Answer to comments Editor

The original comments of the editor are in black color and indicated by “R:”. Replies by the authors (“A”) are colored in green. Actions are introduced by “Action:”, changes done in the manuscript are in italics.

General Comment:

As you can see from the reviewers' report, there are still some concerns on the paper, that I agree need to be addressed before evaluating the paper for publication. Please revise the paper accordingly, with special attention to the structure of the results and discussion sections and to the need to expand the analysis to other periods relevant for regions outside Europe in order to support the focus on global scale.

-- Thank you for your recommendations. 1) We agree that also that in the discussion section of manuscript, we showed results. We therefore followed the review suggestion and have merged the results and discussion section into one section. 2) Regarding the presentation of global-scale results we already had presented Figs. 3 (SMDAI) and 6 (QDAI), which provided the overall DAI indicator behavior globally and for the whole reference period. Additionally, we had already included the analysis of two non-European streamflow drought events are presented in Fig. 9. We do not think that an additional similar analysis for SMDAI would be suitable for the paper for various reasons: 1) The paper is already very long and more figures would rather overwhelm or bore the readers, 2) The main innovation is in QDAI while the inclusion of SMDAI is done to show that a) both soil and streamflow droughts can be assessed by a deficit-anomaly approach in a parallel and consistent manner and b) the important issue of drought propagation from soil moisture drought to streamflow drought can be analyzed with the introduced approach, 3) any presentation of SMDAI results during non-European drought events would just be illustrative and 4) the manuscript is primarily a methodological study on new drought indicator(s) and not a study on the occurrence of historical droughts around the globe. So we have not included any further analysis results for specific regions and drought periods as for the for reasons given, we strongly fell that this would be detrimental for the paper. Otherwise, we have adjusted the paper in many aspects in response to the helpful comments of the reviewer (please see below).

Answer to comments of Anonymous Referee #1

The original comments of Referee #1 are in black color and indicated by “R:”. Replies by the authors (“A”) are colored in green. Actions are introduced by “Action:”, changes done in the manuscript are in italics.

General Comment:

I would like to thank the authors for carefully considering my comments and revisiting the manuscript accordingly. Even if I found the new version significantly improved compared to the original version, I still have some major concerns on some details of the methodology, as well as on the way the results are presented, especially in the framework of a work on global scale drought.The authors introduced some simplifications from the original formulation of the DSI, but in my opinion they fail in highlighting if/how these simplifications impact on the behavior of the index.

-- In Lines 182-85, the comparison between both p_{soil} and p_DSI is mentioned along with the slight difference between the two for Aug 2003 observed in respective world maps represented in FigS2.

Action : To be more precise, we have added the following in line 184:
“For very few gridcells, SMDAI is much larger than DSI and there are some areas where DSI is slightly larger than SMDAI.”

In my understanding, the two main differences are: 1) the d index is compute on the Smax, and 2) the computation of p is performed differently. From the reported results I see a very high fraction of data with high d values (Fig. 2), and I am wondering how much of this is related to the adopted modification.

-- In our global scale model, different from European scale model used in Cammalleri et al (2016), we do not have the information on water content at wilting point and field capacity, which is why we cannot compute the deficit according to Eq. 1 of Cammalleri et al. for the global scale. With our definition of d (Eq. 1 in our manuscript), computed d at small soil moisture saturation values are smaller than with the definition of Cammalleri et al (2016). (their Eq. 1 and their Fig. 2) while d at high soil moisture saturation is not as close. This is due to the S-shaped d-curve of Cammalleri, while our d-curve is a straight line. We do not know whether Camalleri et al. compute smaller d values for August 2003. Anyway, differences would also be caused by the different hydrological models used to compute soil moisture in the two studies. Following the reviewers suggestion below, we added text to section 2.1.1 on how the deficit definitions differ between our approach and the approach of Cammalleri et al (2016).

In term of p, the new Figure S1 show to me a quite different behavior compared to DSI (e.g. for F = 0.87 P is 0.1 and 0.4, respectively).Similarly, Fig. 1 is supposed to show two contrasting examples, as done in Cammalleri et al. for DSI. However, these two cases do not seem opposite example at first glance. If you look at Fig. 3 in the DSI paper, in one case the DSI resembles d while in the other resembles p. In your analogous figure, both cases resemble p. This means that either you selected two cases that are too similar or that your simplifications do a disservice to the index. Please provide better examples or clarify.

When we evaluated Norwegian grid cell shown in Fig. 3 of Cammalleri et al (2016), we found that with our deficit approach and the output of our global hydrological model, deficits were not zero (or very close to zero), different from the results of Cammalleri et al. We think this is both due to our deficit definition (see explanation of differences above) and the applied hydrological model. We think that the Cammalleri deficit equation, where deficit is not a linear function of soil moisture saturation, does lead to more distinct deficit/no-deficit identifications. We do not think that our d definition is better than the one of Cammalleri et al (2016) and do not express this in the paper, but we also think that it is not necessarily worse. For our Fig. 1, we selected a grid cell in India that different from the grid cell in Spain has a very low deficit most of the time but different from the Norwegian cell in Cammalleri et al (2016) there are longer periods with a small (but non-zero) deficit. Therefore, different from the Norwegian cell in Cammalleri et al (2016), p peaks do not completely vanish. However, we do see a stark contrast between the Spanish and the Indian cell, as in the Indian cell the SMDAI is always much smaller than the anomaly p_soil.

Following my main concern in the first round of review, I still found the section on results quite lacking in the context of a global study. Too much emphasis is given to a specific arbitrarily-selected month (August 2003), and the addition of few figures in the supplementary materials does not alleviate the issue. I strongly suggest to the authors to reshape the approach adopted to show the results, in a way that better highlight the results during relevant droughts. As you stated, most of the globe is likely to be in no drought during any given month, so it is meaningless to show the behavior of the index during such period.I much prefer the approach adopted in Fig. 9 for the EFR, and I suggest to expand this event-based approach to the other analyses as well. This is valid for both indices.

We have selected August 2003 for as an example for showing global maps of SMDAI (Fig. 2) and QDAI (Fig. 5) because it was a month with an extreme drought in Central Europe, to which we refer also in the time series plots for SMDAI (Figs. 1) and QDAI (Fig. 4). To cover more than one arbitrary month but show the overall behavior over the whole time series and globally, we prepared Figs. 3 (SMDAI) and 6 (QDAI). We disagree with the reviewer that it is meaningless to show where there is no drought.

We have already included the analysis of two non-European drought events regarding 9) and do not think that an additional similar analysis for SMDAI would be suitable for the paper for various reasons. 1) The paper is already very long and more figures would rather overwhelm or bore the readers, 2) The main innovation is in QDAI while the inclusion of SMDAI is done to show that a) both soil and streamflow droughts can be assessed by a deficit-anomaly approach in a parallel and consistent manner and b) the important issue of drought propagation from soil moisture drought to streamflow drought can be analyzed with the introduced approach, 3) any presentation of SMDAI results during non-European drought events would just be illustrative and 4) the manuscript is primarily a methodological study on new drought indicator(s) and not a study on the occurrence of historical droughts around the globe.

Also, at the moment there is no clear distinction between the “results” and the “discussion” section, with both containing what can be called results. The discussion section should discuss the outcomes reported in the results section, not adding new outcomes. Please improve this structure, even by simply merging the two into a single results and discussion section and homogenize.

-- Thank you for your recommendation. We have merged the results and discussion section.

Overall, I see a lot of potentiality in this paper, and a large amount of valuable data that can really help the drought community in improving the understanding of both soil moisture and river droughts, but I think that these data need to be better presented to the readers to transfer the right message.

Specific Comments

L8-9. remove “the condition of”
---- Action: Thank you for pointing it out. We have adapted the suggested changes.

L11. Please add full reference to the DSI in the abstract.
--- During the initial submission, the editor had advised that it would be better to avoid references in the abstract.

L12. Too many details for an abstract. Please reword “… is based on… mapping function” as something in the line “…as a simplified version of the DSI (ref)”.

--- Thank you for your suggestion. However, the term simplified would not reflect our modification. We feel that it is important to indicate it already in the abstract. What the implemented improvement consists in.

L39. I think that “can be used” is not necessary here.
---- Action: Thank you for pointing it out. We have adapted the suggested changes.

L98. outputs.
---- Action: Thank you for pointing it out. We have adapted the suggested changes.

L100. Please add “:” after “five sectors”.

---- It is already there L100: *“The water use models compute water use in the five sectors: household, manufacturing, cooling of thermal power plants, livestock and irrigation”*

L109-110. I suggest to reword this sentence. At first glance, it seems that “source of water abstraction” is not accounted, and not that this is done by a different module.

--- Action: The new sentence now starts with *“ The water use models themselves …” instead of “The water use models ...”*

L131-133. It is important to mention specific validations made on low-flow/drought (if any). It is well-known that performances on the lower spectrum of the flow regime may differ quite significantly from the average/high regimes. In absence of specific validations on drought, it is worth mentioning this other possible source of error.

--- Action: We added the following sentence in line 135:

"It is found that WaterGAP can simulate the low flow percentile (Q95) very well, but it can also overestimate the return period of low streamflow (Zaherpour et al., 2018). "

L135-136. This sentence needs some rewording in my opinion.
---- Action : We have replaced:

"This study uses simulated data of 30-years (1981 – 2010) monthly time series of WaterGAP gridded (0.5° x 0.5°) output of 67420 land grid cells covering all land areas of the globe except Greenland and Antarctica, for 1) soil moisture (S) *[mm], 2) streamflow* (Q_{ant}) *[km3 month-1], 3) streamflow under naturalized condition* (Q_{nat}) *[km3 month-1], assuming there are no human water abstraction or man-made reservoirs, and 4) total surface water abstractions* (WU_{sw}) *[km3 month-1]."*

By

This study uses 30-years (1981- 2010) monthly time series of WaterGAP gridded (0.5° x 0.5°) outputs for 67420 land grid cells covering all land areas of globe except Greenland and Antarctica. These include 1) soil moisture [mm], 2) streamflow [km3 month-1], 3) streamflow under naturalized condition [km3 month-1], assuming there are no human water abstraction or man-made reservoirs, and 4) total surface water abstractions) [km3 month-1]."

L144. Here I see a major difference between this approach and the one proposed in DSI. If I understand correctly, in DSI the critical water content is used (50% of filed capacity) rather than the field capacity itself. Indeed, the absence of water stress starts in many cases well before field capacity in reached. I suggest to make this difference much more clear (see also discussion 4.1) and to clarify the impact on the high d values observed over most of the globe in fig. 1.

---- We thank you for this suggestion.

Action: We have added the following text to section 2.2.1 where we introduce the computation of the soil moisture deficit d.

"This definition of soil moisture deficit is different from the definition used in Cammalleri et al. (2016, their Eq. 1) because their definition cannot be applied when using the global hydrological model WaterGAP to compute soil moisture. The deficit computation according to Cammalleri et al. (2016) requires data on soil moisture content at wilting point and at field capacity, which is not available in WaterGAP. With our approach, which is consistent with the way of computing actual evapotranspiration from potential evapotranspiration in WaterGAP, d-values at low soil moisture

saturation are lower than those of Cammaleri et al. (2016), while at high saturation they are higher. Consequently, we identify very few months and grid cells with a deficit of zero, likely less than we would do if we would have implemented the deficit definition of Camalleri et al. (2016).”

L211-216. This is a quite key point, that needs to be stressed more. It would be interesting to have some more insight on the differences with anomalies computed on the deficit, even if no detailed analysis is provided.

---Thank you for the suggestion.

Action: We replaced the sentence

“The unusualness of a streamflow drought is better captured by a standard cumulative distribution function that can reproduce the statistical structure of streamflow (Qant) compared to a standard distribution function reproducing the statistical structure of streamflow deficit (dQ) due to the temporal variability of the water demand.”

by

“We select to consider the anomaly of streamflow (Qant) instead of the anomaly of the streamflow deficit (dQ) as the temporal variability including long-term trends of the water demand prevented us, for most grid cells with relevant water demand, from identifying a standard distribution function for the time series of dQ.”

The revised sentence gives some insights into the anomalies of streamflow vs. anomalies of streamflow deficit and explains why we could not consider the anomaly of the deficit.

L245. It would be more consistent to use the ECDF for all the cells, especially since you are not using the mode as reference (which needs to be derived from the theoretical distribution), even if I guess that there is not much difference in the case of a good fitting. I also suggest to integrate figure S4 here rather than as supplementary material.

--- We think that it is better to fit optimal functions where it is possible (as is done in most drought studies). And we agree with you that it will not change much the results in those grid cells where we could identify CDFs. We think that we should keep Fig. S4 in the supplementary material, as it provides just a background explanation and not a central methodological information or a study result.

L252. The relationship between dsoil and what? Please rephrase.

--- Thank you. Action: We change the sentence as:

“ The relations between d_{soil}, mean monthly (d_{soil_mean}), p_{soil} and SMDAI are further clarified by time series of these variables in Figure 1 for two grid cells with rather different characteristics: a grid cell in Germany (42.25N, -121.75 E, left panels in Figure 1) and one in northeast India (88.25 E,27.25 N, right panels in Figure 1).”

L264. “very high soil moisture saturation”. Please reword.

-- Unfortunately, we cannot think of a better term.

L286. Still, very surprising that most of Canada during December 1999 is in high water stress. The high fraction of the world with dsoil > 0.75 is very surprising in general. This need a clear explanation (it seems to happen also in December 1999). Also, I do not think that the analysis of a single month (out of 30+ years) is enough here. You need to come up with a synthetic map that summarize the whole period (or the major events for different areas), not just a single case (see major comments).

---The low soil moisture in Canada in winter is due to sub-zero temperatures, such that all precipitation falls as snow and does infiltrate the soil. In WaterGAP, like in most hydrological models, soil freezing and permafrost is not taken into account, and in case of no liquid water entering from the top of the soil, the soil drains downward and becomes more and more unsaturated.

Action : We added in the following text in italics in line 286:

“but high in most snow-dominated northern high-latitude regions *(as no liquid water enters the soil)*,

Regarding a summary presentation over the whole 30 years, we have done this in Figs. 3 and 6 for SMDAI and QDAI, respectively.

L301-302. Is it realistic to call this soil moisture drought, when most of the deficit is due to snow fall rather than liquid precipitation (hence, the soil is covered in snow)? I see this more of a problem for the indicator rather than a desired behavior. Example: How is a good thing that Australia and Canada behave the same?

--- We expect that trees in Canada during months with below zero temperatures do suffer from some stress from low water availability. While it may be true that they suffer more from the low temperatures than from the low soil water saturation, any vegetation during cooler winters, with e.g. temperature below 5-10 °C, will react to lower temperatures, too. We believe that it is beyond the scope of a drought study to take the combination of temperature and water stress into account.

L308-310. this is an unnecessary introduction.

--- We prefer to keep this sentence to better introduce the readers with respect to the results presented in the following sentences.

L337. The effect of EFR in defining a drought is quite important and needs to be better highlighted in my opinion. At the moment, relegating this analysis to a supplementary material does not give justice to a really key point of transferring this concept from soil moisture to stream flow.

--- Thank you for the comment. We agree that EFR is important for defining streamflow and hence, we have already committed an entire subsection (4.2) on it. With 11 figures, with main text is already unusually high for a research paper

L337-340. I am not following this argument. It seems to me that QDAI is rather similar to p in the cell over US, and I do not see any major clear differences between the two sites in term of "strength" of the droughts. Please elaborate better this concept.

Action: Thank you for your comment. We have adapted the required the required sentences (L 331 – 346)

" Characterized by a high seasonality, anthropogenic surface water demand, WU_{sw} (dashed grey line in center plot) and total surface water demand (i.e., $WU_{sw} + EFR\_0.8$, orange line in center plot) result in very high deficits d_Q (green line of the bottom plot) during almost every summer. However, there are only a few months with drought as identified by the anomaly-based drought hazard p_Q exceeding zero (dark blue line).This occurs because the decade shown in Figure 4 happens to be a very wet decade compared to the whole reference period. Another reason is that more than 20% of the years show zero streamflow in the calendar months August and September such that p_Q is zero in all 30 August and September months of the reference period, i.e. no drought is indicated even in case of zero streamflow (see left panel of Figure S7). Due to the large deficit values, $p_Q$is almost always smaller than d_Q in this US grid cell.

In the German grid cell (right panels in Figure 4), the relatively low anthropogenic surface water abstractions result in almost identical values of Q_{nat_mean} and Q_{ant_mean} (lines overlap in the top plot), and total surface water demand is very similar to EFR (lines overlap in the center plot). Non-zero d_Q values (bottom plot) are mainly computed if Q_{ant} is lower than EFR, such as during the central European drought of 2003. It is sensible to consider this type of situation as a drought hazard as water supply companies would have to stop any surface water abstraction if they wished to protect the river ecosystem. Different from the US grid cell, droughts are rather equally distributed over all decades of the reference period in the German grid cell but the summers of 2003 and 2005 suffer from the most severe droughts of the reference period, in line with expected dryer summer due of climate change. Even if taking into account EFR as 80% of of Q_{nat_mean} ($EFR_{0.8}$), the total surface water demand is so low that in contrast to the US cell, d_Q is always smaller than p_Q.

Assumptions about the magnitude of EFR have a strong impact on d_Q and thus QDAI of all grid cells except those with very high surface water abstractions such as the US cell. If the water demand of the ecosystem were assumed to be only 20% of Q_{nat_mean}($EFR\_0.2$) instead of 80% of Q_{nat_mean}, d_Q decreases somewhat in the US cell but reduces to zero during the whole reference period in the German cell (Figure S6). Therefore, water suppliers in the German grid cell would not suffer from any drought hazard (as indicated by QDAI) and would not have to decrease their surface *water abstractions even during a drought similar to the 2003 central European drought."*

L352. "…is mostly smaller than less than…"

Action: Thank you for pointing it out.

We have changed it to *".... is mostly less than..."*

L354. Cells with recurring zero-flow should be treaded differently, since the deficit (how low is Q compared to the historical data) cannot be used as reliable quantity for drought. The length of dry spells is considered a much better proxy here. I suggest to better clarify that such areas need to be masked from the analysis (as successively discussed in L365-366). Please also use a different color for these areas, since the current color is too similar to the dark red used for extreme drought (it is the case also for Fig. 6, see int vs. high frequency).

-- In Figure 5, when analyzing a QDAI in a specific month, it is not necessary to exclude cells with more than 6 months (of the 30 months for each calendar months) as the ECDF will be indicative of the anomaly p (see Fig. S7 left). This is different from Fig. 6 where we also show frequencies of no-drought conditions. We followed your advice and changed the color of the masked out cells in Figs. 5 and 6.

L361. "…no-drought conditions according to QDAI (Figure 6)…"

---- Action: Thank you for pointing it out. We have adapted the suggested changes.

L399. A clear definition of (semi)arid and humid is needed in the methodology. Also, what about the other climates? Where all the cells classified as either of the two, and how?

---- Action: Thank you for your recommendation. We have added the definition of (semi)arid and humid in supplement.

L407. I do not really see any major differences between the two months, which is also kind of expected if you do a global average. Differences can be related only to the fact that there is more land in the northern hemisphere compared to the southern. Again, I see more useful an analysis on the full dataset (or specific drought events) rather than 2 randomly selected months that give very similar outcomes.

–Action: We have added an additional box-plot in Fig 8. where global distribution of QDAI in August 2003 (left), December 2003 (middle) and for all 360 months of the reference period (right), computed with alternative assumptions about **EFR** for grid cells with humid and (semi)arid conditions. Grid cells where all three **EFR** assumptions result in QDAI = 0 are not included.

L449. These two examples need to be better highlighted in Figure 10. For what I can see, even SSFI has no drought at the end of 2005, or maybe you are referring to end of 2004. What I see is two drought periods in SSFI across the 2004 and 2005 lines. Also, what about the other two cases? No examples where QDAI improves on SSFI?

-- Thank you for your comment. We have already addressed the following in L445 - 452

4.4 I miss the role of this section, which again focuses only on a single point and a single case. Even on the specific case, what is the message that you are trying to pass here on this known phenomenon?

-- We want to suggest that two proposed indicators can be used together to analyse drought propagation.

Fig. 4. There are some inconsistencies in the plots. The legend of the plots on the bottom line seems to be off, also according to the text (d should be in green and p in blue).

Action: Thank you. The required corrections have been made in figure 4.

Also, in the top-left panel Qant seems to be always below Qant_mean, which suggest to me that the opposite occurs before 2000. This can be related to a trend in the data, which should be highlighted and discussed.

Action: Thank you for the comment. We observed that the decade shown in Figure 4 happens to be a very wet decade compared to the whole reference period have adapted the required paragraph (L: 331 – 346) which discusses and highlights the same as follows:

" Characterized by a high seasonality, anthropogenic surface water demand, WU_{sw} *(dashed grey line in center plot) and total surface water demand (i.e.,* $WU_{sw} + EFR\_0.8$*, orange line in center plot) result in very high deficits* d_Q *(green line of the bottom plot) during almost every summer. However, there are only a few months with drought as identified by the anomaly-based drought hazard* p_Q *exceeding zero (dark blue line).This occurs because the decade shown in Figure 4 happens to be a very wet decade compared to the whole reference period. Another reason is that more than 20% of the years show zero streamflow in the calendar months August and September such that* p_Q *is zero in all 30 August and September months of the reference period, i.e. no drought is indicated even in case of zero streamflow (see left panel of Figure S7). Due to the large deficit values,* p_Q*is almost always smaller than* d_Q *in this US grid cell.*

In the German grid cell (right panels in Figure 4), the relatively low anthropogenic surface water abstractions result in almost identical values of Q_{nat_mean} *and* Q_{ant_mean} *(lines overlap in the top plot), and total surface water demand is very similar to* EFR *(lines overlap in the center plot). Non-zero* d_Q *values (bottom plot) are mainly computed if* Q_{ant} *is lower than* EFR*, such as during the central European drought of 2003. It is sensible to consider this type of situation as a drought hazard as water supply companies would have to stop any surface water abstraction if they wished to protect the river ecosystem. Different from the US grid cell, droughts are rather equally distributed over all decades of the reference period in the German grid cell but the summers of 2003 and 2005 suffer from the most severe droughts of the reference period, in line with expected dryer summer due of climate change. Even if taking into account EFR as 80% of of* Q_{nat_mean} $(EFR_{0.8})$*, the total surface water demand is so low that in contrast to the US cell,* d_Q *is always smaller than* p_Q*.*

Assumptions about the magnitude of EFR have a strong impact on d_Q *and thus QDAI of all grid cells except those with very high surface water abstractions such as the US cell. If the water demand of the ecosystem were assumed to be only 20% of* $Q_{nat\_mean}(EFR\_0.2)$ *instead of 80% of* Q_{nat_mean}*,* d_Q *decreases somewhat in the US cell but reduces to zero during the whole reference period in the German cell (Figure S6). Therefore, water suppliers in the German grid cell would not suffer from any drought hazard (as indicated by QDAI) and would not have to decrease their surface water abstractions even during a drought similar to the 2003 central European drought."*