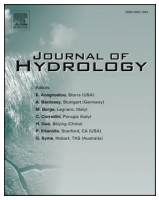
Journal of Hydrology 643 (2024) 131915

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol



Research papers

Effects of grid resolution on regional modelled groundwater salinity and salt fluxes to surface water

Ignacio Farías [a,*], Gualbert H.P. Oude Essink [a,b], Perry G.B. de Louw [b,c], Marc F.P. Bierkens [b,a]

[a] Department of Physical Geography, Utrecht University, PO Box 80125, 3508 TC Utrecht, The Netherlands
[b] Unit Subsurface and Groundwater Systems, Deltares, P.O. Box 85467, 3508 AL, Utrecht, The Netherlands
[c] Soil Physics and Land Management Group, Wageningen University, P.O. Box 47, 6700 AA Wageningen, the Netherlands

ARTICLE INFO

This manuscript was handled by Huaming Guo, Editor-in-Chief

Keywords:
Surface-groundwater interactions
Salt loads
Resolution
Coastal aquifers
Parallel computing

ABSTRACT

Coastal fresh groundwaters face growing threats from rising withdrawals and climate change, with salinization of surface and groundwater as notable impacts. Recent developments in parallel variable-density modelling enable studying these threats at unexplored resolutions and extents. To improve the understanding of ground- and surface water salinization processes, we designed an experiment to measure how spatial resolution affects both groundwater salinity distribution and salt loads to surface waters. An existing parallelized coupled variable-density groundwater flow and salt transport (VD-GWT) model is applied to a region in the Netherlands whereby the surface water drainage network is parameterized at grid resolutions between 10 and 250 m. The results show a strong dependency between resolution and the salinity distribution in shallow groundwater while simulated salt loads are affected by imperfect flux scaling. The computational resources needed for this test suggest that regional high-resolution VD-GWT models are feasible using HPC environments, albeit not practical.

1. Introduction

Over the past decades, fresh groundwater and surface water resources have and will undergo an unprecedented shift in their availability due to climate change and increasing groundwater withdrawal (Forzieri et al., 2014; McDonald et al., 2011; Oude Essink et al., 2010). These changes have significant consequences, particularly in coastal areas like The Netherlands, a low-lying deltaic region highly susceptible to sea-level rise (Faneca Sànchez et al., 2012; Haasnoot et al., 2020; Kwadijk et al., 2010; Sathish et al., 2022), as well as water scarcity during dry seasons (Koopman et al., 2019). In these areas, fresh groundwater and surface water availability are increasingly threatened by saltwater intrusion (J. R. Delsman, 2015; Loáiciga et al., 2012; Michael et al., 2013; Rasmussen et al., 2013; Zamrsky et al., 2024).

To understand, predict and project these changes and their impacts, numerical modeling is commonly employed, specifically coupled variable density groundwater flow and salt transport (VD-GWT) models. However, VD-GWT models introduce computational complexities that often result in longer runtimes compared to traditional groundwater flow models. To address these challenges, simplifications are frequently made in either spatial or temporal resolution or process representation to manage the computational overhead (Werner et al., 2013). This trade-off between spatial resolution, runtimes, and accuracy raises the question of how different grid resolutions and associated scaling approaches for boundary conditions may impact model outcomes (Noorduijn et al., 2021; Vermeulen et al., 2006; Yu & Michael, 2022).

One of the incentives to use VD-GWT models is to predict the spatial change in salt mass to the surface water network by exfiltrating groundwater. This mass flux is commonly understood as the salt loads to the surface which are driven by both salinity and pressure gradients (De Louw et al., 2011a), as well as how the groundwater-surface water (GW-SW) interactions have been conceptualized. Understanding this process is paramount in irrigated areas with shallow groundwater levels as land and surface water salinization could affect agriculture, drinking water safety and others (Bailey & Hosseini, 2023; Christen & Skehan, 2001). In this regard, The Netherlands is particularly vulnerable as the water management strategy of the country relies on periodical flushing of the surface water system to control its salinity levels (Raats, 2015; Van Rees Vellinga et al., 1981). Projections show that due to climate change and sea level rise, in the future surface water salinization will increase, likely increasing the need for flushing with a simultaneous decrease in summer freshwater availability, generating an increased water gap (Mens et al., 2024). This further highlights the need for accurate estimates of these GW-SW interactions.

* Corresponding author at: Department of Physical Geography, Utrecht University, PO Box 80125, 3508 TC Utrecht, The Netherlands.
E-mail address: i.e.fariasgutierrez@uu.nl (I. Farías).

https://doi.org/10.1016/j.jhydrol.2024.131915

Available online 1 September 2024

The MODFLOW family of groundwater modelling and transport software (Harbaugh, 2005; Zheng & Wang, 1999) commonly quantifies GW-SW interactions via the conductance parameter, a physically based measure of exchange effectiveness between mediums. A suitable lumped parametrization of the conductance (also called "scaling") is needed to accommodate several surface water features in coarse grid resolutions. Examples of lumped parametrization schemes have been documented for densely drained regions as The Netherlands (De Lange, 1999; Pauw et al., 2015). Research has shown that when GW-SW interactions are dominant, the parametrization and scaling of surface water boundary conditions (SW-BC) outweigh the scaling of hydrogeological parameters (Vermeulen et al., 2006). One of these scaling methods was developed by De Lange (1999). Pauw et al. (2015) tested the effect of the SW-BC on fluxes and salt loads to the surface using a synthetic VD-GWT model. Their results validate the use of De Lange's methods, reinforcing the importance of advanced parametrization techniques to address scaling effects. Di Ciacca et al., (2019) presented a more generalized scaling approach that was tested using a real-world drainage network with varying horizontal and vertical discretization, albeit limited in size and only considering flow. Despite the numerous efforts documented in the literature, a clear gap is apparent since all cases use either a synthetic model or only a small-scale real-world test. The impact of resolution on SW-BC parametrization has yet to be tested comprehensively, especially in scenarios involving large-scale or densely interconnected surface water drainage networks with significant volume and mass exchanges.

Recent developments have successfully bridged the gap between resolution and runtimes through distributed memory parallelization on the SEAWAT numerical code (Verkaik et al., 2021). This allows for VD-GWT models with a high number of active cells ($>10^6$) while maintaining feasible runtimes (J. R. Delsman et al., 2023; Seibert et al., 2023; Van Engelen et al., 2019; Verkaik et al., 2024). Also, access to higher resolutions and computational resources currently allows for unexplored model parametrizations and sizes for which the expected improvement in results is not clear. The premise is that fine resolutions lead to higher accuracy of results, at the expense of longer model runtimes. We presume that at higher resolutions local processes are better represented, such as salinity upconing between ditches (Pauw et al., 2015). Inversely, coarser resolutions might oversimplify the GW-SW interactions as ditches, canals and rivers on dense drainage networks can be only a few meters apart. As such, the objective of this paper is to assess the effect of changing the spatial resolution (viz. grid size) on fluxes and salt loads to the surface of a dense drainage network at a regional scale. For this purpose, we built a fully scripted workflow in conjunction with a workflow manager that allows to cut and vary the grid size of a regional area of the existing national VD-GWT model of The Netherlands, originally at 250 m grid resolution (J. R. Delsman et al., 2023). This is complemented by a cell-size based redefinition of the drainage network conductance using a lumped parametrization scheme (De Lange, 1999).

The structure of this paper is as follows; section 2.1 describes the general model set up sub-divided into the surface water network (section 2.1.1) and the VD-GWT model used (section 2.1.2), followed by the area definition (section 2.2), particularities of the workflow focused on computational resources (section 2.3), and a description of the statistical indicators used for post-processing (section 2.4). We then compare the results of varied model runs at grid sizes ranging from 10 to 250 m (section 3) focusing on groundwater salinity distributions, salt loads and the consequences of using a scaled-based parametrization. Finally, in section 4, we discuss the benefits and drawbacks of changing the resolution in relation to presumed *accuracy, runtimes and its implications for water management.*

2. Methods

The experiment designed to assess resolution effects on salt loads was built upon the workflow and datasets created for the Dutch nation-wide coupled variable-density groundwater flow model (J. R. Delsman et al., 2023), referred to as LHM fresh-salt. We modified the workflow to enable flexible grid discretization and modelling extent. This was compounded with an existing workflow designed to discretize the surface water system of the Netherlands called "LHM topsysteem" (De Lange et al., 2014).

A model run starts with the selection of a specific area within The Netherlands and LHM fresh-salt. The area in shapefile format and the election of the horizontal grid-sizes to be tested are the main user inputs needed. This information is used to generate clipped LHM fresh-salt models with scaled surface water parameters. The vertical discretization is kept the same as that of LHM fresh-salt.

The modified workflow is run at 10-, 25-, 50-, 100- and 250-meter grid-resolutions for the extent covered by the water management area of the Delfland water board. The main model outputs analyzed are groundwater heads, groundwater salinity distribution and groundwater fluxes and salt loads to the surface water system.

2.1. Model set-up

2.1.1. Surface water system parametrization

The discretization of the surface water network focuses on waterways that are wide enough to be identified. The Netherlands (Stoter et al., 2014) mapped the position of rivers, canals and ditches and generated vectorial datasets that hold positions and a general classification of each feature. This classification is based on feature width; the primary system considers waterways of 3 m width and larger, the secondary system widths between 1 and 3 m and tertiary comprises the rest of the smaller surface water features. A secondary classification is also available that groups the shallow subsurface into "hydro types" according to their geohydrological properties (Massop et al., 2006). The combination of both categorizations yields values of width, depth and seasonal water stages (summer and winter) for each surface water feature. Note that in the western part of the Netherlands in so-called "polder areas" water stages are artificially controlled by pumps and kept at predefined levels.

With the vectorial information classified and spatially located, a buffer is built around each linear feature using its width. Vectorial features that cover an area (polygons) such as ponds, lakes and polders are not buffered and are assigned to the primary water system. The buffered linear features and areal features are then rasterized at a 0.1-meter pixel size, with each pixel inheriting the properties of its vectorial source. The 0.1 m-pixels are aggregated to 5-meter cells returning as output the areas covered by each water system at this resolution. The aggregation can then be repeated to obtain upscaled parameter sets at any resolution that is a multiple of 5.

The main output from the surface water discretization is the size of the surface water features in each cell at different grid resolutions (Fig. 1). With these geometrical parameters, the conductance parameter is calculated to characterize the flow exchange between the subsurface and the surface at 5 m resolution. We consider two ways of calculating the conductance; one corresponds to the classic MODFLOW conceptualization (Harbaugh, 2005) and the other is a lumped parametrization method that accounts for feature overlap on aquifers overlaying an aquitard and which we refer to as De Lange's method (De Lange, 1999).

We will refer to the classic MODFLOW conductance method as "Standard MODFLOW". This method conceptualizes the conductance parameter as a vertical resistance (c) over a transverse area (A_T):

$$C_{MODFLOW} = \frac{A_T}{c} = (L^*W)^*\frac{K_v}{D}$$

where "L" [m] represents the length of the feature, "W" [m] is the width, D [m] is the bed depth, and K_v [m/d] is the vertical hydraulic conductivity of the bed. In our scenario, where multiple surface water features are overlapping, we consider A_T as the area sum of all the features within

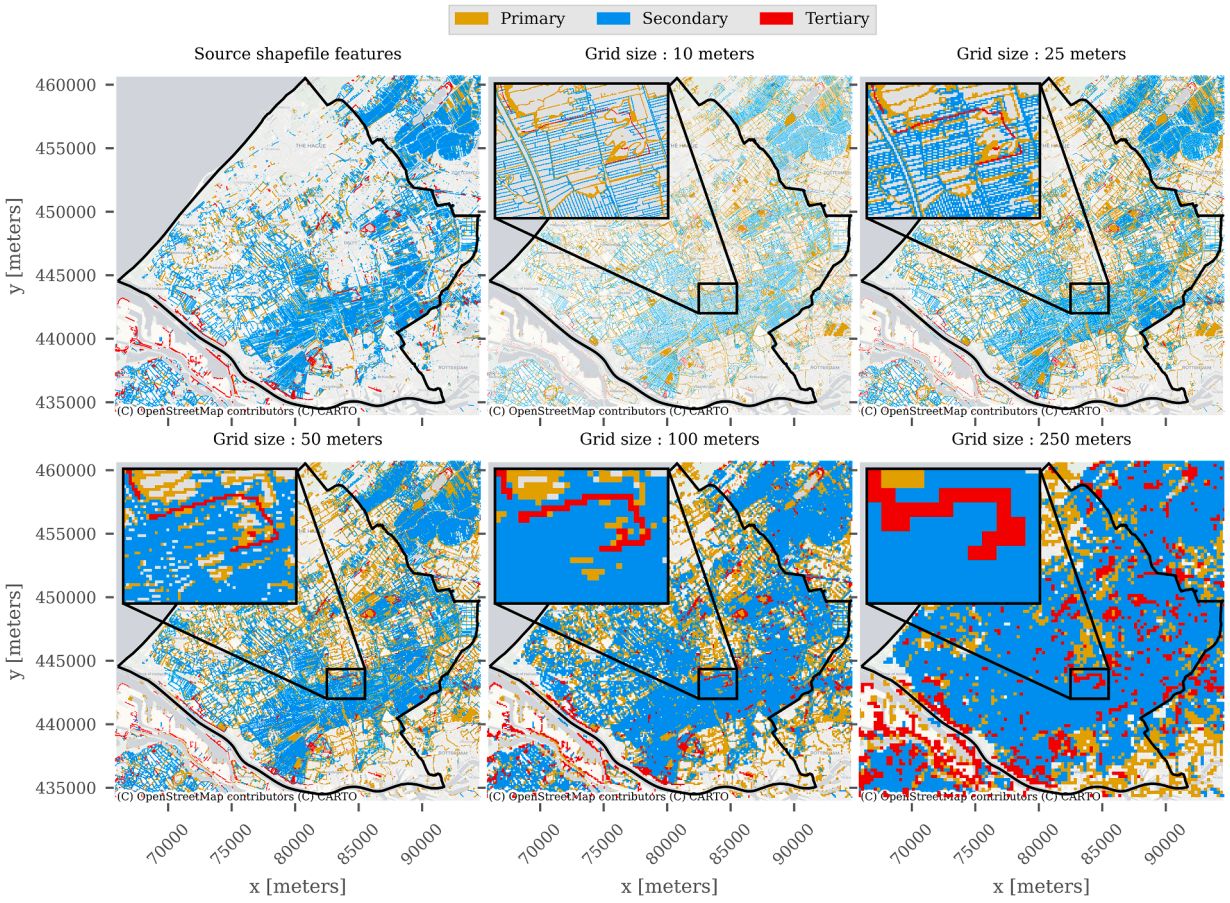


Fig. 1. Drainage system distribution for the superficial extent of the water board of Delfland at various resolutions and compared to the original vectorial source information.

a cell. In other words, the conductance calculated via this method is equivalent to aggregating the conductance of 5-meter cells to the desired resolution. The vertical resistance is part of the source information considered and available at 25-meter resolution for the whole country.

De Lange (1999) derives a conductance term from the analytical solution of a conceptual model of flow to parallel linear drainage features. This simplified concept considers the interaction between surface and groundwater in a shallow aquifer overlaying a regional aquifer with an aquitard between them and takes into account the radial flow patterns around a shallow water course. The result is an expression that quantifies upscaled representative conductance of the surface waters in a cell. This conductance depends on hydrogeological parameters such as the horizontal and vertical hydraulic conductivity of the phreatic aquifer, the vertical resistance of the surface water feature and the vertical resistance of the underlying aquitard. For geometrical parameters it considers the cell size, saturated thickness of the upper aquifer, length of the surface water feature and the wetted perimeter (approximated as the feature width). The conceptual model, the resulting equations and its parameters are described in Appendix A.

Then, calculated conductance parameters (area, length and stages) at 5-meter are upscaled to the desired grid resolutions. For length and area, the cells values are aggregated while for seasonal water stage and depth the average of the 5 m cells is taken. This process is straightforward as all target resolutions are a multiple of 5 which ensures a perfect cell overlap.

2.1.2. *LHM fresh-salt*

LHM fresh-salt is a VD-GWT model built on a parallelized version (Verkaik et al., 2021) of the widely used SEAWAT numerical code (Langevin et al., 2008). It consists of a series of python scripts that use the iMOD-python interface (Vermeulen et al., 2019) to build, run, and post-process the VD-GWT model via a reproducible scripted workflow using the Snakemake workflow manager (Köster et al., 2021).

We base our test on the LHM fresh-salt model of the Netherlands to inherit its datasets with a horizontal resolution of 250x250 meters. This serves as a baseline and provides datasets for hydrogeological parameters, aquifer geometry, topography and boundary conditions (recharge, abstractions, initial conditions, and drainage). Model results of LHM fresh-salt underwent a validation procedure that compared modelled heads, chloride concentrations and salt loads against historical data, previous model results, and water/mass balances of specific well studied areas (America et al., 2021; J. Delsman et al., 2020). LHM fresh-salt is modified to generate and run a model of a certain spatial extent. The area of interest (as a shapefile) is used to crop all the national model datasets. The gridded source datasets are downscaled to the required grid-size using nearest neighbor interpolation. As we only focus on testing the surface water boundary parametrization, we maintain the spatial distribution of all other parameters unchanged as they do not vary across grid resolutions. The only exceptions are the river package, and the drain package. The parameters for the former are based on the discretization and parametrization described in section 2.1.1. The latter is comprised of three parts; tile drainage (shallow buried pipes to avoid

field over-saturation), ditches (small canals not big enough to be parametrized in the "topsystem") and surface overland flow (drains at terrain level to simulate runoff). The conductance of these drainage components is not a part of the "topsystem" schematization, so we maintain the LHM fresh-salt values for conductance but scaled proportional to the cell area. This means that the total conductance of the three components of the drain package sums up to roughly the same value at all resolutions.

The vertical discretization of the model comprises 39 model layers that accommodate 7 aquifer-aquitard pairs and a phreatic aquifer on top. The amount of model layers per aquifer and aquitard depends on the area of the Netherlands with each active cell having varying thickness with a minimum of 1 m. The 10 uppermost layers of the model have a hydraulic conductivity calculated based on distributing the vertical resistance of the Holocene cover layer (consisting of clay, peat and sand). This distribution is done proportional to the cell thickness to calculate a vertical hydraulic conductivity. The horizontal hydraulic conductivity of these 10 layers is set assuming a horizontal anisotropy of 3 and excludes model layer 1 that is set to 0.5 [m/d]. The active modelling extent is set to the inland area of the water board of Delfland. The coastal area is deactivated because of offshore areas without information of layer resistance. Transport parameters are maintained from LHM fresh-salt, with a constant longitudinal dispersivity of 1 [m], horizontal transversal dispersivity of 0.1 [m], vertical transversal dispersivity of 0.01 [m] and a molecular diffusion coefficient of 10^{-9} [m^2/s].

Our modified version of LHM fresh-salt considers 100 yearly stress periods starting from the year 2000. This starting point is the approximate median year of the chloride measurements used to interpolate the initial groundwater salinity distribution (J. R. Delsman et al., 2023). The initial condition for chloride concentrations was interpolated with a 3D Multiple Indicator Kriging and corresponds to the 50th percentile of said interpolation. Recharge rates are steady-state and taken from LHM (De Lange et al., 2014) as a long-term yearly average value calculated from the years 2011 to 2018. The recharge in LHM is computed using the MetaSWAP unsaturated zone hydrological model (van Walsum and Groenendijk, 2008), which is coupled with MODFLOW and applied to the top-most active layer. The vadose zone processes are thus implicitly included within the spatial distribution of the recharge. Time steps were set to be daily. However, the numerical solver controls these steps to ensure numerical stability. In the original LHM fresh-salt, the sea and open estuaries are represented by a constant-head boundary with values from a scenario that predicts a long-term sea-level rise (SLR) of 3 m by the year 2278 and 0.7 m by 2100 (Haasnoot et al., 2018, 2019).

A procedure built into iMOD-python is used to set the results from LHM fresh-salt (groundwater heads and concentration), as a time-varying general head boundary (GHB) that envelops the area of interest. We use the GHB to represent the regional interaction of the groundwater system over the edge of the model. The conductance term for the GHB is calculated using a distance of two kilometers from the edge of the model to reduce boundary effects. Groundwater heads of the time varying GHB are the results from the original LHM fresh-salt run with SLR. The same logic is used for the GHB concentration, this time using the groundwater salinity per year at the model edge.

For numerical stability reasons iMOD-WQ parallel follows the logic defined by SEAWAT and its coupling with MT3DS (Zheng & Wang, 1999), meaning that advection, dispersion, sink/source and Peclet stability constraints are evaluated every time-step to maintain numerical accuracy. In our tests, the advection constraint becomes the most relevant and limiting factor for fine grid-sizes due to its dependency on the horizontal discretization and the flow velocity in each cell. For simplicity, and to reduce the computational burden, we decide to remove high-velocity-inducing features from LHM fresh-salt, this means removing wells and boils (De Louw et al., 2010). In addition, a semi-scripted procedure is implemented to deactivate surface water features that induce high velocities at small grid-sizes for the same reasons depicted for wells/boils (e.g: cells where the topography locally deepens generating a gradient towards it). If a surface water feature is deactivated in the fine grid resolution, the nearest centroid to that cell is also deactivated in all grid resolutions for consistency across scales.

Solver settings for imod-WQ were inherited from LHM fresh-salt. The notable exception is the "rclose" parameter that controls the cell mass balance convergence criterion. As this measure is a function of cell volume, this parameter is inherently scale-dependent since larger cell sizes result in larger volumes. The resolution dependent RCLOSE values utilized are shown in Table 1.

2.2. Area definition: Water board of delfland

We clipped an area from the LHM Fresh-salt model with a surface area of roughly 440 km^2, corresponding to the areal extent of the water management area of the Delfland Water Board. Delfland consists of polder areas with surface water levels below sea level ranging from −9 to + 10 m. In this coastal area, saline groundwater is found at shallow depth and causes the salinization of the surface water systems via upward seepage (De Louw et al., 2013). The current salt loads towards the surface water system are on average 600 kg/ha/yr and it is expected that this will increase in the future due to sea level rise (J. R. Delsman et al., 2023). This area is suitable for testing scaling effects on salt loads as it consists of a very detailed drainage system with varying degrees of surface–groundwater interactions involving salinity, where salinization due to saline seepage is alternated with areas of freshwater infiltration. The groundwater levels in the area are for the most part shallow (1 to 2 m deep) meaning that in nearly all of the area (>90 %) there is perfect hydraulic connection between the surface water system and the saturated zone.

2.3. Workflow specifications

This section describes the procedure and specifications needed to run the consolidated workflow at varying grid resolutions. The workflow is run in the Dutch supercomputer, Snellius (Surf, 2023), due to its extensive and scalable multi-processing power and resources. We require such flexibility because choosing a regional extent at a high resolution (10 m) result in a large (>10 million) number of cells. This translates to high RAM memory usage and storage requirements rendering this test impossible on average desktop computers or servers.

Snellius is equipped with various types of computing nodes, but for our purposes, we exclusively used two types known as "thin" and "fat" nodes. Both types feature 128 cores each, but they differ in their RAM memory capacity, with the "thin" node equipped with 224 GB per node and the "fat" node having 960 GB per node. By default, core allocation provides RAM memory in proportion to the core count, with 'thin' node cores receiving 1.75 GB and 'fat' node cores getting 7.5 GB. We simplify optimization by focusing solely on core count for memory usage. This means that if a specific part of our workflow demands more RAM memory, we increase the core count accordingly to meet those requirements.

First, we apply the LHM topsysteem workflow to the Delfland area, resulting in conductance values, stages, and water levels at varying grid resolutions. This workflow is run on a "thin" computing node utilizing the minimum possible allocation of 32 cores and 56 GB of RAM memory.

Table 1
Grid-size dependent rclose values used.

cell size [m]	RCLOSE [kg/d]
10	3.75
25	23.44
50	93.75
100	375.00
250	2343.75

This is the minimum allocatable node fraction at the time the final runs were set up (currently, Snellius allows for runs using 1/8 of a thin node, 16 cores and 28gb of RAM).

One of the features of using Snakemake as the workflow manager is that resources can be allocated individually per script. We mainly managed the number of cores used per script. Most of the workflow does not require major resources and runs on a single core. The scripts that required multi-core resources, and the associated allocations are shown in Table 2. The "Run Model" rule starts each imod-WQ model run, initiating the required MPI process, meaning that the allocated cores shown in Table 2 are divided between the MPI processes. The 10-meter grid resolution model is the most resource-demanding, not only in terms of computational power, but also in terms of file system usage. One of the major limitations of iMod-WQ is that results are saved in the binary "idf" format per layer and per timestep for each of the outputs. Additionally, the parallel solver does the same per utilized MPI process. A single run encompasses nine outputs, including data related to heads, concentrations, and mass budgets. Consequently, when simulating 100 years with a 39-layer model using 64 MPI processes, this accumulates to nearly 250,000 files per output. This amount of I-O is a challenge even for cluster environments. As a solution, we opted to partition the 10-meter model into four segments, with each segment encompassing 25 years of simulation. These segments utilize the final state of the preceding model as their initial condition. In addition, we implemented a "merge idf" routine to merge each output into a single netCDF file, allowing us to greatly reduce the load on the file system.

2.4. Post-processing

The two main outputs of a model run are concentrations and fluxes to the surface. Salt loads are calculated as the product between the flux to the surface in each cell and the concentration in the same cell and then aggregated vertically. The resulting salt loads and fluxes are analyzed statistically and compared between grid sizes. When the spatial distribution of a result is relevant, we present the values for the last stress period of the simulation. This is done because we expect initial and boundary conditions from LHM Fresh-Salt not necessarily being in equilibrium with the new grid sizes. This may be due to actual autonomous salinization in response to past sea-level rise (Oude Essink et al., 2010) or artificial model drift due to limited calibration. Regardless, to minimize the effect of this change we allow the system to evolve towards a new dynamic equilibrium by simulating salinity for 100 years.

First, basic statistical indicators (average, standard deviation, variance) are calculated per grid size and conductance method. To assess variability, the Inter Quartile Range (IQR) is calculated due to its robustness (Botta-Dukát, 2023). It is defined as:

$$IQR = Q_{75} - Q_{25}$$

With Q_{25} the 25th percentile of the salt loads and Q_{75} being the 75th percentile.

For the comparison between grid sizes to be consistent we also aggregate (upscale) the salt loads of the fine grid sizes to the coarse resolution (250x250). The aggregation considers the cell area for cases where grids misalign, for example 100 m and 250 m resolution, because of this we perform a "weighted sum" aggregation.

With the aggregated salt loads per grid size, we construct histograms to compare the distribution of values. Before comparing histograms, we apply a Gaussian Kernel Density Estimation (KDE) (Rosenblatt, 1956) to each of these to obtain a smoothed non-parametric probability density function. The same procedure is also applied to water fluxes.

3. Results

This section presents the results of the designed experiment to test the effect of grid size on the salt loads to the surface and the groundwater salinity distribution. The workflow was run successfully on the Dutch supercomputer Snellius resulting in 10 simulations for the area of interest (Delfland), 2 per grid size, each using a different method to parametrize the conductance values (De Lange, 1999) or standard MODFLOW method) of the drainage network of the Netherlands. Overall, our findings align with expected outcomes, showing both qualitative and quantitative enhancements when using fine grids, albeit minor in some respects.

Runtimes increase with the number of cells, regardless of if they are active or inactive. Even though there is a clear relationship between model size, runtime and core usage, the increment is not linear (Table 3). The finest resolution (10 m) has 600 more active elements than the coarsest resolution (250 m), while taking 178 times longer to run and using 8 times more cores. RAM memory also follows a similar trend requiring 336 times more resources between finest and coarsest scales.

3.1. Surface water network scaling: Conductance and area coverage

Scaling the parametrization of the surface water network not only implies computing values but also identifying the position of each feature. At coarse resolutions several overlapping features need to be accounted for. This translates to coarse resolutions (250- and 100-meter grid sizes) showing an area coverage of surface water elements higher than 70 % whereas fine grid resolutions are more analogous to reality which translates to area coverages lower than 50 %, going as low as 30 % for 10-meter grid resolution (Table 4). This coverage is independent of the conductance scaling method. As a note, the calculated area of surface water features is 27.8 km^2, this corresponds to roughly 6 % of the active modelling extent as area coverage. Furthermore, even at the highest resolution(10-meters), the area with a river BC is 4.5 times larger than the real value.

Conductance values calculated for the different grid resolutions are,

Table 3
General model info per grid size, run times and resource usage.

Grid size [m]	Cell Number	Active cell number	Run time [hours]	Cores	RAM usage [GB]
10	289,768,050	161,881,280	59	256	1010
25	46,412,457	25,969,074	8.25	96	57
50	11,644,776	6,521,595	2.73	32	14
100	2,932,566	1,645,202	1.07	32	5
250	475,410	269,734	0.33	32	3

Table 2
Core usage for LHM fresh-salt scripts per grid size.

Script	River/Drainage		build model		run model			merge idf	
cell size [m]	cores	node type	cores	node type	cores	MPI processes	node type	cores	node type
10	32	thin	128	fat	256	64	fat	64	fat
25	8	thin	64	fat	96	16	thin	64	fat
50	8	thin	32	fat	32	8	thin	32	fat
100	8	thin	32	thin	32	8	thin	32	thin
250	8	thin	32	thin	32	8	thin	32	thin

Table 4
River BC area coverage in the active modelling extent of Delfland.

Grid size [m]	Average active cells p/ layer	Total cells with river BC	Percent coverage	Area covered by river BC [km^2]
250	6,916	5,792	84 %	362.0
100	42,185	30,398	72 %	304.0
50	167,220	97,345	58 %	243.4
25	665,874	290,450	44 %	181.5
10	4,150,802	1,255,072	30 %	125.5

on average, higher when using the standard MODFLOW formulation across all grid resolutions and river systems when compared to De Lange (Table 5). This is consistent with the conceptualization in De Lange as this formula includes extra resistance parameters that occur due to the radial flow patterns around small surface water features which lead to lower conductance. Secondary and tertiary systems show massive discrepancies in conductance per cell between methods. However, around 99 % of the total conductance of the area is in the primary and secondary systems with the former being the most dominant (~77 % of the total conductance in Standard MODFLOW for all grids and between 96 % and 85 % for De Lange depending on grid size). Regarding changes with grid size, the gap between both methods is consistently reduced when approaching finer grid scales (Table 5). This behavior is expected as De Lange corrects for the actual existence of many surface water features within the area of a cell yielding higher resistances. This implies that higher differences in the computed conductance are expected at coarse scales, illustrating the consistent conductance scaling of this method. For the standard MODFLOW method it was also checked that the total conductance for the whole area remains roughly equal, as this method scales linearly with the cell area.

3.2. Variable density groundwater model results

Outputs of the groundwater model runs per conductance method and grid size comprise head and concentration (salinity) as well as cell budgets per boundary condition which are used to calculate salt loads. We discuss these results in three parts; groundwater salinity distribution, salt loads, and finally the effects of conductance parametrization.

Fig. 2 shows a generic overview of the main results analyzed for the end-of-simulation (year 2100) at 250x250 meter resolution using De Lange's parametrization scheme for conductance. Salinization of the shallow subsurface (b) is visible on the south, south-west portion of the model area while shallow groundwater heads (a) for the most part follow the topography and levels imposed by the drainage network. Salt loads (d) appear to be correlated with the upward seepage fluxes (c) that occur through various areas of the model but mainly concentrated in the coastal zone and its south-west border (river Rhine). The amount of water that exfiltrates to the surface (e) which is a mix of precipitation surplus and saline seepage shows some level of correlation with both salt loads and seepage fluxes.

3.2.1. Groundwater salinity distribution

The groundwater salinity distribution is heterogeneous and depends on the area considered in the model. In general, the greatest variations are observable close to the surface. As an example, Fig. 3 shows a randomly selected cross section within the modelling extent with its saltwater concentration values and normalized flow direction vectors for the last stress period simulated. A saltwater interface (SWI, *iso*-chlorine line at 0.5 [kg/m^3]) is also presented to show the spatial variability incurred by changes in the grid size. This cross section intersects surface water features of varying width and depth. High resolution results show local deepening of the SWI below infiltrating surface water features whereas coarse resolutions show a mostly horizontal interface which can be attributed to an averaged behavior over a larger cell size. This behavior, which leads to the appearance of localized freshwater lenses, is also present in other areas of the models where seepage from an underlying salinized aquifer meets fresher infiltration from surface water features. Up coning is another process only distinguishable at very high resolutions (10 m and to a lesser degree at 25-meter resolution), where a sharply rising SWI develops under surface water features. The processes described can be seen irrespective of the conductance method used and are a clear result of salinization processes occurring at finer resolutions. We further complement the aforementioned with additional cross sections over the modelling extent in Appendix A.

Variations at greater depths of the groundwater salinity distribution were also analyzed (first aquifer and deeper), but their fluctuations with resolution are minor. This can be seen in Fig. 3 where beyond 20 m depth all resolutions show a groundwater salinity distribution that can be considered equivalent. This area corresponds to the first aquifer where the regional flow does not vary significantly across the varying grid sizes.

3.2.2. Salt loads

The salt masses that exfiltrate to the surface (salt loads) are calculated and presented in two ways; on a per cell basis for all grid resolutions expressed in kg/m^2/d (we refer to this as "normalized" salt loads) or aggregated to 250x250 grid cells (we refer to this as "aggregated" salt loads). Aggregated salt loads are presented for the most part in kg/m^2/d but depending on the resulting salt loads, non-normalized values per year are also used. Another metric analyzed is the accumulated salt loads over the 100-year period, which show barely any difference (within 2 %) between resolutions. Based on this result and for practical matters, it is likely that the biggest effect of high-resolution modelling is on the peak mass values and their location

Table 6 presents a summary of salt load values per grid size and conductance method used. Normalized values show a consistent increase in mean salt-loads by decreasing grid size irrespective of the conductance method. This average is calculated considering all cells with values higher than 0, as such, by comparing the area covered by positive salt loads, we can attribute this increase to the smaller model area covered by the surface water network (Fig. 4), eg., the coarsest resolution model presents over 60 % more area with higher-than-zero salt loads than the finest resolution model.

Variability in the form of standard deviation shows the same behavior increasing with resolution and independently of the method used to calculate conductance. Quartile values further corroborate this

Table 5
Average conductance per cell for each river system and calculation method. Percentage difference shows the average increase in conductance by using the standard MODFLOW formulation.

	Primary			Secondary			Tertiary		
Grid size [m]	De Lange [m^2/d]	Modflow [m^2/d]	Percentage difference	De Lange [m^2/d]	Modflow [m^2/d]	Percentage difference	De Lange [m^2/d]	Modflow [m^2/d]	Percentage difference
250	648.72	819.36	26 %	27.90	266.83	856 %	15.35	175.89	1046 %
100	170.53	206.79	21 %	8.59	64.96	656 %	11.23	79.00	603 %
50	71.09	83.33	17 %	4.38	24.57	461 %	8.78	40.27	359 %
25	30.33	34.40	13 %	2.96	10.48	254 %	6.63	19.59	195 %
10	9.09	9.81	8 %	1.98	3.67	85 %	4.31	7.41	72 %

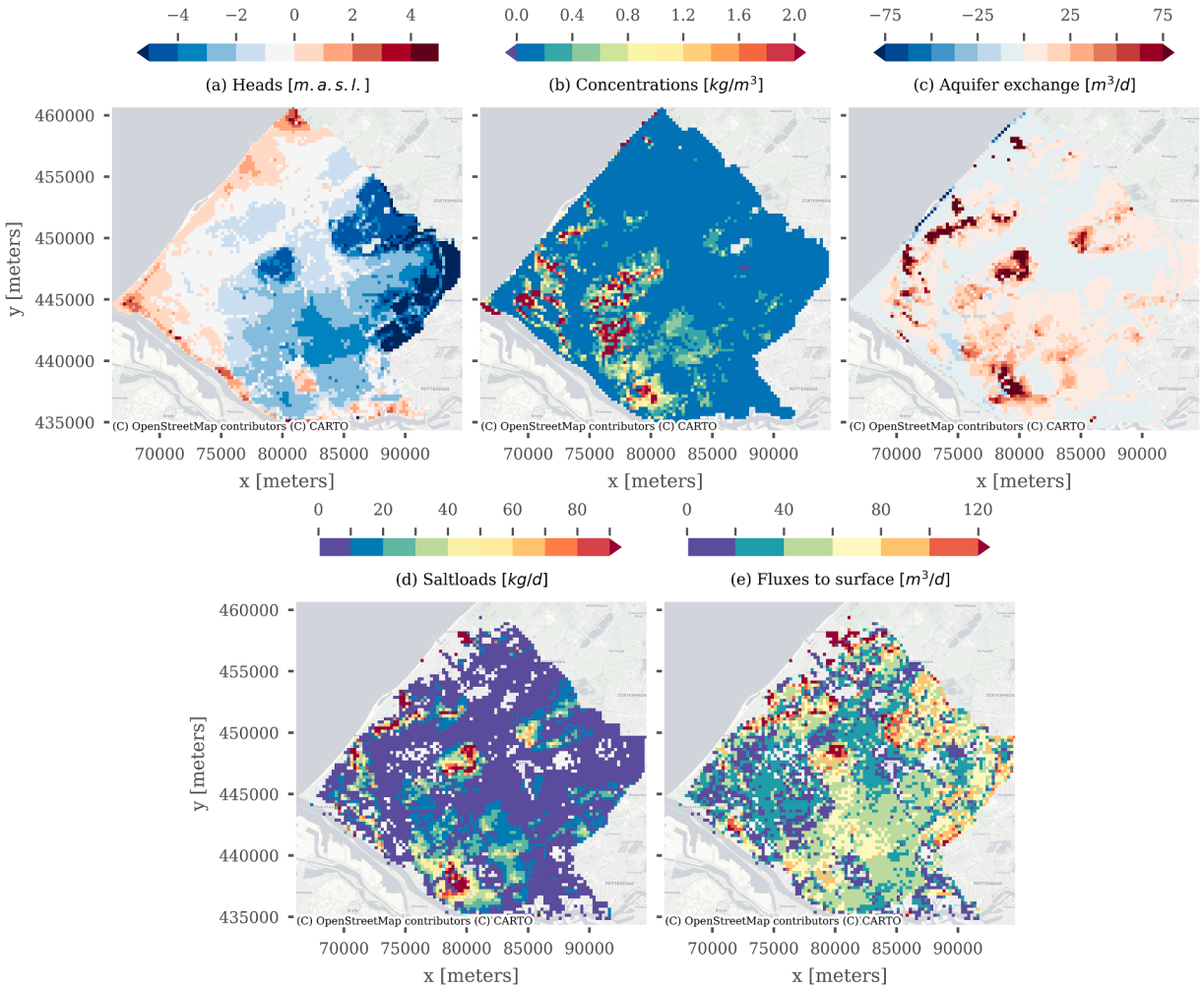


Fig. 2. Model results for stress period 100 at 250x250 meter resolution using De Lange's conductance parametrization. (a) shows the groundwater heads of layer 1, (b) is the groundwater salinity concentrations for model layer 1 (depth ~ -1.5[m], (c) depicts the groundwater exchange from the bottom of the phreatic aquifer, positive values indicate flow entering from the bottom (seepage) and negative values indicate flow to deeper layers (infiltration). (d) shows the salt loads to the surface and (e) are the total fluxes per cell that exfiltrate to the surface.

result, showing higher variability when considering all cells, which also translates to higher Inter Quartile Range values. Aggregated statistics exhibit quartile results that are for the most part homogenous across resolutions, with IQR values that are equivalent over all resolutions. Mean and standard deviation show slightly decreasing values with shrinking grid sizes. The decrease in mean and standard deviation is not considered significant but is an indicator of a marginal overestimation of total salt loads at coarse resolutions.

When comparing the aggregated salt load values with the reference 250-meter grid size, the spatial distribution shows a general trend when reducing grid size (Fig. 5). Particularly, areas that present higher salt loads (red areas) than the reference 250-meter, exfiltrate increasingly larger salt masses as the grid size gets smaller. Generally, the opposite also holds true, as cells with smaller aggregated salt loads (blue) tend to show a consistent decrease in salt mass. To explain this behavior a series of scatter plots are presented in Fig. 6 that show the degree of correlation between the difference in salt loads and the difference of four other model results. In all cases the comparison is done using the reference (250-meters) and the smallest grid size of 10-meters to accentuate the differences. It is clear from Fig. 6(a) that we can attribute the spatial changes of salt loads to the seepage entering from the underlying aquifer as its correlation is almost perfect (~1). The complete opposite can be said from infiltration (b) as the correlation with salt loads is almost non-existent, a logical result since salt loads do not occur in infiltration areas. We deduce from this that salt loads driven by seepage are a very local process originated mainly from vertical fluxes, and the existence of contiguous infiltration areas does not modify the general behavior of the surrounding surface. In other words, the pressure the surface water system exerts over the underlying aquifer defines the salt loads that will exfiltrate and at the same time explain the occurrence of local processes such as upconing. An interesting behavior is presented in panels (c) and (d) of Fig. 6, as fluxes and concentrations are what originate salt loads but their correlation, even though present, is somewhat weak.

3.2.3. Effects of conductance parametrization and resolution on fluxes and salt loads

Groundwater exfiltration values show little to no difference when evaluating the accumulated flows over the whole area across grid sizes and conductance method. This also holds true when comparing fluxes to the surface for the final stress period simulated as by this point the

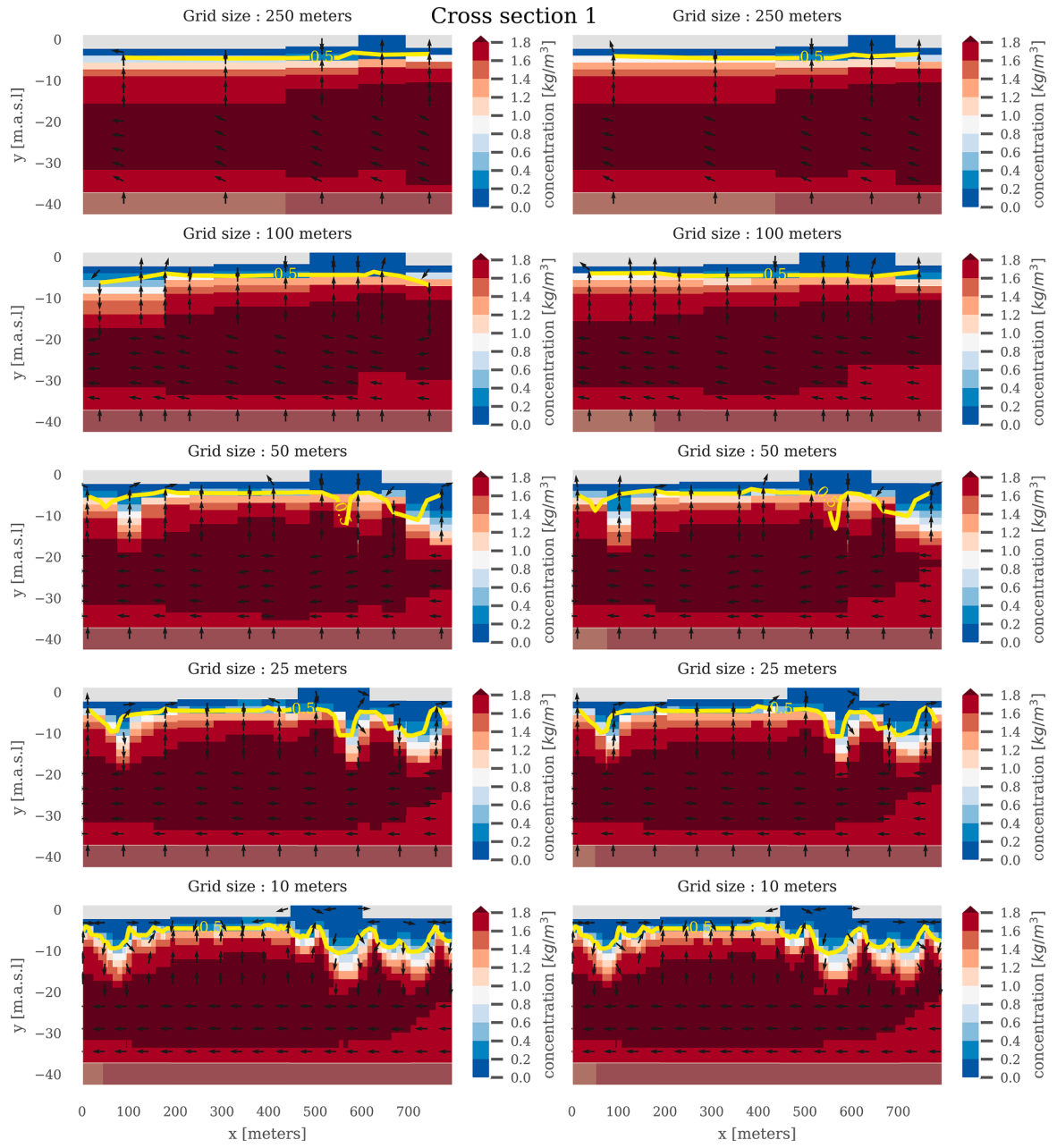


Fig. 3. Example of groundwater salinity distribution for the upper 40 m of a randomly selected 800[m] long cross section at varying grid resolutions. The left column shows model results using De Lange for conductance parametrization while the right column uses the standard MODFLOW conceptualization. Blue colors represent fresh groundwater while red represents salinized groundwater salinized. White and light colors represent the mixing zone between both (brackish groundwater). The yellow line indicates an interpolated concentration isolines at 0.5 [kg/m^3], black arrows show normalized flow directions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

system has reached a dynamic equilibrium. In other words: flows are most likely c by the amount of water entering the model. Hence, recharge is likely the most relevant factor since it is applied to the surface directly interacting with the surface water network rapidly given the shallow nature of groundwater levels in this area.

Via building an approximate representation of the distribution of values as histograms and a kernel density estimation we can detect scaling anomalies across the results. This is presented in Fig. 7 per conductance method and source MODFLOW package (Drain, River and sum of both as Total). We focus on the scaling over (water) fluxes (panels 1 to 6) and salt loads (panels 7 to 12). Perfect scaling would translate to aligned KDEs across all resolutions, this is something that we can only see clearly for flux scaling in the drain package results (2)-(5), irrespective of conductance method. We can attribute this to the fact that drain conductance was not scaled using lumped methods like De Lange and instead was assigned proportionally to the cell area. Consequently, we can deduce that the imperfect scaling of the drain salt loads (8)-(11) is due to differences in the groundwater concentrations due to local processes like saltwater upconing.

River flux scaling (3)-(6) shows lower quality scaling results as their KDEs do not overlap consistently, thus, in our experiment fluxes are affected by grid size. At high values (~10^4 m^3/d) De Lange conceptualization (3) shows better flux scaling at all resolutions except for the native 250 m grid when compared to the standard MODFLOW conceptualization (6). The sum of river and drain fluxes (1)-(3) shows KDEs more akin of the drain results, this is due to total drain fluxes being one

Table 6
Statistics of the cell-area normalized salt loads to the surface. Top table considers all cells per model at their native resolution while bottom table considers all cells aggregated to 250x250 meters and then normalized by the cell area.

	Standard Modflow [kg/m^2/d]					De Lange [kg/m^2/d]				
Grid size [m]	Mean	Standard deviation	Q25	Q75	IQR	Mean	Standard deviation	Q25	Q75	IQR
250	0.0815	0.2262	0.0086	0.0856	0.0770	0.0831	0.2293	0.0088	0.0884	0.0797
100	0.0883	0.2810	0.0092	0.0952	0.0860	0.0878	0.2826	0.0091	0.0943	0.0852
50	0.0981	0.3446	0.0101	0.1036	0.0935	0.0969	0.3451	0.0098	0.1011	0.0913
25	0.1207	0.4546	0.0119	0.1217	0.1097	0.1189	0.4530	0.0112	0.1181	0.1069
10	0.1397	0.5424	0.0119	0.1331	0.1213	0.1382	0.5415	0.0113	0.1295	0.1183
	Standard Modflow [kg/m^2/d]					**De Lange [kg/m^2/d]**				
Grid size [m]	**Mean**	**Standard deviation**	**Q25**	**Q75**	**IQR**	**Mean**	**Standard deviation**	**Q25**	**Q75**	**IQR**
250	0.0815	0.2261	0.0086	0.0856	0.0770	0.0831	0.2293	0.0088	0.0884	0.0797
100	0.0746	0.1662	0.0087	0.0826	0.0739	0.0744	0.1657	0.0087	0.0823	0.0736
50	0.0759	0.1802	0.0090	0.0842	0.0753	0.0753	0.1792	0.0090	0.0835	0.0745
25	0.0750	0.1971	0.0088	0.0820	0.0732	0.0748	0.1967	0.0088	0.0816	0.0728
10	0.0742	0.1931	0.0087	0.0813	0.0726	0.0742	0.1930	0.0087	0.0809	0.0722

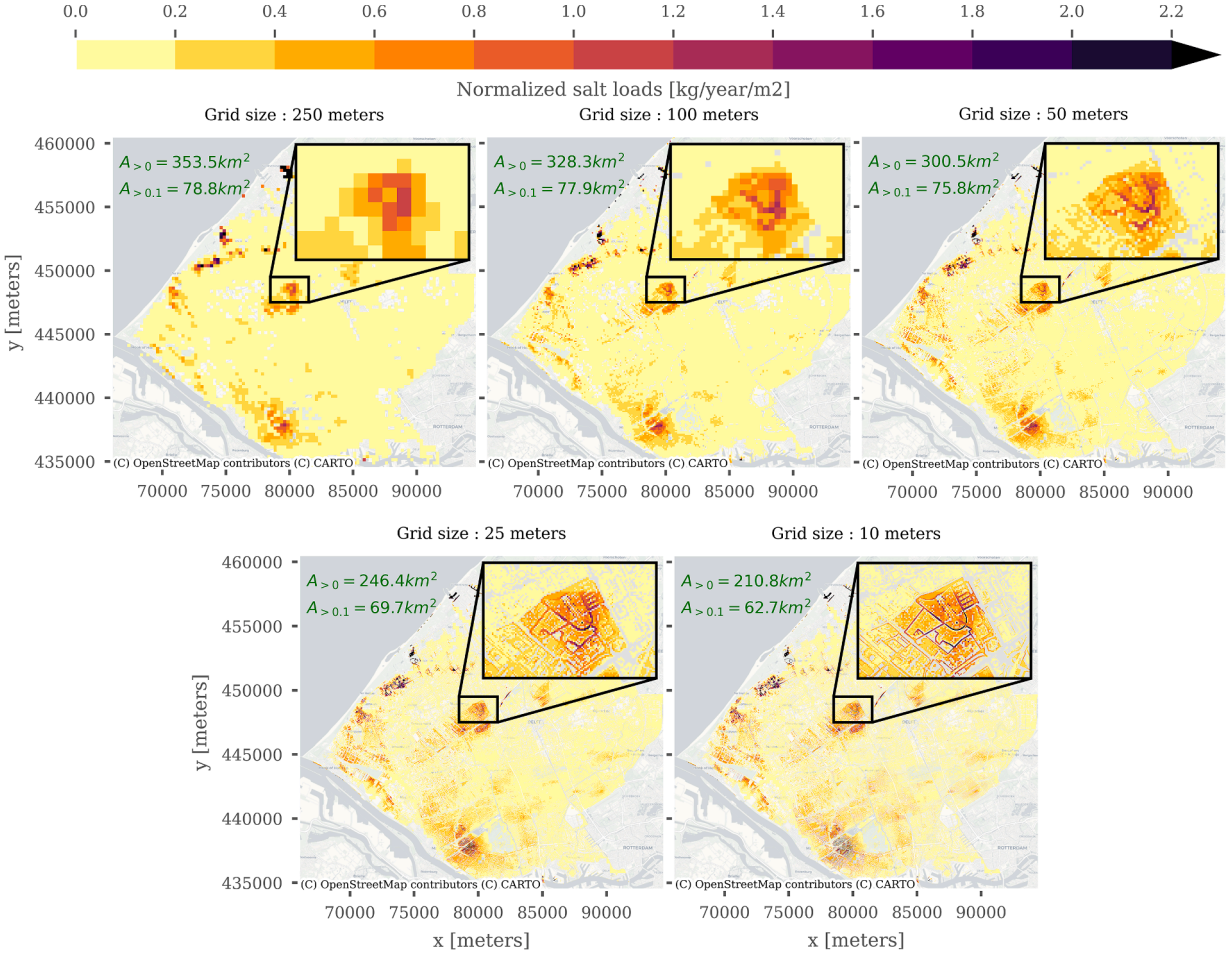


Fig. 4. Normalized salt loads using De Langes conductance method. The area covered by salt loads larger than 0 and 0.1 kg/year/m2 is shown in green for all grid resolutions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

order of magnitude bigger when compared to rivers fluxes.

To summarize, the effect of surface water parametrization on this system translates to imperfect flux scaling of the river system (primary, secondary and tertiary systems) for both methods. Even though it presents deficiencies, De Lange's method improves scaling of the high flux area of the KDEs, which are the ones more likely to drive high salt loads. All in all, and even though the standard MODFLOW formulation has much higher total conductance over the whole system, the effects of parametrization can be described as minor regarding total salt loads and groundwater salinity distribution (for more illustrative examples of the

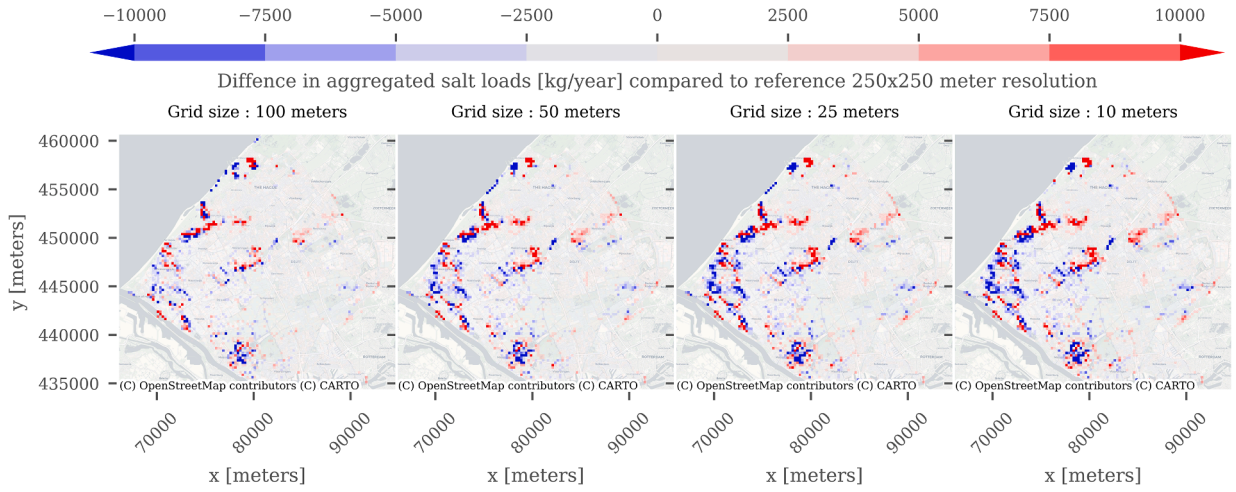


Fig. 5. Difference in aggregated salt loads (De Lange's method) aggregated to 250 m from different resolutions with the 250-meter grid resolution results as a reference. Red(positive) values indicate higher salt loads than the reference and blue values(negative) depict lower salt loads compared to the reference. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

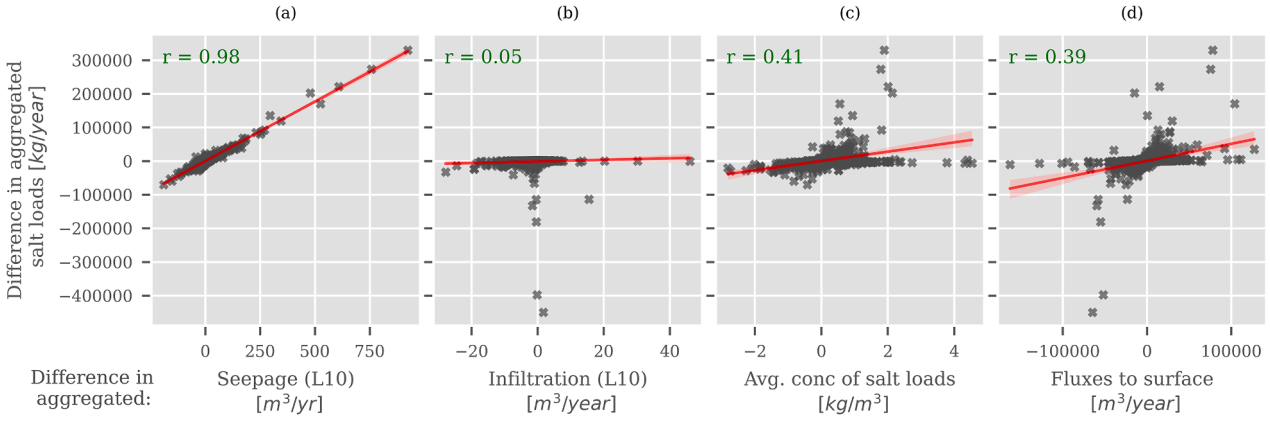


Fig. 6. Correlation (Pearson: "r") between salt loads and different model results using the difference between the reference resolution(250-meters) and the aggregated values at 10-meter grid size. The correlation is used to identify which variables impact differences in salt loads the most.

groundwater salinity distribution changes refer to appendix B). From this we also deduct a general lack of sensitivity to the parametrization method in this area, and as a consequence, model results bear low degrees of uncertainty. The main aspect where conductance parametrization plays a role is in the spatial distribution of the salt loads. Fig. 8 depicts this situation by showing a subtraction of salt loads between parametrization methods. For fine resolutions, where conductance values are similar between both methods, the difference in salt loads are minimal. However, for coarser grid sizes the spatial variations become more apparent. We can directly attribute this to de Lange's considerations for overlapping features within a cell.

Fluxes and salt loads also allow us to derive a secondary result, an average surface water concentration of chloride in the surface water network which we calculate aggregated over Local Surface Water units (LSW's), a water management division that groups areas with homogenous characteristics (Klijn et al., 2012) and water management strategies. The average concentration is calculated dividing the total salt loads by the total exfiltrating flux in each LSW. Average surface water concentrations within a LSW can be used as a direct way of evaluating the impact of salt loads for management purposes, as chloride concentration can be critical for water users. To name a few, the limit for drinking water in The Netherlands is 150 [mg/L], greenhouses often require very low (>50 [mg/L]) chloride concentrations (Stuyt et al., 2011), a sensitive crop like cucumbers only withstands around 100 [mg/L] while some tubers (like potatoes) and grains can grow with chloride concentrations in water higher than 1000 [mg/L] (Stuyt et al., 2016). Water boards, like Delfland, manage the surface water concentration via flushing of the water ways using freshwater from the river Rhine with varying target concentrations depending on the LSW. The average surface water concentration presented in Fig. 9 serves to demonstrate a scenario absent of flushing and how this results changes because of resolution. For perspective, we present Fig. 10 which shows the 20 LSWs with the largest area and their average concentrations across all resolutions. There are marked differences that could potentially mean not requiring flushing at certain resolutions or gross over/under estimations of the flushing volumes required. The general trend is that finer resolutions tend to show lower average concentration, an effect that appears more notably at higher concentrations. From a water management perspective, we see that over half the area is salinized to the point where sensitive crops are threatened. It is important to highlight that these results represent yearly averages, whereas the surface water salinization problem is more prevalent during summer, a period that is becoming

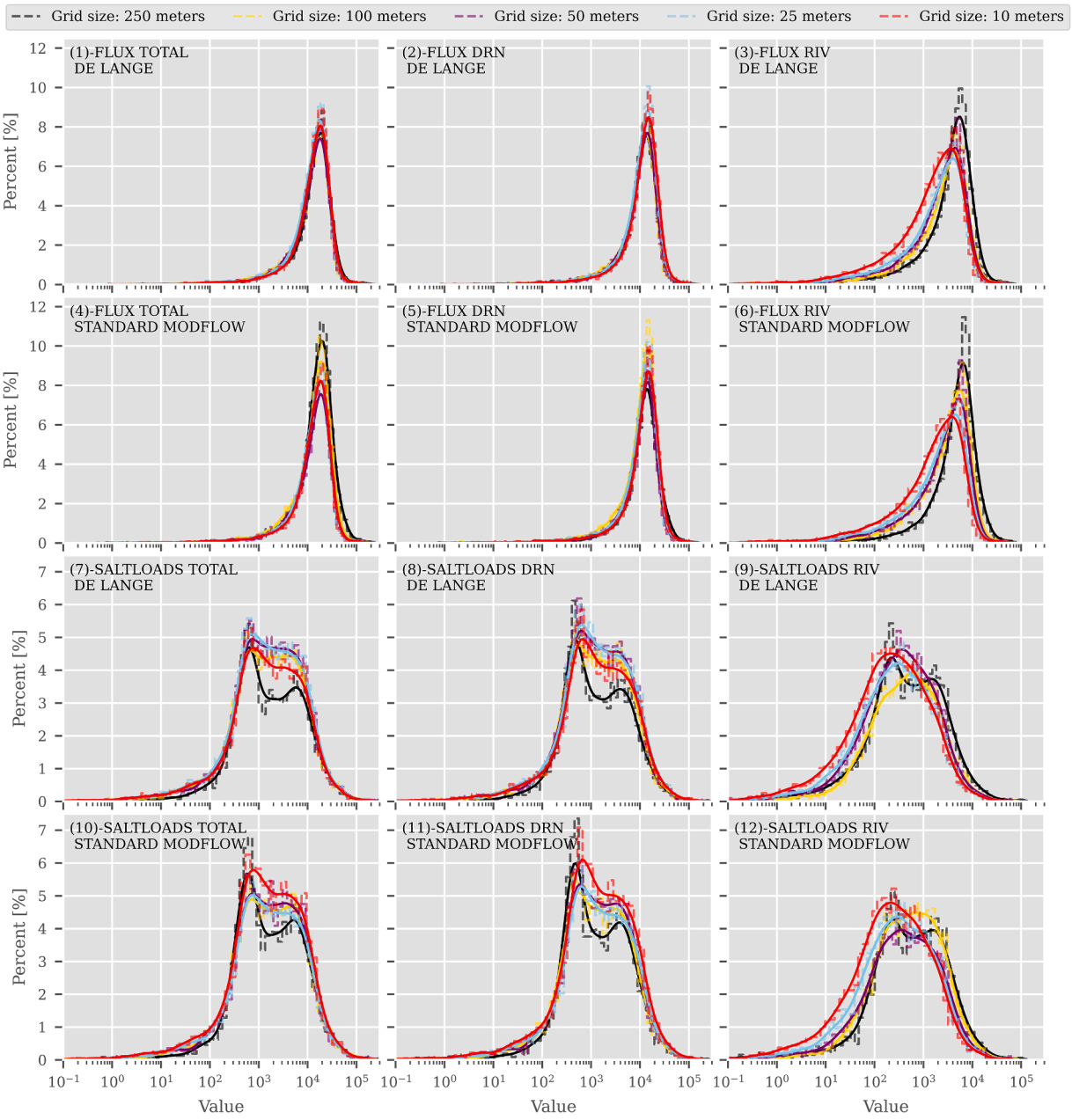


Fig. 7. Histograms and KDE for fluxes (in m^3/d, in panels 1 to 6) and salt-loads (in kg/year, in panels 7 to 12). Values are aggregated to 250x250 meters resolution, and their frequency is shown in the vertical access as a normalized percentage. DRN represents the output of the drain package, RIV is the output of the river package and total represents the sum of both.

increasingly drier due to climate change compounded with the threat of sea level rise. This affects the source of flushing water (river Rhine) augmenting the freshwater gap further highlighting the need for accurate salt loads estimations to correctly predict the future flushing estimates.

4. Discussion

4.1. Scaling effects

The effects of scaling the surface water network on the results of a regional variable density groundwater and transport models are significant. We find that lumped parametrization methods like De Lange's improve salt load and flux calculations on coarse scales as the standard MODFLOW conceptualization tends to overestimate fluxes due to producing larger conductance values. As expected, the effectiveness of De Lange's is greatly reduced with smaller grid sizes as its conceptualization that considers overlapping features within a cell becomes irrelevant and the calculated conductance from both methods converges to similar values. Surprisingly, we found that total salt loads over a large regional area remain essentially constant across resolutions for both scaling methods. We believe this is due to this system not being limited by its drainage capacity (in the form of conductance), but by the fluxes that enter the model, mainly recharge as it occurs close to the surface water network. Spatial variations in the groundwater salinity distribution do occur when changing resolution, as fine scales present local high conductance values that can drive larger fluxes and salt loads causing saltwater upconing. The aggregation of salt loads over large areas does

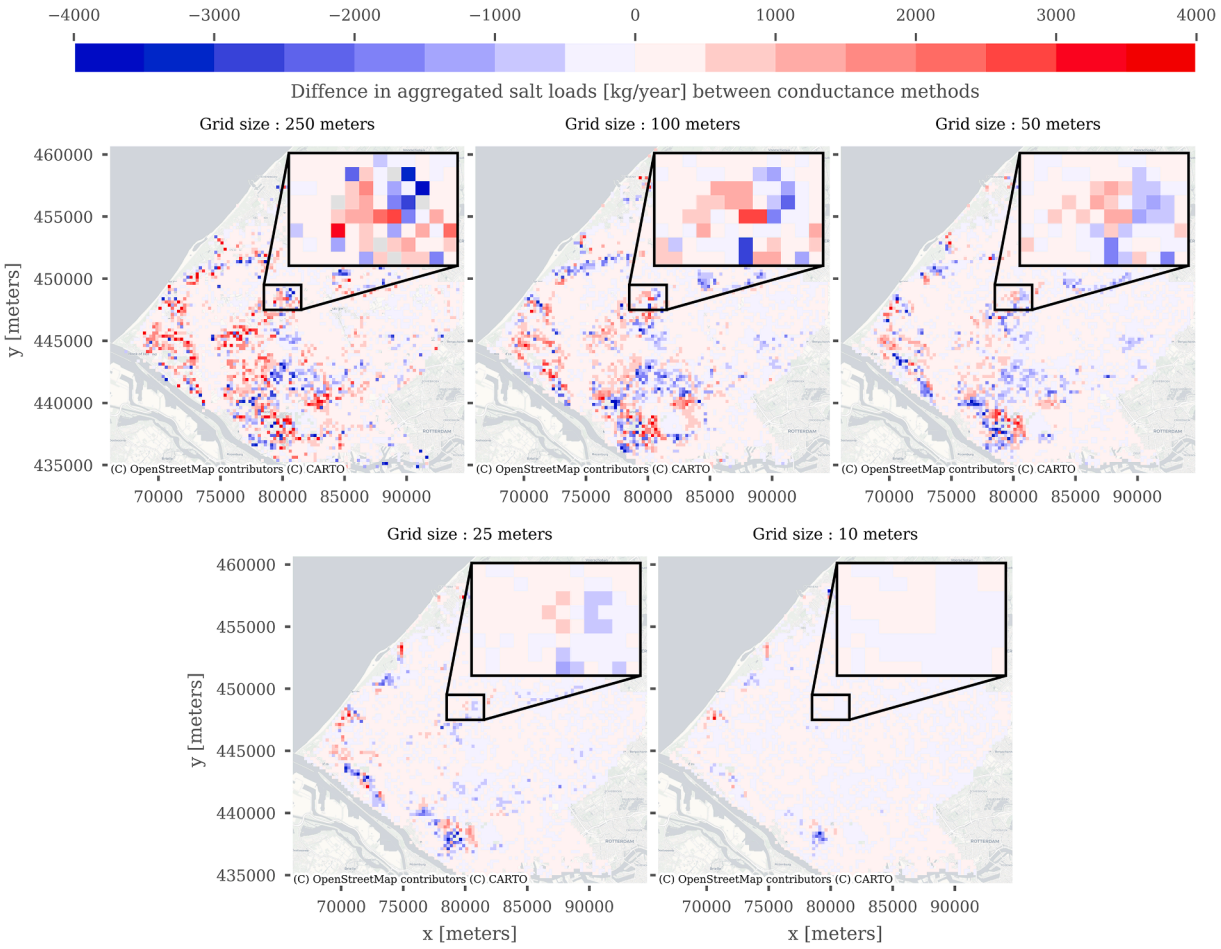


Fig. 8. Salt load comparison in [kg/year] per grid size. Standard MODFLOW values are subtracted from De Lange's. Red colors indicate areas where De Lange's method yields higher salt loads whereas blue colors depict the opposite. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

show spatial variation for which we do not have a definite explanation, but we attribute them mainly to local parametrization singularities induced by an extensive and heterogenous surface water network. For waterboards that manage and control surface water salinization, these results mean that coarse resolution models, even though capable of estimating total masses, lack the precision to estimate peak salt loads (likely underestimating them) and its specific locations. All in all, we confirm the necessity of using more complex scaling methods such as De Lange (1999) for coarse resolution regional and national modelling. The downside is that to capture local processes such as salt water upconing, downconing and the formation of freshwater lenses, finely gridded numerical models are needed which are still computationally costly at larger model extents.

4.2. *Implications for water management: Flushing*

Average surface water concentrations over LSWs are a tell-tale sign of water management units subject to salinization problems. From a water management point of view this means that surface water needs to be allocated to flush a certain LSW to reduce surface water salinity below a given management threshold. At first order, we can estimate the volume of fresh surface water required to "flush" the whole of Delfland to a target concentration as a dilution. For this first-order estimate we define a target concentration of 0.3 [g/L] and assume we will be flushing with water at a concentration of 0.2 [g/L]. Using the coarse resolution model results (250 m) it would require close to 132 [million m^3/yr] of flushing water, while for the high resolution it requires 126 [million m^3/yr]. For reference, the total recharge entering the model is 88.75 [million m^3/yr], in other words, the coarse resolution model overestimates the flushing requirement by 6 [million m^3/yr] or roughly ~ 3 % difference. It is important to note that the aggregated behavior over the whole area masks differences occurring locally per LSW as large as 200 %. To expand on this, Fig. 11 shows the flushing requirements for the 20 largest LSWs with average concentrations above 0.3 [g/L]. Here we can see differences in the flushing requirements as large as millions of cubic meters of water from coarser to finer resolution. We can also see a spatial trend where the largest differences occur in LSWs closer to the coastline. As an example, LSW 991073 requires almost 20 [million m^3 /year] of flushing at 250 m resolution but at 10 m this value is reduced to ~ 13 [million m^3 /year]. On the other hand, LSW 991094 requires close to 3 [million m^3 /year] at the coarse resolution but this value increases to 8.5 [million m^3 /year] at the smallest grid size. In line with our previous results, we can conclude that the key aspect of resolution is not the sum of absolute values over large areas but how they distribute spatially. In other words, for management purposes, grid size effects are not a matter of mass/volume, but rather a matter of location.

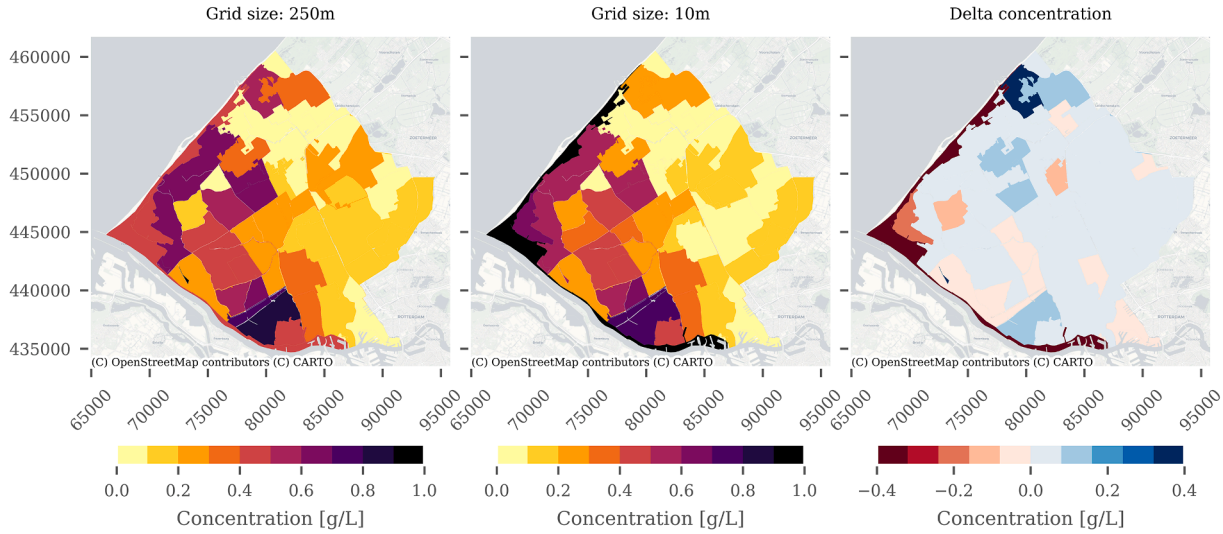


Fig. 9. Average chloride concentration over LSWs for the model results using De Lange's parametrization. The left and middle panels show the concentration for 250- and 10-meters grid size respectively. The rightmost panel shows the difference in concentration between 250 and 10 m, blue colors show higher average concentration in the coarser grid size while red colors indicate higher concentration on the high-resolution model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

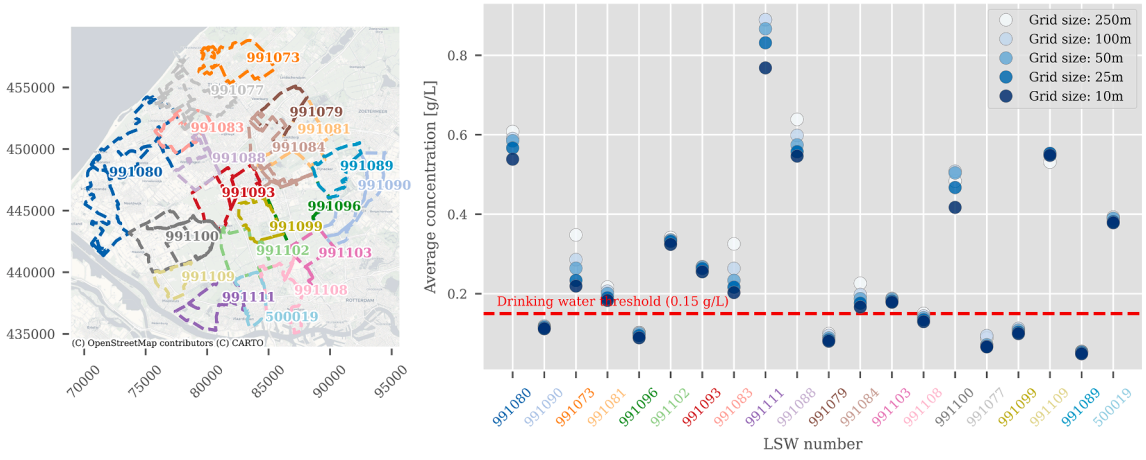


Fig. 10. Average surface water concentration for the 20 largest LSWs in Delfland across all grid resolution (right) and its location (left). Model results using De Lange's conductance method.

4.3. Limitations

The use of a high-performance computing environment like Snellius, in combination with parallelization, allowed for model sizes that would otherwise be not feasible for practical problems. Our intention by using a test case as close as possible to reality was not only to consider as many variables as possible, but also to test the limits and feasibility of 3D variable density models with complex surface water parametrizations. In this regard, an alternative that was not explored was the use of unstructured grids to locally refine around surface water features, which could lead to a significant decrease in cell number, and thus runtimes, while maintaining high resolution in relevant areas. This decision is entirely based on the lack of a MODFLOW-based numerical code freely available with parallel capabilities and support for this type of grids to solve VD-GWT problems (at the time of writing).

Due to runtime limitations, model simplification such as removing wells and boils, where needed due to the advection stability constraint imposed by SEAWAT when using an explicit solver. We believe these simplifications do not hinder the quality of the results, as the main drivers of the surface water dynamics, such as seepage, freshwater infiltration and lateral flows, are all still included. In this area, wells are for the most part located in the deeper layers of the models, separated from the surface water system by at least an aquitard. Even though it is possible for pumping to modify the flow regime in the vicinity of a well, its effect, assuming that the pumped volume is low, should not affect the surface. We do not rule out that wells interfere with the surface water dynamics, however, this complexity is not part of this study and should be further explored. This same logic unfortunately does not apply to boils as they are documented drivers of shallow groundwater and surface water salinization that stems from deeper aquifers (De Louw et al.,

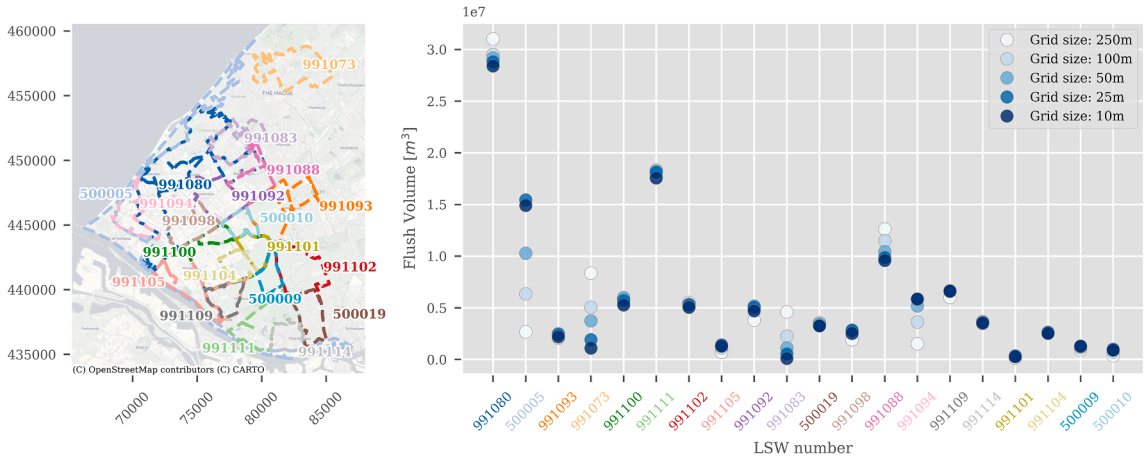


Fig. 11. Flushing requirement to reach 0.3 [g/L] via dilution with water with 0.2 [g/L] of chloride for the 20 largest LSWs in Delfland across all grid resolution (right) and its location (left). Model results using De Lange's conductance method.

2010, 2013; De Louw et al., 2011b).Therefore, future work should focus on effectively including high flow velocity features such as wells and boils.

This work focuses on spatial scaling of the boundary conditions, but other simplifications, such as considering mostly steady state boundary conditions, namely an average recharge, opens a new set of questions regarding the effect of temporal scaling over salt budgets and their spatial distribution. We were unable to test these premises as our scaled models undergoes a stabilization period due to maintaining the same initial condition as LHM fresh-salt. This state is not in dynamic equilibrium with the newly scaled boundary conditions represented by the GHB for edges and conductance of the river package. Delsman et al. (2014) documented that present day groundwater salinization in The Netherlands is mainly driven by changes occurred during the Holocene, as such, downscaling a coarse initial groundwater salinity distribution will not capture locally ongoing salinization processes. As an alternative, further iterations of our models could include a "warm up" period of 5 to 10 years before the start of a simulation. This stems from inspection of the temporal evolution of salt loads (Fig. 12) where it is possible to see an initially sharp decrease that leads to a steady increase in mass to the surface. The former can be linked to the system adjusting to the new boundary conditions and grid size while the latter can be attributed to autonomous salinization and potentially to sea level rise.

4.4. Runtimes

Runtimes and IO requirements are as expected the biggest hindrance for high resolution modelling applications. Utilizing VD-GTW models for decision making requires a certain degree of practicality, as it is often not feasible to wait in the order of weeks to obtain results to evaluate measures in any given area. This leads us to believe that large high-resolution models ($>10^8$ cells) are still not efficient for decision making, as it is often either a matter of reducing the area of focus or increasing the grid size. Based on our findings, we suggest adopting grid sizes ranging from 25 to 50 m for regional applications. This recommendation stems from the fact that an 8-hour runtime allows users to attain conclusions within a single workday, while maintaining a sufficient level of detail of local groundwater processes.

5. Conclusions

We find a direct relationship between grid size, groundwater salinity distribution and salt loads to the surface implying that for accurate regional VD-GWT models, small grid sizes (~10 m) are preferable.

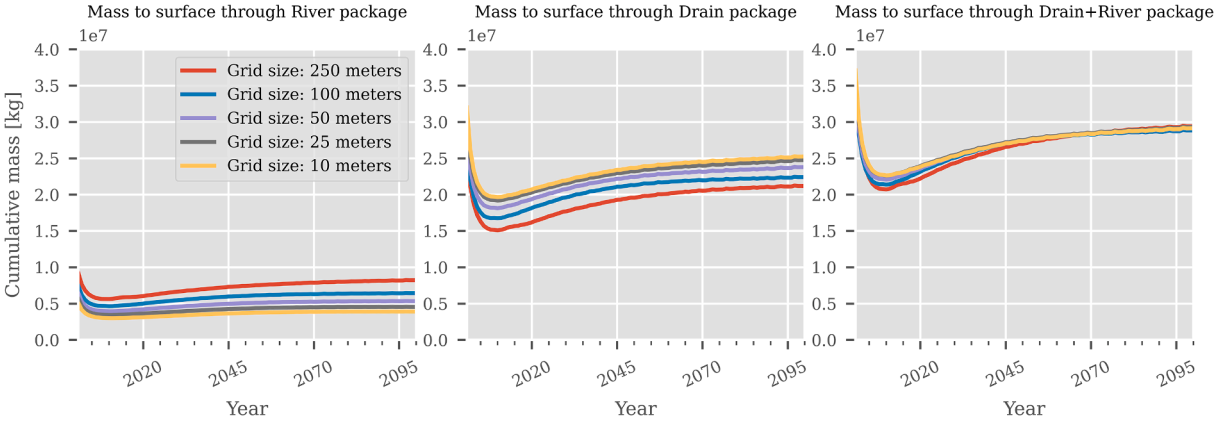


Fig. 12. Temporal evolution of salt loads per grid size for the drain and river packages and the sum of both. The conductance is calculated using De Lange's method.

However, the computational resources needed are still the limiting factor leading us to recommend grid sizes around ~ 25 m. Lumped parametrization schemes, like De Lange's show a limited but quantifiable improvement over the scaling of the conductance parameter that leads to better flux scaling across grid sizes. The use of such methods is then recommended when for a given grid size, several overlapping features end up in the same model cell. We confirm that in this coastal area, the main driver of salt loads to the surface is the vertical seepage flow from underlying aquifers, and most importantly, that these seepage fluxes are favored by the presence of surface water features. Consequently, a drainage network modelled at high resolution will show higher peak salt loads as fluxes are concentrated over a smaller area generating local salinity upconing. This behavior shapes the groundwater salinity distribution and how it varies across resolutions. Local salinity upconing is visible at small grid sizes (10–25 m, and to a lesser extent 50-meter) also favoring the development of freshwater lenses between surface water features. Such local processes are lost at coarse resolutions (>100 m grid size). All scale effects over the groundwater salinity distribution are observable irrespective of the conductance parametrization used.

Software and data availability

Model workflow and code is accessible at https://gitlab.com/ifariasg1/nhi-fresh-salt-clip. The executable of the software code (IMOD-WQ) is available at https://oss.deltares.nl/web/imod for Windows machines. Input data for LHM fresh-salt is available from https://doi.org/10.5281/zenodo.7419219. Complementarily, the NHI data portal (https://data.nhi.nu) gathers several inputs as the 3D groundwater salinity distribution of The Netherlands.

CRediT authorship contribution statement

Ignacio Farías: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Gualbert H.P. Oude Essink:** Writing – review & editing, Methodology, Conceptualization. **Perry G.B. de Louw:** Writing – review & editing, Methodology, Conceptualization. **Marc F.P. Bierkens:** Writing – review & editing, Methodology, Conceptualization, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors thank Joost Delsman and Huite Bootsma for their support during the implementation of the workflow. This work used the Dutch national e-infrastructure with the support of the SURF Cooperative using grant no. EINF-3928. This research was performed within the framework of the research program AquaConnect, funded by the Dutch Research Council (NWO, grant-ID P19-45) and public and private partners of the AquaConnect consortium and coordinated by Wageningen University and Research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2024.131915.

References

America, I., Mulder, T., & Delsman, J. (2021). *Veranderingsrapportage en plausibiliteitstoets LHM zoet-zout 4.1. Deltares report 11206798-013-BGS-0001*.

Bailey, R.T., Hosseini, P., 2023. Comprehensive simulation of salinity transport in irrigated watersheds using an updated version of SWAT-MODFLOW. Environ. Model. Softw. 159, 105566 https://doi.org/10.1016/j.envsoft.2022.105566.

Botta-Dukát, Z. (2023). Quartile coefficient of variation is more robust than CV for traits calculated as a ratio. *Scientific Reports 2023 13:1*, *13*(1), 1–6. 10.1038/s41598-023-31711-8.

Christen, E., Skehan, D., 2001. Design and management of subsurface horizontal drainage to reduce salt loads. J. Irrig. Drain. Eng. 127 (3), 148–155. https://doi.org/10.1061/(asce)0733-9437(2001)127:3(148).

De Lange, W.J., 1999. A Cauchy boundary condition for the lumped interaction between an arbitrary number of surface waters and a regional aquifer. J. Hydrol. 226 (3–4), 250–261. https://doi.org/10.1016/S0022-1694(99)00143-2.

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., Th, L., Kroon, T., 2014. 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. https://doi.org/10.1016/j.envsoft.2014.05.009.

De Louw, P.G.B., Oude Essink, G.H.P., Stuyfzand, P.J., Van Der Zee, S.E.A.T.M., 2010. Upward groundwater flow in boils as the dominant mechanism of salinization in deep polders, The Netherlands. J. Hydrol. 394, 494–506. https://doi.org/10.1016/j.jhydrol.2010.10.009.

De Louw, P.G.B., Eeman, S., Siemon, B., Voortman, B.R., Gunnink, J., Van Baaren, E.S., Oude Essink, G.H.P., 2011a. Shallow rainwater lenses in deltaic areas with saline seepage. Hydrology and Earth System Sciences 15 (12), 3659–3678. https://doi.org/10.5194/HESS-15-3659-2011.

De Louw, P.G.B., Van Der Velde, Y., Van Der Zee, S.E.A., 2011b. Quantifying water and salt fluxes in a lowland polder catchment dominated by boil seepage: a probabilistic end-member mixing approach. Hydrol. Earth Syst. Sci. 15 (7), 2101–2117. https://doi.org/10.5194/HESS-15-2101-2011.

De Louw, P.G.B., Vandenbohede, A., Werner, A.D., Oude Essink, G.H.P., 2013. Natural saltwater upconing by preferential groundwater discharge through boils. J. Hydrol. 490, 74–87. https://doi.org/10.1016/J.JHYDROL.2013.03.025.

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., Bierkens, M.F.P., 2014. Paleo-modeling of coastal saltwater intrusion during the Holocene: an application to the Netherlands. Hydrol. Earth Syst. Sci. 18, 3891–3905. https://doi.org/10.5194/hess-18-3891-2014.

Delsman, J.R., Mulder, T., Verastegui, B.R., Bootsma, H., Huizer, S., Essink, G.H.P.O., Zitman, P., Huizer, S., Oude Essink, G.H.P., 2023. Reproducible construction of a high-resolution national variable-density groundwater salinity model for the Netherlands. Environ. Model. Softw. 105683 https://doi.org/10.1016/J.ENVSOFT.2023.105683.

Delsman, J., Oude Essink, G., Huizer, S., Bootsma, H., Mulder, T., Zitman, P., Verastegui, B., Janssen, G., 2020. Actualisatie zout in het NHI. Toolbox NHI Zoet-Zout Modellering En Landelijk Model. 10.13140/RG.2.2.17077.09447.

Delsman, J. R. (2015). Saline groundwater – Surface water interaction in coastal lowlands. In *Saline Groundwater - Surface Water Interaction in Coastal Lowlands*. IOS Press. 10.3233/978-1-61499-518-0-i.

Di Ciacca, A., Leterme, B., Laloy, E., Jacques, D., Vanderborght, J., 2019. Scale-dependent parameterization of groundwater–surface water interactions in a regional hydrogeological model. J. Hydrol. 576, 494–507. https://doi.org/10.1016/j.jhydrol.2019.06.072.

Faneca Sànchez, M., Gunnink, J.L., Van Baaren, E.S., Oude Essink, G.H.P., Siemon, B., Auken, E., Elderhorst, W., De Louw, P.G.B., 2012. Modelling climate change effects on a dutch coastal groundwater system using airborne electromagnetic measurements. Hydrol. Earth Syst. Sci. 16 (12), 4499–4516. https://doi.org/10.5194/HESS-16-4499-2012.

Forzieri, G., Feyen, L., Rojas, R., Flörke, M., Wimmer, F., Bianchi, A., 2014. Ensemble projections of future streamflow droughts in Europe. Hydrol. Earth Syst. Sci. 18 (1), 85–108. https://doi.org/10.5194/HESS-18-85-2014.

Haasnoot, M., Builder, L., Diermanse, F., Kwadijk, J., Spek, A. van der, Oude Essink, G., Delsman, J., Weiler, O., Man, M., Maat, J. ter, Huismans, Y., Sloff, K., & Musselman, E. (2018). *Mogelijke gevolgen van versnelde zeespiegelstijging voor het Deltaprogramma. Een verkenning*.

Haasnoot, M., Diermanse, F., Kwadijk, J., Winter, R. de, & Winter, G. (2019). *Strategieën voor adaptatie aan hoge en versnelde zeespiegelstijging. Een verkenning* .

Haasnoot, M., Kwadijk, J., Van Alphen, J., Le Bars, D., Van Den Hurk, B., Diermanse, F., Van Der Spek, A., Oude Essink, G., Delsman, J., Mens, M., 2020. Adaptation to uncertain sea-level rise; how uncertainty in Antarctic mass-loss impacts the coastal adaptation strategy of the Netherlands. Environ. Res. Lett. 15 (3), 034007 https://doi.org/10.1088/1748-9326/AB666C.

Harbaugh, A. W. (2005). MODFLOW-2005 : the U.S. Geological Survey modular ground-water model–the ground-water flow process. In *Techniques and Methods*. 10.3133/TM6A16.

Klijn, F., van Velzen, E., ter Maat, J., Hunink, J., Baarse, G., Beumer, V., Boderie, P., Buma, J., Delsman, J., Hoogewoud, J., Hoogvliet, M., Prinsen, G., van Bakel, J., van der Mark, R., van Ek, R., van Sligte, R., Verheij, H., & Zwolsman, G.-J. (2012). *Zoetwatervoorziening in Nederland: aangescherpte landelijke knelpuntenanalyse 21e eeuw* (2e (gecorr). Deltares.

Koopman, J. F. L., Kuik, O., Tol, R. S. J., van der Vat, M. P., Hunink, J. C., & Brouwer, R. (2019). *Distributing Water Between Competing Users in the Netherlands*. 159–192. 10.1007/978-981-13-6101-2_8.

Köster, J., Mölder, F., Jablonski, K. P., Letcher, B., Hall, M. B., Tomkins-Tinch, C. H., Sochat, V., Forster, J., Lee, S., Twardziok, S. O., Kanitz, A., Wilm, A., Holtgrewe, M., Rahmann, S., & Nahnsen, S. (2021). Sustainable data analysis with Snakemake. *F1000Research 2021 10:33*, *10*, 33. 10.12688/f1000research.29032.2.

Kwadijk, J.C.J., Haasnoot, M., Mulder, J.P.M., Hoogvliet, M.M.C., Jeuken, A.B.M., van der Krogt, R.A.A., van Oostrom, N.G.C., Schelfhout, H.A., van Velzen, E.H., van Waveren, H., de Wit, M.J.M., 2010. Using adaptation tipping points to prepare for climate change and sea level rise: A case study in the Netherlands. Wiley Interdiscip. Rev. Clim. Chang. 1 (5), 729–740. https://doi.org/10.1002/wcc.64.

Langevin, C. D., Jr., D. T. T., Dausman, A. M., Sukop, M. C., & Guo, W. (2008). SEAWAT Version 4: A Computer Program for Simulation of Multi-Species Solute and Heat Transport. *Techniques and Methods*. 10.3133/TM6A22.

Loáiciga, H.A., Pingel, T.J., Garcia, E.S., 2012. Sea water intrusion by sea-level rise: scenarios for the 21st century. Groundwater 50 (1), 37–47. https://doi.org/10.1111/J.1745-6584.2011.00800.X.

Massop, H.T.L., van der Gaast, J.W.J., Hermans, A.G.M., 2006. Kenmerken van het ontwateringstelsel in Nederland. Alterra. https://research.wur.nl/en/publications/kenmerken-van-het-ontwateringstelsel-in-nederland.

McDonald, R.I., Green, P., Balk, D., Fekete, B.M., Revenga, C., Todd, M., Montgomery, M., 2011. Urban growth, climate change, and freshwater availability. PNAS 108 (15), 6312–6317. https://doi.org/10.1073/PNAS.1011615108/-/DCSUPPLEMENTAL.

Mens, M., Pouwels, J., Weiler, O., Delsman, J., & Huismans, Y. (2024). *The freshwater balance of the lower Netherlands in a warmer climate. Deltares Report 11210362-000-ZWS-0001 (De zoetwaterbalans van laag Nederland in een warmer klimaat)*. https://www.deltares.nl/expertise/publicaties/zoetwaterbalans-van-laag-nederland-in-een-warmer-klimaat.

Michael, H.A., Russoniello, C.J., Byron, L.A., 2013. Global assessment of vulnerability to sea-level rise in topography-limited and recharge-limited coastal groundwater systems. Water Resour. Res. 49 (4), 2228–2240. https://doi.org/10.1002/WRCR.20213.

Noorduijn, S.L., Refsgaard, J.C., Petersen, R.J., Højberg, A.L., 2021. Downscaling a national hydrological model to subgrid scale. J. Hydrol. 603, 126796 https://doi.org/10.1016/j.jhydrol.2021.126796.

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

Pauw, P.S., Van der Zee, S.E.A.T.M., Leijnse, A., Delsman, J.R., De Louw, P.G.B., De Lange, W.J., Oude Essink, G.H.P., 2015. Low-resolution modeling of dense drainage networks in confining layers. Groundwater 53 (5), 771–781. https://doi.org/10.1111/GWAT.12273.

Raats, P.A.C., 2015. Salinity management in the coastal region of the Netherlands: a historical perspective. Agric. Water Manag. 157, 12–30. https://doi.org/10.1016/J.AGWAT.2014.08.022.

Rasmussen, P., Sonnenborg, T.O., Goncear, G., Hinsby, K., 2013. Assessing impacts of climate change, sea level rise, and drainage canals on saltwater intrusion to coastal aquifer. Hydrol. Earth Syst. Sci. 17 (1), 421–443. https://doi.org/10.5194/HESS-17-421-2013.

Rosenblatt, M. (1956). Remarks on Some Nonparametric Estimates of a Density Function. *10.1214/Aoms/1177728190*, *27*(3), 832–837. 10.1214/AOMS/1177728190.

Sathish, S., Chanu, S., Sadath, R., Elango, L., 2022. Impacts of regional climate model projected rainfall, sea level rise, and urbanization on a coastal aquifer. Environ. Sci. Pollut. Res. 29 (22), 33305–33322. https://doi.org/10.1007/S11356-021-18213-8/TABLES/3.

Seibert, S.L., Greskowiak, J., Bungenstock, F., Freund, H., Karle, M., Meyer, R., Oude Essink, G.H.P., van Engelen, J., Massmann, G., 2023. Paleo-hydrogeological modeling to understand present-day groundwater salinities in a low-lying coastal groundwater System (Northwestern Germany). Water Resour. Res. 59 (4) https://doi.org/10.1029/2022WR033151 e2022WR033151.

Stoter, J., Post, M., Van Altena, V., Nijhuis, R., Bruns, B., 2014. Fully automated generalization of a 1:50k map from 1:10k data. Cartogr. Geogr. Inf. Sci. 41 (1), 1–13. https://doi.org/10.1080/15230406.2013.824637.

Stuyt, L.C.P.M., Blom-Zandstra, M., Kselik, R.A.L., 2016. Inventarisatie en analyse zouttolerantie van landbouwgewassen op basis van bestaande gegevens. Wageningen Environ. Res. 10.18174/391931.

Stuyt, L. C. P. M., van Bakel, P. J. T., & Massop, H. T. L. (2011). *Basic survey zout en joint fact finding effecten van zout: naar een gedeeld beeld van het zoetwaterbeheer in laag Nederland* (Issue 2200). Alterra.

Surf. (2023). *Snellius - SURF User Knowledge Base - SURF User Knowledge Base*. https://servicedesk.surf.nl/wiki/display/WIKI/Snellius.

Van Engelen, J., Verkaik, J., King, J., Nofal, E.R., Bierkens, M.F.P.P., Oude Essink, G.H.P.P., 2019. A three-dimensional palaeohydrogeological reconstruction of the groundwater salinity distribution in the Nile Delta Aquifer. Hydrol. Earth Syst. Sci. 23 (12), 5175–5198. https://doi.org/10.5194/hess-23-5175-2019.

Van Rees Vellinga, E., Toussaint, C.G., Wit, K.E., 1981. Water quality and hydrology in a coastal region of The Netherlands. J. Hydrol. 50 (C), 105–127. https://doi.org/10.1016/0022-1694(81)90063-9.

van Walsum, P.E.V., Groenendijk, P., 2008. Quasi steady-state simulation of the unsaturated zone in groundwater modeling of lowland regions. Vadose Zone J. 7 (2), 769–781. https://doi.org/10.2136/VZJ2007.0146.

Verkaik, J., Van Engelen, J., Huizer, S., Bierkens, M.F.P., Lin, H.X., Oude Essink, G.H.P., 2021. Distributed memory parallel computing of three-dimensional variable-density groundwater flow and salt transport. Adv. Water Resour. 154, 103976 https://doi.org/10.1016/j.advwatres.2021.103976.

Verkaik, J., Sutanudjaja, E.H., Oude Essink, G.H.P., Lin, H.X., Bierkens, M.F.P., 2024. GLOBGM v1.0: a parallel implementation of a 30 arcsec PCR-GLOBWB-MODFLOW global-scale groundwater model. Geosci. Model Dev. 17 (1), 275–300. https://doi.org/10.5194/GMD-17-275-2024.

Vermeulen, P. T. M., Roelofsen, F. J., Minnema, B., Burgering, L. M. T., Verkaik, G. M. C. M. J., Janssen, & Verastegui, B. R. (2019). *iMOD User Manual. Version 5.1, June 15, 2020*. Deltares. http://oss.deltares.nl/web/iMOD.

Vermeulen, P.T.M., Te Stroet, C.B.M., Heemink, A.W., 2006. Limitations to upscaling of groundwater flow models dominated by surface water interaction. Water Resour. Res. 42 (10) https://doi.org/10.1029/2005WR004620.

Werner, A.D., Bakker, M., Post, V.E.A., Vandenbohede, A., Lu, C., Ataie-Ashtiani, B., Simmons, C.T., Barry, D.A., 2013. Seawater intrusion processes, investigation and management: Recent advances and future challenges. Adv. Water Resour. 51, 3–26. https://doi.org/10.1016/J.ADVWATRES.2012.03.004.

Yu, X., Michael, H.A., 2022. Impacts of the scale of representation of heterogeneity on simulated salinity and saltwater circulation in coastal aquifers. Water Resour. Res. 58 (1) https://doi.org/10.1029/2020WR029523 e2020WR029523.

Zamrsky, D., Essink, G.H.P.O., Bierkens, M.F.P., 2024. Global impact of sea level rise on coastal fresh groundwater resources. Earth's Future 12 (1). https://doi.org/10.1029/2023EF003581 e2023EF003581.

Zheng, C., & Wang, P. P. (1999). MT3DMS - A Modular Three-Dimensional Multispecies Transport Model. *Strategic Environmental Research and Development Program*, 1–40. http://ezproxy.library.dal.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2009-10794-007&site=ehost-live%5Cnhttp://mouzas@med.uth.gr.