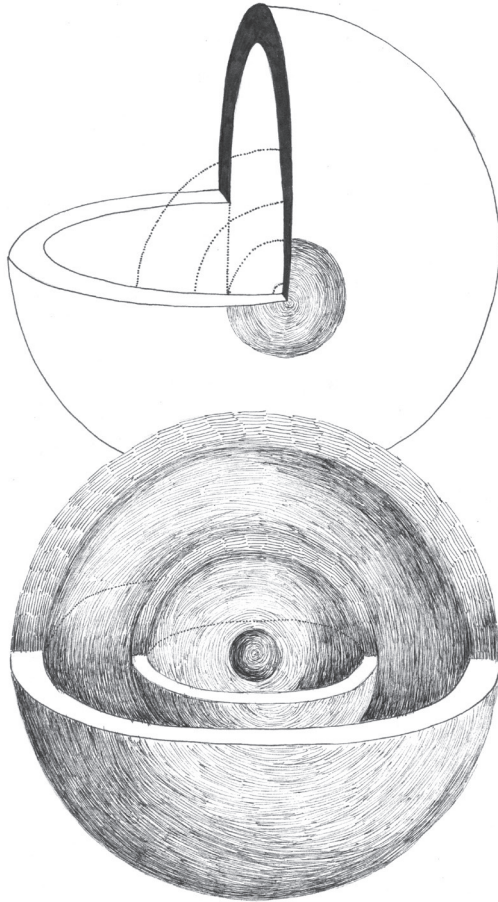# Airborne electromagnetic mapping of coastal groundwater salinity

Quantifying uncertainty and investigating methodological improvements

Jude King

|

Airborne electromagnetic mapping of coastal groundwater salinity

Quantifying uncertainty and investigating methodological improvements

Airborne elektromagnetische kartering van het zoutgehalte van grondwater in kustgebieden

Kwantificering van onzekerheden en onderzoek naar verbetering methodes

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 21 oktober 2022 des middags te 4.15 uur

door

Jude Axon King

geboren op 27 februari 1984
te Grahamstown, Zuid-Afrika

Promotoren:
Prof. dr. ir. M.F.P. Bierkens
Dr. ir. G.H.P. Oude Essink

Beoordelingscommissie:
Prof. dr. J. Carrera
Prof. dr. rer. nat. S. Attinger
Dr. M. Karaoulis
Prof. dr. ir. T.J. Heimovaara
Prof. dr. J. Griffioen

Airborne electromagnetic mapping of coastal groundwater salinity

Quantifying uncertainty and investigating methodological improvements

Jude King

Utrecht 2022

Department of Physical Geography
Faculty of Geosciences
Utrecht University

Promotor:
Prof. dr. ir. M.F.P. Bierkens

Co-Promotor:
Dr.ir. G.H.P. Oude Essink

Examination Committee:
Prof. dr. J. Carrera
Prof. dr. rer. nat. S. Attinger
Dr. M. Karaoulis
Prof. dr. ir. T.J. Heimovaara
Prof. dr. J. Griffioen

The following chapters are either unpublished articles or final author versions of previously published articles, © by Jude King and co-authors. More information and citation suggestions are provided at the beginning of these chapters.

Contents

1

Introduction

1.1   Background

Low elevation coastal zones (LECZs), defined here as areas that are ≤10 m above mean sea-level, currently host nearly 800 million people – about 10% of the world's population on only 2% of the earth's surface (Neumann et al., 2015). Thanks to these resource rich environments, people have been drawn to LECZs for millennia, notably occupying deltaic areas almost in synchronicity with the formation of modern deltas about 7000 yr. B.P (Stanley and Warne, 1997). Popularity of these areas has not changed – in fact current estimates point to significant population growth in the future, with estimates of 949 million and 1.4 billion inhabitants by the years 2030 and 2060 respectively (Neumann et al., 2015). Despite current (and predicted) global popularity, these zones are fraught with environmental challenges. A few famous examples include Jakarta (Indonesia) and the Mekong Delta (Vietnam), where because of anthropogenic stressors, large parts of these cities are sinking below sea-level – causing significant property damage and human displacement (Minderhoud et al., 2017). Unfortunately, what happens on the surface is only part of the picture, and is often either indicative, or indeed caused by what's going on beneath us.
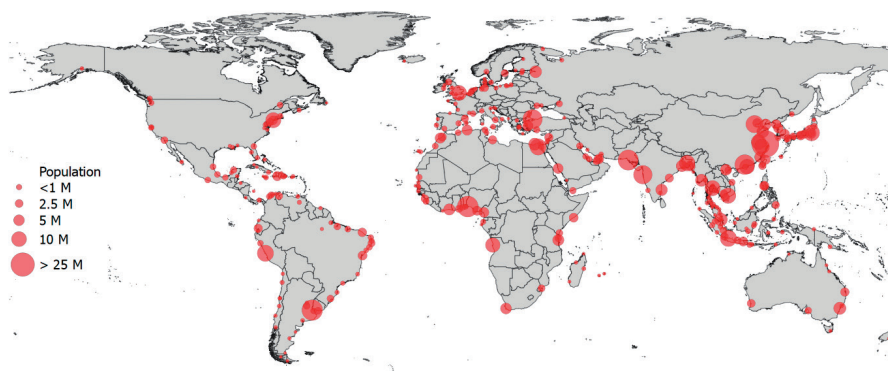
Figure 1.1. Global population centres within low elevation coastal zones (LECZs).
LECZ data from Macmanus et al., (2021), population represented as all settlements with populations > 100,000 inhabitants from ESRI's 'World Cities' database – accessed 2022.

With sustained population growth comes an increased demand for fresh water, including household, industrial and agricultural uses. A major source of fresh water in LECZs is groundwater. Within these coastal aquifers, saltwater intrusion (SWI) can occur, where freshwater is displaced by largely unusable saline groundwater. Worryingly, this phenomenon can worsen with various anthropogenic drivers, such as excessive groundwater extraction, land subsidence, and sea-level rise (Jasechko et al., 2020; Meyer et al., 2019; Oude Essink et al., 2010). Moreover, historical events such as sea-level transgressions and land reclamation

also affect the current availability and distribution of fresh groundwater (Delsman et al., 2014; van Engelen et al., 2018). These challenges necessitate careful management of LECZ aquifers, and in turn effective management requires reliable and accurate subsurface data. Acquiring such data is challenging given the regional scales involved, which often extend 100's of kilometres along coastlines, and 10's of kilometres landward. **Mapping regional scale fresh-saline groundwater distributions economically, and with the required precision is the topic of this thesis**. **More specifically I investigate the use of airborne electromagnetics for this purpose**. In this chapter, I will first describe in more detail the current threats to fresh groundwater, followed by on overview of known mapping methods and an outline of my research.

1.2 Factors affecting the distribution of fresh-saline groundwater

As this research is focussed on mapping, I will explain the drivers that control the spatial distribution of fresh or saline groundwater within coastal groundwater systems. First, I will describe natural processes, followed by human influences, and finally future challenges relating to the availability of fresh groundwater. For brevity, I will focus on primary drivers, rather than a detailed description of hydrogeological properties or transport processes. For more detailed explanations refer to Werner et al. (2013).

1.2.1 Natural Causes of Fresh-Saline Groundwater Distributions
Over time, a combination of tectonic and climatic events caused global changes in sea-levels; this in turn resulted in the deposition of salt layers and saline groundwater within aquifers (e.g., Delsman et al., 2014; van Engelen et al., 2018). With subsequent sea-level regression, fresh-water infiltration (primarily from rain, also driven by climatic events) into these aquifers resulted in the formation of freshwater volumes. Driven by density contrasts, this freshwater lies above saline groundwater, creating freshwater lenses. Along coastlines, this manifests as a saltwater wedge between the land and the sea. Finally, sea-level rise (natural, or human driven) can cause a landward movement of this wedge, causing a gradual intrusion of seawater.

1.2.3 Anthropogenic Causes of Fresh-Saline Groundwater Distributions
Resulting fresh-saline groundwater distributions from natural causes are further affected by humans through various means. In what is analogous to tectonic events causing global sea-level changes, but on a considerably shorter timescale, recent land reclamation can cause the freshening of saline groundwater – in these cases shallow freshwater lenses can form (e.g., Delsman et al., 2014). Freshening can also occur through the artificial recharge of aquifers, through aquifer storage and recovery programmes (Zuurbier et al., 2015). Of course, human intervention can also cause salinisation, primarily via groundwater extraction through several

mechanisms: 1) the upconing of saline groundwater from excessive pumping , 2) a lowering of the land surface and thus the piezometric head, causing saltwater intrusion, 3) surface sealing and the resulting decline of groundwater recharge and 4) enhanced upconing caused by the drainage peat/clay soils and land subsidence (Werner et al., 2013) – all of which have essentially the same effect as sea-level rise, but again often on a much shorter timescale.

1.2.4 Impacts and threats of groundwater salinization

With an understanding that human activity can rapidly affect the availability of fresh groundwater within LECZs, as well as the knowledge that these areas are expected to see explosive population growth – we can begin to piece together the potential costs of poorly managed (or mapped) aquifers in these areas. An overview of these threats globally are presented in van Weert et al., (2009), and summarised in the following.

A clear threat is simply the availability of fresh drinking water. Besides the obvious problems associated with not having access to fresh water, the consumption of even slightly saline water – potentially in doses that are undetectable to humans – have been shown to cause hypertension (Scheelbeek et al., 2017). Furthermore, depending on plant type and salinity levels, saline water can cause plant stress – reducing, or even eliminating growth altogether (Maas and Hoffman, 1977). Besides toxicity to plants, livestock are also intolerant to various levels of saline water (López et al., 2021). Naturally these effects can affect agricultural yields, which in turn could result in food shortages and financial loss to farmers, potentially leading to migration. There are also broader economic costs. Desalinisation plants are expensive compared to simply extracting fresh groundwater (Cooley et al., 2019), the costs of which will be passed on to consumers – who are often located in low-income LECZ areas (Neumann et al., 2015).

Thanks to ongoing research, it's not all bad news. Recent advancements in groundwater management include aquifer storage and recovery, where fresh water is infiltrated into aquifers for later use (Dillon, 2005), a method that is already used extensively in the United States (David and Pyne, 2020). Furthermore, the extraction of groundwater can be optimised to reduce or eliminate upconing of saline groundwater and land subsidence (Wagner, 1995). Another recent example highlights the possibility of using a network of optimised pipelines for decentralised groundwater extractions in the Netherlands (Willet et al., 2020). All these management strategies share a single trait – they require an excellent (often three-dimensional) spatial understanding of fresh-saline groundwater distributions for effective implementation and ongoing use. That is to say: innovations to manage and safeguard fresh groundwater resources require good data.

1.3 Geophysical Methods to Map Groundwater Salinity

In simple terms, most geophysical methods to map groundwater salinity exploit the fact that saline water is electrically conductive. By measuring the electrical conductivity (EC) of a groundwater sample, typically in SI units as siemens per metre (S/m), empirical relationships can be used to quantify the amount of salinity present (Lewis and Perkin, 1981). The resulting data are often expressed as total dissolved solids (TDS), typically in grams per litre (g/l), or as chloride – as chloride is the main conservative anion of seawater. Not all methods measure the salinity of groundwater on its own, frequently the combined conductivity of lithology and groundwater (as pore water) is measured – known as bulk electrical conductivity (ECb). To convert ECb to groundwater EC (ECw) – and ultimately salinity – petrophysical transformations are used (e.g. Archie, 1942; Waxman and Smits, 2003). There are several methods that map groundwater using EC. In the following I will briefly outline some commonly used techniques. All methods follow the basic principles described above.

1.3.1 Intrusive Methods

The simplest way to map groundwater salinity are intrusive methods, typically using a monitoring well – where ECw measurements are directly measured. Other intrusive methods, such as electrical cone penetration testing (ECPT) provide a vertical profile of ECb (Begemann, 1952), thus inferring both salinity and lithology. Intrusive methods provide excellent vertical resolution, but naturally lack horizontal information.

1.3.2 Non-Intrusive Methods

Mapping groundwater salinity using non-intrusive methods are typically undertaken using ground-based or airborne electromagnetic platforms. As the principles are the same for both, the method will first be described in general – followed by a comparison of ground-based and airborne approaches. An overview of electromagnetic methods for subsurface characterisation is given in Swift (1988).

With electromagnetic techniques, electrical signals are first transmitted from an instrument. This in turn produces secondary currents in the subsurface which are recorded by a receiver. The method is sensitive to EC contrasts of the subsurface – in terms of groundwater salinity, these contrasts highlight areas where fresh, brackish, or saline interfaces contact. The resulting observations require additional processing to translate measurements into usable (ECb) data, a process known as inversion (e.g., Farquharson et al., 2003; Vignoli et al., 2015). Inverting data is a great source of uncertainty, given that an infinite number of models can fit the data – otherwise known as non-uniqueness. Two electromagnetic methods are used – frequency and time-domain – referred to in the following as FEM and TEM respectively.

FEM methods transmit and record currents at discrete frequencies, each frequency relates to conductivity contrasts at different depths (Siemon, 2009). FEM methods are generally used

for shallow (10's of metres depth), near surface high-resolution investigations (e.g., Siemon et al., 2018). TEM operates by transmitting a current (the primary field) through a loop, the current is then switched off and a secondary field is measured from a receiver loop, resulting in so-called decay curves. These curves can be windowed (known as gates) for modelling, where the magnitude of the curves over time relate to subsurface conductivity contrasts. Penetration depths vary according to the system and subsurface conditions, but this method is typically used for deeper (~100's of metres) groundwater investigations (e.g., Barfod et al., 2016; Vandevelde et al., 2018).

Ground-based EM is often undertaken on foot, where equipment is carried into the field and measurements are taken on a point-by-point basis (e.g., Bording et al., 2017). Other methods utilise powered transport, such as a quad bike or a boat, to tow the instrument – taking essentially continuous measurements along a path (Auken et al., 2019). Ground-based EM methods offer useful information on a local scale (as 1-D or 2-D measurements), but on regional scales result in a disconnected understanding of the subsurface. Ground-based measurements also require direct land access which is not always practical in difficult terrain, such as areas with surface water.

Airborne electromagnetics (AEM) provide a solution to these shortcomings. Here, the EM instrument is towed beneath a helicopter or fixed-wing aircraft. Surveys are flown in grid-like pattern at low altitude (~50m above the surface), where flightlines are typically 100 – 500m apart (e.g., Delsman et al., 2018). The measurements are high frequency, such that data are typically measured every ~5m along flightpaths – despite survey speeds of up to ~150km/h. Given the high density of measurements and the grid-like coverage, the inverted results lend themselves well to 3D interpolation – where ultimately 3D volumes of estimated groundwater salinity can be delivered to stakeholders (e.g., Delsman et al, 2018). Several contractors have developed AEM systems, including: SkyTEM, Spectrem, VTEM, Xcite, GEOTEM, TEMPEST, HELITEM and RESOLVE – comparisons between systems are given in Mulè et al. (2019) and Viezzoli and Selfe (2018).

Figure 1.2. Ground and airborne electromagnetic systems.

Top left: Time-domain quad-bike system in the Netherlands, as used by the Geoligical Survey of the Netherlands (TNO-GDN) (photo: self). Top right: the Xcite time-domain AEM system operating in Africa (photo courtesy of NRG at https://www.nrgex.co.za). Bottom: the frequency domain AEM system in The Netherlands, as used by the Federal Institute for Geosciences and Natural Resources, Germany (BGR). Photo courtesy of Deltares at https://www.deltares.nl/.

1.4 Estimating Groundwater Salinity using Groundwater Models

Apart from geophysical methods, groundwater models, or more specifically 3D variable-density groundwater flow and coupled salt transport models (necessarily 3D-VDG models for short), have been used since the early 2000's to estimate groundwater salinity distributions (e.g., Oude Essink, 2001). A 3D-VDG model solves the coupled density-dependent groundwater flow and salt transport equations. Here, density differences, along with hydraulic drivers, cause pressure differences, that initiate groundwater flow and associated advective and diffusive-dispersive transport of dissolved salts. Salt transport in turn affects the salinity and thus density distribution, impacting flow and transport etc. 3D-VDG models can be used to estimate past, present, and future groundwater salinity distributions (e.g., Meyer et al., 2019; Oude Essink et al., 2010; Van Engelen et al., 2019). An advantage of 3D-VDG models is that like AEM, it results in a 3D model of groundwater salinity over regional scales. It

does however suffer from great uncertainty – given the large number of parameters and heterogeneity involved. This uncertainty, however, can in part be reduced with the addition of observations, including both intrusive and geophysical data sources.
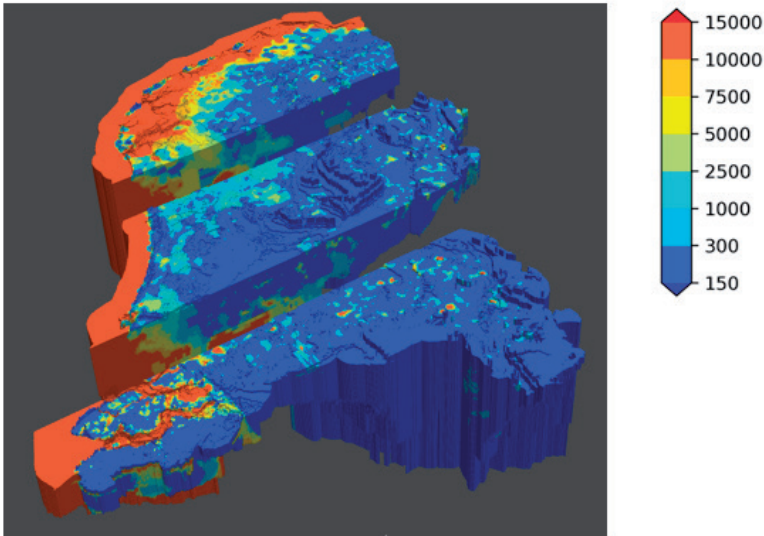
Figure 1.3. A country scale 3D-VDG groundwater salinity model in the Netherlands (Delsman et al., 2020).

1.5 Previous Research

The research presented in this thesis focusses on the use of Airborne Electromagnetics (AEM) to map groundwater salinity. Given the challenges involved, the innovation of this work is sought in combining AEM methods with other data sources such as intrusive well monitoring with 3D-VDG modelling. As a result, in the following I will not focus on technological advancements in the form of equipment (e.g., a new AEM system with improved resolution) – but rather on methods and processing using AEM systems that were available at the time. A general review of using AEM for groundwater exploration is presented in Siemon et al. (2009) and Paine and Minty (2005). For clarity this will be described in chronological order and is, for practical reasons, non-exhaustive – thus will focus only on studies that relate in a broader sense to the work presented in this thesis. More specific details, including an extensive list of references are provided in the chapters that follow.

The first use of AEM to map groundwater salinity was in 1978 on the island of Spiekeroog, Germany in the North Sea (Sengpiel and Meiser, 1981). Here an early three frequency FEM system was used to successfully map the depth of the fresh-saline interface across the island.

The resulting observations were inverted using a so-called resistivity/depth sounding curves, and delivered results were depth-contours of the fresh-saline interface. Later, in 1981 the same AEM system was used to map saline intrusion beneath a riverbed in Pakistan (Sengpiel, 1989). In this study, a centroid-depth algorithm was used to invert the data – an improved approach that was shown to compare well with in-situ data.

A large portion of the Everglades Nation Park, Florida was surveyed in 1998 to delineate coastal saline intrusion (Fitterman and Deszcz-Pan, 1998). The survey covered just over 1000 km2 and was flown by an improved, five frequency FEM system – allowing greater mapping accuracy. The data were inverted using a layered earth method, which was subsequently converted to salinity using petrophysical relationships with help of ground data. This work is one of the first to provide regional, quantitative estimates of groundwater salinity – usefully represented as chloride concentrations rather than just ECb.

In the meantime improved processing methods were developed, including pseudo 2-D inversions (Auken and Christiansen, 2004). This approach links neighbouring model properties to produce laterally coherent results in coastal aquifers and allowed the use of in-situ data as constraints. Initially developed for ground EM surveys, the method was subsequently adapted for AEM surveys (Siemon et al., 2009a) – where it was used to invert FEM data from a survey on Borkum, a North Sea island, in 2008 (Siemon et al., 2009b). The resulting inversions resolved the vertical and lateral extent of a freshwater lens on the island and were subsequently transformed to ECw to determine the volume of fresh groundwater available. Further inversion method developments saw the use of Spatially Constrained Inversions (SCI) for AEM data – where models are also constrained between neighbouring flightlines (Viezzoli et al., 2009) – creating pseudo 3D results. Subsequently, several AEM groundwater salinity mapping programmes utilised these (and other similar) methods effectively (e.g., Ball et al., 2010; Bedrosian et al., 2013; Chongo et al., 2015; de Louw et al., 2011; Faneca Sánchez et al., 2012; Gunnink et al., 2012). A notable example is the almost complete coverage of Denmark by AEM and ground EM measurements, where a subsurface atlas of ECb was effectively created, allowing hydrologically related insights across an entire country (Barfod et al., 2016).

However, there are inherent uncertainties found in the processing of AEM data for groundwater salinity mapping. These are primarily from two sources: 1) the inversion method, given the non-uniqueness of traditional approaches (Minsley, 2011), and 2) error associated from the petrophysical transformation from ECb to ECw (Revil et al., 2017). So far, most AEM investigations had taken a sequential approach to recover salinity of observations – running a geophysical inversion and then performing the petrophysical transformation to recover salinity sequentially. This approach is now known as sequential hydrogeophysical inversion, and has tendency to propagate error from the two steps (Hinnell et al., 2010). Joint, or coupled hydrogeophysical inversions (CHI) were subsequently developed to (in part) try and reduce

these uncertainties and explore more interdisciplinary approaches (e.g., Bauer-Gottwein et al., 2010; Hinnell et al., 2010; Steklova and Haber, 2015). CHI avoids the use of traditional geophysical inversion, whereby a 3D-VDG model is transformed into a physical property that is forward modelled and iteratively compared to AEM observations until a suitably small fit is found. Of course, uncertainty isn't eliminated using this approach as there is still a reliance on a petrophysical transformation – and results are dependent on the choice of conceptual model used for the 3D-VDG model. Delsman et al. (2018) took a probabilistic, sequential hydrogeophysical inversion approach to create a 3D map groundwater salinity across an entire province of the Netherlands. Here, each step of the process was included in a stochastic model, which allowed insights into uncertainties of the whole process – including (amongst others such as interpolation) the inversion and petrophysical transformation. The results of this highlighted that inversions add the greatest source of uncertainty, followed closely by the petrophysical step.

Research presented here highlights that there are an array of available acquisition and processing techniques for mapping groundwater salinity using AEM. Especially encouraging is the notable increasing interest in hydrogeophysics in general (Binley et al., 2015), a sure sign of uptake in the future. However, a quantitative understanding of mapping accuracy over regional scales is not fully understood. This uncertainty includes unknowns about the effect of using different inversion methods, as well as ambiguity regarding the financial benefit of including additional in-situ data to interpret results – or indeed the effect of changing flightline spacing. As groundwater salinity can change rapidly due to anthropogenic stressors, it is also unknown whether AEM surveys are sensitive to these movements over time. If they are sensitive to this, the practical viability of using repeated surveys for time-lapse monitoring is also not understood – and indeed if it is possible, it opens up exciting research avenues. One of them being the potential of using repeated AEM surveys to improve the parameterisation of 3D-VDG models.

1.6 Research Questions and Thesis Outline

The research in this thesis makes use of recently available AEM data, from a survey over the Province of Zeeland, The Netherlands (Delsman et al., 2018). These data are also used to construct synthetic models to help fill research gaps. Flightline locations of these data are illustrated in figure 1.5.
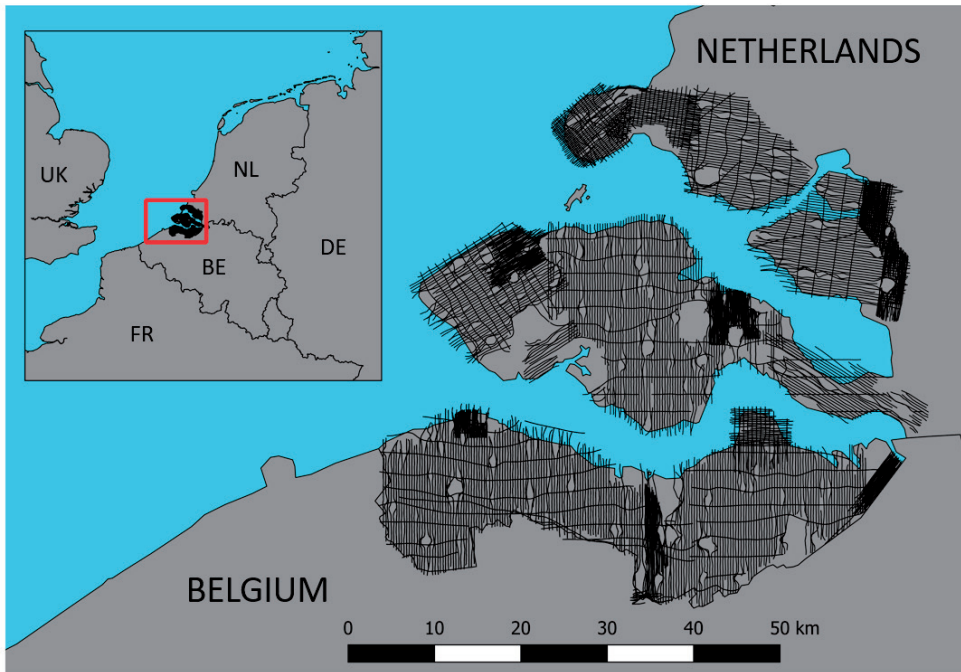
Figure 1.5. Locations of available AEM data for the research presented in this thesis, represented as flightlines.

Understanding uncertainty relating to regional scale AEM groundwater salinity mapping served as an initial target of this PhD research. Based on knowledge obtained from these initial studies, as well as uncertainties identified from previous research, the **main objectives** of this thesis are addressed:

(1) developing improvements on existing processing methods, focussing on combining AEM with intrusive in-situ data;

(2) investigating the potential of flying repeat AEM surveys to monitor groundwater salinity changes over time.

Based on the identified knowledge gaps, research questions were formulated as per the following:

What is the effect of using different inversion methods and parameters on mapping results?
In chapter 2, extensive research highlights the most commonly used inversion methods for groundwater investigations. Using available in-situ data for validation purposes, these inversion methods are then quantitatively tested for a variety of practical mapping outcomes.

How are results affected by different quantities of available data?
Another practical and quantitative study is undertaken in chapter 3, using the results of chapter 2 to guide the inversion method. Here, using a highly realistic synthetic model, the effects of different quantities of in-situ geological data (for estimating formation factors) and flightline spacing are investigated. The results help highlight the relative importance of data and consequently offers financial guidance in regard to survey planning in other LECZs.

What further methodological improvements can we make, based on the results of chapters 2 and 3?
The results from chapter 2 and 3 are revisited, where the usefulness of a customised and novel inversion method to improve mapping accuracy is investigated in Chapter 4. The method is practical as it is based on a commonly used inversion method, but changes the characteristics of the inversion results to better match the groundwater salinity values at the locations of in-situ data.

Are groundwater salinity movements sensitive to repeated AEM surveys?
In chapter 5, AEM surveys and groundwater models are physically coupled to synthetically investigate the sensitivity of AEM data to groundwater movements. Consequently, the time needed for surveys to map groundwater salinity changes are understood, resulting in the development of a novel tool that can improve the parameterisation of groundwater models using time lapse AEM.

2

Quantifying Geophysical Inversion Uncertainty Using Airborne Frequency Domain Electromagnetic Data— Applied at the Province of Zeeland, the Netherlands.

An accurate understanding of the fresh-saline distribution of groundwater is necessary for effective groundwater management. Airborne electromagnetic (AEM) surveys offer a rapid and cost-effective method with which to map this, offering valuable additional information about the subsurface. To convert AEM data into electric conductivity and ultimately groundwater salinity, an inversion is undertaken. A number of algorithms are available for this purpose; however, these are affected by significant uncertainty, owing to inherent nonunique characteristics of this process. The most commonly used inversion codes in hydrogeophysical studies were quantitatively tested using frequency domain AEM and ground data from the province of Zeeland, the Netherlands. These include UBC1DFM code and quasi-2D laterally constrained inversions. Following an investigation of inversion parameter settings, data were inverted for four inversion methods and interpolated into 3-D volumes. Using geological data and empirical electrical conductivity and water salinity relationships, each inversion was converted into groundwater electrical conductivity and split into fresh-brackish-saline regions. For groundwater volume estimates out of a total volume of 2.8 billion m3, a fresh groundwater estimate could differ by as much as 178 million m3, depending on the inversion used. The primary factor here was the choice of model smoothness, which was shown to affect the thickness of the brackish interval. Fresh-brackish-saline interfaces were consistently mapped with an accuracy of ~3 m, the brackish being the most accurately resolved. The few layer method was less successful at resolving smoothly varying salinity distributions but more successful at mapping the brackish interface at greater depth.

2.1 Introduction

Globally, over 600 million people live within low-elevation coastal zones less than 10 m above sea level; this number is estimated to rise to over 1 billion by 2060 (Neumann et al., 2015). As many of these areas are around or below sea level, fresh groundwater resources suffer from saline intrusion—a problem further exacerbated by the threat of sea level rise and land subsidence (Minderhoud et al., 2017; Oude Essink et al., 2010; van Weert et al., 2009). An accurate understanding of the fresh-saline groundwater distribution is therefore required for effective groundwater management. Nonintrusive airborne electromagnetic (AEM) techniques offer a rapid and cost-effective method with which to achieve this, in contrast to conventional ground-based techniques that offer limited spatial resolution at the larger regional scales required. AEM systems were initially developed for mineral exploration (Fraser, 1978) and comprise two primary methods: frequency and time domain—differences and capabilities of each are discussed in Steuer et al., (2009). Early AEM programs include mapping the fresh-saline groundwater interface over a small island in the North Sea (Sengpiel and Meiser, 1981), and later a more advanced AEM system was used to map fresh groundwater in Pakistan (Sengpiel and Fluche, 1992). These methods have since been successfully applied for groundwater mapping campaigns across the world (e.g. Auken et al., 2008; Chongo et al., 2015; Fitterman & Deszcz-Pan, 2001; Siemon et al., 2015). In the Netherlands, the technique was used in 2009 over the island of Schouwen-Duivenland (Faneca Sànchez et al., 2012) and again in 2014–2015 covering the entire province of Zeeland (Delsman et al., 2018).

Frequency domain AEM, commonly referred to as helicopter-borne electromagnetics (HEM), operate by transmitting electrical signals from an instrument towed beneath a helicopter called a bird. These signals produce currents in the subsurface, which are in turn detected by receivers. As a result, the system detects electrical conductivity (EC) within the subsurface—which, in the case of hydrogeophysical studies, relates primarily (but not exclusively) to clay content of the lithology and/or the salinity of the groundwater (McNeil, 1980). To obtain useful quantitative data, AEM measurements need to be converted into an EC distribution, and ultimately a groundwater salinity estimate, or to map hydrostratigraphy—this is undertaken using a process called inversion. In AEM mapping programs, a number of inversion algorithms are available for this purpose, which include 1-D methods (Farquharson et al., 2003; Siemon et al., 2009a; Vignoli et al., 2015), or more complex 2-D (Hermans et al., 2012) and 3-D (Cox et al., 2012; Scheunert et al., 2016) inversions. To validate new inversion techniques (or improvements on existing algorithms), these are typically tested using a synthetic model and then applied to an area where relatively little is known about the subsurface (Auken et al., 2005; Farquharson et al., 2003; Siemon et al., 2009a). Furthermore, inversions used in applied hydrogeophysical studies are often presented without a sensitivity analysis of the effects of available inversion types and inversion parameter settings. Using a probabilistic approach, Minsley, (2011) illustrated that a range of inversion models can sufficiently explain the

measured frequency domain (or observed) data, highlighting the problem of accepting a single result that fits. As academia has observed a year-by-year increase in hydrogeophysical related publications (Binley et al., 2015), it is therefore suggested that the uncertainty introduced by the inversion process on estimated resistivity and derived subsurface properties, including salinity, should be better understood. A comparison of inversions was undertaken by Hodges and Siemon (2008), qualitatively comparing the capabilities of a number of 1-D techniques by means of a synthetic model and real data using two flightlines. This article aims to build on this by quantitatively examining the practical effectiveness and uncertainty of a number of frequently used 1-D inversion algorithms, while making use of 2-D sections and interpolated 3-D models. This is possible due to extensively available HEM and ground data in the province of Zeeland, the Netherlands, where the subsurface distribution of geological and hydrological features is thoroughly understood.

In the following, some of the most commonly used 1-D inversion methods used for AEM were first tested individually for different inversion parameter settings using a test flightline. Second, based on these results, four different inversion methods were selected for further analysis, where remaining flightlines were inverted for each. Third, results were interpolated into 3-D volumes of EC and transformed into groundwater EC using geological information. Finally, inversion uncertainty was analyzed based on practical hydrogeophysical mapping aims, such as groundwater salinity distributions and groundwater interface mapping. Section 2 introduces the study area and available data, with an overview of inversion methods. Section 3 presents the methods and the selection process for inversion parameter settings, followed by the results in section 4 and a discussion in section 5.

2.2 Background

2.2.1 Study Area

Located in southeastern Netherlands within the province of Zeeland, the rhombic-shaped former island of Walcheren (~20 × 20 km2) faces the southern North Sea (Figure 2.1). Along the coast, topography is characterized by elevated sand dunes—reaching a height of around 40 m above sea level. Inland, the so-called creek ridges are apparent. They are the result of topographic inversion by agricultural drainage, which resulted in a larger subsidence of the adjacent peat-dominated low-lying polders than of the former sandy tidal creeks (Pauw et al., 2015). Much of the rest of Walcheren lies at, or below sea level.
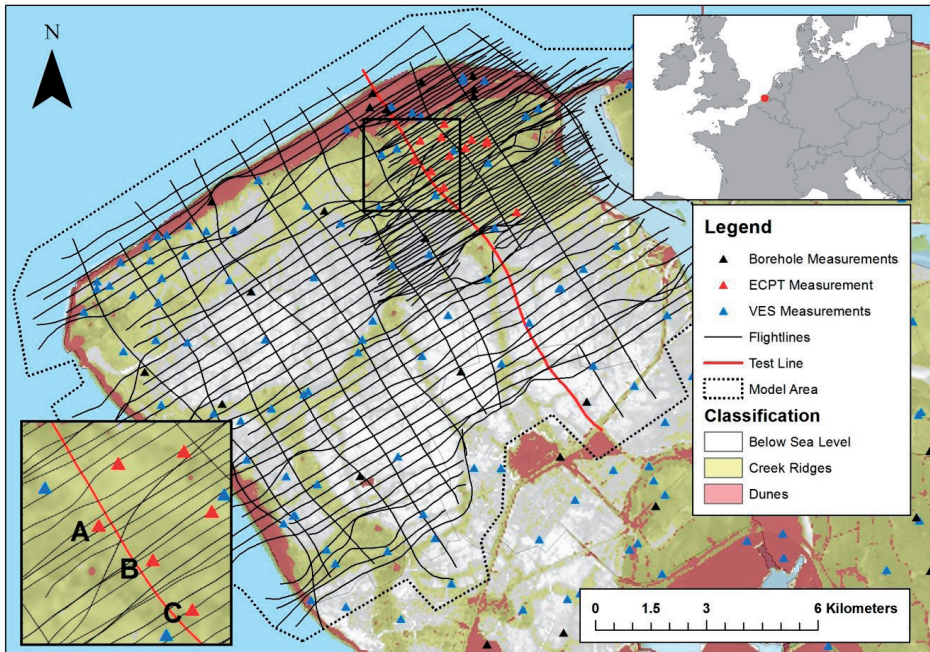
Figure 2.1: *The 2-D horizontal map of the study area, including flightline locations (thin black lines), the test line location (thick red line), ground data (red) triangle: ECPT, black triangle: boreholes, and blue triangle: VES. Dune areas (red) and creek ridges (yellow). Zoomed area highlights labeled ECPT used for analysis. ECPT = Electric Cone Penetration Test; VES = Vertical Electrical Sounding.*

2.2.2 Geology

Geologically, Walcheren is located at the southern edge of the North Sea basin and comprises gently north dipping Neogene and Quaternary sediments (Stafleu et al., 2011). As a result of sequential sea level transgressions throughout the Holocene, these sediments were deposited mostly in shallow marine, estuarine or fluvial conditions (Vos, 2015). Holocene sediments primarily comprise the Naaldwijk formation—consisting of the Wormer and Walcheren members that lie above an eroded peat unit called the Basisveen. The Wormer member is heterogeneous and comprises fine sands and clay. The surficial Walcheren member is separated from the Wormer by a thick peat unit and consists of 1- to 5-m thick clay-rich tidal flats, sand-rich tidal channels that caused up to 45-m deep incisions into the underlying sediments. Along the coast, the Zandvoort and Schoorl members of the Naaldwijk formation form a sandy coastal barrier (Stafleu et al., 2011). The underlying Pleistocene deposits include the sands and silts of the periglacial Boxtel formation deposits, which have in places been eroded by the Walcheren member tidal channels.

2.2.3 Hydrogeology

A current estimate of the fresh-brackish-saline groundwater distribution has been recently undertaken across the Dutch province of Zeeland using the HEM data available for this study (Delsman et al., 2018; Van Baaren et al., 2018). The hydrological situation of Walcheren is largely influenced by early Holocene sea level transgressions and the subsequent construction of man-made coastal defenses (Berendsen, 2005). Sea level transgressions during the Holocene caused extensive salinization, whereby the denser saline water displaced the less dense freshwater in shallow aquifers (Delsman et al., 2014). The construction of coastal defenses and resulting land reclamation prevented further sea water inundation and in turn allowed the freshening of shallow aquifers. Freshening primarily occurs through the formation of rainwater lenses, where less dense fresh groundwater floats on the underlying denser saline groundwater (Goes et al., 2009). In Walcheren, this lensing effect is observed differently in three main areas: (1) 2-m thick lenses in low-lying areas, (2) 5- to 15-m thick lenses underneath creek ridges, and (3) up to ~60-m thick around coastal dunes (Delsman et al., 2018). As freshwater lens thicknesses are controlled by geomorphological features (Goes et al., 2009), they are easily identifiable using topographical data, as observed in Figure 2.1.

2.2.4. Data

The available geophysical and geological data in Walcheren are summarized in Table 2.1 and illustrated in Figure 2.1. The data are described in more detail in the following section.

Table 2.1. *Data Types and Quantities Used in This Study*

Name	Type	Quantity
Helicopter-borne EM (HEM)	Airborne Geophysical	~1000 line km
Electrical Cone Penetration Tests (ECPT)	Ground Geophysical	12 holes
Geo-electrical boreholes	Ground Geophysical	16 holes
GeoTOP	Ground Geological. 3D Geological Model.	From surface down to 50m depth
Vertical Electrical Sounding (VES)	Ground Geophysical	~80 Measurements

Note. EM = electromagnetics.

2.2.4.a

In 2014/2015 a HEM survey was flown over the entire province of Zeeland, the Netherlands, by the Federal Institute for Geosciences and Natural Resources (BGR). The survey was undertaken for the research project FRESHEM Zeeland (fresh salt groundwater distribution by helicopter electromagnetic survey in the province of Zeeland) and a research program involving Deltares, the Geological Survey of the Netherlands (TNO) and BGR. The survey covered an area of ~2,000 km2, totaling over 9,000 line kilometers. The primary aim of FRESHEM was to translate the HEM data into a 3-D volume of fresh-saline water distributions in the subsurface (Delsman et al., 2018).

The survey was flown with the following parameters by BGR's RESOLVE System (manufactured by Fugro Airborne Surveys, Table 2.2).

Table 2.2. *Instrument Parameters and Specifications Used by BGR for the HEM Survey*

Towed bird HEM system specifications	
Frequencies	380, 1770, 5410, 8300, 41000 and 129500 Hz
Coil separation	~8m/9m (depending on conductivity)
Coil orientations	5 x horizontal coplanar, 1 x vertical coaxial
Depth of investigation	50 - 150m max
Measurement spacing	~4m downline
Bird type and manufacturer	RESOLVE, BKS60. Fugro Airborne Systems
Survey parameters	
Flight line spacing	300m (100m or 200m in some areas)
Bird height	~40m

Note. BGR = Geosciences and Natural Resources; HEM = helicopter-borne electromagnetics.

Over the study area in Walcheren, available airborne data totals ~1,000 line kilometers, covering an area of ~270 km2—as reading is taken approximately every 4-m downline, in total about 250,000 individual airborne measurements were available. Line spacing was mostly flown at 300 m across most of the area; however, to the north this was decreased to 100 m over an area of interest called the Waterfarm. Tie lines were flown perpendicularly to these at 1,000-m spacing. Flightline orientation was selected to be perpendicular to the strike of prominent features such as dunes and creek ridges and was therefore oriented NE-SW or NW-SE for tie lines. As inversion results directly depend on data quality, AEM data requires careful processing to minimize influence from measurements that do not reflect the targeted subsurface properties. This process is summarized as follows: (1) data were calibrated in the field to more accurately convert observed data into parts per million (ppm) relative to the primary field; (2) measurement drift was corrected during the survey and postprocessing; (3) microleveling was applied to remove striping in the data; and finally, (4) based on a data quality assessment, selected data were removed—such as those affected by man-made infrastructure or where the altitude of the HEM system was increased for safety reasons. Detailed reviews and processing steps undertaken for HEM surveys are available in Siemon et al., (2011).

BGR's RESOLVE HEM system generates primary magnetic dipole fields through transmitter coils, which in turn induce eddy currents in the subsurface. The coils are housed in a cigar-shaped instrument towed beneath a helicopter called a bird, which typically is about 10-m long. Each of the transmitter coils produces sinusoidally varying currents at discrete frequencies, ranging from 380 Hz to 130 kHz. Low and high frequencies relate to deep and shallow features, respectively. The eddy currents generate secondary magnetic fields based on the conductivity distribution of the subsurface, which are measured by receiver coils. Due to the induction process within the Earth, there is a small phase shift between the primary and secondary field, that is, the relative secondary magnetic field is a complex quantity having in-phase (I) and quadrature (Q) components. The measured secondary field is very weak compared to the primary. Therefore, the primary field is bucked out and receiver coils record the data in parts per million relative to the primary field. Coil orientation is selected according

to survey requirements and is either vertical (VMD: vertical magnetic dipole) or horizontal (HMD: horizontal magnetic dipole); corresponding receiver coil orientation is coupled to that of the transmitter. The system's depth of investigation (DOI) ranges approximately from 50 to 150 m in conductive or resistive conditions respectively. A more detailed review and description of the RESOLVE system are available in Siemon, Christiansen, et al. (2009).

2.2.4.b Electric Cone Penetration Test
Twelve Electric Cone Penetration Test (ECPT) measurements were used for this study, all located in the densely sampled area called the Waterfarm. These were mostly measured at, or adjacent to creek ridges. Data were sampled vertically every 0.5 cm up to a depth of around 25 m below surface. Originally developed for geotechnical surveys, cone penetration testing (CPT) forces a cone-shaped tool vertically into the ground at a controlled rate, whereby resistance and friction are measured to obtain a vertical soil profile (Begemann, 1965). With the addition of an EC sensor (henceforth ECPT), a profile of EC is measured. ECPT has been used extensively as a high vertical resolution tool for hydrological mapping purposes (Gunnink et al., 2012; De Louw et al., 2011).

2.2.4.c Boreholes.
Located throughout the study area, 15 geo-electrical borehole measurements were available. Importantly, a number of these were measured to a greater depth in the dunes areas to the north, where a deeper fresh-brackish groundwater interface is expected. Measurement depths range from 10 m in the central area to ~45 m in the coastal dune areas. The method works by using an electrical probe, where a measurement is taken every 5 cm and measures conductivity of both the groundwater and lithology and, therefore as with ECPT data, measures bulk EC. A detailed description of the method is available in (Spies, 1996).

2.2.4.d Geological Data
The Geological survey of the Netherlands (TNO) hosts a national database of geological data and models (dinoloket.nl—accessed November 2016). TNO's GeoTOP provides a high-resolution 3-D volume of the onshore part of the Netherlands up to 50-m depth. Available as a 3-D voxel model, the volume contains numerous properties for each cell at a horizontal and vertical resolution of 100 and 0.5 m, respectively. Properties used for this study were lithoclasses, where an estimate of the following classes was available for each cell: peat, clay, sandy clay and clayey sand, fine sand, medium-grained sand, coarse-grained sand, and gravel. A more detailed description of the geological model and how it was constructed is available in Stafleu et al. (2011).

2.2.5 Geophysical Inversions
Geophysical inversions calculate a distribution of physical properties based on a set of observations or measured data. In electromagnetics, this is fitting a calculated electromagnetic response to a set of observed values. The resulting physical properties can be presented in

SI units as either conductivity in Siemens per meter (S/m) or the reciprocal—resistivity in ohm-meter (Ω m). For consistency, conductivity as S/m will be used to present results and describe data in this paper. In order to delineate features in the subsurface, a physical property contrast is required—in the case of EC, these are present as clay content, water saturation, water salinity, and temperature (McNeil, 1980).

As hydrogeophysical investigations are often conducted in areas with young, horizontal or subhorizontal substrata, physical property contrasts are generally horizontally smoothly varying. Furthermore, depending on system altitude, system frequency, and subsurface EC (Tølbøll and Christensen, 2007), HEM systems have a relatively small footprint (~100–200 m); therefore, the benefits of 2-D (e.g., Boesen et al., 2018; Li et al., 2016) and 3-D (e.g., Cox et al., 2012; Scheunert et al., 2016) inversions are limited and considered to be impractical compared to the less computationally expensive 1-D techniques. As a result, this study will focus on 1-D inversion methods only. A number of these are available for HEM (e.g, Auken & Christiansen, 2004; Brodie & Sambridge, 2006; Farquharson et al., 2003; Huang & Fraser, 1996; Macnae et al., 1998; Siemon, Auken, et al., 2009; Viezzoli et al., 2008; Vignoli et al., 2015; Yin & Hodges, 2007). However, for practical purposes, it was decided that a thorough analysis of all available 1-D methods was not possible. Instead, a literature review of applied AEM-based hydrogeophysical studies was undertaken in order to highlight the most commonly used methods. Results of this are presented in Figure 2.2.
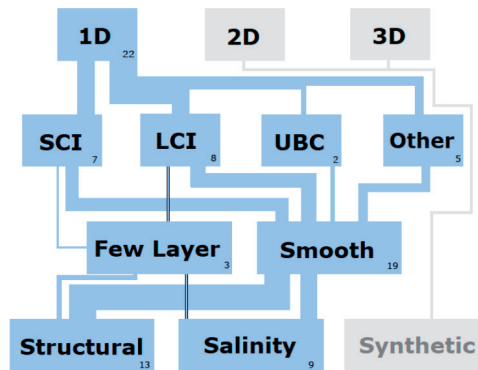
Figure 2.2. Results of a literature review studying commonly used inversion techniques and mapping objectives.

SCI = spatially constrained inversion, LCI = laterally constrained inversion, UBC =University of British Columbia EM1DFM. Other refers to studies that were either unclearly specified or used other older methods such as homogeneous half-space inversions. Structural = mapping of hydrological units. Salinity = mapping of saline distributions or salinity interfaces. (Auken et al., 2008; Ball et al., 2010; Bedrosian et al., 2013; Brodie et al., 2004; Chongo et al., 2015; De Louw et al., 2011; Delsman et al., 2018; Faneca Sanchez et al., 2012; Fitterman & Deszcz-Pan, 2004; Gunnink et al., 2012; Haider et al., 2014; He et al., 2014; Herckenrath et al., 2013; Hill., 2011; Jørgensen et al., 2012; Rasmussen et al., 2013; Sengpiel and Fluche, 1992; Siemon, Auken, et al., 2009; Siemon, Christiansen, et al., 2009; Siemon et al., 2015; Wynn, 2005).

Here it was found that over ~50% used either laterally constrained inversion (LCI) or University of British Columbia (UBC) EM1DFM (Farquharson et al., 2003) inversions, while another ~30% used the spatially constrained inversion (SCI) method (Auken and Christiansen, 2004; Viezzoli et al., 2008).. The SCI process requires a number of neighboring lines to be inverted in order to test measurements along a single flightline, a process that is further complicated because the regularization term needs to be selected according to flightline spacing—where the Waterfarm area has a line spacing of 100 m, in contrast to the 300-m spacing flown for remainder of Walcheren. It was therefore decided that a robust analysis of SCI was impractical for this study.

Both LCI and EM1DFM inversion methods recover a layered, discretized model of the subsurface, where a distribution of EC and corresponding depths are recovered—known as 1-D layered half-space inversions (Sengpiel & Siemon, 2000). Here these inversion algorithms will be subcategorized into two types: (1) smooth, multilayer inversions (Constable, 1987; Farquharson et al., 2003) and (2) layered, few layer inversions (Auken and Christiansen, 2004). Smooth, multilayer inversions (Type 1) have fixed thicknesses, whereby the starting model layer thicknesses are preserved. Layered, few layer inversions (Type 2) invert for both layer thicknesses and conductivity values and have between three and nine layers. Both types take all frequencies into account. Multilayer (Type 1) inversions aim to produce smooth, minimum structure models that lack sharp boundary definition, whereby the smoothest model that fits the data is sought (Farquharson et al., 2003). In contrast, few layer (Type 2) inversions result in sharp boundaries as a result of the inclusion of depths as a parameter and are therefore suitable for areas with more horizontally continuous formation boundaries (Auken & Christiansen, 2004).

In these inversions, there are generally more unknowns than observed data, resulting in an infinite number of possible outcomes, otherwise known as nonuniqueness. In mathematics this is referred to as an ill-posed or underdetermined problem (Tikhonov & Arsenin, 1979). These algorithms therefore need stabilization, otherwise known as regularization (e.g., Constable, 1987; Marquardt, 1963). Regularization is introduced through the model norm, which favors specific properties in the inverted model. The balance between fitting the data and the regularization is approached as an optimization problem—conceptually, this can be expressed as minimizing the objective function (Farquharson and Oldenburg, 2004):

$$\phi(m, p) = \phi_d(m, p) + \beta \phi_r(p) \qquad (1)$$

where the vector p contains parameters of the Earth model, m is a vector with the observed data, ϕ_d is data misfit, ϕ_r is a regularization term based on the difference between current parameter values and the initial estimate of the parameters of the Earth model, and β is the regularization parameter that balances ϕ_d and ϕ_r. If the data are fitted too well, artifacts could

be present in the final model, resulting in unrealistic structures. Conversely, a fit too far from observed data will result in a model that opts to fit data according to the regularization term used instead. This effect is generally referred to as a trade-off, whereby the regularization term can be set to be fixed if noise levels are understood or calculated automatically (Farquharson and Oldenburg, 2004). In smoother, minimum structure methods, a fit that favors the trade-off parameter will result in a smoother model. The addition of lateral and spatial constraints used in LCI and SCI methods allows the adjustment of the inversion input parameter to favor proximity to neighboring models (Auken and Christiansen, 2004; Viezzoli et al., 2008).

As well as a regularization term, layered half-space inversions require a conceptualized model of the subsurface, or a starting model, where the Earth is discretized into a set number of layers, each of which are given property and thickness values. Starting models can be based on prior knowledge, such as known (hydro) geological boundaries (Brodie et al., 2004), or automatically generated based on observed data using sounding curves (Delsman et al., 2018; Gunnink et al., 2012; Sengpiel & Siemon, 2000); mostly, however, these are based on a simplified conceptual model of the subsurface (Auken et al., 2008; Bedrosian et al., 2013; Chongo et al., 2015). The type of inversion, starting model, and inversion parameter settings relating to the regularization term will affect the inversion's outcome. Based on this, the inversion will then iteratively change the starting model until the observed data are sufficiently explained. The response of the current model at each iteration is then calculated by forward modeling and compared to the measured (or observed) data, where a misfit is calculated. The inversion is performed iteratively through a linearized approximation of the nonlinear forward mapping of the model to the data space (Auken et al., 2005; Menke, 1989). The inversion process stops once it reaches a specified trade-off threshold, that is, minimizing a criterion such as equation 1, or until the maximum specified number of iterations is reached. In order for the resulting inversion model to be considered as a potential candidate of the subsurface, a suitable fit to the data is necessary—referred to here as the misfit between observed and predicted data.

UBC inversions comprise four methods, all of which are smooth, multilayer types, and have the same base formulas. However, each method differs in regard to how the trade-off parameter is dealt with (Farquharson et al., 2003; Farquharson and Oldenburg, 2004). These encompass the following: (1) Generalized Cross Validation (GCV), (2) L curve, (3) Fixed Trade-Off, and (4) Line Search (Farquharson & Oldenburg, 2004) and will be described in more detail in that order. By searching for a model that is least affected by any other data point, the GCV method provides an estimated value of the trade-off parameter that sufficiently explains the data. As a result, an estimation of noise is provided (Haber and Oldenburg, 2000). For the L-curve method, if an inversion misfit is plotted against the model norm for a selection of appropriate trade-off values, an L-shaped curve is produced. Here the point of maximum curvature approximately depicts a balance between misfit and model norm and is therefore considered an appropriate trade-off value (Farquharson & Oldenburg, 2004). In contrast to

GCV and L-Curve methods, Fixed Trade-Off and Line Search methods require an inversion input parameter that assumes some knowledge of noise. The Fixed Trade-Off method is the simplest approach, where the user provides a value that remains fixed throughout the inversion. As a result, this method has the fastest calculation times. It may, however, require a time-consuming trial-and-error approach. As the GCV method results in a determination of trade-off values along a flightline, these could be used as values for the Fixed Trade-Off method. The Line Search approach is most useful if the noise in a set of observation is understood. Here the expectation of the misfit equals the number of observations, and the trade-off parameter is chosen according to a target misfit in relation to this (Farquharson & Oldenburg, 2004). A more technical description of regularization estimation for UBC EM1DFM inversions is available in Farquharson and Oldenburg (2004).

Conceptually, the LCI method consists of linked 1-D models along individual flightlines, where the model spacing is determined by downline measurement spacing. As a result, the models are inverted as one system along each flightline (Auken et al., 2005). LCI inversion input parameters (or trade-off parameters) impose lateral and vertical constraints that tie together resistivities and depths of adjacent layers (Auken et al., 2005). As a result, LCI inversions favor laterally continuous physical property distributions, as demonstrated in Hodges & Siemon (2008). The primary inversion parameters are therefore resistivities, or in the case of few layer methods—also depths, as demonstrated in Figure 2.3, (adapted from Auken et al., 2005).
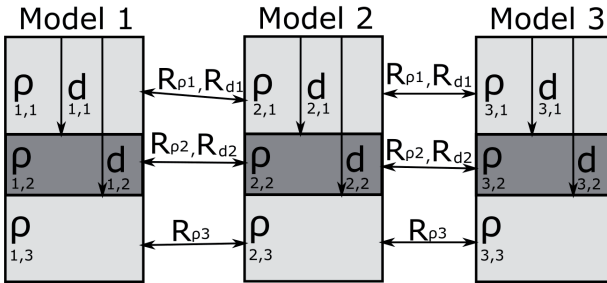
Figure 2.3: *A conceptual figure of the LCI method demonstrating how neighboring models are linked. R = lateral constraints, ρ = resistivity values, and d = depth values. The LCI few layer approach links both depths (d) and resistivities (ρ); the LCI multilayer method links only resistivities (Auken et al., 2005). LCI = laterally constrained inversion.*

Furthermore, LCI inversions allow a sharp option, whereby vertical resistivity changes are permitted to change more rapidly and could be seen as a compromise between few layer and smooth inversions (Vignoli et al., 2015). LCI inversion input parameters are described in detail by Auken et al. (2005) and Vignoli et al. (2015) and will be described and analyzed here as three primary options: (1) horizontal (2) vertical, and (3) sharpness constraints. These inversion input parameters may be chosen based on knowledge of the subsurface, for

example, over Walcheren we know that rapid vertical changes in EC are expected within the scale of a few meters, due to expected fresh-saline groundwater interfaces (Delsman et al., 2018). Horizontal variations are expected to be smoother, in the order of around 10m, due to gently dipping Holocene and Pleistocene sediments (De Louw et al., 2011; Vos, 2015).

In order to effectively assess inversion uncertainty, a misfit threshold between observed and predicted data is needed. This allows differences between the methods and inversion input parameters to be analyzed with the knowledge that they fit the data and thus are all correct, where each quantitatively represents a possible model of the subsurface. Although the data for this study is of high quality and undergone extensive processing (Delsman et al., 2018), field data will nevertheless always be contaminated with some degree of noise (Siemon et al., 2009a). As a result, flexibility in regard to misfit is required, where an estimate of noise can inform the threshold misfit value for inversion results. For this study an approximate value of 5% will be used based on previous estimates (Farquharson et al., 2003), and therefore, in this paper results that do not fit this criterion will be excluded. In the following, misfit is calculated as relative %, using the method of Siemon, Auken, et al. (2009).

2.3 Methodology

Inversions were undertaken using two platforms: (1) UBC's EM1DFM code through the UBC Geophysical Inversion Facility (UBC GIF, Farquharson et al., 2003) and (2) LCI inversion methods using Aarhus Workbench (Auken and Christiansen, 2004; Vignoli et al., 2015). Initially, eight separate inversion types were tested using a single flightline, encompassing each of the four EM1DFM trade-off options, two LCI multilayer inversions, the smooth and sharp inversions, and the few layer LCI method with either five or nine layers.

The process is summarized by an eight-part workflow as illustrated in Figure 2.4 and discussed in order in sections 3.1 to 3.8.
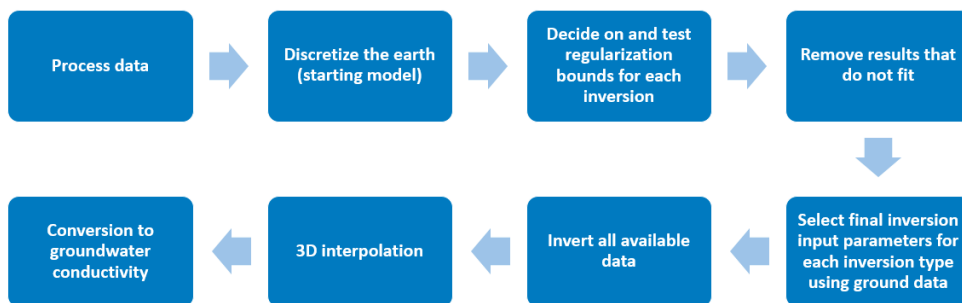
Figure 2.4: *The eight-part methodology workflow used for this study.*

2.3.1 Processing of HEM Data

Standard processing of the HEM data was undertaken by BGR and, for this survey, is summarized in Delsman et al. (2018) and detailed in section 2.4.1. The following postprocessing steps were undertaken for this study: (1) measurements where bird altitudes above 60 m were removed due to low signal to noise ratios, and (2) every second downline measurement was removed, resulting in a measurement every 8 m to facilitate faster inversion times. As the footprint of the HEM system is ~100 m to 200 m (Tølbøll and Christensen, 2007), an 8-m measurement distance is assumed to have no or little effect on the outcome, while allowing practical computation times given the amount of inversion runs required for this study.

2.3.2 Inversion Setup

All data were inverted using measured in-phase (I) and quadrature (Q) secondary field data. To minimize ambiguity, it was decided that the same starting model was to be used for all inversions. For smooth and sharp multilayer inversions, these were chosen to consist of 20 layers, with the thickness of the top layer starting at 0.5 m and subsequent layers increasing logarithmically with depth till 60 m. These layers are parallel to Dutch AHN25 topography data, a high-resolution LIDAR topographical data set of the Netherlands. For few layer inversions, the same criteria were applied but using 5 layers instead of 20. The 60-m depth was chosen according to borehole measurements in the dune areas, where they indicated a probable fresh-saline groundwater interface at a maximum depth of ~50 m, based on observed strong conductivity contrasts. This depth also reflects the DOI capabilities of the HEM system. For inversions undertaken in Aarhus Workbench, a starting conductivity of 1 S/m was selected owing to the highly conductive saline environment of Walcheren. For UBC inversions, the starting conductivity was calculated automatically for each inversion as this was found to improve stability, particularly in areas where apparent shallow, highly conductive features were present (Farquharson and Oldenburg, 2004).

2.3.3 Regularization Parameter Choices

For practical purposes, inversion types and parameter inputs were initially tested on a single flightline. Flightline 9.9 (Figure 2.1) was selected for this purpose, as it covers a range of hydrological settings, including a shallow, highly conductive area to the south-east, and crosses perpendicularly to the more resistive dune area to the north-west. It was also found that bird altitudes were relatively high in some places, likely because of proximity to anthropogenic features, and therefore offers a range of mapping challenges, many of which would typically be found in a HEM survey. Furthermore, this flightline is in close proximity to available ground constraints. Inversion parameter input options for each inversion type are summarized in Table 2.3, alongside the ranges of values used for this study. For each inversion method, a significant number of parameter input combinations are possible; therefore, parameter values were changed one parameter at the time, while keeping the other parameters at their defaults (ceteris paribus). In total over 70 separate inversions were run for this process.

Table 2.3. *Parameter Sets Tested Along With Descriptions and Ranges to be Used.*

Inversion Name	Parameter Name	Description	Ranges	Reference
UBC Fixed Trade	Trade Off	Trade-off value	0.5 - 3	
All UBC Inversions	Decrease	Greatest allowed decrease in the trade-off parameter per iteration	0.011 - 1	Farquharson and Oldenburg, 2004
UBC Line Search	Chifact	Trade off value	0.5 - 10	
LCI Smooth	Later/Vertical Constraints	Amount that neighbouring constraints can change. A higher numerical allows more change and vice versa.	Lateral: 1 – 3, Vertical: 1.5 - 3	Auken et al., 2005
LCI Sharp	Sharp Numeric	Vertical and horizontal sharpness of the model	Vertical: 1 – 100, Horizontal: 1.5 - 150	Vignoli et al., 2015

Note. UBC = University of British Columbia; LCI = laterally constrained inversion.

Initial upper and lower bounds for inversion parameter settings were decided on practical criteria such as inversion stability, observing data misfit and previously used values in published literature (Bedrosian et al., 2016; Gunnink et al., 2012).

2.3.4 Removing Inversions Results From Further Analysis

Out of the initial ~70 inversion runs, a subset of ~25 was selected for further analysis. Inversions were removed for one of three reasons: (1) a noticeable change in the inversion outcomes was not observed (i.e., adjusting the inversion input parameter resulted in almost no difference compared to the next inversion and therefore, for practical reasons, was not included in the analysis); (2) results showed apparent instability, and (3) the inversion did not converge to within the 5% misfit threshold and therefore was excluded as a potential candidate of the subsurface. Figure 2.5 shows observed and predicted data for a LCI smooth inversion result, where misfits are in the 1–3% range; misfit plots for the other three inversions are available as supporting information.
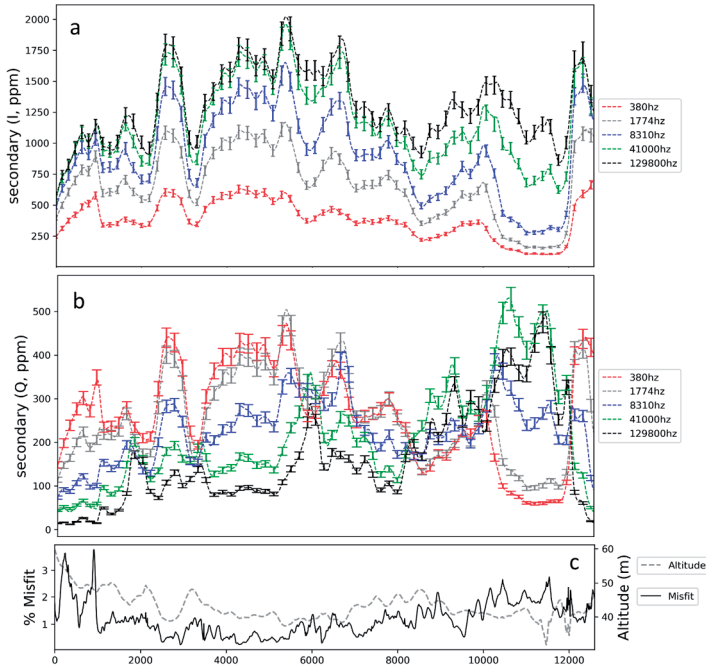
Figure 2.5: *LCI smooth inversion results, lateral and vertical constraints input as 1.3 and 3, respectively. Predicted (dashed lines) and observed (as error bars as 5% of data magnitude) for in-phase (a) and quadrature (b) components. Misfit (c) shown as solid black line as relative % error, and instrument altitude as gray dashed line. LCI = laterally constrained inversion.*

Out of the four UBC trade-off options, the Fixed Trade-Off inversion was consistently more stable and converged successfully. It was therefore decided to only use the UBC Fixed-Trade method for further analysis; however, the GCV was useful at suggesting an appropriate inversion input parameter. The 9 layers used in the LCI few layer approach were discarded despite successfully converging, as conceptually there are between 5–10 degrees of freedom within the data space (encompassing both in-phase and quadrature components) and 18 parameters in the inversion. Therefore, without prior constraints, the inversion could not be trusted.

As a result of removing inversions based on the criteria above, four methods remain: (1) 20-layer LCI smooth, (2) 20-layer LCI sharp, (3) LCI five layers, and (4) 20-layer UBC fixed trade-off.

2.3.5 Testing Regularization Terms

In order to the assess the robustness of the remaining four inversion methods with different regularization parameters, 2-D sections of inversion results with standard deviations of

conductivity were plotted for each of the regularization parameter sets, along with misfits. The variance plots illustrate areas in the inversion result that showed the largest (or least) change in the inversion outcome by changing the regularization term (or inversion input parameter) between the bounds shown in Table 2.3. Figure 2.6 illustrates two examples of this, one from the UBC Fixed Trade parameter and another from the LCI smooth lateral/vertical constraint options. These are plotted with a single inversion result each to assess the overall structure of each model.
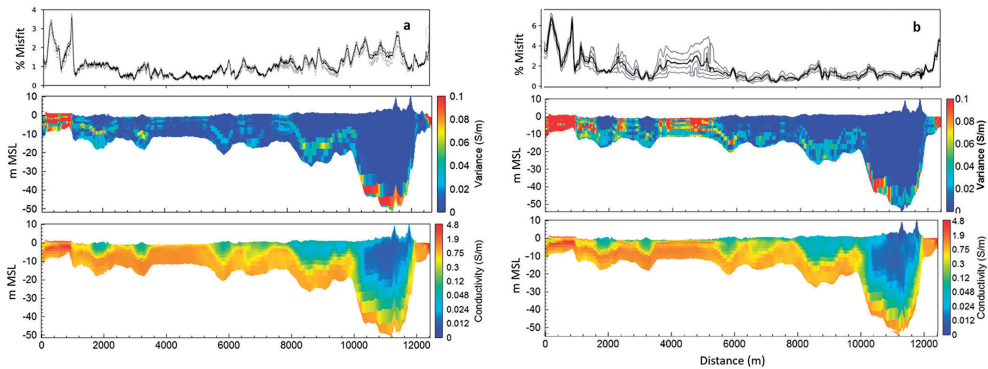

Figure 2.6: *Variance plots testing inversion robustness based on regularization parameters with misfit to observed data. (a) LCI inversion, lateral/vertical constraints variance with inversion result. (b): UBC Fixed Trade variance with inversion plot (Fixed Trade-Off parameter = 2). Misfit is represented relative difference as percentage. The blanked out area is the result of applying the DOI calculation from the LCI smooth result. LCI = laterally constrained inversion; DOI = depth of investigation.*

2.3.6 Selecting Inversion Parameters for Each Code

As observed in the variance plots (Figure 2.6), predictably the inversion parameter settings for each inversion type resulted in easily observable differences, with standard deviations observed at just over 0.1 S/m in some areas. This effect was more prevalent in three groupings: (1) shallow, highly conductive regions on either end of the profile, (2) intermediate zones with distinct linear characteristics, and (3) deeper areas, often at the base of the DOI toward the north-west. For these reasons, final input parameters were selected based on an unbiased quantitative means. This was done by comparing all inversion results to nearby ECPT data. Searching to a maximum distance of 150 m to the test line, three ECPT measurements were used for this purpose (labeled A–C in Figure 2.1), whereby the 10 nearest 1-D inversion models were averaged for each ECPT location to smooth out single anomalous inversion results. Finally, these were compared directly with one another, where absolute differences were calculated by subtracting ECPT data with each inversion result. To demonstrate the method, results from the LCI smooth parameter (lateral/vertical constraints) set are illustrated in Figure 2.7.
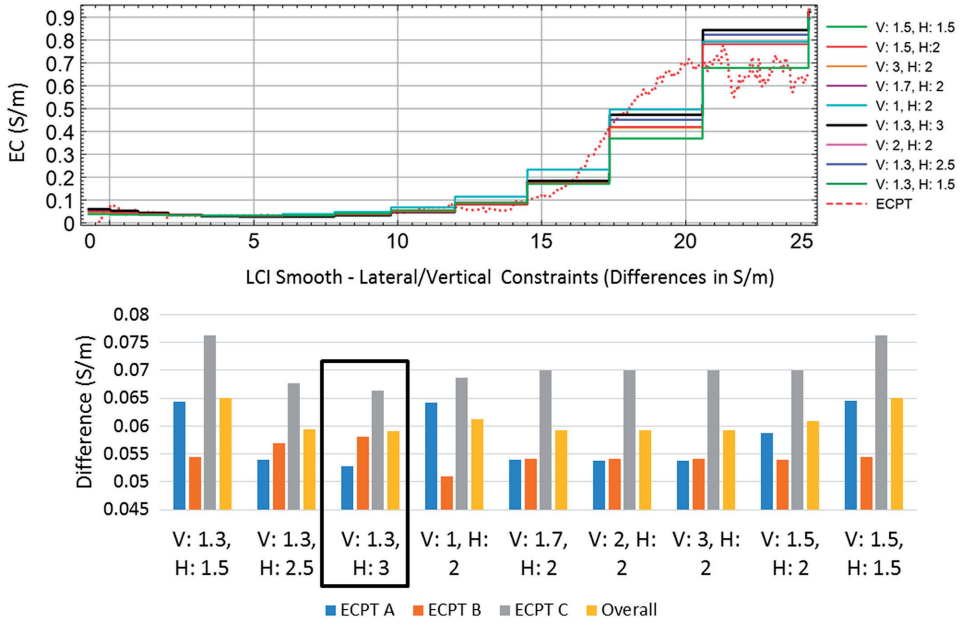
Figure 2.7: *An example vertical section of LCI lateral/vertical constraints input versus ECPT C (ECPT locations illustrated in Figure 2.1), parameter set (V: 1.3 H:3; see Table 2.3 for an explanation of these parameters) that resulted in a closest match is highlighted in black. (bottom) Mean fit between inversion model EC and ECPT data, final choice highlighted by the black box. Misfit is based on average absolute value mismatch in Siemens per meter. LCI = laterally constrained inversion; ECPT = Electric Cone Penetration Test; EC = electrical conductivity.*

To reduce ambiguity and bias, misfit to each individual ECPT was not analyzed in detail. Rather, the best fitting parameters were selected based on an overall best fit, which is the average difference of each of the three ECPTs. The inversion types and corresponding best fitting regularization parameters for each of the finally selected inversion method are listed in Table 2.4. On the assumption that these parameters would be best suited for the entire area, these were chosen for further analysis and inversion of the entire data set. The best regularization results from the LCI multilayer inversions were also applied to the few layer inversion.

Table 2.4. Final Regularization Parameter Choices Based on a Quantitative Match to ECPT Data on Flightline 9.9.

Final Inversion Parameters	Trade Off	Decrease	Chifact	Lateral & Vertical Standard Deviation	Power Law	Vertical & Horizontal Sharp Numeric
UBC Fixed Trade	3	0.5	N/A	N/A	N/A	N/A
LCI Smooth	N/A	N/A	N/A	L: 1.3, V: 3	1	N/A
LCI Sharp	N/A	N/A	N/A	L: 1.3, V: 2	1	V: 30, H: 15
LCI 5 Layer	N/A	N/A	N/A	L: 1.3, V: 3	1	N/A

Note. ECPT = Electric Cone Penetration Test; UBC = University of British Columbia; LCI = laterally constrained inversion.

2.3.7 3-D Interpolation

Having a continuous 3-D distribution of properties permits the use of a versatile environment where properties can be analyzed in a 3-D GIS workspace, allowing the undertaking of a comprehensive analysis such as fresh groundwater volume estimates and fresh-brackish groundwater interface depths. This article's focus is on quantitatively comparing inversion algorithms and parameters thereof, rather than interpolation methods. The interpolation method therefore needed to be simple, unambiguous, and consistent, while honoring the following minimum requirements: (1) conform exactly to data directly beneath sounding locations, avoiding excessive smoothing (except for resolution de-sampling); (2) layered nature of the inversions should be maintained with minimal smoothing; (3) resolution of the 3-D model should closely reflect the resolution of the HEM system, while maintaining a reasonably sized model that is not too bulky; and (4) lightweight enough that all four inversions can be practically interpolated and merged into one voxel model.

For these purposes, a method similar to that of Pryet et al. (2011) was selected where the conductivity distribution of each inversion layer is gridded using 2-D horizontal kriging. Using the depth information from starting models (and layer depths of few layer inversions), each 2-D conductivity grid could be filled vertically into a 3-D voxel. A 2-D ordinary kriging method as coded in Geosoft's Oasis Montaj was used to interpolate conductivity values, where a spherical semivariogram model was used. A 50-m horizontal resolution was deemed appropriate considering the footprint of the HEM system and also conformed to the resolution of the GeoTOP model (i.e., the lithological model used for formation factor (FF) values; see section 3.8). The resulting grids were saved to a database for each inversion, where depth values for the smooth inversions were added from starting models. For the few layer inversions, the inverted depths were gridded using the same kriging method as above; the resulting grids were then sampled into their respective databases. Finally, the depth values for all inversions were corrected for terrain using AHN 25-m resolution data. With each inversion now saved into a database with XYZ coordinates at 50-m horizontal resolution, the data were translated into a 3-D volume. This was done in Paradigm's Gocad, where a $50 \times 50 \times 0.5$m resolution voxel was created over the study area. As the final layer for each inversion is assumed to be infinitely deep, these properties were filled to approximately twice the thickness of the layer before it. To take into account the penetration depth of the HEM system, the standard DOI result from the LCI smooth inversion was applied to the 3-D volume, where values below this depth were removed. For overall consistency, it was decided that the same DOI result was to be applied to all inversions; this includes removing data that were outside of Geotop's geological model, that is, offshore.

The 3-D interpolation result was extensively validated where overall it was found to be appropriate for further analysis and adhered to the initial requirements. A detailed description of the method and validation results are available as supporting information.

2.3.8 Translating the Conductivity Volume Into Salinity Distributions

In order to interpret the results with respect to groundwater quality, it is necessary to relate the conductivity distribution to that of EC groundwater, rather than simply the conductivity of saturated sediments, or bulk EC. To split the signal, bulk EC is multiplied by the FF, which relates to the lithology, and is primarily a function of pore space and pore connectedness (Archie, 1942). This method is further improved by taking into account factors such as surface conductivity as a result of clay particles (Revil and Glover, 1998). Given that this study examines inversion methods only and therefore does not aim to map salinity directly, it was decided that a simplistic approach was suitable for relative comparisons. An apparent formation factor is commonly used for this purpose, whereby the ratio between EC bulk and EC groundwater based on local observations is utilized (Siemon et al., 2009b). De Louw et al. (2011) collected information from seven different soil types over the Province of Zeeland, including measurements in Walcheren that were found to adequately represent the lithology of the study area. The resulting apparent FF values will be used for this study and are listed in Table 2.5.

Table 2.5. *Apparent Formation Factor (FF) Values Used in This Paper, From De Louw et al. (2011)*

| Lithology	Formation Factor (FF)
Peat	2.1
Clay	2.5
Sandy Clay/Clayey Sand	2.8
Fine Sand	3.2
Medium Coarse Sand	4
Coarse Sand	5
Sand with Gravel	6 -7

To determine lithology types within the 3-D model, TNO's Geotop model was utilized (Stafleu et al., 2011), where each of the cells within the 3-D model area was assigned with an appropriate apparent FF based on available lithostratigraphic classifications. The FF within the DOI area of the 3-D model was then applied to each inversion result using FF values from individual voxel cells. Finally, using empirical relationships between salinity and EC, it is possible to subdivide the 3-D volumes of groundwater EC into fresh-brackish-saline regions, based on the published values listed in Table 2.6 (Paine and Minty, 2005). All EC values in the following refer to groundwater EC.

Table 2.6. *Fresh, Saline, Brackish Classifications for Groundwater EC Used in This Paper, From Paine and Minty (2005).*

Conductivities (S/m) or Resistivity ($\Omega \cdot$m)	Classification
0 – 0.18 S/m, 0 – > 5.6 $\Omega \cdot$m	Fresh
0.18 – 1.8 S/m, 5.6 – 0.56 $\Omega \cdot$m (0.54 S/m, 1.85 $\Omega \cdot$m brackish-saline interface)	Slightly to Moderately Saline (Brackish)
> 1.8 S/m (< 0.56 $\Omega \cdot$m)	Very Saline to Brine (Saline)

Note. EC = electrical conductivity.

Figure 2.8 illustrates the 3-D result of the LCI smooth inversion, along with some outputs available for analysis.
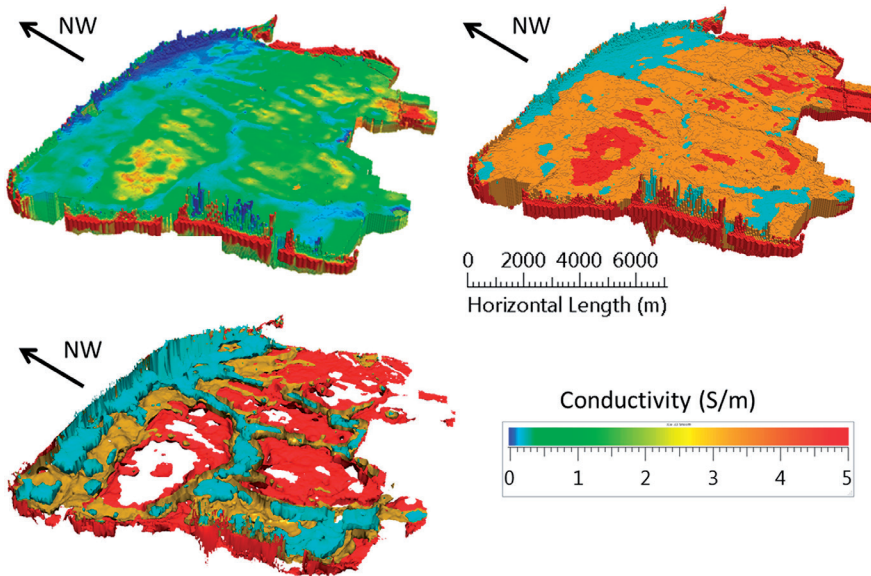
Figure 2.8: *The 3-D interpolation results for the LCI smooth inversion. (top left) EC groundwater (Siemens per meter) low conductivity (blue) and high conductivity (red). (top right) Fresh (blue, < 0.18 S/m), brackish (orange, 0.18–1.8 S/m), and saline (red, > 1.8 S/m) volumes. (bottom left) fresh-brackish (blue, 0.18 S/m), brackish-saline (orange, 0.54 S/m), and saline (red, 1.8 S/m) interfaces extracted as surfaces. Forty times vertical exaggeration. LCI = laterally constrained inversion; EC = electrical conductivity.*
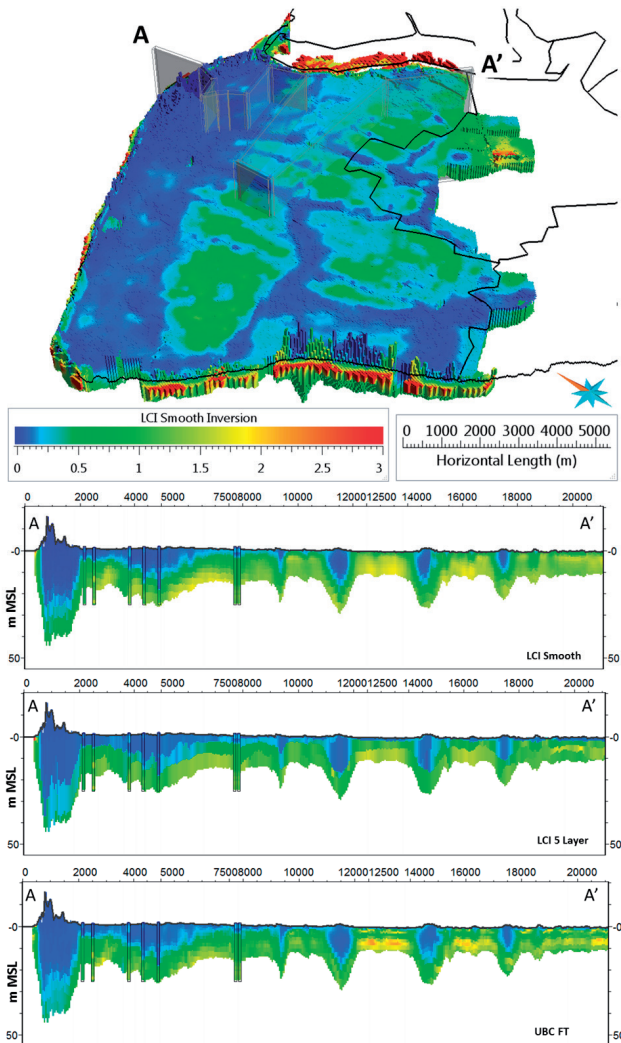
2.4 Results

With a 3-D volume of groundwater EC, as well as fresh-brackish-saline regions available for each inversion method, differences between the algorithms were analyzed in novel ways—with a practical and applied focus on some of the mapping requirements for a hydrogeophysical study. Given that the exact same process was explicitly completed for each inversion, and inversion parameter settings were chosen on a quantitative basis, the following assumes that relative differences noted between algorithms are from the inversion method itself. In regard to the use of direct ground measurements, the following should be noted: (1) these are bulk EC measurements and therefore represent the EC of both lithology and groundwater; (2) EC is sensitive to temperature (~2% increase per 1 °C), and (3) no two methods are exactly comparable—thus depending on the conditions of acquisition at the time and method used, EC values will vary. As a result, while the ground measurements are assumed to provide a more accurate (and more localized) measurement of groundwater EC, these are more quantitatively

useful as a relative comparison between inversion methods, rather than an exact depiction of a groundwater salinity distribution.

Here results will be presented as follows: (1) qualitative observations; (2) fresh, brackish, and saline groundwater volume estimates; (3) accuracy of salinity distributions; and (4) groundwater interface depth mapping.

2.4.1 Qualitative Observations

The 3-D model with representative cross sections is shown in Figure 2.9, comparing LCI smooth, LCI five-layer, and UBC Fixed Trade results to ECPT data. The cross section was designed as irregular in order to illustrate specific features such as creek ridges and to intersect ground constraints.

Figure 2.9: *A qualitative view of general inversion result observations. (top) LCI smooth 3-D model with 2-D section locations. (bottom) Results of LCI smooth, LCI five-layer, and UBC Fixed Trade inversions. Forty times vertical exaggeration. Colored by conductivity in Siemens per meter, same color legend for both 3-D models and 2-D profiles. ECPT data highlighted with black outlines. LCI = laterally constrained inversion; ECPT = Electric Cone Penetration Test; UBC = University of British Columbia.*

Qualitative comparisons suggest the following: (1) inversions results agreed with each other about the general conductivity distributions, (2) EC contrasts were consistently mapped at around the correct depths, (3) the few layer inversions tended to map a thin resistive layer at the surface, this was less apparent in smooth results, and (4) subtle differences in the amount of structure resolved were observed between the smooth 20-layer results—this is particularly noticeable between the EM1DFM and LCI inversions.

2.4.2 Volume Estimates

Fresh-brackish-saline volume calculations were completed for each inversion, and the results of which are illustrated in Figure 2.10. All groundwater volume estimates were calculated as volumes between the AHN25 terrain surface and the base of the DOI. Given that the differences between the terrain surface and the groundwater table are known to be small (Van Baaren et al., 2018), it was decided that using the terrain as an upper boundary was suitable and allowed a straightforward and unambiguous way to compare inversions. As an exact (ideal) 3-D model of the fresh-saline distribution over the study area is unavailable, this quantitative step is useful as a relative comparison only.
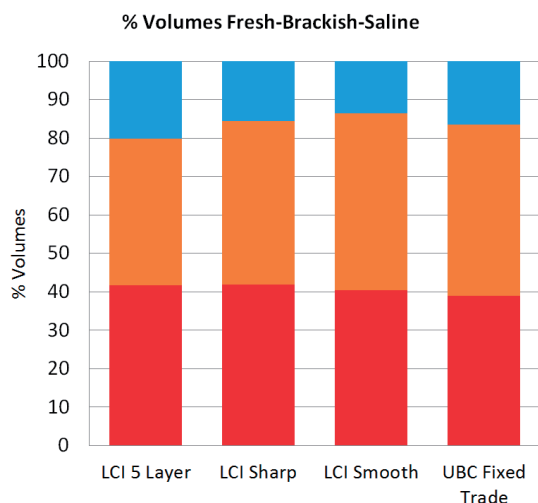

Figure 2.10: *Volume estimates for each inversion. Red (saline, ≥ 1.8 S/m), orange (brackish, 0.18–1.8 S/m), and blue (fresh, ≤ 0.18 S/m). LCI = laterally constrained inversion; UBC = University of British Columbia.*

Each inversion consistently estimated saline volumes, with a maximum difference of ~4% observed, where salinity volumes were all in the 40% range. Larger differences were noted between fresh and brackish estimates, where the LCI five-layer inversion estimated higher freshwater volumes at ~20%, a difference of ~6.5% when compared with the lowest estimate from the LCI smooth (20 layers) result. Quantitatively, out of a total model volume of 2.8 billion m3, a freshwater estimate difference of ~178 million m3 was observed in an area of 20 km × 20 km × ~50 m between the methods. The few layer LCI and LCI sharp inversions favored resolving a smaller brackish (and larger fresh) volume.

### 2.4.3 Overall Accuracy of Groundwater Distributions

By transferring the 3-D model EC values to the locations of the ECPT and borehole data, each result could be directly and quantitatively compared. This was achieved by plotting the EC

from borehole and ECPT data against the inversions and calculating correlation coefficients for each (Figure 2.11).
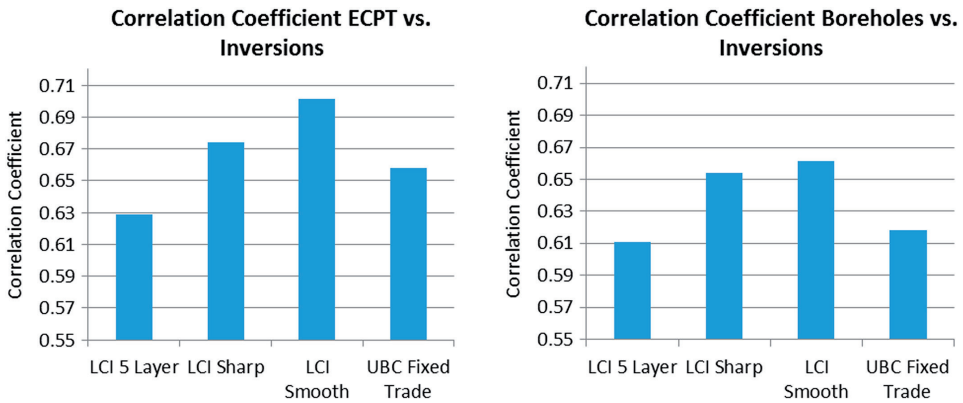
Figure 2.11: *Correlation coefficients based on XY plots of each inversion versus ECPT data. (left) Borehole correlation. (right) ECPT correlation. ECPT = Electric Cone Penetration Test; LCI = laterally constrained inversion; UBC = University of British Columbia.*

Smoother, multilayer results were consistently closer against both ECPT and borehole measurements. The LCI five-layer inversion correlated the least, compared to LCI smooth that was shown to be overall closer. These are compared in Figure 2.12.
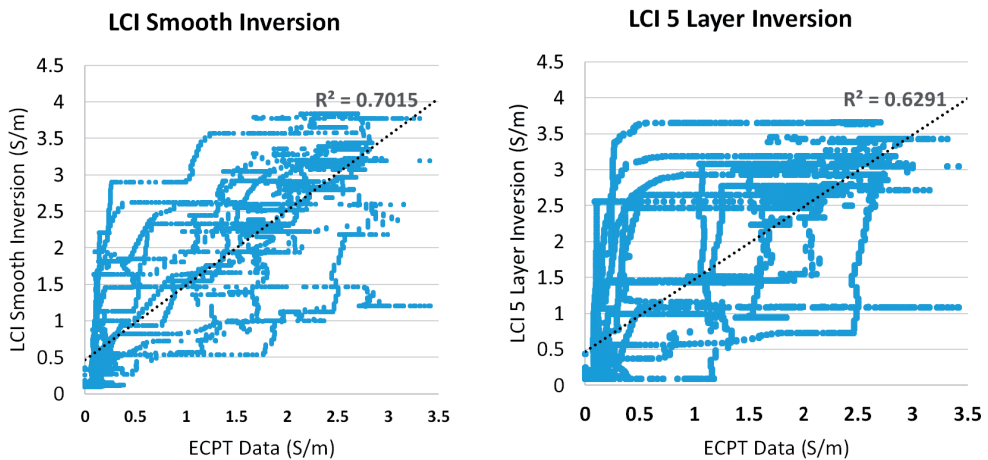

Figure 2.12: *Illustrative XY plots based on inversion results versus ECPT data. (left) LCI smooth inversion result. (right) LCI five-layer inversion result. ECPT = Electric Cone Penetration Test; LCI = laterally constrained inversion.*

The cross-plot distributions show that most of the differences are related to the layered nature of the inversions versus the smoothly varying ECPT and borehole data, as observed in Figure 2.12 by the vertical/horizontal patterns in the LCI five-layer plot.

2.4.4 Interface Depth Accuracy

Interfaces between fresh-brackish, brackish-saline, and the center of the brackish groundwater zone were extracted from each of the inversion results as isosurfaces, as well as corresponding fresh-brackish-saline interface depths from ECPT and borehole measurements using groundwater EC values. Figure 2.13 illustrates the depth to the brackish zone from the LCI smooth result, as well as the 3-D surface with extracted ECPT point values for the brackish interval. For consistency interfaces are presented as fresh-brackish (0.18 S/m), center of brackish (0.54 S/m), and brackish-saline (1.8 S/m), which are based on values listed in Table 6.
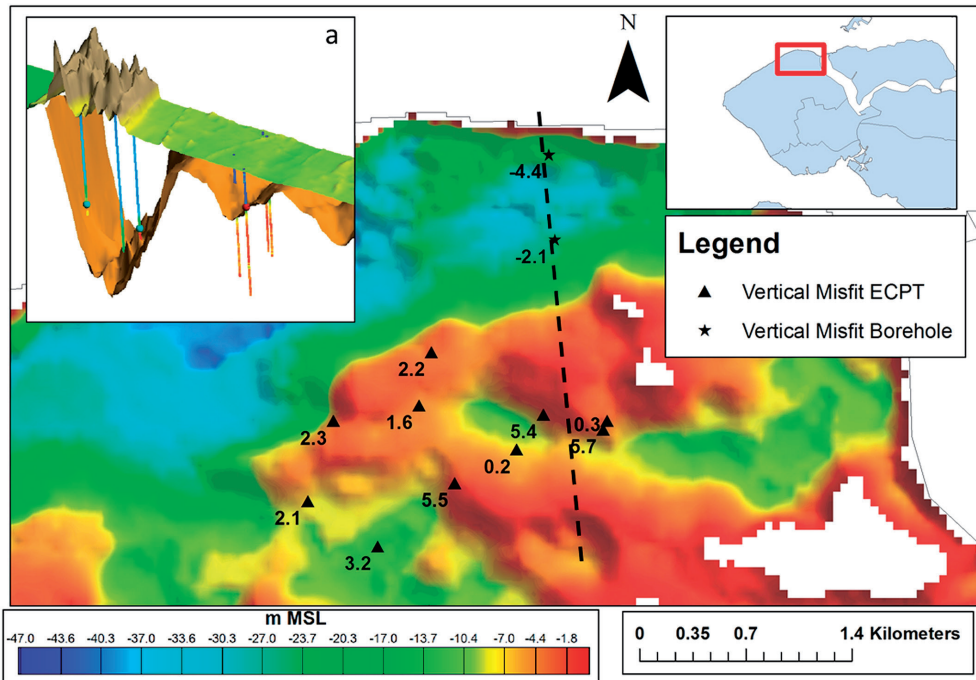

Figure 2.13: *LCI smooth interface mapping result of the brackish interface (0.54 S/m). Black stars (ECPT misfit, meters), black triangles (borehole misfit, meters); misfit is presented as ± values. Background image is colored in depth to the interface with linear color stretch. The black dotted plane is the location of the 3-D inset (a), here red and blue dots show the interface depth extracted from ECPT and borehole data, respectively. ECPT = Electric Cone Penetration Test; LCI = laterally constrained inversion; UBC = University of British Columbia.*

Distances from the ECPT and borehole point data to the corresponding 3-D surfaces were calculated as a vertical misfit as both absolute (Figure 2.14) and ± values (Figure 2.15); negative values correspond to inversions that were deeper than ground data and vice-versa. As the borehole measurements were often deeper than the ECPT, the two are plotted separately.
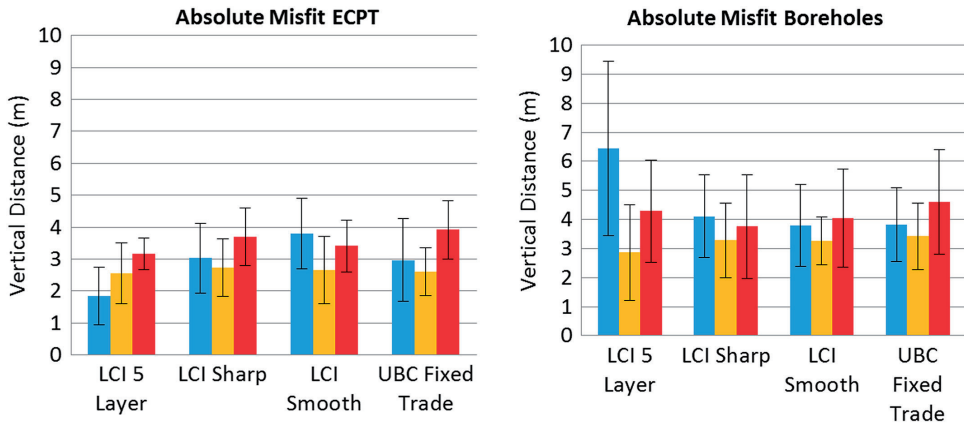
Figure 2.14: *Vertical mismatch (mean absolute differences between observed data) of inversion results to extracted interfaces in meters. Red (saline misfit, 1.8 S/m), orange (brackish misfit, 0.54 S/m), and blue (fresh misfit, 0.18 S/m). ECPT = Electric Cone Penetration Test; LCI = laterally constrained inversion; UBC = University of British Columbia.*
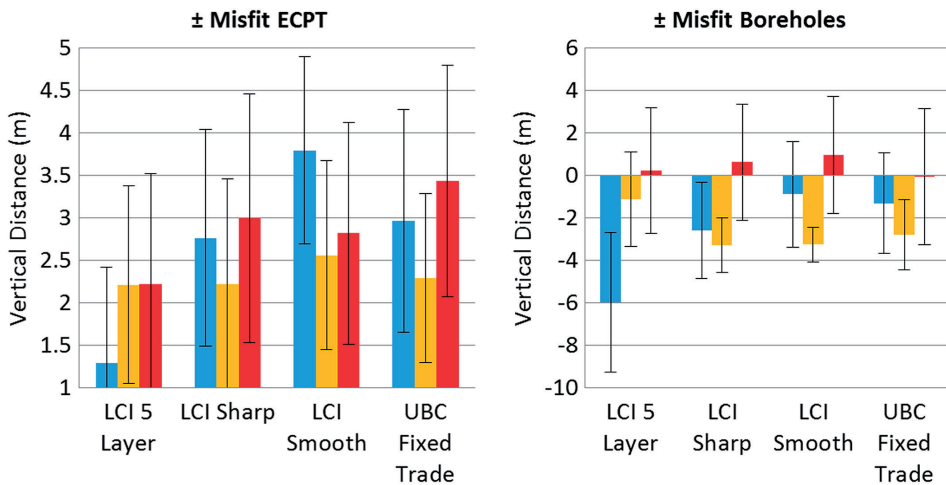
Figure 2.15: *Vertical mismatch (mean ± differences between observed data) of inversion results to extracted interfaces in meters. Red (saline misfit, 1.8 S/m), orange (brackish misfit, 0.54 S/m), and blue (fresh misfit, 0.18 S/m). ECPT = Electric Cone Penetration Test; LCI = laterally constrained inversion; UBC = University of British Columbia.*

From the absolute value misfits in Figure 2.14, it follows that all methods consistently mapped interfaces within an accuracy range of around 3–4 m against both ECPT and borehole measurements. Furthermore, both comparisons suggest that all inversions were better at resolving the center of the brackish zone (0.54 S/m). Overall, smoother, multilayer inversions were consistent with one another compared with ECPT data. A distinction between the methods is noted in the ECPT misfit, where the fresh-brackish groundwater interface was mapped with greater accuracy by the few layer LCI, multilayer LCI sharp, and UBC Fixed

Trade inversions. This is a mapping error difference of just over 2 m between the inversions. Overall, ECPT misfit suggests that fresh groundwater interface mapping in Zeeland area favors sharper techniques; however, less distinction is noted when trying to resolve the center of the brackish zone or the brackish-saline groundwater interface. In contrast to ECPT differences, the deeper borehole measurements suggest that the few layer inversion was less effective at resolving the fresh-brackish groundwater interface, however slightly more accurate in terms of mapping the center of the brackish zone overall.

In order to quantify vertical offset between interfaces and ground data, the above misfits are illustrated as ± values in Figure 2.15.

Similar patterns are observed in Figure 2.14, where ECPT data show that the brackish interface was seen to be the most consistently mapped ($\sim + 2$ m ± 2 m), although borehole data suggest that the saline interface was most accurately resolved at the correct offset ($\sim +$ 0.5 m ± 4 m). The few layer method more accurately resolved the offset of the fresh-brackish interface compared with the ECPT (~ 1–2 ± 2 m) and the center of the brackish zone with the deeper boreholes (-1 m to -2 ± 4 m). A constant offset was observed against the ECPT data, whereby inversion results were all around 2 m shallower than predicted. In the borehole data little offset is observed with the saline interface, although both brackish and fresh interface results suggest that inversions were generally deeper.

2.5 Discussion

Based on the qualitative results presented in Figure 2.9, all inversion results are roughly consistent and visually honor the ground data in terms of groundwater EC distributions and locations of major conductivity contrasts. As expected, more obvious differences are observed between few layer and smooth, multilayer methods. Other aspects are more subtle and thus will be discussed here in the same order as section 4. Observations from the variance plots (Figure 2.6) showed that shallow, highly conductive features were mapped with greater stability by the LCI algorithm. Here a feature between 4 and 5 km showed much greater variability in regard to misfit and resolved features. This is likely because the LCI algorithm was improved to handle high-frequency data found on the Resolve HEM system (Siemon, 2012). It should be noted in the following that the effect of geology in the study area was observed to be relatively insignificant, whereby an obvious and consistent long-wavelength conductivity contrast dominated the area and relates easily to groundwater salinity. This effect is of course site specific. Furthermore, as available ground data are mostly present in or around creek ridges and dunes, which are typical freshwater lens features, it should be further noted that there is a bias on techniques that resolve features well under resistive cover.

2.5.1 Volume Estimates

Fresh-brackish-saline volume estimates primarily illustrate the effect of model smoothness and mapping results. Here it is observed that in the study area, freshwater volume estimates differ by up to 6.5%, or 178 million m3, between the sharpest LCI five-layer and the smoothest LCI smooth results. This effect is primarily caused by the preservation of inversion layers from the 3-D interpolation, resulting in very sharp boundaries (spread over five layers for the few layer inversion) and therefore results in a small brackish zone. As the transition between the fresh-brackish-saline groundwater distributions is naturally smoother (van Baaren et al., 2018), results suggest that for volume estimates it is preferable to use smooth, multilayer inversions. Furthermore, brackish zone estimates between the LCI sharp and LCI smooth results differed by around 3% or a volume of 82 million m3. It is therefore shown that by using a multilayer ($\geq$ 20 layers) inversion and carefully selecting the regularization term based on prior knowledge, this smoothing effect can be controlled according to known conditions. This is therefore the recommended approach for volume mapping.

2.5.2 Accuracy of Salinity Distributions

By plotting inversion results with ECPT and borehole data (Figures 2.11 and 2.12), results suggest that LCI smooth and LCI sharp multilayer methods were the most successful at honoring ground data. Predictably, it is shown that sharp interfaces from the few layer inversions, due to having few layers, result in larger misfits. As the ECPT and borehole data are measured every 0.5 and 5 cm, respectively, and the inversions are layered, it is expected that a good match to salinity distribution mapping favors smoother results. Results therefore do not mean that few layer methods are worse—rather for the purposes of mapping smoothly varying groundwater salinities, they were less successful. Results would likely be different if ECPT and borehole data were averaged for each layer of the inversion; however, the smooth distribution of salinity is a common mapping objective—thus, it was decided that a direct comparison was appropriate. Furthermore, given that the province of Zeeland was flooded with saline water relatively recently, and subsequently recharged with fresh groundwater (Berendsen, 2005), the brackish zone is expected to be sharper (or thinner) than many typical hydrological situations—further highlighting the fact that smoother techniques are more suitable in general.

2.5.3 Interface Mapping

Absolute value results (Figure 2.14) suggest that interface mapping was remarkably consistent across all inversions, with an average accuracy of around 3 m ($\pm$ 3 m). This could be the result of bias, whereby the regularization term was selected based on the same ECPT data, or additionally it could be an effect of the interpolation method or airborne acquisition system itself. However, given that this consistency is also observed across both ECPT and borehole data, it is unlikely that this error can be attributed to the ground constraints. Overall, the center of the brackish zone interface of 0.54 S/m was resolved better. This is more obvious in the ECPT

data, where misfits were observed at between 2.5 and 3 m, compared to the brackish-saline (1.8 S/m) groundwater interface, where these were between 3 and 4.5 m. In terms of resolving the fresh-brackish (0.18 S/m) interface, the ECPT data suggest that few layer and LCI sharp inversions were more successful at a misfit of ~3 m, in contrast to ~4 m from the LCI smooth method. However, borehole data suggest that smoother results from the multilayer inversions worked better. The most likely explanation for this contrasting result may be the fact that the borehole data are located in the dune areas and are deeper. Here the deeper DOI in the dune areas has resulted in thickening of the layers where there are small conductivity contrasts, therefore picking specific interfaces at depth without significant conductivity contrasts becomes more arbitrary for few layer methods. Interestingly, the few layer results are more accurate at resolving the brackish interface in these deeper areas, at ~3 m for the LCI five-layer result. This would suggest that because the few layer inversion changed layer thickness and conductivity in the inversion process, in contrast to the (smoother at depth) 20-layer models, high conductivity contrasts are more likely to be mapped successfully in deeper areas by using few layer methods. As the boreholes are up to 50-m deep, and the multilayer inversion would have a layer thickness of around 10 m at these depths, an error of around ~5 m is expected. It is therefore suggested that unless a strong conductivity contrast is present, trying to map subtle features at greater depth using few layer methods is likely to be less successful. Given that geophysical methods favor strong physical property contrasts, and the dominant contrast in the study area is known to be caused by water salinity (Delsman et al., 2018), this result is therefore expected. In previous hydrogeophysical HEM studies in areas where minimal subsurface information was available, fresh-saline groundwater interfaces were commonly estimated using a conductivity value based on empirical relationships that relate conductivity to salinity (Siemon et al., 2009b). Our study suggests that selecting the strongest conductivity contrast is preferable over this method and should result in more accurate estimates, albeit with limited flexibility in regard to values (and therefore only a single interface) used. This was undertaken in De Louw et al. (2011) where the sharpest vertical conductivity ratio was used to indicate a mixing zone or the brackish groundwater interface as referred to in this study. Furthermore, using the same HEM data, Siemon et al. (2018) found that using this method resulted in a brackish zone mapping error of +0.1 m (± 1.7 m) compared to available ECPT data. Finally, the presence of a strong conductivity contrast is observed in the bimodal distribution of the inversion results, where EC values were grouped mostly between 0–0.3 and > 0.7 S/m.

By using ± values (Figure 2.15), similar patterns were observed—however, most notably ECPT data suggested that inversion results were consistently too shallow at +2 m (±2 m). In contrast, borehole data showed that all inversions were either around 0 m (±2.5 m) for the saline interface or too deep at −1 to −6 m (± 3 m). This effect results from a number of factors, for example, (1) inherent differences between ground measurement methods, where HEM and direct current methods are more sensitive to horizontal and vertical resistivities, respectively (Siemon et al.,

2009b), (2) localized effects on EC measurements such as temperature and lithology, or (3) highly localized features mapped by ground measurements could be smoothed by the HEM system or 3-D interpolation—highlighting resolution differences between the methods. Consistent vertical offsets observed in all inversions highlight that while the ground data were a useful quantitative comparison for this study, further research is required to understand this occurrence.
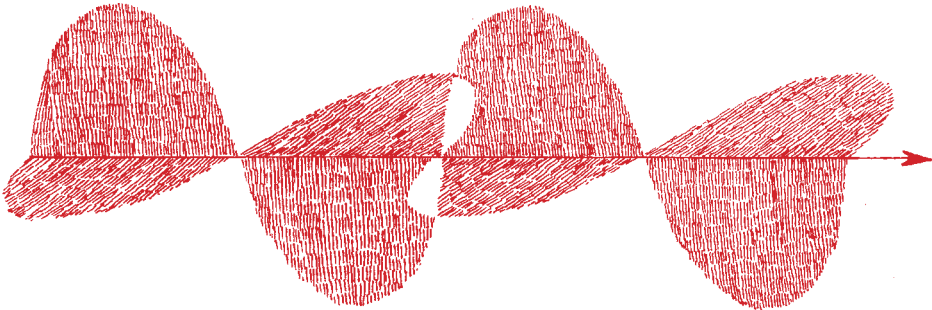
2.6 Conclusions

Using an extensive amount of available data, the results of this study quantitatively highlight the effects of using different inversion techniques and regularization terms, based on the objectives of a typical hydrogeophysical study. Overall, the four inversion methods tested are generally consistent with one another and are shown to be effective at resolving a number of hydrological features. Differences between multilayer and few layer inversions are observed to be the most prominent, and their suitability depends on specific mapping objectives. To simply map a brackish (0.54 S/m) groundwater interface in shallow or deeper areas with minimal ground data, a few layer inversion type is suggested. Here the area with the sharpest conductivity contrast is likely to be an excellent proxy to this boundary, and no information about FFs is required. If both fresh-brackish (0.18 S/m) and brackish interfaces (0.54 S/m) are required, then a smooth, multilayer inversion is suggested, although it should be noted that due to thick layering at depth, the brackish zone will be less effectively mapped here than with the few layer approach. If the objective is to map a smoothly varying volume of conductivity (therefore salinity distributions), a multilayer inversion type is favored. Further, the choice of regularization terms and the type of multilayer inversion used are shown to affect the estimated brackish zone thickness. If ground data are available, then the smooth inversion regularization term could be chosen accordingly to better match the thickness of the brackish zone. Overall, inversion methods should be selected carefully according to mapping objectives—preferably based on a combination of experience and prior knowledge of the subsurface.

2.7 Acknowledgments

3

A practical quantification of error sources in regional-scale airborne groundwater salinity mapping.

Abstract

Hosting over 10% of the world's population, low elevation (<10m above mean sea-level) coastal zones are susceptible to saline groundwater intrusion - making fresh groundwater an often scarce and threatened commodity. To inform suitable coastal groundwater management strategies, regional-scale mapping of fresh and salt groundwater occurrence is extremely beneficial. This mapping is usually based on conventional ground-based methods. However, these are not only slow and expensive, but result in localized and disconnected information which is uneconomical and impractical on the large scales required. Airborne electromagnetic (AEM) surveys have been proven a frugal and rapid way to overcome these shortcomings. Consequently, AEM methods are increasingly being used globally. Little is known about the effects of flightline spacing and additional ground-based data on the quality of mapping results, and in general the accuracy of AEM, other than validation against often sparse ground measurements. Understanding this is therefore invaluable as input to groundwater management strategies, survey planning and decision making. Here, we use a regional scale (900km2), high-resolution (50m x 50m x 0.5m) 3D synthetic model of electrical conductivity and geological properties, to investigate the effects of data availability on the accuracy of regional-scale groundwater salinity mapping. This was undertaken by simulating commonly used AEM parameters and realistic data acquisition methods. Two key data components are considered: (1) the AEM survey itself, and (2) geological information used to convert the AEM results into groundwater salinity. Spatially, different data-densities of these two components are quantitatively compared to highlight ideal geometrical configurations for given accuracy requirements. Our results indicate that in terms of optimising costs versus benefits, the value of additional lithological information is dependent on how well the initial distribution of electrical conductivity is resolved by the acquisition and inversion process.

3.1 Introduction

Despite comprising only ~2% of global land area, low elevation (≤10m above sea-level) coastal zones host ~10% of the world's inhabitants, the majority of which are in developing countries (Neumann et al., 2015). Coastal aquifers are consequently a high-demand and over-stressed water source (Post, 2005). Unfortunately, these are susceptible to groundwater salinization (P. G B De Louw et al., 2011), a problem worsened by excessive groundwater extraction and land-subsidence (Minderhoud et al., 2017), which could be further aggravated by sea-level rise (Oude Essink et al., 2010). Properly informed and vigilant groundwater management strategies are therefore crucial; thus accurate mapping of the regional-scale occurrence of fresh and saline groundwater is needed.

This mapping is traditionally based on conventional ground-based methods, which are not only slow and expensive, but result in localized and disconnected information which is uneconomical and impractical on the large scales required. Airborne electromagnetic (AEM) surveys have been proven a frugal and rapid way to overcome these shortcomings since the 1970s (Sengpiel, 1981) and have since been successfully applied to map the distribution of fresh, brackish or saline resources regional scales (Auken et al., 2008; Delsman et al, 2018; Faneca Sànchez et al., 2012; Fitterman and Deszcz-Pan, 2001; Jørgensen et al., 2012; Meyer et al., 2019; Siemon et al., 2015; Sulzbacher et al., 2012). Primarily, two methods exist: time and frequency domain electromagnetics, referred to here as TEM and FEM respectively, and are discussed in detail and compared in Steuer et al., (2009). Both methods operate by transmitting and receiving electrical signals from an instrument towed beneath a helicopter. The signals produce secondary currents in the subsurface which are in turn measured by receivers. The suitability of the method for coastal groundwater salinity mapping relates to electrical conductivity (EC) contrasts in the subsurface in the form of pore water salinity and clay content of host sediments (McNeil, 1980). Surveys are typically flown in a grid pattern at specified flightline spacing, resulting in observations every few meters downline.

Observations are normally represented as physical properties using inversions (Auken et al., 2005; Brodie and Sambridge, 2006; Farquharson et al., 2003; Hansen and Minsley, 2019; Siemon et al., 2009a; Vignoli et al., 2015), where these are inverted to produce a subsurface distribution of bulk electrical conductivity (ECb) – or rather the combined EC from two sources: lithology and pore water. Although ECb distributions are useful and can be used as a proxy to groundwater salinity distributions – or even used to automatically estimate a fresh-saline groundwater interface (Siemon et al., 2018) – practical groundwater management requires quantitative salinity information. For this, pore water conductivity is needed (ECw), which can subsequently be transformed into chloride concentrations using empirical relationships (e.g., De Louw et al., 2011). To acquire ECw from ECb from inversion results, formation factors (FF) are used (Archie, 1942b) which requires geological information.

This is a serious source of uncertainty, as field measurements have shown that fresh groundwater in clay could have the same ECb values as brackish groundwater in sand (P. G B De Louw et al., 2011). Thus, along with undertaking an AEM survey, potentially expensive and intrusive geological information is required for accurate transformation into groundwater salinity. Unlike many AEM mineral exploration programmes which aim to discern discrete or anomalous features of interest, hydrogeophysical mapping programmes tend to treat the entirety of the survey area with equal value, and thus a large scale and potentially expensive geological model is required.

Relatively few AEM studies have attempted to quantify regional groundwater salinity in 3D using FFs (e.g. Faneca Sànchez et al., 2012, Delsman et al., 2018, Vandevelde et al., 2018). However, the recent success of hydrogeophysical groundwater investigations have resulted in increased attention to the method (Binley et al., 2015). Current large-scale examples include plans for a potential a country-wide AEM survey across the Netherlands, and at the time of writing an unprecedently large AEM survey is currently being undertaken in Australia for mineral and groundwater mapping purposes (Ley-cooper et al., 2018). In 2016 an atlas of subsurface EC was compiled throughout Denmark using primarily time-domain electromagnetics and ground data (Barfod et al., 2016). Furthermore, recent methodological advances have allowed high-resolution provincial-scale 3D voxel models of groundwater salinity using probabilistic lithological and AEM derived ECb information (Delsman et al, 2018). Consequently, it has become pertinent to better understand the groundwater salinity mapping error of AEM in general, as well as the effect of data density on mapping results. Although AEM has been tested and proven to be accurate and repeatable for both TEM (Foged et al., 2013) and FEM systems (Huang and Cogbill, 2006), currently little is known about the effects of flightline spacing and additional ground based data on the quality of 3D groundwater salinity mapping results. In general, model results are typically validated using 1D ground measurements (e.g. Delsman et al, 2018; Faneca Sànchez et al., 2012), but relative to an entire 3D regional model this only encompasses a very small region of the model space. Furthermore, physical differences in measurement techniques and acquisition conditions (e.g. temperature) mean that ground data are often difficult to directly compare with AEM. A quantitative and practical overview of the relative sources of error is therefore essential to inform the typically intra-disciplinary teams involved in AEM survey planning, decision making and ultimately groundwater management.

Consequently, in this paper we investigate relative sources of error based on established salinity mapping techniques using AEM with 3D models. Two key data types and their spatial configurations are considered: (1) the AEM survey itself and allocated flightline spacing, and (2) lithological information in the form of FF values through intrusive, ground-based drilling. This is undertaken using a large-scale (900km^2), high-resolution (50 x 50 x 0.5m) realistic synthetic model extending to 50m depth and comprising ~150 million model cells

and 90 separate model results. To the authors' knowledge, a large-scale, fully 3D quantitative assessment of this nature is unique. We present the accuracy of results relating to three common mapping objectives: (1) error of the 3D model as a whole, (2) volume estimates and (3) fresh-saline interface depths. Results are presented in a practical manner as data densities per km*2* and flightline spacing in meters apart and include current survey cost estimates to include economic considerations.

3.2 Methods

3.2.1. General approach
Our approach is summarized in six steps, and explained in order in sections 2.2 to 2.7: (1) *synthetic data generation* by transforming two existing 3D models (lithological and chloride) into an ECb volume; (2) *simulate AEM acquisition* using the ECb model (forward modelling) and subsequent inversion; (3) *simulate lithological data collectio*n from the 3D lithological model; (4) *3D interpolation of inversion results and lithological data*; and (5) *Conversion to ECw and chloride* by transforming interpolated results into ECw, and finally into 3D chloride models; (6) *Comparison with reference chloride model and error analysis*: finally, the estimated chloride estimates are compared against the reference model. Steps 2-6 are performed for different flightline spacings and different densities of ground-based information (geological drillings), where results are presented in section 3. The process is illustrated in figure B.1. Hereafter, these steps are explained in more detail.

3.2.2. Step 1: Synthetic data generation
The synthetic data is based on two existing high-resolution 3D datasets from southwestern Netherlands (figure 3.1). The latitudinally oriented Zeeuws-Vlaanderen strip (Zeelandic Flanders in English) covers an area of ~60 x 15km and lies within the southern edge of the North Sea basin. The shallow subsurface (up to ~50m) consists of gently northward dipping Neogene and Quaternary sediments (Stafleu et al., 2011). Overall, lithologies comprise younger fine sands, clays and peats and deeper, gently dipping sands and silts (Vos, 2015). Along the North Sea coast sandy coastal barriers are marked by localized topographic highs. Holocene sea-level transgressions caused extensive groundwater salinization (Delsman et al., 2014) however subsequent land reclamation has allowed the formation of shallow freshwater lenses (Berendsen, 2005). Given these events are recent in geological terms, distinct zones of fresh and saline groundwater with a relatively thin brackish zone are present throughout (Delsman et al, 2018).

The chloride model used here (figure 3.1B) is the result of a provincial scale FEM survey that was flown in 2014/2015 in the Netherlands as a part of a research project called FRESHEM Zeeland (Delsman et al, 2018). The resulting 3D groundwater salinity model is publicly

available and has a resolution of 50m horizontal and 0.5m vertical, covering an area of ~2000 km^2. AEM surveys decrease in resolution at depth, therefore it should be noted that the 0.5m resolution throughout the model is the result of a secondary interpolation step. In the study area the survey was flown at mostly 300m line spacing, in areas of interest spacing was decreased to 100m resulting in ~3000 line km of HEM data. Flightlines are shown in figure 3.1A. Flightline spacing is selected according to a compromise between target sizes and budget. Tie-lines are flown perpendicularly to these for data processing at a wider spacing.

A high-resolution 3D lithological voxel model (100 x 100 x 0.5m resolution, up to 50m depth from ground surface) of the area is available from the GeoTOP geological model as part of the Dutch national database of geological data (dinoloket.nl - accessed November 2016), and provided by the Geological Survey of the Netherlands (GSN-TNO). As well as a stratigraphic model, GSN-TNO's GeoTOP model contains the probability of lithoclasses occurring at any voxel and is therefore a probabilistic heterogeneous style model (lithoclasses and formation factors outlined in table 3.1 and shown in figure 3.1C). A more detailed description of the model is available in (J. Stafleu, D. Maljers, J.L. Gunnink, 2011). Apparent FF values were assigned to these lithologies and are based on nearby field measurements (De Louw et al., 2011) (table 3.1). We recognize that the EC of saturated sediment is complex, such as surface conductivity caused by clay particles (Revil and Glover, 1998), however for this work, we neglect the effect of the clay to surface conductivity since the exact values are heavily depended on the area of investigation (Revil et al., 2017). Previous research examines this effect (Delsman et al, 2018), but for this regional analysis we did not consider it an important factor.

Table 3.1. *Lithological formation factor values used in this study, from De Louw et al., (2011).*

Lithology	Formation Factor (FF)
Peat	2.1
Clay	2.5
Sandy clay	2.8
Fine sand	3.2
Medium sand	4
Coarse sand	5

To forward model (and subsequently invert) an AEM survey, a synthetic ECb model (i.e. the combined effect of both groundwater and lithology) is required. Using a combination of the two existing 3D models, a realistic and heterogeneous-style ECb model was created, the result of which is shown in figure 3.1D. This was undertaken in two steps: (1) the 3D chloride model was transformed to ECw using an empirical linear relationship from field measurements (De Louw et al., 2011) and corrected back to a reference groundwater temperature of 11C° from an initial 25C°; (2) ECw was transformed into ECb by applying FF values based on the geological model. The resulting regional-scale 3D ECb synthetic model has a cell resolution

of 50m horizontal and 0.5m vertical, covering an extent of ~15 x 60km down to 50m depth from ground surface.
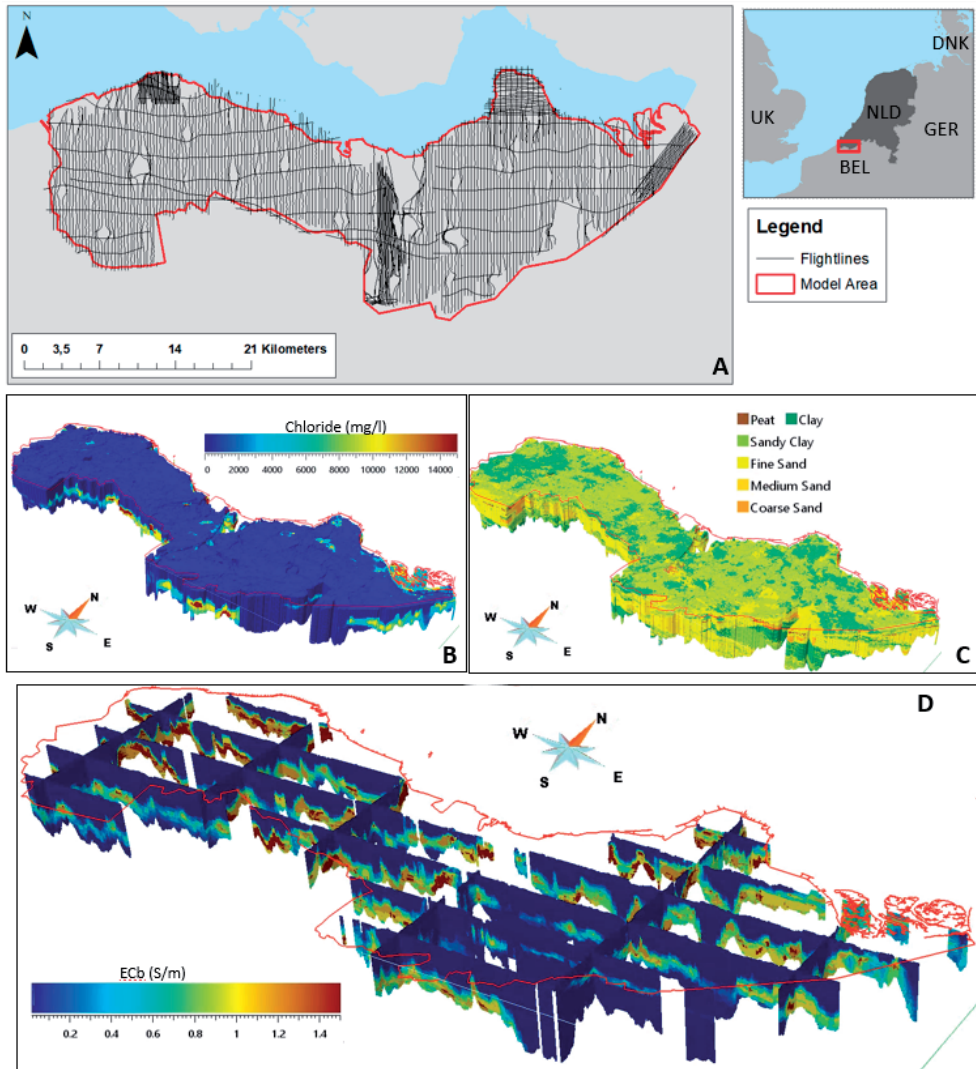
Figure 3.1. *Data used to create the 3D synthetic model. 3.1A: flightlines and model area. 3.1B: 3D chloride model. 3.1C: lithological model. 1D: the resulting 3D ECb model used for forward modelling.*

3.2.3. Step 2: Simulate AEM acquisition

Given that the available 3D chloride reference model was completed using a FEM survey, it made sense to simulate this rather than the TEM method. To simulate a typical survey across the 3D model, a distribution of measurement locations was needed. These were taken from the same survey that produced the 3D salinity model (Delsman et al, 2018) in order to replicate

typical heliborne survey characteristics (figure 3.1A). In total ~800,000 measurements locations over 3000 line km were used from 361 flightlines, these were typically flown at between 100 – 300m flightline spacing depending on the area. To facilitate processing times, the measurements were decimated to one every ~40m from every ~4m (totalling ~100,000 measurements) – this was found to be a good balance between respecting the FEM system's ~100m footprint (Yin et al., 2014) and hardware practicalities. Given that the signal measured by the AEM system is inherently 3D, and the forward modelling code used in this is paper is 1D (Auken et al., 2005) – a pseudo-3D sampling technique was implemented. To achieve this, 100m diameter circles were created around each measurement location to approximate the footprint of the FEM system (Yin et al., 2014). Within each circle ECb was averaged, this was repeated at 0.5m depth-slice intervals down to 50m depth (i.e. at vertical model resolution), resulting in 100,000 synthetically generated AEM acquisition points on flight lines, each with 100 ECb values with depth (every 0.5 m until 50 m). Figure 3.2 illustrates the pseudo-3D sampling technique used, as well as an illustration of the resulting AEM coverage based on different flightline spacing.
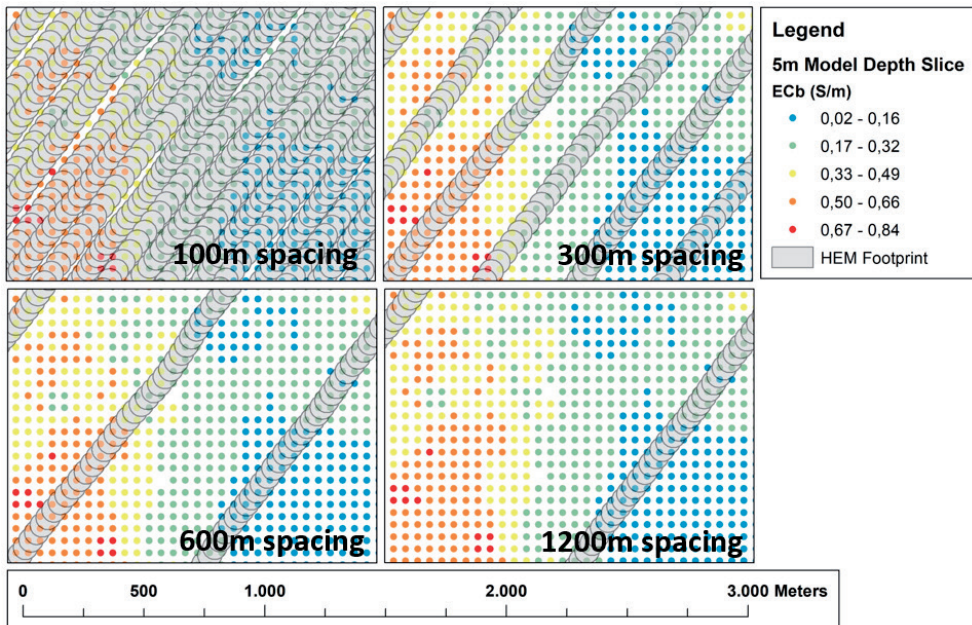
Figure 3.2. *The pseudo-3D sampling technique used to simulate the HEM system footprint. The effects of using different flightline spacing on coverage are illustrated.*

The resulting data were forward modelled using AarhusINV (Auken et al., 2005). Source and receiver spacing and frequencies were selected based on the RESOLVE HEM system; orientation was selected based on typical and previously used hydrogeophysical applications

e.g. (Delsman et al, 2018; Gunnink et al., 2012; Steuer et al., 2009). Finally, 5% random gaussian noise was added to the forward response to approximate HEM noise levels (Farquharson et al., 2003; Green and Lane, 2013).

Table 3.2. *Simulated AEM system, survey and inversion parameters.*

| Towed bird frequency domain system specifications	
Frequencies	380, 1770, 8300, 41000 and 129500 Hz
Coil separation	~8m
Coil orientations	5 x horizontal coplanar
Measurement spacing	~40m downline
Footprint	100m
Survey Parameters	
Bird height	40m
Noise	5%
Inversion Parameters	
Vertical Constraints	1.3
Lateral Constraints	3

The "acquired" data, were inverted using AarhusINV (Auken et al., 2005). Based on a commonly used inversion method (e.g. Auken et al., 2008; Chongo et al., 2015; Delsman et al, 2018) and corresponding input parameters that were found suitable for hydrogeophysical studies (King et al., 2018) (Table 3.1), a laterally constrained inversion (LCI) was used (Auken and Christiansen, 2004; Siemon et al., 2009a). The LCI method is a pseudo-2D method, whereby neighbouring model properties are constrained to produce laterally coherent inversion results along separate flightlines. The minimum-structure inversion style used here inverts for ECb values only, where layer thicknesses and depths remain constant, and has been found to accurately reproduce smoothly varying salinity distributions (Delsman et al, 2018; King et al., 2018). The inversion starting model consisted of 20 layers, the thickness of the top layer was set to 0.68m and subsequent layers increasing in thickness logarithmically until 50m. Below 50m the final layer is assumed to extend infinitely. A starting model conductivity of 1 S/m was selected based on common knowledge that low-lying coastal zones are (electrically) conductive environments.

3.2.4. Step 3: Simulate lithological data collection

Simulating the acquisition of geological information, or indeed realistic quantities of existing data that may be available during hydrogeophysical mapping is not trivial. For example the GeoTOP model used to construct the synthetic geological model in this paper is the result of expertly interpreted boreholes and geophysical data, resulting in a large-scale 3D probabilistic model with different properties (Gunnink et al., 2012; Stafleu et al., 2011). For this reason, a practical approach based on the current distribution of publicly available borehole data was used. Here we sampled synthetic borehole locations and depths from existing borehole data in the area available on Dinoloket, the Dutch Geological Survey (TNO) data portal. It was found

that within the synthetic model area, cumulative vertical borehole depths totalled ~42,000m (or ~60 vertical metres/km^2), with a depth distribution highlighted in figure 3.3.
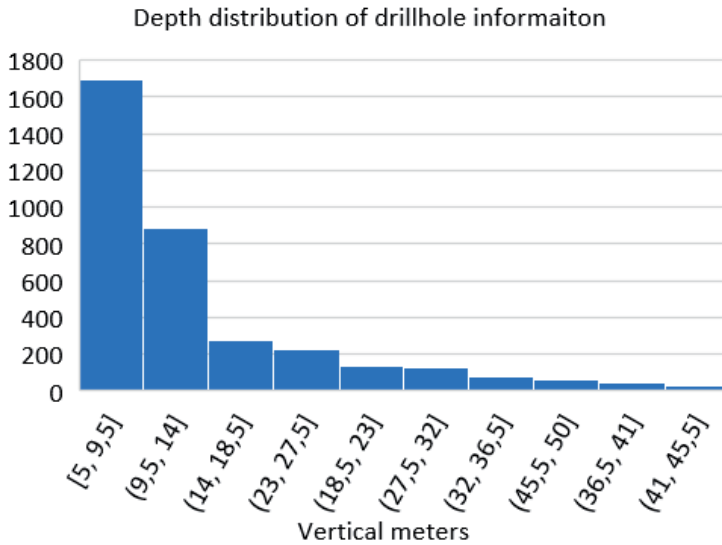
Figure 3.3. *End-depth distribution of actual available lithological information on TNO's Dinoloket data portal.*

With this information, we used inverse transform sampling (ITS) with prescribed total vertical borehole depths and randomised borehole locations (random selection without replacement) to simulate a realistic data-distribution, where typically most available data is very shallow (<5m). Prescribed total drilling length was varied in following six classes as follows: 100m (0.13m/km2), 1000m (1.36m/km^2), 5000m (6.80m/km^2), 10000m (13.61m/km^2), 20000m (27.23m/km^2) and 42000m (57.19m/km^2) – based on practical criteria up to the actual amount available. Five ITS realisations were used for each total length class to account for sampling bias and to quantify uncertainty due to placement location. Finally, FF values were taken from the lithology of the synthetic voxel model sampled by each randomly selected borehole at 0.5m vertical resolution till borehole depth. This resulted in 6x5=30 synthetic borehole datasets containing FF information. Here it is assumed that apparent FF information has been perfectly interpreted for each borehole.

3.2.5. Step 4: 3D interpolations of ECb inversion results and lithological data
In order to assess the effect of different flightline spacing on mapping results, the ECb inversions were subsequently split into three separate databases: (1) 300m spacing, 100m in specific areas (~3000 line km), (2) 600m flightline spacing (~1500 line km), and (3) 1200m spacing (~750 line km). The 100m line-spaced data was kept together with the 300m, as this

reflects a realistic data distribution and is the actual distribution of the original survey. This allows an analysis of the effects of a higher AEM data density in specific areas. Finally, each database was interpolated into 3D volumes of ECb using the method of King et al., (2018) and is based on 2D Kriging. We recognise that there are many other sophisticated methods to interpolate data, however this approach was found to accurately preserve the characteristics of inversion results between flightlines and offered a straightforward and recognised approach. The model base was assigned according to the DOI of the airborne system, which was calculated from the inversion output at each 1D model location (Christiansen & Auken, 2012). Minimum-curvature gridding was then used to interpolate DOI estimates, data below these were removed from further analysis for all further models.

A statistically robust, simple and repeatable 3D interpolation method was utilised to interpolate the FF values at the borehole data throughout the synthetic model. For this an established technique called Sequential Indicator Simulation (SIS) was used (Gomez-Hernandez and Strivastave, 1990), based on Indicator Kriging formulation. Using indicator semivariogram models fitted on the FF values of the boreholes (a separate semivariogram model with different horizontal and vertical ranges for each of the 30 borehole datasets), 10 SIS realisations were run for all 30 FF spatial distributions. The median (or p50) of the 10 simulated FF values at each model cell was used for further analysis. This resulted in 30 3D FF spatial distributions (for each of the six classes of specified total drilling length with five randomised locations for each).

3.2.6. Step 5: Conversion to ECw and chloride

Each of the 30 interpolated FF spatial distributions were multiplied by the three interpolated ECb inversion results to obtain 90 3D ECw spatial distributions (6x3=18 combinations of total borehole and flightline lengths with 5 random borehole sets per combination), which were then converted back to 3D chloride (as per figure B.1).

3.2.7. Step 6: Comparison with reference chloride model and error analysis

With practical hydrogeophysical mapping considerations in mind, the resulting 90 3D ECw spatial distribution models were compared against the reference model in three ways. All evaluations were undertaken over the same domain, where values beneath the calculated DOI from inversion results were removed. First, mean absolute error (MAE) of groundwater salinity estimates as mg/l chloride were calculated as per the following:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - x|$$

where xi is the prediction and x is the true value. Second, groundwater salinity volumes were calculated according to the chloride classifications set out by Stuyfzand (1986) and include the

following: (1) 0 -1000 mg/l, or "fresh", (2) 1000 – 3000 mg/l, or "brackish", and (3) >3000 mg/l, or "saline". Finally, based on this classification, interface depths for the fresh-brackish-saline regions were extracted as 500, 1500 and 3000mg/l chloride respectively. Saline inversion was present in some areas (saline above fresh groundwater), therefore multiple interfaces were frequently encountered at the same horizontal location. As a result, only the shallowest, or first encountered interfaces were extracted as iso-surfaces for further analysis.

3.3. Results

For clarity, *geology* refers to additional borehole information as vertical metres per km*2*; *inversions* refer to error as a result of the AEM acquisition and inversion process; and the *reference model* is the synthetic chloride model. In total all 90 3D chloride model properties of 150 million cells each (thus in total ~13.5 billion data points) are compared in the following, relating to different geometrical data configurations of geological and AEM data. Volumes and interface error are calculated according to specified salinity classes, while MAE calculations focus on the entirety of the 3D models for an overview of error contributions. Uncertainty relating to borehole placement and 3D interpolation is shown as standard error.

3.3.1 MAE of salinity estimates
The mean absolute error (MAE) of mapped chloride is presented in figure 3.4, showing average error for all data configurations tested. To better isolate error per data-type, in figure B.2. we assumed a perfect inversion model and only adjusted the amount of lithological information.
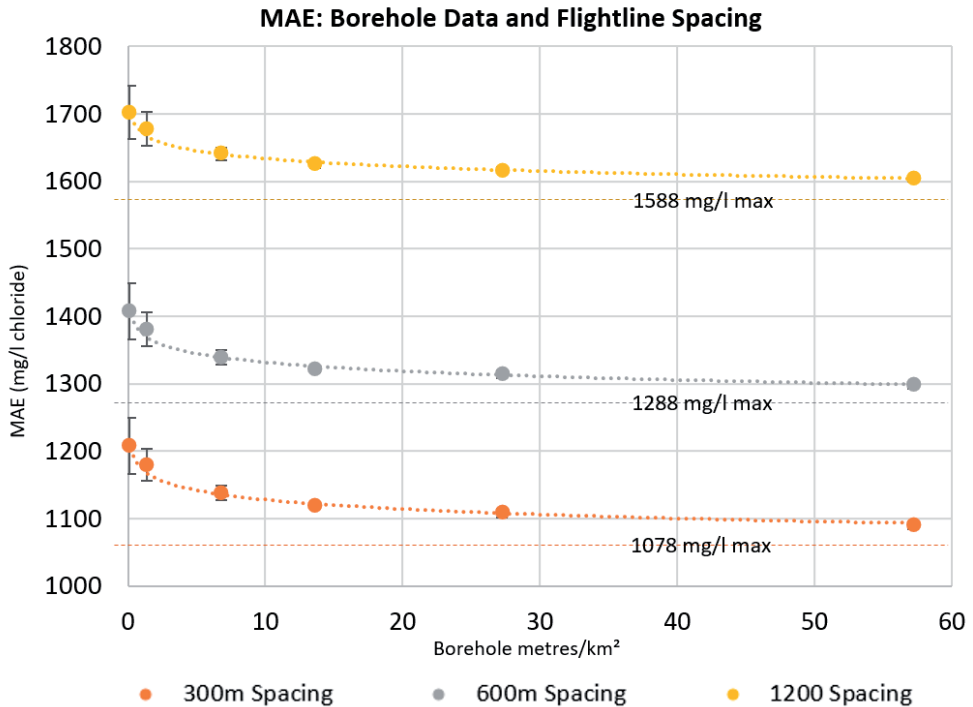
Figure 3.4. *MAE of flightline spacing as meters apart vs. data density of borehole data as vertical meters drilled per km2. Yellow = 1200m spacing, grey = 600m spacing, orange = 300m spacing; the vertical black error bars represent uncertainty range over the five sampled borehole configurations per borehole density class; the horizontal dashed lines are error due to AEM inversion only assuming perfect geological knowledge.*

Figure 3.4 illustrates that increasing borehole density to a value of more than ~10m/km2 results in relatively little improvement of mapped chloride. Geological data reduces mapping error by ~100 mg/l chloride between the highest and lowest data densities for all flightline configurations and at maximum 150 mg/l chloride in case of perfect geological knowledge. Reducing flightline spacing consistently reduces error by ~70 mg/l chloride per 100m reduction. The reduction in salinity estimation error by adding geological information is significant if the AEM inversion is without error (figure B.2., the combined error is dominated by the error in AEM inversion. Thus, the most effective way of reducing the error in chloride concentration mapping is by decreasing flightline spacing.

3.3.2 Volumes

Calculations were undertaken based on the classifications outlined in section 2.6, where a porosity factor of 0.3 was applied based on a sandy aquifer (Abdallatif et al., 2009). The result is a total groundwater volume of ~$5.71 \cdot 10^6$km3, with $3.13 \cdot 10^6$km3 (54%) fresh, $2.2 \cdot 10^6$km3 (39%) salt and $3.82 \cdot 10^5$km3 (6.7%) brackish groundwater volumes within the reference model. The AEM data acquisition and inversion process alone (assuming perfect geological

knowledge) results in a substantial thickening of the brackish zone, adding ~$6.1 \cdot 10^5$ km^3 to brackish zone volume estimates (figure B.4), underestimating the fresh and saline volume by ~$3.27 \cdot 10^5$ km^3 and ~$2.83 \cdot 10^5$ km^3 respectively. Figure 3.5 presents averaged volume differences against the reference model for each geological data density class with standard error (over the six borehole datasets per total borehole length class) and for a flightline distance of 300m. Total error for each estimate has been subtracted by those of the reference model – thus positive values represent an overestimation.
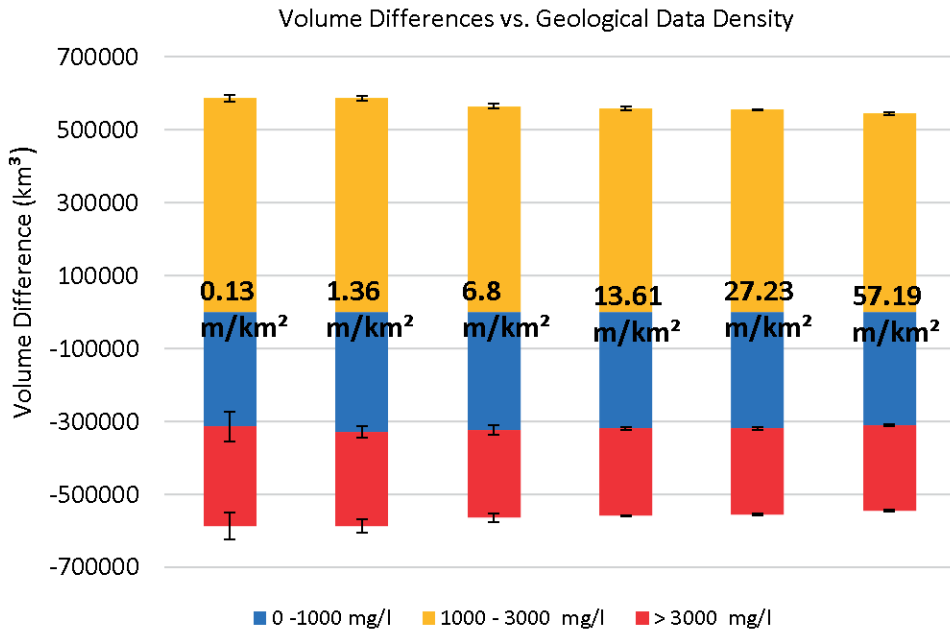
Figure 3.5. *Differences in volume calculations against the reference model of fresh-brackish-saline groundwater regions using the 300m line spaced survey. Km3 calculations are based on applying geological information (FF) to the inverted model (300m line spacing). Blue = fresh, orange = brackish, red = saline. Borehole data density labelled as vertical metres/km2. Positive values represent an overestimation.*

Results for 600m and 1200m flightline distances share a similar magnitude of error. Figure B.5. shows that if no AEM acquisition and inversion error is included, the reduction of errors in estimated volumes are notable. With fewer geological data a trend of overestimating the brackish zone at the expense of saline volumes is seen. Quantitatively, the addition of geological data reduces volume estimate error between the 0.13m/km^2 and 57.19m/km^2 borehole classes by ~$8.0 \cdot 10^3$ km^3, ~$4.0 \cdot 10^4$ km^3, and ~$5.0 \cdot 10^4$ km^3 for the fresh, brackish and saline classes respectively. It is clear however in figure B.4, that that the inversion results in a significant overestimation of the brackish zone, therefore the addition of geological information reduces this error, but only marginally relative to the error introduced by the inversion.

The analysis of errors in volume estimates show that the largest error comes from the AEM acquisition and inversion and that, contrary to errors in salinity concentration estimates, this error is not very sensitive to flightline spacing. Also, the reduction of the volume error from additional geological data is insignificant. Thus, if one is interested in estimating volumes of fresh, saline and brackish groundwater by airborne AEM, large flightline distances and little geological information is the most efficient option with the current methods of AEM acquisition and inversion.

3.3.3 Interfaces

When resolving the 500, 1500 and 3000 mg/l chloride interfaces, possible ECb values may range between 0.06 – 0.13 S/m, 0.11 - 0.25 S/m and 0.2 – 0.4 S/m respectively, in case the FF is unknown and between the values illustrated in figure B.6. The result is that there exists an interface depth uncertainty range of a specified vertical thickness if interfaces are obtained from ECb values. As volume estimates show (figure B.4.), there is a significant smoothing effect from the inversion process - thus it follows that this will affect the thickness of the FF uncertainty range. This effect is demonstrated in cross-section, where the uncertainty limits are presented in figure 3.6 for the reference model and for the 300m flightline spacing inversion model.
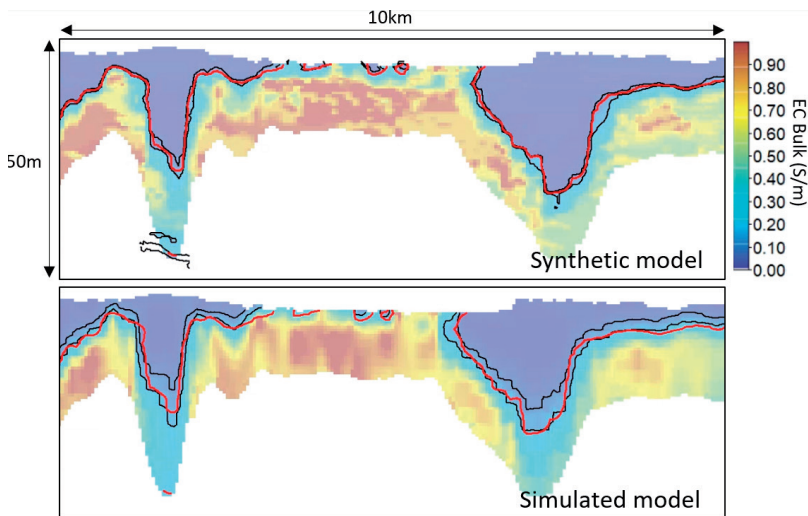
Figure 3.6. *Upper (0.11 S/m) and lower (0.25 S/m) 1500 mg/l interface depth uncertainty ranges (black) for the reference model and inverted model, illustrated as a cross-section. The actual location of the interface in red.*

Figure 3.6 highlights how the vertical uncertainty range is narrow (~1.7m MAE) in the reference model, with substantial thickening (~2.9m MAE) in the inverted model. It is also apparent that the inversion has caused the interface to become shallower overall, which is the result of an overestimation of the brackish zone by the inversion. To understand magnitude of the contribution of flightline distance and borehole density on the error in interface estimation, the MAE of interface positions against the reference model for the 1500 mg/l chloride are shown in figure 3.7. The 500 and 3000 mg/l chloride plots are supplied as supplementary information (figures B.7. and B.8.).
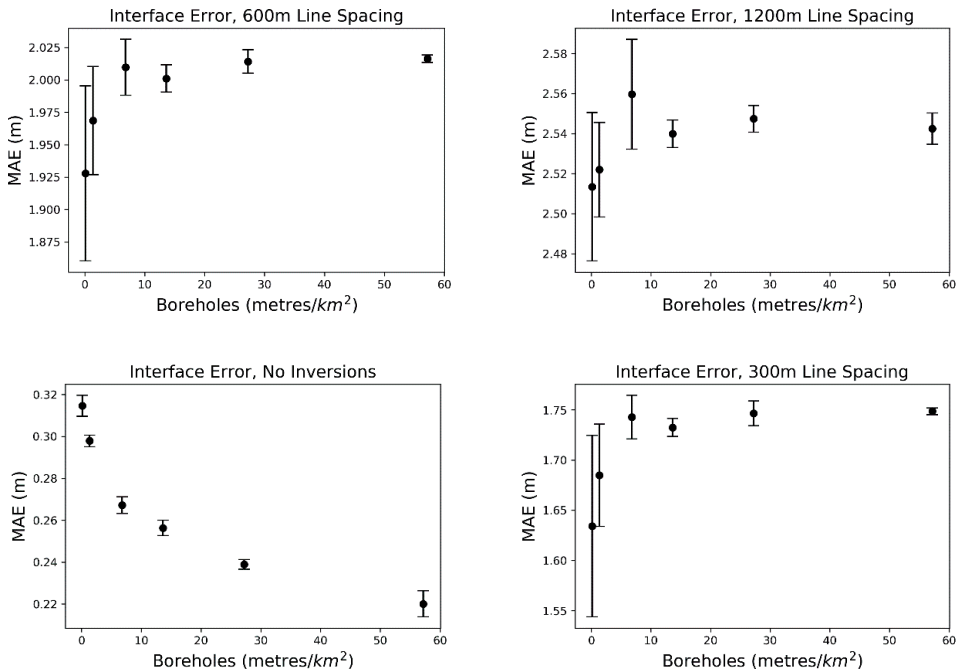
Figure 3.7. *Absolute vertical error of the 1500 mg/l chloride interface for each flightline spacing class, as well as a comparison directly against the reference model.*

Without the inversion, error margins are small (<0.5m) for all interfaces and the addition of geology adds little improvement (between 0.05 - 0.1m). The inversion and increasing flightline spacing however adds a more significant error to the mapping results: ~3, 1.5 and 1m for the 500, 1500 and 3000 mg/l chloride interfaces respectively. For every 300m reduction in flightline spacing, approximately 0.1, 0.25 and 0.4m of accuracy are added for 500, 1500 and 3000 mg/l chloride interfaces respectively – indicating that decreasing flightline spacing is useful for the more accurate delineation of deeper interfaces. However, the addition of geological information does little to help improve error and may even worsen it (for the 500

and 1500 S/m interfaces). This contra-intuitive effect may be caused by the relatively large number of shallower boreholes which pickup more of the clay lithologies that are predominant close to the surface and therefore over-sampled. The relative increase in clay textures and higher FFs results in more fresh groundwater to the unjustifiably classified as brackish and enhances the shallow bias in the 500 and 1000 S/m interface depths.

3.4. Discussion and conclusions

We present a large-scale, quantitative study of groundwater salinity mapping accuracy using AEM. This was undertaken using a high-resolution, 150 million cell synthetic 3D voxel model. We compared 90 possible 3D models, each of which was based on different data configurations of either geological data in the form of lithological logging, and as AEM in the form of flightline spacing. Commonly used and realistic data collection methods were simulated and interpolated into 3D groundwater salinity volumes; these were then compared to the reference model to understand sources and quantities of error. The results are practical and useful as input to regional groundwater management strategies in coastal areas.

It was successfully demonstrated that a careful and systematic handling of the AEM acquisition and inversion process is the most important and cost-effective approach for accurate 3D groundwater salinity mapping using AEM. This supports the findings of Delsman et al., (2018) where the inversion process was the biggest uncertainty contributor to a 3D salinity model. The addition of ground-based geological information will help reduce error, but relatively insignificantly if the inversion process - and therefore the initial distribution of ECb - requires improvement.

As inversions are mathematically ill-posed, countless models can fit the observed data, thus selecting a single inversion that fits the data is not recommended, as demonstrated by Minsley (2011). Inversion error of groundwater salinity mapping was quantified by King et al., (2018), where similar smoothing effects appeared to thicken the brackish zone – resulting in comparable fresh-saline interface mapping errors of ~3m. In this study ~13.5 billion data points across 90 different 3D models were processed, therefore for practical reasons we used a single commonly used inversion method and parameters thereof. For the same reason we used a single interpolation method on the inversion results (King et al., 2018), furthermore we found this suitable as it accurately preserved the thickness of the brackish zone between flightlines. Although testing different inversions and parameters would be interesting, this has been undertaken in a practical manner (Hodges and Siemon, 2008; King et al., 2018), and we feel that this study highlighted the importance of using inversion models carefully in general. Considering this, a number of methodological improvements could be suggested to constrain the thickness of the brackish zone; the simplest of which is to compare inversion methods and

input parameters with accurate and direct ground-based ECb measurements. For example if we know a narrow brackish zone exists from prior knowledge, the LCI sharp method (Vignoli et al., 2015) could be used to try and match this characteristic. Hansen & Minsley, (2019) offer a statistical approach that avoids the use of smoothing regularization constraints – such as the one that caused the thickening of the brackish zone in this study – and would allow the inclusion of explicit a-priori information.

Reducing flightline spacing reliably improved mapping results, particularly for salinity concentration mapping and mapping of salinity interfaces, and seemed to be the most efficient way of trading costs vs. accuracy. The economics involved in AEM surveys however are not straightforward and should be approximated with caution - for example, mobilisation and demobilisation costs mean that the price per line-km is non-linear and highly site-specific. For the sake of discussion however, as a coarse current estimate we could approximate €50,000 for mobilisation and €200 per line-km flown, including inverse modelling and full interpretation (Deltares, personal communication, September 2019). As a result, using these rough metrics the 300, 600 and 1200m line spacing (or ~3000, ~1500, ~750 line-km) survey results could total ~€650k, ~€350k, and ~€200k respectively.

To conclude, we showed that when mapping groundwater salinity using airborne EM (AEM) and ground based geological data, the error from AEM inversions far exceed that from errors in estimated formations factors (FF) obtained from lithological mapping. As a result, and given the estimated costs of AEM and geological drilling, decreasing flightline distances and producing high-quality inversion results are the most effective and efficient way of increasing the accuracy of groundwater salinity mapping, in particular when estimating salinity concentrations and the depth of interfaces between fresh and brackish groundwater. Volume estimates of the entire study area are however less sensitive to flightline distance. Overall results presented here allow a quantitative understanding of mapping error in large scale groundwater salinity mapping campaigns. This information is useful for the planning of regional groundwater mapping programmes using AEM in general. The scale and realistic nature of the synthetic model mean that the results are relevant over a broad range of hydrogeological terranes.

3.5. Acknowledgements
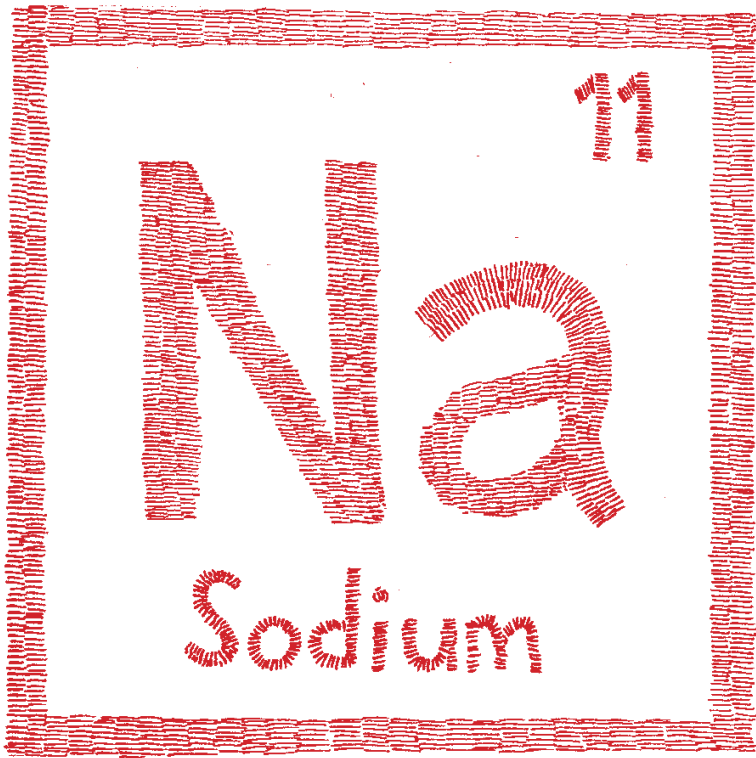
4

Controlling the smoothness of airborne geophysical inversions to improve the accuracy of regional groundwater salinity mapping

Abstract

The ongoing threat of saltwater intrusion into low-lying, coastal aquifers, necessitates a good regional understanding of this phenomenon. In these areas, airborne electromagnetic (AEM) surveys are increasingly used to create maps of fresh-saline groundwater distributions – a process that is essential for effective groundwater management. To be used practically, AEM data needs to be converted from observations to a subsurface distribution of electrical conductivity (EC), where finally petrophysical relationships are used to convert EC into groundwater salinity. This is undertaken using a process called inversion, where AEM observations are typically modelled iteratively until a subsurface distribution of EC is matched to AEM observations to an acceptable fit. Unfortunately, this procedure comes with great uncertainty as an infinite number of models can explain the data. Furthermore, recent research has highlighted that commonly used inversion methods tend to produce overly thick brackish zones in some environments – resulting in inaccuracies, particularly regarding the vertical positions of fresh-brackish-saline interfaces. As a result, in this research we developed an inversion approach to resolve the distributions of fresh, brackish and saline groundwater more accurately by using additional, cheaply available ground data and existing inversion methods. Using in-built inversion parameters of a layered earth inversion method and a customised script, a regularisation term (or penalisation) can be set to favour either the default inversion parameters, or the observed thickness of the brackish zone from in-situ data. The method was tested using the commonly implemented laterally constrained inversion from AarhusINV, using both synthetic and real data from the Netherlands. Results highlight some improved accuracies regarding the mapped location of fresh-brackish-saline interfaces; however, it was found that further development is required for meaningful results.

4.1. Introduction

Airborne electromagnetics methods (AEM) have been used since the late 1970's to map groundwater salinity (Fraser, 1978). Given the increasingly understood risks to the availability of fresh groundwater from salt water intrusion (Barlow and Reichard, 2009; Bocanegra et al., 2009; Custodio, 2009; Oude Essink et al., 2010; Post et al., 2018; Werner, 2009; Werner et al., 2013), the method has since been applied successfully to map this phenomenon in many parts of the world (e.g., Auken et al., 2008; Chongo et al., 2015; Delsman et al, 2018; Fitterman and Deszcz-Pan, 2001; Ley-Cooper et al., 2018; Siemon et al., 2015). The sustained popularity of AEM relates to the ability of the method to rapidly and cost-effectively map regional-scale groundwater salinity (King et al., 2020). Furthermore, compared to the use of in-situ (ground-based monitoring), AEM surveys generally result in a continuous, 3D salinity model of the subsurface – allowing an effective regional understanding of groundwater salinisation.

AEM surveys are sensitive to electrical conductivity (EC) contrasts caused by both clay content of the host sediments and the salinity of pore water (Revil and Glover, 1998), and as such are also used to map the structure of aquifers (e.g. Auken et al., 2008; Gunnink et al., 2012) alongside groundwater salinity distributions (e.g. Delsman et al., 2018; Vandevelde et al., 2018). A comparison of available AEM methods for groundwater mapping purposes is described in Steuer et al. (2009). These include frequency and time-domain electromagnetic methods (referred to here as FEM and TEM respectively). As we are using acquired FEM data in this study, in the following we will refer exclusively to this method and application only, however the results presented here are applicable to both. For an overview of applied groundwater mapping using AEM in general, including survey characteristics and data processing, refer to Siemon et al. (2009).

FEM surveys transmit electrical signals from a towed instrument, referred to as the primary field. These signals produce secondary currents which are in turn recorded by receivers as parts per million relative to the primary field. To extract useful information from AEM measurements, it is converted into a usable EC distribution – or specifically bulk EC (ECb), which comprises signal from the host sediment and pore-water. In this procedure ECb is modelled using the secondary field data, a process referred to inversion which is typically approached using 1D layered earth models (e.g. Auken and Christiansen, 2004; Brodie and Sambridge, 2006; Farquharson et al., 2003; Siemon et al., 2009a; Viezzoli et al., 2008; Minsley, 2011; Vignoli et al., 2015). For an overview of available AEM inversion methods and a quantitative understanding of accuracies in relation to groundwater salinity mapping, we refer to King et al. (2018). Inversion results are then translated into groundwater salinity using petrophysical relationships (e.g. De Louw et al., 2011). For brevity, in the following we refer to groundwater salinity as chloride – as this is the dominant conservative anion found in coastal saline groundwater.

Inversions suffer from non-uniqueness, meaning an infinite number of models can fit the data. As a result, these methods need stabilisation, reducing the degrees of freedom – otherwise known as regularisation (Constable, 1987; Marquardt, 1963). In general, a result of this is that the user will need to parameterise the inversion such that a balance is found between fitting the data and a regularisation term that imposes a degree of simplicity or smoothness to the inversion result. In practical terms, inversion results could either be too smooth (favours the regularisation too strongly) or contain noise present from the acquisition process (it is constrained too weakly). In the case of chloride mapping, fresh groundwater volume estimates could deviate by over 5% depending on parameterisation (King et al., 2018). Using synthetic modelling, King et al. (2020) showed that by using default inversion parameters of a commonly used inversion method, the fresh-saline interface appeared on average almost 3m too shallow. This error has potentially significant consequences for quantitative groundwater management policies. For instance, in the Province of Zeeland, The Netherlands concentrations of 0-300, 300-3000, > 3000mg/l chloride are commonly used to classify fresh, brackish, and saline respectively (Van Baaren et al., 2016), and this classification determines greatly determines the utility and application of water resources. Naturally, it follows that the correctly parameterising inversions are highly important for quantitative groundwater management.

In light of the above, we developed a novel method that uses readily available in-situ data to carefully balance smoothness vs. noise present in inversions. Our approach uses the simple idea that inversion smoothness affects the thickness of the brackish zone and by extension this feature is a good proxy for appropriately parameterised inversions. Given that relatively cheap in-situ information is readily available in low-lying coastal areas, it serves an economical constraint that is likely to be available even in data-poor areas. The amount of noise present is controlled by penalising inversions that are parameterised differently from default parameters. The penalisation, or regularisation term, can be set to favour either the default parameters or the observed thickness of the brackish zone in the in-situ observations. For brevity we refer to traditional (or in-built) inversion parameters as parameters, and our additional regularisation as penalisation. The method was first tested using a synthetic model in order to find the optimal weighting between regularisation and data-fit and then applied to a real case using data from the Province of Zeeland, the Netherlands.

In the following, Section 4.2 provides details of our methods - including a summary of our approach in section 4.1 followed by a description of how the synthetic model was created in section 4.2. Sections 4.3 to 4.5 outline technical details of the optimisation method and 4.6 to 4.7 describe applications to the synthetic and real cases respectively. The results are presented in section 4, where 4.1 and 4.2 describe results of the synthetic and real cases respectively. Finally, a discussion and concluding statements are described in section 4.4.

4.2. Methods

4.2.1 Summary of approach

To find an appropriate balance between sharp and smooth inversion models using additional in-situ data, a customised regularisation strategy was developed. This strategy is based on a minimization criterion that penalizes both the deviation from an initial inversion (without additional in-situ data) and the deviation from in-situ observations. In effect, the approach allows the use of simple information – in this case an estimate of the brackish zone thickness – to parameterise an inversion while still allowing a control on the introduction of noise because of an overly sharp inversion result. The method was adapted to a commonly used layered-earth, deterministic inversion method, but could feasibly be adapted to other inversion methods.

The procedure was first tested using a synthetic model to obtain a quantitative idea of the correct balance between fitting the inversion to default parameters or to in-situ data constraints. Information from the synthetic test was then used as a guide to test the method using real AEM data and known brackish zone thicknesses. The approach is illustrated in figure 4.1, where the both the synthetic study and the real case are shown.

In general, an iterative process runs an inversion, starting with purposefully overly smooth constraints. The estimated thickness of the brackish zone is then extracted and compared against the in-situ estimates and those of the default inversion parameters. The regularisation process then calculates a value of the criterion (say J) that is based on weighting (using a weighting parameter λ), that penalizes either a deviation from the initial inversion using the default (smoother) parameters (λ) or a deviation from the in-situ data (1- λ) leading to generally sharper inversions. The procedure then restarts and applies a pre-factor to the inversion parameters, essentially running the inversion again (as above), but with a slightly sharper model. The process iterates until a given maximum sharpness is reached, which means that the criterion J is minimalized for the given value of λ to extract the optimal inversion parameter. This process thus quantitatively guides the user to select inversion parameters based on a desire to either obtain a sharper or smoother model. We used a realistic synthetic case to determine which value of λ leads to the optimal weighting between smoothness and fit to the in-situ data. The synthetic case was also used to ascertain how much in-situ data is needed to obtain robust results. The resulting value for λ was then used to test the method on a real case, where extensively available ground and airborne data was used to validate the method in general, as well as the appropriateness of the given λ from the synthetic result.
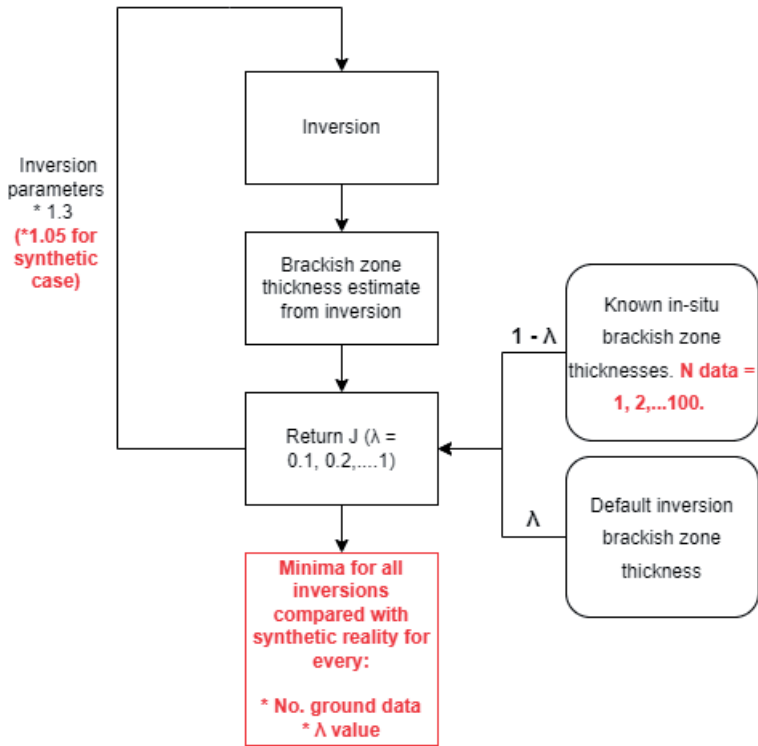
Figure 4.1. *The method outline. Red areas denote additional steps relating the synthetic case, for the real case only steps in black are shown. λ is a weight between deviation from the initial inversion and the in-situ data.*

The process is described in more detail in the following order: a) the synthetic model creation, b) the inversion method, c) the minimization criterion and d) an application to a real case.

4.2.2 Creating the synthetic case

The synthetic model was created from an existing AEM based 3D salinity model to preserve realistic site characteristics. A detailed description of the model, the AEM survey, and subsurface characteristics of the area are described in Delsman et al. (2018). From this survey, we selected a flightline (and subsurface data thereof) that hosted characteristic features of coastal groundwater mapping – such as both shallow and deep freshwater lenses – as well as an area with saline groundwater at the surface. The synthetic model is shown in figure 4.2A. To simulate an AEM survey, the synthetic model was first converted from chloride to ECb using a petrophysical transformation with a procedure described in King et al. (2020). The resulting 2D section of ECb values were then forward modelled using AarhusINV inversion software (Auken and Christiansen, 2004), which is based on the Fugro Resolve frequency domain AEM system used for the real survey. To help simulate realistic survey conditions, flight altitude was variable and taken from the original survey. Finally, 5% gaussian noise was

added to forward modelled data, based on known estimates of acquisition errors (Farquharson et al., 2003).

4.2.3 The inversion method

The commonly used Laterally Constrained Inversion (LCI) was selected for this study (Auken et al., 2015a). The method is a deterministic style, layered earth inversion method and is suitable for the laterally continuous conductivity contrasts found in low-lying coastal areas. Besides site-suitability, the method was selected purposefully as previous research found that these inversions can cause an overly thick estimated brackish zone when incorrectly parameterised (King et al., 2020).

The inversion was discretised using a 20-layer model, with each consecutive layer logarithmically increasing with thickness from surface until 50 m depth. Each layer was assigned a starting conductivity of 1 S/m owing to the highly saline environment of our study area. Data beneath an estimated depth of investigation (DOI) was removed based on the method by Christiansen and Auken. (2012). The inversions were run using AarhusINV executables (Auken and Christiansen, 2004) and a customised script. The inverted synthetic model is illustrated in figure 4.2B, where default inversion parameters were used.

4.2.4 The minimization criterion

To extract optimal inversion parameters the following minimization criterion was used (Equation 4.1):

$$J = (1-\lambda)\frac{1}{n_{in-situ}}\sum_{i=1}^{n_{in-situ}}\left(\frac{D_{it,i}-D_{obs,i}}{\sigma_{obs}}\right)^2 + \lambda\frac{1}{n_x}\sum_{j=1}^{n_x}\left(\frac{D_{it}(x_j)-D_{def}(x_j)}{\sigma_{def}}\right)^2$$

Where $D_{it,i}$ is the estimated brackish zone thickness at the location of the in-situ data and $D_{it}(x_j)$ at the location x_j along the flight line as obtained from a given iteration, $n_{in-situ}$ the number of in-situ data, n_x the number of vertical salinity estimates (and brackish zone thickness estimates) along a flight line, $D_{obs,i}$ is the observed brackish zone thickness from in-situ observation i and $D_{def}(x_j)$ the calculated thickness at location xj for default inversion parameters. Data were normalised using standard deviations of estimated errors in-situ (σ_{obs}), or default (σ_{def}) estimates of brackish zone thickness. The second term in this criterion is the regularization term and its strength is determined by the value of λ changing between 0 (favouring the in-situ data) to 1 (favouring the default inversion parameters).

4.2.5 Details of inversion and minimization procedure

To calculate J for different inversion parameters and λ values, an iterative process was developed in Python. First, an inversion is run with an initial set of parameters. The resulting

inversion (in ECb) is then transformed into the electrical conductivity of groundwater (ECw) using petrophysical relationships in the form of formation factors (FF) (Archie, 1942). For this we used an approximation of 3.13, which is the known value for sandy clay in the local area (De Louw et al., 2011) and the same used in King et al., (2018). To estimate the area of the inversion that represented the brackish zone, we used groundwater salinity classifications set out by Paine and Minty (2005): 0–0.18 S/m (fresh), 0.18–1.8 S/m (brackish) and > 1.8 S/m (saline). Unfortunately, a full brackish range was not present at many AEM measurement locations (e.g., a vertical profile may comprise only 0.18 – 1.2 S/m) – this was particularly the case after applying a depth cut-off based on DOI estimates of the AEM inversion results. As only full sequences of brackish data were permitted in the procedure, the resulting removal of useful data was potentially problematic. We therefore instead used a portion of the brackish range as proxy for a brackish thickness, given by 0.18 – 0.54 S/m. This was found to greatly increase the amount of available data from inversion results, while still being sensitive to inversion parameterisation. The in-situ ECb data for both synthetic and real applications were processed as above to obtain brackish zone thickness estimates. The process then calculates the difference in brackish zone thicknesses between nearby in-situ data and the default inversion parameters.

Results are then appended and the iterator repeats using updated inversion parameters (using a simple grid search) until J is minimized. The inversion parameters are changed at each iteration using the given pre-factor, here inversions are purposefully run from overly smooth to overly sharp to obtain a full range of values based on the given λ. The LCI inversion parameterisation comprise horizontal and vertical constraints that control the smoothness of respective dimensions, for a more detailed explanation refer to Auken et al. (2005). For brevity we will refer to horizontal and vertical constraints as HCON and VCON respectively. For this study, HCON and VCON constraints were not changed individually, but rather linked by changing the values by the same pre-factor, as per published advice in an AarhusINV manual (Aarhus HydroGeophysics Group, 2018).

For the synthetic case, the minimization of J as described above is repeated for different λ using equation 1, with λ ranging from 0 to 1 in 0.1 increments.

4.2.6 Application to the synthetic case

Our approach was applied to a synthetic case by first extracting brackish thicknesses along a salinity profile from an existing AEM estimated salinity distribution, as described in subsection 2.5 The resulting thickness estimates were then used to simulate in-situ data collection (*Dobs*), where 100 measurements were extracted. An inversion was run with default parameters (HCON = 1.3, VCON = 3) using the forward modelling data; the results were used to extract default inversion brackish zone thicknesses (*Ddef*). The procedure was then run iteratively, starting with HCON = 1.075, VCON = 1.75, then using a pre-factor of 1.05 per

iteration until VCON became greater than 50. At the end of each iteration *J* was calculated for every combination of *Dobs* (n=2,3….,100) and λ (n=0,0.1,….,1). In total 120 different inversions were run, each with different parameters, taking about 4 hours to complete on a standard quad-core desktop PC. Three resulting inversions are illustrated in figure 4.2, where a smoother, default setting and sharper inversion are illustrated.
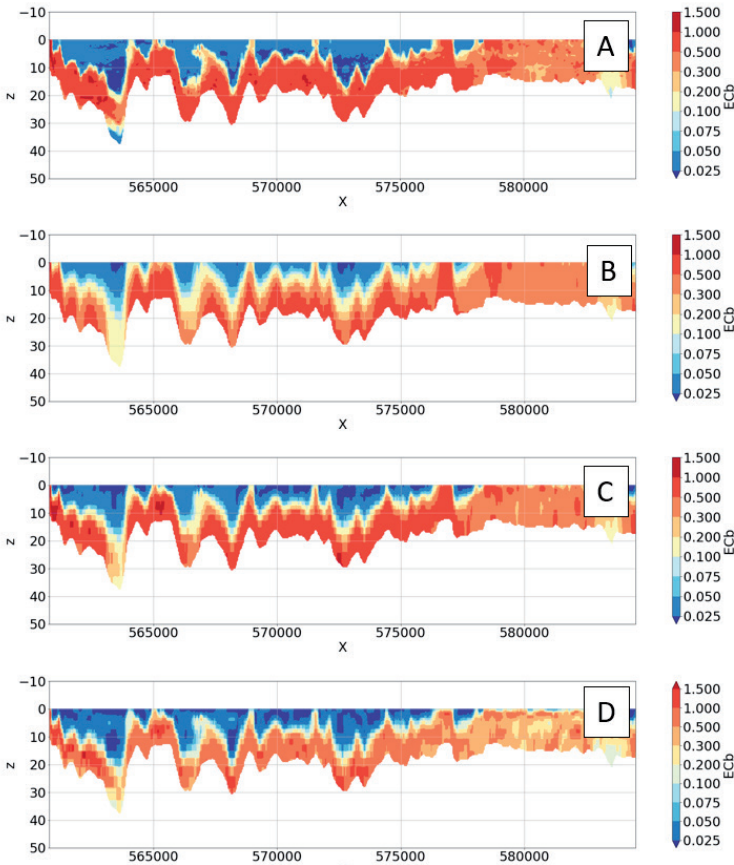
Figure 4.2. *The uninverted, synthetic model (A), the synthetic model inverted with smoother parameters (B) (HCON = 1.08, VCON = 1.5), default parameters (C) (HCON = 1.3, VCON = 3) and sharper parameters (D) (HCON = 1.5, VCON = 10).*

4.2.7 Application to a real case

The real case was tested in an area of the Province of Zeeland, the Netherlands – in an area of around 270 km*2*. Available data comprise 50 flightlines of raw, uninverted AEM data totalling ~1000-line kilometres, and AEM measurements were taken every ~4m. Available in-situ (*Dobs*) data comprise 29 vertical Electrical Cone Penetration Tests (ECPT) profiles of observed ECb, measuring to an average depth of ~25 m below surface and were sampled at a vertically resolution of 0.5 cm. Locations and characteristics of the data are shown and described in detail in King et al. (2018). Out of the 29 available ECPT profiles, five were kept apart solely for validation purposes, the remaining 24 were used to extract brackish zone thickness estimates (*Dobs*) for the inversion and minimization of Eq. 1. Using the optimal value of λ found in the synthetic case, inversions were first run using default parameters and converted to ECw (as described above) to obtain default inversion brackish zone thickness for all flightlines (*Ddef*). The described procedure was then run iteratively for every flightline, starting with HCON = 0.08, VCON = 0.53 and using pre-factor of 1.3 until VCON became greater than 50, – resulting in 18 separate inversions for each flightline. At the end of each inversion, all in-situ data < 500m radius from the flightline was assigned to the nearest AEM measurement and used as input for *Dobs* in Eq 1. As 500m is the upper estimate for the AEM systems footprint (Yin et al., 2014), data collected within this radius should be spatially consistent with the inversions. The process was repeated for all 50 flightlines, totalling 1000 inversions (20 inversions per each flightline), taking a total of ~18 hours to complete on a standard quad-core desktop PC.

Finally, to quantitatively test differences against the validation set of five ECPT data, the default and resulting minimised inversions were then interpolated into a 3D volume of ECb. This was undertaken using 2D ordinary Kriging of each of the 20 inversion layers, as per the method of King et al. (2018). Data below the calculated DOI was removed from further analysis. The resulting 3D estimates of groundwater salinity (after conversion using formation factors) was then compared to the observed salinities in the 5 ECPT profiles kept apart for validation.
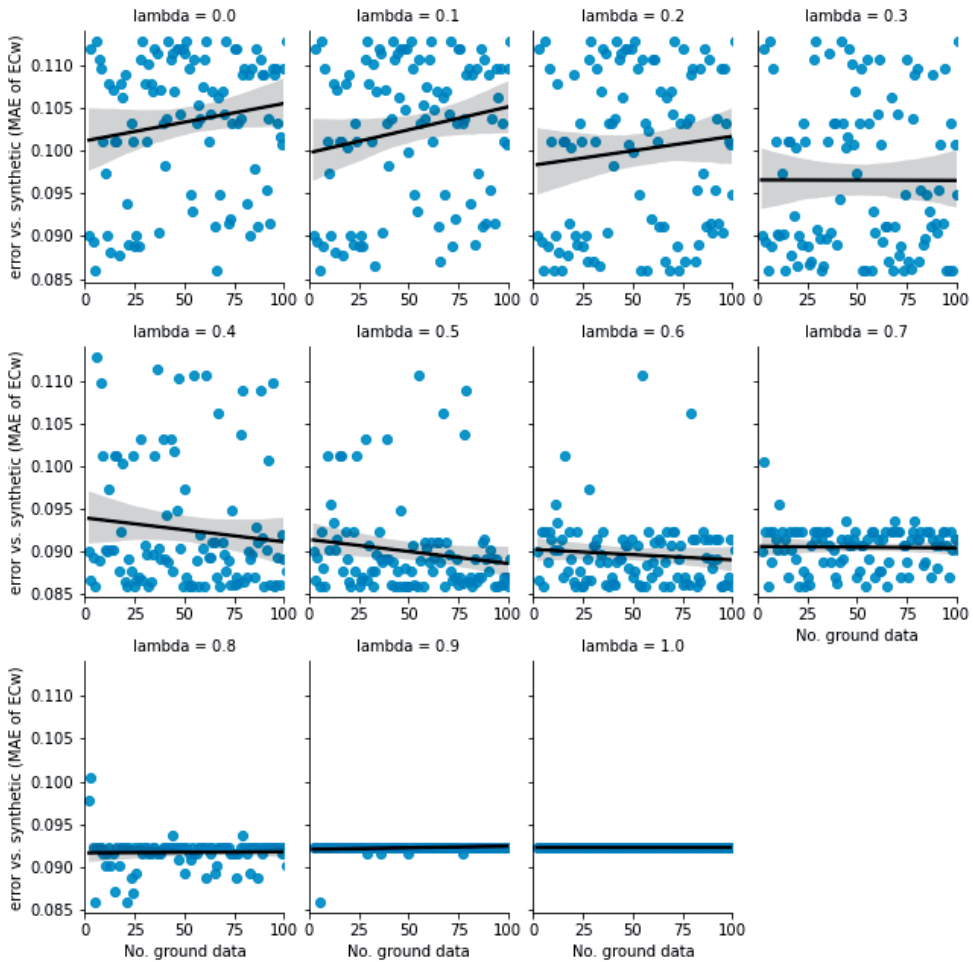
4.3. Results

As results from the synthetic model are used to inform the real case, the synthetic model will be presented first (section 3.1), followed an application of the method to real data (section 3.2).

4.3.1. Synthetic model results

Figure 4.3 shows results from extracting the inversion result that resulted in a minimum J for each number of available in-situ data and for different values of λ. Each point represents a fit to the synthetic model in ECw (S/m) for a combination of λ and number of in-situ of data

used. For each inversion, the resulting mean absolute error (MAE) in estimated ECw (S/m) with respect to the synthetic reality is plotted.

Figure 4.3. *MAE of inversion results against the synthetic reality, represented as S/m difference of ECw vs. number of in-situ data as a result of the favoured inversion sharpness returned by J.*

Here it is observed that inversions that more strongly favoured the in-situ data (~λ 0 to 0.4) displayed a greater degree of uncertainty in finding the optimal inversion result, as they are too noisy. Naturally, inversions that favoured the default parameters were more consistent, with a mean synthetic model error of 0.093 S/m against ECw estimates. The middle to upper λ values (0.6 to 0.8) showed a good balance between obtaining consistent results and achieving a better fit against the synthetic model. Within these ranges, it was also observed that a relatively small quantity of in-situ data is required for reliable results, in the range of two to

five measurements. Although results in this range are somewhat noisy, further improvements using more in-situ data is limited as shown by the regression lines. Figure 4.4. shows the results for each value of λ as averaged over all amounts of in-situ data. The average MAE shows that a λ of 0.6 performed the best overall.
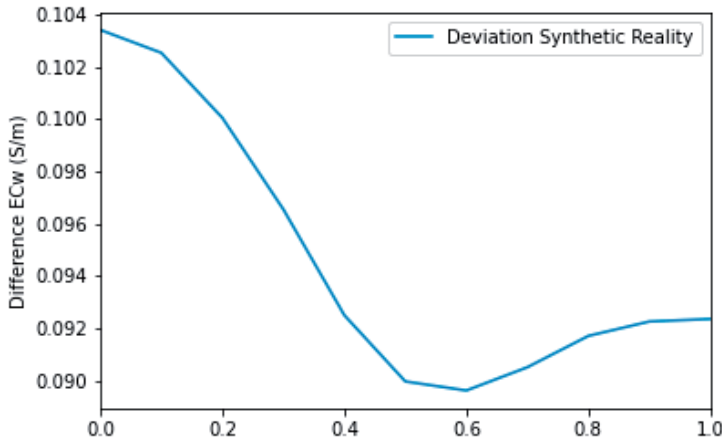
Figure 4.4. *Fit of inversion results vs. the synthetic model as MAE in S/m of ECw for each value of λ averaged over all amounts of in-situ data used.*

From the synthetic model results, it follows that the inversion parameters that relate to λ = 0.6 resulted in the best fitting inversion, and that around two or more estimates of available in-situ measurements of brackish zone thicknesses is required to find the corresponding optimal inversion parameters in a real case.

4.3.2. Real case results
Out of 50 available flightlines, nine were within 500m of in-situ data. For these nine flightlines, the procedure outlined in section 2 with λ = 0.6 was performed. Table 4.1 shows for each of the selected flightlines the MAE (of ECb in S/m) from the default inversion for the inversion with the minimum value of criterion (eq. 1) and the related inversion parameters. Out of these, five favoured the default parameters and four suggested that sharper inversions were more optimal. The availability of in-situ data ranged from two to six for each flightline. Figure 4.5 shows results of the procedure for a single flightline that preferred sharper constraints than defaults, where the balance between fitting in-situ data and default parameters is illustrated.

Table 4.1. *Summary of results from λ = 0.6 is applied to all available in-situ and airborne data. Results are listed individually for each flightline.*

Error in-Situ (in metres of brackish zone thickness)	MAE from default (as ECb in S/m)	No. in-Situ	HCON	VCON	Flightline
0.673145416	0	6	1.3	3	L10
1.460649195	0.138962	6	1.389	3.601	L11
3.087383168	0	4	1.3	3	L12
3.000744495	0	2	1.3	3	L13
9.024223901	0.672868	2	1.506	4.381	L6
2.736390154	0	4	1.3	3	L8
0.809435049	0.497443	4	1.506	4.381	L9
2.845341124	0	3	1.3	3	T109
3.674896976	0.856185	3	1.658	5.395	T111
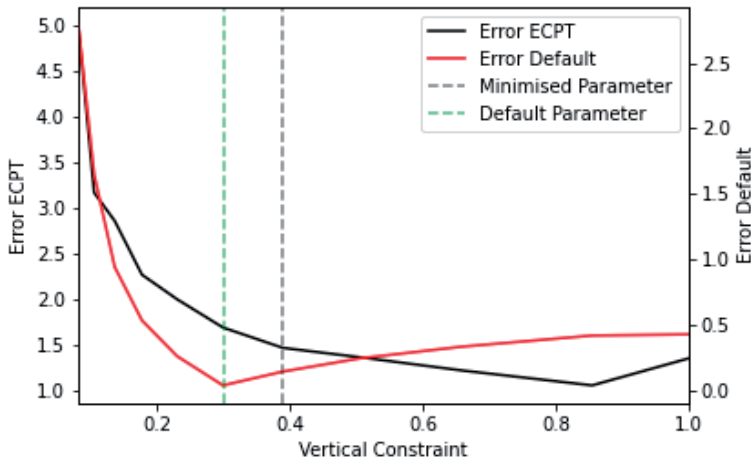
Figure 4.5. *Example of output from minimisation procedure for a single flightline (L11), using λ = 0.6 and six nearby in-situ data for the calculation. The x-axis shows constraints ranging from smoothest to sharpest. Extracted minima and corresponding constraints are highlighted. The y-axis shows brackish-zone thickness differences in metres against in-situ data and the default inversion.*

Using the mean of inversion parameters shown in table 4.1., applying λ = 0.6 resulted in preferred inversion parameters of VCON = 1.4 and HCON = 3.64, slightly higher than that of the defaults. As a result, inversions using these parameters were run on all 50 flightlines. The upper most estimate of inversion parameters from a single flightline, as noted in flightline T111 in table 4.1 with VCON = 1.658 and HCON = 5.395, were also extracted for further analysis for a complete comparison. In the following we refer to the mean (VCON = 1.40 and HCON = 3.64) and upper (VCON = 1.658 and HCON = 5.395) estimated parameters as the "averaged" and "upper estimates" respectively.

4.3.2.a. Comparing Inversions: 2D Sections

Results of these inversions are qualitatively compared in figure 4.6 for a single flightline against default parameters. All flightlines displayed similar qualitative characteristics, therefore only a single flightline is shown for brevity.
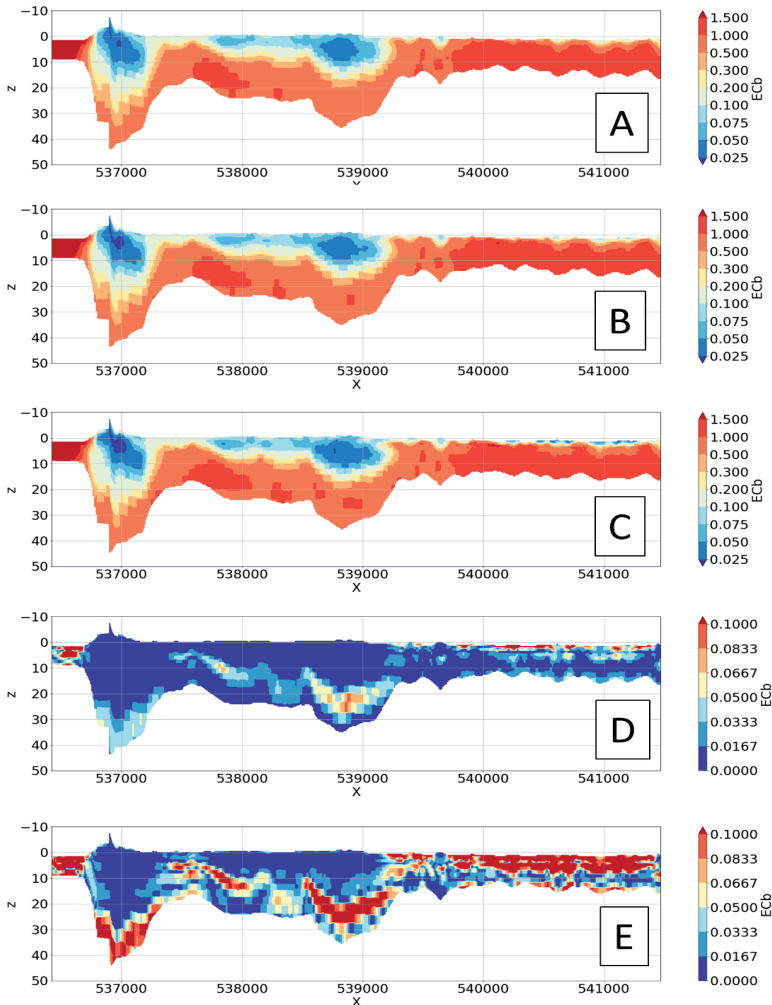

Figure 4.6. *A: Inversion result along a flightline from using default parameters; B: inversion result from averaged minimised parameters; C: inversion result from upper minimised parameters; D: the absolute difference between A and B; and E: the absolute difference between A and C.*

Initial observation shows there is little obvious difference between default parameters (figure 4.6A) and the new inversions from the averaged optimisation (figure 4.6B) and the upper estimate (figure 4.5C). Comparing optimised parameters against defaults (figure 4.6D and 4.6E), it is noted that differences occur mostly in the brackish zone, as expected – these

are mostly in the range of 0.05 to 0.1 S/m for average parameters and up to 0.5 S/m for the upper estimate. The difference between the upper estimated inversion parameters and the defaults are more noticeable in magnitude, though both share similar spatial distributions. The MAE against default parameters for both averaged and upper estimates are 0.02 and 0.06 S/m respectively. In general results show that the fresh-saline transition, and therefore likely the locations of fresh-brackish-saline interface, are different between default and minimised inversion parameters, tending towards more fresh groundwater at shallower depths, i.e., a deeper fresh-brackish interface.

4.3.2.b. Comparing Inversions: Volume Estimates

As the method presented here focusses on the use of brackish-zone thicknesses to improve inversions, we used the 3D interpolated results to compare the total volume of fresh-brackish-saline regions for each. For this we used volumes of ECw, and the classifications set out in section 2. Results of these estimates are presented in figure 4.7.
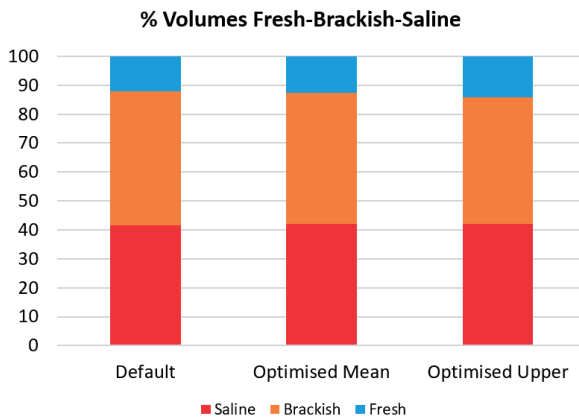
Figure 4.7. *Volume estimates for each inversion. Red (saline, > 1.8 S/m), orange (brackish, 0.18–1.8 S/m), and blue (fresh, ≤ 0.18 S/m).*

Here it was observed that fresh-water volume estimates were smallest for the default parameters and largest for the optimised upper estimates at 12% and 14% of total model volumes respectively. The estimated brackish zone varied in volume from 46%, 45% and 44% for the default, optimised mean and optimised upper inversions respectively. Saline volume estimates were consistent for all inversions, occupying about 42% of the total model volume. This confirms the slightly deeper fresh-brackish interface that was shown in the flightline with different parameters in figure 4.6 (previous section).

4.3.2.c. Comparisons against validation data: Direct Comparisons

The MAE of each inversion against the five in-situ (ECPT) data are shown in table 4.2. Results from using the 3D interpolated model as well as simply fetching the nearest ECPT data to a flightline are shown to quantify potential error as a result of the interpolation process.

Table 4.2. *MAE of inversion results vs. available in-situ data in S/m of ECb.*

	Default	Optimised Mean	Optimised Upper
Error against ECPT, 3D interpolation	0.557	0.567	0.578
Error against ECPT, direct comparison	0.509	0.521	0.527

From table 4.2 it follows that although the optimised inversions reduced the total volume of the brackish zone, the MAE of the evaluation against ECPT data were slightly better for the default parameters. Error estimates are lower for the direct comparisons than the 3D interpolated results – most likely due to resolution downsampling to 50 x 0.5m (from a downline measurement spacing of ~4m). However, they are considered close enough and share the same magnitude in error difference between each inversion. As a result, further comparisons using the 3D interpolation results are considered valid. Figure 4.8 demonstrates two direct (i.e., not from the 3D interpolation) in-situ data comparisons against the inversion results.
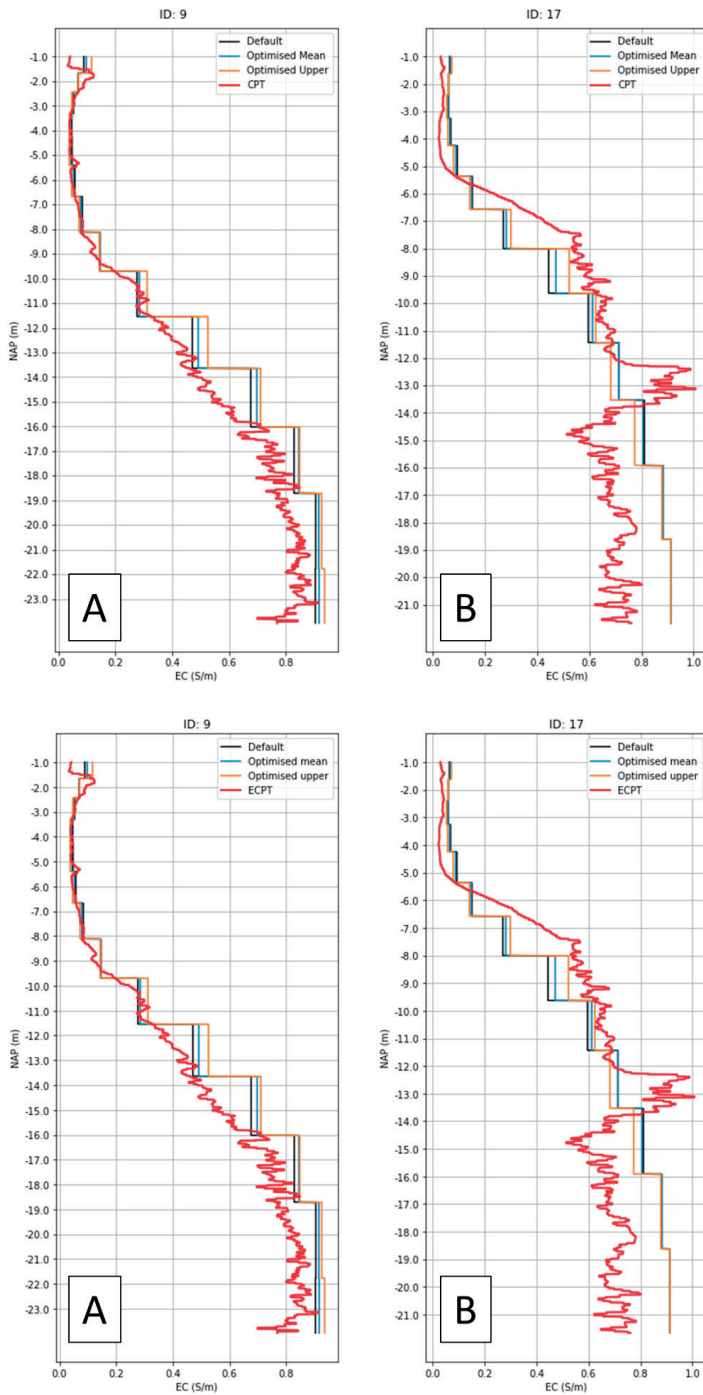
Figure 4.8. *Comparisons between inversion results and two available in-situ data.*

In figure 4.8A it is observed that the optimised inversions in general perform slightly worse than the default inversion, generally overestimating ECb, however subtle improvements over the defaults are also noted in the shallow (1 to 3 m depths) area. In figure 4.8B the optimised inversions appear to match in-situ data better in the fresh-saline transition zone (5 to 12m), with no noticeable differences at greater depths.

4.3.2.d. Comparisons against validation data: Interface Mapping Accuracy
3D iso-surfaces of fresh (0.18 S/m), centre of brackish (0.54 S/m) and saline (1.8 S/m) groundwater interfaces were extracted from the 3D model for all inversion results, according to the classification set out in section 2. Likewise, corresponding point data were extracted from the available ECPT validation data. Vertical location error statistics between the 3D interfaces and validation data was then calculated. Error statistics are presented as MAE to obtain overall error, and as ± values to determine if estimates were either too shallow or too deep - positive values correspond to inversion estimates that are shallower than the validation data.
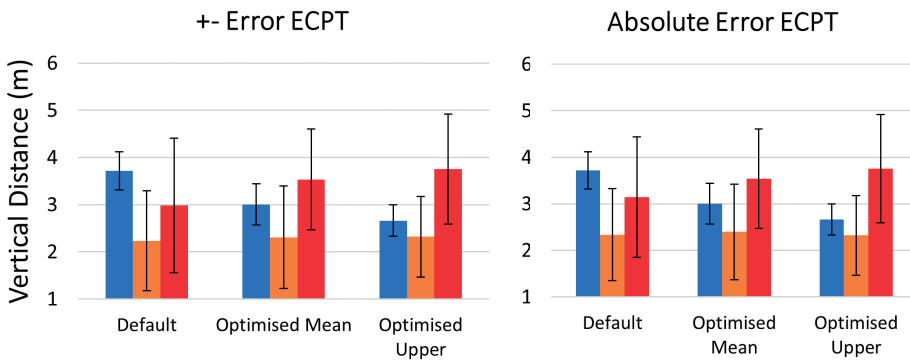
Figure 4.9. *Vertical error of extracted interfaces from inversion results against validation data. Red (saline interface error, 1.8 S/m), orange (brackish interface error, 0.54 S/m), and blue (fresh interface error, 0.18 S/m).*

In figure 4.9, it is observed that MAE and ± error are practically identical, highlighting that interface mapping was consistently too shallow for all classifications. The brackish interface was the most accurately mapped, with an error of ~2.3m for all inversions. The optimised inversions more accurately resolved the fresh interface, with an error of ~3 and 2.5m for the optimised mean and optimised upper, respectively – against a value of ~3.7m for the default inversion. In contrast, the default parameters mapped the saline interface with an accuracy of ~3m, against values of 3.5 and 3.8m for the optimised mean and optimised upper inversions.

4.4. Discussion & Conclusions

In general, while the results from the real case showed some improvements and highlighted the potential usefulness of cheap in-situ data, these were not as drastic as hoped for the problem at hand – despite the robustness of the developed method. As a result, we will examine possible reasons for this, and finally suggest further research needed to build upon and improve the approach discussed here. In the following we will first discuss the results of the synthetic model and the resulting implications for the study in general, followed by an analysis of results obtained in the real case.

Results from the synthetic model suggested that overall, a better fitting inversion could be achieved by sharpening the inversion parameters over default values. This result was based on the observation that a better fit to the known synthetic reality was achieved over the default inversion for λ values between 0.5 to 0.9 (1 being the default). Given that the in-built regularisation constraints used by the LCI inversions are designed to smooth out structure mapped from noise, this conclusion assumes that our estimate of noise (5% in this case) was accurate. A noise estimate that is too high or too low, would result in the synthetic case selecting inversions parameters that are too smooth or too sharp (or noisy) respectively. As the 5% noise threshold selected for this study is considered an upper estimate in actual surveys (Farquharson et al., 2003), in our real case this may lead to selecting inversions that are too smooth – or indeed simply not different enough to those achieved from using default parameters. A further consideration is that the synthetic model was based on inversions itself, and thus the optimised inversion parameters could have favoured those that were used to create it in the first place. As this model was probabilistic (Delsman et al, 2018), it was based on the results of three different inversions. Furthermore, the resulting 3D chloride model was then post processed, whereby chloride values were reclassified – ultimately resulting in a sharper model than the inversions would've created. As a result, we consider that using this synthetic model would not have biased the results of this study.

In the real case, after running the optimisation on all 50 flightlines and available in-situ data using λ = 0.6, it was found that on half the flightlines that were close to in-situ data, the procedure selected default parameters. By averaging the resulting optimised estimates, a slightly sharper inversion was suggested (using VCON = 1.4, HCON = 3.64) – which in this case was only one parameter set higher from the defaults based on the pre-factor method used. Out of these results, we included the upper most estimate (i.e., sharpest suggested inversion) to quantify whether a sharper inversion would be preferable – and potentially highlight if the 5% noise added to the synthetic model was too high – and therefore if a λ of > 0.6 would have been more appropriate. Overall, relatively subtle differences between the default and optimised inversions were observed – both with qualitative observations comparing 2D sections, as well as a quantitative fit to in-situ data. However, MAE against in-situ data between the default

(0.56 S/m ECw) and optimised (0.57 S/m ECw) showed an increasing trend with sharpness. As similar observations were observed between the 3D interpolated and directly compared inversions - ostensibly, this points to towards the sharper inversions introducing unwanted structure from noise and is unrelated to 3D interpolation artefacts. It also indicates that the noise estimate of 5% used in the synthetic case was not too high. However, comparing volumes estimates in section 3, the optimised inversion resulted in a larger volume of mapped fresh groundwater at the cost of a smaller volume of saline groundwater. This seemed to correlate with the optimised inversions also mapping the location of the fresh-brackish interface with greater accuracy (section 3), while most improvement in this regard was seen using the optimised inversion from the upper estimate – suggesting that the greater fresh groundwater volume estimate was valid. This improvement, however, came at the cost of reduced accuracy when resolving the brackish-saline interface, which indicates that spatially, this area is the largest source of error.

These results are similar to those by King et al., (2018), where four different inversion methods (Vignoli et al., 2015; Farquharson et al., 2003) were compared (i.e., not different parameters of the same inversion) using the same data as this study. King et al. (2018) found that the sharper style inversion methods mapped a greater volume of fresh groundwater, along with improved fresh-brackish interface mapping accuracy, at the cost of performing less favourably, overall compared to the LCI default inversion, when mapping the brackish-saline interface. This feature was noted for few-layer LCI inversions (where layer thicknesses as well as resistivities are permitted to change), LCI sharp (Vignoli et al., 2015) and the EM1DF (Farquharson et al., 2003) inversions method.

The results of this study indicate that to preserve the smoothly varying nature of fresh-brackish-saline gradients, while at the same time accurately mapping respective interfaces, other inversion methods should be considered. This could include coupled hydrogeophysical inversion (e.g. Herckenrath et al., 2013; Hinnell et al., 2010), where a groundwater simulation is forward modelled against AEM observations until a suitable fit is found – potentially preserving more realistic groundwater characteristics. This approach could be further developed to jointly match the in-situ brackish zone thickness and AEM data using inverse modelling. However, this process does suffer from uncertainty relating to petrophysical transformations and heterogeneity, which has unfortunately been shown to potentially affect the predicted thickness of the brackish zone (González-Quirós and Comte, 2020). Other possible approaches to use are those that include probabilistic inversion (Minsley et al., 2020), where not only an estimate of uncertainty can be obtained within the brackish zone, but also prior information can be included – such as cheaply available brackish zone thickness estimates.
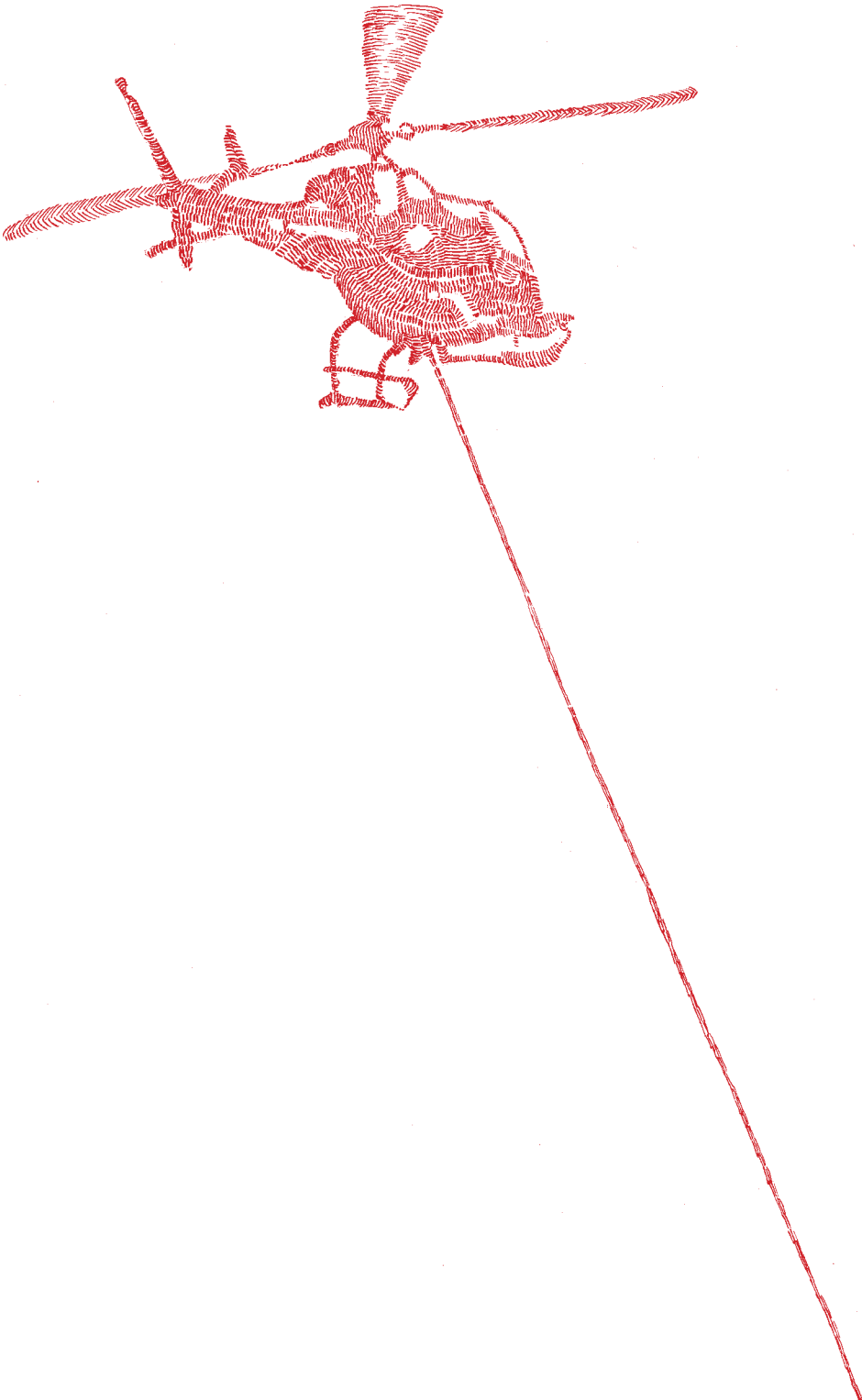
5

Joint estimation of groundwater salinity and hydrogeological parameters using variable-density groundwater flow, salt transport modelling and airborne electromagnetic surveys

Abstract

Freshwater aquifers in low elevation coastal zones are known to be threatened by saltwater intrusion (SWI). As these areas host a significant share of the world's population, an excellent understanding of this phenomenon is required to effectively manage the availability of freshwater. SWI is a dynamic process, therefore saline groundwater distributions can change quickly over time – particularly in stressed areas with anthropogenic drivers. To model these changes, regional 3D variable-density groundwater (3D-VDG) flow and coupled salt transport models are often used to estimate the current (and future distributions) of saline groundwater. Unfortunately, parameterising 3D-VDG models is a challenging task with many uncertainties. Generally, uncertainty is reduced through the addition of observational data – such as Airborne Electromagnetic (AEM) surveys or ground-based information – that offer information about parameters such as salinity and hydraulic head. Recent research has shown the ability of AEM surveys to provide accurate 3D groundwater salinity models across regional scales, as well as highlighting the potential for good survey repeatability. To this end we investigated the novel approach of using repeat AEM surveys (flown over the same area at different points in time) and 3D-VDG models to jointly improve the parameterisation of 3D-VDG models - while simultaneously providing a detailed 3D map of groundwater salinity distributions. Using detailed 3D synthetic models, the results of this study quantitatively highlight the usefulness of this approach, while offering practical information on implementation and further research.

5.1. Introduction

3D variable-density groundwater flow and coupled salt transport models – referred to in the following as 3D-VDG models – are proven and commonly used tools to simulate current and future distributions of saline groundwater (e.g. Faneca Sànchez et al., 2012; Meyer et al., 2019; Van Engelen et al., 2019). The usefulness of 3D-VDG models is particularly clear in highly populated low elevation coastal zones, where saltwater intrusion (SWI) into freshwater groundwater systems poses a significant risk to the availability of fresh water for agricultural, industrial and household uses (Gomaa et al., 2021; Oude Essink et al., 2010; Simmons et al., 2010; Werner et al., 2013). As a result, 3D-VDG models are extensively used to inform groundwater policies in these areas.

The parameterization of 3D-VDG models, i.e., finding values for ground water flow (e.g., hydraulic conductivity, storage coefficients) and salt transport (e.g., porosity, dispersivities) parameters is a daunting task. Particularly as these parameters can be very heterogeneous in space. Typically, the parameterization process starts with a hypothesis about parameter heterogeneity. This is followed by creating a spatial regionalization (e.g., zonation, geostatistical simulation) which reduces the number of free (unknown) parameters. Next, following an initial guess, the free parameters are estimated through model calibration. Calibration is often approached as an inverse problem, where model parameters are iteratively adjusted until the model response fits to observations of models states or outputs. For an overview of calibration approaches we refer to Zhou et al. (2014) and Doherty (2010), for an example of an application to 3D-VDG models refer to Carrera et al. (2010). Typically, 3D-VDG models are calibrated using in-situ head and salinity observations which are often sparse. Adjusted 3D-VDG parameters include, for example, hydraulic conductivity, porosity, dispersivity and recharge. However, 3D-VDG parameterization is prone to uncertainty. First, the large spatial heterogeneity of underground lithology creates errors during the spatial regionalization step. Second, calibration being an inverse problem usually suffers from non-uniqueness as a result of parameter correlation (Carrera et al., 2005), which arises from the sparseness of the in-situ observation used.

Naturally, one way to reduce uncertainty is the addition of more observational data that relate to observing 3D-VDG states (head, salinity) or outputs (e.g., stream discharge and concentration). Data sources are often ground-based, such as wells that monitor salinity and hydraulic head, or indirect geophysical methods such as Electrical Resistivity Tomography that measure electrical conductivity (EC) (Beaujean et al., 2014) which is converted to groundwater salinity estimates using petrophysical relationships. For regional 3D-VDG modelling, often extending across 100's of kms (e.g. Cobaner et al., 2012; Gossel et al., 2010; Mabrouk et al., 2019; Michael et al., 2009; Nocchi and Salleolini, 2013; Oude Essink et al., 2010; Van Engelen et al., 2019), 1D or 2D ground measurements are considered expensive – offering localized

and disconnected information (King et al., 2020b). Airborne Electromagnetics (AEM) offer an indirect, but fast and economical data source to overcome these shortcomings.

The AEM method is sensitive to both lithology and groundwater, where primarily clay content and groundwater salinity offer electrical conductivity (EC) contrasts (Revil and Glover, 1998). As a result, the use of AEM to resolve state variables, such as groundwater salinity, and parameters has proliferated in recent years. AEM methods have been used successfully to map structure such as clay content or lithological units (Auken et al., 2008; Foged et al., 2014; Gunnink et al., 2012; Høyer et al., 2015) and such could be used to help parameterize 3D-VDG models with additional lithological information. In more saline environments however, mapping structure (which can relate to parameters such as hydraulic conductivity) is more challenging because signals are dominated by the strong EC contrasts found at fresh-saline interfaces (King et al., 2020b). Naturally, this makes it highly suitable for mapping salinity instead, where in these environments the method has seen continued success for decades (Delsman et al., 2018; Faneca Sànchez et al., 2012; Fitterman and Deszcz-Pan, 2001; Jørgensen et al., 2012; Rahman et al., 2021; Vandevelde et al., 2018).

Conceptually, calibrating a 3D-VDG model with state variables obtained from geophysical data can be approached using sequential or coupled hydrogeophysical inversion (Herckenrath et al., 2013; A. C. Hinnell et al., 2010). For an overview and history of hydrogeophysical methods see Binley et al. (2015). In sequential methods a deterministic geophysical inversion is traditionally run first, and then translated into groundwater salinity using petrophysical relationships – allowing model calibration (e.g. Faneca Sànchez et al., 2012). Coupled hydrogeophysical inversion involves the transformation of 3D-VDG state variables into a physical property which can then be forward modelled and compared to geophysical observations (e.g. Bauer-gottwein et al., 2010; González-Quirós and Comte, 2021; Steklova and Haber, 2015). As a result, coupled models do not rely on a geophysical inversion. The 3D-VDG model can then be run iteratively until a misfit against geophysical observations are appropriately small.

In the following we present a method that exploits the ability of AEM to map groundwater salinity to improve the parameterization of 3D-VDG model properties, while at the same time providing improved 3D-VDG predictions of groundwater salinity. We use the idea that the distribution of saline and fresh groundwater changes over time and the resulting movement of EC contrasts between fresh and saline groundwater can be resolved by AEM if measured repeatedly at the same location. Furthermore, previous research suggests that if repeated flightline paths are spatially consistent and that inverted data are compared, rather than the electromagnetic response itself, then the AEM method offers good repeatability (Huang and Cogbill, 2006).

Thus, using a similar idea to traditional inverse calibration, where a model is run from the past to match observations at the present, we explore the possibility of using two separate AEM surveys flown at the same location at two different periods in time. Similar to time-lapse inversion that has been used extensively in reservoir engineering (Johnston, 2013) and hydrology (e.g. Karaoulis et al., 2011), we investigate the idea that the observed changes in salinity distributions can infer model 3D-VDG parameters. If a 3D-VDG model with initial groundwater salinity distributions obtained from the first survey, say at t0, it should be able to predict the groundwater salinity distribution at the time of the second AEM survey (say t1).

Given that groundwater in the Netherlands is generally a mixture of seawater and freshwater, chloride is the dominant conservative anion, therefore for consistency we will refer to a distribution of groundwater salinity as the chloride distribution. To test the idea and to see if such inverse estimation would be feasible in theory, we created a highly detailed synthetic reality (3D lithology and chloride distribution over time between t0 and t1) based on real AEM data (Delsman et al., 2018), a detailed lithological model (Stafleu et al., 2011) and 3D-VDG modelling with the computer code SEAWAT (Langevin et al., 2007; Verkaik et al., 2021) and a given set of "real" parameters. We then used a geophysical modelling approach (King et al., 2020b) to simulate an AEM survey at the start and end of the 3D-VDG model run. We subsequently tested an optimization strategy that iteratively runs the 3D-VDG model with unknown parameters between t0 and t1 while comparing this "modelled" chloride distribution with the AEM survey at t1 and adjust the unknown 3D-VDG model hydrogeological parameters to minimize the difference. Using this approach, it was tested if the original "real" parameters could be re-estimated correctly.

In the following we provide a summary of our approach in section 3.1. Section 3.2 describes how the synthetic model was created including the study area it was based on, followed by the 3D-VDG model set up and parameterization in 3.3 and the optimization method in 3.4. Section 4 outlines the results of the optimization, which are followed a discussion and finally concluding remarks in 5 and 6 respectively.

5.2. Methods

5.2.1 General approach
Our approach to estimate hydrogeological parameters is based on performing two sequentially processed AEM surveys, one at time t0 and one at time t1, that are sufficiently spaced apart to detect real changes in the chloride distribution of a groundwater body. This means that the approach is limited to groundwater systems where due to external forcing (extraction fresh groundwater, injecting fresh water, sea-level rise) the chloride distribution changes relatively rapidly. The estimated chloride distribution obtained from the survey at time t0 is used as initial

condition by a 3D-VDG model that is subsequently used to simulate the chloride distribution at time t1. This simulation is first done with an initial estimate of the unknown hydrogeological parameters. We then perform a simulated AEM survey on the 3D-VDG model simulated chloride distribution to obtain an AEM chloride distribution at the same (lower) spatial footprint as the actual AEM survey at t1 (see Figure 5.2). The differences between the actual and simulated AEM chloride distributions are then used to drive an optimization framework, as outlined in Figure 5.1. Here, the unknown hydrogeological parameters are adjusted. The 3D-VDG model and AEM simulation and parameter estimation steps are repeated iteratively while minimizing the differences between the actual and simulated AEM at t1. In effect, we are applying a coupled, optimization driven time-lapse hydrogeophysical inversion. The result of this approach are optimized estimates of the unknown hydrogeological parameters and an optimized estimate of the chloride distribution at the resolution of the 3D-VDG model (generally higher than that of an AEM survey) which is also consistent with physical laws.
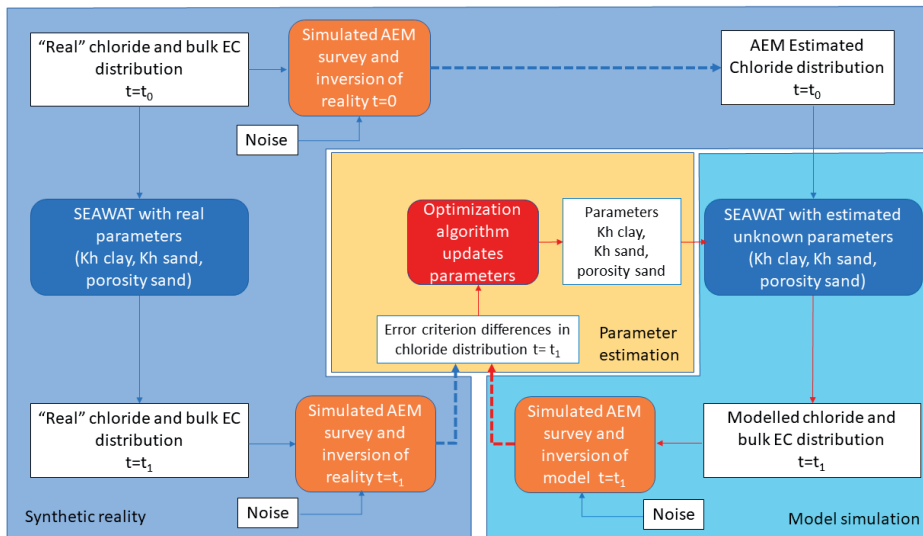
Figure 5.1. *Method outline. Dark blue region (left): creating the synthetic reality, light blue region (right): iterating over simulated models with given parameters, yellow region (centre): estimating subsurface parameters. White squares: 'real' 3D chloride distributions in time, orange squares: geophysical simulations of 'real' 3D chloride distributions, blue squares: 3D-VDG model run with given parameters, red square: parameterization update (driven by optimization strategy). Dashed lines indicate a petrophysical transformation from inverted ECb to chloride.*

At the moment, to our knowledge, regional-scale time-lapse AEM surveys are non-existent, but to be expected in the near future. Thus, to demonstrate the methodology, we used a so-called twin-experiment where we use a realistic synthetic reality and simulated AEM surveys. The complete setup of the twin experiment is shown in Figure 5.1: a) the synthetic reality which consists of a realistic chloride distribution at t0 and t1 simulated with the 3D-VDG model using the "real hydrogeological parameters" and two simulated AEM surveys (in blue);

b) the model simulation (also using the 3D-VDG model) and 3D-VDG model based AEM survey used in the optimization (light blue); and c) the optimization algorithm (orange). We note that the simulated AEM surveys include realistic observation noise and degrading of resolution conform the AEM footprints (King et al., 2020). Using the same 3D-VDG model for reality and model simulation means that we ignore model structural errors.

In the following we describe in more detail: a) the synthetic case, including the simulated AEM in section 5.2.2; b) the model simulation and initial hydrogeological parameter estimates in section 5.2.3; c) optimization routine in section 5.2.4.

5.2.2 Creating a synthetic case

The synthetic case was generated from 3D estimates from an existing AEM survey from the Province of Zeeland in The Netherlands (Delsman et al., 2018). Within this region, data was extracted from an area called Zeeuws-Vlaanderen (or Zeelandic Flanders in English), in southern Zeeland. This east-west oriented strip (~60 x 15 km) is bordered to the North by the Scheldt River estuary and to the south by the Belgian border. At the centre of Zeeuws-Vlaanderen, around the town of Terneuzen, a subset area was selected covering ~7 x 7 km for further processing. The model area and AEM flightlines are highlighted in Figure 5.2.
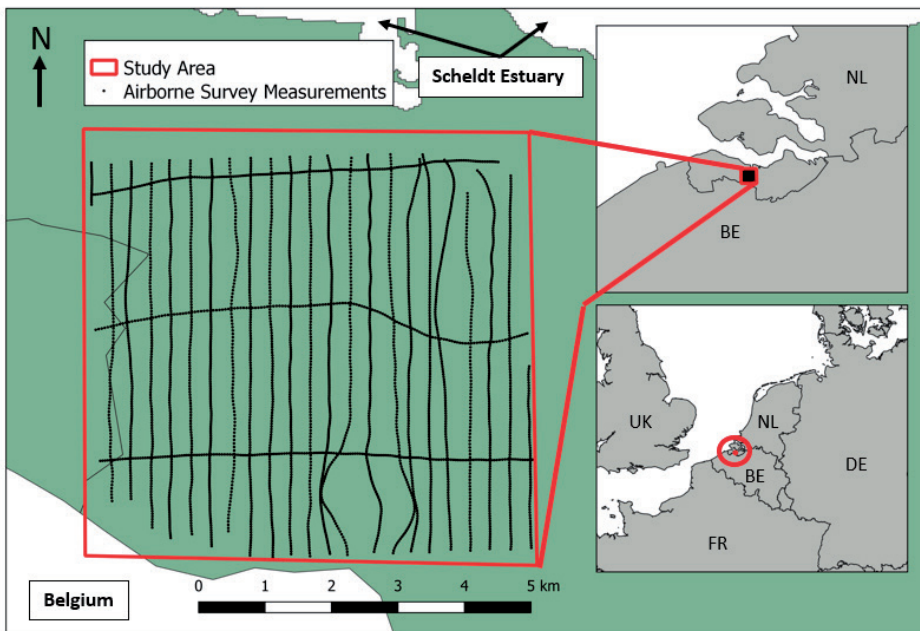
Figure 5.2. *Location of the case study used to create the synthetic case. The flight lines shown are the same as in the original AEM survey.*

The groundwater system characteristics of the area are summarized by the presence of shallow freshwater lenses typically about 10 – 20 m thick (Delsman et al., 2018), caused by Holocene sea-level transgressions and subsequent land reclamation (Berendsen, 2005; Delsman et al., 2014). The hydrogeological system is hosted within lithologies comprising Neogene and Quaternary sediments (Stafleu et al., 2011), including younger fine sands, clays and peats, with deeper northward dipping sand and silts (Vos, 2015). Within the study area, the dipping hydrogeological base extends to ~50 – 70 m below the surface and denotes an impermeable aquitard. To create a synthetic reality, we use the existing detailed 3D chloride distribution that was obtained from the AEM survey performed over the Province of Zeeland (Delsman et al., 2018). This chloride distribution was taken as the "real chloride distribution" at t0. For this area, a very detailed 3D-lithological characterization is available (Stafleu et al., 2011), based on ~1500 borelogs in the study area (see Figure 5.3). In our synthetic example, we assume the lithology to be fully known and focus on estimating the chloride distribution and hydrogeological parameters only.

5.2.2a Translation of chloride to electrical conductivity
The chloride distribution and lithological information were further processed into bulk electrical conductivity (ECb) to allow for the simulation of an AEM survey and subsequent inversion. First, a known empirical relationship from the study area was used to transform the data into the electrical conductivity of the groundwater (ECw) and corrected from a reference groundwater temperature of 25 *o*C to 11 *o*C, as outlined in De Louw et al. (2011). As the ECb is a product of both lithology and pore water salinity, a petrophysical transformation is required to split the information between the two. The simplest approach is using apparent formation factors (FF) which are the ratio of saturated sediment (ECb) to that of the pore water itself (ECw) (Archie, 1942). The electrical properties of clays minerals such as surface conductivity complicate this relationship (Revil et al., 2017). As a result, a more accurate transformation in areas with clay present requires the addition of surface conductivity, such as Waxman and Smits (2003). In this study we decided to use the simpler, apparent FFs (ECb/ECw) – values are shown in Table 5.1. The values for apparent FFs were taken from over 500 samples in an area only 20 km away from the area that was used to create our synthetic model, and therefore offered good representative values.

The 3D distribution of lithological units to derive the FF distribution is the one taken from GeoTOP (Stafleu et al., 2011) (Figure 5.3) and corresponding FF values from nearby field-measurements, as shown in table 5.1. (De Louw et al., 2011).

Table 5.1. *Formation factors used per sediment class for the conversion of ECb data to chloride.*

Lithology	Formation factor (FF)	Std	Samples
Peat	2.1	0.7	41
Clay	2.5	0.6	192
Sandy clay	2.8	0.8	52
Fine sand	3.2	0.4	299
Medium sand*	4		
Coarse sand*	5		

* FF taken from another study (Goes et al., 2009), without uncertainty and sample numbers.

5.2.2b Obtaining the chloride distribution at t1 using a 3D-VDG model

To set up a 3D-VDG model that produces a realistic change in chloride distributions at t*1*, we constructed a model that comprises a combination of typical anthropogenic and natural drivers that result in chloride movements – while at the same time ensuring that the model changes enough to produce sufficient signal. Natural drivers are natural recharge and autonomous freshening and salinization that occurs as a result of past inundation or sea level change (Vos, 2015). We also introduced anthropogenic drivers: extensive groundwater extraction that results in the so-called upconing (or shallowing) of brackish to saline groundwater and areas with enhanced recharge of fresh surface water through Aquifer Storage and Recovery (ASR) (Dillon, 2005), which would likely result in a rapid increase of fresh groundwater volumes (Pauw et al., 2015b; Zuurbier et al., 2015).
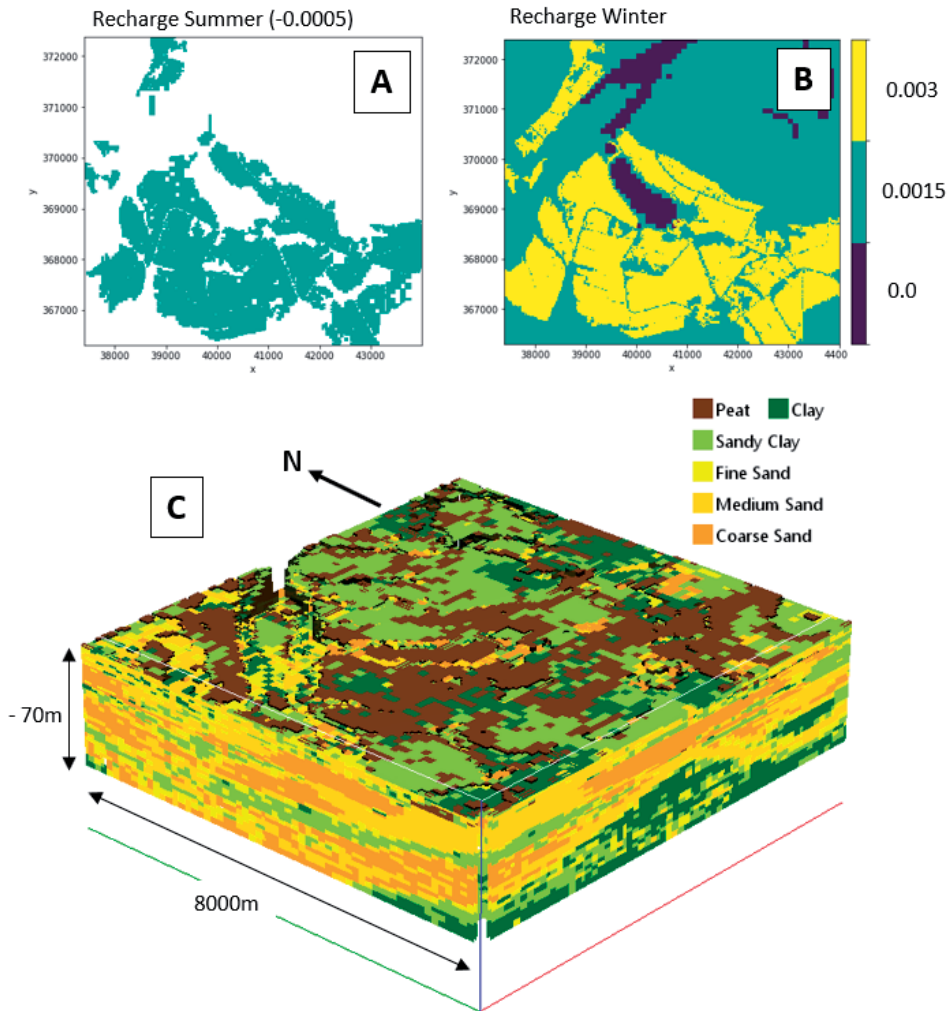
Figure 5.3. *3D-VDG model set up highlighting areas of extraction, represented as summer recharge (A) with enhanced recharge in the extraction areas (B), both as m/day. The lithological model (C) that was used to assign hydraulic conductivities, shown with 25 x vertical exaggeration.*

With this in mind, a 3D-VDG model was discretized into 25 x 25 m horizontal cells across a ~8 x 8 km area and comprised a 0.5 m vertical resolution at shallow depths, increasing in thickness logarithmically with depth to a maximum thickness of 10 m, resulting ~3.5 million active cells over 54 layers. A hydrogeological base was assigned at between 50 and 70 m depth and was derived from the GeoTOP model and REGIS (Vernes et al., 2010). The hydrogeological parameters of the model were similarly taken from the 3D lithology distribution of GeoTOP, assigning the parameter values from Table 5.2. Here, lithology classes

coarse, medium, fine loamy and silty sand were assumed as aquifer material and the finer and organic classes as aquitard material. Well locations were selected based on three conditions, in all model cells that are: (1) at 8 m depth (below surface), (2) within aquifer lithologies, and (3) within freshwater areas (<0.5 g/l chloride).

Table 5.2. *The parameters chosen for the 3D-VDG model*

Parameter	Description	Value (unit)
Kh **Aquifer**	*Horizontal hydraulic conductivity of the aquifer (m/day)*	10 (m/day)
Kh **Aquitard**	*Horizontal hydraulic conductivity of the aquitard (m/day)*	0.01 (m/day)
Kh/Kv	*Anisotropy*	3.3 (aquifer) 2 (aquitard)
Porosity	*Porosity*	35 (%)
Recharge winter	*Recharge in winter, higher values denote ASR areas*	0.003 m/day (ASR areas), 0.0015 m/day (other areas)
Recharge summer	*Recharge in summer, negative value denotes evaporation*	-0.0005 m/day
Well Extraction winter	*Groundwater extraction in winter*	0 m3/day per model cell
Well Extraction summer	*Groundwater extraction in summer*	-0.625 m3/day per model cell

To understand how an AEM survey might respond to groundwater chloride movements across a broad range of time-scales, the model was run for 60 years, longer than was thought necessary for this study. Finally, at every consecutive 6 month time-step, we modelled the geophysical response of the AEM system. In this study, we conclude that after 15 years there was likely sufficient signal available for the AEM system based on a known noise threshold of 5%, (see the Results section).

5.2.2c Simulating an AEM Survey
As highlighted in figure 5.1, to recover a 3D chloride distribution that has the same resolution and physical characteristics of an actual survey at $t0$ and $t1$, an AEM survey needed to be simulated. The process is similar to (King et al., 2020) and is summarized in the following. Existing measurement locations and flightlines (typically with a spacing between 30 and 60 m) were taken from the survey that produced the "real" 3D chloride distribution (section 2.2.1). ECb values, which are obtained using the method explained in section 2.2.1, were then sampled at these points every ~50 m to facilitate fast inversion times while still honouring the 50 m horizontal resolution of the 3D model. Vertical sampling was done at 0.5 m intervals. The horizontal resolution, or footprint of the system, is approximated as a 100 – 200 m diameter circles directly beneath the towed instrument (Reid et al., 2006; Yin et al., 2014). Therefore, at each measurement location, the nearest two model cells were averaged to imitate a 100 m footprint. Data were forward modelled using AarhusINV (Auken et al., 2005). Coil source and receiver spacing, and orientations were selected based on the values used by Fugro's RESOLVE HEM system during the original survey (Delsman et al., 2018). Finally, to approximate the noise levels present during FEM acquisition (Farquharson et al., 2003; Green and Lane, 2013), 5% white noise was added to the forward modelled data.

To recover a distribution of electrical properties from an AEM survey, a geophysical inversion is typically undertaken. For this study, a deterministic geophysical inversion (e.g., Auken et al., 2015; Farquharson et al., 2003; Viezzoli et al., 2009; Vignoli et al., 2015) (referred to in the following as an inversion) using AarhusINV (Auken et al., 2015) was implemented. This inversion approach uses a pseudo-2D laterally constrained approach where neighboring observations are constrained in the regularization process, helping to produce laterally coherent results. The method has been successfully used in many hydrogeophysical studies (e.g. Auken et al., 2008; Chongo et al., 2015; Delsman et al., 2018; Kirkegaard et al., 2011) and is particularly useful in areas with laterally continuous EC contrasts such as our study area.

The inversion starting model consisted of 20 fixed layers, with the first set to 0.5 m thick, increasing logarithmically till 50 m. Below this, the final layer is assumed to extend infinitely. Given the highly saline environment, the model was assigned a starting ECb value of 1 S/m. The lateral and vertical constraint parameters were set to 1.3 and 3 respectively, following recommended and commonly used values (Auken et al., 2005). This minimum-structure style inversion changes ECb only and has been found to accurately reproduce smoothly varying chloride distributions (King et al., 2018). The depth of investigation (DOI) varies according to the EC of the subsurface; in this case by shallow saline areas, and therefore could range from ~5 to 60 m in saline or fresh groundwater areas, respectively. The inversion output has an estimation of DOI for each 1D model location. The method to calculate this procedure is described in Vest Christiansen and Auken (2012). All inversions converged to a misfit of less than 5%, thereby effectively explaining the forward modelled observations.

Finally, as a 3D initial groundwater chloride distribution is required for the 3D-VDG model and the inversion method only produces 1D vertical profiles along flightlines, a 3D interpolation step was required. This was approached using a simple method that exploits the layered nature of the inverted data to efficiently produce 3D ECb models from the inversion results. Each one of the 20 inversion layers for all flightlines were interpolated separately using 2D kriging with automatic variogram modelling. The resulting 2D layers were then compiled to produce a 3D model. The method is described in (King et al., 2018) and has been shown to accurately produce 3D volumes of ECb in similar settings. As a final step, the DOI estimates were interpolated using the same 2D Kriging technique, where ECb values below the DOI were filled to the model base using the last encountered ECb value at that depth. Finally, the resulting ECb volume was transformed back into chloride using the reverse of the approach described in section 2.2a, resulting in simulated AEM surveyed chloride distributions at t*0* and t*1*.

5.2.3. 3D-VDG model simulation and initial parameter estimates
The 3D-VDG model simulation was run iteratively for a 15-year period using the exact same discretization and general parameterization as the model described in section 2.2.2, except for

the unknown parameters. Following extensive experimentation, we conclude that the method was to be used to predict the hydraulic conductivities of the aquifer and aquitard, and porosity. The reasoning behind this choice was two-fold: (1) it was found that the 3D-VDG model was most sensitive to these parameters, and (2) we had to restrict the number of parameters to avoid excessive computation times and identification issues. The ratio of horizontal to vertical conductivity Kh/Kv was kept constant, and thus in effect the optimization was used to resolve three parameters. As this study is unique and therefore a proof of concept, it was felt that while more parameters could be possible, it was important to keep the approach simple. Given that especially higher Kh and Kv values result in longer SEAWAT runtimes, initial estimates were chosen to be too low on purpose to facilitate faster iteration times early on. To test the sensitivity of the optimization to incorrect initial estimates, the starting values were chosen to be ten-times greater than the actual values. Effective porosity was simply assigned by a random guess to test the general robustness of the method. Values used are highlighted in Table 5.3.

Table 5.3. *Initial parameter estimates used in the first iteration.*

Parameter	Initial Estimate (m/day)	Actual Value (m/day)
Kh Aquifer	1	10
Kv Aquifer	0.3	3
Kh Aquitard	0.001	0.01
Kv Aquitard	0.0005	0.005
Effective porosity	0.1	0.35

5.2.4. The optimization

We used a suitable non-gradient global optimizer that works well in the high-dimensional problems typically found in the field of groundwater modeling. For this, the Nelder-Mead downhill-simplex method was selected for its robustness and proven ability to optimize model parameters and find global minima (Nelder and Mead, 1965). Its ability to handle noise is also desirable given the inherent noisiness related to the acquisition of AEM data. The optimization was implemented using the SciPy Optimize package available in Python, that uses an improvement to the original method by adjusting the simplex parameters relative to the dimensionality of the problem (Gao and Han, 2012).

The objective function returns the root mean squared error (RMSE) between t*1* survey and t*1* predicted, given by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(y_{pred,i} - y_{ref,i}\right)^2}{N}} \tag{1}$$

where *pred* at t*1* denote simulated AEM surveyed chloride values based on the 3D-VDG model predictions (with unknown parameters) and *ref* at t*1* the simulated AEM surveyed

chloride values based on the synthetic reality (3D-VDG model with known parameters). This was minimized until a convergence criterion was met, in this case given by the RMSE between iterations. This value was set to 0.1 g/l chloride and was chosen to be unrealistically low which effectively allowed the optimization to run indefinitely to allow a user-based decision on whether it has converged sufficiently. In practice, the convergence criteria could however be set to reflect the levels of noise from the AEM system if known.

5.3. Results

5.3.1 AEM sensitivity to estimated chloride distributions

The 3D-VDG model was first run for 60 years to understand the length of time needed for chloride changes to be detected by the AEM system at each time-step, and indeed the sensitivity of the AEM method to the transient effects in general. To understand this quantitatively, each time step of the resulting 3D-VDG model was geophysically forward modelled using the method outlined in section 2, where successive time-steps were compared quantitatively to the first model as relative mean absolute error (%). The 5% black horizontal line in Figure 5.3 represents an approximation of AEM noise levels thereby providing a rough indication of required groundwater chloride movements before there is enough signal for a second survey to measure meaningful differences.
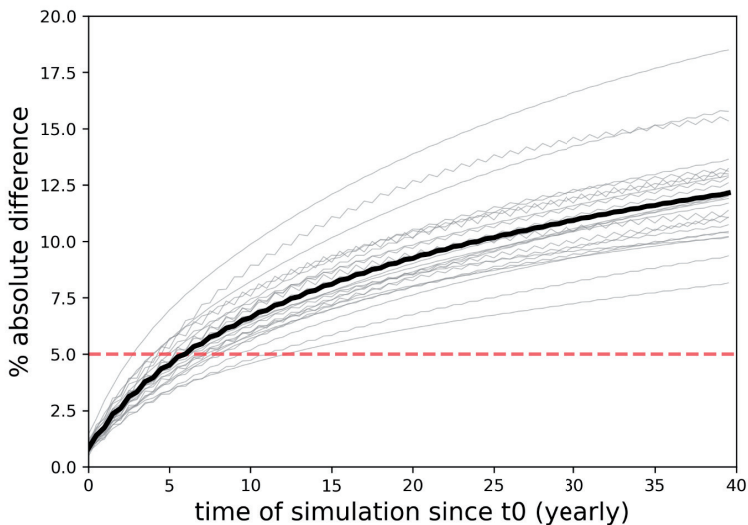
Figure 5.4. *Sensitivity of the AEM system over time, represented per flightline (light grey lines) and averaged across all flightlines (red line). Sensitivities are represented as the mean absolute difference of all five frequencies of the acquisition system over time, including both in-phase and quadrature components from the AEM acquisition system for each flightline.*

From Figure 5.4, it follows that prior to ~5 years a repeat survey would likely fall within the estimated noise range, indicating that a repeat survey in this instance would fail to effectively map differences. After 15 years (or 30 time-steps) the slope gradient decreases after an initial sharp increase to ~7.5% difference. Therefore, in the following the time-period between zero and 15 years was used to simulate the two surveys and ultimately guide the optimization. In a real setting, localized ground measurements over time and geophysical forward modelling could be used to estimate if there is likely to be enough signal for a repeat AEM to be useful.

5.3.2 Estimated Parameters

The optimization ran for 250 iterations for 15 (3D-VDG simulated) years per iteration, taking on average five hours per iteration and ~2 months to complete on a standard desktop PC with four processing cores. The 3D-VDG model simulation was by far the most computationally intensive step, with the geophysical forward modelling, inversion, 3D interpolation and petrophysical transforms step only taking around half an hour in total per iteration.

The optimization results at each iteration are highlighted in Figure 5.4., as the Kh/Kv ratio was kept constant only the Kh values are shown for the aquifers and aquitards.
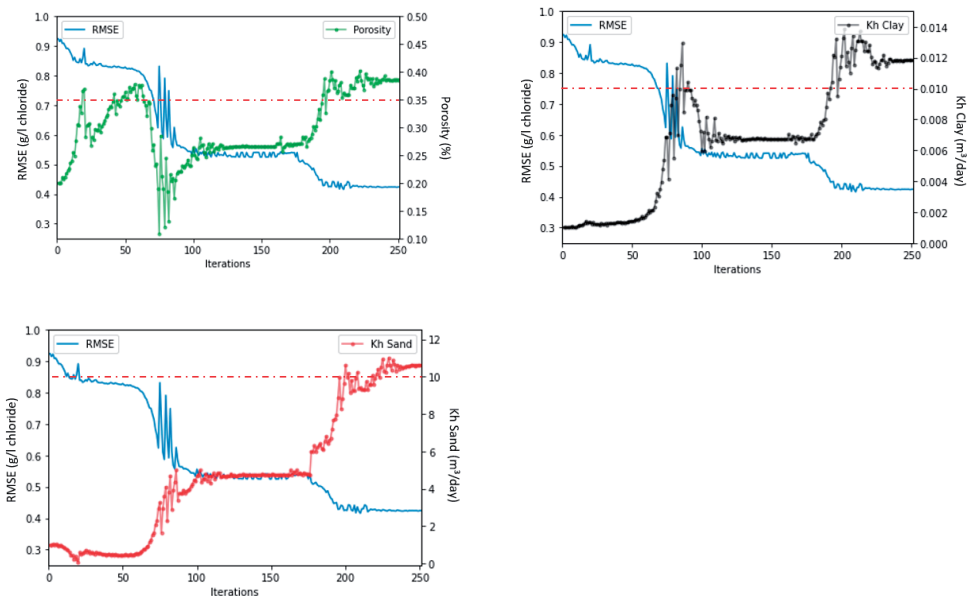
Figure 5.4. *RMSE (in g Cl-/l) for each predicted model parameter over each iteration. Top left: predicted Kh clay (or aquitard) in m3/day. Top right: predicted Kh cand (or aquifer) in m/day. Bottom left: predicted porosity (%). Red line indicates the real, target value.*

Here it is apparent that there were two local minima: (1) between iterations 20 – 50 and (2) between 100 – 170, both of which were successfully avoided by the algorithm. Predictably changes in RMSE correlated well to changes in model parameters and showed that the steps between each iteration were appropriately sized and generally sensitive to the objective function. This is despite the fact that smaller changes in chloride distributions were *not* resolved as a result of the AEM simulation, the effect of which is highlighted in Figure 5.3.

Quantitative results based on the final iteration (Table 5.4), suggest that each of the parameters were either successfully predicted or were improved considerably over the course of the optimization.

Table 5.4. *The optimization results showing the values of initial, actual and predicted model parameters.*

Parameter	Initial Estimate	Actual Value	Predicted Value (difference actual)
Kh Aquifer (m/day)	1	10	10.63291 (0.63291)
Kv Aquifer (m/day)	0.3	3	3.312433 (0.312433)
Kh Aquitard (m/day)	0.001	0.01	0.011838 (0.001838)
Kv Aquitard (m/day)	0.0005	0.005	0.005831 (0.000831)
Porosity	0.1	0.35	0.386181 (0.036181)

5.3.3 Estimated chloride distributions

Qualitative results of chloride distributions at t1 are presented in Figure 5.5, showing the synthetic reality on row 1.a. and 1.b. for the AEM salinity and 3D-VDG models respectively, followed by the results of the optimization, with the results of the first iteration on row 2, the last iteration on row 3 and the difference between the last iteration and the synthetic reality in row 4.

Comparing the chloride distributions between the AEM and the 3D-VDG models in general (columns 'a' and 'b' respectively), it is clear that the AEM simulation successfully mapped the locations of larger features such as upconing of saline groundwater and freshwater lenses. Smaller features (~< 200 m across) were not successfully resolved given the footprint of the AEM system and the loss of resolution during the 3D interpolation of ECb values. In these areas the features have been smoothed out – effectively creating a vertically thicker brackish zone instead. It was also observed that the optimization successfully managed to recreate the synthetic realities' chloride distribution by the final iteration. The difference between the AEM synthetic reality (Figure 5. 1.a.) and the AEM final iteration (Figure 5. 3.a.) is larger than the differences between that of the 3D-VDG models. This is likely due to the 5% noise added in the geophysical modelling step and appears to cause uncertainty regarding the location of the fresh-saline interface. Despite this error, the 3D-VDG model was able to dynamically downscale the smoothed features observed in the AEM surveys.
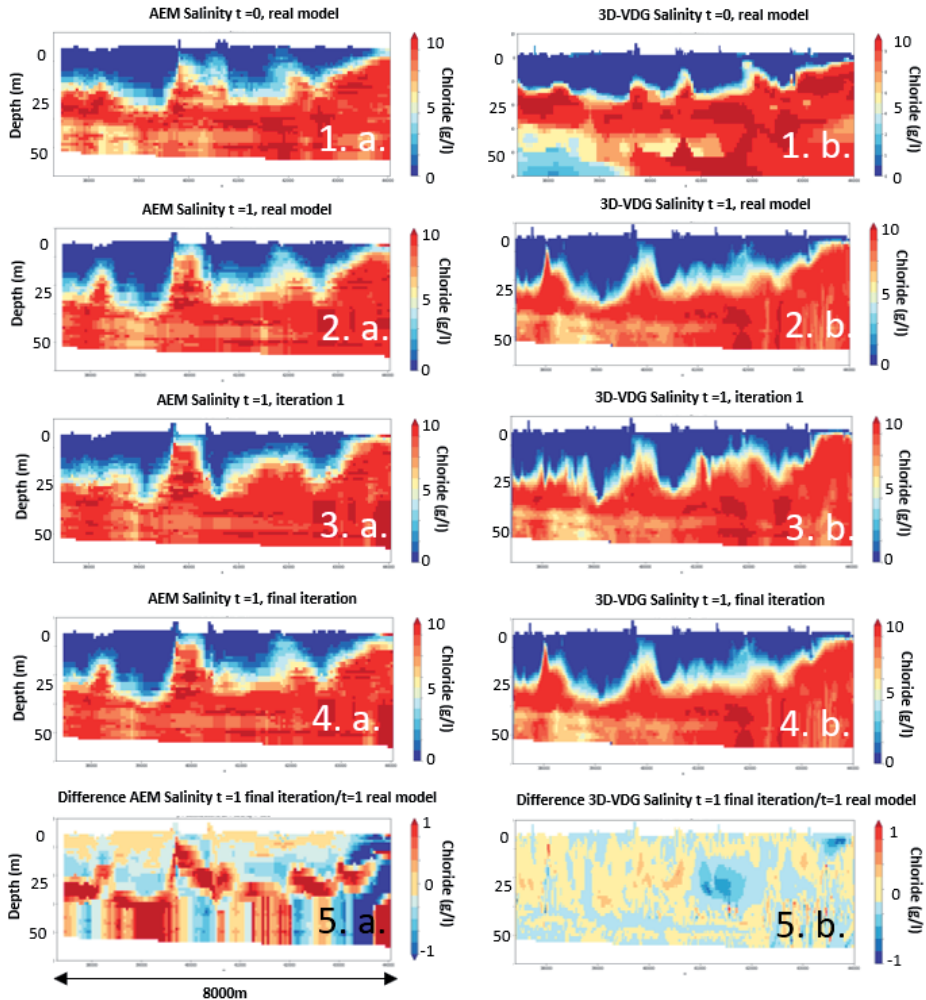
Figure 5.5. *AEM (left column) and 3D-VDG (right column) model results at t1 and initial salinity at t0. Rows 1 and 2 illustrate the 'real' model salinities at t0 and t1 respectively, therefore show our synthetic realities. Row 3 is the optimization result at the first iteration and the 4th row is the final optimization result at the final iteration. The last (5th) row is the difference between the final iteration and the synthetic reality. The section location is shown in figure 5.6 with the red line.*

Figure 5.6 shows a 3D image of the final estimated AEM chloride distribution and the differences between the initial and final iteration thereof. The observed differences in salinity clearly show the zones where the chloride distribution was updated by the calibration process.
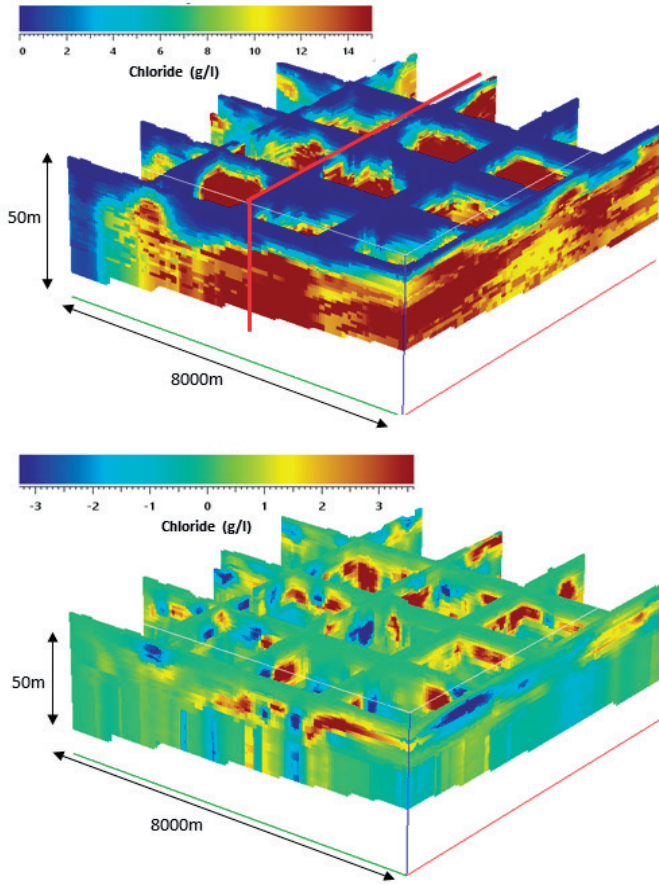
Figure 5.6. *3D view of estimated simulated chloride distributions. Top: Final estimated AEM chloride distribution. Bottom: Differences between the first and final iteration AEM chloride distributions, highlighted zones that were updated by the optimization process.*

5.4. Discussion

Given the exploratory nature of this study, several necessary simplifications were made. As a result of these simplifications, in the following we will first discuss sources of uncertainty and their potential effects on our results, followed by a discussion of the optimization method. Finally, we mention thoughts on practicalities followed by potential for further research.

In this twin-experiment, we assumed that the synthetic model correctly describes reality – this assumption applies to the physics of the simulated AEM survey itself, as well as the magnitude of groundwater transport over time based on the 3D-VDG model. In the case

of the AEM survey, we assumed that the chloride distributions were mapped with similar physical limitations to an actual survey. A major cause of this uncertainty is that we used a 1D geophysical forward and inverse modelling tool for a 3D problem (Auken et al., 2005), effectively removing 3D effects on the signal. Although there are 3D AEM geophysical modelling tools available (e.g., Cox et al., 2012), it was found that using these tools would lead to considerably longer iteration times on an already computationally expensive task. Instead, we added 5% gaussian noise which is considered an upper estimate of noise (Siemon, 2012) as well as simulating a 100 m footprint at measurement location (Reid et al., 2006; Yin et al., 2014). These steps effectively tested the robustness of repeat AEM surveys while maintaining reasonable optimization run-times.

For the 3D-VDG model simulation, we sought a careful balance between a realistic model while ensuring that there was sufficient signal for the optimization to work with. Based on a combination of using real subsurface data and known parameters, which were taken from an applied study (Mulder et al., 2020) as well as the groundwater modelling experience of the authors – we feel that the model describes reality to an acceptable level for meaningful conclusions.

An additional source of uncertainty is that we assume knowledge of parameter values and locations of the other hydrogeological parameters, such as the interaction between groundwater and drainage via the so-called drain resistance and the locations of aquitards and aquifers. We suggest that a simple way to examine the effect of keeping unknown parameter values constant would be to add noise to those that are not part of the optimization – where permitted deviations could be based on known uncertainty levels. The same could be said for the spatial distribution of clay and sand units, allowing the locations thereof to change based on given noise thresholds. The results of implementing this could then be used to determine quantitively the effect of these assumptions. Ideally, all uncertainties should have been included in the optimization as a fully heterogeneous 3D model. In this study, however, we chose to keep our method simple owing in part to the already high computational burden (discussed below) as well the fact that this is presented as a proof-of-concept – thus further complexity is considered subject for further study.

Besides correctly describing reality – uncertainty exists in both the type of geophysical inversion used, as well as the petrophysical transformation (i.e. to transform from ECb to salinity and vice-versa). Recent research highlights this, where it was found that the petrophysical transformation introduced the most uncertainty – and that overall incorrect handling of these two features can result in the mapping of an overly thick brackish zone (González-Quirós and Comte, 2020). The deterministic method used here suffers from non-uniqueness – where an infinite number of models can explain the data. In these cases regularization constrains the inversion, typically resulting in smooth models (Arsenin, 1979; Constable, 1987). One way to

explore this effect is stochastic inversion, where a prior model is used instead of regularization (e.g. Minsley, 2011; Minsley et al., 2020). In this study we used a single inversion method as a necessity, given the computational burden of stochastic inversion. Furthermore, in this synthetic study we assume that the lithology is known – and that we estimate the properties of only the high and low permeability sediments. In real applications the efficacy of the AEM system for mapping similar shallow fresh-saline groundwater distributions, using the same geophysical inversion method, was quantitatively validated recently in Delsman et al. (2018). As a result, we feel that on balance our approach was valid for this study.

With the petrophysical model, we settled on the simpler apparent FF without using surface conductivities for a number of reasons. First, these values were derived from an actual, local field-study, and thus offered a good, realistic approximation for the study area. Second, lithologies such as peat and clay that could introduce error were only present as 1.4 and 6.2% of the total model volume respectively. As a result, we felt that additional uncertainty introduced when finding appropriate values for surface conductivity negated the use of local, representative data. It is however recommended that this step is taken with care in other use cases, given the propensity for incorrectly used petrophysical models to introduce non-physical error into hydrogeophysical models (González-Quirós and Comte, 2020). The ECb gradient is strongly controlled by fresh-saline groundwater contrasts – an effect also observed clearly in local ECb ground measurements. As such, despite potential error introduced in the petrophysical transformation and inversion, the relative magnitude of signal changes over time were considered sufficient to drive the optimization and were a good enough approximation for our purposes.

Overall, as the time-lapse approach presented here is new, we recommend that future research focusses on a global uncertainty analysis. Having a better quantitative grasp on what this means for practical mapping outcomes would be a useful step and potentially highlight methods to reduce uncertainty, for example the use of in-situ salinity data. For the purposes of this study however, we feel that the potential use of repeated airborne surveys was effectively proven given the highly realistic nature of our 3-D synthetic model.

The downhill-simplex optimization method used here (Gao and Han, 2012; Nelder and Mead, 1965) was selected for its simplicity and ability to handle highly-dimensional problems. Despite the successes of using the method in this study, in a practical (rather than exploratory) setting – we suggest that other methods are considered. This is recommended for two reasons: (1) the optimization was impractically slow to find a solution, and (2) the addition of other parameters would increase the likelihood of the downhill-simplex arriving at a local-minimum, especially given its sensitivity to the choice of initial guesses (Wang and Shoup, 2011). As the 3D-VDG model computation step cost ~5 hours per iteration, one *could* simply run this step externally on large computational clusters using iMOD-SEAWAT, which

utilizes distributed memory parallelization for faster computation times (Verkaik et al., 2021). In a practical setting however, we suggest fully parallelizing the optimization itself, using methods such as evolutionary algorithms (Brauer et al., 2002a; Mühlenbein et al., 1991a) and parallel Bayesian optimization (e.g. González et al., 2015; Kandasamy et al., 2017) where function evaluations can be done in parallel rather than sequentially. The latter approach has been successfully used in hydrocarbon reservoir modelling (e.g. Abdollahzadeh et al., 2011), where similar to the optimization used in this study, inverse problems are solved that require multiple flow-simulations, which are more computationally expensive compared to our problem. We consider the implementation of this an important avenue for future research given the possibility to parameterize fully heterogeneous models.

Finally, while the results of this research highlight the usefulness of two AEM surveys in time, we find it necessary to discuss these results against 3D hydrogeological inverse models that utilize different data configurations – for example a single AEM survey, or even multiple sources of in-situ head and salinity data.

Unlike this study, in these cases there isn't a complete picture of a groundwater state at two time intervals in 3D. As a result, an inverse modelling routine relies on a conceptualized version of the past based on prior knowledge, followed by a history matching routine until the chloride distribution is matched with acceptable accuracy to available data. This has the disadvantage that, owing to the large inertia of fresh-salt groundwater systems, simulations have to start far back in the past, sometimes hundreds to thousands of years back (Meyer et al., 2019; Delsman et al., 2014). Apart from being highly computationally expensive (in practice requiring) massively parallel computing, it also requires knowledge of boundary conditions from the distance past, adding considerable uncertainty and requiring extensive research. Nevertheless, we believe that a quantitative comparison between history matching to a single AEM survey and optimizing between two in a similar experiment would be useful as future research.

5.5. Conclusions

Using a twin-experiment involving two highly detailed synthetic realities and simulated airborne electromagnetic (AEM) surveys, it was successfully demonstrated that it is possible to jointly estimate groundwater salinity and hydrogeological parameters. In a real application, this could be achieved using a combination of two AEM surveys (flown across the same area over two periods in time) and a variable density groundwater flow and salt transport (3D-VDG) model. By coupling AEM and 3D-VDG models using geophysical forward modelling methods, it was shown that two AEM surveys are sensitive to changing groundwater chloride distributions over time - despite a significant loss of resolution as a

result of the survey process itself. In doing so, it was also demonstrated that the time needed for chloride movements to be sensitive to AEM can be attained quantitatively. Given the knowledge of sensitivities, it followed that an optimization method could be used to recover hydrogeological properties using 3D-VDG. With this in mind, a simple optimization process was implemented to resolve hydrogeological parameters of interest, while at the same time producing a physically consistent estimate of the chloride distribution at higher spatial resolution then the AEM surveys.

5.6 Acknowledgements

6

Synthesis

6.1 Introduction

The objectives of the research presented here were twofold: **1) to quantitatively understand the uncertainties of regional scale groundwater salinity mapping using Airborne Electromagnetic AEM methods, and 2) to use this knowledge to guide the development of novel mapping approaches.**

To reach these objectives, data acquired from the Province of Zeeland, the Netherlands were used (Delsman et al., 2018). These high-quality data were ideal for the objectives of this thesis, as the study area is known as a highly saline groundwater area. Furthermore, the data provided a framework for the creation of highly detailed synthetic models to test the development of novel mapping methods.

The synthesis that follows effectively answers the research questions posed in chapter 1. These questions are addressed in detail in chapters 2 to 5. Based on the conclusions of this research, this chapter finishes with suggestions for further research.

6.2 Research Questions

6.2.1 What is the effect of using different inversion methods and parameters on mapping results? (Chapter 2)

Geophysical inversion (or simply inversion in the following) is the process of converting AEM observations into usable data in the form of electrical conductivity (EC) at depth. Inversion adds great uncertainty to the EC mapping as, in principle, the inversion problem is ill-posed: infinite possible models can explain the data – as such there are several methods available for this purpose (e.g., Auken and Christiansen, 2004; Brodie and Sambridge, 2006; Farquharson et al., 2003). In order to counteract the uncertainty due to this non-uniqueness, these methods use mathematical constraints to stabilise results – as a result each method also requires user-based parameter inputs. Besides qualitative studies (Hodges and Siemon, 2008), a fully quantitative understanding of the effects of deterministic inversion methods on mapping groundwater salinity had not been undertaken. Generally, AEM groundwater studies utilise a single inversion method that is subsequently validated using often sparse ground (i.e. in-situ) data (e.g., Auken et al., 2005; Farquharson et al., 2003; Siemon et al., 2009).

To quantitatively understand the effect of different inversion methods and different parameter settings for mathematical constraints on groundwater salinity mapping, available AEM data were inverted over a highly saline area, along with extensively available ground data for validation. Using four commonly used layered-earth inversion methods, several combinations of input parameters (controlling constraints) were run for each. Inversion methods comprised

the following: 1) smooth, laterally constrained inversion (LCI), 2) few-layer LCI (Auken and Christiansen, 2004), 3) sharp LCI (Vignoli et al., 2015) and 4) EM1DFM multilayer (Farquharson et al., 2003) types. The most suitable parameters were selected for each through quantitative comparisons with in-situ data. The resulting inversions were converted from EC to groundwater salinity using petrophysical and analytical relationships (Archie, 1942; De Louw et al., 2011; Waxman and Smits, 2003a), and ultimately 3D volumes of groundwater salinity using a novel interpolation method. Subsequently, 3D groundwater salinity estimates for each method were then compared to assess practical groundwater salinity mapping outcomes.

Results were split into four sections: 1) qualitative comparisons, 2) differences in fresh-brackish-saline groundwater volume estimates, 3) overall accuracy vs. in-situ data, and 4) interface mapping accuracy. A qualitative comparison highlighted that all inversions were similar and appeared to agree well with in-situ data, with only subtle structural differences noted. However, large differences in estimated fresh groundwater volumes were noted between the approaches, differing by up to ~7% between the few-layer and smooth LCI methods. In practical terms, this is a difference of ~178 million m3 in a total mapped volume of 2.8 billion m3. Overall accuracy against in-situ data found that smoother inversions were more precise over sharper methods, and that the LCI smooth inversion was preferable in this case. For interface mapping however, fresh-brackish interfaces were consistently mapped too shallow, overestimating the thickness of the brackish zone. In this regard sharper inversions appeared to be better suited.

Generally, this level of uncertainty corroborates with Delsman et al. (2018), where inversions were found to add the greatest amount of uncertainty in a regional scale groundwater salinity AEM study. In conclusion, smooth, multilayer inversions seemed preferable overall, however caution is advised for all methods – as carefully controlling the amount of smoothness is needed to avoid overestimating the thickness of the brackish zone.

6.2.2 How are results affected by different quantities of available data? (Chapter 3)

The effect that different data densities and types have on regional saline groundwater mapping using AEM was quantitatively examined using highly detailed 3D synthetic models. Increasing interest in the field of hydrogeophysics (Binley et al., 2015) has highlighted the potential of using AEM to map groundwater salinity quantitatively in 3D (e.g., Delsman et al., 2018; Faneca Sànchez et al., 2012; Vandevelde et al., 2018). As geophysical inversions – a necessary part of AEM data processing – are most commonly delivered as 1D vertical measurements, a considerable amount of further processing is required to recover a fully 3D, quantitative groundwater salinity model in practical units (such as chloride in g/l). This research question follows on from chapter 2, as those results guided the inversion approach used here.

Inverted data are represented as bulk electrical conductivity (ECb) – i.e., the combined effect of both pore water and lithology. As a result, petrophysical relationships are needed to split the signal to recover the conductivity of pore water (ECw) using formation factors (Archie, 1942; Waxman and Smits, 2003). This step requires the collection of additional, potentially expensive lithological information, e.g. using borehole descriptions or cone penetration tests. Of course, an additional budgetary consideration is the survey itself, where spacing between flightlines (i.e., the tightness of the survey grid) needs to be allocated during survey design according to mapping specifications. Although AEM systems have been shown to be reliable and repeatable (Huang and Cogbill, 2006), little is currently understood about the effect of data density on regional mapping. Consequently, in this chapter we investigated relative sources of error based on established 3D groundwater salinity mapping techniques. Specifically, the impact of the intensity of two data types were considered: 1) flightline spacing of the AEM survey and 2) lithological information from boreholes needed to translate ECb into ECw.

To obtain a quantifiable understanding of these aspects, a regional-scale 3D synthetic model was constructed based on real hydrogeological and geophysical data. In total 90 model realisations were run with different densities and spatial configurations of lithological data (in the form of boreholes) and flightline spacing. The results of these realisations were compared with the 'real' synthetic model in three components: 1) overall error simply as model differences, 2) groundwater volume estimates, and 3) fresh-brackish-saline interface mapping.

In general, it was found that error as a result of the inversion process far exceeded that of poorly understood lithological information. It was also noted that decreasing flightline distance consistently improved results. As a result, simply handling the inversion process carefully and assigning an appropriate flightline spacing are the most important aspects to consider for accurate AEM groundwater salinity mapping. These observations are good news for areas with sparse lithological observations, however the need for improved inversion methods is emphasised. As overly smooth inversions appear to be a large source of error, it is suggested that in this case methods that could offer sharper inversion results are used (e.g., Vignoli et al., 2015), or even a probabilistic approach that results in a number of possible models (Minsley, 2011). Overall, these results highlighted that regional-scale mapping could benefit from an inversion method that offers an appropriate level of smoothing relative to the groundwater characteristics in that particular area.

6.2.3 Based on the results of chapters 2 and 3, what further methodological improvements can we make? (Chapter 4)

Chapter 3 concluded that in regard to groundwater salinity mapping accuracy, the improved accuracy obtained from an abundance of lithological data is relatively negligible *if* the inversion method is unsuitable. This conclusion is drawn despite using the best performing inversion from chapter 2, as well it being a commonly used method in other hydrogeophysical

studies (e.g., Auken et al., 2008; Chongo et al., 2015; Delsman et al., 2018; Gunnink et al., 2012). The observed error in this case appears to stem from the inversion result being overly smooth, and thus predicting a brackish zone that is too thick. Of course, the same inversion could also be too sharp in areas with different subsurface characteristics. In order to assess the correct level of smoothness of AEM inversions, a practical inversion approach was developed that utilises cheaply available ground data to supplement an established inversion method (Auken and Christiansen, 2004).

The method exploits in-built inversion parameters to control the level of smoothness based on known brackish zone thicknesses, which are often readily available from monitoring wells. Here, a regularisation (or penalisation) term can be set to either favour default inversion parameters or the observed thickness of the brackish zone. Simply fitting a model to the brackish zone is not necessarily the optimal approach, as this could introduce unwanted noise into results – and ultimately goes against the point of recovering smooth results, which exist to stabilise models and minimise noise (Arsenin, 1979; Constable, 1987). The procedure was first tested using a realistic synthetic model to ascertain a good balance between fitting the brackish zone and the default (generally less noisy) parameters.

Results from the synthetic case showed that the default parameters were indeed too smooth, but also that favouring the ground data too much introduced considerable noise and uncertainty. The resulting optimal regularisation value obtained from the synthetic case was then applied to real data – comprising ~100 line km of flightlines and ~30 vertical profiles of ECb as ground data. Comparing the default and new (sharper) inversions against validation data showed that results were only marginally different, with some improvements observed when mapping the location of the fresh-brackish groundwater interface. These results highlight potential of the method, but that other inversion approaches should be considered for further improvements. These could include the use of coupled hydrogeophysical inversion, where groundwater simulations are physically modelled against AEM measurements (Herckenrath et al., 2013; A. C. Hinnell et al., 2010). Building on the method presented here, coupled hydrogeophysical modelling could be further developed by jointly matching in-situ brackish-zone thicknesses and AEM data.

6.2.4 Are groundwater salinity movements sensitive to repeated AEM surveys? (Chapter 5)
Chapter 5 was inspired by the conclusions of chapter 4, that highlighted the possible advantages of coupled hydrogeophysical modelling. Here, the novel idea of using repeated AEM surveys over the same location to improve the parameterisation of 3D variable density groundwater flow and salt transport models (3D-VDG) is investigated. This coupled approach also has the advantage of jointly estimating groundwater salinity distributions, potentially honouring subsurface characteristics with greater accuracy than conventional geophysical inversion.

The method builds on traditional inverse model parameterisation of 3D-VDG models, where parameters are iteratively adjusted until a suitable fit is found against observations (Carrera et al., 2005). In this process, uncertainties are reduced through the addition of observational data, such as groundwater salinity or hydraulic head – which are often sparse in regional 3D-VDG studies (e.g., Cobaner et al., 2012; Van Engelen et al., 2019). Repeat AEM surveys have the potential of providing two or more regional 3D estimates of groundwater salinity over time, thus for the purpose of model calibration offer great potential.

Addressing the research question – the first challenge was to understand if AEM surveys are sensitive to groundwater salinity movements over time, and if they are – how much time is needed to pass under a given set of conditions. This question was in part answered by Huang and Cogbill (2006), who showed that AEM surveys offer good repeatability in a static environment – however it was still unknown whether changes over time could be observed (i.e., from groundwater salinity movements). Using a detailed 3-D synthetic model, a 3D-VDG model was physically coupled with AEM observations through a forward modelling routine. The model was run for 60 years with realistic natural and anthropogenic stressors, such as modelled groundwater extraction. Each time-step (half-yearly) was then geophysically forward modelled and compared with the first time-step to assess AEM sensitivity over time. Results of this indicate that in this case study, after ~5+ years a second AEM survey should be sensitive to groundwater salinity differences.

An optimisation routine was subsequently implemented that minimises the observed groundwater salinity distributions from two separate AEM surveys in time. This was done using iterated 3D-VDG model runs over a period of time defined by the now understood sensitivity (time) period between two AEM surveys. At the end of the defined time period (in this case 15 years), the resulting 3D-VDG model is compared with AEM observations. The 3D-VDG model then restarts with updated hydrogeological parameters (e.g., hydraulic conductivity, porosity) until a suitable minimum is found – in our synthetic case study, successfully jointly estimating hydrogeological parameters and groundwater salinity. Our findings highlight the potential of this method, and build on other coupled hydrogeophysical methods (e.g., Herckenrath et al., 2013; Hinnell et al., 2010) by using repeat AEM surveys.

6.3 Recommendations for further research

While using AEM to map groundwater salinity is relatively well-established, there is an increasing interest for combining these with high-resolution, quantitative 3D groundwater salinity models. There also remain known sources of error in the method – in particular those introduced through processing steps. As a result, a number of recommendations for further research are suggested. The topics discussed are based on the main results and conclusions of

chapters 2 to 5, therefore the following will be discussed: 1) deterministic inversion methods, 2) coupled hydrogeophysical modelling, and 3) applying repeat AEM surveys to a real case.

6.3.1 Deterministic inversion methods

Traditional inversion methods are deterministic, whereby a single model is selected that fits the data. As these kinds of inversion are ill-posed (i.e., an infinite number of models can fit the data), selecting a single model is not recommended. The practical limitations to this are noted in chapters 2 and 3 of this thesis. One way to approach this problem is through probabilistic inversion (Minsley, 2011), which was recently further developed to include prior information as a constraint (Hansen and Minsley, 2019). The results of this method would allow the user to select either the most probable model, or indeed one that reflects the characteristics of the study appropriately. While these methods offer great potential, they are complex to implement, including the extensive post-processing, and are computationally expensive such that they require the use of distributed-memory super computers. As a result, these methods are not yet widely applied. However, further research into computational efficiency and the continued development of a more broadly accessible user-interface would almost certainly increase uptake. A less computationally expensive alternative to probabilistic inversion is simply the incorporation of prior information to deterministic methods. This was attempted in chapter 4 of this thesis, through the incorporation of brackish-zone thickness measurements to constrain smoothness. However, a few challenges remain as improvements were not as drastic as hoped. The most straightforward suggestion is to simply apply the method to a different inversion approach. This could be adapting the method to work with a so-called 'sharp' inversion, such as Vignoli et al. (2015), or adapted to an available open source method with published Python scripts (e.g., Cockett et al., 2015). Finally, besides simply constraining the inversion to match the thickness of the brackish zone using inversion parameters, research that examines the potential of incorporating a coupled hydrogeological model into this method is suggested. Here, the brackish zone thickness could be constrained to match in-situ data and run iteratively until a suitable match to AEM observations is found. The potential of coupled hydrogeophysical inversion is expanded upon in the section below.

6.3.2 Coupled hydrogeophysical modelling

Coupled hydrogeophysical modelling offers a promising alternative to stochastic geophysical inversion. Here, the use of deterministic (or indeed probabilistic) inversion is bypassed – where combined with petrophysical transformations, 3D-VDG groundwater models can be forward modelled and made to fit AEM observations (e.g., González and Comte, 2021; Herckenrath et al., 2013; Hinnell et al., 2010). A big advantage of this is that a conceptual hydrogeological model can be incorporated, thus allowing models that are more characteristic to a study area. It also has the advantage of encouraging interdisciplinary research between hydro(geo)logists and geophysicists, which helps produce realistic models and provides more accurate expectation management. However, this process comes with uncertainties of

its own – in particular the petrophysical model used. Here, research has found that due to model heterogeneity, using a low quality petrophysical model has the potential to introduce substantial error (González and Comte, 2020). As a result, further research is recommended to use lithological information more accurately on a regional scale. This includes a better understanding of incorporating data that is collected at different scales and resolutions, such localised lithological data and regional AEM observations. In general, this topic offers great potential and is considered a key research area for future studies.

6.3.3 Applying repeat AEM surveys to a real case

Chapter 5 quantitatively highlighted the potential of flying repeat AEM surveys. Besides the existing challenge noted with coupled hydrogeophysical modelling in the previous section, there are a number of challenges to overcome first before applying this to real data. The method presented in chapter 5 was a necessary simplification of reality – given that it was novel and therefore presented as proof-of-concept. As a result, there are three main simplifications that require further investigation. First, during the optimisation routine we assumed full knowledge of other hydrogeological properties such as the interaction between groundwater and drainage, as well as the locations and hydrogeological makeup of aquifers and aquitards. It is therefore suggested that this effect is examined further by adding noise to parameters that are not part of the optimisation, allowing permitted deviations according to known noise levels. Including these properties in a fully 3D heterogeneous model would allow a quantitative understanding of these assumptions. Second, a simplified, apparent formation factor was used for petrophysical transformations (Archie, 1942). This method does not take into account the surface conductivities of clay (Revil et al., 2017; Waxman and Smits, 2003), and therefore potentially adds error. It is therefore suggested that this effect is quantitatively included in a global uncertainty analysis. Third, from a practical perspective the optimisation used was very slow in terms of computation time, therefore research into using a fully parallelised optimisation is necessary. This could include evolutionary algorithms (Brauer et al., 2002b; Mühlenbein et al., 1991b) or parallel Bayesian optimisation (González et al., 2015; Kandasamy et al., 2017b). Finally, as hydrocarbon reservoir simulations utilise time-lapse geophysical acquisition and multiple flow-simulations (Sambo et al., 2020), it is suggested that a review of the state-of-art in this field should serve as further inspiration for a practical implementation in general in the field of hydrogeology.

A

Appendix to Chapter 2

The 3D interpolation validation

After removing inversions from every 2nd neighboring flightline, the above 3D interpolation method was repeated in order to validate the method. This was done using the LCI smooth inversion results. The validation 3D volume was therefore interpolated using ~500 line km of data, rather than the initial total of ~1000 line km. Considering the flightline spacing, distances between flightlines will mostly be 600m in the validation model, rather than the initial 300m – and 200m over the Waterfarm area where flightline spacing was 100m. A number of steps were taken to compare the two interpolation results. Firstly, the 'between' 2D profiles were resampled in the 'gaps' where the flightlines were removed, and compared for differences against the original interpolation (figure S1). Secondly, conductivity interfaces were extracted directly from the 3D volumes and compared for inconsistencies (figure S2). Thirdly, histogram distributions were calculated for each (figure S3), and finally, a comparison of volume estimates was undertaken (table S1).

An advantage to the 3D interpolation method presented here, is seen in the preservation of sharpness between low to high conductivity values at depth, where original layering from the inversion starting models is still preserved. Laterally less extensive features appear to be lost, as observed in figure S1 (feature 'A').

Predictably, shorter wavelength (therefore likely to be spatially less extensive) features were often lost, as indicated by the white boxes (Z and Y) in figure S2.a - this feature is also observed in figure S1 (feature A). This is likely because this information is often in-between flightlines and is smaller than the footprint of the HEM system. In this case, this is probably caused by flightlines beings perpendicular to smaller 'creek ridge' features. Naturally, features that are laterally less extensive than the flightline spacing used in the validation model, minus the AEM system footprint (i.e. 550m or less) are at risk of not being resolved. More subtle differences are between -0.2 to 0.2 S/m, as observed in the interfaces in figures S2.b and S2.c, this effect could shift an interface depth slightly – particularly at depth where inversion layers are thicker.

3D interpolation can inherently result in a smoothing effect, and in the case of hydrogeophysical mapping, this would result in an over-estimation of the brackish zone. Using histogram distributions of conductivity, this effect can be observed. A bimodal distribution is observed in the original 1D (along flight-line) results, as shown in the histograms in figure S3.

The bimodal nature of the inversion result was seen to be almost identical between the original inversion and that of the interpolation directly beneath inversion models, an indication that distributions were correctly honored directly beneath flightlines. When compared with the entire interpolated volume, the effect of smoothing was observed whereby conductivity values were clustered at around 0.3 to 0.7 S/m, and less so on either end of the curve. Overall however, the bimodal nature can be seen to be preserved effectively. As a result, it can be

assumed that the brackish zone thickness will be preserved beneath flightlines, and slightly smoothed between them.

Based on these histograms, volume estimates were undertaken for conductivity groups: (1) low (< 0.3 S/m), medium (0.3 – 0.7 S/m) and high (> 0.7 S/m) for both the interpolation using all the data, and the validation interpolation. These distributions could be seen as proxies for fresh (0 – 0.3 S/m), brackish (0.3 – 0.7 S/m) and saline (> 0.7S/m) distributions. This is shown in table S1.

Improvements could be made with regard to flightline anisotropy (downline striping), which was observed in the validation interpolation. However, this effect was not qualitatively observed in the final interpolation that used all lines. Furthermore, given that the inversion methods are layered earth 1D models, the 3D interpolation honored the true nature of the inversion results remarkably. Considering this study will extract fresh-saline groundwater interface depths and fresh groundwater volume estimates for analysis, the interpolation method was found to be acceptable.
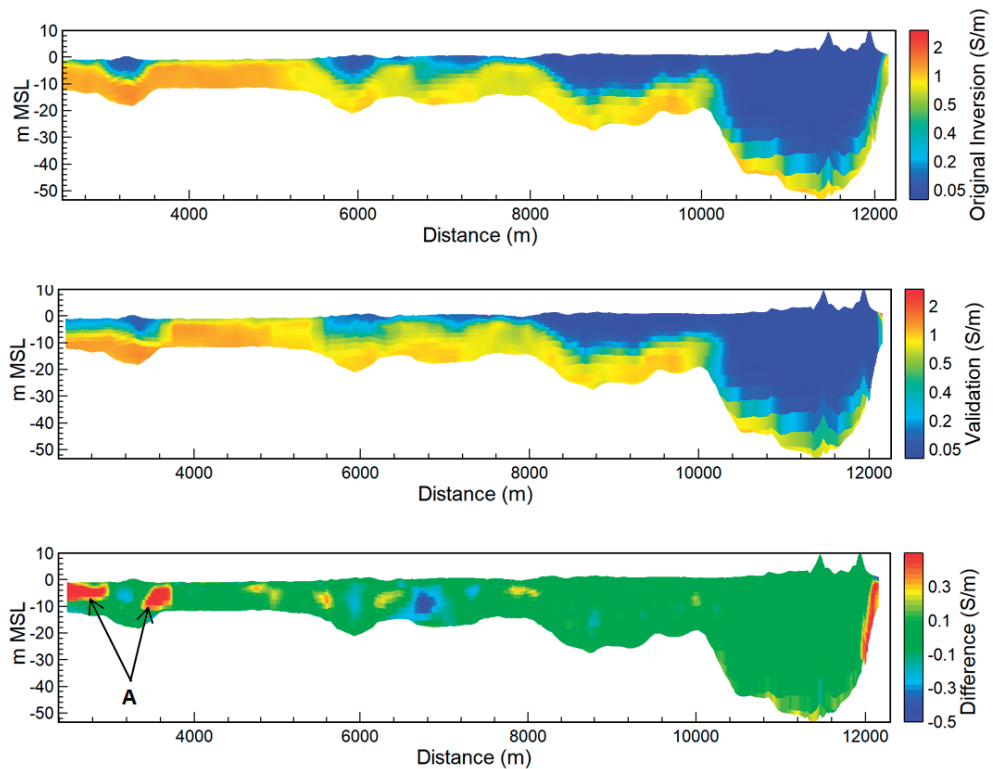
Figure A.1. *Interpolation validation sections. Top: original 1D inversion section. Middle: validation section (data sampled from between flightlines). Bottom: differences in S/m between the original inversion and the interpolation validation*
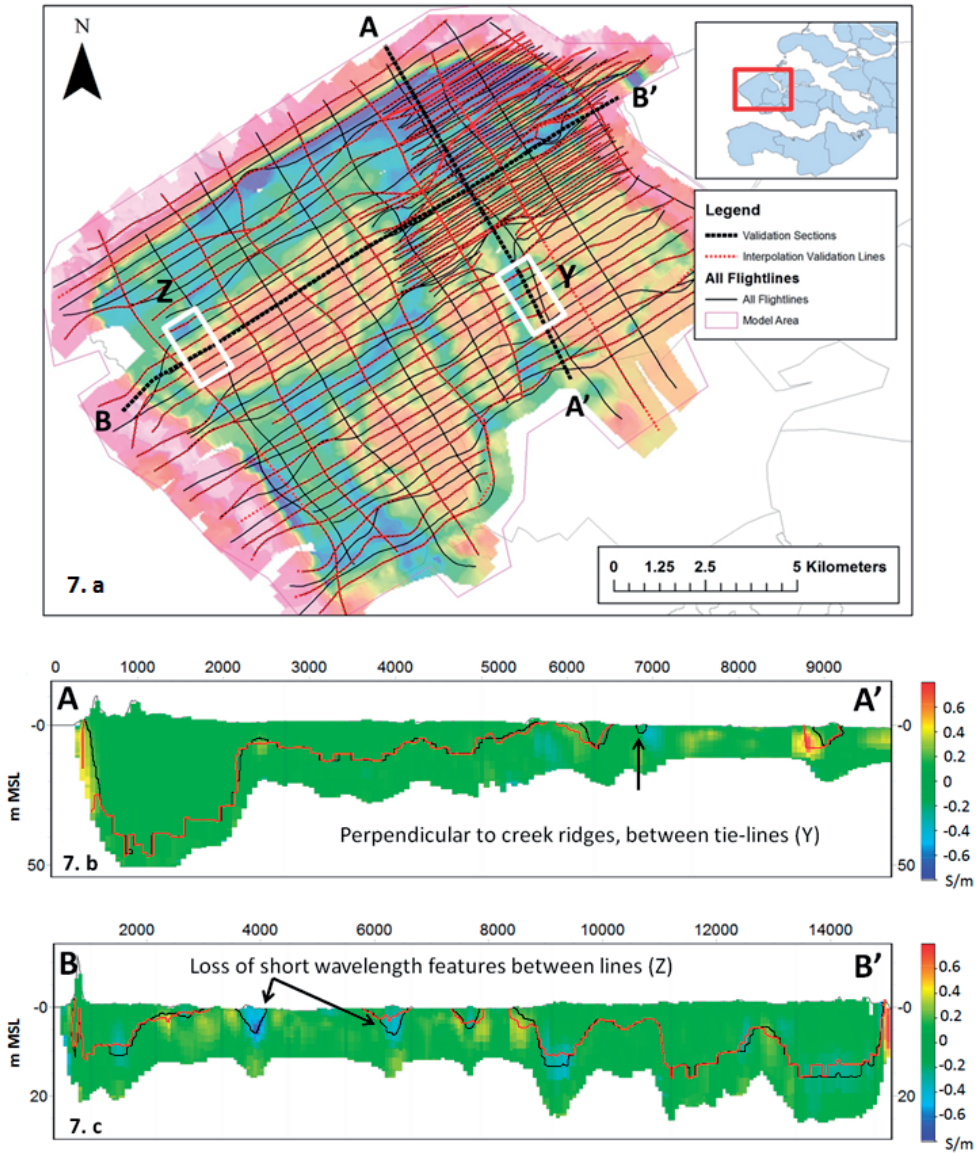
Figure A.2. *Interpolation validation sections taken from the 3D volume. Top (S2.a): horizontal 2D map of data and flightline locations - all flightlines (thin black lines), flightlines used in the validation (dotted red lines), section locations A-A' and B-B' (thick black lines). Features discussed are highlighted in white boxes (Z and Y). Background is a gridded channel of the raw data to illustrate relative high/low conductivity areas. Sections A-A' & B-B' (S2.b and S2.c respectively): differences observed between the interpolation and validation, given as difference in S/m, along with an extracted interface of 0.18 S/m. black = interpolation interface using all flightlines, red = validation interface using every second line.*
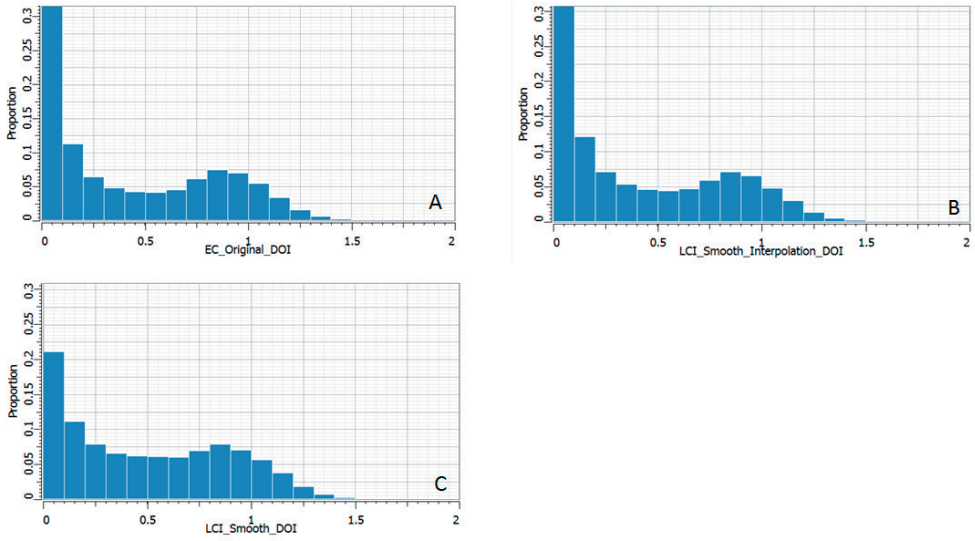
Figure A.3. *Histogram distributions of inversion results in S/m. Top left (A): original inversion without interpolation. Top right (B): interpolated inversion directly beneath flightlines. Bottom left (C): distribution of the entire 3D volume.*
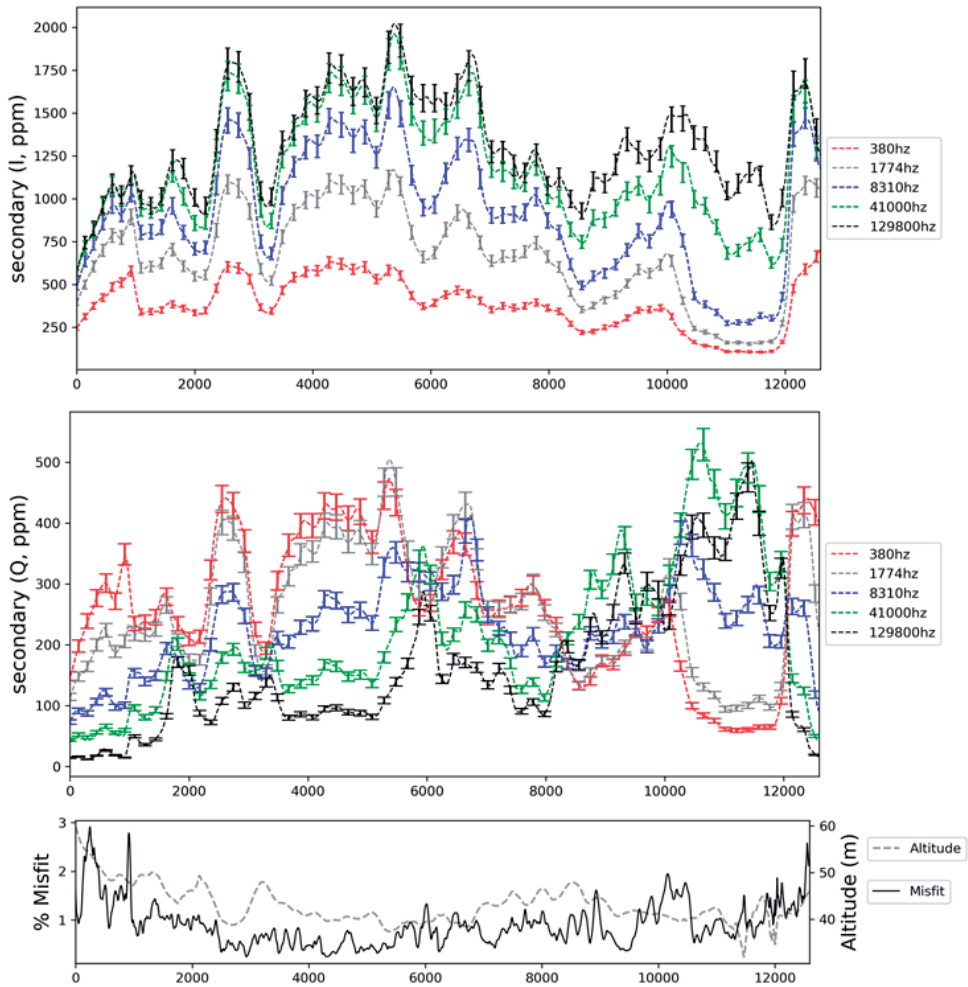
Figure A.4. *LCI Sharp inversion misfit of lateral and vertical constraints input as 1.3 and 2 respectively. Vertical and horizontal sharpness as 30 and 15 respectively. Predicted (dashed lines), observed (as error bars as 5% of data magnitude) for in-phase (top) and Quadrature (middle) components. Misfit (bottom) shown as solid black line as relative % error, instrument altitude as grey dashed line.*
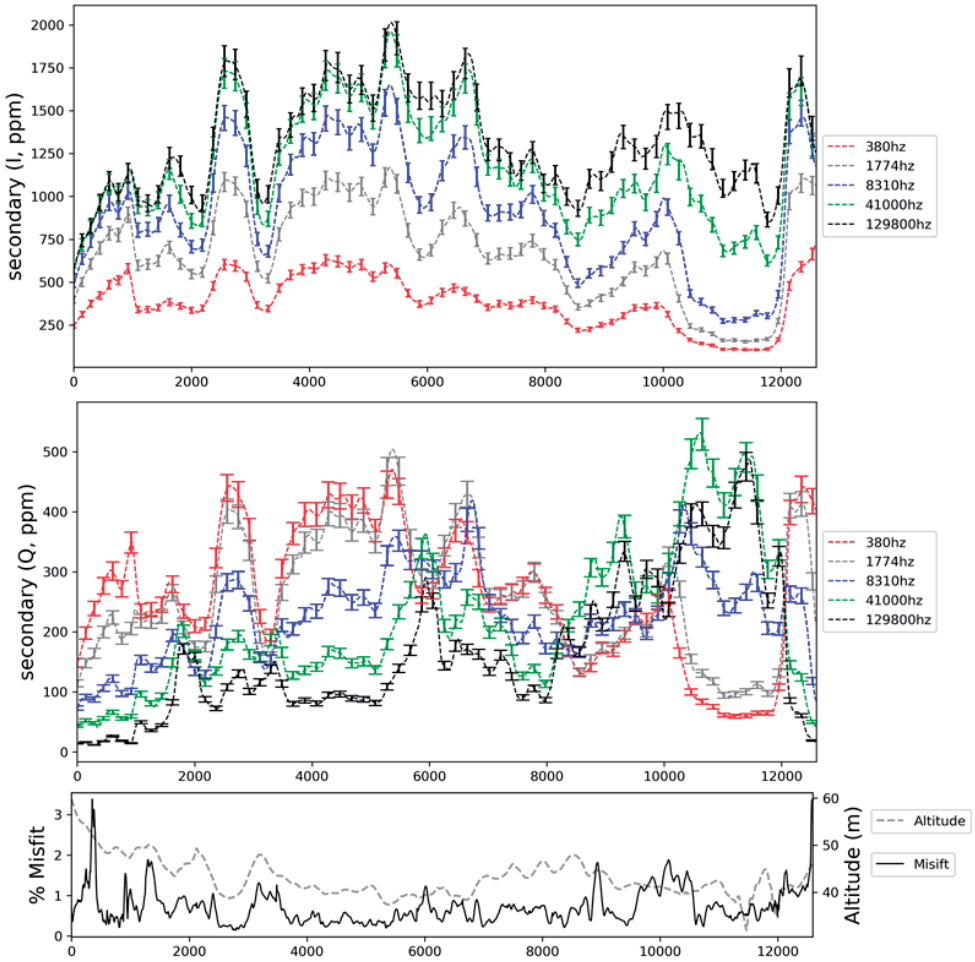
Figure A.5. *LCI 5 layer inversion misfit of lateral and vertical constraints input as 1.3 and 3 respectively. Predicted (dashed lines), observed (as error bars as 5% of data magnitude) for in-phase (top) and Quadrature (middle) components. Misfit (bottom) shown as solid black line as relative % error, instrument altitude as grey dashed line.*
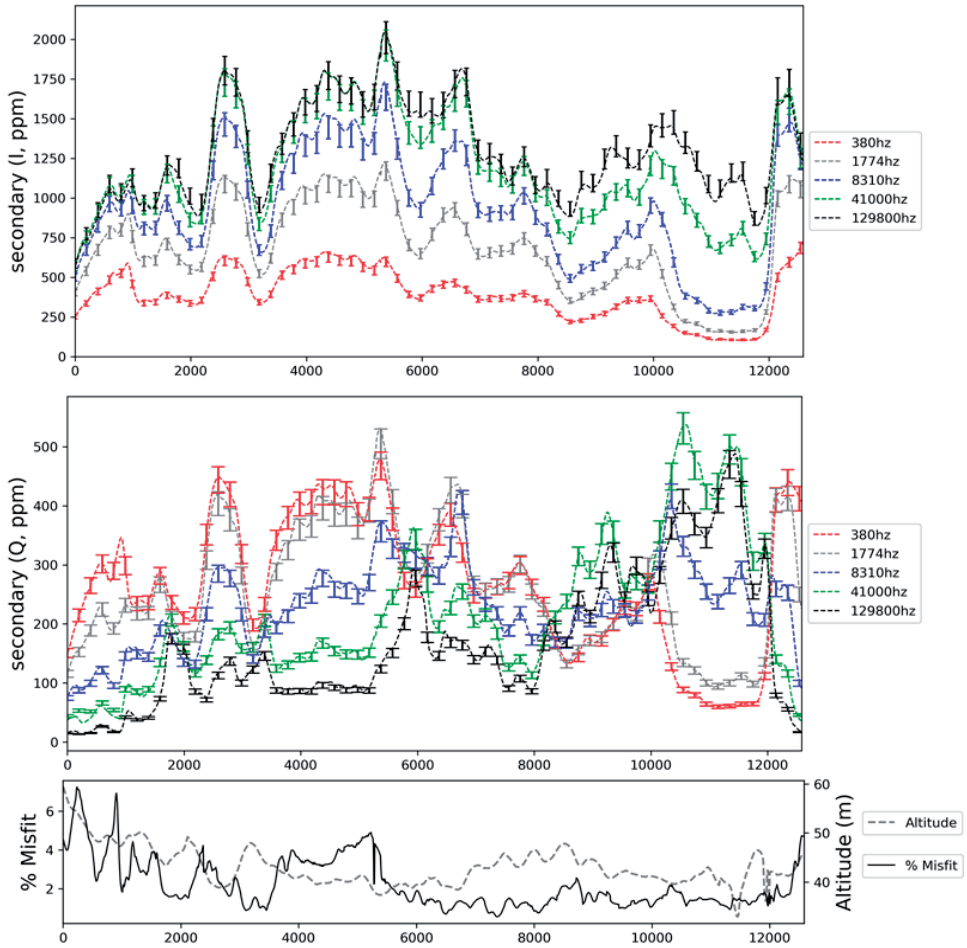
Figure A.6. *UBC Fixed Trade-Off inversion misfit with the trade-off parameter as three. Predicted (dashed lines), observed (as error bars as 5% of data magnitude) for in-phase (top) and Quadrature (middle) components. Misfit (bottom) shown as solid black line as relative % error, instrument altitude as grey dashed line.*

Table A.1. *Volume differences between the interpolation and the validation based on the conductivity distribution, illustrated as percentages.*

	< 0.3 S/m %	0.3 - 0.7 S/m %	> 0.7 S/m %
Original Interpolation	40.14	24.90	34.94
Validation Interpolation	39.52	25.22	35.25

B

Appendix to Chapter 3
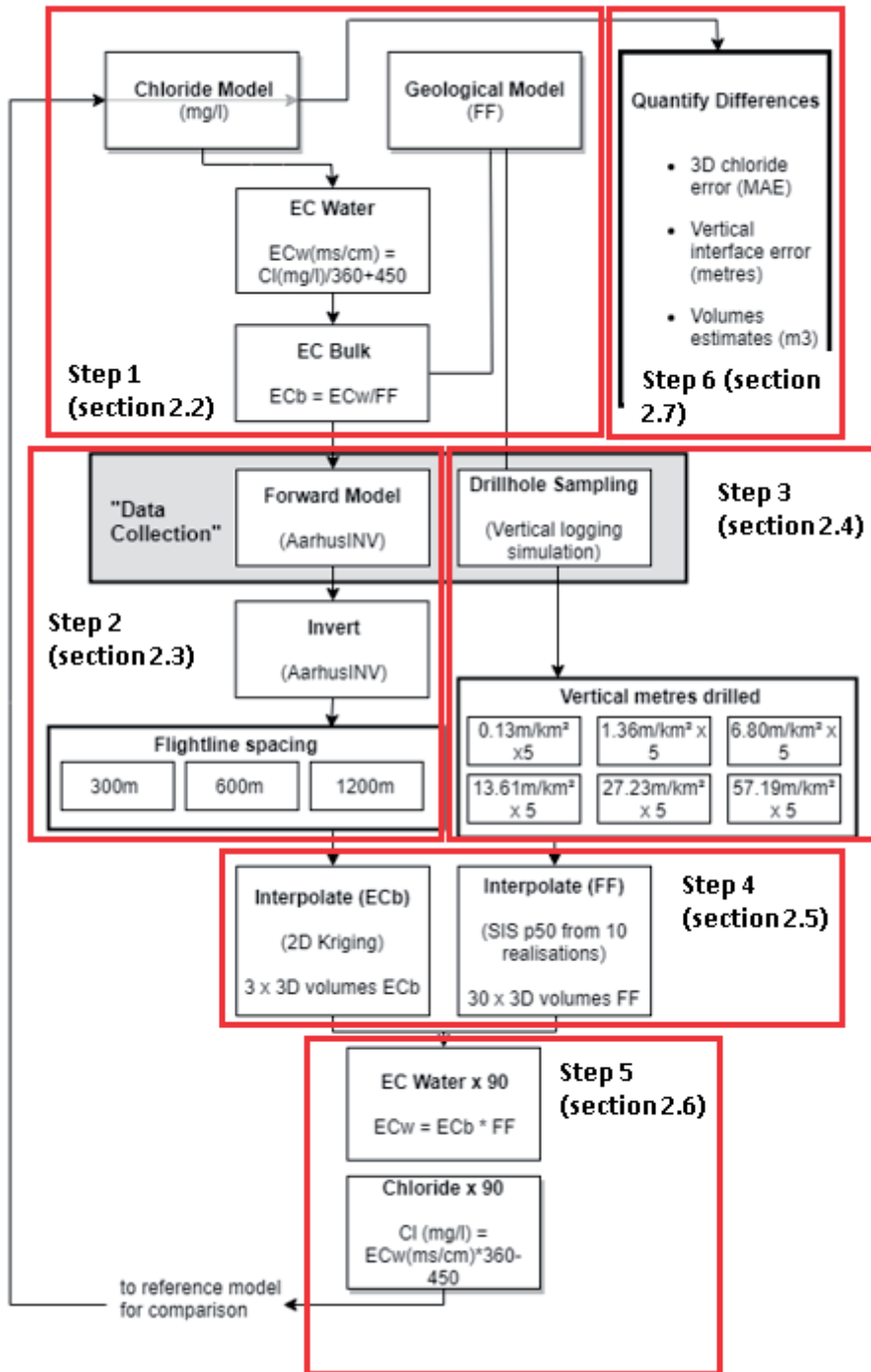
Figure B.1. *Overview of methods to compare various data configurations and resulting 3D salinity mapping error.*

MAE: Borehole Data vs. Reference Model
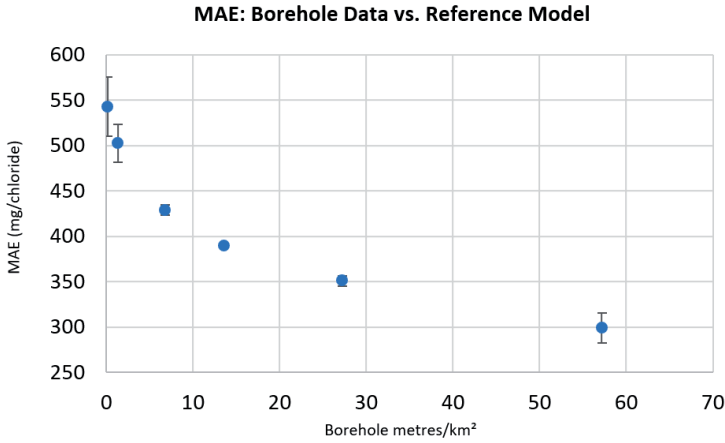
Figure B.2. *Comparing geological data density against the reference mode assuming a perfect AEM survey and inversion. X-axis = geology data density as vertical metres drilled per km2.Y-axis = MAE of chloride distributions as mg/l. Error bars highlight uncertainty as a result of borehole placement and subsequent interpolation.*
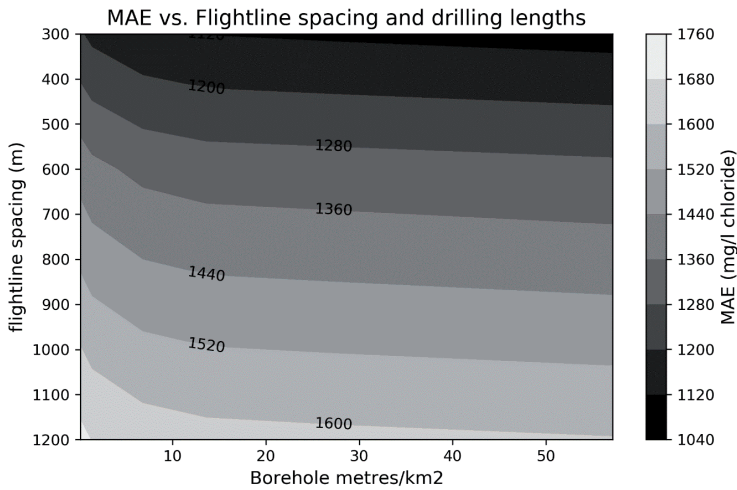
Figure B.3. *A contour plot of all data configurations tested. X-axis = geology data density as vertical metres drilled per km2. Y-axis = data density of the AEM survey represented as flightline spacing (i.e. smaller spacing is more data).*

Total Volumes: Reference and Inverted Models
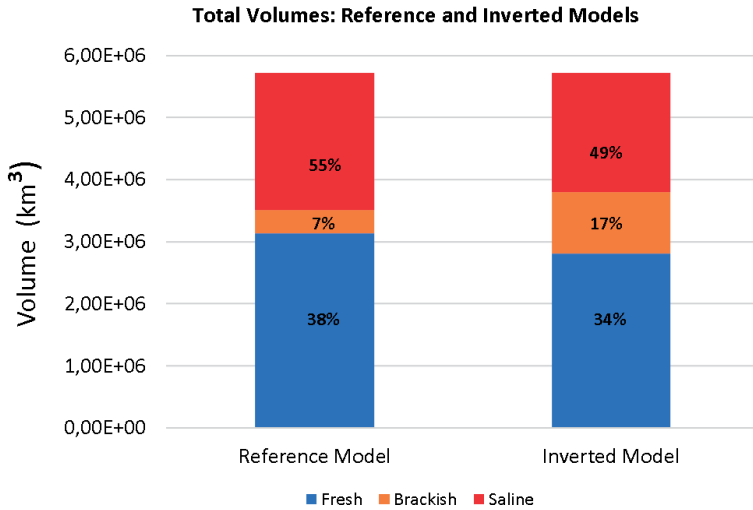
Figure B.4. *Total volumes of fresh-brackish-saline groundwater regions of the reference model and inverted (300m line spacing) model in km3. The reference geological model was applied to the inverted model, thus a perfect geological map is assumed. Blue = fresh, orange = brackish, red = saline.*

Volumes Differences vs. Geological Data Density
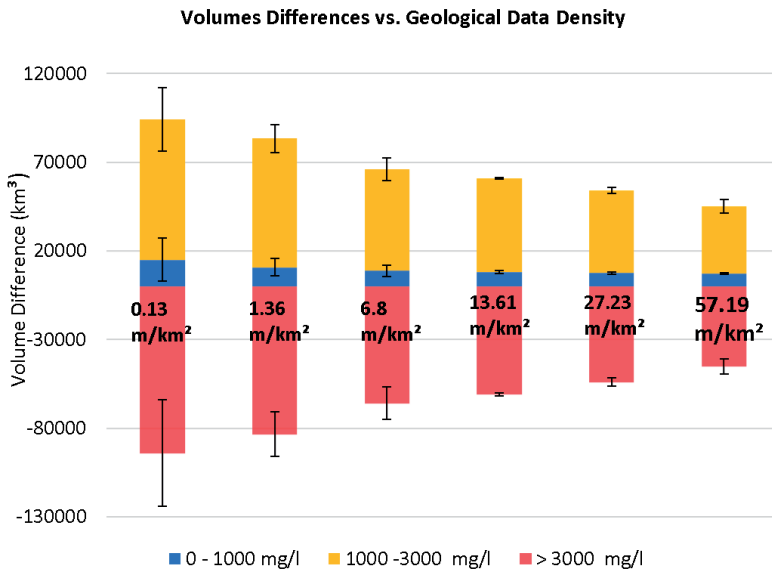
Figure B.5. *Volume difference against the reference model of fresh-brackish-saline groundwater regions in m3 difference. Calculations are based on applying geological information (FF) to the reference model, rather than inverted data. Blue = fresh, orange = brackish, red = saline. Borehole data density labelled as vertical metres/km2. Positive values represent an overestimation.*

To understand the spatial uncertainty with regards to interface mapping, figure 10 illustrates the possible effects of applying different FF values to the inversion model for chloride classes 0 – 5000mg/l.
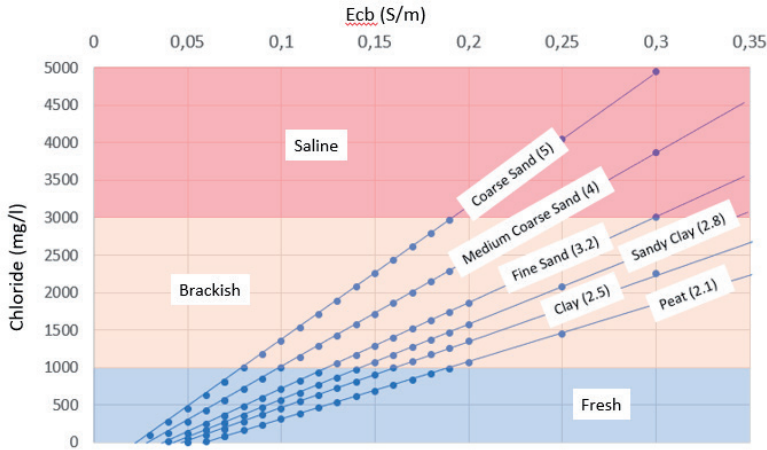
Figure B.6. *Possible ranges of mapped chloride based on the allocated FF value for values between 0 – 5000mg/l. Chloride values are calculated based on the process outlined in figure 1.*
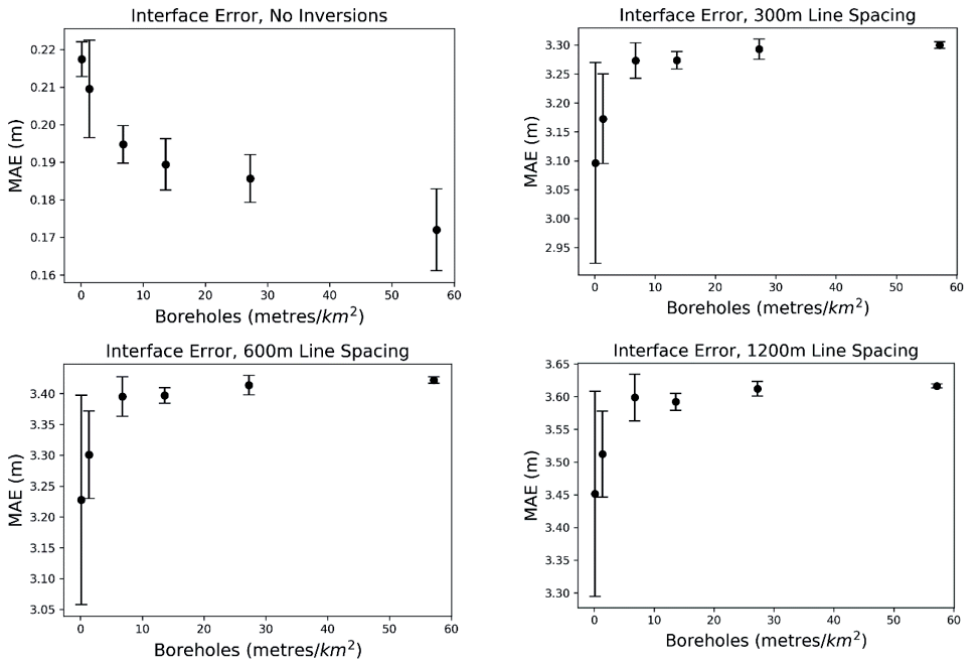
Figure B.7. *Absolute vertical error of the 500mg/l chloride interface for each flightline spacing class, as well as a comparison directly against the reference model.*
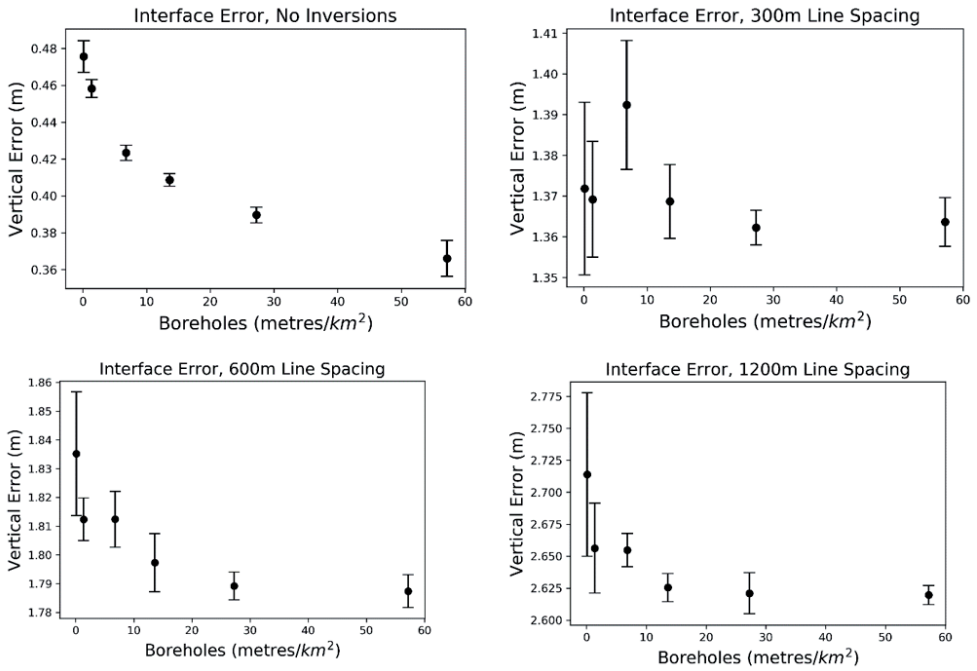
Figure B.8. *Absolute vertical error of the 3000mg/l chloride interface for each flightline spacing class, as well as a comparison directly against the reference model.*

C

Appendix to Chapter 5

Figure B1 illustrates a single flightline showing chloride distributions at t=0 and t=1, along with the geological section and forward models for the in phase component at that location. Plots of airborne electromagnetic forward models based on 3D variable-density groundwater flow and coupled salt transport models (3D-VDG) time-steps are shown below for all flightlines (figures C.2 – C.25).
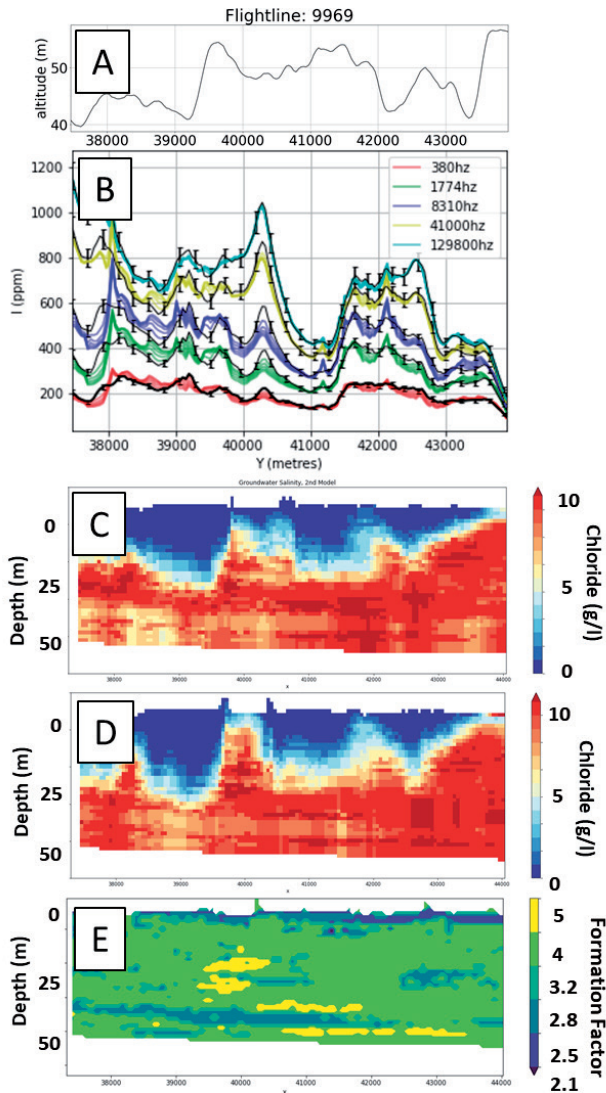
Figure C.1: *A single flightline showing the forward models for the in-phase component at that location as parts per million, coloured lines represent successive forward models at each time step at 5yr intervals (B), as well as flight altitude (A). Inverted chloride distributions, translated into chloride, at t=0 and t=1 (C and D respectively) along with the geological section (E) at that location. The geological section illustrates the formation factors used, therefore: 5 = coarse sand, 4 = medium sand, 3.2 = fine sand, 2.8 = sandy clay, 2.5 = clay, 2.1 = peat.*
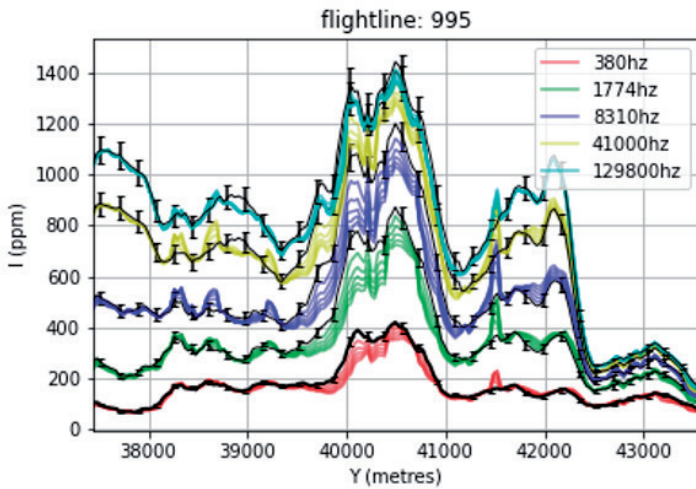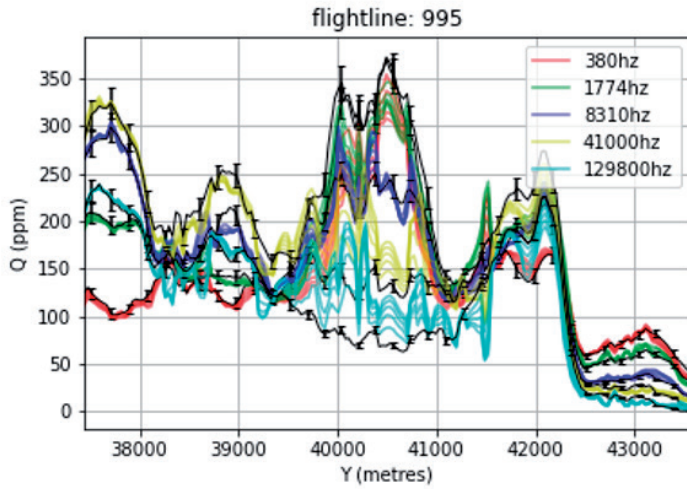
Figures C.2 – C2.5: *Both in-phase and quadrature are shown in parts per million (ppm) for all five frequencies across all flightlines. The dark line represents the initial forward model (at t=0), error bars are 5% of data magnitude and thus represent the error threshold. The coloured lines show change over time in increments of 5 years up to a maximum of 30 years in 3D-VDG simulations.*

flightline: 99

flightline: 99

flightline: 97

flightline: 97

flightline: 93

flightline: 93

flightline: 92

flightline: 92

flightline: 91

flightline: 91

flightline: 89

flightline: 89

flightline: 88

flightline: 88

flightline: 87

flightline: 87

flightline: 86

flightline: 86

flightline: 85

flightline: 85

flightline: 84

flightline: 84

flightline: 79

flightline: 79

flightline: 78

flightline: 78

Bibliography

Abdallatif, T.A., Abdel Rahman, A.A., Abu Ashour, N.M.H., Auken, E., Berthold, S., Blindow, N., Börner, F., Christiansen, A. V., Dietrich, P., Ernstson, K., Gabriel, G., Hertrich, M., Hördt, A., Janik, M., Jørgensen, F., Kirsch, R., Krummel, H., Leven, C., Lykke-Andersen, H., Abbas Mabrouk, M., Rabbel, W., Siemon, B., Sandersen, P.B.E., Sørensen, K.I., Tezkan, B., Wiederhold, H., Yaramanci, U., 2009. Groundwater Geophysics. A Tool for Hydrogelogy. 2nd Edition, Magnetic, geothermal and radioactive methods. https://doi.org/10.1007/978-3-540-88405-7
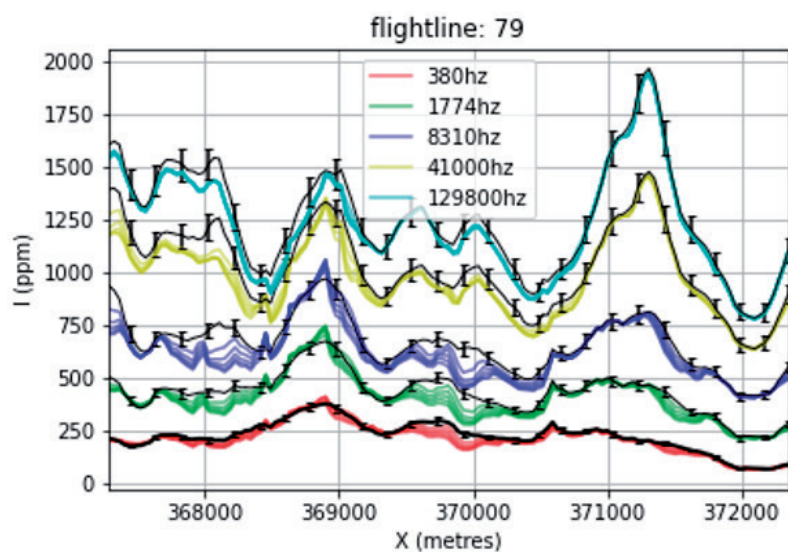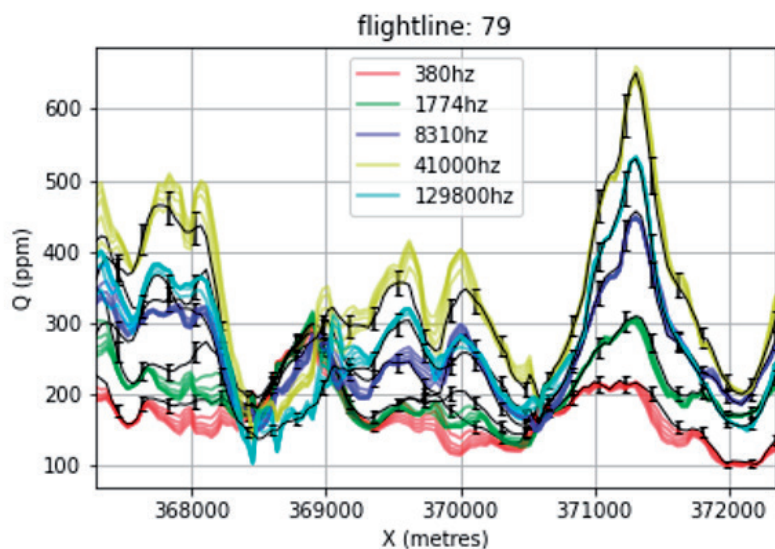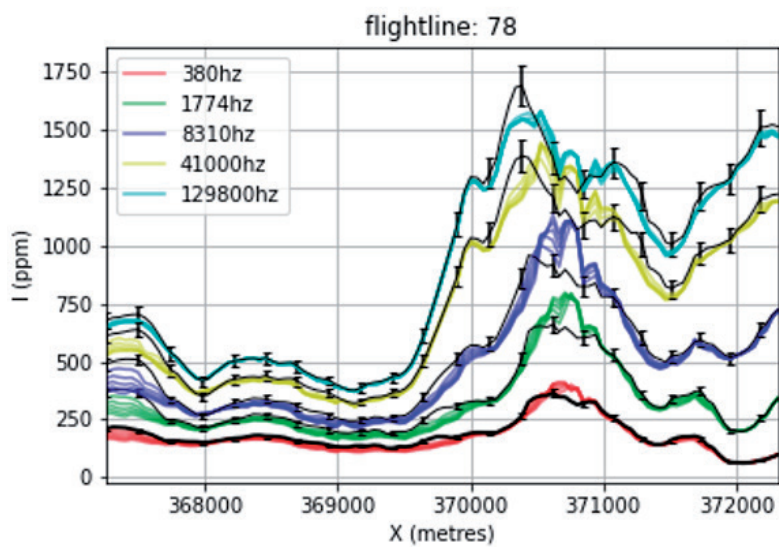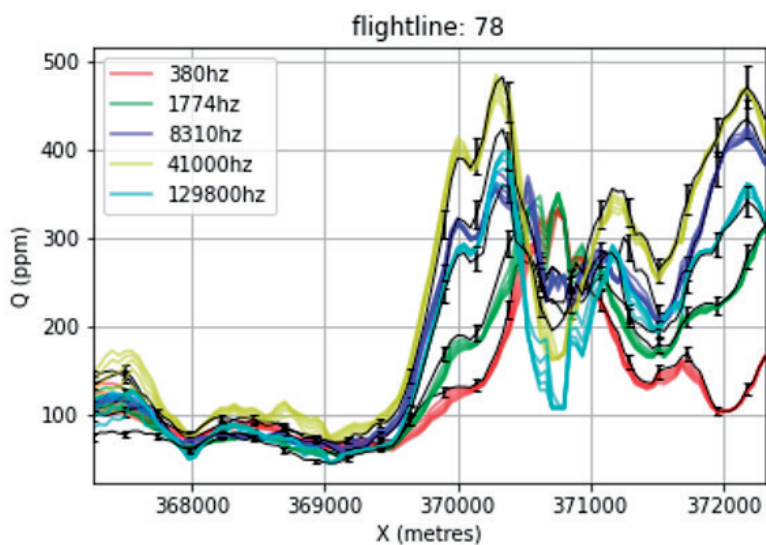
Abdollahzadeh, A., Reynolds, A., Christie, M., Corne, D., Davies, B., Williams, G., 2011. Bayesian optimization algorithm applied to uncertainty quantification, in: 73rd European Association of Geoscientists and Engineers Conference and Exhibition 2011: Unconventional Resources and the Role of Technology. Incorporating SPE EUROPEC 2011. Society of Petroleum Engineers, pp. 2317–2331. https://doi.org/10.2118/143290-ms

Archie, G.E., 1942. The Electrical Resistivity Log as an Aid in Determining Some Reservoir Characteristics. Pet. Technol. 54–62. https://doi.org/10.2118/942054-G

Arsenin, A.N.T. and V.Y., 1979. Solutions of Ill-Posed Problems. Soc. Ind. Appl. Math. 21, 266–267.

Auken, E., Christiansen, A.V., 2004. Layered and laterally constrained 2D inversion of resistivity data. Geophysics 69, 752–761. https://doi.org/10.1190/1.1759461

Auken, E., Christiansen, A.V., Kirkegaard, C., Fiandaca, G., Schamper, C., Behroozmand, A.A., Binley, A., Nielsen, E., Effersø, F., Christensen, N.B., Sørensen, K., Foged, N., Vignoli, G., 2015. An overview of a highly versatile forward and stable inverse algorithm for airborne, ground-based and borehole electromagnetic and electric data. Explor. Geophys. https://doi.org/10.1071/EG13097

Auken, E., Christiansen, A. V., Jacobsen, B.H., Foged, N., Sørensen, K.I., 2005. Piecewise 1D laterally constrained inversion of resistivity data. Geophys. Prospect. 53, 497–506. https://doi.org/10.1111/j.1365-2478.2005.00486.x

Auken, E., Christiansen, A. V., Jacobsen, L.H., Sørensen, K.I., 2008. A resolution study of buried valleys using laterally constrained inversion of TEM data. J. Appl. Geophys. 65, 10–20. https://doi.org/10.1016/j.jappgeo.2008.03.003

Auken, E., Foged, N., Larsen, J.J., Valdemar, K., Lassen, T., Kumar Maurya, P., Dath, S.M., Eiskjaer, T.T., 2019. tTEM — A towed transient electromagnetic system for detailed 3D imaging of the top 70 m of the subsurface. https://doi.org/10.1190/GEO2018-0355.1

Ball, L., Smith, B., Minsley, B., Abraham, J., Voss, C., Astley, B., Deszcz-Pan, M., Cannia, J., 2010. Airborne Electromagnetic and Magnetic Geophysical Survey Data of the Yukon Flats and Fort Wainwright Areas , Central Alaska , June 2010 Airborne Electromagnetic and Magnetic Geophysical Survey Data of the Yukon Flats and Fort Wainwright Areas , Central Al. U.S. Geol. Surv. Rep.

Barfod, A.A.S., Møller, I., Christiansen, A. V., 2016. Compiling a national resistivity atlas of Denmark based on airborne and ground-based transient electromagnetic data. J. Appl. Geophys. 134, 199–209. https://doi.org/10.1016/j.jappgeo.2016.09.017

Barlow, P.M., Reichard, E.G., 2009. Saltwater intrusion in coastal regions of North America. Hydrogeol. J. 2009 181 18, 247–260. https://doi.org/10.1007/S10040-009-0514-3

Bauer-Gottwein, P., Gondwe, B.N., Christiansen, L., Herckenrath, D., Kgotlhang, L., Zimmermann, S., 2010. Hydrogeophysical exploration of three-dimensional salinity anomalies with the time-domain electromagnetic method (TDEM). J. Hydrol. 380, 318–329. https://doi.org/10.1016/J.JHYDROL.2009.11.007

Beaujean, J., Nguyen, F., Kemna, A., Antonsson, A., Engesgaard, P., 2014. Calibration of seawater intrusion models: Inverse parameter estimation using surface electrical resistivity tomography and borehole data. Water Resour. Res. https://doi.org/10.1002/2013WR014020

Bedrosian, B.P.A., Ball, L.B., Bloss, B.R., Survey, U.S.G., 2013. Airborne Electromagnetic Data and Processing within Leach Lake Basin – Fort Irwin , California.

Bedrosian, P.A., Schamper, C., Auken, E., 2016. A comparison of helicopter-borne electromagnetic systems for hydrogeologic studies. Geophys. Prospect. 64, 192–215. https://doi.org/10.1111/1365-2478.12262

Begemann, H. K. S. (1965). The friction jacket cone as an aid in determining the soil profile. Proceedings of the 6th International Conference on Soil Mechanics and Foundation Engineering, ICSMFE, Montreal, September 8 - 15 (Vol. 2, pp. 17 - 20).

Berendsen, H.J.A., 2005. The Rhine-Meuse delta at a glance, 8th International Conference on Fluvial Sedimentology. https://doi.org/ISBN-number: 90-77079-13-0

Binley, A., Hubbard, S.S., Huisman, J.A., Revil, A., Robinson, D.A., Singha, K., Slater, L.D., 2015. The emergence of hydrogeophysics for improved understanding of subsurface processes over multiple scales. Water Resour. Res. 51, 3837–3866. https://doi.org/10.1002/2015WR017016

Bocanegra, E., da Silva, G.C., Custodio, E., Manzano, M., Montenegro, S., 2009. State of knowledge of coastal aquifer management in South America. Hydrogeol. J. 2009 181 18, 261–267. https://doi.org/10.1007/S10040-009-0520-5

Boesen, T., Auken, E., Christiansen, A.V., Fiandaca, G., Kirkegaard, C., Pfaffhuber, A.A., Vöge, M., 2018. An efficient 2D inversion scheme for airborne frequency domain data. Geophysics 83, 1–47. https://doi.org/10.1190/geo2017-0280.1

Bording, T., Christiansen, A. V., Auken, E., Gunnink, J.L., Oude Essink, G.H.P., 2017. Groundbased TEM survey in the subsiding mekong delta. 23rd Eur. Meet. Environ. Eng. Geophys. 2017, 1–5. https://doi.org/10.3997/2214-4609.201701991/CITE/REFWORKS

Brauer, M.J., Holder, M.T., Dries, L.A., Zwickl, D.J., Lewis, P.O., Hillis, D.M., 2002a. Genetic algorithms and parallel processing in maximum-likelihood phylogeny inference. Mol. Biol. Evol. 19, 1717–1726. https://doi.org/10.1093/oxfordjournals.molbev.a003994

Brodie, R., Green, A., & Munday, T. (2004). Constrained inversion of resolve electromagnetic data—Riverland, South Australia. Bentley, WA: CRC LEME.

Brodie, R., Sambridge, M., 2006. A holistic approach to inversion of frequency-domain airborne EM data. Geophysics 71, G301–G312. https://doi.org/10.1190/1.2356112

Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., Slooten, L.J., 2005. Inverse problem in hydrogeology. Hydrogeol. J. https://doi.org/10.1007/s10040-004-0404-7

Carrera, J., Hidalgo, J.J., Slooten, L.J., Vázquez-Suñé, E., 2010. Computational and conceptual issues in the calibration of seawater intrusion models. Hydrogeol. J. https://doi.org/10.1007/s10040-009-0524-1

Chongo, M., Christiansen, A.V., Tembo, A., Banda, K.E., Nyambe, I.A., Larsen, F., Bauer-Gottwein, P., 2015. Airborne and ground-based transient electromagnetic mapping of groundwater salinity in the Machile-Zambezi Basin, southwestern Zambia. Near Surf. Geophys. 13, 383–395. https://doi.org/10.3997/1873-0604.2015024

Cobaner, M., Yurtal, R., Dogan, A., Motz, L.H., 2012a. Three dimensional simulation of seawater intrusion in coastal aquifers: A case study in the Goksu Deltaic Plain. J. Hydrol. 464–465, 262–280. https://doi.org/10.1016/J.JHYDROL.2012.07.022

Cockett, R., Kang, S., Heagy, L.J., Pidlisecky, A., Oldenburg, D.W., 2015. SimPEG: An open source framework for simulation and gradient based parameter estimation in geophysical applications. Comput. Geosci. 85, 142–154. https://doi.org/10.1016/J.CAGEO.2015.09.015

Constable, S.C., 1987. Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. Geophysics 52, 289. https://doi.org/10.1190/1.1442303

Cooley, H., Phurisamban, R., Gleick, P., 2019. The cost of alternative urban water supply and efficiency options in California. Environ. Res. Commun. 1, 042001. https://doi.org/10.1088/2515-7620/AB22CA

Cox, L.H., Wilson, G.A., Zhdanov, M.S., 2012. 3D inversion of airborne electromagnetic data. Geophysics 77, WB59–WB69. https://doi.org/10.1190/geo2011-0370.1

Custodio, E., 2009. Coastal aquifers of Europe: an overview. Hydrogeol. J. 2009 181 18, 269–280. https://doi.org/10.1007/S10040-009-0496-1

David, R., Pyne, G., 2020. Aquifer storage recovery: Recent developments in the United States. Artif. Recharg. Groundw. 257–261. https://doi.org/10.1201/9781003078500-39

de Louw, P.G.B., Eeman, S., Siemon, B., Voortman, B.R., Gunnink, J., van Baaren, E.S., Oude Essink, G.H.P., 2011. Shallow rainwater lenses in deltaic areas with saline seepage. Hydrol. Earth Syst. Sci. 15, 3659–3678. https://doi.org/10.5194/hess-15-3659-2011

De Louw, P. G.B., Eeman, S., Siemon, B., Voortman, B.R., Gunnink, J., Van Baaren, E.S., Oude Essink, G.H.P., 2011. Shallow rainwater lenses in deltaic areas with saline seepage. Hydrol. Earth Syst. Sci. 15, 3659–3678. https://doi.org/10.5194/hess-15-3659-2011

Delsman, J.R., Hu-A-Ng, K.R.M.R.M., Vos, P.C.C., De Louw, P.G.B., Oude Essink, G.H.P., Stuyfzand, P.J.P.J.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., Van Baaren, E.S., Siemon, B., Dabekaussen, W., Karaoulis, M.C., Pauw, P.S., Vermaas, T., Bootsma, H., De Louw, P.G.B., Gunnink, J.L., Wim Dubelaar, C., Menkovic, A., Steuer, A., Meyer, U., Revil, A., Oude Essink, G.H.P., 2018. Large-scale, probabilistic salinity mapping using airborne electromagnetics for groundwater management in Zeeland, the Netherlands. Environ. Res. Lett. 13, 084011. https://doi.org/10.1088/1748-9326/aad19e

Dillon, P., 2005. Future management of aquifer recharge. Hydrogeol. J. 2005 131 13, 313–316. https://doi.org/10.1007/S10040-004-0413-6

Doherty, J., 2010. PEST: Model-Independent Parameter Estimation, User Manual: 5th Edition. PEST Man. 5th Editio, 279.

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, 4499–4516. https://doi.org/10.5194/hess-16-4499-2012

Farquharson, C.G., Oldenburg, D.W., 2004. A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems. Geophys. J. Int. 156, 411–425. https://doi.org/10.1111/j.1365-246X.2004.02190.x

Farquharson, C.G., Oldenburg, D.W., Routh, P.S., 2003. Simultaneous 1D inversion of loop–loop electromagnetic data for magnetic susceptibility and electrical conductivity. Geophysics 68, 1857. https://doi.org/10.1190/1.1635038

Fitterman, D., Deszcz-Pan, M., 2004. Characterization of Saltwater Intrusion in South Florida Using Electromagnetic Geophysical Methods. … 18th Salt Water Intrusion Meet. 405–416.

Fitterman, D. V., Deszcz-Pan, M., 2018. Helicopter EM mapping of saltwater intrusion in Everglades National Park, Florida. https://doi.org/10.1071/EG998240 29, 240–243. https://doi.org/10.1071/EG998240

Fitterman, D. V, Deszcz-Pan, M., 2001. Saltwater Intrusion in Everglades National Park, Florida Measured by Airborne Electromagnetic Surveys.

Foged, N., Auken, E., Christiansen, A.V., Sørensen, K.I., 2013. Test-site calibration and validation of airborne and ground-based TEM systems 78.

Foged, N., Marker, P.A., Christansen, A. V., Bauer-Gottwein, P., Jørgensen, F., Høyer, A.S., Auken, E., 2014. Large-scale 3-D modeling by integration of resistivity models and borehole data through inversion. Hydrol. Earth Syst. Sci. 18, 4349–4362. https://doi.org/10.5194/HESS-18-4349-2014

Fraser, D. C. (1978). Resistivity mapping with an airborne multicoil electromagnetic system. *Society of Exploration Geophysicists*, **43**(1), 144– 172.

Gao, F., Han, L., 2012. Implementing the Nelder-Mead simplex algorithm with adaptive parameters. Comput. Optim. Appl. 51, 259–277. https://doi.org/10.1007/s10589-010-9329-3

Goes, B.J.M., Oude Essink, G.H.P., Vernes, R.W., Sergi, F., 2009. Estimating the depth of fresh and brackish groundwater in a predominantly saline region using geophysical and hydrological methods, zeeland, the netherlands. Near Surf. Geophys. 7, 401–412. https://doi.org/10.3997/1873-0604.2009048

Gomaa, S.M., Hassan, T.M., Helal, E., 2021. Assessment of seawater intrusion under different pumping scenarios in Moghra aquifer, Egypt. Sci. Total Environ. 781, 146710. https://doi.org/10.1016/J.SCITOTENV.2021.146710

Gomez-Hernandez, J.J., Strivastave, R.M., 1990. Isim3D : an Ansi-C Three-Dimensional Multiple Indicator Conditional. Comput. Geosci. 16, 395–440.

González-quirós, A., Comte, J., 2020. Advances in Water Resources Relative importance of conceptual and computational errors when delineating saltwater intrusion from resistivity inverse models in heterogeneous coastal aquifers. Adv. Water Resour. 144, 103695. https://doi.org/10.1016/j.advwatres.2020.103695

González-Quirós, A., Comte, J.C., 2021. Hydrogeophysical model calibration and uncertainty analysis via full integration of PEST/PEST++ and COMSOL. Environ. Model. Softw. 145, 105183. https://doi.org/10.1016/J.ENVSOFT.2021.105183

González, J., Dai, Z., Hennig, P., Lawrence, N.D., 2015. Batch Bayesian Optimization via Local Penalization. Proc. 19th Int. Conf. Artif. Intell. Stat. AISTATS 2016 648–657.

Gossel, W., Sefelnasr, A., Wycisk, P., 2010. Modèelisation d'une paléeo-intrusion d'eau saléee dans la partie Nord du Système Aquifèere Nubien, Nord-Est de l'Afrique. Hydrogeol. J. 18, 1447–1463. https://doi.org/10.1007/s10040-010-0597-x

Green, A., Lane, R., 2013. Estimating Noise Levels in AEM Data. ASEG Ext. Abstr. 2003, 1–5. https://doi.org/10.1071/aseg2003ab093

Gunnink, J.L., Bosch, J.H.A., Siemon, B., Roth, B., Auken, E., 2012. Combining ground-based and airborne em through Artificial Neural Networks for modelling glacial till under saline groundwater conditions. Hydrol. Earth Syst. Sci. 16, 3061–3074. https://doi.org/10.5194/hess-16-3061-2012

Haber, E., Oldenburg, D., 2000. A GCV based method for nonlinear ill-posed problems. Comput. Geosci. 4, 41–63. https://doi.org/10.1023/A:1011599530422

Haider, K., Engesgaard, P., Sonnenborg, T.O., Kirkegaard, C., 2014. Numerical modeling of salinity distribution and submarine groundwater discharge to a coastal lagoon in Denmark based on airborne electromagnetic data. Hydrogeol. J. 23, 217–233. https://doi.org/10.1007/s10040-014-1195-0

Hansen, T.M., Minsley, B.J., 2019. Inversion of airborne EM data with an explicit choice of prior model. Geophys. J. Int. 218, 1348–1366. https://doi.org/10.1093/gji/ggz230

He, X., Koch, J., Sonnenborg, T.O., Flemming, J., Schamper, C., Refsgaard, J.C., 2014. Using Airborne Geophysical Data and Borehole Data. Water Resour. Res. 1–23. https://doi.org/10.1002/2013WR014593.Received

Herckenrath, D., Fiandaca, G., Auken, E., Bauer-Gottwein, P., 2013. Sequential and joint hydrogeophysical inversion using a field-scale groundwater model with ERT and TDEM data. Hydrol. Earth Syst. Sci. 17, 4043–4060. https://doi.org/10.5194/hess-17-4043-2013

Hermans, T., Vandenbohede, A., Lebbe, L., Martin, R., Kemna, A., Beaujean, J., Nguyen, F., 2012. Imaging artificial salt water infiltration using electrical resistivity tomography constrained by geostatistical data. J. Hydrol. 438–439, 168–180. https://doi.org/10.1016/j.jhydrol.2012.03.021

Hill., B., 2011. Helicopter Electromagnetic and Magnetic Geophysical Survey Data , Hunton Anticline , South-Central Oklahoma Helicopter Electromagnetic and Magnetic Geophysical Survey Data , Hunton Anticline , South-Central Oklahoma. USGS.

Hinnell, A C, Ferré, T.P.A., Vrugt, J.A., Huisman, J.A., Moysey, S., Rings, J., Kowalsky, M.B., 2010. Improved extraction of hydrologic information from geophysical data through coupled hydrogeophysical inversion 46, 1–14. https://doi.org/10.1029/2008WR007060

Hodges, G, Siemon, B., 2008. Comparative analysis of 1 dimensional inversions of helicopter-borne frequency-domain electromagnetic data. Proc. AEM2008 5th Int. Conf. Airborne Electromagn. 28–30.

Høyer, A.S., Jørgensen, F., Sandersen, P.B.E., Viezzoli, A., Møller, I., 2015. 3D geological modelling of a complex buried-valley network delineated from borehole and AEM data. J. Appl. Geophys. 122, 94–102. https://doi.org/10.1016/j.jappgeo.2015.09.004

Huang, H., Cogbill, A., 2006. Repeatability study of helicopter-borne electromagnetic data. Geophysics 71, G285–G290. https://doi.org/10.1190/1.2353797

Huang, H., Fraser, D.C., 1996. The differential parameter method for multifrequency airborne resistivity mapping. Geophysics 61, 100–109. https://doi.org/10.1190/1.1574674

Stafleu, J. D. Maljers, J.L. Gunnink, A.M.& F.S.B., 2011. 3D modelling of the shallow subsurface of Zeeland, the Netherlands. Netherlands J. Geosci. 16, 5647. https://doi.org/10.5242/iamg.2011.0076

Jasechko, S., Perrone, D., Seybold, H., Fan, Y., Kirchner, J.W., 2020. Groundwater level observations in 250,000 coastal US wells reveal scope of potential seawater intrusion. Nat. Commun. 2020 111 11, 1–9. https://doi.org/10.1038/s41467-020-17038-2

Johnston, D.H., 2013. Practical Applications of Time-lapse Seismic Data. Pract. Appl. Time-lapse Seism. Data. https://doi.org/10.1190/1.9781560803126

Jørgensen, F., Scheer, W., Thomsen, S., Sonnenborg, T.O., Hinsby, K., Wiederhold, H., Schamper, C., Burschil, T., Roth, B., Kirsch, R., Auken, E., 2012. Transboundary geophysical mapping of geological elements and salinity distribution critical for the assessment of future sea water intrusion in response to sea level rise. Hydrol. Earth Syst. Sci. 16, 1845–1862. https://doi.org/10.5194/hess-16-1845-2012

Kandasamy, K., Krishnamurthy, A., Schneider, J., Poczos, B., 2017. Asynchronous Parallel Bayesian Optimisation via Thompson Sampling.

Karaoulis, M.C., Kim, J.H., Tsourlos, P.I., 2011. 4D active time constrained resistivity inversion. undefined 73, 25–34. https://doi.org/10.1016/J.JAPPGEO.2010.11.002

King, J., Essink, G.O., Karaoulis, M., Bierkens, M.F.P., 2020. A practical quantification of error sources in regional-scale airborne groundwater salinity mapping OPEN ACCESS A practical quantification of error sources in regional-scale airborne groundwater salinity mapping.

King, J., Oude Essink, G., Karaolis, M., Siemon, B., Bierkens, M.F.P., 2018. Quantifying Geophysical Inversion Uncertainty Using Airborne Frequency Domain Electromagnetic Data—Applied at the Province of Zeeland, the Netherlands. Water Resour. Res. 54, 8420–8441. https://doi.org/10.1029/2018WR023165

Kirkegaard, C., Sonnenborg, T.O., Auken, E., Jørgensen, F., 2011. Salinity Distribution in Heterogeneous Coastal Aquifers Mapped by Airborne Electromagnetics. Vadose Zo. J. 10, 125. https://doi.org/10.2136/vzj2010.0038

Langevin, C.D., Thorne Jr., D.T., Dausman, A.M., Sukop, M.C., Guo, W., 2007. SEAWAT Version 4: A Computer Program for Simulation of Multi-Species Solute and Heat Transport. U.S. Geol. Surv. Tech. Methods B. 6 39.

Lewis, E.L., Perkin, R.G., 1981. The practical salinity scale 1978: conversion of existing data. Deep Sea Res. Part A. Oceanogr. Res. Pap. 28, 307–328. https://doi.org/10.1016/0198-0149(81)90002-9

Ley-cooper, A.Y., Richardson, M., Ave, J., Ave, J., 2018. AusAEM ; Australia ' s airborne EM continental scale acquisition program AusAEM ; Australia ' s airborne EM continental scale acquisition program. Exploration Geophysics 51 193–202

Li, W. Ben, Zeng, Z.F., Li, J., Chen, X., Wang, K., Xia, Z., 2016. 2.5D forward modeling and inversion of frequency-domain airborne electromagnetic data. Appl. Geophys. 13, 37–47. https://doi.org/10.1007/s11770-016-0548-y

López, A., Arroquy, J.I., Hernández, O., Nasca, J.A., Juárez Sequeira, A. V., DiLorenzo, N., Distel, R.A., 2021. A meta-analytical evaluation of the effects of high-salt water intake on beef cattle. J. Anim. Sci. 99. https://doi.org/10.1093/JAS/SKAB215

Maas, E. V., Hoffman, G.J., 1977. Crop Salt Tolerance—Current Assessment. J. Irrig. Drain. Div. 103, 115–134. https://doi.org/10.1061/JRCEA4.0001137

Mabrouk, M., Jonoski, A., Oude Essink, G.H.P., Uhlenbrook, S., 2019. Assessing the Fresh-Saline Groundwater Distribution in the Nile Delta Aquifer Using a 3D Variable-Density Groundwater Flow Model. https://doi.org/10.3390/w11091946

Macmanus, K., Balk, D., Engin, H., Mcgranahan, G., Inman, R., 2021. Estimating population and urban areas at risk of coastal hazards, 1990-2015: How data choices matter. Earth Syst. Sci. Data 13, 5747–5801. https://doi.org/10.5194/ESSD-13-5747-2021

Macnae, J., King, A., Stolz, N., Osmakoff, A., Blaha, A., 1998. Fast AEM data processing and inversion. Explor. Geophys. 29, 163–169.

Marquardt, D.W., 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. J. Soc. Ind. Appl. Math. https://doi.org/10.1137/0111030

McNeil, J. D. (1980). Electrical conductivity of soils and rocks, (p. 20). Missisauga, Ont: Geonics Ltd.

Menke, W., 1989. Geophysical data analysis: discrete inverse theory. Acad. Press San Diego.

Meyer, R., Engesgaard, P., Sonnenborg, T.O., 2019. Origin and Dynamics of Saltwater Intrusion in a Regional Aquifer: Combining 3-D Saltwater Modeling With Geophysical and Geochemical Data. Water Resour. Res. 55, 1792–1813. https://doi.org/10.1029/2018WR023624

Michael, H.A., Voss, C.I., Gossel, W., Sefelnasr, A., Wycisk, P., Nocchi, M., Salleolini, M., 2009. A 3D density-dependent model for assessment and optimization of water management policy in a coastal carbonate aquifer exploited for water supply and fish farming. Hydrogeol. J. 18, 200–218. https://doi.org/10.1007/s10040-010-0597-x

Minderhoud, P.S.J., Erkens, G., Van Hung, P., Vuong, B.T., Erban, L.E., Kooi, H., Stouthamer, E., 2017. Impacts of 25 years of groundwater extraction on subsidence in the Mekong delta, Vietnam. Environ. Res. Lett. 12, 13. https://doi.org/10.1088/1748-9326/aa7146

Minsley, B.J., 2011. A trans-dimensional Bayesian Markov chain Monte Carlo algorithm for model assessment using frequency-domain electromagnetic data. Geophys. J. Int. 187, 252–272. https://doi.org/10.1111/j.1365-246X.2011.05165.x

Minsley, B.J., Foks, N.L., Bedrosian, P.A., 2020. Quantifying model structural uncertainty using airborne electromagnetic data. Geophys. J. Int. 224, 590–607. https://doi.org/10.1093/GJI/GGAA393

Mühlenbein, H., Schomisch, M., Born, J., 1991. The parallel genetic algorithm as function optimizer. Parallel Comput. 17, 619–632. https://doi.org/10.1016/S0167-8191(05)80052-3

Mulè, S., Miller, R., Carey, H., Lockwood, R., 2019. Review of three airborne EM systems. https://doi.org/10.1071/ASEG2012ab352

Nelder, J.A., Mead, R., 1965. A Simplex Method for Function Minimization. Comput. J. 7, 308–313. https://doi.org/10.1093/comjnl/7.4.308

Neumann, B., Vafeidis, A.T., Zimmermann, J., Nicholls, R.J., 2015. Future coastal population growth and exposure to sea-level rise and coastal flooding - A global assessment. PLoS One 10. https://doi.org/10.1371/journal.pone.0118571

Nocchi, M., Salleolini, M., 2013. A 3D density-dependent model for assessment and optimization of water management policy in a coastal carbonate aquifer exploited for water supply and fish farming. J. Hydrol. 492, 200–218. https://doi.org/10.1016/j.jhydrol.2013.03.048

Oude Essink, G.H.P., 2001. Salt Water Intrusion in a Three-dimensional Groundwater System in The Netherlands: A Numerical Study. Transp. Porous Media 2001 431 43, 137–158. https://doi.org/10.1023/A:1010625913251

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, 1–16. https://doi.org/10.1029/2009WR008719

Paine, J., Minty, B., 2005. 11 AIRBORNE HYDROGEOPHYSICS JEFFREY G. PAINE 1 and BRIAN R. S. MINTY 2 1. Water 333–357.

Pauw, P.S., van Baaren, E.S., Visser, M., de Louw, P.G.B., Essink, G.H.P.O., 2015a. Increasing a freshwater lens below a creek ridge using a controlled artificial recharge and drainage system: a case study in the Netherlands. Hydrogeol. J. 23, 1415–1430. https://doi.org/10.1007/s10040-015-1264-z

Post, V.E.A., 2005. Fresh and saline groundwater interaction in coastal aquifers: Is our technology ready for the problems ahead? Hydrogeol. J. 13, 120–123. https://doi.org/10.1007/s10040-004-0417-2

Post, V.E.A., Oude Essink, G., Szymkiewicz, A., Bakker, M., Houben, G., Custodio, E., Voss, C., 2018. Celebrating 50 years of SWIMs (Salt Water Intrusion Meetings). Hydrogeol. J. 2018 266 26, 1767–1770. https://doi.org/10.1007/S10040-018-1800-8

Pryet, A., Ramm, J., Chilès, J.P., Auken, E., Deffontaines, B., Violette, S., 2011. 3D resistivity gridding of large AEM datasets: A step toward enhanced geological interpretation. J. Appl. Geophys. 75, 277–283. https://doi.org/10.1016/j.jappgeo.2011.07.006

Rahman, M.A., Zhao, Q., Wiederhold, H., Skibbe, N., Eva González, ·, Deus, N., Siemon, · Bernhard, Kirsch, · Reinhard, Elbracht, J., González, E., Siemon, B., Kirsch, R., n.d. Coastal groundwater systems: mapping chloride distribution from borehole and geophysical data. Grundwasser. https://doi.org/10.1007/s00767-021-00475-1

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, 421–443. https://doi.org/10.5194/hess-17-421-2013

Reid, J.E., Pfaffling, A., Vrbancich, J., 2006a. Airborne electromagnetic footprints in 1D earths. Geophysics 71. https://doi.org/10.1190/1.2187756

Revil, A., Coperey, A., Shao, Z., Florsch, N., Fabricius, I.L., Deng, Y., Delsman, J.R., Pauw, P.S., Karaoulis, M., de Louw, P.G.B., van Baaren, E.S., Dabekaussen, W., Menkovic, A., Gunnink, J.L., 2017. Complex conductivity of soils. Water Resour. Res. 53, 7121–7147. https://doi.org/10.1002/2017WR020655

Revil, A., Glover, P.W.J., 1998. Nature of surface electrical conductivity sandstones, and clays in natural sands. Geophys. Res. Lett. 25, 691–694.

Sambo, C., Iferobia, C.C., Babasafari, A.A., Rezaei, S., Akanni, O.A., 2020. The Role of Time Lapse(4D) Seismic Technology as Reservoir Monitoring and Surveillance Tool: A Comprehensive Review. J. Nat. Gas Sci. Eng. 80, 103312. https://doi.org/10.1016/J.JNGSE.2020.103312

Scheelbeek, P.F.D., Chowdhury, M.A.H., Haines, A., Alam, D.S., Hoque, M.A., Butler, A.P., Khan, A.E., Mojumder, S.K., Blangiardo, M.A.G., Elliott, P., Vineis, P., 2017. Drinking Water Salinity and Raised Blood Pressure: Evidence from a Cohort Study in Coastal Bangladesh. Environ. Health Perspect. 125. https://doi.org/10.1289/EHP659

Scheunert, M., Ullmann, A., Afanasjew, M., Börner, R.U., Siemon, B., Spitzer, K., 2016. A cut-&-paste strategy for the 3-D inversion of helicopter-borne electromagnetic data - I. 3-D inversion using the explicit Jacobian and a tensor-based formulation. J. Appl. Geophys. 129, 209–221. https://doi.org/10.1016/j.jappgeo.2016.03.023

Sengpiel, K.P., and Meiser, P., 1981. Locating the freshwater/ salt water interface on the island of Spiekeroog by airborne EM resistivity/depth mapping. Geol. Jahrb. C29, 255–271.

Sengpiel, K., 1989. Results of airborne EM groundwater exploration in a desert area using the centroid depth algorithm for evaluation 70.

Sengpiel, K., & Fluche, B. (1992). Application of airborne electromagnetics to groundwater exploration in Pakistan. Zeitschrift der Deutschen Geologischen Gesellschaft Band, 143, 254– 261.

Sengpiel, K., Siemon, B., 2000. Advanced inversion methods for airborne electromagnetic exploration. Geophysics 65, 1983–1992. https://doi.org/10.1190/1.1444882

Sengpiel, K.P., 1981. New possibilities for groundwater exploration using airborne electromagnetics. Proc. 9th Salt Water Intrusion Meet.

Siemon, B., 2012. Accurate 1D forward and inverse modeling of high-frequency helicopter-borne electromagnetic data. Geophysics 77, WB71. https://doi.org/10.1190/geo2011-0371.1

Siemon, B., 2009. Electromagnetic methods – frequency domain. Groundw. Geophys. 155–178. https://doi.org/10.1007/978-3-540-88405-7_5

Siemon, B., Auken, E., Christiansen, A.V., 2009a. Laterally constrained inversion of helicopter-borne frequency-domain electromagnetic data. J. Appl. Geophys. 67, 259–268. https://doi.org/10.1016/j.jappgeo.2007.11.003

Siemon, B., Christiansen, A.V., Auken, E., 2009b. A review of helicopter-borne electromagnetic methods for groundwater exploration. Near Surf. Geophys. 7, 629–646. https://doi.org/10.3997/1873-0604.2009043

Siemon, B., Costabel, S., Voß, W., Meyer, U., Deus, N., Elbracht, J., Günther, T., Wiederhold, H., 2015. Airborne and ground geophysical mapping of coastal clays in Eastern Friesland, Germany. Geophysics 80, WB21–WB34. https://doi.org/10.1190/geo2014-0102.1

Siemon, B., Steuer, A., Ullmann, A., Vasterling, M., Voß, W., 2011. Application of frequency-domain helicopter-borne electromagnetics for groundwater exploration in urban areas. Phys. Chem. Earth 36, 1373–1385. https://doi.org/10.1016/j.pce.2011.02.006

Siemon, B., van Baaren, E., Dabekaussen, W., Delsman, J., Dubelaar, W., Karaoulis, M., Steuer, A., 2018. Automatic identification of fresh-saline groundwater interfaces from airborne EM data in Zeeland, the Netherlands. Near Surf. Geophys. 3–25. https://doi.org/10.1002/nsg.12028

Simmons, C.T., Bauer-Gottwein, P., Graf, T., Kinzelbach, W., Kooi, H., Li, L., Post, V.E.A., Prommer, H., Therrien, R., Voss, C.I., Ward, J., Werner, A.D., 2010. Variable density groundwater flow: from modelling to applications, Journal of Chemical Information and Modeling. https://doi.org/10.1017/CBO9781107415324.004

Spies, B.R., 1996. Electrical and electromagnetic borehole measurements: A review. Surv. Geophys. 17, 517–556. https://doi.org/10.1007/BF01901643

Stanley, D.J., Warne, A.G., 1997. Holocene Sea-Level Change and Early Human Utilization of Deltas.

Steklova, K., Haber, E., 2015. Joint Hydrogeophysical Inversion : State Estimation for Seawater Intrusion Models in 3D. J. Hydrol. https://doi.org/10.1007/s10596-016-9595-y

Steuer, A., Siemon, B., Auken, E., 2009. A comparison of helicopter-borne electromagnetics in frequency- and time-domain at the Cuxhaven valley in Northern Germany. J. Appl. Geophys. 67, 194–205. https://doi.org/10.1016/j.jappgeo.2007.07.001

Stuyfzand, P., 1986. A new hydrochemical classification of watertypes: principles and application to the coastal dunes aquifer system of the Netherlands. Proc. Ninth Salt Water Intrusion Meet. 641–655.

Sulzbacher, H., Wiederhold, H., Siemon, B., Grinat, M., Igel, J., Burschil, T., G??nther, T., Hinsby, K., 2012. Numerical modelling of climate change impacts on freshwater lenses on the North Sea Island of Borkum using hydrological and geophysical methods. Hydrol. Earth Syst. Sci. 16, 3621–3643. https://doi.org/10.5194/hess-16-3621-2012

Swift, C.M., 1988. 1. Fundamentals of the Electromagnetic Method. Investig. Geophys. 4–11. https://doi.org/10.1190/1.9781560802631.CH1

Tikhonov, A. N., & Arsenin, V. Y. (1979). Solutions of ill-posed problems. Society for Industrial and Applied Mathematics, 21, 266– 267.

Tølbøll, R. J., & Christensen, N. B. (2007). The sensitivity functions of frequency-domain magnetic dipole-dipole systems. Geophysics, 72(2), F45– F56.

Van Baaren, E.S., Oude Essink, G.H.P., Janssen, G.M.C.M., De Louw, P.G.B., Heerdink, R., Goes, B.J.M., 2016. Verzoeting en verzilting freatisch grondwater in de Provincie Zeeland; Zeeland model: 3D regionaal zoet-zout grondwater, Deltares report 1220185.

Van Baaren, E. S., Delsman, J. R., Karaoulis, M., Pauw, P. S., Vermaas, T., Bootsma, H., et al. (2018). FRESHEM Zeeland - FREsh Salt groundwater distribution by Helicopter ElectroMagnetic survey in the Province of Zeeland. Deltares report 1209220. Utrecht, Netherlands.

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

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

Van Weert, F., van der Gun, J., Reckman, J., 2009. Global Overview of Saline Groundwater Occurrence and Genesis, IGRAC, GP 2009-1.

Vandevelde, D., Van Baaren, E., Delsman, J., Karaoulis, M., Oude Essink, G., De Louw, P., Vermaas, T., Pauw, P., De Kleine, M., Thofte, S., Teilmann, R., Walraevens, K., Van Camp, M., Dominique, H., Dabekaussen, W., Gunnink, J., Vandenbohede, A., 2018. Groundwater salinity mapping of the Belgian coastal zone to improve local freshwater storage availability. E3S Web Conf. 54. https://doi.org/10.1051/e3sconf/20185400040

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

Vest Christiansen, A., Auken, E., 2012. A global measure for depth of investigation. Geophysics 77, WB171. https://doi.org/10.1190/geo2011-0393.1

Viezzoli, A., Auken, E., Munday, T., 2009. Spatially constrained inversion for quasi 3D modelling of airborne electromagnetic data an application for environmental assessment in the Lower Murray Region of South Australia. Explor. Geophys. 40, 173–183. https://doi.org/10.1071/EG08027

Viezzoli, A., Christiansen, A. V., Auken, E., & Sørensen, K. (2008). Quasi-3D modeling of airborne TEM data by spatially constrained inversion. Geophysics, 73(3), F105– F113.

Viezzoli, A., Selfe, G., 2018. Comparing responses from different AEM systems and derived models at the Sunnyside nickel project, Botswana. https://doi.org/10.1071/ASEG2018abW8_2E

Vignoli, G., Fiandaca, G., Christiansen, A.V., Kirkegaard, C., Auken, E., 2015. Sharp spatially constrained inversion with applications to transient electromagnetic data. Geophys. Prospect. 63, 243–255. https://doi.org/10.1111/1365-2478.12185

Vos, P. (2015). Origin of the Dutch coastal landscape: Long-term landscape evolution of the Netherlands during the Holocene, described and visualized in national, regional and local palaeogeographical map series. GRONINGEN: Barkhuis. https://doi.org/10.2307/j.ctt2204s8d

Wagner, B.J., 1995. Recent advances in simulation-optimization groundwater management modeling. Rev. Geophys. 33, 1021–1028. https://doi.org/10.1029/95RG00394

Wang, P.C., Shoup, T.E., 2011. Parameter sensitivity study of the Nelder-Mead Simplex Method. Adv. Eng. Softw. 42, 529–533. https://doi.org/10.1016/j.advengsoft.2011.04.004

Waxman, M.H., Smits, L.J.M., 2003. Electrical Conductivities in Oil-Bearing Shaly Sands. SPE Repr. Ser. 8, 107–122. https://doi.org/10.2118/1863-a

Werner, A.D., 2009. A review of seawater intrusion and its management in Australia. Hydrogeol. J. 2009 181 18, 281–285. https://doi.org/10.1007/S10040-009-0465-8

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

Willet, J., King, J., Wetser, K., Dykstra, J.E., Oude, G.H.P., Rijnaarts, H.H.M., 2020. Water supply network model for sustainable industrial resource use a case study of Zeeuws-Vlaanderen in the Netherlands. Water Resour. Ind. 24, 100131. https://doi.org/10.1016/j.wri.2020.100131

Wynn, J., 2005. Mapping Ground Water in Three Dimensions — An Analysis of Airborne Geophysical Surveys of the Upper San Pedro River Basin , Cochise County , Southeastern Arizona, U.S. Geological Survey Professional Paper 1674, one CD-ROM. available at http://pubs.usgs.gov/pp/2006/1674/.

Yin, C., Hodges, G., 2007. Simulated Annealing For Airborne EM Data Interpretation. Fifth Decenn. Int. Conf. Miner. Explor. 72, 843–846. https://doi.org/10.1190/1.2736195

Yin, C., Huang, X., Liu, Y., Cai, J., 2014. Footprint for frequency-domain airborne electromagnetic systems. Geophysics 79, E243–E254. https://doi.org/10.1190/geo2014-0007.1

Zhou, H., Gómez-hernández, J.J., Li, L., 2014. Advances in Water Resources Inverse methods in hydrogeology : Evolution and recent trends. Adv. Water Resour. 63, 22–37. https://doi.org/10.1016/j.advwatres.2013.10.014

Zuurbier, K.G., Kooiman, J.W., Groen, M.M.A., Maas, B., Stuyfzand, P.J., 2015. Enabling Successful Aquifer Storage and Recovery of Freshwater Using Horizontal Directional Drilled Wells in Coastal Aquifers. J. Hydrol. Eng. 20. https://doi.org/10.1061/(asce)he.1943-5584.0000990
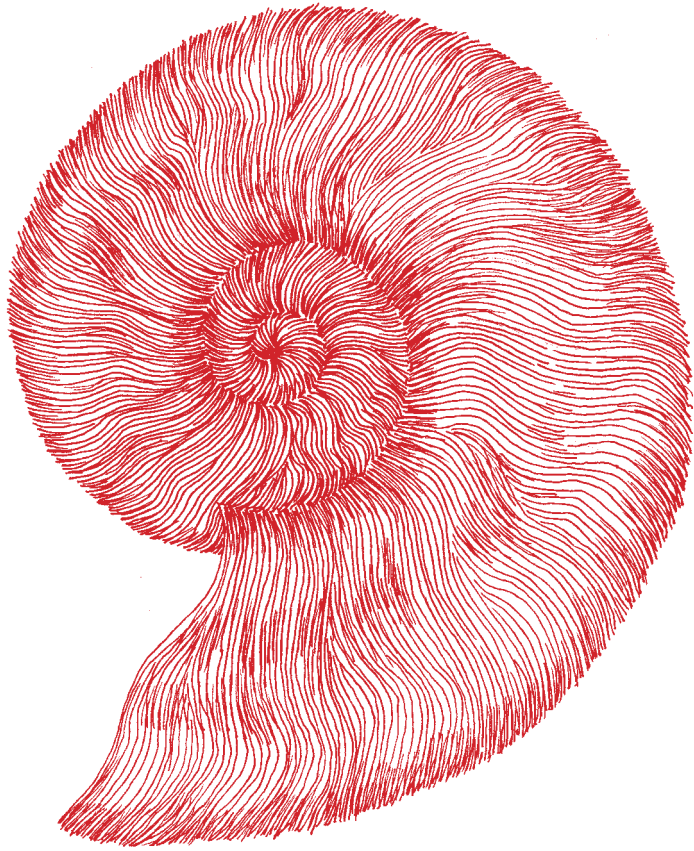
7

Summary

Low elevation coastal zones (LECZs), defined here as areas ≤10 m above mean sea-level, have attracted people for millennia. With their abundant resources and access to trading ports, today these areas host nearly 800 million people – a figure that is predicted to rise to 1.4 billion by 2060. Naturally, it follows that with population growth comes an increased demand for freshwater. Globally, about 50% of the world's population rely on groundwater to satisfy basic requirements – and LECZs are no exception. However, owing to anthropogenic activity, these aquifers are highly stressed and vulnerable to saltwater intrusion – where freshwater can be displaced by saline groundwater. As a result, aquifers within LECZs require effective management, which in turn requires an excellent regional understanding of fresh-saline groundwater distributions. Airborne electromagnetic (AEM) surveys offer a rapid and cost-effective method to map this, and thus are increasingly used for these purposes. Despite increasing popularity, AEM is relatively poorly understand in terms of regional (provincial or country scale) groundwater salinity mapping. The primary objective of this thesis is therefore twofold: 1) to better understand the uncertainties involved and 2), use this understanding to develop novel mapping methods. Consequently, the following research questions were formulated:

1. *What is the effect of using different inversion methods and parameters on mapping results?*
2. *How are results affected by different quantities of available data?*
3. *Based on the results of chapters 2 and 3, what further methodological improvements can we make?*
4. *Are groundwater salinity movements sensitive to repeated AEM surveys?*

To answer these research questions, I used data from the Province of Zeeland, The Netherlands. Zeeland experienced sea-level transgressions in the early Holocene, followed by the construction of by man-made coastal defences – which allowed the recent freshening of shallow aquifers. As a result, much of Zeeland comprises shallow rainwater lenses (often as little as 1 – 2m thick), and therefore offers a fascinating study area for applied groundwater research. Furthermore, a recently undertaken, high quality AEM survey and plentiful ground-based data provide an ideal testing ground. Throughout the thesis, these data were used either directly, or as basis for the construction of highly detailed synthetic models.

The process of converting AEM observations into usable, quantitative data, is referred to geophysical inversion. Typically, this process involves iteratively adjusting a model in an optimisation routine until a suitable fit is found against observations. The results are typically represented in SI units as electrical conductivity (EC), which in terms of groundwater salinity mapping, relate to the locations of fresh-brackish-saline groundwater and lithology. However, inversions are deterministic (i.e., the result is a single estimated model) – this is potentially problematic as an infinite number of models can explain the data. In order to understand these uncertainties, I quantitatively compared four commonly used inversion methods and associated input parameters. Results of each were interpolated into 3D volumes of EC and

transformed into groundwater salinity, where extensively available ground data were used for comparisons. Results indicated that all methods offered similar accuracy, however fresh-brackish-saline groundwater interfaces were frequently mapped too shallow because of smoothing introduced by inversions.

With a quantitative understanding of inversions and groundwater mapping using AEM, it was now possible to examine other areas of uncertainty. Inversions result in models of EC, which is the combined effect of lithology and groundwater salinity, otherwise known as bulk electrical conductivity (ECb). In order to recover the EC of groundwater (ECw), which is ultimately converted to groundwater salinity using empirical formulae, lithological information is required. This data is expensive and is therefore often sparse relative to the regional scale of AEM mapping programmes. As a result, detailed 3D synthetic models were used to quantitatively examine the effect using different amounts of lithological data to recover 3D salinity models. Survey design is also an important financial consideration, therefore the effects of flightline spacing were examined too. By comparing almost 100 different 3D models comprising different levels of lithological information and flightline spacing, it was found that error introduced by the inversion process exceeded that of poorly understood lithology. It was also noted that decreasing flightline spacing consistently improved mapping accuracy.

Based on the understanding that inversions introduced groundwater salinity mapping error because of overly smooth results, a practical inversion approach was developed to try and control the level of smoothing based on available ground data. The method exploits in-built inversion parameters of a commonly used method to match the known vertical thickness of a brackish zone. In order to balance the effect of matching ground data while not introducing noise, a penalisation term can be set to favour either ground data or default inversion parameters. The method was initially tested using a synthetic case, the results of which guided a practical application using real data. Using available validation data, the results of the real case highlighted the potential of the method – however it was concluded that more research is still needed. In particular, the potential of coupled hydrogeophysical inversion was highlighted as a possible area for further research – which is essentially the geophysical modelling of groundwater simulations.

The final chapter of this thesis builds on the conclusion that coupled hydrogeophysical inversion offers an interesting research direction. More specifically, we investigated the novel idea of using repeat AEM surveys to jointly estimate groundwater salinity distributions while improving the parameterisation of 3D variable density groundwater flow and salt transport models (3D-VDG). The approach was developed and tested using a highly detailed 3D synthetic model. Here, AEM data were first coupled with 3D-VDG simulations using a petrophysical transformation and geophysical forward modelling. After running the 3D-VDG model for 60 years it was concluded that after ~5 or more years, it was possible

to monitor groundwater salinity changes over time. An optimisation routine was then implemented that minimises the differences between a simulated and an AEM-observed salinity distribution based on AEM surveys flown over the same area at two points in time. This was done by iterating 3D-VDG model runs with varying hydrogeological parameters over the period between the two surveys, where at the final time step the resulting model results were compared against AEM observations. The 3D-VDG model then restarted with updated hydrogeological parameters until a suitable minima was found. The method was shown to successfully estimate hydrogeological parameters while simultaneously resolving a realistic distribution of groundwater salinity.

Samenvatting

Low elevation coastal zones (LECZ's), hier gedefinieerd als gebieden 10 m boven zeeniveau, trekken al millennia mensen aan. Met hun overvloedige hulpbronnen en toegang tot handelshavens, herbergen deze gebieden tegenwoordig bijna 800 miljoen mensen - een cijfer dat naar verwachting zal stijgen tot 1,4 miljard in 2060. Uiteraard volgt hieruit dat met de bevolkingsgroei een grotere vraag naar zoet water komt. Wereldwijd is ongeveer 50% van de wereldbevolking afhankelijk van grondwater om aan basisbehoeften te voldoen - en LECZ's zijn geen uitzondering. Vanwege antropogene activiteit zijn deze watervoerende lagen kwetsbaar voor zoutwaterindringing - waar zoet water kan worden verdrongen door zout grondwater. Als gevolg hiervan vereisen aquifers binnen LECZ's effectief beheer, wat op zijn beurt een goed regionaal begrip van zoet-zoute grondwaterdistributies vereist. Airborne electromagnetic (AEM) surveys bieden een snelle en kosteneffectieve methode om dit in kaart te brengen en worden daarom steeds vaker voor deze doeleinden gebruikt. Ondanks de toenemende populariteit is AEM relatief slecht begrepen in termen van regionale (provinciale of landelijke schaal) grondwatersaliniteitskaarten. Het primaire doel van dit proefschrift is daarom tweeledig: 1) de onzekerheden beter begrijpen en 2) deze kennis gebruiken om nieuwe karteringmethoden te ontwikkelen. Daarom zijn de volgende onderzoeksvragen geformuleerd:

1. Wat is het effect van het gebruik van verschillende inversiemethoden en parameters op de resultaten van de kartering?
2. Hoe worden resultaten beïnvloed door verschillende hoeveelheden beschikbare gegevens?
3. Welke verdere methodologische verbeteringen kunnen we op basis van de resultaten van hoofdstuk 2 en 3 maken?
4. Zijn grondwaterverziltingsbewegingen gevoelig voor herhaald AEM-onderzoek?

Om deze onderzoeksvragen te beantwoorden, heb ik gegevens uit de provincie Zeeland gebruikt. Zeeland had te maken met zeespiegeloverschrijdingen in het vroege Holoceen, gevolgd door de aanleg van door de mens gemaakte kustverdediging - die de recente verzoeting van ondiepe watervoerende lagen mogelijk maakte. Hierdoor bestaat een groot deel van Zeeland uit ondiepe regenwaterlenzen (vaak slechts 1 à 2 meter dik), en biedt daarmee een boeiend studiegebied voor toegepast grondwateronderzoek. Bovendien bieden een recent uitgevoerd, hoogwaardig AEM-onderzoek en overvloedige grondgebaseerde gegevens een ideale proeftuin. In het proefschrift werden deze gegevens ofwel direct gebruikt, ofwel als basis voor de constructie van zeer gedetailleerde synthetische modellen.

Het proces van het omzetten van AEM-waarnemingen in bruikbare, kwantitatieve gegevens, wordt geofysische inversie genoemd. Meestal houdt dit het proces in van het iteratief aanpassen van een model in een optimalisatieroutine totdat een geschikte fit is gevonden

tegen waarnemingen. De resultaten worden meestal weergegeven in SI-eenheden als elektrische geleidbaarheid (EC), die in termen van het in kaart brengen van het zoutgehalte van het grondwater betrekking hebben op de locaties van zoet-brak-zout grondwater en lithologie. Inversies zijn echter deterministisch (d.w.z. het resultaat is een enkel geschat model) - dit is potentieel problematisch omdat een oneindig aantal modellen de gegevens kan verklaren. Om deze onzekerheden te begrijpen, heb ik vier veelgebruikte inversiemethoden en bijbehorende invoerparameters kwantitatief vergeleken. De resultaten van elk werden geïnterpoleerd in 3D-volumes van EC en omgezet in het zoutgehalte van het grondwater, waarbij uitgebreid beschikbare grondgegevens werden gebruikt voor vergelijkingen. De resultaten gaven aan dat alle methoden een vergelijkbare nauwkeurigheid boden, maar zoet-brak-zout-grondwaterinterfaces werden vaak te ondiep in kaart gebracht vanwege een afvlakking, geïntroduceerd door de inversies.

Met een kwantitatief begrip van inversies en grondwaterkartering met behulp van AEM, was het ook mogelijk om andere oorzaken van onzekerheid te onderzoeken. Inversies resulteren in modellen van EC, wat het gecombineerde effect is van lithologie en zoutgehalte van het grondwater, ook wel bekend als bulk elektrische geleidbaarheid (ECb). Om de EC van grondwater (ECw) te bepalen, die uiteindelijk met empirische formules wordt omgezet in grondwaterzoutgehalte, is lithologische informatie nodig. Deze gegevens zijn duur en daarom vaak schaars in verhouding tot de regionale schaal van AEM-kaartprogramma's. Als gevolg hiervan werden gedetailleerde 3D-synthetische modellen gebruikt om het effect kwantitatief te onderzoeken met behulp van verschillende hoeveelheden lithologische gegevens om 3D-zoutgehaltemodellen te herstellen. Naast lithologische gegevens moeten ook de kosten van de AEM inwinning worden meegenomen. Daarom zijn ook de effecten van vluchtlijnafstand onderzocht. Door bijna 100 verschillende 3D-modellen met verschillende niveaus van lithologische informatie en vluchtlijn -afstand te vergelijken, werd ontdekt dat de fout die door het inversieproces werd geïntroduceerd, groter was dan die van slecht begrepen lithologie. Er werd ook opgemerkt dat het verkleinen van de afstand tussen de vluchtlijnen de nauwkeurigheid van de kaart consequent verbeterde.

Op basis van het inzicht dat inversies een fout in het in kaart brengen van het zoutgehalte van het grondwater hebben geïntroduceerd vanwege te gladde resultaten, werd een praktische inversiebenadering ontwikkeld om te proberen het niveau van afvlakking te beheersen op basis van beschikbare in-situ gegevens. De methode maakt gebruik van ingebouwde inversie parameters van een veelgebruikte methode om de bekende verticale dikte van een brakke zone aan te passen. Om het effect van het matchen van grondgegevens in evenwicht te brengen zonder ruis te introduceren is bij de minimalisatie een penalty-term meegenomen waarmee meer of minder gewicht kan worden gegeven aan de in-situ gegevens of aan de standaard inversieparameters. De methode werd aanvankelijk getest met behulp van een synthetische casus, waarvan de resultaten leidden tot een praktische toepassing met behulp

van echte gegevens. Met behulp van beschikbare validatiegegevens lieten de resultaten van het echte geval het potentieel van de methode zien, maar er werd geconcludeerd dat er nog meer onderzoek nodig is. In het bijzonder werd het potentieel van een gekoppelde hydrogeofysische benadrukt als een mogelijk gebied voor verder onderzoek - dit is in feite de geofysische modellering van grondwatersimulaties.Het laatste hoofdstuk van dit proefschrift bouwt voort op de conclusie dat gekoppelde hydrogeofysiche inversies een interessante onderzoeksrichting bieden. Specifiek hebben we het nieuwe idee onderzocht om herhaalde AEM-onderzoeken te gebruiken om gezamenlijk de saliniteitsverdelingen van het grondwater te schatten en tegelijkertijd de parametrisering van 3D-modellen voor grondwaterstroming met variabele dichtheid en zouttransport te verbeteren (3D-VDG). De aanpak is ontwikkeld en getest met behulp van een zeer gedetailleerd 3D-synthetisch model. Hier werden AEM-gegevens eerst gekoppeld aan 3D-VDG-simulaties met behulp van een petrofysische transformatie en geofysische voorwaartse modellering. Na 60 jaar te hebben gesimuleerd met het 3D-VDG-model, werd geconcludeerd dat het na ~5 jaar of meer mogelijk was om veranderingen in het zoutgehalte van het grondwater in de loop van de tijd te volgen. Vervolgens werd een optimalisatieroutine geimplementeerd waarbij het verschil tussen gesimuleerde en AEM-waargenomen zoutverdelingen worden geminimaliseerd op basis van AEM-onderzoeken die op twee tijdstippen over hetzelfde gebied zijn gevlogen. Dit werd gedaan door 3D-VDG-modelruns met verschillende hydrogeologische parameters over de tijdsperiode tussen de twee AEM surveys te herhalen, waarbij de modelresultaten bij de laatste tijdstap warden vergeleken met AEM-waarnemingen. Het 3D-VDG-model startte vervolgens opnieuw met bijgewerkte hydrogeologische parameters totdat een geschikt minimum werd gevonden. Er werd aangetoond dat de method hydrogeologische parameters met succes schat en tegelijkertijd een realistische verdeling van het zoutgehalte van het grondwater oplevert.

Acknowledgements

Many individuals, institutions, funding bodies etc. are acknowledged for their part in making this possible.

Firstly, my wife, Liza King – because I believe that I wouldn't be writing these acknowledgements without her support. She joined me on this PhD journey in 2016, when we both moved to the Netherlands. Right up to the day that I write this she has been a rock, despite having stresses of her own – including a research masters and the jobs that followed. Stressful days were not only made bearable – they were often fun. All those evenings sitting with a cheap supermarket beer talking through things got me through countless anxiety ridden moments during the PhD. I could go on for a while, never mind the technical help with programming too, it's safe to say she probably deserves to be added as a co-author.

I would like to thank my supervisors Marc Bierkens and Gualbert Oude Essink for putting their trust in me back in 2016, during the early stages of the Water Nexus programme (to whom I also owe my thanks). Their help was consistent and unwavering, despite it taking two years longer than planned – ultimately spanning the entire length of a global pandemic, three American presidents, and Brexit. Marc Bierkens is especially thanked for his regular conceptual chats, particularly concerning chapters 4 and 5 of this thesis. These discussions helped combine hydro(geo)logy and airborne geophysics not just in a conceptual manner but helped formulate a physical link between the disciplines – which I discovered is no easy feat. Gu is thanked, amongst many other things, for his highly detailed feedback on publication text, sometimes resulting in many months of extra work – which in retrospect was necessary for publishing. His detailed explanations and help with Dutch working culture also played a big part in understanding how to navigate collaboration between other Dutch institutions, including Deltares – where I currently work. My co-authors are all thanked for their input, notably Marios Karaoulis, Bernhard Siemon and Tobias Mulder – all of whom provided valuable input into this thesis.

As far as institutions go, Deltares is thanks for their ongoing support throughout the process, including offering a space to work and a laptop. Their assistance in obtaining the data that was used for this project was invaluable – for that I would also like to thank all members of the 'FreshEM Zeeland' project. Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek (TNO) is thanked for their useful discussions throughout the PhD. I would also like to thank individuals at TNO (amongst others - Jan Gunnink and Willem Jan Zaadnordijk) for allowing me the opportunity to work on interesting airborne electromagnetic projects after my PhD funding had expired. This was an invaluable opportunity to finish off the thesis while still doing something related to my research.

About the author

Jude King was born on February 27*th*, 1984 in Makhanda (formerly Grahamstown) in the Eastern Cape, South Africa. He grew up in a small town near Cape Town called Fish Hoek, where he stayed until he moved to the UK at age 14. He then moved back to Cape Town a few years later, where he graduated from high school with disappointing grades in general.

Feeling inspired, he returned to the UK in 2003 to pursue a career in bartending and shop-keeping – eventually finding himself in the unusual position of trying to sell houses in London during the economic crisis of 2008. It was around this time that he realised he was old enough that his high school grades probably no longer mattered, so he applied for university. His plan worked and he settled on an honours degree course in Geology at Brighton University. Here he graduated with marginally more success than high school in 2011 and accepted a place in a master's degree program at the University of Leeds. It was here that he discovered his love of Northern England pubs and geophysical modelling. He graduated in 2012 and promptly returned to Cape Town, where he worked in airborne geophysical exploration for the mineral and hydrocarbon industry around Africa and the Middle East. It was in Cape Town he met his future wife (Liza) in 2013. In 2016 they decided to move to the Netherlands despite having never been there.

In the meantime, he had applied for a PhD position at Utrecht University in mapping groundwater salinity using airborne electromagnetics. His love of rocks had taken him this far, and despite now living in a country without any, he happily accepted the offer of a PhD position and started in 2016. His four-year PhD programme was completed six years later in 2022. He now works as researcher at Deltares in Utrecht, focusing mainly on geophysics and groundwater modelling. He currently resides in the Utrecht with his wife and cat.

Utrecht University
Faculty of Geosciences
Department of Physical Geography