Anonymous Referee #1

I carefully read the paper titled "Using NDII pattern for a semi-distributed rainfall-runoff model in tropical nested catchments", by Sriwongsitanon et al. This paper completes a previous paper by Sriwongsitanon et al., 2016, showing that NDII is correlated with the dynamics of the water level in the root zone reservoir of the FLEXL lumped model, suggesting that NDII could be used as a proxy of the water content on the one hand, and could improve the model performance on the other hand.

Here, the purpose is to assess the efficiency of FLEX-derived semi-distributed models based on distribution of the travel times (FLEX-SD) and on the distribution of both the travel times and the water storage in the root zone, according NDII spatial variability for the latter (FLEX-SD-NDII). The SD models were tested in several sub-catchments of a nearly 6,000 km^2 catchment in Thailand, at a daily time step. The semi-distributed model URBS was also compared to FLEX-SD and FLEX-SD-NDII. The method consists in calibrating all the models at the outlet of the larger catchment P.1, then in performing each model without calibration at the outlet of 5 sub-catchments. The lumped FLEX model was however calibrated in each sub-catchment in order to give a reference of goodness in each sub-catchment. Four error functions were considered: Nash-Sutcliffe Efficiency (NSE), Kling-Gupta Efficiencies (KGE) for high flows, low flows, and the flow duration.

Finally, the SD-FLEX models were found to perform correctly in the sub-catchments, without recalibration. NSE and KGE mean values over all the sub-catchments were respectively 0.74, 0.79, 0.66 and 0.84 for FLEX-SD, 0.74, 0.81, 0.66, 0.87 for FLEXSD-NDII. In addition, the relationships between the 8-days NDII values and the water levels in the root zone reservoir (Su) are better for FLEX-SD-NDII than for FLEX-SD. So that FLEX-SD-NDII does not perform better than FLEX-SD in terms of accuracy of runoff estimates, but as the authors say, FLEX-SD-NDII "has gained realism and hence predictive power". It is also shown that the relationships between Su and NDII are poor during dry seasons, highlighting the limitation of NDII to be considered as a powerful proxy of the root zone water content. Auxiliary other results are that first, the FLEX-SD models perform better than the SD URBS model, and second, that the Su levels in the root zone reservoir are better correlated with the SWI index than with the NDII index.

My opinion is that the study shows interesting results which are worth to be published in the review. It could help in applying widely-used models such as FLEX or URBS, and show the interest of semi-distributed models for reproducing internal flows within a given catchment. However, there is sometimes a lack of clarity, and the paper would gain in my sense by giving some more details. Although the paper is concise, well written and well structured, I feel that at least the general strategy and the main steps of the method should be reinforced and summarized earlier in the paper, in order that the reader could have a complete view of the construction as soon as the end of the introduction. In addition, some more information could be presented on model calibration and on the strategy for the comparison of the models. The gain obtained by the spatial variability of the rainfall itself did not seem to be discussed in the paper, although it could be a major benefit.

Answer: We would like to thank referee#1 for this very detailed and constructive review. In the detailed comments below, we address all his/her concerns

Introduction

I would appreciate that the introduction would give a more comprehensive review of the NDII interest for hydrological modelling. What were the main results which were obtained by Sriwongsitanon et al., 2016? Are there more studies dealing with using NDII in hydrological modelling? How does this paper complete the previous ones?

Answer: The referee #1 is right in that more background could be given on other authors who used the NDII to interpret root zone moisture or to use the NDII as a constraint for root zone soil moisture dynamics. The following texts will be included in the introduction of the revised manuscript, starting from line 25 of page 2.

"This corresponds to the study carried out by Castelli et al. (2019) who found reasonable correlations between Landsat 7 NDII values and measured root-zone soil moisture contents of rainfed olive trees growing in the arid regions of south eastern Tunisia, allowing the use of NDII as an indicator for soil water content.

Mao and Liu (2019) developed the Water And ecosYstem Simulator (WAYS) which is a distributed model based on FLEXL to simulate discharge as well as root zone water storage (RZWS) at a global scale accross 10 major basins, comprising the Congo, Nile, Niger, Yangtze, Ganges, Parana, Amazon, Mississippi, Murray-Darling, and Mekong. The WAYS model appeared to be capable to simulating both evaporation and discharge in this wide range of river basins. It also was able to simulate RZWS in most of the regions through comparison with NDII (with correlation ranging from 0.95 to 0.71 with an average of 0.88). This is with the exception of some basins comprising the Amazon, Murray-Darling and Mississippi (with correlation ranging from 0.55 to 0.68) which have a large percentage of relatively moist areas with low moisture stress, where the NDII is known to have difficulty reflecting RZWS dynamics.

Regarding the use of the Soil Water Index (SWI) as a proxy for root zone soil moisture, Paulik et al. (2014) found reasonable correlations between in-situ soil moisture data from 664 stations - available through the international Soil Moisture Network (ISMN) - and the SWI produced from ASCAT SSM estimates. The average of Pearson correlation coefficients was shown to be 0.54 with 64.4% of all time series greater than 0.5. SWI could be another index to indicate soil moisture of the basin with or without moisture stress situation."

In addition, the authors could summarize at the end of the introduction the main steps of the strategy they intend to develop in order to prove the interest of the SD models, for giving the reader a global view of the structure of the paper. I feel that I was discovering the method step by step while going on and reading the different sections of the paper.

Answer: This is a good suggestion. In fact, the referee gave a very clear summary of the approach in his/her comment. We have revised the introduction and will add the following paragraph on the approach at the end of the Introduction:

The main steps of the approach followed in the following sections are the following:

1. *To test the effect of runoff timing in a catchment with multiple sub-catchments, the travel times to the outfall of each individual sub-catchment are computed on the basis of topographical indicators and the routing of the discharge from the sub-catchment outfall to stations further downstream are computed using the Muskingum method. These time lags are then applied both in the FLEX-SD model system and in the well-established URBS model, for the purpose of comparison. These two semi-distributed models only account for timing, but not for the distribution of the moisture storage capacity, a crucial parameter in runoff generation.*
2. *Subsequently, the effect of distribution of the root zone moisture storage is studied in the FLEX-SD model, making use of the distribution pattern of the maximum seasonal range of NDII values, a proxy for root zone moisture storage.*
3. *Finally, as a validation of the model and to check if the models are capable of representing the internal states, the simulated root zone moisture storage is compared to the independent data set of the Soil Wetness Index (SWI).*

Rainfall-Runoff data

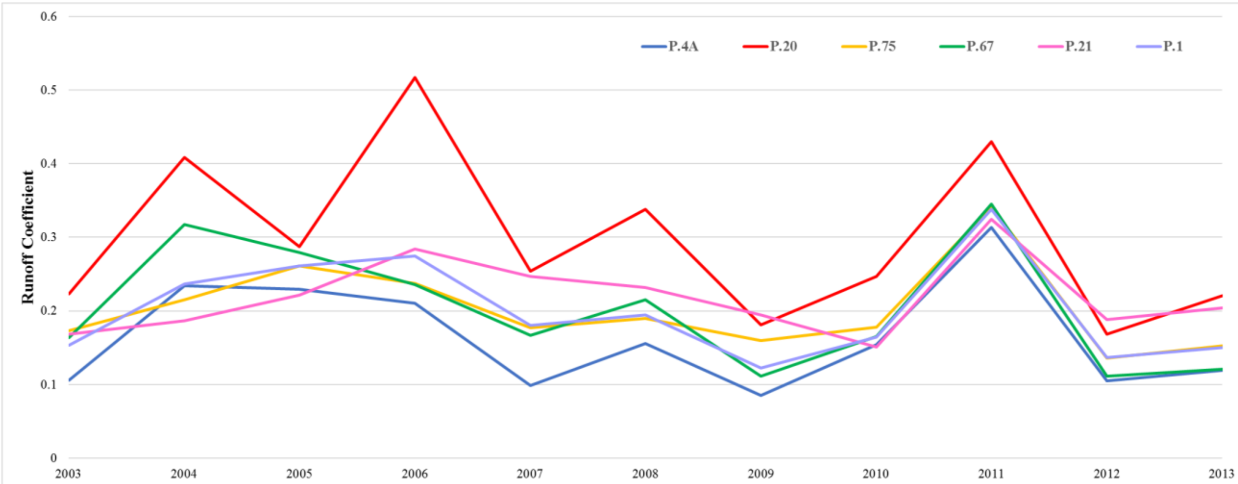
4, 2-3: The Royal Irrigation Department (RID) operates 7 daily runoff stations in the study area between 2003 and 2013 as shown in Fig. 1. Catchment P.56A was rejected from the study because it is located upstream of Mae Ngat reservoir. Why is it a problem?

Answer: It is possible to calibrate FLEXL model at P.56A. However, it is not possible to determine the runoff upstream of Mae Ngat reservoir using any semi-distributed rainfall-runoff model at P.1 because the regulated reservoir is within the network. Therefore, we have to use the outflow from Mae Ngat Dam as input for the calibrations of FLEX-SD and FLEX-SD-NDII at P.1.

4, 5-6: The data have been checked for their accuracy by comparing them with average rainfall data covering their catchment areas at the same periods. What were the test and the results of the test?

Answer: The below sentences are included in Section 2.3 of the revised manuscript.

Runoff data have been checked for their accuracy by comparing the annual runoff coefficient between all stations. The comparison revealed that the runoff coefficients at P.20 in 2006 and 2011 are overestimated, while the runoff coefficient at P.21 in 2004 is underestimated and in 2007 and 2009 are overestimated due to incorrect rating curves (see the below figure). These inaccurate data would affect the results of model calibration, and may be reasons why these stations do not always show good performance.



Runoff Station	P.20	P.75	P.4A	P.67	P.21	P.1
Area (km^2)	1,309	3,029	1,954	5,333	516	6,142
Altitude range (m)	993	1,035	686	1,058	581	1,067
Length main channel (km)	89	126	143	155	52	185
Average channel slope	0.006	0.005	0.004	0.004	0.01	0.004
Average rainfall (mm/yr)	1,250	1,243	1,199	1,225	1,230	1,229
Rainfall Range (mm/yr)	1,050 – 1,640	1,074 – 1,643	984 – 1,449	1,051 – 1,570	922 – 1,606	1,037 – 1,565
Average runoff (mm/yr)	376	256	206	257	275	254
Runoff Range (mm/yr)	186.2 – 672.4	145.7 – 478.8	91.5 – 453.9	120.0 – 494.1	154.9 - 521.0	133.1 – 492.7
Irrigated Area (%)	15.7	18.1	9.4	15.1	17.4	15
% Runoff Range	16.9 - 51.7	13.6 - 34.1	8.5 - 31.3	11.1 - 34.5	15.1 - 32.4	12.3 - 33.8
% Runoff Average	29.8	20.2	16.5	20.3	21.8	20.1

Performing/calibrating the SD-models

Figure 1 shows the location of the 6 gauging stations, as well as the number and the boundaries of the 10 sub-catchments which were designed upstream of the gauging stations. So, 2 sub-catchments seem to have been considered upstream the gauging station P.4A, 2 upstream of P.20, 4 upstream of P.75, 1 upstream of P.21 … This should be mentioned explicitly in the paper. Did the SD models actually run over several sub-catchments within a given sub-catchment: for example, did the SD models run over 2 sub-catchments upstream from P.4A, over 4 sub-catchments upstream from P.75, etc… In this case, how were selected the number and the area of the sub-catchments upstream of each gauging station? How is the model dependent of the number of sub-catchments, and could this dependence alter the interpretation of the performance of the SD-models?

Answer: The semi-distributed rainfall-runoff model can be used to estimate runoff for a larger number of sub-catchments. However, a larger number of sub-catchments will only result in a better performance if the rainfall amounts in these sub-catchments is sufficiently variable. In this case, rainfall amounts in these 10 sub-catchments only vary between 1,171 and 1,369 mm/yr, which are not much different. We did test a higher resolution of sub-catchments for FLEX-SD and FLEX-SD-NDII. After dividing the catchment of P.1 into 32 sub-catchments, the average NSE values for the 7 stations were not significantly different from the results obtained when using only 10 sub-catchments. This is probably due to the relatively low variability of rainfall between the sub-catchments.

We will modify the original manuscript between the lines 18 and 19 of page 5 *"URBS is categorized as a semi-distributed rainfall-runoff model that can provide runoff estimates not only at the calibrated station but also at any required location upstream"* into:

"URBS is categorized as a semi-distributed rainfall-runoff model that can provide runoff estimates not only at the calibrated station ***but also at the outlet of every sub-catchment at any required location upstream***."

3, 29-31: These data have been validated for their accuracy on monthly basis using double mass curve and some inaccurate data were removed from the time series before spatially averaging using an inverse distance square (IDS) to be applied as the forcing data of URBS, FLEXL, and FLEX-SD.

I suppose that the mean areal rainfalls were calculated for each sub-catchment. That means that the gain of the SD-models could be due to taking into account the rainfall variability. Could the authors give some information about this variability? Furthermore, could they assess the effect of the distribution of the rainfall on the gains that bring the SD-models? For example, what would be the results of a semi-distributed FLEX model using only the spatial variability of the rainfall, and keeping the same TlagF and TlagS for all the sub-catchments.

Answer: The information of rainfall distribution of each sub-catchment is presented in the following table. To test the effect of merely rainfall distribution, we rerun the models using the same TlagF and TlagS for all sub-catchments. A recalibration with these constant time lags produced similar performance indicators on the runoff, some indicators were slightly better, others were slightly worse. The fact that there was not much difference is mainly due to the low spatial variability of the rainfall in this part of Thailand.

P.1	**P.4A**	**P.20**	**P.21**	**P.67**	**P.75**	**Sub-ID**	**Outlet**	**Area**		

									Average rainfall (mm/yr)	Areal Average rainfall (mm/yr)
[redacted]	[redacted]			[redacted]		Sub1	-	1,256	1,171.06	-
[redacted]	[redacted]			[redacted]		Sub2	P.4A	698	1,249.05	1,198.93
						Sub3	-	1,280	-	***Mae Ngat Reservoir***
[redacted]		[redacted]		[redacted]	[redacted]	Sub4	-	622	1,247.85	-
[redacted]		[redacted]		[redacted]	[redacted]	Sub5	P.20	687	1,252.76	1,250.42
[redacted]				[redacted]	[redacted]	Sub6	P.75	440	1,368.98	1,280.24
[redacted]				[redacted]		Sub7	-	155	1,234.02	-
[redacted]				[redacted]		Sub8	P.67	195	1,200.79	1,235.44
[redacted]						Sub9	-	57	1,175.76	-
[redacted]			[redacted]			Sub10	P.21	516	1,229.58	1,229.58
[redacted]						Sub11	P.1	237	1,322.74	1,238.38

Modelling results

9, 15-16: Table 4 surprisingly shows that FLEX-SD and FLEX-SD-NDII provide better or equally good runoff estimates at P.1, P.67 and P.75 stations located along the main Ping River, compared to those provided by the calibrated FLEXL model.

This does not seem so surprising, at least for the P.1 gauging station, where all the models have been calibrated. In addition, the SD-models count with a higher number of parameters, which could allow higher scores in reproducing runoff characteristics when calibrating the models.

Answer: FLEXL was separately calibrated at each runoff station. FLEX-SD and FLEX-SD-NDII were calibrated only at P.1, however, they can provide runoff estimates at all upstream stations. There is a significant difference in the time required to prepare the data and for running the models. The individual FLEXL models are far more time consuming. FLEX-SD and FLEX-SD-NDII have a great advantage over the individual FLEXL models, especially since the accuracy of runoff estimation between these two types of models is not much different or even better in some stations. Moreover, runoff estimates can be acquired at all 10 sub-catchments using FLEX-SD and FLEX-SD-NDII, while FLEXL can only provide runoff estimates at the calibrated stations.

The correlation between NDII and the maximal storage capacity in the root zone appears to be poor in the wet season. The R^2 values were around 0.30-0.40 with the calibrated FLEX L model in the different sub-catchments. In the previous paper published in 2016, such values were higher, around 0.4-0.5 even more. Could the authors comment on this difference?

Answer:

The period of model calibration carried out by Sriwongsitanon et al. (2016) for P.4A and P.21 was between 1995 and 2009, and for P.20 between 1995 and 2011. In this paper, the calibration period of these 3 stations was 2003-2013.

Table 4: bold values are missing for P.1.

Answer:

We correct this in the revised manuscript.

Conclusion

11, 21-22: "The lag time from storm to peak flow and the lag time of recharge from the root zone to the groundwater have been distributed among sub-catchments using their catchment areas and reach lengths."

I think that the lag times are not from storm to peak flow, but from storm to runoff generation

Answer: The referee is correct that this formulation was incorrect. It should read: "The lag times from rainfall to surface runoff and to groundwater recharge (see Figure 2) have been distributed among sub-catchments using their catchment areas and reach lengths."

The correlation between NDII and the maximal storage capacity in the root zone appears to be poor in the wet season. The authors recognize that NDII cannot be considered as a reliable proxy of the soil water content. This should be recalled in the conclusion.

Answer: We will include the following sentences in the conclusion of the revised manuscript.

"This study collaborates with the study carried out by Sriwongsitanon el al. (2016) who concluded that NDII can be used as a proxy for catchment-scale root zone moisture deficit when plants are exposed to water stress. However, during the wet season when soil moisture is replenished as a result of rainfall, NDII values are no longer well correlated with soil moisture. However, the – partly model-based – SWI proved to be a reliable index to estimate soil moisture both under water stressed and wet conditions."

In conclusion, due to the scientific interest of the paper and the good value of the study, I recommend that the paper would be published in the review, after modifications which can be considered as minor revision.

Answer: Thank you for your kind consideration and valuable comments and suggestions.