ON THE INFLUENCE OF ERROR MODEL IN THE GOOD PERFORMANCE OF THE HYDROLOGICAL MODEL FOR THE RIGHT REASONS

HERNÁNDEZ M.R. (1) AND FRANCÉS F. (1)

(1): Research Institute of Water and Environmental Engineering, Universitat Politècnica de València, Spain

Hydrological models provide extrapolations or predictions, which are not lacking of uncertainty. One phase of the hydrological implementation process, significantly contributing to that uncertainty, is the calibration phase in which values of the unknown model parameters are tuned by optimizing an objective function. Traditionally, the most commonly used fitting criterion has been the simple least squares (SLS), regardless of the SLS criterion involves strictly assumptions about the probability distribution of the errors. Failure of these assumptions introduces noise into the estimation of the parameters. In the present research it has been carried out an estimation of the parameters of TETIS in an exercise of Bayesian inference, with three error model assumptions. The main motivations for the development of this work are: (1) verify whether and to what extent a distributed conceptual model calibrated with the SLS criterion become a data-driven model, yielding sets of unrealistic parameter values but with outstanding performance, (2) contrast model divergence phenomenon between the inference layouts, and (3) compare the reliability of the uncertainty assessment for each of the error models. The analysis of the results suggests that to achieve the objective of having a calibrated hydrological model, which works well for the right reasons, it is necessary to draw the inference of its parameters using an appropriate error model, and the traditional SLS calibration criterion, is not adequate for this purpose.

INTRODUCTION

For at least the last 50 years, there has been a significant effort in modeling the hydrologic cycle and understanding its processes. Data-driven models disregard de facto the results of these research efforts, in contrast with the Bayes principle which would combine the observations with all possible a priori knowledge on the hydrological processes and possibly on the parameter values to obtain less uncertain a posteriori forecasts (E. Todini [1]). Today is recognized both data-driven and conceptual models have advantages and disadvantages, nevertheless the conceptual models should not be used as if they were data-driven. Hydrologic knowledge is typically embodied in a deterministic catchment model (R. Krzysztofowicz [2]). However, this statement is meaningless if the model parameters yield values that are far from their physical meaning.

Moreover, hydrological models provide extrapolations or predictions, which are not lacking of uncertainty. One phase of the hydrological implementation process, significantly contributing to that uncertainty, is the calibration phase in which values of the unknown model parameters are tuned by optimizing an objective function. Traditionally, the most commonly used fitting criterion has been the simple least square (SLS) residual errors, regardless of the SLS criterion involves strictly assumptions about the probability distribution of the errors. Failure of these assumptions introduces noise into the estimation of the parameters, which leads to the phenomenon called model divergence (S. Sorooshian, J. Dracup [3]), where the errors variance of the (spatially and temporally) forecasted flows, far exceeds the errors variance obtained in the fitting period.

In the present work it has been carried out an estimation of the parameters of TETIS, a conceptual distributed hydrological model with a particular split structure of the effective model parameters (F. Francés et al. [4]). Such an estimate has been performed with the aid of a Markov Chain Monte Carlo (MCMC) algorithm called Dream-ZS (E. Laloy, J. Vrugt [5],G. Schoups, J. Vrugt [6]). MCMC algorithm quantifies the uncertainty of the parameters by getting the posterior probability distribution, conditioned on the observed flows. The calibration process is performed with three error model assumptions wherein each one attempts to improve the previous.

The main motivations for the development of this work are: (1) verify whether and to what extent a distributed conceptual model calibrated with the SLS criterion become a data-driven model, yielding sets of unrealistic parameter values but with outstanding performance, (2) contrast model divergence phenomenon between the inference layouts, and (3) compare the reliability of the uncertainty assessment for each of the error models.

METHODS

Bayesian Inference

Model parameters inferences are based on a likelihood function that quantifies the probability that the observed data were generated by a particular parameter set (G. Box, G. Tiao [7]). To derive the appropriate form for the likelihood function, first it must have been hypothesized the functional form of the joint probability density function of the residual errors, namely the statistical model for the residual errors. Traditionally it has been assumed that the errors are independent and identically distributed according to a normal distribution with zero mean and a constant variance, which results in SLS approach for parameter estimation. Nevertheless, violation of the SLS assumptions may introduce bias in estimated parameter values and affect parameter and predictive uncertainty (M. Thyer et al. [8]) making the model itself and its parameters lose their physical meaning.

A key question is to understand why an inference process based on a wrong error model assumption can yield biased parameter values. For example, assumptions made by the SLS method are only valid when residuals resulting from the inference process are exclusively due to random measurement errors (Gaussian white noise). However, residuals of hydrologic models can arise from other sources of uncertainty, in addition to the measurements. In fact, model inputs (mainly forcing data) and model structural errors are the most influential sources of uncertainty. When the error model doesn’t include all the possible effects of all uncertainty sources in an explicit way, the projection of forcing and structural uncertainty onto the model parameters will occur, corrupting its “true value” and converting the hydrologic model into something similar to a data-driven model. The use of a formal Bayesian inference method has the advantage of being able to verify (a posteriori) the hypotheses assumed in the error model.

Inference Layouts

Three layouts of inference were defined (Table 1). The first layout (E1) is based on the same assumptions for the residuals, as SLS method (namely normal residuals i.i.d.). Second and third layouts are based on the general likelihood function developed in (G. Schoups, J. Vrugt [6]) which allows modeling with correlated, heteroscedastic and non-Gaussian errors. Both layouts E2 and E3, consider the residuals auto-correlation through a second order autoregressive linear model AR(2), and those residuals can exhibit nonstationary variance (linear regarding the simulated flow), whereas standardized value of the innovations are described by a skew exponential power (SEP) density distribution. Finally, the difference between error models E2 and E3 is that the latter explicitly considers the existence of a bias in the residuals of the hydrological model, the mean of which departs systematically from the zero value. Thus on the E3 layout, bias model is merely described by a double linear formulation: are defined two different linear models, according to the value of the simulated flow be less or greater than a flow threshold. The value of this threshold is another parameter to be calibrated.

Table 1. Inference Scenarios

ID	RESIDUALS DISTRIBUTION	RESIDUALS VARIANCE	RESIDUALS AUTOCORRELATION	BIAS	Nº PARAMETERS (HM + EM)
E1	GAUSSIAN	HOMOSCEDASTIC	NO	NO	9 + 1
E2	SEP	HETEROSCEDASTIC	AR(2)	NO	9 + 7
E3	SEP	HETEROSCEDASTIC	AR(2)	YES	9 + 12

Computational Bayesian Algorithm

The Differential Evolution Adaptive Metropolis algorithm Dream-ZS has been used in the present work to sample the posterior distribution of the parameters. Original DREAM (J. Vrugt et al. [9]) is an efficient MCMC sampler that simultaneously runs multiple Markov chains. During the process the whole parameter space is explored and the algorithm automatically adjusts the scale and orientation of the proposal distribution. Dream-ZS algorithm is a modification of DREAM algorithm. It uses an archive of past states to generate candidate points in each individual chain. Such sampling is more efficient than optimal random walk Metropolis (C. Ter Braak, J. Vrugt [10]), maintains efficient convergence and needs less parallel Markov Chains than the original DREAM. The R-statistic (A. Gelman, D. Rubin [11]) is used to check whether the chains have converged (threshold value of R is 1.2). In this research, the number of chains used in all three inference configurations was 10.

RESULTS AND DISCUSSION

Case Study

In this study, it has been used the conceptual distributed watershed model TETIS (F. Francés, I. Vélez, J. Vélez [4],M. Puricelli [12]) which has nine parameters that need to be estimated. Inputs to the model include time series at defined station points (observed discharge, rainfall, evapotranspiration). TETIS model uses the inverse distance method to interpolate spatially temporal series. Cartographic information uses raster format maps. Digital Terrain Model (cell size selected was 500 m×500 m) and soil properties (available water and saturated hydraulic conductivities) are obtained based on soils studies, land use, geological maps, edaphologic information, hydrogeological data and other environmental topics that could be interesting and are available at the study area.

The split-parameter structure of TETIS, refers to the way the model estimate the parameter maps. Initially, the maps are estimated a priori using environmental and available information. Then, correction factors (also named regularization factors in P. Pokhrel, H. V Gupta [13]) are used to modify globally the previously estimated maps. In this way, the spatial variability captured in the initial estimated maps is kept.

It has been used historical daily data from Oria River, at the Basque Country Region (Spain). The average annual rainfall in this hydrologic unit is 1534 mm, the average temperature of 13.0 º C, and the annual PET is 837 mm. The most common lithology is alternating detritic strata. Regarding vegetation, grasslands and forests cover much of the surface (almost 60%). Control point of the analyzed Oria subcatchment is named “C5Z1” (337 km^2). For this subcatchment, 2 years of data were used for model calibration (October 1, 1998 to September 30, 2000), with an additional “warm up” period of one year. For the model validation were employed 3 years of data (October 1, 1995 to September 30, 1998), with an additional “warm up” period of one year.

Fulfillment of hypothesis about errors

After having carried out the inference process, is necessary to check the fulfillment of the hypothesis about the error model assumed a priori. Figure 1 displays the observed and theoretical distribution of errors. Only layouts E3 and E2 (E2 not shown) fulfill the distribution hypothesis. With regard to the hypothesis about the error variance and errors dependence, the same fact occurs (results not shown).

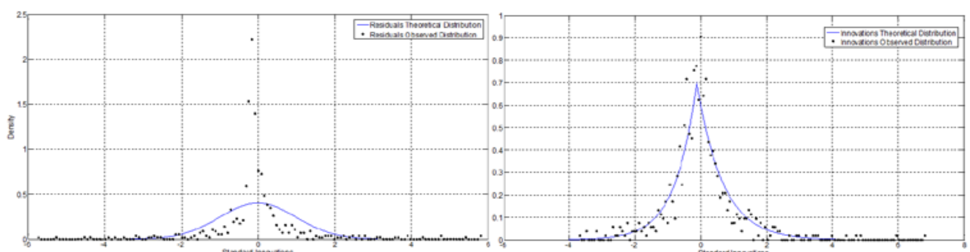

Figure 1. Residuals distribution in E1 layout (left) and Innovations distribution in E3 layout (right)

“Credibility” of the best point parameter estimates

The analysis of the results indicates important differences in the posterior distribution, as well as in the best point estimation of parameter values, between the three error models. Similarly can be appreciated a great difference in terms of water balance. We set the focus on the F1 parameter which is the corrector factor of the soil water storage capacity, a key parameter in the model for separating effective rainfall into evapotranspiration, runoff and infiltration. Parameter estimate under E1 layout, lead to nonphysical value (figure 2). So, although E1-calibrated model it’s the best fitted model, it's not for the right reasons.

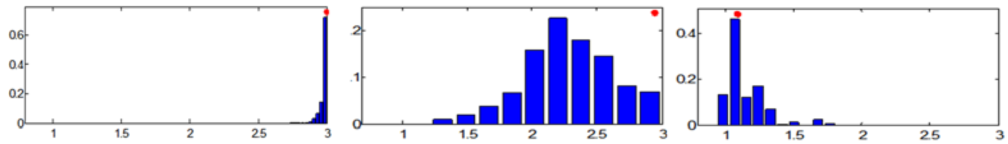

Figure 2. Posteriors of Parameter F1. From left to the right: E1, E2 and E3

The model divergence phenomenon

Regarding the aforementioned model divergence phenomenon, Table 2 shows how the performance of hydrological model in layouts E1 and E2 are the best for the calibration scenario (NSE 0.93 and 0.91 respectively). The performance worsening in validation scenario for the three layouts seems to be similar in terms of NSE (around -6%). Nevertheless if we focus on the RMSE, the worsening in E1 is much greater than in E3 (+33% versus +8% respectively). In the same way, if we focus on the volume relative error (VE) we can see an increment in E1 and E2 layouts, whereas the E3 layout shows that VE decreases. It can be concluded that divergence exists in E1 and E2 layouts, whereas that E3 layout performance in validation scenario seems to be solely affected by peak flows (to which the NSE is very sensitive) rather than by medium or low flows.

Table 2. Performance of the Hydro Model, Mean Prediction and Predictive Distribution

			CALIBRATION SCENARIO			VALIDATION SCENARIO		
			E1	**E2**	**E3**	**E1**	**E2**	**E3**
HYDRO MODEL		**NSE**	0.93	0.91	0.79	0.863	0.852	0.745
		RMSE	2.62	2.97	4.42	3.48	3.62	4.76
		VE (%)	2.4	3.1	10.6	-4.5	-6	3.9
MEAN PREDICTION		**NSE**	0.93	0.91	0.91	0.863	0.852	0.853
		RMSE	2.62	2.97	2.86	3.48	3.62	3.611
		VE (%)	2.4	3.1	-0.7	-4.500	-6.000	-5.3
PREDICTIVE DISTRIBUTION	**RELIABILITY**	**ALFA**	0.77	0.86	0.92	0.77	0.89	0.91
		ETA	0.98	0.99	0.99	0.96	0.93	0.96
	RESOLUTION (sharpness)		2.7	4.6	5.3	2.2	4.8	5.4
	95% UNCERTAINTY BANDS	**%OBS INSIDE**	94	94	95	93	86	89
		MAX. BANDWITH	10	197	127	10	242	158
		MIN. BANDWITH	10	0.3	0.2	10	0.3	0.1
		MEAN BANDWITH	10	10	8	10	8	7

Reliability of the Predictive Distribution

Since hydrological models provide extrapolations or predictions, which are not lacking of uncertainty, is essential to perform an estimate of such uncertainty in a reliable way. In the elaboration of that assessment, the assumed error model is the key.

Figure 3 displays the evolution in the refinement of the 95% confidence bands, from the E1 to E3 error model. Moreover, Table 2 evidences that E2 full predictive distribution is better than E1 one, and similarly E3 outperforms E2. The concepts of reliability and resolution of the predictive distribution can be revised in F. Laio, S. Tamea [14], B. Renard et al. [15].

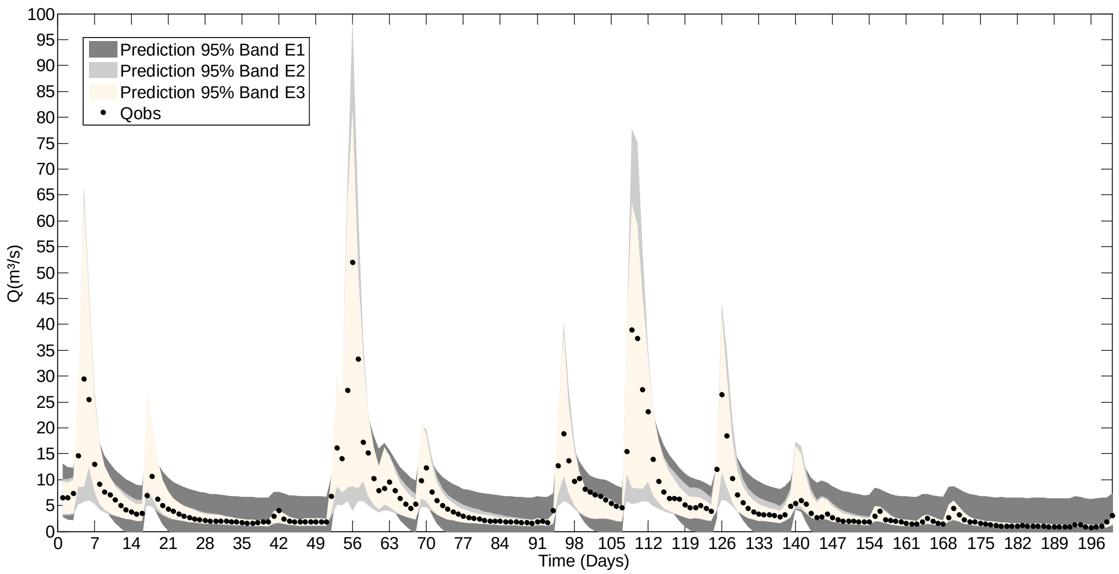

Figure 3. Reliability of Prediction 95% Confidence Bands. First Day: December 24, 1999

CONCLUSIONS

In the present work it has been carried out an estimation of the parameters of TETIS, a conceptual distributed hydrological model. Such an estimate has been performed with the aid of a MCMC algorithm called Dream-ZS. The inference process is performed with three error model assumptions wherein each one attempts to improve the previous one.

The first and already known finding is that SLS criterion (E1 layout) is not a valid calibration criterion for the hydrological model parameters inference, and its calibrated model is clearly a data-driven model. Moreover, any error model is not good to fulfill the goal of achieving unbiased (or slightly biased) parameter values. For example, the E2 layout is better than E1 layout. Nevertheless, E2 also produces bias in the parameters, which is observed through the model divergence phenomenon.

Furthermore even the E3 error model, which complies in a higher degree the assumed hypothesis, and exhibits a more reliable predictive distribution, surely is not the most appropriate for the estimated hydrologic model. However, be closer to the best possible error model, allows to obtain parameter values giving a hydrologic model that works as it is, not as a data-driven model.

The authors believe that it is important that the model, works better or worse in a given watershed, do it for the right reasons. That is the key to identify the presence of possible problems in the model, as well as the key to understand the true hydrological performance of the analyzed system.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the Spanish Ministry of Economy and Competitiveness through the research projects ECO-TETIS (CGL2011-28776-C02-01) and SCARCE (CSD2009-00065). Jasper Vrugt is acknowledged for providing the code for the Dream-ZS-algorithm.

REFERENCES

[1] E. Todini, Hydrological catchment modelling: past, present and future, *Hydrol. Earth Syst. Sci.* 11 (2007) 468–482. doi:10.5194/hess-11-468-2007.

[2] R. Krzysztofowicz, Bayesian theory of probabilistic forecasting via deterministic hydrologic model, *Water Resour. Res.* 35 (1999) 2739–2750.

[3] S. Sorooshian, J. Dracup, Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: Correlated and heteroscedastic error cases, *Water Resour. Res.* (1980).

[4] F. Francés, I. Vélez, J. Vélez, Split-parameter structure for the automatic calibration of distributed hydrological models, *J. Hydrol.* 332 (2007) 226–240. doi:10.1016/j.jhydrol.2006.06.032.

[5] E. Laloy, J. Vrugt, High-dimensional posterior exploration of hydrologic models using multiple-try DREAM (ZS) and high-performance computing, *Water Resour. Res.* 48 (2012) 1–18. doi:10.1029/2011WR010608.

[6] G. Schoups, J. Vrugt, A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors, *Water Resour. Res.* 46 (2010) 1–17. doi:10.1029/2009WR008933.

[7] G. Box, G. Tiao, *Bayesian inference in statistical analysis*, Wiley, 1992.

[8] M. Thyer, B. Renard, D. Kavetski, G. Kuczera, S.W. Franks, S. Srikanthan, Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis, *Water Resour. Res.* 45 (2009) n/a–n/a. doi:10.1029/2008WR006825.

[9] J. Vrugt, C. Ter Braak, C. Diks, B. Robinson, J. Hyman, D. Higdon, Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling, …*Numerical Simul.* 10 (2009) 271–288. doi:10.1515/IJNSNS.2009.10.3.273.

[10] C. Ter Braak, J. Vrugt, Differential Evolution Markov Chain with snooker updater and fewer chains, *Stat. Comput.* 18 (2008) 435–446. doi:10.1007/s11222-008-9104-9.

[11] A. Gelman, D. Rubin, Inference from iterative simulation using multiple sequences, *Stat. Sci.* (1992).

[12] M. Puricelli, Parameter extrapolation to ungauged basins with a hydrological distributed model in a regional framework, *Hydrol. Earth*. (2009) 229–246.

[13] P. Pokhrel, H. V Gupta, On the use of spatial regularization strategies to improve calibration of distributed watershed models, *Water Resour. Res.* 46 (2010) 1–17. doi:10.1029/2009WR008066.

[14] F. Laio, S. Tamea, Verification tools for probabilistic forecasts of continuous hydrological variables, *Hydrol. Earth Syst. Sci.* 11 (2007) 1267–1277. doi:10.5194/hess-11-1267-2007.

[15] B. Renard, D. Kavetski, G. Kuczera, M. Thyer, S.W. Franks, Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, *Water Resour. Res.* 46 (2010) W05521. doi:10.1029/2009WR008328.