Proceeding Paper

Surrogate-Enhanced Parameter Inference for Function-Valued Models †

Christopher G. Albert [1,2,*] , Ulrich Callies [3] and Udo von Toussaint [1]

1   Max-Planck-Institut für Plasmaphysik, 85748 Garching, Germany; udt@ipp.mpg.de
2   Institute of Theoretical and Computational Physics, Technische Universität Graz, 8010 Graz, Austria
3   Helmholtz-Zentrum Hereon, 21502 Geesthacht, Germany; ulrich.callies@hereon.de
*   Correspondence: albert@alumni.tugraz.at
†   Presented at the 40th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, online, 4–9 July 2021.

Abstract: We present an approach to enhance the performance and flexibility of the Bayesian inference of model parameters based on observations of the measured data. Going beyond the usual surrogate-enhanced Monte-Carlo or optimization methods that focus on a scalar loss, we place emphasis on a function-valued output of a formally infinite dimension. For this purpose, the surrogate models are built on a combination of linear dimensionality reduction in an adaptive basis of principal components and Gaussian process regression for the map between reduced feature spaces. Since the decoded surrogate provides the full model output rather than only the loss, it is re-usable for multiple calibration measurements as well as different loss metrics and, consequently, allows for flexible marginalization over such quantities and applications to Bayesian hierarchical models. We evaluate the method's performance based on a case study of a toy model and a simple riverine diatom model for the Elbe river. As input data, this model uses six tunable scalar parameters as well as silica concentrations in the upper reach of the river together with the continuous time-series of temperature, radiation, and river discharge over a specific year. The output consists of continuous time-series data that are calibrated against corresponding measurements from the Geesthacht Weir station at the Elbe river. For this study, only two scalar inputs were considered together with a function-valued output and compared to an existing model calibration using direct simulation runs without a surrogate.

Keywords: parameter inference; Monte Carlo; surrogate model; Gaussian process regression; dimensionality reduction

1. Introduction

Delayed acceptance [1,2] can accelerate Markov chain Monte Carlo (MCMC) sampling up to a factor of one over the acceptance rate. In order to do so, it requires a surrogate of the posterior that contains the cost function inside the likelihood in the case of model calibration. The simplest way to implement delayed acceptance relies on a surrogate with scalar output built for this cost function or for the likelihood. Here, we take an intermediate step and construct a surrogate for the functional output of a blackbox model to be calibrated against reference data. Typical examples are numerical simulations that output time-series or spatial data and depend on tunable input parameters.

There exist numerous related works treating blackbox models with functional outputs with surrogates. Campbell et al. [3] used an adaptive basis of principal component analysis (PCA) to perform global sensitivity analysis. Pratola et al. [4] and Ranjan et al. [5] used GP regression for sequential model calibration in a Bayesian framework. Lebel et al. [6] modeled the likelihood function in an MCMC model calibration via a Gaussian process. Perrin [7] compared the use of a multi-output GP surrogate with a Kronecker structure to an adaptive basis approach.

The present contribution relies on the adaptive basis approach in principal components (Karhunen–Lóeve expansion or functional PCA) to reduce the dimensions of the functional

output, while modeling the map from inputs to weights in this basis via GP regression. We demonstrate the application of this approach on two examples using usual and hierarchical Bayesian model calibration. In the latter case, a surrogate beyond the L_2 cost function is required if the likelihood depends on additional auxiliary parameters. As an example, we allow variations of the (fractional) order of the norm, thereby, marginalizing over different noise models, including Gaussian and Laplacian noise.

2. Gaussian Process Regression and Bayesian Global Optimization

Gaussian process regression [8–10] is a commonly used tool to construct flexible non-parameteric surrogates. Based on the observed outputs $f(x_k)$ at training points x_k and a covariance function $k(x, x')$, the GP regressor predicts a Gaussian posterior distribution at any point x^*. For a single prediction $f(x^*)$, the expected value and variance of this distribution are given by

$$\bar{f}(x^*) = m(x^*) + K^*(K + \sigma_n I)^{-1} y, \tag{1}$$

$$\mathrm{var}[f(x^*)] = K^{**} - K^*(K + \sigma_n I)^{-1} K^{*T}, \tag{2}$$

where $m(x^*)$ is the mean model, the covariance matrix K contains entries $K_{ij} = k(x_i, x_j)$ based on the training set, $K_i^*(x^*, x_i)$ are entries of a row vector, and $K^{**} = k(x^*, x^*)$ is a scalar. The unit matrix I is added with the noise covariance σ_n, which regularizes the problem and is usually estimated in an optimization loop together with other kernel hyperparameters.

Such a surrogate with uncertainty information can be used for Bayesian global optimization [11–13] of the log-posterior as a cost function. Here, we apply this method to reach the vicinity of the posterior's mode before sampling. As an acquisition function, we use the expected improvement (see, e.g., [12]) at a newly observed location x^* given existing training data $\mathcal{D}$,

$$
\begin{aligned}
a_{\mathrm{EI}}(x^\star) &= E[\max(0, \bar{f}(x^*) - \hat{f})|x^*, \mathcal{D}] \\
&= (\bar{f}(x^*) - \hat{f})\Phi(\hat{f}; \bar{f}(x^*), \mathrm{var}[f(x^*)]) + \mathrm{var}[f(x^*)]\mathcal{N}(\hat{f}; \bar{f}(x^*), \mathrm{var}[f(x^*)]),
\end{aligned} \tag{3}
$$

where $\hat{f}$ is the optimum value for $f(x)$ observed thus far. Due to the non-linear transformation from the functional blackbox output to the value of the cost function, it is more convenient to realize Bayesian optimization with a direct GP surrogate of the cost function that is constructed in addition to the surrogate for the functional output for the KL expansion coefficients described below.

3. Delayed Acceptance MCMC

Delayed acceptance MCMC builds on a fast surrogate for the posterior $\tilde{p}(x|y)$ to reject unlikely proposals early [1,2]. Following the usual Metropolis–Hastings algorithm, the probability to accept a new proposal x^* in this first stage in the n-the step of the Markov chain is, as usual,

$$\tilde{P}_{\mathrm{acc}}^n = \frac{\tilde{p}(x^*|y)}{\tilde{p}(x_{n-1}|y)} \frac{g(x_{n-1}|x^*)}{g(x^*|x_{n-1})}, \tag{4}$$

where g is a transition probability that has been suitably tuned during warmup. The true posterior $p(x|y)$ is only evaluated if the proposal "survives" this first stage and enters the final acceptance probability

$$P_{\mathrm{acc}}^n = \frac{p(x^*|y)}{p(x_{n-1}|y)} \frac{\tilde{p}(x_{n-1}|y)}{\tilde{p}(x^*|y)}. \tag{5}$$

Actual computation is typically performed in the logarithmic space with a cost function

$$\ell(x|y) \equiv -\log p(x|y). \tag{6}$$

If this function is fixed, it is most convenient to directly build a surrogate $\tilde{\ell}(x|y)$ for the log-posterior $\ell(x|y)$ including the corresponding prior.

4. Bayesian Hierarchical Models and Fractional Norms

One application of modeling the full functional output instead of only the cost function is the existence of additional distribution parameters θ in the likelihood in addition to the original model inputs x. Such dependencies appear within Bayesian hierarchical models [14], where θ are again subject to a certain (prior) distribution with possibly further levels of hyperparameters. There are essentially two ways to construct a surrogate with support for additional parameters θ: Building a surrogate for the cost function that adds θ as independent variables or constructing a surrogate with functional output for $f_k(x)$ and keeping the dependencies on θ exact. Here, we focus on the latter, and apply this surrogate within delayed acceptance MCMC with both, x and θ as tunable parameters.

As an example, we use a more general noise model than the usual Gaussian likelihood that builds on arbitrary ℓ^θ norms [15–17] with real-valued θ not fixed while traversing the Markov chain. We allow members of the exponential family for observational noise and specify only its scale, but keep θ as a free parameter. Namely, we model the likelihood for observing y in the output as

$$p(y|x,\theta) = \frac{1}{2\sqrt{2}\sigma\,\Gamma(1+\theta^{-1})}e^{-\ell(y;x,\theta)},\tag{7}$$

with the normalized ℓ^θ norm to the power of θ,

$$\ell(y;x,\theta) \equiv \frac{1}{D}\sum_{i=1}^{D}\left|\frac{y_i - f_i(x)}{\sqrt{2}\sigma}\right|^\theta\tag{8}$$

as the loss function between observed data y_i and blackbox model $f_i(x)$. Choosing the usual L_2 norm leads to a Gaussian likelihood for the noise model, whereas using the L_1 norm means Laplacian noise. To maintain the relative scale when varying θ, it is important to add the term $\log\Gamma(1+\theta^{-1})$ from (7) to the negative log-likelihood. In the following use cases, we are going to compare the cases of fixed and variable θ.

5. Linear Dimension Reduction via Principal Components

Formally, the blackbox output for given input x can be a function $f(t) \in \mathbb{H}$ in an infinite-dimensional Hilbert space (though sampled at a finite number of points in practice). Linear dimension reduction in such a space means finding the optimum set of basis functions $\varphi_k(t)$ that spans the output space $f(t;x)$ for any input x given to the blackbox. The reduced model of order r is then given by

$$f(t;x) \approx \sum_{k=1}^{r} z_k(x)\varphi_k(t).\tag{9}$$

This approach is known as the Karhunen–Loéve (KL) expansion [18] in case $f(t;x)$ are interpreted as realizations of a random process, or as the functional principal component analysis (FPCA) [19]. For our application, this distinction does not matter. The KL expansion boils down to solving a regression problem in the non-orthogonal basis of N observed realizations to represent new observations. Then, an eigenvalue problem is solved to invert the $N \times N$ collocation matrix with entries

$$M_{ij} = \langle f(t;x_i), f(t;x_j)\rangle.\tag{10}$$

Here, the inner product in Hilbert spaces and its approximation for a finite set of support points is given by

$$\langle u, v \rangle = \int_{\Omega} u(t)v(t)\, dt \approx \frac{1}{N_t} \sum_{k=1}^{N_t} u(t_k)v(t_k). \tag{11}$$

If $N_t \gg N$ (many support points, few samples), solving the eigenvalue problem of the collocation matrix M is more efficient than the dual one of the covariance matrix C with $C_{ij} = \sum_k f(t_i, x_k)f(t_j, x_k)$ in the usual PCA (see [9] for their equivalence via the singular value decomposition of $Y_{ij} = f(t_i, x_j)$). The question of at which r to truncate the eigenspectrum in (9) depends on the desired accuracy in the output, which is briefly analyzed in the following paragraph.

Error Estimate

Here, we justify why we can assume an L_2 truncation error of the order of the ratio λ_r / λ_1 between the smallest eigenvalue considered in the approximation and the largest one. The truncated SVD can be shown to be the best linear approximation $M^{(r)}$ of lower rank r to an $N \times N$ matrix M in terms of the Frobenius norm $||M||_F$ (see, e.g., [20]). Its value is simply computed from the L_2 norm of singular values,

$$||M||_F = \left(\sum_{k=1}^{N} \sigma_k^2 \right)^{1/2}, \tag{12}$$

where $\sigma_k^2 = \lambda_k$ in the case of real eigenvalues λ_k of a positive semi-definite matrix as for the covariance or collocation matrix. The truncation error is given by

$$||M^{(r)} - M||_F = \left(\sum_{k=r+1}^{N} \lambda_k \right)^{1/2}. \tag{13}$$

The error estimate for the KL expansion uses this convenient property together with the fact that the Frobenius norm is compatible with the usual L_2 norm $|x|$ of vectors y, i.e.,

$$|My| \leq ||M||_F |y|. \tag{14}$$

Representing y via the first r eigenvalues of the collocation matrix yields a relative squared reconstruction error of

$$|(M^{(r)} - M)y|^2 / |y|^2 \leq \sum_{k=r+1}^{N} \lambda_k \leq (N - r)\lambda_r. \tag{15}$$

The last estimate is relatively crude if $N \gg r$, and the spectrum decays fast with the index variable k. If one assumes a decay rate α with

$$\lambda_k \approx \lambda_r (k - r)^{-\alpha}, \tag{16}$$

one obtains

$$\sum_{k=r+1}^{N} \lambda_k \approx \sum_{k=r+1}^{\infty} \lambda_r (k - r)^{-\alpha} = \lambda_r \sum_{k=1}^{\infty} k^{-\alpha} = \lambda_r \zeta(\alpha), \tag{17}$$

where ζ is the Riemann zeta function. This function diverges for a spectral decay of order $\alpha = 1$ and reaches its asymptotic value $\zeta(\infty) = 1$ relatively quickly for $\alpha \geq 2$ (e.g., $\zeta(3) = 1.2$). The spectral decay rate α can be fitted in a log–log plot of λ_k over index k and takes values between $\alpha = 3$ and 5 in our use case. The underlying assumptions are violated if the spectrum stagnates at a large number of constant eigenvalues for higher indices k.

6. Implementation and Results

The idea behind the realization of MCMC with a function-valued surrogate is quite simple. Instead of directly using the surrogate for the cost ℓ with fixed θ, we take a step in-between. Multiple surrogates $\tilde{z}_k(x)$ are built, where each maps the input x to one weight $z_k(x)$ in the KL expansion. A surrogate $\tilde{f}_i(x) \equiv \tilde{f}(t_i; x)$ for the model output is then given by replacing $z_k(x)$ by $\tilde{z}_k(x)$ in (9). The according surrogate $\tilde{\ell}(y; x, \theta)$ for the cost function uses $\tilde{f}_i(x)$ instead of $f_i(x)$ in (8). Dependencies on θ are kept exact in this approach. The main algorithm proceeds in the following steps:

1. Construct a GP surrogate for the L_2 cost function on a space-filling sample sequence over the whole prior range.
2. Refine the sampling points near the posterior's mode through Bayesian global optimization with the L_2 cost surrogate.
3. Train a multi-output GP surrogate for the functional output $z(x)$ on the refined sampling points.
4. Use the function-valued surrogate for delayed acceptance in the MCMC run.

For all GP surrogates, we use a Matern 5/2 kernel for $k(x, x')$ together with a linear mean model for $m(x)$, as realized in the Python package GPy [21]. For step 4, we use Gibbs sampling and the surrogate for $z(x)$, yielding the full output $y(t, x)$ rather than only the L_2 distance to a certain reference dataset. The idea to refine the surrogate iteratively during MCMC had to be abandoned early. The problem is that detailed balance is violated as soon as the surrogate proposal probabilities change when modifying the GP regressor with a new point. In the following application cases, we compare a usual MCMC evaluation using the full model to MCMC with delayed acceptance using the GP surrogate together with the KL expansion/functional PCA (GP+KL) in the output function space.
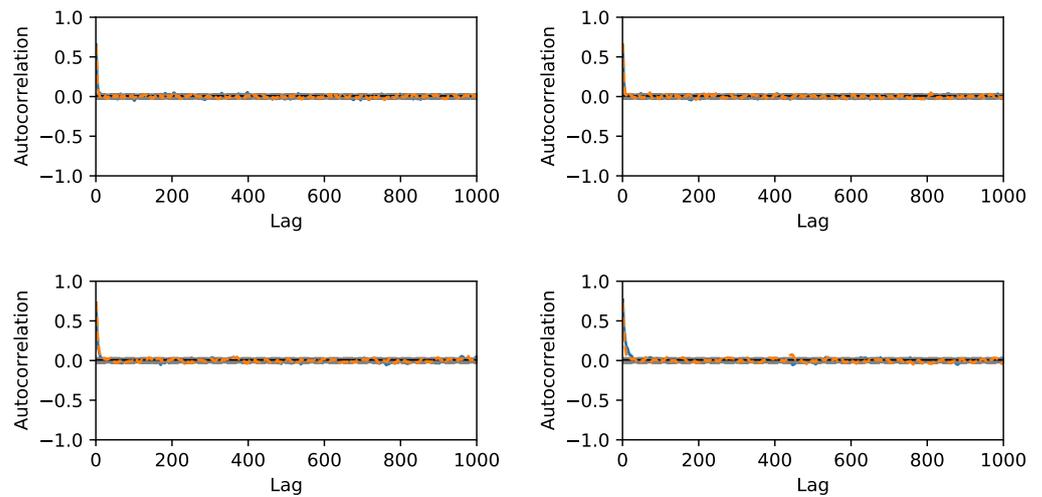
6.1. Toy Model

First, we test the quality of the algorithm on a toy model given by

$$y(t, x) = x_1 \sin((t - x_2)^3). \tag{18}$$

We choose reference values $x_1 = 1.15$, $x_2 = 1.4$ to test the calibration of x against the according output $y^{\text{ref}}(t) \equiv y(t, x^{\text{ref}})$ and add Gaussian noise of amplitude $\sigma = 0.05$. A flat prior is used for x. For the hierarchical model case (7), we choose a starting guess of $\theta = 2$ for the norm's order and a Gaussian prior with $\sigma_\theta = 0.5$ around this value together with a positivity constraint. The initial sampling domain in the square $x_1, x_2 \in (0, 2)$. The comparison between MCMC and delayed acceptance MCMC is made once for fixed $\theta = 2$ (Gaussian likelihood) and then for a hierarchical model with a random walk also in θ. The respective Markov chain with 10,000 steps has a correlation length of ≈ 10 steps (Figure 1) and yields a posterior parameter distribution for (x_1, x_2) depicted in Figure 2.

The results in Figure 2 show good agreement in the posterior distributions of full MCMC and delayed acceptance MCMC. Compared to the case with fixed $\theta = 2$, the additional freedom in θ in the hierarchical model leads to further exploration of the parameter space. The posterior of θ according to the Markov chain is given in Figure 3. The similarity to the prior distribution shows that the data does not yield new information on how to choose θ.

Figure 1. Autocorrelation over lag in MCMC steps for inputs x_1 (solid) and x_2 (dashed) in the toy model. **Top**: Gaussian likelihood, and **bottom**: hierarchical model. **Left**: full MCMC, and **right**: delayed acceptance MCMC with GP+KL surrogate.
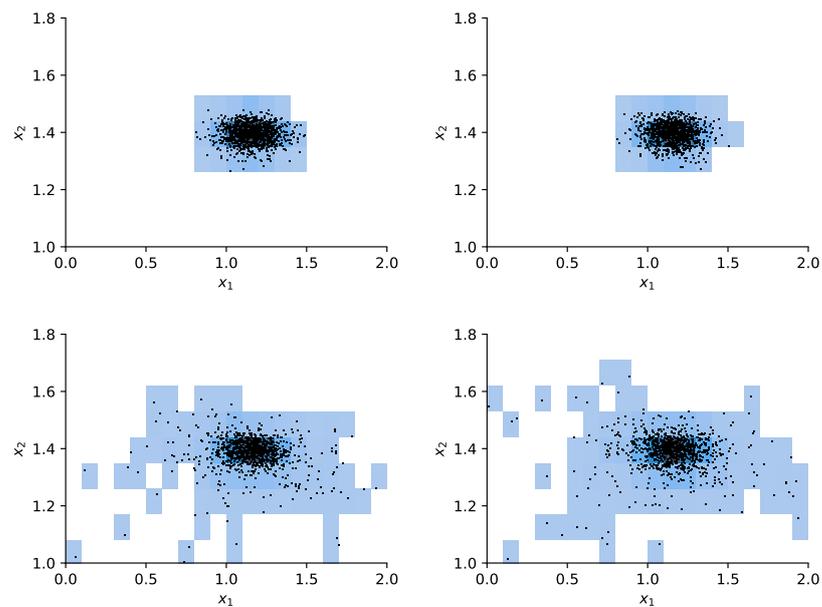
Figure 2. Posterior distribution of the calibrated parameters x in (18). **Top**: Gaussian likelihood, **bottom**: hierarchical model. **Left**: full MCMC, **right**: delayed acceptance MCMC with GP+KL surrogate.
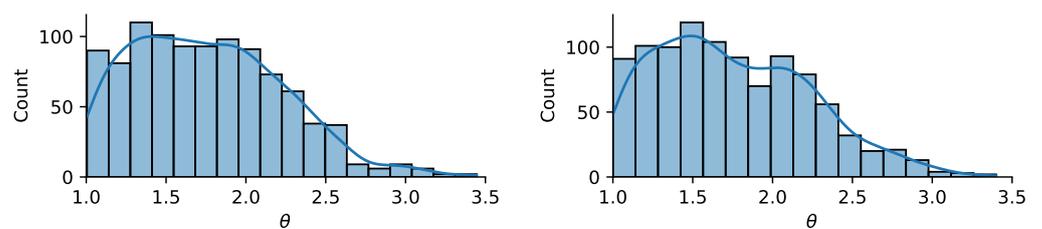
Figure 3. Posterior distribution of the fractional order θ in the loss function with ℓ^θ norm. **Left**: full MCMC, **right**: delayed acceptance MCMC with GP+KL surrogate.

6.2. Riverine Diatom Model

The final application of the described method is on a riverine diatom model [22,23]. This model predicts the chlorophyll *a* concentration at an observation point at the Elbe river as a time series and depends on several input parameters. For simplicity, and to limit computational resources, we select only two of the six scalar inputs and use fixed values for the remaining four. Namely, the chosen parameters $x_1 = K_{\text{light}}$ and $x_2 = \mu_0$ appear in the growth rate inside the diatom model. The latter is given by the "Smith formula" [24] for photosynthesis,

$$\mu(t) \propto \mu_0 \frac{1}{D} \int_0^D \frac{I(t)e^{-\lambda(t)z}}{\sqrt{K_{\text{light}}^2 + I^2(t)e^{-2\lambda(t)z}}} dz,$$

where D is the water depth, and $I(t)$ is the radiation intensity prescribed at the water surface. Light attenuation $\lambda(t) \equiv \lambda_S C_{\text{chl}}(t)$ is modeled to be proportional to the chlorophyll *a* concentration $C_{\text{chl}}(t)$. Equations are solved within a Lagrangian setup, following water parcels that travel down the Elbe river. Data points of the local chlorophyll time series simulated at Geesthacht Weir are made up by chlorophyll *a* values at the Lagrangian trajectory end points. These values are the functional model output $y(t)$ for which the model is calibrated with respect to measurements $y_{\text{ref}}(t)$. As the parameters are positive and limited by reasonable maximum values from domain knowledge, we use a half-sided Cauchy (Lorentz) prior

$$p(x_k) = \frac{2}{\pi} \frac{b_k}{b_k^2 + x_k^2} \text{ for } x_k > 0, \quad p(x_k) = 0 \text{ for } x_k \leq 0. \tag{19}$$

Here, we choose a scale value x_k^* for which $P^* = 90\%$ of the probability volume is contained within $x_k < x_k^*$. Considering the cumulative distribution, we have to set

$$b_k = \frac{x_k^*}{\tan\left(\frac{\pi}{2} P^\star\right)} \tag{20}$$

to realize this condition.

As in the case of the toy model, we use 10,000 steps in the Markov chain. The results for autocorrelation and posterior samples using the full model versus delayed acceptance are shown in Figures 4 and 5. The correlation time of $\approx$500 steps is much larger than in the toy model, and the decay of the autocorrelation over the lag roughly matches between the two approaches. Delayed acceptance sampling produces similar posterior samples in Figure 5 at about one third of the overall computation time. There, one also sees the issue of high correlation between K_{light} and μ_0 in the posterior of the calibration, making Gibbs sampling inefficient in that particular case.
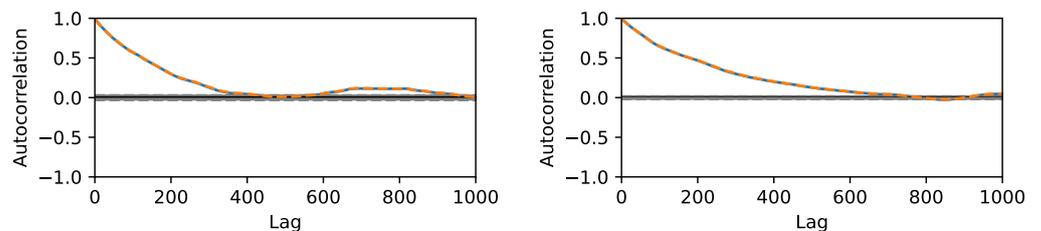
Figure 4. Autocorrelation over lag in MCMC steps for inputs K_{light} (solid) and μ_0 (dashed) in the riverine diatom model. **Top**: Gaussian likelihood, **bottom**: hierarchical model. **Left**: full MCMC, **right**: delayed acceptance MCMC with GP + KL surrogate.
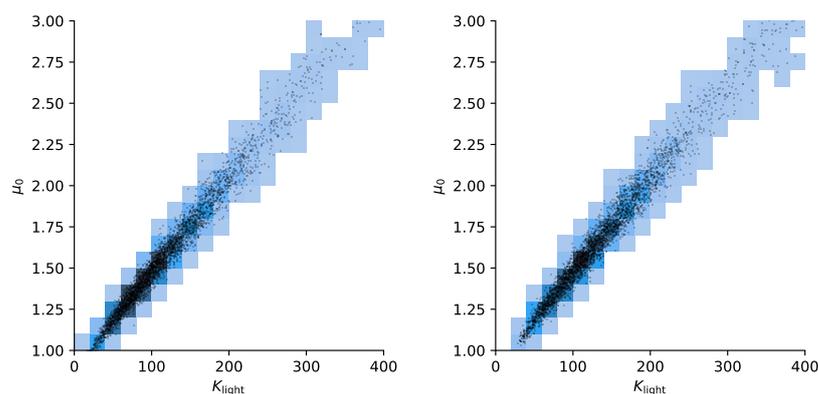
Figure 5. Posterior distribution of calibrated parameters for the riverine diatom model. **Left**: full MCMC, **right**: delayed acceptance MCMC with GP + KL surrogate.

7. Conclusions and Outlook

We illustrated the application of function-valued surrogates to delayed acceptance MCMC for parameter calibration in simple as well as hierarchical Bayesian models. Using a surrogate for the functional output rather than a cost function or likelihood is useful for several reasons. Conceptually, it allows introducing additional distribution parameters in Bayesian hierarchical models. Our results demonstrate that it is possible and efficient to perform MCMC with delayed acceptance on such models while keeping the dependencies in these additional parameters exact. In particular, the fractional order of the norm appearing in the cost function was left free, which is useful for robust model calibration.

The method was applied to a toy model and an application case of a riverine diatom model. In both cases, using delayed acceptance with a surrogate for the functional output produced results comparable to using the full model at only about one third of the actual model evaluations. Compared to direct surrogate modeling of the cost function, we could also observe an increase in the quality of the predicted cost. This is likely connected to the higher flexibility of modeling weights to multiple principal components with Gaussian processes with individual hyperparameters.

The described approach is not immune to the curse of dimensionality. On the one hand, the number of required GP regressors grows linearly with the effective dimensions of the *output* function space. Since evaluation is fast and parallelizable, this is a minor issue in practice. On the other hand, increasing the dimension of the *input* space soon prohibits the construction of a reliable surrogate due to the required number training points to fill the parameter space. In such cases, the preprocessing overhead is expected to outweigh the speedup of delayed acceptance MCMC for either functional or scalar surrogates. More detailed investigations will be required to give quantitative estimates on this trade off.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available in a publicly accessible repository. The data presented in this study are openly available in Zenodo at https://doi.org/10.5281/zenodo.5773865, reference number [25].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Christen, J.A.; Fox, C. Markov Chain Monte Carlo Using an Approximation. *J. Comput. Graph. Stat.* **2005**, *14*, 795–810. [CrossRef]
2. Wiqvist, S.; Picchini, U.; Forman, J.L.; Lindorff-Larsen, K.; Boomsma, W. Accelerating Delayed-Acceptance Markov Chain Monte Carlo Algorithms. *arXiv* **2019**, arXiv:1806.05982.
3. Campbell, K.; McKay, M.D.; Williams, B.J. Sensitivity Analysis When Model Outputs Are Functions. *Reliab. Eng. Syst. Saf.* **2006**, *91*, 1468–1472. [CrossRef]
4. Pratola, M.T.; Sain, S.R.; Bingham, D.; Wiltberger, M.; Rigler, E.J. Fast Sequential Computer Model Calibration of Large Nonstationary Spatial-Temporal Processes. *Technometrics* **2013**, *55*, 232–242. [CrossRef]
5. Ranjan, P.; Thomas, M.; Teismann, H.; Mukhoti, S. Inverse Problem for a Time-Series Valued Computer Simulator via Scalarization. *Open J. Stat.* **2016**, *6*, 528–544. [CrossRef]
6. Lebel, D.; Soize, C.; Fünfschilling, C.; Perrin, G. Statistical Inverse Identification for Nonlinear Train Dynamics Using a Surrogate Model in a Bayesian Framework. *J. Sound Vib.* **2019**, *458*, 158–176. [CrossRef]
7. Perrin, G. Adaptive Calibration of a Computer Code with Time-Series Output. *Reliab. Eng. Syst. Saf.* **2020**, *196*, 106728. [CrossRef]
8. O'Hagan, A. Curve Fitting and Optimal Design for Prediction. *J. R. Stat. Soc. Ser. B* **1978**, *40*, 1–24. [CrossRef]
9. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: Berlin, Germany, 2006.
10. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006. [CrossRef]
11. Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R.P.; de Freitas, N. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proc. IEEE* **2016**, *104*, 148–175. [CrossRef]
12. Osborne, M.A.; Garnett, R.; Roberts, S.J. *Gaussian Processes for Global Optimization*; Learning and 3rd International Conference on Learning and Intelligent Optimization, LION 3; Springer: Trento, Italy, 2009.
13. Preuss, R.; von Toussaint, U. Global Optimization Employing Gaussian Process-Based Bayesian Surrogates. *Entropy* **2018**, *20*, 201. [CrossRef] [PubMed]
14. Allenby, G.M.; Rossi, P.E.; McCulloch, R.E. *Hierarchical Bayes Models: A Practitioners Guide*; SSRN Scholarly Paper ID 655541; Social Science Research Network: Rochester, NY, USA, 2005. [CrossRef]
15. Aggarwal, C.C.; Hinneburg, A.; Keim, D.A. *On the Surprising Behavior of Distance Metrics in High Dimensional Space*; Database Theory—ICDT 2001; Lecture Notes in Computer Science; Van den Bussche, J.; Vianu, V., Eds.; Springer: Berlin/Heidelberg, Germany, 2001; pp. 420–434. [CrossRef]
16. Dose, V. Bayesian Estimate of the Newtonian Constant of Gravitation. *Meas. Sci. Technol.* **2006**, *18*, 176–182. [CrossRef]
17. Flexer, A.; Schnitzer, D. Choosing Lp Norms in High-Dimensional Spaces Based on Hub Analysis. *Neurocomputing* **2015**, *169*, 281–287. [CrossRef]
18. Newman, A.J. *Model Reduction via the Karhunen-Loeve Expansion Part I: An Exposition*; Institute for Systems Research Technical Reports; Univ. Maryland: College Park, MD, USA, 1996.
19. Shang, H.L. A Survey of Functional Principal Component Analysis. *AStA Adv. Stat. Anal.* **2014**, *98*, 121–142. [CrossRef]
20. Cadzow, J.A. Spectral Analysis. In *Handbook of Digital Signal Processing*; Elsevier: Amsterdam, The Netherlands, 1987; pp. 701–740.
21. GPy. *GPy: A Gaussian Process Framework in Python*; Software Publication, Univ. Sheffield: Sheffield, UK, 2012.
22. Callies, U.; Scharfe, M.; Ratto, M. Calibration and Uncertainty Analysis of a Simple Model of Silica-Limited Diatom Growth in the Elbe River. *Ecol. Model.* **2008**, *213*, 229–244. [CrossRef]
23. Scharfe, M.; Callies, U.; Blöcker, G.; Petersen, W.; Schroeder, F. A Simple Lagrangian Model to Simulate Temporal Variability of Algae in the Elbe River. *Ecol. Model.* **2009**, *220*, 2173–2186. [CrossRef]
24. Smith, E.L. Photosynthesis in Relation to Light and Carbon Dioxide. *Proc. Natl. Acad. Sci. USA* **1936**, *22*, 504–511. [CrossRef] [PubMed]
25. Albert, C.G.; von Toussaint, U.; Callies, U. Dataset for article "Surrogate-Enhanced Parameter Inference for Function-Valued Models", v1.0. Zenodo, 2021.