Abstract# 568188

Noninformative priors in hydrological models: is uniform prior truly a noninformative prior?

Abhinav Gupta[1] and Rao S. Govindaraju[1,2]

[1]Lyles School of Civil Engineering, Purdue University, West Lafayette, IN

[2]Bowen Engineering Head of Civil Engineering and Christopher B. and Susan S. Burke Professor of Civil Engineering

1. Background and Objectives

In order to use Bayes theorem in hydrological models, one needs to specify a likelihood function and a prior distribution. Many studies have addressed the problem of choosing an appropriate likelihood function to model the information content in the data. However, relatively fewer studies have addressed the problem of prior specification. Non-informative priors do not seem to have been addressed in the hydrologic modeling literature.

Generally, uniform prior are assumed to be non-informative priors; but the validity of this assumption is never questioned. This study has the following two objectives:

(1) To find out if the uniform priors over model-parameters are appropriate in hydrologic applications, and

(2) to explore the principle of maximum gain to determine non-informative priors.

Uniform prior over parameters may introduce bias in the inference.

2. Principle of maximum information gain to derive non-informative priors

Principle of maximum information gain (Bernardo, 1979) was used to derive non-informative priors. The main idea is as follows:

(1) The asymptotic posterior distribution of parameters represents the maximum information about the parameters that one can obtain through data;

(2) The relative entropy between prior and asymptotic posterior distributions quantifies the missing information about the parameters before any data are observed, and thus quantifies the information gain as one moves from prior to posterior;

(3) The main idea of principle of maximum information gain is to choose a non-informative prior such that the missing information between asymptotic posterior distribution and the prior distribution is maximized. The prior distribution thus obtained is defined as the "best" prior in this regard.

4. Conclusions

- Uniform prior may not always be the non-informative prior.
- Bernardo priors may be used to derive non-informative priors.

References

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference. Journal of the Royal Statistical Society. Series B (Methodological), 113-147.

Chow, V. T., Maidment, D. R., & Mays, L. W. (1988). Applied hydrology.

Stillman, J. S., Haws, N. W., Govindaraju, R. S., & Rao, P. S. C. (2006). A semi-analytical model for transient flow to a subsurface tile drain. Journal of Hydrology, 317(1-2), 49-62.

PURDUE UNIVERSITY

AGU 100 ADVANCING EARTH AND SPACE SCIENCE

3. Example Applications

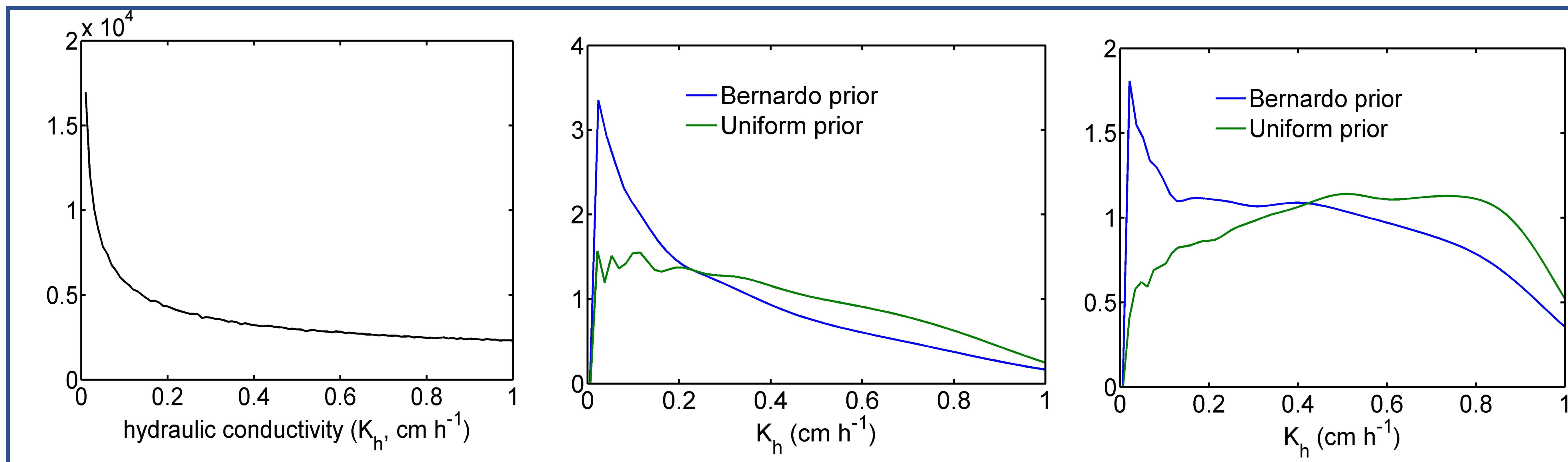

Fig. Bernardo prior over hydraulic conductivity, K_h, in Green-Ampt model (left), posterior distribution over K_h using Bernardo prior and uniform prior given that the true value of observed cumulative infiltration is 1.2 cm (middle) and 3.17 cm (right).

Example 1

- The noninformative prior obtained is not uniform
- Since only one observation was used, the posterior distributions is heavily affected by using two different priors, uniform and Bernardo.

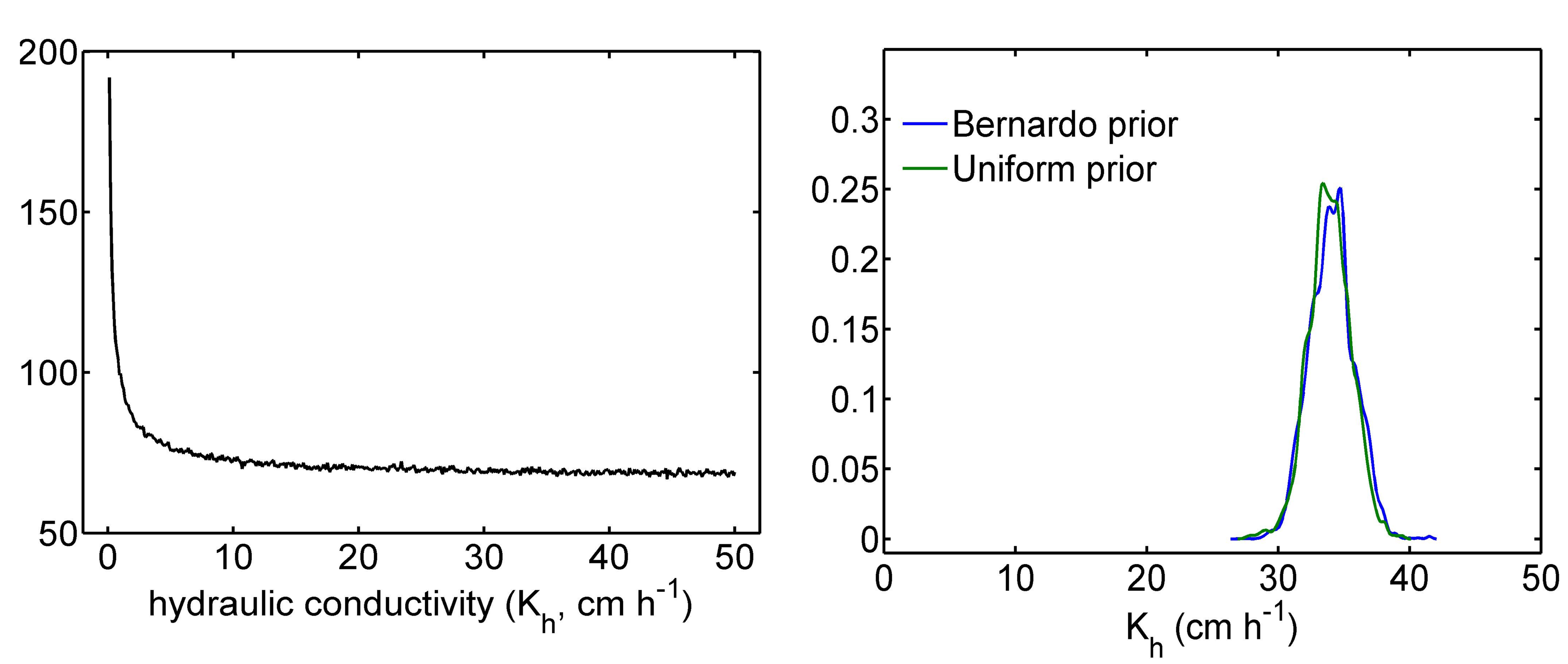

Fig. Bernardo prior over hydraulic conductivity, K_h, in falling head Green-Ampt model (left), posterior distribution over K_h using Bernardo prior and uniform prior (right).

Example 2

- The non-informative prior obtained is not uniform.
- After a value of approximately $10\ \mathrm{cm\ h^{-1}}$, the prior is equivalent to a uniform prior.
- The true value of K_h is $33\ \mathrm{cm\ h^{-1}}$; therefore, posterior distributions obtained by using Bernardo and uniform priors are similar.

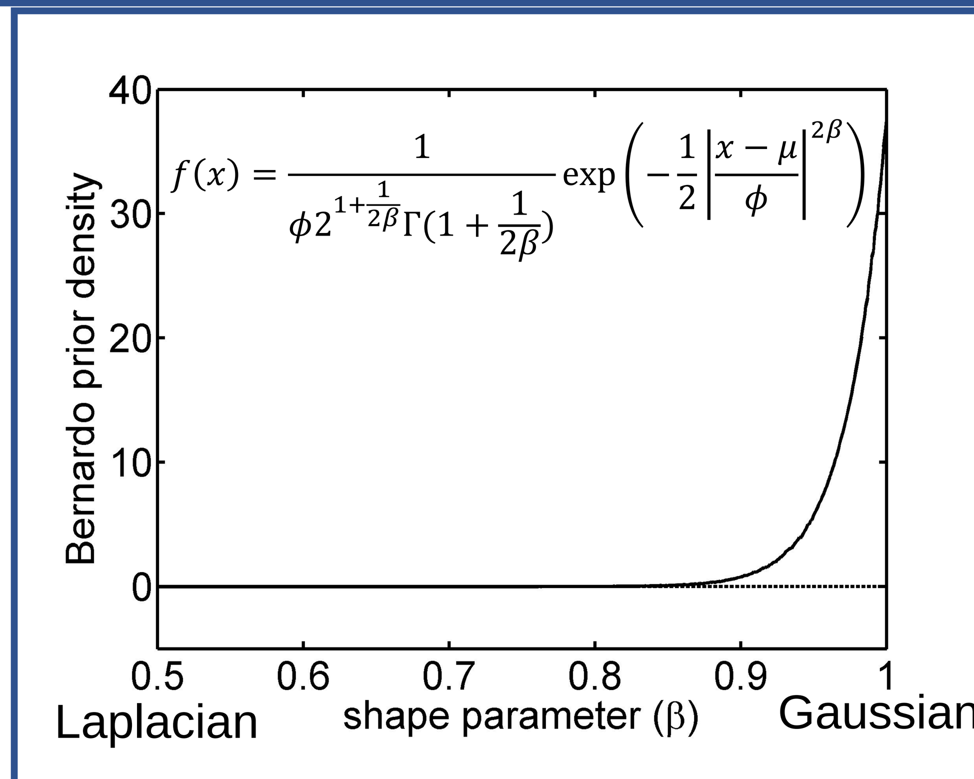

Fig. Bernardo prior over shape parameter, β, in generalized Gaussian distribution

Example 3

- Again, the uniform prior is not the best non-informative prior
- The non-informative prior seems to be biased towards a Gaussian distribution
- If one uses Bernardo prior to infer the value of β, a large amount of data would be required to infer a Laplacian distribution ($\beta = 0.5$), in line with the fact that the large amount of data are required to accurately compute high-order moments, i.e., characterize a Laplacian distribution.