Helmholtz-Zentrum Potsdam
Deutsches GeoForschungsZentrum

Setti, S., Barik, K. K., Merz, B., Agarwal, A., Rathinasamy, M. (2022): Investigating the impact of calibration timescales on streamflow simulation, parameter sensitivity and model performance for Indian catchments. - Hydrological Sciences Journal - Journal des Sciences Hydrologiques, 67, 5, 661-675.

https://doi.org/10.1080/02626667.2022.2036340

Institional Repository GFZpublic: https://gfzpublic.gfz-potsdam.de/

Investigating the impact of calibration timescales on streamflow simulation, parameter sensitivity and model performance for Indian catchments

Sridhara Setti[1], Kamal Kumar Barik[1], Bruno Merz[2], Ankit Agarwal[3, *] and Maheswaran Rathinasamy[4]

[1]Department of Civil Engineering, Centurion University of Technology & Management CUTM, Bhubaneswar, Odisha, 752050 India

[2]GFZ German Research Centre for Geosciences, Section 4.4: Hydrology, Telegrafenberg, Potsdam, 14473 Germany

[3]Department of Hydrology, Indian Institute of Technology Roorkee, 247667 Uttarakhand, India

[4]Department of Civil Engineering, Indian Institute of Technology Hyderabad, Kandi, 502284 India

*Corresponding author: ankit.agarwal@hy.iitr.ac.in

Abstract

Hydrological model calibration is a quintessential step in model development and the time scale of calibration depends on the application. However, the implications of choice of time scale of calibration have not been explored extensively. Here, we evaluate the effect of the timescale of calibration on model sensitivity, best parameter ranges, and predictive uncertainty for three river basins using the SWAT model. Multiple models were setup for three different catchments from southern India. Our results showed that the sensitivity of the parameters, best parameter ranges, and model performance is conditioned on the timescale of calibration. The models calibrated at coarser time scales marginally outperformed the models calibrated at fine time scale in terms of Nash-Sutcliffe Efficiency and percentage bias. Transfer of parameters across scales (both from coarse to fine and fine to coarse) have general tendency to worsen the model performance in all three catchments, leaving for few exceptions.

Keywords – Timescale of streamflow calibration, SWAT Model, sensitivity analysis, transferability of parameters

1. Introduction

Hydrological models are essential tools for various purposes, such as understanding the water balance, estimating the impacts of anthropogenic activities, designing watershed management strategies, and flood warning and risk reduction (Wu et al., 2017; Sleziak et al., 2015; Zanon et al. 2010). Hydrological models are generally classified into black-box models, conceptual models, and physics-based models (Beven 2001). Notwithstanding the type of model, their application requires calibration, i.e. estimating the model parameters so that the model closely matches the behaviour of the real system it represents (Gupta et al. 1998). Some parameters can be determined through field measurements; however, most model parameters (particularly for conceptual and black-box models) need to be estimated through calibration. In most cases, this is done by adjusting the important model parameters so that simulated and observed streamflow agree sufficiently well. Calibration methods can be classified into trial-and-error and automatic procedures. The former involves numerous trial runs with different parameter values for reducing the error between simulation and observed data. Auto-calibration uses mathematical methods, such as optimisation, to find the optimal parameter set (Abbaspour et al., 2012). The trial-and-error method becomes highly cumbersome and complex when there are numerous parameters, and it is highly subjective. In this case, auto-calibration is more efficient and effective (Madsen et al., 2000; Getirana et al., 2010).

Generally, auto-calibration involves estimating the best parameter values for the sensitive parameters by minimising the objective function, which measures the closeness of the model to the observed data at a specified time and spatial scale. One of the crucial factors in model calibration is the temporal scale, i.e. the temporal resolution at which the simulations are compared with the observations. Using a conceptual model, Schake et al. (1996) suggested that the model should be calibrated at the computational timescale, i.e. at the timescale at which it is operated. However, models are frequently calibrated at coarser timescales owing to i) the lack of fine time scale data for calibration (for example, the daily streamflow data is not made available to the public for certain important rivers), ii) model driving input data may not be robust (for example, the climate model simulations which drive the hydrological models are less accurate at daily scales), iii) the simulations are often required at coarser timescales, e.g. monthly or yearly simulations are desired for planning studies and iv) calibration at coarser timescale is computationally less intensive (Sudheer et al. 2007, Wang et al. 2011). Some examples studies include Novotny and Stefan 2007; Lotz et al. 2017; Da Silva et al. 2018; Ang and Oeurng 2017 and Setti et al. 2017 where the model is calibrated at the coarser time scale, but the model simulation and result analysis is performed at a finer scale. This approach indirectly assumes that models mimic the process dynamics even at a smaller time step than the one they have been calibrated for. Even though this assumption is acceptable as long as the model is used simultaneously, the model results cannot be extended to other lower scales without investigation. SWAT model, a popular model, works at daily time steps, but a large fraction of studies does calibration at the monthly timescale (White and Chaubey, 2005, Adla et al., 2019, Lerat et al., 2020). Adla et al. (2019) identified more than 500 papers using SWAT that assume that a good performance at one timescale will translate into a similar performance at other timescales.

Even though several studies (Finnerty et al., 1997; Littlewood and Croke 2008; Wang et al., 2009; Cho et al., 2009; Remesan et al., 2010; Kavetski et al., 2011 and Reynold et al. 2018) investigated the influence of the timescale of the input data on model parameters and model performance. However, the effect of the calibration timescale remains poorly understood, as very few studies have investigated this effect. One of these studies includes Sudheer et al. (2007), who concluded that the model's performance could not be ensured at the finer timescale (such as daily) by calibrating at the monthly timescale. Troy et al. (2008) studied the impact of transferring parameters across scales using the VIC model and concluded that it is possible to calibrate at coarser time steps to save computational time. Daggupati et al. (2015) evaluated the parameter transfer across spatial and temporal scales for the West Lake Erie Basin andfound that transferring parameters from monthly to yearly and daily time steps performed well. On the other hand, Adla et al. (2019) reported that the SWAT model calibrated at the monthly scale failed to characterise the streamflow simulation at the daily time scale for the Punpun River Basin, India. The results from some of the prior studies (Duggupart et al., (2015), Adla et al., (2019) and Sudheer et al.(2007)) suggest that there is a deterioration in coarse to fine transition but these studies were based on one catchment and analysis considered only transfer across daily and monthly.

Apart from the above studies on the parameter transfer from one scale to another scale, researchers (Atkinson et al. 2002; Reusser et al. 2011; Herman et al. 2013 and Xi et al. 2017) have worked on investigating the parameter sensitivity with time and have shown model parameters sensitivities changes through time and provide a basis for the present study in understanding the how the parameters are scale dependent.

In this study, three river basins of various sizes and characteristics are considered in this study to address the following questions: 1) How does the timescale of calibration affect the sensitivity, the model parameters, and the streamflow prediction? (2) Can we transfer the parameters calibrated at one timescale to other timescales for the simulation? The answers to these questions can be very useful for regions where high temporal data is scarce (unavailability and quality).

To understand the impact of time scale on calibration and the implications of parameter transfer on model performance, we have considered the most widely used hydrological model, SWAT, which performs water budgeting daily, but in general, calibrated monthly or yearly time scale. Adla et al. (2019) report that most studies (around 50%) calibrate the model at a monthly scale and do not report the results of daily calibration and validation statistics. Therefore, most of the studies inherently assume the model performing well at coarse scale will perform well at finer scale also. Further, it is also interesting to understand how the transfer of parameters from finer scale to coarser scale will be applicable in situations where the fine scale streamflow data is only available for a certain time period. Therefore, to evaluate whether parameters obtained by calibration at one timescale can be transferred to other timescales, we applied the Soil and Water Assessment Tool (SWAT) model for three watersheds, namely Vamshadhara, Kagna, and Kharkai watersheds, India, at three different timescales (daily, monthly, and yearly) and developed different scenarios to study the changes in parameters across time scales.

2. Material and Methods

2.1 Study Area

We have selected three different river basins from the southern part of the Indian subcontinent for this study. These basins have been selected owing to the following reasons, i) past studies at these basins using the SWAT model have reported good performance in terms of streamflow simulation, ii) there is no significant impact on the hydrological systems in terms of water resources infrastructure and diversions upstream of gauging points, iii) the three selected basins similar hydroclimate gradient but different kind land use and human influence and iv) further, the size of the catchments are different.

The Vamshadhara River Basin, with an area of 10,448 km^2, is located between the Godavari and Mahanadi major river basins in Southeast India (Fig.1) and is contained by geographical co-ordinates of 18°20'59" N latitude and 84°07'59" E longitude. The Vamshadhara River originates in the Kalahandi district of Odisha state and flows around 254 km before joining the Bay of Bengal at Kalingapatnam, Andhra Pradesh. VRB receives an average annual rainfall of 1400 mm, with 75% of the rainfall falling during the south-west monsoon months of June to

September. 78% of its area is covered by forest and 20 % by irrigated crops. Fig.2 shows the topography, the spatial distribution of land use and soils, and the location of the Kasinagar gauge station used for calibration and validation. The river basin is covered majorly by clay soils (67%) and loam soils (34%).

The Kharkai Watershed (KW) is located in the Subarnarekha River Basin near Jamshedpur town. It has 6,267 km^2 area extending between 21°59'56" N latitude and 86°25'29" E longitude. The Kharkai River originates at Gobardhansahi village of Mayurbhanj and flows through Jharkhand and Odisha states. The average annual rainfall is 1400 mm, of which 79% is received in the monsoon months. Major land use/land cover classes are forest (41%) and irrigated crops (57%). The watershed is covered by clay soils (43%), loam soils (38%), and sandy-clay-loam soils (19%). The calibration discharge gauge is Adityapur (Fig. 2).

The Kagna Watershed (KW), with an area of 1,909 km^2, is located in the Krishna River Basin and near Tanuru Mandal of Telangana state between 17°01'3" N latitude and 77°57'30" E longitude. It receives an annual rainfall of around 800 mm, with 80% during the monsoon months. 73% of the Kagna Watershed is covered by clay soils and 27% by clay-loam. The majority of land use/land cover classes are irrigated crops (82%) and forest (12%). We used the Lewangi gauge discharge data for model calibration and validation (Fig. 2).

Spatial data, i.e. Digital Elevation Model, land use/land cover, soil properties, and temporal data, i.e. gauge discharge and meteorological data, were used as input for the SWAT model and detailed source of each dataset is given in Table S1 (in supplementary material).

2.2 SWAT Model

SWAT is a continuous and semi-distributed hydrological model. It was developed by the Agricultural Research Service of the United States Department of Agriculture (USDA-ARS) (Arnold et al. 1998 & 2012) to assist water resources management and planning. SWAT requires input data like weather, topographical, soil properties, land use, and land cover for simulating surface runoff and sediment yield of the river basin at daily time steps following the water balance equation (Neitschet et al., 2002):

$$\mathrm{SW}_t = \mathrm{SW}_O + \sum_{i=1}^{t}(R_{day,i} - Q_{surf,i} - E_{a,i} - W_{seep,i} - Q_{gw,i}) \quad (1)$$

where SWt denotes the final soil water content (in mm), SW_o represent the initial soil water content on day i (in mm H_2O), and t represents a simulation period (in days). $Q_{sur,i}$, $R_{day,i}$ and $E_{a,i}$ denote the amount of surface runoff, precipitation, and evapotranspiration (in mm H_2O) on any day i, respectively. $Q_{gw,i}$ and $W_{seep,i}$ represent the amount of groundwater return flow and percolation on a day i (in mm H_2O), respectively.

SWAT model simulates canopy storage, infiltration, surface runoff, lateral subsurface flow, percolation, groundwater flow, soil water content, evapotranspiration, pond recharge, snowmelt, and transmission losses (Arnold et al., 2012; Spruill et al., 2000). Surface runoff can be modelled by either i) CREAMS runoff model (Knisel, 1980), which includes SCS curve number method, ii) Green and Ampt infiltration method and iii) the modified rational formula method. In this study, we derived surface runoff from the Soil Conservation Service (SCS) - Curve Number (CN) method:

$$S = 254\left(\frac{CN}{100}\right) - 1 \quad (2)$$

where S represents the retention parameter (in mm), and CN represents the Curve Number that depends on the soil, land use, and soil moisture conditions. Since the CN method is an infiltration loss model that does not account for evaporation and evapotranspiration, its use was restricted to modelling storm losses. However, the parameter S should be linked with the soil moisture accounting module for continuous streamflow simulation. The SWAT model links S with available soil moisture and for using the CN method for continuous simulation.

Manning's formula is used for estimating the watershed time of concentration (considering both overland and channel flow). SWAT uses a storage routing technique to model the percolation and flows through each soil layer in the root zone (Spruill et al., 2000), also calculates lateral subsurface flow and recharge beyond the lowest soil layer. In SWAT, the plant growth model used for estimating water and nutrients uptake from the root zone, transpiration, and bio-mass production (Arnold et al. 2012). SWAT provides three methods for estimating the Potential Evapotranspiration (PET): Hargreaves method (Hargreaves and Samani, 1985), Priestley Taylor method (Priestley and Taylor, 1972), and Penman-Monteith method (Monteith, 1965). In this study, we applied the Penman-Monteith method. Groundwater flow is estimated by routing the shallow aquifer storage to the streams (Arnold et al., 1993). We used the QGIS interface (QGIS 2.6.1) and SWAT 2012 to process the input data and run the model, respectively. A detailed description of the different hydrological processes and the corresponding model parameters are shown in Table S2 (supplementary material).

2.3 Model Setup

The streams and the sub-basin boundaries were delineated by adopting a minimum sub-basin area of 100 Sq. Km. Each sub-basin was further disaggregated into several Hydrological Response Units (HRUs) based on a unique combination of soil properties, land use/land cover, and slope (SWAT 2012). This resulted in 26 sub-basins and 1460 HRUs, 20 sub-basins and 500 HRUs, 14 sub-basins and 300 HRUs for Vamsadhara, Kharkai, and Kagna watersheds, respectively. Based on the cropping pattern in each of these watersheds, the model's management options were modified accordingly. Apart from minor variations, the main crops in the three watersheds were paddy and pulses during the summer and winter cropping seasons, respectively.

2.4 Performance Measures

We used the coefficient of correlation (Willmott, 1981), the Nash-Sutcliffe efficiency coefficient (Nash and Sutcliffe, 1970), and Percent bias (Pbias) (Yapo et al., (1996) to evaluate the streamflow simulations:

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{sim}\right)^2}{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_{mean}^{obs}\right)^2}\right] \tag{3}$$

$$R^2 = \left(\frac{\sum_{i=1}^{n}\left(y_i^{obs}-y_{mean}^{obs}\right)\left(y_i^{sim}-y_{mean}^{sim}\right)}{\sqrt{\sum_{i=1}^{n}\left(y_i^{obs}-y_{mean}^{obs}\right)^2\left(y_i^{sim}-y_{mean}^{sim}\right)^2}}\right) \tag{4}$$

$$PBIA\,S = \left[\frac{\sum_{i=1}^{n}\left(y_i^{obs}-y_i^{sim}\right)\mathrm{X}100}{\sum_{i=1}^{n}\left(y_i^{obs}\right)}\right] \tag{5}$$

Here, Y_i^{obs} and Y_i^{sim} denote the ith observed and simulated data, respectively, Y_{mean}^{obs} is the mean of the observed data during the simulation period, and n denotes the number of observations. For both criteria, values close to 0 indicate unsatisfactory model performance. If NSE and R^2 are close to 1, then the model is ideal. The ideal value is zero for Pbias; if Pbias is negative, the simulated streamflow overestimates the observed streamflow and vice versa, the model is an underestimation. These performance measures were applied at different time steps daily, monthly, or annual; thus, allowing calibration across temporal scales.

2.5 Sensitivity Analysis

Sensitivity analysis is used to identify the most important model parameters, reducing the number of parameters used in the calibration process (Arnold et al., 2012). It can be divided into two types: local and global sensitivity analysis (Abbaspour, 2015). Local sensitive analysis (OAT - One At a Time) is limited as it does not consider the simultaneous variation of parameters and thus cannot represent interactions between parameters. Hence, we used a global sensitivity analysis (AAT - All At a Time) which typically leads to more robust results (Arnold et al., 2012). Here we estimate the parameter sensitivity is using the multiple regression system, which regresses the Latin hypercube generated parameters against the objective function values as shown in Eq. (6)

$$h = a + \sum_{j=1}^{m} b_j \gamma_j \qquad (6)$$

where h is the objective function value (in this study NSE (Eq.3) is used), γ_j indicates the parameter vector, a is the regression constant, and b is the regression coefficient vector. The sensitivities obtained to estimate the average changes in the objective function resulting from changes in each parameter, while all other parameters are changing (Abbaspour et al., 2015). It is important to note that the above method does not consider the interaction between the model parameters such as those possible in the SOBOL method, which can be explored in future studies.

A t-test is used for estimating the relative significance of the parameter γ_{i} (Abbaspour et al., 2007). The t-statistic is obtained by dividing the coefficient of a parameter by its standard error. It measures the precision with which the regression coefficient is measured. If the coefficient value is large compared to its standard error, the value will be different from zero, and the parameter is sensitive (Abbaspour et al., 2015). The smaller the p-values and the larger the t-test absolute values, the more sensitive is the parameter.

2.6 Model Calibration and Validation

We used an auto-calibration procedure by applying the Sequential Uncertainty Fitting Algorithm Version 2 (SUFI-2) of the SWAT-CUP (Calibration and Uncertainty Programmes) software. The procedure accounts for interactions between calibration parameters, as it assesses the performance of parameter sets and not the performance of individual parameters during the calibration. The SUFI-2 procedure results in the best range of parameters rather than individual values (Abbaspour et al., 2004). The propagation of the uncertainties in the parameters leads to the uncertainties in the model output (here streamflow), expressed as 95% probability distributions (95-PPU). The 95-PPU has estimated at the 2.5% and 97.5% levels of the output variable obtained from the *n* simulations using the *n* set of parameters. The resulting 95-PPU envelop is the output obtained from the SUFI approach. The P-factor and R-factor measure the fit between the observed data and the output from SUFI (expressed in terms of 95-PPU).

The P-factor indicates the fraction of the observed data falling within the 95% confidence limits. For instance, a P-factor of 1 indicates that 100% of the observed data fall within the 95% band. The R-factor indicates the average width of the 95-PPU band. It is calculated as the average 95-PPU thickness divided by the standard deviation of the corresponding observed variable (Abbaspour et al., 2015). Theoretically, the P-factor ranges from 0 to 1, and R-factor ranges from 0 to ∞. A simulation with P-factor =1 and R-factor=0 exactly corresponds to the observed data. The extent to which the values of the P-factor and R-factor are near to these numbers will help us understand the calibration level. The larger value of the P-factor will be achieved at the cost of the R-factor. While we would like to capture the observed data within the 95-PPU, we would like to have a small uncertainty envelop; therefore, a compromise between the two is required. The SUFI algorithm performs several iterations, and in each iteration, the parameter ranges get narrower, zooming on the region of the parameter space, where the previous iteration obtained good results.

As a consequence, the 95-PPU becomes smaller, resulting in a smaller P-factor and R-factor. Generally, R-factor values near 1.5 are considered satisfactory (Abbaspour, 2011). When satisfactory P- and R-factors are obtained, the final parameter ranges are defined as the best ranges (for details, see Abbaspour et al., 2007). This study has chosen the Nash Sutcliffe Efficiency (Yang et al., 2016) as the objective function (for Eq. 6) for calibrating the models and R^2 and Pbias for model assessment calibrated models.

Since discharge data was not available for a common time window at all three catchments, we have adopted different windows for calibration and validation for the different catchments (Table 1). The time windows are selected based on the availability of continuous discharge data. We have divided the entire time period into three parts: warmup period, calibration period, and validation period for three watersheds based on available discharge data (as shown in Table 1).

2.7 Impact of Timescale on Calibration and Parameter Transfer Scenarios

To the answer the two research questions raised, we developed nine parameter transfer scenarios (D, M, Y, DM, DY, MD, MY, YD, YM) as shown in Table.2. The SWAT model was calibrated using daily, monthly and yearly streamflow data, denoted as D, M, and Y, respectively. Then these calibrated models were validated at these three timescales, creating nine scenarios as shown in Table.2. We followed two-letter notations for each scenario, wherein the first letters denote the scale of calibration and the second letter scale of validation. For example, DY denotes a Model calibrated daily scale and applied at a yearly timescale for validation.

3. Results and Discussion

The results and subsequent discussion in Section 3.1 to 3.3 would answer the question, how does the timescale of calibration affect the sensitivity analysis, the model parameters, and the streamflow prediction?

3.1 Parameter Sensitivity

We analysed the sensitivity of the model response, i.e. catchment runoff, to variations in 18 parameters (Table S2). These parameters were selected based on previous literature (Abbaspour et al., 2017; Narasimlu et al., 2015; Murthy et al., 2014; Abbaspour et al., 2007). The initial parameter range was obtained from the SWAT database and is consistent with the physics of the process modelled. Using the global sensitivity analysis approach, we determined the sensitivity of these parameters and the corresponding ranks when calibrated at the three timescales for each watershed (Table 3).

Significant variations in the sensitive parameter ranking at each timescale and for each watershed can be observed. The ranks are highly dependent on the calibration timescale. For example, soil evaporative demand (ESCO) was a dominant parameter for all watersheds at the yearly and monthly timescale but not at the daily scale. In contrast, the alpha base flow factor (ALPHA_BF) was one of the sensitive parameters in all watersheds at daily and monthly timescales but not at the yearly timescale. There is also variation between watersheds. For instance, for the Vamshadara river basin, the curve number (CN2) was sensitive only at the daily and yearly timescales.

On the other hand, for the Kharkai basin, CN2 was sensitive at all timescales. Overall, the results suggest a significant impact of the calibration timescale on the parameter sensitivity. However, there is no clear pattern emerging from the results for these three watersheds.

3.2 Best Parameter Range

The calibration using the SUFI-2 optimisation algorithm starts with a wide parameter range and ends with a narrower range, i.e. the best parameter range. We have used 18 parameters for calibration of the model. For the first iteration, the parameter uncertainty is expressed by a uniform distribution. The optimal parameter value and the best parameter ranges resulting from the calibration at different timescales varied significantly (Figure 3). Analysis of some of the parameters which directly affect the water balance is given below.

ESCO

For the Vamsadhara river basin, the ESCO values best parameter value range between 0-0.5. ESCO is the coefficient can be used to alter the depth distribution which is linked with the evaporative demand. Lower the values of ESCO, the deeper layers can contribute to the evaporation resulting in more evaporation and a decrease in stream flow. From the results, the best parameter values were lower for daily and higher for monthly and yearly time scales, indicating the fine scale model allows evaporation from deeper levels of soils than the models at monthly and yearly scales.

In the Kharkai river basin case, there is a significant difference in the best parameter and the range of ESCO values. The daily scale values are close to 1, indicating lower evaporative demand, and the monthly and yearly scale values were close to 0.2 indicating higher demand. For the Kagna river basin, daily scales values are higher than the coarser time scales, indicating the lower demand.

The difference in the ESCO values can be attributed to the idea that when we calibrate a hydrological model at coarse time steps, say monthly or annual, the model only needs to reproduce the total streamflow correctly and overall water balance. This would not give importance to variations of the other processes, such as evapotranspiration. Further, the difference in the pattern across the catchments can be attributed to the corresponding dominant land use and land cover. For example, the Vamshadara river basin covered by forests (80%)

allows higher evaporative demand from deeper soil layers. In contrast, Kagna is dominantly covered by agricultural land (75%) allows lower evaporative demand from the deep soil layers.

CN2

CN2 is an important parameter as they directly control the amount of excess runoff generated and its travel through the system. The values in Figure 3b are represented in the percentage increase or decrease concerning the initial CN value. For example, 0.1 indicates a 10% increase in the CN in comparison with the initial value. In the Vamsadhara watershed, the values are in the low range (-0.1 to 0.02) at daily calibration, 0.0-0.07 at monthly and -0.02 to 0.01 at yearly calibration. The decrease in the CN2 value from the default value shows that the model allowed more surface runoff. In the case of Kharkai, the range of CN2 is similar at daily, monthly, and yearly calibrations, and the range is positive, indicating the model was underestimating the runoff at all three windows. For the Kagna watershed, the range of CN2 is similar for both monthly and yearly calibration; however, the range of CN2 is different and negative in the daily calibration, which indicates that the model overestimates the surface runoff at the daily scale.

ALPHA-BF

The ALPHA-BF has a smaller value for the Vamsadhara river basin than in the monthly and yearly calibration, indicating a quicker baseflow recession at the coarse scale than the daily calibrated model. For the Kharkai watershed, daily and yearly calibrations have a similar range of parameters, and the values are higher, indicating a quick recession. Still, at the monthly calibration, the range is 0-0.5, indicating the slow movement of baseflow and sustained flow in the river. For the Kagna watershed, the ALPHA-BF is sensitive at daily, monthly, and yearly calibration but has a dissimilar range indicate the baseflow movement varies at all three time scales.

The results discussed here are directly dependent on the choice of objective function used for calibration and sensitivity analysis. If one was to use NSE on box-cox transformed streamflow time series to give equal weightage to high and low flows or any another statistical metric, results may become quite different.

3.3 Performance of Models calibrated at different Timescales

The SWAT model was individually calibrated for the three watersheds and the three timescales. The results for the calibration periods are summarised in Table 4 (performance measures) and Figures 4-6 (discharge time series). In general, good to very good results were obtained at all timescales.

The model performance, quantified by R^2 and NSE, improved at coarser timescales for all three basins. It is essential to understand that even while the model is calibrated at coarser scales, SWAT simulates the flow at a daily time step. The daily values are then time-averaged to monthly and annual values, respectively. The improvement can be attributed to this time-averaging, as overestimations may compensate for underestimations and vice versa. Time delay errors at the daily timescale do typically not play a role at monthly and annual timescales (Adla et al., 2019).

Another observation is that the model performance increased with increasing catchment area for all calibration timescales. For instance, NSE at the daily timescale increased from 0.60 for the Kagna watershed (1,902 km^2) to 0.63 for the Kharkai watershed (6,267 km^2) and 0.75 for Vamsadhara watershed (10,448 km^2). A similar conclusion was drawn by Poncelet et al. (2017) and Merz et al. (2009). Using conceptual lumped models on hundreds of catchments in Europe, they found an increase in modeling efficiency with increasing catchment size. Based on the results from a lumped data-driven model, Maheswaran and Khosa (2012) showed that the nonlinearity and complexity of the catchment processes are lower for a larger catchment due to damping effects.

Interestingly, the improvement in model performance with spatial scale has been observed for very different modeling concepts, from lumped data-based (Maheswaran and Khosa, 2012) through lumped conceptual-model-based (Poncelet et al., 2017, Merz et al., 2009) to semi-distributed process-based (our study). However, when we compare the uncertainty (P- and R-factors) in the simulation across the three basins, it is observed that the uncertainty is lower for the Kagna watershed (smallest area, less variability in land use) and higher for the Vamsadhara river basin (largest area and higher variations in land use). The uncertainty in streamflow simulation seems to be a function of the catchment size and the variability in soil, land use land cover, and topography. Hence, although larger catchments tend to have good model performance, they still can show large uncertainties in their estimations.

The P-factor increases for all three catchments from the finer to the coarser times scales. This is explained by the smoother variation of streamflow at coarser timescales; hence, it is easier to capture the variations within the 95-PPU. R-factor values, which represent the thickness of the 95 PPU curves and generally lower values, are desired. The pattern of variation of the R-factor concerning the calibration scale is similar for the three basins considered. For example, for all the basins, the R-value is highest for daily and lowest for monthly. This could probably be due to the higher levels of uncertainty at the daily scale calibration.

The calibration at different timescales resulted in different sensitive parameters and best parameter ranges. Overall, the model performance in terms of R^2, NSE, P-factor, and R-factor improves when models are calibrated and validated at coarser timescales. This is due to the smaller streamflow variability due to time averaging and the fact that model errors tend to cancel out each other at coarser timescales.

3.4 Effect of Pparameter Transfer across Timescales

The influence of transferring the best calibration parameter set across timescales is summarized in Table 5 for the nine transfer scenarios. The results and subsequent discussion would answer the question, can we transfer the parameters calibrated at one timescale to other timescales for the simulation?

Self-validation of the models

When the models are validated at the timescale for which they have been calibrated, the validation results (Table 5) are close to the calibration results (Table 4) for all basins. This indicates that the models are calibrated adequately.

From finer to coarser timescales

Referring to Table 5, calibrating the model at fine scale and validating at the coarser scale resulted in deterioration of the model results compared to the model calibrated and validated at coarser scale. For example, in Vamshadara River Basin, the NSE for calibration and validation was found to be 0.91 and 0.72, respectively; however, NSE for the MY and DY scenarios was found to be 0.24 and 0.44, respectively. Similar behaviour was observed other two river basins. It is interesting to note that transfers from daily to monthly have better performance than the transfer from daily to yearly. For example, in the Vamshadara basin, DM scenarios yielded NSE=0.71, whereas DY produced results with NSE=0.24. Overall, it is observed that the good performance of the fine scale calibrated models does not warrant the similar performance of the coarse scale.

One possible reason for this behavior could be arising from the choice of objective function used for calibration. In this study, we have used the widely used NSE as the objective function; however, Schaefli and Gupta (2007) caution that in the case of monthly timescale, a model that only captures the seasonal features but not the small fluctuations will still have good NSE values, however, for predictions at the daily timescale, this (high) value will be misleading. Lerat et al. (2020), based on their study using four different objective functions, found that the performance of monthly scale models at the daily time step is a function of the objective function used and reported in models using NSE is a loss of information. In another related study, Rathinasamy et al. (2014) emphasised the importance of model assessment and calibration using scale-wise decomposition of the observed discharge rather than using the single scale observation. From this, it is clear that the choice of the objective function will add another dimension of uncertainty not only in model performance, as shown by Sridhar et al. (2020) but also in the transfer of parameters.

From Coarser to finer timescales

Transferring parameters from the coarser to the finer scale (YD, MD, and YM) also reduced NSE in 8 out of the 9 cases considering all the basins; however, comparatively better results were obtained. For example, in the Vamsadhara river basin, when the model was calibrated at the monthly scale and applied at the daily scale, only a negligible difference (NSE from 0.56(DD) to 0.46(MD)) in performance was obtained. Further, the scenario YD led to a comparatively better result than MD for this catchment. For the Kharkai basin, the loss in performance when transferring parameters from the coarser to the finer scale is relatively small, with values between -0-01 to -0.05. Surprisingly, for Kagna, the smallest catchment, YD and MD scenarios yielded very poor results (NSE for YD: 0.15 and MD: 0.31) than the DD scenario (NSE: 0.55).

Interestingly, for the Kharkai and Vamshadara basins, the application of parameters from the yearly model produced better results than those obtained using the monthly model. We analysed the parameter ranking, and the parameter ranges for the Kharkai basin to understand this effect. We computed the correlation between the parameter ranks of the timescales (shown in Table 3), which is 0.37 between the yearly and daily timescale and 0.10 between the monthly and daily scale for Kharkai. Further, from Figure 3, we observe that for CN2, ALPHA_BF, GWQMN, the parameter ranges were closer between the daily and yearly scales than between the daily and monthly scales.

Figure 7 shows the hydrographs for the scenarios MD and YD for the Vamsadhara River Basin. MD underestimates the peak values, and there is a lag in the peak runoff. YD captures better the timing and magnitude of the peaks. Similar observation can be seen in the other two basins from Figures S1 and S2. One possible reason for this difference might stem up from the study by Kumarasamy and Belmont (2018), wherein the authors investigate the parameter sensitivity to the periods and scales using wavelet coherence analysis. In that study, they show that certain parameters influence only a specific scale, and other than that timescale, there is no impact of that parameter.

The transfer of parameters from finer to coarser scales and the parameter transfer from coarser to finer scales mostly aggravated model performance. The lower performance of the coarser scales model at finer scales could be attributed to its inability to capture the variability in the streamflow. When the monthly scale calibrated models were used for simulating the flow at a daily time step, the results in terms of NSE were lower or similar compared to daily timescale calibration for all three basins. When yearly scale models were used to generate the flow daily, the results were better than the MD (monthly-daily) transfer but were closer to the daily calibration results.

It is important to note that a large percentage of SWAT modeling studies do not report the model performance on a daily scale when they have calibrated at monthly or yearly scales. Our results suggest that a good model performance at a coarse timescale may not ensure good performance at smaller timescales, and therefore, caution must be exercised in such cases. The results from this study strengthen the understanding provided by Adla et al. (2019) based on one river basin. Since our study is limited to three catchments and SWAT model, further studies can be directed towards understanding the transferability of parameters from one scale to another as a function of several factors like catchment size and its complexity, spatial variability of rainfall, model complexity, and also the objective function used for calibration. For a more generalised understanding of the transferability of parameters across timescales, particularly for detailed distributed models like SWAT, studies along similar lines must be conducted for hundreds of catchments of varying sizes and characteristics.

4. Conclusions

This paper investigated the effect of the timescale of calibration on hydrologic model calibration results, sensitivity analysis, and parameter uncertainty using the SWAT model for three catchments, Vamshadara, Kharkai, and Kagna, in India. The sensitivity of the model parameters and the best parameter range varies for different calibration timescales. Therefore, the decision about the timescale of calibration has implications for the sensitivity analysis stage in the hydrologic model calibration. Finally, the model performance was higher for the coarser scale models than the finer scale models.

A SWAT model, which has been calibrated at a finer timescale, achieves lower model performance at coarser timescales when compared to a model calibrated directly at the coarse scales. However, the reduction in performance seems to be modest, with a mean NSE reduction of 0.04 for our catchments. In contrast, when the parameters were transferred from the coarser to the finer timescales, the performance declined in almost all cases. The decline was particularly substantial for the Kagna river basin, i.e. the smallest catchment. To understand whether there are systematic influences of catchment size and other characteristics on the gain or loss of performance during the parameter transfer would require similar studies with hundreds of catchments.

Overall, these observations indicate that careful attention must be exercised while assuming the validity of coarser scale parameters for fine-scale simulation. The implicit assumption that such models mimic the process dynamics even at a smaller time step than the one they have been calibrated for may not be valid. Instead, our results suggest that the SWAT model should be calibrated at a time scale at which model results are required.

Acknowledgement

The financial assistance provided by SERB, Government of India under the scheme Early Career Research Award held by Dr. Maheswaran is gratefully acknowledged. AA and BM acknowledge the joint funding support from the University Grant Commission (UGC) and DAAD under the Indo-German Partnership in Higher Education (IGP) framework for the COPREPARE project.

References

Abbaspour, C.K. SWAT Calibrating and Uncertainty Programs; A User Manual; Eawag: Dübendorf, Switzerland, 2015.

Abbaspour, K.C., Vaghefi, S.A. & Srinivasan, R. 2017 A guideline for successful calibration and uncertainty analysis for soil and water assessment: A review of papers from the 2016 international SWAT conference. *Water (Switzerland)*. 10 (1).

Abbaspour, K.C., Yang, J., Maximov, I., Srinivasan, R. & Zobrist, J. 2006 Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *Journal of Hydrology*. 333 (2–4), 413–430.

Abbaspour, K.C.; Johnson, C.A.; Van Genuchten, M.T. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. Vadose Zone J. 2004, 3, 1340–1352. [CrossRef]

Abbaspour, K.C.; Johnson, C.A.; Van Genuchten, M.T. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. Vadose Zone J. 2004, 3, 1340–1352. [CrossRef].

Adla, S.; Tripathi, S.; Disse, M. Can We Calibrate a Daily Time-Step Hydrological Model Using Monthly Time-Step Discharge Data? *Water* **2019**, *11*, 1750.

Atkinson, S.E., Woods, R.A. and Sivapalan, M., 2002. Climate and landscape controls on water balance model complexity over changing timescales. Water Resources Research, 38(12), pp.50-1.

Arnold, J.G., Moriasi, D.N., Gassman, P.W. & Abbaspour, K.C. 2012 SWAT: Model use, calibration, and validation. *Biological Systems Engineering: Papers and Publications*. *Transactions of the ASABE* Vol. 55(4): 1491-1508

Arnold, J.G., Srinivasan, R., Muttiah, R.S. & Williams, J.R. 1998 Large area hydrologic modeling and assessment part I: model development. *J AWRA* 34(1)

Arnold, J.G.; Moriasi, D.N.; Gassman, P.W.; Abbaspour, K.C.; White, M.J.; Srinivasan, R.; Santhi, C.; Harmel, R.D.; Van Griensven, A.; Van Liew, M.W.; et al. SWAT: Model use, calibration, and validation. Trans. ASABE 2012, 55, 1491–1508. [CrossRef]

Arnold, J.G.; Moriasi, D.N.; Gassman, P.W.; Abbaspour, K.C.; White, M.J.; Srinivasan, R.; Santhi, C.; Harmel, R.D.; Van Griensven, A.; Van Liew, M.W.; et al. SWAT: Model use, calibration, and validation. Trans. ASABE 2012, 55, 1491–1508. [CrossRef]

Arnold, J.G.; Moriasi, D.N.; Gassman, P.W.; Abbaspour, K.C.; White, M.J.; Srinivasan, R.; Santhi, C.; Harmel, R.D.; Van Griensven, A.; Van Liew, M.W.; et al. SWAT: Model use, calibration, and validation. Trans. ASABE 2012, 55, 1491–1508. [CrossRef]

Beven, K.; Binley, A. The future of distributed models: Model calibration and uncertainty prediction. Hydrol. Process. 1992, 6, 279–298. [CrossRef]

Cho, J., Mostaghimi, S., Kang, M.S & Chun, J.A. 2009 Sensitivity to grid and time resolution of hydrology components of Dansat. *Transactions of the ASABE*. 52(4), 1121–1128.

Da Silva, V.P.R., Singh, V.P., de Souza, E.P., de Holanda, R.M. & Almeida, R.S.R. 2017 Simulation of streamflow and hydrological response to land-cover changes in a tropical river basin. *Catena*. 162 (December), 166–176.

Eckhardt, K. & Arnold, J.G. 2001 Automatic calibration of a distributed catchment model. *Journal of Hydrology*. 251 (1–2), 103–109.

Finnerty, B., Smith, M., Seo, D.J., Koren, V. & Moglen, G. 1997 Space–timescale sensitivity of the Sacramento model to radar– gauge precipitation inputs. *Journal of Hydrology*. 203, 21–38.

Getirana, A.C.V. 2010 Integrating spatial altimetry data into the automatic calibration of hydrological models. *Journal of Hydrology*. 387 (3–4), 244–255.

Gupta, H.V. & Sorooshian, S. & Yapo, P.O. 1999 Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel expert Calibration. J. Hydrol. Eng. 4(2), 135-143.

Gupta, H.V. & Sorooshian, S. 1998 Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resources Research*. 34 (4), 751–763.

Herman, J.D., Reed, P.M. and Wagener, T., 2013. Time-varying sensitivity analysis clarifies the effects of watershed model formulation on model behavior. Water Resources Research, 49(3), pp.1400-1414.

ISSN 0022-1694,Littlewood, I.G & Croke, B.F.W. 2008 Data time-step dependency of conceptual rainfall–streamflow model parameters: an empirical study with implications for regionalisationregionalization. *Hydrological Sciences Journal*. 53 (4), 685–695.

Julien Lerat, Mark Thyer, David McInerney, Dmitri Kavetski, Fitsum Woldemeskel, Christopher Pickett-Heaps, Daeyhok Shin, Paul Feikema. 2020 A robust approach for calibrating a daily rainfall-runoff model to monthly streamflow data, *Journal of Hydrology*, 591 125-129.

Kavetski, D., Fenicia, F. & Clark, M.P. 2011 Impact of temporal data resolution on parameter inference and model identification in conceptual hydrological modeling: Insights from an experimental catchment. *Water Resources Research*. 47 (5), 1–25.

Klemeš, V. Operational testing of hydrological simulation models. Hydrol. Sci. J. 1986, 31, 13–24. [CrossRef]

Knisel, W.G. CREAMS: A field scale model for Chemicals, Runoff, and Erosion from Agricultural Management Systems. In Conservation Research Report No. 26; US Dept. of Agriculture, Science and Education Administration: Washington, DC, USA, 1980.

Lenhart, T., Eckhardt, K., Fohrer, N. & Frede, H.G. 2002 Comparison of two different approaches of sensitivity analysis. *Physics and Chemistry of the Earth*. 27 (9–10), 645–654.

Lerat J, Mark Thyer, David McInerney, Dmitri Kavetski, Fitsum Woldemeskel, Christopher Pickett-Heaps, Daeyhok Shin, Paul Feikema. 2020 A robust approach for calibrating a daily rainfall-runoff model to monthly streamflow data, Journal of Hydrology, 591 125-129.

Liu, X.; Yang, M.; Meng, X.; Wen, F.; Sun, G. Assessing the Impact of Reservoir Parameters on Runoff in the Yalong River Basin using the SWAT Model. *Water* **2019**, *11*, 643.

Lotz, T., Opp, C. & He, X. 2018 Factors of runoff generation in the Dongting Lake basin based on a SWAT model and implications of recent land cover change. *Quaternary International*. 475, 54–62.

Madsen, H. 2000 Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *Journal of Hydrology*. 235 (3–4), 276–288.

Merz, R., J. Parajka, and G. Blöschl (2009), Scale effects in conceptual hydrological modeling, Water Resour. Res., 45, W09405, doi:10.1029/2009WR007872.

Moriasi, D.N., Arnold, J.G., Liew, M.W. Van. & Bingner, R.L. 2007 *Model Evaluation guidelines*. 50 (3), 885–900.

Murty, P.S., Pandey, A. & Suryavanshi, S. 2014 Application of semi-distributed hydrological model for basin-level water balance of the Ken basin of Central India. *Hydrological Processes*. 28 (13), 4119–4129.

Narsimlu, B., Gosain, A.K., Chahar, B.R. & Singh, S.K. 2015 SWAT Model Calibration and Uncertainty Analysis for Streamflow Prediction in the Kunwari River Basin, India, Using Sequential Uncertainty Fitting. *Environmental Processes*. 2 (1), 79–95.

Nash, J.E. & Sutcliffe, J. V. 1970 River Flow Forecasting through Conceptual Models Part I-A discussion of principles. *Journal of Hydrology*. 10, 282–290.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R. & Srinivasan, R. 2002 Soil and Water Assessment Tool User's Manual. *TWRI Report TR-192*. 412.

Novotny, E. V. & Stefan, H.G. 2007 Streamflow in Minnesota: Indicator of climate change. *Journal of Hydrology*. 334 (3–4), 319–333.

Poncelet, C., Merz, R., Merz, B., Parajka, J., Oudin, L., Andréassian, V., and Perrin, C.: Process-based interpretation of conceptual hydrological model performance using a multinational catchment set, Water Resour. Res., 53, 7247–7268, https://doi.org/10.1002/2016WR019991, 2017.

Remesan, R., Shamim, M. A., Ahmadi, A., & Han, D. 2010 Effect of data time interval on real-time flood forecasting. *Journal of Hydroinformatics.* 12(4), 396-407.

Reusser, D.E., Buytaert, W. and Zehe, E., 2011. Temporal dynamics of model parameter sensitivity for computationally expensive models with the Fourier amplitude sensitivity test. Water Resources Research, 47(7).

Reynolds, J.E., Halldin, S., Seibert, J. and Xu, C.Y., 2018. Definitions of climatological and discharge days: do they matter in hydrological modeling?. *Hydrological Sciences Journal*, *63*(5), 836-844.

Santhi, C., Arnold, J.G., Williams, J.R. & Dugas, W.A. 2001 Validation of the SWAT model on a large river basin with point and nonpoint sources. *Journal of the American Water Resources Association*. 37 (5), 1169–1188.

Schaake, J.C., Koren, V.I., Duan, Q.Y., Mitchell, K. & Chen, F. 1996 Simple water balance model for estimating runoff at different spatial and temporal scales. *Journal of Geophysics Research.* 101, 7461–7475.

Sleziak, Patrik & Szolgay, Jan & Hlavčová, K & Parajka, Juraj. (2015). Assessment of the effectiveness of calibrating a conceptual hydrological model in relation to selected catchment characteristics.

Spruill, C. A., Workman, S. R. & Taraba, J. L. 2013 Simulation of Daily and Monthly Stream Discharge From Small Watersheds Using the Swat Model. *Transactions of the ASAE*. 43 (6), 1431–1439.

Spruill, C.A.; Workman, S.R.; Taraba, J.L. Simulation of daily and monthly stream discharge from small watersheds using the SWAT model. Trans. ASAE 2000, 43, 1431–1439. [CrossRef]

Sudheer, K.P., Chaubey, I., Garg, V. & Migliaccio, K.W. 2007 Impact of the timescale of the calibration objective function on the performance of watershed models. J Hydrological Process.

Tang, Y., Reed, P., Wagener, T. & Van Werkhoven, K. 2007 Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. *Journal of Hydrological Earth System Science*. 11,793–817.

Troy, T.J., Wood, E.F. and Sheffield, J., 2008. An efficient calibration method for continental-scale land surface modeling. Water Resources Research, 44(9).

van Griensven, A., Meixner, T., Grunwald, S. & Bishop, T. 2006 A global sensitivity analysis tool for the parameters of multi-variable catchment models. *Journal of Hydrology*. 324 (1–4), 10–23.

Wang, Y.I., He, B. & Takase, K. 2009 Effects of temporal resolution on hydrological model parameters and its impact on prediction of river discharge / Effets de la résolution temporelle sur les paramètres d'un modèle hydrologique et impact sur la prévision de l'écoulement en rivière. *Hydrological Sciences Journal*. 54(5), 886-898.

White, K.L. & Chaubey, I. 2005 Sensitivity Analysis, Calibration, and Validations for a Multisite and Multivariable Swat Model. *Journal of the American Water Resources Association*. 41 (5), 1077–1089.

Willmott, C.J. 1981 On the validation of models. Physical Geography. 2 (2), 184–194.

Xie, H., Shen, Z., Chen, L., Qiu, J. and Dong, J., 2017. Time-varying sensitivity analysis of hydrologic and sediment parameters at multiple timescales: Implications for conservation practices. Science of the Total Environment, 598, pp.353-364.

Yap, P., Gupta, H.V., Sorooshian, S. 1996 Automatic calibration of conceptual rainfall-runoff sensitivity to calibration data models. J Hydrology 181(1-4):23–48.https://doi.org/10.1016/0022-1694(95)02918-4Zanon, F., Borga, M., Zoccatelli, D. & Marchi, L. 2010 Hydrological analysis of a flash flood across a climatic and geologic gradient: The September 18, 2007 event in Western Slovenia. Journal of Hydrology. 394 (1–2), 182–197.

Table 1: Time windows used for calibration and validation.

Watershed	Gauge Location	Warm-up Period	Calibration period (Validation period)
Vamsadhara	Kasinagar	1979-1981	1982-2000 (2001-2010)
Kharkai	Adityapur	1982-1984	1985-2000 (2001-2010)
Kagna	Jewangi	1979-1981	1982-1992 (1993-2000)

Table 2: Details about the different model scenarios generated in this study to evaluate the impact of the time scale of calibration for three watersheds.

Scenario	Remark
D	Signifies the self-validation at daily time scale.

M	Self-validation at monthly time scale.
Y	self-validation at yearly time scale
DM	A Model calibrate at daily time scale then validate at monthly time scale
DY	A Model calibrate at daily time scale then validate at yearly time scale
MD	A Model calibrate at monthly time scale then validate at daily time scale
MY	A Model calibrate at monthly time scale then validate at yearly time scale
YD	A Model calibrate at yearly time scale then validate at daily time scale
YM	A Model calibrate at yearly time scale then validate at monthly time scale

***Table* 3 Results of Sensitivity Parameters *rank* and p(t) values for Vamsadhara, Kharkai and Kagna watersheds corresponding to three time scales during 1979 – 2010.**

S.no	Parameter_Name	Vamsadhara						Kharkai						Kagna					
		Daily		Monthly		Yearly		Daily		Monthly		Yearly		Daily		Monthly		Yearly	
		R	p(t)	R	p(t)	R	p(t)	R	p(t)	R	p(t)	R	p(t)	R	p(t)	R	p(t)	R	p(t)
1	1:V__ESCO.hru	7	0.00(-4.61)	2	0.00(-5.81)	1	0.00(-29.44)	15	0.73(-0.34)	2	0.00(-2.97)	5	0(-8.53)	8	0(-3.11)	2	0(-6.51)	1	0(-9.74)
2	15:R__CN2.mgt	1	0.00(-18.84)	12	0.38(-0.89)	2	0.00(-16.27)	4	0(20.24)	1	0.00(30.99)	1	0(21.76)	7	0(-3.51)	3	0(-5.06)	2	0(4.58)
3	5:V__SLSUBBSN.hru	6	0.00(5.94)	10	0.14(1.49)	3	0.00(6.88)	6	0(-15.4)	5	0.00(-2.83)	7	0(-3.35)	4	0(-12.52)	17	0.96(-0.05)	8	0.37(0.9)
4	4:R__SOL_AWC(..).sol	14	0.04(2.09)	6	0.01(2.61)	4	0.00(6.04)	10	0.19(1.3)	17	0.39(0.86)	12	0.89(-0.14)	16	0.53(0.63)	8	0.09(1.7)	4	0(3.56)
5	3:V__HRU_SLP.hru	8	0.00(-3.92)	8	0.10(-1.64)	5	0.00(-6.08)	7	0(12.64)	3	0.00(3.35)	4	0(6.1)	6	0(6.24)	13	0.39(0.85)	13	0.63(0.48)
6	9:V__OV_N.hru	16	0.19(1.31)	13	0.45(-0.76)	6	0.04(-2.06)	5	0(-15.96)	9	0.69(0.4)	17	0.24(-1.19)	3	0(-12.65)	12	0.38(-0.88)	11	0.44(0.78)
7	7:V__EPCO.hru	15	0.10(1.66)	18	0.97(0.03)	7	0.08(1.75)	17	0.93(0.09)	15	0.1(1.66)	8	0.81(0.24)	15	0.49(0.7)	7	0.04(2.03)	5	0(3.54)
8	14:V__SURLAG.bsn	5	0.00(-9.97)	15	0.73(0.35)	8	0.41(0.83)	13	0.65(-0.45)	16	0.49(0.69)	15	0.83(0.21)	14	0.48(0.71)	16	0.94(-0.08)	18	0.97(-0.04)
9	12:V__CH_K2.rte	2	0.00(14.09)	7	0.01(-2.50)	9	0.39(-0.86)	2	0(-32.92)	12	0.00(-4.06)	3	0.35(0.93)	1	0(-37.65)	6	0.03(-2.19)	10	0.43(-0.78)
10	6:V__RCHRG_DP.gw	9	0.00(3.26)	1	0.00(-9.01)	10	0.32(1.04)	16	0.9(-0.13)	4	0.21(1.25)	11	0(5.24)	11	0.24(-1.17)	5	0.02(-2.35)	3	0(-3.9)
11	2:V__LAT_TTIME.hru	10	0.00(3.15)	16	0.81(-0.24)	11	0.21(1.25)	9	0.15(-1.43)	11	0.19(-1.32)	10	0.33(0.97)	13	0.34(0.95)	11	0.33(0.97)	17	0.97(-0.04)
12	16:V__ALPHA_BF.gw	3	0.00(-12.53)	3	0.00(5.69)	12	0.94(0.07)	1	0(46.91)	7	0.00(25.63)	2	0.19(1.32)	2	0(19.67)	1	0(19.74)	6	0.03(2.15)
13	10:V__GW_REVAP.gw	13	0.02(2.35)	5	0.01(2.80)	13	0.42(-0.81)	14	0.67(0.43)	6	0.4(-0.83)	13	0(-3.01)	9	0.1(-1.67)	9	0.1(-1.65)	16	0.85(0.19)
14	11:V__CH_N2.rte	4	0.00(11.68)	11	0.22(1.22)	14	0.33(-0.97)	3	0(-21.05)	13	0.00(-2.83)	6	0.44(-0.76)	5	0(-11.91)	4	0(-4.04)	7	0.1(-1.67)
15	18:V__GWQMN.gw	12	0.02(2.43)	4	0.00(3.95)	15	0.59(0.55)	11	0.3(1.05)	8	0.59(0.54)	16	0.23(-1.2)	10	0.2(-1.29)	15	0.92(0.1)	9	0.4(0.85)
16	13:R__SOL_K(..).sol	17	0.20(1.29)	9	0.13(1.53)	16	0.73(-0.35)	18	0.99(0.01)	18	0.45(0.76)	14	0.93(-0.09)	12	0.33(0.98)	10	0.17(1.36)	12	0.54(0.61)
17	17:V__GW_DELAY.gw	11	0.01(2.65)	17	0.94(0.08)	17	0.30(1.04)	12	0.59(0.54)	10	0.94(-0.07)	18	0.26(-1.14)	17	0.76(-0.3)	14	0.69(-0.39)	14	0.69(0.4)
18	8:V__REVAPMN.gw	18	0.86(0.17)	14	0.60(-0.52)	18	0.70(0.39)	8	0.13(1.51)	14	0.11(1.61)	9	0.71(-0.37)	18	0.87(0.17)	18	1(0)	15	0.79(-0.26)

Table 4: NSE, R2, P- and R-factor values for Vamsadhara, Kharkai and Kagna watersheds for the calibration period for three timescales.

Watershed /Timescale		Daily	Monthly	Yearly
Vamsadhara	R^2	0.77	0.9	0.92
	NSE	0.75	0.9	0.91
	Pbias	-31.4	-4.1	-1.7
	P-factor	0.49	0.82	0.95
	R-factor	0.47	0.57	0.91
Kharkai	R^2	0.66	0.86	0.84
	NSE	0.63	0.84	0.79
	Pbias	27.9	-3.1	-0.7
	P-factor	0.65	0.61	0.81
	R-factor	0.11	0.46	0.58
Kagna	R^2	0.65	0.82	0.79
	NSE	0.63	0.79	0.76
	Pbias	19.1	5.2	4.6
	P-factor	0.29	0.2	0.45
	R-factor	0.18	0.17	0.34

Table 5: R^2 and NSE for the self-validation and different transfer scenarios performed at three catchments. For example, M to D (scenario MD) indicates that the calibration was performed at the monthly time scale and the validation at the daily scale.

Watershed	Statistic metrics	Self-Validation (1993-2000)			Parameter values transfer from one-time scale to another time scale					
		DD	MM	YY	DM	DY	MD	MY	YD	YM
Vamsadhara	R2	0.63	0.77	0.74	0.78	0.65	0.47	0.64	0.49	0.67
	NSE	0.56	0.76	0.72	0.71	0.24	0.46	0.44	0.48	0.62
Kharkai	R2	0.75	0.87	0.78	0.88	0.78	0.57	0.78	0.75	0.89
	NSE	0.65	0.75	0.74	0.74	0.23	0.5	0.16	0.69	0.75
Kagna	R2	0.6	0.85	0.91	0.74	0.89	0.34	0.9	0.17	0.69
	NSE	0.55	0.76	0.72	0.58	0.44	0.31	0.72	0.15	0.6

Figure 1: Index Map showing the geographical location of the three catchments, namely Vamsadhara, Kharkai, and Kagna.

Figure 2: Topography (a), land use/land cover (b) and soil classes (c) for the three watersheds of Vamsadhara, Kharkai and Kagna, respectively.

Figure 3: Best parameter values and best parameter ranges resulting from the calibration at different time scales for Vamsadhara, Kharkai and Kagna watersheds using auto-calibration procedure by applying the Sequential Uncertainty Fitting Algorithm Version 2 (SUFI-2). Line and Dot represent the best parameter range and best-fitted values, respectively.

Figure 4: Stream flow simulation for Vamshadara watershed during the calibration period daily, monthly and yearly time scale (top to bottom) using the best parameter compared with the observed flow and the 95-PPU bands shown in green. For clarity, the results at the daily scale are shown only for a small time window.

Figure 5: Stream flow simulation for Kharkai watershed during the calibration period daily, monthly and yearly time scale (top to bottom) using the best parameter compared with the observed flow and the 95-PPU bands shown in green. For clarity, the results at the daily scale are shown only for a small time window.

Figure 6: Stream flow simulation for Kagna watershed during the calibration period daily, monthly and yearly time scale (top to bottom) using the best parameter compared with the observed flow and the 95-PPU bands shown in green. For clarity, the results at the daily scale are shown only for a small time window.

Figure 7: Comparison of the daily runoff generated from parameter transfer scenarios (a) Monthly to Daily scale (MD) and (b) Yearly to Daily scale (YD) scenarios with the observed runoff for the Vamsadhara River basin.