SLGP-based surrogates for spatially dependent discrete outputs:

modelling, uncertainty quantification, and data acquisition

Athénaïs Gautier
DTIS, ONERA, Université Paris-Saclay, 91120, Palaiseau, France
athenais.gautier@onera.fr

mODa14 June 17, 2026

Why GPs are so popular for active learning ?

Flexibility

GPs fits complex response surfaces with very few assumptions.
GPs can be defined over plenty of different inputs.
GPs can incorporate prior knowledge.

Probabilistic prediction

A full posterior, hence **uncertainty** at every input - not just a point estimate.
$\rightarrow$ probabilistic prediction!

Why GPs are so popular for active learning ?

Flexibility

GPs fits complex response surfaces with very few assumptions.
GPs can be defined over plenty of different inputs.
GPs can incorporate prior knowledge.

Probabilistic prediction

A full posterior, hence **uncertainty** at every input - not just a point estimate.
$\rightarrow$ probabilistic prediction!

$\rightarrow$ That **probabilistic prediction** is exactly what is leveraged for active learning: level-set estimation, Bayesian optimisation, probabilistic inversion, ...

Why GPs are so popular for active learning ?

Flexibility

GPs fits complex response surfaces with very few assumptions.
GPs can be defined over plenty of different inputs.
GPs can incorporate prior knowledge.

Probabilistic prediction

A full posterior, hence **uncertainty** at every input - not just a point estimate.
$\rightarrow$ probabilistic prediction!

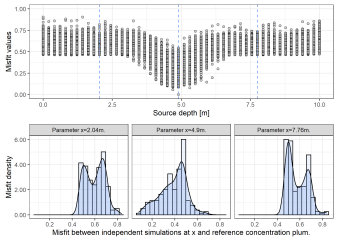
$\rightarrow$ That **probabilistic prediction** is exactly what is leveraged for active learning: level-set estimation, Bayesian optimisation, probabilistic inversion, ...

But: a GP assumes a *Gaussian* output. What if the response is not Gaussian: a density, a class, a count?

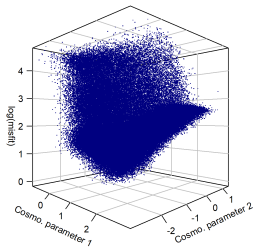
PhD Motivation: Stochastic inverse problems & surrogate modelling

The challenge: Precise statistical inference in complex systems with index $\mathbf{x}$ and response $T_{\mathbf{x}}$.

- Complex spatial dependence of $T_{\mathbf{x}}$
- Costly model evaluations
- Heterogeneously scattered data



1D quantity of interest in an hydrogeological inverse problem



2D quantity of interest for a cosmological inverse problem

1 Modelling for probabilistic inference

2 Spatial Logistic Gaussian Process

3 Active learning

4 Applications and proof of concept
- Case 1: binomial distribution with fixed n_{bin} and varying p
- Case 2: binomial distribution with varying n_{bin} and p

Spatial Logistic Gaussian Process : SLGP

Let the latent GP Z depend on the index $\mathbf{x}$ as well as the response $t \in [a, b]$:

$$\begin{cases} (Z_{\mathbf{x},t})_{\mathbf{x}\in D,\ t\in[a,b]} \sim \mathcal{GP} \\ Y_{\mathbf{x},t} := \dfrac{e^{Z_{\mathbf{x},t}}}{\int_{[a,b]} e^{Z_{\mathbf{x},u}}\, du} \sim SLGP \end{cases}$$

Spatial Logistic Gaussian Process : SLGP

Let the latent GP Z depend on the index $\mathbf{x}$ as well as the response $t \in [a, b]$:

$$\begin{cases} (Z_{\mathbf{x},t})_{\mathbf{x}\in D,\ t\in[a,b]} \sim \mathcal{GP} \\ Y_{\mathbf{x},t} := \dfrac{e^{Z_{\mathbf{x},t}}}{\int_{[a,b]} e^{Z_{\mathbf{x},u}}\, du} \sim SLGP \end{cases}$$

At each $\mathbf{x}$, $Y_{\mathbf{x},\cdot}$ is, informally, a "random density" w.r.t. Lebesgues measure.

Spatial Logistic Gaussian Process : SLGP

Let the latent GP Z depend on the index $\mathbf{x}$ as well as the response $t \in [a, b]$:

$$\begin{cases} (Z_{\mathbf{x},t})_{\mathbf{x}\in D,\ t\in[a,b]} \sim \mathcal{GP} \\ Y_{\mathbf{x},t} := \dfrac{e^{Z_{\mathbf{x},t}}}{\int_{[a,b]} e^{Z_{\mathbf{x},u}}\, du} \sim SLGP \end{cases}$$

At each $\mathbf{x}$, $Y_{\mathbf{x},\cdot}$ is, informally, a "random density" w.r.t. Lebesgues measure.

Discrete-valued SLGP, for a response k in $\{1, ..., K\}$:

$$\begin{cases} (Z_{\mathbf{x},k})_{\mathbf{x}\in D,\ k\in\{1,\dots,K\}} \sim \mathcal{GP} \\ Y_{\mathbf{x},k} := \dfrac{e^{Z_{\mathbf{x},k}}}{\sum_{i=1}^{K} e^{Z_{\mathbf{x},i}}} \sim SLGP \end{cases}$$

Likelihood of the observations

For locations $(x_i)_{i=1}^n \in D^n$, observations $\mathbf{k_n} := (k_i)_{i=1}^n \in \{1, ..., K\}^n$, assuming independent k_i's we have the likelihood:

$$\pi(\mathbf{k_n}|Y) = \prod_{i=1}^{n} Y_{\mathbf{x}_i,k_i} = \prod_{i=1}^{n} \frac{e^{Z_{\mathbf{x}_i,k_i}}}{\sum_{j=1}^{K} e^{Z_{\mathbf{x}_i,j}}} \tag{1}$$

Likelihood of the observations

For locations $(x_i)_{i=1}^n \in D^n$, observations $\mathbf{k_n} := (k_i)_{i=1}^n \in \{1, ..., K\}^n$, assuming independent k_i's we have the likelihood:

$$\pi(\mathbf{k_n}|Y) = \prod_{i=1}^{n} Y_{\mathbf{x}_i,k_i} = \prod_{i=1}^{n} \frac{e^{Z_{\mathbf{x}_i,k_i}}}{\sum_{j=1}^{K} e^{Z_{\mathbf{x}_i,j}}} \tag{1}$$

We focus on *finite-rank* GPs:

$$Z(\mathbf{x}, k) = \mu(\mathbf{x}, k) + \sum_{j=1}^{p} f_j(\mathbf{x}, k)\varepsilon_j = \boldsymbol{\varepsilon}^\top \boldsymbol{F}_(\mathbf{x}, k), \ \forall \mathbf{x} \in D, \ k \in \{0, ..., K\}, \tag{2}$$

where $\boldsymbol{\varepsilon} = (\varepsilon_1, ..., \varepsilon_p)^\top$ is a random vector of p i.i.d. $\mathcal{N}(0, \sigma^2)$ random variables, and μ is a deterministic known trend function.

Posterior of a finite-rank SLGP

For locations $(x_i)_{i=1}^n \in D^n$, observations $\mathbf{k_n} := (k_i)_{i=1}^n \in \{1, ..., K\}^n$, assuming independent k_i's and a finite-rank SLGP, we have the posterior:

$$\pi[\boldsymbol{\epsilon}|\mathbf{k_n}] \propto \phi_p(\boldsymbol{\epsilon}) \prod_{i=1}^{n} \frac{e^{\mu(\mathbf{x}_i,k_i)+\sum_{j=1}^{p}\epsilon_j f_j(\mathbf{x}_i,k_i)}}{\sum_{\ell=1}^{K} e^{\mu(\mathbf{x}_i,\ell)+\sum_{j=1}^{p}\epsilon_j f_j(\mathbf{x}_i,\ell)}} \tag{3}$$

where ϕ_p the pdf of the p-variate standard normal distribution

Posterior of a finite-rank SLGP

For locations $(x_i)_{i=1}^n \in D^n$, observations $\mathbf{k_n} := (k_i)_{i=1}^n \in \{1, ..., K\}^n$, assuming independent k_i's and a finite-rank SLGP, we have the posterior:

$$\pi[\boldsymbol{\epsilon}|\mathbf{k_n}] \propto \phi_p(\boldsymbol{\epsilon}) \prod_{i=1}^{n} \frac{e^{\mu(\mathbf{x}_i,k_i)+\sum_{j=1}^{p}\epsilon_j f_j(\mathbf{x}_i,k_i)}}{\sum_{\ell=1}^{K} e^{\mu(\mathbf{x}_i,\ell)+\sum_{j=1}^{p}\epsilon_j f_j(\mathbf{x}_i,\ell)}} \tag{3}$$

where ϕ_p the pdf of the p-variate standard normal distribution

This posterior is **log-concave**.

- Fast optimization (MAP)
- Good Laplace approximation
- Fast MCMC sampling. (Dwivedi et al., 2018)

The main properties of discrete SLGP

Inference yields N posterior draws $\boldsymbol{\epsilon}^{(1)}, \ldots, \boldsymbol{\epsilon}^{(N)}$ (via Laplace or MCMC), each inducing a pmf field $\widehat{Y}^{(i)}_{\mathbf{x},k}$. The average is $\bar{Y}_{\mathbf{x},k} = \frac{1}{N} \sum_{i=1}^{N} \widehat{Y}^{(i)}_{\mathbf{x},k}$.

The main properties of discrete SLGP

Inference yields N posterior draws $\boldsymbol{\epsilon}^{(1)}, \ldots, \boldsymbol{\epsilon}^{(N)}$ (via Laplace or MCMC), each inducing a pmf field $\widehat{Y}^{(i)}_{\mathbf{x},k}$. The average is $\bar{Y}_{\mathbf{x},k} = \frac{1}{N} \sum_{i=1}^{N} \widehat{Y}^{(i)}_{\mathbf{x},k}$.

✔ Exact

No quadrature error; the normaliser is a finite softmax sum. Inference is *cheaper* than in the continuous case.

Can be evaluated at any $\mathbf{x}$

The finite-rank representation allows the surrogate can be *queried at unseen indices* - essential for design.

Ordinal for free

One *shared* latent process on $D \times \{1, \ldots, K\}$ couples neighbouring categories - natural for counts.

Connections to familiar models

Multinomial logistic regression

- feature map $\check{\boldsymbol{F}}(\mathbf{x})$ *shared* across classes
- coefficients $\check{\varepsilon}_k$ vary with k
- up to $K \times P$ free parameters

Multiclass GP classifier

- one latent GP *per class*
- no intrinsic ordering of labels

Discrete-output SLGP

- *single* shared ε
- class-specific *basis* $\boldsymbol{F}(\mathbf{x}, k)$
- one process on $D \times \{1, \ldots, K\}$
- structured, input-dependent cross-class correlation
- ordinal structure built in
- latent dimension fixed

1 Modelling for probabilistic inference

2 Spatial Logistic Gaussian Process

3 Active learning

4 Applications and proof of concept
- Case 1: binomial distribution with fixed n_{bin} and varying p
- Case 2: binomial distribution with varying n_{bin} and p

Active data acquisition with SLGP

Goal. Select new design points $\mathbf{x} \in D$ that are informative for the entire conditional probability mass function $(Y_{\mathbf{x},1}, \ldots, Y_{\mathbf{x},K}) \mid \mathcal{D}_n$.

$$p(T_{\mathbf{x}} = k|\mathcal{D}_n) = \int Y_{\mathbf{x},k}(\varepsilon)\, p(\varepsilon|\mathcal{D}_n)\, d\varepsilon. \tag{4}$$

Active data acquisition with SLGP

Goal. Select new design points $\mathbf{x} \in D$ that are informative for the entire conditional probability mass function $(Y_{\mathbf{x},1}, \ldots, Y_{\mathbf{x},K}) \mid \mathcal{D}_n$.

$$p(T_{\mathbf{x}} = k|\mathcal{D}_n) = \int Y_{\mathbf{x},k}(\varepsilon)\, p(\varepsilon|\mathcal{D}_n)\, d\varepsilon. \tag{4}$$

Posterior-draw approximation

Given posterior draws $\widehat{Y}^{(i)}_{\mathbf{x},k}$, $i = 1, \ldots, N$:

$$p(T_{\mathbf{x}} = k|\mathcal{D}_n) \approx \bar{Y}_{\mathbf{x},k} = \frac{1}{N} \sum_{i=1}^{N} \widehat{Y}^{(i)}_{\mathbf{x},k}.$$

Active data acquisition with SLGP

Goal. Select new design points $\mathbf{x} \in D$ that are informative for the entire conditional probability mass function $(Y_{\mathbf{x},1}, \ldots, Y_{\mathbf{x},K}) \mid \mathcal{D}_n$.

$$p(T_{\mathbf{x}} = k|\mathcal{D}_n) = \int Y_{\mathbf{x},k}(\varepsilon)\, p(\varepsilon|\mathcal{D}_n)\, d\varepsilon. \tag{4}$$

Posterior-draw approximation

Given posterior draws $\widehat{Y}^{(i)}_{\mathbf{x},k}$, $i = 1, \ldots, N$:

$$p(T_{\mathbf{x}} = k|\mathcal{D}_n) \approx \bar{Y}_{\mathbf{x},k} = \frac{1}{N} \sum_{i=1}^{N} \widehat{Y}^{(i)}_{\mathbf{x},k}.$$

$$\mathbf{x}_{n+1} \in \arg\max_{\mathbf{x} \in D} a(\mathbf{x})$$

Considered strategies

- **Random sampling** *(baseline)*.

Considered strategies

- **Random sampling** *(baseline)*.
- **Probability variance**:

$$\widehat{a}_{\mathrm{VarProb}}(\mathbf{x}) = \sum_{k=1}^{K} \frac{1}{N-1} \sum_{i=1}^{N} \left(\widehat{Y}_{\mathbf{x},k}^{(i)} - \bar{Y}_{\mathbf{x},k} \right)^2$$

Direct ℓ_2-type dispersion in probability space.

Considered strategies

- **Random sampling** *(baseline)*.
- **Probability variance**:

$$\widehat{a}_{\mathrm{VarProb}}(\mathbf{x}) = \sum_{k=1}^{K} \frac{1}{N-1} \sum_{i=1}^{N} \left(\widehat{Y}_{\mathbf{x},k}^{(i)} - \bar{Y}_{\mathbf{x},k} \right)^2$$

  Direct ℓ_2-type dispersion in probability space.
- **Predictive entropy** (MacKay, 1992; Cohn et al., 1994):

$$\widehat{a}_{\mathrm{Ent}}(\mathbf{x}) = -\sum_{k=1}^{K} \bar{Y}_{\mathbf{x},k} \log \bar{Y}_{\mathbf{x},k}$$

  Total predictive uncertainty, mixes epistemic and aleatoric.

Considered strategies

- **Random sampling** *(baseline)*.
- **Probability variance**:

$$\widehat{a}_{\mathrm{VarProb}}(\mathbf{x}) = \sum_{k=1}^{K} \frac{1}{N-1} \sum_{i=1}^{N} \left(\widehat{Y}_{\mathbf{x},k}^{(i)} - \bar{Y}_{\mathbf{x},k} \right)^2$$

  Direct ℓ_2-type dispersion in probability space.
- **Predictive entropy** (MacKay, 1992; Cohn et al., 1994):

$$\widehat{a}_{\mathrm{Ent}}(\mathbf{x}) = -\sum_{k=1}^{K} \bar{Y}_{\mathbf{x},k} \log \bar{Y}_{\mathbf{x},k}$$

  Total predictive uncertainty, mixes epistemic and aleatoric.
- **BALD** (Houlsby et al., 2011):

$$\widehat{a}_{\mathrm{BALD}}(\mathbf{x}) = -\sum_{k=1}^{K} \bar{Y}_{\mathbf{x},k} \log \bar{Y}_{\mathbf{x},k} + \frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{K} \widehat{Y}_{\mathbf{x},k}^{(i)} \log \widehat{Y}_{\mathbf{x},k}^{(i)}$$

  Mutual information; targets disagreement between pmf draws.

1 Modelling for probabilistic inference

2 Spatial Logistic Gaussian Process

3 Active learning

4 Applications and proof of concept
- Case 1: binomial distribution with fixed n_{bin} and varying p
- Case 2: binomial distribution with varying n_{bin} and p

1 Modelling for probabilistic inference

2 Spatial Logistic Gaussian Process

3 Active learning

4 Applications and proof of concept
- Case 1: binomial distribution with fixed n_{bin} and varying p
- Case 2: binomial distribution with varying n_{bin} and p

Case 1: binomial conditional pmf

Controlled benchmark. We consider a one-dimensional conditional distribution with known ground truth:

$$T_p \mid p \sim \mathrm{Bin}(10, p), \qquad p \in (0, 1),$$

so that $K = 11$ categories and

$$\mathbb{P}(T_p = k) = \binom{10}{k} p^k (1-p)^{10-k}, \qquad k = 0, \ldots, 10.$$

Evaluation metric. Prediction quality is measured using integrated KL divergence:

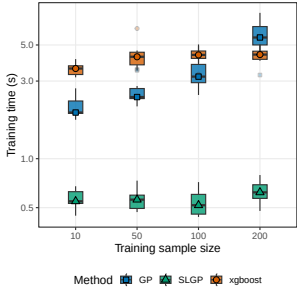
$$\mathrm{IKL}(f, g) = \int_0^1 \sum_{k=0}^{10} f_p(k) \log \frac{f_p(k)}{g_p(k)} \, dp,$$

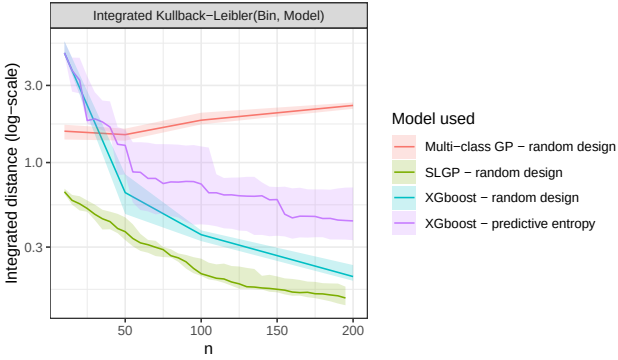
approximated numerically on a dense grid.

Case 1: binomial conditional pmf

Compared methods

- SLGP for conditional pmf estimation;
- GP-based multiclass model;
- random forest baseline.





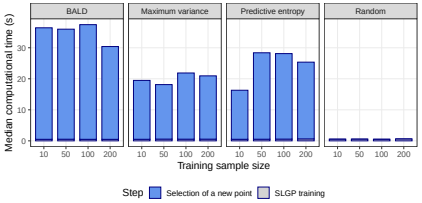
Predictive performance and computational time over 10 repetitions.

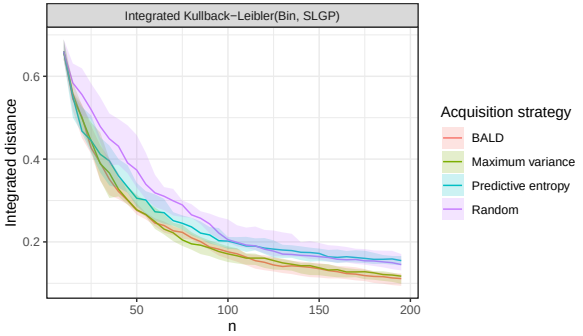
Sequential design: learning the pmf field efficiently

Protocol. The SLGP is initialized with $n_0 = 10$ random observations. At each iteration, a new location p_{new} is selected and 5 independent observations are generated from $\text{Bin}(10, p_{\text{new}})$.

Acquisition strategies

- random sampling;
- predictive entropy;
- BALD;
- posterior variance of class probabilities.





Median IKL across 10 repetitions; ribbons show interquartile ranges.

Finite-rank SLGP (RFF, Matérn-5/2, 100 frequencies $\rightarrow$ 200 basis functions).

Case 2: binomial with varying n_{bin} and p

Two-dimensional benchmark. We now index by $\mathbf{x} = (n_{bin}, p) \in \{5, \dots, 20\} \times (0, 1)$, with

$$T_{n_{bin},p} \mid \mathbf{x} \sim \mathrm{Bin}(n_{bin}, p),$$

so that $K = 21$ categories and the *support itself varies* with n_{bin}:

$$\mathbb{P}(T_{n_{bin},p} = k) = \mathbb{1}_{k \le n_{bin}} \binom{n_{bin}}{k} p^k (1-p)^{n_{bin}-k}, \qquad k = 0, \dots, 20.$$

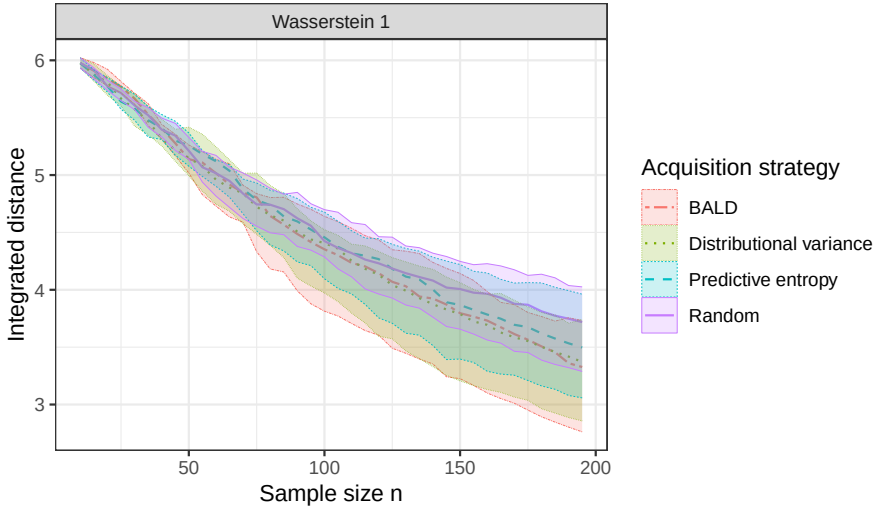
Evaluation metric. KL is unstable under support shifts, so we use the Wasserstein-1 distance:

$$W_1(F, G) = \sum_{k=0}^{K} \big| F(k) - G(k) \big|,$$

on the CDFs F, G, integrated over the input domain.

Same finite-rank SLGP as before (RFF, Matérn-5/2, 100 frequencies $\to$ 200 basis functions). The varying support makes inference and design more involved.

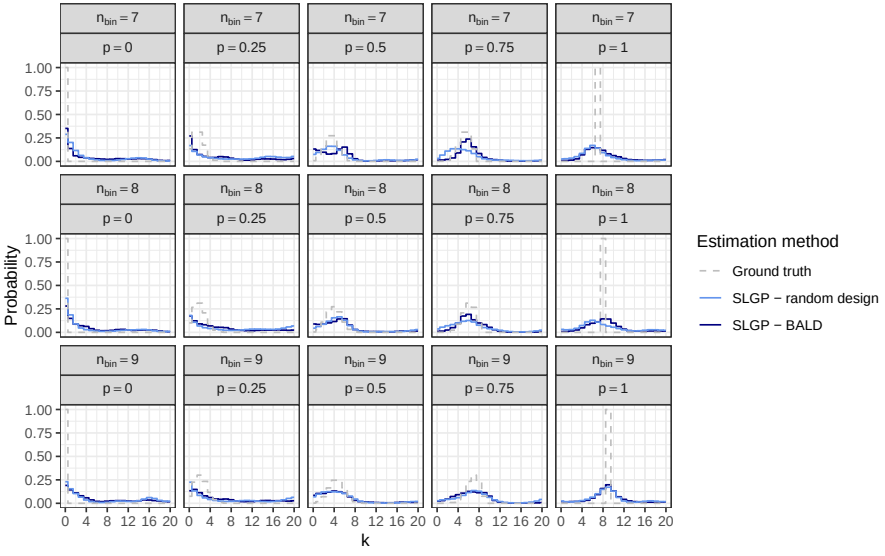
Case 2: results without structural trend



Integrated W_1

vs. dataset size; ribbons show interquartile ranges (10 repetitions).

Case 2: results without structural trend



Truth vs. SLGP (random) vs. SLGP (BALD). ⇒ purely smooth latent representations adapt poorly to

Encoding support with a structured trend

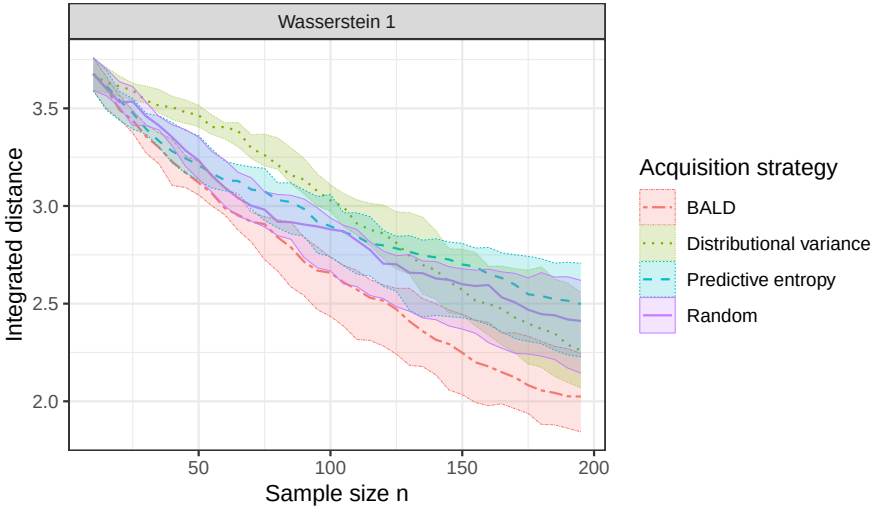
Idea. Inject structural prior knowledge into the latent field via a deterministic trend that penalises out-of-support categories:

$$\mu(n_{bin}, p, k) = -10 \cdot \mathbf{1}_{\{k > n_{bin}\}}.$$

Plugging into the discrete SLGP gives probabilites:

$$Y_{n_{bin},p}(k) \propto \begin{cases} \exp\{\varepsilon^\top \mathbf{F}(n_{bin}, p, k)\}, & k \leq n_{bin}, \\ 4.5 \times 10^{-5} \exp\{\varepsilon^\top \mathbf{F}(n_{bin}, p, k)\}, & k > n_{bin}. \end{cases}$$

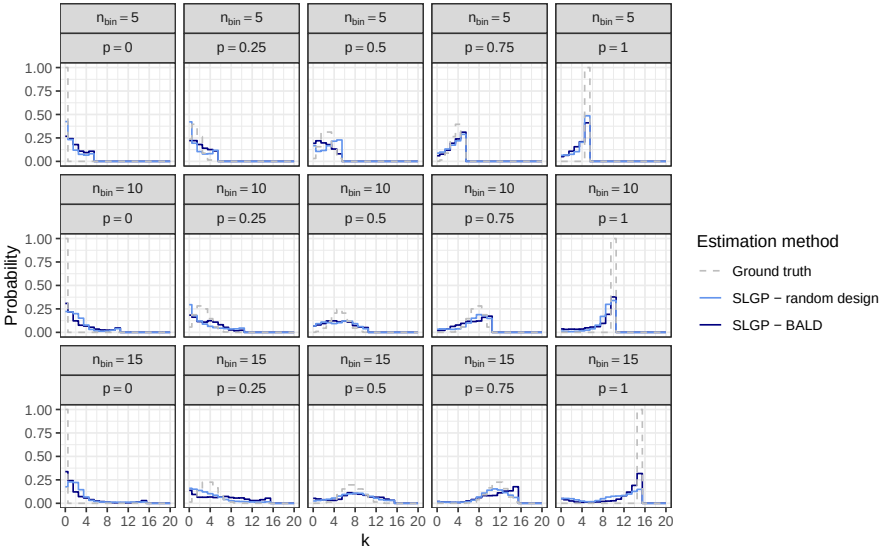
Case 2: results with structural trend



Integrated W_1

vs. dataset size; ribbons show interquartile ranges (10 repetitions).

Case 2: results with structural trend



Truth vs. SLGP (random) vs. SLGP (BALD).

Conclusion

What we did. A discrete-output **SLGP** for spatially indexed categorical/count responses $\rightarrow$ full posterior on the pmf at any location, with coherent uncertainty.

Takeaways.

- The continuous SLGP extends to finite support $\rightarrow$ normalized w.r.t. atomic reference measure, normaliser is a finite softmax (no quadrature error).
- Epistemic-uncertainty criteria (especially **BALD**) beat random sampling for distributional reconstruction.
- A **structured trend** encodes support constraints and markedly improves the varying-n case.

Next. Estimating the trend jointly, variational / augmentation-based inference for large K, SUR-type criteria on distributional functionals, proper scoring rules when ground truth is unknown.

References I

Previous SLGP literature

Tokdar, Surya T. et al. (2010). "Bayesian density regression with logistic Gaussian process and subspace projection". In: *Bayesian analysis* 5.2, pp. 319–344.

Pati, Debdeep et al. (2013). "Posterior consistency in conditional distribution estimation". In: *Journal of multivariate analysis* 116, pp. 456–472.

Theory and methodology

Gautier, Athénaïs (2023). "Modelling and predicting distribution-valued fields with applications to inversion under uncertainty". PhD Thesis. Institute of Mathematical Statistics and Actuarial Science, University of Bern.

Gautier, Athénaïs and David Ginsbourger (2025). *Continuous Logistic Gaussian Random Measure Fields for Spatial Distributional Modelling*. To appear in Annals of the Institute of Statistical Mathematics. arXiv: `2110.02876 [stat.ME]`.

Gautier, Athenaïs (2026). *SLGP-based surrogates for spatially dependent discrete outputs: modelling, uncertainty quantification, and data acquisition*. Under review for the Proceedings of mODa 14.

References for the criterion

- MacKay, David JC (1992). "Information-based objective functions for active data selection". In: *Neural computation* 4.4, pp. 590–604.
- Cohn, David et al. (1994). "Active learning with statistical models". In: *Advances in Neural Information Processing Systems 7*.
- Houlsby, Neil et al. (2011). *Bayesian Active Learning for Classification and Preference Learning*. arXiv: `1112.5745 [stat.ML]`.

References for the SLGP implementation, and code

- Gautier, Athénaïs (2025b). *Vignette - Introduction to the SLGP package, and application to a dataset*. URL: https://htmlpreview.github.io/?https://github.com/AthenaisGautier/SLGPImplementation/blob/main/IntroductionSLGP.html (visited on 02/18/2025).
- — (2026a). *Github repository - Code for the numerical experiments*. URL: https://github.com/AthenaisGautier/SLGP-discrete (visited on 05/30/2026).
- — (2026b). *SLGP: Spatial Logistic Gaussian Process for field density estimation*. R package version 1.0.1. URL: https://CRAN.R-project.org/package=SLGP.

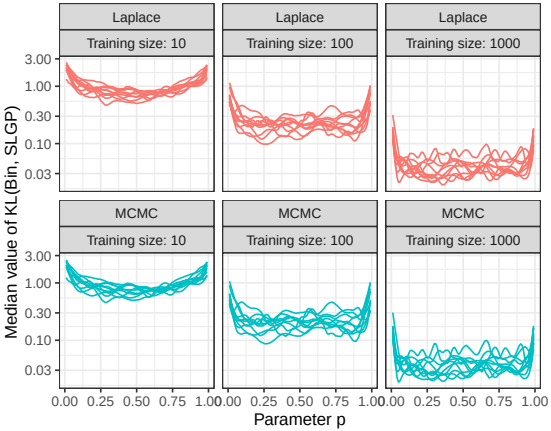
Other application of the SLGP for goal-oriented learning

- Gautier, Athénaïs et al. (2020). "Probabilistic ABC with Spatial Logistic Gaussian Process modelling". In: *Third Workshop on Machine Learning and the Physical Sciences*. URL: https://ml4physicalsciences.github.io/2020/files/NeurIPS_ML4PS_2020_112.pdf.
- — (2021). "Goal-oriented adaptive sampling under random field modelling of response probability distributions". In: *ESAIM: Proceedings and Surveys* 71, pp. 89–100. DOI: 10.1051/proc/202171108. URL: https://doi.org/10.1051/proc/202171108.
- Gautier, Athénaïs (2023). "Modelling and predicting distribution-valued fields with applications to inversion under uncertainty". PhD Thesis. Institute of Mathematical Statistics and Actuarial Science, University of Bern.

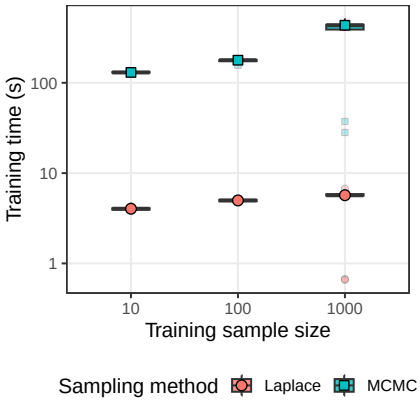
5 Appendix

Posterior approximation: Laplace vs. MCMC

The SLGP likelihood is log-concave and the prior Gaussian $\Rightarrow$ the posterior is log-concave, so a Laplace approximation around the mode is well justified.



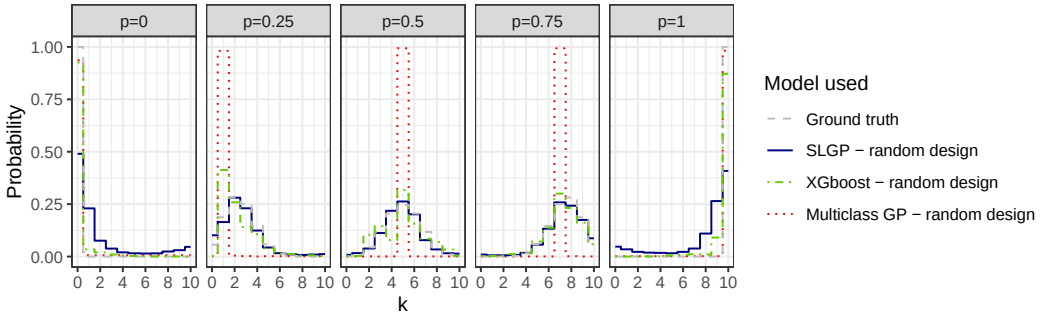
Pointwise KL to the ground-truth pmf, $n \in \{10, 100, 1000\}$, 10 repetitions.



Training time on a laptop.

Posterior approximation: SLGP vs GP VS XGBoost

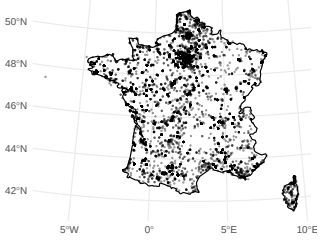
Case : Binomial with varying p, estimation of SLGP vs of multi-class GP VS XGboost



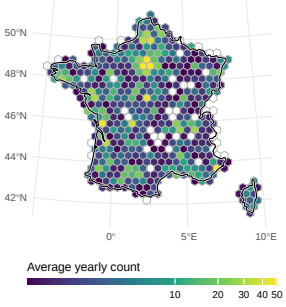
Real data: *Bombus terrestris* occurrences in France

Data. GBIF occurrence records of *Bombus terrestris* in metropolitan France, 2019–2023, with valid coordinates and dates. Spatial index $\mathbf{x} = (\mathrm{lon}, \mathrm{lat}) \in D \subset \mathbb{R}^2$.

- Regular **hexagonal tessellation** (isotropic neighbourhoods, less directional bias).
- Per cell h and year y, a truncated count response: $t_{h,y} = \min\big(50,\ \#\{\text{occurrences in } h \text{ during year } y\}\big)$.
- Truncation limits the influence of densely sampled urban cells.



Raw occurrence points.

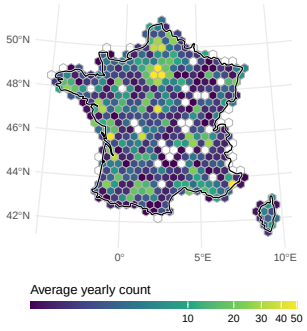


Hexagonal tessellation.

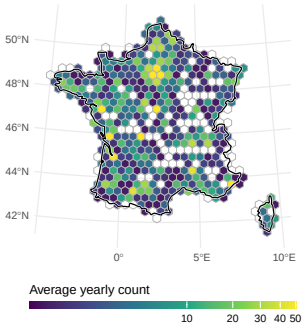
Real data: training protocol and predictive summary

Protocol. Train on counts from **2020–2021** (each year an independent spatial replication), RFF Matérn-5/2, 200 frequencies → 400 basis functions.

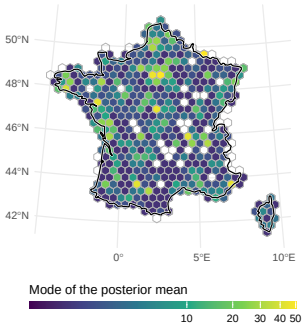
- Reference: empirical average yearly count over the longer window 2019–2023, $\bar{t}_h^{\mathrm{ref}} = \frac{1}{5}\sum_{y=2019}^{2023} t_{h,y}$.
- Compared to the **mode of the SLGP posterior predictive** fitted on 2020–2021.



Full reference (2019–2023).



Training data (2020–2021).



SLGP posterior mode.