Sensitivity analysis of the SWAT model to spatial distribution of precipitation in streamflow simulation in an Arctic watershed

View the article online for updates and enhancements.

Sensitivity analysis of the SWAT model to spatial distribution of precipitation in streamflow simulation in an Arctic watershed

M T Bui[1], J Lu[1] and L Nie[2]

1 Department of Technology and Safety, Faculty of Science and Technology, UiT The Arctic University of Norway, Tromsø, Norway
2 Centre for Sustainable Development and Innovation of Water Technology, CSDI WaterTech, Oslo, Norway
E-mail: minh.t.bui@uit.no

Abstract. This study approached a physically based, semi-distributed SWAT model to test the model sensitivity to the spatial distribution of precipitation. Ten scenarios of precipitation from five scattered rain gauges in an Arctic watershed Målselv in northern Norway were used as inputs to run the SWAT model. Streamflows were simulated. The model runs at monthly time interval based on the historical data of precipitation from 1979-2012. The study used statistical parameters, values of long-term average monthly streamflow and streamflow hydrograph between simulated and observed data for sensitivity analysis. The study found that the result of streamflow simulation is highly sensitive with spatial distribution of rain gauges input. For instance, the scenarios integrating rain gauge number 3, locating inside the watershed with lower precipitation amount than average of selected rain gauges, provided model unsatisfactory (statistical coefficient NSE<0.5) in streamflow simulation. However, streamflow simulation is satisfactory (NSE: 0.5-0.6) at hydro-gauging station Lundberg far away from rain gauge 3. The hydrograph showed underestimated streamflow simulation in scenario 3,5,6-10 that integrated rain gauge 3, while scenario 1,2,4 that excluded rain gauge 3 showed reasonable agreement between simulated and observed flow. Underestimated streamflow was only found in scenario 3 and 5 at Lundberg. Moreover, the curves of average monthly streamflow showed that the simulated peak discharge in scenario 1,2,4 was performed better than the remaining scenarios.

1. Introduction

The SWAT-Soil and Water Assessment Tool [1] is a state-of-the-art tool for environmental and water resources management in a river basin. To perform the simulation, various input data in both spatial and temporal dimension are necessary to be collected, of which precipitation is considered as the primary input for the hydrological model [2, 3]. However, the major challenge of the hydrological model is to accurately simulate the spatial variations of precipitation across the whole river basin [2]. Normally, areal precipitation for the whole river basin is calculated based on monitoring data from existing rain gauges [2]. However, number of available rain gauges are often inadaquate to accurately represent the heterogeneous distribution of precipitation in many river basin [2, 4]. This is especially true in the Arctic region, where rain gauges network are usually scattered. It is, therefore, spatial variation of selected rain gauges in a data sparse region could influence the hydrological simulation [3], particularly streamflow which is considered as the most important hydrological variable since it is an integrated output of the water cycle in the river basin [5, 6]. The influence of spatial variation of precipitation on streamflow simulation has been proven in some previous studies around the world.

For instance, Zhao et al. (2011) [7] studied the effects of spatial rainfall variation on streamflow prediction in the Orara River, Bawden Bridge catchment in south-eastern Australia. They stated that the large spatial variability of precipitation provided underestimation of the total streamflow volume as well as the entire daily streamflow. In another study, Aouissi et al. (2013) [8] examined the sensitivity of SWAT model to spatial distribution of rainfall in streamflow simulation in a watershed locating in Mediterranean region namely the Joumine in Tunisia. They concluded that SWAT model was extremely sensitive to spatial distribution of rain gauges network over the watershed in simulation of daily and monthly streamflow. Recently, Sirisena et al. (2018) [2] investigated four different precipitation dataset with different spatial resolution of rain gauges network to simulate streamflow in the Irrawaddy River basin in Myanmar. They found that the four input datasets of precipitation significantly impacted on model performance, estimation of model parameters, especially uncertainty in streamflow simulation. Moreover, other studies also took into account the the effect of spatial rainfall variation on streamflow simulation such as Arnaud et al. (2002) [9], Segond et al. (2007) [10] and Arnaud et al. (2011) [11]. Most of these studies focused on the influences of spatial rainfall variation on simulating the flood events, and they stated that prediction of runoff volume, peak flow and the timing of hydrographs were impacted by spatial rainfall variation. As many studies were carried out in various regions around the world including tropical region or arid region, it raises a question that whether spatial variation of precipitation could influence streamflow simulation in the Arctic condition or not. This study was conducted to answer this question.

2. Study area

The Målselv river basin located in the north of Norway was selected as the study area (Figure 1). The river basin covers an area of 5,912.8 km^2. The elevation of ground surface over the whole area varies from 0 to 1,718 m. According to long-term historical data from the Norwegian Water Resources and Energy Directorate [12], the average annual precipitation in the study area varies from below 500 mm to 1,000 mm. The average annual air temperature fluctuates from -5 ^{o}C to 6 ^{o}C.
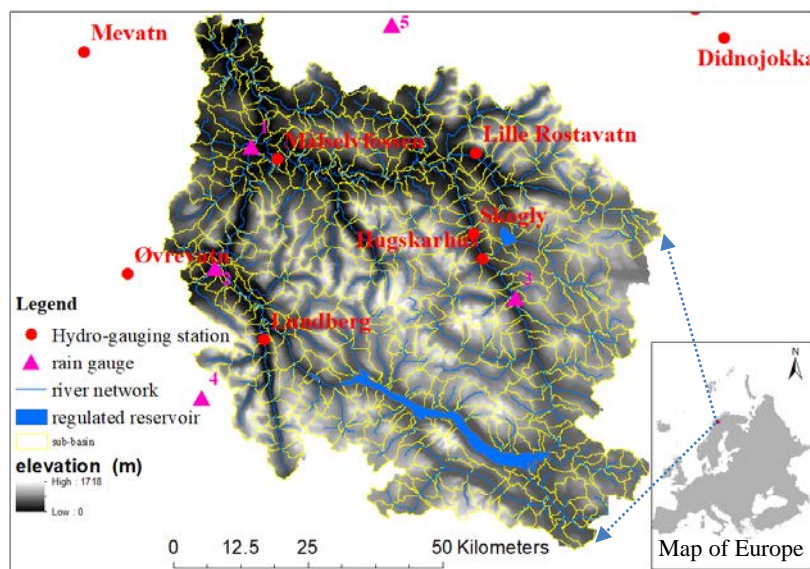

Figure 1. Map of study area and distribution of the rain gauges and hydro-gauging stations

3. Material and methodologies

3.1. SWAT model

The SWAT (Soil and Water Assessment Tool) is a physically based, semi-distributed model that was developed to simulate the impacts of anthropogenic activities on water resources and associated environmental issues in a large-scale watershed with complex conditions over long period [13, 14].

Also, the SWAT model is used to investigate the impacts of climate change scenarios [15]. The model works based on the following water balance equation [1]:

$$SW_t = SW_0 + \sum_{i=1}^{t} (R_{day} - Q_{surf} - E_a - w_{seep} - Q_{gw})$$
(1)

where:

- SW_t is the final soil water content in mm H_2O,
- SW_0 is the initial soil water content on day i in mm H_2O,
- t is time in days,
- R_{day} is amount of precipitation on day i in mm H_2O,
- Q_{surf} is amount of surface runoff on day i in mm H_2O,
- E_a is amount of evapotranspiration on day i in mm H_2O,
- w_{seep} is amount of water entering the vadose zone from the soil profile on day i in mm H_2O,
- Q_{gw} is amount of return flow on day i in mm H_2O.

3.2. Data acquisition

To run the SWAT model, several input data from different sources were collected. Digital Elevation Map (DEM) with 10 m x 10 m resolution, raster of soil type and land use were collected from WATERBASE (http://www.waterbase.org/download_data.html). Daily precipitation at five rain gauges, coded from 1 to 5 (Figure 1), was collected from ECA&D-The European Climate Assessment & Dataset project (https://www.ecad.eu//dailydata/index.php). The other data such as maximum and minimum air temperature, solar radiation, humidity, wind speed were collected from CFSR-Climate Forecast System Reanalysis (https://globalweather.tamu.edu/). Totally 34 years (1979-2012) of precipitation data and other climate data were collected. Measured streamflow from five hydro-gauging stations namely Hugskarhus, Skogly, Lille Rostavatn, Målselvfossen, Lundberg (Figure 1) and other datasets such as regulated reservoirs, river systems were collected from NVE-The Norwegian Water Resources and Energy Directorate [12].

3.3. Sensitivity analysis

3.3.1. Scenarios of precipitation inputs

Ten scenarios of precipitation inputs were developed in this study (Table 1) based on the spatial distribution of totally five rain gauges scattered over the watershed and its surroundings. The rain gauges numbered 1-3 are located inside the watershed, while rain gauges 4 and 5 are outside. The first seven scenarios used the rain gauges inside the watershed, of which three scenarios used single rain gauge, three scenarios combined two rain gauges and one scenario combined three rain gauges. The next two scenarios used four rain gauges that combined three rain gauges inside the watershed and one station outside. The last scenario combined all five rain gauges. The scenarios have stations outside the watershed aim to examinate the interaction between stations inside and outside the watershed.

Table 1. Scenarios of precipitation inputs

Scenario (sc)	Number of rain gauges integration	Rain gauge ID
1	1	1
2	1	2
3	1	3
4	2	1 & 2
5	2	1 & 3
6	2	2 & 3
7	3	1 & 2 & 3
8	4	1 & 2 & 3 & 4
9	4	1 & 2 & 3 & 5
10	5	1 & 2 & 3 & 4 & 5

3.3.2. Model performance evaluation

The model ran from 1979-2012, with a nine-year warming up period to let the model to reach an optimal state from the estimated initial condition [16]. The simulated and observed results were compared using several statistical parameters such as the Nash-Sutcliffe coefficient of efficiency-NSE (equation (2)), and Percentage bias-PBIAS (equation (3)) which was recommended by Moriasi et al. (2007) [17]. Also, Pearson Correlation Coefficient-CC (equation (4)) was used to measure the fitness of the linear relationship between simulated and observed data/results [17].

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(Y^{obs}_i - Y^{sim}_i\right)^2}{\sum_{i=1}^{n}\left(Y^{obs}_i - Y^{mean}\right)^2} \tag{2}$$

$$PBIAS = \frac{\sum_{i=1}^{n}\left(Y^{obs}_i - Y^{sim}_i\right) * 100}{\sum_{i=1}^{n} Y^{obs}_i} \tag{3}$$

$$CC = 1 - \frac{\sum_{i=1}^{n}\left(Y^{obs}_i - Y^{obs}_{mean}\right)\left(Y^{sim}_i - Y^{sim}_{mean}\right)}{\left[\sum_{i=1}^{n}\left(Y^{obs}_i - Y^{obs}_{mean}\right)^2\right]^{1/2}\left[\sum_{i=1}^{n}\left(Y^{sim}_i - Y^{sim}_{mean}\right)^2\right]^{1/2}} \tag{4}$$

where

- Y^{obs}_i and Y^{sim}_i are the observed and simulated value at time i,
- Y^{obs}_{mean} and Y^{sim}_{mean} are mean observed and simulated data for the entire evaluation period,
- n is the total number of observations/simulations.

Table 2 illustrates the threshold values of NSE and PBIAS for evaluating the model performance.

Table 2. NSE and PBIAS thresholds for evaluating the hydrological model performance

Model performance	NSE	PBIAS
Very good	$0.75 < NSE \leq 1.00$	$PBIAS \leq \pm 1.0$
Good	$0.65 < NSE \leq 0.75$	$\pm 1.0 \leq PBIAS \leq \pm 15$
Satisfactory	$0.50 < NSE \leq 0.65$	$\pm 15 \leq PBIAS \leq \pm 25$
Unsatisfactory	$NSE \leq 0.50$	$PBIAS \geq \pm 25$

CC value should be equal to or greater than 0.5 to indicate that two variables have correlation [18].

4. Results and discussion

4.1. Model evaluation based on value of Pearson Correlation Coefficient (CC)

All ten scenarios of precipitation inputs provided a good linear relationship between simulated and observed data (Figure 2a). The maximum of CC value was 0.88 and minimum was 0.59. Values of CC at Lundberg and Målselvfossen hydro-gauging stations were higher than the remaining stations. At Høgskarhus, the linear relationship between simulation and observation for all scenarios was not strong compared to other stations.

4.2. Model evaluation based on the Nash- Sutcliffe coefficient of efficiency (NSE)

In general, the first two scenarios sc1 and sc2 provided higher model performance at five hydro-gauging stations compared to other scenarios, with highest is 0.6 at Skogly (in sc1) and 0.62 at Lille Rostavatn (in sc2). Scenario 3 and scenario 5 provided unsatisfactory results at all five hydro-gauging stations. The scenario 4 provided satisfactory result at Målselvfossen and closely to satisfactory at four remaining stations as value of NSE is close to 0.5. The scenarios 6, 7, 8, 9 and 10 provided satisfactory results at Lundberg and unsatisfactory result at four remaining stations (Figure 2b).

4.3. Model evaluation based on Percentage bias (PBIAS)

The scenarios 5-10 provided unsatisfactory results at all hydro-gauging stations. Scenario 1 provided good performance at Høgskarhus and Skogly, and scenario 2 provided satisfactory results at Høgskarhus, Skogly, and Målselvfossen. The scenario 3 showed good performance at Målselvfossen, while the scenario 4 showed good performance at Skogly and satisfactory at Høgskarhus (Figure 2c).
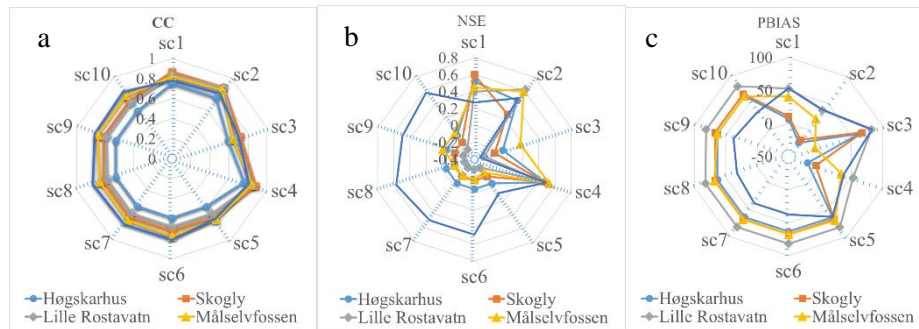

Figure 2. CC, NSE and PBIAS for different scenarios in monthly time step over 1988-2012 period

4.4. Model evaluation in prediction of streamflow hydrograph

4.4.1. At Høgskarhus (subbasin 408)

Høgskarhus hydrology station locates in the upstream of Målselv compared to other. Also, this station is very close to rain gauge number 3. The modelling result showed that using precipitation at two stations 1 and 2 resulted in better streamflow hydrograph, in scenario 1, 2 and 4 (Figure 3a), compared to other rain gauges, in scenario 3 and scenarios 5-10 (Figure 3b).

4.4.2. At Skogly (subbasin 412)

Skogly station is in downstream of Høgskarhus. These two stations are very close and near rain gauge station 3. Hence, precipitation input from rain gauge number 3 resulted in similar behavior of streamflow hydrograph at Skogly and Høgskarhus. For instance, results from scenario 1, 2 and 4 (Figure 3c) were better than results from others scenarios (Figure 3d).

4.4.3. At Lille Rostavatn (subbasin 402)

Lille Rostavatn is in another tributary compared to Skogly and Høgskarhus. However, this station also received most of precipitation input from rain gauge number 3 like Skogly and Høgskarhus as this station is also close to rain gauge number 3 compared to remaining rain gauges. Therefore, the streamflow hydrograph from scenario 1, 2 and 4 (Figure 3e) were better than remaining scenarios (Figure 3f).

4.4.4. At Målselvfossen (subbasin 444)

Målselvfossen is located almost downstream of Målselv River and close to rain gauge number 1. The scenarios that used precipitation from this rain gauge station resulted in low model performance (Figure 3g). However, due to influence of precipitation from rain gauge number 3, like at other hydrological stations, prediction of streamflow hydrograph at Målselvfossen from scenario 1,2 and 4 (Figure 3h) were better than other scenarios.

4.4.5. At Lundberg (subbasin 381)

Lundberg is located in upstream of Bardu River, which is a tributary of Målselv River. This hydrological station is in the middle of rain gauge number 2 and 4. Unlike other hydrological stations, Lundberg is less influenced by precipitation from rain gauge number 3. The model performances, therefore, were better. For instance, eight scenarios (Figure 3i) performed better than two remaining senarios (Figure 3j).
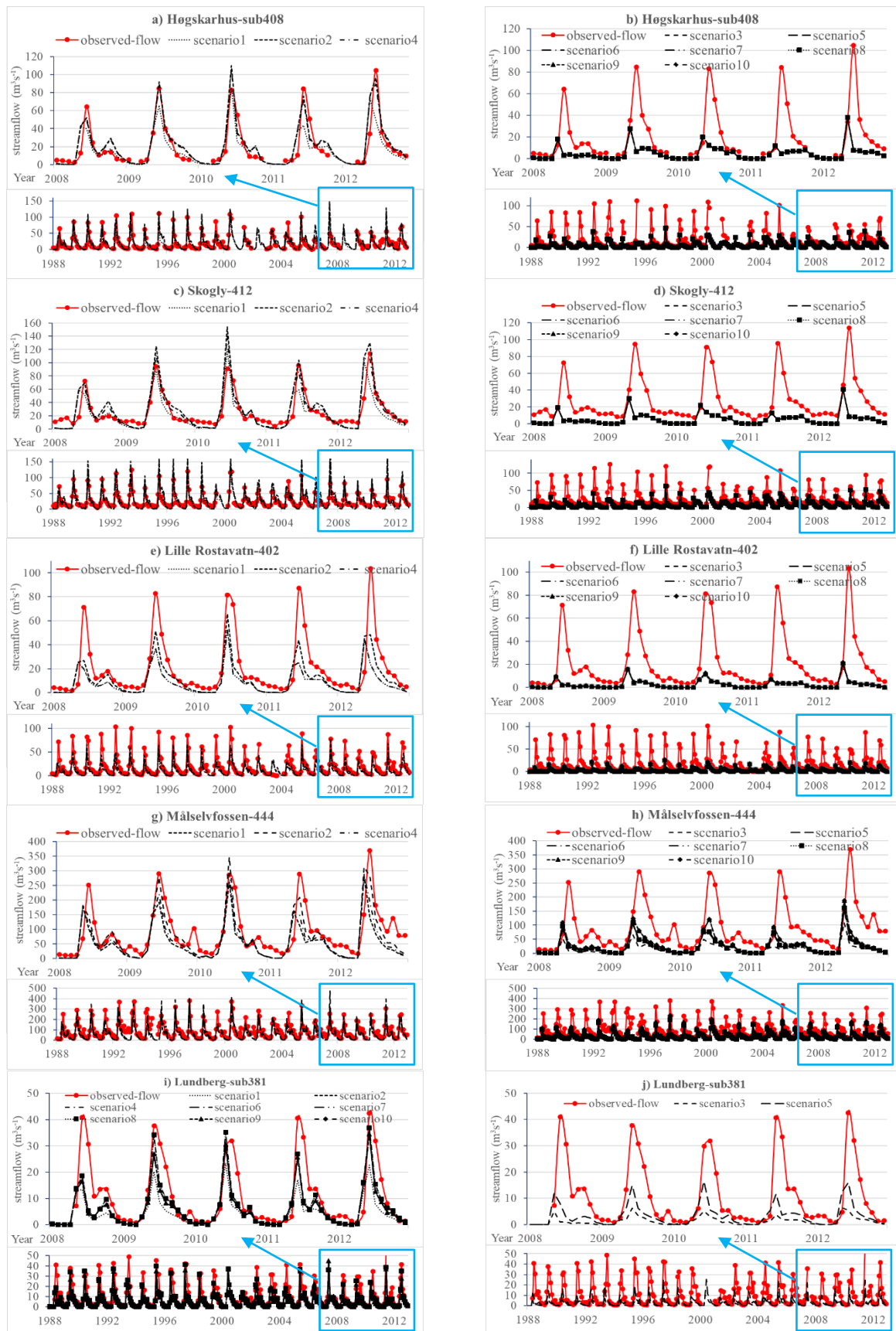
Figure 3. Hydrograph of observed & simulated streamflow at Høgskarhus (a,b), Skogly (c,d), Lille Rostavatn (e,f), Målselvfossen (g,h) and Lundberg (i,j) over 1988-2012 period

4.5. Model evaluation in simulation of average monthly streamflow

Monthly streamflow over 25 years from the model was averaged for all five hydro-gauging stations (Figure 4). In general, scenario 1, 2 and 4 simulated peak flow better than other scenarios. Particularly, at upstream station, Høgskarhus, the simulated peak discharge was close to observed data at scenarios 1, 2 and 4. It showed similar behavior for scenario 1 at Skogly. At downstream station, Målselvfossen, peak discharge between simulation and observation was met perfectly at scenario 2, and followed by scenario 4 and 1. However, at Lille Rostavatn and Lundberg, peak discharge was underestimated for all scenarios.
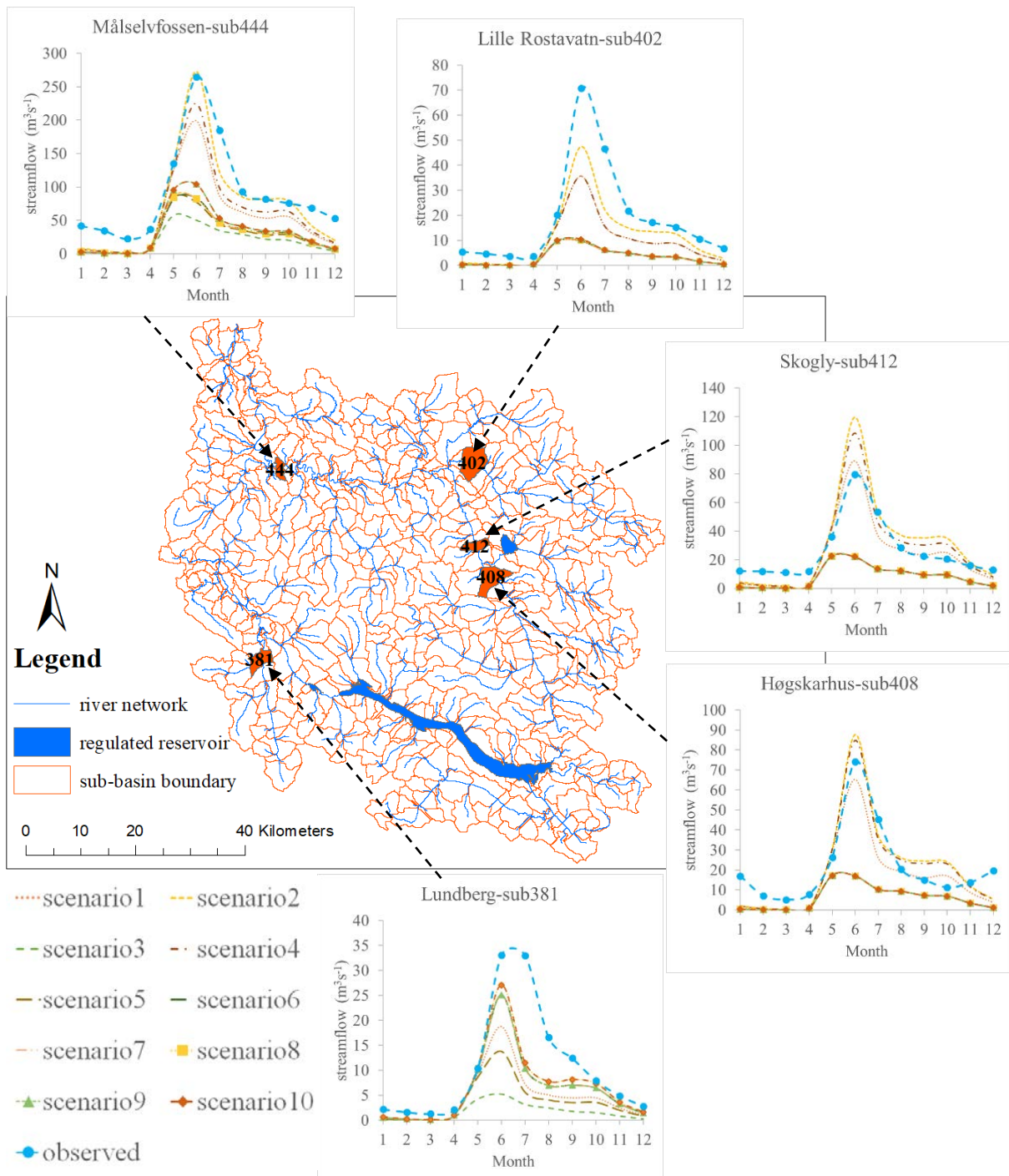

Figure 4. Average monthly streamflow between prediction and observation data at five hydro-gauging stations over 1988-2012 period

4.6. Evaluating the influence of number of rain gauges integration on model performance in streamflow simulation

The study found that increasing number of rain gauges integration did not improve model performance in streamflow simulation since the present of an unrepresentative rain gauge, particularly rain gauge number 3 close to Høgskarhus and Skogly hydro-gauging stations. Precipitation amount at rain gauge number 3 is much smaller than other rain gauges. For instance, the average annual precipitation over 1979-2012 period at this rain gauge is around 344 mm, while it is 683 mm, 922 mm, 855 mm, and 831 mm at rain gauge 1, 2, 4 and 5 respectively (Figure 5). Therefore, the selected rain gauge, rain gauge number 3, could not represent the precipitation of the nearby subbasins or for the whole watershed. Especially, in SWAT model, the nearest neighbor search (NNS) approach was used to calculate the areal precipitation for each sub-basin, and this could result in inaccurate representation of precipitation input for each sub-basin. Therefore, the subbasins, which received precipitation from rain gauge number 3 or integrated precipitation from such rain gauge with other stations, have inaccurate results of streamflow simulation. The hydrograph in Figure 3 illustrated reasonable agreement between simulated and observed streamflow at the monthly time step for five hydro-gauging stations when using precipitation from rain gauge number 1 and 2, but it was unsatisfactory when using precipitation from rain gauge number 3 and in scenarios integrating it with other rain gauges. It showed the same behavior for prediction of peak flow. Hence, it could conclude that rain gauge number 3 only represents for local precipitation and could not represent for the whole watershed.
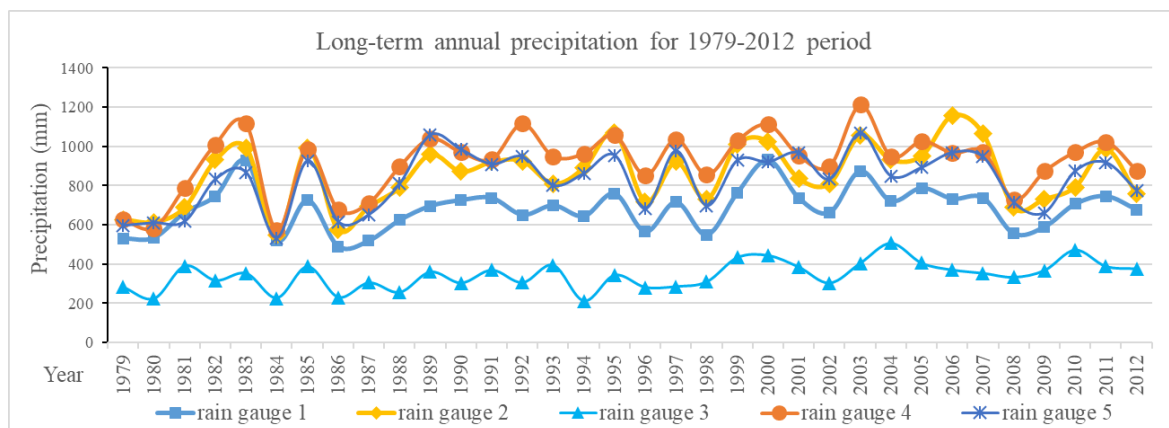

Figure 5. Long-term annual precipitation at five rain gauges over 1979-2012 period

5. Conclusion

Precipitation is the primary input variable to the hydrological model. Several previous studies have confirmed the influence of the spatial distribution of precipitation on streamflow simulation in different regions around the world. This study again tested such hypothesis but in the Arctic condition. Ten scenarios of precipitation input were set up for the investigation in this study. The scenarios were built based on five scattered rain gauges locating both inside and outside the watershed. The study detected one rain gauge (rain gauge number 3) in totally five rain gauges that highly impacted the modelling results as well as model performance. Precipitation from rain gauge 3 itself and integration of this rain gauges with others leaded to model unsatisfactory. The reason is because of precipitation amount at rain gauge 3 is much smaller than its in other stations. Therefore, it leaded to unaccuracy in calculating the representative areal precipitation for the nearby sub-basins, and then influenced the simulation results of runoff and streamflow. However, streamflow simulation at the hydro-gauging station (at Lundberg) that locates far away from rain gauge 3 has less influenced by precipitation input from this rain gauge. This obviously indicated the effect of the spatial distribution of rain gauges across the watershed. Simulation of the average monthly streamflow, especially peak flow of the hydrograph was highly influenced by the present of rain gauge number 3. Herein, peak flow of the hydrograph was much underestimated compared to observed data for the scenarios using precipitation input from rain gauge number 3.

6. References

[1]   SWAT: Soil and Water Assessment Tool Available from: https://swat.tamu.edu/docs/ (Accessed 22 November 2019)

[2]   Sirisena, T.A.J.G., et al. 2018 Effects of different precipitation inputs on streamflow simulation in the Irrawaddy River Basin, Myanmar. *Journal of Hydrology-Regional Studies* **19** 265-78.

[3]   Strauch, M., et al. 2012 Using precipitation data ensemble for uncertainty analysis in SWAT streamflow simulation *Journal of Hydrology* **414** 413-24

[4]   Miao, C.Y., et al. 2015 Evaluation of the PERSIANN-CDR Daily Rainfall Estimates in Capturing the Behavior of Extreme Precipitation Events over China *Journal of Hydrometeorology* **16** 1387-96

[5]   Shamir, E., et al. 2005 The role of hydrograph indices in parameter estimation of rainfall-runoff models *Hydrological Processes* **19** 2187-2207

[6]   Sun, W.C., et al. 2015 Estimating daily time series of streamflow using hydrological model calibrated based on satellite observations of river water surface width: Toward real world applications *Environmental Research* **139** 36-45

[7]   Zhao, F.F., et al. 2011 The effect of spatial rainfall variability on streamflow prediction for a south-eastern Australian catchment *19th Int. Congress on Modelling and Simulation (Modsim2011)* p 3684-90

[8]   Aouissi, J., et al. 2013 Sensitivity analysis of SWAT model to the spatial rainfall distribution and watershed subdivision in streamflow simulations in the Mediterranean context: a case study in the Joumine watershed. Tunisia *5th Int. Conf. on Modeling, Simulation and Applied Optimization (Icmsao)* p 1-6

[9]   Arnaud, P., et al. 2002 Influence of rainfall spatial variability on flood prediction *Journal of Hydrology* **260** 216-30

[10]   Segond, M.L., H.S. Wheater, and C. Onof 2007 The significance of spatial rainfall representation for flood runoff estimation: A numerical evaluation based on the Lee catchment, UK *Journal of Hydrology* **347** 116-31

[11]   Arnaud, P., et al. 2011 Sensitivity of hydrological models to uncertainty in rainfall input *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* **56** 397-410

[12]   *The Norwegian Water Resources and Energy Directorate.* (Accessed 18 November 2019)

[13]   Gassman, P.W., et al. 2007 The soil and water assessment tool: Historical development, applications, and future research directions *Transactions of the Asabe* **50** 1211-50

[14]   Ndomba, P., F. Mtalo, and A. Killingtveit 2008 SWAT model application in a data scarce tropical complex catchment in Tanzania *Physics and Chemistry of the Earth* **33** 626-32

[15]   Dile, Y.T., R. Berndtsson, and S.G. Setegn 2013 Hydrological Response to Climate Change for Gilgel Abay River, in the Lake Tana Basin - Upper Blue Nile Basin of Ethiopia *Plos One* **8** 1-13

[16]   Kim, K.B., H.H. Kwon, and D.W. Han 2018 Exploration of warm-up period in conceptual hydrological modelling J*ournal of Hydrology* **556** 194-210

[17]   Moriasi, D.N., et al. 2007 Model evaluation guidelines for systematic quantification of accuracy in watershed simulations *Transactions of the Asabe* **50** 885-900

[18]   Mukaka, M.M. 2012 Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research *Malawi Medical Journal* **24** 69-71

Acknowledgments