

Review of the paper entitled "*Integrated assessment of the impacts of climate and land use changes on groundwater quantity and quality in Mancha Oriental, Spain*" by Pulido-Velazquez et al., submitted to HESS

Dear reviewer,

Thanks for your constructive and helpful review. We have now included our responses (plain font) to the comments from the referee (bolded) and the proposed changes in the manuscript (in italics and under quotation marks). The new text is now included in the marked-up manuscript version with track changes.

Summary:

This paper presents a combined assessment of the impact of climate and land use change the groundwater resource of the Mancha Oriental aquifer in Spain. Three modeling tools are used for this simulation exercise: SWAT for the soil water balance, runoff and recharge, crop growing and nitrate leaching, MODFLOW for groundwater dynamics simulation and the MT3D for the Nitrate transport in the aquifer. Sequential coupling is applied, the outputs of the first two being used as inputs in the third. After calibrating the coupled models, 3 climate and four land use scenarios were applied by different future periods (short, middle and long term) and their results compared in terms of recharge, piezometric level, crop yield and nitrate concentration.

General comments:

The paper begins with a rather clear presentation of context and the modeling tools, even if some comments can be done (see specific comments below), for what is an interesting question of comparing the competing effects of climate change and land use evolution on both quantitative and qualitative aspects related to groundwater resources in Spain on a temporal scale.

Regrettably, when starting to present results, the quality of the description falls, likewise the figures legibility. Together with some editing errors, omissions, questionable methodological choices and some lack of logical organization of the argumentation, it becomes hard for the reader to follow the author's demonstration in the rest of the paper.

Following the specific comments provided below, I recommend a thorough revision of the results presentation section in order to be able to clearly see what is really resulting from this interesting coupling exercise. The comments aim at proposing some scientific discussion points to the authors that should enhance the relevance of their interesting work.

Apart from these specific comments it seems to me that two additional short discussions could improve the paper: 1) a more systematic analysis on the differences of impacts of CC and LUCS, 2) about the uncertainties related to all the modeling tools, the DD method, the climate model, the SRES A1B scenario, the different runoff/discharge simulation methods in SWAT.

Thanks for your comments. Unfortunately, this review was somehow done on a very preliminary version of the paper, which was modified and resubmitted to the journal for its review. In any case, although most comments refer to that previous version, we have done our best to accommodate them to the last version, although some issues (specially those regarding edition, format, organization and presentation), we think they were mostly corrected in that version of the manuscript.

Minor and major specific comments:

- 1. lines 48, 54, 61, 313: missing references.** The missing references have been added in the References section, corresponding to Jyrkama et al (2007), Klöve et al (2012), Holman et al (2012), and CHJ (2013)

2. **I. 594-618: references in the list not cited in the paper.** Non-cited references have been removed from the References section
3. **I. 605-607: journal name not provided.** The reference was not cited, so it has been removed from the References section
4. **I. 58-59: I agree with this sentence but it seems to definitive for me, because of the difficulty of estimating physical parameters that need to be calibrated anyway and also to the problem of the equifinality of over parameterized models (like SWAT for example...). Maybe moderate a bit this sentence to be more careful.** The sentence has been rewritten. Modified text: *“Numerical simulation models representing the spatial heterogeneity and temporal variability, if properly designed and calibrated, provide the most adequate way to estimate the impacts of climate and land use changes on groundwater systems”* (section 1, Introduction).
5. **I. 63-66: some reference could be cited for the sequential coupling concept.** Some references have been added in the References section, corresponding to Candela et al, 2009; Narula and Gosain, 2013. Modified text: *“The sequential coupling of numerical models can be used as an alternative to adequately assess climate and global change impacts. Those approaches generally employ hydrological models capable of representing the land phase of the hydrological cycle (infiltration, recharge, river flow, surface nitrate flow, nitrate leaching, etc), while groundwater quantity and quality impacts are assessed through groundwater flow and transport models (e.g., Candela et al, 2009; Narula and Gosain, 2013). Their linkage is made through an output-input scheme, in which some of the hydrological models’ outputs (groundwater recharge, nitrate leaching, etc.) are used as inputs in the groundwater models. Therefore, sequential coupling is able to keep the necessary detail in the key processes in the system while providing a comprehensive description of the way those processes are interconnected”* (section 1, Introduction).
6. **about the entire introduction: short presentations of the state of the art of CC and LU impacts on groundwater and in particular in the Mediterranean context, could be provided for a more precise description of the work presented in the paper with respect to the literature.** Some description of the state-of-the-art has been added to section 1, Introduction. Modified text: *“Recent climate and land use change studies in the Mediterranean region pointed at temperature and evapotranspiration increase, especially during summer (Chaouche et al, 2010; Candela et al, 2012; Molina-Navarro et al, 2014; Ertürk et al, 2014). Although the existence of a decreasing trend in the yearly precipitation has not been unanimously reported in all the studies, they agree in pointing at a change in the in-year precipitation pattern, with higher precipitation during fall and winter and lower rainfall during summer (Chaouche, 2010; Molina-Navarro et al, 2014). Those modified patterns, combined with land use changes, will decrease the amount of water recharge to the aquifers (Candela et al, 2009; Ertürk et al, 2014). This will provoke lower water tables and decreased groundwater discharge, which may consequently reduce stream baseflow and impact water supply and groundwater dependent ecosystems (Klove et al., 2014). These effects will be higher in heavily committed groundwater bodies, and will be intensified if water abstraction is increased to meet a growing demand for water. Predicting recharge and discharge conditions under future climatic and land use changes is essential for integrated water management and adaptation.”*
7. **about the case study presentation: a short description of the aquifer is needed, at least to let the reader know which type it is (alluvial, sedimentary, ???) and if it is confined or not, what is the pumping withdrawals total annual amount, etc.** A short description has been added. Modified text: *“The Mancha Oriental aquifer is made of three main layers over a Triassic (Keuper) impervious base. The bottom layer is a confined Jurassic limestone aquifer with 250-meter*

thickness, which emerges, getting unconfined, in certain areas in the south and west of the study zone. The mid layer is formed with Cretaceous limestone mainly confined but emerging in the north and east. The upper layer, which holds the majority of the pumping wells, is a Miocene limestone unconfined aquifer with some clay intercalations, up to 150 meters thick. A detailed geological description can be found in Sanz (2003) and Sanz et al (2009 and 2011)." (section 2.1, Case Study: the Mancha Oriental System in Spain).

8. Fig 1: it appears hard to distinguish between surface and groundwater water demand areas. Surface irrigated areas are small and located exclusively in the borders of the rivers, making it difficult to distinguish them from the river itself. In order to simplify the representation, the surface and groundwater areas have been merged in Fig 1.

9. L. 110-111: any reference can be cited? In any case, it is necessary to describe a bit the DD method, was it a statistical or dynamical one? The DD method applied was a dynamical one, using a regional climate model developed by the Swedish Meteorological and Hydrological Institute (SMHI). The paragraph has been added in the text and its references were included in the References section, corresponding to Kjellstrom et al., 2011; and Nikulin et al., 2011. Modified text: *"The climate change scenarios rely on the SRES A1B emission scenario for Europe. Climate data were obtained from three different GCM drivers: CNRM (National Centre of Meteorological Research), ECHAM5-r3 (European Centre for Medium-Term Weather Forecast), and HADCM3-Q0 (Hadley Centre). These scenarios have been downscaled using the SMHIRCA 3.0 RCM (Regional Climate Model) of the Swedish Meteorological and Hydrological Institute (SMHI). This study is one of the 16 case studies in the EU GENESIS Project, which deals with climate and land use impacts on groundwater and dependent ecosystems. For all case studies, the same dynamic downscaling method was applied by SMHI, which provided the meteorological forcing time series for the climate change scenarios (Kjellstrom et al., 2011; Nikulin et al., 2011). Daily time series of the relevant meteorological variables were provided for the 1961-2100 period, corresponding to a control period (1961-1990), and the short-term (2010-2040), medium-term (2040-2070), and long-term (2070-2100) scenarios."* (section 2.2, Climate change scenarios).

10.L. 121: it is the maximum temperature that is presented. Yes, it is. The minimum temperature was not included. Several minor corrections were made in the text. Modified text: *"Regarding monthly standard deviations of maximum temperature, all scenarios showed lower values, especially during summer, with no remarkable differences between them."* (section 2.2, Climate change scenarios).

11.L. 125-126: it is maybe a bit exaggerated to tell that as only this is mostly true in November. Some minor modifications were made in that paragraph to fix that exaggeration. Modified text: *"Fig. 2 shows some differences on the mean precipitation monthly pattern, especially in April and November. Some deviations are also found in the comparison of the standard deviations."* (section 2.2, Climate change scenarios).

12. Fig 2 is toll little to be really legible and legend must be completed with maximum temperature. The legend and axis titles and values' size were increased to be legible. The temperature showed was already the maximum, as indicated in the modified graph title.

13. Fig 3: not really legible and you should keep the same range for the y-axis bar of the first two graphs in order to compare the respective slopes of both temperature series. For the precipitation graph, the scale bar must be incorrectly labeled?. less than 50 mm of rainfall/year seems really arid. The requested changes were made in the graphs to make them legible. The temperature y-axis bar was set equal in both plots. The precipitation graph shows monthly values, so the unit is mm/month instead of mm/year, as indicated in the plot's title.

14.L. 144-147: you should be clear whether you are describing the SRES A1 model (and thus cite Nakicenovich) or a modeling performed within the GENESIS or other project. As it is written, it is not clear. The SRES A1B scenario was used. A short comment about it was made in the text. Modified text: *“Finally, CO2 concentrations were obtained derive from the carbon cycle models ISAM and BERN, both used in the IPCC 2007 4th Assessment Report climate projections, for the SRES A1B emission scenario (Nakicenovic and Swart, 2000).”* (section 2.2, Climate change scenarios).

15.L. 149-152: this sentence seems too long and not easy to understand what is explained. That sentence was replaced by another in order to clarify it. Modified text: *“Four land use change (LUC) scenarios have been considered in this case study.”* (section 2.3, Land use change projections).

16. For the LUCS presentation, maybe you could use some names to identify more easily the scenarios rather than number, such as: “baseline”; high irrigation”, “no irrigation”,... it could then be easier to the reader to understand to which scenario you refer in the results presentation and discussion. On another hand, even in a lot of information may be available in the cited references (l. 157), you cannot ask the reader to accept and understand what is considered inside your scenarios without explaining them in more details. You can at least provide some numbers about for example % of variation of irrigated areas or impervious surfaces, or any kind of measures about fertilization or adaptation (if any). The land use change scenario identifiers have been modified to be more understandable by “No Land Use Change (NLUC)”, “Short Term Scenario (ST)”, “Medium and Long-Term Scenario Increased Irrigation (MLTII)”; and “Medium and Long Term Scenario Decreased Irrigation (MLTDI)”. The increased and decreased percentages were added. No measures regarding fertilization or adaptation were considered, since they not regard to land use, but to land management (which was not considered when defining the scenarios).

Modified text:

- *“NLUC: no land use change or baseline scenario. This scenario has been included to be compared against the other ones, allowing the identification of synergic effects of climate and land use changes. It has been defined by all the three time periods (from 2010 to 2100), and its land use pattern corresponds to the current one.*
- *ST: short-term scenario based on a multi-temporal analysis of historical LUC changes and their key drivers, future EU scenarios and a combination of LUC allocation techniques. This scenario has been defined for the 2010-2040 period.*
- *MLTII: medium and long-term scenario with increased irrigation. It considers the LUC trend observed in the last 20 years, characterized by a change from non-irrigated to irrigated crops of about 10% of the agricultural area with respect to the ST scenario. We can assume that this is partially supported by subsidies coming from the EU Common Agricultural Policy (CAP). This scenario has been defined for 2040-2070 and 2070-2100.*
- *MLTDI: medium and long-term scenario with decreased irrigation. It considers the policies recently set up by the Farmers’ User Association of Mancha Oriental (JCRMO), and potential energy and water price increases, the latter coming from the application of the pricing policies required by the EU Water Framework Directive (WFD). These driving forces would cause a decrease in the irrigated area of about 20% with respect to the ST scenario, which would return to be operated without irrigation. This scenario has been defined for 2040-2070 and 2070-2100.”*

(section 2.3, Land use change projections)

17.L. 165: as this is a Spanish reference, maybe hard to find, you may provide additional information for non Spanish readers.

The additional information requested was added right after the scenario description. Modified text:

“For the short-term LUC scenario, a multi-temporal analysis of land use change (Oñate- Valdivieso and Bosque- Sendra, 2010) was carried out, based on the study of historical trends in crop patterns (derived from remote sensing images annually processed for the area; Calera et al., 1999, 2005 and 2012) and key driving forces for explaining LUC transitions. A spatially explicit model was used for that purpose, the Land Change Modeler Module (Eastman, 2006). This model provides a set of tools to perform historical analysis, trends and future scenario projections based on GIS techniques.

Data sets include series of remote sensing images, driving forces and scenarios from regional projects. LUC images from CORINE Land Cover Project were used as baseline (1990 and 2006; Fig. 4). Corine Land Cover (CLC) images have 100 m resolution and provide a LUC classification scheme widely used in all Europe (Feranec et al., 2010). Based on these images, historical analysis was performed to identify, evaluate and select the most significant LULC transitions. To develop LUC future scenarios all transitions need to be modeled according to the most likely driving forces. Driving forces were chosen from the available literature and statistical data with spatial representativeness (INE, 2011) from a wide range of biophysical, social, economic and political factors. A set of 20 driving forces were finally selected and spatially represented in raster format (Table 1), using a Cramer’s V test to select relevance of each driving force in every transition with a threshold value of 0.15 (Eastman, 2006).

Markov chains and European scenarios and projections (Eururalis 2.0 and Image 2.2) have been used to define trends in future land use in Europe. The EU project Eururalis 2.0 (Klijn et al, 2005) was chosen because its focus on rural areas. Eururalis developed Europe’s future land-use scenarios for 2010, 2020 and 2030 (Westhoek et al, 2006; Rienks, 2007); all available on-line to assist decision makers on the most likely scenarios of the Common Agricultural Policy (CAP). The predicted LUC scenarios were obtained from a combination of suitability maps and transition probability matrices extracted from all data sets described. A multicriteria Evaluation (MCE) method using Artificial Neural Networks (ANN) was applied to obtain suitability maps (Oñate-Valdivieso and Bosque Sendra, 2010) where change could happen for each transition selected. The Multilayer Perceptron (MLP) with three layers, backpropagation of error algorithm and the sigmoid transformation function were used to train every ANN. The transition probability matrices (TPM), based on the Markov’s chain theory, express the probability of one LUC to change to another in a determined period of time. TPM were obtained from historical data and from the Eururalis land-use scenarios. Applying the Markov process theory to the historical data (2000-2006), transition matrices for the 2000-2010 period were derived. For the 2000-2020 period, transition matrices were obtained from the tendencies showed in the Eururalis scenario. Finally, the cellular automata algorithm fed with the suitability maps of every transition and the transition probability matrices produced the final suitability maps and the land-use scenarios. A detailed explanation of the method can be found in Henriquez-Dole (2012).”

(section 2.3, Land use change projections)

18.L. 176: giving the number 4 for a non modification scenario is confusing, you could call it LUCS 0 if you really want to keep numbers or even call it LUCS “Refence” (LUCS REF) to be more specific. The name of the scenario has been changed and it has been placed the first in the description (as in response to comment 16).

- 19.L. 190-191: the reader must trust you as there is no information provided about the type of aquifer nor the piezometric data.** Some comments about the piezometric level evolution was placed in the case study description. Modified text: *“The aquifer has been subject to an intensive groundwater overexploitation since the 80’s, which has resulted in a continued drop of groundwater levels, especially in the southern area where irrigated crops concentrate. Groundwater levels have declined up to 60 m with respect to the levels in the 1980’s, with up to 3000 Mm³ of groundwater storage depletion from 1980 to 2005 (Sanz et al, 2011).”* (section 2.1, Case Study: the Mancha Oriental System in Spain)
- 20.L. 236-238: this correlation quality is not shown, it could be useful to provide some numbers in order to let the reader evaluate the uncertainty linked to it?** The correlation analysis was carried out in the process, although not included in the paper. A summary of this analysis was included. Modified text: *“In order to complete the data gaps, a spatial correlation analysis has been carried out, using the available data in all the weather stations. Correlation coefficients were obtained with respect to the Albacete – Los Llanos for estimating the unrecorded values as a function of the observed ones through regression analysis. Humidity correlation coefficients ranged between 0.82 and 0.97; wind speed, between 0.59 and 0.96; and solar radiation, between 0.90 and 0.99.”* (section 2.5, Watershed modelling).
- 21.L. 239-242: which one of the 3 types of soil mentioned above have not been found in the SWAT Database and how have its characteristics been attributed/estimated?** All of them were found in the SWAT database, although several properties were modified to adapt to the specificities of the case study area, according to the information provided by previous studies and field analyses. The sentence in which this checking was described was eliminated.
- 22.L. 245: a bit more information could be provided here (objective of the ERMOT project?) as the Henriquez-Dole will not be easy to find and read.** The ERMOT project was an estimation by remote sensing of the Mancha Oriental irrigated areas, and was used to validate the CORINE-provided irrigated areas. Other projects have been developed in the area with the same goal. In order to provide the reader with references easier to find, the text has been modified and the included references have been added to the References Section, corresponding to Calera et al, 2005; and Calera et al, 2012. Modified text: *“Land use data has been obtained from the CORINE Land Cover (CLC) project for year 2000. CLC data was obtained at 1:100,000 scale, with a 100x100 m resolution mesh and a minimum land use unit area about 25 Ha. The CLC-obtained data has been compared with land images (Calera et al, 2005 and 2012; Henriquez-Dole, 2012), in order to check its availability.”* (section 2.5, Watershed modelling).
- 23. L. 248: could the 12 type of land use types be specified? And why 7 soil type categories are mentioned here while only 3 are presented in I. 240?** The 12 land use types are the ones provided by the CORINE land cover (urban areas, water bodies, scrubland, forest, irrigated areas, etc), Fig 4 depicts them (some minor-presented areas were aggregated to other ones for increasing the figure readability, so 9 instead of 12 types are shown). There are 3 soil types presented but with different textures, so one soil type can lead to more than 1 soil categories if it is presented with different textures. An explanation has been included in the manuscript. Modified text: *“In the case study area Glegyc Cambisol (Bg), Chromic Luvisol (Lc), and Calcic Cambisol (Bk) soil types have been found with different textures, representing a total of 7 soil categories (1 for Bg, 1 for Lc and 5 for Bk).”* (section 2.5, Watershed modelling).
- 24.L. 252: this lead to a decrease of the LU categories from 12 to what?** It does not imply any reduction; it was an internal SWAT operation, which reduces the number of HRU’s, not the number of categories; an HRU (unique combination between soil type and land use) is filtered if

its extension is less than 10% of the total area of a sub-basin, being assigned to another HRU according to SWAT criteria.

25.L. 256-257: how this information has been integrated in SWAT and where it comes from? This information corresponds to the standard practices followed by the farmers, which are known by the team members and used in previous studies; it has been integrated in SWAT using the ArcSWAT graphic interface for ArcGIS

26. For the MODFLOW presentation, could it be provided additional information about the withdrawals points considered in the model? Maybe the number of cells and layers? The whole section has been redone to provide the requested and additional information.

Modified text: *“The MOS groundwater model consists of 7 hydrogeological units (HU), three of them are considered as aquifers (HU2, HU3 and HU7) and the other as aquitards (Sanz et al., 2011, 2009). The hydrogeological unit 7 is present throughout the MOS and is composed of limestone and fractured dolostone. The HU3 is only present in the northeast part of the study area and is composed of fractured limestone and dolostone. The upper aquifer, the HU2, which is located in the central part is composed of an alternate sequence of marl-lime and marl. Six hydrogeological domains can be identified in the MOS: Northern (ND), Central (CD), El Salobral-Los Llanos (SLD), Moro-Nevazos (MND), PozoCañada (PCD) and Montearagón-Carcelén (MCD) domains. According to Sanz (2011), there is hydraulic connection among the ND, CD and SLD domains, but not among MND, PCD and MCD or among these three and ND, CD and SLD. The Jucar River is the most important surface body and it is hydraulic connected to the aquifer, mainly to the HU2.*

The groundwater flow was simulated using MODFLOW (McDonald and Harbaugh, 1988). The model was discretized into 114 columns, 129 rows and 6 layers, with a cell size of 1 km². The main aquifers are represented by layers 2, 4 and 6 ((HU7/HU6/HU5, HU3 and HU2), while the other layers are semipermeable units (HU4, lower HU1 and upper HU1). The total number of pumping wells considered was 1776, most of them located in layers 2 and 6. To the west of the studied area, the Western Mancha System is assumed to contribute with groundwater discharge into the MOS. This was simulated by a constant-head boundary condition. The initial heads, hydraulic conductivity and storage coefficients were obtained from Sanz (2005), although the latter values were further modified during the calibration process. The hydraulic conductivity varies between 0.05-500 m/day and the storage coefficient, between 1e-4 and 1e-5. Recharge and pumping values were obtained from the SWAT model outputs, assuming that farmers pump the optimal amount of water required by their crops.” (section 2.6, Groundwater flow model).

27.L. 278-281: how do the parameters have been estimated? Through calibration? They were estimated through literature, a sentence has been added in order to clarify that. Modified text: *“Nitrate transport parameters for the Mancha Oriental aquifer were estimated using the values reported in the literature. Decay and sorption in the saturated zone were not considered.” (section 2.7, Groundwater nitrate transport model)*

28. About the calibration section: there is no presentation of the validation methods and results? Since the length of the period with available data for both SWAT and MODFLOW was short (1994-2006), and no formal validation was possible for MT3D, no specific comments about validation were included. However, the validation for SWAT was done for the 2005-2006 period with successful results (daily correlation coefficients equal to 0.80 in Los Frailes gauging station, and to 0.52 in Alcalá del Júcar; while daily Nash-Sutcliffe indexes were equal to 0.68 in Los Frailes and 0.33 in Alcalá del Júcar).

- 29.L. 291: is this coefficient controlling superficial aquifers drainage during low flows?** We have now included in the text that the alpha coefficient refers to groundwater discharge into surface waters (ie., baseflow). Modified text: *“the groundwater discharge coefficient (α) into surface waters”* (section 3.1, SWAT model).
- 30.L. 293: it is the flow that is lost and not the coefficient.** True, the text has been corrected according to that. Modified text: *“the coefficient of water finally lost due to evaporation”* (section 3.1, SWAT model).
- 31.L. 296: this sentence is important but a question arises here about the order of the priority of the computed processes as from as far as I can understand what is written: is the runoff simulated before the evaporation or not? If it is the case, could it be the main reason of the CN2 sensitivity?** Both processes are calculated separately in SWAT, so there is no interference between them in terms of sensitivity. When the runoff is computed, the initial abstractions are included specifically, separating both components.
- L. 302: why such a large calibration period and a short validation one?** A 5 to 1 relationship was kept for calibration and validation periods (10 for calibration, 2 for validation), given that the available period with a complete data set was short.
- 32.L. 306-308: these values correspond to calibration, but the same corresponding to the validation period could be provided?** Now provided in the text (section 3.1):
“The validation of SWAT for streamflow was done for the 2005-2006 period with acceptable results: daily correlation coefficients equal to 0.80 in Los Frailes gauging station, and to 0.52 in Alcala del Jucar; while daily Nash-Sutcliffe indexes were equal to 0.68 in Los Frailes and 0.33 in Alcala del Jucar”
- 33.Fig. 5: hardly legible, it’s been getting harder... the legend could be completed by mentioning that left graph is for the center of the basin and right graph for the basin outlet. When zooming in, some contrasted differences appear especially during low flows between 1997- 2001 for the basin outlet ; could any explanation be provided?** The plot was updated including the requested info. An explanation of the accuracy decrease between the left and the right is provided in the manuscript. Modified text: *“The lower performance shown by the calibration coefficients in Alcala station seems to be related to the higher drainage area, increasing the complexity for the calibration.”* (section 3.1, SWAT model).
- 34. L. 311-315: it may be Sanz 2011? Moreover, could any insight on the seasonal values be provided ?** Yes, it corresponds to 2011, it has been corrected. The seasonality of the recharge mostly follows the seasonality of the precipitation, although the seasonal variations in the river stage (depending on reservoir releases upstream) will also imply seasonal variations in the stream-aquifer interaction. This was all included in the SWAT simulations.
- 36. L. 318: any reference for the ITAP data?**
The reference for the ITPA data has been added in the References section, corresponding to López- Urrea et al (2003). It has been properly cited in section 3.1, SWAT model.
- 37. L. 324: there are some differences that must be at least mentioned;**
We have realized we had some mistakes in Table 2 (section 3.1, SWAT model), which have been corrected in the current version. We can now see no significant differences, and the values found with SWAT are within the interval reported in the literature.

CROP TYPE	IRRIGATION (mm)		YIELD (Tn/ha)		LEACHING (Kg NO ₃ /ha)	
	SWAT	ITAP	SWAT	ITAP	SWAT	GEPIC
Wheat	363	350 – 420	7.2	7 - 8	53	45 – 85
Onion	572	550 – 620	5.1	4.5 – 5.5	114	100 – 120
Corn	552	500 – 600	12.2	12 -13	100	90 – 115
Sugarbeet	827	800 – 900	12.8	12.5 – 13.5	74	65 – 90
Barley	282	250 – 350	8.6	8 - 9	34	30 – 45
Alfalfa	782	750 – 850	12.1	11.5 – 12.5	12	10 – 30

Table 2: crop irrigation, crop and nitrate leaching management calibration

38. Table 1: mention that the values are at the yearly time scale in the titles and the SWAT presented values are averaged over the calibration period in the legend, by the way, as I assume that all that processes may vary from one year to another during the simulated period, why not provide the variation range as for the ITAP and GEPIC data? For yield, the similarity between the SWAT and the ITAP column is strange...why ITAP values are not varying for this variable?

This table has been corrected (in the current version is Table 2, section 3.1, SWAT model). The range in ITAP represents variations over time and space, so it is difficult to link the limits of the interval to the values provided by the simulation of the model. This is why we compare the average simulated with the provided range.

39. L. 329: provide some explanation about the work done by IGME and what type of institution is as it is not necessarily known by non-Spanish readers.

The IGME is the Geological Survey of Spain. Its work and functions are the same as for the geological surveys of any other country. A short comment has been added to the text. Modified text: *“the Geological Survey of Spain (IGME).”* (section 3.1, SWAT model).

40. L. 330-333: already said in the previous section (l. 320), remove it here or there;

This text has been removed.

41. About the the § on table 1 comments, I suggest to reorganize it by commenting the values presented in the table immediately after presenting each column to avoid the repetitions and facilitate the reading;

We have reorganized the text that explains Table 2 (in previous version was table 1):

“The average net value obtained with SWAT for stream-aquifer interaction was also compared to those reported by the Geological Survey of Spain (IGME) for the area. The values previously reported by the IGME expressed a range from 40 to 60 Mm³/year (IGME/DGA, 2010), where SWAT reported an average of 45 Mm³/year.

Unlike other previous studies using SWAT (e.g. Narula and Gossain, 2013), SWAT calibration has

been also based on the crop simulation component. The simulated irrigation volumes and crop yields have been compared with the ones reported from crop surveys and experimental data in the zone. These data have been obtained from the deliverables of the Agronomic Technical Institute of the Province (Instituto Técnico Agronómico Provincial, ITAP) (López Urrea et al, 2003). As shown in Table 2, the SWAT model performance is consistent with the expected values.

Regarding nitrate leaching, there is no possible comparison with historical data, as there are no historical records. Given the usual absent of data, the calibration of nitrate loads is often based on observed nitrate concentrations in the existing piezometers and the crop growth component. Nitrate leaching calculations uses SWAT leaching equations, information on fertilizer types and application, soil characteristic (including initial NO_3 concentration), and other factors such as the percolation factor affecting NO_3 transport.” (section 3.1, SWAT model).

42. L. 334-337: as far as I understand, this has not been done here isn't it? You should state it clearly and explain why or better not mention it at all.

This text has been eliminated, since we have not calibrated nitrate concentrations in surface stations. Modified text:

“Given the usual absent of data, the calibration of nitrate loads is often based on observed nitrate concentrations in the existing piezometers and the crop growth component.” (section 3.1, SWAT model).

43. About fig. 6: hardly legible too! some piezometric series show seasonal fluctuations, could this be linked to pumping? As pumping point locations are not presented elsewhere it is hard to have an idea of their possible influence. It should be better to use separate labels (plain / dotted lines) for the simulated and observed curves to avoid similarities after B/W printing; could it be interesting to provide details about the mismatching observed for 690-698-677 (no observed data after one point...), 656-685 piezometric gauges...

We have modified figure 7, (section 3.2, MODFLOW model), (in the previous version that was figure 6) in order to improve the understanding of the figure. Yes, the seasonal fluctuations are mainly due to seasonal pumping (because the pumping is produced during irrigation season), although seasonal variations in recharge will also affect. For the sake of clarity of the figures, we did not include all the wells in the map, but Fig 1 shows the irrigated areas, therefore helping to visualize the main pumping areas of the region.

44. L. 354: isn't it possible to provide some results comparison anyway? It could help discuss about the uncertainty issue associated to this model

We have included figure 8 that visualizes the simulated groundwater nitrate concentrations (section 3.3, MT3DMS model)

45. About Table 2: the logical organization of this table is a bit hard to catch, it should be easier to understand to attribute numbers depending on the considered period (GC 01-06 for short term and GC15-24 for long term for example). In fact the logical organization of the demonstration starts to be hard to follow from this point...

We have now changed the way we refer to the scenarios, according to your suggestion. The new table (now Table 3) is (Section 4, Scenario results):

Table 3: Analysis scenarios

SCENARIO	CLIMATE CHANGE SCENARIO	LAND USE SCENARIO
GC01	ECHAM5	ST
GC02	CNRM	ST
GC03	HADCM3	ST
GC04	ECHAM5	MLTII
GC05	CNRM	MLTII
GC06	HADCM3	MLTII
GC07	ECHAM5	MLTDI
GC08	CNRM	MLTDI
GC09	HADCM3	MLTDI
GC10	ECHAM5	NLUC
GC11	CNRM	NLUC
GC12	HADCM3	NLUC

46. L. 361-363: this sentence could be revised to tell that the coupling make It necessary to assess the recharge

We have modified the sentence:

“With changing precipitation and temperature patterns and values, climate change is expected to have a significant impact on groundwater recharge”

(section 4.1, Impacts on groundwater recharge)

47. About Fig 7, 10 and 12, the adopted presentation is not easy to understand, as different kinds of information are mixed. I don’t understand why there is no result presented for the present situation? It could be useful to see the magnitude of the impacts of the different scenarios. Moreover, presenting only some of the LUCS scenarios by period is confusing, why not present all the scenarios and all periods on the same graph, for instance assembling them by climate model and scaling them temporally? The adopted presentation is really hard to understand and moreover it raises concern about the interest of using LUCS scenarios for different periods or, alternatively to consider three future periods instead of one (let’s say by 2050), as I assume there is not strong differences between LUCS2 or LUCS3 between the middle and the long term

In the previous version of the manuscript, figures 7, 10 and 12 were “Groundwater recharge results obtained with SWAT”, “Crop yield results” and “Nitrate leaching results obtained” respectively.

We have modified the following figure titles (sections 4.1, Impacts on groundwater recharge, and 4.3, Impacts on groundwater quality):

➤ *Fig 9: Groundwater recharge results obtained with SWAT for the short-term (up), medium-term (middle) and long-term (down)*

- *Fig 11: Nitrate leaching results obtained with SWAT for the short-term (up), medium-term (middle) and long-term (down)*

We have eliminated the figure regarding crop yield results.

Now all LUC scenarios are represented in the figure. The average recharge and leaching values for the present situation are now given for comparison.

- 48. L. 390-394: the argument here is hard to follow from the provided data and figures and should be further discussed as it seems that whatever the CC scenario, the future recharge values are higher than present ones? (considering values given in L. 311-315)**

We have eliminated that argument (section 3.1).

We have redone the current Fig. 9 (groundwater recharge) and compare the values with the average recharge for the present situation in the text as follows (section 4.1, Impacts on groundwater recharge):

“The results for all climate change scenarios agree in a reduction of the mean recharge over the 21st century, specially in the long-term (Fig. 9), and a reduction with respect to the historical recharge in the last decades, estimated as 310 Mm³/year. The values of the short-term mean recharge are similar for the CNRM and HADCM3 scenarios, despite the fact that the precipitation is higher in CNRM than in HADCM3 (Fig 3). In this case, the difference in precipitation is not fully transferred to recharge. However, in the ECHAM5 scenario a lower precipitation results in a lower groundwater recharge. A remarkable recharge reduction can be noticed in the long-term for the ECHAM5 and CNRM scenarios, associated to the steep decrease in precipitations (Fig 3); whereas the HADCM3 scenarios show no decrease in either average recharge or precipitation.”

- 49. L. 397-401: This is not possible to assess from the provided figures, maybe a better representation of the data could help it?**

We have modified Fig 9 (section 4.1, Impacts on groundwater recharge) *“Groundwater recharge results obtained with SWAT for the short-term, medium-term and long-term”*, in order to improve the understanding of the obtained results.

- 50. L. 412-413: what about pumping? I wonder whether this means that you could feed crops with no precipitation at all, without considering any groundwater depletion risk?**

We have now included the values of average pumping for each scenario and period in Fig. 10. It shows how pumping varies depending on the climate scenario (because of the combined effect of precipitation and temperature affecting actual ET) and the combined land use scenario (with no change, increase or decrease of land use).

- 51. L. 423: please clarify why only LU-2 and 3 are presented**

Regarding this question we have added (section 4.2, Impacts on groundwater quantity):

“Fig. 10 shows the evolution of the groundwater levels associated to the aquifer layers 2 and 6, which hold the largest amount of pumping wells, considering all the possible combinations of the climate and land use scenarios except from HADCM3 on climate and NLUC on land uses (with recharge between MLTII and MLTDI).”

52. L. 424: please clarify what the combination of scenarios intend to represent

We have changed the nomenclature in order to improve the understanding of the results (section 2.3 and table 3):

SCENARIO	CLIMATE	
	CHANG E SCENAR IO	LAND USE SCENAR IO
GC01	ECHAM5	ST
GC02	CNRM	ST
GC03	HADCM3	ST
GC04	ECHAM5	MLTII
GC05	CNRM	MLTII
GC06	HADCM3	MLTII
GC07	ECHAM5	MLTDI
GC08	CNRM	MLTDI
GC09	HADCM3	MLTDI
GC10	ECHAM5	NLUC
GC11	CNRM	NLUC
GC12	HADCM3	NLUC

(see response 16)

53. Fig. 11: this figure is strictly illegible; maybe plot the most interesting results in a greater size?

We have changed this figure by a new one, *Fig 10: Groundwater level evolution during the 21st century of layers 2 and 6 of the Mancha Oriental Aquifer*. This figure shows the groundwater level evolution in the most important layers of the aquifer (2 and 6), providing a better understanding of the whole aquifer status.

54. L. 427-435: because of the poor quality of the fig. 11, the reader is not able the follow the arguments of this paragraph...

We have changed this figure by *Fig 10: Groundwater level evolution during the 21st century of layers 2 and 6 of the Mancha Oriental Aquifer* (see previous response)

55. L. 441-442: “the higher precipitation that originate the higher groundwater recharge”, wouldn’t “the lower precipitation decrease that originate the lower groundwater recharge decrease” apply best?

We now better explain this in (section 4.1, Impacts on groundwater recharge) with this added text:

“The values of the short-term mean recharge are similar for the CNRM and HADCM3 scenarios, despite the fact that the precipitation is higher in CNRM than in HADCM3 (Fig 3). In this case, the difference in precipitation is not fully transferred to recharge. However, in the ECHAM5 scenario a lower precipitation results in a lower groundwater recharge. A remarkable recharge reduction can be noticed in the long-term for the ECHAM5 and CNRM scenarios, associated to the steep decrease in precipitations (Fig 3); whereas the HADCM3 scenarios show no decrease in either average recharge or precipitation. “

56. L. 445-446: It is difficult to follow the authors in their results presentation: I suggest comparing historical to LUCS 4 in order to describe the CC impact and then LUCS 1–CC for short term and then LUCS 2-3 for the middle and long term. It seems to me that the use of 3 periods with different LUCS is confusing. I wonder if using only two periods (short term and long term for instance) could be enough to support all the results presented and more easy to follow for the reader?

We have now compared the results in Fig. 11 (nitrate leaching) with the historical mean nitrate leaching in the text (section 4.3). Modified text:

“With regard to nitrate leaching, all the future scenarios show an increase with respect to the average value reported by SWAT for the calibration period, which was equal to 56 Kg NO₃/ha. This situation is consistent with recent studies in which an increase of nitrate leaching in groundwater was predicted for the 21st century (Stuart et al., 2011). Regarding Fig. 11, the scenarios show higher differences among scenarios when advancing through the 21st century. While similar leaching values were found for the short term (between 75 and 85 Kg NO₃/year), the results for the medium and long-term show a broader range, from 65 to 105 Kg NO₃/year. The scenarios associated with the MLTII land use scenario (the largest irrigated area) are generally the ones with higher nitrate leaching, while the MLTDI scenarios offer the lower bound of leaching values (Fig. 11).”

57. L. 454-456: Do scenarios consider any evolution in terms of fertilizer use? Otherwise is it just like: more crops = more NO₃...

We opted for autofertilization, although the optimal fertilization amount will depend on the climatic and crop conditions.

58. L. 464: please clarify which specificities they are?

This has been eliminated in the last version. Those specificities consisted in stream-aquifer interaction near the piezometer, so the nitrate concentrations were influenced by the stream-aquifer interactions (section 4.3).

59. L. 469: the decline is not unequivocal in fig 11...

We have changed this figure by Fig 10: Groundwater level evolution during the 21st century of layers 2 and 6 of the Mancha Oriental Aquifer.

60. L. 471: is the stabilization visible in the observed period? Could a time series be provided somewhere?

This stabilization was reported by the Jucar River Management Authority for about the last 10 years. However, the calibration period was 1993-2004, when this stabilization was not so evident (section 5).

61. L. 473: not sure to have seen that the infiltration rate simulated elsewhere in the paper?

This has been eliminated, as this process was not analyzed (section 5).

62. L. 480-482: maybe providing some synthetic details about the LUCS impacts on type and temporal perspective could be interesting?

We have now include an improved description of the effect of LUCs as compared to climate change (section 4.1, 4.2 and Discussion)

63. L. 491, this conclusion seem a bit too general, couldn't some additional elements be provided for the coupling exercise?

We have now reformulated these conclusions, with more insight on the interpretation of the climate and land use contributions to the recharge and groundwater quantity and quality results (section 5).

64. L. 492: adaptation measures? All the measures aim at meeting the WFD requirements as far as I understood?

Yes, although no adaptation measures were explicitly simulated in this work.

65. L. 512-513: but these measures are not focusing on adaptation to CC...

Yes, but although they were not specifically defined to do so, they can play a role in climate change adaptation.

66. L. 519: the mentioned results without GC impacts are not presented elsewhere in the paper?

Those adaptation measures were not simulated. Their impact assessment was carried out by the Jucar River Management Authority (CHJ), which presented them in the Jucar River Basin Management Plan.

67. L. 523-525, to be more specific and based on the LUCS, maybe the authors could provide some insights about the practical consequences of some of the considered measures?

No specific adaptation measures were simulated, since the focus was on a first assessment of potential impacts of climate and land use changes.

68. L. 527-529: see comment 65 and above, LUCS are not really planned for CC adaptation.

No, LUC responds to possible evolution trends of land use (section 2.3)

69. L. 594-617: all these references are not cited in the main text?

Non-cited references have been eliminated.