

General comments

The paper deals with a very far reaching topic (generation of climate and land use change projections; establishing a model chain for a regional 3-D transient case study) with a lot of related issues which leads to the paper losing its focus (no clear research question recognizable). Major points include:

- Methodological weakness since there is no calibration on nitrate measurements so that the goodness of the groundwater quality predictions can't be evaluated
- Uncertainty aspect due to climate change and land use transition models once mentioned (line 69) but not further discussed
- No clear paper message – instead in the short discussion and conclusions section some points are introduced that were only slightly or nowhere mentioned before (reevaluation of MOS management plan, stakeholder involvement, MAVT)

Specific comments

L 58: add river stages

L 60: use temporal and spatial variability instead of heterogeneity

L 67: add in spatial and temporal detail at end of sentence

L 74: add it before sometimes

L 73 – L 78: unclear as is, needs to be explained in more detail; e.g. it is unclear how conceptual models can be used to adequately assess climate and global change impacts – please provide an example or reference

L 142-143: What about the period between 1990 and 2010

L 150-152: What is your conclusion / explanation from the strongly diverging standard deviation of monthly averaged temperatures; in Figure 2 please refer to GCMs instead of using the number codes

L 152-156: How do you assess the discrepancies between historical monthly mean precipitation and the simulated values for the quality of predicted values?

L 169: use combining instead of coupling

L 173-174: include references for EU scenarios and LUC allocation techniques; it might be an option to move L 188-192 right to that bullet point

L 175+179: use long instead of large

L 180: latter instead of latest

L 195-220: Land use change projections parts very hard to follow; consists of a lot of technical terms and would benefit of a restructuring and simplification (maybe a flowchart might help) with concentration on main results used in further progress of the study

L 239; insert simulating after properly

L 242: delete second and

L243-245: what is the difference between the existing couplings between SAT and MODFLOW and your approach

L 275: If SWAT is a single plant growth model you simulated only the dominating crop in each of the 445 HRUs – is that correct?

L 300-303: Is there no Spanish soil database? What is the spatial resolution of the FAO's digital soil map within the study area? Is the parametrization useful given the local conditions?

L 305-306: I don't understand minimum land use unit area about 25 ha

L 307: do you mean validity

L 310: how many HRUs are excessive amount?

L 320-326: can be deleted and covered by references; please use just one MODFLOW reference instead of 3 (L338, L247-248)

L 334: reference missing after according to

L327-337: a cross section or a 3D image would be helpful to better understand the geological features of the MOS aquifer

L 347-348: it is unclear if pumping values used in MODFLOW are taken from SWAT outputs (assuming optimal irrigation) or from real farmer data like stated in l 318

L 350-359: can be deleted and covered by reference

L 360: model instead of mode

L 371-375: This sentence is obsolete since in the following one the most sensitive parameters are repeated. However, the explanation of the CN2 parameter in relation to climate and land use needs to be enhanced; it is unclear as it is

L 386: which table?

L 390-393: Can you also give us some information about the changing spatial and temporal patterns of recharge since this is important for pollutant input.

L 393: replace offer by in

L 399-421: Please combine and condense discussion about nitrate leaching. Avoid use of the term calibration since you compare SWAT results with results from another modeling study (btw, it should read table 2 in L 416)

L 411: what are the SWAT predefined functions, are they applicable to the MOS aquifer?

L 423-426: Can you give some numbers on the goodness of your calibration on heads? In Fig. 7 please use individual y-axis scales and enlarge the plots for readability. For some wells the dynamics between observed and computed time series is very different.

L 428-434: Please show in example for you procedure of matching observed nitrate groundwater concentrations. How did you assess the temporal development of your computed nitrate concentration time series? What transport parameters did you use? Again, I think the term calibration might be misleading.

L 435-441 delete since repeat from L 428-434

Please integrate in fig. 8 and 10 in the bars some internal shading or structures so that one can right away identify the displayed combinations of cc and lulc scenario; as is the discussion is difficult to follow

L 449-454: please try to provide a more detailed discussion of the precipitation-recharge nexus, e.g. to what extent do you find the difference in precipitation between the three cc scenarios and in the 3 different time periods also in the computed recharge patterns (in particular since you conclude that cc seems to be the main driver of change)

L 466: insert as after much

L 459-462: see comments on L 347-348: did you use optimal model irrigation or real irrigation values of the corresponding farmers? Return flow from irrigation might be lower in the latter case since farmers might consider weather forecasts.

L 472: did you experience the same crop yield differences for the various crops you simulated?

L 473: assigned and but instead of by

L 491: How does the different main driver for groundwater quantity relate to the driver identified for recharge? Is this due to pumping for irrigation?

L 493: no instead of to

L 494: here you mention uncertainty due to climate variability the only time as a side aspect; I believe you have to work more on this issue

L 503: Why does LUCS-4 necessarily lead to the lowest nitrate leaching values?

In L504-505 you mention that difference between luc scenarios become for nitrate leaching become significant in the medium and long term, however, in terms of groundwater nitrate concentration luc scenarios do not show a significant effect. Please elaborate how these findings fit together.

L 510-512: against the end of the 21st century recharge declines but this also means that there will be less nitrate leachate and thus lower nitrate groundwater concentrations. In the next but one sentence the authors state exactly that since the CNRM scenarios have the highest precipitation and thus recharge (I assume). If I'm correct please resolve this inconsistency

It should generally be made clear that the findings related to future groundwater quality development are more uncertain than that for groundwater quantity since there are only little data

available for validating the model chain with respect to nitrate leaching and groundwater nitrate concentrations.

The discussion in general does not very much relate to the specific results shown in the previous chapter

L 523-526: You did not calibrate or validate on irrigation and nitrate concentrations

L 528-531: If the overall goal of the paper is to improve insight into the systems vulnerability and potential adaptation options this needs to be better elaborated. Until the last page of the manuscript only L 124-128 cover this issue

L 532-535: As the authors state in chapter 4.1 the groundwater level decline is only predicted until 2020, then it stabilizes and starts to oscillate. This is not what you write in these lines. You also never really mentioned predicted streamflow depletion in the Jucar river.

L 536+543: threat instead of threaten

L 546-549: reads like a repeat to the previous sentences since here you state that the new management plan includes measures to avoid potential future threats

L 550-566: the conclusions appear to be rather general but include some very important issues which are not fully exploited. They mainly deal with the necessity for reassessing the MOS management plan and options and techniques to do so. However, this was not the focus of the paper (see also above comments)

L712-714: Why are yield results for SWAT and ITAP data identical? Did I overlook something? Since SWAT is a single crop model are the values per crop type averages of the HRUs where the particular crop was planted?

L 727-729: What is the real message of these diagrams?

L730-732: The figures must be increased, nothing can be read

Fig7: use individual scales and enlarge comparisons of time series

Fig8+10: enlarge and use same scales

Fig11: enlarge time series comparisons for readability. Did you use new codes for combinations of cc and luc scenarios? From what I can read only in one observation well groundwater nitrate concentrations exceed 40 mg/l – is this correct?