A Conditional Approach to Multivariate Spatial Modelling

Noel Cressie
and
Andrew Zammit-Mangion

National Institute for Applied Statistics Research Australia
University of Wollongong

ncressie@uow.edu.au

- Univariate spatial model
  - "Marginal" behaviour of a single spatial variable
  - Optimally predict at all spatial locations: Kriging
- Multivariate spatial model
  - Two or more interacting spatial variables
  - Optimally predict one of the variables by using the observations on all variables: Cokriging
  - Determine which variable caused the observed phenomenon: Source separation (not considered in this talk)

Gneiting, Kleiber, and Schlather (2010), Pacific Northwest of North America (Left panel: First variable is forecast temperature errors. Right panel: Second variable is forecast pressure errors)

Other Examples

- Royle and Berliner (1999), Midwestern USA, centred on Illinois (First variable: Maximum temperature in degC. Second variable: Tropospheric ozone concentrations in ppb.)
- Jin, Carlin, and Banerjee (2005), Minnesota, USA (First variable: lung cancer death rates. Second variable: esophagus cancer death rates.)
- Genton and Kleiber (2015), Colorado, USA (First variable: Minimum temperature residuals in degC. Second variable: Maximum temperature residuals in degC.)

NIASRA
NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

- Statistical Modelling: Given a bivariate process $(Y_1(\cdot), Y_2(\cdot))$, we say that the *cross-covariance function matrix* (CCFM),

$$\begin{pmatrix} C_{11}(\cdot,\cdot) & C_{12}(\cdot,\cdot) \\ C_{21}(\cdot,\cdot) & C_{22}(\cdot,\cdot) \end{pmatrix},$$

  is nonnegative-definite (nnd) if any covariance matrix derived from it is nnd. In the CCFM,

$$C_{ij}(\mathbf{s},\mathbf{u}) \equiv \mathrm{cov}(Y_i(\mathbf{s}), Y_j(\mathbf{u}));\ \mathbf{s},\mathbf{u} \in \mathbb{R}^d,$$

  and note that $C_{12}(\mathbf{s},\mathbf{u}) = C_{21}(\mathbf{u},\mathbf{s})$, since $\mathrm{cov}(W, X) = \mathrm{cov}(X, W)$.
- Computational: Sometimes we have computational difficulty with kriging (univariate models) – how do such algorithms scale to multivariate modelling and multivariate spatial prediction (including cokriging)?

NIASRA
NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

- Linear model of co-regionalisation, or LMC (Journel and Huijbregts, 1978; Wackernagel, 1995): Define

$$Y_1(\cdot) \equiv a_{11}\widetilde{Y}_1(\cdot) + a_{12}\widetilde{Y}_2(\cdot),$$
$$Y_2(\cdot) \equiv a_{21}\widetilde{Y}_1(\cdot) + a_{22}\widetilde{Y}_2(\cdot),$$

where, independently,

$$\widetilde{Y}_1(\cdot) \sim \mathcal{N}(\mu_1(\cdot), C_1(\cdot,\cdot)), \quad \widetilde{Y}_2(\cdot) \sim \mathcal{N}(\mu_2(\cdot), C_2(\cdot,\cdot)).$$

The CCFM is nnd for any $\{a_{ij} : i,j = 1,2\}$, and

$$C_{ij}(\cdot,\cdot) = a_{i1}a_{j1}C_1(\cdot,\cdot) + a_{i2}a_{j2}C_2(\cdot,\cdot).$$

Hence, $C_{ij}(\mathbf{s},\mathbf{u}) = C_{ij}(\mathbf{u},\mathbf{s})$. This symmetry constraint can be inappropriate. In general, $C_{ij}(\mathbf{s},\mathbf{u}) \neq C_{ji}(\mathbf{s},\mathbf{u})$.

Multivariate Matérn models can be built from the assumption that $\{C_{ij}(\mathbf{h}) : \mathbf{h} \in \mathbb{R}^d\}$ are each proportional to a univariate Matérn correlation function; that is,

$$C_{ij}(\mathbf{h}) \propto 2^{1-\nu_{ij}}\Gamma(\nu_{ij})^{-1}(\kappa_{ij}\|\mathbf{h}\|)^{\nu_{ij}}K_{\nu_{ij}}(\kappa_{ij}\|\mathbf{h}\|),$$

where $\{\nu_{ij}\}$ are smoothness parameters, and $\{\kappa_{ij}\}$ are spatial-scale parameters. The proportionality constants are given by a covariance matrix $\{\tau_{ij}\}$. Notice that the symmetry constraint, $C_{ij}(\cdot) = C_{ji}(\cdot)$, and restrictions on parameters are needed to obtain a CCFM that is nnd.

Consider now the bivariate Matérn models:

Gneiting et al. (2010) defined bivariate Matérn models. However, as noted above, they satisfy the symmetry constraint.

- Bivariate parsimonious Matérn model: Suppose that $\nu_{ij} \equiv (\nu_i + \nu_j)/2$ and $\kappa_{ij} \equiv \kappa$; $i, j = 1, 2$. In $\mathbb{R}^2$, the CCFM is nnd iff

$$\frac{\tau_{12}^2}{\tau_{11}\tau_{22}} \leq \frac{\nu_1\nu_2}{((\nu_1 + \nu_2)/2)^2}.$$

- Bivariate full Matérn model: Here assumptions on smoothness and spatial-scale parameters are relaxed, but finding the parameters for which the CCFM is nnd is much more involved.

Symmetry is usually inappropriate

Often one process ($Y_1(\cdot)$) is potentially causative of the other ($Y_2(\cdot)$). If this is the case, we should not use models that have the symmetry constraint.

- $Y_1(\cdot)$: precipitation at present
- $Y_2(\cdot)$: precipitation in 5-minutes' time

The symmetry constraint, $C_{12}(\mathbf{s}, \mathbf{u}) = C_{21}(\mathbf{s}, \mathbf{u})$, is inappropriate here.

Bivariate spatial models with asymmetry

- An easy way to introduce asymmetry is to consider $Y_1(\cdot)$ and $Y_2(\cdot)$ modelled with the symmetry constraint, and then shift one of the processes by $\boldsymbol{\Delta}$ (e.g., fit the model $Y_1(\cdot)$ and $Y_2(\cdot - \boldsymbol{\Delta})$).
  References: Ver Hoef and Cressie (1993); Christensen and Amemiya (2001); **?**); Li and Zhang (2011)

- Another approach is to introduce latent spatial dimensions in the index space. Then asymmetry in the full-dimensional space implies asymmetry in the original space. Reference: Apanasovich and Genton (2010)

A conditional approach

This approach is valid regardless of whether $Y_1(\cdot)$ is the "baseline" process or whether $Y_2(\cdot)$ is. It is often obvious which is which (e.g., Y_1 is pollution and Y_2 is cancer-incidence rates). Here we consider $Y_1(\cdot)$ to be the baseline with covariance function $C_{11}(\cdot,\cdot)$. Write:

$$E\left(Y_2(\mathbf{s}) \mid Y_1(\cdot)\right) = \int_D b(\mathbf{s},\mathbf{v}) Y_1(\mathbf{v})\,\mathrm{d}\mathbf{v}; \quad \mathbf{s} \in D,$$

$$\mathrm{cov}\left(Y_2(\mathbf{s}), Y_2(\mathbf{u}) \mid Y_1(\cdot)\right) = C_{2|1}(\mathbf{s},\mathbf{u}); \quad \mathbf{s},\mathbf{u} \in \mathbb{R}^d.$$

Building blocks:

- $C_{11}(\cdot,\cdot)$ (univariate covariance); nnd function
- $C_{2|1}(\cdot,\cdot)$ (univariate covariance); nnd function
- $b(\cdot,\cdot)$ (interaction function); any integrable function

- The CCFM is easy to find:

$$\begin{bmatrix} C_{11}(\mathbf{s},\mathbf{u}) & \int_D C_{11}(\mathbf{s},\mathbf{v})b(\mathbf{u},\mathbf{v})\mathrm{d}\mathbf{v} \\ \int_D b(\mathbf{s},\mathbf{v})C_{11}(\mathbf{v},\mathbf{u})\mathrm{d}\mathbf{v} & C_{22}(\mathbf{s},\mathbf{u}) \end{bmatrix},$$

where

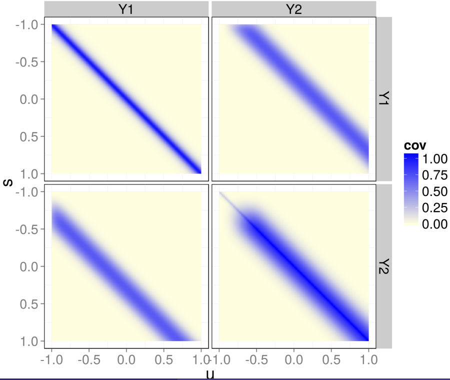
$$C_{22}(\mathbf{s},\mathbf{u}) = C_{2|1}(\mathbf{s},\mathbf{u}) + \int_D \int_D b(\mathbf{s},\mathbf{v})C_{11}(\mathbf{v},\mathbf{w})b(\mathbf{w},\mathbf{u})\mathrm{d}\mathbf{v}\mathrm{d}\mathbf{w},$$

and it is always nnd (the proof is given later).

- Asymmetry (i.e., $C_{12}(\mathbf{s},\mathbf{u}) \neq C_{21}(\mathbf{s},\mathbf{u})$) is guaranteed if $b(\cdot,\cdot)$ is not symmetric (i.e., $b(\mathbf{s},\mathbf{u}) \neq b(\mathbf{u},\mathbf{s})$).

A simple example of asymmetry in $\mathbb{R}^1$:

- $s, u \in D \equiv \{-1, -0.9, \ldots, 0.9, 1\}$.
- Define $b_o(s - u) \equiv b(s, u)$ that is "off-centre" (i.e., not symmetric about 0).
- From the figure below, $C_{22}(\cdot, \cdot)$ has edge effects and $C_{12}(s, u) \neq C_{21}(s, u)$.



NIASRA
NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

The conditional approach to multivariate spatial modelling:

- We can have a very heterogeneous CCFM, since $C_{11}(\cdot,\cdot)$, $C_{2|1}(\cdot,\cdot)$ need not be stationary, and $b(\mathbf{s},\mathbf{u})$ need not be symmetric in $\mathbf{s}$ and $\mathbf{u}$.
- We can have stationarity if we want.
- We are not restricted to Matérn-type covariance functions. The bivariate parsimonious Matérn model is a special case of our conditional approach.
- $Y_2(\cdot)$ can be smoother than $Y_1(\cdot)$, and it can have a different spatial scale, depending on $C_{2|1}$.

Example in $\mathbb{R}^1$ (see earlier slide)

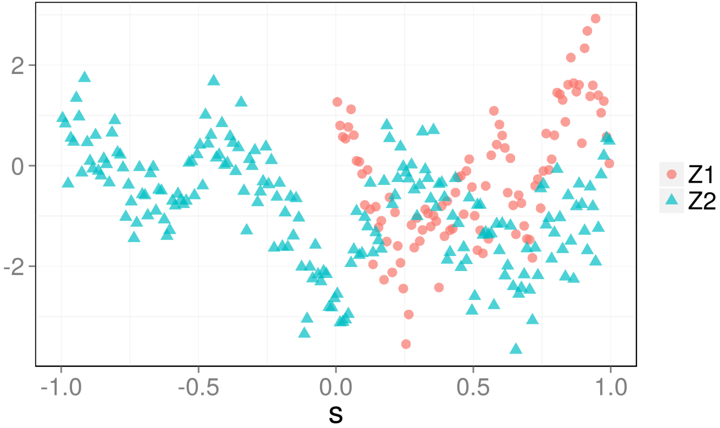
In this simple example, recall that $d = 1$ (i.e., $\mathbb{R}^1$), $D = \{-1, -0.9, \ldots, 0.9, 1\}$, and the interaction function b_o is not symmetric about 0.

- For simplicity, assume all parameters are known. Assume $Y_1(\cdot)$ is only partially observed and with measurement error.
- $Z_j(s) = Y_j(s) + \varepsilon_j(s)$, for all observation locations s, where $\{\varepsilon_j(\cdot)\}$ are independent white-noise components.
- Use both simple cokriging and simple kriging to estimate $Y_1(\cdot)$:

$$\hat{Y}_1(s_0) \equiv E(Y_1(s_0) \mid \mathbf{Z}_1, \mathbf{Z}_2) \qquad \text{simple cokriging predictor,}$$

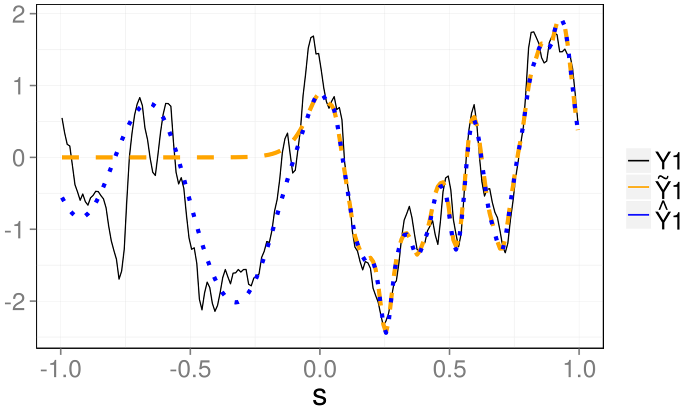
$$\widetilde{Y}_1(s_0) \equiv E(Y_1(s_0) \mid \mathbf{Z}_1) \qquad \text{simple kriging predictor.}$$

There are no observations on the first variable in the left-hand half of D.



NIASRA
NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

There are no observations on the first variable in the left-hand half of D, but cokriging of Y_1 based on all observations (blue dotted line) captures the spatial variability over all of D (true process is the **black line**).



Is the bivariate model always valid?

- $C_{11}(\cdot,\cdot)$ and $C_{2|1}(\cdot,\cdot)$ are nnd. Then $C_{22}(\cdot,\cdot)$ is nnd (recall its expression as a quadratic form).
- $C_{12}(\mathbf{s},\mathbf{u}) = C_{21}(\mathbf{u},\mathbf{s})$; this is not symmetry, and it trivially holds for all $\mathbf{s},\mathbf{u}$ (recall $\mathrm{cov}(W,X) = \mathrm{cov}(X,W)$).
- On the next slide, we show that the CCFM is nnd, and hence the model is always valid. That is, for any n_1, n_2 such that $n_1 + n_2 > 0$, any locations $\{\mathbf{s}_{1k}\}, \{\mathbf{s}_{2l}\}$, and any real numbers $\{a_{1k}\}, \{a_{2l}\}$,

$$
\begin{aligned}
&\mathrm{var}\left(\sum_{k=1}^{n_1} a_{1k} Y_1(\mathbf{s}_{1k}) + \sum_{l=1}^{n_2} a_{2l} Y_2(\mathbf{s}_{2l})\right)\\
&= \sum_{k=1}^{n_1}\sum_{k'=1}^{n_1} a_{1k}a_{1k'} C_{11}(\mathbf{s}_{1k},\mathbf{s}_{1k'}) + \sum_{l=1}^{n_2}\sum_{l'=1}^{n_2} a_{2l}a_{2l'} C_{22}(\mathbf{s}_{2l},\mathbf{s}_{2l'})\\
&+ \sum_{k=1}^{n_1}\sum_{l'=1}^{n_2} a_{1k}a_{2l'} C_{12}(\mathbf{s}_{1k},\mathbf{s}_{2l'}) + \sum_{l=1}^{n_2}\sum_{k'=1}^{n_1} a_{2l}a_{1k'} C_{21}(\mathbf{s}_{2l},\mathbf{s}_{1k'}) \geq 0.
\end{aligned}
$$

CCFM is nonnegative-definite: Proof

- It is straightforward to show that

$$\operatorname{var}\left(\sum_{k=1}^{n_1} a_{1k} Y_1(\mathbf{s}_{1k}) + \sum_{l=1}^{n_2} a_{2l} Y_2(\mathbf{s}_{2l})\right)$$
$$= \sum_{l=1}^{n_2}\sum_{l'=1}^{n_2} a_{2l}a_{2l'} C_{2|1}(\mathbf{s}_{2l}, \mathbf{s}_{2l'}) + \int_D \int_D a(\mathbf{s})a(\mathbf{u}) C_{11}(\mathbf{s}, \mathbf{u})\, \mathrm{d}\mathbf{s}\mathrm{d}\mathbf{u},$$

where for $\delta(\cdot)$ the Dirac delta function,

$$a(\mathbf{s}) \equiv \sum_{k=1}^{n_1} a_{1k}\delta(\mathbf{s} - \mathbf{s}_{1k}) + \sum_{l=1}^{n_2} a_{2l} b(\mathbf{s}_{2l}, \mathbf{s}); \quad \mathbf{s} \in \mathbb{R}^d.$$

Since $C_{11}(\cdot,\cdot)$ and $C_{2|1}(\cdot,\cdot)$ are nnd by assumption, the right-hand side is ≥ 0.

NIASRA
NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

- For $p \geq 2$, $[Y_1(\cdot), \ldots, Y_p(\cdot)]$ can be decomposed as,

$$[Y_1(\cdot)] \times [Y_2(\cdot)|Y_1(\cdot)] \times \ldots \times [Y_p(\cdot) \mid Y_{p-1}(\cdot), Y_{p-2}(\cdot), \ldots, Y_1(\cdot)].$$

- For $p \geq 2$, $[Y_1(\cdot), \ldots, Y_p(\cdot)]$ can be decomposed as,

$$[Y_1(\cdot)] \times [Y_2(\cdot)|Y_1(\cdot)] \times \ldots \times [Y_p(\cdot) \mid Y_{p-1}(\cdot), Y_{p-2}(\cdot), \ldots, Y_1(\cdot)].$$

- Assume the conditional expectation for the q-th term is, for $\mathbf{s} \in D$,

$$E(Y_q(\mathbf{s}) \mid \{Y_r(\cdot) : r = 1, \ldots, q-1\}) \equiv \sum_{r=1}^{q-1} \int_D b_{qr}(\mathbf{s}, \mathbf{v}) Y_r(\mathbf{v}) \mathrm{d}\mathbf{v},$$

where $\{b_{qr}(\cdot, \cdot) : r = 1, \ldots, q-1; q = 2, \ldots, p\}$ are integrable.

- For $p \geq 2$, $[Y_1(\cdot), \ldots, Y_p(\cdot)]$ can be decomposed as,

$$[Y_1(\cdot)] \times [Y_2(\cdot)|Y_1(\cdot)] \times \ldots \times [Y_p(\cdot) \mid Y_{p-1}(\cdot), Y_{p-2}(\cdot), \ldots, Y_1(\cdot)].$$

- Assume the conditional expectation for the q-th term is, for $\mathbf{s} \in D$,

$$E(Y_q(\mathbf{s}) \mid \{Y_r(\cdot) : r = 1, \ldots, q-1\}) \equiv \sum_{r=1}^{q-1} \int_D b_{qr}(\mathbf{s}, \mathbf{v}) Y_r(\mathbf{v}) \mathrm{d}\mathbf{v},$$

where $\{b_{qr}(\cdot, \cdot) : r = 1, \ldots, q-1; q = 2, \ldots, p\}$ are integrable.

- Assume the conditional covariance for the q-th term is, for $\mathbf{s}, \mathbf{u} \in \mathbb{R}^d$,

$$\mathrm{cov}(Y_q(\mathbf{s}), Y_q(\mathbf{u}) \mid \{Y_r(\cdot) : r = 1, \ldots, (q-1)\}) \equiv C_{q|(r<q)}(\mathbf{s}, \mathbf{u}),$$

which is a univariate nnd function.

We show that the p-variate process is valid, by induction:

- For nnd C_{11} and $C_{2|1}$, the bivariate process is valid.
- Assume that the $(p-1)$-variate process is valid.
- Show that the p-variate process is valid: For any n's, $\mathbf{s}$'s, and a's,

$$\mathrm{var}\left(\sum_{q=1}^{p}\sum_{m=1}^{n_q} a_{qm} Y_q(\mathbf{s}_{qm})\right) = \sum_{m=1}^{n_p}\sum_{m'=1}^{n_p} a_{pm}a_{pm'} C_{p|(q<p)}(\mathbf{s}_{pm},\mathbf{s}_{pm'}) + \sum_{q=1}^{p-1}\sum_{r=1}^{p-1}\int_D\int_D a_q(\mathbf{s})a_r(\mathbf{u})C_{qr}(\mathbf{s},\mathbf{u})\mathrm{d}\mathbf{s}\mathrm{d}\mathbf{u},$$

where for $\delta(\cdot)$ the Dirac delta function and for $\mathbf{s}\in\mathbb{R}^d$,

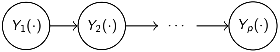
$$a_q(\mathbf{s}) \equiv \left(\sum_{k=1}^{n_q} a_{qk}\delta(\mathbf{s}-\mathbf{s}_{qk}) + \sum_{m=1}^{n_p} a_{pm} b_{pq}(\mathbf{s}_{pm},\mathbf{s})\right).$$

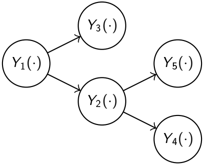
The following families of multivariate spatial processes contain classes that are special cases of models defined by the conditional approach:

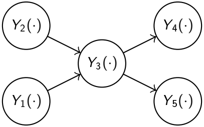
- Royle and Berliner (1999) defined a conditional approach on random vectors of data and predictor rather than on random processes.
- Cressie and Wikle (2011) discretised D and defined a conditional approach on the resulting vectors of the random processes.
- The parsimonious Matérn model of Gneiting et al. (2010).
- The linear model of coregionalisation, used for example by Wackernagel (1995).
- The moving-average model of Ver Hoef and Barry (1998).
- The shifted models of Ver Hoef and Cressie, Christensen and Amemiya, and Li and Zhang (see earlier slide: "Bivariate spatial models with asymmetry").

- Directed acyclic graphs (DAGs) on the variables define a partial order that allows a more parsimonious p-variate model. The conditional covariances, $\{C_{q|(r<q)}\}$, are replaced by the parsimonious set $\{C_{q|pa(q)}\}$, where $pa(q)$ denotes the "parents" of the q-th variable.
- Computationally efficient algorithms are available for DAGs.

Examples of DAGs include:







Model flexibility

$C_{11}(\cdot,\cdot)$	
$C_{21}(\cdot,\cdot)$	$C_{22}(\cdot,\cdot)$

Bivariate system: We need to specify three marginal/cross-covariance functions that result in a nnd CCFM. Recall that $C_{12}(\mathbf{s},\mathbf{u}) = C_{21}(\mathbf{u},\mathbf{s})$, which is not the symmetry constraint.

Building blocks for the conditional approach: Three functions, nnd $C_{11}(\cdot,\cdot)$, nnd $C_{2|1}(\cdot,\cdot)$, integrable $b(\cdot,\cdot)$, specified independently.

Model flexibility, ctd

NIASRA
NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

$C_{11}(\cdot,\cdot)$		
$C_{21}(\cdot,\cdot)$	$C_{22}(\cdot,\cdot)$	
$C_{31}(\cdot,\cdot)$	$C_{32}(\cdot,\cdot)$	$C_{33}(\cdot,\cdot)$

Trivariate system: Need to specify six marginal/cross-covariance functions that result in a nnd CCFM. Recall that $C_{ij}(\mathbf{s},\mathbf{u}) = C_{ji}(\mathbf{u},\mathbf{s})$.

Building blocks for the conditional approach: Six functions, $C_{11}(\cdot,\cdot)$, $C_{2|1}(\cdot,\cdot)$, $C_{3|1,2}(\cdot,\cdot)$, $b_{21}(\cdot,\cdot)$, $b_{31}(\cdot,\cdot)$, $b_{32}(\cdot,\cdot)$, specified independently.

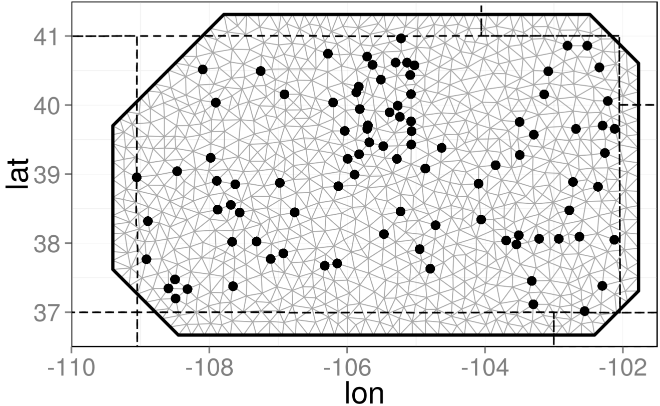
- Datatset: Minimum and maximum temperatures taken on September 19, 2004 in the state of Colorado, USA (Genton and Kleiber, 2015).
- Data come from 94 measurement stations (collocated measurements); our data are residuals $\mathbf{Z}_1$ (min. temp.) and $\mathbf{Z}_2$ (max. temp.) obtained by subtraction of the respective statewide means.
- The maximum-temperature residual process, occurring later in the afternoon ($Y_2(\cdot)$), is highly dependent on the minimum-temperature residual process, occurring in the early-morning hours ($Y_1(\cdot)$).
- We fit three models and compare them using DIC:

Model 1: $b_o(\mathbf{h}) \equiv 0$ (i.e., independence)

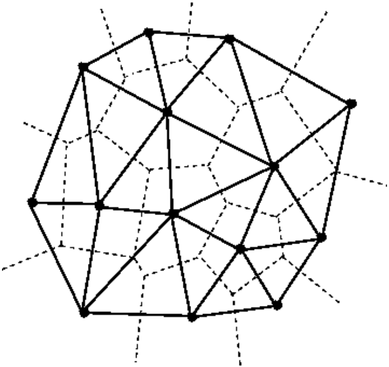
Model 2: $b_o(\mathbf{h}) \equiv A\delta(\mathbf{h})$ (i.e., pointwise dependence)

Model 3: $$b_o(\mathbf{h}) \equiv \begin{cases} A\{1 - (\|\mathbf{h} - \boldsymbol{\Delta}\|/r)^2\}^2, & \|\mathbf{h} - \boldsymbol{\Delta}\| \leq r \\ 0, & \text{otherwise.} \end{cases}$$

- Consider a discretisation of $Y_1(\cdot)$ and $Y_2(\cdot)$; call the resulting n-dimensional ($n = 968$) vectors, $\mathbf{Y}_1$ and $\mathbf{Y}_2$, respectively, and define $\mathbf{Y} \equiv (\mathbf{Y}_1', \mathbf{Y}_2')'$. The 188-dimensional ($m = m_1 + m_2 = 94 + 94 = 188$) data vector is $\mathbf{Z} \equiv (\mathbf{Z}_1', \mathbf{Z}_2')'$ at 94 locations:



Numerical integrations



Approximate $E\left(Y_2(\mathbf{s}) \mid Y_1(\cdot)\right) = \int_D b(\mathbf{s}, \mathbf{v}) Y_1(\mathbf{v})\, \mathrm{d}\mathbf{v};\ \mathbf{s} \in D,$ by

$$E(Y_2(\mathbf{s}_l) \mid Y_1(\cdot)) \simeq \sum_{k=1}^{n} A_k b(\mathbf{s}_l, \mathbf{v}_k) Y_1(\mathbf{v}_k),$$

where $\{A_k : k = 1, \ldots, 968\}$ are the polygonal-tessellation areas.

- **Data model**:

$$\begin{pmatrix}\mathbf{Z}_1\\ \mathbf{Z}_2\end{pmatrix}\Bigg|\begin{pmatrix}\mathbf{Y}_1\\ \mathbf{Y}_2\end{pmatrix},\boldsymbol{\theta}\sim\mathcal{N}\left(\begin{pmatrix}\mathbf{D}\mathbf{Y}_1\\ \mathbf{D}\mathbf{Y}_2\end{pmatrix},\sigma_\varepsilon^2\begin{pmatrix}\mathbf{I} & \rho_\varepsilon\mathbf{I}\\ \rho_\varepsilon\mathbf{I} & \mathbf{I}\end{pmatrix}\right),$$

where $\mathbf{D}$ is a 94×968 incidence matrix and $\boldsymbol{\theta}$ includes σ_ε^2 and ρ_ε.

- **Process model**:

$$\begin{pmatrix}\mathbf{Y}_1\\ \mathbf{Y}_2\end{pmatrix}\Bigg|\boldsymbol{\theta}\sim\mathcal{N}\left(\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix},\begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{11}\mathbf{B}'\\ \mathbf{B}\boldsymbol{\Sigma}_{11} & \mathbf{B}\boldsymbol{\Sigma}_{11}\mathbf{B}'+\boldsymbol{\Sigma}_{2|1}\end{pmatrix}\right),$$

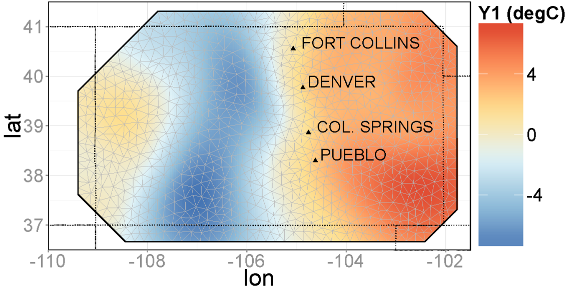
where $\mathbf{B}$ (interaction matrix), $\boldsymbol{\Sigma}_{11}$ (marginal covariance matrix), and $\boldsymbol{\Sigma}_{2|1}$ (conditional covariance matrix) are 968×968 matrices that depend on parameters included in $\boldsymbol{\theta}$.

- **Parameter model**: see `https://github.com/andrewzm/bicon`

Assume $C_{11}(\cdot)$ and $C_{2|1}(\cdot)$ are equally smooth Matérn covariance functions with parameters $(\nu_{11} = 1.5, \kappa_{11}, \sigma^2_{11})$ and $(\nu_{2|1} = 1.5, \kappa_{2|1}, \sigma^2_{2|1})$, respectively. Here we show our (Bayesian) inference on the scale parameters only; notice the change in $[\kappa_{2|1} | \mathbf{Z}_1, \mathbf{Z}_2]$ for Model 3.

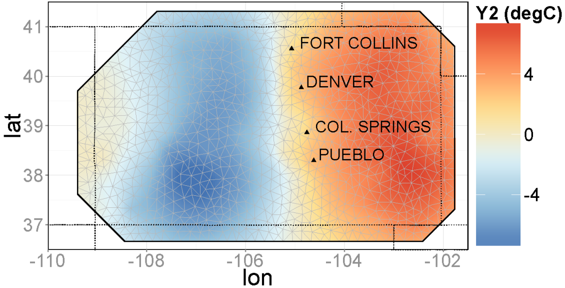
Parameter	Model 1	Model 2	Model 3
σ^2_ε	x	x	x
ρ_ε	x	x	x
σ^2_{11}	x	x	x
$\sigma^2_{2\|1}$	x	x	x
κ_{11}	0.98 (0.76, 1.22)	1 (0.8, 1.26)	1.03 (0.83, 1.25)
$\kappa_{2\|1}$	0.76 (0.56, 1)	0.62 (0.46, 0.81)	3.65 (1.16, 6.72)
A		x	x
r			x
Δ_1			x
Δ_2			x
DIC	992.45	985.17	982.45

Posterior mean field: Minimum temperature

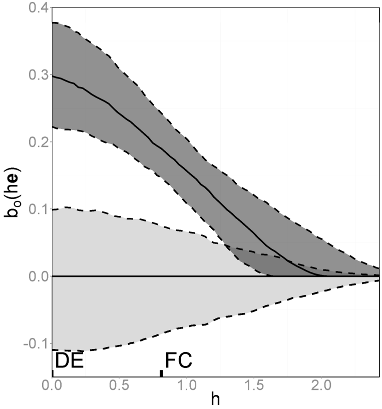


Optimal (cokriging) map of predicted (residual) minimum temperature, $E(\mathbf{Y}_1 \mid \mathbf{Z}_1, \mathbf{Z}_2)$, in degrees Celsius (degC).

Posterior mean field: Maximum temperature



Optimal (cokriging) map of predicted (residual) maximum temperature, $E(\mathbf{Y}_2 \mid \mathbf{Z}_1, \mathbf{Z}_2)$, in degrees Celsius (degC).



Prior (light grey) and posterior (dark grey) median (solid line) and interquartile ranges (enclosed by dashed lines) of $b_o(\cdot)$ from Model 3, along a unit vector **e** originating at Denver (DE) in the direction of Fort Collins (FC).

Conclusions

NIASRA NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

- Bivariate and multivariate spatial models often appear in environmental studies. For convenience, one or more of these variables are often "explained away" prior to commencing a univariate spatial analysis. We wish to avoid this by providing a methodology for building flexible (e.g., no symmetry constraint; easy-to-verify nnd conditions) multivariate spatial models.
- The conditional approach allows for a (very) flexible model class through the specification of integrable interaction functions that can be arbitrarily complex.
- One way to handle non-Gaussian multivariate data is as follows: A generalised linear model for the data model; a transformed multivariate Gaussian process within the process model; and the conditional approach applied to the Gaussian process.
- Slides and reproducible code available at `https://github.com/andrewzm/bicon`.

Apanasovich, T. V. and Genton, M. G. (2010). "Cross-covariance functions for multivariate random fields based on latent dimensions." *Biometrika*, 97(1): 15–30.

Christensen, W. F. and Amemiya, Y. (2001). "Generalized shifted-factor analysis method for multivariate geo-referenced data." *Mathematical Geology*, 33(7): 801–824.

Cressie, N. and Wikle, C. K. (2011). *Statistics for Spatio-Temporal Data*. Hoboken, NJ: John Wiley and Sons.

Genton, M. G. and Kleiber, W. (2015). "Cross-covariance functions for multivariate geostatistics (with discussion)." *Statistical Science*, 30: 147–163.

Gneiting, T., Kleiber, W., and Schlather, M. (2010). "Matérn cross-covariance functions for multivariate random fields." *Journal of the American Statistical Association*, 105: 1167–1177.

Jin, X., Carlin, B. P., and Banerjee, S. (2005). "Generalized hierarchical multivariate CAR models for areal data." *Biometrics*, 61: 950–961.

Journel, A. G. and Huijbregts, C. J. (1978). *Mining Geostatistics*. London: Academic Press.

Li, B. and Zhang, H. (2011). "An approach to modeling asymmetric multivariate spatial covariance structures." *Journal of Multivariate Analysis*, 102: 1445–1453.

Royle, J. A. and Berliner, L. M. (1999). "A hierarchical approach to multivariate spatial modeling and prediction." *Journal of Agricultural, Biological, and Environmental Statistics*, 4: 29–56.

Ver Hoef, J. M. and Barry, R. P. (1998). "Constructing and fitting models for cokriging and multivariable spatial prediction." *Journal of Statistical Planning and Inference*, 69: 275–294.

Ver Hoef, J. M. and Cressie, N. (1993). "Multivariable spatial prediction." *Mathematical Geology*, 25: 219–240. Errata: 1994, Vol. **26**, pp. 273–275.

Wackernagel, H. (1995). *Multivariate Geostatistics*. Berlin, DE: Springer.