Assessing reliability of 6-hourly and disaggregated 1-hourly Canadian Precipitation Analysis (CaPA) data

Pedram Darbandsari, Tara Razavi, Paulin Coulibaly

NSERC Canadian FloodNet , Water Resources and Hydrological Modeling Lab, McMaster University

Introduction

Canadian Precipitation Analysis (CaPa) system is designed by Environment Canada to combine various sources of precipitation data to produce a near real-time gridded accumulated 6-hourly precipitation estimates. In order to use the 6 -hourly CaPA data for flood forecasting in areas where rain gauges are sparse or have poor quality records, assessing the reliability of the original sub-daily CaPA data and producing reliable 1-hour disaggregated form of CaPa data are crucial.

Objectives

- ✓ Comparing original 6-hourly CaPA precipitation data with the available gauge data in the Don River and Humber River watersheds
- ✓ Applying a combination of two different bias correction techniques and cascade-based disaggregation technique to generate a more reliable and proper hourly continuous precipitation time series for the aim of flood forecasting.

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 Humber River watershed spans 903 Km2 from the headwaters on the Niagara Escarpment and Oak Ridges Moraine down through fertile clay plains to the marshes and river mouth on Lake Ontario.

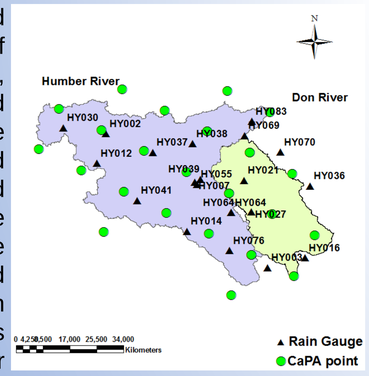

Figure 1. Location of the study areas, rain gauges and CaPA points

CaPA precipitation Data

Figure 1 shows the CaPA points located in and near Humber and Don river watersheds, the 6-hourly precipitation data for all of these points are available from 2002 to 2015 with some missing values in 2012.

TRCA rain gauges

Observed precipitation data set are available from 20 rain gauges which are located in the two watersheds. Hourly precipitation data with some missing values, especially in the winter, is available in these stations for various periods.

Methodology

Figure2 shows the bias-correction and dis-aggregation procedure used for assessing the reliability of CaPA data in the study regions. First, two different bias correction techniques are used to correct the 6-hourly and Daily CaPA data , then the cascade-based disaggregation technique is used to develop hourly precipitation data. Finally using the bias correction techniques, the difference between observed and the developed disaggregated hourly data is reduced. The comparison between CaPA data and the observed precipitation data of the nearest rain gauge are made in different steps of the proposed flow chart.

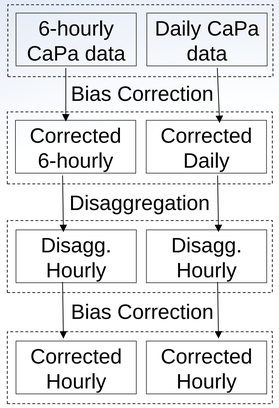

Figure 2. Flow chart of the methodology

Bias-Correction techniques:

In this study, two Bias correction methods are used. These methods are briefly explained:

a) Frequency and intensity correction [1]

In this method, first, the frequency of the CaPA rainfall is corrected by truncating its distribution based on the threshold value ($\tilde{x}_C$) being calculated by using the minimum observed rainfall amount, $\tilde{x}$ (Equation 1).

$$\tilde{x}_C = F_C^{-1}(F_{obs}(x)) \quad (1)$$

In the next step, the truncated CaPA intensity distribution (e.g. a fitted gamma or the empirical distribution of the truncated CaPA data) is mapped to a gamma distribution fitted to observed intensity distribution (Equation 2).

$$x_C^{Cor} = \begin{cases} F_{obs}^{-1}(F_C(x_i)) & x_i \geq \tilde{x}_C \\ 0 & x_i < \tilde{x}_C \end{cases} \quad (2)$$

This method is applied for each of the 12 calendar months separately.

b) Local intensity scaling (LOCI) [2]

The main steps of the LOCI method is correcting the wet time-step frequency and intensity, respectively, therefore, first the threshold for the month m of CaPA data ($\tilde{x}_{C,m}$) is determined to ensure that the threshold exceedance matches the wet-day frequency of the observation and then the following equation is used for correcting the intensity of CaPA data in each month:

$$x_{C,m,i}^{Cor} = \begin{cases} x_{C,m,i} \times \dfrac{\mu(x_{m,i} | x_{m,i} > 0)}{\mu(x_{C,m,i} | x_{C,m,i} > 0)} & x_{C,m,i} \geq \tilde{x}_{C,m} \\ 0 & x_{C,m,i} < \tilde{x}_{C,m} \end{cases} \quad (3)$$

Disaggregation technique:

A random cascade disaggregation model [3] is used for disaggregating the 6-hourly and daily rainfall data into hourly rainfall data. In this method, between two cascade levels each rainfall amount is divided into two equal sub-time step and the value of precipitation for each of them are determined by multiplying the total interval rainfall amount by weights W_1 and W_2 being calculated by:

$$W_1, W_2 = \begin{cases} 0 \; and \; 1 & with \; probability \; P(0/1) \\ 1 \; and \; 0 & with \; probability \; P(0/1) \\ W_{x/x} \; and \; 1 - W_{x/x} & with \; probability \; P(0/1) \end{cases} \quad (2)$$

Results

Table 1 and 2 present the performance measures (mean absolute error and R-squared) of different bias correction methods for precipitation data of CaPA points near stations HY021 and HY027 of Don River and HY037 and HY041 of Humber River watersheds for 6-hourly and daily precipitation data. It can be seen that the LOCI bias correction method has the best performance in all stations. Furtheremore, the scatter plots of the LOCI bias corrected 6-hourly and Daily rainfall CaPA data in comparison with the original CaPA data for all four stations are illustrated in Figures 2 and 3, respectively.

Table 1. Performance measurements of different bias correction methods for 6-hourly data

Stations	No Corr.		GG		EG		LOCI	
	MAE	R2	MAE	R2	MAE	R2	MAE	R2
HY021	0.45	0.49	0.44	0.49	0.43	0.50	0.40	0.51
HY027	0.51	0.32	0.44	0.32	0.45	0.34	0.41	0.34
HY037	0.46	0.56	0.48	0.53	0.48	0.56	0.45	0.60
HY041	0.49	0.44	0.50	0.44	0.49	0.45	0.47	0.49

*GG and EG represent the first presented bias correction method (a) by considering gamma and empirical distribution for truncated CaPA data , respectively.

Table 2. Performance measures of different bias correction methods for daily data

Stations	No Corr.		GG*		EG*		LOCI	
	MAE	R2	MAE	R2	MAE	R2	MAE	R2
HY021	1.14	0.70	1.11	0.73	1.16	0.70	0.97	0.76
HY027	1.48	0.48	1.24	0.53	1.29	0.51	1.15	0.54
HY037	1.29	0.68	1.37	0.68	1.36	0.67	1.23	0.73
HY041	1.30	0.63	1.35	0.65	1.33	0.65	1.24	0.70

*GG and EG represent the first presented bias correction method (a) by considering gamma and empirical distribution for truncated CaPA data , respectively.

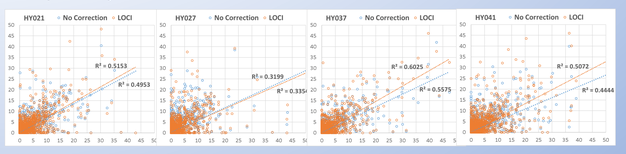

Figure 3. Comparison of original and Bias Corrected 6-hourly CaPA data using LOCI method (units of both axes: mm)

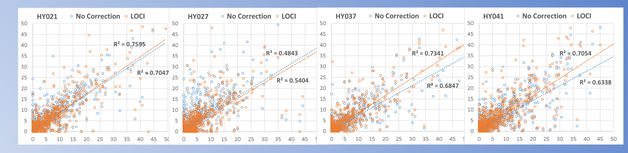

Figure 4. Comparison of original and Bias Corrected Daily CaPA data using LOCI method (units of both axes: mm)

Finally, the LOCI bias-corrected 6-hourly and Daily CaPA data for HY021 and HY027 of Humber River watershed which have rainfall time series from 2008 to 2013 are disaggregated to hourly precipitation using the random cascade disaggregation model and then different bias correction methods are used for the disaggregated hourly data. The disaggregated hourly and the best bias corrected results are presented in Figure 5.

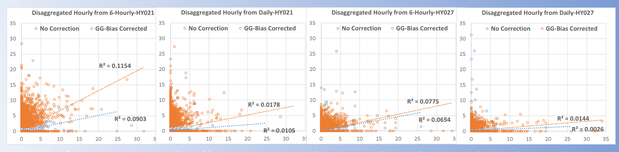

Figure 5. Hourly disaggregated CaPA data from 6-hourly and Daily bias-corrected CaPA rainfall data and the best bias corrected disaggregated hourly data (units of both axes: mm)

Discussion and Conclusion

The bias correction results show that the two quantile mapping bias correction methods are not appropriate for 6-hourly and daily CaPA data and the reason is that the distribution of original data is close to the observed data in various stations, although, the magnitudes of 6-hourly data does not match enough to the corresponding observed values. On the other hand, although LOCI bias correction method somehow improves the quality of 6-hourly and daily CaPA precipitation data, it still needs more improvement for deriving a reliable hourly disaggregated data set.

Therefore, the results of the disaggregation are not reliable enough, however, it can be concluded that hourly disaggregated rainfall data based on 6-hourly bias-corrected CaPA data is better than the one derived from daily bias-corrected CaPA data. Moreover, the first bias correction method has better performance than LOCI for disaggregated hourly data.

Overall, it can be concluded that for using the CaPA data as a reliable source of data for flood forecasting, other bias correction methods or the combination of various methods should be investigated for correcting 6-hourly as well as disaggregated hourly precipitation data.

References

[1] Ines, A. V., & Hansen, J. W. (2006). Bias correction of daily GCM rainfall for crop simulation studies. Agricultural and forest meteorology, 138(1), 44-53.

[2] Schmidli, J., Frei, C., & Vidale, P. L. (2006). Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. International journal of climatology, 26(5), 679-689.

[3] Olsson, J. (1998). Evaluation of a scaling cascade model for temporal rain-fall disaggregation. Hydrology and Earth System Sciences Discussions, 2(1), 19-30.

Acknowledgment : Thanks to Canadian Natural Sciences and Engineering Research Council (NSERC) for providing the funding required for this analysis.