

Oct 14, 2014

Dear Prof. Ralf Merz,

We have revised the manuscript “hess-2014-157: Analyzing runoff processes through conceptual hydrological modelling in the Upper Blue Nile basin, Ethiopia “for Hydrology and Earth System Sciences (HESS). We wish to thank the editor and the reviewers for their thorough and useful suggestions. In our revision, we tried to account for every remark or suggestion made by the editor and reviewers. Below you find the comments from the editor and reviewers, the reply to the comments by the authors and a marked-up manuscript version showing the changes made. The manuscript has been updated according to the comments and changes are also highlighted in the manuscript by track changes. We hope that this revised version may satisfy the reviewers and the editor.

Kind regards,

Mekete Dessie

Comments from editors and reviewers and the corresponding feedback by the authors

I. Comments from the editor

Comment

1) Reviewer One questioned your objective of the paper and you answered:

„However, we believe that the innovation of this paper should not be sought in these aspects but rather in the fact that through a simplified model (i.e. a conceptual model), we are able to assess hydrological processes in the Lake Tana basin, which would be complex to do if a more sophisticated model were used (one would immediately encounter problems due to data scarcity). The insights that are gained learn a lot about the hydrology of this system, which up till now was less understood."

If you want to demonstrate that a simple model works better than a complex one, you should apply a complex model and compare that to your simple model. It would be nice to see, but I do not see that in the paper.

You agreed to the reviewer that type of model used does not fully allow to gain additional insight in hydrological responses but you argue that you gained (at least) some insights about the hydrology of this system, (i.e. catchment). This is not convincing for me but for me the main interesting point of the study and the reason why I would see a revised paper published is the following: In data scarce regions we have the problem that we have not enough data to run complex physically based models. But also it is difficult to run reliable conceptual models, because the lack of data for calibration or verification. So we need clever simple models that try to use all available information on hydrology. One such available information on hydrology is topography and I would guess that in region such as the upper Nile basin topography is a good proxy for the variability of most of the catchment characteristics. Hence I would like to see the story of the paper directed towards the question if a simple topography driven model structure is a good choice for such regions. What is the difference to a benchmark model that is based on different basic assumptions?

Reply from authors

We thank the editor for his constructive inputs to the paper and we tried to follow his suggestions.

a) Comparison with other models

We made comparisons with two bench mark models. A relatively complex model of Soil Water Assessment Tool (SWAT), developed by the United States Department of Agriculture (USDA) having many parameters, and the lumped model (FlexB) by Fenicia et al. (2008) were used. A brief description of the models and the performance comparisons are included in the revised manuscript: shown by track changes in the

manuscript: Pages 17 and 18, Line numbers 473 to 517; Fig.9 to Fig.12 on page 42, 43 and 44.

b) Objective of the paper

We fully agree with your suggestion. The objective of the paper is to analyze rainfall-runoff processes in the Upper Blue Nile basin, considering topography as a proxy for the variability of most of the catchment characteristics and to show that topography driven model structure and use of all available information on hydrology based on topography is a good choice for the Upper Blue Nile basin. We want also the paper to get published from this perspective. Accordingly, we have included additional clarification of the previous reply to Reviewer One, who questioned the objective of the paper (see details below in Section II under the comments from the Anonymous Referee #1 and the reply from authors).

Comment

2) Concerning the quality of the model and the analysis of model errors: Please add some more model performance indicators and parameter uncertainty studies, but just add more indicators is not too helpful as many indicators are highly inter-correlated. I think it would be helpful to discuss the simulation performance of some single events in much more detail based on hydrological understanding and to discuss some simulation results such as the modelled soil moisture variability. Are the patterns plausible? Are there some observations of hydrological variables which may support the model results? (e.g. variability of groundwater level which give hint to simulate groundwater recharge.)

Reply from authors

We accept the comment and we considered the following additional model performance indicators and parameter uncertainty studies.

a) Flow duration curves to illustrate the model performance on flow frequency simulation This was recommended by Hongkai Gao (reviewer 2) and we made a comparison of simulated and observed flow duration curves for the various models. This is shown in the revised manuscript on page 44 (Fig.11 and Fig.12).

b) Global sensitivity analysis

In addition to the local parameter sensitivity analysis, in the revised version global sensitivity analysis of the model parameters is included based on the comments of Anonymous Referee #1. This is shown on page 46 (Fig.15) in the revised manuscript.

c) Percent bias (PBIAS) as an additional model performance indicator

We have included percent bias (PBIAS) as an additional model performance indicator in the revised version of the manuscript. It measures the average tendency of the simulated data to be larger or smaller than the observations (Gupta et al., 1999). This is shown by track changes in the revised manuscript on pages 16 and 17, line numbers 458 to 463. The different model performance results are shown on the revised Table 3, page 36.

d) Verification of model results with some observations of hydrological variables and previous studies

The editor suggested to investigate some observations of hydrological variables which may support the model results. In line with this, we tried to find groundwater level variation data. However, such data generally lacks in the study region. But studies by Kebede (2013) show that one of the study catchments (Gilgel Abay catchment) is identified as high groundwater recharge catchment. We have also witnessed this from several big springs in the catchment, for example one of such big springs used as a source of water supply for Bahir Dar town is in this catchment (Photo shown in the manuscript, page 45, Fig.13). This is in line with the results of this model for this catchment (baseflow takes the larger proportion in this catchment based on the model result). Therefore, these additional observations are included in the revised manuscript, shown by track changes on page 22, line numbers 634 to 637.

II) Comments from the Anonymous Referee #1

General comments

The paper develops a rainfall-runoff model of medium complexity, distinguishing between groundwater, direct runoff and interflow; and splitting the catchments into three using topography. The parameter estimation uses a combination of calibration and estimation of parameters based on soil properties. The work is a brave attempt to develop and test a model for an area that suffers from limited flow, precipitation and hydrological properties data. However the paper does not really provide significant advances in

understanding hydrological responses or innovation in modeling. The quality of model outputs is declared good, but this is arguable and a detailed analysis of model errors has not been reported.

Reply from Authors

The framework used may not be innovative in modeling, as many conceptual models have been developed and used. However, we believe that the innovation of this paper should not be sought in these aspects but rather in the fact that topography driven model structure and use of all available information on hydrology based on topography is a good choice for the Upper Blue Nile basin, and through a simplified topography driven model (i.e. a conceptual model), we are able to show and assess hydrological processes in the Lake Tana basin and compare such topography driven model with other models (SWAT and Flex_B models). The insights that are gained learn a lot about the hydrology of this system, which up till now was not well understood. We hope that the reviewer may agree and that the objective of the paper is sufficiently novel to be published from this perspective.

Concerning the quality of the model outputs and the errors associated with them, they can be evaluated based on model performance indicators like the Nash-Sutcliffe Efficiency (NSE), Root Mean Squared Error (RMSE), the coefficient of determination (R^2), Bias, etc. The visual comparison (Plots) can also give an overall judgment. The authors used NSE, RMSE, PBIAS and R^2 in addition to the plot comparison of the observed and simulated outputs. The statistical results for NSE and R^2 were greater than 0.7, which shows the good performance of the model.

Model uncertainty arises from a variety of sources, such as model parameterization, process representation, equifinality, and calibration accuracy. We agree on the importance of a detailed analysis of model errors. In this regard, we carried out more than 2500 iterations using the Particle Swarm Optimization (PSO) technique to reach to the optimal model parameters and minimize calibration errors. We made sensitivity analysis of the model parameters to identify the important parameters and rank parameters that have significant impact on the model outputs (Fig.14 and Fig. 15 in the revised version of

the paper). Each and every step of the model development has been discussed with the possible limitations.

We agree with the reviewer to further investigate model errors and model clarity based on the relevant comments suggested by the reviewer that will be shown on the response to the more detailed comments of the reviewer and the manuscript is revised based on these and other comments.

Comment

While recognizing the data issues, the authors claim that periods of data are relatively high quality; however this has not been shown, for example the reader cannot judge the degree of rainfall and flow data errors. The model is rather complex given the data restrictions, shown by the sensitivity analysis. The conclusions about hydrological processes cannot be justified given the data issues, the use of text book values of parameters of unknown applicability here, and the apparent limited performance of the model.

Reply from authors

The authors' claim of relatively high quality discharge data for the calibration of the models emancipates from the data acquisition methodologies used. Unlike the previous water level measurement of twice a day using staff gauges, in this case, the water level measurements were made using Mini-Divers, automatic water level recorders (every 10 min.), and manual readings from a staff gauge (three times a day, at 7 a.m., 1 p.m. and 6 p.m.). Moreover, rating curves were produced using a recent survey at river cross sections. We do not dare to say that the data is absolutely perfect. However, there is a significant improvement to what was available before. We agree that we did not show the rating curve plot in the manuscript together with the range of levels to which the rating curves were applied. In the revised version, we have shown and elaborated more on the accuracy of the rating curves that were established (see below under the detailed comments).

With respect to the rainfall data, the authors used the available rainfall data from the rain gauges in and around the study catchments and discussed the accuracy of the data. In the revised version, we have included a figure to show the location of the rain gauges (see below under the detailed comments).

The reviewer commented on the limitations of the use of text book values of parameters of unknown applicability. We agree with the comment and understand the limitation. Unfortunately, the study area faces high scarcity of data. In such instances, it is normal to consider data from relevant sources with caution. Accordingly, porosity and field capacity of the soils were derived from the study area soil texture data based on literature recommendation. Similar procedures are followed for the saturated hydraulic conductivity for the deep percolation by identifying the likely aquifer formation of the study area. We tried to optimize the literature recommended values to get better performance of the model (again, we refer to answers to more detailed comments of the reviewer).

Comment

There are various gaps in the description of the method, as I explain in my comments below. Overall, I am not confident that this model or the conclusions made about processes are justified, and all the evidence points to the model being over-complicated. The authors may have been better using a stricter application of the methods of Fenicia et al. 2008 to gradually build up the complexity of the model to the justified level, with more explicit attention to errors in inputs and outputs. Below are a few more detailed comments that may help in a revised version; however in my opinion the aims and approach need re-thought.

Reply from authors

We note the need for the clarification of some of the descriptions of the method that were not clear to the reviewer and to the other readers. These will be addressed in reply to the detailed comments of the reviewer point by point. We appreciate the suggestion of the reviewer to use a stricter application of the methods of Fenicia et al. (2008) in the building up of the model and we have considered it as a bench mark model to compare the performance of our topography driven model and this is accounted in the revised version of the paper.

More detailed comments (Specific comments)

Comment

5293, 3. Model is modified from what? Not clear what is being modified. 5293, 10.

Differently from what?

Reply from authors

As explained on page 5293 in the discussion paper, our model is developed based on the works of Jothityangkoon et al. (2001), Krasnostein and Oldham (2004) and Fenicia et al. (2008). However, we made modifications on some of their model concepts and equations. The major modification (variations) made in this paper can be seen from three cases.

i. The catchment bucket representation concept

The works of Jothityangkoon et al. (2001), Krasnostein and Oldham (2004) and Fenicia et al. (2008) considered the catchment bucket to consist of the soil reservoir and the groundwater reservoir. In our model, in addition to the soil and groundwater reservoirs, we included the other component that considers the impermeable part of the catchment. So, the catchment is divided into the soil and groundwater reservoirs part and the impermeable part. As we know that our study catchments have impermeable surfaces (with little or no soil cover), we needed to consider this separately in the rainfall-runoff process representation of the conceptual model.

ii. Soil surface Catchment characterization

Catchment characterization was made based on topography. Hence, the catchment was divided as steeply, hilly and level and the input data to the model were determined accordingly. This is because the model is not a fully distributed model and hence topography is considered as a major landscape characteristics to determine the other catchment features required for the model.

iii. Percolation to the groundwater table and hydraulic conductivity for the interflow

In the soil and groundwater reservoirs, we modified the equations of deep percolation and hydraulic conductivity for the interflow component of the soil reservoir.

For example, in the case of Fenicia et al. (2008), percolation to groundwater reservoir is modelled as:

$$P_s = P_{\max} \left(\frac{S_u}{S_{fe}} \right)$$

where P_{\max} is maximum percolation, S_u soil storage and S_{fe} is maximum soil storage. For details, refer to Fenicia et al. (2008). In our paper, this is conceptualized differently as given in the paper. Moreover, we made a distinction between the upper and deep soil hydraulic conductivities such that the hydraulic conductivity for the interflow component of the soil reservoir is dealt separately in our modelling approach. Details of the equations are shown in the paper.

It was to reflect these aforementioned points that the authors used the word “modified” on page 5293 in the discussion paper. But taking the comment into consideration, we have revised the manuscript and rewritten to be clearer. This is shown by track changes in the revised manuscript on page 6, line numbers 150 to 160. The two figures (Fig.2 and Fig.3 in the first manuscript have now been merged into one figure (Fig.2 in the revised manuscript).

Comment

Eq 8 and 9. Equations applicable at hill-slope scale? Needs some further justification.

Reply from authors

Equations 8 and 9 are universal equations. Equation 8 is a universal equation for velocity.

$$velocity = \frac{Displacement}{Time}$$

This equation is applicable anywhere as long as the displacement and time are determined accurately. In our case, the displacement is assumed to be the average slope length of the catchment (distance subsurface flow travels) and the time is the subsurface flow response time (the time the subsurface flow takes to reach to its exit).

Equation 9 is Darcy’s equation that describes the flow of a fluid through a porous medium. It is applicable for a porous medium as long as the flow is laminar (which generally is the case in the case of a natural groundwater flow). Similar application of Darcy’s Law to the groundwater aquifer within a planar hillslope has been indicated in Jothityangkoon et al. (2001).

Comment

How can all these parameters be justified? Why are there only seven – they need estimated for each of the three slope classifications?

Reply from authors

The model parameters are justified from calibration, validation, sensitivity analysis and performance studies of the model. From the model development, we identified seven parameters and these were calibrated using the Particle Swarm Optimization (PSO) technique. From the local model sensitivity analysis, we showed that three of the seven model parameters are hardly sensitive and there is little confidence in the model's correspondence with these parameters and they can be reduced without appreciable impact on the model (This is shown on pages 5306 and 5307 in the discussion paper under Section 6.3). However, the global sensitivity analysis (Fig.15 in the revised manuscript) still shows that all the parameters are sensitive.

The seven parameters for the three slope classifications are reached as follows.

Parameters for the recharge (α_1 and α_2)

In the three slope classification, α_1 is to consider for the recharge from the steeply slope into the medium slope surface and α_2 is for the recharge from the medium slope surface into the flat slope surface. There is no parameter for the steeply slope surface since there is no surface that recharges it. So, there are two parameters for the three slope classifications.

Parameter for the impermeable surface of the catchment (λ)

In this case, the catchment is divided into two surfaces (impermeable surface with no or little soil cover and the soil surface). The parameter λ is introduced to represent the fraction of impermeable surface within the total catchment and this part of the catchment is not classified as steeply, medium and flat slopes since the classification of this part of the catchment into such classes is not important. So we have one parameter.

The parameters β , γ , k_1 and $K_{s,u}$

These parameters β and γ are introduced to account variability of permeability and deep percolation of soil with soil water storage. k_1 relates discharge and storage for the ground water and $K_{s,u}$ is the saturated hydraulic conductivity in the upper soil layer. We assumed that these parameters are less influenced by topography and each model parameter is assumed to be same for each slope classification of the catchment. Moreover, it looks quit inconsistent to separate the groundwater system in the catchment and we preferred all the three slope based classified catchments to share the same groundwater reservoir.

In this perspective, we will have in total seven parameters for the three slope classifications. We agree with the reviewer that we did not provide this explanation in the paper. In the revised version of the paper, clarifications have been added. These are shown by track changes in the revised manuscript on pages 11 and 12, line numbers 304 to 328.

Comment

5300, 1-4. Local relevance of the text book values? Really the textbook should provide ranges, which are fed into calibration (further increasing the calibration problem).

Reply from authors

We estimated porosity and field capacity of the soils and the saturated hydraulic conductivity for the deep percolation from literature recommendations. We agree with the comment and understand the limitation. However, owing to the high scarcity of data in the study area, it still remains necessary to consider data from relevant sources with caution. Hence, the soil texture class data of the catchments were used to estimate porosity and field capacity of the soils. From studies by Cosby et al. (1984), we note that soil texture is closely related to the variability in soil moisture parameters (porosity and field capacity of the soils). Similar procedures are followed for the saturated hydraulic conductivity for the deep percolation in that its value is estimated by identifying the likely aquifer formation of the study area. In fact, the literatures provide a range of values. In such instances, we considered average values and we tried to optimize the values by iterating to get the best model performance results.

Comment

5300, 12. We need to see location map of these gauges – as precipitation is the key input –and know something about their accuracy. Was the PE spatially variable?

What assumptions have been made about stream flow routing and stream-groundwater interactions?

Reply from authors

The location map of the rain gauges have been provided in the revised version of the manuscript. This is shown in the manuscript on page 41, Fig.7.

Generally rainfall data are obtained on daily basis. The data for most of the stations are consistent and continuous, particularly for first class stations like Dangila, Adet and Debretabor. However, we encountered gaps in some stations like Sekela Station for some periods in the year. In such instances, only the rainfall data from the other stations is considered. As discussed in the paper, most of the rainfall stations in Gilgel Abay catchment are installed at the water divides and there is no station in the middle of the catchment. In this regard, the Gumara catchment is with higher density of rainfall stations. PE is also spatially variable. Based on the comment, the above discussions on the rainfall data have been included in the revised manuscript to let the readers know the accuracy of the rainfall data. This is shown by track changes in the revised manuscript on page 14, line numbers 384 to 391.

In this paper of hydrological modeling, stream-groundwater interactions are assumed to be minimal and the groundwater is assumed to recharge the streams from deep percolation of rainfall on the catchments that produces baseflow of the rivers/streams. The storage effect of the streams when considered on the basis of average daily flows of the streams is assumed to be negligible and hence streamflow routing was not considered for such smaller streams. This clarification has been shown in the revised manuscript by track changes on page 12, line numbers 324-328.

Comment

Figure 6. Does not look like great performance to me. Needs some more insightful plots to elucidate magnitude and nature of errors.

Reply from authors

The quality of the model outputs and the errors associated with them are usually evaluated based on model performance indicators like the Nash-Sutcliffe Efficiency (NSE), Root Mean Squared Error (RMSE), the coefficient of determination (R^2), Bias, etc. The visual comparison (Plots) can also give an overall judgment. The authors used NSE, RMSE and R^2 in addition to the plot comparison of the observed and simulated outputs. The statistical results for NSE and R^2 were greater than 0.7, which shows the good performance of the model. The plots of the simulated and observed discharges of Fig. 6 in the paper can show this, but it is true that there are some deviations of the simulated discharge from the observed ones at some points in the time series. There can be various reasons for this, as explained in the paper. One instance can be the rainfall data. As can be seen from Fig.7 in the revised manuscript, there are no rain gauges in the middle of the Gilgel Abay catchment and given the high spatial variability of the rainfall in the whole Blue Nile basin, this can create its own uncertainty on the model performance. Fig. 6 (now Fig.10 in the revised manuscript) has been updated to include comparisons with other benchmark models and also to give better visualization for the magnitude and nature of errors. This is shown in the revised manuscript on page 43, Fig.10.

Comment

Eq 20, 21. Authors claim that the gauged flow data are high accuracy – it would be useful for the reader to see the rating curves, together with the range of levels to which the rating curves were applied.

Stochastic optimization implies the stochastic nature of the input errors were considered? How are rainfall errors considered? Stochastic optimization gives stochastic outputs, which is misrepresented, or at least under-utilized, by reporting optimal parameter values.

Reply from authors

The rating curves together with the range of levels to which the rating curves were applied have been provided in the revised version of the manuscript (shown in Fig.8 in the revised manuscript on page 41).

In the model calibration, we did not use stochastic optimization that depends on one or more of the input data subject to randomness. The input data (for example rainfall) are observed data (soil data have been estimated from relevant sources when observed data are absent). For the model calibration, we used the particle Swarm Optimization (PSO) technique. PSO optimizes a problem by having a population of candidate solutions, here particles, and moving these particles around in the search-space according to simple mathematical formulae over the particle's position and velocity. Each particle's movement is influenced by its local best known position but, is also guided toward the best known positions in the search-space, which are updated as better positions are found by other particles.

Comment

5302, 5. Why only 7 parameters? Each catchment was split into different runoff production units to represent variation in catchment properties using topography, so why not 21 parameters?

Can the splitting into three areas be shown on a map, e.g. using color coding?

Reply from authors

This comment, “why 7 parameters? “, is similar to a comment given by the reviewer above (Page 10). The explanation is given there.

Using color coding, the splitting of the study catchments into three using topography is shown in the revised manuscript on page 39, Fig.4.

Comment

5304, 5. Figures 5 and 6 do not show this very well. Some more insightful plots about the errors are needed. In Figure 5, it seems there are some rather serious errors. E.g. the wetting up period deserves some discussion, In Fig 6, I cannot really see the nature or magnitude of the errors; however there are clearly some systematic errors that need

critical discussion. The flow regime / climate in the validation period seems quite similar to the calibration period, so comparable performance is expected.

Validation should ideally test the model to breaking point.

Reply from authors

This comment is similar to one of the comments above. Figures 5 and 6 in the discussion paper are plots of predicted and observed discharge and precipitation of the Gumara and the Gilgel Abay catchments for the calibration and validation periods. We still believe that the plot simulates well the general behavior of the observed streamflow hydrographs. Generally, the errors do not seem to have a trend. However, we notice that the model errors tend to increase during wetting up periods in most instances. Initially, the soils are relatively dry and most of the rainfall during the beginning of the rainy season is not effective to produce runoff in the model as the soil reservoir has to be filled first to generate the faster component of the runoff. In the model, mostly average conditions prevail owing to average input data (rainfall, soil, catchment characteristic, etc.). Besides model uncertainties, the rainfall data quality can also affect the model performance, mainly in the case of the Gilgel Abay catchment. The comments and discussions on the model errors have been incorporated in the revised manuscript (shown by track changes on page 19, line numbers 542-548).

The flow data used for validation is from 2000-2005 (6 years data) and for calibration is the 2011 and 2012 years data for Gumara and the 2012 data for Gilgel Abay. Each year data is different, depending on the climate of the year and catchment conditions. However, the trend is similar each year such that there is high discharge in the rainy season (June to September) and a decreasing trend of discharge after September in line with the dry season. The 6 years discharge data are considered sufficient to run validation tests.

Comment

5304, 14. I didn't follow what this meant. Which data are averaged over the year?

Reply from authors

The modelled discharges appear to be less variable over time than the observed discharges. Therefore, the sentence on page 5304, line 14 in the discussion paper is to explain this. We used average daily rainfall data, average soil data (e.g. porosity, field capacity, and soil depth), average catchment characteristics data (e.g. slope, slope length) to mention some for the model inputs. Hence, this averaged condition may be one source of error such that the model may not exactly mimic extremes like peak discharges. We have included these clarifications in the revised paper (shown by track changes on page 20, line numbers 554-558).

Comment

5305-5306. I don't see how these observations are meaningful given the errors in the model. There seem to be large errors in the flow peaks, so the model cannot be used as a basis for concluding upon importance of direct runoff.

Reply from authors

Generally, the model performance and the model errors have to be explained based on commonly employed model performance indicators. In modeling, the usual practice is that if the model performance indicator results are above a recommended value (for example > 0.5 for NSE and R^2), then the model is taken as acceptable and model results are considered meaningful. Our modeling approach is not different from this. The authors used NSE, RMSE and R^2 in addition to the plot comparison of the observed and simulated outputs. The statistical results for NSE and R^2 were greater than 0.7, which shows the good performance of the model. The plots of the simulated and observed discharges of Fig.9 and Fig.10 (in the revised paper) can show this, but it is true that there are some deviations of the simulated discharge from the observed ones at some points in the time series. This is the limitation of the model. To build more confidence on the model results, we have also used benchmark models (SWAT and Flex_B models) to evaluate the performance of the model. All models show good performance in the calibration period. We understand that the model results are very important clues to understand the runoff processes in this data scarce region of the Upper Blue Nile basin

and for the general water resources planning in the area. But we do not dare to say that the results are absolutely perfect.

Comment

Figure 7. Local sensitivity analysis – value of this is unclear give high uncertainty in parameter values. Global analysis would be more useful.

5307 - Sensitivity analysis results support the view that the model is too complex; or at least components of it are too complex

Reply from authors

The optimal model parameters are obtained using the particle Swarm Optimization (PSO) algorithm, which performs global analysis. In Figure 7 (in the discussion paper), we investigated the local sensitivity of each model parameter when the parameter value is different from the optimal one, keeping the other model parameter values constant (equal to the global optimal value). Based on the comment, we also made global sensitivity analysis and results are depicted in Fig.15 in the revised manuscript. Further discussions are also shown in the revised manuscript (shown by track changes on pages 23 and 24, line numbers 666-669 and 673-681).

Comment

5307, 23. This is not an encouraging performance. Probably a two or three-parameter model could achieve this.

Reply from authors

As shown in the paper, the results of NSE and R^2 , for the direct parameter transferability test to other catchment were 0.58 and 0.6 respectively. The authors' suggestion of encouraging performance is based on these results. As it can be seen, the results are not bad. But the authors still stressed the need for further tests on similar catchments, as shown in the paper. We understand that various types of models with different number of parameters can be considered. Probably a two or three-parameter model could also give acceptable performance results, but such types of models are black box types and may not help for understanding the runoff processes in a particular catchment.

Comment

5309, 10. This conclusions is not justified from the results. The effect of the topographically-based division of the catchment has not been explored at all?

Reply from authors

In the paper, the effect of the topographically-based division of the catchment is reflected mainly with respect to the input data to the model. Since the model was not a fully distributed model, it was necessary to use average catchment data. For this, we used topography as a proxy for the variability of most of the catchment characteristics like soil data (soil depth, porosity and field capacity) and undertake catchment classification. The explanations on page 5309, lines 9 and 10 in the discussion paper are to emphasize this role of topography in the model. Moreover, we also showed the effects of topography on runoff and we obtained that hillslopes (medium and steep slope areas) generated almost no direct runoff as saturated excess flow.

In line to the comment, we have further elaborated the discussion of the effect of the topographically-based division of the catchment in relation to the benchmark models. This is shown by track changes in the revised manuscript on pages 20 and 21, line numbers 560-585.

III) Comments from Hongkai Gao (Reviewer 2)

General comments

This manuscript is very interesting for me. The writing is clear and concise. The authors applied a new modelling framework (Savenije, 2010) to do runoff production area classification by topography information. Slope was used as criteria to do the classification.

The model structure is simple but reasonable. The number of free parameters is also limited to 7, which dramatically reduces the equaifinality. Although the model did not apply the normally used curve in soil reservoir to represent the distribution of water storage capacity (Zhao, 1992), the results are also excellent, which is intriguing for me.

The authors cooperated topographic information and soil texture information into the model. The average slope gradient and slope length are parameterized into the conceptual model by semi-empirical relations. The porosity and field capacity of soil are used to

determine the storage capacity. All the functions are clear and reasonable for me. But there are still several things needs to be clarified.

Comment

1. Following the comments from Prof. Merz and Anonymous Referee #1, I also think a benchmark model is necessary in this paper to illustrate the better performance or transferability of this modelling approach than traditional lumped models which neglect the heterogeneity of catchments. Not only hydrograph, but also the flow duration curve shall be shown to illustrate the model performance on flow frequency simulation.

Reply from authors

We accepted the comment to use a benchmark model as an alternative to compare the results with our topography driven model. We chose the lumped model with lumped input data (FlexB) by Fenicia et al. (2008) and a relatively complex model of Soil Water Assessment Tool (SWAT), developed by the United States Department of Agriculture (USDA) having many parameters, as benchmark models to assess the benefits and performance of topography-driven semi-distributed modelling of this paper. A brief description of the models and the performance comparisons are included in the revised manuscript: shown by track changes in the manuscript: Pages 17 and 18, line numbers 473-517 and pages 20 and 21, line numbers 560-585. Model simulation comparisons have been also shown in Fig.9 and Fig.10.

We thank Hongkai Gao for his suggestion to show the flow duration curve to illustrate the model performance on flow frequency simulation. Accordingly, we have shown the flow duration curve of the models (including the benchmark models). The flow duration curves for the different models have been shown in the revised manuscript on page 44, Fig. 11 and Fig.12.

Comment

2. The model structure is not very clear for me, although it is mentioned in the text and shown in Figure 2 and Figure 3. I suggest the authors show the inter-link between different runoff production areas in one figure, which could be clearer and easier to follow.

Reply from authors

We agree with this comment and we have worked further on the model structure to make it clearer, and the clarifications have been provided in the revised version of the manuscript (shown by track changes on page 6, line numbers 150-161). The link between different runoff production areas (Figure 2 and Figure 3 in the discussion paper) have been combined into one figure in the revised manuscript (Fig.2).

Comment

3. Please show the slope map, classification map obtained by topography criteria, and the soil map, from which we can easily see the heterogeneity among different catchments.

Reply from authors

We agree with this comment and we have included the slope and soil maps in the revised version of the manuscript. The slope map is shown on page 39, Fig.4 and the soil map on page 40, Fig.5 and Fig.6 in the revised manuscript.

Comment

4. In Section 6.3, for the transferability test, I think the authors should do more discussion to clarify why this modelling approach can get good transferability. The authors could refer our newly published paper (Gao et al., 2014) about the application of the FLEX-Topo modelling approach in the Heihe river basin in China, in which paper the model performance comparison and transferability with several benchmark models are test.

Reply from authors

We agree with Hongkai Gao suggestions to do more discussions to clarify why this modelling approach can get good transferability. We also thank Hongkai Gao for his recommendation to the valuable reference (Gao et al., 2014). We have included the additional discussions in the revised version of the paper (shown by track changes on pages 24 and 25, line numbers 704-716).

Minor comments:

Comment

1. Perhaps I have missed something, do the different hydrological components have isolated groundwater or they share the same groundwater reservoir?

Reply from authors

They share the same groundwater reservoir since it looks quit inconsistent to separate the groundwater system in these relatively small catchments and we preferred all the three slope based classified sub-catchments to share the same groundwater reservoir. For this clarifications are added in the revised manuscript (shown by track changes on pages 11 and 12, line numbers 304-328).

Comment

2. Equation 12. Why the saturated hydraulic conductivity of deep soil layer ($K_{s,e}$) is not a free parameter in Table 2? How did the authors determine the $K_{s,e}$?

Reply from authors

It can be a free parameter. But the authors' interest is to reduce the number of parameters in the model formulation and use more knowledge available from observation and data as much as possible to reduce the equaifinality and increase chance of model transferability. In this perspective, we preferred to estimate the saturated hydraulic conductivity of deep soil layer ($K_{s,e}$) by identifying the likely aquifer formation of the study area (colluvial mantle on top of the igneous rock) and using ranges of conductivities given by Domenico and Schwartz (1990) for the different aquifers.

Comment

3. Equation 23 and 24. Where is i in these equations?

Reply from authors

We thank Hongkai Gao for this remark. Equations 22, 23 and 24 are updated to show i in the equations in the revised version of the manuscript (shown on page16, Equations (22, 23 and 24).

Comment

4. Table 1. Is 'flat' more suitable than 'level'?

Reply from authors

We think that the two words are synonymous.

Comment

5. Table 1. Are field capacity and porosity parameters or input data? If they are input data, how did you get these information in catchment scales? Please clarify this point.

Reply from authors

Field capacity and porosity are input data. As we tried to explain in Section 4.2 in the manuscript, page 13 in the revised manuscript, the porosity and field capacity of the soils were derived from the soil texture based on the work of McWorter and Sunada (1977). We determined the dominant soil textures of the study catchments (Table 1) from soil map of the Abay River Basin integrated master plan study BCEOM (1998a). Average values of the porosity and field capacity of the soils were considered at catchment scale from the ranges of values recommended by McWorter and Sunada (1977) based on the relevant soil texture in each catchment category classified based on slope.

Comment

6. Table 2. Why lambda is a parameter? To my point view, you can determine the proportion of impermeable surface by soil map. Is it possible?

Reply from authors

We agree with Hongkai Gao's idea of the possibility of determining the proportion of impermeable surface from the soil map. However, currently the available soil map of the study areas is not with such details. The available soil maps (Fig.5 and Fig.6 in the revised manuscript on page 40) do not differentiate the impermeable portions of the catchments, making difficult to know the impermeable surfaces from such maps. So, we have to represent it through a model parameter lambda.

Analyzing runoff processes through conceptual hydrological modelling in the Upper Blue Nile basin, Ethiopia

Mekete Dessie^{1,2*}, Niko E.C. Verhoest², Valentijn R.N. Pauwels⁵, Teshager Admasu^{3,4}, Jean Poesen⁶, Enyew Adgo³, Jozef Deckers⁶ and Jan Nyssen⁴

¹ School of Civil & Water Resources Engineering, Bahir Dar University, P.O.Box 430, Ethiopia

² Laboratory of Hydrology and Water Management, Ghent University, Coupure links 653, B- 9000 Gent, Belgium

³ College of Agriculture & Environmental Sciences, Bahir Dar University, P.O.Box 79, Ethiopia

⁴ Department of Geography, Ghent University, Krijgslaan 281 (S8), B-9000 Gent, Belgium

⁵ Department of Civil Engineering, Monash University, Clayton, Victoria, Australia

⁶ Department of Earth and Environmental Sciences, KU.Leuven, Belgium

* Corresponding author: MeketeDessie.Wossenie@UGent.be

Abstract

Understanding runoff processes in a basin is of paramount importance for the effective planning and management of water resources, in particular in data scarce regions of the Upper Blue Nile. Hydrological models representing the underlying hydrological processes can predict river discharges from ungauged catchments and allow for an understanding of the rainfall-runoff processes in those catchments. In this paper, such a conceptual process-based hydrological model is developed and applied to the upper Gumara and Gilgel Abay catchments (both located within the Upper Blue Nile basin, the Lake Tana sub-basin) to study the runoff mechanisms and rainfall-runoff processes in the basin. Topography is considered as a proxy for the variability of most of the catchment characteristics. We divided the catchments into different runoff production areas using topographic criteria. Impermeable surfaces (rock outcrops and hard soil pans, common in the Upper Blue Nile basin) were considered separately in the conceptual model. Based on model results, it can be inferred that about 65% of the runoff appears in the form of interflow in the Gumara study catchment, and baseflow constitutes the larger proportion of runoff (44-48%) in the Gilgel Abay catchment. Direct runoff represents a smaller fraction of the runoff in both catchments (18-19% for the Gumara, and 20% for the Gilgel Abay) and most of this direct runoff is generated through infiltration excess runoff mechanism from the impermeable rocks or hard soil pans. The study reveals that the hillslopes are recharge areas (sources of interflow and deep percolation) and direct runoff

20 as saturated excess flow prevails from the flat slope areas. Overall, the model study suggests that identifying the catchments into different runoff production areas based on topography and including the impermeable rocky areas separately in the modeling process mimics well the rainfall-runoff process in the Upper Blue Nile basin and brings a useful result for operational management of water resources in this data scarce region.

25 **Key words:** interflow, direct runoff, baseflow, rainfall-runoff, Blue Nile

1 Introduction

The Upper Blue Nile basin, the largest tributary of the Nile River, covers a drainage area
30 of 176 000 km² and contributes more than 50 percent of the long term river flow of the Main Nile (Conway, 2000). The basin (Fig.1a) drains the central and south-western highlands of Ethiopia. The Ethiopian government is pursuing plans and programs to use the water resource potential of the basin for hydropower and irrigation in an effort to substantially reduce poverty and increase agricultural production. The Grand Ethiopian
35 Renaissance Dam near the Ethiopian–Sudan border is currently under construction and several other water resource development projects are underway in its sub-basins.

Owing to such rapidly developing water resource projects in the basin, there is an increasing need for the management of the available water resources in order to boost
40 agricultural production and to meet the demand for electrical power. Sustainable planning and development of the resources depend largely on the understanding of the interplay between the hydrological processes and the availability of adequate data on river discharges in the basin. However, the available hydrological data are limited (for example, presently about 42% of the Lake Tana sub-basin, source of the Blue Nile, is
45 gauged by the Ministry of Water Resources of Ethiopia). Furthermore, research efforts performed so far in the Upper Blue Nile basin with respect to the basin characteristics, hydrology and climatic conditions are scanty and fragmented (Johnson and Curtis, 1994; Conway, 1997; Mishra and Hata, 2006; Antar et al., 2006). Hydrological models that allow for a description of the hydrology of the region play an important role in predicting
50 river discharges from ungauged catchments, and understanding the rainfall-runoff processes in the catchments to enhance hydrological and water resources analysis. As

such, a number of models have been developed and applied to study the water balance, soil erosion, climate and environmental changes in the Blue Nile basin (e.g. Johnson and Curtis, 1994; Conway, 1997; Mishra and Hata, 2006; Kebede et al., 2006; Kim and Kaluarachchi, 2008; Collick et al., 2009; Steenhuis et al., 2009; Tekleab et al., 2011
55 Tilahun et al., 2013).

The Soil and Water Assessment Tool (SWAT) and the Hydrologiska Byråns Vattenbalansavdelning Integrated Hydrological Modelling System (HBV-IHMS) models have been applied in the basin (Setegn et al., 2008; Wale et al., 2009, Uhlenbrook et al.,
60 2010). The SWAT model is based on the Soil Conservation Service (SCS) runoff curve number approach, where the parameter values are obtained empirically from plot data in the United States with a temperate climate. Liu et al. (2008) studied the rainfall–runoff relationships for the three Soil Conservation Research Project (SCRCP) watersheds (Hurni, 1984) in the Ethiopian highlands and showed the limitations of using such models,
65 developed in temperate climates, in monsoonal Ethiopia. Adjusted runoff curve numbers for steep slopes with natural vegetation in north Ethiopia were reported by Descheemaeker et al. (2008).

Using a simple runoff-rainfall relation to estimate inflows to the Lake Tana from ungauged catchments, Kebede et al. (2006) computed the water balance of Lake Tana.
70 However, hills and floodplains were not differentiated in their simplified runoff-rainfall relations. Mishra et al. (2004) and Conway (1997) developed grid-based water balance models for the Blue Nile basin, using a monthly time step, to study the spatial variability of flow parameters and the sensitivity of runoff to climate changes. In both models, the role of topography was not incorporated, and in the model of Conway (1997) soil
75 characteristics are assumed spatially invariant. Very few of the models discussed above attempted to identify the catchments into different hydrological regimes based on the relevant landscape characteristics to study the runoff mechanisms and the hydrological processes in the basin. Landscape characteristics can lead into conceptual structures and relationships or the conceptual hydrological models can benefit from them (Beven,
80 2001). Istanbuloglu and Bras (2005) considered topography as a template for various landscape processes that include hydrologic, ecologic, and biologic phenomena. This is more appealing to the Ethiopian highlands, in particular to the Upper Blue Nile basin, as

farming and farm drainage methodologies, soil and water conservation works, soil properties, vegetation, drainage patterns and density, and even rainfall are much linked to topography in the Ethiopian highlands. Therefore, it remains necessary to investigate the hydrological processes in the Blue Nile basin taking topography as a proxy for the variability of most of the catchment characteristics. The objective of this paper is to study runoff mechanisms in the Upper Blue Nile basin using topography as the dominant landscape component and classify a catchment (as steep, medium and flat slope areas) into different runoff production areas. The study tries to identify the dominant rainfall-runoff mechanism on the hillslopes (steep and medium slop areas) and the valley bottoms (flat areas). A considerable portion of the mountainous areas in the Upper Blue Nile basin consists of impermeable rocks and hard soil pans leading to a different runoff processes. This paper further investigates the contribution of such landscapes in the rainfall-runoff process by including a class for these impermeable rock and hard soil surfaces in the conceptual hydrological model. This approach is not so far tested in the Upper Blue Nile basin. However, similar methodologies to the conceptual hydrological model development are discussed by Savenije (2010). Furthermore, it is necessary to obtain better quality river discharge data in the basin. In this paper, we will face all these challenges. The conceptual hydrological model for the rainfall-runoff studies of the basin is calibrated using good quality discharge data obtained from recently established measurement stations. These outcomes positively add to the existing knowledge and contribute to the development of water resources plans and decision making in the basin.

2 Description of study catchments

The study catchments (Fig. 1b), where the model developed is applied, are located in the Lake Tana basin, the source of the Blue Nile River. The Lake Tana basin, located in the north-western Ethiopian highlands, with a catchment area of 15077 km² (including the lake area), consists predominantly of the Gilgel Abay, Gumara, Rib and Megech Rivers. About 93% of the annual inflow to Lake Tana is believed to come from these rivers (Kebede et al., 2006), and better understanding of the hydrology of these rivers plays a crucial role for an efficient management of the lake and its basin. Two of the sub-

115 catchments (Gumara and Gilgel Abay) were selected for this study, in order to represent the hilly and mountainous lands of the southern and eastern parts of the sub-basin as the bulk of it is located here (figure 1), and to optimally use the available data. For both sub-catchments, large parts of their territory are intensively cultivated. The lower floodplains in these catchments with their buffering capacity are not considered by this study, but were discussed by Dessie et al. (2014).

120 The Gilgel Abay catchment (Fig. 1) covers an area of 1659 km² at the gauging station near Picolo, with elevations ranging between 1800 and 3524 m a.s.l. Soils are characterized by clay, clay loam and silt loam textures, each texture sharing similar proportions of the catchment area (Bitew and Gebremichael, 2011). The majority of the catchment is a basalt plateau with gentle slopes, while the southern part has a rugged
125 topography.

The Gumara catchment covers part of the eastern side of the Lake Tana basin. At its upper and middle portion, it has mountainous, highly rugged and dissected topography with steep slopes. The lower part is a valley floor with flat to gentle slopes. Elevation in the catchment varies from 1780 to 3700 m a.s.l. At the upper gauging station (Fig. 1), the
130 catchment area is 1236 km². Two independent studies found very homogeneous textures of the soils in this catchment. BCEOM (1998a) described it as dominantly clay with sandy clay soil at some places in the catchment, while soil data collected by Miserez (2013) show that texture is clay and clay loam. In the hilly catchments, clay soils are essentially Nitisols which do not present cracking properties as opposed to lowland
135 Vertisols (Miserez, 2013).

Based on rainfall data from the Dangila and Bahir Dar stations, observed in the period 2000 to 2011, mean annual rainfall is ca. 1500 mm, with more than 80% of the annual rainfall concentrated from June to September. Geologically, the catchments consist of Tertiary and Quaternary igneous rocks, as well as Quaternary sediments. The rivers in the
140 hilly areas are generally bedrock rivers, whereas in the floodplain the rivers meander and sometimes braid (Poppe et al., 2013).

Fig. 1 approximately here

145

3 Model development

The model developed is based on a simple water balance approach and the studies by Jothityangkoon et al. (2001), Krasnostein and Oldham (2004) and Fenicia et al. (2008).

150 The setup of this model is shown in Fig. 2. In this modeling approach, the catchment is first split into soil surface and impermeable surface (these are areas with little or no soil cover and bedrock outcropping in the catchment as well as soils with well-developed tillage pans). The runoff from the presumed impermeable areas is modeled as infiltration excess (Hortonian flow) runoff and is represented as Q_{se2} . The other component of the
155 catchment, recognized as the soil surface, is further divided into three using topographic criteria (slope), considering topography as a proxy for the variability of most of the catchment characteristics. Here, two reservoirs are introduced (the soil reservoir and the groundwater reservoir). The slow reacting reservoir (or the groundwater reservoir) is set to be common to all of the three slope based divisions of the catchment as it looks quit
160 inconsistent to separate the groundwater system in the catchment. The catchment buckets (reservoirs) and the conceptual runoff processes are depicted in Fig.2 (b) and (c). However, the model is modified to reflect the actual catchment conditions of the study areas, such that an additional component that accounts for surface runoff production from impermeable surfaces (with little or no soil cover) in the catchments is included.
165 Topography is considered as a proxy for the variability of most of the catchment characteristics. We divided the catchments into different runoff production areas using topographic criteria. Moreover, the percolation to the groundwater table and the hydraulic conductivity for the interflow are modelled differently and the formation of saturated areas at the bottom of slopes as a result of interflow from steep hills in the catchments is
170 considered.

Jothityangkoon et al. (2001) conceptualized the upper soil layer (further referred to as the soil reservoir) as a ‘leaky bucket’. By adding a groundwater reservoir (Krasnostein and Oldham, 2004), the conceptual model for modelling the runoff discharge at the catchment outlet was developed. ~~The setup of this model is shown in Fig. 2.~~

175 ***Fig. 2 approximately here***

In figure 2, Q_I [mm/day] is the sum of direct runoff and interflow in the soil reservoir, Q_2 [mm/day] is the baseflow from the groundwater reservoir, Q_{Se2} is the direct runoff from impermeable surface of the catchment and the sum of Q_I ~~and~~ Q_2 ~~and~~ Q_{Se2} forms the total river discharge, Q [mm/day], at the outlet of a catchment.

The water storage at any time t within the soil reservoir, $S(t)$ in mm, is determined by the precipitation (P , in mm/day), evapotranspiration (E_a , in mm/day), and other catchment controlled outputs (cfr. Fig. 2-2c(i-iii) and 3a). When the storage depth exceeds the field storage capacity (S_f , in mm), precipitation is assumed to be partly transformed into subsurface runoff, to represent inter- or subsurface flow (Q_{ss} , in mm/day), and partly into deep percolation or recharge (R , in mm/day) to the groundwater (Fig. 2cii3b). When the soil reservoir fills completely, and the inflows exceed the outflows, surface runoff (Q_{se1} , in mm/day) is generated.

Fig. 3 approximately here

Quantitatively, the depth of water stored in the soil, $S(t)$, evolves over time using the water balance:

$$S(t) = S(t - \Delta t) + (P - E_a - Q_{ss} - Q_{se1} - R)\Delta t \quad (1)$$

where P is the precipitation [mm/day], E_a is the actual evapotranspiration [mm/day], $S(t - \Delta t)$ is the previous time step storage [mm], Q_{ss} is the interflow or subsurface runoff [mm/day], Q_{se1} is the direct or overland flow from the soil reservoir [mm/day], R is deep percolation or recharge to the substrata and groundwater [mm/day], and Δt is the time step equal to one day.

Different studies show that some part of the interflow water from the steep hills appears at the hill bottoms during wet periods in the form of increased moisture content or overland flow (Frankenberger et al., 1999; Bayabil et al., 2010; Mehta et al., 2002; Tilahun et al., 2013). These findings reveal that the hill bottoms receive additional inputs to the soil reservoir from the steep upper parts of the hills besides the rainfall. In this modelling approach, it is assumed that steep hills first recharge the medium slope sections, and consequently the medium slope surfaces recharge the flat regions (valley bottoms). The magnitude of the recharge (Q_r , in mm/d) is modelled as:

$$Q_r = \alpha Q_{ss} \quad (2)$$

where α (-) is interflow partitioning parameter and Q_{ss} is as defined above. Equation (1)

210 is, therefore, modified for the medium slope and flat surfaces as

$$S(t) = S(t - \Delta t) + (P + Q_r - E_a - Q_{ss} - Q_{se1} - R)\Delta t \quad (3)$$

3.1 Actual evapotranspiration

215 During wet periods, when the depth of available water exceeds the maximum available soil storage capacity (S_b , in mm), the actual evapotranspiration is equal to the potential evapotranspiration (E_p , in mm/day). When $S(t)$ is lower than S_b , E_a is assumed to decrease linearly with moisture content as follows (Steenhuis and van der Molen, 1986):

$$E_a = E_p \left(\frac{S(t)}{S_b} \right) \quad (4)$$

$$220 \quad S_b = D\phi \quad (5)$$

where D is the soil depth [mm] and ϕ is the soil porosity (-).

3.2 Subsurface runoff

225 Subsurface runoff, Q_{ss} [mm/day], occurs only when the storage depth exceeds the field storage capacity (S_f , in mm). It is calculated as the difference between the storage and the field storage capacity, divided by the response time (T_r) of the catchment with respect to subsurface flow (Jothityangkoon et al., 2001):

$$Q_{ss} = \frac{S(t) - S_f}{T_r}, \text{ when } S(t) > S_f \quad (6)$$

$$Q_{ss} = 0, \text{ when } S(t) \leq S_f$$

230 The field storage capacity of the soil reservoir, S_f [mm], is calculated by

$$S_f = F_c D \quad (7)$$

where F_c (-) is the field capacity of the soil (dimensionless).

The catchment response time is the time taken by the excess water in the soil to be released from the soil and drained out from the catchment. This response time depends on the properties of the soil and the topography of the system, and the subsurface flow velocity (V_b , in mm/day) can be expressed as

$$V_b = \frac{L}{T_r} \quad (8)$$

where L is the average slope length of the catchment [mm]. From Darcy's law in saturated soils, V_b is also given as

$$V_b = K_s i \quad (9)$$

Where K_s is the saturated hydraulic conductivity of the soil [mm/day] and i is the hydraulic gradient, which is approximated by the average slope gradient (G) of the catchment.

Brookes et al. (2004) analyzed the variability of saturated hydraulic conductivity with depth, and they found large K_s values near the surface or root zone layer and the transmissivity that decreases exponentially with depth. Accordingly, a variation is made between the upper soil layer (that affects interflow) and deep soil layer (percolation to groundwater) hydraulic conductivities. The permeability (K , in mm/day) of the upper soil layer for the interflow under different soil water conditions is modelled as:

$$K = K_{s,u} \left(1 - e^{-\beta \frac{S(t)}{S_b}}\right) \quad (10)$$

where β is a dimensionless parameter, and $K_{s,u}$ [mm/day] is the saturated hydraulic conductivity of the upper soil layer, both of which are to be calibrated.

The response time (T_r) in Equation (6) is, hence, approximated from Equations (8), (9) and (10) as

$$T_r = \frac{L}{GK} \quad (11)$$

Where L and K are as defined in Equations (8) and (10) and G is average slope gradient of the catchment.

The deep percolation or recharge to groundwater (R , in mm/day) under varying soil water content conditions is modelled as:

$$260 \quad R = K_{s,e} \left(1 - e^{-\gamma \frac{S(t)}{S_b}} \right) \quad (12)$$

Where γ a dimensionless parameter, and $K_{s,e}$ [mm/day] is the saturated hydraulic conductivity of the deep soil layer, which is to be estimated from the aquifer properties of the catchments. This equation is identical to Equation (10) such that in both cases it is assumed that conductivities vary exponentially under varying soil water content conditions but with different magnitudes.

3.3 Saturated excess runoff

Saturated excess runoff or surface runoff (Q_{se1} , in mm/day) is calculated as the depth of water that exceeds the total water storage in the soil reservoir at each time step (Jothityangkoon et al., 2001; Krasnostein and Oldham, 2004).

$$Q_{se1} = \frac{S(t) - S_b}{\Delta t}, \text{ when } S(t) > S_b$$

$$Q_{se1} = 0, \text{ when } S(t) \leq S_b \quad (13)$$

3.4 Surface runoff from the impermeable areas

Field visits on the Upper Blue Nile basin (including the study catchments) revealed the existence of exposed surfaces that cause strong runoff response. These are areas with little or no soil cover and bedrock outcropping in some parts of the catchment as well as soils with well-developed tillage pans (Melesse Temesgen et al., 2012a, 2012b) (Fig. 3). Hence, runoff from these almost impermeable areas is modeled as infiltration excess (Hortonian flow) runoff with a very small amount of retention before runoff occurs (Steenhuis et al., 2009). The surface runoff from these areas (Q_{se2} , in mm/day) is calculated as

$$Q_{se2} = P - E_p, \text{ when } P > E_p$$

$$Q_{se2} = 0, \text{ when } P \leq E_p \quad (14)$$

285 Where P and E_p [mm/day] are as defined above. The impermeable portion of the catchment area (A_r , in km^2) is modelled from the total catchment area (A_t , in km^2) as

$$A_r = \lambda A_t \quad (15)$$

where λ is the fraction of impermeable surface within the catchment.

290 ***Fig. 3 approximately here***

3.5 Groundwater reservoir and baseflow

The introduction of a deep groundwater storage (Fig. 2b) helps to improve low flow predictions. This baseflow reservoir is assumed to act as a non-linear reservoir (Wittenberg, 1999) and its outflow, Q_2 [mm/day], and storage, S_g [mm], are related as

$$Q_2 = \frac{S_g^{k_1}}{\Delta t} \quad (16)$$

where k_1 is a dimensionless model parameter. The water balance of the slow reacting reservoir (groundwater reservoir) is given by

$$300 \quad S_{g(t)} = S_{g(t-\Delta t)} + (R - Q_2)\Delta t \quad (17)$$

where $S_{g(t)}$ [mm] is the groundwater storage at the given time step, $S_{g(t-\Delta t)}$ [mm] is the previous time step groundwater storage, R [mm/day] is the deep percolation, as given by Equation (12).

In total the model has seven parameters:

305 (i) Parameters for the recharge (α_1 and α_2): In the three slope classification, α_1 is to consider for the recharge from the steeply slope into the medium slope surface and α_2 is for the recharge from the medium slope surface into the flat surface. There is no parameter for the steeply slope surface since there is no surface that recharges it. So, there are two parameters for the three slope classifications.

310 (ii) Parameter for the impermeable surface of the catchment (λ)

The catchment is divided into two surfaces (impermeable surface with no or little soil cover and the soil surface). The parameter λ is introduced to represent the fraction of impermeable surface within the total catchment and this part of the catchment is not

classified as steeply, medium slopes and flat surfaces since the classification of this part of the catchment into such classes is not important. So we have one parameter.

(iii) The parameters β , γ , k_I and $K_{s,u}$

These parameters β and γ are introduced to account variability of permeability and deep percolation of soil with soil water storage. k_I relates discharge and storage for the ground water and $K_{s,u}$ is the saturated hydraulic conductivity in the upper soil layer. We assumed that these parameters are less influenced by topography and each model parameter is assumed to be same for each slope classification of the catchment. Moreover, it is quit inconsistent to separate the groundwater system in the catchment. Therefore, all the three slope based classified sub-catchments share the same groundwater reservoir.

In this modeling approach, stream-groundwater interactions are assumed to be minimal and the groundwater is assumed to recharge the streams from deep percolation of rainfall on the catchments that produces baseflow of the rivers/streams. The storage effect of the streams when considered on the basis of average daily flows of the streams is assumed to be negligible and hence streamflow routing was not considered for such smaller streams.

3.6 Total river discharge

The total river discharge (Q_t , in mm/day) at the outlet of the catchments is given by:

$$Q_t = Q_{ss} + Q_{se1} + Q_{se2} + Q_2 \quad (18)$$

4 Data inputs

The data needed for the model are classified into three types: topographical, soil, and hydrological data.

4.1 Topographical data

Steenhuis et al. (2009) found that overland flow in the Blue Nile basin is generated from saturated areas in the relatively flatter areas and from bedrock areas, while in the rest of

345 the catchment all the rainfall infiltrates and is lost subsequently as evaporation, interflow or baseflow. Topographical processes have been found to be the dominant factors in affecting runoff in the Blue Nile Basin (Bayabil et al., 2010). We used topography of catchments as the main criterion to divide the catchment into different runoff production surfaces. Based on slope criteria (FAO, 2006), each study catchment was divided into
350 three sub-catchments as steep (slope gradient > 30%), hilly or medium (slope gradient between 8 and 30%) and flat (slope gradient < 8%) to consider spatial variability in catchment properties and runoff generation mechanisms [\(Fig.4\)](#).

[***Fig.4 approximately here***](#)

The 30 m by 30 m resolution Global Digital Elevation Model (GDEM) was used to
355 define the topography (downloaded from the ASTER website, <http://earthexplorer.usgs.gov/>). The GDEM (figure 1) was used to delineate and calculate the average slope gradient and average slope length of the catchments (topography-related inputs to the model).

360 4.2 Soil data

The model requires data on depth, porosity and field capacity of the soils. Soil depth and soil [types data \(Fig.5 and Fig.6\)](#) were obtained from the Abay River Basin integrated master plan study BCEOM (1998a).

365 [***Fig.5 approximately here***](#)

[***Fig.6 approximately here***](#)

In this modeling philosophy, the soil depth is meant to represent the depth of water stored in the topmost layer (root zone) of the soil (Fig. ~~2~~~~and~~~~3~~). The porosity and field capacity of the soils were derived from the soil texture based on the work of McWorter and Sunada (1977). From this, we determined the soil textures of the study catchments
370 (Table 1). The saturated hydraulic conductivity for the deep percolation (Equation 12) was estimated using ranges of conductivities given by Domenico and Schwartz (1990) for the saturated hydraulic conductivities of a deep soil layer (colluvial mantle on top of the igneous rock). A summary of the topographic, soil and saturated hydraulic conductivity
375 data for the study catchments is provided in Table 1.

Table 1 approximately here

4.3 Weather data

Daily precipitation is the key input meteorological data for the model. Other meteorological data like minimum and maximum air temperature, humidity, wind speed and duration of sunshine hours were also used to calculate the potential evapotranspiration, the other input variable to the model. All weather data were obtained from the Ethiopian National Meteorological Agency (NMA) for 13 stations located within and around the catchments (www.ethiomet.gov.et). [The location map of the rain gauge stations used for this study are depicted in Fig.7. The data for most of the stations are consistent and continuous, particularly for the first class stations like Dangila, Adet and Debretabor. However, we encountered gaps in some stations like Sekela Station for some periods in the year. In such instances, only the rainfall data from the other stations were considered. Most of the rainfall stations in Gilgel Abay catchment are installed at the water divides and there is no station in the middle of the catchment. In this regard, the Gumara catchment has a higher density of rainfall stations.](#) The areal rainfall distribution over the catchments was calculated using the Thiessen Polygon method, and the potential evapotranspiration was calculated using the FAO Penman-Monteith method (Allen et al., 1998).

[Fig.7 approximately here](#)

4.4 River discharge

Starting from July 2011 water level was measured at the Wanzaye station (11.788073°N, 37.678266°E) on Gumara River and from December 2011 at the Picolo station (11.367088°N, 37.037497°E) on Gilgel Abay River. The water level measurements were made using Mini-Divers, automatic water level recorders (every 10 minutes), and manual readings from a staff gauge (three times a day, at 7 AM, 1 PM and 6 PM), following the procedures described by Amanuel et al. (2013).

Discharges were computed from the water levels using rating curves (Equations 19 and 20) for each station. The rating curves ([Fig.8](#)) were calibrated based on detailed surveys

of the cross-sections of the rivers and measurements of flow velocity at different flow stages, using the following commonly used expression:

$$Q = ah^b \quad (19)$$

410 where a and b are fitting parameters and Q [m^3/s] and h [m] are discharge and water level respectively. The resulting rating curve equation for the Gumara catchment at the gauging station (Wanzaye Station) is:

$$Q = 44.1h^{1.965} \quad (R^2=0.997, n=12) \quad (20)$$

and for Gilgel Abay catchment at Pico Station:

415 $Q = 70.39 h^{2.105} \quad (R^2=0.985, n=14) \quad (21)$

Fig.8 approximately here

Compared to the discharge data that have been gathered in the past, the discharge data that are acquired for this study are of superior quality, since a high time resolution during the measurement has been used. This minimizes the risk of missed peaks, particularly during the night. Furthermore, frequent supervision was also made during the data collection campaign. Hence, these data were used for the model calibration. Discharge data collected before December 2011 were obtained for nearby stations from the Hydrology Department of the Ministry of Water Resources of Ethiopia, which has a long data record (since 1960) for these stations. However, the latter measurements were made using staff gauge readings twice a day, with many data gaps and discontinuities, particularly at the end of the observation window. The discharge data from 2000-2005 are relatively better and are used to validate the model.

420
425
430 The 2012 discharge data for Dirma catchment (outlet at 12.427194°N, 37.326209°E), collected in the same way as those of Gilgel Abay and Gumara, were used to assess the transferability of the model parameters.

5 Calibration and validation

435 The model calibration and validation were performed at a daily time step, and the hydrological datasets of 2012 and 2011-2012 were used to calibrate the Gilgel Abay and Gumara catchments, respectively. Discharge data of 2000-2005 were used for validation.

There are 7 calibration parameters in this model (Table 2), and the calibration was performed using the Particle Swarm Optimization (PSO) algorithm. PSO is a population based stochastic optimization technique inspired by social behavior of bird flocking or fish schooling (Kennedy and Eberhart, 1995). The advantages of PSO are that the algorithm is easy to implement and that it is less susceptible to getting trapped in local minima (Scheerlinck et al., 2009). We carried out 50 iterations and 50 repetitions, in total 2500 runs for each catchment to search for the optimal value of the model parameters (Table 2) and 30 particles were used in the PSO. The criterion in the search for the optimal value was to minimize the root mean squared error (*RMSE*) as the objective function, given by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{n}} \quad (22)$$

where Q_{obs} is observed discharge [mm/day], Q_{sim} is simulated or modelled discharge [mm/day], and n is the number of data points. The parameter values corresponding to the minimum *RMSE* were considered as optimum. From the optimal model parameters, the performance of the model was also evaluated using (i) the Nash-Sutcliffe Efficiency (*NSE*) according to Nash and Sutcliffe (1970), and (ii) the coefficient of determination (R^2).

$$NSE = 1 - \frac{\sum_{i=1}^n (Q_{sim,i} - Q_{obs,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \bar{Q}_{obs})^2} \quad (23)$$

$$R^2 = \left[\frac{\sum_{i=1}^n (Q_{sim,i} - \bar{Q}_{sim})(Q_{obs,i} - \bar{Q}_{obs})}{\sqrt{\sum_{i=1}^n (Q_{sim,i} - \bar{Q}_{sim})^2} \sqrt{\sum_{i=1}^n (Q_{obs,i} - \bar{Q}_{obs})^2}} \right]^2 \quad (24)$$

where \bar{Q}_{obs} [mm/day] and \bar{Q}_{sim} [mm/day] are the mean observed and simulated discharges, respectively.

[Percent bias \(PBIAS\) is used as an additional model performance indicator. It measures the average tendency of the simulated data to be larger or smaller than the observations](#)

460 [\(Gupta et al., 1999\)](#). The optimal value of PBIAS is 0.0, with lower values indicating better model simulation (positive values indicate overestimation, whereas negative values indicate model underestimation bias).

$$PBIAS = \frac{\sum_{i=1}^n (Q_{sim,i} - Q_{obs,i})}{\sum_{i=1}^n Q_{obs,i}} * 100\% \text{-----(25)}$$

Table 2 approximately here

465 The impacts of model parameters on the output of the model when their values are different from the calibrated optimal values were evaluated with respect to the Root Mean Squared Error for Gumara catchment. The sensitivity analysis was made by randomly selecting parameter values in the region of the optimal values obtained from PSO and calculating *NSE* for each selected value. The applicability of the model to other
 470 ungauged catchments outside the study catchments in the Lake Tana basin was also tested using direct parameter transferability.

6. Soil and Water Assessment Tool (SWAT) and (FlexB) models as benchmarks for comparison with this paper model

475 [The two models are used as benchmark models to assess the performance of the model of this paper \(hereafter named as Wase -Tana model, in favor of the project name that funded this study\), which tries to use all available information and considers topography as a good proxy for the variability of most of the catchment characteristics in the Upper Nile basin.](#)

6.1 SWAT Model

485 [SWAT is a basin-scale and continuous-time model, used to simulate the quality and quantity of surface and ground water and predict the environmental impact of land use, land management practices, and climate change \(Arnold et al., 1998\). The hydrological model is based on the water balance equation](#)

$$SW_t = SW_0 + \sum_{i=1}^t (R_i - Q_i - ET_i - P_i - QR_i) \text{-----(26)}$$

490 Where: SW_t is the soil water content at time t , SW_0 is the initial soil water content, t is the
time step in days and R_i , Q_i , ET_i , P_i and QR_i respectively are the daily amounts of
precipitation, runoff, evapotranspiration, percolation and return flow. All units are in mm.
In SWAT, a watershed is divided into homogenous hydrologic response units (HRUs)
based on elevation, soil, management and land use, whereby a distributed parameter such
as hydraulic conductivity is potentially defined for each HRU. Hence, an analyst
495 confronts with the difficult task of collecting or estimating a large number of input
parameters, which are usually not available for regions like the Upper Blue Nile basin.
Details of the model can be accessed at the SWAT website (<http://swatmodel.tamu.edu>).
Automatic calibration and validation of the model was made using SWAT-CUP. It is an
interface that has been developed for SWAT automatic calibration and model uncertainty
500 analysis (Abbaspour et al., 2007). Coefficient of determination (R^2) and Nash-Sutcliffe
Efficiency (NSE) were used as objective functions during the calibration process of the
search for the optimal value.

505 6.2 FlexB Model

This model is a lumped conceptual type and it is characterized by three reservoirs as
described by Fenicia et al. (2008): the unsaturated soil reservoir (UR), the fast reacting
reservoir (FR) and the slow reacting reservoir (SR). The model has eight parameters: a
shape parameter for runoff generation β [-], the maximum UR storage S_{fc} [mm], the
510 runoff partitioning coefficient D [-], the maximum percolation rate P_{max} [mm/h], the
threshold for potential evaporation L_p [-], the lag times of the transfer functions N_{lag} [h],
and the timescales of FR and SR: K_f [h] and K_s [h]. Details of the model and the various
equations of the model can be referred to Fenicia et al. (2008).

Calibration of this model was made using the particle Swarm Optimization (PSO)
515 technique, following similar procedures of the Wase –Tana model calibration algorithm.
The same objective function, root mean squared error (RMSE), is also used in the search
for the optimal value.

520 **76 Results and discussion**

76.1 The daily hydrograph and model performance indicators

a) Wase –Tana model performance

525 Figures 95 and 106 show a comparison of the modeled with the observed discharge data for the two study catchments and for both the calibration and validation periods.

Fig. 95 approximately here

530 Despite the possible spatial variability of some input data (average soil and rainfall data are considered) and the simplicity of the model, discharge is reasonably well simulated during both the calibration and validation periods. This can be seen from the visual inspection of the hydrographs and from the model performance indicators (Table 3).

535 ***Fig. 106 approximately here***

Table 3 approximately here

540 The Nash-Sutcliffe efficiency of the model is high for both catchments. In the calibration period, *NSE* equals 0.86 for Gumara catchment and 0.84 for Gilgel Abay catchment, while they are 0.78 and 0.7, respectively, during the validation period. Figures 95 and 106 also show that the model simulates well the overall behavior of the observed streamflow hydrographs. However, an overestimation of the large flood peaks for the Gilgel Abay catchment is found for the validation period. In the calibration period for this catchment, the model errors tend to increase during wetting up periods almost for all the models. Initially, the soils are relatively dry and most of the rainfall during the beginning of the rainy season is not effective to produce runoff in the model as the soil reservoir has to be filled first to generate the faster component of the runoff. Besides model uncertainties, the rainfall data quality can also affect the model performance, mainly in the case of the Gilgel Abay catchment. The R^2 values for the time series of daily streamflow between simulated and observed values were 0.80 to 0.86 for the Gumara catchment, and from 0.79 to 0.85 for the Gilgel Abay catchment, for the validation and

545

550

555 calibration periods, respectively. Generally, the modelled discharges appear to be less variable over time than the observations, as shown by the standard deviations in Table 3. This is likely due to the fact that data used in the model are averaged over the year, while observed river discharges are highly seasonal. We used average daily rainfall data, average soil data (e.g. porosity, field capacity, and soil depth), average catchment characteristics data (e.g. slope, slope length) to mention some for the model inputs. Hence, this averaged condition may be one source of error such that the model may not exactly mimic extremes like peak discharges.

560 b) Performance in comparison with the benchmark models

565 For the calibration period, almost all the three models performed pretty well (Table3). However, an appreciable decrease in model performance has been noticed for the validation period in Gilgel Abay catchment for the benchmark models. SWAT is a physically-based complex model, requiring extensive input data which is a challenge for data scarce regions like the Upper Blue Nile basin. The model simulations can only be as accurate as the input data. This suggests that the coarser data input used for the model in the study catchments might have affected significantly the calibration and consequently the validation simulations. On the other hand, the likely reason for a decreased performance of the Flex_B model for the Gilgel Abay catchment is the oversimplification of the catchment heterogeneity, since it is a lumped one and the impact is more when the catchment gets bigger (Gilgel Abay catchment is bigger than Gumara catchment).

570 A look at the flow duration curves (Fig.11 and Fig.12) indicates the higher uncertainty of the two benchmark models (mainly SWAT model) with respect to low flow predictions.

575 ***Fig.11 approximately here***

Fig.12 approximately here

580 In relative terms, Wase-Tana model offers more flexibility in adapting the model to the catchments based on the validation simulation performances. This can be attributed to the consideration of topography driven landscape heterogeneity analysis and catchment information extraction for the model, which strengthens the hypothesis that topography driven model structure and use of all available information on hydrology based on

585 [topography is a good choice for the Upper Blue Nile basin. From a comparison of four model structures on the Upper Heihe in China, Gao et al. \(2014\) also confirmed that topography-driven model reflects the catchment heterogeneity in a more realistic way.](#)

76.2 The hydrograph components and hydrological response of the catchments

590 This hydrological model ([Wase-Tana model](#)) is based on the generation of direct runoff from saturated and impermeable (degraded surfaces and rock outcrops with little or no soil cover) areas, interflow from the soil storage in the root zone layer and baseflow from the deeper layer as groundwater storage. The understanding of the relative importance of these processes on the hydrological response of each catchment is still unknown. The mean annual surface runoff (Q_{se} , sum of Q_{se1} and Q_{se2}), interflow or subsurface flow 595 (Q_{ss}) and baseflow (Q_2) components of the total daily hydrograph computed by the model for the calibration and validation periods are given in Table 4.

Table 4 approximately here

600 The total mean annual runoff generated by the model is in line with the observations for both catchments in the calibration period (Table 4), while an appreciable difference is noticed in the values for the Gilgel Abay catchment in the validation period. One of the problems in accurate modelling of the discharge is that precipitation measurements do not cover well the catchments. This is particularly the case for the Gilgel Abay catchment, where the rainfall stations are poorly distributed as most of the meteorological stations lie 605 near the water divides. The calibration results are better, since the data from the recently established precipitation stations (e.g. Durbetie) could be used. There are also doubts on the representativeness of the discharge data used for the validation of the model, because the water level measurements were made manually and twice daily (in the morning and late afternoon), leading to the possibility of missing flash floods at other moments of the 610 day as the stream discharge is very variable. This can be clearly seen from the mean annual observed flows during the calibration and validation periods for Gilgel Abay. The mean annual observed flow in the validation period was found to be much smaller than the corresponding flow during the calibration period (Table 4). The closer total mean

615 annual runoff values and the better model performance indicators for the Gumara catchment during the calibration period suggest that the model can perform satisfactorily with better input discharge and precipitation data.

From PBIAS results (Table 3), Flex_B model has showed overestimated bias and SWAT model behaved the opposite for both catchments during the calibration period.

620 Despite the variations in mean annual runoff generated by the Wase-Tana model, the partitioning of the total runoff into the different components (Table 4) in each period is almost identical for each catchment, as expected. About 65% of the runoff appears in the form of interflow for the Gumara catchment, and baseflow takes the larger proportion for Gilgel Abay catchment (44 - 48%). Uhlenbrook et al. (2010) obtained the baseflow to be about 32% from similar model study results for Gilgel Abay catchment. Vogel and Kroll
625 (1992) have showed that baseflow is a function of catchment area, and geomorphological, geological and hydrogeological parameters of the catchment have a linear incidence on the discharges. The difference between the baseflow of the two catchments is high, despite their comparable catchment sizes, suggesting rather the different structure, functioning and hydrodynamic properties of the two catchments. Hence, the model
630 results reveal that the groundwater in the Gilgel Abay catchment receives more recharge and makes a greater contribution to the river flow. This is in line with Kebede (2013) and Poppe et al. (2013), who show that the largest part of the Gilgel Abay catchment consists of pumice stones and fractured quaternary basalts with a high infiltration capacity and hydraulic properties, which clarifies the large groundwater potential. In line with this, several big springs exist in the catchment, including one that is used as a source of water supply for Bahir Dar town (Fig.13).

635 ***Fig.13 approximately here***

640 The other interesting result is that direct runoff is the smallest fraction of the total runoff for both catchments (18-19% for Gumara and 20% for Gilgel Abay) and almost all peak flow incidences are associated with direct runoff. More than 90% of this direct runoff is found to be from the relatively impermeable (degraded areas, plough pans or rock outcrops with little or no soil cover) surfaces. The calibrated result shows that this type of runoff production area covers 15% of the Gumara and 17% of the Gilgel Abay catchments, respectively. In a similar study, Steenhuis et al. (2009) mention that the rock

645 outcrops occupy 20% of the total catchment area in the Abay (Blue Nile) catchment at the Ethiopia–Sudan border upstream of the Rosaries Dam, which is very similar to the result of Gilgel Abay catchment in this study.

The remaining direct runoff is generated from the flat slopes of the catchments as saturated excess runoff, probably near the valley bottoms. The hillslopes (medium and steep slope source areas in this paper) generated almost no direct runoff as saturated excess flow. Similar results were obtained by different researchers in the Blue Nile Basin, who identified hillslopes as main recharge areas (Steenhuis et al., 2009, Collick et al., 2009, Tilahun et al., 2013). Our results contribute to the debate on the relative importance of saturated excess runoff versus infiltration excess runoff (Hortonian overland flow) mechanisms in the Upper Blue Nile Basin, showing that the rainfall-runoff processes are better represented by the soil reservoir methodology. Yet, further research is necessary that involves rainfall intensity and event-based analysis of hydrographs.

660 **7.6.3 Transferability of model parameters to other ungauged catchments and sensitivity**

The sensitivity analysis was performed on model parameters for Gumara catchment with respect to the Root Mean Squared Error.

Fig. 14-7 approximately here

665 ***Fig. 15 approximately here***

The parameters β , $\alpha 1$ and γ show poor sensitivity for a wide range of values with respect to the local sensitivity analysis. The local sensitivity analysis shows the sensitivity of a variable to the changes in a parameter if all other parameters are kept constant at some value (optimal value in this case). An increase in the value of β beyond 1.4 showed almost no sensitivity, while the model efficiency decreased slightly after an increase in the value of γ from the optimum. This means that there is little confidence in the model's correspondence with these parameters and they can be reduced without appreciable impact on the model (Fenicia et al., 2008). $k1$, $K_{s,u}$ and λ are very sensitive parameters in this model and the model performance drops abruptly if the parameters exceed beyond some threshold value (Fig.14-7).

680 The global sensitivity analysis (Fig.15), however, shows interactions among all the input parameters of the model. Although global sensitivity analysis reveals details of the model behavior in a more general sense through random parameter sampling and that the parameters are all sensitive, the local sensitivity analysis indicates that moderate variations of the parameter values for some parameters can still drastically change the model performance.

685 The model parameter transferability to other ungauged catchments in the basin has been tested by analyzing the variability among the calibrated parameters of the two catchments. Table 2 shows that the calibrated parameters are nearly identical for both catchments, except for γ and λ , which are related to deep percolation and impermeable fraction of the catchment, respectively. As described above, they affect the baseflow and direct runoff contributions to the total river flow. However, we showed that the contributions of these components to the total runoff are relatively small and γ is poorly sensitive to a wide range of values. Thus the influence of these parameters is expected to be minimal. This is verified by generating flows using the average of the calibrated parameters of the two catchments and analyzing the effect on the model performance indicators (Table 5). The model performance obtained using the average model parameter values is similar to the results found using the optimal model parameters (Table 3). To further verify the adaptability of the average calibrated model parameter values outside 690 the study catchments and see the impacts of scale, we applied the average parameter values to another catchment (Dirma catchment in the northern part of the Lake Tana sub-basin, Fig.1) with an area of 162.6 km². Encouraging model efficiency could be obtained, with *NSE* and *R*² values of 0.58 and 0.6 respectively (Table 5). This is to be elaborated further in the future, involving more catchments and more years of data.

700 ***Table 5 approximately here***

In general, transferability results showed good performance of the daily runoff model in the two study catchments and an average performance in the test catchment (Dirma catchment). This can be explained by the fact that emphasis was made to incorporate more knowledge in the model structure to increase model realism. We based strongly on the soil storage characterization of the soil reservoir in the rainfall-runoff process and

710 [representation of the maximum storage of the unsaturated reservoir at the catchment](#)
[scale, which is closely linked to rooting depth and soil structure and strongly depends on](#)
[the ecosystem. Transferability of the model has benefited from this in that we were able](#)
[to derive most of the input data from the test catchments. The consideration of](#)
[topography driven landscape heterogeneity analysis and catchment information extraction](#)
[based on topography \(slope\) for the model is another reason for the better performance of](#)
[the model transferability. The role of topography in controlling hydrological processes](#)
[and its linkage to geology, soil characteristics, land cover, and climate through](#)
715 [coevolution have been indicated in different studies \(Sivapalan, 2009, Savenije, 2010,](#)
[Gao, 2014\).](#) The results suggest the possibility of directly using the average model
parameter values for other ungauged catchments in the basin, even though further tests on
such catchments is still recommended. However, we believe that this is a useful result for
operational management of water resources in this data scarce region.

720

87 Conclusion

In this paper, a simple conceptual semi-distributed hydrological model was developed
and applied to the Gumara and Gilgel Abay catchments in the Upper Blue Nile basin,
725 Lake Tana sub-basin, to study the runoff processes in the basin. Good quality discharge
data were collected through a field campaign using automatic water level recorders with
high time resolution. We used the topography and soil texture data of the catchments as
the dominant catchment characteristics in the rainfall-runoff process. In the model, a
distinction is made between impermeable surfaces (degraded surface or exposed rock
730 with little or no soil cover) and permeable (soil) surfaces, as different types of source
areas for runoff production. The permeable surfaces were further divided into three
subgroups using topographic criteria such as flat, medium, and steep slope areas. The
rainfall-runoff processes were represented by two reservoirs (soil and groundwater
reservoirs) and the water balance approach was used to conceptualize the different
735 hydrological processes in each of the two reservoirs. Such a detailed form of modelling,
using topography as a dominant landscape characteristics to classify a catchment into

different hydrological regimes, has not been applied yet in the Upper Blue Nile, Lake Tana sub-basin.

740 We demonstrated that the model performs well in simulating river discharges, irrespective of the many uncertainties. Model validation indicated that the Nash–Sutcliffe values for daily discharge were 0.78 and 0.7 for the Gumara and Gilgel Abay catchments, respectively.

745 We were able to partition the total runoff into a fast component (direct runoff and interflow) and a slow component (baseflow) and estimated the contributions of each component for the catchments. About 65% of the runoff appears in the form of interflow for the Gumara catchment, and baseflow is responsible for the larger proportion of the discharge for the Gilgel Abay catchment (44-48%). Direct runoff generates the lower fraction of runoff components in both catchments (18-19% for the Gumara and, 20% for the Gilgel Abay) and almost all peak flow incidences are associated with direct runoff.
750 More than 90% of this direct runoff is found to be from the relatively impermeable (plough pan or rock outcrops with little or no soil cover) source areas. The hillslopes (medium and steep slope source areas) are recharge areas (sources of interflow and deep percolation) and generated almost no direct runoff as saturated excess flow.

755 The results of this study, [with comparisons to two benchmark models](#), clearly demonstrate that topography is a key landscape component to consider when analyzing runoff processes in the Upper Blue Nile basin. Generally, runoff in the basin is generated both as infiltration and saturation excess runoff mechanisms. A considerable portion of the landscape in the Upper Blue Nile basin consists of impermeable rock outcrops and hard soil surfaces (15%-17% of the total catchment area as per the results of this study)
760 and they are the sources of most of the direct runoff. This conceptual model, developed to study the runoff processes in the Upper Blue Nile basin, may help to predict river discharge for ungauged catchments for a better operation and management of water resources in the basin, owing to its simplicity and parsimonious nature with respect to parameterization. The runoff processes in the basin are also found to be affected much by
765 the rainfall, as the performance of the model was better for those study catchments where coverage of rainfall stations was good. Hence a better spatial and temporal resolution of

rainfall data is required to further improve the model performance and to further enhance the understanding of the runoff processes in the basin.

770 **Acknowledgments**

We wish to acknowledge the efforts of the field staff and data collectors for their help during the installation of monitoring stations and data recording. We thank the project staff and MSc students for the logistic help and valuable field inputs. We are grateful to the Ministry of Water and Energy and National Meteorological Agency of Ethiopia for
775 making data available. This research was supported by the Belgian Development Cooperation (VLIR-UOS, WASETANA project).

References

- 780 [Abbaspour, K.C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J., and Srinivasan, R.: Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT, Journal of Hydrology, 333:413-430, 2007.](#)
- Allen, R.G., Pereira, L.S., Raes D., and Smith, M.: Crop evapotranspiration. Guide- lines for computing crop water requirements, FAO Irrigation and Drainage Paper 56,
785 FAO, Rome, 1998.
- Amanuel Zenebe, Vanmaercke, M., Poesen, J., Verstraeten, G., Nigussie Haregeweyn, Mitiku Haile, Kassa Amare, Deckers, J., and Nyssen, J.: Spatial and temporal variability of river flows in the degraded semi-arid tropical mountains of northern Ethiopia, *Zeitschrift für Geomorphologie*, 57(2): 143-169, 2013.
- 790 Antar, M.A., Ellassiouti, I., and Allam, M.N.: Rainfall-runoff modeling using artificial neural networks technique: a Blue Nile catchment case study, *Hydrological Processes* 20: 1201–1216, 2006.

795 [Arnold, J.G., Srinivasin, R., Muttiah, R.S., and Williams, J. R.: Large Area Hydrologic Modeling and Assessment: Part I. Model Development, JAWRA Journal of the American Water Resources Association 34\(1\):73-89, 1998.](#)

Bayabil, H.K., Tilahun, S.A., Collick, A.S., and Steenhuis, T.S.: Are runoff processes ecologically or topographically driven in the Ethiopian Highlands? The case of the Maybar, *Ecohydrology*, 3, 457–466, doi:10.1002/eco.170, 2010.

800 BCEOM. : Abay River Basin Integrated Development Master Plan – Phase 2 – Land Resources Development – Reconnaissance Soils Survey, Ministry of Water Resources: Addis Ababa; 208 pp, 1998a.

Beven, K.: On landscape space to model space mapping, *HP today, Hydrol. Process.*, 15, 323–324, 2001.

805 Bitew, M.M., and Gebremichael, M.: Assessment of satellite rainfall products for streamflow simulation in medium watersheds of the Ethiopian highlands, *Hydrol. Earth Syst. Sci.*, 15, 1147–1155, doi: 10.5194/hess-15-1147-2011, 2011.

Brooks, E.S., Boll, J., McDaniel, and P.A.: A hillslope-scale experiment to measure lateral saturated hydraulic conductivity, *Water Resour. Res.*, 40, W04208, doi: 10.1029/2003WR002858, 2004.

810 Collick, A.S., Easton, Z.M., Ashagrie, T., Biruk, B., Tilahun, S., Adgo, E., Awulachew, S.B., Zeleke, G., and Steenhuis, T.S.: A simple semidistributed water balance model for the Ethiopian highlands, *Hydrological Processes* 23: 3718–3727, 2009.

Conway, D.: A water balance model of the upper Blue Nile in Ethiopia, *Hydrological Sciences* 42: 265–282, 1997.

815 Conway, D.: The Climate and hydrology of the Upper Blue Nile River, *The Geographical Journal* 166: 49-62, 2000.

820 Descheemaeker, K., Poesen, J., Borselli, L., Nyssen, J., Raes, D., Mitiku Haile, Muys, B., and Deckers, J.: Runoff curve numbers for steep hillslopes with natural vegetation in semi-arid tropical highlands, northern Ethiopia, *Hydrological Processes* 22:4097-4105. (DOI: 10.1002/hyp.7011), 2008.

Domenico, P.A., and Schwartz, F.W.: *Physical and Chemical Hydrogeology*, John Wiley & Sons, New York, 824 p, 1990.

- 825 Fenicia, F., Savenije, H.H.G., Matgen, P., and Pfister, L.: Understanding catchment behavior through stepwise model concept improvement, *Water Resour. Res.*, 44, W01402, doi: 10.1029/2006WR005563, 2008.
- Food and Agriculture Organization of the United Nations, Rome (FAO): Guidelines for soil description, Rome, Italy, 2006.
- Frankenberger, J.R., Brooks, E.S., Walter, M.T., Walter, M.F., Steenhuis, T.S.: A GIS-Based Variable Source Area Hydrology Model, *Hydrological Processes*. 13(6):804-822, 1999.
- 830 [Gao, H., Hrachowitz, M., Fenicia, F., Gharari, S., and Savenije, H. H. G.: Testing the realism of a topography-driven model \(flex-topo\) in the nested catchments of the upper heihe, china, *Hydrol. Earth Syst. Sci.*, 18, 1895-1915, 10.5194/hess-18-1895-2014, 2014.](#)
- 835 [Gupta, H. V., Sorooshian, S., and Yapo, P.O.: Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration, *J. Hydrologic Eng.* 4\(2\): 135-143, 1999.](#)
- Hurni, H.: Third Progress Report, Soil Conservation Research Project, Vol. 4. University of Berne and the United Nations University, Ministry of Agriculture, Addis Ababa, 1984.
- 840 Istanbuluoglu, E., and Bras R. L.: Vegetation-modulated landscape evolution: Effects of vegetation on landscape processes, drainage density, and topography, *J. Geophys. Res.*, 110, F02012, doi: 10.1029/2004JF000249, 2005.
- 845 Johnson, P.A., and Curtis, P.D.: Water balance of Blue Nile River Basin in Ethiopia, *Journal of Irrigation and Drainage Engineering*: ASCE 120: 573–590, 1994.
- Jothityangkoon, C., Sivapalan, M., and Farmer, D.L.: Process controls of water balance variability in a large semi-arid catchment: downward approach to hydrological model development, *Journal of Hydrology* 254 (2001) 174-198, 2001.
- 850 Kebede, S., Travi, Y., Alemayehu, T., and Marc, V.: Water balance of Lake Tana and its sensitivity to fluctuations in rainfall, Blue Nile basin, Ethiopia, *Journal of Hydrology* 316: 233–247, 2006.

- Kebede, S.: Groundwater in Ethiopia: Features, Numbers and Opportunities, ISBN 3642303919, Springer, 297 p. DOI 10.1007/978-3-642-30391-3, 2013.
- 855 Kennedy, J., and Eberhart, R.: Particle Swarm Optimization, Proceedings of IEEE International Conference on Neural Networks, IV. pp. 1942–1948, 1995.
- Kim, U., and Kaluarachchi, J.J.: Application of parameter estimation and regionalization methodologies to ungauged basins of the Upper Blue Nile River Basin, Ethiopia, *Journal of Hydrology* 362: 39–52, 2008.
- 860 Krasnostein, A. L, and Oldham, C. E.: Predicting wetland water storage, *Water Resources Research*, vol. 40, W10203, doi: 10.1029/2003WR002899, 2004.
- Liu, B.M., Collick, A.S., Zeleke, G., Adgo, E., Easton, Z.M., and Steenhuis, T.S.: Rainfall-discharge relationships for a monsoonal climate in the Ethiopian highlands, *Hydrological Processes* 22(7): 1059–1067, 2008.
- 865 McCartney, M., Alemayehu, T., Shiferaw, A., and Awulachew, S. B.: Evaluation of current and future water resources development in the Lake Tana Basin, Ethiopia. Colombo, Sri Lanka: International Water Management Institute. 39p. (IWMI Research Report 134). doi:10.3910/2010.204, 2010.
- McWorter, D.B., and Sunada, D.K.: Groundwater Hydrology and hydraulics, *Water Resources Publications*, Ft. Collins.Co., 1997.
- 870 Mehta, V.K., Walter, M.T., Brooks, E.S., Steenhuis, T.S., Walter, M.F., Johnson, M., Boll, J., and Thongs, D.: Application of SMR to modeling watersheds in the Catskill Mountains, *Environmental Modeling and Assessment*. 9:77-89, 2004
- Dessie, M., Verhoest, N., Admasu, T., Pauwels, V., Poesen, J., Adgo, E., Deckers, J., and
875 Nyssen, J.: Effects of the floodplain on river discharge into Lake Tana (Ethiopia), *Journal of Hydrology*, 519, Part A: 699-710, 2014. DOI: 10.1016/j.jhydrol.2014.08.007.
- Melesse Temesgen, Savenije, H.H.G, Rockström, J, and Hoogmoed, W.B.: Assessment of strip tillage systems for maize production in semi-arid Ethiopia: Effects on
880 grain yield, water balance and water productivity, *Physics and Chemistry of the Earth, Parts A/B/C* 47–48(0): 156-165, 2012a
- Melesse Temesgen, Uhlenbrook, S., Simane, B., van der Zaag, P., Mohamed, Y., Wenninger, J., and Savenije, H.H.G: Impacts of conservation tillage on the

- hydrological and agronomic performance of Fanya juus in the upper Blue Nile
885 (Abbay) river basin, *Hydrol. Earth Syst. Sci.*, 16, 4725–4735, 2012b.
- Miserez, A.: Soil erodibility and mapping in different hydrological land systems of Lake
Tana basin, Ethiopia, M.Sc. thesis study, Department of Soil and Water
Management, KULeuven, Leuven, Belgium, 2013.
- Mishra, A., Hata, T., and Abdelhadi, A.W.: Models for recession flows in the upper Blue
890 Nile River, *Hydrological Processes* 18: 2773–2786, 2004.
- Mishra, A., and Hata, T.: A grid-based runoff generation and flow routing model for the
upper Blue Nile Basin, *Hydrological Sciences Journal* 51: 191–206, 2006.
- Nash, J.E., and Sutcliffe, J.V.: River flow forecasting through conceptual models part I: a
discussion of principles, *Journal of Hydrology* 10: 282–290, 1970.
- 895 Poppe, L., Frankl, A., Poesen, J., Teshager Admasu, Mekete Dessie, Enyew Adgo,
Deckers, J., and Nyssen, J.: Geomorphology of the Lake Tana basin, Ethiopia.
Journal of Maps, Vol. 9, Iss. 3, DOI:10.1080/17445647.2013.801000, 2013
- Savenije, H. H. G.: HESS Opinions "Topography driven conceptual modelling (FLEX-
Topo)", *Hydrol. Earth Syst. Sci.*, 14, 2681-2692, doi: 10.5194/hess-14-2681-2010,
900 2010.
- Scheerlinck, K., Pauwels, V. R. N., Vernieuwe, H., and Baets, B. De.: Calibration of a
water and energy balance model: Recursive parameter estimation versus particle
swarm optimization, *Water Resour. Res.*, 45, W10422, doi:
10.1029/2009WR008051, 2009.
- 905 Setegn, S.G., Ragahavan, S., and Bijan, D.: Hydrological modelling in the Lake Tana
Basin, Ethiopia using SWAT model, *The Open Hydrology Journal* 2: 49–62,
2008.
- [Sivapalan, M.: The secret to 'doing better hydrological science': Change the question!.
Hydrol. Process., 23, 1391–1396, doi:10.1002/hyp.7242, 2009.](#)
- 910 Steenhuis, T.S., and van der Molen, W.H.: The Thornthwaite-Mather procedure as a
simple engineering method to predict recharge, *Journal of Hydrology* 84: 221–
229, 1986.
- Steenhuis, T. S., Collick, A. S., Easton, Z. M., Leggesse, E. S., Bayabil, H. K., White, E
.D., Awulachew, S. B., Adgo, E., and Ahmed, A. A.: Predicting discharge and

- 915 erosion for the Abay (Blue Nile) with a simple model, *Hydrol. Process.* 23, 3728–
3737, 2009.
- Tekleab, S., Uhlenbrook, S., Mohamed, Y., Savenije, H.H.G, Temesgen, M., and
Wenninger, J.: Water balance modeling of the upper Blue Nile catchments using a
top-down approach, *Hydrology and Earth System Science* 15: 2179–2193. DOI:
920 10.5194/hess-15-2179, 2011.
- Tilahun, S. A., Guzman, C.D., Zegeye, A.D., Engda, T.A., Collick, A.S., Rimmer, A.,
and Steenhuis, T.S.: An efficient semi-distributed hillslope erosion model for the
subhumid Ethiopian Highlands, *Hydrol. Earth Syst. Sci.*, 17, 1051–1063, 2013,
doi: 10.5194/hess-17-1051-2013, 2013.
- 925 Uhlenbrook, S., Mohamed, Y., and Gragne, A.S.: Analyzing catchment behavior through
catchment modeling in the Gilgel Abay, Upper Blue Nile River Basin, Ethiopia,
Hydrol. Earth Syst. Sci., 14, 2153–2165, doi: 10.5194/hess-14-2153-2010, 2010.
- Vogel, R.M., and Kroll, C.N.: Regional geohydrologic-geomorphic relationships for the
estimation of low-flow statistics, *Water Resources Research* 28(9): 2451–2458,
930 1992.
- Wale, A., Rientjes, T.H.M., Gieske, A.S.M., and Getachew, H.A.: Ungauged catchment
contributions to Lake Tana’s water balance, *Hydrological Processes* 23: 3682–
3693, 2009.
- Wittenberg, H.: Base flow recession and recharge as nonlinear storage processes, *Hydrol.*
935 *Process.* 13, 715-726, 1999.

940

945

Figure captions

- 950 **Fig. 1.** The Upper Blue Nile basin and the Lake Tana sub-basin (a) and the study catchments and the gauging stations in the Lake Tana sub-basin georeferenced on the SRTM DEM (b)
- 955 **Fig. 2.** The modeling approach showing (a) divisions of a catchment into different runoff production areas, (b) conceptual model configuration of the soil surface at an outlet of a catchment and (c) Inflows and outflows for the soil reservoir when the soil water storage capacity is (i) below field storage capacity, (ii) greater than field storage capacity and (iii) greater than the maximum soil water storage (after Krasnostein and Oldham, 2004).
- 960 **Fig. 3.** Typical surfaces with poor infiltration on hillslopes in the Gumara catchment: (a) shallow soil overlying bedrock, and (b) plough pan with typical plough marks. The occurrence of high runoff response on these surfaces is evidenced by the presence of rill erosion (Photos: Elise Monsieurs)
- 965 **Fig. 4.** The three slope categories for the Gilgel Abay and Gumara catchments
- Fig. 5** [Major soil types in the Lake Tana basin and the study catchments](#)
- Fig.6.** [Soil depth in the Lake Tana basin and the study catchments](#)
- Fig.7.** [Location map of rainfall stations for the study catchments](#)
- 970 **Fig.8.** [Stage-Discharge relationship \(Rating curves\) for Gilgel Abay at Picolo and Gumara at Wanzaye Stations](#)
- Fig.9.** [Comparison of predicted and observed discharge and precipitation of the Gumara and the Gilgel Abay catchments for the calibration period](#)
- Fig.10.** [Predicted and observed discharges and precipitation of the Gumara and the Gilgel Abay catchments for the validation period](#)
- Fig.11.** [Predicted and observed flow duration curves of the Gumara and the Gilgel Abay catchments for the calibration period](#)
- Fig.12.** [Predicted and observed flow duration curves of the Gumara and the Gilgel Abay catchments for the validation period](#)
- 975 **Fig.13.** [One of the springs in Gilgel Abay catchment used as a water supply source for Bahir Dar town](#)

[Fig.14. Local model parameter sensitivity analysis for Gumara catchment. Parameters are explained in Table 2](#)

[Fig.15. Global model parameter sensitivity analysis results for Gumara catchment. Parameters are explained in Table 2](#)

980

Tables

985

Table 1. Input data on topography, soil and saturated hydraulic conductivities for the study catchments as classified into different hydrological regimes using topography

Catchment	Slope class	Average slope (%)	Coverage from the total area (%)	Average Soil depth (m)	Dominant soil texture	Porosity	Field capacity	Saturated hydraulic conductivity $K_{s,e}$ (m/s)
Gilgel Abay	Level ($\leq 8\%$)	3.4	54	0.92	clay	0.46	0.36	9.26×10^{-8}
	Hilly ($8\% < \text{slope} \leq 30\%$)	15.9	38	1.29	Clay to clay loam	0.42	0.32	
	Steeply ($>30\%$)	41.4	8	1.49	Clay loam to Silty loam	0.4	0.26	
Gumara	Level ($\leq 8\%$)	4.0	24	1.5	clay	0.46	0.36	1.16×10^{-8}
	Hilly ($8\% < \text{slope} \leq 30\%$)	17.2	60	1.24	Loam , Silty clay	0.42	0.26	
	Steeply ($>30\%$)	41.5	16	1.2	Sandy loam	0.25	0.1	

990

995

1000

Table 2. Model parameters, their ranges, and calibrated values found in 2500 iterations in the PSO calibration

Parameter	Explanation	units	Minimum	Maximum	calibrated values		Average value of both catchments
					Gumara	Gilgel Abay	
β	parameter to account variability of permeability of soil with soil water storage	–	1	3	2.445	2.314	2.380
k1	relates discharge and storage for the ground water	–	0.1	2	0.971	1.012	0.992
$K_{s,u}$	Saturated hydraulic conductivity in the upper soil layer	m/s	0.001	0.1	0.016	0.05	0.033
γ	parameter to account variability of deep percolation with soil water storage	–	0.5	2	1.409	0.9	1.155
λ	coefficient that represents part of catchment that is impermeable	–	0.05	0.5	0.149	0.173	0.161
α_1	interflow partitioning coefficient for the steep slope surface	–	0.05	0.8	0.653	0.575	0.614
α_2	interflow portioning coefficient for the medium slope surface	–	0.05	0.8	0.065	0.152	0.109

1005

1010

1015

1020

1025

Table 3. Statistical comparison and model performance of the modelled and observed river discharge (Q) for the two catchments

		Model performance indicators					
		Mean Q [mm/day]	Standard Deviation [mm/day]	RMSE ¹ [mm/day]	NSE ^{2*}	R ²	PBIAS ³
Observed data	Gumara						
	calibration (2011-2012)	2.31	3.79	—	—	—	—
	validation (2000-2005)	2.3	3.75	—	—	—	—
	Gilgel Abay						
	calibration (2012)	3.89	5.05	—	—	—	—
	validation (2000-2005)	2.33	3.4	—	—	—	—
Wase - Tana model	Gumara						
	calibration (2011-2012)	2.37	3.56	1.34	0.86	0.86	3.30
	validation (2000-2005)	1.95	3.05	1.37	0.78	0.8	-11.75
	Gilgel Abay						
	calibration (2012)	3.85	4.7	1.85	0.84	0.85	-21.61
	validation (2000-2005)	3.14	3.71	1.67	0.7	0.8	34.06
SWAT model	Gumara						
	calibration (2011-2012)	1.91	3.33	1.55	0.77	0.78	-17.50
	validation (2000-2005)	1.62	3.11	1.63	0.72	0.75	-29.48
	Gilgel Abay						
	calibration (2012)	2.02	3.20	1.40	0.60	0.79	-44.01
	validation (2000-2005)	2.45	3.86	2.30	0.55	0.63	5.45
Flex_B model	Gumara						
	calibration (2011-2012)	2.43	3.64	1.54	0.82	0.82	5.30
	validation (2000-2005)	2.01	3.35	1.47	0.80	0.81	-12.67
	Gilgel Abay						
	calibration (2012)	3.81	4.03	1.62	0.80	0.84	5.64
	validation (2000-2005)	4.13	4.33	2.15	0.50	0.75	77.67

1030

1. *RMSE* : Root Mean Squared Error as defined in Equation (22)

2*. *NSE* : Nash-Sutcliffe Efficiency as defined in Equation (23)

3. *PBIAS*: Percentage Bias as defined in Equation (25)

1035

1040

Table 4. Model results on the hydrograph components of the catchments

Runoff components	unit	For the calibration period		For the validation period	
		Gumara	Gilgel Abay	Gumara	Gilgel Abay
Total mean annual runoff predicted (Q_{pr})	mm/year	864	1405	713	1146
Total mean annual runoff observed (Q_{ob})	mm/year	843	1420	841	938
Mean annual surface runoff (Q_{se})	mm/year	161	280	129	234
	% from the total Q_{pr}	19	20	18	20
Mean annual interflow (Q_{ss})	mm/year	574	508	458	369
	% from the total Q_{pr}	66	36	64	32
Mean annual baseflow (Q_2)	mm/year	128	617	126	548
	% from the total Q_{pr}	15	44	18	48

1045

Table 5. Comparison of model performance between the optimal and average model parameters of the three catchments

catchment		Model performance for the optimal model parameters			Model performance for the average of the optimal model parameters of the two catchments		
		RMSE [mm/day]	<i>NSE</i>	R^2	RMSE [mm/day]	<i>NSE</i>	R^2
Gumara	Calibration period	1.34	0.86	0.86	1.48	0.84	0.86
	Validation period	1.37	0.78	0.80	1.82	0.76	0.77
Gilgel Abay	Calibration period	1.85	0.84	0.85	1.98	0.83	0.84
	Validation period	1.67	0.70	0.80	1.93	0.68	0.78
Dirma	For the 2012 discharge	-	-	-	1.79	0.58	0.60

1050

1055

Figures

1060

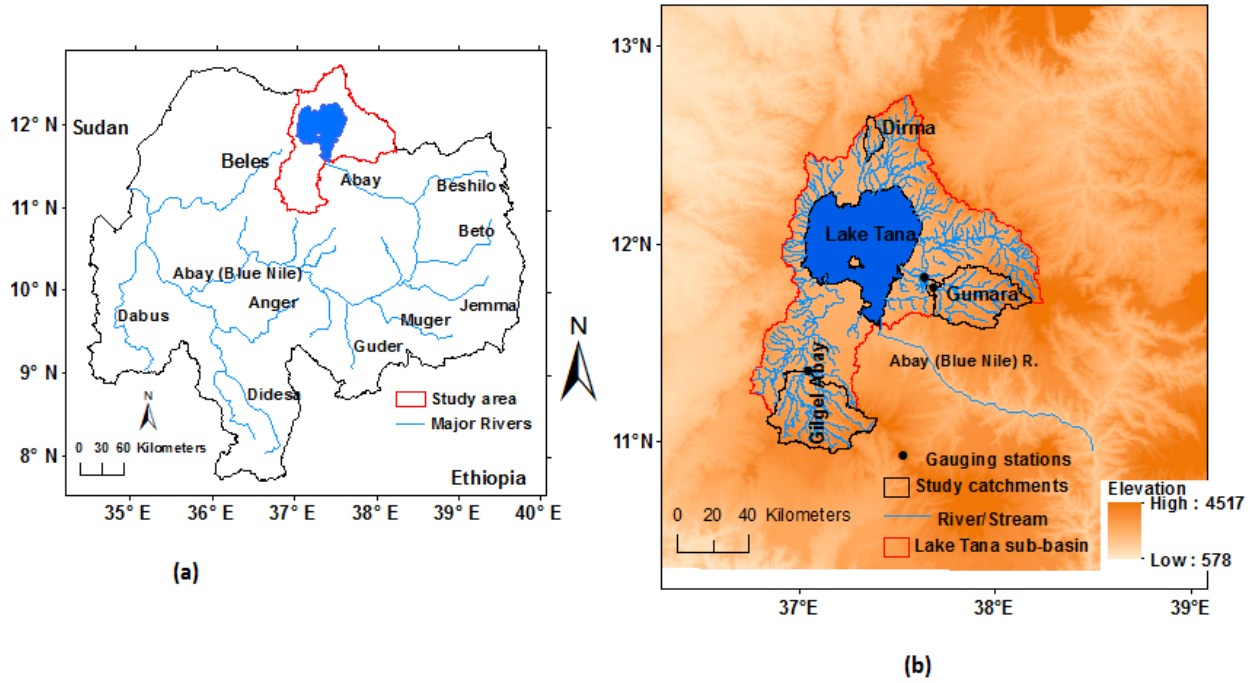


Fig. 1. The Upper Blue Nile basin and the Lake Tana sub-basin (a) and the study catchments and the gauging stations in the Lake Tana sub-basin georeferenced on the SRTM DEM (b)

1065

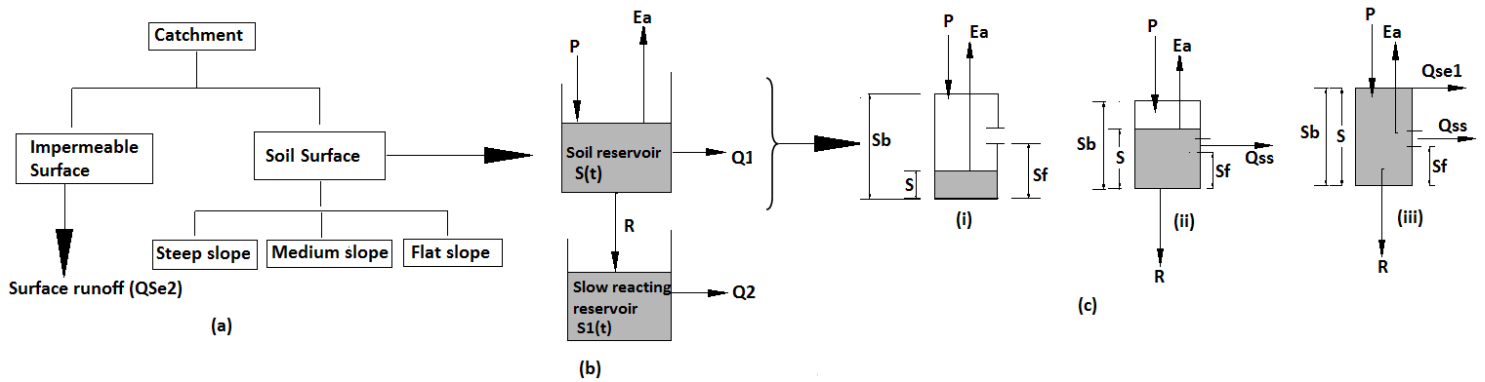
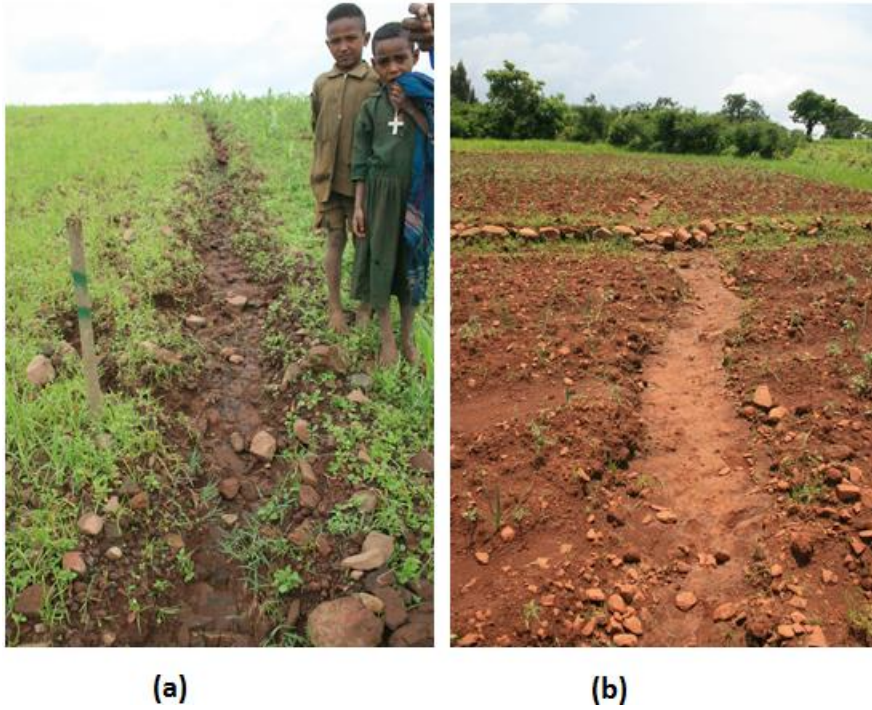


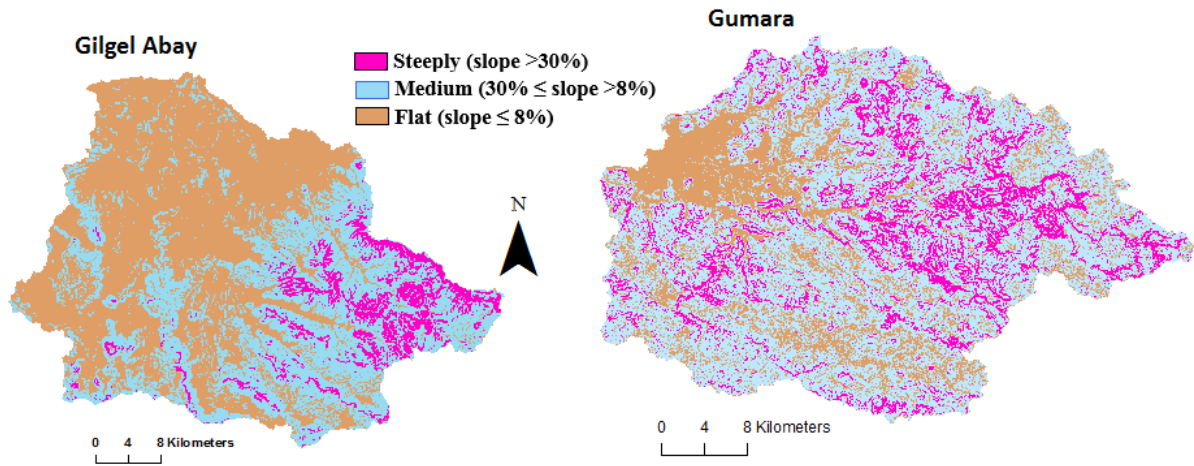
Fig.2. The modeling approach showing (a) divisions of a catchment into different runoff production areas, (b) conceptual model configuration of the soil surface at an outlet of a catchment and (c) Inflows and outflows for the soil reservoir when the soil water storage capacity is (i) below field storage capacity, (ii) greater than field storage capacity and (iii) greater than the maximum soil water storage (after Krasnostein and Oldham, 2004).

1070



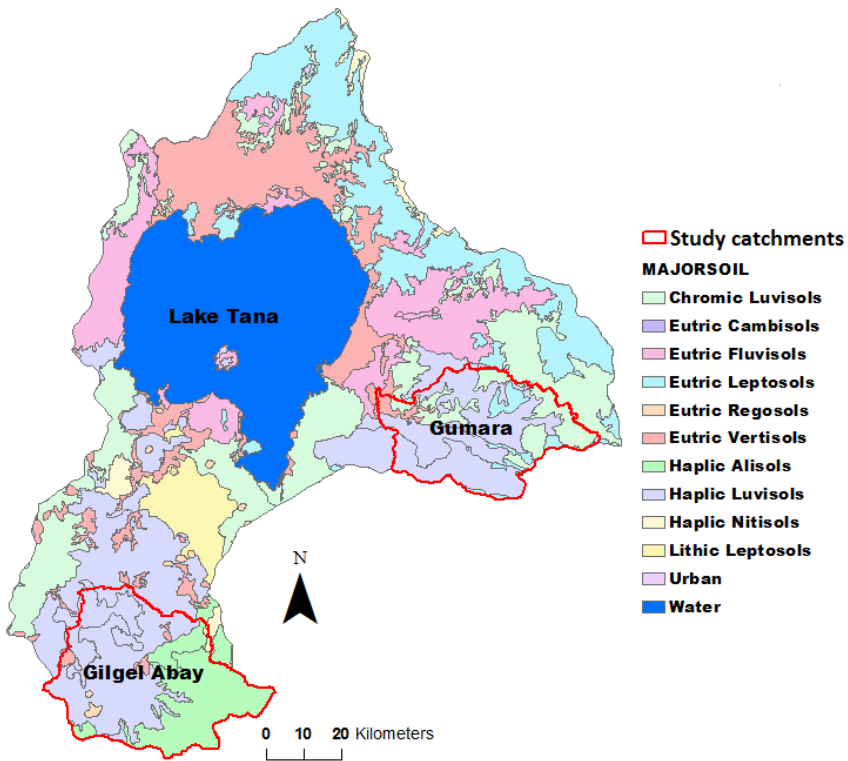
1075

Fig. 3. Typical surfaces with poor infiltration on hillslopes in the Gumara catchment: (a) shallow soil overlying bedrock, and (b) plough pan with typical plough marks. The occurrence of high runoff response on these surfaces is evidenced by the presence of rill erosion (Photos: Elise Monsieus)

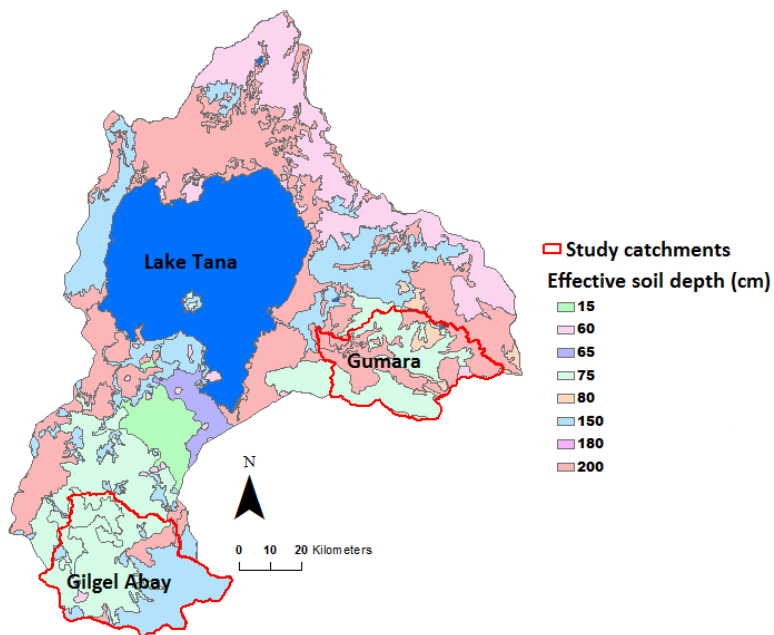


1080

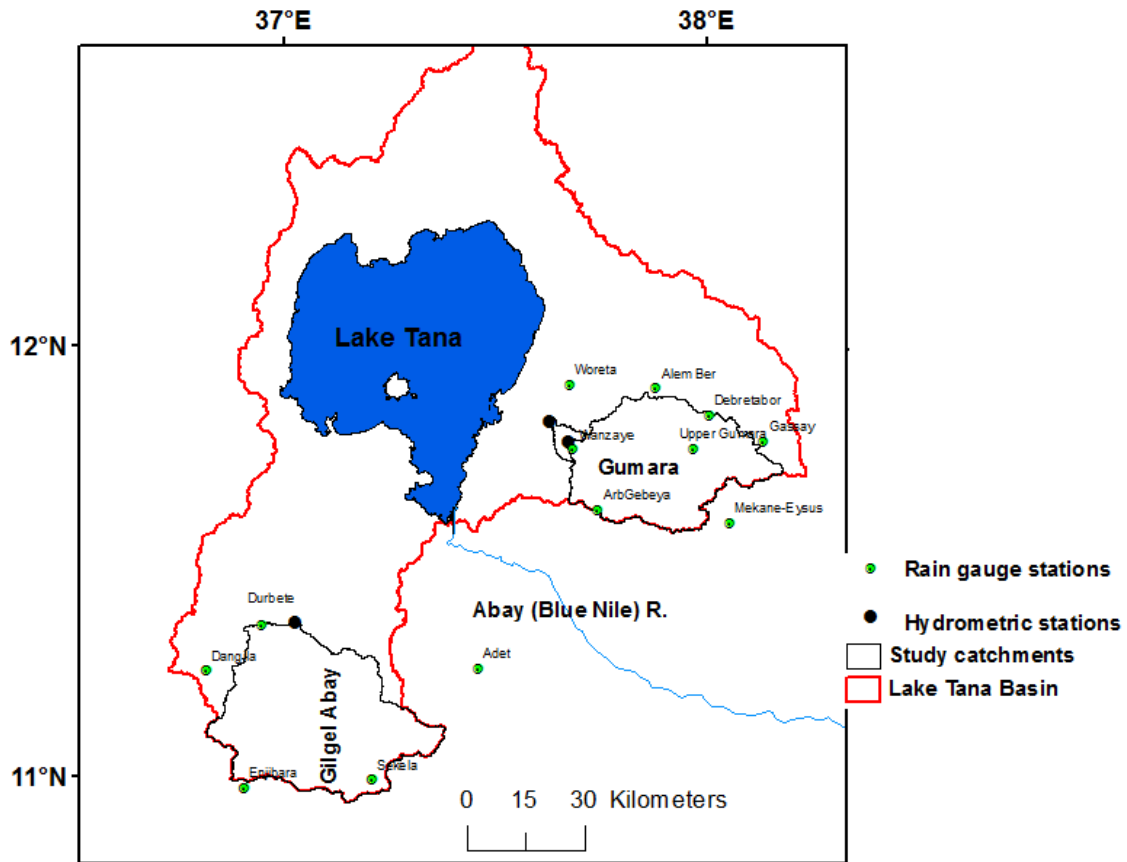
Fig.4. The three slope categories for the Gilgel Abay and Gumara catchments



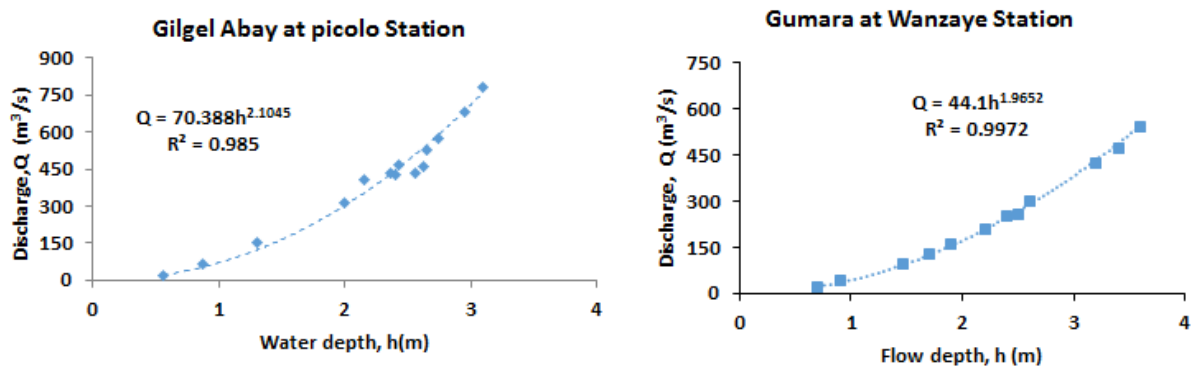
1085 **Fig.5.** Major soil types in the Lake Tana basin and the study catchments



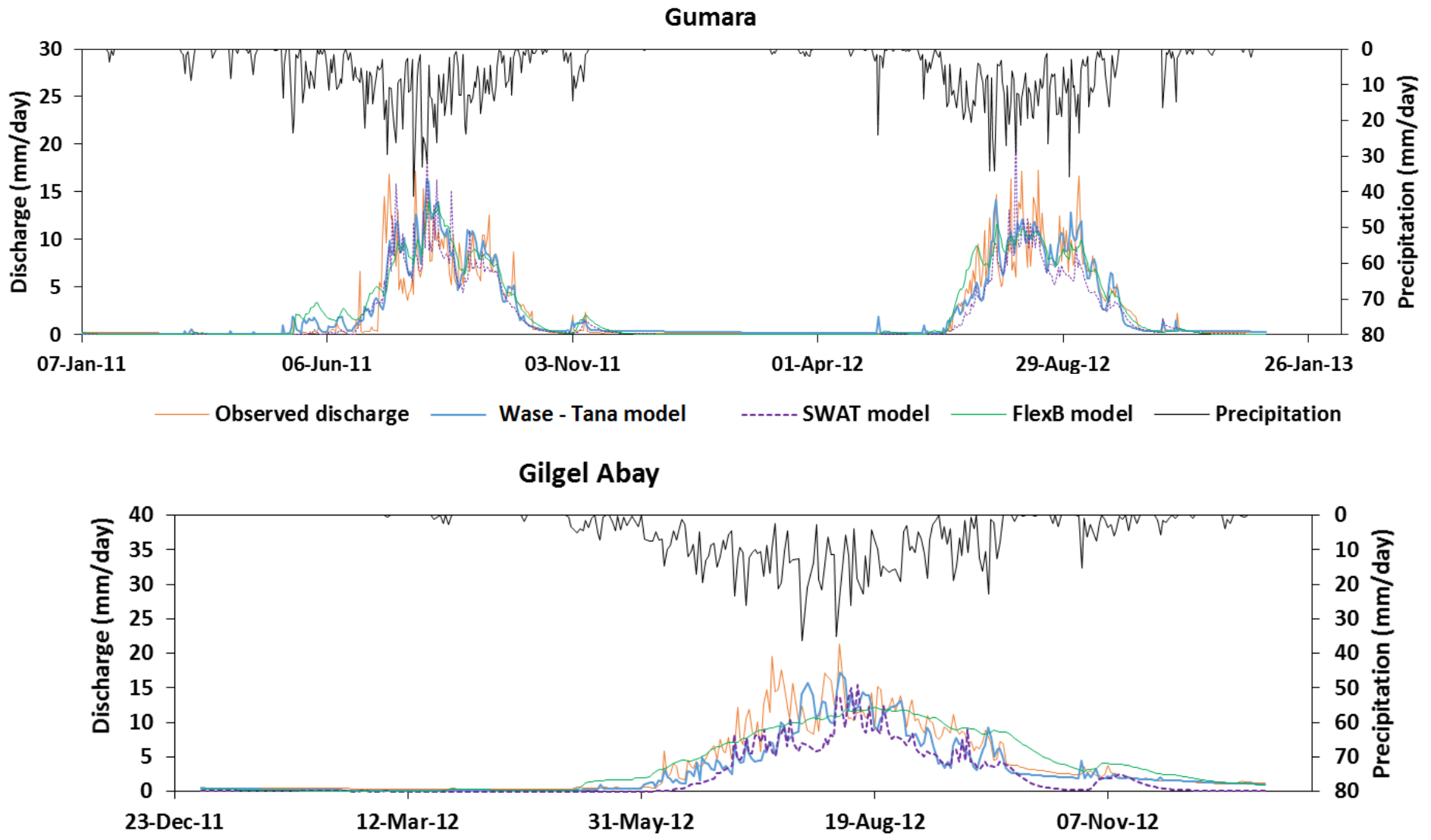
1090 **Fig.6.** Soil depth in the Lake Tana basin and the study catchments



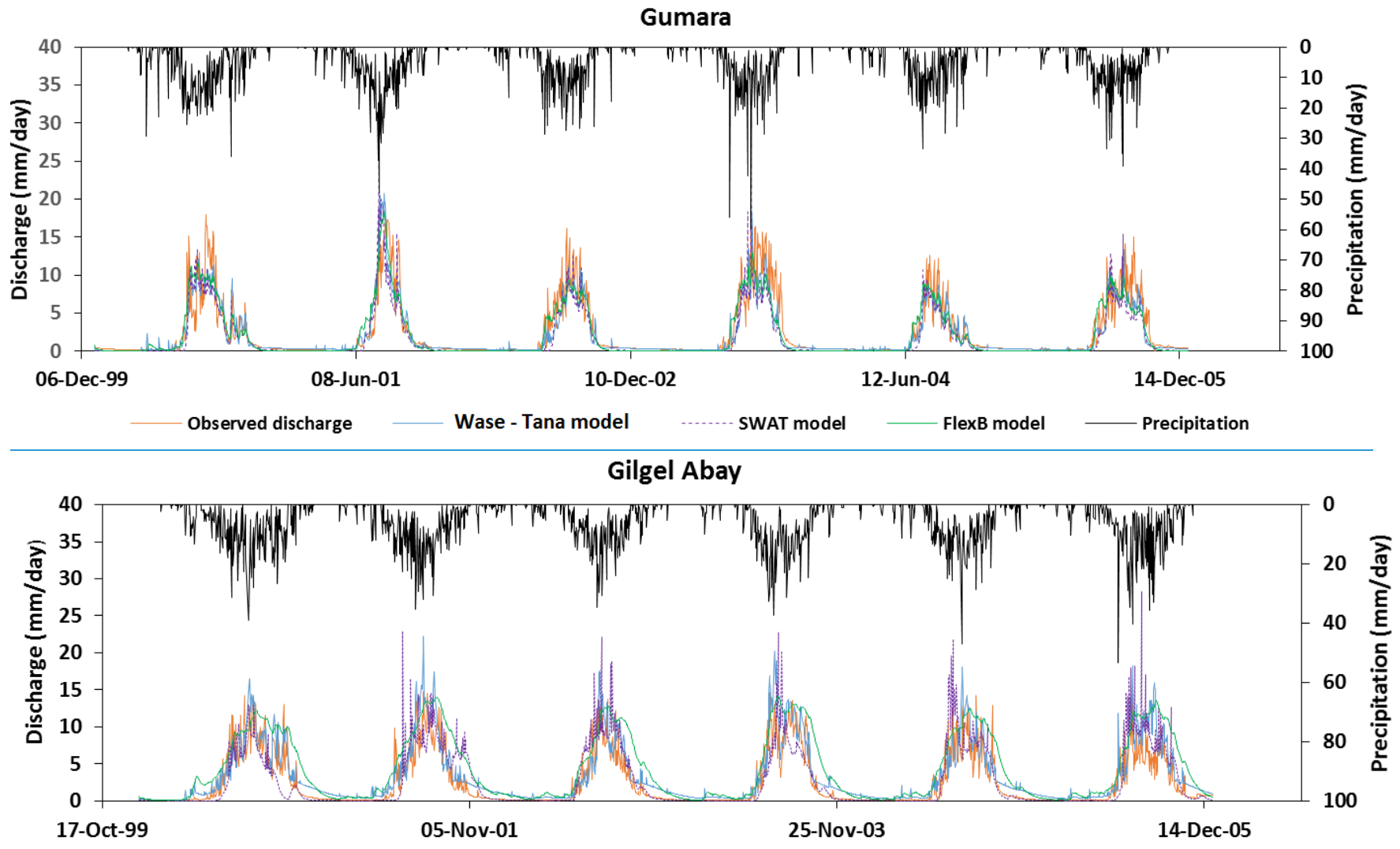
1095 Fig.7. Location map of rainfall stations for the study catchments



1100 Fig.8. Stage-Discharge relationship (Rating curves) for Gilgel Abay at Picolo and Gumara at Wanzaye Stations



1105 [Fig. 95. Comparison of predicted and observed discharge and precipitation of the Gumara and the Gilgel Abay catchments for the calibration period](#)



1110

Fig.10-6. Predicted and observed discharges and precipitation of the Gumara and the Gilgel Abay catchments for the validation period

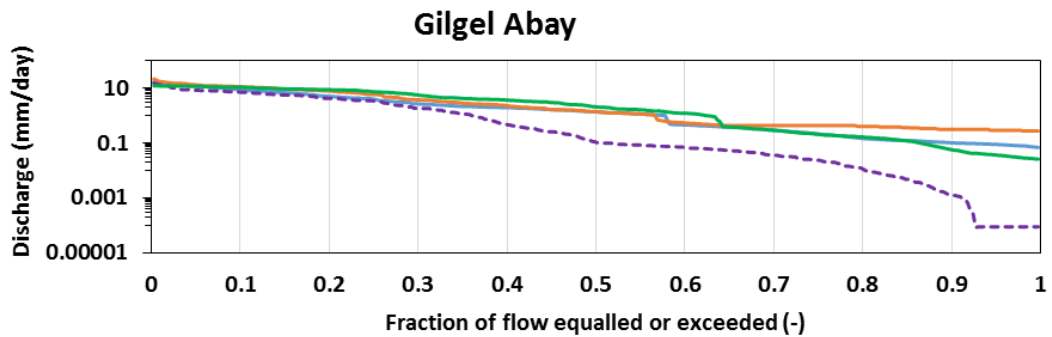
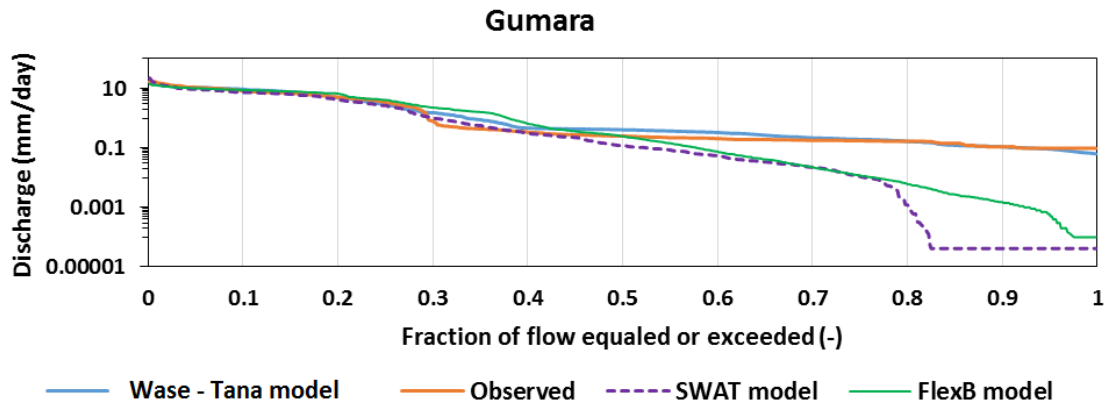


Fig. 11. Predicted and observed flow duration curves of the Gumara and the Gilgel Abay catchments for the calibration period

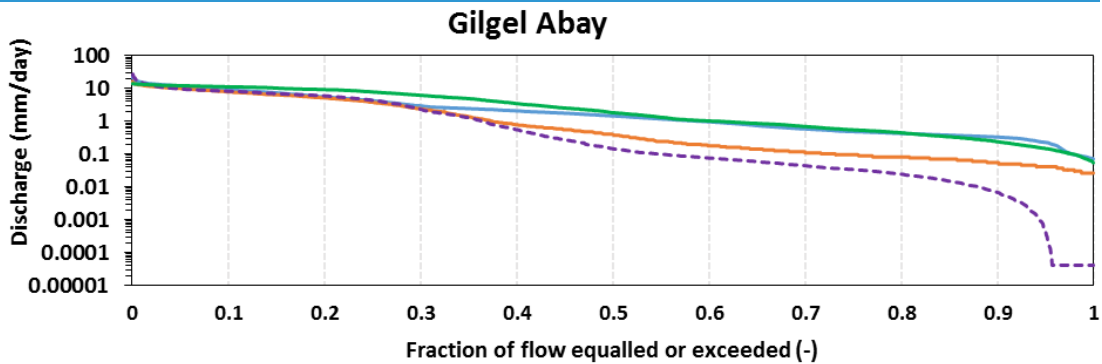
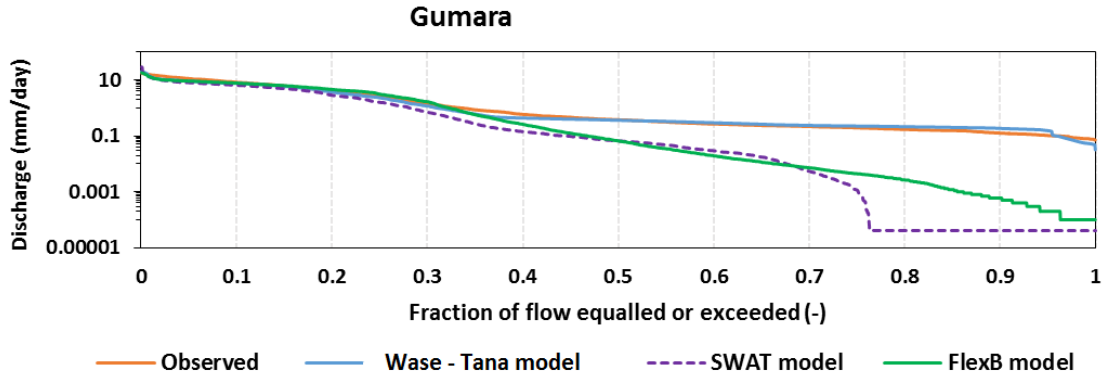


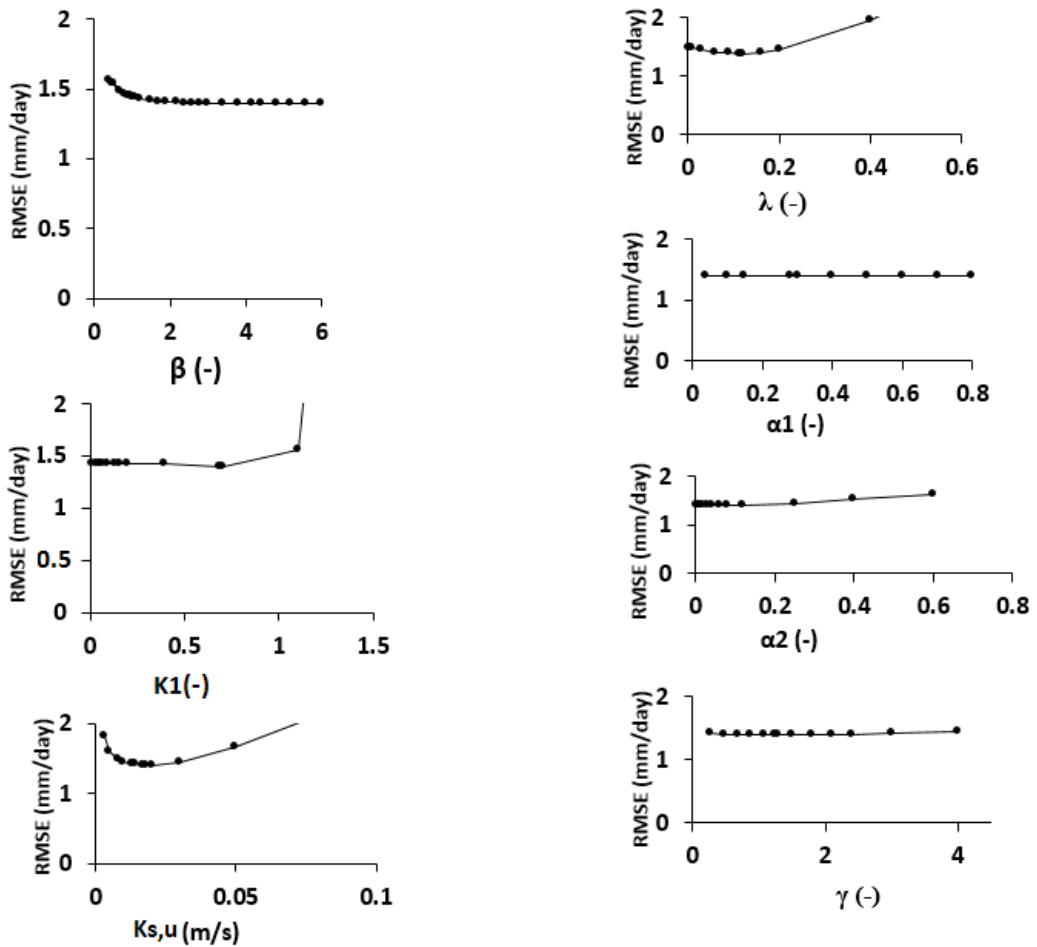
Fig.12. Predicted and observed flow duration curves of the Gumara and the Gilgel Abay catchments for the validation period

1115



1120

Fig.13. One of the springs in Gilgel Abay catchment used as a water supply source for Bahir Dar town



1125

Fig. 147. Local model parameter sensitivity analysis for Gumara catchment. Parameters are explained in table 2.

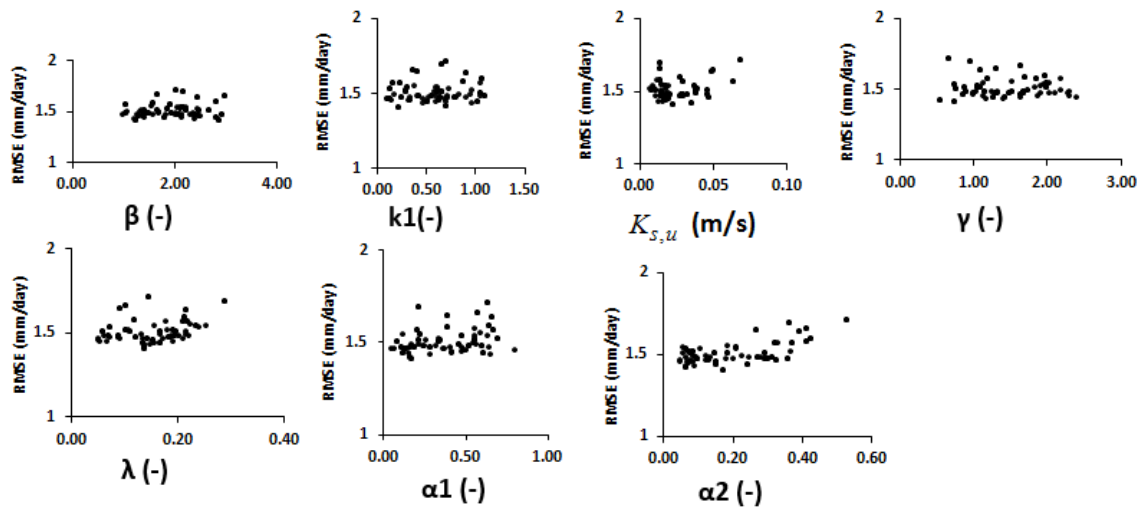


Fig.15. Global model parameter sensitivity analysis results for Gumara catchment. Parameters are explained in table 2.

1130