Interactive comment on "Comparing TRMM 3B42, CFSR and ground-based rainfall estimates as input for hydrological models, in data scarce regions: the Upper Blue Nile Basin, Ethiopia" by A. W. Worqlul et al.
Response to Anonymous Referee #1

We would like to thank the Anonymous Referee #1 for his time and effort spent on reviewing our manuscript. We appreciate that the referee likes the clarity, that the structure is easy to follow and "most figures and tables are appropriated (and) express what authors want to demonstrate" However, it is unfortunate that some of statements in this review are based on personal knowledge that are contrary to what generally assumed as valid in the literature. In the section below, we comment particularly on the accuracy of observed rainfall gage data as model input and on the choice of the models. In addition we discuss the choice of the satellite rainfall products. Finally we clarify the objective of our study since the referee states that:

> "although the objective of the paper itself is clear, I failed to understand the methodology authors presented: using a modeling approach to test the suitability of the precipitation products"

1. Paper objectives

We agree with the reviewer that it is difficult to understand the connection between the objective and the methodology. Our objective was (Page 2083 line 5 to 12):

> "The objective of this study is to compare the advantages and the limitation of commonly used high-resolution satellite rainfall products as input to hydrological models as compared to sparsely populated network of rain gauges. For this comparison we use two semi-distributed hydrological models Hydrologiska Byråns Vattenbalansavdelning (HBV) and Parameter Efficient Distributed (PED) that performed well in Ethiopian highlands in two watersheds: the Gilgel Abay with relatively dense network and Main Beles with relatively scarce rain gauge stations. Both are located in the Upper Blue Nile Basin."

Inherently we assumed (but not clearly stated in the manuscript) that the weather data that best correlate with watershed streamflow is the best representation of the weather occurring over the watershed. Similarly to the methods proposed in Podobnik and Stanley (2008) and employed in Fuka et al (2013) rainfall products can be evaluated though the use of hydrological model acting as a filter between the physical forcing variables and the resulting streamflow response. In other words the basin can be considered one large rain gage and we use the model to unscramble the outflow signal to evaluate if the daily amount of water in the rain gage is appropriate. More properly stated we evaluate the mutual correlation between the weather forcing variables and the resultant streamflow though the use of models

We recognize that fully or partly distributed models would result in many calibration degrees of freedom. Thus, we selected two models (HBV and PED) that had proven to give good prediction in the Ethiopian highlands with the number of calibration parameters or degrees of freedom in the calibration, and thus reduce over parameterization or over-fitting issues.

Because we did not explain very well in the discussion the underlying assumptions, it is understandable (but invalid) that the referee states (at the end of his review).

> "Results and discussion p2094 Line 7 – Line 8: This is exactly an example of parameter uncertainty. "Fraction of hillside area" in Table 2 and FC and PERC in Table 3 shows that different parameters pair can result in similar model preference. Therefore, you won't be able to quantify the true effect of input rainfall theoretically"

Although it is not completely clear what the referee means, it is very likely related to the question of equifinality as introduced by Beven and Freer (2001). Equifinality in our case is an advantage because when the model are insensitive to the characteristics of the watershed, the parameters cannot be changed in such a way that it appears that the rainfall product is more accurate than it is in reality.

In summary as stated in the objectives we are interested in evaluating the three rainfall products for the Ethiopian landscape with poor coverage of ground based. We fully recognize that a model with many parameters can make up for the shortcoming of the rainfall data but that would be inconsistent with the physical meaning of the parameters.

2. Accuracy of ground based rainfall data

The reviewer states (at the end of his comment) that

> "Why is it a surprise that ground based data performed better than satellite-based products? Ground based data should be considered as truth data that satellite-based products try to calibrate to".

The implication of this statement is that ground based products are "the gold standard" for predicting stream discharge. This is contrary to the literature findings we are familiar with, For example Mehta et al. (2004) reported that in modeling the watershed output several peak flow were underpredicted in a watershed in the Catskills that was located near two rain gages 30 km apart. They linked the poor prediction to problems with the ground gauges observation because the daily rainfall for these two stations had a regression coefficient of 23%. Thus proving ground rainfall measurements are point measurements which are not representative of the spatial pattern. In another study Fuka et al. (2013) indicated that CFSR precipitation data performed as good as or better than models forced by gauging stations especially when stations are located more than 10 km from the watershed. Ground rainfall measurement lacks areal representation, the gauged measurements cannot be treated as the ground truth reference for the areal-average rainfall (Cohen Liechti et al., 2012;Wang and Wolff, 2010).

The World Meteorological Organization (WMO, 1994) made an extensive study on the desired density of ground based precipitation networks and recommended that the minimum rainfall station network density recommended for tropical regions is 100 to 250 km^2 per station for mountainous regions such as both watershed are located in. Rainfall gauging stations in both study sites are scares (Figure 2 page 2104). The 1650 km^2 Gilgel Abay watershed has only four rainfall gauging stations in and around the watershed; three of them are outside of the watershed. The 3212 km^2 Main Beles has three station and all of them are outside of the watershed. Thus in both watershed the ground based network according to the well-researched WMO standard is not representative for the actual rainfall falling in the basin.

In summary, it is understandable but unfortunate that due the missing underlying model assumption in the discussion paper writes

> "In my opinion, authors should design their experiment in the following steps: 1) choose physically-based models; 2) calibrate model only used observation data and 3) validate model preferences with satellite-based rainfall products. In this way, both model structure uncertainty and parameter uncertainty are controlled so the different results are for sure coming from input data uncertainty"

Thus the referee's statement that ground based gage data are by definition better than satellite based products is not supported by findings in the literature and should not be considered a reason to reject our publication as suggested by the referee.

3. Choice of models:

We were taken back by the comment that we should have used "a physical model". What model does the reviewer think we should use? The fact is that we have tried several models and have seen the performance of others. Why we chosen the particular model is described clearly in the introduction part (Page 2084 and 2085 line 24 to 26 and line 1 to 12, respectively). These two models predicted in most cases the outflow much better on a daily basis than other models used in the Ethiopian highlands (Steenhuis et al., 2009;Steenhuis et al., 2013;Tilahun et al., 2013b;Tilahun et al., 2013a;Abdo et al., 2009;Collick et al., 2009;Wale et al., 2009;Bitew and Gebremichael, 2011).

4. Choice of Rainfall products

The reviewer suggest that TRIMM is not good choice

> "A simple data analysis (as authors did in Figure 4) already demonstrates that TRMM is not a good input data for hydrologic models. The long-term climatology showed that TRMM has different monthly pattern and different annual volume in precipitation than observation, and one do not need to run the model to know that TRMM is not suitable. For other case study area this might be different, then I would suggest authors change their study area."

Although this paper proves that the use of TRIMM is not recommended for the Ethiopian highlands , in South America Collischonn et al. (2008) and Su et al. (2008) showed that TRMM 3B42 rainfall estimate can be used as an alternative as input to rainfall-runoff modelling.

We do not understand the last part of the critics "For other case study area this might be different, then I would suggest authors change their study area." We would like to understand in order to respond to this comment. Why we should change our study area. We describe clearly that we choose one basin with a relative dense network and one basin with a sparse network. Moreover both station had good quality outflow data.

5. Automatic calibration vs manual calibration

Automatic calibration might yields as good as or better result than manual calibration. But for a parsimonious semi-distributed hydrological model with a limited number of parameters a manual calibration can capture reasonably the observed flow. We have to remember that automated calibration approaches cannot replace the insights of the modular physical understanding and knowledge of the effects of parameters on the system response (Abbaspour et al., 2007). It is also true that, there is no unique model parameter set able to optimize all objective functions simultaneously in automatic calibration (Madsen, 2000).

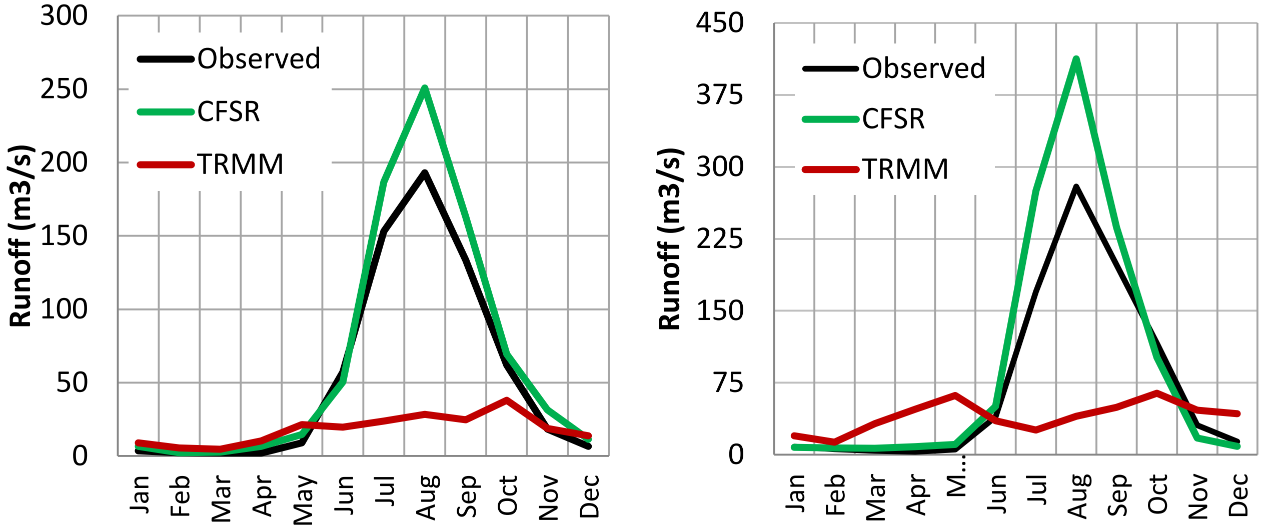
6. Use ground based observations for calibrating satellite products

As cited above already the referee suggested to validate the satellite rainfall using the calibrated model parameter sets of gauged rainfall. Although we do agree with this procedure we have used the calibrated and validated model parameter sets using gauged rainfall for PED and HBV models to simulate the satellite rainfall estimates.

In the simulation, the optimized model parameter sets for the gauged rainfall are used to simulate CFSR and TRMM rainfall estimate for both watersheds. This procedure is done by substituting the gauged rainfall by CFSR and TRMM rainfall estimate and keeping the optimised model parameters of the gauged rainfall. The performance of the simulated flow for both watersheds is shown under Table S1 and Figure S1. We clearly see that the performance of the simulated flow has deteriorated, the relative volume difference between simulated and observed flow has increased as indicated by PBIAS. If we compare the performance of both watersheds, Gilgel Abay performed better for both TRMM and CFSR data. It is obvious that the difficulty with the approach suggested by the referee that the gauged rainfall used to calibrate the watershed is not representative of the watershed especially for Main Beles, where none of the gauged rainfall stations used to calibrate the watershed are inside the watershed.

Table S1. Performance of CFSR and TRMM simulated by calibrated and validated model parameters for gauged flow.

	Gilgel Abay		Main Beles	
	TRMM	CFSR	TRMM	CFSR
PBIAS	55.0	-24.0	53.4	-36.0
NSE	-0.18	0.57	-0.07	0.29
R^2	0.2	0.7	0.01	0.61



(a) Gilgel Abay (b) Main Beles

Figure S1. Flow simulated TRMM and CFSR data using calibrated and validated model parameters of gauged rainfall using HBV model, (a) Gigel Abay and (b) Main Beles.

We thank the referee for his efforts in reviewing this paper.

Tammo Steenhuis and Abeyou Worqlul

References

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.

Abdo, K., Fiseha, B., Rientjes, T., Gieske, A., and Haile, A.: Assessment of climate change impacts on the hydrology of Gilgel Abay catchment in Lake Tana basin, Ethiopia, Hydrological Processes, 23, 3661-3669, 2009.

Beven, K., and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, Journal of hydrology, 249, 11-29, 2001.

Bitew, M., and Gebremichael, M.: Assessment of satellite rainfall products for streamflow simulation in medium watersheds of the Ethiopian highlands, Hydrology and Earth System Sciences, 15, 1147-1155, 2011.

Cohen Liechti, T., Matos, J., Boillat, J.-L., and Schleiss, A.: Comparison and evaluation of satellite derived precipitation products for hydrological modeling of the Zambezi River Basin, Hydrology and Earth System Sciences, 16, 489-500, 2012.

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 semi-distributed water balance model for the Ethiopian highlands, Hydrological processes, 23, 3718-3727, 2009.

Collischonn, B., Collischonn, W., and Tucci, C. E. M.: Daily hydrological modeling in the Amazon basin using TRMM rainfall estimates, Journal of Hydrology, 360, 207-216, 2008.
Fuka, D. R., Walter, M. T., MacAlister, C., Degaetano, A. T., Steenhuis, T. S., and Easton, Z. M.: Using the Climate Forecast System Reanalysis as weather input data for watershed models, Hydrological Processes, 2013.
Madsen, H.: Automatic calibration of a conceptual rainfall–runoff model using multiple objectives, Journal of hydrology, 235, 276-288, 2000.
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 & Assessment, 9, 77-89, 2004.
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 sediment for the Abay (Blue Nile) with a simple model, Hydrological processes, 23, 3728-3737, 2009.
Steenhuis, T. S., Tilahun, S. A., Elkamil, M., Atanaw, F., Worqlul, A., Ayana, E. K., Abiy, A. Z., Mulukken, M., Tesemma, Z. K., and Mohamad, Y. A.: Simulating discharge and sediment concentrations in the increasingly degrading Blue Nile Basin, Proceedings conference on science and technology towards the Development of East Africa, 2013, 291-299,
Su, F., Hong, Y., and Lettenmaier, D. P.: Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin, Journal of Hydrometeorology, 9, 622-640, 2008.
Tilahun, S., Guzman, C., Zegeye, A., Engda, T., Collick, A., Rimmer, A., and Steenhuis, T.: An efficient semi-distributed hillslope erosion model for the subhumid Ethiopian Highlands, Hydrology and Earth System Sciences, 17, 1051-1063, 2013a.
Tilahun, S., Mukundan, R., Demisse, B., Engda, T., Guzman, C., Tarakegn, B., Easton, Z., Collick, A., Zegeye, A., and Schneiderman, E.: ASATURATION EXCESS EROSION MODEL, 2013b.
Wale, A., Rientjes, T., Gieske, A., and Getachew, H.: Ungauged catchment contributions to Lake Tana's water balance, Hydrological processes, 23, 3682-3693, 2009.
Wang, J., and Wolff, D. B.: Evaluation of TRMM ground-validation radar-rain errors using rain gauge measurements, Journal of Applied Meteorology and Climatology, 49, 310-324, 2010.