RC1 - Anonymous Referee #1 - May 8, 2022

Review of the manuscript egusphere-2022-22
"Influence of intensive agriculture and geological heterogeneity on the recharge of an arid aquifer system (Saq-Ram, Arabian Peninsula)"
by Seraphin P, Gonçalvès J, Hamelin B, Stieglitz T, Deschamps P
May 8, 2022

1 General comments

This manuscript describes an interesting analysis of GRACE data to estimate the components of the water budget for the Saq-Ram Aquifer System.
I appreciated the careful analysis of data in order to assess which are the best products of satellite data processing to perform a water budget analysis for a wide, moderately or poorly gauged, basin. I think this is the most important and innovative aspect of the work. The remarks on effects of intensive agriculture and geological heterogeneity are also interesting, but, in my opinion, they are not so deeply considered in the work as to make them the central focus of the title.
The scientific quality of the work is good, but I would suggest to provide further discussion of aspects related to scaling, which is somehow hidden in the work-flow. The manuscript is generally well written and well organized. Some minor technical comments are listed below. I think that this manuscript deserves publication on a public repository. With some corrections and/or integration, it could be submitted to a high-quality international scientific journal.

We agree with Reviewer #1 that data analysis and the choice of satellite products are important, especially given that this GRACE-derived approach is now widespread in applied hydrology. This detailed analysis for such a large-scale integrative method is especially pivotal when the objectives are, as proposed in this study, to characterize finer processes poorly discussed in the literature (i.e. differential recharge rates depending on geological heterogeneities, and relative disconnection between the recharge front and the water table decline at the outskirt of intensive agricultural areas).
Reviewer #1 judges that the two features ("effects of intensive agriculture and geological heterogeneity") that are put forward in the title *"are not so deeply considered in the work".* These two points are however original outcomes of this study that are also highlighted by Reviewer #2. As an example, we clearly show that intensive agriculture is by far the main output of the water budget of the Saq aquifer and that the associated irrigation return flow ends up to be an input as significant as the natural recharge. Such process was not accounted for by the previous GRACE-derived water budget studies (see for example Fallatah et al 2019).
So, we think that the title should let it show these original features specific to the studied site. However, to take into consideration this Reviewer #1 comment, we will slightly modify the conclusion of the article in the revised version to better put forward them. We agree that the tool/method itself could appear verbatim in the title along with the main results and processes studied. We proposed to modify the title such a way to take account of these two aspects as follow: "Influences of intensive agriculture and geological heterogeneity on the recharge estimate of an arid aquifer system (Saq-Ram, Arabian Peninsula) inferred from GRACE data". Maybe the editorial board would have advices regarding this matter?

2 Specific comments

1. Lines 87 to 97. The resolution which can be reached with the remote sensing approach should be discussed. Also the extension of the basin which can be properly studied with satellite data should be discussed. For instance, at line 96, it would be useful to explicitly state how large are the mentioned regional aquifers.

"Without the use of specific downscaling approaches, these selected aquifers are usually larger than 0.1×10^6 km² given the 3°×3° native spatial resolution of the GRACE data (filtered up to 1°×1° for some products; Landerer and Swenson, 2012; Wiese et al., 2016)." will be added Line 91, as well as the surfaces of the mentioned regional aquifers Line 96: "from 0.6×10^6 km² to 2×10^6 km²".

2. Line 214. The averaged water demand for agricultural purposes is given with three or four significant digits. I do not think this is physically acceptable, also taking into account the estimates of uncertainties which are explicitly given later for similar quantities.

The significant digits were indeed irrelevant due to a simplification oversight after reporting the exact figures. The modifications have been addressed.

3. Figure 3. These data deserves further comments. Any remarks about the increasing discrepancies among the products by JPL, CSR and GSFC for TWS, which is clearly apparent after 2012? Which regression is applied to TWS time series? The linear trends for SWS are not very appropriate, as these time series show, at least, a three-phase behavior: great variability from 2002 to 2006; small variations and values close to zero or poorly negative from 2007 to 2018; high variability and positive values after 2019.

We agree that these data should have been commented. The following paragraph will be added Line 274:

"While the time series of the three GRACE solutions (Figure 3a) clearly show a significant decrease of the TWS chiefly due to an increasing groundwater deficit (no permanent surface water bodies over the studied domain), the GLDAS products reveal a three-phase behavior of the SWS signal (Figure 3b) which can be related to variations of the climatic inputs of the respective models: a great variability from 2002 to 2006; small variations and values close to zero, or poorly negative, from 2007 to 2018; and high variability and positive values after 2019. The increasing discrepancies among the GRACE-JPL, -CSR and -GSFC products for TWS, which is clearly apparent after 2012, are mainly due to the diverse shape and size of the Mascons, and the various methods of eliminating signal leakage effects used by the three respective computing centers. Lenczuk et al. (2020) acknowledge that the GRACE-GSFC product differs the most from the others."

Regarding the regressions presented Figure 3, these were just to better illustrate the trends but, in fact, the GWS signal (i.e. TWS minus SWS) is computed on a monthly basis, and then a polynomial fit is applied on the resulting GWS time-series in order to assess the long-term trend. Thus, regressions which introduce unnecessary confusion have been removed from the Figure 3. The polynomial fit applied to the resulting GWS signal was mentioned in the following section (2.4 Methods), but have been more clearly commented Line 311: "Upon identification of a long-term trend (18 years in our case) of the *GWS* signal (using a 4^{th} order polynomial regression to filter out the seasonal signal), ΔGWS is calculated".

4. Sections 4.3 & 4.4. I appreciate this discussion. However, the two sections could be merged in a better way. At the beginning of section 4.3 it is stated that the transit time of water through the vadose zone is about 350 years; taking into account porosity and saturation, the recharge times should be even longer. This immediately rises the concern about the fact that the dynamics of the unsaturated zone and of the aquifer recharge seems to be much slower than the variability of the gravity field. Furthermore, the time series considered here are 20 years long, which is a time interval much shorter than 350 years. At line 541, it is stated that 12 years of GRACE data are required to obtain the long-term average recharge. This seems to be in contrast with the previous comments, doesn't it?

Indeed, the transit time of the infiltration through the vadose zone compared to the time period considered for the gravity data is a real concern. That is why we proposed these two discussions with a local application (section 4.3) and a general conceptual representation of the 'standard' GRACE application to hydrogeology (section 4.4) which cannot be merged since the considered scales and resulting interpretations are very contrasted.
As stated section 4.3: "the transit time of water across the vadose zone in this area (i.e. 70 m) is about 350 years". This is valid only for this sand dune area studied by Al-Sagaby and Moallim (2001) submitted to natural recharge only. However, this transit time would be much lower in agricultural areas since the water content and the total recharge (artificial + natural) would be much greater. Since we do not know what is the domain averaged water content of the vadose zone for the whole Saq-Ram aquifer, this transit time cannot be obtained at this large scale. Hence, the 350 years of transit time cannot be directly compared to the time series of gravity data (20 years).
On the other hand, section 4.4 answers your concerns by demonstrating that with large transit times of the water in the vadose zone (presumably in between 0 and 350 years) it is difficult to attribute an annual GRACE-derived recharge to a specific year since the GRACE-GLDAS approach on annual time-scale does not yield annual recharges, but annual infiltrations at the bottom of the GLDAS models (i.e. 2m depth). That is why we advise to use only a long-term trend of the Gravity signal (15 years) to obtain the infiltration rate characteristic of a steady-state (> 10 years) in the case of arid aquifers with thick vadose zone and mostly diffuse recharge process. This steady-state infiltration rate can be considered as a long-term recharge rate since it buffers the annual variations of the recharge. Indeed, even if the rain water takes hundreds of years to reach the water table surface, there is still older water reaching the it during the GRACE time frame (i.e. a sort of piston effect). One can assume that if the resulting natural recharge is valid now, it was more or less the same 350 years ago.

3 Technical comments

1. I recommend to follow as carefully as possible the guidelines by BIPM about the use of SI units and the format with which values of physical quantities are written. In particular, I recommend to use parentheses to avoid confusion. I list here some examples.
(a) A space should be left between the value and the percentage symbol, e.g., 1 %.
done

(b) Line 81. "3-4%" should be substituted, possibly with "approximately 3.5 %". Similarly, at line 107.
done

(c) Line 109. "1.8 ± 0.3%" should be substituted with "(1.8 ± 0.3) %". Similarly, at lines 115, 345, 388, 418, 419, 445, 454, 459, 562.
done

(d) Lines 187 & 188. "30 and 90 × 106m3yr−1" should be substituted, possibly with "30 × 106 m3 · yr−1 and 90m3 · yr−1". Similarly at line 220.
done

(e) Line 312. "210 ± 30 × 106m3yr−1" should be substituted with "(210 ± 30) × 106 m3 · yr−1". Similarly at lines 325, 330, 355, 387, 462, 463, 465, 488, 489, 541, 559 to 561.
done

2. Line 51. Although Bierkens and Wada (2019) is a review paper, it could be fair to add citations of some of the seminal papers on this subject.
We agree. The citation of the first global assessment of groundwater depletion (Wada et al 2010) have been added.
Wada, Y., van Beek, L. P. H., van Kempen, C. M., Reckman, J. W. T. M., Vasak, S., and Bierkens, M. F. P.: Global depletion of groundwater resources, Geophys. Res. Lett., 37, https://doi.org/10.1029/2010GL044571, 2010.

3. Line 54. Since the resolution of GRACE data and of the products used in this work are given in degrees, it would be useful to include the extension of the basin in latitude and longitude. This can be done here or in section 2.1.
"spanning 10° longitude and 8° latitude" have been added at the beginning of section 2.1.

4. Lines 151 & 152. A verb is missing in the sentence "The Climatic Research Unit... (over the period 1901-2019)", isn't it?
The verb "is" was not missing, but the sentence has been modified for clarity purpose:
"Extracted from the Climatic Research Unit database (CRU; Harris et al., 2020), the long term annual average rainfall (AAR) over the studied domain is 55 mm yr-1 (over the period 1901-2019), with a maximum of 80 mm in 1982, and a minimum of about 40 mm in 1978 and 2009 (Figure 2)."

5. Lines 212, 318; legend of figure 2. "et al." should be substituted with "and McCluskey".
These citations within the manuscript are correct, however an error was present in the references (due to my citation software). The reference has been corrected to:
"Alhassan, A. A., McCluskey, A., Alfaris, A., and Strzepek, K.: Scenario Based Regional Water Supply and Demand Model: Saudi Arabia as a Case Study, Int. J. Environ. Sci. Dev., 7, 46–51, https://doi.org/10.7763/IJESD.2016.V7.739, 2016."

6. Figure 2. Why only one time series is shown with error bars? How is the uncertainty estimated?
This is explained in the dedicated section (3.1 Groundwater pumping):
"Using previously published data (see Sect. 2.2), we can also estimate a continuous agricultural pumping time series for the Saudi part of Saq-Ram Aquifer System. First, we computed linear regressions between the Othman et al. (2018) time series and (i) BRGM and Abunayyan Trading Corp. (2008) data, (ii) Alhassan et al. (2016), and Chowdhury and Al-Zahrani (2013) data given per KSA region. Subsequently, as shown in Figure 2, the mean agricultural withdrawals and their uncertainties were obtained by averaging these two regressions."

The data previously published were not given with associated uncertainties since it originally comes from the Ministry of Agriculture.

7. Lines 231 & 232. Expression "its 2004.0-2009.9 average value" should be substituted, possibly with "its average value from January 2004 to September 2009"
That is the way the NASA-JPL literature defines gravity anomalies. But it is indeed confusing, it has been substituted by "average value from January 2004 to December 2009" Line 231-232, as well as Line 271.

8. Figure 3. For both plots, I could not find an explicit definition of the reference equivalent water thickness, which corresponds to a null value.
GRACE does not provide absolute gravity measurements, only measurements relative to an average value. Line 231-232 have been modified into: "gravity anomalies (i.e. gravity value of a given month minus the average value from January 2004 to December 2009)". We computed GLDAS SWS the same way to be consistent (as stated Line 270-271). Hence, a zero Equivalent Water Thickness means that the gravity value of a given month equals the average from January 2004 to December 2009. The term "anomalies" was also added to the caption of Figure 3.

9. Lines 264 & 265. "(GLDAS data are available at https://ldas.gsfc.nasa.gov/gldas)" should be moved before, possibly at the end of the first sentence of this paragraph.
Done

10. Line 290. Which polynomial fit is used? I mean, second, third, or higher degree? I do not see ΔGWS in Figure 3.
"(using a 4th order polynomial regression to filter out the seasonal signal)" has been added Line 209. Adding the GWS signal to Figure 3 would overlap the TWS signal and therefore require another graph which we don't think is essential.

11. Line 343. In which sense "needed"?
This is a common practice in agriculture: an excess of the dose of irrigation required by crops is necessary to avoid the concentration of pollutants (most of the time salts) which reduces soil fertility.