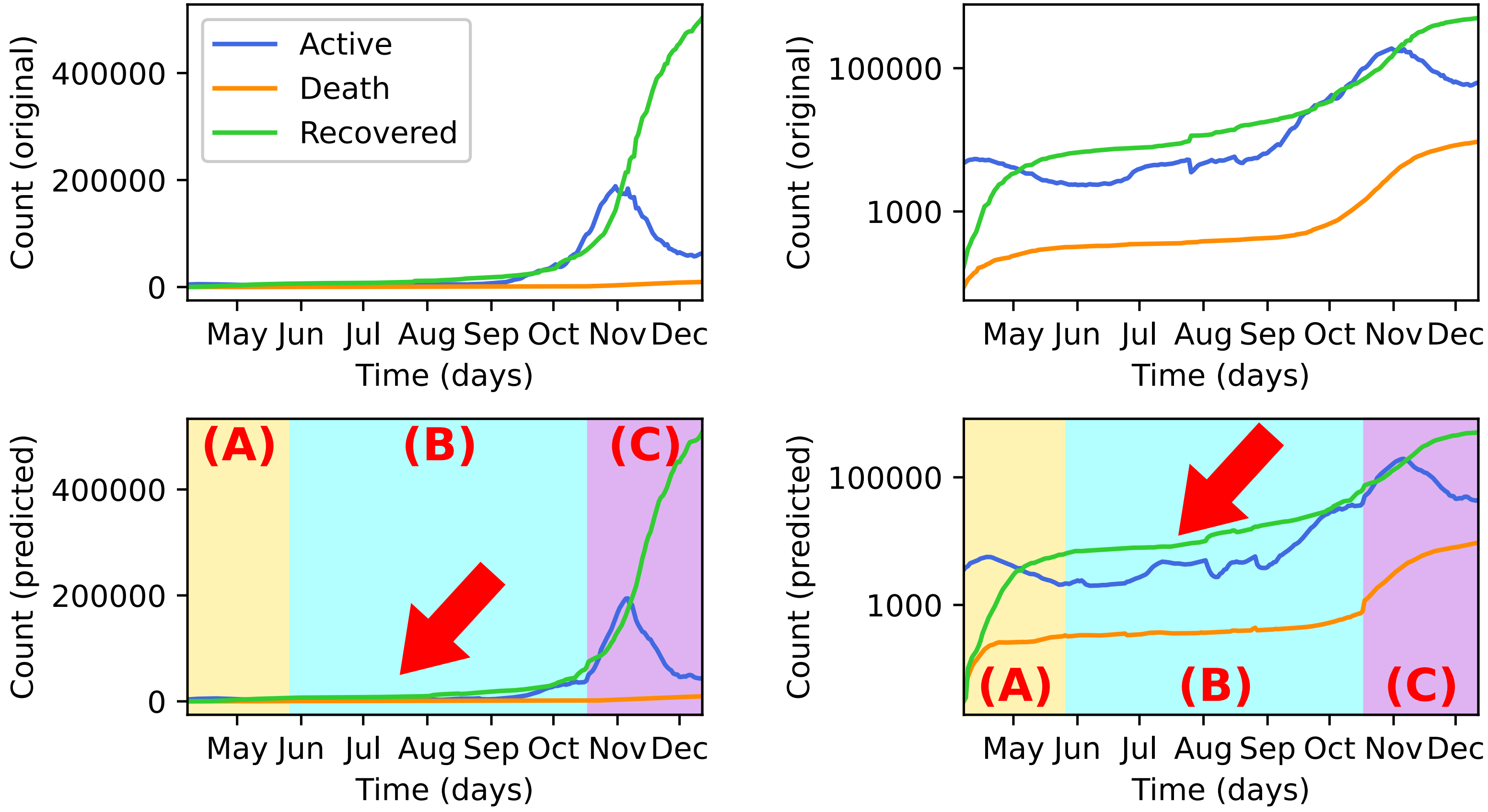

**linear-log scale**



# Q1. Effectiveness

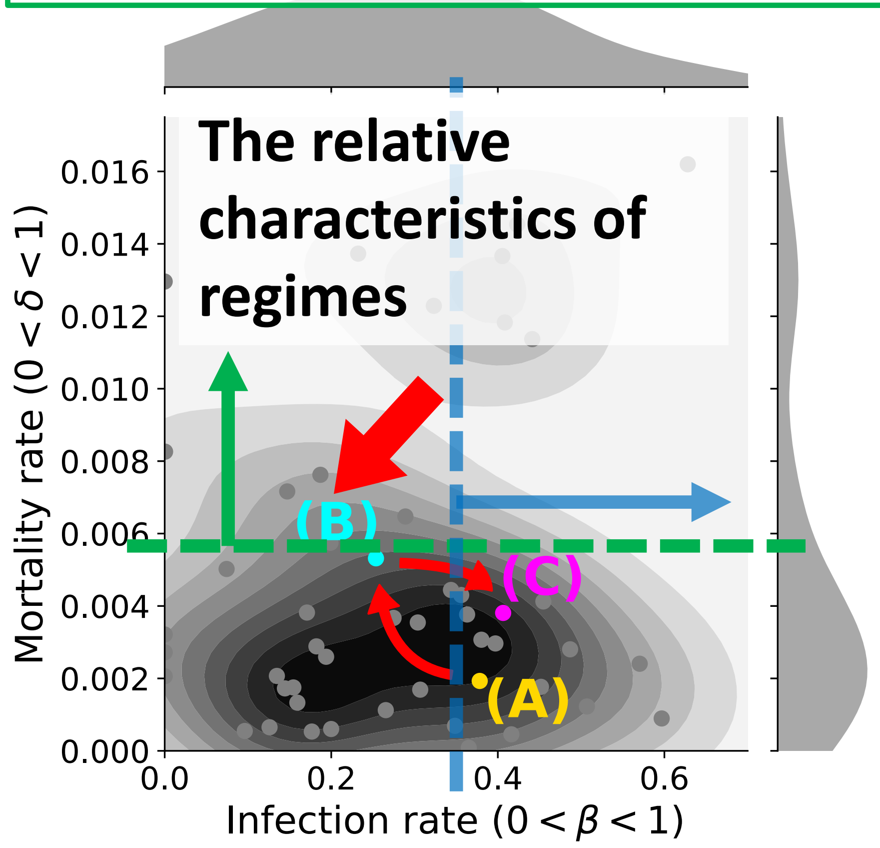
**Regimes A and B have high infection rates $\beta$**

**Regime B has high mortality rates $\gamma$**




**linear-linear scale**

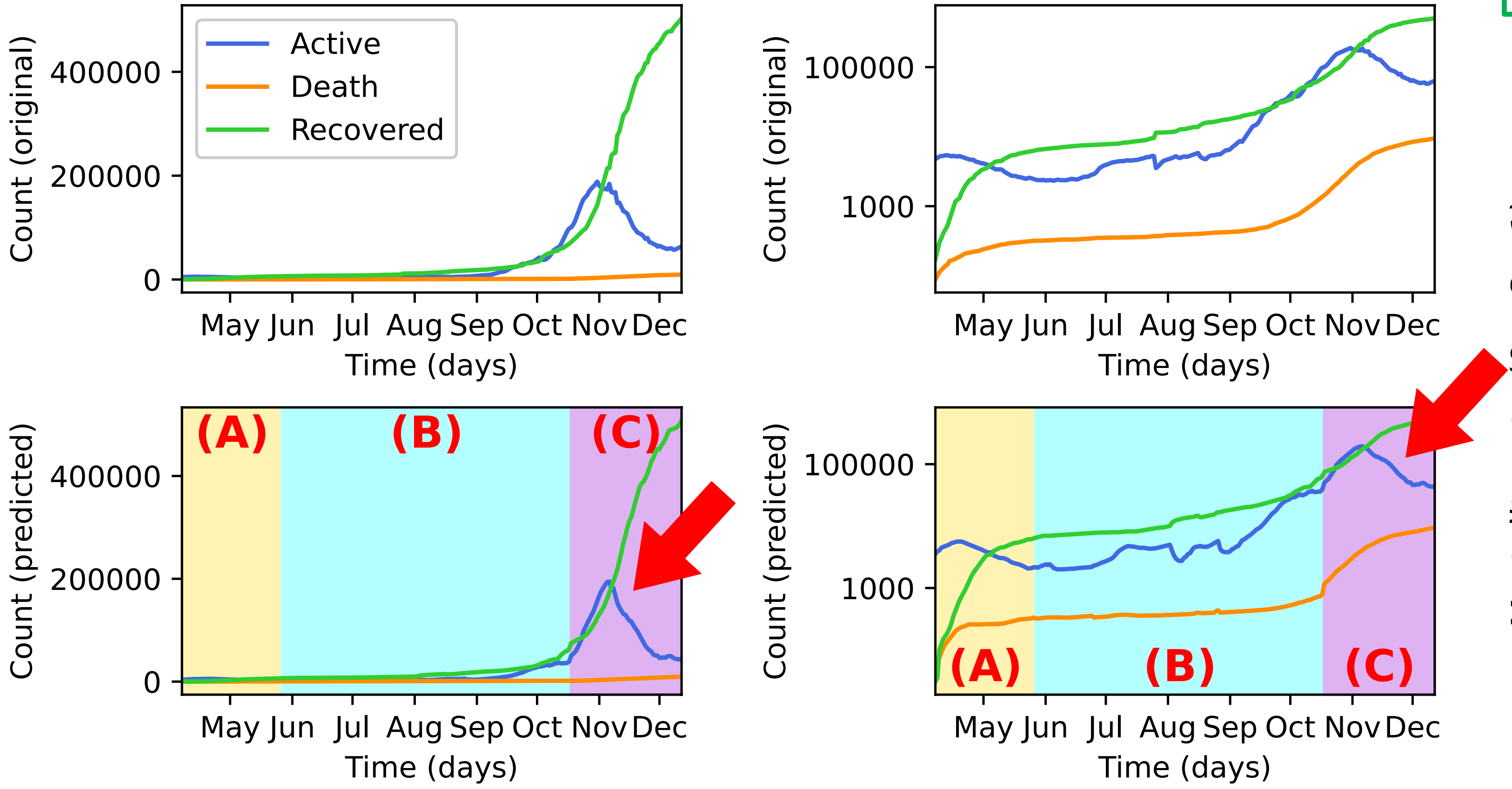
**linear-log scale**



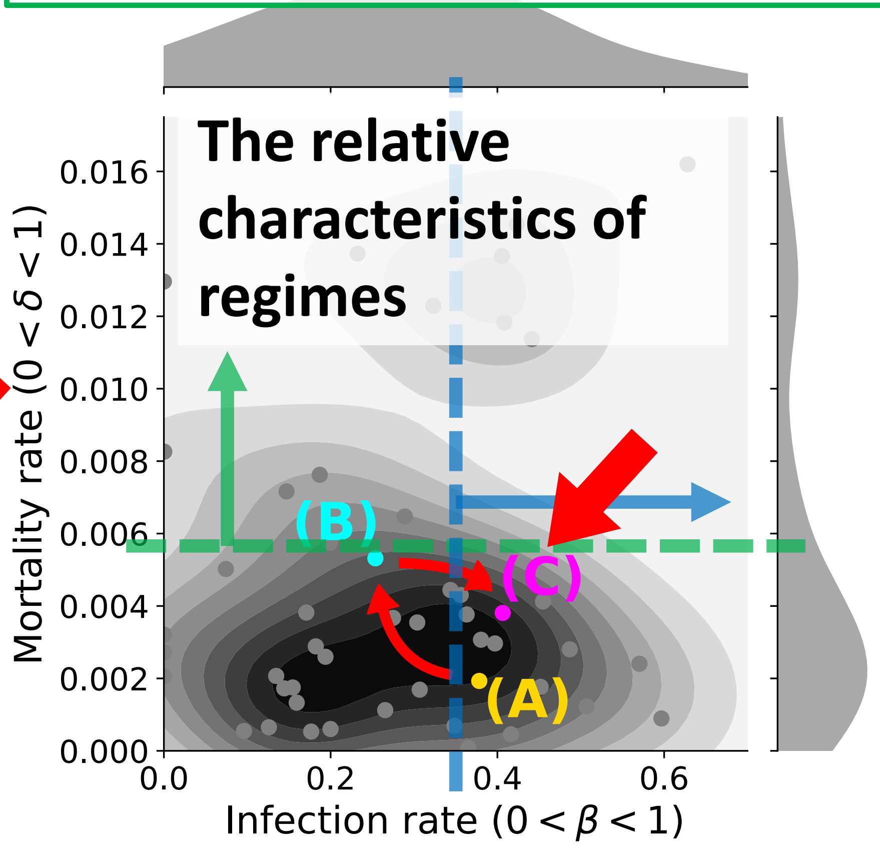
# Q1. Effectiveness

**Regimes A and B have high infection rates $\beta$**

**Regime B has high mortality rates $\gamma$**



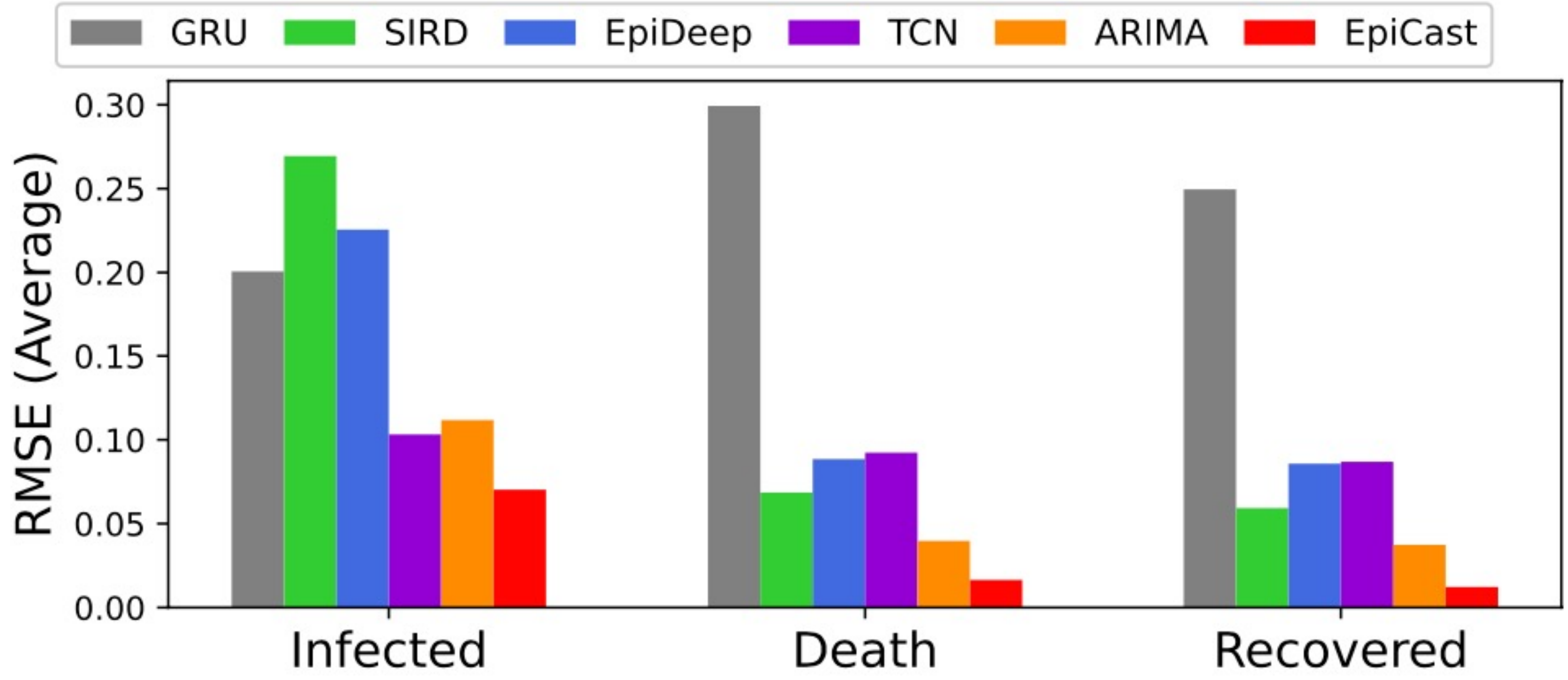

**linear-linear scale** **linear-log scale**



# Q2. Accuracy

Forecasting accuracy (Lower is better)

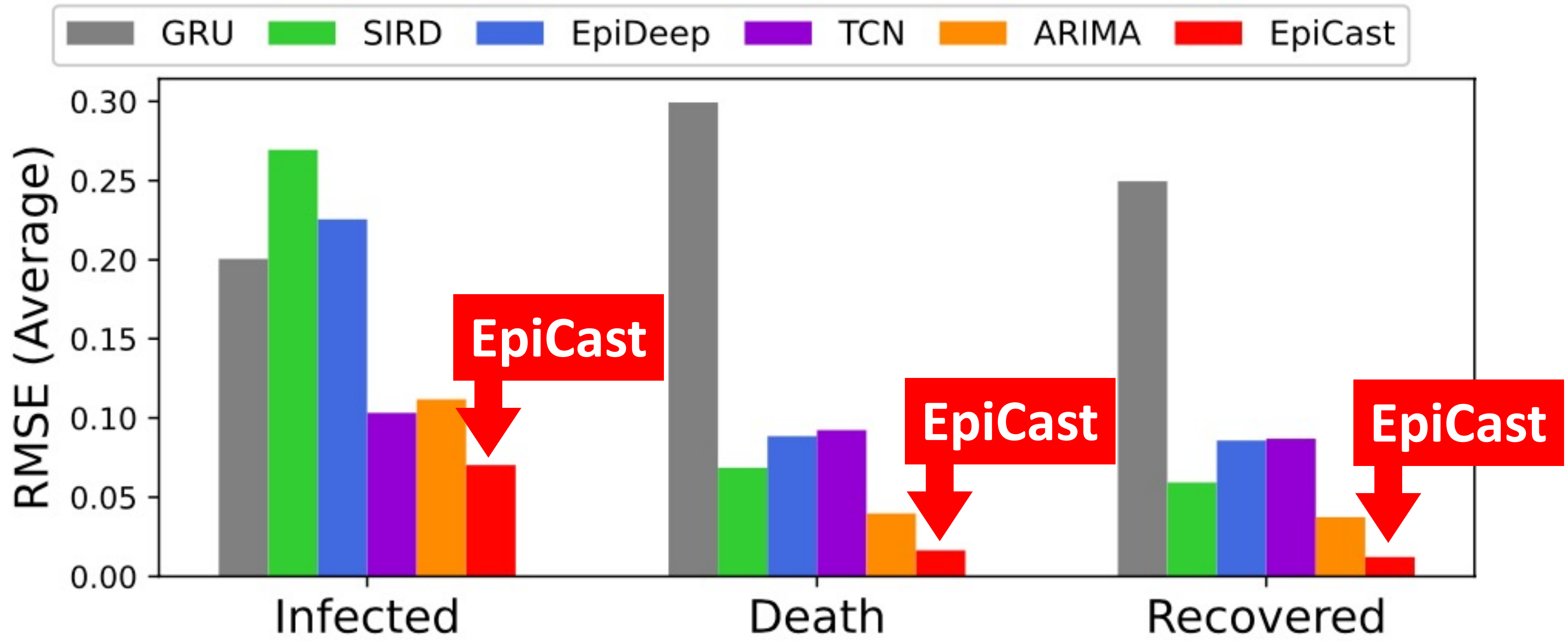
**7-days-ahead forecasting**



# Q2. Accuracy

Forecasting accuracy (Lower is better)

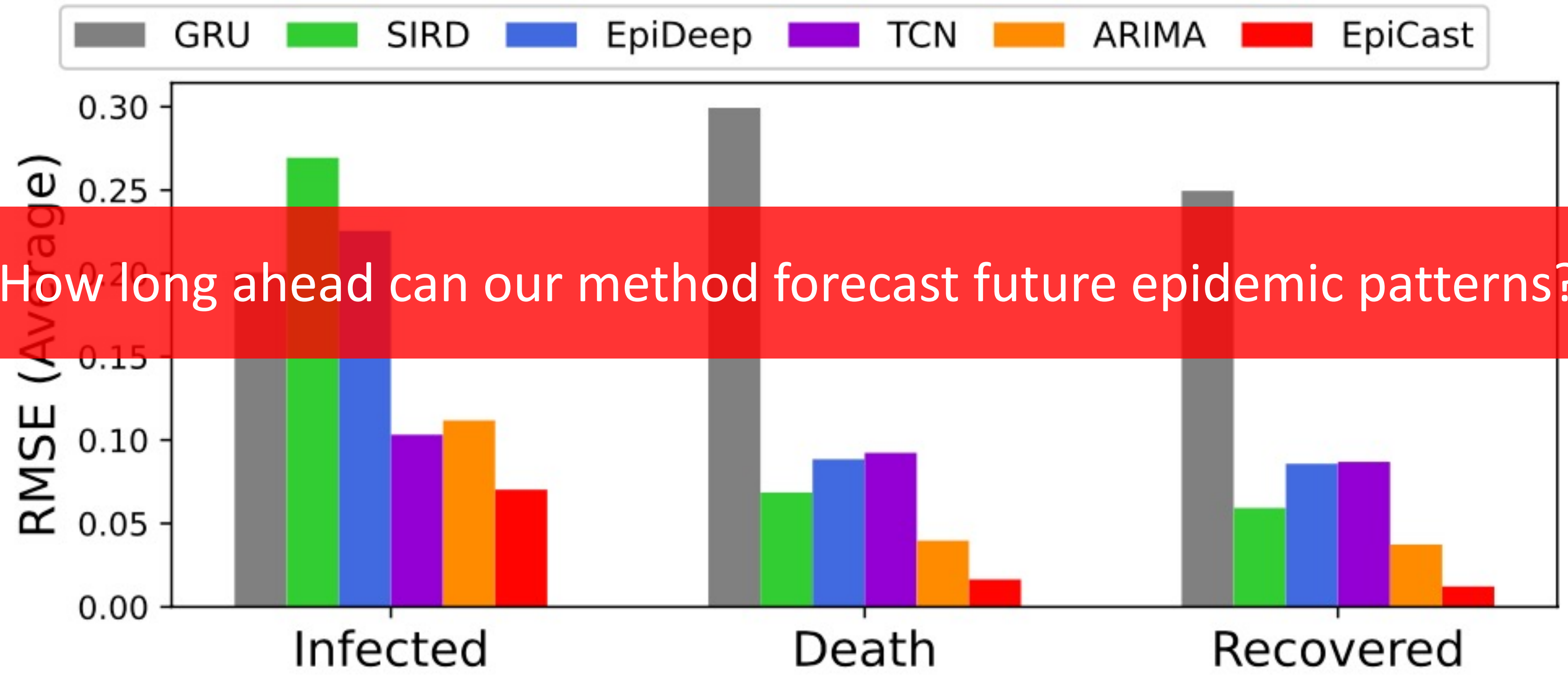
**7-days-ahead forecasting**



# Q2. Accuracy

Forecasting accuracy (Lower is better)

**7-days-ahead forecasting**

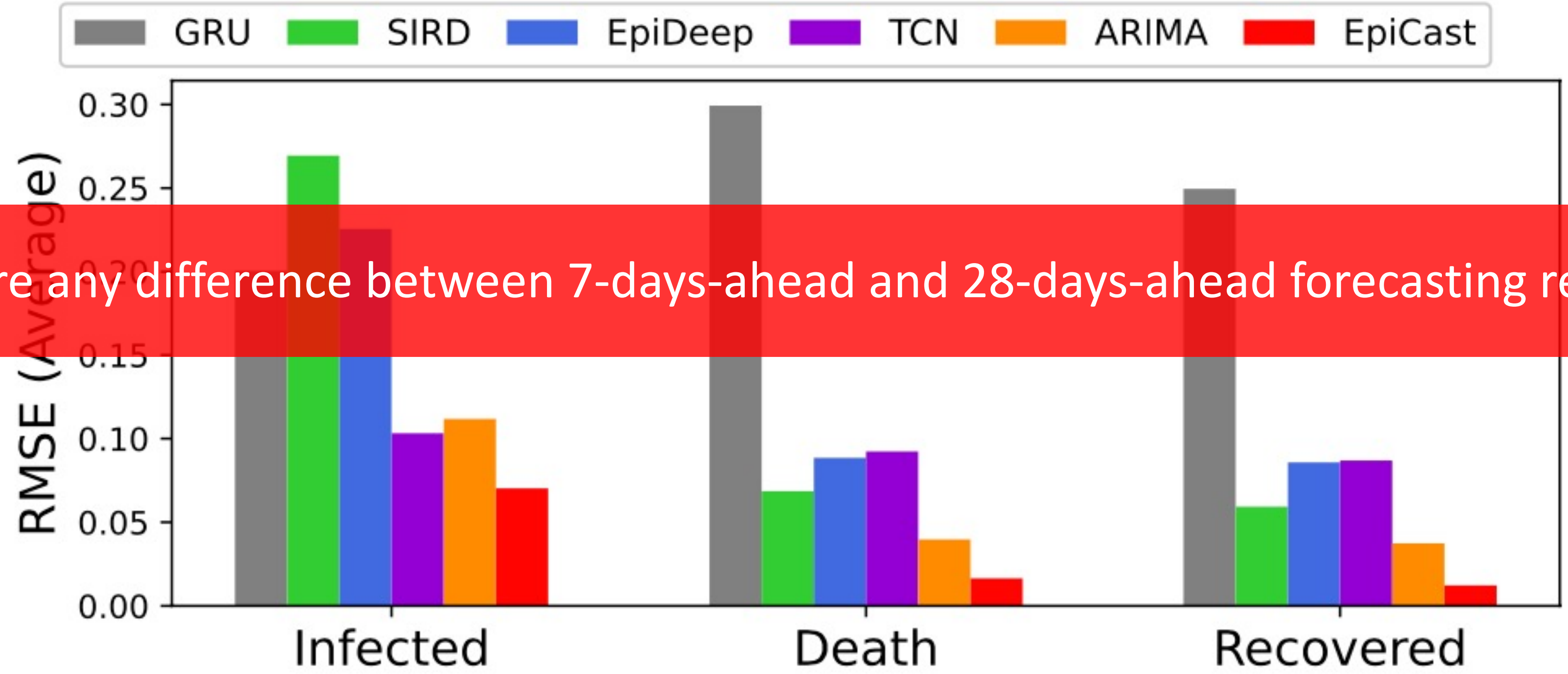



How long ahead can our method forecast future epidemic patterns?

# Q2. Accuracy

Forecasting accuracy (Lower is better)

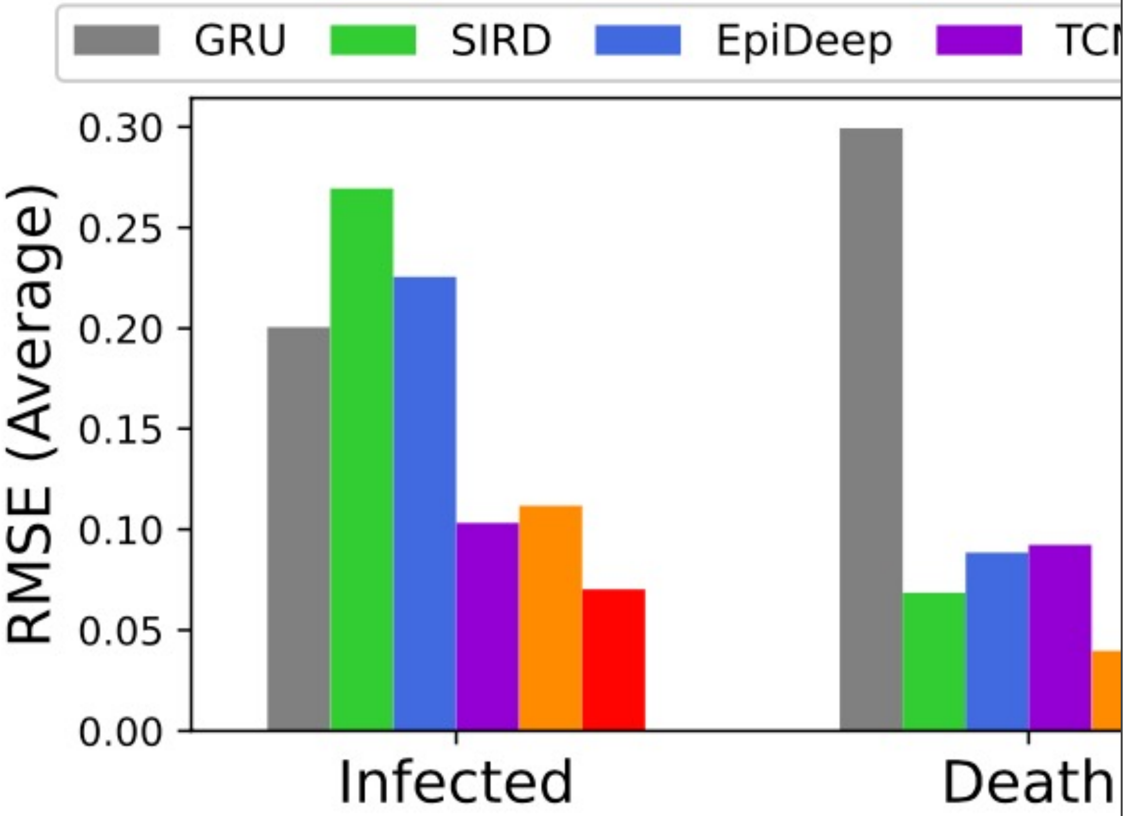
**7-days-ahead forecasting**




Is there any difference between 7-days-ahead and 28-days-ahead forecasting results?

# Q2. Accuracy

Forecasting accuracy (Lower is better)




## Details in appendix of paper

**Table 2: Comparison of marginal/average forecasting error: RMSE between original and $l_s$-steps-ahead values of EpiCast and its competitors at each current window length $l_c$ (lower is better).**

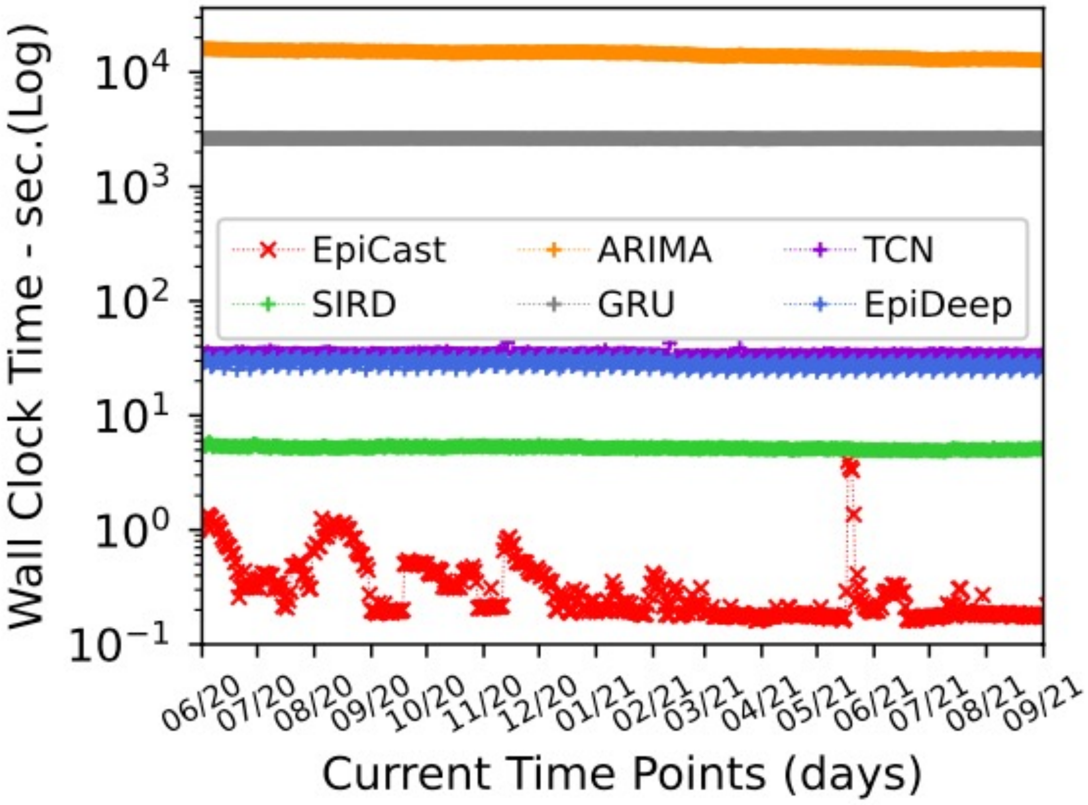
| $l_c$ | $l_s$ | EpiCast | SIRD | ARIMA | GRU | TCN | EpiDeep |
|---|---|---|---|---|---|---|---|
| 14 | 7 | .0330 ± .0327 | .1325 ± .1166 | .0630 ± .0398 | .2498 ± .0925 | .0942 ± .0171 | .1333 ± .0826 |
| | 14 | .0583 ± .0587 | .1977 ± .1716 | .0912 ± .0551 | .2760 ± .1080 | .0856 ± .0442 | .1586 ± .1141 |
| | 21 | .0841 ± .0833 | .2541 ± .2100 | .1172 ± .0682 | .2947 ± .1267 | .1035 ± .0622 | .1845 ± .1827 |
| | 28 | .1082 ± .1040 | .2992 ± .2335 | .1408 ± .0801 | .3123 ± .1542 | .1235 ± .0739 | .2187 ± .3827 |
| 21 | 7 | .0376 ± .0364 | .1313 ± .0969 | .0761 ± .0463 | .2855 ± .1092 | .0900 ± .0259 | .1332 ± .0827 |
| | 14 | .0632 ± .0642 | .1844 ± .1383 | .1032 ± .0603 | .3093 ± .1292 | .0902 ± .0520 | .1584 ± .1144 |
| | 21 | .0882 ± .0858 | .2323 ± .1685 | .1285 ± .0736 | .3274 ± .1540 | .1102 ± .0668 | .1844 ± .1834 |
| | 28 | .1114 ± .1040 | .2725 ± .1886 | .1517 ± .0893 | .3432 ± .1849 | .1306 ± .0768 | .2187 ± .3846 |
| 28 | 7 | .0426 ± .0419 | .1288 ± .0922 | .0884 ± .0519 | .3038 ± .1347 | .0927 ± .0312 | .1361 ± .0981 |
| | 14 | .0679 ± .0700 | .1671 ± .1180 | .1149 ± .0660 | .3294 ± .1635 | .0950 ± .0543 | .1634 ± .1579 |
| | 21 | .0926 ± .0927 | .2080 ± .1465 | .1397 ± .0828 | .3480 ± .1963 | .1152 ± .0684 | .1936 ± .2859 |
| | 28 | .1160 ± .1177 | .2462 ± .1716 | .1630 ± .1101 | .3660 ± .2515 | .1352 ± .0781 | .2191 ± .3948 |

**Table 3: Comparison of individual forecasting error: RMSE between original and $l_s$-steps-ahead values of EpiCast and its competitors at each current window length $l_c$ (lower is better). Please also see text for more observations.**

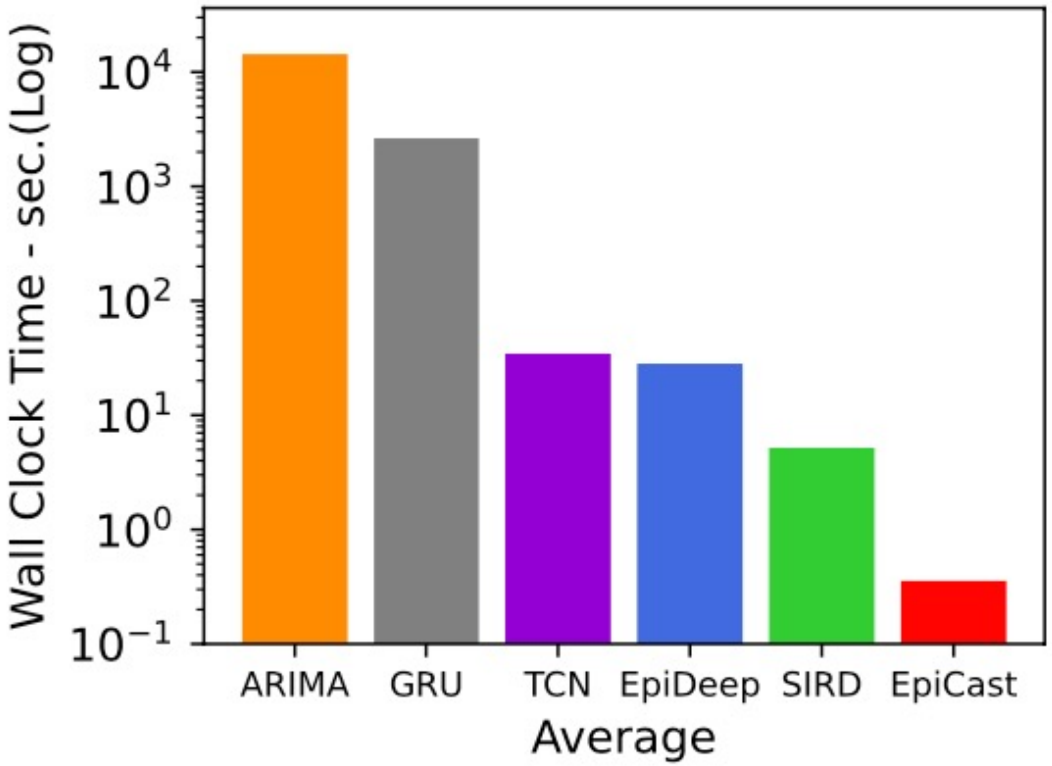
| $l_c$ | $l_s$ | dimension | EpiCast | SIRD | ARIMA | GRU | TCN | EpiDeep |
|---|---|---|---|---|---|---|---|---|
| 14 | 7 | Infected | .0703 ± .0266 | .2695 ± .1038 | .1118 ± .0301 | .2006 ± .0506 | .1032 ± .0220 | .2256 ± .0691 |
| | | Death | .0165 ± .0173 | .0686 ± .0379 | .0398 ± .0130 | .2993 ± .1085 | .0923 ± .0120 | .0885 ± .0340 |
| | | Recovered | .0123 ± .0108 | .0593 ± .0188 | .0374 ± .0102 | .2496 ± .0821 | .0869 ± .0110 | .0858 ± .0421 |
| | 14 | Infected | .1247 ± .0412 | .3852 ± .1670 | .1576 ± .0421 | .2357 ± .0549 | .1371 ± .0414 | .2639 ± .1311 |
| | | Death | .0305 ± .0418 | .1311 ± .0763 | .0602 ± .0217 | .3271 ± .1382 | .0596 ± .0099 | .1072 ± .0386 |
| | | Recovered | .0197 ± .0149 | .0768 ± .0229 | .0557 ± .0157 | .2652 ± .0944 | .0601 ± .0087 | .1046 ± .0628 |
| | 21 | Infected | .1790 ± .0569 | .4723 ± .2081 | .1972 ± .0521 | .2685 ± .0655 | .1772 ± .0546 | .3058 ± .2660 |
| | | Death | .0449 ± .0597 | .1974 ± .1076 | .0802 ± .0315 | .3424 ± .1664 | .0658 ± .0164 | .1252 ± .0426 |
| | | Recovered | .0283 ± .0192 | .0926 ± .0271 | .0741 ± .0248 | .2732 ± .1157 | .0674 ± .0141 | .1225 ± .0800 |
| | 28 | Infected | .2272 ± .0701 | .5315 ± .2306 | .2301 ± .0636 | .2996 ± .0988 | .2109 ± .0640 | .3742 ± .6305 |
| | | Death | .0599 ± .0742 | .2590 ± .1319 | .0993 ± .0392 | .3535 ± .1854 | .0788 ± .0217 | .1424 ± .0462 |
| | | Recovered | .0374 ± .0225 | .1072 ± .0311 | .0930 ± .0413 | .2838 ± .1598 | .0807 ± .0184 | .1393 ± .0940 |
| 21 | 7 | Infected | .0801 ± .0247 | .2503 ± .0670 | .1320 ± .0351 | .2305 ± .0627 | .1158 ± .0299 | .2256 ± .0693 |
| | | Death | .0194 ± .0227 | .0769 ± .0449 | .0497 ± .0176 | .3378 ± .1309 | .0783 ± .0081 | .0883 ± .0335 |
| | | Recovered | .0133 ± .0112 | .0667 ± .0180 | .0465 ± .0131 | .2882 ± .0966 | .0760 ± .0076 | .0856 ± .0421 |
| | 14 | Infected | .1332 ± .0373 | .3436 ± .1143 | .1745 ± .0456 | .2665 ± .0728 | .1510 ± .0472 | .2639 ± .1317 |
| | | Death | .0352 ± .0575 | .1280 ± .0691 | .0701 ± .0273 | .3613 ± .1656 | .0590 ± .0138 | .1070 ± .0381 |
| | | Recovered | .0213 ± .0157 | .0816 ± .0209 | .0649 ± .0208 | .3000 ± .1158 | .0607 ± .0116 | .1044 ± .0629 |
| | 21 | Infected | .1856 ± .0488 | .4179 ± .1467 | .2122 ± .0570 | .3019 ± .1089 | .1889 ± .0590 | .3059 ± .2675 |
| | | Death | .0490 ± .0704 | .1833 ± .0870 | .0895 ± .0348 | .3727 ± .1867 | .0696 ± .0194 | .1250 ± .0420 |
| | | Recovered | .0301 ± .0196 | .0956 ± .0240 | .0840 ± .0363 | .3077 ± .1495 | .0721 ± .0158 | .1223 ± .0801 |
| | 28 | Infected | .2312 ± .0643 | .4720 ± .1677 | .2441 ± .0802 | .3329 ± .1634 | .2208 ± .0672 | .3747 ± .6338 |
| | | Death | .0631 ± .0771 | .2361 ± .1006 | .1081 ± .0410 | .3813 ± .2014 | .0840 ± .0242 | .1422 ± .0456 |
| | | Recovered | .0398 ± .0229 | .1095 ± .0285 | .1028 ± .0551 | .3154 ± .1854 | .0870 ± .0194 | .1391 ± .0941 |
| 28 | 7 | Infected | .0905 ± .0285 | .2468 ± .0603 | .1500 ± .0390 | .2529 ± .0732 | .1261 ± .0336 | .2332 ± .1078 |
| | | Death | .0220 ± .0291 | .0679 ± .0258 | .0595 ± .0230 | .3640 ± .1746 | .0767 ± .0083 | .0884 ± .0308 |
| | | Recovered | .0153 ± .0121 | .0716 ± .0151 | .0558 ± .0181 | .2944 ± .1138 | .0753 ± .0081 | .0868 ± .0478 |
| | 14 | Infected | .1425 ± .0429 | .3130 ± .0836 | .1905 ± .0503 | .2931 ± .1078 | .1582 ± .0495 | .2783 ± .2245 |
| | | Death | .0375 ± .0647 | .1030 ± .0475 | .0792 ± .0302 | .3852 ± .2004 | .0625 ± .0152 | .1070 ± .0349 |
| | | Recovered | .0237 ± .0165 | .0853 ± .0176 | .0749 ± .0324 | .3100 ± .1569 | .0645 ± .0130 | .1048 ± .0630 |
| | 21 | Infected | .1943 ± .0634 | .3815 ± .1163 | .2271 ± .0732 | .3295 ± .1685 | .1953 ± .0606 | .3341 ± .4599 |
| | | Death | .0507 ± .0754 | .1441 ± .0644 | .0982 ± .0363 | .3959 ± .2150 | .0736 ± .0206 | .1249 ± .0386 |
| | | Recovered | .0327 ± .0200 | .0983 ± .0203 | .0938 ± .0496 | .3185 ± .1976 | .0766 ± .0169 | .1218 ± .0742 |
| | 28 | Infected | .2413 ± .1032 | .4391 ± .1498 | .2600 ± .1248 | .3663 ± .2823 | .2266 ± .0686 | .3776 ± .6538 |
| | | Death | .0642 ± .0817 | .1880 ± .0809 | .1163 ± .0409 | .4039 ± .2225 | .0876 ± .0254 | .1422 ± .0425 |
| | | Recovered | .0427 ± .0235 | .1115 ± .0243 | .1129 ± .0722 | .3277 ± .2455 | .0914 ± .0204 | .1374 ± .0801 |

# Q3. Scalability

Wall clock time vs. data stream length

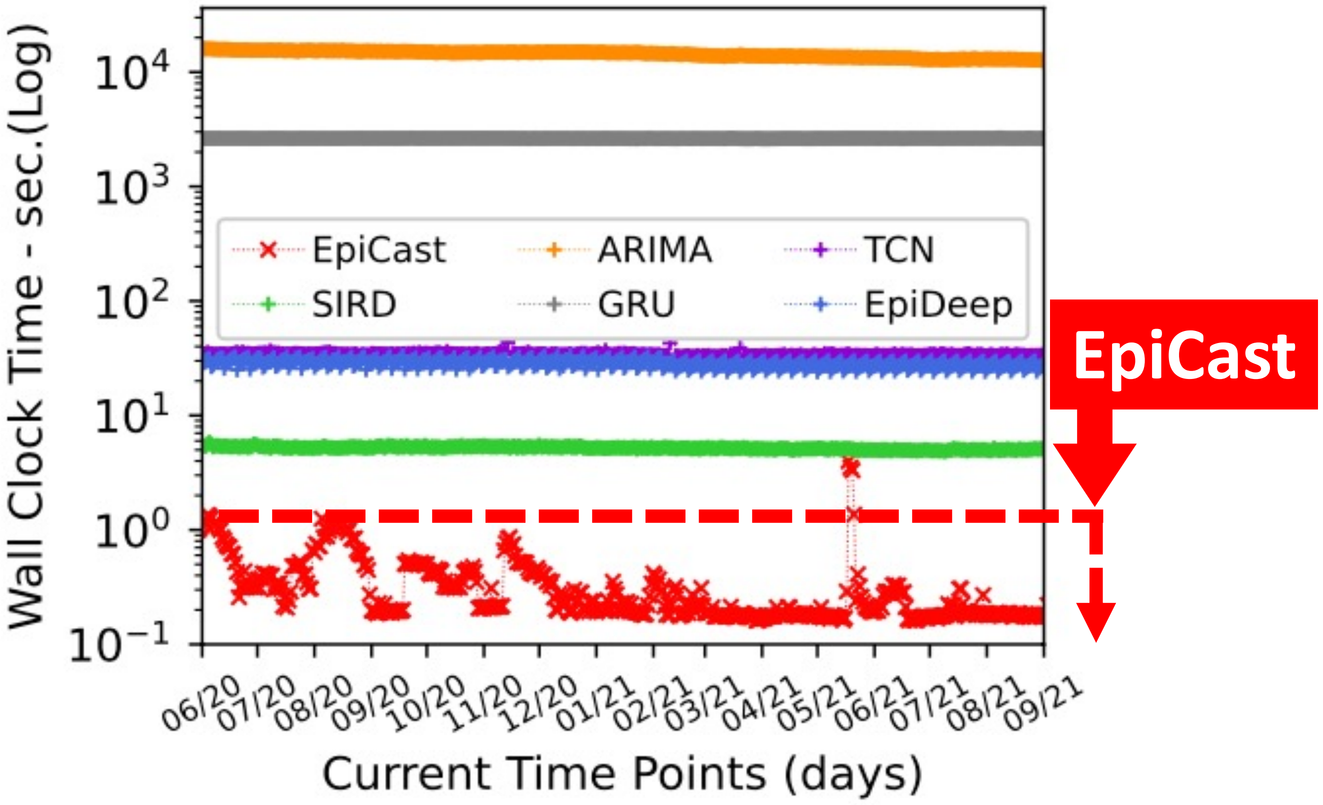



Wall clock time (average time consumption)

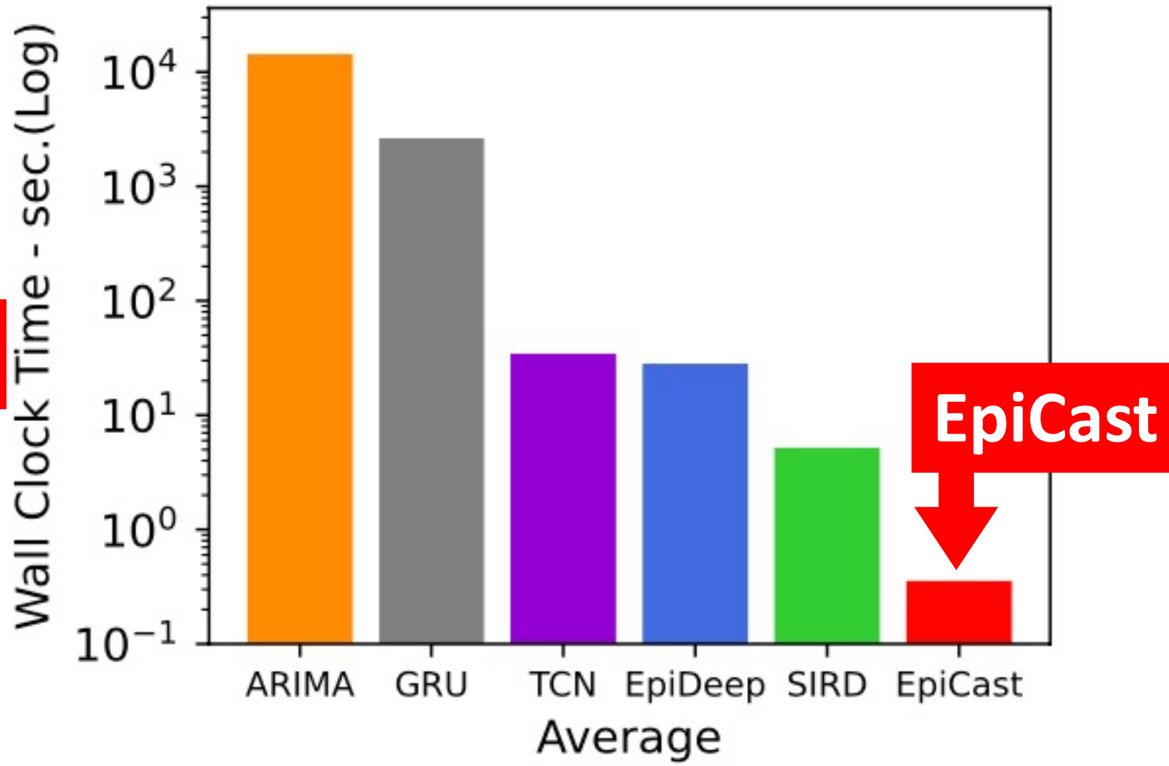


# Q3. Scalability

Wall clock time vs. data stream length




Wall clock time (average time consumption)



# Outline

~~Motivation~~

~~Modeling power of EpiCast~~

~~Problem definition~~

~~Proposed model~~

~~Algorithms~~

~~Experiments~~

▶ **Conclusions**

# Conclusions

EpiCast has the following properties:

**Effective**

It captures **important dynamical epidemic patterns** in data streams
And provides **long-range forecasting** at any time

**Adaptive**

It can **dynamically and adaptively capture current regimes**,
by sharing non-linear models among multiple locations

**Scalable**

It does **not depend on data stream size**

# Thank you for your kind attention!