Answer to Reviewer 1

We thank reviewer 1 for his/her thoughtful review that greatly helped to improve our paper. The feedback and suggestions were very motivating for us to improve our paper. In this document, we have entered your comments in italics, added our response as regular text, and subsequently added suggested changes to the paper in red. After our responses you can find the suggested changes made to the text in red. We also like to add that we added supplementary information to this document, as it spares you the effort of looking for an appendix of a previous paper we wrote. Line numbers are written as "**[P… L…]**", indicating page number, line number in the <u>no markup</u> document.

General comments

First of all though, with regards to the Introduction section, you can build a stronger justification for this study by first discussing the studies that have been conducted in other areas (presently starting on page 3, line 25), which demonstrate that over-simplified models of large-scale aquifer systems are conceptually flawed. Then you can bring in the Nile delta system, and argue that there is also a need there to analyze the palaeo-geographical evolution in order to understand the present-day conditions. By starting off with a focus on the Nile straight away, you present it too much as a regional problem, not a scientific analysis that yields outcomes that can be transferred to other study regions.

This is a good suggestion, as it shifts the focus of the paper to the scientific problem instead of the regional problem. This will also help with Reviewer 2's main concern. We have added sentences of introduction to the broader problem **[P2 L1-7]** and consequently followed your suggestions. Given the severity of the local problems, we subsequently moved the former first paragraph to a dedicated section under "Area description" **[P4 L2-21]** .

The paper now starts as follows:

Palaeohydrogeological conditions have influenced groundwater quality in the majority of large-scale groundwater systems, since groundwater below 250 m depth is dominated by groundwater with an age of over 12 ka (Jasechko et al., 2017). These conditions can especially be found in deltaic areas, where the effects of marine transgressions are often still observed in groundwater salinities (Larsen et al., 2017). More specifically, their low elevation allowed for far reaching marine transgressions, leading to a large vertical influx of sea water, and hampered subsequent flushing with fresh water after the marine regression. This hypothesis is supported by hydrogeochemical research in several deltas (e.g. Colombani et al., 2017; Fass et al., 2007; Faye et al., 2005; Manzano et al., 2001; Wang and Jiao, 2012). The physical justification for this hypothesis, however, often still has to be tested, with a few notable exceptions (Delsman et al., 2014; Larsen et al., 2017; Van Pham et al., 2019), and can be provided by palaeohydrogeologic modelling, which has recently provided important insights for several cases. Gossel et al. (2010) created a large-scale 3D variable-density groundwater model of the Nubian Aquifer System and showed that seawater intrusion has occurred since the Pleistocene Lowstand towards the Qattara Depression (North-West Egypt). Later, Voss and Soliman (2013) showed with a

1

parsimonious 3D model of the same groundwater system that groundwater tables are naturally lowering during the Holocene, since they receive limited recharge and are drained into oases or sabkhas. Moreover, the authors used an inventive validation method by comparing the position of discharge areas in the model with a dataset of oases or sabkha locations. Delsman et al. (2014) conducted a detailed palaeohydrogeological reconstruction of the last 8.5 ka over a cross-section in the Netherlands to show that the system has never reached a steady state. They showed that the Holocene transgression caused substantial seawater intrusion, from which the system is still recovering. Using a combination of geophysics and 2D numerical models, Larsen et al. (2017) showed that during the Holocene transgression sea water preferentially intruded in former river branches in the Red River Delta, Vietnam. Van Pham et al. (2019) showed that most of the fresh groundwater in the Mekong Delta (Vietnam) was likely recharged during the Pleistocene and preserved by the Holocene clay cap. Despite being in a humid climate, recharge to the deeper Mekong Delta groundwater system is very limited and freshwater volumes are still declining naturally.

Moreover, the paper as a whole, but the Discussion section in particular, is a bit of a
confusing mix of a number of problems. There is (A) the scientific problem of understanding
the evolution of the groundwater salinity distribution in large delta systems
over long timescales. This is mixed with (B) the local management problem/question
of how much freshwater there is in the Nile delta. And then there is (C) the problem
that previous models have assumed steady-state conditions. For a publication in
HESS, the local management problem (B) is not the most important. The scientific
problem (A) is, and it should be made clear from the onset and throughout the paper
that this is the main focus. The implications for the local management problem can
be mentioned toward the end (it is especially interesting to note that depending on
the conceptualization, the locations where freshwater occurs will differ), but should not
feature prominently anywhere else.

Following one of your suggestions in the specific comments (on P14L2), we discuss the scientific problem at the start of our discussion **[P15 L1-30]** and the local management issues to a later part of the discussion **[P16 L19-30]**. Following the suggestion in the first general comment, the scientific problem is now also discussed at the start in the introduction **[P2 L1-7]** . Furthermore, the start of the abstract is changed. This has aided a lot in changing the focus of this paper **[P1 L11-12]** .

With regards to the assumption of the system being in steady-state (C), a considerable
portion of the paper is devoted to a comparison between dynamic (i.e., models
considering the palaeo-geographical evolution) and steady-state conditions (equivalent
simulations for the same geology). It is concluded that the dynamic models are
a better fit to the data. But given that the data set has severe limitations (there are
relatively few data points, and the quality of the available data is also not assured, at
least the paper does not describe the QC/QA procedures), one could wonder if that
is really a such strong criterion. I think a much stronger argument could be made by
looking at the time required to reach steady state. This information is not presented for
all model scenarios, but in line 28 on page 11 it is mentioned that it took 60 ka for the
B-model scenarios. Doesn't this automatically invalidate the steady-state assumption,

without having to perform anymore detailed comparisons between the dynamic models and their steady-state equivalents? I am not sure what time is required for the other simulations, but I am guessing it will be on the same order, except maybe for the most unrealistic representations of the delta's lithology. I would encourage the authors to present the timescale aspect in more detail and use it to build the argument against steady-state being a realistic assumption. Much of the detailed comparisons such as those presented in figures 9, 10 and 11 could then be omitted.

This would also have the benefit that the paper becomes more easy to follow, because as it stands, one quickly gets lost in the many different scenarios.

We think the suggestion to discuss the time to the steady state is good. This, in combination with changing the order in which we discuss the results and discussion, will improve the readability of the paper. Omitting much of the detailed comparisons though, would negate one of the strong points of this study though, namely that we ran a lot of scenarios for a complex 3D model. These scenarios provide a lot of unique information that is relevant to our discussion, as it makes this study more applicable to other deltas. So here we don't fully agree. Still, we have removed what was Fig. 10 in the initial submission, as we barely discussed it in the text. Furthermore, we removed Fig. 5 and accompanying text as the data is limited.

We have added the time to reach a steady state to "3.7 model evaluation", as this section appeared to be the most fitting. We added the following lines to the methods section **[P11 L10-12]:**

To assess the validity of the steady-state assumption, we checked for all equivalent steady-state models the time they reached a steady state. This time was determined by calculating the derivative of the fresh water volume over time. If this did not change more than 1E-04% of the total volume, we considered the model to have reached a steady state.

To the results section we added **[P13 L4-8]:**

All equivalent steady-state scenarios required at least several thousands of years to reach a steady state (Table 4) from an initial Pleistocene steady state. Most notable are the B-scenarios, where the hypersaline groundwater caused the system to respond very slowly, over tens of thousands of years, thus exceeding the extent of the Holocene. The shortest scenarios were the N-T scenarios, as they did not include HGw and due to the lack of clay layers the salt water did not experience any resistance during its flow upwards from its initial Pleistocene state to the Holocene steady state.

To the discussion we added **[P15 L6-9]:**
Our equivalent steady-state scenarios required at least 5500 years to reach a steady state (Table 4), a period in which already considerable changes occurred to the boundary conditions (Fig. 3). This increased to tens of thousands of years for the more complex models. We therefore doubt that using a steady-state approach with current boundary conditions results in a reasonable estimate of the current fresh-salt groundwater distribution for such a complex, large-scale system.

To the conclusion we added **[P18 L2-L7]**:

It was found that large timescales are involved, as steady-state model scenarios required at least 5500 years to reach equilibrium. Hence, none of the evaluated paleohydrogeological scenarios reached a steady state over the last 9000 years, meaning that the transient boundary conditions definitely had an influence on current groundwater salinity. Given the large range variation in delta-architectures analyses, we can conclude that steady-state models are not likely to result in realistic FGw distributions in deltaic areas.

To the abstract **[P1 L24-L25]**::

Furthermore, the time required to reach a steady state under current boundary conditions exceeded 5500 years for all scenarios.

The Results section
could start with what is now subsection 4.2, which could be expanded and/or merged
with subsection 4.4. This would give the reader a much better overview of the actual
processes before diving into the more detailed analysis of model performance and
freshwater volume.

We followed this suggestion as follows: We have started the results sections with a statement that for readability we first discuss the results of a few scenarios that were selected based on a later discussed model performance assessment **[P12 L2-L5]**. Then we start with the former section 4.2 **[P12 L6]**, which describes the spatial distribution. We continue with the former section 4.4, describing the contribution of several boundary conditions through time **[P12 L14]**. After that, we discuss the model performance **[P12 L23]**.

So the outline changed to:

4 Results

Based on a comparison with observations, five scenarios were selected that showed the best match with the observations. These are called the "acceptable" scenarios. To keep the results comprehensible, we start with discussing the results of these five selected scenarios through space (section 4.1) and time (section 4.2), before discussing this actual selection procedure (section 4.3).

Specific comments
Page 1

10 *lines 15-16, "observed by hydrogeochemists": No need to suggest a disciplinary*
bias

We have removed this in **[P1 L10-L14]**.

15 *line 17, "palaeo-reconstruction": This suggests your reconstruction itself is*
ancient. Choose better wording

This is indeed confusing. We changed this to "palaeohydrogeological reconstruction" (just as there are for
example palaeoclimate reconstructions) throughout the complete document.
20

line 18: Insert "a" between "using" and "state-of-the-art model"

We have corrected this in **[P1 L17]**

25 *line 23: You use both "sea water" and "seawater" in your manuscript. Pick*
one, and check for consistency.

We were indeed inconsistent and moreover grammatically incorrect in the original text. However, according to
the Cambridge Dictionary "saltwater" should be used as adjective (saltwater intrusion) and "salt water" as a
30 noun.

Noun:
https://dictionary.cambridge.org/dictionary/english/salt-water
Adjective:
35 https://dictionary.cambridge.org/dictionary/english/saltwater

5

We assumed here that sea water and seawater should be used the same as saltwater and salt water, and hence changed "sea water intrusion" to "seawater intrusion" and use "sea water" as a noun. Thus, to stay grammatically correct, we have to stay inconsistent. The same holds for "fresh water" and "freshwater".

5 *line 29: add "s" to "distribution"*

We corrected this in **[P1 L29]**.

Page 2

10

line 6: insert "that they were" between "indicated" and "pumping". More generally, the language usage needs some attention, the paper is generally well written, but every now and then it lacks some attention to detail. I will not focus on these issues from this point onward, but the authors should do a careful proofread of their revised
15 *manuscript to resolve them.*

We have corrected this specific point **[P4 L7]** and conducted extra proofreading (by a native speaker). We hope this improves our attention to detail sufficiently.

Page 4

20 *line 31: delete "hypersaline and", the fact that they are hypersaline should not be a reason to discard them. On the next page you talk about hypersaline groundwaters as well (or at least, salinities greater than seawater)*

This is indeed confusing, we have deleted these two words in **[P5 L9-10]**.
25

Page 5

line 1: I am not sure if this reasoning holds true. A 1000y old groundwater can still move appreciable distances over a couple of decades if it is near a large well
30 *field*

Good point, our reasoning is flawed here. We have added the different measurement dates as an extra source of the spatial variation in observed salinities.

35 Thus changing the list of potential causes to **[P5 L11-13]**:

This variability in measured salinity can be explained with 1) the different measurement depths, 2) different dates of the measurement campaigns, 3) heterogeneity in the hydraulic conductivity of the subsoil resulting in heterogeneous salt transport, and 4) heterogeneous evapoconcentration.
40

line 26: Some additional information is required here to explain the choices and rationale behind these 9 different lithological models

We have extended this section in **[P6 L11-12].**

The height of the clayey sediments determines how disconnected the deeper groundwater system is from the sea, and thus the ability of the system to preserve denser hypersaline groundwater in its aquifers (van Engelen et al., 2018). The hydraulic conductivity of the onshore-reaching clay layers is varied to get a first-order approximation of the effect of clay layers on regional groundwater flow. We assigned a continuous hydraulic conductivity to these clay layers, based on three different lithologies (in order of decreasing hydraulic conductivity): sand, fluvial clay and marine clay (Table 3). The rationale behind this is that small clay lenses have negligible effect on regional groundwater flow, thus are assigned a hydraulic conductivity of sand. Fluvial clay layers are assigned a hydraulic conductivity of the current confining Holocene clay layer, as this was deposited under fluvial conditions (Pennington et al., 2017). Marine clay layers present continuous layers of low conductive with a big influence on the regional groundwater flow and thus have the lowest hydraulic conductivity.

Page 6

line 5: Replace "its'" with "its"

We corrected this in line **[P6 L24]**

lines 10, 11: Not sure what you mean by "keeping memory locally".

We meant here that each subdomain is assigned its own private memory at the computational node, instead of fitting all subdomains into one big shared memory. We changed the sentence, to clear this up. **[P6 L29-30]**

using the Message Passing Interface (The MPI Forum, 1993) to exchange data between subdomains, where each subdomain has its own private memory assigned at a computational node.

line 15 a.f.: You need to include a map with the model area.

We have added this map as Appendix A1. **[P40 L1-5]**

lines 23, 24, "as the hills above this height have no important contribution to the groundwater flow": Without the aforementioned map it is hard to assess the validity of this statement. Where are these hills? How high are they? More importantly, what is the recharge and the water table elevation. The elevation of the hill itself is not so important, its hydrogeological characteristics much more so, of course...

This was indeed an offhand remark. We do not have actual groundwater tables of this specific area, but we know that the amount of rainfall is very low and that there is no surface water located there. This is supported by Geirnaert and Laeven (1992), who found that shallow groundwater in these locations is >5000 years,

meaning very limited recent recharge. We assumed that this leads to groundwater tables well below surface level such that loss through evaporation is negligible.

We added extended this sentence to provide support **[P7 L12-14]**:

The top of the NDA was clipped off above 20 m AMSL, as the hills above this height have no important contribution to the groundwater flow (Geirnaert and Laeven, 1992), because there is very limited rainfall in the south and no surface water here.

Page 7

line 3, "time slices": This sounds like what we would normally call a stress period in groundwater modelling. Why the confusing terminology? And why with a space in the title, and without in this sentence. Please pay more attention to detail.

This is indeed what we call a stress period in groundwater modelling. Still, we decided to stick with the term "time slices", following Delsman et al. (2014). "Stress period" makes sense from a modelling perspective, but can be confusing for non-modellers (like paleogeographers, hydrogeochemists), hence why we prefer this term. We added a sentence of explanation in **[P7 L23-25]**

A time slice is also known as "stress period" by groundwater modellers (Harbaugh, Arlen, 2005)

line 12: Replace "announced" with "at"

We corrected this. **[P7 L32]**

Next, time slice 3 covered the marine transgression, which was a period of rapid sea-level rise at the start of the Holocene

line 22, "100 day resistance" (better would be "a resistance of 100 days" BTW): How sensitive are the modelling results to this assumption? A single value is of course highly unrealistic, and citing studies from the Rhine-Meuse Delta does not provide any justification, because this parameter is just as uncertain and spatially (and even temporally at this timescale) variable there. But you've got to work with what you have, I understand that, but in the end, some sensitivity analysis is required to test to what extent the study outcomes might be affected by this modelling choice.

We have conducted a local sensitivity analysis of this resistance on one scenario. We initially wanted to vary the resistance with a factor 10 or 5 but this led to numerical convergence issues for the scenarios with a decreased resistance. Therefore, we stayed at a factor 2. To not repeat ourselves in this answer, the text we added as Appendix B should explain the rest **[P18 L26]**:

5

10

line 31: replace "acquiring" with "achieving"

Corrected in **[P8 L20]**.

15

Page 8

lines 3-6: Did you do this via SEAWAT's density options in the CHD package?
A range of 2-18 g TDS/L gives quite a range in density. What value did you adopt? And
20   *again, how sensitive is it?*

We used the RIV package for the lagoons. We estimated the mean salinities from the Flaux et al. (2013) *A 7500-year strontium isotope record from the northwestern Nile delta (Maryut lagoon, Egypt)* for each time slice. Of the total of seven time slices, the last four have lagoons, since lagoons started to form from 7.5 ka. To be
25   specific, the salinities assigned to time slices 3 to 7 are respectively: 18, 9, 4.5, 2.5 g TDS/l. Despite these decreasing salinities, we can still observe quite high salinities at the (former) locations in lagoons that fall in between 5-15 g TDS/l (see Figure 4), so past lagoonal salinities are still locally present in our model.
We do not, however, think that different salinities will have a large effect on our conclusions. Since the mixing zones are relatively small, changing the lagoonal salinity will only have a small effect on the fresh groundwater
30   volumes, as long as this salinity does not get lower than 1 g TDS/l. This is presumably more controlled by the location of the lagoons. Which model scenarios were deemed "acceptable" was in the end mainly assessed based on the location of hypersaline groundwater, so this will also not change strongly, since the density of HGw is much higher (up to 120 g/l) than that of the lagoons (up to 18 g/l). This brackish groundwater will exert limited force on the HGw.
35

We added the following text that is more specific here **[P8 L23-27]**:

time slices 3 to 7 are respectively: 18, 9, 4.5, 2.5 g TDS/l. These salinities show strong variation through time, because the inflow of the Nile varied through time.

Page 10

line 7: You could cite the following article here: Sanford, W.E. & Pope, J.P. Hydrogeol J (2010) 18: 73. https://doi.org/10.1007/s10040-009-0513-4

This is a very fitting suggestion, thanks. Also relevant for the discussion.

We added the reference to method section 3.7 **[P10 L29]** and added an extra sentence to the discussion **[P16 L1-2]**:

Similar conclusions were drawn by Sanford and Pope (2010) for the Eastern Shore of Virginia (USA), an area with a similar observation density.

line 10 a.f.:
This is somewhat hard to follow and it might be worth adding a sketch that illustrates the principle (could be in a supplementary document). Other authors may choose to adopt this methodology, hence it could be worthwhile doing this.

We have added a sketch as Fig A2 **[P41 L1-5]**.

Page 11

line 19: what do you mean with "behavioural"? And what is the justification for using 0.07?

With the term "behavioural" we mean the ability of models to reproduce certain patterns observed, following Beven and Binley (1992) *The future of distributed models: Model calibration and uncertainty prediction.* Whatever these patterns specifically are, is up to the researcher to decide. So, on second thought, perhaps the word "acceptable" captures the inherent arbitrariness of this decision better. The value of 0.07 was based on a visual inspection of the figure, as around that value there is a separation visible. The scenarios with $Md|\Lambda| < 0.07$, predict the location of the HGw with similar skill as to with which they predict the location of saline groundwater. These scenarios we call "acceptable" .

In the text **[P12 L29 – P13 L3]**:

More striking differences are observed in the hypersaline zone, where we observe a division around $\Lambda = 0.07$ into two groups. There are the scenarios with $Md|\Lambda| < 0.07$, that predict the location of the HGw with similar skill as to with which they predict the location of saline groundwater. We call these scenarios "acceptable". Specifically, these are the following five model scenarios: C-M-B-P, C-N-T-P, H-M-T-P, H-F-T-P, H-N-T-P. The other scenarios perform considerably worse in predicting the location of the HGw.

lines 19-21: Sentence does not seem to flow well due to a grammar error,
not sure what you are trying to say here.

This should be clear now **[P12 L29 – P13 L3]**.

Page 12

line 3: Not clear what you mean here with "behavioural"

Hopefully cleared up now with the changes made to the description of the model evaluation **[P12 L29 – P13 L3]**.
See response comments of P11L19.

line 4: Up until
this point the difference between fluvial and marine clay layers has not been explicitly
discussed.

Now expanded on in section 3.1 **[P6 L11-12]**. So see answer to comment for P5L26.

lines 7, 8: replace nondescript terms like "more 3D patterns are visible in
the salinities" and "partly has a conical" with more accurate descriptions

We are more specific here now **[P12 L10-13]**

Regardless of the differences between scenarios, in all realizations the fresh-salt interface roughly follows the coastline, except in the west where there is far extending seawater intrusion visible towards Wadi El Natrun (Fig. 1). Next to this depression, (former) lagoons are visible as shallow brackish zones and (former) dune areas are visible as freshwater lenses.

lines 24 a.f., "This table also shows...": I could not follow what you are trying to say here

This sentence could be more to the point. Changed the sentence to **[P13 L22-L24]**:

This table also shows that these parts of the model are also the most uncertain, since disregarding potential deep and offshore fresh groundwater volumes decreases the spread from 74% to 32%.

Page 14

line 2 a.f.: I think the point you want to make here is that you can come up with
multiple models that fit the observations equally well and yet, the volume of freshwater
varies a lot. What do you mean with "The variance in the results should also affect
management decisions."? I think the management decisions will not be based on the
total volume of freshwater, but on the possibility to be able to extract groundwater in
a particular region. See also general remark made before about the relevance of the

freshwater volume issue for this paper. I would not start the Discussion section with
this paragraph (see next point)

We have removed these sentences now, since it shifts the focus too much to the regional problem. We tried to say that policy makers and managers should take into mind the large uncertainty of groundwater models in these complex aquifers, as shown in this research, but since we do not provide any suggestion to do so this is a bit of a weak statement.

However, even though it has no effect on the changes made to this paper, we would like to express our different view on how management decisions are made in the Nile Delta though. Currently, there is an ongoing political discussion between Egypt and the upstream countries on the effects of the large dams that are being built. Discharge of the Nile will decrease and thus the volume of fresh groundwater that can serve as a "strategic stock" during periods of low-flow is becoming of increasing interest, from the perspective of delta-scale management.

Page 15

line 6: Start the Discussion section with this paragraph, it is much more relevant
to a broad readership than local aspects such as the discussion of the total
freshwater volume (also see general comment) and flow to Wadi El Natrun.

Good suggestion, we have brought this paragraph to the start of the Discussion section. **[P15 L2-30]**

line 32: Also include a discussion of the representation of free convection phenomena
in your model. The large grid size is prohibitive for an accurate process representation.
How confident are you that this does not harm the general conclusions drawn from the
model outcomes?

We are confident that the general conclusions are not harmed, since we have investigated the errors made in modelling with a crude resolution for the Nile Delta in a previous paper. This was published as an appendix, so easily overlooked. We have added a deliberation on this in the discussion of this paper in lines **[P17 L10-18]**. For your convenience we have added this appendix of our previous paper as supplementary information to this response to authors.

To the discussion we added:
In addition, for a proper physical representation of free convection, a finer grid is required. A coarse horizontal cell size results in a delay in the onset of free convection, while a coarse vertical cell size results in an onset of free convection even for situations that are expected to be stable (Kooi et al., 2000). van Engelen et al. (2018, Appendix D) investigated the errors caused by coarse model cells for the Nile Delta and found that especially the crude horizontal grid size had an influence. They found that this resulted in similar downward fluxes, but a delay in the onset of free convection. This effect, however, was negligible after ~50 years and thus dwarfed by the timescale of our time slices. The coarse vertical grid size was not an issue, since the marine transgression occurred over sand with a high hydraulic conductivity, meaning there is a very instable situation and free

convection has to occur. We thus think that the errors made in modelling free convection do not influence our conclusions.

Figure 1: It could just be because of the pdf, but the resolution is very poor. Figure
needs an inset showing the location of the area within Egypt/the Mediterranean region.
Add north arrow and scale bar! In this figure all areas outside the delta are white, better
to make the Mediterranean blue and the desert yellow(-ish).

Our maps indeed lacked these crucial map features, so therefore we added them to this figure **[P30 L1-2]**. The resolution presumably deteriorated in converting the vector file first to a .png to add in MS Word and consequently to a pdf. We plan on submitting the original vector files as pdfs for the final version, which should solve this.

Figure 2: Somewhere we will need a map with the model boundary. On this map you
will need to indicate which part of the model is shown here.

We added a dedicated figure to this (Fig. A1), since adding these features to Fig. 1 resulted in a too complicated, cluttered map.

Figures 4, 5: Take the reader by the hand here as the figures are complex. Explicitly
mention in the caption that a, b, c and d reflect the salinity classes, and explain the n
value.

We changed the caption to **[P35 L1-8]**:

Figure 6: Goodness of fit boxplots for all palaeohydrogeological reconstructions, binned into four salinity classes. The higher the value of Λ, the worse the fit. Codes indicate model scenarios (Table 1). TDS values are binned in the classes [0, 1], [1, 5], [5, 35], [35, 100] g TDS/L for respectively "fresh", "brackish", "saline", "hypersaline", these are respectively plotted in panels A, B, C, D. "n" indicates the amount of observations available in the TDS bin to evaluate model results with. Diamonds indicate outliers, defined as values separated from the first or third quartile at 1.5 times the interquartile range. When no box is plotted for a scenario, 75% of the measurements equal zero, which consequently causes the interquartile range to be zero, rendering every non-zero value an outlier.

Supplementary material Answer to Reviewer 1

Effects grid size on free convection

Context:

In this Appendix we conducted a grid convergence test for a simple sandbox model. "NDA-f model" means "Nile Delta Aquifer free convection model". There was also a model without free convection, and for this model we conducted a grid convergence test on the original model. Therefore, the text below starts with the words "different approach". In addition, we removed all less relevant parameters to this research from Table 1.

The text below is taken from Appendix D of :

van Engelen et al. (2018) *On the origins of hypersaline groundwater in the Nile Delta aquifer*

For the NDA-f model, a different approach was taken, to assess the effect of a wider range of Δx, since this dimension had the coarsest resolution (1 km). We modeled a simple 800 m thick sandbox with a hypersaline lake on top. This allowed us to stretch the domain as much as necessary in the horizontal direction. The chosen parameters for this model are the same as in Table 1. The amount of cells was kept at 100 in the horizontal direction. Ξ is the normalized fluid mass increase, which is calculated by dividing the increase in fluid mass across the top boundary by the maximum mass storage increase:

$$
\text{(D.1)} \quad \Xi \; = \; \frac{\iint \frac{Q}{\Delta x} \rho \mathrm{d}x \mathrm{d}t}{V_{tot}\left((\rho_{max} - \rho_0)n + \frac{1}{2} S_f \frac{\rho_{max} - \rho_0}{\rho_0} z_{max}\rho_{max} \right)}
$$

where Q is the volumetric flux through a boundary cell (m²/d), V_{tot} is the total volume of the model domain (m²), ρ_{max} is the concentration at the top boundary which was set at 1078 kg/m³, ρ_0 was the initial density of the groundwater in the domain (1000 kg/m³), S_f is the specific storage in terms of fresh water head (d⁻¹), z_{max} the depth of the domain (m), and n the porosity (-). For the test with Δx, Δz was kept at 10 m; in testing Δz, Δx was kept at 10 m. As the density at the top boundary of the model was perturbed randomly, 10 simulations were started for each discretization, of which the minimum, maximum and mean of Ξ are plotted. It can be seen in figure D.2 that there is some spread in Ξ, which is caused both by integration errors and differences between realizations. Furthermore, a larger Δx causes a delay in the onset of free convection and causes a slower downward movement of fingers. However, all errors are unimportant on timescales larger than ~100 years, and thus also on our timescales of interest, which is over 1000 years.

Table 1. Parameters for the NDA-f model. We are purposely inconsistent with units, as it allows for easier comparison with values found in the literature.

| Parameter	Description	Value	Unit
$K_{h, sand}$	Horizontal hydraulic conductivity sand. Compaction case and free convection case	75	m/d
$\dfrac{K_h}{K_v}$	Anisotropy	10	-
α_l, α_t, α_v	Longitudinal, transversal and vertical dispersion length	10, 1, 0.1	m
n_e	Effective porosity	0.10	-
$\dfrac{\Delta h_{riv}}{\Delta x}$	River gradient	9.375e-2	m/km
S_f	Specific storage in terms of fresh water head	1e-5	1/m
$\dfrac{\partial \rho}{\partial C}$	Slope linear equation of state	0.71	(kg/m³)/(g/l)
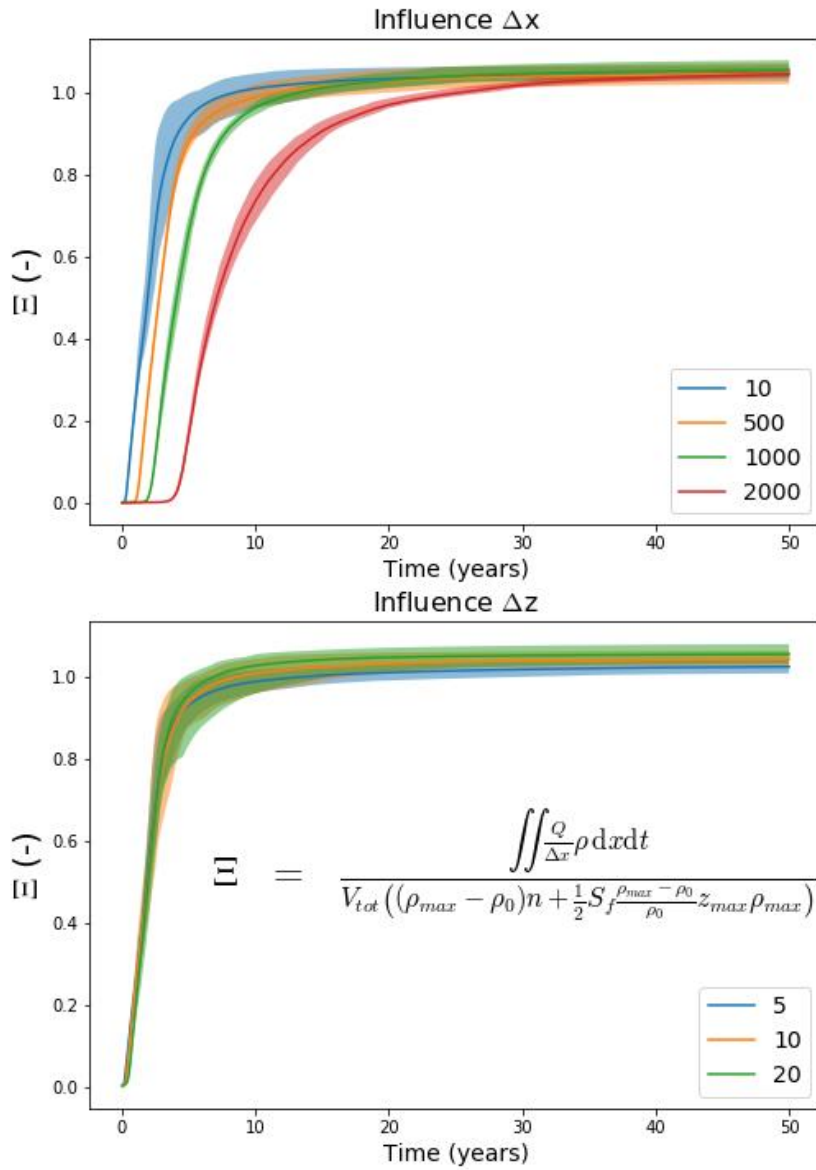
Figure D.2: Results of the test of the errors that are introduced in the free convection model by the large grid size: a. shows the influence of Δx and b. shows the influence of Δz.

Answer to Reviewer 2

We thank reviewer 2 for his/her thoughtful review that greatly helped to improve our paper. In this document, we have entered your comments in italics, added our response as regular text, and subsequently added suggested changes to the paper in red. After our responses you can find the suggested changes made to the
5   text in red. Line numbers are written as "**[P…L…]**", indicating page number, line number in the <u>no markup</u> document.

General comments

General comments: This paper describes numerical modelling used to reconstruct the salinity distribution in the Nile Delta aquifer over the past 32 ka. The model simulations
10   *include the Holocene marine transgression and various scenarios relating to possible aquifer lithology, and sources of hypersaline seawater. Through consideration of palaeo-environments, the authors aim to better represent the Nile Delta aquifer and an observed freshening. The authors state that simulation of the freshening of the aquifer has not been possible previously using realistic model parameters and steady-state*
15   *modelling. The objectives of this modelling study are stated as being (p4, L8) to: 1) investigate the physical plausibility of the Holocene transgression hypothesis for the Nile Delta; 2) investigate the influence of the uncertain geology; 3) provide estimates of the present-day fresh groundwater [FGW] volumes; 4) assess the importance of using palaeo-reconstructions compared to less expensive steady-state modelling. These*
20   *objectives have been addressed within the study and the description of how this was achieved is generally clearly detailed. There are not very many palaeo-reconstruction type studies looking at saltwater-freshwater dynamics in coastal aquifers and therefore there is some novelty in this contribution. However the significance is not clear. The study relates specifically to the Nile Delta aquifer and the more generalised extension*
25   *of the science is not apparent. As such the study is primarily a case study, albeit with some interesting insights for other similar deltaic systems. For this manuscript to be suitable for publication in HESS the authors need to indicate how the study provides a substantial contribution, and is not just a case study.*

30   We think with the help of also the general comments of Reviewer 1, we managed to change the focus of this paper from the regional problem to the scientific problem. We start the introduction **[P2 L1-7]** and abstract now with the general scientific problem. Furthermore, we discuss the scientific problem now at the start of our discussion **[P15 L1-30]** and the local management issues are kept as a later part of the discussion **[P16 L19-30]**. Furthermore, the start of the abstract is changed **[P1 L11-12]**. This adaptation should more clearly aid in
35   stressing the scientific significance of this research.
Furthermore, given the wide range of lithological scenarios, several of the findings are also applicable to other deltas: The severe decrease in FGw volume after the marine transgression (at the start of the Holocene), the large variation between the different palaeohydrogeological reconstructions compared to steady-state scenarios, and the lithologies that favor offshore fresh groundwater volumes in deltas.

To show this stress this more, we added this to the conclusion **[P18 L6-L9]**:

Given the large range variation in delta-architectures analyses, we can conclude that steady-state models are not likely to result in realistic FGw distributions in deltaic areas. Our results also show that the occurrence of past marine transgressions constitute a valid hypothesis explaining the occurrence of the extensive saline zone land inward.

Specific comments

Specific comments: The use of iMOD-SEAWAT has presumably assisted in reducing run times and making modelling of the various scenarios used in the study tractable. The paper would benefit from more information describing the run times and benefits of using iMOD-SEAWAT. I expect that other researchers will find this interesting for similar palaeo-reconstruction type work.

The added benefits of this code were already described (namely faster runtimes on computers with multiple CPUs), we however could add some information on run times **[P6 L32 – P7 L1]**.

Simulations were conducted on the Dutch national computational cluster Cartesius (Surfsara, 2014), using Intel Xeon E5-2690 v3 processors. With this new code the model scenarios had a wall clock time ranging of 44 hours to 108 hours on 48 cores, depending on model complexity.

The study involved numerous scenarios and I found many of these confusing as I read through the paper. The section describing the coding of the scenarios needs improvement. For example, I was not able to discern the meaning of P in the scenario H-F-B-P (p10, L1), despite that this was used as part of the explanation of how the coding system works.

We can understand that this might be confusing, as it was very difficult to come up with an appropriate coding system. To reduce the confusion we expanded the example and added each feature-member combination behind each piece of the sentence where it was specified **[P10 L19-23]**:

For each feature, the corresponding letters in table 1 are converted to a code as follows: *{sea}-{clayer}-{prov}-{temp}*. For example, the palaeohydrogeological reconstruction (*temp*: P) of an aquifer with a half-open sea connection (*sea*: H), fluvial horizontal clay layers (*clayer*: F), and HGw seeping in from the bottom (*prov*: B) gets the following code: H-F-B-P. Its equivalent steady-state model is attributed the code H-F-B-S.

Additionally, the term 'behavioural' model scenarios (p. 12, L1) seems odd and I suggest changing.

Copying from the response to comment to P11L19 by Reviewer 1:

With the term "behavioural" we mean the ability of models to reproduce certain patterns observed, following Beven and Binley (1992) *The future of distributed models: Model calibration and uncertainty prediction.* Whatever these patterns specifically are, is up to the researcher to decide. So, on second thought, perhaps the word "acceptable" captures the inherent arbitrariness of this decision better. Hence, we have changed the word 5 "behavioural" to "acceptable" in the paper.

In the text **[P12 L29 – P13 L3]**:

More striking differences are observed in the hypersaline zone, where we observe a division around $\Lambda = 0.07$ into two groups. There are the scenarios with $\text{Md}|\Lambda| < 0.07$, that predict the location of the HGw with similar skill as to with which they predict the location of saline groundwater. We call these scenarios "acceptable".
10 Specifically, these are the following five model scenarios: C-M-B-P, C-N-T-P, H-M-T-P, H-F-T-P, H-N-T-P. The other scenarios perform considerably worse in predicting the location of the HGw.

19

A three-dimensional palaeo~~hydrogeological~~ -reconstruction of the groundwater salinity distribution in the Nile Delta Aquifer

Joeri van Engelen[1,2], Jarno Verkaik[1,2], Jude King[1,2], Eman R. Nofal[3], Marc F.P. Bierkens[1,2], Gualbert H.P. Oude Essink[1,2]

[1]Department of Physical Geography, Utrecht University, Utrecht, 3584 CB, The Netherlands
[2]Unit Subsurface and Groundwater systems, Deltares, Utrecht, 3584 BK, The Netherlands
[3]Research Institute for Groundwater, National Water Research Center (El-Kanater El-Khairiya), Egypt

Correspondence to: Joeri van Engelen (joeri.vanengelen@deltares.nl)

Abstract.

~~The Nile Delta is an important agricultural area with a fast-growing population. Though traditionally irrigated with surface water, the delta increasingly relies on groundwater. However, saline groundwater extends far land inward, rendering groundwater close to the coastal zone useless for consumption or agriculture. To aid groundwater management decisions,~~ ~~The~~ Holocene marine transgression~~s are~~is often put forward~~used~~ to explain observed groundwater salinities that extend far land inwards in deltas. This hypothesis was also ~~brought forward~~proposed in -the hydrogeochemical literature to explain the large land-inward extent of saline groundwater in the Nile Delta. Despite that, ~~hydrogeologists reconstructed~~ the variable-density groundwater models previously built for the area used very large dispersivities to reconstruct this saline and brackish groundwater zone~~using variable density groundwater models with very large dispersivities~~. However, this approach cannot explain the observed freshening of this zone. ~~as observed by hydrogeochemists, who hypothesize that the coastal saline zone is the effect of the Holocene transgression.~~ Here, we investigated physical plausibility of the Holocene transgression -~~this~~ hypothesis to explain observed salinities by conducting a palaeohydrogeological- reconstruction of groundwater salinity for the last 32 ka with a complex 3D variable-density groundwater flow model, using a state-of-the-art version of the computer code SEAWAT ~~model code~~that allows for parallel computation. Several scenarios with different lithologies and hypersaline groundwater provenances were simulated, of which five were selected that showed the best match ~~with the~~against observations. Amongst these selections, total fresh-water volumes varied strongly, ranging from 1526 to 2659 km^3, mainly due to uncertainties in the lithology offshore and at larger depths. This range is smaller (1511-1989 km^3) when we only consider the volumes of onshore fresh groundwater within 300 m depth. Regardless of ~~the~~this~~e~~ variation~~nce~~~~the variance~~, in all cases the total volume of hypersaline groundwater exceeded that of sea water. We also show that during the last 32 ka, the total fresh~~water~~ ~~groundwater~~ volumes significantly declined, with a factor ranging from 1.9 to 5.4, due to the rising sea-level. ~~Compared to a steady state solution with present day boundary conditions, the palaeo reconstruction improved our validation for the saline zone (5 g/L – 35 g/L TDS).~~ Furthermore, the time required to reach steady state under current boundary conditions exceeded 5500 years for all scenarios. ~~Also~~Finally, under highly permeable conditions the marine transgression simulated with the palaeohydrogeological- reconstruction led to a steeper fresh-salt interface compared to its

steady-state equivalent, while low permeable clay layers allowed for the preservation of volumes of fresh groundwater. This shows that long-term transient simulations are needed when estimating present-day fresh-salt groundwater distributions in large deltas. The insights of this study are also applicable to other major deltaic areas, given the wide-range of lithological model scenarios used in this study and since many deltas also experienced a Holocene marine transgression.

5

1 Introduction

Palaeohydrogeological conditions have influenced groundwater quality in the majority of large-scale groundwater systems, since groundwater below 250 m depth is dominated by groundwater with an age of over 12 ka (Jasechko et al., 2017). These conditions can especially be found in deltaic areas, where the effects of marine transgressions are often still observed in groundwater salinities (Larsen et al., 2017). More specifically, their low elevation allowed for far reaching marine transgressions, leading to a large vertical influx of sea water, and hampered subsequent flushing with fresh water after the marine regression. This hypothesis is supported by hydrogeochemical research in several deltas (e.g. Colombani et al., 2017; Fass et al., 2007; Faye et al., 2005; Manzano et al., 2001; Wang and Jiao, 2012). The physical justification for this hypothesis, however, often still has to be tested, with a few notable exceptions (Delsman et al., 2014; Larsen et al., 2017; Van Pham et al., 2019), and can be provided by palaeohydrogeologic modelling, which has recently provided important insights for several cases. Gossel et al. (2010) created a large-scale 3D variable-density groundwater model of the Nubian Aquifer System and showed that seawater intrusion has occurred since the Pleistocene Lowstand towards the Qattara Depression (North-West Egypt). Later, Voss and Soliman (2013) showed with a parsimonious 3D model of the same groundwater system that groundwater tables are naturally lowering during the Holocene, since they receive limited recharge and are drained into oases or sabkhas. Moreover, the authors used an inventive validation method by comparing the position of discharge areas in the model with a dataset of oases or sabkha locations. Delsman et al. (2014) conducted a detailed palaeohydrogeological reconstruction of the last 8.5 ka over a cross-section in the Netherlands to show that the system has never reached a steady state. They showed that the Holocene transgression caused substantial seawater intrusion, from which the system is still recovering. Using a combination of geophysics and 2D numerical models, Larsen et al. (2017) showed that during the Holocene transgression sea water preferentially intruded in former river branches in the Red River Delta, Vietnam. Van Pham et al. (2019) showed that most of the fresh groundwater in the Mekong Delta (Vietnam) was likely recharged during the Pleistocene and preserved by the Holocene clay cap. Despite being in a humid climate, recharge to the deeper Mekong Delta groundwater system is very limited and freshwater volumes are still declining naturally.

~~The Nile Delta is vital for Egypt. It is an important bread basket, as it has been through history (Dermody et al., 2014), since it is the main area suitable for agriculture in an otherwise water scarce region (WRI, 2008). Furthermore, the area is densely populated (Higgins, 2016), and the population is only expected to increase over the coming decades (World Bank, 2018). Although the area is traditionally irrigated with surface water, agriculture is increasingly relying on groundwater (Barrocu and Dahab, 2010). For instance, a majority of interviewed farmers in the central delta indicated pumping groundwater almost continuously during the summer (El Agha et al., 2017). This groundwater pumping will increase in the future, as large dams are planned or built upstream, e.g. the Grand Ethiopian Renaissance Dam (GERD), which will reduce the discharge of the~~

Nile and will consequently hamper flushing of the irrigation system. This is expected to further deteriorate the quality of surface water in the irrigation system, which currently already is in a critical state; the water that currently reaches the sea is mostly saline and highly polluted (Stanley and Clemente, 2017). For example, the nitrate concentrations of the water inflow into the Manzala lagoon have increased by a factor 4.5 in the past 20-25 years, as an effect of the intensified agriculture and

5 the reduced inflow of river water (Rasmussen et al., 2009). Given the increasing stress on the groundwater system, it is important to assess the status of the current and future fresh groundwater volumes.

The Nile Delta Aquifer (NDA) is very thick, especially near the coast, where thicknesses up to 1000 m are observed (Sestini, 1989). For reference, only 0.3% of the coastal aquifers in the world is estimated to exceed this thickness (Zamrsky et al., 2018). Despite its large groundwater volume, the amount of groundwater suitable for domestic, agricultural or industrial

10 water supply is limited. A considerable fraction of the groundwater volume is saline (Kashef, 1983; Sefelnasr and Sherif, 2014) or even hypersaline (van Engelen et al., 2018), and unsuitable for most uses. Saline groundwater can be detrimental to Egypt's already critical food supply. For example, when increased irrigation caused the level of the brackish groundwater to rise in Tahrir, Cairo, wheat yields reduced by 41% within a four year period as soil salinity increased (Biswas, 1993). The saline groundwater issue has led to a continuous line of research into fresh and saline groundwater occurrence in the

15 NDA. The Nile Delta is one of the deltas where the physical grounds of the marine transgression hypothesis has s not been tested by physical analysis yet, even though there aredespite severe problems with groundwater salinity (section 2.1). This saline groundwater issue has led to a continuous line of research into fresh and saline groundwater occurrences in the Nile Delta Aquifer (NDA). These studies can be divided into hydrogeochemical and groundwater mechanical studies. Starting with the former, Geirnaert and Laeven (1992) provided the first conceptual palaeohydrological model (dating back to 20 ka

20 (=20,000 years BP)). They used groundwater dating to show that the shallow groundwater was likely recharged as fresh river water around 3.5 ka. Barrocu and Dahab (2010) extended this conceptual model to up to 180 ka and concluded that the saline groundwater in the north is old connate groundwater trapped in the Early Holocene, since it is freshening. Geriesh et al. (2015) further supported this conceptual model with additional measurements.

As an example of a study based on groundwater mechanics, Kashef (1983) used the Ghyben-Herzberg relationship to show

25 that the salt seawater wedge reached far inland, despite that the aquifer was gaining fresh water at that time. This was further confirmed by the 2D numerical model constructed by Sherif et al. (1988), who concluded that the width of the dispersion zone may be considerable. Sefelnasr and Sherif (2014) showed the sensitivity of the area to sea-level rise with a 3D numerical model, as the low topography of the Nile Delta allowed for a far reaching land surface inundation that can be detrimental to fresh groundwater water volumes (Ketabchi et al., 2016; Kooi et al., 2000). Mabrouk et al. (2018) showed that

30 groundwater extraction is a larger threat to freshwater groundwater volumes than sea-level rise, under the assumption that land surface inundation with sea water is prevented. Van Engelen et al. (2018) investigated the origins of hypersaline groundwater that is found starting from 400 m depth, which was overlooked in preceding studies. They tested four hypotheses of which two remained valid: 1) free convection of hypersaline groundwater formed in the Late-Pleistocene

coastal sabkha deposits and 2) upward compaction-induced flow of hypersaline groundwater, formed during the Messinian Salinity Crisis. Moreover, they showed the influence of ~~the~~ uncertain lithology on the groundwater salinity distribution.

When comparing the results of the hydrogeochemical studies with those of the groundwater mechanical studies, a discrepancy can be observed that is difficult to reconcile. Many of the groundwater-mechanical studies that use physically-based variable-density groundwater flow models tend to neglect the influences of palaeohydrogeological conditions that were inferred by hydrogeochemists. Instead, they attribute the large brackish zone in the groundwater system (Fig. 1) to a large influence of hydrodynamic dispersion (Sefelnasr and Sherif, 2014; Sherif et al., 1988), with longitudinal dispersivities exceeding 70 m. The hydrogeochemical studies, however, attribute it to the Holocene transgression, since a large fraction of the brackish zone is presumably freshening (Geriesh et al., 2015; Laeven, 1991) and the extent of this brackish zone coincides with palaeo-shorelines (Stanley and Warne, 1993) (Fig. 1). Yet a variable-density groundwater flow model with present-day boundary conditions and a large amount of hydrodynamic dispersion cannot support the freshening of the brackish zone, since hydrodynamic dispersion would only lead to progressive salinization. In addition, currently no reliable field data exists that supports ~~—~~ longitudinal dispersivities beyond 10 m (Zech et al., 2015). This discrepancy in conceptualization still exists, ~~—despite~~ ~~that~~ the large influence of palaeohydrogeologcial conditions on salinities ~~was~~ previously shown by applying variable-density groundwater flow models ~~for~~ to other cases (Kooi et al., 2000; Meisler et al., 1984).

The lack of including palaeohydrogeological development ~~conditions~~ ~~s reconstructions~~ in variable-density groundwater flow and coupled salt transport models is understandable as it requires vast amounts of computational resources, especially when 3D reconstructions are considered. ~~Yet these reconstructions can deliver important insights on the groundwater system, as follows from some recent examples. Gossel et al. (2010) created a large scale 3D variable-density groundwater model of the Nubian Aquifer System and showed that sea water intrusion occurred already during the Pleistocene Lowstand towards the Qattara Depression, which lays roughly 200 km West from the Nile Delta. Later, Voss and Soliman (2013) showed with a parsimonious 3D model of the same groundwater system that groundwater tables are naturally lowering during the Holocene, since they receive limited recharge and are drained in oases or sabkhas. Moreover, the authors used an inventive validation method by comparing the position of discharge areas in the model with a dataset on oases or sabkha locations. Delsman et al. (2014) conducted a detailed palaeo reconstruction over a cross-section in the Netherlands and showed that the system has never reached steady-state. They showed that the Holocene transgression caused substantial salt water intrusion, from which the system is still recovering. Larsen et al. (2017) showed with a combination of geophysics and 2D numerical models that during the Holocene transgression salt water preferentially intruded in former river branches in the Red River Delta, Vietnam. Van Pham et al. (2019) showed that most of the fresh groundwater in the Mekong Delta Vietnam was likely recharged during the Pleistocene and preserved by the Holocene clay cap. Despite being in a humid climate, recharge to the deeper Mekong Delta groundwater system is very limited and fresh groundwater volumes are still declining naturally.~~

~~Yet, given~~ The ~~Given the stress on the groundwater system, the recent impetus of numerical palaeo reconstructions, and~~ ~~the~~ availability of a newly developed model code that allows for high-performance computing (Verkaik et al., 2017), ~~we~~

24

~~deem~~has now made it ~~necessary and~~ possible to create the first numerical palaeohydrogeological -reconstruction of the Nile Delta Aquifer in 3D. Therefore, in this ~~research~~ study we have constructed a variable-density groundwater flow and coupled salt transport model of the Nile Delta Aquifer and use it for a palaeohydrogeological -reconstruction of Nile Delta Aquifer salinity over the last 32 ka. Specifically, we use the model to 1) investigate the physical plausibility of the Holocene transgression hypothesis for the Nile Delta; 2) investigate the influence of the uncertain geology; 3) provide volume estimates of ~~the volume of the~~ present-day fresh groundwater [FGW]~~ volumes~~; 4) assess the relative importance of using palaeohydrogeological -reconstructions ~~compared to~~against -less expensive steady-state modelling.

2 Area description

2.1 Area relevance and vulnerability

The Nile Delta is vital for Egypt. It is an important bread basket, as it has been through history (Dermody et al., 2014), since it is the main area suitable for agriculture in an otherwise water-scarce region (WRI, 2008). Furthermore, the area is densely populated (Higgins, 2016), and the population is only expected to increase over the coming decades (World Bank, 2018). Although the area is traditionally irrigated with surface water, agriculture ~~is~~ increasingly relies~~ying~~ on groundwater (Barrocu and Dahab, 2010). For instance, a majority of interviewed farmers in the central delta indicated that they are pumping groundwater almost continuously during the summer (El-Agha et al., 2017). This groundwater pumping will increase in the future, as large dams are planned or built upstream, e.g. the Grand Ethiopian Renaissance Dam (GERD), which will reduce the discharge of the Nile and will consequently hamper flushing of the irrigation system. This is expected to further deteriorate the quality of surface water in the irrigation system, which currently already is in a critical state; the water that currently reaches the sea is mostly saline and highly polluted (Stanley and Clemente, 2017). For example, the nitrate concentrations of the water inflow into the Manzala lagoon have increased by a factor 4.5 in the past 20-25 years, as an effect of the intensified agriculture and the reduced inflow of river water (Rasmussen et al., 2009). Given the increasing stress on the groundwater system, it is important to assess the status of the current and future freshwater volumes.

Despite the large groundwater volume of the Nile Delta Aquifer, the amount of groundwater suitable for domestic, agricultural or industrial water supply is limited. A considerable fraction of the groundwater volume is saline (Kashef, 1983; Sefelnasr and Sherif, 2014) or even hypersaline (van Engelen et al., 2018), and unsuitable for most uses. Saline groundwater can be detrimental to Egypt's already critical food supply. For example, when increased irrigation caused the level of the brackish groundwater to rise in Tahrir, Cairo, wheat yields reduced by 41% within a four year period as soil salinity increased (Biswas, 1993).

2.~~1~~2 Hydrogeology

The Nile Delta Aquifer (NDA) is very thick, reaching up to 1 km depth at the coast (Sestini, 1989). For reference, only 0.3% of the coastal aquifers in the world ~~is~~are estimated to exceed this thickness (Zamrsky et al., 2018). The NDA ~~The Nile Delta~~

25

~~Aquifer reaches up to 1 km depth and~~ consists of the Late-Pliocene El Wastani formation and the Pleistocene Mit Ghamr formation. These ~~formations~~ both contain mainly coarse sands with a few clay intercalations (Sestini, 1989). It is capped by the Holocene Bilqas formation, which has recently been mapped extensively (Pennington et al., 2017). The Pliocene Kafr El Sheikh formation is generally assumed to be the hydrogeological base as it is 1.5 km thick and consists mainly of marine clays, though it is possible that there is compaction-driven salt transport through this layer (van Engelen et al., 2018). Hydrogeological research dealing with this area generally assumed that the connection between aquifer and sea is completely open, but there exists ample evidence against this assumption from both seismics (Abdel-Fattah, 2014; Abdel Aal et al., 2000; Samuel et al., 2003) as well as from the fact that offshore borelogs contain more clay than sands (Salem et al., 2016). These observations can be explained by the fact that in the seaward direction marine clays are more commonly observed in boreholes. These marine clays can even form large vertical clay barriers when deposited during aggradational or retrogradational phases (Nichols, 2009) with a major effect on the groundwater system (van Engelen et al., 2018).

2.~~2~~ 3 Groundwater salinity

Figure 1 shows all groundwater Total Dissolved Solids (TDS) measurements above 250 m depth that we could find in ~~the~~ literature (Geriesh et al., 2015; Laeven, 1991; Salem and El-Bayumy, 2016), combined with data of the Research Institute for Groundwater (RIGW) (Nofal et al., 2015). Other publications are not included as 1) these did not report measurement depths, 2) did not report measurement locations, or 3) only showed isohalines, from which the actual measurement locations are impossible to discern. The few available salinity measurements deeper than 250 m are not shown here as these are ~~hypersaline and~~ likely not influenced by the Holocene transgression (van Engelen et al., 2018). ~~The different dates of the measurement campaigns in Fig. 1 matter little as groundwater ages exceeding 1000 years have been determined at similar depths at most locations (Geirnaert and Laeven, 1992). Despite this, the~~The ~~shown~~depicted measurements ~~still show~~highlight considerable spatial variation, especially in the brackish zone. This variability in measured salinity can be explained with 1) the different measurement depths, 2) different dates of the measurement campaigns, ~~2~~3) heterogeneity in the hydraulic conductivity of the subsoil resulting in heterogeneous salt transport, and ~~3~~4) heterogeneous evapoconcentration. The latter is inferred from the observed salinities that exceed the salinity of sea water (35 g TDS/L), which presumably are pockets of evapoconcentrated Holocene groundwater (Diab et al., 1997). Despite the spatial variability in salinity, we observe a trend: the extent of the brackish zone seems to conform in the east to the coastline during the maximum transgression at 8 ka. Westwards, the extent of this zone can be explained with the location of the former Maryut lagoon, which had several periods where its salinities approached that of sea water (Flaux et al., 2013).

It is important to note that our dataset is biased towards areas with low soil salinities, as measurements have been preferentially taken in the most productive agricultural areas and thus of higher economic interest to monitor. This becomes ~~evident~~apparent from a comparison with soil salinity maps (Kubota et al., 2017) and ~~. This~~is particularly evident from the data gathered by Salem et al. (2016b). These authors conducted research in a specific coastal area with low soil salinities, which is ~~the~~an area that has predominantly consisted of coastal dunes from 7.5 ka until present (Sestini, 1989; Stanley and

Warne, 1993), thus providing the hydrogeological circumstances for the development of a freshwater lens. The other, former dune areas (Fig. 3) were either eroded naturally or removed by humans (El Banna, 2004; Malm and Esmailian, 2013; Stanley and Clemente, 2014) and have high soil salinities, presumably causing former freshwater lenses to salinize. Given this bias in our dataset towards fresh groundwater, the extent of the saline groundwater problem is likely underexpressed in Fig. 1.

5

3 Methods

3.1 Lithological model

A dataset of 159 borelogs was compiled ~~of~~ using georeferenced data from~~in~~ different literature sources (Coleman et al., 1981; Nofal et al., 2016; Salem et al., 2016; Summerhayes et al., 1978). These borelogs were used to constrain a 3D

5 lithological model using SKUA-GOCAD's implicit modelling engine (Paradigm, 2017). There are two main uncertainties in the lithological model thus constructed. The first ~~uncertainty~~ pertains to the question of whether the continental slope is covered completely with low-permeable clayey sediments or not (section 2.2) and ~~the~~ second whether the clay layers observed in the NDA are continuous, forming low permeable structures, or are disconnected with only a limited effect on regional groundwater flow. To account for these uncertainties, nine different lithological model scenarios were constructed

10 (Fig. 2), where we varied the height of the clayey sediments on the continental slope and the hydraulic conductivity of the onshore-reaching clay layers.

The height of the clayey sediments determines how disconnected the deeper groundwater system is from the sea, and thus the ability of the system to preserve denser hypersaline groundwater in its aquifers (van Engelen et al., 2018). The hydraulic conductivity of the onshore-reaching clay layers is varied to get a first-order approximation of the effect of clay layers on

15 regional groundwater flow. We assigned a continuous hydraulic conductivity to these clay layers, based on three different lithologies (in order of decreasing hydraulic conductivity): sand, fluvial clay and marine clay (Table 3). The rationale behind this is that small clay lenses have negligible effect on regional groundwater flow, thus are assigned a hydraulic conductivity of sand. Fluvial clay layers are assigned a hydraulic conductivity of the current confining Holocene clay layer, as this was deposited under fluvial conditions (Pennington et al., 2017). Marine clay layers present continuous layers of low

20 conductivity~~e~~ with a big influence on the regional groundwater flow and thus have the lowest hydraulic conductivity.

28

3.2 Model code and computational resources

We applied the newly developed iMOD-SEAWAT code (Verkaik et al., 2017) which is based on SEAWAT (Langevin et al., 2008), but supports distributed memory parallelization that allows for a significant reduction in computation times. The SEAWAT code is the industry standard for solving variable-density groundwater and coupled solute transport problems, therefore the reader is referred to its~~²~~ manuals for an extensive explanation (Guo and Langevin, 2002; Langevin et al., 2003, 2008). The main improvement of the iMOD-SEAWAT code is that it replaces the original solver packages for variable-density groundwater flow and solute transport (respectively PCG and GCG) with the Parallel Krylov Solver package (PKS). The PKS linear solver is largely based on the unstructured PCGU-solver for MODFLOW-USG (Panday et al., 2013) and solves the global linear system of equations with the additive Schwarz parallel preconditioner (Dolean et al., 2015; Smith et al., 1996) using the Message Passing Interface (The MPI Forum, 1993) to exchange data between subdomains, ~~while keeping memory locally~~ where each subdomain has its own private memory assigned at a computational node. The variable-density flow problem and the solute transport problem are solved in parallel using respectively the additive Schwarz preconditioned Conjugate Gradient and BiConjugate Gradient Stabilize linear accelerators (Barrett et al., 2006; Golub and Van Loan, 1996). Simulations were conducted on the Dutch national computational cluster Cartesius (Surfsara, 2014), using Intel Xeon E5-2690 v3 processors. With this new code the model scenarios had a wall clock time ranging of 44 hours to 108 hours on 48 cores, depending on model complexity.

3.3 Model domain and numerical discretization

The study area encompasses the complete deltaic plain, the coastal shelf, coastal slope, and the Eastern desert and Western desert fringes (near the Suez Canal and Wadi El Natrun area). Its Western boundary is close to Alexandria, its Eastern the Suez Canal, its Southern near Cairo, and the Northern boundary laid 70 km offshore, resulting in a rectangle of 240 km meridional (W-E) distance and of nearly 260 km zonal (S-N) distance (see also Fig. A1). This area was discretized into 1 by 1 km cells. To determine the vertical dimension, borelogs and bathymetrical/topographical data were used to respectively determine the bottom and top of the NDA. For the topography of our model, we used a combination of the Shuttle Radar Topography Mission (NASA, 2014) for the onshore part and the General Bathymetric Chart of the Oceans (GEBCO, 2014) for the offshore part. The top of the NDA was clipped off above 20 m AMSL, as the hills above this height have no important contribution to the groundwater flow (Geirnaert and Laeven, 1992), because there is very limited rainfall in the south and no surface water there. The model domain was discretized into 35 model layers, from top to bottom: 21 model layers of 20 m thickness, 10 model layers of 40 m thickness, and 4 model layers of 50 m thickness. In total, the model has more than 2 million cells. We simulated a time period of 32,000 years, covering the Late-Pleistocene and Holocene. This period was chosen as 1) no palaeogeographic maps were available of the preceding period, and 2) this research only focused on the potential effects of the Holocene transgression, not previous Pleistocene transgressions.

3.4 Boundary conditions

3.4.1 Time slices

The palaeohydrogeological -reconstruction consisted of several consecutive time slices (Fig. 3), following Delsman et al. (2014), in which boundary conditions were kept constant. A time slice is also known as "stress period" by groundwater
5   modellers (Harbaugh, Arlen, 2005). Each time slice consisted of a palaeogeographic map (Pennington et al., 2017; Stanley and Warne, 1993) that defines the location of geographical classes. Boundary conditions were assigned to five of these classes, namely sea, lagoon, sabkha, river, and dune/beach. These boundary conditions are explained more extensively in sections 3.4.2–3.4.5; a summary of their data sources can be found in table 2. The classes "clay" and "sand" indicate where the Holocene confining clay layer was respectively located and absent in each time slicetimeslice. Firstly, time slices 1 and 2
10   covered the Late-Pleistocene, where sea-levels were low and therefore the coastline was located ~70 km further to the North and the hydraulic gradient was high. The area consisted of an alluvial plain with braided rivers. In between these rivers, sabkha (salt flat) deposits were found in local depressions, which stayed fixed in location (Stanley and Warne, 1993). Next, time slice 3 covered the marine transgression, which was a period of rapid sea-level rise that at announced the start of the Holocene. In time slice 4 the delta started prograding in the west and centre, and lagoons were formed. Of particular interest
15   is the large extent of the Maryut lagoon in the west, filling up what is now known as the Maryut depression (Warne and Stanley, 1993). In time slice 5, progradation started in the east, leading to the symmetrical shape in time slice 6. The eastwards progradation was caused by the decreased hydraulic gradient that led to finer sediments being deposited, which were then transported eastwards with the dominantly eastwards flowing sea currents. During time slice 7, humans converted most river branches to a system of irrigation channels, leaving only the Rosetta and Damietta river branches to persist.

20   #### 3.4.2 Surface waters

The surface water systems, (the sea, rivers, and lagoons)., were incorporated as a CauchyRobin boundary condition (Jazayeri and Werner, 2019), requiring a specified head, salinity, and bed resistance, the latter being the resistance exerted by the bed sediments on the surface-groundwater interaction. Of these three inputs, the least is known about the bed resistance. A 100 day resistance was used for all surface water systems (Table 3), which is a common value for models of the Rhine-Meuse
25   Delta (De Lange et al., 2014; Timmerman and Hemker, 1993). In Appendix B we discuss the effects of this assumption in more detail.

Sea boundary cells were placed on the edge of the coastal shelf and slope, following the bathymetry (GEBCO, 2014) and, where possible, palaeogeographic maps that specified the coastline (Fig. 3). This meant that for the Late-Pleistocene (time slices 1 and 2) we had to recourse to the present-day bathymetry and the contemporary sea-level to approximate the
30   coastline. We specified the head using palaeo sea-level curves. No specific sea-level curves are available for the Nile Delta (Pennington et al., 2017). Hence we resorted to eustatic sea-levels for the Pleistocene (Spratt and Lisiecki, 2016) and a more local curve for the Holocene (Sivan et al., 2001), as this was the best information available. The minimum sea-level was

fixed at -100 m, as below that ~~acquiring~~ achieving numerical convergence was cumbersome. This only slightly affected results as land surface inundation at these sea-levels was ~~the~~ nearly~~near~~ equal due to the steep coastal slope. ~~The~~ The salinity of the sea was set at a constant 35 g TDS/L.

Of the total of seven time slices, the last four have lagoons, since lagoons started to form from 7.5 ka. The lagoon stage was
5   set such that it was in hydrostatic equilibrium with the sea, so that its~~is~~ pressure is corrected for salinity (Post et al., 2007). Lagoonal palaeo-salinities were estimated from the published strontium isotope ratios from the Maryut lagoon (Flaux et al., 2013) for each time slice.~~-~~ To be specific, the salinities assigned to time slices 3 to 7 are respectively: 18, 9, 4.5, 2.5 g TDS/l. These salinities show strong variation through time, ~~varying between 2-18 g TDS/L, as~~because the inflow of the Nile varied through time.

10   River stages were specified by creating linear profiles from the apex to the coastline. The location of the delta apex and its river stage were fixed through time at the present-day conditions. This is a simplification, as in reality the apex' location varied through time. For instance, the delta apex was located up to 65 km south of its current location somewhere in the last 6000 years (Bunbury, 2013), but due to the lack of data on apex migration for the rest of the modelled time domain we chose this simplification. Since the coastline is irregular in shape and the location of rivers was not fixed through time, river stages
15   were determined in the following manner: Firstly, lines were drawn from the apex to points at the coastline for each raster cell at the coastline. Along these lines, river stages declined linearly to the contemporary sea-level at the coastline. Secondly, these profiles were subsequently interpolated to a surface using Inverse Distance Weighting. Finally, the contemporary river branches were clipped out of this surface. Rivers were assigned a salinity of 0 g TDS/L, since the Nile Delta is not tide dominated (Galloway, 1975), meaning that saline water intrusion in river branches is ~~small~~limited.

20   ### 3.4.3 Dunes and beaches

Locations where dunes or beaches occurred were assigned a fixed recharge of 200 mm/year, equal to the present-day average precipitation near Alexandria (WMO, 2006) and also the current recharge along the Levant coast (Yechieli et al., 2010). This is reasonable, as the climate was mainly wetter throughout the Holocene than it currently is (Geirnaert and Laeven, 1992). This recharge allowed for the formation of freshwater lenses underneath dunes and beaches, which ~~have already been~~were
25   observed in the area during ancient times (Post et al., 2018).

3.4.4 Extractions

Groundwater extractions mainly occur in the South-West. This area, near Wadi El Natrun, was appointed a reclamation area by the Egyptian Government in 1990 and since then has seen a rapid increase in extraction rates (King and Salem, 2012; Switzman et al., 2015). This was implemented in the numerical model by including extraction wells (locations and rates) that
30   were in the RIGW database (Nofal et al., 2018) for the last 30 years of the simulation.

3.4.5 Hypersaline groundwater provenances

Below 400 m depth, hypersaline groundwater (HGw) has been observed in the Nile Delta Aquifer. Van Engelen et al. (2018) investigated potential origins of this hypersaline groundwater and identified two potential sources that are included in this model. Firstly, the Late-Pleistocene sabkhas could have been sources of the hypersaline groundwater. Therefore, we set the concentrations of these areas at 120 g/L for certain scenarios that are specified further in section 3.4.6, so that hypersaline groundwater could be formed in these areas. We refer to these as scenarios where HGw originates from the "top".

Secondly, another potential source of hypersaline groundwater could be seepage of hypersaline groundwater expulsed from the low-permeable Kafr Eel Sheikh formation due its compaction. This seepage flux was included in the model as fixed fluxes at the bottom. The magnitude and salt concentration of these fluxes were set at 6E-08 d^{-1} and 120 g TDS/L (van Engelen et al., 2018). These scenarios are referred to as those where HGw originates from the "bottom".

3.5 Initial salinity distributions

Initial 3D salinity distributions were set with a dynamic spin-up time for 160 ka using Late-Pleistocene boundary conditions. Although the 160 ka exceeds a glaciation cycle, we observed that no significant changes in salinity concentrations occurred after 4 ka for the homogeneous and 80 ka for the heterogeneous cases, which is well within the range of a glaciation cycle. For the majority of the model scenarios we started the spin-up with a completely fresh Nile Delta Aquifer, except for the model scenarios receiving HGw from the bottom, where the output of a simulation from previous research (model "NDA-c" in van Engelen et al., 2018) was used to set the initial conditions. This was a simulation of the effect of 2.5 million years of compaction-induced upward flow solute transport through the Kafr El Sheikh formation into the aquifer under interglacial conditions (low hydraulic gradient).

3.6 Model scenarios

Van Engelen et al. (2018) identified two sensitive inputs for hydrogeological models in predicting the distribution of hypersaline groundwater, which are the geological model and the source of hypersaline groundwater. Additionally, they found that when the hypersaline groundwater originates from compaction-induced upward salt transport, the bottom of the aquifer should be closed off. Even though there are 18 possible combinations between the nine lithological model scenarios and the two HGw provenances (top and bot), only 13 were calculated. Based on previous findings in Van Engelen (2018), three combinations could be excluded. These were the combinations where HGw was coming from the bottom and the sea boundary was completely open. The other two excluded scenarios are those with a closed sea boundary, horizontal clay layers (fluvial or marine), and HGw coming from the top. The free-convective plumes in these scenarios created entrapped volumes of fresh groundwater in between clay layers. Pressures in these zones increased fast, which, in combination with the heterogeneity, made it impossible to maintain numerical convergence for these two model scenarios, despite our best efforts.

Additionally As an addition to the palaeo hydrological hydrogeological reconstruction, we formulated for each geology-HGw provenance combination, an equivalent steady-state model which used only the present day hydrological forcings until its salinity distribution reached a steady- state, comparable to previous 3D models made for the area (Mabrouk et al., 2018; Sefelnasr and Sherif, 2014). This all resulted in 26 scenarios in total. To distinguish different model scenarios, we

5   introduced a coding system. For each feature, the corresponding letters in table 1 are converted to a code as follows: *{sea}-{clayer}-{prov}-{temp}*. For example, the palaeohydrogeological reconstruction *(temp*: P) of an aquifer with a half-open sea connection *(sea*: H), fluvial horizontal clay layers *(clayer*: F), and HGw seeping in from the bottom *(prov*: B) gets the following code: H-F-B-P. Its equivalent steady-state model is attributed the code H-F-B-S. Furthermore, all model scenarios that share the same feature are noted as "{feature}-model scenarios", e.g. "T-model scenarios" means all model scenarios

10 where HGw originates from the top, i.e. the Pleistocene sabkhas. The scenarios with a "Homogeneous" lithological model (Fig. 2) get the scenario codes "O-N".

3.7 Model evaluation

We compared our modelled salinity distributions to all the available TDS measurements (n=293), of which the majority is shown in Fig. 1. Comparing measurements with model salinities at the point location can be deceiving as sharp transition

15 zones often occur in salinity distributions (Sanford and Pope, 2010). Therefore, we assessed our model's ability to reproduce measured salinity patterns as follows. First, TDS measurements were binned into classes as shown in Fig. 1, similar to what is common in the validation of hydrogeophysical products (e.g. Delsman et al., 2018). Second, isosurfaces were determined in our model output for all bin edges with the Marching Cubes algorithm (van der Walt et al., 2014). If these classes did not correspond at the measurement location, the minimum displacement (Λ) to the observed class in the model output was

20 determined by calculating the minimum Euclidean distance in 3D:

$$\Lambda = \min \left| \sqrt{\sum_{i=1}^{3}(o_i - \boldsymbol{l_i})^2} \ , \ \sqrt{\sum_{i=1}^{3}(o_i - \boldsymbol{u_i})^2} \right| \qquad (1)$$

where o_i is the location of the observation in dimension i and $\boldsymbol{l_i}$ and $\boldsymbol{u_i}$ the locations of the isosurface vertices in dimension i

25 for respectively the lower and upper bin edge. (Fig. A2). All locations were normalized to a range of [0,1] since the model domain was a rectangular cuboid, in other words not a perfect cube, and thus unequal in size across each dimension.

To assess the validity of the steady-state assumption, we checked for all equivalent steady-state models the time they reached a steady state. This time was determined by calculating the derivative of the freshwater volume over time. If this did not

30 change more than 1E-04% of the total volume, we considered the model to have reached a steady state.
Subsequently, Δ Λ was calculated to assess the improvement in validation of the palaeo reconstructions over their steady-state equivalents:

5   **3.8 Comparison palaeohydrogeological -reconstruction with its equivalent steady-state model**

The palaeohydrogeological- reconstruction was compared with its equivalent steady-state model in two manners. Firstly, we calculated:

$$\Delta C = C_p - C_s \tag{3}$$

10

where C is the 3D salinity matrix at the last timestep, and p and s respectively are the palaeohydrogeologicalpalaeo and steady-state model scenarios. The For locations where $\Delta C > 0$, the palaeohydrogeologocial -reconstruction is saltier than its steady-state equivalent and vice versa.

Secondly, we calculated isohalines for a set of TDS values (3, 10, 20, 30 g/L) for both model scenarios and then

15   lookingassessed at the distance between these isohalines in the y-dimension (North-South). As our coastline was irregular, these distances varied in the x-dimension (West-East), therefore the median (Md) was calculated across this dimension as a conservative measure of isosurface separation (ω), which for a given depth and concentration is defined as:

$$\omega = \text{Md} \left| y_{p,x} - y_{s,x} \right|_x \tag{4}$$

20

where y are the isohaline locations in the y-dimension, and p and s respectively are the palaeohydrogeological and steady-state model scenarios.

34

4 Results

Based on a comparison with observations, five scenarios were selected that showed the best match with the observations. These are called the "acceptable" scenarios. To keep the results comprehensible, we start with discussing the results of these five selected scenarios through space (section 4.1) and time (section 4.2), before discussing this actual selection procedure (section 4.3).

4.1 Model evaluation

Figure 4 shows the evaluation results for all palaeo reconstructions. Note that the 4 classes above 5 g/L in Fig. 1 were grouped into 2 classes: "saline" (5–35 g TDS/L) and "hypersaline" (35–100 g TDS/L), as the number of TDS measurements was low for these classes. All model scenarios matched the location of fresh groundwater correctly (A = 0) at a majority of the observation points. However, their prediction quality worsens with increasing salinity. Despite having sometimes quite different salinity distributions (Fig. 6), the majority of the model scenarios predict the location of the brackish and saline zone with the same error. More striking differences are observed in the hypersaline zone, where major differences occur amongst model scenarios. Therefore, the latter class is the sole criterium in assessing model scenario behaviour. Model scenarios with ᴹᵈ|A| < 0.07 were considered behavioural, which are the following five model scenarios: C-M-B-P, C-N-T-P, H-M-T-P, H-F-T-P, H-N-T-P.

Figure 5 shows the improved skill of a palaeo reconstruction compared to a steady state model. For 11 out of 13 model scenarios, the palaeo reconstruction improved the goodness of fit over the interquartile range of the saline measurements. For the other two (C-M-B and H-F-T), the palaeo-reconstruction led to no clear improvement with a median ΔAAA at zero. The palaeo reconstruction had no clear effect on goodness of fit to fresh or brackish observation points. For four B model scenarios with a bottom HGw provenance (except C-M-B), a steady state simulation improved the fit, as the HGw had virtually unlimited time to flow upwards and get into steady-state with the present day hydraulic gradient. This gradient is roughly an order of magnitude lower than the Pleistocene gradient, the starting situation of the palaeo reconstruction. These four steady-state equivalent model scenarios required more than 60 ka to reach steady-state conditions, which is six times the extent of the Holocene period. In contrast, the C-M-B-P model scenario had a better fit to the observations than its steady-state equivalent, as the low permeable structures allowed preservation of HGw even under steep gradients.

4.1̲2̶ Current spatial TDS distribution of acceptable ̶a̶s̶ ̶p̶r̶e̶d̶i̶c̶t̶e̶d̶ ̶b̶y̶ ̶b̶e̶h̶a̶v̶i̶o̶u̶r̶a̶l̶ model scenarios

The acceptable̶b̶e̶h̶a̶v̶i̶o̶u̶r̶a̶l̶ model scenarios show different salinity distributions (Fig. 4), ̶d̶e̶s̶p̶i̶t̶e̶ ̶a̶l̶l̶ ̶h̶a̶v̶i̶n̶g̶ ̶a̶ ̶n̶e̶a̶r̶l̶y̶ ̶s̶i̶m̶i̶l̶a̶r̶ ̶f̶i̶t̶ ̶t̶o̶ ̶t̶h̶e̶ ̶d̶a̶t̶a̶ despite similar model performance (Fig. 6)(̶F̶i̶g̶.̶ ̶6̶)̶. The model scenarios without clay layers (N) or fluvial clay layers (F) have ̶s̶a̶l̶i̶n̶i̶t̶y̶ freshwater distributions very similar to the O-N-T-P model scenario. In the model scenarios with marine clay layers (M), however, fresh groundwater is preserved in between low-permeable clay layers, especially in the centre. Regardless of the differences between scenarios, in all realizations the fresh-salt interface roughly follows the

coastline, except in the west where there is far extending ~~salt~~ seawater intrusion visible towards Wadi El Natrun (Fig. 1). Next to ~~the influence of~~ this depression, (former) lagoons are visible as shallow brackish zones and (former) dune areas are visible as freshwater lenses. ~~more 3D patterns are visible in the salinities, especially near the (former) lagoons and the fresh groundwater lenses at (former) dune areas (Fig. 3). In general, the fresh-salt interface partly has a conical shape.~~

4.~~2~~4 Salt sources over time

Figure 8~~5~~ shows the provenance of the different groundwater types as fraction of the total modelling domain.~~In~~ In all four T-model scenarios, the model domain starts initially mainly fresh, dominated by infiltrated Pleistocene river water~~.~~, T, ~~this~~ water is then replaced by hypersaline groundwater. As sea-level rises, this hypersaline groundwater is in turn replaced with sea water. The C-M-B-P model scenario takes a slightly different course, as the model is in steady-state during time slice 1, dominated by river water. As sea-level rises and the hydraulic gradient decreases, the hypersaline groundwater and infiltrated sea water volumes increase at the expense of river water volumes. Regardless of differences amongst model scenarios, the total volume of sea groundwater is lower than the total volume of hypersaline groundwater and river water in all model scenarios. The total amount of dune water remains small for all five model scenarios over the entire modelled time.

4.~~1~~3 Model evaluation

Figure 6~~4~~ shows the evaluation results for all palaeo~~hydrogeological~~ reconstructions. Note that the four~~4~~ classes above 5 g/L in Fig. 1 were grouped into 2~~two~~ classes: "saline" (5–35 g TDS/L) and "hypersaline" (35–100 g TDS/L), as the number of TDS measurements was low for these classes. All model scenarios matched the location of fresh groundwater correctly ($\Lambda = 0$) at a majority of the observation points. However, their prediction quality worsens with increasing salinity. Despite having sometimes quite different salinity distributions (Fig. 4~~6~~), the majority of the model scenarios predict the location of the brackish and saline zone with the same error. More striking differences are observed in the hypersaline zone, where ~~major differences occur amongst model scenarios.~~ we observe a division around $\Lambda = 0.07$ into two groups. There are the scenarios with $Md|\Lambda| < 0.07$, that predict the location of the HGw with a similar ~~performance~~ skill as ~~to with which they predict~~ the location of saline groundwater~~.~~, W~~w~~e call these scenarios "acceptable". Specifically, these are the following five model scenarios: C-M-B-P, C-N-T-P, H-M-T-P, H-F-T-P, H-N-T-P. The other scenarios perform considerably worse in predicting the location of the HGw. ~~Therefore, the latter class is the sole criterium in assessing model scenario behaviour. Model scenarios with Md|Λ| < 0.07 were considered behavioural, which are the following five model scenarios: C M B P, C N T P, H M T P, H F T P, H N T P.~~

All equivalent steady-state scenarios required at least several thousands of years to reach a steady state (Table 4) from an initial Pleistocene steady state. Most notable are the B-scenarios, where the hypersaline groundwater caused the system to

respond very slowly, over tens of thousands of years, thus exceeding the extent of the Holocene. The shortest scenarios were the N-T scenarios, as they did not include HGw and due to the lack of clay layers the salt water did not experience any resistance during its flow upwards from its initial Pleistocene state to the Holocene steady state.

~~Figure 5 shows the improved skill of a palaeo-reconstruction compared to a steady-state model. For 11 out of 13 model~~
5 ~~scenarios, the palaeo reconstruction improved the goodness of fit over the interquartile range of the saline measurements. For the other two (C-M-B and H-F-T), the palaeo reconstruction led to no clear improvement with a median A at zero. The palaeo reconstruction had no clear effect on goodness of fit to fresh or brackish observation points. For four B model scenarios with a bottom HGw provenance (except C-M-B), a steady-state simulation improved the fit, as the HGw had virtually unlimited time to flow upwards and get into steady state with the present day hydraulic gradient. This gradient is~~
10 ~~roughly an order of magnitude lower than the Pleistocene gradient, the starting situation of the palaeo reconstruction. These four steady state equivalent model scenarios required more than 60 ka to reach steady state conditions, which is six times the extent of the Holocene period. In contrast, the C-M-B-P model scenario had a better fit to the observations than its steady-state equivalent, as the low permeable structures allowed preservation of HGw even under steep gradients.~~

4.~~3~~ Fresh~~water~~ ~~groundwater~~ volume dynamics and sensitivity analyis

15 Figure 7a shows that fresh~~water~~ ~~groundwater~~ reserves have declined strongly throughout the Late-Pleistocene and Holocene. In case of the T-model scenarios, first by free convection of hypersaline groundwater in time slice~~timeslice~~ 1, next, for all model scenarios, by sea-level rise (time slice~~timeslice~~ 2) and from 13.5 ka by the marine transgression. The total FGw ~~fresh groundwater~~ volumes drop considerably with a factor ranging from 1.9 (C-M-B-P) up to 5.5 (C-N-T-P). There is also considerable variance amongst ~~behavioural~~ acceptable model scenario results, with the C-M-B-P having 74% more
20 FGw ~~fresh groundwater~~ than C-N-T-P. A peculiar observation for the model scenarios C-N-T-P, H-N-T-P, and H-F-T-P is that during the marine transgression (time slice~~timeslice~~ 3) the total FGw ~~fresh groundwater~~ volume recovers after a quick drop. This is caused by the disappearance of the sabkhas, stopping inflow of HGw, whereas outflow of HGw continues, allowing the total FGw volume to recover (see video supplement).

The sensitivity analysis (Fig. 7b) shows a clear influence of horizontal clay layers. For the M-model scenarios, more FGw
25 ~~fresh groundwater~~ is maintained in the transient model scenarios than in the steady-state model scenarios. In contrast, the opposite is visible for the N-model scenarios: FGw ~~fresh groundwater~~ volumes are lower in the transient simulations than in their steady-state counterparts.

In the model scenarios with marine clay layers there are still considerable fresh groundwater reserves available below 300 m depth and in one case even offshore [C-M-B-P] (Table 5~~4~~). This table ~~also~~ shows that ~~there is less spread amongst model~~
30 ~~scenarios when looking at the shallow onshore fresh groundwater volume~~these parts of the model are ~~also~~ the most uncertain as well~~.~~, since ~~D~~disregarding potential deep and offshore fresh groundwater volumes decreases the spread from 74% to 32%.

4.5 Fresh-salt distribution: of the palaeohydrogeological reconstruction in comparison withagainst its equivalent steady-state distribution

ΔC (eq. 3 in sect. 3.8) is generally negative in the marine clay (M) comparisons (snapshots 2 and 4 in Fig. 98), as the low permeable clays block vertical intrusion during the marine transgression, concurrent with previous research (Kooi et al., 2000; Post et al., 2013). In the no clay (N) comparison (snapshots 1, 3, and 6 in Fig. 98) however, we observe somethinga different observation is made: the shallow groundwater (< 200 m depth) is saltier, whereas the toe of the wedge of the steady-state equivalents lays further land inward. This pattern is the most clearly visible in the homogeneous model (Fig. 89, plot 1) and is obscured more as heterogeneity increases, as also visible in Fig. 10. In the C-N-T scenario, the palaeohydrogeological reconstruction is more saline nearly everywhere, as there is a larger volume of hypersaline groundwater that is influencing the toe of the salt seawater wedge (see also Fig. 64, snapshot 3).

In all comparisons, (former) lagoons are visible (Fig. 3 and Fig. 89). In the west, the former Maryut lagoon is clearly visible in all model scenarios as a red zone, stretching all the way towards the Wadi El Natrun depression. The central and eastern lagoons, however, are coloured blue. At these areas, the palaeo-hydraulic gradients were the steepest as the lagoons lay more land inward up till 0.2 ka (Stanley and Warne, 1993), resulting in a quicker freshening of the groundwater at these locations.

Figure 911 shows the isosurface separation ω for the model scenarios where the previous described pattern was visible. An S-shaped trend in ω is visible with depth, where there is a maximum isosurface separation at around -130 m and minimum at around -350 m depth. The latter is not visible in C-N-T as the hypersaline groundwater influences the fresh-salt interface. This trend means that, for these four model scenarios, using a steady-state model results in an underestimation of salt seawater intrusion length above -250 m depth and overestimation of the intrusion of the "toe" of the wedge. The median of this over- and underestimation of the salt seawater wedge top and toe ranges up to 10 km. The isosurface seperation increases with concentration, because the denser surface has to rotate further to reach a steady state.

5 Discussion

The efforts of conducting 3D palaeohydrogeological -reconstructions were rewarded with new insights. We quantitatively show for the first time that a strong reduction in FGw volume presumably occurred in the Nile Delta Aquifer during the last 32 ka. As sea-level rose, FGw volumes were reduced with a factor ranging from 1.9 to 5.4 (Fig. 7). Just as in the Netherlands

5   (Delsman et al., 2014) and the Mekong (Van Pham et al., 2019), the system neverdid not reached a steady -state in the last 9000 years. Our equivalent steady-state scenarios required at least 5500 years to reach a steady state (Table 4), a period in which already considerable changes occurred to the boundary conditions (Fig. 3). This increased to tens of thousands of years for the more complex models. We therefore doubt that using a steady-state approach with current boundary conditions results in a reasonable estimate of the current fresh-salt groundwater distribution for such a complex, large-scale system.

10   Moreover, the variance in FGw volumes of the palaeohydrogeological -reconstructions was larger than that of the steady-state model scenarios. This implies that conducting sensitivity and uncertainty analyses on steady-state model scenarios, can lead to grave underestimations of the uncertainty. In addition, the palaeo-reconstruction improved simulations of the location of the saline zone (Fig. 5). Furthermore, we have shown that it is physically possible that the influences of the Holocene marine transgression can be still observed in the delta. Either as freshwater -groundwater volumes in between clay layers in

15   the northern part of the NDA, concurrent with earlier research (Kooi et al., 2000), or as a steeper fresh-salt interface than the steady-state interface, which to the authors' knowledge has never been shown before in this detail. The implications of this steeper interface are that using the steady-state approach can lead to an underestimation of the amount of salt groundwater above roughly 250 m depth and an overestimation of the toe length of the interface (Fig. 911). Likewise, it could explain the observed freshening of the shallow NDA as observed in hydrogeochemical studies (Barrocu and Dahab, 2010; Geriesh et al.,

20   2015), since only the zone above 250 m depth has been sampled. It was however difficult to compare the observed freshening with model results directly, since the time scale over which this observed freshening has developed is often unknown (Stuyfzand, 2008). WhichThe sign of the error is made whenin simulating salinity distributions of deltaic areas with steady-state scenarios largely depends on the lithology of their subsurface. Deltas with abundant low permeable clay layers are expected to possess larger quantities of deep fresh groundwater than would be approximated with steady-state

25   scenarios. Examples of such deltas are the Chao Praya Delta, Thailand (Yamanaka et al., 2011), Red River Delta, Vietnam (Larsen et al., 2017), and the Pearl Delta, China (Wang and Jiao, 2012). In deltaic groundwater systems consisting of mainly coarse material, a steady-state approximation may underestimate the slope of the fresh-salt interface. Examples of these deltaic areas are the Niger Delta, Nigeria (Summerhayes et al., 1978), and the Tokar Delta, Sudan (Elkrail and Obied, 2013). Note that the Holocene transgression in the above-mentioned Asian deltas reached more land inward than it did in the Nile

30   Delta (Sinsakul, 2000; Tanabe et al., 2006; Zong et al., 2009), thus their salt groundwater distributions are expected to even more strongly deviate even more from from from steady-state approximations than in the research case in this study.

There is a large ~~variance~~ variation in estimated FGw ~~fresh groundwater~~ volumes between acceptable model scenarios in the NDA, mainly located offshore and at larger depths. This ~~variance~~ variation is mainly attributed to the limited availability of salinity observations in the saline zone. Although the number of observations may look adequate on a 2D map, it is still by far insufficient in 3D. Similar conclusions were drawn by Sanford and Pope (2010) for the Eastern Shore of Virginia (USA), an area with a similar observation density. Though it is generally assumed for the sake of simplicity that the northern part of the delta is completely saline, here we show that there might ~~be just as well large~~also be large, overlooked quantities of offshore fresh groundwater. Still, the FGw ~~fresh groundwater~~ volume ~~estimations~~estimates~~ionsestimations~~ are on the large side. For instance, Mabrouk et al. (2018) estimated a total amount of 1290 km^3 and Sefelnasr and Sherif (2014) 883 km^3. We stress that the latter authors disregarded the 80 km of coastal shelf offshore and used a lower porosity value. Reducing the effective porosity from our 25% to their 17%, reduces our estimated FGw volume with 32%; disregarding the increased groundwater flow velocities this would result in, which in turn would slightly increase the extent of the FGw ~~fresh groundwater~~ zone. ~~The variance in the results should also affect management decisions. Three dimensional groundwater model scenarios are becoming increasingly important for management decisions, as they are currently the only available tool to estimate the fresh groundwater volumes on a deltaic groundwater system scale (~200 km * 200 km * 1 km) (Delsman et al., 2014). As Egypt's reliance on groundwater increases (World Bank, 2018), more hydrogeological data should be obtained and made available to better constrain model results, which should aid management decisions.~~

Of all groundwater constituents, hypersaline groundwater in the NDA is the main contributor to total salt mass (Fig. 5~~8~~). We can limit the amount of potential combinations between lithology and HGw provenance that were posed in van Engelen et al (2018), as the current study also incorporated a steep glacial hydraulic gradient, whereas van Engelen et al (2018) kept a constant low (interglacial) hydraulic gradient. Apparent from the model evaluation is that the only ~~behavioural~~ acceptable model where HGw originates from the bottom includes a closed off continental slope and extending, low-permeable horizontal clay layers, resulting in a compartmentalization of the bottom of the aquifer. When HGw originates from the top, ~~also~~ a closed off bottom of the coastal slope is also necessary to preserve the observed amount of hypersaline groundwater since the last Glacial. This implies that regardless of HGw provenance, the interaction at the aquifer-sea connection needs to be very limited at the lower half of the aquifer to explain the observed salinities, as otherwise all HGw would have flowed out to the sea under the steep Late-Pleistocene (32 – 13.5 ka) hydraulic gradients.

Next to estimating FGw volumes, the three-~~dimensionality~~dimensional nature~~itydimensionality~~ of our model showed us a vulnerable zone in the area: The Wadi El Natrun depression. This area might have started attracting, and eventually draining, sea water from the Maryut lagoon as sea-level rose above the depression height. The amount of evidence of this actually occurring is not overwhelming, but there are some observations in support of this hypothesis (Ibrahim Hussein et al., 2017). Nevertheless, it is concerning that the area with the most far reaching ~~salt~~ seawater intrusion in our model scenarios, is coincidentally in reality also the area with the most intense groundwater pumping. Even though the 30 years of groundwater extraction in our model scenarios did not seem to show a large influence on ~~salt~~ seawater intrusion, this influence can

increase in the coming century (Mabrouk et al., 2018), since the Nile Delta Aquifer is a slow responding groundwater system. In addition, the large cell sizes of our model can also be the cause of the lack of visible quick responses to groundwater pumping, as large model cells tend to negate local salt-water upconing effects around wells (Pauw et al., 2015). In addition, iIt should also be kept in mind that for future hydrogeochemical analyses of the Wadi El Natrun area that the former Maryut lagoon used to extend a lot more land inward towards this depression than currently, and that the composition of the water in this lagoon fluctuated strongly over time (Flaux et al., 2013).

The efforts of conducting 3D palaeo-reconstructions were rewarded with new insights. We quantitatively show for the first time that a strong reduction in FGw volume presumably occurred in the Nile Delta Aquifer during the last 32 ka. As sea-level rose, FGw volumes were reduced with a factor ranging from 1.9 to 5.4 (Fig. 7). Just as in the Netherlands (Delsman et al., 2014) and the Mekong (Van Pham et al., 2019), the system never reached steady-state in the last 9000 years. Moreover, the variance in FGw volumes of the palaeo-reconstructions was larger than that of the steady state model scenarios. This implies that conducting sensitivity and uncertainty analyses on steady-state model scenarios, can lead to grave underestimations of the uncertainty. In addition, the palaeo-reconstruction improved simulations of the location of the saline zone (Fig. 5). Furthermore, we have shown that it is physically possible that the influences of the Holocene marine transgression can be still observed in the delta. Either as fresh groundwater volumes in between clay layers in the northern part of the NDA, concurrent with earlier research (Kooi et al., 2000), or as a steeper fresh-salt interface than the steady-state interface, which to the authors' knowledge has never been shown before in this detail. The implications of this steeper interface are that using the steady state approach can lead to an underestimation of the amount of salt groundwater above roughly 250 m depth and an overestimation of the toe length of the interface (Fig. 11). Likewise, it could explain the observed freshening of the shallow NDA as observed in hydrogeochemical studies (Barrocu and Dahab, 2010; Geriesh et al., 2015), since only the zone above 250 m depth has been sampled. It was however difficult to compare the observed freshening with model results directly, since the time scale over which this observed freshening has developed is often unknown (Stuyfzand, 2008). Which error is made in simulating salinity distributions of deltaic areas with steady state scenarios largely depends on the lithology of their subsurface. Deltas with abundant low permeable clay layers are expected to possess larger quantities of deep fresh groundwater than would be approximated with steady-state scenarios. Examples of such deltas are the Chao Praya Delta, Thailand (Yamanaka et al., 2011), Red River Delta, Vietnam (Larsen et al., 2017), and the Pearl Delta, China (Wang and Jiao, 2012). In deltaic groundwater systems consisting of mainly coarse material, a steady-state approximation may underestimate the slope of the fresh-salt interface. Examples of these deltaic areas are the Niger Delta, Nigeria (Summerhayes et al., 1978), and the Tokar Delta, Sudan (Elkrail and Obied, 2013). Note that the Holocene transgression in the above-mentioned Asian deltas reached more land inward than it did in the Nile Delta (Sinsakul, 2000; Tanabe et al., 2006; Zong et al., 2009), thus their salt groundwater distributions are expected to more strongly deviate from from steady-state approximations.

Despite our efforts at increasing the realism of our model scenarios, some processes were only incorporated in a very simplistic manner. The most prominently simplified process in this arid area is salinization due to evapoconcentration and

redissolution of precipitated salts. Though this does not currently seem to influence groundwater salinities, it is hypothesized to have been of influence in the past (Diab et al., 1997; Geirnaert and Laeven, 1992). In our T-scenarios, the Pleistocene sabkhas were assigned a fixed concentration (120 g/L), meaning that free convection here was unconstrained by available salt and water mass. Our T-scenario results indeed show a very rapid change in the first 5000 years (Fig. 5 and 7 and 8), after which they reach equilibrium conditions until the next time slice. This may seem very rapid, but experiments with smaller scale models which fully accounted for salt precipitation and dissolution, evapo-concentration, variable-density groundwater flow and the unsaturated zone, showed that higher salinities (> 200 g/L) are possible under evaporation rates similar to those in the Nile Delta (Shen et al., 2018; Zhang et al., 2015). Given that the Late-Pleistocene sabkhas did not migrate (Stanley and Warne, 1993), we deem our T-model scenarios to be conservative. Likewise, rock dissolution is thought to influence salinities around the borders of the delta (Ibrahim Hussein et al., 2017). Either of these processes can explain the saline observations in the South-East (Fig. 1), which our model cannot.

In addition, for a proper physical representation of free convection, a finer grid is required. A coarse horizontal cell size results in a delay in the onset of free convection, while a coarse vertical cell size results in an onset of free convection even for situations that are expected to be stable (Kooi et al., 2000). van Engelen et al. (2018, Appendix D) investigated the errors caused by coarse model cells for the Nile Delta and found that especially the crude horizontal grid size had an influence. They found that this resulted in similar downward fluxes, but a delay in the onset of free convection. This effect, however, was negligible after ~50 years and thus dwarfed by the timescale used for of our time slices. The coarse vertical grid size was not an issue, since the marine transgression occurred over sand with a high hydraulic conductivity, meaning there is a very instable situation and free convection has to occur. We thus think that the errors made in modelling free convection are not of strong influence will not impact to our conclusions.

Future research for this area would greatly benefit from an enhanced geological model, especially of the lithology of the delta shelf and slope. A lot of measurements were conducted for the petroleum industry (Sestini, 1989), but most of this data is still inaccessible for external researchers. Furthermore, the available publications on offshore geology are written mainly for the petroleum industry, focusing on the area >2.5 km depth, rendering them unsuitable for hydrogeological research of the Nile Delta Aquifer. Therefore, we think research for this area would benefit from a re-analysis of the available data with a hydrogeological perspective. In addition, a larger salinity dataset, especially in the saline zone would help validating groundwater models that are used more frequently in management decisions, e.g. Nofal et al. (2016).

6 Conclusions

A 3D variable-density groundwater flow and coupled salt transport model was used for a palaeohydrogeological -reconstruction of the salinity distribution of the Nile Delta Aquifer. It was found that large timescales are involved, as steady-state model scenarios required at least 5500 years Compared to a steady-state analysis, the palaeo-reconstruction was better in reproducing salinities in the saline zone. to reach equilibrium. Hence, allnone of the evaluated paleohydrogeological scenarios never reached a steady state inover the last 9000 years, meaning that the transient boundary conditions definitely had an influence on current groundwater salinity. Given the large range variation in delta-architectures analysesTherefore, we can conclude that steady-state models are not likely to result in realistic FGw distributions in deltaic areas. Therefore,Our results also show that the occurrence of -past we deem the marine transgressions constitute atransgression as a valid hypothesis explaining the occurrence of the extensive saline zones land inward. Nevertheless, the estimated FGw volumes were subject to considerable uncertainty, due to the lack of constraining data in this salineis zone, mainly due to uncertainty in the permeability of the horizontal clay layers. Regardless, the palaeohydrogeological- reconstruction provided several insights. Firstly, during the Late-Pleistocene the total FGw volumes declined strongly, with estimates ranging from a factor 1.9 to 5.4. Secondly, ignoring past boundary conditions lead to an underestimation of FGw uncertainty resulting from a lack of knowledge on geological schematizations and HGw provenances. Thirdly, the differences between the fresh-salt distribution of the palaeohydrogeological -reconstruction and the steady-state scenarios varied, depending on the geology. In case of low permeable clay layers, the palaeohydrogeological- reconstruction resulted in considerable volumes of fresh groundwater stored underneath these clay layers. However, with more permeable or no clay layers, the fresh-salt interface was steeper than in the steady-state equivalents. Therefore, using steady-state boundary conditions is likely to result in erroneous fresh-salt distributions and an underestimation of the uncertainty in FGw volumes.

Video supplement

Supplementary videos show the development of the groundwater fresh-salinity distribution through time for the five validated acceptable models scenarios. These can be found under the following DOI: 10.5281/zenodo.2628427

Appendix A: Additional figures methods

-Figure A1 here-

-Figure A2 here-

Appendix B: Sensitivity analysis of the resistance of the boundaries

To assess the effects of our assumption of the boundary resistance, we ran two alternative versions of the "H-N-T-P" scenario with a different resistance. This scenario was chosen as it presumably is the "acceptable" scenario that would be affected the most by the resistance value, as it has no horizontal clay layers that resist changes in boundary conditions and the sea boundary has the most open connection with the sea. We multiplied the resistance with a factor 0.5 and 2. Lowering the resistance more than with a factor 0.5 lead to numerical convergence issues. Fig. B1 shows that throughout the Pleistocene the resistance influences the groundwater types, as the lower resistance allows more river water to be replaced with hypersaline groundwater. The groundwater types of the different models quickly converge, however, through the Holocene. We therefore think that the choice of boundary resistance has limited effect on our results and conclusions, despite that we only varied the resistance to a limited extent in this sensitivity analysis.

References

Abdel-Fattah, M. I.: Petrophysical characteristics of the messinian abu madi formation in the baltim east and north fields, offshore Nile delta, Egypt, J. Pet. Geol., 37(2), 183–195, doi:10.1111/jpg.12577, 2014.

Abdel Aal, A., El Barkooky, A., Gerrits, M., Meyer, H., Schwander, M. and Zaki, H.: Tectonic evolution of the Eastern Mediterranean Basin and its significance for hydrocarbon prospectivity in the ultradeepwater of the Nile Delta, Lead. Edge, 19(10), 1086, doi:10.1190/1.1438485, 2000.

El Banna, M. M.: Nature and human impact on Nile Delta coastal sand dunes, Egypt, Environ. Geol., 45(5), 690–695, doi:10.1007/s00254-003-0922-y, 2004.

Barrett, R., Berry, M., Chan, T. F., Demmel, J., Donato, J., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C. and van der Vorst, H.: Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods., Math. Comput., 64(211), 1349, doi:10.2307/2153507, 2006.

Barrocu, G. and Dahab, K.: Changing climate and saltwater intrusion in the Nile Delta , Egypt, in Groundwater Response to a changing Climate, edited by T. Makato and I. Holman, pp. 11–25, Taylor & Francis., 2010.

Biswas, A. K.: Land Resources for Sustainable Agricultural Development in Egypt, Ambio, 22(8), 556–560, 1993.

Bunbury, J.: Geomorphological development of the Memphite floodplain over the past 6,000 years, Stud. Quat., 30(2), 61–67, doi:10.2478/squa-2013-0005, 2013.

Coleman, J. M., Roberts, H. H., Murray, S. P. and Salama, M.: Morphology and dynamic sedimentology of the eastern nile delta shelf, Dev. Sedimentol., 32(C)I, 301–326, doi:10.1016/S0070-4571(08)70304-2, 1981.

Colombani, N., Cuoco, E. and Mastrocicco, M.: Origin and pattern of salinization in the Holocene aquifer of the southern Po Delta (NE Italy), J. Geochemical Explor., 175, 130–137, doi:10.1016/j.gexplo.2017.01.011, 2017.

Delsman, J., Van Baaren, E. S., Siemon, B., Dabekaussen, W., Karaoulis, M. C., Pauw, P., Vermaas, T., Bootsma, H., De Louw, P. G. B., Gunnink, J. L., Dubelaar, W., Menkovic, A., Steuer, A., Meyer, U., Revil, A. and Oude Essink, G. H. P.: Large-scale, probabilistic salinity mapping using airborne electromagnetics for groundwater management in Zeeland, the Netherlands, Environ. Res. Lett., 13, doi:10.1088/1748-9326/aad19e, 2018.

Delsman, J. R., Hu-A-Ng, K. R. M., Vos, P. C., De Louw, P. G. B., Oude Essink, G. H. P., Stuyfzand, P. J. and Bierkens, M. F. P.: Paleo-modeling of coastal saltwater intrusion during the Holocene: An application to the Netherlands, Hydrol. Earth Syst. Sci., 18(10), 3891–3905, doi:10.5194/hess-18-3891-2014, 2014.

Dermody, B. J., van Beek, R. P. H., Meeks, E., Klein Goldewijk, K., Scheidel, W., van der Velde, Y., Bierkens, M. F. P., Wassen, M. J. and Dekker, S. C.: A virtual water network of the Roman world, Hydrol. Earth Syst. Sci., 18(12), 5025–5040, doi:10.5194/hess-18-5025-2014, 2014.

Diab, M. S., Dahab, K. A. and El-Fakharany, M. A.: Impact of the Paleohydrogeological conditions on the groundwater quality in the Northern Part of the Nile Delta, Egypt, Egypt. J. Geol., 41(2B), 779–796, 1997.

Dolean, V., Jolivet, P. and Nataf, F.: An Introduction to Domain Decomposition Methods : algorithms , theory and parallel

implementation To cite this version : An Introduction to Domain Decomposition Methods : algorithms , theory and parallel implementation, , 289 [online] Available from: https://hal.archives-ouvertes.fr/cel-01100932v2, 2015.

Domenico, P. A. and Schwartz, F. W.: Physical and Chemical Hydrogeology, First edit., John Wiley & Sons., 1990.

El-Agha, D. E., Closas, A. and Molle, F.: Below the radar: the boom of groundwater use in the central part of the Nile Delta in Egypt, Hydrogeol. J., (6), doi:10.1007/s10040-017-1570-8, 2017.

Elkrail, A. B. and Obied, B. A.: Hydrochemical characterization and groundwater quality in Delta Tokar alluvial plain, Red Sea coast-Sudan, Arab. J. Geosci., 6(8), 3133–3138, doi:10.1007/s12517-012-0594-6, 2013.

van Engelen, J., Oude Essink, G. H. P., Kooi, H. and Bierkens, M. F. P.: On the origins of hypersaline groundwater in the Nile Delta aquifer, J. Hydrol., 560, 301–317, doi:10.1016/j.jhydrol.2018.03.029, 2018.

Farid, M.: Nile Delta Groundwater Study, Cairo University., 1980.

Fass, T., Cook, P. G., Stieglitz, T. and Herczeg, A. L.: Development of saline ground water through transpiration of sea water, Ground Water, 45(6), 703–710, doi:10.1111/j.1745-6584.2007.00344.x, 2007.

Faye, S., Maloszewski, P., Stichler, W., Trimborn, P., Faye, S. C. and Gaye, C. B.: Groundwater salinization in the Saloum (Senegal) delta aquifer: Minor elements and isotopic indicators, Sci. Total Environ., 343(1–3), 243–259, doi:10.1016/j.scitotenv.2004.10.001, 2005.

Flaux, C., Claude, C., Marriner, N. and Morhange, C.: A 7500-year strontium isotope record from the northwestern Nile delta (Maryut lagoon, Egypt), Quat. Sci. Rev., 78(August), 22–33, doi:10.1016/j.quascirev.2013.06.018, 2013.

Gallichand, J., Marcotte, D., Prasher, S. O. and Broughton, R. S.: Optimal sampling density of hydraulic conductivity for subsurface drainage in the Nile delta, Agric. Water Manag., 20(4), 299–312, doi:10.1016/0378-3774(92)90004-G, 1992.

Galloway, W. D.: Process Framework for describing the morphologic and stratigraphic evolution of deltaic depositional systems, Deltas, Model. Explor., 86–98, 1975.

GEBCO: GEBCO Dataset, [online] Available from: http://www.gebco.net/data_and_products/gridded_bathymetry_data/gebco_30_second_grid/ (Accessed 5 October 2016), 2014.

Geirnaert, W. and Laeven, M. P.: Composition and history of ground water in the western Nile Delta, J. Hydrol., 138(1–2), 169–189, doi:10.1016/0022-1694(92)90163-P, 1992.

Gelhar, L. W., Welty, C. and Rehfeldt, K. R.: A Critical Review of Data on Field-Scale Dispersion in Aquifers, Water Resour. Res., 28(7), 1955–1974, doi:10.1029/92WR00607, 1992.

Geriesh, M. H., Balke, K.-D., El-Rayes, A. E. and Mansour, B. M.: Implications of climate change on the groundwater flow regime and geochemistry of the Nile Delta, Egypt, J. Coast. Conserv., 19(4), 589–608, doi:10.1007/s11852-015-0409-5, 2015.

Golub, G. and Van Loan, C. F.: Matrix Computations, 3rd ed., The Johns Hobkins University Press, London., 1996.

Gossel, W., Sefelnasr, A. and Wycisk, P.: Modelling of paleo-saltwater intrusion in the northern part of the Nubian Aquifer System, Northeast Africa, Hydrogeol. J., 18(6), 1447–1463, doi:10.1007/s10040-010-0597-x, 2010.

Guo, W. and Langevin, C. D.: User's Guide to SEAWAT: A computer program for simulation of three-dimensional variable-density ground-water flow, USGS, Tallahassee., 2002.

Harbaugh, Arlen, W.: MODFLOW-2005 , The U . S . Geological Survey Modular Ground-Water Model — the Ground-Water Flow Process, U.S. Geol. Surv. Tech. Methods, 253, 2005.

5   Higgins, S. a.: Review: Advances in delta-subsidence research using satellite methods, Hydrogeol. J., 587–600, doi:10.1007/s10040-015-1330-6, 2016.

Ibrahim Hussein, H. A., Ricka, A., Kuchovsky, T. and El Osta, M. M.: Groundwater hydrochemistry and origin in the south-eastern part of Wadi El Natrun, Egypt, Arab. J. Geosci., 10(7), 1–14, doi:10.1007/s12517-017-2960-x, 2017.

Jasechko, S., Perrone, D., Befus, K. M., Bayani Cardenas, M., Ferguson, G., Gleeson, T., Luijendijk, E., McDonnell, J. J.,
10 Taylor, R. G., Wada, Y. and Kirchner, J. W.: Global aquifers dominated by fossil groundwaters but wells vulnerable to modern contamination, Nat. Geosci., 10(April), 425–430, doi:10.1038/ngeo2943, 2017.

Jazayeri, A. and Werner, A. D.: Boundary condition nomenclature confusion in groundwater flow modelling, Groundwater, gwat.12893, doi:10.1111/gwat.12893, 2019.

Kashef, A.-A. I.: Salt-Water Intrusion in the Nile Delta, Groundwater, 21(2), 160–167, doi:10.1111/j.1745-
15 6584.1983.tb00713.x, 1983.

Ketabchi, H., Mahmoodzadeh, D., Ataie-Ashtiani, B. and Simmons, C. T.: Sea-level rise impacts on seawater intrusion in coastal aquifers: Review and integration, J. Hydrol., 535, 235–255, doi:10.1016/j.jhydrol.2016.01.083, 2016.

King, C. and Salem, B.: A socio-ecological investigation of options to manage groundwater degradation in the western desert, Egypt, Ambio, 41(5), 490–503, doi:10.1007/s13280-012-0255-8, 2012.

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

Kubota, A., Zayed, B., Fujimaki, H., Higashi, T., Yoshida, S., Mahmoud, M. M. A., Kitamura, Y. and Hassan, W. H. A. El: Chapter 7: Water and salt movement in Soils of the Nile Delta, in Irrigated Agriculture in Egypt., 2017.

Laeven, M. P.: Hydrogeological Study of the Nile Delta and Adjacent Desert Areas Egypt with emphasis on hydrochemistry
25 and isotope hydrology, Amsterdam., 1991.

De Lange, W. J., Prinsen, G. F., Hoogewoud, J. C., Veldhuizen, A. A., Verkaik, J., Oude Essink, G. H. P., Van Walsum, P. E. V., Delsman, J. R., Hunink, J. C., Massop, H. T. L. and Kroon, T.: An operational, multi-scale, multi-model system for consensus-based, integrated water management and policy analysis: The Netherlands Hydrological Instrument., Environ. Model. Softw., 59, 98–108, doi:10.1016/j.envsoft.2014.05.009, 2014.

30 Langevin, C. D., Shoemaker, W. B. and Guo, W.: MODFLOW-2000, the U.S. Gelogical Survey Modular Ground-Water Model-Documentation of the SEAWAT-2000 version with the Variable-Density Flow Process (VDF) and the Integrated MT3DMS Transport Process (IMT). [online] Available from: http://water.usgs.gov/ogw/seawat/, 2003.

Langevin, C. D., Thorne Jr., D. T., Dausman, A. M., Sukop, M. C. and Guo, W.: SEAWAT Version 4: A Computer Program for Simulation of Multi-Species Solute and Heat Transport, U.S. Geol. Surv. Tech. Methods B. 6, 39, 2008.

Larsen, F., Tran, L. V., Van Hoang, H., Tran, L. T., Christiansen, A. V. and Pham, N. Q.: Groundwater salinity influenced by Holocene seawater trapped in incised valleys in the Red River delta plain, Nat. Geosci., 10(May), doi:10.1038/ngeo2938, 2017.

Mabrouk, M., Jonoski, A., Oude Essink, G. H. P. and Uhlenbrook, S.: Impacts of sea level rise and increasing fresh water demand on sustainable groundwater management, Water, 10(1690), 1–14, doi:10.3390/w10111690, 2018.

Mabrouk, M. B., Jonoski, A., Solomatine, D. and Uhlenbrook, S.: A review of seawater intrusion in the Nile Delta groundwater system – the basis for assessing impacts due to climate changes and water resources development, Hydrol. Earth Syst. Sci. Discuss., 10(8), 10873–10911, doi:10.5194/hessd-10-10873-2013, 2013.

Malm, A. and Esmailian, S.: Ways In and Out of Vulnerability to Climate Change: Abandoning the Mubarak Project in the Northern Nile Delta, Egypt, Antipode, 45(2), 474–492, doi:10.1111/j.1467-8330.2012.01007.x, 2013.

Manzano, M., Custodio, E., Loosli, H., Cabrera, M. C., Riera, X. and Custodio, J.: Palaeowater in coastal aquifers of Spain, Geol. Soc. London, Spec. Publ., 189(1), 107–138, doi:10.1144/gsl.sp.2001.189.01.08, 2001.

Meisler, H., Leahy, P. P. and Knobel, L. L.: Effect of Eustatic Sea-Level Changes on Saltwater-Freshwater in the Northern Atlantic Coastal Plain, in USGS Water Supply Paper: 2255, p. 28, U.S. Geological Survey., 1984.

NASA: Shuttle Radar Topography Mission, 2014.

Nofal, E. R., Amer, M. A., El-Didy, S. M. and Akram, M. F.: Sea Water Intrusion in Nile Delta in Perspective of New Configuration of the Aquifer Heterogeneity Using the Recent Stratigraphy Data, J. Am. Sci., 11(6), 567–570, 2015.

Nofal, E. R., Mostafa, H. F. A., van der Linden, W. and Oude Essink, G. H. P.: Enhanced Water Resources Management Project Egypt, Technical Study 4: Groundwater Modelling, Nile Delta., Cairo., 2016.

Nofal, E. R., Fekry, A. M., Ahmed, M. H. and El-Kharakany, M. M.: Groundwater: Extraction versus recharge; vulnerability assessment, Water Sci., (2017), doi:10.1016/j.wsj.2018.07.002, 2018.

Oude Essink, G. H. P., Van Baaren, E. S. and De Louw, P. G. B.: Effects of climate change on coastal groundwater systems: A modeling study in the Netherlands, Water Resour. Res., 46(10), 1–16, doi:10.1029/2009WR008719, 2010.

Panday, S., Langevin, C. D., Niswonger, R. G., Ibaraki, M. and Hughes, J. D.: MODFLOW-USG Version 1: An Unstructured Grid Version of MODFLOW for Simulating Groundwater Flow and Tightly Coupled Processes Using a Control Volume Finite-Difference Formulation, U.S. Geol. Surv. Tech. Methods, (6-A45), 66, 2013.

Paradigm: SKUA-GOCAD, [online] Available from: http://www.pdgm.com/products/skua-gocad/, 2017.

Pauw, P. S., van der Zee, S. E. A. T. M., Leijnse, A. and Oude Essink, G. H. P.: Saltwater Upconing Due to Cyclic Pumping by Horizontal Wells in Freshwater Lenses, Groundwater, (1997), doi:10.1111/gwat.12382, 2015.

Pennington, B. T., Sturt, F., Wilson, P., Rowland, J. and Brown, A. G.: The fluvial evolution of the Holocene Nile Delta, Quat. Sci. Rev., 170, 212–231, doi:10.1016/j.quascirev.2017.06.017, 2017.

Van Pham, H., van Geer, F. C., Tran, V. B., Dubelaar, W. and Oude Essink, G. H. P.: Paleo-hydrogeological reconstruction of the fresh-saline groundwater distribution in the Vietnamese Mekong Delta since the late Pleistocene, J. Hydrol. Reg. Stud., 2019.

Post, V., Kooi, H. and Simmons, C.: Using hydraulic head measurements in variable-density ground water flow analyses, Ground Water, 45(6), 664–671, doi:10.1111/j.1745-6584.2007.00339.x, 2007.

Post, V. E. A., Groen, J., Kooi, H., Person, M., Ge, S. and Edmunds, W. M.: Offshore fresh groundwater reserves as a global phenomenon., Nature, 504(7478), 71–8, doi:10.1038/nature12858, 2013.

5 Post, V. E. A., Houben, G. J. and Engelen, J. Van: What is the Ghijben-Herzberg principle and who formulated it?, Hydrogeol. J., (September 1912), 2018.

Rasmussen, E. K., Svenstrup Petersen, O., Thompson, J. R., Flower, R. J. and Ahmed, M. H.: Hydrodynamic-ecological model analyses of the water quality of Lake Manzala nile delta, northern Egypt, Hydrobiologia, 622(1), 195–220, doi:10.1007/s10750-008-9683-7, 2009.

10 Salem, Z. E. and El-Bayumy, D. A.: Hydrogeological, petrophysical and hydrogeochemical characteristics of the groundwater aquifers east of Wadi El-Natrun, Egypt, NRIAG J. Astron. Geophys., 5(1), 124–146, doi:10.1016/j.nrjag.2015.12.001, 2016.

Salem, Z. E., Al Temamy, A. M., Salah, M. K. and Kassab, M.: Origin and characteristics of brackish groundwater in Abu Madi coastal area, Northern Nile Delta, Egypt, Estuar. Coast. Shelf Sci., 178, 21–35, doi:10.1016/j.ecss.2016.05.015, 2016.

15 Samuel, A., Kneller, B. and Al., E.: Prolific deep marine slope channels of the Nile Delta, Egypt, Am. Assoc. Pet. Geol. Bull., 87(4), 541–560, doi:10.1306/1105021094, 2003.

Sanford, W. E. and Pope, J. P.: Current challenges using models to forecast seawater intrusion: lessons from the Eastern Shore of Virginia, USA, Hydrogeol. J., 18(1), 73–93, doi:10.1007/s10040-009-0513-4, 2010.

Sefelnasr, A. and Sherif, M.: Impacts of Seawater Rise on Seawater Intrusion in the Nile Delta Aquifer, Egypt, Groundwater, 20 52(2), 264–276, doi:10.1111/gwat.12058, 2014.

Sestini, G.: Nile Delta: a review of depositional environments and geological history, Geol. Soc. London, Spec. Publ., 41(1), 99–127, doi:10.1144/GSL.SP.1989.041.01.09, 1989.

Shen, C., Zhang, C., Xin, P., Kong, J. and Li, L.: Salt Dynamics in Coastal Marshes: Formation of Hypersaline Zones, Water Resour. Res., 54(5), 3259–3276, doi:10.1029/2017WR022021, 2018.

25 Sherif, M. M., Singh, V. P. and Amer, A. M.: A two-dimensional finite element model for dispersion (2D-FED) in coastal aquifers, J. Hydrol., 103(1–2), 11–36, doi:10.1016/0022-1694(88)90003-0, 1988.

Sinsakul, S.: Late Quaternary geology of the Lower Central Plain, Thailand, J. Asian Earth Sci., 18(4), 415–426, doi:10.1016/S1367-9120(99)00075-9, 2000.

Sivan, D., Wdowinski, S., Lambeck, K., Galili, E. and Raban, A.: Holocene sea-level changes along the Mediterranean coast 30 of Israel, based on archaeological observations and numerical model, Palaeogeogr. Palaeoclimatol. Palaeoecol., 167(1–2), 101–117, doi:10.1016/S0031-0182(00)00234-0, 2001.

Smith, B. F., Bjørstad, P. E. and Gropp, W. D.: Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations., Cambridge University Press, NY, USA., 1996.

Spratt, R. M. and Lisiecki, L. E.: A Late Pleistocene sea level stack, Clim. Past, 12(4), 1079–1092, doi:10.5194/cp-12-1079-

2016, 2016.

Stanley, D. J. and Warne, A. G.: Nile Delta: Recent Geological Evolution and Human Impact, Science (80-.)., 260(5108), 628–634, doi:10.1126/science.260.5108.628, 1993.

Stanley, J.-D. and Clemente, P. L.: Clay Distributions, Grain Sizes, Sediment Thicknesses, and Compaction Rates to
5   Interpret Subsidence in Egypt's Northern Nile Delta, J. Coast. Res., 30(1), 88–101, doi:10.2112/JCOASTRES-D-13-00146.1, 2014.

Stanley, J. D. and Clemente, P. L.: Increased land subsidence and sea-level rise are submerging Egypt's nile delta coastal margin, GSA Today, 27(5), 4–11, doi:10.1130/GSATG312A.1, 2017.

Stuyfzand, P. J.: Base exchange indices as indicators of salinization or freshening of (coastal) aquifers, [2]0th Salt Water
10   Intrusion Meet. Naples, Florida, USA, (1941), 262–265 [online] Available from: http://www.swim-site.nl/pdf/swim20/file281-284.pdf, 2008.

Summerhayes, C. P., Sestini, G., Misdorp, R. and Marks, N.: Nile Delta: Nature and Evolution of Continental Shelf Sediments, Mar. Geol., 27(3890), 43–65, 1978.

Surfsara: Description of the Cartesius system, [online] Available from:
15   https://userinfo.surfsara.nl/systems/cartesius/description (Accessed 14 May 2018), 2014.

Switzman, H., Coulibaly, P. and Adeel, Z.: Modeling the impacts of dryland agricultural reclamation on groundwater resources in Northern Egypt using sparse data, J. Hydrol., 520, 420–438, doi:10.1016/j.jhydrol.2014.10.064, 2015.

Tanabe, S., Saito, Y., Lan Vu, Q., Hanebuth, T. J. J., Lan Ngo, Q. and Kitamura, A.: Holocene evolution of the Song Hong (Red River) delta system, northern Vietnam, Sediment. Geol., 187(1–2), 29–61, doi:10.1016/j.sedgeo.2005.12.004, 2006.

20   The MPI Forum: MPI: A Message Passing Interface, 1993.

Timmerman, P. H. A. and Hemker, C. J.: van de uiterwaarden en polders nabij Langerak (Alblasserwaard), H20, 26(1), 1993.

Verkaik, J., Van Engelen, J., Huizer, S. and Oude Essink, G. H. P.: The New Parallel Krylov Solver for SEAWAT, in AGU Fall Meeting 2017., 2017.

25   Voss, C. I. and Soliman, S. M.: The transboundary non-renewable Nubian Aquifer System of Chad, Egypt, Libya and Sudan: classical groundwater questions and parsimonious hydrogeologic analysis and modeling, Hydrogeol. J., 22(2), 441–468, doi:10.1007/s10040-013-1039-3, 2013.

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E. and Yu, T.: scikit-image: image processing in Python, PeerJ, 2, e453, doi:10.7717/peerj.453, 2014.

30   Wang, Y. and Jiao, J. J.: Origin of groundwater salinity and hydrochemical processes in the confined Quaternary aquifer of the Pearl River Delta, China, J. Hydrol., 438–439, 112–124, doi:10.1016/j.jhydrol.2012.03.008, 2012.

Warne, A. G. and Stanley, D. J.: Late Quaternary Evolution of the Northwest Nile Delta and Adjacent Coast in the Alexandria Region, Egypt, J. Coast. Res., 9(1), 26–64, 1993.

WMO: Weather Information for Alexandria, [online] Available from: https://worldweather.wmo.int/059/c01268.htm

(Accessed 11 September 2018), 2006.

World Bank: Population, total, Databank [online] Available from: https://data.worldbank.org/indicator/SP.POP.TOTL?end=2016&locations=EG&start=1960&type=points&view=chart (Accessed 23 March 2018), 2018.

5 WRI: World Resources 2008 Data Tables, Washington D.C., 2008.

Yamanaka, T., Shimada, J., Tsujimura, M., Lorphensri, O., Mikita, M., Hagihara, A. and Onodera, S.: Tracing a confined groundwater flow system under the pressure of excessive groundwater use in the lower central plain, Thailand, Hydrol. Process., 25(17), 2654–2664, doi:10.1002/hyp.8007, 2011.

Yechieli, Y., Shalev, E., Wollman, S., Kiro, Y. and Kafri, U.: Response of the Mediterranean and Dead Sea coastal aquifers
10 to sea level variations, Water Resour. Res., 46(12), doi:10.1029/2009WR008708, 2010.

Zamrsky, D., Oude Essink, G. H. P. and Bierkens, M. F. P.: Estimating the thickness of unconsolidated coastal aquifers along the global coastline, Earth Syst. Sci. Data, 10(3), 1591–1603, doi:10.5194/essd-10-1591-2018, 2018.

Zech, A., Attinger, S., Cvetkovic, V., Dagan, G., Dietrich, P., Fiori, A., Rubin, Y. and Teutsch, G.: Is unique scaling of aquifer macrodispersivity supported by field data?, Water Resour. Res., 51, 7662–7679, doi:10.1002/ 2015WR017220, 2015.

15 Zhang, C., Li, L. and Lockington, D.: A physically based surface resistance model for evaporation from bare soils, Water Resour. Res., 51, 1084–1111, doi:10.1002/2014WR015490, 2015.

Zong, Y., Huang, G., Switzer, A. D., Yu, F. and Yim, W. W. S.: An evolutionary model for the Holocene formation of the Pearl River delta, China, Holocene, 19(1), 129–142, doi:10.1177/0959683608098957, 2009.

20

Table 1: Letters that are used to create model scenario codes.

Feature	Scenario member	Letter
Sea connection {*sea*}	Open	O
	Half-Open	H
	Closed	C
Horizontal clay layer {*clayer*}	No clay	N
	Fluvial clay	F
	Marine clay	M
HGw provenance {*prov*}	Top	T
	Bot	B
Temporal change {*temp*}	Palaeo	P
	Steady	S

Table 2: References to data used for spatially- and/or time varying input to the model.

Data	Temporal	Spatially	Reference	Remarks
Late Pleistocene sea-level (Eustatic)	yes	no	Spratt and Lisiecki (2016)	
Holocene sea-level (Mediterranean Sea)	yes	no	Sivan et al. (2001)	No detailed sea-levels for the Holocene exist for the Nile Delta, the closest data is from the Levant (Pennington, 2017)
Palaeosalinities lagoons	yes	no	Flaux et al. (2013)	Data for Maryut lagoon, taken as representative for other lagoons
Palaeogeography	yes	yes	Pennington et al. (2017); Stanley and Warne (1993)	Palaeogeography determines the spatial-temporal occurrence of the sea, dunes, rivers, and lagoons
Bathymetry	no	yes	GEBCO (2014)	

DEM	no	yes	NASA (2014)	
Groundwater extraction wells	yes	yes	Nofal et al. (2018)	
Geological borelogs	no	yes	Nofal et al. (2016)	

Table 3: Parameters fixed throughout simulation.

Parameter	Value	Unit	Reference	Remarks
Effective porosity	0.25	-	Mabrouk et al. (2018)	
Longitudinal dispersivity	2	m	Oude Essink et al. (2010)	Value for Rhine Meuse Delta, a similar system
Transverse dispersivity	0.2	m	Gelhar et al. (1992)	Reference used for anisotropy between transverse and longitudinal dispersivity
Vertical dispersivity	0.02	m	Gelhar et al. (1992)	Reference used for anisotropy between vertical and longitudinal dispersivity
Molecular diffusion coefficient	8.64e-5	m2/d	-	
Horizontal hydraulic conductivity Holocene fluvial alluvial clay	0.2	m/d	Gallichand et al. (1992)	Average value taken here
Horizontal hydraulic conductivity Quaternary sands	75	m/d	Barrocu and Dahab (2010); Laeven (1991); Mabrouk et al. (2013)	
Horizontal hydraulic conductivity Quaternary marine clays	1e-4	m/d	Domenico and Schwartz (1990)	Upper limit taken here, as these values stem from lab scale measurements and we model on a large scale.
Horizontal hydraulic conductivity Quaternary clayey sands	75	m/d	Farid (1980)	Author states that kh clayey sands is close to kh of sands
Anisotropy (horizontal divided by the vertical hydraulic conductivity)	10 (100 for Holo clay)	-	Farid (1980)	
Resistance surface water beds	100	d	De Lange et al. (2014); Timmerman and Hemker (1993)	Value for Rhine Meuse Delta, a similar system
River stage apex delta	15	m	NASA (2014)	
Salinity sea	35	g TDS/L	-	

Salinity sabkha	120	g TDS/L	Van Engelen et al. (2018)	Based on previous model in reference
Salinity inflow from hydrogeological base	120	g TDS/L	Van Engelen et al. (2018)	Based on previous model in reference
Flux from hydrogeological base	6e-8	d^{-1}	Van Engelen et al. (2018)	Based on previous model in reference
Recharge in dunes	5.5e-4	m/d	WMO (2006)	Precipitation near Alexandria

Table 4: Time to a reach steady state for all equivalent steady-state scenarios.

Code	Time (years)
C-F-B-S	35000.0
C-M-B-S	90500.0
C-N-B-S	23500.0
C-N-T-S	22500.0
O-N-T-S	5500.0
O-F-T-S	10500.0
O-M-T-S	10000.0
H-F-B-S	37500.0
H-F-T-S	7500.0
H-M-B-S	44500.0
H-M-T-S	10000.0
H-N-B-S	19500.0
H-N-T-S	5500.0

Table 45: End state FGw fresh groundwater volumes for behavioural acceptable model scenarios.

Code	Total fresh groundwater (km3)	Fresh groundwater < 300 m (km3)	Fresh groundwater < 300 m and onshore (km3)
H-M-T-P	1974	1853	1829
C-M-B-P	2659	2325	1989
C-N-T-P	1526	1512	1511
H-N-T-P	1728	1642	1641
H-F-T-P	1765	1670	1668

8 ka

Salinity

TDS (g/L)
- △ 0 - 1
- △ 1 - 5
- △ 5 - 15
- △ 15 - 25
- △ 25 - 35
- △ 35 - 55
- △ 55 - 75
- △ 75 - 100

Depth (m)
- △ 50
- △ 100
- △ 150
- △ 200
- △ 250

Data Source
- ● 80's campaign in Laeven (1991)
- ■ 90's campaign in Laeven (1991)
- ▲ RIGW (2013)
- ⬠ Geriesh (2015)
- ◣ Salem (2016)

7.5 - 6 ka

Maryut lagoon

Paleogeography
- ▢ Lagoons
- ▢ Coastal Marshes
- ▢ Holocene Clay (Bilqas 2)
- ▢ Coastal barrier-breaches/Strand plains
- ▢ Pleistocene Sands (Mit Ghamr & Gezira cover)
- ▢ Sabhkas & Carbonate ridges
- ▢ Wadi El-Natrun
- ⋯ Current Coastline

Figure 1: Palaeogeography and groundwater salinity measurements up to 250 m depth of the Nile Delta. Palaeogeographical data from Pennington et al. (2017). <u>**The inset shows the location of the delta in Egypt.**</u>
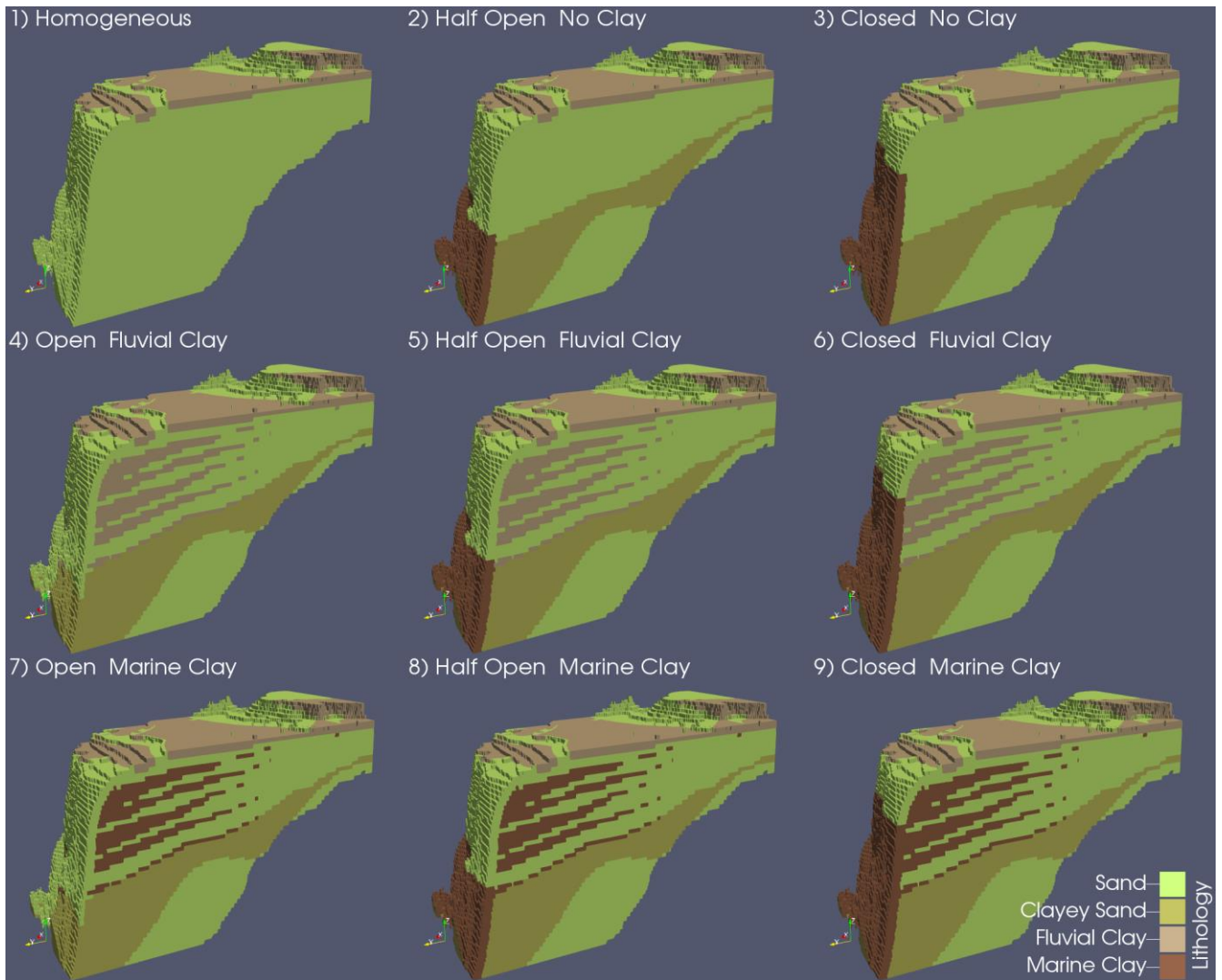
Figure 2: Snapshots of the lithological model scenarios that are used as input for the numerical variable-density groundwater flow model. The eastern half of the model is plotted here (See Fig. A1 for extent). The first word refers to the connection to the sea of the deeper part of the NDA, the second word to the assigned hydraulic conductivity of the onshore-reaching clay layers.

5
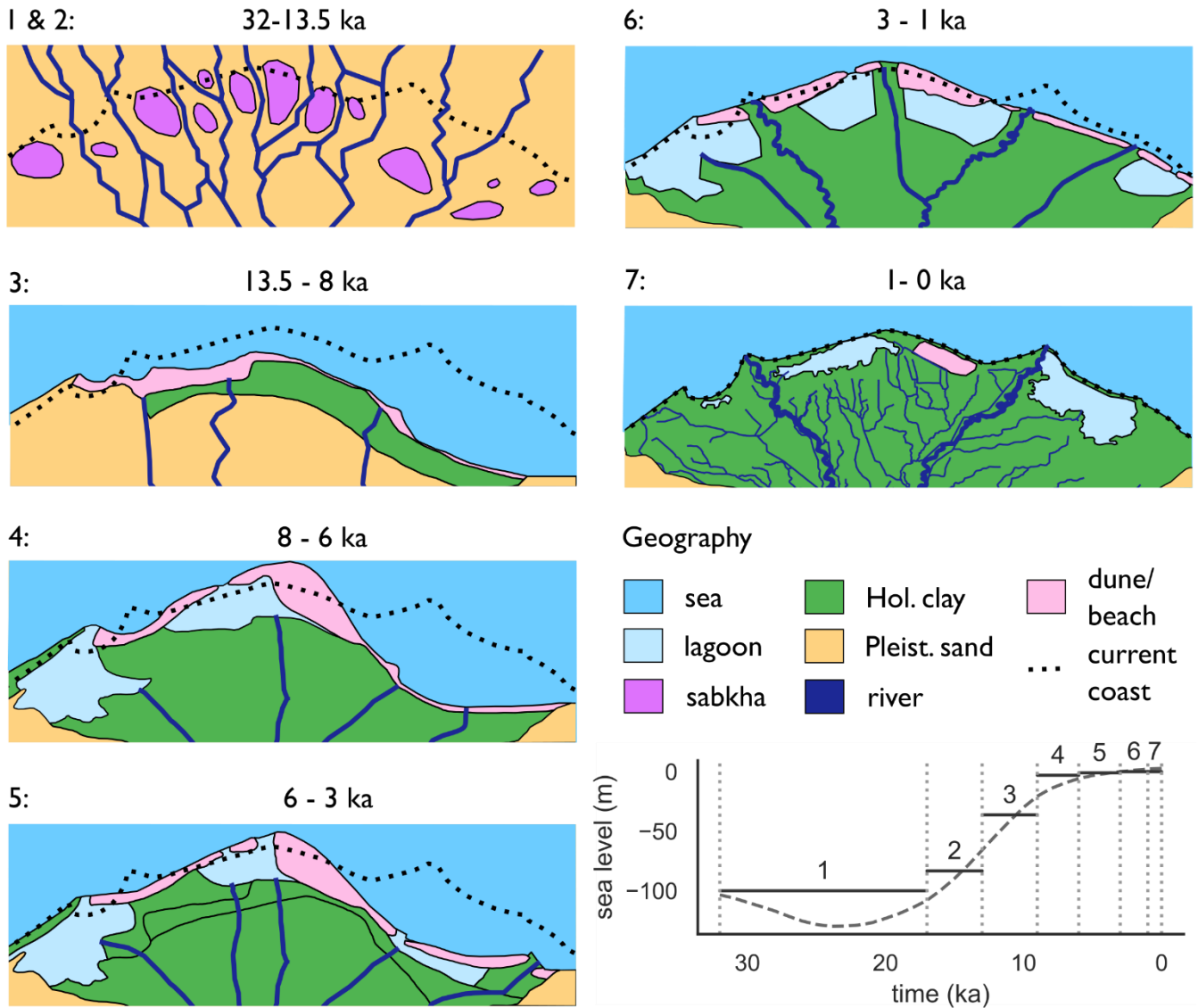
Figure 3: Time slices with simplified palaeogeography used as boundary conditions. Figures modified after (Pennington et al., 2017; Stanley and Warne, 1993). The panels focus on the area around the present-day coastline (see Fig. A1 for extent), where upper boundary conditions varied the most. Actual model domain extended beyond the extents of these panels. In the bottom right: The median eustatic sea-level curve (Spratt and Lisiecki, 2016) and the selected sea-level for each time slice.
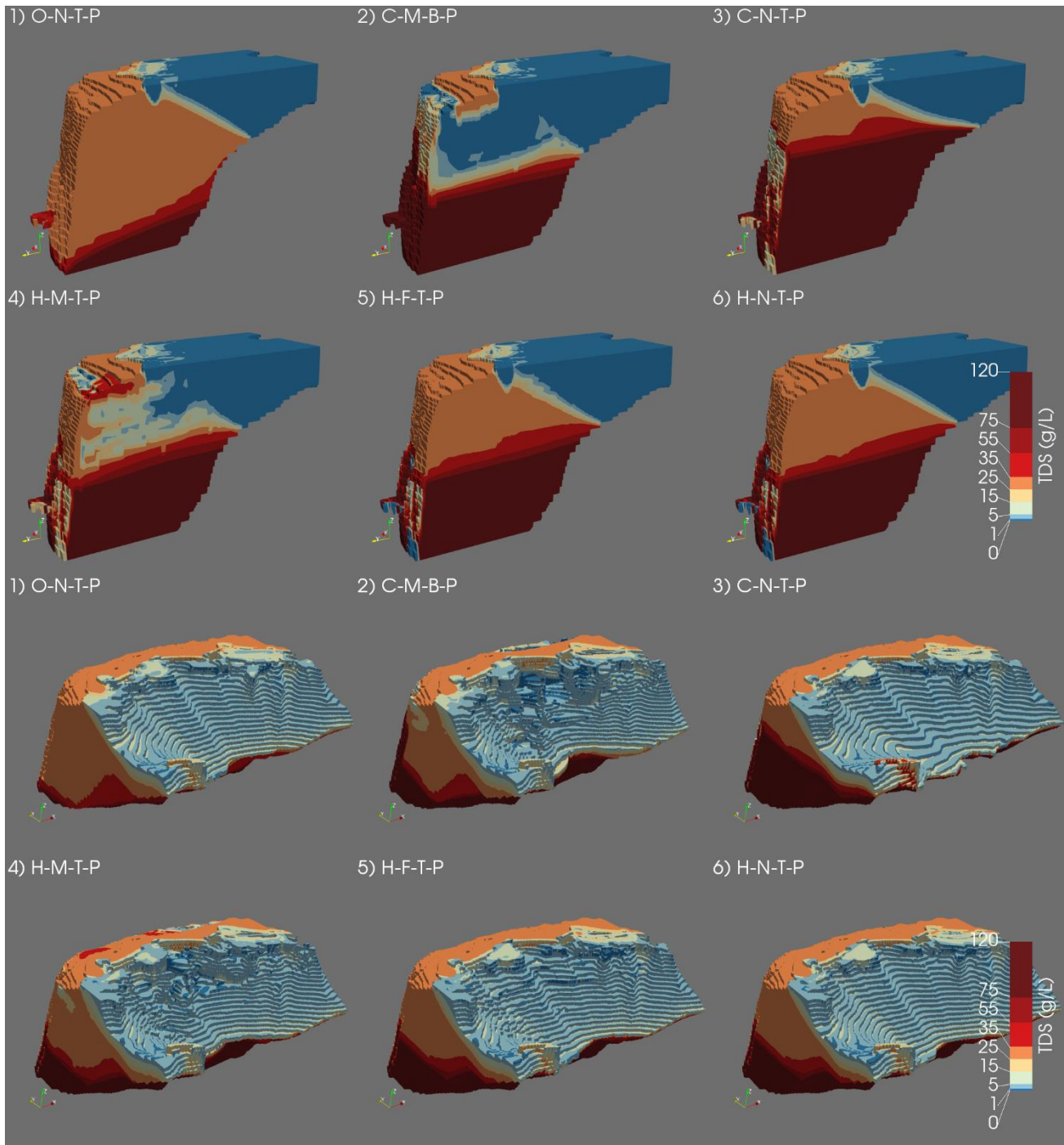
Figure 4: Salinity distributions of the five selected model scenarios after the validation; the results of the homogeneous model (viz. O-N-T-P) are also added as first plot for reference. Note that the results are stretched in the z-direction by a factor 150. In the top six plots the salinity distribution is sliced in half over the model domain and camera is pointed in the south-east direction. In the bottom six plots the full delta is visible, but the fresh groundwater is made fully transparent. Camera is pointed in downwards, in north-easterly direction.
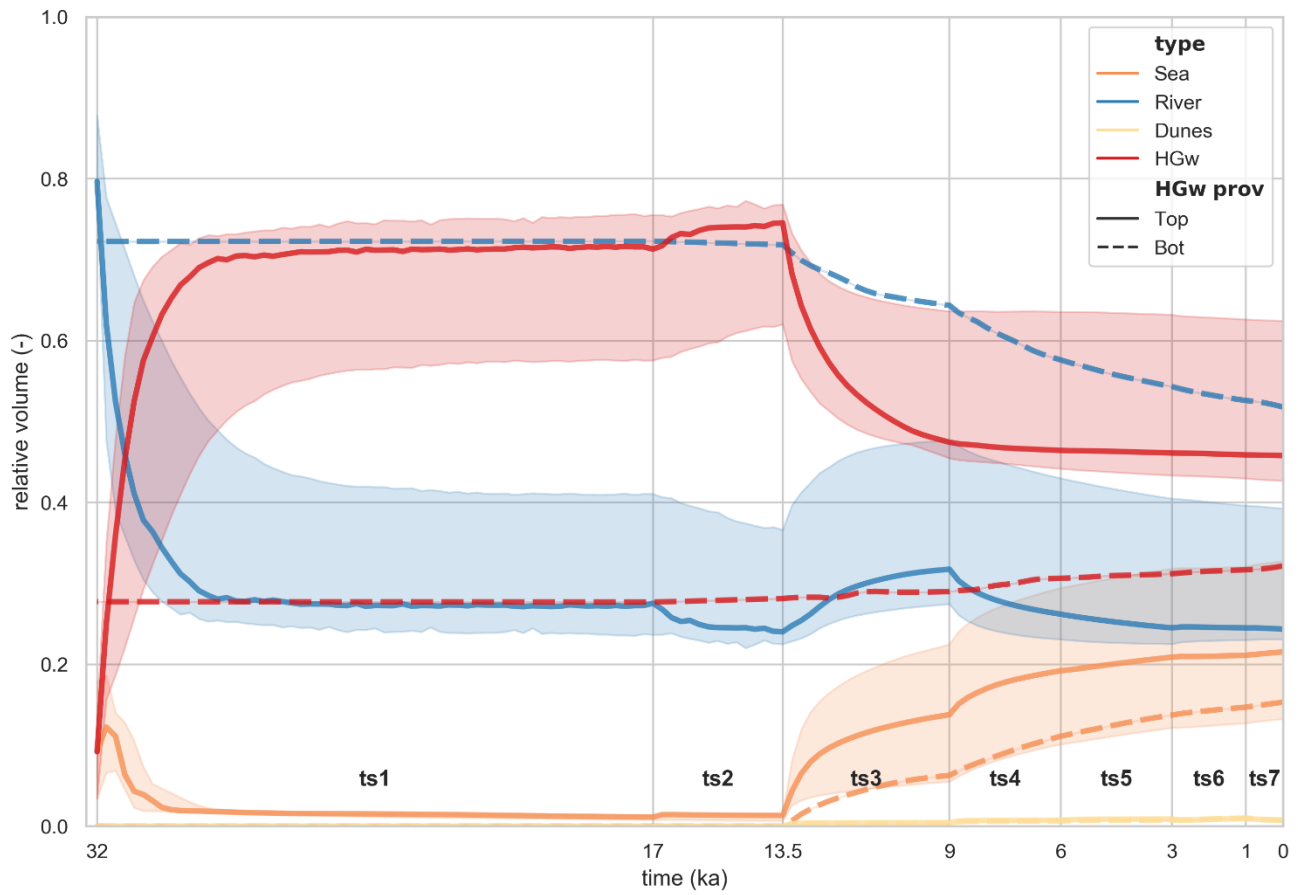
Figure 5: Groundwater provenance of different water types as fraction of the total modelling domain for the five acceptable model scenarios. The dashed lines represent the C-M-B-P model. Shaded area indicates the range of the acceptable T-model scenarios (C-N-T-P, H-N-T-P, H-F-T-P, H-M-T-P), the thick line their median. Time slice numbers are indicated in bold, preceded by "ts".

5

Figure 6: Goodness of fit boxplots for all palaeohydrogeological -reconstructions, binned in to four salinity classes. The higher the value of Λ, the worse the fit. Codes indicate model scenarios (Table 1). TDS values are binned in the classes [0, 1], [1, 5], [5, 35], [35, 100] g TDS/L for respectively "fresh", "brackish", "saline", "hypersaline", these are respectively plotted in panels A, B, C, D. "n" indicates the amount of observations available in the TDS bin to evaluate model results with. Diamonds indicate outliers, defined as values separated from the first or third quartile at 1.5 times the interquartile range. When no box is plotted for a scenario, 75% of the measurements equal zero, which consequently causes the interquartile range to be zero, rendering every non-zero value an outlier.

Figure 5: Difference in goodness of fit between palaeo-reconstruction and equivalent steady-state model scenarios. Codes indicate model scenarios. TDS values are binned in the [0, 1], [1, 5], [5, 35], [35, 100] g TDS/L for respectively "fresh", "brackish", "saline", "hypersaline". Diamonds indicate outliers, defined as values separated from the first or third quartile at 1.5 times the interquartile range. When no box is plotted for a scenario, 75% of the measurements equal zero, which consequently causes the interquartile range to be zero, rendering every non-zero value an outlier. Note that the hypersaline class was left out for the T-model scenarios, as these had no steady-state equivalent with hypersaline groundwater.

1) O-N-T-P   2) C-M-B-P   3) C-N-T-P

4) H-M-T-P   5) H-F-T-P   6) H-N-T-P

1) O-N-T-P   2) C-M-B-P   3) C-N-T-P

4) H-M-T-P   5) H-F-T-P   6) H-N-T-P
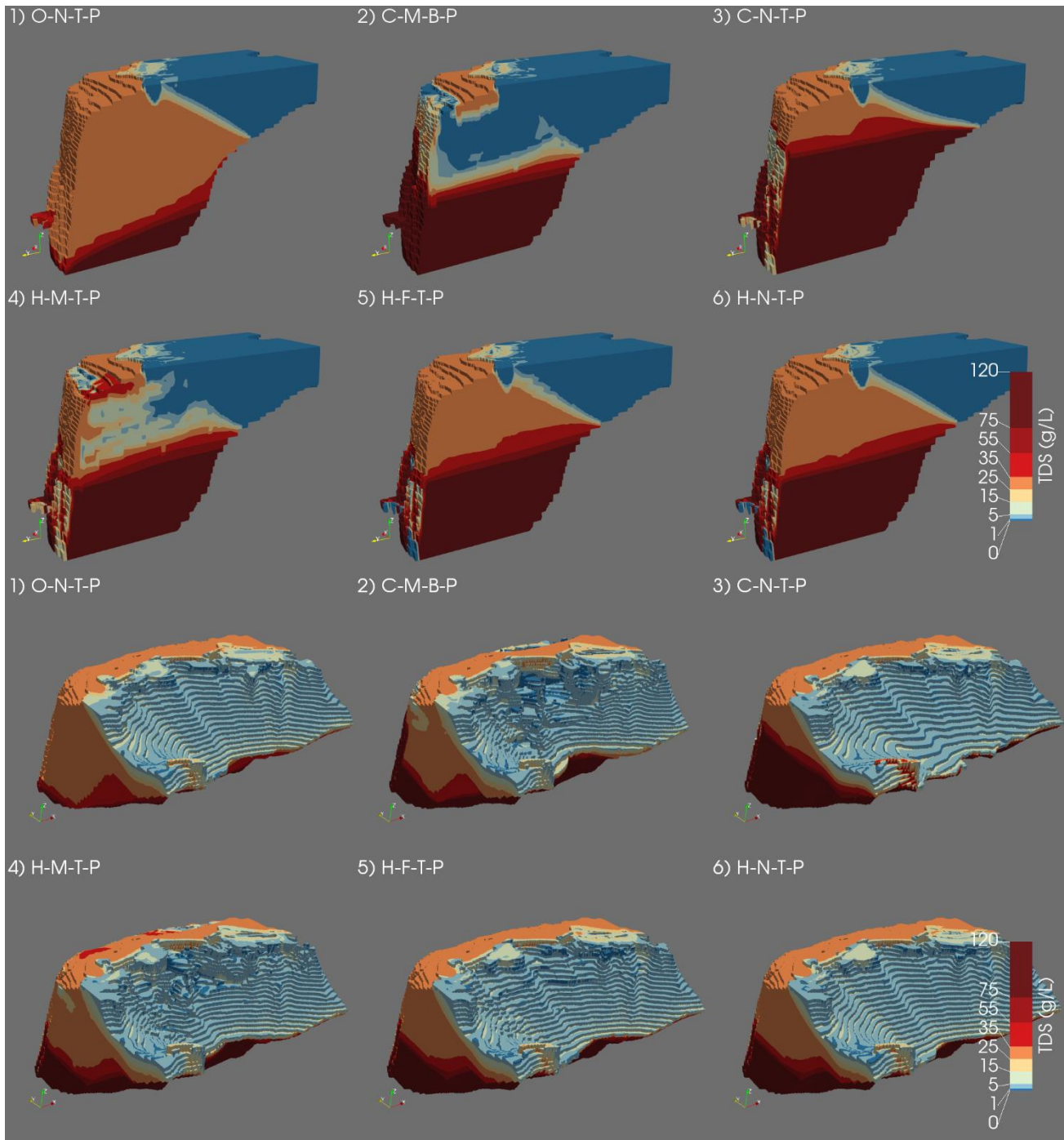
Figure 6: Salinity distributions of the five selected model scenarios after the validation; the results of the homogeneous model (viz. O-N-T-P) are also added as first plot for reference. Note that the results are stretched in the z-direction by a factor 150. In the top six plots the salinity distribution is sliced in half over the model domain and camera is pointed in the south-east direction. In the bottom six plots the full delta is visible, but the fresh groundwater is made fully transparent. Camera is pointed in downwards, in north-easterly direction.

67
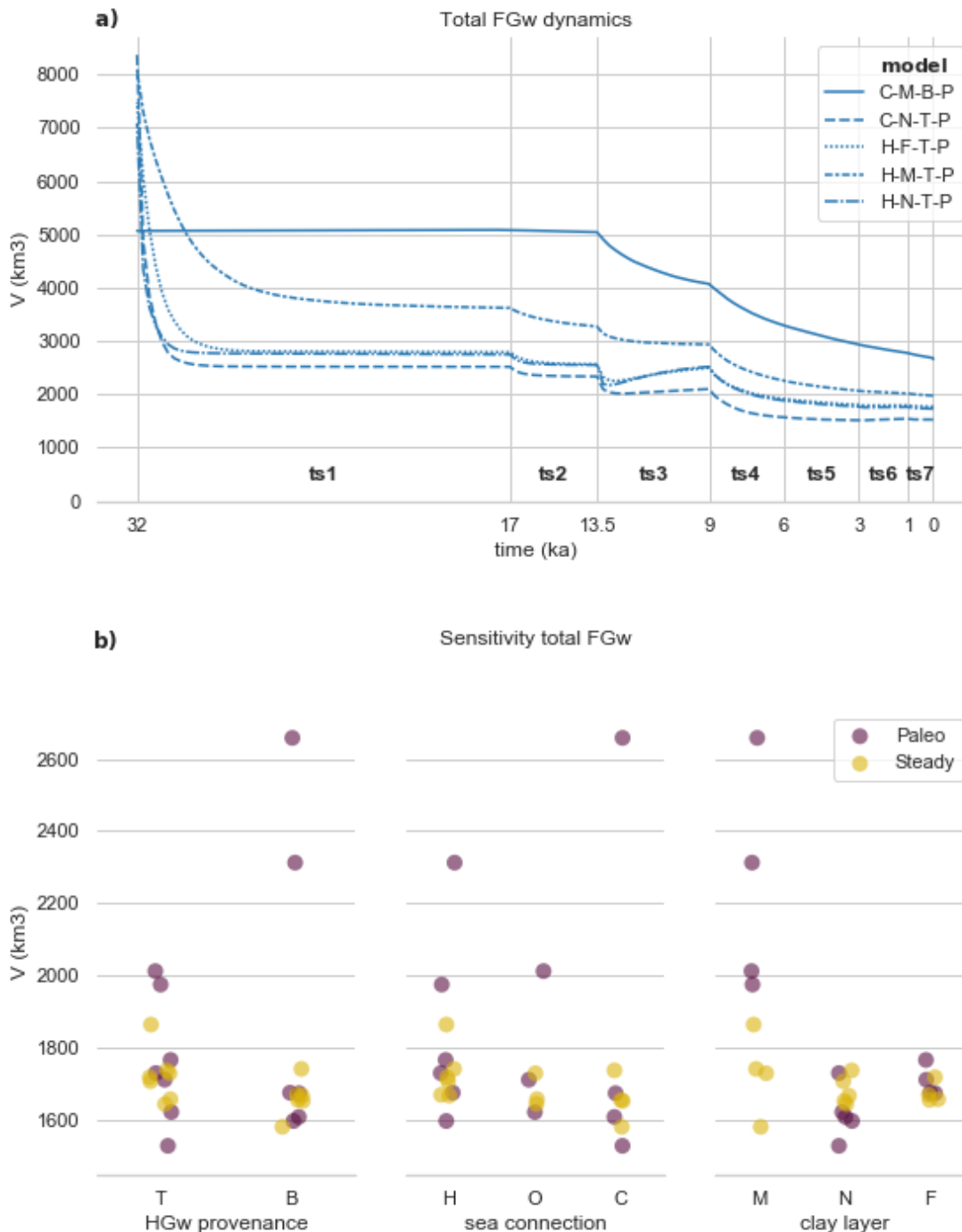
Figure 7: (a) the development of the total ~~fresh groundwater~~FGw volume through time for the five ~~behavioural~~ acceptable model scenarios. Time slice~~Timeslice~~ numbers are indicated in bold with "ts". **(b)** -End state total FGw ~~fresh groundwater~~ volumes of all model scenarios, grouped for different inputs. Colours of the dots indicate a steady -state or palaeohydrogeological -reconstruction result. Dots are jittered to better show overlapping dots.

68
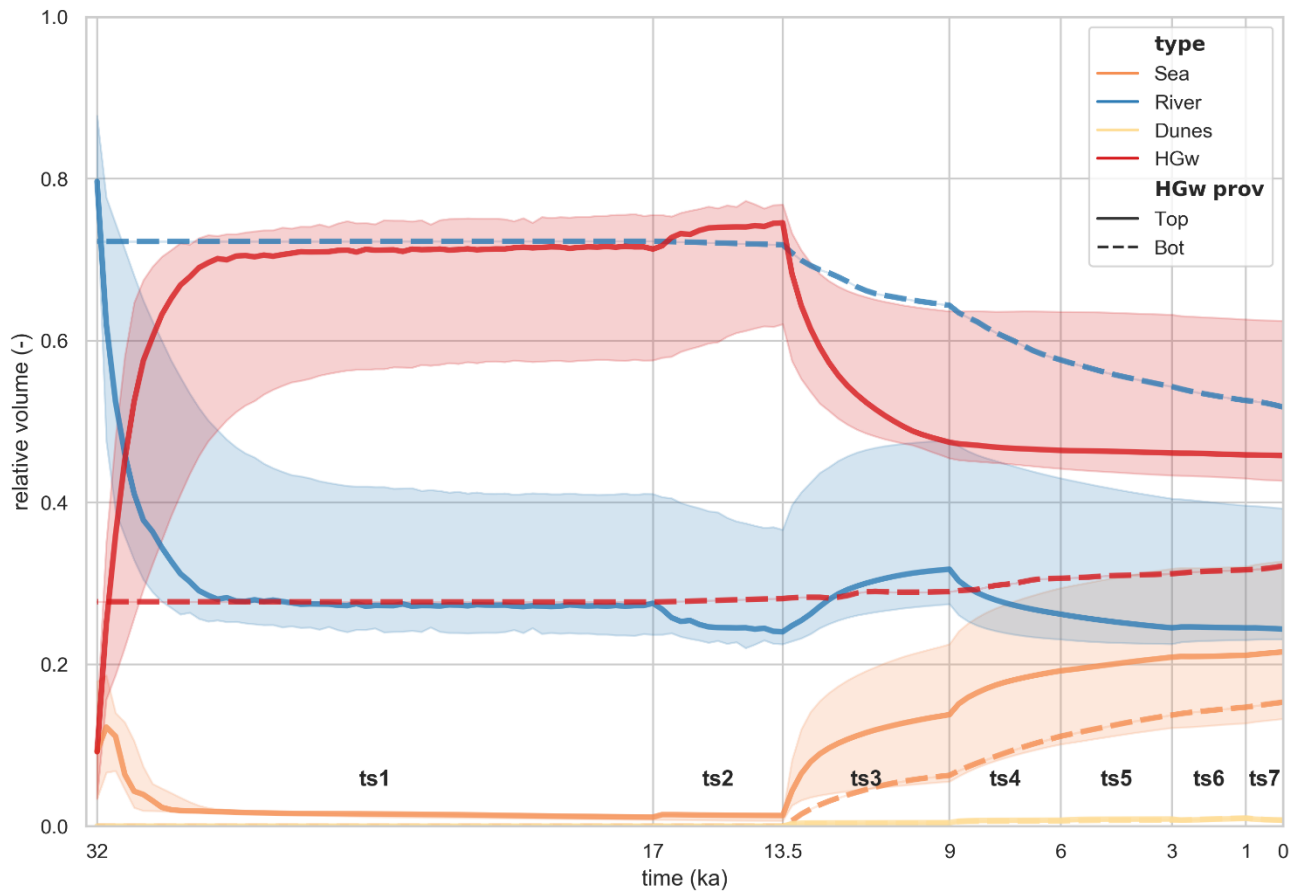
Figure 8: Groundwater provenance of different water types as fraction of the total modelling domain for the five behavioural model scenarios. The dashed lines represent the C-M-B-P model. Shaded area indicates the range of the behavioural T-model scenarios (C-N-T-P, H-N-T-P, H-F-T-P, H-M-T-P), the thick line their median. Timeslice numbers are indicated in bold, preceded by "ts".
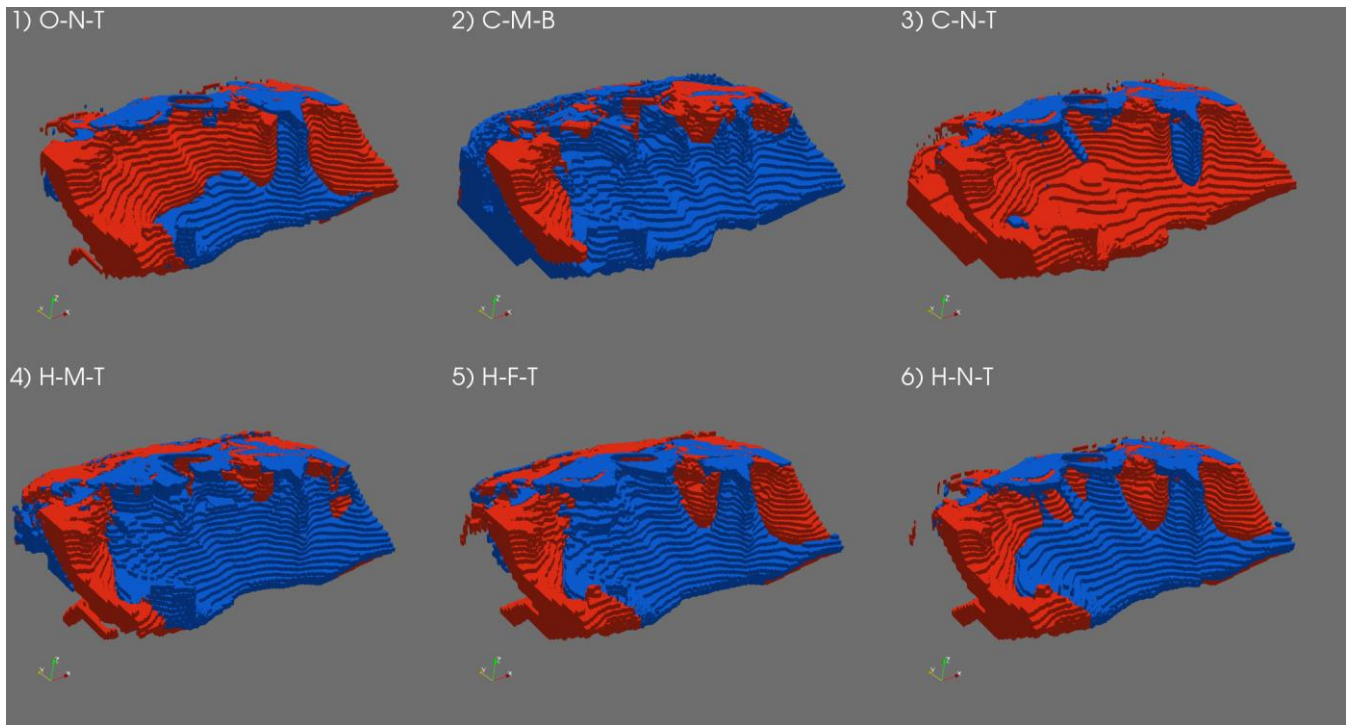
Figure 8: Comparison between the palaeohydrogeological- reconstructions and the steady-state model scenarios. ~~Everywhere~~ Red colours indicate ~~that red,~~ $\Delta C > 1$, meaning that the palaeohydrogeological -reconstruction is saltier than its steady-state equivalent. Vice versa for blue, where $\Delta C < -1$. All values in between -1 and 1 are made transparent to focus purely on major differences. Groundwater below 450 m depth is not plotted, as the present-day fresh-salt interface was not located below these depths. Camera angle is the same as in the six bottom plots of Fig. ~~6~~4.
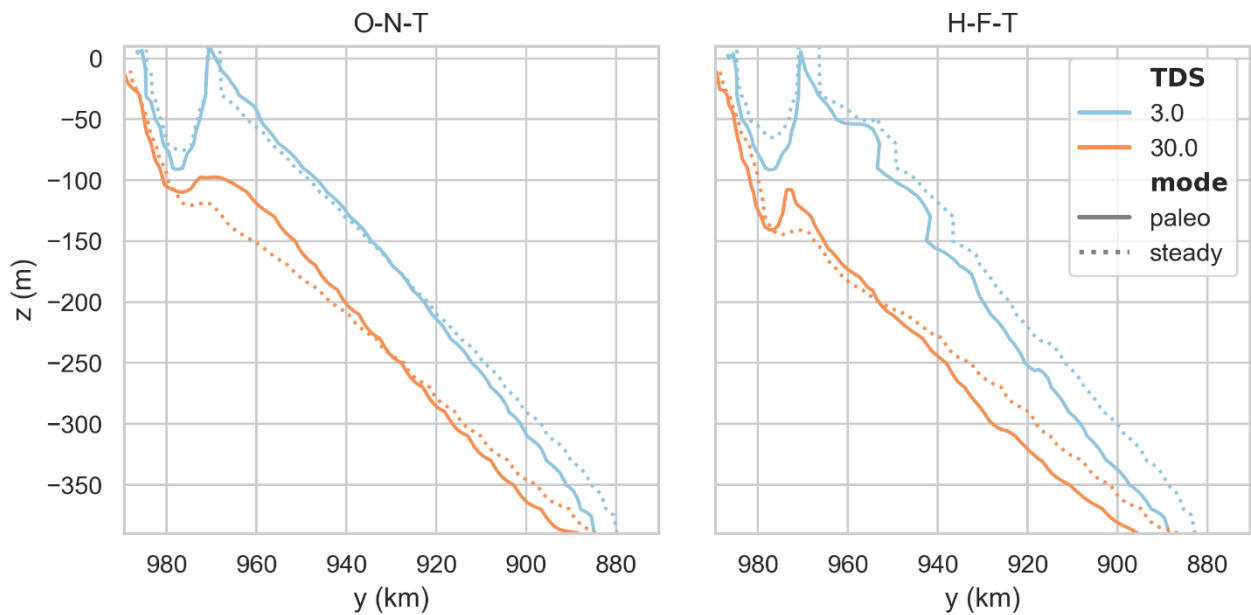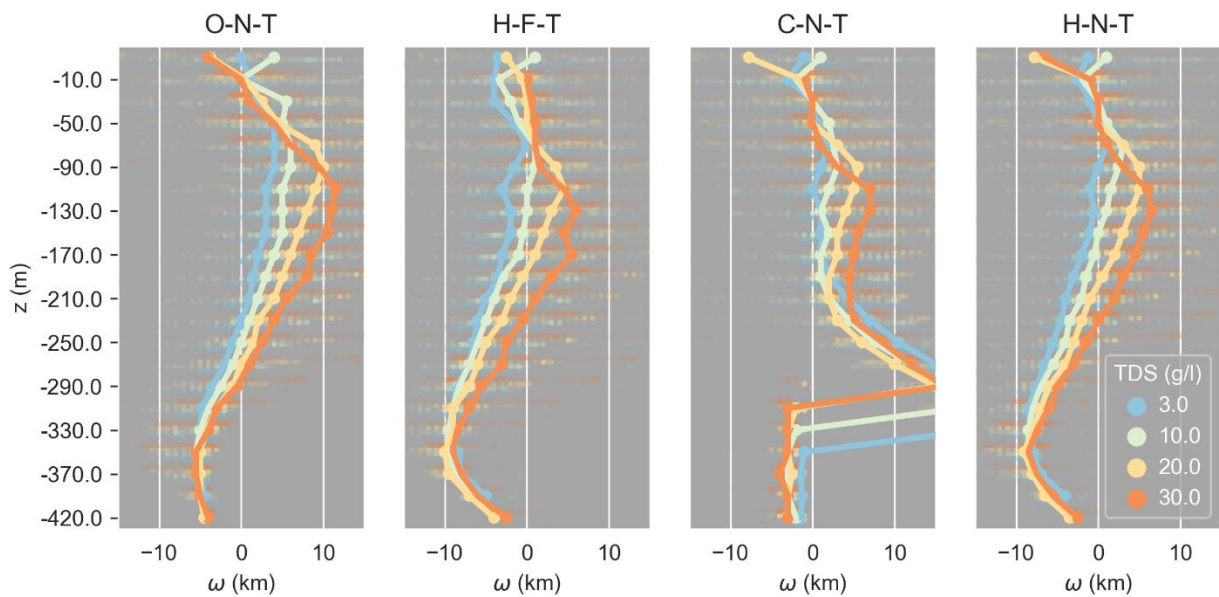
5   **Figure 9: Isosurface separation ω with depth for different TDS values and selected model scenarios. The line indicates the median isosurface separation, the near transparent dots all data points. Positive ω means the isohaline of the palaeohydrogeological reconstruction is located more ~~landinwards~~land inwards than its' steady-state equivalent and vice versa.**
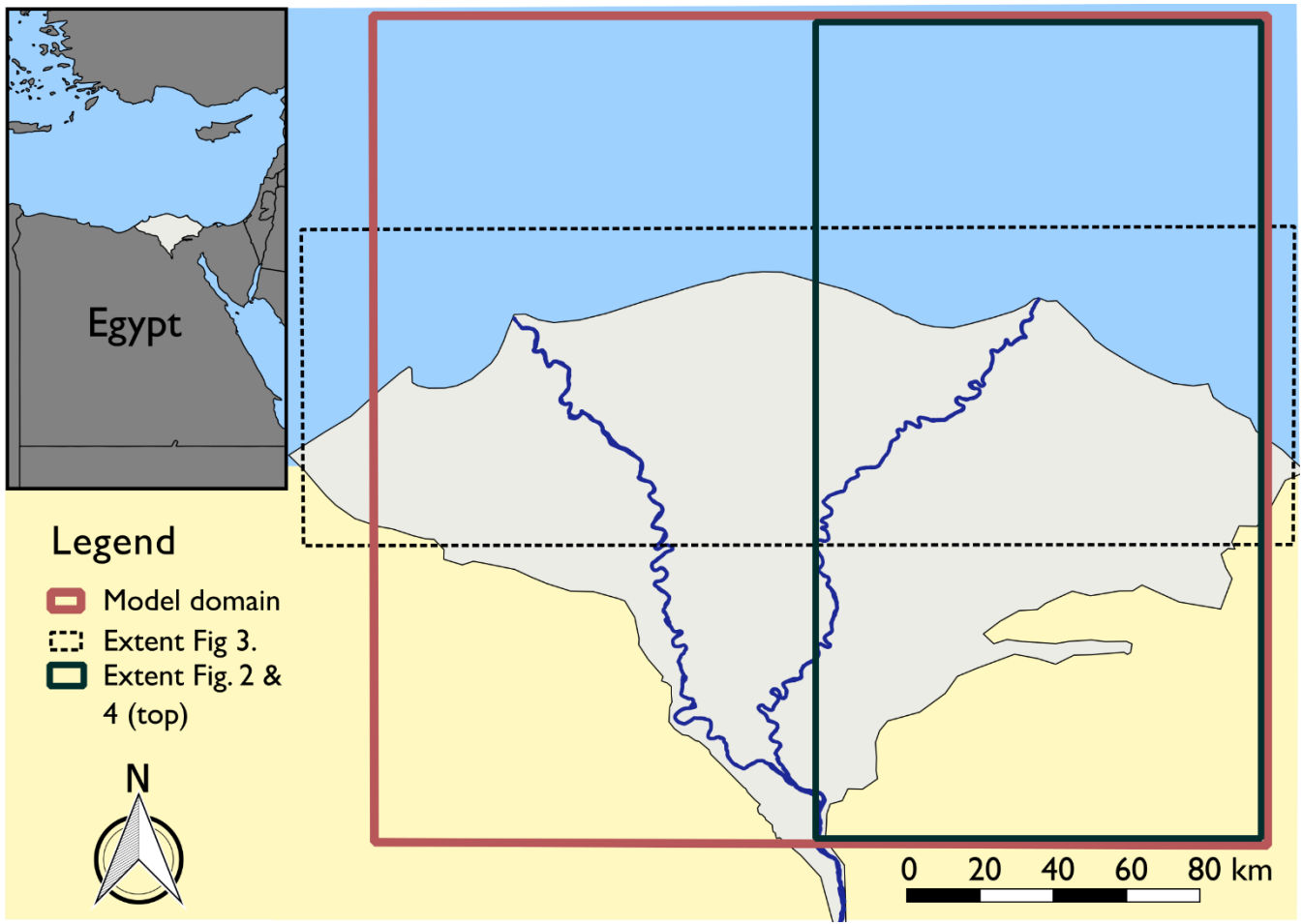
Figure A1: Overview of the Nile Delta with boxes indicating the model domain and the extents of several figures. The main Nile river branches are plotted as a geographical reference. The red rectangle shows the extent of our model domain, the dark green box the plotted area in Figure 2 and the top half of Figure 4, and the dotted box the plotted area in Figure 3. The inset on the upper left corner shows the location of the Nile Delta in Egypt.
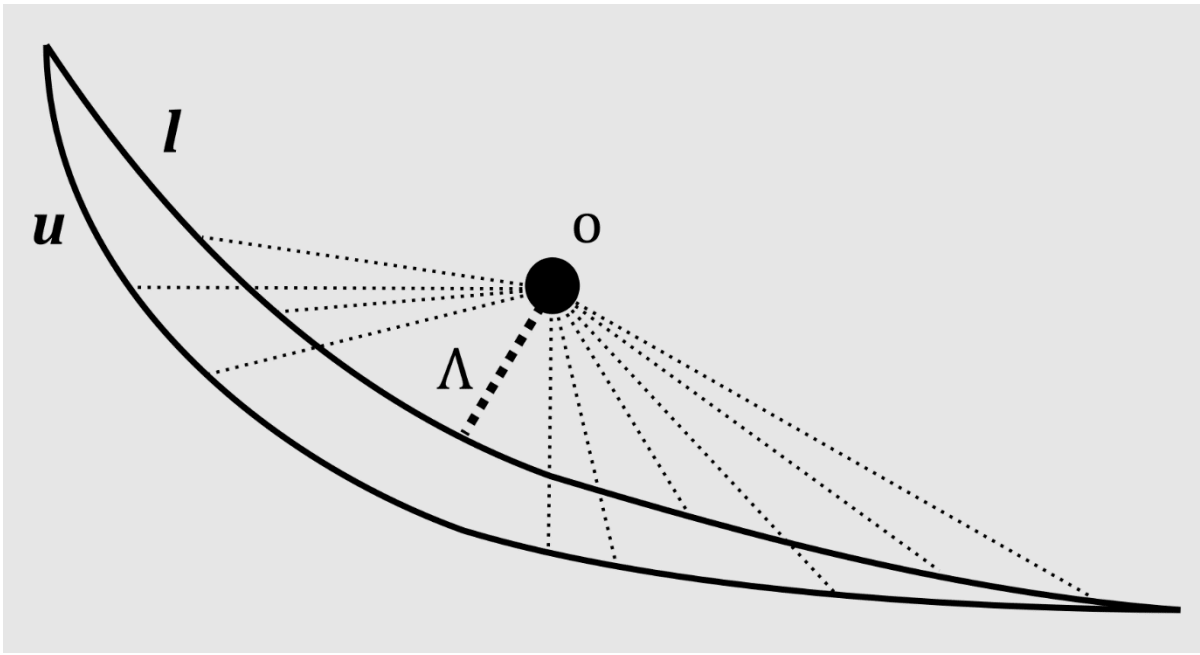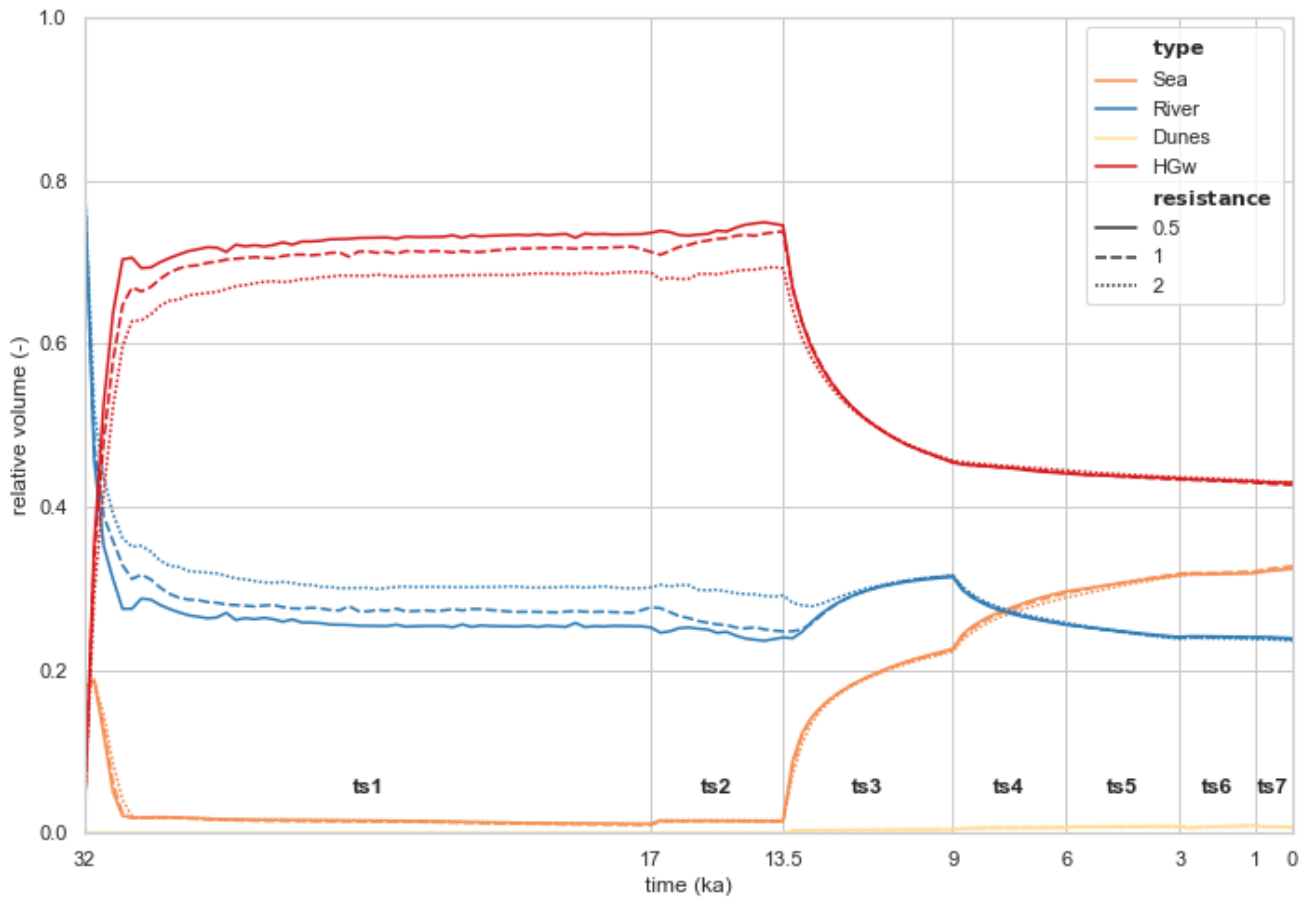
Figure A2: Sketch of how Λ is determined with equation 1 for a 2-dimensional example. The circle represents the location of the observation. l **and** u**, the isosurfaces of the lower and upper bound, respectively. The thick dotted line represents the minimum Eucledian distance to an isosurface bound, Λ. The thinner dotted lines some other Euclidian distances.**

Figure B1: Groundwater origins for scenario "H-N-T-P" with three resistances. The colour indicates the groundwater type, the linestyle the value the original resistance is multiplied with.