Impact of different types of meteorological data inputs on predicted hydrological and erosive responses to projected land use changes

Suman Bhattacharyya and Joy Sanyal*

Department of Geography, Presidency University, 86/1 College Street, Kolkata 700 073, India.
*Corresponding author. e-mail: joy.geog@presiuniv.ac.in

Hydrological responses to land use/land cover (LULC) changes are complex in nature and tend to have an impact on the hydrological cycle, affecting the livelihood of the inhabitants. Rainfall–runoff models, such as the Soil and Water Assessment Tool, were used in the past to unravel the interactions between the impacts of climate and land use changes. However, the sensitivity of the model outcome, regarding the hydrological and erosive response to climatic data derived with different methods, has not been fully understood. We carried out a hydrological simulation using (a) Climate Forecast System Reanalysis data set, which synthesises outputs of global climate models along with gauged weather information and has a global coverage, and (b) purely weather station-based gridded climate data provided by Indian Meteorological Department. A possible LULC scenario for the year 2020 was created using the combined Cellular Automata–Markov model. Application of both climate data sets resulted in a modest increase in the predicted streamflow and sediment yield as a response to the probable development scenario in 2020. However, the marked variations emerged in the location and monthly pattern of significant changes in the surface runoff and sediment yield in response to the likely LULC scenario for 2020 vis-à-vis 2010.

Keywords. Land use change prediction; SWAT; hydrological response; CFSR.

1. Introduction

Vegetation is an important component of a natural environment, as it distributes water and solar energy on the earth's surface through various hydrological pathways. Land use change or land conversion is a dynamic process, which varies spatially and temporally depending on the economic, political and social needs of a region. It has been shown to affect the catchment hydrological processes such as precipitation, evapotranspiration, infiltration, groundwater recharge, base flow, surface runoff and sediment yield (Li *et al.* 2011; Baker and Miller 2013; Yira *et al.* 2016) by increasing or decreasing the flood/drought frequency (Schilling *et al.* 2014; Tarigan 2016), soil erosion and nutrient loss (McGrath *et al.* 2001) leading to land degradation and lower agricultural productivity (Zucca *et al.* 2010). Destructive forest cover changes for a shorter period (Khoi and Suetsugi 2014) and expansion in farmlands and built-up surfaces were found to increase surface runoff and sediment loads in river basins (Yan *et al.* 2013).

In India, the population increased from 200 million to 1200 million between 1880 and 2010, forest cover decreased from 89 million ha to 63 million ha, while the cropland increased by ∼50 million ha, or 56%, during this period (Tian *et al.* 2014).

This makes it necessary to gain certain knowledge about the possible land use/land cover (LULC) changes and the impact they may have on the surface runoff and sediment fluxes in the large river basins of India. Markov modelling has been used to project the trajectory of land use (Iacono *et al.* 2015). This model views LULC at any given time as a discrete state and considers it to be a function of its previous state only. The probability of transition between each LULC condition and its previous state is incorporated as an element of a transition probability matrix. However, the use of Markov chain in LULC prediction has the weakness of limited spatial details (Myint and Wang 2006). On the other hand, a cellular automaton is capable of simulating the dynamic nature of land use change with better spatial accuracy by controlling the spatial pattern change using transition rules and neighbourhood 'cell state' (López *et al.* 2001). A coupled Markov and cellular automaton model performs better than applying these methods separately and have been employed successfully to predict LULC scenario in urban (Guan *et al.* 2011; Zheng *et al.* 2015) and regional scales (Arsanjani *et al.* 2011; Zare *et al.* 2017; Etemadi *et al.* 2018).

Soil and Water Assessment tool (SWAT) has been particularly useful for studying the effect of environmental changes on the quality and the quantity of water (Azari *et al.* 2016; Francesconi *et al.* 2016). Zhang *et al.* (2016) attempted to quantify the hydrological responses to land use change under constant and possible climate change scenarios in the Heihe River basin of northwest China using SWAT and observed that land use changes have a greater impact on streamflow and sediment yield at the sub-basin levels than at the main basin outlet. The possibility of climate-induced land use changes in affecting catchment hydrology and fluvial erosion has also been predicted (Serpa *et al.* 2015). Wagner *et al.* (2016) integrated the LULC modelling as well as climatic change scenarios and evaluated its impact on the hydrological response of a small and highly urbanised watershed in western India. However, we felt the need of further investigations in this direction, especially in large rural river basins, where expansion of arable land rather than the built-up surface is likely to be the main driver of LULC changes. As India is still an agro-economy, the agricultural water management is expected to be the main focus of long-term hydrological studies (Garg *et al.* 2012).

Weather data is always an important driver of rainfall–runoff processes but limited measurement of atmospheric variables over space and time and their availability sometimes restrict the practice of hydrological modelling for research and management purposes (Ciach 2003). Climate Forecast System Reanalysis and Reforecast (CFSR) provided by the National Center for Environmental Prediction at the National Center for Atmospheric Research is the most widely used weather data in SWAT models (Fuka *et al.* 2014; Auerbach *et al.* 2016). This dataset has a resolution of ∼38 km with a near-global coverage (Saha *et al.* 2010). In the Indian context, an alternative regional daily weather dataset is available in gridded format from the Indian Meteorological Department (IMD). Unlike the CFSR data set, which is partially derived from the outputs of global climate models, the IMD gridded data is entirely derived from the network of Indian weather stations and interpolated at a cell size of ∼36 km or $1° \times 1°$ (Rajeevan *et al.* 2006) and ∼18 km or $0.5° \times 0.5°$. The $1° \times 1°$ IMD data has been successfully used in the SWAT-based hydrological modelling (Singh and Gosain 2011) but the sensitivity of SWAT to CFSR *vs.* IMD weather inputs in predicting hydrological and erosion response to future LULC scenarios has never been evaluated, especially in large rural river basins. The current study aims to examine the spatial and temporal dimensions of the predicted hydrologic responses to possible future LULC scenarios that may arise due to the use of a global (CFSR) *vs.* regional (IMD) weather product.

2. Study area

The Indravati river basin (basin size: 40,525 km²) is a typical example of a large, primarily agricultural river basin in India (figure 1) with a modest amount of river controls. Information about the operation of the reservoirs is often necessary for rainfall–runoff modelling. This information is scant in India. There are no large dams in the basin except one at the extreme upstream portion. It is characterised by a tropical monsoon climate, with high seasonal variability in precipitation and temperature. The average annual precipitation is about 1288 mm in the Indravati basin (Vemu and Pinnamaneni 2011). Most of the rainfall occurs between late June and October due to the southwest monsoon. For the rest of the months, the
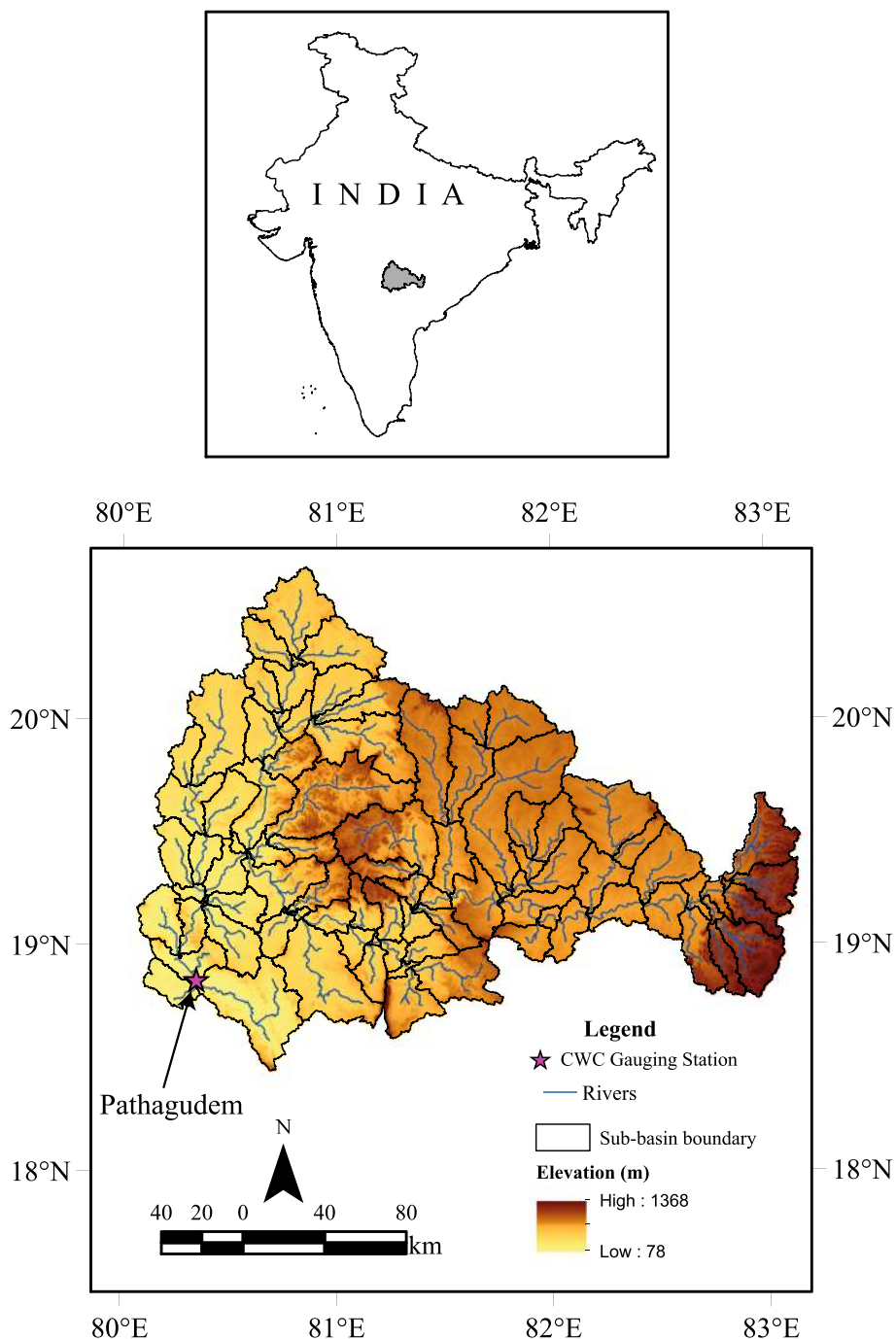
Figure 1. Location and description of the Indravati river basin, India.

annual precipitation is very low, which makes the river non-perennial.

According to GLOBELAND30 Project of China (http://www.globallandcover.com/GLC30 Download/index.aspx) the forest and agricultural land are two dominant land uses in this area, accounting for 43% and 46%, respectively, in 2010. Rest is covered by grassland, shrub land, waterbody and built up area. The central highland and western part of the basin have high forest cover whereas agricultural land is mainly confined in the upper basin in the southeast. According to the National Atlas and Thematic Mapping Organization (NATMO), Government of India, the two main types of soils in the Indravati basin are red sandy soils and red loamy soils. About 74% of the basin area is red sandy soil and the rest is red loamy soil. Moreover, some smaller tracts are dominated by clay loam, sandy clay loam and clayey soils.

Table 1. *Input dataset for LULC change simulation.*

Factors	Source
Elevation	SRTM DEM
Slope	SRTM DEM
Annual mean precipitation	IMD
Soil organic matter content	FAO Harmonized World Soil Database v 1.2
Population density	Census of India
Euclidian distance to rivers	SRTM DEM
Euclidian distance to present LULC (2010) classes	GLOBELAND30, China

3. Data and methods

3.1 *Land use change prediction*

Possible future LULC scenarios of 2020 for the Indravati basin were created in this research by integrating the Markov model, cellular automata and multi-criteria evaluation (MCE) analysis. The CA_MARKOV module in IDRISI SELVA 17.0 was used for this purpose. Initially, the transfer area matrix and the transition probability matrix were computed to determine the transition rules between the LULC state of 2000 and 2010. MCE was employed to construct suitability images that account for many important factors and constrains of land use change (Eastman 2012). For example, the area under water body and urban land use was considered as constrains while predicting the probability of these LULC categories to change into agricultural land. Factors such as elevation, slope, rainfall, organic matter in the soil, population density, distance from the river and distance from the present LULC classes (table 1) were considered for generating the suitability maps. In the next step, each factor

was standardised (0–255) using the fuzzy membership functions and their relative weights were determined using the analytic hierarchy process following the methods reported by Saaty (2003). Consequently, factor images with their relative weights were aggregated into a single suitability map for one land use category through the aggregation method. Finally, a 5×5 contiguity filter was applied to define the neighbourhood function for the land use change simulation.

The ability of the cellular automata and Markov model to predict the LULC for 2020 in the study area was tested by validating the simulated LULC of 2010 with the actual map of 2010. Pontius and Millones (2011) proposed a spatial adaptation of the κ index that is conventionally used for accuracy assessment of LULC classification as the quantity disagreement and allocation disagreement. As per these indices, we found an overall agreement of 0.89 with a simulation error of 0.109; 2.52% of the error was attributed to the quantity and 8.48% occurred due to allocation.

The computed κ index of agreement ($\kappa_{\text{no}} = 0.87$, $\kappa_{\text{location}} = 0.85$, $\kappa_{\text{standard}} = 0.82$) indicates a reasonable concurrence between the observed and simulated maps of 2010. All κ indices were above 0.82. It indicated that our simulated classified map of 2010 was 82% better than the one that would result from a random chance agreement. κ statistics calculated for each land use type (table 2) shows higher simulation accuracy for the agricultural land ($\kappa - 0.92$) and forest ($\kappa - 0.86$). As more than 90% of the basin area was under forest and agricultural land, the result indicated a high accuracy of our method to predict the LULC condition in the Indravati river basin. With this impression, we simulated the LULC map of 2020 for the study area from the map of 2010 using the same methodology.

Table 2. *Comparison of areas under different LULC classes in the simulated land use map of 2010 with the observed land use map of the same year.*

Year	LULC 2010 observed		LULC 2010 simulated			Relative
Land use	Area (km^2)	Percentage (%)	Area (km^2)	Percentage (%)	κ	error (%)
Agriculture	17434.92	43.02	17642.21	43.53	0.9168	0.51
Forest	18682.84	46.10	18502.03	45.65	0.8651	−0.45
Grassland	3521.19	8.69	3503.43	8.64	0.5515	−0.04
Shrub land	338.66	0.84	368.57	0.91	0.2619	0.07
Wet land	2.89	0.01	3.58	0.01	0.8558	0.00
Water	262.78	0.65	235.13	0.58	0.7603	−0.07
Urban	258.06	0.64	251.91	0.62	0.8016	−0.02
Bare land	24.44	0.06	18.92	0.05	0.545	−0.01

Table 3. *Input data for SWAT model in the Indravati river basin.*

Input data	Resolution	Source
DEM	30 m	SRTM DEM (https://earthexplorer.usgs.gov/)
Land use map	30 m	GLOBELAND30, China (http://www.globallandcover.com/GLC30Download/index.aspx)
Soil map	5 km	FAO Digital soil map of the world (http://www.fao.org/geonetwork/srv/en/metadata.show?id=14116)
Meteorological data	Daily	CFSR real analysis data (http://swat.tamu.edu/)
Hydrological data	Daily	CWC, India–WRIS project, India (http://www.india-wris.nrsc.gov.in/)

3.2 *Hydrological modelling*

3.2.1 *Setting up the SWAT model*

SWAT requires various spatial information (topographical, land use, soil) and meteorological inputs to operate. Topographic and channel information (reach length, longest flow path, etc.) were extracted from SRTM digital elevation model (DEM) (30 m×30 m) using Arc-SWAT tool. LULC data (2000, 2010) was obtained from the GLOBELAND30 Project. Soil maps including physical properties of the soil were obtained from the Digital Soil Map of the World (DSMW, FAO). More details about the source of the data are given in table 3. The Indravati river basin (figure 1) was divided into 94 sub-basins. The slope map of the basin was prepared from the DEM. After superimposing the physical attributes such as LULC, soil and slope map, each sub-basin was further divided into a number of hydrologic response units (HRUs) with a unique combination of LULC, soil and slope classes. A threshold for each of the physical attributes was applied to avoid the creation of unnecessarily smaller HRUs. In the next step, we incorporated weather parameters into the model domain. The Penman–Monteith method was used to calculate the potential evapotranspiration.

The primary objective of this research is to examine the spatial and temporal differences of the predicted hydrological and erosive response of a basin to likely future land use change that may arise due to the use of IMD vis-à-vis CFSR climate data. The IMD climate data was assumed to be more reliable than CFSR as the former was exclusively derived from gauged records. On the other hand, reanalysis weather data sets are produced on a large spatial scale, assimilating the ground observation and remotely sensed measurements embedded with atmospheric model 'hindcasts' to provide estimates of atmospheric variables worldwide with
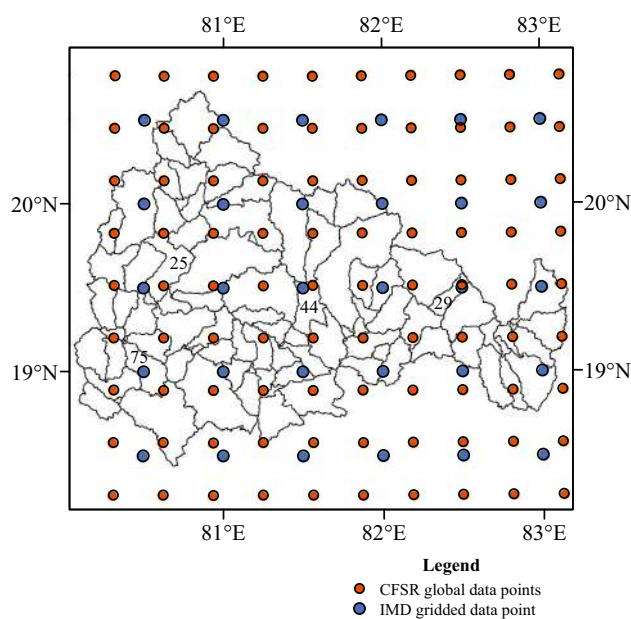

Figure 2. Distribution of gridded weather data points (CFSR and IMD) over the sub-catchments of the Indravati river basin.

a continuous based record for several decades (Saha *et al.* 2010). CFSR (at a resolution of $0.25° \times 0.25°$) and IMD gridded data (at a resolution of $0.5° \times 0.5°$) are available from the SWAT homepage (https://swat.tamu.edu/) in a format readily usable in SWAT. We used these two climate products in our study. When a gridded climatic data is fed into SWAT, the grid with its centroid nearest to the centroid of a sub-basin is taken into consideration as the climatic data for that sub-basin. SWAT is a semi-distributed hydrological model and hence lumps the climatic data at the sub-basin level.

The reliability of the CFSR data was measured using the IMD data as the benchmark. Superimposing the IMD and the CFSR grids reveal that in sub-basins nos. 25, 29, 44, 75 (figure 2) the CFSR and IMD grid centroids are very close and

Table 4. *List of the evaluation indices used to compare the CFSR and IMD precipitation data in the Indravati river basin.*

Statistical index	Equation	Unit	Perfect value
Correlation of determination (R^2)	$\sqrt{\dfrac{\sum_{i=1}^{n}(G_i-\bar{G})(S_i-\bar{S})}{\sqrt{\sum_{i=1}^{n}(G_i-\bar{G})^2}\times\sqrt{\sum_{i=1}^{n}(S_i-\bar{S})^2}}}$	–	1
RMSE	$\sqrt{\dfrac{\sum_{i=1}^{n}(S_i-G_i)^2}{n}}$	mm	0
PBIAS	$\left(\dfrac{\sum_{i=1}^{n}(S_i-G_i)}{\sum_{i=1}^{n}G_i}\right)\times 100$	mm	0
NS efficiency	$1-\dfrac{\sum_{i=1}^{n}(G_i-S_i)^2}{\sum_{i=1}^{n}(G_i-\bar{G}_i)^2}$	–	1

Note: S_i, reanalysis precipitation (CFSR); G_i, gauged observation (IMD).

hence suitable for computing the bias of the CFSR data sets in comparison to the IMD data. We used (i) degree of agreement (R^2 and Nash–Sutcliffe (NS) efficiency), (ii) error (root mean square error (RMSE)) and (iii) bias index (percentage bias) to quantify the bias in the CFSR data in our study area in daily and monthly temporal scales (table 4). Pearson's coefficient of determination (R^2) was calculated to evaluate the degree of collinearity between the CSFR and IMD rainfall figures. RMSE is an efficient way of measuring the error as it assigns a higher weight to larger error values. Percentage bias (PBIAS) measured the average tendency of the CFSR dataset to be smaller or larger to their observed (IMD) counterparts. The NS efficiency measures the relative magnitude of residual variance compared to measured data variance. The result (figure 3) shows that the CFSR data overestimated the average monthly precipitation for each of these sub-basins especially during the rainy season (July–September). Comparison of CFSR precipitation product at a daily scale shows very low agreement (R^2 0.16–0.33; NS -0.49 to 0.09) with the IMD data. The PBIAS for daily comparison ranges between -38% and -19% indicating a high overestimation of rainfall by the CFSR product. At a monthly time scale, CFSR data performed better with reported R^2 and NS values between 0.66–0.78 and 0.21–0.72, respectively, with the IMD (figure 3). PBIAS ranges were between -19% and -38% and the highest and lowest bias were associated with sub-basin nos. 29 and 25, respectively.

Rainfall–runoff routing was calculated using the SCS curve number method. After running the model for a 2-year spin-up time (1995 and 1996), the rainfall–runoff process was simulated for the

period 1997–2003. The LULC map of the year 2000 was used as the baseline LULC condition to setup the SWAT model and the IMD weather records (in SWAT format) are available for 1971–2005. In order to calibrate and validate the baseline scenario, it was necessary to compare the simulated model outputs with observed records such as gauged stream discharge and suspended sediment load. Hence, we have chosen a period (i.e., 1997–2005) close to our baseline LULC data (i.e., 2000).

3.2.2 *Calibration and uncertainty assessment*

Being a semi-distributed hydrological model, SWAT is associated with large uncertainties. The coarse resolution soil maps and weather data inherently increased the degree of uncertainty in the model outcomes. In order to calibrate the model and quantify the model uncertainties, the sequential uncertainty fitting (SUFI-2) algorithm (Abbaspour *et al.* 2007) was applied using the SWAT-CUP software.

Before proceeding to model calibration and model validation, it is essential to perform a sensitivity analysis to identify the parameters that are likely to have significant impacts on the model outputs. We used the latin hypercube one at a time sampling procedure for identifying the sensitive parameters (table 5) that may have an effect on the predicted discharge (Q) and suspended sediment concentration in the Indravati river basin. The predicted streamflow (discharge) was found to be influenced most significantly by curve number (CN-2), followed by the other parameters shown in table 5 which are ordered by their sensitivity to the model outcome. The
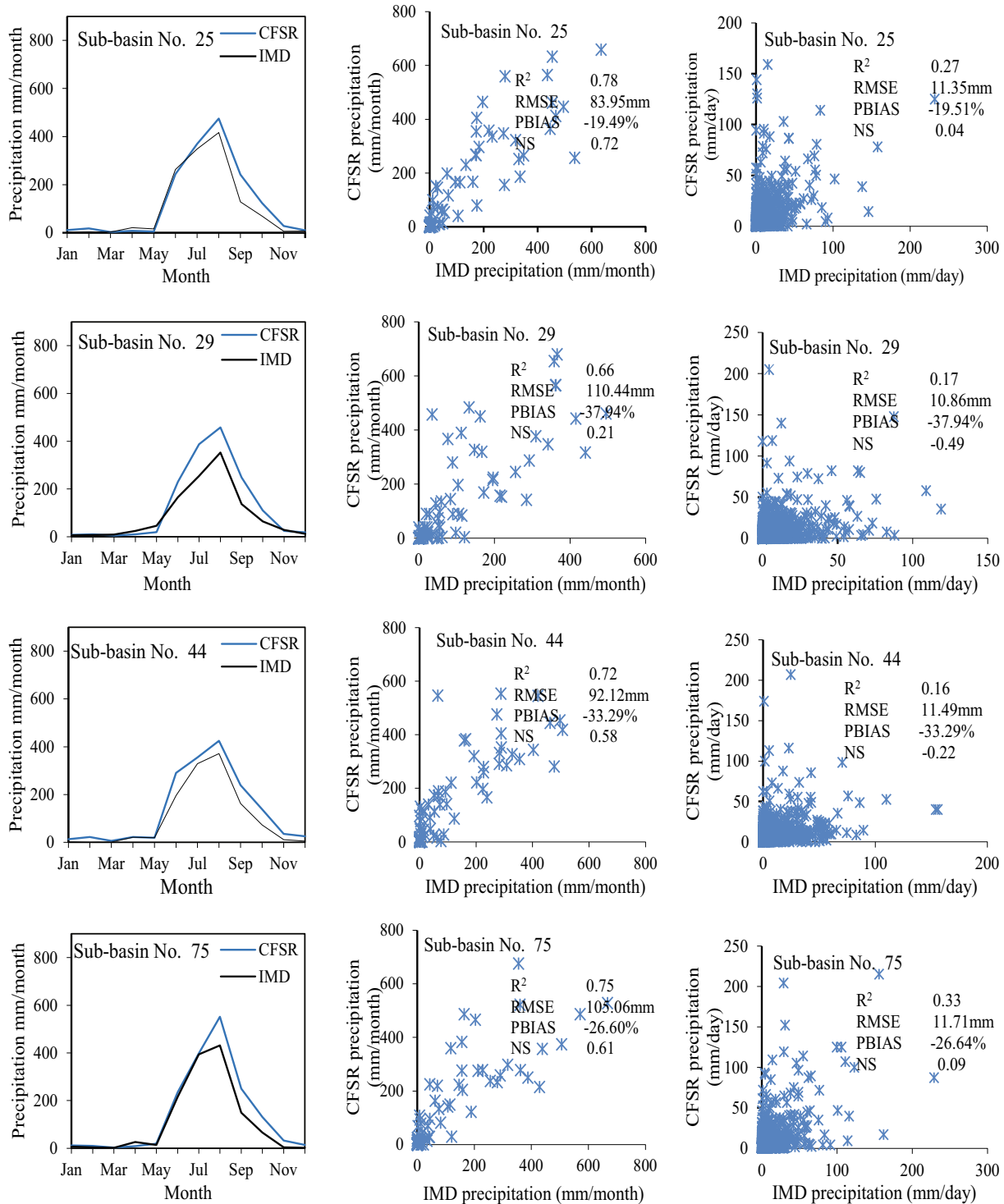
Figure 3. Comparison of CFSR data against the IMD weather dataset at the monthly and daily temporal scales at four selected sub-basins.

parameters related to the calculation of erosion and sediment export are also ordered in table 5 according to their sensitivity to the predicted sediment yield.

After identifying sensitive parameters, they were optimised using SUFI-2 by setting the NS efficiency as the objective function with a threshold value of 0.5. It was achieved by updating the initial parameter ranges after each iteration, taking into account the corresponding *p*- and *r*-factors. Initially, parameters related to surface runoff were set constant and the sensitive parameters to predict sediment yield were updated. Subsequently, the entire set of sensitive parameters was allowed

Table 5. *Results of the parameter sensitivity analysis for the SWAT model of the study area.*

Parameter name	Processes	Min.	Max.	Final range	
Parameters sensitive to flow					
CN2[a]	Surface runoff	35	98	− 0.14265	0.1726
GWQMN[b]	Soil	0	5000	800.675	4000.785
ALPHA_BF[b]	Groundwater	0	1	0.2394	0.7466
CH_K2[b]	Channel	−0.01	150	0.6750	56.0749
CH_N2[b]	Channel	0.01	0.03	0.0029	0.2096
GW_DELAY[b]	Groundwater	0	500	4.1841	48.1159
ESCO[b]	Evaporation	0	1.0	0.3804	0.8616
SOL_AWC[a]	Soil	0	1.0	− 0.2162	0.1944
OV_N[b]	Overland flow	0.01	0.3	− 0.2653	0.2361
Parameters sensitive to sediment					
SPEXP[b]	Channel	1	1.5	1.2197	1.483458
PRF_BSN[b]	Channel	0.00	2	0.5444	0.954426
SPCON[b]	Channel	0.0001	0.01	0.00017	0.000320

[a]Parameter value is multiplied by (1+ a given value), [b]Parameter value is replaced by a given value.

to vary within their respective realistic ranges to achieve the best fit with the observed surface runoff and sediment yield records available at Pathagudem gauging station (figure 1). It is mentioned in section 3.2.1 that the IMD climate data is considered to be more reliable than the CFSR products. Hence, this dataset was used for carrying out the sensitivity analysis and calibration. In this process, the calibrated parameter ranges were determined on the basis of the IMD climatic records. In the later stage, the already identified calibrated parameter ranges were used with the CFSR test cases. The period 1997–2000 was used for model calibration; validation was carried out for 2001–2003. The streamflow and sediment load data available at Pathagudem gauging station were obtained from Water Resources Information System of India (WRIS) for the purpose of model calibration and validation.

4. Results

4.1 *LULC change in the Indravati basin and future trends*

LULC changes in the Indravati river basin have been presented in table 6. Figure 4 illustrates the spatial pattern of LULC changes from 2000 to 2010 and the projected 2020 scenario across the sub-basins. Forest and agricultural land were two dominant LULC covering, about 47% and 41%, respectively, of the total area in 2000. Grassland was sparsely distributed over the entire basin (9.4% of the total area) and decreased by 0.7%

by 2010. Rest of the area is generally occupied by shrub land, urban area, bare land and water body. Forests mainly cover the western part of the basin and occupy the higher reaches, which decreased by ∼1% between 2000 and 2010. Agricultural land dominated in the eastern part and the lowland areas adjacent to the river and increased 1.6% between 2000 and 2010. LULC changes in the Indravati river basin have been primarily a result of the conversion of forest to grassland and agricultural land and grassland to agricultural land and forest. Since it is a rural catchment, deforestation and agricultural expansion were the most important drivers of land use changes in the study area. The simulated LULC scenario of 2020 vis-à-vis the prevailing condition in 2010 indicates expansion of the agricultural and urban areas with their share increasing to 2.14% and 0.06%, respectively. On the other hand, forest and grassland have been predicted to decrease by 1.45% and 0.77%, respectively (table 6).

4.2 *Runoff and sediment load simulation*

Different precipitation inputs can introduce significant uncertainties in streamflow and sediment load simulation. The calibration and validation of the model results, derived using the CSFR and IMD weather inputs, showed good agreement with the observed flow and suspended sediment concentration. Both weather data sets were able to capture the observed trend in the discharge and suspended load during calibration and validation

Table 6. *Actual and predicted land use/land cover change in the Indravati river basin (2000–2020).*

| Year	2000 Area		2010 Area		2020 Area		2010–2020 Area	
LULC	(km^2)	(%)	(km^2)	(%)	(km^2)	(%)	(km^2)	(%)
Agriculture	16772.8	41.39	17434.92	43.02	18301.64	45.16	866.72	2.14
Forest	19083.1	47.09	18682.84	46.10	18094.30	44.65	− 588.54	− 1.45
Grassland	3800.51	9.38	3521.19	8.69	3210.19	7.92	− 311.00	− 0.77
Shrub land	354.97	0.88	338.66	0.84	353.75	0.87	15.09	0.04
Wet land	2.94	0.01	2.89	0.01	3.04	0.01	0.15	−
Water	254.44	0.63	262.78	0.65	261.22	0.64	− 1.56	−
Urban	233.25	0.58	258.06	0.64	281.67	0.70	23.61	0.06
Bare land	23.79	0.06	24.44	0.06	19.97	0.05	− 4.47	− 0.01

'−' Indicates negligible change.
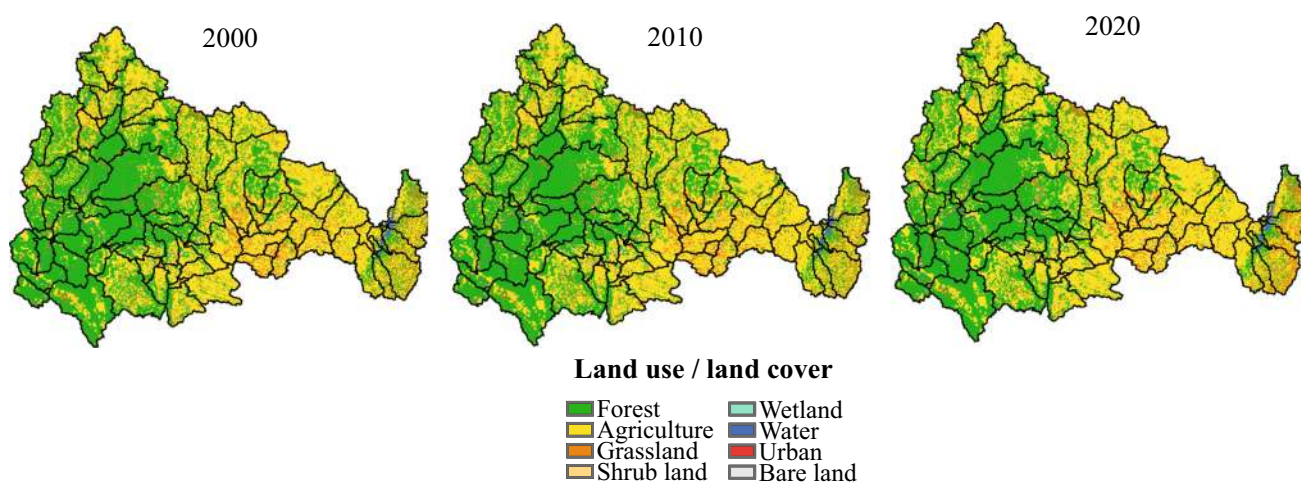
Figure 4. Historical and projected LULC patterns in the Indravati river basin.

periods (figure 5). During the streamflow calibration, the IMD inputs produced better model performance than the CFSR data set. It was evident from the computed *p*-factor, which was 10.8% and 23.6% higher than the CFSR data during streamflow calibration and validation, respectively (figure 5). In addition, the *r*-factor obtained by the IMD weather product was 11.3% (calibration) and 11.7% (validation) lower than the CFSR weather data.

During the suspended sediment load calibration, both weather products (IMD and CFSR) captured 79% of the observation but *r*-factor was 11.5% lower for the IMD data as compared to the CFSR product. NS and R^2 were above 0.73 for both data sets during the discharge and suspended sediment calibration. A similar trend was observed during the validation period and the model was able to reproduce streamflow more accurately with IMD inputs than the CFSR data set. However,

for suspended sediment load, the modelled outputs were found to be slightly better when the CFSR data was used. It was noted that 58% (*r*-factor 0.84) and 61% (*r*-factor 0.94) of the suspended sediment load observations were captured by the IMD and CFSR data sets, respectively, during the validation period.

4.3 *Effect of future LULC change*

4.3.1 *Impact on runoff*

In the near future (2020), Indravati basin is likely to experience LULC changes mainly in the form of agricultural expansion at the cost of forest and grassland. We simulated the rainfall–runoff process using the daily weather data from 1995 to 2003 obtained from the IMD and CFSR data sets and used them to predict the LULC condition of 2020.
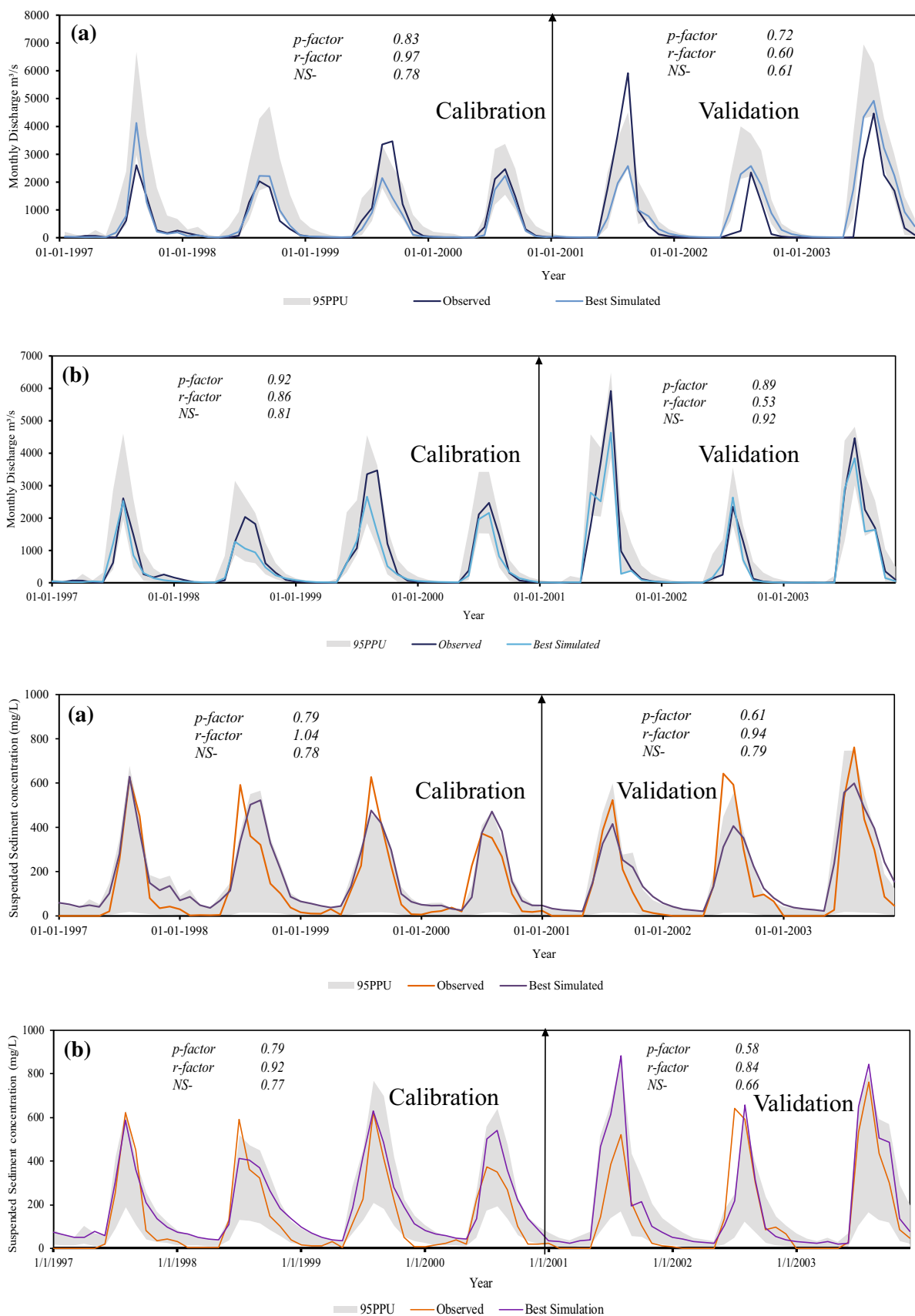
Figure 5. Monthly hydrograph of discharge (m^3/s) (two figures from the top) and suspended sediment concentration (mg/l) (third and fourth figured from the top) during calibration (1997–2000) and validation period (2001–2003) at Pathagudem gauging station. (**a**) and (**b**) represents results obtained using the CFSR and IMD weather data sets, respectively.

The model outputs pertaining to flow and sediment load were averaged for the monsoon months as well as annually. The overall results indicate an increase in surface runoff and sediment yield from the LULC state of 2010 in response to the predicted changes in 2020 (table 7). Under the predicted LULC scenario of 2020, we could expect an increase of 4.28 mm (1.24%) in the surface runoff if the IMD dataset is used as the weather input. On the other hand, an increase of 4.41 mm (1.07%) from 2010 level could be expected if the CFSR data is employed as the meteorological input in the SWAT model. The maximum increase is likely to occur during the rainy season when the region receives its maximum rainfall (table 7). Particularly during the monsoon, we can expect an average increase in the surface runoff by ~4 mm or 1.23% and 3.93 mm or 1.01% from the 2010 LULC condition using the IMD and CFSR data sets, respectively.

The changes in the surface runoff in response to the possible LULC condition of 2020 were predicted to have significant spatial variation (figure 6a). At the sub-basin scale, the highest average decrease in the surface runoff under the LULC scenario of 2020 from 2010 was found to be 12.5 mm/yr or 5.16% (sub-basin 43) using the IMD weather inputs. While applying the CFSR data, the highest average decrease was found to occur in the same sub-basin (24.84 mm/yr or 5.02%). The maximum increase of IMD and CFSR data were found to be 36.8 mm/yr or 7.69% (sub-basin 69) and 34.5 mm/yr or 3.5% (sub-basin 91), respectively. We would like to highlight that different weather inputs identified sub-basin 43 as the place to experience a significant decrease in the surface runoff in future but the predicted location of a substantial increase in the surface runoff differed according to the weather inputs. On the monthly scale, the IMD data estimated a maximum increase of 1.73% (September) in surface runoff whereas CFSR data produced the maximum increase (1.85%) in October (figure 6b).

4.3.2 *Impact on sediment yield*

Use of IMD data resulted in the prediction of an overall average increase in the sediment yield under the LULC scenario of 2020 from 2010 by 1.19 t/ha (1.13%). The figure obtained using CFSR data was 1.13 t/ha (0.47%). The season-wise breakdown of the changes is presented in table 7. Results indicate 1.16% and 0.45% possible increase under IMD

and CFSR data, respectively, during monsoon. Figure 6(b) depicts that the predicted overall average increase in the sediment yield from 2010 to 2020 is likely to be highest in September (1.28% and 1.14% for IMD and CFSR, respectively).

The maximum increase in the sediment load from 2010 to 2020 was not predicted at the same location using the IMD and CFSR data. The model predicted a maximum increase (13.95 t/ha or 7.28%) at sub-basin 60 when IMD data was used. CFSR input resulted in the identification of sub-basin 22 as the location of maximum increases (29.02 t/ha or 9.2%). Similarly, the maximum decrease in the sediment load was predicted to occur at sub-basin 58 (11.30 t/ha or 6.39%) and sub-basin 81 (20.8 t/ha or 3.57%) for the IMD and the CFSR datasets, respectively. The higher magnitude of maximum increase/decrease predicted with the CFSR dataset indicates a possible over estimation of rainfall in the CFSR data.

5. Discussion

In this study, we analysed the effect of possible LULC change on runoff and sediment yield in the Indravati river basin by integrating a land use change model with a hydrological model. Although several studies were published recently on the future land use simulation and hydrological response to future land use scenarios, they mostly focused on medium- to small-sized urban and semi-urban river basins. Such studies are particularly few in the Indian context, which is going to be the most populous country in the world in the near future with a high potential for LULC changes.

In the current study, we selected the Indravati river basin as a typical example of a large rural watershed in India. Minimum available data on past population growth and road development was derived from public domain information such as Google Earth to assist land use simulation through an integrated CA–Markov model. The overall agreement error in the simulated land use could be kept under control (0.109) by achieving a low allocation disagreement. As allocation disagreement is a function of cell resolution and can be improved by transforming the cell resolution from fine to coarse (Eastman 2012), we were able to attain a low allocation disagreement value by resampling the original LULC maps from 30 to 60 m resolution. For a basin the size of ~40,000 km^2,

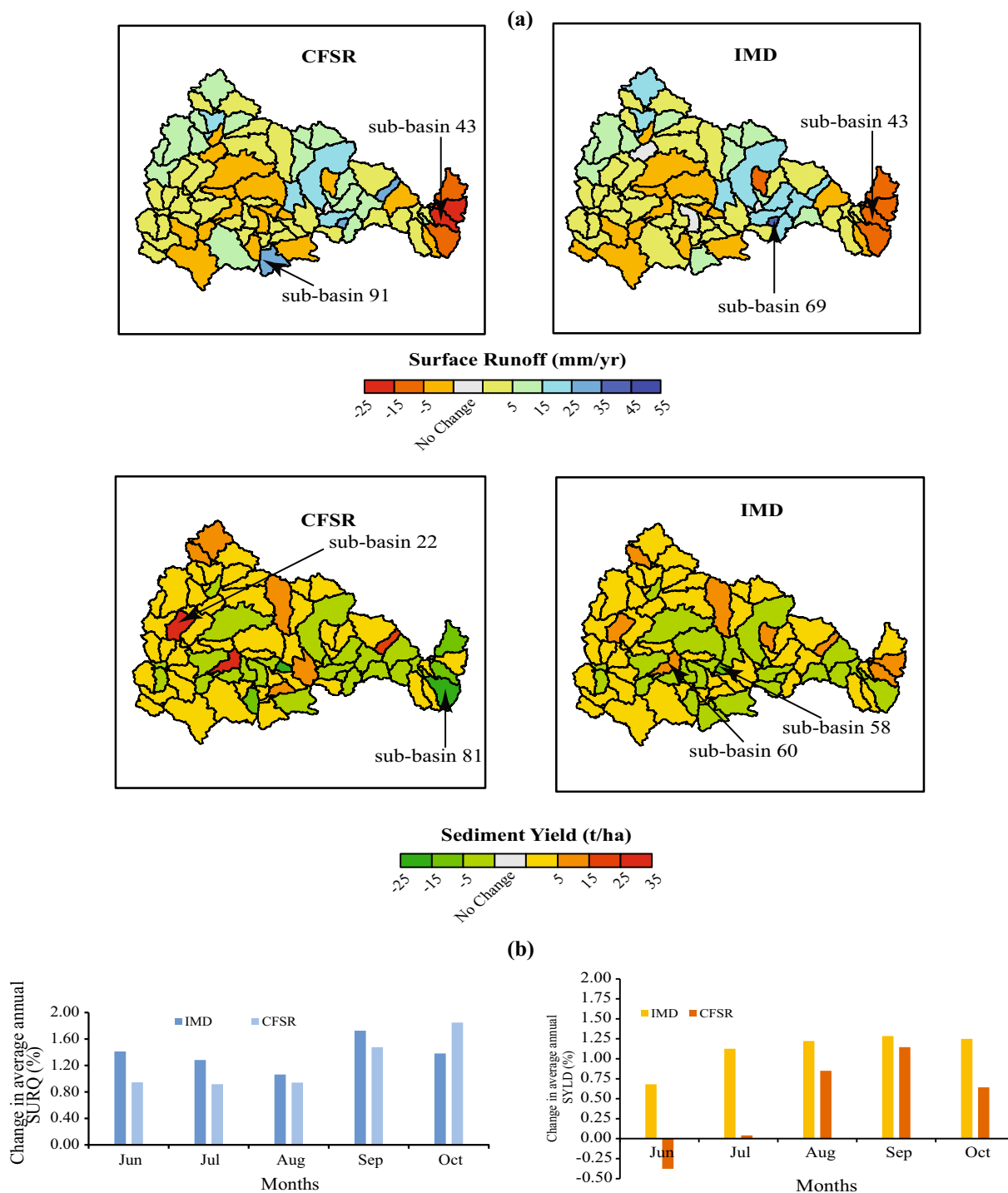Figure 6. (**a**) Changes in runoff and sediment yield in response to LULC changes (2010–2020) at the sub-basin scale. (**b**) Monthly variation in annual average runoff (left) and sediment yield (right) in the monsoon season as a response to LULC changes (2010–2020).

this aggregation of the LULC data is hardly likely to have any impact on the model outcomes.

The overall result indicates an increase in runoff and sediment yield possibly due to expanding agricultural land. Farmlands generally have a higher surface runoff coefficient than grassland and forests partly due to lower infiltration rate and reduced friction offered to the overland flow by the row crops in comparison to forest and grasslands. The tillage of land in the rainy season also contributes to the higher degree of soil erosion and consequent increase in the sediment yield at the basin outlet.

Table 7. *Variation in simulated annual average in the Indravati river basin values under different land use scenario.*

Season	Weather data	Precipitation (mm)	Surface runoff (mm)				Sediment yield (t/ha)			
			Year		Change		Year		Change	
			2010	2020	(mm)	(%)	2010	2020	(mm)	(%)
Winter	CFSR	47.82	1.72	1.78	0.06	3.49	0.96	0.97	0.01	1.04
	IMD	18.66	0.6	0.6	0.00	0.00	0.67	0.66	− 0.01	− 1.49
Pre-monsoon	CFSR	299.31	1.54	1.58	0.04	2.60	0.74	0.72	− 0.02	− 2.70
	IMD	255.33	0.12	0.12	0.00	0.00	0.05	0.05	0.00	0.00
Monsoon	CFSR	1100.11	389.72	393.65	3.93	1.01	231.11	232.16	1.05	0.45
	IMD	888.97	325.15	346.05	3.99	1.23	93.59	94.68	1.09	1.16
Post-monsoon	CFSR	161.48	20.68	21.06	0.38	1.84	12.04	12.15	0.11	0.91
	IMD	93.29	20.18	20.47	0.29	1.44	10.57	10.68	0.11	1.04
Annual average	CFSR	1608.72	413.66	418.07	4.41	1.07	244.85	246	1.15	0.47
	IMD	1256.25	346.05	350.33	4.28	1.24	104.88	106.07	1.19	1.13

Our results, on the whole, have been found in line with the previous studies (Ma *et al.* 2009; Khoi and Suetsugi 2014). The complexity of large catchments could mask the effects of land use changes that are more identifiable on smaller sub-basins. It was found in this study that the impact of land use change on annual average runoff and sediment yield was more pronounced in the smaller sub-basins compared to the whole catchment.

The SWAT model was able to capture the observed discharge and suspended sediment load with reasonable accuracy using the IMD and CFSR daily weather products. The current investigation demonstrated that regional and global weather data might lead to a remarkably different prediction about the hydrologic response to LULC changes. These differences may include variation in the seasonal pattern of flow and sediment yield and variations in the response of individual sub-basins to the altered LULC condition. The later finding is in line with the results reported by Sanyal *et al.* (2014). The significant difference in the highest amount of increase/decrease in the runoff and sediment loads also indicates the very important role of the weather inputs in examining the hydrological response to land use changes. The results are quite reliable as they represent the average model output for 6 years of daily weather data (1997–2003).

A more pronounced change in the sediment yield was predicted than surface runoff for the CFSR test case than the IMD one in response to the projected LULC conditions of 2020 (table 7). We maintain that Reanalysis rainfall products, such as CFSR data overestimates precipitation with too many wet days and underestimates high magnitude rainfall events (Shah and Mishra 2014;

Blacutt *et al.* 2015) which are not the case for the IMD data. More temporally, well distributed rainfall events (e.g., more common in CFSR) result in greater abstraction of rainfall through infiltration and evapotranspiration, generating less runoff. Hence, it was noted that the CFSR rainfall estimates are consistently higher than the IMD data in all seasons (table 7) but changes in the runoff due to the changes in the projected LULC condition (2020) are mostly higher for the IMD test case (figure 6b, left panel). The sediment yield for the CFSR test case recorded negative or negligible positive changes in comparison to the IMD test case (figure 6b, right panel). SWAT estimates sediment yield depending on the peak runoff volume (Arnold *et al.* 1998). Thus, underestimation of high magnitude rainfall events in CFSR data reduces the occurrences of sediment transporting flow events, resulting in a decrease in sediment yield even in the scenario of higher precipitation and agricultural expansion at the expense of forests.

In the validation stage of the SWAT, the main difference between the output of the CFSR and IMD data was found for the year 2001 (the first year of validation in figure 5); for 2002–2003 the differences are negligible for both runoff and sediment yield. For 2001, the CFSR data underestimated the observed runoff but predicted the observed sediment yield better than IMD which showed a marked overestimation. We argue that some parts of a catchment are generally more sensitive, in terms of sediment yield, to rainfall than others. This behaviour is primarily controlled by the slope of the land. In the Indravati river basin, the sub-basins located in the east are part of the Eastern Ghats hills, an area characterised by high relief and steep slope (figure 1). This region is the primary

sediment source area for the study basin. The spatial distribution of bias in the CFSR rainfall data adds some element of uncertainty in the model outputs at the basin outlet (Pathagudem gauging station) which are difficult to track. The better performance of the IMD data in simulating the runoff and CFSR in reproducing the sediment yield for 2001 could be attributed to this uncertainty.

6. Conclusion

The present study used SWAT to establish a long-term rainfall–runoff and sediment yield modelling framework in a large river basin in peninsular India. The aim was to examine the sensitivity of the predicted surface runoff and sediment yield to different types of meteorological inputs, generated using dissimilar methods, especially in the context of possible LULC changes in the study area. Purely in-situ observation-based regional daily weather product in the form of IMD gridded data and partially model deduced weather product with global coverage (CFSR) were considered for assessing the impact of meteorological inputs on SWAT outputs. The overall results point towards a possible increase in the surface runoff volume and sediment load in response to expansion of agricultural land at the expense of forest and grassland. The impact of future LULC conditions would likely to be more pronounced at individual sub-basin level than the entire basin. Most importantly, it was demonstrated that significant differences can arise in the predicted hydrological response to possible LULC scenario depending on the choice of weather data. These differences can have seasonal as well as spatial variations. This investigation demonstrated a robust methodology for predicting the future LULC condition in large rural basins with minimum available data, which is typical in the developing countries.

Acknowledgement

References

Abbaspour K C, Yang G, Maximov I, Siber R, Bogner K, Mieleitner J, Zobrist J and Srinivasan R 2007 Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT; *J. Hydraul. Eng.* **33** 413–430.

Arnold J G, Srinivasan R, Muttiah R S and Williams J R 1998 Large area hydrologic modeling and assessment part I: Model development; *J. Am. Water. Resour. A* **34(1)** 73–89.

Arsanjani J J, Kainz W and Mousivand A J 2011 Tracking dynamic land-use change using spatially explicit Markov Chain based on cellular automata: The case of Tehran; *Int. J. Image Data Fusion* **2(4)** 329–345.

Auerbach D A, Easton Z M, Walter M T, Flecker A S and Fuka D R 2016 Evaluating weather observations and the Climate Forecast System Reanalysis as inputs for hydrologic modelling in the tropics; *Hydrol. Process.* **30(19)** 3466–3477.

Azari M, Moradi H R, Saghafian B and Faramarzi M 2016 Climate change impacts on streamflow and sediment yield in the North of Iran; *Hydrol. Sci. J.* **61(1)** 123–133.

Baker T J and Miller S N 2013 Using the Soil and Water Assessment Tool (SWAT) to assess land use impact on water resources in an East African watershed; *J. Hydrol.* **486** 100–111.

Blacutt L A, Herdies D L, de Gonçalves L G G, Vila D A and Andrade M 2015 Precipitation comparison for the CFSR, MERRA, TRMM3B42 and Combined Scheme datasets in Bolivia; *Atmos. Res.* **163** 117–131.

Ciach G J 2003 Local random errors in tipping-bucket rain gauge measurements; *J. Atmos. Ocean. Tech.* **20(5)** 752–759.

Eastman J R 2012 IDRISI selva; Clark University, Worcester, MA.

Etemadi H, Smoak J M and Karami J 2018 Land use change assessment in coastal mangrove forests of Iran utilizing satellite imagery and CA–Markov algorithms to monitor and predict future change; *Environ. Earth. Sci.* **77(5)** 208.

Francesconi W, Srinivasan R, Pérez-Miñana E, Willcock S P and Quintero M 2016 Using the Soil and Water Assessment Tool (SWAT) to model ecosystem services: A systematic review; *J. Hydrol.* **535** 625–636.

Fuka D R, Walter M T, MacAlister C, Degaetano A T, Steenhuis T S and Easton M Z 2014 Using the Climate Forecast System Reanalysis as weather input data for watershed models; *Hydrol. Process.* **28(22)** 5613–5623.

Garg K K, Bharati L, Gaur A, George B, Acharya S, Jella K and Narasimhan B 2012 Spatial mapping of agricultural water productivity using the SWAT model in Upper Bhima Catchment India; *Irrig. Drain.* **61(1)** 60–79.

Guan D, Li H, Inohae T, Su W, Nagaie T and Hokao K 2011 Modeling urban land use change by the integration of cellular automaton and Markov model; *Ecol. Model.* **222(20)** 3761–3772.

Iacono M, Levinson D, El-Geneidy A and Wasfi R 2015 A Markov chain model of land use change; *TeMA J. Land Use Mobility Environ.* **8(3)** 263–276.

Khoi D N and Suetsugi T 2014 Impact of climate and land-use changes on hydrological processes and sediment yield – A case study of the Be River catchment, Vietnam; *Hydrol. Sci. J.* **59(5)** 1095–1108.

Li Q, Cai T, Yu M, Lu G, Xie W and Bai X 2011 Investigation into the impacts of land-use change on runoff generation characteristics in the upper Huaihe River Basin, China; *J. Hydraul. Eng.* **18(11)** 1464–1470.

López E, Bocco G, Mendoza M and Duhau E 2001 Predicting land-cover and land-use change in the urban fringe: A case in Morelia city, Mexico; *Landscape. Urban. Plan.* **55(4)** 271–285.

Ma X, Xu J, Luo Y, Prasad Aggarwal S and Li J 2009 Response of hydrological processes to land cover and climate changes in Kejie watershed south-west China; *Hydrol. Process.* **23(8)** 1179–1191.

McGrath D A, Smith C K, Gholz H L and Oliveira F D 2001 Effects of land-use change on soil nutrient dynamics in Amazonia; *Ecosystems* **4(7)** 625–645.

Myint S W and Wang L 2006 Multicriteria decision approach for land use land cover change using Markov chain analysis and a cellular automata approach; *Can. J. Remote. Sens.* **32(6)** 390–404.

Pontius Jr R G and Millones M 2011 Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment; *Int. J. Remote. Sens.* **32(15)** 4407–4429.

Rajeevan M, Bhate J, Kale J D and Lal B 2006 High resolution daily gridded rainfall data for the Indian region: Analysis of break and active; *Curr. Sci.* **91(3)** 296–306.

Saaty T L 2003 Decision making in complex environments – The analytic hierarchy process (AHP) and the analytic network process (ANP) for decision making with dependence and feedback (superdecisions tutorial); www.superde-cisions.com.

Saha S, Moorthi S, Pan H L, Wu X, Wang J, Nadiga S, Tripp P, Kistler R, Woollen J, Behringer D and Liu H 2010 The NCEP climate forecast system reanalysis; *B. Am. Meteorol. Soc.* **91(8)** 1015–1057.

Sanyal J, Densmore A L and Carbonneau P 2014 Analysing the effect of land-use/cover changes at sub-catchment levels on downstream flood peaks: A semi-distributed modelling approach with sparse data; *Catena* **118** 28–40.

Schilling K E, Gassman P W, Kling C L, Campbell T, Jha M K, Wolter C F and Arnold J G 2014 The potential for agricultural land use change to reduce flood risk in a large watershed; *Hydrol. Process.* **28(8)** 3314–3325.

Serpa D, Nunes J P, Santos J, Sampaio E, Jacinto R, Veiga S, Lima J C, Moreira M, Corte-Real J, Keizer J J and Abrantes N 2015 Impacts of climate and land use changes on the hydrological and erosion processes of two contrasting Mediterranean catchments; *Sci. Total. Environ.* **538** 64–77.

Shah R and Mishra V 2014 Evaluation of the reanalysis products for the monsoon season droughts in India; *J. Hydrometeorol.* **15(4)** 1575–1591.

Singh A and Gosain A K 2011 Climate-change impact assessment using GIS-based hydrological modelling; *Water. Int.* **36(3)** 386–397.

Tarigan S D 2016 Land cover change and its impact on flooding frequency of Batanghari Watershed Jambi Province Indonesia; *Procedia Environ. Sci.* **33** 386–392.

Tian H, Banger K, Bo T and Dadhwal V K 2014 History of land use in India during 1880–2010: Large-scale land transformations reconstructed from satellite data and historical archives; *Global. Planet. Change* **121** 78–88.

Vemu S and Pinnamaneni U B 2011 Estimation of spatial patterns of soil erosion using remote sensing and GIS: A case study of Indravati catchment; *Nat. Hazards* **59(3)** 1299–1315.

Wagner P D, Bhallamudi S M, Narasimhan B, Kantakumar L N, Sudheer K P, Kumar S, Schneider K and Fiener P 2016 Dynamic integration of land use changes in a hydrologic assessment of a rapidly developing Indian catchment; *Sci. Total. Environ.* **539** 153–164.

Yan B, Fang N F, Zhang P C and Shi Z H 2013 Impacts of land use change on watershed streamflow and sediment yield: An assessment using hydrologic modelling and partial least squares regression; *J. Hydrol.* **484** 26–37.

Yira Y, Diekkrüger B, Steup G and Bossa A Y 2016 Modeling land use change impacts on water resources in a tropical West African catchment (Dano Burkina Faso); *J. Hydrol.* **537** 187–199.

Zare M, Panagopoulos T and Loures L 2017 Simulating the impacts of future land use change on soil erosion in the Kasilian watershed, Iran; *Land Use Policy* **67** 558–572.

Zhang L, Nan Z, Yu W and Ge Y 2016 Hydrological responses to land-use change scenarios under constant and changed climatic conditions; *Environ. Manag.* **57(2)** 412–431.

Zheng H W, Shen G Q, Wang H and Hong J 2015 Simulating land use change in urban renewal areas: A case study in Hong Kong; *Habitat. Int.* **46** 23–34.

Zucca C, Canu A and Previtali F 2010 Soil degradation by land use change in an agropastoral area in Sardinia (Italy); *Catena* **83(1)** 46–54.

Corresponding editor: SUBIMAL GHOSH