

Online Resource A: Derivation of marginal posteriors and posterior predictive distributions

Belongs with: Model-based geostatistics from a Bayesian perspective: Investigating area-to-point kriging with small datasets
Mathematical Geosciences Luc Steinbuch, Tom Orton, Dick Brus ✉ luc.steinbuch@wur.nl, Wageningen UR

A1 Introduction

In this additional material, we will – additional to the main manuscript – explain the mathematics of the provided marginal posteriors and posterior predictive distribution, and connect the mathematics to the implementation in the BAK (Bayesian areal kriging) R code.

We assume that the point-support variable is multivariate Gaussian with mean $\mathbf{X}\beta$ (design matrix times regression coefficients vector) and covariance matrix entries $\sigma^2 C(\mathbf{h}; \phi)$ where $C(\mathbf{h}; \phi)$ is the correlation function of separation distance $\mathbf{h}$ with the single distance parameter ϕ . Note that since the nugget provides no contribution to the likelihood function, it is not included in this work.

The full likelihood function for these parameters, with areal-support rather than point-support data, is given by:

$$f_l(\bar{\mathbf{z}}|\beta, \sigma^2, \phi) = \frac{1}{(2\pi)^{\frac{m}{2}} (\sigma^2)^{\frac{m}{2}} |\bar{\mathbf{C}}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2\sigma^2} (\bar{\mathbf{z}} - \bar{\mathbf{X}}\beta)^T \bar{\mathbf{C}}^{-1} (\bar{\mathbf{z}} - \bar{\mathbf{X}}\beta)\right\}, \quad (\text{A1})$$

with: $\bar{\mathbf{z}}$ the m areal averages, $\bar{\mathbf{X}}$ the design matrix based on area average covariates, $\bar{\mathbf{C}}$ the mean correlation matrix, with entries given by the means of the point-support correlations – or the approximations provided by Gaussian quadrature or some other method.

We assume a prior that represents a priori independence between the trend, variance and correlation range parameters, which can therefore be decomposed as

$$f_0(\beta, \sigma^2, \phi) \propto f_0(\beta) f_0(\sigma^2) f_0(\phi) \quad (\text{A2})$$

The assumed prior is unlimited uniform, and thus improper, in the regression coefficient vector β (therefore $f_0(\beta)$ can be disregarded), is a user-specified proper distribution for ϕ , and is proportional to an inverse-gamma for σ^2 with shape and scale parameters α_0 and β_0 respectively:

$$f_0(\beta, \sigma^2, \phi) \propto f_0(\phi) \frac{1}{(\sigma^2)^{\alpha_0+1}} \exp\left(\frac{-\beta_0}{\sigma^2}\right). \quad (\text{A3})$$

Note that the inverse-gamma scale parameter, β_0 here and in general β , is not to be confused with regression coefficients vector β , which is set in boldface, or with its elements β_q . For the inverse-gamma prior for the variance, we allow hyper-parameter values $\alpha_0 \geq -1, \beta_0 \geq 0$ (in contrast to the usual bounds for inverse-gamma parameters of $\alpha_0 > 0, \beta_0 > 0$), which permits improper priors for σ^2 ; in particular, with $\beta_0 = 0$ and $\alpha_0 = 0$ we obtain a prior that is proportional to the inverse of the variance, a common (improper) uninformative prior and is equivalent to a uniform prior for $\log(\sigma^2)$ – it is this prior that we adopt in the main manuscript, and these hyper-parameters therefore do not appear in the text there. Also, with $\beta_0 = 0$ and $\alpha_0 = -1$ we obtain a uniform improper prior for the variance – not such a common prior for the variance, but worth noting as a special case.

In the numerical implementation, all values for β , σ^2 and ϕ are limited and discrete.

According to Bayes' theorem, the multivariate posterior is the proportional product of prior and likelihood:

$$f_p(\beta, \sigma^2, \phi|\bar{\mathbf{z}}) \propto f_l(\bar{\mathbf{z}}|\beta, \sigma^2, \phi) f_0(\beta, \sigma^2, \phi) \quad (\text{A4})$$

However, we are mainly interested in the following marginal posteriors:

$f_p(\phi|\bar{\mathbf{z}})$, so we have to integrate out β and σ^2 ;

$f_p(\sigma^2|\bar{\mathbf{z}})$, so we have to integrate out β and ϕ ;

$f_p(\beta_q|\bar{\mathbf{z}})$ for each β_q in β , so we have to integrate out σ^2 and ϕ , and subsequently find the marginal $f_p(\beta_q|\bar{\mathbf{z}})$'s.

We are also interested in the posterior predictive distribution $f_p(\mathbf{z}^*|\bar{\mathbf{z}})$, with $\mathbf{z}^*$ the predictions on the point locations of interest, and the associated prediction variance $\mathbf{v}^*$.

In the remainder of this document, the following topics are discussed:

- In section A2, *Marginal posterior for ϕ and σ^2 (via Integral a):*

To be able to calculate $f_p(\phi|\bar{\mathbf{z}})$ and $f_p(\sigma^2|\bar{\mathbf{z}})$, first we analytically integrate out β from the full joint posterior distribution, eqn. (A4), to get the joint posterior for ϕ and σ^2 . The result involves eqn. (A17), which we name *Integral a*, and which is equivalent to the REML function.

- Section A3, *Marginal posterior for ϕ (via Integral b)*:
To calculate $f_p(\phi|\bar{z})$, we analytically integrate out σ^2 from *Integral a* multiplied by the prior for σ^2 . The final result is eqn. (A29) (*Integral b*). Numerically, we create a one-dimensional, not necessarily regular, grid for ϕ in parameter space, calculate the marginal posterior for each ϕ , and normalise it to a distribution that integrates to one.
- Section A4, *Marginal posterior for σ^2* :
To calculate $f_p(\sigma^2|\bar{z})$, the marginal posterior distribution for σ^2 , we use *Integral a*. We express it as an inverse-gamma distribution conditional on ϕ , see eqn. (A35). Numerically, we extend the vector with ϕ to a 2D grid in the parameterspace of ϕ and σ^2 and calculate the marginal posterior for σ^2 for every gridpoint; then we calculate a weighted sum over ϕ to get the marginal distribution for σ^2 .
- Section A5, *Marginal posteriors for β_q (via Integral c)*:
To calculate $f_p(\beta_q|\bar{z})$'s, the marginal posterior distribution for the β_q , we analytically integrate σ^2 out of the posterior (eqn. (A4)) to arrive at eqn. (A64): *Integral c*. Numerically, we extend the vector with ϕ to a grid in parameterspace of ϕ and β and calculate the marginal posterior for β for every gridpoint; then we calculate a weighted sum over ϕ to get the marginal distribution for the β_q 's.
- Section A6, *Posterior predictive distribution $f_p(z^*|\bar{z})$* :
We use an analytically derived multivariate t-distribution to provide a prediction, eqn. (A82), and the associated uncertainty, eqn. (A86), for every spatial prediction point, for every ϕ in the grid vector. We use this information to numerically calculate the final prediction and uncertainty, weighted by $f_p(\phi|\bar{z})$.
- Finally, section A7, *Connect to R code* contains a pseudo-code overview to explain the functionality of the main function `postc1abetaz0`, which is closely related to, but – especially in sequence – not the same as the above provided overview.

A2 Marginal posterior for ϕ and σ^2 (via *Integral a*)

The joint posterior for ϕ and σ^2 is given by integrating out β from the full joint posterior

$$\begin{aligned} f_p(\sigma^2, \phi|\bar{z}) &\propto \int f_l(\bar{z}|\beta, \sigma^2, \phi) f_0(\beta, \sigma^2, \phi) d\beta \\ &= f_{ILa}(\bar{z}|\phi, \sigma^2) f_0(\sigma^2) f_0(\phi) \end{aligned} \quad (A5)$$

where we define *Integrated Likelihood a*, or shortly *Integral a*, as:

$$f_{ILa}(\bar{z}|\phi, \sigma^2) = \int_{\beta} f_l(\bar{z}|\beta, \phi, \sigma^2) d\beta. \quad (A6)$$

Eqn. (A6) applied on eqn. (A1) gives

$$f_{ILa}(\bar{z}|\phi, \sigma^2) = \int_{\beta} \frac{1}{(2\pi)^{\frac{m}{2}} (\sigma^2)^{\frac{m}{2}} |\bar{C}|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2\sigma^2} (\bar{z} - \bar{X}\beta)^T \bar{C}^{-1} (\bar{z} - \bar{X}\beta) \right\} d\beta. \quad (A7)$$

Next, we apply the following relationship (Harville, 1974) (just preceding his Eq. 2):

$$(\bar{z} - \bar{X}\beta)^T \bar{C}^{-1} (\bar{z} - \bar{X}\beta) = (\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + (\beta - \hat{\beta})^T (\bar{X}^T \bar{C}^{-1} \bar{X}) (\beta - \hat{\beta}). \quad (A8)$$

Combining eqns. (A7) and (A8), and moving some constant terms out of the integral, gives

$$\begin{aligned} f_{ILa}(\bar{z}|\phi, \sigma^2) &= \\ &\frac{1}{(2\pi)^{\frac{m}{2}} (\sigma^2)^{\frac{m}{2}} |\bar{C}|^{\frac{1}{2}}} \int_{\beta} \exp \left\{ -\frac{1}{2\sigma^2} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + (\beta - \hat{\beta})^T (\bar{X}^T \bar{C}^{-1} \bar{X}) (\beta - \hat{\beta}) \right] \right\} d\beta, \end{aligned} \quad (A9)$$

where $\hat{\beta}$ is the generalised least squares estimator for β , and implicitly a function of ϕ ¹:

¹Note that most commonly, the covariance matrix instead of the correlation matrix is used in eqn (A10), however σ^2 cancels out:

$$\begin{aligned} \hat{\beta} &= (\bar{X}^T (\sigma^2 \bar{C})^{-1} \bar{X})^{-1} \bar{X}^T (\sigma^2 \bar{C})^{-1} \bar{z} \\ &= (\bar{X}^T (\sigma^2)^{-1} \bar{C}^{-1} \bar{X})^{-1} \bar{X}^T (\sigma^2)^{-1} \bar{C}^{-1} \bar{z} \\ &= (\bar{X}^T \bar{C}^{-1} \bar{X})^{-1} (\sigma^2)^{-1} \bar{X}^T \bar{C}^{-1} \bar{z} \\ &= (\bar{X}^T \bar{C}^{-1} \bar{X})^{-1} \bar{X}^T \bar{C}^{-1} \bar{z}, \end{aligned}$$

thus $\hat{\beta}$ does not depend on σ^2 .

$$\hat{\beta} = (\bar{\mathbf{X}}^T \bar{\mathbf{C}}^{-1} \bar{\mathbf{X}})^{-1} \bar{\mathbf{X}}^T \bar{\mathbf{C}}^{-1} \bar{\mathbf{z}}. \quad (\text{A10})$$

Then, we split up the exponent (using the general rule $e^{(a+b)} = e^a e^b$), and move additional constant terms out of the integral:

$$f_{ILa}(\bar{\mathbf{z}}|\phi, \sigma^2) = \frac{1}{(2\pi)^{\frac{m}{2}} (\sigma^2)^{\frac{m}{2}} |\bar{\mathbf{C}}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2\sigma^2} [(\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})^T \bar{\mathbf{C}}^{-1} (\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})]\right\} \int_{\beta} \exp\left\{-\frac{1}{2\sigma^2} [(\beta - \hat{\beta})^T (\bar{\mathbf{X}}^T \bar{\mathbf{C}}^{-1} \bar{\mathbf{X}}) (\beta - \hat{\beta})]\right\} d\beta. \quad (\text{A11})$$

Define the mean and covariance parameters for a multivariate normal distribution, respectively:

$$\boldsymbol{\mu}_{m\nu n} = \hat{\beta} \quad (\text{A12})$$

and

$$\boldsymbol{\Sigma}_{m\nu n} = [\bar{\mathbf{X}}^T (\sigma^2 \bar{\mathbf{C}})^{-1} \bar{\mathbf{X}}]^{-1}. \quad (\text{A13})$$

Combining eqns. (A11), (A12) and (A13), and multiplying by a constant term before the integral and its reciprocal inside the integral brings us this equation:

$$f_{ILa}(\bar{\mathbf{z}}|\phi, \sigma^2) = \frac{1}{(2\pi)^{\frac{m}{2}} (\sigma^2)^{\frac{m}{2}} |\bar{\mathbf{C}}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2\sigma^2} [(\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})^T \bar{\mathbf{C}}^{-1} (\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})]\right\} (2\pi)^{\frac{k}{2}} |\boldsymbol{\Sigma}_{m\nu n}|^{\frac{1}{2}} \times \int_{\beta} \frac{1}{(2\pi)^{\frac{k}{2}} |\boldsymbol{\Sigma}_{m\nu n}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} [(\beta - \boldsymbol{\mu}_{m\nu n})^T \boldsymbol{\Sigma}_{m\nu n}^{-1} (\beta - \boldsymbol{\mu}_{m\nu n})]\right\} d\beta, \quad (\text{A14})$$

where the integral part equals the k -variate normal distribution:

$$f_{MVN}(\beta; \boldsymbol{\mu}_{m\nu n}, \boldsymbol{\Sigma}_{m\nu n}) = \frac{1}{(2\pi)^{\frac{k}{2}} |\boldsymbol{\Sigma}_{m\nu n}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} (\beta - \boldsymbol{\mu}_{m\nu n})^T \boldsymbol{\Sigma}_{m\nu n}^{-1} (\beta - \boldsymbol{\mu}_{m\nu n})\right\}. \quad (\text{A15})$$

By definition of a well-defined probability distribution, eqn. (A15) integrates, over all possible values for β , to one². Taking the remaining part of eqn. (A14):

$$f_{ILa}(\bar{\mathbf{z}}|\phi, \sigma^2) = \frac{1}{(2\pi)^{\frac{m}{2}} (\sigma^2)^{\frac{m}{2}} |\bar{\mathbf{C}}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2\sigma^2} [(\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})^T \bar{\mathbf{C}}^{-1} (\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})]\right\} (2\pi)^{\frac{k}{2}} |\boldsymbol{\Sigma}_{m\nu n}|^{\frac{1}{2}}, \quad (\text{A16})$$

and undoing the parametrisation in $\boldsymbol{\Sigma}_{m\nu n}$, and cancelling out wherever possible, provides us with *Integral a*, also known as the REML function (Laird and Ware, 1982) (which is more commonly presented after taking the logarithm):

$$f_{ILa}(\bar{\mathbf{z}}|\phi, \sigma^2) = \frac{\left((\sigma^2)^k |\bar{\mathbf{X}}^T \bar{\mathbf{C}}^{-1} \bar{\mathbf{X}}|^{-1}\right)^{\frac{1}{2}}}{(2\pi)^{\frac{m-k}{2}} (\sigma^2)^{\frac{m}{2}} |\bar{\mathbf{C}}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2\sigma^2} [(\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})^T \bar{\mathbf{C}}^{-1} (\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})]\right\} \\ = \frac{\exp\left\{-\frac{1}{2\sigma^2} [(\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})^T \bar{\mathbf{C}}^{-1} (\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})]\right\}}{(2\pi)^{\frac{m-k}{2}} (\sigma^2)^{\frac{m-k}{2}} |\bar{\mathbf{C}}|^{\frac{1}{2}} |\bar{\mathbf{X}}^T \bar{\mathbf{C}}^{-1} \bar{\mathbf{X}}|^{\frac{1}{2}}}. \quad (\text{A17})$$

In the last equation, note the power in $(\sigma^2)^k$, where k is also the size of $\bar{\mathbf{X}}^T \bar{\mathbf{X}}$, based on the following equality:

$$|\boldsymbol{\Sigma}_{m\nu n}| = |\bar{\mathbf{X}}^T (\sigma^2 \bar{\mathbf{C}})^{-1} \bar{\mathbf{X}}|^{-1} = (\sigma^2)^k |\bar{\mathbf{X}}^T \bar{\mathbf{C}}^{-1} \bar{\mathbf{X}}|^{-1}, \quad (\text{A18})$$

which follows from the general rule that $\det(b\mathbf{A}) = b^k \det(\mathbf{A})$, with b a scalar, $\mathbf{A}$ a square matrix and k the dimensions of $\mathbf{A}$.

²We reserve the term “proper” for any probability distribution that integrates to a finite value rather than exactly to one.

A3 Marginal posterior for ϕ (via *Integral b*)

To arrive at the marginal posterior for ϕ , we have to integrate out β and σ^2 from the likelihood $\times$ prior in eqn. (A4):

$$\begin{aligned} f_p(\phi|\bar{z}) &\propto \int_{\beta, \sigma^2} f_l(\bar{z}|\beta, \sigma^2, \phi) f_0(\beta, \sigma^2, \phi) d\beta d\sigma^2 \\ &= f_{ILb}(\bar{z}|\phi) f_0(\phi) \end{aligned} \quad (A19)$$

where we define *Integrated Likelihood b* as

$$f_{ILb}(\bar{z}|\phi) = \int_{\beta, \sigma^2} f_l(\bar{z}|\beta, \sigma^2, \phi) \frac{1}{(\sigma^2)^{\alpha_0+1}} \exp\left(\frac{-\beta_0}{\sigma^2}\right) d\beta d\sigma^2 \quad (A20)$$

This integral, with $\alpha_0 = \beta_0 = 0$, with an alternative form of the prior for ϕ , is used in Berger et al. (2001). Applying eqn.(A17) (*Integral a*) gives:

$$f_{ILb}(\bar{z}|\phi) = \int_{\sigma^2} \frac{1}{(2\pi)^{\frac{m-k}{2}} (\sigma^2)^{\frac{m-k}{2}} |\bar{C}|^{\frac{1}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2\sigma^2} (\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta})\right\} \frac{1}{(\sigma^2)^{\alpha_0+1}} \exp\left(\frac{-\beta_0}{\sigma^2}\right) d\sigma^2 \quad (A21)$$

Combining, and subsequently moving the elements that do not depend on σ^2 out of the integral gives:

$$f_{ILb}(\bar{z}|\phi) = \frac{1}{(2\pi)^{\frac{m-k}{2}} |\bar{C}|^{\frac{1}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}}} \int \sigma^2 \frac{1}{(\sigma^2)^{\frac{m-k}{2} + \alpha_0 + 1}} \exp\left\{-\frac{1}{2\sigma^2} \left\{(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0\right\}\right\} d\sigma^2. \quad (A22)$$

Next, we define the two parameters for an inverse-gamma distribution:

$$\alpha = \frac{1}{2}(m-k) + \alpha_0 \quad (A23)$$

$$\beta = \frac{1}{2}(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + \beta_0, \quad (A24)$$

and noting the inverse-gamma distribution:

$$f(\sigma^2; \alpha, \beta) = \frac{(\beta)^\alpha}{\Gamma(\alpha)} (\sigma^2)^{-\alpha-1} \exp\left\{-\frac{\beta}{\sigma^2}\right\}, \quad (A25)$$

we can write eqn. (A22) as:

$$f_{ILb}(\bar{z}|\phi) = \frac{1}{(2\pi)^{\frac{m-k}{2}} |\bar{C}|^{\frac{1}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}}} \frac{\Gamma(\alpha)}{(\beta)^\alpha} \int_{\sigma^2} \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma^2)^{-\alpha-1} \exp\left\{-\frac{\beta}{\sigma^2}\right\} d\sigma^2. \quad (A26)$$

Noting that the integral part equals eqn. (A25), and when integrated over all possible values of σ^2 integrates to one (assuming that $\alpha > 0$ and $\beta > 0$), we arrive at

$$f_{ILb}(\bar{z}|\phi) = \frac{1}{(2\pi)^{\frac{m-k}{2}} |\bar{C}|^{\frac{1}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}}} \frac{\Gamma(\alpha)}{(\beta)^\alpha}. \quad (A27)$$

Undoing the parametrisation by α and β

$$f_{ILb}(\bar{z}|\phi) = \frac{\Gamma\left(\frac{m-k+2\alpha_0}{2}\right)}{(2\pi)^{\frac{m-k}{2}} |\bar{C}|^{\frac{1}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}} \left[\frac{1}{2}(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + \beta_0\right]^{\frac{m-k+2\alpha_0}{2}}}, \quad (A28)$$

brings us the final formulation of the integrated likelihood (*Integral b*):

$$\begin{aligned}
f_{ILb}(\bar{z}|\phi) &= \frac{\Gamma\left(\frac{m-k+2\alpha_0}{2}\right)}{(2\pi)^{\frac{m-k}{2}} |\bar{C}|^{\frac{1}{2}} \left|\bar{X}^T \bar{C}^{-1} \bar{X}\right|^{\frac{1}{2}} \left[\frac{1}{2}(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + \beta_0\right]^{\frac{m-k+2\alpha_0}{2}}} \\
&= \frac{\Gamma\left(\frac{m-k+2\alpha_0}{2}\right)}{(2)^{\frac{m-k}{2}} (\pi)^{\frac{m-k}{2}} |\bar{C}|^{\frac{1}{2}} \left|\bar{X}^T \bar{C}^{-1} \bar{X}\right|^{\frac{1}{2}} \left(\frac{1}{2}\right)^{\frac{m-k+2\alpha_0}{2}} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0\right]^{\frac{m-k+2\alpha_0}{2}}} \\
&= \frac{2^{\alpha_0} \Gamma\left(\frac{m-k+2\alpha_0}{2}\right)}{(\pi)^{\frac{m-k}{2}} |\bar{C}|^{\frac{1}{2}} \left|\bar{X}^T \bar{C}^{-1} \bar{X}\right|^{\frac{1}{2}} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0\right]^{\frac{m-k+2\alpha_0}{2}}}.
\end{aligned} \tag{A29}$$

The marginal posterior for ϕ is calculated by multiplying *Integral b* by the prior $f_0(\phi)$, as shown in eq. (A19); in this final step, we also remove all constants, to arrive at:

$$f_p(\phi|\bar{z}) \propto f_0(\phi) \frac{1}{|\bar{C}|^{\frac{1}{2}} \left|\bar{X}^T \bar{C}^{-1} \bar{X}\right|^{\frac{1}{2}} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0\right]^{\frac{m-k+2\alpha_0}{2}}} \tag{A30}$$

Setting $\alpha_0 = \beta_0 = 0$ leads to the form presented in the main manuscript:

$$f_p(\phi|\bar{z}) \propto f_0(\phi) \frac{1}{|\bar{C}|^{\frac{1}{2}} \left|\bar{X}^T \bar{C}^{-1} \bar{X}\right|^{\frac{1}{2}} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta})\right]^{\frac{m-k}{2}}} \tag{A31}$$

Numerically, BAK creates an one-dimensional, regularly spaced grid for ϕ in parameter space, calculates the marginal posterior for each ϕ , and normalises the marginal posterior to a proper distribution.

A4 Marginal posterior for σ^2

The posterior distribution for σ^2 is given by marginalising the full joint posterior, eq. (A4):

$$f(\sigma^2|\bar{z}) \propto \int_{\beta, \phi} f_l(\bar{z}|\beta, \sigma^2, \phi) f_0(\beta, \sigma^2, \phi) d\beta d\phi. \tag{A32}$$

Splitting the prior into its constituent parts (via eq. (A2)), and taking terms outside of the integrals where appropriate, we have:

$$f(\sigma^2|\bar{z}) \propto \int_{\beta, \phi} f_l(\bar{z}|\beta, \sigma^2, \phi) d\beta f_0(\sigma^2) f_0(\phi) d\phi. \tag{A33}$$

The integral with respect to β in the above equation is our *Integral a*, already presented in eqn. (A17). Combining *Integral a* with the prior for σ^2 gives:

$$f(\sigma^2|\bar{z}) \propto \int_{\phi} \frac{1}{(\sigma^2)^{\alpha_0+1}} \exp\left(\frac{-\beta_0}{\sigma^2}\right) \frac{\exp\left\{-\frac{1}{2\sigma^2} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta})\right]\right\}}{(2\pi)^{\frac{m-k}{2}} (\sigma^2)^{\frac{m-k}{2}} |\bar{C}|^{\frac{1}{2}} \left|\bar{X}^T \bar{C}^{-1} \bar{X}\right|^{\frac{1}{2}}} f_0(\phi) d\phi. \tag{A34}$$

Combining and moving, and removing constants gives:

$$f(\sigma^2|\bar{z}) \propto \int_{\phi} \frac{\exp\left\{-\frac{1}{2\sigma^2} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0\right]\right\}}{(\sigma^2)^{\frac{m-k+2\alpha_0+2}{2}} |\bar{C}|^{\frac{1}{2}} \left|\bar{X}^T \bar{C}^{-1} \bar{X}\right|^{\frac{1}{2}}} f_0(\phi) d\phi. \tag{A35}$$

We can set $\alpha_0 = \beta_0 = 0$ and express the marginal posterior for σ^2 as an integral:

$$f_p(\sigma^2|\bar{z}) \propto \int_{\phi} f_0(\phi) \frac{\exp\left\{-\frac{1}{2\sigma^2} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta})\right]\right\}}{|\bar{C}|^{\frac{1}{2}} \left|\bar{X}^T \bar{C}^{-1} \bar{X}\right|^{\frac{1}{2}} (\sigma^2)^{\frac{m-k+2}{2}}} d\phi. \tag{A36}$$

We note an alternative form for expressing the posterior distribution, eq. (A35), in terms of an inverse-gamma distribution for σ^2 and the posterior distribution for ϕ . Redefine the inverse-gamma parameters for this context:

$$\alpha = \frac{m-k+2\alpha_0}{2} \tag{A37}$$

and

$$\beta = \frac{(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0}{2}. \quad (\text{A38})$$

Reminder: the inverse-gamma distribution is, using the above parameters, defined as:

$$f_{IG}(\sigma^2; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \sigma^{2-\alpha-1} \exp\left(-\frac{\beta}{\sigma^2}\right). \quad (\text{A39})$$

Then, we can write:

$$\begin{aligned} f(\sigma^2|\bar{z}) &\propto \int_{\phi} \frac{\exp\left\{-\frac{\beta}{\sigma^2}\right\}}{(\sigma^2)^{\alpha+1} |\bar{C}|^{\frac{1}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}}} f_0(\phi) d\phi \\ &= \int_{\phi} \frac{\Gamma(\alpha)}{\beta^\alpha} \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\exp\left\{-\frac{\beta}{\sigma^2}\right\}}{(\sigma^2)^{\alpha+1} |\bar{C}|^{\frac{1}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}}} f_0(\phi) d\phi \\ &= \int_{\phi} \frac{\Gamma(\alpha)}{\beta^\alpha} \frac{1}{|\bar{C}|^{\frac{1}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}}} f_0(\phi) f_{IG}(\sigma^2; \alpha, \beta) d\phi. \end{aligned} \quad (\text{A40})$$

Explicitly substituting the expressions for α and β into eq. (A40) gives:

$$f(\sigma^2|\bar{z}) \propto \int_{\phi} \frac{\Gamma\left(\frac{m-k+2\alpha_0}{2}\right)}{\left[\frac{(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0}{2}\right]^{\frac{m-k+2\alpha_0}{2}}} \frac{1}{|\bar{C}|^{\frac{1}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}}} f_0(\phi) f_{IG}(\sigma^2; \alpha, \beta) d\phi. \quad (\text{A41})$$

and dropping constant terms leads to:

$$f(\sigma^2|\bar{z}) \propto \int_{\phi} \frac{1}{|\bar{C}|^{\frac{1}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}} \left[\frac{(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0}{2}\right]^{\frac{m-k+2\alpha_0}{2}}} f_0(\phi) f_{IG}(\sigma^2; \alpha, \beta) d\phi. \quad (\text{A42})$$

Comparing this expression with eq. (A30), we can see that:

$$f(\sigma^2|\bar{z}) \propto \int_{\phi} f_p(\phi|\bar{z}) f(\sigma^2|\bar{z}, \phi) d\phi. \quad (\text{A43})$$

where

$$f(\sigma^2|\bar{z}, \phi) = f_{IG}\left(\sigma^2; \frac{m-k+2\alpha_0}{2}, \frac{(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0}{2}\right) \quad (\text{A44})$$

Numerically, BAK extends the vector with ϕ to a 2D grid in parameterspace of ϕ and σ^2 , and calculates the joint posterior (the integrand) for σ^2 and ϕ for every gridpoint; then it performs a trapezoidal integration over ϕ , and normalises, to arrive at the marginal distribution for σ^2 .

A5 Marginal posteriors for β_q (via *Integral c*)

To arrive at $f_p(\beta_q|\bar{z})$, we first derive $f_p(\beta|\bar{z}, \phi)$ by again marginalising the full posterior distribution, eq. (A4):

$$f_p(\beta|\bar{z}) \propto \int_{\sigma^2, \phi} f_l(\bar{z}|\beta, \sigma^2, \phi) f_0(\beta, \sigma^2, \phi) d\sigma^2 d\phi. \quad (\text{A45})$$

Again splitting the prior into its constituent parts (via eq. (A2)), and taking terms outside of the integrals where appropriate, we have:

$$\begin{aligned} f_p(\beta|\bar{z}) &\propto \int_{\sigma^2, \phi} f_l(\bar{z}|\beta, \sigma^2, \phi) f_0(\sigma^2) d\sigma^2 f_0(\phi) d\phi \\ &= \int_{\phi} f_{ILc}(\bar{z}|\beta, \phi) f_0(\phi) d\phi \end{aligned} \quad (\text{A46})$$

where we define *Integrated Likelihood c* (ILc) as:

$$f_{ILc}(\bar{z}|\beta, \phi) = \int_{\sigma^2} f_l(\bar{z}|\beta, \sigma^2, \phi) f_0(\sigma^2) d\sigma^2. \quad (A47)$$

Focusing now on the expression for the *Integrated Likelihood c*, we rearrange, and move the constant part out of the integral to give:

$$\begin{aligned} f_{ILc}(\bar{z}|\beta, \phi) &= \int_{\sigma^2} \frac{1}{(2\pi)^{\frac{m}{2}} |\bar{C}|^{\frac{1}{2}}} \frac{1}{(\sigma^2)^{\frac{m}{2}}} \exp\left\{-\frac{1}{2\sigma^2} \left[(\bar{z} - \bar{X}\beta)^T \bar{C}^{-1} (\bar{z} - \bar{X}\beta)\right]\right\} \frac{1}{(\sigma^2)^{\alpha_0+1}} \exp\left(\frac{-\beta_0}{\sigma^2}\right) d\sigma^2 \\ &= \frac{1}{(2\pi)^{\frac{m}{2}} |\bar{C}|^{\frac{1}{2}}} \int_{\sigma^2} \frac{1}{(\sigma^2)^{\frac{m}{2} + \alpha_0 + 1}} \exp\left\{-\frac{1}{2\sigma^2} \left[(\bar{z} - \bar{X}\beta)^T \bar{C}^{-1} (\bar{z} - \bar{X}\beta) + 2\beta_0\right]\right\} d\sigma^2. \end{aligned} \quad (A48)$$

Now we redefine the inverse-gamma parameters for this context:

$$\alpha = \frac{m}{2} + \alpha_0, \quad (A49)$$

$$\beta = \frac{1}{2} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0 \right], \quad (A50)$$

and - for readability - we repeat the inverse-gamma distribution with above parameters:

$$f_{IG}(\sigma^2; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma^2)^{-\alpha-1} \exp\left\{-\frac{\beta}{\sigma^2}\right\}. \quad (A51)$$

Thus we can write eqn. (A48) as eqn. (A52) in which the integrand is replaced by eqn. (A51):

$$f_{ILc}(\bar{z}|\beta, \phi) = \frac{1}{(2\pi)^{\frac{m}{2}} |\bar{C}|^{\frac{1}{2}}} \frac{\Gamma(\alpha)}{\beta^\alpha} \int_{\sigma^2} \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma^2)^{-\alpha-1} \exp\left\{-\frac{\beta}{\sigma^2}\right\} d\sigma^2. \quad (A52)$$

Applying that the integral integrates to one, and reversing the parametrisation

$$\begin{aligned} f_{ILc}(\bar{z}|\beta, \phi) &= \frac{\Gamma\left(\frac{m}{2} + \alpha_0\right)}{(2\pi)^{\frac{m}{2}} |\bar{C}|^{\frac{1}{2}} \left[\frac{1}{2}(\bar{z} - \bar{X}\beta)^T \bar{C}^{-1} (\bar{z} - \bar{X}\beta) + \beta_0\right]^{\frac{m}{2} + \alpha_0}} \\ &= \frac{\Gamma\left(\frac{m}{2} + \alpha_0\right)}{(\pi)^{\frac{m}{2}} (2)^{\frac{m}{2}} |\bar{C}|^{\frac{1}{2}} \left[(\bar{z} - \bar{X}\beta)^T \bar{C}^{-1} (\bar{z} - \bar{X}\beta) + 2\beta_0\right]^{\frac{m}{2} + \alpha_0} \left(\frac{1}{2}\right)^{\frac{m}{2} + \alpha_0}} \\ &= \frac{2^{\alpha_0} \Gamma\left(\frac{m}{2} + \alpha_0\right)}{(\pi)^{\frac{m}{2}} |\bar{C}|^{\frac{1}{2}} \left[(\bar{z} - \bar{X}\beta)^T \bar{C}^{-1} (\bar{z} - \bar{X}\beta) + 2\beta_0\right]^{\frac{m}{2} + \alpha_0}} \end{aligned} \quad (A53)$$

Next, returning to the joint marginal posterior for β , eq. (A46), applying the relationship in eqn. (A8) where $\hat{\beta}$ is again the GLS estimate and thus a function of ϕ (and C), and further removing all constants gives:

$$\begin{aligned} f_p(\beta|\bar{z}) &\propto \int_{\phi} f_{ILc}(\bar{z}|\beta, \phi) f_0(\phi) d\phi \\ &\propto \int_{\phi} \frac{2^{\alpha_0} \Gamma\left(\frac{m}{2} + \alpha_0\right)}{(\pi)^{\frac{m}{2}} |\bar{C}|^{\frac{1}{2}} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + (\beta - \hat{\beta})^T (\bar{X}^T \bar{C}^{-1} \bar{X}) (\beta - \hat{\beta}) + 2\beta_0\right]^{\frac{m}{2} + \alpha_0}} f_0(\phi) d\phi \\ &\propto \int_{\phi} \frac{1}{|\bar{C}|^{\frac{1}{2}} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + (\beta - \hat{\beta})^T (\bar{X}^T \bar{C}^{-1} \bar{X}) (\beta - \hat{\beta}) + 2\beta_0\right]^{\frac{m+2\alpha_0}{2}}} f_0(\phi) d\phi. \end{aligned} \quad (A54)$$

We rearrange the sum in eqn. (A54) [equivalent to $a + b \propto 1 + b/a, a \neq 0$], extend the power with $-k + k$, and multiply the main fraction inside the denominator by $\frac{m-k+2\alpha_0}{m-k+2\alpha_0}$ respectively:

$$\begin{aligned}
f_p(\beta|\bar{z}) &\propto \int_{\phi} \frac{1}{|\bar{C}|^{\frac{1}{2}} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0 \right]^{\frac{m+2\alpha_0}{2}} \left[1 + \frac{(\beta - \hat{\beta})^T (\bar{X}^T \bar{C}^{-1} \bar{X}) (\beta - \hat{\beta})}{(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0} \right]^{\frac{m+\alpha_0}{2}}} f_0(\phi) d\phi \\
&\propto \int_{\phi} \frac{1}{|\bar{C}|^{\frac{1}{2}} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0 \right]^{\frac{m+2\alpha_0}{2}} \left[1 + \frac{(\beta - \hat{\beta})^T (\bar{X}^T \bar{C}^{-1} \bar{X}) (\beta - \hat{\beta})}{(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0} \right]^{\frac{m-k+k+2\alpha_0}{2}}} f_0(\phi) d\phi \\
&\propto \int_{\phi} \frac{1}{|\bar{C}|^{\frac{1}{2}} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0 \right]^{\frac{m+2\alpha_0}{2}}} \\
&\quad \times \frac{1}{\left[1 + \frac{1}{m-k+2\alpha_0} (\beta - \hat{\beta})^T \frac{(\bar{X}^T \bar{C}^{-1} \bar{X})}{((\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0)/(m-k+2\alpha_0)} (\beta - \hat{\beta}) \right]^{\frac{m-k+k+2\alpha_0}{2}}} f_0(\phi) d\phi.
\end{aligned} \tag{A55}$$

Note that $(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta})/(m-k)$ (i.e. the denominator with $\alpha_0 = \beta_0 = 0$) is the REML estimate of σ^2 given $\bar{C}^{-1}$. Also note that the form is now comparable to that of the multivariate t distribution (MVT) (Roth, 2013; Gelman et al., 2013), with degrees of freedom ν^3

$$\nu = m - k + 2\alpha_0, \tag{A56}$$

location vector

$$\mu = \hat{\beta}, \tag{A57}$$

and shape matrix Σ , where

$$\Sigma^{-1} = \frac{(\bar{X}^T \bar{C}^{-1} \bar{X})}{((\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0)/(m-k+2\alpha_0)}. \tag{A58}$$

The resulting MVT is defined as:

$$t_{\nu}(\beta; \hat{\beta}, \Sigma) = \frac{\Gamma[(\nu+k)/2]}{\Gamma(\nu/2) \nu^{k/2} \pi^{k/2} |\Sigma|^{1/2}} \left[1 + \frac{1}{\nu} (\beta - \hat{\beta})^T \Sigma^{-1} (\beta - \hat{\beta}) \right]^{-(\nu+k)/2} \tag{A59}$$

and therefore (A55) can be written as:

$$f_p(\beta|\bar{z}) \propto \int_{\phi} \frac{1}{|\bar{C}|^{\frac{1}{2}} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0 \right]^{\frac{m+2\alpha_0}{2}}} \frac{\Gamma(\nu/2) \nu^{k/2} \pi^{k/2} |\Sigma|^{1/2}}{\Gamma[(\nu+k)/2]} t_{\nu}(\beta; \hat{\beta}, \Sigma) f_0(\phi) d\phi \tag{A60}$$

We note that the determinant of Σ is given by

$$|\Sigma| = \frac{\left[((\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0)/(m-k+2\alpha_0) \right]^k}{|\bar{X}^T \bar{C}^{-1} \bar{X}|}, \tag{A61}$$

and therefore dropping constant terms and rearranging leads to:

³Don't confuse the degrees of freedom ν of a t distribution with the smoothness parameter of the Matérn covariance function in the main paper, represented by ν_{sim} or ν_M .

$$\begin{aligned}
f_p(\beta|\bar{z}) &\propto \int_{\phi} \frac{1}{|\bar{C}|^{\frac{1}{2}} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0 \right]^{\frac{m+2\alpha_0}{2}}} |\Sigma|^{1/2} t_v(\beta; \hat{\beta}, \Sigma) f_0(\phi) d\phi \\
&\propto \int_{\phi} \frac{\left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0 \right]^{k/2}}{|\bar{C}|^{\frac{1}{2}} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0 \right]^{\frac{m+2\alpha_0}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}}} t_v(\beta; \hat{\beta}, \Sigma) f_0(\phi) d\phi \\
&\propto \int_{\phi} \frac{1}{|\bar{C}|^{\frac{1}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}} \left[(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta}) + 2\beta_0 \right]^{\frac{m-k+2\alpha_0}{2}}} f_0(\phi) t_v(\beta; \hat{\beta}, \Sigma) d\phi
\end{aligned} \tag{A62}$$

Comparing this expression with eq. (A30), we can write:

$$f_p(\beta|\bar{z}) \propto \int_{\phi} f_p(\phi|\bar{z}) t_v(\beta; \hat{\beta}, \Sigma) d\phi. \tag{A63}$$

According to Roth (2013) the marginal distributions of the multivariate t distribution are also t (his eqn. (4.3)):

$$t_v(\beta_q; \hat{\beta}_q, \Sigma_q) = \frac{\Gamma[(v+k_q)/2]}{\Gamma(v/2) v^{k_q/2} \pi^{k_q/2} |\Sigma_q|^{1/2}} \left[1 + \frac{1}{v} (\beta_q - \hat{\beta}_q)^T \Sigma_q^{-1} (\beta_q - \hat{\beta}_q) \right]^{-(v+k_q)/2}. \tag{A64}$$

with k_q the dimension of β_q , in our case one, $\hat{\beta}_q$ the q th element of $\hat{\beta}$ as the location parameter and $\Sigma_q = v[\hat{\beta}]_{q,q}$ the q th diagonal of matrix $v[\hat{\beta}] = \Sigma^{-1}$, as the scale parameter. Therefore in our case:

$$t_v(\beta_q; \hat{\beta}_q, \Sigma_q) = \frac{\Gamma[(v+1)/2]}{\Gamma(v/2)(v\pi)^{1/2} |\Sigma_q|^{1/2}} \left[1 + \frac{1}{v} (\beta_q - \hat{\beta}_q)^T \Sigma_q^{-1} (\beta_q - \hat{\beta}_q) \right]^{-(v+1)/2}, \tag{A65}$$

and the final marginal posterior for β_q is given by:

$$f_p(\beta_q|\bar{z}) \propto \int_{\phi} f_p(\phi|\bar{z}) t_v(\beta_q; \hat{\beta}_q, \Sigma_q) d\phi \tag{A66}$$

Similarly to Section A4, BAK creates 2-dimensional grids covering the parameter spaces of ϕ and β_q (for all q) and applies the trapezoidal rule to calculate the integral over ϕ in Eq. (A65); finally it normalises to get the marginal distributions for the individual β_q 's.

A6 Posterior predictive distribution $f_p(z^*|\bar{z})$

We consider our spatial phenomenon of interest a Gaussian Random field, therefore we can define our n^* point predictions (z^*) together with m area observations ($\bar{z}$) as one $n^* + m$ -dimensional multivariate normal distribution:

$$\begin{bmatrix} z^* \\ \bar{z} \end{bmatrix} | \beta, \sigma^2, \phi \sim MVN_{n^*+m} \left(\begin{bmatrix} X^* \beta \\ \bar{X} \beta \end{bmatrix}, \sigma^2 \begin{bmatrix} C^{**} & \bar{C}^* \\ (\bar{C}^*)^T & \bar{C} \end{bmatrix} \right), \tag{A67}$$

with C^{**} the correlation matrix between prediction points, dimension $n^* \times n^*$, and $\bar{C}^*$ the mean correlation between observation area and prediction points, dimension $n^* \times m$.

The posterior predictive distribution is defined by:

$$f_p(z^*|\bar{z}) = \int_{\phi} f(z^*|\bar{z}, \phi) f_p(\phi|\bar{z}) d\phi. \tag{A68}$$

Taking out the first part under the integral

$$\begin{aligned}
f(z^*|\bar{z}, \phi) &= \int_{\beta, \sigma^2} f(z^*, \beta, \sigma^2|\bar{z}, \phi) d\beta d\sigma^2 \\
&= \int_{\beta, \sigma^2} f(z^*|\bar{z}, \beta, \sigma^2, \phi) f(\beta, \sigma^2|\bar{z}, \phi) d\beta d\sigma^2 \\
&= \int_{\beta, \sigma^2} f(z^*|\bar{z}, \beta, \sigma^2, \phi) f(\beta|\bar{z}, \sigma^2, \phi) f(\sigma^2|\bar{z}, \phi) d\beta d\sigma^2 \\
&= \int_{\sigma^2} \left(\int_{\beta} f(z^*|\bar{z}, \beta, \sigma^2, \phi) f(\beta|\bar{z}, \sigma^2, \phi) d\beta \right) f(\sigma^2|\bar{z}, \phi) d\sigma^2
\end{aligned} \tag{A69}$$

The integral with respect to β results in a Gaussian pdf (Kitanidis, 1986, chapter 4) for z^* , known as universal kriging (here formulated for the area-to-point context) with mean:

$$\hat{z}^*|\sigma^2, \phi = \bar{C}^* C^{-1} (\bar{z} - \bar{X}\hat{\beta}) + X^* \hat{\beta}, \tag{A70}$$

and corresponding variance-covariance matrix:

$$\text{var}[\hat{z}^*|\sigma^2, \phi] = \sigma^2 A[\hat{z}^*|\phi], \tag{A71}$$

where

$$A[\hat{z}^*|\phi] = C^{**} - \bar{C}^* C^{-1} (\bar{C}^*)^T + (X^* - \bar{C}^* C^{-1} \bar{X}) (\bar{X}^T C^{-1} \bar{X})^{-1} (X^* - \bar{C}^* C^{-1} \bar{X})^T. \tag{A72}$$

Note that we write $\hat{z}^*|\sigma^2, \phi$ and $\text{var}[\hat{z}^*|\sigma^2, \phi]$ here to make explicit the dependence of the distribution that these values parameterise on the parameters σ^2 and ϕ . Also note that we have cancelled out σ^2 wherever possible, in particular so that $\hat{z}^*|\sigma^2, \phi$ does not actually depend on σ^2 , thus we write $\hat{z}^*|\sigma^2, \phi = \hat{z}^*|\phi$, and that the diagonal elements of matrix $\text{var}[\hat{z}^*]$ are the prediction variances, also known as kriging variances, as often presented in other literature.

Now we have:

$$f(z^*|\bar{z}, \phi) \propto \int_{\sigma^2} \frac{1}{|\sigma^2 A[\hat{z}^*|\phi]|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2\sigma^2} (z^* - \hat{z}^*|\phi)^T A[\hat{z}^*|\phi]^{-1} (z^* - \hat{z}^*|\phi) \right\} f(\sigma^2|\bar{z}, \phi) d\sigma^2 \tag{A73}$$

For $f(\sigma^2|\bar{z}, \phi)$, we refer back to the REML equation in Equation (A17) preceeded by the prior for σ^2 , expressed as an inverse-gamma distribution with parameters α_0 and β_0 :

$$f(\sigma^2|\bar{z}, \phi) \propto \frac{1}{(\sigma^2)^{\alpha_0+1}} \exp \left(-\frac{\beta_0}{\sigma^2} \right) \times \frac{\exp \left\{ -\frac{1}{2\sigma^2} [(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta})] \right\}}{(2\pi)^{\frac{m-k}{2}} (\sigma^2)^{\frac{m-k}{2}} |\bar{C}|^{\frac{1}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}}} \tag{A74}$$

Combining the previous two equations, and taking out the constant $(2\pi)^{\frac{m-k}{2}}$:

$$\begin{aligned}
f(z^*|\bar{z}, \phi) &\propto \int_{\sigma^2} \frac{1}{|\sigma^2 A[\hat{z}^*|\phi]|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2\sigma^2} (z^* - \hat{z}^*|\phi)^T A[\hat{z}^*|\phi]^{-1} (z^* - \hat{z}^*|\phi) \right\} \\
&\quad \times \frac{1}{(\sigma^2)^{\alpha_0+1}} \exp \left(-\frac{\beta_0}{\sigma^2} \right) \frac{\exp \left\{ -\frac{1}{2\sigma^2} [(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta})] \right\}}{(\sigma^2)^{\frac{m-k}{2}} |\bar{C}|^{\frac{1}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}}} d\sigma^2
\end{aligned} \tag{A75}$$

Separating (and subsequently combining) exponents and denominators:

$$\begin{aligned}
f(z^*|\bar{z}, \phi) &\propto \int_{\sigma^2} \frac{1}{|\sigma^2 A[\hat{z}^*|\phi]|^{\frac{1}{2}} (\sigma^2)^{\alpha_0+1} (\sigma^2)^{\frac{m-k}{2}} |\bar{C}|^{\frac{1}{2}} |\bar{X}^T \bar{C}^{-1} \bar{X}|^{\frac{1}{2}}} \\
&\quad \times \exp \left\{ -\frac{1}{2\sigma^2} (z^* - \hat{z}^*|\phi)^T A[\hat{z}^*|\phi]^{-1} (z^* - \hat{z}^*|\phi) - \frac{\beta_0}{\sigma^2} - \frac{1}{2\sigma^2} [(\bar{z} - \bar{X}\hat{\beta})^T \bar{C}^{-1} (\bar{z} - \bar{X}\hat{\beta})] \right\} d\sigma^2
\end{aligned} \tag{A76}$$

Next, we separate elements without σ^2 from the integral, and reorganise. Note that the dimension of $A[\hat{z}^*|\phi]$ is $n^* \times n^*$.

$$f(\mathbf{z}^*|\bar{\mathbf{z}}, \phi) \propto \frac{1}{|\mathbf{A}[\hat{\mathbf{z}}^*|\phi]|^{\frac{1}{2}} |\bar{\mathbf{C}}|^{\frac{1}{2}} |\bar{\mathbf{X}}^T \bar{\mathbf{C}}^{-1} \bar{\mathbf{X}}|^{\frac{1}{2}} \sigma^2} \int \frac{1}{(\sigma^2)^{\frac{n^*+m-k}{2} + \alpha_0 + 1}} \times \exp \left\{ -\frac{1}{\sigma^2} \left[\frac{1}{2} (\mathbf{z}^* - \hat{\mathbf{z}}^*|\phi)^T \mathbf{A}[\hat{\mathbf{z}}^*|\phi]^{-1} (\mathbf{z}^* - \hat{\mathbf{z}}^*|\phi) + \frac{1}{2} [(\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})^T \bar{\mathbf{C}}^{-1} (\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})] + \beta_0 \right] \right\} d\sigma^2 \quad (\text{A77})$$

Now, the integral part can be rewritten as proportional to an inverse-gamma distribution with shape

$$\alpha = \frac{n^* + m - k}{2} + \alpha_0, \quad (\text{A78})$$

and scale

$$\beta = \frac{1}{2} [(\mathbf{z}^* - \hat{\mathbf{z}}^*|\phi)^T \mathbf{A}[\hat{\mathbf{z}}^*|\phi]^{-1} (\mathbf{z}^* - \hat{\mathbf{z}}^*|\phi)] + \frac{1}{2} [(\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})^T \bar{\mathbf{C}}^{-1} (\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})] + \beta_0. \quad (\text{A79})$$

Integrating out σ^2 , keeping the equation correct, and then dropping constant term gives:

$$\begin{aligned} f(\mathbf{z}^*|\bar{\mathbf{z}}, \phi) &\propto \frac{1}{|\mathbf{A}[\hat{\mathbf{z}}^*|\phi]|^{\frac{1}{2}} |\bar{\mathbf{C}}|^{\frac{1}{2}} |\bar{\mathbf{X}}^T \bar{\mathbf{C}}^{-1} \bar{\mathbf{X}}|^{\frac{1}{2}}} \frac{\Gamma(\alpha)}{\beta^\alpha} \\ &\propto \frac{1}{|\mathbf{A}[\hat{\mathbf{z}}^*|\phi]|^{\frac{1}{2}} |\bar{\mathbf{C}}|^{\frac{1}{2}} |\bar{\mathbf{X}}^T \bar{\mathbf{C}}^{-1} \bar{\mathbf{X}}|^{\frac{1}{2}} \beta^\alpha}. \end{aligned} \quad (\text{A80})$$

Dropping more terms that do not depend on $\mathbf{z}^*$:

$$f(\mathbf{z}^*|\bar{\mathbf{z}}, \phi) \propto \frac{1}{\beta^\alpha}. \quad (\text{A81})$$

Substituting back α and β , and rearranging:

$$\begin{aligned} f(\mathbf{z}^*|\bar{\mathbf{z}}, \phi) &\propto \frac{1}{\left[\frac{1}{2} (\mathbf{z}^* - \hat{\mathbf{z}}^*|\phi)^T \mathbf{A}[\hat{\mathbf{z}}^*|\phi]^{-1} (\mathbf{z}^* - \hat{\mathbf{z}}^*|\phi) + \frac{1}{2} [(\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})^T \bar{\mathbf{C}}^{-1} (\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})] + \beta_0 \right]^{\frac{n^*+m-k}{2} + \alpha_0}} \\ &\propto \frac{1}{\left[1 + \frac{(\mathbf{z}^* - \hat{\mathbf{z}}^*|\phi)^T \mathbf{A}[\hat{\mathbf{z}}^*|\phi]^{-1} (\mathbf{z}^* - \hat{\mathbf{z}}^*|\phi)}{(\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})^T \bar{\mathbf{C}}^{-1} (\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta}) + 2\beta_0} \right]^{\frac{n^*+m-k}{2} + \alpha_0}} \\ &\propto \frac{1}{\left[1 + \frac{1}{m-k+2\alpha_0} \frac{(\mathbf{z}^* - \hat{\mathbf{z}}^*|\phi)^T \mathbf{A}[\hat{\mathbf{z}}^*|\phi]^{-1} (\mathbf{z}^* - \hat{\mathbf{z}}^*|\phi)}{((\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})^T \bar{\mathbf{C}}^{-1} (\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta}) + 2\beta_0)/(m-k+2\alpha_0)} \right]^{\frac{n^*+m-k+2\alpha_0}{2}}} \end{aligned} \quad (\text{A82})$$

This final expression is proportional to the multivariate t distribution (Roth, 2013) with – in this context – degrees of freedom

$$\nu_t = m - k + 2\alpha_0, \quad (\text{A83})$$

mean

$$\mu = \hat{\mathbf{z}}^*|\phi, \quad (\text{A84})$$

and shape matrix parameter

$$\Sigma_t = \frac{(\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta})^T \bar{\mathbf{C}}^{-1} (\bar{\mathbf{z}} - \bar{\mathbf{X}}\hat{\beta}) + 2\beta_0}{m - k + 2\alpha_0} \mathbf{A}[\hat{\mathbf{z}}^*|\phi]. \quad (\text{A85})$$

Its variance is given by $\frac{v_t}{v_t-2}\Sigma_t$ (A86). Both mean and variance are functions of ϕ . Note that Σ_t with $\alpha_0 = \beta_0 = 0$ equals the UK variance with REML estimated $\hat{\sigma}^2$. Equation (A86) shows that the variance slightly increases now the uncertainty in σ^2 is considered, as one would expect.

We will approximate Equation (A68) by a finite sum using trapezoidal integration, that is:

$$f_p(z^*|\bar{z}) \approx \sum_{i=1}^G \frac{1}{2} [f(z^*|\bar{z}, \phi_i) f_p(\phi_i|\bar{z}) + f(z^*|\bar{z}, \phi_{i-1}) f_p(\phi_{i-1}|\bar{z})] \delta_{\phi_i}, \quad (\text{A87})$$

where $\delta_{\phi_i} = \phi_i - \phi_{i-1}$ is the i th spacing between consecutive values in the 1D ϕ parameter grid, $\phi_i; i = 0, \dots, G$. This can be rewritten as:

$$f_p(z^*|\bar{z}) \approx \sum_{i=0}^G w_i f(z^*|\bar{z}, \phi_i), \quad (\text{A88})$$

where we define the weights:

$$w_i = \begin{cases} \frac{1}{2} f_p(\phi_0|\bar{z})(\phi_1 - \phi_0) & \text{if } i = 0 \\ \frac{1}{2} f_p(\phi_G|\bar{z})(\phi_G - \phi_{G-1}) & \text{if } i = G \\ \frac{1}{2} f_p(\phi_i|\bar{z})(\phi_{i+1} - \phi_{i-1}) & \text{otherwise} \end{cases} \quad (\text{A89})$$

This defines a mixture distribution for z^* , with weights given by w_i . Its mean is

$$\hat{z}^* = \sum_{i=0}^G w_i E[f(z^*|\bar{z}, \phi_i)] \quad (\text{A90})$$

and its variance is

$$\text{var}[\hat{z}^*] = \sum_{i=0}^G w_i \{E[f(z^*|\bar{z}, \phi_i)]^2 + \text{var}[f(z^*|\bar{z}, \phi_i)] - \hat{z}^*\} \quad (\text{A91})$$

Numerically, BAK first creates for each prediction point s^* a vector of predictions and a vector of corresponding prediction variances, both as function of ϕ . Finally, the algorithm calculates the mean and variance of the posterior predictive distribution (or more formally of a finite mixture distribution that approximates this distribution, with weights defined based on $f_p(\phi|\bar{z})$ and the spacings of the ϕ parameter grid, eq. (A89)).

A7 Connect to R code

Function `postc1abetaz0` as pseudo-code

- Create parametergrids for
 - ϕ (`aVec`)
 - σ^2 (`c1Vec`)
 - β (`betaMat`)
- Create discretized log prior for ϕ (`lnf0a`)
- For each ϕ in `aVec`:
 - Calculate, using function `intLikb`, the log likelihood `lnfla`. This function also returns other variables to be used in the following calculations, such as: `betahat`, `sigma2hat`, `Ac` and `resinvAres`
 - Calculate log posterior σ^2 (as grid: `lnfpc1a`) for each σ^2 in `c1Vec`
 - Calculate the marginal log posterior for each β_q (as grid: `lnfpbetaa`), using function `margbetab`, for each β in `betaMat`
 - Calculate posterior prediction (`z0IFa`) and the associated uncertainty (`v0IFa`) using function `fz0IFa`
- For σ^2 , using the 2 dimensional variable `lnfpc1a`, backtransform the (log) posterior probabilities, then marginalize and normalize
- For ϕ : Backtransform and normalize vector `lnfla`
- For β_q : Backtransform, marginalize and normalize the $1 \times p$ dimensional variable `margbetab`
- Marginalize predictions $z^*(s)$ over ϕ and calculate corresponding prediction uncertainties $v^*(s)$, using `z0IFa` and `v0IFa`.

References

- Berger, J. O., de Oliveira, V., and Sansó, B. Objective bayesian analysis of spatially correlated data. *Journal of the American Statistical Association*, 96(456):1361–1374, 2001. ISSN 01621459. URL <http://www.jstor.org/stable/3085905>.
- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. *Bayesian data analysis*. Texts in statistical science. Chapman and Hall/CRC, Boca Raton, FL [etc.], 3rd edition edition, 2013. ISBN 9781439840955.
- Harville, D. A. Bayesian inference for variance components using only error contrasts. *Biometrika*, 61(2):383–385, 1974. ISSN 00063444. doi: 10.2307/2334370. URL <http://www.jstor.org/stable/2334370>.
- Kitanidis, P. K. Parameter uncertainty in estimation of spatial functions: Bayesian analysis. *Water Resources Research*, 22(4):499–507, 1986. doi: 10.1029/WR022i004p00499. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/WR022i004p00499>.
- Laird, N. M. and Ware, J. H. Random-effects models for longitudinal data. *Biometrics*, 38(4):963–74, 1982. ISSN 0006-341X (Print) 0006-341x.
- Roth, M. On the multivariate t distribution. Report, Department of Electrical Engineering, Linköpings universitet, 2013. URL <http://users.isy.liu.se/en/rt/roth/student.pdf>.