Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-121-SC2, 2016

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

Interactive comment on "A global-scale two-layer transient groundwater model: development and application to groundwater depletion" *by* Inge E. M. de Graaf et al.

I.E.M. de Graaf

idegraaf@mines.edu

Received and published: 3 May 2016

Reviewer 1

We thank Reviewer 1 for the thoughtful and extensive evaluation of our work. He/she raises a number of valid points and concerns, which we hope to address in the answers below. The reviewer's points have certainly been very useful in improving the paper. The reviewer comments are in italics and our rebuttal in roman.

This paper presents an ambitious attempt at hydrologic modeling at the global scale. The work builds on a previous model of the author by adding confined aquifer units and using a transient model. Given the scarcity of physical subsurface data available at the

global scale, large assumptions were made about aquifer structure and parameters. The work is clearly a step in the right direction, and we need to test our ability to model these systems, but the usefulness of the results is not clear.

We thank the reviewer for his/her assessment of the merits of our paper. We agree that large assumptions had to be made. This is yet another step towards a global groundwater model, while we certainly admit that many uncertainties remain.

I have two specific concerns:

1) To calibrate a global model with observations only from the United States and from one delta in Europe doesn't seem reasonable. Especially, given that a major value of the model lies in it's ability to parameterize subsurface systems or predict groundwater level changes for the remainder of the world, where we happen to have few observations. At this point, perhaps the model should just be applied to the US and part of Europe, where the model structure can be better tested?

We appreciate the concerns about the fact that we only used observations from the US and Europe. However, we must remark that an earlier steady-state version of the model (De Graaf et al., 2015) was already validated on all observations provided by Fan et al. (2013), which also consist of observations from Spain, Brazil and Australia. However, many of these observations are not time series but single or time-average numbers. We therefore used these in our previous steady-state version (De Graaf et al., 2015) but not in this transient version. We need time series to calibrate a transient model and these are only freely available in the US and Europe (i.e. in Europe it is not only the delta but the entire Rhine-Meuse basin including Germany, Switzerland and parts of France). We appreciate that we have to extrapolate over the entire earth from this but there are a number of reasons to attempt this: 1) There is no way around it if one seeks to establish a global groundwater model; 2) as stated, the steady state version of the model has been validated against a wider set; 3) Even though only in parts of the world, we use data from large range of climates, geological settings and landscapes in a split-sample test and are thus reasonably confident that our results make sense; 4) we have supporting validation (from GRACE and reported depletion data from India) that our model produces results are reasonable outside the calibration areas. We will provide an additional discussion on these points in the Discussion part of the paper.

2) I'm concerned about the overall discrepancy between the representative model depth, and the system it aims to simulate. This stems from using surface geology to infer aquifer properties, and similarly, using surface geology to infer the presence and properties of confining units. Many primary aquifer systems are multi-layered with numerous confining units and aquifers with varying properties with depth. The objective of the paper could be to just model the near-subsurface system. However including groundwater extraction values, which in many cases are drawn from deep systems, may force the calibration process beyond reasonable limits.

This is indeed only a second attempt to assemble global hydrogeological model of the world. We know that local groundwater models provide more confining layers in sedimentary basins and we lump these all together in one confining unit and one big aquifer. Using the assumption that productive aquifers coincide with sedimentary basins and sediments below river valleys, the distinction is made between (1) mountain ranges, and (2) sediment basins representing the aquifers. Subsequently, aquifer thicknesses were estimated by relating these to terrain attributes (e.g. curvature) after calibration of these relations with reported aquifer thicknesses from U.S. groundwater modeling studies (de Graaf et al., 2015). These thicknesses are those of productive aquifers until the first impermeable basis. This means that our aquifers lump together aquifers separated by semi-permeable layers into one big aquifer. This may under-estimate head decline. For the confining layer properties those belonging to fine grained or mixed-grained (with and without layers) sediments were taken and for the underlying aquifers the properties belonging to coarse-grained sediments. The storage coefficient of the confined aquifers and the horizontal and vertical conductivities of the confining

layer and aquifers were calibrated. This resulted in an anisotropy ration of kh:kv =10:1 , which effectively corrects for neglecting semi-permeable layers and lumping multi-aquifer systems into one. This is indeed an approximation and will result in a first-order estimate of average head decline over the entire aquifer set. We realize this and hope that by incorporating more regional to local studies in the future our model will slowly grow more mature. We will provide a more thorough discussion of these uncertainties in the Discussion part.

Overall, it's clear that a significant amount of work went into this, and it moves us closer to having a global groundwater model. Addressing some of the comments here about model structure, calibration, and uncertainty in storage change will help clarify the value of the model and it's results.

We thank the reviewer for these nice words and will try to answer the specific questions below. Whenever possible we will add additional explanations and discussion to the manuscript to clarify methods and the value of our results.

Specific comments:

1) It would be helpful to conceptually explain the model and assumptions a bit more clearly in the methods section. Obviously the data required to model these deep aquifers is rare, certainly at the global scale – so the current project is making reasonable assumptions in order to move the understanding forward. Given that, it should be clear early in the paper what system and dynamics it expects to model reasonably well, given the data input restrictions. Broadly, the model improves upon a previous version, which modeled all aquifers as unconfined. Is the current model explicitly modeling the most surficial aquifer and most surficial confining units only? The permeability values represent the surface geology, and the confining unit permeability also seems to be based only on the shallowest layer of material.

See the answer above under item 2. We will more explicitly state in the Introduction what the model supposed to represent and in the discussion what the restriction and uncertainties are of the approach and what data would be needed to improve estimate in the future.

2) Most of the large-scale (irrigation, industrial, municipal, etc.) groundwater usage is drawn from deep wells, whose regional aquifer characteristics may not be well represented in this model. Can the authors discuss how calibrating the model with relatively shallow aquifer input parameters to fit potentially deep system extraction rates may impact the model performance? It seems like there may be a discrepancy between the system modeled and the one it is calibrated to.

As stated above, we presume to represent the entire stack of upper aquifers until the first impermeable layer (hydrologic basis) without explicitly resolving the semi-permeable layers explicitly but modeling their effects using anisotropic effective conductivity. It means assuming that some surface lithological properties are representative for deeper layers (lacking any other information). But effective properties are calibrated based on the head measurements, some of which are also from deeper sediments. Also, its should be noted that in many areas of the world (India, China, Iran) groundwater abstraction occurs from small agricultural wells that are not deep at all.

3) Given a two-layer model, are interactions with a shallow unconfined aquifer (e.g. alluvial aquifer overlying a confining unit) lost? Are there specific areas where surface water - groundwater dynamics were not well represented, perhaps useful for guiding future research to improve our subsurface parameterization capacity in these areas?

We don't believe so. We estimate the width and depth of the surface water systems using geomorphologic relationships and bankfull discharge. If these dimensions are such that they penetrate the confining layer, which is the case for large rivers and rivers at the edge of coastal confining layers, we have as expected intensive interaction between the surface water system and the aquifer. Otherwise, if the bottom of the surface water is positioned within the confining layers, surface water groundwater interaction is

limited.

4) Were groundwater observations from all well depths used to calibrate the model? For the confined aquifer areas, it is highly possible that groundwater observations are being made in multiple aquifers, where deeper layers would not be expected to have a direct connection with the surface as is being modeled. I understand we cannot expect this level of detail to be included in the model, I'm just curious how fitting a model to these data will impact your results.

We used all the available time series of sufficient length in the calibration. This means indeed that there may be head observations that are not or barely influenced by e.g. surface water levels or recharge while our model does calculate this influence. The lack of information on the precise vertical hydrogeological structure of the aquifers around the observation well screens in this case results in poorer calibration results.

5) The brief description of how aquifer thicknesses were calculated (in addition to the citation to the 2015 paper) is helpful. Can a similar one be provided for how thickness of the confining unit were calculated?

We have added a description of this. For the description of the coastal aquifers we refer to 2.3.2 and for the others (10% of the total thickness) 2.3.3.

6) Were the parameters for the confining units assumed based on the surface unit texture? Were any measurements (or regional model parameters) used to inform individual aquifer confining unit permeability, or were they set uniformly across the globe?

They were set uniformly given the surface texture (fine grained, mixed grained and mix grained layered). Then a single prefactor was used to calibrate them further. So no regional information was used.

7) In the methods section 2.1.2, Does "Next to the river levels" mean proximally adjacent to? Or "next" figuratively? It sound like there are fixed head boundaries being specified at sea level adjacent to all the rivers. If this is correct, can you justify why you

chose to do this? Can you explain this decision with respect to Figure 8? The depth to groundwater appears to follow topography (as you say in the paper), and is simulated quite a bit deeper than observed (e.g. much of western US and Mexico).

This should have stated “apart from the river levels”. We will re-write it like that. Furthermore. We state in section 3.3 “Also, for mountain regions deep groundwater tables are simulated. In these areas local aquifers in sedimentary pockets in mountain valleys are smaller than the grid resolution (< 10 km) and are therefore not captured. As a result, groundwater heads in these regions are likely underestimated (de Graaf et al., 2015).” However, we realize that we should state this more clearly in the introduction already and will do so in the next version of the paper.

8) If it took 10 years for the model to reach equilibrium, does that say something about the degree of disequilibrium in the groundwater system in 1960? Do you think years is reasonable? If so, or not, can you infer something about how the model is functioning?

We don’t believe that it is a matter of model functioning. It takes just quite a few years to warm-up the model. This has to do with the large volume of groundwater in the model causing considerable inertia. So we start with a steady state, then run 1960-1970 first and start with 1960 again.

9) There are two periods of rapid groundwater depletion in Figure 12 early 1980s and 2000s. You explain the first as being delayed despite overall abstraction > recharge (is that right?) by stream capture. Is this a process that would be included in the model, without having feedback from groundwater level on surface water?

This process is still there indeed. Although the effect may be somewhat overstated. What happens is that we have the larger rivers connected to the groundwater system in MODFLOW (through the RIV package) and the smaller rivers by the drain package. When one starts to pump more than is being recharged part of it will come out of storage, but in the beginning part will come from reduced discharge (to rivers and drains). After the drains fall dry, part of it may still be supplied by the rivers (river bed infiltration)

and this part increases as the groundwater head drops. After the head drops below the river bottom this infiltration flux becomes constant. So after that most of the additional pumping must come from storage and cause increased depletion rates. Of course, this effect may be overstated because river levels themselves are kept constant in our approach, while they would also decline in reality causing a more gradual increase of depletion rates.

10) The total groundwater depletion is given with 4 significant figures. Can you justify this precision? Can you provide an estimate of uncertainty on the depletion estimate based on errors associated with the groundwater level simulations and storage values?

No we cannot justify this. Thanks for this insight. We should change this to 2 significant figures and use scientific notation.

11) The conclusion that model performance is only slightly better with the inclusion of the confined systems suggest that we do not need to model the confined systems? Or that they should be modeled another way?

That is a good point. The model performance in terms of heads is not better. However, our estimates in terms of depletion rates are closer to previous volume-based estimates (e.g. Konikow, 2011) and thus believe that these are better. We will add this observation to the paper.

12) Many of the figures can be tightened up: they could use subfigure labels (A,B,C,etc.), and make sure the axis labels are final (some say "data", several are missing), and that for figures with subplots that the axis line up for all figs. There are a handful of typos, but those can be corrected with minor effort.

We will provide updated figures in the next version.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-121, 2016.