We would like to thank the editor and all referees for their helpful comments. With these comments, we were able to improve the paper significantly. One of the major changes is that the main goal and key messages were more clearly focused. For this purpose, and as advised by the editor, the areal rainfall analysis was excluded throughout the paper. In addition, the method description were restructured and explained more detailed. These improvements have changed the manuscript significantly and therefor it is quite hard to indicate all individual changes. However, besides detailed answer to all comments of editor and referees (starting from page 2), we summarize the main changes we made below (linked to referees comments).

<u>Objective, key messages and title</u>

All reviewers and the editor indicated that the paper would benefit from clearer objectives and a more appropriate title. As a result, the main goal, key objectives and title were adjusted to:

- Title: Rainfall-discharge modelling using river stage time series in the absence of reliable discharge information: a case study in the semi-arid Mara River Basin.
- Main goal: The goal of this study is to illustrate the potential of water level time series for model calibration by incorporating the hydraulic equation describing the rating curve within the model.
- Key objectives: 1) present an important data set for the Mara River Basin, 2) illustrate a hydrological modelling methodology where the model is calibrated using river water levels instead of discharge.

As a result of reformulating the key objectives, the analyses on the areal rainfall estimates was excluded as suggested by the editor to give the paper a clearer focus. Changes in the main objective and key messages were applied throughout the entire paper, for example in the introduction by reformulating the goal (page 2, lines 57-62), methods (rainfall analysis removed), results and discussion (rainfall analysis removed), and conclusion by reformulating the goal and key messages (page 17, lines 352-362).

In response to:

- Editor: "*As all the referees highlight, and as you already acknowledge in your replies, the main objective should be made clear and the presentation should be structured along such line.*"
- Referee #1: "*Section 1. The paper states that the "The goal is to develop a reliable hydrological model for the semi-arid and poorly gauged Mara". In my opinion, this is not the kind of objectives that warrants a publication. I am convinced that the authors can identify a set of more appealing objectives for their work.*"
- Referee #2: "*1. The main aims and goals of the paper are poorly stated. Developing a hydrological model for a particular region as stated as the main goal of the paper is not a 'Cutting edge case study'. Similarly, the key research contributions from the paper are not clearly highlighted within the conclusions. The authors need to think about the novel aspects of the paper and two-three key messages they want the reader to take away. 2. The title of the paper is currently misleading – I would remove 'with data uncertainty' as you do not consider uncertainties in stage data and the analysis of precipitation uncertainties is limited.*"

<u>Methods</u>

We agree with the editor and reviewers that the methods section needed to be explained more detailed and restructured. For example, the sensitivity analysis for the HAND and slope thresholds was added (Section 3.1, page 6, lines 136-141); the procedure on applying process and parameter constraints was described more detailed and formulas were added (Section 3.3, page 10, lines 199-202); the model calibration approach was explained more detailed using a new flow chart (Figure 7) and restructured significantly (Section 3.4, page 12, line 214); and a sub-section on the rating curve analysis was added (Section 3.5, page 14, lines 242-252). In the sub-section on the model calibration, additional formulas and parameter values were included, as also an explanation on why the model was calibrated using FDCs (Section 3.5, page 12, lines 216-219). The sub-section on the rainfall analysis was excluded.

This is in response to:

- Editor: "*The proposed approach for replacing the use of unreliable rating curves and how you use the Strickler equation is the core of the work and should definitely be explained much better. The description of the available data but especially of the calibration/validation procedure (very complex per se: three river sections, multiple objective functions, ..) also need to be thoroughly revised and integrated with a number of important clarifications also on the choices (necessarily subjective in some cases) that you made.*"
- Referee #1: "*Section 3.3. This section, which is key to explain what was done in the paper, is very convoluted, and impossible to understand. […] In general every paragraph contains a lot of information in a very convoluted way. It is necessary to describe the methodology in a much more streamlined way. […] The procedure for calibration using h and the procedure for evaluation using Q needs to be clearly distinguished.*"
- Referee K. Keshavarz: "*There is no information about the calibration process of either the FLEX-Topo model or the Manning-Strickler formula. What were the initial ranges of parameters? How many parameters have been calibrated? Did the authors used an optimization algorithm or an uncertainty-based method? What were the final ranges/values of parameters?*"

Response to the Editor

Thank you very much for your comments. These were taken into account to improve the paper.

Comments of the editor:
I do believe that it's very important providing more and more information that may be used in regions where it is difficult to obtain reliable and long times series of meteo-hydrological data for calibrating rainfall-runoff models is crucial and proposing procedures tailored to improve the model implementation in such regions is definitely an important topic for our journal. As all the referees highlight, and as you already acknowledge in your replies, the main objective should be made clear and the presentation should be structured along such line.
I would suggest to remove the analysis on the areal rainfall estimates: I believe that focussing on the first one of the key-messages you list in your reply to Ref#1 and Ref#2 (and that automatically implies addressing also the second one) is already an ambitious objective and more than enough for the paper.
In fact I am afraid that adding also information and a cursory analysis on the important issue of the rainfall spatial field (that alone would need a separate detailed analysis, probably focussing in detail also on the features of the typical rainfall events that are expected in this specific part of the world if you want to add something to the vast literature on the subject) would not help to improve the clarity of the work, and you already need to add a number of details and clarifications, especially on the proposed procedure and on the model implementation, as required by the referees.
The proposed approach for replacing the use of unreliable rating curves and how you use the Strickler equation is the core of the work and should definitely be explained much better. The description of the available data but especially of the calibration/validation procedure (very complex per se: three river sections, multiple objective functions, ..) also need to be thoroughly revised and integrated with a number of important clarifications also on the choices (necessarily subjective in some cases) that you made.

Response to the comments:
We agree the paper would benefit from a clearer objectives and key messages. We also agree that removing the areal rainfall analysis would result in a clearer focus. Therefore, the main goal is reformulated as follows: *The goal of this study is to illustrate the potential of water level time series for model calibration by incorporating a hydraulic equation describing the rating curve within the model*. The key messages are 1) present an important data set for the Mara River Basin, and 2) illustrate a hydrological modelling methodology where the model is calibrated using river water levels instead of discharge. This was applied throughout the entire paper; hence, the rainfall analysis was removed accordingly. The importance of the rainfall was only mentioned briefly in the limitations and recommendations. In addition, the calibration procedure was explained more detailed and restructured taking into account all comments.

Response to Anonymous Referee #1

Thank you very much for your review. Your detailed comments will be taken into consideration to improve the paper.

Regarding the major comments:
Section 1. The paper states that the "The goal is to develop a reliable hydrological model for the semi-arid and poorly gauged Mara". In my opinion, this is not the kind of objectives that warrants a publication. I am convinced that the authors can identify a set of more appealing objectives for their work.
All three reviewers pointed out that the paper could benefit from clearer objectives and subsequently a more appropriate title. We agree with the reviewers that the paper needs improvement here. We have submitted our study as a "cutting edge case study". According to HESS "**Cutting-edge case studies** report on case studies that require (a) broadening the knowledge base in hydrology as well as (b) sharing the underlying data and models. These case studies should be cutting edge with respect to the quality and diversity of data provided the soundness of the models employed, and the importance of the study objective."

We present both 1) an important and high quality data set for the data-poor Mara River Basin after detailed analysis of the available rainfall, river stage and discharge measurements and 2) an innovation in rainfall-runoff modelling using river water level time series for model calibration in absence of reliable discharge data which is often encountered in African river basins. In our opinion the latter contributes to the knowledge base of hydrology, in particular rainfall-runoff modelling. In addition, we analysed the influence of rainfall data averaging in semi-arid basins where the rainfall typically has a high spatial and temporal variability.

The main goal was not to merely develop a hydrological model, but to develop a modelling methodology which can help increasing the hydrological understanding in this poorly gauged semi-arid region using water level time series for calibration instead of discharge since the rating curve was of very poor quality. Hence, the challenge was to assess the water availability despite the poor data quality. In the Mara River Basin, there is limited data available, let alone a complete assessment of the data availability and quality. In addition, there are only limited hydrological models of this basin, therefore the understanding of the local hydrological processes is quite limited. Moreover, the absence of good quality discharge time series is not unique to this area, therefore assessing the possibility of calibrating on water levels instead of discharge is very useful for poorly gauged areas and should be explored more detailed in future studies. The advantage of water level time series is the higher availability as it is easier to measure and higher reliability since there is no calculation step in between (using a rating curve). In the future this could be combined with remotely sensed altimetry data.

In short, our key objectives are: 1) present an important data set for the Mara River Basin, 2) illustrate a hydrological modelling methodology where the model is calibrated using river water levels instead of discharge and 3) show the difference between input averaging of the rainfall as typically done and output averaging of the modelled discharge. The latter allows the inclusion of the non-linear behaviour of the rainfall-discharge relation in river basins.

Therefore the key messages for the reader to take away are:

1. In poorly gauged river basins, calibration on water level time series is more reliable than on discharge time series since additional uncertainties arise from fitting rating curves on scarce discharge measurements.
2. In this methodology, the water level-discharge relation is implicitly included in the model; the power exponent of this relation is related to the geometrical data which is observable in the field.
3. The method for dealing with highly spatially distributed rainfall in hydrological modelling is significant to obtain reliable results.

To take this comment into account and highlight these key objectives more clearly, this division into these main topics will be applied throughout the article. In combination with a clearer title, we hope the key messages will be clearer. We suggest changing the title into: **Rainfall-discharge modelling using river stage time series in the absence of reliable discharge information: a case study in the semi-arid Mara River Basin.**

Section 1. Can the authors clarify why using water levels for model calibration avoids the effect of discharge uncertainties? This is presented as a fact, with no references to previous literature, and no explanations. I do not find the explanation obvious. Do they imply that rating curves are constant, and that the whole procedure of updating rating curves, as commonly done, is flawed and useless?
Thank you for this comment, this indeed should be explained more explicitly.

It is important to make a distinction between well and poorly gauged river basin. In well gauged basins, sufficient discharge measurements can be available for fitting a rating curve more reliably and updating it regularly. In that case, discharge time series are indeed reliable and useful for model calibration. However, in poorly gauged areas, discharge measurements are generally very scarce. As a result, rating curves are fitted to scarce data and not updated regularly resulting in high uncertainties especially when extrapolating. As a result, there are significant uncertainties in discharge time series. Water level time series however are direct measurements which are therefore more reliable.
For this specific case of the Mara River, data analysis indicated that there are indeed high uncertainties in the discharge data (section 2). Therefore, here water level time series were more reliable than the discharge.
In short, using water levels for model calibration instead of discharge is only an improvement if the rating curve is indeed of poor quality, as often the case in poorly gauged areas.

132: to further delimit HRUs. Which HRUs? Even reading the paragraph further, it is unclear how many HRUs are used. You say 4, but then mention "are mainly cropland and forest, whereas further south the land use is dominated by grassland", which are not in the 4 HRUs.
The HRUs were defined in Line 134: "This resulted in four HRUs in the sub-basin of the Mara River Basin: forested hill slopes, shrubs on hill slopes, agriculture and grassland".

Section 3.3. This section, which is key to explain what was done in the paper, is very convoluted, and impossible to understand. The first sentence states "Parameters and process constraints have been applied to eliminate unrealistic model results". Which model results? "For example, the maximum storage" – why for example? I want to know exactly what was done and how it was done. Instead, there are just a few sentences of how the methodology was carried out, relegating the essential details to even more unclear supplementary materials. "The model was calibrated and evaluated" how was this done? What is the difference between calibration and evaluation? "For the evaluation of this calibration", why this? Is there another calibration? In general every paragraph contains a lot of information in a very convoluted way. It is necessary to describe the methodology in a much more streamlined way.
As the reviewer pointed out, the section on model methodology is indeed quite concise and could benefit from more elaboration. Therefore more details will be added in this section and in the supplements. A table of all the constraints was already included in the supplement (Table S1 and S2).
The formulation of "unrealistic model results" is indeed confusing, "unrealistic parameter sets" is more accurate as constraints were applied to eliminate unrealistic parameter sets rather than unrealistic model results; for instance forest interception should be greater than cropland interception. Furthermore, the model evaluation step consisted of several elements: first the model was evaluated by means of validation (which is what is meant in this section), later on by analysing the discharge on sub-catchment level, analysing the rating curves and the influence of the rainfall (in the discussion).

205. Needs to be expanded and clarified.
1) The procedure for calibration using h and the procedure for evaluation using Q needs to be clearly distinguished.
2) You write that you use d for model calibration and flow duration curves for model evaluation. Flow means Q, but all the plots show d duration curves. Where is the flow used?
3) The value d_{mod} is not present in the Strickler formula (there is A and R). What is the relation to d? It should be written explicitly.
4) What is the relation between the Strickler formula and Q = a â´L°U (h ⍰ h0)b?
5) What is the value of b?
6) If I understand well, the observed and modelled water discharge are obtained using the same formula with the same parameters. Why is then Qrec needed? Trying to explain 3 essential things (model calibration, evaluation and evaluation of rating curves) in the same paragraph does not work.
Reply to 1) There are indeed multiple steps in the use of water level and discharge for calibration and validation that could confuse the reader and should therefore be explained more clearly. First, the model was calibrated on water level (line 202), then the modelled discharge ($Q_{Strickler}$ and Q_{mod}) were compared to the recorded discharge (line 211) for model evaluation.
Reply to 2) The reviewer is right, this will be corrected. Instead of flow duration curve, the duration curves of the water depths were used for calibration. The flow was not used for the calibration.
Reply to 3) Thank you for this comment; this indeed is not written explicitly and should be included:
The cross-sections were simplified as a trapezium with a river width B and two different river bank slopes i_1 and i_2; these coefficients (Table 1) were estimated based on available cross-section data (Supplement S2). In addition, the water depth *d* was calculated from the water level *h* and reference level h_0.

$$A = B * d + \frac{1}{2} * d * (i_1 + i_2) * d$$
$$R = \frac{A}{B + d * \left((1 + i_1^2)^{\frac{1}{2}} + (1 + i_2^2)^{\frac{1}{2}}\right)}$$
$$d = h - h_0$$

Table 1: Coefficients used for the simplification of the river cross-section

	River width B [m]	**River bank slope i_1 [-]**	**River bank slope i_2 [-]**	**Reference level h_0 [m]**
Amala	10.0	3.50	1.83	0
Nyangores	19.05	2.65	5.56	0
Mines	43.81	3.53	3.66	10

Reply to 4) Both equations estimate the discharge using water level data. In the rating curve (Q = a * (h-h0)^b), parameter *a* includes information on the cross-section, roughness and slope; parameter *b* information on the cross-section. This information is more direct in the Strickler formula.
Reply to 5) The value *b* varies for each cross-section. In line 206/7, this information could be included such as: Note that by using the Strickler formula the exponent of the rating curve is fixed; $Q = a * (h - h_0)^b$, with b = 1.71 at Amala, b = 1.71 at Nyangores and b = 1.70 at Mines using the same water level time series as for the calibration and validation.
Reply to 6) Thank you for this comment, this indeed needs to be explained more clearly. In contrast to what the reviewer stated, the modelled and observed discharge were obtained using different formulas. The modelled discharge Q_{mod} was obtained through the FLEX-Topo model. Q_{Str} was calculated using the Strcikler formula, a calibrated roughness/slope parameter *c* and water level time series. Q_{rec} was obtained from the water department and was calculated locally probably by using a rating curve and the water level time series.
This discharge Q_{Str} was compared to Q_{rec} to compare the recorded and modelled rating curves with each other.

215. Does the model provide simultaneously the output at the 3 stations? Was it calibrated simultaneously to the 3 gauging stations? Or was it calibrated individually to each station? If it was calibrated individually to each station, shouldn't the parameters of the same HRU in different catchment be the same? How was this ensured?
The reviewer has a good point here. The model was calibrated for all three stations individually using the same parameter ranges and constraints. As a result, the parameters were similar, yet slightly different for each station. After calibration, the "best" parameter sets were used for cross-validation. The model performed well when validating at Nyangores using the parameter set based on Mines ($NS_{FDC, log}$ = 0.94 and NS_{FDC} = 0.83) whereas vice versa resulted in poor performance ($NS_{FDC, log}$ = 0.29 and NS_{FDC} = 0.00). This is not surprising as all HRUs were represented when calibrating at Mines and only two HRUs when calibrating at Nyangores, namely forest and agriculture.

230. It appears that the model was calibrated using FDCs. But the objective is to simulate streamflow. Are FDCs sufficient to represent streamflow time series? E.g. I can imagine that information about seasonality as well as timing of peaks is lost when calibrating to FDCs. How were these problems addressed?
This is a good question. By calibrating on FDCs, the focus is on the flow statistics (e.g. how often high flows occur). This information is also in the streamflow, only the exact timings are not included when calibrating on FDCs. However, in this case, the timings were off anyway due to the limited number of rainfall stations available which was insufficient to capture the spatial heterogeneity well. Therefore, in this case the FDCs were good for model calibration.

230. Was it multi objective calibration leading to a Pareto-front? Needs to be clarified.
Thank you for this comment. Instead of analysing a Pareto-front, the values for the objective functions were ordered and the ones with the highest values were considered as "good" parameter sets.

Response to Anonymous Referee #2

Thank you very much for your review. Your detailed comments will be taken into consideration to improve the paper.

Regarding the major comments:

1. The main aims and goals of the paper are poorly stated. Developing a hydrological model for a particular region as stated as the main goal of the paper is not a 'Cutting edge case study'. Similarly, the key research contributions from the paper are not clearly highlighted within the conclusions. The authors need to think about the novel aspects of the paper and two-three key messages they want the reader to take away.

2. The title of the paper is currently misleading – I would remove 'with data uncertainty' as you do not consider uncertainties in stage data and the analysis of precipitation uncertainties is limited.

Reply to 1-2): All three reviewers pointed out that the paper could benefit from clearer objectives and subsequently a more appropriate title. We agree with the reviewers that the paper needs improvement here. We have submitted our study as a "cutting edge case study". According to HESS "**Cutting-edge case studies** report on case studies that require (a) broadening the knowledge base in hydrology as well as (b) sharing the underlying data and models. These case studies should be cutting edge with respect to the quality and diversity of data provided the soundness of the models employed, and the importance of the study objective."

We present both 1) an important and high quality data set for the data-poor Mara River Basin after detailed analysis of the available rainfall, river stage and discharge measurements and 2) an innovation in rainfall-runoff modelling using river water level time series for model calibration in absence of reliable discharge data which is often encountered in African river basins. In our opinion the latter contributes to the knowledge base of hydrology, in particular rainfall-runoff modelling. In addition, we analysed the influence of rainfall data averaging in semi-arid basins where the rainfall typically has a high spatial and temporal variability.

The main goal was not to merely develop a hydrological model, but to develop a modelling methodology which can help increasing the hydrological understanding in this poorly gauged semi-arid region using water level time series for calibration instead of discharge since the rating curve was of very poor quality. Hence, the challenge was to assess the water availability despite the poor data quality. In the Mara River Basin, there is limited data available, let alone a complete assessment of the data availability and quality. In addition, there are only limited hydrological models of this basin, therefore the understanding of the local hydrological processes is quite limited. Moreover, the absence of good quality discharge time series is not unique to this area, therefore assessing the possibility of calibrating on water levels instead of discharge is very useful for poorly gauged areas and should be explored more detailed in future studies. The advantage of water level time series is the higher availability as it is easier to measure and higher reliability since there is no calculation step in between (using a rating curve). In the future this could be combined with remotely sensed altimetry data.

In short, our key objectives are: 1) present an important data set for the Mara River Basin, 2) illustrate a hydrological modelling methodology where the model is calibrated using river water levels instead of discharge and 3) show the difference between input averaging of the rainfall as typically done and output averaging of the modelled discharge. The latter allows the inclusion of the non-linear behaviour of the rainfall-discharge relation in river basins.

Therefore the key messages for the reader to take away are:

1. In poorly gauged river basins, calibration on water level time series is more reliable than on discharge time series since additional uncertainties arise from fitting rating curves on scarce discharge measurements.
2. In this methodology, the water level-discharge relation is implicitly included in the model; the power exponent of this relation is related to the geometrical data which is observable in the field.
3. The method for dealing with highly spatially distributed rainfall in hydrological modelling is significant to obtain reliable results.

To take this comment into account and highlight these key objectives more clearly, this division into these main topics will be applied throughout the article. In combination with a clearer title, we hope the key messages will be clearer. We suggest changing the title into: **Rainfall-discharge modelling using river stage time series in the absence of reliable discharge information: a case study in the semi-arid Mara River Basin**

3. There needs to be a broader introduction to data uncertainty in the introduction including rainfall uncertainty as this is considered later in the paper. Furthermore, there should also be a larger section devoted to model

calibration and model diagnostics and particularly how to perform robust model evaluation in the face of data uncertainties.
Thank you for this comment. One of the objectives is indeed on the uncertainty caused by the rainfall heterogeneity, more specific: the difference between averaging of the input precipitation in contrast to averaging the output modelled discharge. Therefore, the introduction should indeed also include rainfall variability. However, in this study, uncertainties in the data for the Mara River Basin were pointed out rather than performing a complete uncertainty analysis to assess the influence of data uncertainty on the modelling results. Therefore, we feel that a section on model uncertainty analysis in the introduction is outside the scope of this article.

4. A separate section on data would be useful. At the moment, different datasets are introduced at lots of different points throughout section 2 and section 3.
The reviewer makes a good point here. All data should be introduced in section 2. Those newly mentioned in section 3 (DEM, land cover map, NDVI and remotely sensed precipitation) should have been introduced in section 2 as well. This will be done by subdividing section 2 into multiple sub-sections: Section 2.1 Site description (lines 77-85), Section 2.2 Ground measurements (lines 87-112) and Section 2.3 Remotely sensed data. The latter will be added to introduce remote sensing data that are now newly mentioned in section 3:

Section 2.3 Remotely sensed data
Besides ground measurements, also remotely sensed data were used for the model development. The catchment classification was based on the topography and the land cover. For the topography, a digital elevation map (SRTM) with a resolution of 90 m and vertical accuracy of 16 m was used (U.S. Geological Survey, 2014). The land cover was based on Africover, a land cover database based on ground truth and satellite images (FAO, 1998). Moreover, NDVI maps were used to define parameter constraints.

New information mentioned in section 2.3 will then be excluded from section 3 to avoid repetition.

5. One of the reasons for calibrating the model to water level is to 'avoid' uncertainties in water discharge. However, by then calibrating the 'c' parameter for the Strickler formula surely you just replace one source of uncertainty with another. As stated in the paper, it is likely that this parameter is also compensating for large sources of uncertainty in your precipitation data so I wonder how robust the results are given all these different sources of uncertainty. This needs to be better discussed in the limitations.
This indeed is a limitation of this methodology. However, in contrast to the discharge uncertainties, this is a parameter uncertainty that could be quantified more accurately. This is a recommendation for future studies. Therefore, a new section will be added in the discussion to highlight more clearly the short comings of this methodology (e.g. compensation of the slope-roughness parameter *c* for non-closure effects) and recommendations for future studies (e.g. quantification of uncertainties in the parameter *c*, methodologies to constrain or estimate parameter *c*, analysis of the potential of water level based model calibration in well gauged basins to assess the uncertainties more reliably, determination of suitable objective functions for calibrating on water levels instead of flow etc.).

6. Section 3.3 is really difficult to follow and certain model choices need to be better justified –
a. Why was NSE chosen for model evaluation? How appropriate is NSE for calibrating water levels?
Thank you for this comment. In this case, NSE was chosen for model calibration and validation. However, it was not analysed how appropriate this is for calibrating on water levels. Therefore this is a good recommendation for future studies!
b. Strickler formula on line 205 needs to be presented as a separate equation – what do 'k' and 'i' denote?
Thank you for this comment. This will be clarified as such (line 205): […], where R is the hydraulic radius, A the cross-sectional area, k the roughness and i the slope; […]

7. Results
a. Section 4.1. The authors state at a couple of points that 'the observed and modelled water depth were quite similar to each other'. How similar is similar!? It would be better here to state NSE values as a quantitative measure of how similar they are.
A quantitative measure is indeed useful here. This was done in Table 6.

b. Section 4.2. How many point discharge measurements were taken? While these can be useful in model calibration and evaluation – I don't think comparing a single point measurement to a whole month of modeled discharge was useful and the fact that the modeled results were 'within an order of magnitude of the point *measurement' not a particularly persuasive argument that the model was performing well. I think these could be incorporated much better into the model evaluation framework.*

In total, five point measurements were taken at three locations (see section 2). As there are only a few measurements and a significant time difference between the measurements (2014) and the model (1970s/1980s), it is not possible to use these measurements for model evaluation other than comparing the order of magnitude. If there would have been more measurements, then more accurate comparison methodologies would have been possible.

8. I was surprised that given the amount of effort that went into defining HRUs and different model structures for the basin based on field observations and interviews, no results or analysis was presented on these different model structures. Was it just data uncertainty that lead to poor model performance or also the definition of model processes? How were model simulations improved by using two different model structures rather than one?

Thank you for this comment. Analyses on the effect of using different model structures (lumped vs. semi-distributed) were done in an early stage yet the results were indeed not included in the paper. The during the model development, a lumped model structure was compared with a semi-distributed model using two different model structures. This comparison showed applying multiple model structures significantly improved the model, especially during validation (Table 2).

Table 2: Model comparison: semi-distributed vs lumped for calibration (1988-1991) and validation (1985-1987)

	Calibration		**Validation**	
	$NS_{FDC,log}$	**NS_{FDC}**	**$NS_{FDC,log}$**	**NS_{FDC}**
Semi-distributed	0.91	0.71	0.74	0.93
Lumped (SSF model structure)	0.87	0.42	0.00	0.41
Lumped (HOF model structure)	0.90	0.62	0.09	0.12

Literature

FAO: Africover, GLCN, 1998.

Digital Elevation Map: www.earthexplorer.usgs.gov, 2014.

Response to K. Keshavarz

Thank you very much for your review. Your detailed comments will be taken into consideration to improve the paper.

Regarding the major comments:
"As far as I have understood, the authors have used water level observations (dobs) to calibrate the model. So, Manning-Strickler formula has been implemented in the model (line 60) to simulate both discharge, Qmod, and water levels, dmod. Moreover, the authors have produced discharges based on dobs using Manning-Strickler formula and named it QStrickler. Then they compared the recorded observed discharge, Qrec, with QStrickler and Qmod in Figure 12. How did the authors produce QStrickler? Have you had information about the cross-section details at three locations indicated in Figure 12? The research method explanation is hard to follow and understand."
Thank you for this comment. The methodology indeed is explained quite concisely and could benefit form more elaboration. Nevertheless, the reviewer understood the methodology and answered the question correctly: the discharge $Q_{Strickler}$ was indeed calculated by using the Manning-Strickler equation, cross-section data and a calibrated parameter for the roughness and slope. This was explained briefly in lines 204 – 210.

"There is no information about the calibration process of either the FLEX-Topo model or the Manning-Strickler formula. What were the initial ranges of parameters? How many parameters have been calibrated? Did the authors used an optimization algorithm or an uncertainty-based method? What were the final ranges/values of parameters? Have you tried any other objective function rather than Nash-Sutcliffe? Why have the authors used two validation periods for Mines (lines 221-222)?"
This is a good point as this was indeed not mentioned in the paper and should be included. For the calibration, the MOSCEM-UA algorithm was applied (Vrugt et al., 2003). No other objective functions have been tested, however several signatures were tested such as the hydrograph, logarithm of the hydrograph and slope of the flow duration curve. As no major differences were found in this case, this was not tested more detailed. Two validation periods were used for Mines to use as much data as possible taking into account the limited data availability.
To address this issue more detailed in the paper, the Table 3 and Table 4 will be added in the supplement and the sentence in line 201 will be adjusted to:
After having set up the model and defined the constraints, the model was calibrated applying the MOSCEM-UA algorithm (Vrugt et al., 2003) and validated.

Table 3: Parameter ranges and optimal parameter sets

Parameters	Parameter ranges	Unit	Optimal parameter set		
			Nyangores	Amala	Mines
$I_{max, F}$	0.2 - 2.7	mm	1.26	0.60	2.34
$I_{max, A}$	0.6 - 6.0	mm	1.10	0.60	1.51
$I_{max, G}$	0.7 - 3.6	mm	0.78	0.64	1.56
$I_{max, S}$	0.3 - 2.0	mm	1.24	0.62	1.71
β	0.5 - 2.0	-	1.88	0.62	1.32
T_{lag}	0.5 - 1.5	D	1.45	1.44	1.46
$K_{f,H}$	1 - 28	d	27.24	3.01	6.01
$K_{f,T}$	1 - 28	d	12.02	2.01	3.05
F	0 - 15	mm/d	0.42	12.77	1.71
c	Mines : 0 - 2.6 Nyangores: 0.4 - 1.6 Amala : 3.2 - 4.1	$m^{1/3}/s$	0.89	3.40	1.31
$S_{s,max}$	50 - 150	mm	99.87	106.04	141.98
$S_{F/S}$	0 - 0.5	-	0.27	0.30	0.22
$S_{A/G}$	0 - 0.5	-	0.09	0.33	0.24

Table 4: Fixed parameters

Parameters	Parameter value	Unit
K_s	28	d
C_e	0.5	-
S_{umaxF}	122	mm
S_{umaxA}	94	mm
S_{umaxG}	83	mm
S_{umaxS}	89	mm
$S_{max,Amala}$	46	mm
$S_{max,Nyangores}$	74	mm
$S_{max,Middle}$	122	mm
$S_{max,Lemek}$	119	mm
$S_{max,Talek}$	69	mm
$S_{max,Sand}$	29	mm
$S_{max,Lower}$	48	mm

"The time-step of the model seems to be neglected. The information about the timestep is not discussed in the paper expect a minor reference under Table 4 caption. Have you tried different time-steps? Could results improve if you use a smaller time-step?"
The model was run on a daily time scale. A smaller time-step was not possible due to data limitations. As this was indeed not mentioned clearly in the paper, the sentence in line 201 will be adjusted to:
After having set up the model and defined the constraints, the model was calibrated on a daily time scale applying the MOSCEM-UA algorithm (Vrugt et al., 2003).

"One of the main purposes of hydrological models is producing the hydrographs at different locations. Although authors have tried to indicate the water level time series (Figures 7, 8, 9 and 14), the hydrographs are missing."
Discharge time series were indeed not shown as the focus was on simulating the water depth instead of the discharge. For comparison sake, this will be included in the supplement.

The details of sensitivity analysis to produce thresholds of different landscape slopes and HAND values are missing. Is the HAND model based on the research of Nobre et al. (2011)? Have you used any specific sensitivity analysis algorithm/approach?
Thank you for this comment. The thresholds influence the area contribution of the different landscapes, for instance a higher slope threshold could result in less hillslope areas. In the sensitivity analysis, this influence of the thresholds on the change of the area contribution was analysed. It was found that these area contributions behave asymptotically to changes in the thresholds. Therefore thresholds were chosen where changes in area contributions become insignificant. This asymptotical behaviour was strongly visible for the slope threshold. As there were no wetlands (based on field observations), the HAND threshold was set to zero; this will be corrected in the paper.

Are calibrated roughness values in accordance with the streambed material for Manning-Strickler formula?
Yes, they are. Natural channels with short grass typically have a Strickler coefficient between 25 and 45 $m^{1/3}/s$. The calibrated Strickler parameter was within this range assuming a slope between 10^{-2} and 10^{-4} which is realistic as it is a flat area with multiple rapids.

How did the authors specify the average flow velocity (line 165)? Would changing this parameter value impact the overall results? Does it change the hypothesis of using Manning-Strickler formula?
Thank you for this comment. The average flow velocity was an assumption which agreed with the point measurements in the river. With this velocity, the maximum delay from the sub-catchment furthest away was 4 days. Changing this velocity would change the timing of the flow from a specific sub-catchment. However, this timing uncertainty was insignificant compared to timing uncertainties caused by the highly heterogeneous rainfall which was poorly represented with the available stations.

Regarding the minor comments:
Thank you for those comments, they will be taken into consideration. These comments included: correcting English language, adding a separate section introducing the different data sources used, checking the literature

referencing to avoid missing or faulty references, renaming “Strickler formula” to “Strickler-Manning formula” to make it more general, referring to specific tables and figures in the supplement (e.g. see Table S1) instead of the supplement in general (“see supplement”) and adjusting some figures. For the figures, sub-figures can be indicated more clearly through numbering/letters, months written out instead of numbers, sub-catchment boundaries included in the legend, figure adjusted such that the number of stations are consistent with the text.

In addition, to respond more detailed to some of the minor comments:
The title of research seems awkward. What does ‘modeling [: : :] with data uncertainty’ mean? Where did the uncertainty of streamflow, either water level or discharge, come into consideration?
Also other reviewers have stated that the title needs to be improved. The following title is suggested: Rainfall-discharge modelling using river stage time series in the absence of reliable discharge information: a case study in the semi-arid Mara River Basin.

Equation 1 indicating Nash-Sutcliffe formula is wrong (lines 230 to 233).
Unfortunately, it is not clear to the authors what is wrong with this equation. The caption however will be removed.

Equation 1: Formulas for the Nash-Sutcliff objective function. The indices mod and obs indicate modelled and observed values, respectively. In all cases, sorted data was used for the calculation of the objective function therefore the flow duration curve was calibrated.

$$\mathrm{NS}_{\log(d)} = 1 - \frac{\Sigma\left(\log\left(d_{\mathrm{mod,sorted}}\right) - \log\left(d_{\mathrm{obs,sorted}}\right)\right)}{\Sigma\left(\log\left(d_{\mathrm{obs,sorted}}\right) - \log\left(d_{\mathrm{obs,avg}}\right)\right)} \qquad \mathrm{NS}_{d} = 1 - \frac{\Sigma\left(d_{\mathrm{mod,sorted}} - d_{\mathrm{obs,sorted}}\right)}{\Sigma\left(d_{\mathrm{obs,sorted}} - d_{\mathrm{obs,avg}}\right)}$$

What is the time period of discharge data indicated in Figure 12?
In this figure, the model calibration results were plotted (see caption), therefore the time periods used were the ones used for calibration (lines 219-229).

Section 4.4 needs more discussion as no general suggestion to future research is made. Moreover, it is not apparent whether these strategies have improved the results of calibration.
Thank you for this comment. Indeed, a section will be added to the discussion to include details on limitations of this methodology (e.g. compensation of the slope-roughness parameter *c* for non-closure effects) and recommendations for future studies (e.g. quantification of uncertainties in the parameter *c*, methodologies to constrain or estimate parameter *c*, analysis of the potential of water level based model calibration in well gauged basins to assess the quality and uncertainties more reliably, determination of suitable objective functions for calibrating on water levels instead of flow etc.).

The conclusions need to be considered again as many ideas have been repeated from the introduction/abstract part. It could have been more concise and explicit.
Thank you for this comment. The conclusion will be reformulated such that it is more concise and that key messages are stated more clearly.

Literature
Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., and Sorooshian, S.: Effective and efficient algorithm for multiobjective optimization of hydrologic models, Water Resources Research, 39, n/a-n/a, 10.1029/2002WR001746, 2003.

Marked-up manuscript

~~Modelling the Mara River Basin with data uncertainty using water levels for calibration~~

Rainfall-runoff modelling using river stage time series in the absence of reliable discharge information: a case study in the semi-arid Mara River Basin

Petra Hulsman[1], Thom A. Bogaard[1], Hubert H.G. Savenije[1]

[1]Water Resources Section, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN Delft, the Netherlands

Correspondence to: Petra Hulsman (p.hulsman@tudelft.nl)

Abstract. Hydrological models play an important role in Water Resources Management. ~~In hydrological modelling,~~ These models generally rely on discharge data ~~is generally required~~ for calibration. ~~To obtain continuous~~Discharge time series~~,~~, are normally derived from observed water levels ~~are usually converted into discharge~~ by ~~using~~use of a rating curve. However~~ with~~, this ~~methodology,~~ method suffers from many uncertainties ~~are introduced in the discharge data~~ due to insufficient observations, inadequate rating curve fitting procedures, rating curve extrapolation~~ or~~, and temporal changes in the river geometry. Unfortunately, this problem is ~~often the case~~prominent in many African river basins. In this study, an alternative calibration method is presented using water level time series instead of discharge, applied to a semi-distributed rainfall runoff model ~~has been applied to~~for the semi-arid and poorly gauged Mara River Basin ~~for the assessment of the water availability. To reduce the effect of discharge uncertainties in this model, water levels instead of discharge time series were used for calibration. In this model, seven sub-catchments are distinguished and four hydrological response units: forest, shrubs, cropland and grassland. To calibrate the model on water level data,~~in Kenya. The modelled discharges ~~have been~~were converted into water levels using ~~cross-section observations and~~ the Strickler-Manning formula. ~~In addition, new~~ This method produces an additional model output: a “geometric rating ~~curves have been obtained based on~~ curve equation” which relates the modelled discharge~~,~~ to the observed water level ~~and~~using the Strickler-Manning formula and a calibrated slope-roughness parameter. This procedure resulted in good and consistent model results during calibration and validation. The hydrological model was able to reproduce the water ~~depths~~levels for the entire basin as well as for the Nyangores sub-catchment in the north. The newly derived geometric ~~and recorded (i.e. existing)~~ rating curves were ~~significantly different at Mines,~~ subsequently compared to the existing rating curves. At the catchment outlet~~, probably~~ of the Mara, these differed significantly, most likely due to uncertainties in the recorded discharge time series. ~~At Nyangores however~~However, at the ‘Nyangores’ sub-catchment, the geometric and recorded discharge were almost identical. ~~In addition, it has been found that the precipitation estimation methodology influenced the model results significantly. Application of a single station for each sub-catchment resulted in flashier responses whereas Thiessen averaged precipitation resulted in more dampened responses.~~ In conclusion, ~~by using water level time series for calibrating the hydrological model of the Mara River Basin promising model results were~~ the results obtained~~. For this~~ for the Mara river basin~~, the main limitation for obtaining an accurate hydrograph~~

~~representation was~~ illustrate that with the ~~inadequate knowledge on~~proposed calibration method the ~~spatial distribution of~~water level time series can be simulated well, and that also the ~~precipitation.~~ discharge-water level relation can be derived, even in catchments with uncertain or lacking rating curve information.

1 Introduction to rating curve uncertainties

Hydrological models play an important role in Water Resources Management. In hydrological modelling, discharge time series are of crucial importance. For example, discharge is used when estimating flood peaks (Di Baldassarre et al., 2012; Kuczera, 1996), calibrating models (Domeneghetti et al., 2012; McMillan et al., 2010) or determining the model structure (Bulygina et al., 2011; McMillan et al., 2015). Discharge is commonly measured indirectly through interpolation of velocity measurements over the cross-section (Di Baldassarre et al., 2009; WMO, 2008). However, to obtain frequent or continues discharge data, this method is time consuming and cost-inefficient. ~~In~~Moreover, in African river catchments, the quantity and quality of the available discharge ~~measurement~~measurements is unfortunately often inadequate for reliable calibration of hydrological ~~modelling~~models (Hrachowitz et al., 2013; Shahin, 2002).

There are several sources of uncertainty in discharge data when using rating curves that cannot be neglected. First, measurement errors in the individual discharge measurements affect the estimated continuous discharge data, for example in the velocity-area method uncertainties in the cross-section and velocity can arise due to poor sampling (Pelletier, 1988; Sikorska et al., 2013). Second, these measurements are usually done during normal flows, however during floods the rating curve needs to be extrapolated. Therefore, the uncertainty increases for discharges under extreme conditions (Di Baldassarre et al., 2011; Domeneghetti et al., 2012). Thirdly, the fitting procedure does not always account well for irregularities in the profile, particularly when banks are overtopped. Finally, the river is a dynamic, non-stationary system which influences the rating curve: such as changes in the cross-section due to sedimentation or erosion, backwater effects or hysteresis (Petersen-Øverleir, 2006). The lack of incorporating such temporal changes in the rating curve increases the uncertainty in discharge data (Guerrero et al., 2012; Jalbert et al., 2011; Morlot et al., 2014). As a result, the rating curve should be regularly updated to take such changes into account. The timing of adjusting the rating curve relative to the changes in the river affects the number of rating curves and the uncertainty (Tomkins, 2014).

~~The goal of this study is to develop a reliable hydrological model for the semi-arid and poorly gauged Mara River Basin in Kenya. Previous studies have~~Previous studies focused on assessing the uncertainty of rating curves (Clarke, 1999; Di Baldassarre et al., 2009) and their effect on model predictions (Karamuz et al., 2016; Sellami et al., 2013; Thyer et al., 2011). ~~In this study however, the effects of discharge uncertainties are avoided by using water level instead of discharge~~ However, in the absence of reliable discharge data, water level time series provide reliable and valuable information on the flow dynamics (Seibert et al., 2016) and therefore could be a good alternative for hydrological model calibration. In general, water levels time series are more reliable than discharge data as these are direct measurements and not processed data. However, the potential of calibrating models on water level time series has not been studied in detail, especially in combination with a hydraulic equation, and in poorly gauged semi-arid areas.

The goal of this study is to illustrate the potential of water level time series for model calibration by incorporating ~~the~~a hydraulic equation describing the rating curve within the model. ~~The~~This calibration method is applied to the semi-arid and poorly gauged Mara River Basin in Kenya. For three gauging stations within this basin, the quality of the recorded rating curves have been analysed and compared to the model results~~ are verified using a few high quality discharge measurements. In previous studies, water level time series are found to provide valuable information on the flow dynamics for model calibration, especially in wet catchments whereas in dry catchments additional information is needed to constrain the flow volume (Jiang et al., 2017; Seibert et al., 2016)~~. For this purpose, a semi-distributed rainfall runoff model has been developed on a daily timescale applying the FLEX-Topo modelling concept (Savenije, 2010).

2 Site description of the Mara River Basin and data availability

The Mara River originates in Kenya in the Mau Escarpment and flows through the Masai Mara National Reserve in Kenya into Lake Victoria in Tanzania. The main tributaries are the Nyangores and Amala Rivers in the upper reach and the Lemek, Talak and Sand in the middle reach (Figure 1). The first two tributaries are perennial while the remaining tributaries are ephemeral, which generally dry out during dry periods. In total, the river is 395 km long (Dessu et al., 2014) and its catchment covers an area of about 11,500 km^2 (McClain et al., 2013) of which 65% is located in Kenya (Mati et al., 2008).

Within the Mara River Basin, there are two wet seasons linked to the annual oscillations of the ITCZ (Inter-tropical Convergence Zone). The first wet season is from March to May and the second from October to December (McClain et al., 2013). The precipitation varies spatially over the catchment following the local topography. The largest annual rainfall can be found in the upstream area of the catchment: between 1000 and 1750 mm/yr. In the middle and downstream areas, the annual rainfall is between 900 and 1000 mm/yr and between 300 and 850 mm/yr, respectively (Dessu et al., 2014).

The elevation of the river basin varies between 3000 m above sea level at the Mau Escarpment, 1480 m at the border to Tanzania and 1130 m at Lake Victoria (McClain et al., 2013). In the Mara River Basin, the main land cover types are agriculture, grass, shrubs and forests. The main forest in the catchment is the Mau Forest, which is located in the north. Croplands are mainly found in the north and in the south, whereas the middle part is dominated by grasslands.

2.1 Data availability

2.1.1 In situ monitoring data

In the Mara River Basin, long term daily water level and discharge time series are available for 44-60 years between 1955 and 2015 at the downstream station near Mines and in the two main tributaries: the Nyangores and Amala. In addition, precipitation and air temperature is measured at ~~29~~27 and ~~5~~7 stations, respectively (Figure 1 and ~~Table 1~~Table 5). However, the temporal coverage of these data is poor as there are many gaps.

~~Also, there~~There are many uncertainties in the discharge and precipitation data~~.~~ in the Mara River Basin. Discharge data analyses indicated that the time series were unreliable due to various inconsistencies in the data, ~~for example changing rating curves at~~ especially at Mines and Amala~~, unrealistic rating curve compared to cross-section based estimations at Nyangores and high~~ . At Mines, a high scatter in the discharge-water level graph ~~at Mines;~~was observed (Figure 2); also back-calculated cross-section average flow velocities were ~~much lower than~~ below 1 m/s (Figure S1) whereas in 2012 the measured velocity was 2.13 m/s and discharge 529.3 m^3/s (GLOWS-FIU, 2012). At Amala, the rating curves were adjusted multiple times affecting mostly the low flows. Only the rating curve at Nyangores was stable and consistent with field measurements ~~at Mines~~. The precipitation data analysis showed a high spatial variability between the ~~rainfall stations. This could be a result of high heterogeneity which is poorly represented by the~~ limited number of rainfall stations available. ~~See~~ More information can be found in the supplement ~~for more details.~~“S1 Data quality”.

~~As a result of using this precipitation data for hydrological modelling, significant errors and uncertainties will occur in the modelled discharge which are required for a solid water resources allocation plan. The uncertainties in the measured rating curve and precipitation need to be taken into account in the evaluation of the hydrological model performance. In contrast to previous studies where discharge time series were used to calibrate the hydrological model of the Mara River Basin using the Soil Water Assessment Tool (SWAT) (Dessu et al., 2012; Mwangi et al., 2016), in this study water level time series are used to avoid the uncertainties in the discharge data.~~

During field trips, ~~some~~ point discharge measurements were done in September/October 2014 at Emarti Bridge, Serena Pump House and New Mara Bridge, see ~~Table 2 and Figure 2.~~Table 6 and Figure 3. At each location, the discharge was derived ~~from cross-section and velocity measurements done with a RiverSurveyor, a small boat that was pulled across the river and on which was mounted~~ using an Acoustic Doppler Profiler~~,~~ (Sontek RiverSurveyor M9) mounted on a portable raft which is also equipped with a Power Communications Module and a DGPS antenna (Rey et al., 2015).

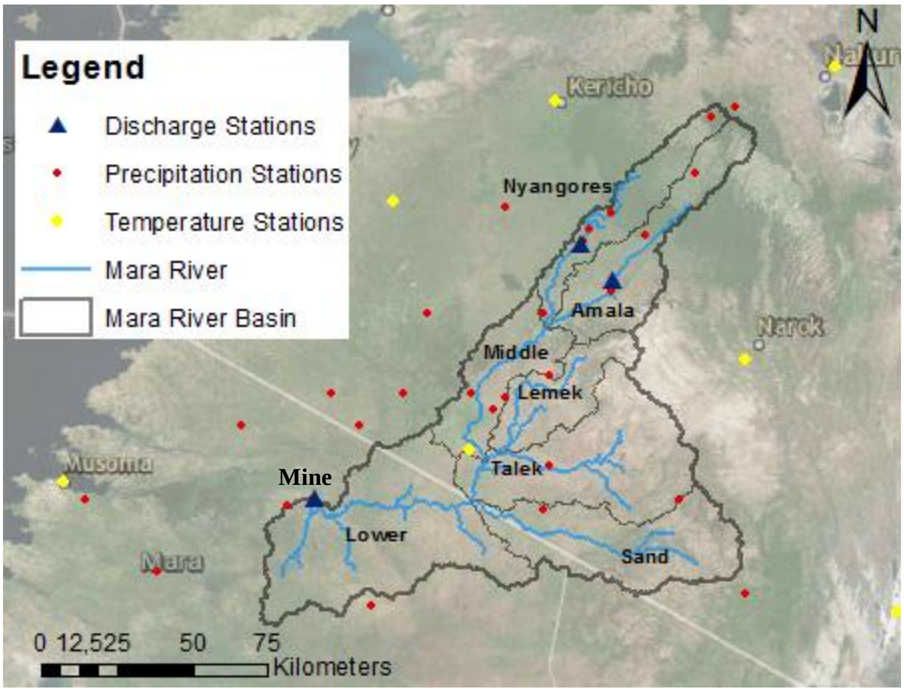


Figure 1: Map of the Mara River Basin and the hydro-meteorological stations for which data is available

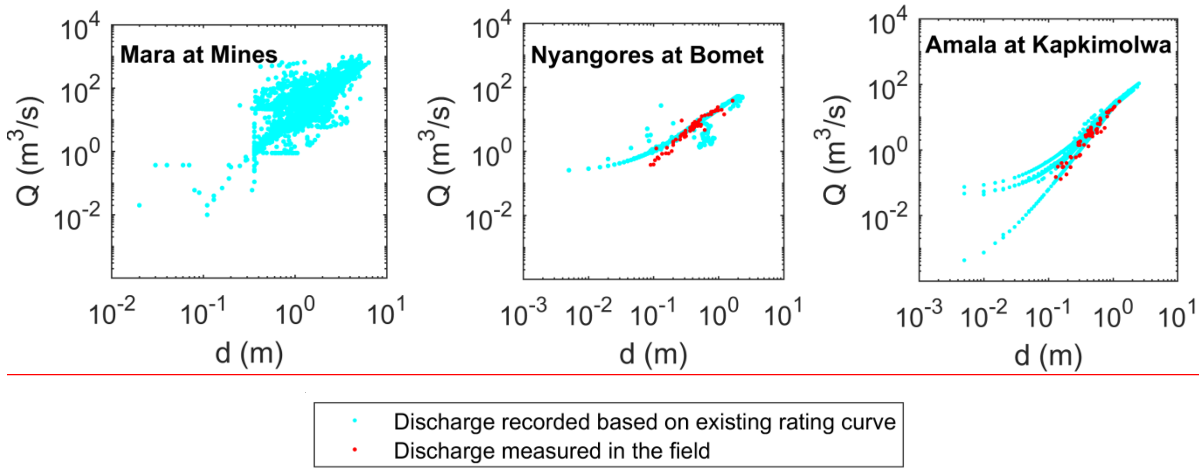


Figure 2: Discharge - water depth graphs for the three main river gauging stations in the Mara River Basin: Mara at Mines, Nyangores at Bomet and Amala at Kapkimolwa. 1) Recorded discharge and water level time series between 1960 and 2010 (light blue), 2) discharge field measurements from the Nile Decision Support Tool (NDST) for the time period 1963 - 1989 (Nyangores) and 1965 - 1992 (Amala), no data was available for Mines (red)

Table 5: Hydro-meteorological data availability in the Mara River Basin. The temporal coverage for water level and discharge can be different due to poor administration.

	Precipitation	Temperature	Water level, discharge		
Number of stations	~~29~~28	~~5~~7	3		
Station ID	-	-	1LA03	1LB02	5H2
Station location	-	-	Nyangores at Bomet	Amala at Kapkimolwa	Mara at Mines
Time range	1959 -2011	1957 - 2014	1963-2009	1955-2015	1969-2013
Duration [years]	0 - 43	3 - 57	46	60	44
Coverage	8 - 100%	30 -100%	Discharge: 85% Water level: 85%	Discharge: 72% Water level: 70%	Discharge: 53% Water level: 61%

Table 6: Discharge measured in the field using ~~a RiverSurveyor at three locations in the Mara River Basin. A RiverSurveyor is a small boat on which~~ an Acoustic Doppler Profiler~~,~~ (Sontek RiverSurveyor M9) mounted on a portable raft which is also equipped with a Power Communications Module and a DGPS ~~antenna was mounted~~ (Rey et al., 2015)

Station name	Date	Mean discharge	Standard deviation
Emarti Bridge	13 Sep 2014	19.2 m^3/s	0.7 m^3/s
	4 Oct 2014	13.4 m^3/s	0.6 m^3/s
Serena Pump House	9 Oct 2014	16.6 m^3/s	0.4 m^3/s
New Mara Bridge	19 Sep 2014	19.6 m^3/s	0.6 m^3/s
	6 Oct 2014	21.9 m^3/s	0.4 m^3/s

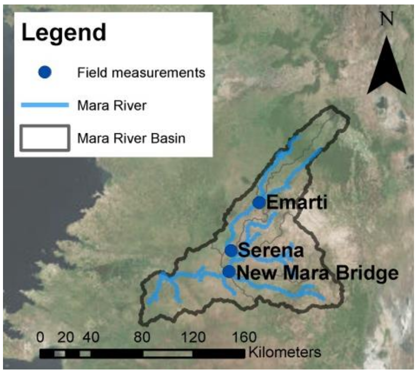


Figure 3: Map of discharge measurement locations during field trips in September/October 2014

2.1.2 Remotely sensed data

Besides ground observations, also remotely sensed data were used for setting up the rainfall-runoff model. Catchment classification was based on topography and land cover. For the topography, a digital elevation map (SRTM) with a resolution of 90 m and vertical accuracy of 16 m was used (U.S. Geological Survey, 2014). The land cover was based on Africover, a land cover database based on ground truth and satellite images (FAO, 1998). For the climate, remotely sensed precipitation was used from FEWSNET on a daily timescale from 2001 to 2010 and monthly actual evaporation from USGS from 2001 to 2013. Moreover, NDVI maps derived from Landsat images were used to define parameter constraints.

3 Hydrological model setup for the Mara River Basin

3.1 Catchment classification based on landscape and land use

For this study, the modelling concept of FLEX-Topo ~~was~~has been used~~.~~ (Savenije, 2010). It is a semi-distributed rainfall runoff modelling framework that distinguishes hydrological response units (HRUs) based on landscape features. The ~~landscapes~~landscape classes were identified based on the topographical ~~indexes~~indices HAND (Height Above Nearest Drain) and slope ~~(Savenije, 2010)~~ using a digital elevation map~~ (SRTM) with a resolution of 90 m and vertical accuracy of 16 m (U.S. Geological Survey, 2014).~~. Hillslopes are defined by a strong slope ~~(more than 12.9%)~~ and high HAND~~ (more than 5.9 m);~~, wetlands by a low HAND~~;~~, and terraces by a high HAND and mild slope. The ~~thresholds~~threshold for the slope ~~and HAND were~~(21.9%) was based on a sensitivity analyses within the Mara Basin~~.~~ which revealed that the area of a hillslopes changed asymptotically with the threshold. Therefore, the slope threshold was chosen at the point where changes in the sloped area become insignificant. As the wetland area was insignificant based on field observations, the HAND threshold was set to zero. In the Mara River Basin, there are mainly terraces and ~~hill slopes.~~ hillslopes.

To further delimit ~~HRUs,~~these two main landscape units, the land cover is taken into account ~~based on Africover, a land cover database based on ground truth and satellite images (FAO, 1998). This resulted in four HRUs in the sub-basin of the Mara River Basin: forested hill slopes, shrubs on hill slopes, agriculture and grassland (Figure 3, Figure 4 and Table 3).~~as well. In the upper sub-catchments, there are mainly ~~cropland~~croplands and ~~forest~~forests, whereas further south the land use is dominated by ~~grassland~~grasslands. In the lower sub-catchment, there ~~is~~are mostly ~~cropland.~~croplands and grasslands. This resulted in four HRUs within the sub-basin of the Mara River Basin: forested hillslopes, shrubs on hillslopes, agriculture and grassland (Figure 4, Figure 5 and Table 7).

Table 7: Classification results: area percentage of each hydrological response unit per sub-catchment in the Mara River Basin

Sub-catchment	Agriculture	Shrubs on ~~hill slopes~~hillslopes	Grassland	Forested ~~hill slopes~~hillslopes
Amala	67%	0%	0%	33%
Nyangores	61%	0%	0%	39%
Middle	19%	16%	65%	0%
Lemek	10%	39%	51%	0%
Talek	0%	21%	79%	0%
Sand	0%	42%	58%	0%
Lower	26%	23%	52%	0%

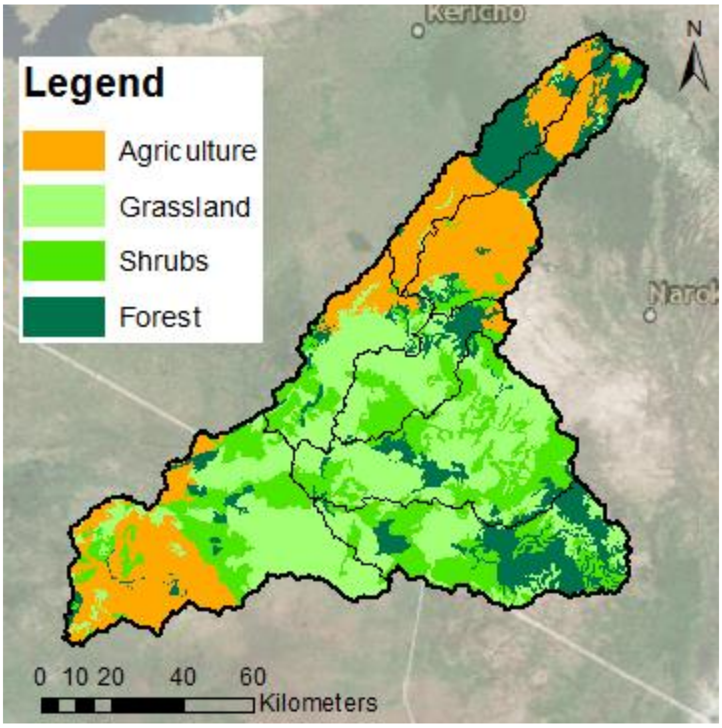


Figure 4: Classification of the Mara River Basin into four hydrological response units for each sub-catchment based on land use and landscape

3.2 Hydrological model structure

Each HRU is represented by a lumped conceptual model; the model structure is based on the dominant flow processes observed during field trips or deducted from interviews with local people. For example, in forests and shrub lands, Shallow Subsurface Flow (SSF) was seen to be the dominating flow mechanism: Rainwater infiltrates into the soil and flows through preferential flow paths to the river. In contrast, grassland and cropland generate overland flow. The observed soil compaction, due to cattle trampling and ploughing, reduces the preferential infiltration capacity resulting in overland flow during heavy rainfall. Consequently, ~~in these land~~there Hortonian Overland Flow (HOF) occurs at high rainfall intensities excessing the maximum infiltration capacity. The perception of the dominant flow mechanisms (~~Figure 4~~Figure 5) was then used to ~~identify a suitable~~develop the model structure (Figure 6). This approach of translating a perceptual model into a model concept (Beven, 2012) was applied successfully in previous FLEX-Topo applications (Gao et al., 2014a; Gharari et al., 2014).

The model structure contains multiple storage components schematised as reservoirs (Figure 6). For each reservoir, the inflow, outflow and storage are defined by water balance equations, see Table 8. Process equations determine the fluxes between these reservoirs as a function of input drivers and their storage. HRUs function in parallel and independently from each other. However, they are connected through the groundwater system and the drainage network. To find the total runoff at the sub-catchment outlet $Q_{m,sub}$, the outflow $Q_{m,i}$ of each HRU is multiplied by its relative area ~~percentage~~ and then added up together with the groundwater discharge Q_s. The relative area ~~percentage~~ is the area of a specific HRU divided by the entire sub-catchment area. Subsequently, the modelled discharge at the catchment outlet is obtained by using a simple river routing technique where a delay from sub-catchment outlet to catchment outlet was added assuming an average river flow velocity of 0.5

m/s. In the Sand sub-catchment, it is schematised that runoff can percolate to the groundwater from the river bed and that moisture can evaporate from the groundwater through deep rooting or riparian vegetation.

Table 8: Equations applied in the hydrological model. The formulas for the unsaturated zone are written for the hydrological response units: *Forested hillslopes* and *Shrubs on hillslopes*; for grass and agriculture, the inflow P_e changes to Q_F. The modelling time step is Δt = 1 day. Note that at a time daily step, the transfer of interception storage between consecutive days is assumed to be negligible.

Reservoir system	Water balance equation	Process functions
Interception	$\frac{\Delta S_i}{\Delta t} = P - P_e - E_i \approx 0$	$E_i = \cancel{\min\left(E_p, \min\left(P, \frac{I_{max}}{\Delta t}\right)\right)} \min\left(E_p, \min(P, I_{max})\right)$
Surface	$\frac{\Delta S_o}{\Delta t} = P_e - Q_F - Q_{HOF} - E_o$	$Q_F = \min(\frac{S_o}{\Delta t}, F_{\max})$
		$Q_{HOF} = \max\frac{(0, S_o - S_{max})}{\Delta t}$
		$E_o = \max(0, \min\left(E_p - E_i, \frac{S_o}{\Delta t}\right))$
Unsaturated zone	$\frac{\Delta S_u}{\Delta t} = (1 - C) * P_e - E$	$C = 1 - \left(1 - \frac{S_u}{S_{u,max}}\right)^{\beta}$
		$E = \min((E_p - E_i), \min\left(\frac{S_u}{\Delta t}, (E_p - E_i) * \frac{S_u}{S_{u,max}} * \frac{1}{C_e}\right))$
Groundwater recharge		$R_s = W * C * P_e$
Fast runoff	$\frac{\Delta S_f}{\Delta t} = R_{fl} - Q_f$	$R_{fl} = T_{lag}(C * P_e - R_s)$ → in a linear delay function T_{lag}
		$Q_f = \frac{S_f}{K_f}$
Groundwater	$\frac{\Delta S_s}{\Delta t} = R_{s,tot} - Q_s - E_s + Q_{inf}$	$R_{s,tot} = \sum_{i=1}^{i=4} R_{s;HRU_i}$
		$Q_s = \frac{S_s}{K_s}$
		$E_s = 0 \; and \; Q_{inf} = 0 \; for \; all \; sub-basins \; except \; Sand$
		$Q_{inf} = \min\left(\frac{S_{s,max} - S_s}{\Delta t}, Q_f\right) \; for \; Sand \; sub-basin$
		$E_s = \max\left(0, \min\left(E_p - E_i - E_o - E, \frac{S_s}{\Delta t}\right)\right) for \; Sand \; sub-basin$
Total runoff		$Q_m = Q_s + \sum_{i=1}^{i=4} Q_{f;HRU_i}$

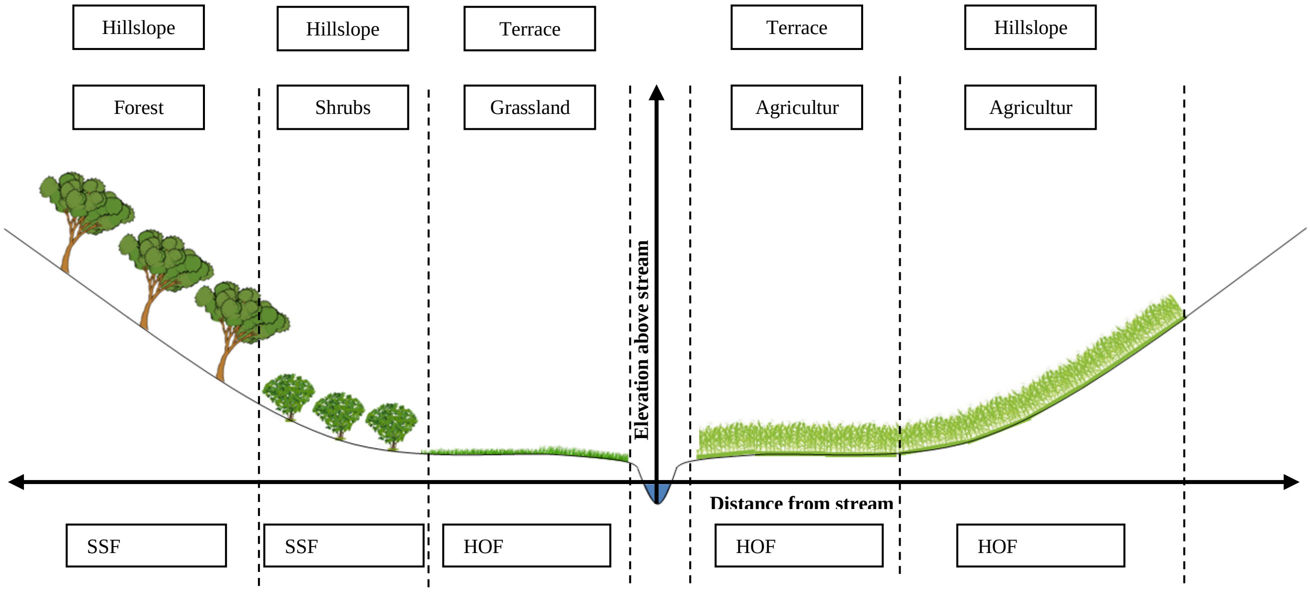


Figure 5: Schematization of the landscape and land use based classification

235

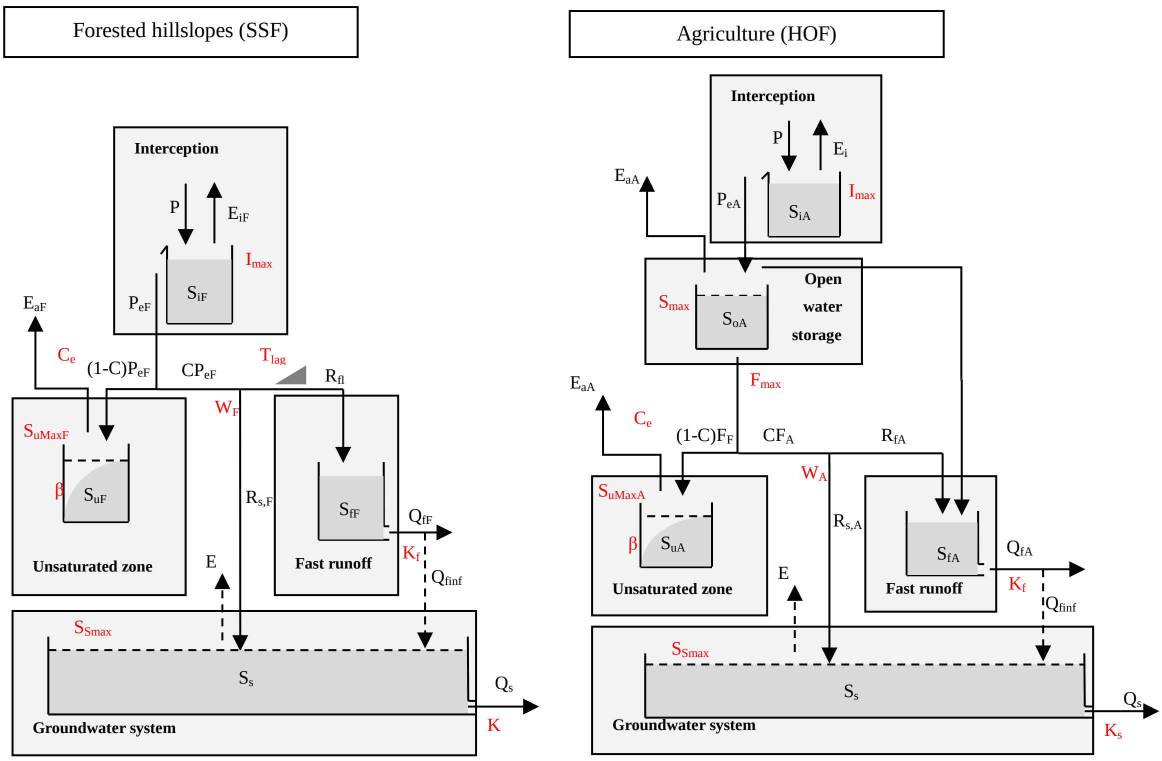


Figure 6: Model structure of the HRUs: *Forested* ~~hill slopes~~*hillslopes* (left) and *Agriculture* (right). The structure for *Shrubs on* ~~hill slopes~~*hillslopes* is similar to the left one replacing the indices F with S. The structure for *Grassland* is similar to the right one replacing the indices A with G. Parameters are marked in red, storages and fluxed in black. Symbol explanation: Fluxes: precipitation (*P*), evaporation of the interception zone (E_i), actual evaporation (E_a), evaporation from groundwater only applied in the sub-catchment Sand (E_s), effective precipitation (P_e), infiltration into the unsaturated zone (F_A), discharge from unsaturated zone to the fast runoff zone (R_f), groundwater recharge (R_s), discharge from the fast runoff (Q_f), infiltration into groundwater system only applied in the sub-catchment Sand ($Q_{f,\,inf}$), discharge from the slow runoff (Q_s). Storages: storage in the interception zone (S_i), open water storage (S_{oA}), storage in the root zone (S_u), storage for the slow runoff (S_s), storage for the fast runoff (S_f). Remaining symbols: splitter (*W*), splitter (*C*), soil moisture distribution coefficient (*β*), transpiration coefficient (C_e = 0.5), reservoir coefficient (*K*); indices f and s indicate the fast and slow runoff. Units: fluxes [mm/d], storages [mm], reservoir coefficient [d], remaining parameters [-].

3.3 Model constraints

Parameters and process constraints ~~have been~~were applied to eliminate unrealistic ~~model results~~parameter combinations and constrain the flow volume. Parameter constraints were applied to the maximum interception, reservoir coefficients, the storage capacity in the root zone or on the surface, and the slope-roughness parameter, Table 9. Process constraints were applied to the runoff coefficient, groundwater recharge, interception and infiltration, Table 10. The effect of including these parameter and process constraints is illustrated in Figure S5. For ~~example~~instance, the maximum storage in the unsaturated zone $S_{u,max}$~~, equal to~~ equals the root zone storage capacity~~, has been~~ and was estimated ~~based on~~using the method of Gao (2014) ~~using remote~~based on remotely sensed precipitation and evaporation ~~data~~ (Gao et al., 2014b; Wang-Erlandsson et al., 2016). The dry season evaporation has been derived from the actual evaporation using the NDVI. ~~In addition, the total evaporation has been constrained using the Budyko curve (Gharari et al., 2014). In the supplement, a list~~

Table 9: Overview of all parameter constraints applied in the hydrological model for the Mara River Basin

Parameter	**Symbol**	**Formula**	**Comment**
Interception	I_{max}	$I_{max,forest} > I_{max,grass}, I_{max,shrubs}, I_{max,cropland}$ $I_{max,shrubs} > I_{max,grass}, I_{max,cropland}$	Based on perception
Reservoir coefficient	K_s, K_f	$K_s > K_f$	Based on perception
Storage capacity in unsaturated zone	$S_{u,\,max}$	$S_{R,y_i} = \int P_e - E_d \, dt$ With: $\frac{E_d}{E_a} = \frac{\mathrm{NDVI_D}}{\mathrm{NDVI_A}}$ thus: $E_d = E_a * \frac{\mathrm{NDVI_D}}{\mathrm{NDVI_A}}$	Based on NDVI, equivalent to the root zone storage capacity (Gao et al., 2014b) $S_{R,\,yi}$: required storage for year i P_e: effective rainfall over dry season E_d: annual mean dry season evaporation, calculated assuming a linear relation between the evaporation and the NDVI E_a: actual mean annual evaporation $\mathrm{NDVI_D}$: annual mean dry season NDVI $\mathrm{NDVI_A}$: annual mean actual NDVI Through a statistical analysis of S_R using the Gumbel distribution, the storage capacity $S_{u,\,max}$ with a return period of 20 years is calculated.
Reservoir coefficient for groundwater system	K_s	$Q_s = Q_{t=0} * exp\left(-\frac{t}{K_s}\right)$	Based on hydrograph recession analysis Q_s: groundwater discharge
Maximum surface water storage	S_{max}	-	Based on DEM assuming S_{max} is equal to the sink volumes
Slope-roughness parameter	c	$Q = c * A * R^{\frac{2}{3}} = u * A$ $u = c * R^{\frac{2}{3}} \rightarrow c_{calculated} = \frac{u}{R^{\frac{2}{3}}}$ $c_{calculated,-25\%\,error} < c < c_{calculated,+25\%\,error}$	Based on Strickler formula, cross-section data and a single discharge and velocity measurement at Mines allowing a wide error margin of 25%

Table 10: Overview of all process constraints applied in the hydrological model for the Mara River Basin

Process	**Symbol**	**Formula**	**Comment**

Average annual runoff coefficient	C	$C = 1 - \frac{E}{P} = e^{-\frac{E_p}{P}}$	Based on the Budyko curve using the 95% percentile, hence the modelled average annual runoff coefficient should be below the 95-percentile of the observations
Groundwater recharge	R_s	$R_{s,F} > R_{s,C}, R_{s,G}$	Based on the assumption that deeper rooting vegetation creates preferential drainage patterns
Annual interception	E_i	$E_{i,F} > E_{i,G}, E_{i,S}$	Based on the assumption that the interception is higher in forests than in grassland and shrublands
Fast runoff infiltration	-	$f_{Q_{river}} < 3\ \mathrm{yr}^{-1}$	Frequency of river runoff. Based on interviews, locals seldom observed runoff more than 3 times a year.

~~3.3 Calibration and validation strategy using water level data~~

~~After having set up the model and defined the constraints, the model was calibrated and evaluated. The hydrological model was calibrated on water levels due to lack of reliable discharge data. For the evaluation of this calibration, the Nash Sutcliffe coefficient was used on the flow duration curve and its logarithm, see Eq. (1). The modelled water depth d_{mod} was calculated from the modelled discharge Q_{mod} using the Strickler formula and the cross-sectional geometry ($Q = k * i^{\frac{1}{2}} * A * R^{\frac{2}{3}} = c * A * R^{\frac{2}{3}}$), where R is the hydraulic radius and A the cross-sectional area; the unknown parameter c was calibrated. Note that by using the Strickler formula the exponent of the rating curve is fixed; $Q = a * (h - h_0)^b$. Also note that the parameter c compensates for non-closure of the water balance. Therefore the calibrated c values have to be checked whether they are in a feasible range of roughness and slope values. Subsequently, the discharge was estimated with the same Strickler formula, but now using the observed water depth d_{obs} which is the water level subtracted by the reference level. This discharge $Q_{Strickler}$ was then compared to the modelled discharge Q_{mod} and the recorded discharge Q_{rec}. As a result new geometric rating curves were obtained (relation between $Q_{Strickler}$ and d_{obs}) and compared to the recorded rating curves (Table 5 for a schematisation of the methodology).~~

~~The model was run for the entire catchment using the station Mines, and for the sub-catchments Nyangores and Amala. For each simulation, the obtained water depth was evaluated by the flow duration curve, the water level time series and the logarithm of the time series. The selected time periods for each simulation were:~~

~~Mines (5H2):~~

- ~~Calibration 1970-1974~~
- ~~Validation 1 1980-1981~~
- ~~Validation 2 1982-1983~~

~~Nyangores (1LA03):~~

- ~~Calibration 1970-1980~~
- ~~Validation 1981-1992~~

Amala (1LB02):

- Calibration 1991-1992
- Validation 1985-1986

Equation 1: Formulas for the Nash-Sutcliff objective function. The indices mod and obs indicate modelled and observed values, respectively. In all cases, sorted data was used for the calculation of the objective function therefore the flow duration curve was calibrated.

$$NS_{\log(d)} = 1 - \frac{\Sigma\left(\log\left(d_{mod,sorted}\right) - \log\left(d_{obs,sorted}\right)\right)}{\Sigma\left(\log\left(d_{obs,sorted}\right) - \log\left(d_{obs,avg}\right)\right)} \qquad NS_{d} = 1 - \frac{\Sigma\left(d_{mod,sorted} - d_{obs,sorted}\right)}{\Sigma\left(d_{obs,sorted} - d_{obs,avg}\right)}$$

Table 5: Schematisation of the methodology

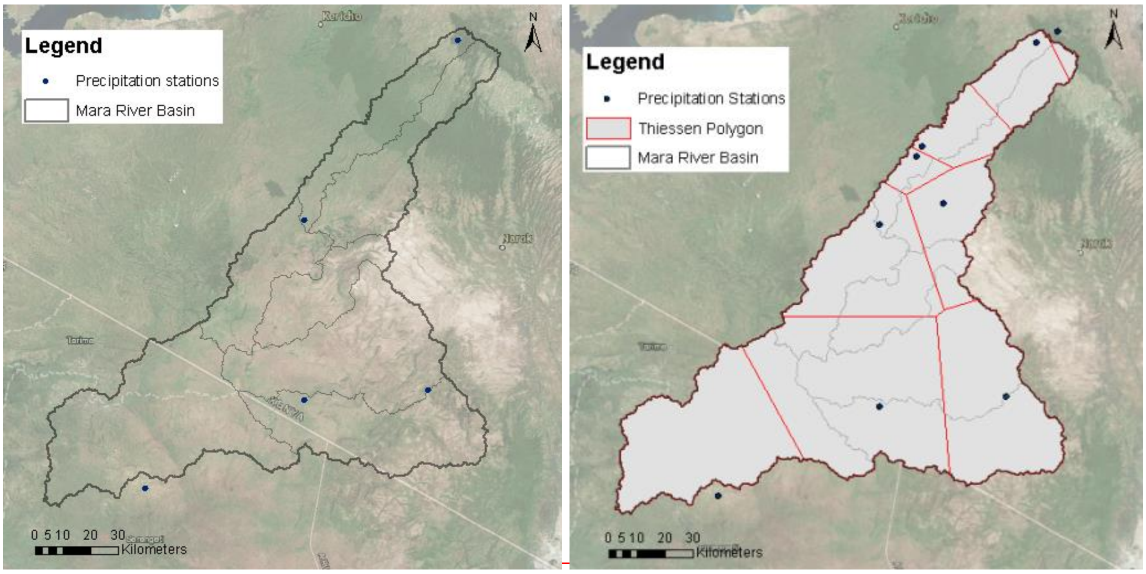
	Model input data		**Model output data**		
Model	Precipitation Temperature	FLEX-Topo	Q_{mod} (Modelled discharge, output from FLEX-Topo)	Strickler	d_{mod} (Modelled water depth calculated with Q_{mod} using the Strickler formula, parameter c is calibrated)
			$Q_{Strickler}$ (Discharge calculated with d_{obs} using the Strickler formula, parameter c is calibrated)	Geometric rating curve	Calibration
Observation			Q_{rec} (Recorded discharge)	Recorded rating curve	d_{obs} (Observed water depth)

3.4 Precipitation input data

For the precipitation input data, a single station was chosen for each sub-catchment assuming it was representative for the entire area (Figure 6 A). However, the representative average precipitation in an sub-catchment can also be estimated using Thiessen polygons (Figure 6 B). Alternatively, multiple precipitation stations can be used within a single sub-catchment by splitting it up into different areas with the same precipitation based on Thiessen polygons. Therefore, the following three methods were used to estimate the representative precipitation for the hydrological model:

- Method 1: Selection of a single station for each sub-catchment assuming it was representative for the entire area
- Method 2: Calculation of the representative average precipitation for each sub-catchment using Thiessen polygons
- Method 3: Sub-division of each sub-catchment into areas with equal rainfall using Thiessen polygons

Method 1 was used as reference and the remaining two methods to assess the model sensitivity to areal rainfall estimates.



~~**Figure 6: Map of the precipitation stations used for modelling based on A) Method 1 and B) Method 2 and 3 for areal rainfall estimates. Method 1: Single precipitation station for each sub-catchment; Method 2: Representative average precipitation for each sub-catchment using Thiessen polygons; Method 3: Sub-division of each sub-catchment into areas with equal rainfall using Thiessen polygons**~~

3.4 Model calibration method using water levels

The hydrological model was calibrated on a daily timescale applying the MOSCEM-UA algorithm (Vrugt et al., 2003) with parameter ranges and values as indicated in Table S1 and S2. For the calibration, the Nash-Sutcliffe coefficient was calculated on the water level duration curve (Eq.1 linear, and Eq.2 log-scale). By calibrating on the duration curve, the focus is on the flow statistics and not on the timing of individual flow peaks. This information is also in the time series This is justified since there were high uncertainties in the timings of floods events due to the limited number of available rainfall stations to capture the spatial variability of the rainfall input well. Therefore, duration curves were considered as a good signature for calibrating this model; this was also concluded in previous studies (Westerberg et al., 2011; Yadav et al., 2007).

$$NS_d = 1 - \frac{\Sigma(h_{mod,sorted} - h_{obs,sorted})}{\Sigma(h_{obs,sorted} - h_{obs,avg})} \quad (1)$$

$$NS_{\log(d)} = 1 - \frac{\Sigma(\log(h_{mod,sorted}) - \log(h_{obs,sorted}))}{\Sigma(\log(h_{obs,sorted}) - \log(h_{obs,avg}))} \quad (2)$$

For the water level based calibration, the modelled discharge needs to be converted to modelled water level. This calculation was done with the Strickler-Manning formula in which the discharge is a function of the water level (Eq. (3)), where R is the hydraulic radius (Eq. (6)), A the cross-sectional area (Eq. (5)), i the slope, k the roughness and c the slope-roughness parameter (Eq. (4)). The hydraulic radius and cross-section are a function of the water depth d which is the water level subtracted h by the reference level h_0 (Eq. (7)). The cross-sections were simplified as a trapezium with river width B and two different river bank slopes i_1 and i_2; these coefficients (Table 1) were estimated based on the available cross-section information (Figures S6 – S8). Since the slope and roughness are unknown, the slope-roughness parameter c was calibrated.

$$Q = k * i^{\frac{1}{2}} * A * R^{\frac{2}{3}} = c * A * R^{\frac{2}{3}} \quad (3)$$

$$c = k * i^{\frac{1}{2}} \quad (4)$$

$$A = B * d + \frac{1}{2} * d * (i_1 + i_2) * d \quad (5)$$

$$R = \frac{A}{B + d * \left((1+i_1^2)^{\frac{1}{2}} + (1+i_2^2)^{\frac{1}{2}} \right)} \quad (6)$$

$$d = h - h_0 \quad (7)$$

Table 11: Coefficients used for the river cross-section

	River width B [m]	River bank slope i_1 [-]	River bank slope i_2 [-]	Reference level h_0 [m]
Amala	10.0	3.50	1.83	0
Nyangores	19.05	2.65	5.56	0
Mines	43.81	3.53	3.66	10

This model calibration method, illustrated graphically in Figure 7, was applied to three basins individually: the entire river basin using the station Mines, and for the sub-catchments Nyangores and Amala. At each location, the model was calibrated and validated for time periods indicated in Table 12; at Mines two time periods were used for validation to maximise the use of the available ground measurements.

Table 12: Time periods used for the calibration and validation at three basins: Mines, Nyangores and Amala

	Mines	Nyangores	Amala
Calibration time period	1970-1974	1970-1980	1991-1992
Validation time period	1980-1981	1981-1992	1985-1986
	1982-1983		

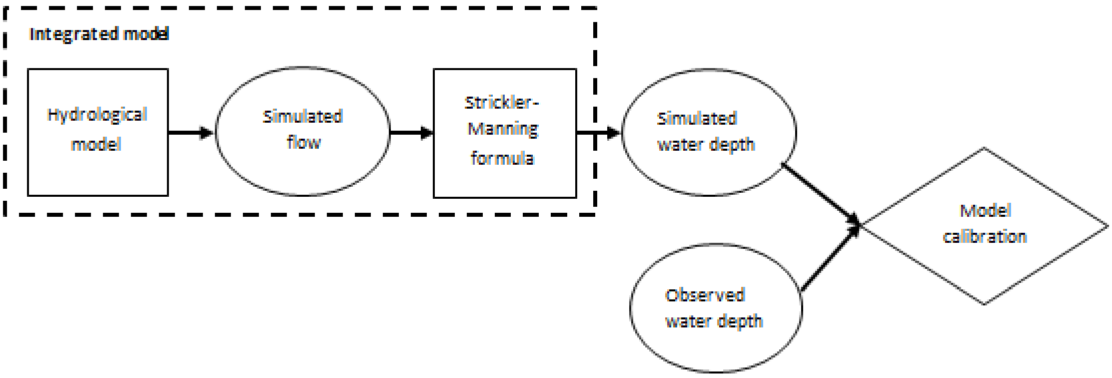


Figure 7: Flow chart of the proposed calibration method

3.5 Rating curve analysis

After calibration, the modelled water levels and discharges were analysed. For the model calibration and validation, the modelled and recorded water levels were compared at basin level, focusing on the time series and the duration curves. Hereafter, water level – discharge relations were analysed taking two rating curves into consideration:

- "Recorded rating curve", relating Q_{rec} to h_{obs},
- "Geometric rating curve", relating $Q_{Strickler}$ to h_{obs}.

The geometric rating curve relates the modelled discharge $Q_{Strickler}$ to the observed water level h_{obs}. This discharge $Q_{Strickler}$ was calculated with the Strickler-Manning formula using the calibrated slope-roughness parameter *c*, cross-section data, and the observed water level h_{obs}. Therefore, the equation behind the geometric rating curve basically is the Strickler-Manning formula (Eq. (3)) instead of the traditional rating curve equation (Eq. (8)). The advantage of the Strickler-Manning formula is that only one parameter is unknown (river bed slope and roughness *c*, Eq. (4)), instead of two (fitting parameters *a*, and *b*). However, the Strickler-Manning rating curve approach requires additional information on the cross-section.

$$Q = a * (h - h_0)^b \tag{8}$$

4 Results and discussion

4.1 Water ~~depth~~level time series and ~~flow~~ duration curve

Model results were analysed graphically (Figure 8 to Figure 10 and Figure S9 to Figure S19) and numerically based on the Nash-Sutcliffe values for the objective functions (Table 13). The results of the objective functions indicate that at Nyangores and Mines the calibration and validation results were ~~more~~ consistent ~~(**Fout! Verwijzingsbron niet gevonden.**)~~. At Mines, the ~~observed and~~ modelled water ~~depth were quite similar to each other~~level was simulated well, particularly with regard to the duration curve (~~Table 6 and Figure 7).~~Figure 8). At individual events, there were substantial differences~~, but this could be due to the spatial heterogeneity of the rainfall that were not represented well~~. In some years, for example in 1974, the ~~forcing~~observed data~~. On the other hand, the year 1974 was~~ were very well represented~~.~~ by the model outcome, however, in other years this was not the case. In general, the model captured the dynamics in the water ~~depth~~level well. This was the case during both calibration and validation (see ~~supplement~~Figure S12 and S13).

At Nyangores the observed and modelled water ~~depths~~levels were also similar during calibration and validation, extreme high flows excluded (Figure 9). However, at Amala, the observed and modelled water ~~depths~~levels differed significantly during calibration (Figure 10) and validation~~.~~ (Figure S15). The model missed several ~~rain~~discharge events completely, likely ~~linked~~related to missing rain fall events in the input data due to the high heterogeneity in precipitation. ~~Also there seemed to be backwater effects raising the water level, possibly due to a river blockage such as a weir, sand dam or dunes.~~

Table 13: Overview of the values of the objective functions for each model simulation. Calibration was done based on the water ~~depth~~level: $NS_{log(\cancel{d}h)}$ and ~~NSd~~NS_h; for comparison, objective functions using the discharge were added here as well

	Nyangores		Amala		Mines		
	Calibration	Validation	Calibration	Validation	Calibration	Validation 1	Validation 2
$NS_{log(d)}$	0.92	0.75	0.92	-0.23	0.97	0.81	0.93
NS_d	0.80	0.69	0.26	0.37	0.97	0.92	0.89
$NS_{log(Q)}$	0.92	0.69	0.57	0.63	0.97	0.81	0.93
NS_Q	0.55	0.37	0.08	-1.67	0.90	0.76	0.77

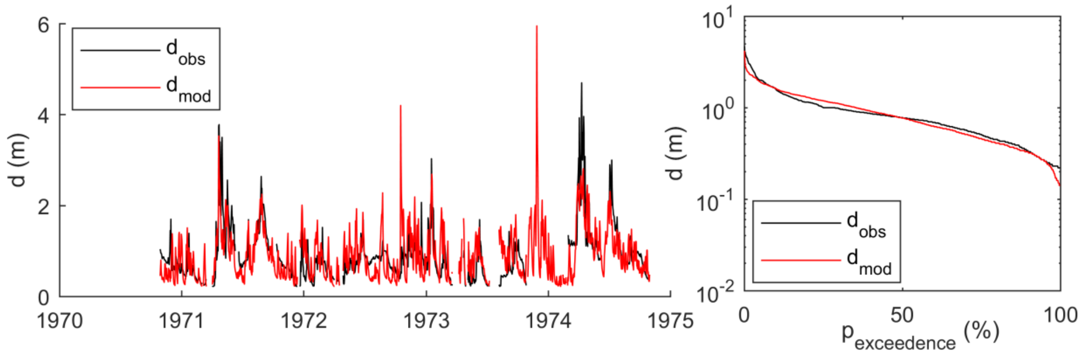


Figure 8: Model results at Mines during calibration: water depth time series and water depth exceedance

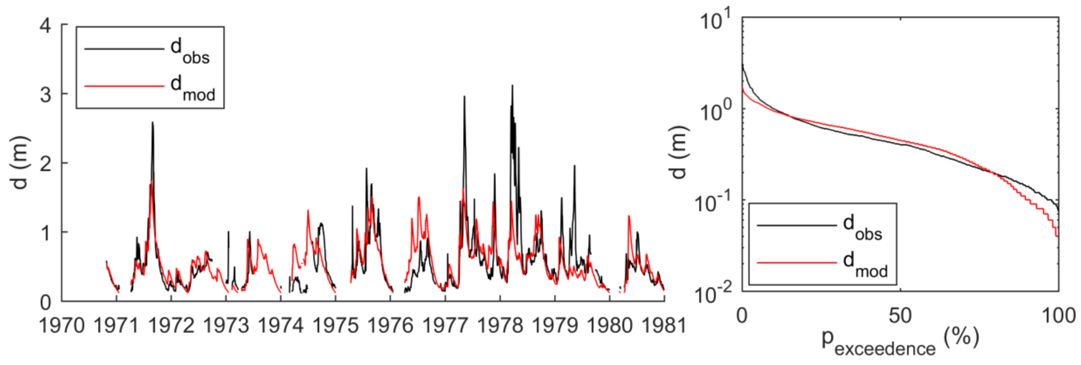


Figure 9: Model results at Nyangores during calibration: water depth time series and water depth exceedance

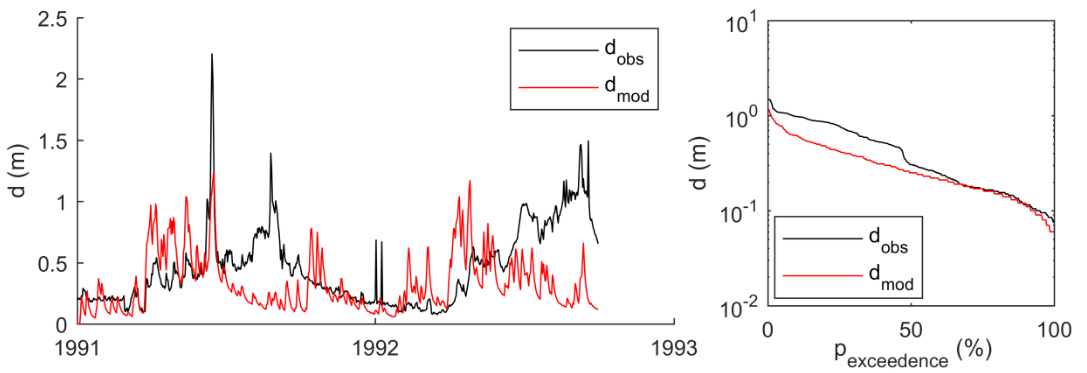


Figure 10: Model results at Amala during calibration: water depth time series and water depth exceedance

4.2 Discharge at sub-catchment level

At Mines, the discharge originates from seven different sub-catchments, each with a different contribution. Based on field observations, the ~~upper~~mountainous upstream sub-catchments from the north should have the largest contribution whereas the contribution from the relatively drier and flatter Lemek and Talek tributaries from the eastern part of the catchment should be relatively low. The contribution of each sub-catchment to the total modelled discharge was assessed on a monthly timescale and compared with observations.

As shown in ~~Figure 10~~Figure 11, the contribution varied throughout the year. In the summer (July-September), the modelled discharge mainly originates from the ~~upper~~northern sub-catchments, Nyangores and Amala~~, just as expected.~~. However, in the winter (November-April), the modelled discharge mainly originates from the Sand and Lower sub-~~catchment~~catchments. The eastern Middle, Talek and Lemek sub-catchments have the lowest discharge throughout the entire year just as observed.

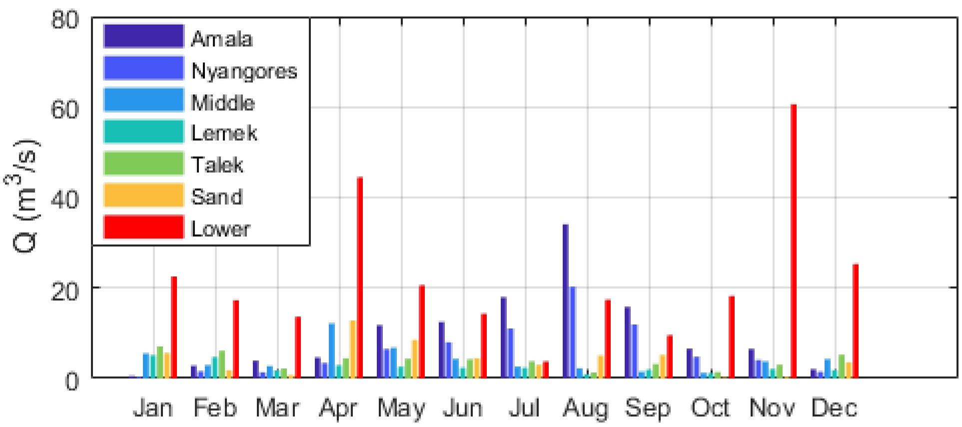


Figure 11: Monthly averaged modelled discharge for each sub-catchment

~~To validate the model at sub-catchment level, model results were compared with discharge measurements done during field trips in September/October 2014 at Emarti Bridge, Serena Pump House and New Mara Bridge. At all three locations, the modelled discharge in the same month was of the same order of magnitude as the point measurement, see Figure 11.~~ In previous studies, it ~~was~~has been shown that only a few discharge measurements can contain sufficient information to constrain model predictive uncertainties effectively (Seibert et al., 2009). To evaluate the model at sub-catchment level, model results were compared with discharge measurements done during field trips in September/October 2014 at Emarti Bridge, Serena Pump House and New Mara Bridge. At all three locations, the point measurements fitted well within the range of the modelled discharge (see Figure 12).

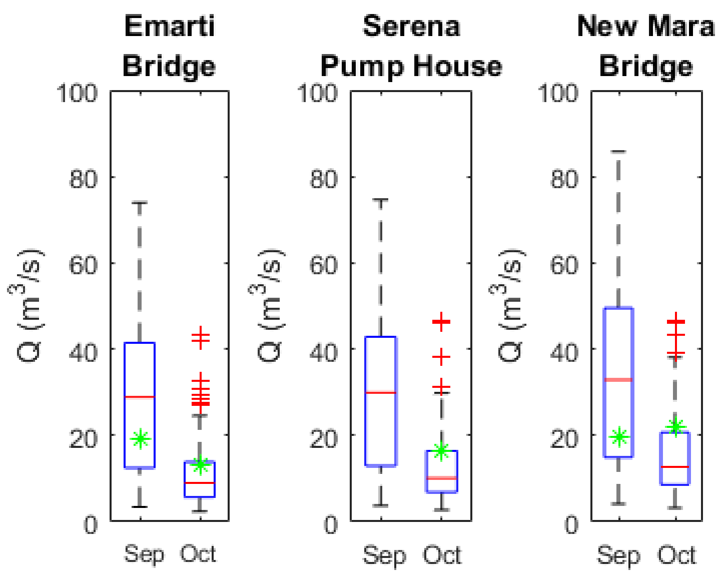


Figure 12: Boxplot of the modelled discharge at three locations; the green asterix represents the measured discharge in Sep/Oct 2014

4.3 Rating curve analysis

~~The discharge and rating curves have been evaluated by analysing discharge – water depth graphs which basically plot the rating curve. In this study, two different rating curves are distinguished:~~

- ~~“Geometric rating curve”, relating $Q_{Strickler}$ to d_{obs}, and~~
- ~~“Recorded rating curve”, relating Q_{rec} to d_{obs}~~

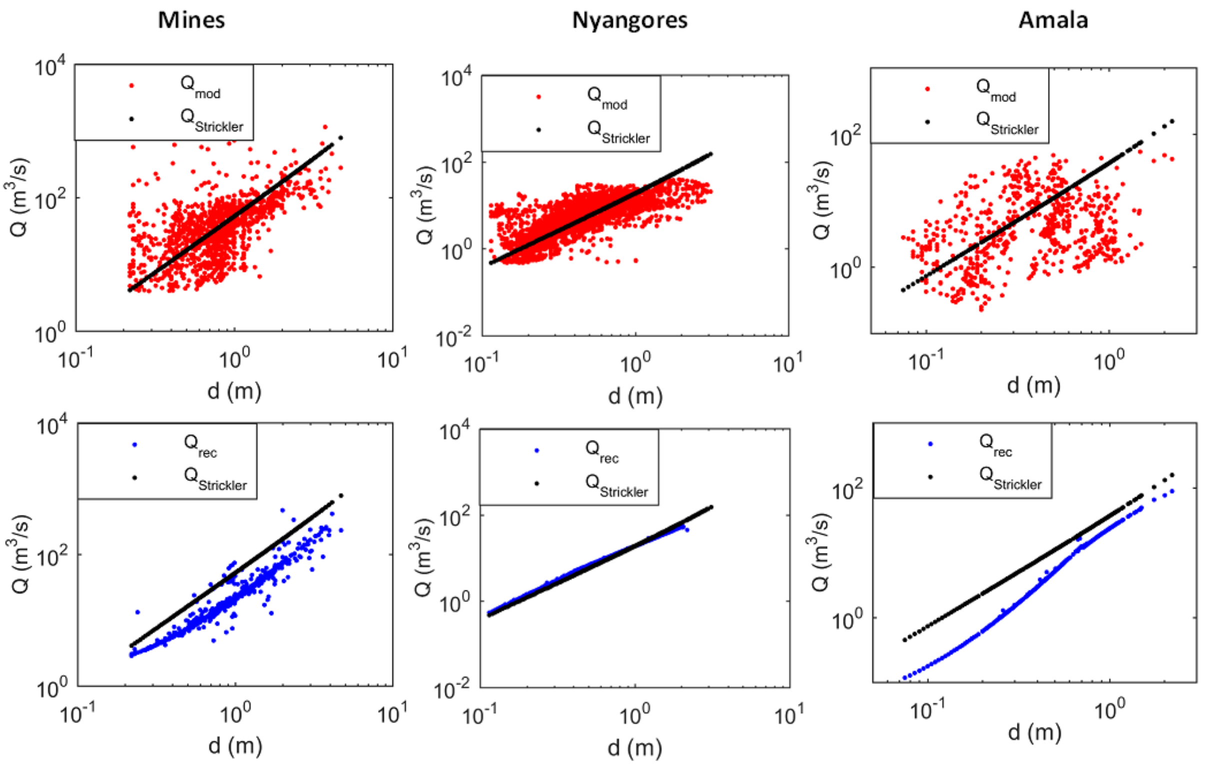
~~At Mines, the modelled discharge correlated with $Q_{Strickler}$, but with considerable scatter (~~In this study, the recorded and geometric (Strickler-Manning) rating curves were compared (Figure 13). ~~Comparison of the recorded discharge and $Q_{Strickler}$ however, revealed that the recorded discharge was lower. Therefore also the recorded and geometric rating curves were~~ At Mines, these two rating curves differed significantly ~~different from each other. However, for~~. For medium to high flows, both rating curves, recorded and geometric, ~~were~~ run parallel ~~to each other~~ indicating similar cross-sectional properties~~. This observation reoccurred~~; only the off-set differed through changing river bed levels. On the other hand, the simulated cross-section average flow velocity were realistic compared to the point measurements at Mines indicating that velocities are greater than 2 m/s during ~~validation as well.~~

~~The~~ high flows (see Figure 13). At Nyangores, the recorded and geometric rating curves were almost identical, while there were significant differences at Amala gauging station, especially in the low flows. Interestingly, these observations also hold for the validation period for all three stations.

The difference between the recorded and geometric rating ~~curve~~curves at Mines ~~can be a result of~~probably resulted from uncertainties in the available recorded discharge data~~, hence the recorded rating curve.~~. In the complete discharge ~~-~~– water ~~depth~~level graphs ~~at Mines (see supplement), a~~for all available data (Figure S2), large scatter ~~is found in the observation which should not be the case assuming one rating curve was used, compared with Nyangores where there is no scatter. This scatter~~ was found. This could be the result of natural variability in ~~for example~~ the reference water level h_0 in the rating curve equation ~~for example due to sand banks and bed forms.~~which was not taken into account. A sensitivity analysis of the recorded rating curve equation at

Mines showed that a deviation of 0.1 m in the reference water level ~~altered~~could alter the discharge with 4% ~~- 46%, lowest~~ for high flows and ~~highest~~46% for low flows. However, a deviation of 0.5 m ~~in the reference water level~~ resulted in a 19% ~~-~~ – 325% change in the discharge. Therefore, ~~variability~~unnoticed variations in the ~~reference water~~river bed level ~~increases~~strongly affects the uncertainty in the recorded rating curve~~. The uncertainties in the discharge data can also be seen in the calculated cross-section average flow velocity based on the recorded discharge and water level data: this was below 1 m/s (see supplement) whereas for example the measured velocity in 2012 was 2.13 m/s (GLOWS-FIU, 2012).~~ at Mara Mines, which is located in morphologically dynamic section of the river (Stoop, 2017).

~~At Nyangores and Amala, the modelled discharge correlated with~~ ~~$Q_{Strickler}$,~~ ~~also with considerable scatter (Figure 12). The recorded and geometric rating curves were almost identical at Nyangores, but not at Amala.~~



At Amala, the difference between both rating curves could be related to the effect of missing rain events in the input data as result of the short time series for calibration and validation. This resulted in absent discharge peaks and hence an underestimation of the flow; most extremely at Amala. During model calibration, this was compensated by increasing the parameter c in the Strickler-Manning formula (Eq. (4)). As a result, discharge values during missed events were increased, but also for all other days. The compensation effect was limited though since the model was calibrated on the duration curves instead of the time series. As parameter c is linearly related to the geometric rating curve (Eq. (3)), the latter was overestimated as well. Therefore, missing rain events in the input data resulted in the overestimation of the geometric rating curve.

In short, at the two stations with inconsistent rating curves, Amala and Mines, the geometric rating curve deviated significantly from the recordings. Strikingly, the deviations were observed at the same flow magnitudes where large inconsistencies were found in the observations, for instance in the low flows at Amala. However, at the gauging station with a reliable rating curve, Nyangores, the geometric and recorded discharge-water level relations were almost identical.

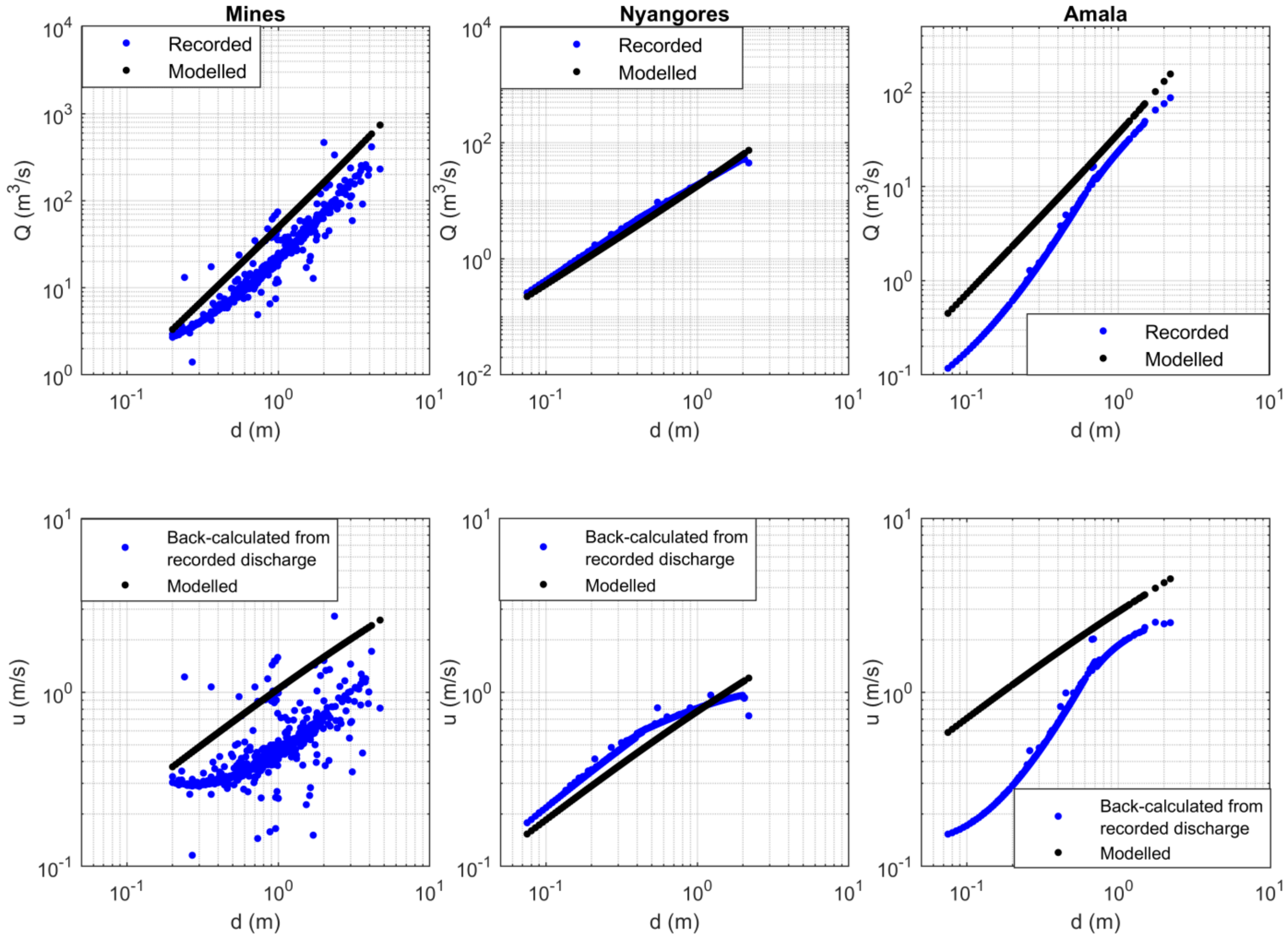


Figure 13: Model calibration results at Mines, Nyangores and Amala: Discharge ~~-~~ – water depth graphs (upper) and velocity – water depth graphs (lower).

~~**4.4 Sensitivity to areal rainfall estimates**~~

~~In the previous sections, it was shown that water level data can be used instead of discharge data to calibrate a model and to establish a rating curve equation. However, how sensitive is this method to areal rainfall estimation methodologies? This was analysed by comparing three different methods of representative rainfall estimates for each sub-catchment: 1) single station, 2) average of multiple stations based on Thiessen polygons, 3) sub-division into areas with equal rainfall based on Thiessen polygons. All three methods resulted in different daily or monthly rainfall values; the maximum difference was 86 mm/month at Amala in August (Figure 13). In general, there were more dry days when using a single station for each sub-catchment (method 1). Also, when using Thiessen polygons (methods 2 and 3), rainfall events were more dampened as a result of averaging multiple stations.~~

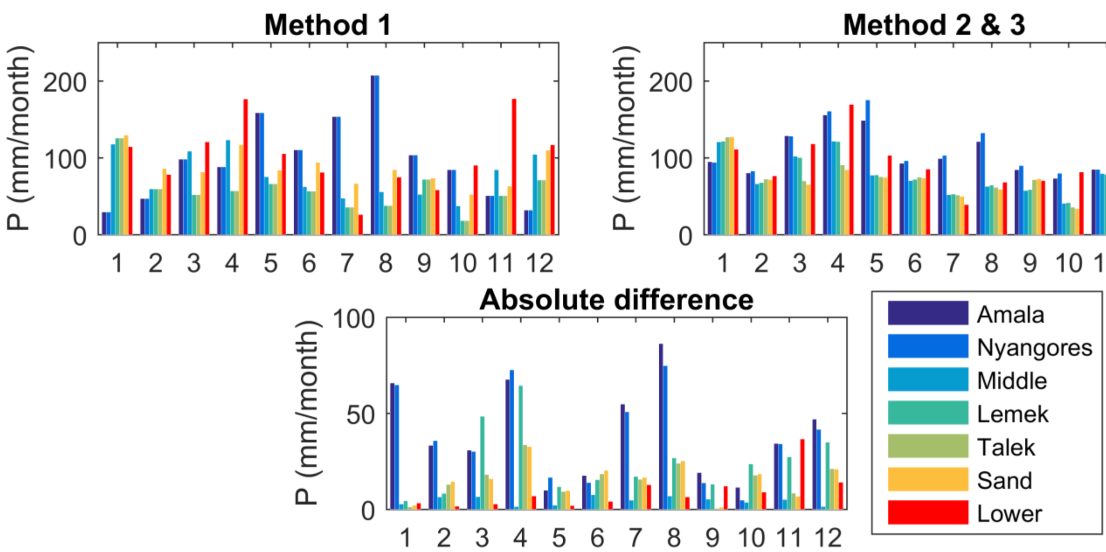
~~These differences in the precipitation data were reflected in the modelled water depth. Compared to the observation, method 1 resulted in very flashy responses and method 2 very dampened ones whereas method 3 was a combination of both (Figure 14). The change in precipitation input data also influenced the geometric~~

~~rating curve as shown in Table 7: the constant, parameter a, in the rating curve equation~~ $\cancel{Q = a * (h - h_0)^b}$ ~~increased with 45% and 35% for methods 2 and 3 respectively. This difference was within the modelling uncertainty bounds which was 75% in this case. However, this change in the rating curve constant indicates that the model compensated errors in the rainfall data by closing the water balance.~~

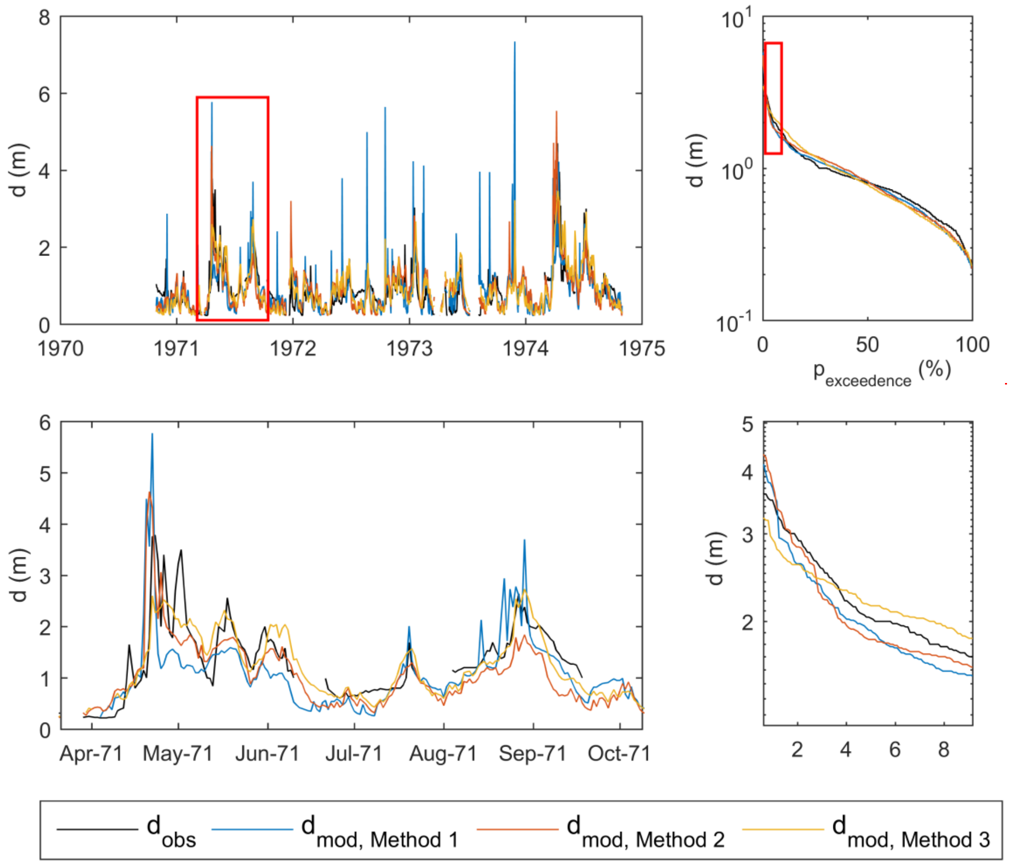
~~Besides altering the geometric rating curve equation, the precipitation estimation method also influenced the modelled annual averaged runoff coefficient (Figure 15). Averaged over the entire river catchment, this difference in runoff coefficient was insignificant, however on sub-catchment level, the largest variation was found in the Sand sub-catchment: the runoff coefficient changed from 5% with method 1 to 1% with method 2.~~

~~Table 7: Recorded rating curve and model results for the geometric rating curves using three different methods for areal rainfall estimates. Method 1: Single precipitation station for each sub-catchment; Method 2: Representative average precipitation for each sub-catchment using Thiessen polygons; Method 3: Sub-division of each sub-catchment into areas with equal rainfall using Thiessen polygons~~

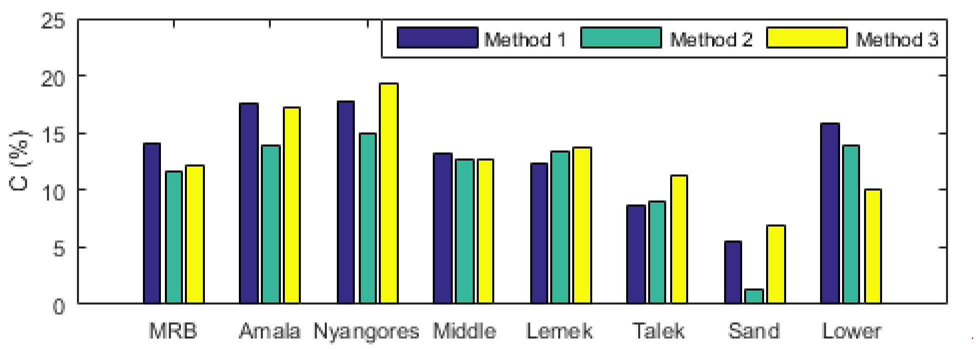
	~~**Method 1**~~	~~**Method 2**~~	~~**Method 3**~~
~~**Recorded rating curve**~~	$\cancel{Q_{obs} = 23.1 * (h - h_0)^{1.54}}$	$\cancel{Q_{obs} = 23.1 * (h - h_0)^{1.54}}$	$\cancel{Q_{obs} = 23.1 * (h - h_0)^{1.54}}$
~~**Geometric rating curve**~~	$\cancel{Q_{str} = 52.5 * (h - h_0)^{1.70}}$	$\cancel{Q_{str} = 47.4 * (h - h_0)^{1.70}}$	$\cancel{Q_{str} = 46.3 * (h - h_0)^{1.70}}$



~~Figure 13: Monthly average precipitation per sub-catchment. A) Method 1, B) Method 2 and 3, C) Absolute difference. Method 1: Single precipitation station for each sub-catchment; Method 2: Representative average precipitation for each sub-catchment using Thiessen polygons; Method 3: Sub-division of each sub-catchment into areas with equal rainfall using Thiessen polygons~~



~~**Figure 14: Modelled water depth for the Mara River Basin at Mines: time series (left) and flow duration curve (right) for the entire modelled time series (upper) and zoomed in a section marked in the red boxes (lower) using the model parameters obtained with three methods for areal rainfall estimates. Method 1: Single precipitation station for each sub-catchment; Method 2: Representative average precipitation for each sub-catchment using Thiessen polygons; Method 3: Sub-division of each sub-catchment into areas with equal rainfall using Thiessen polygons**~~



~~**Figure 15: Modelled runoff coefficient for the entire Mara River Basin (MRB) and each sub-catchment with the three methods for areal rainfall estimates. Method 1: Single precipitation station for each sub-catchment; Method 2: Representative average precipitation for each sub-catchment using Thiessen polygons; Method 3: Sub-division of each sub-catchment into areas with equal rainfall using Thiessen polygons**~~

4.5 Limitations

This study illustrates the potential of water level time series for model calibration, also in semi-arid river basins with insufficient discharge data. However, there are several limitations to this method. First, the slope-roughness parameter compensates for non-closure effects in the water balance, for instance due to errors in the precipitation which is extremely heterogeneous in semi-arid Mara basin. Unfortunately, this heterogeneity is poorly described in our study area with the available rain gauges (see section S7.2 on the precipitation data analysis) influencing the modelling results. Therefore, this parameter should be constrained to minimize this compensation as much as possible. Second, the cross-section was assumed to be constant during the modelling time period. Data analyses indicated that expected changes in the river width or slope cannot affect the rating curve significantly. However, if this is not the case, then this cross-section change should be included during the model calibration.

5 Summary and Conclusion

~~Hydrological models play an important role in Water Resources Management. Unfortunately, the quantity and quality~~The goal of ~~the available discharge measurements are often inadequate for reliable hydrological modelling in African river catchments. There are various sources of uncertainty in discharge time series when using rating curves due to extrapolation to estimate flood peaks or non-stationarity due to sedimentation or erosion altering the cross section. To cope with these uncertainties during model calibrations, there are two options: 1) assess the uncertainty in discharge data and its effect on model predictions, or 2) avoid these uncertainties by~~this paper was to illustrate a new calibration method using water level ~~data~~time series instead~~.~~

~~In this study, a hydrological model is developed for the~~ of discharge in a semi-arid and poorly gauged ~~Mara River Basin~~basin. This method offers a potential alternative for calibration on discharge data, as ~~a case study. The effects of the discharge data uncertainties are avoided by using water level instead of discharge time series by incorporating the hydraulic equation describing the rating curve within the model. A~~is common practice also in poorly gauged catchments. The semi-distributed rainfall runoff modelling framework ~~called~~ FLEX-Topo was ~~used to model the Mara River Basin.~~applied. The catchment was ~~split~~divided into four hydrological response units (HRUs) and seven sub-catchments based on the river tributaries. For each HRU, a unique model structure was defined based on the ~~expected~~observed dominant flow processes. By constraining the parameters and processes, unrealistic ~~results~~parameter sets were excluded from the calibration parameter set and the flow volume was constrained. This model was ~~then~~ calibrated based on water ~~depths~~levels to capture the flow dynamics~~; ~~. For this purpose, the modelled discharge was converted to water ~~depths were calculated from modelled discharges with cross-section data and~~levels using the Strickler-Manning formula. The unknown slope-roughness parameter was calibrated.

~~The hydrological model simulated the water depths well for the entire basin and the Nyangores sub-catchment in the north. In addition, a new geometric rating curve was calibrated based on the modelled discharge, observed water level and the Strickler formula. The geometric and recorded rating curve were slightly different at Mines, the catchment outlet, probably due to uncertainties in the recorded discharge data. At Nyangores however, the modelled and recorded discharge were almost identical. In addition, it was found that the precipitation estimation methodology influenced the model results significantly: application of a single station for each sub-catchment resulted in flashier responses whereas Thiessen averaged precipitation resulted in more dampened responses. The inadequate knowledge of the spatial distribution of the precipitation was the main limitation for accurate rainfall-runoff modelling. Therefore rapidly improving precipitation monitoring methods from space offer promising approximations for improving rainfall-runoff modelling in poorly gauged basins. Note that by calibrating the unknown parameter of the hydraulic equation, a combination of slope and roughness, the non-closure of the water balance is compensated as also errors in the rainfall data. Therefore, calibrated parameter values should be verified and if possible constrained.~~

~~In conclusion, promising results have been obtained when using water level time series for calibrating the hydrological model of the Mara River Basin in combination with process controls to constrain the flow volume.~~
An important output of this calibration approach is the “geometric rating curve equation” which relates the discharge to the water level using the Strickler-Manning formula. The geometric and recorded rating curves were significantly different at the following two gauging stations: Mines, the catchment outlet, and Amala, a sub-catchment outlet. At both locations, the deviations were with the same flow magnitudes where large inconsistencies were found in the observations. However, at the gauging station with a reliable rating curve, Nyangores, the recorded and geometric discharge-water level relations were almost identical. In conclusion, this calibration method allows reliable simulations of the discharge-water level relation, even in a data poor region.
In addition, this paper analysed the current status of the hydro-meteorological network in the Mara River Basin focusing on the data availability and quality. Moreover, a hydrological model and an improved geometric rating curve equation were developed for this river. All three aspects contribute to improving the assessment of the water resources availability in the Mara River Basin.

6 Recommendations

This paper illustrated that the proposed water level calibration method simulated the discharge-water level relation well for the gauging station where consistent rating curve information was available. It would be interesting to apply this calibration method to other study river basins with different climatic conditions and better data availability. Furthermore, it is recommended to assess the effect of rainfall uncertainties on this calibration method. Moreover, the hydrological model was calibrated on two signatures only. However, it has not been analysed whether these signatures provide sufficient information for calibration. Therefore, the procedures for water level based calibration should be analysed in more detail.

Acknowledgement

This research was part of the MaMaSe project (Mau Mara Serengeti) led by IHE Delft. Station data (discharge, water level and precipitation) was provided by the Water Resource Management Authority (WRMA) in Kenya. Temperature and additional precipitation data was obtained from NOAA online database (Menne et al., 2012).

References

Beven, K. J. (2012). Rainfall-runoff modelling : the primer. Retrieved from http://public.eblib.com/choice/publicfullrecord.aspx?p=822562

Bulygina, N., & Gupta, H. (2011). Correcting the mathematical structure of a hydrological model via Bayesian data assimilation. *Water Resources Research, 47*(5), n/a-n/a. doi:10.1029/2010WR009614

Clarke, R. T. (1999). Uncertainty in the estimation of mean annual flood due to rating-curve indefinition. *Journal of Hydrology, 222*(1-4), 185-190. doi:10.1016/S0022-1694(99)00097-9

Dessu, S. B., ~~& Melesse, A. M. (2012). Modelling the rainfall–runoff process of the Mara River basin using the Soil and Water Assessment Tool. *Hydrological Processes, 26*(26), 4038-4049. doi:10.1002/hyp.9205~~

~~Dessu, S. B.,~~ Melesse, A. M., Bhat, M. G., & McClain, M. E. (2014). Assessment of water resources availability and demand in the Mara River Basin. *CATENA, 115*(0), 104-114. doi:http://dx.doi.org/10.1016/j.catena.2013.11.017

Di Baldassarre, G., & Claps, P. (2011). A hydraulic study on the applicability of flood rating curves. *Hydrology Research, 42*(1), 10-19. doi:10.2166/nh.2010.098

Di Baldassarre, G., Laio, F., & Montanari, A. (2012). Effect of observation errors on the uncertainty of design floods. *Physics and Chemistry of the Earth, Parts A/B/C, 42–44*, 85-90. doi:http://dx.doi.org/10.1016/j.pce.2011.05.001

Di Baldassarre, G., & Montanari, A. (2009). Uncertainty in river discharge observations: a quantitative analysis. *Hydrol. Earth Syst. Sci., 13*(6), 913-921. doi:10.5194/hess-13-913-2009

Domeneghetti, A., Castellarin, A., & Brath, A. (2012). Assessing rating-curve uncertainty and its effects on hydraulic model calibration. *Hydrology and Earth System Sciences, 16*(4), 1191-1202. doi:10.5194/hess-16-1191-2012

~~FAO (Cartographer). (1998). Africover~~

Gao, H., Hrachowitz, M., Fenicia, F., Gharari, S., & Savenije, H. H. G. (2014a). Testing the realism of a topography-driven model (FLEX-Topo) in the nested catchments of the Upper Heihe, China. *Hydrol. Earth Syst. Sci., 18*(5), 1895-1915. doi:10.5194/hess-18-1895-2014

Gao, H., Hrachowitz, M., Schymanski, S. J., Fenicia, F., Sriwongsitanon, N., & Savenije, H. H. G. (2014b). Climate controls how ecosystems size the root zone storage capacity at catchment scale. *Geophysical Research Letters, 41*(22), 2014GL061668. doi:10.1002/2014GL061668

Gharari, S., Hrachowitz, M., Fenicia, F., Gao, H., & Savenije, H. H. G. (2014). Using expert knowledge to increase realism in environmental system models can dramatically reduce the need for calibration. *Hydrol. Earth Syst. Sci., 18*(12), 4839-4859. doi:10.5194/hess-18-4839-2014

GLOWS-FIU. (2012). Environmental Flow Recommendation for the Mara River, Kenya and Tanzania: Global Water for Sustainability program (GLOWS). Miami, FL.

Guerrero, J. L., Westerberg, I. K., Halldin, S., Xu, C. Y., & Lundin, L. C. (2012). Temporal variability in stage-discharge relationships. *Journal of Hydrology, 446-447*, 90-102. doi:10.1016/j.jhydrol.2012.04.031

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., et al. (2013). A decade of Predictions in Ungauged Basins (PUB)—a review. *Hydrological Sciences Journal, 58*(6), 1198-1255. doi:10.1080/02626667.2013.803183

Jalbert, J., Mathevet, T., & Favre, A. C. (2011). Temporal uncertainty estimation of discharges from rating curves using a variographic analysis. *Journal of Hydrology, 397*(1-2), 83-92. doi:10.1016/j.jhydrol.2010.11.031

~~Jiang, L., Schneider, R., Andersen, O. B., & Bauer-Gottwein, P. (2017). CryoSat-2 altimetry applications over rivers and lakes. *Water (Switzerland), 9*(3). doi:10.3390/w9030211~~

Karamuz, E., Osuch, M., & Romanowicz, R. J. (2016) The influence of rating curve uncertainty on flow conditions in the River Vistula in Warsaw. *Vol. none. GeoPlanet: Earth and Planetary Sciences* (pp. 153-166).

Kuczera, G. (1996). Correlated rating curve error in flood frequency inference. *Water Resources Research, 32*(7), 2119-2127. doi:10.1029/96WR00804

Mati, B. M., Mutie, S., Gadain, H., Home, P., & Mtalo, F. (2008). Impacts of land-use/cover changes on the hydrology of the transboundary Mara River, Kenya/Tanzania. *Lakes & Reservoirs: Research and Management, 13*(2), 169-177.

McClain, M. E., Subalusky, A. L., Anderson, E. P., Dessu, S. B., Melesse, A. M., Ndomba, P. M., Mtamba, J. O. D., Tamatamah, R. A., & Mligo, C. (2013). Comparing flow regime, channel hydraulics and biological communities to infer flow-ecology relationships in the Mara River of Kenya and Tanzania. *Hydrological Sciences Journal, 59 (3-4)*, 1-19. doi:10.1080/02626667.2013.853121

McMillan, H., Freer, J., Pappenberger, F., Krueger, T., & Clark, M. (2010). Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions. *Hydrological Processes, 24*(10), 1270-1284. doi:10.1002/hyp.7587

McMillan, H. K., & Westerberg, I. K. (2015). Rating curve estimation under epistemic uncertainty. *Hydrological Processes, 29*(7), 1873-1882. doi:10.1002/hyp.10419

Menne, M. J., Durre, I., Korzeniewski, B., McNeal, S., Thomas, K., Yin, X., Anthony, S., Ray, R., Vose, R. S., Gleason, B. E., & Houston, T. G. (2012). Global Historical Climatology Network - Daily (GHCN-Daily), Version 3.12. doi:http://doi.org/10.7289/V5D21VHZ

Morlot, T., Perret, C., Favre, A. C., & Jalbert, J. (2014). Dynamic rating curve assessment for hydrometric stations and computation of the associated uncertainties: Quality and station management indicators. *Journal of Hydrology, 517*, 173-186. doi:10.1016/j.jhydrol.2014.05.007

~~Mwangi, H. M., Julich, S., Patil, S. D., McDonald, M. A., & Feger, K. H. (2016). Modelling the impact of agroforestry on hydrology of Mara River Basin in East Africa. *Hydrological Processes, 30*(18), 3139-3155. doi:10.1002/hyp.10852~~

Pelletier, P. M. (1988). Uncertainties in the single determination of river discharge: a literature review. *Canadian journal of civil engineering, 15*(5), 834-850. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-0024163837&partnerID=40&md5=7f974b9677db876af486ab1cf405ef00

Petersen-Øverleir, A. (2006). Modelling stage-discharge relationships affected by hysteresis using the Jones formula and nonlinear regression. *Hydrological Sciences Journal, 51*(3), 365-388. doi:10.1623/hysj.51.3.365

Rey, A., de Koning, D., Rongen, G., Merks, J., van der Meijs, R., & de Vries, S. (2015). Water in the Mara Basin: Pioneer project for the MaMaSe project. *Unpublished MSc project report, Delft University of Technology*.

Savenije, H. H. G. (2010). Topography driven conceptual modelling (FLEX-Topo). *Hydrol. Earth Syst. Sci., 14*, 2681-2692.

Seibert, J., & Beven, K. J. (2009). Gauging the ungauged basin: how many discharge measurements are needed? *Hydrol. Earth Syst. Sci., 13*(6), 883-892. doi:10.5194/hess-13-883-2009

Seibert, J., & Vis, M. J. P. (2016). How informative are stream level observations in different geographic regions? *Hydrological Processes, 30*(14), 2498-2508. doi:10.1002/hyp.10887

Sellami, H., La Jeunesse, I., Benabdallah, S., & Vanclooster, M. (2013). Parameter and rating curve uncertainty propagation analysis of the SWAT model for two small Mediterranean catchments. *Hydrological Sciences Journal, 58*(8), 1635-1657. doi:10.1080/02626667.2013.837222

Shahin, M. (2002). *Hydrology and Water Resources of Africa* (Vol. 41): Springer Netherlands.

Sikorska, A. E., Scheidegger, A., Banasik, K., & Rieckermann, J. (2013). Considering rating curve uncertainty in water level predictions. *Hydrology and Earth System Sciences, 17*(11), 4415-4427. doi:10.5194/hess-17-4415-2013

Stoop, B. M. (2017). *Morphology of the Mara River: Assessment of the long term morphology and the effect on the riverine physical habitat.* Delft University of Technology.

Thyer, M., Renard, B., Kavetski, D., Kuczera, G., & Clark, M. (2011). Improving hydrological model predictions by incorporating rating curve uncertainty (pp. Proceedings of the 34th IAHR World Congress, 2011: pp.1546-1553). Australia: Engineers Australia.

Tomkins, K. M. (2014). Uncertainty in streamflow rating curves: Methods, controls and consequences. *Hydrological Processes, 28*(3), 464-481. doi:10.1002/hyp.9567

U.S. Geological Survey. (2014). Digital Elevation Map. Retrieved from www.earthexplorer.usgs.gov

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., & Sorooshian, S. (2003). Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Water Resources Research, 39*(8), n/a-n/a. doi:10.1029/2002WR001746

Wang-Erlandsson, L., Bastiaanssen, W. G. M., Gao, H., Jägermeyr, J., Senay, G. B., van Dijk, A. I. J. M., Guerschman, J. P., Keys, P. W., Gordon, L. J., & Savenije, H. H. G. (2016). Global root zone storage capacity from satellite-based evaporation. *Hydrol. Earth Syst. Sci., 20*(4), 1459-1481. doi:10.5194/hess-20-1459-2016

Westerberg, I. K., Guerrero, J. L., Younger, P. M., Beven, K. J., Seibert, J., Halldin, S., Freer, J. E., & Xu, C. Y. (2011). Calibration of hydrological models using flow-duration curves. *Hydrol. Earth Syst. Sci., 15*(7), 2205-2227. doi:10.5194/hess-15-2205-2011

WMO. (2008). Guide to Hydrological Practices; Volume I Hydrology – From Measurement to Hydrological Information.

Yadav, M., Wagener, T., & Gupta, H. (2007). Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. *Advances in Water Resources, 30*(8), 1756-1774. doi:https://doi.org/10.1016/j.advwatres.2007.01.005