SWAT-modeling of the effects of the construction of a dam on streamflow in the Karkheh basin, Iran

Majid Fereidoon[1,*] and Manfred Koch[1]

[1]Department of Geotechnology and Geohydraulics,
University of Kassel, Kassel 34125, Germany

Abstract

Hydrology and sustainable water development manner needs effective management of water resources, so it is necessary to understand the various watershed components with their governing processes. Therefore, in the present application of the semi distributed physically based Soil and Water Assessment Tool (SWAT) model in the Karkheh river basin (KRB) in the semiarid region of Iran, the filling of the gaps in the data series required to drive and calibrate the model, has been the first major task. Two methods, linear regression with the nearest station (LRN) and inverse distance weighting (IDW) are investigated for filling in gaps in the precipitation and temperature series at 10 weather stations. SWAT-CUP (SWAT-Calibration and Uncertainty Programs) is used for model calibration and sensitivity analysis, following the Sequential Uncertainty Fitting (SUFI-2) technique. Calibration and validation of the SWAT-model is performed on measured stream flows at 8 gauge stations for the 1985-1999- and 2000-2004 time periods, respectively. The calibrated SWAT-model performs fine for the prediction of the monthly streamflow at the outlets, as witnessed by the four following statistical measures: (i) P-factor, defining the percentage of observation data bracketed by 95 percentage prediction uncertainty (95PPU), (ii) R-factor: which is the relative width of the PPU, (iii) Nash Sutcliffe coefficient (NSE), and (iv) coefficient of determination of (R2). The P-factor values for the various gauge stations range from 0.62 to 0.94 and 0.60 to 0.88 for calibration and validation, respectively, which, according to literature, identify reasonable accuracy. NSE values are also acceptable, ranging from 0.52 to 0.82 and from 0.62 to 0.80 for calibration and validation, respectively. The same holds for the R2 -values which are also within an acceptable range. The Karkheh dam started its operation in August 2002, i.e. during the SWAT-model's validation period, and its effect is clearly indicated by a significant reduction of the average annual streamflow in the two subsequent years 2003-2004, compared with that of the pre-operation period (1985-2001). Thus, the average annual flow volumes for the downstream gauge stations Pay-e-Pol and Hamidiyeh are, respectively, 2.19×10^3 m^3 and 1.93×10^3 m^3 in the 1985-2002- and 1.49×10^3 m^3 and 9.66×10^2 m^3 in the 2002-2004 time period, which corresponds to a reduction of 30% and 50%, respectively.

1. Introduction

Water use efficiency, especially, in the arid and semi-arid regions of Iran, because of increasing water use and lengthening of dry periods, requires understanding of the hydrological process to develop a suitable model. To that avail the use of hydrological models is nowadays an invaluable tool. However, for the successful build-up and assessment, i.e. calibration and validation of such a hydrological model, the availability of complete and continuous climatological and hydrological time series is indispensable. This may become a particular problem in developing countries where financial and logistic constraints often inhibit complete data recordings, i.e. gaps are common. Several authors [1, 2] have indicated that incomplete datasets increase the level of complexity and uncertainty in hydrological modeling. Especially, big gaps in datasets time series may hide the pattern of the real data, and they may considerably distort the

results of any statistical analysis or climatological/ hydrological modeling. Thus, before the later task can be endeavored, it is necessary to use some optimal method for infilling the missing data.

Numerous gap-filling approaches [2 - 4] have been developed for hydro-meteorological time series that use some statistical information on the structure of the time series itself, or information from near neighbor stations. However, the jury is still out which method is the most suitable in a particular application.

In this paper which deals with the evaluation of the applicability and performance of the well-known SWAT hydrological model [5] to assess the role of the Karkheh dam, in southwest Iran, on the stream flow, numerous incomplete precipitation and temperature time series had to be dealt with before the modeling task proper. To that avail, linear regression with the

*Corresponding author; e-mail: majid.fereidoon@gmail.com
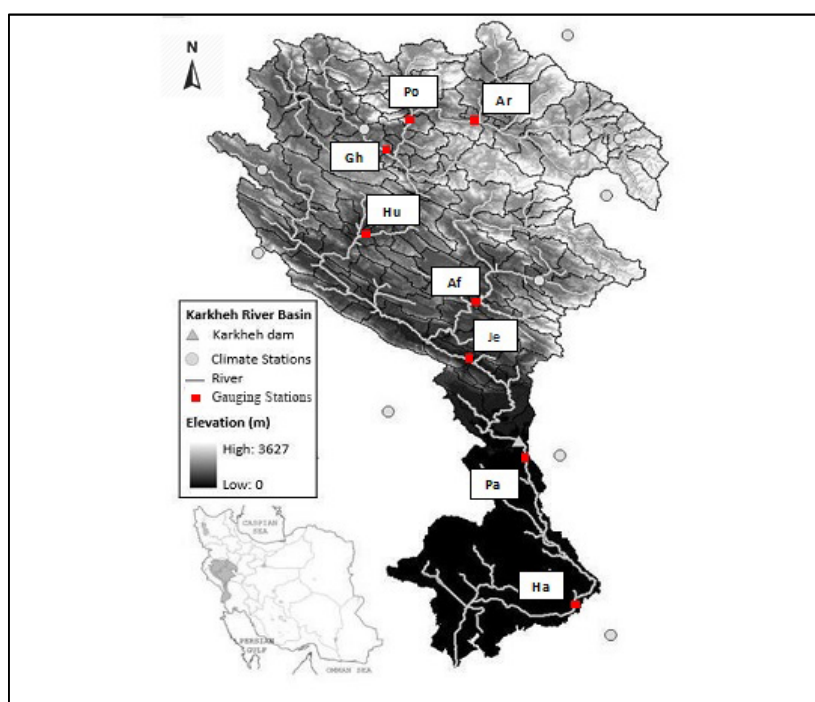
Figure 1 Map of the Karkheh river basin with weather stations and the Karkheh dam

nearest station (LRN) and inverse distance weight (IDW) methods for filling in the gaps in time series will be used. The filled-in time series are then employed used as drivers in the SWAT-model to assess the effects of climatological variability and other hydrologic parameters on the streamflow and the monthly water yield at the outlets of this watershed.

Study area

The Karkheh river basin (KRB) located in southwest Iran, between 30°58–34°56 N latitude and 46°06–49°10 E longitude (Figure 1) is the third largest watershed in Iran and has the third highest average annual outflow after the Karoon and Dez rivers. Nearly two thirds of the basin lies in the mountains (minimum altitude of 3m above sea level in the south to a maximum of 3645 m in the north), and surface and ground water resources are replenished from winter snow falls in the high Zagros mountain change ranges. The river becomes progressively more saline as it flows downstream of the Karkheh dam. The basin's area is approximately 50700 km^2 and a mean annual ground water recharge of 3.4 km^3. Population growth has negative influence on land use changes, and has put pressure on the water resources and productivity in KRB during the last few decades.

The climate of KRB is arid and semiarid, with large differences in the average annual precipitation that ranges between 150 mm in the south and 7500 mm in the upper parts of the KRB [6].

The multipurpose Karkheh dam which is located in the northwestern province of Khuzestan became operational in August 2002. The objectives of the dam are the storage and regulation of water for irrigation of 320,000 ha of the agricultural lands in the downstream plains, as well as hydroelectric power generation, amounting to 934 GWh per year [7] And, of course, the Karkheh dam serves also for the prevention of destructive floods in the downstream sections of the river basin [8 - 11].

2. Materials and methods

2.1 Data

The hydroclimate data consists of daily precipitation and temperature time series from 10 climate stations shown in Figure 1, as well as outflow data from the Karkheh reservoir, and streamflow records at 8 gauging stations, all recorded between 1982-2004. As mentioned, most of the meteorological series have gaps of up to 20%. Examples of two infilled sections – by the Inverse Distance Weighting method to be discussed further down - of the precipitation series in a particular 18-month time window, as well for the fully incomplete month of January 1988, are illustrated in Figure 2.

The SWAT model uses a digital elevation model (DEM) map at a resolution of 90 m (NASA) to delineate the watershed and to extract the stream network in the catchment. Land-use-maps with a resolution of 900 m was obtained from MahabGhods Engineering company and soil data further required in SWAT from FAO (1995) for two depth layers (0-30 cm and 30-100 cm depth) at a spatial resolution of 10 km. Maps of this data are shown in Figure 2.

(a) DEM   (b) Land Use Map   (c) Soil Map
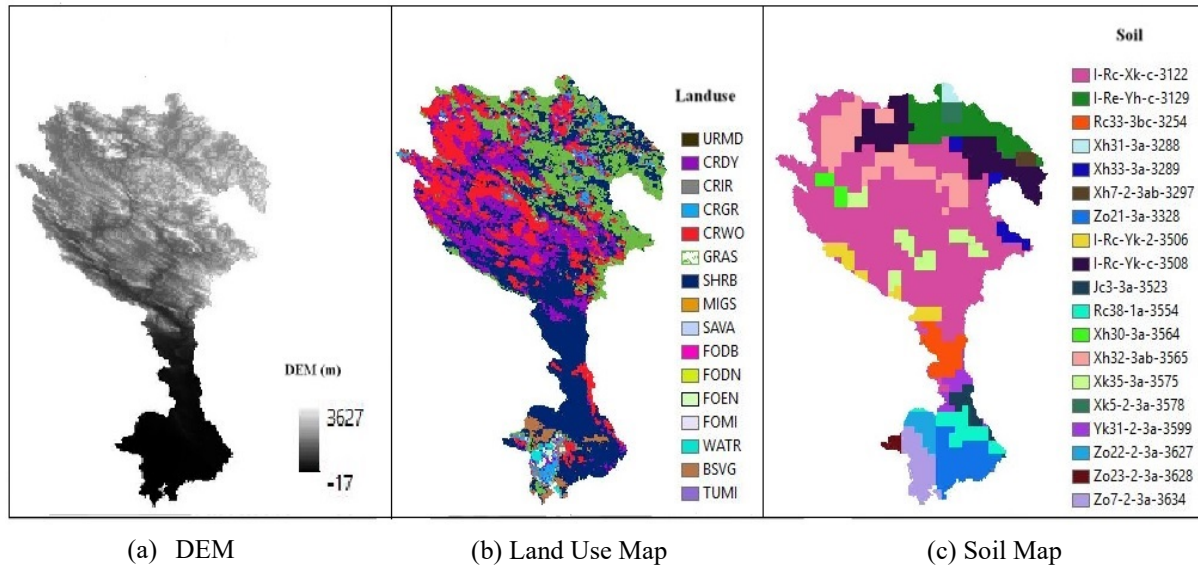
Figure 2 Digital elevation model (a), land-use (b) and soil map of the Karkheh basin used in SWAT

2.2 Gap-infilling methods

Two different gap-filling techniques for estimating daily precipitation and temperature series were evaluated; Inversed Distance Weighting (IDW) and Linear Regression with the Nearest station (LRN) methods, were assessed and applied to SWAT model. These are described in the following two sub-sections.

1) Inversed distance weighting (IDW)

In the IDW-method missing values in an observational dataset at a particular weather station are estimated by applying the following equation [12].

$$V_{est} = \frac{\sum_{i=1}^{n} v_i * d_i^{-k}}{\sum_{i=1}^{n} d_i^{-k}} \qquad (1)$$

where V_{est} is the unknown target weather station value, V_i is value in the weather station i, n is the number of stations and k is the power of distance, referred to as a friction distance ranging between 0.5 and 2.

2) Linear regression with the nearest station (LRN)

In the LRN- gap filling method a linear regression between the unknown target value V_{est} and those from a selected, nearby station V_i is performed, i.e.

$$V_{est} = b * V_i + a \qquad (2)$$

where a and b are the estimated regression parameters. That time series of the nearby station is selected whose Pearson correlation coefficient with the series of the target station is the highest [13].

The statistical measures, Root Mean Square Error, RMSE, and the coefficient of determination, R2, are used to evaluate the performance of the two gap-filling methods:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(F_i - Obs_i)^2}{n}} \qquad (3)$$

$$R^2 = 1 - \left(\frac{\sum_{i=1}^{n}(Obs_i - F_i)^2}{\sum_{i=1}^{n}(Obs_i - \overline{Obs_i})^2}\right) \qquad (4)$$

where F_i is the filled-in daily value, Obs_i is the average for a day for all years with observed data, for all years, and $\overline{Obs_i}$ is the average of Obs_i for the n-missing values.

2.3 SWAT hydrological model

The SWAT model is a basin-scale hydrological model that operates on a daily time step to quantify the impact of land management practices, agricultural chemical yields, sediment in large and complex watersheds with varying soils, land use and management conditions [14]. The model is process based, computationally efficient, and capable of continuous simulation over long periods. Major model components include weather, hydrology, soil temperature and properties, plant growth, nutrients, pesticides, bacteria and pathogens, and land management.

In SWAT the total watershed area, based on the DEM, is divided first in sub-basins (198 sub-basins in the present basin) which are further divided in so-called hydrologic response units (HRUs) which are defined as areas having the homogeneous land use, management, topographical, and soil characteristics. The HRUs are represented as a percentage of the sub-basin area and may not be contiguous or spatially identified within a SWAT simulation. Alternatively, a watershed can be subdivided into only sub-basins that are characterized by dominant land use, soil type, and management. SWAT simulates the hydrological cycle based on the water balance equation:

$$SW_t = SW_0 + \sum_{i=1}^{t}\left(R_{day} - Q_{surf} - E_a - W_{seep} - Q_{gw}\right) \qquad (5)$$

where SW_t is final soil moisture content value, SW_0 is initial soil moisture content value, R_{day} is amount of precipitation, Q_{surf} is amount of surface runoff, E_a is amount of evapotranspiration, W_{seep} is amount of water
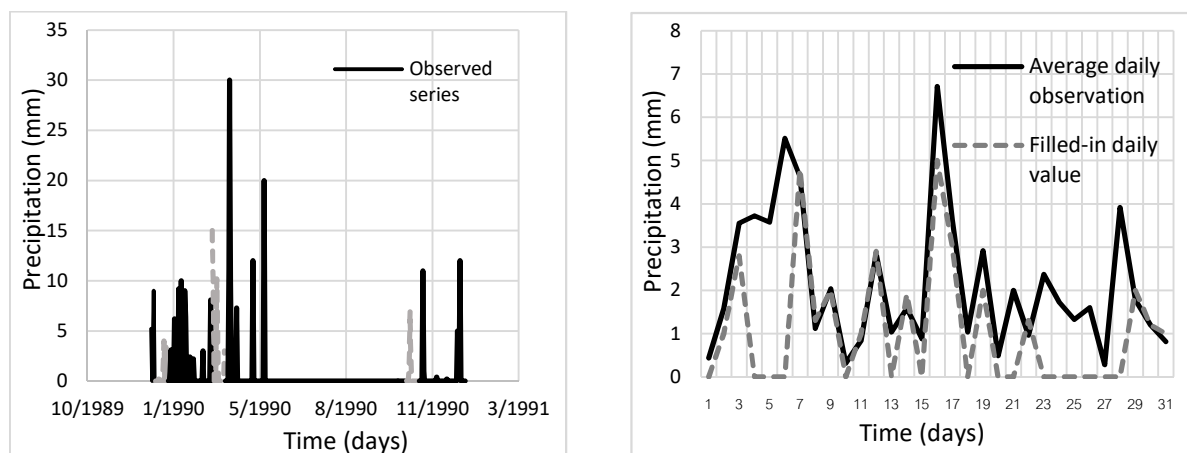
Figure 3 Left: Example of a gap-infilled (dashed) daily precipitation series at the Khorramabad climate station, Right: Average daily observation and filled-in daily values time series for January 1988

entering the vadose zone from the soil profile, Q_{gw} is amount of return flow and t is daily time step. These can be specified in Arc SWAT, the ArcGIS interface for SWAT, by means of appropriate threshold levels, to keep the number of HRUs at a manageable size. Nevertheless, because of the large size and complexity of the study basin, the number of HRU's amounts to more than 11000.

SWAT's modeled sub-basin components which are of interest in the present study, are the surface runoff, subsurface runoff, groundwater, percolation, infiltration and soil water storage, return flow, actual and potential evapotranspiration, snowmelt, transmission losses from streams and water storage and losses from ponds [15].

Because SWAT requires, in addition to the named hydro-meteorological, topography, soil and land-use data, the specification of numerous other process parameters, which are usually not well-known, calibration and subsequent validation of the SWAT-model on observed streamflow is a major task of the modeling process. Although calibration of a simple model may be somewhat achieved by a cumbersome trial- and error method, for the complex Karkheh watershed, an automatic calibration is advocated.

Two calibration/optimization methods commonly used in SWAT are the deterministic Parameter Solution (ParaSol) method [16, 17], and the more recent, stochastic Sequential Uncertainty Fitting 2 (SUFI-2, version 2) within SWAT-CUP (Calibration and Uncertainty Program) [17]. The latter has been employed here. SUFI quantifies the uncertainty of a calibrated parameter by the 95% prediction uncertainty band (95PPU) calculated at the 2.5% and 97.5% levels of the cumulative distribution function obtained through Latin hypercube sampling of the output objective function [18].

Finally the performance of the calibration model is evaluated by two indices: P-factor and R-factor. The P-factor is the percentage of observations covered by the 95PPU band. Its value ranges between 0 and 1. The R-factor denotes the relative width of the 95PPU band divided by the standard deviation of the measured variable. It ranges between 0 and infinity and a value less than 1 is stated to be desirable for a parameter [19]. The quality of the fit of the model output, namely, streamflow, to the observed one is measured by the RMSE, the R^2 or the Nash–Sutcliff efficiency (NSE) [18].

3. Results and discussion
3.1 Evaluation of the gap-filling methods

The two gap filling methods described above have only been applied to those meteorological time series which have more than 20% of the data missing. Time series of 7 out of the total 10 climate stations fulfil this criteria. In all other cases, SWAT's internal weather generator has been used for short-term gap filling. Using the two performance measures RMSE (3) and R^2 (4) of the gap filling procedures for the daily precipitation the RMSE ranges between 2.35-2.55 mm for the IDW- and between 2.48-3.12 mm for the LRN method.

For the temperatures the corresponding values are 4-8.3°C for IDW and 4.8-9 for LRN. R2, on the other hand, ranges, depending on the method and the time series considered, between 0.48-0.73. Overall, it can be stated that the IDW performed slightly better than the LRN- method. Figure 3 shows the performance of the IDW-method for the fill-in of two incomplete time sections of a precipitation series in the basin.

3.2 SWAT-modeling calibration and validation

Calibration and validation of the SWAT-model was performed on measured stream flows from 8 gauge stations for the 1985-1999 and 2000-2004 time periods, respectively. Calibration followed by sensitivity and uncertainty analysis for 26 parameters was carried out for monthly time steps for 8 gauging stations (Figure 1). The results obtained for the 7 most sensitive

Table 1 Initial and final ranges of the 7 most sensitive SWAT calibration parameters

| Parameter name	Definition	Initial range	Final range
r__CN2	SCS runoff curve number	-0.4 to 0.4	-0.32 to -0.16
v__GWQMN	Threshold depth of water in shallow aquifer required for return flow	1500 to 3500	1520 to 2538
v__ALPHA_BF	Base flow alpha factor	0.4 to 1	0.56 to 0.87
v__EPCOPlant uptake compensation factor		0.2 to 0.7	0.27 to 0.55
r__SOL_BD	Moist bulk density	0 to 0.35	0 to 0.23
v__RCHRG_DPDeep aquifer percolation fraction		0.1 to 0.7	0.25 to 0.64
v__SHALLSTInitial	depth of water in shallow aquifer	2500 to 4000	2930 to 3849

Table 2 Statistical measures for monthly outflows at the 8 gauging stations for calibration and validation

| Station	P-factor cal./val	R-factor cal/val	R^2 cal/val	NSE cal/val
Aran	0.76/0.78	1.18/1.09	0.62/0.52	0.59/0.56
Polchehr	0.81/0.85	1.19/1.03	0.68/0.64	0.52/0.63
Ghurbaghestan	0.88/0.88	1.37/1.05	0.61/0.81	0.54/0.80
Hulian	0.63/0.63	0.97/0.77	0.77/0.76	0.73/0.74
Afarineh	0.94/0.88	1.26/0.82	0.70/0.69	0.62/0.52
Jelogir	0.78/0.82	1.09/0.87	0.83/0.81	0.82/0.80
Pay-e-Pol*	0.72/0.88	1.00/1.08	0.78/0.48	0.78/0.66
Hamidiyeh*	0.62/0.63	1.21/0.98	0.79/0.59	0.72/0.61

*Stations located downstream of the Karkheh dam
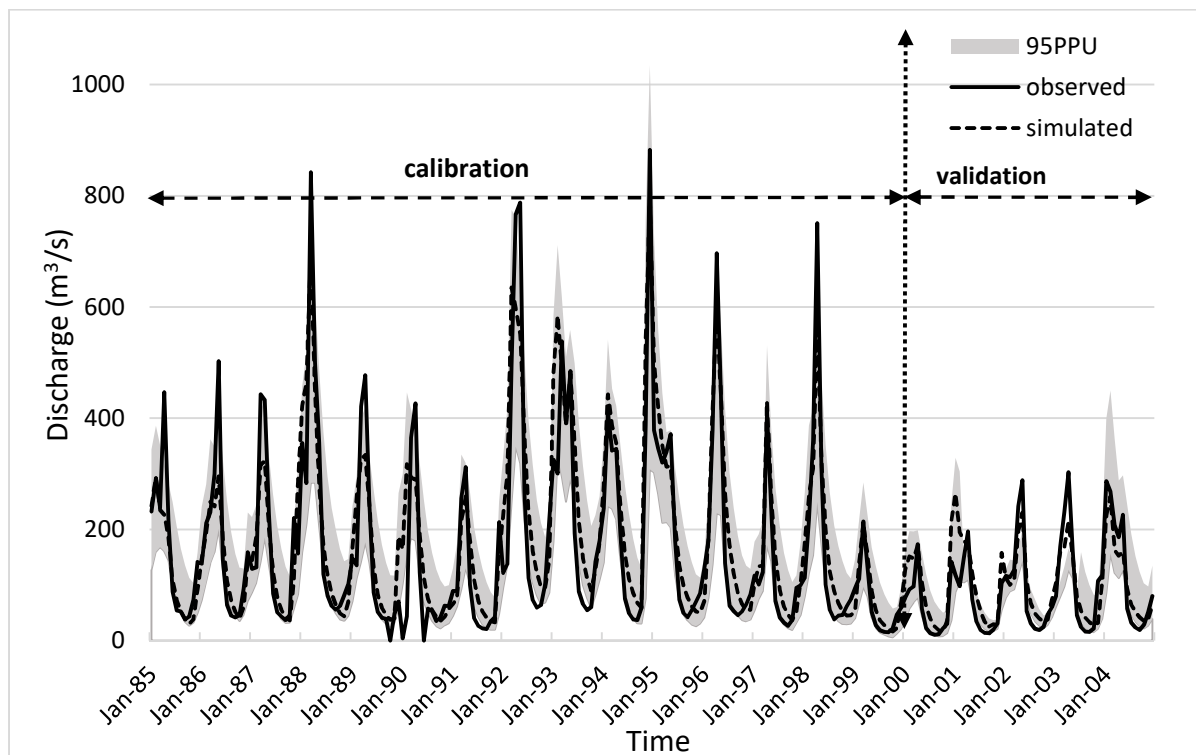
Figure 4 Observed and simulated monthly outflow (m^3/s) for calibration and validation at Jelogir station, upstream of the Karkheh dam
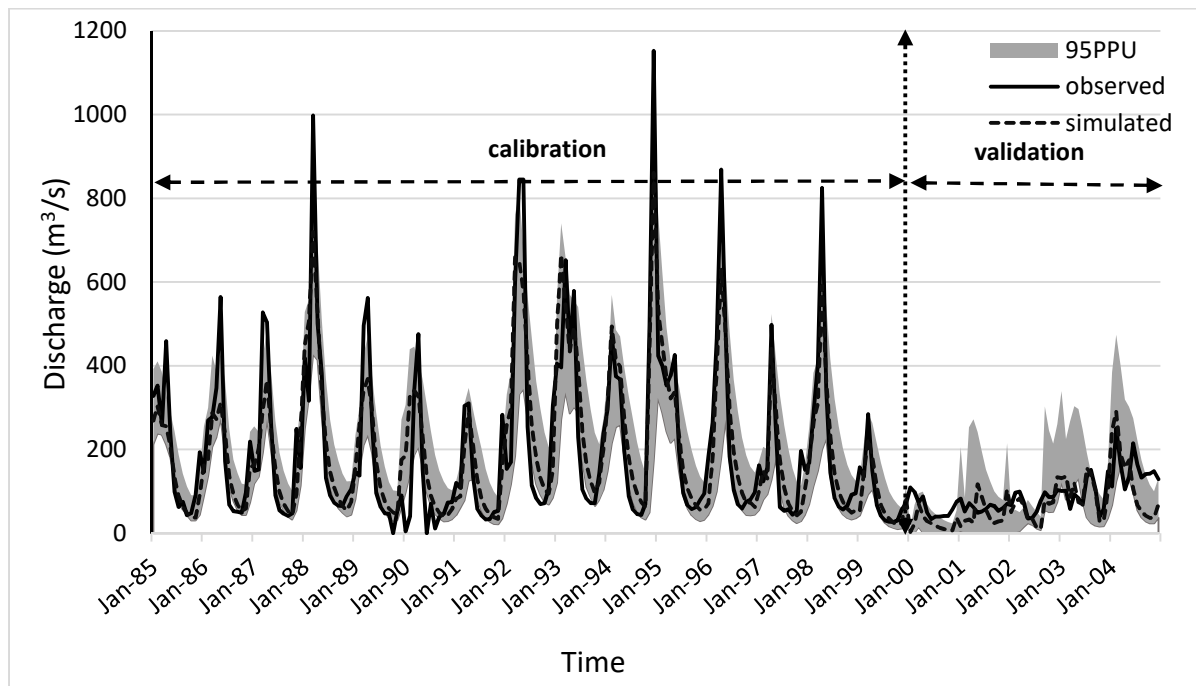
Figure 5 Similar to Figure 3, but for Pay-e-Pol station, downstream of the Karkheh dam
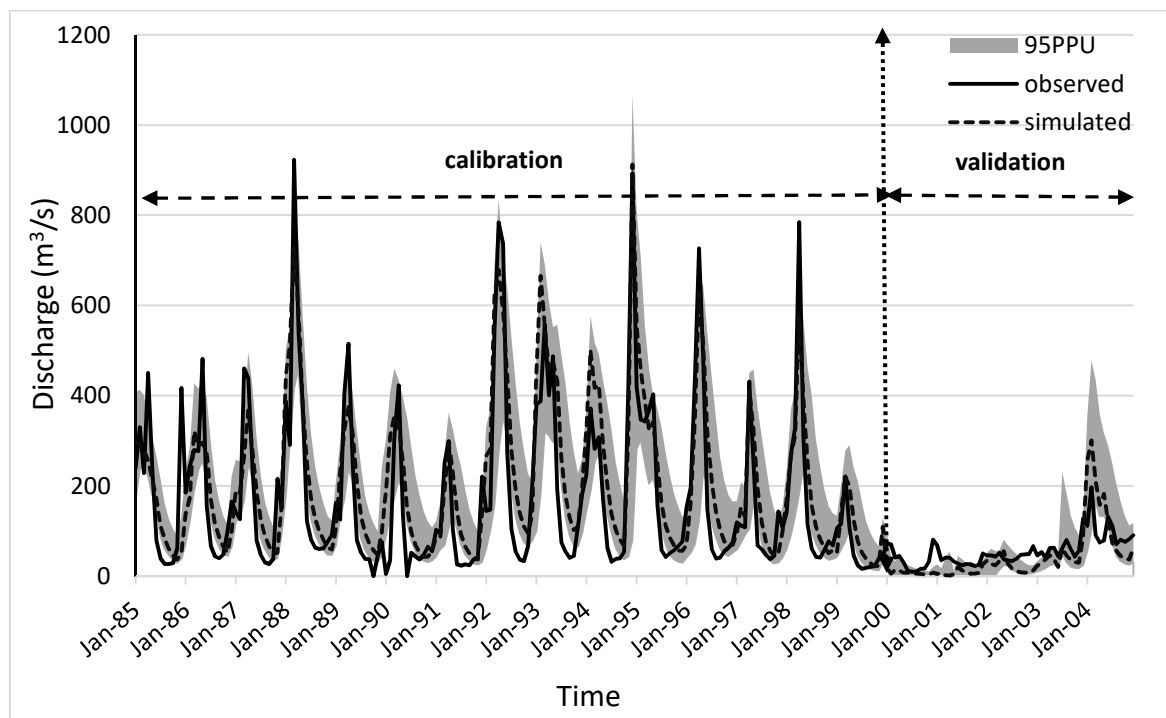
Figure 6 Similar to Figure 3, but for Hamidiyeh station, downstream of the dam

parameters are listed in Table 1. Thus, the Curve Number at moisture condition II (CN2) followed by the threshold depth of water in the shallow aquifer required for return flow to occur (GWQMN) and the base flow alpha factor (ALPHA_BF) tur out to be the most sensitive parameters.

Generally, ground water parameters are more sensitive than other parameters. SWAT model's performance was evaluated by the SUFI2 optimization procedure in SWAT-CUP, wherefore for the generation of the distribution of the output objective function several iterations with about 500 simulations each are performed.

The calibrated SWAT-model performs fine for the prediction of the monthly streamflow at the eight outlets, for both calibration and validation periods, as witnessed by the four statistical measures, P- and R-factor, NSE, R^2, all listed in Table 2.

The observed and simulated monthly streamflow time series hydrographs are shown in Figure 4 for the Jelogir station which is still located upstream of the Karkheh dam, in Figure 5 for Pay-e-Pol station located downstream of the Karkheh dam and in Figure 6 for the even further-downstream-located station Hamidiyeh for both the calibration- (1985-1999) and validation (2000-2004) periods.

The Karkheh dam started its operation in August 2002, i.e. during the SWAT- model's validation period, and its effect is to reduce the average annual streamflows in the two subsequent years 2003-2004 at these two downstream stations, compared with the pre-operation period (1985-2001). Nevertheless, for some months, especially, after the Karkheh dam became operational, with its impact on regulating the river discharge, validation is not very accurate, given that the dam's outflow characteristics were not always well documented for correct use in SWAT's reservoir module.

Figure 4 to 6 show that the monthly outflows of the main river before the Karkheh dam became operational have the same seasonal trends. However, once the dam became operational, notable reductions of the monthly discharge of more than 50% for the two stations Pay-e-Pol and Hamidiyeh downstream of the Karkheh dam (Figures 5 and 6), whereas the streamflow at the upstream station Jelogir (Figure 4) is barely affected. As a matter of fact, one of the reasons for the large reduction of the Karkheh river's discharge after year 2000 was the severe and extended drought lasting from 1999 to 2004 in Iran which coincided with the completion, filling and operation of the Karkheh Dam and which decreased the inflow and outflow of the Karkheh reservoir correspondingly. Besides, storing the river discharge in the Karkheh reservoir for the dry-season irrigation increased the evaporation rate from the reservoir which is nowadays the largest artificial lake in Iran.

4. Conclusions

Filling-in gaps in data-missing hydro-meteorological time series for use in deterministic hydrological models continues to be a challenge. In the present study, IDW and LRN, have been used successfully on daily time series of precipitation and temperature. The statistical measures applied indicate that IDW performs slightly better than LNR.

Subsequently, the SWAT-model was set up to simulate the hydrologic processes in the study region and to the predict water discharge at the 8 gauge stations in the basin. The model SWAT was calibrated using SWAT-CUP to match monthly river discharge for the period 1985-1999, with the 26 input parameters and then validated for the period of 2000-

2004. The sensitivity analysis by means of SWAT-CUP indicate that the 7 most sensitive parameters determining the streamflow in the watershed are CN2, GWQMN, ALPHA_BF, EPCO, SHALLST, RCHRG_DP and SOL_BD. Based on the four statistical measures, P-factor, R-factor, Nash Sutcliffe coefficient (NSE) and the coefficient of determination (R^2), a good performance of the SWAT model is indicated for both calibration and validation period. Finally, the impact of the Karkheh dam on the downstream river discharge was assessed. A signification reduction of the streamflow in the two following years 2003-2004 after the dam became operational, compared with that of the pre-operation period (1985-2001), is obtained.

References

[1] Beauchamp JJ, Dowing DJ, Railsback S.F. Comparison of regression and time-series methods for synthesizing missing streamflow records. **Water Resour. Bull.** 1989, **25**: 961-975.

[2] Ng WW, Panu U.S. Infilling missing daily precipitation data at multiple sites using the multivariate truncated normal distribution model for weather generation. **Water**. 2010; 8 pp.

[3] Kim JJ, Ryu JH. Quantifying a threshold of missing values for gap filling processes in daily precipitation series. **Water Resour. Manag**. 2015; **29 (11)**: 4173-4184.

[4] Kim JJ, Ryu JH. A heuristic gap filling method for daily precipitation series. **Water Resour. Manag.** 2016; **30 (11)**: 2275-2294.

[5] Arnold JG, Srinivasan R, Muttiah RR, Williams JR. Large area hydrologic modeling and assessment part I: model development. **J. Am. Water Resour. Assoc**. 1998; **34**: 73-89.

[6] Fereidoon M, Koch M. A sequential Simulation-Optimization Model for Water Allocation from the multi-Reservoir System in the Karkheh River Basin System, Iran. **11th International Conference on Hydroscience & Engineering**, ICHE Hamburg, 2014, September 28-October 2.

[7] Jamali S, Abrishamchi A, Madani K, Climate change and hydropower planning in the Middle East: implications for Iran's Karkheh hydropower systems. **J Energy Eng**. 2013; **139 (3)**: 153-160.

[8] Ahmad MD, Giordano M. The Karkheh River Basin: the food basket of Iran as aunder pressure. **Water International**. 2010; **35 (5)**: 522-544.

[9] Marjanizadeh S, de Fraiture C, Loiskandl W. Food and water scenarios for the Karkheh River Basin, Iran. **Water International**. 2010; **35 (4)**: 409-424.

[10] Masih I, Uhlenbrook S, Maskey S, Ahmad MD. Regionalization of a conceptual rainfall-runoff model based on similarity of the flow duration curve: A case study from the semi-arid Karkheh basin, Iran. **Journal of Hydrology**. 2010; **391**: 188-201.

[11] Salajegheh A, Razavizadeh S, Khorasani N, Hamidifar M, Salajegheh S. Land use changes and its effects on water quality (case study: Karkheh River Basin). **Journal of Environmental Studies**. 2011; **37 (58)**: 22-25.

[12] Teegavarapu RSV, Tufail M, Ormsbee L. Optimal functional forms for estimation of missing precipitation data. **J. Hydrol**. 2009; **374**: 106-115.

[13] Dumedah G, Coulibaly P. Evaluation of statistical methods for infilling missing values in high-resolution soil moisture data. **J. Hydrol**. 2011; **400 (1-2)**: 95-102.

[14] White KL, Chaubey I. Sensitivity analysis, calibration, and validation for a multisite and multivariable SWAT model. **J. Amer. Water Resour. Assoc**. 2005; **41**: 1077-1089.

[15] Neitsch SL, Arnold JC, Kiniry JR, Williams JR, King KW. **Soil and Water Assessment Tool Theoretical Documentation**. Version 2000. Texas Water Resources Institute, College Station, Texas, USA; 2002.

[16] Van Griensven A, Meixner T. Dealing with unidentifiable sources of uncertainty within environmental models, **The International Environmental Modelling and Software Society Conference**, Osnabrück, 2014, June 14-17.

[17] Koch M, Cherie N, SWAT Modeling of the Impact of future Climate Change on the Hydrology and the Water Resources in the Upper Blue Nile River Basin, Ethiopia. In: **6th Internat. Conf. on Water Resour. and Environ. Res**., Koblenz, 2013, June 3-7.

[18] Abbaspour KC, Johnson CA, van Genuchten MTh. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. **J. Vadose Zone**. 2004; **3 (4)**: 1340-1352.

[19] Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. **Trans. Am. Soc. Agric. Biol. Eng**. 2007; **50 (3)**: 885-900.