Potential improvement of Don River PCSWMM model with Hydrologic Data Assimilation

Tara Razavi, Ph.D. ; James Leach, M.Sc.
NSERC Canadian FloodNet , Water Resources and Hydrological Modeling Lab, McMaster University

Introduction

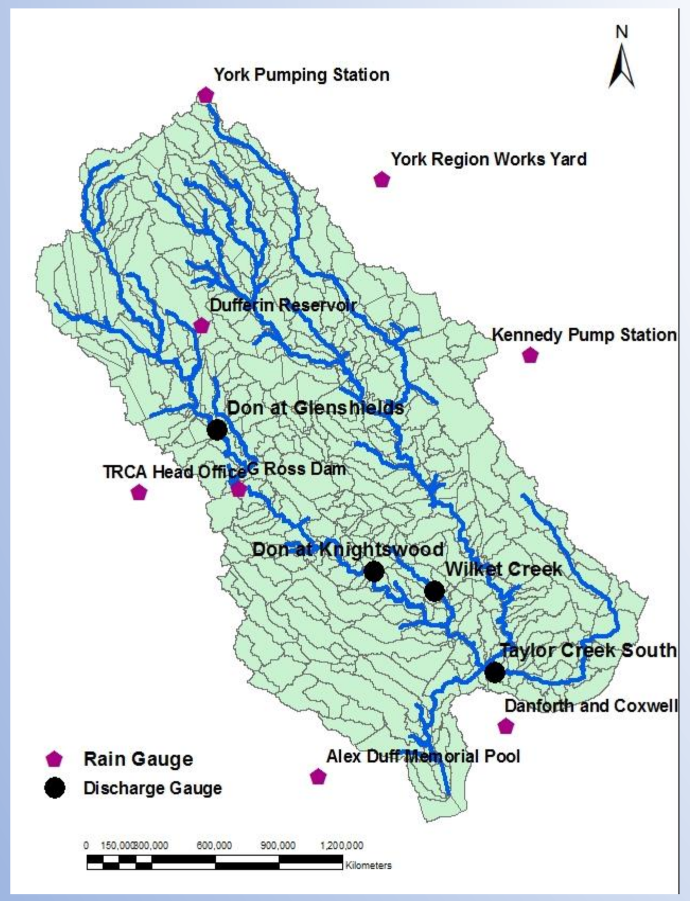
Data assimilation (DA) optimally merges information from model simulations and observations with appropriate uncertainty modeling. It accounts for various types of uncertainty such as errors in model state, variables, and hydrometeorologic forcing data. Numerous research studies have shown that DA can be used to improve prediction accuracy while also quantifying uncertainty. The application of DA with a distributed hydrologic model with the purpose of improving flood forecasting and shortening lead times, as well as application in improved operational forecast systems have been challenging in this context [1].

Objectives

- Evaluating Don River watershed PCSWMM model with the current calibration.
- Evaluating the potential of hydrologic data assimilation in PCSWMM model improvement.
- Analyzing further improvement of Don River PCSWMM model in predicting peak flows with hydrologic data assimilation.

Case Study and Data

Don River watershed covers municipalities of Toronto, York, Markham , Richmond Hill, and Vaughan. The area of the watershed is 358 km2 and length of major tributaries are 9–43 Km with monthly mean streamflow of 4 m3/s. Don river watershed land use is 96% urban with 8% forest , 6% meadow , 1% successional and 0% wetland [2].



Available observed precipitation and discharge data form TRCA

Station ID	LocationName	Monitoring Network	Region	Municipality	Elevation_masl	Length of data
HY003	Alex Duff Memorial Pool	Precipitation	Toronto	Toronto	109	4/8/2008_12/31/2013
HY016	Danforth and Coxwell	Precipitation	Toronto	Toronto	115	5/22/2008-12/6/2013
HY021	Dufferin Reservoir	Precipitation	York	Vaughan	226	4/8/2008_12/31/2013
HY027	G Ross Dam	Precipitation	Toronto	Toronto	175	4/8/2008_12/31/2013
HY036	Kennedy Pump Station	Precipitation	Toronto	Toronto	190	5/8/2008-12/5/2013
HY064	TRCA Head Office	Precipitation	Toronto	Toronto	188	4/11/2008_12/6/2013
HY069	York Pumping Station	Precipitation	York	Richmond Hill	325	4/8/2008_12/5/2013
HY070	York Region Works Yard	Precipitation	York	Richmond Hill	217	4/8/2008_12/5/2013
HY017	Don at Glenshields	Stream	York	Vaughan	182	1/1/2008_12/31/2013
HY018	Don at Knightswood	Stream	Toronto	Toronto	121	1/1/2009_12/31/2013
HY062	Taylor Creek South	Stream	Toronto	Toronto	90	1/1/2008_12/31/2013
HY068	Wilket Creek	Stream	Toronto	Toronto	122	1/1/2010_12/31/2013

Methods

a) PCSWMM model

PCSWMM model has the same hydrologic and hydraulics engine as the U.S. EPA SWMM5 model. It consists of 1D & 2D analysis comprehensive river modeling tools and real-time control analysis time series management. It has the capability to perform hydrologic, hydraulic and water quality modeling.

b) Hydrologic data assimilation : Ensemble Kalman Filtering (EnKF)

Data assimilation is a process that quantifies errors in both the hydrological model and observations, and update hydrological model states in a way that optimally combines model forecast with observations. Each individual observation is then updated based on the relative error in both the model and observations. The model forecast is made in the EnKF for each ensemble member as follows [3]:

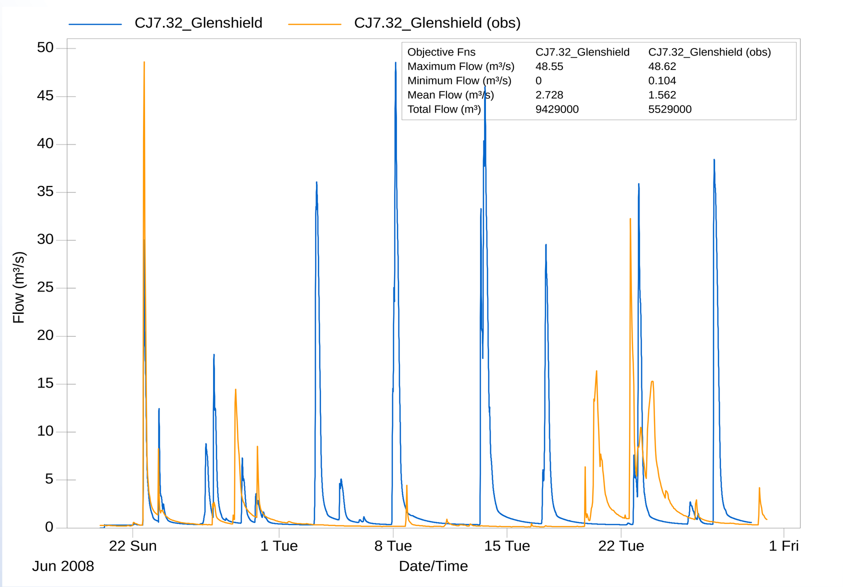
$x_{t+1}^{i-} = f(x_t^{i+}, u_t^i, \theta, t) + \omega_t^i$, where:

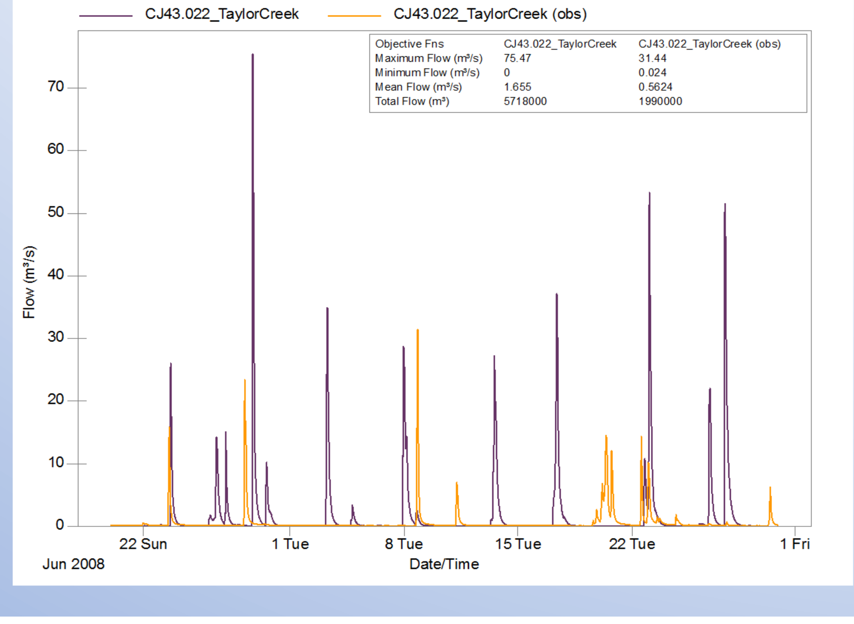
x_{t+1}^{i-} : The i_{th} ensemble member forecast of model state variable at time t+ 1.
x_t^{i+} : The i_{th} updated ensemble member at time t.
u_t^i: Perturbed forcing data.
θ: Model parameter.
ω_t^i : The model error that represents all the uncertainties related to model structure and the forcing Data.
$x_{t+1}^{i+} = x_{t+1}^{i-} + K_{t+1}(y_{t+1}^i - \hat{y}_{t+1}^i)$
x_{t+1}^{i+} : The i_{th} member of updated states at time t+1.
y_{t+1}^i : The i_{th} member of perturbed observation at time t+1.
$\hat{y}_{t+1}^i$: The ith predictive variable at time t+1.
K_{t+1} : Kalman gain matrix: $K_{t+1} = \Sigma_{t+1}^{xy-}[\Sigma_{t+1}^{yy} + \Sigma_{t+1}^{y}]^{-1}$
Σ_{t+1}^{yy} : The forecast error covariance matrix of the prediction $\hat{y}_{t+1}^i$
Σ_{t+1}^{xy-} :The forecast cross covariance of the state variables x_{t+1}^{i-} and prediction $\hat{y}_{t+1}^i$

Results

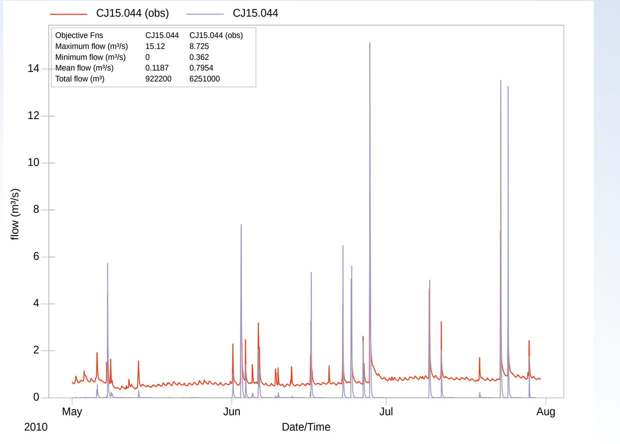
Performance of current calibrated PCSWMM model (last calibrated in March 2014 by TRCA)

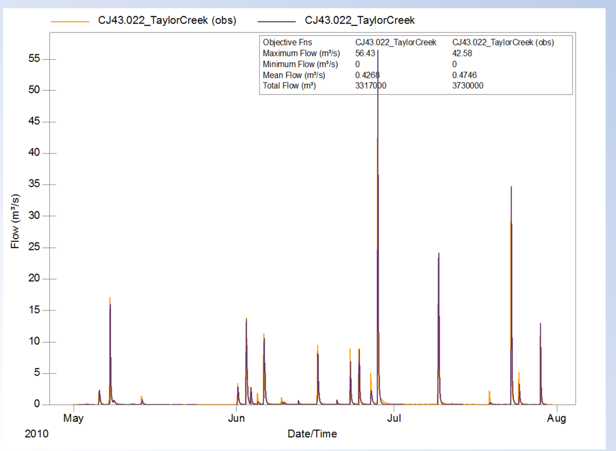
1- Calibration
June 20,2008-July 30,2008 (40 days)

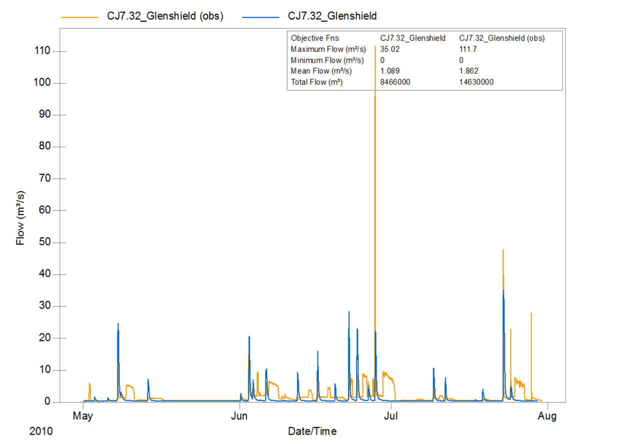


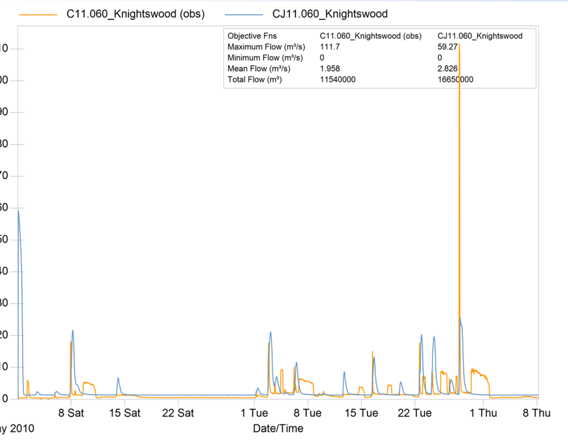


2- Validation: May 1st ,2010- July 30th,2010 (90 days)

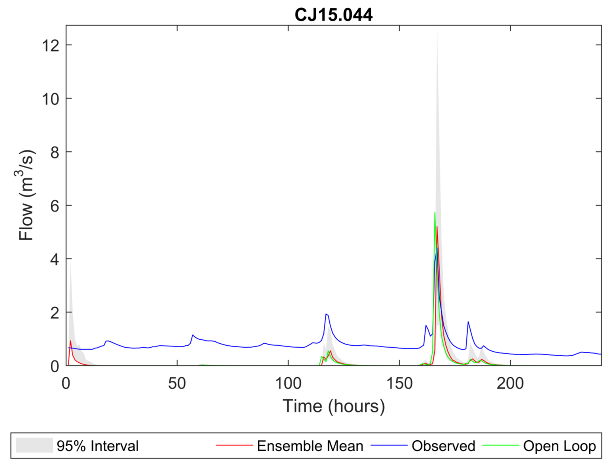


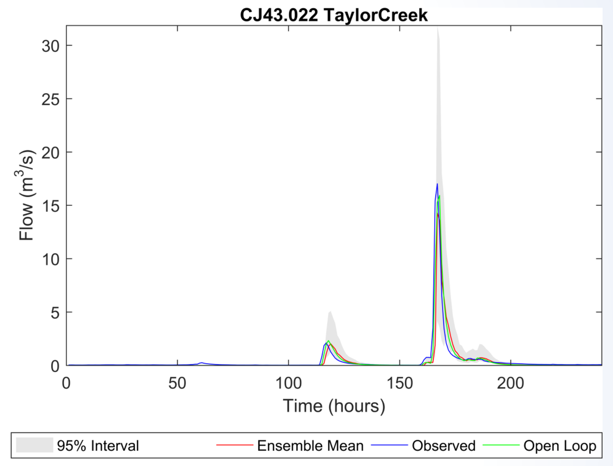


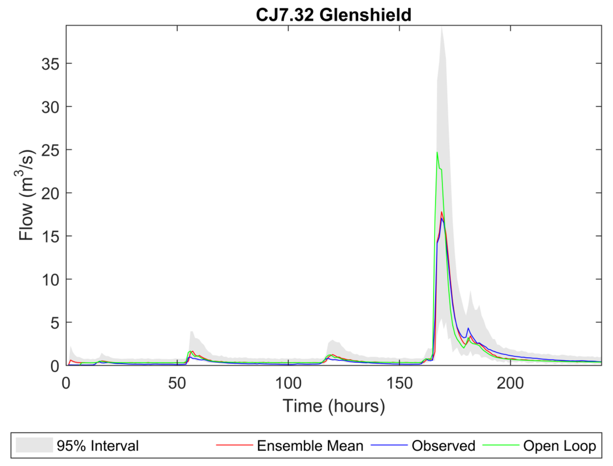


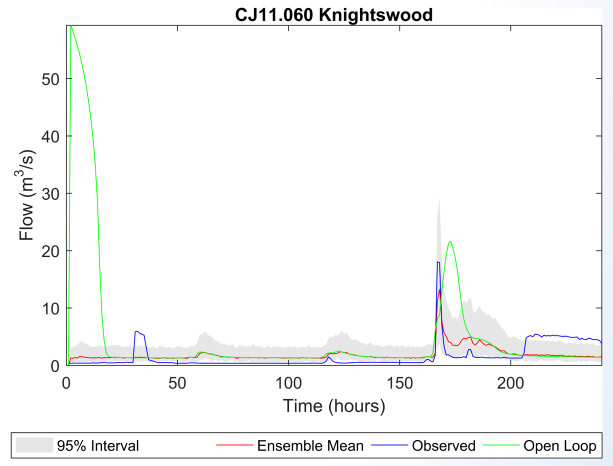


Data Assimilation results









Open Loop: No assimilation
Ensemble Mean: Average of 48 ensemble members

Updated state variables include:

Sub-catchment states: Current initial soil moisture deficit, current cumulative infiltrated volume, current upper zone infiltrated volume.
Link states: new flow, new depth

Model Performance Statistics (for 250 hours)

	HY018/CJ11.060 Knights wood		HY068/CJ15.044 Wiket Creek		HY062/CJ43.022 Taylor Creek		HY017/CJ7.32 Glenshield	
	Open Loop	DA	Open Loop	DA	Open Loop	DA	Open Loop	DA
NSE	-0.17	-1.00	-1.46	-1.80	0.88	0.56	0.84	0.98
RMSE	12.02	1.88	0.72	0.74	0.61	0.99	1.25	0.37
VE	2.25	0.20	-0.89	-0.88	-0.07	-0.15	0.12	0.05
BIAS	3.52	0.32	-0.67	-0.67	-0.03	-0.06	0.12	0.05

Discussion

The current PCSWMM model for the Don River watershed for calibration period (June 20th,2008 – July30th,2008) at the two streamflow monitoring locations of TRCA indicated that the model failed to estimate most of the peak flow values correctly. For the validation period (May 1st 2010 – July 30th , 2010) only at one stream gauge location (i.e. Taylor Creek) the model can capture most of the peak flows correctly in terms of time to peak and amount of peak flow. This evaluation indicates that the current calibrated PCSWMM model needs re-calibration against the four stream gauge locations of TRCA. Different calibration algorithms such as Particle Swarm Optimization can help to improve the current calibration.

Data assimilation results indicates that assimilating both the sub-catchment and link state variables of SWMM5 model has the potential to improve the model performance.
Further improvements can be achieved with a better calibrated model along with data assimilation.

Some challenges and Next steps

During this study, some challenges were encountered, such as:

- Large number of sub-catchments
- Few observational data
- Large computation time
- Poorly calibrated model

To overcome these challenges and improve the results, we are planning on doing the following:

- Re-calibrating the Don River PCSWMM/SWMM5 model
- Adapting EnKF to work better with distributed models:
  - Updating more sub-catchments and state variables
  - Trying different updating approaches including log transformations
- Using additional data assimilation methods
- Transitioning into real-time operational flood forecasting

References

[1] Liu, Yuqiong, et al. "Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities." *Hydrology and Earth System Sciences, 16 (10), 2012* (2012).
[2] Don River watershed, report card 2013, TRCA.
[3] Moradkhani, Hamid, et al. "Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter."*Water Resources Research* 41.5 (2005).

Acknowledgment : Thanks to Canadian Natural Sciences and Engineering Research Council (NSERC) for providing the funding and Toronto and Region Conservation Authority (TRCA) for providing the data required for this analysis.