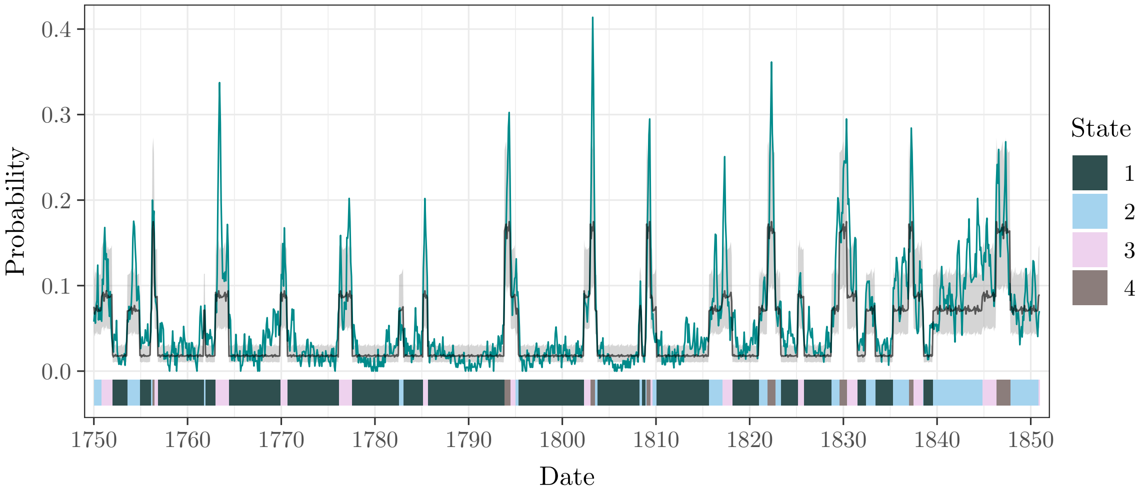
Results from the hidden Markov model with four states

Supplementary Table 1: Posterior means and 95% posterior intervals of the transition probabilities in the transition matrix A when $S = 4$. Rows represent the states to transfer from and the columns the states to transfer to.

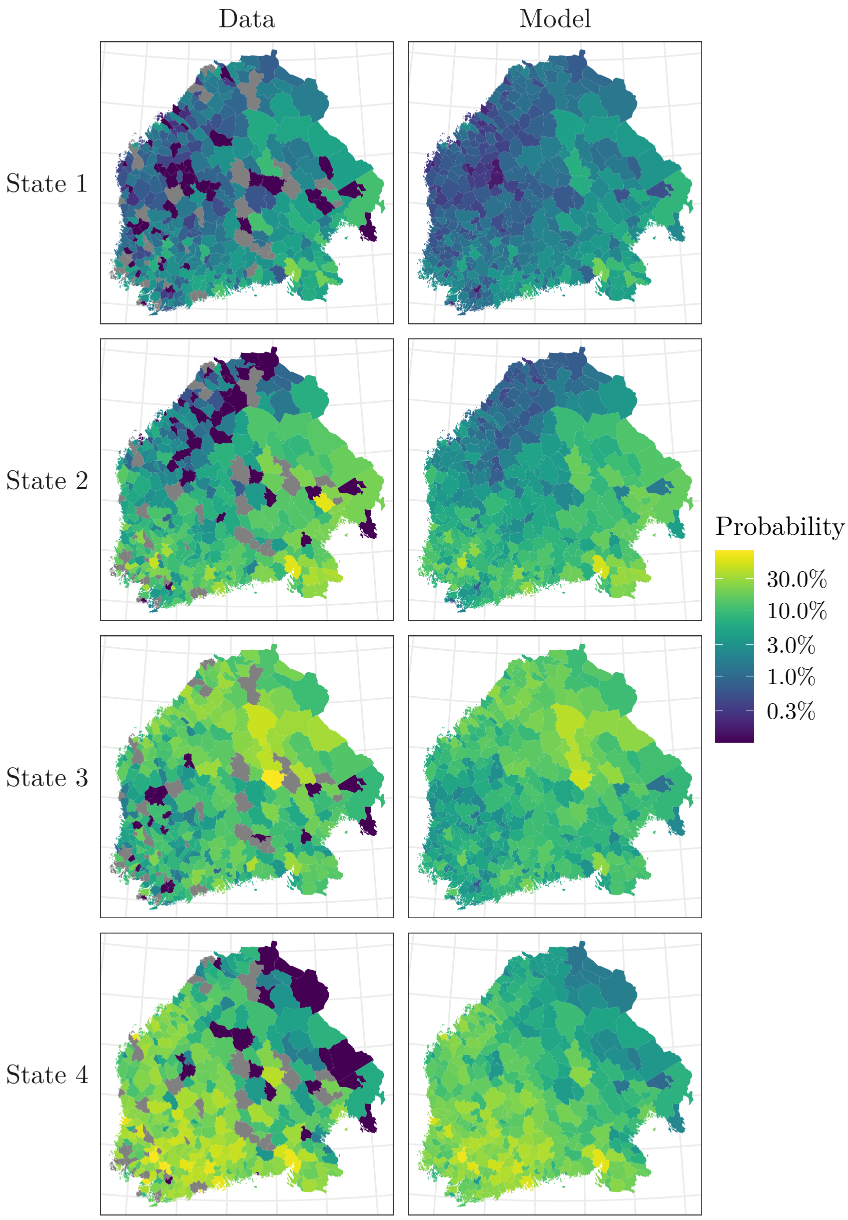
	State 1	State 2	State 3	State 4
State 1	0.97 (0.96, 0.98)	0.02 (0.01, 0.03)	0.01 (0.00, 0.02)	0.00 (0.00, 0.01)
State 2	0.04 (0.02, 0.08)	0.91 (0.87, 0.95)	0.02 (0.01, 0.05)	0.02 (0.01, 0.05)
State 3	0.06 (0.03, 0.11)	0.02 (0.00, 0.05)	0.90 (0.84, 0.94)	0.02 (0.00, 0.04)
State 4	0.01 (0.00, 0.03)	0.05 (0.01, 0.11)	0.07 (0.02, 0.14)	0.87 (0.79, 0.94)

Supplementary Table 2: Posterior means and standard deviations, 95% posterior intervals, bulk and tail effective sample sizes (ess), and $\widehat{R}$ statistics of the initial probabilities ρ_s, the state specific constants μ_s, the deviation of local constants σ_λ, and the deviation parameter τ of the ICAR component for four-state HMM.

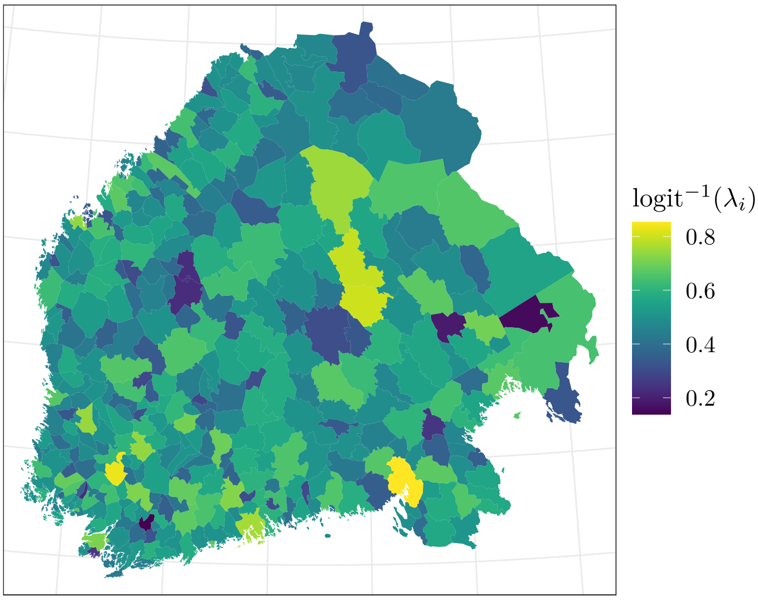
	mean	sd	2.5%	97.5%	ess bulk	ess tail	$\widehat{R}$
ρ_1	0.20	0.17	0.01	0.61	18148	5248	1.00
ρ_2	0.40	0.20	0.06	0.81	22665	6353	1.00
ρ_3	0.20	0.17	0.01	0.61	18114	5805	1.00
ρ_4	0.20	0.16	0.01	0.60	14960	6157	1.00
μ_1	-4.25	0.03	-4.31	-4.20	6044	7379	1.00
μ_2	-2.84	0.03	-2.90	-2.78	3700	7363	1.00
μ_3	-2.40	0.03	-2.46	-2.35	6028	6355	1.00
μ_4	-1.69	0.03	-1.75	-1.63	4175	7243	1.00
σ_λ	0.58	0.03	0.52	0.64	2527	5624	1.00
τ	0.76	0.03	0.70	0.83	1287	2718	1.00



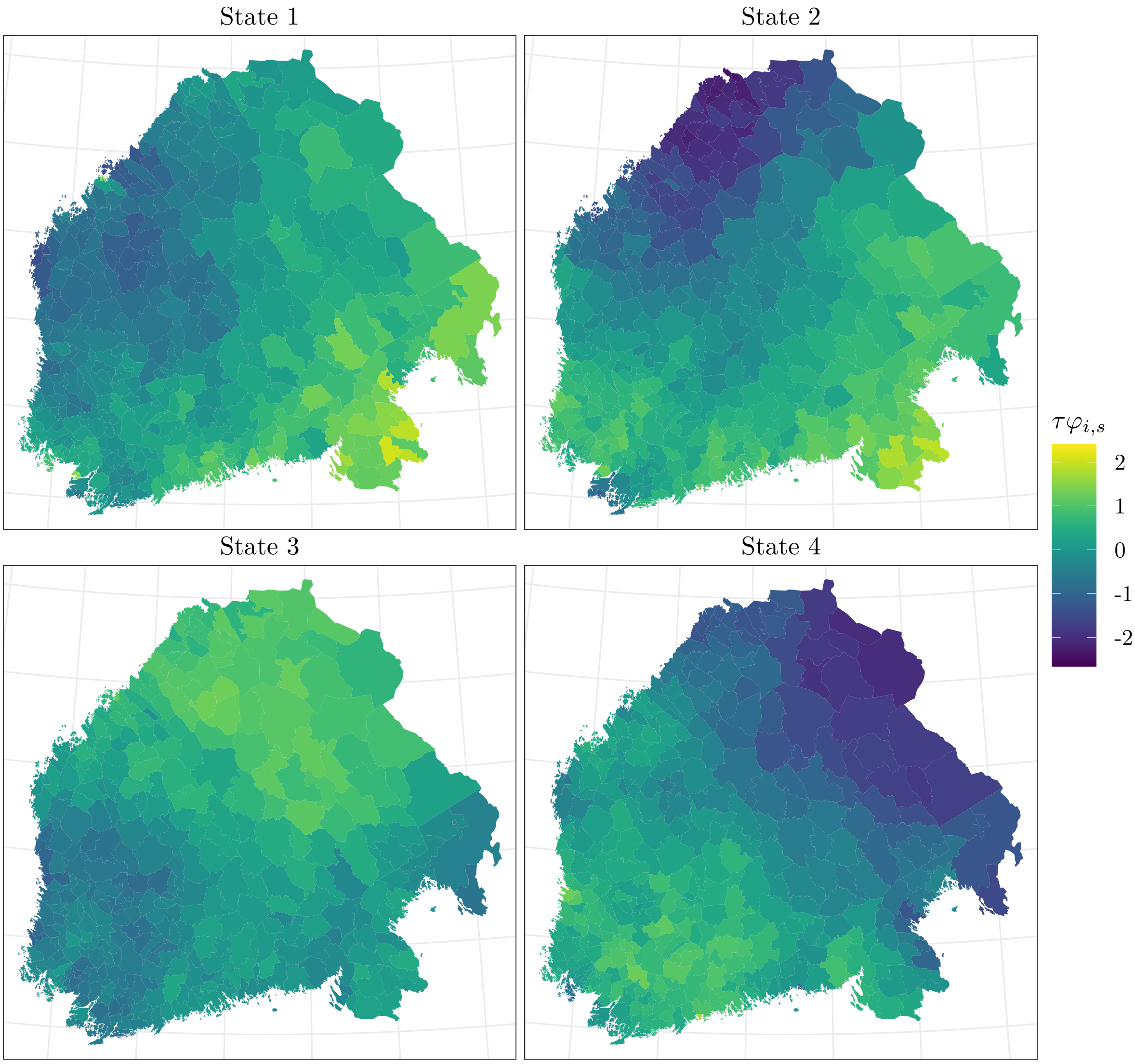
Supplementary Figure 1: The turquoise line denotes the proportion of towns where at least one death caused by measles was observed in that month according to the data. The black line represents the nationwide average over local probabilities to observe at least one death caused by measles given by the four-state HMM: $\text{logit}^{-1}(\mu_{x_t} + \lambda_i + \tau\varphi_{i,x_t} + \gamma_t)$. The grey area around the black line shows the 95% posterior interval of the predicted average. The colors below the curve indicate in which state it was most likely to be at that time.



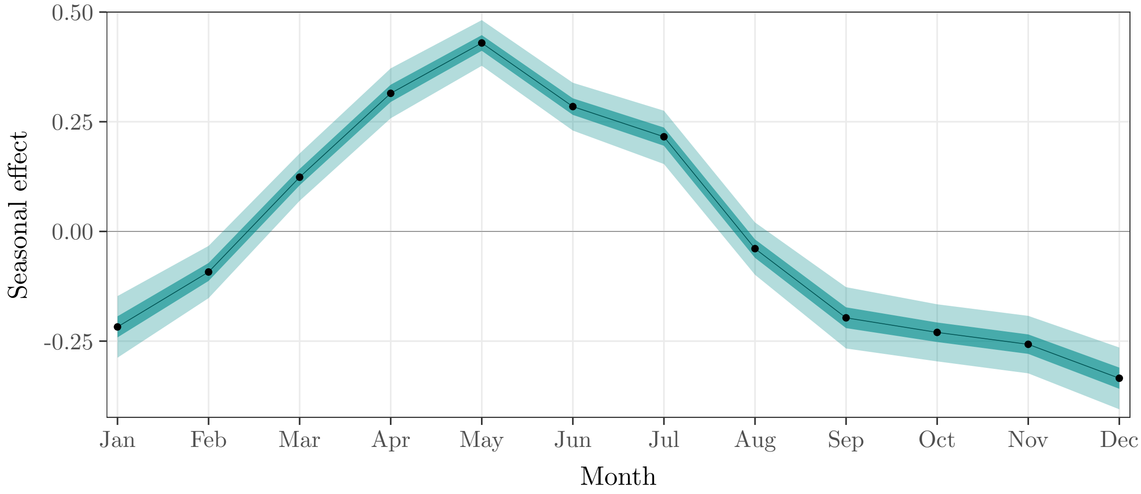
Supplementary Figure 2: On the left column are the probabilities to to observe at least one death caused by measles in each town as means over the months when each state was most likely, and on the right column are the corresponding estimated probabilities $\text{logit}^{-1}(\mu_{x_t} + \lambda_i + \tau\varphi_{i,x_t} + \gamma_t)$ from the four-state HMM. The rows represent different states. Gray area indicates a site with only missing observations in months corresponding to that state. Note that the color scale is logarithmic.



Supplementary Figure 3: Inverse logit transformations of the posterior means of the local constants λ_i.



Supplementary Figure 4: Posterior means of the spatial terms $\tau\varphi_{i,s}$ by state.



Supplementary Figure 5: The black dots and line show the posterior mean of the monthly seasonal term γ_t. The dark and pale turquoise areas represent the 50% and 95% posterior intervals, respectively.

Leave-one-out cross validation for hidden Markov models

Exact leave-one-out cross validation (LOO-CV) is not possible for many Bayesian models, such as one considered here, as re-estimation of the model hundreds or thousands of times is computationally infeasible. Instead, approximate LOO-CV using Pareto smoothed importance sampling allows computation of LOO-CV metrics based on the posterior samples obtained from the full data model. For this, for each posterior sample of the model parameters, we need to compute the leave-one-out log-likelihood for observation y_{it}, $i = 1, \dots, N$, $t = 1, \dots, T$. For the hidden Markov model this is

$$\begin{aligned} \log P(y_{it} \mid Y^{-it}) &= \log \sum_{s=1}^{S} P(y_{it} \mid x_t = s) P(x_t = s \mid Y^{-it}) \\ &= \log \sum_{s=1}^{S} \omega_{s,t}(i) \gamma_t^{-it}(s), \end{aligned} \tag{1}$$

where Y^{-it} is the full data without $y_{i,t}$, $\omega_{s,t}(i) = P(y_{it} \mid x_t = s)$ is the emission probability of y_{it}, and $\gamma_t^{-it}(s) = P(x_t = s \mid Y^{-it})$ is the posterior probability (as in HMM literature) of state x_t without considering y_{it}.

By Bayes' rule and conditional independence of observations given the hidden states, we have

$$\begin{aligned} \gamma_t(s) &= P(x_t \mid Y^{-it}, y_{it}) \\ &= \frac{P(y_{it} \mid x_t = s) P(x_t = s \mid Y^{-it})}{P(y_{it} \mid Y^{-it})} \\ &= \frac{\omega_{s,t}(i) \gamma_t^{-it}(s)}{P(y_{it} \mid Y^{-it})}, \end{aligned} \tag{2}$$

from which we can solve

$$\begin{aligned} \gamma_t^{-it}(s) &= \frac{\gamma_t(s) P(y_{it} \mid Y^{-it})}{\omega_{s,t}(i)} \\ &\propto \frac{\gamma_t(s)}{\omega_{s,t}(i)}. \end{aligned} \tag{3}$$

Then,

$$\begin{aligned} P(y_{it} \mid Y^{-it}) &= \sum_{s=1}^{S} \omega_{s,t}(i) \frac{\frac{\gamma_t(s)}{\omega_{s,t}(i)}}{\sum_{s'=1}^{S} \frac{\gamma_t(s')}{\omega_{s',t}(i)}} \\ &= \sum_{s=1}^{S} \frac{\gamma_t(s)}{\sum_{s'=1}^{S} \frac{\gamma_t(s')}{\omega_{s',t}(i)}} \\ &= \frac{1}{\sum_{s'=1}^{S} \frac{\gamma_t(s')}{\omega_{s',t}(i)}}. \end{aligned} \tag{4}$$

Therefore, the log-likelihood for the left out observation can be computed using the posterior probabilities $\gamma_t(s)$ from the full data model, and the relevant emission probabilities $\omega_{s,t}(i)$. If any $\omega_{s,t}(i)$ are zero, those terms should be omitted from the sum.

Supplementary Table 3: Differences of the ELPDs and the standard errors of the ELPD differences for the leave-one-out cross-validation of the four-state and five-state models.

States	$\mathrm{ELPD}_{\mathrm{diff}}$	$\mathrm{SE}_{\mathrm{diff}}$
5	0.0	0.0
4	-566.8	42.7