A new Bayesian approach to modelling of short and long term earthquake forecasting using inlabru

Mark Naylor

Researchers: Kirsty Bayliss, Francesco Serafini, Farnaz Kamranzad

Faculty: Finn Lindgren, Ian Main

Inla + inlabru

https://www.r-inla.org:

- *The integrated nested Laplace approximation (INLA) is a method for approximate Bayesian inference.*
  - *An alternative to other methods such as MCMC because of its speed and ease of use via the R-INLA package*

https://inlabru-org.github.io/inlabru/:

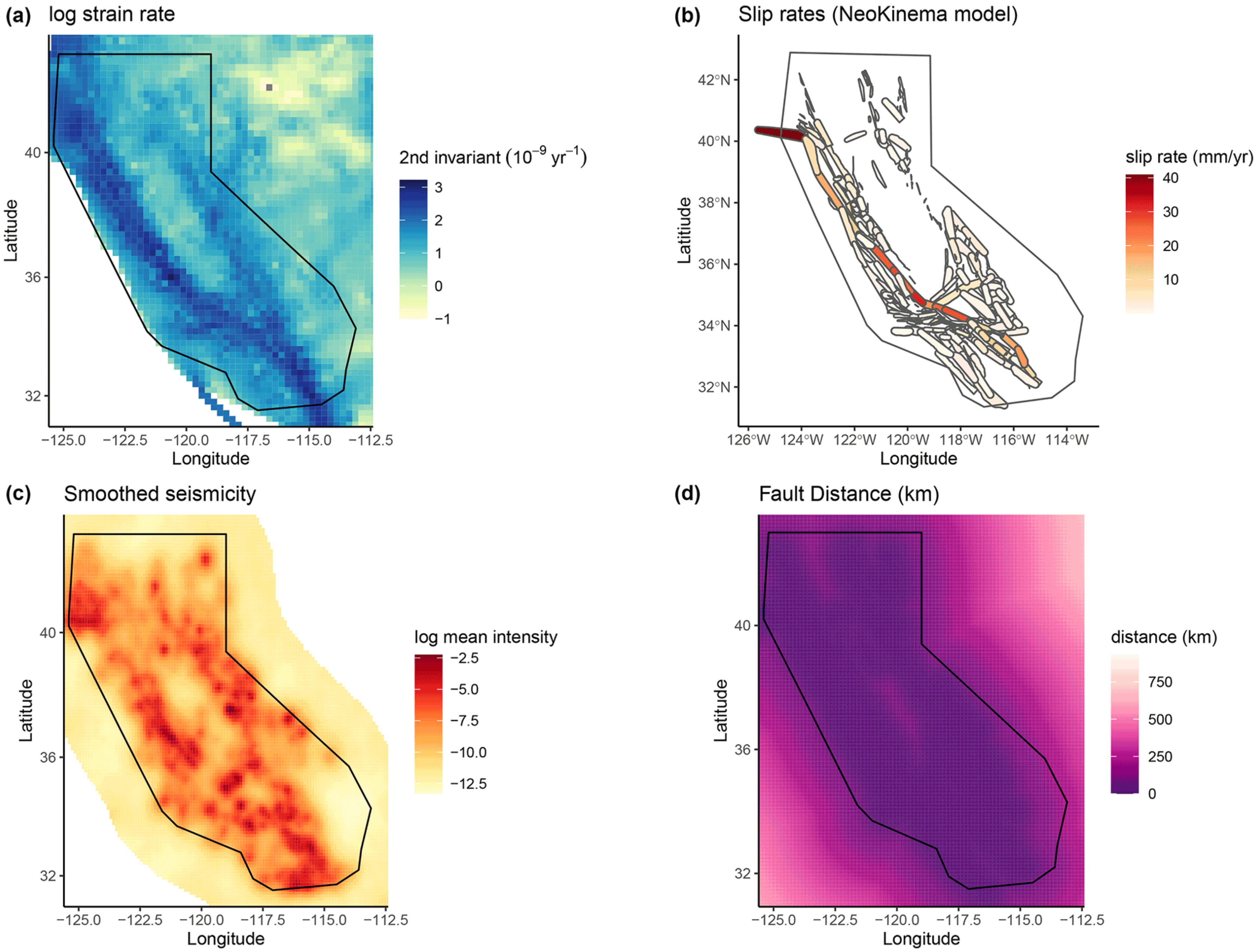
- *The goal of inlabru is to facilitate spatial modeling using integrated nested Laplace approximation via the R-INLA package.*

1. Time independent spatial models: California (Log-Gaussian Cox process)

Kirsty Bayliss

Engineering and Physical Sciences Research Council

1. Bayliss, Naylor, Illian & Main (2020). Data-driven optimization of seismicity models using diverse data sets: Generation, evaluation, and ranking using inlabru. *Journal of Geophysical Research: Solid Earth*, *125*, e2020JB020226. https:// doi.org/10.1029/2020JB020226
2. Bayliss, Naylor, Kamranzad, and Main (2022) Pseudo-prospective testing of 5-year earthquake forecasts for California using inlabru, *Nat. Hazards Earth Syst. Sci.*, 22, 3231–3246, 2022 https://doi.org/10.5194/nhess-22-3231-2022



Intensity is function of a linear predictor

$$\lambda(s) = \exp(-\eta(s))$$

Linear predictor combines:

- Intercept
- Weighted sum of spatial covariates
- Gaussian Markov Random Field

$$\eta(s) = \beta_0 + \sum_{m=1}^{M} \beta_m x_m(s) + \zeta(s)$$

Gaussian Markov Random Field describes correlation in the spatial residuals

- Unmodelled clustering
- Location uncertainty
- Unseen trigger

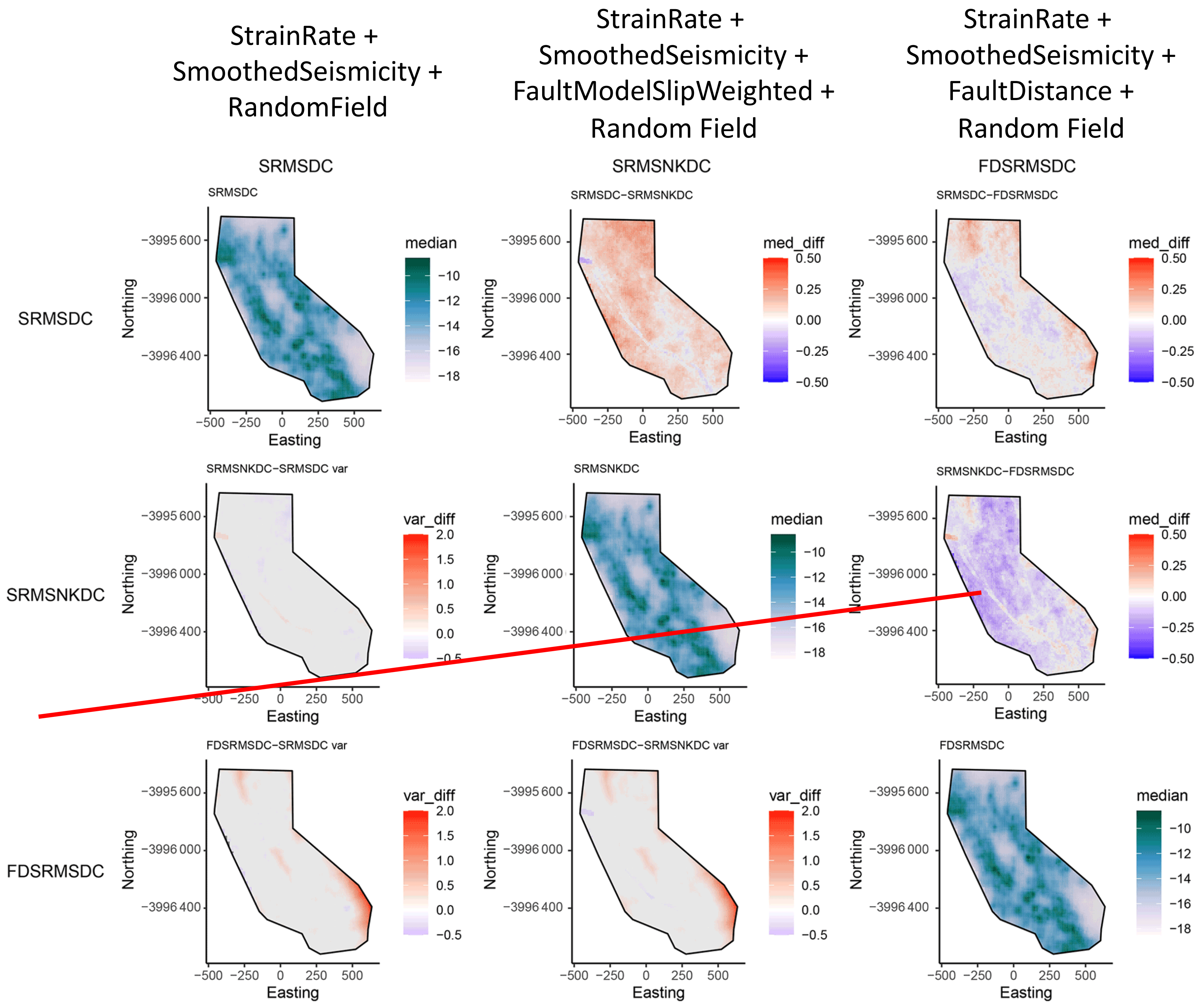
Catalogue input from the UCERF3

- California earthquakes
- 20 years (1985–2005)
- magnitude ≥4.95

- *Declustered*

Forecasts of rates for M>4.95

Median difference shows influence of fault network



Catalogue input from the UCERF3

- California earthquakes
- 20 years (1985–2005)
- magnitude ≥4.95

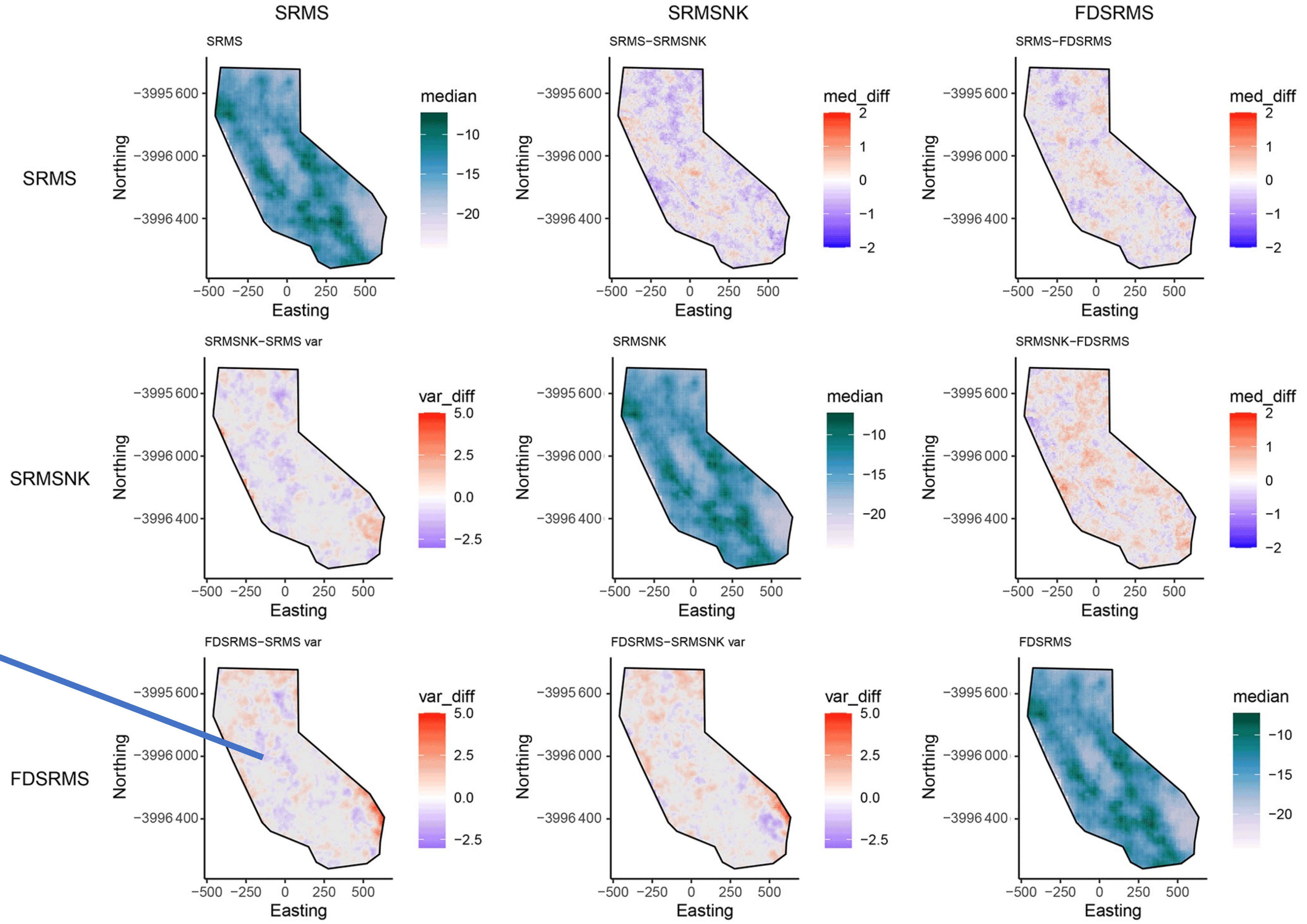
- *Full Catalogue*
    - *NOT declustered*

Forecasts of rates for M>4.95

StrainRate + SmoothedSeismicity + RandomField

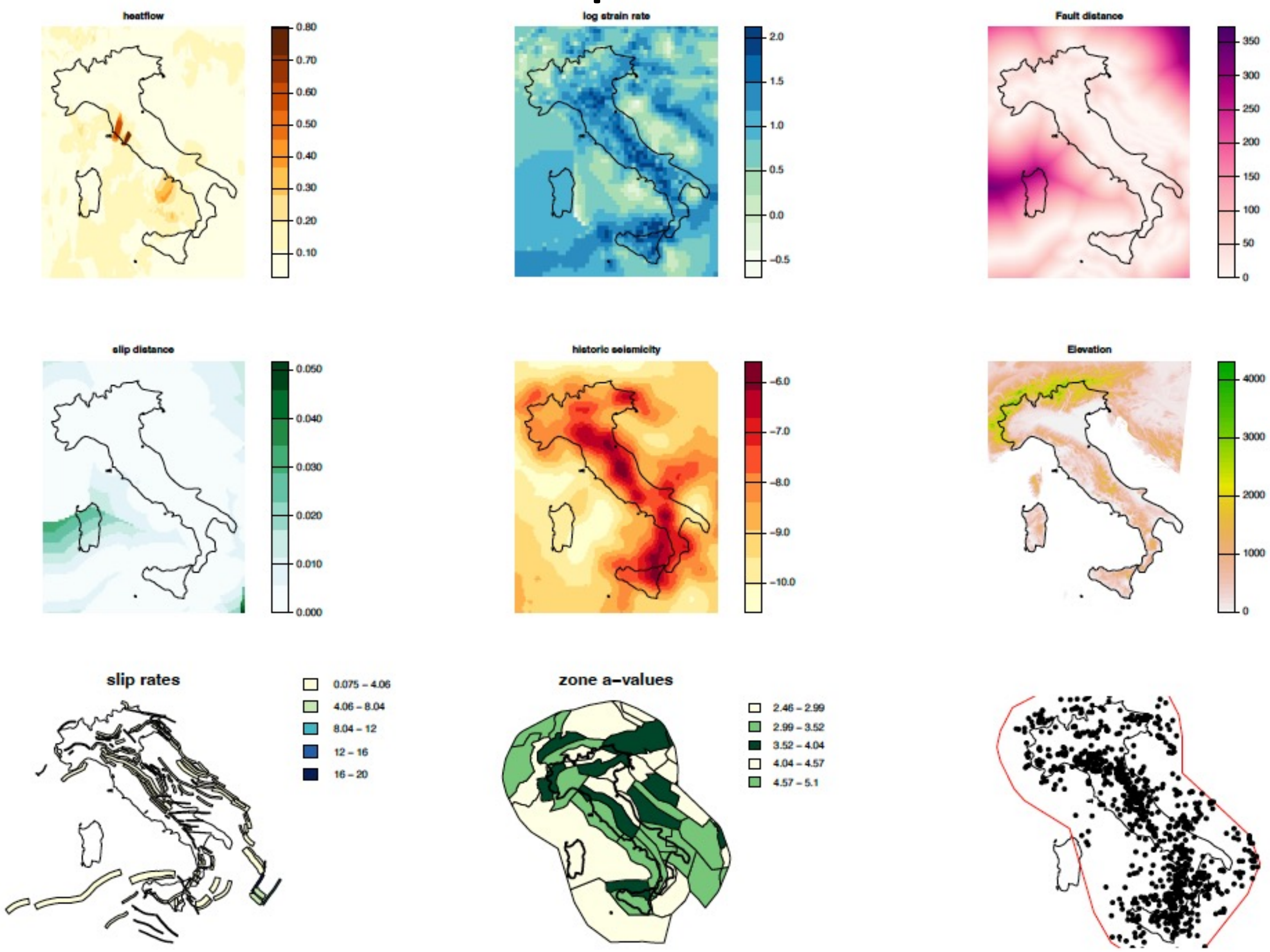
StrainRate + SmoothedSeismicity + FaultModelSlipWeighted + Random Field

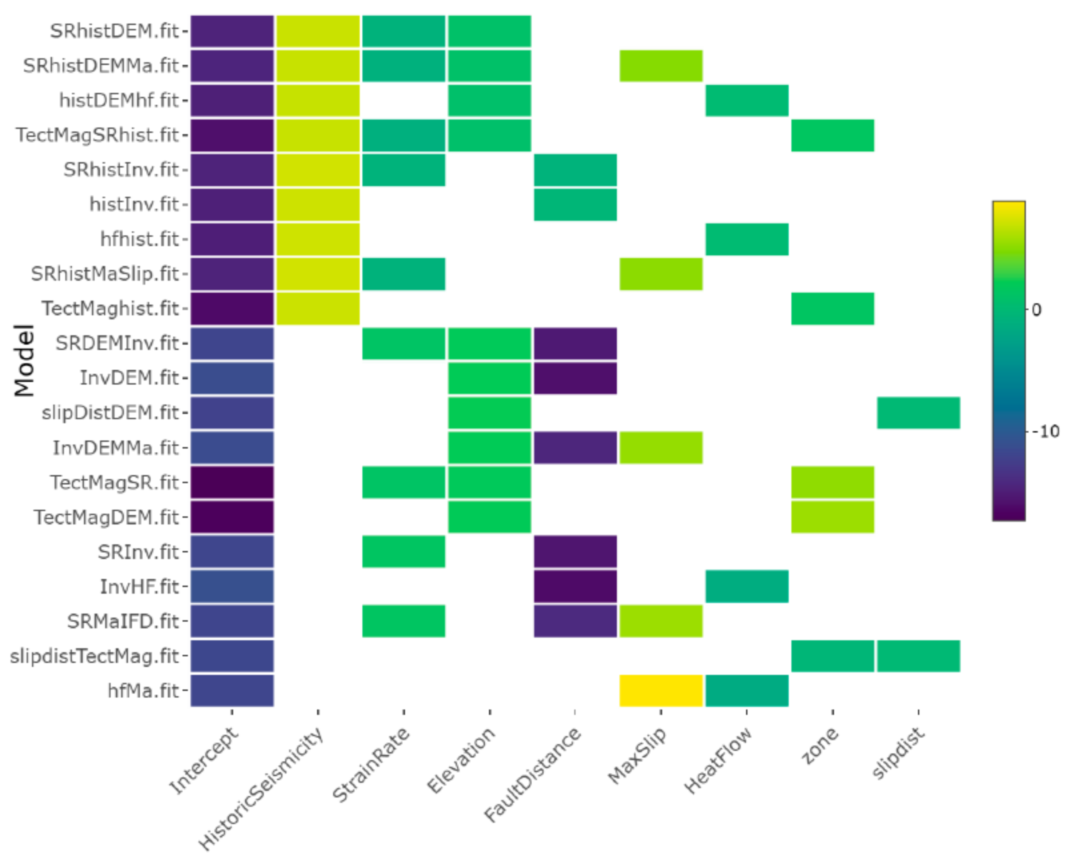
StrainRate + SmoothedSeismicity + FaultDistance + Random Field



Variance difference shows how random field is trying to account for spatial clustering
(some latent structure)

Time independent – whole Italy





3. Bayliss, Naylor and Main (In Prep) Developing a background seismicity model for operational earthquake forecasting in Italy

But this is not the best solution because…

… I’m not a fan of declustering.

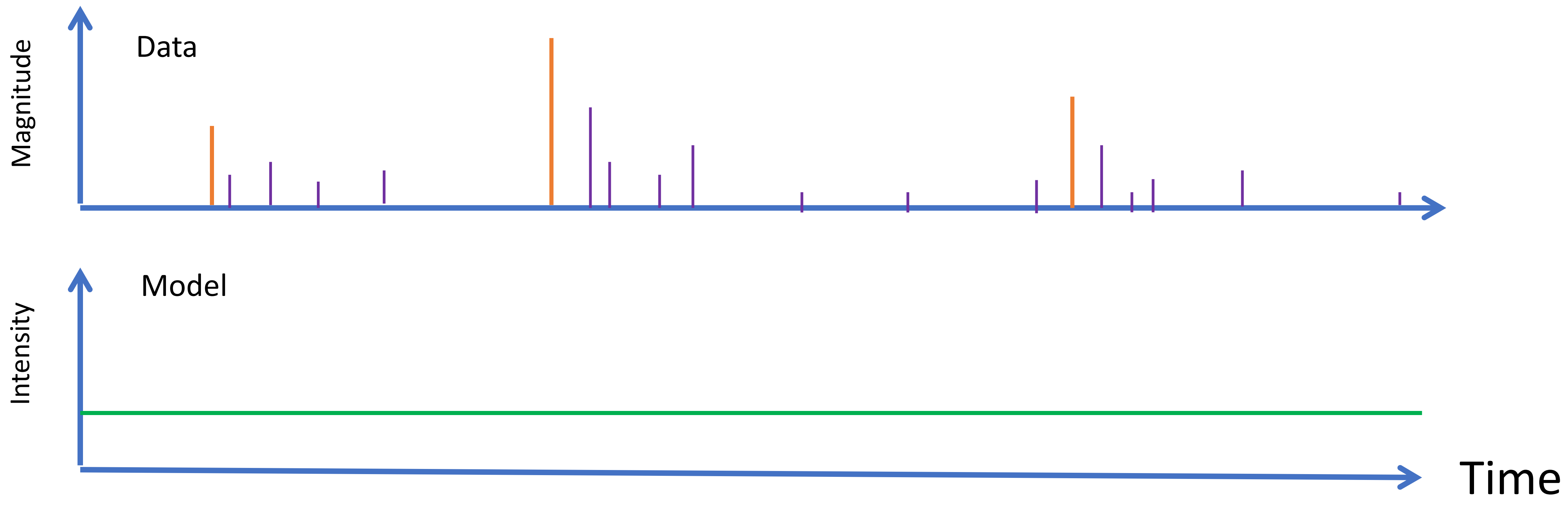
I’d rather have a background map with an estimate of uncertainty in the clustering process.

So…

2. Time dependent ETAS like model (Amatrice)

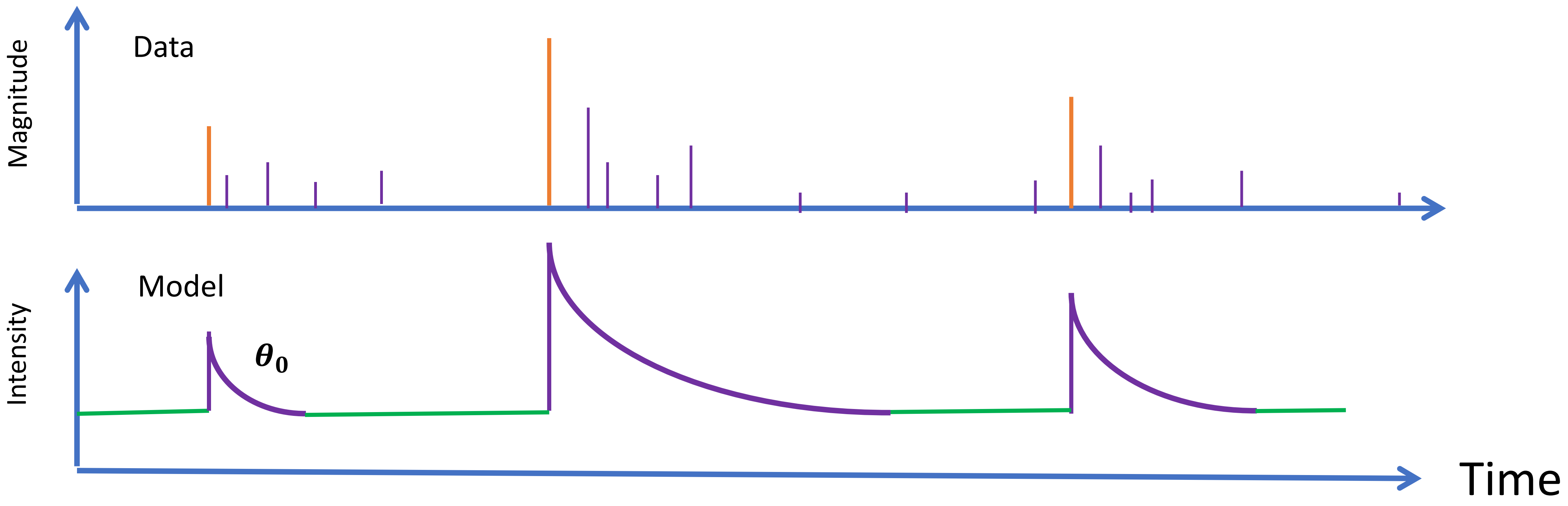
Francesco Serafini

- Inlabru only had functionality for Log-Gaussian Cox process
- We extended to Hawkes processes (See Francesco's poster)
  - Self exciting point process
  - Trick: lgcp with numerical integration for a sample of ETAS parameters



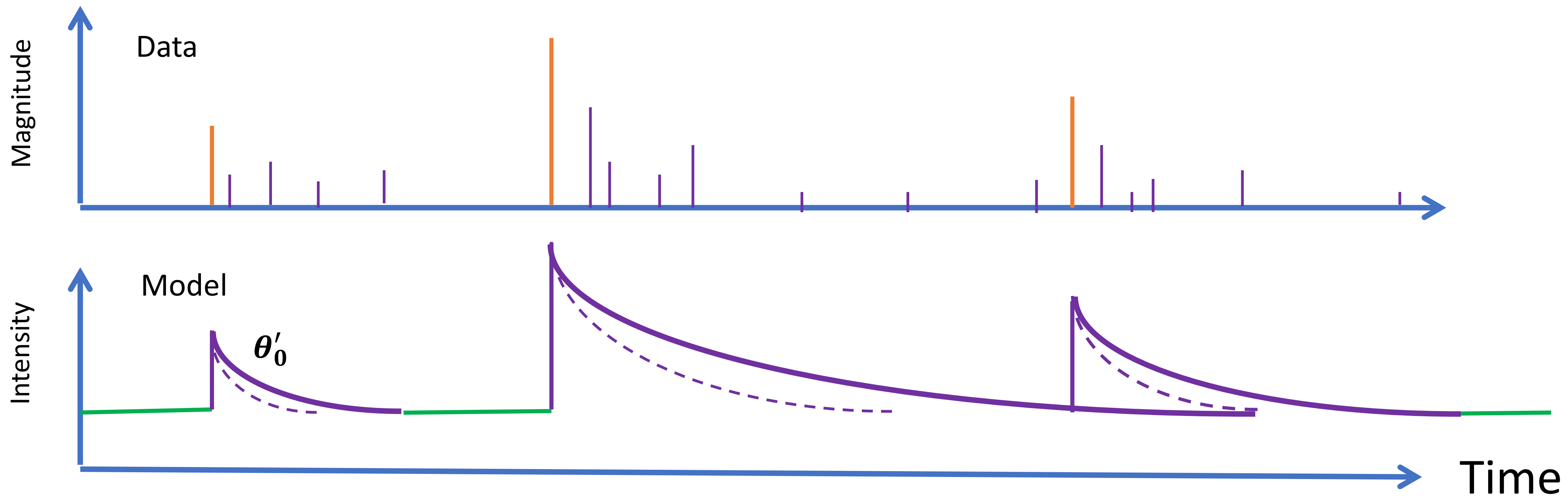
2. Time dependent ETAS like model (Amatrice)

- Inlabru only has functionality for log-Gaussian Cox process
- We needed to extend to Hawkes processes (See Francesco’s poster)
  - Self exciting point process
  - Trick: lgcp with numerical integration for a sample of ETAS parameters



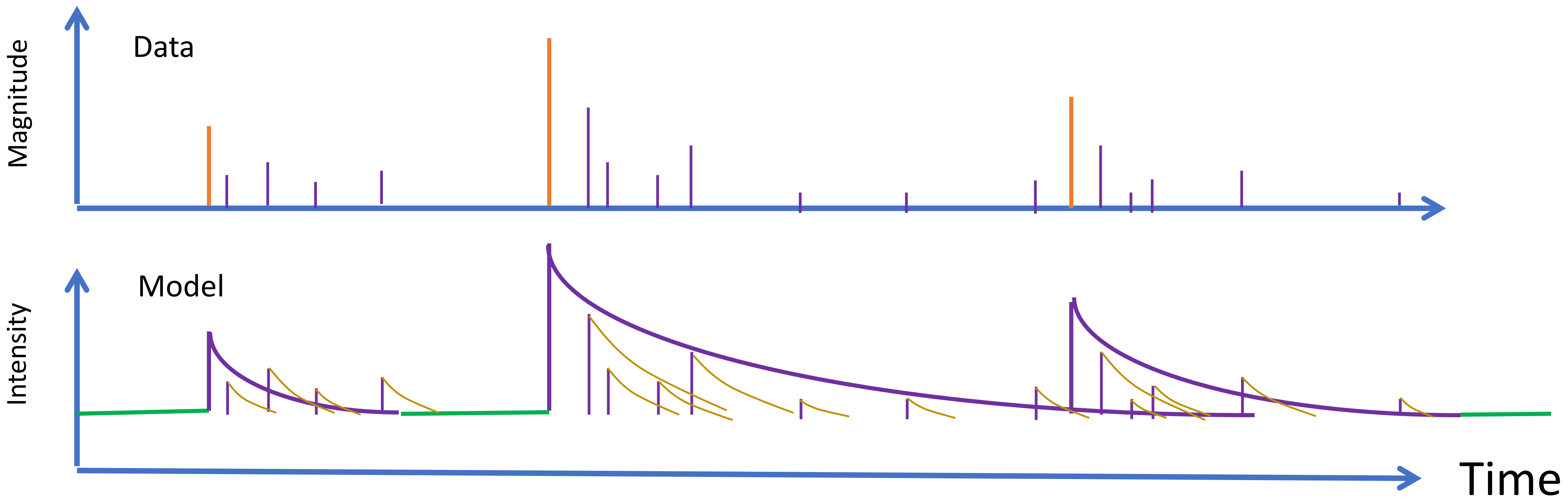
2. Time dependent ETAS like model (Amatrice)

- Inlabru only has functionality for log-Gaussian Cox process
- We needed to extend to Hawkes processes (See Francesco's poster)
  - Self exciting point process
  - Trick: lgcp with numerical integration for a sample of ETAS parameters



2. Time dependent ETAS like model (Amatrice)

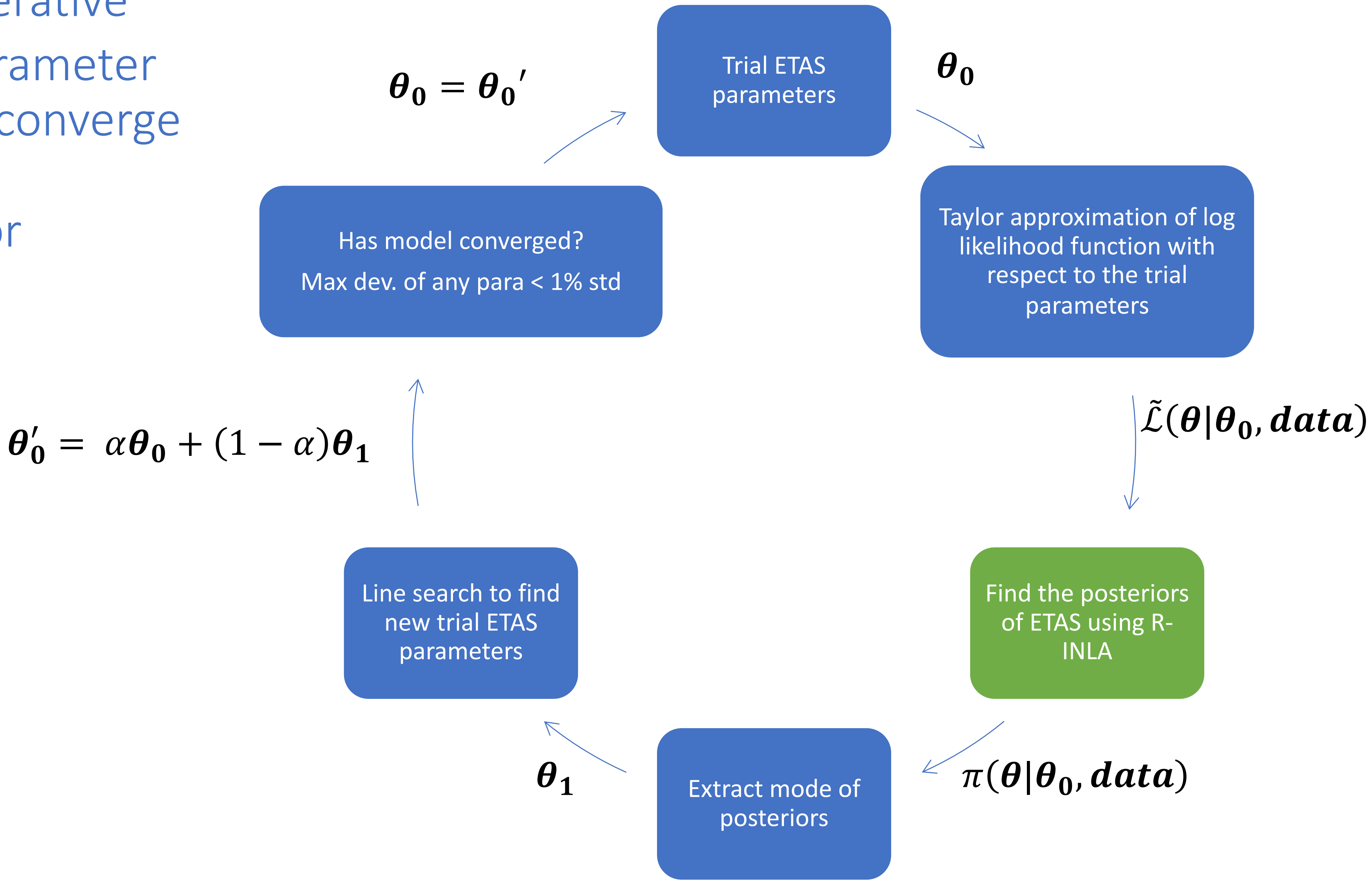
- Inlabru only has functionality for log-Gaussian Cox process
- We needed to extend to Hawkes processes (See Francesco's poster)
  - Self exciting point process
  - Trick: lgcp with numerical integration for a sample of ETAS parameters

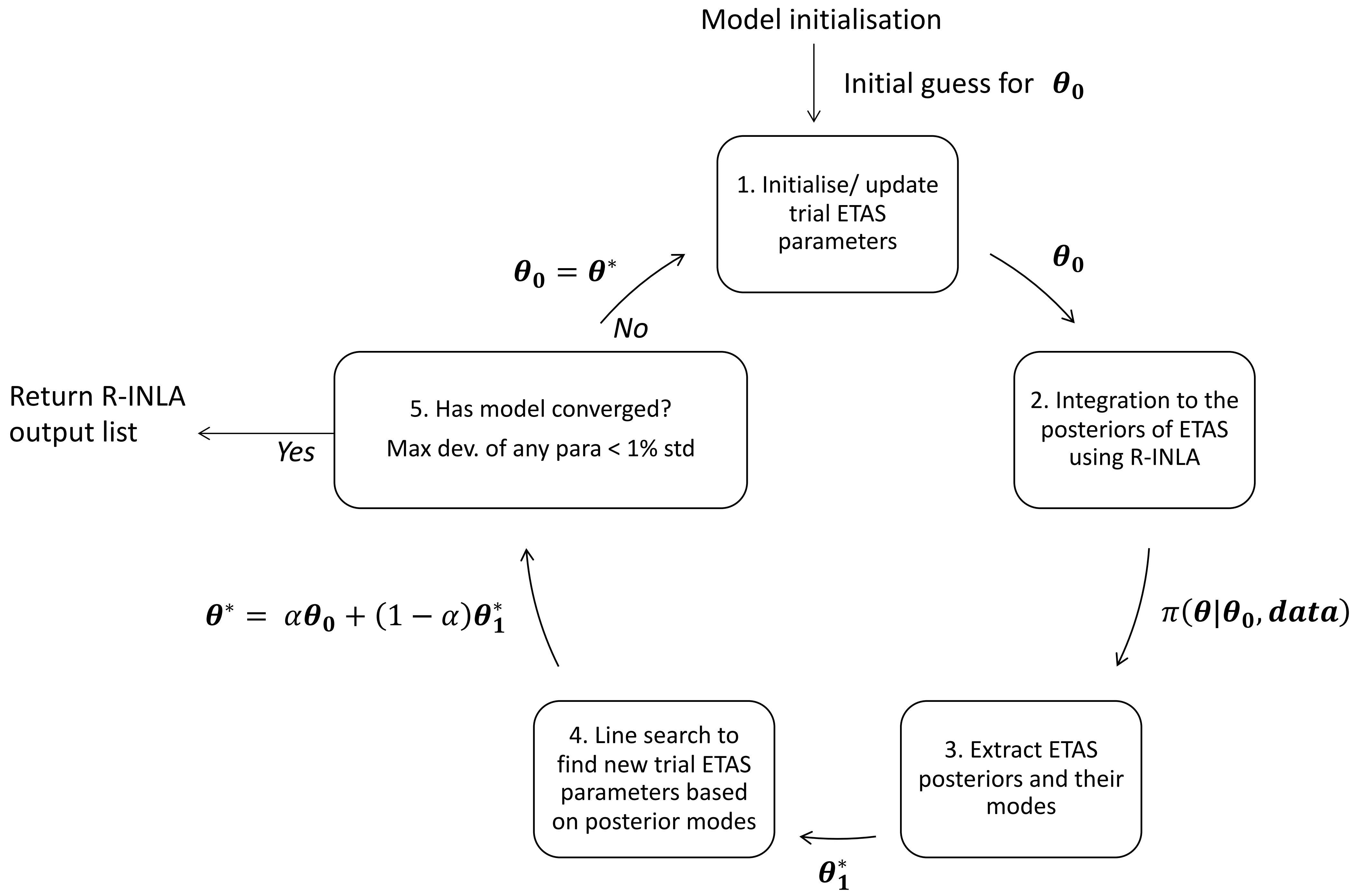


Inlabru

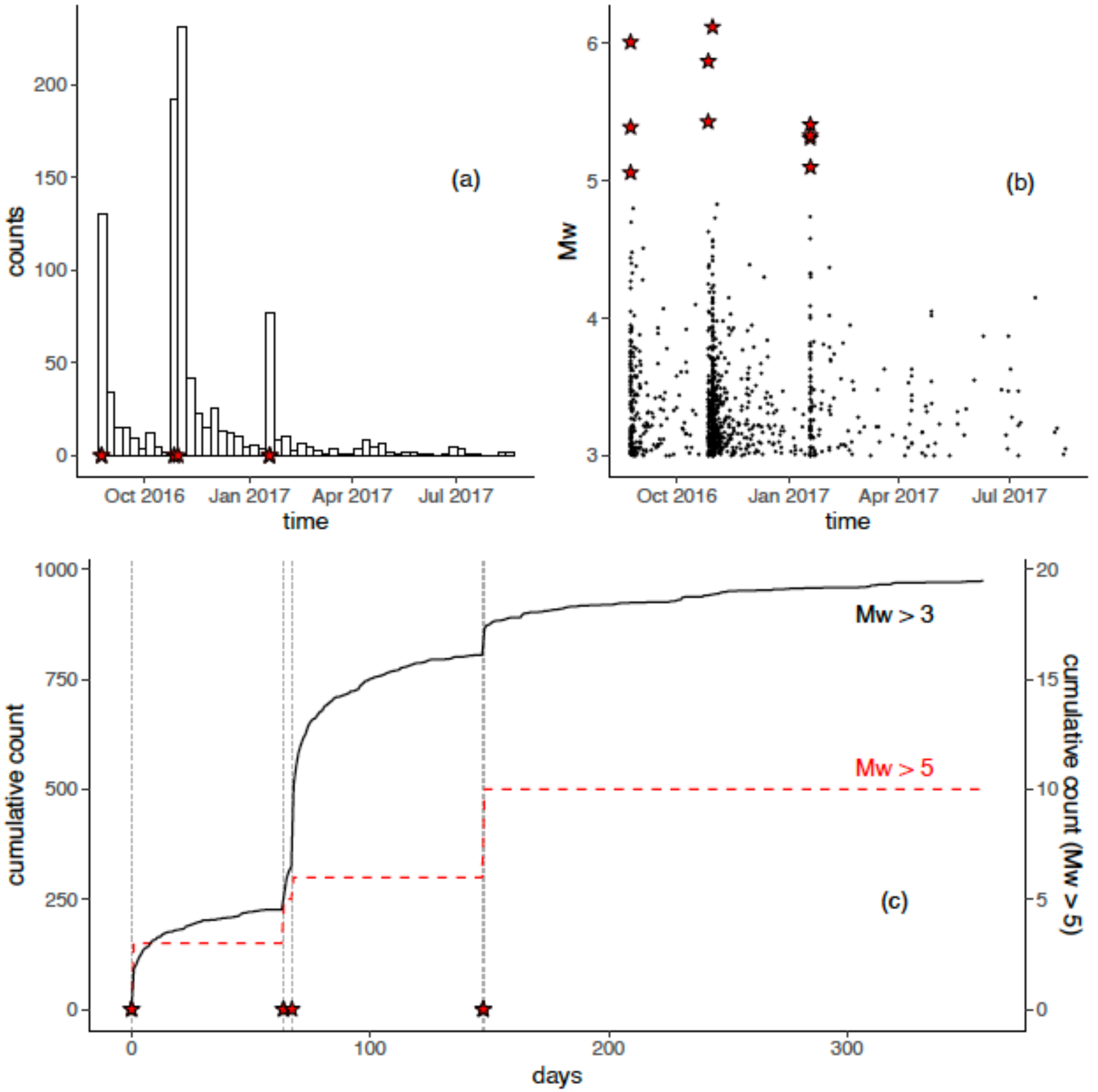
- Manages the iterative update of trial parameter modes until they converge

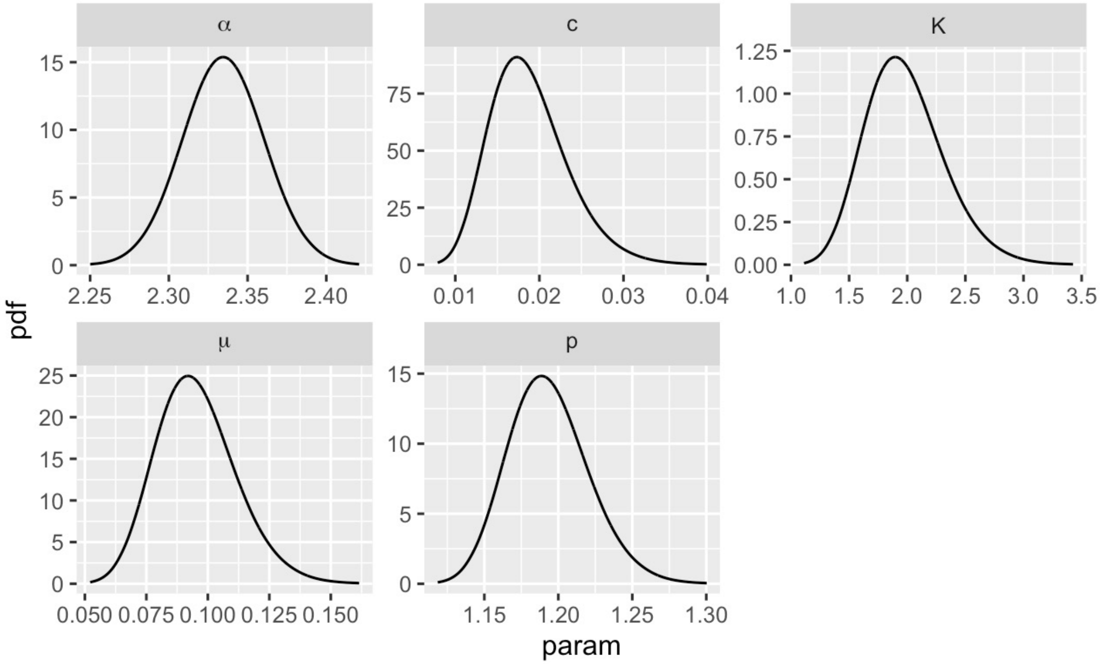
- Returns posterior distributions




Model initialisation
Initial guess for $\boldsymbol{\theta_0}$
1. Initialise/ update trial ETAS parameters
$\boldsymbol{\theta_0}$
2. Integration to the posteriors of ETAS using R-INLA
$\pi(\boldsymbol{\theta}|\boldsymbol{\theta_0}, \boldsymbol{data})$
3. Extract ETAS posteriors and their modes
$\boldsymbol{\theta_1^*}$
4. Line search to find new trial ETAS parameters based on posterior modes
$\boldsymbol{\theta}^* = \alpha\boldsymbol{\theta_0} + (1-\alpha)\boldsymbol{\theta_1^*}$
5. Has model converged?
Max dev. of any para < 1% std
Yes
Return R-INLA output list
No
$\boldsymbol{\theta_0} = \boldsymbol{\theta}^*$

Sequence: Amatrice



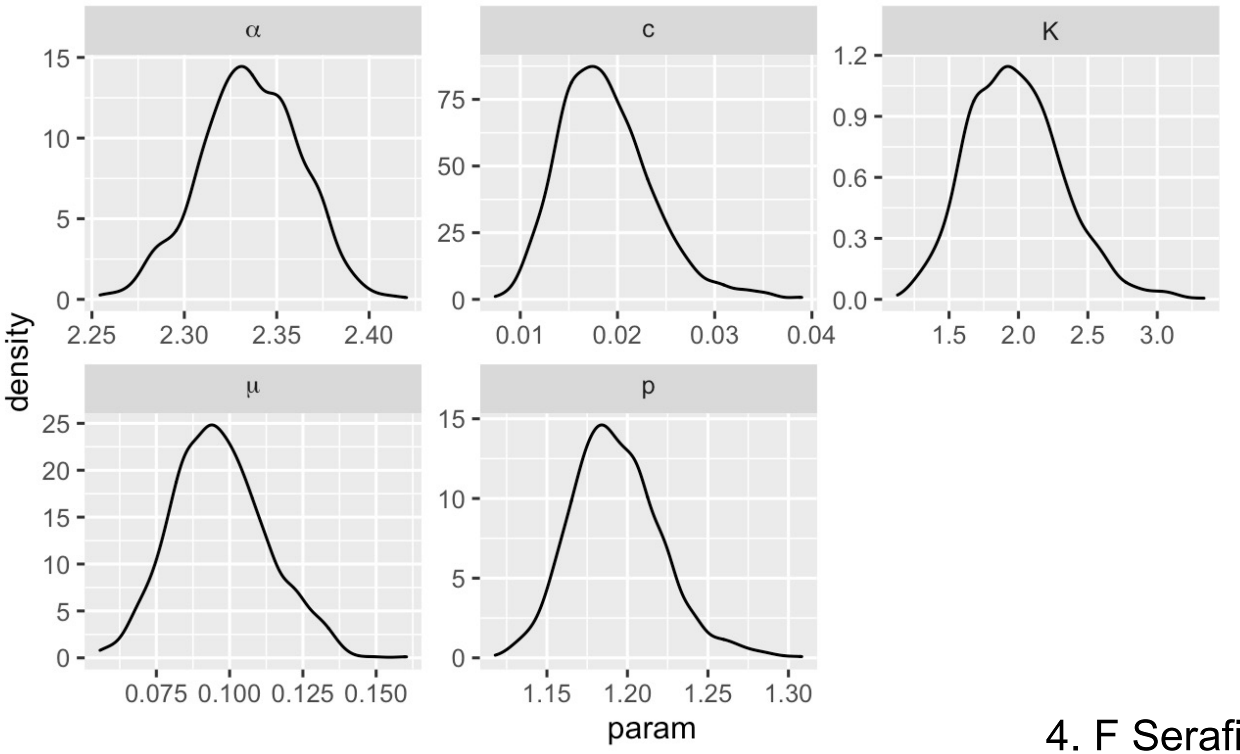


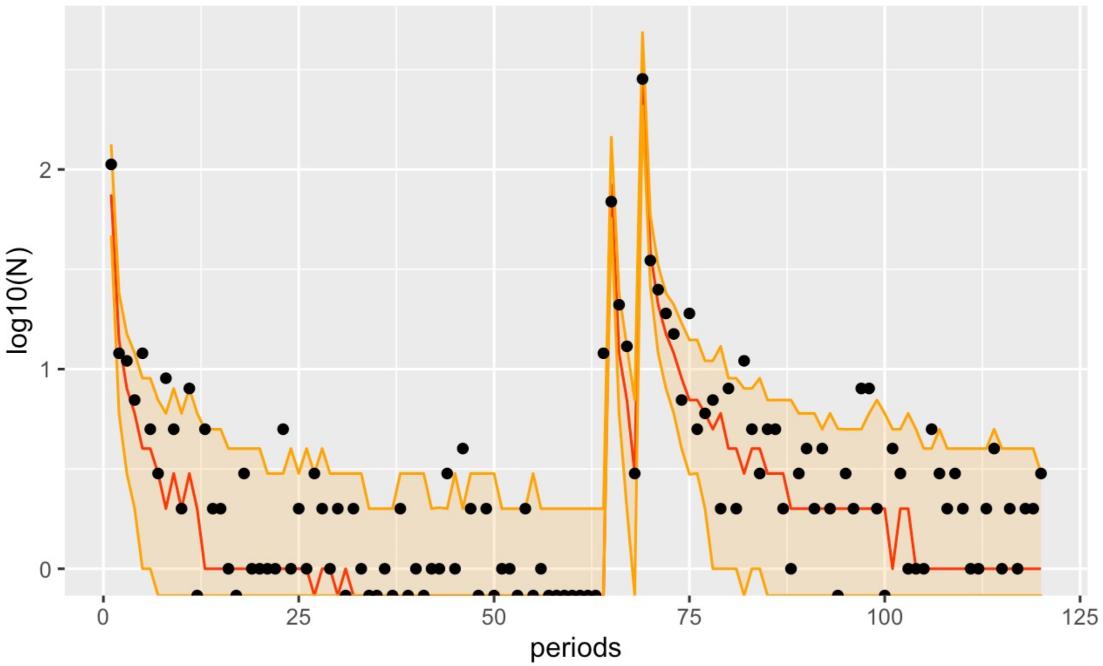
INLA derived posteriors for ETAS parameters

No declustering, jointly optimized

(MCMC fitting often uses a latent variable to partition background and clustered events)

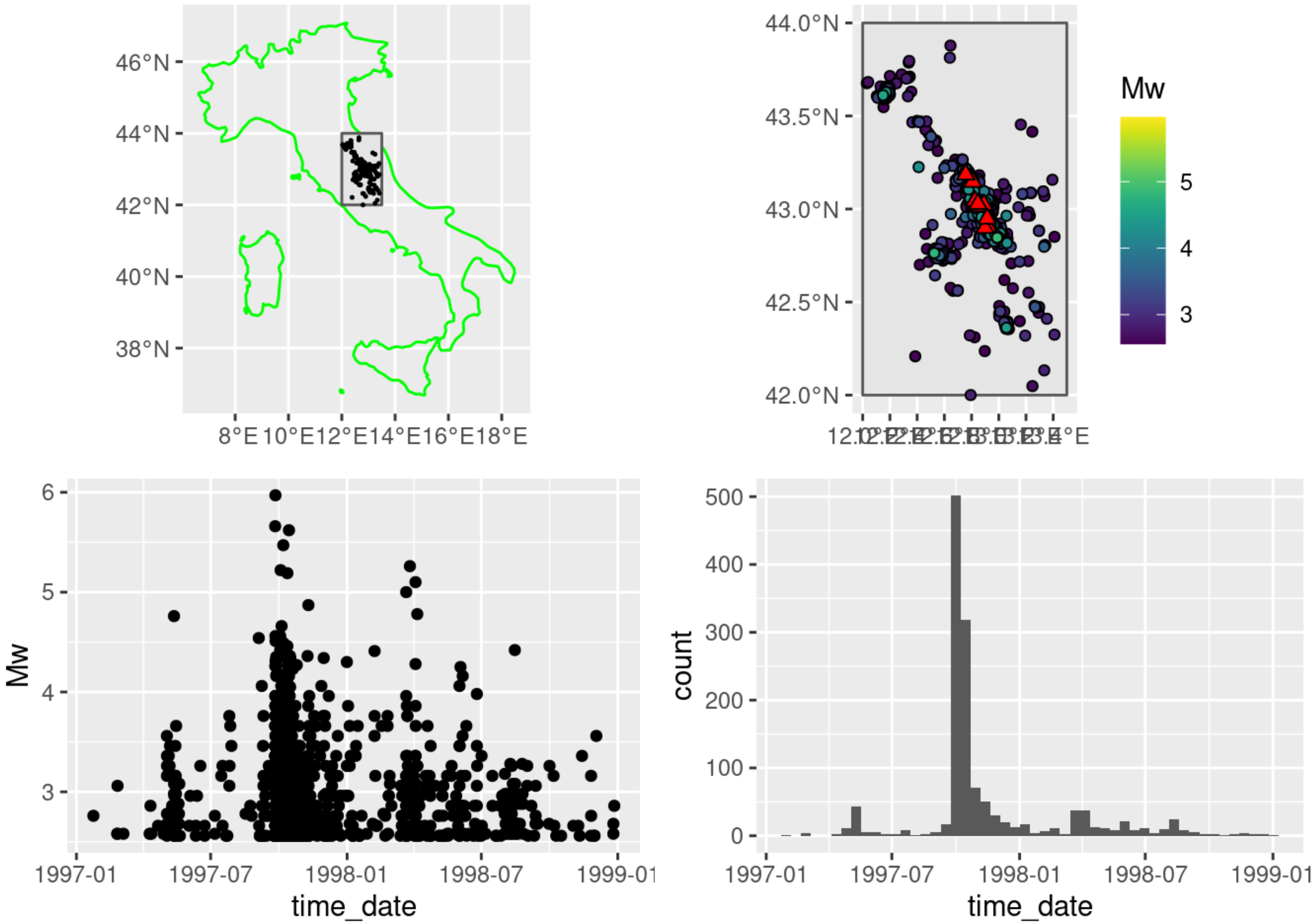
1000 samples from the joint posteriors to generate daily forecasts of following day for M>3 events





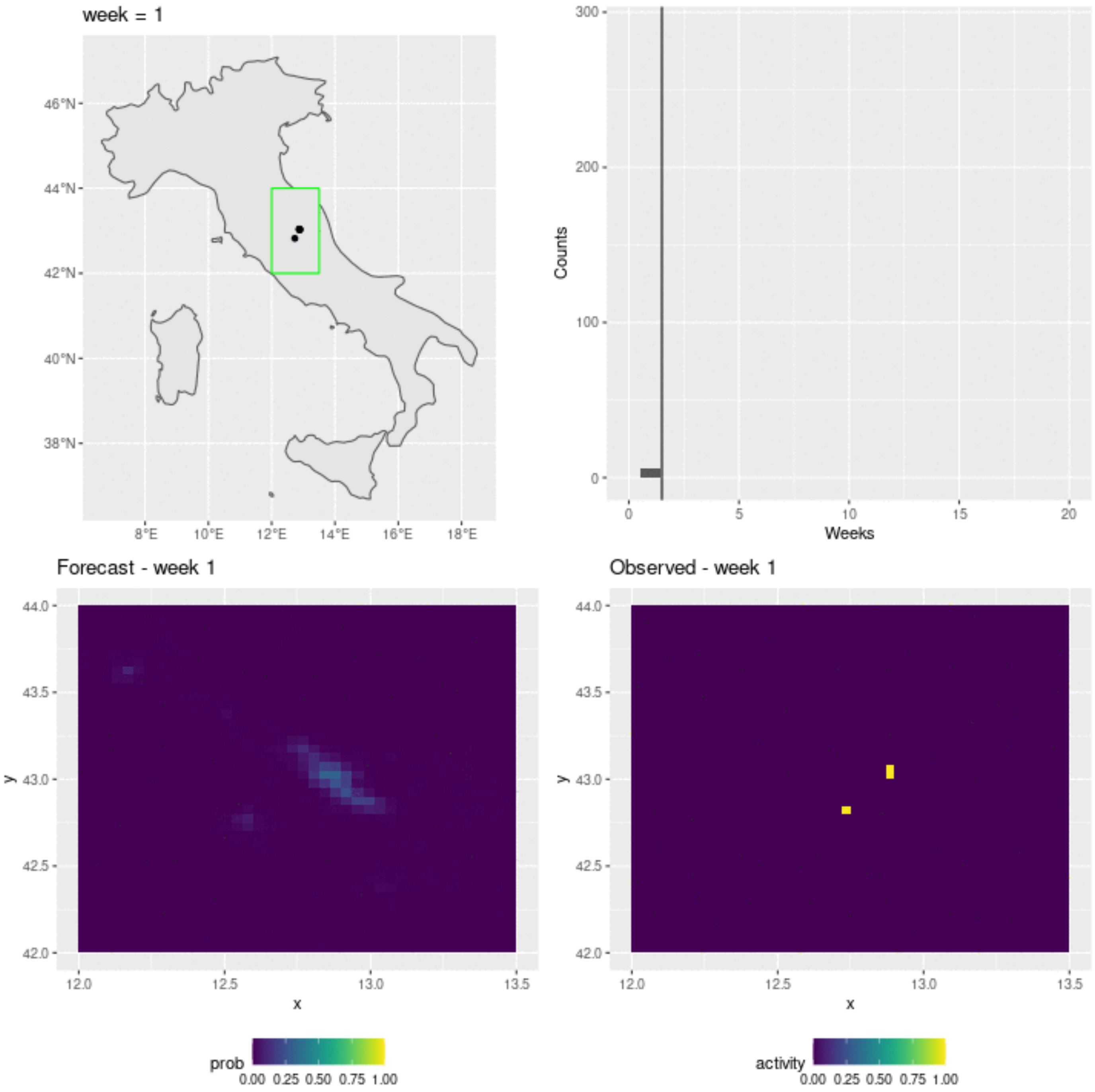
4. F Serafini, F Lindgren, M Naylor (Submitted) Approximation of bayesian Hawkes process models with Inlabru, arXiv preprint arXiv:2206.13360

3. Full spatio-temporal modelling:
Spatial (*Kirsty*) + Temporal (*Francesco*)



Sequence: Colfiorito

- Horus Catalogue, M>2.5
- 2 years of data
- ~ 1,380 events
- Not (yet) pseudo-prospective



Weekly retrospective forecast

- Here, the the spatial model is just the random field (work in progress)
- Summary statistics for 10,000 simulations

Some of you might be interested in speed

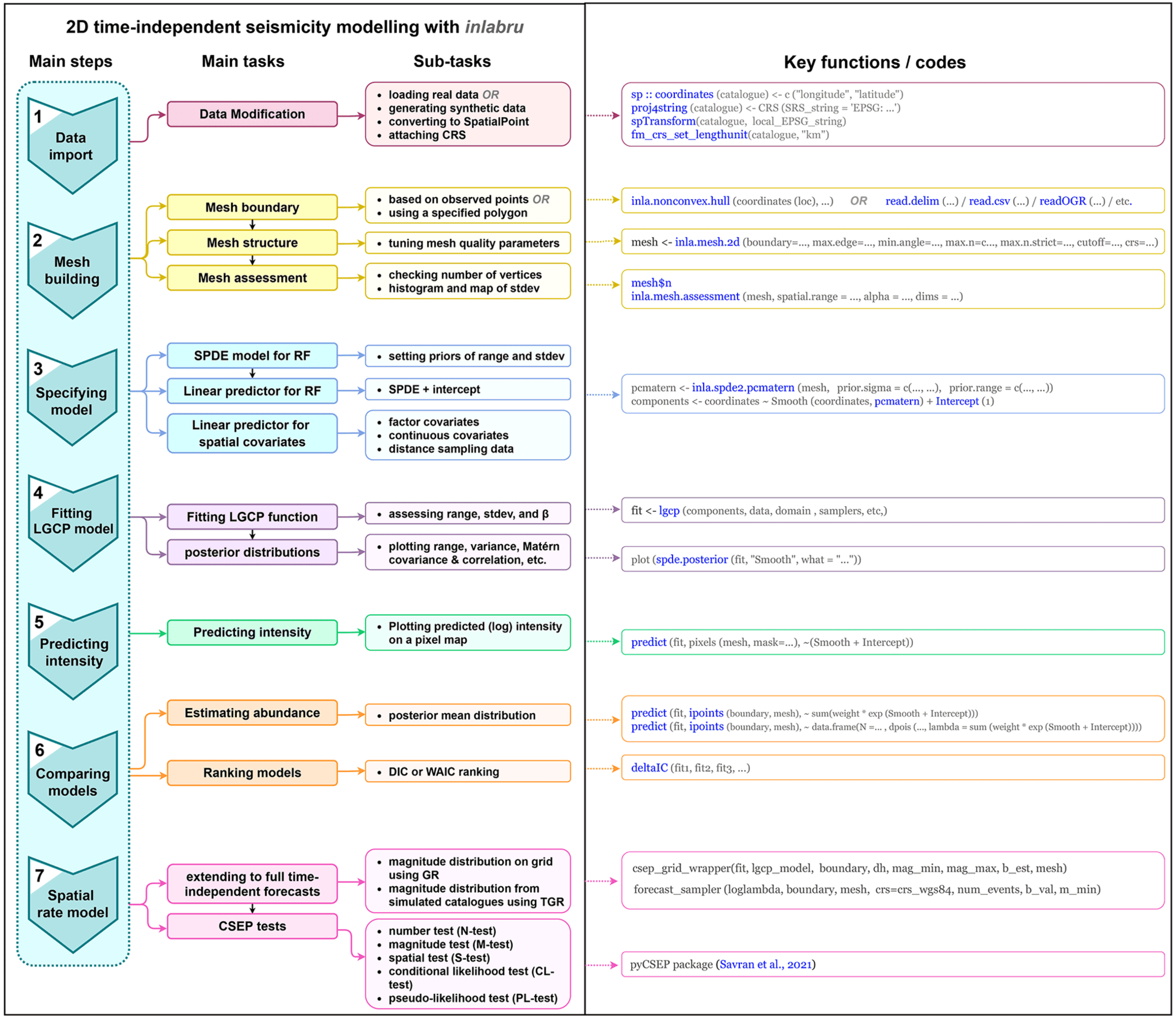
- Fit retrospective model ~13 minutes on laptop
- 10,000 simulations ~1.5hrs on laptop but…

…it is trivially parallel

Looking forward

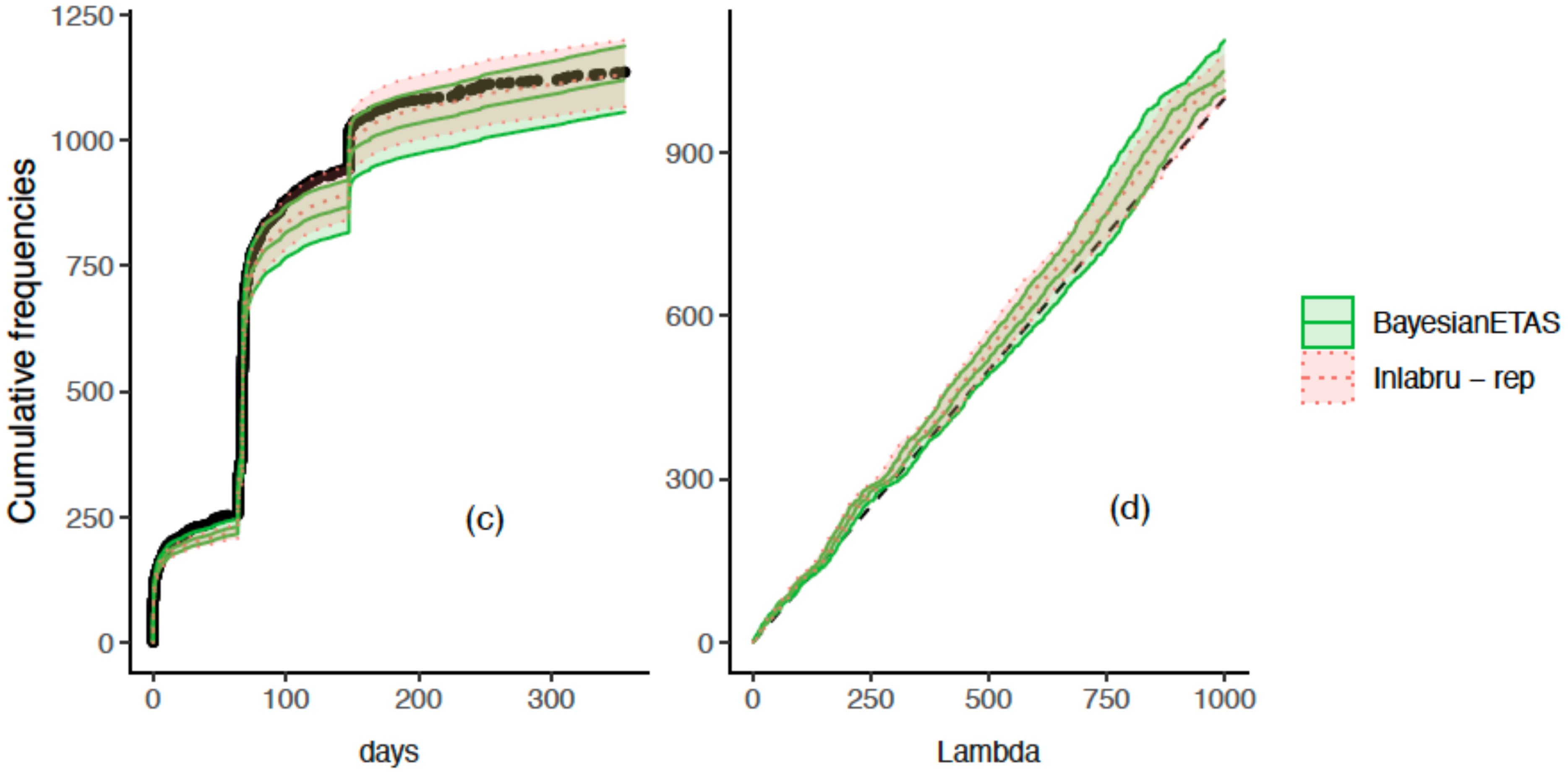
- OEF
  - Spatial structure derived from spatial covariates, random field and locally anisotropic kernel (E.g. Ross's talk)
  - Work with available data types and perform hypothesis testing within a consistent framework
  - Forecasts sampling ETAS and spatial covariate posteriors
- PSHA
  - Generate background maps which include uncertainty in the clustering process
  - Expert judgement goes into the priors…
- Exciting opportunities
  - We can let the latent field describe differences rather than drawing boxes around data
  - ETAS parameters and spatial anisotropy linear functions of covariates and random field
  - Surface geology shape files to distinguish evidence of no faults from absence of evidence for faults under basins
  - Explore how quality of forecasts degrades when only some covariates available
  - Censoring for temporal incompleteness
  - Test time dependent covariates (e.g. Coulomb maps)
  - …

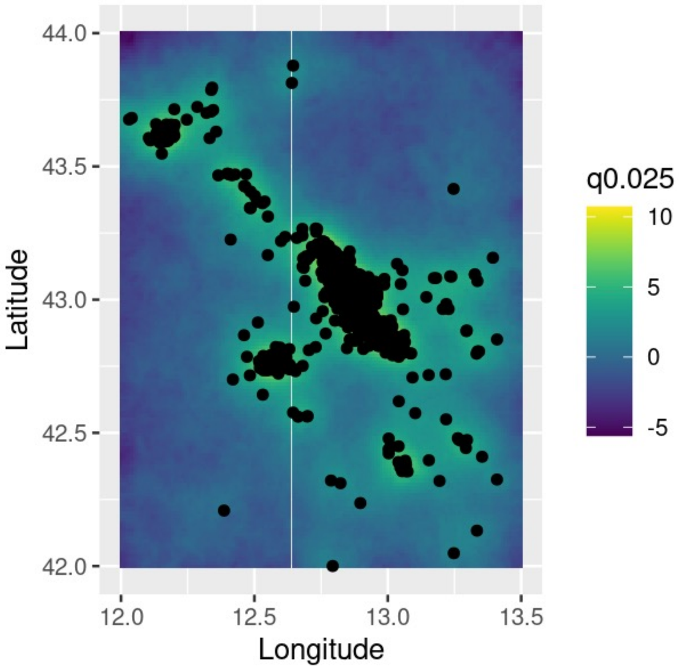
Workflow

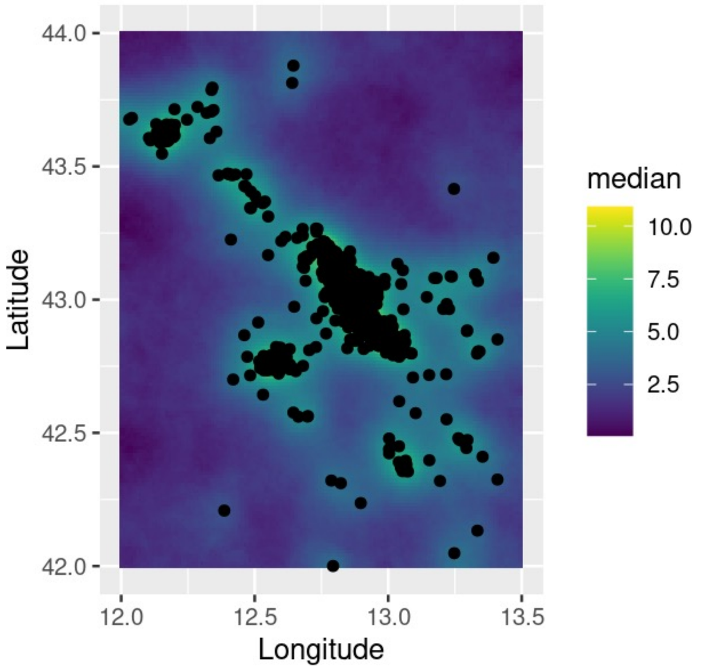


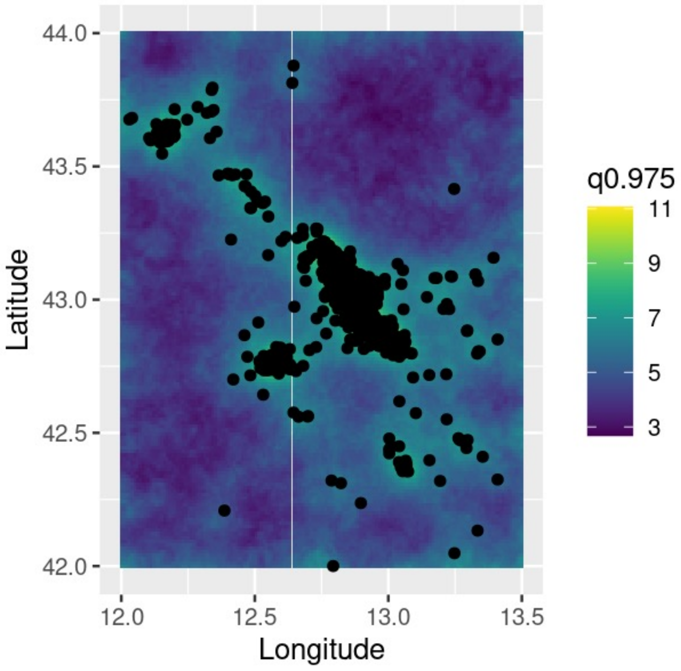
2D time-independent seismicity modelling with inlabru
Main steps
Main tasks
Sub-tasks
Key functions / codes
1 Data import
Data Modification
loading real data OR
generating synthetic data
converting to SpatialPoint
attaching CRS
sp :: coordinates (catalogue) <- c ("longitude", "latitude")
proj4string (catalogue) <- CRS (SRS_string = 'EPSG: ...')
spTransform(catalogue, local_EPSG_string)
fm_crs_set_lengthunit(catalogue, "km")
2 Mesh building
Mesh boundary
Mesh structure
Mesh assessment
based on observed points OR
using a specified polygon
tuning mesh quality parameters
checking number of vertices
histogram and map of stdev
inla.nonconvex.hull (coordinates (loc), ...) OR read.delim (...) / read.csv (...) / readOGR (...) / etc.
mesh <- inla.mesh.2d (boundary=..., max.edge=..., min.angle=..., max.n=c..., max.n.strict=..., cutoff=..., crs=...)
mesh$n
inla.mesh.assessment (mesh, spatial.range = ..., alpha = ..., dims = ...)
3 Specifying model
SPDE model for RF
Linear predictor for RF
Linear predictor for spatial covariates
setting priors of range and stdev
SPDE + intercept
factor covariates
continuous covariates
distance sampling data
pcmatern <- inla.spde2.pcmatern (mesh, prior.sigma = c(..., ...), prior.range = c(..., ...))
components <- coordinates ~ Smooth (coordinates, pcmatern) + Intercept (1)
4 Fitting LGCP model
Fitting LGCP function
posterior distributions
assessing range, stdev, and β
plotting range, variance, Matérn covariance & correlation, etc.
fit <- lgcp (components, data, domain , samplers, etc,)
plot (spde.posterior (fit, "Smooth", what = "..."))
5 Predicting intensity
Predicting intensity
Plotting predicted (log) intensity on a pixel map
predict (fit, pixels (mesh, mask=...), ~(Smooth + Intercept))
6 Comparing models
Estimating abundance
Ranking models
posterior mean distribution
DIC or WAIC ranking
predict (fit, ipoints (boundary, mesh), ~ sum(weight * exp (Smooth + Intercept)))
predict (fit, ipoints (boundary, mesh), ~ data.frame(N =... , dpois (..., lambda = sum (weight * exp (Smooth + Intercept))))
deltaIC (fit1, fit2, fit3, ...)
7 Spatial rate model
extending to full time-independent forecasts
CSEP tests
magnitude distribution on grid using GR
magnitude distribution from simulated catalogues using TGR
number test (N-test)
magnitude test (M-test)
spatial test (S-test)
conditional likelihood test (CL-test)
pseudo-likelihood test (PL-test)
csep_grid_wrapper(fit, lgcp_model, boundary, dh, mag_min, mag_max, b_est, mesh)
forecast_sampler (loglambda, boundary, mesh, crs=crs_wgs84, num_events, b_val, m_min)
pyCSEP package (Savran et al., 2021)

Temporal inlabru vs BayesianETAS (Amatrice)

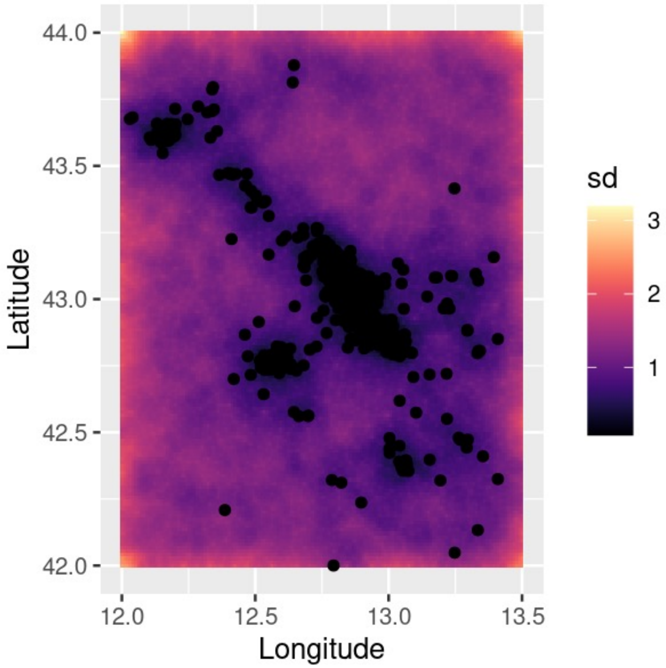




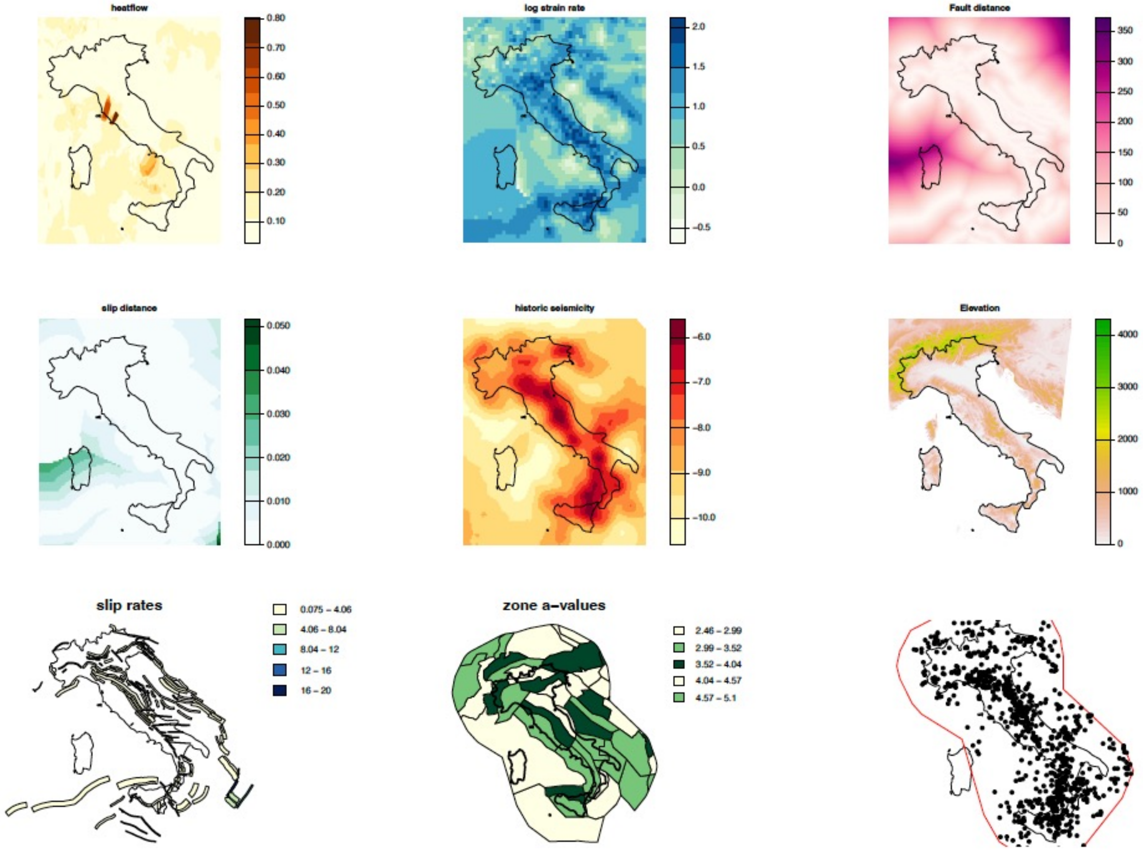


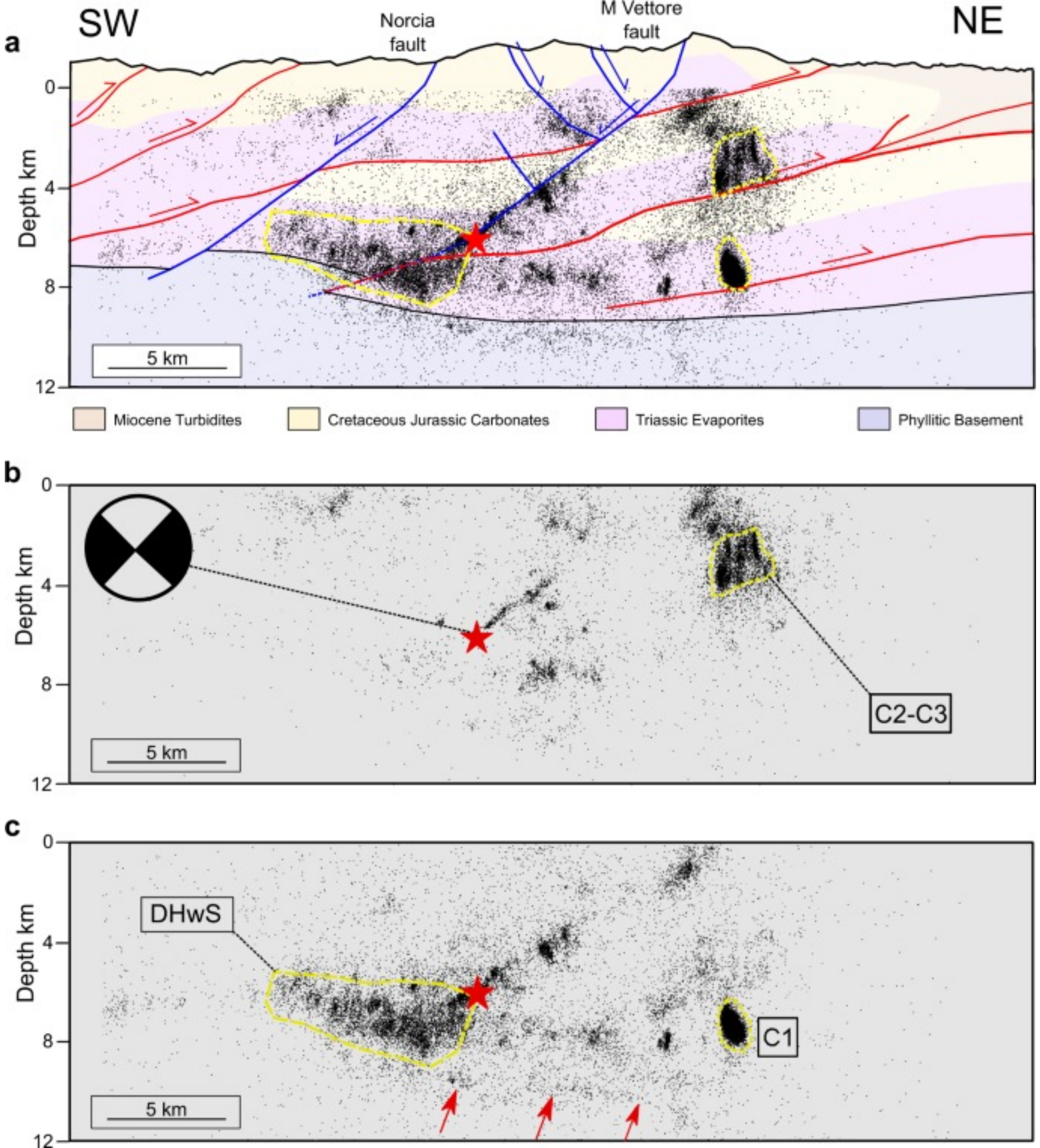


Retrospective Spatio-temporal Colfiorito



Amatrice





Collettini, C., Barchi, M.R., De Paola, N. *et al.* Rock and fault rheology explain differences between on fault and distributed seismicity. *Nat Commun* **13**, 5627 (2022). https://doi.org/10.1038/s41467-022-33373-y