

Lecture 22

Spatio-temporal Models

Colin Rundel

11/28/2018

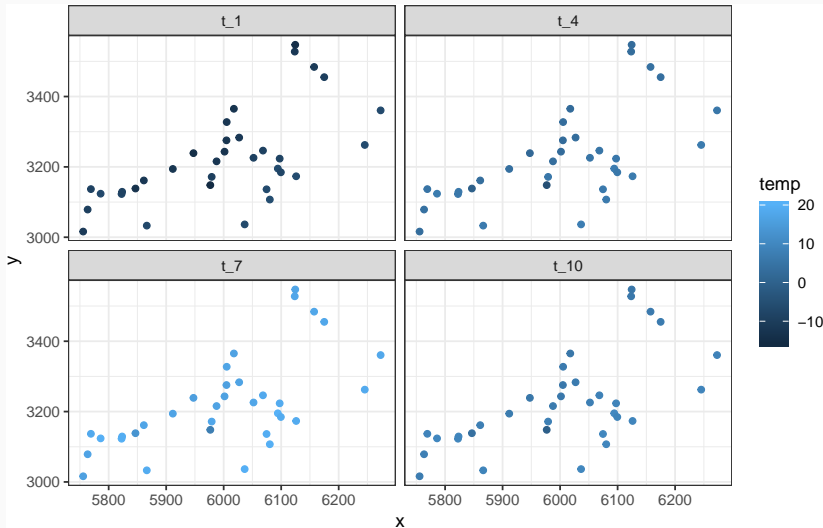
Spatial Models with AR time dependence

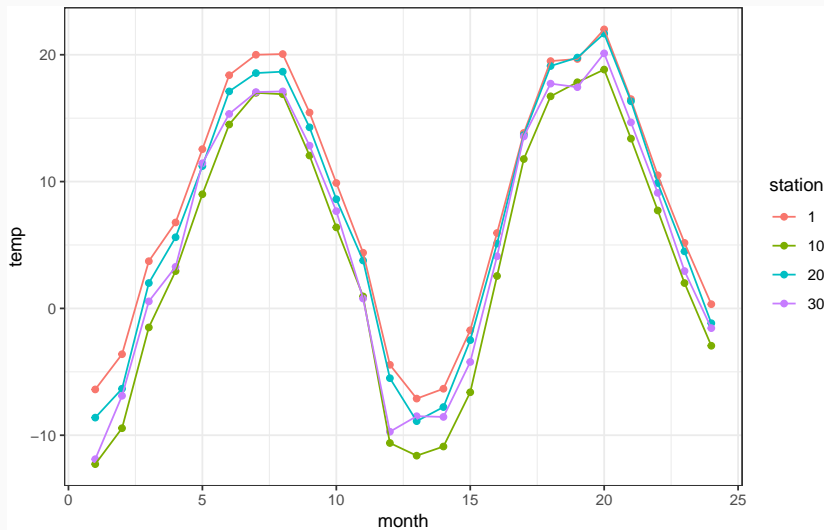
Example - Weather station data

Based on Andrew Finley and Sudipto Banerjee's notes from National Ecological Observatory Network (NEON) Applied Bayesian Regression Workshop, March 7 - 8, 2013 Module 6

NETemp.dat - Monthly temperature data (Celsius) recorded across the Northeastern US starting in January 2000.

```
## # A tibble: 34 x 27
##       x     y elev  t_1  t_2  t_3  t_4  t_5  t_6  t_7  t_8
##   <dbl> <dbl> <int> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1  6094. 3195.   102  -6.39 -3.61  3.72  6.78 12.6 18.4 20   20.1
## 2  6245. 3262.    1  -6.28 -4.11  2.61  6.56 11.4 16.8 18.4 18.7
## 3  6157. 3484.   157 -11.1  -9.44 -0.389  3.94  9.89 15.4 17.5 17.4
## 4  6124. 3528.   176 -11.6  -9.72 -1.17  2.89  9.67 14.8 17.4 16.9
## 5  6005. 3275.   400 -12.6  -9.06 -1.61  2.56  8.56 14.3 15.9 15.8
## 6  6052. 3226.   133  -9.11  -6.39  1.22  4.94 10.9 15.9 17.3 17.6
## 7  6099. 3185.    56  -7.94  -6.06  2.06  5.56 11.1 17   18.6 18.8
## 8  6075. 3136.    59  -6.56  -3.5  3.17  6.17 11.5 17.4 19.1 19.4
## 9  6175. 3455.   160  -9.94  -8.94 -0.278  3.56  9.61 15.3 17.7 17.3
## 10 6005. 3327.   360 -12.3  -9.44 -1.5  2.94  9   14.5 17   16.9
## # ... with 24 more rows, and 16 more variables: t_9 <dbl>, t_10 <dbl>,
## #   t_11 <dbl>, t_12 <dbl>, t_13 <dbl>, t_14 <dbl>, t_15 <dbl>,
## #   t_16 <dbl>, t_17 <dbl>, t_18 <dbl>, t_19 <dbl>, t_20 <dbl>,
## #   t_21 <dbl>, t_22 <dbl>, t_23 <dbl>, t_24 <dbl>
```





$$y_t = \mathbf{F}'_t \boldsymbol{\theta}_t + v_t \quad \text{observation equation}$$

$1 \times p$ $p \times 1$

$$\boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t \quad \text{evolution equation}$$

$p \times 1$ $p \times p$ $p \times 1$ $p \times 1$

$$\mathbf{v}_t \sim \mathcal{N}(0, \mathbf{V}_t)$$

$$\boldsymbol{\omega}_t \sim \mathcal{N}(0, \mathbf{W}_t)$$

ARMA / ARIMA are a special case of a dynamic linear model, for example an $AR(p)$ can be written as

$$F'_t = (1, 0, \dots, 0)$$

$$G_t = \begin{pmatrix} \phi_1 & \phi_2 & \dots & \phi_{p-1} & \phi_p \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

$$\omega_t = (\omega_1, 0, \dots, 0), \quad \omega_1 \sim \mathcal{N}(0, \sigma^2)$$

ARMA / ARIMA are a special case of a dynamic linear model, for example an $AR(p)$ can be written as

$$F'_t = (1, 0, \dots, 0)$$

$$G_t = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

$$\omega_t = (\omega_1, 0, \dots, 0), \quad \omega_1 \sim \mathcal{N}(0, \sigma^2)$$

$$y_t = \theta_t + v_t, \quad v_t \sim \mathcal{N}(0, \sigma_v^2)$$

$$\theta_t = \sum_{i=1}^p \phi_i \theta_{t-i} + \omega_1, \quad \omega_1 \sim \mathcal{N}(0, \sigma_\omega^2)$$

Dynamic spatio-temporal model

The observed temperature at time t and location s is given by $y_t(s)$ where,

$$y_t(\mathbf{s}) = \mathbf{x}_t(\mathbf{s})\boldsymbol{\beta}_t + u_t(\mathbf{s}) + \epsilon_t(\mathbf{s})$$

$$\epsilon_t(\mathbf{s}) \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \tau_t^2)$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t$$

$$\boldsymbol{\eta}_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \boldsymbol{\Sigma}_\eta)$$

$$u_t(\mathbf{s}) = u_{t-1}(\mathbf{s}) + w_t(\mathbf{s})$$

$$w_t(\mathbf{s}) \stackrel{i.i.d.}{\sim} \mathcal{N}(\mathbf{0}, \Sigma_t(\phi_t, \sigma_t^2))$$

Dynamic spatio-temporal model

The observed temperature at time t and location s is given by $y_t(s)$ where,

$$y_t(\mathbf{s}) = \mathbf{x}_t(\mathbf{s})\boldsymbol{\beta}_t + u_t(\mathbf{s}) + \epsilon_t(\mathbf{s})$$

$$\epsilon_t(\mathbf{s}) \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \tau_t^2)$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t$$

$$\boldsymbol{\eta}_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \boldsymbol{\Sigma}_\eta)$$

$$u_t(\mathbf{s}) = u_{t-1}(\mathbf{s}) + w_t(\mathbf{s})$$

$$w_t(\mathbf{s}) \stackrel{i.i.d.}{\sim} \mathcal{N}(\mathbf{0}, \Sigma_t(\phi_t, \sigma_t^2))$$

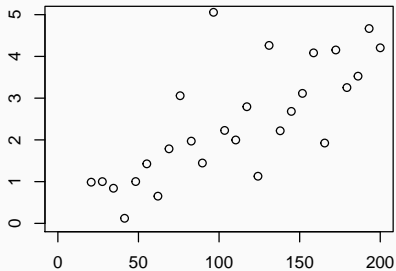
Additional assumptions for $t = 0$,

$$\boldsymbol{\beta}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$$

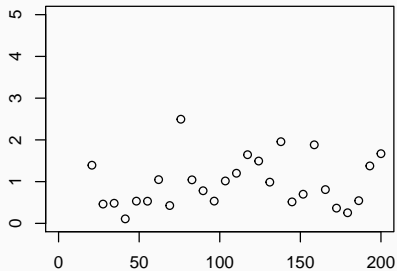
$$u_0(\mathbf{s}) = 0$$

Variograms by time

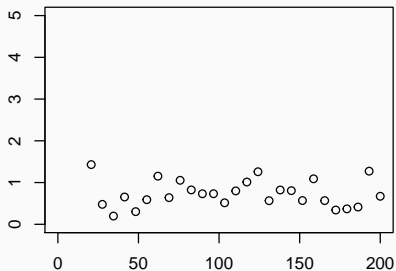
Jan 2000



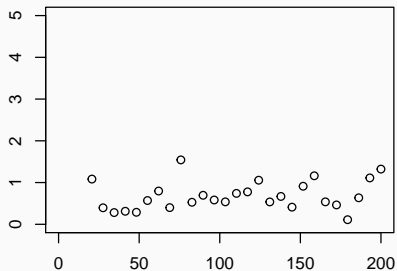
Apr 2000



Jul 2000



Oct 2000



Data:

```
max_d = coords %>% dist() %>% max()
n_t = 24
n_s = nrow(ne_temp)
```

Parameters:

```
n_beta = 2
starting = list(
  beta = rep(0, n_t * n_beta), phi = rep(3/(max_d/4), n_t),
  sigma.sq = rep(1, n_t), tau.sq = rep(1, n_t),
  sigma.eta = diag(0.01, n_beta)
)
tuning = list(phi = rep(1, n_t))
priors = list(
  beta.0.Norm = list(rep(0, n_beta), diag(1000, n_beta)),
  phi.Unif = list(rep(3/(0.9 * max_d), n_t), rep(3/(0.05 * max_d), n_t)),
  sigma.sq.IG = list(rep(2, n_t), rep(2, n_t)),
  tau.sq.IG = list(rep(2, n_t), rep(2, n_t)),
  sigma.eta.IW = list(2, diag(0.001, n_beta))
)
```

Fitting with spDynLM from spBayes

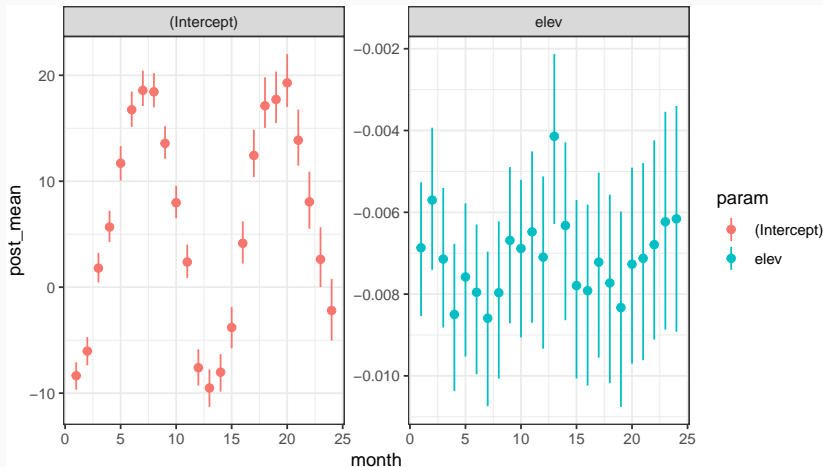
```
n_samples = 10000
models = lapply(paste0("t_",1:24, "~elev"), as.formula)

m = spBayes::spDynLM(
  models, data = ne_temp, coords = coords, get.fitted = TRUE,
  starting = starting, tuning = tuning, priors = priors,
  cov.model = "exponential", n.samples = n_samples, n.report = 1000)

m = clean_spdynlm(m, n_samples/2+1, n_samples, (n_samples/2)/1000)
save(m, file="dynlm.Rdata")

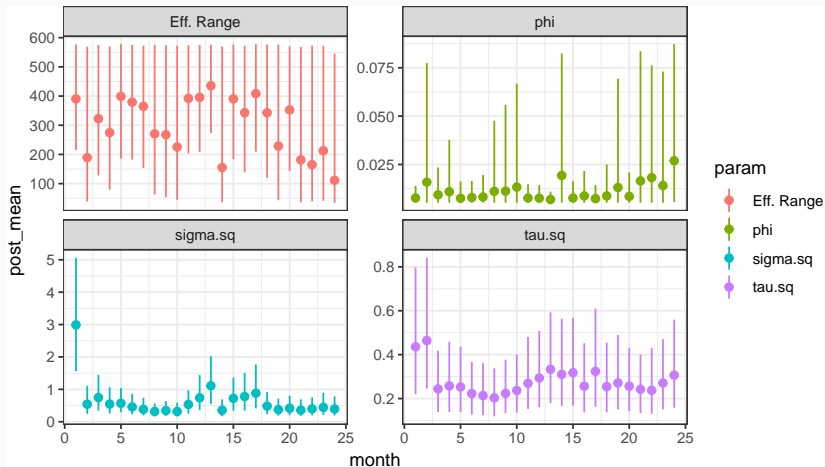
## -----
##      General model description
## -----
## Model fit with 34 observations in 24 time steps.
##
## Number of missing observations 0.
##
## Number of covariates 2 (including intercept if specified).
##
## Using the exponential spatial correlation model.
##
## Number of MCMC samples 10000.
##
## ...
```

Posterior Inference - β s

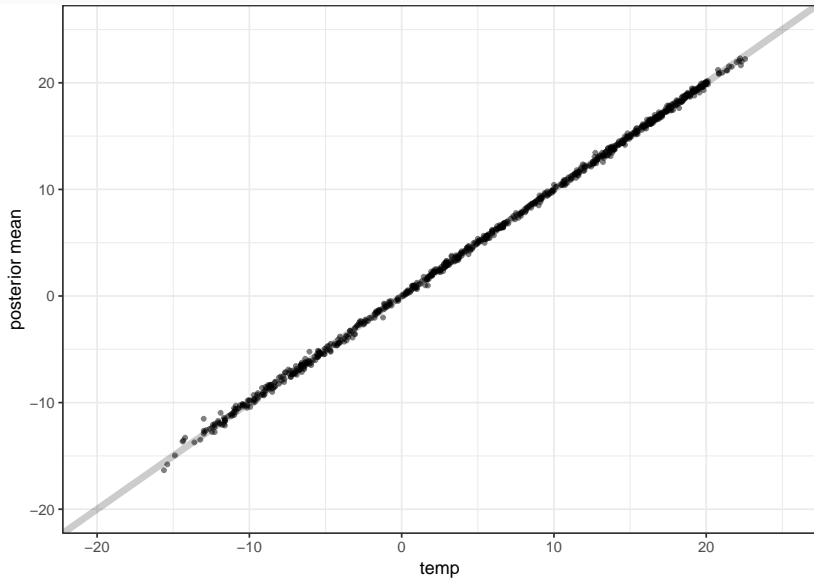


Lapse Rate $\approx -9.8^\circ\text{C}/\text{km}$.

Posterior Inference - θ



Posterior Inference - Observed vs. Predicted



`spPredict` does not support `spDynLM` objects but it will impute missing values.

```
r = raster(xmn=5750, xmx=6300, ymn=3000, ymx=3550, nrow=20, ncol=20)

pred = xyFromCell(r, 1:length(r)) %>%
  as.data.frame() %>%
  mutate(type="pred") %>%
  bind_rows(
    ne_temp %>% mutate(type = "obs"),
    .
  )
```

```

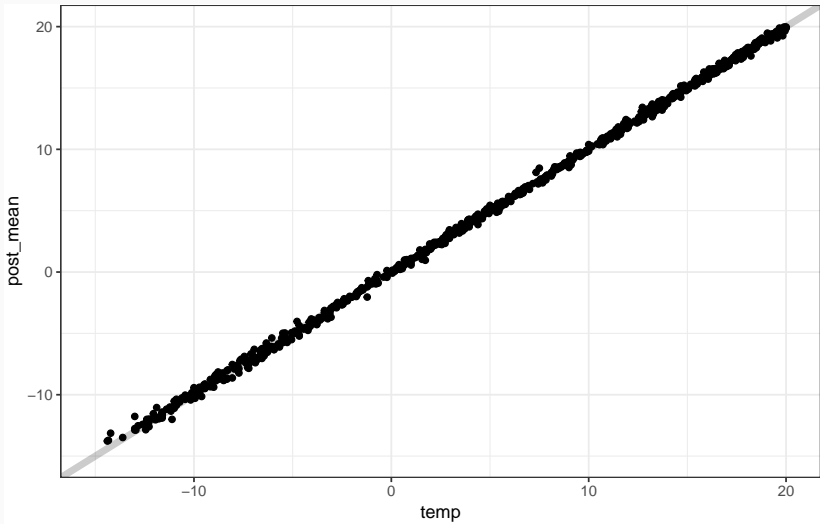
models_pred = lapply(paste0("t_",1:n_t, "~1"), as.formula)

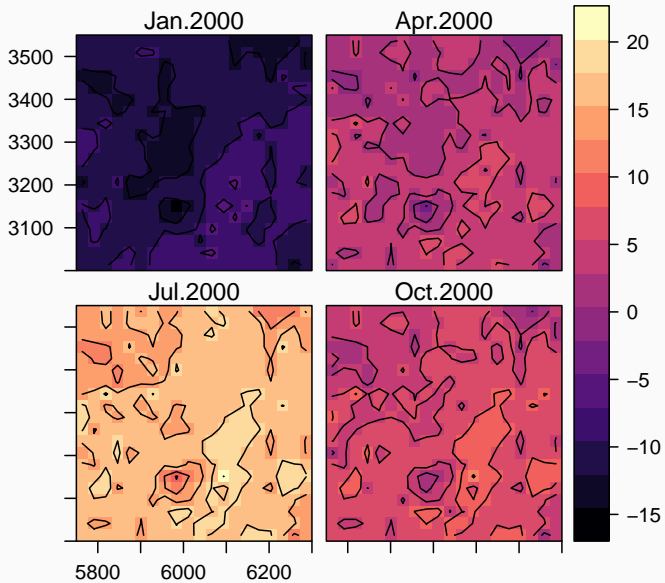
n_samples = 5000
m_pred = spBayes::spDynLM(
  models_pred, data = pred, coords = coords_pred, get.fitted = TRUE,
  starting = starting, tuning = tuning, priors = priors,
  cov.model = "exponential", n.samples = n_samples, n.report = 1000)

m_pred = clean_spdynlm(m_pred, n_samples/2+1, n_samples, thin = 5)
save(m_pred, file="dynlm_pred.Rdata")

## -----
## General model description
## -----
## Model fit with 434 observations in 24 time steps.
##
## Number of missing observations 9600.
##
## Number of covariates 1 (including intercept if specified).
##
## Using the exponential spatial correlation model.
##
## Number of MCMC samples 5000.

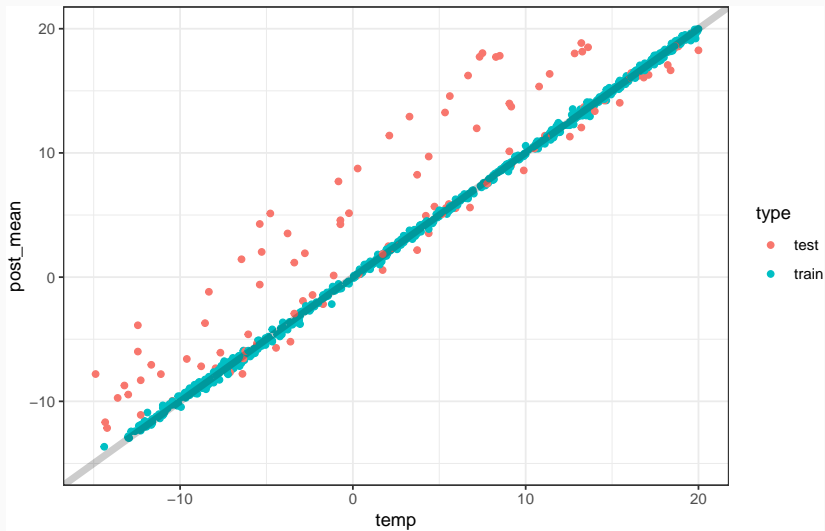
```





Out-of-sample validation

```
## # A tibble: 34 x 29
##       x     y elev type station  t_1 t_10 t_11 t_12 t_13 t_14
##   <dbl> <dbl> <int> <chr> <int> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 6094. 3195.  102 test     1  NA   NA   NA   NA   NA   NA
## 2 6245. 3262.    1 train    2 -6.28 8.89 3.89 -4.22 -7.11 -5.06
## 3 6157. 3484.  157 train    3 -11.1 6.44 1.94 -8.72 -11.6 -10.4
## 4 6124. 3528.  176 train    4 -11.6 5.94 1.67 -9.17 -11.8 -11
## 5 6005. 3275.  400 train    5 -12.6 5.67 0.278 -10.7 -11.9 -11
## 6 6052. 3226.  133 train    6 -9.11 7.56 2.44 -7.11 -9.44 -8.61
## 7 6099. 3185.   56 test     7  NA   NA   NA   NA   NA   NA
## 8 6075. 3136.   59 train    8 -6.56 9.61 4.17 -4.89 -6.06 -5.06
## 9 6175. 3455.  160 train    9 -9.94 6.67 1.72 -8.44 -12.1 -10.7
## 10 6005. 3327.  360 train   10 -12.3 6.39 0.944 -10.6 -
11.6 -10.9
## # ... with 24 more rows, and 18 more variables: t_15 <dbl>, t_16 <dbl>,
## # t_17 <dbl>, t_18 <dbl>, t_19 <dbl>, t_2 <dbl>, t_20 <dbl>, t_21 <dbl>,
## # t_22 <dbl>, t_23 <dbl>, t_24 <dbl>, t_3 <dbl>, t_4 <dbl>, t_5 <dbl>,
## # t_6 <dbl>, t_7 <dbl>, t_8 <dbl>, t_9 <dbl>
```



Spatio-temporal models for continuous time

In general, spatiotemporal models will have a form like the following,

$$\begin{aligned} y(\mathbf{s}, t) &= \underbrace{\mu(\mathbf{s}, t)}_{\text{mean structure}} + \underbrace{e(\mathbf{s}, t)}_{\text{error structure}} \\ &= \underbrace{\mathbf{x}(\mathbf{s}, t) \boldsymbol{\beta}(\mathbf{s}, t)}_{\text{Regression}} + \underbrace{w(\mathbf{s}, t)}_{\text{Spatiotemporal RE}} + \underbrace{\epsilon(\mathbf{s}, t)}_{\text{White Noise}} \end{aligned}$$

In general, spatiotemporal models will have a form like the following,

$$\begin{aligned} y(\mathbf{s}, t) &= \underbrace{\mu(\mathbf{s}, t)}_{\text{mean structure}} + \underbrace{e(\mathbf{s}, t)}_{\text{error structure}} \\ &= \underbrace{\mathbf{x}(\mathbf{s}, t) \boldsymbol{\beta}(\mathbf{s}, t)}_{\text{Regression}} + \underbrace{w(\mathbf{s}, t)}_{\text{Spatiotemporal RE}} + \underbrace{\epsilon(\mathbf{s}, t)}_{\text{White Noise}} \end{aligned}$$

The simplest possible spatiotemporal model is one where we assume there is no dependence between observations in space and time,

$$w(\mathbf{s}, t) = \alpha(t) + \omega(\mathbf{s})$$

these are straight forward to fit and interpret but are quite limiting (no shared information between space and time).

Spatiotemporal Covariance

Lets assume that we want to define our spatiotemporal random effect to be a single stationary Gaussian Process (in 3 dimensions*),

$$\mathbf{w}(\mathbf{s}, \mathbf{t}) \sim \mathcal{N}(\mathbf{0}, \Sigma(\mathbf{s}, \mathbf{t}))$$

where our covariance function depends on both $\|s - s'\|$ and $|t - t'|$,

$$\text{cov}(\mathbf{w}(\mathbf{s}, \mathbf{t}), \mathbf{w}(\mathbf{s}', \mathbf{t}')) = c(\|s - s'\|, |t - t'|)$$

Spatiotemporal Covariance

Lets assume that we want to define our spatiotemporal random effect to be a single stationary Gaussian Process (in 3 dimensions*),

$$\mathbf{w}(\mathbf{s}, \mathbf{t}) \sim \mathcal{N}(\mathbf{0}, \Sigma(\mathbf{s}, \mathbf{t}))$$

where our covariance function depends on both $\|s - s'\|$ and $|t - t'|$,

$$\text{cov}(\mathbf{w}(\mathbf{s}, \mathbf{t}), \mathbf{w}(\mathbf{s}', \mathbf{t}')) = c(\|s - s'\|, |t - t'|)$$

- Note that the resulting covariance matrix Σ will be of size $n_s \cdot n_t \times n_s \cdot n_t$.
 - Even for modest problems this gets very large (past the point of direct computability).
 - If $n_t = 52$ and $n_s = 100$ we have to work with a 5200×5200 covariance matrix

Separable Models

One solution is to use a separable form, where the covariance is the product of a valid 2d spatial and a valid 1d temporal covariance / correlation function,

$$\text{cov}(\mathbf{w}(\mathbf{s}, \mathbf{t}), \mathbf{w}(\mathbf{s}', \mathbf{t}')) = \sigma^2 \rho_1(\|\mathbf{s} - \mathbf{s}'\|; \boldsymbol{\theta}) \rho_2(|\mathbf{t} - \mathbf{t}'|; \phi)$$

Separable Models

One solution is to use a separable form, where the covariance is the product of a valid 2d spatial and a valid 1d temporal covariance / correlation function,

$$\text{cov}(\mathbf{w}(\mathbf{s}, \mathbf{t}), \mathbf{w}(\mathbf{s}', \mathbf{t}')) = \sigma^2 \rho_1(\|\mathbf{s} - \mathbf{s}'\|; \boldsymbol{\theta}) \rho_2(|\mathbf{t} - \mathbf{t}'|; \phi)$$

If we define our observations as follows (stacking time locations within spatial locations)

$$\mathbf{w}(\mathbf{s}, \mathbf{t}) = (w(\mathbf{s}_1, t_1), \dots, w(\mathbf{s}_1, t_{n_t}), \dots, w(\mathbf{s}_{n_s}, t_1), \dots, w(\mathbf{s}_{n_s}, t_{n_t}))^t$$

then the covariance can be written as

$$\underset{n_s n_t \times n_s n_t}{\boldsymbol{\Sigma}_w(\sigma^2, \theta, \phi)} = \sigma^2 \underset{n_s \times n_s}{\mathbf{H}_s(\theta)} \otimes \underset{n_t \times n_t}{\mathbf{H}_t(\phi)}$$

where $\mathbf{H}_s(\theta)$ and $\mathbf{H}_t(\phi)$ are correlation matrices defined by

$$\{\mathbf{H}_s(\theta)\}_{ij} = \rho_1(\|\mathbf{s}_i - \mathbf{s}_j\|; \theta)$$

$$\{\mathbf{H}_t(\phi)\}_{ij} = \rho_2(|t_i - t_j|; \phi)$$

Definition:

$$\mathbf{A}_{[m \times n]} \otimes \mathbf{B}_{[p \times q]} = \begin{pmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{pmatrix}_{[m \cdot p \times n \cdot q]}$$

Definition:

$$\underset{[m \times n]}{\mathbf{A}} \otimes \underset{[p \times q]}{\mathbf{B}} = \underset{[m \cdot p \times n \cdot q]}{\begin{pmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{pmatrix}}$$

Properties:

$$\begin{aligned} \mathbf{A} \otimes \mathbf{B} &\neq \mathbf{B} \otimes \mathbf{A} \quad (\text{usually}) \\ (\mathbf{A} \otimes \mathbf{B})^t &= \mathbf{A}^t \otimes \mathbf{B}^t \end{aligned}$$

$$\begin{aligned} \det(\mathbf{A} \otimes \mathbf{B}) &= \det(\mathbf{B} \otimes \mathbf{A}) \\ &= \det(\mathbf{A})^{\text{rank}(\mathbf{B})} \det(\mathbf{B})^{\text{rank}(\mathbf{A})} \end{aligned}$$

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$$

If we have a spatiotemporal random effect with a separable form,

$$\mathbf{w}(\mathbf{s}, \mathbf{t}) \sim \mathcal{N}(\mathbf{0}, \Sigma_w)$$

$$\Sigma_w = \sigma^2 \mathbf{H}_s \otimes \mathbf{H}_t$$

then the likelihood for \mathbf{w} is given by

$$\begin{aligned} & -\frac{n}{2} \log 2\pi - \frac{1}{2} \log |\Sigma_w| - \frac{1}{2} \mathbf{w}^t \Sigma_w^{-1} \mathbf{w} \\ = & -\frac{n}{2} \log 2\pi - \frac{1}{2} \log [(\sigma^2)^{n_t \cdot n_s} |H_s|^{n_t} |H_t|^{n_s}] - \frac{1}{2\sigma^2} \mathbf{w}^t (\mathbf{H}_s^{-1} \otimes \mathbf{H}_t^{-1}) \mathbf{w} \end{aligned}$$

- Additive and separable models are still somewhat limiting
- Cannot treat spatiotemporal covariances as 3d observations
- Possible alternatives:
 - Specialized spatiotemporal covariance functions, i.e.

$$\gamma(\mathbf{s}, \mathbf{s}', t, t') = \sigma^2(|t-t'|+1)^{-1} \exp(-\|\mathbf{s}-\mathbf{s}'\|(|t-t'|+1)^{-\beta/2})$$

- Mixtures of separable covariances, i.e.

$$w(\mathbf{s}, t) = w_1(\mathbf{s}, t) + w_2(\mathbf{s}, t)$$