A Bayesian inversion of hydrological and thermal parameters in the hyporheic zone

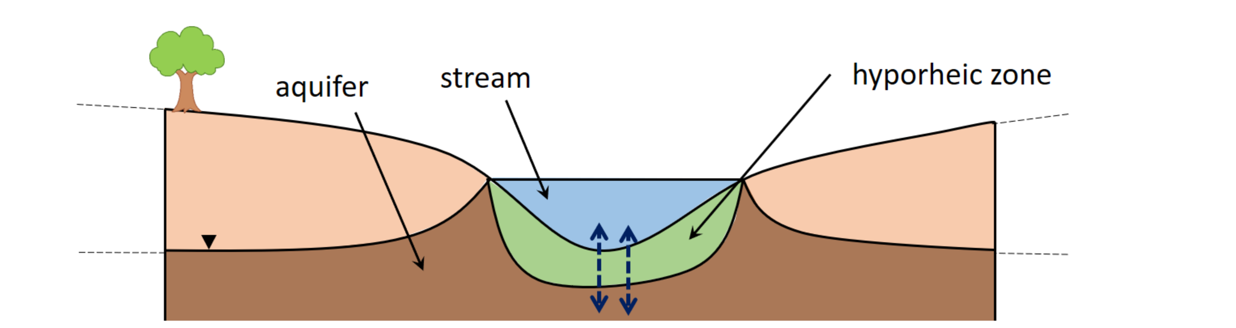
AGU Fall 2015
H53C-1670

Karina Cucchi[1,2], Nicolas Flipo[2], Agnès Rivière[2], Yoram Rubin[1]
[1] University of California, Berkeley, Department of Civil and Environmental Engineering, USA
[2] Centre de Géosciences, MINES ParisTech, PSL Research University, France

Objective of study

We develop a Bayesian inverse modeling framework for estimating hydrological and thermal parameters in the hyporheic zone. The end goal is to obtain posterior probability density functions of these parameters.

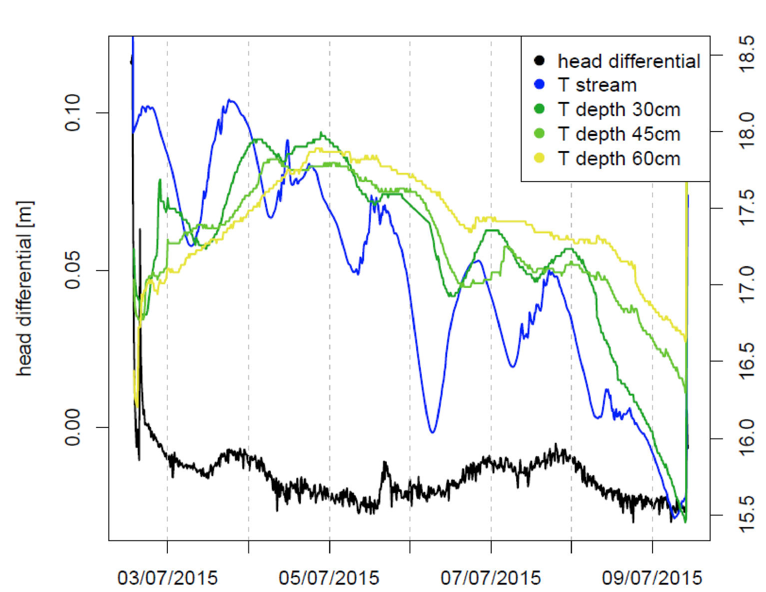
In this framework, parameters are treated as random variables. Using pressure measurements and heat as a tracer of the flow, we seek to characterize their statistical probability distribution. One benefit is that this framework handles non-sinusoidal timeseries. The posterior distributions can then be used in a Monte-Carlo framework in order to simulate uncertainty-quantified stream-aquifer exchanges timeseries.

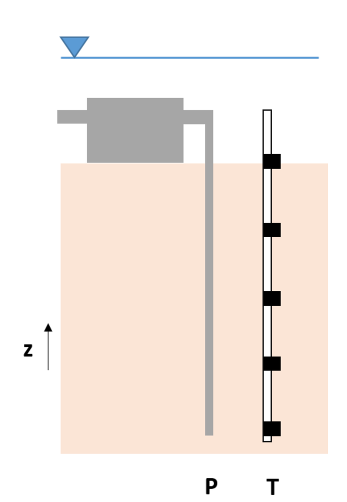


Available data

- 5 vertically-distributed temperature timeseries
- 1 hydraulic head differential timeseries

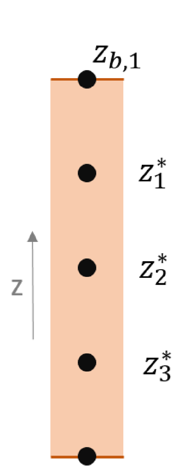
15-min sampling period





Avenelles basin, France - Mouhri et al. (2013)

Target parameters and forward model



We use a one-dimensional finite volume model (Ginette) to simulate water and heat exchanges.

Boundary conditions

- Stream and deepest temperature timeseries
- Head differential timeseries

Data for inversion (model output)

- Other temperature timeseries (Type-B)

$$\frac{\partial T}{\partial t} = -\frac{\rho_w c_w}{n\rho_w c_w + (1-n)\rho_s c_s}\, q\, \frac{\partial T}{\partial z} + \frac{\left(n\sqrt{\lambda_w} + (1-n)\sqrt{\lambda_s}\right)^2}{n\rho_w c_w + (1-n)\rho_s c_s}\frac{\partial^2 T}{\partial z^2}$$

$$q = -\frac{k\,\rho_w g}{\mu}\frac{dH}{dz}$$

$\mathbf{y}$: target parameter vector
$\mathbf{y} = (n, k, \lambda_s, c_s, \rho_s)$

Bayesian inversion algorithm

Structural model definition

We characterize the physical parameters $y = (n, k, \lambda_s, c_s)$ by a statistical distribution function,

ie. we know what family of distribution represent the distribution of y in the field.

$$n \sim \mathcal{TN}(\mu_n, \sigma_n^2, 0.01, 0.99)$$
$$\log(k) \sim \mathcal{N}(\mu_k, \sigma_k^2)$$
$$\lambda_s \sim \mathcal{TN}(\mu_\lambda, \sigma_\lambda^2, 1, 100)$$
$$c_s \sim \mathcal{TN}(\mu_c, \sigma_c^2, 10, 100000)$$
$$\boldsymbol{\beta} = (\mu_n, \sigma_n^2, \mu_k, \sigma_k^2, \mu_\lambda, \sigma_\lambda^2, \mu_c, \sigma_c^2)$$

We seek to characterize $\boldsymbol{\beta}$, parameters of the statistical distribution, using information from measurements $\boldsymbol{z}^*$.

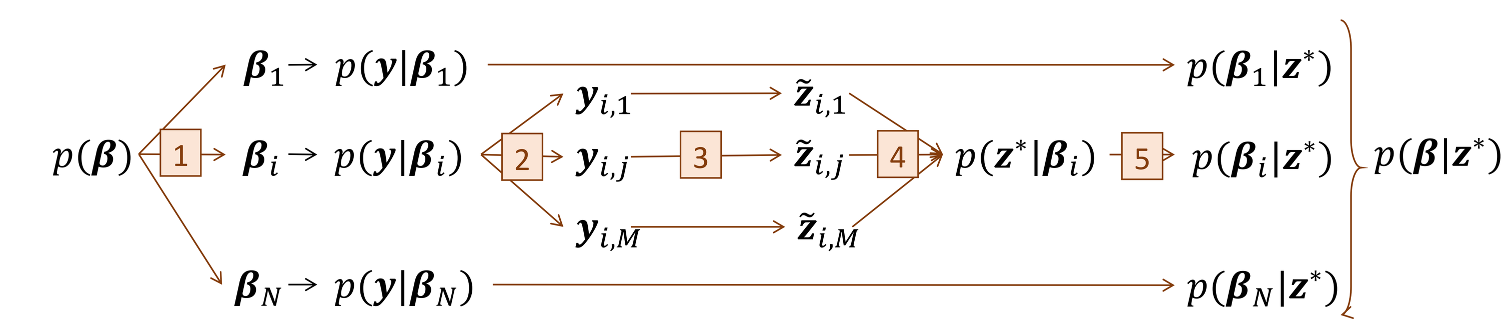
$$p(\boldsymbol{\beta}|\boldsymbol{z}^*) \propto p(\boldsymbol{z}^*|\boldsymbol{\beta}) \times p(\boldsymbol{\beta})$$

posterior ∝ likelihood × prior

We use the nested simulation procedure following the Method of Anchored Distributions (MAD) and implemented in MAD# (Osorio et al., 2015).

We choose to work with flat non-informative prior distributions. 1

1. Sample from prior	2. Sample from structural model	3. Run forward model	4. Evaluate likelihood	5. Evaluate posterior



$\boldsymbol{\beta}$ – structural parameters
$\boldsymbol{y}$ – physical parameters
$\tilde{\boldsymbol{z}}$ – model outputs at measurement locations
$\boldsymbol{z}^*$ – measurements

i – index of sample
j – index of realization

Osorio et al. (2015); Over et al. (2015); Rubin et al. (2010)

Testing the likelihood function on a synthetic case

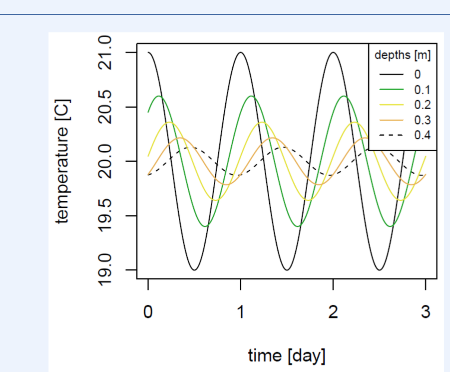
The analytical solution is the forward model. 3

$$\mathbf{T(z,t) = T_m + Ae^{-az}\cos(\omega t - bz)}$$

a, b are functions of $n, k, \lambda_s, c_s, \rho_s$.

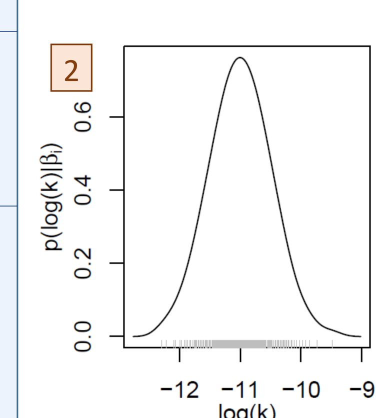
$n = 0.15$
$k = 10^{-11} m^2$
$\lambda_s = 2.3\ \mathrm{W\,m^{-1}K^{-1}}$
$c_s = 10^3\ \mathrm{J\,kg^{-1}K^{-1}}$
$\rho_s = 2.9 \cdot 10^3\ \mathrm{kg\,m^{-3}}$

Parameters of synthetic truth



Idealized timeseries used for the synthetic case study. 0 and 0.4m depths are used as boundary conditions of the model, the other timeseries are used in the inversion.

2

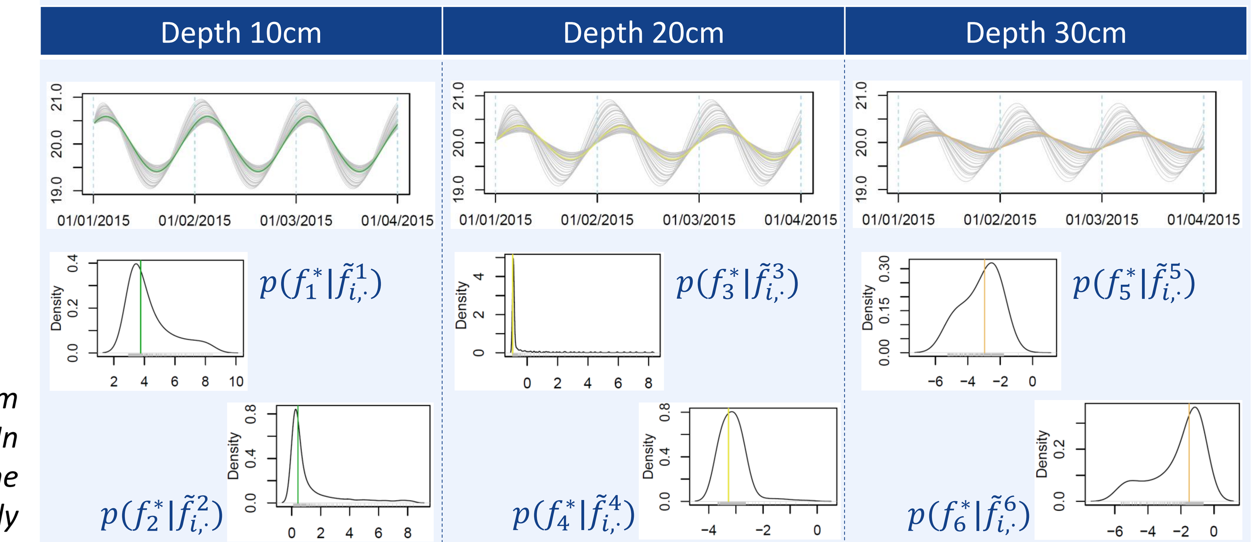


Sampling parameters from the structural model. In order to ensure that the realizations are equally plausible, we use latin hypercube sampling.

4 Evaluation of likelihood function at $\boldsymbol{\beta}_i$

$$p(\boldsymbol{z}^*|\boldsymbol{\beta}_i) = p(f^*|\tilde{f}^1_{i,\cdot}, \ldots, \tilde{f}^6_{i,\cdot})$$

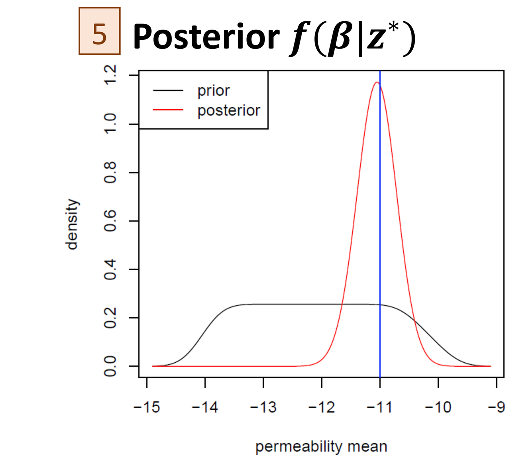
- Data reduction using 6 Fourier coefficients (2 per timeseries)
- Nonparametric density estimation



Fourier coefficients of realizations $\tilde{f}_{i,\cdot}$ (grey) and synthetic observation f^* (vertical line). The likelihood value is the density of the realizations $\tilde{f}_{i,\cdot}$ at the synthetic observation f^*.

Combining likelihoods for each Fourier coefficient yields to a likelihood value for β_i and to the posterior distribution.

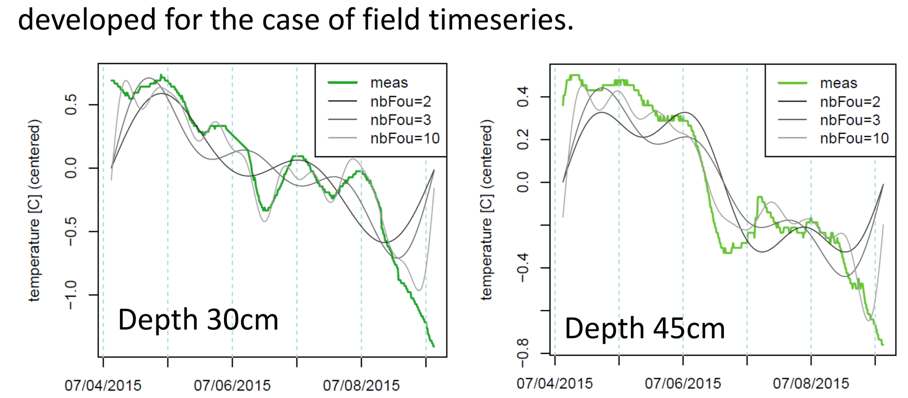
5 **Posterior $f(\boldsymbol{\beta}|\boldsymbol{z}^*)$**



Prior and posterior density functions for varying μ_m. The vertical line represents the underlying true parameter. The posterior peaks at the underlying true parameter.

Data reduction strategy for field temperature timeseries

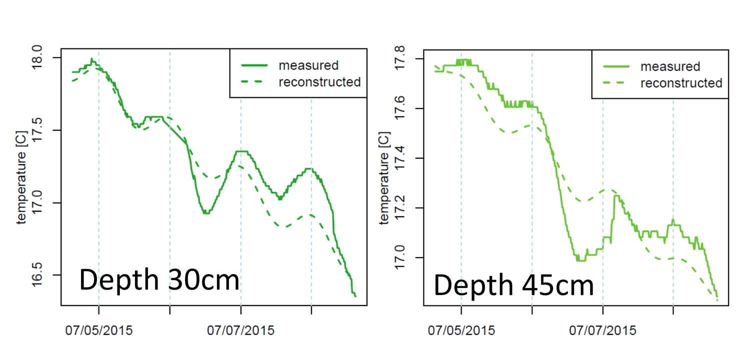
4 When working with field data, measurements exhibit a high dimensionality. For the purpose of nonparametric density estimation, it is recommended to work with a low number of dimensions, ideally less than 6 (Scott and Sain, 2004). While 2 Fourier coefficients per timeseries allow to reconstruct perfectly the shape of the timeseries for the analytical case, a different decomposition needs to be developed for the case of field timeseries.



Reconstruction of field temperature timeseries using the Fourier transformation. nbFou corresponds to the number of Fourier coefficients used in the decomposition. The shape of the timeseries is challenging to characterize using a Fourier decomposition only.

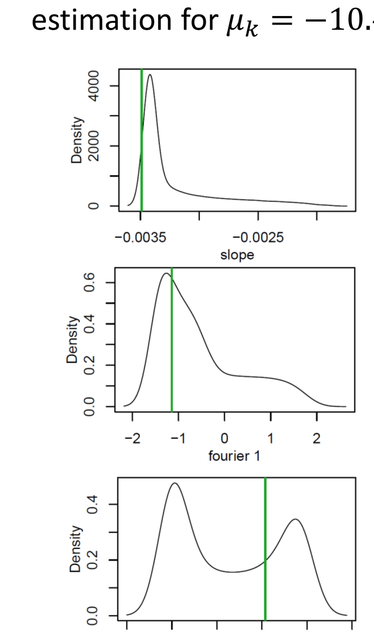
We decompose each timeseries using 3 coefficients:

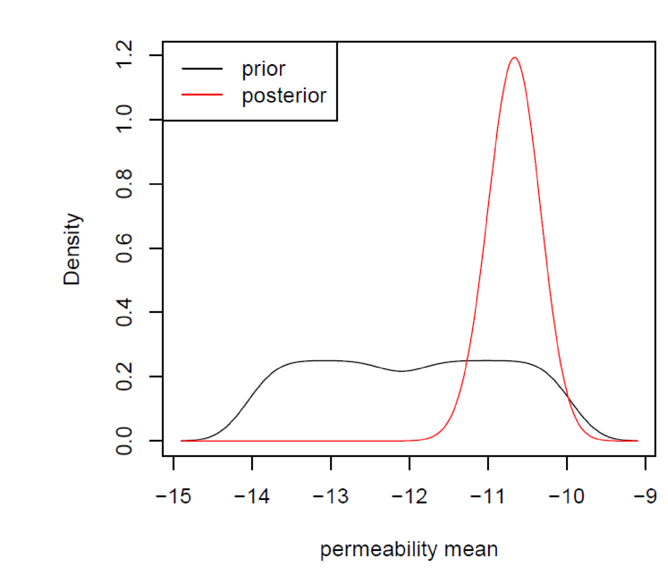
- 1 coefficient for the slope
- 2 Fourier coefficients of the residuals after removing a moving average



Reconstruction of the temperature timeseries using a combination of linear trend and Fourier coefficients.

Example of likelihood estimation for $\mu_k = -10.4$

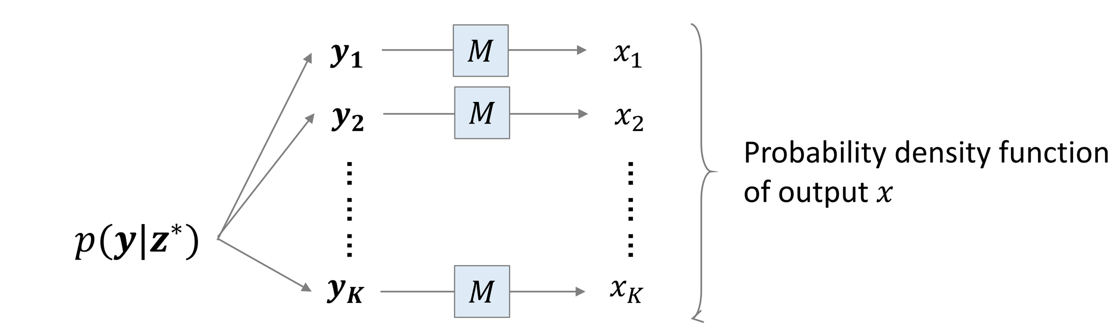




Prior and posterior density functions for varying μ_k. This time, field measurements were used to compute the likelihood, so there is no underlying true parameter.

Application: uncertainty-quantified stream-aquifer exchanges

The posterior distribution of physical parameters obtained from the inversion can be used in a Monte-Carlo analysis to obtain uncertainty-quantified model outputs. This analysis can be used to predict another variable of interest x that is an output from a physical model involving parameters $\boldsymbol{y}$.



$\boldsymbol{y}$ – physical parameters
$\boldsymbol{z}^*$ – measurements
x – other quantity that we want to predict
M – forward model

This strategy allows to fully take into account uncertainty in parameters $\boldsymbol{y}$ in the estimation of stream-aquifer exchanges.

Summary

- We present a data-driven framework for estimation of hyporheic hydrothermal properties. We use a combination of pressure and temperature measurements and use heat as a tracer of water exchanges.
- Thanks to the specification of a structural model, we don't need to make assumptions on the shape of the likelihood function.
- The synthetic study allows to test the algorithm for a low-dimensional timeseries. However, on field data, effective data reduction strategies are needed to keep the dimensionality of the likelihood function under 6.
- In the framework of a Monte-Carlo uncertainty analysis, physical parameters can be directly from the posterior distributions to estimate stream-aquifer exchanges and the associated uncertainty.
- One main challenge is that this algorithm is computationally expensive. We used parallel-computing to increase the computation time of the study.

References and Acknowledgments

Mouhri, A., Flipo, N., Rejiba, F., de Fouquet, C., Bodet, L., Goblet, P., Kurtulus, B., Ansart, P., Tallec, G., Durand, V., Jost, A. (2013). Designing a multi-scale sampling system of stream-aquifer interfaces in a sedimentary basin, *J. Hydrol.*, 504, 194–206, doi:10.1016/j.jhydrol.2013.09.036

Osorio-Murillo C. A., M. W. Over, H. Savoy, D. P. Ames, Y. Rubin (2015). Software framework for inverse modeling and uncertainty characterization, *Environmental Modelling & Software*, 66, 98-109, doi:10.1016/j.envsoft.2015.01.002.

Over, M. W., U. Wollschläger, C. A. Osorio-Murillo, and Y. Rubin (2015). Bayesian inversion of Mualem-van Genuchten parameters in a multilayer soil profile: A data-driven, assumption-free likelihood function, *Water Resour. Res.*, 51, 861–884, doi:10.1002/2014WR015252

Rubin, Y., X. Chen, H. Murakami, and M. Hahn (2010). A Bayesian approach for inverse modeling, data assimilation, and conditional simulation of spatial random fields, *Water Resour. Res.*, 46, W10523, doi:10.1029/2009WR008799

Scott, D. W., & Sain, S. R. (2004). Multidimensional Density Estimation. In Handbook of Statistics (Vol. 24, pp. 229–261). Elsevier Masson SAS. doi:10.1016/S0169-7161(04)24009-3

This work has been supported by the Jane Lewis Fellowship and the Chateaubriand Fellowship.