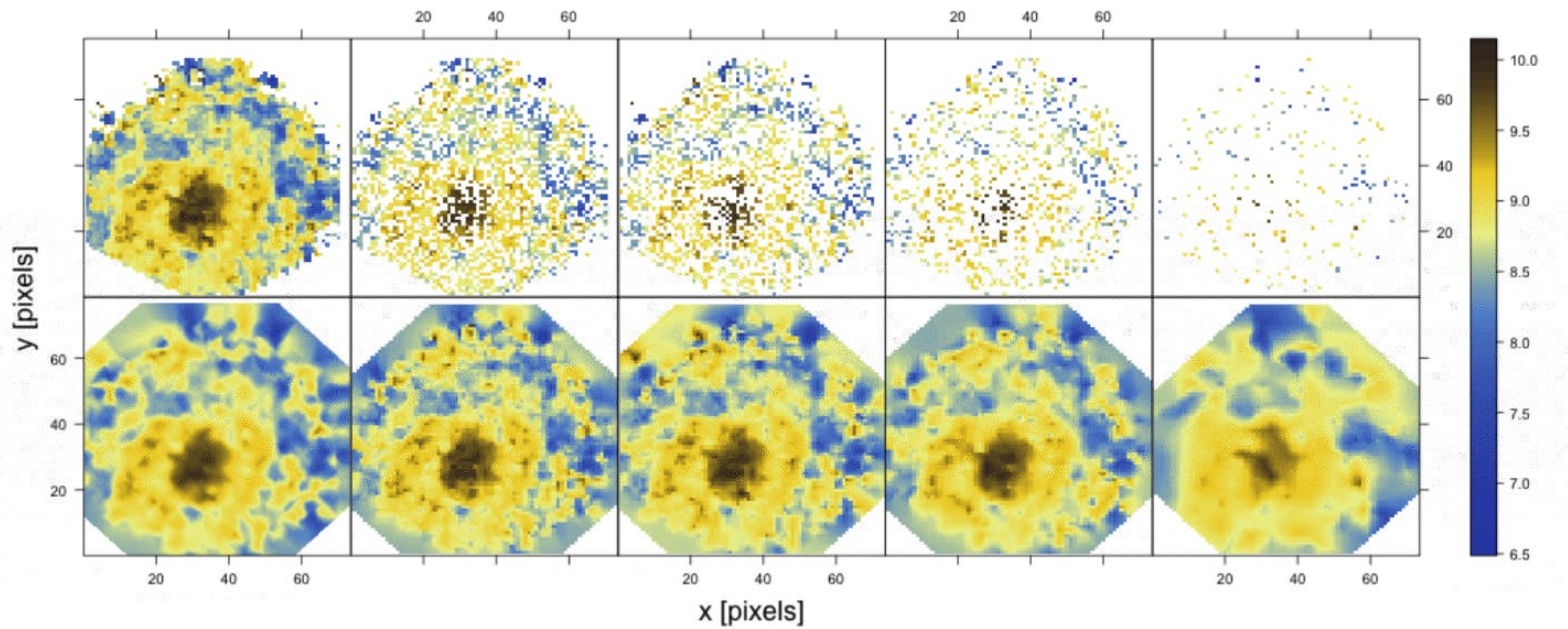
Spatial Field Reconstruction with INLA

João Rino-Silvestre, Centra-Sim, University of Lisbon

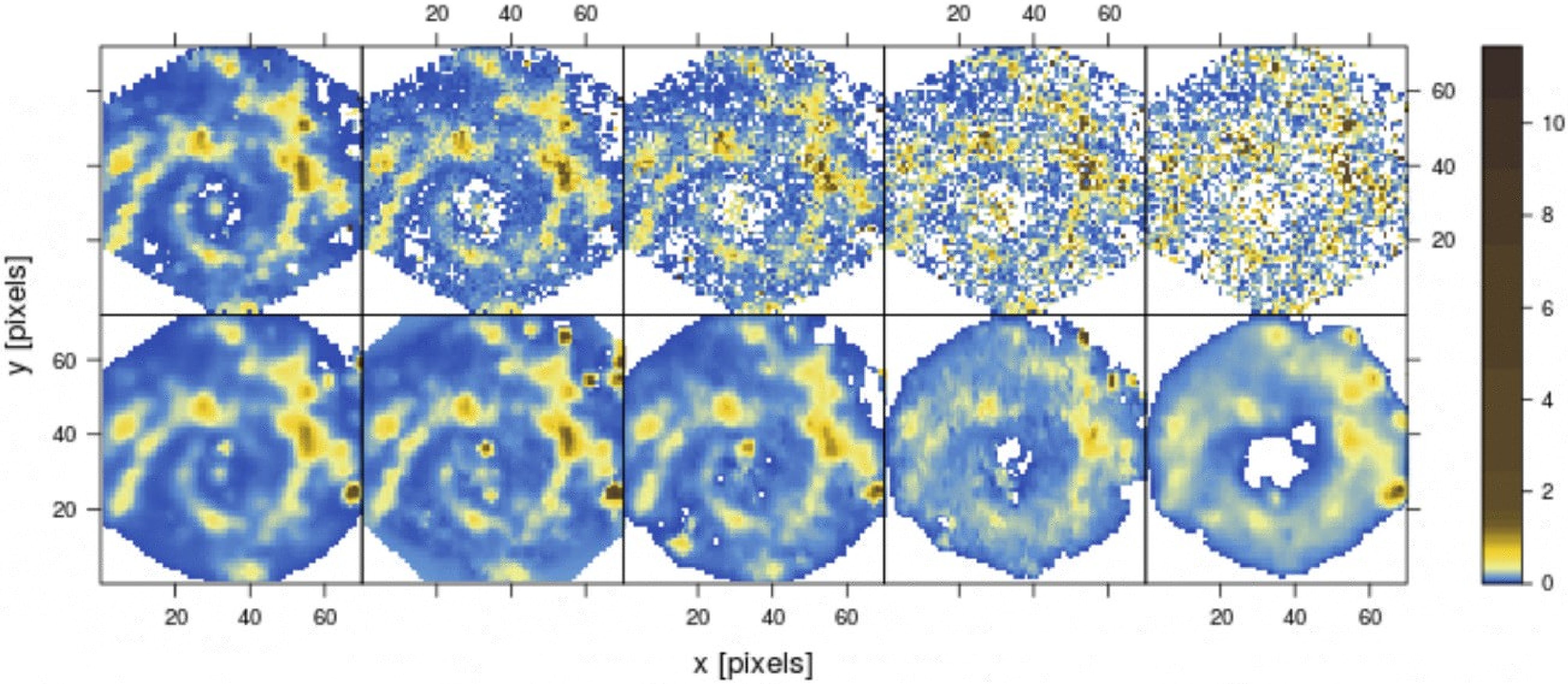
November 12th, 2019

Examples



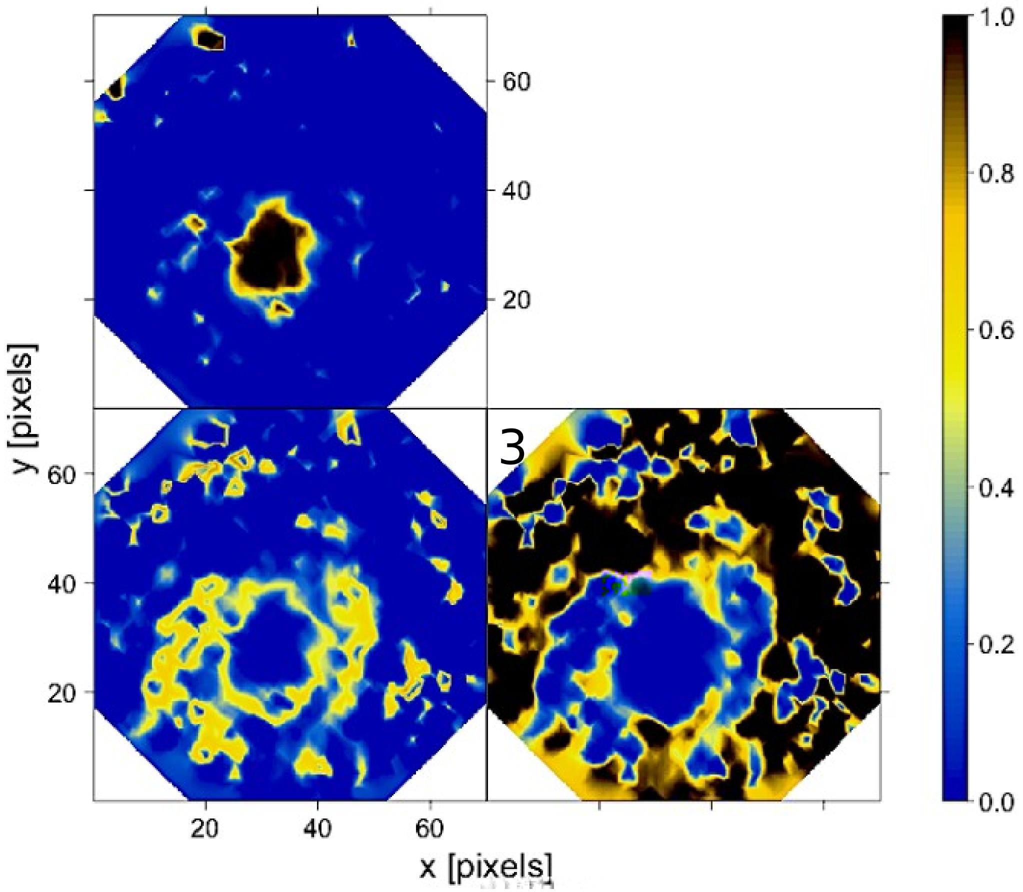
Predictions from INLA for input starlight age of NGC 0309 when 100, 75, 50, 25 and 5% (left to right) of the data is used. Upper panels show the starlight input, bottom the INLA prediction (source, González-Gaitán et al. (2018)).

Examples



Predictions from INLA for Hα EW map of NGC 0309 with S/N of 10, 2, 1, 0.5 and 0.3 (left to right). Upper panels show input and bottom panels INLA predictions (source, González-Gaitán et al. (2018)).

Examples



Probability maps of NGC 0309 for three ranges of age, log(t[yr]), arranged anticlockwise: log(t[yr])>9.5, 9.3>log(t[yr])>9.0 and log(t[yr]<8.9. The bins were chosen to represent bottom (<2.5%), middle (32%−68%) and top (>97.5%) quantiles of the reconstructed population age map (source, González-Gaitán et al. (2018)).

Structure

- **Introduction**
  - Bayesian Statistics
  - Latent Gaussian Models (LGMs)
    - Notation & Properties
  - Inference
- **INLA**
  - A different approach
  - The Method
  - Applications in Astronomy
- **Present**

Introduction

Bayesian Statistics

Let $\boldsymbol{x}$ be a latent field and $\boldsymbol{y}$ an observable

- Prior density: $\pi(\boldsymbol{y}|\boldsymbol{x})$
- Posterior density: $\pi(\boldsymbol{x}|\boldsymbol{y})$
- Joint density: $\pi(\boldsymbol{x},\boldsymbol{y}) = \pi(\boldsymbol{x})\cdot\pi(\boldsymbol{y}|\boldsymbol{x}) = \pi(\boldsymbol{y})\cdot\pi(\boldsymbol{x}|\boldsymbol{y})$
- Marginal posterior density: $\pi(x_i|\boldsymbol{y}) = \int \pi(\boldsymbol{x}|\boldsymbol{y})\mathrm{d}\boldsymbol{x}_{-i}$

Latent Gaussian Models (LGMs)

- Bayesian Additive Models (BAMs)
  - y_i is assumed to belong to an exponential family with mean μ_i
  - y_i is linked to a structured additive predictor, η_i, via a link function $g(.)$, such that $g(\mu_i)=\eta_i$, where

$$\eta_i = \alpha + \sum_{j=1}^{n_f} f^{(j)}(u_{ji}) + \sum_{k=1}^{n_\beta} \beta_k z_{ki} + \varepsilon_i \qquad (1)$$

- LGMs are a subset of BAMs, with a predictor as (1) and which assign a Gaussian prior to α, $\{f^{(j)}(.)\}$, $\{\beta_k\}$ and $\{\varepsilon_i\}$.
- **Applications:** relaxation of regression models, dynamic models (u_t), spatial models (u_s), ...

Latent Gaussian Models (LGMs)

- **Notation**
  - $\pi(.|.)$ - conditional density of its arguments
  - $\boldsymbol{x}$ - all n Gaussian variables $\{\eta_i\}$, $\{f^{(j)}(.)\}$, $\{\beta_k\}$ and $\{\varepsilon_i\}$
  - $\pi(\boldsymbol{x}|\boldsymbol{\theta}_1)$ - is Gaussian with assumed zero mean, precision matrix $\boldsymbol{Q}(\boldsymbol{\theta}_1)$, and hyperparameters $\boldsymbol{\theta}_1$
  - $N(\boldsymbol{x}; \boldsymbol{\mu},\boldsymbol{\Sigma})$ - Gaussian density $N(\boldsymbol{\mu},\boldsymbol{\Sigma})$ at configuration $\boldsymbol{x}$
  - $\pi(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta}_2)$ - the distribution for the n_d observables $\boldsymbol{y}$ (assumed conditionally independent given $\boldsymbol{x}$ and $\boldsymbol{\theta}_2$)
  - $\boldsymbol{\theta}=(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2)^T$, with $\dim(\boldsymbol{\theta})=m$

Latent Gaussian Models (LGMs)

The posterior then reads:

- $\pi(\boldsymbol{x}, \boldsymbol{\theta} | \boldsymbol{y}) \propto \pi(\boldsymbol{\theta}) \pi(\boldsymbol{x} | \boldsymbol{\theta}) \prod_i \pi(y_i | x_i, \boldsymbol{\theta})$

$$\propto \pi(\boldsymbol{\theta}) |\boldsymbol{Q}(\boldsymbol{\theta})|^{1/2} \exp\left(-\frac{1}{2} \boldsymbol{x}^{\boldsymbol{T}} \boldsymbol{Q}(\boldsymbol{\theta}) \boldsymbol{x} + \sum_i \log(\pi(y_i | x_i, \boldsymbol{\theta}))\right)$$

- **Properties** (satisfied by many LGMs but not all)
  - Latent field $\boldsymbol{x}$ admits conditional independence properties, making it a Gaussian Markov random field with a sparse precision matrix $\boldsymbol{Q}(\boldsymbol{\theta})$
  - The number of hyperparameters, m, is small ($m \leq 6$)

Both are usually required to produce fast inference

Inference

- **Aim:** infer posterior marginals for $\pi(x_i|\boldsymbol{y})$, $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ and $\pi(\ _j|\boldsymbol{y})$

- **Possibilities:**
  - Markov Chain Monte Carlo
    - Poor performance when applied to LGMs
  - Deterministic approximations
    - Better computational cost

A different approach

- The posterior marginals of interest can be written as

$$\pi(x_i|\boldsymbol{y}) = \int \pi(x_i|\boldsymbol{\theta}, \boldsymbol{y})\pi(\boldsymbol{\theta}|\boldsymbol{y})\,d\boldsymbol{\theta}$$

$$\pi(\ \ _j|\boldsymbol{y}) = \int \pi(\boldsymbol{\theta}|\boldsymbol{y})\,d\boldsymbol{\theta}_{-j}$$

- INLA (Rue et al. (2009)) uses this form to construct nested approximations

$$\tilde{\pi}(x_i|\boldsymbol{y}) = \int \tilde{\pi}(x_i|\boldsymbol{\theta}, \boldsymbol{y})\tilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y})\,d\boldsymbol{\theta}$$

$$\tilde{\pi}(\ \ _j|\boldsymbol{y}) = \int \tilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y})\,d\boldsymbol{\theta}_{-j}$$

where $\tilde{\pi}(.|.)$ is an approximated density of its arguments and the integrations are performed numerically. The Laplace approximation of $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ is given by

$$\tilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y}) \propto \frac{\pi(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{y})}{\tilde{\pi}_{!}(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{y})}\Big|_{\boldsymbol{x}=\boldsymbol{x}*(\boldsymbol{\theta})} \qquad (2)$$

The Method

1) Exploring $\tilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y})$

a) locate mode of $\tilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y})$, $\boldsymbol{\theta^*}$: using the difference between successive gradient vectors, approximate second derivatives of $\log(\tilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y}))$;

b) at $\boldsymbol{\theta^*}$ compute the negative Hessian matrix $\boldsymbol{H} > 0$ and let $\boldsymbol{\Sigma} = \boldsymbol{H}^{-''}$; use standardized variables $\boldsymbol{z}$ instead of $\boldsymbol{\theta}$, using the form

$$\boldsymbol{\theta}(\boldsymbol{z}) = \boldsymbol{\theta}* + \boldsymbol{V}\boldsymbol{\Lambda}^{1/2}\boldsymbol{z}$$

c) explore $\log(\tilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y}))$: start from the mode ($\boldsymbol{z} = \boldsymbol{0}$); go in the positive direction of z_1 with step δ_z, while

$$\log(\tilde{\pi}(\boldsymbol{\theta}(\boldsymbol{0})|\boldsymbol{y})) - \log(\tilde{\pi}(\boldsymbol{\theta}(\boldsymbol{z})|\boldsymbol{y})) < \#_{\pi} \qquad (3)$$

then switch direction; treat the remaining coordinates in the same way (fig. 1)

The Method

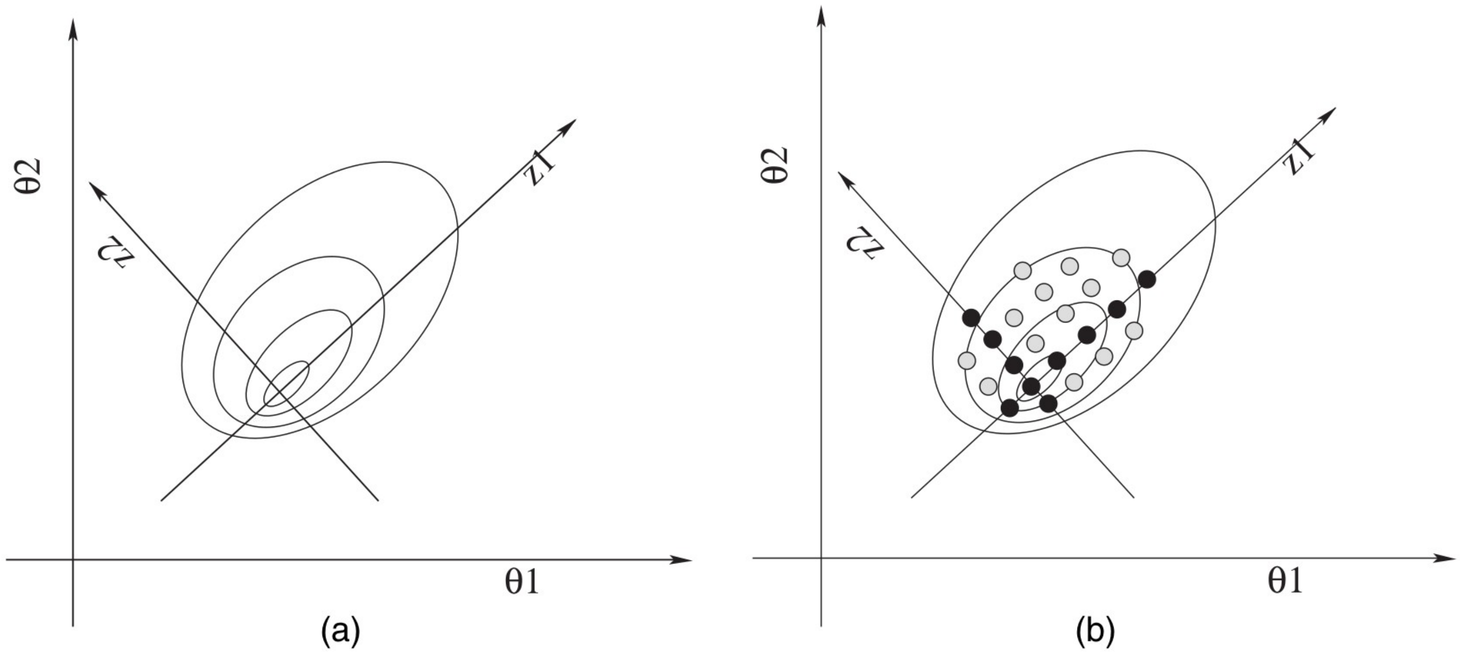


Fig. 1 - Illustration of the posterior marginal for $\boldsymbol{\theta}$: in (a) the mode is locate and the Hessian and co-ordinate system for **z** are computed; in (b) each co-ordinate direction is explored (•) while (3) prevails; new points (•) are explored combining coordinates of (•) (source: Rue et al. (2009)).

The Method

d) use computed before to construct an interpolant to $\log(\tilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y}))$ and compute marginals using numerical integration from this interpolant

2- Approximating $\pi(x_i|\boldsymbol{\theta},\boldsymbol{y})$

a) Approximate the modal configuration

$$\boldsymbol{x}_{-i}*(x_i,\boldsymbol{\theta}) \approx \%_{\tilde{\pi}_{!}}(\boldsymbol{x}_{-i}|x_i) \qquad (4)$$

b) Define a ROI around i, $\$_i(\boldsymbol{\theta})$, for only those x_j ‘close’ to x_i should have an effect on its marginal;

c) Consider the Laplace approximation

$$\tilde{\pi}_{\&'}(x_i|\boldsymbol{\theta},\boldsymbol{y}) \propto \frac{\pi(\boldsymbol{x},\boldsymbol{\theta},\boldsymbol{y})}{\tilde{\pi}_{!}(\boldsymbol{x}_{-i}|x_i,\boldsymbol{\theta},\boldsymbol{y})}\Bigg|_{\boldsymbol{x}_{-i}=\boldsymbol{x}_{-i}*(x_i,\boldsymbol{\theta})} \qquad (5)$$

The Method

d) derive a simplified Laplace approximation $\tilde{\pi}_{(\&'}(x_i|\boldsymbol{\theta}, \boldsymbol{y})$ by doing a series expansion of $\tilde{\pi}_{\&'}(x_i|\boldsymbol{\theta}, \boldsymbol{y})$ around $x_i = \mu_i(\boldsymbol{\theta})$

e) expanding the log densities of both numerator and denominator in (5) around $x_i = \mu_i(\boldsymbol{\theta})$, we get

$$\log\left(\tilde{\pi}_{(\&'}(x_i^s|\boldsymbol{\theta}, \boldsymbol{y})\right) =)*nst+nt - \frac{1}{2}\left(x_i^s\right)^2 + ,_i^{(1)}(\boldsymbol{\theta})\, x_i^s + \frac{1}{6}\left(x_i^s\right)^3 ,_i^{(3)}(\boldsymbol{\theta}) + \dots \qquad (6)$$

The Method

where

$$,_i^{(1)}(\boldsymbol{\theta})=\frac{1}{2}\sum_{j\in -\backslash i} ._j^{2}(\boldsymbol{\theta})\{1-)*//_{\tilde{\pi}_!}(x_i,x_j)^2\}d_j^{(3)}\{\mu_i(\boldsymbol{\theta}),\boldsymbol{\theta}\}\ ._j(\boldsymbol{\theta})+_{ij}(\boldsymbol{\theta})$$

$$,_i^{(3)}(\boldsymbol{\theta})=\sum_{j\in -\backslash i} d_j^{(3)}\{\mu_i(\boldsymbol{\theta}),\boldsymbol{\theta}\}\{._j(\boldsymbol{\theta})+_{ij}(\boldsymbol{\theta})\}^3$$

$$d_j^{(3)}(x_i,\boldsymbol{\theta})=\frac{\partial^3}{\partial x_j^3}\log\{\pi(y_j|x_j,\boldsymbol{\theta})\}\Big|_{x_j=\%_{\tilde{\pi}_!}(x_j|x_i)}$$

$$x_i^s=\frac{x_i-\mu_i(\boldsymbol{\theta})}{._i(\boldsymbol{\theta})}$$

The Method

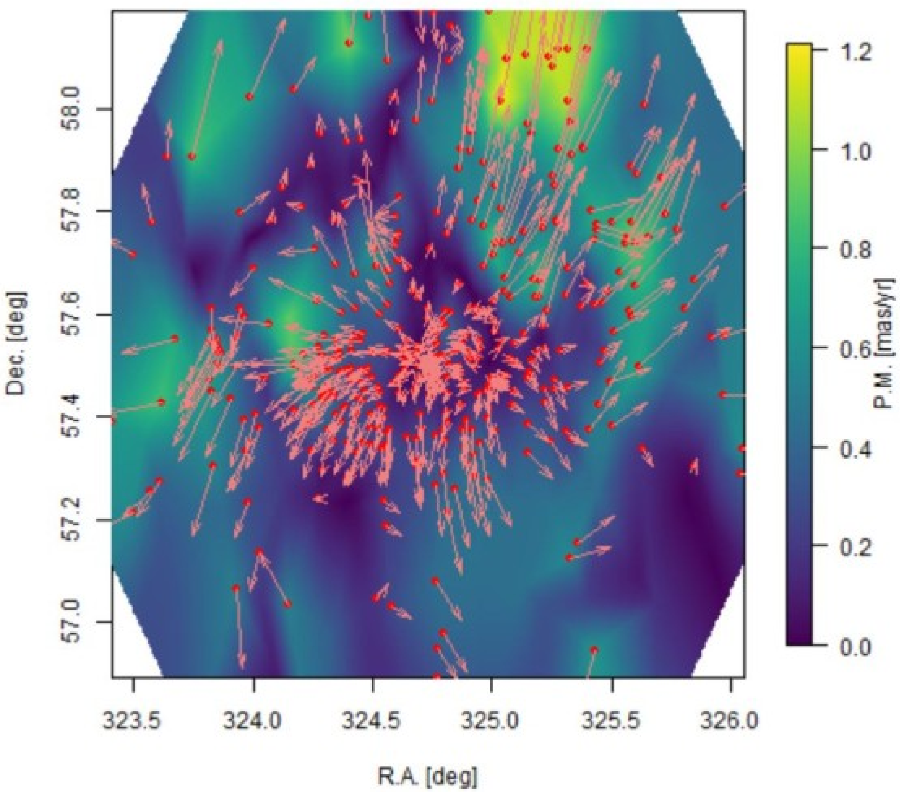
f) finally, fit a skew normal distribution of the form (7) to (6) so that the third derivative at the mode is $,_{i}^{(3)}$, the mean is $,_{i}^{(1)}$ and the variance is 1.

$$\pi_{(N}(z) = \frac{2}{2} \, 0\left(\frac{z-3}{2}\right) 1\left(+\frac{z-3}{2}\right) \qquad (7)$$

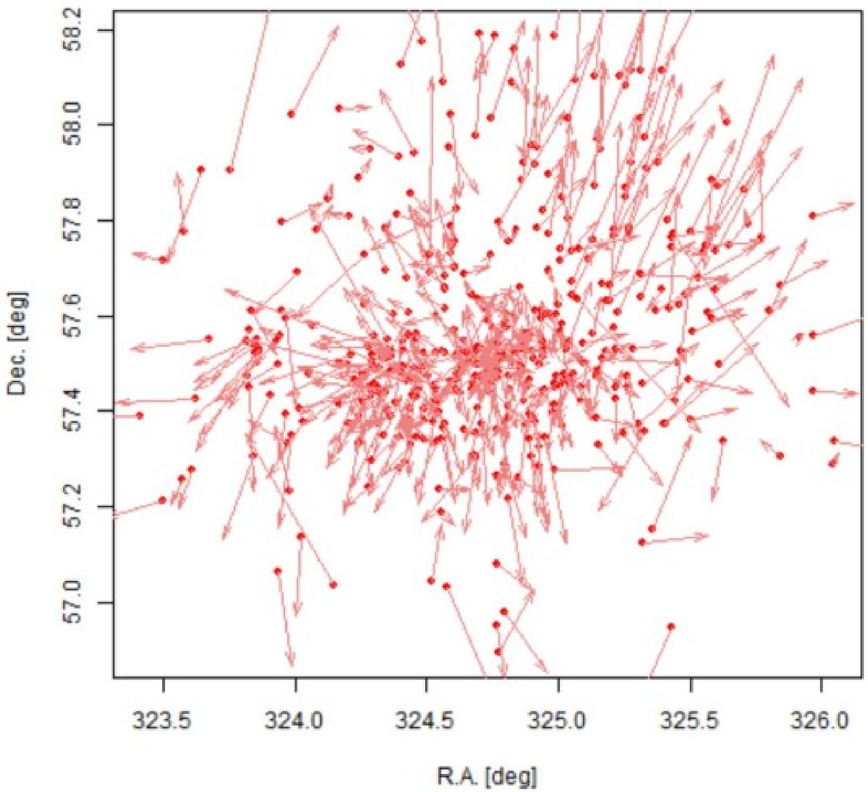
0(.) - density function
1(.) - distribution function
a - skewness parameter
ξ - location parameter
ω - scale parameter

INLA

Applications in Astronomy



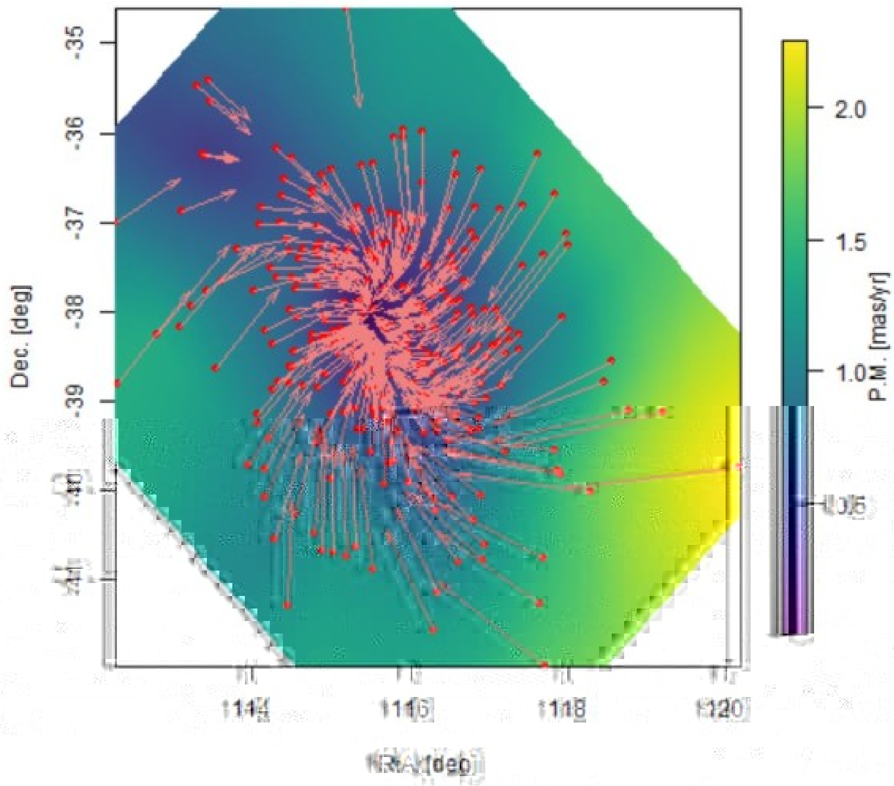
IC 1396, inferred data



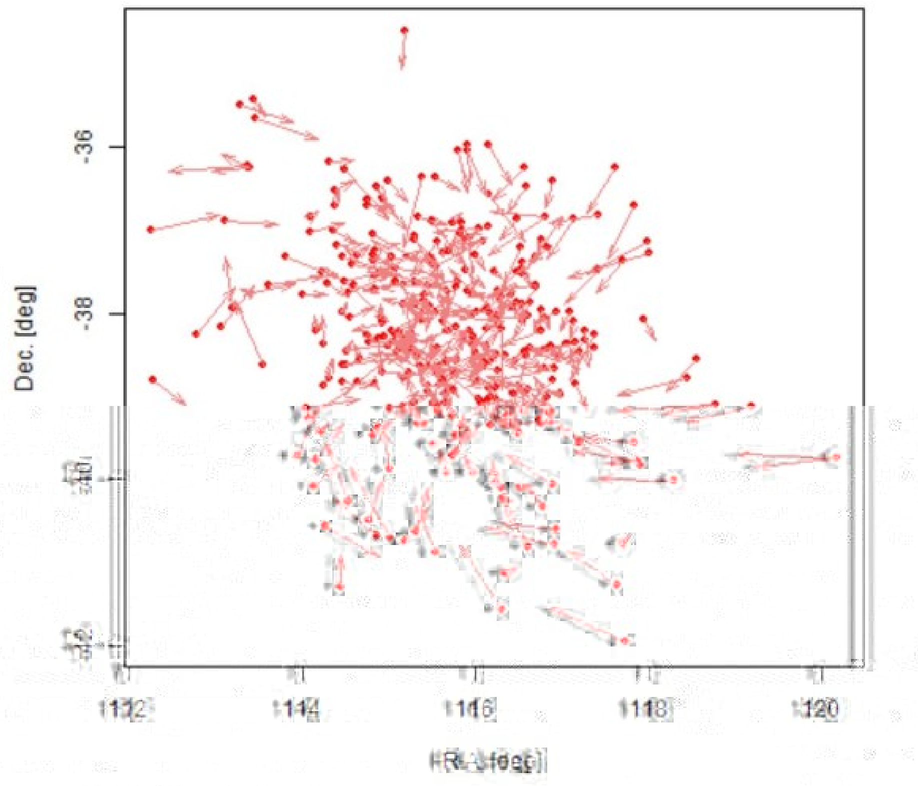
IC 1396, real data

(source, Garcia et al. (2020))

Applications in Astronomy



NGC 2451A, inferred data

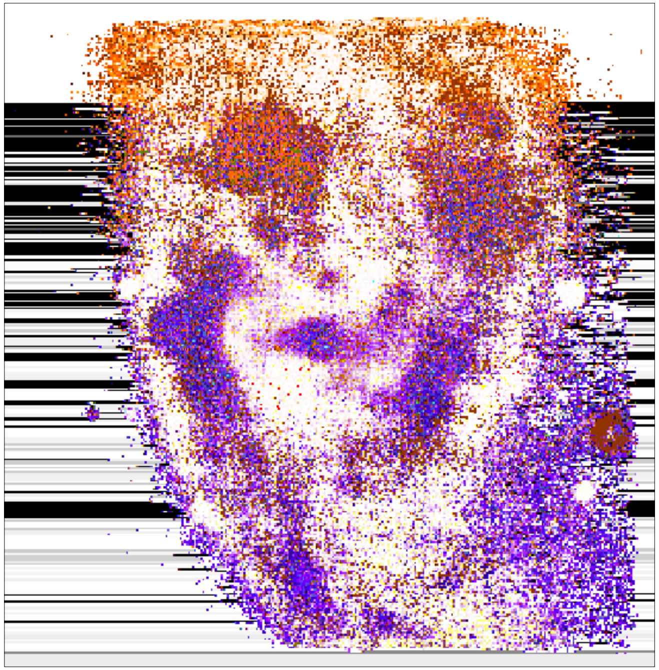


NGC 2451A, real data

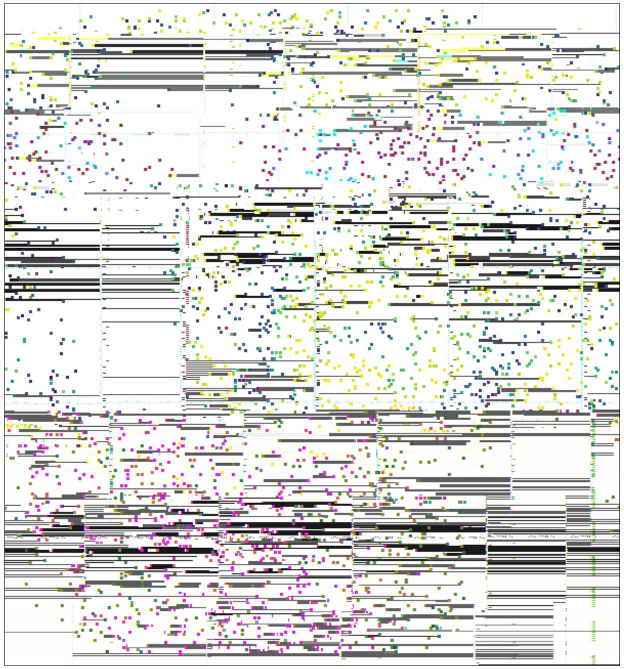
(source, Garcia et al. (2020))

INLA

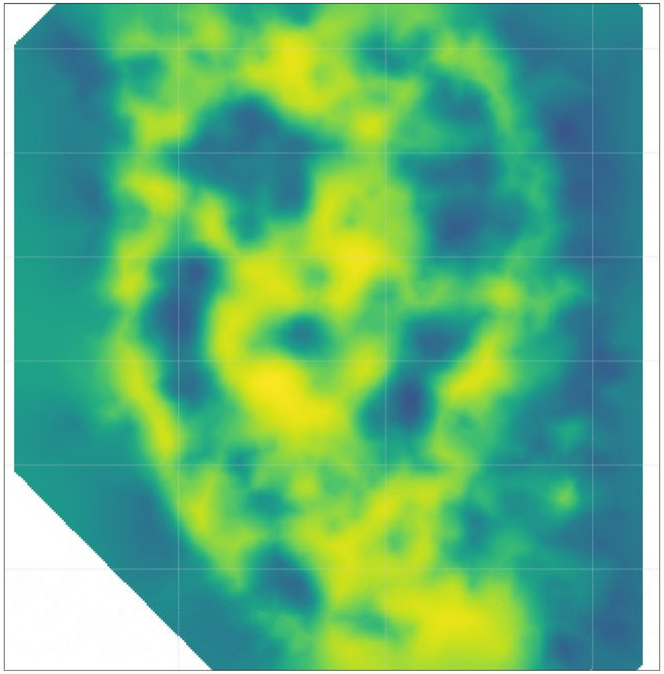
Applications in Astronomy



ASASSN15db_agel
real data

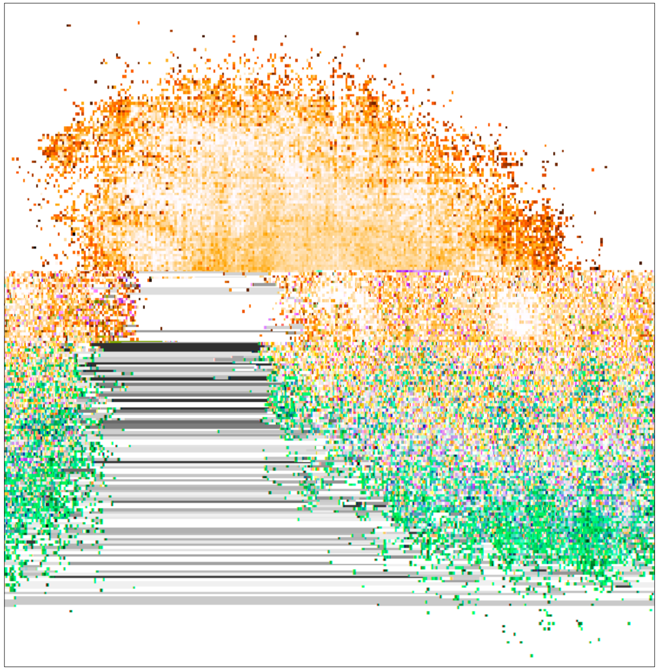


ASASSN15db_agel
5% sampling of real data

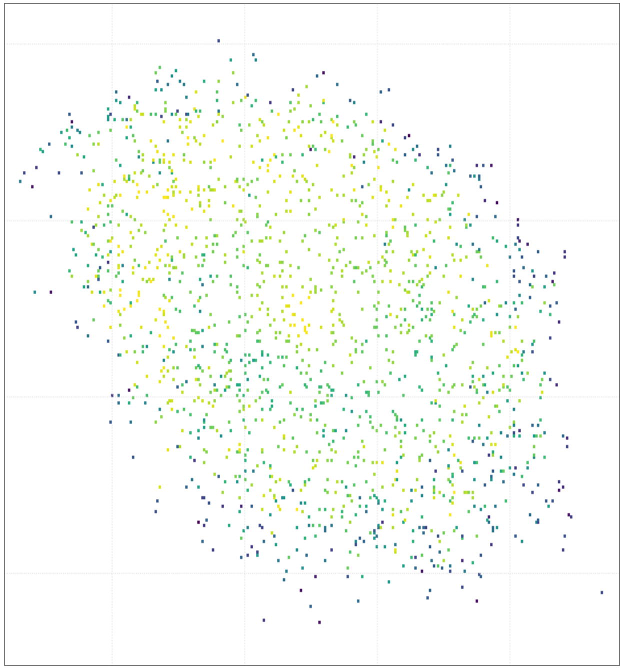


ASASSN15db_agel
INLA reconstruction

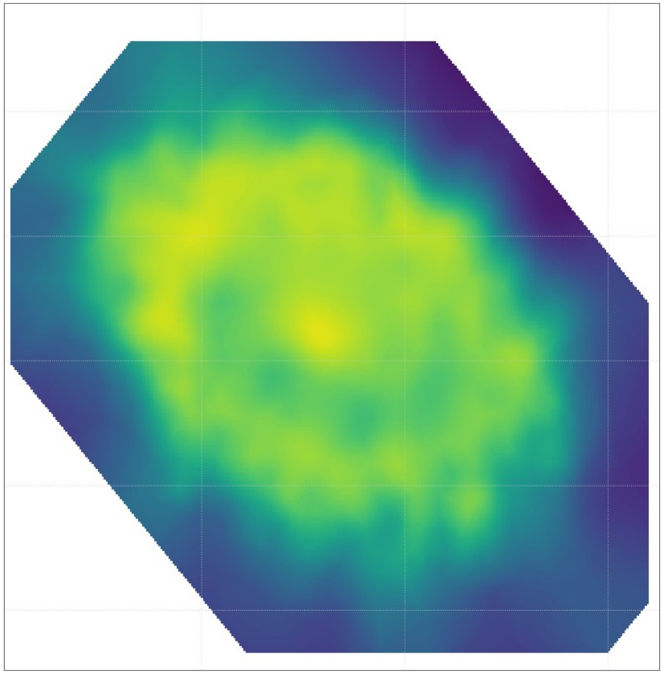
Applications in Astronomy



PTF11qnr_agel
real data



PTF11qnr_agel
5% sampling of real data



PTF11qnr_agel
INLA reconstruction

Present

- **INLA + Monte Carlo Radiative Transfer (MCRT)**

1) Generate low resolution simulations of radiative transfer using MC
2) Preprocess output files
3) Feed results as priors to INLA
4) Get high resolution posteriors in a fraction of the time

Thank you