Author Responses to Reviewer Comments

Katoria Lekarkar*

We would like to thank the reviewers for the time and effort to review our manuscript. The comments provided by the reviewer will undoubtedly improve our manuscript. In the sections that follow, we provide point-by-point responses to the comments on the manuscript. The responses to individual comments are written in this **color**. The revisions made in the manuscript are italicized and written in this ***color***.

1: Responses to Reviewer #1:

Author response:

Thank you for providing feedback to improve our manuscript. We have followed all the suggestions and implemented the changes as detailed in the sections below.

Reviewer Comment #1: Overall assessment

The manuscript reconstructs multi-decadal root-zone soil moisture over Belgium with mHM and characterizes drought events using an SMI-based framework, comparing them with precipitation-based indicators (SPEI-1/3). The central finding—that 2011–2020 is the driest decade since at least 1971—is relevant for Belgian drought monitoring and well aligned with broader European trends.

However, the presentation is currently hard to follow due to (i) too many metrics without a clear hierarchy for ranking drought severity. (ii) Several methodological elements (MPR, KDE→SMI percentiles, Fisher-z, NSE definition, event splitting/merging rules) need one-line clarifications so readers can reproduce and interpret results.

Author response:

We appreciate this constructive assessment. We agree that the manuscript would benefit from clearer criteria on how drought severity is ranked and from concise clarifications of several methodological components. In the revised manuscript, we have (i) explicitly defined a hierarchy of drought-severity metrics, (ii) added short clarifying sentences where methods are introduced, and (iii) improved the structure of the Results section to enhance readability and reproducibility.

*Corresponding author: katoria.lesaalon.lekarkar@vub.be

Reviewer Comment #2:

The title can be read as if only 2011–2020 is analyzed, yet the study reconstructs 1970–2020 and concludes that 2011–2020 is the driest decade of the five. Please revise the title.

Author response:

We have revised the title to reflect the full reconstruction period (since 1970) and to clearly state the main decadal conclusion. We have revised the title to:

Reconstructed soil moisture droughts in Belgium reveal that 2011–2020 was the driest decade since 1970.

Reviewer Comment #3:

Provide a simple, explicit severity ranking protocol. At present the Results toggle between TDM, SMI$\leq \tau$ area, exposure months, peak area, duration, SPEI-1/3, etc., without a decision rule. Readers cannot tell which event is "most severe." Please state a clear hierarchy.

Add a Table listing the top events with: TDM, peak area (%), duration, exceptional-class exposure.

Annotate TDM values directly in Fig. 5 and state in the caption which tie-breaker decided final ranks when two events are similar.

Author response:

We agree that at present the hierarchy of drought classification is not very clear and that can confuse the reader. In the revised manuscript, we now provide a clear ranking protocol. Events are ranked primarily by Total Drought Magnitude (TDM), and when two events have similar TDM (difference $\leq$ 1%), ties are resolved using peak affected area, then duration, then exceptional-class exposure. We have also added **Table 2** to manuscript, which list the top-ranked events with TDM, peak area, duration, and exceptional-class exposure. As recommended, we have also annotated TDM values directly in Fig. 5. The caption in Fig. 5 also now states the tie-breaker hierarchy used to determine final ranks. The updated Fig. 5 and Table are shown below.

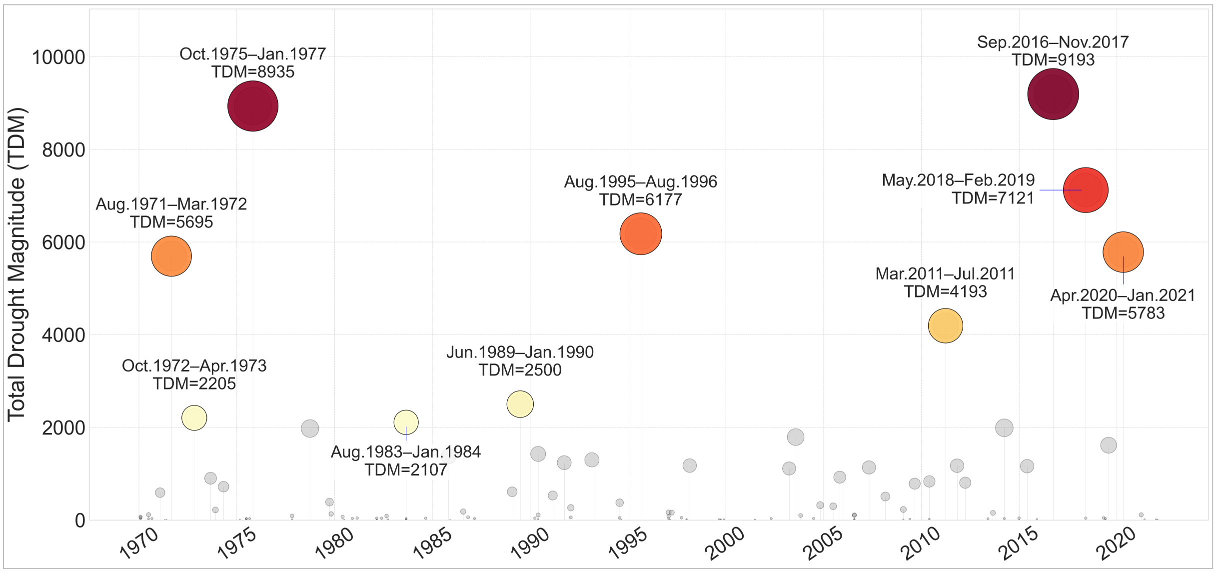

Figure 1: Duration and magnitude of drought events from 1970 to 2020. Each circle represents a drought event, positioned according to its start date (x-axis). The circle size is proportional to the Total Drought Magnitude (TDM) of each event. The ten most severe droughts, ranked by TDM, are highlighted with coloured markers, annotated with their corresponding periods and TDM values. Events are ranked primarily by TDM and when two events have similar TDM (difference $\leq 1\%$), ranks are determined by peak affected area, then duration, then exceptional-drought exposure (defined as SMI ≤ 0.02).

Table 1: The ten biggest soil moisture drought events in Belgium ranked by Total Drought Magnitude.

Rank	Event period	TDM	Average affected area (%)	Duration (months)	Exceptional class exposure (%·mo)
1	Sep 2016–Nov 2017	9193.14	64.0	15	182.9
2	Oct 1975–Jan 1977	8934.97	62.6	16	103.4
3	May 2018–Feb 2019	7120.88	73.1	10	108.3
4	Aug 1995–Aug 1996	6177.27	60.3	13	69.8
5	Apr 2020–Jan 2021	5782.83	58.0	10	27.7
6	Aug 1971–Mar 1972	5694.68	72.9	8	57.3
7	Mar 2011–Jul 2011	4192.92	81.6	5	84.9
8	Jun 1989–Jan 1990	2500.37	51.5	8	8.0
9	Oct 1972–Apr 1973	2204.63	35.4	7	11.2
10	Aug 1983–Jan 1984	2107.37	47.0	6	22.4

Reviewer Comment #4:

Lines 363-365, is 2016–2017 or 1975–1977 drought bigger? And based on which indicators? Lines 411-414, based on drought persistence, 2011-2020 is the biggest one.

Author response:

We understand that the language can be confusing to the reader. In the revised manuscript, we have explicitly ranked the droughts in Table 2 and clarified that TDM is the primary ranking metric. As for lines 363-365, we have now clarified that the 2016-2017 event is bigger than the 1975-1977, based on the TDM and affected area. In lines 411-414, we are discussing all droughts occurring in a particular decade, not individual drought events. So we mean that droughts were most frequent in the 2011-2020 decade. This frequency combines the occurrence of all the droughts in a particular decade. (For 2011-2020 for example, it includes the droughts in 2011, 2016–17, 2018–19, plus all the other smaller droughts).

Reviewer Comment #5:

Lines 365 and 367, when you talk about area percentage, provide the figure reference.

Author response:

Thank you for this suggestion. We have now cross-referenced Table 2 in the main text.

Reviewer Comment #6:

Lines 368–372 discuss 2022–2023, but your decadal analysis ends in 2020. Please remove or move to Discussion/SI with an explicit caveat.

Also line 341 says Fig. 5 covers 1970–2023, but the figure appears to show 1970–2020; please make the figure and caption consistent with the text.

Author response:

Thank you for pointing this out and we agree. Since the 2022-23 drought is outside the period of our analysis, we have now removed that section entirely. To maintain consistency, we have also corrected Fig. 5 accordingly.

Reviewer Comment #7:

Add a short description about each subsection at the start of Results. Two–three sentences will prevent readers from getting lost.

Author response:

Thank you for this suggestion. Indeed it is needed to improve readability. We have now added a short introductory section at the start of the Results section where e discuss the dynamocs of droughts during the five decades. We have also explained that the sections complement each other and why each is needed. Below we provide the excerpt of the short description to the Results section.

To summarize how soil-moisture drought behaviour evolves across decades, we use three complementary metrics. First, we quantify the magnitude of each event using the Total Drought Magnitude (TDM), which integrates drought severity over space and time and thus allows drought events to be ranked consistently (Section 3.2.1). Second, in Section 3.2.2, we describe how drought severity is distributed by quantifying the fraction of drought-affected area falling into different severity classes (moderate, severe, extreme, and exceptional) within each decade. These classes capture shifts in the composition of drought conditions beyond just the total magnitude. Third, in Section 3.2.3, we quantify cumulative drought exposure as the total number of months in which each grid cell experiences drought per decade (months need not be consecutive). This metric summarizes how frequently drought conditions recur at a given location over a decade. For decadal summaries, we defined decades starting from 1971 (i.e., 1971–1980) since SPEI construction requires accumulated water-balance anomalies over preceding months (January 1970 will thus not have SPEI-1 values, while January–March 1970 lacks SPEI-3 values. The first year with complete SPEI values is 1971).

Reviewer Comment #8:

Lines 196-199, explain why resolutions differ.

Author response:

We have described in detail how the model harmonizes these differing resolutions in Section 2.2. mHM distinguishes between Level-0 (L0) datasets, which define the static morphological datasets (e.g. land use, soils, DEM), and Level-2 (L2) datasets, which represent the meteorological inputs. Since gridded meteorological inputs are often available at coarser resolutions than morphological data, mHM allows the two datasets to be provided in different resolutions. We have explained this in Lines 158–170. The model harmonizes the data internally using the Multiscale Parameter Regionalization (MPR) technique. With MPR, fields of parameters at a given modelling scale are obtained by upscaling their corresponding estimates at the scale of the input data, based on either the arithmetic mean, the geometric or harmonic means, or the majority operator (Kumar et al., 2013). This upscaling leads to quasi scale-invariant parameters that enable mHM to preserve the spatial variability of state variables, conserve mass balance and reduce overparameterization.

Section 2.2 of the manuscript now reads as follows;

We used the mesoscale Hydrologic Model (mHM; Samaniego et al., 2010; Kumar et al., 2013) (version v-5.13.2-dev0) to simulate domain-wide root zone (0-2 m) soil moisture conditions and streamflow, which we used as an additional hydrologic constraint for validating basin-scale hydrology at major outlets. mHM is a spatially distributed hydrological model based on numerical representations of dominant hydrological processes. The model is driven by hourly to daily meteorological forcings, which include precipitation, temperature, and potential evapotranspiration, and accounts for major hydrological processes like snowmelt and accumulation, canopy storage, evapotranspiration, surface runoff and flood routing, three-layer soil moisture content, and subsurface storage. To represent spatial variability of inputs and state variables, the model uses three different spatial resolutions, namely (in order of fine to coarse resolution) Level-0 (L_0: small-scale morphology) to represent the main terrain features, geological features, land cover, and soil properties; Level-1 (L_1: mesoscale hydrology) to represent the dominant hydrological processes; and Level-2 (L_2: large-scale meteorology) to describe the variability of meteorological forcings. The model harmonizes the data internally using the multiscale parameter regionalization (MPR; Samaniego et al., 2010). MPR links model parameters at L_1 to their corresponding ones at L_0 using non-linear transfer functions that couple catchment characteristics with global (calibration) parameters to regionalize model hydrologic parameters at L_0 and link them to their corresponding values at L_1 using upscaling operators such as arithmetic mean, geometric mean, and harmonic mean (MPR; Livneh et al., 2015). With this technique, mHM achieves quasi scale-invariant parameters that enable the model to preserve the spatial variability of state variables and conserve mass balance (Samaniego et al., 2010; Samaniego et al., 2011; Kumar et al., 2013; Samaniego et al., 2013).

Reviewer Comment #9:

Define NSE on first use (Results 3.1.2). Give the range and interpretation (≈1 perfect; ≈0 equals mean-flow benchmark; < 0 worse than mean).

Author response:

Thank you for this suggestion. In line with a similar comment from Reviewer #2, we have moved this section to the supplementary text so that we give a detailed description of the model performance statistics without it occupying a large part of the main manuscript. Therein, we describe NSE and how it is calculated as well as the interpretation of different NSE values mean. In the main article, we refer the readers to the supplementary text for more details about the model performance in simulating streamflow. We have thus collapsed Section 3.1.2 into a paragraph is Section 3.1.1 which reads as follows:

Regarding streamflow performance, the model shows good and spatially consistent skill across the entire modelling domain and thus provides a reliable basis for analysing soil moisture dynamics. We evaluated daily discharge at 168 gauging stations. During calibration, the mean Nash-Sutcliffe Efficiency (NSE) across stations was 0.62, with 80% of stations achieving NSE ≥ 0.5 (a commonly used benchmark for satisfactory streamflow simulation). Model performance during the validation period was also comparatively good, with a mean NSE of 0.63 and 83% of stations recording NSE ≥ 0.5. The full details of the streamflow evaluation, including the NSE definition, are provided in Supplementary Text S2 (Figure S1).

Reviewer Comment #10:

Lines 470-475, restate the minimum overlap area rule used to merge adjacent months into one multi-temporal event. Or how do you define duration. This clarifies whether a brief wet interlude (e.g., March–April 2017) splits or does not split an event.

Author response:

Indeed. We have added this paragraph for added context.

However, the wet spell did not split the event because the month-to-month overlap in the drought area still exceeded the 640 km^2 merging threshold; thus, the drought remained as a single multi-temporal event.

Reviewer Comment #11:

MPR, what is the full name.

Author response:

MPR is an abbreviation for Multiscale Parameter Regionalization. It appears in line 164 of the manuscript.

Reviewer Comment #12:

Terminology clarity. The manuscript uses three different concepts that contain the word "calibration":

- A 5-year warm-up: is it 1965–1969, if yes, why exclude 1970 as a calibration year for drought analysis?
- Excluding 1970 as a calibration year for drought analysis when forming decades (hence using 1971–1980),
- Streamflow parameter calibration (2000–2023) vs validation (1970–1999).

Please clarify these three meanings to avoid confusion.

Author response:

We agree that the current use of the term "calibration" may cause confusion. The warm up period is indeed 1965–1969. The exclusion of 1970 from the drought analysis is not related to model calibration, but to the construction of the SPEI. SPEI is based on accumulated water-balance anomalies over preceding months. As a result, January 1970 does not contain valid SPEI-1 values, and January–March 1970 do not contain valid SPEI-3 values. The year 1971 is therefore the first year with complete SPEI values for all months. For this reason, decades are defined starting from 1971 (i.e. 1971–1980).

Accordingly, we have updated the language in the revised manuscript to distinguish the use of these terminologies.

2: Responses to Reviewer #2:

Reviewer Comment #1: Overall assessment

This study analyzes the root-zone soil moisture dynamics in Belgium from 1970 to 2020, focusing on the severity and persistence of droughts during the 2011–2020 period. It highlights the unprecedented nature of these droughts and evaluates the limitations of precipitation-based drought indices (such as SPEI) in drought assessments. The paper proposes using root-zone soil moisture as a more effective drought monitoring indicator and underscores the increasing frequency and persistence of drought events in the context of climate change. The work is valuable and substantial, but its scientific novelty is limited, and some methodological and analytical aspects need further clarification or strengthening.

Specifically, the manuscript currently lacks critical quantitative evidence to support two core claims: (1) that root-zone soil moisture provides added operational value over precipitation-based indices (SPEI), and (2) that the chosen reconstruction approach (mHM) is preferable to widely used soil-moisture products (e.g., ERA5-Land). I recommend the authors (a) perform an event-based contingency analysis comparing mHM-RZSM and SPEI drought events (onset, termination, duration, severity) and report hit/miss/false-alarm statistics, and (b) provide a direct inter-comparison with at least one widely used soil-moisture product to quantify time-series agreement and event-detection differences.

Author response:

The manuscript presents the first high-resolution, multi-decadal reconstruction of root-zone soil moisture droughts for Belgium, where long-term in situ soil-moisture observations or any studies of the kind we did are not currently available. The primary scientific contribution lies in combining this reconstruction with a spatially explicit drought-event analysis and a mechanistic comparison between precipitation-based and soil-moisture-based drought indicators. This combined approach allowed us to determine how the droughts experienced in recent years compared with those experienced in the last 50 years and enabled us to understand their rarity. No study of this kind has been done in Belgium.

We agree that distinguishing drought types and explaining the conceptual advantages of root-zone soil moisture is important. We would like to point out that the current manuscript already includes a detailed discussion of these points (Lines 88–103), where we explicitly describe the limitations of precipitation and temperature-based indices such as SPEI for capturing agricultural drought conditions. In the manuscript we report that the precipitation-based indices are not able to represent the vertical distribution of water in the root zone or vegetation water stress, the nonlinear and lagged response of soil moisture to precipitation, and the soil-moisture memory effect that often leads to prolonged drought persistence. We feel that these points reinforce the added operational value that soil moisture provides. As for the suggestion to perform a contingency analysis (onset, duration, severity) our manuscript already compares mHM-RZSM and SPEI droughts in Section 3.5 of the manuscript. Our analysis show that: (i)

SPEI-based droughts terminate earlier than soil moisture droughts which could result in premature conclusions about the termination of droughts (ii) SPEI-based droughts underestimate the severity of droughts compared to soil moisture droughts. By analyzing the three largest droughts during the period of analysis, our analysis also showed that SPEI at short accumulation times (e.g., SPEI-1) is highly responsive to short-lived rainfall deficits and surpluses that may not immediately alter root-zone storage; this sensitivity captures meteorological conditions which differ from soil moisture conditions that integrate past deficits through slow infiltration and plant uptake.

In Section 3.5 of the revised manuscript, we have included other metrics like the number of months under drought conditions during the major drought events. We discuss these changes in our response to Comment #7.

On the merits of using mHM over other soil moisture products, we used mHM in our study to generate soil-moisture fields at high spatial resolution. We generated soil moisture fields at $1/32^0$, which is a resolution high enough to account for the heterogeneities that characterize land surface conditions that affect soil moisture. This higher resolution is better suited for drought analysis since it takes into account the heterogeneity in soils, land use and land cover all of which vary strongly at short distances. ERA5-Land is indeed a valuable reanalysis product but it is better suited to for large-scale applications (continental to global scales). The $\approx$9 km resolution of ERA5-Land limits its suitability for local-scale drought monitoring as the resolution generalizes land surface conditions. Due to this generalization of land surface heterogeneity, coarse-scale products like ERA5-Land can produce biases in surface water and energy fluxes (Crow et al., 1999). To illustrate, if we used the 9 km ERA5-Land over Belgium (about 30,000 sq. km in area), we would obtain about 370 grid cells. On the other hand, our model resolution generates 3,100 grid cells, almost ten times more than ERA-5 Land. Our model resolution is thus better at differentiating fine-scale heterogeneities in soil moisture conditions and deriving local-scale variability.

One would also ask why we did not use remote sensing (RS) soil moisture products. While RS offers an alternative source of soil moisture, RS-derived soil moisture only measures water content in upper few centimetres of the soil profile and have low quality under certain surface conditions such as dense vegetation, frozen soils and mountainous terrain. The data can also often be missing due to retrieval conditions or satellite revisit times (Wang et al., 2011; Peng et al., 2017).

Reviewer Comment #2:

Please clarify the definition of the root-zone layer used in this study (0–0.5 m). Why was this depth chosen, and how representative is it across different vegetation types and land-cover conditions in Belgium? Considering the variability of underlying surfaces could help assess the robustness of the results.

Author response:

In the manuscript we have clarified that the root-zone depth was limited to 0–0.5 m because in most areas of the country, groundwater occurs just below this depth (Lines 263-264), thus including deeper layers would risk mixing soil-moisture dynamics with groundwater storage. We agree that representativeness across vegetation types can vary. For clarity, We have added a sentence justifying that this depth corresponds to the dominant rooting depth of shallow-rooted croplands, grasslands, and some vegetation types that are disproportionally sensitive to moisture availability in the unsaturated zone and thus more susceptible to soil moisture droughts.

Reviewer Comment #3:

The Introduction successfully establishes the severity of drought in Belgium and correctly identifies the scientific gap regarding the limitations of precipitation-based indices. However, the section's structure and balance require revision to maximize clarity and scientific impact.

- The lengthy, detail-heavy descriptions of the 2011, 2018–2019, and 2022 droughts (Lines 46–74) read like an event chronicle. This narrative must be significantly condensed to focus only on the key messages that motivate the need for a long-term assessment, ensuring the scientific gap is presented more prominently.
- Conceptual Distinctions: The discussion on drought monitoring (Lines 79–103) should be enhanced by more explicitly distinguishing meteorological/hydrological drought from agricultural drought. This includes clarifying why traditional indices like SPEI are limited and emphasizing the superior conceptual role of root-zone soil moisture (RZSM) for capturing plant water stress.
- The final paragraph (Lines 104–116) prematurely introduces technical specifics (e.g., mHM model, offline forcings, SMI derivation via percentile ranking). These details should be relocated entirely to the dedicated Methods section.

Author response:

Thank you for these comments. We have rewritten the introduction to describe the drought events more concisely without losing important details. On the second point, we would like to point out that the conceptual distinction between meteorological and agricultural drought is already addressed in Lines 84–103, where we explicitly describe the limitations of precipitation and temperature-based indices such as SPEI for capturing agricultural drought conditions. In the manuscript we report that the following limitations of precipitation-based indices:

- They do not represent the vertical distribution of water in the root zone or vegetation water stress

- The response of soil moisture to rainfall is lagged and non-linear, which is not reflected by precipitation-based indices
- They do not account for soil-moisture memory effect that often leads to prolonged drought persistence

We feel that these points reinforce the added operational value that soil moisture provides.

On the final comment, we have modified the last paragraph of the introduction to mainly focus on the objectives of the study, and moved the other details to the Methods section.

Reviewer Comment #4:

The authors provide a compelling and well-referenced justification for prioritizing the Pearson correlation coefficient to assess the temporal agreement of standardized soil moisture anomalies. However, two critical components are missing for a complete validation of the model's skill in the context of this drought study. To fully characterize the model's performance beyond just temporal consistency, the authors should consider reporting an appropriate error metric, such as the Unbiased Root Mean Square Error (ubRMSE). The ubRMSE is ideally suited for this validation context, as it quantifies the error component related to the model's random fluctuations and timing errors, while excluding the systematic absolute bias that is deliberately factored out by the standardization approach. Crucially, given that the study's goal is drought analysis, the evaluation should include an explicit assessment of the model's ability to accurately represent drought conditions as observed by the in situ data. For instance, comparing the model's and in situ data's ability to correctly classify dry/drought days based on an established threshold (e.g., the 20th percentile), assessing the correlation or error between the model-derived soil moisture index (SMI) and an index derived from the in-situ data.

Author response:

Thank you for the suggestion. We agree that correlation alone may not fully describe model performance and that an error-based metric and an explicit drought-detection evaluation can strengthen the validation in the context of drought analysis. However, the suggested use of ubRMSE computed in physical units is not directly applicable here because the in situ sensors report volumetric soil moisture (m^3/m^3), whereas the model output used in this study is soil-water storage (mm per soil layer). This is why we first standardized both datasets, as we explain in Section 2.2.3. A direct ubRMSE would therefore compare quantities with different units. We applied correlation to the standardized values since it directly assesses temporal agreement and anomaly timing, which is what we are interested in from a drought analysis perspective. To make our approach consistent with good-practice recommendations for soil-moisture validation, we have added an explicit drought classification assessment comparing the model and in situ data in terms of their ability to identify droughts using a percentile threshold (e.g., the 20th percentile) on daily values. Here we assess the model in terms of drought hits, misses, false alarms and correct negatives. To preserve the flow of the main manuscript and to avoid using too many metrics and to address the point from Reviewer #1 who commented that we have used quite a number of statistical metrics in the manuscript, we have implemented these analyses in the Supplementary text.

We have added the following text to specifically assess the skill of the model to reproduce observed drought days.

To evaluate how well the model simulates drought conditions, we investigated the drought-day detection skill (when the observed standardized anomaly fell below its 20th percentile) by counting hits (H; days when both model and observations indicate drought), misses (M; observed drought days not flagged by the model), false alarms (F; days flagged as drought by the model but not by the observations), and correct negatives (C; days when both indicate non-drought). (The methodology is described in more detail in Supplementary Text S1). From this analysis, we found that the model shows high skill in reproducing observed drought conditions, as it was able to detect 74% of observed drought days from the 21 stations. The false alarm rate was also only 5%, while the mean F_1 score ($2H/(2H + F + M)$) which summarizes the balance between misses and false alarms) was 75%. These metrics indicate that the model can be applied to study droughts. We attribute the differences in detecting droughts to the scale mismatch between mHM soil moisture, which represents average conditions over a grid cell, and the highly localized nature of point in situ measurements.

Reviewer Comment #5:

The detailed presentation of streamflow simulation performance (Section 3.1.2) is robust, but its inclusion as a standalone subsection immediately following the core Soil Moisture evaluation (3.1.1) is structurally misleading and requires clarification. Based on the Methods section (Lines 249–257), the streamflow analysis serves primarily as an internal calibration and performance check of the hydrological modeling framework, not a direct validation of the primary variable of interest (soil moisture). Therefore, the detailed streamflow analysis (Section 3.1.2) should be relocated and significantly condensed. This content belongs logically in a dedicated subsection within the Methods (e.g., Model Calibration) to briefly demonstrate the adequacy of the modeling framework, rather than occupying a prominent position in the main Results section. If the authors insist on keeping the streamflow results prominent, they must establish a clear logical link showing how the successful streamflow calibration improves or validates the soil moisture simulations.

Author response:

Thank you for this comment. As also remarked by the first Reviewer, we have condensed the section and moved the figure and the description and formulation of NSE to the Supplementary Text S1). The description of streamflow calibration now only appears as a paragraph at the end of Section3.1. On establishing a clear link between streamflow calibration, we respectfully note that the manuscript already states that soil moisture and streamflow are intrinsically linked through the catchment water balance (Methods, Section 2.2.3, Lines 251–269). The streamflow evaluation is therefore presented as an independent consistency check of the simulated water balance, which implicitly includes soil moisture storage and depletion processes. The section is included to add credibility to the modelled soil moisture results.

Reviewer Comment #6:

While Section 3.2 effectively uses the Total Drought Magnitude (TDM) index to characterize the evolution of drought events and identifies a compelling increase in frequency and severity post-2011, the claims of 'three distinct drought regimes' and a 'significant shift' are primarily descriptive, relying heavily on the narrative of the top ten events (Figure 5). This conclusion lacks robust, decade-spanning quantitative evidence. The author should consider providing a concise table or figure reporting key summary statistics for each complete decade (e.g., 1971–1980, 1981–1990, etc.), such as the total cumulative TDM or the mean annual number of drought days. This quantitative evidence is essential to lend statistical rigor to the core finding of the significant shift.

Author response:

We would like to note that the manuscript already includes a dedicated statistical decadal analysis in Sections 3.3 and 3.4, which provide the decadal-scale quantitative evidence supporting the identification of the distinct drought regimes. This includes cumulative drought duration (months per decade), drought severity composition, spatial drought exposure, and bootstrapped confidence intervals for drought months per decade (Figures 6–8), which show that 2011–2020 is statistically distinct from preceding decades. We acknowledge that this evidence is presented later in the Results and is not explicitly referenced in Section 3.2. To ensure the connection is clearer to readers, we have added an introductory paragraph at the beginning of the Results section so describing what each section of the manuscript describes.

Reviewer Comment #7:

The central finding that SMI exhibits stronger persistence and SPEI underestimates deficits (L448, L455) is crucial but currently relies on qualitative descriptions (e.g., 'more persistent,' 'underestimate'). It is better to supplement the discussion with quantified persistence metrics. For example, for the three major events analyzed (1975–1977, 2016–2017, 2018–2019), the author should consider reporting the following metrics: (a) the maximum duration (in months) of the moderate drought for SMI, SPEI-1, and SPEI-3; and (b) the total number of months under moderate drought for each index during the event period. These data are essential for empirically validating the conclusion that SMI has higher inertia than the precipitation-based indices.

Author response:

We thank the reviewer for this comment. We agree that explicit quantitative persistence metrics strengthen this comparison. In the revised manuscript we now report two metrics for each of the three major events (1975–1977, 2016–2017, 2018–2019) and for SMI, SPEI-1, and SPEI-3: (i) drought persistence as the maximum number of *consecutive* months under at least moderate drought; and (ii) cumulative drought exposure as the total number of months meeting the moderate-drought criterion within the event window, regardless of whether those months are consecutive. These metrics directly quantify drought inertia and empirically support the conclusion that SMI exhibits stronger persistence than precipitation-based indices.

To examine how precipitation-based drought indicators reflect land-surface moisture stress, we compared SMI and SPEI patterns during the three most severe soil moisture drought events ranked by total drought months (TDM): 1975–1977, 2016–2017, and 2018–2019. Because SMI is computed on a monthly timescale, we derived the climatic water balance (precipitation minus potential evapotranspiration) from E-OBS and calculated SPEI at one- and three-month accumulation periods. Pixel-wise SPEI-1 and SPEI-3 were computed using the `SPEI` *package Vonk, 2024. We limited the accumulation period to three months because this timescale is currently used in operational drought monitoring in Belgium. Since SPEI is anomaly-based rather than percentile-based, we associated SPEI =* -1.0 *with SMI =* 0.2 *to represent the threshold for at least moderate drought, following the drought severity guidelines of Svoboda et al., 2002. We evaluate differences between indices in terms of (i) anomaly magnitude, (ii) drought persistence (maximum number of* consecutive *months under at least moderate drought within an event), and (iii) cumulative drought exposure (total number of months, not necessarily consecutive, under at least moderate drought within the event window).*

In terms of anomaly magnitude, SMI generally indicated stronger and longer-lasting deficits than SPEI-1 and, to a lesser extent, SPEI-3 (Figure 8). SPEI-1 responds strongly to short-lived precipitation anomalies that may not immediately translate into changes in root-zone storage. By design, SPEI-3 smooths some of the short-term variability in SPEI-1 and more closely resembles the temporal evolution of soil-moisture anomalies, but still tends to underestimate deficit magnitude relative to SMI over our domain (Figure 8). Among the three events, SMI indicated the strongest soil moisture deficits during 2016–2017 (with SMI approaching zero), which is not reflected in either SPEI-1 or SPEI-3. Although the 2016–2017 drought was interrupted by intermediate wet conditions during March and April 2017, leading to partial recovery, this wet spell did not split the event because the month-to-month overlap in drought area remained above the 640 km^2 *merging threshold, and the drought therefore remained a single multi-temporal event.*

We also found that SMI-based droughts exhibited higher persistence than SPEI-based droughts. Median persistence for SMI was 9 months in 1975–1977, 6 months in 2016–2017, and 7 months in 2018–2019 (Table **??***). In comparison, SPEI-1 shows much shorter median* *persistence (3 months in 1975–1977 and 2 months in both 2016–2017 and 2018–2019), while SPEI-3 is closer to SMI but remains lower (7 months in 1975–1977 and 5 months in both 2016–2017 and 2018–2019).*

The same pattern is evident for cumulative drought exposure. Median cumulative exposure for SMI was 10 months in 1975–1977 and 2016–2017 and 8 months in 2018–2019, compared with 4, 6, and 3 months for SPEI-1 and 7, 8, and 6 months for SPEI-3 (Table 2, additional maps in Supplementary Text S3, Figures S3 and S4. Figures S3 and S4 correspond to Figures 2 and 3 below).

The same pattern is evident for cumulative drought exposure. Median cumulative exposure for SMI was 10 months in 1975–1977 and 2016–2017 and 8 months in 2018–2019, compared with 4, 6, and 3 months for SPEI-1 and 7, 8, and 6 months for SPEI-3 (Table 2). This systematically longer persistence and exposure in SMI indicate higher soil-moisture inertia, suggesting sensitivity to the sequencing of meteorological anomalies as well as their magnitude. Additional analysis on the patterns of SMI and SPEI recovery is presented in Supplementary Text S3.

Table 2: Quantified drought persistence and cumulative drought exposure during the three major events (1975–1977, 2016–2017, and 2018–2019) for SMI, SPEI-1, and SPEI-3. Values represent the median across all grid cells in the domain. For cumulative exposure, we additionally report the spatial maximum (in brackets), which represents the time until the last grid cells recover above the moderate-drought threshold within the event window.

	Max. consecutive months (persistence)			**Total drought months (cumulative exposure)**		
Index	1975–1977	2016–2017	2018–2019	1975–1977	2016–2017	2018–2019
SMI	9	6	7	10 (16)	10 (15)	8 (10)
SPEI-1	3	2	2	4 (7)	6 (8)	3 (5)
SPEI-3	7	5	5	7 (13)	8 (12)	6 (9)

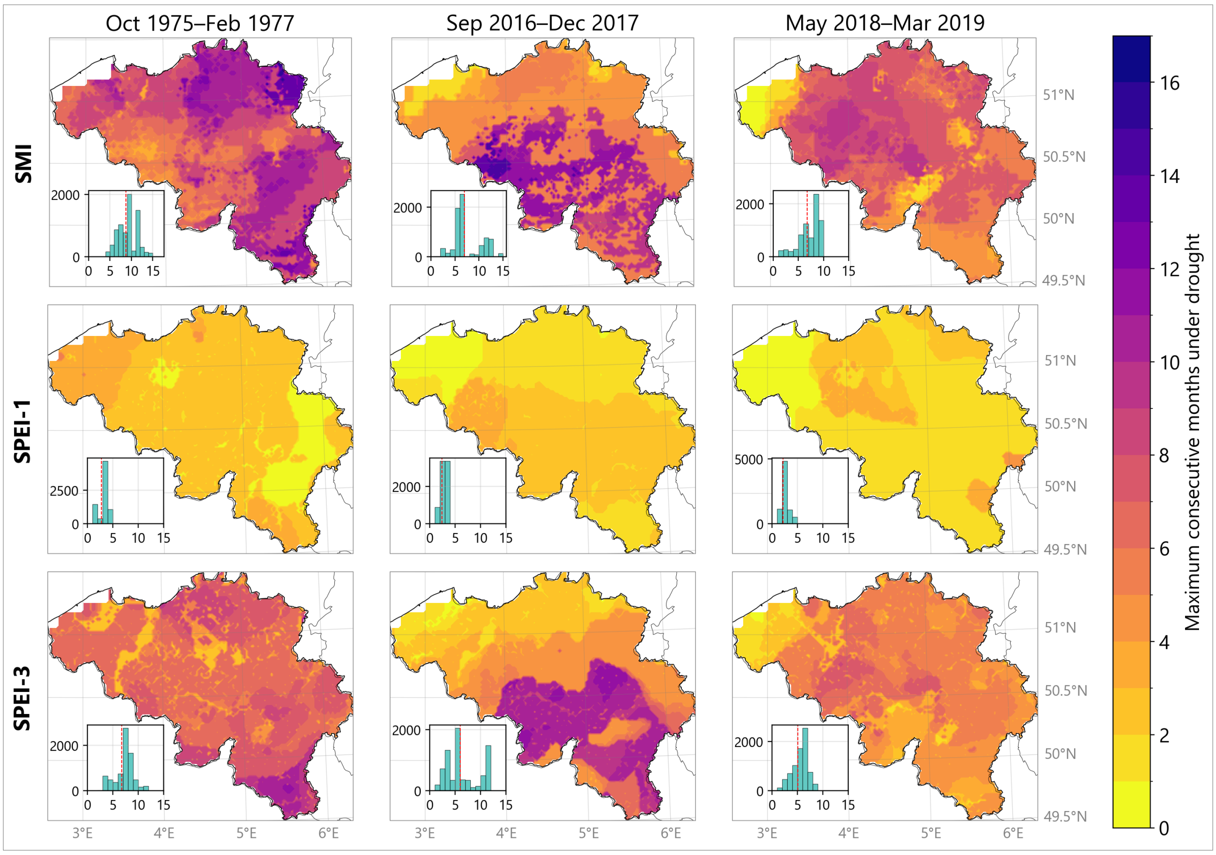

Figure 2: Pixel-wise drought persistence for the three largest droughts during the period of analysis for SMI, SPEI-1 and SPEI-3. Persistence is computed as the longest consecutive sequence of months where at least moderate conditions exist in each gridcell. The inset histogram shows the distribution per-pixel drought persistence. The vertical red line shows the mean persistence.

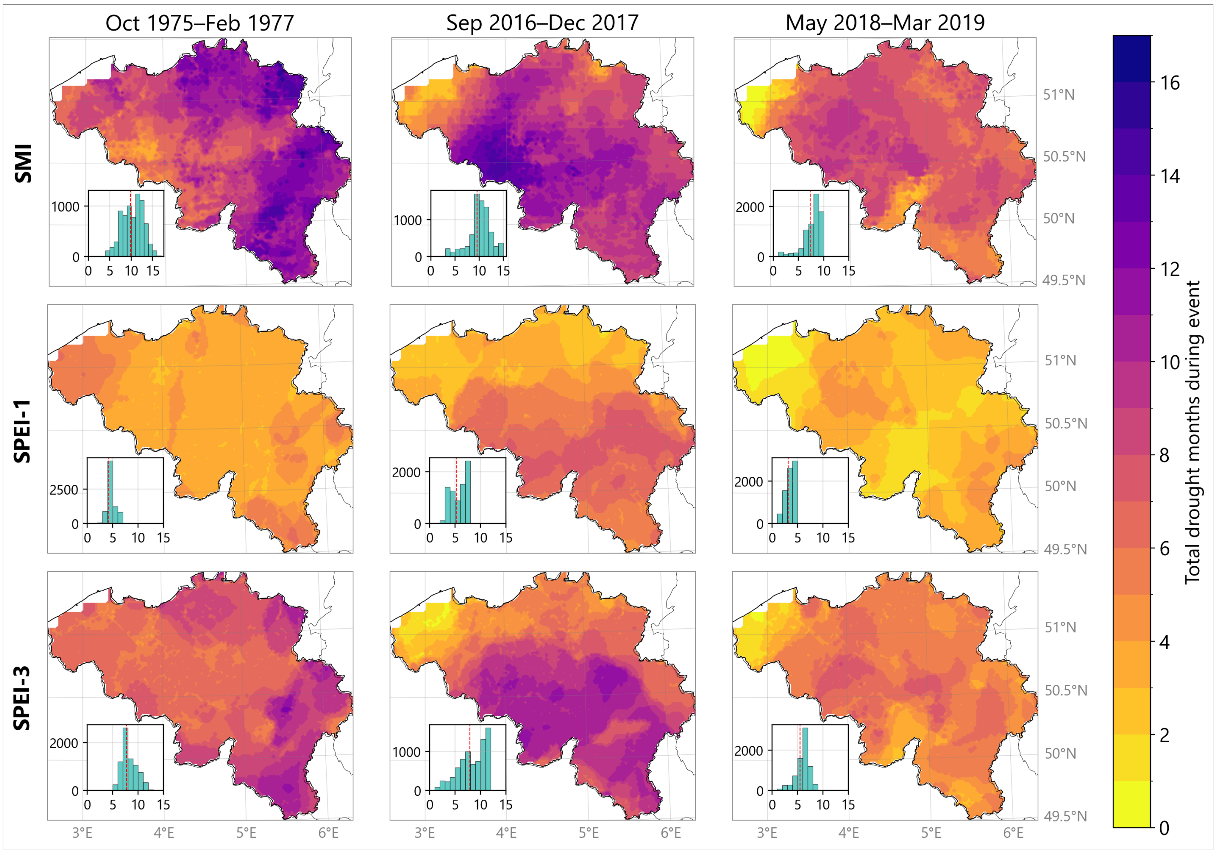

Figure 3: Pixel-wise cumulative drought exposure for the three largest droughts during the period of analysis for SMI, SPEI-1 and SPEI-3. Persistence is computed as the total number of months, not necessarily consecutive, under at least moderate drought drought in each grid cell. The inset histograms are similar to Figure 2 but for cumulative exposure.

Reviewer Comment #8:

The Discussion section could be strengthened by elaborating on the broader implications of this work. How can the findings be applied to larger regions or integrated into operational drought monitoring and management strategies?

Author response:

We have added more depth to the Discussion to incorporate the suggestions. In the revised manuscript we have added a section on broader implications and how the work can be integrated into operational management. We presented the updated text below.

Our findings are relevant beyond Belgium because the workflow used in this study can be transferred to other regions provided that the meteorological forcing is available at appropriate resolution, a hydrological or land-surface model is parameterized to represent soil-water storage, and consistent long-term simulations can be produced. Extending the analysis to other domains would allow the same drought dynamics addressed in this study to be evaluated under different climate gradients, soil, land-cover conditions, and management regimes. From an operational perspective, the results support a monitoring strategy that complements precipitation-based indices with soil-moisture-based indicators rather than interchanging them. As we have shown, precipitation-based indices are useful for tracking meteorological anomalies and can provide early signals of emerging drought risk, but they may not capture persistent impacts when land-surface memory sustains root-zone deficits after rainfall resumes. In an operational system, precipitation-based indices can be used for early warning, while a root-zone soil moisture drought indicator is better utilized to track agricultural drought development and recovery and to assess when conditions have returned to normal in the soil profile. These outputs can be integrated into management decisions by linking drought phase and persistence to sector-relevant decisions. For example, soil-moisture drought persistence is directly relevant for agricultural advisories that inform planting and irrigation planning and signaling crop yield risk and the risks associated with the occurrence of wildfires or floods that can occur due to seasonally saturated soils. Slow recovery in soil and catchment storage after meteorological drought can also inform water supply preparedness and groundwater management, since water resources often show a delayed return to normal conditions (Yang et al., 2017). For inland navigation and low-flow management, combining soil moisture drought information with streamflow indicators can help distinguish short, transient precipitation deficits from longer-lasting, storage-driven drought conditions. In practice, monthly updates of a root-zone soil moisture drought map, paired with precipitation-based indices, would support earlier identification of drought evolution and lead to more realistic expectations for recovery following intermittent wet periods (Van Loon et al., 2024).

Reviewer Comment #9: Minor comment

The content in Section 2.2.1 regarding the lack of long-term in situ soil moisture data in Belgium and the subsequent expansion of the model domain is more appropriate for Section 2.2.3. The Input data section should focus primarily on datasets used to drive the model. Standardize the capitalization of section titles, as the current usage is inconsistent. Please standardize the terminology and use 'in situ' consistently throughout the manuscript; some occurrences currently use 'in-situ'. While the term 'Drought Persistence' is used in the text, the definition corresponds more closely to cumulative drought exposure or duration. The Conclusion should focus more on synthesizing key findings and scientific contributions, and discussion-type content should be reduced. Please standardize the terminology and use 'root-zone' consistently throughout the manuscript. It is recommended to present the other drought indices, currently discussed in the Discussion section, within the Methods section.

Author response:

Thank you for this important comment. We believe it will improve the readability of the manuscript. We have shifted the description of the in situ data to the last paragraph of Section 2.2.1 since it also describes some aspects of the model set up. We have also corrected the terminologies as suggested and made the Conclusions section much more compact.

Reviewer Comment #10: Special comments

- Line 36: Please correct 'agricultural' to 'agricultural'.
- Line 160: Please remove the extraneous semicolon ';'.
- Line 425 and Line 429: Please correct 'Figure ??(a)' and 'Figure ??(b)'.
- Line 501: Correct 'cosysytems' to 'ecosystems'.

Author response:

Thank you for pointing out these errors. We have corrected typographical and formatting issues, and explicitly specified that temperature refers to air temperature.

References

Crow, Wade T and Eric F Wood (1999). "Multi-scale dynamics of soil moisture variability observed during SGP'97". In: *Geophysical Research Letters* 26.23, pp. 3485–3488.

Kumar, Rohini, Luis Samaniego, and Sabine Attinger (2013). "Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations". In: *Water Resources Research* 49.1, pp. 360–379.

Livneh, Ben, Rohini Kumar, and Luis Samaniego (2015). "Influence of soil textural properties on hydrologic fluxes in the Mississippi river basin". In: *Hydrological Processes* 29.21, pp. 4638–4655.

Peng, Jian, Alexander Loew, Olivier Merlin, and Niko EC Verhoest (2017). "A review of spatial downscaling of satellite remotely sensed soil moisture". In: *Reviews of Geophysics* 55.2, pp. 341–366.

Samaniego, Luis, Rohini Kumar, and Sabine Attinger (May 2010). "Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale". en. In: *Water Resources Research* 46.5, 2008WR007327. ISSN: 0043-1397, 1944-7973. DOI: `10.1029/2008WR007327`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2008WR007327` (visited on 02/13/2025).

Samaniego, Luis, Rohini Kumar, and Conrad Jackisch (2011). "Predictions in a data-sparse region using a regionalized grid-based hydrologic model driven by remotely sensed data". In: *Hydrology Research* 42.5, pp. 338–355.

Samaniego, Luis, Rohini Kumar, and Matthias Zink (Feb. 2013). "Implications of Parameter Uncertainty on Soil Moisture Drought Analysis in Germany". en. In: *Journal of Hydrometeorology* 14.1, pp. 47–68. ISSN: 1525-755X, 1525-7541. DOI: `10.1175/JHM-D-12-075.1`. URL: `http://journals.ametsoc.org/doi/10.1175/JHM-D-12-075.1` (visited on 02/13/2025).

Svoboda, Mark, Doug LeComte, Mike Hayes, Richard Heim, Karin Gleason, Jim Angel, Brad Rippey, Rich Tinker, Mike Palecki, David Stooksbury, David Miskus, and Scott Stephens (Aug. 2002). "THE DROUGHT MONITOR". en. In: *Bulletin of the American Meteorological Society* 83.8, pp. 1181–1190. ISSN: 0003-0007, 1520-0477. DOI: `10.1175/1520-0477-83.8.1181`. URL: `https://journals.ametsoc.org/doi/10.1175/1520-0477-83.8.1181` (visited on 03/27/2025).

Van Loon, Anne F, Sarra Kchouk, Alessia Matanó, Faranak Tootoonchi, Camila Alvarez-Garreton, Khalid EA Hassaballah, Minchao Wu, Marthe LK Wens, Anastasiya Shyrokaya, Elena Ridolfi, et al. (2024). "Drought as a continuum—memory effects in interlinked hydrological, ecological, and social systems". In: *Natural Hazards and Earth System Sciences* 24.9, pp. 3173–3205.

Vonk, Martin A. (2024). *SPEI: A simple Python package to calculate and visualize drought indices*. DOI: 10.5281/ZENODO.10816741. URL: https://zenodo.org/doi/10.5281/zenodo.10816741 (visited on 07/02/2025).

Wang, Aihui, Dennis P Lettenmaier, and Justin Sheffield (2011). "Soil moisture drought in China, 1950–2006". In: *Journal of Climate* 24.13, pp. 3257–3271.

Yang, Yuting, Tim R McVicar, Randall J Donohue, Yongqiang Zhang, Michael L Roderick, Francis HS Chiew, Lu Zhang, and Junlong Zhang (2017). "Lags in hydrologic recovery following an extreme drought: Assessing the roles of climate and catchment characteristics". In: *Water Resources Research* 53.6, pp. 4821–4837.

We thank the reviewer for providing feedback to improve our manuscript. Below we provide a point-by-point response to the minor comments.

Standardize and refine the formatting of all tables to ensure consistency, including the format of table titles.

We have now reformatted all tables to maintain consistency. This includes shortening and standardizing table titles, removing colour, and applying a uniform layout across all tables.

The first three paragraphs of the Introduction mainly describe drought events in Belgium and their associated impacts. While the background and motivation of the study are important, this part appears somewhat lengthy and occupies a relatively large proportion of the Introduction. It may be worth considering streamlining this section to improve the overall balance and conciseness of the manuscript.

We have now condensed the narrative of the drought events in Belgium to a maximum of one paragraph. We have retained only text that provides context to the reader.

While the authors discuss the limitations of precipitation-based indices for representing agricultural drought, it would improve clarity to briefly define meteorological, agricultural, and hydrological drought in the Introduction. Since these three types of drought are mentioned in the manuscript, a short clarification of their differences would help readers better understand the context.

We agree. We have now added definitions of meteorological, agricultural, and hydrological drought to the Introduction to clarify the distinction between these drought types.

4. The first part of the Discussion appears somewhat lengthy and contains a substantial amount of background information. Streamlining this section and focusing more directly on the interpretation of the study's findings would improve the clarity and readability of the manuscript.

In the revised manuscript we have shortened this section and retained only the broader European context needed to situate the results.

Lines 525–535: The manuscript notes that the results are consistent with the intensification of droughts across Europe in the 21st century and then lists several individual studies in detail. It may not be necessary to describe each study separately in the text. If the authors wish to retain these references,

presenting them in a concise summary (e.g., a table or a shorter synthesis) could make the information more accessible while reducing the textual burden.
We agree. We have replaced the study-by-study discussion with a shorter synthesis that retains the key references while reducing the textual burden.

Lines 540–558: This section discusses large-scale atmospheric circulation patterns and anthropogenic warming as drivers of drought dynamics. While this information is relevant in a broader context, its direct connection to the results of this study appears somewhat limited. The authors may consider shortening this part or relocating some of the background discussion to the Introduction to maintain a clearer focus in the Discussion section.

We have also shortened this section and reframed so that it provides only the broader physical context needed to interpret the recent sequence of severe droughts identified in Belgium.

With these revisions, we believe that the revised manuscript addresses all the comments raised.