Response letter to RC1

In this response letter, we put the text of the reviewers in bold, and the questions are numbered as Q1, Q2 etc. Our answer starts with Answer and is in normal font-weight; The text in the manuscript is put in quotation marks. An empty line separates the questions.

This paper discusses ways to determine heatwaves, droughts, and combination of them in a systematic way and at daily scale allowing to better capture the start and end date of the events. They use 120 years of data over Uccle (point) and suggested an objective method to remove minor spells and merge mutually dependent ones and discussed four ways to define a CDHW event. They eventually used the outcomes to study the seasonality of the drought, heatwaves, and combination of them. They found that seasonal patterns, with the occurrences of droughts and heatwaves being negatively dependent in winter but positively dependent in the other three seasons at Uccle.

I believe the paper is well structured and well written in almost all aspects: the science within the scope of the journal, relatively novel in the method and concepts, important conclusion, good literature review, and good discussion of the outcomes and associated interpretation. The conclusions are backed up with proper studies, the title and abstract are fine. I only have few minor points:

Answer: We are truly grateful for the reviewer's positive feedback and thoughtful evaluation of our manuscript. Your comments are highly valuable to us. Here, we address the points raised and provide our responses to each of them.

Q1: The study is performed using a point measurement that has 120 years of daily temperature and precipitation data. Yet, the number of events, especially compound events, are not much. Is there any recommendation for the minimum length of the data record if someone wants to apply the method over other regions?

Answer: In the proposed method, three steps are based on fitting probability distributions. Because of this, the longer the time series, the better the fitting: (1) calculating SPI and SHI, (2) events' severity assumed to follow a GEV distribution, and (3) the events' arrivals assumed to follow a Poisson process.

Generally, 30 is widely suggested and used as the minimum length or number for fitting probability distributions in the field of hydrology. Based on this, step (1) asks for at least 30 years of observation. Steps (2) and (3) ask for the final number of identified events to be larger than 30. The tricky part is that the final number of identified events number also depends on the accumulation period and merging threshold besides the observation length.

If we are requested to give a minimum length for the data record, at least 30 years of continuous precipitation/temperature data are needed, but longer records would be preferable. And at the same time, it is important to check the goodness-of-fit performance in steps (2) and (3).

Q2: Line 26: remove “more” before regionally

Answer: We sincerely appreciate the reviewer’s meticulous review of our manuscript. The suggestion to remove the word ’more’ before ’regionally’ on Line 26 is duly noted. We will promptly make the adjustments in the manuscript.

Q3: Line 39: what is (7d)? Fig 7d in their paper?

Answer: We apologize for the unclear expression: (7 d) in Line 3 means 7 days. We will change the “(7 d)” to “(a week)” to avoid any misunderstanding. Here are the revised sentences (to provide the context, we also copied two sentences before Line 39):

“This coarse scale of the drought index and the mismatch of scales between the drought and heatwave indices could entail some bias on the actual start and/or end dates as well as the severity of droughts, further affecting the identification of CDHW events. A recent study by Mukherjee and Mishra (2021) upscaled the resolution to weekly droughts and daily heatwaves. They identified a drought week when the self-calibrated Palmer Drought Severity Index (Wells et al., 2004) falls below the 10th percentile, thus regarding short dry spells (a week) as drought events. Although the mismatch of scales is partially solved this way, too short dry spells might be identified as drought events.”

Q4: Line 98: not sure how you come up with 91 days? Is the 91 day window for both 30yr and entire record analysis? I think it is better to further clarify this.

Answer: In Line 98, the 91 refers to 91 windows, resulting from the 30-year moving window approach from 1901 to 2020. Considering the possible non-stationarity of climate variables, we use the past 30 years as the historical reference period, then this method generates windows (1901-1930, 1902-1931, . . . , 1991-2020), resulting in 91 windows in total.

We will promptly clarify this in the manuscript, shown below (to provide the context, we also copied three sentences before Line 98):

“Accounting for the fact that people and ecosystems adapt to a changing climate, we use the past 30 years as the historical reference period and average the values on every day during this period. This average is then regarded as the expected normal temperature on this day, instead of the average during the longest climatology (the period of record). This 30-year moving window approach has been suggested to deal with the climate non-stationarity bias (Hoylman et al., 2022). To examine the impact of this approach, we apply the Mann–Kendall (MK) test for the daily mean temperature in the period 1901–2020 (the period of record) and in the 30-year moving windows (1901-1930, 1902-1931,. . . , 1991-2020; 91 windows in total per day) based on the data described in Subsection 2.1.”

Q5: Line 224: shouldn’t it be the sum of “positive” SHI?

Answer: In the context of our calculation of marginal drought and heatwave severity, we understand the point you’ve raised about the sign of SPI and SHI values within CDHW events. While most SPI values are indeed negative within CDHW events, they may not always be exclusively so. Similarly, SHI values are typically positive within CDHW events, but exceptions may exist.

To ensure consistency and clarity in our calculations, we have modified the text from

"The marginal drought severity is computed as **the sum of the negative SPI values** within the CDHW event, while the marginal heatwave severity is computed as the sum of the SHI values within the CDHW event."

to:

"The marginal drought severity is calculated as **the negative of the sum of the SPI values** within the CDHW event, while the marginal heatwave severity is calculated as the sum of the SHI values within the CDHW event."

Q6: Line 254: Check if the orders is right. I think it should be 1.3, 1, and 0.5 instead.

Answer: Thanks for pointing out this error. The right order should be 1.3, 1, and 0.5. Here is the corrected sentence:

"In the case N_{h} = 3, warm spells shorter than 6, 9, and 13 d are removed when $\mathrm{SHI}_{\mathrm{h}}$ ranges over 1.3, 1, and 0.5."

Q7: Explain why no merging is carried out in some cases in Table 1.

Answer: In the method we proposed, first the removal procedure is carried out, and subsequently the merging procedure. After the minor spells are removed, the inter-arrival time between spells correspondingly becomes larger, and the neighboring spells are less likely to be mutually dependent. One case that could happen is that after removing minor spells, the inter-arrival time already follows an exponential distribution and thus, there is no more need to carry out the merging procedures. So, no merging occurs in some cases in Table 1. We will add a short explanation in the manuscript.

Q8: For N_{d}=60 and SPI =-1.3, what does Md=51 mean? why all other Mds are much less? Are these results stable?

Answer: For $N_{\mathrm{d}} = 60$ and SPI =-1.3, $M_{\mathrm{d}} = 51$ means that neighboring spells are merged into one longer event if the proximity is less than 51. The proximity is defined as below:

This study introduces two terms: *total deficiency* and *proximity*. The total deficiency of a time interval $[a,b]$ corresponds to the area enclosed between the SPI (resp. SHI) curve and the pre-identification threshold line. The total drought deficiency (TD_{d}) is calculated as:

$$\mathrm{TD}_{\mathrm{d}} = \sum_{i=a}^{b} (\mathrm{SPI}_{m,i} - \mathrm{SPI}_{\mathrm{d}}),$$

where $\mathrm{SPI}_{\mathrm{d}}$ is the pre-identification threshold for droughts. Similarly, the total heat deficiency (TD_{h}) is calculated as:

$$\mathrm{TD}_{\mathrm{h}} = \sum_{i=a}^{b} (\mathrm{SHI}_{\mathrm{h}} - \mathrm{SHI}_{m,i}),$$

where $\mathrm{SHI}_{\mathrm{h}}$ is the pre-identification threshold for heatwaves.

To measure how close two neighboring spells are, we define the *proximity* as the total deficiency of the time window between the spells. The proximity provides a comprehensive view of the period between two spells, accounting for both duration and how far it is from being dry (resp. warm). For example, a small proximity value means the interval is short and/or the SPI (resp. SHI) values are near the pre-identification threshold. In that case, the neighboring spells are more likely to be mutually dependent, suggesting to merge the neighboring spells and the interval between them into one longer spell.

Table 1. Removal thresholds (R_d), merging thresholds(M_d), ΔC, and the ratio of ΔC to C_1 , for droughts (a) and heatwaves (b).

(a)	R_d			M_d			ΔC			ΔC /C_1		
SPI$_\mathrm{d}$	-0.5	-1	-1.3	-0.5	-1	-1.3	-0.5	-1	-1.3	-0.5	-1	-1.3
N_d=15	23	14	10	1	15	11	2	15	5	0.015	0.111	0.035
N_d=30	42	22	20	8	46	1	9	15	1	0.115	0.185	0.015
N_d=45	44	31	25	1	-Inf	1	5	0	2	0.053	0	0.039
N_d=60	58	45	24	0	-Inf	51	4	0	7	0.058	0	0.149
N_d=90	81	59	32	1	9	17	2	4	6	0.038	0.133	0.176

(b)	R_h			M_h			ΔC			ΔC /C_1		
SHI$_\mathrm{h}$	0.5	1	1.3	0.5	1	1.3	0.5	1	1.3	0.5	1	1.3
N_h=3	13	9	6	0	0	3	6	1	9	0.027	0.005	0.040
N_h=5	16	10	8	0	-Inf	-Inf	4	0	0	0.019	0	0.000
N_h=7	27	11	10	4	-Inf	6	3	0	10	0.051	0	0.076
N_h=10	23	13	12	-Inf	-Inf	0	0	0	2	0	0	0.015
N_h=15	29	17	18	2	-Inf	11	9	0	1	0.073	0	0.014

why all other Mds are much less? Are these results stable?

To answer the above two questions, we did additional calculations to show the effects of the merging procedure. The number of events after removal and merging thresholds is denoted as C_1, and the number of events when carrying out the merging procedures is denoted as C_2; the difference between C_2 and C_1 is denoted as ΔC, $\Delta C = C_2 - C_1$. Table 1 shows the removal threshold, merging threshold, ΔC, and the ratio of ΔC to C_1 for droughts (a) and heatwaves (b).

We can see from Table 1 that a larger M_d doesn't always mean a larger ΔC. For example, M_d=51 (N_d=60 and SPI =-1.3) results in ΔC = 7. M_d=15 (N_d=15 and SPI =-1) and M_d=46 (N_d=30 and SPI =-1) both result in ΔC = 15.

We also can see from the tables that $\Delta C/C_1$ is upper bounded, *i.e.* ΔC /C_1 <0.2 in drought identification and <0.1 in heatwave identification. The merging process is more common in drought identification than in heatwave identification, with larger $\Delta C/C_1$. From this, one could see that the results of M_d are relatively stable.