Visual Analysis of Multi-scale Trends of COVID-19

Huan He*, Sijia Liu, Liwei Wang, Andrew Wen, Ming Huang, Yanshan Wang, Hongfang Liu**

Division of Digital Health Sciences, Department of Health Sciences Research, Mayo Clinic

* He.Huan@mayo.edu
** Liu.Hongfang@mayo.edu

Outline

- Background

- Task Analysis and Design Requirements

- Data Abstraction

- System Architecture

- Visual Designs

- Case Studies

- Discussion

- Summary

Background

- COVID-19

In 2019, a new coronavirus was identified as the cause of a disease outbreak. The virus is now known as the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).

The disease it causes is called coronavirus disease 2019 (COVID-19). In March 2020, the World Health Organization (WHO) declared the COVID-19 outbreak a **pandemic**.
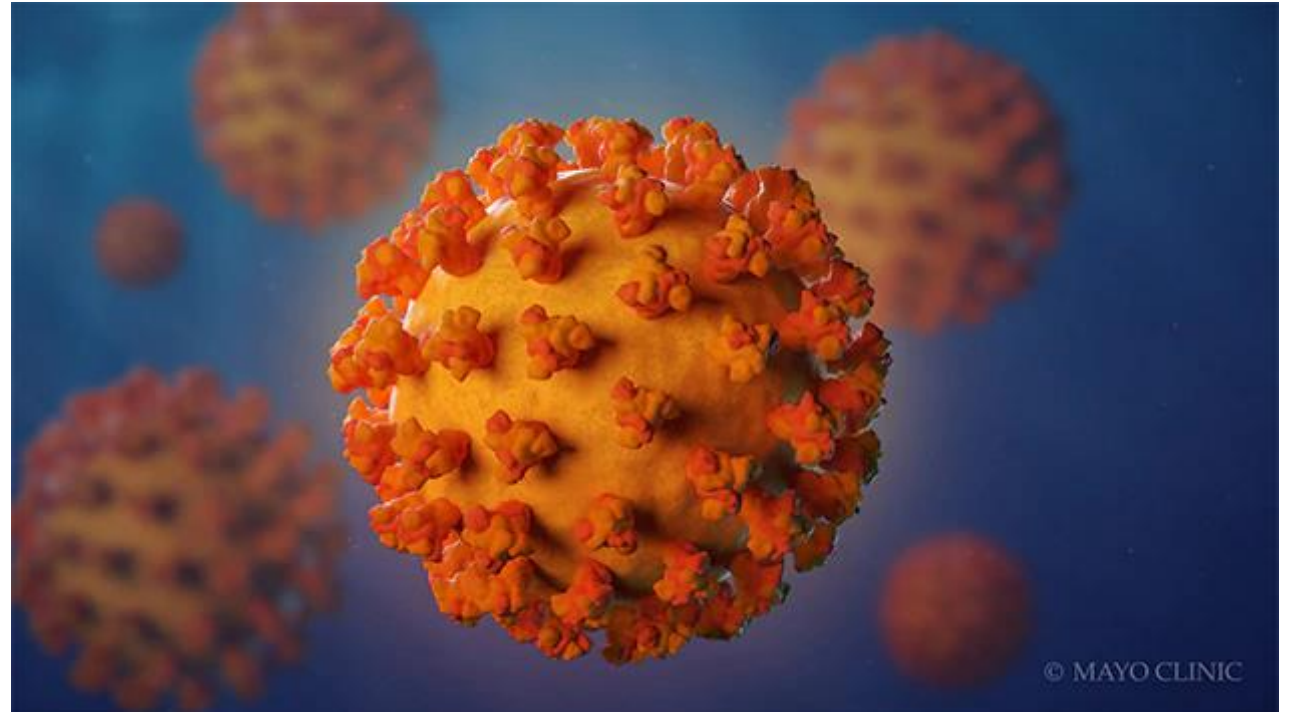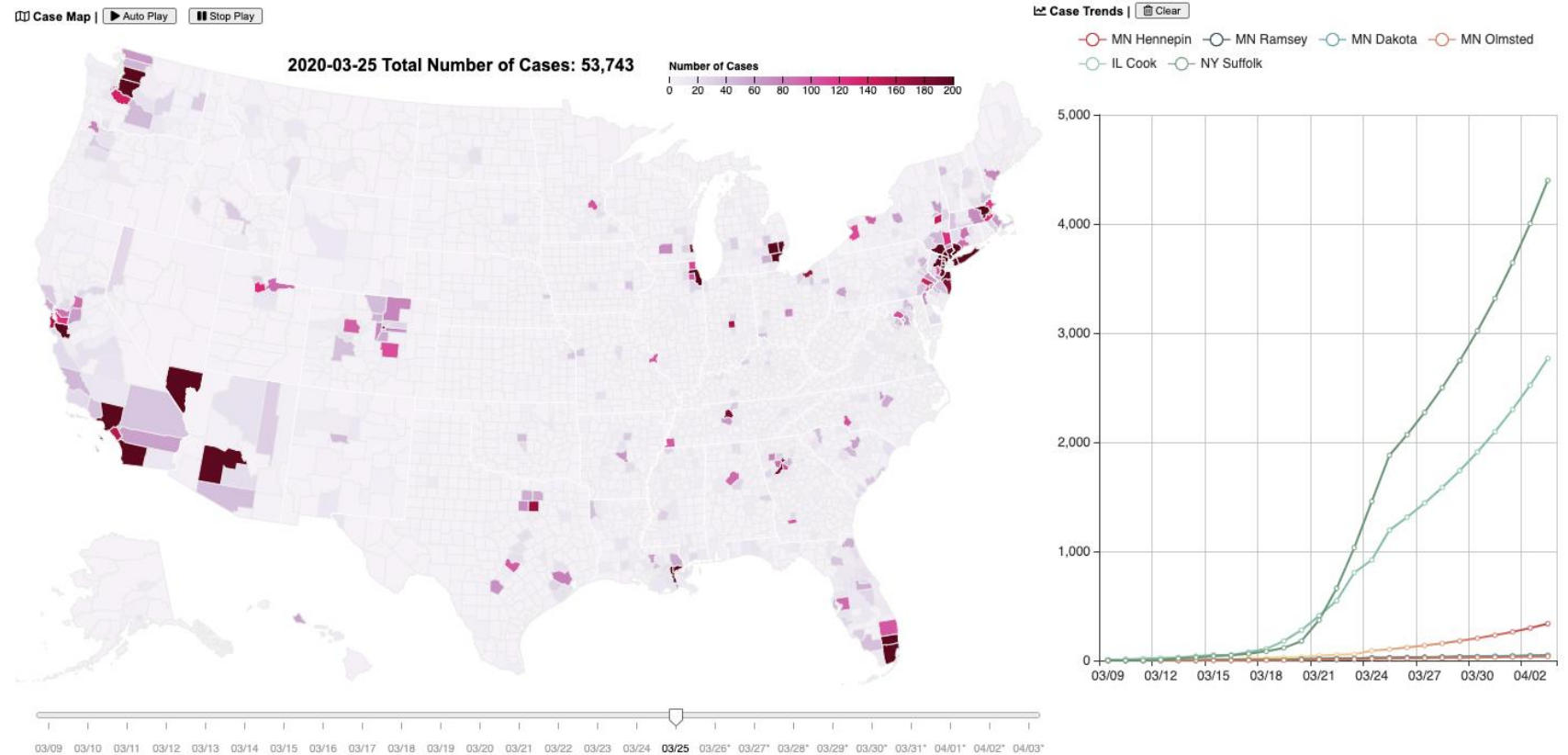
As the pandemic spreads, the public and researchers became more concerned about the potential trends of the COVID-19 spreading in their local communities.

Task Analysis – Evolving Needs

- ## Show the data

Prototype system for monitoring the total cases in each county and state

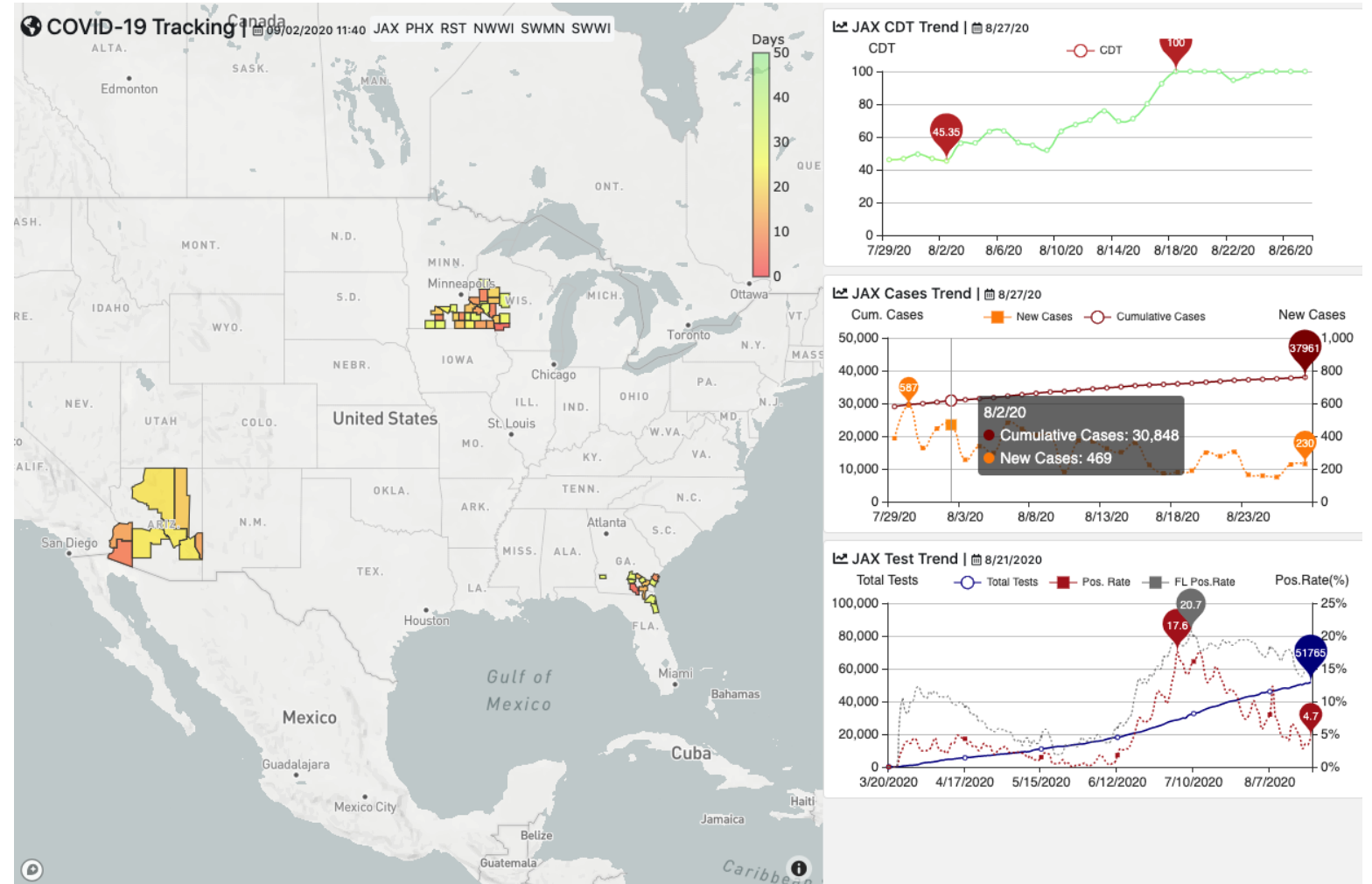- Choropleth map
- Line chart

The screenshot of the first prototype system showing the county-level distribution of the confirmed cases

Task Analysis – Evolving Needs

- ## Show more data

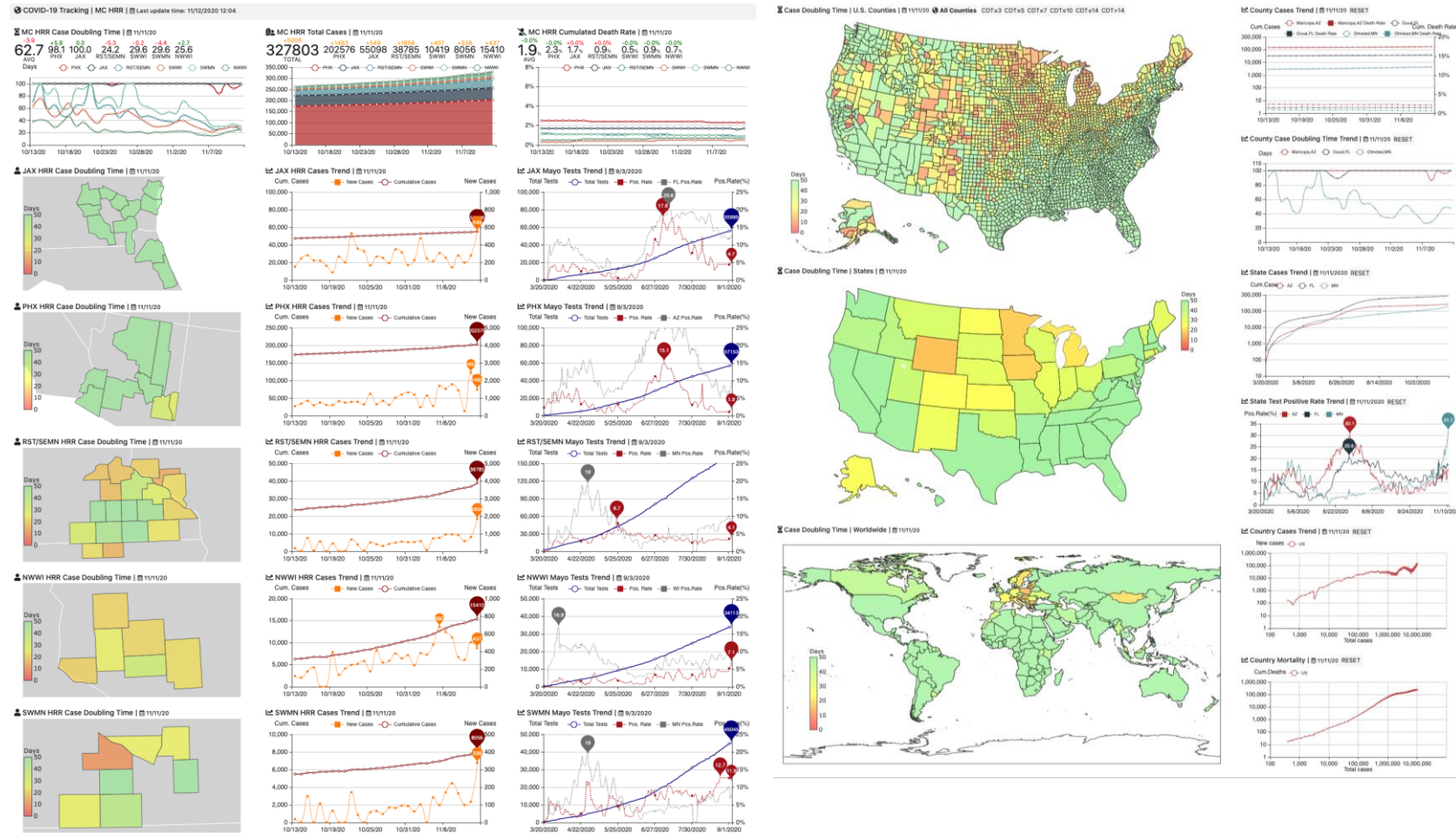Prototype system for showing more aspects of the regional situation

- New indicators
- More line charts

The screenshot of the second prototype system showing more indicators of a specific region

Task Analysis – Evolving Needs

- ## Show more regions

Prototype system for combining both the regional and global situation
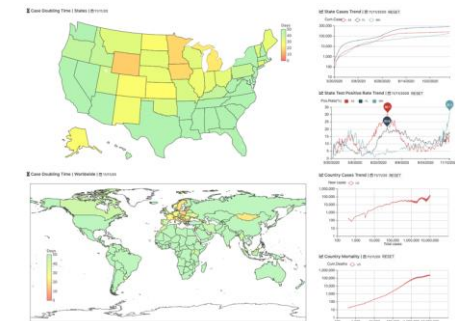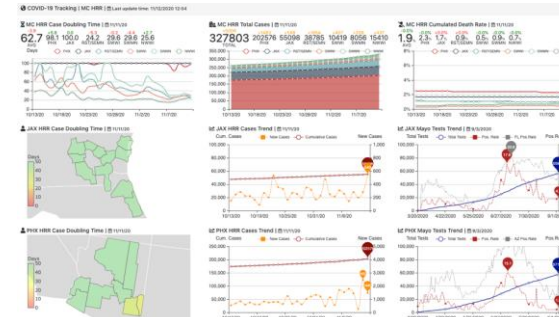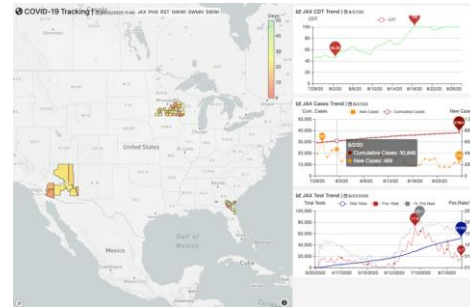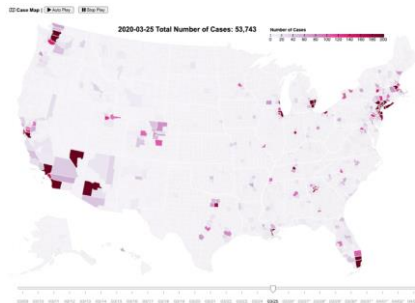
- Dashboard

- Regional panels

The screenshot of the third prototype system showing more regions for comparison

Task Analysis – Tasks and Design Requirements

- T.1 What is the **geographic distribution** of the COVID-19 pandemic?

- T.2 How does the regional COVID-19 pandemic **change over time**?

- T.3 What is the **regional difference** between COVID-19 outbreaks?

- R.1 **Multi-scale exploration** of regional COVID-19 pandemic state.

- R.2 Exploration of **temporal trends** of regional COVID-19 pandemic state.

- R.3 Exploration among **different regions**.

- R.4 **Interactive** exploration.

Data Abstraction

$$CDT_d = 4 \times \frac{log(2)}{log(N_d + 0.5)/N_{d-4}}$$

- Indicators
  - Case Doubling Time (CDT)

    The number of days taken for the number of coronavirus cases to double.

  - Cr7d100k

    7-day smoothed average daily case rate per 100k capita

  - RW_Cr7d100k

    the ratio of this week's Cr7d100k comparing to the week before

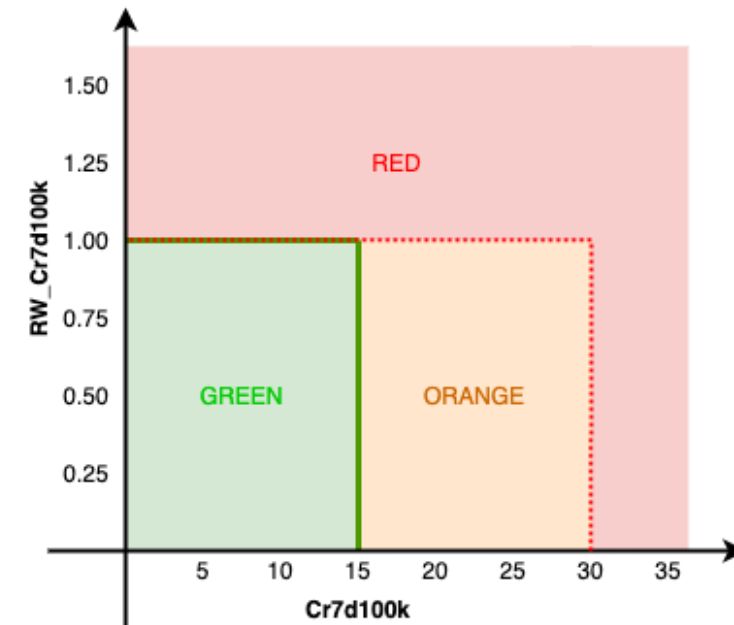 - CrRW Status

    GREEN: Cr7d100k<15 and RW<1 for 7 days

    RED: Cr7d100k>30 or RW>1 for 7 days

    ORANGE: everything else

$$Cr7d100k_d = \frac{1}{7} \times \frac{100,000}{Population_{region}} \times \sum_{i=d-7}^{d} n_i$$

$$RW\_Cr7d100k_d = \frac{Cr7d100k_d}{Cr7d100k_{d-7}}$$

System Architecture – Data Pipeline

- Data sources
  - USAFacts
  - COVID-19 Tracking Project
  - CDC COVID Data Tracker
  - Johns Hopkins Coronavavirus Resource Center

- Data watcher

- Data pre-processing

- Data calculation

- Output multi-scale regional JSON files
  - County-level
  - State-level
  - Country-level
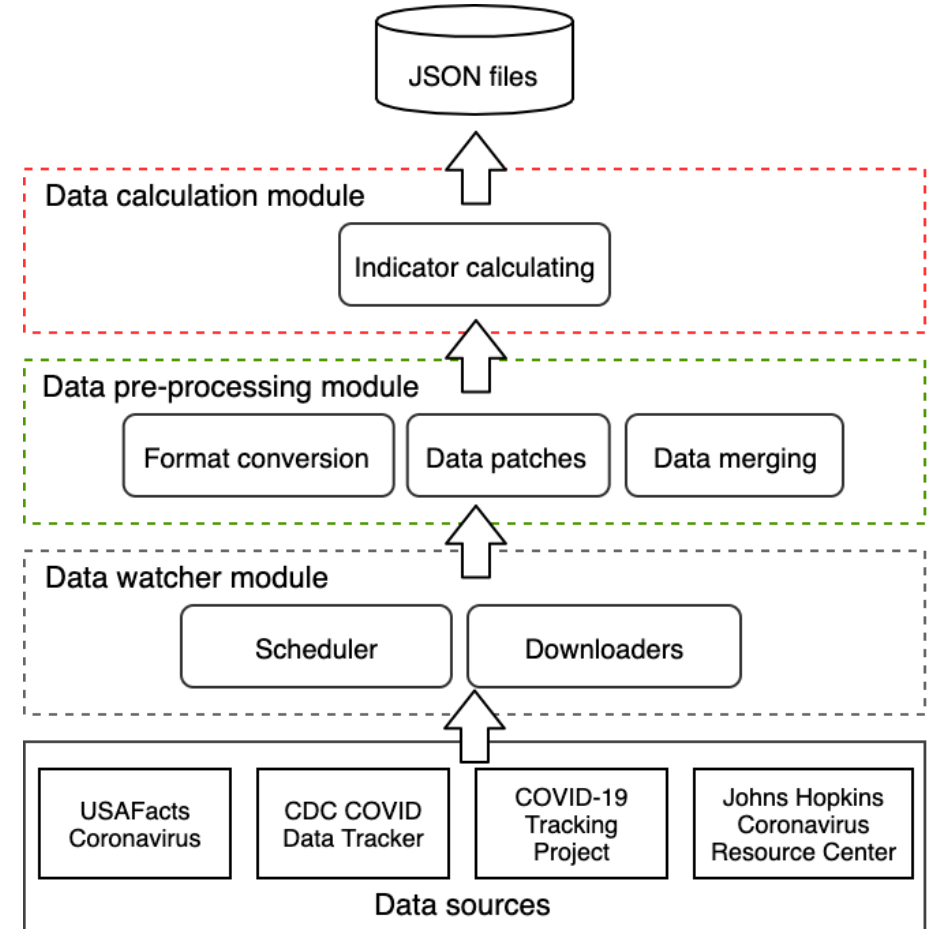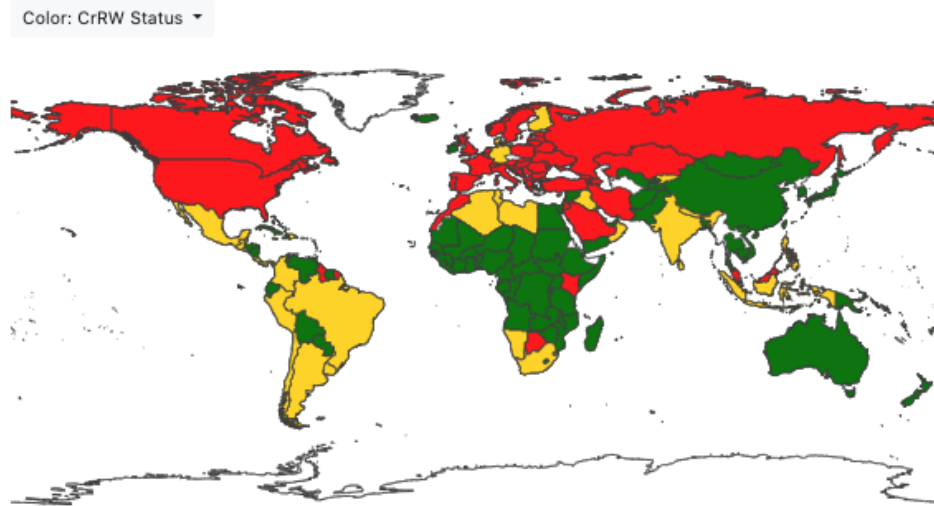
System Architecture

- Data Pipeline ➡ Visualization

COVID-19 Dashboard [1]   COVID-19 Map [2]

Visualization Module

Based on Vue.js, Plotly, Mapbox, and ECharts

[1] https://ohnlp.github.io/covid19tracking/
[2] https://www.mayoclinic.org/coronavirus-covid-19/map
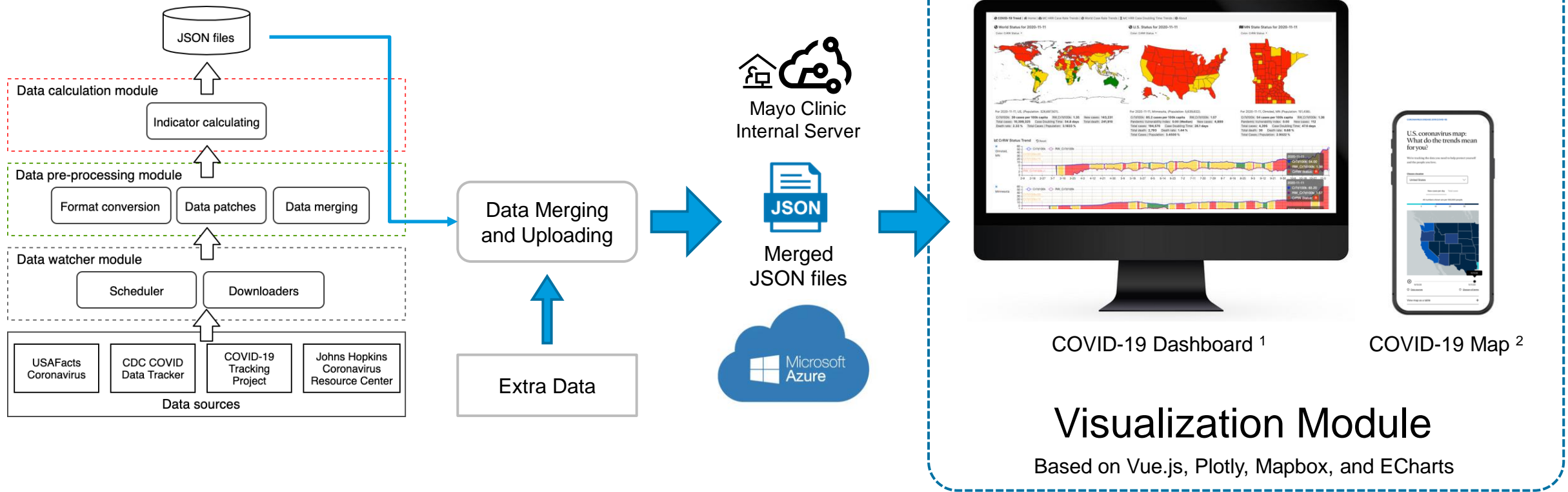
Visual Design – Map View

- Three maps

World-level   State-level   County-level
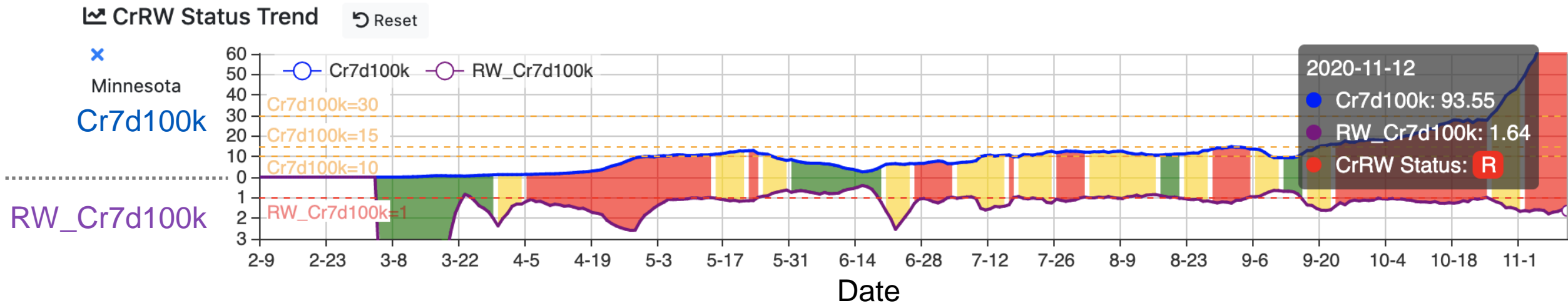
Visual Design – Map View

- Color encoding
  - CrRW Status
  - CDT
  - Total Cases / Population
  - CDC PVI *

- Interactive designs
  - Hover and click events
  - Linked map
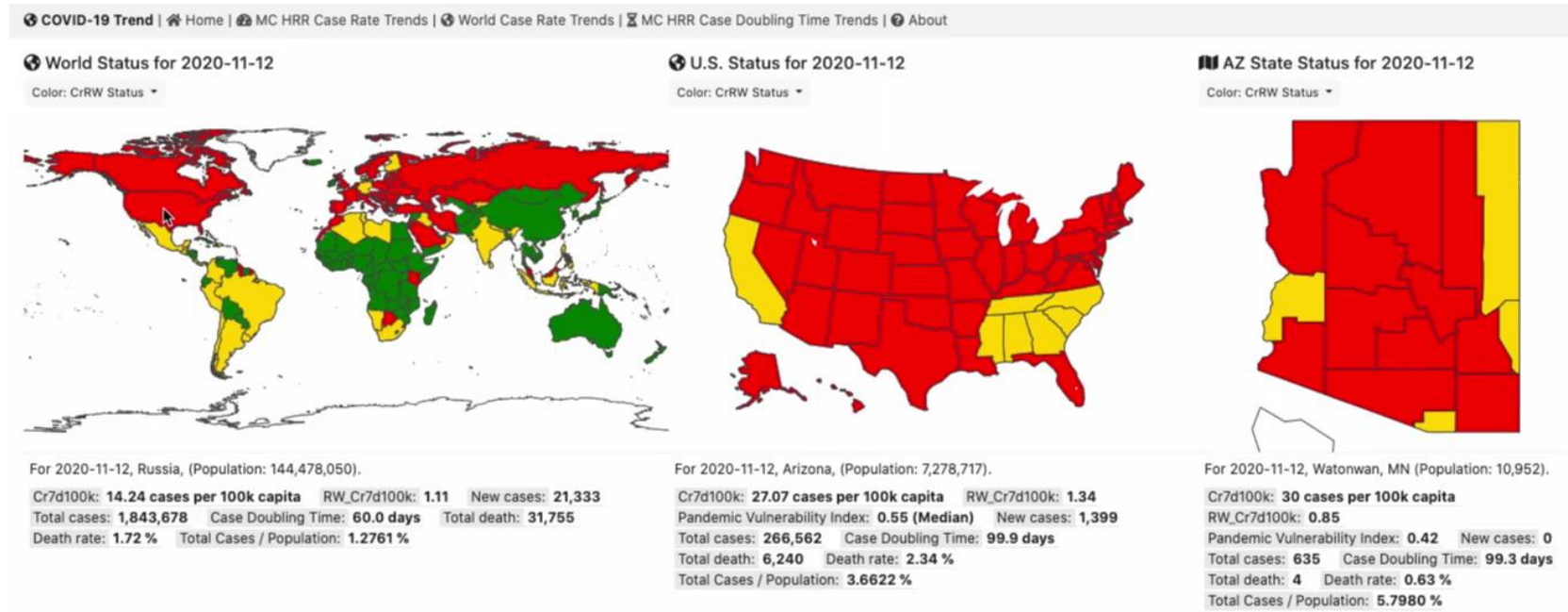
Visual Design – Trend View

- Trend chart
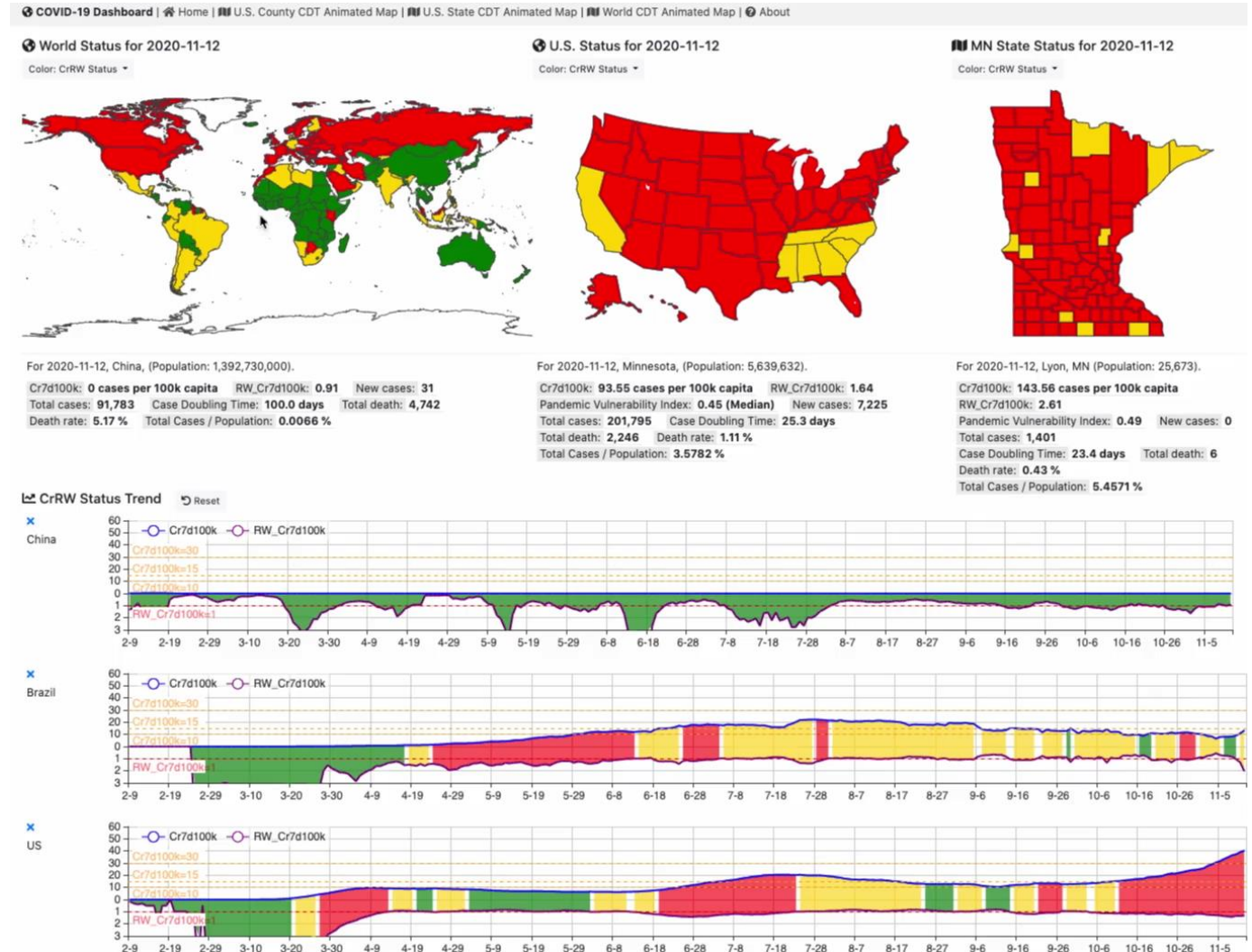 - How to show the changes of the multiple indicators over time intuitively?

- Combining line chart and area chart shows the pandemic
- Show the full history of the regional Cr7d100k, RW_7d100k, and CrRW status at the same time
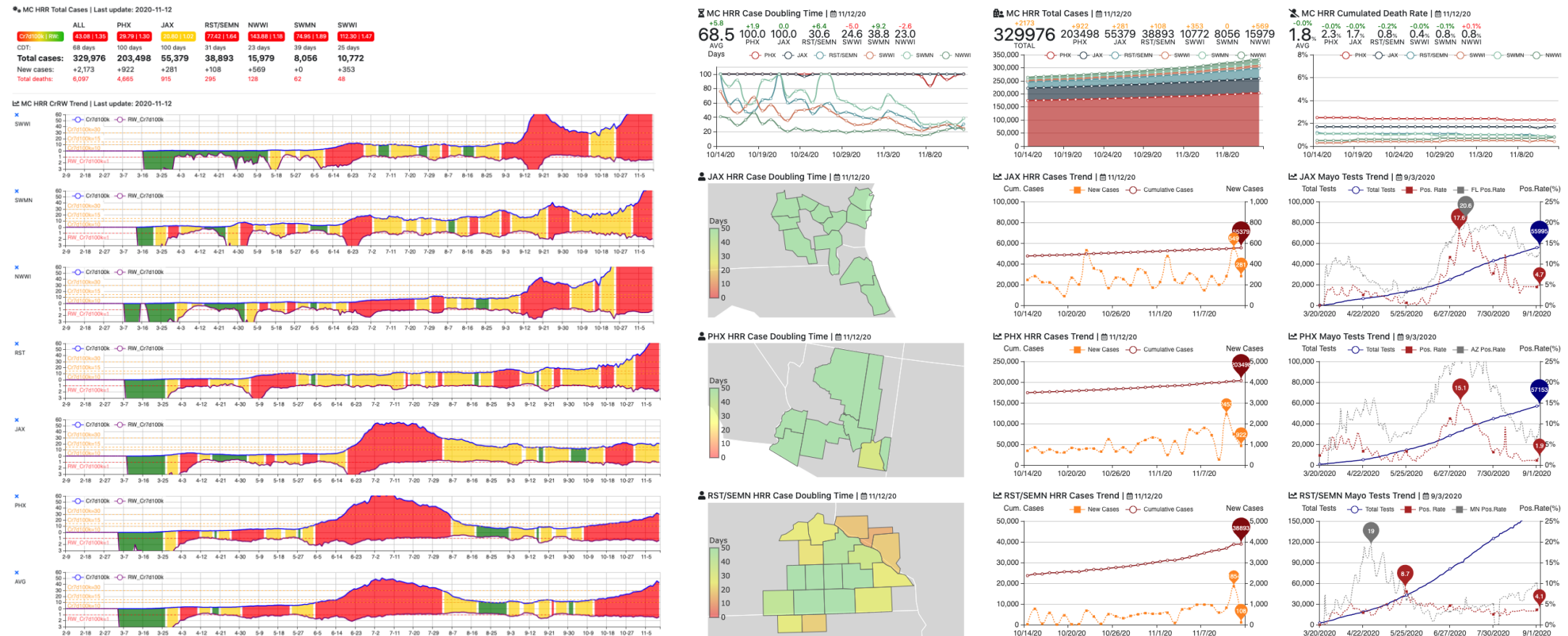
Visual Design – Trend View

- Color encoding
  - CrRW status

- Interactive designs
  - Linked with the map view
 - Hover events
  - All trend charts are linked

Visual Design – Region Summary View

- Combined trend view and other regional information

Case Study 1 – Geographical distribution

- ## January to February
  - a few counties on the west coast etc.

- ## March to May
  - Whole country

- ## June to August
  - South regions

- ## September to October
  - Spreading in north regions
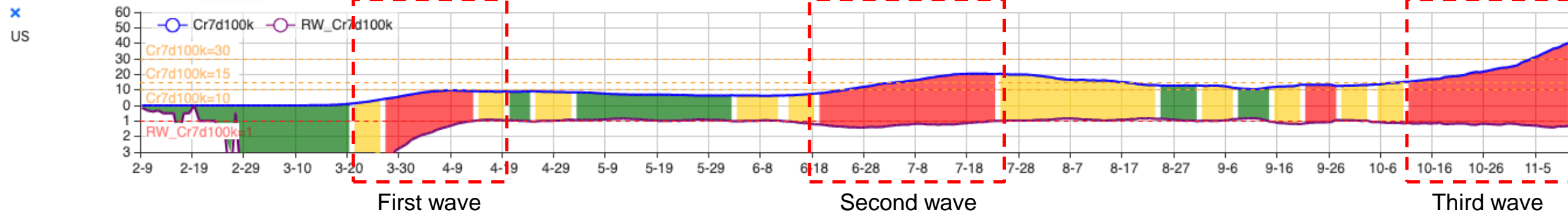
- ## November to now
  - Centre and north regions
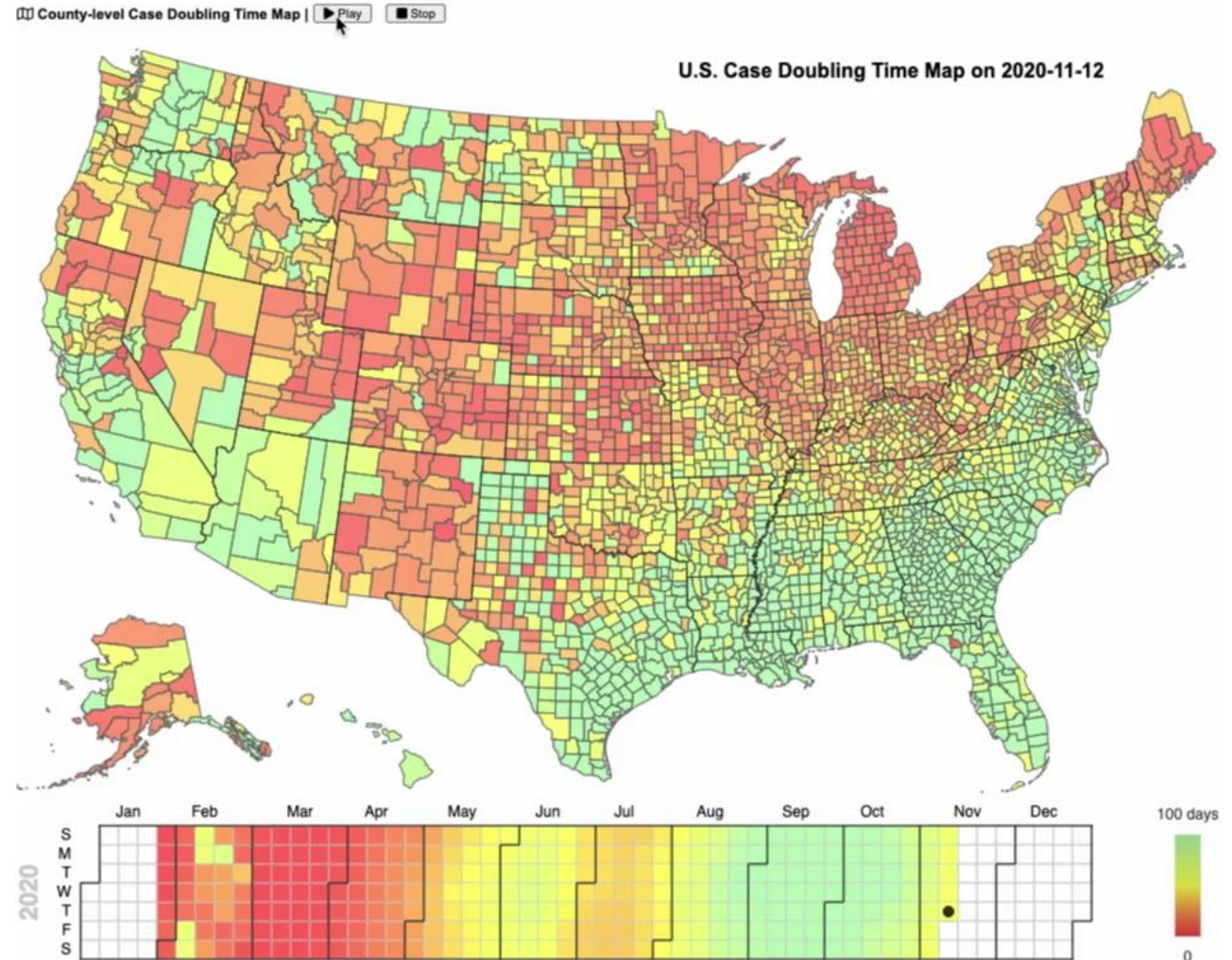
Case Study 2 – Temporal patterns
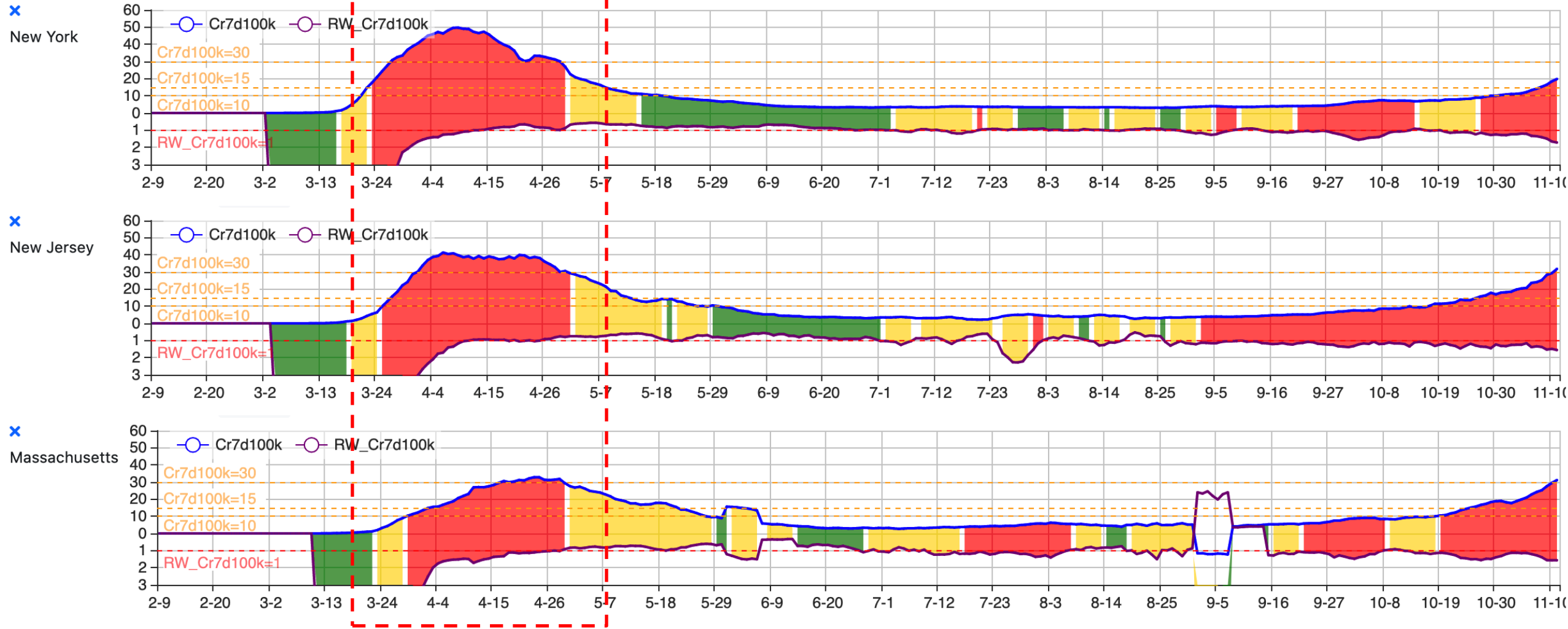
- Overall stable: few states

- Three waves:

First wave   Second wave   Third wave
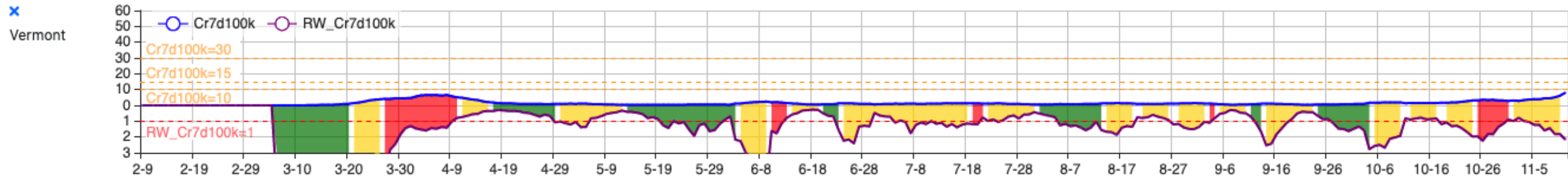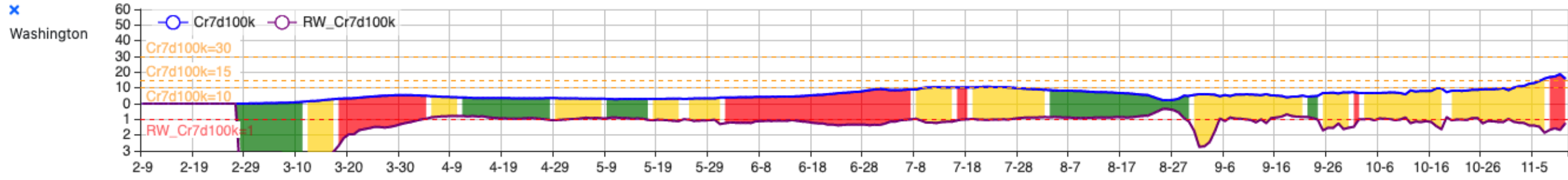
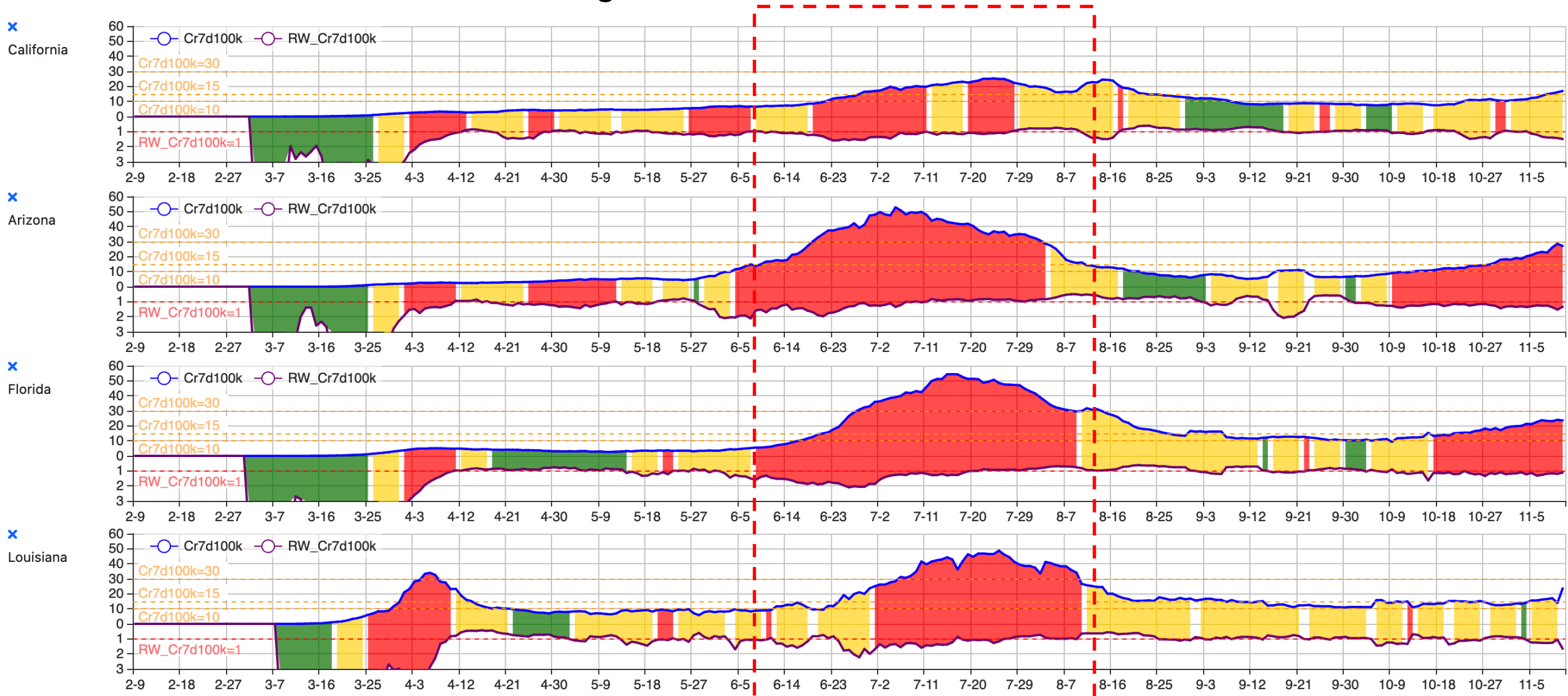# Case Study 2 – Temporal patterns: three waves

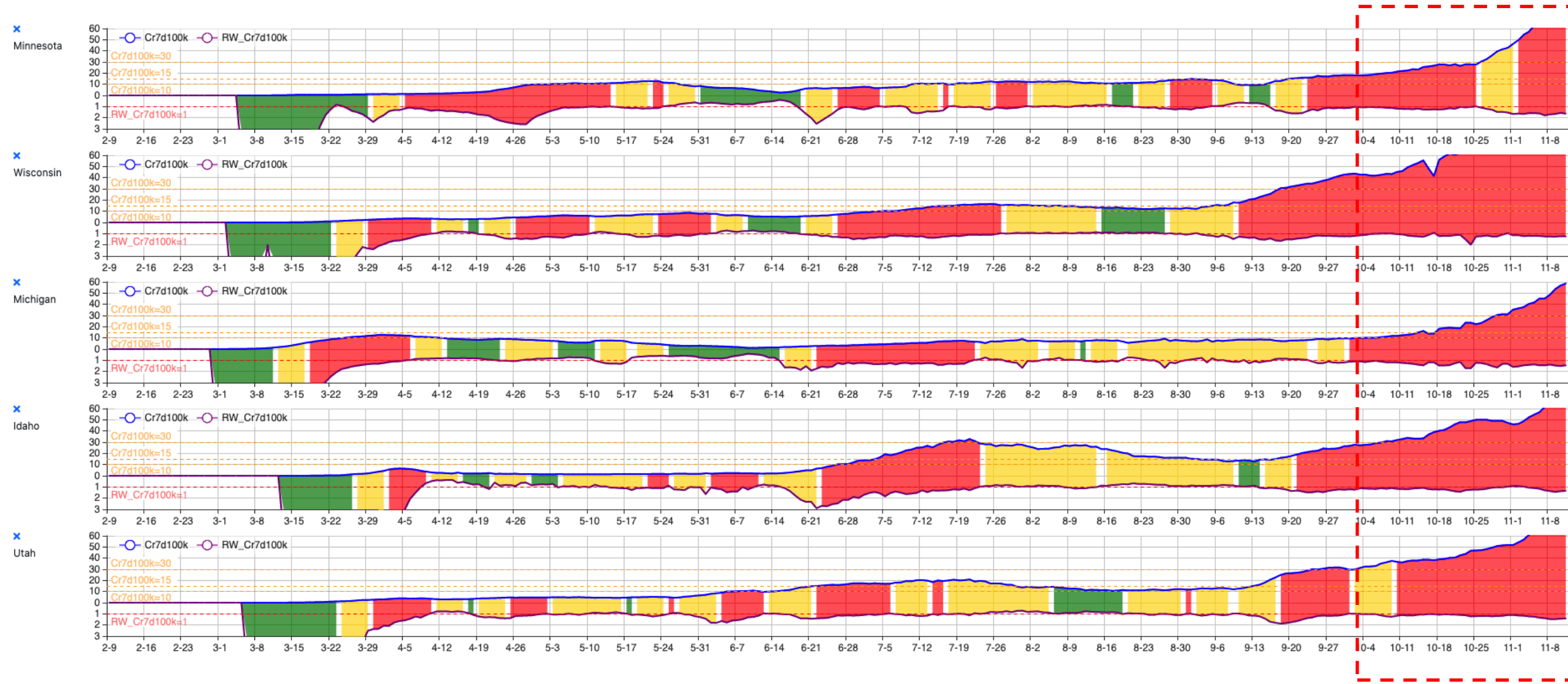- First wave: March to April, a few states

Case Study 2 – Temporal patterns: three waves

- Second wave: June to August, more states

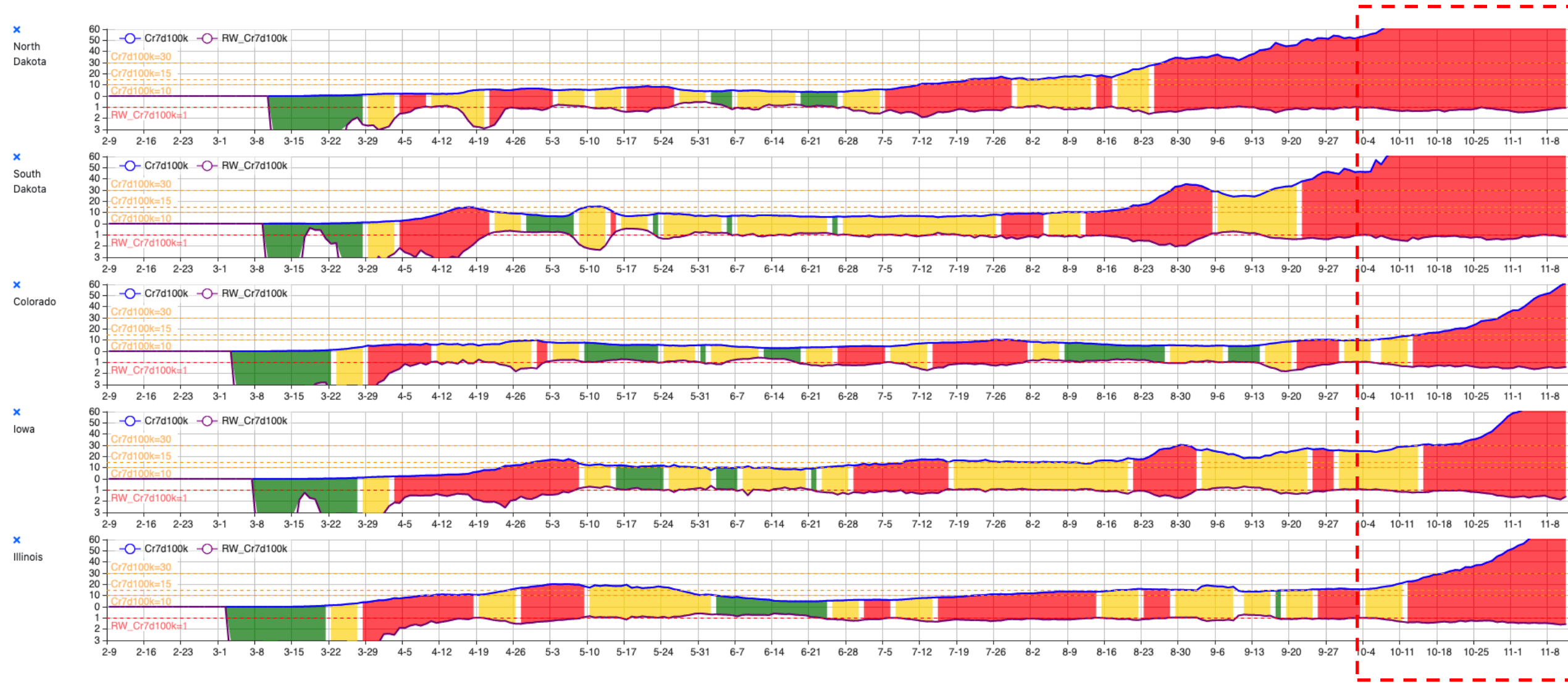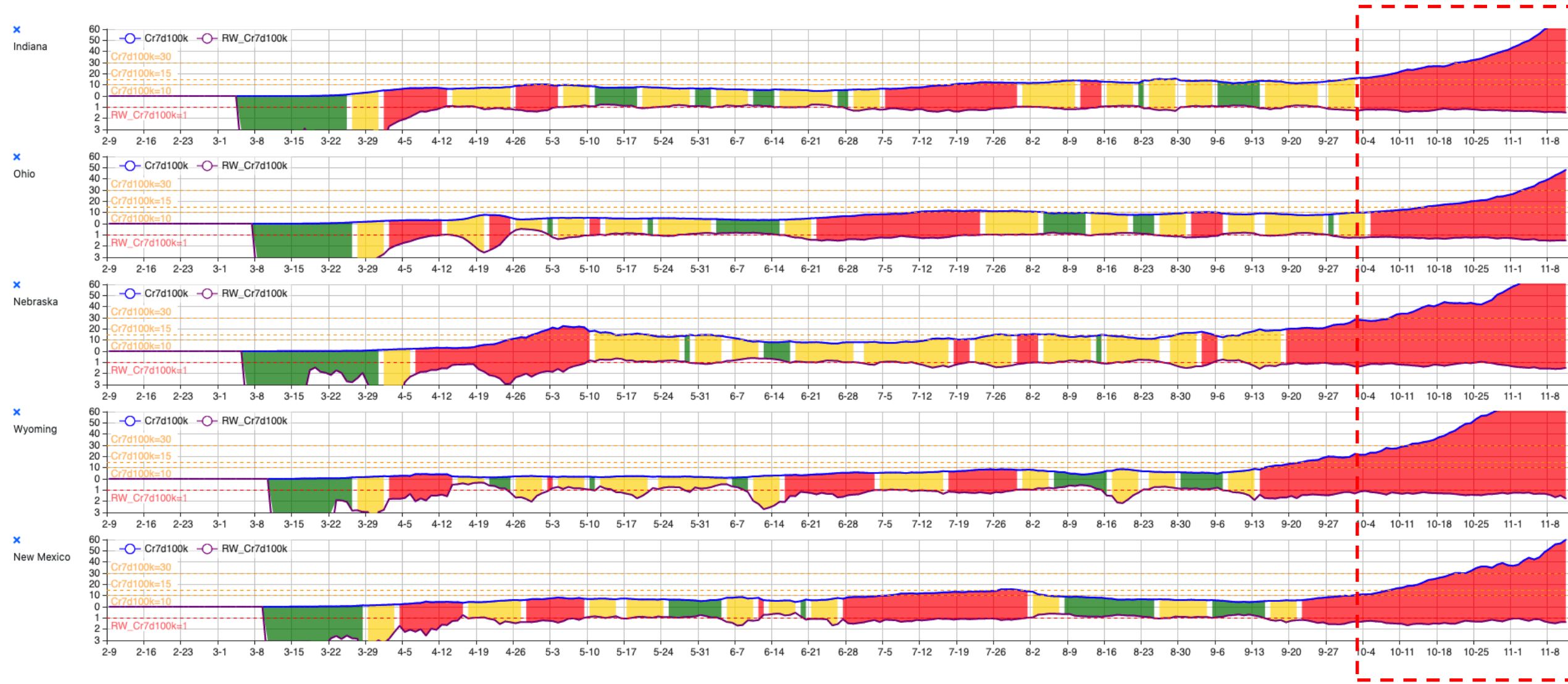# Case Study 2 – Temporal patterns: three waves

- Third wave: October to now, a lot of states

Case Study 2 – Temporal patterns: three waves

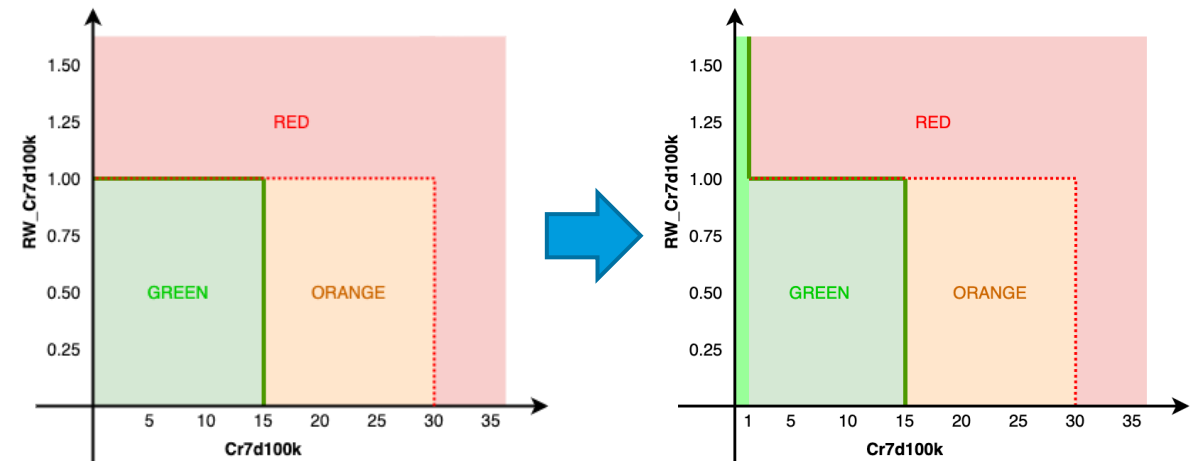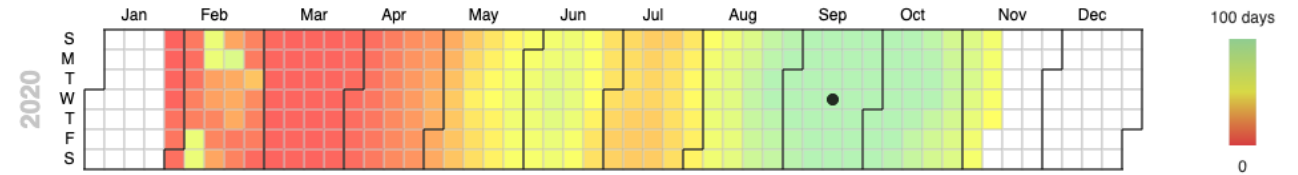- Third wave: October to now, a lot of states

Case Study 2 – Temporal patterns: three waves

- Third wave: October to now, a lot of states

Discussion

- CDT is sensitive in early stage
  - Low CDT indicates potential outbreak, but high CDT doesn't mean it's safe

- CrRW status
  - Sensitive to the RW_Cr7d100k
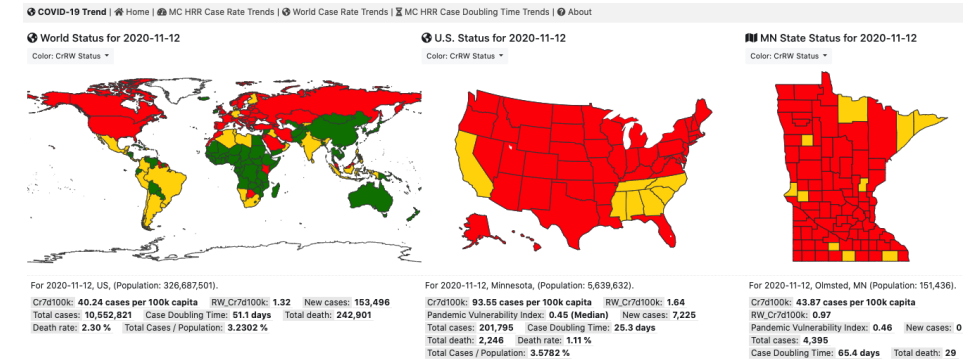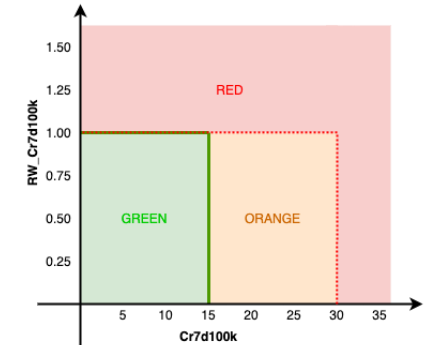 - Thresholds may need to be adjusted

Summary

- CrRW status for representing the current status of the pandemic as well as recent trends

- Trend chart design and other views for analyzing the regional pandemic

- Two case studies that show the distribution and the waves

Stay healthy, stay safe :)
Thank you for watching!

Questions?
He.Huan@mayo.edu
Liu.Hongfang@mayo.edu

Source code and online live demo:
https://ohnlp.github.io/covid19tracking/

$$Cr7d100k_d = \frac{1}{7} \times \frac{100,000}{Population_{region}} \times \sum_{i=d-7}^{d} n_i$$

$$RW\_Cr7d100k_d = \frac{Cr7d100k_d}{Cr7d100k_{d-7}}$$