Quantifying the uncertainty of contour maps

David Bolin University of Gothenburg

joint work with Finn Lindgren

Banff, July 12, 2017



References and the connection to Peter



Seamocs workshop, Malta 2009

 B. and Lindgren: Excursion and contour uncertainty regions for latent Gaussian models, JRRS Series B (2015): 77(1):85-106.
Acknowledgements: The authors are grateful to ... Peter Guttorp for highlighting the need for a thorough treatment of the subject.

References and the connection to Peter

Email from Peter in August 2013: I realised that one might be able to construct simultaneous confidence bands for the uncertainty in sea level predictions using David's method. What do you think?

References and the connection to Peter

Email from Peter in August 2013: I realised that one might be able to construct simultaneous confidence bands for the uncertainty in sea level predictions using David's method. What do you think?

This resulted in two joint papers with Peter:

- Guttorp et al: Assessing the uncertainty in projecting local mean sea level from global temperature, Journal of Applied Meteorology and Climatology, 53, 2163-2170 (2014)
- B. et al: *Statistical prediction of global sea level from global temperature*, Statistica Sinica, 25, 351-367 (2015)

References and the connection to Peter

Email from Peter in August 2013: I realised that one might be able to construct simultaneous confidence bands for the uncertainty in sea level predictions using David's method. What do you think?

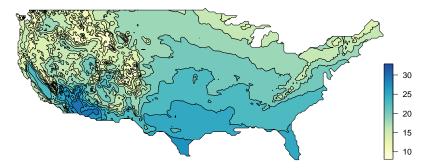
This resulted in two joint papers with Peter:

- Guttorp et al: Assessing the uncertainty in projecting local mean sea level from global temperature, Journal of Applied Meteorology and Climatology, 53, 2163-2170 (2014)
- B. et al: *Statistical prediction of global sea level from global temperature*, Statistica Sinica, 25, 351-367 (2015)

Contour maps

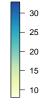
 B. and Lindgren: Quantifying the uncertainty of contour maps, J of Computational and Graphical Statistics (2016).
Acknowledgements: The authors ... wish to acknowledge the importance of Prof Peter Guttorp, who has been a strong champion of the topic.

Contour map of US summer mean temperature



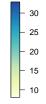
• Can we trust the apparent detail of the level crossings?

Contour map of US summer mean temperature



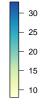
- Can we trust the apparent detail of the level crossings?
- How many contours should we use?

Contour map of US summer mean temperature



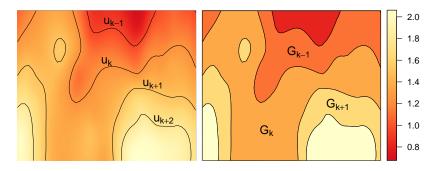
- Can we trust the apparent detail of the level crossings?
- How many contours should we use?
- Can we put a number on the statistical quality of the contour map?

Contour map of US summer mean temperature



- Can we trust the apparent detail of the level crossings?
- How many contours should we use?
- Can we put a number on the statistical quality of the contour map?
- Fundamental question: What *is* the statistical interpretation of a contour map?

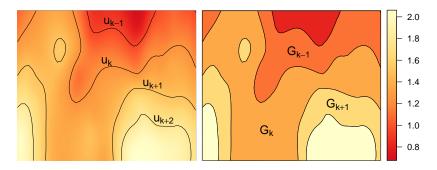
Contour maps



For a function f, a contour map $C_f(u_1,\ldots,u_K)$ with K contour levels $u_1 < u_2 < \ldots < u_K$ is the collection of

• The contour curves, and the

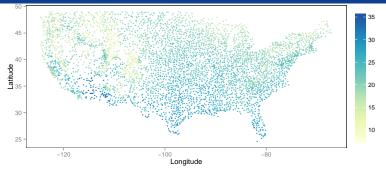
Contour maps



For a function f, a contour map $C_f(u_1,\ldots,u_K)$ with K contour levels $u_1 < u_2 < \ldots < u_K$ is the collection of

- The contour curves, and the
- Level sets $G_k = \{s : u_k < f(s) < u_{k+1}\}.$

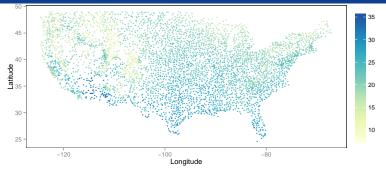
Latent Gaussian models



$$\begin{split} \xi(\mathbf{s}) &\sim \text{Gaussian random field} \\ x(\mathbf{s}) &= \mathbf{z}(\mathbf{s})\boldsymbol{\beta} + \xi(\mathbf{s}) \\ y_i | x(\cdot) &\sim \pi(y_i | x(\cdot), \boldsymbol{\theta}), \quad \text{e.g. } \mathsf{N}(x(\mathbf{s}_i), \sigma^2) \end{split}$$

where $\mathbf{z}(\cdot)$ are explanatory variables and y_i are observations.

Latent Gaussian models

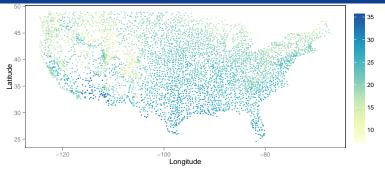


$$\begin{split} \xi(\mathbf{s}) &\sim \text{Gaussian random field} \\ x(\mathbf{s}) &= \mathbf{z}(\mathbf{s})\boldsymbol{\beta} + \xi(\mathbf{s}) \\ y_i | x(\cdot) &\sim \pi(y_i | x(\cdot), \boldsymbol{\theta}), \quad \text{e.g. } \mathsf{N}(x(\mathbf{s}_i), \sigma^2) \end{split}$$

where $\mathbf{z}(\cdot)$ are explanatory variables and y_i are observations.

• A contour map is often reported for $\hat{x}(\mathbf{s}) = \mathsf{E}(x(\mathbf{s})|\mathbf{y})$.

Latent Gaussian models

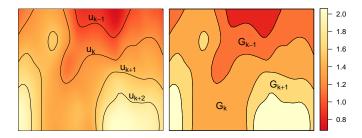


$$\begin{split} \xi(\mathbf{s}) &\sim \text{Gaussian random field} \\ x(\mathbf{s}) &= \mathbf{z}(\mathbf{s})\boldsymbol{\beta} + \xi(\mathbf{s}) \\ y_i | x(\cdot) &\sim \pi(y_i | x(\cdot), \boldsymbol{\theta}), \quad \text{e.g. } \mathsf{N}(x(\mathbf{s}_i), \sigma^2) \end{split}$$

where $\mathbf{z}(\cdot)$ are explanatory variables and y_i are observations.

- A contour map is often reported for $\hat{x}(\mathbf{s}) = \mathsf{E}(x(\mathbf{s})|\mathbf{y})$.
- We interpret the contour map as being informative about x itself.

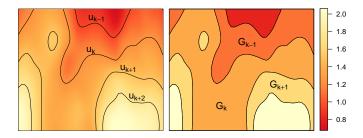
Interpreting a contour map $C_{\hat{x}}(\mathbf{u})$



• Intuitively, one might interpret a contour map as having

 $\mathsf{P}(u_k < x(\mathbf{s}) < u_{k+1}, \text{for all } \mathbf{s} \in G_k \text{ and all } k) pprox 1$

Interpreting a contour map $C_{\hat{x}}(\mathbf{u})$

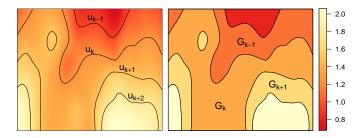


Intuitively, one might interpret a contour map as having

 $\mathsf{P}(u_k < x(\mathbf{s}) < u_{k+1}, \text{for all } \mathbf{s} \in G_k \text{ and all } k) \approx 1$

• This probability will nearly always be close to or equal to zero!

Interpreting a contour map $C_{\hat{x}}(\mathbf{u})$



Intuitively, one might interpret a contour map as having

 $\mathsf{P}(u_k < x(\mathbf{s}) < u_{k+1}, \text{for all } \mathbf{s} \in G_k \text{ and all } k) \approx 1$

- This probability will nearly always be close to or equal to zero!
- Polfeldt (1999), *On the quality of contour maps*, Environmetrics, instead considered the marginal probabilities

 $p(\mathbf{s}) = \mathsf{P}(u_k < x(\mathbf{s}) < u_{k+1}, \text{for } k \text{ such that } \mathbf{s} \in G_k)$

and argued that if p(s) is close to 1 in a large proportion of the region, the contour map is not overconfident.

Contour avoiding sets

We construct an alternative to p(s) that has a joint probability statement.

Contour avoiding sets

We construct an alternative to p(s) that has a joint probability statement.

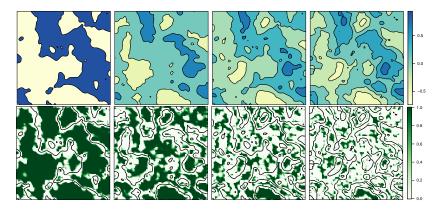
Contour avoiding sets

Let $A_k = \{s : u_k < x(s) < u_{k+1}\}$. The joint $\mathbf{u} = (u_1, \dots, u_K)$ contour avoiding set is $C_{\mathbf{u},\alpha}(x) = \bigcup_k M_{u_k,\alpha}$, where $M_{\mathbf{u},\alpha} = (M_{u_1,\alpha}, \dots, M_{u_K,\alpha})$ is given by

$$M_{\mathbf{u},\alpha} = \underset{(D_1,\dots,D_K)}{\operatorname{arg\,max}} \left\{ \sum_{k=1}^K |D_k| : D_k \subseteq G_k, \mathsf{P}\left(\bigcap_k \{D_k \subseteq A_k\}\right) > 1 - \alpha \right\},$$

The contour avoiding set is the largest set so that, with probability $1 - \alpha$, the intuitive contour map interpretation holds for $\mathbf{s} \in C_{\mathbf{u},\alpha}(X)$.

The contour map function



Given $C_{\bar{u},\alpha}(X)$ we define the *contour map function*

$$F_u(s) = \sup\{1 - \alpha; s \in C_{\bar{u},\alpha}\},\$$

as a joint probability extension of the Polfeldt idea.

Contour maps

Contour map quality measures

Recall that we would like to put a number on the statistical quality of the contour map, we call such a number a contour map quality measure.

Contour map quality measures

Recall that we would like to put a number on the statistical quality of the contour map, we call such a number a contour map quality measure.

The P_0 quality measure

Given the contour map function, a simple contour map quality measure, $P_{\rm 0}$ is given by

$$P_0(x, C_f) = \frac{1}{|\Omega|} \int_{\Omega} F_{\mathbf{u}}(\mathbf{s}) \, \mathrm{d}\mathbf{s}.$$

Loosely speaking, this is the percentage of the total area for which the intuitive interpretation of the contour map holds.

Contour map quality measures

Recall that we would like to put a number on the statistical quality of the contour map, we call such a number a contour map quality measure.

The P_0 quality measure

Given the contour map function, a simple contour map quality measure, $P_{\rm 0}$ is given by

$$P_0(x, C_f) = \frac{1}{|\Omega|} \int_{\Omega} F_{\mathbf{u}}(\mathbf{s}) \,\mathrm{d}\mathbf{s}.$$

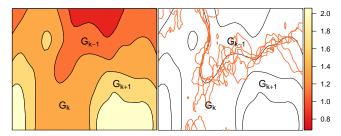
Loosely speaking, this is the percentage of the total area for which the intuitive interpretation of the contour map holds.

For the contour maps in the example above, we have

 $P_0 = 0.613, \quad P_0 = 0.440, \quad P_0 = 0.394, \quad P_0 = 0.148$

The "intuitive" interpretation is not the only global interpretation of a contour map!

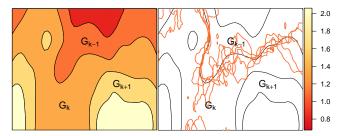
The P_1 quality measure



Five realisations of u_k contour curves from the posterior for x.

 Another natural interpretation of C_{x̂} is that G_{k-1} ∪ G_k should contain all level u_k crossings of the process x.

The P_1 quality measure

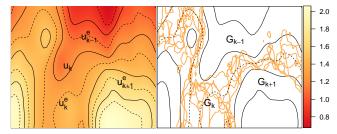


Five realisations of u_k contour curves from the posterior for x.

- Another natural interpretation of $C_{\hat{x}}$ is that $G_{k-1} \cup G_k$ should contain all level u_k crossings of the process x.
- Define P_1 as the probability for this occurring:

$$P_1(X, C_f(u_1, \dots, u_K)) = \mathsf{P}\left(\bigcap_{k=0}^K \{u_{k-1} < x(s) < u_{k+2}, s \in G_k\}\right)$$

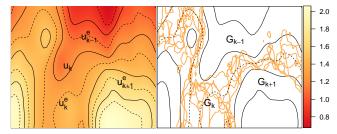
The P_2 quality measure



Five realisations of u_k^e contour curves from the posterior for x.

• When drawing contour maps, the set G_k is associated with the level $u_k^e = (u_{k+1} + u_k)/2$.

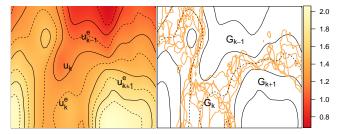
The P_2 quality measure



Five realisations of u_k^e contour curves from the posterior for x.

- When drawing contour maps, the set G_k is associated with the level $u_k^e = (u_{k+1} + u_k)/2$.
- A third natural interpretation of C_{x̂} is that G_k should contain all level u^e_k crossings of the process x.

The P_2 quality measure

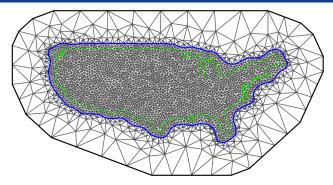


Five realisations of u_k^e contour curves from the posterior for x.

- When drawing contour maps, the set G_k is associated with the level $u_k^e = (u_{k+1} + u_k)/2$.
- A third natural interpretation of C_{x̂} is that G_k should contain all level u^e_k crossings of the process x.
- Define P_2 as the probability for this occurring:

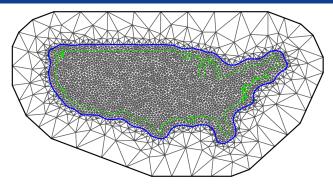
$$P_2(X, C_f(u_1, \dots, u_K)) = \mathsf{P}\left(\bigcap_{k=0}^K \{u_{k-1}^e < x(s) < u_{k+1}^e, s \in G_k\}\right)$$

Computational details



- To compute the quality measures, high dimensional joint posterior probabilities need to be evaluated.
- We consider the situation where the random field can be discretised with weights for piecewise linear local basis functions.
 - Common contour plotting methods are based on variations of such linear interpolation, e.g. contour in R and Matlab.

Computational details



- To compute the quality measures, high dimensional joint posterior probabilities need to be evaluated.
- We consider the situation where the random field can be discretised with weights for piecewise linear local basis functions.
 - Common contour plotting methods are based on variations of such linear interpolation, e.g. contour in R and Matlab.
 - SPDE-based spatial models satisfy this by construction.

• Let x denote the discretisation of $x(\mathbf{s})$, y the data, and $\boldsymbol{\theta}$ the model parameters.

Posterior probabilities

- Let x denote the discretisation of x(s), y the data, and θ the model parameters.
- Assuming that $\pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})$ is Gaussian, the marginal posterior probability $\mathsf{P}(\mathbf{a} < \mathbf{x} < \mathbf{b}|\mathbf{y}, \boldsymbol{\theta})$ is a Gaussian integral.

- Let x denote the discretisation of $x(\mathbf{s})$, y the data, and $\boldsymbol{\theta}$ the model parameters.
- Assuming that $\pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})$ is Gaussian, the marginal posterior probability $\mathsf{P}(\mathbf{a} < \mathbf{x} < \mathbf{b}|\mathbf{y}, \boldsymbol{\theta})$ is a Gaussian integral.
- Quantile corrections can be used if $\pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})$ is not Gaussian.

- Let x denote the discretisation of x(s), y the data, and θ the model parameters.
- Assuming that $\pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})$ is Gaussian, the marginal posterior probability $\mathsf{P}(\mathbf{a} < \mathbf{x} < \mathbf{b}|\mathbf{y}, \boldsymbol{\theta})$ is a Gaussian integral.
- Quantile corrections can be used if $\pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})$ is not Gaussian.
- In a Bayesian setting, one can numerically approximate the posterior probability as

$$\mathsf{P}(\mathbf{a} < \mathbf{x} < \mathbf{b} | \mathbf{y}) = \mathsf{E}(\mathsf{P}(\mathbf{a} < \mathbf{x} < \mathbf{b} | \mathbf{y}, \boldsymbol{\theta})) \approx \sum_{k} w_k \mathsf{P}(\mathbf{a} < \mathbf{x} < \mathbf{b} | \mathbf{y}, \boldsymbol{\theta}_k)$$

where $w_k \propto \pi(\theta_k | \mathbf{y})$ and θ_k are cleverly chosen parameter configurations (for example as done in INLA).

- Let x denote the discretisation of x(s), y the data, and θ the model parameters.
- Assuming that $\pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})$ is Gaussian, the marginal posterior probability $\mathsf{P}(\mathbf{a} < \mathbf{x} < \mathbf{b}|\mathbf{y}, \boldsymbol{\theta})$ is a Gaussian integral.
- Quantile corrections can be used if $\pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})$ is not Gaussian.
- In a Bayesian setting, one can numerically approximate the posterior probability as

$$\mathsf{P}(\mathbf{a} < \mathbf{x} < \mathbf{b} | \mathbf{y}) = \mathsf{E}(\mathsf{P}(\mathbf{a} < \mathbf{x} < \mathbf{b} | \mathbf{y}, \boldsymbol{\theta})) \approx \sum_{k} w_k \mathsf{P}(\mathbf{a} < \mathbf{x} < \mathbf{b} | \mathbf{y}, \boldsymbol{\theta}_k)$$

where $w_k \propto \pi(\theta_k | \mathbf{y})$ and θ_k are cleverly chosen parameter configurations (for example as done in INLA).

• Often only a few configurations are needed for accurate results.

Gaussian integrals

Computing the Gaussian probability is done by computing an integral

$$\mathbf{I} = \frac{|\mathbf{Q}|^{1/2}}{(2\pi)^{d/2}} \int_{\mathbf{a} \le \mathbf{x} \le \mathbf{b}} \exp(-\frac{1}{2}\mathbf{x}^{\top}\mathbf{Q}\mathbf{x}) \,\mathrm{d}\mathbf{x},$$

• There are several methods for doing this (see e.g. Genz and Bretz (2009), Computation of Multivariate Normal and t Probabilities, Lecture Notes in Statistics, Springer).

Gaussian integrals

Computing the Gaussian probability is done by computing an integral

$$\mathbf{I} = \frac{|\mathbf{Q}|^{1/2}}{(2\pi)^{d/2}} \int_{\mathbf{a} \le \mathbf{x} \le \mathbf{b}} \exp(-\frac{1}{2}\mathbf{x}^{\top}\mathbf{Q}\mathbf{x}) \,\mathrm{d}\mathbf{x},$$

- There are several methods for doing this (see e.g. Genz and Bretz (2009), Computation of Multivariate Normal and t Probabilities, Lecture Notes in Statistics, Springer).
- For GMRFs, we want to use the sparsity of Q.

Gaussian integrals

Computing the Gaussian probability is done by computing an integral

$$\mathbf{I} = \frac{|\mathbf{Q}|^{1/2}}{(2\pi)^{d/2}} \int_{\mathbf{a} \le \mathbf{x} \le \mathbf{b}} \exp(-\frac{1}{2}\mathbf{x}^{\top}\mathbf{Q}\mathbf{x}) \,\mathrm{d}\mathbf{x},$$

- There are several methods for doing this (see e.g. Genz and Bretz (2009), Computation of Multivariate Normal and t Probabilities, Lecture Notes in Statistics, Springer).
- For GMRFs, we want to use the sparsity of \mathbf{Q} .
- We use a method based on sequential importance sampling.

Gaussian integrals

Computing the Gaussian probability is done by computing an integral

$$\mathbf{I} = \frac{|\mathbf{Q}|^{1/2}}{(2\pi)^{d/2}} \int_{\mathbf{a} \le \mathbf{x} \le \mathbf{b}} \exp(-\frac{1}{2} \mathbf{x}^\top \mathbf{Q} \mathbf{x}) \, \mathrm{d} \mathbf{x},$$

- There are several methods for doing this (see e.g. Genz and Bretz (2009), Computation of Multivariate Normal and t Probabilities, Lecture Notes in Statistics, Springer).
- For GMRFs, we want to use the sparsity of \mathbf{Q} .
- We use a method based on sequential importance sampling.
 - It is based on that a GMRF can be viewed as a non-homogeneous AR-process defined backwards in the indices of **x**:

$$x_i | x_{i+1}, \dots, x_n \sim \mathsf{N}\left(\mu_i - \frac{1}{L_{ii}} \sum_{j=i+1}^n L_{ji}(x_j - \mu_j), L_{ii}^{-2}\right),$$

where \mathbf{L} is the Cholesky factor of \mathbf{Q} .

The excursions package

The methods are implemented in the R package excursions

- Available on CRAN.
- Manual: B. and Lindgren: *Calculating probabilistic excursion sets and related quantities using excursions*, ArXiv preprint (2017):1612.04101.

The excursions package

The methods are implemented in the R package excursions

- Available on CRAN.
- Manual: B. and Lindgren: *Calculating probabilistic excursion sets and related quantities using excursions*, ArXiv preprint (2017):1612.04101.

Important functions in the package:

- excursions: Compute uncertainty regions for individual contour curves, excursion sets, and excursion functions.
- contourmap: Compute contour maps, quality measures, and contour map functions.
- simconf: Compute simultaneous credible bands.

The excursions package

The methods are implemented in the R package excursions

- Available on CRAN.
- Manual: B. and Lindgren: *Calculating probabilistic excursion sets and related quantities using excursions*, ArXiv preprint (2017):1612.04101.

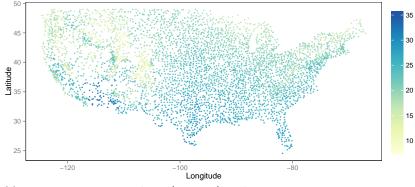
Important functions in the package:

- excursions: Compute uncertainty regions for individual contour curves, excursion sets, and excursion functions.
- contourmap: Compute contour maps, quality measures, and contour map functions.
- simconf: Compute simultaneous credible bands.

Specialized versions of the functions:

- Interface to R-INLA: e.g. excursions.inla
- Functions to analyze MCMC output: e.g. excursions.mc

Back to the US temperatures

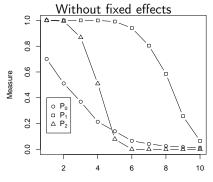


Measurements at approximately 8000 locations.

Estimate the true temperature surface using the model:

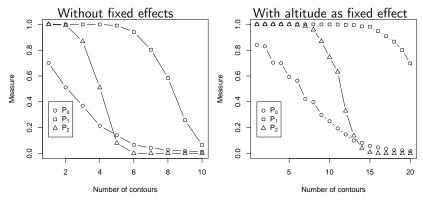
- Likelihood: $Y_i \sim \mathsf{N}(X(s_i), \sigma^2)$.
- Latent temperature model: x(s) is a Gaussian Matérn field.

Contour map quality measures



Number of contours

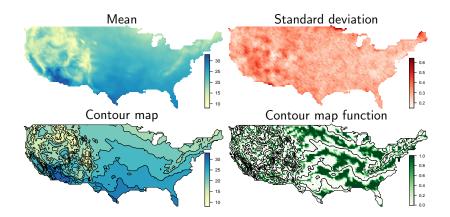
Contour map quality measures



The spatial predictions are more uncertain in a model without spatial explanatory variables (left) than in a model using elevation (right).

Without explanatory variables, use 3 contours. With elevation, use 10.

Results



With 10 contour levels the contour map above has $P_2 \approx 0.95$.

Alternative methods for uncertainty visualization



Is a contour map the best way of visualizing the uncertainty?

Alternative methods for uncertainty visualization



Is a contour map the best way of visualizing the uncertainty?

Methods for drawing contours with uncertainty:

• Add credible band to the contours.

Alternative methods for uncertainty visualization



Is a contour map the best way of visualizing the uncertainty?

- Add credible band to the contours.
 - G. Lindgren and I. Rychlik (1995): *How reliable are contour curves? Confidence sets for level contours*, Bernoulli

Alternative methods for uncertainty visualization



Is a contour map the best way of visualizing the uncertainty?

- Add credible band to the contours.
 - G. Lindgren and I. Rychlik (1995): *How reliable are contour curves? Confidence sets for level contours*, Bernoulli
 - We can use excursions to compute joint credible bands.

Alternative methods for uncertainty visualization



Is a contour map the best way of visualizing the uncertainty?

- Add credible band to the contours.
 - G. Lindgren and I. Rychlik (1995): *How reliable are contour curves? Confidence sets for level contours*, Bernoulli
 - We can use excursions to compute joint credible bands.
- Draw "Fuzzy" contours.

Alternative methods for uncertainty visualization



Is a contour map the best way of visualizing the uncertainty?

- Add credible band to the contours.
 - G. Lindgren and I. Rychlik (1995): *How reliable are contour curves? Confidence sets for level contours*, Bernoulli
 - We can use excursions to compute joint credible bands.
- Draw "Fuzzy" contours.
- Show "spaghetti plots".

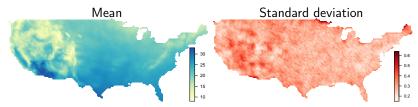
Alternative methods for uncertainty visualization



Is a contour map the best way of visualizing the uncertainty?

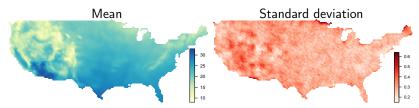
- Add credible band to the contours.
 - G. Lindgren and I. Rychlik (1995): *How reliable are contour curves? Confidence sets for level contours*, Bernoulli
 - We can use excursions to compute joint credible bands.
- Draw "Fuzzy" contours.
- Show "spaghetti plots".
- +++

Alternative methods for uncertainty visualization



Perhaps we should not use contour maps at all to visualize uncertainty?

Alternative methods for uncertainty visualization

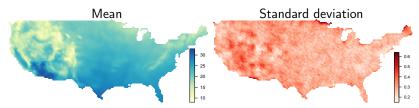


Perhaps we should not use contour maps at all to visualize uncertainty?

There are many other ways to visualize the uncertainty:

- Two maps: One of a point estimate and one of posterior standard deviations.
- Point estimate with opacity given by standard deviations.
- +++

Alternative methods for uncertainty visualization



Perhaps we should not use contour maps at all to visualize uncertainty?

There are many other ways to visualize the uncertainty:

- Two maps: One of a point estimate and one of posterior standard deviations.
- Point estimate with opacity given by standard deviations.
- +++

Is there a best alternative?

Questions for the breakout session

Join the discussion in Breakout session G tomorrow!

- What is the best way of visualizing estimates of spatial fields and their uncertainties?
- Ø How should one do visualization for more complicated scenarios:
 - problems in three spatial dimensions
 - spatio-temporal applications
 - hierarchical models
- S Visualization for model validation.
- Visualization for data uncertainty.

Questions for the breakout session

Join the discussion in Breakout session G tomorrow!

- What is the best way of visualizing estimates of spatial fields and their uncertainties?
- Ø How should one do visualization for more complicated scenarios:
 - problems in three spatial dimensions
 - spatio-temporal applications
 - hierarchical models
- S Visualization for model validation.
- Visualization for data uncertainty.

Thanks for your attention!