Hydrol. Earth Syst. Sci. Discuss., 11, C6533–C6537, 2015
www.hydrol-earth-syst-sci-discuss.net/11/C6533/2015/

Hydrology and Earth System Sciences
Open Access
Discussions

Replies to comments by Anonymous Reviewer #II

[*italics* for comments and regular for reply]

Interactive comment on "**Investigation of variable threshold level approaches for hydrological drought identification**" by B. S. Beyene et al.

Received and published: 20 February 2015

1) General Comments:

This paper investigates different threshold level methods and their suitability in characterizing droughts from a range of catchments in Europe. The work is well-referenced and finds that the choice of threshold level method may not be as important as has been previously considered. However, there are some significant issues in the interpretation of the information presented in Table 1. The number and scale of the misinterpretations are such that I urge the authors to thoroughly review the data and interpretation presented in this paper before re-submission. In addition, the content of the discussion section needs to be addressed, and there is an over-reliance on one specific reference. Considerable work is required on major revisions to address these issues before the paper is ready for potential publication.

We found the general comments very much constructive. We have thoroughly reviewed our data on Table 1 of the discussion paper and appropriate corrections are done with the data interpretations. Besides, we revised the discussion part of the paper as shown in the submitted manuscript **[from page 14 line 8 to page 15 line 29]**.

There are some misinterpretations derived from the table with results (Table 1). We understand that it is a big concern and we made the necessary corrections in the revised manuscript. In addition, we included an additional graph displaying the frequency of how often highest and lowest values of drought characteristics were identified. However, the general conclusion that frequent meteorological droughts give rise to fewer soil moisture, groundwater and runoff droughts holds. It is hypothesized that drought numbers should decrease, mean drought duration should increase and drought severity should decrease moving from meteorological drought through soil moisture drought to hydrological drought. In our finding, the hypothesis is met in almost all the four approaches for the five catchments. The Results for of Narsjo catchment are presented in the following figures.

We used here the Narsjø as an example catchment as it has the longest record time and thus we expect that the results derived from such a prolonged data are most reliable. As can be clearly seen from Figure RII_1, the most frequent meteorological droughts propagate into prolonged, less severe soil moisture and groundwater droughts, which are later followed by longer sustaining, and more severe discharge droughts.

In addition to examining our hypothesis (see above), we also performed a rigorous drought propagation analysis in terms of the sensitivity of drought characteristics (number of droughts, mean duration and deficit volume) to the choice of the threshold calculation in each catchment. We applied the proposed threshold calculation methods to hydrometeorological variables derived from the five selected catchments. We then inter-compared how magnitudes of these characteristics in each catchment for the different look like under these methods.

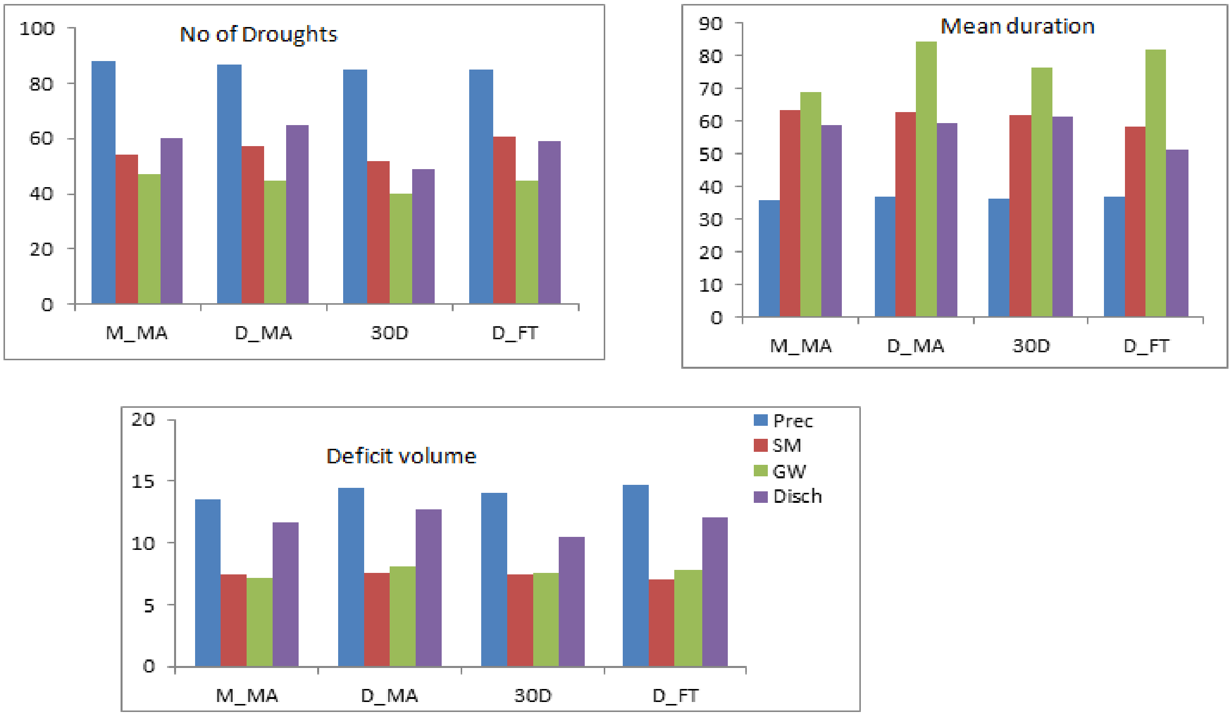

Fig RII_1: Graphical representation of results displayed on Table 1 of the manuscript; showing the number of droughts, mean duration and deficit volume quantified using the proposed four approachesfor the Narsjø catchment. Similar results can be reproduced using data on Table 1 of the manuscript.

Please refer to detail responses in sections 2A-2C (below).

2) Specific Comments:

A) Interpretation of data

There are a number of statements in Section 4.1 that are not supported by the results in Table 1. These include:

~~"the M_MA threshold method has given the least number of droughts in most catchments (Table 1)" (Page 12776, Lines 2-3)

The only rows in Table 1 for which M_MA produces the least number of droughts (n) of all of the four threshold methods is for GW in Nedožery and GW and Q for Guadiana (there are some other rows for which M_MA is tied for the lowest, but not enough to validate the statement above). In the case of the Narsjø, for P and GW the M_MA method actually identifies more droughts than any other method.

More specifically, we made inter-comparisons of 20 cases (5 catchments and 4 variables),which the reviewer referred as rows, and counted the minimum and maximum number of records. For example, when the same threshold level is applied to all catchments, we expect four cases (standing for precipitation, soil moisture, groundwater and discharge) for each catchment and the same for five different catchments (5 catchments and 4 variables). We have then counted how often (x out of 20) a particular identification method has the lowest number of drought, lowest mean duration, lowest deficit volume (i.e. minimum). This has been repeated for the highest number (i.e. maximum). In some circumstance, these values may be counted up to four times because two or more threshold levels may have the same lowest or highest value. For example, the number of drought for Nedožery catchment is 43 when using three threshold methods (M_MA, D_MA and 30D) and 44 (with D_FF). In such circumstances, the number of cases was assumed to be more as we counted these lowest or highest values independently for comparison purposes.

Table RII_1: The table displays the number of times in 20 (or more) caseslowest and highest value of magnitudes of drought characteristics as identified by the four approaches.

		No. of drought	Mean duration	Deficit volume
Minimum (Lowest) values	M_MA	7	6	10
	D_MA	2	1	2
	30D	16	4	4
	D_FF	1	9	4
Total		26	20	20
Maximum (highest) values	M_MA	2	3	2
	D_MA	4	8	9
	30D	1	4	1
	D_FF	15	5	8
Total		22	20	20

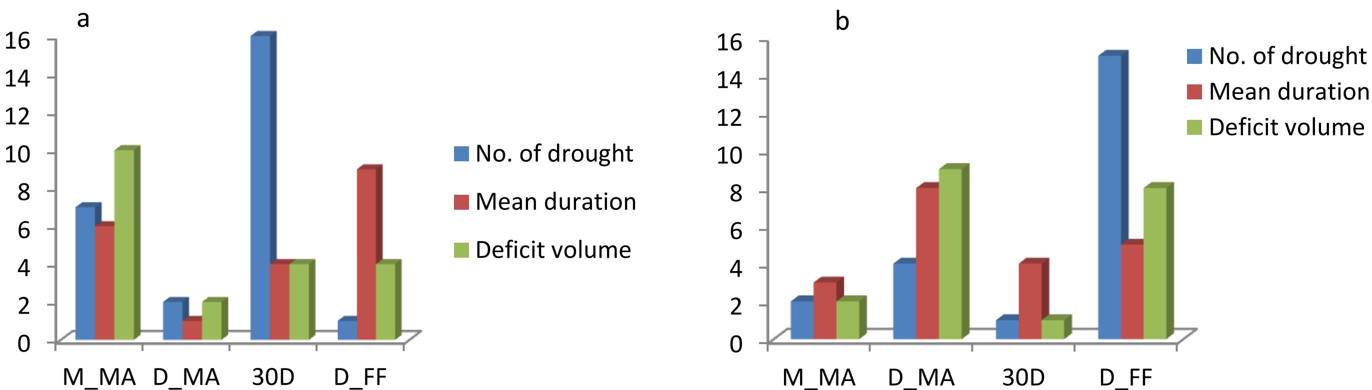

Fig RII_2: Histogram showing the numbers of times a) minimum values were calculated, b) Maximum values were calculated.

There are not physical arguments to choose whether the minimum or maximum values are appropriate dimensions quantities to indicate relative performance of the threshold level approaches. In this context, maximum values imply the maximum number of droughts, and longest mean durations and vice versa. Among the threshold level approaches, D_MA and D_FF threshold methods have given the least number of droughts, in most catchments (Table RII_1). However, the most comprehensive argument for general conclusion should be made based on how frequent (large number of) meteorological droughts propagate into few (least number of) and prolonged soil moisture and hydrological droughts.

In the particular case mentioned by the anonymous reviewer #II mentioned in 2A (above), it can be seen that least number of droughts were most often identified when using 30D threshold level approach (Table RII_1 and FigureRII_2 a). The number of drought was identified as minimum 16 times out of in 20 possible circumstances. This may be due to the relative stability of the threshold level determined by 30D threshold approach. We addressed the above issues in the revised manuscript **[page 14 lines 21-23]**.

~~"the method [presumably still referring to M_MA] generates longer mean drought duration" (Page 12776, Line 10)

The section preceding this line is a little difficult to follow, but if I have understood correctly, this sentence refers to M_MA producing longer mean drought duration. If this is the case, Table 1 does not support this claim. D for M_MA is less than D for D_MA in 13/20 rows, and is the same for another 4 rows. In addition, D for M_MA is less than D for 30D in 11/20 rows, and is the same for another 2 rows. For three quarters of the rows, D_MA or 30D (or both) actually produce longer mean durations than M_MA.

As can be referred from Table RII_1 and its visual presentation (FigureRII_2 b), shows we can see that longest mean durations were most often identified when using D_MA threshold level approach. This implies that mean drought duration is often longest when using this approach than any of the other three approaches. In this study, we have identified that in 8 out of 20 circumstances, the mean duration was longest when it is calculated using D_MA followed by 5 circumstances by D_FF. We addressed the issue in the revised manuscript **[page 15 line 9-13]**.

~~"the computed mean drought duration for these hydrometeorological variables longer than those computed with other methods for the rest of the catchments"(Page 12776, Lines 23-25)

I assume that the sentence refers to D_FF as it appears in the section on D_FF. There are only 4 rows (out of 20) in Table 1 for which D for D_FF is larger than for any other threshold method (another 2 rows are tied largest with another method). In 7 out of 20rows, D for D_FF is actually the lowest D value compared to the other three methods. There is also inconsistency between this statement and the one quoted above, with one claiming D_FF computes longer mean duration and the other claiming M_MA does the same.

The statement on Page 12776, Lines 23-25 explains the mean duration calculated by D_FF; as you assumed. In this context, "*these hydrometeorological variables*" stands for precipitation and discharge series for the aforementioned catchments, where the threshold level determined using D_FF becomes a fixed threshold level. The example drought event mentioned following the statement is presented to support our conclusion that the fixed threshold gives more extended droughts than one would expect when compared to the mean duration as calculated by the other threshold level approaches. We understand that the addition of the phrase "*for the rest of the catchments*" leads to a misunderstanding. Therefore, we removed and restated the discussion on the revised manuscript **[page 16 lines 17-26]**.

~~"Mean calculated deficit volume is often higher when using the D_FF and D_MA threshold methods than using the M_MA and D30 threshold methods" (Page 1277,Lines 1-3)

The values for V or H are generally very similar across all four threshold methods; where they are different, there is only 1-5 mm difference between the highest and lowest of the four. There is so little variation, it might not even be worth mentioning (except to say that there is so little variation). With reference to the phrase above, I can only count 4 rows out of 20 where V or H for both D_FF and D_MA are higher than V or H for both M_MA and 30D.

The statement on Page 12777, Lines 1-3 discusses the mean deficit volume from all the five catchments as calculated using four threshold level approaches. We understand that we should clearly state that this is not connected to the previous cases and its example (stated from Page 12776, Line 25 to Page 12777, Line 1). Therefore, we revised the statement as "*The mean deficit volume calculated from variables in the five catchments is often ....*". As can be referred from Table RII_I and FigureRII_2 b, the number of times the maximum deficit volume is computed is higher when using D_MA (9 out of 20 circumstances) and D_FF (8 out of 20 circumstances). These results are added in the revised manuscript **[page 15 lines 9-15]**.

The exact numbers representing the magnitude of the deficit may sound insignificant. However, the cumulative deficit (in soil moisture or groundwater level) over a catchment area or over an extended period (for precipitation and discharge) could have a significant implication for the impacts and in the subsequent analysis of hydrological processes; particularly the drought propagation analysis. In addition, the numbers presented in Table 1 of the manuscript are all rounded to the nearest whole number. We understand that this has brought some misinterpretation in inter-comparing the relative performance of the threshold approaches for deficit volume calculation. Therefore, we presented the numbers in Tables 1 of the manuscript in 2 decimal places such that even very small differences could be taken into account. These corrections are added in the revised manuscript **[page29 Table 1]**.

B) Content of discussion section

More than half of the discussion section (Page 12778, Line 22 to Page 12780, Line 8) focuses on different options for threshold level approaches and how these have been applied in other research. Although well-referenced, this part reads like a literature review of threshold level approaches, and could probably be moved earlier in the paper into the Introduction. The first detailed discussion specifically on this paper starts at Page 12780, Line 9 and runs until Page 12781, Line 2 (less than a full page). Perhaps moving the more literature-based first half of the Discussion into the Introduction will allow more thorough interpretation of the results in this paper.

We accepted you comment and we revised the discussion section as of your suggestion. We adapted discussion from Page 12778, Line 22 to Page 12780, Line 8 of the discussion paper **[from page 18 line 26 to page 19 line 8 of revised manuscript]** into the Introduction section **[from page 4 line 2 to page 6 line 12]**.

C) Over-reliance on specific reference

The work of Van Loon & Van Lanen (2012) is referenced twelve times in this paper. Figures 1 and 3 of this paper are reproduced directly, and the modelling work underpinning the research in this paper was also conducted for the earlier work. Given the common model outputs underpinning both studies, it is surprising to note that the data in the left-hand columns of Table 1 are marginally different (but almost identical) to those given in Table 3 of Van Loon & Van Lanen (2012), if I have understood correctly the methodologies of the two papers (perhaps the authors could clarify this). I think Equations 8, 9 and 10 in this paper also appear in the earlier work, although Equation 8 may be incorrectly written (Di+1 has been written twice). Whilst the analysis of different thresholds is novel in this paper, there is perhaps an over-reliance on Van Loon & Van Lanen (2012) in general.

Van Loon & Van Lanen (2012) discussed the threshold level approach for the purpose of drought identification and typology. Similar threshold level approach has been discussed in many papers (see references in manuscript). However, we found that Van Loon & Van Lanen (2012) are more explanatory and have exhaustively discussed how magnitude and severity of meteorological drought is translated to drought in soil moisture, groundwater and discharge. We took their approach as a foundation to this work and used their model outputs as discussed on Page12770, Lines 9-23. We have clearly stated that we used their simulation results. For these reasons, we used their paper more frequently that may have resulted in over reliance. This main reason is incorporated in the revised edition **[page 7 line 21-23]**.

Please note, that there could be some discrepancies between our study and the work of Van Loon & Van Lanen (2012) because of the following reasons. In both works, the 30 days centered moving average has been used. Normally these 30 days are separated into to 15 days of equal duration. Here we have two options to take the centered day; the last day of the first half or the first day of the second half. Equation (2) (first option) was applied in this paper. In their paper, Van Loon and Van Lanen (2012) have chosen the second option given as:

$Thr_{\_MAM}(m) = average(Q_{\_MAM}(m-15):Q_{\_MAM}(m+14))$

Whereas in this research work, we used

$Thr_{\_MAM}(m) = average(Q_{\_MAM}(m-14):Q_{\_MAM}(m+15))$

The choice for one of these two options cannot be justified as both of them represent a 30-days centered moving average, and therefore, the decision is highly subjective. However, the implication is that we moved the annual threshold level to the left by one day compared to the threshold level computed by Van Loon & Van Lanen (2012). When this modified (displaced) threshold level is compared against long-term time series of a variable, it would normally result in a slightly higher or lower number of droughts, mean duration or deficit volume. We would like to mention to the reviewer/editor that such discrepancies may be attributed to the fact that displacing the threshold level to the left. The modified (displaced) threshold level:

a) may have reproduced few drought events that were pooled together in the earlier studies, but separated in this study as the choice resulted in increased inter-event time, or the other way.

b) may have reproduced few events with increased duration that met the minimum drought duration to be identified as droughts. For example, the events with duration of 14 days may have now 15 days with the modified threshold level and, therefore, increases the number of droughts. The opposite could also happen in such circumstances. This is likely to happen when longer period is taken into account.

3) Technical Corrections:

Page 12766, Lines 25-26: Floods can have a "very complex development pattern" too

The two phenomena (drought and flooding) have some complexity. However, the degree of complexity is not comparable as a result of catchment responses (e.g. subsurface processes and storage). For example the effect of flooding can be easily estimated from the amount of precipitation and the nature of the landscape. The subsequent impacts are also visible in a relatively short period of time, which might decrease the degree of complexity of its impact. However, drought is a large-scale phenomenon that involves various hydro-climatic processes and catchment characteristics. Its impact also depends on meteorological, hydrological and land surface factors and vulnerability to these factors may also depend on the water demand and water management systems. In addition, Drought and floods have different spatial and temporal scales, as indicated by Stahl and Hisdal (2004)), which makes these very different, also in terms of complexity. Therefore, it is acceptable to take drought phenomena as more complex than flooding.

Stahl, K. and Hisdal, H, (2004): Hydroclimatology. In: Tallaksen, L.M. & van Lanen, H.A.J. (Eds.) (2004) Hydrological Drought. Processes and Estimation Methods for Stream flow and Groundwater. Developments in Water Science, 48, Elsevier Science B.V., pg. 19-51.

Page 12767, Lines 15-16: "Holling et al., 1978" is given as "Holling (1978)" in references

We agree with your comment and we have revised the manuscript **[page 3 line 12]**.

Page 12768, Line 1: It might not be clear what "SD" means to the reader without writing out in full

We agree with your comment. We have revised the manuscript accordingly **[page 3 line 26, page 16 line 7, page 30 line 1]**.

Page 12772, Lines 15-16: What happens to data / thresholds for 29th February?

We removed hydrometeorological data corresponding to 29th, February as we have only the required data only once in four years. All the threshold level calculation methods take into account only the 28 days of the month of February throughout the observation period. We have added a line to explain this in the revised manuscript **[page 8 lines 21-25]**.

Page 12774, Lines 2-4: It would be worth stating the cutoff frequency used in this study

[**Also in replies to anonymous reviewer I**]Any time series signal that is represented by non-periodic functions (in this case say precipitation or discharge series) can be reproduced as a linear sum of many discrete sinusoidal frequency components. These discrete frequencies can be obtained using Fourier Transform; the method that converts the time series data into frequency components. For discrete time series signal data, the conversion is done by using large number of complex multiplications. Fast Fourier Transform uses special algorithm that accelerates the conversion process by reducing the number of such multiplications (Kimball, 1974; Knuth, 5 1998; Johnson and Frigo, 2007).

The conversion enables us to apply piecewise mathematical manipulations such as attenuation and removal on the frequency components above a predefined frequency called cut-off frequency. This manipulation results in smoothed series and when inverse Fourier Transform is applied, it provides smooth time series signal (smoothed time series of hydrometeorological variables), which is one used in this work as D_FF threshold level.

The question here might be ‘how is the cut-off frequency selected?’. We adapted (optimized) the cut-off frequency until when the inverse Fourier Transform is applied it best fits the 30D threshold level. We chose the 30D threshold level for optimization just because it does not require secondary smoothing (quantile calculation followed by application of smoothing techniques). The above explanation is included in the revised manuscript **[from page 10 line 20 to page 11 line 14]**

Page 12774, Line 19: "magnitude" is a drought characteristic itself; this phrase could confuse readers

We agree with the comment. We have revised the text in the manuscript **[page 12 line 20]**.

Page 12775, Line 22: Can you justify the hypothesis "drought severity should decrease"?

Differences between the precipitation time series and the threshold are always larger than those for groundwater and river flow. This is because precipitation fluctuates more than groundwater and river flow. This implies that the drought severity expressed as deficit volume is larger in the precipitation than in groundwater or river flow. We will add lines to clarify this in the revised manuscript **[page 20 lines 9-17]**.

Page 12776, Line 13: Text says "130 days" but Table 1 seems to suggest this value is 756 days. If the correct value is 130 days, this would make it considerably smaller than for the other three methods (unless those data are incorrect as well). Are these differences between tabulated data and data quoted in text the cause of the issues (mentioned above) surrounding the interpretation of Table 1?

Thank you for the comment. The values written are both correct. However, the text needs to be revised as "For example, the M_MA threshold level approach alone generated mean groundwater drought duration of 130 days longer than the average of mean drought durations computed using the other three threshold approaches applied to Upper Guadiana catchment". The average of the three mean durations in groundwater level (average of 614, 620 and 614 days) is shorter than mean duration calculated using M_MA in the above mentioned catchment. This information is added in the manuscript **[page 16 lines 3-7].**

Page 12776, Lines 25-29: It is not clear which catchment this sentence refers to?

Sazava catchment. We understand that it should be clearly stated as requested **[page 16 lines 24-29]**.

Page 12777, Line 3: It should be "30D" rather than "D30" written here

Yes, we agree and rewrite them in the revised manuscript **[page 8 line 6 and page 17 line 2]**.

Page 12777, Line 22: The artefact drought is only for a small part of "December 1984to June 1985"

Yes we agree. The artifact event sustained for 24 days from 1985-04-17 to 1985-05-11. Therefore, we have to rephrase that it only sustained for only this short duration **[page 17 lines 21-22]**.

Page 12778, Lines 15-17: Looking at Figure 8, there seems to be a break in the drought across June-July 1976 for discharge in the Narsjø catchment (as well as other smaller breaks later in 1976). This seems to contradict the phrase: "the discharge anomaly persisted for 308 days since 7 March 1976"

We applied pooling of minor drought events with inter-event period of 10 days and excluded minor droughts that sustain less than 15 days (Page 12774, lines 23-26). Figures display the threshold level counterplotted with the daily simulated data before these modifications (that we set up to define drought events). Therefore, such breaks and minor droughts appear in the figures. However, these breaks and minor droughts are taken into account in calculating the drought statistics (Table 1; number of droughts, mean duration and deficit volume). We will add a line in the revised manuscript to clarify this **[page 18 lines 15-23]**.

Page 12781, Line 1: Abrupt changes in discharge can also occur in temperate climates

Yes, that is true. We just forwarded examples.

Page 12781, Line 2: I am not sure that analyses on five European catchments means that conclusions can be drawn on the most suitable threshold level method for "global scale drought analysis"

We believe that the threshold level is an indicator for patterns in the inter-annual variability of a predefined domain in the hydrological cycle (precipitation, soil moisture, groundwater level or discharge). As a semi-distributed conceptual rainfall-runoff model, these variables are generated based on hydrological processes that occur in any catchment around the world. The catchments included in this study have diverse climate condition and catchment characteristics; ranging from most snow-dominated and fast responding catchment with sub-arctic climate characterized by mild summers and very cold winters (Narsjo) to slow responding catchment with Mediterranean and semiarid climate condition characterized by very warm summers and mild winters. Besides, the drought definition/identification in all the four approaches follows similar procedures (such as setting inter-event time period and elimination minor drought) are proposed based on tests of representative sample catchments around the world. Therefore, we expect that there will be little difference between approaches in global scale without pronounced seasonality. This argument is included in the manuscript **[page 7 lines 23-28].**

Page 12788, Table 1: Vertical lines could be used to separate the threshold methods into four sections (otherwise the headers "V (mm)" and "n" are too close together, and the "/" links them together)

We accept your comment. We have implemented this in the revised manuscript **[page 29 Table 1]**.

Page 12789, Table 2: "SD" should be written out in full in a figure caption

We accept your comment. We have implemented this in the revised manuscript **[page 30 Table 2]**.

Page 12791, Figure 2: It should state in the figure caption which catchment these curves represent

We accept your comment. We have implemented this in the revised manuscript **[page 32 Figure 2]**.