Dear Xing Yuan,

Thank you for the opportunity to submit a revised version of our manuscript *Asymmetric impact of groundwater use on groundwater droughts*, to be considered for publication in Hydrology and Earth System Sciences. This manuscript has Manuscript ID hess-2020-22. We are very grateful to you and the reviewers for their careful reading and insightful comments. Incorporating these changes has improved the clarity of the manuscript greatly.

We have made some significant changes to address the key concerns of both reviewers requesting additional explanation in the method and results sections. The concerns are associated with the uncertainty when using standardised indices for drought studies and in particular for groundwater studies. To address these comments, we have gathered a number of explanatory examples and rephrased sections in order to improve the descriptions and discussion of results. The major changes we have incorporated are as follows:

1. Further and detailed explanations of the used correlation analysis and identification of human influence in standardised groundwater time series including an illustrated example.
2. Addressing the variability in presented results that included a new data analysis, improved the description, and discussion of the old and new results.

In addition, we have carefully considered the suggested alternative methods provided by the reviewers and we have made every attempt to address their concerns in a revised manuscript. We provide a detailed point-by-point response to each of the reviewer's comments below. Line numbers refer to the revised manuscript.

Thank you for considering a revised version of our manuscript and we hope to hear from you soon.

Yours sincerely,

Doris Wendt (on behalf of co-authors)

Editor comments to the Author:

Dear Authors,

I would like to thank for your responses to the comments. It would be useful to upload the revised manuscript, including a track changes version.

Looking forward to your revisions.

Regards
Xing Yuan

Thanks for the opportunity to upload the revised manuscript. We have addressed the reviewers' comments and revised the manuscript in line with their suggestions. Please find the detailed point-by-point response below.

Reviewer(s)' Comments to Author:

Reviewer: 1

This study uses a framework that consists of two approaches, and conducted a case study in UK to investigate the impact of groundwater use on groundwater droughts. Generally, the manuscript is well organized with clear logic, before I recommend it for publication, major improvements are still needed, particularly for the method they used for recognizing the presence or absence of human-influence on groundwater. Please find my specific comments below:

Thank you for your comments. We are relieved to hear that the structure and logic of the paper was well-received. We thank you for your careful reading and have modified the manuscript along the lines of your suggestions that greatly improved the clarity of the manuscript.

General comments:

R1C1: Lines 171-172: 'In this study, the presence or absence of human-influence on groundwater was determined in relation to the lowest SPIQ-SGI correlation of each near-natural reference cluster'. I think it is questionable to determine the presence or absence of human influence depending on the correlation analysis. For example, for a certain site, SGI is best correlated to SPI at short time scales. Due to human interference, the drought duration indicated SGI may become longer, leading to SGI best correlated with SPI at longer time scales. The increased time scale of SGI does not necessarily corresponds to reduced correlation of SPIQ-SGI, and the correlation may also increase. Moreover, considering the significant spatial heterogeneity of groundwater (e.g., groundwater of the monitoring sites may show different patterns from reference sites), it would be better to recognize human influences by analyzing the temporal variation of groundwater for the same site (e.g., compare the statistics of groundwater among different decades). The uncertainty derived from the method for recognizing the presence or absence of human-influence on groundwater needs to be discussed.'

We understand the concern raised in comment 1 regarding the correlation analysis, but disagree with the assertion that the inference of abstraction effects cannot be inferred through analysis of correlations between SPI_Q and SGI, and their specific assertion that correlations may increase with increased abstraction effects. We set out our arguments in more detail below and have extended section 3.2.2 as shown below (Lines 164:180) to reflect our response. We agree that uncertainty in recognising the presence or absence of abstraction effects needs to be discussed: Reviewer 2 has raised similar points (see R2C3 and R2C4). We now provide an example (in the new Supplementary Information, S2) that illustrates the method and more discussion of uncertainties at Lines L251:262 in the results section. More detailed justification for these changes is given below.

> [*L164:180]: Under near-natural conditions, the optimum correlation between standardised precipitation and groundwater indices (SPI_Q-SGI) is generally high in unconfined aquifers (Bloomfield and Marchant, 2013). Anomalies in precipitation propagate with a relatively constant delay in recharge to the groundwater, which is due to, subsurface controls on recharge, the antecedent condition of the land surface, and non-linear response of groundwater systems (Eltahir and Yeh, 1999; Peters et al., 2006; Tallaksen et al., 2009). This constant delay is included by the optimal precipitation accumulation period in the calculated SPI_Q-SGI correlation represents a long-term relationship for a certain site, as both the SPI and*

SGI were calculated for a continuous 30-year period including all seasons and both anomalously dry and wet periods.

The SPI_Q-SGI correlation can be reduced when groundwater level response becomes disconnected from driving precipitation under confined conditions (Bloomfield et al., 2015; Kumar et al., 2016; Lee et al., 2018) or when groundwater abstraction changes groundwater storage and levels independent from driving precipitation (Bloomfield et al., 2015; Lorenzo-Lacruz et al., 2017; Haas and Birk, 2017). In this study, the impact of confined conditions on reducing the SPI_Q-SGI correlations is expected to be minimal, as only a small selection of Chalk sites are located in the semi-confined Chalk in South Lincolnshire (Table 1). On the other hand, the impact of dynamic groundwater use on SPI_Q-SGI correlations is expected to be significant, as long-term changes in groundwater use in the water management units resulted in a spatially heterogeneous pattern of irregular, decreasing, or increasing influence of abstraction on groundwater storage. For example, Ohdedar (2017) shows that groundwater use in the UK increased until the late 1980s and reduced afterwards with a large redistribution of where water is taken from to minimise the impacts on low flows.

There are three main reasons why we believe that our approach is appropriate, as follows: 1) the definition and nature of SGI and SPI, and high SPI_Q-SGI correlations based on long-term average relationships for all groundwater levels under near-natural conditions, 2) the irregular and dynamic nature of groundwater abstraction in the water management units, and 3) consistency with the results of previous studies.

First, we would like to emphasize that correlations between standardized precipitation and groundwater time series are generally high in unconfined systems for near-natural conditions (Bloomfield and Marchant, 2013). Long-term standardised groundwater and precipitation indices (SGI and SPI respectively) were calculated for a continuous period including all seasons and both anomalously dry and wet periods. The calculated SPI_Q-SGI correlations represented thus a long-term average relationship between precipitation and groundwater response, not just the relationship during episodes of drought. Consequently, the suggestion of Reviewer 1 that anthropogenic influences during droughts might increase SPI_Q-SGI correlations at longer accumulation periods would only occur if abstraction effects were sustained for the majority of the analysis period, not just during droughts since the correlation is based on a 30-year record. We found no evidence that this has occurred in the four investigated water management units, in fact for all units if anything a decrease in overall groundwater use was found (see Table 1 in manuscript).

Secondly, there are two main reasons why the long term average SPI_Q-SGI correlations may be reduced. The first reason is when groundwater level response becomes disconnected from driving precipitation under confined conditions (Bloomfield et al., 2015; Kumar, et al. 2016; Lee et al., 2018). For our sites, this is not considered to be a significant issue, as only a few sites are located in sections that become increasingly confined (Southern Lincolnshire; see Table 1). The second reason for reduced SPI_Q-SGI correlation is the effect of groundwater abstraction. In this study, groundwater abstraction is conceptualised as exerting change in groundwater storage, and therefore groundwater levels, independent of natural changes in groundwater storage associated with driving precipitation. These changes are considered highly dynamic in both space and time, as multiple abstraction wells in a water management unit (i.e. well field) are typically used to meet water demand. We don't have quantitative information about either detailed operational practices, but there is sufficient evidence that abstraction and management practises have changed during the period of investigation.

The amount of abstracted groundwater depends on the dynamic water demand and management policies in place. For example, water demand is often seasonal with higher abstraction in spring and summer. This seasonal change in water use was previously found to reduce long-term correlations between precipitation and groundwater (Lorenzo-Lacruz et al., 2017). In addition to seasonal variation, long-term changes in groundwater abstraction show an increase up until the late 1980s nationally, since when legislation has resulted in a general reduction in abstraction, but with a redistribution of where water is taken from to minimise the impacts of surface flows (Whitehead and Lawrence, 2006; Environment Agency, 2010; Shepley et al., 2008; Shepley and Streetly, 2007; Ohdedar, 2017). Both short-term (seasonal) and long-term changes in abstractions are likely to result in a spatially heterogeneous pattern of irregular, decreasing, or increasing influence of abstraction on groundwater storage.

Thirdly, our hypothesis that this highly dynamic pattern of groundwater abstraction will result in reduced SPI_Q-SGI correlations matches previous research, for example by Bloomfield et al., 2015, who found lower SPI_Q-SGI correlations for wells that are influenced by groundwater abstractions (clusters 3 and 6). Another example of disturbance of this relationship is given by Haas and Birk (2017), who showed that correlations between precipitation, streamflow, and groundwater observations are reduced due to the interference of power plants.

Lastly, the complexity and irregularity of management practices across the study sites combined with the lack of quantitative information on abstractions have also meant that we could not do an analysis of temporal variations in groundwater response to abstraction.

Consequently, based on these considerations we feel that our working hypothesis that the varying influence of abstraction will contribute to a reduction in the long-term SPI_Q-SGI correlation is reasonable. We have adjusted L164:180 in the new version of the manuscript to also clarify this for the reader.

R1C2: Section 3.2.4 Lines 185-196. The authors use the statistic variable 'Z' of the Mann Kendall test to judge whether the groundwater of the monitoring sites involves human influences. I think the statistic variable 'Z' can indicate the significance level (e.g., when |Z|>2.56, it suggests a significant trend), however, it seems arbitrary to conclude that the detected trend becomes more significant with increased value of |Z|. Fortunately the authors mentioned that both PET and precipitation present no significant trend, while groundwater level presents significant trend. This inconsistent pattern between PET /precipitation and groundwater level may imply the existence of human influence. I suggest the authors use additional methods (e.g., linear regression) to confirm the existence of human influence.

We agree with Reviewer 1 that the description of the trend Z indicator could be improved. We have addressed this point by extending our description of the trend test methodology at Section 3.2.4 [L201:206]. We have also added additional detail to the results section 4.3 [L301:308] and modified Figure 3 to include significant and non-significant trend values. We have also changed the significance level as suggested by Reviewer 1 that resulted in slight rephrasing of the results' section (see below).

> *[L201:206] Trends were quantified by the trend Z value showing positive or negative deviations from the null hypothesis (no trend). Positive/negative Z values indicated increasing/decreasing trend directions. |Z| values over |2:56| (α = 0.01) were considered significant. Trends in groundwater level time series were tested using a modified Mann-Kendall trend test (Mann, 1945; Kendall, 1948), which includes a modification developed by Yue and Wang (2004) to account for significant auto-correlation in the annual groundwater*

data (Hamed, 2008). Trends in climate time series were also calculated from annual data using a standard Mann-Kendall trend test.

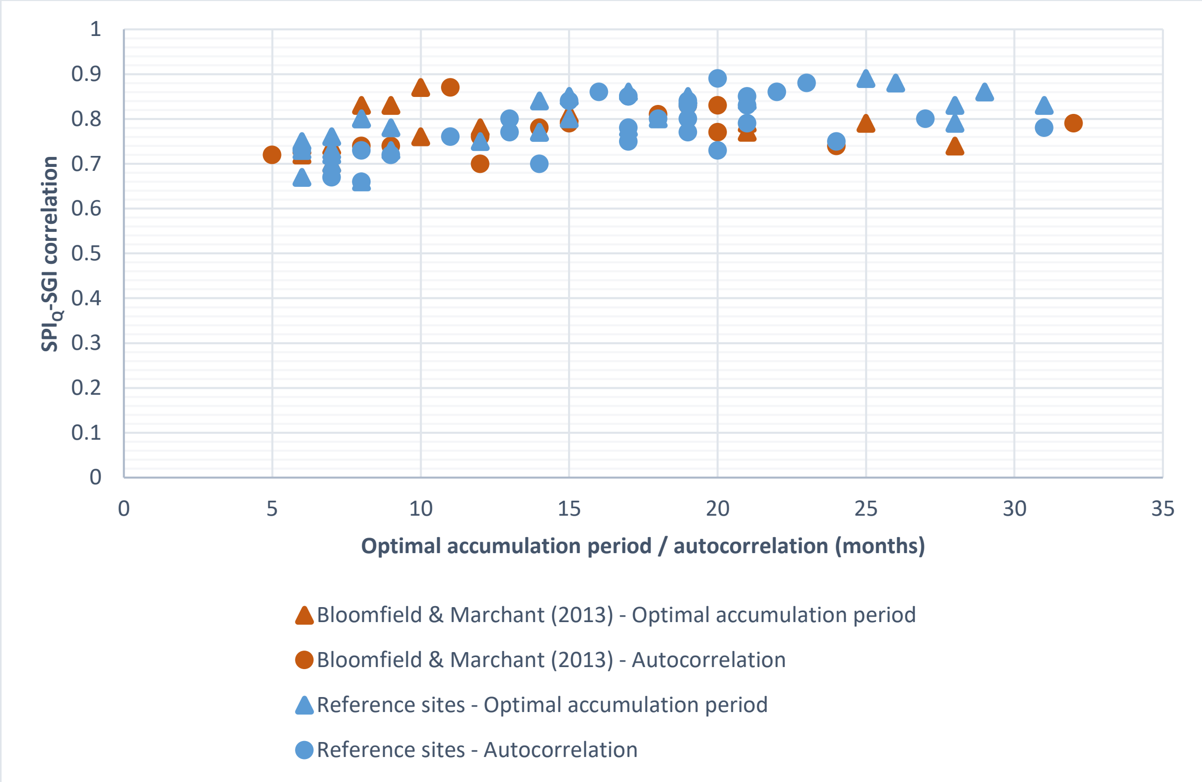
[L301:308] Significant trends in groundwater level were detected in 38% of all monitoring sites in the water management units. Of these 38%, half of the trends are upward (positive) and the other half is downward (negative) trends (Figure 3). Overall, upward trends are dominating (61% of sites including significant and non-significant trends), indicating a sustained rise in the 30-year groundwater level time series. Fewer (39% including significant and non-significant) downward trends are detected indicating sustained lowering of groundwater levels. The presence of these significant trends in groundwater is notable given the weak, non-significant, trends in the 30-year precipitation and potential evapotranspiration data (P: Z =−0.75-1.53, PET: Z=0-0.65).

We mention in the manuscript that the significant auto- and serial correlation in the groundwater time series [L204:205] limits the application of parametric methods, such as linear regression, which is only applicable to normally distributed independent data. We tested the annual groundwater level time series and only 5 out of 170 time series are normally distributed (Shapiro-Wilk Normality Test). All others (165 time series) deviate from a normal distribution, which was also found by Bloomfield and Marchant (2013) for their groundwater time series. Therefore it seems unsuitable to apply linear regression to the majority of the groundwater dataset.

R1C3. The time series of SGI for reference wells in Figure 1 (Section 4.1) show significant spatial heterogeneity, and their time scales vary from one site to another. For example, C2 #5 presents long time scales, while C4#9 presents short time scales. This may lead to the higher correlation between SPIQ-SGI for C2 #5 than C4#9 (see comment 1). I think the way of using correlation to judge the human influence is worth thinking

We acknowledge the noted spatial heterogeneity by Reviewer 1. This is consistent with previously documented spatial variations in the characteristics of (autocorrelation structure and record of hydro-climatic extremes) of groundwater level time series in the Chalk by Marchant and Bloomfield (2018). However, there is no systematic evidence for higher SPI_Q-SGI correlations between sites with longer (C2 #5) or shorter (C4 #9) SPI accumulation periods (or autocorrelations in SGI), in fact quite the contrary, SPI_Q-SGI correlation has been shown to be broadly independent of SPI accumulation period. We illustrate this in the Figure below which shows SPI_Q-SGI correlation co-efficient as a function of SPI accumulation period (months) and SGI autocorrelation for data from this study (blue triangles and dots respectively) and for data from Bloomfield and Marchant (Table 2, 2013) (orange triangles and dots respectively) from Bloomfield and Marchant (Table 2, 2013). Consequently, we have rephrased the current explanation [L222:226] to highlight the consistency of SPI_Q-SGI correlations for these near-natural sites. Below this inserted text, we provide some additional evidence for this revision to the text.

[L222:226] The optimal SPI_Q-SGI correlations of the near-natural wells are high on average (0.79) with a range of 0.66 to 0.89. These correlations are found using the optimal accumulation period, which accounts for delay in recharge that is different for each reference cluster. High SPI_Q-SGI correlations are found for both short and long accumulation periods and there was no systematic relationship between the SPI_Q-SGI correlation and the SPI accumulation period Q or SGI autocorrelation in the near-natural wells.



R1C4: Lines 254-264 and 315-318: The authors mentioned that 'The first pattern, apparent in Lincolnshire, Chilterns, and Shropshire, shows an increase in short drought events often found before a major drought event or during hot summers, which is probably related to an increase in water use'. However, from Fig.2 it shows that 'minor droughts C2 before major drought events' are not limited to influenced sites, similar phenomena are also observed in uninfluenced sites. Other factors such as the drought identification method, and the spatial heterogeneity of groundwater may also generate such minor droughts. It seems arbitrary to attribute such events to the increased water use and there is much uncertainty on the results.

We agree with Reviewer 1 that the paper would benefit from an additional analysis and justification of statements related to the interpretation of the occurrence of minor droughts. Consequently, we have provided more text in the results section at Lines 268:283, including a new data analysis and figures in the Supplementary material. In Lines 284:295 we provide additional contextual information from other studies. The discussion section has, consequently, also been rephrased [344:350].

> *[268:270] Categorised influenced sites (those with SPI$_Q$-SGI correlations lower than the cluster minimum) had typically shorter drought events of a lower magnitude. The distribution of drought duration in Figure S6 shows that the majority of these additional droughts is recorded in influenced sites compared to uninfluenced sites in Lincolnshire, Chilterns, and Shropshire.*

> *[279:283] However, there was no consistency between the study areas in relation to the timing of these shorter drought events. In Lincolnshire, minor droughts occur more often during reference droughts. In the Chilterns and Shropshire, more droughts are detected prior to reference droughts (Table S8). All minor droughts are shorter than the groundwater memory (auto-correlation) suggesting that these minor droughts are less likely to be related*

to propagated precipitation deficits, but instead are probably related to groundwater abstraction.

[284:295] Drought descriptions in the literature show an increase in water demand during the 1995-97, 2003-06 and 2010-12 drought (Walker and Smithers, 1998; Marsh et al., 2013; Durant, 2015). For example, Durant (2015) found that during the 1988-93 drought event evapotranspiration was exceptionally high. Impacts were mostly felt in the Chalk, particularly in regions where groundwater is the principal source of water supply where abstractions amplified the drought effects. An extreme rise in water use was reported by Walker and Smithers (1998) during the 1995-1997 drought event putting strain on drinking water supply systems in North East England. Sections of the Permo-Triassic sandstone were amongst the worst affected with drought conditions until 1998 (Durant, 2015). During the 2003-06 and 2010-12 droughts, a sudden increase in groundwater use was found that was attributed to dry weather and hot summers in the work of Marsh et al. (2007, 2013) and Durant (2015). In the work of Rey et al. (2017), low SPI_3 values were found in summer months for 1995, 1996, 2003-2006, and 2010-2011 highlighting exceptional dry weather that led to surface water use restrictions prior to droughts to maintain low flows. Consequently, the reduced surface water abstractions were replaced by groundwater, for which use was rarely restricted (Rey et al., 2017) resulting in lowered groundwater levels and potentially aggravating groundwater droughts.

[344:350] The first pattern, apparent in Lincolnshire, Chilterns, and Shropshire, shows an increase in short drought events in influenced sites that sometimes occur before a major drought event or during unusual dry period that results in a rapid increase in both surface water and groundwater use (Walker and Smithers, 1998; Marsh et al., 2013; Durant, 2015) and/or complementary groundwater use due to surface water use restrictions (Rey et al., 2017; Rio et al., 2018). We see the effect of this local increase in water use in our data in the temporarily lowered groundwater levels resulting in additional drought events. The majority of these events occur in influenced sites, but some of the (on average) uninfluenced sites also show minor droughts. Given the high correlation in these uninfluenced sites, the minor droughts seem to not disturb the long-term average correlation.

In the new analysis described below, we show the distribution of drought duration for influenced and uninfluenced sites (see Figure 1 included in Supplementary material S6.). Table 2 in the manuscript shows that mean drought duration is lower in influenced sites and Figure 1 shows the distribution of these shorter droughts (3-5 months) for influenced sites. The spike in the distribution plot confirms the increased occurrence of minor droughts in influenced sites in Lincolnshire, Chilterns and Shropshire. These minor droughts vary slightly in duration, as expected, given the different hydrogeological settings of the water management units.

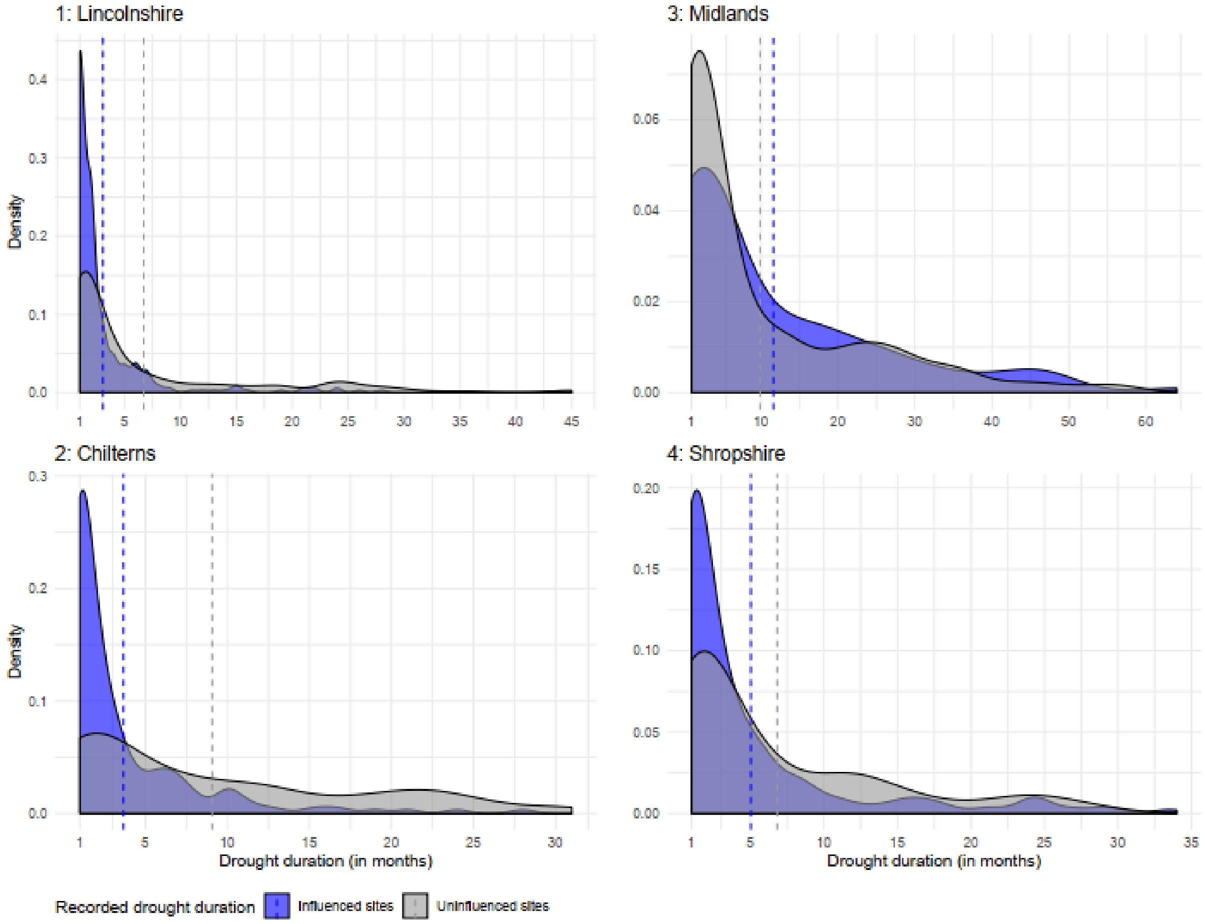


Figure 1: Distribution of drought duration for classified influenced (blue) and uninfluenced (grey) sites in the four water management units. The average drought duration is highlighted by the striped vertical line in the graph (colours are matching).

The additional data analysis also shows that not the majority, but 27 to 43 percent of the shorter droughts occur 1-24 months before the reference droughts (fourth column Table 2). In Lincolnshire, 60 percent of the shorter droughts occurs during the reference droughts compared to a smaller percentage in the Chilterns (27%) and Shropshire (23%; in fifth column Table 2). We have changed the wording in lines 278:281 to highlight the relative occurrence of the shorter drought events.

Table 1: Duration and occurrence of minor droughts in influenced sites in Lincolnshire, Chilterns, and Shropshire. Results show that the average during is shorter than the average groundwater memory (auto-correlation).

Water management units	**Average duration of minor droughts (in months)**	**Average autocorrelation (in months)**	**Occurrence of minor droughts 24 months before reference droughts (%)**	**Occurrence of minor droughts during reference droughts (%)**
1: Lincolnshire	3.1	11.6	27	60
2: Chilterns	3.7	17.3	34	27
4: Shropshire	5.0	15.1	43	23

The minor droughts recorded in the influenced sites are also shorter than the groundwater memory (auto-correlation in third column Table 2) for all water management units suggesting that these droughts are not related to a natural deficit in groundwater due to a propagated precipitation deficit, but to abstraction influence. Contextual information shows that increased water use has amplified existing droughts and that increased groundwater use was related to periods of exceptional dry weather. We have rephrased the text in order to clarify the sources of this contextual information [L279:295].

R1C5: Lines 322-323, The authors mentioned 'We see the effect of this local increase in water use in our data in the temporarily lowered groundwater levels, resulting in additional drought events'. Could you provide additional information on the evolution process of water use and droughts, e.g., show the time series of both water use data and groundwater levels in one figure.

We agree with Reviewer 1 that it would be interesting to analyse both water use data and groundwater level variations in one figure. Unfortunately, this is not possible given the unavailability of abstraction records [L59:60]. The unavailability of detailed, time-varying records oft abstraction is the primary reason for developing the methods here to infer abstraction influence [L62:63]. We have, however, provided an example of four groundwater hydrographs in the Chalk for which the first is categorised as near-natural, and the other three are groundwater monitoring sites (see S2 in R2C3). Out of these three monitoring sites, the first site shows a high correlation with the accumulated SPI, hence classified as uninfluenced. The other two sites have a low correlation with accumulated SPI and have an irregular, spikey hydrograph that also shows the temporarily lowered SGI values despite normal or above-normal SPI. These two wells are assumed to be continuously influenced by abstraction resulting in a lower SPI_Q-SGI correlation.

R1C6: Lines 115-118, The authors failed to illustrate how they calculate SGI clearly. For example, which probability distribution was employed to fit the groundwater series. Whether the impact of data seasonality was considered when calculating SGI? More details on the computation of SGI should be added.

We have changed the phrasing in lines 133-136 explaining the assigned SGI value and calculation of SGI. Note that SGI is calculated here using the non-parametric method of Bloomfield and Marchant (2013) so no assumptions about distributions were made.

> *[L133:136] Monthly groundwater observations were grouped for each calendar month and within each month observations were ranked and assigned a SGI value based on an inverse normal cumulative distribution of the data. No distribution was fitted, but SGI values were assigned to monthly observations accounting for seasonal variation within the calendar year.*

R1C7: Lines 120-121: '208 sites have been included in the analysis, 39 are reference sites and 170 monitoring sites. ' Here '208 sites' should be '209 sites (170+39=209)'.

We thank Reviewer 1 for spotting this mistake. It should be 209. In total, there are 39 near-natural reference wells areas (9 in the PT sandstone and 30 in the Chalk). There are 170 groundwater monitoring sites divided over the four water management units (see Table 1 first column).

> *[L118:119] 209 sites have been included in the analysis, 39 are reference sites and 170 monitoring sites.*

R1C8: Lines 130-131: How do you fill the missing sequences, using the time series of adjacent sites? Details on the linear interpolation method should be supplemented. Besides, sites with missing data more than 6 months would be removed directly?

We interpolated the missing data from the last measured groundwater observation to the next linearly if that missing sequence was not longer than 6 months, as previously applied in the work of

Tallaksen and Van Lanen, (2004) and Thomas et al., (2016). Groundwater sites with missing sequences longer than 6 months were indeed removed from the dataset prior to the analysis. The text has been revised as such at L127-129.

> *[L127-129] Missing data were linearly interpolated from the last observation to the next observation in case of short sequences of missing data (less than 6 months) (Tallaksen and Van Lanen, 2004; Thomas et al., 2016).*

R1C9: The current form of Fig. 2 makes it difficult to judge the impact of human influences. The authors could add the time series of SGI for the monitoring sites so that readers can easily find human influenced periods

We thank Reviewer 1 for this final comment, but in light of the newly added examples in S2 showing four SGI time series to illustrate the method (R2C3), we don't think is necessary to add more SGI time series to Figure 2. This is because, the design of Figure 2 is so that the timing and magnitude of groundwater droughts can be overviewed at glance. We wanted to highlight that groundwater droughts observed in sites with a reduced SPI_Q-SGI correlation differ in timing and magnitude. To transform this graph into time series would require a figure capturing 170 time series that would, in our opinion, result into more confusion than the highlighted droughts occurring in these 170 time series.

In addition to this, we would like to highlight that there are no specific 'human-influenced periods' identified in the investigation period. We have contextual information about the overall water use that changes in time showing that the aquifers have continuously regulated from the 1960s until now (Ohdedar, 2017).

Reviewer: 2

This paper investigates the impact of groundwater use on groundwater drought for a case study in the UK.

Overall, I found the paper to be well-written, with some interesting results supported by some nice figures. The work represents a useful contribution to better understand how groundwater use affects groundwater drought and how groundwater levels have changed over time in the UK.

My main suggestions for the paper are to improve the clarity of the methods and reflection of the variability in groundwater levels. I agree with reviewer 1 that there is a lot of uncertainty in the results and some of the links between groundwater use and groundwater drought are somewhat arbitrary. I sympathise with the authors as I know how difficult it is to gain groundwater use data that may help make these findings more robust, nevertheless, I believe the authors could make more efforts to discuss the limitations of their results and report the uncertainty/variability in their results

We thank Reviewer 2 for their careful reading and constructive comments. We have pleased to hear that the contribution of the study is useful and have addressed uncertainty and limits of the current methods. Please find our point-by-point responses to comments below.

R2C1: In the methods/discussion please add some comment on the choice of gamma distribution used to calculate SPI. Other studies have shown that this is often not the

most appropriate distribution for precipitation data and it would be good to discuss the impacts of this (see Svensson et al. 2017 for example).

We thank Reviewer 2 for their comment. We now tested the alternative distributions for the SPI calculation and have included additional phrases to comment on the choice of gamma distribution used to calculate the SPI. The uncertainty has now been addressed more explicitly by improving the phrasing in L110:116.

However, we would like to emphasise that the SPI is primarily used in combination with the SGI to find the optimal correlation between the SPI and SGI. For this correlation, primarily long accumulation periods (> 12months) of the SPI were used (see the mean of optimum accumulation periods in L226:230 for near-natural wells and the range of accumulation periods in S3 in the supplementary material). Considering the use of long accumulation periods, the `best' fitting distribution varies (Svensson et al. 2017). High rejection rates are found for multiple distributions (Stagge et al. 2015), which suggests we need to test which distribution performs best. We have tested different distributions for a subset of the data, shown below the inserted text.

> *[L110-116] Precipitation estimates were converted into standardised precipitation indices (SPI) following the method of McKee et al. (1993). A gamma distribution was fitted to precipitation estimates and alternative distributions were also tested (Normal, Pearson III, and Logistic). Considering the use of SPI to account for delayed recharge, a large range of accumulation periods of precipitation (1 to 100 months) was calculated in order to find the optimal correlations between precipitation and groundwater time series. For this particular use of the SPI, the 'best' fitting distribution varies (Svensson et al., 2017). Alternative distributions showed minimal differences in the computed correlations between standardised precipitation and groundwater time series, hence we decided therefore to use the gamma distribution.*

The additional test is performed on a subset of the total dataset (45 precipitation grids matching to groundwater monitoring sites in the Chilterns). Three alternative distributions were tested: Normal, Pearson III, and Logistic distribution and results are presented in a similar way to Figure 5d in Svensson et al., (2017) (Figure 2). Figure 2 shows an example of a SPI_{15} in which a slight variation is seen in the calculated SPI values during droughts using different distributions. This variation did, however, not result in higher or lower SPI_{Q}-SGI correlation. For the subset of the data (45 monitoring sites), the range of correlations using the Gamma distribution was 0.41-0.89 with a mean of 0.794. The mean of the calculated correlation remained the same when using different distributions (Normal, Pearson III, and Logistic distribution). The range of the 45 correlations showed minimal changes compared to Gamma distribution (0.41-0.89): 0.40-0.89 (Normal & Logistic), 0.40-0.90 (Pearson III). Reviewing the minimal change in SPI_{Q}-SGI correlation, we think that the use of alternative distributions instead of the current distribution (gamma) would not change the results of this study given the use of the SPI only in the correlation analysis.

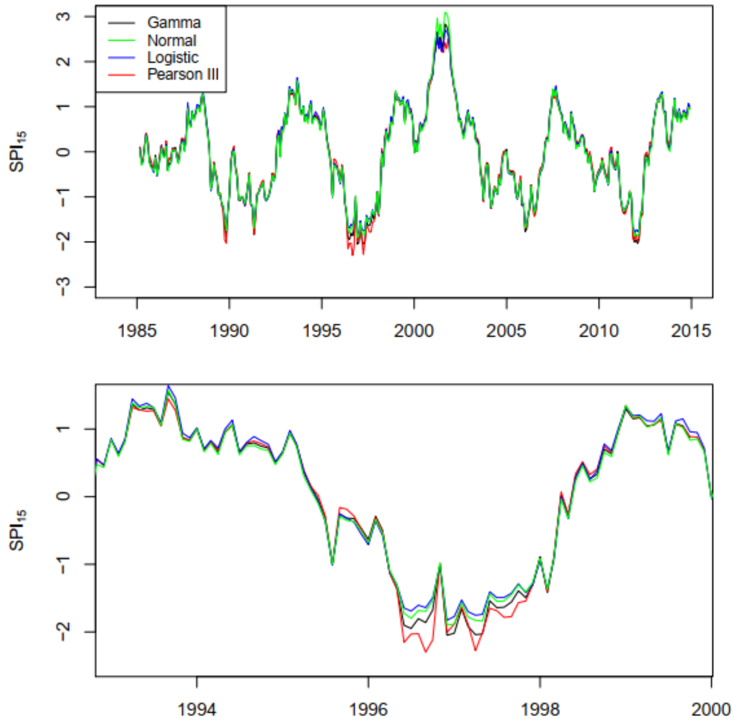


Figure 2: SPI_{15} computed using different distributions for a precipitation estimate located in the Chilterns (the location corresponding to groundwater site (SP90.27).

R2C2: **From the methods section, it seems that you compare the SPI from a single grid cell with the corresponding groundwater well location (this should be clarified in the text). It would be good to add to the discussion the impacts of comparing a 1km2 grid cell of SPI with SGI that is a product of a regional groundwater aquifer system and regional rainfall patterns.**

We agree with Reviewer 2 and rephrased the text to clarify our approach (L104:109) and provide additional context below.

> *[L104:109] We aggregated daily potential evapotranspiration estimates to monthly sums. For both gridded datasets (GEAR and CHESS) grid cells were extracted corresponding to groundwater well locations. The 1km^2 gridded precipitation and potential evapotranspiration sums were compared to the monthly groundwater observations of the same location. This point-scale comparison assumes that the influence of precipitation is largest surrounding the groundwater monitoring site (Bloomfield and Marchant, 2013; Bloomfield et al., 2015; Li and Rodell, 2015; Kumar et al., 2016).*

The regional extent of groundwater recharge varies and the precise extent of this recharge area associated with a given observation borehole is often unknown. In contrast to surface water boundaries, there is no consistent source of information regarding the recharge area for groundwater monitoring sites in the UK in either the Hydrometric Register (Marsh and Hannaford, 2008) or the water management units. The unknown recharge area is a common uncertainty for groundwater studies and other studies have either used a regional aggregate to overcome this unknown recharge area (Haas et al., 2018) or used a point-scale analysis under the assumption that the influence of precipitation is largest surrounding the groundwater monitoring site (Bloomfield and Marchant, 2013; Bloomfield et al., 2015; Li and Rodell, 2015; Kumar et al., 2016). Even though a regional precipitation product would potentially result in a more accurate correlation, it could also introduce larger uncertainties given the unknown extent. In addition to this, high correlations between SPI_Q-SGI have been previously obtained using the point-based precipitation estimates in different climate regions and by different authors (Bloomfield and Marchant, 2013; Bloomfield et al., 2015; Li and Rodell, 2015; Kumar et al., 2016). Also in our study, high correlations are found for near-natural wells and the majority of groundwater monitoring wells [L222:226 and L231:235]. Therefore, considering the unknown recharge area and reasonable results, we don't propose to modify our methodology, which is consistent with previous studies using point-based precipitation estimates.

R2C3: The methods (in places) were not clear – in particular, the SPI_SGI correlations and the use of the near-natural wells, uninfluenced and influenced monitoring sites. It would be useful to have a worked example of how the SPI-SGI correlations work in practice (showing an example for two sites – one influenced and one non-influenced and how they compare to the near natural reference cluster). It would also be useful to have a map of the influenced and non-influenced wells (this is maybe already included in Figure 1 but this figure is quite busy so it is hard to tell!)

In response to R2C3, we have added an illustration of the SPI_Q-SGI correlation methodology to the Supplementary information (S2) using four wells to show the SPI_Q-SGI correlations in a single water management unit for a near-natural reference site and three groundwater monitoring sites (influenced and uninfluenced) and referred to this illustrated example in the section 3.2.2.

> *[L185:186] An illustrated example is provided in Figure S2 showing SGI time series of a near-natural reference site and three groundwater monitoring sites.*

We agree with Reviewer 2 that it would be very interesting to show spatial patterns of detected influenced wells. We had considered mapping locations of influenced wells, but analysing and explaining spatial patterns in such maps would require detailed knowledge of the hydrogeological setting of each monitoring well and records of abstraction wells close by – information that we don't have. There is no consistent spatial pattern based on annual maximum abstraction licences. We expect that some wells are used episodically or not at all, while others are used regularly resulting in a highly variable picture. Given the unexplained spatial patterns, complex local hydrogeological structure, and the unknown use of the abstraction wells, we have not included these manuscript, as more information is required to explain the spatial patterns.

R2C4: Reflection of uncertainty/variation
There is a lot of variation in the groundwater levels between sites and this needs to be better reflected in the results. I suggest that the authors report the min/max or 5th/95th percentile of their results alongside the average in Table 2 and elsewhere in the text.

As suggested by Reviewer 2, we have amended Table 2 to include the 5th and 95th percentiles of the duration, magnitude and frequency of groundwater droughts at the uninfluenced and influenced sites for each water resource management unit.

We have also included primary reasons for the variation in the groundwater levels [L241-245], as the groundwater level observations are set in a range of different hydrogeological settings and drought events vary in timing, intensity and duration, as groundwater droughts are episodic. On top of the spatial and temporal differences, human-influence on groundwater level variations change in time, which results in the variation in Table 2. We improved the phrasing in the results section describing these three different facets in the drought characteristics [L251-253].

> *[L241-245] Groundwater droughts observed in the reference clusters reflect both spatial and temporal variation due to driving precipitation and hydrogeological setting. In general, the*

four UK-wide droughts (1988-1993, 1995-1998, 2003-2006, and 2010-2012) are reflected in near-natural groundwater time series. Spatial patterns in driving precipitation, however, result in variable groundwater drought occurrence (Figure 1).

[L251-253] On a smaller scale in the water management units, average drought characteristics (duration in months, magnitude in accumulated SGI over the drought period, and frequency) for monitoring sites show differences due to abstraction influence, which we have classified in, on average, uninfluenced and influenced sites, see Table 2.

In addition, to better illustrate the variability in the drought characteristics between uninfluenced and influenced sites we have introduced a new set of distribution figures in the Supplementary Information (S5-S7) and provided additional explanatory text at Lines 255:262.These distribution figures show the difference and overlap between influenced and uninfluenced sites.

[L255-262] Droughts are observed twice as often in influenced compared to uninfluenced sites in Lincolnshire and Chilterns, but this difference is smaller in Shropshire. The distribution of recorded drought frequency (Figure S5) shows that the difference between on average influenced and uninfluenced sites is less pronounced in Lincolnshire and Shropshire. Table 2 shows that the average drought duration of influenced sites exceeds the duration in uninfluenced sites in the Midlands. Longer and more intense groundwater droughts occurred less often in influenced sites, which is in contrast with the other water management units. The distribution of recorded drought frequency (Figure S5) shows a majority of sites recording fewer droughts and some sites that record a higher frequency. On average, this results in a small difference between the influenced and uninfluenced sites.

R2C5: Like Reviewer 1, I am somewhat sceptical of attributing the shorter droughts in Lincolnshire, Shropshire and the Chilterns to water use and/or hotter Summers. Firstly the years that were identified in L263-265 did not have particularly hot summers (or this is certainly not consistent for these years) and many of these drought events can also be identified in the uninfluenced wells. These uncertainties need to be reflected in the discussion or the methods for identification need to be more robust.

We understand the concern raised regarding the attribution of shorter droughts to increased water use by Reviewer 2 and earlier by Reviewer 1 (R1C4). We have rephrased the result and discussion section [267:295, and 344:350, see R1C4] and have included the new data analysis in the supplementary material (S6 and Figure 1 in this rebuttal). Reviewer 2 is right to note that shorter (or minor) droughts are also observed in uninfluenced sites. However, the distribution graphs of recorded drought frequency and duration in groundwater monitoring sites show that the majority is in uninfluenced sites. We have also provided additional contextual information regarding the reported increased water use [284:295, see R1C4].

Minor Comments

R2MC1. The abstract is quite long – I would shorten it and just highlight the key results. Currently, your more interesting results get a little lost in all the text.

We agree with Reviewer 2 and we have shortened the Abstract into the following:

[L1:19] Groundwater use affects groundwater storage continuously, as the removal of water changes both short-term and long-term groundwater level variation. This has implications for groundwater droughts, i.e. a below-normal groundwater level. The impact of groundwater use on

groundwater droughts, however, remains unknown. Hence, the aim of this study is to investigate the impact of groundwater use on groundwater droughts in the absence of actual abstraction data adopting a methodological framework that consists of two approaches. The first approach compared groundwater droughts at monitoring sites that are potentially influenced by abstraction to groundwater droughts at sites that are known to be near-natural. Observed groundwater droughts were compared in terms of drought occurrence, magnitude, and duration. The second approach investigated long-term trends in groundwater levels in all monitoring wells. This framework was applied to a case study of the UK using four regional water management units, in which groundwater is monitored and abstractions are licensed. Results show two, asymmetric, responses in groundwater drought characteristics due to groundwater use. The first response is an increase of shorter drought events, and is found in three water management units where long-term annual average groundwater abstractions are smaller than recharge. The second response, seen in one water management unit where groundwater abstractions temporarily exceeded recharge, is a lengthening and intensification of groundwater droughts. Analysis of long-term (1984-2014) trends in groundwater levels shows mixed, but generally positive trends, while trends in precipitation and potential evapotranspiration are not significant. The generally rising groundwater levels are consistent with changes in water use regulations and with an overall reduction in abstractions during the period of investigation. We summarised our results in a conceptual typology that illustrates the asymmetric impact of groundwater use on groundwater drought occurrence, duration, and magnitude. The long-term balance between groundwater abstraction and recharge plays an important role in this asymmetric impact, which highlights the relation between long-term and short-term sustainable groundwater use.

R2MC2. P3 L80 It would be good to name these four water management units in the text

Agreed. We have included the names of the water management units in now L74:75.

> *[L74:75] The UK case study consists of four water management units (1: Lincolnshire, 2: Chilterns, 3: Midlands, 4: Shropshire) across the Chalk and Permo-Triassic sandstone aquifers that are the two main aquifers in the UK (Figure 1).*

R2MC3. P5 L117 – What accumulation periods did you calculate SPI over, you need to be more specific here.

Agreed. We have included the full range of accumulation periods in the new version of the manuscript.

> *[L112:114] Considering the use of SPI to account for delayed recharge, a large range of accumulation periods (1 to 100 months) in order to find the optimal correlations between precipitation and groundwater time series.*

R2MC4. Table 1 – what time period were the long term precipitation and PET calculated over? It would be good if this was consistent with the time periods used in your study.

We thank Reviewer 2 for pointing this out, as the long-term precipitation and PET was taken from the Mansour and Hughes (2018) study and based on daily data from 1962 to 2016. We have now clarified that in the table caption.

R2MC5. P7 L194 – Were these the climate time series from a single grid cell?

This data is indeed from the same climate datasets using the same extracted grid cells. We have now clarified this by rephrasing L197:200.

> *[L197:200] Hence, an additional trend test was introduced to compare trends in annual (averaged for each calendar year) groundwater levels to climate data (precipitation and evapotranspiration) that were extracted for grid cells corresponding to groundwater well locations from the GEAR and CHESS datasets (Tanguy et al., 2016; Robinson et al., 2016)*

R2MC6. Table 2 - I was a little surprised that the average drought frequency for the Midlands cluster is significant when the values are quite similar (9.5 for uninfluenced and 9 for influenced) – is this correct?

We agree with Reviewer 2 that this is an interesting result and checked this before submission. The averaged difference in drought frequency is indeed statistically significant, which is now clearer when looking at the distribution the spread of the data in S5.

R2MC7. Section 4.3 – in this section you don't distinguish between 'influenced' and 'uninfluenced' wells. It would be useful know whether the strong trends are just in the 'influenced' wells? If they are not, then your 'uninfluenced' wells may be more affected than suggested.

We thank reviewer 2 for noting this difference between sections and we acknowledge that this topic did not receive much attention in the manuscript. We have explicitly stated this now in the abstract, methods and results [L7:8, 194, 301].

We have indeed not distinguished between influenced and uninfluenced sites and this is because the methods to categorise influence of abstraction are not designed for the identification of trends or long-term changes in groundwater levels. We have added this to the method section now [L196:198].

> [*L196:198] This trend test contributes to the first approach, as the SGI and SPI_Q-SGI correlation analysis do not specifically account for trends in groundwater time series that could result in significant trends going unnoticed.*

The trend results correspond with the categorisation of influenced and influenced sites. Most uninfluenced sites (75%) have a non-significant trends compared to most influenced sites (72%) that have a significant trend. From the uninfluenced sites, only a small percentage (5%) has a negative trend. These sites indicate an indirect influence of abstraction nearby the groundwater monitoring and time series show both a downward trend and episodic drought events that align with an accumulated SPI signal. Interestingly, 20% of the uninfluenced sites have a significant positive trend. Investigating the drivers of these significant positive trends in groundwater levels would be interesting, although beyond the scope of the current study.

Additional references in rebuttal:

Bloomfield, J. P. and Marchant, B. P.: Analysis of groundwater drought building on the standardised precipitation index approach, Hydrology and Earth System Sciences, 17, 4769–4787, https://doi.org/10.5194/hess-17-4769-2013, 2013.

Bloomfield, J. P., Marchant, B. P., Bricker, S. H., and Morgan, R. B.: Regional analysis of groundwater droughts using hydrograph classification, Hydrology and Earth System Sciences, 19, 4327–4344, https://doi.org/10.5194/hess-19-4327-2015, 2015.

Environment Agency: Vale of St. Albans Numerical Groundwater Model Final Report, Tech. rep., 2010.

Haas, J. C. and Birk, S.: Characterizing the spatiotemporal variability of groundwater levels of alluvial aquifers in different settings using drought indices, Hydrology and Earth System Sciences, 21, 2421–2448, https://doi.org/10.5194/hess-21-2421-2017, 2017.

Kumar, R., Musuuza, J. L., Van Loon, A. F., Teuling, A. J., Barthel, R., Ten Broek, J., Mai, J., Samaniego, L., and Attinger, S.: Multiscale evaluation of the Standardized Precipitation Index as a groundwater drought indicator, Hydrology and Earth System Sciences, 20, 1117–1131, https://doi.org/10.5194/hess-20-1117-2016, 2016

Lee, J. M., Park, J. H., Chung, E., Woo, N. C.: Assessment of Groundwater Drought in the Mangyeong River Basin, Korea, Sustainability, 10, 831, 2018

Li, B. and Rodell, M. Evaluation of a model-based groundwater drought indicator in the conterminous U.S. Journal of Hydrology, 526, 78-88, 2015

Mansour, M.M. and Hughes, A.G.. 2018 *Summary of results for national scale recharge modelling under conditions of predicted climate change.* Nottingham, UK, British Geological Survey, 136pp. (OR/17/026) (Unpublished)

Marsh, T. and Hannaford, J., eds.: UK hydrometric register. A catalogue of river flow gauging stations and observation wells and boreholes in the United Kingdom together with summary hydrometric and spatial statistics, Hydrological Data UK, Centre for Ecology & Hydrology, 2008.

Ohdedar, B.: Groundwater law, abstraction, and responding to climate change: assessing recent law reforms in British Columbia and England, Water International, 42, 691–708, https://doi.org/10.1080/02508060.2017.1351059, 2017

Shepley, M. and Streetly, M.: The estimation of 'natural' summer outflows from the Permo-Triassic Sandstone aquifer, UK, Quarterly Journal of Engineering Geology and Hydrogeology, 40, 213–227, 2007.

Shepley, M., Pearson, A., Smith, G., and Banton, C.: The impacts of coal mining subsidence on groundwater resources management of the East Midlands Permo-Triassic Sandstone aquifer, England, Quarterly Journal of Engineering Geology and Hydrogeology, 41, 425–438, https://doi.org/10.1144/1470-9236/07-210, 2008

Stagge, J.H., Tallaksen, L.M., Gudmundsson,L. Van Loon, A.F. and Stahl K. Candidate Distributions for Climatological Drought Indices (SPI and SPEI) INTERNATIONAL JOURNAL OF CLIMATOLOGY **35**: 4027–4040, 2015

Svensson, C., Hannaford, J., and Prosdocimi, I. Statistical distributions for monthly aggregations of precipitation and streamflow in drought indicator applications, Water Resour. Res., 53, 999– 1018, doi:10.1002/2016WR019276. 2017

Tallaksen, L. M. and Van Lanen, H. A. J.: Hydrological drought : processes and estimation methods for streamflow and groundwater, Elsevier, 2004.

Thomas, B., Landerer, F., Wiese, D., and Famiglietti, J.: A comparison of watershed storage trends over the eastern and upper Midwestern regions of the United States, 2003–2015, Water Resources Research, 52, 6335–6347, https://doi.org/10.1002/2016WR018617, 2016.

Whitehead, E. and Lawrence, A.: The Chalk aquifer system of Lincolnshire, Tech. rep., Keyworth, Nottingham, contributors: J P Bloomfield, P J McConvey, J E Cunningham, M G Sumbler, D Watling, M Hutchinson, 2006.

Asymmetric impact of groundwater use on groundwater droughts

Doris E Wendt[1], Anne F Van Loon[1], John P Bloomfield[2], and David M Hannah[1]

[1]University of Birmingham, Birmingham, UK

[2]British Geological Survey, Wallingford, UK

Correspondence: Doris Wendt (dew637@bham.ac.uk)

Abstract. Groundwater use affects groundwater storage continuously, as the removal of water changes both short-term and long-term ~~variation in groundwater level~~ groundwater level variation. This has implications for groundwater droughts, i.e. a below-normal groundwater level. The impact of groundwater use on groundwater droughts, however, remains unknown. Hence, the aim of this study is to investigate the impact of groundwater use on groundwater droughts in the absence of actual abstraction data adopting a methodological framework that consists of two approaches. The first approach ~~compares groundwater~~ compared groundwater droughts at monitoring sites that are potentially influenced by abstraction to ~~uninfluenced sites~~ groundwater droughts at sites that are known to be near-natural. Observed groundwater droughts ~~are~~ were compared in terms of drought occurrence, magnitude, and duration. The second approach ~~consists of a groundwater trend test that investigates the impact of groundwater use on~~ investigated long-term ~~groundwater level variation~~trends in groundwater levels in all monitoring wells. This framework was applied to a case study of the UK ~~. Four~~ using four regional water management units~~in the UK were used~~, in which groundwater is monitored and abstractions are licensed. ~~The potential influence of groundwater use was identified on the basis of relatively poor correlations between accumulated standardised precipitation and standardised groundwater level time series over a 30-year period from 1984 to 2014. Results of the first approach show twomain patterns~~ Results show two, asymmetric, responses in groundwater drought characteristics due to groundwater use. The first ~~pattern shows~~ response is an increase of shorter drought events, ~~mostly during heatwaves or prior to a long drought event for influenced sites compared to uninfluenced sites. This pattern~~ and is found in three water management units where ~~the~~ long-term ~~water balance is generally positive and~~ annual average groundwater abstractions are smaller than recharge. The second ~~pattern is found~~ response, seen in one water management unit where ~~temporarily groundwater abstractions exceeded recharge. In this case,~~ groundwater ~~droughts are lengthened and intensified in influenced sites. Results of the second approach show that nearly half of the groundwater time series have a significant trend, whilst trends~~abstractions temporarily exceeded recharge, is a lengthening and intensification of groundwater droughts. Analysis of long-term (1984-2014) trends in groundwater levels shows mixed, but generally positive trends, while trends in precipitation and potential evapotranspiration ~~time series are negligible. Detected significanttrends are both positive en negative, although positive trends dominate in most water management units. These positive trends, indicating~~ are not significant. The generally rising groundwater levels ~~, align~~ are consistent with changes in water use ~~regulation. This suggests that groundwater abstractions have reduced~~ regulations and with an overall reduction in abstractions during the period of investigation. ~~Further research is required to assess the impact of this change in groundwater abstractions on drought characteristics. The overall impact of groundwater use is summarised~~ We summarised our

results in a conceptual typology that illustrates the asymmetric impact of groundwater use on groundwater drought occurrence, duration, and magnitude. The long-term balance between groundwater abstraction and recharge ~~appears to be influencing~~ plays an important role in this asymmetric impact, which highlights the relation between long-term and short-term sustainable groundwater use.

1 Introduction

Groundwater is an essential source of water supply, as it provides almost half the global population with domestic water (Gun, 2012), 43% of the irrigation water (Siebert et al., 2010), and 27% of industrial water use (Döll et al., 2012), as well as sustaining ecologically important rivers and wetlands (de Graaf et al., 2019). The usage and dependency on groundwater resources has grown in the past decades (Famiglietti, 2014), particularly during meteorological droughts, when groundwater is used frequently (Taylor et al., 2013; AghaKouchak, 2015).

Meteorological droughts propagate through the hydrological cycle~~and the deficit in precipitation results in~~ , ultimately resulting in a groundwater drought (Wilhite, 2000; Van Lanen, 2006), defined as below-normal groundwater levels that are associated with short-term reductions in storage (Chang and Teoh, 1995; Tallaksen and Van Lanen, 2004; Mishra and Singh, 2010). Increased use of groundwater before or during meteorological droughts can also lower groundwater levels and ~~can~~ thereby aggravate groundwater droughts (Wada et al., 2013; Christian-Smith et al., 2015). Managing groundwater use during droughts is therefore important, as overexploitation of groundwater has disastrous consequences (Custodio, 2002; Famiglietti, 2014; Russo and Lall, 2017; Mustafa et al., 2017). However, to date groundwater droughts have been studied under primarily near-natural conditions and there is limited conceptual understanding of the impact of groundwater use on groundwater droughts despite this being of interest to water regulators and policy makers.

Under near-natural conditions, the propagation of meteorological droughts to groundwater droughts depends on the antecedent condition of the land surface, subsurface controls on recharge, and non-linear response of groundwater systems ~~(Peters et al., 2006; Tallaksen et al., 2009; Eltahir and Yeh, 1999)~~(Eltahir and Yeh, 1999; Peters et al., 2006; Tallaksen et al., 2009) . These processes determine the spatial distribution, duration, magnitude, and recovery of near-natural groundwater droughts (Van Lanen et al., 2013; Van Loon, 2015; Parry et al., 2018). However, in human-modified environments, groundwater droughts are also impacted or driven by water use (Van Loon et al., 2016b). This type of groundwater drought is therefore distinguished from a natural drought and referred to as human-modified or human-induced drought (Van Loon et al., 2016a).

In human-modified environments, understanding the influence of groundwater use on groundwater drought requires information related to ~~both~~ the natural propagation of a drought and groundwater ~~abstraction. Droughts can be~~ use in time. Droughts are influenced by historical and recent abstractions, as these change both short-term and long-term groundwater storage (Gleeson and Richter, 2017; Thomas and Famiglietti, 2015; Jackson et al., 2015). Unfortunately, information on groundwater abstraction, if available at all, is often considered commercially confidential. Abstraction records are usually unavailable for research, although these records are ~~an important component of groundwater models that are~~ included in groundwater models developed for commercial and regulatory purposes (Shepley et al., 2012). Consequently, qualitative information about groundwater use and management regulations often is invaluable to investigate the influence of groundwater abstraction on groundwater droughts (Döll et al., 2014; Panda et al., 2007). Management regulations are organised on regional or even national scale. This scale differs from the much smaller scale at which groundwater droughts are often studied. For example, physically-based groundwater models, which are developed for regulatory purposes or research, rarely cover the entire drought-impacted area or the entire aquifer that is affected by a drought event (Peters et al., 2006; Tallaksen et al., 2009; Shepley et al., 2012). Studying groundwater droughts in human-modified environments would thus require a regional approach to align the scale of a groundwater drought study with the scale at which management decisions are made.

The aim of this study is to investigate the impact of groundwater use on regional groundwater droughts in the absence of actual abstraction data. For doing so, a methodological framework is designed to investigate groundwater droughts in water management units under a broad range of conditions, i.e. from where groundwater use is a small proportion of the long-term annual average recharge to where it is (temporarily) a significant proportion of the long-term annual average recharge. A case study from the United Kingdom (UK) is used , consisting of four water management units over the two main aquifers in the UK. As is common elsewhere, no data is freely available on actual abstractions for the four water management units in the case study area. However, information indicating the annual maximum abstraction according to the groundwater abstraction licence licences is available and groundwater level observations are provided for 170 sites in the four water management units. Consequently, inferential approaches are used to assess the potential impact of abstraction on groundwater droughts. Here, we used two complementary approaches. Firstly, given the typically good correlation between precipitation and groundwater level time series under near-natural conditions (Bloomfield and Marchant, 2013; Bloomfield et al., 2015; Kumar et al., 2016), we used correlations defined by a limited number of near-natural groundwater hydrographs as reference. Deviations from this reference correlation are then used to qualitatively subdivide sites in on average uninfluenced and influenced by abstraction, and characterise the modifying effect . This subdivision is used to characterise the impact of groundwater abstraction on regional groundwater droughts. Secondly, we investigated long-term effects of abstraction abstraction influence is investigated through the spatial distribution of trends in groundwater level time series in relation to the distribution of licensed abstractions. The results Results are discussed in terms of the role groundwater abstraction plays in modifying near-natural groundwater droughts. A conceptual figure is proposed suggesting that long-term groundwater abstraction may modify drought frequency, duration, and magnitude depending on the proportion of abstraction and recharge.

2 Study area

The UK case study consists of four water management units across the (1: Lincolnshire, 2: Chilterns, 3: Midlands, 4: Shropshire) across Chalk and Permo-Triassic sandstone aquifers that are the two main aquifers in the UK (Figure 1). The two aquifers have contrasting hydrogeological characteristics. Regional groundwater flow and storage in the Chalk aquifer are dominated by its primary fracture network (Bloomfield, 1996) and secondary solution-enhanced fractures (Downing et al., 1993; Maurice et al., 2006). The response of Chalk groundwater hydrographs to driving meteorology is a function of regional variations in the nature of the fracture network, extent of karstification, nature of overlying superficial deposits amongst other factors (Allen

Table 1. Regional features of the four water management units summarising the ~~1)~~ area size, ~~2)~~ long-term precipitation (P) and potential evapotranspiration (PET) as calculated by Mansour and Hughes (2018) based on daily data from 1962 to 2016, ~~3)~~ hydrogeological features and ~~4)~~ main groundwater use changes in time. All water management units are shown in Figure 1. In Figure S1, the purpose and locations of recent abstraction licences are shown. Hydrogeological information and groundwater use is based on Allen et al. (1997) and complemented with additional references (see last column).

Water management unit & number of monitoring wells	Area (km^2)	Annual average (mm/yr)	Hydrogeological features	Groundwater use	Additional literature
1: Lincolnshire	1310	P: 589	Highly permeable outcrop due to dissolved fractures and weathering	Abstraction peaked in 1970 and reduced since 2000	Whitehead and Lawrence (2006)
38 wells		PET: 454	South-East of aquifer increasingly confined by superficial deposits	Abstractions exceed average recharge only during droughts	Bloomfield et al. (1995), Hutchinson et al. (2012)
2: Chilterns	1650	P: 674	Chalk aquifer partly covered by superficial deposits	Abstractions increased during 1970-2003 and decreased after 2003	Jones (1980), Jackson et al. (2011)
45 wells		PET: 485	karstification in valleys	recent abstraction is estimated on 50% of average recharge	Environment Agency (2010)
3: Midlands	1100	P: 630	Varying aquifer thickness from 120-300m	Abstraction exceeded the average recharge rates by 25% in 1980-90	Zhang and Hiscock (2010)
36 wells		PET: 476	Confined by superficial deposits in the East	Abstraction reduced in 2000 to meet average recharge	Shepley et al. (2008)
4: Shropshire	1400	P: 722	Highly variable aquifer thickness: 30-1400m	Abstraction represented 40-50% of recharge in 1970-90 and reduced after 2000.	Cuthbert (2009), Voyce (2008)
51 wells		PET: 471	Major faults interrupt groundwater flow across sandstone layers	River augmentation scheme increases abstractions during dry periods	Shepley and Streetly (2007)

et al., 1997). In the Permo-Triassic sandstone aquifer, groundwater flow and storage are influenced by variations in the matrix porosity, ~~variable~~ aquifer thickness, and to ~~a lesser extent by~~ some extend on fracture characteristics (Shepley et al., 2008; Allen et al., 1997). Faults divide the Permo-Triassic sandstone in separate sections ~~. The effect of faults varieswidely. Some~~ and the impact on regional groundwater flow varies: some faults act as hydraulic barriers ~~whilst others~~ and other times enhance permeability resulting in increased recharge (Allen et al., 1997). Hydrographs in the Permo-Triassic sandstones typically respond more slowly to driving meteorology than those in the Chalk (Bloomfield and Marchant, 2013) and are influenced by local variation in aquifer thickness and confinement by superficial deposits.

Regional hydrological features of the four water management units in the ~~two main~~ aquifers are summarised in Table 1. Two of the water management units are situated in eastern England (Lincolnshire, unit 1) and central southern England (the Chilterns, unit 2) and are underlain by the Chalk aquifer, and two of the water management units are situated in central England (East Midlands, unit 3) and north west England (Shropshire, unit 4) and are underlain by the Permo-Triassic sandstone aquifer. The largest groundwater ~~user~~ use sector in these management units is drinking water, followed by ~~industrial water use, agricultural~~ industry, agriculture and environmental water use (BGS, 2015). Groundwater use is regulated using abstraction licences, which have changed since their introduction in 1963 (Ohdedar, 2017). Since the implementation of the Water Framework Directive in 2000, ~~groundwater~~ abstraction licences follow a water balance approach to ensure 'good groundwater status' ~~. This resulted~~ resulting in a reduction of licensed groundwater ~~abstractions~~ use (Environment Agency, 2016). Specific information regarding the change in water use in these water management units was found in previous groundwater studies ~~(Whitehead and Lawrence, 2006; Environment Agency, 2010; Shepley and Streetly, 2007; Shepley et al., 2008). In all water management units, a dense network of groundwater monitoring sites and physically-based models are in use by water managers to observe the groundwater status at catchment scale~~(see fifth column of Table 1).

3 Data and Methods

3.1 Data

The analysis has been undertaken for a 30-year period (1984-2014) using precipitation, evapotranspiration, and groundwater level time series. This time period includes at least four major ~~groundwater~~ droughts with national spatial extent, namely: 1988-1994, 1995-1997, 2003-2006, and 2010-2012 (Durant, 2015).

Precipitation and potential evapotranspiration data were obtained from the GEAR dataset (Tanguy et al., 2016) and the CHESS dataset (Robinson et al., 2016). The gridded (~~1km~~1 km^2) GEAR dataset contains interpolated monthly precipitation estimates derived from the UK rain gauge network. The CHESS dataset is also gridded (~~1km~~1 km^2) and contains climate data, from which potential evapotranspiration estimates are computed using the Penman-Monteith equation. We aggregated daily potential evapotranspiration estimates to monthly sums~~for grid cells~~. For both gridded datasets (GEAR and CHESS) grid cells were extracted corresponding to groundwater well locations. ~~The precipitation~~ The 1 km^2 gridded precipitation and potential evapotranspiration sums were compared to monthly groundwater observations of the same location. This point-scale comparison assumes that the influence of precipitation is largest surrounding a groundwater monitoring site (Bloomfield and Marchant, 201
.

Precipitation estimates were converted into standardised precipitation indices (SPI) following the method of McKee et al. (1993). A gamma distribution was ~~fit to the~~ fitted to precipitation estimates and ~~standardised indices were computed after converting to a normal distribution. The SPI was calculated for a number~~ alternative distributions were also tested (Normal, Pearson III, and Logistic). Considering the use of SPI to account for delayed recharge, a large range of accumulation periods ~~(in months)~~ of precipitation (1 to 100 months) was calculated in order to ~~define the optimal correlation between~~ find the optimal correlations between precipitation and groundwater time series. For this particular use of the SPI, the 'best' fitting distribution varies (Svensson et al., 2017). Alternative distributions showed minimal differences in the computed correlations between standardised precipitation and groundwater time series~~.~~, hence we decided therefore to use the gamma distribution.

Groundwater level time series were obtained from the national groundwater database in the UK, which contains time series for both reference wells and (regular) monitoring wells. ~~208~~ 209 sites have been included in the analysis, of which 39 are reference sites and 170 (regular) monitoring sites. Reference ~~wells~~ sites were taken to represent near-natural conditions in the 30-year time period. These ~~wells~~ sites were selected from the Index and Observation wells listed in the UK Hydrometric Register (Marsh and Hannaford, 2008) ~~. All these Index and Observation wells~~ and have previously been assessed by the British Geological Survey~~and well~~. Well descriptions indicate near-natural or possible (intermittent) influence of groundwater abstraction. Wells ~~have been~~ selected for this study are categorised as ~~being~~ near-natural reflecting regional variation in groundwater levels with minimal abstraction impacts. This selection of reference wells includes 30 wells in the Chalk and 9 wells in the Permo-Triassic sandstone. ~~In the~~ Regular monitoring sites are part of the monitoring network in place in the four water management units. Initially, 660 monitoring sites were originally considered for the regional groundwater drought analysis ~~. These groundwater level time series~~ that were truncated to the 30-year analysis period ~~, after which all groundwater level observations were quality-checked~~and quality checked. Unrealistic observations were cross-validated with available meta-data, and if unexplained, removed from the dataset. Short Missing data were linearly interpolated from the last observation to the next observation in case of short sequences of missing data (less than 6 months) in the time series were filled using linear interpolation (Tallaksen and Van Lanen, 2004; Thomas et al., 2016). Time series with (Tallaksen and Van Lanen, 2004; Thomas et al., 2016). Sites with records containing longer sequences of missing data were removed from the dataset prior to the analysis leaving a total of 170 (out of the original 660) groundwater level time series that were deemed of good quality, of which 38 were located in Lincolnshire, 45 in Chilterns, 36 in Midlands, and 51 in Shropshire.

Groundwater All groundwater level time series from the reference wells and monitoring sites were standardised into the Standardised Groundwater level Index (SGI) (Bloomfield and Marchant, 2013), which is briefly explained here. Monthly groundwater observations were grouped for each calender month . The rank of each monthly observation within the 30-year time period was determined by a non-parametric fitting. The ranked observations were standardised by calendar month and within each month observations were ranked and assigned a SGI value based on an inverse normal cumulative distribution to calculate the SGI value. of the data. No distribution was fitted, but SGI values were assigned to monthly observations accounting for seasonal variation within the calendar year. The resulting SGI time series represent extremely low to below-normal $(-3 < SGI < 0)$ and higher than normal above-normal to extremely high $(0 > SGI > 3)$ monthly groundwater levels in the groundwater time series. Groundwater level observations are physically constrained by length of the screened interval of the borehole. Therefore, the lowest SGI value might indicate that the groundwater level groundwater levels fell below the borehole screen and the highest SGI value can indicate the groundwater level groundwater levels reached the surface.

Qualitative information about groundwater use was provided for each water management unit by the national regulator (the Environment Agency (EA) in England). Detailed maps were made available regarding the purpose and recent (dated at 2015) licensed abstraction volumes (see Figure S1). In addition, reports describing the EA's regional groundwater resource models and location specific groundwater studies were used as reference material to estimate indicate changes in groundwater use for each water management unit (Table 1).

3.2 Methods

The methodological framework that was developed used developed methodological framework consists of two approaches to investigate the impact of groundwater use on groundwater droughts. The first approach starts with uses a regional near-natural groundwater drought reference based on the reference wells. The SGI time series of the reference wells are clustered to identify common spatial and temporal patterns in the near-natural groundwater levels of the two aquifers. The reference wells are Reference wells were taken to represent regional groundwater variation that is primarily driven by climate and hydrogeology. Then, monitoring wells in each of the four water management units were paired to these regionally-coincident clusters of reference wells (Figure 1). The occurrence and characteristics of droughts in the monitoring wells were compared with those in the paired reference clusters to assess the potential effects of abstraction on groundwater droughts. The second approach consisted of a groundwater trend test that quantified the strength of long-term trends as a consequence of continuous impact of groundwater use in the water management units. The spatial distribution of identified trends was evaluated according to the annual abstraction licences in the water management units.

3.2.1 Time series clustering

Three hierarchical clustering methods: single linkage, complete linkage, and Ward's minimum were tested to find the most suitable and least biased approach for clustering ~~the~~ SGI time series of the reference wells (Haaf and Barthel, 2018). In each method, Euclidean distance was used as measure of similarity and cluster compositions that showed the least overlap between clusters were selected (Aghabozorgi et al., 2015). Criteria for the clusters were set by previous studies (~~for the~~ Chalk aquifer only) and known hydrogeological differences in the aquifers. For both aquifers, the minimum number of hydrograph clusters was sought that produced ~~spatially coherent~~ spatially-coherent clusters.

3.2.2 Correlation between SPI_Q-SGI

Under near-natural conditions, the ~~maximum~~ optimum correlation between standardised precipitation and groundwater indices (SPI_Q-SGI) is generally high ~~, given an~~ in unconfined aquifers (Bloomfield and Marchant, 2013). Anomalies in precipitation propagate with a relatively constant delay in recharge to the groundwater, which is due to, subsurface controls on recharge, the antecedent condition of the land surface, and non-linear response of groundwater systems (Eltahir and Yeh, 1999; Peters et al., 2006; Tallak . This constant delay is included by the optimal precipitation accumulation period ~~(Q) that accounts for delayed recharge (Bloomfield and Marchant, 2013). In previous studies, this correlation was used to demonstrate either the absence (Haas and Birk, 2017; ?; , or presence of human-influence on groundwater level observations (Lorenzo-Lacruz et al., 2017; Lee et al., 2018)~~ in the calculated SPI_Q-SGI correlation represents a long-term relationship for a certain site, as both the SPI and SGI were calculated for a continuous 30-year period including all seasons and both anomalously dry and wet periods.

The SPI_Q-SGI correlation can be reduced when groundwater level response becomes disconnected from driving precipitation under confined conditions (Bloomfield et al., 2015; Kumar et al., 2016; Lee et al., 2018) or when groundwater abstraction changes groundwater storage and levels independent from driving precipitation (Bloomfield et al., 2015; Lorenzo-Lacruz et al., 2017; Haas and Birl . In this study, the ~~presence or absence of human-influence on groundwater was determined in relation to the lowest~~ impact of confined conditions on reducing the SPI_Q-SGI ~~correlation of each near-natural reference cluster. This assumes that any deviation from a correlation between the driving precipitation and the resulting groundwater level time series is primarily due to the effects of groundwater abstraction . Strongly non-linear processes in the unsaturated zone that may reduce the correlation with groundwater levels were accounted for by using the optimal precipitation accumulation period. Based on these assumptions, the lowest~~ correlations is expected to be minimal, as only a small selection of Chalk sites are located in the semi-confined Chalk in South Lincolnshire (Table 1). On the other hand, the impact of dynamic groundwater use on SPI_Q-SGI correlations is expected to be significant, as long-term changes in groundwater use in the water management units resulted in a spatially heterogeneous pattern of irregular, decreasing, or increasing influence of abstraction on groundwater storage. For example, Ohdedar (2017) shows that groundwater use in the UK increased until the late 1980s and reduced afterwards with a large redistribution of where water is taken from to minimise the impacts on low flows.

The presence or absence of human-influence on groundwater observations in the water management units was determined on the basis of the SPI_Q-SGI ~~correlation of the~~ in each near-natural reference cluster~~is taken~~. For each cluster, the lowest

SPI$_Q$-SGI correlation was used as a threshold to differentiate long-term *influenced* from *uninfluenced* groundwater monitoring sites. Monitoring wells with high or higher SPI$_Q$-SGI correlations are regarded as (on average over the 30-year investigation period) uninfluenced and those with lower correlations as potentially human-influenced. ~~The monitoring wells are thus separated into two groups of *uninfluenced* wells and *influenced* wells.~~ An illustrated example is provided in Figure S2 showing SGI time series of a near-natural reference site and three groundwater monitoring sites. Statistical differences between the categorised uninfluenced and influenced wells were computed using a non-parametric Wilcox test.

3.2.3 Drought analysis

Groundwater droughts were defined using a threshold approach applied to the SGI series. Groundwater droughts are considered to occur when the SGI value is at or below -0.84, which corresponds to a 80th percentile ~~as used by Yevjevich (1967), Tallaksen and Van Lanen (2004), Tallaksen et al. (2009),~~ or a 'once every 5 year drought event' (Yevjevich, 1967; Tallaksen and Van Lanen . Drought characteristics were compared between ~~the reference~~ reference clusters and monitoring sites focusing on drought occurrence, frequency, duration, and magnitude.

3.2.4 Trend test

~~The last step of the analysis was~~ The second approach consisted of a monotonic trend test ~~for annual groundwater level time series. Monthly groundwater level readings were averaged using the calender year. These annual groundwater level time series were tested for monotonic trends~~ applied to all monitoring sites given the previously identified trends in human-modified groundwater systems (Thomas and Famiglietti, 2015; Sadri et al., 2016; Bhanja et al., 2017; Pathak and Dodamani, 2018). This trend test contributes to the first approach, as the SGI and SPI$_Q$-SGI correlation analysis do not specifically account for trends in groundwater time series that could result in significant trends going unnoticed. Hence, an additional trend test was introduced to compare trends in annual (averaged for each calendar year) groundwater levels to climate data (precipitation and evapotranspiration) that were extracted for grid cells corresponding to groundwater well locations from the GEAR and CHESS datasets (Tanguy et al., 2016; Robinson et al., 2016).

Trends were quantified by the trend Z value showing positive or negative deviations from the null hypothesis (no trend). Positive/negative Z values indicated increasing/decreasing trend directions. $|Z|$ values over $|2.56|$ (α = 0.01) were considered significant. ~~Our assumption was that human-influenced groundwater systems show more persistent trends compared to natural conditions. This has been shown in the literature (Thomas and Famiglietti, 2015; Sadri et al., 2016; Bhanja et al., 2017; Pathak and Dodam , when applying a trend test to relatively short time periods, e. g. 10-30 year time series.~~ Trends in groundwater level time series were tested using a modified Mann-Kendall trend test (Mann, 1945; Kendall, 1948), which includes a modification developed by Yue and Wang (2004) to account for significant ~~autocorrelation~~ auto-correlation in the annual groundwater data (Hamed, 2008). ~~The trend Z statistics (Z) indicated increasing (Z>0) or decreasing (Z<0) trend direction. Z values over |2| were considered significant (Panda et al., 2007). Climate time series (P and PET) were also tested for trends using annual data to compare groundwater trends with trends in climate data. Trends in the non-autocorrelated climate data were tested using the~~ Trends in climate time series were also calculated from annual data using a standard Mann-Kendall trend test.

4 Results

4.1 Near-natural groundwater reference clusters

The near-natural groundwater reference clusters, based on ~~the clustering of SGI time series~~ SGI clusters of the reference wells and the clustering criteria, were defined by Ward's minimum clustering technique. The Ward's minimum cluster composition shows the least overlap between clusters of the three ~~clustering techniques that were used (Figure S2~~tested clustering techniques (Figure S3). Eight clusters are identified, of which five clusters are located in the Chalk (C1-5) and three in the Permo-Triassic Sandstone (S1-3) (Figure 1). The spatial distribution of Chalk clusters (C1, C3, C4) is consistent with clusters ~~previously~~ identified by Marchant and Bloomfield (2018). ~~A separate cluster is identified~~Two additional clusters are identified, of which one is located in East Anglia ~~for~~ (5 reference wells ~~(~~in C2) ~~. The smallest cluster is C5~~ and one in South East England (2 wells ~~), for which the~~ in C5). The cluster dendrogram shows a small difference in similarity between C4 and ~~the 2 reference wells in~~ C5~~that are~~ , which is located close to the coastline (cluster dendrogram result not shown; difference between C4 and C5 is shown in Fig. S2). C1 and C3 are coincident with water management unit 1 and 2, and are used as near-natural reference for monitoring sites in those units. In the Permo-Triassic sandstone aquifer, only one spatially coherent cluster (S2) is found when all nine SGI time series are clustered (Figure 1). The cluster composition of the other two smaller clusters (S1 and S3) is not spatially coherent and there is no evidence of previous clustering studies available that can confirm these two clusters. Hence, only S2 is used as near-natural reference for monitoring sites in water management units 3 and 4.

~~The maximum~~ The optimal SPI_Q-SGI correlations of ~~the reference wells vary between~~ near-natural wells are high on average (0.79) with a range of 0.66 ~~and~~ to 0.89. These correlations are found using the optimal accumulation period, which accounts for delay in recharge that is different for each reference cluster. High SPI_Q-SGI correlations are found for both short and long accumulation periods and there was no systematic relationship between the SPI_Q-SGI correlation and the SPI accumulation period Q or SGI autocorrelation in the near-natural wells. C1 represents a relatively fast-responding section of the Chalk and has a short Q of 12.6±5.4 months. The Q of C2 and C3 is higher, respectively ~~24±8.6 and~~ 18.2±~~4~~ 4.3 and 24±8.6 months. This corresponds to the delay in groundwater recharge due to the Quaternary deposits present in these regions (Allen et al., 1997). In the South East, the Chalk is highly fractured, which is reflected by a short Q of 8±~~2.1~~ 2.2 months for C4 and C5. In the Permo-Triassic sandstone, the Q of S2 is 35±4.5 months, which confirms a slow-responding groundwater system (Allen et al., 1997).

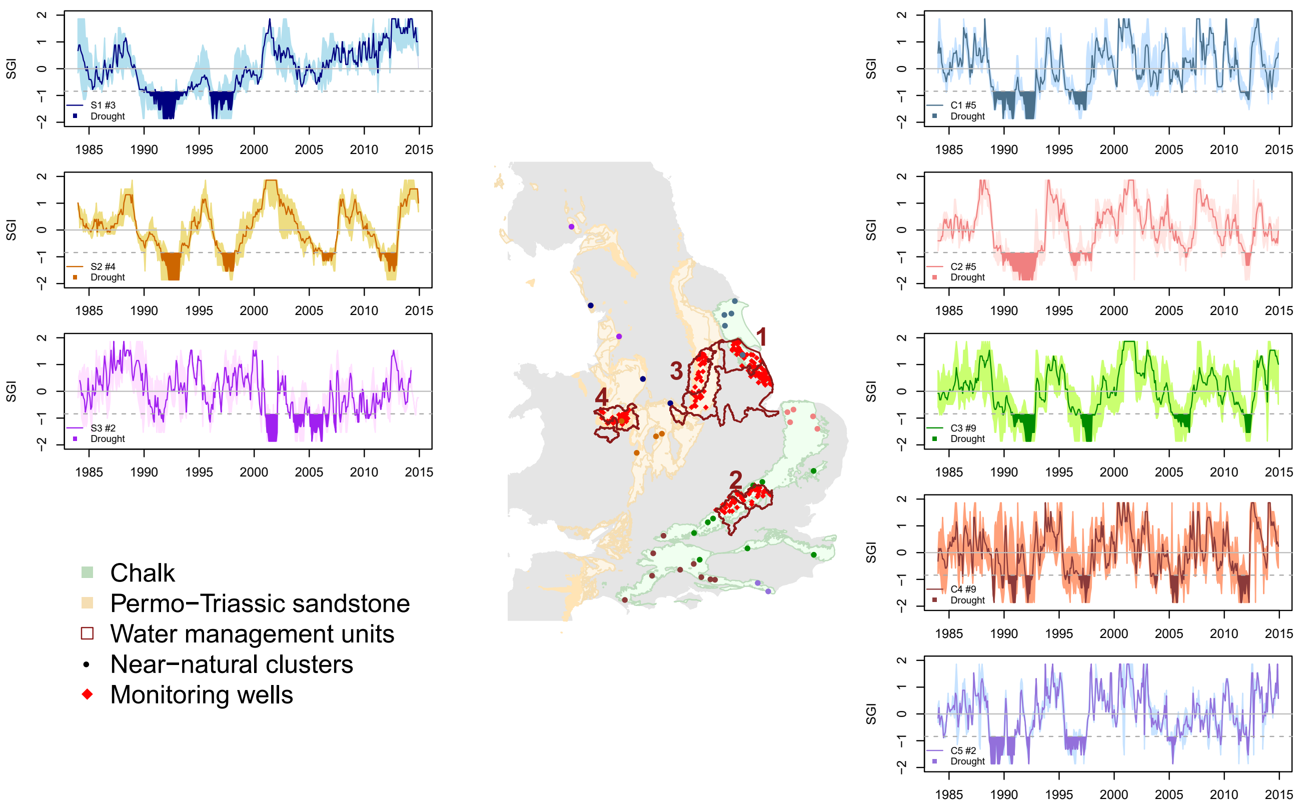
In the monitoring sites, the majority of the SPI_Q-SGI correlations are as high or higher than the minimum correlation of paired reference clusters. ~~These~~ Hence, these monitoring sites are ~~therefore considered~~considered, on average, uninfluenced by abstraction~~and the range in optimal correlations between them is most likely related to local hydrogeological settings (e.g. aquifer depth and semi-confined sections). The accumulation periods for the monitoring sites~~ . The percentage of uninfluenced sites varies between the water management units. The largest percentage is found in the Chilterns (71%), followed by the Midlands (63%), Shropshire (53%), and Lincolnshire (31%). Monitoring sites with a SPI_Q-SGI correlation below the minimum correlation of the paired reference cluster are treated as possibly influenced by abstraction.

The found optimal precipitation accumulation periods within the management units is variable and appears to be in part a function of aquifer depth and the local nature of aquifer confinement (Figure S3S4). For example, shorter accumulation periods are found in shallow sections of the aquifer (East Shropshire and West Chilterns), and in outcrops (East Lincolnshire). Longer accumulation periods are found in deep sections of the Permo-Triassic aquifer (West Shropshire) and semi-confined sections of the Permo-Triassic (Midlands) and Chalk aquifer (East Chilterns, and South East Lincolnshire). The percentage of uninfluenced sites varies between the water management units. The largest percentage is found in the Chilterns (71%), followed by the Midlands (63%), Shropshire (53%), and Lincolnshire (31%). Monitoring sites with a SPI_Q-SGI correlation below the minimum correlation of the paired reference cluster are treated as possibly influenced by abstraction.

4.2 Groundwater droughts

Groundwater droughts observed in the reference clusters show variation due to spatial patterns in both reflect both spatial and temporal variation due driving precipitation and hydrogeology . The NW-SE precipitation gradient in England results in different precipitationpatterns, which might be reflected in the variation in setting. In general, the four UK-wide droughts (1988-1993, 1995-1998, 2003-2006, and 2010-2012) are reflected in near-natural groundwater time series. Spatial patterns in driving precipitation, however, result in variable groundwater drought occurrence (Figure 1). For example, in C1 groundwater levels are low in 2003-06, but not below the drought threshold. In C2, groundwater levels are slightly lower and a short drought event is observed in the SGI cluster mean. In C3-5 and S2, however, the 2003-06 drought event was a major drought event. Spatial variation in the hydrogeology results in varying drought duration for the Chalk clusters. In central England, longer drought durations are found in clusters C2 and C3. This region is partly covered by Quaternary deposits , which delays rechargeand prolongs droughts for the reference wellsthat delays recharge. Shorter (and more frequent) events are observed in C4 and C5, which are located in highly fractured Chalk.

The On a smaller scale in the water management units, average drought characteristics (duration in months, magnitude in accumulated SGI over the drought period, and frequency) for monitoring sites in each water management unit show differences between show differences due to abstraction influence, which we have classified in, on average, uninfluenced and influenced sites, see Table 2. Shorter and less intense, but more frequent drought events are observed in the influenced sites in Lincolnshire, Chilterns, and Shropshire. In these water management units, the difference in average drought duration and frequency is significant. Droughts are observed twice as often in the influenced compared to the uninfluenced sites in Lincolnshire and Chilterns, but this difference is smaller in Shropshire. In the Midlands, The distribution of recorded drought frequency (Figure S5) shows that the difference between on average influenced and uninfluenced sites is less pronounced in Lincolnshire and Shropshire. Table 2 shows that the average drought duration of influenced sites exceeds the drought duration in uninfluenced sites in the Midlands. Longer and more intense groundwater droughts occur occurred less often in influenced sites, which is in contrast with the other water management units. However, only the difference in frequency is statistically significant. The distribution of recorded drought frequency (Figure S5) shows a majority of sites recording fewer droughts and some sites that record a higher frequency. On average, this results in a small difference between the influenced and uninfluenced sites.



~~H~~

Figure 1. Eight clusters based on the 39 reference groundwater sites in the Permo-Triassic sandstone and Chalk aquifer are shown, representing long-term near-natural groundwater level variation. All time series are standardised for the 30-year time period (1984-2014). In the centre, locations of the reference wells are shown marked by the dots in different colours for all eight clusters. The four water management units are indicated in dark red (groundwater monitoring sites in triangles). Three of these units coincide with reference clusters: 1: Lincolnshire (C1), 2: Chilterns (C3), and 4: Shropshire (S2). S2 is also used to compare water management unit 3 (Midlands) as this is the nearest reference cluster in the Permo-Triassic sandstone. In the panels left (Permo-Triassic sandstone) and right (Chalk), SGI time series are shown for each cluster, showing the cluster mean (thick line), the range of all reference wells in the cluster (shaded coloured area) and reference droughts of the cluster mean (filled area).

Table 2. Average drought characteristics ~~:~~ (duration, magnitude, and frequency) of all monitoring sites in the four water management units. 5^{th} - 95^{th} percentile of the drought characteristics are in parentheses. Distribution plots for all drought characteristics can be found in S5,S6,S7. The monitoring sites are separated using the lower limit of the cluster SPI$_Q$-SGI into on average *uninfluenced* and *influenced*. Differences between the two groups are tested for significance using a Wilcox test. Tests for which the p<0.05 are in **bold**.

	Uninfluenced wells (%)	Duration (in months)		Magnitude (from SGI)		Frequency	
		Uninfluenced Average	Influenced Average	Uninfluenced Average	Influenced Average	Uninfluenced Average	Influenced Average
1: Lincolnshire	31	**7.6 (1 - 28)**	**3.3 (1 - 12)**	-3.4 (-19 - -0.05)	-1.5 (-6.1 - -0.05)	**11.0 (4 - 17)**	**24.9 (12 - 36)**
2: Chilterns	71	**8.67 (1 - 24)**	**3.4 (1 - 11)**	**-3.9 (-15 - -0.05)**	**-1.54 (-6.5 - -0.05)**	**10.0 (5 - 18)**	**25.4 (9 - 34)**
3: Midlands	63	9.89 (1 - 36)	11.6 (1 - 45)	-4.5 (-22 - -0.05)	-5.3 (-26 - -0.05)	**9.5 (3 - 16)**	**9.0 (4 - 20)**
4: Shropshire	53	**6.8 (1 - 24)**	**5.0 (1 - 24)**	-3.1 (-14 - -0.05)	-2.3 (-12 - -0.05)	**11.9 (5 - 17)**	**15.7 (10 - 24)**

~~The drought~~ Drought characteristics in Table 2 suggest that drought events vary widely within and between water management units. These differences are shown in a combined time series plot in ~~Fig. 2 . For each water management unit, there are two plots. The upper plot shows the SGI hydrograph of the reference cluster with the cluster mean and drought events highlighted. The lower plot shows periods of drought~~ Figure 2 capturing reference droughts and drought recorded in monitoring sites that are colour-coded by the drought intensity~~at individual monitoring sites. The~~ . These monitoring sites are sorted ~~from high to low~~ based on their SPI$_Q$-SGI correlation (high to low). The cluster minimum SPI$_Q$-SGI correlation is indicated with a dashed line, i.e. 0.75 for Lincolnshire, 0.71 in the Chilterns, and 0.69 in the Midlands and Shropshire. ~~Figure 2 shows that the timing of droughts in uninfluenced wells~~ Below this minimum correlation, drought occurrence in uninfluenced sites aligns mostly with ~~droughts of~~ that of droughts in the reference clusters. ~~In Lincolnshire and Chilterns, sites with a SPI$_Q$-SGI correlation higher than the cluster minimum (uninfluenced sites) had drought events similar to the reference sites. Influenced sites (~~Categorised influenced sites (those with SPI$_Q$-SGI correlations lower than the cluster minimum) had typically shorter, but more drought events of a lower magnitude ~~. In Shropshire, additional droughts are found before and after drought events in the reference wells. However, these additional events are not exclusively observed~~ in Lincolnshire, Chilterns, and Shropshire. The distribution of drought duration in Figure S6 shows that the majority of these additional droughts is recorded in influenced sites ~~. In nearly all monitoring sites , additional drought events are found in 1984, 1989-90, 1995-96, 2005-06, and 2009, which is prior to a long drought event for all cases, except for 1984.~~ compared to uninfluenced sites in Lincolnshire, Chilterns, and Shropshire. Contrastingly, longer and more intense droughts are observed in all Midland sites in 1990-95. Droughts observed in influenced sites are also longer in 1984-1986, 1997-2001, and 2005-06 compared to the reference cluster and fewer droughts are observed in 2010-12.

The additional events in influenced sites coincide with low SGI values in the reference wells that sometimes occur prior to a long drought event. For example, additional droughts are observed in 1984, 1995-96, 2005-06, and 2014 in Lincolnshire, and in 1984-86, 2004, and 2009-10 in the Chilterns. In those periods~~in both water management units~~, the reference cluster mean was below 0, but not below the drought threshold. In the case of 1995-96, 2004, and 2009-10, these additional drought events occurred prior to a long drought event. ~~It could be that a sudden increase of groundwater use pushes groundwater level~~

~~below the drought threshold in influenced sites.~~ However, there was no no consistency between the study areas in relation to the timing of these shorter drought events. In Lincolnshire, minor droughts occur more often during reference droughts. In the Chilterns and Shropshire, more droughts are detected prior to reference droughts (Table S8). All minor droughts are shorter than the groundwater memory (auto-correlation) suggesting that these minor droughts are less likely to be related to propagated precipitation deficits, but instead are probably related to groundwater abstraction.

Drought descriptions in the literature show an increase in water demand during the 1995-97, 2003-06 and 2010-12 drought (Walker and Smithers, 1998; Marsh et al., 2013; Durant, 2015). ~~Hot summers, heatwaves or dry conditions can increase the local groundwater use. Another explanation for increased groundwater use could be related to~~ For example, Durant (2015) found that during the 1988-93 drought event evapotranspiration was exceptionally high. Impacts were mostly felt in the Chalk, particularly in regions where groundwater is the principal source of water supply where abstractions amplified the drought effects. An extreme rise in water use was reported by Walker and Smithers (1998) during the 1995-1997 drought event putting strain on drinking water supply systems in North East England. Sections of the Permo-Triassic sandstone were amongst the worst affected with drought conditions until 1998 (Durant, 2015). During the 2003-06 and 2010-12 droughts, a sudden increase in groundwater use was found that was attributed to dry weather and hot summers in the work of Marsh et al. (2007, 2013) and Durant (2015). In the work of Rey et al. (2017), low SPI_3 values were found in summer months for 1995, 1996, 2003-2006, and 2010-2011 highlighting exceptional dry weather that led to surface water use restrictions ~~(voluntarily or mandatory) that might be in place before a major groundwater drought (Rey et al., 2017; Rio et al., 2018). The~~ prior to droughts to maintain low flows. Consequently, the reduced surface water ~~availability is then replaced with groundwater,~~ abstractions were replaced by groundwater, for which use was rarely restricted (Rey et al., 2017) resulting in lowered groundwater levels and potentially aggravating ~~a groundwater drought.~~ groundwater droughts.

~~Overall~~ Over the whole investigation period, drought magnitude seems to be decreasing since the 1995-1997 drought event. ~~The droughts~~ Droughts observed in 2003-2006 and 2010-12 are shorter and of lower magnitude than the 1995-97 drought in most sites. This is seen most convincingly in Lincolnshire, Chilterns and the Midlands, where the magnitude of droughts decreases dramatically over the 30-year time period. In Shropshire, this tendency is less strong, as the 2010-12 drought was of a similar magnitude as the 1995-1997 drought.

4.3 Trends in groundwater

Significant trends in groundwater level have been detected in ~~48~~38% of all monitoring ~~wells~~ sites in the water management units. ~~The strength of these significant ($Z > |2|$) trends varies between and within management units, and both upward and downward trends have been identified~~ Of these 38%, half of the trends are upward (positive) and the other half is downward (negative) trends (Figure 3). Overall, ~~27% of the significant trends are upward (positive) ,~~ upward trends are dominating (61% of sites including significant and non-significant trends) indicating a sustained rise in the 30-year groundwater level time series~~, compared to 21% of significant downward (negative) trends that indicate~~. Fewer (39% including significant and non-significant) downward trends are detected indicating sustained lowering of groundwater levels. The presence of these

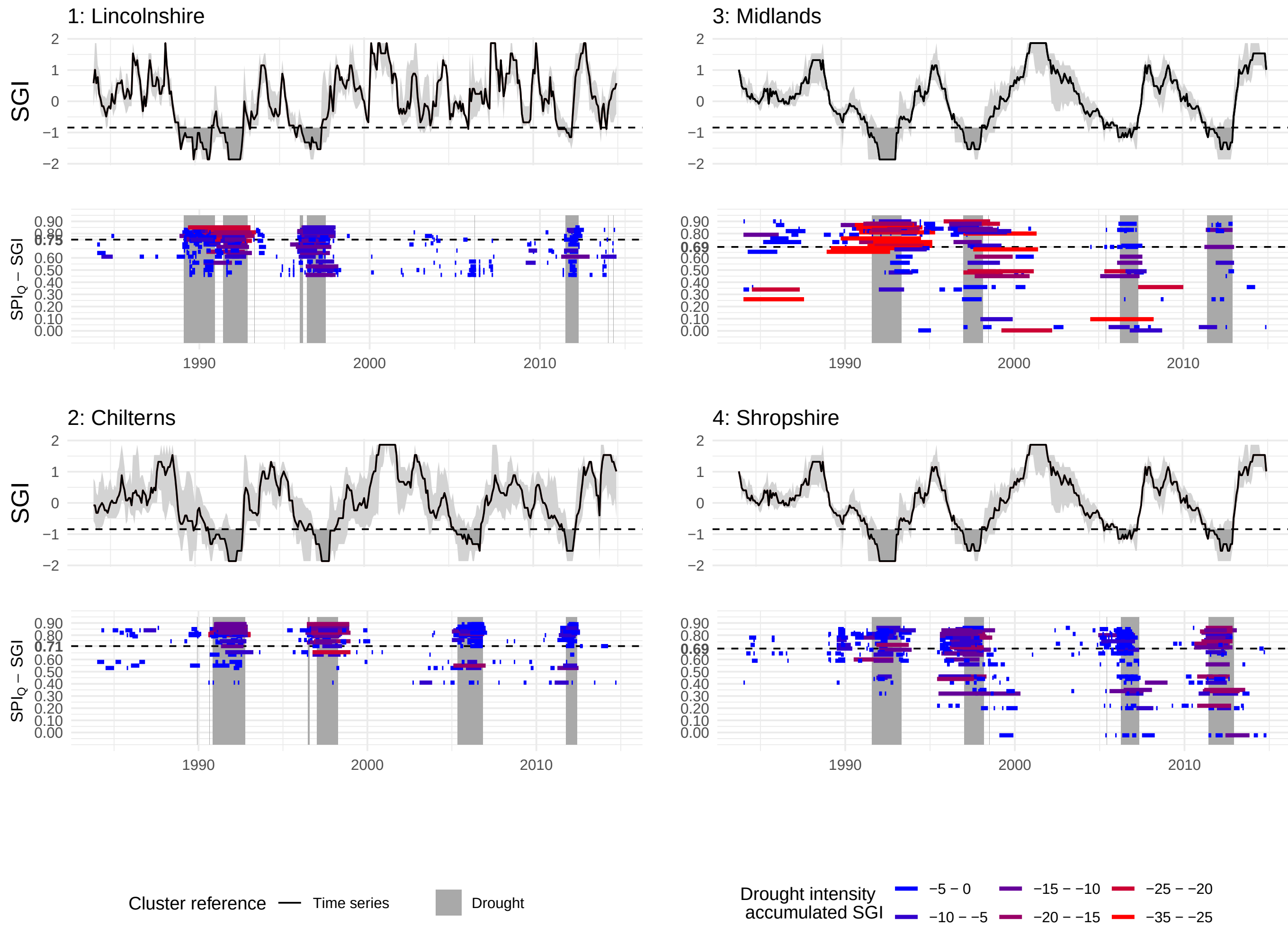


Figure 2. Drought occurrence visualised for all four water management units: 1: Lincolnshire, 2: Chilterns, 3: Midlands and 4: Shropshire. The top panel shows the SGI hydrograph of the reference cluster mean based on reference wells (see Figure 1 for the location). The range of ~~the~~ reference ~~cluster~~ clusters is coloured in grey. The dotted line represents the drought threshold for the cluster mean with shaded areas for the reference drought events. These reference drought events are also shown in long grey panels in the lower plot that shows the individual droughts as found in monitoring sites in each water management unit. The length of coloured bars indicates the drought duration ~~, whereas~~ and the colour represents ~~the~~ drought magnitude of each drought in blue-red scale for accumulated SGI.

significant trends in groundwater is notable given the weak, non-significant, trends in the 30-year precipitation and potential evapotranspiration data (P: Z = -0.75 - 1.53, PET: Z = 0 - 0.65).

The direction of trends in groundwater and their spatial coherence within the water management units show different patterns (Figure 3). In the Chalk water management units, positive trends dominate. In Lincolnshire, ~~9~~ 5 out of the total 25 positive trends are significant, compared to ~~9~~ 3 out of 32 in Chilterns. There are fewer ~~sites with a significant negative trend~~ negative trends detected in both water management units~~. This is respectively 4 out of 11~~ , but more of these are significant, respectively 7 out of 13 in Lincolnshire and 4 out of 12 in Chilterns. In Lincolnshire, sites with a negative trend are, all but one, located in the semi-confined Chalk. This is in sharp contrast with the semi-confined Chalk in Chilterns, where mainly (significant) positive trends are found. In the Permo-Triassic sandstone ~~, the majority of monitoring sites have a significant trend (69~~water management units, more significant trends are detected compared to the Chalk (63% in Midlands and ~~53~~43% in Shropshire). In the Midlands, more positive than negative trends are detected. In total, ~~18~~ 17 out of 25 positive trends are significant, compared to ~~7~~ 6 out of 11 significant negative trends. ~~Positive~~ Negative trends are mainly found in the centre of the water management unit. ~~Negative~~ Positive trends are found north and south of that. In Shropshire, more negative than positive trends are detected. 31 sites have a negative trend, of which ~~17~~ 15 significant. These trends are mainly detected in the west of the water management unit. Positive trends are mainly located east in between two fault lines (Ollerton and Childs Ercall Fault (Voyce, 2008)). ~~Half~~ Seven of these positive trends (20 in total) are significant. In Fig. 3, the maximum licensed abstraction ~~volume is~~ volumes are also shown. These licences show in which aquifer sections groundwater is primarily abstracted. However, without a record of the actual use of these licences it is impossible to directly relate ~~the~~ detected trends to these abstraction locations.

5 Discussion

~~The presented~~ Presented results of the ~~two main aquifers in the UK~~ UK case study show that groundwater droughts in the Chalk and Permo-Triassic sandstone aquifer are primarily driven by precipitation, and modified by the hydrogeology setting and groundwater use. The precipitation gradient was the primary driver for regional variation in near-natural groundwater droughts in 1989-1992 and 2003-06, which is confirmed by the work of Bryant et al. (1994) and Marsh et al. (2007). This explains the absence of a groundwater drought in ~~the~~ 2003-06 ~~period~~ in the northern Chalk (C1), compared to the southern Chalk (C2-C5). Regional variation of near-natural droughts within the different hydrogeological units was linked to the hydrogeological setting, as accumulation period varied in each reference cluster. These accumulation periods align with previous findings of Bloomfield and Marchant (2013). On a smaller scale, accumulation periods varied gradually within the water management units, as a function of aquifer depth and confinement of the aquifer, which was also found by Kumar et al. (2016), Van Loon et al. (2017) and Haas and Birk (2017). The relation between accumulation period and groundwater drought duration, as observed in the reference clusters, corresponds to the relation between groundwater memory and drought duration for near-natural observations, as found by Bloomfield and Marchant (2013).

~~Influence~~ Impact of groundwater use on groundwater droughts is detected in a subset of monitoring sites in ~~each of the~~ all four water management units. This subset often represents a minority of monitoring sites~~in the water management unit~~. Two

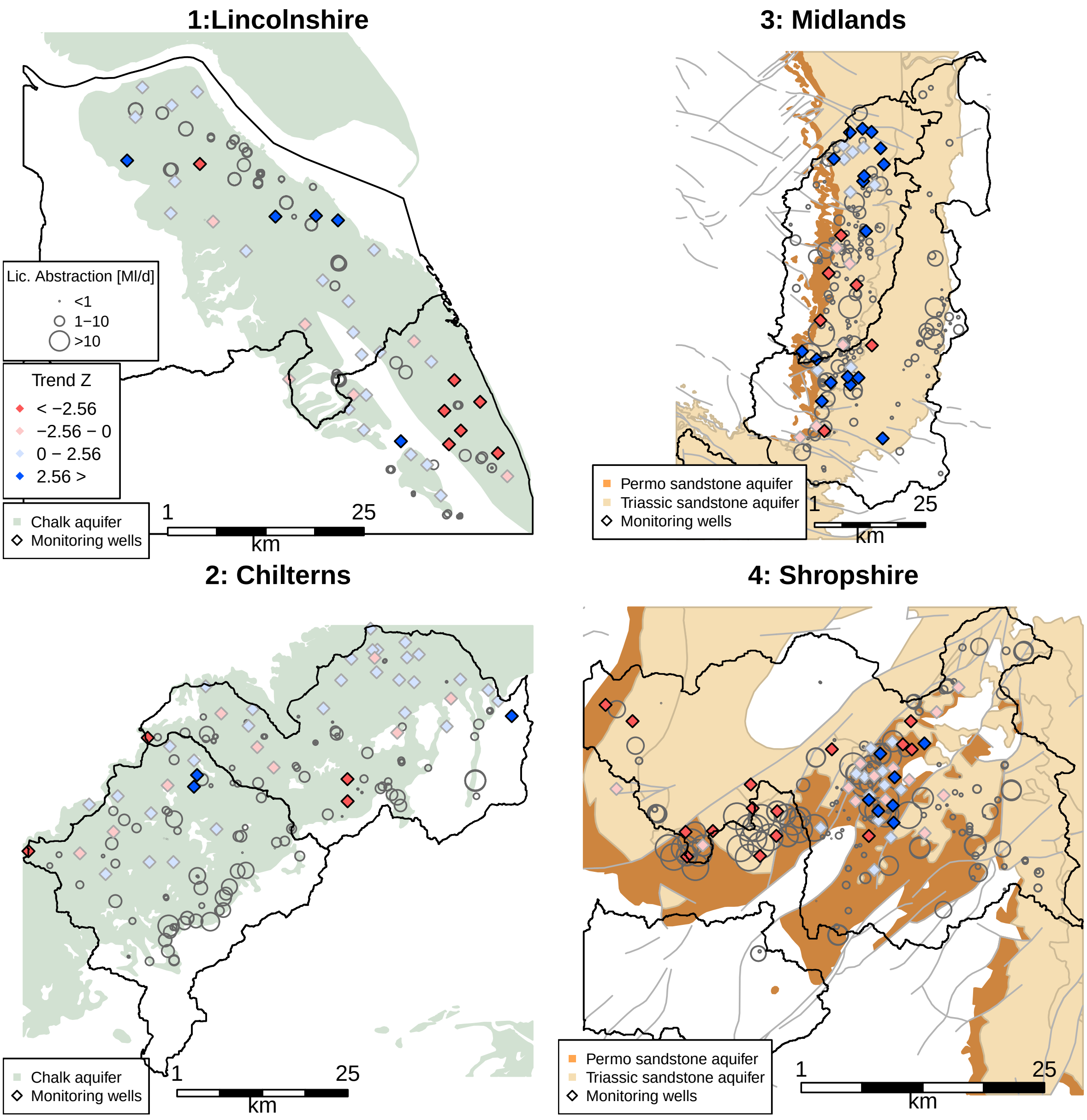


Figure 3. Trend values for monitoring wells in the four water management units (1: Lincolnshire, 2: Chilterns, 3: Midlands, 4: Shropshire). The red and blue diamonds indicate the positive or negative Z values for the Modified Mann-Kendall trend test for each monitoring well. Z values over ~~|2|~~ |2.56| indicate a significant trend in the 30-year (1984-2014) groundwater time series.

patterns are found ~~in the water management units~~ that illustrate an asymmetric impact of water use on groundwater droughts. The first pattern (found in three water management units) is that of more, but shorter and less intense droughts that are primarily observed in the~~influenced~~ , on average, influenced sites compared to uninfluenced sites. The second pattern (found in one water management unit) shows the opposite impact with less, but longer groundwater droughts in, on average, influenced compared to uninfluenced sites. Both patterns are inferred as a direct consequence of groundwater use in the water management units.

The first pattern, apparent in Lincolnshire, Chilterns, and Shropshire, shows an increase in short drought events ~~often found~~ in influenced sites that sometimes occur before a major drought event or during ~~hot summers, which is probably related to an increase in water~~ unusual dry period that results in a rapid increase in both surface water and groundwater use (Walker and Smithers, 1998; Marsh et al., 2013; Durant, 2015) and/or complementary groundwater use due to surface water use restrictions (Rey et al., 2017; Rio et al., 2018). We see the effect of this local increase in water use in our data in the temporarily lowered groundwater levels ~~,~~ resulting in additional drought events. The majority of these events occur in influenced sites, but some of the (on average) uninfluenced sites also show minor droughts. Given the high correlation in these uninfluenced sites, the minor droughts seem to not disturb the long-term average correlation. The short duration and low intensity of these additional droughts suggests that local groundwater levels recover quickly. Whether groundwater was removed from groundwater storage or capture (impacting environmental flows) remains unknown (Konikow and Leake, 2014), although the short duration and rapid recovery suggest that an equilibrium was established soon after the abstractions. Regional groundwater model studies in these three water management units show that the annual average *actual* abstractions are smaller than modelled recharge for Lincolnshire, Chilterns, and Shropshire. The ratio abstraction to recharge is 0.67 (Hutchinson et al., 2012), 0.5 (Environment Agency, 2010), 0.5 (Shepley and Streetly, 2007) for the three water management units respectively. Even though these ratios are calculated using data from different regional groundwater models, the results show that the long-term balance between groundwater use and recharge is positive, which might be ~~the reason that the overall influence of abstraction on groundwater droughts is relatively minor with a~~ related to the overall reduced drought intensity and duration for influenced sites.

The second pattern, apparent in the Midlands, shows intensified groundwater droughts that occur less often. Most of the ~~intense~~ intensified drought events are observed prior to 2001 with lengthened droughts in 1984-1986, 1990-95, 1997-2001. Lengthening of droughts is a common phenomenon in overused groundwater systems (Custodio, 2002). In the Midlands, prior to 2000, groundwater abstraction exceeded ~~the~~ modelled recharge by 25% (Shepley et al., 2008). The overabstraction resulted in lower ~~streamflows~~ streamflow in the area (Shepley et al., 2008) ~~,~~ suggesting that the balance between water removed from capture and storage was disrupted (Konikow and Leake, 2014). Reforms of water allocations in 2000 have reduced groundwater abstractions to meet the long-term water balance. These long-term changes in groundwater abstractions match with the majority of significant positive groundwater trends in the Midlands.

~~The long-term~~ Long-term influence of groundwater use ~~is~~ was inferred from identified trends in the groundwater level time series. Large spatial differences are found in the ~~strength and~~ direction of groundwater trends in both aquifers, ~~whilst~~ while trends in precipitation and potential evapotranspiration are negligible. Positive groundwater trends dominate in the water management units, which may be a result of ~~overall rising groundwater levels due to a~~ the reduction of groundwater use since 1984 (start of the investigation period of this study). A gradual or immediate reduction of water use can restore the balance between groundwater use and recharge (Gleeson et al., 2010; Konikow, 2011), although it can take decades before an equilibrium is reached (Gleeson et al., 2012). ~~Overall, groundwater droughts show a~~ This slow rise or recovery to pre-development groundwater levels is not specifically included in the classification of influenced and uninfluenced monitoring sites, as a (slow) rise in groundwater level might not disturb the propagation of precipitation anomalies. SGI and SPI anomalies could in this case synchronise well resulting in a high linear correlation, while a long-term positive trend is observed as groundwater levels slowly recover. Over longer time periods, these rising groundwater levels could also buffer precipitation anomalies. In our results, groundwater droughts show an overall reduction in magnitude and duration from 1984 to 2014. Most intense droughts are found during in the first two decades (1984-2004) of the time period. Even though this coincides with a reduction of groundwater use, more research is required to distinguish ~~the~~ climate-driven droughts from ~~the~~ human-modified droughts.

A conceptual typology is presented in Figure 4 summarising near-natural drought, two types of human-modified droughts as found in the water management units, and an extreme condition of human-modified drought. Under near-natural conditions, groundwater droughts occur given the climate forcing and hydrogeological setting (upper panel in Figure 4). ~~Under human-influenced conditions~~In human-modified environments, the impact of groundwater use on groundwater droughts is asymmetric. In regions where the annual average groundwater use is smaller than the annual average recharge, the frequency of groundwater droughts increases ~~,~~ resulting in shorter events of a lower magnitude (second panel in Figure 4). This corresponds to the 'dynamic sustainable range' as presented in the conceptual model of ~~?~~Gleeson et al. (2020). In regions where the annual average groundwater use approaches annual average recharge, the opposite is found with less, but prolonged droughts of higher magnitude and duration (third panel in Figure 4) corresponding to strategic aquifer depletion, when meeting the dynamic sustainable range over a long time scale ~~(?)~~(Gleeson et al., 2020). The last panel shows ~~the~~ extreme conditions of groundwater depletion, in which groundwater droughts are not recovering by the average annual recharge and groundwater levels tend to fall consistently. These extremes conditions are not identified in the UK, but heavily intensified and lengthened droughts are found in California (He et al., 2017), Australia (Leblanc et al., 2009), Spain (Van Loon and Van Lanen, 2013), Bangladesh (Mustafa et al., 2017) and India (Asoka et al., 2017).

Further research is required to analyse the ~~effects of water use changing over timeto groundwater droughts~~modifying effects on droughts of a change in water use over time. In this study, we have investigated the overall long-term impact of groundwater use using monotonic trends in groundwater. ~~A~~ However, a different methodology is required to evaluate the impact of new water regulations on groundwater droughts (Bhanja et al., 2017). For example, an observation-modelling or conceptual modelling approach can be used to differentiate pre- and post-regulation groundwater droughts (Van Loon et al., 2016b; Kakaei et al., 2019; Liu et al., 2016). This future modelling work could also provide long-term context for water management effects, natural variability, non-stationary effects of anthropogenic climate change (specifically warming) on changes in groundwater drought characteristics (Bloomfield et al., 2019).

Further applications of this study could be beneficial for water regulators and scientists alike, as the presented conceptual typology can be used to investigate the impact of groundwater use without having to obtain time series of actual groundwater abstractions. The developed methodology shows how qualitative information on groundwater use and annual long-term averages aid to get a better understanding of asymmetric impact of groundwater use on groundwater droughts. Considering the

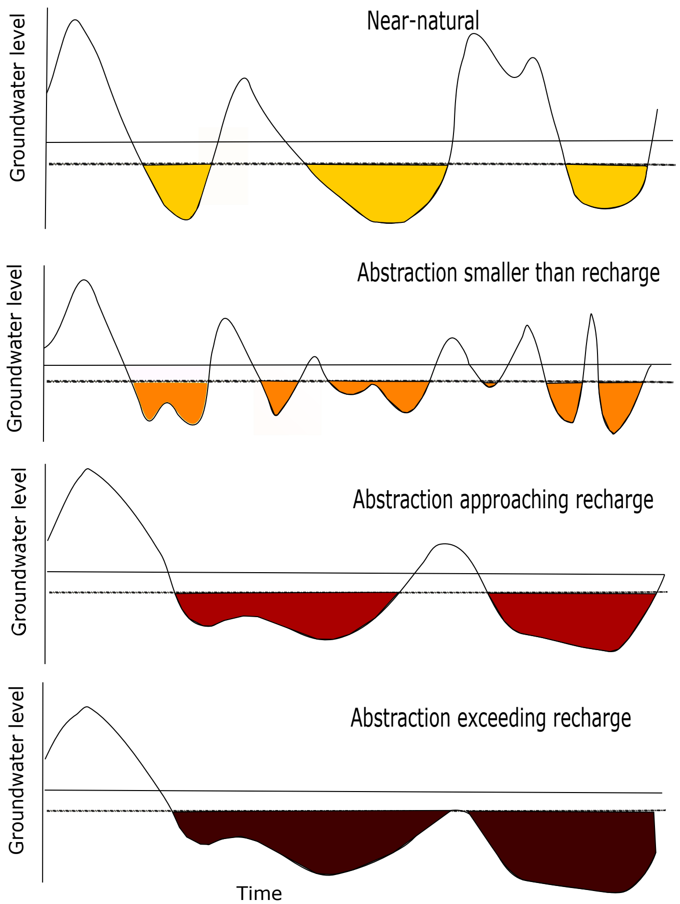


Figure 4. Conceptual figure summarising the near-natural groundwater droughts , two identified patterns in the four water management units, and one extreme scenario of groundwater depletion. The (toppanel shows groundwater droughts under near-natural conditions, the lower) and three panels show human-modified groundwater droughts under human-influenced conditions with increasing intensity of impact of groundwater use. The second top panel shows typical an example of near-natural groundwater droughts, followed by human modified droughts when annual average abstractions are smaller than the annual average groundwater recharge (second panel; identified in the three water management units in the UK). The third panel illustrates modified groundwater droughts when annual average abstractions approaches recharge (identified in one water management unit in the UK), and the last panel shows extreme groundwater drought conditions when average annual abstractions exceed recharge.

large-scale modification of the hydrological cycle and the consequences for droughts (Van Loon et al., 2016a), it is important to further this approach and investigate the sustainable use of groundwater resources ~~(?)~~(Gleeson et al., 2020).

6 Conclusions

The impact of groundwater use on groundwater droughts is investigated based on a comparison of potentially influenced groundwater monitoring sites and near-natural ~~, or largely uninfluenced reference sites~~ reference sites in the UK. Results show that long-term groundwater use has an asymmetric impact on groundwater droughts for a subset of influenced groundwater ~~sites~~monitoring sites in water management units in the UK. A conceptual typology summarises these different patterns in groundwater drought occurrence, duration, and magnitude. The first type (identified in three water management units) shows an increase in groundwater droughts with a low magnitude, of which the timing sometimes coincides with periods of a high water demand~~, for instance during heatwaves~~. This is found in three water management units where the long-term water balance is positive and annual average groundwater abstractions are less than ~~the~~ groundwater recharge. The second type is marked by lengthened, more intense groundwater droughts. This is found in one water management unit where annual average groundwater abstractions temporarily exceeded recharge. The balance between long-term groundwater use and recharge seems to explain the asymmetric impact of groundwater use on groundwater droughts. However, more research is required to investigate the impact of changes in water use. During the period of investigation, regulated groundwater abstractions have reduced and our results show a majority of rising groundwater trends based on 30 years of data. Further research could potentially indicate how droughts are affected by these changes in water use.

In conclusion, this study presents a conceptual typology to analyse groundwater droughts under human-modified conditions. We found that human-modified droughts differ in frequency, magnitude, and duration dependent on the long-term proportional groundwater use compared to recharge. This highlights the relation between long-term and short-term groundwater sustainability.

Code availability. The code is available upon request.

Data availability. The raw groundwater time series and abstraction locations can be obtained via the Environment Agency. Standardised groundwater level time series is available upon request.

Author contributions. DW, AVL, BP, DH conceived and designed the study. DW performed the analysis and wrote the paper, supervised by AVL, BP, DH. All authors contributed to the manuscript.

Competing interests. The authors declare no conflict of interest.

Acknowledgements. We would like to thank Richard Morgan and Catriona Finch for providing data and their valuable feedback in the initial stages of this study, and Michael Kehinde, Vicky Fry, Alex Chambers, and Kevin Voyce for providing groundwater monitoring data and background material. The study has benefited from valuable discussions during meetings and workshops of the 'Drought in the Anthropocene' working group of the IAHS Panta Rhei network, and we would like to thank Henny Van Lanen in particular. Financial support for DW was provided by a CENTA NERC grant (NE/lL002493/1) and CASE studentship of British Geological Survey (GA/16S/023). JPB publishes with permission of the Director, British Geological Survey (NERC, UKRI).

References

Aghabozorgi, S., Shirkhorshidi, A. S., and Wah, T. Y.: Time-series clustering – A decade review, Information Systems, 53, 16 – 38, 2015.

AghaKouchak, A.: Recognize anthropogenic drought, Nature, 524, 409–411, 2015.

Allen, D., Brewerton, L., Coleby, L., Gibbs, B., Lewis, M., MacDonald, A., Wagstaff, S., and Williams, A.: The physical properties of major aquifers in England and Wales, 1997.

Asoka, A., Gleeson, T., Wada, Y., and Mishra, V.: Relative contribution of monsoon precipitation and pumping to changes in groundwater storage in India, Nature Geoscience, 10, 109–117, 2017.

BGS: Current UK groundwater use, Website, https://www.bgs.ac.uk/research/groundwater/waterResources/GroundwaterInUK/2015.html, 2015.

Bhanja, S. N., Mukherjee, A., Rodell, M., Wada, Y., Chattopadhyay, S., Velicogna, I., Pangaluru, K., and Famiglietti, J. S.: Groundwater rejuvenation in parts of India influenced by water-policy change implementation, Scientific reports, 7, 7453, 2017.

Bloomfield, J.: Characterisation of hydrogeologically significant fracture distributions in the Chalk: an example from the Upper Chalk of southern England, Journal of Hydrology, 184, 355 – 379, https://doi.org/10.1016/0022-1694(95)02954-0, 1996.

Bloomfield, J. P. and Marchant, B. P.: Analysis of groundwater drought building on the standardised precipitation index approach, Hydrology and Earth System Sciences, 17, 4769–4787, https://doi.org/10.5194/hess-17-4769-2013, 2013.

Bloomfield, J. P., Brewerton, L. J., and Allen, D. J.: Regional trends in matrix porosity and dry density of the Chalk of England, Quarterly Journal of Engineering Geology and Hydrogeology, 28, S131–S142, https://doi.org/10.1144/GSL.QJEGH.1995.028.S2.04, 1995.

Bloomfield, J. P., Marchant, B. P., Bricker, S. H., and Morgan, R. B.: Regional analysis of groundwater droughts using hydrograph classification, Hydrology and Earth System Sciences, 19, 4327–4344, https://doi.org/10.5194/hess-19-4327-2015, 2015.

Bloomfield, J. P., Marchant, B. P., and McKenzie, A. A.: Changes in groundwater drought associated with anthropogenic warming, Hydrology and Earth System Sciences, 23, 1393–1408, https://doi.org/10.5194/hess-23-1393-2019, 2019.

Bryant, S., Arnell, N., and Law, F.: The 1988–92 drought in its historical perspective, Water and Environment Journal, 8, 39–51, 1994.

Chang, T. and Teoh, C.: Use of the kriging method for studying characteristics of ground water droughts, Water Resources Bulletin, 31, 1001–1007, 1995.

Christian-Smith, J., Levy, M. C., and Gleick, P. H.: Maladaptation to drought: a case report from California, USA, Sustainability Science, 10, 491–501, https://doi.org/10.1007/s11625-014-0269-1, 2015.

Custodio, E.: Aquifer overexploitation: what does it mean?, Hydrogeology Journal, 10, 254–277, https://doi.org/10.1007/s10040-002-0188-6, 2002.

Cuthbert, M. O.: An improved time series approach for estimating groundwater recharge from groundwater level fluctuations, Water Resources Research, 46, https://doi.org/10.1029/2009WR008572, 2009.

de Graaf, I. E., Gleeson, T., van Beek, L. R., Sutanudjaja, E. H., and Bierkens, M. F.: Environmental flow limits to global groundwater pumping, Nature, 574, 90–94, 2019.

Döll, P., Hoffmann-Dobrev, H., Portmann, F., Siebert, S., Eicker, A., Rodell, M., Strassberg, G., and Scanlon, B.: Impact of water withdrawals from groundwater and surface water on continental water storage variations, Journal of Geodynamics, 59, 143–156, https://doi.org/10.1016/j.jog.2011.05.001, 2012.

Döll, P., Müller Schmied, H., Schuh, C., Portmann, F. T., and Eicker, A.: Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites, Water Resources Research, 50, 5698–5720, https://doi.org/10.1002/2014WR015595, 2014.

Downing, R. A., Price, M., and Jones, G.: The hydrogeology of the Chalk of north-west Europe, Clarendon Press, 1993.

Durant, M.: Description of groundwater droughts in the UK: 1890 to 2015, Nottingham, UK, British Geological Survey, (OR/15/007), 52, 2015.

Eltahir, E. A. B. and Yeh, P. J. F.: On the asymmetric response of aquifer water level to floods and droughts in Illinois, Water Resources Research, 35, 1199–1217, https://doi.org/10.1029/1998WR900071, 1999.

Environment Agency: Vale of St. Albans Numerical Groundwater Model Final Report, Tech. rep., 2010.

Environment Agency: Managing water abstraction, Tech. Rep. May, EA, Bristol, 2016.

Famiglietti, J. S.: The global groundwater crisis, Nature Climate Change, 4, 945–948, https://doi.org/10.1038/nclimate2425, 2014.

Gleeson, T. and Richter, B.: How much groundwater can we pump and protect environmental flows through time? Presumptive standards for conjunctive management of aquifers and rivers, River Research and Applications, 34, 83–92, 2017.

Gleeson, T., VanderSteen, J., Sophocleous, M. a., Taniguchi, M., Alley, W. M., Allen, D. M., and Zhou, Y.: Groundwater sustainability strategies, Nature Geoscience, 3, 378–379, https://doi.org/10.1038/ngeo881, 2010.

Gleeson, T., Alley, W. M., Allen, D. M., Sophocleous, M. A., Zhou, Y., Taniguchi, M., and VanderSteen, J.: Towards sustainable groundwater use: setting long-term goals, backcasting, and managing adaptively, Groundwater, 50, 19–26, 2012.

Gleeson, T., Cuthbert, M., Ferguson, G., and Perrone, D.: Global Groundwater Sustainability, Resources, and Systems in the Anthropocene, Annual Review of Earth and Planetary Sciences, 48, 431–463, https://doi.org/10.1146/annurev-earth-071719-055251, https://doi.org/10.1146/annurev-earth-071719-055251, 2020.

Gun, J.: Groundwater and global change: trends, opportunities and challenges. United Nations World Water Assessment Programme, UNESCO, SIDE Publications Series, 1, 1–38, 2012.

Haaf, E. and Barthel, R.: An inter-comparison of similarity-based methods for organisation and classification of groundwater hydrographs, Journal of Hydrology, 559, 222 – 237, http://www.sciencedirect.com/science/article/pii/S0022169418301112, 2018.

Haas, J. C. and Birk, S.: Characterizing the spatiotemporal variability of groundwater levels of alluvial aquifers in different settings using drought indices, Hydrology and Earth System Sciences, 21, 2421–2448, https://doi.org/10.5194/hess-21-2421-2017, 2017.

Hamed, K. H.: Trend detection in hydrologic data: The Mann–Kendall trend test under the scaling hypothesis, Journal of Hydrology, 349, 350 – 363, https://doi.org/10.1016/j.jhydrol.2007.11.009, 2008.

He, X., Wada, Y., Wanders, N., and Sheffield, J.: Intensification of hydrological drought in California by human water management, Geophysical Research Letters, 44, 1777–1785, 2017.

Hutchinson, M., Ingram, R., Grout, M., and Hayes, P.: A successful model: 30 years of the Lincolnshire Chalk model, Geological Society, London, Special Publications, 364, 173–191, 2012.

Jackson, C. R., Meister, R., and Prudhomme, C.: Modelling the effects of climate change and its uncertainty on UK Chalk groundwater resources from an ensemble of global climate model projections, Journal of Hydrology, 399, 12 – 28, https://doi.org/https://doi.org/10.1016/j.jhydrol.2010.12.028, 2011.

Jackson, C. R., Bloomfield, J. P., and Mackay, J. D.: Evidence for changes in historic and future groundwater levels in the UK, Progress in Physical Geography: Earth and Environment, 39, 49–67, https://doi.org/10.1177/0309133314550668, 2015.

Jones, D. K.: The shaping of southern England, Academic Press, 1980.

Kakaei, E., Moradi, H. R., Moghaddam Nia, A., and Van Lanen, H. A.: Quantifying Positive and Negative Human-Modified Droughts in the Anthropocene: Illustration with Two Iranian Catchments, Water, 11, https://doi.org/10.3390/w11050884, 2019.

Kendall, M. G.: Rank correlation methods. 1948, London: Charles Griffin, p. 160, 1948.

Konikow, L. F.: Contribution of global groundwater depletion since 1900 to sea-level rise, Geophysical Research Letters, 38, 2011.

Konikow, L. F. and Leake, S. A.: Depletion and capture: revisiting "the source of water derived from wells", Ground water, 52, 100–111, https://doi.org/10.1111/gwat.12204, 2014.

Kumar, R., Musuuza, J. L., Van Loon, A. F., Teuling, A. J., Barthel, R., Ten Broek, J., Mai, J., Samaniego, L., and Attinger, S.: Multiscale evaluation of the Standardized Precipitation Index as a groundwater drought indicator, Hydrology and Earth System Sciences, 20, 1117–1131, https://doi.org/10.5194/hess-20-1117-2016, 2016.

Leblanc, M. J., Tregoning, P., Ramillien, G., Tweed, S. O., and Fakes, A.: Basin-scale, integrated observations of the early 21st century multiyear drought in southeast Australia, Water Resources Research, 45, https://doi.org/10.1029/2008WR007333, 2009.

Lee, J. M., Park, J. H., Chung, E., and Woo, N. C.: Assessment of Groundwater Drought in the Mangyeong River Basin, Korea, Sustainability, 10, 831, 2018.

Li, B. and Rodell, M.: Evaluation of a model-based groundwater drought indicator in the conterminous US, Journal of Hydrology, 526, 78–88, 2015.

Liu, Y., Ren, L., Zhu, Y., Yang, X., Yuan, F., Jiang, S., and Ma, M.: Evolution of hydrological drought in human disturbed areas: a case study in the laohahe catchment, Northern China, Advances in Meteorology, 2016.

Lorenzo-Lacruz, J., Garcia, C., and Morán-Tejeda, E.: Groundwater level responses to precipitation variability in Mediterranean insular aquifers, Journal of Hydrology, 552, 516 – 531, https://doi.org/https://doi.org/10.1016/j.jhydrol.2017.07.011, 2017.

Mann, H. B.: Nonparametric Test Against Trend, Econometrica, 13, 245–259, https://doi.org/10.2307/1907187 , 1945.

Mansour, M. and Hughes, A.: Summary of results for national scale recharge modelling under conditions of predicted climate change, http://nora.nerc.ac.uk/id/eprint/521605/, British Geological Survey 136pp. (OR/17/026) (Unpublished), 2018.

Marchant, B. and Bloomfield, J.: Spatio-temporal modelling of the status of groundwater droughts, Journal of Hydrology, 564, 397 – 413, https://doi.org/https://doi.org/10.1016/j.jhydrol.2018.07.009, 2018.

Marsh, T. and Hannaford, J., eds.: UK hydrometric register. A catalogue of river flow gauging stations and observation wells and boreholes in the United Kingdom together with summary hydrometric and spatial statistics, Hydrological Data UK, Centre for Ecology & Hydrology, http://nora.nerc.ac.uk/id/eprint/3093/, 2008.

Marsh, T., Cole, G., and Wilby, R.: Major droughts in England and Wales, 1800-2006, Royal Meteorological Society, 62, 87–93, https://doi.org/10.1002/wea.67, 2007.

Marsh, T., Parry, S., Kendon, M., and Hannaford, J.: The 2010-12 drought and subsequent extensive flooding: a remarkable hydrological transformation, Tech. rep., Wallingford, 2013.

Maurice, L., Atkinson, T., Barker, J., Bloomfield, J., Farrant, A., and Williams, A.: Karstic behaviour of groundwater in the English Chalk, Journal of Hydrology, 330, 63 – 70, https://doi.org/https://doi.org/10.1016/j.jhydrol.2006.04.012, 2006.

McKee, T. B., Doesken, N. J., and Kleist, J.: The relationship of drought frequency and duration to time scales, AMS 8th Conference on Applied Climatology, pp. 179–184, https://doi.org/citeulike-article-id:10490403, 1993.

Mishra, A. K. and Singh, V. P.: A review of drought concepts, Journal of Hydrology, 391, 202–216, https://doi.org/10.1016/j.jhydrol.2010.07.012, 2010.

Mustafa, S. M. T., Abdollahi, K., Verbeiren, B., and Huysmans, M.: Identification of the influencing factors on groundwater drought and depletion in north-western Bangladesh, Hydrogeology Journal, 25, 1357–1375, https://doi.org/10.1007/s10040-017-1547-7, 2017.

Ohdedar, B.: Groundwater law, abstraction, and responding to climate change: assessing recent law reforms in British Columbia and England, Water International, 42, 691–708, https://doi.org/10.1080/02508060.2017.1351059, 2017.

Panda, D. K., Mishra, A., Jena, S., James, B., and Kumar, A.: The influence of drought and anthropogenic effects on groundwater levels in Orissa, India, Journal of Hydrology, 343, 140–153, https://doi.org/10.1016/j.jhydrol.2007.06.007, 2007.

Parry, S., Wilby, R., Prudhomme, C., Wood, P., and McKenzie, A.: Demonstrating the utility of a drought termination framework: prospects for groundwater level recovery in England and Wales in 2018 or beyond, Environmental Research Letters, 13, https://doi.org/10.1088/1748-9326/aac78c, 2018.

Pathak, A. A. and Dodamani, B. M.: Trend Analysis of Groundwater Levels and Assessment of Regional Groundwater Drought: Ghataprabha River Basin, India, Natural Resources Research, https://doi.org/10.1007/s11053-018-9417-0, 2018.

Peters, E., Bier, G., Van Lanen, H. A. J., and Torfs, P. J. J. F.: Propagation and spatial distribution of drought in a groundwater catchment, Journal of Hydrology, 321, 257–275, https://doi.org/10.1016/j.jhydrol.2005.08.004, 2006.

Rey, D., Holman, I. P., and Knox, J. W.: Developing drought resilience in irrigated agriculture in the face of increasing water scarcity, Regional Environmental Change, 17, 1527–1540, 2017.

Rio, M., Rey, D., Prudhomme, C., and Holman, I. P.: Evaluation of changing surface water abstraction reliability for supplemental irrigation under climate change, Agricultural Water Management, 206, 200–208, https://doi.org/10.1016/j.agwat.2018.05.005, 2018.

Robinson, E., Blyth, E., Clark, D., Comyn-Platt, E., Finch, J., and Rudd, A.: Climate hydrology and ecology research support system potential evapotranspiration dataset for Great Britain (1961-2015), https://doi.org/10.5285/8baf805d-39ce-4dac-b224-c926ada353b7, 2016.

Russo, T. A. and Lall, U.: Depletion and response of deep groundwater to climate-induced pumping variability, Nature Geoscience, 10, 105, 2017.

Sadri, S., Kam, J., and Sheffield, J.: Nonstationarity of low flows and their timing in the eastern United States, Hydrology and Earth System Sciences, 20, 633–649, https://doi.org/10.5194/hess-20-633-2016, 2016.

Shepley, M. and Streetly, M.: The estimation of ‘natural’summer outflows from the Permo-Triassic Sandstone aquifer, UK, Quarterly Journal of Engineering Geology and Hydrogeology, 40, 213–227, 2007.

Shepley, M., Pearson, A., Smith, G., and Banton, C.: The impacts of coal mining subsidence on groundwater resources management of the East Midlands Permo-Triassic Sandstone aquifer, England, Quarterly Journal of Engineering Geology and Hydrogeology, 41, 425–438, https://doi.org/10.1144/1470-9236/07-210, 2008.

Shepley, M. G., Whiteman, M., Hulme, P., and Grout, M.: Introduction: groundwater resources modelling: a case study from the UK, Geological Society, London, Special Publications, 364, 1–6, 2012.

Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P., and Portmann, F. T.: Groundwater use for irrigation – a global inventory, Hydrology and Earth System Sciences, 14, 1863–1880, https://doi.org/10.5194/hess-14-1863-2010, 2010.

Svensson, C., Hannaford, J., and Prosdocimi, I.: Statistical distributions for monthly aggregations of precipitation and streamflow in drought indicator applications, Water Resources Research, 53, 999–1018, https://doi.org/10.1002/2016WR019276, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016WR019276, 2017.

Tallaksen, L. M. and Van Lanen, H. A. J.: Hydrological drought : processes and estimation methods for streamflow and groundwater, Elsevier, 2004.

Tallaksen, L. M., Hisdal, H., and Lanen, H. A. V.: Space–time modelling of catchment scale drought characteristics, Journal of Hydrology, 375, 363–372, https://doi.org/10.1016/j.jhydrol.2009.06.032, 2009.

Tanguy, M., Dixon, H., Prosdocimi, I., Morris, D. G., and Keller, V. D. J.: Gridded estimates of daily and monthly areal rainfall for the United Kingdom (1890-2015), https://doi.org/10.5285/33604ea0-c238-4488-813d-0ad9ab7c51ca, 2016.

Taylor, R. G., Scanlon, B., Doell, P., Rodell, M., van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J. S., Edmunds, M., Konikow, L., Green, T. R., Chen, J., Taniguchi, M., Bierkens, M. F. P., MacDonald, A., Fan, Y., Maxwell, R. M., Yechieli, Y., Gurdak, J. J., Allen, D. M., Shamsudduha, M., Hiscock, K., Yeh, P. J. F., Holman, I., and Treidel, H.: Ground water and climate change, Nature Climate Change, 3, 322–329, https://doi.org/10.1038/NCLIMATE1744, 2013.

Thomas, B., Landerer, F., Wiese, D., and Famiglietti, J.: A comparison of watershed storage trends over the eastern and upper Midwestern regions of the United States, 2003–2015, Water Resources Research, 52, 6335–6347, https://doi.org/10.1002/2016WR018617, 2016.

Thomas, B. F. and Famiglietti, J. S.: Sustainable Groundwater Management in the Arid Southwestern US: Coachella Valley, California, Water Resources Management, 29, 4411–4426, https://doi.org/10.1007/s11269-015-1067-y, 2015.

Van Lanen, H. A.: Drought propagation through the hydrological cycle, IAHS publication, 308, 122, 2006.

Van Lanen, H. A. J., Wanders, N., Tallaksen, L. M., and Van Loon, A. F.: Hydrological drought across the world: impact of climate and physical catchment structure, Hydrol. Earth Syst. Sci, 17, 1715–1732, https://doi.org/10.5194/hess-17-1715-2013, 2013.

Van Loon, A. F.: Hydrological drought explained, Wiley Interdisciplinary Reviews: Water, 2, 359–392, https://doi.org/10.1002/wat2.1085, 2015.

Van Loon, A. F. and Van Lanen, H. A. J.: Making the distinction between water scarcity and drought using an observation-modeling framework, Water Resources Research, 49, 1483–1502, https://doi.org/10.1002/wrcr.20147, 2013.

Van Loon, A. F., Gleeson, T., Clark, J., Van Dijk, A. I. J. M., Stahl, K., Hannaford, J., Di Baldassarre, G., Teuling, A. J., Tallaksen, L. M., Uijlenhoet, R., Hannah, D. M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., Rangecroft, S., Wanders, N., and Van Lanen, H. A. J.: Drought in the Anthropocene, Nature Geoscience, 9, 89–91, https://doi.org/10.1038/ngeo2646, 2016a.

Van Loon, A. F., Stahl, K., Di Baldassarre, G., Clark, J., Rangecroft, S., Wanders, N., Gleeson, T., Van Dijk, A. I. J. M., Tallaksen, L. M., Hannaford, J., Uijlenhoet, R., Teuling, A. J., Hannah, D. M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., and Van Lanen, H. A. J.: Drought in a human-modified world: reframing drought definitions, understanding, and analysis approaches, Hydrology and Earth System Sciences, 20, 3631–3650, https://doi.org/10.5194/hess-20-3631-2016, 2016b.

Van Loon, A. F., Kumar, R., and Mishra, V.: Testing the use of standardised indices and GRACE satellite data to estimate the European 2015 groundwater drought in near-real time, Hydrology and Earth System Sciences, 21, 1947–1971, https://doi.org/10.5194/hess-21-1947-2017, 2017.

Voyce, K.: Groundwater Management: the Shropshire Groundwater Scheme, Proc. Shropsh. Geol. Soc, 14, 20–29, 2008.

Wada, Y., van Beek, L. P. H., Wanders, N., and Bierkens, M. F. P.: Human water consumption intensifies hydrological drought worldwide, Environmental Research Letters, 8, 034 036, https://doi.org/10.1088/1748-9326/8/3/034036, 2013.

Walker, S. and Smithers, H.: A review of the 1995–96 drought in the North West, Water and Environment Journal, 12, 273–279, 1998.

Whitehead, E. and Lawrence, A.: The Chalk aquifer system of Lincolnshire, Tech. rep., Keyworth, Nottingham, contributors: J P Bloomfield, P J McConvey, J E Cunningham, M G Sumbler, D Watling, M Hutchinson, 2006.

Wilhite, D. A., ed.: Droughts: A Global Assesment, Routledge, 2000.

Yevjevich, V. M.: An objective approach to definitions and investigations of continental hydrologic droughts, Hydrology papers (Colorado State University); no. 23, 1967.

Yue, S. and Wang, C.: The Mann-Kendall Test Modified by Effective Sample Size to Detect Trend in Serially Correlated Hydrological Series, Water Resources Management, 18, 201–218, https://doi.org/10.1023/B:WARM.0000043140.61082.60, 2004.

Zhang, H. and Hiscock, K.: "Modelling the impact of forest cover on groundwater resources: A case study of the Sherwood Sandstone aquifer in the East Midlands, UK", "Journal of Hydrology", 392, 136 – 149, https://doi.org/https://doi.org/10.1016/j.jhydrol.2010.08.002, 2010.