Journal of Civil Engineering (IEB), 42 (1) (2014) 1-37

Calibration of hydrologic model in a data-limited watershed: Tay Basin, Canada

Ferdous Ahmed

Rideau Valley Conservation Authority
3889 Rideau Valley Drive, Manotick, Ontario K4M 1A5, Canada

Received 16 April 2012

Abstract

An integrated hydrologic/hydraulic model of the Tay River watershed has been constructed using the Mike11 modeling system of the Danish Hydraulic Institute. This watershed system is complex, comprising of channels, local drainage areas, lateral inflows, numerous lakes, and three regulated dams. The availability of streamflow data is rather limited and its quality is poor, making model calibration challenging. The model was calibrated using measured streamflow data for four years. A wide range of methods – both qualitative and quantitative – were used to evaluate the model performance. It was found that the model performance was acceptable by industry standard but relatively inferior compared to other models in adjacent areas. Future steps necessary to improve the model have been identified.

Keywords: Hydrologic model, Mike11, Tay River, Rideau Valley, Calibration, Validation, Hydraulic model, Watershed model

---

1. Introduction

The Tay River is one of the major rivers within the Rideau Valley Conservation Authority (RVCA), which is one of the 36 watershed-based agencies in Ontario, Canada (Fig. 1). These agencies have as their mandate the overall management of the watershed, including the water resources, forestry, fisheries, ecology and land use planning. In order to carry out its mandate with a more scientific and quantitative rigor, RVCA has recently developed comprehensive watershed models (RVCA 2007, Ahmed 2010). On of those models is for the Tay River watershed (Fig. 2), a complex system of streams, contributing drainage areas, lateral inflows, numerous lakes, and three actively operated dams. The Mike11 modeling platform of the Danish Hydraulic Institute (DHI 2003, 2004) was used in this study. The development and calibration of the model are described here.

The full description of the initial model development and other background information are available in an unpublished RVCA report (RVCA 2007). This original model, with some modifications and corrections, is called 2007B here. The original model was later extended and refined using additional surveyed data on cross-sections and river crossings; it is called 2008B.

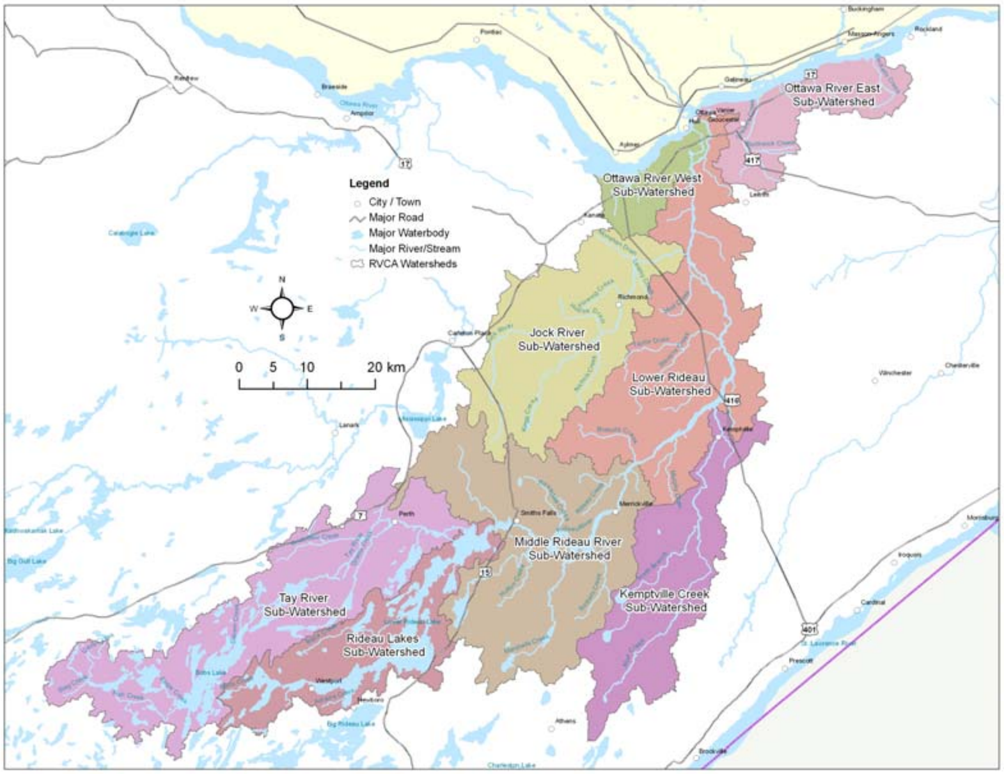


Fig. 1. Rideau valley watershed

It is rather ironic that numerical models are more useful and thus in greater demand in watersheds with sparse data sets, where they are harder to calibrate and validate. However, such is the fate of practical engineers who are charged with the task of producing hydrological estimates in the absence of adequate data. Moreover, hydrologic data collection in most jurisdictions started long before anyone thought about its use in numerical models. Now, with available computing power and GIS-based data of watershed characteristics, the lack of long enough streamflow data for calibration has suddenly become the bottleneck of successful hydrologic modeling.

With abundant computing power, a good number of well-established modeling software, and data in geospatial format, the modeler now has the option to build elaborate models of large and complex systems in great detail. Therefore, as Guzha and Hardy (2010) has pointed out, the current thrust in hydrological modeling is the use of distributed models together with GIS-based data, which enables subdividing the watershed into as small units as necessary to utilize available geospatial data. Also, in recognition of the relative exactness of hydraulic computation compared to the considerable uncertainty in hydrologic computation, modelers now try to maximize overall model performance by making more elaborate use of hydraulic components.

In this case of the Tay River, streamflow data was short in record length with many gaps; so was the dam operation data. All these made the calibration challenging, although model 2008B showed some improvement over 2007B. We looked for ways to improve the model performance, mainly with finer discretization of the basin and more river cross-sections. We report it an example of practical modeling exercise – for the benefit of practicing engineers – and in the spirit of Andréassian et al. (2010) who advocated analyzing and reporting challenging modeling exercise in an effort to learn from them.

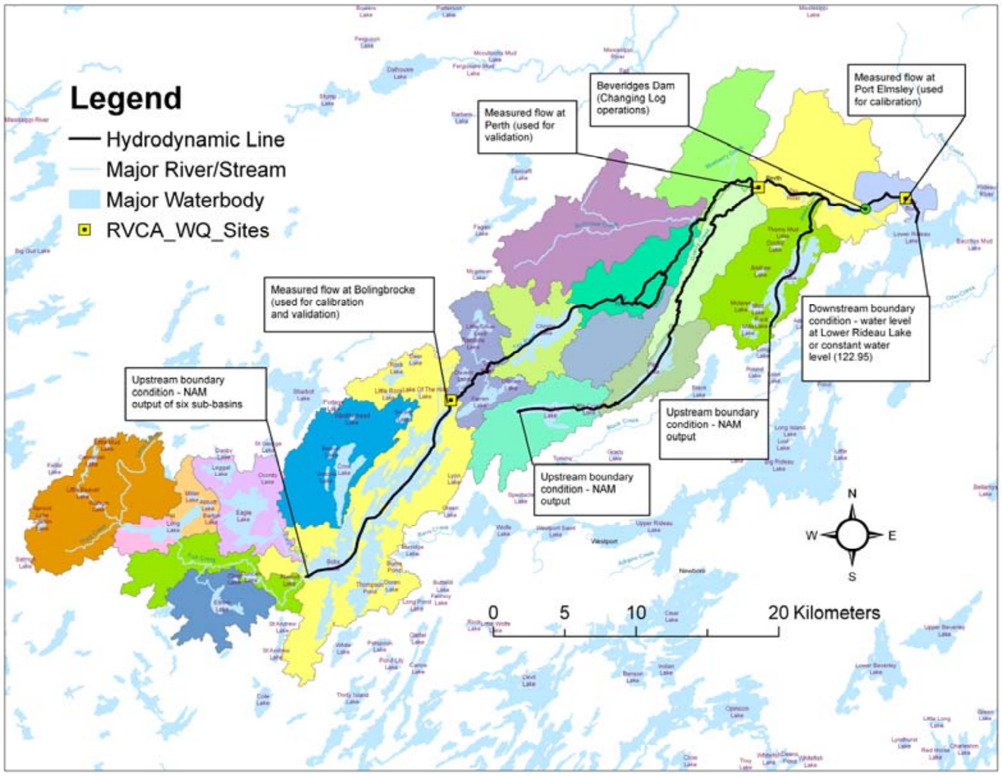


Fig. 2. Tay River basin model schematic

2. Description of the watershed

The Tay River basin is one of the eight sub-watersheds within the jurisdiction of RVCA (Fig. 1 and 2). It has an area of 797 km^2 and drains to the Lower Rideau Lake, which is a part of the Rideau River which in turn joins the Ottawa River further downstream. The Tay basin is part of the Canadian Shield, characterized by numerous rock outcrops (Precambrian bedrock) and thin overburden (< 2m). Surface geology is predominantly loam (about 25% of the basin area) and loamy sand (75%). Land use is a mix of cropland (9%), pasture (9%), forests (60%), wetlands (6%) and open water lakes (14%) (Table 1). The wetlands are situated mainly along the streams. The lakes are numerous and vary in size from large to small. The lakes and wetlands make the hydrologic response rather subdued.

Table 1
Land use in Tay basin

Land Use	Area (km2)	% of Total Area
Alvar	0.38	0.05
Conifer swamp	8.59	1.08
Cropland	69.01	8.66
Deciduous swamp	31.17	3.91
Dense coniferous forest	18.28	2.29
Dense deciduous forest	162.15	20.34
Freshwater inland marsh	5.14	0.64
Quarries and bedrock outcrop	4.86	0.61
Mixed forest mainly coniferous	134.49	16.87
Mixed forest mainly deciduous	82.90	10.40
Open fen	4.00	0.50
Pasture and abandoned fields	71.18	8.93
Sparse coniferous forest	36.42	4.57
Sparse deciduous forest	47.22	5.92
Treed bog	13.14	1.65
Water	108.33	13.59
Total	797.21	100

Table 2
Climate normals in RVCA

Month	Mean Monthly Temperature (°C)	Mean Monthly Precipitation (mm)	Mean Monthly Potential Evapotranspiration (mm)	Mean Monthly Actual Evapotranspiration (mm)
January	-10.2	72.8	0	0
February	-8.7	56.0	0	0
March	-2.6	71.3	0	0
April	5.3	70.1	28.4	28.4
May	12.6	76.3	80.3	80.3
June	17.5	79.5	114.5	114.5
July	20.1	79.9	133.5	130.2
August	18.8	81.6	115.3	110.3
September	14.0	88.6	72.8	72.8
October	7.6	76.7	34.7	34.7
November	1.1	79.3	4.2	4.2
December	-6.4	79.7	0	0
ANNUAL	5.8	912	584	575

The climate is the typical Canadian cold climate (Table 2), with below freezing temperature from December through March. January is the coldest month with an average temperature of -10 °C, although the temperature frequently dips down to -30 °C. Annual precipitation is on average 912 mm, which is relatively uniformly distributed over the year. Temperature rises above freezing in early April and causes spring freshet. During winter months, snow accumulates on the ground, sometimes reaching 60 mm in water equivalent. Major floods usually occur during spring freshet, with occasional minor flooding during the summer months. More information on the climatology and hydrology is available in RVCA (2002). Stream flow data of the Tay River at Perth (02LA024) is available since 1994; however data gaps were prevalent at the beginning and good quality year-round data is available only since 2003. Further downstream at Port Elmsley (02LA016), continuous flow data is available for a few years (1982-88). The estimated 1:100 year floods at Perth and Port Elmsley are 95 and 111 cms respectively.

3. Model setup

As shown on Fig. 2, the integrated model consists of the main hydrodynamic (HD) network along the Tay River, Jebbs and Grants creeks, and five large lakes (Bobs, Christie, Crossby, Littele Crossby and Otty). The prevalent (constant) water level of Rideau River Lake was taken as the downstream boundary condition. The inflows from sub-basins, computed by the rainfall-runoff module (RR or NAM), constitute upstream boundary conditions.

For the 2007B model, the Tay River watershed was divided into 13 basins; all of them were separately modeled by the RR/NAM module and were connected to the HD network in a suitable fashion – either as point inflow or as distributed lateral inflow over a stream length. The NAM parameters (Table 3) were calculated in a two-stage calibration process (Ahmed 2010).

Table 3
NAM parameters for drainage units within Tay River watershed

Catchment Name	Area sq-km	Umax mm	Lmax mm	CQOF (-)	CKIF hr	CK1,2 hr	TOF (-)	TIF (-)	TG (-)	CKBF hr
OTTY	52.1572	34.5	184	0.172	931.3	10.5	0.990	0.00145	0.229	501
BLUEBERRY	44.4607	34.4	209	0.203	898.3	10.1	0.989	0.00053	0.316	500
TAY B	58.0986	34.3	232	0.158	755.7	10.8	0.990	0.00026	0.406	502
TAY A	12.9046	34.4	237	0.265	747.9	10.0	0.990	0.00235	0.428	505
RUDSDALE	63.2647	34.4	213	0.439	939.8	10.1	0.990	0.00226	0.332	500
TAY C †	55.1307	33.7	188	0.148	912.1	11.1	0.978	0.00400	0.306	499
CHRISTIE	32.1770	30.2	191	0.274	991.8	10.0	0.990	0.00008	0.237	498
PIKE & CROSBY	62.7641	20.2	197	0.293	886.9	10.1	0.989	0.00010	0.159	495
GRANTS CREEK	30.6967	34.4	229	0.278	785.5	10.2	0.989	0.00218	0.399	501
LEFE4SUBS ††	173.7360	25.1	209	0.492	861.1	10.3	0.990	0.00038	0.229	502
TAY D	29.5800	34.4	227	0.227	912.5	10.2	0.989	0.00050	0.374	503
BOBS LAKE	132.5220	22.1	182	0.107	756.2	10.5	0.988	0.00122	0.428	504
CROW LAKE	49.7206	34.3	208	0.180	864.6	11.9	0.989	0.00289	0.268	499

† Tay C was split into two for the 2008B model.
†† This area was split into six basins for the 2008B model.
$U_{\max}$ maximum water content in surface storage, $L_{\max}$ maximum water content in root zone storage, $CQOF$ overland flow runoff coefficient, $CKIF$ time constant for interflows, $CK_{1,2}$ time constants for overland flow routing, TOF root zone threshold value for overland flow, TIF root zone threshold value for inter flow, TG root zone threshold value for ground water recharge, $CKBF$ time constant for routing base flow; see DHI (2003, 2004) for details.

In this process, after the appropriate values of the NAM parameters were determined through auto-calibration, separate NAM models for all 13 basins were set up using these values, with minor adjustments to account for local conditions. These NAM models were then connected to the HD model setup for the river network. Cross-sectional data for the HD model was taken from previous studies and was also generated from available digital elevation models. In total about 110 km of river reach, five lakes, 133 cross-sections, 11 bridges, one weir and two dams were included in the 2007B model (Table 4).

Both dams have movable stop-log sections, which are manipulated to control the flow and water level at the dams. The dams are operated by the Rideau Canal office of Parks Canada (PC) to optimize various demands on the water such as the maintenance of lake water levels, instream flow needs and flood control. Historical data related to dam operation was obtained from PC and used in the model.

For the 2008B model, the Tay River watershed was divided into 19 basins. In total about 110 km of river reach, five lakes, 136 cross-sections, 25 bridges, two culverts, two weirs and two dams were included in the model (Table 4). Moreover, 25 cross-sections were resurveyed. The main improvement of 2008B model therefore lies in the better definition of cross-sections and the addition of 14 bridges in the model. The hydrodynamic (HD) module was thus significantly improved.

Table 4
Salient features of the models

	2007B model	2008B model
Number of drainage basins	13	19
River length (km)	110	110
Number of lakes	5	5
Number of cross-sections	133	136
Number of bridges	11	25
Number of culverts	--	2
Numbers of weirs	1	2
Number of dams	2	2

4. Methodology

Calibration of the rainfall-runoff model (RR or NAM) was based on the entire drainage basin contributing to the gage at Port Elmsley (02LA016), and was done by adjusting nine NAM parameters and by finding appropriate initial conditions parameters. The parameter values found this way were assigned to all basins within the Tay River subwatershed, and fine tuned as warranted by local conditions. As described by Ahmed (2010), the NAM auto-calibration was done prior to connecting them to the HD model. The auto-calibration was done to optimize two objective functions: (a) agreement between the average simulated and observed runoff, by minimizing the water balance error ($\%WBL$); and (b) overall agreement of the shape of the hydrograph, by minimizing the Root Mean Square Error ($RMSE$) for high flows (Madsen 2000, 2003). When solving multi-objective calibration, the problem is usually transformed into a single-objective optimization problem by defining a scalar that aggregates the various objective functions. The NAM autocalibration is implemented by giving all objectives equal weightage and by searching the solution by the shuffled complex evolution algorithm (Madsen 2000, 2003). Therefore, the present calibration is expected to give good results for runoff volume and for the high end of the flow spectrum. The auto-calibration (Fig. 3) yielded $\%WBL$ and R^2 values of -1.8% and 0.651 respectively.

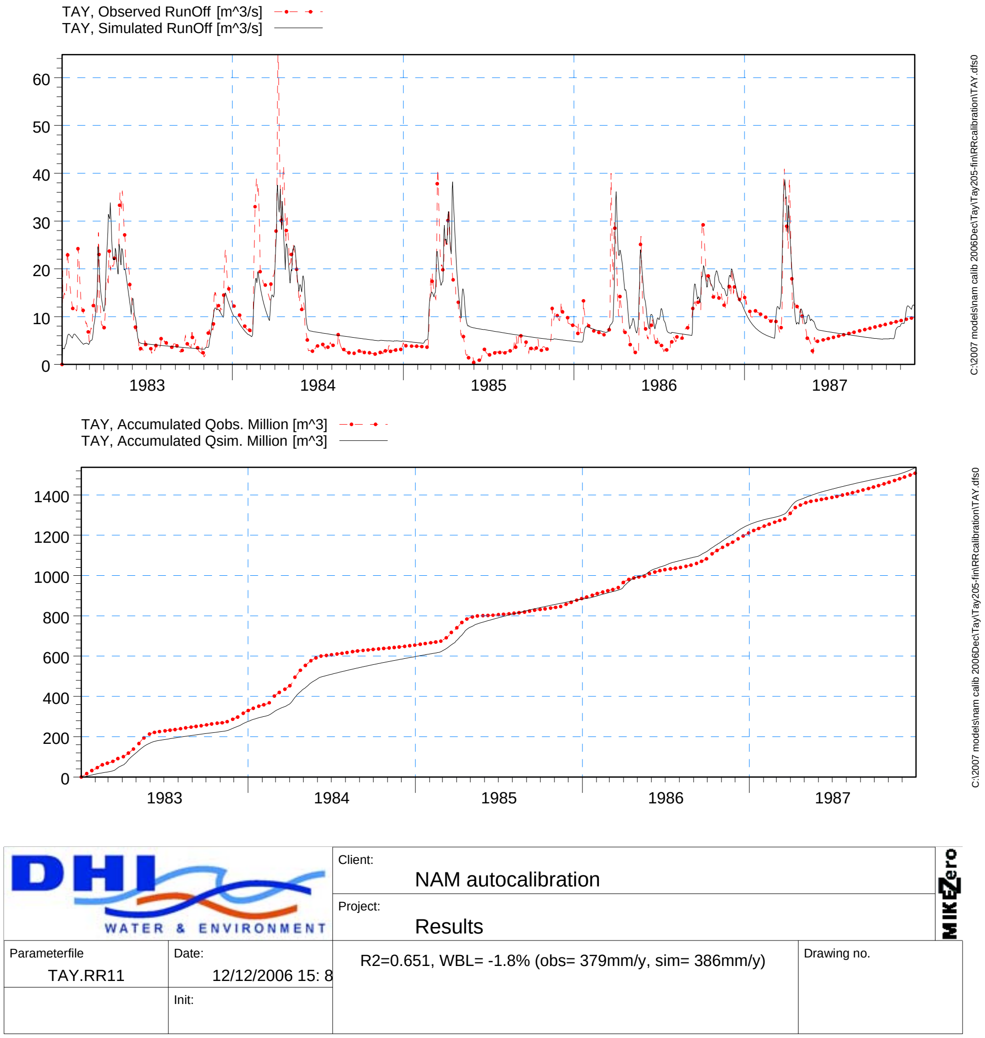


Fig. 3. NAM autocalibration at Port Elmsley gage (02LA016)

After the NAM calibration was done, the 13 RR modules were connected to the HD network. Calibration of the HD model was then done for a four year period from January 1st, 1983 through December 31st, 1986, using measured flow data at Port Elmsley (02LA016). The model was actually run from January 1st, 1982, but the first year was not used in calibration to minimize the effects of initial conditions.

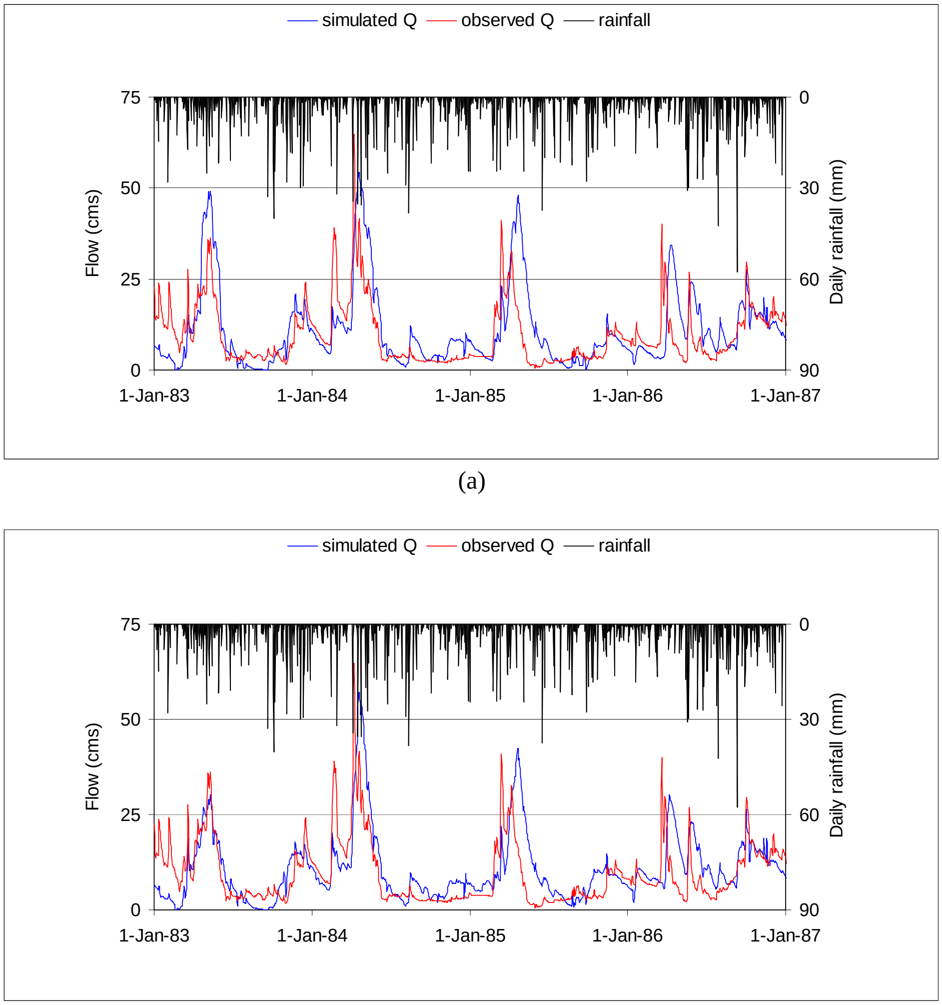
The HD calibration was done by adjusting the roughness coefficient (Manning's n) of the channels and floodplains. Typically the Manning's n was found to be around 0.035 for the main channel and ranged from 0.05 to 0.08 for the floodplain areas depending on the land use and vegetation.

The rainfall and temperature data from Ottawa Airport (ID 6105976), some 80 km from the Tay basin was used in this study.

For the 2008B model, 19 RR modules were used and new and improved surveyed cross-sections and additional bridges were incorporated (Table 4). The methodology remained the same.

5. Calibration

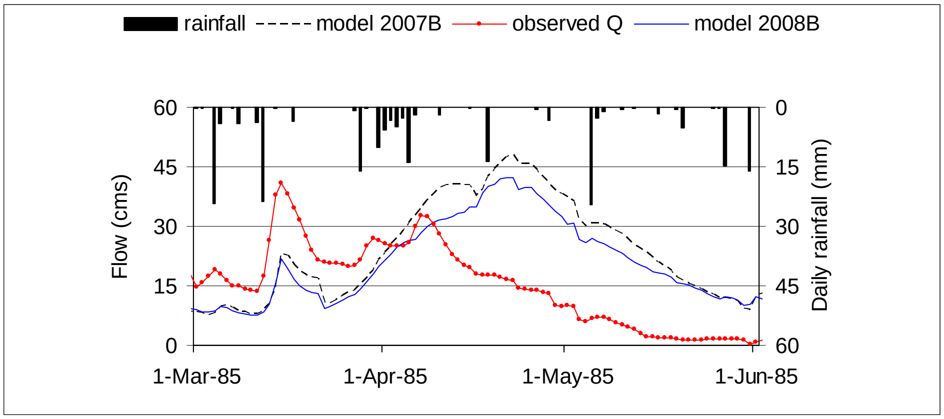
As in most modeling exercises, the main objective here was to predict stream flow as accurately as possible. Figures 4a, b show the overall simulated and measured flow data at the Port Elmsley gage location. The overall simulation during the entire four years of calibration period indicates the model's ability to capture the major seasonal variations in the hydrograph and to simulate the watershed response to spring freshet and intense summer storms. However, some phase lag, perhaps due to using a distant rainfall station and mismatch during low flows are also apparent. The improvement from model 2007B to 2008B is not readily discernable from these figures, except for a couple of instances like 1983 freshet or 1986-87 winter.



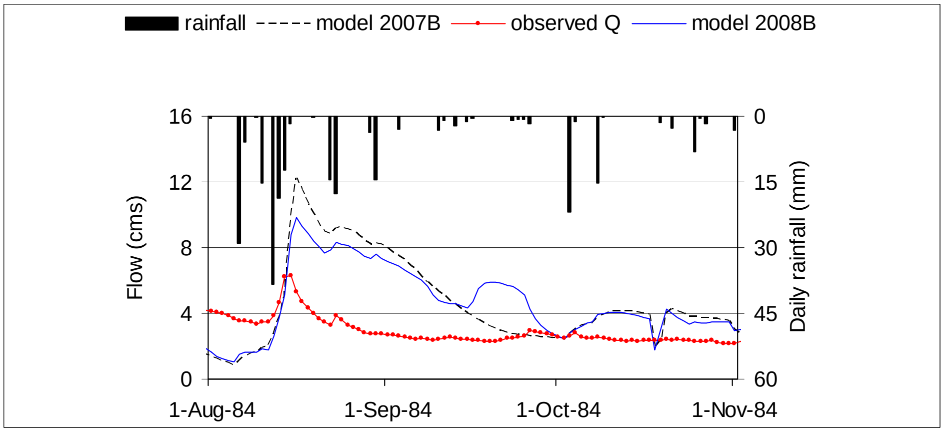
(a)

(b)

Fig. 4. Overall hydrograph during calibration – (a) 2007B model; (b) 2008B model



(a)



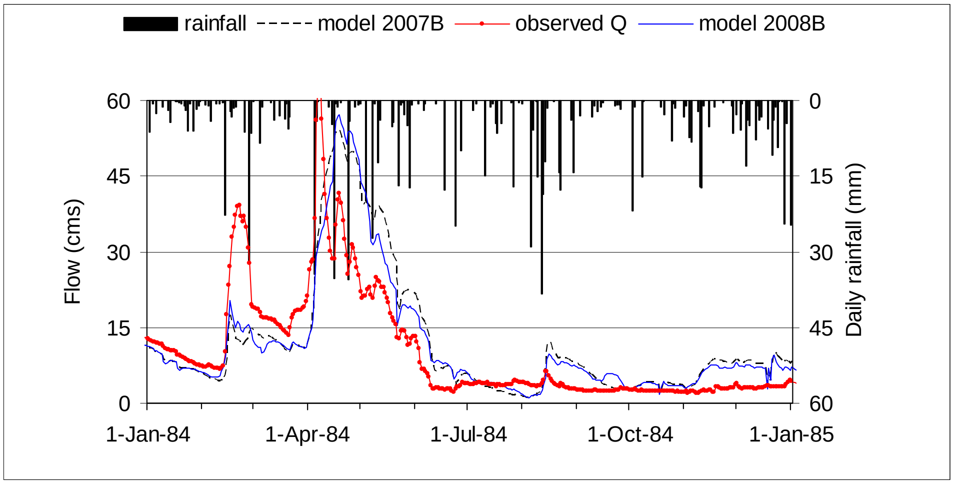
(b)

Fig. 5. Daily flows at Port Elmsley during calibration – (a) 1985 spring freshet; (b) 1984 summer months

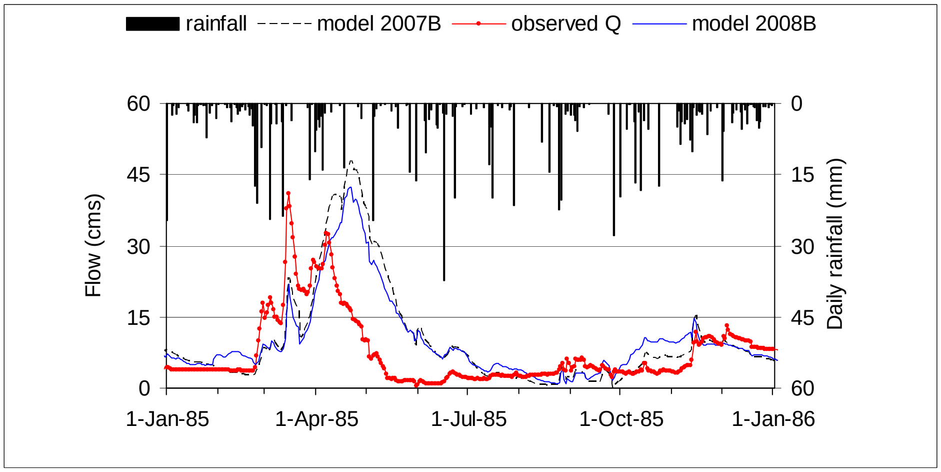
When we look at the results at a scale that allows day to day comparison of flow during individual storm or snowmelt events (Fig. 5a, b), the model behavior can be analyzed in a different way revealing a different set of model characteristics. We tried in vein to improve the obvious mismatch of modeled and observed flow observable at this scale. Failing that, we tried to find the reason for this problem. This kind of discrepancy cannot be explained by a distant rainfall station or an inaccurate representation of baseflow. A strong possibility is inaccurate record keeping of dam operation; however, there is no way to be sure. But during the course of modeling various watersheds within RVCA, we have come across occasional instances of inconsistent record of water control structures; when compared to other data sets used in this model, this appears to be most prone to errors.

Another challenge was the difficulty in calibrating the model to simulate well both the high and low flows. Our experience indicates that low flow calibration warrants higher structural consistency in the model, due mainly to the fact that low flow response is equally dependent on surface and subsurface flow components. High flow, on the other hand, is predominantly dependent on the surface runoff component, and therefore requires only a good capture of surface flow phenomenon. Efforts were therefore given to calibrate the model for the high flows at this stage; a separate calibration for low flow will be done in future.

When the entire year is considered (Figure 6a, b), the seasonal trends are seen to be simulated pretty well, with some time lag during spring freshet caused perhaps by using a distant rainfall station, and possible inaccuracy of dam operation data. Phase lag aside, magnitude of simulated high flows is within 20% of observed values. For the low flow, it is within roughly 50%. Slight improvement from model 2007B to 2008B can be discerned in the annual hydrographs; however, it is also apparent that there are physical processes which are not accurately captured in this model.

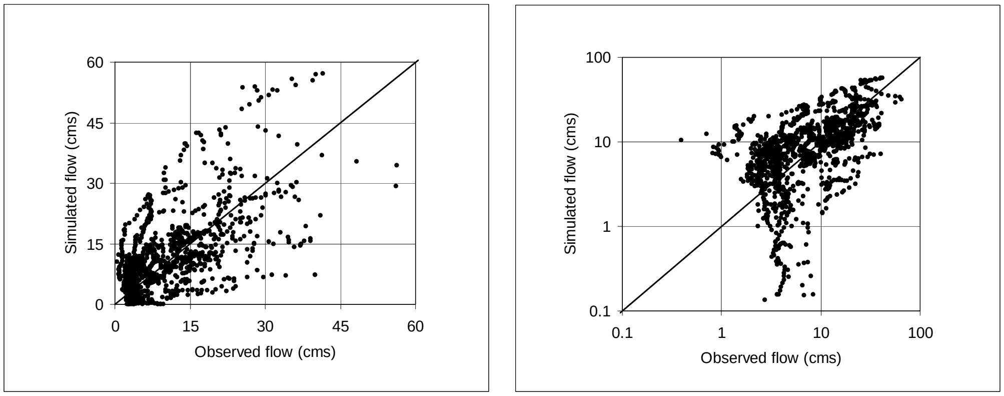


(a)



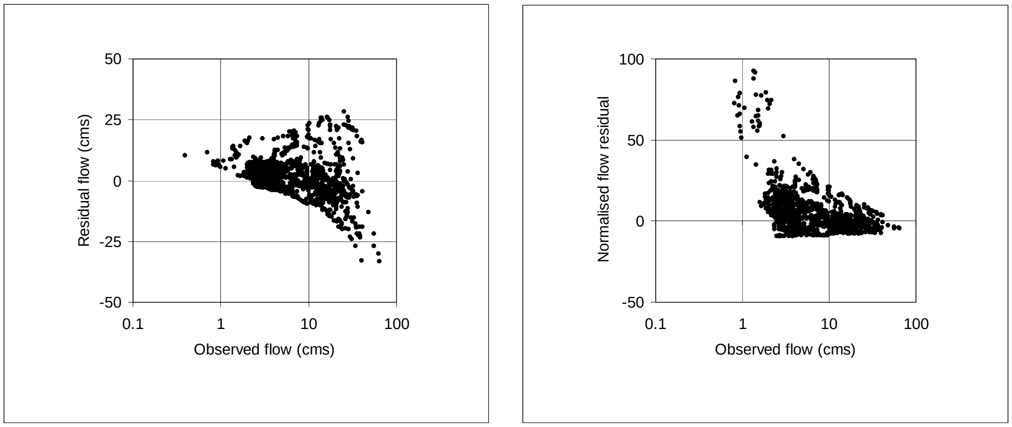
(b)

Fig. 6. Annual hydrograph comparison at Port Elmsley during calibration – (a) 1984; (b) 1985



(a) (b)

Fig. 7. Scatter plots of daily flows for 2008B model during calibration – (a) normal scale; (b) log scale



(a) (b)

Fig. 8. Residual plots of daily flows for 2008B model during calibration – (a) absolute residuals; (b) relative residuals

Scatter plots of daily flows show considerable scatter (Fig. 7a, b), especially for flows less than 10 cms. Residuals plots, as expected, show higher absolute deviation for high flows, but lower relative deviation (Fig. 8a, b), again indicating poor representation of lower flows.

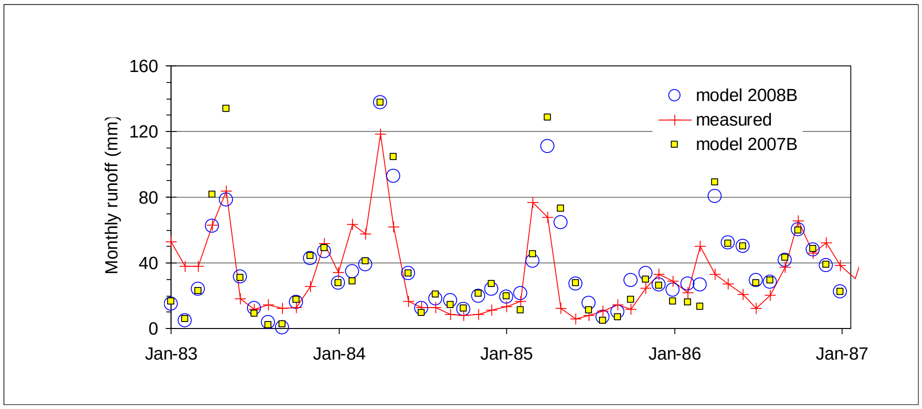


Fig. 9. Monthly runoff volume during calibration

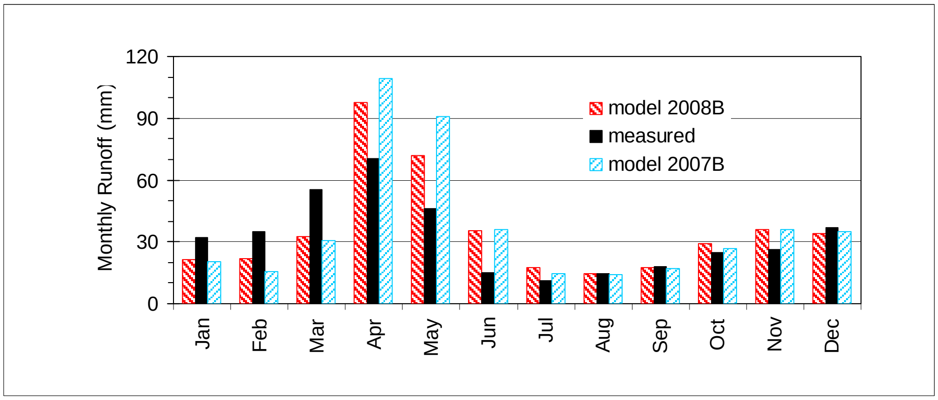


Fig. 10. Four-year average of monthly runoff volume during calibration

During calibration, the monthly flows were looked at from various angels: a simple time series of monthly total runoffs (Fig. 9), bar graph of the 4-year average of monthly volumes (Fig. 10), and a flow duration curve based on from monthly runoff volumes (Fig. 11). The time series of the monthly runoff volumes, calculated from the simulated daily flows, captures the seasonal trend reasonably well (Fig. 9). The snowmelt volume during April/May was consistently overestimated (Fig. 10). The prediction during the summer and autumn months with lower flows was good; however the winter month flows were somewhat underestimated. The monthly flow duration curves (Fig. 11) indicate better matching for low-yielding months (< 50mm) than high yielding months, an inference opposite to what we formed in case of the daily flows. Also noticeable is the slight improvement of model 2008B compared to 2007B for low-yielding months, but substantial improvement during high-yield months (> 50 mm).

This observation is important, especially in light of the fact that monthly runoff volumes are indicative of the slower hydrological processes such as interflow and baseflow and are useful for specific purposes (e.g., drinking water availability and water quality), as opposed to daily flows that reflect the immediate response of the basin and is useful for other purposes (e.g., flood risk mapping or flood forecasting). The present model appears to produce accurate water availability during drier months, which are critical in terms of water availability and water quality.

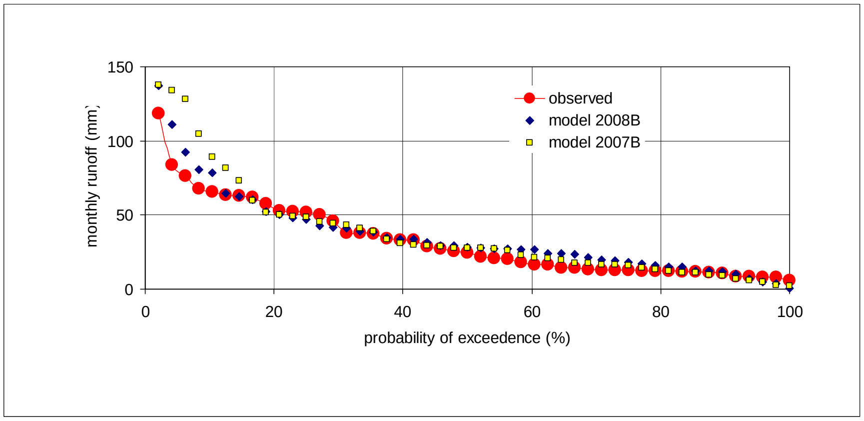


Fig. 11. Flow duration curve based on monthly runoff during calibration

The ratio of the simulated to measured daily flow ranges mainly from 0.1 to 10, while the flow varies by two orders of magnitude (Figure 12). This plot also shows the general tendency of the model to overestimate low flows and underestimate high flows. These observations are in general congruent with similar plots reported by Singh et al. (2005) for the Inroquois River in Illinois, although in our case the scatter band is rather wide. Similar plots for the monthly runoff (Figure 13) show a narrower band of the simulated to observed ratio (0.1 to 6) while the runoff varies by only one order of magnitude (8 to 80 mm). The spread of the ratio is rather similar for all runoff values for the Tay river, as opposed the findings of Singh et al. (2005). In their case, the spread was narrower for higher runoffs and higher for lower runoffs; e.i., the high runoffs were predicted more accurately. In the case of Tay River, the control by reservoirs and dams has narrowed the difference between high and low flows, thus eliminating this distinction.

Compared to other basins under RVCA jurisdiction where calibration and validation were done for 5 years each, the limited data at Tay basin allowed only a 4-year calibration and no validation at all. Currently, efforts to collect more data are underway, and we hope to re-calibrate the model in future with additional data.

6. Quantitative model assessment

As usual, both graphical and numerical techniques were used to evaluate the model performance. Although the graphical plots described above are excellent in visualizing and interpreting overall model results and can guide the calibration/validation process, they are nonetheless qualitative assessment and are thus somewhat subjective. Therefore, for more rigorous and objective model testing, several quantitative numerical criteria have been used. Some of them are based on elementary statistical principles that are used for establishing relationship between two time series, and the others have been devised in relation to hydrologic modeling, and thus have more hydrologic meaning. Since they are based on numerical manipulation of data in a prescribed fashion, they are considered to be objective and unbiased indicators of model performance. Ahmed (2010) compiled a list of performance indicators widely used in hydrologic modeling. Some of them, listed below, are used in this study.

1. Mean Absolute Error (MAE)
2. Percentage Bias ($\%BIAS$) or water balance error ($\%WBL$)
3. Root Mean Square Error ($RMSE$)
4. Relative Root Mean Square Error ($RRMSE$)
5. Coefficient of Determination (R^2)
6. Pearson Correlation Coefficient (PCC)
7. Nash-Sutcliffe Coefficient of Efficiency (NSE)
8. Flow Duration Error Index (EI)
9. Index of Agreement (IA)
10. Ratio of $RMSE$ to the standard deviation of measured data (RSR)

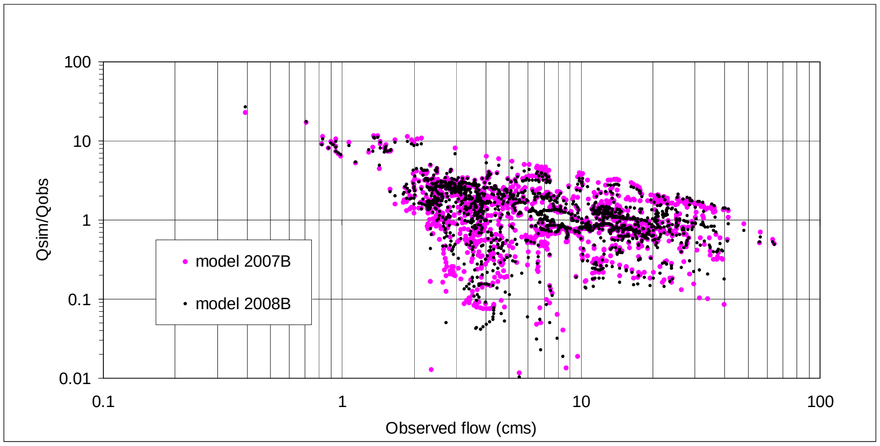


Fig. 12 Ratio of simulated to observed daily flows during calibration (1983-1986)

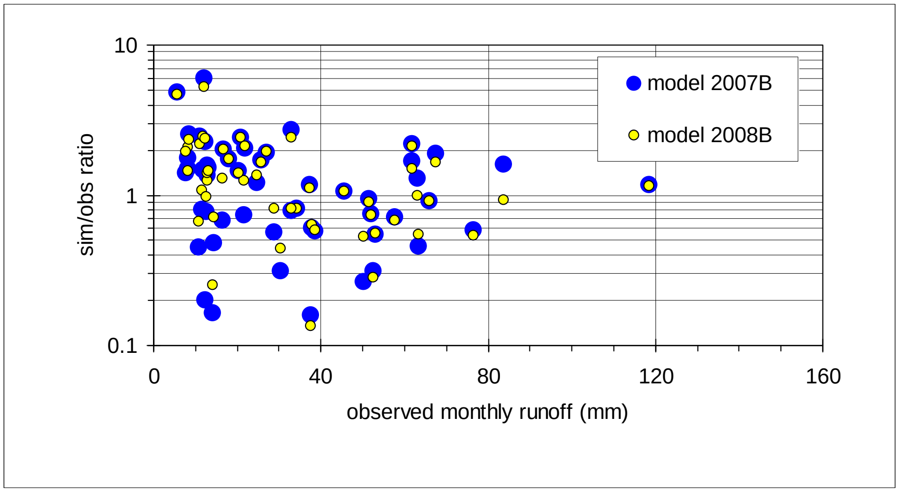


Fig. 13 Ratio of simulated to observed monthly runoffs during calibration (1983-1986)

The last indicator (RSR) was recently developed by Moriasi et al. (2007), and ranges from zero upwards, with lower values indicating better model performance. All the above indicators were calculated for the Tay River models (2007B and 2008B) and are presented in Table 5 for the calibration period. The wide range of indicators is expected to enable us to assess the model performance from different vantage points, and would, it is hoped, provide a more comprehensive evaluation of the models.

It is noted that the performance of the Tay models (both 2007B and 2008B) are below average compared to other models within the Rideau Watershed (RVCA 2007, Ahmed 2010); certainly, better performance was achieved elsewhere using similar methodology and expending comparable effort. Even after the update with recently surveyed cross-sectional and bridge/culvert data, the model performance (2008B) has improved only marginally. In integrated watershed models with many different constitutive elements (rivers, lakes, drainage basins, water control structures, etc.), all play important but separate roles and contribute differently towards model performance; hence it is important that all are given due consideration during model building and calibration.

Table 5
Model performance during calibration

Gage location	Performance indicator	Model 2007B Calibration (1 Jan 1983 to 31 Dec 1986)				Model 2008B Calibration (1 Jan 1983 to 31 Dec 1986)			
		Daily Flow	unit	Monthly Runoff	unit	Daily Flow	unit	Monthly Runoff	unit
Port Elmsley (02LA016)	MAE	5.863	cms	17.508	mm	5.239	cms	14.881	mm
	%BIAS	15.370	[-]	15.394	[-]	11.306	[-]	11.326	[-]
	RMSE	8.517	cms	23.564	mm	7.442	cms	19.670	mm
	RRMSE	0.885	[-]	0.732	[-]	0.773	[-]	0.611	[-]
	NSE	-0.005	[-]	0.081	[-]	0.233	[-]	0.360	[-]
	R^2	0.425	[-]	0.523	[-]	0.420	[-]	0.531	[-]
	PCC	0.652	[-]	0.723	[-]	0.648	[-]	0.728	[-]
	IA	0.780	[-]	0.821	[-]	0.793	[-]	0.844	[-]
	EI	0.853	[-]	0.824	[-]	0.856	[-]	0.824	[-]
	RSR	1.003	[-]	0.959	[-]	0.876	[-]	0.800	[-]

MAE mean absolute error, $\%BIAS$ percentage bias ($\%BIAS$) or water balance error ($\%WBL$), $RMSE$ root mean square error, $RRMSE$ relative root mean square error, R^2 coefficient of determination, PCC Pearson correlation coefficient, NSE Nash-Sutcliffe coefficient of efficiency, EI flow duration error index, IA index of agreement; see Ahmed (2010) for definitions and equations.; RSR RMSE-observation standard deviation ratio as defined by Moriasi et al. (2007)

During the combined RR/HD calibration process (model 2008B), a $\%BIAS$ of about 11.3% was obtained, which is higher than the 1.8% obtained during the NAM auto-calibration process (note that Mike11 autocalibration module computes $\%WBL$ as $\%BIAS$, but with a negative sign). This implies no improvement (rather a deterioration) of annual water budget at the gage location as a result of incorporating the HD module, which requires much effort to implement. It also shows the relative ease with which a rainfall-runoff model (in this case NAM) can be autocalibrated, which is largely a data fitting exercise. Similar conclusion can also be drawn from the value of R^2, which was higher during NAM autocalibration (0.651) than during RR/HD calibration (0.420). This means that it is easier to calibrate a lumped model than an integrated model which consists of various elements and thus various hydrological processes. What is, then, the purpose of expending substantially more effort if the result cannot be improved? The answer is that it is not enough for the model to perform well; it must perform well for the right reason. Besides, there is the more utilitarian reason of being able to simulate flow and water level at various locations along the streams.

Relatively lower values of NSE and R^2 (Table 5) imply a rather modest model performance. However, the moderately high values of other statistical measures, including PCC, EI and IA, indicate some promise assessed from different vantage points. This demonstrates the advantage of using difference indicators. The indicators collectively point to a slightly better performance of 2008B than 2007B, implying the marginal improvement resulting from improved cross-sectional and structures data.

In terms of some parameters (NSE, PCC R^2 and IA), the model performed better on the basis of monthly runoff rather than the daily flow. This implies better correlation between observed and predicted values for longer-duration quantities that damp out short-duration fluctuations. The same was found for other watersheds in the region (RVCA 2007).

A number of observations can be made on the performance indicators (Table 5). For example, the statistics indicate that the model predicts higher flows. A MAE value in the order of 5.2 cms indicates that the flow is overestimated on average by 5.2 cms, which is also reflected in a $\%WBL$ value of 11.3%. $RMSE$ and $RRMSE$ also point to the same direction.

Some authors have suggested setting *a priori* performance targets based on anticipated purpose of the model that can be used to accept or reject models. This has not been done for the models built by RVCA, since the aim is to build models in as detailed a fashion as permitted by available data and to use them for various watershed management purposes as opportunities arise. However, it is useful to compare the model performance against what others have considered acceptable standard.

Lorup et al. (1988) consider a model valid only when three specific criteria were met: $\%WBL < 10\%$; $NSE > 0.80$; and $EI > 0.70$. Our model meets the third criterion and would therefore not be considered acceptable according to this test. According to the performance rating used by Hendriksen et al. (2003), our model can be rated poor (so defined when NSE = 0.2 to 0.5, and $\%WBL$ = 10% to 20%). Moriasi et al. (2007) consider a model unacceptable when $NSE < 0.50$, and $RSR > 0.70$; our model would thus be considered unacceptable. Overall, the Tay model is yet to reach a satisfactory level of performance by industry standard. Considering the limitations of data, this is rather a foregone conclusion; however, we present this challenging modeling exercise in the spirit of learning from challenges we face in the real world. Moreover, most of the performance thresholds have not been related to intended use, thus missing the all-important dependence of performance to utility.

7. What now?

While it is clear that the Tay model performs unsatisfactorily according to industry standard, its reasons are by no means clear. Lack of streamflow data to do an adequate validation and the uncertainty regarding the dam operation data have been mentioned already. Another issue is that the Tay basin, as opposed to rest of the Rideau watershed, is part of the Canadian Shield with fractured bed rock and shallow overburden. The characteristics of subsurface flow and how it affects the surface flow are largely unknown at this time, although some research is ongoing (by others). The presence of a large number of lakes and their outlet control is another issue affecting both the water balance (through evaporation) and flood hydraulics. Work is underway to incorporate more lakes into the hydrodynamic network; however, appropriate care has to be taken to model outlet controls properly.

While it is easy (relatively speaking – actually to do it right is not so easy) to figure out what needs to be done to improve the model, to make it happen requires more than scientific acumen. In this specific case, we expect that better data and additional information will be in place for building a model of acceptable performance. However, sustained commitment by and collaboration among several agencies is necessary to achieve this end. This will be no small feat considering the present-day uncertainties associated with shifting agency priorities, budgetary constraints and short attention span. That, however, is another story; suffice it to say here that there are important societal and institutional factors that contribute to the success of modeling projects.

8. Closure

A numerical model of the Tay River watershed, which consists of a number of drainage basins, rivers, wetlands, and actively operated dams, has been constructed using the Mike11 software. The model was calibrated against four years of stream flow data; lack of data did not allow validation. A wide range of performance indicators were used to assess the model. When compared to published literature, the model performed moderately well, which can be attributed to a number of factors (distant rainfall gage, inadequate representation of lakes, and probable inaccuracy of dam operation record). Efforts are currently underway to collect more data which would allow a more satisfactory calibration and validation of the model.

Acknowledgement

This paper is based on the initial work done at RVCA; however, the author has substantially extended it. Bruce Reid, Director of Watershed Science and Engineering Services, provided encouragement and advice. S. Khan and N. Howlader did much of the model building and data analysis. All these are gratefully acknowledged. Notwithstanding the contribution of others in the original study, the opinions presented in this paper are based on further analyses conducted by and are the sole responsibility of the author, and do not necessarily represent the official position of RVCA.

References

Ahmed F (2010) Numerical Modeling of the Rideau Valley Watershed. Natural Hazards 55:63-84

Andréassian V et al (2010) The Court of Miracles in Hydrology: can failure stories contribute to hydrological science? Hydrological Sciences Journal 55:849-856

DHI (2003) Mike 11 – A Modeling System for Rivers and Channels, Reference Manual, Danish Hydraulic Institute, Hørsholm, Denmark

DHI (2004) NAM – Technical Reference and Model Documentation Draft, Danish Hydraulic Institute, Hørsholm, Denmark

Guzha AC and Hardy TB (2010) Application of a Distributed Hydrological Model, TOPNET, to the Big Darby Creek Watershed, Ohio, USA. Water Resources Management 24:979-1003

Henriksen HJ et al (2003) Mehtodology for construction, calibration and validation of a national hydrological model for Denmark. Journal of Hydrology 280:52-71

Lorup JK et al (1998) Assessing the effect of land use change on catchment runoff by combined use of statistical tests and hydrological modelling: Case studies from Zimbabwe. Journal of Hydrology 205:47-163

Madsen H (2000) Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. Journal of Hydrology 235:276-288

Madsen H (2003) Parameter estimation in distributed hydrological catchment modeling using automatic calibration with multiple objectives. Advances in Water Resources 26:205-216

Moriasi DN et al (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the American Society of Agricultural and Biological Engineers 50(3):885-900

RVCA (2002) Tay River watershed management plan. Rideau Valley Conservation Authority, Manotick, Canada

RVCA (2007) Rideau River Watershed Modeling Using Mike11, Draft Report. Rideau Valley Conservation Authority, Manotick, Canada

Singh J et al (2005) Hydrological modeling of the Iroqouis river watershed using HSPF and SWAT. Journal of the American Water Resources Association 41:343-360