Supplementary Material for “Bayesian analysis of measurement error models using INLA”

Stefanie Muff
Division of Biostatistics, Institute of Social and Preventive Medicine, University of Zurich, Switzerland.
Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Switzerland.

Andrea Riebler
Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway.

Leonhard Held
Division of Biostatistics, Institute of Social and Preventive Medicine, University of Zurich, Switzerland.

Håvard Rue
Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway.

Philippe Saner
Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Switzerland.

1. R-code for the three applications in the main text

In this section we guide the reader through the r-inla code and technical details of the three examples discussed in the main text. On www.r-inla.org/examples/case-studies/muff-etal-2014 selected data and R-code are provided for download. The r-inla package can be installed by typing the following command line in the R terminal:

source("http://www.math.ntnu.no/inla/givemeINLA.R")
upgrade.inla(testing=TRUE)

Using
inla.version()

information regarding the actual installed version is shown. Here, we used the r-inla version built on December 4, 2013. For more information regarding the installation process we refer to www.r-inla.org.

Address for correspondence: Leonhard Held, Division of Biostatistics, Institute of Social and Preventive Medicine, University of Zurich, Hirschengraben 84, 8001 Zurich, Switzerland.
E-mail: leonhard.held@ifspm.uzh.ch

* These authors contributed equally to this work.
1.1. **Inbreeding in Swiss ibex populations (classical error)**

Let all variables be defined as in Section 5.1 of the main text. Recall that the model is Gaussian and contains five covariates \((x, z_1, z_2, z_3, z_4)\). The covariate \(x\) is not directly observed, but only a proxy \(w\) following a classical heteroscedastic error model \(w \mid x \sim N(x, \tau_u D)\). The prior distributions are elicited from expert/prior knowledge, see main text, and are defined as:

- \(x \sim N(0, \tau_x I)\).
- \(\beta_0, \beta_x, \beta_{z_1}, \beta_{z_2}, \beta_{z_3}, \beta_{z_4} \sim N(0, \tau_\beta)\), with \(\tau_\beta = 0.0001\),
- \(\tau_x \sim G(\alpha_x, \beta_x)\), with \(\alpha_x = 1.194\) and \(\beta_x = 0.00085\),
- \(\tau_y \sim G(\alpha_y, \beta_y)\), with \(\alpha_y = 0.903\) and \(\beta_y = 0.0014\),
- \(\tau_u \sim G(\alpha_u, \beta_u)\), with \(\alpha_u = 8.5\) and \(\beta_u = 7.5\).

The object `data` consists of seven columns:

\[
\begin{align*}
  y & \quad w \quad z_1 \quad z_2 \quad z_3 \quad z_4 \quad \text{error.prec}
\end{align*}
\]

They contain (for \(n = 26\)):

- \(y_1 \ldots y_n\): The populations' intrinsic growth rates.
- \(w_1 \ldots w_n\): The estimated inbreeding coefficients (proxies for \(x_1 \ldots x_n\); centered).
- \(z_{11} \ldots z_{1n}\): Length of the time series (centered).
- \(z_{21} \ldots z_{2n}\): Average precipitation in summer (centered).
- \(z_{31} \ldots z_{3n}\): Average precipitation in winter (centered).
- \(z_{41} \ldots z_{4n}\): Interaction between \(z_1\) and \(z_2\).
- \(\text{error.prec}_1 \ldots \text{error.prec}_n\): The error precisions in the estimates \(w\).

Start with the prior specification process as described above and in the main text:

```r
data <- read.table("ibex_data4supp.txt", header=T) attach(data) prior.beta <- c(0, 0.0001) prior.prec.x <- c(1.194, 0.00085) prior.prec.y <- c(0.903, 0.0014) prior.prec.u <- c(8.5, 7.5)
```

# initial values (mean or mode of prior)
```r
prec.x = 1.194/0.00085
prec.y = 0.903/0.0014
prec.u = 1
```
Next, we define the INLA model formula. There are four fixed effects \((\beta_1, \beta_2, \beta_3, \beta_4)\) and one hyperparameter \(\beta_x\) belonging to the error-prone covariate \(x\), where the new \texttt{mec} model is employed for the latter. Note that the heteroscedasticity in the error in \(w\) is encoded by assigning the vector of error precisions \texttt{error.prec} to the \texttt{scale} option. In the \texttt{values} option, all values of \(w\) must be listed. The model contains four hyperparameters:

- \texttt{beta} corresponds to \(\beta_x\), the slope coefficient of the error-prone covariate \(x\), with a Gaussian prior.
- \texttt{prec.u} is the error precision \(\tau_u\) with gamma prior.
- \texttt{prec.x} is the precision \(\tau_x\) of \(x \sim N(\alpha_0, \tau_x)\) with gamma prior.
- \texttt{mean.x} corresponds to the mean \(\alpha_0\), which is fixed at zero due to covariate centering.

The prior settings are defined in the different entries of the list \texttt{hyper}. The option \texttt{fixed} specifies whether the corresponding quantity should be estimated or fixed at the initial value. The field \texttt{param} captures the prior parameters of the corresponding prior distribution. Gaussian prior distributions are the default for \texttt{beta} and \texttt{mean.x}, while log-gamma distributions are used for the log-transformed precisions \texttt{prec.u} and \texttt{prec.x}. Note hereby that if a variable \(\tau\) is gamma distributed with shape parameter \(a\) and rate parameter \(b\) leading to the mean \(a/b\) and variance \(a/b^2\), then \(\log(\tau)\) is log-gamma distributed with the same parameters \(a\) and \(b\).

```r
library(INLA)

formula <- y ~ f(w, model = "mec", scale = error.prec, values = w,
hyperl = list(  
  beta = list(    
    param = prior.beta,  
    fixed = FALSE
  ),
  prec.u = list(    
    param = prior.prec.u,    
    initial = log(prec.u),    
    fixed = FALSE
  ),
  prec.x = list(    
    param = prior.prec.x,    
    initial = log(prec.x),    
    fixed = FALSE
  ),
  mean.x = list(    
    initial = 0,    
    fixed = TRUE
  )
),
) + z1 + z2 + z3 + z4
```

The call of the \texttt{inla} function includes the specifications for \(\tau_x\), the hyperparameter of the Gaussian regression model. These can be controlled via the \texttt{control.family} entry. The
prior distributions for the intercept $\beta_0$ and the fixed effects of the other covariates $z_1, \ldots, z_4$ are specified in the control.fixed entry.

```r
r <- inla(formula, data = data.frame(y, w, z1, z2, z3, z4, error.prec),
  family = "gaussian",
  control.family = list(
    hyper = list(
      prec = list(param = prior.prec.y,
                   initial = log(prec.y),
                   fixed = FALSE)
    )
  ),
  control.fixed = list(
    mean.intercept = prior.beta[1],
    prec.intercept = prior.beta[2],
    mean = prior.beta[1],
    prec = prior.beta[2]
  )
)
```

The last command improves the estimates of the posterior marginals for the hyperparameters of the model. The call is optional, but a slightly better agreement with the MCMC posterior marginals was found in this example. Thereby, $dz$ denotes the step length in a standardised scale used for the numerical integration of the hyperparameters. The default is 0.75 and smaller values lead to better approximations. The second tuning parameter, `diff.logdens`, is used to stop the numerical integration. It represents the maximally allowed difference between the log posterior marginal distribution of $\theta$ evaluated at the mode and at a potential support point $\theta_k$ (in standardised scale). The default is 7 and larger values give better approximations, as more density mass is taken into account. Note that both parameters have a significant influence on computational speed. To get a quick overview of the results, use the `summary` command.

```
summary(r)
```

1.2. Influence of systolic blood pressure on coronary heart disease (classical error)

Let all variables be defined as in Section 5.2 of the main text. The outcome is binary in $\{0, 1\}$, and assumed to be binomial distributed, i.e. $p(y_i) = \binom{N}{y_i} \pi_i^{y_i} (1 - \pi_i)^{1-y_i}$, with $N = 1$, $\pi_i = \exp(\eta_i)/(1 + \exp(\eta_i))$ and $\eta_i = \beta_0 + \beta_x x_i + \beta_z z_i$. We have a classical error structure, where the covariate $x = (x_1, \ldots, x_n)^\top$ is not directly observed, but two replicates, $w_1 = (w_{11}, \ldots, w_{1n})^\top$ and $w_2 = (w_{12}, \ldots, w_{n2})^\top$ are used as proxy, where $w_1 \sim N(x, \tau_x I)$ and $w_2 \sim N(x, \tau_x I)$. The prior distributions are elicited from expert/prior knowledge, see main text, and are defined as:
• $x \sim N(\alpha_0 + \alpha_2 z, \tau_x I)$.
• $\beta_0, \beta_x, \beta_z \sim N(0, \tau_\beta)$, with $\tau_\beta = 0.01$.
• $\alpha_0 \sim N(\mu_{\alpha_0}, \tau_{\alpha_0})$, with $\mu_{\alpha_0} = 0$ and $\tau_{\alpha_0} = 1$.
• $\alpha_2 \sim N(\mu_{\alpha_2}, \tau_{\alpha_2})$, with $\mu_{\alpha_2} = 0$ and $\tau_{\alpha_2} = 1$.
• $\tau_x \sim G(\alpha_x, \beta_x)$, with $\alpha_x = 10$ and $\beta_x = 1$.
• $\tau_u \sim G(\alpha_u, \beta_u)$, with $\alpha_u = 100$ and $\beta_u = 1$.

The object `data` consists of four columns:

\[
\begin{array}{cccc}
  y & w_1 & w_2 & z \\
\end{array}
\]

They contain (for $n = 641$):

• $y_1 \ldots y_n$: The binary response $y_i \in \{0, 1\}$.
• $w_{11} \ldots w_{n1}$: log(SBP - 50) at examination 1 (centered).
• $w_{12} \ldots w_{n2}$: log(SBP - 50) at examination 2 (centered).
• $z_1 \ldots z_n$: Smoking status $z_i \in \{0, 1\}$.

As described in the main text, the hierarchial model of this example is formulated in INLA as a joint model by applying the `copy` feature. The full model can be written as

\[
\begin{bmatrix}
  y_1 & \text{NA} & \text{NA} \\
  \vdots & \vdots & \vdots \\
  y_n & \text{NA} & \text{NA} \\
  \text{NA} & \text{NA} & \text{NA} \\
  \vdots & \vdots & \vdots \\
  \text{NA} & \text{NA} & w_{11} \\
  \vdots & \vdots & \vdots \\
  \text{NA} & \text{NA} & w_{21} \\
  \vdots & \vdots & \vdots \\
  \text{NA} & \text{NA} & w_{2n}
\end{bmatrix} =
\begin{bmatrix}
  1 \\
  \vdots \\
  n \\
  \text{NA} \\
  \vdots \\
  \text{NA} \\
  \text{NA} \\
  \vdots \\
  \text{NA} \\
  \text{NA}
\end{bmatrix}
\begin{bmatrix}
  \beta_0 \\
  \vdots \\
  \beta_n \\
  \alpha_0 \\
  \vdots \\
  \alpha_z \\
  \alpha_0 \\
  \alpha_z \\
  \alpha_z
\end{bmatrix}
\]

The reader is guided through the `r-inla` code for this joint model formulation in the following. The terms below the brackets indicate the names as they will be employed in the code. Start with the prior specification process, as described in the main text:

```r
data <- read.table("fram_data4supp.txt", header=T)
attach(data)
n <- nrow(data) #641
```
prior.beta <- c(0, 0.01)
prior.alpha0 <- c(0, 1)
prior.alphaz <- c(0, 1)
prior.prec.x <- c(10, 1)
prior.prec.u <- c(100, 1)

# initial values (mean of prior)
prec.u <- 100
prec.x <- 10

Second, the response matrix $Y$ and the data vectors are filled according to the naming of the above joint model equation:

\[
Y \leftarrow \text{matrix}(\text{NA}, 4*n, 3)
Y[1:n, 1] \leftarrow y
Y[n+(1:n), 2] \leftarrow \text{rep}(0, n)
Y[2*n+(1:n), 3] \leftarrow w1
Y[3*n+(1:n), 3] \leftarrow w2
\]

\[
\begin{align*}
\text{beta.0} & \leftarrow c(\text{rep}(1, n), \text{rep}(\text{NA}, n), \text{rep}(\text{NA}, n), \text{rep}(\text{NA}, n)) \\
\text{beta.x} & \leftarrow c(1:n, \text{rep}(\text{NA}, n), \text{rep}(\text{NA}, n), \text{rep}(\text{NA}, n)) \\
\text{idx.x} & \leftarrow c(\text{rep}(\text{NA}, n), 1:n, 1:n) \\
\text{weight.x} & \leftarrow c(\text{rep}(1, n), \text{rep}(-1, n), \text{rep}(1, n), \text{rep}(1,n)) \\
\text{beta.z} & \leftarrow c(z, \text{rep}(\text{NA}, n), \text{rep}(\text{NA}, n), \text{rep}(\text{NA}, n)) \\
\text{alpha.0} & \leftarrow c(\text{rep}(\text{NA}, n), \text{rep}(1, n), \text{rep}(\text{NA}, n), \text{rep}(\text{NA}, n)) \\
\text{alpha.z} & \leftarrow c(\text{rep}(\text{NA}, n), z, \text{rep}(\text{NA}, n), \text{rep}(\text{NA}, n)) \\
\text{Ntrials} & \leftarrow c(\text{rep}(1, n), \text{rep}(\text{NA}, n), \text{rep}(\text{NA}, n), \text{rep}(\text{NA}, n))
\end{align*}
\]

\[
data\text{-joint} \leftarrow \text{data.frame}(Y=Y, \\
\text{beta.0}=\text{beta.0}, \text{beta.x}=\text{beta.x}, \text{beta.z}=\text{beta.z}, \\
\text{idx.x}=\text{idx.x}, \text{weight.x}=\text{weight.x}, \\
\text{alpha0}=\text{alpha.0}, \text{alpha.z}=\text{alpha.z}, \\
\text{Ntrials}=\text{Ntrials})
\]

The next step contains the definition of the INLA formula. There are four fixed effects ($\beta_0, \beta_z, \alpha_0$ and $\alpha_z$) and two random effects. The latter are needed to encode that the values of $x$ in the exposure (7) and error model (8) are assigned the same values as in the regression model (6), where $\beta_x x$ represents a product of two unknown quantities. The two random effects terms are:

- $f(\text{beta.x},...)$: The \text{copy}="\text{idx.x}" \text{ call guarantees the assignment of identical values to } x \text{ in all components of the joint model. As discussed in the main text, } \beta_x \text{ is treated as a hyperparameter, namely the scaling parameter of the copied process } x'.$

- $f(\text{idx.x},...)$: \text{idx.x} \text{ contains the } x \text{ values, encoded as an i.i.d. Gaussian random effect, and weighted with } \text{weight.x} \text{ to ensure correct signs in the joint model. The } \text{values} \text{ option contains the vector of all values assumed by the covariate for which the effect is estimated. It must be a numeric vector, a vector of factors or NULL. The precision } \text{prec} \text{ of the random effect is fixed at } \tau = \exp(-15). \text{ This is necessary as the uncertainty in } x \text{ is already modelled in the second level (column 2 of } Y \text{) of the joint model, which defines the exposure component.}$
library(INLA)

formula <- Y ~ f(beta.x, copy = "idx.x", 
    hyper = list(beta = list(param = prior.beta, fixed = FALSE))) +
  f(idx.x, weight.x, model = "iid", values = 1:n,
    hyper = list(prec = list(initial = -15, fixed = TRUE))) +
  beta.0 - 1 + beta.z + alpha.0 + alpha.z

Since there is no common intercept in the joint model, it has to be explicitly removed using -1. The call of the inla function is given next. The following options need some explanation:

- **family**: There are three different likelihoods here, namely the binomial likelihood of the regression model and two Gaussian likelihoods, one for the exposure and one for the error model. They correspond to the different columns in the response matrix Y.

- **control.family**: Specification of the hyperparameters for the three likelihoods, in the same order as given in family. The binomial likelihood does not contain any hyperparameters, thus the respective list is empty. In the second and third likelihoods the hyperparameters $\tau_x$ and $\tau_u$ need to be specified, respectively.

- **control.fixed**: Prior specification for the fixed effects.

r <- inla(formula, Ntrials = Ntrials, data = data.joint,
           family = c("binomial", "gaussian", "gaussian"),
           control.family = list(
               list(hyper = list()),
               list(hyper = list(
                   prec = list(initial = log(prec.x),
                   param = prior.prec.x,
                   fixed = FALSE))),
               list(hyper = list(
                   prec = list(initial=log(prec.u),
                   param = prior.prec.u,
                   fixed = FALSE)))),
           control.fixed = list(
               mean = list(beta.0=prior.beta[1], beta.z=prior.beta[1],
                           alpha.z=prior.alphaz[1], alpha.0=prior.alpha0[1]),
               prec = list(beta.0=prior.beta[2], beta.z=prior.beta[2],
                           alpha.z=prior.alphaz[2], alpha.0=prior.alpha0[2]))
)

The last call (inla.hyperpar) is not required. It is used to improve the estimates of the posterior marginals for the hyperparameters using a finer grid in the numerical integration. In this application, only the marginal of $\tau_z$ changes slightly by this correction.

1.3.  **Seedling growth across different light conditions (Berkson error)**
Let all variables be defined as in Section 5.3 of the main text. Recall that the model is Poisson and contains the two covariates $x$ and $z$, and one independent, normal random
effect γ to account for potential overdispersion. The covariate x is not directly observed, but only a proxy w following a Berkson error model x | w ∼ N(w, τuI). The prior distributions are elicited from expert/prior knowledge, see main text, and are defined as:

- β0, βx, βz ∼ N(0, τβ), with τβ = 0.01,
- τγ ∼ G(αγ, βγ), with αγ = 1 and βγ = 0.005,
- τu ∼ G(αu, βu), with αu = 1.12 and βu = 0.0203.

**Analysis with the meb model**

The object `data` consists of three columns:

\[ y \quad w \quad z \]

They contain (for n = 60):

- \( y_1 \ldots y_n \): The number of new leaves.
- \( w_1 \ldots w_n \): log(\%light) for the target light intensities under dark, middle and light conditions (i.e., only three different values; centered).
- \( z_1 \ldots z_n \): Degree of defoliation (0\%, 25\%, 50\%, 75\%; centered).

Let us start again with prior specification process in accordance to the main text:

```r
# initial values (mean of prior)
prior.beta <- c(0,0.01)
prior.tau <- c(1,0.005)
prior.prec.u <- c(1.12,0.0203)
prec.tau <- 1/0.005
prec.u <- 1.12/0.0203

The fourth line contains a trick to ensure that the light values w from the s = 15 shadehouses are not completely identical, because in the new meb model only the unique values of w are used. Thus, if two or more elements of w are identical, then they refer to the same element in the covariate x, which is not desired here. Next, we define the meb model formula. The model contains two hyperparameters:

- **beta** corresponds to \( β_x \), the slope coefficient of the error-prone covariate x, with a Gaussian prior.
- **prec.u** is the error precision \( τ_u \) with gamma prior.
The prior settings are defined in the different entries of the list `hyper`. The option `fixed` specifies whether the corresponding quantity should be estimated or fixed at the `initial` value. The field `param` captures the prior parameters of the corresponding prior distribution. A Gaussian prior distribution is the default for `beta`, while a gamma distribution is used for `prec.u` (again defined as log-gamma distribution for the log-precision).

The model contains as additional fixed effect the degree of defoliation `z`, plus an additional i.i.d. random effects term per individual to account for unspecified heterogeneity, specified in `f(individual,...)`, which extends the GLM to a GLMM:

```r
library(INLA)
formula <- y ~ f(w, model="meb", hyper = list(
  beta = list(
    param = prior.beta,
    fixed = FALSE
  ),
  prec.u = list(
    param = prior.prec.u,
    initial = log(prec.u),
    fixed = FALSE
  )
)) +
  z +
  f(individual, model = "iid", values = 1:n, hyper = list(prec = list(
    initial = log(prec.tau),
    param = prior.tau
  )
  )
)

r <- inla(formula, data = data.frame(y, w, z, individual),
  family = "poisson",
  control.fixed = list(
    mean.intercept = prior.beta[1],
    prec.intercept = prior.beta[2],
    mean = prior.beta[1],
    prec = prior.beta[2])
)

r <- inla.hyperpar(r)
summary(r)
```

**Analysis with the `copy` feature**

As described in the main text, as an alternative to the use of the new `meb` model, the same results can be obtained by employing the `copy` feature in INLA. The approach is similar to...
the one taken in Section 1.2. Recall though that in case of Berkson measurement error, the use of the copy feature does not add to the generality of the model and is presented here only for completeness.

The object data now contains an additional fourth column:

\[ y \quad w \quad z \quad sh \]

Column \( sh \) contains the values \( sh_1, \ldots, sh_n \), where \( sh_i \) is the index of the shadehouse of seedling \( i \). Note that the \( n = 60 \) seedlings are distributed over \( s = 15 \) shadehouses \( (1 \leq sh_i \leq 15) \), whereas always five shadehouses belong to the same light condition (dark, middle, light). There are thus 15 different correct light intensities \( (x, \text{one value per shadehouse}) \), but only 3 different target light intensities \( (w, \text{one value per light condition}) \). As the error model in this example is Berkson, the joint model simplifies to two equations and the response matrix has only two columns. The model can be represented as

\[
\begin{bmatrix}
y_1 \\
\vdots \\
y_n \\
w_1 \\
w_s \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
NA \\
The definition of the INLA formula is almost analogous to the one in Section 1.2. The main difference is the additional i.i.d. random effects term per individual $\gamma_{ijk}$, specified in $f(\gamma,...)$, which extends the GLM to a GLMM:

```r
library(INLA)
formula <- Y ~ beta.0 - 1 +
    f(beta.x, copy = "idx.x",
        hyper = list(beta = list(param = prior.beta, fixed = FALSE))) +
    f(idx.x, weight.x, model = "iid", values = 1:s,
        hyper = list(prec = list(initial = -15, fixed = TRUE))) +
    beta.z +
    f(gamma, model = "iid", values = 1:n,
        hyper = list(prec = list(initial = log(prec.tau), param = prior.tau)))
```

As in Section 1.2 we have to explicitly remove the common intercept using -1. The call of the INLA function is as well in analogy to Section 1.2, but there are only two likelihoods involved here: the Poisson likelihood for the regression model and the Gaussian likelihood for the error model. The former has no additional hyperparameters, while in the latter the error precision $\tau_u$ needs specification.

```r
r <- inla(formula, data = data.joint,
    family = c("poisson", "gaussian"),
    control.family = list(
        list(hyper = list()),
        list(hyper = list(
            prec = list(
                initial=log(prec.u),
                param = prior.prec.u,
                fixed = FALSE)))))

control.fixed = list(
    mean.intercept = prior.beta[1],
    prec.intercept = prior.beta[2],
    mean = prior.beta[1],
    prec = prior.beta[2])
)
```

```r
r <- inla.hyperpar(r)
summary(r)
```
2. Supplementary Figures

2.1. The effect of measurement error

Fig. 1. Least squares estimates in the linear model with ME. Left: Classical ME. Two effects can be seen: 1) The absolute value of the covariate estimate is biased (attenuated); 2) The variability around the regression line in the data with ME (black circles) is much larger than in the case of the truly observed data (grey squares). Right: Berkson ME. The absolute value of the covariate estimate is unbiased in the linear model, while the variability around the regression line is larger for the data with ME.

3. Supplements to Sections 5.1 – 5.3 in the main text

3.1. Full conditionals for the MCMC sampler of Section 5.2

Let all variables be defined as in Section 5.2 of the main text, see also the beginning of Section 1.2 in this Supplementary Material for a compact review. The full conditionals for the unknowns in the regression model are given in the following. For $\beta = (\beta_0, \beta_x, \beta_z)^T$ we have

$$\beta | \text{rest} \propto \pi(y | x, z) \cdot \pi(\beta)$$

$$\propto \exp\left(\sum_{i=1}^{n} y_i \eta_i - \sum_{i=1}^{n} \log(1 + e^{\eta_i}) - \frac{\tau_\beta}{2} \beta^T \beta \right).$$
The $\alpha = (\alpha_0, \alpha_z)^\top$ coefficients can be sampled from a Gaussian distribution. Let $D$ be the matrix with rows $D_i^\top := (1 \ z_i^\top)$, $R = \begin{pmatrix} \tau_{\alpha_0} & 0 \\ 0 & \tau_{\alpha_z} \end{pmatrix}$, and $\mu = (\mu_{\alpha_0}, \mu_{\alpha_z})^\top$. Then

$$\alpha \mid \text{rest} \propto \pi(x \mid \text{rest}) \cdot \pi(\alpha)$$
$$\propto \exp \left( -\frac{\tau_x}{2} (x - D \alpha)^\top (x - D \alpha) - \frac{1}{2} (\alpha - \mu)^\top R (\alpha - \mu) \right)$$
$$\sim \mathcal{N}\left( (\tau_x D^\top D + R)^{-1} \left( \tau_x D^\top x + R \cdot \mu \right), \tau_x D^\top D + R \right),$$

where the second argument in the last expression again denotes the precision matrix. To sample from the distribution of the latent variable $x$, full conditionals for $x_i$ are needed:

$$x_i \mid \text{rest} \propto \pi(y_i \mid x_i, z_i) \cdot \pi(w_{1i} \mid x_i) \cdot \pi(w_{2i} \mid x_i) \cdot \pi(x_i \mid z_i)$$
$$\propto \exp \left( y_i \eta_i - \log(1 + e^{\eta_i}) - \frac{\tau_u}{2} \left((w_{1i} - x_i)^2 + (w_{2i} - x_i)^2\right) - \frac{\tau_x}{2} (x_i - \alpha_0 - \alpha_z z_i)^2 \right).$$

Finally, the precisions can be sampled from gamma distributions

$$\tau_x \mid \text{rest} \propto \pi(x \mid z) \cdot \pi(\tau_x)$$
$$\sim \mathcal{G}\left( a_x + \frac{n}{2}, b_x + \frac{1}{2} (x - D \alpha)^\top (x - D \alpha) \right),$$

and

$$\tau_u \mid \text{rest} \propto \pi(w \mid x) \cdot \pi(\tau_u)$$
$$\sim \mathcal{G}\left( a_u + n, b_u + \frac{1}{2} (w_1 - x)^\top (w_1 - x) + \frac{1}{2} (w_2 - x)^\top (w_2 - x) \right).$$
3.2. MCMC and INLA posterior marginals

![Graph showing MCMC and INLA posterior marginals for ibex data](image)

**Fig. 2.** Comparison of the MCMC samples (histograms) with the INLA posterior marginal densities (lines) for the ibex data (Section 5.1). The $z$-covariates are those treated as error-free. They are the length of the time series ($z_1$), average precipitation in summer ($z_2$), average precipitation in winter ($z_3$) and the interaction $z_4 = z_1 z_2$. 
Fig. 3. Comparison of the MCMC samples (histograms) with the INLA posterior marginals (lines) for the Framingham data (Section 5.2).

Fig. 4. Comparison of the MCMC samples (histograms) with the INLA posterior marginals (lines) for the seedling growth data (Section 5.3).