**RC1**

Dear editor, dear authors,

I reviewed the paper in detail, and the tutorial cursorily. The case for the model introduced by the authors is made convincingly, and the model clearly falls within the scope of GMD. Although the model consists of existing components that have been augmented and linked, there is enough new material in the paper to justify publication in GMD. The relevance for practical applications is explained well by the authors.
Response: Thank you for providing a review of this manuscript, and for your comments.

The tutorial could use a few more references (including links) to established methods/techniques/software that the authors used. Examples of established software used are Notepad++ and QGIS, which are given without link or reference. Several links to other useful resources are given in the text of the tutorial, but not included in the reference list, which is a bit inconvenient.
Response: Thank you for pointing this out. We have included the following note on **Page 2** of the tutorial: "Note: This tutorial references QGIS, a geographic information system software that is free and open-source, version 3.42.1-Münster." We have edited the text on **Page 30**, in reference to text editors: "You can open up each file inside a text editor (e.g., Notepad++, a free source code editor for use with Microsoft Windows; EditPlus, a text editor for Windows) to view the contents."

The tutorial explains the working principle of PEST very well, which was quite helpful. But the Morris method used for the global sensitivity analysis used by PEST++ appears without explanation or reference.
Response: Thank you. We have now included the following explanation of the Morris method, with accompanying reference:
**Lines 414-418**:
*"In the Morris Method (Morris, 1991), the influence of a selected model parameter is assessed by quantifying the ratio of the change in the simulated response (e.g., streamflow, groundwater head) to the change in parameter value. This ratio is calculated for multiple base values of the parameter, with the collection of ratios then averaged to provide an overall sensitivity index for the parameter. This process is repeated for all selected model parameters, with the parameters then ranked according to their sensitivity index. Parameters with the highest index have the highest control on model output."*

All in all, I have no major comments. A few unclarities and inconsistencies appear in the text, highlighted in the minor comments directly inserted into the text. I suspect the inconsistencies are textual, rather than substantial, and can therefore be removed by a suitable rewriting of the text. I do not think the model code itself needs corrections. If this assessment proves to be correct, minor revisions should be sufficient.
Response: Thank you for these comments. We have addressed them as follows:

- QGIS reference: this is now included.
- Comma delimiter for thousands separator: I did not find this in the submission guidelines, so at this point we will leave the commas in the manuscript text and the tutorial. We can change at the technical editing stage, if requested.

- Line 100, handling salts: Yes, we have developed a salt module for SWAT+. However, it has not yet been published, so there is no reference for this.
- Lines 104-105 (network of channels): the network can be identified in either manner: the channel network can be "burned in" to the digital elevation model; or conversely, the channel network is identified using solely the topography from the digital elevation model.
- HRU: HRUs are defined in the preceding paragraph, as computational units with distinct land use, soil, and slope composition. For these two study regions, a portion of these are cultivated fields. To clarify this delineation, we have included the following text:
  **Lines 103-104**: *"JMR has 101 subbasins, 1,324 channels, and 10,611 HRUs, of which 5,101 are assigned to boundaries of individual, cultivated fields."*
- Line 112: the original SWAT+ model is not capable of handling this process (channel seepage to aquifer, if the water table is lower than channel stage). However, with the coupling of MODFLOW, this process can now be handled.
- Figure 1 caption (NHD+ channels): this is now defined in the Figure 1 caption (*"NHD+ channels = stream segments from the digital stream network of the United States (Moore and Dewald, 2016).")*
- HUC8: we apologize for not defining this. We have now defined it in the Table 1 caption.
- Line 144: the unconfined aquifer is divided into a set of small sections (grid cells), which are connected via groundwater lateral flow.
- Paragraph 170: correct, the MODFLOW model only included the unconfined (phreatic) aquifer. Based on previous studies in these two basins, there is no significant exchange with deeper, confined aquifers.
- Line 178: this refers to the vertical discretization (13 layers), whereas the discretization on line 144 refers to the horizontal discretization, using 500 m grid cells.
- Line 182: yes; we have now clarified this in the text.
- Line 182: yes, they are both MODFLOW models. We have now clarified this in the text.
- Line 206: thank you; we have corrected this.
- Line 221 (Figure 3): yes, we understand that this might be confusing. We have now changed the figure, so that the crops and tree (on the right) are positioned in the "birds-eye" view.
- Line 232: yes, this is a good point. We have changed this to "to".
- Figure 4: thank you; we have included a "volume transfer" designation in Figure 4B, to show the volume of groundwater that is transferred to the soil profile.
- Line 248: thank you for pointing this out; according to the submission guidelines, the correct style is to use only numbers: (1), (2), etc. We have corrected this for each of the 5 equations.
- Figure 5: thank you; we have now outlined the cell and labeled it clearly in Figure 5.
- Line 275: correct, that capillary fringe is not included in the simulation; therefore, the volume of groundwater transferred to the soil profile is only the depth of groundwater above the base of the soil profile.
- Lines 303-304: thank you for noticing this duplicate sentence. We have deleted it.
- Line 323: thank you. We have now defined soil water stress in the text: *"defined as the fraction of potential plant growth achieved due to soil water deficit"*
- Lines 324-325: we have clarified this in the text: *"Multiple irrigation sources can be specified for each demand object, with the order and demand fraction specified for each source."*

- Lines 326-327: these are arbitrary values, based on what is expected in these two regions. Typically, 1 inch (25.4 mm) or water is applied per irrigation event, and irrigation is applied when soil water stress occurs.
- Line 331: depth of water is specified by the user, and in our models we have set it to 25.4 mm (1 inch), as indicated on line 327.
- Line 339: yes, evaporation can happen. We do not account for this flux in the model, as seepage is often much higher in magnitude than evaporation.
- Line 360: thank you. We have indicated in the figure caption that these are cultivated fields.
- Line 365: SWAT+ has a routine that uses DRAINMOD, calculating a shallow water table and determining if it is above the elevation of the drain.
- Lines 367-368: correct.
- Line 399: The "command" subroutine calls the subroutines for HRU, routing unit, reservoir, point sources, and channel routing. These routes are then followed by a set of subroutines that write out results for each object type. This has now been added (with an arrow) in Figure 8.
- Line 403: we apologize for neglecting this reference. We have now included the White et al. (2020) reference.
- Lines 415-417: thank you; we agree that this sentence is long and confusing. We have now divided it into two sentences: *"Therefore, including MODFLOW for these two regions increases model run-time by 250% and 100%, respectively. These increases seem to be acceptable for model applications such as calibration, sensitivity analysis, and uncertainty analysis which often require hundreds or thousands of model runs."*
- Line 422: Line 416 indicates that run-times for calibration would be acceptable. We did not actually perform a calibration with PEST++.
- Line 424: we have not investigated the effect of including/neglecting evaporation from open ditches. This is a great point, and we will assess in future studies.
- Line 438: we agree that including north arrows and scale bars is helpful; we have included this in Figures 10 and 11.
- Line 470: D indicates seepage from canals, which are shown in Figure 1B as red lines; seepage occurs for cells that intersect these canal lines.
- Line 480: positive values indicate a decrease in head, which is a result of groundwater pumping. We have now clarified this in the figure caption.
- Line 485: we have now indicated this in the figure caption. Positive values indicate a decline in head, negative values indicate an increase in head.
- Line 510: thank you; we have changed this to "unsaturated flow".

Sincerely,
Gerrit de Rooij

**We thank Reviewer #1 for the helpful comments and suggestions.**

**RC2**

This paper presents a new coupling of two existing models, SWAT+ (watershed hydrology) and MODFLOW-NWT (groundwater flow). The coupled model is more hydrologically complete than either of the models on their own, and would be a useful tool for many regional hydrological studies. The paper falls within the scope of GMD, and I believe that it will be fit for publication after minor revisions have been completed. I think that this paper and the tool it describes are useful, but certain explanations need to be clarified and figures in particular need some work to aid the reader's understanding.
Response: Thank you for providing the review and for the positive response. We have addressed all comments and suggestions.

This paper provides an updated coupling after a 2020 paper (referenced Bailey et al., 2020a in the paper) already linked SWAT+ and MODFLOW. The motivation given for the new paper is that SWAT+ has been modified extensively over the past 5 years. One question that I have is whether the new linkage will be able to absorb future improvements to SWAT+, or will brand new couplings be required every few years as modifications continue?
Response: This is a great point. The new linkage will be able to include future improvements to SWAT+. I am in constant communication with the SWAT+ development team, as new routines are included in SWAT+, necessitating a new merge of the code. So, no new couplings (the basic coupling code is set), but we will need to include SWAT+ modifications as they come.

The paper links to a tutorial document, executable file, and sample input data to attempt reconstructions of the simulations presented. I appreciate this effort and the inclusion of the sample data. However, as a Linux user, I was unable to run the .exe file. Source code is provided; it would be extremely helpful to include compilation instructions.
Response: Thank you, this is an excellent point. We have now included a document that outlines the procedure for compiling SWAT+ in Intel Visual Studio. I am not familiar with Linux, and therefore do not have instructions on compiling using a different Fortran compiler.

**Line 90:** The paragraph starting at this line seems repetitive and unnecessary. Can it be integrated into the above paragraph?
Response: Thank you. We deleted the first sentence, and integrated the second sentence into the above paragraph.

**Figure 1:** This figure needs some work, e.g. both the Arkansas River and the San Joaquin River have the same symbology, so maybe it would work better to use the same symbol and then just specify on which panel does it refer to which river. What are the small black dots on panel B? Why does panel C have the additional aquifer information while panel A does not?
Response: Thank you. We have attempted to be consistent between the maps. The small black dots in B are irrigation pumping wells. We neglected to include this in Figure 1B, but have now corrected this. Panel C has additional information because of the existing MODFLOW model (CVHM2) which is providing many of the aquifer property values for the smaller model used in this study.

Line 163 states that for the JMR, the aquifer boundary is the same as the watershed boundary (Fig 1B), but Fig 1B does not actually show this – it would be clearer if explicitly shown or mentioned in the caption.
Response: Thank you; we have now included this in the caption to Figure 1.

Are aquifer boundaries shown in this figure and mentioned in the text from an external dataset?
Response: For the John Martin Reservoir watershed, the aquifer boundary is set to the same boundary as the watershed, with geologic units (see Figure 2B) obtained from a national dataset, as specified in Table 2 ("geologic units"). For the San Joaquin – Lower Chowchilla Watershed, the boundary is obtained from the CVHM2 model.

Table 2 includes aquifer thickness but it is not clear if this dataset also defines the boundaries.
Response: the datasets of aquifer thickness and geologic units do not provide the boundaries; the boundaries are explained now in the caption to Figure 1 (for the JMR model). The aquifer extent for the MSJ-LC model is provided in **lines 166-167**: *"For the MSJ-LC, only the 33,012 cells within the delineation of the Central Valley aquifer (see Fig. 1C for the aquifer boundary) are active."*

**Line 173** begins a description of population grid cell properties for the MSJ-LC model. Why are these needed for only this model, and not for JMR? Conversely, why is JMR MODFLOW data included in Table 2, but not MSJ-LC MODFLOW data?
Response: Properties for the MSJ-LC model are provided by the regional MODFLOW model CVHM2, a model previously developed by the USGS. Conversely, there is no existing MODFLOW model for the JMR, and therefore datasets are required to populate aquifer properties and geologic unit boundaries.

**Line 181:** What is meant by a 'geographic connection'? Was the data reprojected using GIS?
Response: The term "geographic connection" refers to the geographic intersection between the grid cells of the larger model (CVHM2) and the smaller model, using a shape file for both grids. Both grids used the same geographic projection. This has now been clarified in the text.

**Line 187:** How were the averages weighted? Based on layer thickness, or something else?
Response: Yes, by layer thickness. We have now clarified this in the text.

**Figure 2:** Why are we seeing geologic units for one aquifer, but hydraulic conductivity for the other? I'm wondering if the intention is to show two different types of setup/ simulation depending on your aims, but if so, that has not been made clear in the text.
Response: Yes, the intention is to show two different types of setup. Also, because the CVHM2 model has calibrated hydraulic conductivity values that we can map to the smaller grid, and the JMR model has uncalibrated values, and hence we show the geologic units that are used to define the boundaries of the hydraulic conductivity zones. This is now clarified in **lines 191-192**: *"Fig. 2D shows K values whereas Fig. 2B shows geologic unit boundaries, as the aquifer properties for the JMR model have not yet been calibrated."*

**Figure 3:** This figure is very blurry.

Response: This is likely because the high-resolution figure was not downloaded or included in the review version. The original figure is high-resolution and not blurry.

**Line 233-235 and Figure 4:** I am confused by the differentiation between options A and B. A makes sense since water will percolate downwards from the unsaturated soil profile; in B, why is exchange only possible from aquifer to soil profile? This implies that the water table can only rise in this case?
Response: Yes, in condition "B", groundwater is transferred to the soil profile, and for that time step percolation from the soil to the aquifer is not simulated. When the water table drops below the soil profile (e.g., on a subsequent day), then percolation from the soil to the aquifer can once again resume. We have modified Figure 4 to show the volume of groundwater transferred to the soil profile.

**Lines 252-261 and Figure 5:** On what scale are the HRUs constructed? Are they vectors? Is each agricultural field its own HRU? What else defines an HRU?
Response: HRUs are polygon. For these two models, each agricultural field has been designated as a unique HRU. There are other HRUs within the model boundary, based on intersections between soil type, topography, and land use.

Additionally, I thought that the cell size was 500 m (from line 162), but the quoted HRU sizes in this paragraph don't match with that.
Response: Yes, the size of the MODFLOW cells are 500 m. However, the HRU size depends on the agricultural field boundaries and the intersection polygons between soil type, topography, and land use.

**Figure 8:** This figure has helpful information, but I feel like I, as the reader, am having to do too much work. Having a reference for what all of these abbreviations mean closer to the image e.g. in the caption would help a lot.
Response: Yes, this is a great point. We have tried to provide meaningful information regarding the model input/output and code structure, but there are many abbreviations. We have included abbreviation references for MODFLOW and SWAT+ files and subroutines in the figure caption.

**Lines 403-406:** What is PEST++? Is it different from PEST? While references are provided, a very brief explanation would be helpful.
Response: PEST++ is a new version of PEST that allows for sensitivity analysis and uncertainty analysis. For this study, we have used PEST++ to include sensitivity analysis. We have included the correct reference (White et al., 2020), a short description of the Morris method, and a brief explanation of the extension from the original PEST.

**Figure 10:** What is recharge in all of the areas that are left white?
Response: These are areas of no recharge (0 flux). This is now clarified in the figure caption; also for irrigation pumping and canal seepage.

**Figure 12:** Is this the change from the model or from the well data? How do the two compare to one another?

Response: This is the change in the observed groundwater head. We have now clarified this in the figure caption. These changes should be compared to the average yearly change in head as simulated by the model, in Figure 11B. This is described in the text, **lines 466-468**.

**We thank Reviewer #2 for the helpful comments and suggestions.**