Interactive comment on "Comparing CFSR and conventional weather data for discharge and sediment loss modelling with SWAT in small catchments in the Ethiopian Highlands" *by* V. Roth and T. Lemann

V. Roth and T. Lemann

vincent.roth@cde.unibe.ch

[hess, manuscript]copernicus

tablefootnote threeparttable [normalem]ulem color

[1,2,3]VincentRoth [1,2,3]TatendaLemann

[1]Centre for Development and Environment (CDE), University of Bern, Switzerland
[2]Sustainable Land Management Research Group, University of Bern, Switzerland

[3]Water and Land Resource Centre (WLRC), Addis Abeba, Ethiopia

Comparing CFSR and conventional weather data for discharge and soil loss modelling with SWAT

Vincent Roth

Vincent Roth (vincent.roth@cde.unibe.ch)

57401

Comparing CFSR and conventional weather data for discharge and sediment soil loss modelling with SWAT in small catchments in the Ethiopian Highlands

22 December 2015

Abstract

Accurate rainfall data are the key input parameter for modelling river discharge and sediment soil loss. Remote areas of Ethiopia often lack adequate precipitation data and where it is available, there might be substantial temporal or spatial gaps. To counter this challenge, the Climate Forecast System Reanalysis (CFSR) of the National Centers for Environmental Prediction (NCEP) readily provides weather data for any geographic location on earth between 1979 and 2014. This study assesses the applicability of CFSR weather data to three watersheds in the Blue Nile Basin in Ethiopia. To this end, the Soil and Water Assessment Tool (SWAT) was set up to simulate discharge and sediment soil loss, using CFSR and conventional weather data, in three small-scale watersheds ranging from 102 112 to 477 ha. Calibrated simulation results were compared to observed river discharge and observed sediment soil loss over a period of 32 years. The conventional weather data resulted in very good discharge outputs for all three watersheds, while the CFSR weather data resulted in unsatisfactory discharge outputs for all of the three gauging stations. Sediment soil loss simulation with conventional weather inputs

yielded satisfactory outputs for two of three watersheds, while the CFSR weather input resulted in three unsatisfactory results. Overall, the simulations with the conventional data resulted in far better results for discharge and sediment soil loss than simulations with CFSR data. The simulations with CFSR data were unable to adequately represent the specific regional climate for the three watersheds, performing even worse in climatic areas with two rainy seasons. Hence, CFSR data should not be used lightly in remote areas with no conventional weather data where no prior analysis is possible.

Accurately represented, spatially distributed rainfall is one of hydro-meteorological and hydro-climatic data are the most important input parameters for hydrological modelling with the Soil and Water Assessment Tool(SWAT), called SWAT hereafter (???) . Although a great deal of effort is being invested into rainfall and climatic data collection, many areas of Ethiopia have no adequate precipitation data, and where such data are available, the monitoring network contains substantial temporal and spatial gaps. This makes it necessary to use other sources of modeled rainfall data for SWAT modelling. The Global Weather Data for SWAT website Climate Forecast System Reanalysis (?) readily provides, for any coordinates coordinated on the globe, a Climate Forecast System Reanalysis (CFSR) data set for downloadclimate data set adapted to SWAT. This data set is the result of the close cooperation between two United States organizations, the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR), which have completed a global climate data reanalysis over 35 36 years from 1979 through 2014. The CFSR data is based on a spectral model which includes the parametrisation of all major physical processes as described in detail in ?; ? ; , ? , and ?.

However, a first comparison the applicability of the CFSR data for small-scale catchments in the Ethiopian Highlands has not been adequately investigated yet. Aforementioned studies did focus on large basins with numerous CFSR stations, which tend to balance errors in rainfall patterns. A first evaluation, carried out by our research group, of CFSR-modelled rainfall data with that measured by the Water and Land Re-

source Centre (WLRC, formerly the Soil Conservation Research Programme [SCRP]) in Ethiopia has shown substantial differences in daily, monthly, and annual rainfall. So far, few studies have been conducted in the Ethiopian context on the impact of rainfall data on streamflow simulations. The impact of spatial variability of precipitation on model run-off showed that standard uniform rainfall assumptions can lead to large uncertainties in run-off estimation (**?**) . Several studies evaluating the CFSR data set have suggested that climatic models tended to overestimate interannual variability but underestimate spatial and seasonal variability (**?**) Fuka et al. (**?**) used CFSR data in a 1200 km^2 watershed in Ethiopia with SWAT suggesting CFSR data performs as good or even better than conventional precipitation. Worqlul et al. (**?**) correlating conventionally recorded rainfall with CFSR data over the Lake Tana basin (15'000 km^2). They suggested that seasonal patterns could adequately be captured although the CFSR data did uniformly overestimate and underestimate measured rainfall. A recent study (**?**) from **?** evaluated the use of CFSR data for hydrological prediction using SWAT in the Lake Tana basin, Ethiopia. The study achieved satisfactory results in its simulations for both CFSR and conventional data. While the outcome was better with conventional weather data, the study concludes that CFSR could be a valuable option in data-scarce regions. Other studies using CFSR data not in the Ethiopian context (**??**) and with large to very large catchments (13'750 to 73'000 km^2) concluded that CFSR data gave good to very good results and the SWAT model responded reasonably to the data set. One CFSR application in China (**?**) with meso-scale watersheds (366 to 1098 km^2) concluded that CFSR data was significantly different and that the CFSR data spatial distribution might be the cause for the weak performance.

The impact of spatial variability of precipitation on model run-off showed that standard uniform rainfall assumptions can lead to large uncertainties in run-off estimation (**?**) . Several studies evaluating the CFSR data set have suggested that climatic models tended to overestimate interannual variability but underestimate spatial and seasonal variability (**?**) . In another study, Cavazos and Hewitson (**?**) performed statistical downscaling of daily CFSR data with Artificial Neural Networks, and their predictions showed

C5744

low performance in near-equatorial and tropical locations, which led them to conclude that the CFSR data is most deficient in locations where convective processes dominate. Another study found the CFSR data set performed well on a continental scale but that it failed to adequately reproduce some regional features (**?**). A study in China performed streamflow simulations by SWAT using different precipitation sources in a large arid basin using rain gauge data combined with Tropical Rainfall Measuring Mission (TRMM) data (**?**). The study established that streamflow modelling performed better using a combination of TRMM and rain gauge, as opposed to rain gauges only. Different interpolation schemes with the use of univariate and covariate methods showed that Kriging and Inverse Distance Weighting performed similarly well when used with the SWAT model (**?**).

In this paper, WLRC and SCRP rainfall data (hereafter called WLRC data) are compared to CFSR data over a period of 35 maximum period of 34 years from 1981 to 2014. 2014 (Maybar, 33 years for Andit Tid and 32 years for Anjeni). The main objective of this paper is to compare the two data sets for annual, interannual, and seasonal cycles and subsequently to compare the effects on discharge and sediment soil loss modeling when using these data sets in three locations in the Ethiopian highlands (see figure 1). Calibrated CFSR modeled discharge and sediment soil loss is then compared to calibrated WLRC modelled discharge and sediment soil loss, and the applicability of the CFSR data in small-scale catchments for hydrological predictions is statistically evaluated and compared.

1 Methods and materials

The effects of spatial and temporal variability in the CFSR rainfall data set for the study area were examined in several steps. First the CFSR data were statistically compared to measured WLRC rainfall data for accurate representation of annual, interannual,

C5745

and seasonal cycles. This is important because temporal occurrence of rainfall has a great impact not only on discharge but moreover on sediment yield generation. Many crop types are sowed at the beginning of the rainy season(s), which implies extensive extensive ploughing beforehand, which leaves fields unprotected for the first few rainfall events. Hence, is clear that temporal occurrences of annual, interannual and seasonal cycles play a crucial role for the validation of a data set like the CFSR climatic data. Second, the impact of spatial and temporal variability of rainfall on hydrology and soil loss was assessed by modeling discharge and sediment soil loss with the SWAT model. The SWAT model was calibrated for discharge once using WLRC rainfall climatic data and once using the CFSR rainfall climatic data set. Afterwards sediment soil loss was calibrated for each catchment. In a last step discharge and sediment soil loss on a monthly basis were statistically and visually compared using performance ratings established by Moriasi et al. (**?**).

1.1 Study area

The study areas of the three micro-scale catchments are located in the eastern and central part of the Blue Nile Basin. The Anjeni (AJ) and the Andit Tid (AT) are sub-basins of the Blue Nile Basin, which drains towards the west into the main Nile at Khartoum. The Maybar (MA) catchment drains into the Awash river to the East of the Ethiopian highlands. The catchment sizes range from 104 ha to 447 112 ha to 477 ha and their altitudinal ranges extend from 2400 to 3548 2406 to 3538 masl (see table 1 on page C5764 for details). The catchments have a sub-humid to humid climate with an annual temperature ranging from 12 C to 16 C and a mean annual rainfall ranging from 1211 mm to 1690 mm. The rainy seasons are divided into two seasons for Anjeni has a unimodal rainfall pattern with a main rainy season from June to September while Andit Tid and Maybar and into one for Anjenihave a bimodal rainfall regime with a small rainy season from April to May (belg) and a main rainy season from June to September (kremt) followed by a long dry season from October to March. Land use

C5746

is dominated by smallholder rain-fed farming-systems with grain-oriented production, ox-plough farming, and uncontrolled grazing practises.

1.2 Hydrometeorological data

The hydrometeorological data consists of two sets. The conventional or measured data contains daily rainfall and maximum and minimum temperature from one climatic station for each watershed. These climatic stations have been installed in the early 1980s and span the period until 2014 with some larger gaps (see Table 1 for details) mainly from 2000 to 2010. The CFSR data (The Texas A&M University spatial sciences website, globalweather.tamu.edu) was obtained for the entire Blue Nile Basin (Bounding box: latitude $8.60 - 12.27$ N and longitude $33.94 - 40.40$ E) before choosing the four closest stations for each watershed. It includes daily rainfall, maximum and minimum temperature as well as wind speed, relative humidity, and solar radiation for 12 locations, 4 for each watershed (see Figure 1 for details).

1.2.1 Hydrologic model

ArcSWAT (Version 2012.10.14

1.3 Hydrologic model

SWAT (SWAT2012 rev. 620) was used to assess the impact of different rainfall patterns on run-off and sediment loss dynamics (**?**) soil loss dynamics through the ArcSWAT interface (Version 2012.10_1.14). Here, we present the SWAT model only briefly, as it has been widely used in the past, with extensive review of its performance and parameterization in Ethiopia and other regions (**???????**). SWAT is a physically-based river basin or watershed modelling tool. The SWAT model requires specific information

C5747

about weather, soil properties, topography, vegetation, and land management practices occurring in the watershed (**?**). ArcSWAT divides the catchment into hydrological response units (HRUs) based on unique combinations of soil type, land use, and slope classes that allow for a high level of spatial detail simulation. Runoff is predicted separately for each HRU and routed at subbasin level to obtain the total runoff for the watershed (Neitsch et al., 2011). The surface run-off is estimated in the model using one of two options (1) the Green and Ampt method (**?**) or (2) the Natural Resources Conservation Service Curve Number (SCS-CN) method (**?**). The flow routing is estimated using the variable storage coefficient method (**?**), or the Muskingum method (**?**). Sediment soil loss for each HRU is calculated through the Modified Universal Soil Loss Equation (MUSLE). Sediment routing in channels is estimated using stream power (**?**) and deposition in channels is calculated through fall velocity (**??**).

1.4 Spatial data

The spatial data used in ArcSWAT for the present study included the digital elevation model (DEM), land use data, and soil data (see table 1 for details). The DEM for the three WLRC watersheds was developed by the Centre for Development and Environment (CDE) of the University of Bern, Switzerland, for the former SCRP (**???**) and has a resolution of 2 m. The spatial distribution of soils for Anjeni was adapted from a soil survey carried out by the SCRP (**?**) and a PhD dissertation by Gete Zeleke (**?**). The physical and chemical parametrisation of the soil was adapted from the soil database in Zeleke's thesis and from Kejela's report. The soil characteristics for Maybar were adapted from the SCRP's Soil Conservation Research Report 7 (**?**) and for Andit Tid from the SCRP's Research Report 3 (**?**). Land use data were adapted from yearly surveys carried out by SCRP and WLRC through land use mapping and interviews and by own surveys in 2008 and 2012. To adapt to annually changing land use patterns, a generic map was adapted from the WLRC land use maps of 2008, 2012, 2014 (Anjeni), and 2010, 2012, 2014 (Andit Tid, Maybar). The planting and harvesting times were av-

eraged over the entire period and planted at similar dates for the entire simulation. To simulate crop growth we used the heat unit function in ArcSWAT. Teff, for example, was planted beginning of July and harvested beginning of December with several tillage operations preceding planting. Tillage operations were adapted to the usage of the traditional Ethiopian plough called "Maresha" according to Temesgen et al. (**?**).

1.5 SWAT model setup

The watersheds were delineated using the Arc–SWAT delineation tool and its stream network compatibility was checked against the stream network from satellite images . The sub-basin sizes were fixed at 2 ha. (one satellite image for each watershed). SWAT compiled 1038 HRUs for Anjeni, 1139 HRUs for Maybar, and 728 HRUs for Andit Tid respectively. All HRUs were defined using a zero percentage threshold area, which means that all land use, soil, and slope classes were used in the process. Daily precipitation and minimum and maximum temperature data at three WLRC stations were used to run the model with conventional weather inputs. All three WLRC stations The CFSR time series were complete from 1979 to 2014. The WLRC data had substantial gaps in the time series, mostly in the early 1990s and after 2000 (see Table 1 on page C5764 1 for details). The SWAT weather generator was used to fill the gaps for rainfall, temperature, solar radiation, and relative humidity in the WLRC data set for rainfall and temperature. Otherwise daily precipitation and minimum and maximum temperature data were used to run the model. Potential evapotranspiration (PET) was estimated using the Hargreaves method (Hargreaves et al., 1985). Daily river flow and sediment concentration data were measured at the outlet of the three WLRC watersheds. The flow observations are available throughout the entire year while sediment concentrations calculated sediment concentrations from grab samples are only available during rainstorm events , when sediment concentrations are visible in the river and are extrapolated over the whole time period. Personnel at the research station are instructed to take grab samples only during rainfall events, when the river is turning

brown. The planting and harvesting times were averaged over the entire period and planted at similar dates for the entire simulation. To simulate crop growth we used the heat unit function in ArcSWAT. Teff (eragrostis teff), a widely cultivated and highly nutritional crop native to Ethiopia, was planted beginning of July and harvested beginning of December with several tillage operations preceding planting. Tillage operations were adapted to the usage of the traditional Ethiopian plough called Maresha according to Temesgen et al. (**?**) with a tillage depth of 20 cm and a mixing efficiency of 0.3. During the dry season and outside rainfall events the monitored rivers are assumed sediment free.

The model was run for 32 years from 1983 to 2014 with daily data inputs but monthly outputs. Calibration and validation periods were chosen equally balanced regarding high-flow and low-flow years in all three catchments. The model was first calibrated and validated for discharge and then calibrated and validated for sediment soil loss (see Table 1 on page C5764 for details).

SWAT Swat parameters used for discharge and soil loss calibration with initial ranges and fitted value final parameter ranges.

Variable	Parameter name	Definition	Initial range	Fitted parameter ranges		
				Andit Tid	Anjeni	Mayba
Discharge	a*a__CN2.mgt	Curve number	−25 to 15	16.7 to 18.7	−7 to −4	15 to 2
	**v__GWQMN.gw	Threshold depth of water in shallow aquifer required for return flow to occur	0 to 5000	4761 to 4990	0 to 1611	2500 to 500
	a__ESCO.hru	Soil evaporation compensation factor	0 to 1	-0.0038 to 0.046	0.0023 to 0.067	0 to 0.3
	v__GW_REVAP.gw	Groundwater "revap" coefficient	0.02 to 2	0.18 to 0.19	0.17 to 0.21	0.15 to 0
	a__CH_K2.rte	Effective hydraulic conductivity in channel	−0.01 to 500	6 to13to 13	−11 to 58	−0.01 to 1
	a__CH_N2.rte	Manning's "n" value for the main channel	-0.01 to 0.3	0.0012 to 0.067	−0.15 to 0.062	0.025 to 0.06
	a__SURLAG.bsn	Surface runoff lag time	0.05 to 24	-0.084 to 3.98	0 to 6.63	0.05 to 1
	a__RCHRG_DP.gw	Deep aquifer percolation fraction	0 to 1	0.36 to 0.66	−0.51 to 0.23	0 to
	v__EPCO.hru	Plant uptake compensation factor	0 to 1	0.78 to 1.55	0.22 to 0.745	0 to
	v__SOL_AWC(1).sol	Available water capacity of the soil layer	0 to 1	0.13 to 0.22	0.19 to 0.47	0 to
Sediment	a__SLSUBBSN.hru	Average slope length	−10 to 45	8.85 to 42.34	−6.24 to -4.60	−5 to
	a__HRU.SLP_SLP.hru	Average slope steepness	−0.1 to 0.4	−0.16 to −0.04	−0.12 to −0.09	−0.5 to 0.7
	a__USLE_K(1).sol	USLE equation soil erodibility (K) factor	−0.11 to 0.24	0.079 to 0.14	0.44 to 0.49	0.04 to 0.3
	a__USLE_C.plant.dat	Min value of USLE C factor applicable to the land cover/plant	0.04 to 0.24	0.0009 to 0.004	0.48 to 0.5	0.34 to 0.62
	a__USLE_P.mgt	USLE equation support practice	0.42 to 0.79	−0.41 to 0.19	0.16 to 0.26	0.09 to 0.9
	v__SPCON.bsn	Linear parameter the maximum amount of sediment that can be reentrained	0.0001 to 0.01	0.005 to 0.007	0.0067 to 0.010	−0.01 to 0.0
	v__SPEXP.bsn	Exponent parameter for calculating sediment reentrained	1 to 1.5	1.27 to 1.5	1.32 to 1.37	1.23 to 1.3
	v__CH_COV1.rte	Channel cover factor	−0.05 to 0.6	0.2 to 0.39	0.057 to 0.099	−0.02 .05 to 0.0
	v__PRF_BSN.bsn	Peak rate adjustment factor for sediment routing in the main channel	0 to 2	0.9 to 1.1	1.2 to 1.6	0.89 to 1

a__ means a given value is added to the existing parameter value

v__ means the existing parameter value is to be replaced by a given value

1.6 Calibration setup, parameterization validation, and uncertainty sensitivity analysis

The SUFI-2 algorithm (**??**) in SWAT-Cup (**??**) was used for the calibration and validation procedure and for sensitivity, and uncertainty analysis. SWAT-Cup calculates the 95% prediction uncertainty band (95PPU) in a iterative process. For the goodness of fit two indices called "p-factor" and "r-factor" are used. The P-factor is the fraction of measured data inside the 95PPU band, and varies from 0 to 1 where 1 indicates perfect model simulation. The r-factor is the ratio of the average width of the 95PPU band and the standard deviation of the measured variable. There are different approaches regarding balance of p-factor and r-factor. The p-factor should preferably be above 0.7 for discharge and the r-factor value should be below 1.5 (**?**), but when measured

data are of lower quality other values apply. Once acceptable p-factor and r-factor are reached statistical parameters for time series analysis are compared.

For this study we used the Nash-Sutcliff Nash-Sutcliffe Efficiency (NSE), the standardized Root Mean Square Error (RSR), and the Percent Bias (PBIAS). All are very commonly used statistical parametersThese are well-known statistical parameters, which are often used for comparison of time-series especially in hydrological modelling (?????) and therefore help others to compare our modeling results to previous studies. This study refers to the model evaluation techniques described by Moriasi et al. (?), who established guidelines for the proposed statistical parameters (see table 2 below for details). The NSE is a normalised statistic that indicates how well a plot of observed versus simulated data fits the 1:1 line and determines the relative magnitude of the residual variance compared to the measured data variance (?). NSE ranges from $-\infty$ (negative infinity) to 1, with a perfect concordance of modelled to observed data at 1, a balanced accuracy at 0 and a better accuracy of observations below zero. The RSR is a standardized RMSE, Root Mean Square Error (RMSE, standard deviation of the model prediction error), which is calculated from the ratio of the RMSE and the standard deviation of measured data. RSR incorporates the benefits of error index statistics and includes a scaling factor. RSR varies from the optimal value of 0, which indicates zero RMSE or residual variation, which indicates perfect model simulation to a large positive value, which indicates a large residual value and therefore worse model simulation performance (?).

The PBIAS measures the average tendency of the simulated values to be larger or smaller than their observed counterparts. The optimal value of PBIAS is zero. PBIAS is the deviation of data being evaluated, expressed as a percentage. A positive PBIAS value indicates the model is under–predicting measured values, whereas negative values indicate over–predicting.

For this article the recommendations for reported values were strictly applied for discharge calibration and lowered for sediment loss soil loss calibration.

The model performance was also evaluated using the hydrograph visual technique, which allows a visual model evaluation overview. As suggested by Legates and McCabe (?) ? this should typically be one of the first steps in model evaluation. Adequate visual agreement between observed and simulated data was compared on discharge and sediment soil loss plots on a monthly basis.

2 Results and discussion

2.1 General comparison of CFSR and WLRC rainfall data

The raw CFSR and WLRC rainfall input data showed significantly different patterns and rainfall amounts. For Andit Tid, situated on the eastern escarpment of the Blue Nile Basin, the belg and kremt rainfall seasons were temporally adequately represented; i.e., the timely occurrences of the rainy seasons were correctly represented through the CFSR data. However, total CFSR rainfall amounts were far from adequately representedmeasured values: while the belg rainfall season in the CFSR data showed some overestimation, the total rainfall and length of the kremt rainy season were strongly underestimated. WLRC data distinctly show a main rainy season from July to September and a light rainy season from March to May, while the CFSR data only show mildly increased rainfall in March, April, July, and August but no distinct rainy season (see figure 2 on page C5763 for comparison).

The CFSR data for Anjeni highly overestimated rainfall in the region. While WLRC data showed a clear trend towards only one main rainy season from May/June to September with average monthly rainfall ranging from 100 mm (May) to 380 mm (July), the CFSR data showed a pronounced main rainy season with monthly averages ranging from 400 mm to 1000 mm from June to September and a distinct small rainy season from March to May with monthly averages three times as high as the WLRC rainfall data. The total

annual CFSR rainfall was three times the WLRC annual rainfall.

WLRC Maybar data showed a clear seasonality, with two rainy seasons, one in March and April, and one from July to August. The belg rainy season showed only mild increase of average rainfall to around 75 mm/month and the kremt rainy season showed a distinct increase of rainfall to an average of 270 mm/month. From the CFSR rainfall data, no clear distinction could be made between the belg and the kremt rainy season – both showed a rainfall increase to around 150 mm/month and the total annual rainfall was strongly underestimated.

In general, all CFSR rainfall patterns showed a similar composition: data variability was more uniformly distributed and the distinct seasonality of the WLRC data was not well represented. CFSR data underestimated the bimodal rainfall climates and strongly overestimated the unimodal rainfall climate. The WLRC data has a highly variable rainfall range in the bimodal rainfall locations, which is not reflected by the CFSR data. In general, the CFSR rainfall data does not represent the high variability of rainfall measured by WLRC data.

2.1.1 Seasonal comparison of rainfall data

The seasonal components of the CFSR rainfall were assessed for the three stations by breaking the monthly data into seasons (dry season from October to March, small rainy season (belg) from April to May, and large rainy season (kremt) from June to September) and by comparing only these. The comparison of measured rainfall to modelled rainfall for the dry season from October to March was unsatisfactory (NSE < 0.50) with negative NSEs for three stations (AT: –1.92, AJ: –12.19, MA: –0.77). The PBIAS indicated model underestimation for Anjeni and Maybar (AJ: 134.2, MA: 30.7) and an overestimation of the rainfall for Andit Tid (AT: –55.2). The RSR showed large positive values (AT: 1.68, AJ: 3.55, MA: 1.3) indicating a low model simulation performance and again an unsatisfactory rating (see table).

For the belg rainy season from April to May the model performed badly. Surprisingly, the model performed worst in Anjeni, where no small rainy season occurs. The CFSR model performance for Anjeni was unsatisfactory, with an NSE of –5.42, a PBIAS of 106.1, and an RSR of 2.48. The CFSR model overestimated the monthly rainfall in all but 5 out of 22 years. Andit Tid and Maybar were slightly more adequate but still unsatisfactory. NSE was –0.79 and –0.24 respectively, indicating unsatisfactory performance. PBIAS was –39.4 and 24.3, respectively. RSR was 1.31 and 0.85, which again indicates an unsatisfactory result.

The kremt rainy season from June to September is the season with the heaviest rainfall throughout the year. On average some 77% of the yearly rain falls within this time period. This is also the time period where the heaviest soil erosion occurs induced by rainfall. For Anjeni, Andit Tid, and Maybar the CFSR model performed unsatisfactorily (see Table 7 and Figure 3 in appendix A) with NSEs below 0.50 (AT: –9.79, AJ: –50.09, MA: –3.28), RSRs above 0.70 (AT: 3.23, AJ: 7.0, MA: 2.03), and PBIAS values ranging from –69.2 (AT) and –47.1 (MA) to +128 (AJ).

The kremt rainy season was underestimated by the CFSR model for the bimodal rainfall pattern in Andit Tid and Maybar, while the unimodal rainfall pattern was heavily overestimated by the CFSR model.

2.2 Discharge modeling with WLRC and CFSR data

The performance ratings for each of the three catchments including SWAT–Cup *p–factor* and *r–factor* are summarised in Table 4 on page C5766. The table is divided into discharge comparison and sediment soil loss comparison. Each model was calibrated with one to five iterations using 500 simulations each. The data was split into calibration and validation periods, which contained similar amplitudes (see figure 3 for further details) over their respective periods. Parameters initially contained original ranges , which were gradually adapted according to modeling results. The final ranges Final

2.2.1 Andit Tid

Calibration of Andit Tid with WLRC rainfall data yielded *very good* results. With an *p–ractor* of 0.71 and a *p–factor* of 0.53 (see Chapter 1.6 on page C5751 for performance rating) the statistical parameters RSR, NSE and PBIAS yielded *"very good"* results (0.46, 0.79, 3.1 respectively). Validation for Andit Tid yielded in satisfactory results with The CFSR rainfall data, which underestimated the WLRC rainfall pattern, yielded *unsatisfactory* results with RSR, NSE, and PBIAS of 0.80, 0.36, and 31.4. Parameter ranges settings were maximised, but still inside SWAT absolute values ((**?**) . The hydrograph on page C5763 shows that the underestimation of rainfall amounts for Andit Tid did result in a constant underestimation of peak flows and of base flows throughout the whole time period.

Validation of discharge for Andit Tid with WRLC data showed *very good* results with RSR: 0.46, NSE: 0.79 and PBIAS 9.6 and marginally *unsatisfactory* results for the CFSR dataset (RSR: 0.74, NSE: 0.45, PBIAS: 37.9).

2.2.2 Anjeni

Anjeni showed *very good* result for calibration with WLRC rainfall data. RSR, NSE and PBIAS were well inside the optimal performance ratings (0.39, 0.85, and 3.7 respectively), see table 2 on page C5765 and figure 3 on page C5763 for comparison.

Calibration Satisfactory calibration could not be reached with CFSR data , where the CFSR rainfall data did strongly overestimate the measured rainfall data proved impossible. With parameter ranges set to maxima, and neither baseflow, nor peaks could be adequately represented. With a *p–factor* of 0.49 and an *p–ractor* of 1.91 the statisti-

cal parameters were *unsatisfactory* (RSR: 2.70, NSE:-6.27, and PBIAS: -226.0). The hydrograph (Figure 3 on page C5763) shows that the strong overestimation of CFSR rainfall data during belg lead to a modelled discharge with extreme peaks during kremt, which do not correspond to the discharge regime of measured WLRC data.

Validation of discharge for Anjeni with WRLC data showed *very good* results with RSR: 0.41, NSE: 0.83 and PBIAS −6.7 and *unsatisfactory* results for the CFSR dataset with RSR: 1.24, NSE: −0.53, and *very good* PBIAS: 8.1.

2.2.3 Maybar

Calibration of Maybar with WLRC rainfall data proved to be less straight forward than Anjeni and Andit Tid. The rugged topography of Maybar combined with a inadequate cross-section proved challenging to model. Nonetheless, *satisfactory* result were achieved for discharge with RSR, NSE, and PBIAS of 0.63, 0.60, and −23.4 respectively.

The CFSR rainfall data yielded an *unsatisfactory* discharge simulation result with RSR: , NSE: , NSE, and PBIAS: . As the CFSR modelled rainfall shows two similar rainy seasons where WLRC rainfall data has distinct belg and kremt rainy season, SWAT modelled discharge showed similar trends. The hydrograph with CFSR data on page C5763 shows Figure 3 shows regular discharge peaks from February to April for every year, when there are none measured while showing only small CFSR peaks for the main rainy season from June to September, when measured discharge is significantly increasing. Again, the SWAT modelled discharge reflected the March, in accordance to rainfall pattern deviation as seen on Figure 2, when no increase of discharge was measured at the research station. The SWAT model reflected input rainfall pattern adequately, which lead to discharge peaks during belg, when there are none in the measured data. At the same time it lead to reduced discharge peaks during kremt, when the measured WLRC data are clearly pronounced.

Validation of discharge for Maybar with WRLC data showed *good* results with RSR: 0.56, NSE: 0.74 and PBIAS 17.3 and *unsatisfactory* results for the CFSR dataset with RSR: 0.98, NSE: 0.04, and *very good* PBIAS: −1.9.

2.3   Sediment Soil loss modelling with WLRC and CFSR data

Sediment Soil loss modelling was calibrated using the same set of 9 parameters for each catchment (see Table 1.5 2 on page C5751 for description). Calibration of soil loss was conducted using the parameter ranges for discharge calibration, and adapting the sediment parameters while leaving discharge parameters untouched. Performance ratings for each of the three catchments including SWAT–Cup *p–factor* and *p–ractor* are summarised in table 4 on page C5766 and visually represented on Figure 4 on page C5763. Performance rating levels were considerably lowered for sediment soil loss modeling. Threshold for the *p–factor* was set at 0.40 with an *r–factor* below 1.80 and standard performance ratings for RSR, NSE and PBIAS.

2.3.1   Andit Tid

The good results from WLRC discharge modeling facilitated sediment soil loss calibration and resulted in *satisfactory* performance ratings for RSR, NSE (0.69, 0.65), and an *unsatisfactory* PBIAS, which was slightly below threshold with −56.3. Graphic representation showed good visual results (see figure 4 on page C5763) in general, but also showed constant overestimation of the modelled data except for three years 1988, 1989, and 1994.

Sediment loss modelling with CFSR data reflected the results from discharge modelingValidation of sediment yield for Andit Tid with WRLC data showed a marginally *satisfacroy* result with RSR: 0.68, NSE: 0.51 and *unsatisfactoy* PBIAS −64.3 indicating a general overestimation and *unsatisfactory* results for the CFSR dataset with RSR: 1.39,

C5758

NSE: −0.94, and *satisfactory* PBIAS: −11.9 indicating underestimation.

2.3.2   Anjeni

Sediment Soil loss modeling with WLRC rainfall data and calibrated discharge yielded *satisfactory* results.With a *P–factor* of 0.40 and an *r–factor* of 0.65, and statistical parameters RSR: 0.67, NSE: 0.55, and PBIAS: -19.9 the model was just *satisfactory*. The graphic showed adequate results with a constant overestimation of the model except for two years in the early nineties. Modelling with CFSR data, resulted in strongly unsatisfactory results (RSR: 1.01, NSE: −0.02, and PBIAS: -33.9), which can easily be explained with the strong model overestimation of rainfall and subsequently discharge. Parameters could not be adapted further to achieve better results as they were already set to the edge of the possible ranges.

Validation of sediment yield for Anjeni with WRLC data showed *satisfactory* results with RSR: 0.67, NSE: 0.64 and PBIAS −14.1 indicating a general overestimation and *unsatisfactory* results for the CFSR dataset with RSR: 1.02, NSE: −0.03, and *satisfactory* PBIAS: −1.9 indicating underestimation.

2.3.3   Maybar

Sediment Soil loss calibration with WLRC rainfall data and calibrated discharge resulted in *unsatisfactory* statistical results (RSR: 1.24, NSE: −0.54, PBIAS: −34.1). *P–factor* and *r–factor* were 0.42 and 0.60, respectively.

Calibration in Maybar with CFSR rainfall data yielded *unsatisfactory* results (RSR: 1.02, NSE: −0.03, PBIAS: 54.4). As described in the discharge calibration section (Section 2.2.3), CFSR rainfall data in Maybar tended towards overestimation of belg and underestimation of kremt, which resulted in overestimation of monthly discharge during

C5759

belg and underestimation during kremt. This trend was redrawn with sediment calibration resulting in small but distinct peaks during belg and smaller peaks than measured during kremt. There was no satisfactory calibration possible with CFSR rainfall data.

Validation of sediment yield for Maybar with WRLC data showed *satisfactory* results for both data sets with a very strong overestimation from the CFSR data set and an equally strong overestimation from the WLRC data set.

In this paper we studied the applicability of CFSR weather data to three small-scale watersheds in the Ethiopian highlands with the goal of assessing the usability for future modelling in data-scarce regions. First, we compared CFSR and WLRC rainfall data at three stations in the Ethiopian Highlands and therefore rainfall data was compared on a monthly basis with boxplots. Second, we modelled discharge with the SWAT model; once with WLRC data and once with CFSR rainfall data. Third, we modelled sediment soil loss for the three stations with the SWAT model and compared calibrated results from CFSR rainfall and conventional rainfall to measured data.

The rainfall data comparison for CFSR and WLRC data showed strong discrepancies in seasonal and monthly rainfall amounts for all three catchments. For Andit Tid, both, belg and kremt rainy season were levelled downwards resulting in *unsatisfactory* results for each season with strongest deviations for kremt (see Tabel 3 on page C5765 for details). Anjeni rainfall data from the CFSR model overestimated the measured WLRC rainfall very strongly. This resulted in strong deviations with performance ratings well below *satisfactory* thresholds. Maybar rainfall data from CFSR showed the highest deviation for the representation of seasonality. Neither belg, nor kremt or the dry season were adequately modelled. Deviation ranged from slight (dry season) to overestimation of belg season and a strong underestimation of kremt season. All in all the CFSR model could not adequately render rainfall patterns for Maybar.

Discharge simulation comparisons with WLRC data produced very good results: the three catchments could be modelled with *very good* performance ratings for RSR and

C5760

NSE except for the PBIAS, which was only *satisfactory* in the case of Maybar and *very good* for Andit Tid and Anjeni.

Discharge simulations with CFSR The WLRC rainfall data set resulted in three calibrated and validated discharge models while the CFSR data resulted in *unsatisfactory* performance ratings for the three catchments. Discharge modelling results yielded results in line with rainfall data comparison: Anjeni discharge was highly overestimated, Andit Tid discharge was underestimated and Maybar discharge had overestimation of belg discharge and underestimation of kremt discharge.

Sediment loss modeling with none. For the soil loss modeling the WLRC rainfall data and calibrated discharge resulted in two *satisfactory* (Andit Tid and Anjeni) and one *unsatisfactory* (Maybar) calibrations. For Andit Tid the model could render sediment loss adequately except for some peaks in the mid and late nineties. For Anjeni the model performed even better with a slight overestimation over the whole period. For Maybar calibration failed. out of three calibrated and validated models while none could be adequately calibrated or validated for the CFSR data set. The SWAT modelling showed that CFSR rainfall pattern and rainfall yearly total amount variations were so significant that SWAT model calibration could not adequately represent measured discharge and sediment yield.

Sediment loss performance ratings from simulations with CFSR rainfall data and calibrated discharge yielded in *unsatisfactory* results for the three catchments. The same deviation patterns observed in discharge calibration ensued in sediment loss calibration. Catchments with high discharge model overestimation resulted in high sediment loss overestimation and catchments with displaced seasonal discharge patterns resulted in displaced sediment loss patterns.

Andit Tid sediment loss modelling with CFSR data resulted in *unsatisfactory* results. The underestimation of the discharge modelling did not allow for satisfactory sediment loss calibration. The hydrograph (see Figure 4 on page C5763) shows that the general

C5761

underestimation of rainfall data lead to underestimation of discharge, which lead to reduced sediment loss modelling. Sediment loss modeling with CFSR data in Anjeni resulted in a constant overestimation of sediment loss and performance ratings were *unsatisfactory*. For Maybar the misplaced seasonal rainfall lead to higher discharge for belg and lower discharge for kremt, which resulted in a shift of sediment loss peaks from kremt to belg. Performance ratings were *unsatisfactory*.

Our results clearly show that adequate discharge and sediment soil loss modelling was not possible, in present case, with the CFSR datain present case. This suggests that SWAT simulations in small-scale watersheds in the Ethiopian highlands do not perform well with CFSR data in every case, and that sometimes there is no substitute for high-quality conventional weather data. Such weather data – with high spatial and temporal climatic data resolution – were available for the three small-scale catchments used in the study but are not in many other cases. In these other cases one should carefully check CFSR data against similar climatic stations with conventionally measured data. In addition, discharge and sediment soil loss modelling showed that usage of CFSR weather data not only resulted in substantial deviation in both total discharge and total sediment soil loss, but also in the seasonal rainfall pattern. The seasonal weather pattern is one of the major drivers of sediment soil loss and is especially pronounced in the Blue Nile Basin, with one long rainy season occurring as fields are ploughed and sowed. Thus, contrary to Dile and Srinivasan (**?**)previous studies for the Ethiopian Highlands, this study suggests that CFSR data may not be applicable in any case for small-scale modelling in data-scarce regions: the authors even suggest that outcomes of SWAT modelling with CFSR data alone for small-scale catchments may yield erroneous results which cannot be verified and may lead to wrong conclusions. Nonetheless, the advantage of CFSR data is its completeness over time, which would allow for comprehensive watershed modelling in regions with no conventional weather data or with longer gaps in conventionally recorded rainfall records.

C5762

Fig. 1. Map overview of Blue Nile (Abbay) Basin with the WLRC research stations, agro-ecological zones according to Hurni **?** and emplacements of CFSR stations.

Fig. 2. Monthly CFSR and WLRC rainfall distribution of all stations (1979–2010), Andit Tid, Anjeni, Maybarstation as boxplots with monthly rainfall distribution. CFSR data from 1979 to 2014 and WLRC data from 1981/1982/1984 to 2014. See Table 1 for details.

Fig. 3. Calibration and validation of Modelled SWAT discharge with compared to measured discharge (blue) for WLRC (violet) and CFSR (pink) input data and the 95 Percent Prediction Uncertainty (light blue). Data Each sub-figure contains the calibration and the validation period. Results are given in m^3/s.

Fig. 4. Calibration and validation of sediment Modelled SWAT soil loss with compared to measured soil loss (blue) for WLRC (red) and CFSR (gree) input data and the 95 Percent Prediction Uncertainty (light blue). Data Each sub-figure contains the calibration and the validation period. Results are given in tons (t).

C5763

Table 1. Study Description of study sites, model input data , sources and available time series and gaps. The subdivision of data relates to calibration and validation periods.

	Andit Tid	Anjeni	Maybar
Year of constructiona	1982	1983	1981
Location	9.815 N	10.678 N	10.996 N
	37.711 E	37.530 E	39.657 E
Size WLRC	477.3 hab	113.4 hac	112.8 hab
Size SWAT-delineation	466.78 ha	105.23	101.98
Altitudinal range	3040–3538 masl	2406–2506 masl	2530–2857 masl
		Data resolution	
DEM		2m	
Land use mapd		field scale	
Soil mape		5x5m	
Climatic data		Daily precipitation	
		Daily min. and max. temperature	
Hydrology data		Daily discharge	
Soil loss data		Daily soil loss	
Sources		SCRP/WLRC/CDE/own	
		Data availability	
	Andit Tid	Anjeni	Maybar
Precipitation data	1982–2004	1984–2004	1981–2001
	2006	2010–2014	2004–2006
	2010–2014		2010–2014
Temperature	1982–1993	1984–1993	1981–1993
	1997–2002	1998–2004	1995–1998
	2010–2013	2010–2013	2010–2013
Discharge	1982–1993	1984–1993	1981–1993
	1995–1997	1995–2000	1997–2006
		2011–2014	2010–2014
Sediment	1982–1993	1984–1993	1981–1991
	1995–1997	1995–1998	1995–2006
	2011–2014	2011–2014	2011–2014
		Subdivision of data	
Calibration	1984 – 1993	1986 – 1998	1983 – 2006
Validation	1994 – 1997	2010 – 2014	2008 – 2014

Table 2. General performance ratings recommended by Moriasi et al. (?)

Performance Rating	RSR	NSE	PBIAS	
			Streamflow	Sediment
Very good	$0.00 \leq RSR \leq 0.50$	$0.75 < NSE \leq 1.00$	$PBIAS < \pm 10$	$PBIAS \leq \pm 15$
Good	$0.50 < RSR \leq 0.60$	$0.65 < NSE \leq 0.75$	$\pm 10 \leq PBIAS < \pm 15$	$\pm 15 \leq PBIAS < \pm 30$
Satisfactory	$0.60 < RSR \leq 0.70$	$0.50 < NSE \leq 0.65$	$\pm 15 \leq PBIAS < \pm 25$	$\pm 30 \leq PBIAS < \pm 55$
Unsatisfactory	$RSR > 0.70$	$NSE \leq 0.50$	$PBIAS \geq \pm 25$	$PBIAS \geq \pm 55$

Table 3. Seasonal comparison of rainfall datatime series of daily rainfall amounts. Satisfactory performance ratings are highlighted in bold. Details for duration and gaps can be found in table 1

	Andit Tid (1982-2014)	Anjeni (1984-2014)	Maybar (1981-2014)
		Dry season	
		Oct-Nov-Dec-Jan-Feb-Mar	
RSR	1.68	3.55	1.3
NSE	−1.92	−12.9	−0.77
PBIAS	55.2	134.2	30.7 **30.7**
		Belg	
Andit Tid Anjeni Maybar		Apr-May	
RSR	1.31	2.48	0.85
NSE	−0.79	−5.42	−0.24
PBIAS	−39.4	106.1	24.3 **24.3**
		Kremt	
Andit Tid Anjeni Maybar		Jun-Jul-Aug-Sep	
RSR	3.23	7.0	2.03
NSE	−9.79	−50.09	−3.28
PBIAS	−69.2	128	−47.1

Table 4. Calibration and validation results , of monthly CFSR and WLRC modelled discharge and soil loss. Values that meet at least the "satisfactory" criteria are highlighted in bold

	Andit Tid		Anjeni		Maybar	
	CFSR	WLRC	CFSR	WLRC	CFSR	WLRC
Discharge - Calibration						
p-factor	0.49	0.71	0.49	0.92	0.41	0.61
r-factor	0.20	0.53	1.91	0.46	0.54	0.96
RSR	0.83	**0.46**	2.70	**0.37**	1.16	**0.53**
NSE	0.31	**0.79**	–6.27	**0.86**	–0.35	**0.72**
PBIAS	46.1	**3.1**	–226.0	**2.0**	29.6	**1.5**
Discharge - Validation						
p-factor	0.30	0.66	0.69	0.69	0.38	0.61
r-factor	0.29	0.54	1.41	0.57	0.52	1.11
RSR	0.74	**0.46**	1.24	**0.41**	0.98	**0.56**
NSE	0.45	**0.79**	–0.53	**0.83**	0.04	**0.74**
PBIAS	37.9	**9.6**	8.1	**–6.7**	–1.9	**–17.3**
Soil loss - Calibration						
p-factor	0.33	0.45	0.32	0.40	0.44	0.28
r-factor	0.19	0.59	1.30	0.65	4.47	0.28
RSR	1.02	**0.67**	1.01	**0.67**	2.55	0.84
NSE	-0.03	**0.64**	-0.02	**0.55**	–5.51	0.29
PBIAS	54.4	**–14.1**	**–33.9**	**–19.9**	180.5	**39.2**
Soil loss - Validation						
p-factor	0.30	0.39	0.38	0.38	0.23	0.15
r-factor	0.51	1.60	1.61	1.10	2.67	0.06
RSR	1.39	**0.68**	1.08	**0.62**	2.24	0.98
NSE	–0.94	**0.51**	-0.17	**0.62**	-4.04	–0.03
PBIAS	11.9	–64.3	**–30.5**	**–31.3**	–94.7	92.8

C5766