Supplementary materials: A Robust calibration/validation protocol of a hydrological model using hidden Markov states

Etienne Guilpart [1], Vahid Espanmanesh [1], Amaury Tilmant [1], and François Anctil [1]

[1]Département de génie civil et de génie des eaux, Université Laval, Québec, Canada

Correspondence: E. Guilpart (etienne.guilpart@gmail.com)

S1 Calibrations and validations on the wet historical episode $T^{1945-1971}$

Here, we apply our protocol on the period 1945-1971 ($T^{1945-1971}$), which can be consider as a wet historical episode of the SRB. The results are displayed in Figure S1, Table S1 and Table S2.

The wet historical episode $T^{1945-1971}$

Pettitt test	Basins	p value	Year break		
	Daka Saidou	0.694	-		
	Oualia	0.399	-		
	Bakel	0.646	-		
	Basins	μ	σ	δ	M
2-states-HMM	Daka Saidou	366 ; 273.3	25.6;43.6	1,0	$\begin{bmatrix} 0.864 & 0.136 \\ 0.413 & 0.587 \end{bmatrix}$
	Oualia	201.8;124	32.1; 15	0,1	$\begin{bmatrix} 0.344 & 0.656 \\ 0.287 & 0.713 \end{bmatrix}$
	Bakel	1073;713	65.4;125.4	0,1	$\begin{bmatrix} 0.684 & 0.316 \\ 0.535 & 0.465 \end{bmatrix}$
3-states-HMM	Daka Saidou	361.6;289.8;217	15.3;20.9;27.2	0,1,0	$\begin{bmatrix} 0.589 & 0.158 & 0.252 \\ 0.141 & 0.696 & 0.163 \\ 0.154 & 0.252 & 0.594 \end{bmatrix}$
	Oualia	127.7;192.9;243.9	17;19.5;6.5	0,1,1	$\begin{bmatrix} 0.397 & 0.281 & 0.322 \\ 0.482 & 0.435 & 0.083 \\ 0 & 0.835 & 0.165 \end{bmatrix}$
	Bakel	243.9;192.9;127.7	6.5;19.5;17	0,1,0	$\begin{bmatrix} 0.397 & 0.281 & 0.322 \\ 0.482 & 0.435 & 0.083 \\ 0 & 0.835 & 0.165 \end{bmatrix}$

Table S1. Pettitt test results and Hidden Markov Model parameters (N=2 and N=3) for Daka Saidou, Oualia, and Bakel sub-basins, on the wet sub-sequence $T^{1945-1971}$.

For the three sub-basins, Pettitt's tests are inconclusive, meaning that there is no trend in those sub-sequences. The transition probability matrices for the 2-states HMM and 3-states HMM are diverging from an identity matrix, indicating that the temporal persistence is less pronounced. However, in a such situation, HMM classification remains a useful tool to divide a period

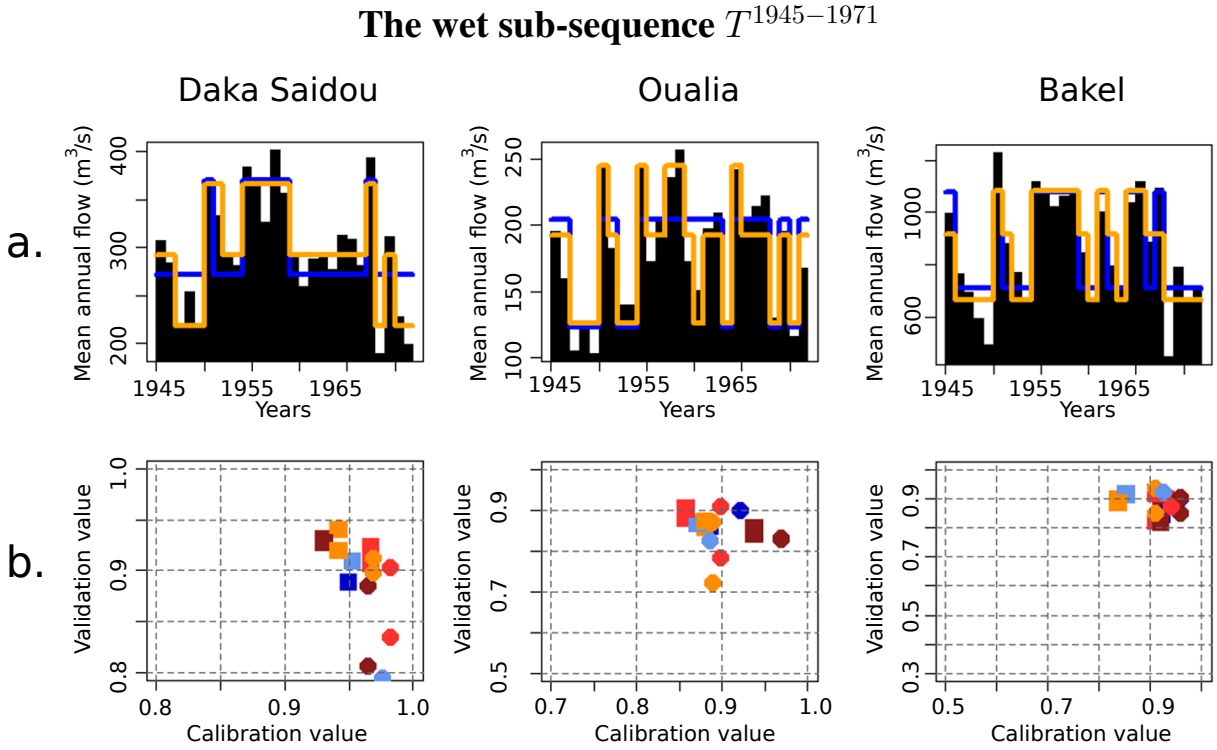

Figure S1. a. Years classifications of $T^{1945-1971}$ according to the Pettitt test (vertical red lines), 2-states-HMM (in blue) and 3-states-HMM (in orange); b. Scatter-plot of NSE (squares) and KGE (dots) calibration/validation values. Light green refers to the case 1 (Pettitt test, calibration on $T_{pettitt.wet}$ and validation on $T_{pettitt.dry}$); dark green to the case 2(Pettitt test, calibration on $T_{pettitt.dry}$ and validation on $T_{pettitt.wet}$); Light blue to the case 3 (2-states-HMM, calibration on $T_{2HMM.dry}$ and validation on $T_{2HMM.wet}$.), and dark blue to the opposite (case 4); Orange to the case 5 (3-states-HMM, calibration on $T_{3HMM.dry}$ and validation on $T_{3HMM.nor}$ and $T_{3HMM.wet}$); red to the case 6, and dark red to the seventh case.

into climate sub-sequences, and thus to apply a protocol following the differential split-sample test.

With no surprise, we note that the mean annual flow of dry sub-sequences ($T^{1945-1971}_{2HMM.dry}$ and $T^{1945-1971}_{3HMM.dry}$) are relatively high (in comparison with $T^{1940-1998}_{2HMM.dry}$ and $T^{1940-1998}_{3HMM.dry}$). We note that the HMM could provide a year classification in which the number of years is very small (for example: $T^{1945-1971}_{3HMMM.wet}$ for Oualia (5 years)).

In addition, transition probabilities values indicate that climate states for $T_{1945-1971}$ are less well distinct (close to 0.5) than for $T_{1940-1998}$ (close to 1 or 0).

The wet sub-sequence $T^{1945-1971}$

Daka Saidou

		Pettitt's Test		2-states HMM		3-states HMM		
	Sub-sequences(s)	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
Calibration	Dry		*/*	0.95/0.98(19y)		0.94/0.97(6y)		
	Normal						0.97/0.98(12y)	
	Wet	*/*			0.95/0.97(7y)			0.93/0.96(8y)
Validation	Dry	*/*			0.89/0.75(19y)		0.92/0.9(6y)	0.93/0.88(6y)
	Normal					0.94/0.91(12y)		0.93/0.81(12y)
	Wet		*/*	0.91/0.79(7y)		0.92/0.9(8y)	0.91/0.83(8y)	

Oualia

Calibration	Dry		*/*	0.87/0.88(8y)		0.88/0.89(9y)		
	Normal						0.86/0.9(12y)	
	Wet	*/*			0.89/0.92(18y)			0.94/0.97(5y)
Validation	Dry	*/*			0.86/0.9(8y)		0.89/0.91(9y)	0.85/0.83(9y)
	Normal					0.85/0.87(12y)		0.84/0.83(12y)
	Wet		*/*	0.87/0.82(18y)		0.87/0.72(5y)	0.9/0.78(5y)	

Bakel

Calibration	Dry		*/*	0.85/0.92(16y)		0.84/0.91(12y)		
	Normal						0.91/0.94(5y)	
	Wet	*/*			0.92/0.96(10y)			0.92/0.96(9y)
Validation	Dry	*/*			0.85/0.9(16y)		0.82/0.9(12y)	0.81/0.85(12y)
	Normal					0.9/0.94(5y)		0.91/0.91(5y)
	Wet		*/*	0.91/0.92(10y)		0.88/0.85(9y)	0.92/0.87(9y)	

Table S2. Table of NSE/KGE calibration and validation scores according to the seven cases for the wet sub-sequence $T^{1945-1971}$. As the Pettitt's test is not conclusive here, no calibration/validation scores are given (symbols $*/*$).

S2 Calibrations and validations on the dry historical episode $T^{1972-1998}$

Here, we apply our protocol on the period 1972-1998 ($T^{1972-1998}$), which can be consider as a dry historical episode of the SRB. The results are displayed in Figure S2, Table S3 and Table S4.

Likewise in S1, there is no clear trend and poor p-values are found with the Pettitt test while the examination of the transition probability matrices reveal a lower temporal persistence compared to the cases with the entire streamflow record.

Here, we are in the opposite situation than before: the wet states $T^{1972-1998}_{2HMM.wet}$ and $T^{1972-1998}_{3HMMM.wet}$ have relatively dry mean annual flows. Again, the HMM classification could yield to a number of years too small in some situations (example: $T^{1972-1998}_{3HMMM.dry}$ for Daka Saidou (5 years), $T^{1972-1998}_{3HMMM.nor}$ for Oualia (5 years)). This could constitute an issue for the calibration or the validation phase. These two points are addressed in the discussion section. Similarly to $T_{1945-1971}$, transition probabilities values indicate that climate states for $T_{1972-1998}$ are less well distinct (close to 0.5) than for $T_{1940-1998}$ (close to 1 or 0).

The dry historical episode $T^{1972-1998}$

Pettitt test	Basins	p value	Year break		
	Daka Saidou	0.277	-		
	Oualia	0.399	-		
	Bakel	0.474	-		
	Basins	μ	σ	δ	M
2-states-HMM	Daka Saidou	210.1;162.4	18.1;22.8	0,1	$\begin{bmatrix} 0.933 & 0.067 \\ 0.128 & 0.872 \end{bmatrix}$
	Oualia	88.4; 37.8	11.4; 25.8	1,0	$\begin{bmatrix} 0.781 & 0.219 \\ 0.185 & 0.814 \end{bmatrix}$
	Bakel	668.2;356.9	111.3;81.3	1,0	$\begin{bmatrix} 0.903 & 0.097 \\ 0.487 & 0.513 \end{bmatrix}$
3-states-HMM	Daka Saidou	210.8; 173.9;132.7	12.3; 12.8;18	0,0,1	$\begin{bmatrix} 0 & 0.757 & 0.243 \\ 0.307 & 0.693 & 0 \\ 0.129 & 0 & 0.871 \end{bmatrix}$
	Oualia	105;69.2;37.9	3.1;11.3;21.4	1,0,0	$\begin{bmatrix} 0.799 & 0 & 0.201 \\ 0.263 & 0.736 & 0 \\ 0.141 & 0.276 & 0.583 \end{bmatrix}$
	Bakel	686.1; 421.3; 315.6	44.2; 71.1; 93.5	1,0,0	$\begin{bmatrix} 0.482 & 0.518 & 0 \\ 0.756 & 0 & 0.244 \\ 0 & 0.5 & 0.5 \end{bmatrix}$

Table S3. Pettitt test results and Hidden Markov Model parameters (N=2 and N=3) for Daka Saidou, Oualia, and Bakel sub-basins, on the dry sub-sequence $T^{1972-1998}$.

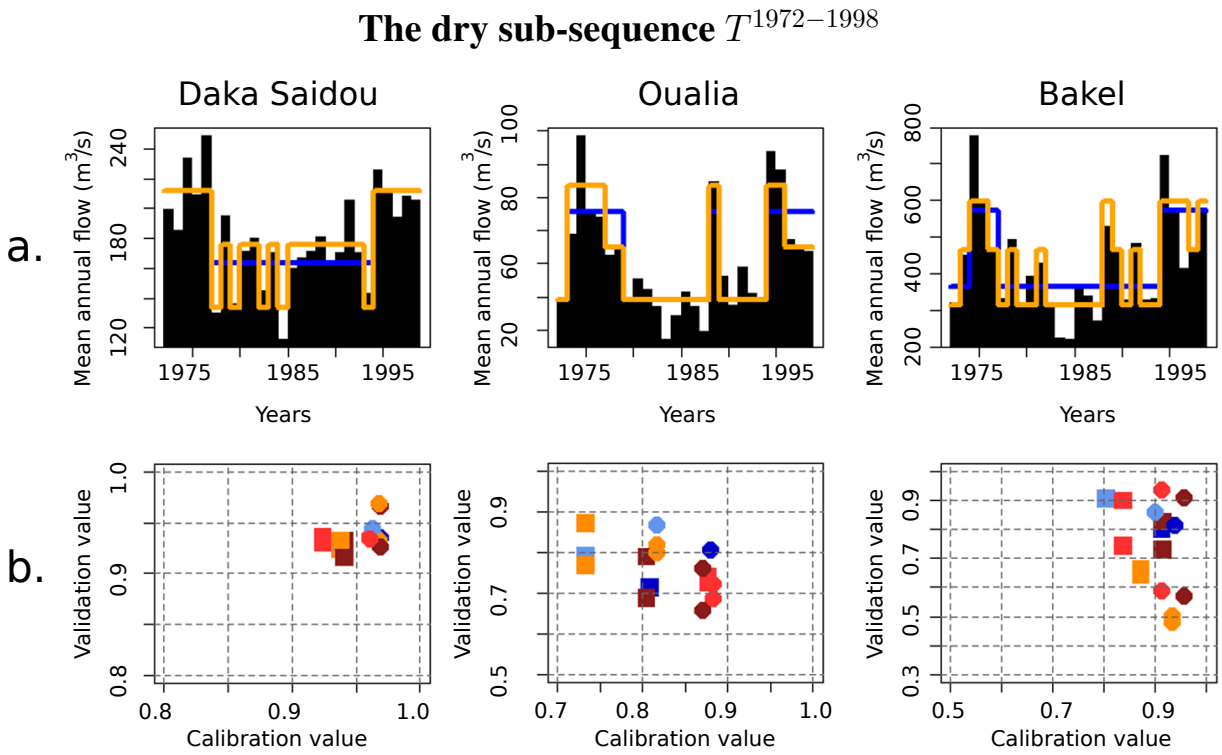

Figure S2. Same than Figure S1, but for $T^{1972-1998}$.

The dry sub-sequence $T^{1972-1998}$

Daka Saidou

		Pettitt's Test		2-states HMM		3-states HMM		
	Sub-sequence(s)	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
Calibration	Dry		*/*	0.93/0.96(17y)		0.94/0.97(5y)		
	Normal						0.92/0.96(12y)	
	Wet	*/*			0.94/0.97(9y)			0.94/0.97(9y)
Validation	Dry	*/*			0.92/0.93(17y)		0.94/0.94(5y)	0.93/0.97(5y)
	Normal					0.92/0.93(12y)		0.93/0.81(12y)
	Wet		*/*	0.93/0.94(9y)		0.93/0.97(9y)	0.93/0.93(9y)	

Oualia

Calibration	Dry		*/*	0.73/0.82(14y)		0.73/0.82(14y)		
	Normal						0.88/0.88(5y)	
	Wet	*/*			0.81/0.88(12y)			0.8/0.87(7y)
Validation	Dry	*/*			0.71/0.8(14y)		0.72/0.72(14y)	0.69/0.76(14y)
	Normal					0.87/0.8(5y)		0.79/0.66(5y)
	Wet		*/*	0.79/0.87(12y)		0.77/0.82(7y)	0.72/0.68(7y)	

Bakel

Calibration	Dry		*/*	0.73/0.82(18y)		0.73/0.82(12y)		
	Normal						0.88/0.88(7y)	
	Wet	*/*			0.81/0.88(8y)			0.8/0.87(7y)
Validation	Dry	*/*			0.71/0.8(18y)		0.72/0.72(12y)	0.69/0.76(12y)
	Normal					0.87/0.8(7y)		0.79/0.66(7y)
	Wet		*/*	0.79/0.87(8y)		0.77/0.82(7y)	0.74/0.68(7y)	

Table S4. Table of NSE/KGE calibration and validation scores according to the seven cases for the dry sub-sequence $T^{1972-1998}$. As the Pettitt's test is not conclusive here, no calibration/validation scores are given (symbols $*/*$).