Author comment to Reviewers comment #2

Helena Gerdener, Olga Engels and Jürgen Kusche

The author's answers are indicated in red color, as well as old text passages. New text passages are indicated in green color.

General comments

In this paper the authors developed a framework that potentially contributes to the understanding of how drought signals propagate through various GRACE drought indicators. By applying three methods (GRACE-based indicators), the authors assessed the skills of newly derived GRACE drought indicators under rather more controlled conditions. This work is significant, as the study is a considerable addition to the existing literature about drought identification methods. Also, the topic is within the scope of Hydrology and Earth System Sciences. Overall, the experimental design is clear, and for the most part, the authors' conclusion are supported by their findings. However, I outline several general concerns, followed by a range of specific comments, which prevent me from recommending this manuscript for publication in its current form. I do hope through that the authors will be able to adequately address my comment and when that is done, this paper should be acceptable for publication.

Response:
Thank you very much for your positive assessment and for your helpful feedback. We hope that we found good solutions to adequately address your comments and to improve the manuscript.

Comment 1.

The paper is relatively poorly written. There is a significant number of grammatical/syntactic errors that are present throughout the entire body of the manuscript. I specify several of these in the "Specific Comments" section below, but the authors need to thoroughly check the entire text, as similar or other mistakes may exist elsewhere.

Response:
We thank the reviewer for this comment. The comments in the "Specific Comment" section will be addressed (see below), and we will thoroughly double check the entire text for revision.

Comment 2.

Page 3 Line 14 "As can be expected, TWSC and 6 months SPI appear moderately similar (correlation 0.43), characterised by positive peaks at the beginning of 2013. This motivates us to modify common GRACE indicators…" I find the evidence not supportive enough to safely conclude that this link/association between TWSC and SPI is always (or everywhere) the case. The authors should test this on several different regions characterized by varying hydro-climatic conditions. Making such conclusive statements using only one example is scientifically inaccurate.

Response:

We agree with the reviewer that one example is not sufficient to warrant such a conclusive link association between TWSC and the SPI. In fact, we tested this link for other regions, and indeed we found considerable correlations between TWSC and SPI (e.g. Missouri river basin, South Africa, Maharashtra in West India). This was not illustrated (with figures) in the previous version due to space limitations, but we realize we should at least mention these results. Thus, a short sentence about some other regions including correlations is added.

Old text:

As can be expected, TWSC and 6 months SPI appear moderately similar (correlation 0.43), characterised by positive peaks e.g. at the beginning of 2004 and at the end of 2009, and negative peaks at the beginning of 2013. This motivates us to modify common GRACE indicators … .

New text

As can be expected, TWSC and 6 months SPI appear moderately similar (correlation 0.43), characterised by positive peaks e.g. at the beginning of 2004 and at the end of 2009, and negative peaks at the beginning of 2013. We also found correlations between TWSC and 6 months SPI in regions with different hydro-climatic conditions, among others, for the Missouri river basin (0.31), Maharashtra in West India (0.46), and South Africa (0.45). This motivates us to modify common GRACE indicators … .

Comment 3.

More information is required for the cluster identification. How exactly were the three clusters determined? The authors also need to clearly specify their exact geographic location.

Response:

We believe that a detailed description of the EM-clustering is given in the literature, so we would like to avoid explaining the EM-algorithm in the main part of the paper. However, we would like to follow the reviewer's suggestion to provide some information to interested readers so we add the main idea and equations of the EM-clustering to the appendix.

Thanks for pointing it out, the information about the polygons can indeed easily be missed out. We adjusted the text and changed the color of the polygons to make them better detectable. We also added the global distribution of all clusters to Fig. B1 in the appendix.

Old text1:

As a result of this procedure, we chose three clusters located in East Brazil (EB), South Africa (SA), and West India (WI), which were also affected by droughts in the past (e.g. Parthasarathy et al., 1987; Rouault and Richard, 2003; Coelho et al., 2016).

New text1:

As a result of this procedure, we identified three clusters located in East Brazil (EB), South Africa (SA), and West India (WI), which were indeed affected by droughts in the past (e.g. Parthasarathy et al., 1987; Rouault and Richard, 2003; Coelho et al., 2016). Location and shape of the three chosen clusters are shown in Fig. 3 and a global map of all clusters is provided in Fig. B1.

Old text2:
The EM algorithm by Chen (2018) is modified to identify regional clusters by maximizing the likelihood of the data (Alpaydin, 2009).

New text2:
The EM algorithm by Chen (2018) is modified to identify regional clusters. The EM-algorithm alternates expectation and a maximization steps to maximize the likelihood of the data (e.g. Dempster, 1977; Redner, 1984; Alpaydin, 2009). More details about EM-clustering are provided in App. B.

Appendix B: EM-Clustering
Expectation maximization (EM) represents a popular iterative algorithm that is widely used for clustering data. EM partitions data into cluster of different sizes and aims at finding the maximum likelihood of parameters of a predefined probability distribution (Dempster, 1997). In case of a Gaussian distribution the EM-algorithm maximizes the Gaussian mixture parameters, which are the Gaussian mean μ_k, covariance Σ_k and mixing coefficients π_k (Szeliski 2010). The algorithm then iteratively applies two consecutive steps to maximize the parameters: the expectation step (E-step) and the maximization step (M-step). Within the E-step we estimate the likelihood that a data point x_t is generated from the k-th Gaussian mixture by
E-step:

$$z_{ik} = \frac{1}{Z_i} \pi_k N(x|\mu_k, \Sigma_k) \quad ,$$

The M-step then re-estimates the parameters for each Gaussian mixture:
M-step:

$$\mu_k = \frac{1}{N_k} \sum_i z_{ik} x_i$$

$$\Sigma_k = \frac{1}{N_k} \sum_i z_{ik} (x_i - \mu_k)(x_i - \mu_k)^T$$

$$\pi_k = \frac{N_k}{N}$$

by using the number of points assigned to each cluster via

$$N_k = \sum_i z_{ik} \quad .$$

Using the maximized parameters EM assigns each data point to a cluster. The final global distributed clusters of the AR-parameters (Fig. 3) are shown in Fig. B1. These clusters were derived by modifying and applying an EM-algorithm provided by Chen (2018).

This appendix section contains a new reference, which is added to the reference list as follows:
Szeliski, R.: Computer Vision: Algorithms and Applications, Springer Science and Business Media, 2010

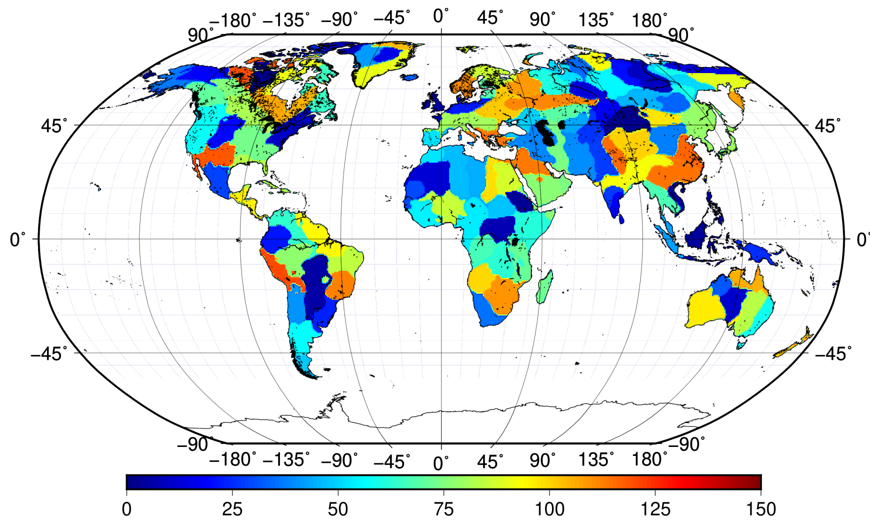

Figure B1. Clusters based on Expectation Maximization (EM) clustering applied to the global autoregressive model (AR)-model coefficients.

<u>Comment 4.</u>

The authors should provide more detailed information (characteristics) about specific droughts mentioned in their methodology section.

Response:
To elucidate the chosen drought events, we added a table containing the specific regions and the corresponding considered drought year and TWSC months.

Old text:
Searching for drought duration and magnitude (step 3) led to four droughts seen in GRACE-TWSC: The 2005 and 2010 droughts in the Amazon (e.g. Chen et al., 2009; Espinoza et al., 2011), the 2011 drought in Texas (e.g. Long et al., 2013), and the 2003 drought in Europe (e.g. Seitz et al., 2008).

New text:
Performing literature research for duration and magnitude (step 3) led to four droughts seen in GRACE-TWSC (Tab. 4): The 2005 and 2010 droughts in the Amazon (e.g. Chen et al., 2009; Espinoza et al., 2011), the 2011 drought in Texas (e.g. Long et al., 2013), and the 2003 drought in Europe (e.g. Seitz et al., 2008).

Table 4. Drought events in Europe, Amazon river basin and Texas with corresponding duration taken from literature.

Region	Year of drought	Considered TWSC months	Examples of literature
Europe	2003	June to August	Andersen et al. (2005) Rebetez et al. (2006) Seitz et al. (2008)
Amazon river basin	2005	May to September	Chen et al. (2009) Frappart et al. (2012)
	2010	June to September	Espinoza et al. (2011) Frappart et al. (2013) Humphrey et al. (2016)
Texas	2011	February to October	Humphrey et al. (2016) Long et al. (2013)

This table contains two new references, which is added to the reference list as follows:
Frappart, F., Papa, F., Santos da Silva, J., Ramillien, G., Prigent, C., Seyler, F. and Calmant, S.: Surface freshwater storage and dynamics in the Amazon basin during the 2005 exceptional drought, Environmental Research Letters, 7(4), 044010, doi:10.1088/1748-9326/7/4/044010, 2012.

Rebetez, M., Mayer, H., Dupont, O., Schindler, D., Gartner, K., Kropp, J. P. and Menzel, A.: Heat and drought 2003 in Europe: a climate synthesis, Annals of Forest Science, 63(6), 569–577, doi:10.1051/forest:2006043, 2006.

Specific Comments

Abstract
“Thus, this study aims at a better understanding of how drought signals, in the presence of trends and GRACE-specific spatial noise, propagate through GRACE drought indicators”: This phrase is perhaps the essence of the abstract; therefore it should be able to provide the necessary information on its own. The authors need to specify which trends they are referring to.

Response:
Thanks, we are referring to linear trends and constant accelerations in the paper, which are described with a linear term $a_1 (t - t_0)$ and a quadratic term $a_2 \frac{1}{2} (t - t_0)^2$ in Eq. 18. Linear trends and possible constant accelerations in GRACE TWSC can result from many different hydrological processes, for example, accelerations can results from linear trends in the fluxes precipitation, evapotranspiration and runoff. To specify the terms, we added linear trend and constant accelerations to the abstract.

New text:
Thus, this study aims at a better understanding of how drought signals propagate through GRACE drought indicators in the presence of linear trends, constant accelerations and GRACE-specific spatial noise.

According to this comment, we specified the meaning of trends and accelerations for the subsequent usage of the terms.

Page 7 Line 16
O: The signal is computed by … at time t with a constant a_0, linear trend a_1 and acceleration a_2 terms, an annual signal b_1 and b_2, and similar for a semi-annual signal c_1 and c_2.

N: The signal is computed by … at time t with a constant a_0, linear trend a_1 and constant acceleration a_2 terms, an annual signal b_1 and b_2, and similar for a semi-annual signal c_1 and c_2. Trends and possible accelerations in GRACE TWSC can result from many different hydrological processes. For example, accelerations can results from trends in the fluxes precipitation, evapotranspiration and runoff (e.g. Eicker et al. 2016). In the following, the linear trends are denoted as trends and constant accelerations are denoted as accelerations.

Line 10 application-dependent Yes, corrected, thanks.
Line 10 large differences Corrected.
Line 11 particularly Addressed.
Line 12 We show that trend and accelerations – what do the authors mean by "accelerations"?

Response:
We mean possible constant accelerations contained in the analysed time series that is described by the quadratic term $a_2 \frac{1}{2} (t - t_0)^2$ in Eq. 18. We hope this is more clear now by specifying the trends, as the reviewer recommended in the first comment of the "Specific Comments" section (above).

Page 1
Line 17 affect the Done, thanks.
Line 18 replace "reach" with "range" Done.
Line 24 led Yes, thanks, corrected.

Page 2
Line 4 depends on the accumulation period considered – unclear

Response:
Yes, we see that the term accumulation period leads to confusion here, because it is introduced at a later point. We remove this part of the sentence.

Old text:
For South Africa, due to a complex rainfall regime, areas and percentage of land surface affected by drought can vary strongly (Rouault and Richard, 2005) and their identification depends on the accumulation period considered.

New text:
For South Africa, due to a complex rainfall regime, areas and percentage of land surface affected by drought can vary strongly (Rouault and Richard, 2005).

Line 16 Much fewer Done.

Line 23 and the first data are expected

Response:
We updated this sentence, because the first data is now available and not “expected to become available in May 2019”.

Old text:
Meanwhile, GRACE has been continued with the GRACE-FO mission and the first data are expected to become available in May 2019.

New text:
Meanwhile, GRACE has been continued with the GRACE-FO mission and the first data are available.

Line 27 they found good agreement to net precipitation minus evaporation. - unclear

Response:
We agree this needs clarification. The agreement between TWSC and the combination of the net precipitation and evaporation is meant.

Old text:
For example, Seitz et al. (2008) investigated the 2003 heat wave over seven Central European basins using GRACE timeseries; they found good agreement to net precipitation minus evaporation.

New text:
For example, Seitz et al. (2008) investigated the 2003 heat wave over seven Central European basins using GRACE timeseries; they found a good agreement between TWSC and the combination of net precipitation and evaporation.

Line 34 without utilizing external information – please specify

Response:
Separating a specific compartment from GRACE TWSC data requires knowledge from other observation techniques or model outputs, because GRACE can only measure the sum of all compartments.

Old text:
However, neither GRACE nor GRACE-FO enable one to separate different compartments such as groundwater storage without utilizing external information, and their spatial (about 300 km for GRACE) and temporal (nominally one month) resolution are limited.

New text:
However, neither GRACE nor GRACE-FO enable one to separate different storage compartments such as groundwater storage without utilizing additional (e.g. compartment-specific) observations or model outputs, and their spatial (about 300 km for GRACE) and temporal (nominally one month) resolution are limited.

Page 3
Line 4 delete "e.g." Done, thanks.
Line 7 "smoothing" Done.
Line 17 What are "differencing periods"

Response:
We agree that the term here is confusing, because it was not introduced before. We change the sentence.

Old text:
This motivates us to modify common GRACE indicators to account for accumulation and differencing periods.

New text:
This motivates us to modify common GRACE indicators to account for accumulation periods of input data, which are used with e.g. 6 months SPI, but also periods that are based on differences of input data.

Line 21 spatially averaged Done, thanks.
Line 26 will complete the paper Done.

Page 4
Line 2 explore Thanks, corrected.
Line 10 more regularly Corrected.

Page 8
Line 10 we construct Done, thanks.
Line 13 including the introduced (in Sec. 2.3) signal … Done.
Line 26 … following A et al. (2013) … is there something missing here?

Response:
Indeed it might lead to confusion but A is the full last name.

Page 11
Line 8 drought onset and end Corrected.
Lines 10-14 these thresholds are rather arbitrarily made. It seems to me that a single value for the drought duration and magnitude should not be used for different hydrologic regimes.

Response:
We do not agree with the reviewer that these values for the threshold are arbitrary because we identified these values by analysing different historical droughts that were detected in literature using GRACE TWSC. Of course, one can not assume that one value for drought duration and magnitude can be detected in different hydrological regimes, but this is not what we intended with this analysis. We aim at simulating a signal that is similar to existing drought signals contained in GRACE, which is able to show up as exceptional drought in at least one indicator.

Page 12
Line 5 inappropriate use of English for a scientific paper Corrected.

Old text
However, seen these difficulties, we decided to stick to the most simple TWSC drought model, i.e. a constant water storage deficit within a given time span.

New text
However, due to these difficulties, we decided to use the most simple TWSC drought model, i.e. a constant water storage deficit within a given time span.

Page 13
Line 10 delete "would" Corrected, thanks.

Page 14
Line 17 for the 3, and 6 months differenced DSID Sorry we do not see a difference.

Page 20
Line 24 climatic phenomenon Yes, thanks, corrected.
Line 24 delete "related to climatic conditions" as it is redundant Corrected.

Page 21
Line 9 in the northeastern Thanks, we changed it to "Northeastern".

Due to this comment we also changed following text:

Old text:
Fig. 3 shows the estimated AR-model coefficients, which represent the temporal correlations, ranging from very low up to 0.3, e.g. over the Sahara or in South West Australia, to about 0.8, for example in Brazil or in South Eastern U.S. EM-clustering is then based on these coefficients.

New text:
Fig. 3 shows the estimated AR-model coefficients, which represent the temporal correlations, ranging from very low up to 0.3, e.g. over the Sahara or in South West Australia, to about 0.8, for example in Brazil or in Southeastern U.S. EM-clustering is then based on these coefficients.

Page 23
Line 22 particularly Done.
Line 25 the onset and end Done.