Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-120-AC1, 2018

Geoscientific Model Development Discussions

Interactive comment on "Beyond the bucket – Developing a global gradient-based groundwater model (G^{3}M v1.0) for a global hydrological model from scratch" *by* Robert Reinecke et al.

Robert Reinecke et al.

reinecke@em.uni-frankfurt.de

Received and published: 24 July 2018

We thank both reviewers for the thoughtful comments and questions. They helped us in particular to improve our explanation of the conceptual approach to gradient-based groundwater modeling that is necessary for global-scale groundwater modeling with a coarse spatial resolution, including the choice of simulating unconfined conditions and the conceptual difficulties in defining depth to groundwater table.

Referee #1

#1.1 A fundamental problem with these models is that they are difficult to verify, but this

is not at all reflected in the discussion of the results.

Reply: We have added a new paragraph to section 4 Discussion.

Changes to manuscript: We added the following paragraph to section 4 Discussion (second paragraph) "It is difficult to assess performance of the presented steady-state G^{3}M results. Model performance assessment is hindered by data availability and the coarse model resolution. (1) To our knowledge the data collection of depth to groundwater by Fan. et al (2013) is unique. However, they do not represent steady-state values. Apart from depth to groundwater observations, hardly any relevant data is available at the global scale. Especially exchange between surface water and groundwater is difficult to measure even at the local scale. Therefore, we compared G3M results with the results from other large-scale models. Comparison to the results of catchment-scale groundwater flow models is planned for transient runs that will be possible after the integration into WaterGAP. (2) Scale differences make the comparison to point observations of depth to groundwater difficult. Multiple local observations within a 5' cell may strongly vary, maybe just due to land surface elevation variations within the approximately 80 km2 large cells (compare Fig. S1 and S8). Often, observations are biased towards alluvial aquifers in valleys. The calculated hydraulic head of the grid cell may represent the average groundwater level per grid cell correctly but can be still far off the local observations of depth to groundwater. As the current model only presents an uncalibrated natural steady-state, a comparison to observations only provides a first indicator where the model and the performance measurements needs to be improved as we move to a fully transient model."

#1.2 The authors state on line 24 page 1 that these models are useful in areas with little or no data, as they allow to generate robust information. How can anything robust be generated (and how do we know its true?) in the absence of data. The hydrogeological literature is full of examples where even in the data -rich regions different models produce different outcomes.

Reply: In the revised version, we have deleted the statement about robust information, and explain now in more detail (on page 3 lines 24 ff) the purpose of our research effort, i.e. global-scale gradient-based groundwater flow modeling.

Changes to manuscript: Page 3, Line 24 ff now reads: "Our model development approach was to learn from existing large-scale regional models (Faunt, 2009; Vergnes et al., 2014, Maxwell et al., 2015; Dogrul et al., 2016) to gain insights into how the coarse spatial resolution, incomplete data, and conceptual model design affects model outcome. We want to find out whether we can use gradient-based groundwater modelling at the global scale, when later integrated into a global hydrological model, to improve estimation of flows between SW and GW (affecting both e.g. streamflow and groundwater recharge and thus water availability for humans and ecosystems) and capillary rise (affecting evapotranspiration)."

#1.3 We are presented with plots, numbers and graphs and some interpretation, but there is no credible discussion on the reliability of the result obtained. The only indication of model performance is that there is essentially no correlation between simulated and observed depth to groundwater. To me this means simply that the model cannot be used to make these types of predictions.

Reply: Comparison between depth to GW derived from simulated steady-state hydraulic head and point-scale observations of (non-steady state) depth to GW is not straightforward at all, such that clear conclusions about the model performance are difficult. The model performance assessment is hindered by two factors: data availability and scale. (1) To our knowledge the data collection of depth to groundwater by Fan. 2013 is unique. We try to extend that picture with large scale regional models as base for our comparison. We do acknowledge that comparison to a model is not the same as a comparison to observations. Apart from depth to groundwater observations hardly any relevant data is available. Especially exchange between surface water and groundwater is inherently hard to measure even at the local scale and thus often a calibration parameter in small scale models. (2) Scale differences make the comparison to depth

to groundwater observations difficult. (1) Multiple local observations within a 5 arcmin cell can vary by a large range (2) and they may have been observed at location very different from the average groundwater characteristic (e.g. average hydraulic head) within a grid cell - often biased towards alluvial aquifers in valleys. The calculated hydraulic head of the grid cell may represent the average groundwater level per grid cell but can be still far of the local observation. Furthermore, we observe that the depth to groundwater is highly influenced by the location of the surface water bodies (swb) and perception of depth to groundwater changes if calculated heads are compared to swb elevation and not to the average cell elevation.

Changes to manuscript: To respond to this comment, we added the following paragraph to section 4 Discussion (second paragraph) (refer to comment #1.1) To show the conceptual difficulty of calculating "simulated" depth to GW from the simulated 5' grid cell hydraulic head and an effective or mean land surface elevation at this scale, we revised section 3.1 and added, as Fig. 3b, a map showing the difference between P30 (the 30th percentile of the 30" land surface elevations, the assumed elevation of the surface water body water table) and the computed hydraulic head.

#1.4 It is not useful to plot observed and simulated hydraulic heads over such large scales, even if its just for the sake of model comparison. It is true that other authors have also presented simulated vs observed hydraulic heads over such large scales, but this is simply misleading. Depth to groundwater is the variable that counts for calculating exchanges with surface, amongst many other processes. In this sense none of the available models on a global scale is ready yet. This must not necessarily be a problem, as long as the results are not oversold, as is unfortunately rather often the case.

Reply: Hydraulic head is the main model output which is a good reason for showing it as such. And while the simulated heads might not match the observed very well in terms of absolute quantities, there are insights to be gained by looking at trends. Even local-scale models often do not match heads very well, but can be useful to understand the system response (i.e., how/where do aquifer heads change with other changes in the system stresses (pumping increases, recharge decreases, stream flows change, etc.). Depth to groundwater table is only derived from the model output using some estimate of a representative land surface elevation (see response to the above comment and the ensuing revisions of the manuscript). Calculated depth to groundwater highly depends on how a DEM is used to account for inter cell variability. On the other hand, this is also true for the derived head observations. Plots of simulated head vs. observed head are heavily influenced by the DEM signal and deviations due to difference in depth to water table are obfuscated due to the plot scales. Furthermore, the interaction of surface water bodies and the groundwater is driven by the gradients between heads. We do agree, however, that simulation of capillary rise requires a good estimate of local-scale depth to GW. Currently the model outcomes are not suitable to perform such a calculation. We already stated in the original manuscript (p.14 line 19-20) that there is almost no correlation between depth to GW observations and simulated values. So we are transparent about this and think that we do not "oversell" our results.

Changes to manuscript: none

#1.5 The formulation of the equation 2 is for a confined aquifer. The authors justify this conceptually wrong choice on line 20, page 6: "Flow equations are for confined aquifer because it reduces convergence time. "This is a very poor argument, purely based on convenience. To what extent the model should capture the relevant physics should cannot be a question on how difficult it is to solve equations. The goal of this modelling approach is to advance the interaction between the surface and the subsurface across very large scales. Given that the direct interaction with the surface always happens with unconfined aquifers the fundamental basis of the approach is flawed on the most basic level. While for steady state simulations the term falls out of the equation it is still very concerning that that a model is developed with inadequate flow equations.

Reply: The paper presents a conceptual model that differs in many aspects to tradi-

tional regional GW models due to the required coarse spatial resolution. Using the flow equation for unconfined conditions, which is typically done for the upper layer of groundwater models (unless confined by aquitards) is done to represent that in case of the same hydraulic gradient, less water can be transported if the hydraulic head and thus the saturated thickness drops. When looking at depth of GW in Fig. (old)3, one may think that in particular in mountainous terrain, the 100 m thick upper layer of the aquifer has fallen (almost) dry and does therefore in reality transfer no more groundwater. However, as shown in section 3.1, the high depth to GW is mainly related to the large land surface elevation differences within the 5' grid cell. , in almost all cells, the groundwater table is above the elevation of the water table in the surface water bodies (while land surface elevation per se is not part of the flow equation). Thus, given the high uncertainty of assumed hydraulic conductivity values and unknown actual aquifer depth, the assumption of fixed transmissivities seems to be appropriate for our global 5' model. Using the equation of unconfined conditions cannot be expected to improve the simulations significantly. Conceptually, at the applied coarse spatial resolution of the GW model, model layers should not be considered to be fixed to a land surface elevation. The model layers can be rather thought to be vertically (somewhat) aligned with the elevation of the surface water body table, and the flow equation rather governs the lateral and vertical fluxes over a thickness of 200 m.

Changes to manuscript: To clarify the difficult but important aspect of the relation between model layers and surface elevation in steep terrain at the spatial resolution of 5', we revised Figure 1 and added to section 2.1: "In addition, due to the coarse spatial scale and the possible large variations of land surface elevations within each grid cell, the upper model layers should not be considered to be aligned with an average land surface elevation. The model layers can be rather thought to be vertically aligned with the elevation of the surface water body table, as this prescribed elevation is, together with the sea level, the only elevation included in the groundwater flow equation (Eq. 2)." We added to the second paragraph of section 2.3: "We choose to simulate confined flow conditions in both layers even though the upper layer can be expected to decrease in depth and thus in transmissivity (hydraulic conductivity times saturated depth). Every unconfined aquifer can have an equivalent confined representation assuming a correct saturated thickness (Sheets et al., 2015). However, given the large uncertainties regarding hydraulic conductivities (possibly an order of magnitude) and the lack of knowledge about aquifer thickness, it is appropriate to choose the computationally more efficient assumption of confined conditions."

#1.6 I did not understand why the authors develop a new model in the first place. They rightfully acknowledge that models such as MODFLOW exist, and these model could potentially do the job. Their argument is that MODFLOW models typically integrate geological data that is not available on a global scale. Therefore, a simplified model is developed. But this is a strange way of reasoning, as with MODFLOW one is not obliged to integrate all the geological complexity. It would have been perfectly possible to use MODFLOW for this project, with several significant advantages: for example, an unconfined aquifer (see below) could have been simulated. In this sense the novelty of the aspects concerning model development is questionable.

Reply: The main reason for not using MODFLOW directly but just implementing the MODFLOW approach is an efficient coupling to the existing global hydrological model WGHM. The structure of MODFLOW does not allow an efficient in memory coupling that also account for the two different scales without too much computational overhead. The new model allows a more flexible extension of new components and adaptions to the conceptual nature of the model like an alternative capillary rise or dynamic recalculation of surface waterbody conductance. Additionally, the model framework is indeed capable of simulating unconfined conditions. Nevertheless, we believe the decision against it is a reasonable assumption - see response to #1.5.

Changes to manuscript: Page 8, Line 25-27 "The main motivation to develop a new model framework is the efficient in-memory coupling to the GHM and more flexible adaptation to the specific requirements of global-scale modelling."

#1.7 There are many other problems working on a global scale which are not even mentioned here but will even further undermine the credibility of the model. The three most important ones are: (1) Elevation is the wrong parameter for such a model. The data that should be used is not an ellipsoid- DEM but rather a geoid as the geoundulation is significant. (2) The density of sea-water is different, therefore there should be a density correction. (3) Steady-state conditions are inappropriate assumption that is not justified sufficiently well.

Reply: (1) As far as I understand the SRTM-based DEM it is based on a reference ellipsoid (WGS84) and a reference geoid that should already account for geoundulation. We assume that on a 5 arcmin resolution differences in the gravitational field are negligible. Furthermore, other inputs present a much higher uncertainty. (2) As the model is not intended to be used for studying specifically groundwater-ocean interactions, and given the cell size of 9 km, we assume that the difference in density can be neglected at this scale as other parameterization introduce a higher level of uncertainty. (3) Presenting a steady-state model is one of the first steps to understand model behaviour before moving towards a fully transient and fully coupled model. This represents a well-established method in developing groundwater models - regional as well as large-scale models. A steady-state model (1) limits the degrees of freedom and thus model complexity as no time-variation needs to be taken into account and no storage changes need to be tracked. (2) A steady-state clearly uncovers dominant processes and trends that otherwise might have been obfuscated in a transient model due to the slow changing nature of groundwater. It is evident that not all processes can be observed in a steady-state and model behaviour will change as we move towards a fully transient model. It represents a first step in the model development process. Furthermore, (3) generated steady-state hydraulic heads can be used as initial state for a transient model spin-up phase in a fully coupled model. It is true however that surface water bodies do not have a steady-state and that aquifers are ever changing. This is why the presented steady-state represents a first step into the model development as we move towards fully transient and coupled model. We think that it is not meaningful to move to a transient model directly with a completely new model without looking at the steady-state behaviour first.

Changes to manuscript: We added to the last paragraph of the introduction: "Steady-state simulations are a well-established first step in groundwater model development to understand the basic model behavior limiting model complexity and degrees of freedom, thus providing insights into dominant processes and uncovering possible model-inherent characteristics impossible to observe in a fully coupled transient model. A transient model might obfuscate model inherent trends due to the slow changing nature of groundwater processes. A fully coupled model furthermore adds complexity and uncertainty to the model outcome. In addition, the steady-state solution can be used as initial condition for future fully coupled transient runs. " (2) Page 5, Line 16

#1.8 Validation is done with other macro-scale models. This is a not an ideal strategy, as these large-scale models suffer from similar deficiencies (even though on less fundamental level). For a solid assessment of model performance, a detailed catchment scale hydrogeological model should be used for a benchmark comparison.

Reply: Validation has been achieved by a comparison to global groundwater observations, assumed naturalized conditions in a well-studied area (Central Valley) and by an additional comparison to other large-scale models (Maxwell et al./Fan et al.). Goal of the model development was not the replication of regional groundwater characteristics - at this scale this is not a reasonable goal. Comparison is furthermore likely to be very challenging or impossible as a catchment might span only a couple of cells of the global model. The comparison to other large-scale models however enable a comparison based on similar input data (and input data deficiencies) uncovering how model decisions at this scale affect model outcome.

Changes to manuscript: See changes in response to comment #1.3 Page 17, Line 29 ff. "The presented comparison to other large-scale models is based on the assumption that same model deficiencies e.g. in available data and scale issues can uncover

differences in model decision. A comparison to catchment scale models is challenging as scales can differ by multiple magnitudes. As the model is further developed towards a transient model the presented comparison to simulations in data-rich regions need to be extended and temporal changes in interactions with surface water investigated."

#1.9 On line 28,page 7 the authors highlight that this is ok –" . . .without losing important model behavior. " Transient and steady state is significantly different in both spatial and temporal dynamics.

Reply: Reviewer refers to line 28 on page 3. We agree with the reviewer

Changes to manuscript: We revised the sentence. See changes in response to comment #1.2 and #1.7.

#1.10 The description of the conductance is confusing. In MODFLOW L is not the length of the river itself, but the length of the river within a grid cell. But this might just be an imprecise formulation.

Reply: This is correct (See table 1). Manuscript has been changed accordingly.

Changes to manuscript: Page 6, Line 5

#1.11 Other aspects also require more justification and discussion. Why only 8 % of wetland surfaces? Where does this number come from? What are the numerical convergence criteria, as well as a wide range of additional model parameters?

Reply: Manuscript is describing 80% of wetland area. Available maps of wetland areas show the maximum spatial extent of surface water bodies. As the maximum extent is seldom reached we reduce the extent for the steady-state model to 80% of the area shown in maps. In the fully transient model the wetland area will be adjusted in each time step as a function of wetland water storage.

It is not clear to us what the referee meant by "as well as a wide range of additional parameters". Parameters including convergence criteria are shown in Table 1.

Changes to manuscript: Page 8, Line 12

———————————————————————————————-

Referee #2

#2.1 Is 5' an appropriate resolution at which to simulate groundwater flow? The analysis by Krakauer et al may be useful in determining the appropriate resolution.

Reply: Kraukauer et al. (2014) suggests that a grid spacing smaller than 0.1° (6') for lateral groundwater processes is favourable for models running at a finer resolution than 1°. Thus a 5' seems to be reasonable even though our results suggest that the scale properties of surface water elevation need to be investigated further and that information from subgrid scales might need to be accounted for to improve overall results.

Changes to manuscript: Page 4, Line 12,13

#2.2 The work is coupled to WaterGap at 0.5deg, this is a really large scale discrepancy. How do you think this might alter the model results?

Reply: As groundwater recharge is mainly driven by climate inputs that are only available at coarse scales the presented steady-state model is not affected by the scale differences. Moving towards a fully coupled model scale differences between the two models play an important role especially for surface water body coupling. For example, it is not reasonable to calculate a river head change in the 0.5° model and apply that change equally to all 5' grid cells to recalculate the interaction between the surface water and the groundwater. The (future) presentation of a fully transient coupled model needs to discuss this more extensively.

Changes to manuscript: none

#2.3 The comparisons between this study and Fan et al and Maxwell et al are interesting. While pressure head is important, I think the bias from these scatterplots, basically

water table depth, is more meaningful (as plotted in Fan et al / Maxwell et al too). The statistics will really be driven by topography which can occlude model performance and differences.

Reply: Please refer to our responses and changes to manuscript in response to comments #1.3 and #1.4.

#2.4 The diagram for how the model handles topographic breaks (Fig 1) is super confusing. Basically is water moved between cells even if there is a disconnect?

Reply: Yes, this is due to the coarse lateral discretization where in a 5' grid cell with approx. 80 km^2 area, the elevation differences can be larger than 200 m (as described in the text). Lateral interaction between neighbouring cells is always calculated in the model even if large topographic breaks are present. In order to avoid confusion, we modified Fig. 1 and text in section 2.1 to clarify that the top layer in the model should not be thought of as being located right at the land surface elevation.

Changes to manuscript: To clarify the difficult but important aspect of the relation between model layers and surface elevation in steep terrain at the spatial resolution of 5', we revised Figure 1 and added to section 2.1: "In addition, due to the coarse spatial scale and the possible large variations of land surface elevations within each grid cell, the upper model layers should not be considered to aligned with an average land surface elevation. The model layers can be rather thought to be vertically aligned with the elevation of the surface water body table, as this prescribed elevation is, together with the sea level, the only elevation included in the groundwater flow equation (Eq. 2)."

#2.5 The assumption of confined conditions really seems hard to justify. This is effectively what de Graaf et al (2015, 2017) do with their two layer MODFLOW model with a stream package connection to PCRGLOB. There are so many assumptions present I think more careful discussion of how sensitivities in these assumptions (e.g. parameters in what amounts to the stream package used here) and feedback back to the WaterGap (which I think is just one-way at this point) would be really important.

Reply: Regarding the assumption of confined conditions, we now explain the rationale for it (see changes to manuscript). A sensitivity analysis is beyond the scope of this paper. We are currently preparing a paper that presents an extensive sensitivity analysis of the steady-state G^3M presented here.

Changes to manuscript: Regarding the assumption of confined conditions, we added to the second paragraph of section 2.3: "We choose to simulate confined flow conditions in both layers even though the upper layer can be expected to decrease in depth and thus in transmissivity (hydraulic conductivity times saturated depth). Every unconfined aquifer can have an equivalent confined representation assuming a correct saturated thickness (Sheets et al., 2015). However, given the large uncertainties regarding hydraulic conductivities (possibly an order of magnitude) and the lack of knowledge about aquifer thickness, it is appropriate to choose the computationally more efficient assumption of confined conditions."

#2.6 From Figure 2 it appears that not all the features are implemented in this model, or perhaps not all the features are activated except for recharge. Since the abstract discusses capillary subsides for plant water use but this feature is not described (nor is it entirely clear how that would be implemented as a simple flux), I think a thorough re-working of this discussion and assumptions are needed. Unfortunately, this figure begs the question why is a methods paper in GMD incomplete and not presenting all the model features?

Reply: The intention of Fig. 2 was to show how the gradient-based groundwater model G3M is planned to be coupled with/integrated into the global hydrological model WaterGAP. This information is necessary to understand the modelling choices made for the steady-state G3M presented in the manuscript, as a first step towards a fully coupled transient model. We think that a steady-state model is an important first step to justify a newly developed groundwater model and needs to be presented to the scientific community before moving further along to a fully coupled transient model. The steady-state model alone shows the difficulties of simulating groundwater flows at the coarse spatial

resolution required for global-scale modelling. The model feature capillary rise is not presented as it cannot work without coupling to the soil compartment of WaterGAP.

Changes to manuscript: We added the following sentence to the last paragraph of the 1 Introduction: "Capillary rise is not included in the presented steady-state simulation as simulation of capillary rise requires information of soil moisture that is only available when G^3M is fully integrated into WGHM."

#2.7 The maps of water table depth seem to have a tremendous shallow bias. It is hard to say because of low figure resolution, but perhaps most of Eastern N America, most of Australia, half of Europe and all of Tropical Africa are under water. I think additional discussion is needed here at least. Could this be due to the steady state assumptions? Confined conditions? The stream aquifer package? Resolution and slope? ET feedbacks?

Reply: The visual impression is wrong, only the darkest blue means "under water", and this happens only in 2.1% of all cells. As we write (already in the first manuscript version) in section 3.1, "In 2.1 % of all cells, GW head is simulated to be above the land surface elevation, by more than 1 m in 0.3 % and by more than 100 m in 0.004 % of the cells.". Still, areas in Eastern N-America, Australia, Europe and tropical Africa present very shallow groundwater tables. This is mainly due to large wetland extends in these areas in connection with the steady-state approach. The extent of all wetlands (global already reduced by 20%) likely is overestimated as the data represents a maximum extend that is rarely reached in reality. Additionally, wetlands don't have a steady-state (or rather no surface water body) thus the interaction with the groundwater is likely overestimated and leads to the observed flooding.

Changes to manuscript: none

#2.8 It's hard to tell what the difference is here between the PRCGlob-MODFLOW model and this current model. More discussion is needed to clarify this distinction. I actually feel it's okay if there are many similar models out there (and both can be good models or bad models, it's not a competition), I would like more dissection of the differences in approach.

Reply: Already in the first version, we wrote in the abstract "Together with an appropriate choice for the effective elevation of the SW table within each grid cell, this enables a reasonable simulation of drainage from GW to SW such that, in contrast to the GW model of de Graaf et al. (2015, 2017), no additional drainage based on externally provided values for GW storage above the floodplain is required in G^{3}M. Comparison of simulated hydraulic heads to observations around the world shows better agreement than de Graaf et al. (2015)." More explanation about this additional drainage required by PCR-GLOBWB but not G^{3}M is given in the introduction: "The first global gradient-based GW model that was run for both steady-state (de Graaf et al., 2015) and transient conditions (de Graaf et al., 2017) was driven by GW recharge and SW data of the GHM PCR-GLOBWB (van Beek et al., 2011). However, there is not yet a two-way coupling of a GW flow model and a GHM. This may be due to the way de Graaf et al. (2015, 2017) modelled river-GW interaction. To achieve plausible hydraulic head results, they found it necessary to add an additional drainage flux to GW drainage driven by the hydraulic head difference between GW and river. This additional drainage, which accounts for about 50% of global GW drainage, is simulated as a function of GW storage above the floodplain, the values of which are computed externally by the linear GW reservoir model of PCR-GLOBWB (Equation 3 of de Graaf et al. (2017) – the model component that the gradient-based model was intended to replace. This prevents a full integration of the global GW flow model of de Graaf et al. (2017) into a GHM, as then, the linear GW reservoir model would be replaced by the GW flow model." The section in the discussion read "De Graaf et al. (2015) set their SW head (h_swb) to the land surface elevation of the 6' grid cells minus river depth at bankfull conditions plus water depth at average river discharge. Together with the missing interaction between lakes and wetlands and a different approach to river conductance, this might be a reason for the additional drainage above the floodplain that was necessary to avoid excessive flooding. On the other hand, this adaption allows the drainage of water even if the hydraulic

head is below the SW elevation that might have led to the global underestimation of hydraulic heads. Thus, the difference in model heads seems to be closely related to the sensitivity of SW body elevation."

Changes to manuscript: We modified the section in the discussion on the comparison to the gw model for PCR-GLOBWB by adding (see bold words): "De Graaf et al. (2015) set their SW head (h_swb) to the land surface elevation of the 6' grid cells minus river depth at bankfull conditions plus water depth at average river discharge. Together with the missing interaction between lakes and wetlands and a different approach to river conductance, this might be a reason for the additional drainage above the floodplain that was necessary to avoid excessive flooding, and that is not needed in G3M. On the other hand, this adaption allows the drainage of water even if the hydraulic head is below the SW elevation that might have led to the global underestimation of hydraulic heads. Thus, the difference in model heads seems to be closely related to the sensitivity of SW body elevation.

#2.9 The current model is also completely different from the Central Valley model. This strikes me as odd too. Is it water use? Boundary conditions?

Reply: The presented Central Valley model plot show the initial state of the CVHM model and not computed model results. The initial condition represents the close to natural conditions in the early 1960s in the Central Valley with a very shallow groundwater table and large wetlands. Scale is most likely the main driver for the different results. Except for the scale differences G^{3}M correctly computes shallow conditions close to the values assumed by CVHM with groundwater above the surface in the north and partially in the south of the valley. Furthermore, the depth to groundwater decrease towards the Sierra Nevada. Other differences are likely due to the steady-state and the connected assumptions on surface water bodies.

Changes to manuscript: Page 16, Line 14-17 "G^{3}M correctly computes the shallow conditions with groundwater above the surface in the north, partially in the south of the valley and decreasing towards the Sierra Nevada. The difference in the extend of flooded area could be due to large wetlands areas still present in the early 60s which are not represented in this extent in the data used by G^3M." Page 18, Line 3-6 "The comparison to the initial state (based on historical observations) of the CVHM model presents a first comparison within a data-rich region which provides also the future possibility of comparing transient model results and human impact on a regional scale. G^3M is able to reproduce the shallow groundwater table in the early 1960s. Differences are likely due to the steady-state approach and the connected assumptions on surface water bodies."

Please also note the supplement to this comment:
https://www.geosci-model-dev-discuss.net/gmd-2018-120/gmd-2018-120-AC1-supplement.pdf

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-120, 2018.

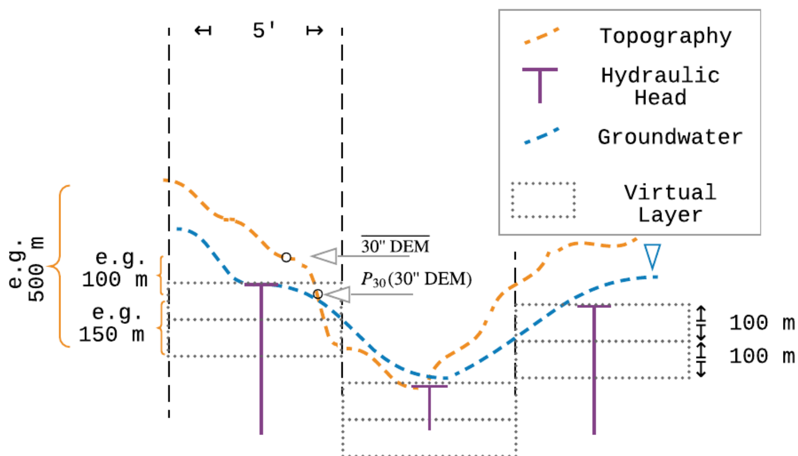

Fig. 1. revised figure 1

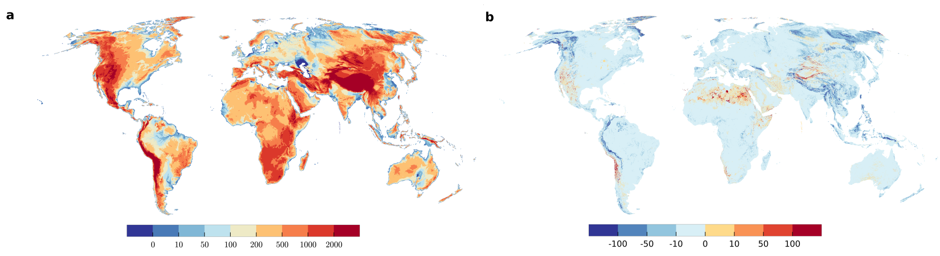

Fig. 2. revised figure 3