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

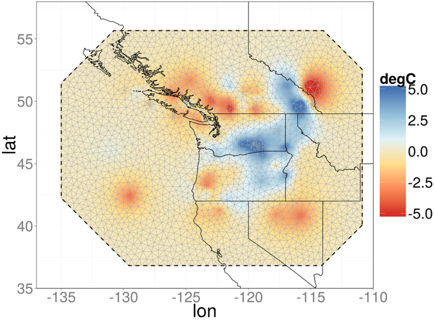
Introduction

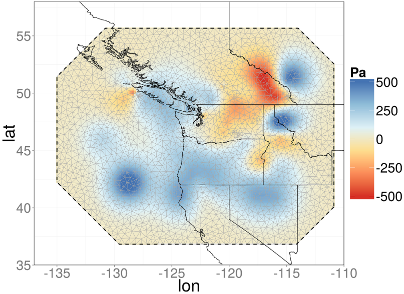
NIASRA NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

- 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)

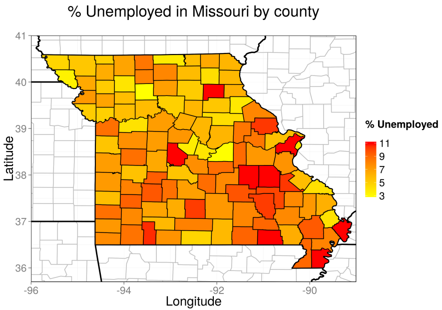
Example: Temperature and Pressure

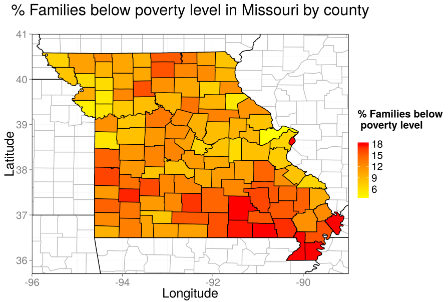
Gneiting, Kleiber, and Schlather (2010), Pacific Northwest of North America (Left panel: First variable is forecast temperature errors. Right panel: Second variable is forecast pressure errors)





Porter, Wikle, and Holan (2015), Counties of Missouri, USA (Left panel: First variable is percentage of unemployed individuals 16 years or older. Right panel: Second variable is percentage of families below the poverty level)





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)?

- 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

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 $\mathbf{\Delta}$ (e.g., fit the model $Y_1(\cdot)$ and $Y_2(\cdot - \mathbf{\Delta})$). References: Ver Hoef and Cressie (1993); Christensen and Amemiya (2001, 2002); 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)

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:

$$\begin{aligned} E\left(Y_2(\mathbf{s}) \mid Y_1(\cdot)\right) &= \int_D b(\mathbf{s},\mathbf{v}) Y_1(\mathbf{v})\,\mathrm{d}\mathbf{v}; \quad \mathbf{s} \in D, \\ \operatorname{cov}\left(Y_2(\mathbf{s}), Y_2(\mathbf{u}) \mid Y_1(\cdot)\right) &= C_{2|1}(\mathbf{s},\mathbf{u}); \quad \mathbf{s},\mathbf{u} \in \mathbb{R}^d. \end{aligned}$$

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 model generalises to $Y_1(\cdot)$ on D_1 and $Y_2(\cdot)$ on D_2.

- 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

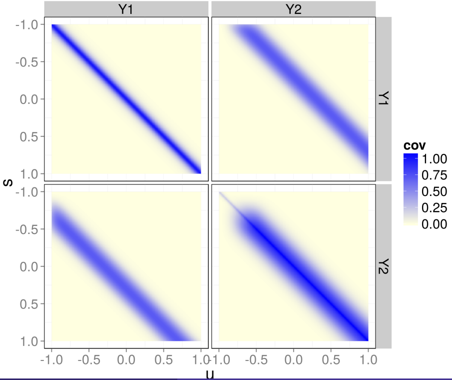
$$C_{22}(\mathbf{s},\mathbf{u}) = C_{2|1}(\mathbf{s},\mathbf{u}) + \int_D \int_D b(\mathbf{s},\mathbf{v}) C_{11}(\mathbf{v},\mathbf{w}) b(\mathbf{w},\mathbf{u}) \mathrm{d}\mathbf{v}\mathrm{d}\mathbf{w},$$

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 $C_{12}(s,u) \neq C_{21}(s,u)$.



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 $Y_1(\cdot)$:

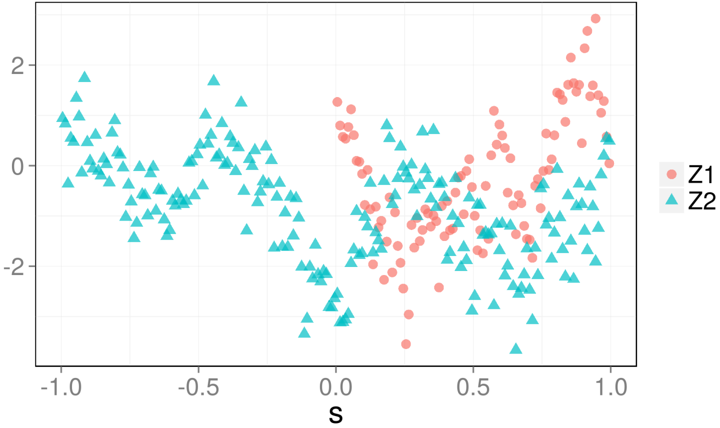
$$\hat{Y}_1(s_0) \equiv E(Y_1(s_0) \mid \mathbf{Z}_1, \mathbf{Z}_2) \qquad \text{simple cokriging predictor,}$$

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

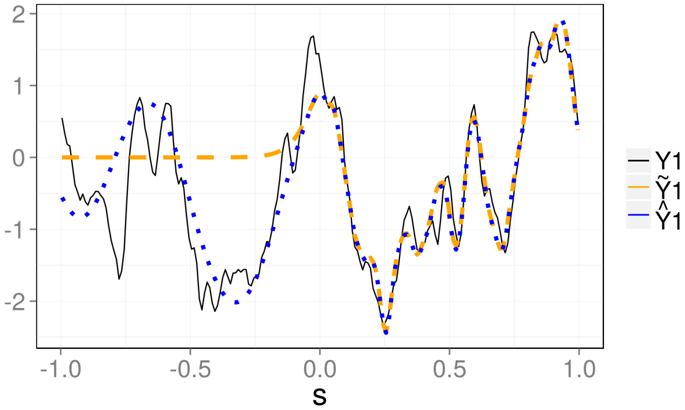
Example in $\mathbb{R}^1$, ctd

NIASRA NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

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



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**).



NIASRA NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

- $C_{11}(\cdot,\cdot)$ 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 $\{a_{1k}\}, \{a_{2l}\}$,

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

- It is straightforward to show that

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

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.

Beyond bivariate

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

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

NIASRA
NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

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

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

- 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 $\mathbf{s} \in \mathbb{R}^d$,

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

Relationship to other multivariate models

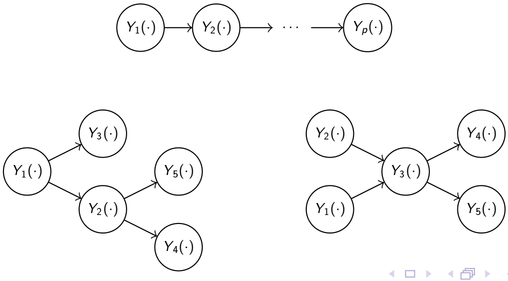
The following families of multivariate spatial processes contain classes that are special cases of models defined by the conditional approach:

- 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").

Graph structure

- 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.

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

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

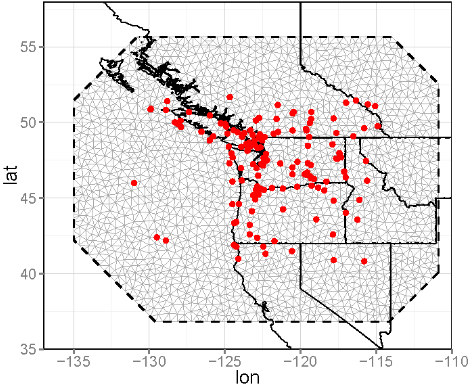
Building blocks for the conditional approach: Six functions, $C_{11}(\cdot,\cdot)$, $C_{2|1}(\cdot,\cdot)$, $C_{3|1,2}(\cdot,\cdot)$, $b_{21}(\cdot,\cdot)$, $b_{31}(\cdot,\cdot)$, $b_{32}(\cdot,\cdot)$, specified independently.

Temperature- and pressure-forecast errors

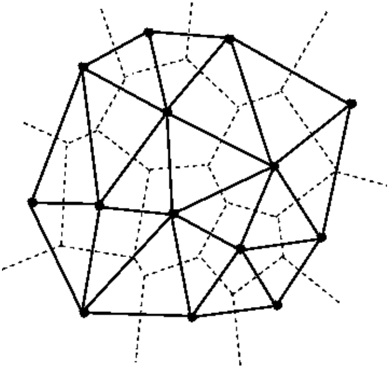
- Datatset: Temperature- and pressure-forecast errors at station locations in the Pacific Northwest of North America on December 18, 2003 at 4p.m. (Gneiting et al., 2010).
- Data come from 157 measurement stations (co-located measurements).
- The pressure-forecast error process, $(Y_2(\cdot))$, is highly dependent on the temperature-forecast error process $(Y_1(\cdot))$.
- We fit four models and compare them using cross-validation and AIC:

$$\begin{aligned}
&\text{Model 1:} \quad b_o(\mathbf{h}) \equiv 0 \text{ (i.e., independence)}\\
&\text{Model 2:} \quad b_o(\mathbf{h}) \equiv A\delta(\mathbf{h}) \text{ (i.e., pointwise dependence)}\\
&\text{Model 3:} \quad b_o(\mathbf{h}) \equiv \begin{cases} A\{1-(\|\mathbf{h}\|/r)^2\}^2, & \|\mathbf{h}\| \leq r\\ 0, & \text{otherwise,}\end{cases}\\
&\text{Model 4:} \quad b_o(\mathbf{h}) \equiv \begin{cases} A\{1-(\|\mathbf{h}-\boldsymbol{\Delta}\|/r)^2\}^2, & \|\mathbf{h}-\boldsymbol{\Delta}\| \leq r\\ 0, & \text{otherwise.}\end{cases}
\end{aligned}$$

- Consider a discretisation of $Y_1(\cdot)$ and $Y_2(\cdot)$; call the resulting n-dimensional ($n = 2063$) vectors, $\mathbf{Y}_1$ and $\mathbf{Y}_2$, respectively, and define $\mathbf{Y} \equiv (\mathbf{Y}_1', \mathbf{Y}_2')'$. The 314-dimensional ($m = m_1 + m_2 = 314$) data vector is $\mathbf{Z} \equiv (\mathbf{Z}_1', \mathbf{Z}_2')'$ at 157 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 $\{A_k : k = 1, \ldots, 2063\}$ are the polygonal-tessellation areas.

NIASRA
NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

We have observations, $O_q(s_i)$, and forecasts, $F_q(s_i)$, $q = 1, 2$. The data are

$$Z_q(s_i) = Y_q(s_i) = F_q(s_i) - O_q(s_i) \ (i = 1, \ldots, 157; q = 1, 2).$$

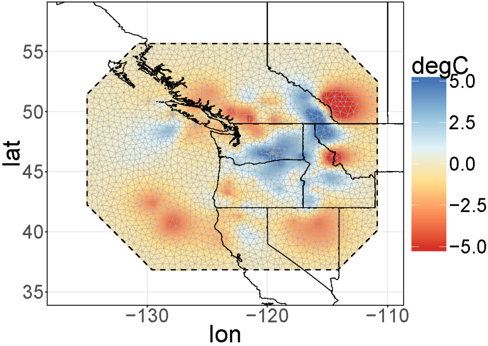
The data Z_q and the process Y_q at their respective locations are the same.

- **Data model**:
$$\mathbf{Z}_1 = \mathbf{D}\mathbf{Y}_1, \qquad \mathbf{Z}_2 = \mathbf{D}\mathbf{Y}_2,$$
where $\mathbf{D}$ is a 157×2063 incidence matrix.
- **Process model**:
$$\begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix} \Bigg| \boldsymbol{\theta} \sim \mathcal{N}\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11} + \tau_1^2 \mathbf{I}_{m_1} & \boldsymbol{\Sigma}_{11}\mathbf{B}' \\ \mathbf{B}\boldsymbol{\Sigma}_{11} & \mathbf{B}\boldsymbol{\Sigma}_{11}\mathbf{B}' + \boldsymbol{\Sigma}_{2|1} + \tau_2^2 \mathbf{I}_{m_2} \end{pmatrix} \right),$$
where $\tau_q^2, q = 1, 2$, are fine-scale variance parameters, and $\mathbf{B}$ (interaction matrix), $\boldsymbol{\Sigma}_{11}$ (marginal covariance matrix), and $\boldsymbol{\Sigma}_{2|1}$ (conditional covariance matrix) are 2063×2063 matrices that depend on the parameter vector $\boldsymbol{\theta}$.

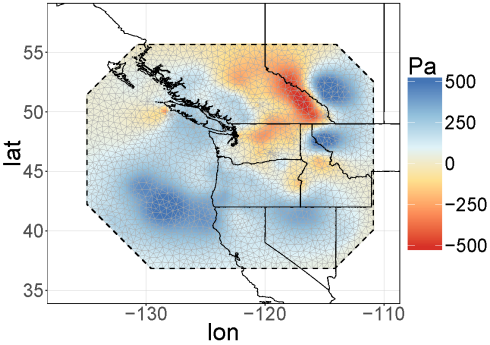
Assume $C_{11}(\cdot)$ and $C_{2|1}(\cdot)$ are equally smooth Matérn covariance functions with parameters $(\nu_{11}, \kappa_{11}, \sigma_{11}^2)$ and $(\nu_{2|1}, \kappa_{2|1}, \sigma_{2|1}^2)$, respectively. Notice how the log likelihood and the estimate for A increase, while the AIC and the estimate for $\sigma_{2|1}$ decrease overall, with the model number.

	τ_1	τ_2	σ_{11}	$\sigma_{2\|1}$	κ_{11}	$\kappa_{2\|1}$	ν_{11}	$\nu_{2\|1}$
Model 1	0.00	68.47	2.60	275.34	0.011	0.010	0.60	1.56
Model 2	0.00	67.78	2.60	242.04	0.011	0.011	0.60	1.58
Model 3	0.00	70.16	2.68	243.77	0.011	0.010	0.61	1.84
Model 4	0.01	69.79	3.02	199.86	0.007	0.004	0.56	1.24

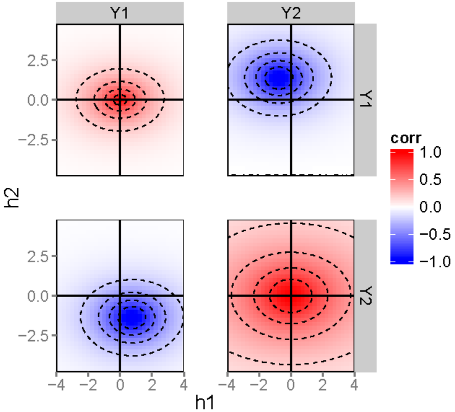
	A	r	Δ_1	Δ_2	Log-lik.	AIC
Model 1					-1276.77	2569.54
Model 2	-14.30				-1269.92	2557.84
Model 3	-40.83	1.46			-1264.90	2549.80
Model 4	-65.58	1.18	0.76	-1.42	-1258.21	2540.43



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



Optimal (cokriging) map of predicted pressure-forecast error, $E(\mathbf{Y}_2 \mid \mathbf{Z}_1, \mathbf{Z}_2)$, in Pascal (Pa).



Correlation and cross-correlation functions estimated from Model 4.

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.
- 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." *Math. Geol.*, 33(7): 801–824.

Christensen, W. F. and Amemiya, Y. (2002). "Latent variable analysis of multivariate spatial data." *J. Amer. Statist. Assoc.*, 97(457): 302–327.

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)." *Statist. Sci.*, 30: 147–163.

Gneiting, T., Kleiber, W., and Schlather, M. (2010). "Matérn cross-covariance functions for multivariate random fields." *J. Amer. Statist. Assoc.*, 105: 1167–1177.

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

NIASRA NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

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

Li, B. and Zhang, H. (2011). "An approach to modeling asymmetric multivariate spatial covariance structures." *J. Multiv. Anal.*, 102: 1445–1453.

Porter, A. T., Wikle, C. K., and Holan, S. H. (2015). "Small area estimation via multivariate Fay-Herriot models with latent spatial dependence." *Australian & New Zealand Journal of Statistics*, 57: 15–29.

Royle, J. A. and Berliner, L. M. (1999). "A hierarchical approach to multivariate spatial modeling and prediction." *J. Agric., Biol., and Environm. Statist.*, 4: 29–56.

Ver Hoef, J. M. and Barry, R. P. (1998). "Constructing and fitting models for cokriging and multivariable spatial prediction." *J. Statist. Plann. Infer.*, 69: 275–294.

Ver Hoef, J. M. and Cressie, N. (1993). "Multivariable spatial prediction." *Math. Geol.*, 25: 219–240.

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