# Measuring and modelling to optimise a salinity monitoring network for use in the optimal control of flushing

## Case study: Lissertocht catchment

H.P. Hagedooren

TU Delft

# Measuring and modelling to optimise a salinity monitoring network for use in the optimal control of flushing

## Case study: Lissertocht catchment

by

H.P. Hagedooren

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday June 27, 2018 at 13:00 PM.

| | | |
|---|---|---|
| Student number: | 4147898 | |
| Project duration: | September, 2017 – June, 2018 | |
| Thesis committee: | Prof. dr. ir. N.C. van de Giesen, | TU Delft |
| | Dr. ir. M.M. Rutten, | TU Delft |
| | Dr. ir. E. Abraham, | TU Delft |
| | Ir. B.E. Aydin, | TU Delft |
| | Dr. B.M. van Breukelen, | TU Delft |
| | Ir. B. Schilperoort, | TU Delft |

*This thesis is confidential and cannot be made public until June 27, 2018.*



TU Delft

# Preface

This thesis is the result of several months working in the office and doing measurements in the field (see cover image) in order to obtain the master degree Civil Engineering at the Delft University of Technology in the track Water Management. I became very motivated once I clearly defined what I wanted to do in this research.

First, I would like to thank my committee members Boran Aydin, Edo Abraham, Martine Rutten, Nick van de Giesen, Boris Breukelen and Bart Schilperoort for giving feedback and always being available for discussing my work. The door of Boran's office was always open for a good discussion about salinity and flushing polders, but we have also had a lot of fun going into the field throwing a 1000 meter cable in a channel. Edo was always available to give feedback or help with solving mathematical problems. Martine stayed involved in my project even after she left the TU Delft, and stayed very interested in my work. Nick and Boris I did not speak weekly to, but when I asked for feedback they provided me with new insights and ideas. Bart introduced me in the field of Distributed Temperature Sensing and was willing to show the equipment and go with me the first time I installed it.

I would like to thank Gualbert van Oude Essink and Joost Delsman of Deltares for helping me and willing to share their knowledge about salinity problems and the Lissertocht catchment.

I would like to thank Mark Kramer from Hoogheemraadschap Rijnland for giving advice, providing data and helping me with arranging the field work. I would also like to thank Henk Sloof, who helped me arrange the fieldwork. I would like to thank Timo Steenwijk for letting me conduct measurements in his fields.

I would like to thank Tom Bik, who helped me with the DTS measurements and who has done the EC routing also as part of his graduation project at the Hogeschool of Rotterdam. I also would like to thank Martijn Mulder and Ivar Abas, who helped to deploy the DTS cables even when it was freezing outside. A special thanks to my friends and family for their support, lively discussions and coffee breaks.

***H.P. Hagedooren***
*Delft, June 2018*

# Abstract

In deep polders in the Netherlands the dominant mechanism of salinization is via intense seepage through boils. These are local connections between the surface water and a deep saline aquifer. The study area of this research is the Lissertocht. This catchment, which is located inside the Haarlemmermeerpolder, contains a lot of boils that can cause problems for the main land use in this catchment agriculture. During summer the salinity is lowered by flushing the catchment with fresh water from outside the polder.

The objective of this thesis is to measure and model the salt concentrations and their sources in order to optimise the placement of sensors in a salinity monitoring network for optimal flushing in the Lissertocht catchment.

This thesis has provided insights into the detection of boils using different methods (Distribute Temperature Sensing (DTS), EC routing and visual inspection). The measurements show that EC routing is not able to detect all the boils and it has to be combined with DTS measurements or visual inspection. The DTS method was the only method that is able to detect boils located at the bottom of the bigger main channels. Moreover, DTS is able to characterize the boil flux on a qualitative level and also on a quantitative level for a boil that is located near a pump; both have been done in this research. The boil flux is calculated with an energy balance and an 1D advection-diffusion model, the results correspond to values found in literature but the uncertainty of these calculations is high. This was due missing information about the distribution of the temperature and salinity over depth.

Data from CTD divers and sensitivity analysis have shown the influence of various processes on the spatial and temporal distribution of the salinity in the catchment. The sensitivity analysis in SOBEK also shows the areas which are sensitive to flushing, the flushing speed and the main pathways which the water from the inlets take. The complexity of the spatial and temporal distribution of the salinity is reduced to a small system of linear components using the Principal Component Analysis (PCA). For this case study, it was shown that over 90% of the total variance in the system can be captured with three orthogonal linear components. A predictive model is built based on the PCA, this model is able to test the performance of a layout of sensors by reconstructing the principal components with virtual measurements. This shows that with a small number of sensors, a salinity monitoring network can be built that is able to predict the spatial and temporal distribution of the salinity in the catchment. Based on the knowledge from the measurements and modelling, three categories of key locations to monitor can be identified.

The optimisation of the sensor locations shows that a greedy algorithm is a very efficient way to solve the sensor placement problem. When more sensors are added with the greedy algorithm the performance of the monitoring network improves, but after six sensors the additional gain of an extra sensor is small in this catchment. The exhaustive algorithm is not able to create a monitoring network with more than three sensors because, the amount of possibilities quickly becomes too big to compute. However, this algorithm is able to find the optimal solution for a network of three sensors. This solution is used to validate the greedy algorithm and the identified categories of key locations to monitor. The robustness analysis against disturbances shows that a monitoring network with more than three sensors is more robust. Moreover, the robustness analysis shows that a monitoring network is most sensitive to the locations of the boils.

The new insights into methods to detect and measure boils and calculate their fluxes, the spatial and temporal distribution of salinity, reduce complexity of a system with PCA and solving a sensor placement problem can be used to implement a monitoring network in the Lissertocht and for other deep polders in the Netherlands.

# Contents

# 1 Introduction

26% of the Netherlands is located below sea level (Planbureau voor de Leefomgeving), which makes it necessary to reclaim, protect and control areas which are lying below sea level with dikes and other structures. With the invention of the windmill in the $16^{th}$ century, Dutch people were able to create big polders by reclaiming lakes (Hoes and Van de Giesen, 2015). Hoes and Van de Giesen (2015, p. 15) gives the following definition of a polder: "a polder is defined as a level area which has originally been subject to a high groundwater or surface water, and is separated from the surrounding hydrological regime, to be able to control the water levels in the polder". Polders typically have a lower elevation than the surrounding area, this results in a vertical pressure gradient between the polder and the surrounding area. This pressure gradient drives the seepage fluxes which flow into the polder and these seepages fluxes can be fresh or saline. In coastal areas the groundwater is often saline because of sea water intrusion, marine transgressions or sea spray (Stuyfzand and Stuurman, 1994). This can result in saline seepage which can lead to salinization of surface water in polders when both types of water mix. Coastal zones are often the most densely populated areas in the world, in these areas the demand for fresh water resources can easily exceed the availability of these resources (Custodio and Bruggeman, 2013). The saline intrusion rates are expected to rise in the future due to land subsidence, climate change and sea level rise, which will accelerate the salinization of shallow ground water and surface water in polders (Oude Essink et al., 2010). This will make low lying areas in the Netherlands even more vulnerable to salinization. This problem occurs also in other areas of the world, such as in the Godavari Delta in India (BOBBA, 2002), coastal regions in America (Barlow and Reichard, 2010) and in the Po river plane in Italy (Giambastiani et al., 2007).

In deep polders in the Netherlands the dominant mechanism leading to surface water salinization is upward groundwater flow through boils (De Louw et al., 2010). Boils are localized preferential pathways of groundwater. Boils connect (saline) aquifers to the surface which allow groundwater to flow to the surface (De Louw, 2013). Some examples of deep polders where boils contribute for more than 50% of the total salt load are the Schermerpolder, the Noordplaspolder and the Haarlemmermeerpolder (De Louw et al., 2004, 2010; Goudriaan et al., 2011). In De Louw et al. (2010) three different types of seepage in the Noordplaspolder are identified and their concentration is measured:

- Diffuse seepage with an average concentration of 100 mg/L
- Seepage through paleochannel belts with an average concentration of 600 mg/L
- Intense seepage through localized boils with an average concentration of 1100 mg/L

The salt concentrations of the different types of seepage correspond to the depth from where the seepage water originates. In case of the Noordplaspolder, the groundwater from deeper aquifers is more saline because the salinity of the groundwater increases with depth (De Louw et al., 2010). The seepage through boils has an higher concentration because this groundwater originates from deeper and more saline aquifers. The

temperature of the seepage water from the boils is constant over the year and equal to the mean annual groundwater temperature of 11 °C (Goudriaan et al., 2011) because the water originates from a deep aquifer. This temperature signature can be used to detect boils (Goudriaan et al., 2011; Hoes et al., 2009; Vandenbohede et al., 2014). Figure 1.1 shows an schematic overview of the different types of seepage.

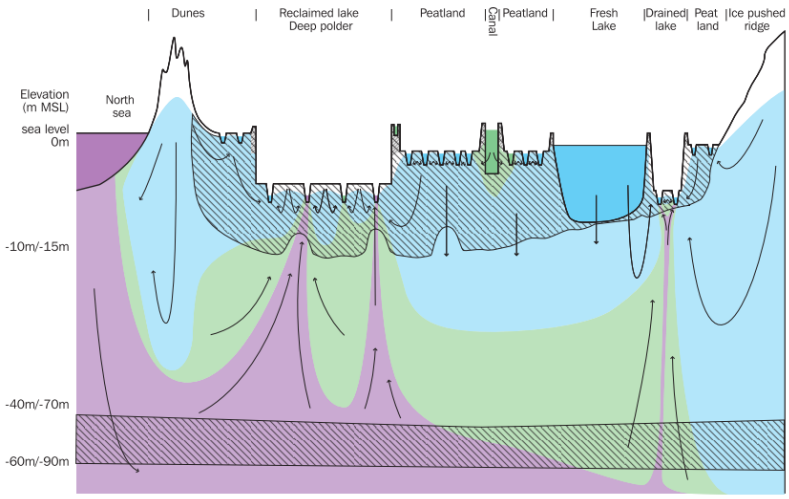



Figure 1.1: Schematisation of seepage through boils in an east-west profile of the Haarlemmermeerpolder. In purple the saline groundwater which originates from the sea ($> 1000 mg/L$), in green the brackish groundwater ($150 - 1000 mg/L$) and in blue fresh groundwater ($< 150 mg/L$). Retrieved from (De Louw, 2013, p. 142)

Boils are formed if there is a vertical pressure gradient between a deep aquifer and the surface which is bigger than the weight of the subsoil between the surface and the deep aquifer, this can lead to the development of a flow path from the deep aquifer to the surface (De Louw, 2013). The locations of boils are related to the following aspects (De Louw et al., 2010):

- The magnitude of the vertical pressure gradient, this means that areas with lower elevation within the polder are more likely to contain boils. This means that it is more likely that boils occur at the bottom of a channel than on land because the bottom of a channel has a lower elevation. Areas near the border of a polder are also more vulnerable because a big difference in elevation.
- Geology of the subsoil, it is most likely that boils will be formed at a location where the resistance of the subsoil is lowest. Areas with locally weaker layers or for example an old river bed (paleochannel) will be more likely to contain boils.

In Figure 1.2a and 1.2b some examples of boils in the study area can be seen. Boils can be identified based on the following properties (De Louw et al., 2010):

- Sediments discharged from boils can form a sand volcano around the boil
- Change in or no vegetation around the boil
- Banks of channels can collapse if a boil is located nearby
- Bubbles
- No ice or holes in the ice because the seepage water from the boils has an higher temperature than the surface water during the winter.
- Area around a boil can be wet and muddy

(a) (b)

Figure 1.2: Example of two boils which are observed in the study area

It is possible to close a boil by injecting gel or bioseal, but field tests have shown that this method is expensive and hard to implement (De Louw et al., 2012). Moreove,r this method only moves the problem to another location because a new boil will be formed after closing one boil (De Louw et al., 2007). To prevent salinization of the surface water many Dutch polders are flushed between the 1st of April and the 1st of October (growing season). The polders are flushed by diverting fresh water from the Meuse and Rhine to the surrounding canal (boezem) of the polder from where the fresh water will enter the polder network via inlet structures. This fresh water will flow through the polder network to the outlet of the polder where a pump pumps the water back into the boezem. Another approach to prevent damage from salinization can be to stop cultivating salt sensitive crops in polders which suffer from saline groundwater exfiltration and cultivate salt tolerant crops instead.

In the Water Nexus project (Water Nexus) a PhD research focusses on the development of a Model Predictive Control (MPC) scheme to efficiently flush a polder (Aydin et al., 2015, 2016). This MPC scheme controls the water levels and salt concentrations in a polder based on real time information from a monitoring network in the polder. Based on this information a MPC scheme can take operational decisions to control the water quantity and water quality in the polder.

This master thesis investigates the possibilities to set up a monitoring network in a polder to support the MPC scheme of the Lissertocht catchment. Based on measurements and modelling results, the sensor placement will be optimised. Moreover, the knowledge about the system will be improved by conducting measurements to update the model. This will give more information and new insights about the salinity problems in the Lissertocht catchment.

## 1.1. Study area

This research will focus on the Lissertocht catchment (peilvlak 9 and 19) which is located in the Haarlemmermeer polder, see Figure 1.3. The waterboard Rijnland is responsible for this catchment. The target level of the surface water is -6.42 meter NAP in summer (April - October) and -6.59 meter NAP during winter (October - April) (Hoogheemraadschap Rijnland). The size of the catchment is 10 $km^2$ and the total length of the network of channels is 72 km (Nationaal Georegister, 2017). The mean annual sum of precipitation in this area is 808 millimeter (KNMI). This research will focus on peilvlak 19 because this part has problems with salinity. The other part 'peilvlak 9' is added because one of the inlets is located here, the two catchments are connected with a weir. The smaller (0.6 $km^2$) part does not have any problems with salinities because there are no boils in this part. The catchment is controlled with two pumps; Dr. J.P. Heyepad (capacity of 0.42 $m^3/s$) and Dr. J.P. Heye (capacity of 1.48 $m^3/s$). Next to the pump J.P. Heye a sensor is located which measures the water level, the pump switches on and off automatically based on the readings of this sensor, typically twice a day. The pump J.P. Heyepad is only used less frequently, only if there is a lot of precipitation. The area has four inlets which can be used to divert fresh water from the Ringvaart to the polder.

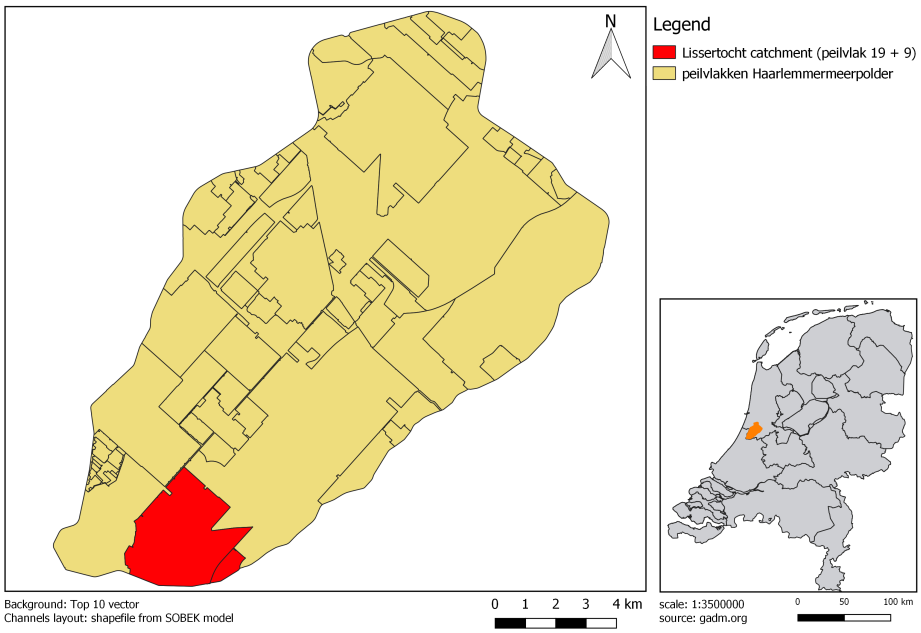



Figure 1.3: Location of Lissertocht catchment

The reclamation of the Haarlemmermeerpolder was completed in July 1852 with steam powered pumps. Until World War 2 the main land use in the Haarlemmermeerpolder was agriculture. After World War 2 the Haarlemmermeerpolder residential and business area developed and agriculture decreased; an important driver for this change was the opening of Schiphol airport in 1967 (Gemeente Haarlemmermeer). The Lissertocht is an exception to this trend, nowadays the main land use in this catchment is still agriculture. Crops which are being cultivated are potatoes, grassland, corn, beets and flowers (Alterra, 2015). This is one of the main reasons why this research is very relevant for this catchment. The elevation ranges between -5.6 meter NAP and -4.4 meter NAP, see Figure 1.4, this is lower than the rest of the Haarlemmermeerpolder. This is a factor which increases the risks on development of boils.

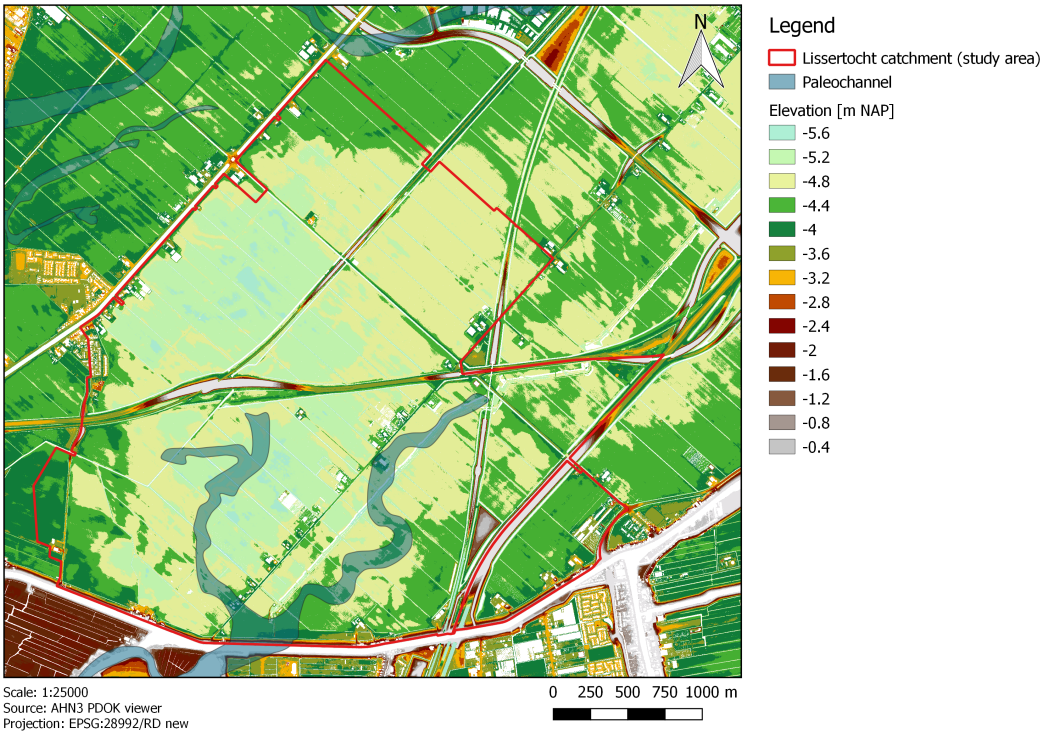



Figure 1.4: Elevation of the Lissertocht catchment and location of the paleochannel (old river bed)

The history of the geology of the Haarlemmermeerpolder explains the origin of the saline aquifer below the study area. Delsman et al. (2014) describes the evolution of this saline aquifer from the the end of the Pleistocene (13000 BC), when the area was characterized by sandy plains with braided rivers and ice-pushed ridges, until present-day (2000 AD). In this time period paleo-geographical changes resulted in marine transgression which formed the deep saline aquifer below the Haarlemmermeerplder.

## 1.2. Current knowledge

Earlier studies led to models which predict the chloride concentrations and seepage in this catchment. A Rapid Saline Groundwater Exfiltration Model (RSGEM) is developed to calculate groundwater exfiltration into the fields and ditches (Delsman et al., 2017). This is a lumped model which describes the flow of water and salt from agricultural fields to the surface water. The model was first calibrated and applied to the Schermer polder but later also to the Lissertocht catchment (Kelderman, 2015). Kelderman (2015) developed a SOBEK model which calculates the water levels and salt concentrations in the area. Delsman et al. (2013) conducted an end member mixing analysis on daily samples taken at the outlet (pump) of the Lissertocht catchment. This resulted in five different water types, see Figure 1.5. Kelderman (2015) modelled the different fractions of the these different water types with the SOBEK model which makes it possible to trace the different sources of water through the catchment.

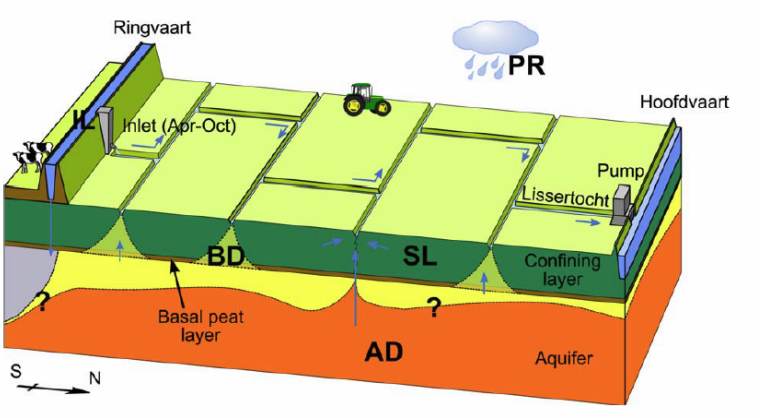



Figure 1.5: Schematic scheme of flow paths of different contributions in the Lissertocht catchment.
IL: inlet water, PR: precipitation, BD: roundwater below ditches, SL: shallow, phreatic groundwater, AD: deep aquifer groundwater (boils)
Source: (Delsman et al., 2013)

Research has been done to optimise sensor placement in sewer networks and drinking water distribution networks (Krause et al., 2008), (Eliades et al., 2014), (Banik et al., 2015), (Dorini, G., Jonkergouw, P., Kapelan, Z., di Pierro, F., Khu, S., and Savic, 2006) and (Hart and Murray, 2010). These papers all apply the optimisation of sensor placement to pressurized pipes (drinking water distribution network) or unpressurized pipes (sewer network). There are two papers written about optimizing a monitoring network of water level stations in the Delfland polder (Alfonso et al., 2010a,b). In this research the information theory is used to calculate entropy and total correlation in the catchment to determine the best locations to place sensors which monitor the water level.

To the knowledge of the author there is no literature on optimizing a monitoring network to control the salinity of surface water in a polder system. In Table 1.1 this knowledge gap is summarised.

Table 1.1: Knowledge gap optimisation

| **Optimisation of sensor placement** | **Polder system** | **Water distribution network** |
|---|---|---|
| Water quality | Knowledge gap | Yes, contamination detection |
| Water quantity | Yes, water level monitoring | Yes, leakage detection |

Before optimizing the sensor placement in a monitoring network it is necessary to have a good understanding of the dynamics and process in the Lissertocht catchment. Because the boils are the dominant source of salinization of the surface water it is important to locate the boils and place them in the model. In this research different methods are used and will be compared to locate the boils in a catchment.

## 1.3. Objective and research questions

The objective of this study is:

*Measuring and modelling of the salt concentrations and their sources in order to optimise the placement of sensors in a salinity monitoring network for optimal flushing in the Lissertocht catchment.*

The research questions are:

1. How can the boils in this catchment be detected and measured?
2. What is the spatial and temporal distribution of the salt concentrations in this polder and which processes influence these patterns?
3. Is it possible to identify key locations for monitoring the salt concentrations based on modelling and measuring results?
4. Is it possible to optimise sensor placement for a monitoring system in the Lissertocht catchment?
5. Can the results of this research be used for the design of a salinity monitoring network in another catchment?

## 1.4. Thesis outline

In Chapter 2 the measurement methods will be selected to detect and measure the boils. The results of the methods EC routing, CTD divers, DTS measurements and visual inspection will be compared and combined. This will answer the first research question and gives information which helps to answer research questions two and three later in this thesis.

The measurements provide input for the modelling in Chapter 3. The models SOBEK and RSGEM will be explained, with these models the flow of water and salt concentrations are calculated in the catchment under different scenarios. Afterwards the results from the SOBEK model will be analysed with the PCA method, with which a new system of linear components is composed to represent the behaviour of the salinity in the catchment. This leads to a predictive model which can be used to test the performance of a monitoring network. This chapter provides the answers to the second and third research question

The predictive model from Chapter 3 is used in the Chapter 4 to optimise the sensor placement for a efficient monitoring network. This is done by using different strategies such as expert opinion, an exhaustive algorithm and a greedy algorithm to place the sensors. Finally, the robustness of the solutions is tested, the business case is explored and a guideline for applying the knowledge from this research in other catchments is formulated. This will answer the last two research questions.

In the final chapter conclusions are drawn and recommendations will be given. In the appendices additional modelling results can be found, the set-up of the DTS measurements and additional data obtained from the measurement.

# 2

# Measurements

To obtain more knowledge about the spatial and temporal distribution of the salt concentrations more measurements of the boil locations and their fluxes are needed. The first research question regarding detecting and measuring boils will be answered in this chapter. In the first section different methods to detect boils are discussed and the methods EC routing, DTS and visual inspection are chosen afterwards. Next, the results of the EC routing, visual inspection, DTS measurements and CTD divers will be presented. Finally, the different measurement methods will be compared and the results will be combined. The measurements have given more knowledge about the catchment which will help to answer the second and third research question in Chapter 3. The new knowledge about the boil locations and their fluxes will be used to update the SOBEK model in Section 3.2

## 2.1. Selection of measurement methods

As explained in the introduction the boils are the dominant source of saline water in this catchment, therefore it is important to know the locations of the boils to model the salt concentration of the surface water accurately.

The following methods are available to locate the boils in a catchment:

- Detect boils based on visual properties such as bubbles, sand volcanoes, holes in ice, change in vegetation or collapsed banks (De Louw et al., 2013). This is explained in Section 2.2.
- Seepage water from boils has a high salt concentration (above 1000 mg/L) (De Louw et al., 2010); by measuring the electrical conductivity locations with high salt concentrations can be identified. The disturbances in electrical conductivity can be measured with an EC meter or a CTD diver. In bigger channels it is possible to deploy the equipment behind a boat or canoe but for the smaller channels it is easier to drag the equipment through the channels with a (fishing) rod. Apart from the location of the boils, this method also gives an overview of the spatial distribution of the electrical conductivity. This method is called EC routing in this research and is applied in Section 2.2.
- The seepage water from boils has a constant temperature of 11 °C (Goudriaan et al., 2011); this makes it possible to detect them by measuring the water temperature when the temperature of the surface water is lower (mid-winter) or higher (mid-summer). The following methods based on the difference in temperature are available:
  - The research of Hoes et al. (2009) shows that boils can be detected by measuring the water temperature with a DTS cable. This method is applied in this thesis and the results are presented in Section 2.3.

  - Detect disturbances in temperature with a thermal infra-red camera under a drone or a plane. Goudriaan et al. (2011) were able to detect boils on land with this method. Because of the no fly zone around Schiphol airport, this method was not feasible for the Lissertocht catchment.

- Identify areas which are more likely to contain boils based on available information about the polder. These areas are already identified in Figure 1.4 in the introduction:
  - Use Algemeen Hoogtebestand Nederland 3 (AHN3) data to identify areas in the catchment with a lower elevation than the rest of the polder. It will be more likely that these areas contain boils because the upper soil layer will be thinner which increases the risk of boil development.
  - Analyse geology of the area based on data in the DINO loket. For example areas with a paleochannel are more likely to contain boils (del Val Alonso, 2011).

## 2.2. EC routing

This catchment has an area of 10 square kilometers and all channels together have a length of 71,66 kilometers (Hoogheemraadschap Rijnland). EC routing is an easy and cheap method to locate boils. The method consists of a CTD diver attached with a rope to a fishing rod. By walking along the channels while dragging the CTD diver through the channels, it is easy to measure the whole catchment with this method. If the CTD diver is synchronized with the GPS a map can be made of the electrical conductivity values in the catchment. In this section EC maps from previous research will be analysed and compared with maps of new EC measurements. The difference between the EC maps from previous research and the new measurements will be discussed at the end of Section 2.2.

### 2.2.1. EC routing in previous research

The EC routing in May 2011 by Joost Delsman (Figure 2.1) shows that the drainage channels in the middle and in the north west part of the catchment have higher electrical conductivity. The channel which leads to the outlet also has increased values of electrical conductivity. The rest of the main channels and the drainage channels which border to the south east border of the catchment are almost all fresh. The EC routing in April 2011 and February 2013 show significant lower electrical conductivity values. In Figure 2.2b it can be seen that in February 2013 the only channel with a very high electrical conductivity (in red) is one of the channels, which is chosen as the first location for the DTS measurements in Section 2.3. The difference in electrical conductivity values between May 2011, April 2011 and February 2013 can be explained by natural and artificial flushing of the polder. In both April 2011 and February 2013 there was much more precipitation than in May 2011 and this precipitation causes natural flushing of the catchment. This natural flushing dilutes the saline water, which results in lower electrical conductivity values. Therefore, it is harder to detect the boils with EC routing during a wet period.

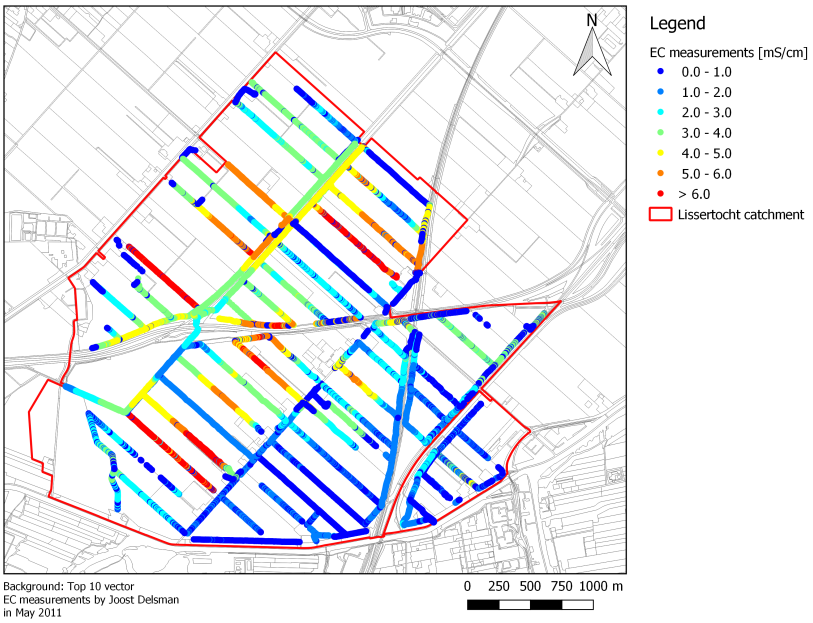
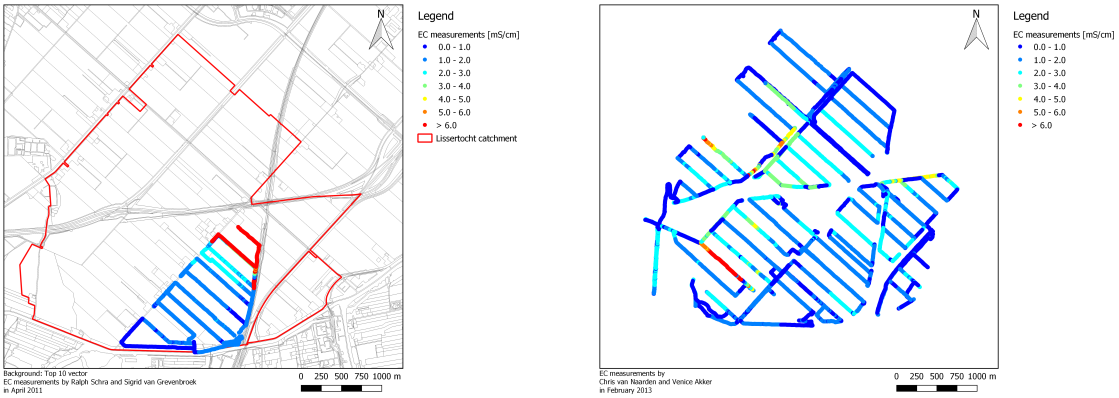



Figure 2.1: EC map of the catchment in May 2011, these measurements are conducted by Joost Delsman




(a) (b)

Figure 2.2: (a) shows EC map of the catchment in April 2011 (measurement conducted by Schra (2013) and Van Grevenbroek (2013)) and (b) shows the EC map in February 2013 (measurements conducted by Akker and Van Naarden (2013)

### 2.2.2. EC routing during this research

During this research, the EC routing was applied again to find out whether it is possible to detect additional so far undetected boils and confirm the boil locations in the SOBEK model. This method is applied twice to see whether are differences between a wet period (Figure 2.3a) and a dryer period (Figure 2.3b). In some areas it was not possible to compare the two EC routing measurements because it was not possible to measure all channels. Some channels were completely dry during the first attempt in March 2018 and some other channels were completely overgrown in May 2018, this explains the channels without values in Figure 2.3a and Figure 2.3b. Because EC routing in the whole catchment takes multiple days the weather conditions also change within in the measurement period, this makes it even harder to compare the two measurements. The measurements with CTD divers in Section 2.4 show the effects of precipitation and a dry period on the salt concentrations. During summer there is less natural flushing via drainage but from the $1^{st}$ of April the inlets

are opened to decrease the salinity in the catchment. In some areas these two types of flushing replaces it each other which makes it more difficult to see a difference between these two periods.

An addition in comparison with the previous measurements is GPS tagging of locations where visual signs such as collapsed banks, seepage on land and less vegetation of boils were observed. Only with this observation the EC map can be interpreted correctly. Locations with an increased electrical conductivity value and visual signs of boils are most likely boils. Locations where the surface water has a higher electrical conductivity value and where there are no visual signs of boils can contain a boil but it is also possible that the saline water originates from other sources or channels. Besides, places that contain visual signs of boils but no increased electrical conductivity value have to be investigated further to be sure the boil adds saline water to the system.


(a) (b)

Figure 2.3: (a) shows the EC map of the catchment in a wet period (March 2018) and (b) shows the EC map of a dryer period (May 2018)

The EC map of May 2011 (Figure 2.1 was used to locate the boils in the current SOBEK model. The boils were located by looking at the EC map and then reasoning back where there should be a boil in the model to reproduce the same EC map. This approach seems logical but it can introduce errors because boils will be located based on the impact they have on the salt concentrations locally. When a boil does not have a big impact locally because of a bigger cross sectional area and higher flow velocities, it will be hard to identify these kind of boils with only an EC map. The boils have relatively a much bigger impact on the electrical conductivity in small stagnant drainage channels, these boils will be detected more easily with EC routing. EC routing combined with visual inspection confirms that some boils in the main channels will not be detected with EC routing while there are clear visual signs of a boil, like collapsed banks. Thus, EC routing alone does not detect all boils.

## 2.3. Distributed Temperature Sensing (DTS)

To detect boils in the Lissertocht catchment a Distributed Temperature Sensing method is used. With the DTS method it is possible to measure the temperature along a cable with a high spatial and temporal resolution. With the DTS measurement it is possible to detect the boils directly by measuring the disturbances in temperature. The deep groundwater from the boils has a constant temperature around 11 °C while the surface water has a temperature between the 2 °C and 5 °C in the winter period. Therefore, the boils can be detected by looking at peaks in temperature along the cable. During the summer these measurements are also possible, in this case the temperature of the surface water is higher than the temperature of the groundwater. The DTS measurements will also be used to calculate the boil flux in Section 2.5.

### 2.3.1. Applications of DTS

The DTS technique was originally developed for the oil and gas industry. Since the late 1980s this technique has been used to monitor pipelines and mine shafts (Kersey, 2000). Later, this technique was also applied in environmental studies to measure the temperature in lakes, glaciers and streams (Selker et al., 2006b) (Selker et al., 2006a). It shows that DTS is a useful technique to measure environmental processes with heat exchange or when temperature can be used as a tracer.

Hoes et al. (2009) showed that DTS can also be applied to measure seepage in polders. In this research they deployed a DTS cable in some channels in the Wieringermeerpolder, Bovenkerkerpolder and Groot Mijdrecht Noord polder to identify seepage along the channels. By applying this technique it was clearly visible where seepage enters those channels. Vandenbohede et al. (2014) tried to combine DTS measurements and measurements of the groundwater temperature around a boil to build a numerical model which can characterize the boil. The conclusion of this research was that it was not possible to quantify the boil based on the DTS measurements because of variation over time (boil flux and stability of the boil conduit), uncertainty on the 3D construction of the boil conduit and heterogeneity of the subsoil (Vandenbohede et al., 2014). The measurements of Hoes et al. (2009) and Vandenbohede et al. (2014) were conducted during summer when the groundwater has a lower temperature than the surface water.

Briggs et al. (2012) compared conventional methods such as dye dilution, acoustic Doppler velocimeter differential gauging and geochemical end-member mixing with the DTS method to determine the spatial distribution and magnitude of groundwater input to streams. The results of this research showed that DTS heat tracing gives the same magnitude of groundwater input as the conventional methods, but with DTS it is possible to describe the groundwater inflow at a higher spatial resolution (Briggs et al., 2012). Westhoff et al. (2011) compared DTS measurement of a stream in the Maisbich catchment in Luxembourg by Selker et al. (2006b) with an advection-dispersion model coupled with an energy balance model to quantify the spatial and temporal dynamics in a stream during a summer rainstorm.

Hilgersom et al. (2018) measured 3D temperature profiles around boils by placing a cage of DTS cables on top of a boil. With this data they built a model that simulates the transport of heat and salt around a boil. The authors discovered that the double diffusive phenomena complicates the calculation of a boil flux (Hilgersom et al., 2016). Because of this phenomena it is hard to model the locations of the convection scales and the boundary conditions, this makes the calculation of a boil flux from a measured water temperature profile less accurate (Hilgersom et al., 2018).

### 2.3.2. Principle of DTS

The concept behind DTS is to observe the backscatter from a laser pulse which is sent through a glass fibre optic cable. The pulse is backscattered in the fibre optic cable as a result of the disordered structure of the glass fibre. The three primary nodes of scattering are: Rayleigh, Brillouin and Raman. Rayleigh scattering is elastic, this means that there is no shift in wavelength between the original signal and the scattered signal. The Brillouin and Raman scattering are non-elastic. These scattered signal exist of signals with a longer wavelength, the Stokes signal, and ones with a shorter wavelength, the anti-Stokes signal, than the original signal (Selker et al., 2006b), see Figure 2.4. To calculate the temperature the Brillouin or Rayleigh band are filtered out of the reflected signal. The wavelength shift ( 40 nm) of the Raman backscatter is big enough in order to isolated this band from the other bands (van der Spek, 2003).

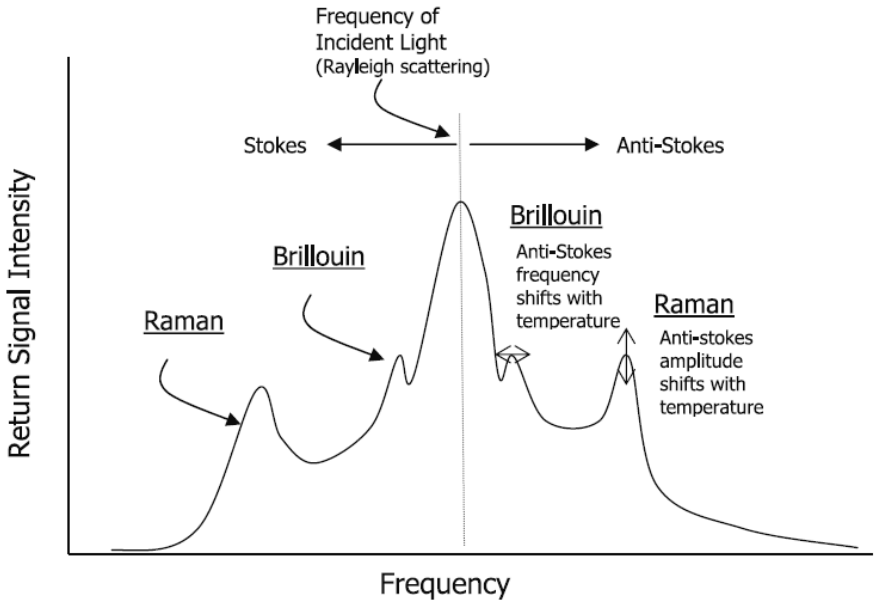



Figure 2.4: Diagram of Rayleigh, Raman and Brillouin backscatter. Stokes and anti-Stokes backscatter.
Source: (Selker et al., 2006b)

In case of the Raman backscatter the amplitude (intensity) of the anti-Stokes backscatter strongly depends on the temperature of the cable. The amplitude of the Anti-Stokes backscatter is bigger if the temperature of the cable is higher. The amplitude of the Stokes backscatter is weakly dependent on the temperature. By observing the ratio between Stokes and anti-Stokes backscatter, it is possible to calculate the temperature along the cable. By measuring the return time of the light, it is also possible to calculate the distance from the place where the light is reflected.

The following formula is used to calculate the temperature from the backscatter:

$$T(z,t) = \frac{\gamma}{ln(\frac{P_S}{P_{aS}}) + C(t) - \int_0^z \Delta\alpha(z')dz'} \tag{2.1}$$

With:
$\gamma = \frac{\delta E}{k}$ = the shift of energy between a photon at the wavelength of the incident signal and the scattered Raman photon $[K]$
$C$ = dimensionless calibration parameter $[-]$
$\Delta\alpha$ = the differential attenuation between the anti-Stokes and Stokes signal $[\Delta m^{-1}]$
$\frac{P_S}{P_{aS}}$ = the ratio of the power of the Stokes signal divided by the power of the anti-stokes signal [ - ]
$T(z)$ = temperature along the cable at position z at time t $[K]$

### 2.3.3. Set-up, locations, equipment and calibration

In this subsection the two selected measurement locations and the set-ups are discussed. Besides the equipment and calibrations is presented here. Pictures of the set-up, equipment and the locations can be found in Appendix A.

The following material has been used for the DTS measurements:

- Cable: OD 6.0mm Armored FO PBT patch cord 50/1252 fibers with 2 x E2000/APC – 2 x E2000/APCconnector
- Silixa Ultima DTS device (Silixa) with the following specifications:
  - The range is 2 kilometer
  - Spatial resolution of minimal 30 cetimeter
  - Sampling resolution of 12.5 centimeter
  - Temporal resolution of minimal 1 second
  - 4 channels which allows to measure 4 fibers simultaneously
- Calibration bath to calibrate the absolute offset, this calibration bath contains:
  - Box with water
  - Thermometer which measures the temperature of the water
  - Aquarium bubler to prevent stratification in the calibration bath
  - 20 meter of the DTS cable
- Tipping bucket with temperature sensor which measures the precipitation and air temperature at the DTS measurement locations.

The first measurement location at the Kaagweg is an agricultural field of 1000 meter by 200 meter which is known to contain boils according to the previous EC routing and visual inspection. These boils are located on land, at the banks of the ditches and in the ditches themselves. At each side of the agricultural field a drainage ditch is located, in both ditches a cable of 1000 meter is deployed. From here on the two ditches will be called the west ditch and the east ditch based on their location. The cables were at the first measurement location from the $31^{st}$ of January 2018 to the $13^{th}$ of February 2018.

At the beginning of the cable a calibration bath is located. This bath is located next to the Ultima Silixa device which is located in the barn. From the calibration bath one cable has to cross the field of the farmer and then enters the west ditch and another cable directly enters the eastern ditch. At the end of each ditch a coil is located with the last meters of the cable and a splice, this splice connects the ends of the fibres. In the west ditch the cable was only able to measure for roughly two days because the cable was broken by an excavator on the $2^{nd}$ of February. In the east ditch the cable was able to measure for the full period of 14 days. In the east ditch the cable had to cross three small dams; at those places the water flows through a culvert.

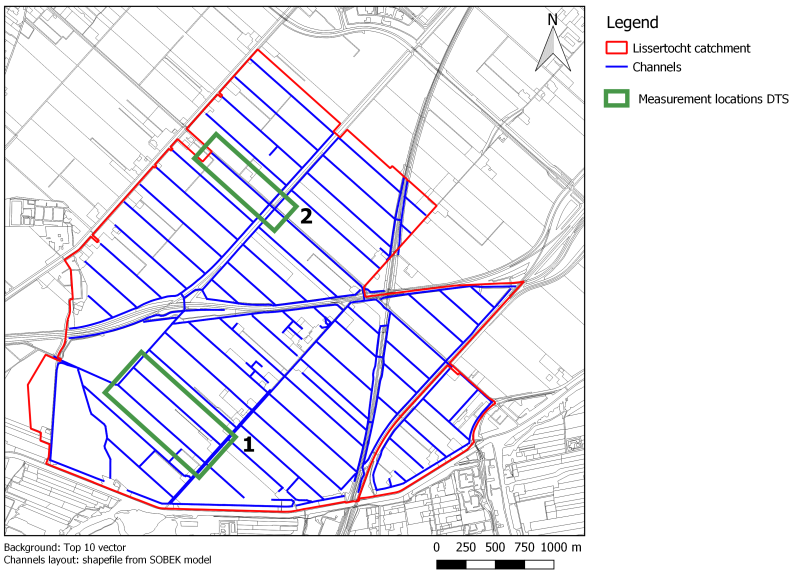



Figure 2.5: Measurement locations of DTS

The second measurement location is located in the channel that leads to the outlet of the catchment, the pump J.P. Heye. At this location it is hard to detect boils with only EC measurement because the cross section of this channel is bigger and there are disturbances if the pump switches on and off. In addition, the saline water from the whole catchment flows through this channel to the pump, which makes it more difficult to determine if the higher salinity at this location is caused by boils in this channel or by boils which are located somewhere else in the catchment. At this location one cable of 1000 meter is deployed in a loop, 500 meter along one bank and 500 meter back along the other bank of the channel. Because the width of this channel (8 meter) was bigger than the width of the channels at the first measurement location (2 meter) it is necessary to measure at multiple points along the width of the channel. The DTS device and calibration bath were installed at the pump. The cable has been measuring at this location from $13^{th}$ of February until $15^{th}$ of March. The set-up at the second location is shown at Figure A.2. in Appendix A.1

The measurements are calibrated by comparing the temperature measurements of the thermometer and of the DTS cable in the calibration bath. The offset (difference) between the two is the absolute offset of the DTS measurements. This offset was in a range of 0.1 to 0.5 °C in some channels but was in the range of 1 °C in other channels. The signal strength in the cables with a low offset was much stronger than in the cables with a bigger offset. The loss in signal strength can be explained by dirt in the connectors or connectors which have a lower quality. It can also be caused by the quality of the splice at the end of the cable. In Figure 2.6 the offset of one of the channels with a strong signal is plotted. The Ultima Silixa DTS device was only able to measure over 1800 meter of fiber while the total length of the fibers was 2000 meter, 1000 meter from the device at the beginning of the cable to the splice at the end of the cable and 1000 meter back again to the device. Without measuring the full length of the fiber it is not possible to apply the single end calibration method (Hausner et al., 2011) or double end calibration method (van de Giesen et al., 2012). These methods need two measurement points at each location along the cable.

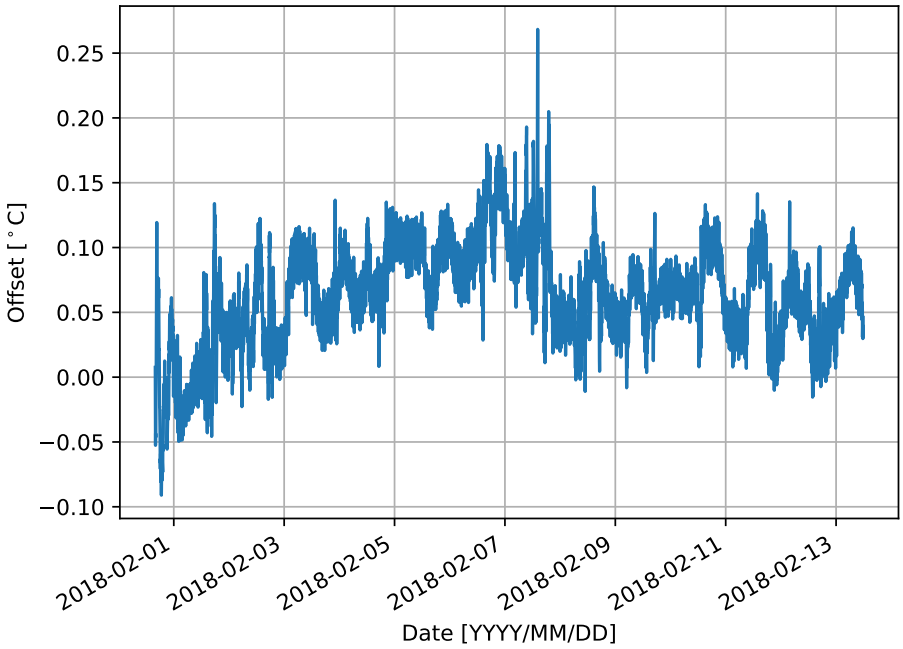



Figure 2.6: Absolute offset of temperature measurements at the first location

### 2.3.4. Results location 1: Kaagweg

Figure 2.7 shows the temperate along the cable in the western ditch until the time when the cable was broken. During the measurement period warmer areas can be seen around the 300 meter, 350 meter, 450 meter and 550 meter marks from the barn. These are the locations where warmer and more saline water (around 11 °C) enters the ditch via boils. Visual observations show that at this location the bank of the ditch has collapsed and that seepage is entering the ditch (see Figure 2.9). Measurements with an electrical conductivity meter at those locations show salinity values up to 4780 mg/L (electrical conductivity is converted with formula from Deltares (2015)) and temperatures up to 9,8 °C at the biggest boils (see Figure 2.8). These values are higher than measured by De Louw et al. (2010) of boils in the Noordplaspolder; this difference can be caused by the difference in depth of the origin of the seepage water and by different characteristics of the deep saline aquifer. So visual observations and EC measurements confirm that the locations where the DTS cable measures a higher temperature are boils. However, it is hard to identify the precise locations of the boils with an EC meter or visual observations only. Only with the DTS measurements it is possible to identify 4 boils and their precise locations. With visual observations it is only possible to say that there is an area with boils along this ditch. With the DTS results it is also possible to make a comparison between the boils on a qualitative level. The peak in temperature of the two boils at 290 meter and 350 meter is higher than the peaks in temperature which correspond to the two others more downstream, this gives an indication of the flux and size of the boils.

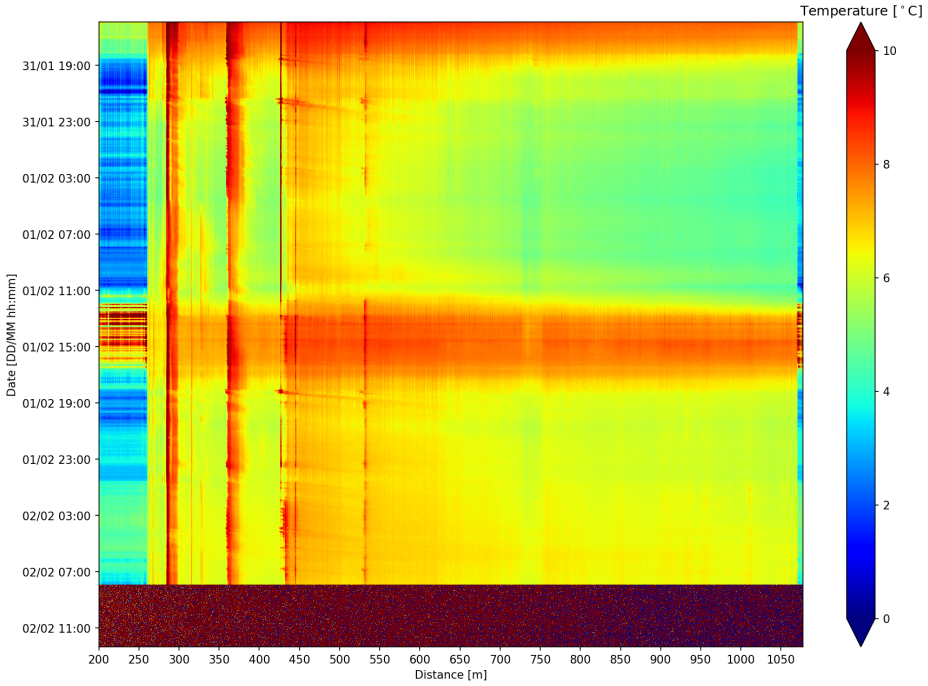



Figure 2.7: Temperatures measured along the cable in the west ditch over the period from the $31^{th}$ of January 2018 until the $2^{nd}$ of February 2018 (when the cable has been broken)

The temperature and salinity decrease when the water flows towards the end of the ditch (right side of Figure 2.7). At this side the ditch drains into a bigger channel through a culvert which at the end drains the water from the catchment via the pump. When it rains the surface water in the ditch also dilutes when it flows towards the end of the ditch because drainage pipes drain rainwater into the ditch.

Figure 2.8: One of the boils in the west ditch (at x = 350 meter). A temperature of 9.8 C° and a electrical conductivity which corresponds to 4780 mg/L has been measured here

Figure 2.9: Section of the ditch with the boils

The cable in the east ditch was able to measure for a longer period ($31^{th}$ of January until $13^{th}$ of February). Figure 2.10 shows all the data which is measured with the second cable. This cable enters the ditch near an artificial boil. This boil exits from a pipe which has been used in the past to extract gas. In the measurements

it is visible that warmer groundwater with a constant temperature of 10 °C still flows through this artificial boil into the ditch (at x = 50 meter). The salinity measured next to this artificial boil is 1250 mg/L. After this part of the channel the cable has to cross a small dam (at x = 90 meter). Thereafter it enters the main ditch again, here the cable has to cross two more dams at 480 to 500 meter and 600 to 650 meter before the cable reaches the end of the ditch. At the end of the ditch the water flows through a culvert into the Kagertocht. The dam which is located at 600 meter is used for a temporary road for construction of a tower for a high voltage electricity cable. This side is located on the field between the two ditches.

The seepage enters this ditch around 250 meter and 540 meter from the barn. The DTS measurements show an increase in temperature in these areas. EC measurements show that the salinity at the first boil is 1250 mg/L and a temperature of 5.4 C°, this boil next to the ditch is also clearly visible. In comparison with the boils in the west ditch these boils give a weaker heat signal. This corresponds with the manual measurements of electrical conductivity and temperature. This means that the boils in the west ditch should have a bigger flux than the ones in the east ditch. This also means that it is possible to make a qualitative comparison between the two ditches and the different boils.

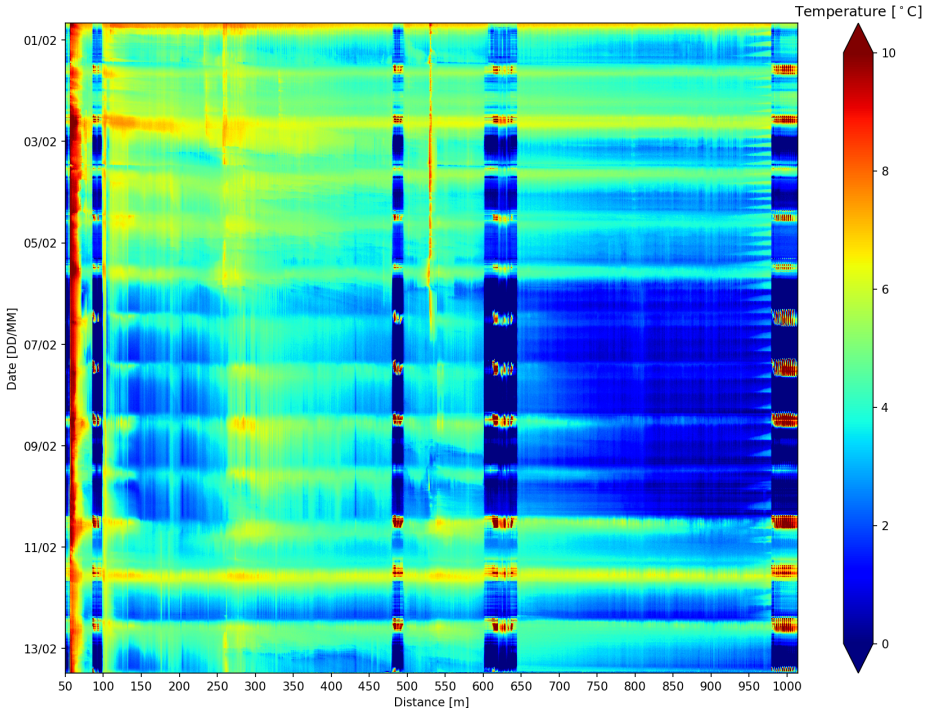



Figure 2.10: Temperatures measured along the cable in the east ditch over the period from the $31^{th}$ of January 2018 until the $13^{th}$ of February 2018

The temperature measurements from both ditches show a clear day-night pattern. During the day the ditch warms up due to the solar radiation and during night it cools down. This pattern is also visible at the points where the cable is outside of the ditch. At the beginning of the measurement period there were some warmer days and near the end of the period there were some days when the temperature dropped below 0 °C. During this colder period visual observations show that there is a thin layer of ice on the ditch. The west ditch was almost completely free of ice and the east ditch had only some small areas without ice near the boils. This confirms the data from the DTS measurements, this dataset shows that the west ditch is warmer (more saline) than the east ditch. So the DTS measurement can be used to compare the two ditches and the boils on a qualitative level. It is not possible to quantify the boils with this dataset because the temperature around the boils is almost constant over time which makes it difficult to determine how much water is added by the boils.

### 2.3.5. Results location 2: pump J.P. Heye

The second measurement location is located in the Lissertocht channel, this is the channel which leads to the pump (J.P. Heye) of this catchment. In Figure 2.11 the temperature measurements at this location are shown between the $13^{th}$ of February and the $15^{th}$ of March. In this figure the first 40 meter along the x-axis shows the temperature in the calibration bath. From here the cable is in the channel for the next 500 meter along the bank of the channels (eastern side). After a loop at 550 meter it runs along the other bank of the channel (western side) for approximately 500 meter. The temperature at x = 140 meter, x = 355 meter and at x = 380 meter is always higher. This is a clear indication that there are boils located at those locations. The same disturbances in temperature are measured along the cable when it passes the boils again along the other bank of the channel. So there are three boils located at this location. The first boil is located in the middle of the channel because the temperature signal of this boil can be seen at both cables and no visual signs of the boil are visible on the banks of the channel. The other two boils are located near the bank of the channel because the disturbance in the temperature signal is stronger at one of the cables then at the other one. The bank of the channel near these boils has also collapsed, see Figure 2.12. The DTS measurements also shows that the first boil is bigger than the second and third. The second and third boil are located near each other so it is possible that there are two conduits from one boil instead of two independent boils.

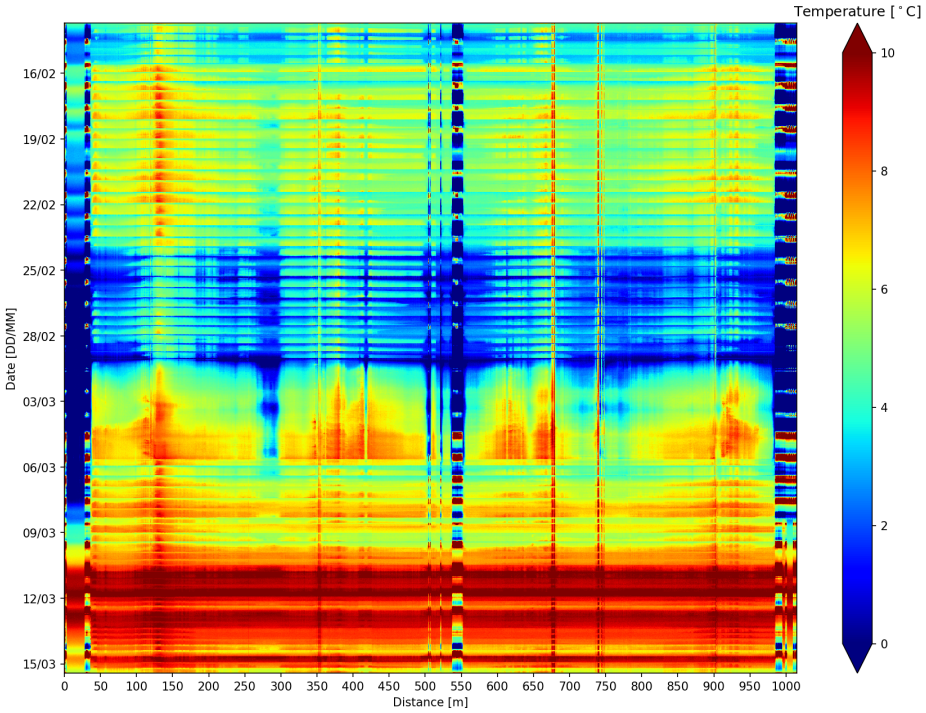



Figure 2.11: Temperatures measured along the cable in the channel next to the pump J.P. Heye over the period from $13^{th}$ of February 2018 until the $15^{th}$ of March 2018. The calibration bath is located from 0 to 40 meter, the cable is deployed in the channel along one bank from 40 meter until 540 meter. A loop is located at 550 meter after which the cable runs along the opposite side of the channel back to the starting point

During the measurement period there was ice on the channels for some days around the $1^{st}$ of March. During this cold period the temperature on the bottom of the channel has risen up to 8 °C while the outside air temperature was -5 °C. The temperature in the channel was able to rise during this period because the pump was switched off.

Figure 2.12: A collapsed bank caused by one of the boils which is detected in the Lissertocht

At this location it was not possible to detect the boils with only EC routing. The salt concentrations in this channels are high due to the boils and because of saline water from the whole catchment which flows toward the outlet. Therefore, it is harder to determine the source of the saline water. The cross section of this channel is bigger than other channels in the area. Therefore, the chance of stratification is bigger because the salt water has to spread over a bigger area. In Section 2.5.2 the stratification of heat and salt are estimated. The EC routing is less reliable when there is more stratification because it is hard to measure at the same depth along all channels. At this location a disturbance in the electrical conductivity values can also mean that the CTD diver has been measuring at a different depth. Along this channel there are visual signs that there are one or multiple boils at the location of the second and third boil (around 370 meters from the pump), these boils can also be seen at the DTS results. This means that the first boil has only been detected with DTS measurements. This is a boil which is located at the bottom of a main channel.

## 2.4. CTD divers

To measure the effect of the precipitation, flushing and pumping on the salt concentrations three CTD divers are installed. This instrument measures the electrical conductivity, temperature and pressure at one point with a given time interval. In Figure 2.13 a picture of one of the CTD divers can be found. This CTD diver is installed in a PVC tube with holes and a filter to protect the diver from dirt and sediment. This tube is pushed in the bottom of the channel. The sampling interval of the CTD diver was set to 5 minutes.

Figure 2.13: CTD diver

The three CTD divers were installed at the locations indicated in red in Figure 2.14 from the $29^{th}$ of March until the $30^{th}$ of April. Figure 2.15 shows the electrical conductivity ($1^{st}$ axis) and the precipitation at Schiphol ($2^{nd}$ axis) during this period, the purple blocks indicate when the pump was working. This measurements clearly show that the electrical conductivity in the location Lissertocht decrease when there is a lot of precipitation, while the electrical conductivity at the locations Kaagweg and J.P. Heyepad do not change. The last two locations are near the inlets and more upstream so almost no drainage channels drain into this point. The other location is at the outlet, so a large amount of the fresh precipitation flows to these points via drains and drainage channels. From the $18^{th}$ of April to the $23^{th}$ of April there is almost no precipitation and the pump has not been used, during this period the electrical conductivity increases at the CTD diver in the Lissertocht and at the CTD diver at the Kaagweg. This clearly shows the influence of precipitation and of a dryer period on the electrical conductivity. It also shows that the dynamics can be different for different locations but also that the locations Lissertocht and J.P. Heyepad behave in the same pattern.

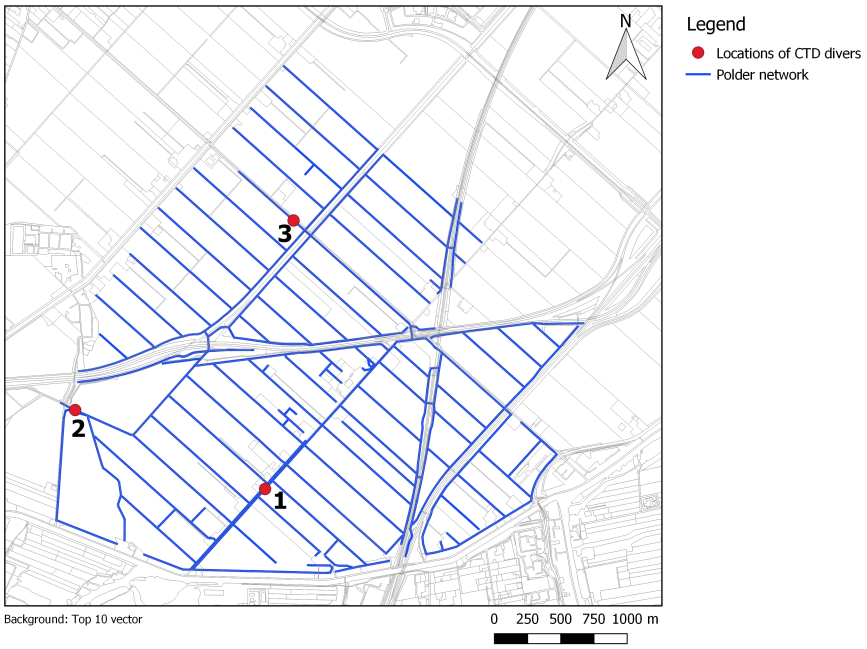



Figure 2.14: Locations of installed CTD divers. Location 1 = Kaagweg, Location 2 = J.P. Heyepad, location 3 = Lissertocht

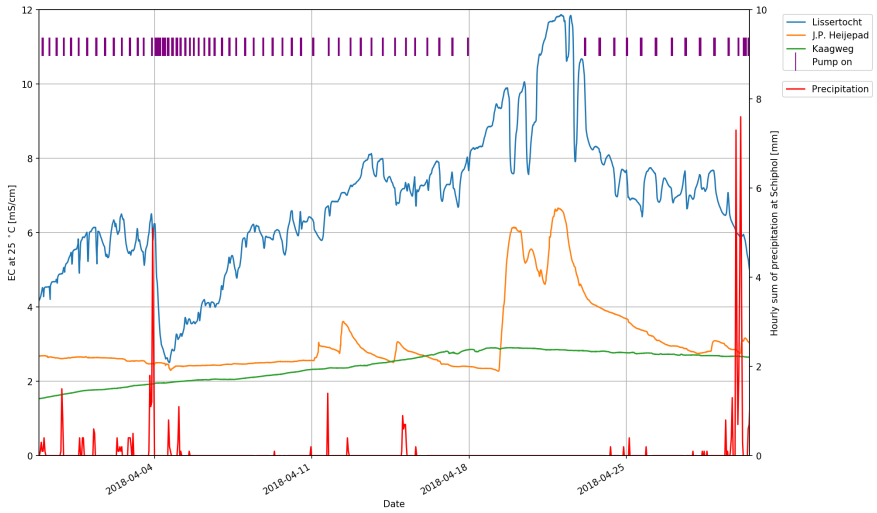



Figure 2.15: EC measurements with CTD divers and hourly precipitation

After the first measurement period the CTD divers were installed again in the channel near the pump J.P. Heye (location 3 in Figure 2.14), where they have been recording from $30^{th}$ of April until the $22^{th}$ of May. This is the same locations where multiple boils have been measured with the DTS method in Section 2.3.5. The goal of these measurements is to compose a salt balance over the channel next to the outlet to determine the fluxes of the boils in this channel. In Figure 2.16 it can seen the result of the three CTD divers installed near the pump. Between the diver at the pump and the middle one there is a big boil (see Figure 2.11 at x = 140 meter). Between the one which is located 380 meters from the pump and the middle one (215 meter from the pump) there are two smaller boils located. In the second axis the precipitation is plotted, this shows that it is a dry period. This Figure also shows that when the pump is turned the divers observe first a peak in electrical conductivity and later the EC value decrease. In Section 2.5.4 the data will be used for determining the boil flux with a 1D advection-diffusion model.

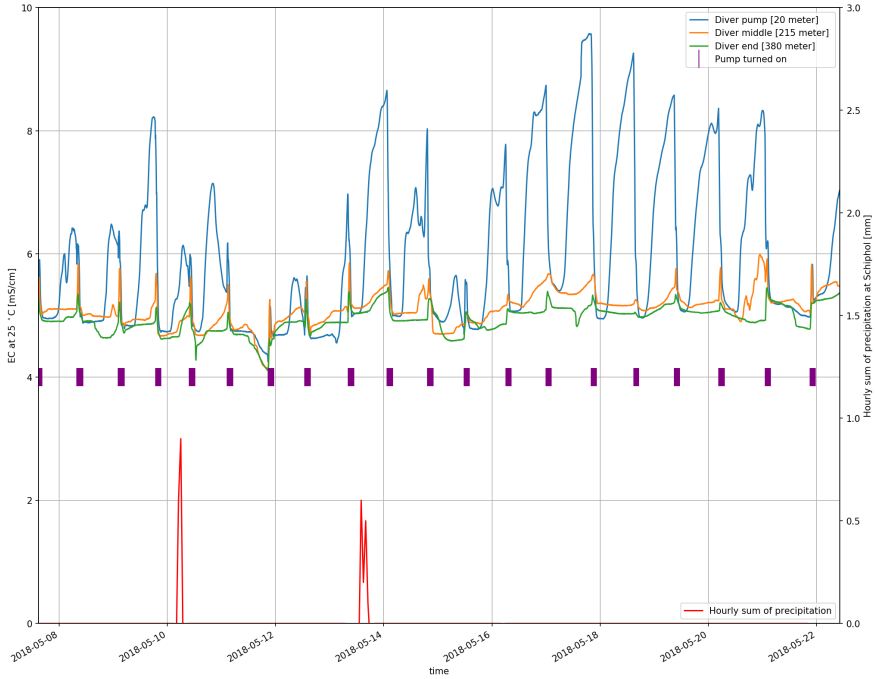



Figure 2.16: Three divers installed in the channel next to the pump

## 2.5. Determining the boil flux

In previous research of Delsman et al. (2013) the total seepage flux of this area is determined to be 0.32 mm/day. This seepage flux is obtained by taking daily samples at the outlet of the catchment and making a water balance for the whole catchment. The flux of the individual boils is obtained by dividing the total seepage flux with a standard deviation of 10% over the 20 boils which are located in the catchment (Kelderman, 2015). This results in boil fluxes which vary between 0.00172 $m^3/s$ to 0.00218 $m^3/s$.

In this chapter the boil flux will be calculated based on the DTS measurements, flow measurements and CTD divers at the channel next to the pumping station (the Lissertocht). Several approaches to calculate the boil flux in this channel will be explained and their limitations will be discussed. First the the flow speed is measured and discharge is calculated, then the stratification in the channel is investigated by calculating dimensionless numbers and some measurements, next an energy balance is computed based on the DTS measurements. Finally, a 1D advection-diffusion model will be solved using the data from the measurements.

### 2.5.1. Flow measurements

At the Lissertocht flow measurements were conducted to measure the flow velocity when the pump is switched off and when it is on. These flow velocities are used to calculate the Reynolds number and to calculate the discharge in the channel, the discharge will be used in the 1D advection-diffusion model. When the pump is switched off the flow velocities are too small to use a propeller current meter or an ADCP to measure the flow velocity, so a float with a small weight and some resistance at the middle is chosen to measure the flow velocity. The travel time of the float between two point is measured to calculate the flow velocity. When the pump was off the float was not moving so this is an indication that the flow velocities are very small at that moment. When the pump was switched on a flow velocity of 0.25 m/s was measured. When multiplying this flow velocity with the cross sectional area the discharge can be calculated: $0.25m/s * 5m^2 = 1.25m^3/s$. The capacity of the pump is given by the waterboard Rijnland to be $1.48m^3/s$, so that is in the same order as the measured discharge in the channel which leads to the pump.

### 2.5.2. Stratification

The boils are adding more saline water and water of with a higher temperature than the surface water to the system, this can result in temperature and salinity gradients over depth, this is called a thermohaline system. For the calculation of the boil flux it is important to know more about the flow regime, dominant processes in heat and salt transfer and the stratification in the channel.

A first step to determine the kind of mixing of salt and heat is to use dimensionless numbers. Ruddick (1983) proposed to use the Turner angle $|Tu|$ to determine the type of mixing in a thermohaline systems, the Turner angle is defined as:

$$Tu = arctan(\frac{N_T^2 - N_S^2}{N_T^2 + N_S^2}) + \frac{180^\circ}{\pi} \tag{2.2}$$

When the Turner angle is smaller than 45° the thermohaline system is stable, when the Turner angle is bigger than 90° the system is gravitational unstable, double diffusive processes occur for a Turner angle between $-90^\circ$ and $-45^\circ$ and salt fingers can occur for a Turner angle between 45° and 90°.
With the local temperature stability $N_T^2 = -g * \alpha_v * \partial T/\partial z$ and the local salinity stability $N_S^2 = g * \beta_v * \partial S/\partial z$ (Turner, 1973). The values for the volumetric expansion coefficient $\alpha_v$ ($^\circ C^{-1}$) and the expansion coefficient $\beta_v$ ($10^3 kgkg^{-1}$) can be calculated with (Wright, 1997):

$$\alpha_v = -2.285097 * 10^{-5} + 1.324876 * 10^{-5} * T - 9.288537 * 10^{-8} * T^2 + 1.563353 * 10^{-6} * S \tag{2.3}$$

$$\beta_v = 7.998742 * 10^{-4} - 2.774404 * 10^{-6} * T + 3.188185 * 10^{-8} * T^2 - 4.151510 * 10^{-7} * S \tag{2.4}$$

This results in a value for $\alpha_v$ in the order of $10^{-5}$ and a value for $\beta_v$ in the order of $10^{-4}$ in the Lissertocht. The salinity gradient $\partial S/\partial z$ and temperature gradient $\partial T/\partial z$ are not measured. So these gradients are estimated based on the temperature and salinity of the boil and the measured temperature and salinity of the surface water. Depending on these estimated gradients, the Turner angle is approximately between $-48°$ and $-58°$, this is an indication that double diffusive processes occur.

With the flow velocity the dimensionless Reynolds number (Re) can be calculated. This number gives an indication if the flow in a channel is laminar ($Re < 2100$), turbulent ($Re > 10000$) or in the transitional regime between the two ($2100 < Re < 10000$). The Reynolds number is defined as (Mills, 1999):

$$Re = \frac{\rho * v * L}{\mu} = \frac{v * R}{v} \tag{2.5}$$

Where $\rho$ is the density of the water ($10^3 kg/m^3$), $v$ is the flow velocity ($m/s$), $\mu$ is the dynamic viscosity of water ($10^{-4} kg/ms$) and $v$ is the kinematic viscosity of water ($10^{-6} m^2/s$) (Mills, 1999). In the case of open channel flow the characteristic length $L$ is equal to the hydraulic radius $R$ of the channel. The hydraulic radius is equal to the area divided by the wetted perimeter $P$, for the second measurement location the hydraulic radius is equal to $5/1.9 = 2.6$ meter. The kinematic viscosity $v$ is equal to the dynamic viscosity $\mu$ divided by the density $\rho$, for water this is equal to $10^{-6} m^2/s$. So when the pump is on the Reynolds number is in the order of $\frac{10^{-1} * 1}{10^{-6}} \approx 10^5$, this means that the flow is turbulent. If the pump is switched off the flow velocity is too small too measure with a float so it has to be smaller then $10^{-3} m/s$, in the SOBEK model the flow velocities in this channel are in the order of $10^{-5} m/s$, this means that the Reynolds number is lower than 2100 and that the flow is laminar if the pump is switched off.

The Rayleigh number indicates if the heat transfer in a fluid is mainly in the form of conduction or in the form of convection. When the Rayleigh number is below the critical value conduction is dominant, otherwise convection is the dominant form of heat transfer in the system. The Rayleigh number is defined as (Mills, 1999):

$$Rayleigh = Gr * Pr = \frac{g * \beta * \Delta T * L^3}{\alpha * v} \tag{2.6}$$

This equation holds as long as the Prandtl number is bigger than 1, this number is defined as (Mills, 1999):

$$Pr = \frac{v}{\alpha} \tag{2.7}$$

The Pandtl number for water is between 1 and 10 (Mills, 1999), dependent on the temperature. In the Lissertocht the Prandtl number is around 10. This means that the momentum diffusivity is bigger than the thermal diffusivity. The Grashof number approximates the ratio of buoyancy to viscous force acting on a fluid, this number is defined as (Mills, 1999):

$$Gr = \frac{g * \beta * \Delta T * L^3}{v^2} \tag{2.8}$$

In these equations $g$ is the gravitational acceleration ($m/s^2$), $\alpha$ is the thermal diffusivity ($m^2/s$) ,$\beta$ is the thermal coefficient of volume expansion ($K^{-1}$). $v$ the kinematic viscosity ($m^2/s$),$L$ is the characteristic length ($m$) and $\Delta T$ is the difference in temperature between the surface water and the seepage water ($K$). In this case the Grashof number is in the order of $10^7$, this means that the Rayleigh number has a magnitude of approximately $10^7 * 10 \approx 10^8$. This is smaller than the critical Rayleigh number of $10^9$ (Mills, 1999) which means that heat transfer will be primary in the form of conduction.

Table 2.1: Summary of values and meaning of dimensionless numbers at the second measurement location (pump J.P. Heye) when the pump is switched off

| **Dimensionless numbers** | **Value** | **Meaning** |
|---|---|---|
| Turner angle | $-48°$ to $-58°$ | Double diffusive processes can occur |
| Reynolds number | 10 to $10^3$ | Laminar |
| Rayleigh number | $10^8$ | conduction if main form of heat transfer |

The dimensionless numbers for the channel next to the pump J.P. Heye are summarised in Table 2.1. The Reynolds number shows that there is laminar flow if the pump is switched off and turbulent flow when it is turned on. When the pump is turned off the Turner angle shows that double diffusive processes can occur. The Rayleigh number shows that conduction is the main form of heat transfer in this system.

On the $22^{th}$ of May the stratification is measured in the channel near the pump J.P. Heye with an EC meter. This showed that the electrical conductivity is highest in the middle near the bottom (10 to 12 mS/cm), lower at the sides near the bottom (6 to 8 mS/cm) and lowest near the water surface (3 to 4 mS/cm), this is presented visually in Figure 2.19. When a sensor is placed in a bigger main channel, the stratification has to be taken into account. This is because, when a sensor is placed at the bottom of the channel a higher electrical conductivity will me measured than with a sensor which is placed near the water surface.

### 2.5.3. Energy balance

When the pump is turned on, colder water from the rest of the catchment will flow through this channel to the pump, which pumps it on a bigger channel which leads to one of the main pump ('gemaal Leeghwater') of the Haarlemmermeer polder. This colder water replaces the warmer water above the boils. This results in a uniform temperature along the cable. When the pump is turned off the boils start heating up the water again until a steady state is reached. The warmer water is spreading along the channel from the location of the boils. This process is repeated every time the pump starts pumping, on average twice a day. This phenomena offers an opportunity to calculate the boil flux. In Figure 2.17 the temperature is plotted at 150 meter from the start of the cable, this shows the described phenomena.

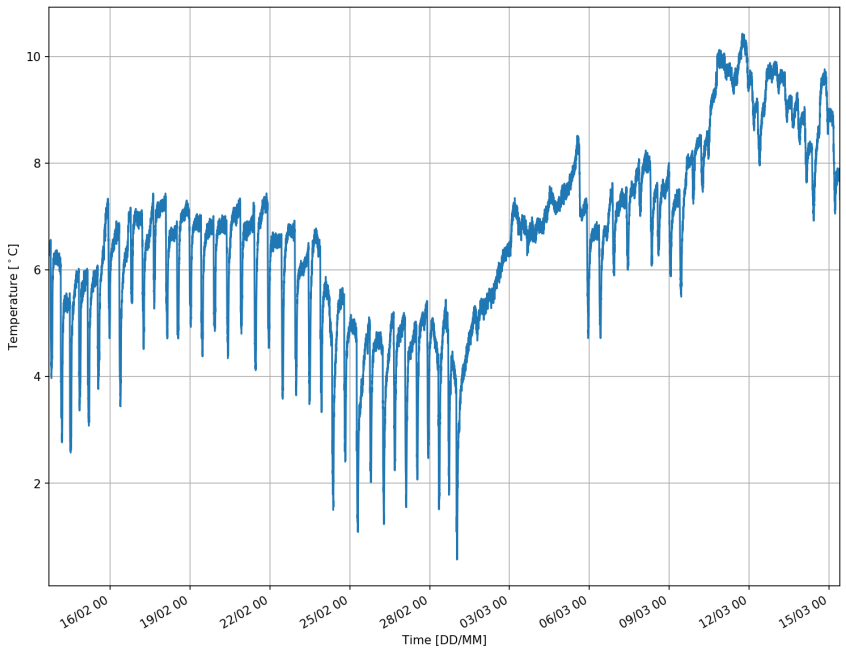



Figure 2.17: Temperature in the Lissertocht at 150 meter from the start of the cable. Shows the effect of the pump on the temperature over the whole measurement period

Figure 2.17 also shows the period around the $1^{st}$ until the $6^{th}$ of March when there was ice. During this period the pump has been off and the temperature at the bottom of the channel increased. When the ice was gone and the pump started working again the same pattern can be seen again. Figure 2.18 shows the temperature near the first boil over 3 pumping cycles. This shows that the temperature drops 3 to 4 °C in less than one hour. The temperature increases more gradually again when the pump is turned off. This pattern repeats itself on average twice a day, this reduces the error when it is used to determine the boil flux. However, the pattern is not always constant because the water temperature varies due to changing weather conditions. The pump cycles are dependent on the water level, which is dependent on the precipitation mainly. Therefore the pattern of the temperature near a boil is not constant over time because the pump influences the water temperature near the boil.

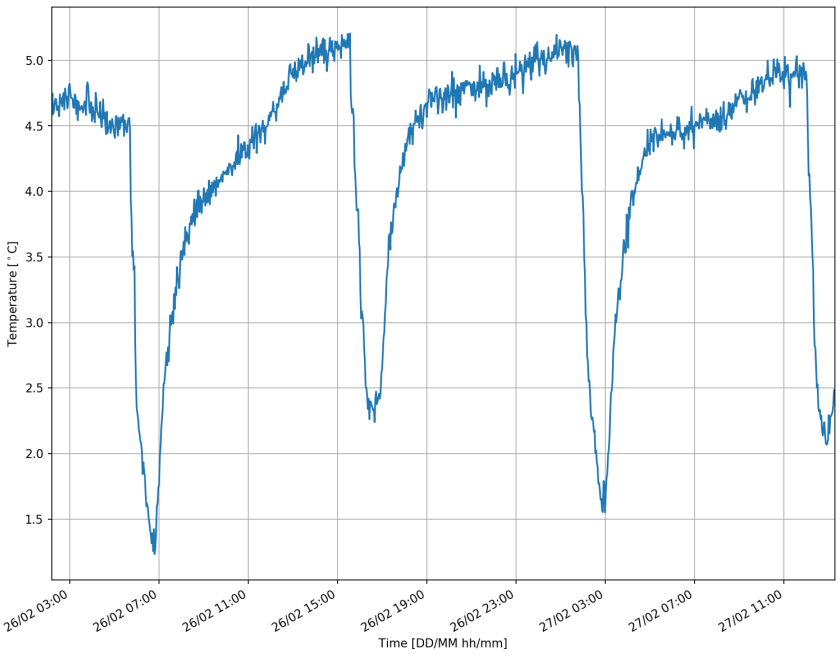



Figure 2.18: Temperature in the Lissertocht at 150 meter from the start of the cable. Shows the effect of the pump on the temperature over some pumping cycles

The energy which is necessary to heat the water can only be provided by the warmer water from the boils. Heat transfer with the air or thermal radiation is much slower than the phenomena observed here. Besides this phenomenon is also observed during night when the other processes will normally cool down the water instead of heating. A final reason is that the measurements show that the temperature starts increasing first near the locations of the boils and then spread along the channel like a wave, if other processes play a significant role in heating the water they will not only occur at the locations of the boils. In this case the energy balance can be simplified to:

$$\Delta E_{surfacewater} + E_{boils} = 0 \tag{2.9}$$

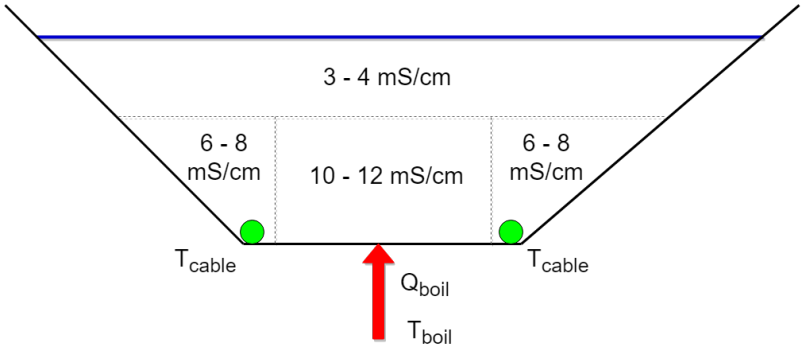



Figure 2.19: Schematisation of energy balance and EC measurements near the pump. The DTS cables are indicated with the green circles. The dotted lines present the EC measurement near the pump J.P. Heye (second measurement location), this gives an indication of the stratification at this location. The red arrow represents the boil flux

The dimensionless Rayleigh number in the previous section has showed that the dominant form of heat transfer is in the form of conduction. The formula for heat transfer via conduction is defined as (Moran and Shapiro, 2010):

$$E_{added} = m * c * \Delta T \tag{2.10}$$

With this equation and the energy balance the volume of water added by the boils can be defined as:

$$V_{boil} = \frac{A * h * \rho_w * c_w * \Delta T}{\rho_{boil} * c_{boil} * \Delta T_{boil}} \tag{2.11}$$

This volume can be converted to a discharge by dividing it with the time interval over which the calculation is done. In these equations $m$ is the mass of the volume of water which has to be heated in $kg$, $c$ is the specific heat capacity in $kJ/kg°C$ and $\Delta T$ is the difference in temperature in °C. The main sources of uncertainties will be:

- In this case there is no data available about the temperature profile over the depth of the channel. As a result it is unknown if the water from the boils stays in a layer at the bottom or if it mixes over the total profile. For this reason it is not possible to determine the location of the double diffusive interface which is described in Hilgersom et al. (2016).
- The boil flux will vary over time because the water depth above the boil varies due to pumping. This will cause a difference in pressure, however, this effect is limited. Only 10 to 15 cm of water level variations, this corresponds with 1,5 kPA.
- Energy exchange between the surface water and the atmosphere. During the night the channel will give heat to the colder air and during the day there will be solar radiation when the sun is shining. These terms are not included in the energy balance.

The DTS measurements at the second measurement location are used to calculate the amount of heat which is added after some pumping cycles. With this measurements at each 30 cm and at two sides of the channel a temperature is known. Thus, the mean of these two temperatures is taken. Still, no temperature over depth is known, only at the bottom. So the two most extreme situations are calculated. The first one assumes uniform temperature and heating over the depth, in this case the temperature measured at the bottom is taken as the temperature over the whole cross section. This approach results in a boil flux of approximately 0.015 $m^3/s$. The other extreme situation is a highly stratified profile where the temperature measured at the bottom of the channel only corresponds with the temperature of the lower 10 centimeters of the profile, in this situation the salt water stays at the bottom of the channel. With this assumption the boil flux is equal to approximately 0.0015 $m^3/s$. Because the distribution of the temperature over depth is it only possible to say that the boil flux is somewhere between 0.0015 and 0.015 $m^3/s$.

### 2.5.4. 1D advection - diffusion model

Another approach to determine the boil flux is simulating the transport of salt in the channel. This can be simulated with a 1D advection - diffusion model. This model simulates the transport of water and salt in a 1D pool, for a schematisation see Figure 2.20 (Xu et al., 2010). The water level is calculated with the Saint Venant equations and continuity equation. The salt concentrations are calculated by solving the advection-diffusion equation. This equation contains an advection term, a diffusion term and a source/sink term (in this case the boils are a source term which add saline water).

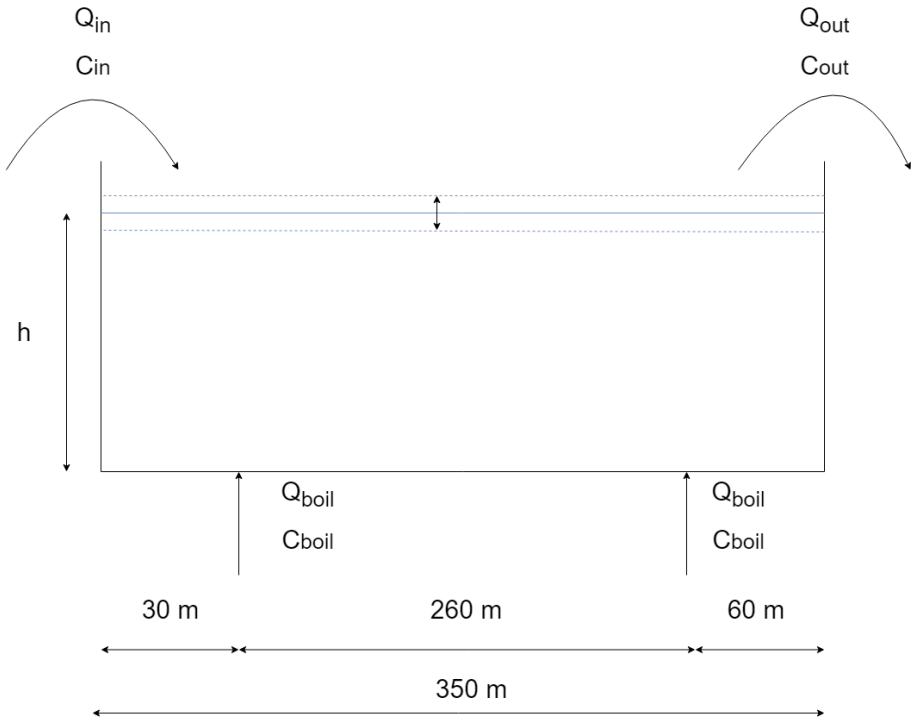



Figure 2.20: Schematisation of the Lissertocht with h = water depth, Qin = discharge into the channel, Cin = salinity of water flowing into channel, Qout = water pumped out of the channel, Cout = salinity of water pumped out of the channel, Qboil = discharge of the boil and Cboil = salinity of the boil

In this model the discharge $Q_{out}$ is equal to the capacity of the pump ($1.48 m^3/s$) which is given by Rijnland. The waterboard also register when the pump is turned on and the duration of each pumping cycle. This information is used for the discharge $Q_{out}$ out of the model. The inflow $Q_{in}$ is determined by solving the water balance over the channel. In this way the $Q_{in}$ corresponds to the rise in water level until the switch on level is reached and the water level is lowered until the switch off level. The water depth in this channel is equal to 0.85 meter (Hoogheemraadschap Rijnland). The concentration of the boils $C_{boil}$ is equal to 5453 mg/L, this is the same as in the SOBEK model (Kelderman, 2015). The measured salt concentration with a CTD diver at the beginning of the channel is used as $C_{in}$. Different values of $Q_{boil}$ are used until the modelled concentration at the outlet $C_{out}$ is equal to the measured salt concentration at the outlet.

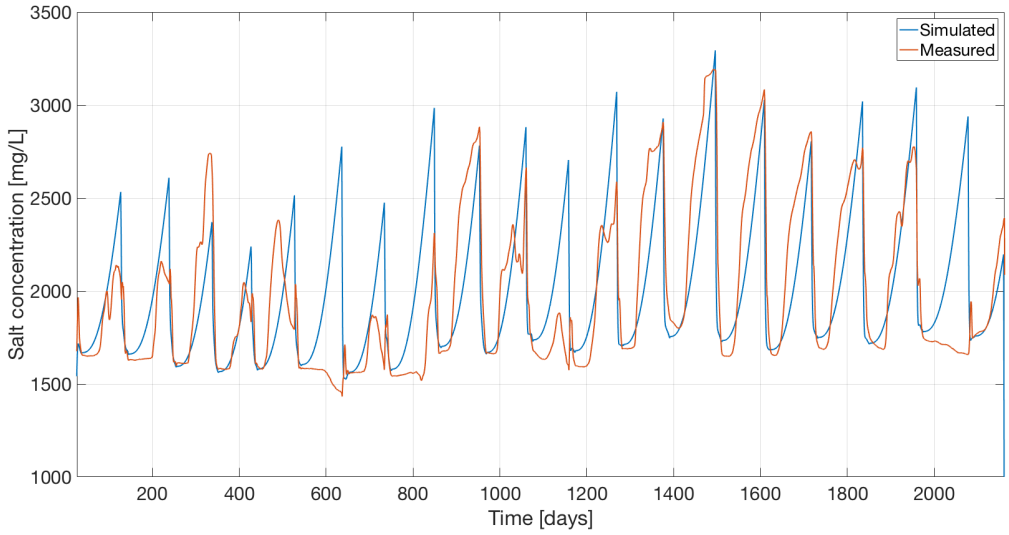



Figure 2.21: Simulation of advection-diffusion model compared with measurements in the Lissertocht

This results in a boil discharge of 0.0027 $m^3/s$, in Figure 2.21 the result of this simulation is shown. The modelled salt concentrations almost matches the measured salt concentrations at the outlet. The simulations in the model do not take into account extra water added by precipitation or other sources, in this case this does not have a big influence because it was a dry period (see Figure 2.16). In the model the salt concentrations quickly decreases when the distance gets bigger from the boil locations, in reality there seems to be more diffusion and advection of the saline water to locations which are downstream of the boils.

## 2.6. Summary measurement results

The results of the fieldwork campaign are an updated map of the locations of the boils (see Figure 2.22), detailed information about the boils and their behaviour at the three channels measured with DTS, influence of different factors on the electrical conductivity and a quantification of the boil fluxes.

A comparison has been made between the EC routing method, visual observations and DTS measurements. The EC routing method is a quick and cheap method to detect boils in a catchment but some boils are hard to detect with this method, especially in the main channels. With this method it is only possible to identify areas with high salt concentrations, which can be an indication that there are boils in those areas. However, visual observations or additional methods are needed to confirm the presences of boils and precisely locate them. EC routing and visual observations also do not give enough information on their own to determine the boil flux on a qualitative or quantitative level.

The DTS method is a very robust method and has a very high spatial and temporal resolution. During midsummer or mid-winter when the temperature difference between surface water and the seepage water is big the DTS method will always be able to detect the boils. Moreover, it is possible to compare the boil fluxes in a channel on a qualitative level and identify the influence of precipitation and pumping on the warmer seepage water from the boils. The DTS method can also be used to quantify the boil flux near a pump when the temperature distribution over the depth is also measured, which is not done in this research. However, the equipment for DTS measurements is expensive and it is laborious to install the cables.

Figure 2.22 shows the boils which were in the original SOBEK model and boils located in this thesis with EC routing, visual observations and DTS measurements. The red triangles indicate the boil locations in the original SOBEK model, these boils are identified based on the EC map of 2011. The orange squares indicate

channels where a disturbance in electrical conductivity is measured during the two EC routing campaigns in 2018, during this thesis. The purple diamonds show the boils which are identified during the EC routing in 2018 based on visual indications of boils. The green circles show the boils which are identified with the DTS measurements during this research at the two measurement locations. This map shows that there are many areas in the catchment where a disturbance in electrical conductivity is measured during one of the EC routing campaigns in April, February and May 2011 or in March and May 2018. However, it is not always possible to identify which boil is responsible for this disturbance. Some of these disturbances in EC are confirmed by visual observations, DTS measurements and/or a second EC routing measurement, in this case there must be a boil at that location.

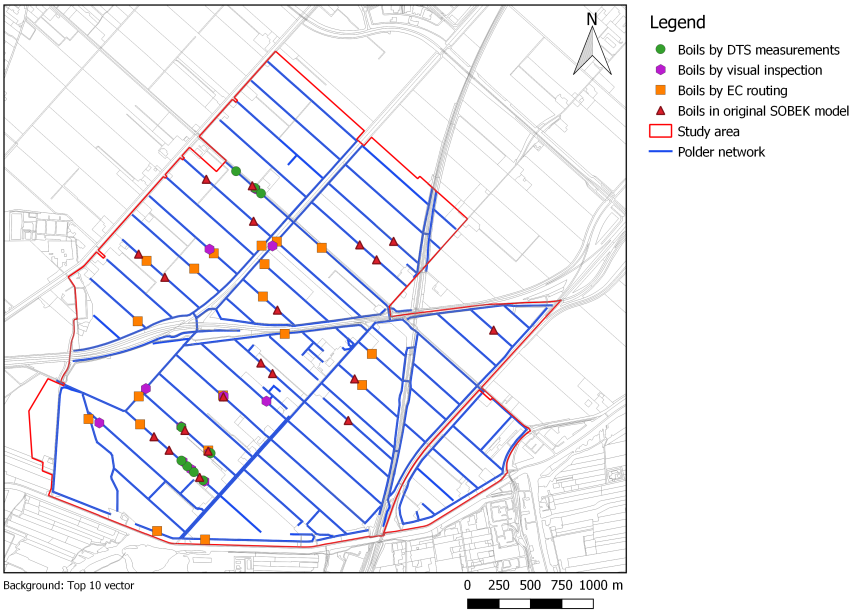



Figure 2.22: Map of the boils in the catchment, identified during this thesis with EC routing, DTS and visual inspection

The measurement with the CTD divers show that the electrical conductivity values in the channel which leads to the outlet (the Lissertocht) are influenced by precipitation in the catchment, when it rains the EC values decrease. At other locations more upstream this relationship is not visible because less drainage channels drain into these channels. When there is a dry period the pump will not be turned on and the catchment can only be flushed by opening the inlets. The CTD divers show that the electrical conductivity values increase in this situation. In this thesis there are new boil locations identified and some boil locations from the SOBEK model have been confirmed.

The mean boil flux in the original SOBEK model was set to $0.002 m^3/s$ per boil. The dimensionless numbers indicated that the flow is laminar when the pump is turned off, double diffusive process can occur and conduction is the main form of heat transfer in the system. There was not enough information to determine the boil flux accurately with an energy balance because the temperature distribution over depth is unknown, therefore it was only possible to calculate that the boil flux should be in a range from $0.001 m^3/s$ to $0.015 m^3/s$. The results of the CTD divers, DTS measurements and flow measurements have been used to make a 1D advection-diffusion model of a channel with two boils. With this model a boil flux of $0.0027 m^3/s$ has been calculated. These values for the boil flux are all in the same range but to determine the precise boil flux more measurements are necessary.

# 3

# Modelling spatial and temporal distribution of salinity

In this chapter existing models for the transport of water and salt in the Lissertocht catchment are used to analyse the spatial and temporal distribution of the salinity. In the sensitivity analyses different scenarios are simulated to discover the relations between different factors such as the boil locations, boil flux, flushing discharge, flushing scheme and the pump on the salt concentrations in the catchment. Next a Principal Component Analysis (PCA) will be performed on the results of the SOBEK model; this shows that the distribution of salt concentrations can be represented with a system of only a few linear components. Finally, a predictive model is built that can be used to test the performance and compare different layouts of monitoring networks. The results from the measurements in Chapter 2 will be used to update the existing SOBEK model in Section 3.2.5.

## 3.1. Set-up of the models

The already developed models to calculate the salt concentrations, flow and water levels in the catchment consist of a SOBEK model (Kelderman, 2015) and the RSGEM model (Delsman et al., 2017). The SOBEK model first calculates the flow and water levels and then calculates the salt concentrations in the network of channels. The RSGEM model is used to calculate the groundwater exfiltration which enters the channels via the lateral fluxes in the SOBEK model. These models will be described in the next paragraphs.

### 3.1.1. SOBEK model

The SOBEK model consists of 2 modules: the SOBEK Rural module for the water quantities and the SOBEK 1DWAQ module for the water quality; these two models run in series. The SOBEK Rural module solves the Saint Venant equations and continuity equation to calculate the flow of water in the network of channels. The SOBEK 1DWAQ module uses these results to solve the advection-diffusion-reaction equation(Deltares, 2016). The grid of the SOBEK rural and the 1DWAQ module exists of 755 segments and the simulation period is from the $1^{st}$ of January 2011 until $1^{st}$ of January 2014.

The input for the SOBEK model consists of:

- Precipitation and evaporation data from the weather station located at Schiphol airport
- Lay out of the network with all structures. This is based on the administration of the waterboard Rijnland (Hoogheemraadschap Rijnland)
- Lateral fluxes from RSGEM model
- Location of the boils and their fluxes

In Figure 3.1 the layout of the SOBEK model can be seen with the most important nodes and structures. The main pump (J.P. Heye) has a capacity of $1.48 m^3/s$. The other pump (J.P. Heyepad) has a capacity of $0.42 m^3/s$; this pump is not used in the SOBEK simulations because this pump is only used to prevent flooding in extreme situations when there is a lot of precipitation. In this case the salt concentrations are not critical because the polder will be flushed naturally. There are five inlets at the southern border of the catchment and these inlets are normally opened from the $1^{st}$ of April until the $1^{st}$ of October to flush the polder. There are 20 boils in the original SOBEK model, which are placed to reproduce the EC map of May 2011, see Figure 2.1. 'peilvak 19' is connected with 'peilvak 9' with a weir, which allows fresh water to flow from the smaller 'peilvak 19' to 'peilvak 9' but not in the opposite direction. The SOBEK model has been calibrated with the EC map of May 2011 and electrical conductivity measurements at five locations in the catchment over 2011 and 2012 (Kelderman, 2015).

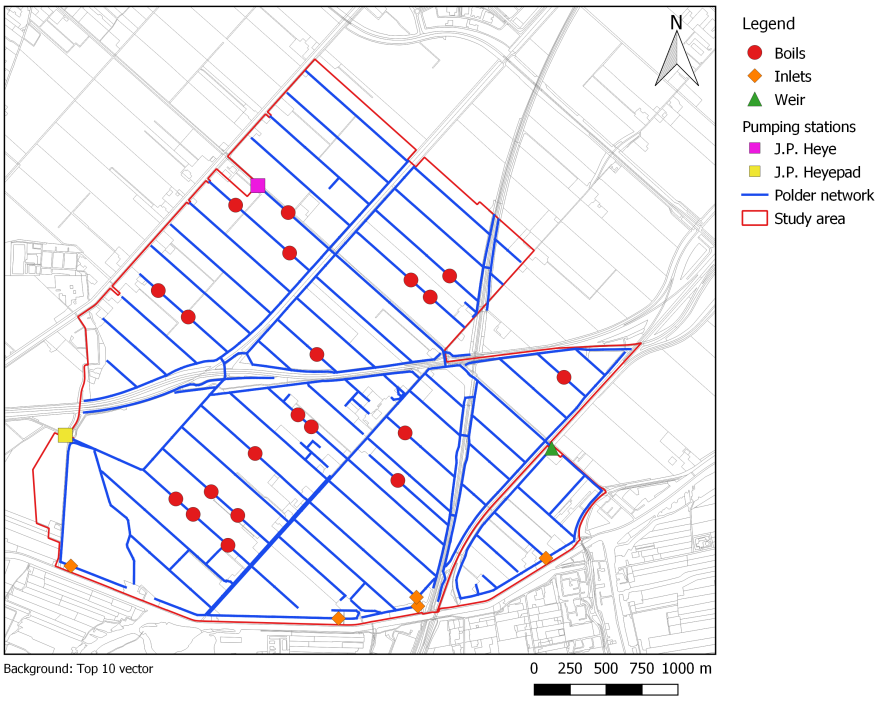



Figure 3.1: Layout of the SOBEK model of the Lissertocht catchment

### 3.1.2. RSGEM model

Delsman et al. (2017) developed a Rapid Saline Groundwater Exfiltration Model (RSGEM) to simulate groundwater exfiltration in densely drained coastal lowlands. This is a lumped water balance model which describes the flow of salt and water from an agricultural field to surface water. The main drivers which are included in this model are:

- Regional hydraulic gradient
- Precipitation
- Evaporation

The most important assumptions which are made in this model are that:

- the gradual interface between fresh and saline groundwater is assumed to be a sharp interface, and
- the, influence of density variations on groundwater flow are considered negligible.

The salinity of the exfiltrated water is determined by the ratio of groundwater from above and below the sharp interface. When there are changes in the the recharge of drainage there can be head variations. In this model the sharp interface moves instantaneously and vertically when there are head variations (Delsman et al., 2017).

This model is originally developed and calibrated based on measurements in the Schermer polder and later also calibrated and applied for the Lissertocht catchment by Kelderman (2015). In the SOBEK model of the Lissertocht catchment the fluxes from RSGEM are implemented as two different lateral fluxes into the drainage channels at the long side of each agricultural field. The first lateral flux is the exfiltration of freatic groundwater via tile drains ($Q_{drain}$) with a concentration of 75 mg/L and the second flux describes the diffuse exfiltration directly into the ditch ($Q_{ditch}$) with a concentration of 336 mg/L (Kelderman, 2015). The lateral fluxes are calculated before running the SOBEK simulations so there is no direct interaction between the SOBEK model and the RSGEM. The RSGEM model causes natural flushing of this channels when it rains because the exfiltration from the agricultural fields is fresh in the catchment. So in this case it is not a source of saline water but a source of fresh water which dilutes the saline water from the boils. In the SOBEK model the lateral fluxes $Q_{drain}$ and $Q_{ditch}$ have a constant salt concentration.

The model input (forcing data) is given by:

- Precipitation ($P$)
- Evaporation ($ET$)
- Regional seepage ($Q(h)$)
- Boundary conditions
- Model parameters, some of which can be measured such as drain level, ditch width and concentration of regional groundwater. Some other parameters such as hydraulic conductivity, effective porosity and specific yield has to be estimated and/or calibrated

In Figure 3.2 a conceptual scheme of the model is shown. Interested readers can refer to Delsman et al. (2017) for all details of RSGEM.

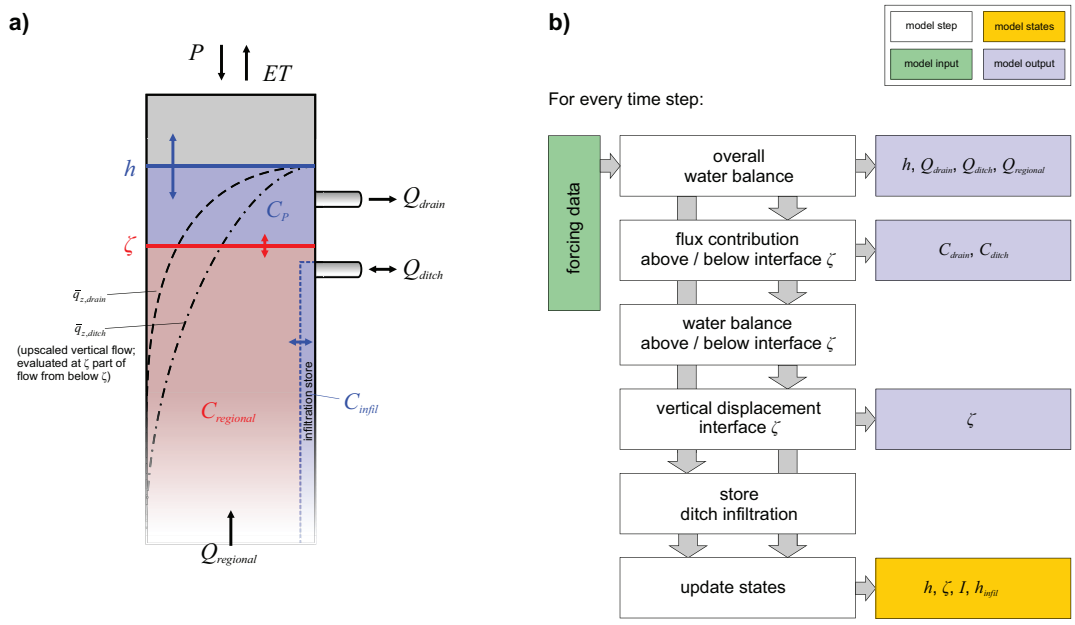



Figure 3.2: Schematisation of RSGEM model, retrieved from Delsman et al. (2017, p. 326)

## 3.2. Sensitivity analysis

In this analysis the following characteristics of the system will be analysed:

- Impact of different processes such as the boil flux, boil location and flushing on the distribution of salt concentrations in the catchment
- Travel time and preferential pathways of the flushing water, these are the channels that can be controlled by flushing
- Locate stagnant ditches which are barely interacting with the rest of the system, these ditches are also harder to flush by opening the inlets
- Search for locations which are interesting to monitor based on variance and correlation with other channels

This will be done by simulating different scenarios in the SOBEK model, which will provide information about the temporal and spatial distribution of the salt concentrations. This sensitivity analysis will give a first indication of interesting locations to monitor.

Figure 3.3 shows an overview of the modelling steps and the Python framework to generate results. This framework is made by the author to efficiently process, analyse and use the data from the different scenarios. The schematic overview shows the input for the SOBEK model and how the results of simulations from SOBEK are processed into graphs to analyse the temporal distribution and maps to analyse the spatial distribution of the salt concentrations, this to answer the second research question. Via the same framework of Python scripts the principal component analysis, variance, correlation and optimization algorithms are conducted, these analyses use the same SOBEK simulations as input data. The results of all these analysis are presented with the same templates to visualize the spatial distributions in maps and the temporal distributions in graphs.

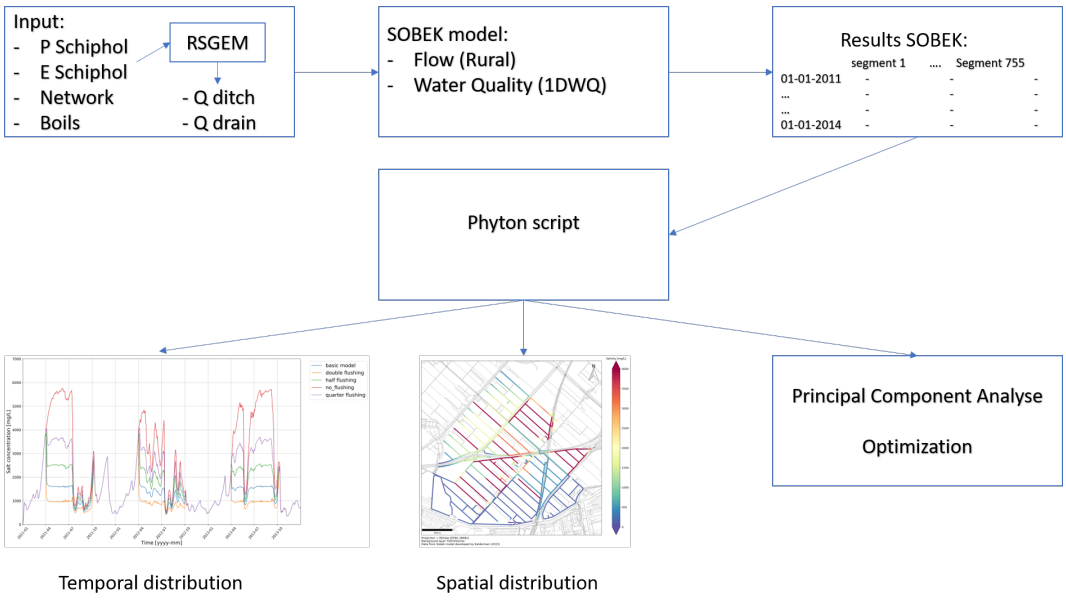


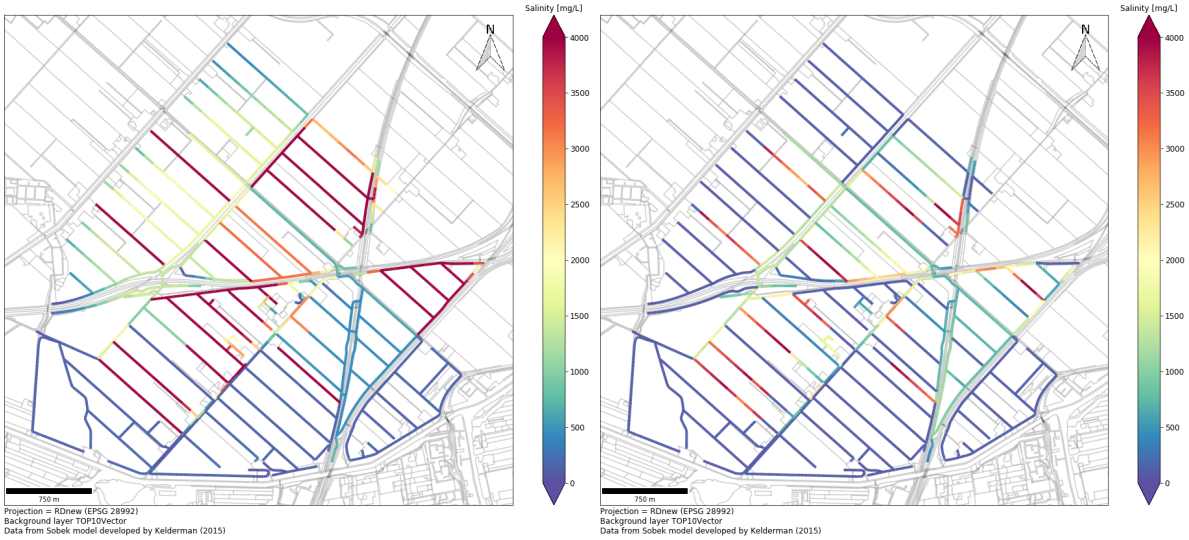

Figure 3.3: Overview of the model structure

### 3.2.1. Reference scenario

For the reference scenario the original SOBEK model developped by Kelderman (2015) is used with the layout of the boils and structures as indicated in Figure 3.1. Figure 3.4a shows that the salt concentrations are high during summer in the reference scenario. In the south-east part and along the southern border the ditches stay fresh because these areas are located next to the inlets and there are no boils located in this area. The

salinity in the main channels can increase up to 2000 mg/L. Those channels are connected to the inlets but also saline water from the boils has to flow through these channels to reach the outlet (pump) of the catchment. The drainage channels which border the agricultural fields become very saline during summer, salinity can rise up to 5000 mg/L. The drainage channels, which have very small dimensions (width 1.5 to 2.5 meter and depth of 0.3 to 0.4 meter and length of 1000 meter), are mainly flushed by drainage water (when it rains) and only a limited amount of flushing water from the inlets reaches those ditches. These factors result in low flow velocities, small water depths and limited or no supply of fresh water during the summer. Some of those ditches contain one or multiple boils.

During winter the salt concentrations are lower because of natural flushing, see Figure 3.4b. The precipitation causes natural flushing through exfiltration through the drains. There are still some areas with high salinities near the locations of the boils, those areas will always remain saline. All maps which present the results of the different scenarios show the change in salt concentrations in comparison with the reference scenario.




(a) (b)

Figure 3.4: Sobek simulation of the reference scenario: (a) shows the salt concentrations during summer and (b) shows the salt concentrations during winter

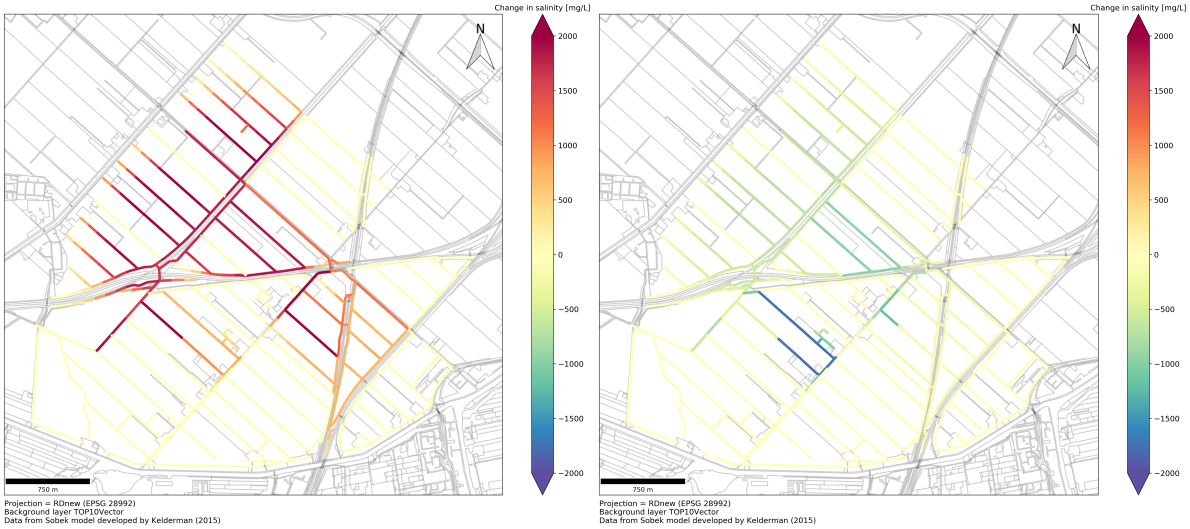
### 3.2.2. Flushing

There are 5 inlets which can flush the polder with fresh water from the Ringvaart. These 5 inlets are normally opened between the 1st of April and the 1st of October and have altogether a discharge of $0.0956 m^3/s$ (normal flushing discharge). In this analysis the flushing discharge and scheme is varied to see what the effect of flushing is and in which areas the chloride concentrations can be controlled by flushing. The scenarios which are calculated with SOBEK are:

- No flushing
- 25% of the flushing discharge compared to the normal flushing discharge
- 50% of the flushing discharge compared to the normal flushing discharge
- Double the flushing discharge compared to the normal flushing discharge
- Different flushing scheme

Figure 3.5a show the change in salt concentration during the summer when the flushing discharge is decreased to 25% of the normal flushing discharge. This shows that the salt concentrations in some areas do not change due to flushing (yellow in Figure 3.5a), these areas stay saline even when the flushing discharge through the inlets is doubled (see Figure 3.5b). These areas are at parts of the catchment which are poorly connected with the rest network of channels; in these parts, the main channels and drainage channels are

dead ends. This is caused by the two highways and the railway which crosses the catchment. In combination with a small cross section this results in almost stagnant flow. The main channels from the inlets to the outlet of the catchment are affected by the amount of flushing discharge, this can be seen in Figure 3.5a and Figure 3.5b. The salt concentrations in most of the drainage channels which are directly connected to these main channels will also decrease if the inlets are open. When the flushing discharge is doubled the salt concentrations are decreased in locations which are sensitive to flushing and still have a high salinity, see Figure 3.5b.




(a) (b)

Figure 3.5: (a) shows the change in salt concentrations during summer (three months after opening the inlets) when the catchment is only flushed with 25% of the original discharge (b) shows the change in salt concentrations during summer when the flushing discharge is doubled compared to the reference scenario

In Figure 3.6 the effect of different flushing discharges on the salt concentrations at the outlet (pump J.P. Heye) can be seen. This also shows the effects of changing the flushing discharge over time and that the effect of flushing is not proportional to the amount of fresh water which enters the catchment via the inlets.

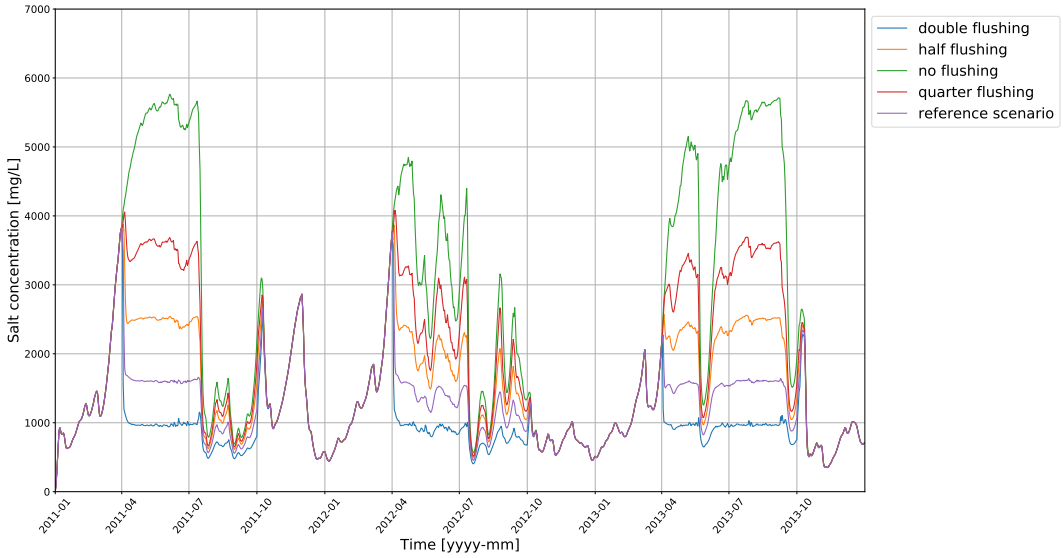
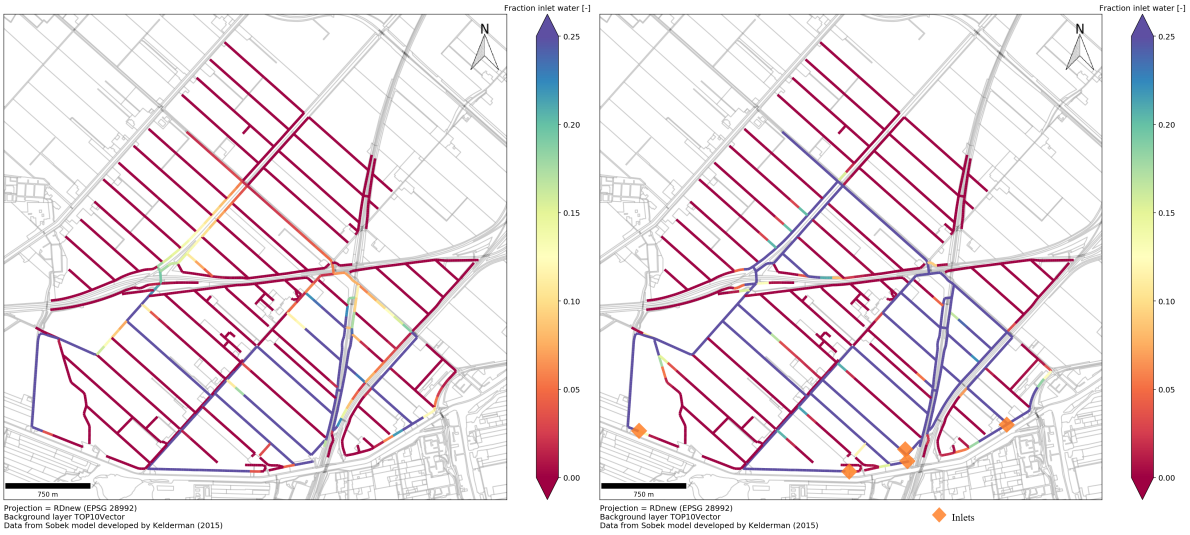



Figure 3.6: Effect of flushing on the salt concentrations at the outlet

During earlier research Kelderman (2015) and Delsman et al. (2013) modelled the different water fractions from different sources. They divided the water based on their origin: boils, flushing, precipitation, drainage and exfiltration from the shallow groundwater. Figure 3.7b shows how big the fraction inlet water is after 10 days of flushing. This shows which pathway the water takes from the inlets to the outlet of the catchment. The fractions also show that it takes approximately 3 days before the water from the inlets reaches the outlet, see Figure 3.7a. A part of the drainage channels are sensitive to flushing but it takes more than 10 days before the flushing water reaches those channels. Also the fractions from other sources are modelled by Kelderman (2015), Figure B.6 in Appendix B.1 shows that the fraction from the drainage flux (lateral flows calculated with RSGEM model) is dominant during the winter when there is enough precipitation to flush the catchment naturally. Because the drainage flux is added at all drainage channels there are no areas which are not flushed by natural flushing during the winter. The fractions also show that during the summer approximately 60% to 70% of the water at the outlet is water which originates from the inlet and approximately 30% originates from boils.




(a) (b)

Figure 3.7: (a) shows the spatial distribution of the fractions corresponding to water which originates from the inlets after 3 days of flushing (b) shows the distribution of the fractions 10 days after opening the inlets

In the scenario in Figure 3.8 the flushing scheme is changed from opening the inlets constantly from the $1^{st}$ of April until the $1^{st}$ of October to a scheme where the inlets are opened for one week and then closed for one week and then opened for one week again. This shows that the flushing water needs some days to lower the salt concentrations at the outlet and that if during summer the inlets are closed the salt concentrations will start increasing again.

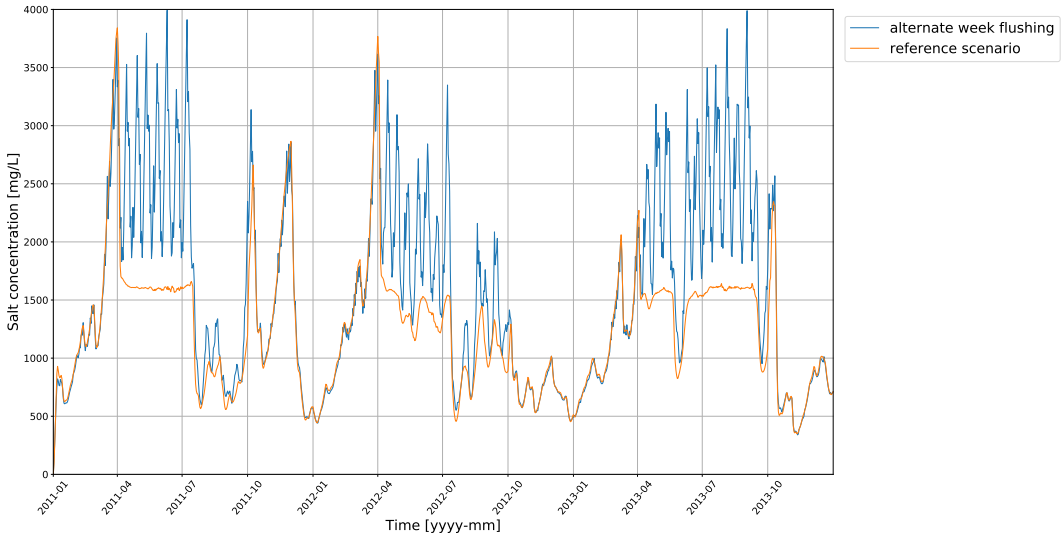



Figure 3.8: Effect of changing the flushing scheme on the salt concentrations at the outlet

### 3.2.3. Boil location and flux

In the future the flux of the boils is expected to increase due to sea level rise and land subsidence, this can increase up to 100% in 2100 in the Haarlemmermeer polder (Oude Essink et al., 2010). To see the relation between the boil flux and the salinity the following scenarios are simulated:

- Reduce flux of boils by 50%
- Increase the flux of the boils from $0.002 m^3/s$ to $0.0027 m^3/s$, this is the boil flux which is calculated in Section 2.5
- Double the flux of the boils

The results of these three scenarios can be found in Figure B.3, Figure B.4 and Figure B.5 in Appendix B.1. These maps show that changing the boil flux results in a change in salt concentrations that is proportional. At the outlet the salt concentrations increase when the boil fluxes increase, this relation is linear. The patterns or flow paths of the saline water does not change when the boil flux is changed.

Figure B.1 in Appendix B.1 shows the effect of adding one boil to a drainage channel. This boil will increase the salt concentrations locally with 2000 mg/L during a dry period. This has a big local effect because the cross section is small and the flow almost stagnant. When two boils are added to a main channel the effect on the spatial distributions of the salt concentrations is marginal (see Figure B.2 in Appendix B.1) because the flow velocity is higher and the main channels have a bigger cross section. This corresponds with the results of EC routing which also show that it is difficult to (reliable) measure the effects of boils in the main channels.

### 3.2.4. Modelling of the pump

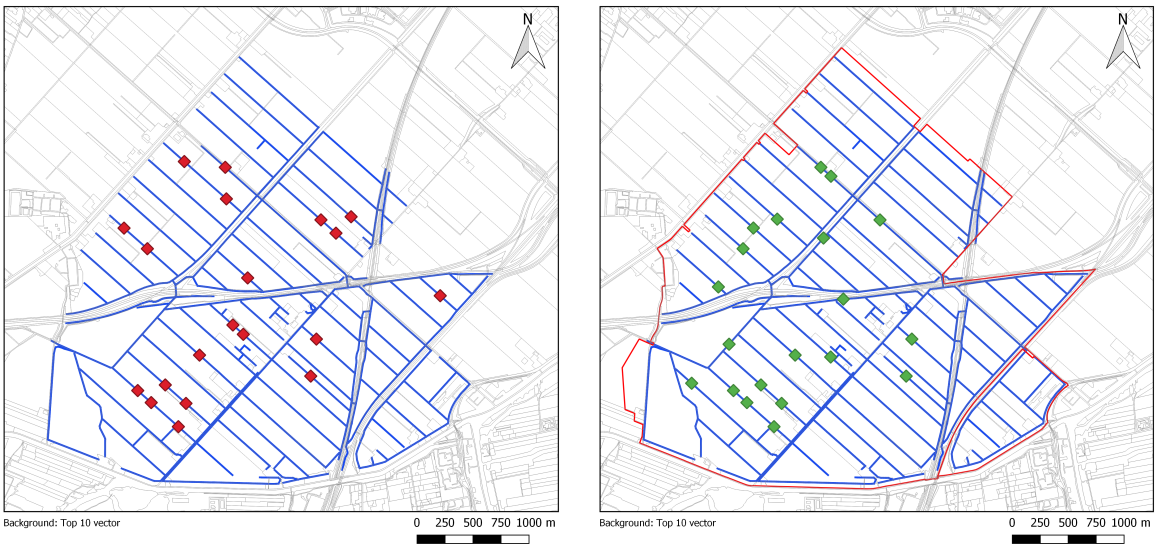
In the original model the main pump of the catchment is modelled as a boundary condition with constant water level. In reality the water level at the outlet fluctuates as the pump is switches on and off depending on the thresholds set and the measured values from the (water level) sensor next to the pump. The capacity of the pump is 1.48 $m^3/s$ and the discharge to the pump is calculated to be 1.25 $m^3/s$ (see Section 2.5.1).

When a pump is added to the SOBEK model the trend and magnitude of the salt concentrations in the catchment does not change. It will only cause small disturbances on daily time scales, so when looking at the seasonal trend of salt concentrations the line will be less smooth. At locations near the pump temporarily a decrease will be seen in salinity when the pump is turned on. These disturbances are only very small. At some small drainage channels near the pump the small disturbance has a big effect on the salinity; due to this

disturbance the water level drops by 5 to 10 centimeter in these ditches. This also causes a disturbance in the salinity in those channels. This only happens at the beginning of the drainage channels where they are connected with the main channel. In the SOBEK model the drainage channels are all directly connected to the main channels while in reality almost all drainage channels are connected with a culvert to the main channels. This is the reason why the disturbance caused by the pump is so big at the beginning of the drainage ditches. If looked at the daily or hourly time-scale the pump operation has an influence. However, if looked at the seasonal or the yearly time-scale using a boundary condition with a constant water level gives a smoother and more realistic result. This gives the advantage that there are no big disturbances at the beginning of the drainage channels.

### 3.2.5. Updated model

Figure 3.9a shows the boil locations in the original SOBEK model, in Figure 3.9b the boil locations are updated based on the measurements results (see Figure 2.22. All boils which are identified with the DTS measurements are added. Boils which are detected with the EC routing method are only added if they are confirmed by the DTS measurement, visual inspection or another EC routing. The total number of boils and the total seepage flux is kept the same as in the original model, so the salt concentrations at the outlet do not change. The salt concentration at the outlet have been calibrated with measurements when the SOBEK model was developed by Kelderman (2015).




(a) (b)

Figure 3.9: (a) shows the locations of the boils in the original SOBEK model and (b) shows the boil locations in the updated version of the SOBEK model

The updated model is simulated with SOBEK; the results are presented in Figure 3.10. This shows that some parts of the polder do not have any problem with salinity anymore when a boil is moved to another location. Some other locations have become more saline if a boil is placed nearby. The areas in the north-east part of the polder contain no boils anymore in the updated mode, Figure 3.10 shows that these channels (indicated in blue) are not saline anymore. However, the channels where boils have been added have become more saline, these channels are indicated in red in Figure 3.10.

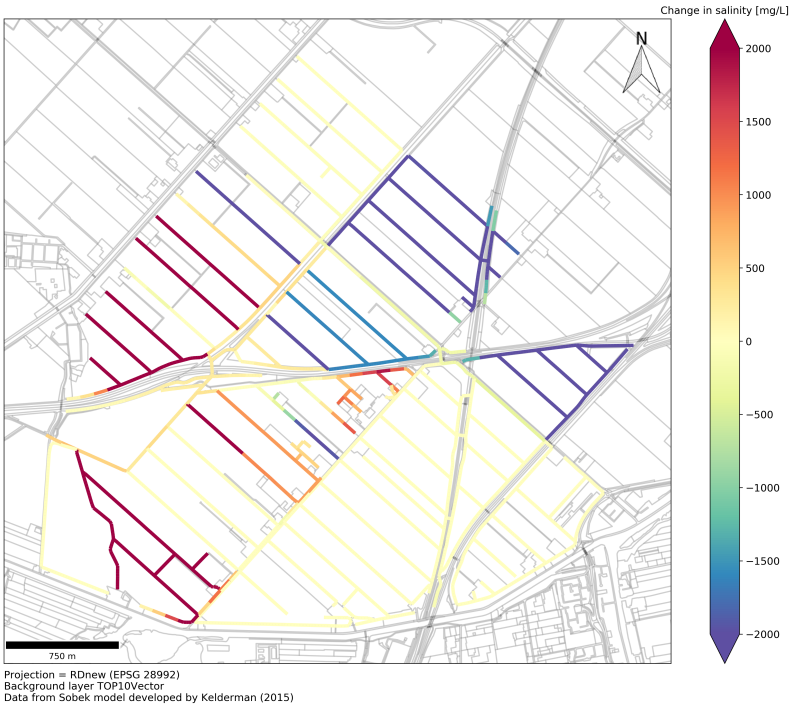



Figure 3.10: Simulation of changes in salt concentrations after updating the boil locations

## 3.3. Principal component analysis

In this section the Principal Component Analysis (PCA) is used to derive a new system of linear components which represents most of the variance in the system. The PCA is used to identify interesting locations to monitor. The PCA makes use of the variance in the system and the correlation between different locations, this is discussed by calculating the variance and correlations in this system. Finally, a predictive model is built which helps to discover the optimal locations for sensor placement and test the performance of a network of sensors.

### 3.3.1. Principle of PCA

The PCA is a method to reduce the dimensionality of a dataset while at the same time preserving as much variation as possible which is present in the original dataset. This is done by identifying the directions with the largest amount of variance, the principal components (PCs) (Jolliffe, 2002). In this way the dataset is transformed to a new set of variables (the principal components) in a lower dimensional space, which has two important properties:

- The first few components contain the most variation from the original dataset
- The PCs are perpendicular to each other

The first principal component (PC 1) is the direction along the largest amount of variance occurs, the second principal component (PC 2) is a direction orthogonal to PC 1 and contains the second largest amount of variance, this procedure continues until a new set of linear components is produced which contains enough of the variance in the original dataset. The principle of PCA is showed graphically in Figure 3.11.

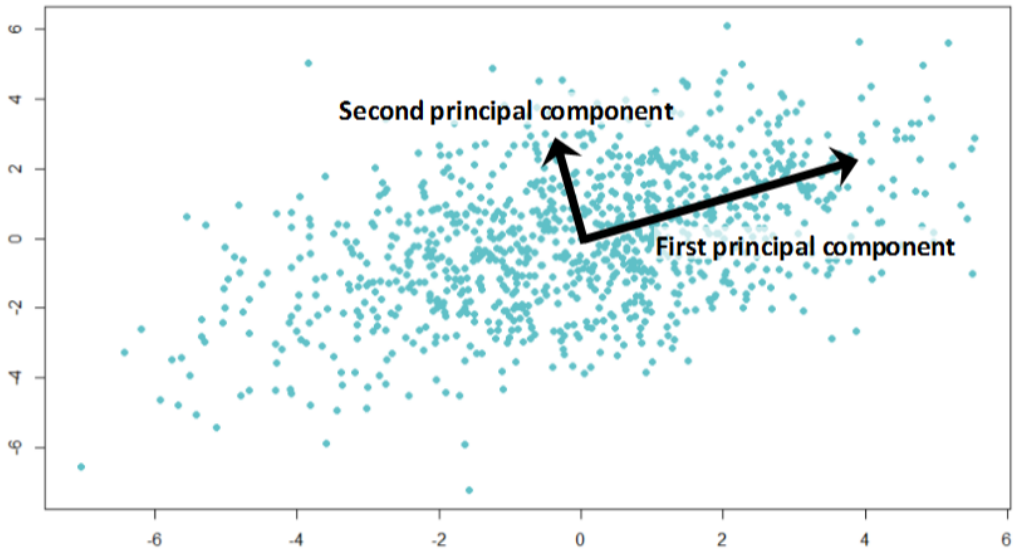



Figure 3.11: Graphical explanation of PCA method,
retrieved from Analytics Vidhya

The first step of the PCA is to search for a linear function $\alpha_1^{'}x$ having maximum variance for the dataset of $m$ x $n$ variables:

$$\alpha_1^{'}x = \alpha_{11}x_1 + \alpha_{12}x_2 + ... + \alpha_{1p}x_p = \sum_{j=1}^{n} \alpha_{1j}x_j \tag{3.1}$$

In the next step the PCA searches for a second component which is uncorrelated with $\alpha_1^{'}$ x and which has maximum variance. At the end the PCA has transformed the dataset in a new set of uncorrelated linear components:

$$\alpha_1^{'}x + \alpha_2^{'}x + ... + \alpha_k^{'}x \tag{3.2}$$

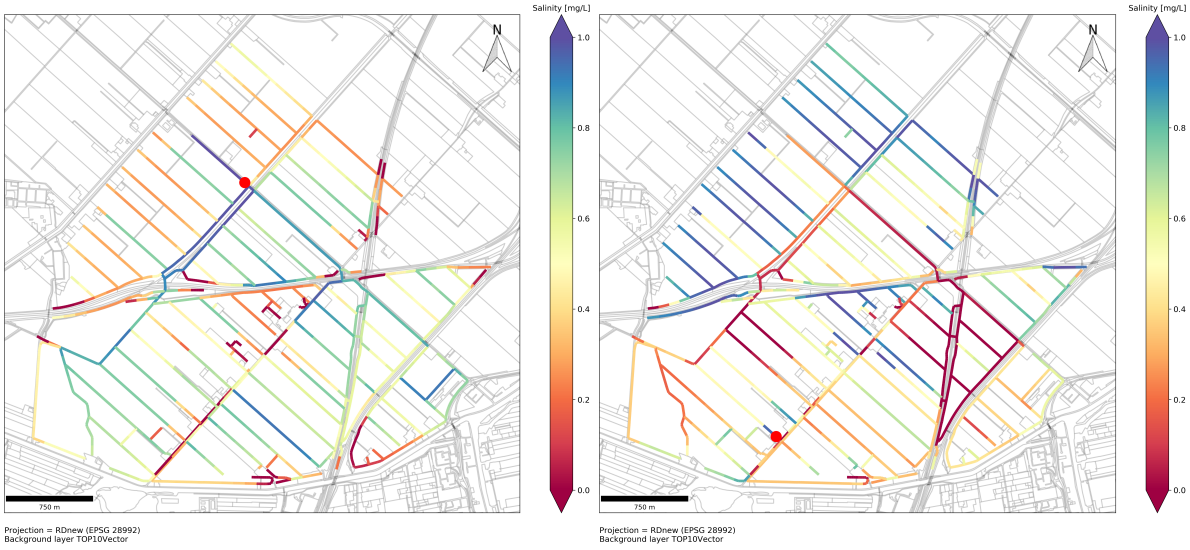
In which $k$ corresponds to the $k^{th}$ principal components which is smaller than the amount of variables $n$. The goal of the PCA is that most of the variance in $x$ can be represented with $k$ PCs, where $k << n$. The matrices of coefficient $\alpha_1^{'}, \alpha_2^{'}, ..., \alpha_k^{'}$ are location dependent and constant over time; these indicate what combination of the $k$ principal components is needed to reconstruct the value at a certain location. The vectors $x_1, x_2, ..., x_k$ are the time signals of the principal components and vary over time(Jolliffe, 2002).

### 3.3.2. Variance and correlation

Two important aspects of the PCA method are variance and correlation. The amount of variance at a location explains how much the salt concentration is different from the mean salt concentration at that location, so it is a measure of the amount of dynamics. This is also an important factor when a monitoring system is designed. A location with a lot of variance is more interesting to constantly monitor than a location which does not have any variation over the year. Figure 3.12 shows the variance in the system for the reference scenario. The locations which have the biggest variance in this figure correspond with the locations where the first principal component is dominant (see Figure 3.16a), these are the drainage channels with boils. The variance at location $X$ is defined as: $Var(X) = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2$ (Statistics How To, 2018), in which $N$ is equal to the number of values in $X$, $X_i$ is the $i^{th}$ value in $X$ and $\mu$ is the mean value of $X$.


Figure 3.12: Variance in the reference scenario

The second important factor of PCA is correlation, this is an indication of the relation between two variables, here locations. This is also an important aspect in solving a sensor placement problem because it is not interesting to measure at two locations if they behave very similarly. In this case one location can be measured to know the behaviour of both locations. In Figure 3.13a and 3.13b some examples of the correlation between different places in the polder network is presented. The Pearson correlation coefficient between location $X$ and location $Y$ is defined as: $\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X * \sigma_Y}$ (The SciPy community, 2018), in which $\sigma_X$ is the standard deviation of $X$ and $Cov(X,Y)$ the covariance between $X$ and $Y$.




(a) (b)

Figure 3.13: Correlation between location (a) (indicated with a red circle) and the rest of the system and correlation between location (b) (indicated with a red circle) and the rest of the system. These correlations are calculated for the reference scenario (see Section 3.2.1

The PCA tries to maximise the variance by composing a new set of linear components which are orthogonal to each other. Therefore, it also takes correlation between locations into account because the components are orthogonal to each other. So if two locations contain a lot of variance but the behaviour at those locations

is not correlated the PCA will make at least two principal components to construct a new system. However, if both locations are strongly correlated, it is possible to represent it with only one principal component and neglect the second principal component.

### 3.3.3. Classification based on PCA

From SOBEK a dataset is obtained of salt concentrations at all 755 segments for 1097 days (3 years). This results in a dataset of dimension 1097 by 755. The linear relations between the different modelled segments are explored by applying the PCA method. The goal is to reduce this multidimensional space to a space with less dimension which still captures most of the temporal variance in the dataset.

The PCA on this dataset results in 755 new principal components which are sorted based on the amount of variance they represent. The first principal component explains 72% of the variance in the dataset, the second 13% and the third 6%. This shows that with 3 principal components 91% of the variance is explained, see Figure 3.14. The rest of the principal components will be neglected because they explain less than 9% of the variance all together. This results in a reduction from a 755 to a 3 dimensional space.

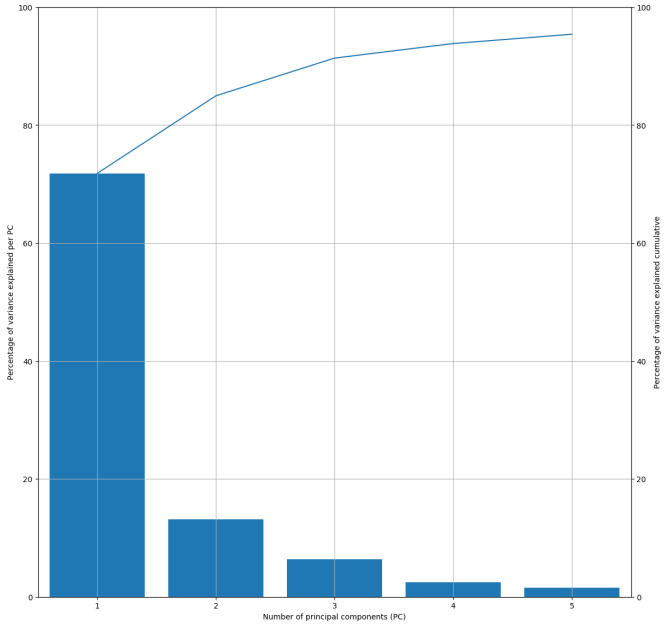



Figure 3.14: Variance represented by the first 5 principal components

Figure 3.15 shows the time dependent component x of the first three principal components. The first principal component increases during the summer (dry period) and decreases during the winter (wet period). Figure 3.16a shows that the location dependent coefficient (loading) of the first principal component is dominant in drainage channel. These have problems with salinity because they contain a boil or a channel nearby contains a boil. This behaviour corresponds to the increase in salt concentrations during the summer and a decrease in salt concentrations when these channels are naturally flushed when it rains.

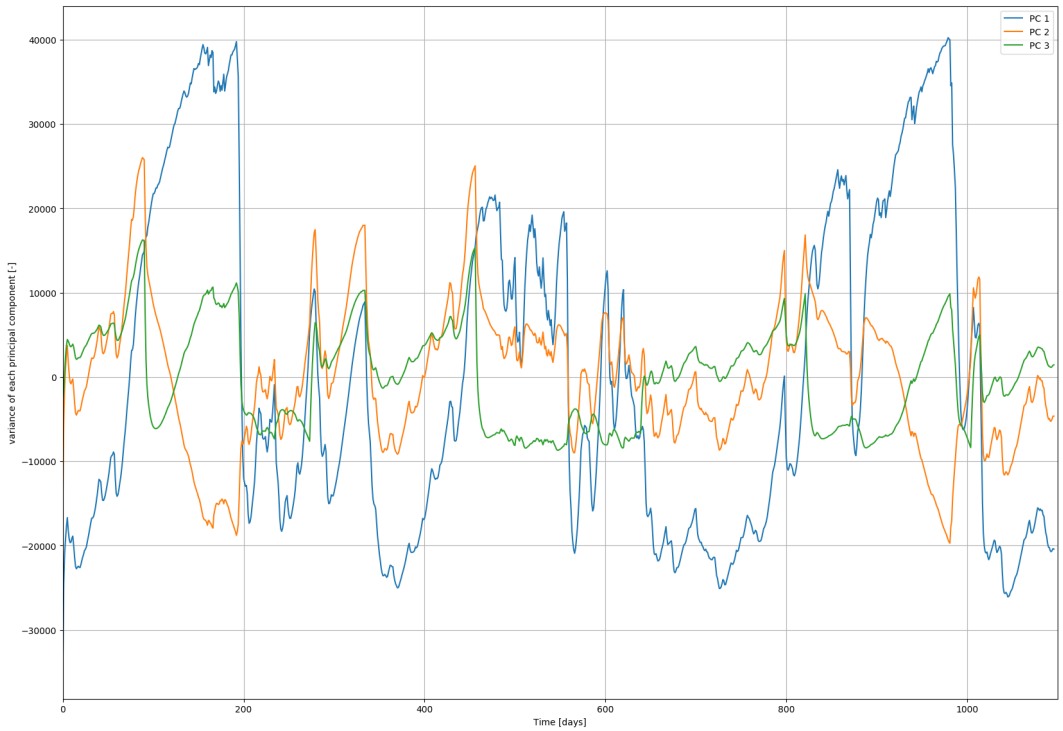
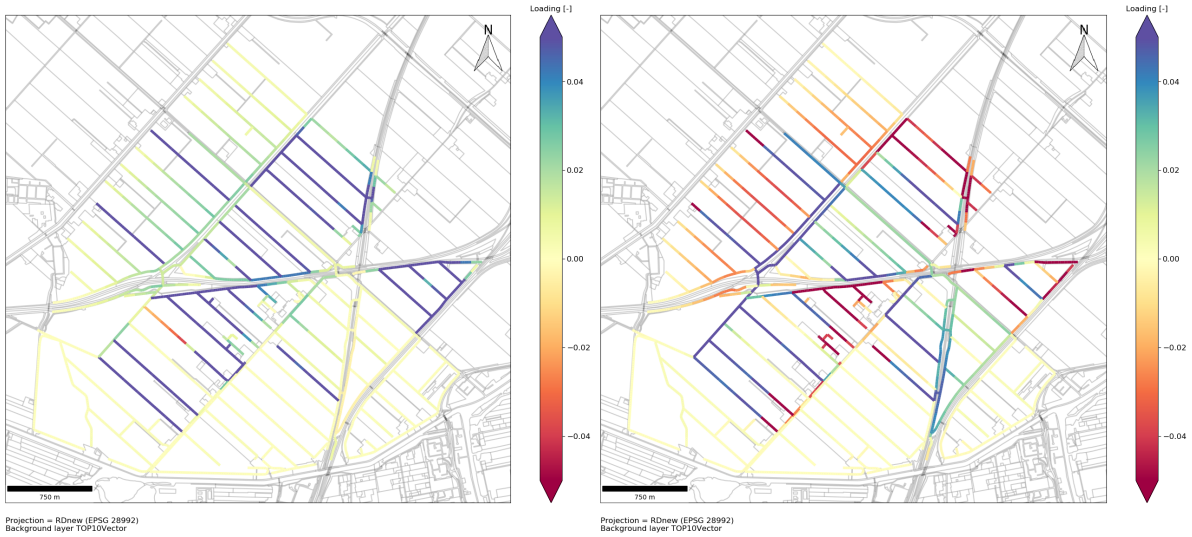



Figure 3.15: Time signal of the principal components

The second principal component decreases when the inlets are opened at the $1^{st}$ of April, this corresponds to the behaviour of main channels which are flushed during the summer by fresh water from the inlets. The loadings in Figure 3.16b show that the second principal component is big in main channels and in drainage channels which are sensitive to flushing. At these channels the second principal component lowers the first one, the same thing is happening with the salt concentrations in the real system. If the inlets will be closed during the summer, the salt concentrations will be high in these channels (represented by the first principal component), however, when the inlets are open the salt concentrations will decrease due to the supply of fresh water (represented by the second principal component).




(a) (b)

Figure 3.16: (a) shows the loadings of the first principal component and (b) shows loadings of the second principal component

The PCA has showed that most of the variance (90%) of the variance in the system can be represented with a system of three linear components. Moreover, the first two principal components can also be physically interpreted, where the first is dominant in drainage channels with high salinity during a dry period and the second is present in channels which are flushed if the inlets are opened. This shows that the PCA method can be used to make a classification between different categories of channels.

### 3.3.4. Predictive model

If it is possible to represent the system with three linear components, then it may also be possible to use this system of linear components to test the performance of the sensor placement in a measurement network. This predictive method makes use of the location dependent coefficients (loadings) which are derived with the PCA at each location. Each location has one loading for each principal component of the PCA, the magnitude of these loadings gives an indication which combination of components are dominant at this location. If in a new dataset, the salinity is measured at three different locations it is possible to derive the time signals (time dependent component) of the first three principal components using measurements. The time signals together with the coefficients can be used to reconstruct the salinity levels at every point in the system for the new time period.

Pseudo algorithm for the predictive model:

1. Split dataset in train and test set:
   $y_{train}$ = 548 x 1097
   $y_{test}$ = 548 x 1097

2. Perform PCA analysis on training dataset ($y_{train}$) and decide how many new principal components ($n$) are taken into account. Retrieve time signals (scores) for 548 time steps for $n$ components of the PCA on the dataset $y_{train}$:
   score = 548 x $n$ :

   $$\left[\bar{x}_1, \bar{x}_2, \bar{x}_3, \cdots, \bar{x}_n\right] \tag{3.3}$$

   Retrieve coefficients (loadings) for all locations for $n$ components of the PCA on the dataset $y_{train}$:
   loading = $n$ x 755

   $$\left[\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_n\right] \tag{3.4}$$

   In the next steps only the loadings for each component at all location will be used.

3. Choose $m$ virtual measurement locations. At the measurement locations the salt concentrations from the dataset $y_{test}$ will be taken.

4. Derive $n$ time signals based on those $m$ sensor locations:

   $$y_{i,t} = \sum_{n=1}^{n} \alpha_{1,n} * \bar{x}_n(t) \tag{3.5}$$

   With least square method find a solution for time signals ($\bar{x}$) for $n$ components.
   If $m = n$, the problem is determined and can also be solved by multiplying the inverse of $\alpha$ by y.
   If $m > n$, the problem is overdetermined.

5. Calculate the salinities at all locations by multiplying predicted time signals with coefficients from PCA on training set (step 2):

   $$\tilde{y}_{predicted} = \sum_{n=1}^{n} \alpha_n * \bar{x}_n \tag{3.6}$$

6. The predicted salt concentrations are compared with the modelled salt concentrations (true values). The measure of fit $R^2$ is used to test the performance of the sensor network, $R^2$ is formulated as (Scikit Learn developers, 2017):

$$R^2 = 1 - \frac{\sum(y_i - \tilde{y}_i)^2}{\sum(y_i - \bar{y})^2} \tag{3.7}$$

With:

- $\tilde{y}$ = predicted value by the predictive model
- $\bar{y}^2$ = mean value of all true values
- $y$ = true value

In Figure 3.17 an example is presented of this model, in this case sensors are placed at location 89, 744 and 661. These sensor locations are chosen based on the outcome of the classification by PCA, the modelling results and measurement results. This results in a sensor in a drainage channel with a boil, a sensor in a main channel which connect the inlets with the outlet and a sensor near the outlet.

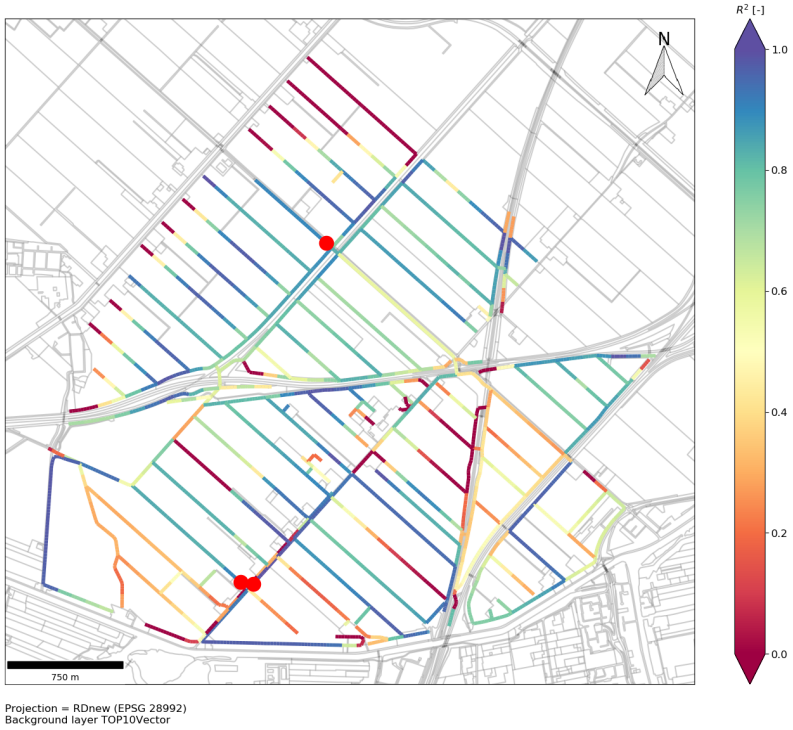
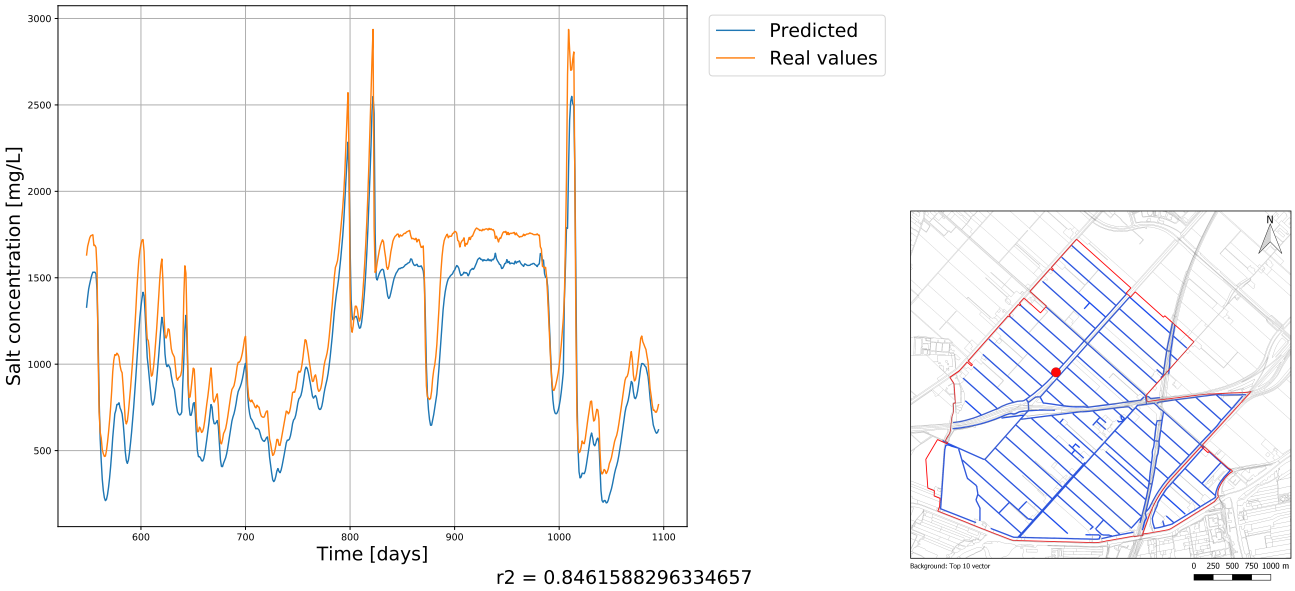



Figure 3.17: Performance of a monitoring network with sensors at locations indicated with red dots

The model shows that with the variance which is measured at these locations the salt concentrations can be predicted at some other areas reasonable well but that their is still room for improvement. In Figure 3.18a the fit at location 542 can be seen. This shows that the predicted salt concentrations are not exactly the same but the trend is predicted well.




(a) (b)

Figure 3.18: Figure (a) shows the performance of the monitoring network to predict the salt concentrations at another location (indicated with red dot in Figure (b)

One of the limitations of testing the performance of a monitoring network with this method is that both the dataset for the training period as for the test period are modelled salt concentrations by the same SOBEK model. Therefore, errors in the SOBEK model will also exist in this model. The virtual measurements which are used to reconstruct the system of linear equations are obtained by revealing the salt concentrations in the test dataset at the chosen sensor locations. However, this model gives a lot of knowledge where the variance is present in the system and how different locations are correlated. The model is a tool which says how good the chosen sensor locations are capable of capturing this variance and correlation. The model will be good in predicting the patterns in salt concentrations but the exact prediction is as good as the accuracy of the SOBEK model. Additional calibration and validation of the SOBEK model will improve the prediction of the absolute salt concentrations. If this model will be conducted with real measurements instead of virtual measurements and by comparing it with a real testing set which is obtained by measuring at multiple locations where there are no sensors it is possible to discover if the variance and correlations also exist in the real catchment. In the Chapter 4 optimisation algorithms are used to find the optimal locations.

## 3.4. Summary modelling results

In this chapter the spatial and temporal distribution of the salt concentrations has been modelled with the existing SOBEK and RSGEM model of the area. The simulations of different scenarios shows that the two highways and the railway which cross the catchment creates parts of the polder where the flow is stagnant and which are not sensitive to flushing. However, the sensitivity analysis also shows the main pathway of the flushing water, these main channels and the channels directly connected to them are sensitive to flushing. The analysis of the fractions shows that it takes 2 to 3 days before the flushing water from the inlets reaches the outlet. The results from the EC routing, visual observations and DTS measurements has resulted in an updated reference scenario, in this scenario the boil locations have changed.

The PCA shows that the catchment can be represented with a system of three linear components, this represents 90% of the variance in the system. The time signals and spatial distribution of the coefficients of these three linear components shows that the first component is dominant in drainage channels with salinity problems. The second component is high in channels where flushing has an effect. This proves that the PCA can be used to classify the channels based on variance and correlation, this is confirmed by calculating and

plotting the variance and correlation separately.

With this new system of linear equations, a model is built that can be used to test different sensor layouts. This model uses the coefficients of each linear component that is determined with a training set of SOBEK results. Three or more visual measurements can be taken from the testing set, with these visual measurements and the PCA coefficients determined using the training set, the salinity levels at all locations can be reconstructed. With this system of linear equations the salt concentrations in the whole catchment can be reconstructed (predicted). The model gives feedback on how much variance and correlation can be captured by placing sensors at several locations. However, errors between the SOBEK model and the reality will exist, so the prediction of the absolute values of salinity is (optimally) as good as the simulations in SOBEK. This model will be used in the next chapter to optimise the sensor placement.

With the modelling results of this chapter and the measurement results of the previous chapter, already some key locations for sensor placement can be located (research question three). One important location is near the outlet of the catchment at the pump J.P. Heye, the salt concentration at this location has a high variance and is influenced by the precipitation, flushing scheme and discharge and operation of the pump. This has become clear with the measurements with the CTD divers which showed the effect of the precipitation and pump operation at this location. The sensitivity analysis in SOBEK has showed that changing the total boil flux, flushing discharge and changing the flushing scheme has effect on the salinity at this location. Moreover, the PCA, variance and correlation coefficient show that this location is important. A second interesting location is a drainage channel which contains a boil, at this type of locations the effect of natural flushing via drainage in a wet period and salinization in dry periods is clearly visible. A third category is locations near the inlets and that become more saline if the catchment is not flushed. These locations clearly show the signal related to flushing the polder by opening the inlets. In Chapter 4 this knowledge will be used for the exact placement of sensors.

# 4

# Optimisation of a monitoring network

The predictive model built in the Chapter 3 is a tool to test the performance of the sensor placement. With the measurement result, simulations in SOBEK and classification with PCA three categories of interesting locations were derived. However, within these categories there are still a lot of different combinations of locations possible. In this chapter, three different strategies are used for the placement of a sensor in a monitoring system. The first one is expert opinion of the author of this thesis; this is based on the knowledge from fieldwork and modelling. The second strategy is trying all possible combinations with an exhaustive algorithm and the third strategy uses a greedy algorithm to find the optimal sensor layout. At the end of the chapter the business case is discussed and guidelines for setting up a monitoring network in a new catchment are discussed. This chapter will provide the answers on the last two research questions.

## 4.1. Objective of the optimisation

The predictive model from the previous chapter will be used to calculate the $R^2$ error of the predicted salt concentrations based on the virtual measurements on the locations where a sensor is placed. In Figure 4.1 the main channels are indicated in green, it is assumed that the farmers will take water from these channels. In this way all farmers will have access to fresh water at the short sides of their agricultural fields. The drainage channels at the long sides (1000 meter) of the agricultural fields have a small cross section, small water depth and low flow velocities. When a farmer puts a pump in one of those channels they will be empty quickly. The cross section, water depth and flow velocities of the main channels are bigger, which makes them more suitable for the purpose of irrigation. The objective of the optimisation strategies is formulated as: *Maximise the mean $R^2$ error over all the main channels.*

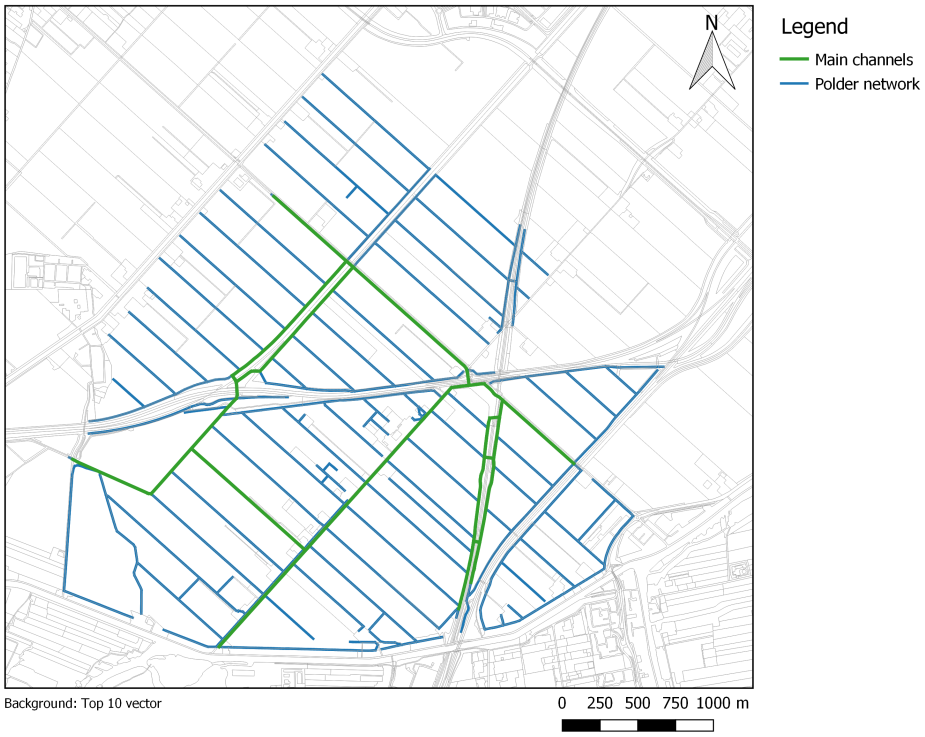



Figure 4.1: The main channels in the polder (indicated in green), where the performance of the monitoring network will be assessed

In reality, some farmers located near the border of the catchment will use water from outside the catchment. Some other farmers in the catchment which cultivate potatoes are not allowed to irrigate their crops to prevent potato diseases. However, the farmers who are cultivating flowers are the most vulnerable for salinization of the surface water in the polder. Moreover, because of different types of crops, the demand for fresh water for irrigation will vary during the year. These factors are not taken into account in this research because precise information is not available and these topics are outside the scope of this research.

## 4.2. Strategies to optimise sensor placement

The PCA has showed that it is possible to build a system of linear relations of only three components to be representative for the results of the SOBEK model. In addition, it was also possible to use the PCA to construct a model which tests the performance of the sensor placement. In this section, optimal sensor placement will be searched and the predictive model will be used as a tool to test the performance of the discovered sensor placements. An expert opinion, exhaustive algorithm and a greedy algorithm will indicate the optimal sensor layouts. Finally, the robustness of the monitoring network will be tested.

### 4.2.1. Expert opinion

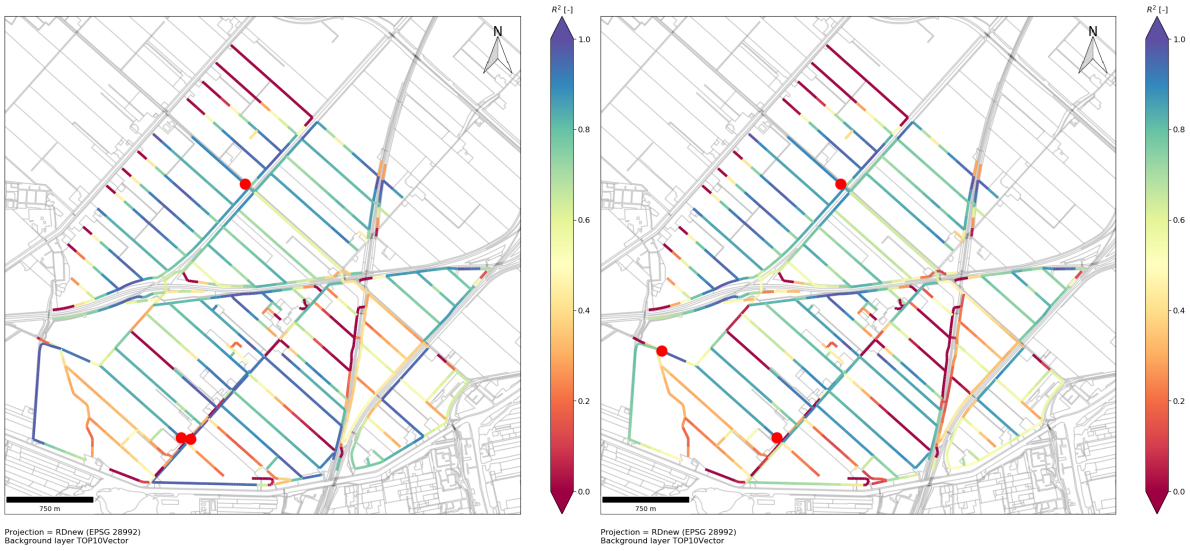
The measurements with DTS show that the temperature profile near the pump is very dynamic, this can be an indication that the variance of the salt concentrations is also high at this location. The measurements with CTD divers show that the electrical conductivity is dynamic at this location and that there is a relation between the electrical conductivity and the precipitation and the pump operation. So from an expert opinion view it is logical to place a sensor near the outlet of the catchment.

The PCA analysis has showed that the variance over the years is the biggest at the drainage channels with boils, in those channels, the first principal component is dominant. At those channels the salinity is high during summer because those channels contain boils and the flushing does not affect these locations. However, when it rains the salinity will decrease in these channels because drainage water from the agricultural

fields dilutes the saline water from the boils. So a second sensor location will be in one of the drainage channels which contains a boil, for example the first measurement location of the DTS measurement.

The last sensor should be located in the pathway of the fresh water, which travels from the inlets to the outlet. This sensor will measure the signal that corresponds to flushing the catchment. In Figure 3.7b the pathway of the water from the inlets is presented. This path corresponds to the main channels of the catchment. The last sensor should be placed in one of these channels which is saline when there is no flushing, otherwise it will not measure a change in salt concentration when the polder is flushed. In the southern part of the polder the channels are almost all fresh; so it is better to place the third sensor in a main channel in the northern part of the catchment.

In Figure 4.2a and Figure 4.2b two monitoring networks are designed by choosing one location from each category. The figure shows that these monitoring networks perform reasonable compared with the reference model, this model is defined in Section 3.2.1.

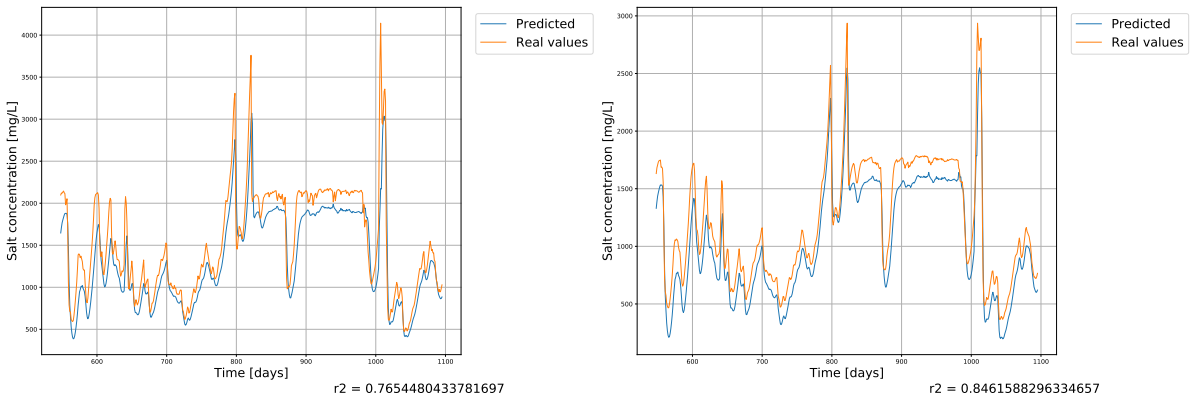



(a) Sensor layout: 89 - 744 - 661

(b) Sensor layout: 240 - 744 - 661

Figure 4.2: [Performance of sensor network based on expert opinion](4.2a) and (4.2b) shows the performance (expressed in $R^2$) of two different sensor layouts, that are chosen with expert opinion and based on the three formulated categories. The red dots indicate the locations of the sensors

In Figure 4.2a and Figure 4.2b two possible layouts of sensor placements are analysed with the predictive model. With this layout the PCA predicts the trends of salinity well for almost the entire catchment, see for example the fit at two different locations in Figure 4.3a and Figure 4.3b.




(a) (b)

Figure 4.3: Shows the predicted salt concentrations at two different locations (both main channels). This prediction is based on measurements with the sensor layout presented in Figure 4.2a. These predicted values are compared with the modelled salt concentrations

By using an expert opinion method for the sensor placement the sensor placement is intuitive and well explained. The drawback of this method is that it cannot explain why another sensor layout from the same three categories will perform better or worse. It ends at the placement of three sensors, it does not give an indication on how much the network will improve if more than three sensors are placed and where they should be placed.

### 4.2.2. Exhaustive algorithm with reduced search space

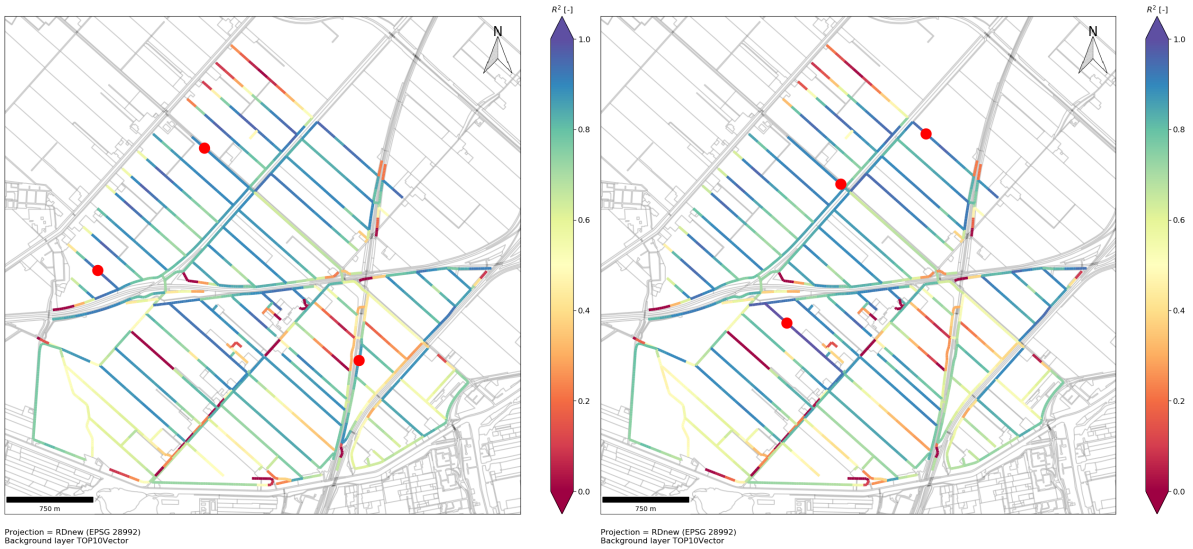
An exhaustive algorithm calculates and tests all possible combinations, it is also called the brute force method (Weisstein). This is an efficient method when the amount of possible combinations is limited.

In this case there are 71.159.504 combinations possible if three sensors are placed out of 755 possible locations. This number is too big, so it is necessary to reduce the search space. By taking only the $5^{th}$ segment the amount of segments can be reduced to 151, this results in 562.475 possible combinations. The results of doing an exhaustive algorithm with these combinations is presented in Table 4.1 and the performance of the best two layouts is presented in Figure 4.4a and Figure 4.4b. The results show that the layouts of sensors which perform good have at least sensors from two different categories which are formulated as key locations earlier, some even have one sensor in each category.

Table 4.1: Results of exhaustive algorithm

| **Layout** | $R^2$ |
|---|---|
| 271 - 666 - 517 | 0.659 |
| 356 - 661 - 691 | 0.659 |
| 356 - 666 - 517 | 0.658 |

To optimise the sensor placement with more than three sensors the amount of combinations will increase rapidly and it will not be possible anymore to use the exhaustive algorithm for solving this problem.




(a) Sensor layout: 271 - 666 - 517

(b) Sensor layout: 356 - 661 - 691

Figure 4.4: Shows performance for the top two layouts from the exhaustive algorithm

### 4.2.3. Greedy algorithm

A greedy algorithm is an algorithm which chooses the locally optimal solution, this is not necessarily the global optimal solution (Moore and Ross). A greedy algorithm always chooses the most beneficial solution locally without looking ahead or backwards. For a lot of problems this is not the optimal optimisation strategy but if the problem is defined in the right format it can be a very efficient algorithm.

A greedy algorithm will first search for the best sensor network if it is only allowed to place one sensor, in that case there are 755 candidate locations. The algorithm will calculate all 755 possibilities and then choose the one which performs best. In the second optimisation round the greedy algorithm will add one extra sensor, in this round there are 754 candidate locations left because one location is already occupied with a sensor. In the case of a sensor network with 3 sensors the algorithm needs: $755 + 755 - 1 + 755 - 2 = 2.262$ calculations while the exhaustive algorithm needs to calculate 562.475 combinations.

Table 4.2: Results of greedy algorithm for the reference scenario

| **Location** | **$R^2$ main channels** |
|---|---|
| 67 | 0.448 |
| 67 - 173 | 0.52 |
| 67 - 173 -496 | 0.655 |
| 67 - 173 -496 - 381 | 0.717 |
| 67 - 173 -496 - 381 - 257 | 0.744 |
| 67 - 173 -496 - 381 - 257 - 367 | 0.76 |

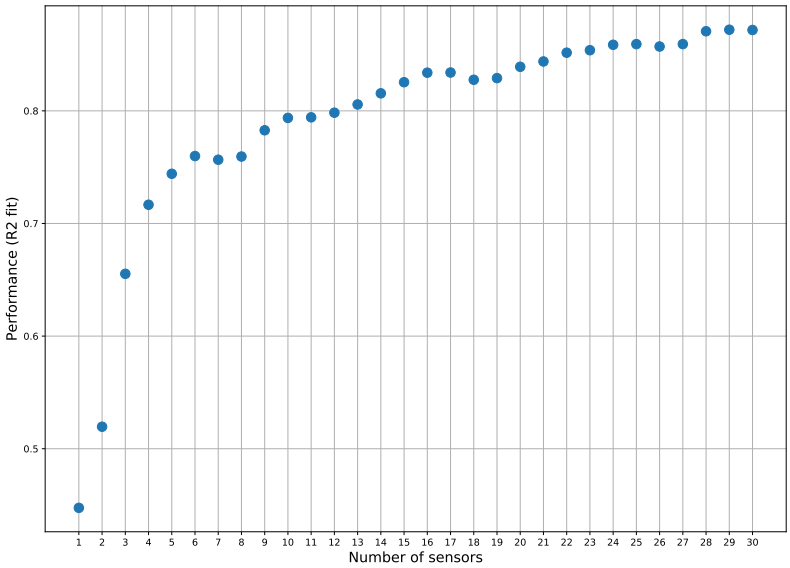
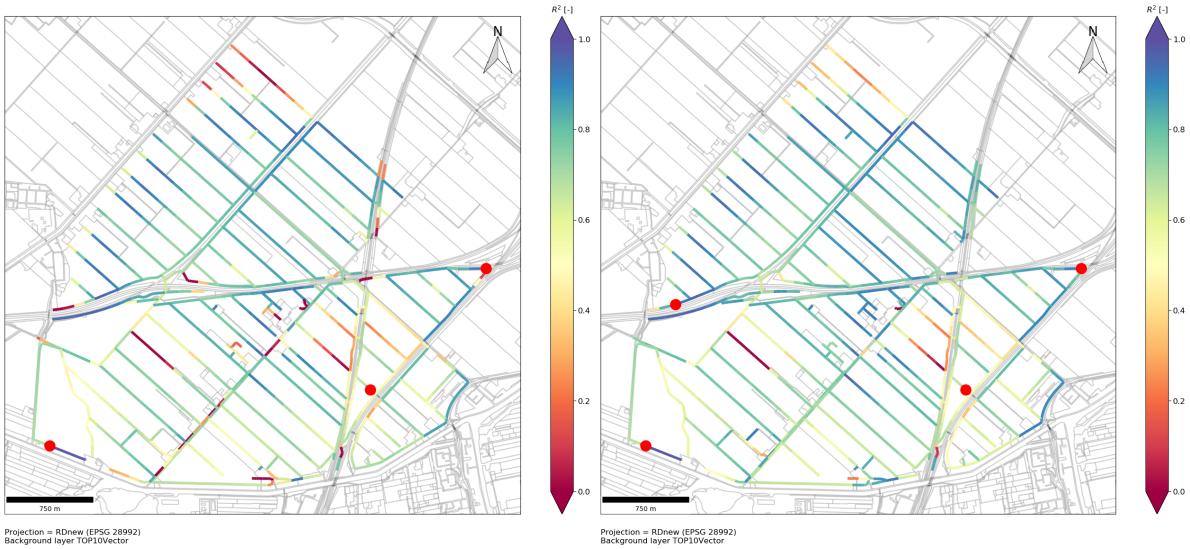



Figure 4.5: The performance of the monitoring network when additional sensors are chosen by the greedy algorithm

The optimal layouts of sensors found with the greedy algorithm and their performance for 1 to 10 sensors are presented in Table 4.2. Figure 4.6a and Figure 4.6b show the performance of two results spatially. Figure 4.5 shows that the performance of the monitoring network continues to improve when more sensors are added but the improvement is biggest in the first 6 sensors. When adding an extra sensor to the monitoring network also an extra principal component is added to the model, in this way it stays a determined problem. Another option is to stop adding extra principal components and solve an overdetermined problem when more sensors than components are present. The first option has been chosen to get the results in Figure 4.5. With this method as much as possible variance is taken into account, when only three principal components are taken into account 9% of variance will always be neglected.

A great advantage of the greedy algorithm is that it is able to calculate the sensor placement for more than three locations. With an exhaustive algorithm this is not possible because there are too many combinations. The greedy algorithm is an efficient way to solve this sensor placement problem because the problem is ordered, the variance decreases when adding more principal components so in this case it is important to choose the local optimum from the beginning. With expert opinion it is hard to place more than three sensors because the additional locations have to be guessed. The greedy algorithm confirms that the three categories of key locations to monitor are correct. In the solution of the greedy algorithm sensors are placed in at least two different categories. However, the third category also has a high performance in the greedy algorithm, but because another solution are marginal better this category is not always chosen.




(a) Sensor layout: 67 - 173 - 496

(b) Sensor layout: 67 - 173 - 496 - 381

Figure 4.6: Shows performance for the top two layouts from the greedy algorithm

### 4.2.4. Summary results of optimisation strategies

In Table 4.3 the results of the three optimisation strategies are summarised. This table shows that while the sensor placement based on expert opinion performs reasonable but the sensor placement with the exhaustive or greedy algorithm performs much better. The greedy algorithm gives almost the same results as the exhaustive algorithm, although the number of calculations is much smaller and the greedy algorithm is the only algorithm which gives insight in the extra information gained when installing more than three sensors.

Table 4.3: Results of sensor placement methods

| | **Calculations** | **#1** | $R^2$ | **#2** | $R^2$ | **#3** | $R^2$ |
|---|---|---|---|---|---|---|---|
| **Expert opinion** | 1 | 89-661-744 | 0.581 | 240-661-744 | 0.534 | - | - |
| **exhaustive** | 562475 | 271 - 666 - 517 | 0.659 | 356 - 661 - 691 | 0.659 | 356 - 666 - 517 | 0.658 |
| **Greedy algorithm** | 7505 | 67 - 173 -496 | 0.655 | 67 - 173 -496 - 381 | 0.717 | 67 - 173 -496 - 381 - 257 - 367 | 0.76 |

### 4.2.5. Robustness of sensor placement

The robustness of the layouts of sensor networks is tested by calculating the performance with different scenarios from the SOBEK simulations. The following scenarios from the sensitivity analysis (see Section 3.2) have been used:

- Reference scenario
- Scenario where the boil flux is doubled
- Scenario where the boil flux is reduced
- Scenario where the individual boil fluxes are changed, but the total seepage stays the same
- Scenario where the flushing discharge is increased
- Scenario where the flushing discharge is decreased
- Scenario where 4 boil locations are changed
- Updated boil locations based on measurements

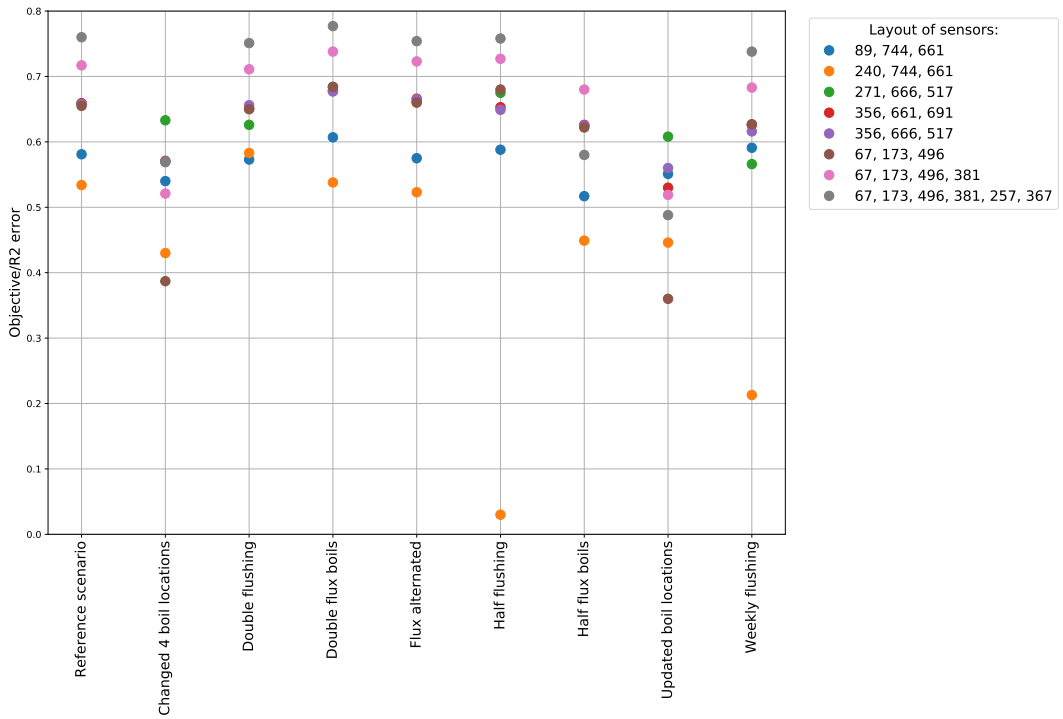
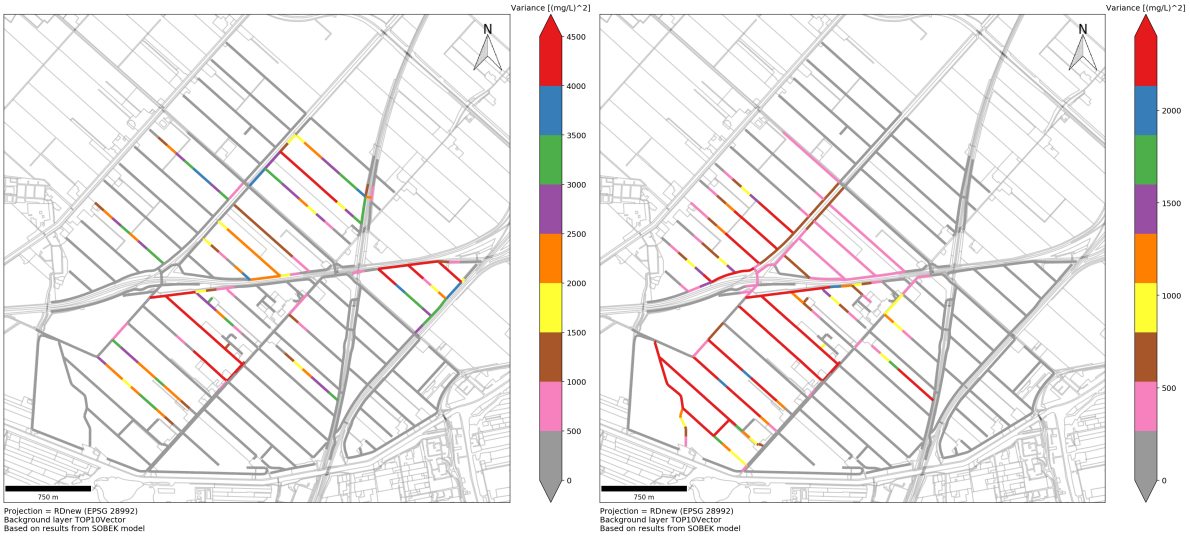



Figure 4.7: Robustness of different sensor layouts visually

Figure 4.7 shows the results of the robustness analysis, the values behind the plot can be found in Table B.1 in Appendix B.2. This shows that all layouts of sensors perform worse at the scenarios where the locations of the boils are changed. When the locations of the boils change the variance is also distributed differently over the catchment. If a sensor is placed at a location which has no variance the performance of the sensor network will decrease because some of the variance will not be measured. However, when a monitoring network exist of more than three sensors it is more robust to a change of the boil locations, in this case another sensor is able to measure the variance. Figure 4.8a and Figure 4.8b show the change in variance when the boil locations change. In the case of the updated model (Figure 4.8b) this leads to a decrease in variance in the north-eastern part of the catchment and an increase in variance in the western part of the catchment. Figure 4.7 also shows that the sensor layout with sensors at the locations: 240, 744 and 661 performs worse than the other solutions and is less robust, this layout is chosen with expert opinion. This layout especially perform worse when the amount of flushing is changed. This is caused by the location 240 which is located near one of the inlets and will capture less information when flushing has a smaller influence on the salinity in the area.




(a) (b)

Figure 4.8: Comparison of the variance between the reference scenario (a) and the scenario with the updated boil locations (b)

The optimisation strategies have provided a set of sensor layouts which perform good at the reference scenario. Th robustness analysis has showed that most of these sensor placements are also robust when processes will be different in reality. Moreover, it shows that the monitoring networks from the greedy algorithm with more than three sensors are more robust, the overall performance of these networks is higher. The conclusion of this chapter is that a sensor network has to have a minimum of three sensors but to improve the performance and robustness of the network 1 to 3 sensors have to be added using the greedy algorithm. From the set of optimal sensor locations the most optimal solution has to be chosen in further research based on the model predictive control scheme for optimal flushing, the business case and field tests.

## 4.3. Implementation

In this section the business case of a salinity monitoring network in the Lissertocht catchment is discussed, only a rough outline of the business case is given because this topic is outside the scope of this thesis. Finally, some important steps before setting up a salinity monitoring network are summarised in a guideline.

### 4.3.1. Business case

Before setting up the monitoring network it is important to investigate if the (financial) benefits of the monitoring network weigh out the (financial) costs. The costs of the monitoring network exist of:

- EC sensors which measure the electrical conductivity and sends the data to the waterboard
- Costs to install the EC sensor into a housing in the channel
- Maintenance of the equipment
- ICT infrastructure to manage the data from the sensors and send commands to structures for flushing
- Automatic inlets, currently there are 2 automatic inlets in the Lissertocht

The fresh water from the Ringvaart which is used for flushing has to be pumped out of the polder again. To pump the water back to the Ringvaart the pumps need electricity. In the SOBEK model roughly 1.5 million $m^3$ fresh water is used per year for flushing. If losses of water are neglected this means that also 1.5 million $m^3$ has to be pumped by the pump J.P. Heye on the Hoofdvaart, from here the pump Leeghwater has to pump the water on the Ringvaart. The waterboard Rijnland has provided an estimate of the electricity costs, these are 600 euros per million $m^3$ for the pump J.P. Heye and 6.000 euros per million $m^3$ for the pump Leeghwater. This makes the total yearly electricity cost for flushing equal to: $1.5 * (600 + 6.000) = 9.900$ euros.

Without knowing the exact amount of the costs that can be saved by a salinity monitoring network combined with a model predictive control scheme it is clear that the financial benefits will be marginal.

However, if other social-economic aspects are taken into account there can be a business case, some of these factors are:

- Sustainability: it is not sustainable to use energy to pump water out of a catchment if it is not needed. Use less electricity is better for the environment, less fossil fuels are used.
- Water scarcity: during dry summers the waterboard Rijnland is dependent on surrounding waterboards for the supply of fresh water. If the amount of fresh water is limited it is very important to allocate the fresh water efficiently. This is only possible with a robust monitoring network and efficient controllers of the inlets and pumps.
- Prevent damage to crops of the farmers.
- Upscaling: There are multiple polders which face the same problem, such as the Noordplas and Shermerpolder. Moreover, the whole Haarlemmermeerpolder uses 30 to 100 million $m^3$ during a dry year (Alterra, 2005). If this system is scaled up the business case can be more positive and maybe the electricity costs in other polders are higher and/or the potential benefits higher. When it is applied to multiple catchments the costs of installing the sensors, the ICT infrastructure and additional research can be divided over all these polders.

In the Lissertocht there are already two EC sensors installed and two inlets are automated. However, these systems are not all communicating with each other and there are sometimes some practical problems, such as inlets that are blocked, the two sensors are of a different brand and use different software and one of the sensors has randomly periods of several days to a week where there is no data or unrealistic data measured.

### 4.3.2. Guidelines for setting up a monitoring network

This section summarises the most important steps from this thesis to design a salinity monitoring network for optimal flushing. These steps are based on the detection and measurements of boils in Chapter 2, the simulation in SOBEK, results from the PCA and predictive model in Chapter 3 and the optimisation of sensor placement in Chapter 4. The guideline (shown below) gives an idea where to start and which lessons can be learned from the case study of the Lissertocht catchment. This guideline has not been tested yet in another polder, this is a topic for further research. To set up a salinity monitoring network in another catchment the following steps can be taken:

1. Locate the sources of saline water in the catchment, in this thesis the tools EC routing, visual inspection and DTS measurements are presented to detect and measure boils. Moreover, DTS is able to characterize the boil flux on a qualitative level and also on a quantitative level for a boil that is located near a pump; both have been done in this research (see Section 2.3 and Section 2.5).
2. Set up a SOBEK model of the catchment to model the transport of water and salt in the catchment. The simulations in SOBEK will show the temporal and spatial distribution of the salinity and the processes that influence the salinity.
3. The modelled salt concentrations can be used to perform a principal component analysis (PCA) to analyse the variance and correlation in the catchment. The PCA will show if most of the variance in the system can be captured with a limited amount of linear components, this gives an indication of the minimal amount of sensors which are necessary.
4. From the results of the measurements, SOBEK simulations and PCA categories of key locations to monitor can be identified. In case of the case study of the Lissertocht the categories are:
   - Near the outlet(s) of the catchment
   - At the path the flushing water takes
   - Near one of the boils in a stagnant channel

   As a first guess a sensor layout which consists of at least one sensor from each category can be used. This layout of sensors can be tested with the predictive model which is based on the PCA. These categories are identified based on the case study of the Lissertocht, in other polders other processes can be dominant which can results in other categories of key locations to monitor.
5. Use a greedy algorithm to find a set of optimal sensor locations. This algorithm will give insight how much the monitoring network improves when additional sensors are placed.
6. Test the robustness against disturbances of the set of optimal layouts of sensors.
7. Before implementing the salinity monitoring network field tests and further research about the business case and combining it with a model predictive control scheme for optimal flushing is needed. A field test can be performed by installing EC sensors at the chosen locations for the sensors in the monitoring network and place some EC sensors at other locations for validation. The sensors for validation should measure approximately the same salinity levels and patterns as predicted by the predictive model based on the measurements of the first EC sensors.

The first six steps are taken in this research for the Lissertocht, for step seven only a outline is given. Step seven contains topics for further research, see Section 5.2.

# 5

# Conclusions and recommendations

In this chapter the conclusions will be drawn, the research questions will be answered and the objective of this thesis will be evaluated. Finally, recommendations and topics for further research will be presented. (Kelderman, 2015)

## 5.1. Conclusions

In this section the conclusions are drawn and the research questions from Chapter 1 (shown below) are addressed.

1. How can the boils in this catchment be detected and measured?
2. What is the spatial and temporal distribution of the salt concentrations in this polder and which processes influence these patterns?
3. Is it possible to identify key locations for monitoring the salt concentrations based on modelling and measuring results?
4. Is it possible to optimise sensor placement for a monitoring system in the Lissertocht catchment?
5. Can the results of this research be used for the design of a salinity monitoring network in another catchment?

**1. How can the boils in this catchment be detected and measured?**
The EC routing, DTS measurements and visual inspection are tools to detect and measure boils. The EC routing is a good method for detecting boils in big areas but is sensitive to weather conditions and it is hard to locate the exact source of the saline water. Areas which have high electrical conductivity values in the EC routing can contain boils. To confirm the results of EC routing and find the exact location of the boils a second method is needed, like DTS or visual inspection.

The DTS method is the only method which can detect boils which are located at the bottom of the bigger main channels. Moreover, DTS is able to characterize the boil flux on a qualitative level and also on a quantitative level for a boil that is located near a pump; both have been done in this research. The boil flux is calculated near the pump J.P. Heye with an energy balance and an 1D advection-diffusion model; the results are in the same range and correspond to values found in literature but the uncertainty of these calculations is high. This was due missing information about the distribution of the temperature and salinity over depth. At the first measurement location it was not possible to quantify the boil fluxes because it is difficult to isolate the influence of the boil on the temperature in the channel when the temperature does not reset due to a pump. However, it was possible to compare the fluxes of the boils in the two drainage channels at the first

measurement locations, which are measured simultaneously, on a qualitative level. This comparison shows that there is variation between the boil fluxes of different boils. The new information about the boil locations and their fluxes is used in the simulations in SOBEK.

**2. What is the spatial and temporal distribution of the salt concentrations in this polder and which processes influence these patterns?**
The spatial and temporal distribution of the salinity is measured and modelled. The CTD divers show the relationship between the salinity at several locations and the effect of pumping, precipitation and a dryer period. Simulations in SOBEK show which parts of the catchment face salinity problems, besides, it shows where the salt concentrations can be decreased by opening the inlets. The sensitivity analysis provides more insight in the influence of the amount of flushing, the flushing scheme, boil locations and boil flux on the spatial and temporal distribution of the salinity. The variance and correlation are important parameters for the sensor placement problem. The goal of a sensor network for monitoring is to measure as much as variance as possible while placing sensors at locations which have a high correlation with other locations in the catchment. These two concepts are taken into account in the principal component analysis, which shows that more than 90% of the variance of the system can be represented with three orthogonal linear components.

**3. Is it possible to identify key locations for monitoring the salt concentrations based on modelling and measuring results?**
From the classification with PCA, measurement results and simulation in SOBEK three categories are formulated, for the sensor placement it is important to have at least one sensor in each categories:

- Near the outlet of the catchment, simulations in SOBEK and measurements with CTD divers show that the variance is big at this location and that changes in the polder have impact on the salt concentrations at this location.
- In one of the main channels between an inlet and the outlet of the catchment. In this channels the effect of flushing can be identified if this channel is saline without flushing. It is important to measure the effects of flushing because the purpose of the monitoring network is to optimise the control of flushing.
- In a drainage channel with a boil. These channels shows the biggest variance over the year, during the summer the salinity stays high at these locations because the flushing is not effective. During the winter the channels are flushed naturally by the fresh precipitation which flows into these channels via the drains.

In addition a predictive model is built using the system of linear equations from the PCA. This predictive model is capable of testing the performance of the monitoring network virtually. This gives a good indication on how much of the variance can be captured with a certain sensor layout and if this information can be used to predict salinity values at locations to be controlled with flushing.

**4. Is it possible to optimise sensor placement for a monitoring system in the Lissertocht catchment?**
The principal component analysis and the predictive model shows that at least three sensors are necessary to reconstruct the three principal components and capture most of the variance in the system. With the predictive model as a tool to test the performance of a big amount of sensors layouts it is possible to solve the sensor placement problem with optimisation strategies, the following has been used:

- Expert opinion
- Exhaustive algorithm
- Greedy algorithm

The expert opinion is able to find a layout of a monitoring network which performs reasonable when one sensor out of each category is chosen, however, the amount of combinations is too big to choose the optimal sensor placement with reasoning only. The exhaustive algorithm tests all possible combinations without

taking into account the categories which are identified, this prevents a bias. The results validate the three formulated categories of key locations to monitor. However, the amount of possible combinations is big and will increase even more when more sensors are added. The greedy algorithm is able to deal with extra sensors in an efficient way. The results of this algorithm show that the quality of the sensor network can even be improved when extra sensors are added but after the sixth sensor the additional gain of an additional sensor is small. The robustness analysis of the found sensor networks shows that the sensor placement is most sensitive to the locations of the boils. Moreover, it show that the monitoring networks with more than three sensors are more robust against disturbances.

**5. Can the results of this research be used for the design of a salinity monitoring network in another catchment?**
The results of the measurements, modelling and optimisation of sensor placement in the Lissertocht provides a guideline for designing a salinity monitoring network for optimal control in other polders.

This thesis has showed that the boil locations are important to design a robust salinity monitoring network, the boil locations will determine the variance and correlations in the catchment. Therefore, the first step should be to detect and measure the boils in a catchment, this thesis has presented several tools for this purpose and discussed their limitations and benefits. The biggest advantage of DTS measurements over visual inspection and EC routing is the possibility to characterize the boil fluxes on a qualitative level and under certain circumstances it is even possible to quantify the boil fluxes.

A second step is to model the transport of water and salt in the catchment with a SOBEK model. The simulations in SOBEK will show the temporal and spatial distribution of the salinity and the processes that influence the salinity. The modelled salt concentrations can be used to reduce the complexity of the system with the PCA. The PCA will show if most of the variance in the system can be captured with a limited amount of linear components, this gives an indication of the minimal amount of sensors which are necessary. From the results of the measurements, SOBEK simulations and PCA categories of key locations to monitor can be identified.

Finally, different optimisation strategies can be used to solve the sensor placement problem. The greedy algorithm is the most efficient way to find out how much sensors are needed. This algorithm will give insight in the additional gain of an extra sensors. A predictive model based on the linear components from the PCA can be used to test the performance of the different sensor layouts.

To implement the salinity monitoring network further research is needed to field tests, the business case and combining the monitoring network with a model predictive control for optimal flushing.

## Objective of the research

The main objective of this thesis is: *Measuring and modelling of the salt concentrations and their sources in order to optimise the placement of sensors in a salinity monitoring network for optimal flushing in the Lissertocht catchment.*

The conclusions show that the salt concentrations and their sources (the boils) are measured with multiple methods. This made it possible to determine the boil locations more accurately, calculate the boil flux and identify processes which have impact on the salt concentrations. This information is used in the modelling part of this thesis where a sensitivity analysis is performed with SOBEK. Moreover, a PCA is conducted which shows the variance in the catchment and how channels are correlated. This can be used to classify different types of channels. Based on the PCA a predictive model is built which allows the user to test the performance of a sensor placement for a salinity monitoring network. Finally, the sensor placement is optimised using expert opinion, an exhaustive algorithm and a greedy algorithm as optimisation strategies. This has resulted in a set of sensor layout which can be used for optimal control of flushing. Because of this, the objective of this research has been achieved.

## 5.2. Recommendations and further research

This thesis has provided new insights into methods to detect and measure boils and calculate their fluxes, the spatial and temporal distribution of salinity, reduce complexity of a system with PCA and solving a sensor placement problem can be used to implement a monitoring network in the Lissertocht and for other deep polders in the Netherlands. Some recommendations for further research are:

- Perform field tests of the monitoring network. Install minimal three sensors which function as a monitoring network and install some more sensors to validate the predictions based on the measurements from the monitoring network. The results of the predictions based on real measurements can be compared with the predictions that are obtained with the virtual measurements. This can be used for validation of the predictive model.
- Combine the monitoring network with the model predictive control scheme for optimal flushing.
- The optimisation of the sensor placement problem gives an idea where to place sensors and which combinations of different locations perform well based on the models. However, objective of the optimisation in this thesis is very general, namely: *Maximise the mean $R^2$ error over all the main channels.* If information about types of crops, fresh water demand and irrigation schemes in combination with the requirements for the implementation of a model predictive control scheme will be taken into account the objective can be formulated more specifically. With a more specific objective it is also possible to test different measurements of fit such as the Nash-Sutcliffe efficiency, Bias or Mean squared error.
- Install DTS cables on different depths in combination with CTD divers at different depths at the second measurement location during mid-winter or mid-summer. Mid-summer is preferred because the seepage water is colder than the surface water. With these measurements the salt and heat transport can be modelled in 3D. This can result in more knowledge about the stratification and an accurate calculation of the boil flux.
- Investigate if a salinity monitoring network can be designed for another deep polder by following the steps in the guideline in Section 4.3.2. If the monitoring network will be upscaled to multiple polders the benefits can even be bigger and the allocation of fresh water for flushing within a waterboard can be optimal.

# A

# Fieldwork and dataset

## A.1. Set-up DTS

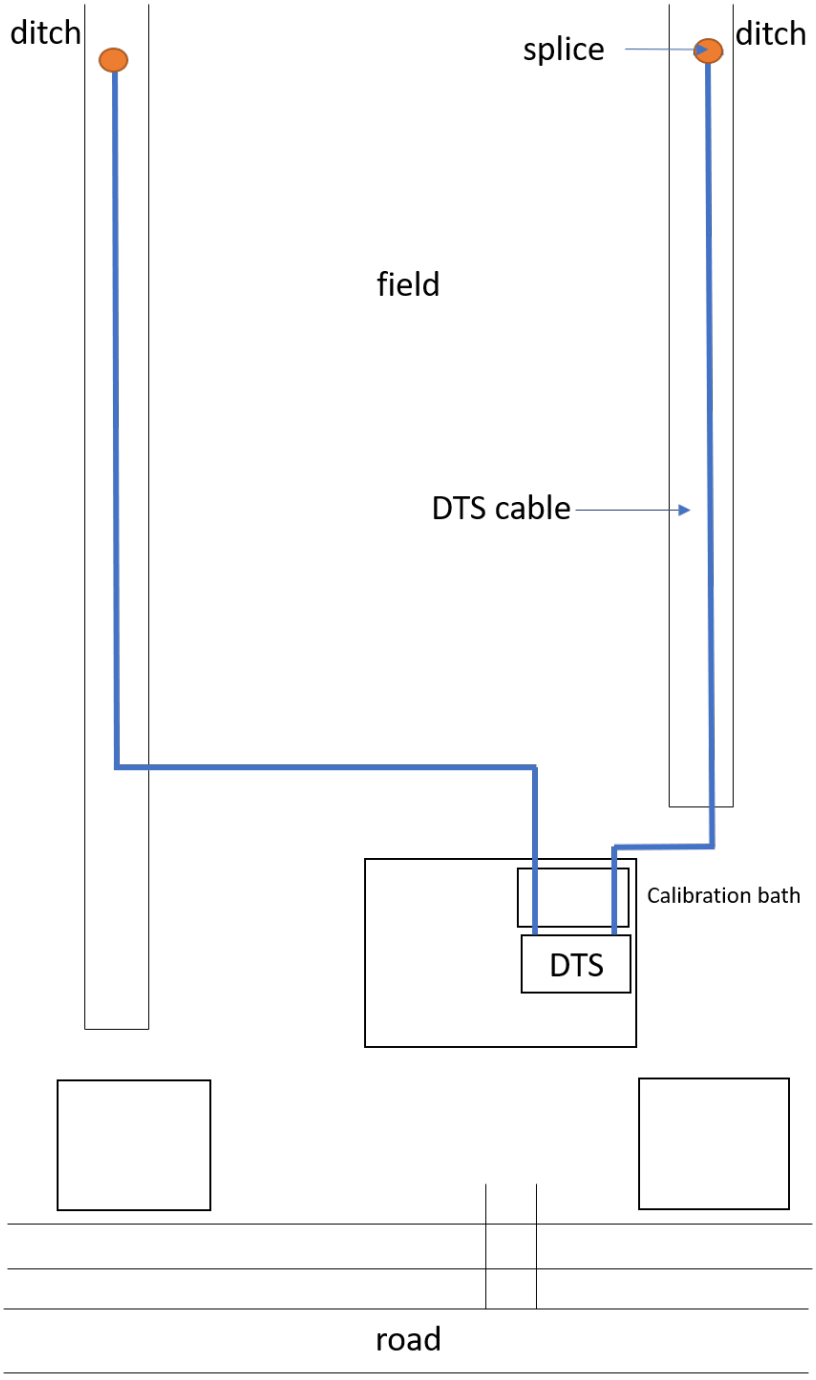



Figure A.1: Set-up at first DTS measurement location, this location is indicated in Figure 2.5 in Section 2.3

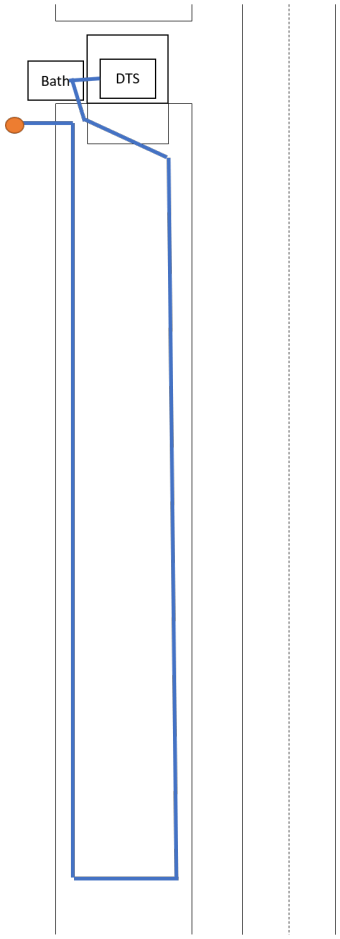
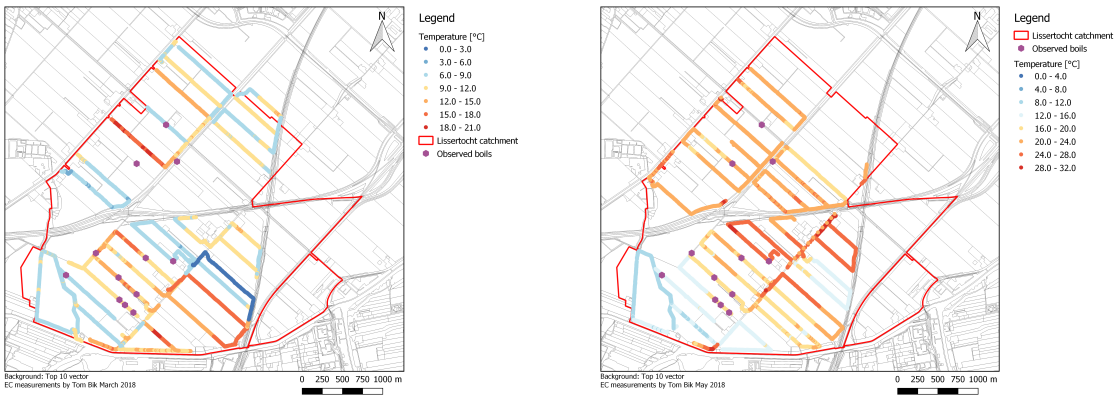



Figure A.2: Set-up at second DTS measurement location, this location is indicated in Figure 2.5 in Section 2.3

## A.2. Temperature measurements EC routing

Temperature measured by the CTD diver during EC routing. It is hard to identify the boils based on these images.




(a) (b)

Figure A.3: Temperature measured during EC routing in may is presented in (a) and the temperature in May is presented in (b)

For the complete dataset contact the author of this thesis Hugo Hagedooren.

# B

# Modelling results

## B.1. Scenarios sensitivity analysis

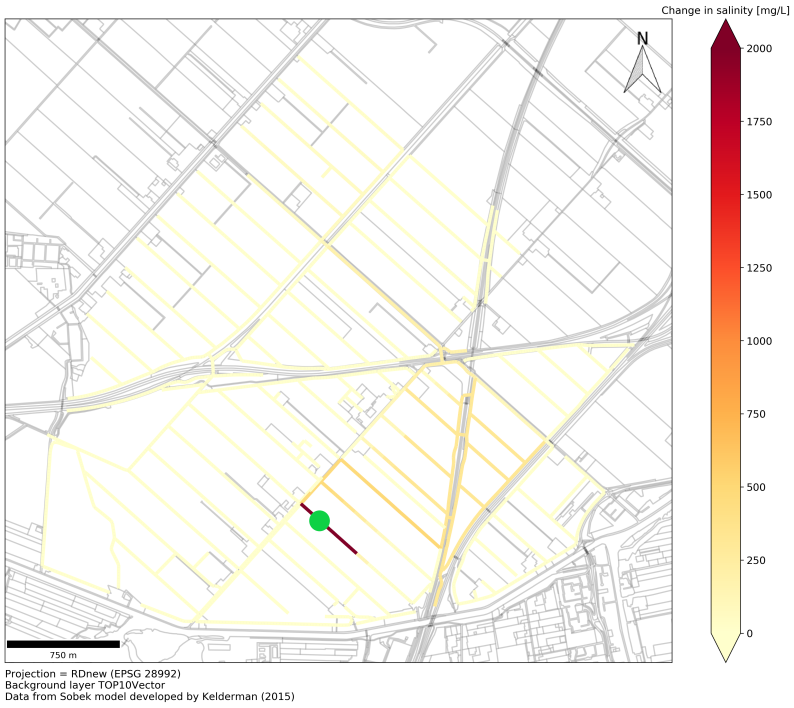



Figure B.1: Effect of adding one extra boil (indicated with green circle) in a drainage channel

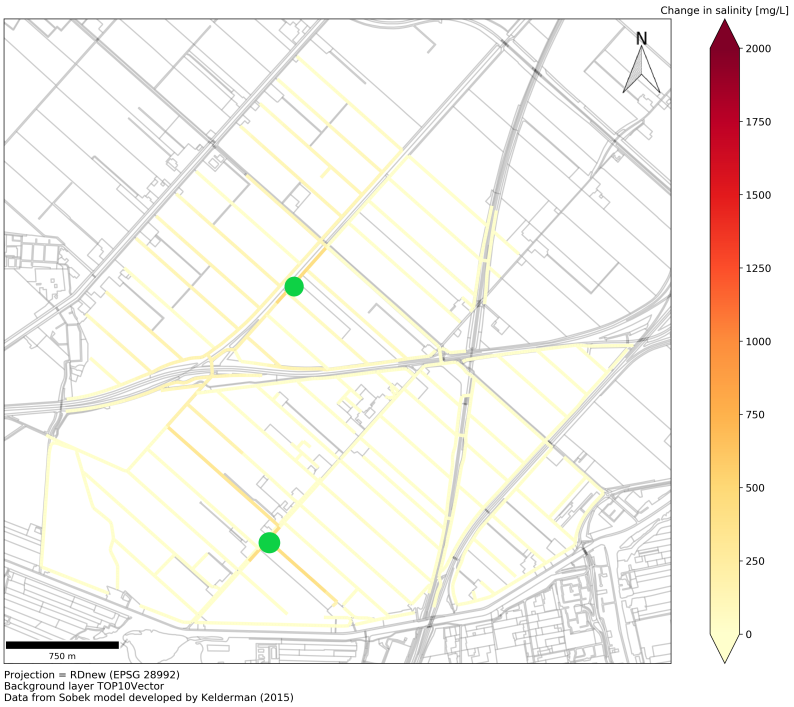



Figure B.2: Effect of adding two extra boils (indicated with green circle) in a main channel

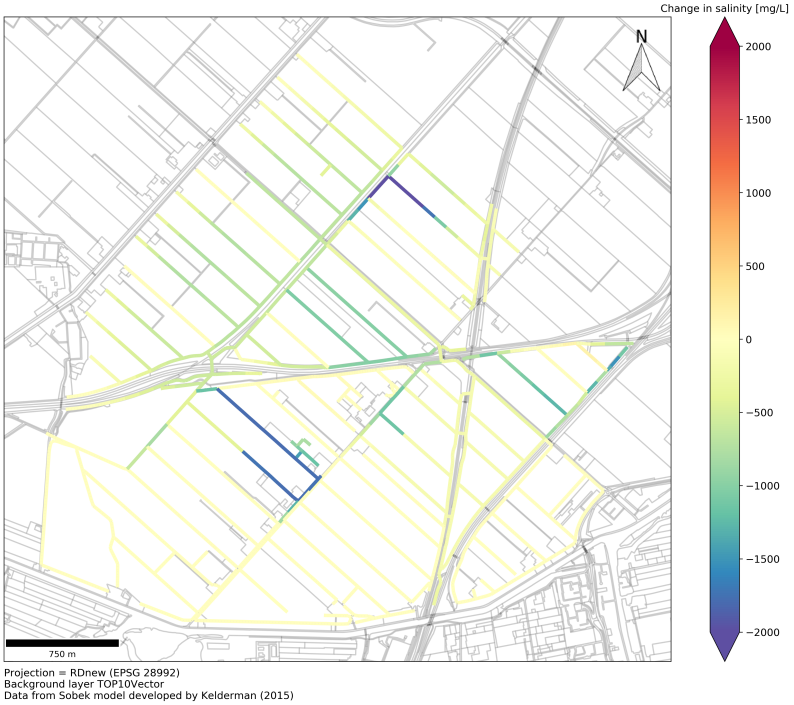



Figure B.3: Effect of decreasing the boil flux with 50% on the spatial distribution of the salt concentrations during the summer

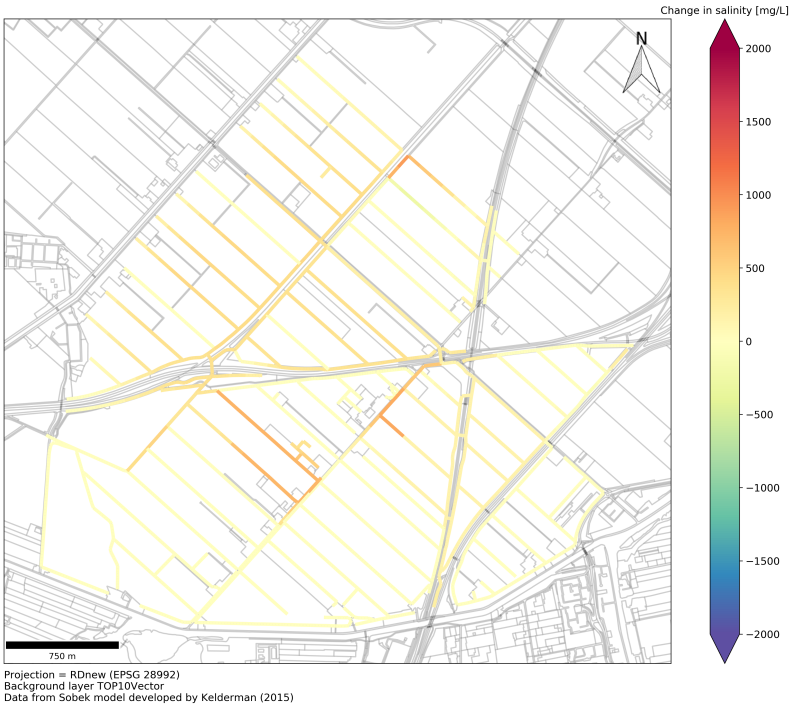



Figure B.4: Effect of changing the boil flux of all boils from $0.002 m^3/s$ to $0.0027 m^3/s$ on the spatial distribution of the salt concentrations during the summer

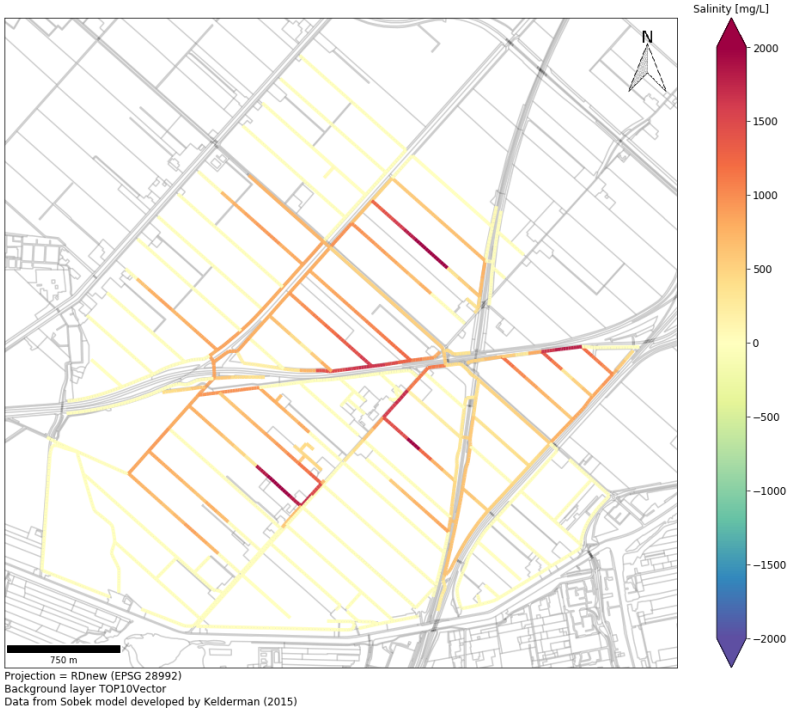



Figure B.5: Effect of increasing the boil flux by 100% on the spatial distribution of the salt concentrations during the summer

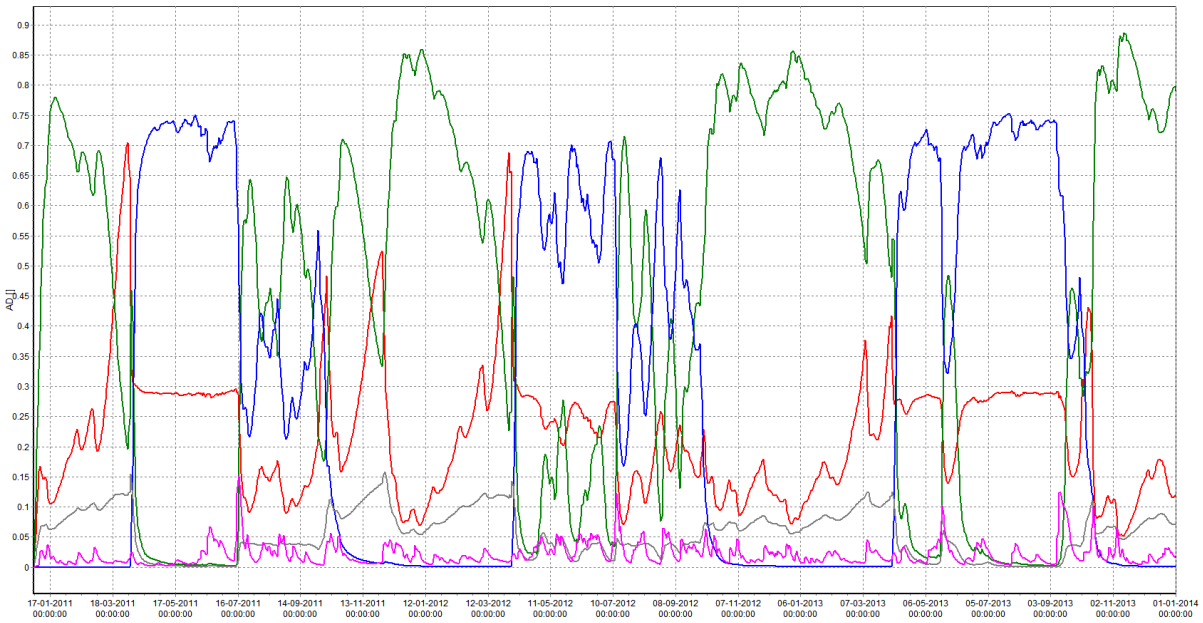



Figure B.6: Overview of all fractions at the outlet (pump J.P. Heye) over the entire simulation period (2011 - 2014). Red line is fraction seepage water from boils, green line is fraction of drainage water (shallow freatic groundwater exfiltration through tile drains), blue line is fraction of inlet water (flushing), gray line is fraction of diffusive seepage under ditches, purple line is fraction of the precipitation

## B.2. Optimisation

Table B.1 shows the results of the robustness analysis, this data is plotted in Figure 4.7.

Table B.1: Robustness of different sensor layouts tested by calculating the mean $R^2$ error

| | **89, 744, 661** | **240, 744, 661** | **271, 666, 517** | **356, 661, 691** | **356, 666, 517** | **67, 173, 496** | **67, 173, 496, 381** | **67, 173, 496, 381, 257, 367** |
|---|---|---|---|---|---|---|---|---|
| **Reference scenario** | 0.581 | 0.534 | 0.659 | 0.659 | 0.658 | 0.655 | 0.717 | 0.76 |
| **Changed 4 boil locations** | 0.54 | 0.43 | 0.633 | 0.571 | 0.569 | 0.387 | 0.521 | 0.57 |
| **Double flushing** | 0.573 | 0.583 | 0.626 | 0.651 | 0.656 | 0.65 | 0.711 | 0.751 |
| **Double flux boils** | 0.607 | 0.538 | 0.684 | 0.678 | 0.677 | 0.684 | 0.738 | 0.777 |
| **Flux alternated** | 0.575 | 0.523 | 0.661 | 0.666 | 0.665 | 0.66 | 0.723 | 0.754 |
| **Half flushing** | 0.588 | 0.03 | 0.675 | 0.653 | 0.649 | 0.68 | 0.727 | 0.758 |
| **Half flux boils** | 0.517 | 0.449 | 0.624 | 0.626 | 0.625 | 0.622 | 0.68 | 0.58 |
| **Updated boil locations** | 0.551 | 0.446 | 0.608 | 0.53 | 0.56 | 0.36 | 0.519 | 0.488 |
| **Weekly flushing** | 0.591 | 0.213 | 0.566 | 0.626 | 0.616 | 0.627 | 0.683 | 0.738 |

# List of Figures

# List of Tables

# Bibliography


Vénice Akker and Chris Van Naarden. Verzilting in West-Nederland. Bachelor honours thesis, Vrije Universiteit Amsterdam, 2013.

Leonardo Alfonso, Arnold Lobbrecht, and Roland Price. Optimization of water level monitoring network in polder systems using information theory. *Water Resources Research*, 46(1):1–13, 2010a. ISSN 00431397. doi: 10.1029/2009WR008953.

Leonardo Alfonso, Arnold Lobbrecht, and Roland Price. Information theory-based approach for location of monitoring water level gauges in polders. *Water Resources Research*, 46(10), 2010b. ISSN 00431397. doi: 10.1029/2009WR008101.

Alterra. Onderbouwing wateropgave Haarlemmermeerpolder. Technical report, Alterra, Wageningen, 2005.

Alterra. LGN 7 database, 2015.

Analytics Vidhya. Practical Guide to Principal Component Analysis (PCA) in R & Python. URL https://www.analyticsvidhya.com/blog/2016/03/practical-guide-principal-component-analysis-python/. Accessed 2018-06-16.

Boran Ekin Aydin, Martine Rutten, Edo Abraham, G. H. P. Oude Essink, and J.R. Delsman. Model Predictive Control of Salinity in a Polder Ditch Under High Saline Groundwater Exfiltration Conditions : A Test Case. pages 3215–3219, 2015.

Boran Ekin Aydin, Martine Rutten, Gualbert H.P.Oude Essink, and J.R. Delsman. Polder Flushing: Model Predictive Control of Flushing Operations to Effective and Real Time Control of Salinity in Polders. *Procedia Engineering*, 154:94–98, 2016. ISSN 18777058. doi: 10.1016/j.proeng.2016.07.424. URL http://dx.doi.org/10.1016/j.proeng.2016.07.424.

B. K. Banik, Jose Leonardo Alfonso Segura, A. S. Torres, A. Mynett, C. Di Cristo, and A. Leopardi. Optimal placement of water quality monitoring stations in sewer systems: An information theory approach. *Procedia Engineering*, 119(1):1308–1317, 2015. ISSN 18777058. doi: 10.1016/j.proeng.2015.08.956. URL http://dx.doi.org/10.1016/j.proeng.2015.08.956.

Paul M. Barlow and Eric G. Reichard. Saltwater intrusion in coastal regions of North America. *Hydrogeology Journal*, 18(1):247–260, 2010. ISSN 14312174. doi: 10.1007/s10040-009-0514-3.

A. GHOSH BOBBA. Numerical modelling of salt-water intrusion due to human activities and sea-level change in the Godavari Delta, India. *Hydrological Sciences Journal*, 47(sup1):S67–S80, 2002. ISSN 0262-6667. doi: 10.1080/02626660209493023. URL http://www.tandfonline.com/doi/abs/10.1080/02626660209493023.

Martin A. Briggs, Laura K. Lautz, and Jeffrey M. Mckenzie. A comparison of fibre-optic distributed temperature sensing to traditional methods of evaluating groundwater inflow to streams. *Hydrological Processes*, 26(9):1277–1290, 2012. ISSN 08856087. doi: 10.1002/hyp.8200.

E. Custodio and G. A. Bruggeman. Groundwater problems in coastal areas. *IHP Working Group*, 53(9):1689–1699, 2013. ISSN 1098-6596. doi: 10.1017/CBO9781107415324.004.

P.G.B. De Louw. *Saline seepage in deltaic areas.* PhD thesis, Vrije Universiteit Amsterdam, 2013.

P.G.B. De Louw, R. Bakkum, H. Folkerts, and H. Van Hardeveld. Het effect van waterbeheer op de chloride- en nutrientenbelasting van het oppervlaktewater in Polder de Noordplas. Technical report, TNO, 2004.

P.G.B. De Louw, Gualbert H.P.Oude Essink, Piet Maljaars, Hoogheemraadschap Van Rijnland, and Bennie Minnema. Achtergrondstudie kwelreductietechnieken. 2007.

P.G.B. De Louw, G. H.P. Oude Essink, P. J. Stuyfzand, and S. E.A.T.M. van der Zee. Upward groundwater flow in boils as the dominant mechanism of salinization in deep polders, The Netherlands. *Journal of Hydrology*, 394(3-4):494–506, 2010. ISSN 00221694. doi: 10.1016/j.jhydrol.2010.10.009. URL http://dx.doi.org/10.1016/j.jhydrol.2010.10.009.

P.G.B. De Louw, P.J. Doornenbal, and D.M.D. Hendriks. Veldonderzoek naar het dichten van wellen. *Deltares rapport 1201949*, page 56, 2012.

P.G.B. De Louw, A. Vandenbohede, A. D. Werner, and G. H. P. Oude Essink. Natural saltwater upconing by preferential groundwater discharge through boils. *Journal of Hydrology*, 490:74–87, 2013. ISSN 00221694. doi: 10.1016/j.jhydrol.2013.03.025. URL http://dx.doi.org/10.1016/j.jhydrol.2013.03.025.

L del Val Alonso. Geochemical interpretation of the relation between ground and surface water systems at small subcatchment in the Haarlemmermeer polder. Technical Report July, VU university, Amsterdam, 2011.

J.R. Delsman, G. H. P. Oude Essink, Keith J. Beven, and Pieter J. Stuyfzand. Uncertainty estimation of end-member mixing using generalized likelihood uncertainty estimation (GLUE), applied in a lowland catchment. *Water Resources Research*, 49(8):4792–4806, 2013. ISSN 00431397. doi: 10.1002/wrcr.20341.

J.R. Delsman, K. R M Hu-A-Ng, P. C. Vos, P.G.B. De Louw, G. H. P. Oude Essink, P. J. Stuyfzand, and M. F P Bierkens. Paleo-modeling of coastal saltwater intrusion during the Holocene: An application to the Netherlands. *Hydrology and Earth System Sciences*, 18(10):3891–3905, 2014. ISSN 16077938. doi: 10.5194/hess-18-3891-2014.

J.R. Delsman, P.G.B. De Louw, Willem J. de Lange, and Ou. Fast calculation of groundwater exfiltration salinity in a lowland catchment using a lumped celerity/velocity approach. *Environmental Modelling and Software*, 96:323–334, 2017. ISSN 13648152. doi: 10.1016/j.envsoft.2017.07.004. URL http://dx.doi.org/10.1016/j.envsoft.2017.07.004.

Deltares. Conversion EC to Chloride, 2015. URL https://publicwiki.deltares.nl/display/ZOETZOUT/Conversion+EC+to+Chloride. Accessed 2018-06-21.

Deltares. *SOBEK user manual*. 2016.

D. Dorini, G., Jonkergouw, P., Kapelan, Z., di Pierro, F., Khu, S., and Savic. An Efficient Algorithm for Sensor Placement in Water Distribution Systems. *Water Distribution Systems Analysis Symposium*, 40941(February 2016):1–13, 2006. doi: 10.1061/40941(247)101.

D. G. Eliades, M. Kyriakou, and M. M. Polycarpou. Sensor placement in water distribution systems using the S-PLACE Toolkit. *Procedia Engineering*, 70(2010):602–611, 2014. ISSN 18777058. doi: 10.1016/j.proeng.2014.02.066. URL http://dx.doi.org/10.1016/j.proeng.2014.02.066.

Gemeente Haarlemmermeer. History | HLMRMeer. URL https://hlmrmeer.nl/nl/node/7637. Accessed 2018-05-26.

Beatrice M S Giambastiani, Marco Antonellini, Gualbert H.P.Oude Essink, and Roelof J. Stuurman. Saltwater intrusion in the unconfined coastal aquifer of Ravenna (Italy): A numerical model. *Journal of Hydrology*, 340(1-2):91–104, 2007. ISSN 00221694. doi: 10.1016/j.jhydrol.2007.04.001.

Ruben Goudriaan, P.G.B. De Louw, and Mark Kramer. Lokaliseren van zoute wellen in de Haarlemmermeerpolder. *H20*, 03:29–32, 2011.

William E. Hart and Regan Murray. Review of Sensor Placement Strategies for Contamination Warning Systems in Drinking Water Distribution Systems. *Journal of Water Resources Planning and Management*, 136 (6):611, 2010. ISSN 07339496. doi: 10.1061/(ASCE)WR.1943-5452.0000081. URL http://link.aip.org/link/JWRMD5/v136/i6/p611/s1{&}Agg=doi.

Mark B. Hausner, Francisco Suárez, Kenneth E. Glander, Nick van de Giesen, J. Selker, and Scott W. Tyler. Calibrating single-ended fiber-optic raman spectra distributed temperature sensing data. *Sensors*, 11(11): 10859–10879, 2011. ISSN 14248220. doi: 10.3390/s111110859.

K. P. Hilgersom, N. C. van de Giesen, P.G.B. De Louw, and M. Zijlema. Three-dimensional dense distributed temperature sensing for measuring layered thermohaline systems. *Water Resources Research*, 52(8):6656–6670, 2016. ISSN 19447973. doi: 10.1002/2016WR019119.

Koen Hilgersom, Marcel Zijlema, and Nick Van De Giesen. An axisymmetric non-hydrostatic model for double-diffusive water systems. *Geoscientific Model Development*, 11(2):521–540, 2018. ISSN 19919603. doi: 10.5194/gmd-11-521-2018.

O. A. C. Hoes and N. C. Van de Giesen. *Lecture notes Polders (CT4460)*. Number April. TU Delft, Delft, 2015. ISBN 06917300084.

O. A C Hoes, W. M J Luxemburg, M. C. Westhoff, N. C. Van De Giesen, and J. Selker. Identifying seepage in ditches and canals in polders in the netherlands by distributed temperature sensing. *Lowland Technology International*, 11(2):21–26, 2009. ISSN 13449656.

Hoogheemraadschap Rijnland. Legger oppervlaktewater 2014. URL http://rijnland.webgispublisher.nl/?map=Legger-watergangen. Accessed 2018-05-27.

I T Jolliffe. Principal Component Analysis, Second Edition. *Encyclopedia of Statistics in Behavioral Science*, 30(3):487, 2002. ISSN 00401706. doi: 10.2307/1270093. URL http://onlinelibrary.wiley.com/doi/10.1002/0470013192.bsa501/full.

I. Kelderman. Slimmer inlaten in de Haarlemmermeerpolder. Technical report, Wageningen University, 2015.

Alan D Kersey. Optical fiber sensors for permanent downwell monitoring applications in the oil and gas industry. *IEICE transactions on electronics*, 83(3):400–404, 2000. ISSN 09168524. doi: 10.1117/12.2302132. URL http://search.ieice.org/bin/summary.php?id=e83-c{_}3{_}400{%}5Cnpapers3://publication/uuid/DEB96EFA-D28A-47DB-A18F-C2C6ACA51E25.

KNMI. KNMI - Maand- en jaarwaarden. URL https://www.knmi.nl/nederland-nu/klimatologie/maandgegevens. Accessed 2018-06-06.

Andreas Krause, Jure Leskovec, Carlos Guestrin, Jeanne M. VanBriesen, and Christos Faloutsos. Efficient sensor placement optimization for securing large water distribution networks. *Journal of Water ...*, 134 (6):516–526, 2008. ISSN 0733-9496/2008/6-516–526. doi: 10.1061/(ASCE)0733-9496(2008)134:6(516). URL http://ascelibrary.org/doi/abs/10.1061/(ASCE)0733-9496(2008)134:6(516).

A.F. Mills. *Heat transfer*. Prentice Hall, second edi edition, 1999. ISBN 0139476245.

K. Moore and E. Ross. Greedy Algorithms. URL https://brilliant.org/wiki/greedy-algorithm/. Accessed 2018-06-15.

M.J. Moran and H.N. Shapiro. *Fundamentals of Engineering Thermodynamics*. John Wiley & Sons, sixth edit edition, 2010. ISBN 9780470540190.

G. H. P. Oude Essink, E. S. Van Baaren, and P.G.B. De Louw. Effects of climate change on coastal groundwater systems: A modeling study in the Netherlands. *Water Resources Research*, 46(10):1–16, 2010. ISSN 00431397. doi: 10.1029/2009WR008719.

Planbureau voor de Leefomgeving. Correctie formulering over overstromingsrisico Nederland in IPCC-rapport - PBL Planbureau voor de Leefomgeving. URL http://www.pbl.nl/dossiers/klimaatverandering/content/correctie-formulering-over-overstromomgsrisico. Accessed 2018-06-12.

Barry Ruddick. A practical indicator of the stability of the water column to double-diffusive activity. *Deep Sea Research Part A, Oceanographic Research Papers*, 30(10):1105–1107, 1983. ISSN 01980149. doi: 10.1016/0198-0149(83)90063-8.

Ralph Schra. *Will Precipitation Have an Effect on the Salinity Concentration in the Haarlemmermeer Polder ?* Bachelor thesis, Vrije Universiteit Amsterdam, 2013.

Scikit Learn developers. 3.3. Model evaluation: quantifying the quality of predictions — scikit-learn 0.19.1 documentation, 2017. URL http://scikit-learn.org/stable/modules/model{_}evaluation.html{#}r2-score. Accessed 2018-06-13.

J. Selker, Luc Thévenaz, Hendrik Huwald, Alfred Mallet, Wim Luxemburg, Nick Van De Giesen, Martin Stejskal, Josef Zeman, M. C. Westhoff, and Marc B. Parlange. Distributed fiber-optic temperature sensing for hydrologic systems. *Water Resources Research*, 42(12):1–8, 2006a. ISSN 00431397. doi: 10.1029/2006WR005326.

J. Selker, Nick C. van de Giesen, M. C. Westhoff, Wim Luxemburg, and Marc B. Parlange. Fiber optics opens window on stream dynamics. *Geophysical Research Letters*, 33(24):27–30, 2006b. ISSN 00948276. doi: 10.1029/2006GL027979.

Silixa. ULTIMA™ DTS. URL http://silixa.com/products/ultima-dts/. Accessed 2018-06-18.

Statistics How To. Variance: Simple Definition, Step by Step Examples - Statistics How To, 2018. URL http://www.statisticshowto.com/probability-and-statistics/variance/. Accessed 2018-06-12.

P J Stuyfzand and R J Stuurman. Recognition and genesis of various brackish to hypersaline groundwaters in the Netherlands. In *Proceedings 13th Salt Water Intrusion Meeting (SWIM)*, pages 125–136, 1994.

The SciPy community. numpy.corrcoef — NumPy v1.14 Manual, 2018. URL https://docs.scipy.org/doc/numpy-1.14.0/reference/generated/numpy.corrcoef.html. Accessed 2018-06-12.

J.S. Turner. *Buoyancy effects in fluids*. Cambridge University Press, 1973.

Nick van de Giesen, Susan C. Steele-Dunne, Jop Jansen, O. A C Hoes, Mark B. Hausner, Scott Tyler, and John Selker. Double-ended calibration of fiber-optic raman spectra distributed temperature sensing data. *Sensors (Switzerland)*, 12(5):5471–5485, 2012. ISSN 14248220. doi: 10.3390/s120505471.

James J. Smolen & Alex van der Spek. Distributed Temperature Sensing A DTS Primer for Oil & Gas Production. *Proc.10th ISH*, page 97, 2003. URL http://w3.energistics.org/schema/witsml{_}v1.3.1{_}data/doc/Shell{_}DTS{_}Primer.pdf.

S.M. Van Grevenbroek. *Influence of Fresh Water Inlet on the Spatial and Temporal Variation of Electrical Conductivity of Haarlemmermeer Polder Drainage System*. Bachelor thesis, Vrije Universiteit Amsterdam, 2013.

A. Vandenbohede, P.G.B. De Louw, and P. J. Doornenbal. Characterizing preferential groundwater discharge through boils using temperature. *Journal of Hydrology*, 510:372–384, 2014. ISSN 00221694. doi: 10.1016/j.jhydrol.2014.01.006. URL http://dx.doi.org/10.1016/j.jhydrol.2014.01.006.

Water Nexus. Water Nexus: Saline water when possible, fresh water when needed | Water Nexus. URL http://waternexus.nl/. Accessed 2018-05-26.

Eric W. Weisstein. Exhaustive Search. URL http://mathworld.wolfram.com/ExhaustiveSearch.html. Accessed 2018-06-15.

M. C. Westhoff, T. A. Bogaard, and H. H. G. Savenije. Quantifying spatial and temporal discharge dynamics of an event in a first order stream, using distributed temperature sensing. *Hydrology and Earth System Sciences*, 15(6):1945–1957, 2011. ISSN 1607-7938. doi: 10.5194/hess-15-1945-2011. URL http://www.hydrol-earth-syst-sci.net/15/1945/2011/.

D.G. Wright. An Equation of State for Use in Ocean Models: Eckart's Formula Revisited. pages 735–741, 1997. ISSN 0739-0572. doi: 10.1175/1520-0426(1997)014<0735:AEOSFU>2.0.CO;2.

M. Xu, P. J. Van Overloop, N. C. Van De Giesen, and G. S. Stelling. Real-time control of combined surface water quantity and quality: Polder flushing. *Water Science and Technology*, 61(4):869–878, 2010. ISSN 02731223. doi: 10.2166/wst.2010.847.