Extending the toolbox –
inlabru – what can it do?

Janine Illian

CREEM
Centre for Research into Ecological and Environmental Modelling,
University of St Andrews, Scotland, UK

November 9, 2018

joint work with: Fabian Bachl, Finn Lindgren, David Borchers

model family:
latent Gaussian model with (hyper)parameters $\boldsymbol{\theta}$, latent Gaussian variables $\boldsymbol{x}$, and observations $\boldsymbol{y}$:

model family:
latent Gaussian model with (hyper)parameters $\boldsymbol{\theta}$, latent Gaussian variables $\boldsymbol{x}$, and observations $\boldsymbol{y}$:

$$\begin{aligned}
\boldsymbol{\theta} &\sim p(\boldsymbol{\theta}), \\
\boldsymbol{x}|\boldsymbol{\theta} &\sim p(\boldsymbol{x}|\boldsymbol{\theta}) \sim \mathsf{N}(\mathbf{0}, \boldsymbol{Q}_x^{-1}), \\
\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta} &\sim p(\boldsymbol{y}|\boldsymbol{\eta}(\boldsymbol{x}),\boldsymbol{\theta})
\end{aligned}$$

model family:
latent Gaussian model with (hyper)parameters $\boldsymbol{\theta}$, latent Gaussian variables $\boldsymbol{x}$, and observations $\boldsymbol{y}$:

$$\begin{aligned}\boldsymbol{\theta} &\sim p(\boldsymbol{\theta}),\\ \boldsymbol{x}|\boldsymbol{\theta} &\sim p(\boldsymbol{x}|\boldsymbol{\theta}) \sim \mathsf{N}(\boldsymbol{0}, \boldsymbol{Q}_x^{-1}),\\ \boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta} &\sim p(\boldsymbol{y}|\boldsymbol{\eta}(\boldsymbol{x}),\boldsymbol{\theta})\end{aligned}$$

the **linear predictor** $\boldsymbol{\eta}(\cdot)$ controls a location parameter of the likelihood:

$$g[\mathsf{E}(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta})] = \boldsymbol{\eta}(\boldsymbol{x})$$

INLA – optimisation

INLA nests optimisation over $\boldsymbol{x}$ within optimisation over $\boldsymbol{\theta}$, followed by numerical integration over $\boldsymbol{\theta}$

INLA – optimisation

INLA nests optimisation over $\boldsymbol{x}$ within optimisation over $\boldsymbol{\theta}$, followed by numerical integration over $\boldsymbol{\theta}$

- conditional posterior mode of $p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{y})$ is

$$\boldsymbol{x}^*(\boldsymbol{\theta}) = \underset{\boldsymbol{x}}{\operatorname{argmax}} \{p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\theta})\}$$

INLA – optimisation

INLA nests optimisation over $\boldsymbol{x}$ within optimisation over $\boldsymbol{\theta}$, followed by numerical integration over $\boldsymbol{\theta}$

- conditional posterior mode of $p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{y})$ is

$$\boldsymbol{x}^*(\boldsymbol{\theta}) = \underset{\boldsymbol{x}}{\operatorname{argmax}} \{p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\theta})\}$$

- parameter posterior density:

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto \left. \frac{p(\boldsymbol{\theta})p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\theta})}{p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{y})} \right|_{\boldsymbol{x}=\boldsymbol{x}^*(\boldsymbol{\theta})}$$

INLA – optimisation

INLA nests optimisation over $\boldsymbol{x}$ within optimisation over $\boldsymbol{\theta}$, followed by numerical integration over $\boldsymbol{\theta}$

- conditional posterior mode of $p(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{y})$ is

$$\boldsymbol{x}^*(\boldsymbol{\theta}) = \underset{\boldsymbol{x}}{\operatorname{argmax}} \{p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta})\}$$

- parameter posterior density:

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto \left. \frac{p(\boldsymbol{\theta})p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta})}{p(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{y})} \right|_{\boldsymbol{x}=\boldsymbol{x}^*(\boldsymbol{\theta})}$$

Laplace approximation:

Taylor series matching of $p(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{y})$ to a multivariate Gaussian density $\widetilde{p}(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{y})$

INLA – optimisation

INLA nests optimisation over $\boldsymbol{x}$ within optimisation over $\boldsymbol{\theta}$, followed by numerical integration over $\boldsymbol{\theta}$

INLA – optimisation

INLA nests optimisation over $\boldsymbol{x}$ within optimisation over $\boldsymbol{\theta}$, followed by numerical integration over $\boldsymbol{\theta}$

- conditional posterior mode of $p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{y})$ is

$$\boldsymbol{x}^*(\boldsymbol{\theta}) = \underset{\boldsymbol{x}}{\operatorname{argmax}} \{p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\theta})\}$$

INLA – optimisation

INLA nests optimisation over $\boldsymbol{x}$ within optimisation over $\boldsymbol{\theta}$, followed by numerical integration over $\boldsymbol{\theta}$

- conditional posterior mode of $p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{y})$ is

$$\boldsymbol{x}^*(\boldsymbol{\theta}) = \underset{\boldsymbol{x}}{\operatorname{argmax}} \{p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\theta})\}$$

- parameter posterior density:

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto \left. \frac{p(\boldsymbol{\theta})p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\theta})}{p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{y})} \right|_{\boldsymbol{x}=\boldsymbol{x}^*(\boldsymbol{\theta})}$$

INLA – optimisation

INLA nests optimisation over $\boldsymbol{x}$ within optimisation over $\boldsymbol{\theta}$, followed by numerical integration over $\boldsymbol{\theta}$

- conditional posterior mode of $p(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{y})$ is

$$\boldsymbol{x}^*(\boldsymbol{\theta}) = \underset{\boldsymbol{x}}{\operatorname{argmax}} \{p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta})\}$$

- parameter posterior density:

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto \left. \frac{p(\boldsymbol{\theta})p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta})}{p(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{y})} \right|_{\boldsymbol{x}=\boldsymbol{x}^*(\boldsymbol{\theta})}$$

Laplace approximation:

Taylor series matching of $p(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{y})$ to a multivariate Gaussian density $\widetilde{p}(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{y})$

What INLA cannot do...

INLA assumes that entire pattern has been observed

What INLA cannot do...

INLA assumes that entire pattern has been observed
$\Rightarrow$ not realistic in many ecological applications

What INLA cannot do...

INLA assumes that entire pattern has been observed
$\Rightarrow$ not realistic in many ecological applications
interpret observed data as **thinned point pattern**

What INLA cannot do...

INLA assumes that entire pattern has been observed
$\Rightarrow$ not realistic in many ecological applications
interpret observed data as **thinned point pattern**

`inlabru`

What INLA cannot do...

INLA assumes that entire pattern has been observed
$\Rightarrow$ not realistic in many ecological applications
interpret observed data as **thinned point pattern**

`inlabru`

$\Rightarrow$ *wrapper around* `R-INLA` *+ extra functionality*

What INLA cannot do...

INLA assumes that entire pattern has been observed
$\Rightarrow$ not realistic in many ecological applications
interpret observed data as **thinned point pattern**

inlabru

$\Rightarrow$ *wrapper around* `R-INLA` *+ extra functionality*

- *R-library that makes INLA more accessible*
- *in particular: easier coding for spatial point processes*

What INLA cannot do...

INLA assumes that entire pattern has been observed
$\Rightarrow$ not realistic in many ecological applications
interpret observed data as **thinned point pattern**

inlabru

$\Rightarrow$ *wrapper around `R-INLA` + extra functionality*

- *R-library that makes INLA more accessible*
- *in particular: easier coding for spatial point processes*
- *takes observation process into account (e.g. for distance sampling, plot sampling)*

What INLA cannot do...

INLA assumes that entire pattern has been observed
$\Rightarrow$ not realistic in many ecological applications
interpret observed data as **thinned point pattern**

inlabru

$\Rightarrow$ *wrapper around* `R-INLA` + *extra functionality*

- *R-library that makes INLA more accessible*
- *in particular: easier coding for spatial point processes*
- *takes observation process into account (e.g. for distance sampling, plot sampling)*

in essence: integrated fitting of observation process (detection) and ecological process of interest

What INLA cannot do...

INLA assumes that entire pattern has been observed
$\Rightarrow$ not realistic in many ecological applications
interpret observed data as **thinned point pattern**

`inlabru`

$\Rightarrow$ *wrapper around* `R-INLA` *+ extra functionality*

- *R-library that makes INLA more accessible*
- *in particular: easier coding for spatial point processes*
- *takes observation process into account (e.g. for distance sampling, plot sampling)*

in essence: integrated fitting of observation process (detection) and ecological process of interest

issue: needs to be linear in the latent variables

INLA – non-linear predictors

what if $\boldsymbol{\eta}(\boldsymbol{x})$ is non-linear (and parametric)?

INLA – non-linear predictors

what if $\boldsymbol{\eta}(\boldsymbol{x})$ is non-linear (and parametric)?

example: non-linear detection function, detection depends on size

INLA – non-linear predictors

what if $\boldsymbol{\eta}(\boldsymbol{x})$ is non-linear (and parametric)?

example: non-linear detection function, detection depends on size

$\Rightarrow$ use Taylor series to linearise:

$$\boldsymbol{\eta}(\boldsymbol{x}) = \boldsymbol{\eta}_k + \boldsymbol{A}_k(\boldsymbol{x} - \boldsymbol{x}_k) + \mathcal{O}(\|\boldsymbol{x} - \boldsymbol{x}_k\|^2),$$

where $\boldsymbol{A}_k$ is the Jacobian

INLA – non-linear predictors

what if $\boldsymbol{\eta}(\boldsymbol{x})$ is non-linear (and parametric)?

example: non-linear detection function, detection depends on size

$\Rightarrow$ use Taylor series to linearise:

$$\boldsymbol{\eta}(\boldsymbol{x}) = \boldsymbol{\eta}_k + \boldsymbol{A}_k(\boldsymbol{x} - \boldsymbol{x}_k) + \mathcal{O}(\|\boldsymbol{x} - \boldsymbol{x}_k\|^2),$$

where $\boldsymbol{A}_k$ is the Jacobian

Iterate over linearisation points, $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, \dots\}$,

1. Start at some $\boldsymbol{x}_0$, and repeat the following for $k = 0, 1, 2, \dots$
2. Let $\overline{\boldsymbol{\eta}}_k(\boldsymbol{x}) = \boldsymbol{\eta}_k + \boldsymbol{A}_k(\boldsymbol{x} - \boldsymbol{x}_k)$

INLA – non-linear predictors

what if $\boldsymbol{\eta}(\boldsymbol{x})$ is non-linear (and parametric)?

example: non-linear detection function, detection depends on size

$\Rightarrow$ use Taylor series to linearise:

$$\boldsymbol{\eta}(\boldsymbol{x}) = \boldsymbol{\eta}_k + \boldsymbol{A}_k(\boldsymbol{x} - \boldsymbol{x}_k) + \mathcal{O}(\|\boldsymbol{x} - \boldsymbol{x}_k\|^2),$$

where $\boldsymbol{A}_k$ is the Jacobian

Iterate over linearisation points, $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, \dots\}$,

1. Start at some $\boldsymbol{x}_0$, and repeat the following for $k = 0, 1, 2, \dots$
2. Let $\overline{\boldsymbol{\eta}}_k(\boldsymbol{x}) = \boldsymbol{\eta}_k + \boldsymbol{A}_k(\boldsymbol{x} - \boldsymbol{x}_k)$
3. Run INLA on the linearised problem, generating the mode $(\boldsymbol{\theta}_k^*, \boldsymbol{x}_k^*) = (\boldsymbol{\theta}_k^*, \boldsymbol{x}_k^*(\boldsymbol{\theta}_k^*))$
4. Let $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k^*$

in practice... ecology

interested in individuals (in space and time)

in practice... ecology

interested in individuals (in space and time)

in practice:

- need to gain information on individuals – given practical limitations

in practice... ecology

interested in individuals (in space and time)

in practice:

- need to gain information on individuals – given practical limitations

$\Rightarrow$ specific observation process

in practice... ecology

interested in individuals (in space and time)

in practice:

- need to gain information on individuals – given practical limitations

$\Rightarrow$ specific observation process

$\Rightarrow$ specific data structure

in practice... ecology

interested in individuals (in space and time)

in practice:

- need to gain information on individuals – given practical limitations

⇒ specific observation process

⇒ specific data structure

⇒ specific statistical methodology

in practice... ecology

interested in individuals (in space and time)

in practice:

- need to gain information on individuals – given practical limitations

$\Rightarrow$ specific observation process

$\Rightarrow$ specific data structure

$\Rightarrow$ specific statistical methodology

here:

- "think" in terms of the underlying structure, the point process
- observation process is operation on the underlying data structure

$\Rightarrow$ more general methodology and software

in practice... ecology

in practice... ecology

for distance sampling

- convenient integrated model fitting

in practice... ecology

for distance sampling

- convenient integrated model fitting
- flexible detection functions
- detection may depend on marks (size etc.)

in practice... ecology

for distance sampling

- convenient integrated model fitting
- flexible detection functions
- detection may depend on marks (size etc.)

BUT: what about if you are not interested in distance sampling...?

in practice... ecology

for distance sampling

- convenient integrated model fitting
- flexible detection functions
- detection may depend on marks (size etc.)

BUT: what about if you are not interested in distance sampling...?

- user-friendly software *inlabru* for complex models

in practice... ecology

for distance sampling

- convenient integrated model fitting
- flexible detection functions
- detection may depend on marks (size etc.)

BUT: what about if you are not interested in distance sampling...?

- user-friendly software *inlabru* for complex models
- other observation processes may be seen as different types of "thinnings"

in practice... other users

in practice... other users

$\Rightarrow$ unified approach, general software

- can fit general spatial models (no thinning) elegantly

in practice... other users

$\Rightarrow$ unified approach, general software

- can fit general spatial models (no thinning) elegantly
- dropping linearity assumption – applicable in many contexts

in practice... other users

$\Rightarrow$ unified approach, general software

- can fit general spatial models (no thinning) elegantly
- dropping linearity assumption – applicable in many contexts
- can interpret (univariate) function as one-dimensional LGCP

in practice... other users

- ⇒ unified approach, general software
  - can fit general spatial models (no thinning) elegantly
  - dropping linearity assumption – applicable in many contexts
  - can interpret (univariate) function as one-dimensional LGCP
- ⇒ use *inlabru* for function estimation (detection function, pdfs, K-functions...)

in practice... other users

$\Rightarrow$ unified approach, general software

- can fit general spatial models (no thinning) elegantly
- dropping linearity assumption – applicable in many contexts
- can interpret (univariate) function as one-dimensional LGCP

$\Rightarrow$ use *inlabru* for function estimation (detection function, pdfs, K-functions...)

in essence:
very general and flexible methodology and associated software