Single-objective vs. multi-objective autocalibration in modelling total suspended solids and phosphorus in a small agricultural watershed with SWAT

Santatriniaina Denise Rasolomanana, Paul Lessard and Peter A. Vanrolleghem

ABSTRACT

To obtain greater precision in modelling small agricultural watersheds, a shorter simulation time step is beneficial. A daily time step better represents the dynamics of pollutants in the river and provides more realistic simulation results. However, with a daily evaluation performance, good fits are rarely obtained. With the Shuffled Complex Evolution (SCE) method embedded in the Soil and Water Assessment Tool (SWAT), two calibration approaches are available, single-objective or multi-objective optimization. The goal of the present study is to evaluate which approach can improve the daily performance with SWAT, in modelling flow (Q), total suspended solids (TSS) and total phosphorus (TP). The influence of weights assigned to the different variables included in the objective function has also been tested. The results showed that: (i) the model performance depends not only on the choice of calibration approach, but essentially on the influential parameters; (ii) the multi-objective calibration estimating at once all parameters related to all measured variables is the best approach to model Q, TSS and TP; (iii) changing weights does not improve model performance; and (iv) with a single-objective optimization, an excellent water quality modelling performance may hide a loss of performance of predicting flows and unbalanced internal model components.

Key words | calibration, parameter estimation, sensitivity analysis, water quality modeling

Santatriniaina Denise Rasolomanana (corresponding author)
Paul Lessard
Peter A. Vanrolleghem
Département de génie civil et de génie des eaux,
Université Laval, Québec,
Pavillon Adrien-Pouliot,
1605 Avenue de la Médecine,
Québec (QC), G1V0A6,
Canada
E-mail: *santatriniaina-d.rasolomanana.1@ulaval.ca*

INTRODUCTION

Small agricultural watersheds need to be modelled for better management of water resources, although usually only few data are available. Here, the model used for simulating the fate of pollutants and identifying the best management practices is the Soil and Water Assessment Tool (SWAT) (Arnold *et al.* 1998). To achieve better precision in modelling small watersheds, a shorter simulation time step is useful. A daily time step better represents the dynamics of pollutants in the river and provides more realistic simulation results. However, case studies on small agricultural watersheds using daily time steps are rare, given that SWAT was originally developed for large watersheds with large amounts of data. In addition, river water quality modelling performance is generally carried out on a monthly or yearly time step and rarely on a daily time step. Among the literature reviewed (Gassman *et al.* 2007; Moriasi *et al.* 2007), only a few case studies show a good daily performance on water quality.

Given this context, the goal of the present study is to find how to improve water quality modelling performance with SWAT on a daily time step. There are two calibration approaches for modelling flow (Q), total suspended solids (TSS) and total phosphorus (TP): single-objective and multi-objective optimization. The present study compares these two approaches in terms of their performance in modelling TSS and TP at a daily time step in small rural watersheds. Given that there are too many parameters due to the complexity of the model in comparison with the amount of data, a sensitivity analysis is necessary to identify the most important parameters. Each variable is sensitive to different parameters and in case of many variables some parameters appear in multiple subsets. So, two types of parameters will be considered: (i) those only related to the variable of interest; and (ii) all those influencing all variables. In addition, the influence of weights assigned to the different

doi: 10.2166/wst.2012.895

objective functions in the case of multi-objective optimization has been tested. Indeed, the calibration algorithm prioritizes fitting the most numerous data and the higher valued data that can induce large global errors. In that sense, the phosphorus data are the most disadvantaged, as they are small in number and magnitude, explaining the difficulty of the model to fit phosphorus data. In this study, the weights will be chosen according to the typical measurement errors and the model fitting errors.

MATERIALS AND METHODS

Study area description

The study was conducted on the Ruisseau du Portage watershed, a 21.41 km^2 small agricultural watershed located in the Boyer river basin in Québec, Canada. Based on the bacteriological and physico-chemical index, the water quality in the watershed is described as 'bad' to 'very poor' due to high turbidity and excessive enrichment of its water by nutrients (nitrogen and phosphorus) (Ministère du Développement Durable, Environnement et Parcs, Québec or MDDEP). The major sources that can affect its water quality originate from agricultural activities taking place in the lower reaches of the basin. The territory is composed of 48% forest, 44% agriculture (6.88% cereals, 0.13% corn, 36.97% grassland and pasture) and 8% wetlands.

This study focuses on data collected between October 1999 and December 2002 for Q, TSS and TP. Precipitation and temperature of the site average, over a year, respectively 1,300 mm and 5.25 °C. The climate is temperate continental. The topography is relatively flat, the altitude ranging from 46 to 117 m, with an average of 86 m. The slopes range from 1.6 to 3.1%, those closest to the outlet being most pronounced. The soil characteristics vary according to the area occupied, the major ones being stony sandy loam, gravelly sandy loam and gravelly loam (Baril & Rochefort 1957; Marcoux 1966; Pageau 1976; Ouellet *et al*. 1995).

Input data

Input data used for modelling are the following:

1. Digital Elevation Model (DEM): produced by Geobase (www.GeoBase.ca), 1:50,000, grid 23.17 m.
2. Soil map: from IRDA (Institut de Recherche et de Développement en Agroenvironnement, Québec, Canada), 1:20,000.
3. River map: produced by BDTQ (Base de Données Topographiques du Québec), 1:20,000.
4. Land use: provided by Canards Illimités, grid 25 m.
5. Hydrometeorological data: from the MDDEP and Service Météorologique Canada (SMC).

Observed data

Streamflow data (Figure 2(a)) were collected by the Centre d'Expertise Hydrique du Québec (CEHQ) while water quality data (Figure 2(b) and (c)) were obtained from the Centre d'Expertise en Analyse Environnementale du Québec (CEAEQ).

Hydrological records

Measured daily streamflow data show an interannual average of 0.35 m^3/s. The flow variations closely follow the variations of precipitation, showing that the basin responds quickly to rainfall due to its small size. The peak flows occur during snowmelt (April–May), causing 46% of the runoff of the entire study period (October 1999 to December 2002), while the flows are lowest in winter (January to March) and summer (July to September).

Water quality records

Water quality data were discontinuous grab samples that are therefore not representative of the whole day, especially when agricultural activities, such as manure spreading, or rainfall occur. The median TSS value of 7 mg/L is slightly less than the magnitude of the median measured in 16 small agricultural catchments (9.25 mg/L), but 3–4 times higher than those measured in 30 forested catchments in Québec (2 mg/L) (Gangbazo & Babin 2000). A concentration of 4 to 5 mg/L of TSS persists throughout the year, which is harmful to aquatic life because a standard of 5 mg/L has been set for chronic toxicity (Gangbazo & Le Page 2005). The median TP, 0.05 mg/L, is slightly above the criterion for the prevention of eutrophication set at 0.03 mg/L in Quebec (Menviq 1990, rév 1992). The daily concentrations of TP fluctuate much during the year, with peaks occurring in April–May and August–September. Around 55% of the TP is soluble. During rain events, given the various land use types, the level of phosphorus in the river is not necessarily high.

The watershed model

The watershed model used in this study is the 2005 version of the Soil and Water Assessment Tool (SWAT) (Arnold et al. 1998), which is one of the most widely used watershed models in the world.

For the simulations, the study site is divided into 5 subbasins and 33 Hydrologic Response Units (HRU) (64 ha on average). A sensitivity analysis of all parameters related to the 3 variables (flow, TSS, TP) was done. Each variable is sensitive to different parameters and in case multiple variables are considered, some parameters appear in multiple subsets: flow and TSS, flow and TP, flow and TSS and TP. The parameters used for calibration are presented in the appendix (Tables A.1 to A.4). Calibrations were performed by using these different sets of parameters (see Results section).

After implementation of the model, simulations were carried out from January 1, 1998 to December 9, 2002, including:

- January 1, 1998 to October 3, 1999: warm-up,
- October 4, 1999 to July 31, 2001: calibration period, and
- August 1, 2001 to December 9, 2002: validation period.

The simulations were conducted throughout the year but the calibration focused on summer (June–October 2000) because of the higher water quality standard requirements applying at this time of the year to meet the increased use of water for recreational and domestic activities during the summer months. The parameter intervals were defined based on the recommendations in Neitsch et al. (2005), except for the parameters related to the sediment re-entrained during channel sediment routing (SPCON, SPEXP), the sediment concentration in lateral and groundwater flow (LAT_SED) and the depth to the subsurface drain (DDRAIN), which all needed adjustments. The predefined SPCON and SPEXP parameter range were not adequate for small watersheds and low flows, the LAT_SED parameter interval was too large causing excess export of sediment and the DDRAIN parameter range was narrowed based on drainage data.

Optimization in SWAT 2005: shuffled complex evolution algorithm-uncertainty analysis (SCE-UA)

SWAT2005 includes an automatic multi-objective calibration and uncertainty analysis in a single run, called Parasol (Parameter Solutions method), developed by van Griensven & Bauwens (2003). The calibration procedure, based on the 'Shuffled Complex Evolution' algorithm or SCE, is a global search algorithm for the minimization of a single function (Duan et al. 1993).

The optimization can be single-objective or multi-objective. For single-objective optimization, there is only one objective function (OF) that needs to be optimized. For multi-objective optimization problems, a series of OFs need to be taken into account simultaneously. The most commonly used OF is the Sum of the Squares of the Residuals (SSQ):

$$\mathrm{SSQ} = \sum_{i=1}^{n} (O_i - S_i)^2$$

where n is the number of pairs of observed (O_i) and simulated (S_i) variables.

For multi-objective calibration, a single global optimization criterion (GOC), defined as an aggregation of several objective functions, is computed as follows:

$$\mathrm{GOC} = \sum_j \frac{\mathrm{SSQ}_j * n_{\mathrm{obs},j}}{\mathrm{SSQ}_{\mathrm{min},j}}$$

with j the number of objective functions. Thus, OFs get weights that are equal to the number of observations ($n_{\mathrm{obs},j}$) divided by the minimum of the objective function ($\mathrm{SSQ}_{\mathrm{min},j}$) (van Griensven 2006).

The methodology used for the two calibration approaches is described below.

Calibration approaches

The most common single-objective approach is to successively calibrate flow, TSS and TP, while the second, the multi-objective approach is to calibrate several components in a single optimization run. A general calibration procedure chart for both single and multi-objective optimization for flow, sediment and total phosphorus is presented in Figure 1. The single-objective calibration techniques are summarized on the SWAT website (http://www.brc.tamus.edu/swat/publications/swat-calibration-techniques_slides.pdf). The multi-objective optimization procedure differs after flow calibration. The flow is re-calibrated with the TSS and TP and the performance criteria are readjusted.

The daily model evaluation limits, which are less strict than the monthly ones because of the lack of averaging over multiple data (Engel et al. 2007), have been adjusted from the monthly evaluation guidelines proposed by Moriasi et al. (2007). Each step is evaluated using two criteria, the

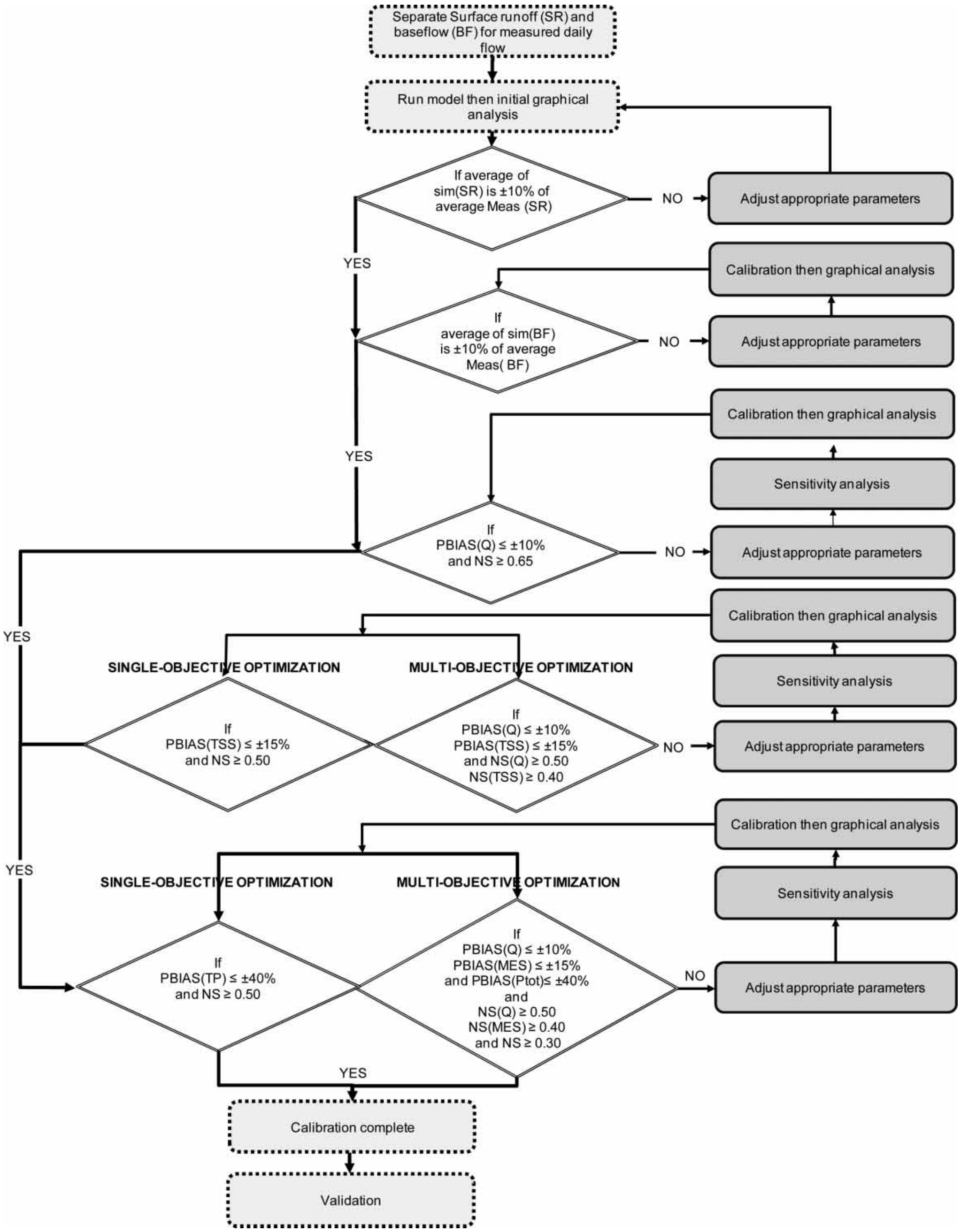

Figure 1 | General single and multi-objective calibration procedure for flow, sediment and total phosphorus in the watershed model.

Nash Sutcliffe Efficiency (NSE) (Nash & Sutcliffe 1970) and the Per cent of Bias (PBIAS) (Moriasi *et al.* 2007).

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n} (O_i - S_i)^2}{\sum_{i=1}^{n} (O_i - \bar{O})^2}$$

$$\mathrm{PBIAS} = \frac{\sum_{i=1}^{n} (O_i - S_i) * 100}{\sum_{i=1}^{n} O_i}$$

where n is the number of pairs of observed (O_i) and simulated (S_i) variables.

The NSE values range from $-\infty$ to 1, with 1 being the optimal value. Negative values indicate that the average of the observed values is a better model than the model fitted to the data, leading to rejection of the model. As for PBIAS, it measures the average trend of the simulated data to be above or below the observed data. The optimum value of PBIAS is zero, indicating a perfect model simulation. A positive PBIAS indicates an underestimation of the model while a negative PBIAS represents an overestimation of the model. This test is recommended because of its ability to clearly demonstrate the poor performance of the model (Gupta *et al*. 1999).

RESULTS AND DISCUSSION

Table 1 and Figure 2 compare some of the results obtained after the single-objective and the multi-objective calibrations, before and after adjusting the internal mass balance.

Mass balance

Among the components of an internal mass balance, surface runoff, baseflow, sediment and phosphorus export were considered as very important in flow, TSS and TP calibration. Internal hydrological components are unbalanced when their per cent bias exceeds the evaluation criteria defined in Figure 1, based on reference data. As reference, surface runoff and baseflow data were obtained with the baseflow filter program and sediment and phosphorus loads were estimated with the Flux 5.0 software (Walker 1998).

In both automatic calibrations, a poor mass balance for Q, TSS and TP was noted: the bias of the baseflow and sediment export is in general too high and phosphorus export too low (Figure 3(a)). To get a more realistic model, manual adjustments of certain parameters were undertaken, and only realistic changes of the parameters were allowed during calibration (e.g. little change by percentage for the parameters related to the geomorphology of the basin). The new results are shown in Table 1(b) and Figure 2(d). The number of influencing parameters is not the same, given that some parameters were fixed.

After adjusting the mass balance, the model performance is worse but the results are more realistic. Figure 3(b) shows that almost the same mass balance is obtained for all cases.

Parameters

Four types of parameters were considered:

- parameters only influencing TSS
- parameters only influencing TP
- parameters influencing Q-TSS and
- parameters influencing Q-TSS-TP.

Please note that the parameters influencing the flow have been set to the values obtained by fitting to the flow data, unless specified otherwise.

Model performance

Single-objective calibration

With only parameters relating to TSS being estimated, the performances obtained are capped at a certain threshold (NSE 0.18 before mass balance adjustment and 0.17 after). Subsequently, by reconsidering the parameters influencing flow in the calibration of TSS (column 3 in Table 1), the performance could be improved significantly for TSS (Figure 2(b)) (NSE 0.49 and 0.32 before and after adjustment, PBIAS less than 5%) at the expense of the flow's performance: NSE dropping from 0.65 to 0.15, PBIAS 53% (Figure 2(a)). After adjustment of parameters to make the mass balance fit, the NSE dropped to −0.28 and PBIAS to −13%.

Multi-objective calibration

On the other hand, these calibrations were also conducted in the multi-objective way using the same sets of influential parameters. With only TSS-influencing parameters, the performance of Q is a little bit improved but the fit to TSS is worse than in the single-objective optimization. With Q-TSS-influencing parameters, the Q performance was kept at the expense of the TSS performance. For TP, the multi-objective approach leads to better results with TP-influencing parameters, especially the TP performance was enhanced very much (Figure 2(c)) (from NSE −0.29 in single-objective to +0.29 in multi-objective optimization) with all influencing parameters. Unfortunately, such good results couldn't be obtained when the mass balance was adjusted for.

In both calibrations, and after adjusting the mass balance, the following results were obtained:

1. when considering only TSS or TP influencing parameters, both calibration approaches gave the same performance given that flow parameters are no longer touched;
2. by considering all influencing parameters, the TP performance deteriorated significantly. Indeed, for a better mass balance, too many parameters had to be fixed and SWAT was no longer able to optimize for TSS and TP. Considering only TP influencing parameters, gave better results after adjusting for the mass balance.

Table 1 | Comparison of results obtained after single-objective and multi-objective calibration, with different combinations of parameters and objective functions

	Reference	Single-objective optimization				Multi-objective optimization			
OF		TSS	TSS	TP	TP	Q-TSS	Q-TSS	Q-TSS-TP	Q-TSS-TP
Parameters influencing		TSS	Q-TSS	TP	Q-TSS-TP	TSS	Q-TSS	TP	Q-TSS-TP
(a) Before adjusting mass balance									
Number of parameters		10	21	13	19	10	21	13	19
NSE[a](Q)	≥0.50	0.65	0.15	0.53	0.56	0.69	0.67	0.53	0.59
NSE[a] (TSS)	≥0.40	0.18	0.49	0.38	0.31	0.20	0.11	0.39	0.34
NSE[a] (TP)	≥0.30			−0.36	−0.29			−0.38	0.29
PBIAS[a] (Q)(%)	≤± 10%	6.72	53.19	3.8	15.53	−10.57	11.94	15.16	11.71
PBIAS[a] (TSS)(%)	≤± 15%	6.36	0.74	7.81	10.48	16.25	7.27	7.95	11.12
PBIAS[a] (TP)(%)	≤± 40%			13.77	10.05			3.58	3.29
(b) After adjusting mass balance									
Number of parameters		10	25	11	21	10	25	11	21
NSE[a] (Q)	≥0.50	0.54	−0.28	0.70	0.62	0.54	0.65	0.70	0.64
NSE[a] (TSS)	≥0.40	0.17	0.32	0.21	−9E07	0.17	0.19	0.21	0.1
NSE[a] (TP)	≥0.30			−0.56	−1.17			−0.56	−1.19
PBIAS[a] (Q)(%)	≤± 10%	−15.77	−13.31	0.68	−1.31	−15.77	−11.74	0.68	−3.68
PBIAS[a] (TSS)(%)	≤± 15%	10.47	4.85	17.74	−5E05	10.47	11.03	17.74	12.97
PBIAS[a] (TP)(%)	≤± 40%			43.30	30.04			43.30	31.80

[a]NSE: Nash Sutcliffe Efficiency; PBIAS: Per cent of Bias.

Validation

We tried to validate the model using the parameters obtained using the multi-objective approach with all influential parameters. The validation run is shown in Figure 2 and the performances obtained are not very good (NSE (Q) 0.25, NSE (TSS) −0.11, NSE (TP) −1.07, PBIAS (Q) 2.07%, PBIAS (TSS) 24.71%, PBIAS (TP) −30.22%). However, this result is not unexpected given the fact that the performance in calibration was only good for the flow predictions and poor performance was obtained for TP and TSS. Further work will be needed to simultaneously get good fitting performance and a mass balance that holds. For instance, data on surface runoff and baseflow could be used in an extended multi-objective setting, i.e. fitting to three flow data series, TSS and TP.

Influence of weights in multi-objective optimization

In addition, to improve the multi-objective calibration performance, the influence of the weights assigned to the individual objective functions has been tested. Indeed, the algorithm prioritizes the most numerous data and the higher valued data that can induce large global errors. The phosphorus data are the most disadvantaged, as they are small in number and magnitude, explaining the difficulty of the model to fit phosphorus data. Given that the user-defined choice of weights with SCE in SWAT2005 is not operational, we have manually tried to add weights calculated according to the measurement errors. Smaller weights were given to variables that were accepted to be less important in the strategy of optimum search (van Griensven & Bauwens 2003). The measurement errors taken into account were 5% for Q, 15% for TSS and 10% for TP. The estimation methodology adopted is as follows: m influential parameters were selected after sensitivity analysis among all parameters related to Q, TSS and TP; n initial parameter estimates were produced with these m influential parameters by using latin hypercube sampling (van Griensven 2006). Thereafter, multi-objective calibrations were carried out, each with a maximum of 20,000 tries. The GOC was computed by trying various weights and evaluating the objective functions for each of the large number of simulations (in total, 400,000 simulations were carried out, $n = 20$, 20,000 tries each), with their corresponding parameter values. After ranking, the minimal GOC was identified and the corresponding parameters were the optimal ones for a particular set of weights.

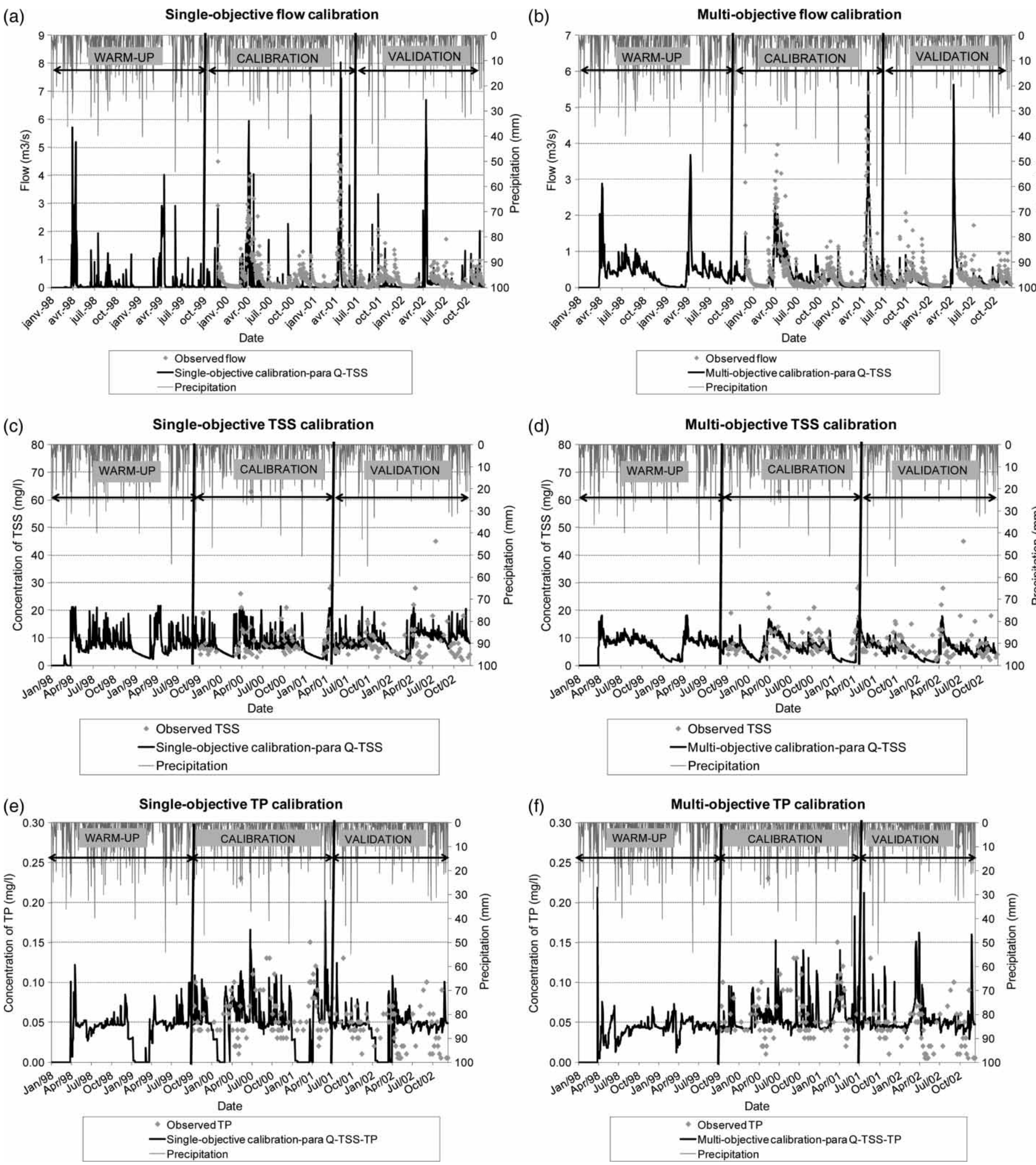

Figure 2 | Some results of single-objective versus multi-objective calibration before mass balance adjustment: (a) and (b) flow; (c) and (d) TSS; (e) and (f) TP.

Through this weighted multi-objective optimization:

- the flow was very well simulated, with NSE between 0.62 and 0.75;
- TSS-performance was good (NSE 0.11 to 0.35);
- the best TP-performance was a NSE of 0.06;
- the internal hydrologic components were very unrealistic: the surface runoff was too high or non-existant, sediment loads were uncontrolled and TP loads very low.

Therefore, we can conclude that the two multi-objective optimization approaches tested, one with weights based on

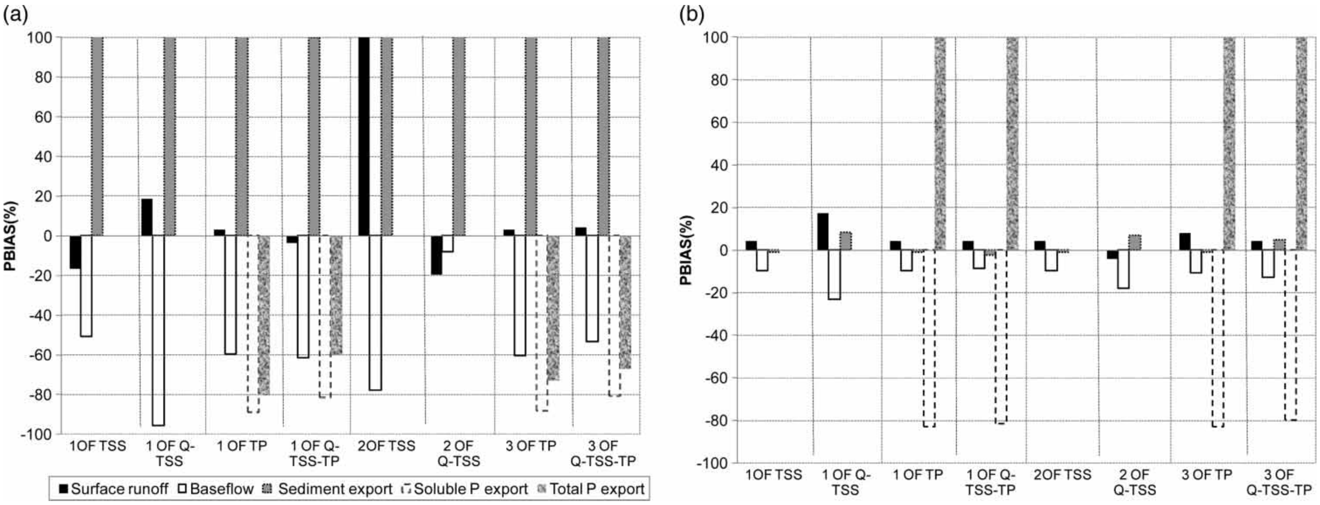

Figure 3 | Mass balance: (a) before adjustment; (b) after adjustment.

measurement errors and the other one with the number of observations divided by the minimum objective function (weights imposed by SWAT2005), lead to a calibrated model with the same performance for Q and TSS. However, for TP, with which it was so difficult to get good performance with a simultaneous good model fit of Q and TSS, the multi-objective optimization gives worse performance when using the search strategy with weights based on measurement errors.

Other optimization algorithms such as SUFI2, NSGA-II, coupled with SWAT, may lead to better performance (Abbaspour *et al*. 2007; Zhang *et al*. 2010). Moreover, subjectivity in the choice of weights is one of the main challenges in multi-objective optimization.

In addition, a good fit of the hydrograph and good values of the performance criteria of the model do not guarantee a correct distribution of the internal components of the model, namely surface runoff, groundwater flow, tile drainage, export of sediment and nutrients. The optimizer does not care about the realism of the parameters and internal components of the model. More data may be needed on these internal components. That is the reason why manual adjustments of parameters or routines in the source code play a crucial role before, during and after calibration.

CONCLUSION

A multi-objective optimization using a modified SCE-UA algorithm, is incorporated in SWAT2005. Two calibration approaches are possible: single-objective and multi-objective optimization. The obtained model performance depends on the choice of calibration approach, but essentially on the selected influencing parameters. Indeed, each variable is sensitive to different parameters and in case of many variables, some parameters appear in multiple subsets: flow and TSS, flow and TP, flow and TSS and TP. Considering them all for calibration improved the obtained water quality fitting performance very much. Based on the results obtained in this study, even if the user-defined choice of weights with SCE in SWAT2005 is not operational, the multi-objective calibration remains the best approach to model TSS and TP, with a daily evaluation performance in the small agricultural Ruisseau du Portage watershed. The following conclusions can be drawn:

1. The multi-objective optimization considering all parameters related to the variables is the best approach to enhance the daily water quality simulation with SWAT2005. The performance of describing flow data is maintained and the water quality prediction performance, especially that of TP, is very much improved.
2. Excellent results on the whole watershed may hide unrealistic mass balances for Q, TSS and TP for each HRU. Forcing a correct mass balance for each HRU leads to a worse daily performance, and fixing certain parameters to impose the mass balance hinders the search for a parameter set that gives adequate model performance.
3. Despite the normalization of the objective functions, the SCE algorithm incorporated in SWAT2005 prioritizes the most numerous data among the variables considered. To overcome this problem, the choice of other weights

assigned to objective functions can be a solution but this choice is not operational in SWAT2005. The attempt to change the weights manually did not improve the performance to describe TP data.

4. With single-objective optimization, the excellent water quality performance that can be achieved may hide a loss of flow fitting performance and unbalanced internal hydrological components.
5. For both calibration approaches, manual adjustments based on good insight into the SWAT model remain crucial.

ACKNOWLEDGEMENTS

This study was supported by the Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT). Special thanks to the «SWAT team»: Etienne Lévesque, Annie-Claude Parent (Université Laval), Isabelle Beaudin and Aubert Michaud (Institut de Recherche et de Développement en Agroenvironnement or IRDA), Flora Umuhire and Ann van Griensven (UNESCO-IHE Institute for Water Education) for their valuable help. Peter Vanrolleghem holds the Canada Research Chair in Water Quality Modelling.

REFERENCES

Abbaspour, K. C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J. & Srinivasan, R. 2007 Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *J. Hydrol.* **333**, 413–430.

Arnold, J. G., Srinivasan, R., Muttiah, R. S. & Williams, J. R. 1998 Large-area hydrologic modeling and assessment: part I. Model development. *JAWRA* **34** (1), 73–89.

Baril, R. & Rochefort, B. 1957 *Étude pédologique du comté de Lotbinière dans la province de Québec*, Service des fermes expérimentales, Ministère fédéral de l'Agriculture en collaboration avec le Ministère de l'Agriculture de Québec et l'École supérieure d'Agriculture, Saint-Anne-de-la-Pocatière. 115 pp+1 carte couleur à l'échelle du 1:63 360 (in French).

Duan, Q. D. 2003 Global optimization for watershed model calibration. In: *Calibration of Watershed Models* (Q. Duan, S. Sorooshian, H. V. Gupta, A. N. Rousseau & R. Turcotte, eds), AGU, Washington DC.

Duan, Q. Y., Gupta, V. K. & Sorooshian, S. 1993 Shuffled complex evolution approach for effective and efficient global minimization. *J. Optim. Theor. Appl.* **76** (3), 501–521.

Engel, B., Storm, D., White, M. & Arnold, J. G. 2007 A hydrologic/water quality model application protocol. *JAWRA* **43** (5), 1223–1236.

Gangbazo, G. & Babin, F. 2000 Pollution de l'eau des rivières dans les bassins versants agricoles. *Vect. Env.* **33**, 47–57 (in French).

Gangbazo, G. & Le Page, A. 2005 *Détermination d'objectifs relatifs à la réduction des charges d'azote, de phosphore et de matières en suspension dans les bassins versants prioritaires*, Ministère du Développement Durable, Environnement et Parcs (in French).

Gassman, P. W., Reyes, M. R., Green, C. H. & Arnold, J. G. 2007 The soil and water assessment tool: historical development, applications and future research directions. *Trans. ASABE* **50** (4), 1211–1250.

Gupta, H. V., Sorooshian, S. & Yapo, P. O. 1999 Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. *J. Hydrol. Eng.* **4** (2), 135–143.

Holland, J. H. 1975 *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor.

Marcoux, R. 1966 *Étude pédologique des comtés de Bellechasse et de Montmagny*, Ministère de l'agriculture et de la colonisation du Québec. Division des sols, Service de la recherche. Bulletin technique no 12.72 pp=2 cartes couleurs à l'échelle du 1: 63 360 (in French).

Menviq. 1990, rév. 1992 *Critères de qualité de l'eau*, Direction de la qualité des cours d'eau, ministère de l'Environnement du Québec, Sainte-Foy, Québec, Canada. 423pp (in French).

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D. & Veith, T. L. 2007 Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* **50** (3), 885–900.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models; part I – a discussion of principles. *J. Hydrol.* **10**, 282–290.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R. & Williams, J. R. 2005 *Soil and Water Assessment Tool Theoretical Documentation, Version 2005*. Grassland, Soil and Water Research Laboratory – Agricultural Research Service. Blackland Research Center – Texas Agricultural Experiment Station.

Nelder, J. A. & Mead, R. 1965 A simplex method for function minimization. *Comput. J.* **7** (4), 308–313.

Ouellet, L., Rompré, M., Carrier, D. & Laflamme, G. 1995 *Étude pédologique du comté de Beauce, Québec*, Service des sols, Direction de la recherche et du développement, Ministère de l'Agriculture, des Pêcheries et de l'Alimentation du Québec. 81 pp. + 2 cartes couleurs à l'échelle du 1:50,000 (in French).

Pageau, E. 1976 *Étude pédologique du comté de Dorchester, Agriculture Québec*. Division des Sols. 94 pp. +4 carte couleur à l'échelle du 1:50 000 (in French).

van Griensven, A. 2006 *Sensitivity, Auto-calibration, Uncertainty and Model Evaluation in SWAT2005*. Technical Report, Ghent University, Belgium.

van Griensven, A. & Bauwens, W. 2003 Multi-objective autocalibration for semidistributed water quality models. *Water Resour. Res.* **39** (12), SWC 9.1–SWC 9.9.

Walker, W. 1998 *Flux, Stream Loads Computation, Version 5.0*. Environmental Laboratory USAE Waterways Experiment Station. Vichsburg, Mississipi, USA.

Williams, J. R. 1995 The EPIC model. In: *Computer Models of Watershed Hydrology* (V. P. Singh, ed.). Water Resources Publications, Highlands Ranch CO, pp. 909–1000.

Williams, J. R. & Berndt, H. D. 1977 Sediment yield prediction based on watershed hydrology. *Trans. ASAE* **20** (6), 1100–1104.

Zhang, X., Srinivasan, R. & Van Liew, M. W. 2010 On the use of multi-algorithm, genetically adaptive multi-objective method for multi-site calibration of the SWAT model. *Hydrol. Process.* **24**, 955–969.

First received 14 June 2011; accepted in revised form 4 October 2011