Reply to comments by reviewer 1 on “Palaeo-modeling of coastal salt water intrusion during the Holocene: an application to the Netherlands” by J. R. Delsman et al.

We thank reviewer 1 for his/her kind words and helpful suggestions. Questions raised by the reviewer are in bold face, our answers in regular face.
On behalf of all coauthors,

Joost Delsman

__General comment__

This paper discusses a very interesting topic of complicated salinity distribution in coastal aquifer and provides a unique approach of hydrological modelling combined with hydrochemical considerations. I, therefore, strongly recommend that it will be published in this journal provided some significant modifications will be made. The main problem is that the paper is quite difficult to follow and some parts can’t be judged by the readers due to lack of explanations. Some of these modifications are relatively easy to make, such as showing some of the figure in colors so that it will easier for the readers to notice the specific details. Other will require more work, such providing more information about the different geochemical data so that the readers will not need to go to another previous paper in order to understand the current one. I am sure that some of specific comments stems from my lack of understanding that can be dealt with by farther explanations.

We thank reviewer 1 again for his/her kind words and helpful suggestions and will do our best to improve our paper accordingly.

__Specific comments__

__Abstract__

1) The abstract is well written but a little too general. It will be nice to have more specific results of this study in the abstract, possibly relating to the use of the geochemical groups.

We pretty much completely revised the abstract and included more specific results.

2) The first several lines are kind of introduction and can be shortened.

We shortened the first four lines of the abstract.

3) I am not sure that the term “perceptual tool” is the correct one here – why not “conceptual tool”?

We changed perceptual to conceptual throughout the paper.

4) Where is the information for “pre-Holocene” coming from? Is it from radiocarbon dating of this water? If yes, then you need to change the text where it is written that radiocarbon is not

accurate and can't be used. If this time frame is obtained from the simulations, it should be specified as such.

The original statement was based on our modeling results, where water initially present in the model domain (termed pre-Holocene for easier recognition) required a minimum chloride concentration to match its current location. We changed the text in the abstract to emphasize that this statement is based on our modeling results.

On a side note: our understanding of this "current location" is indeed partly based on radiocarbon dating, but allowable given the large age contrasts considered (Post, 2004). Our statement on the accuracy of radiocarbon dating in the Methods sections was directed at groundwater dating with a much higher resolution; see also our answer to the reviewer's specific question relating to the accuracy of the radiocarbon measurements.

5) The last sentence is too general – I suggest that you elaborate and be more specific or delete

We included the gist of the concluding sentence in the reworking of the abstract and deleted the last sentence.

<u>**Introduction**</u>

This chapter is well written but I missed some needed information with regards to the details of this studied area.

First, I suggest to have a chapter of "site description" as a second part of the introduction and move what you write in the "method" chapter to this chapter.

We moved the description of the study area and its palaeogeographic development from the methods chapter to a separate 'site description' chapter, immediately following the Introduction.

Second, you need to give a better summary about the concept of "hydrosome" and hydrochemical facies and characteristics of each of them from the published papers of Stuyfzand. The description in the "Hydrochemical facies analysis (HYFA)" chapter is not elaborated enough. The lack of such information make this paper difficult to read – the readers should be able to understand the paper as is and not to be forced to read the previous papers.

We moved the description of HYFA to the site description chapter, and elaborated the section to make it more self-explanatory, by briefly describing the 6 hydrosomes in the area and the 4 hydrochemical facies parameters, which in combination define the hydrochemical zones within a hydrosome.

<u>**Methods**</u>

It will be better to give more specific details with regards to the recharge coefficients and not just to mention "precipitation surplus (precipitation minus evaporation)". I suggest to give the actual estimations of precipitation and evaporation here and not only the surplus. Is the recharge the same in the whole studied area? Is there no different between the dune area and the other areas some of which seems of lower permeability? Is there no surface flow, which could affect the precipitation surplus?

We split net precipitation in the manuscript in total precipitation (840 mm/y) and evapotranspiration (590 mm/y). We applied uniform recharge over the entire model area, given the lack of data on historical vegetation patterns. Recharge differences related to palaeo-vegetation differences occurring in the Netherlands are in the order of 10% (Van Loon et al., 2009), well within other uncertainties in the model. Overland flow by infiltration excess is very rare in the Netherlands given the low relief and permeable soils and therefore neglected. We did, however, apply infinite drainage at the model surface, to simulate small-scale drainage structures and saturation-excess overland flow. We better clarified this in the text.

Are the Tertiary clay layers the same as Maassluis Formation?

The Maassluis Formation is a formation deposited in the early Pleistocene and situated just above the Tertiairy clay layers. We changed Fig.1b and the text on the model boundary conditions to clarify this.

Is there separation between the different sub-aquifers? Are all parts connected to the sea in the same manner? The system seems to behave as one with only local effects of the less permeable layers. Is this something specific to the studied site or this is a good representation of the Dutch coastal aquifer?

It is indeed typical of the Dutch coastal groundwater system that aquitards are discontinuous, and the system functions, on a large scale, essentially as one (Van der Meulen et al., 2013). In our paper we also show however, that the discontinuous aquitards do still have a pronounced effect on the groundwater salinity distribution.

It is not clear how you used the model of Goode for the age distribution. Moreover, there is no output of such simulation to show the readers what you mean in this part of your work and how do the simulated ages fit the tritium data. As you know, the tritium data give information only for ~50 years which is very small part of this study.

We extended the explanation on the use of Goode's method of age modeling in our paper. We added that we implemented the method of (Goode, 1996) by using MT3DMS in a post-processing step after SEAWAT (saving the flow results), using a zero-th order decay term (of – 1/365, to obtain age in years) in MT3DMS (Zheng, 2009).

Tritium data can only really be used to discern pre-1950s water from younger water, which is indeed only a small part of our study. This information however permits the delineation of infiltrating and exfiltrating parts of the model transect. Earlier versions of this manuscript included a figure of the comparison between the calculated 1952 contour and the limited available tritium samples (Figure R1 in this reply). Unfortunately, the comparison was hampered by the orthogonal projection of samples from several kilometers away (more so than the more abundant head and chloride measurements, used in the paper). The delineation of infiltrating and exfiltrating areas is very localized in this part of the Netherlands, resulting mainly from (anthropogenic) differences in surface level and geology (compaction) and variation in surface water / drainage levels, not only along the model transect, but also perpendicular to it. This variation is obviously not included in the employed 2D modeling approach. As an example, tritium ages show very young water below the Horstermeer (the small reclaimed lake at x-coordinate 135 km attracting saline groundwater). These samples are however taken just north of and outside the Horstermeer, and are more representative of the recharge areas along both sides of the Horstermeer. A similar explanation explains the mismatch around km 122. Given these explanations, we felt that the calculated 1952 age contour (and hence the modeled infiltration / exfiltration patterns) lined up very well with the available data. After some consideration, we excluded output from the age calculations for sake of brevity, the age calculation is however shown in upper right hand panel of the movie included as supplementary information. We feel however that the general good correspondence is worth mentioning and like to keep it in the manuscript. We will change the manuscript to refer readers to the supplementary information for the modeled age results, and better highlight the model simplification resulting from the 2D approach in the discussion.

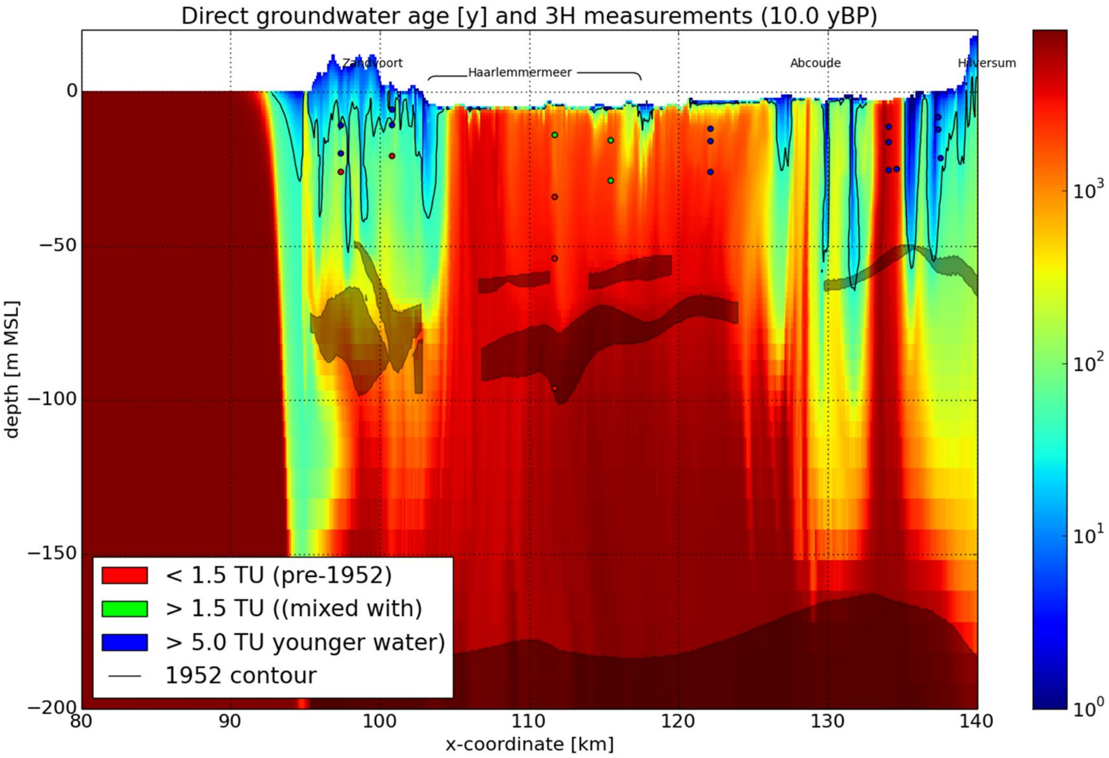

Figure R1: Modeled direct groundwater age (Goode, 1996) at 1990 AD compared to decay-corrected ^{3}H measurements. ^{3}H samples with concentrations less than 1.5 TU are considered older than 1952 AD, higher concentrations are younger or are mixed with younger water. The modeled 1952 contour is shown in black.

Are the radiocarbon ages of Post not accurate or not good at all? If the problem is with accuracy, then you can still use them for this work since, as you state, is more conceptual than quantative.

On the applicability of radiocarbon ages Post (2004) says, in his own words:

"Due to the large contribution of sedimentary carbon sources to groundwater DIC and the heterogeneous age distribution of carbon-bearing materials in the subsurface, accurate radiocarbon dating of the brackish and saline groundwater proved impossible. Carbon-14, however, could be used to distinguish a Holocene from a Pliocene to Early-Pleistocene seawater contribution." (Post, 2004, p. 45)

We opted not to use this data as we are mostly interested in the within-Holocene age variation, and Post (2004) only presents one sample with dates with depths.

Results

As written before, it is impossible for the readers to judge the comparison between the tritium ages and the results of the simulation. If this is important to be mentioned, then you need to give specific results (either as another figure or some specific data)

See response under methods

The explanation for the discrepancy between the results of the simulation and the measured heads at about elevation zero need to be better explained.

We are not sure what discrepancy the reviewer is referring to here, as the fit between modeled and measured heads at elevation zero even seems better than at other elevations. At least, for the measurements excluding the coastal dune area. Discrepancies in the coastal dune area are explained by the fact that measured heads in that area are much influenced by stresses not well included in our model (well fields, artificial recharge). Nevertheless, we changed the manuscript to better clarify discrepancies in the coastal dunes area.

The term "model conservative tracers" is not clear in this work. It seems that in this work it means specific geochemical group, that you define as "hydrosomes". But these need to be better defined. It is not clear how you dealt with them in the simulations. Are there a specific parameter that defines each group, either in the model or in the field or in both ? Is there mixing between the different groups? You also need to define better the other related terms such as "infiltrating transgression" and classic seawater intrusion – is it mostly about the time of the intrusion or there are other differences between these terms? Again, an elaboration in the second part of the "introduction" will be very helpful.

The two terms ('modeled conservative tracer' and 'hydrosome') are two related, but different concepts. The term hydrosome specifically refers to the work of (Stuyfzand, 1993), in which he delineated different groundwater origins based on environmental tracers. With modeled conservative tracers we refer to our approach to model the evolution of specific origins of groundwater through modeling fictitious tracers that represent the fraction of groundwater with a specific origin (e.g., the origin 'surface water' is modeled by applying a concentration of 1 of the tracer 'surface water' at river boundary conditions). These two are obviously related, which allows us to compare the model result with the hydrosome delineation of (Stuyfzand, 1993). The modeled origin tracers do however not map one to one to the hydrosomes. Next to elaborating the description of hydrosomes and the HYFA analysis as previously mentioned, we will better explain the modeling approach in our manuscript, change the term 'conservative tracer' to 'origin tracer' and expand Table 1 (description of origin tracers) with their relation with hydrosomes. Specific references to origin tracers have been formalized by referring to their names in Table 1 (instead of vaguer process-descriptions like infiltrating transgression water) with capitalized first letters.

Discussion

The differences in the thickness of the mixing zones in the different parts of the section are interesting. Can you be more specific and mention what you mean by wide or narrow zone? Also, can you elaborate on the reasons for this difference? Could the flow velocity (of the fresh water) be an important factor?

We specified the widths in the discussion (narrow: some meters, wide: several tens of meters) and added several lines in the Discussion to elaborate on the reasons for this difference.

We argue that there exists a distinct difference between the process driving mixing underneath the coastal dunes and that more inland. Beneath the coastal dunes exists a more or less stable BGH lens, where a shallow mixing zone results from transversal dispersivity and diffusion where the fresh and saline waters meet. More inland however, large scale mixing occurred during transgression, where transgrading seawater rapidly infiltrated by means of unstable fingering (Kooi et al., 2000; Post and Kooi, 2003). Kooi et al. (2000) describe the mixing process, and the resulting fairly stable zone of intermediate salinity, by two occurring phases: (1) the development of small-scale fingers (fingers with widths in the order of meters), protruding quickly all the way to the aquifer bottom with limited dispersion, and (2) negligible flow after the fingers reached the bottom, after which diffusion dominates the exchange between the saline fingers and the intermediate fresh water. Our larger model cell size prohibited, however, the modeling of the small-scale fingering and the accompanying mixing observed by these authors. A larger dispersivity value (α_L = 1m) was needed to mimic the mixing result of these small-scale processes in our (necessarily) coarser model cells. This larger value in turn resulted in an apparent overestimation of the mixing zone underneath the coastal dunes. A dual porosity approach (Lu et al., 2009) could perhaps yield better results, but was outside the scope of the present paper.

I am not sure that this approach is so different from the others that start with steady state condition. It seems to me that the conditions that you use for 6500 BC could be defined as steady state condition and the results will not be very different. The interesting point in this work is that you also change the hydrogeological configuration with time.

The main point that we wished to convey regarding steady-state, is that in order to create a transport model of a similar groundwater system for current conditions or future predictions, one shouldn't start with a steady-state that is conditioned on present-day boundary conditions. We will make this clearer in the manuscript. But even in our case, starting with a steady state would explicitly entail neglecting the salinity in the Maassluis formation (groundwater system would be completely fresh except for saltwater intrusion along the coastline), while we showed that the salinity of the Maassluis formation is important in correctly modeling the salinity evolution.

The actual way that this is done and its effect on the modeling is not elaborated enough. I understood that the end condition of each step is the initial condition for the next step – but, did such exercise create problems since these conditions do not fit the new configuration? In real life, the geological and hydrological condition are changed gradually and not in in jumps, some of which are quite big.

The jumps in boundary conditions (time slices) are another aspect of simplification of the complex reality. Like we chose spatial and temporal discretizations etc., we chose ten big steps in geological and hydrological condition. The jumps in geology and/or boundary conditions were indeed sometimes large, but we did not encounter any problems on the time-slice transitions, in convergence or otherwise. We elaborated on the time-slice transition procedure by noting that we filled 'new' cells with the uppermost concentration that existed below the new cell. We added a line to the results section stating that we did not run into any convergence problems at the time slice transitions, and we added a line to the discussion section stating that we simplified gradually changing conditions to discrete jumps.

Figures

Figure 1 – (a) – it is very difficult to see the details without colors – is there a reason for giving this figure in black and white? Can you show the bathymetry? It is difficult to see the ridge in the sea, otherwise (b) Again, it will be much better in colors. It is very difficult to see the details especially in the upper parts What do you mean here by heterogeneous (in the legend)? Which parameters?

We added color to Fig 1, and also show the bathymetry of the North Sea and Lake IJssel. We are not sure what the reviewer means by the 'ridge in the sea', but suspect it to stem from possible confusion between the modeled transect (A-B) and the geological profile (A'-B). We more clearly marked the extent of the geological profile in Fig 1.

'Heterogeneous' in Fig 1 are not further lithologically differentiated formation members of the geological classification used by the Geological Survey of the Netherlands, due to a lack of borehole information. (Uniform) Hydraulic parameters are however assigned to these formations by the Geological Survey (Van der Meulen et al., 2013; Vernes and Van Doorn, 2005; Weerts et al., 2005). We changed the label from 'heterogeneous' to 'undifferentiated' in Fig 1b to avoid confusion.

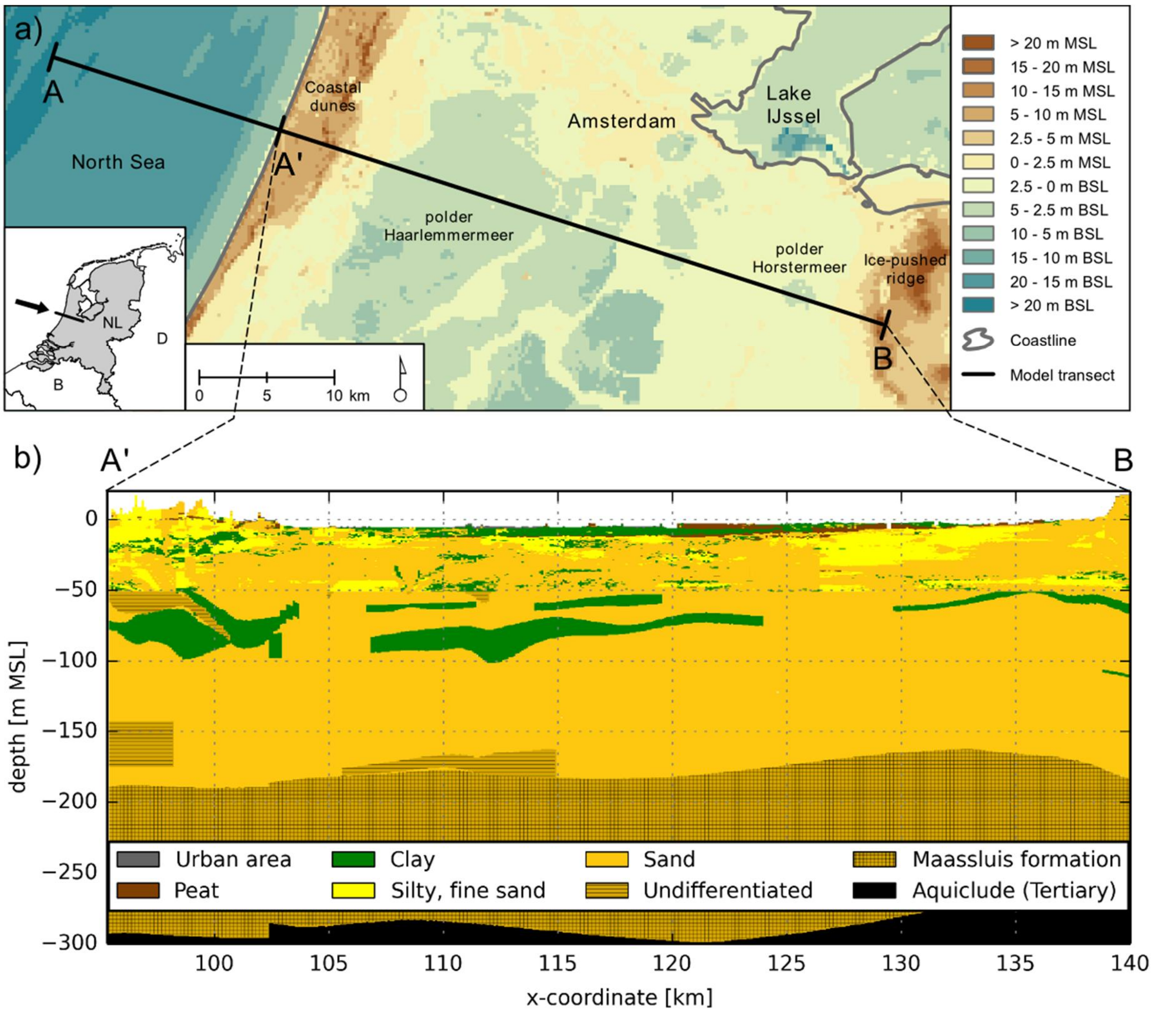

Figure 1. Location of studied transect (A – B), elevation and main topographical features (a), and a lithological cross-section along the transect (A' – B) (b).

Figure 2 – again, please give in colors. I think that a small figure of sea level changes during the whole studied period will be very helpful

We colored Fig 2 and added sea level rise insets to each palaeo-geographical map.

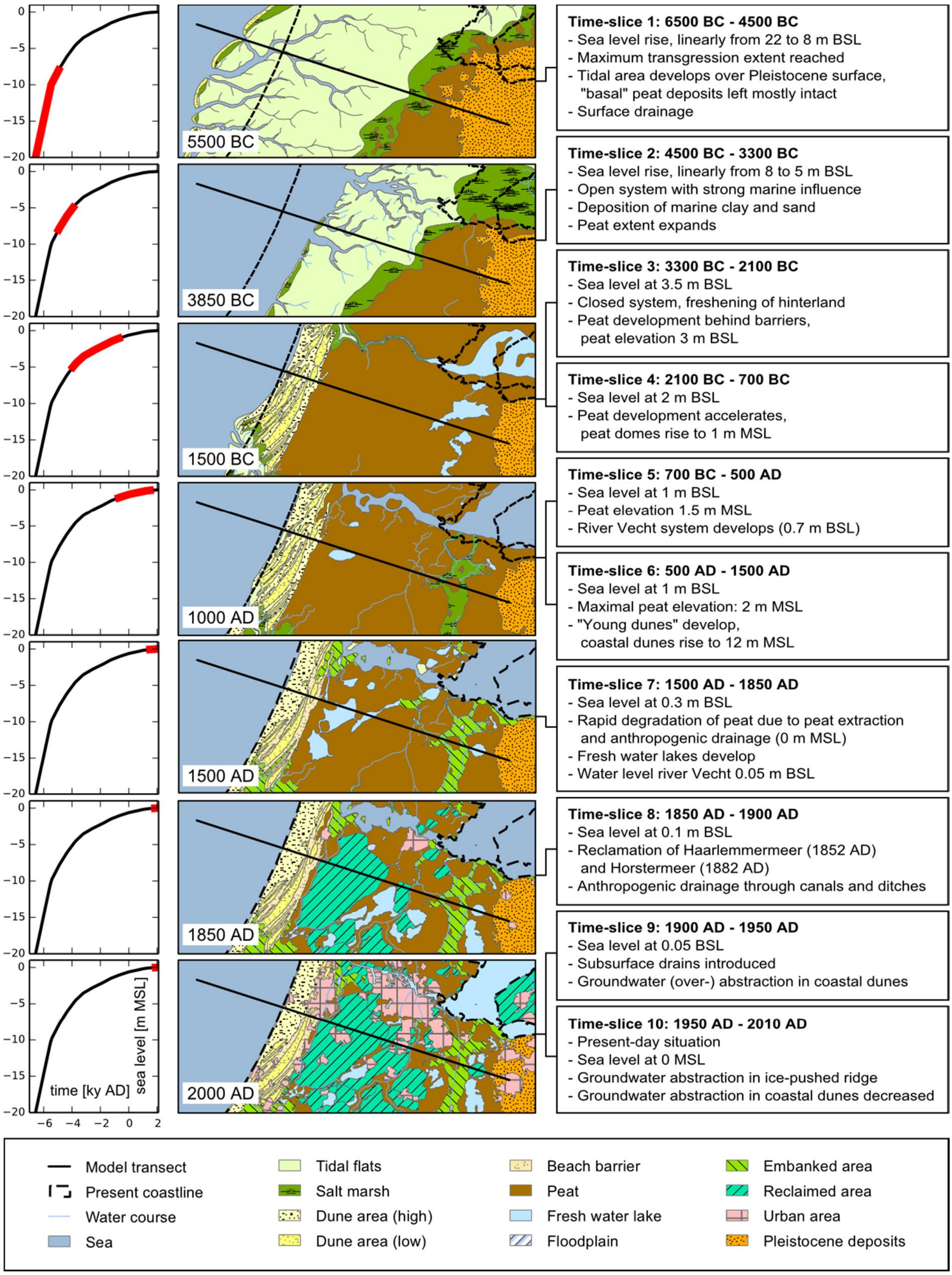

Figure 2. Overview of Holocene palaeo-geographical development (center), sea level rise (left, adapted from Plassche (1982), red curve denotes sea level rise during palaeo-geographical development phase) and description of model time slices (right). For reference, note that the extent of the palaeo-geographical maps equals the extent of Fig. 1a.

Figure 3 – I am not sure that averages are good in this case. If there are big variations with time, then it seems better to give data from specific times.

We specifically chose to compare model results with averages of measured head time series, as we did not model the short-term weather variations reflected in the variations around the measured averages. The chosen time period (1990-2010) was short enough to not include overly large changes to boundary conditions, and long enough to average out weather effects. We therefore respectfully disagree with the reviewer and will leave Fig 3. as is.

Figure 4 – The field data seem to fit the results of the simulation very well which is a great support for this work. Can you elaborate on the locations of disagreement and try to explain them? The effect of the clay layers seems surprisingly small given the fact that you run a transient simulation for not a very long period. Do you have specific data for the configuration and the permeability of these layers? I imagine that a change in these layers (e.g. significantly smaller permeability) will make a difference.

We elaborated the results section on Fig. 4 to include a short discussion on the locations of disagreement. As explained in the methods section, both the location and hydraulic properties are derived from the REGIS database, based on numerous borelogs and conductivity tests on bore samples. Uncertainty is however large, especially in the hydraulic properties, as measured samples range several orders of magnitude in k values, while we used only median values. That is of course also one of the reasons why we present our model as a conceptual tool. We added hydraulic parameters specifically as one of the model simplifications in the Discussion section.

I do not see any evidence for SGD in the lower sub-aquifer. Is there no evidence for relatively fresh groundwater at depth of 100-150 meters near the shore, like there is in some coastal aquifers in other countries? It depends, of course, also on the continuity of the clay layer. Again, it this typical for the whole Dutch coast or specific for this area?

We consider the modeled situation typical for the whole Dutch coast. Drainage and land reclamation caused inland surface elevations to be considerably below MSL. Groundwater flow across the coastline at greater depths is therefore directed inland, contrary to more undulating coastlines in other parts of the world experiencing seaward groundwater flow. We therefore do not expect SGD at depths of around 100-150m. However, in previous periods, SGD is modeled, but not visible in the modeled chloride distribution due to the lack of contrast between the inland brackish/saline water and the sea. It is however visible in Fig. 8, especially 8f. We added a sentence in the results section to highlight this point.

Figure 5 – this seems like a very interesting result and therefore it is a pity that it is not readable, even if I looked at it at 200%. You need to enlarge some of the writing from the work of Stuyfzand. Again, although the comparison is very nice in many cases, it will be nice to elaborate also on the differences. There are some areas that seem without data and still divided to groups – can you explain how this was done? Other data that are not given? Conceptual knowledge? Some of the names are different in (a) and (b) – Is there a reason for that? Can you explain these differences? How did you get the difference in ages?

We colored Fig. 5 and improved its overall readability by removing distracting clutter. We elaborated on the differences in the model validation section, and added the locations of boreholes used in the HYFA analysis to Fig. 5a. The section on HYFA analysis is elaborated to better describe the delineation of hydrosomes. Slight differences in names result from the fact that the origin tracers allow the delineation of more mixing zones than recognized by (Stuyfzand, 1993), e.g., LC/L^{m2}. The differences in ages are determined by the age calculation (Goode, 1996) explained in the modeling section.

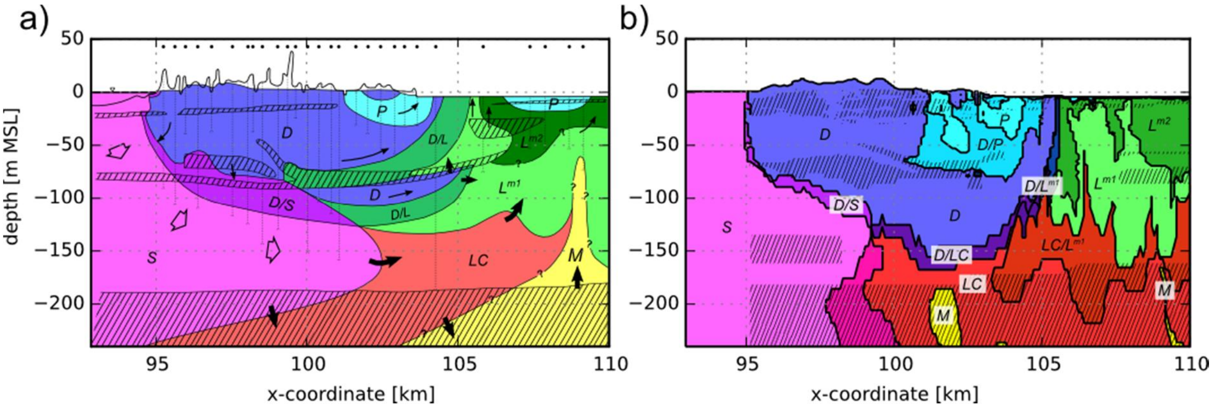

Figure 5. Position of hydrosomes, inferred from hydrochemical facies analysis (adapted from Fig. 4.6 in Stuyfzand (1993)) (a) and from modeled origin tracers (b). Capitals denote discerned hydrosomes: D = Dune (also containing nested artificial recharge hydrosome; not shown), LC = Holocene transgression (L) - Coastal type, Lm1 = L - Ancient Marsh type, Lm2 = L - Young Marsh type, M = Maassluis, P = Polder, S = (actual) North Sea. See Section 2.3 for a description of discerned hydrosomes, and Table 1 for the mapping of origin tracers to hydrosomes.

Figure 6 – What is the salinity of the brown color? I am not sure that I see it in the legend. Is the recharge the same in the whole section? If not, can you show that? It is difficult to see why there are the GH lenses since it is difficult to see the hydraulic heads. You may want to give few inserts of more detailed area in order to show what is missing in the large scale sections.

We checked that all used colors are indeed in the legend. We are not sure what the reviewer means by the brown color, we suspect where an aquitard (shaded area) overlaps the chloride distribution? We extended the legend showing legend items both with and without overlapping ‘aquitard-shading’.

Given the uncertainties surrounding recharge estimates for historic time periods, we opted to apply uniform recharge to the entire model domain, and for the entire modeled period. As stated before, we better explained this in the methods section.

We opted not to show heads for the obvious reason that using heads as a proxy for groundwater flow can be very wrong when dealing with large density differences. Instead, we show the streamfunction in Fig. 6 (and added it to Fig. 8) to visualize flow patterns. We agree with the reviewer that a comparison between the transient model results and the stationary BGH approximation is interesting, and added a line in the results section signifying deviations from the BGH approximation.

Figure 7 – are the SP units the same in the whole model? Why is there a change in SP in 1500? Is the rise in SP in the end a result of pumping?

SP units are the same in the whole model (kg Cl), but we only calculated the SP in the landward side of the model. The (gradual) change in SP after 1500 is caused by anthropogenic alterations to the landscape (drainage and hence land subsidence), causing a landward flux of groundwater over the sea boundary. This rise accelerates after 1850, when the Haarlemmermeer was reclaimed, further lower inland hydraulic heads. Pumping (from 1900 onwards) contributes, but is not the main cause of the increase in SP. We added a line in the results section explaining the rise in SP at 1500 AD.

Figure 8 –I do not see much change in the last 100 years although there were big anthropogenic changes. Did I miss something important? If there are changes, please refer the readers to them (possibly by a more detailed figure)

To start: groundwater flow is a slow process, and solute moves accordingly, even in 100 years. As noted in the results section, the effects of the reclamation of the Haarlemmermeer (the most important anthropogenic change) are effectively limited to the area approximately between x coordinates 103 and 108 km, where the low permeable layer at a depth of around 80 m is absent. Here, upward groundwater flow transported solutes upward at a rate of approximately 0.5 m/y, momentarily slowed down between 1900 and 1950 by the over-extraction in the dune area. Changes are indeed minor where the low permeable layer is present. We emphasized this point in the results section. Note that the upward transport is more clearly visible in the chloride results than in the results for the origin tracers, mainly because of the discrete legend versus the continuous blending legend of Fig. 8, and also differences in dispersion (see added line in discussion).

References

Goode, D. J.: Direct simulation of groundwater age, Water Resour. Res., 32(2), 289–296, 1996.

Kooi, H., Groen, K. and Leijnse, a.: Modes of seawater intrusion during transgressions, Water Resour. Res., 36(12), 3581–3589, doi:10.1029/2000WR900243, 2000.

Van Loon, A. H., Schot, P. P., Griffioen, J., Bierkens, M. F. P. and Wassen, M. J. J.: Palaeo-hydrological reconstruction of a managed fen area in The Netherlands, J. Hydrol., 378(3-4), 205–217, doi:10.1016/j.jhydrol.2009.09.014, 2009.

Lu, C., Kitanidis, P. K. and Luo, J.: Effects of kinetic mass transfer and transient flow conditions on widening mixing zones in coastal aquifers, Water Resour. Res., 45(W12402), 1–17, doi:10.1029/2008WR007643, 2009.

Van der Meulen, M., Doornenbal, J., Gunnink, J. L., Stafleu, J., Schokker, J., Vernes, R. W., Van Geer, F. C., Van Gessel, S. F., Van Heteren, S., Van Leeuwen, R. J. W., Bakker, M. A. J., Bogaard, P. J. F., Busschers, F. S., Griffioen, J., Gruijters, S. H. L. L., Kiden, P., Schroot, B. M., Simmelink, H. J., Van Berkel, W. O., Van der Krogt, R. A. A., Westerhoff, W. E. and Van Daalen, T. M.: 3D geology in a 2D country: perspectives for geological surveying in the Netherlands, Netherlands J. Geosci. / Geol. en Mijnb., 92(4), 217–241, 2013.

Post, V. E. A.: Groundwater salinization processes in the coastal area of the Netherlands due to transgressions during the Holocene, PhD thesis, Faculty of Earth Sciences, VU University Amsterdam, 154 pp., 2004.

Post, V. E. A. and Kooi, H.: Rates of salinization by free convection in high-permeability sediments: insights from numerical modeling and application to the Dutch coastal area, Hydrogeol. J., 11(5), 549–559, doi:10.1007/s10040-003-0271-7, 2003.

Stuyfzand, P. J.: Hydrochemistry and hydrology of the coastal dune area of the Western Netherlands, PhD thesis, Faculty of Earth Sciences, VU University Amsterdam, 362 pp., 1993.

Vernes, R. W. and Van Doorn, T. H. M.: Van Gidslaag naar Hydrogeologische Eenheid - Toelichting op de totstandkoming van de dataset REGIS II, TNO report, Netherlands Inst. of Applied Geoscience TNO, Utrecht, the Netherlands, 2005.

Weerts, H. J. T., Westerhoff, W. E., Cleveringa, P., Bierkens, M. F. P., Veldkamp, J. G. and Rijsdijk, K. F.: Quaternary geological mapping of the lowlands of The Netherlands, a 21st century perspective, Quat. Int., 133-134, 159–178, doi:10.1016/j.quaint.2004.10.011, 2005.

Zheng, C.: Recent Developments and Future Directions for MT3DMS and Related Transport Codes, Ground Water, 47(5), 620–625, doi:10.1111/j.1745-6584.2009.00602.x, 2009.