Comment:
The reviewers looked through the revisions and were satisfied with most of the responses except the one regarding sigma in Eqn. (21). The equation and interpretation assume that RMSE should follow N(0,sigma^2). But the reviewer don't see how this can be derived from the assumption that errors follow N(0, sigma_e^2), given that RMSE is a nonlinear function of errors. Please respond before it can be finally accepted for publication.

Response:
The following shows the detailed derivation of Eq (21) from Eq (18).

$$\begin{aligned}\log L(\boldsymbol{y}|\boldsymbol{X}) &= -\sum_{i=1}^{N}\frac{\left(y_i-\mathcal{M}_i(\boldsymbol{X})\right)^2}{2\sigma^2}-\frac{N}{2}\log(2\pi\sigma^2)\\ &= -\frac{N\times\frac{1}{N}\sum_{i=1}^{N}\left(y_i-\mathcal{M}_i(\boldsymbol{X})\right)^2}{2\sigma^2}-\frac{N}{2}\log(2\pi\sigma^2)\\ &= -\frac{N\times RMSE^2}{2\sigma^2}-\frac{N}{2}\log(2\pi\sigma^2)\\ &= -\frac{N(0-RMSE)^2}{2\sigma^2}-\frac{N}{2}\log(2\pi\sigma^2)\end{aligned}$$

where σ represents the standard deviation of the error between simulation and reference runoff based on the following assumption:

$$\epsilon_i \sim \mathcal{N}(0,\sigma^2), i = 1, 2, \dots, N$$

Here subscript *i* represents *ith* month from 1991-2010, $N = 240$, and the underlying assumption is the error of simulated runoff has the same distribution for each month. It is not clear how to estimate σ with runoff data because model exhibits different error distributions for different months. Specifically, Figure 1 shows that the σ estimated with model simulations and observation for all the 240 months (i.e., 20 year time series) varies significantly. However, focusing on Eq (21), the error term can be further assumed as the difference between RMSE and 0 as RMSE is selected as QoI, and one can consider RMSE is the model simulation (i.e., $\mathcal{M}(\boldsymbol{X}) = RMSE$) and 0 is the observation (i.e., $y = 0$). So, we estimate σ in Eq (21) using the standard deviation of difference between the simulated RMSEs and 0. In this way, we include runoff data from all months for estimating σ. We acknowledge there can be other approximations for σ, but our estimation yields reasonable posterior of parameters and runoff. It is beyond the scope of this study to investigate the impacts of σ on the parameter inference. We added the following statement in Line 221 – Line 227 to clarify this issue.

"The standard deviation of error between model simulated runoff and observation exhibits significant monthly variation. To provide a reasonable value of σ, we further assume σ in Eq (21) has a different meaning than that in Eq (18) by taking RMSE as model simulation, and 0 to as the target. Therefore, σ is approximated as the standard deviation of the difference between 0 and RMSEs, where each RMSE was calculated using simulated runoff and observation for a given training simulation. Our estimation of σ leads to a reasonable posterior (see Sec 5.4), though other methods can also be used to estimate σ. We acknowledge that the value of σ may have an impact on the parameter posterior, but investigating the sensitivity of σ on the posteriors is beyond the scope of this study."

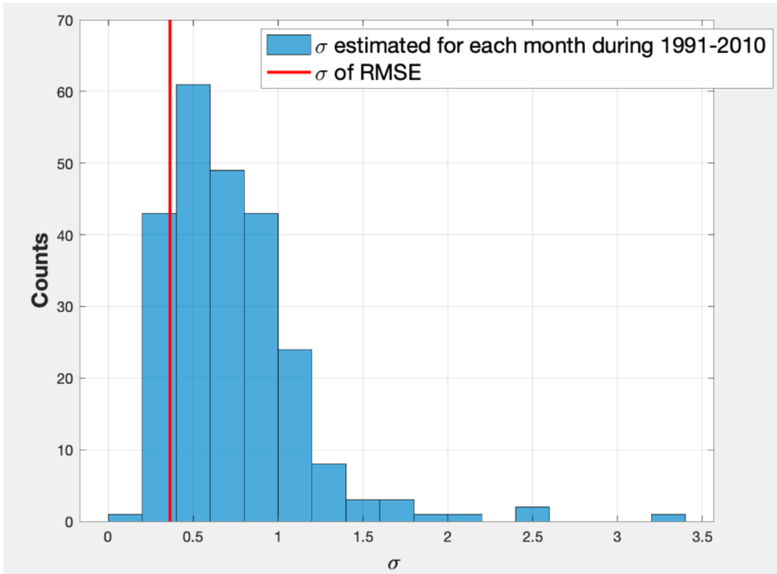

Figure 1. Histogram of σ estimated from error between simulated runoff and GRUN for each month during 1991-2010 at the example grid cell (e.g., Figure 6 in main text). The red line represents the σ estimated with the standard deviation of RMSEs used in this study.