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 1 and Table 2 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 1: 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 2: 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.

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

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.