Le SLGP

ou comment utiliser les champs de densité pour modéliser les phénomènes complexes

Athénaïs Gautier*

athenais.gautier@stat.unibe.ch

*Université de Bern

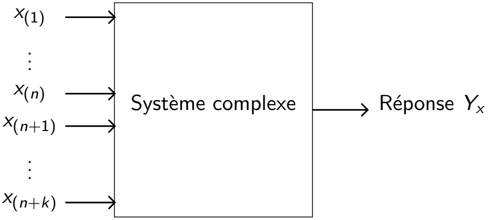
PhD Supervisor: Prof. David Ginsbourger

Swiss National Science Foundation project number 178858

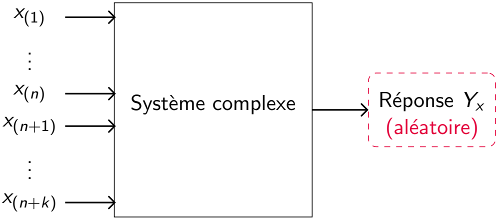
Rencontres chercheur·euse·s et ingénieur·e·s Phimeca

18 Novembre 2021

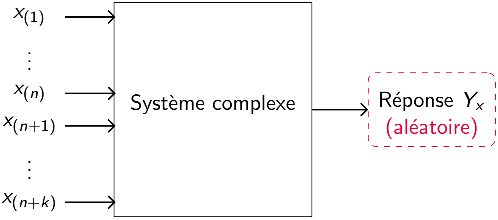
Motivation: modéliser les systèmes complexes



Motivation: modéliser les systèmes complexes

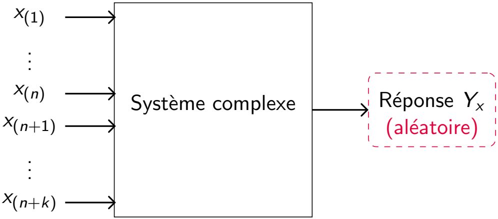


Motivation: modéliser les systèmes complexes



Causes possibles:

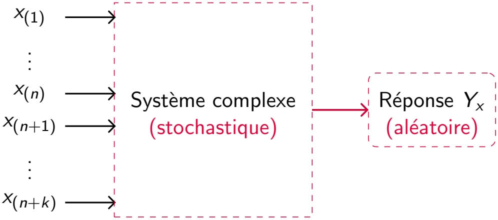
Motivation: modéliser les systèmes complexes



Causes possibles:

- Bruit d'observation sur Y_x

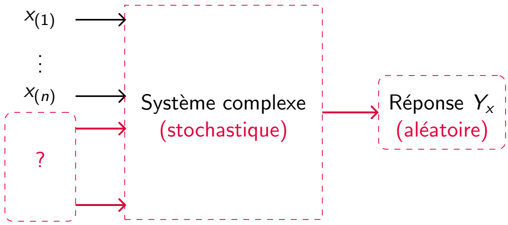
Motivation: modéliser les systèmes complexes



Causes possibles:

- Bruit d'observation sur Y_x
- Système stochastique par nature

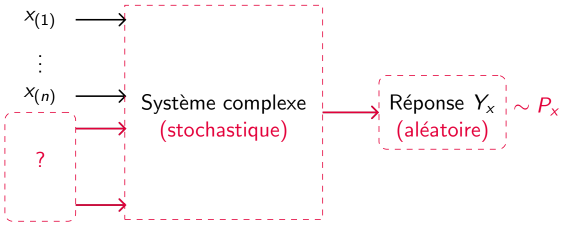
Motivation: modéliser les systèmes complexes



Causes possibles:

- Bruit d'observation sur Y_x
- Système stochastique par nature
- Incertitude sur certains paramètres

Motivation: modéliser les systèmes complexes



Causes possibles:

- Bruit d'observation sur Y_x
- Système stochastique par nature
- Incertitude sur certains paramètres

Motivation: estimer des champs de densité de probabilité

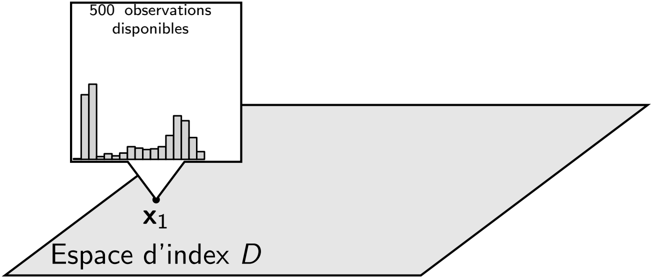


Figure: Distribution sous-jacente (courbe rouge) versus données observées (histogramme)

Motivation: estimer des champs de densité de probabilité

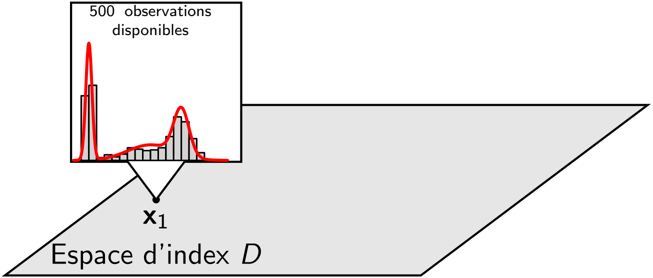


Figure: Distribution sous-jacente (courbe rouge) versus données observées (histogramme)

Motivation: estimer des champs de densité de probabilité

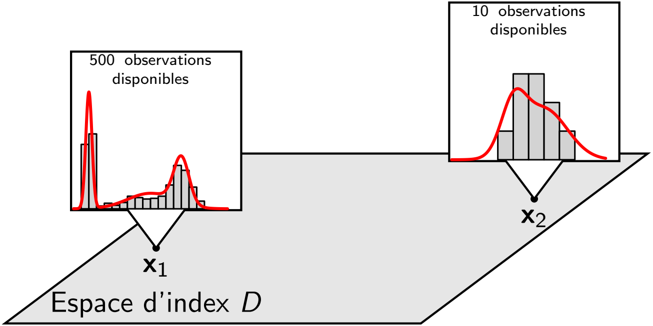


Figure: Distribution sous-jacente (courbe rouge) versus données observées (histogramme)

Motivation: estimer des champs de densité de probabilité

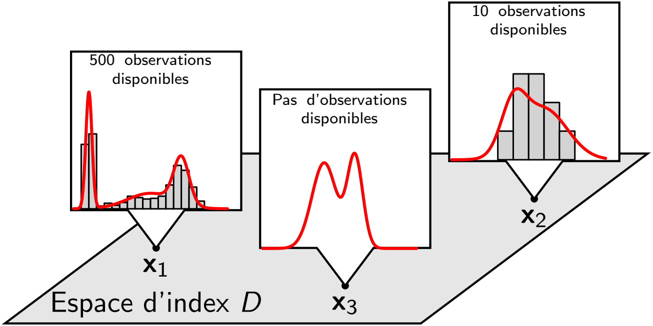


Figure: Distribution sous-jacente (courbe rouge) versus données observées (histogramme)

Motivation: estimer des champs de densité de probabilité

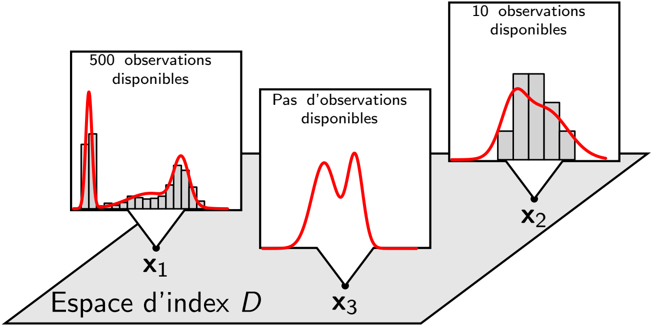


Figure: Distribution sous-jacente (courbe rouge) versus données observées (histogramme)

Notre but: Modéliser le champ de distribution $P_{\mathbf{x}}$ indexé par $\mathbf{x}$, en se basant sur des échantillons localisés.

Quelques idées

Objectif: Estimer des distributions indexées par $\mathbf{x}$

[1] Zhu and Sudret, *Emulation of stochastic simulators using generalized lambda models*.

Quelques idées

Objectif: Estimer des distributions indexées par $\mathbf{x}$

- **Modèle paramétrique:** $P_{\mathbf{x}}$ vient d'une famille paramétrique donnée, on cherche les "meilleurs" paramètres.[1]

[1] Zhu and Sudret, *Emulation of stochastic simulators using generalized lambda models*.

Quelques idées

Objectif: Estimer des distributions indexées par $\mathbf{x}$

- **Modèle paramétrique:** $P_{\mathbf{x}}$ vient d'une famille paramétrique donnée, on cherche les "meilleurs" paramètres.[1]
- **Régression distributionelle:** Pour chaque $\mathbf{x}_i$ séparément on estime la densité locale, puis on interpole les courbes.

[1] Zhu and Sudret, *Emulation of stochastic simulators using generalized lambda models*.

Les difficultés

Objectif: Estimer des distributions indexées par $\mathbf{x}$

Les difficultés

Objectif: Estimer des distributions indexées par $\mathbf{x}$

- Pour des variations de pdf en forme et modalités

Les difficultés

Objectif: Estimer des distributions indexées par $\mathbf{x}$

- Pour des variations de pdf en forme et modalités
- Avec des tailles d'échantillon modérées

Les difficultés

Objectif: Estimer des distributions indexées par $\mathbf{x}$

- Pour des variations de pdf en forme et modalités
- Avec des tailles d'échantillon modérées
- Avec peu de réplications

Les difficultés

Objectif: Estimer des distributions indexées par $\mathbf{x}$

- Pour des variations de pdf en forme et modalités
- Avec des tailles d'échantillon modérées
- Avec peu de réplications
- Avec quantification de l'incertitude

Les difficultés

Objectif: Estimer des distributions indexées par $\mathbf{x}$

- Pour des variations de pdf en forme et modalités
- Avec des tailles d'échantillon modérées
- Avec peu de réplications
- Avec quantification de l'incertitude

Notre approche: Utiliser un modèle bayésien non paramétrique

1 Motivation: l'estimation de champs de densité de probabilité

2 Le SLGP

3 Quelques applications
- Optimisation stochastique
- Problèmes inverses stochastiques

L'idée du modèle

Extension spatiale du LGP[2]

[2]Lenk, "Towards a Practicable Bayesian Nonparametric Density Estimator"; Tokdar and K. Ghosh, "Posterior consistency of logistic Gaussian process priors in density estimation".

L'idée du modèle

Extension spatiale du LGP[2]

1. On considère un processus Gaussien $Z_{\mathbf{x},y}$ (rôle *latent*).

[2] Lenk, "Towards a Practicable Bayesian Nonparametric Density Estimator"; Tokdar and K. Ghosh, "Posterior consistency of logistic Gaussian process priors in density estimation".

L'idée du modèle

Extension spatiale du LGP[2]

1. On considère un processus Gaussien $Z_{\mathbf{x},y}$ (rôle *latent*).
   $\rightarrow$ fonctions aléatoires de $\mathbf{x}$ et y.

[2] Lenk, "Towards a Practicable Bayesian Nonparametric Density Estimator"; Tokdar and K. Ghosh, "Posterior consistency of logistic Gaussian process priors in density estimation".

L'idée du modèle

Extension spatiale du LGP[2]

1. On considère un processus Gaussien $Z_{\mathbf{x},y}$ (rôle *latent*).
   $\rightarrow$ fonctions aléatoires de y, indexées par $\mathbf{x}$.

[2] Lenk, "Towards a Practicable Bayesian Nonparametric Density Estimator"; Tokdar and K. Ghosh, "Posterior consistency of logistic Gaussian process priors in density estimation".

L'idée du modèle

Extension spatiale du LGP[2]

1. On considère un processus Gaussien $Z_{\mathbf{x},y}$ (rôle *latent*).
   $\rightarrow$ fonctions aléatoires de y, indexées par $\mathbf{x}$.
2. On le rend positif en considérant $e^{Z_{\mathbf{x},y}}$

[2] Lenk, "Towards a Practicable Bayesian Nonparametric Density Estimator"; Tokdar and K. Ghosh, "Posterior consistency of logistic Gaussian process priors in density estimation".

L'idée du modèle

Extension spatiale du LGP[2]

1. On considère un processus Gaussien $Z_{\mathbf{x},y}$ (rôle *latent*).
   $\rightarrow$ fonctions aléatoires de y, indexées par $\mathbf{x}$.
2. On le rend positif en considérant $e^{Z_{\mathbf{x},y}}$
   $\rightarrow$ fonctions aléatoires positives de y, indexées par $\mathbf{x}$.

[2] Lenk, "Towards a Practicable Bayesian Nonparametric Density Estimator"; Tokdar and K. Ghosh, "Posterior consistency of logistic Gaussian process priors in density estimation".

L'idée du modèle

Extension spatiale du LGP[2]

1. On considère un processus Gaussien $Z_{\mathbf{x},y}$ (rôle *latent*).
   $\rightarrow$ fonctions aléatoires de y, indexées par $\mathbf{x}$.
2. On le rend positif en considérant $e^{Z_{\mathbf{x},y}}$
   $\rightarrow$ fonctions aléatoires positives de y, indexées par $\mathbf{x}$.
3. On renormalise selon y: $\dfrac{e^{Z_{\mathbf{x},y}}}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}}\, du}$

[2] Lenk, "Towards a Practicable Bayesian Nonparametric Density Estimator"; Tokdar and K. Ghosh, "Posterior consistency of logistic Gaussian process priors in density estimation".

L'idée du modèle

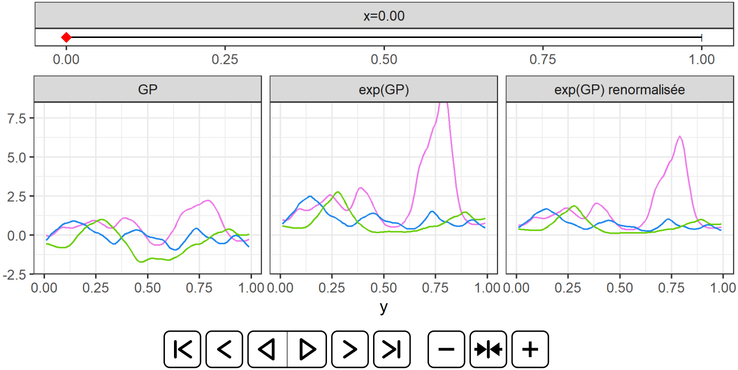
Extension spatiale du LGP[2]

1. On considère un processus Gaussien $Z_{\mathbf{x},y}$ (rôle *latent*).
   $\rightarrow$ fonctions aléatoires de y, indexées par $\mathbf{x}$.
2. On le rend positif en considérant $e^{Z_{\mathbf{x},y}}$
   $\rightarrow$ fonctions aléatoires positives de y, indexées par $\mathbf{x}$.
3. On renormalise selon y: $\dfrac{e^{Z_{\mathbf{x},y}}}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}}\, du}$
   $\rightarrow$ fonctions aléatoires positives de y, s'intégrant à 1 et indexées par $\mathbf{x}$.

[2] Lenk, "Towards a Practicable Bayesian Nonparametric Density Estimator"; Tokdar and K. Ghosh, "Posterior consistency of logistic Gaussian process priors in density estimation".

Le SLGP en images

On regarde 3 *réalisations* d'un SLGP en faisant varier **x**.



[1] Moyenne: 0; Noyau: Matérn 5/2 variance 1, lengthscale 0.1 en **x** et y.

Les principaux aspects du SLGP

Le SLGP: $\dfrac{e^{Z_{\mathbf{x},y}}}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}}\,du}$

Les principaux aspects du SLGP

Le SLGP: $\dfrac{e^{Z_{\mathbf{x},y}}}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}}\, du}$

Avantages:

- modèle flexible, pas limité à une famille paramétrique

Les principaux aspects du SLGP

Le SLGP: $\dfrac{e^{Z_{\mathbf{x},y}}}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}}\,du}$

Avantages:

- modèle flexible, pas limité à une famille paramétrique
- permet d'intégrer des données (même sans réplications)

Les principaux aspects du SLGP

Le SLGP: $\dfrac{e^{Z_{\mathbf{x},y}}}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}}\, du}$

Avantages:

- modèle flexible, pas limité à une famille paramétrique
- permet d'intégrer des données (même sans réplications)
- approche Bayésienne: prédiction probabiliste

Les principaux aspects du SLGP

Le SLGP: $\dfrac{e^{Z_{\mathbf{x},y}}}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}}\,du}$

Avantages:

- modèle flexible, pas limité à une famille paramétrique
- permet d'intégrer des données (même sans réplications)
- approche Bayésienne: prédiction probabiliste
- permet d'encoder des informations via le GP.

Les principaux aspects du SLGP

Le SLGP: $\dfrac{e^{Z_{\mathbf{x},y}}}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}}\,du}$

Avantages:

- modèle flexible, pas limité à une famille paramétrique
- permet d'intégrer des données (même sans réplications)
- approche Bayésienne: prédiction probabiliste
- permet d'encoder des informations via le GP.

Inconvénients:

- computationellement assez lourd

Les principaux aspects du SLGP

Le SLGP: $\dfrac{e^{Z_{\mathbf{x},y}}}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}}\, du}$

Avantages:

- modèle flexible, pas limité à une famille paramétrique
- permet d'intégrer des données (même sans réplications)
- approche Bayésienne: prédiction probabiliste
- permet d'encoder des informations via le GP.

Inconvénients:

- computationellement assez lourd
- peut demander d'estimer des hyper-paramètres.

1 Motivation: l'estimation de champs de densité de probabilité

2 Le SLGP

3 Quelques applications
- Optimisation stochastique
- Problèmes inverses stochastiques

En optimisation stochastique

Objectif: Pour un système (complexe) stochastique dont la sortie $Y_{\mathbf{x}}$ suit une distribution $P_{\mathbf{x}}$ inconnue, on veut optimiser une fonctionelle g:

$$\min_{\mathbf{x} \in D} g(P_{\mathbf{x}})$$

En optimisation stochastique

Objectif: Pour un système (complexe) stochastique dont la sortie $Y_{\mathbf{x}}$ suit une distribution $P_{\mathbf{x}}$ inconnue, on veut optimiser une fonctionelle g:

$$\min_{\mathbf{x}\in D} g(P_{\mathbf{x}})$$

Exemples de fonctionnelles: la moyenne, médiane, quantile donné, variance...

En optimisation stochastique

Objectif: Pour un système (complexe) stochastique dont la sortie $Y_{\mathbf{x}}$ suit une distribution $P_{\mathbf{x}}$ inconnue, on veut optimiser une fonctionelle g:

$$\min_{\mathbf{x} \in D} g(P_{\mathbf{x}})$$

Exemples de fonctionnelles: la moyenne, médiane, quantile donné, variance...

Avec le SLGP:

- On estime $P_{\mathbf{x}}$, et donc $g(P_{\mathbf{x}})$

En optimisation stochastique

Objectif: Pour un système (complexe) stochastique dont la sortie $Y_{\mathbf{x}}$ suit une distribution $P_{\mathbf{x}}$ inconnue, on veut optimiser une fonctionelle g:

$$\min_{\mathbf{x} \in D} g(P_{\mathbf{x}})$$

Exemples de fonctionnelles: la moyenne, médiane, quantile donné, variance...

Avec le SLGP:

- On estime $P_{\mathbf{x}}$, et donc $g(P_{\mathbf{x}})$
- On utilise le SLGP pour guider l'acquisition de données et répondre à la question:
  Où evaluer le système pour gagner de l'information?

En optimisation stochastique

Objectif: Pour un système (complexe) stochastique dont la sortie $Y_{\mathbf{x}}$ suit une distribution $P_{\mathbf{x}}$ inconnue, on veut optimiser une fonctionelle g:

$$\min_{\mathbf{x}\in D} g(P_{\mathbf{x}})$$

Exemples de fonctionnelles: la moyenne, médiane, quantile donné, variance...

Avec le SLGP:

- On estime $P_{\mathbf{x}}$, et donc $g(P_{\mathbf{x}})$
- On utilise le SLGP pour guider l'acquisition de données et répondre à la question:
  Où evaluer le système pour gagner de l'information?

Athénaïs Gautier, David Ginsbourger, and Guillaume Pirot. *Goal-oriented adaptive sampling under random field modelling of response probability distributions*. 2021. arXiv: 2102.07612 [stat.ME]

En problème inverse stochastique

Problème direct: trouver la réponse d'un système à une entrée $\mathbf{x}$.

$$\text{Paramètre } \mathbf{x} \in D \longrightarrow \text{Modèle direct } \mathcal{F} \longrightarrow \text{Réponse } Y_x$$

En problème inverse stochastique

Problème direct: trouver la réponse d'un système à une entrée $\mathbf{x}$.

$$\text{Paramètre } \mathbf{x} \in D \longrightarrow \text{Modèle direct } \mathcal{F} \longrightarrow \text{Réponse } Y_x$$

Problème inverse: Trouver le $\mathbf{x}^*$ qui a généré y_{ref}.

$$\text{Paramètre } \mathbf{x} \in D \overset{?}{\longleftarrow\cdots\cdots} \text{Réponse } y_{ref}$$

En problème inverse stochastique

Problème direct: trouver la réponse d'un système à une entrée $\mathbf{x}$.

$$\text{Paramètre } \mathbf{x} \in D \longrightarrow \text{Modèle direct } \mathcal{F} \longrightarrow \text{Réponse } Y_x$$

Problème inverse: Trouver le $\mathbf{x}^*$ qui a généré y_{ref}.

$$\text{Paramètre } \mathbf{x} \in D \xleftarrow{\quad ? \quad} \text{Réponse } y_{ref}$$

Approche probabiliste:

- Mettre un prior sur $\mathbf{x}$: $\pi[\mathbf{x}]$
- Théorème de Bayes: $\pi[\mathbf{x} | Y_{\mathbf{x}} = y_{ref}] \propto \pi[\mathbf{x}]\pi[Y_{\mathbf{x}} = y_{ref} | x]$

En problème inverse stochastique

Problème direct: trouver la réponse d'un système à une entrée $\mathbf{x}$.

$$\text{Paramètre } \mathbf{x} \in D \longrightarrow \text{Modèle direct } \mathcal{F} \longrightarrow \text{Réponse } Y_x$$

Problème inverse: Trouver le $\mathbf{x}^*$ qui a généré y_{ref}.

$$\text{Paramètre } \mathbf{x} \in D \xleftarrow{?} \text{Réponse } y_{ref}$$

Approche probabiliste:

- Mettre un prior sur $\mathbf{x}$: $\pi[\mathbf{x}]$
- Théorème de Bayes: $\pi[\mathbf{x}|Y_{\mathbf{x}} = y_{ref}] \propto \pi[\mathbf{x}]\pi[Y_{\mathbf{x}} = y_{ref}|x]$

$\rightarrow$ demande la vraisemblance $\pi[Y_{\mathbf{x}} = y|\mathbf{x}]$.

Approximate Bayesian Computation (ABC).

On simule une réponse $y_{\mathbf{x}}$ associée à $\mathbf{x}$ et la compare avec $\Delta(y_{ref}, y_{\mathbf{x}})$

Approximate Bayesian Computation (ABC).

On simule une réponse $y_{\mathbf{x}}$ associée à $\mathbf{x}$ et la compare avec $\Delta(y_{ref}, y_{\mathbf{x}})$

L'approximation:

$$\pi\left[\mathbf{x} | Y_{\mathbf{x}} = y_{ref}\right] \approx \pi\left[\mathbf{x} | \Delta(y_{ref}, Y_{\mathbf{x}}) \leq \epsilon\right] \text{ pour } \epsilon \text{ petit}$$

Approximate Bayesian Computation (ABC).

On simule une réponse $y_{\mathbf{x}}$ associée à $\mathbf{x}$ et la compare avec $\Delta(y_{ref}, y_{\mathbf{x}})$

L'approximation:

$$\pi\left[\mathbf{x} | Y_{\mathbf{x}} = y_{ref}\right] \approx \pi\left[\mathbf{x} | \Delta(y_{ref}, Y_{\mathbf{x}}) \leq \epsilon\right] \text{ pour } \epsilon \text{ petit}$$

Postérieur ABC:

$$\pi\left[\mathbf{x} | \Delta(y_{ref}, Y_{\mathbf{x}}) \leq \epsilon\right] \propto \pi[\mathbf{x}] \pi\left[\Delta(y_{ref}, Y_{\mathbf{x}}) \leq \epsilon | \mathbf{x}\right]$$

Approximate Bayesian Computation (ABC).

On simule une réponse $y_{\mathbf{x}}$ associée à $\mathbf{x}$ et la compare avec $\Delta(y_{ref}, y_{\mathbf{x}})$

L'approximation:

$$\pi\left[\mathbf{x}|Y_{\mathbf{x}} = y_{ref}\right] \approx \pi\left[\mathbf{x}|\Delta(y_{ref}, Y_{\mathbf{x}}) \leq \epsilon\right] \text{ pour } \epsilon \text{ petit}$$

Postérieur ABC:

$$\pi\left[\mathbf{x}|\Delta(y_{ref}, Y_{\mathbf{x}}) \leq \epsilon\right] \propto \pi[\mathbf{x}]\pi\left[\Delta(y_{ref}, Y_{\mathbf{x}}) \leq \epsilon|\mathbf{x}\right]$$

ABC rejection sampling algorithm: Approximer le postérieur ABC avec des simulations.

- Demande beaucoup de simulations dont la plupart sont rejetées.

Approximate Bayesian Computation (ABC).

On simule une réponse $y_{\mathbf{x}}$ associée à $\mathbf{x}$ et la compare avec $\Delta(y_{ref}, y_{\mathbf{x}})$

L'approximation:

$$\pi\left[\mathbf{x} | Y_{\mathbf{x}} = y_{ref}\right] \approx \pi\left[\mathbf{x} | \Delta(y_{ref}, Y_{\mathbf{x}}) \leq \epsilon\right] \text{ pour } \epsilon \text{ petit}$$

Postérieur ABC:

$$\pi\left[\mathbf{x} | \Delta(y_{ref}, Y_{\mathbf{x}}) \leq \epsilon\right] \propto \pi[\mathbf{x}]\pi\left[\Delta(y_{ref}, Y_{\mathbf{x}}) \leq \epsilon | \mathbf{x}\right]$$

ABC rejection sampling algorithm: Approximer le postérieur ABC avec des simulations.

- Demande beaucoup de simulations dont la plupart sont rejetées.
- Le choix des $\mathbf{x}$ n'est pas libre mais doit être en rapport avec le prior.

Accélérer ABC I

- Markov chain Monte Carlo ABC
  - Marjoram et al., "Markov chain Monte Carlo without likelihoods"
- Sequential/population MC ABC
  - Bonassi and West, "Sequential Monte Carlo with Adaptive Weights for Approximate Bayesian Computation"
  - Del Moral, Doucet, and Jasra, "An adaptive sequential Monte Carlo method for approximate Bayesian computation"
  - Lenormand, Jabot, and Deffuant, "Adaptive approximate Bayesian computation for complex models"
  - Sisson, Fan, and Tanaka, "Sequential monte carlo without likelihoods"
  - Toni et al., "Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems"

Accélérer ABC II

- Synthetic likelihood
  - Price et al., "Bayesian synthetic likelihood"
  - Wood, "Statistical inference for noisy nonlinear ecological dynamic systems"
  - Wilkinson, "Accelerating ABC methods using Gaussian processes"
  - Järvenpää et al., "Efficient acquisition rules for model-based approximate Bayesian computation"
  - Järvenpää, Vehtari, and Marttinen, "Batch simulations and uncertainty quantification in Gaussian process surrogate-based approximate Bayesian computation"
- Quantile-based
  - Kacprzak et al., "Accelerating Approximate Bayesian Computation with Quantile Regression: application to cosmological redshift distributions"

Une application en hydrogeologie

Situation: Un contaminant est relaché à profondeur $\mathbf{x}$ (0m à 10m.) et se répand dans un acquifère de structure inconnue. En se basant sur des courbes de concentrations observées, on cherche $\mathbf{x}$.[3]

[3] Gautier, Ginsbourger, and Pirot, "Probabilistic ABC with Spatial Logistic Gaussian Process modelling".

Une application en hydrogeologie

Situation: Un contaminant est relaché à profondeur **x** (0m à 10m.) et se répand dans un acquifère de structure inconnue. En se basant sur des courbes de concentrations observées, on cherche **x**.[3]

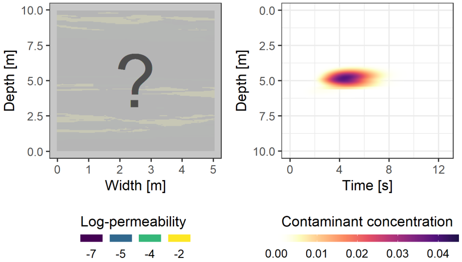


Figure: Courbes de référence

[3] Gautier, Ginsbourger, and Pirot, "Probabilistic ABC with Spatial Logistic Gaussian Process modelling".

Une application en hydrogeologie 2/2

Nouvelles simulations pour une source **x** par:

Une application en hydrogeologie 2/2

Nouvelles simulations pour une source $\mathbf{x}$ par:

1. Simuler une structure géologique plausible.

Une application en hydrogeologie 2/2

Nouvelles simulations pour une source $\mathbf{x}$ par:

1. Simuler une structure géologique plausible.
2. Simuler la réponse associée à cette géologie et profondeur $\mathbf{x}$.

Une application en hydrogeologie 2/2

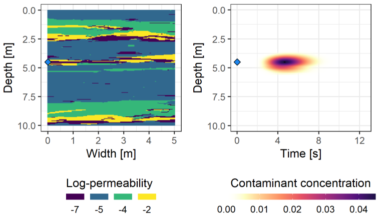
Nouvelles simulations pour une source **x** par:

1. Simuler une structure géologique plausible.
2. Simuler la réponse associée à cette géologie et profondeur **x**.
3. Calculer la dissimilarité entre courbe simulée et observée.

Une application en hydrogeologie 2/2

Nouvelles simulations pour une source **x** par:

1. Simuler une structure géologique plausible.
2. Simuler la réponse associée à cette géologie et profondeur **x**.
3. Calculer la dissimilarité entre courbe simulée et observée.

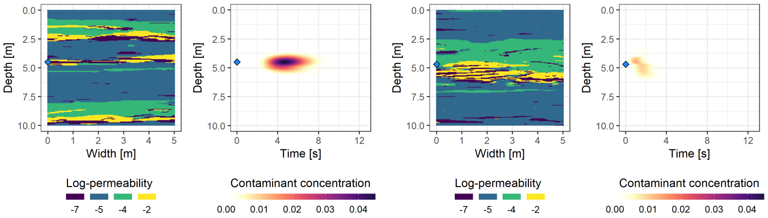


(a) $x = 4.49$m, misfit ≈ 0.09

Une application en hydrogeologie 2/2

Nouvelles simulations pour une source **x** par:

1. Simuler une structure géologique plausible.
2. Simuler la réponse associée à cette géologie et profondeur **x**.
3. Calculer la dissimilarité entre courbe simulée et observée.

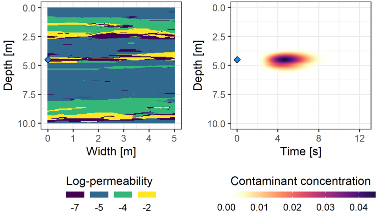


(a) $x = 4.49$m, misfit ≈ 0.09 (b) $x = 4.69$m, misfit ≈ 0.35

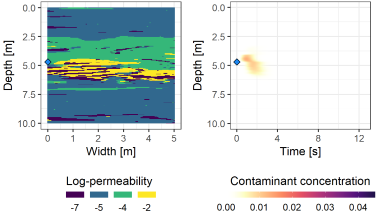
Une application en hydrogeologie 2/2

Nouvelles simulations pour une source **x** par:

1. Simuler une structure géologique plausible.
2. Simuler la réponse associée à cette géologie et profondeur **x**.
3. Calculer la dissimilarité entre courbe simulée et observée.



(a) $x = 4.49$m, misfit ≈ 0.09



(b) $x = 4.69$m, misfit ≈ 0.35

$\rightarrow$ Apprendre la distribution de la dissimilarité

En hydrogéologie: les résultats

But: Trouver la source qui donne les courbes de référence

En hydrogéologie: les résultats

But: Trouver la source qui donne les courbes de référence

Postérieur ABC:

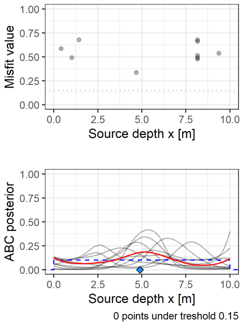
$$\pi\left[\mathbf{x} | \Delta(y_{ref}, y_{\mathbf{x}}) \leq \epsilon\right] \propto \pi[\mathbf{x}]\pi\left[\Delta(y_{ref}, Y_{\mathbf{x}}) \leq \epsilon | \mathbf{x}\right]$$

En hydrogéologie: les résultats

But: Trouver la source qui donne les courbes de référence

Postérieur ABC:

$$\pi\left[\mathbf{x} \middle| \Delta(y_{ref}, y_{\mathbf{x}}) \leq \epsilon\right] \propto \pi[\mathbf{x}]\pi\left[\Delta(y_{ref}, Y_{\mathbf{x}}) \leq \epsilon \middle| \mathbf{x}\right]$$



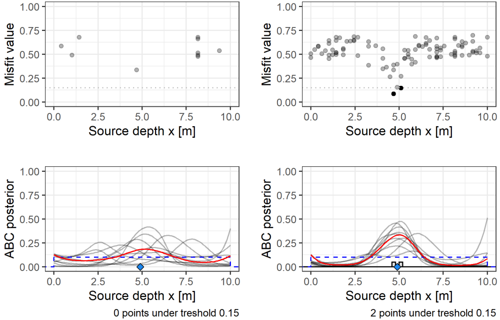
(a) n=10

En hydrogéologie: les résultats

But: Trouver la source qui donne les courbes de référence

Postérieur ABC:

$$\pi\left[\mathbf{x} | \Delta(y_{ref}, y_{\mathbf{x}}) \leq \epsilon\right] \propto \pi[\mathbf{x}] \pi\left[\Delta(y_{ref}, Y_{\mathbf{x}}) \leq \epsilon | \mathbf{x}\right]$$



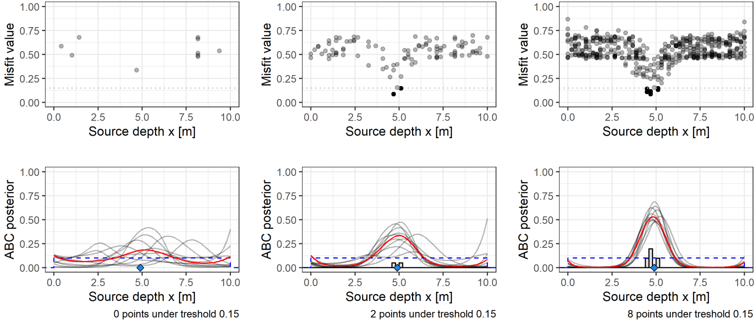
(a) n=10 (b) n=100

En hydrogéologie: les résultats

But: Trouver la source qui donne les courbes de référence

Postérieur ABC:

$$\pi\left[\mathbf{x} \middle| \Delta(y_{ref}, y_{\mathbf{x}}) \leq \epsilon\right] \propto \pi[\mathbf{x}]\pi\left[\Delta(y_{ref}, Y_{\mathbf{x}}) \leq \epsilon \middle| \mathbf{x}\right]$$



(a) n=10 (b) n=100 (c) n=500

Merci de votre attention!

4 Annexe

- Aspects théoriques
- Implémentation
- Problème inverse stochastique “simple”

La caractérisation du processus

Une remarque: Pour un GP $(Z_{\mathbf{x},t})_{(\mathbf{x},t)\in D\times\mathcal{T}}$ et un champ aléatoire $(R_{\mathbf{x}})_{\mathbf{x}\in D}$:

$$\frac{\int_B e^{Z_{\mathbf{x},u}(\omega)}\,d\lambda(u)}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}(\omega)}\,d\lambda(u)} = \frac{\int_B e^{[Z_{\mathbf{x},u}+R_{\mathbf{x}}](\omega)}\,d\lambda(u)}{\int_{\mathcal{T}} e^{[Z_{\mathbf{x},u}+R_{\mathbf{x}}](\omega)}\,d\lambda(u)} \tag{1}$$

La caractérisation du processus

Une remarque: Pour un GP $(Z_{\mathbf{x},t})_{(\mathbf{x},t)\in D\times\mathcal{T}}$ et un champ aléatoire $(R_{\mathbf{x}})_{\mathbf{x}\in D}$:

$$\frac{\int_B e^{Z_{\mathbf{x},u}(\omega)}\,d\lambda(u)}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}(\omega)}\,d\lambda(u)} = \frac{\int_B e^{[Z_{\mathbf{x},u}+R_{\mathbf{x}}](\omega)}\,d\lambda(u)}{\int_{\mathcal{T}} e^{[Z_{\mathbf{x},u}+R_{\mathbf{x}}](\omega)}\,d\lambda(u)} \qquad (1)$$

Comment caractériser les SLGPs ?

La caractérisation du processus: cas "continu"

Soit un processus de SLGP $(Y_{\mathbf{x},t})_{(\mathbf{x},t)\in D\times\mathcal{T}}$, avec $Y_{\mathbf{x},\cdot}$ presque surement une pdf continue et strictement positive pour tous les $\mathbf{x}$ simultanément. Y est indistinguable d'un SLGP si et seulement si

$$\Delta \log Y_{\mathbf{x},t,t'} = \log Y_{\mathbf{x},t} - \log Y_{\mathbf{x},t'}$$

est indistinguable d'un GP continu sur $D \times \mathcal{T}^2$.

La caractérisation du processus: cas "continu"

Soit un processus de SLGP $(Y_{\mathbf{x},t})_{(\mathbf{x},t)\in D\times\mathcal{T}}$, avec $Y_{\mathbf{x},\cdot}$ presque surement une pdf continue et strictement positive pour tous les $\mathbf{x}$ simultanément. Y est indistinguable d'un SLGP si et seulement si

$$\Delta \log Y_{\mathbf{x},t,t'} = \log Y_{\mathbf{x},t} - \log Y_{\mathbf{x},t'}$$

est indistinguable d'un GP continu sur $D \times \mathcal{T}^2$.

Dans le cas continu, le SLGP se caractérise par

$$m([\mathbf{x}, t_1, t_2]) = \mathbb{E}\left[\log Y_{\mathbf{x},t_1} - \log Y_{\mathbf{x},t_2}\right] \tag{2}$$

$$k([\mathbf{x}, t_1, t_2], [\mathbf{x}', t_1', t_2']) = \mathrm{Cov}\Big(\log Y_{\mathbf{x},t_1} - \log Y_{\mathbf{x},t_2}, \log Y_{\mathbf{x}',t_1'} - \log Y \tag{3}$$

Régularité spatiale du SLGP

Question: On sait que la régularité spatiale d'un GP $Z \sim \mathcal{GP}(0, k)$ peut être imposée en prenant un noyau k *régulier*.
Qu'en est-il du SLGP ?

Régularité spatiale du SLGP

Question: On sait que la régularité spatiale d'un GP $Z \sim \mathcal{GP}(0, k)$ peut être imposée en prennant un noyau k *régulier*.
Qu'en est-il du SLGP ?

Si le noyau de covariance sous-jacent est régulier, le SLGP hérite d'une partie de cette régularité.

Régularité spatiale du SLGP en images

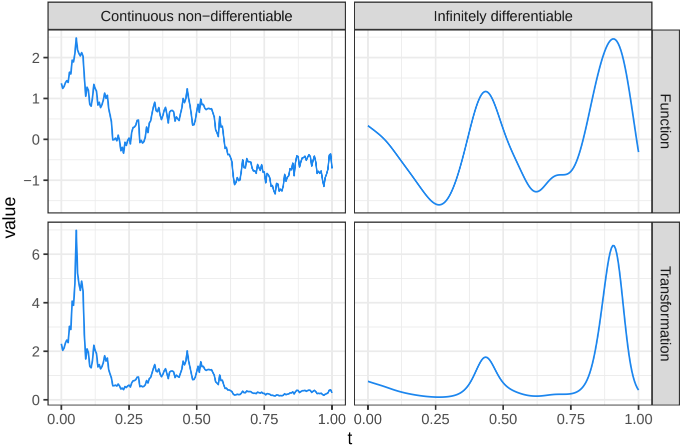


Figure: Logistic density transformation of functions with varying smoothness

La régularité spatiale du processus

S'il existe $C, \alpha_1, \alpha_2 > 0$ tels que $\forall(\mathbf{x}, \mathbf{x}') \in D^2, (t, t') \in \mathcal{T}^2$:

$$d_k^2([\mathbf{x}, t], [\mathbf{x}', t']) \leq C \cdot \max(\|\mathbf{x} - \mathbf{x}'\|_\infty^{\alpha_1}, \|t - t'\|_\infty^{\alpha_2}) \qquad (4)$$

avec d_k^2 la semi-distance canonique de k:

$$d_k^2([\mathbf{x}, t], [\mathbf{x}', t']) = k\left([\mathbf{x}, t], [\mathbf{x}, t]\right) + k\left([\mathbf{x}', t'], [\mathbf{x}', t']\right) - 2k\left([\mathbf{x}, t], [\mathbf{x}',\right.$$

La régularité spatiale du processus

S'il existe $C, \alpha_1, \alpha_2 > 0$ tels que $\forall(\mathbf{x}, \mathbf{x}') \in D^2, (t, t') \in \mathcal{T}^2$:

$$d_k^2([\mathbf{x}, t], [\mathbf{x}', t']) \leq C \cdot \max(\|\mathbf{x} - \mathbf{x}'\|_\infty^{\alpha_1}, \|t - t'\|_\infty^{\alpha_2}) \qquad (4)$$

avec d_k^2 la semi-distance canonique de k:

$$d_k^2([\mathbf{x}, t], [\mathbf{x}', t']) = k\left([\mathbf{x}, t], [\mathbf{x}, t]\right) + k\left([\mathbf{x}', t'], [\mathbf{x}', t']\right) - 2k\left([\mathbf{x}, t], [\mathbf{x}',\right.$$

Pour tout $\gamma > 0, \delta > 0$,il existe $K_{\gamma,\delta}$ tel que pour tout $\mathbf{x}, \mathbf{x}' \in D^2$:

$$\begin{aligned} \mathbb{E}\left[d_H(Y_{\mathbf{x},\cdot}, Y_{\mathbf{x}',\cdot})^\gamma\right] &\leq K_{\gamma,\delta}\|\mathbf{x} - \mathbf{x}'\|_\infty^{\gamma\alpha_1/2-\delta} \\ \mathbb{E}\left[KL(Y_{\mathbf{x},\cdot}, Y_{\mathbf{x}',\cdot})^\gamma\right] &\leq K_{\gamma,\delta}\|\mathbf{x} - \mathbf{x}'\|_\infty^{\gamma\alpha_1-\delta} \\ \mathbb{E}\left[d_{TV}(Y_{\mathbf{x},\cdot}, Y_{\mathbf{x}',\cdot})^\gamma\right] &\leq K_{\gamma,\delta}\|\mathbf{x} - \mathbf{x}'\|_\infty^{\gamma\alpha_1-\delta} \end{aligned} \qquad (5)$$

Les défis de l'implémentation

Le modèle $Y_{\mathbf{x},t} = \dfrac{e^{Z_{\mathbf{x},t}}}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}}\, d\lambda u} \quad (\mathbf{x}, t) \in D \times \mathcal{T}$

Les défis de l’implémentation

Le modèle $Y_{\mathbf{x},t} = \dfrac{e^{Z_{\mathbf{x},t}}}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}}\, d\lambda u} \quad (\mathbf{x}, t) \in D \times \mathcal{T}$

- Le SLGP est intrinsèquement infini-dimensionnel en raison de l’intégrale au dénominateur.

Les défis de l'implémentation

Le modèle $Y_{\mathbf{x},t} = \dfrac{e^{Z_{\mathbf{x},t}}}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}}\, d\lambda u} \quad (\mathbf{x}, t) \in D \times \mathcal{T}$

- Le SLGP est intrinsèquement infini-dimensionnel en raison de l'intégrale au dénominateur.
- Comment la calculer ?

Les défis de l'implémentation

Le modèle $Y_{\mathbf{x},t} = \dfrac{e^{Z_{\mathbf{x},t}}}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}}\, d\lambda u} \quad (\mathbf{x}, t) \in D \times \mathcal{T}$

- Le SLGP est intrinsèquement infini-dimensionnel en raison de l'intégrale au dénominateur.
- Comment la calculer ?
- Comment conditionner sur des données?

Les défis de l'implémentation

Le modèle $Y_{\mathbf{x},t} = \dfrac{e^{Z_{\mathbf{x},t}}}{\int_{\mathcal{T}} e^{Z_{\mathbf{x},u}}\, d\lambda u} \quad (\mathbf{x}, t) \in D \times \mathcal{T}$

- Le SLGP est intrinsèquement infini-dimensionnel en raison de l'intégrale au dénominateur.
- Comment la calculer ?
- Comment conditionner sur des données?
- Comment estimer les hyper-paramètres du modèle?

Une implémentation en rang fini

On considère des processus de rang fini:

$$\forall \mathbf{x} \in D,\ t \in \mathcal{T},\ Z_{\mathbf{x},t} = \sum_{j=1}^{p} \sqrt{\lambda_j} f_j(\mathbf{x}, t)\varepsilon_j$$

$p \in \mathbb{N}$, $\lambda_i > 0$,les f_i sont des fonctions de base $D \times \mathcal{T}$ et ε_i's sont i.i.d. $\mathcal{N}(0, 1)$.

Une implémentation en rang fini

On considère des processus de rang fini:

$$\forall \mathbf{x} \in D,\ t \in \mathcal{T},\ Z_{\mathbf{x},t} = \sum_{j=1}^{p} \sqrt{\lambda_j} f_j(\mathbf{x}, t)\varepsilon_j$$

$p \in \mathbb{N}$, $\lambda_i > 0$,les f_i sont des fonctions de base $D \times \mathcal{T}$ et ε_i's sont i.i.d. $\mathcal{N}(0, 1)$.

$\rightarrow$ Permet d'approximer l'intégrale avec une quadrature.

Choix des fonctions de base: les RFF

Un noyau k stationnaire peut être approximé par k_{RFF}:

$$k_{RFF}(x, y) := \frac{\sigma^2}{2\pi p} \sum_{i=1}^{p} \cos(\omega_i^T x + u_i) \cos(\omega_i^T y + u_i) \tag{6}$$

où $\sigma^2 = k(0, 0)$, les ω_i's sont des tirages i.i.d. de la densité spectrale de k et les u_i's sont des tirages i.i.d. $\mathcal{U}(0, 2\pi)$.

Choix des fonctions de base: les RFF

Un noyau k stationnaire peut être approximé par k_{RFF}:

$$k_{RFF}(x, y) := \frac{\sigma^2}{2\pi p} \sum_{i=1}^{p} \cos(\omega_i^T x + u_i) \cos(\omega_i^T y + u_i) \qquad (6)$$

où $\sigma^2 = k(0, 0)$, les ω_i's sont des tirages i.i.d. de la densité spectrale de k et les u_i's sont des tirages i.i.d. $\mathcal{U}(0, 2\pi)$.

Il satisfait $\mathbb{E}[k_{RFF}] = k$.

Choix des fonctions de base: les RFF

Un noyau k stationnaire peut être approximé par k_{RFF}:

$$k_{RFF}(x,y) := \frac{\sigma^2}{2\pi p} \sum_{i=1}^{p} \cos(\omega_i^T x + u_i) \cos(\omega_i^T y + u_i) \qquad (6)$$

où $\sigma^2 = k(0,0)$, les ω_i's sont des tirages i.i.d. de la densité spectrale de k et les u_i's sont des tirages i.i.d. $\mathcal{U}(0, 2\pi)$.

Il satisfait $\mathbb{E}[k_{RFF}] = k$.

$W_{RFF}(x) := \frac{\sigma}{\sqrt{2\pi p}} \sum_{i=1}^{p} \epsilon_i \cos(\omega_i^T x + u_i)$ où ϵ_i sont i.i.d. $\mathcal{N}(0,1)$, est un GP $(0, k_{RFF}(x,y))$.

Conditionnement sur des données

Données: des emplacements $(\mathbf{x}_i)_i \in D^n$, et les données observées correspondantes $(t_i)_i \in \mathcal{T}^n$.

Conditionnement sur des données

Données: des emplacements $(\mathbf{x}_i)_i \in D^n$, et les données observées correspondantes $(t_i)_i \in \mathcal{T}^n$.

En assumant que les données sont indépendantes, le théorème de Bayes donne:

$$\pi(\varepsilon|\mathbf{T}_n) \propto \pi(\varepsilon) \prod_{i=1}^{n} \frac{e^{\sum_{j=1}^{p} \sqrt{\lambda_j} f_j(\mathbf{x}_i, t_i)\varepsilon_j}}{\int_{\mathcal{T}} e^{\sum_{j=1}^{p} \sqrt{\lambda_j} f_j(\mathbf{x}_i, u)\varepsilon_j} \, du} \tag{7}$$

Conditionnement sur des données

Données: des emplacements $(\mathbf{x}_i)_i \in D^n$, et les données observées correspondantes $(t_i)_i \in \mathcal{T}^n$.

En assumant que les données sont indépendantes, le théorème de Bayes donne:

$$\pi(\varepsilon|\mathbf{T}_n) \propto \pi(\varepsilon) \prod_{i=1}^{n} \frac{e^{\sum_{j=1}^{p} \sqrt{\lambda_j} f_j(\mathbf{x}_i, t_i)\varepsilon_j}}{\int_{\mathcal{T}} e^{\sum_{j=1}^{p} \sqrt{\lambda_j} f_j(\mathbf{x}_i, u)\varepsilon_j} \, du} \tag{7}$$

$\rightarrow$ Implémentable en MCMC

Données météo en Suisse

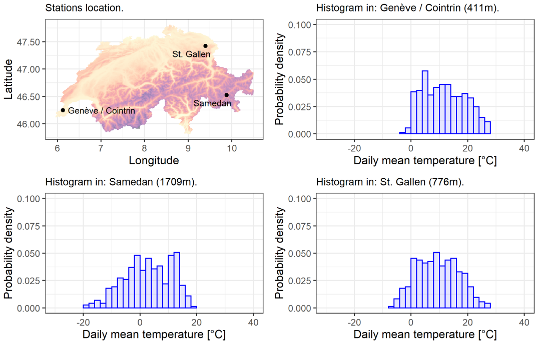


Figure: Distribution marginale des temperatures moyennes quotidiennes à 3 stations météo en Suisse.[4]

[4] MeteoSwiss, *Climatological Network - Daily Values.*

SLGP entrainé

On append $\pi[\underbrace{\text{Temperature}}_{y \text{ pour le SLGP}} \mid \underbrace{\text{emplacement en Suisse}}_{x \text{ pour le SLGP}}]$

[5] **datameteoslgp**.

SLGP entrainé

On apprend $\pi[\underbrace{\text{Temperature}}_{y \text{ pour le SLGP}} | \underbrace{\text{emplacement en Suisse}}_{x \text{ pour le SLGP}}]$

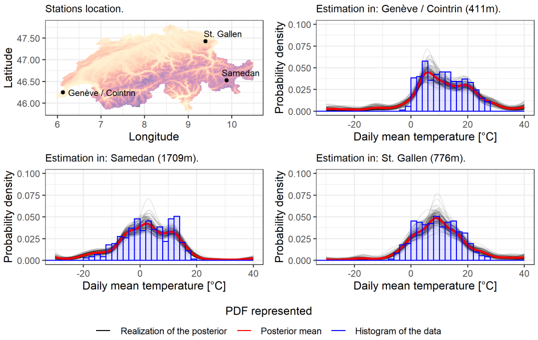


Figure: Predictions of marginal distribution of daily mean temperatures at 3 stations in Switzerland.[5]

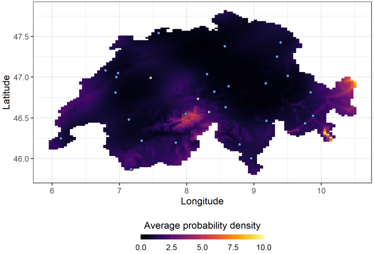
[5] **datameteoslgp**.

Probabilité d'excursion

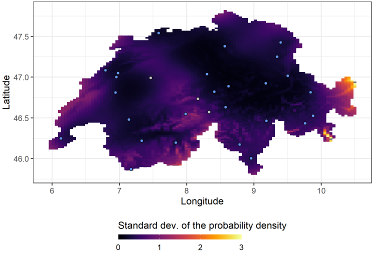
Si je sais que la temperature y_{ref} est plus faible que $-15°$C, où suis-je?
(Avec un prior uniforme sur la Suisse)

Probabilité d'excursion

Si je sais que la temperature y_{ref} est plus faible que $-15°C$, où suis-je?
(Avec un prior uniforme sur la Suisse)



(a) Average the probability



(b) Std of the probability

Figure: Probability of my location in Switzerland. The blue circles indicate the stations locations.