General response to Reviewer 1

We thank the Reviewer for their careful reading and constructive comments to improve the clarity of the submitted manuscript. We will address each comment in detail in the rebuttal, but we have addressed the first five major comments in this Author's response to clarify important aspects of the paper and show how we will include the provided suggestions in the revised manuscript.

Title and Introduction:
We agree with the Reviewer that the title could be stronger once reflecting the findings and implications, rather than the impact of drought policies. We will re-evaluate our choice of title and consider rephrasing.

Regarding the introduction, we will rephrase the scope to clarify the reasoning behind the focus on hydrological droughts, including both baseflow and groundwater droughts. Because of the specific focus on base flow and groundwater, we considered different hydrogeological settings that are associated with different types of drought characteristics.

Comment 1: 'High, medium, low groundwater storage systems (L70, L143) are crucial definitions to understand the analysis. I wonder if the authors described/used here large, medium and small (or shallow) groundwater storage systems, i.e., characterizing the ability of the system to store more (large) or less (small) water in the subsurface. High/low is rather confusing here as high/low is often used for high or low permeability of the aquifer, i.e., the degree of infiltration of water into the aquifer. Has a high groundwater system (also) a large storage? This should be clarified (or changed) and a distinct definition of the three systems is needed in a prominent way in the manuscript. Wording should be revised also to other sections in the manuscript, e.g., "companies with access to principal aquifers might depend more on groundwater compared to companies with access to shallow, less productive aquifers" (L83-84). '

We acknowledge the confusion caused by the naming of the three types of groundwater storage systems. We thank Reviewer 1 for highlighting this and we will rename these groundwater systems as suggested. The reasoning behind the high, medium, low naming of groundwater systems is indeed characterising the overall availability of groundwater storage given the modelled groundwater storage-outflow equations. We will include a detailed description in the suggested 'virtual catchment' section to define the large, medium and small groundwater storage systems.

Comment 2: 'Furthermore, I found the only linkage between hydrogeology and groundwater systems in L139-149. Is the linkage only for some kind of justification for different GW model boxes in HBV or was the aim really to quantify the impact of integrated drought policies on hydrological droughts in different hydrogeological settings? For me it is not clear why the hydrogeological features of the virtual catchment (karstic, porous and fractured) are linked 1:1 to high, medium and low groundwater systems? I guess this could be clarified, however, the hydrogeological context could be better integrated in the study. To be honest, my first concern reading the manuscript was on the added value of the hydrogeological settings (i.e., karstic, porous, fractured). See also point (4) below.'

We agree with Reviewer 1 that this should be highlighted earlier, i.e. in the revised introduction. The reason for representing different aquifer types is driven by 1) the aquifer-dependent delay in groundwater storage-outflow (L23-28), 2) the increased dependency on groundwater during droughts (L28-36) and 3) the absent groundwater component in recent drought policy modelling (L52-60). We will summarise these arguments in the last paragraph of the introduction to emphasise the need for different groundwater systems in this study.

Previous hydrological drought modelling by Van Lanen et al. (2013) applied a modified the standard HBV model to simulate hydrological droughts globally and changed the response time (in days) to represent different groundwater systems, finding that the responsiveness of groundwater systems has a large impact on drought characteristics. Stoelzle et al. (2015) extended the representation of different aquifer structures for a lumped model approach by identifying superior groundwater simulation structures. They tested and recommended a range of alternative model structures for five aquifer types (see our response to Comment 4 for details). Out of these five aquifer types, we modelled three aiming to show the impact of drought management strategies for different groundwater systems.

The three hydrogeological settings (karstic, porous and fractured aquifers) were selected to broadly representative for catchments in England and therefore excluded the ‘mixed’ and ‘combined’ aquifer types of Stoelzle et al. (2015). The first hydrogeological setting is generally associated with a large groundwater storage availability and non-linear drainage as found in karstic aquifers (Bloomfield and Marchant, 2013; Hartmann et al., 2014). Medium groundwater storage availability was found when modelling the porous aquifer type with slow drainage and possible leakage (Shepley et al., 2008; Allen et al., 1997). The last setting represented smaller groundwater storage availability with short response times for fractured or weathered aquifers (Allen et al., 1997). In modelling these three types of aquifers, none of the groundwater storage potential was constrained and different baseflow groundwater storage resulted from the different groundwater storage-outflow equations. In the revised manuscript, we will rephrase the introduction of these three groundwater systems and highlight the link to these different hydrogeological settings.

Comment 3: ‘I suggest to have a separated section “virtual catchment” where the modelling approach is explained (in single steps) using HBV (with a specific model structure) a set of average forcing data for England. Do I understand it correctly that no calibration was done in HBV as fixed parameter values were derived from literature etc. and there is no observed runoff? This should be mentioned more clearly! Consistent terms would be beneficial (at the moment idealized, simplified and virtual catchment is used). Perhaps this could also be done with an extension of Fig.1 showing (a) the HBV model structure (+ extension with three different GW boxes) and forcing data and (b) the sociohydrological model approach next to each other.’

We appreciate the suggested section with the heading ‘virtual catchment’, as this will indeed clarify the scope of the paper and associated modelling assumptions related to the three hydrogeological settings. As stated in L144-149, modelled groundwater storage-outflow parameters were based on the tested parameter range by Stoelzle et al. (2015) and mean aquifer characteristics in England (Allen et al., 1997). A wider range of parameters was tested in the sensitivity analysis, which results can indeed be better included in earlier sections, as suggested in Comment 5. We will rephrase this in the Results section and include findings of the sensitivity analysis were relevant.
As noted by Reviewer 1, there is no comparison to observed discharge or groundwater storage in this idealised hydrological system (L64-66). We used the term ‘idealised’ referring to the simplified hydrological system that can be seen as a stand-alone simplified example used for analysis, hence the term ‘virtual catchment’. We will review the use of these terms for consistency. Given the stand-alone example and modelling exercise, there is no validation of the performance of the different aquifer structures in this virtual modelling study that is merely building on from the validated model structure to assess hydrological droughts (Van Lanen et al., 2013).
Lastly, we have revised Fig. 1 to reflect both validated model structure, different groundwater modules and environmental and anthropogenic water demand. All three components have been included in the revised Figure 1 (see below).

Comment 4: ‘What is the advantage to have three different GW model representations for the three different groundwater systems if also variation of the parameterization of the same model structure could do this job? Stoelzle et al. (2015) showed that the FLEX GW structure outperforms the three structures POW, 1LBY and 2PA used in this study and the hydrogeological clustering of catchments is also better/clearer for FLEX than for other structures (Fig. 4 in Stoelzle et al., 2015). Wouldn’t it be easier to compare the effects of larger and smaller groundwater systems on different drought policies if the model structures across those different GW systems stay the same?’

It is a valid point raised by Reviewer 1 that when model validation is possible, all three groundwater systems can be well-simulated using the FLEX GW structure. However, in this virtual catchment modelling we aimed to represent different groundwater systems without the need for validation to find a representative share of fast and slow responding groundwater discharge. Therefore, the FLEX GW structure is not ideal given the dependency on the threshold-controlled storage outflow (h) to determine either primarily fast or slow groundwater discharge response. As an alternative, we applied suggested conceptual model structures to represent different kinds of groundwater storage-outflow characteristics that typically result in different drought characteristics for karstic, porous and fractured aquifer types. The non-linear groundwater discharge release, as observed in karstic aquifers, is best represented by a power law (POW) (Wittenberg, 2003). Slow porous flow could be represented by a

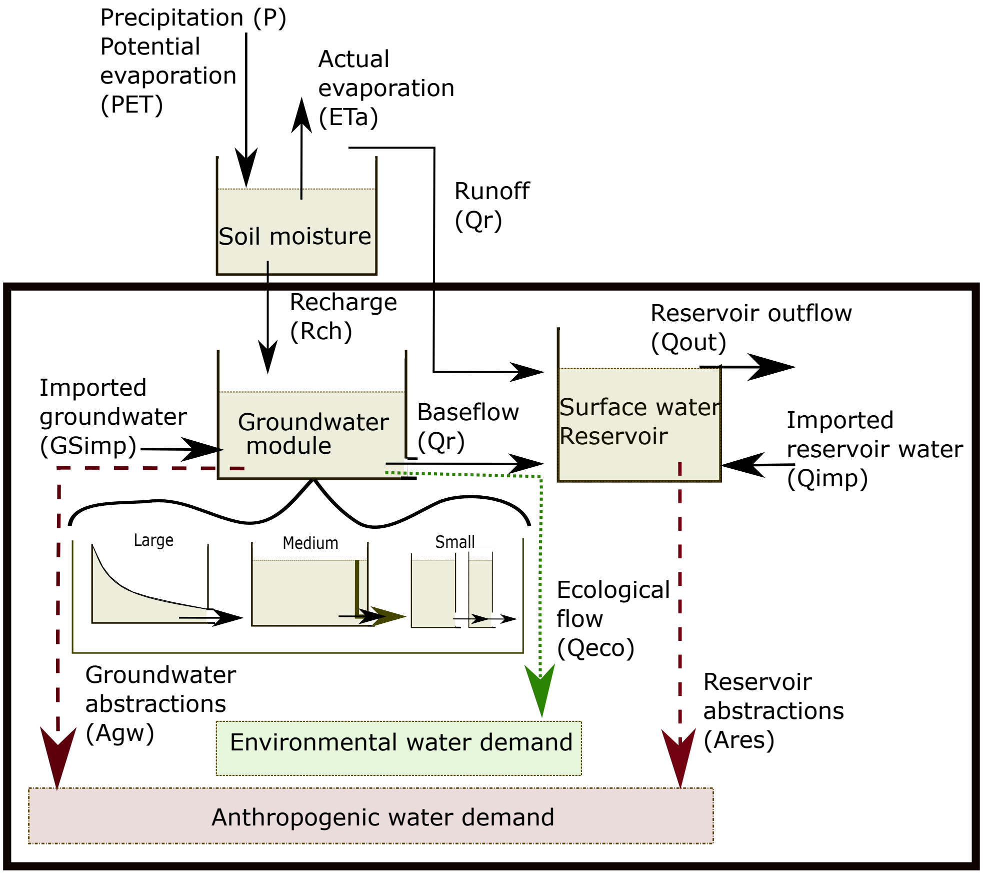

Figure 1. Socio-hydrological model consisting of a soil moisture balance driven by precipitation (P in mm d^{-1}) and potential evaporation (PET in mm d^{-1}), a surface water reservoir storing runoff (Qr mm d^{-1}), and a groundwater module that consists of three groundwater system options (large, medium, small groundwater availability) driven by groundwater recharge (Rch in mm d^{-1}). Anthropogenic water demand is met by reservoir abstractions (Ares in mm d^{-1}) and groundwater abstractions (Agw in mm d^{-1}), both in striped dark red arrows. Natural water demand is represented by ecological flow requirements (Qeco in mm d^{-1}; dotted green arrow) and abstracted as part of the baseflow (Qb in mm d^{-1}). Remaining baseflow is routed to the reservoir. Additional water is imported in the model when reservoir or groundwater storage is insufficient (Qimp and GSimp both in mm d^{-1}). Drought management scenarios apply to the surface water reservoir, groundwater module, and environmental and anthropogenic water demand (all model components in the thick black box).

number of structures, but given the overall high performance by 1LBY (Figure 5 in Stoelzle et al. (2015)) and the reported slow flow and possibly leakage in English Permo-Triassic sandstone aquifers, we selected the 1LBY. For fractured aquifers, the overall recommendation to apply parallel reservoirs and the overall high performance of 2PA (with groundwater recharge input) were the reasons for selecting 2PA for the last groundwater system.

By selecting these three conceptual model structures, we aimed to represent three different kinds of groundwater storage-outflow release whilst testing a range of representative parameters for English catchments. In the revised manuscript, we will include the reasoning behind the selection of model structures in a virtual catchment.

Comment 5: ‘I like the section 5.3 model limitations as this discussion is really important to understand the results of the study. Beside the fact that this section could be incorporated into other sections of the manuscript or at least should not be placed at the end of the discussion section (to gain a more positive ending of the paper), I asked myself

what would have been happened if another forcing than an England's average was used? Are average conditions a good starting point for such drought analysis like here? Additional analysis could shed light on this, however, at least more discussion is needed to evaluate how representative the average approach is for the different regions in England or different water companies.'

We thank Reviewer 1 for these two suggestions. We will include relevant results of the sensitivity analysis in earlier sections if relevant (also see response to comment 2). Regarding the second suggestion, our intention was to select a representative precipitation record for England. The HadUKP data consists of weighted observations, including extremely wet and dry periods within its record (Alexander and Jones, 2001). However, we will verify if other forcing English data, i.e. a representative location in the CEH-GEAR dataset (Tanguy et al., 2016) changes the overall water balance and outcomes. We will include these additional analyses in the full rebuttal and amend the results in revised manuscript if necessary.

References

Alexander, L. and Jones, P.: Updated precipitation series for the UK and discussion of recent extremes, Atmospheric science letters, 1, 142–150, 2001.

Allen, D., Bloomfield, J., and Robinson, V., eds.: The physical properties of major aquifers in England and Wales, British Geological Survey (WD/97/034), 1997.

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.

Hartmann, A., Goldscheider, N., Wagener, T., Lange, J., and Weiler, M.: Karst water resources in a changing world: Review of hydrological modeling approaches, Reviews of Geophysics, 52, 218–242, https://doi.org/https://doi.org/10.1002/2013RG000443, 2014.

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.

Stoelzle, M., Weiler, M., Stahl, K., Morhard, A., and Schuetz, T.: Is there a superior conceptual groundwater model structure for baseflow simulation?, Hydrological processes, 29, 1301–1313, 2015.

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.

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, Hydrology Earth System Sciences, 17, 1715–1732, https://doi.org/10.5194/hess-17-1715-2013, 2013.

Wittenberg, H.: Effects of season and man-made changes on baseflow and flow recession: case studies, Hydrological Processes, 17, 2113–2123, https://doi.org/https://doi.org/10.1002/hyp.1324, 2003.