

## Quantile hydrologic model selection and model structure deficiency assessment: 2. Applications

Saket Pande<sup>1</sup>

Received 12 October 2012; revised 5 July 2013; accepted 13 July 2013; published 18 September 2013.

[1] Quantile hydrologic model selection and structure deficiency assessment is applied in three case studies. The performance of quantile model selection problem is rigorously evaluated using a model structure on the French Broad river basin data set. The case study shows that quantile model selection encompasses model selection strategies based on summary statistics and that it is equivalent to maximum likelihood estimation under certain likelihood functions. It also shows that quantile model predictions are fairly robust. The second case study is of a parsimonious hydrological model for dry land areas in Western India. The case study shows that an intuitive improvement in the model structure leads to reductions in asymmetric loss function values for all considered quantiles. The asymmetric loss function is a quantile specific metric that is minimized to obtain a quantile specific prediction model. The case study provides evidence that a quantile-wise reduction in the asymmetric loss function is a robust indicator of model structure improvement. Finally a case study of modeling daily streamflow for the Guadalupe River basin is presented. A model structure that is least deficient for the study area is identified from nine different model structures based on quantile structural deficiency assessment. The nine model structures differ in interception, routing, overland flow and base flow conceptualizations. The three case studies suggest that quantile model selection and deficiency assessment provides a robust mechanism to compare deficiencies of different model structures and helps to identify better model structures. In addition to its novelty, quantile hydrologic model selection is a frequentist approach that seeks to complement existing Bayesian approaches to hydrological model uncertainty.

**Citation:** Pande, S. (2013), Quantile hydrologic model selection and model structure deficiency assessment: 2. Applications, *Water Resour. Res.*, 49, 5658–5673, doi:10.1002/wrcr.20422.

### 1. Introduction

[2] Quantile regression has been extensively applied [Koenker and Basset, 1978; Koenker, 2005; Ma and Koenker, 2006; Keyzer and Pande, 2009] in parametric and nonparametric statistics. Quantile regression has also been applied in context of hydrologic forecasting [Weerts *et al.*, 2011]. A particular case of quantile model selection in the form of model selection based on flow duration curves (FDCs) has also been extensively studied in hydrology [Yu and Yang, 2000; Son and Sivapalan, 2007; Westerberger *et al.*, 2011; Blazkova and Beven, 2009].

[3] However, the extension of quantile regression to hydrological model selection is nontrivial. For example, in context of hydrological forecasting (such as Weerts *et al.* [2011]), linear quantile forecasts make one critical assumption

that a forecasting model (or parts of it) can be linearized at a point in space or time. Such an assumption has several other implicit subassumptions on differentiability of the forecasting model and allowed perturbations that may not hold when the forecasting temporal range is large or when the temporal resolution of the forecasting model is coarse. This may lead to inaccurate quantile predictions with the quantile estimates of the parameters no more meaningful than being partial derivatives of the nonlinear forecasting model.

[4] Pande [2013] provides a theoretical foundation for its extension to hydrological model selection. The study reveals some interesting properties based on which structural deficiencies can be diagnosed and structure improvements can be assessed. A quantile (specific) model is obtained by minimizing a quantile specific loss function. The loss function is called the asymmetric loss function. The quantile model can then be used to predict the observed quantile. However, it is not limited to quantile model selection and prediction. Model structure deficiencies can lead to the crossing of two quantile model predictions because of the biases in predicting observed quantiles that structure deficiencies introduce. The optimal value of loss function at a quantile contains full information of the bias and thus structure deficiency at that quantile. The loss function values of a set of candidate model structure at a

Companion to Pande [2013], doi:10.1002/wrcr.20411

<sup>1</sup>Department of Water Management, Delft University of Technology, Delft, Netherlands.

Corresponding author: S. Pande, Department of Water Management, Delft University of Technology, NL-2628CN Delft, Netherlands. (s.pande@tudelft.nl)

given quantile thus order the structures in terms of its deficiency at that quantile. The asymmetric loss function across quantiles can therefore be used to assess deficiencies of a set of candidate structures.

[5] Quantile hydrological model selection (or estimation) encompasses some standard model selection techniques based on summary statistics, for e.g., based on the minimization of the mean absolute error. Quantile model selection at a quantile value of 0.5 provides a model that minimizes the mean absolute error. In some cases, quantile model selection is also “equivalent” to maximum likelihood model selection. A model selected by minimizing mean absolute error is equivalent to assuming a Laplace likelihood function. Model selection based on minimizing the mean square error (Ordinary Least Squares, OLS) (a “mean” model) is a weighted sum of models selected at quantiles in the neighborhood of (and including) 0.5, such as at 0.4, 0.5, and 0.6 [Koenker and Basset, 1978], though the assignment of weights is less formal. Yet another way is to estimate a mean model is to take the mean of quantile models with quantile values ranging between 0 and 1.

[6] The model selection strategy that is applied in this paper is unique in following aspects: (1) the method is applicable to time series of any variable of interest (flux or storage), (2) model selection is based on one quantile at a time, unlike strategies based on FDC that identify a model such that its FDC closely matches the observed FDC, (3) a loss function of Koenker and Basset [1978] based on absolute deviations is employed as an objective function that removes the need to identify quantiles of an observed time series, and (4) linearization of a hydrological model is not undertaken unlike linear quantile hydrological forecasting methods. This paper presents applications that demonstrate the methodology and the utility of the theory presented in Pande (2013).

[7] First, the nature of bias in predicting an observed quantile due to model structure deficiency is exhaustively studied alongside the performance of variants of quantile hydrologic prediction. A complex model structure and French Broad river basin data are used for this purpose. Quantile model selection method is then implemented for a case of a parsimonious hydrologic model developed for western India where it is shown that quantile model selection enables the detection of model structure improvement. Finally, several variants of the model structure used in the first case study are applied to the Guadalupe river basin. Multiple model structures are compared using quantile model selection and inferred model structure deficiencies are discussed.

[8] All the properties proved for a generic hydrologic quantile model selection problem (QE2 problem) in Pande [2013] also hold for the model structures of the applications, provided that assumptions 1–7 hold. Appendix A provides relevant definitions for QE2 specification and the assumptions.

[9] These assumptions are valid for the cases studied here. The set  $\mathbf{K}$  then represents the feasible space for model specific parameter set (the values that the parameter sets are allowed to take). Assumptions 1, 4, 5, and 7 hold for these cases (we allow for initial burn in which leads to positive initial storages in case of flexible model structures).

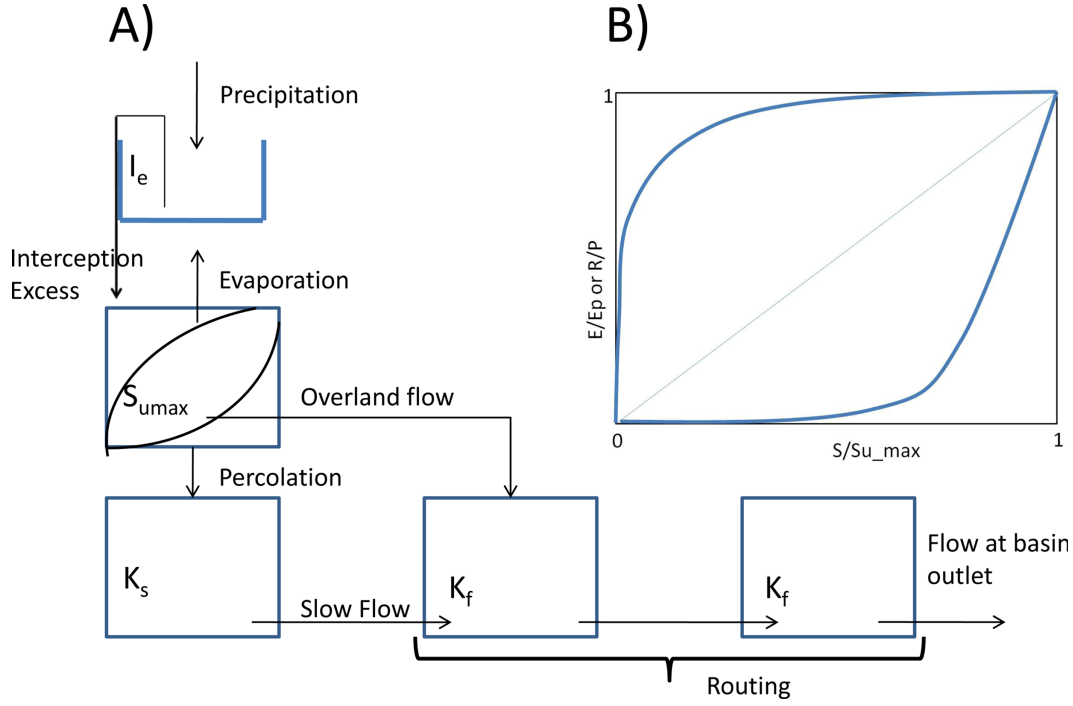
Assumptions 2 and 6 on differentiability can be relaxed without affecting the propositions (not done in the text to maintain brevity). If the functions in these assumptions are not differentiable, their gradients can be replaced by sub-gradients at the optimum. Continuity is still required, but all hydrologic models are continuous in parameters. Monotonicity in at least one parameter in assumption 2 holds in general, for example, flow is monotonic in recession parameters. The gradients of the predictive equation with respect to the parameters are independent of each other as required by Assumption 2. The only restrictive assumption may be that input forcings are nonzero for any time  $t$  (Assumption 3). It is shown to hold for the parsimonious dry land modeling case study, though not for the other two case studies. The sensitivity of propositions to this assumption is left for future work (at present it does not appear to contribute to major steps within the proofs). Different optimizers for each of the case studies are used, a gradient-based minimizer in the parsimonious model case study and a complex evolution-based minimizer in others. Both are initiated with multiple starting points to ensure that a solution is near a global minimum (Assumption 7). The parsimonious model case study is a constrained problem to tame the ill effects of model complexity while other case studies have reasonably large sample for calibration. Thus, the main result of Pande [2013] that the asymmetric loss function is a measure of model structure deficiency (and its utility to rank model structures based on asymmetric loss function at a particular quantile value) is not diluted by these more general case studies.

[10] The paper is organized as follows. Section 2 studies and validates the performance of quantile model selection on French Broad river basin data using a flexible model structure. Section 3 then assesses deficiency of a parsimonious dry land model developed for western India [Pande et al., 2010, 2012a]. Section 4 orders various model structures for the Guadalupe river basin in terms of structural deficiency. Finally, section 5 concludes the paper.

## 2. Performance of Quantile Model Selection

[11] The performance of quantile model selection using a flexible model structure is tested on the French Broad river basin streamflow data set. Quantile models are selected over different but overlapping calibration time periods to judge the robustness of the method. However, note that any such test assumes stationarity in streamflow. Model parameter distribution both over years and quantiles are studied. A split-sample test is also performed, wherein the performance of quantile models selected over a certain calibration period are tested on another nonoverlapping period. It is shown that the behavior of bias estimates across quantiles over different calibration periods yields insights into model structure deficiency, supporting its use in structure deficiency assessment as outlined in corollary 4 of Pande (2013).

[12] The model structure setup [Schoups and Vrugt, 2010] is composed of reservoirs to model interception, unsaturated zone, saturated zone, and river routing, referred here as “crr0” (Figure 1). Precipitation  $P(t)$  in excess of interception capacity ( $I_e$ ),  $P_e(t) = \max(P(t) - I_e, 0)$ , contributes to the unsaturated zone. Evaporation,  $E(t)$ , overland



**Figure 1.** Model structure “crr0” used in *Schoups and Vrugt* [2010] for the Guadalupe River Basin: (a) various process components of the model structure. It contains an interception reservoir with maximum capacity  $I_e$  (mm), interception excess precipitation flows into the unsaturated zone with storage capacity  $S_{u\max}$  (mm). The soil moisture storage in the unsaturated zone drives evaporation, overland flow, and percolation. Percolation is to the saturated zone, which in turn yields slow flow. Finally, slow flow and overland flow are routed by linear reservoirs to the outlet. (b) Nonlinear relationships governing overland flow and evaporation fluxes.

flow,  $R(t)$ , and percolation to the saturated zone are generated from the unsaturated zone as nonlinear functions of  $S_u(t)/S_{u\max}$ , where  $S_u(t)$  is the storage in the unsaturated zone and  $S_{u\max}$  is its storage capacity. Evaporation and overland flow are modeled as:

$$\begin{aligned} E(t) &= E_p [1 - \exp(-\alpha_E S_u(t)/S_{u\max})] / [1 - \exp(-\alpha_E)] \\ R(t) &= P_e(t) [1 - \exp(-\alpha_F S_u(t)/S_{u\max})] / [1 - \exp(-\alpha_F)] \end{aligned} \quad (1)$$

where parameters  $\alpha_E$  and  $\alpha_F$  are nonlinear controls and  $E_p$  is the potential rate of evaporation. Percolation ( $Q_p(t)$ ) is linearly related to  $S_u(t)/S_{u\max}$  as,

$$Q_p(t) = Q_{p\max} S_u(t)/S_{u\max}$$

[13] The slow flow,  $Q_s(t)$ , is a linear function of saturated zone storage,  $S_s(t)$ ,

$$Q_s(t) = S_s(t)/K_s$$

where  $K_s$  is the slow flow time constant.

[14] Finally, overland flow  $R(t)$  and slow flow  $Q_s(t)$  are routed through two linear reservoirs each with time constant  $K_f$ . Table 1 summarizes all the model quantities.

[15] The modeled flow at the outlet,  $Q_{\text{mod}}(t)$ , is estimated and subtracted from observed flows  $Q_{\text{obs}}(t)$  to obtain two types of absolute residuals,  $\{\tilde{\zeta}_t^+, \tilde{\zeta}_t^-\}$ , such that

$$\tilde{\zeta}_t^+ = \begin{cases} |Q_{\text{mod}}(t) - Q_{\text{obs}}(t)| & Q_{\text{mod}}(t) \leq Q_{\text{obs}}(t) \\ 0 & Q_{\text{mod}}(t) > Q_{\text{obs}}(t) \end{cases}$$

$$\tilde{\zeta}_t^- = \begin{cases} 0 & Q_{\text{mod}}(t) \leq Q_{\text{obs}}(t) \\ |Q_{\text{mod}}(t) - Q_{\text{obs}}(t)| & Q_{\text{mod}}(t) > Q_{\text{obs}}(t) \end{cases}$$

[16] A  $\tau$ -quantile specific model for a given model structure (but one model structure at a time; here it is “crr0”) is obtained by minimizing the asymmetric loss function  $\rho_\tau$ ,

$$\left[ \sum_{t=1}^N \left( (1-\tau)\tilde{\zeta}_t^+ + \tau\tilde{\zeta}_t^- \right) \right] / N, \text{ with respect to model parameters.} \quad (2)$$

[17] The Shuffled Complex Evolution global optimization algorithm of University of Arizona (SCE-UA) is implemented to minimize the objective function in equation (2). SCE-UA searches for a global optima by independently (but periodically shuffled) evolving  $m$  complexes each containing  $p$  parameter sets based on operations such as expansion, contractions, and reflection. Readers are referred to *Duan et al.* [1992] for additional details. For this study,  $m$  is fixed at 20,  $p = 41$  with a convergence criteria of 0.1% (change in objective function), and the search is terminated after 100,000 objective function evaluations if no convergence is achieved.

[18] Daily streamflow, precipitation, and potential evapotranspiration data are used for assessing the performance of quantile model selection (estimation). It spans 15 years

**Table 1.** Description of Parameters (to Estimate), Variables, Coefficients, and Indices Used in “crr0” Model Structure

Symbol [Units]	Description	Min	Max
<i>Parameters</i>			
$I_e$ [mm]	Maximum interception	0	10
$S_{u,max}$ [mm]	Top layer/unsaturated zone moisture parameter	0	1000
$Q_{p,max}$ [mm/d]	Maximum percolation rate	0	100
$\alpha_E$ [-]	Curvature parameter for evaporation	0	100
$\alpha_F$ [-]	Curvature parameter for overland flow	-100	0
$\alpha_s$ [-]	Curvature parameter for percolation	-10	10
$K_s$ [day]	Base flow time constant	1	150
$K_r$ [day]	Routing time constant	1	10
<i>Variables</i>			
$S_u(t)$ [mm]	Upper layer/unsaturated zone soil water storage		
$S_s(t)$ [mm]	Lower layer/saturated zone soil water storage		
$E(t)$ [mm/d]	Evaporation		
$R(t)$ [mm/month]	Overland flow		
$Q_p(t)$ [mm/d]	Percolation		
$Q_s(t)$ [mm/d]	Base flow		
$P_e(t)$ [mm/d]	Effective precipitation (after interception)		
<i>Others</i>			
$P(t)$ [mm/d]	Precipitation		
$T$	Day index, {1,..,T}		

from 1961 to 1975. Ten overlapping 6 year periods (such as 1961–1966; 1962–1967 . . . ,1970–1975) are considered for calibration. The quantile models are estimated for each such period at nine quantiles values,  $\tau = 0.1, 0.2 \dots, 0.9$ . The first 2 years of each period are used for burnin and the performance on the remaining 4 years is used in model calibration. Thus,  $10 \times 9 = 90$  models are identified spanning 10 calibration periods at nine different quantile values. The performances of these models are then analyzed on 15 years of validation data set that spans from 1976 to 1990.

[19] Figure 2 shows the calibrated range of selected model parameters over different periods at different quantile values. The ranges of the parameters are tight relative to its allowable range (the  $y$  axis is scaled to the range = [min, max]), except for  $Q_{p,max}$  and  $\alpha_E$ . This suggests that the performance of quantile model selection is robust at each quantile. High variance in the estimates of  $Q_{p,max}$  and  $\alpha_E$  may indicate weak discriminatory power of the model in identifying percolation and evaporation fluxes. The parameter estimates also suggest, as expected, that the models that predict high flows (i.e., high quantile models) have lower upper zone storage capacity, smaller flow time constants, and higher (and less thresholded) percolations rates.

[20] Figure 3 analyzes the distribution of parameter estimates over different quantiles for each calibration period. All parameters have “nearly” constant ranges over time, which suggests that the selected quantile models are “nearly” consistent over years. These ranges are similar to the ones reported by *Schoups and Vrugt* [2010], except for  $K_s$ ,  $\alpha_E$ , and  $Q_{p,max}$  for the French Broad river basin. These differences may be due to the absence of an interception component in the model structure used for this analysis.

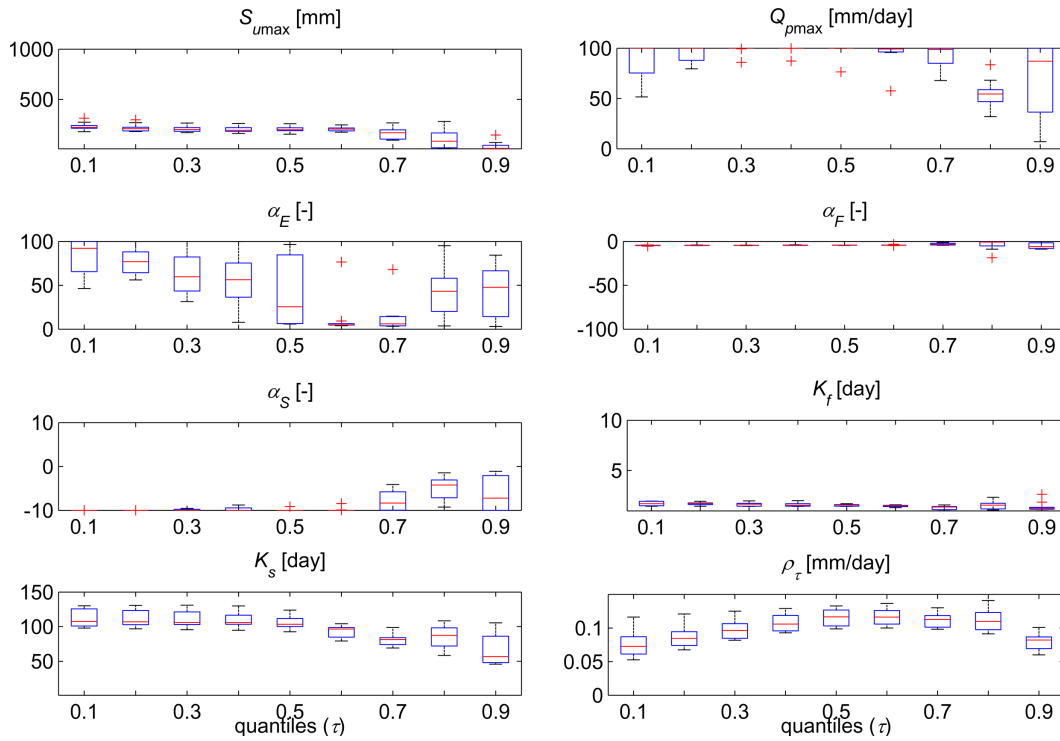
[21] Figure 4 compares the performance of quantile models predictions (resulting from quantile model estimation) with other estimators on a calibration period of 1970–1975. Only the last 75 days of the period are shown. Since the quantiles are varied between 0.1 and 0.9 with increments of 0.1 (lower and higher quantile values than 0.1 and 0.9, respectively, are not considered for robust estimation; see for example, the discussion on the robustness of trimmed estimators in *Vapnik* [2002]), the 10–90% inter quantile range is shown as the 80% quantile confidence interval. A model estimated by minimizing the mean square error (SCE-UA is used as the solver) is shown as the “MSE minimizer.” A median predictor is obtained by minimizing the mean absolute error. It is also a quantile model estimated at  $\tau = 0.5$ . Finally, Gastwirth and quant-mean model predictions are considered that are obtained as a certain weighted combination of quantile models. A “Gastwirth” type prediction  $\hat{y}_g$  is obtained as  $\hat{y}_g = 0.3 * \hat{y}_{\tau=0.3} + 0.4 * \hat{y}_{\tau=0.5} + 0.3 * \hat{y}_{\tau=0.7}$ , where  $\hat{y}_{\tau=0.3}$  is a quantile model prediction for quantile  $\tau = 0.3$  (quantiles  $\tau = 0.33, 0.5, 0.66$  were instead used in *Koenker and Basset* [1978]), while “quant-mean” prediction  $\hat{y}_{qm}$  is the mean of all quantile predictions, i.e.,  $\hat{y}_{qm} = \sum_{\tau=0.1:0.1:0.9} \hat{y}_{\tau}/9$ . These

are also called “inefficient” estimators [*Koenker and Basset*, 1978] because they are less efficient than the maximum likelihood estimators (MLE) with likelihood functions that match the underlying distributions. However, several studies such as *Koenker and Basset* [1978], and references within, have shown that such “inefficient” estimators are asymptotically more efficient across a wide variety of distributions (such as Gaussian, Gaussian mixture, Laplace, Logistic, and Cauchy) than MLE with a particular likelihood function and are almost as efficient as MLE for conventional parametric models. Since the predictions are a certain weighted mean of quantile model predictions, it is akin to predicting from a combination of models, where predictions from models that are estimated at different quantiles are combined.

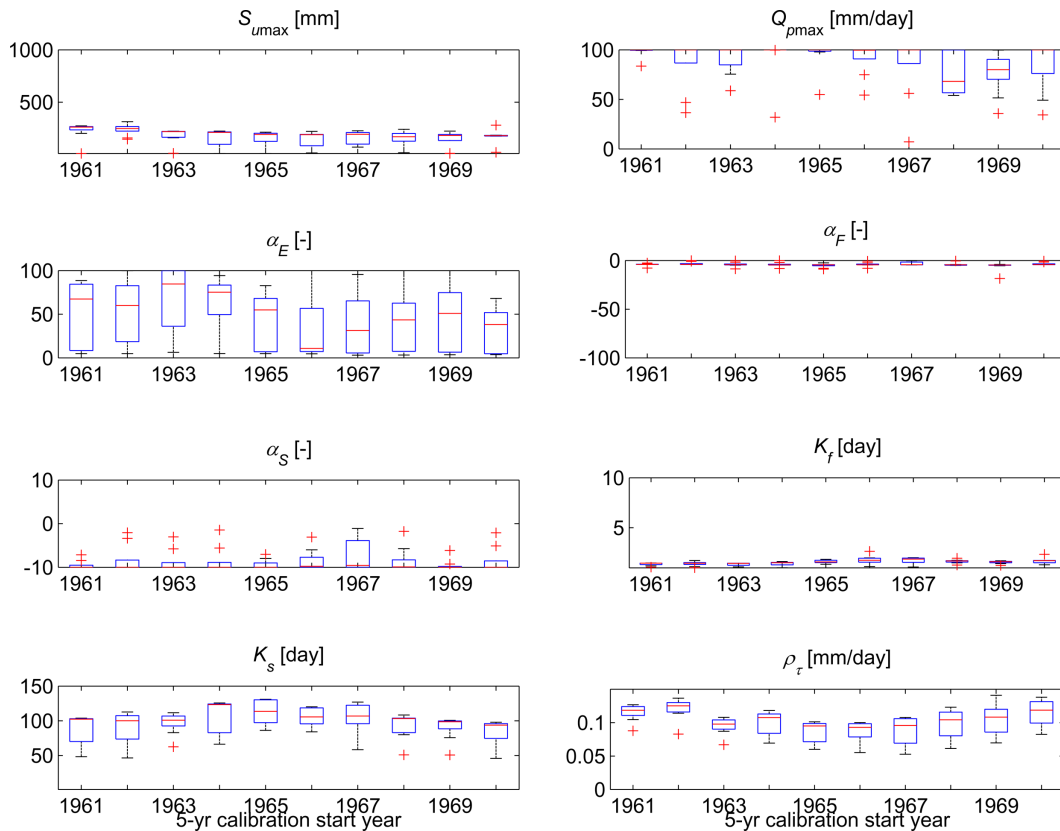
[22] The median model prediction tends to remain at the lower end of the 80% confidence interval. The Gastwirth and Quant-mean predictions lie in the lower to middle part of the confidence interval. This suggests that lower to median quantile predictions are closely spaced and that model structure “crr0” is rigid in predicting low flows when the performance metric is mean absolute deviation (note that quantile models are estimated by minimizing the asymmetric loss function which is a weighted mean of mean absolute deviations). All the predictions miss observed low flow at several locations (time indices between 1420 and 1450). The MSE minimization-based predictions appear to underpredict and overpredict observed streamflow at several locations, possibly as a result of the sensitivity of MSE to outliers. The 80% quantile confidence interval brackets most of the observations except the extremely low flows. Thus, in general it appears that low flow prediction is difficult for the given model structure.

[23] These observations are corroborated by the performance of these models on a separate validation data set. Its last 75 days are shown in Figure 5. The MSE minimizer-based prediction model is more variable, overpredicting

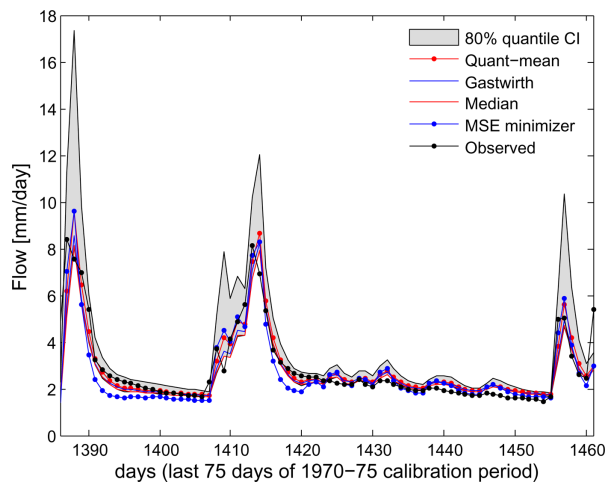
PANDE: QUANTILE MODEL SELECTION



**Figure 2.** Variation in parameters of the models selected over 10 overlapping 6 year periods (1961–1965, 1970–1975) at different quantile values. For parameter definitions, see Table 1. The y axes for the parameters are scaled to its range.  $\rho_\tau$  is the asymmetric loss function.



**Figure 3.** Variation in parameters of the models selected over 9 different quantiles ( $\tau = 0.1, \dots, 0.9$ ) in different 6 year calibration periods. For parameter definitions, see Table 1. The y axes for the parameters are scaled to its range.  $\rho_\tau$  is the asymmetric loss function.



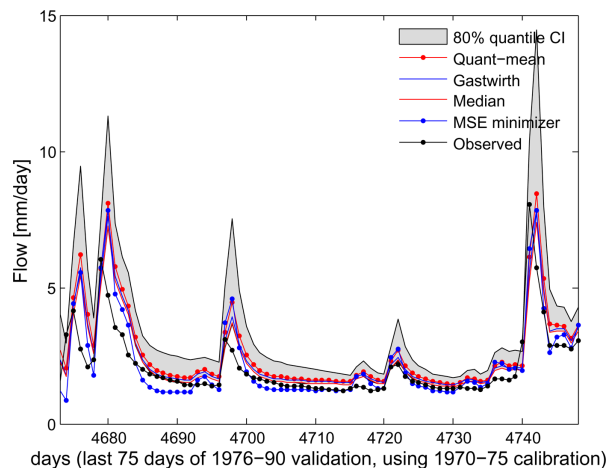
**Figure 4.** Comparison of quantile model prediction (resulting from quantile model estimation) with other related estimators on a calibration period of 1970–1975. The 80% quantile CI is the prediction range between  $\tau=0.1$  and 0.9. Quant-mean is defined as the mean of quantile model predictions made at quantiles  $\tau=0.1:0.1:0.9$ . Gastwirth is a weighted quantile prediction given by  $\hat{y}_g = 0.3 * \hat{y}_{\tau=0.3} + 0.4 * \hat{y}_{\tau=0.5} + 0.3 * \hat{y}_{\tau=0.7}$ , where  $\hat{y}_{\tau=0.3}$  is a quantile model prediction for quantile  $\tau=0.3$ . Median is a model obtained by minimizing mean absolute error while MSE minimizer is a model obtained by minimizing mean squared error.

higher flow and while underpredicting low flows. The median, Gastwirth, Quant-mean, and 80% confidence interval miss low flows to the extent that it gives an impression that quantile-based methods are less adept at handling low flow predictions than MSE minimization-based prediction. However, this is not the case. Table 2 shows the efficiency (Nash-Sutcliffe coefficient) and standard bias (mean of the difference of the observed from the predicted) for the same data set. The coverage frequencies (fraction of observations covered by a prediction interval) of 80% quantile confidence interval on the calibration and the validation data sets are also shown.

[24] Table 2 shows that median prediction is the worst performer on the calibration data set while the MSE minimizer-based model prediction is the worst performer on the validation data set. Quantile models combination-based predictions, i.e., Quant-mean and Gastwirth, are the best performers both on the calibration as well as the validation data, both in terms of efficiency as well as standard bias. Model structure deficiency is evident in the coverage frequency of 80% quantile confidence interval since only 52% of the observations on the calibration data set and 60% on the validation data set, instead of 80%, are covered.

### 3. Assessing Deficiency in a Parsimonious Model for Dry Land Areas

[25] Figure 6 outlines a parsimonious dry land model structure [Pande *et al.*, 2012a] for Western India (states: Gujarat and Rajasthan). The study area is subdivided into basins wherein each basin is further subdivided into a set of



**Figure 5.** Comparison of quantile model prediction (resulting from quantile model estimation) with other related estimators on a validation period of 1976–1990 using models estimated on calibration period of 1970–1975. The definitions of the legend entries are the same as in Figure 4.

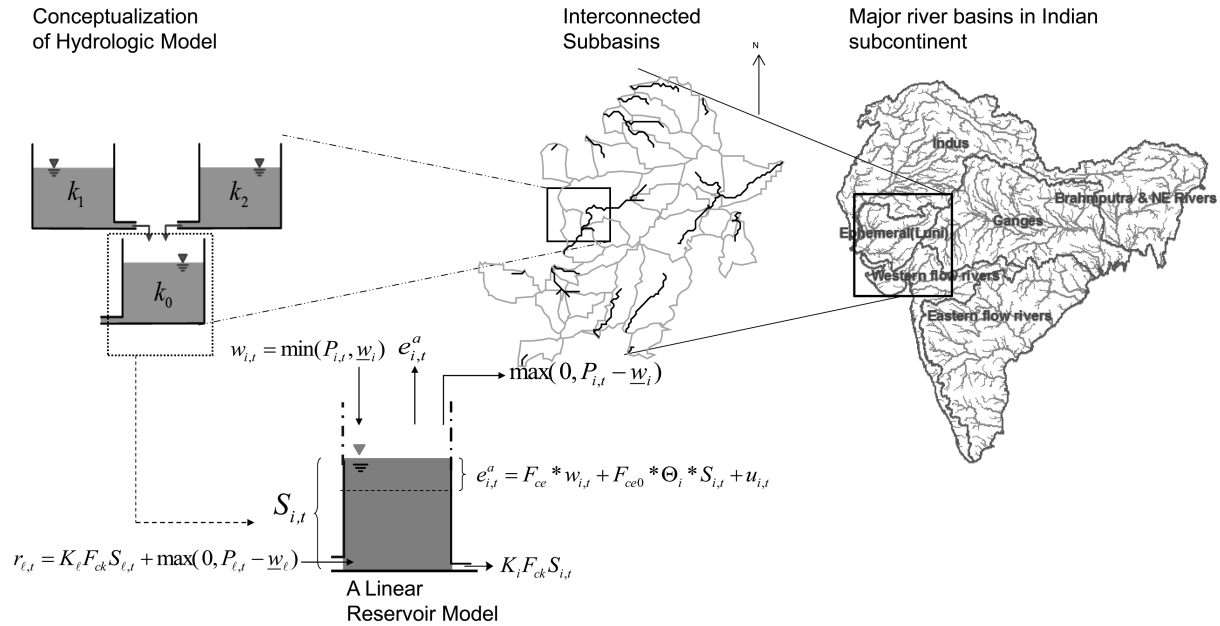
interconnected subbasins. Each such subbasin is conceptualized by a thresholded linear reservoir, with connectivity between the subbasins governed by relative elevation differences between the subbasins. Thus each such reservoir or a store conceptualizes water stored in the subsurface as well in streams within a subbasin. The model runs at monthly time steps and assumes steady state (cyclo-stationarity) at annual time scale (time steps  $t=1, \dots, 12$ ). This means that the model storages are constrained to return to month 1 storage after the 12th month.

[26] Figure 6 also shows the mass balance for  $i$ th reservoir in time  $t$ . The  $i$ th store receives rainfall  $P_{it}$  in time  $t$ , has effective hydraulic conductivity  $K_i$  and effective field capacity  $\Theta_i$ . The latter two subbasin characteristics  $K_i$  and  $\Theta_i$  are used to regionalize parameters that transform storage  $S_{it}$  into subsurface flow and vaporization from soil. The store conceptualizes three processes: overland flow ( $\max\{0, P_{it} - \underline{w}_i\}$ , where  $\underline{w}_i$  is the reservoir threshold), evaporation, and base flow ( $K_i F_{ck} S_{it}$ , where  $K_i F_{ck}$  transforms storage to base flow). By evaporation, we imply

**Table 2.** The Efficiency (Nash-Sutcliffe Coefficient) and Standard Bias (Mean of the Difference of the Observed From the Predicted) Performance Metrics of the Model Predictions in Figures 5 and 6 on a Calibration and Validation Data Set<sup>a</sup>

Prediction	Calibration: 1970–1975 Period		Validation: 1976–1990 Period	
	Efficiency (NS) (-)	Bias (mm/d)	Efficiency (NS) (-)	Bias (mm/d)
Median	0.78	-0.18	0.82	-0.15
Quant-mean	0.81	0.04	0.82	0.09
Gastwirth	0.80	-0.09	0.83	-0.06
MSE minimizer	0.81	-0.18	0.81	-0.20
80% quantile CI (Coverage frequency)	0.52		0.60	

<sup>a</sup>Also shown is the coverage frequency of the 80% quantile prediction range on calibration and validation data set.



**Figure 6.** Study area and model conceptualization [from *Pande et al.*, 2011]. (a) The study area is in the western semiarid to arid area of western India. It is delineated into a set of basins, each of which is described as a set of interconnected subbasins. Each subbasin is conceptualized as a linear reservoir with threshold and with parameter set  $k_i$ . Each such reservoir model conceptualizes base flow, overland flow, and vaporization either due to soil evaporation or land cover transpiration.

vaporization from effective rainfall  $F_{ce} w_{it}$  (where  $w_{it} = \min\{P_{it}, \underline{w}_i\}$ ), moisture in soil ( $F_{ce0} \Theta_i S_{it}$ , either due to evaporation through soil pores or due to land cover specific transpiration, where  $\Theta_i F_{ce0}$  transforms storage to vaporization), and irrigation water use ( $u_{it} = \eta E_{it}^0 a_i^{irri}$ , where  $E_{it}^0$  is subbasin specific evaporation demand and  $a_i^{irri}$  is the fraction of  $i$ th basin area under irrigation). Since the model is conceptualized at monthly time steps, upstream flows are assumed to comprise of two components: base flow  $K_i F_{ck} S_{it}$  generated by the upstream reservoir  $\ell$  as well as its overland flow,  $\max\{0, P_{it} - \underline{w}_i\}$ . The parameters of the model are  $\{F_{ce0}, F_{ce}, F_{ck}, \eta\}$  and  $\underline{w}_i, i = 1, \dots, I\}$  where  $I$  is the number of subbasins in the study area. For further discussion on model conceptualization, study area and data used and how the data are reconciled with the model, readers are referred to *Pande et al.* [2012a].

[27] The model concept is simple with  $S_{it}$  representing the sum of water stored in surface water bodies, saturated (as well as water in confined aquifers if present), and unsaturated zones. The vaporization from soil pores (in mostly unsaturated zone), which is partially due to land cover specific transpiration, is represented by  $F_{ce0} \Theta_i S_{it}$ . The parameter  $F_{ce0}$  is assumed to scale the role of field capacity in plant vaporization as well as to define the unsaturated zone as a fraction of total storage. Meanwhile, the irrigation water contribution to vaporization,  $u_{it}$ , is assumed to have been extracted from the ground water. No seasonality in irrigation water use has been assumed. Thus, the simplifications detailed above entail certain assumptions. Nonetheless we here note that such a locally linear model construct (at subbasin scale) is globally nonlinear at the corresponding basin scale [Pande et al., 2012b]. We explore potential

deficiencies in modeling total vaporization using quantile model selection as a result of the assumptions that have been implicitly made.

### 3.1. Quantile Model Selection on Evaporation Flux

[28] The equations governing the model structure are [Pande et al., 2012a]:

[29] (overland flow event)

$$w_{it} = \min\{P_{it}, \underline{w}_i\} \quad (3)$$

(actual evaporation for  $i$ th subbasin at time  $t$ )

$$e_{it}^a = F_{ce} w_{it} + F_{ce0} \Theta_i S_{it} + u_{it} \quad (4)$$

( $j$  land use specific evaporation demand at time  $t$ )

$$e_{jt}^d = k_{jt}^c E_{it}^0 \quad (5)$$

( $i$  specific irrigation applied at time  $t$ )

$$u_{it} = \eta E_{it}^0 a_i^{irri} \quad (6)$$

(water balance for  $i$ th subbasin at time  $t$ )

$$e_{it}^a - w_{it} = S_{it} - S_{it+1} + \sum_{\ell \in U(i)} r_{\ell t} - r_{it} \quad (7)$$

(constraint on actual evaporation)

**Table 3.** Description of Parameters (to Estimate), Variables, Coefficients, and Indices Used in the Parsimonious Model for a Dry Land Area [From Pande et al., 2012a]

Symbol [Units]	Symbol Description
<i>Description of Parameters</i>	
$F_{ce}$ [-]	Fraction of residual rainfall that evaporates
$F_{ce0}$ [1/month]	Multiplier on $\Theta_i$ , fraction of storage that evaporates
$F_{ck}$ [h/(in* month)]	Multiplier on $K_i$ , fraction of storage contributing to slow flow
$\eta$ [-]	Multiplier on $E_{i,t}^0 a_{ij}^{irri}$ , fraction of maximum irrigation demand
$w_i$ [mm/mo]	Hortonian overland flow threshold parameter, rainfall above this threshold is conceptualized as overland flow contribution
<i>Description of Coefficients</i>	
$K_i$ [in/h]	Effective hydraulic conductivity
$\Theta_i$ [-]	Effective porosity
$k_{i,j,t}^c$ [-]	Crop coefficients based on FAO guidelines for $j$ th land cover type in $i$ th subbasin and month $t$
$a_{ij}^{irri}$ [-]	Fraction of area irrigated in the $i$ th subbasin
$P_{i,t}$ [mm/month]	Monthly rainfall for the $i$ th subbasin in month $t$
$a_{ij}$ [-]	Fraction of $i$ th subbasin covered by $j$ th land cover type
<i>Description of Variables</i>	
$S_{i,t}$ [mm]	Store levels in subbasin $i$ and month $t$
$r_{i,t}$ [mm/month]	Slow flow out from $i$ th reservoir in month $t$
$E_{i,t}^0$ [mm/month]	Reference evaporation calculated using 1985 Hargreaves equation for $i$ th subbasin in month $t$
$u_{i,t}$ [mm/month]	Estimated irrigation demand for $i$ th subbasin in month $t$
$w_{i,t}$ [mm/month]	Residual rainfall in $i$ th subbasin and month $t$ after subtracting Hortonian overland flow
$e_{i,t}^d$ [mm/month]	Evaporation demand of $j$ th crop in month $t$
$e_{i,t}^a$ [mm/month]	Actual total evaporation from $i$ th subbasin in month $t$
<i>Description of Indices</i>	
$U(i)$	Set of subbasins upstream to the $i$ th subbasin
$I$	Subbasin index, $\{1, \dots, N\}$
$J$	Land cover index, $\{1, \dots, J\}$
$T$	Month index, $\{1, \dots, T\}$
$I$	Number of subbasins, 34 in this study
$J$	Number of land cover types, 18 in this study (15 croptypes, three other land cover types)
$T$	Number of months, 12 in this study

$$e_{it}^a \leq \sum_j a_{ij} e_{jt}^d \quad (8)$$

(subsurface flow equation from subbasin  $i$  to its downstream subbasins)

$$r_{it} = K_i F_{ck} S_{it} + \max \left\{ 0, P_{it} - w_i \right\} \quad (9)$$

(T-period steady state constraint)

$$S_{iT+1} = S_{i1} \quad (10)$$

[30] Table 3, reproduced from Pande et al. [2012a], provides a description of various symbols used in the above equations. A solution for the variables is obtained by solving the above equations (3–10) simultaneously for given values of parameters and coefficients.

[31] Consider further that observations for monthly changes in storage,  $\Delta \tilde{S}_{it}$ , and actual evaporation,  $\hat{e}_{it}^a$ , are available. Let the residuals from predicting storage change and actual evaporation be defined as,

$$\Delta \tilde{S}_{it} - (S_{it+1} - S_{it}) = \xi_{it} \quad (11)$$

$$\tau (\hat{e}_{it}^a - e_{it}^a) \leq \zeta_{it\tau}^+ \quad (12)$$

$$(1 - \tau) (-\hat{e}_{it}^a + e_{it}^a) \leq \zeta_{it\tau}^- \quad (13)$$

[32] The following minimization program (called QE3) then implements a  $\tau$ -quantile specific parameter estimation for the model structure given by equations (3–10) with evaporation flux as the variable of interest,

$$\min \left[ 0.5 \sum_{i=1}^I \sum_{t=1}^T |\xi_{it}| + \sum_{i=1}^I \sum_{t=1}^T (\zeta_{it\tau}^+ + \zeta_{it\tau}^-) \right] / NI$$

with respect to  $F_{ce}, F_{ce0}, F_{ck}, \eta, w_i, S_{it}, \zeta_{it\tau}^+, \zeta_{it\tau}^-, e_{it}^a, r_{it}, r_{it}, u_{it}$ ,

[33] subject to constraints (3)–(13) and

$$F_{ce} \geq 0, F_{ce0} \geq 0, F_{ck} \geq 0, \eta \geq 0, w_i \geq 0, S_{it} \geq 0, \zeta_{it\tau}^+ \geq 0 \text{ and } \zeta_{it\tau}^- \geq 0$$

for all  $i=1, \dots, I, \ell \in U(i)$  and  $t=1, \dots, T$ .

[34] We consider two linear reservoirs conceptualization for each subbasin as an improvement over the current model structure. These two linear reservoirs are vertically connected with the top reservoir conceptualizing an unsaturated zone while the second reservoir conceptualizes a saturated zone. Subsurface flow is only produced by the second (bottom) reservoir. Apart from the absence of base flow, the top reservoir is conceptualized in the same manner as a linear reservoir case with threshold-based overland flow and vaporization conceptualizations. The following equations along with equations (3–6 and 8) describe the conceptualization of each subbasin:

[35] (water balance for an  $i$ th subbasin at time  $t$ )

$$e_{it}^a - w_{it} = S_{it} - S_{it+1} + S_{it}^b - S_{it+1}^b + \sum_{\ell \in U(i)} r_{\ell t} - r_{it} \quad (14)$$

(subsurface flow equation from subbasin  $i$  to its downstream subbasins)

$$r_{it} = K_i F_{ck} S_{it}^b + \max \left\{ 0, P_{it} - w_i \right\} \quad (15)$$

[36] ( $T$ -period steady state constraint)

$$S_{iT+1} + S_{iT+1}^b = S_{i1} + S_{i1}^b \quad (16)$$

[37] Note that the equations (14–16) differ from (7, 9, 10) due to an additional storage variable  $S_{it}^b$  for the bottom store. The base flow is a fraction of the bottom storage level. The total storage at time  $t$  is now defined as  $S_{it} + S_{it}^b$ . The minimization program QE4 implements a  $\tau$ -quantile specific parameter estimation for the more complex model



structure given by equations (3–6, 8, 14–16) with evaporation as the variable of interest,

$$\min \left[ 0.5 \sum_{i=1}^I \sum_{t=1}^T |\tilde{\xi}_{it}| + \sum_{i=1}^I \sum_{t=1}^T (\tilde{\zeta}_{it\tau}^+ + \tilde{\zeta}_{it\tau}^-) \right] / NI$$

with respect to  $F_{ce}, F_{ce0}, F_{ck}, \eta, w_i, S_{it}, S_{it}^b, \tilde{\zeta}_{it\tau}^+, \tilde{\zeta}_{it\tau}^-, e_{it}^a, r_{it}, r_{lt}, u_{it}$ ,

[38] subject to constraints (3–6, 8, 14–16),

$$F_{ce} \geq 0, F_{ce0} \geq 0, F_{ck} \geq 0, \eta \geq 0, w_i \geq 0, S_{it}, S_{it}^b \geq 0, \tilde{\zeta}_{it\tau}^+ \geq 0, \tilde{\zeta}_{it\tau}^- \geq 0,$$

and with storage and vaporization residuals defined as:

$$\Delta \tilde{S}_{it} - (S_{it+1} - S_{it} + S_{it+1}^b - S_{it}^b) = \tilde{\xi}_{it} \quad (17)$$

$$\tau(\hat{e}_{it}^a - e_{it}^a) \leq \tilde{\zeta}_{it\tau}^+ \quad (18)$$

$$(1 - \tau)(-\hat{e}_{it}^a + e_{it}^a) \leq \tilde{\zeta}_{it\tau}^- \quad (19)$$

### 3.2. Quantile Model Selection Results

[39] Programs QE3 and QE4 are solved in General Algebraic Modeling System (GAMS) using MINOS5 Discrete NonLinear Programming (DNLP) solver [GAMS, 2008]. The threshold parameters  $w_i$  are set to calibrated values using an equally weighted mean absolute errors, i.e., using program QE3 with  $\tau = 0.5$  [see Pande *et al.*, 2012a] to direct the attention to the parameters that directly affect the estimation of vapor flux. Further, the storage levels are initialized (for the solver) at the levels estimated by the equally weighted mean absolute error in Pande *et al.* [2012a]. In order to address the sensitivity of the solver to parameter initialization by the user, a  $3 \times 3 \times 3 \times 3$  mesh-grid for remaining parameters  $\{F_{ce0}, F_{ce}, F_{ck}, \eta\}$  is created with lower bound at  $\{0,0,0,0\}$  and upper bound at  $\{0.1,0.1,0.1,0.1\}$  for QE3 and QE4. The upper bounds are set based on median parameter estimates of QE3, i.e., solution of QE3 with  $\tau = 0.5$  [Pande *et al.*, 2012a]. A solution that is a minimum of local minima (approximating a global minimum) is thus obtained. The results shown in the following for QE3 and QE4 are for the initialization that has the minimum objective function.

[40] Figures 7a and 7b show quantile parameter estimates relevant to vapor flux for programs QE3 and QE4, respectively. All parameters for the two reservoir model set up of QE4 are nearly constant for all the quantile values. Further, the parameter  $\eta$  corresponding to the evaporation due to irrigation is 0 for all quantile values except the 90th percentile. In the case of QE3, two of the three parameters,  $F_{ce}$ ,  $\eta$ , are also nearly constant for all quantiles except the 90th percentile. However,  $F_{ce0}$  gradually increases with quantiles. The improvement of the model structure in QE3 to that in QE4 reduces the need to explicitly model irrigation (as done in QE3) as the vertically connected model structure implicitly models the vertical flux between the two reservoirs.

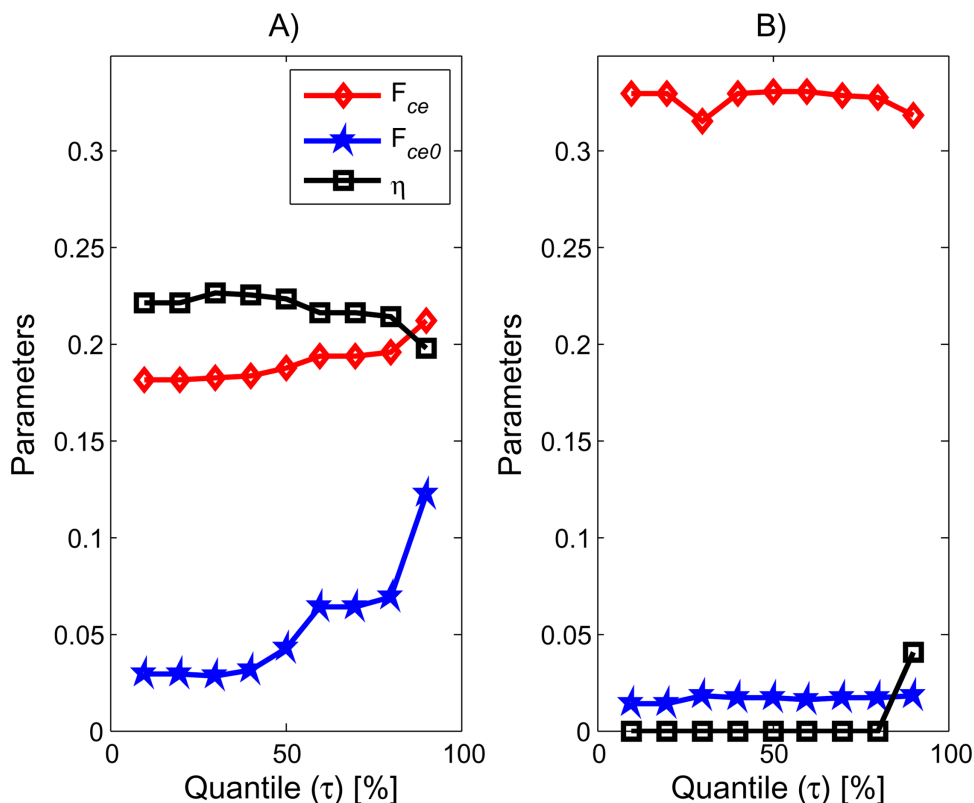
[41] Figure 8 further elaborates on the reduction of model structure deficiency when the model structure is improved. The loss curve for the thresholded single reser-

voir model structure (for each subbasin) in program QE3 increases with quantiles. The loss curve for the two reservoir model structure (for each subbasin) in QE4 first increases and then becomes constant with increasing quantiles. Further, the latter curve is quantilewise closer to 0 than the former. Both these observations reveal a reduction in model structure deficiency when storages in unsaturated and saturated zones are better distinguished via the two reservoir model structure. However, as Figure 8 suggests, neither of the two model structures remove the structural error entirely and yield poor prediction of higher vapor fluxes or there are errors in evaporation flux (reanalysis) data that cannot be accommodated in the model structures.

[42] Figure 9 shows the performance of two quantile models (10th and 90th percentile) with the reanalyzed data used for two subbasins within the study area and for the two model structures considered in programs QE3 and QE4. Note that quantile specific parameter values that are obtained from programs QE3 and QE4 are applicable to all the subbasins within the study area. These parameters are locally scaled by subbasin specific hydrologic properties (such as field capacity, hydraulic conductivity) or variables (such as rainfall) to yield subbasin specific vapor fluxes.

[43] The model structure deficiency in representing the underlying processes is evident in Figures 9a and 9b (though also in Figure 9c). The 80% confidence intervals of the model structure are not able to cover the vapor flux observations of the two subbasins. The model structure corresponds to program QE3. The inter quantile model performance for subbasin 11 is improved by considering the model structure in QE4 (Figure 9d) as it now brackets more observations. However, this appears not to be the case for subbasin 4. The improvement in the model structure as envisaged in QE4 is limited. This may be due to errors in the reanalyzed evaporation data, the 80% confidence interval or the limited model structure improvement in QE4. The interannual decline in the water table that the study area is currently witnessing [Tiwari *et al.*, 2009] has also been ignored. This can be important in explaining the part of observed reanalysis data that are not bracketed by the 80% quantile confidence interval. Note here that the model structures in QE3 and QE4 are both deficient such that the 10th and 90th percentile models do not cross, which supports the arguments presented in Pande *et al.* [2012a] that structural deficiency is a necessary but not a sufficient condition for quantile predictions to cross.

[44] Models based on traditional statistics such as mean absolute error or mean square errors can be obtained from quantile models for each of the structures. A median model is a quantile model for quantile value  $\tau = 0.5$ . Similarly a “mean” model that corresponds to a model obtained by minimizing mean square error statistics is a weighted sum of quantile models for quantiles around  $\tau = 0.5$ . These models are also equivalent to models obtained by maximizing a Laplace or a Gaussian likelihood function. These traditional statistics-based model selection would also suggest that the two reservoir model structure is an improvement over the single reservoir model structure. This is because the asymmetric loss function values for the former structure is quantile-wise closer to 0 than the latter and the statistics for a model selected based on traditional statistics can be

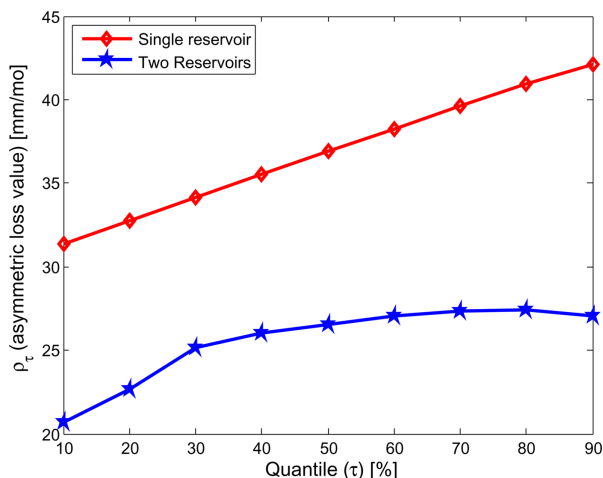


**Figure 7.** Vaporization relevant quantile parameter estimates  $F_{ce}$ ,  $F_{ce0}$ ,  $\eta$  from (a) program QE3 and (b) program QE4, equations (4) and (6).

represented as weighted sums of asymmetric loss function values at various quantiles.

#### 4. Deficiency Ordering of Model Structures for the Guadalupe River Basin

[45] The Guadalupe river basin data set is obtained from the Model Parameter Estimation Experiment (MOPEX) data set [Duan *et al.*, 2006]. The basin has a quick response



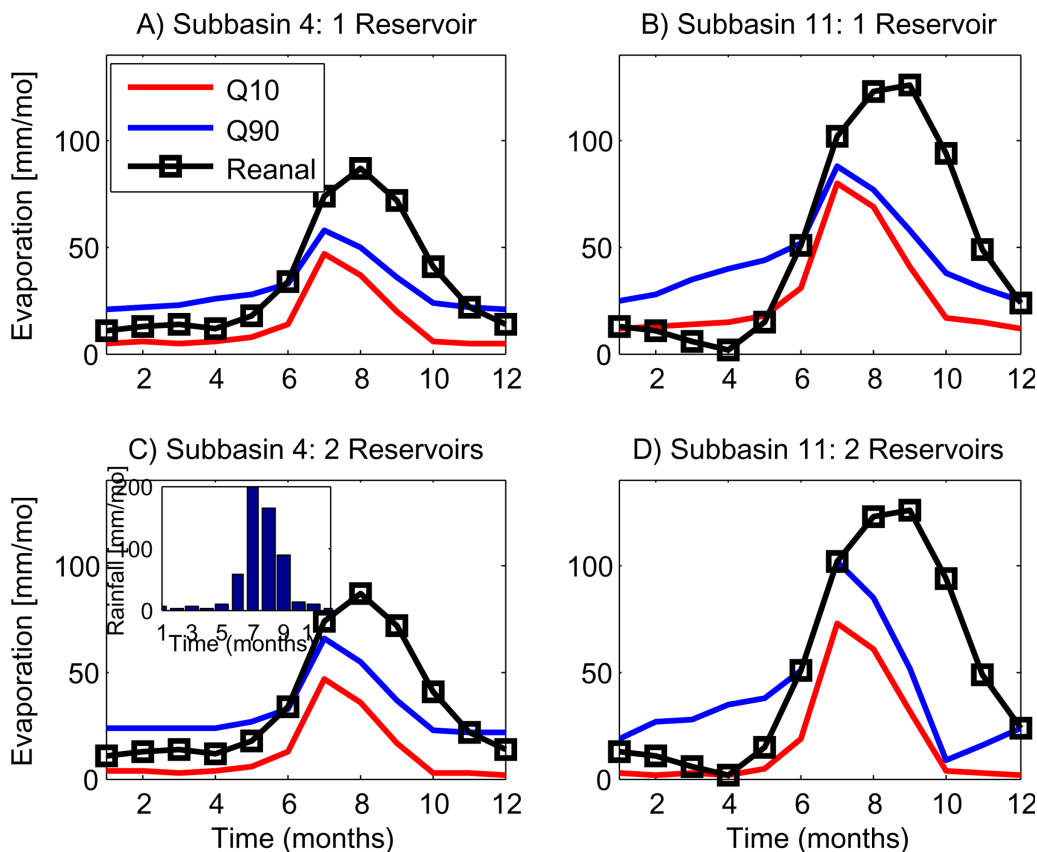
**Figure 8.** Comparison between the two asymmetrically weighted objective functions for model structure in QE3 (single reservoir) and QE4 (two reservoirs).

to precipitation with low flows for nonrainy days. This may indicate that the basin has low percolation rate, quick response of the basin to precipitation, and small residence time for overland flow. Further discussion on the river basin can be found in Clark *et al.* [2008].

[46] The modifications are made to model structure “crr0” (Figure 1) to generate nine different model structures. Three model structures are generated by altering the number of routing reservoirs from 3 to 1 (called crr1, crr2, and crr3). The next three model structures are created by removing the interception component from crr1, crr2, and crr3 yielding crr4, crr5, and crr6, respectively. The final three model structures are based on crr5 with alterations to the remaining components, i.e., the unsaturated and saturated zones. The first of these three structures, crr7, is such that parameter  $\alpha_F$  in equation (1) is an inverse function of storage  $S_s$  in the saturated zone. This mirrors a model structure conceptualization in Clark *et al.* [2008], wherein the saturated area is controlled by an inverse function of moisture in the lower layer. The overland flow equation then becomes:

$$R(t) = P_e(t) \frac{1 - \exp(-\{\alpha_F/S_s(t)\}S_u(t)/S_{u,max})}{1 - \exp(-\alpha_F/S_s(t))} \quad (20)$$

[47] The model structure crr8 is created by replacing the surface runoff generating mechanism in equation (1) with an infiltration excess mechanism based on Moore [1985]. It is assumed that the probability distribution function of the infiltration capacity,  $i$ , is a reflected power within the basin,



**Figure 9.** Comparison between reanalyzed evaporation data and 10th and 90th quantile models for two subbasins: (a and c) for subbasin 4 and (b and d) for subbasin 11. Figures 9a and 9b are for model structure within program QE3 and Figures 9b and 9d are for model structure within program QE4. Monthly rainfall values are shown in the inset.

where  $i_{max}$  is the maximum infiltration capacity present within the basin and  $\alpha_F$  now controls the curvature of this distribution,

$$F(i) = 1 - (1 - i/i_{max})^{\alpha_F} \quad 0 \leq i \leq i_{max} \quad (21)$$

[48] For a given amount of precipitation  $P(t)$  at time  $t$ , the infiltration excess flow is given by Moore [1985]:

$$R(t) = \begin{cases} \int_0^P F(i) di & P(t) \leq i_{max} \\ P(t) - i_{max} + \int_0^{i_{max}} F(i) di & P(t) > i_{max} \end{cases} \quad (22)$$

[49] The solution to equation (22) with the distribution specification of (21) yields:

$$R(t) = \begin{cases} \max \left\{ P(t) + \frac{i_{max}}{\alpha_F + 1} \left[ (1 - P(t)/i_{max})^{\alpha_F + 1} - 1 \right], 0 \right\} & P(t) \leq i_{max} \\ \max \left\{ P(t) - i_{max} + P(t) - \frac{i_{max}}{\alpha_F + 1}, 0 \right\} & P(t) > i_{max} \end{cases} \quad (23)$$

[50] Finally, the model structure crr9 is created from crr5 by introducing an upper bound,  $S_{smax}$  on the second layer

moisture  $S_s$  and the base flow is conceptualized as a nonlinear (power) function of the lower layer soil moisture (a power function with additional parameter  $b_F$  controlling the curvature). The second layer moisture that is in excess of its upper bound  $S_{smax}$  is transferred to the upper layer (similar in conceptualization to Sacramento Soil Moisture Accounting (SAC-SMA) model, Burnash [1995]). Thus,

$$Q_s(t) = (S_s(t)/S_{smax})^{b_F} / K_s \text{ for } S_s(t) \leq S_{smax} \quad (24)$$

[51] Table 4 summarizes all the above model structures and associated additional or redefined parameters.

#### 4.1. Quantile Model Selection on Streamflow

[52] Table 1 summarizes the basic model structure “crr0” which is the basis for model structures crr1 to crr9. The modeled flow,  $Q_{mod}(t)$ , is estimated for each of the model structures crr1 to crr9 and subtracted from observed flows  $Q_{obs}(t)$  to obtain two types of absolute residuals,  $\{\tilde{\zeta}_t^+, \tilde{\zeta}_t^-\}$ . It is then used to estimate a  $\tau$ -quantile specific model for each model structure by minimizing the asymmetric loss function in (2).

[53] The Shuffled Complex Evolution global optimization algorithm of University of Arizona (SCE-UA) is used to minimize the objective function in equation (2). The parameters of the algorithm are kept the same as in section 2.

**Table 4.** Summary of Model Structures crr1 to crr9 Used in This Study<sup>a</sup>

Model Structure	Modified Equation/Description	Additional/Adjusted Parameter (units)
Crr1	Three routing reservoirs	-
Crr2	Two routing reservoirs	-
Crr3	One routing reservoir	-
Crr4	Crr1 without interception	-
Crr5	Crr2 without interception	-
Crr6	Crr3 without interception	-
Crr7	Crr5 with overland runoff parameter an inverse function of lower layer water storage (see equation (20))	$\alpha_F$ (mm): proportionality parameter corresponding to $S_{s(t)}$
Crr8	Crr5 with overland flow generated by infiltration excess mechanism (see equation (23))	$i_{max}$ (mm/d): maximum infiltration capacity within the basin, $\alpha_F$ (-): curvature parameter of reflected power distribution function
Crr9	Nonlinear lower/saturated zone with upper bound on its storage (see equation (24))	$b_F$ (-): exponent, $S_{smax}$ (mm): upper bound on second layer storage

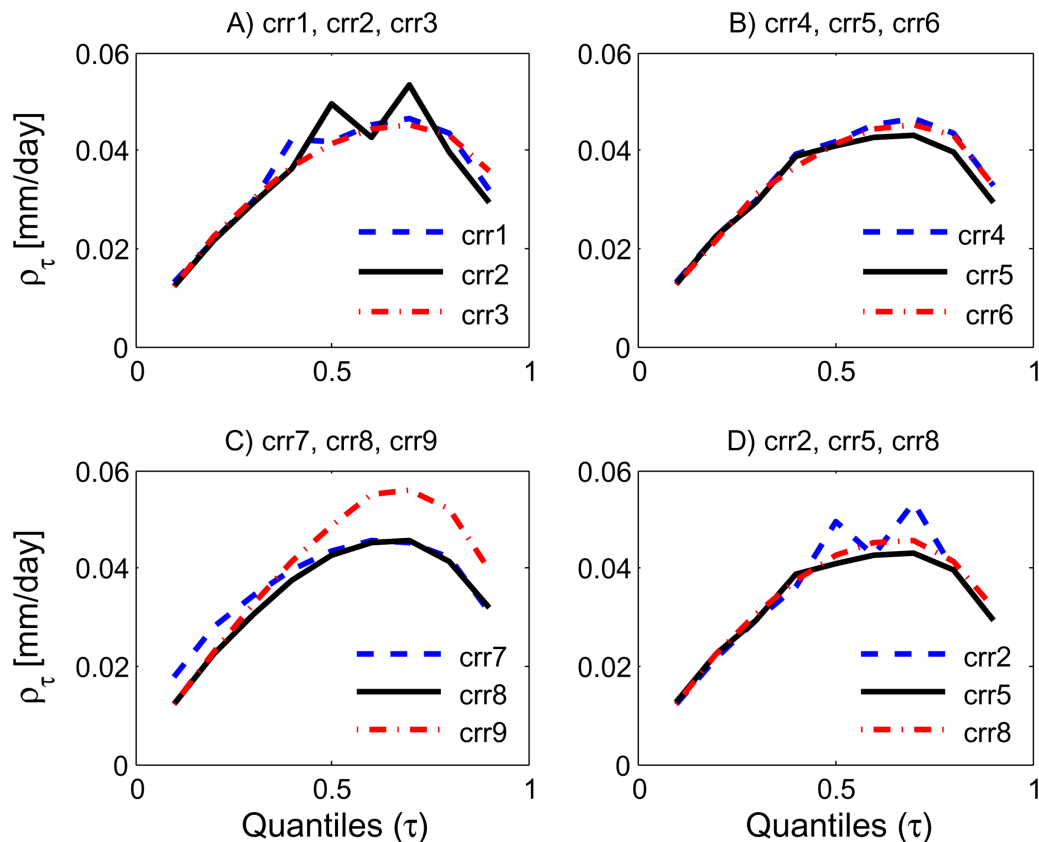
<sup>a</sup>All model structure changes are with reference to “crr0” unless mentioned otherwise.

[54] Daily streamflow, precipitation, and potential evapotranspiration data that are used for assessing the nine model structures span 6 years (1959–1964). It is assumed that the data size is large enough for the estimation of any model from the nine model structures to converge (to an infinite sample estimation), thereby removing the need to test the performance of a selected model over unseen data [Pande et al., 2009, 2012b]. Runga Kutta integrator is used to solve the differential equations involved in the model structures [Schoups and Vrugt, 2010]. SCE-UA is run five times with different initialization for each quantile and model structure. The reported results are quantile-wise best

performing models within each model structure amongst five different SCE-UA initializations.

**4.2. Quantile Model Selection Results for the Guadalupe River Basin**

[55] Figure 10 shows that quantile-wise best performing model (in terms of its asymmetric loss function) can be selected from each of the nine model structures. The model structures crr1–crr3 model low flows (represented by lower quantiles) in the same manner. The differences between the model structures appear in modeling larger quantiles with the model structure crr2 with two routing reservoirs



**Figure 10.** Quantile-wise comparison of asymmetric loss functions for nine model structures summarized in Table 4. Figure 12d shows best performing model structures crr2, crr5, and crr8 from Figures 12a–12c, respectively.

capable of performing the best. The spikes in the asymmetric loss functions of crr2 and crr3 are potentially due to the threshold (interception) estimation. Once the interception component from the model structure is removed, the asymmetric loss functions for the three model structures are more stable as shown in Figure 10b (structures crr4–crr6). All the three model structures, which are dissimilar only in the number of routing reservoirs, are capable of modeling the low flows in the same manner but differ in the way they can model the high flows. The model structure crr5 with two routing reservoirs is capable of providing the best quantile-wise performance amongst the three structures. Figure 10c shows that the model structure crr8 with an infiltration excess runoff generating mechanism yields the best quantile-wise performing model amongst the model structures crr7–crr9. The structure crr9 with a nonlinear and thresholded saturated zone conceptualization is capable of modeling low flows better but it is incapable of modeling high flows. It is the opposite for the model structure crr7 that conceptualizes runoff generation as dependent on the soil water storage in the lower/saturated zone reservoir. The model structure crr7 is not capable of mimicking low flows as good as the other two model structures, though it models high flows nearly as well as the model structure with an infiltration excess runoff generation mechanism.

[56] Figure 10d finally compares the best performing model structures from groups {crr1, crr2, crr3}, {crr4, crr5, crr6}, and {crr7, crr8, crr9}. These three model structures, crr2, crr5, and crr8, are also the best three model structures from the set of all nine model structures. These three model structures are similar in performance for low flows. Surprisingly, the model structure crr5 with two routing reservoirs and no interception (and the remaining structure the same as “crr0”) is better suited to model high flows than the model structure (crr8) with similar structure as crr5 except with infiltration excess as the runoff generating mechanism. Probably, a model structure with infiltration excess mechanism but with different number of routing reservoir may yield a better performing model structure.

[57] Three Bayesian model selection (detailed in Appendix B) criteria are also used to assess the nine model structures. The three criteria find crr1, crr2, crr5, and crr8 as the best of the nine model structures. It finds crr7 (BIC = 5.9742e + 004, HM1 = 5.9761e + 004, HM2 = 5.9661e + 004) and crr9 (BIC = 5.8411e + 004, HM1 = 5.8422e + 004, HM2 = 5.8347e + 004) as the worst structures. However, amongst these four model structures, the criteria find crr8 as the best (BIC = 6.4227e + 004, HM1 = 6.4248e + 004, HM2 = 6.4182e + 004), followed by crr2 (BIC = 6.2744e + 004, HM1 = 6.2760e + 004, HM2 = 6.2678e + 004), crr1 (BIC = 6.2594e + 004, HM1 = 6.2613e + 004, HM2 = 6.2520e + 004), and then crr5 (BIC = 6.2502e + 004, HM1 = 6.2522e + 004, HM2 = 6.2434e + 004). Thus, Bayesian criteria yield a more intuitive ordering of model structures than quantile model selection. It finds infiltration excess as an important process in Guadalupe river basin, followed by interception.

[58] Figure 11 shows the parameter values common to the model structures crr2, crr5, and crr9 as shown in Figure 10d. Even though crr2 and crr5 are nearly similar in performance, the kinks on asymmetric error of crr2 can be seen due to similar kinks in its parameter estimates. This is

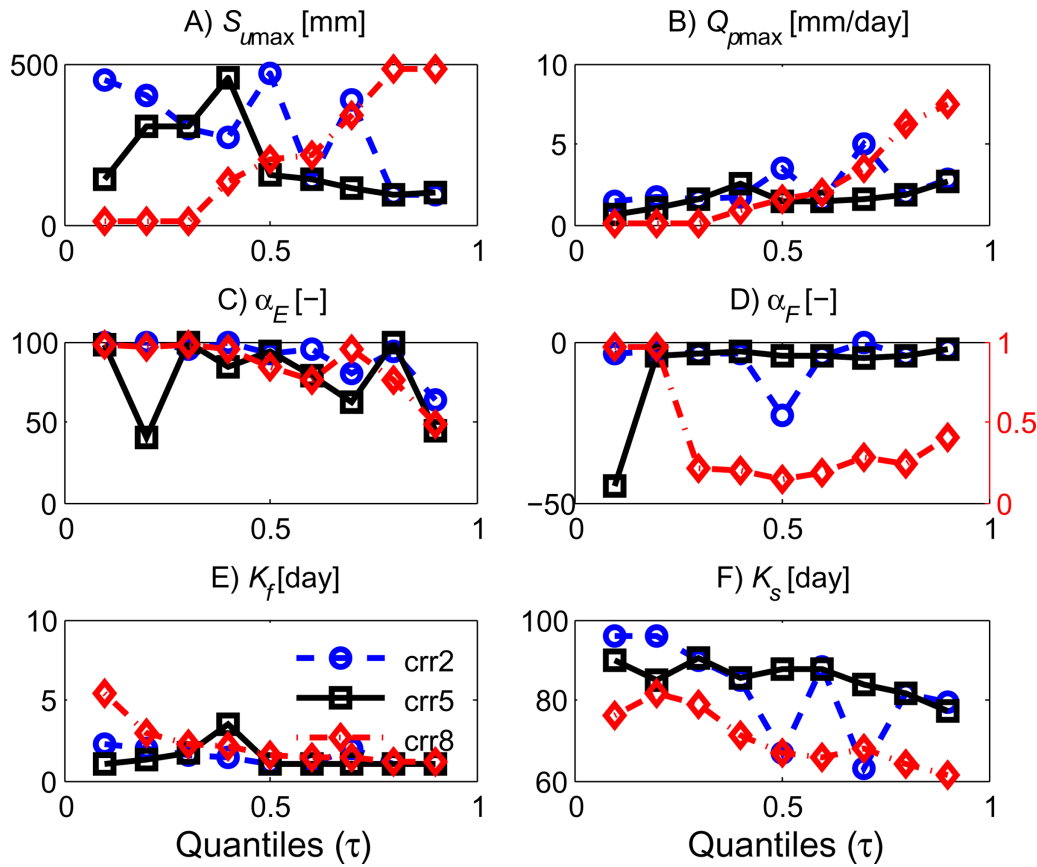
credited to searching for a threshold when it is not smoothed in the model equations [Clark *et al.*, 2008]. The presence of interception allows higher upper layer soil water storage capacity. Similarly, the estimates of maximum percolation rate  $Q_{pmax}$ , base flow and routing time constants,  $K_s$  and  $K_f$ , are lower for crr5 than crr2 for low quantiles values, indicating a compensation for the absence of interception zone in the case of the former. The evaporation relevant parameter estimates for crr5 are more erratic with quantiles compared to the other two model structures.

[59] The parameter estimates of all the model structures except for  $S_{umax}$  and  $\alpha_F$  show similar trend with quantiles (either increasing or decreasing) in Figure 11. The runoff related parameter  $\alpha_F$  has a different interpretation for crr8 ( $\alpha_F$  is the curvature of reflected power distribution function unlike for crr2 or crr5), thus a different  $y$  axis. Further note (from Figures 11a, 11b, and 11e) that the quantile specific parameter estimates for the structure crr8 are much more variable across quantiles. If the true model (of nature) is embedded within a model structure being considered, there would exist an optimal model that (i) has zero asymmetric loss function value and (ii) has same parameter values at different quantiles. These two conditions are akin to the conditions that are necessary and sufficient for a model structure “not” to be deficient (see also its discussion in Pande [2013], section 4.1). Thus, if the true model is not a member of the model structure (i.e., the model structure is deficient as would be the case almost surely), neither of the two consequences may hold. Hence, one may conclude that variability in parameter estimates across quantiles is a consequence of structure deficiency. However, the variability in parameter estimates across quantiles may also be due to different sensitivity of parameters at different quantiles in addition to model structure deficiency and may hide the deficiency effect. This along with the evidence from Figure 10d (worst asymmetric error amongst the three model structures) indicates that structure deficiency of crr8 is possibly higher than crr2 or crr5.

[60] In general, all three model structures show high positive values of  $\alpha_E$  and negative values for  $\alpha_F$  consistently across all the quantiles. Thus, all three model structures suggest that evaporation tends to occur at the potential rate and that the overland flow is initiated after the average basin moisture condition relative to the maximum crosses a certain threshold (for crr2 and crr5). Since the overland flow in the structure crr8 does not depend on the unsaturated zone soil moisture, the estimated  $S_{umax}$  increases with quantiles unlike the case of crr2 or crr5 wherein the unsaturated zone is parameterized shallower for higher quantiles in order to accommodate streamflow peaks. The routing time constant  $K_f$  is close to 1 day, the base flow time constant  $K_s$  decreases with higher quantiles for nearly all quantiles for all the model structures, suggestive of the quick response nature of the basin.

## 5. Conclusions

[61] Applications of quantile model selection and model structure deficiency assessment were presented in this paper. An exhaustive study using a flexible model structure and the French Broad river basin data set demonstrated that quantile model selection and prediction accommodates



**Figure 11.** Quantile specific parameter estimates (that are common across model structures) for three best performing model structures crr2, crr5, and crr8.  $\alpha_F$  values for structure crr8 are plotted on the right-hand side axis. The legend for the model structures is provided in (e).

other types of predictors such as those based on minimization of summary statistics as well and showed that that a weighted sum of quantile predictions is a robust predictor.

[62] Two additional case studies were then studied to demonstrate that the asymmetric loss function embeds model structural deficiency as proposed by Pande [2013]. In the parsimonious dry land modeling study, an expected improvement due to an addition of a lower reservoir in the model structure was reflected in lower asymmetric loss function values for all considered quantiles. In the Guadalupe river basin case study, the asymmetric loss function was employed to infer least deficient model structure out of nine model structures spanning different interception, routing, overland flow, and base flow conceptualizations. It also demonstrated that the asymmetric loss function can be used to order various model structures in terms of its deficiency. The Bayesian criteria revealed the same best four and the worst two model structures. However, it differed from quantile structural assessment in the identification of the least deficient model structure. The Bayesian criteria found infiltration excess and interception as important processes that reduce model structure deficiency in the Guadalupe river basin, unlike quantile model selection.

[63] No assumptions on error structure were made by quantile model selection. The quantile estimates of parameters thus might have absorbed some of the effects of measurement errors in addition to that of structural deficiency.

This however does not limit the conclusion that an improvement in a model structure leads to a reduction in bias in predicting a quantile. Naturally, the bias would not vanish if a structure for measurement errors is not incorporated in addition to the structure of predictive equations. But given that the absence of a structure for measurement errors equally holds for any improvement in a model structure, the bias in quantile prediction would reduce with any reduction in model structure deficiency. This was also shown to hold in the synthetic case study of Pande [2013] that studied the effect heteroscedastic additive noise on quantile model selection.

[64] Quantile model selection assesses structural deficiency across quantiles without making any strong a priori assumption about the “true” model structure. It is an approach that informs a modeler on where and to what degree her model structure is deficient in not being able to model the underlying processes. It is thus a mechanism to approach the “truth” in an efficient manner while acknowledging that the “truth” may never exactly be modeled.

### Appendix A: QE2 Problem Specification and the Assumptions

[65] Let  $S_t$  denotes the storage of a reservoir and let its outflux be a function of the storage denoted by  $f(S_t, k)$ . Here,  $k \in K$  represents a set of parameters (for example

slow and fast runoff coefficients),  $\mathbf{K}$  represents the range of parameters and corresponds to a particular model structure. Let  $\{y_t, x_t\}_{t=1, \dots, N}$  represent observed data set where  $y_t \in \mathbb{R}_+$  and  $x_t \in \mathbb{R}_+$  represent observed outflux and input forcing, respectively, at time  $t$ . Let  $\mathbf{x} = \{x_t\}_{t=1, \dots, N}$  represent the input forcing vector and  $S_o$  is the initial soil moisture condition. A  $\tau$ -quantile specific function and the corresponding parameters based on outflux observations can be estimated by the program (QE2):

$$W(\tau; y, \mathbf{x}) = \min_{\mathbf{k} \in \mathbb{K}, \varepsilon_t^-, \varepsilon_t^+, S_t} \sum_{t=1}^N [(1 - \tau)\varepsilon_t^- + \tau\varepsilon_t^+]$$

$$st$$

$$y_t - f(S_t, \mathbf{k}) - \varepsilon_t^+ \leq 0$$

$$-y_t + f(S_t, \mathbf{k}) - \varepsilon_t^- \leq 0$$

$$S_{t+1} - S_t = x_t - f(S_t, \mathbf{k})$$

$$\varepsilon_t^+, \varepsilon_t^- \geq 0$$

[66] The following assumptions should hold for the propositions provided in Pande [2013] to be valid for a given hydrological model selection problem.

[67] **Assumption 1:** The parameter set  $\mathbf{K}$  that defines the model structure for a given model equation is compact.

[68] **Assumption 2:** The model equation  $f(S_t; \mathbf{k})$  is differentiable, is monotonic in at least one element of  $\mathbf{k}$ , and increasing in  $S_t$ . Further,  $\frac{\partial f}{\partial k_j}$  is independent of  $\frac{\partial f}{\partial k_j}$  where  $k_j$  and  $k_{j'}$  are two distinct elements of  $\mathbf{k}$ .

[69] **Assumption 3:** Input forcing vector is nonzero, i.e.,  $\mathbf{x} = \{x_1, \dots, x_t, \dots, x_N\} \neq 0$ .

[70] **Assumption 4:** Initial model storage is sufficiently greater than 0, i.e.,  $S_o \gg 0$ .

[71] **Assumption 5:** The observed variable of interest is bounded, i.e.,  $\underline{y} \leq y \leq \bar{y}$ .

[72] **Assumption 6:** The cumulative probability density  $F(\mathbf{y}|\mathbf{x})$  is differentiable and

[73]  $F(\underline{y}|\underline{y}^{-t}, \mathbf{x}) = 0, \forall \underline{y}^{-t} \leq \underline{y}^{-t} \leq \bar{y}^{-t}$  where  $\underline{y}^{-t} = \{y_1, \dots, y_{t-1}, y_{t+1}, \dots, y_N\}$ .

[74] **Assumption 7:** We avail of a global optimizer that can identify a minimum of a quantile model selection problem.

## Appendix B: Bayesian Criteria

[75] Three Bayesian criteria are used to approximate the marginal log likelihood of a model structure [Pande, 2013].

[76] 1. Bayesian Information Criteria (BIC) [Kass and Raftery, 1995]:

$$\log m(y|M) \approx \log f(y|M, \theta^*) - 0.5 \log(N)|\theta|$$

[77] 2. Harmonic mean of the log likelihood values of the posterior distribution (HM1) [Kass and Raftery, 1995]:

$$\log m(y|M) \approx \frac{1}{m} \left( \sum_{i=1}^m (\log f(y|M, \theta_i^*))^{-1} \right)^{-1}$$

[78] 3. A variant of Chib and Jeliazkov [2001] (HM2):

$$\log m(y|M) \approx \log f(y|M, \theta^*) + \log \pi(\theta^* | M) - \log \pi(\theta^* | y, M)$$

[79] Here,  $m(y|M)$  is the marginal likelihood that data  $y$  are from a model structure  $M$ ,  $f(y|M, \theta^*)$  is the likelihood

that the data  $y$  are from a model that is from a structure  $M$  and parameterized by  $\theta^*$ ,  $\theta^*$  represents the maximum likelihood parameter estimate (MLE) for a given model structure  $M$ ,  $|\theta^*|$  is the dimensionality of the parameter set,  $\pi(\theta^* | M)$  is the prior probability of the MLE  $\theta^*$ ,  $\pi(\theta^* | y, M)$  is the posterior probability of  $\theta^*$ ,  $N$  is the sample size, and  $m$  is the size of parameter sets sampled from the posterior distribution  $\pi(\theta^* | y, M)$ . The General Likelihood function is used (see section 2.3.1) for  $f(y|M, \theta)$ .

[80] For the Bayesian criteria HM2,  $\pi(\theta|y, M)$  is nonparametrically estimated using multivariate kernel density estimation. For the case studies,  $N = 2192$  days and  $m = 600$ .

[81] **Acknowledgments.** The author is grateful to Michiel A. Keyzer for his critical review and suggestions, to Gerrit Schoups for providing C code for model structure used for Guadalupe River Basin and many discussions on the topic, and to Mojtaba Shafiei, Huub Savenije, and Ashvani K. Gosain for discussions on a previous version of the manuscript. Thanks are due to several referees including Nataliya Bulygina and Jasper A. Vrugt for their critical review of the manuscript. The author also thanks the AE and the Editor for their patience with previous versions of the manuscript.

## References

- Blazkova, S., and K. Beven (2009), A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic, *Water Resour. Res.*, 45, W00B16, doi:10.1029/2007WR006726.
- Burnash, R. J. C. (1995), The NWS river forecast system-catchment modeling, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, pp. 311–366, Water Resour. Publ., Highlands Ranch, Colo.
- Clark, M. P., A. G. Slater, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gupta, T. Wagener, and L. E. Hay (2008), Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, *Water Resour. Res.*, 44, W00B02, doi:10.1029/2007WR006735.
- Duan, Q., S. Sorooshian, and V. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, 28(4), 1015–1031, doi:10.1029/91WR02985.
- Duan, Q., et al. (2006), The Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, *J. Hydrol.*, 320, 3–17, doi:10.1016/j.jhydrol.2005.07.031.
- GAMS (2008), *GAMS-A user's guide: Tutorial by Richard E. Rosenthal*, p. 273, GAMS Dev. Corp., Washington, D. C.
- Keyzer, M. A., and S. Pande (2009), Instrumentalization using quantiles in semiparametric support vector regression, Cent. World Food Stud. Working Pap. WP 09-04, 61 pp., Stichting Onderzoek Wereldvoedselvoorziening van de Vrije Universiteit (SOW-VU), Amsterdam.
- Koenker, R. (2005), *Quantile Regression, Econ. Soc. Monogr. Ser.*, vol. 38, 366 pp., Cambridge Univ. Press, New York, NY.
- Koenker, R., and G. Basset (1978), Regression quantiles, *Econometrica*, 46(1), 33–50.
- Ma, L., and R. Koenker (2006), Quantile regression methods for recursive structural equation models, *J. Econ.*, 134, 471–506.
- Moore, R. J. (1985), Probability-distributed principle and runoff production at point and basin scales, *Hydrol. Sci. J.*, 30, 273–297.
- Pande, S. (2013), Quantile hydrologic model selection and model structure deficiency assessment: 1. Theory, *Water Resour. Res.*, 49, 5658–5673, doi:10.1002/wrcr.20411.
- Pande, S., M. McKee, and L. A. Bastidas (2009), Complexity-based robust hydrologic prediction, *Water Resour. Res.*, 45, W10406, doi:10.1029/2008WR007524.
- Pande, S., H. H. G. Savenije, L. A. Bastidas, and A. K. Gosain (2010), A parsimonious modeling approach to water management in dryland areas, in S. Khan, H. H. G. Savenije, S. Demuth and P. Hubert, eds., *Hydrocomplexity: New Tools for Solving Wicked Water Problems*, vol. 338, Int. Assoc. of Hydrol. Sci. Publ., IAHS Press, Wallingford, U. K.
- Pande, S., B. van den Boom, H. H. G. Savenije, and A. K. Gosain (2011), Water valuation at basin scale with application to western India, *Ecol. Econ.*, 70, 2416–2428, doi:10.1016/j.ecolecon.2011.07.025.

- Pande, S., H. H. G. Savenije, L. A. Bastidas, and A. K. Gosain (2012a), A parsimonious hydrological model for a data scarce dryland region, *Water Resour. Manage.*, 26, 909–926, doi:10.1007/s11269-011-9816-z.
- Pande, S., L. A. Bastidas, S. Bhulai, and M. McKee (2012b), Parameter dependent convergence bounds and complexity measure for a class of conceptual hydrological models, *J. Hydroinformatics* 14(2), 443–463.
- Schoups, G., and J. A. Vrugt (2010), A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors, *Water Resour. Res.*, 46, W10531, doi:10.1029/2009WR008933.
- Son, K., and M. Sivapalan (2007), Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data, *Water Resour. Res.*, 43, W01415, doi:10.1029/2006WR005032.
- Tiwari, V. M., J. Wahr, and S. Swenson (2009), Dwindling groundwater resources in northern India, from satellite gravity observations, *Geophys. Res. Lett.*, 36, L18401, doi:10.1029/2009GL039401.
- Vapnik, V. (2002), *The Nature of Statistical Learning Theory*, 2nd ed., Springer, New York.
- Weerts, A. H., H. C. Winsemius, and J. S. Verkade (2011), Estimation of predictive hydrological uncertainty using quantile regression: Examples from the National Flood Forecasting System (England and Wales), *Hydrol. Earth Syst. Sci.*, 15, 255–265.
- Westerberg, I. K., J.-L. Guerrero, P. M. Younger, K. J. Beven, J. Seibert, S. Halldin, J. E. Freer, and C.-Y. Xu (2011), Calibration of hydrological models using flow duration curves, *Hydrol. Earth Syst. Sci.*, 15, 2205–2227.
- Yu, P.-S., and T.-C. Yang (2000), Using synthetic flow duration curves for rainfall-runoff model calibration at ungauged sites, *Hydrol. Processes*, 14, 117–133.