Supplementary Information: A data-driven semiparametric model of SARS-CoV-2 transmission dynamics in the United States

2022-08-05

John M. Drake, Andreas Handel, Éric Marty, Eamon B. O'Dea, Tierney O'Sullivan, Giovanni Righi, and Andrew T. Tredennick,

Model diagram

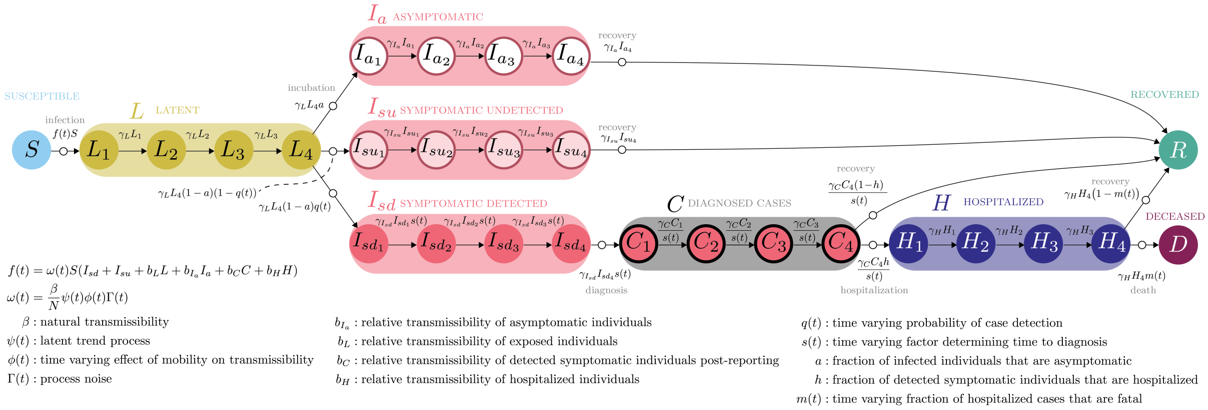


Figure S1: Model diagram.

Model parameters

Table S1: Fixed model parameters.

Parameter definition	Parameter symbol	Value	Source
Reproduction number	$\mathcal{R}_0$	7	Sanche et al. (2020)
Baseline transmission rate of symptomatic individuals	β	$\mathcal{R}_0 \times 0.1 \times \frac{1}{N}$	Assumption
Fraction of infected individuals that are asymptomatic	a	0.18	Mizumoto et al. (2020);
Fraction of diagnosed cases that are hospitalized	h	0.12	Verity et al. (2020)
Relative transmissibility of L to $I_{su/sd}$	b_L	0.12	Li et al. (2020)
Relative transmissibility of I_a to $I_{su/sd}$	b_{I_a}	0.5	Li et al. (2020)
Relative transmissibility of C to $I_{su/sd}$	b_C	0.27	Assumption
Relative transmissibility of H to $I_{su/sd}$	b_H	4.5×10^{-5}	Moghadas et al. (2020)
Duration of time in L stages	$1/\gamma_L$	4 days	Li et al. (2020)
Duration of time in I_a stage	$1/\gamma_{I_a}$	3.5 days	Li et al. (2020)
Duration of time in I_{su} stages	$1/\gamma_{I_{su}}$	6 days	Li et al. (2020)
Duration of time in I_{sd} stage	$1/\gamma_{I_{sd}}$	$0.5 \times 1/\gamma_{I_{su}}$	Assumption
Duration of time in C stages	$1/\gamma_C$	$0.5 \times 1/\gamma_{I_{su}}$	Assumption
Duration of time in H stages	$1/\gamma_H$	6 days	Ohsfeldt et al. (2020)
Minimum detection probability	$q_{\min}$	0.1	Assumption
Maximum detection probability	$q_{\max}$	0.4	Assumption
Day at which detection probability is halfway between $q_{\min}$ and $q_{\max}$	q_{half}	30	Assumption
Rate of increase from $q_{\min}$ to $q_{\max}$	q_{r}	1.1	Assumption
Maximum factor by which diagnosis speed increases	$s_{\max}$	1.0	Zhang et al. (2020)
Day at which diagnosis speed-up factor (s) is halfway between 0 and $s_{\max}$	s_{half}	30	Assumption
Rate of increase from 0 to $s_{\max}$	s_r	1.1	Assumption
Initial size of susceptible pool	$S(t=1)$	N_{state}	Assumption

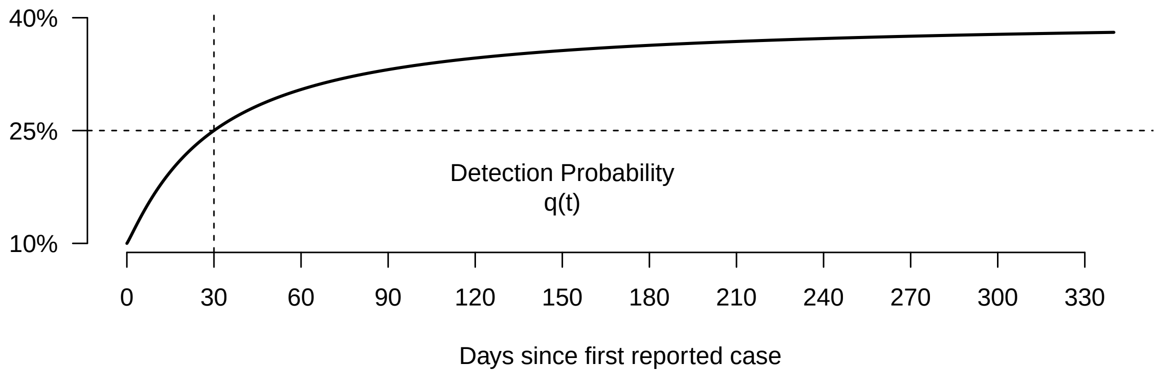
Table S2: Estimated parameters and starting ranges for MIF estimation procedure. The "expit" function refers to back-transforming the parameter from the logit scale, which was used for estimation.

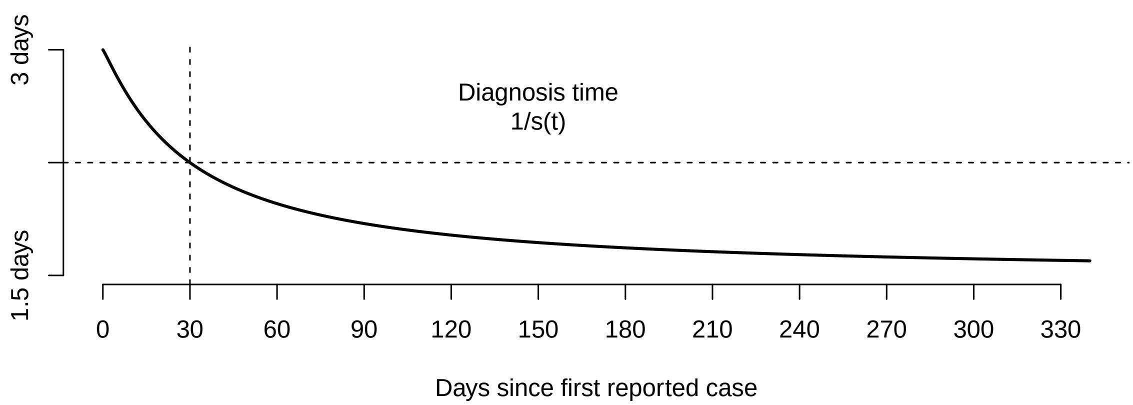
Parameter definition	Parameter symbol	Start range
Baseline fraction of hospitalizations that result in death	m_{base}	[expit(-6), expit(6)]
Minimum fraction of hospitalizations that result in death	m_{min}	[expit(-6), expit(6)]
Day at which death fraction is halfway between m_{base} and m_{min}	m_{half}	[exp(-5), exp(5)]
Extra-demogaphic process noise	σ	[exp(-5), exp(5)]
Case reporting dispersion	θ_c	[exp(-5), exp(5)]
Death reporting dispersion	θ_d	[exp(-5), exp(5)]
Initial size of latent compartment	$L(t=0)$	[exp(0), exp(10)]
Initial size of asymptomatic infectious compartment	$I_a(t=0)$	[exp(0), exp(10)]
Initial size of undetected infectious compartment	$I_{su}(t=0)$	[exp(0), exp(10)]
Initial size of detected infectious compartment	$I_{sd}(t=0)$	[exp(0), exp(10)]
B-spline coefficients	g_i	[-10, 10]

Detection probability and diagnosis speed-up functions

The fraction of exposed individuals that are detected and flow into the I_{sd} compartments (q) starts at a low level of 0.1 and increases sigmoidally towards a maximum of 0.4, reaching the half way point of 0.25 on the 30th day since the first case notification. The time to diagnosis ($1/s$) decreases sigmoidally over time from 3 days on day 0 towards a minimum of 1.5 days, reaching the half way point of 2.25 days on the 30th day since the first case notification.

Both q, s follow a Hill function with Hill coefficient 1.1 and time to half maximum of 30 days:





Mean absolute scaled errors (MASE)

The MASE is a ratio of mean absolute error (MAE) in the model to MAE of an in-sample naïve model (a random walk forecast from the last observation, adjusted for weekly seasonality). MASE values were calculated for each of the 500 particle filter replicates for each state using the **yardstick** package in R. Most states had a mean MASE less than one for cases, while MASE was greater than one for deaths for most states. Thus, our model does not always outperform the benchmarking model, but it does offer mechanistic insight that the random walk model does not.

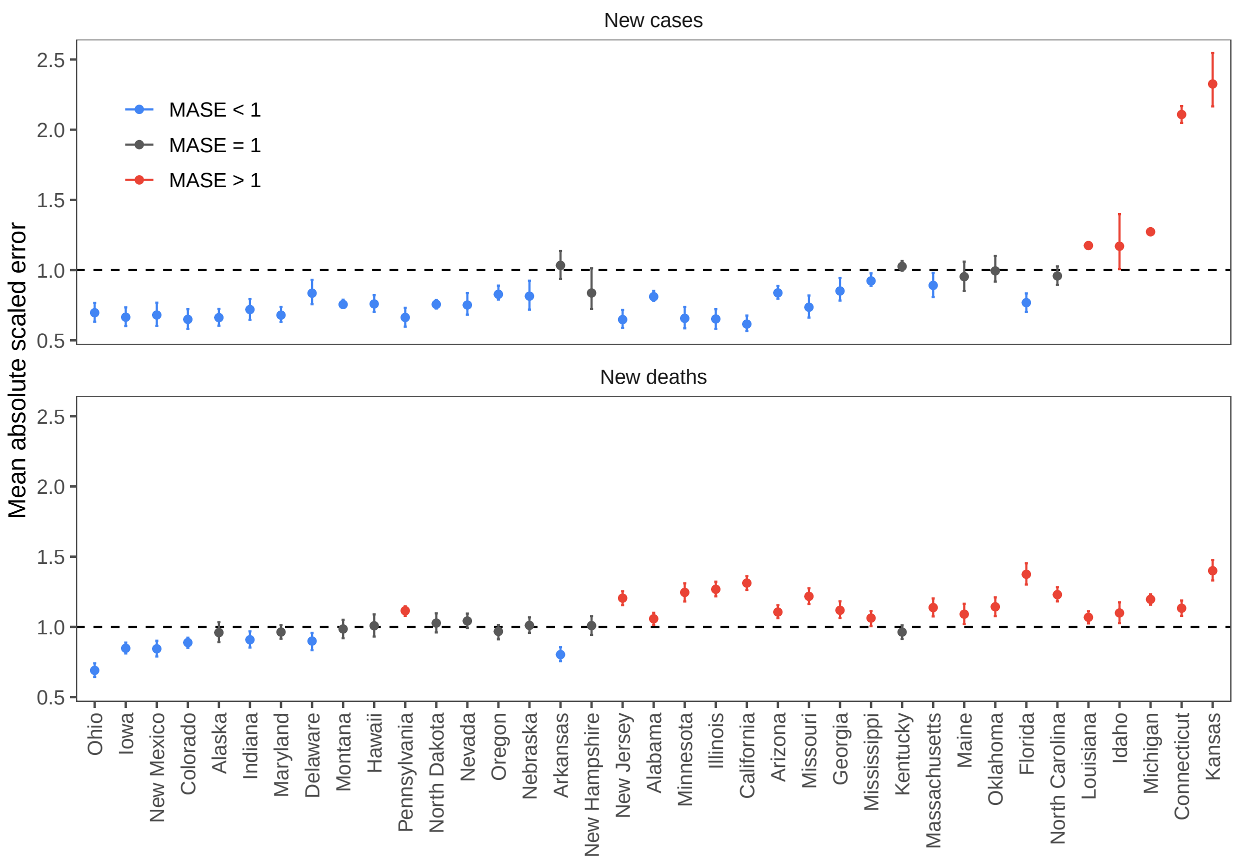
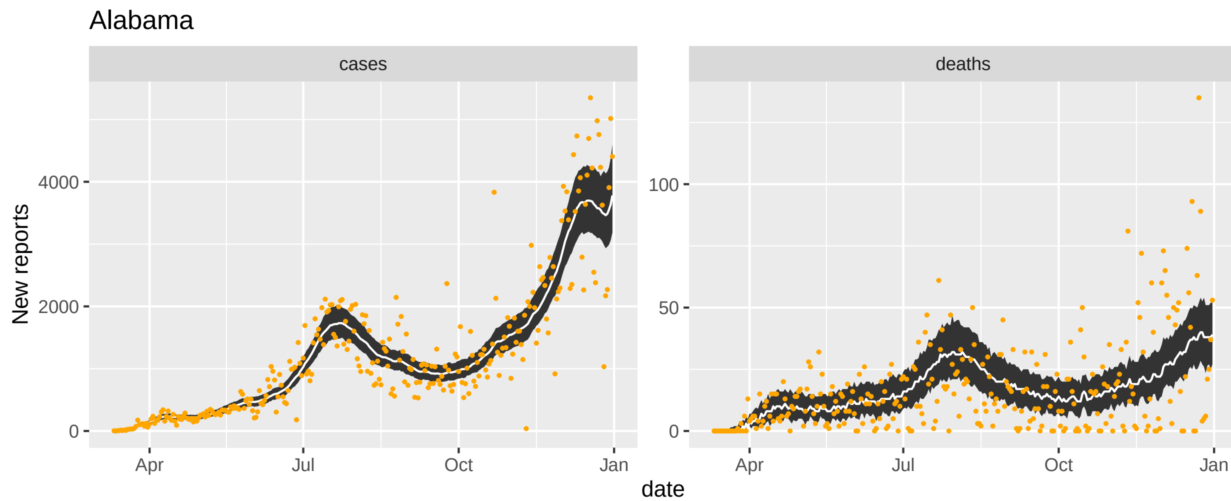
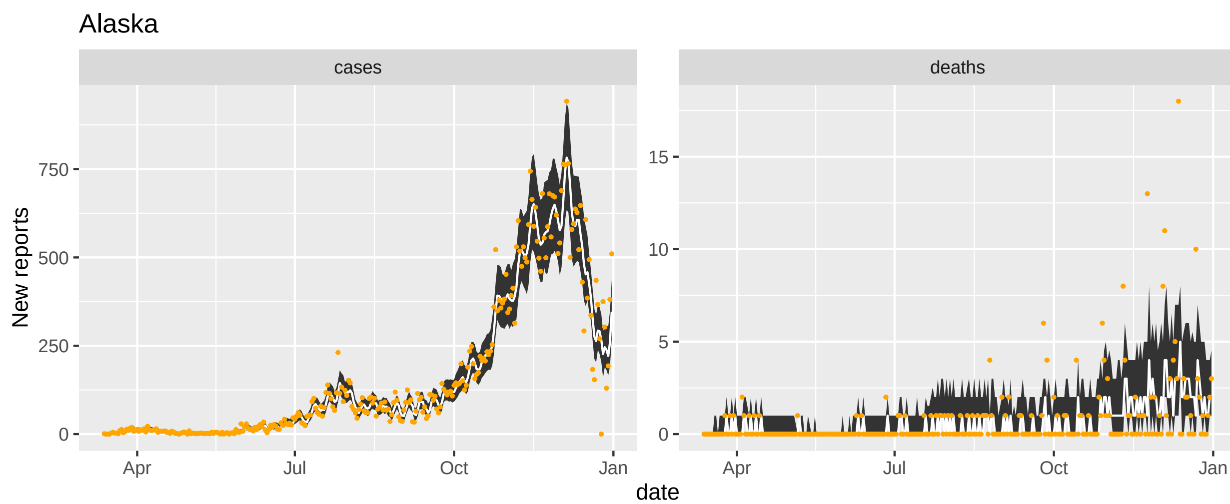


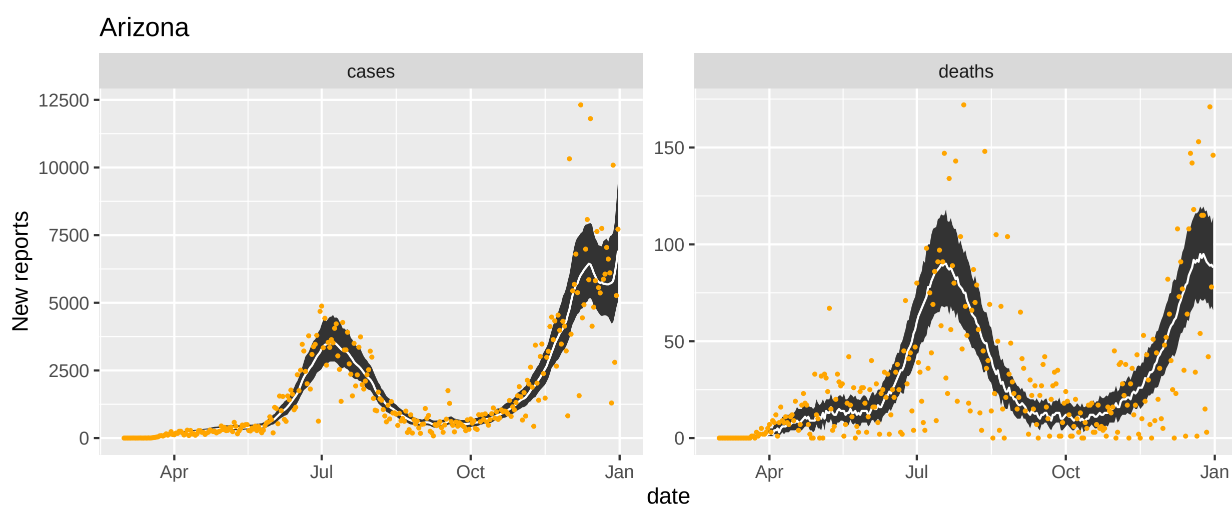
Figure S2: Mean and 95% intervals of mean absolute scaled errors (MASE) for each state. MASE was calculated for each of the 500 particle filter replicates for each state, yielding a distribution of MASE values. The point shows the mean of the MASE distribution and the errorbars show the 95% interval (lower 0.025 quantile to upper 0.975 quantile). States are ranked approximately from lowest MASE to highest MASE across both cases and deaths.

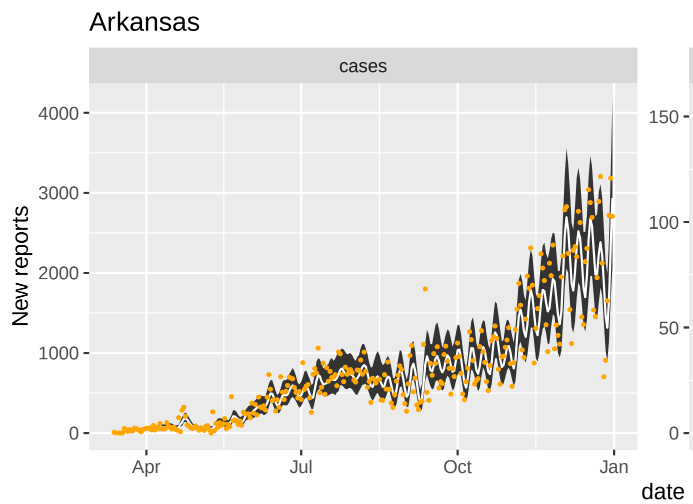
Time series of incident case and death reports

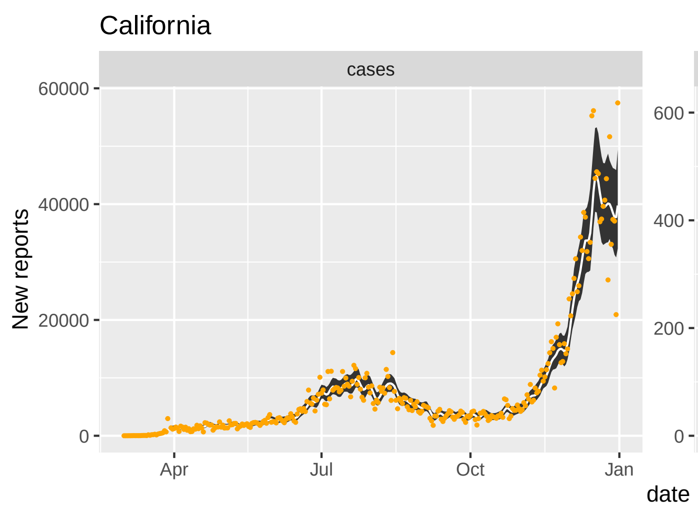
In the following plots, orange points are the reported daily cases or deaths, white line shows the median of the smoothed filtering distributions, and the grey ribbon bounds the 95% prediction interval.

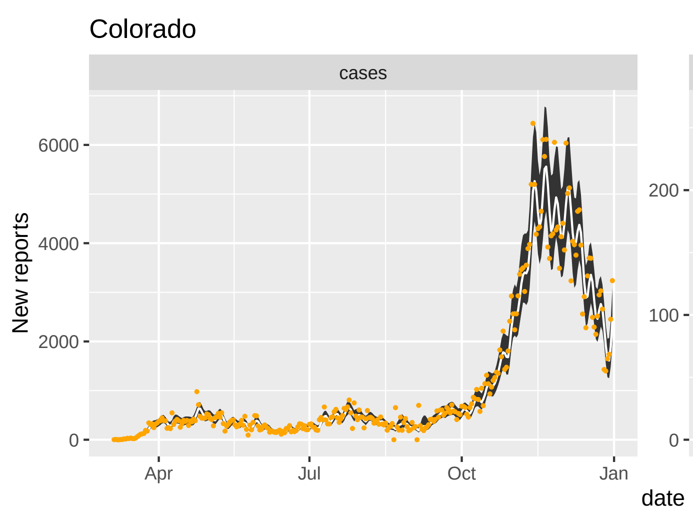


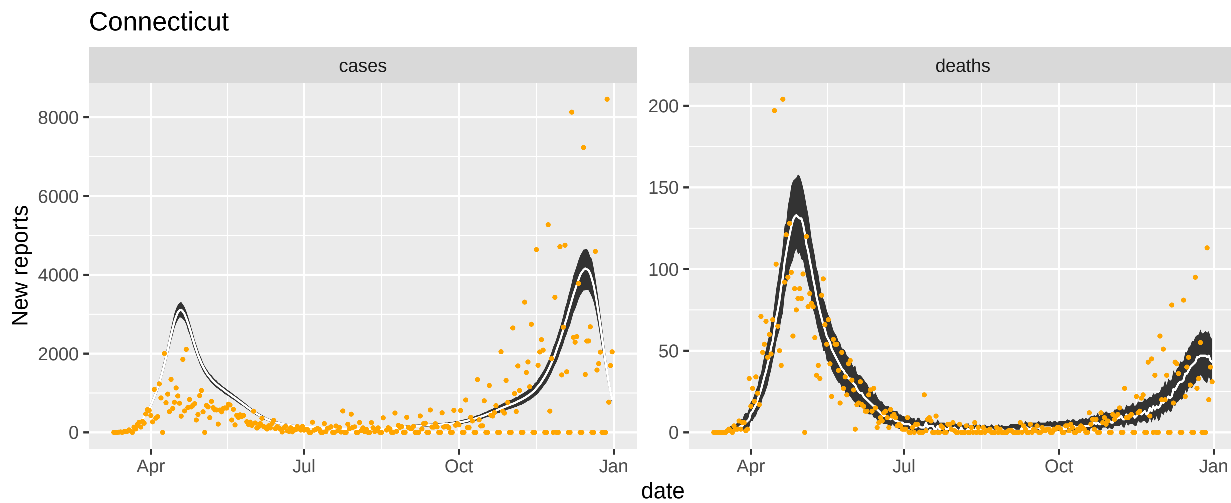


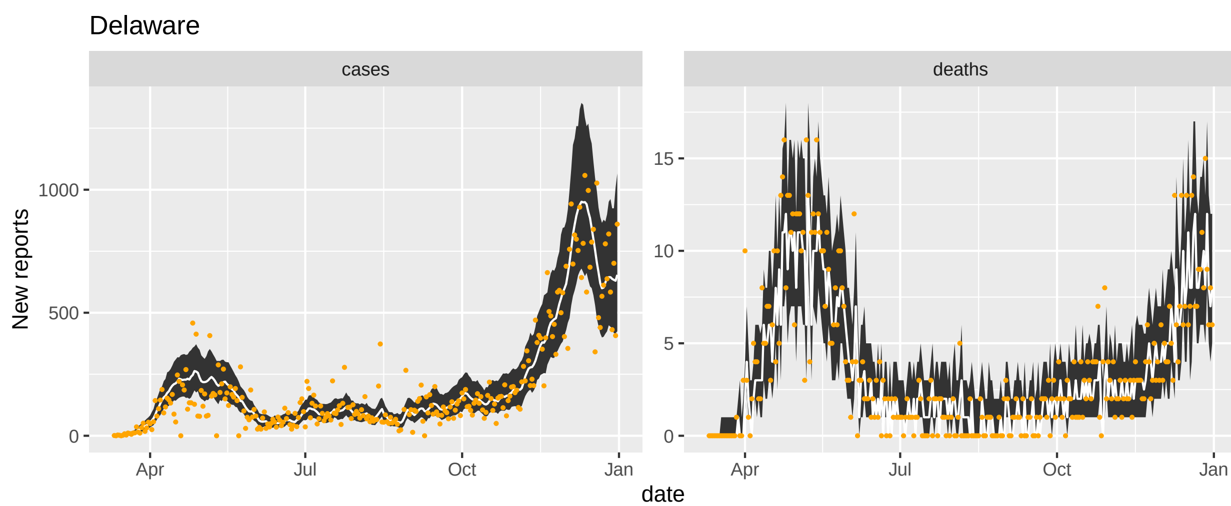


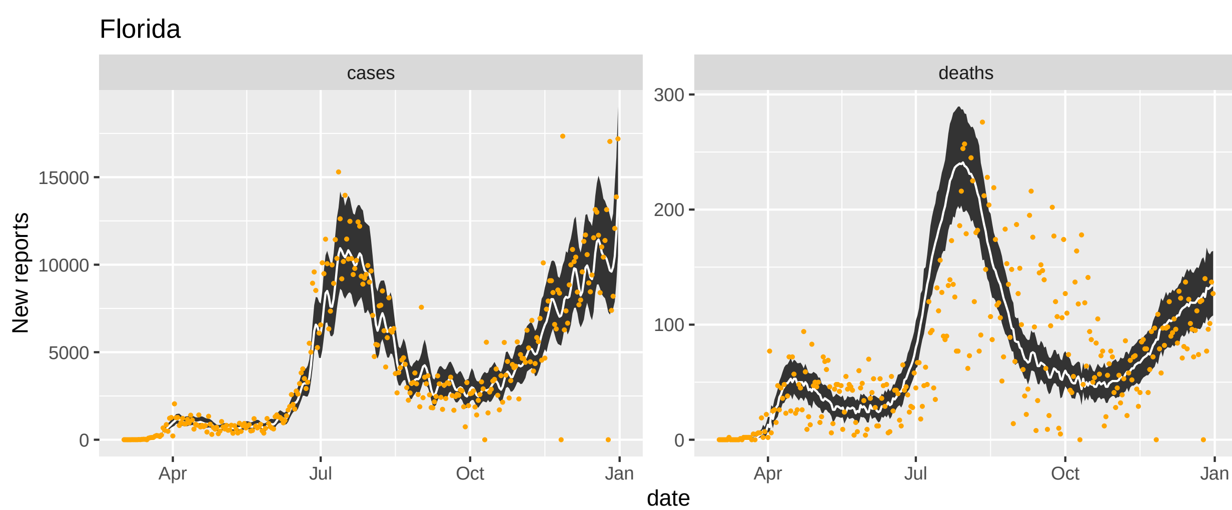

Arkansas
cases
deaths
New reports
0
1000
2000
3000
4000
0
50
100
150
Apr
Jul
Oct
Jan
date

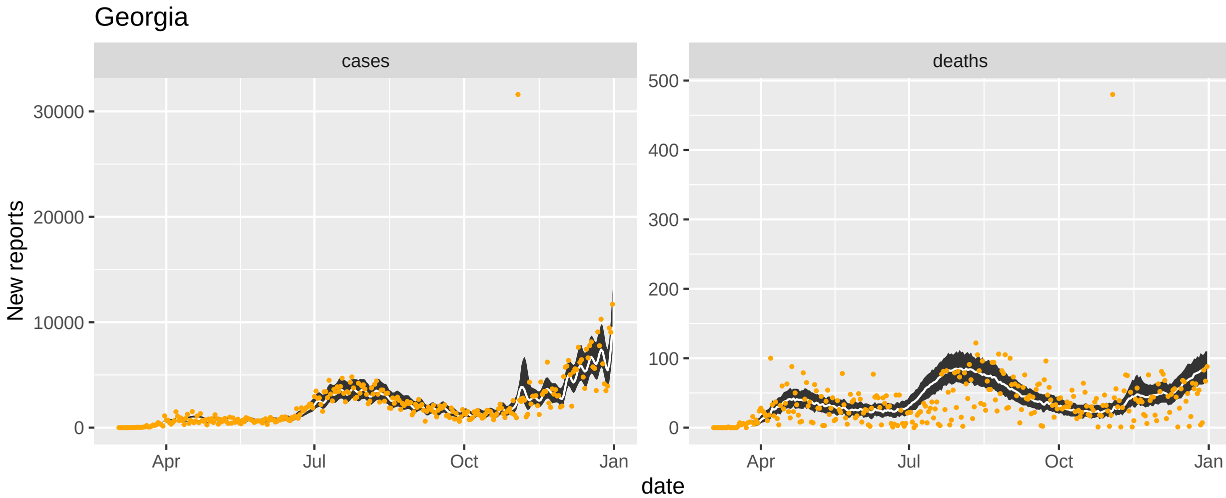

California
cases
deaths
New reports
0
20000
40000
60000
0
200
400
600
Apr
Jul
Oct
Jan
date

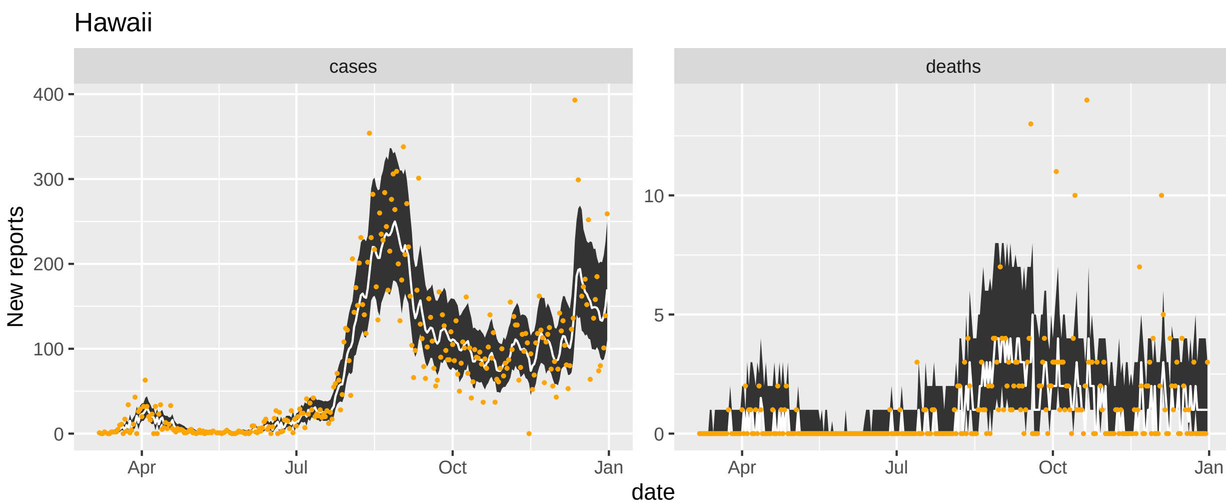

Colorado
cases
deaths
New reports
0
2000
4000
6000
0
100
200
Apr
Jul
Oct
Jan
date

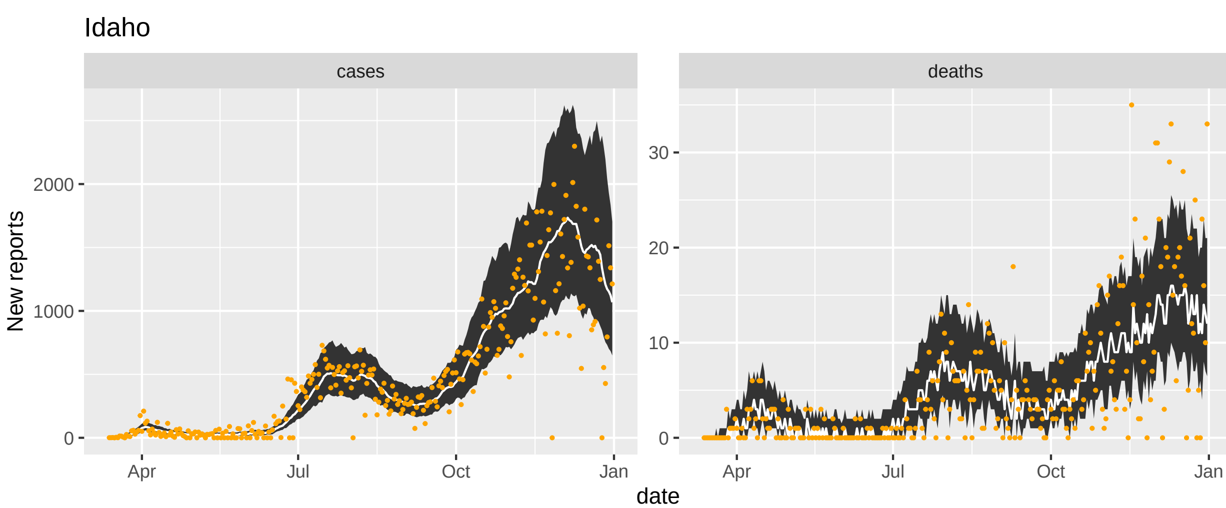

Connecticut
cases
deaths
New reports
8000
6000
4000
2000
0
200
150
100
50
Apr
Jul
Oct
Jan
date

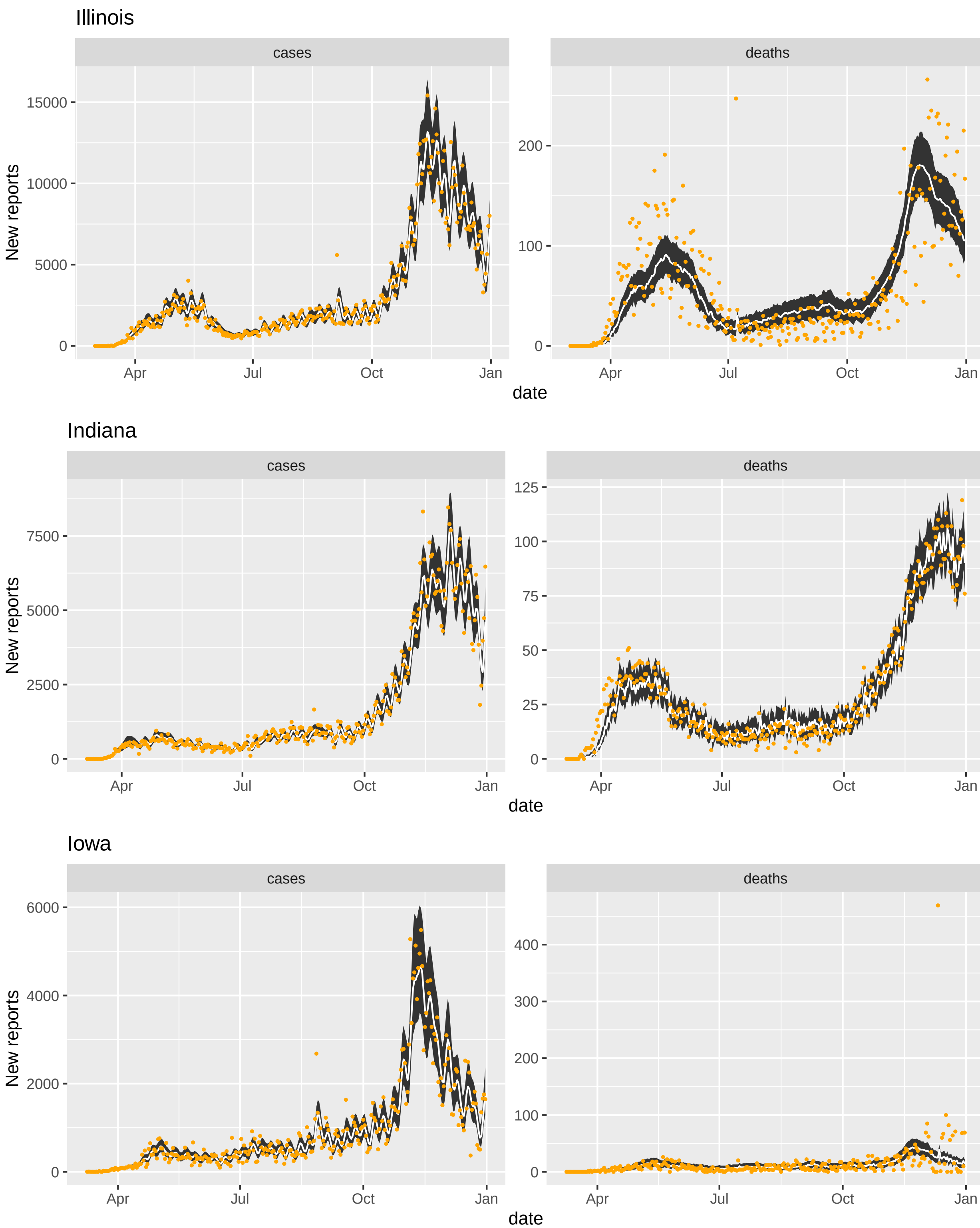

Delaware
cases
deaths
New reports
1000
500
0
15
10
5
Apr
Jul
Oct
Jan
date

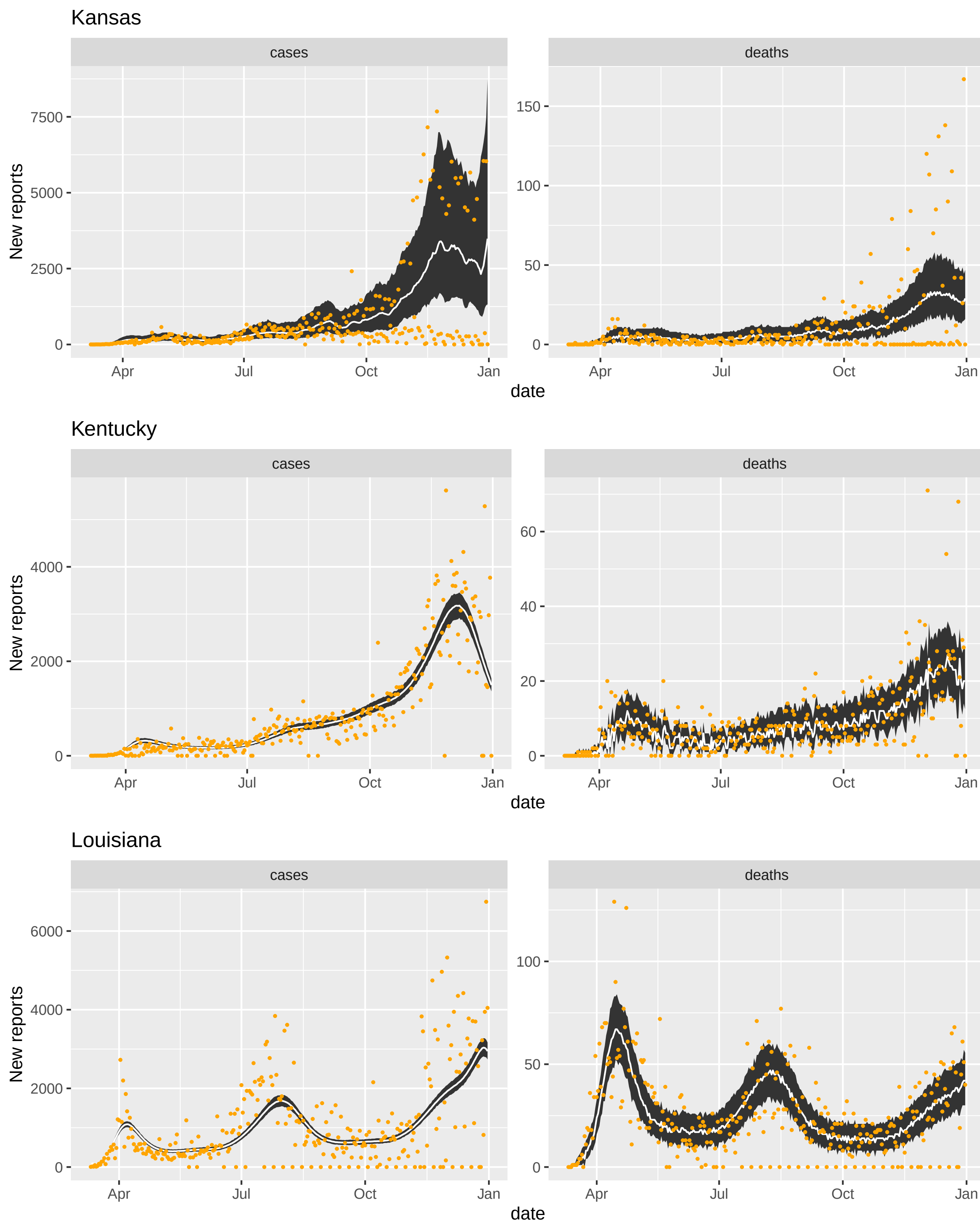

Florida
cases
deaths
New reports
15000
10000
5000
0
300
200
100
Apr
Jul
Oct
Jan
date

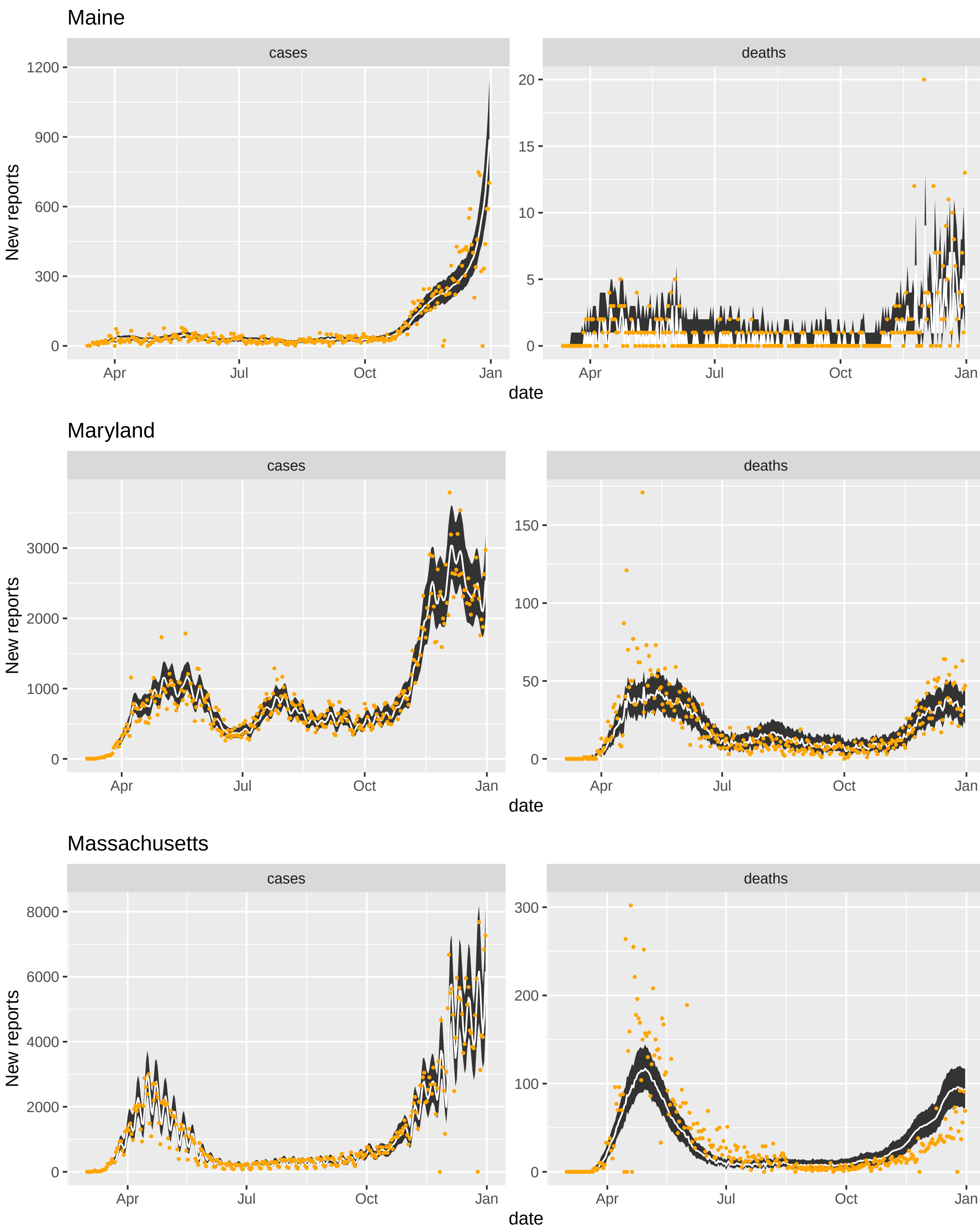

Georgia
cases
deaths
New reports
30000
20000
10000
0
500
400
300
200
100
0
Apr
Jul
Oct
Jan
date

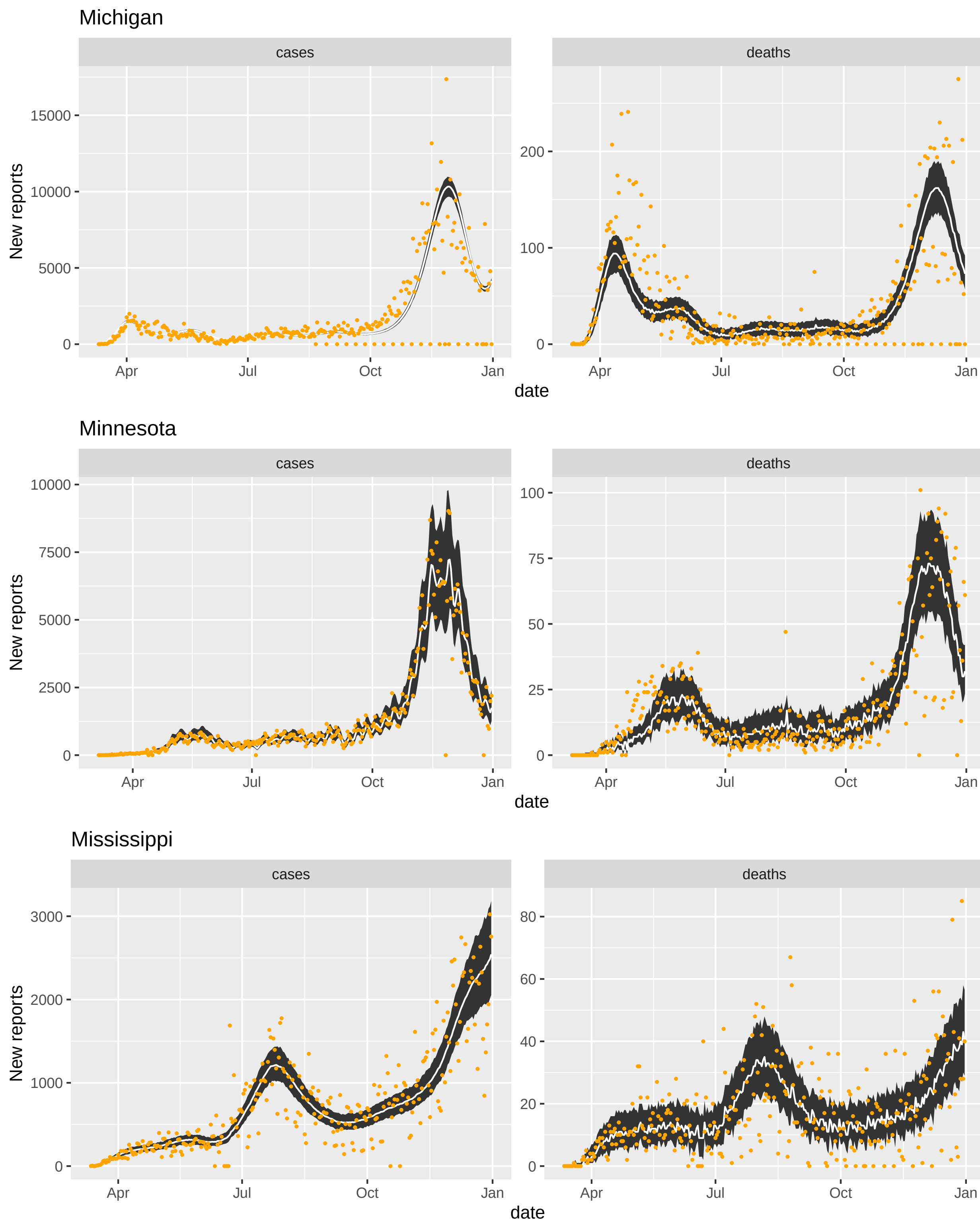

Hawaii
cases
deaths
New reports
400
300
200
100
0
10
5
0
Apr
Jul
Oct
Jan
date

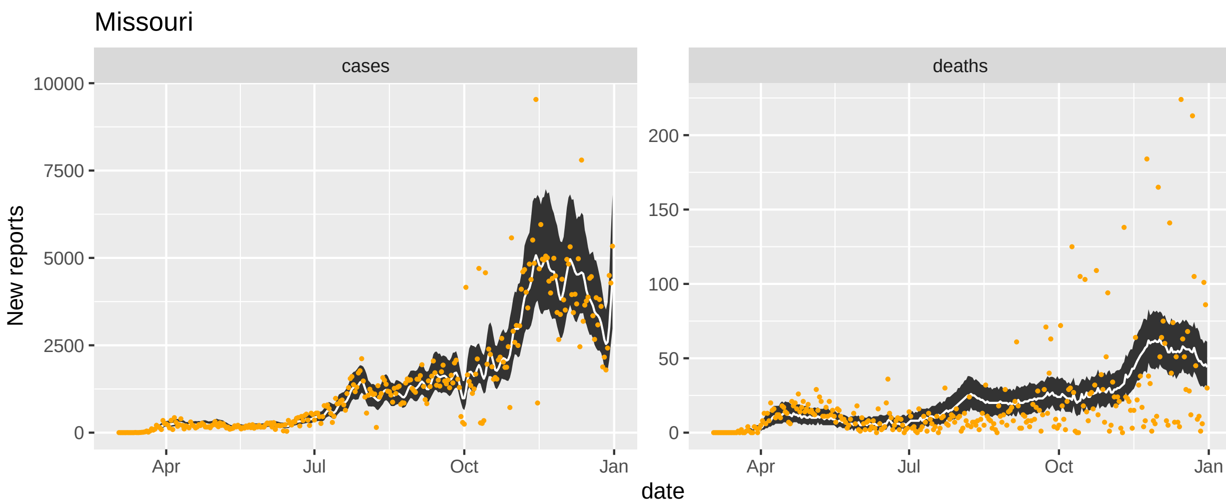

Idaho
cases
deaths
New reports
2000
1000
0
30
20
10
0
Apr
Jul
Oct
Jan
date

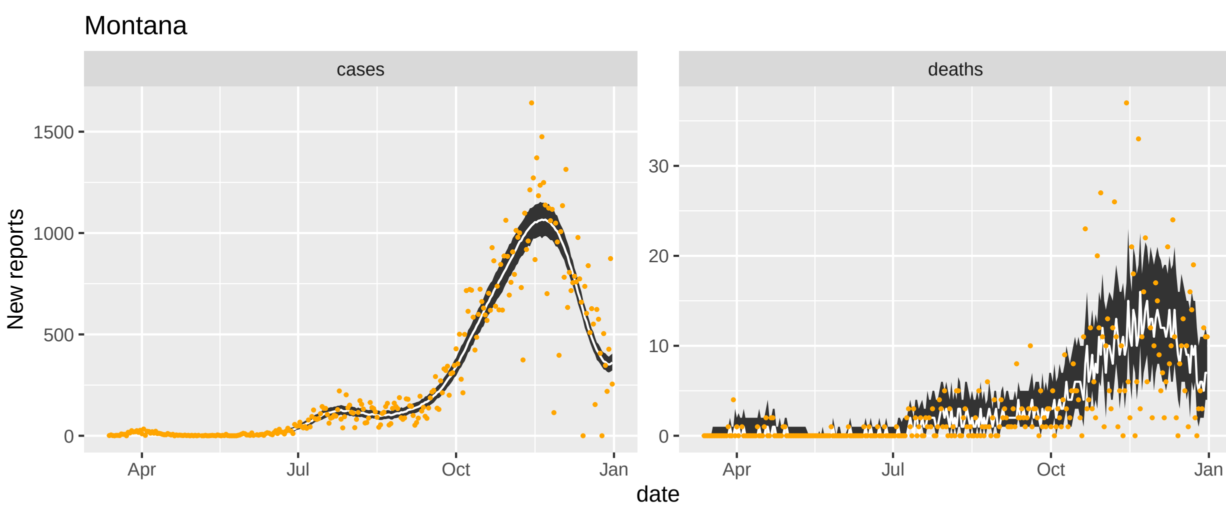

Illinois
cases
deaths
New reports
15000
10000
5000
0
200
100
Apr
Jul
Oct
Jan
date
Indiana
cases
deaths
New reports
7500
5000
2500
125
75
50
25
Iowa
cases
deaths
New reports
6000
4000
2000
400
300

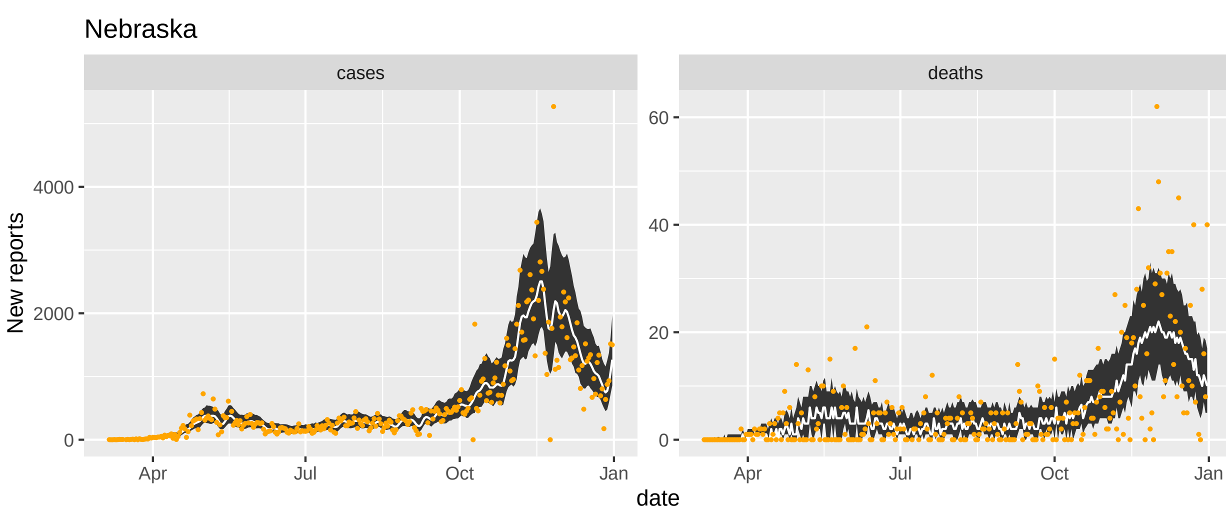

Kansas
cases
deaths
New reports
7500
5000
2500
0
150
100
50
Apr
Jul
Oct
Jan
date
Kentucky
cases
deaths
New reports
4000
2000
0
60
40
20
Apr
Jul
Oct
Jan
date
Louisiana
cases
deaths
New reports
6000
4000
2000
0
100
50
Apr
Jul
Oct
Jan
date

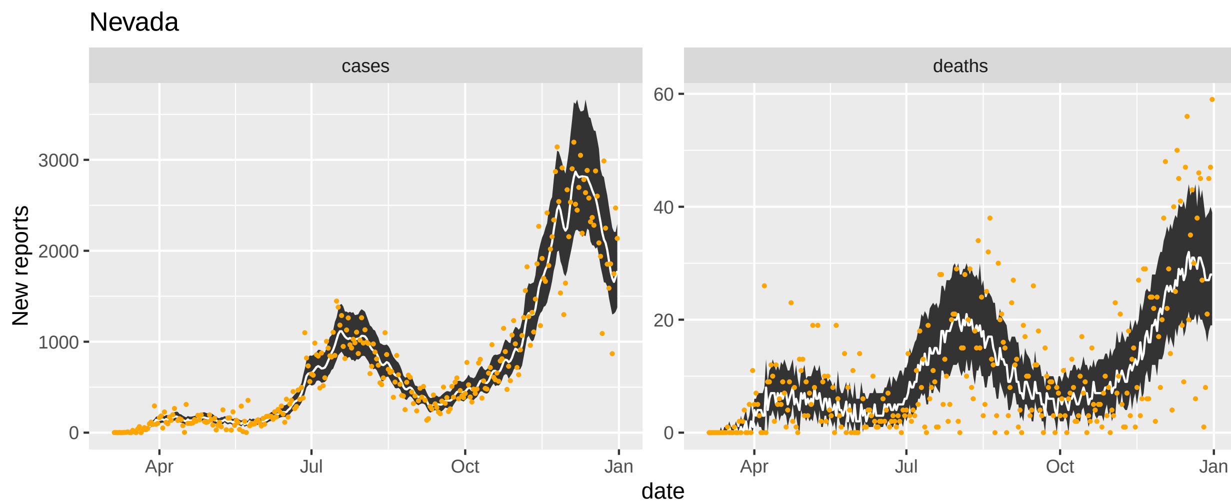

Maine
cases
deaths
New reports
1200
900
600
300
0
20
15
10
5
0
Apr
Jul
Oct
Jan
date
Maryland
cases
deaths
New reports
3000
2000
1000
0
150
100
50
0
Apr
Jul
Oct
Jan
date
Massachusetts
cases
deaths
New reports
8000
6000
4000
2000
0
300
200
100
0
Apr
Jul
Oct
Jan
date

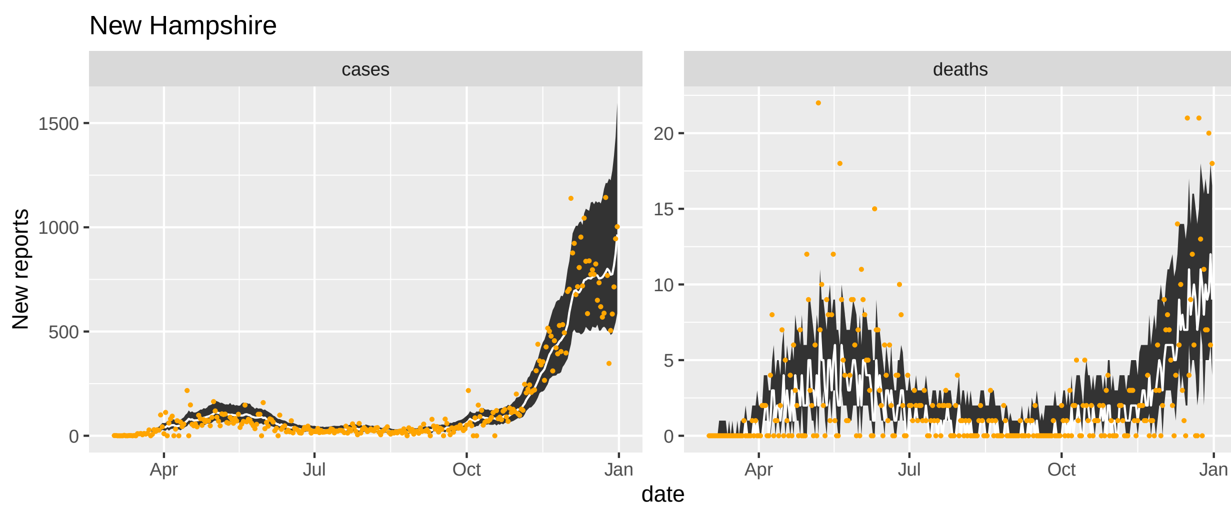

Michigan
cases
deaths
New reports
15000
10000
5000
0
200
100
Apr
Jul
Oct
Jan
date
Minnesota
10000
7500
5000
2500
100
75
50
25
Mississippi
3000
2000
1000
80
60
40
20

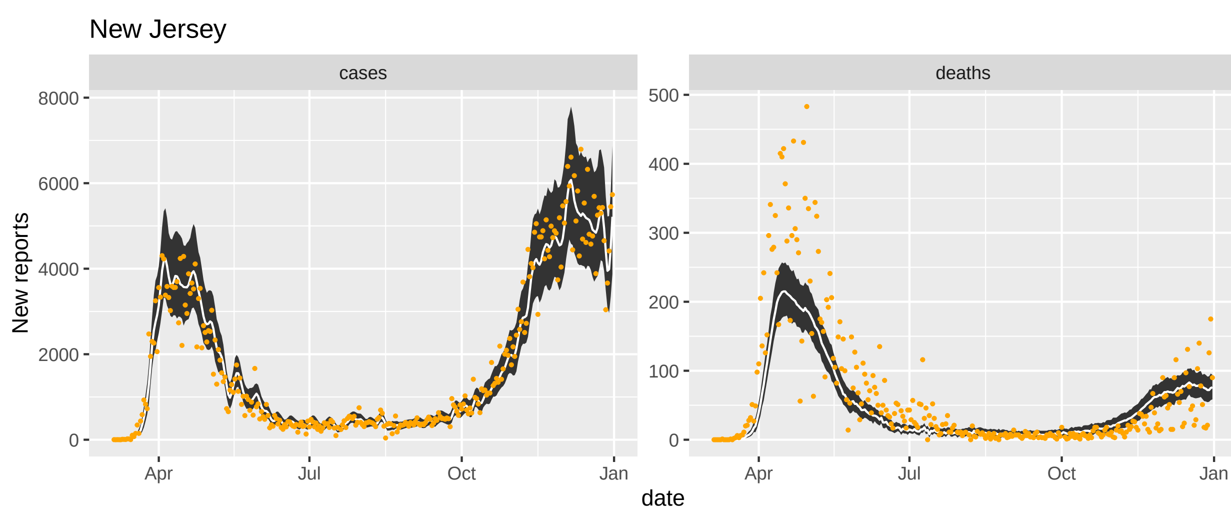

Missouri
cases
deaths
New reports
10000
7500
5000
2500
0
200
150
100
50
0
Apr
Jul
Oct
Jan
date

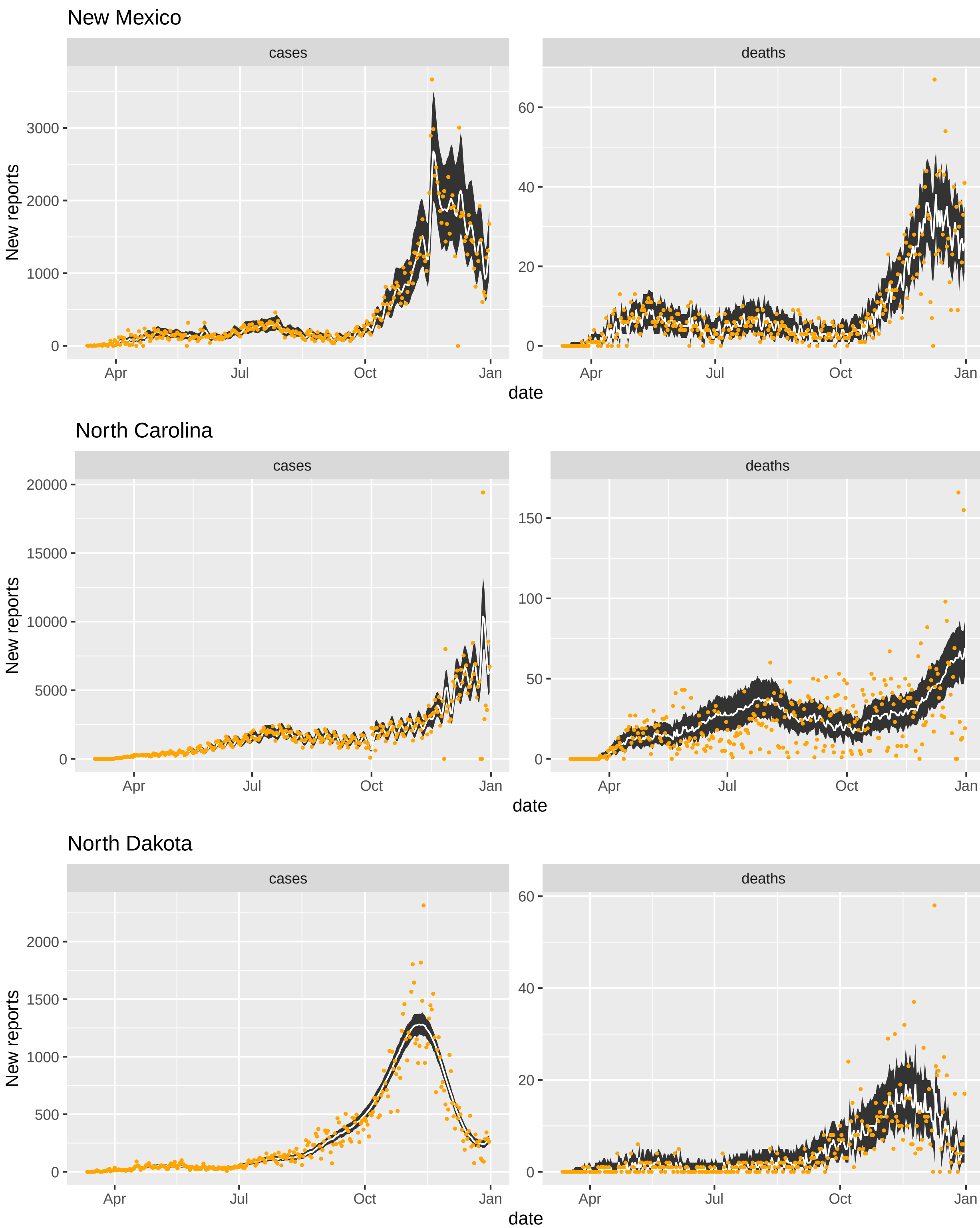

Montana
cases
deaths
New reports
1500
1000
500
0
30
20
10
0
Apr
Jul
Oct
Jan
date

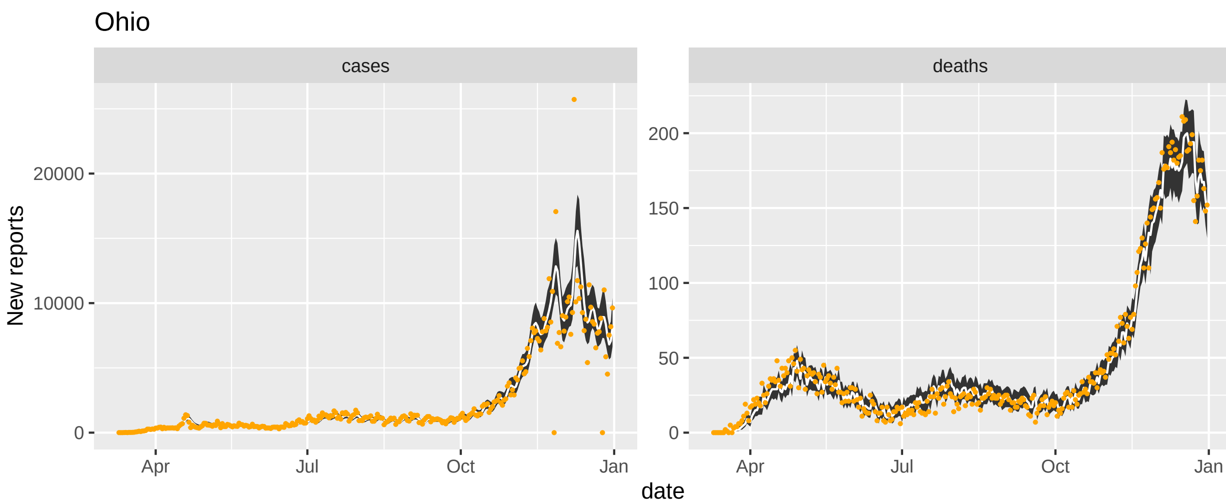

Nebraska
cases
deaths
New reports
4000
2000
0
60
40
20
0
Apr
Jul
Oct
Jan
date

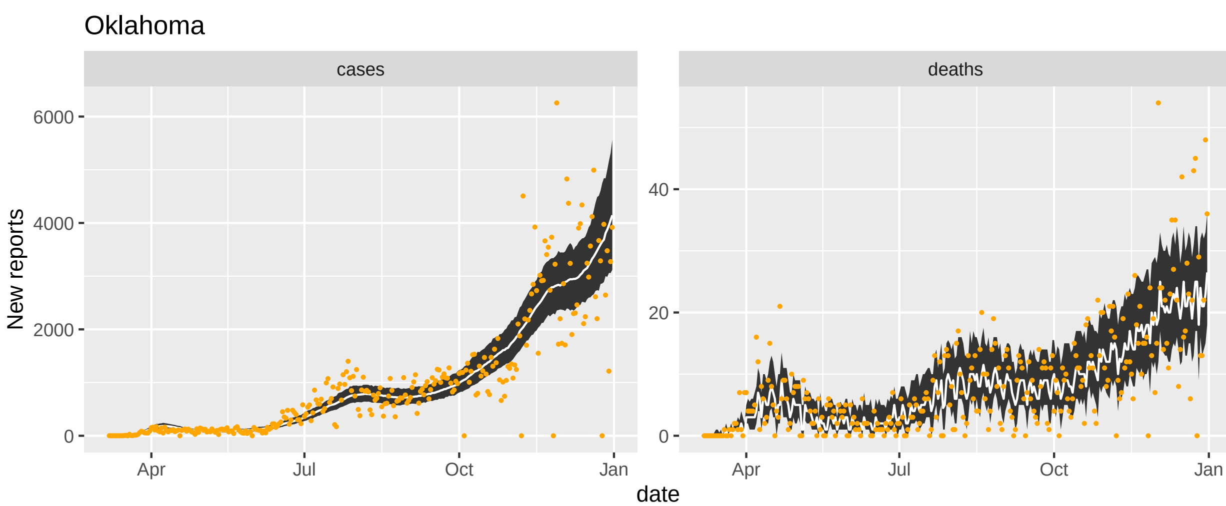

Nevada
cases
deaths
New reports
3000
2000
1000
0
60
40
20
0
Apr
Jul
Oct
Jan
date

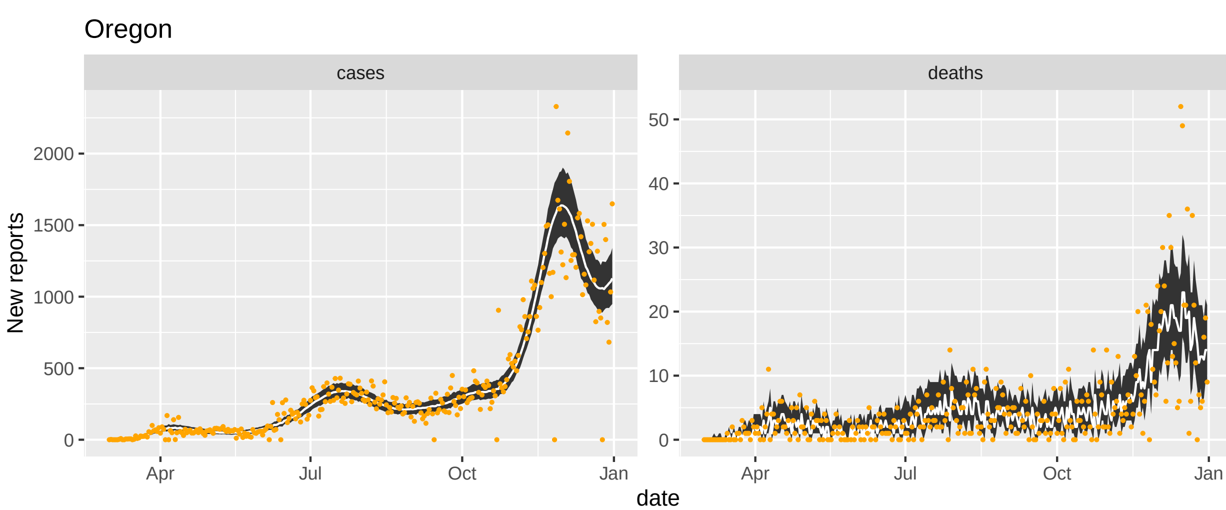

New Hampshire
cases
deaths
New reports
1500
1000
500
0
20
15
10
5
0
Apr
Jul
Oct
Jan
date

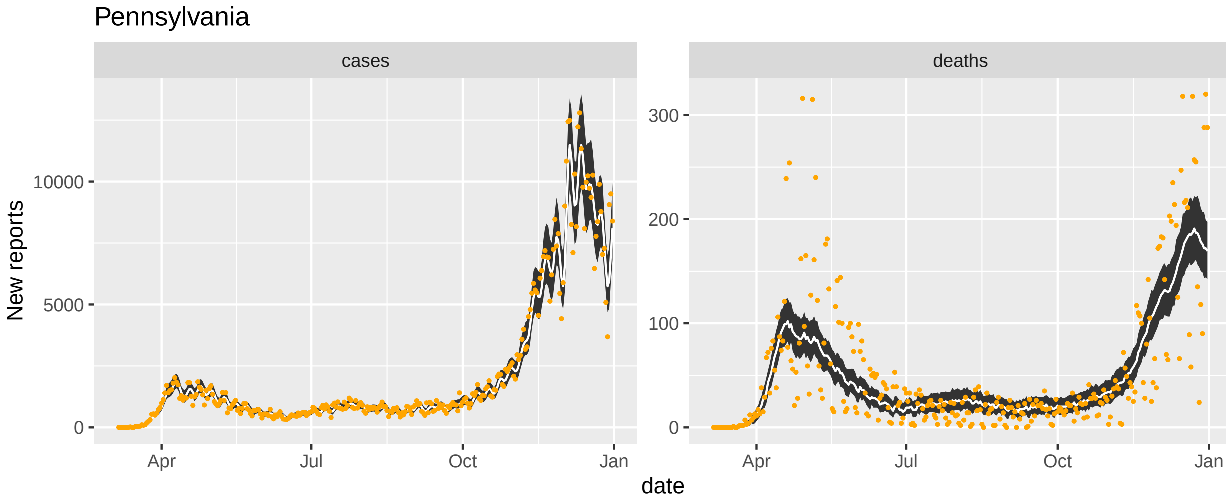

New Jersey
cases
deaths
New reports
8000
6000
4000
2000
0
500
400
300
200
100
0
Apr
Jul
Oct
Jan
date


New Mexico
cases
deaths
New reports
0
1000
2000
3000
0
20
40
60
Apr
Jul
Oct
Jan
date
North Carolina
cases
deaths
New reports
0
5000
10000
15000
20000
0
50
100
150
Apr
Jul
Oct
Jan
date
North Dakota
cases
deaths
New reports
0
500
1000
1500
2000
0
20
40
60
Apr
Jul
Oct
Jan
date


Ohio
cases
deaths
New reports
0
10000
20000
0
50
100
150
200
Apr
Jul
Oct
Jan
date


Oklahoma
cases
deaths
New reports
0
2000
4000
6000
0
20
40
Apr
Jul
Oct
Jan
date

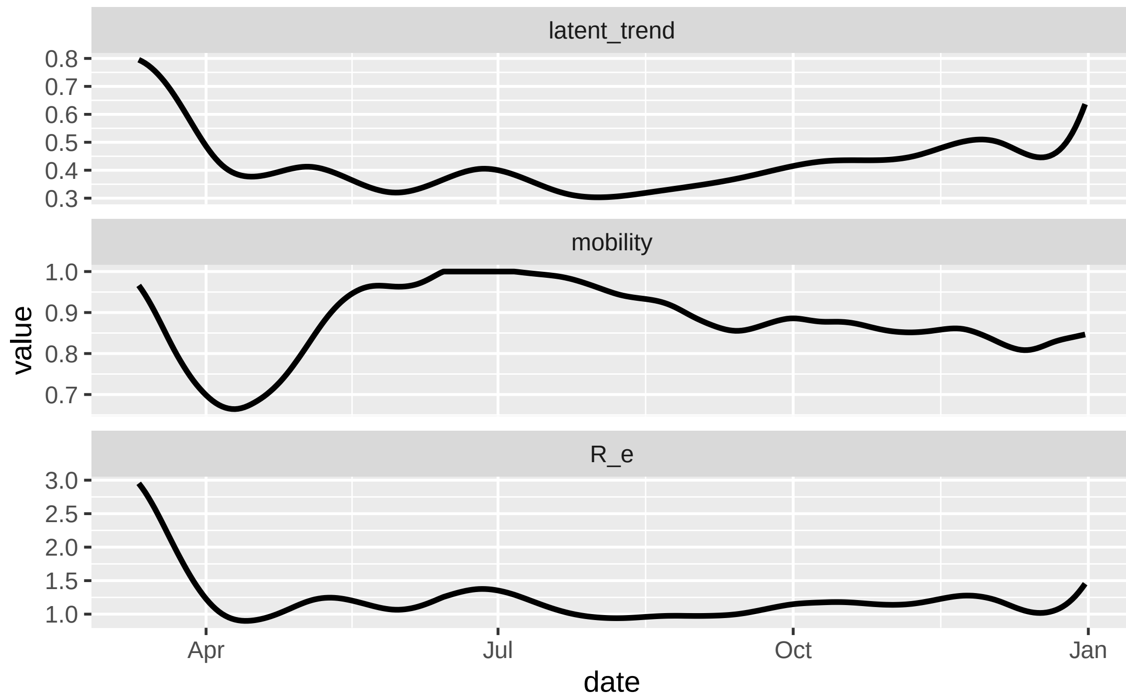

Oregon
cases
deaths
New reports
0
500
1000
1500
2000
0
10
20
30
40
50
Apr
Jul
Oct
Jan
date


Pennsylvania
cases
deaths
New reports
10000
5000
0
300
200
100
Apr
Jul
Oct
Jan
date

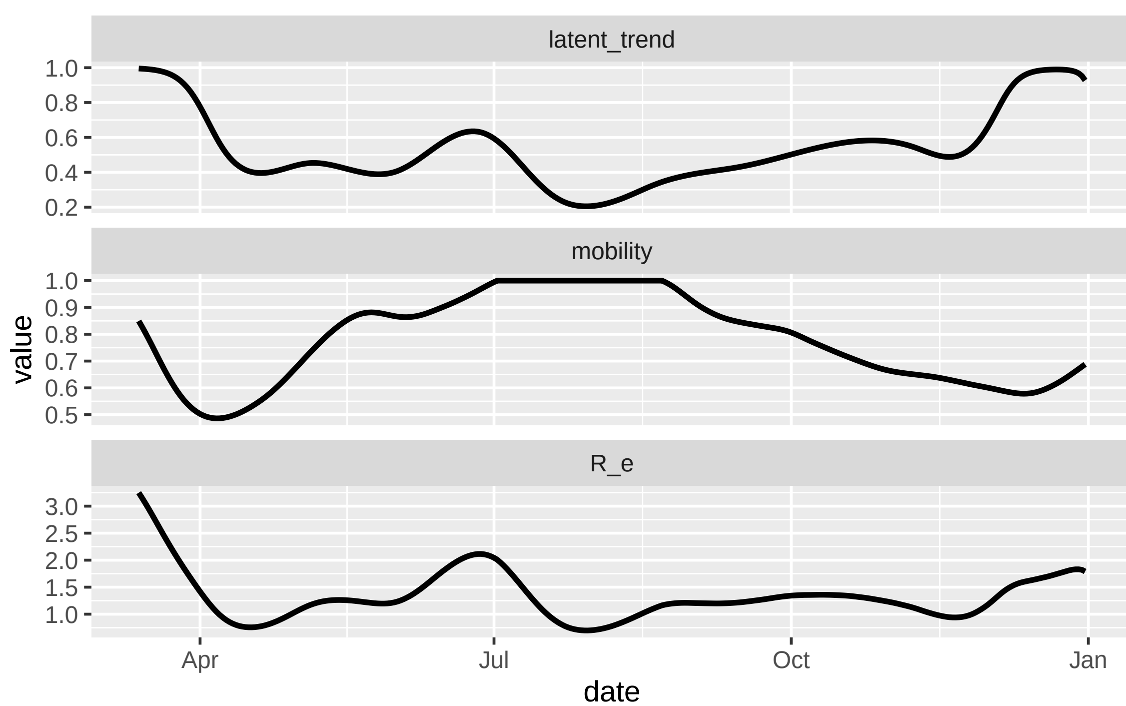
Time series of mobility, estimated latent trend, and $\mathcal{R}_e$

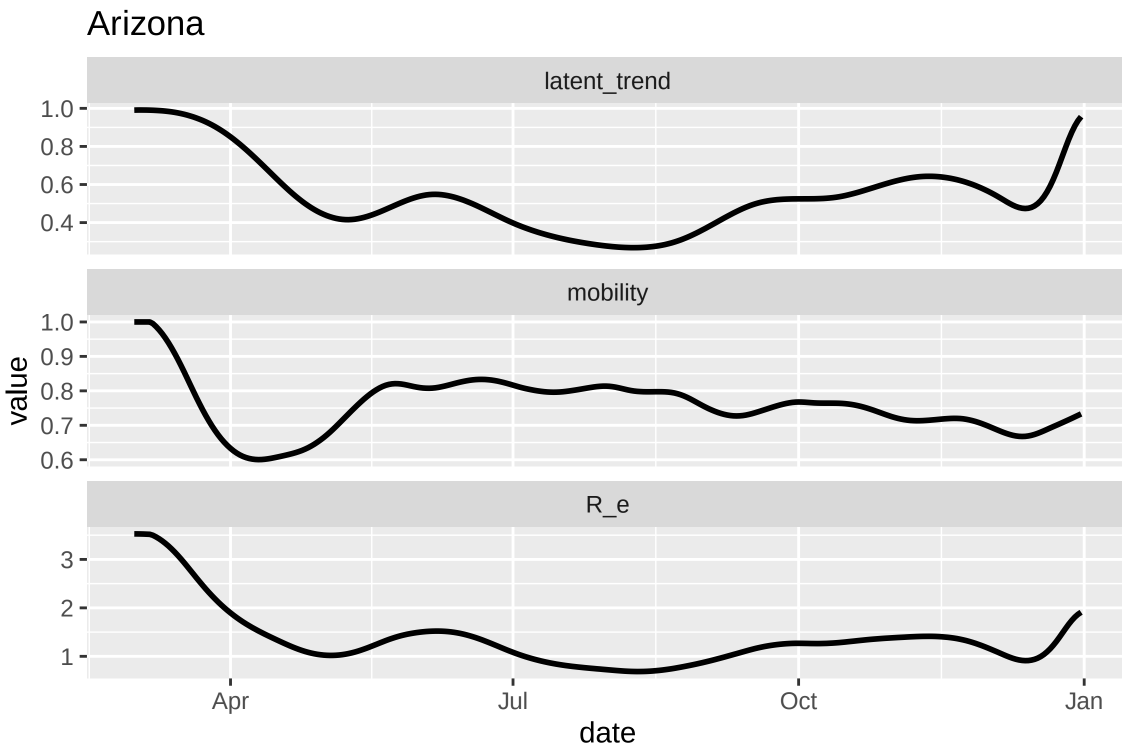
The following plots show the estimated latent trend, relative mobility, and estimated $\mathcal{R}_e$ over time for each state. The latent trend is estimated using the maximum likelihood parameter estimates for the g_i coefficients of the B-spline (see Materials and Methods in main text).

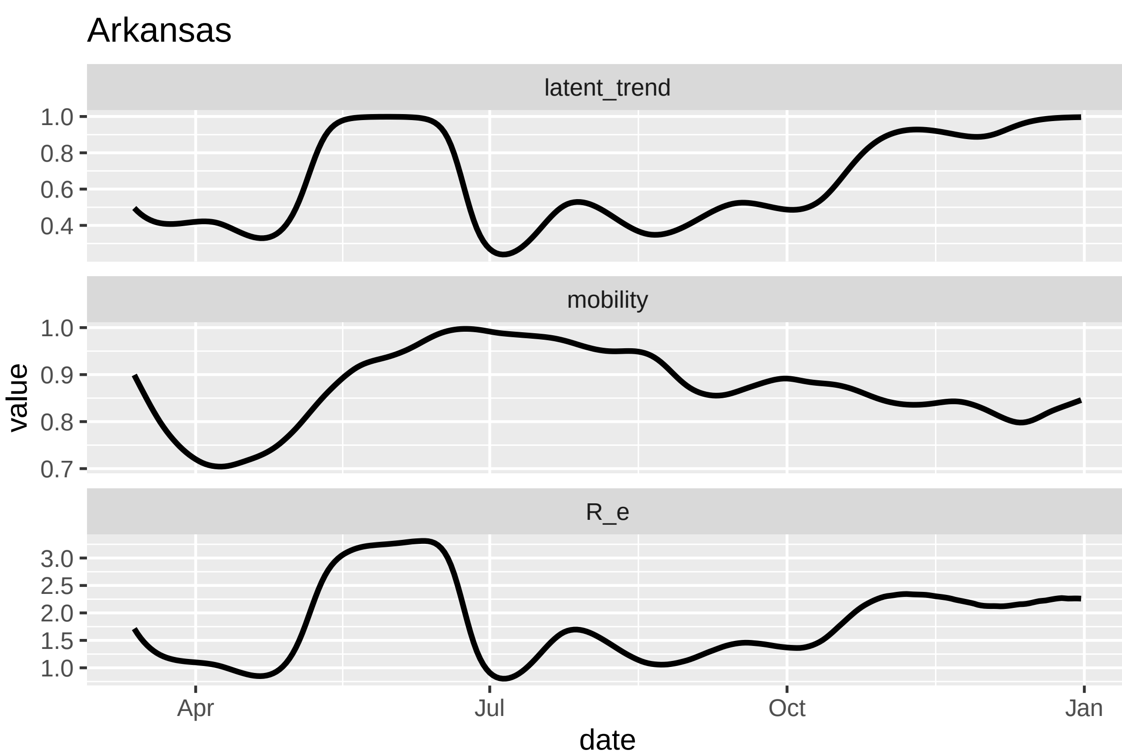
Alabama

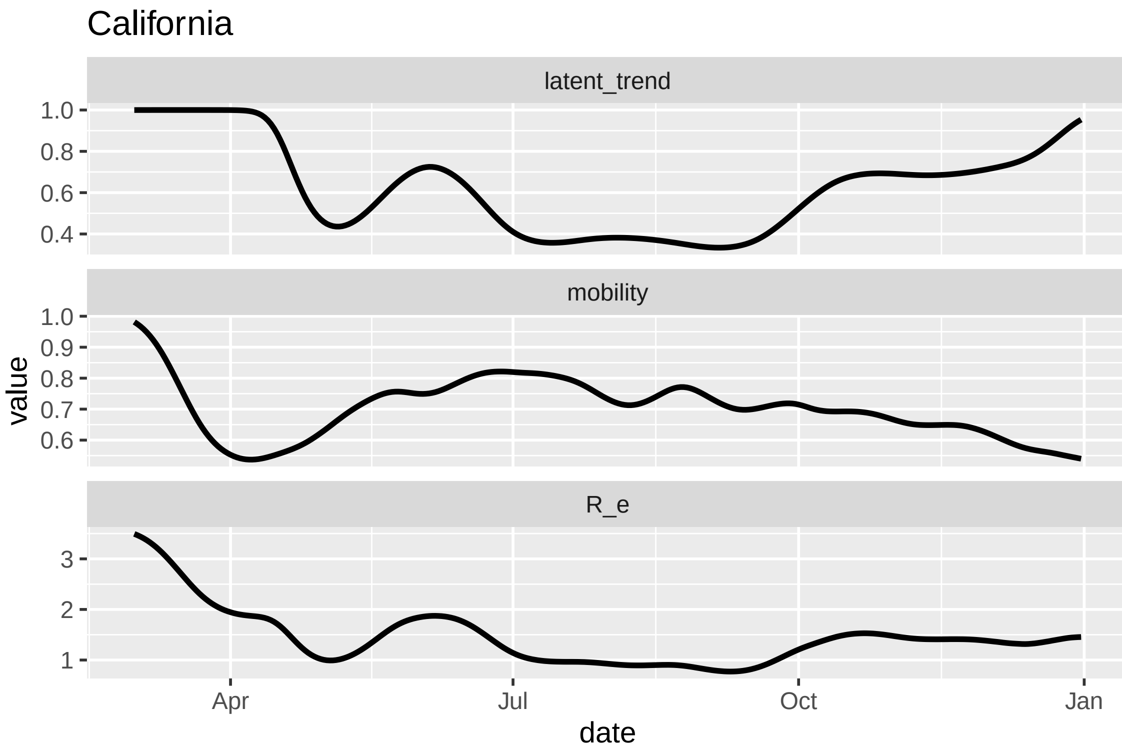


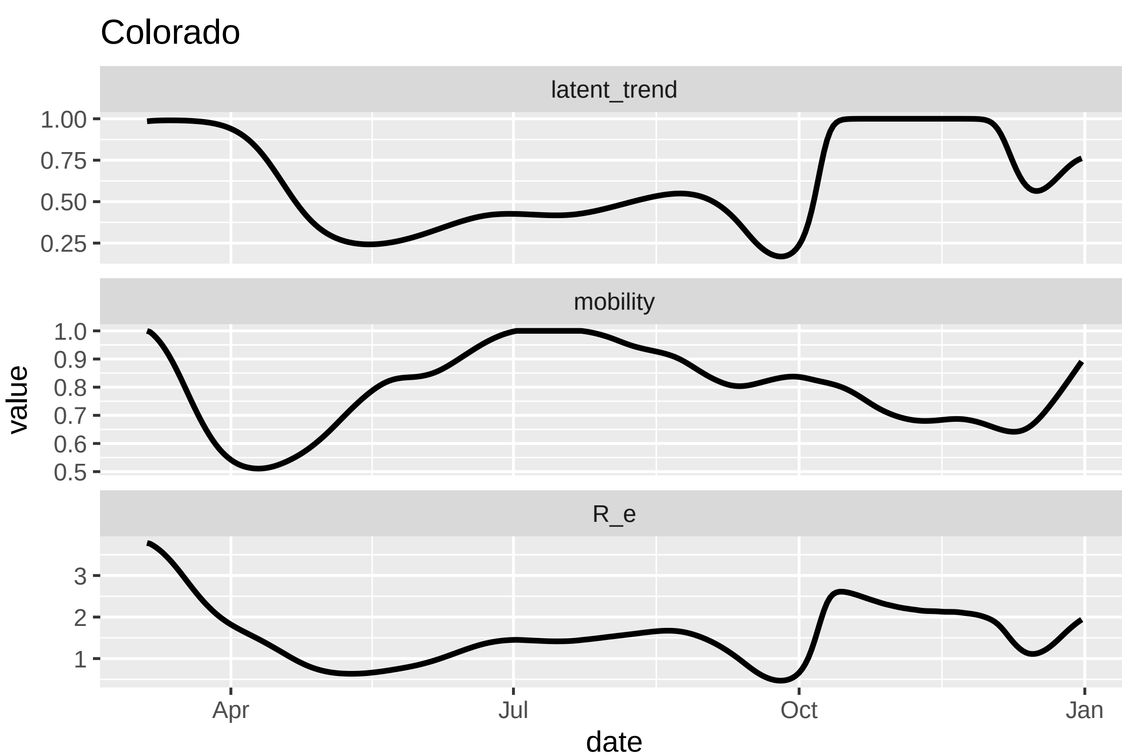
Alaska

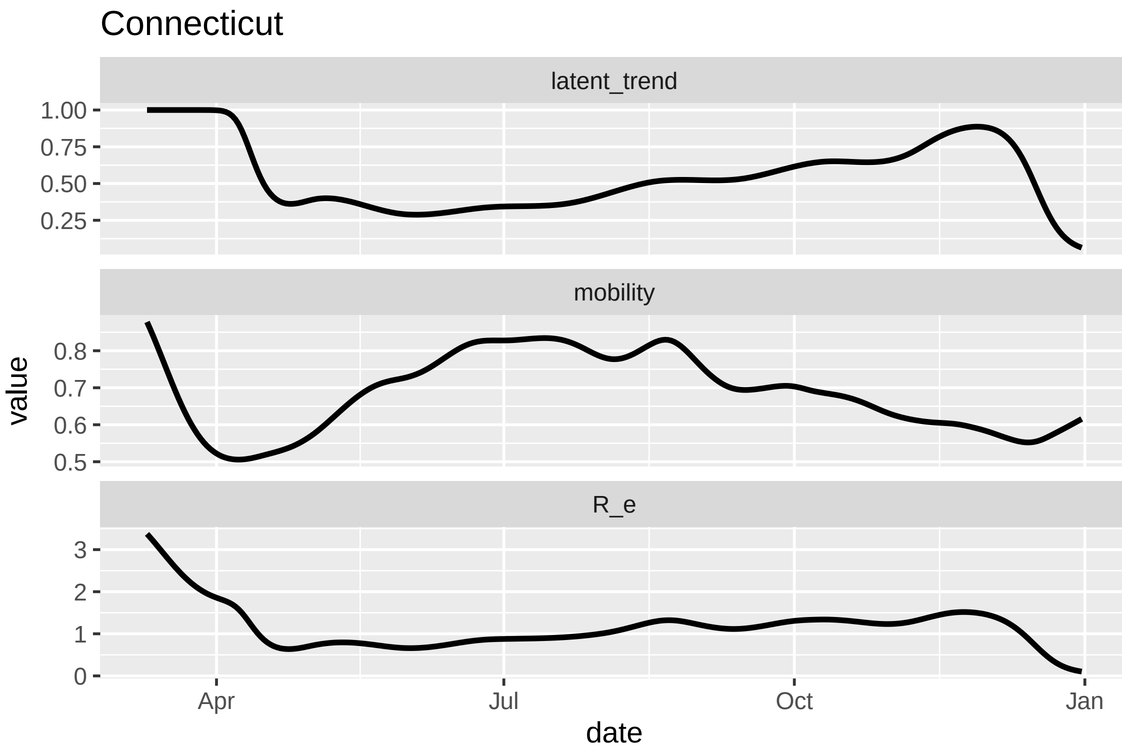


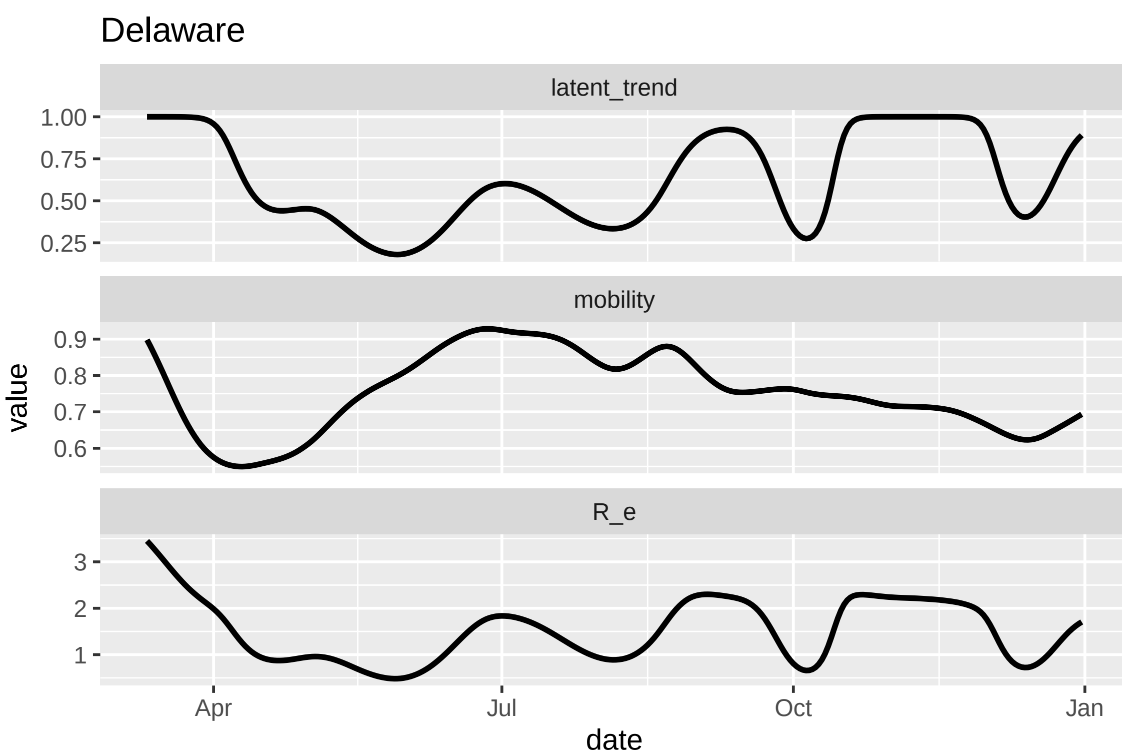

Arizona
latent_trend
mobility
R_e
value
1.0
0.8
0.6
0.4
0.9
0.7
3
2
1
Apr
Jul
Oct
Jan
date

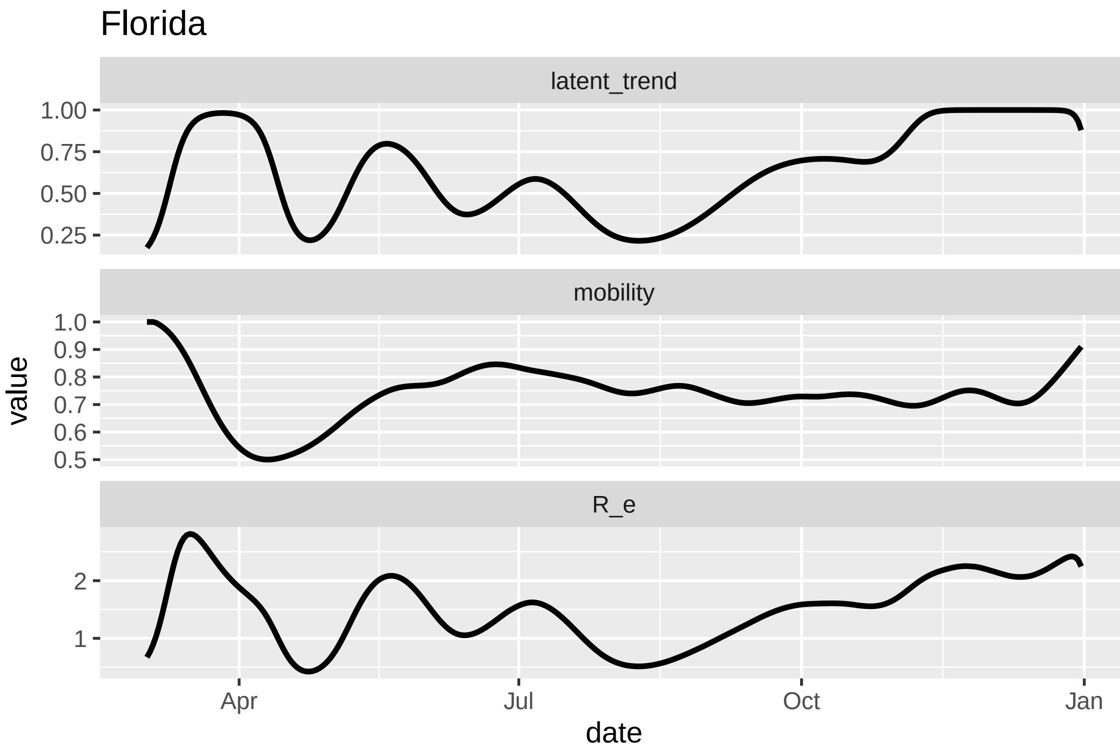

Arkansas
latent_trend
mobility
R_e
value
1.0
0.8
0.6
0.4
0.9
0.7
3.0
2.5
2.0
1.5
Apr
Jul
Oct
Jan
date

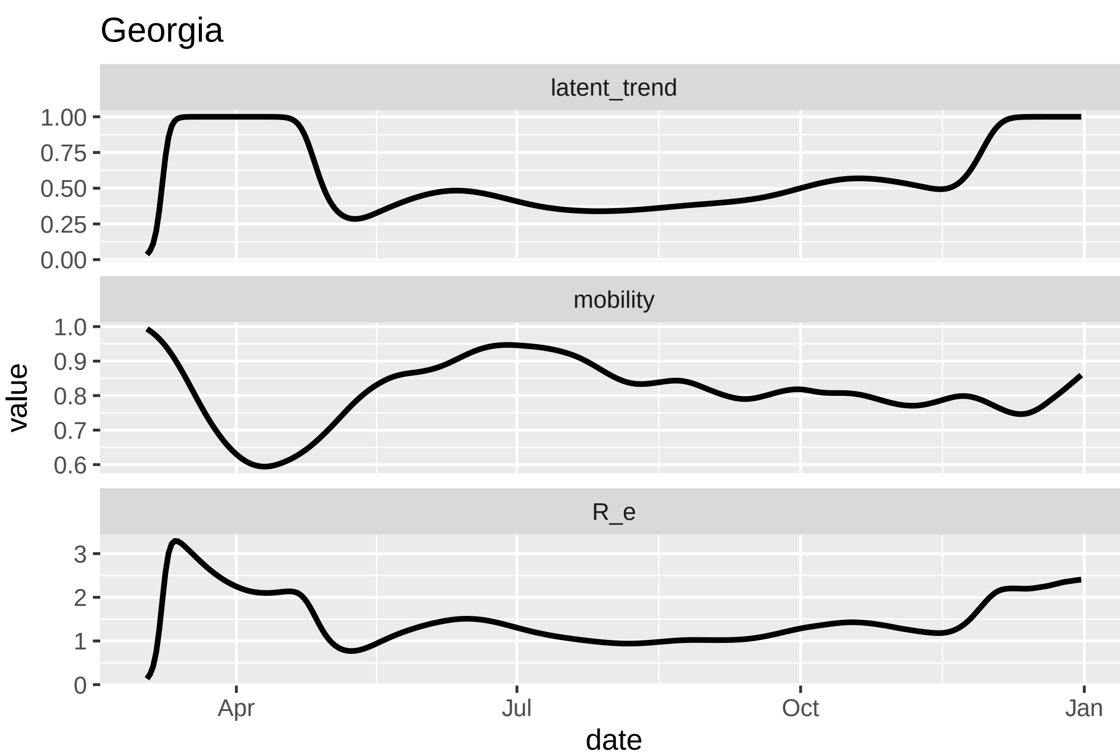

California
latent_trend
mobility
R_e
value
1.0
0.8
0.6
0.4
0.9
0.7
3
2
1
Apr
Jul
Oct
Jan
date

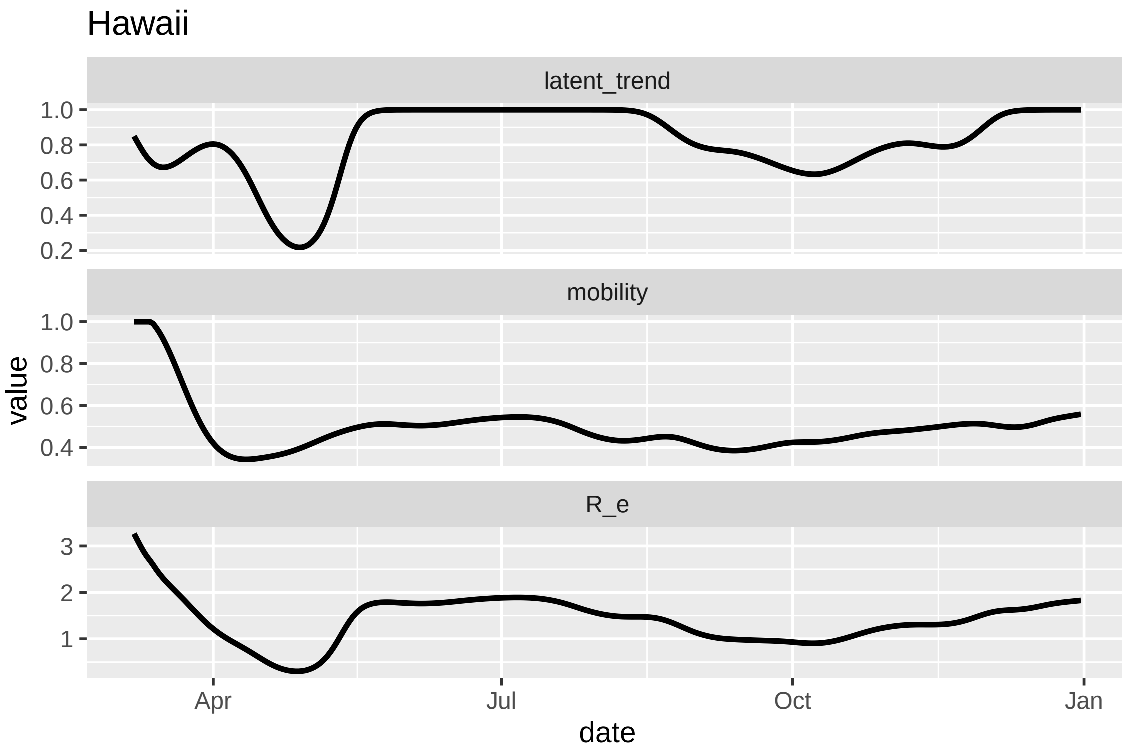

Colorado
latent_trend
mobility
R_e
value
1.00
0.75
0.50
0.25
1.0
0.9
0.8
0.7
0.6
0.5
3
2
1
Apr
Jul
Oct
Jan
date

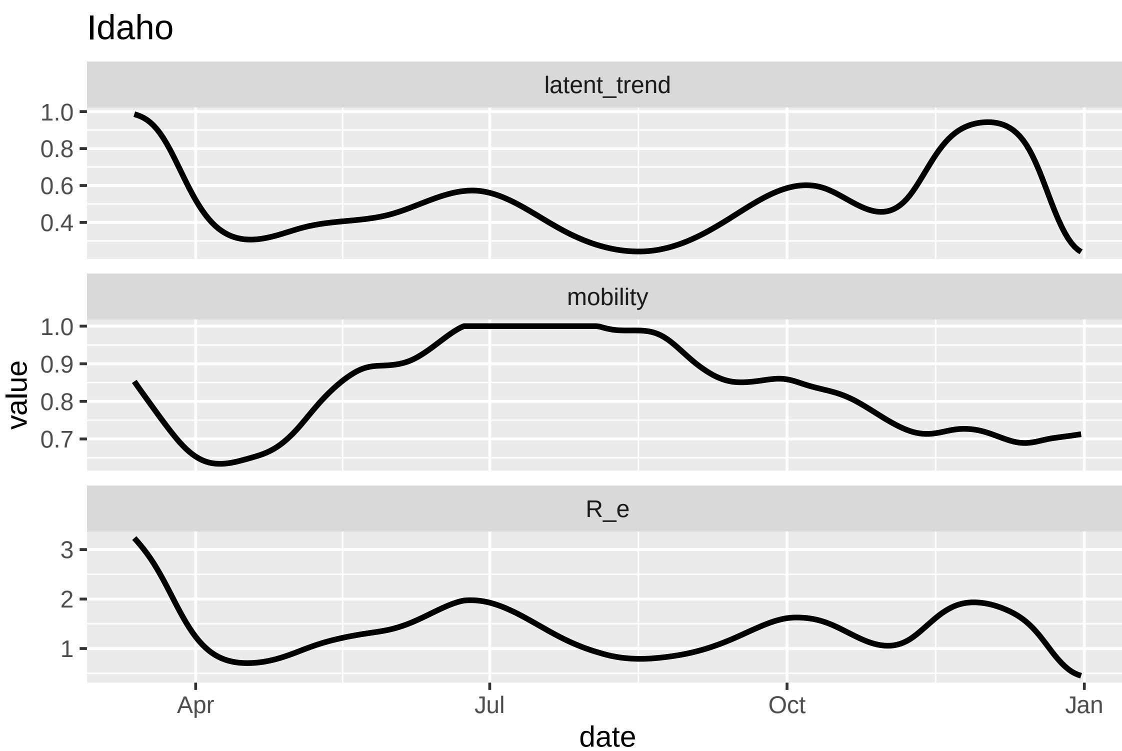

Connecticut
latent_trend
mobility
R_e
value
date
Apr
Jul
Oct
Jan
1.00
0.75
0.50
0.25
0.8
0.7
0.6
0.5
3
2
1
0

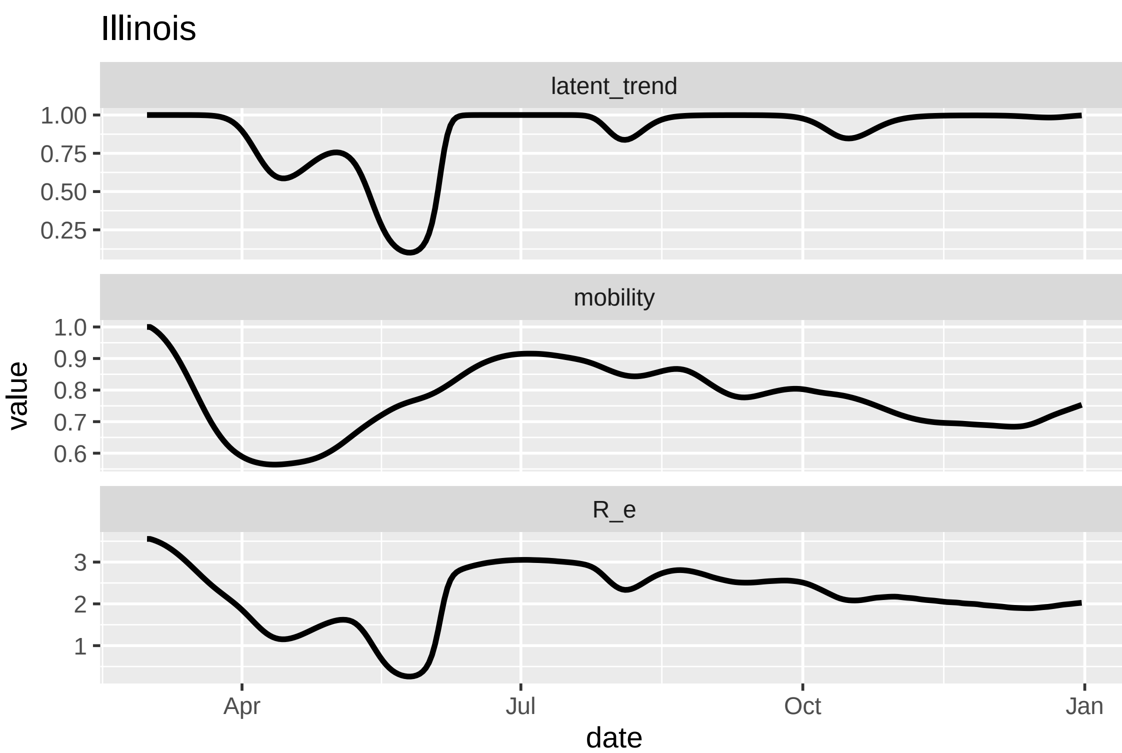

Delaware
latent_trend
mobility
R_e
value
date
Apr
Jul
Oct
Jan
1.00
0.75
0.50
0.25
0.9
0.8
0.7
0.6
3
2
1

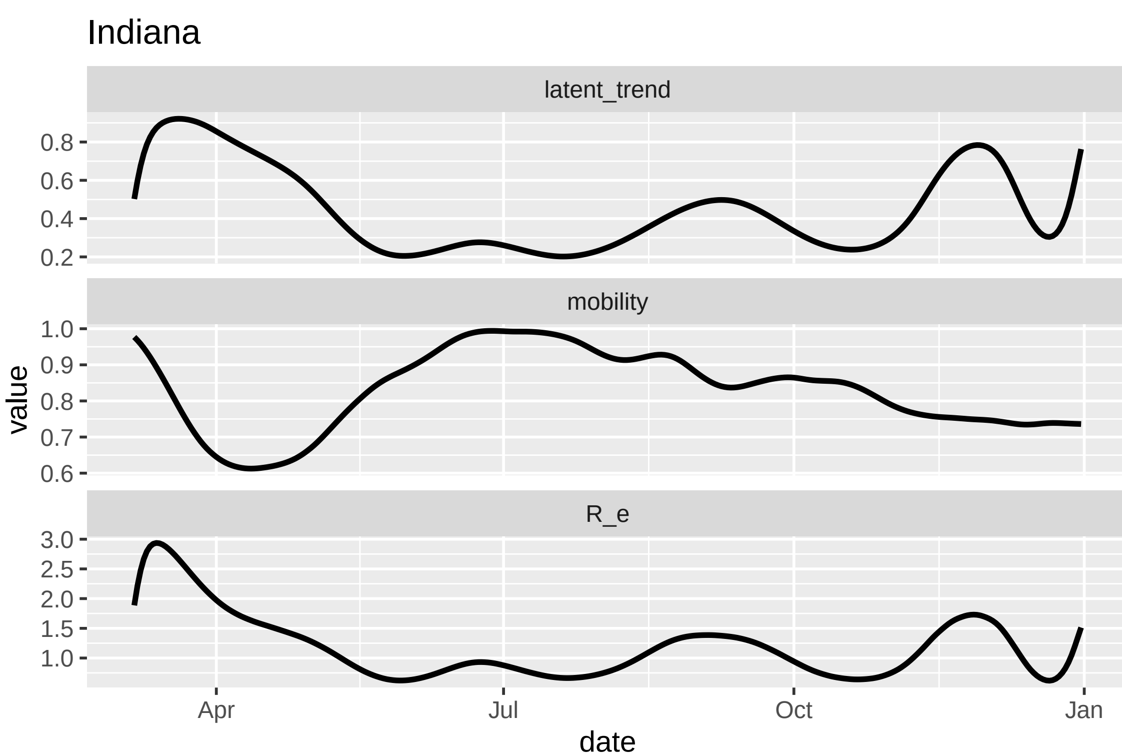

Florida
latent_trend
mobility
R_e
value
1.00
0.75
0.50
0.25
1.0
0.9
0.8
0.7
0.6
0.5
2
1
Apr
Jul
Oct
Jan
date

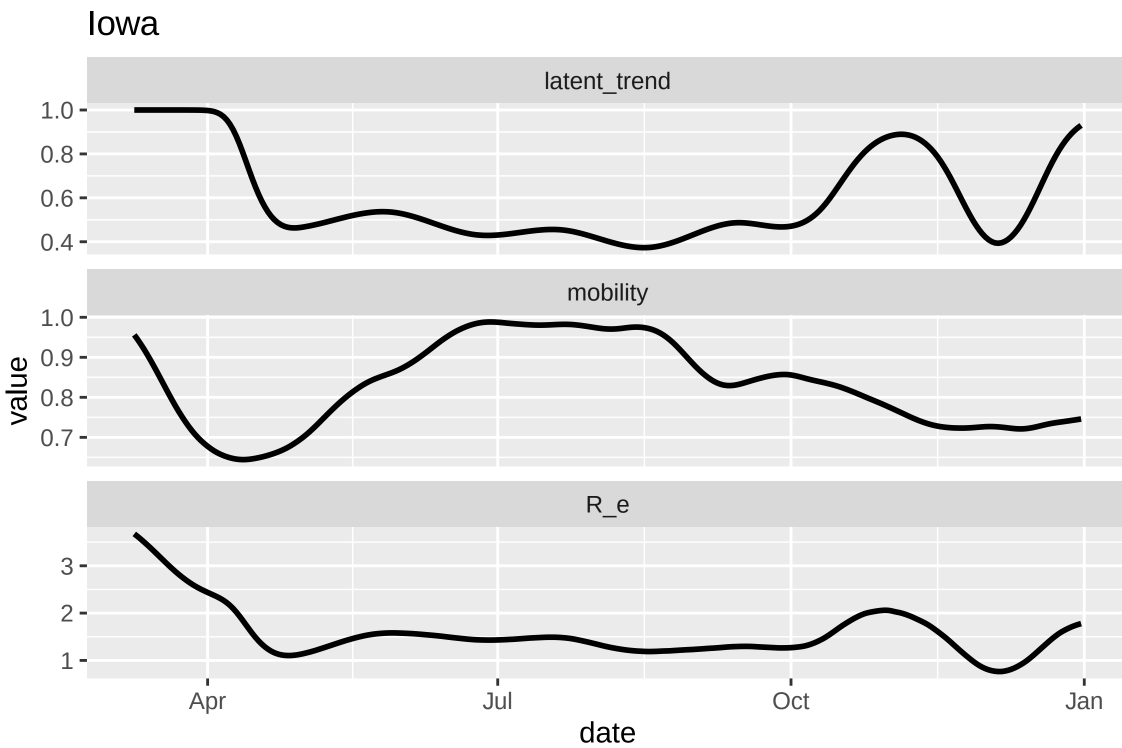

Georgia
latent_trend
mobility
R_e
value
1.00
0.75
0.50
0.25
0.00
1.0
0.9
0.8
0.7
0.6
3
2
1
0
Apr
Jul
Oct
Jan
date

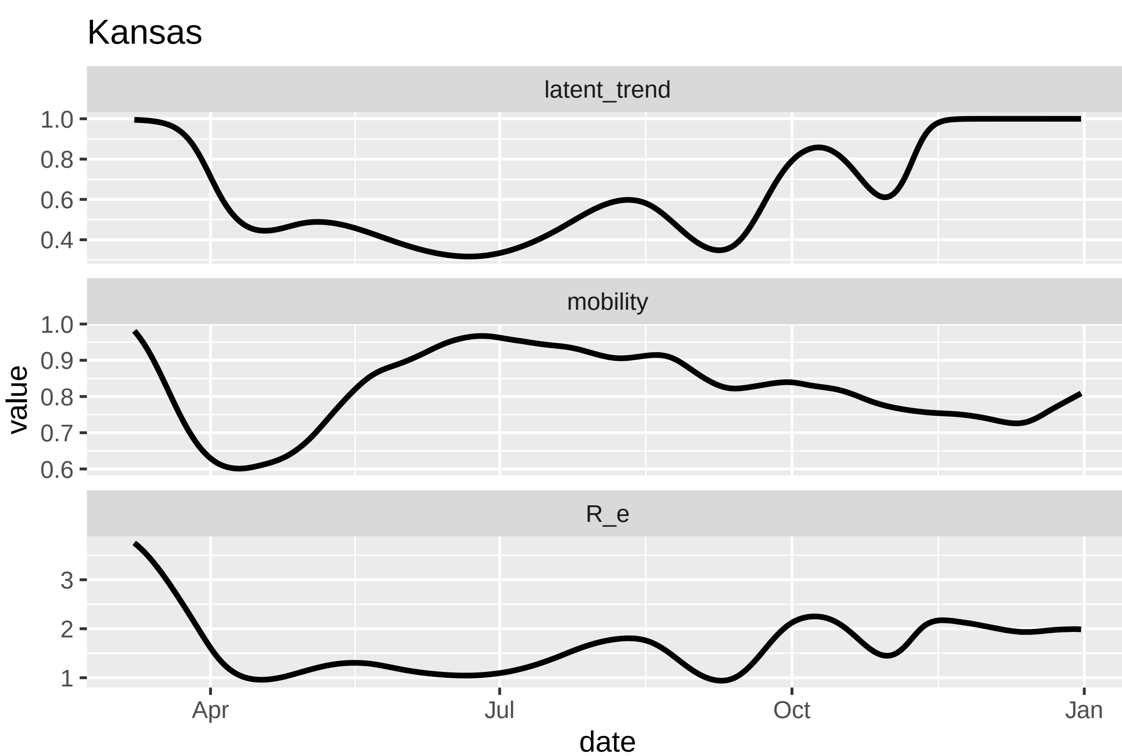

Hawaii
latent_trend
1.0
0.8
0.6
0.4
0.2
mobility
1.0
0.8
0.6
0.4
R_e
3
2
1
value
Apr
Jul
Oct
Jan
date

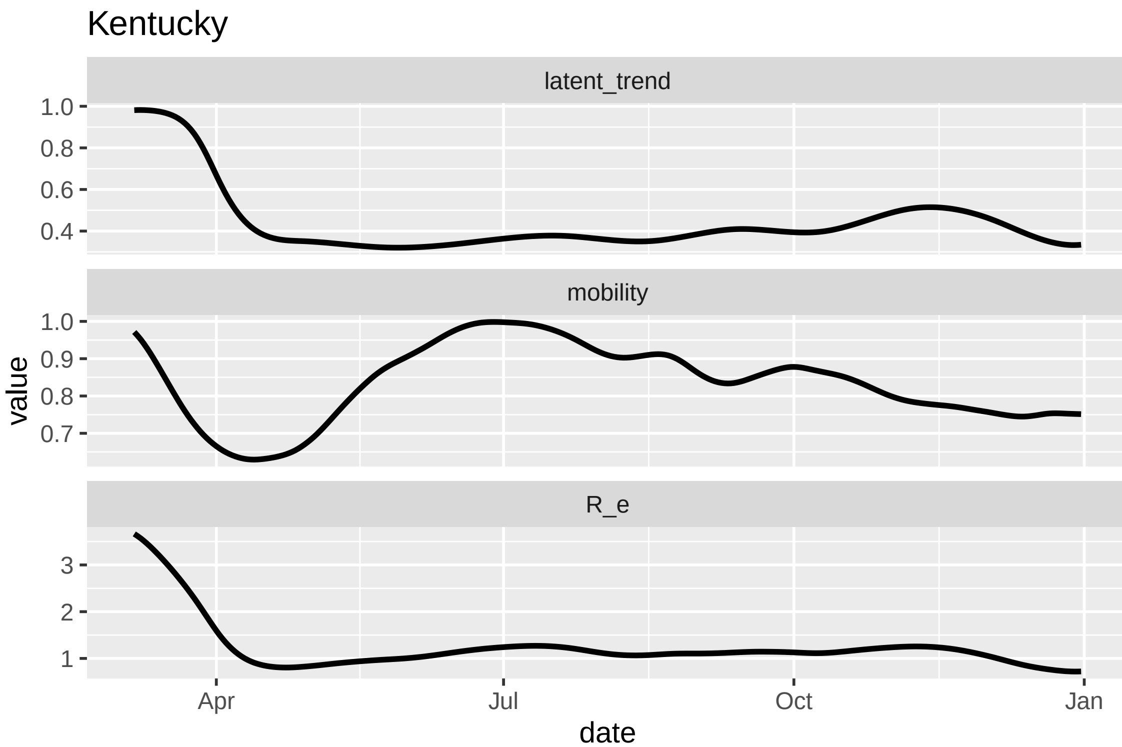

Idaho
latent_trend
1.0
0.8
0.6
0.4
mobility
1.0
0.9
0.8
0.7
R_e
3
2
1
value
Apr
Jul
Oct
Jan
date

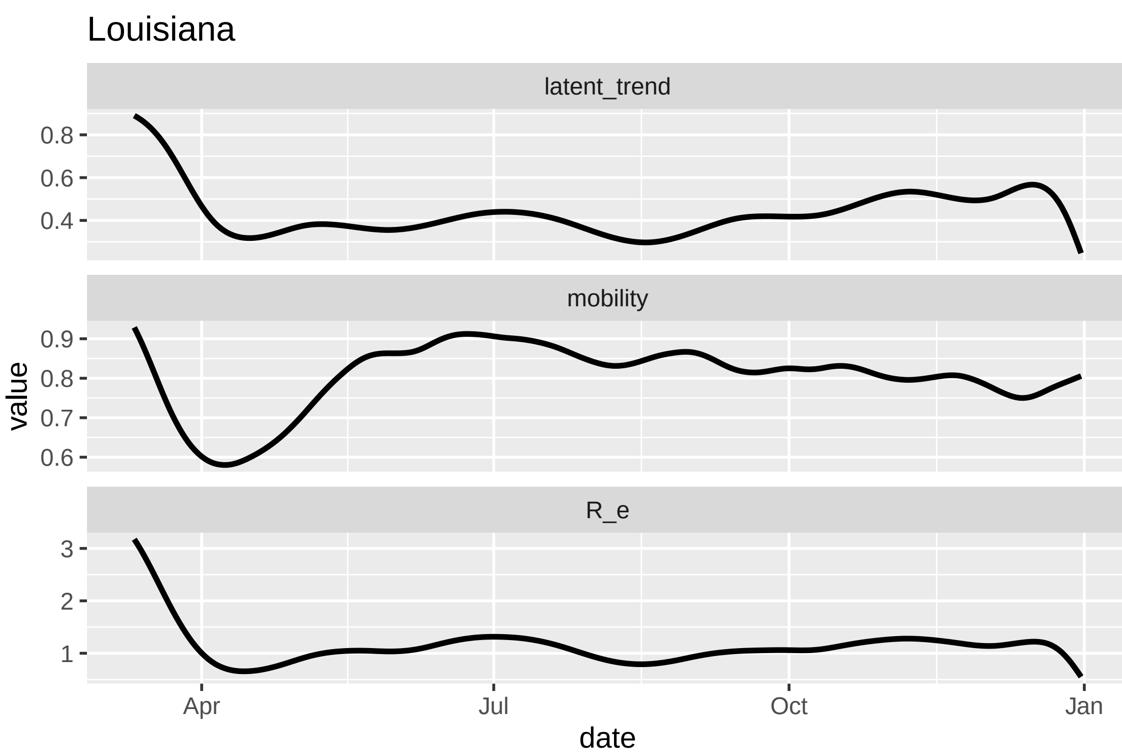

Illinois
latent_trend
1.00
0.75
0.50
0.25
mobility
1.0
0.9
0.8
0.7
0.6
value
R_e
3
2
1
Apr
Jul
Oct
Jan
date

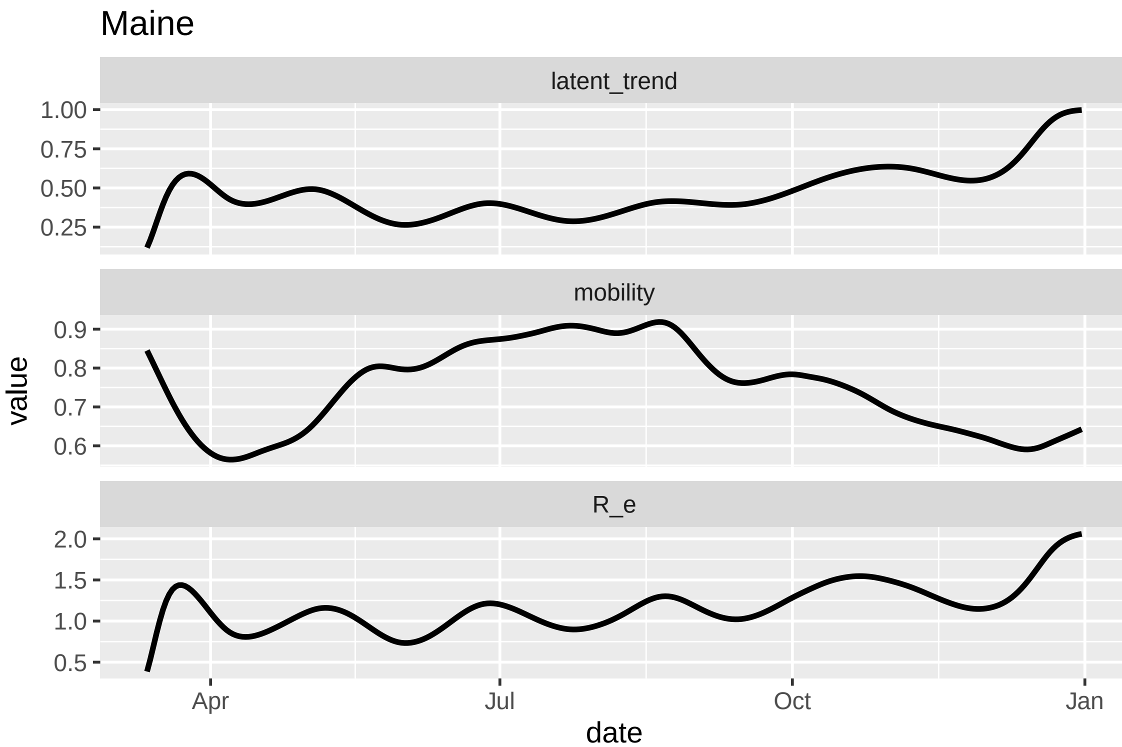

Indiana
latent_trend
0.8
0.6
0.4
0.2
mobility
1.0
0.9
0.8
0.7
0.6
value
R_e
3.0
2.5
2.0
1.5
1.0
Apr
Jul
Oct
Jan
date

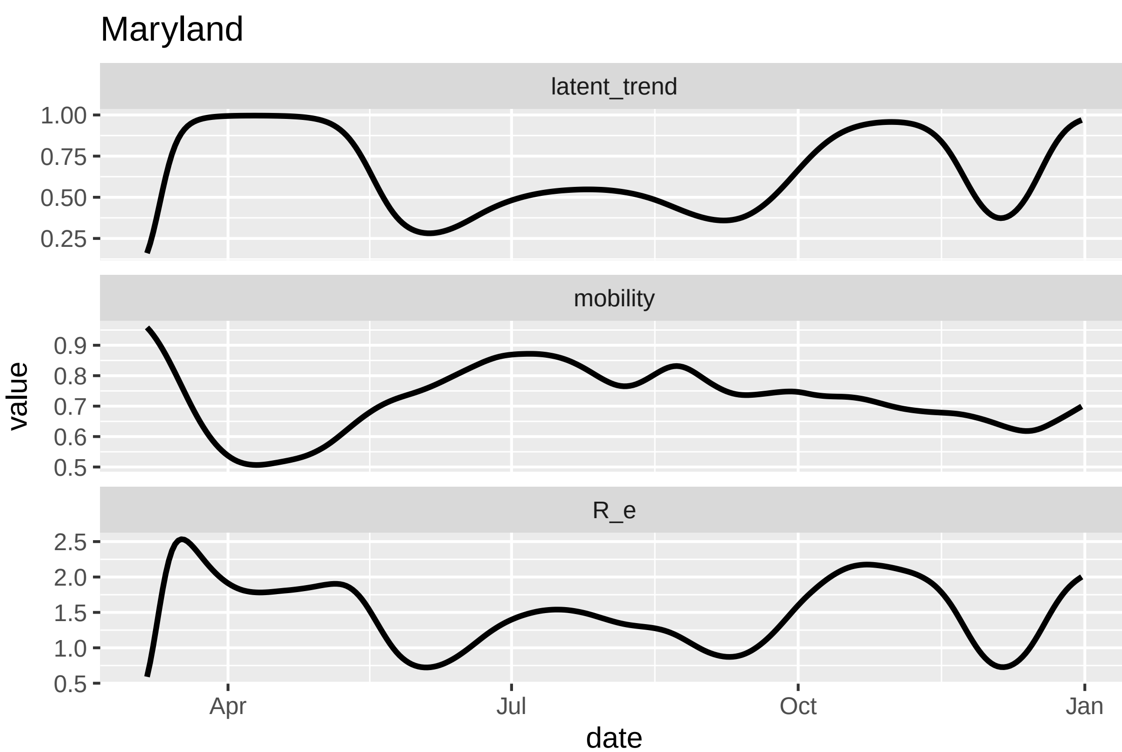

Iowa
latent_trend
1.0
0.8
0.6
0.4
mobility
1.0
0.9
0.8
0.7
value
R_e
3
2
1
Apr
Jul
Oct
Jan
date

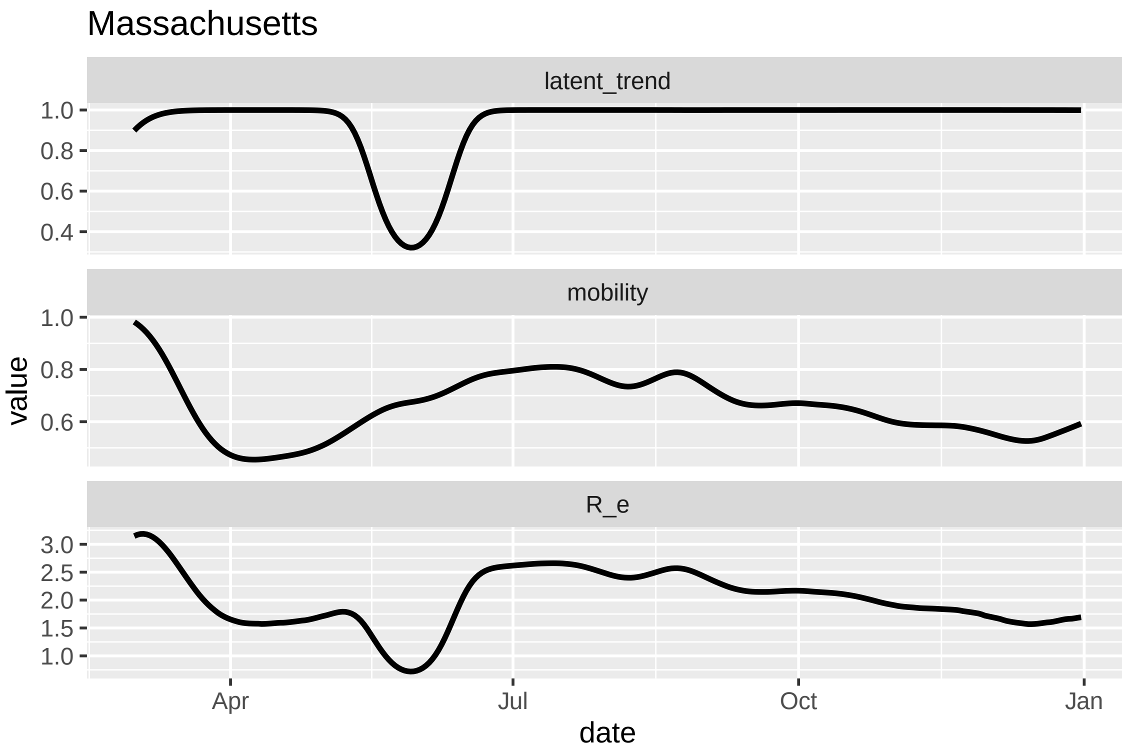

Kansas
latent_trend
1.0
0.8
0.6
0.4
mobility
1.0
0.9
0.8
0.7
0.6
value
R_e
3
2
1
Apr
Jul
Oct
Jan
date

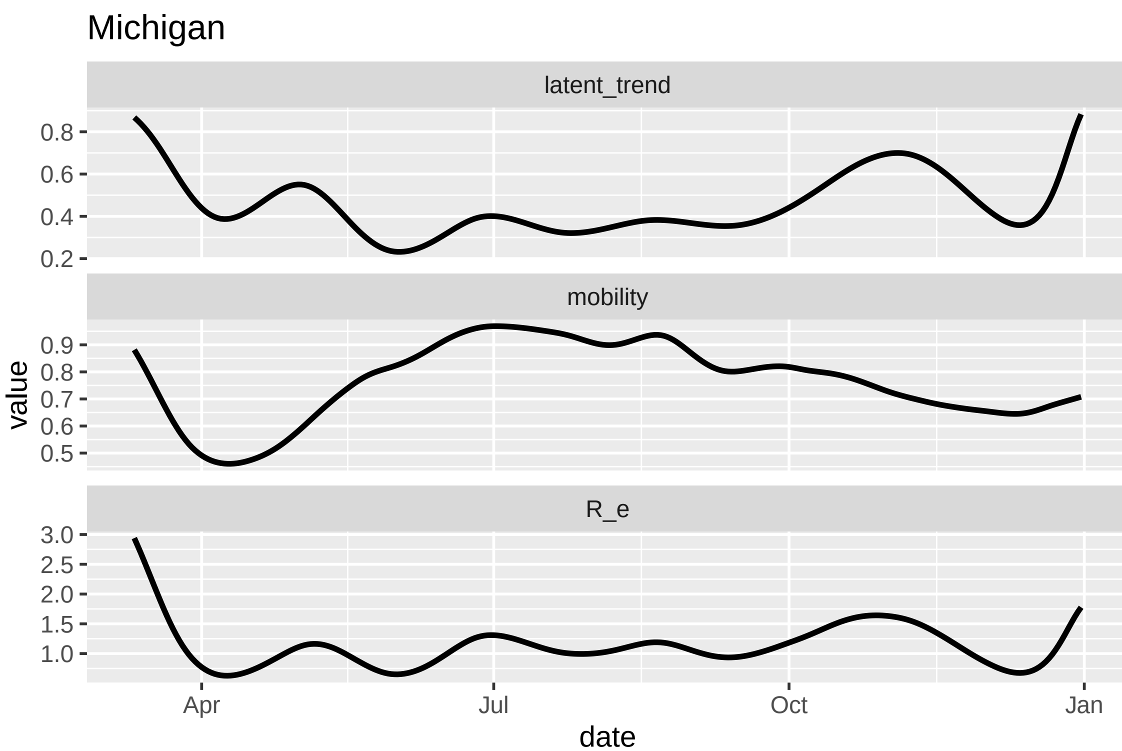

Kentucky
latent_trend
mobility
R_e
value
date
Apr
Jul
Oct
Jan

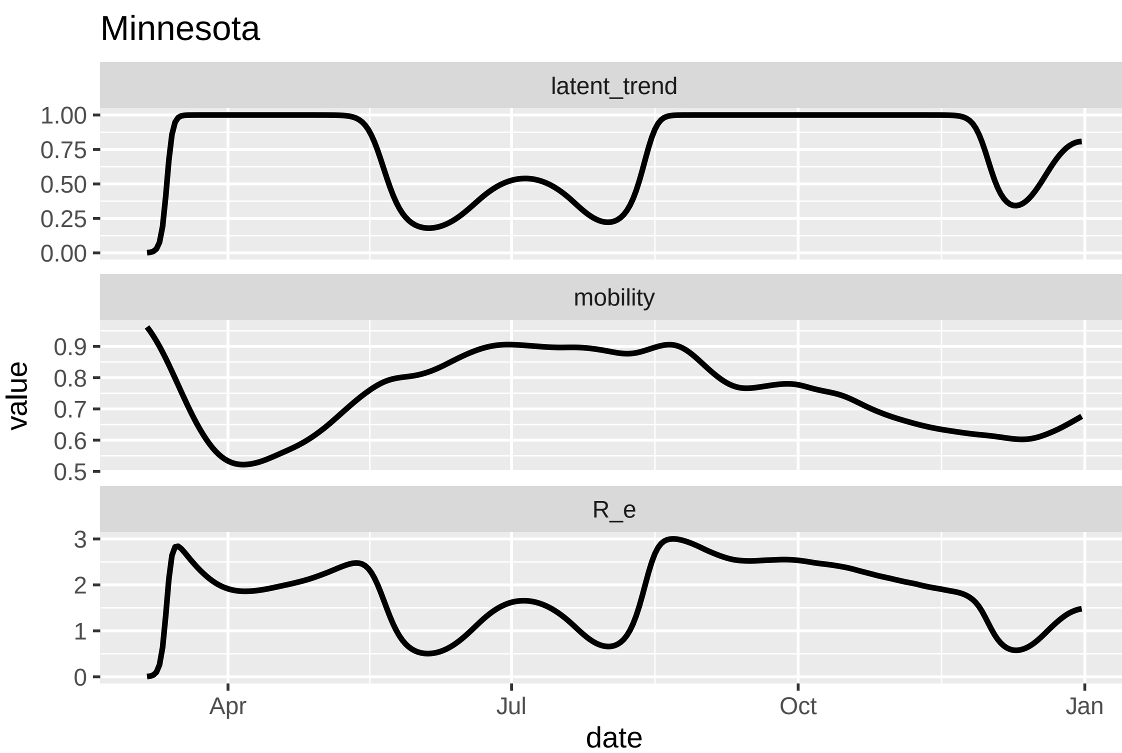

Louisiana
latent_trend
mobility
R_e
value
date
Apr
Jul
Oct
Jan

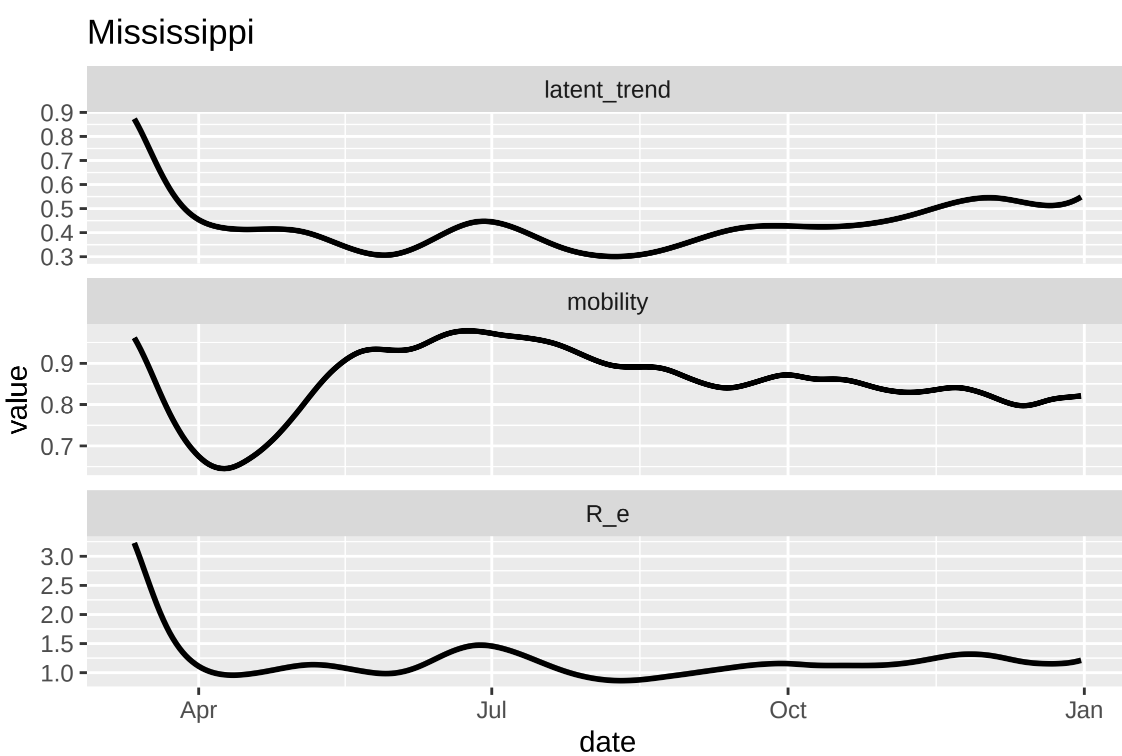

Maine
latent_trend
1.00
0.75
0.50
0.25
mobility
0.9
0.8
0.7
0.6
R_e
2.0
1.5
1.0
0.5
value
Apr
Jul
Oct
Jan
date

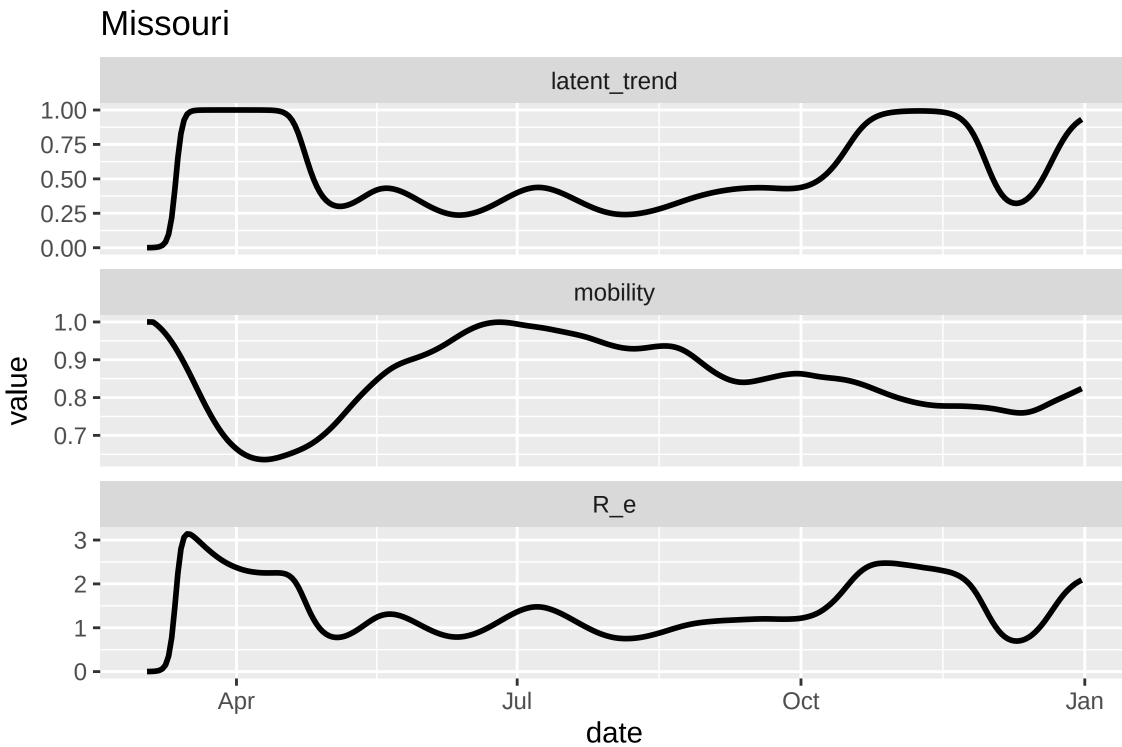

Maryland
latent_trend
1.00
0.75
0.50
0.25
mobility
0.9
0.8
0.7
0.6
0.5
R_e
2.5
2.0
1.5
1.0
0.5
value
Apr
Jul
Oct
Jan
date

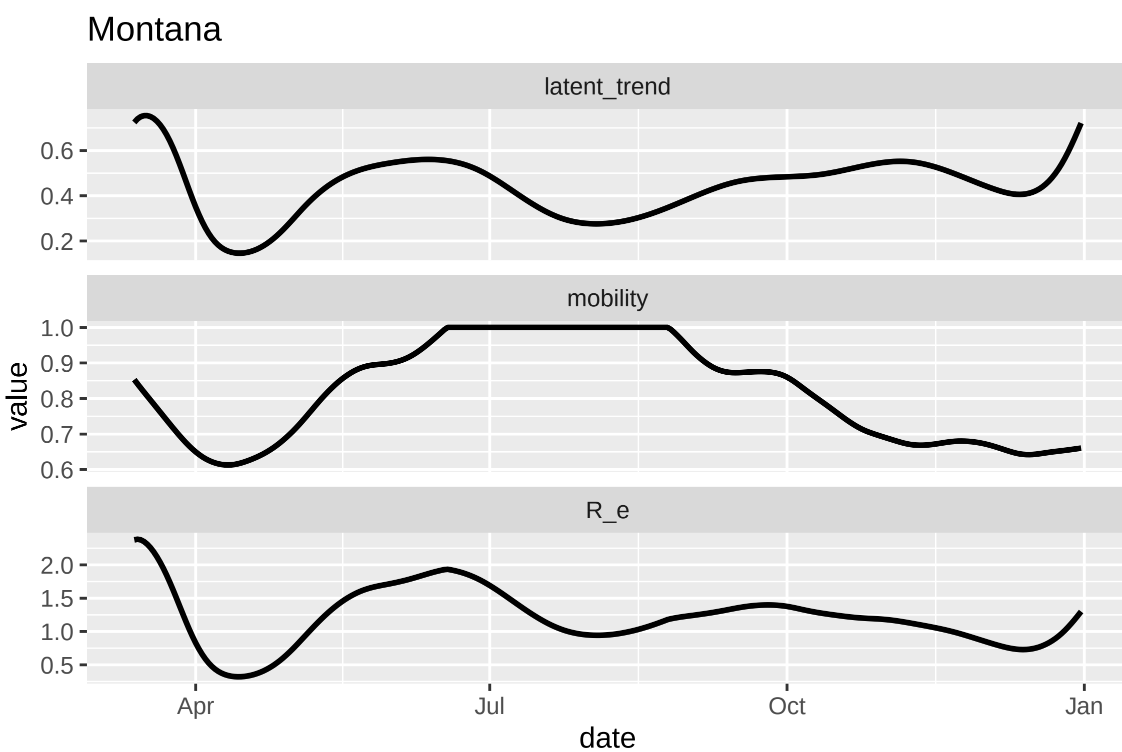

Massachusetts
latent_trend
mobility
R_e
value
1.0
0.8
0.6
0.4
3.0
2.5
2.0
1.5
Apr
Jul
Oct
Jan
date

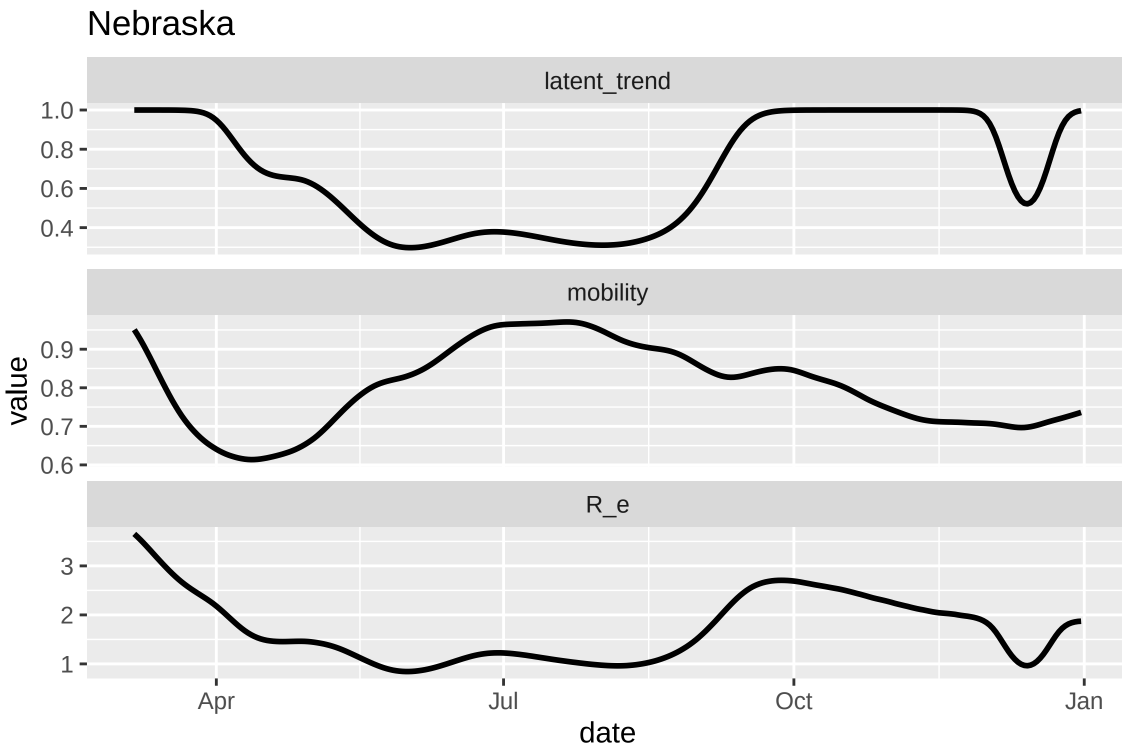

Michigan
latent_trend
mobility
R_e
value
0.8
0.6
0.4
0.2
0.9
0.7
0.5
3.0
2.5
2.0
1.5
1.0
Apr
Jul
Oct
Jan
date

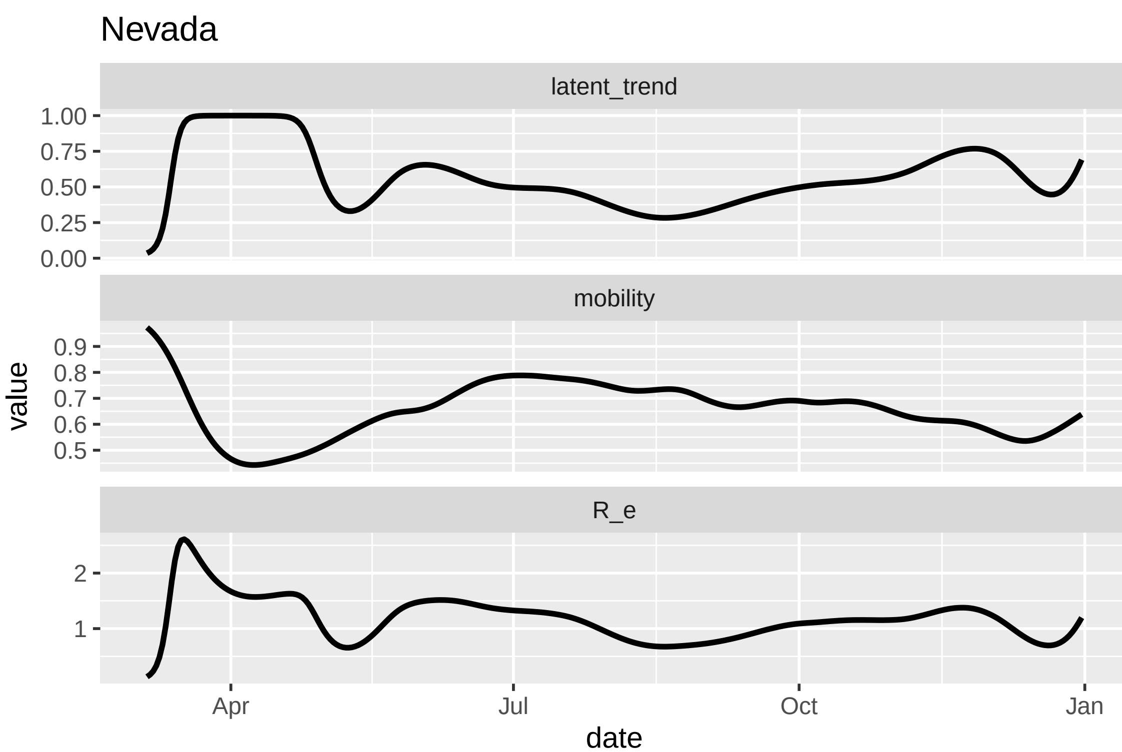

Minnesota
latent_trend
1.00
0.75
0.50
0.25
0.00
mobility
0.9
0.8
0.7
0.6
0.5
R_e
3
2
1
0
value
Apr
Jul
Oct
Jan
date

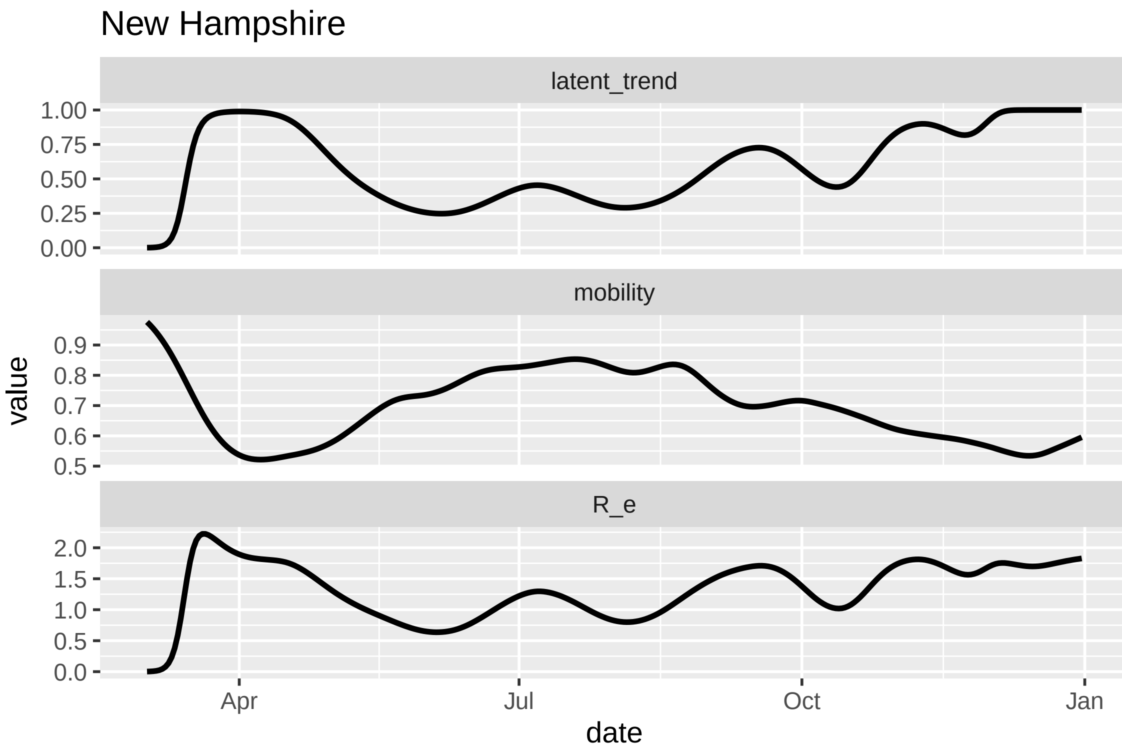

Mississippi
latent_trend
0.9
0.8
0.7
0.6
0.5
0.4
0.3
mobility
0.9
0.8
0.7
R_e
3.0
2.5
2.0
1.5
1.0
value
Apr
Jul
Oct
Jan
date

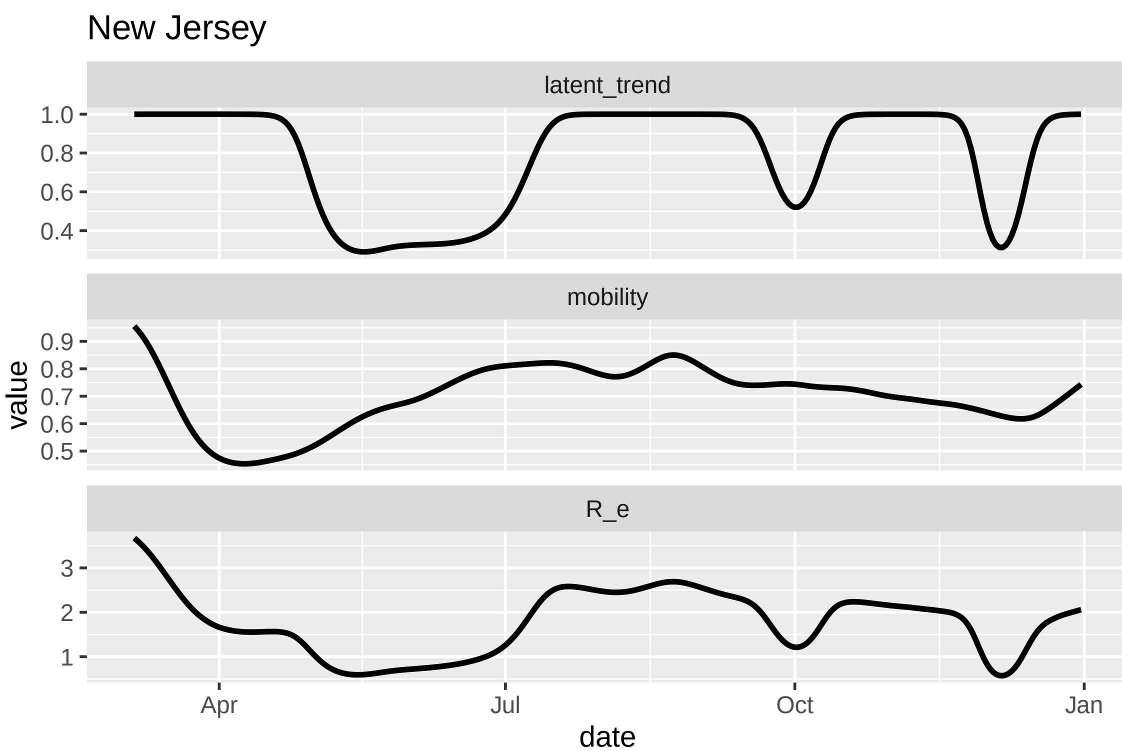

Missouri
latent_trend
1.00
0.75
0.50
0.25
0.00
mobility
1.0
0.9
0.8
0.7
value
R_e
3
2
1
0
Apr
Jul
Oct
Jan
date

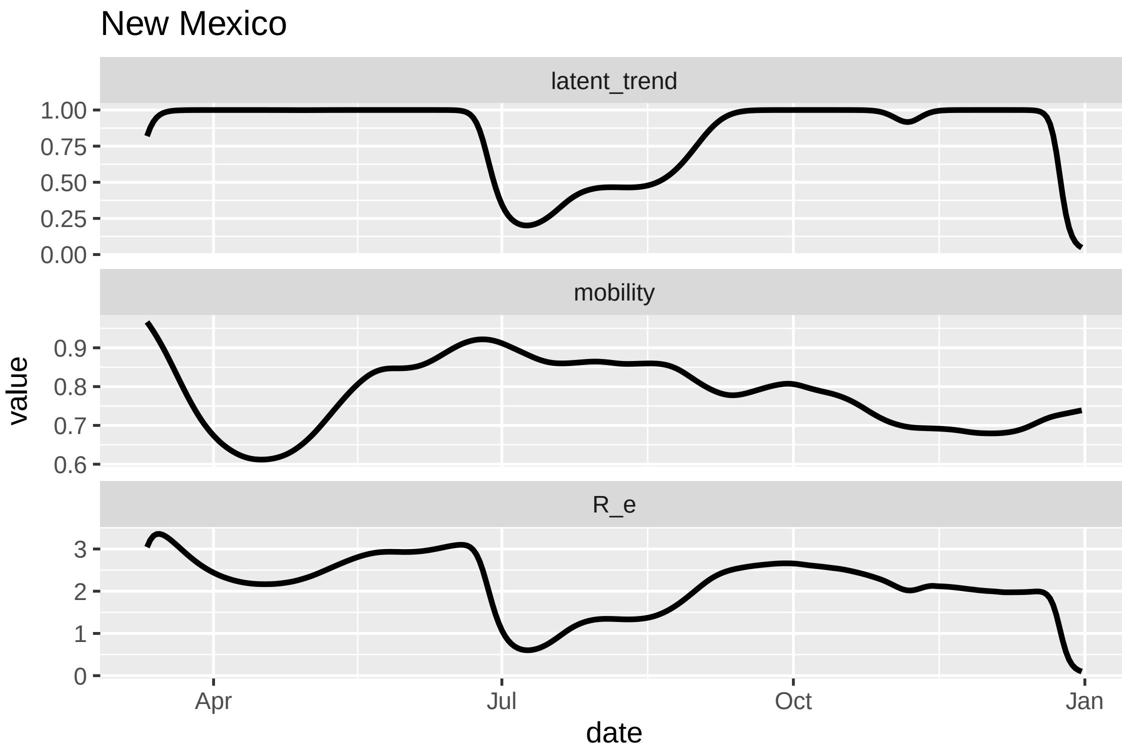

Montana
latent_trend
0.6
0.4
0.2
mobility
1.0
0.9
0.8
0.7
0.6
value
R_e
2.0
1.5
1.0
0.5
Apr
Jul
Oct
Jan
date

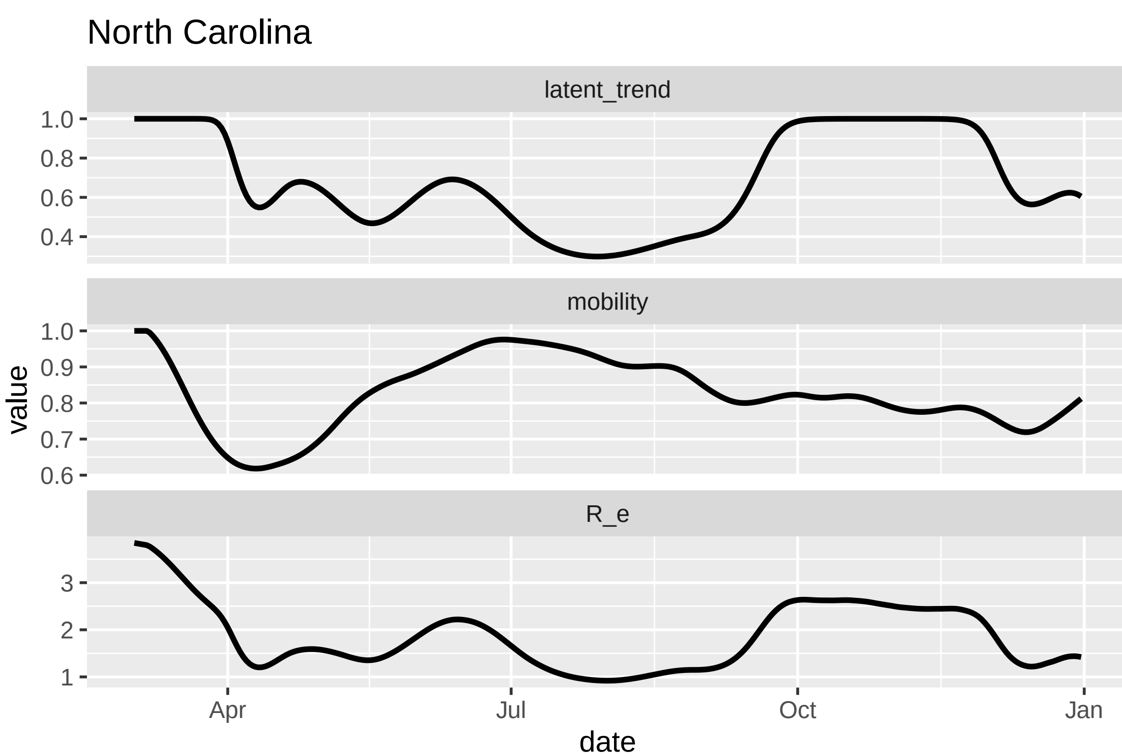

Nebraska
latent_trend
mobility
R_e
value
1.0
0.8
0.6
0.4
0.9
0.8
0.7
0.6
3
2
1
Apr
Jul
Oct
Jan
date

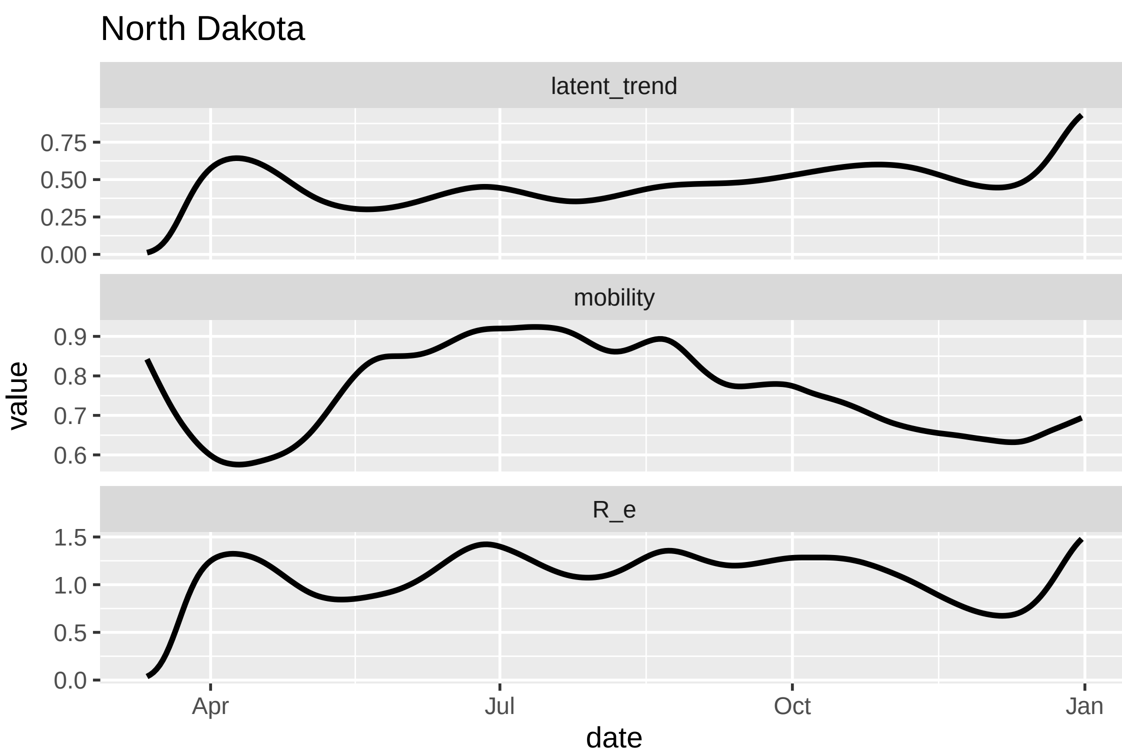

Nevada
latent_trend
mobility
R_e
value
1.00
0.75
0.50
0.25
0.00
0.9
0.8
0.7
0.6
0.5
2
1
Apr
Jul
Oct
Jan
date

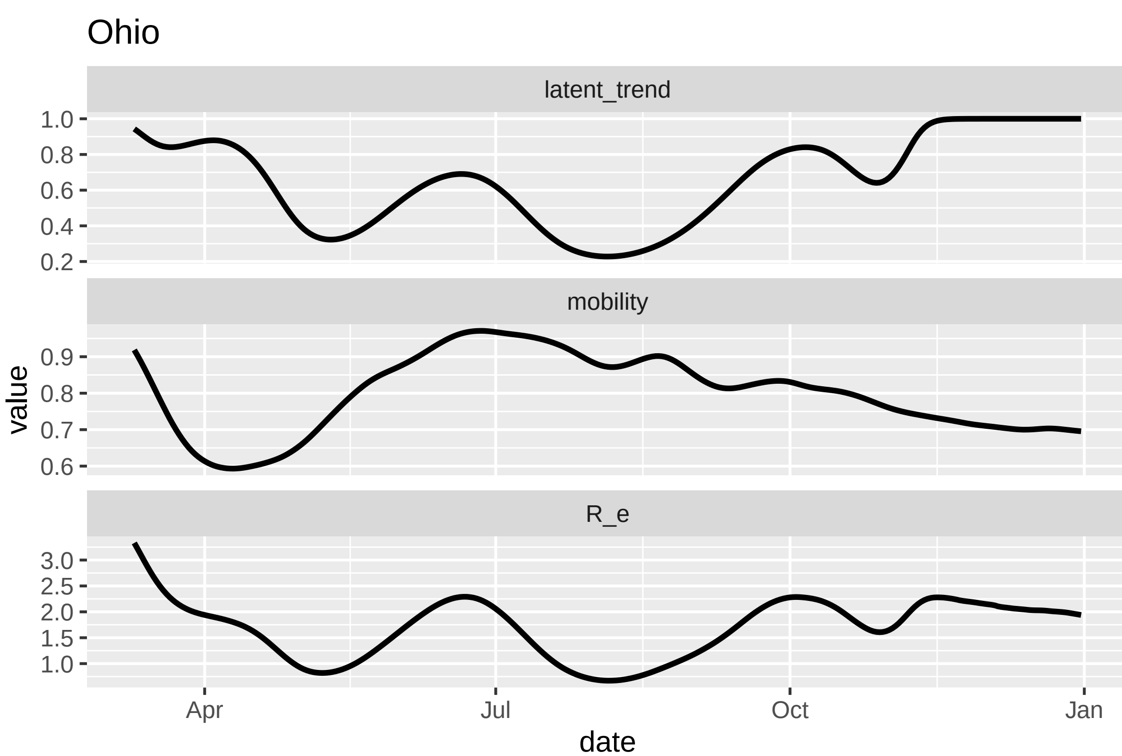

New Hampshire
latent_trend
mobility
R_e
value
1.00
0.75
0.50
0.25
0.00
0.9
0.8
0.7
0.6
0.5
2.0
1.5
1.0
0.5
0.0
Apr
Jul
Oct
Jan
date

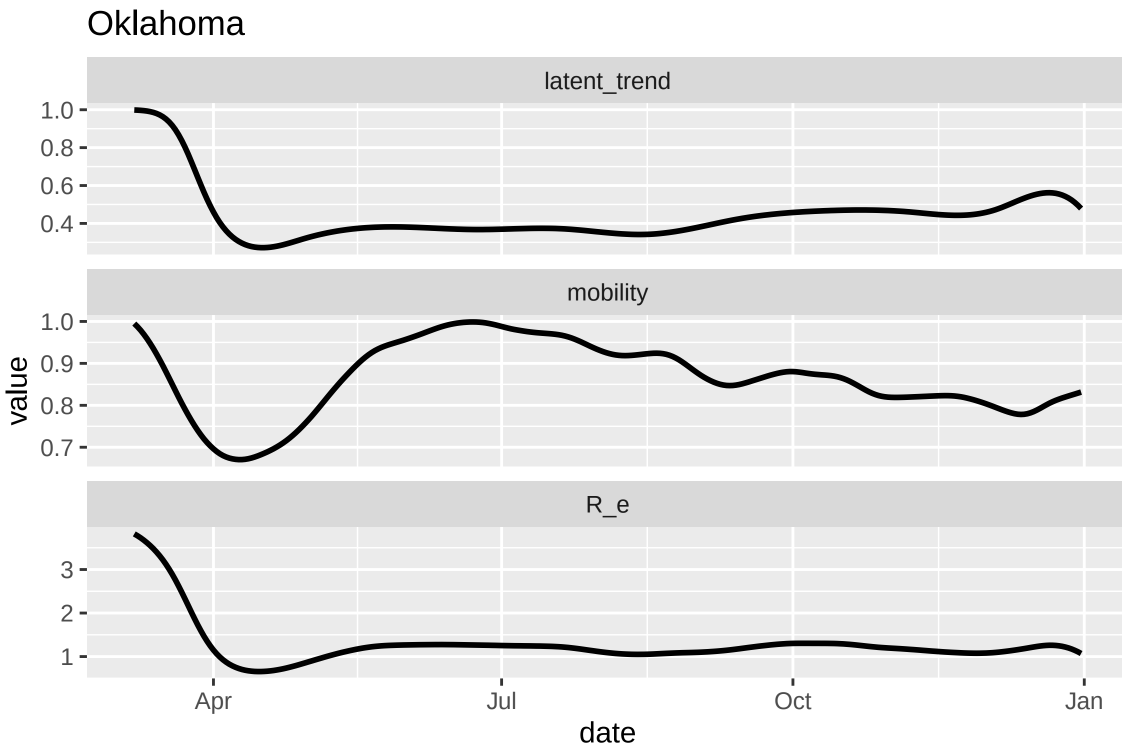

New Jersey
latent_trend
mobility
R_e
value
1.0
0.8
0.6
0.4
0.9
0.8
0.7
0.6
0.5
3
2
1
Apr
Jul
Oct
Jan
date

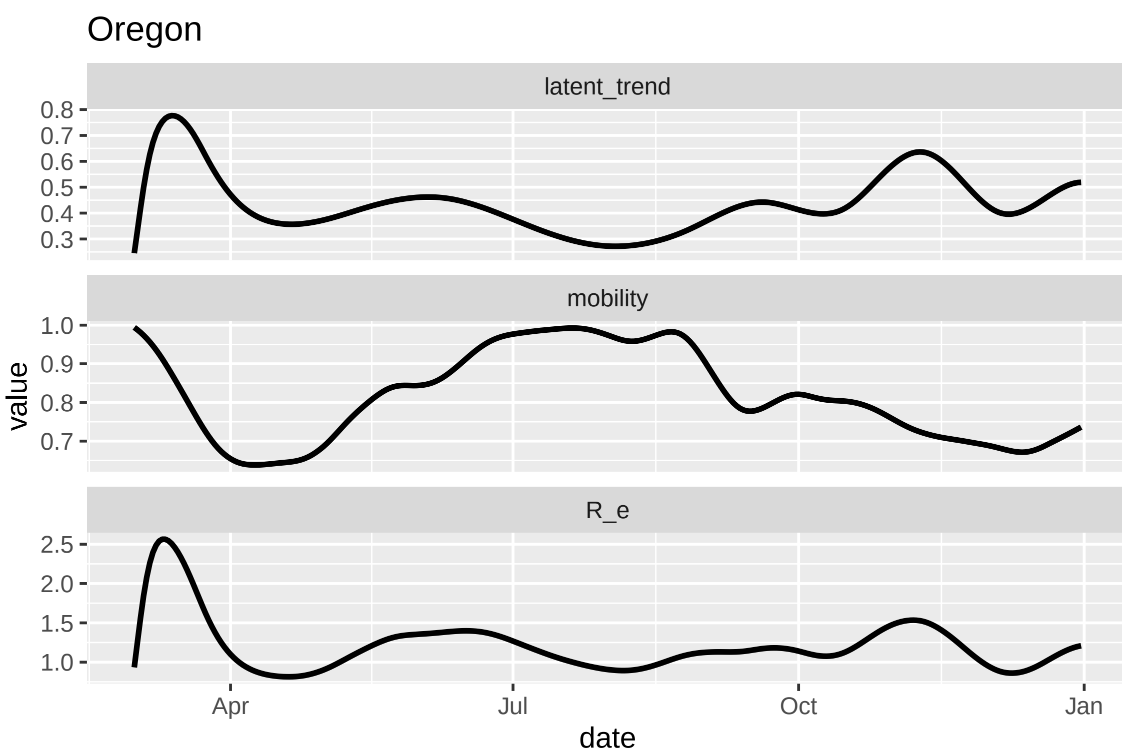

New Mexico
latent_trend
1.00
0.75
0.50
0.25
0.00
mobility
0.9
0.8
0.7
0.6
value
R_e
3
2
1
0
Apr
Jul
Oct
Jan
date

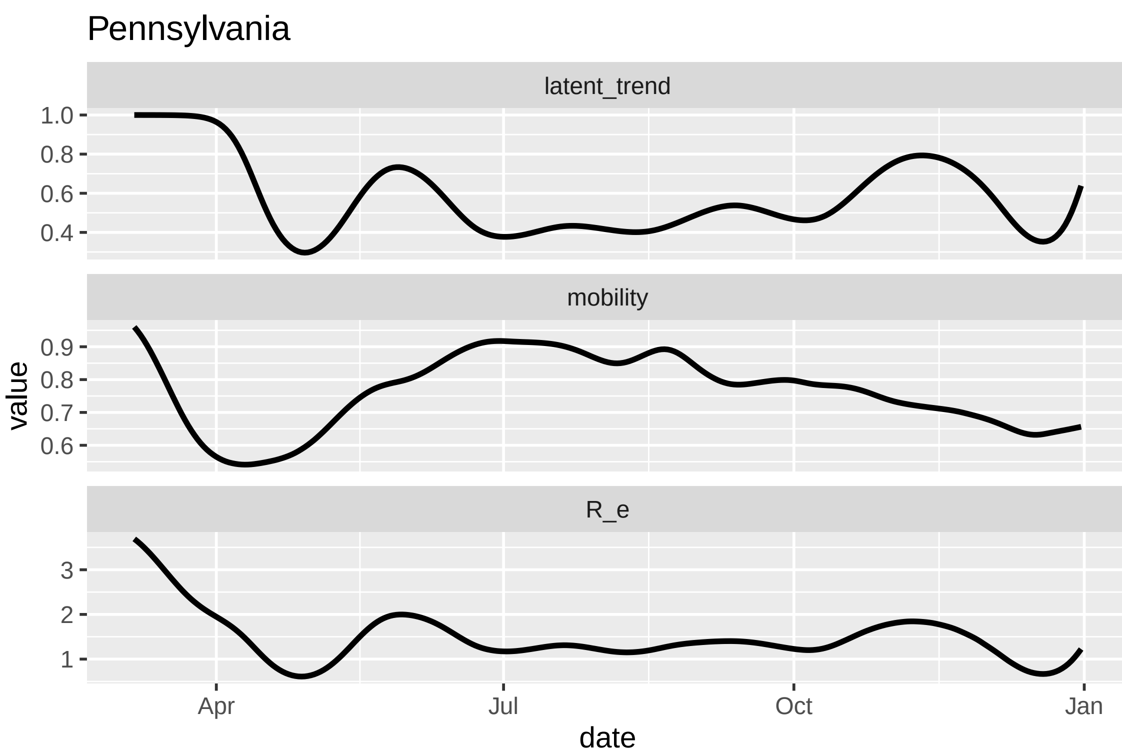

North Carolina
latent_trend
1.0
0.8
0.6
0.4
mobility
1.0
0.9
0.8
0.7
0.6
value
R_e
3
2
1
Apr
Jul
Oct
Jan
date


North Dakota
latent_trend
0.75
0.50
0.25
0.00
mobility
0.9
0.8
0.7
0.6
value
R_e
1.5
1.0
0.5
0.0
Apr
Jul
Oct
Jan
date


Ohio
latent_trend
1.0
0.8
0.6
0.4
0.2
mobility
0.9
0.8
0.7
0.6
value
R_e
3.0
2.5
2.0
1.5
1.0
Apr
Jul
Oct
Jan
date


Oklahoma
latent_trend
mobility
R_e
value
1.0
0.8
0.6
0.4
1.0
0.9
0.8
0.7
3
2
1
Apr
Jul
Oct
Jan
date


Oregon
latent_trend
mobility
R_e
value
0.8
0.7
0.6
0.5
0.4
0.3
1.0
0.9
0.8
0.7
2.5
2.0
1.5
1.0
Apr
Jul
Oct
Jan
date


Pennsylvania
latent_trend
1.0
0.8
0.6
0.4
mobility
0.9
0.8
0.7
0.6
value
R_e
3
2
1
Apr
Jul
Oct
Jan
date

References

Li, Ruiyun, Sen Pei, Bin Chen, Yimeng Song, Tao Zhang, Wan Yang, and Jeffrey Shaman. 2020. "Substantial Undocumented Infection Facilitates the Rapid Dissemination of Novel Coronavirus (SARS-CoV-2)." *Science* 368 (6490): 489–93. https://doi.org/10.1126/science.abb3221.

Mizumoto, Kenji, Katsushi Kagaya, Alexander Zarebski, and Gerardo Chowell. 2020. "Estimating the Asymptomatic Proportion of Coronavirus Disease 2019 (COVID-19) Cases on Board the Diamond Princess Cruise Ship, Yokohama, Japan, 2020." *Eurosurveillance* 25 (10): 2000180. https://doi.org/10.2807/1560-7917.ES.2020.25.10.2000180.

Ohsfeldt, Robert L., Casey Kar-Chan Choong, Patrick L. Mc Collam, Hamed Abedtash, Kari A. Kelton, and Russel Burge. 2021. "Inpatient Hospital Costs for COVID-19 Patients in the United States." *Advances in Therapy* 38 (11): 5557–95. https://doi.org/10.1007/s12325-021-01887-4.

Sanche, Steven, Yen Ting Lin, Chonggang Xu, Ethan Romero-Severson, Nick Hengartner, and Ruian Ke. 2020. "High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2 - Volume 26, Number 7—July 2020 - Emerging Infectious Diseases Journal - CDC." https://doi.org/10.3201/eid2607.200282.

Verity, Robert, Lucy C. Okell, Ilaria Dorigatti, Peter Winskill, Charles Whittaker, Natsuko Imai, Gina Cuomo-Dannenburg, et al. 2020. "Estimates of the Severity of Coronavirus Disease 2019: A Model-Based Analysis." *The Lancet Infectious Diseases* 20 (6): 669–77. https://doi.org/10.1016/S1473-3099(20)30243-7.

Zhang, Juanjuan, Maria Litvinova, Wei Wang, Yan Wang, Xiaowei Deng, Xinghui Chen, Mei Li, et al. 2020. "Evolving Epidemiology and Transmission Dynamics of Coronavirus Disease 2019 Outside Hubei Province, China: A Descriptive and Modelling Study." *The Lancet. Infectious Diseases* 20 (7): 793–802. https://doi.org/10.1016/S1473-3099(20)30230-9.