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

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.

Minor and major specific comments:

1. lines 48, 54, 61, 313: missing references ;
2. l. 594-618: references in the list not cited in the paper ;
3. l. 605-607: journal name not provided ;
4. l. 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 ;
5. l. 63-66: some reference could be cited for the sequential coupling concept;

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;
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..
8. Fig 1: it appears hard to distinguish between surface and groundwater water demand areas…
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?
10. L. 121: it is the maximum temperature that is presented;
11. L. 125-126: it is maybe a bit exaggerated to tell that as only this is mostly true in November;
12. Fig 2 is toll little to be really legible and legend must be completed with maximum temperature ;
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..
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;
15. L. 149-152: this sentence seems too long and not easy to understand what is explained;
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)…
17. L. 165: as this is a Spanish reference, maybe hard to find, you may provide additional information for non Spanish readers..
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…
19. L. 190-191: the reader must trust you as there is no information provided about the type of aquifer nor the piezometric data..
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?
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?
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;
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 l. 240?

24. L. 252: this lead to a decrease of the LU categories from 12 to what?
25. L. 256-257: how this information has been integrated in SWAT and where it comes from?
26. For the MODLFOW presentation, could it be provided additional information about the withdrawals points considered in the model? Maybe the number of cells and layers?
27. L. 278-281: how do the parameters have been estimated? Through calibration?
28. About the calibration section: there is no presentation of the validation methods and results?...
29. L. 291: is this coefficient controlling superficial aquifers drainage during low flows? Please detail as it is not clear enough;
30. L. 293: it is the flow that is lost and not the coefficient…
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?
32. L. 302: why such a large calibration period and a short validation one?
33. L. 306-308: these values correspond to calibration, but the same corresponding to the validation period could be provided?
34. 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?
35. L. 311-315: it may be Sanz 2011? Moreover, could any insight on the seasonal values be provided ?
36. L. 318: any reference for the ITAP data?
37. L. 324: there are some differences that must be at least mentioned;
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?
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.
40. L. 330-333: already said in the previous section (l. 320), remove it here or there;
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;
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..
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…

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;
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…
46. L. 361-363: this sentence could be revised to tell that the coupling make It necessary to assess the 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
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)
49. L. 397-401: This is not possible to assess from the provided figures, maybe a better representation of the data could help it?
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?
51. L. 423: please clarify why only LU-2 and 3 are presented
52. L. 424: please clarify what the combination of scenarios intend to represent
53. Fig. 11: this figure is strictly illegible; maybe plot the most interesting results in a greater size?
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…
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?
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 log term for instance) could be enough to support all the results presented and more easy to follow for the reader?
57. L. 454-456: Do scenarios consider any evolution in terms of fertilizer use? Otherwise is it just like: more crops = more NO3…
58. L. 464: please clarify which specificities they are?
59. L. 469: the decline is not unequivocal in fig 11…

60. L. 471: is the stabilization visible in the observed period? Could a time series be provided somewhere?
61. L. 473: not sure to have seen that the infiltration rate simulated elsewhere in the paper?
62. L. 480-482: maybe providing some synthetic details about the LUCS impacts on type and temporal perspective could be interesting?
63. L. 491, this conclusion seem a bit too general, couldn't some additional elements be provided for the coupling exercise?
64. L. 492: adaptation measures? All the measures aim at meeting the WFD requirements as far as I understood?
65. L. 512-513: but these measures are not focusing on adaptation to CC…
66. L. 519: the mentioned results without GC impacts are not presented elsewhere in the paper?
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?
68. L. 527-529: see comment 65 and above, LUCS are not really planned for CC adaptation..
69. L. 594-617: all these references are not cited in the main text?