Approximate Bayesian computation for forecasting in hydrological models

Jonathan Romero-Cuéllar

Joint work with Antonino Abbruzzo, Giada Adelfio and Félix Francés
Universitat Politècnica de València

SIS 2018: 49th Scientific Meeting of the Italian Statistical Society

jorocue1@doctor.upv.es

22/06/2018

Motivations and Aims

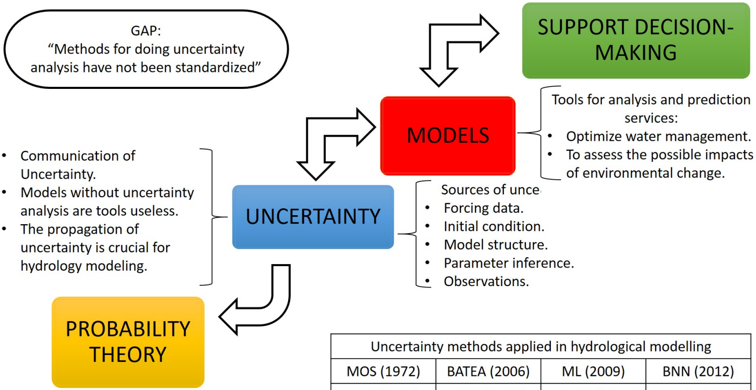
Motivations:

- Hydrological predictions are valuable for risks assessment, water resources management, and ecological issues [1].
- Quantifying the uncertainty of predictions are essential for decision-making [2].

Aims:

- To introduce a new hydrological post-processor based on summary statistics and free-likelihood function.
- To compare the performance of the new Approximate Bayesian Computation (ABC) post-processor with the MCMC post-processor.

Uncertainty in Environmental Models
Why should we be interested in uncertainty?



Uncertainty methods applied in hydrological modelling			
MOS (1972)	BATEA (2006)	ML (2009)	BNN (2012)
HUP (2000)	MCP (2008)	GLMPP (2011)	Copulas(2013)
BMA (2005)	BJP (2009)	QR (2011)	ABC (2013)

Hydrologic Post-processing

- Hydrologic post-processors are statistical models that relate observations with hydrological predictions [3].
- We select the linear model post-processor

$$y_t = \beta_0 + \beta_1 \hat{y}_t + \varepsilon_t, \tag{1}$$

- The ABC produces draws from an approximation of the posterior distribution of $\theta = (\beta_0, \beta_1, \sigma^2)$, i.e.

$$p(\theta|\mathbf{y}) \propto p(\mathbf{y}|\theta)p(\theta)$$

- We assume flat uniform priors for β_0, β_1, and σ^2 and $Y_t|\theta \sim N(\mu_t = \beta_0 + \beta_1 \hat{y}_t, \sigma^2)$ **(NQT)**
- The approximate predictive uncertainty formally defined as

$$g(y_{T+1}|\hat{y}) = \int_\Theta p(y_{T+1}|\theta, \hat{y}) p_\epsilon(\theta|\eta(\hat{y}))d\theta \tag{2}$$

Basic Approximate Bayesian Computation (ABC) algorithm

ABC is probably the most important likelihood-free methodology [4].

Algorithm 1 ABC accept/reject algorithm

1: θ^i, $i = 1, \ldots, N$ from $p(\theta)$
2: $\mathbf{z}^i = (z_1^i, z_2^i, \ldots, z_T^i)^\top$, $i = 1, \ldots, N$, from the likelihood, $p(\cdot|\theta^i)$
3: Select θ^i such that:

$$d\{\eta(\mathbf{y}), \eta(\mathbf{z}^i)\} \leqslant \epsilon$$

where $\eta(\cdot)$ is a vector statistic, $d\{\cdot\}$ is a distance criterion, and, given N, the tolerance level ϵ is chosen to be small.

What make valid predictions?

- **Reliable:** Predictions statistically consistent with observed data
- **Precise:** Small uncertainty in predictions
- **Unbiased:** Predictions not showing an unfair tendency

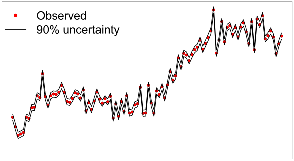
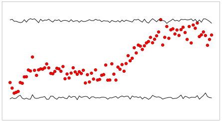
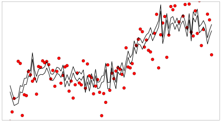


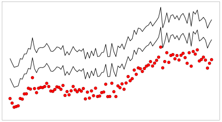
Figure: Reliable, precise, and unbiased



(a) Reliable but imprecise

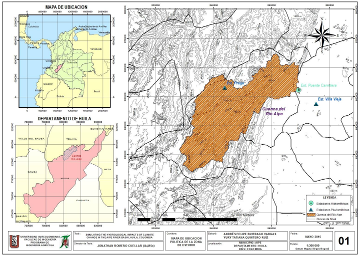


(b) Precise but unreliable

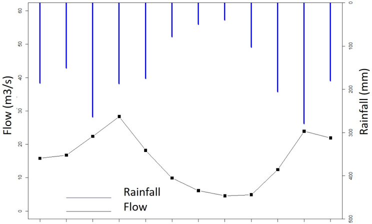


(c) Biased

Study case



(a) The location of the Aipe catchment, Colombia.[5]



(b) The Water balance of Aipe catchment.

Study case, Time series

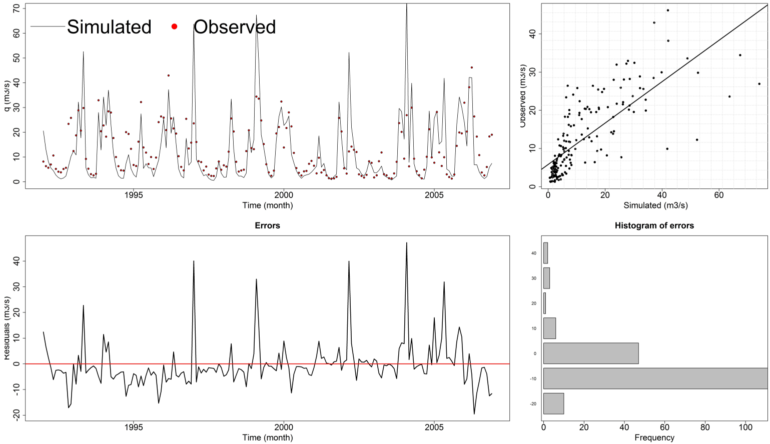


Figure: Time series from Aipe catchment

Performance metrics

Table: Deterministic and probabilistic performance metrics of the raw prediction, MCMC and ABC post-processor for the Aipe catchment.

Performance metric	Calibration			Validation		
	Raw prediction	Post-processing MCMC	Post-processing ABC	Raw prediction	Post-processing MCMC	Post-processing ABC
NSE	0.165	0.669	0.671	0.571	0.777	0.773
KGE	0.527	0.769	0.764	0.637	0.757	0.744
Reliability		0.996	0.996		0.993	0.993
Precision		2.403	2.306		2.581	2.500
K-S test (p-value)		0.465	0.750		0.132	0.223
95% exceed ratio (ER95)		88.33	88.89		94.44	94.44

Uncertainty Band MCMC post-processor

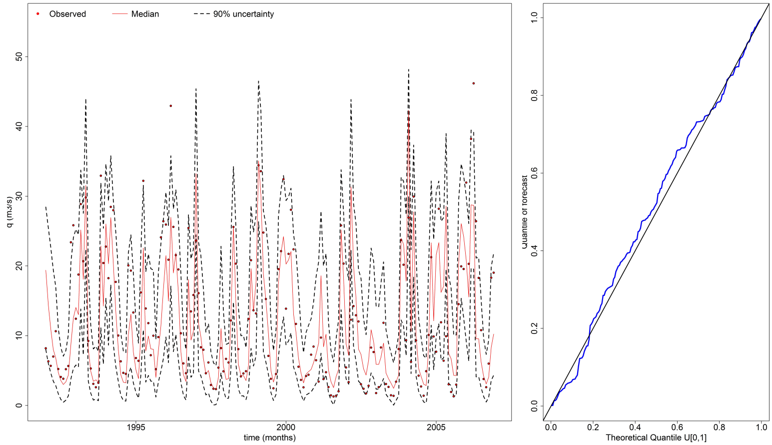


Figure: Conditional predictive uncertainty from MCMC post-processor on the Aipe catchment (left). PP-plot of the conditional predictive distribution (right).

Uncertainty Band ABC post-processor

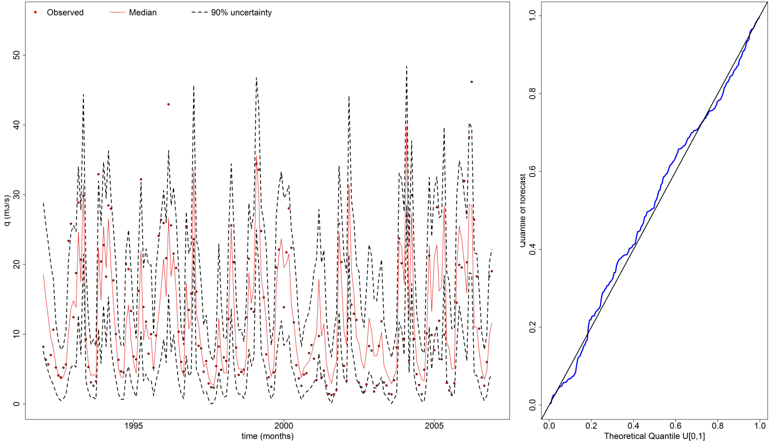


Figure: Conditional predictive uncertainty from ABC post-processor on the Aipe catchment (left). PP-plot of the conditional predictive distribution (right).

Conclusion

- The results show that ABC post-processor as similar performance as the MCMC algorithm regarding forecasting metrics. However, the ABC post-processor just used a summary statistics to quantify the conditional predictive uncertainty. Therefore, ABC post-processor has potential in situations where we do not have a hydrological time series. For example, ungauged catchments or climate change impact studies (work in progress).

References I

Michael B. Butts, Jeffrey T. Payne, Michael Kristensen, and Henrik Madsen.
An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation.
Journal of Hydrology, 298(1):242–266, 2004.

Alberto Montanari and Demetris Koutsoyiannis.
A blueprint for process-based modeling of uncertain hydrological systems.
Water Resources Research, 48(9):W09555, sep 2012.

References II

Introduction
Methodology
Results
Conclusion

Aizhong Ye, Qingyun Duan, Xing Yuan, Eric F. Wood, and John Schaake.
Hydrologic post-processing of MOPEX streamflow simulations.
Journal of Hydrology, 508:147–156, jan 2014.

Fenicia Fabrizio, Kavetski Dmitri, Reichert Peter, and Albert Carlo.
Signaturedomain calibration of hydrological models using approximate bayesian computation: Empirical analysis of fundamental properties.
Water Resources Research, 0(ja):Accepted Author Manuscript, 2018.

References III

Jonathan Romero-Cullar, Andres Buitrago-Vargas, Tatiana Quintero-Ruiz, and Flix Francs.
Simulacin hidrolgica de los impactos potenciales del cambio climtico en la cuenca hidrogrfica del ro aipe, en huila, colombia.
Ribagua, 5(1):63–78, 2018.