Statistical investigation of trends and spatial distribution of Danish drinking water quality

Irene Cavasin

DTU

Università
degli Studi
di Padova

March 2014

Technical University of Denmark
Environmental Engineering
Building 113, DK-2800 Kongens Lyngby, Denmark

Preface

This Master thesis was prepared at the department of Environmental Engineering at the Technical University of Denmark in fulfillment of the requirements for acquiring a M.Sc. degree in Civil Engineering. This masters thesis was carried out from September 2013 to February 2014 under the supervision of Assistant Professor Martin Rygaard from the Department of Environmental Engineering of Danish Technical University and with the support of Full Professor Andrea Defina from the department of Civil and Environmental Engineering of University of Sudies of Padua.

Lyngby, March 2014

Irene Cavasin

Acknowledgements

At the end of my (amazing) Erasmus period and (unfortunately) also of my student life, I would like to say:

THANK YOU!! to Martin. You supervised me in a really professional way: giving me the room to do my own choices in the thesis but also giving me the right motivation.

THANK YOU!! to my parents Paul and Lui. Apart the economical support, you backed me up in all my crazy ideas with your young mind.

THANK YOU!! to my big bro Ame. For the Coronas brought up to Lagazuoi's lake.

THANK YOU!! to my friends in Padua, in particular Marghe, Silvy, Jacopo, Esad, Stefano, Simon, Neven, Ale, Carla, Fede e Lidia for the numerous holidays and trips together, the awesome Wednesday nights, the tisanes in the evening and the long breaks in the "Suite".

THANK YOU!! to the professors in the Hydraulics department of Padua for having passed on your enthusiasm for the subject to me and Carlo for the house for the BBQ.

THANK YOU!! to the "databar-mates" for the unconventional moments of fun of 5.30pm.

THANK YOU!! for all the friends I met during this Eramus: it's great to get to know such good people with different interests, languages and background.

THANK YOU!! to Alberto to bear this really demanding girl with all your love.

Abstract

Danish water supplies are entirely based on groundwater and in Denmark the only treatments applied before the distribution are aeration, filtration and pH adjustment. For that reason, to keep the water source under control is particularly important. Groundwater is often considered a stable and consistent resource for drinking water supply but different treatment practices at the water works influence the final product. Nevertheless, no research were led about the long-term changing of drinking water after treatment at the water works. This study focuses on the changes over time and the average levels of the main inorganic water compounds with a particular attention on their effects on taste, corrosion and dental health.

Based on 163 785 measurements in 1 665 data sets for the period 1980-2011, we have conducted a statistical analysis of temporal trends in the water produced from 100 major Danish water works representing 168 million m^3 per year and about 56% of domestic consumption in Denmark. Trends were investigated using a combination of the Kendall test and a linear regression for 17 parameters: O_2, NH_4^+, Ca^{2+}, Fe^{2+}, Mg^{2+}, Mn^{2+}, K^+, Cl^-, F^-, HCO_3^-, NO_2^-, NO_3^-, SO_4^{2-}, pH, TDS, NVOC and temperature.

The number of recorded measurements in the official database (Jupiter) is not in compliance with the Danish law requirements for 28 out of 100 plants. Only two water works were found to do not have significant trend in any compound for the 32 -year period. With trends observed in 54 to 59% of the water works for oxygen, sulfate, pH, nitrate and chloride, these five parameters are most often associated with a trend over time.

The relatively highest change per decade of measurements was found for the compounds ammonium (<545%), iron (<163%) and nitrate (<159%). Relative change for each of the 17 compounds was further analyzed in comparison with current levels and water quality criteria. It was found that especially ammonium, pH, iron, nitrite and NVOC had noticeable variation and levels of concern for selected plants. Ammonium and iron in particular have high reliable trends with steep slope. Manganese, ammonium and iron experience a sharp decrease during the 1980s in some plants.

46 water works delivered water with a Larson Ratio below 0.5 indicating no potential for corrosion of steel and iron pipes. 11 water works had a Larson Ratio above 1 indicating a strong potential for corrosion of steel and iron. Looking also at water hardness and alkalinity, a general predisposition to scale forming in the chemical stability of water supplied is noticed. The area around København shows high levels of hardness (up to 530 mg/L), while high values of alkalinity (above 350 mg/L) characterize the region of Nordjylland.

Dental health was evaluated based on model's predictions of the content of calcium and fluoride, and 10 waterworks are found to have a decayed, missing and filled surfaces (DMF-S) below 2, indicating an optimal combination of calcium and fluoride in drinking water. Five plants had DMF-S above 4, indicating a sub-optimal water composition for caries prevention. The regions Hovedstaden and Sjælland have the highest number of plants closer to the optimum number of DMF-S of 2, mainly owing to the high fluoride content (0.24-1.24 mg/L).

Basing on the results of a chemometric study about drinking water taste, the inorganic compounds affecting taste were analyzed in addition to pH and TDS. Furthermore, a taste index was built with the intent to gather the inorganic compounds' effect on taste. The numerous opposite influences of the said parameters make the the choice of a "good water taste area" difficult. In general, the effect of high TDS levels (40 plants have meanly TDS above 400 mg/L) are compensated by water rich in inorganic compounds positive at taste.

In conclusion, our results show how drinking water quality of groundwater based water supplies undergoes significant change over time, and due to the large data set employed they give a detailed long-term overview of the state of a nationwide groundwater based drinking water quality.

Sommario

L'approvigionamento idrico in Danimarca è interamente basato su risorse sotterranee e gli unici trattamenti implementati prima della distribuzione sono aerazione, filtrazione e aggiustamento del pH. Per questa ragione, il monitoraggio delle risorse idriche riveste un'importanza fondamentale. L'acqua di falda è generalmente considerata una risorsa stabile e costante ai fini dell'approvigionamento idrico ma diverse pratiche di trattamento all'acquedotto vanno ad influire sul prodotto finale. Ciò nonostante, fino ad ora nessuna ricerca è stata condotta per valutare i cambiamenti a lungo termine dell'acqua potabile post trattamento all'impianto. Il presente studio è incentrato sui cambiamenti nel tempo e sui livelli medi dei principali componenti inorganici, con particolare attenzione ai loro effetti su gusto, corrosione e salute dentale.

Sulla base di 168 785 misure riunite in 1 665 insiemi di dati raccolti nel periodo 1980-2011, è stata condotta un'analisi statistica sugli andamenti temporali dell'acqua prodotta dai 100 maggiori acquedotti danesi, che, erogando annualmente 168 milioni di m^3, rappresentano circa il 56% del consumo domestico in Danimarca. Gli andamenti sono stati valutati attraverso l'insieme di test di Kendall e regressione lineare per 17 parametri: O_2, NH_4^+, Ca^{2+}, Fe^{2+}, Mg^{2+}, Mn^{2+}, K^+, Cl^-, F^-, HCO_3^-, NO_2^-, NO_3^-, SO_4^{2-}, pH, TDS, NVOC e temperatura.

Su 28 dei 100 impianti il numero di misure registrate nel database ufficiale (Jupiter) risulta non conforme con i requisiti minimi previsti dalla legge danese. Solo due acquedotti risultano avere nessun andamento significativo durante il periodo di 32 anni. Ossigeno, solfato, pH, nitrato e cloruro sono i cinque parametri associati più frequentemente a cambamenti temporali con un numero di andamenti osservati variabile dal 54 al 59% degli acquedotti considerati. Il cambiamento relativamente maggiore in un decade di misure è stato rilevato per ammonio (<545%), ferro (<163%) e nitrato (<159%). La pendenza relativa della regressione lineare per ciascuno dei 17 componenti è stata analizzata ulteriormente confrontandola con i livelli attuali e i criteri previsti sulla qualità dell'acqua. Si è notato che specialmente ammonio, pH, ferro, nitrite e NVOC hanno

riscontrato evidenti variazioni e livelli d'allerta per alcuni impianti selezionati. Ammonio e ferro in particolare hanno sperimentato un forte decremento durante gli anni Ottanta in alcuni impianti.

46 acquedotti distribuiscono acqua con Larson Ratio minore di 2, ad indicare l'assenza di potenziale corrosivo per tubi in ferro e acciaio. 11 acquedotti hanno avuto in media un Larson Ratio maggiore di 1, segno di forte potenziale corrosivo per tubi in ferro e acciaio. Osservando inoltre la durezza e l'alcalinità dell'acqua, si nota una generale predisposizione alla formazione di deposito nella stabilità dell'acqua. L'area attorno København riscontra alti livelli di durezza dell'acqua (fino a 530 mg/L), mentre alti valori di alcalinità (fino a 350 mg/L) caratterizzano la regione del Nordjylland.

La salute dentale è stata valutata attraverso previsioni da modello basato sul contenuto medio di calcio e fluoruro, quindi 10 acquedotti risultano avere un numero di superfici dentali cariate, mancanti o otturate (DMF-S) minore di 2, indicante un'ottima combinazione di calcio e fluoruro nell'acqua potabile. Cinque impianti hanno avuto mediamente un numero di DMF-S maggiore di 4, segno di non un'ottima combinazione di calcio e fluoruro. Le regioni Hovedstaden e Sjælland hanno il più alto numero di impianti aventi numero di DMF-S prossimo all'ottimo di 2, principalmente a causa dell'alto livello di fluoruro (0.24-1.24 mg/L).

Sulla base di uno studio chemometrico sul gusto dell'acqua, sono atati analizzati i componenti inorganici che influenzano il gusto oltre a pH e TDS. Inoltre, è stato formulato un "indice del gusto" nel tentativo di riunire l'effetto dei componenti inorganici. Le numerose e opposte influenze dei suddetti parametri rendono difficile la scelta di un'area con migliore gusto dell'acqua. In generale, l'effetto di alti livelli in TDS (40 impianti hanno mediamente TDS maggiore di 400 mg/L) sono compensati da un'acqua ricca in componenti inorganici positivi al gusto.

In conclusione, i risultati ottenuti mostrano come la qualità dell'acqua potabile basata su risorse sotterranee subisce canbiamenti significaivi nel tempo; e, a causa del grande insieme di dati impiegato, questi forniscono una dettagliata visione d'insieme sullo stato dell'acqua potabile di falda su base nazionale.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Drinking water quality affects everyone in several aspects of everyday life. This master thesis analyses the most important water compounds and parameters from samples taken in the largest plants located all over Denmark. The main purpose of this thesis is to find a repeatable method to obtain significant trends in time and spatial variation in drinking water quality. Furthermore, the consequences of the water quality are evaluated with regard to taste, pipe corrosion and health.

1.1 Motivation

Denmark's water supply is wholly based on groundwater sources. The Danish government's approach is to supply drinking water entirely with groundwater which requires only aeration, filtration and pH adjustment, apart from a few plants. Consequently, the maintenance of the quality of water resources is pursued rather than the water treatment after catchment. For this reason it is particularly important to preserve the groundwater sources throughout Denmark from pollution and monitor them.

The industrialization of farming during the 20th century led to the approval of the Statutory Order on Water Quality and Supervision of Water Supply Plants by the Ministry of Environment in 1988. It provides for quality monitoring of abstracted groundwater, which needs a permit to be renewed at specific times. Moreover, the annual abstraction measured must be reported to the authorities once a year.

From 1991 to 2005, 1306 wells were closed because of the content of pesticides or metabolites. Many minor water works closed owing to nitrate pollution while pesticide pollution caused the closure of mainly major water works. In order to preserve the quality of the groundwater quality, Danish Nitrates Directives (91/676/EEC) imposed the nitrate level below 50 mg/l and pesticide levels below 0.1 μg/l.

In 1998 the Danish water supply act established that all water supply data collected must be reported to the national groundwater database Jupiter managed by GEUS.

In 2004 the NOVANA subprogramme was initiated to monitor groundwater quality to meet Denmark's expected monitoring obligations pursuant to the Water Framework Directive (2000/60/EC) and the forthcoming Groundwater Directive. A wider range of variables started to be analyzed, and the current programme monitors a total of 97 substances: 26 remain chemical elements, 14 inorganic trace elements, 23 organic micropollutants and 34 pesticides or pesticide degradation products. In the study carried out, all the inorganic compounds have been included. The current Danish law about drinking water quality is *Bekendtgørelse om vandkvalitet og tilsyn med vandforsyningsanlæg* (n. 1024 of 31/10/2011). This law gives the compulsory frequency and the parameters to check, which will be referred to later on (Chapter 6).

1.2 Thesis aim

The main purpose of this thesis is to develop a method to compare the average levels and variations of a selection of the main inorganic water compounds found in drinking water. The chosen elements and parameters to analyze are pH, temperature, TDS, oxygen, NVOC, nitrite, nitrate, manganese, magnesium, iron, hydrogen carbonate, fluoride, chloride, calcium, ammonium and potassium. Each of the previous compounds influences at least one water characteristic among taste, health and corrosion in the distribution system. Consequently, the impacts of the current water quality on some health aspects, taste and corrosion will be evaluated through the following procedure.

The specific aim of the thesis is to give an overall evaluation of the status and trends of the drinking water quality for 100 major water supplies in Denmark during the years 1980-2012. The steps are to:

- Develop a method for filtering data from the Jupiter database for the effluents of selected water works and conduct a statistical analysis of the measured drinking water quality
- Calculate the averages over time and the slope of time trends for all compounds to point out possible noticeable variation and levels of concern among the selected water works

- Make a comparison between the mean compound levels between the five regions of Denmark
- Evaluate the results against optimum drinking water criteria suggested in literature and the law prescriptions, in reference to the distribution system and appliance corrosion, health and taste.

1.3 Thesis structure

This thesis presents a long-term evaluation of the levels and trends of all inorganic compounds and other parameters like temperature, pH and TDS in the 100 major water supply plants in Denmark. First, a data analysis explores whether a significant linear trend could be seen over the last thirty years at a given level of significance. Twenty plants for each region in Denmark will be considered. The average values will be used to assess the mean water quality effects on corrosion, dental health and taste. The results will be inserted in maps to highlight some possible spatial variations and then evaluated to see if they match the optimum drinking water criteria founded in literature.

In the second chapter, *The data*, the database features are described in addition to the procedure with which data have been filtered. A paragraph focuses on the output of the queries.

In the third chapter *Method* the topic is the implemented statistics. Firstly, it deals with the hypothesis testing led to assess which trends are significant, named Kendall's test. Secondly, some observations about the linear regression are noted.

In the fourth chapter, *Implemented program*, there is the description of the implemented code, and in the fifth chapter, *Results*, the discussion of the maps, noticeable levels, trends and results will be presented.

Chapter 2

Background

2.1 Water supply in Denmark

Denmark has an area of about 43000 km^2 and 5.5 million of inhabitants. It is divided into 5 regions and 98 municipalities after the municipal reform of 2007.

At present, there are about 2500 waterworks in Denmark which are completely fed with underground water. Water coverage, sanitation coverage and continuity of supply are fully assured by the Danish water supply system. Indeed, 90% of the population is connected to piped water services and all urban and rural households have access to safe water provisions and sewage services. In total, about 800 million of m^3 of water are abstracted annually (GEUS [2008]). There are approximately 250 large extraction plants (i.e. plants which obtain more than 350000 m^3 per year). These plants account for the recovery of about 250 million m^3 annually. In addition to the public water supply system, between 70000 and 80000 m^3 per year are supplied by non-public water systems (i.e. supplying less than 10 properties (Naturstyrelsen [2012])).

As far as water utilization is concerned, the average household water consumption is 41.4 m^3 per per person per day. Nevertheless, water consumption is down to 3.8% relatively to 2008, owing to a decrease in all areas of consumption such as households, industry institutions and unregistered consumption. The decrease was caused by the increment of taxes and water saving campaigns. In 2010, the water consumption levels corresponded to 67.7 m^3 per person per year (Dansk Vand-og Spildevandforening (DANVA) [2011]).

Concerning the administration of water supply, service provision is granted only by public and cooperative providers. Seven state environmental centers are responsible for the planning of water under the Water Framework Directives. Regions work in matters of remediation of pollution from old waste dumps and monitoring groundwater. Moreover, Regions undertake the regional reporting and submission of data to the Topic Centre for Groundwater and Wells at the Geological Survey of Denmark and Greenland. Water works are responsible for the control of the groundwater that is used to produce drinking water (well control) pursuant to the Statutory Order on water quality. Municipalities manage the water abstraction permits and the protection of water resources against pollution.

The current law establishes three monitoring points of drinking water quality: at the water works, in the water mains and in the homes of consumers. Each plant treats groundwater abstracted from several plants and sometimes from several aquifers. A change of wells used as a source mirrors a change in water compounds' content.

2.2 Water quality - criteria and optimum range

The recent research in desalination of water supplies has increased the interest into the optimum levels of compounds concerning water quality, thus numerous articles are available in literature about this topic (Lahav and Birnhack [2007],Rygaard et al. [2009]) . The fact that water quality effects health, corrosion in pipes and taste is well-known, and in Rygaard et al. [2011] for some water compounds which effects one of those three aspects, the optimum range is given.

2.2.1 Law requirements

This thesis considers three different legal paradigms: Danish law in matter of water quality and monitoring of water supply systems (Ikrafttr [2014]), WHO guideline for drinking-water quality (Gorchev and Ozolins [1984]) and Australian drinking water guidelines (water quality management [2011]). The averages of the compounds over time are compared with the Danish law requirements. Danish law is taken as reference for the tolerable levels of compounds because it is stricter than the WHO guidelines; even though, WHO guidelines give a comprehensive explanation of the effects on health, taste and distribution system. The Australian guidelines were compared with the European requirements and their summary is the content of appendix F. The Danish law establishes quality requirements for drinking water at three locations in the distribution network: at treatment works, at the entrance to the property and at the consumer's tap. For the parameters that are not affected by the distribution system, the quality standards are the same for all three checkpoints. In this thesis, the official data at the effluent of the plant will be considered; the reason of this choice is the intention to evaluate the drinking water quality after treatment but without consideration of the possible influence of the pipeline network.

Table 2.1 reports the law prescriptions for all evaluated compounds considering the Danish law, the WHO guidelines and the Australian law.

TABLE 2.1: Limits for each compound given by the Danish law (Ikrafttr [2014]), WHO guidelines (Gorchev and Ozolins [1984]) and Australian law (water quality management [2011])

Symbol	**Danish law**	**WHO guidelines**	**Australian law**
O_2	>8	No guidelines but very high levels of dissolved oxygen may exacerbate corrosion of metal pipes.	>85% (aesthetic guideline)
NH_4	0.05	Occurs in drinking-water at concentrations well below those of health concern (ammonia)	no guidelines but aesthetic considerations <0.5 mg/L (ammonia)
Ca^2+	200	see hardness, taste threshold 100−300 mg/L	see hardness
Fe	0.1	Not of health concern at levels causing acceptability problems in drinking water	<0.3 mg/L (aesthetic considerations)
Mg^2+	50	see hardness	see hardness
Mn	0.02	Not of health concern at levels causing acceptability problems in drinking-water	<0.1 mg/L at the tap, <0.5 mg/L for health considerations
K^+	10	Occurs in drinking-water at concentrations well below those of health concern	−
Cl^-	250	Not of health concern at levels found in drinking-water	no guidelines but aesthetic considerations <250 mg/L
F^-	1.5	1.5	<1.5 mg/L (aesthetic considerations)
HCO_3^-	>100	−	−
NO_2	0.017	3	3
NO_3^-	50	50	50 (bottle-fed infants), 100 (people being older than 3 months)
SO_4^{2-}	250	Not of health concern at levels found in drinking-water	<250 mg/L (taste considerations),<500 (health considerations)
pH	7−8.5	Not of health concern at levels found in drinking-water	6.5−8.5 (corrosion and encrustation)
TDS	1500	Not of health concern at levels found in drinking-water	no guidelines, but <600 for good palatability
NVOC	4	−	−
T	>12°C	−	−
TH	5°−30° dH	Not of health concern at levels found in drinking-water	<200 mg/L as $CaCO_3$

2.2.2 Dental health

Regarding dental health, Bruvo et al. inferred a strong correlation between calcium (Ca^{2+}) and fluoride (F^{-}) and the variations in the numbers of decayed, filled, and missing tooth surfaces (DMF-S) of inhabitants who use public drinking water (Bruvo et al. [2008]). DMF-S is a measure of dental caries counting the damaged teeth surfaces. Molars and premolars are considered to have 5 surfaces, the front tooth 4 surfaces. The maximum value for DMF-S comes to 128 for 28 teeth. Bruvo et al. calculated a model giving the number of DMF-S, based on the calcium and chloride content in drinking water (with all concentrations expressed in mg/L):

$$DMF - S = \exp\left\{1.05 - \frac{0.18([F^{-}] - 0.33)}{0.25} - \frac{0.11([Ca^{2+}] - 83.5)}{25.63}\right\}$$

From an analysis based on the Jupiter database (for the chemical data of water) and DMF-S data for Danish schoolchildren (Danish National Board of Health), a statistical model based only on calcium and fluoride was able to explain 45% of the variations in DMF-S among 15-year-old Danish adolescents (Bruvo et al. [2008]). Other variables affecting dental health are for example the personal income in the families or mothers' education but with weaker correlation (Ekstrand et al. [2003]).

Dental health improves if the content of calcium and fluoride in drinking water increases. In particular, there is a moderate relationship between increasing levels of fluoride and low mean DMF-S levels up to about 0.3 mg/L. For fluoride concentrations above 0.3 mg/L, that association seems to disappear. Fluoride makes the tooth surface more resistant to acids, but excessive fluoride levels cause fluorosis. Precisely, at fluoride level of 1 mg/L, 12.5% of the exposed people have fluorosis causing aesthetic concerns (McDonagh [2000]). Calcium favors remineralization, reduces demineralization in early caries lesions and also it diffuses into the plaque providing extra binding sites for fluoride (Bruvo et al. [2008],Ekstrand et al. [2003]).

Bruvo et al. recommend the optimal content of calcium and fluoride to be respectively 90 mg/L and 0.75 mg/L, which leads to an estimated 2 DMF-S. In comparison, in København an average of 3.04 DMF-S/person was observed in 2005 (Rygaard et al. [2009]), and a detailed analysis in 1999 from the National Board of Health showed that at the municipality level mean caries experience among 15-year-olds ranged from slightly lower than 1 DMF-S to about 9 DMF-S (Ekstrand et al. [2003]).

2.2.3 Corrosion

Corrosion is a huge topic, so in this analysis only a few aspects will be evaluated in order to have an overall view about corrosion in iron and steel pipes. Firstly, the average alkalinity of water is evaluated and then the Larson index. Furthermore, water hardness is calculated in order to compare it with the Danish directives and evaluate the possible occurrence of a scale build-up.

Alkalinity provides buffering capacity to aqueous systems. It is mathematically calculated as:

$$Alk = [HCO_3^-] + 2[CO_3^{2-}] + [OH^-] - [H^+]$$

In freshwater the amount of carbonate can be neglected (see Figure 2.1), as well as hydroxyl (OH^-) and hydron (H^+)(as an example if pH is 7.5, $[H^+]$=0.000032 mg/L and $[OH^-]$=0.000316 mg/L). Hence, the final equation is:

$$Alk = (EW(CaCO_3^-))/(EW(HCO_3^-))[HCO_3^-] = 50/61[HCO_3^-]$$

where the alkalinity is expressed in mg/L as $CaCO_3$ and EW means equivalent weight.

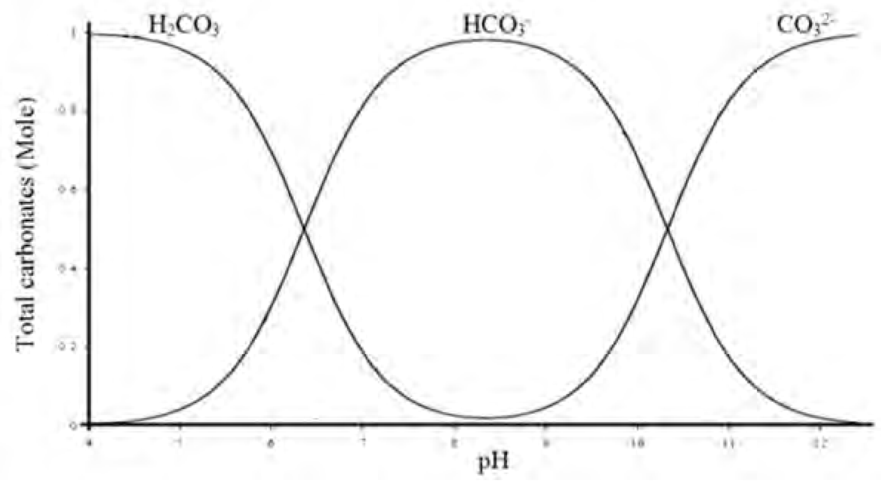


Figure 2.1: Relationship between pH scale and Total Inorganic Carbon in water. In the pH range of drinking water (around 7-8.5) hydroxyl (OH^-) and hydron (H^+) can be neglected.

Low values of alkalinity, below 75 mg/L as $CaCO_3$, may make the pH change and water become more corrosive, causing the dilution of damaging metal from the plumbing to contaminate the drinking water. Alkalinity above 500 mg/L as $CaCO_3$ is associated with high values of dissolved solids and hardness which may cause scale build up, especially in hot water. Scale in plumbing systems make the costs to heat water higher and increase power consumption (University of Rhode Island). In the same source, the following classification is given (see Table 2.2):

Table 2.2: General guideline of alkalinity values for drinking water (University of Rhode Island)

Alkalinity (as $CaCO_3$)	**Corrosivity**
0-100 mg/L and pH<7	Corrosive water
100-200 mg/L	Ideal range
>200 mg/L	Possible deposit o scale build up

Similarly, Lahav proposed alkalinity levels above 80 mg/L as $CaCO_3$ for preventing the release of metal ions to the water (Lahav and Birnhack [2007]). The Larson Ratio (LR) measures the cast iron and steel corrosiveness in relation to chlorides, sulphate and alkalinity (Rygaard et al. [2011]). Chloride and sulphate are aggressive anions which make the general corrosion of carbon steel and galvanized steel to accelerate. It is computed as the ratio between the sum of chloride and sulphate and the sum of bicarbonate and carbonate, all concentrations expressed in equivalent weight. When pH<9, as for the alkalinity, carbonate content is neglected. Thus:

$$LR = \frac{[SO_4^{2-}] + [Cl^-]}{[HCO_3^-]}$$

with all the concentrations expressed in equivalent per liter. Different limits for LR are given, but water with LR below 0.5 can be considered not corrosive for steel and cast iron. In particular, the corrosive potential of water can be classified according to the LR (see Table 2.3) (Delion et al. [2004]).

Table 2.3: Classification of water corrosiveness basing on the Larson ratio (Delion et al. [2004])

LR	**Corrosive potential**
<0,2	no potential
0.2 to 0.4	small potential
0.4 to 0.5	slight potential
0.5 to 1	average potential
>1	strong potential

Lahav and Birnmark consider pH, the Larson ratio and alkalinity in their assessment of desalinated water quality after the post-treatment stage. Quality criteria consider water quality in terms of chemical stability, bio-stability, effects on waste water treatment, and water palatability, as well as health and economic effects. (Lahav and Birnhack [2007]). Alkalinity should be close to 100 mg/L as $CaCO_3$ at a pH value that is as closed as possible to pH 8.0. Alkalinity concentration above 80 mg/L as $CaCO_3$ is the most important individual parameter for preventing the release of metal ions into the water. A high pH is advantageous in terms of buffering capacity, but it may yield a build-up of excessive $CaCO_3$ scales on pipes, pumps, etc. In accordance to the previous table, Rygaard suggest a LR values below 0.5 (Rygaard et al. [2011]).

2.2.4 Water taste

With regard to taste, Platikanov led a chemometric approach to evaluate which compounds influence the palatability of water (Platikanov et al. [2013]). This survey is referred to because it has a higher reliability in comparison to the studies based on evaluations made by untrained assessors. The investigation on the water samples tasted by trained panelists, showed that the preferred waters were associated with moderate contents of total dissolved solids (TDS) and with relatively high concentrations of HCO_3^-, SO_4^{2-}, Ca^{2+}, and Mg^{2+} as well as with relatively high pH values. In contrast, panelists disliked water samples with high concentrations of K^+, Na^+, Cl^- , NO_3^- and Si. The given range for TDS is 200-400 mg/L and pH between 7.5 and 8.1. TDS is the most important factor which influences the taste. A WHO report about TDS (WHO [1996]) gives the following classification for the impact of TDS in water taste (Table 2.4):

Table 2.4: Effect of TDS content on water taste given by WHO (WHO [1996]).

level of TDS [mg/L]	Rating
< 300	Excellent
300-600	Good
600-900	Fair
900-1200	Poor
> 1200	Unaccettable

Furthermore, TDS deteriorates materials and should be maintained as low as possible; the proposed content is below 200 mg/L (Rygaard et al. [2011]).

The taste index is a parameter created in order to quantitatively evaluate the taste based on the results of the survey led by Platikanov et al. (Platikanov et al. [2013]). The taste index is the fraction of the compounds with a positive influence on taste divided by the total amount of compounds that influence taste, all concentration expressed in mg/L. Hence, the equation results:

$$TI = \frac{\frac{[Ca^{2+}]}{20.05} + \frac{[Mg^{2+}]}{12.15} + \frac{[SO_4^{2-}]}{40.03} + \frac{[HCO_3]}{61.02}}{\frac{[Ca^{2+}]}{20.05} + \frac{[Mg^{2+}]}{12.15} + \frac{[SO_4^{2-}]}{40.03} + \frac{[HCO_3^-]}{61.02} + \frac{[K^+]}{39.10} + \frac{[Cl^-]}{35.45} + \frac{[NO_3^-]}{62.00}}$$

Water taste is evidently subjective, thus it is difficult to find a numeric parameter able to represent it on an objective scale. Consequently, the numeric values have to be taken as indicators more than exact values.

The effect of each singular compound which effects taste is then evaluated in a radar graph separately. Firstly, it is presented as the percentages given by the ratio between the average in the region of the concentrations of the compound in the region divided by the average concentration in all Denmark. Then, the percentages of those compounds are weighted with the effluent discharge from each plant normalized by the total effluent per each region. The final percentages can be summarized with the following formula:

$$percentage = \frac{C_{plant}}{C_{DK}} \frac{Q_{plant}}{Q_{region}}$$

Where:

C_{plant} is the average concentration over time in the plant
C_{DK} is the average concentration over time in all of Denmark
Q_{plant} is the average effluent over time in the plant
Q_{region} is the average effluent over time in all regions

Finally, the percentages are averaged within each region and they are presented in a radar graph.

2.2.5 General water quality recommendations

In addiction to the said parameters, hardness was also calculated, very hard water being common in Denmark (Hilbert et al. [2010]). “Hard” water is high in dissolved minerals, in particular calcium and magnesium. Those minerals dissolve in groundwater during its passage through soil and rock. High values of hardness does not pose any health risk, but may cause nuisances, such as (University of Rhode Island):

- Build-up of scales in the internal surface of pipes and fixtures, shortening the life of plumbing systems
- The efficiency of electric water heaters is reduced
- Detergent and soap produce less foam
- Brewed coffee and tea taste more bitter

A limit of 250 mg/L of $CaCO_3$ is suggested by Rygaard, because it ensures low soap consumption and possibly reduce eczema risk. Australian guidelines identify four categories for water hardness (see Table 2.5) based on the corrosion effects.

Table 2.5: Hardness effects on plumbing system according to the Australian Guidelines (water quality management [2011]).

Hardness	**Corrosion effect**
<60 mg/L $CaCO_3$	Soft but possibly corrosive
60-200 mg/L $CaCO_3$	Good quality
200-500 mg/L $CaCO_3$	Increasing scaling problems
>500 mg/L $CaCO_3$	Severe scaling

In addition to the optimum calcium level suggested by Bruvo (80 mg/L) according to dental issues, other calcium ranges are proposed in literature: Lahav gives 80 to 120 mg/L as optimum level for calcium considering its effect on chemical stability, biostability and economic and health issues; Rygaard propose calcium levels beetween 40-50 mg/L to maintain low water hardness and sufficient magnesium. Furthermore, Rygaard extend the optimum level for fluoride given by Bruvo (0.75 mg/L) to the range 0.5-1 mg/L on the basis of considerations about dental health expenses. Since all Danish drinking water comes from groundwater wells, it is rich in mineral content, if compared with country in which drinking water comes from surface sources (Grossman and Yarger [1963]).

The precedent optimal ranges or levels from the literature are reported (see Table 2.6).

Table 2.6: Optimal levels of the analyzed compounds from literature.(Bruvo et al. [2008],Rygaard et al. [2011],Lahav and Birnhack [2007])

.

Reference	**Parameter or compound**	**Optimum level or range**
(Bruvo M., 2008)^1, (M. Rygaard, 2011)^2,	Fluoride	0.75 mg/L^1, 0.5-1 mg/L^2
(Bruvo M., 2008)^1, (M. Rygaard, 2011)^2, (Lahav O., 2006)^3	Calcium	90 mg/L^1, 40-50 mg/L^2, 80-120 mg/L^3
(Lahav O., 2006)	Alkalinity	>80
(Lahav O., 2006)	pH	<8.5
(M. Rygaard, 2011)	Larson Ratio	<0.5
(M. Rygaard, 2011)	Magnesium	>10
(M. Rygaard, 2011)	Hardness	<150 mg/L as $CaCO_3$
(M. Rygaard, 2011)	TDS	<200

Chapter 3

The data

3.1 GEUS and Jupiter database

The analysis is based on data from Jupiter, the national borehole database, maintained by GEUS (Geological Survey of Denmark and Greenland). The database contains data on wells, geology, water level and groundwater chemistry, as well as data about water supply, water abstraction licenses, yearly abstraction, and the exchange of drinking water between waterworks installations in addition to the chemistry of drinking water. Data are provided by the drilling companies and waterworks. Data are collected following a timetable depending on the annual discharge of treated or distributed water in accordance with Danish legislation (see Table 3.1). In the normal controls, the compounds analyzed are NVOC, ammonium, iron, manganese, chloride, fluoride, nitrate, nitrite, sulfate, temperature and pH; in the extended controls TDS, calcium, magnesium and hydrogen carbonate are also analyzed. The choice analyze the data from the effluent of the treatment plants was taken owing to the following reasons:

- To investigate the actual drinking water quality that one Danish citizen takes. In fact, one water work can be supplied by more wells, also located far away from the water distribution system. Thus, sometimes it can be difficult to establish where a certain user's water is taken. However, every user can determine which water works the water is treated by just looking at the bill.
- The water quality should be steady over time after treatment, but no investigations have been led so far to check it. Sometimes the water quality is not so stable as it is supposed to be owing to several causes (such as the change of wells from which the water is taken, changes in water treatment operation, the introduction of new law requirements, etc.).

TABLE 3.1: Frequency of compulsory drinking water sampling at the effluent from the Danish law (Ikrafttr [2014]). 11 among the chosen compounds are measured in the normal controls and 15 in the extended controls. Oxygen and potassium measurements are not mandatory at the water works.

Water volume	**Normal controls**	**Extended controls**
m^3 per year	**studies per year**	**studies per year**
3000-10000	0.5	0.5
10000-35000	0.5	0.5
35000-350000	1	1
35000-1500000	1	1
1500000-2600000	2	1
2600000-3500000	3	1
3500000-7000000	4	1
7000000-10500000	4	2
10500000-14000000	5	2

Water supply facilities with a treated or distributed water volume between 14 and 35 million m^3 water per year have a different frequency of measurements, which depends on the effluent as well (Ikrafttr [2014], see Table 3.1). All data are then updated by both GEUS and local authorities (see Figure 3.1). Laboratories must deliver analyses of drinking and groundwater conducted for waterworks directly to the database, back to the water-supply plants and to the municipality; data have to be approved by the municipalities, and after quality control they are immediately available online. Due to those transparency requirements, data are documented with the corresponding drilling method, drilling company, who performed the geologic description, and a description of which measurement method was used. There is no information regarding the observational error.

Jupiter contains data of about 2800 waterworks all around Denmark. Jupiter database is available using Access, Firebird, Oracle, SQL server-format at the website `http://data.geus.dk/JupiterWWW/downloadpcjupiter.jsp?xl=1`. The database is made of about 60000 records of measurements, which are stored according to the number of the sample; each sample is then related with one plant and with one date. Each plant is characterized by a plant tracking number (field PLANTID in Jupiter database) that refers to a specific town (field LOKALITET). The number must be unique (or, at least, should be) on a country level. Jupiter is composed of two main categories of measurements: one deals with the samples from the water abstraction (it contains data about geology, quality of water at the caption, use of water), while the other deals with the quality of water for consumption (after treatment in the plant). In this context the attention is on the latter part. In Jupiter both organic and inorganic compounds are considered, but the present survey will select only the main inorganic compounds, some

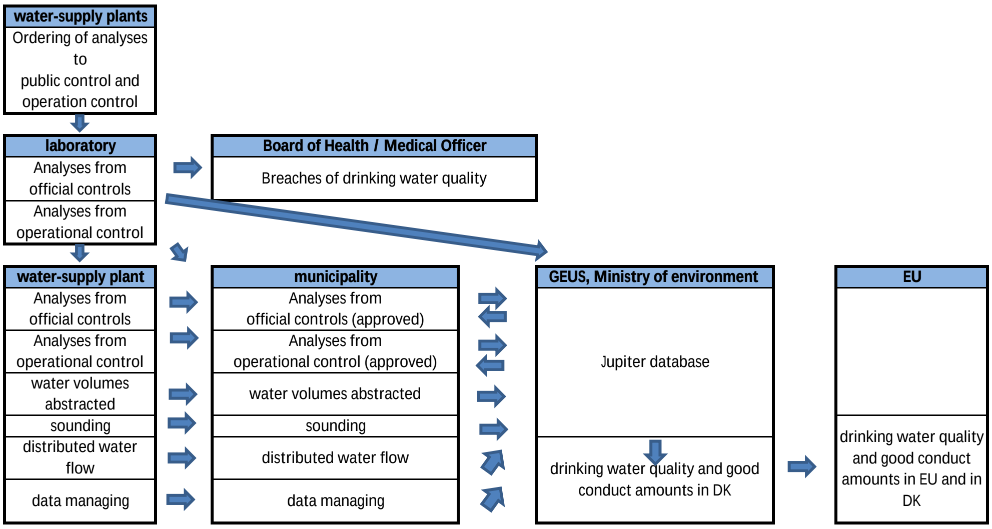


Figure 3.1: Control procedure applied to the data collected from the water works to Jupiter database. All data have to be check and approved from the municipalities (Ikrafttr [2014]).

water characteristics as temperature, pH, TDS (total evaporation residue) and a total index summarizing all organic compound amounts (NVOC, organic carbon). On the GEUS website one can download all data region by region, because each of them has a different database.

3.1.1 Selection of data (queries)

In order to select the measurements averaged by month relating to only the wanted plants and compounds, two queries are necessary:

- A query which selects the main twenty plants for each region separately (named largest_plants_[region_name])
- A query that extracts all the measurements for the given compounds and plants and organizes them in a pivot table, calculating their monthly averages from 1980 until 2012 (named PivotTable_[Region_name]).

In Jupiter all measurements and plant characteristics are within tables connected to each other. Following is the list of the used tables:

- DRWPLANT
- WRRCATCHMENT
- DRWPLANTCOMPANYTYPE
- DRWCHEMANALYSIS
- DRWCHEMSAMPLE

The first three tables are used in the first query to select the main ten plants for each region of Denmark respecting the following criteria:

1. The plants are ranked based on the average produced volume reported (AMOUNT (m^3 per year) in WRRCATCHMENT).
2. The query selects only the plants not referring to well fields (that is those ones which do not have “kilde”as part of the plant name, PLANTNAME in DRW-PLANT).
3. The query filters out all plants not related to drinking water production, only those belonging to the company type V01 and V02 (COMPANYTYPE in DRW-PLANTCOMPANYTYPE).
4. All non-active plants were excluded (through the field ACTIVE in DRWPLANT).

The three tables are related to each other through the field PLANTID, which is the tracking number of the plant. It has done the choice of selecting the main twenty plants for each region and not the main 100 in all of Denmark directly to find plants well distributed all over Denmark. The selected plants are reported in the appendix A. The second query filters the data out in the remaining tables considering only the presented plants and the following compounds:

The link between the tables DRWCHEMSAMPLES and DRWCHEMANALYSYS is based on the field SAMPLEID. It is necessary to manually insert the tracking numbers of the filtered plants in the design of the query. The outputs of the query are the date in which the sample is taken and the correspondent amount reported. They are organized in a pivot table in which all measurements are averaged per month. In addition, all data are grouped by plant and by year in rows, and they are grouped by compound in columns. It is important to notice that most of the cells are empty, because data are highly unevenly distributed in the time range from 1980 to 2012, as can be seen in figure 3.2. That table is the input for the Matlab program.

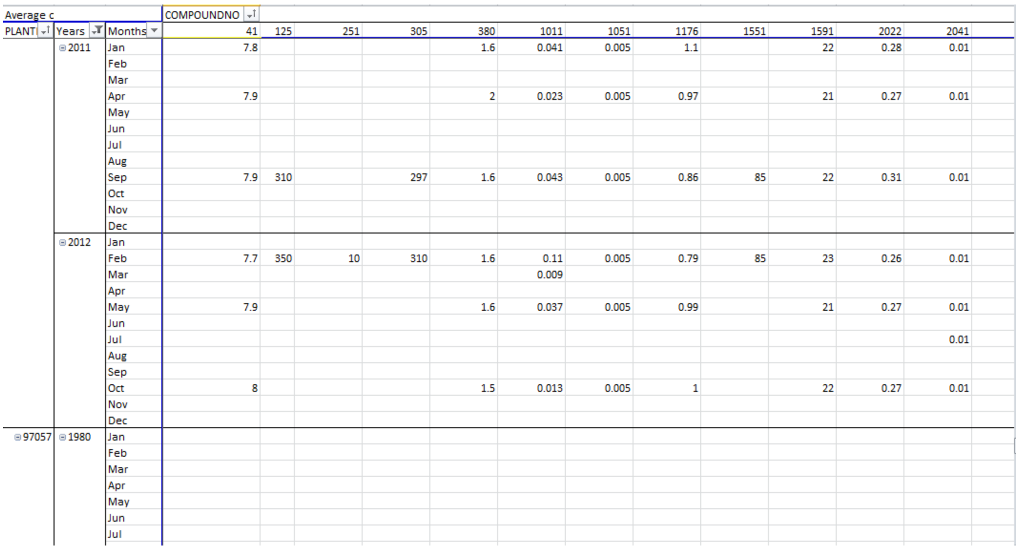
Table 3.2: The chosen compounds include most of the inorganic ones in drinking water, the Total Dissolved Solid, Total Organic Carbon, pH and temperature.

Name	**Symbol**	**Number**
Oxygen	O_2	251
Ammonium	NH_4	1011
Calcium	Ca^2+	1551
Iron	Fe	2041
Magnesium	Mg^{2+}	2081
Manganese	Mn	2086
Potassium	K^+	2056
Chloride	Cl^-	1591
Fluoride	F^-	2022
Hydrogen carbonate	HCO_3^-	305
Nitrite	NO_2	1051
Nitrate	NO_3^-	1176
Sulphate	SO_4^{2-}	2142
pH	-	41
TDS	-	125
NVOC	-	380
temperature	-	9902

To use the query as input for the program it is necessary to export it to Excel (in PivotTable Tools select Export to Excel) and save it in the same folder in which the Matlab file is stored. In the appendix D the implemented SQL statements of both the queries can be found with the instructions for possible adjustment to select different compounds or plants.

3.1.2 Output and final tables

The output of the first query is basically the plant tracking numbers to consider in the survey. The output of the second query is a pivot table which has in the first columns all plants' ID listed in ascending order, for each of them in the second column are the years and in the third the months. Thus there is a line referring to a single month and to a single year for each plant in which all compound measurements are stored and organized in columns depending on the number of the compound. In order to find an automatic procedure to implement the statistical analysis over all the twenty major plants for each region of Denmark, a Matlab code has been written. The structure of the program is easily suitable to different compounds and plants.



Average c			COMPOUNDNO											
PLANT	Years	Months	41	125	251	305	380	1011	1051	1176	1551	1591	2022	2041
	2011	Jan	7.8				1.6	0.041	0.005	1.1		22	0.28	0.01
		Feb												
		Mar												
		Apr	7.9				2	0.023	0.005	0.97		21	0.27	0.01
		May												
		Jun												
		Jul												
		Aug												
		Sep	7.9	310		297	1.6	0.043	0.005	0.86	85	22	0.31	0.01
		Oct												
		Nov												
		Dec												
	2012	Jan												
		Feb	7.7	350	10	310	1.6	0.11	0.005	0.79	85	23	0.26	0.01
		Mar						0.009						
		Apr												
		May	7.9				1.6	0.037	0.005	0.99		21	0.27	0.01
		Jun												
		Jul												0.01
		Aug												
		Sep												
		Oct	8				1.5	0.013	0.005	1		22	0.27	0.01
		Nov												
		Dec												
97057	1980	Jan												
		Feb												
		Mar												
		Apr												
		May												
		Jun												
		Jul												

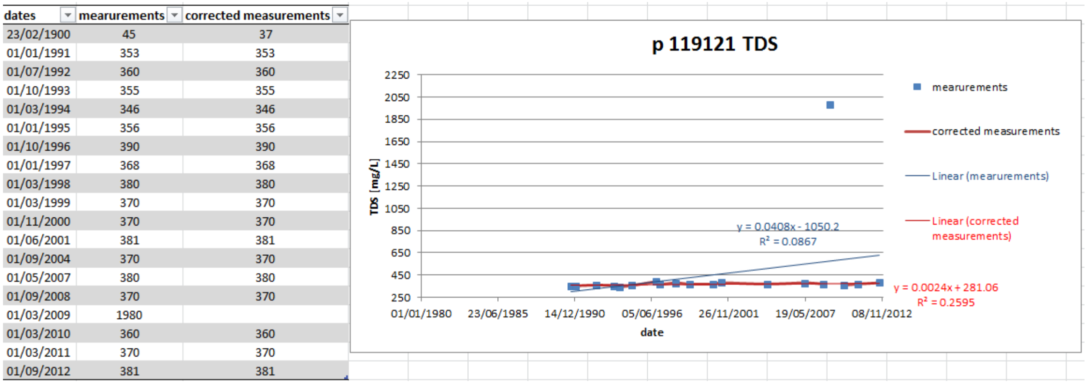
Figure 3.2: Example of a part of a pivot table gained from the final query in Access, as it appears in Excel after the format adjustment. It is important to notice that data are highly unevenly distributed over time.

3.1.3 Note about the plant selection

It is important to note in the selection of the plants, that plants 2065 and 104361 are included both in the Sjælland's and Hovedstaden's lists. Plant 104361 has the same mean yearly effluent, while plant 2065 has an effluent of 4.3 million m^3 per year if it is considered in Hovedstaden, instead of 1.7 in Sjælland. Although, 2065 has the same recorded value, 104361 has some different measurements. In particular, the latter region has more measurements than the first one during the recent years (2010-2012). Because of the previous difference, the choice to keep the plants in both the regions was chosen. This obviously affects the results, because the said plants influence two regions instead of one, but not with the same value.

3.1.4 Outliers

The analysis went through the slopes of the regressions of all compounds in all selected plants. Observing the relative slopes, some particularly noticeable trends are present. If the graphs with the trend and the measurement are analyzed the cause of the high slopes often is the presence of outliers. They can be a mistake in the transcription of the measure, an error in typing the decimal point or simply an atypical value of the compound. Apart from some cases in which the cause of the outlier is evident, for the rest of the time it is difficult to decide if the particular data should be neglected (deleting



dates	mearurements	corrected measurements
23/02/1900	45	37
01/01/1991	353	353
01/07/1992	360	360
01/10/1993	355	355
01/03/1994	346	346
01/01/1995	356	356
01/10/1996	390	390
01/01/1997	368	368
01/03/1998	380	380
01/03/1999	370	370
01/11/2000	370	370
01/06/2001	381	381
01/09/2004	370	370
01/05/2007	380	380
01/09/2008	370	370
01/03/2009	1980	
01/03/2010	360	360
01/03/2011	370	370
01/09/2012	381	381

FIGURE 3.3: Example of the linear regression of TDS in one plant: the outlier highly influences the slope. In a future analysis a criterion to filter out the outliers should be applied.

it from the dataset) or not. However, the high presence of outliers suggests the necessity of a more careful transcription of the measurement. As an example, the relative slope of TDS in plant 119121 (Syddanmark) is taken. The calculated regression is compared with the one based on the same dataset excluding the outlier (in red in the table). As the graph shows, the impact of the outlier on the linear regression is high because its value is about five times bigger than the average. Being that the average value of TDS in the plant 119121 is of 452.21 mg/L, the relative slope based on the original dataset is 32.98%, while not considering the outlier it is 2.34%. The latter slope is aligned with the TDS slopes for other plants. The effect of the outliers on the slope depends on how much the value departs from the average, which is difficult to predict. Furthermore, it also influences the average value, thus when the relative slope is calculated, two causes of errors are introduced. The point is that only one outlier can heavily influence the overall trend, but sometimes the imprecise slope is difficult to determine because it still assumes a plausible value. A high presence of outliers characterizes iron, and it is possible that the scatter in its values is caused not by measurement mistakes but by a chemical cause. In eventual further works based on this dataset, the use of a statistical method to handle the outliers is suggested. A possible criterion should be excluding data out of a range that is centered on the mean and a certain number of standard deviation wide.

Chapter 4

Method

4.1 Choosing the statistical method for testing trends in data

The implemented statistics leads to several results. First of all, the correlation between the monthly average levels of each compound with time, for each singular plant. Once the compounds with significant trends have been identified, the slope of the best fit linear trend is calculated using linear regression. The method also calculates the correspondent probability. That analysis will be implemented to each of the 17 inorganic compounds, to each of the main 20 selected plants from all 5 regions of Denmark; thus an automatic and easy repeatable procedure is needed. In fact, one of the intents is to implement a method that also fits well with different input tables, in order to find spatial variation of trends changing the selected plants (for instance, only the plants near the coast should be selected to note if there is a connection with the salinity of the water-bearing stratum). In order to obtain the previous results it has to be considered that the shape of data distribution is unknown, because it would be computationally too complicated to adjust each analysis based on the respective data histograms. Thus a non-parametric statistical method is chosen among the tests to identify trends in time series. As no information of distribution of the series is available, rank-based methods are a valid alternative to the correlograms. Furthermore, all rank-based tests keep the significance level better than parametric test, even under normality (Ferguson et al. [2011]). The two main rank-based tests are Spearman's rho and Kendall's tau, which are asymptotically equivalent, but in the quoted article a reason to prefer the latter one is given based on Monte Carlo simulations: Kendall's tau outperforms Spearman's rho in detecting first-order autoregressive dependence. Another reason is that Kendall's tau can be easily interpreted and it has the possibility to evaluate if a seasonal trend is present. In conclusion Kendall's tau is chosen as the statistical method.

4.2 Mann-Kendall analysis

4.2.1 Trend statistics for data unevenly distributed in time

The source database has data unevenly distributed in time, with lots of days (sometimes years) without any measurement but also many days containing more than one measure. That means that a kind of averaging is necessary to have a treatable data series. In this case, an average per month is chosen and it is located in time at the first day of each month. In *Testing for trends in data unevenly distributed in time* (Huth [1999]) a comparison can be found between seasonal means, means of individual events and daily events. In this article the authors suggest using the Monte Carlo test for calculating daily trends, because those values are strongly autocorrelated, but in the other cases the conventional test can be used. It is also noticed that there are some differences depending on the way of averaging is chosen, but there is no criteria to choose which one is better. Seasonal mean, individual event mean and daily values are all statistically and physically sound and represent the time effect on data from different perspectives. In this thesis, the monthly average is taken owing to the number of data: a daily mean would have had too many empty days, while a mean per year would have had a too large loss of information as a consequence.

4.2.2 Kendall's tau

Let (x_i, y_i) with $i = 1, 2 \ldots n$ be a set of observations from the respective variables X and Y. The Kendall rank correlation coefficient, named τ, is a measure of the monotonic association between the two variables. It is said a monotonic test because monotonic changes in X and Y (that are alterations in data that preserve the given order) don't change the value of tau. Kendall's tau is a non-parametric coefficient; that is, no assumption on the variable distribution is needed. That is important in this study because the shape of data distribution is unknown at the beginning. Kendall's tau is defined as:

$$\tau = (C - D)/(C(n, 2))$$

Where:

C is the number of the concordant pairs; (x_i, y_i) and (x_j, y_j) are a concordant pair if the ranks of both arguments agree ($x_i > x_j \wedge y_i > y_j$ or $x_i < x_j \wedge y_i < y_j$).

D is the number of the discordant pairs; (x_i, y_i) and (x_j, y_j) are a discordant pair if the ranks of both arguments disagree ($x_i > x_j \wedge y_i < y_j$ or $x_i < x_j \wedge y_i > y_j$).

$C(n, 2) = \frac{1}{2}n(n-1)$ is the number of all possible pair combinations in n elements of the set, that is the maximum value for $C - D$.

In the presents case the variable X represents the set of date serial numbers, so $x_i < x_j$ with $i < j$ always. The Kendall's tau has a range $[-1; 1]$ and can be interpreted as the following:

- If $\tau = 1$ it means that X and Y have the same rankings, so there is perfect positive correlation between them.
- If $\tau = -1$ it means that X has the opposite ranking from that of Y, so there is perfect negative correlation between them.
- If $\tau = 0$ it means that, at given level of significance, X and Y are statistically independent.

Thus, if the module of τ is close to 1, a trend of the analyzed compound in time is likely.

Another interpretation of Kendall's tau is: with the Kendall test one counts the number of different pairs between two ordered sets on the same set of objects, which is a distance between sets, known as symmetric difference distance. The notation corresponding to the symmetric difference distance is $d_\Delta(\wp_1, \wp_2)$ for two sets of ordered pairs $\wp_1$ and $\wp_2$. Considering the maximum number of pairs which can differ between two sets is $C(n, 2)$, the Kendall parameter measures how much the two set rankings differ from one another, thus $C(n, 2) - d_\Delta(\wp_1, \wp_2)$. In order to obtain a Kendall's coefficient that takes values between -1 and +1, the previous difference is normalized by $C(n, 2)$. Lastly, one gets $\tau = (C(n, 2) - d_\Delta(\wp_1, \wp_2))/C(n, 2) = (C - D)/(C(n, 2)) = S/(C(n, 2))$. So, τ can be interpreted as *"the difference between the probability of these objects to be in the same order [denoted P(same)] and the probability of the same objects being in a different order [denoted P(different)]"* Abdi [2007],in formula

$$\tau = P(same) - P(different)$$

4.2.3 Statistic hypothesis testing

The interpretation of the Kendall test is made through a statistic hypothesis testing. Its null hypothesis is $H_0 : \tau = 0$, which means the independence of Y from X (of the compound level from the date). The nominal level chosen is $\alpha = 5\%$, so the compound trend in time is neglected every time that the probability associated to the Kendall's tau is below 0.05. Since the distribution of τ asymptotically converges to a standard normal distribution, the τ-value can be normalized and the linked probability can be assessed. The theoretical background of the statistic method is presented in Appendix (E).

4.2.4 Accounting for ties

It is known that the coefficient is strongly affected from the presence of ties. A pair (x_i, y_i), (x_j, y_j) is said to be tied if $x_i = x_j$ or $y_i = y_j$. If there are no ties in the series and if the sample size is less than 50 the previous tau can be used, otherwise the Kendall's tau is corrected for the ties as follows:

$$\tau_B = \frac{(C - D)}{\sqrt{(C(n,2) - T)(C(n,2) - U)}}$$

Where:

τ_B is the value of the Kendall's tau adjusted to consider the ties.

$T = \sum \frac{t_i(t_i-1)}{2}$, being t_i the number of tied values in the i-th group for X. Given that the considered X is composed by unique values, this variable has no ties, thus $T = 0$.

$U = \sum \frac{u_j(u_j-1)}{2}$, being u_j the number of tied values in the $j-th$ group of Y.

With regard to the hypothesis test, the correction is applied to the z-value:

$$z_B = \frac{(C - D)}{\sqrt{v}}$$

And:

- $v = (v_0 - v_t - v_u)/(18 + v_1 + v_2)$
- $v_0 = n(n-1)(2n-5)$
- $v_t = \sum \frac{t_i(t_i-1)}{2t_i+5}$ as the first sum operator is null.
- $v_u = \sum \frac{u_i(u_i-1)}{2u_i+5}$
- $v_1 = \frac{\sum t_i(t_i-1) \sum u_j(u_j-1)(u_j-2)}{2n(2n-1)}$ equal to zero in this case, for the same reason as above.
- $v_2 = \frac{\sum t_i(t_i-1)(t_i-2) \sum u_j(u_j-1)(u_j-2)}{9n(n-1)(n-2)}$

Given z_B, the correspondent p-value can be calculated. Hypothesizing the test statistic is really distributed as it would be under the null hypothesis, the p-value is the probability of observing a test statistic as extreme as or more extreme than the one actually observed. In the code, the statistical adjustment for the ties was implemented.

4.3 Linear regression

In order to find the several trends for all different compounds a least-square regression is used. The reasons of this choice are that it is the simplest and most used, furthermore it has an evident interpretation. Moreover the correspondent p-value is calculated to evaluate if the range of data effectively follows a linear trend or not. (Also the R^2 is calculated, just to make a comparison with a descriptive parameter and check the founded p-value). The difference between the p-value referred to the Kendall's test and the p-value of the linear regression line is that the former expresses if a significant trend exists but it gives no information about the kind of relationship between the two variables, while the latter says if the possible trend is likely linear or not. All the founded slopes are multiplied by 10 years to find the variation of one compound unit in a decade. Only non-null values were considered in the regression, simply ignoring the missing data.

Chapter 5

Implemented program

5.1 Logic of the program

The code is composed of a main program and two functions: one calculates the Kendall's tau and the relative p-value and the other joins all the tables containing the averages over the years of the five regions in one table, used afterwards as input for GIS. The code should be run in parts as explained later on. It is written in such a way that it fits data even if the number of plants or the compounds analyzed changes as well as the order in which the plants' identity number appear in the different tables. The main folder should contain five folders, one for each region, and the file *Union.m*. Each folder referring to a region has to include:

- A folder named *GISinput* in which all the outputs used to build the maps are stored.
- The main program *time_trend.m*, modified as described later on, depending on the name of the region.
- The function *Kendf.m*, which calculates the statistical parameters.
- The tables *PivotTable_[region_name].xlsx* and *pl_selection_[region_name].xlsx* gained as described previously.
- A text file *data.txt*, having the name of the region in the first row, the number of the selected plants (20) in the second row, the data range of the pivot table in the worksheet in the third row and the names of the selected compounds in the following rows.

The program *time_trend.m* is made up of different parts separated by breakpoints:

- In the first part there is the reading of the pivot table given as input, the initialization of all variables, the definition of the level of significance of the Kendall test and the count of how many years of measurements are included for each plant.

- In the second part, the main part, two "for" loops are used to consider all the ten plants and all the 17 compounds for each of them. They extract two columns of data: one within the not-NaN measurements and the other with the relative dates; the results of Kendall's statistics in terms of tau and p-value, the slopes of the linear regression with the correspondent p-value (and the r-squared); the maximum and minimum values for each column of measurements; and the figures showing all those results (one for each plants within as many graphs as the number of compounds).

- In the third part the path to save the outputs is defined (inside *GISinput* folder); then some text files are written:

  - **Q_pumped.txt**: made from the table *pl_selection_[region_name].xlsx* within the coordinates of the plants of the given region and the averages of the effluent discharges over the years;
  - **Avercoord.txt**: has in the first two columns the coordinates, in the third the plant tracking numbers and in the remaining columns the averages of all compounds over time, followed by hardness, alkalinity, Larson Index, taste index and DMF-S as from the model.
  - **Absslopes.txt**: contains the plants' coordinates and all the slopes calculated through the linear regression for each compound and referred as a period of 10 years.
  - **Relslopes.txt**: the same data of the previous file are written but all the slopes are divided by the average of the correspondent compound and plant.

  Furthermore, a table is stored in the workspace as extremesl, and it has the compound number in the first column, its minimum and maximum values reported among the plants, and its average over all plants and all years.

- In the fourth part all histograms showing the data distributions are drawn in a figure and all the normal probability plots are collocated in another figure. Those two images refer to one single plant, so this part of the code should be run separately for each plant just changing the index j in the range from 1 to 20 (if twenty are the plants to analyse). This part is only a reference to have a broad overview of the distributions of data.

It is important to notice that the Kendall's tau (and consequently the related p-value) can be a not-a-number-value when a dataset consists only of 1 or 0 measurements or because $C(n, 2) < T$. Similarly, the slope of the linear regression is null in the following cases:

- When the p-vale and the Kendall's tau are NaN-values (there are one or no data at all or only one, thus no linear regression can be calculated).
- When the observed test statistic is outside the critical region ($p - value > \alpha$), i.e. the trend is not statistically significant (as from the Kendall test result) so the calculated slope can be considered as null.

The probability that the calculated slope of the linear regression corresponds to the real one is evaluated through the p-value relating to the first coefficient of the linear regression. The t tests are used to conduct hypothesis tests on the regression coefficients obtained in simple linear regression. Assuming that the desired significance level is 0.5, since $p - value < 0.5$, $H_0 : \beta_1 = 0$ is rejected indicating that a relation exists between time and compound level. The lower its value the stronger is the refusal of the null hypothesis. Using this result along with the scatter plot, drawn running Matlab code, it can be concluded that the relationship between time and compound level is linear.

5.2 Kendall function

The Kendall test is executed through the file *kendf.m*, that because otherwise an excessive amount of variables would be created in the workspace. Thus, in order to explain what happens running just one line in the main code, the single commands are justified step by step as follows:
An if-statement is inserted to evaluate if the array of measurement A has at least two components; if not, NaN is given as output for $p - value_b$ and *taub*.

```
if length(A)>1
```

N is the number of measurements and comb is $C(n, 2)$ which is all possible pairs in n elements.

```
N = length( A );
comb = N*( N-1 )/2;
U = zeros(1,N);
```

```
S = U;
ties = U;
vucomp = U;
```

First for-loop which considers all the N measurements, apart from the last, to calculate:

- the difference (S) between the concordant (C) and discordant (D) pairs of the array A.
- the total number of ties in the array A ($Tietot$).
- U_{tot} that is $U = \sum \frac{u_j(u_j-1)}{2}$

```
for j = 1:(N-1)
    sumi = zeros( 1,N );
    ui=0;
```

The second for-loop takes all rows following the i-th one and compares that value with the i-th cell value. So s before computes the signs of the differences of j-th and i-th elements both for A and B then it makes the product. Dates are always in ascendant order, thus $sign(B(j) - B(i))$ is always negative while the other sign determines if the two pairs are concordant or discordant. For example, if all pairs will be concordant for all j values it means that the ranks of the two arrays are the same and so there is a monotonic relation between them.

```
for i=(j+1):N
        s=sign(A(j)-A(i))*sign(B(j)-B(i));
        if abs(A(i)-A(j))<1e-10
            ui=ui+1;
        end
```

sumi gives as i-th value 1 if the i-th A-pair is concordant, -1 if it is discordant and 0 if it is a tie.

```
sumi(i)=s;
   end
   ties(j)=ui;
   prod=ui*(ui-1);
   S(j)=sum(sumi(:));
```

```
   U(j)=prod/2;
   vucomp(j)=ui*(ui-1)/(2*ui+5);
end
Tietot=sum(ties);
Utot=sum(U);
Stot=sum(S);
```

The Kendall statistic is adjusted for the ties, which implies that $\sqrt{C(n,2)(C(n,2)-U)}$ has to be calculated. Thus if the number of ties is as high as to make U greater than all possible combinations, the statistics returns NaN values. If the computation of the statistic is possible tau_b, v, z_b than the p-value has the expression given in the previous pages.

```
v0=N*(N-1)*(2*N+5);
if comb > Utot
    taub=Stot/realsqrt((comb-Utot)*comb);
    vu=sum(vucomp);
    v=(v0-vu)/18;
    zb=Stot/sqrt(v);
    p_value_b=normcdf(-abs(zb))*2;
else
    p_value_b=NaN;
    taub=NaN;
end
else
        p_value_b=NaN;
        taub=NaN;
end
end
```

The working of the program was checked replying the same calculations in Excel for some of the arrays. An example with a small dataset $A = [20; 23; 24; 23; \ldots]$: if one takes i-index=2 with the same data $A(i) = 23$, so all the following elements will be compared with it in the i-th loop for $j > 2$

$$j = 3, A(j) = 24, sig((A(j) - A(i)) * -1 = s = -1$$
$$j = 4, A(j) = 23, sig((A(j) - A(i)) * -1 = s = 0$$
$$\ldots$$

5.3 Considerations about the output

The first relevant matrix as output of the program is P, which contains the number of rows correspondent at each plant (given as the identity number). That table is used by the program to read the correct number of rows, including the empty rows, in the *input_table* and it can be useful to check if the input reading is correct (if any plant is reduced as one row in the pivot table in Excel, it could be seen from here). Each plant has 396 rows because there are 12 rows (one for each month) for all the 33 analyzed years (from 1980 to 2012 included). Another matrix useful to check is MAT, in which all extracted and analyzed arrays without empty cells are placed and each pair of columns correspond to a compound: the first one contains the serial number dates and the second has the measurements. Actually, MAT corresponds to only one plant, so if the n-plant has to be checked, only the n-th loop (q=n) has to be run. The table is not reported here due to space limitations.

Chapter 6

Results

This chapter presents the results of the analysis. It firstly includes the presentation of the basic data on the investigated water samples (represented by water volumes and the number of samples considered in relation to each compound and each region). Then follows a summary of the identified trends and average conditions of the drinking water quality. Finally I present indixes related to corrosion, health and taste, and I show the geographic distribution of water quality parameters.

6.1 Basic data from the datasets

The selected waterworks are 100, thus only a small part of the total number of waterworks (about 2500). Nevertheless, the quantity of water supplied by those plants is an important part of the total because of their size. The aim is to get an overview of how much supplied water the analysis represents in relation to the total, how it is subdivided into the five regions, and if there is a relationship between the effluent and the number of measurements taken.

The effluent of the plants taken as a whole is 168 million of m^3 per year on average during the 32 years of measurements. It represents only the treated water leaving the waterworks, not considering water directly used without any treatment. The water is supplied mainly for domestic purposes and only a small part for agricultural and industrial purposes.

The size of the plant selected changes region by region because the criteria of taking the largest 20 plants is applied to each region. Consequently, even if the smallest plant has 0.42 million of m^3 per year (in Nordjylland), it does not mean that all plants considered in all Denmark which have outgoing discharge higher than that one are studied, but only in Region Nordjylland.

To give an order of magnitude, in total Denmark is supplied meanly with about 300 million m^3 per year for domestic use. Considering that most of selected water works supply water for domestic use, the effluent considered represent about half of the total distribution in Denmark, despite the fact that the selected water works are only around 1/25 of the total Danish plants. It is difficult to find the average effluent from all the water works over the years 1980-2012, thus a precise percentage is not given.

The 20 plants selected in the Region of København (Hovedstaden) represent the largest amount of water (around 40 million m^3/year), followed by Midtjylland (38.9 million m^3/year), Sjælland (35.5 million m^3/year), Syddanmark (31 million m^3/year) and Nordjylland (22.2 million m^3/year). All the selected plants treat an amount of water which varies from 0.5 up to 8 million m^3/year, and particular ranges for each region can be seen from the figure (Figure 6.1). The lower limit of the range is the one that was imposed in the query to get the largest 20 plants for each region, consequently it differs region by region.

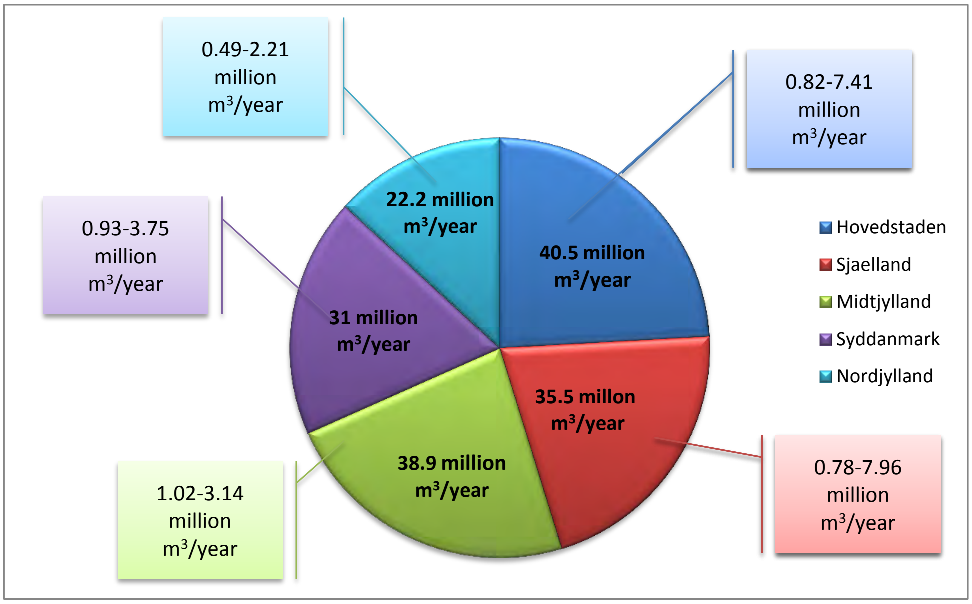


Figure 6.1: Total effluent of all water works considered in the analysis (168 million m^3 per year) divided into the five regions. The text boxes include the ranges of effluent for each region.

According to Danish law requirements, the drinking water quality has to be monitored when leaving the water works and at regular intervals. Based on the law requirements a minimum number of required analyses can be computed according to the produced water volume (see Ikrafttr [2014]):

$$s = 32(11n + 15e)$$

Where:

s is the total number of expected samples

n is the number of samples to take in the normal controls per year according to the effluent discharge (see Table 3.1). It is multiplied by 11 which is the number of compounds to measure in a normal control (temperature, pH, NVOC, ammonium, iron, manganese, chloride, fluoride, nitrate, nitrite and sulphate).

e is the number of samples to take in the extended controls per year according to the effluent discharge (see Table 3.1). It is multiplied by 15, which corresponds to the number of compounds to measure in an extended control (TDS, calcium, magnesium, hydrogen carbonate in addition to the previous one).

32 is the number of the selected years: from 1980 to 2012.

The number of analyses recorded in the Jupiter database is compared with the minimum requirements (see Figure 6.2). To do this comparison not only the monthly mean measurements are considered, but the total number of measurements taken.

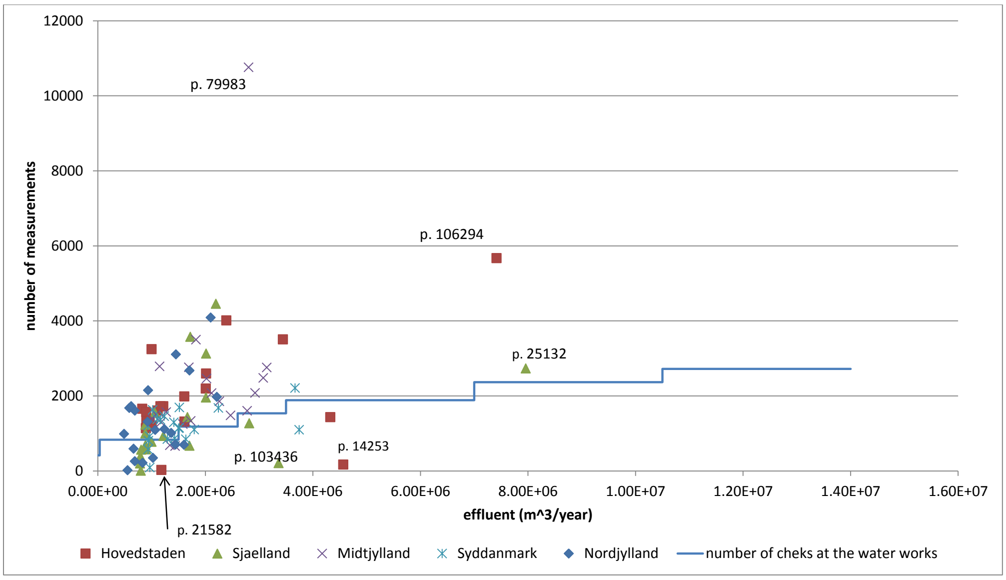


FIGURE 6.2: Number of measurements for all 100 plants in relation to the effluent. The blue line represents the minimum number of measurements required by the Danish law (Ikrafttr [2014]), depending on the discharge (see Table 3.1).

A relationship between the number of measurements and the treated discharge is expected (see the law requirements in 2.1). The water work 79983 (Hvinningdal Vandværk) presents a high number of measurements owing to some non-compliances of some compounds with the law requirements. Despite that the plant Sjælso Vandværk (106294) is second for number of samples taken and also for water supplied, the two variables are not strictly positively related. Plants 14253 and 103436 have less than 200 months of measurement, despite the average effluent being 3 million m^3/year. In fact, checking the list of the samples taken, plant Sødersø Øst (14253) has data only from 1980 to 1991 and they have not been approved by the municipality. Interestingly, plant Kalundborg Vandforsyning A/S (103436) has recorded measurements only from 1992 (which were approved by the Sønderso municipality only in 2007).

Other plants have even fewer measurements but they treat less water. For example, plant Søndersø Vest (21582) has only four samples, one has been taken in 1991 and the others in 1994, but all of them have been approved in 2007. In fact, the total count of measurements is 26, and the compounds analyzed are those that are mandatory for the extended controls only in one sample and in other one only organic micro-pollutants have been measured. Other water works have a number of measurements below the blue line: sometimes there are some years missing, the compounds analyzed may not all be requested or some data may not be recorded in the Jupiter database because they have not been approved yet. Another possible cause is the merging or splitting of water works, which is mirrored in the database as missing years because the plant changed the tracking number, but it has not been possible to gain that information within this project.

In order to see if there are some differences in the number of samples between the different regions or the compounds, the sum of the measurement of each compound is calculated, region by region (see Figure 6.3).
Most investigated compounds have a higher number of measurements in Hovedstaden and Midtjylland, as expected, because the overall effluents of these regions are the largest. However, the number of measurements does not differ a lot from one region to another, apart from temperature and iron, which have more recorded values in Midtjylland and Hovedstaden respectively.

Summing up:

- The evaluated drinking water discharge is about 168 million of m^3. It represents approximately almost half of the total.
- The selected dataset does not always have the expected number of measurements. Even though a rough filter of datasets with only 1 or 2 measurements was led,

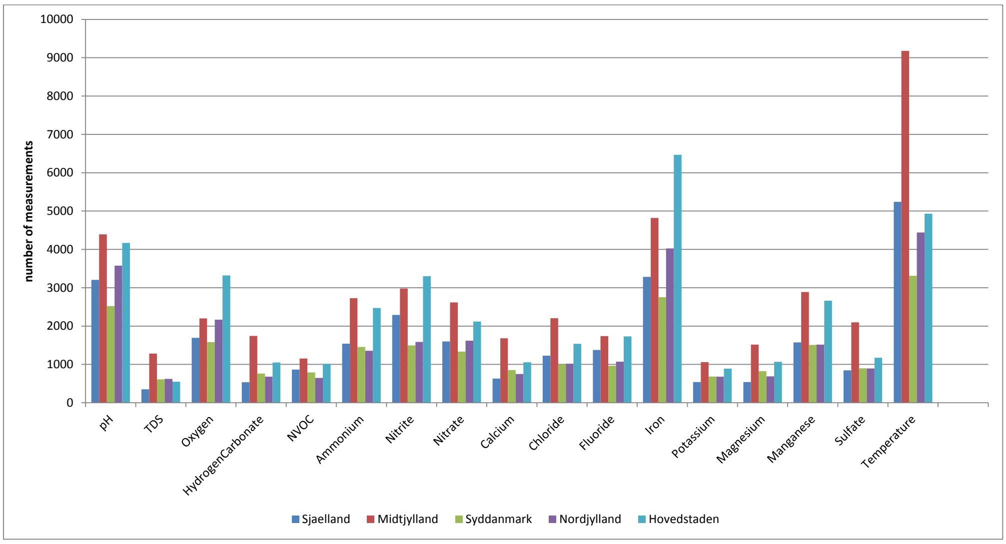


FIGURE 6.3: Number of measurements for all 100 plants divided per compound and region.

a more careful selection is suggested in future developments. For example, only datasets with at least the number of measurements equal to the number of years should be chosen.

- The selected samples contain a considerable number of outliers, thus a criterion to neglect the outlier would lead to more reliable results. Despite of that, the entire set of measurement considered in this work counts about 1700 samples with varying size and it covers a long time period. It must be considered one of the largest and in general most reliable databases covering drinking water quality. The reason for that is Jupiter's data is checked from the municipalities and data quality is documented to some degree; furthermore it was established via paper more than 100 years ago (Refsgaard et al. [2010]).

6.2 Presentation of the averages

A first approach to the database was led through the analysis of the average values over time for all 17 compounds and 100 plants. The aim of this paragraph is to point out the most interesting compounds to show in detail among all the 17 analyzed. Some consideration on the gained averages in relation to the mean levels given in literature are presented.

In appendix C the tables having the averages over time for all selected compounds and plants are reported. Comparing the averages made for all Denmark all over the thirty years with the ones given by Bruvo (Bruvo et al. [2008]), the bigger differences are noticeable in ammonium, iron and nitrite. It has to be considered that Bruvo analyzed all Danish municipalities (275) with a different size of the water works and a more narrow time period (1995-2004). In fact, the averages given by Bruvo are still in the ranges of the values calculated for the averages. Those ranges are quite wide, because the water works have mean compound levels that vary a lot from one to another (see Table 6.1).

From the tables reported in appendix C, the compounds which have more average values beyond those prescribed by law are oxygen, ammonium, iron and manganese, and most of them are concentrated in Midtjylland plants. For ammonium, iron and manganese, Midtjylland is the region with the largest number of unacceptable averages (respectively 11, 15 and 14) and approximately in the same plants. In particular, ammonium has mean values which cross the limit of 0.05 mg/L up to almost 450%.

In fact, some non-compliance values for iron and manganese were already pointed out in a survey in 2002 led by the Danish Environmental Protection Agency (Danish EPA). The Danish EPA implemented the project Investigations of water treatment methods at selected Danish waterworks (in brief The Water Treatment Project), which investigated water treatment at 30 smaller waterworks and gave possible explanations for the waterworks problems in complying with drinking water requirements. The project also provided recommendations for improving water treatment (Landskabsstyrelsen [2009]).

It has to be noticed that a non-compliance in average does not imply a non-compliance during the singular years, because during the almost thirty years of measurements the allowed limit by the law sometimes changed. The introduction of a new requirement or a change in treatment operation is reflected in a change in the compound trend, which is not represented by the average. In conclusion, ammonium, iron and nitrite are the compounds whose averages differ most from the values given in literature. Moreover, oxygen, ammonium, iron and manganese are the compounds with mean values over time higher than the current limit given by the Danish law. Oxygen is not presented in the section 6.4 because the Danish law gives only a guideline and not a limit for it.

6.3 Temporal trends

The aim of this paragraph is to individuate the compounds and plants which have more interesting trends to discuss in the Section 6.4.

Table 6.1: Extreme values of the averages for each compound among all the selected plants. The mean value all over Denmark and the mean values given by Bruvo (Bruvo et al. [2008]) are also reported. The difference refers to the ratio between the difference of the two averages divided by the average all over Denmark. The compounds having the bigger differences in the averages are marked in bold.

Compound	**Minimum average**	**Maximum average**	**Average all over Denmark**	**average in Bruvo et al. [2008]**	**difference in %**
pH	7.29	8.09	7.7	7.69	0
TDS (mg/L)	173.8	923.33	406.4	396	3
Oxygen (mg/L)	0	11.08	8.72	8.8	1
Hydrogen Carbonate (mg/L)	69.04	453	268.82	263	2
NVOC (mg/L)	0.35	19.44	1.84	1.7	8
Ammonium (mg/L)	0	0.22	**0.04**	**0.07**	**75**
Nitrite (mg/L)	0	0.12	**0.01**	**0.03**	**200**
Nitrate (mg/L)	0.02	36.22	5.08	5.7	12
Calcium (mg/L)	35.06	154.05	84.04	83.5	1
Chloride (mg/L)	15.82	670	59.63	47.8	20
Fluoride (mg/L)	0.08	1.29	0.32	0.33	3
Iron (mg/L)	0.01	5.29	**0.19**	**0.08**	**58**
Potassium (mg/L)	0.88	11.78	3.46	3.35	3
Magnesium (mg/L)	2.45	69	13.51	11.9	12
Manganese (mg/L)	0	1.04	0.03	0.03	0
Sulfate (mg/L)	1.43	414.5	46.01	43.3	6
Temperature (°C)	7.78	12.25	9.82	-	-

6.3.1 Significant trends

The first step was to filter the data series with a statistically significant correlation with the date of the measurement, i.e. only the measurements which present a significant trend in time evaluated by the Kendalls test. The results of the Kendall test for the selected plants and for all the compounds are reported in the appendix B. The chosen level of significance for the hypothesis testing was 5% (see Tables B.1,B.2,B.3,B.4,B.5, appendix B). The regions which have more significant trends among all the main 20 plants and 17 compounds are Hovedstaden and Midtjylland, respectively they have 39% and 43% of significant trends compared to the total number of samples for each region ($17 * 20 = 340$). When considering the individual plants, Bagterp Vandværk (71361, Nordjylland), Greve Vandværk (104361, Hovedstaden and Sjælland) and SjælsøVandværk (106294, Hovedstaden) have the highest number of compounds with a significant trend over time.

If the result of the Kendall test is 2, the related compound at the selected plant can be considered steady over time. Consequently, plant Grindsted Vandværk 3 (51183, Syddanmark) can be considered stationary in all compound levels during the time studied. Plant Østvendsyssel Råvandsforsyning (71977, Nordjylland) does not have enough data for any of the selected compound to calculate the statistic. Those two plants are the only ones which do not have at least one significant trend among the 17 compounds.

6.3.2 Evaluating the relative importance of trends

In this paragraph the relationship between the number of significant trends for each compound is discussed along with the discharge outgoing from the correspondents plants. It is useful to understand how much water is affected by not stationary levels in the selected compound.

The linear regression was calculated for each sample to begin with. The slope of the linear regression was multiplied by the number of days in ten years, getting the compound variation in a decade (named absolute slope later on). Although the absolute slopes of different compounds cannot be compared to each other because the unit of measurement or the measurement ranges can differ from one to another.

In order to compare the trends over time of all the 17 compounds and select which ones are most interesting to analyze, the relative slopes are computed. Relative slope is defined as the linear change found by regression per 10 years, divided by the average of the measurements of the given compound at the correspondent plant.

Firstly, the number of significant trends is counted for each compound. After that, also the effluent discharge from each plant has been considered. The sum of the overall effluent from Danish waterworks, averaged over the almost 30 years, is about 168 million of m^3 of drinking water per year. For each compound, all the discharges from the plants which have significant trends are summed up, and divided for the total effluent. In that way it is possible to evaluate the impact of each compound in terms of amount of water supplied. The following table summarizes the results:

Table 6.2: Number of significant trends for each compound. For each compound, the effluent outgoing from the plants with significant trend of the respective compound is reported

Compound	**Significant trends**	**Effluent [million m^3per year]**
pH	57	108.9
TDS	36	71.8
Oxygen	59	105.0
Bicarbonate	41	80.9
NVOC	29	50.2
Ammonium	31	63.5
Nitrite	11	21.4
Nitrate	55	94.8
Calcium	50	82.9
Chloride	54	99.1
Fluoride	45	86.8
Iron	37	56.9
Potassium	42	83.1
Magnesium	43	74.9
Manganese	13	19.7
Sulfate	58	104.9
Temperature	39	72.8

The compounds which have more relevant trends are: pH, oxygen, sulfate, chloride, and nitrate ordered by discharge or oxygen, sulfate, pH, nitrate and chloride if they are ordered by number of significant trends. Looking at the whole table having all the relative slopes, it can be noticed:

- The compounds with the highest slopes are ammonium, nitrate and iron. Iron has significant trends in waterworks which supply in total 56.9 million m^3/year, ammonium influences a total effluent of 63.5 million m^3/year and nitrate 94.8 million m^3/year. Nitrate is not presented in detail in the section 6.4 because its levels are largely under the limit of 50 mg/L.
- The plant Kalundborg Overfladevandforsyning A/S. Tissø (number 103436) in Sjælland has a strong positive slope in manganese values, 1093% in 10 years. Likely this is due to the few data points that are available and to the presence of an outlier. Not considering the outlier the slope is 408.3% in ten years and the mean value is 0.050 mg/L instead of 1.04 mg/L. Despite the correction, it remains one of the highest relative slopes.

- The maximum positive slope is in plant Holbæk Vand A/S (Knabstrup Enge) (103982, Sjælland) for ammonium (54503% in 10 years), which has in general the largest amount of high relative slopes.
- The maximum negative slope is in plant Skindermarkens Vandværk (53083, Syddanmark) for sulfate (-167.9% in 10 years).

6.4 Trends over time and averages of a selection of compounds

Some compounds are selected to carry a deeper analysis of their trend over time. The selected compounds and the reason for their choice are the following:

- **Ammonium**: ammonium in drinking water originates from metabolic, agricultural and industrial processes and is often an important indicator of pollution (Gorchev and Ozolins [1984], water quality management [2011]). Furthermore, some average values of this compound are often out of the Danish law requirements and it is one of the compounds with the higher number of significant slopes. Considering ammonium, nitrate and nitrite result as a consequence of nitrification of ammonia during the treatment process, but they were checked as all the other compounds.
- **Chloride**: chloride is a conservative compound in water, thus it influences the water supply system to a high degreee. The presence of chloride is an effect of waste dumps, the salt spreading on the streets during winter time, seawater intrusion and also of the raise of the salty cone into the freshwater lens (DELTA Laboratories [2008]).
- **NVOC**: it is interesting to analyze because it sums up the overall content of organic compounds. However, its average value only exceeds the guideline value in one plant (Bagsværd vv, 106311).
- **pH**: pH is the compound whose trends influence the drinking water most in terms of discharge. In other words, the sum of all effluents from plants which have significant trend in pH is the biggest in comparison with the other analyzed compounds.
- **Iron**: iron is another compound with many significant trends and with lots of averages which seem not to comply with the requirements.

- **Nitrite**: nitrite has several trends with the same pattern of ammonium. It has average levels which differ from those given by Bruvo. It is one toxic element resulting from the pesticides in agriculture.
- **Manganese**: manganese has high average values in several plants.

It has to be mentioned that to compare the same compound in different plants the absolute slope (change in compound unit over 10 years) has to be considered. That is because it is not influenced by the average value. In fact, taking two samples with the same absolute slope but with different averages, the sample with the lower average has higher relative slope (due to how the relative slope is calculated). But actually the sample with the bigger average value has to be taken in higher consideration. This risk of misleading the trend comparing the same compound in different plants is avoided looking at the absolute slope and the average value separately, in addiction to the graph with the measurements.

6.4.1 Ammonium

The limit for ammonium given by the Danish law is 0.05 mg/L. Midtjylland has 11 water works whose ammonium average is above the limit. Hovedstaden has lots of plants near the threshold and Nordjylland has the highest mean values over time of ammonium. Syddanmark has all ammonium averages which comply with the requirements on ammonium (see Figure 6.4). Looking at the complete series of measurements, ammonium level is generally around 0.01 mg/L until the end of the 1990s for most plants of Syddanmark, and then data start scattering. From the map (see Figure 6.5) high levels of ammonium can be noticed around Horsens and Aarhus. The regions which have more significant trends in ammonium are Midtjylland and Sjælland; Midtjylland has all negative trends in ammonium. Furthermore, the few plants which go over the limit have weak slopes. The graphs with the higher positive trends are reported in the following paragraphs.

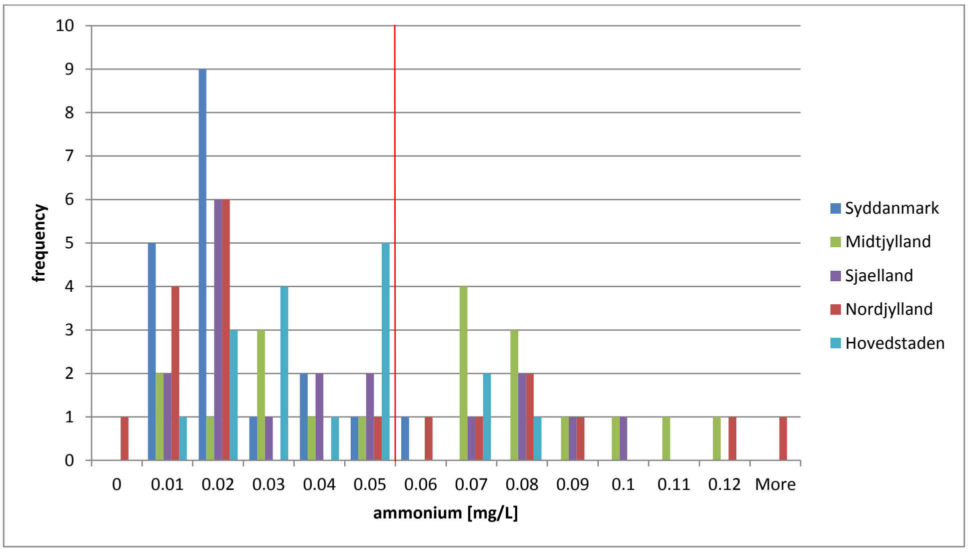


FIGURE 6.4: Histograms of ammonium for all the selected plants divided by region. The red line represent the ammonium limit given by the Danish law (0.05 mg/L). The region with more water works having ammonium averages above 0.05 mg/L is Midtjylland.

In the Sjælland plants in Tissø (Kalundborg, 103436) and Holbæk Vand A/S (Knastrub Enge)(103982), data are all concentrated in a few years, although a trend is present it could be hasty to extend it for many years. The plant in Tissø (Kalundborg, 103436) is the water works with the highest positive absolute slope: its variation is 0.32 mg/L with average value of 0.095 mg/l, but data are available only from 2002 to 2007. The ammonium average in Holbæk water works is 0.044 mg/L, but the decrease computed

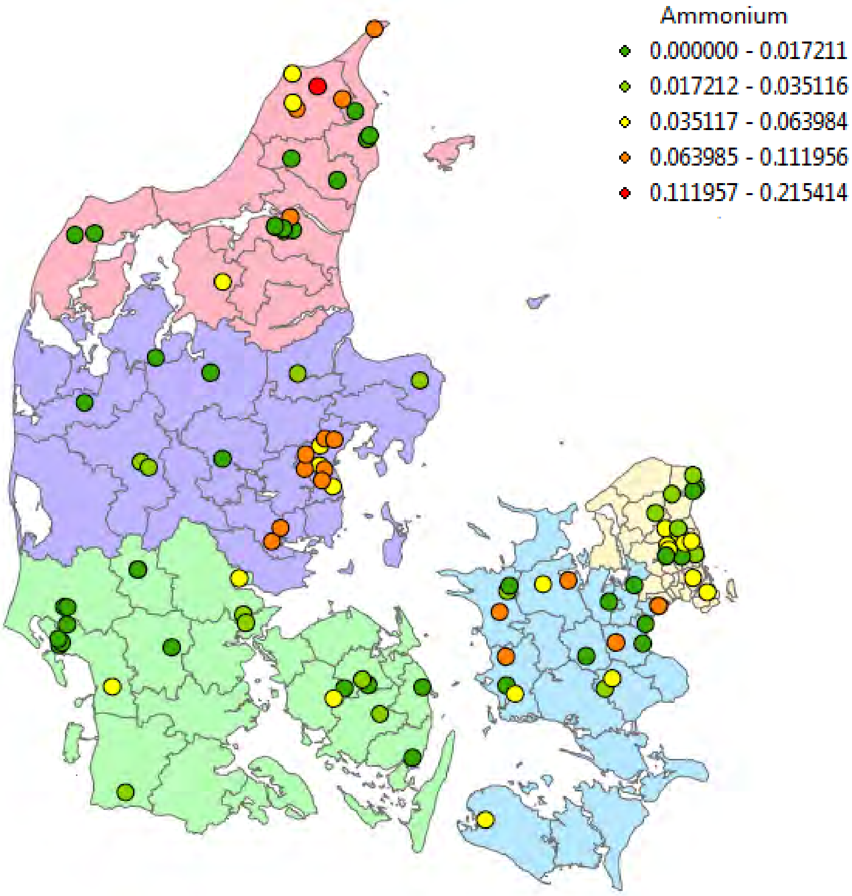
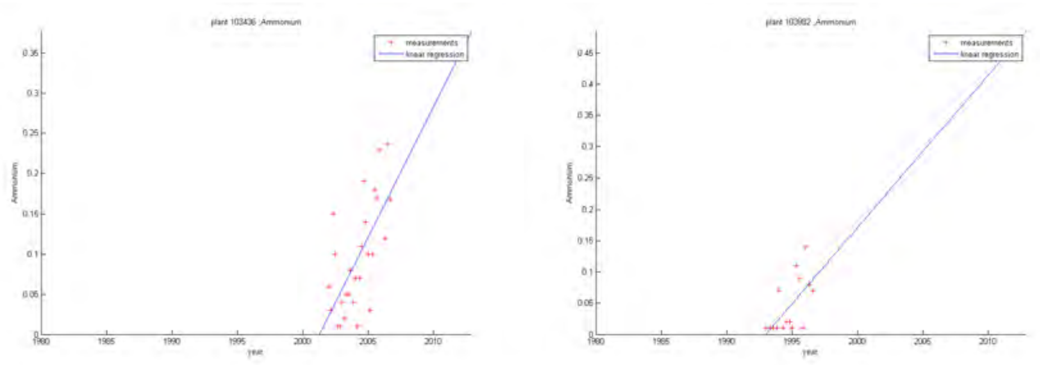


Figure 6.5: Map of ammonium averages espressed in mg/L. The region with more high averages in ammonium is Midtjylland, in fact a concentration of ammonium levels above 0.05 mg/L is present near Aarhus.

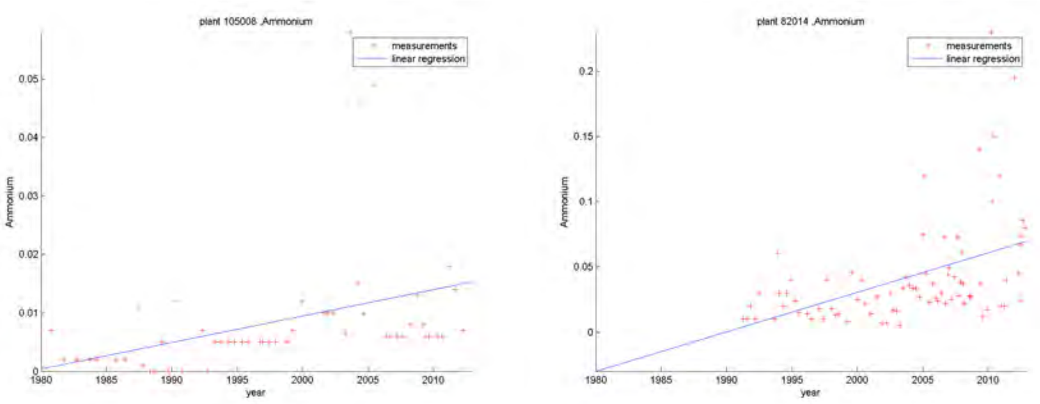
(0.243 mg/L in 10 years) is based only on about 4-5 years. The other plants have a relative trend in 10 years that is quite flat. In general, most values are below 0.02 mg/L (see Figure 6.6).

Considering the plant in Solrød (105008, Sjælland), the measurements are spreading all over the 30 years, the increase is slow (0.005 mg/L in 10 years) and the values reach maximum 0.05 mg/L. The other plants in Sjælland have almost the same trend as Solrød. In Syddanmark, the plant in Odense (82014) has an evident increase of ammonium over time and it likely has a nonlinear dependence. In fact, despite that the absolute slope is 0.030 mg/L over 10 years, ammonium values vary about up to 0.2 mg/L. Only four on the 20 plants in this region have a trend in ammonium.



(a) Kalundborg Overfladevandforsyning A/S. Tissø(103436), Sjælland

(b) Holbæk Vand A/S (Knastrub Enge, 103982), Sjælland



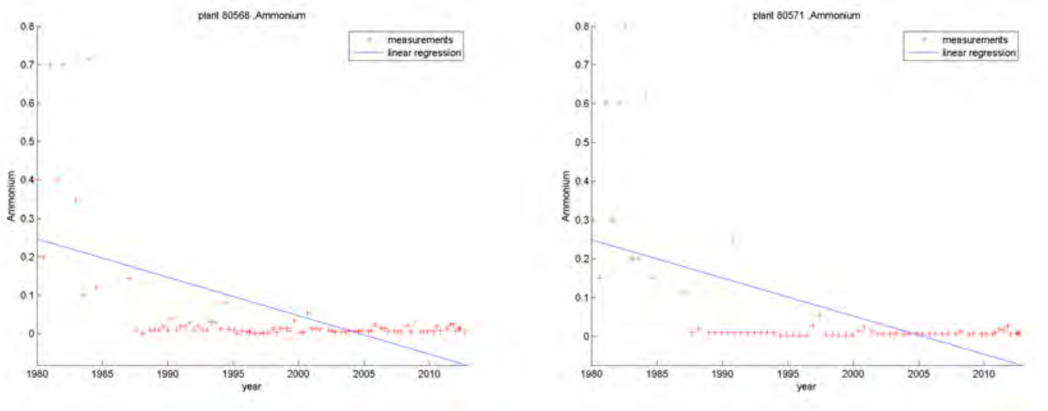
(c) Solrød Vandværk a.m.b.a. (105008), Sjælland

(d) Odense Vandselskab A/S: Holmehaveværket (82014), Syddanmark

Figure 6.6: Measurements of ammonium in mg/L for three water works in Sjælland and one in Syddanmark

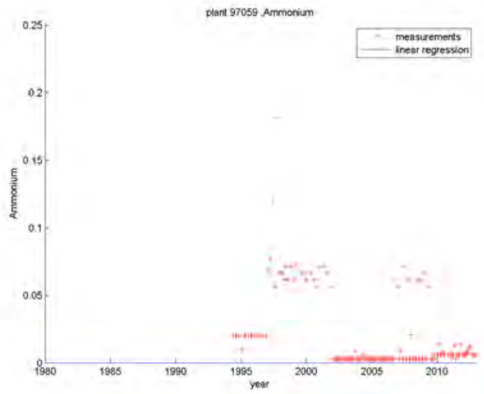
For Midtjylland all the plants with significant trend (8) show descendent values in ammonium. In particular, there is a step in data around 1987-88, in which the level of ammonia become nearly null (as for Bederværket and Lyngbyværket plants). Moreover, plant Østre Vandværk (97059) shows values above 0.05 mg/L during the periods 1997-2002 and 2007-2009 (see Figure 6.7).

In Hovedstaden, plant Hellebæk Vandværk (83344) has lots of values above 0.05 mg/L and a particular increase during the last three years, and plant Frederiksgade Vandværk (83381) has a big increase after 1995. Both of them have no significant trend from the statistic test likely because the trend changes in time (see figure 6.8). In Syddanmark, most of plants have no trend in ammonium, they keep the ammonium level around 0.01 mg/L until about the 1990s and then their ammonium values start scattering. There is only one positive in plant Odense Vandselskab A/S: Holmehavaværket (82014) with levels up to 0.23 mg/L (see Figure 6.9).



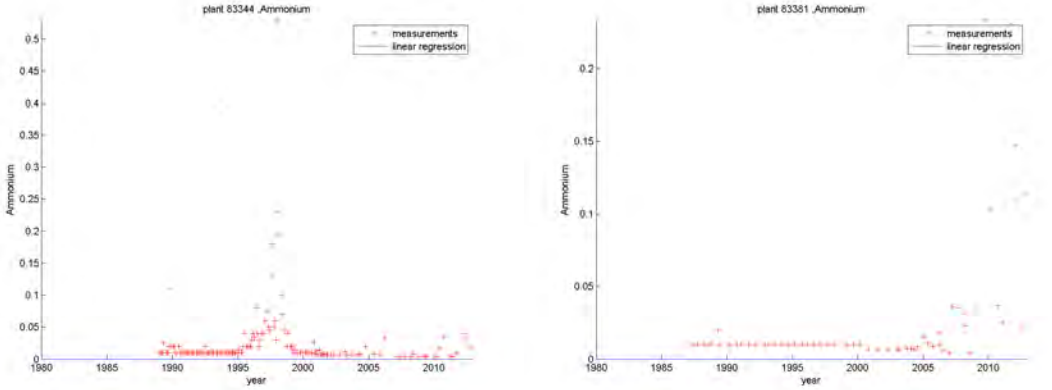
(a) Bederværket, (80568) Midtjylland

(b) Lyngbyværket, (80571) Midtjylland



(c) Østre Vandværk (97059), Midtjylland

Figure 6.7: Measurements of ammonium in mg/L for three water works in Midtjylland

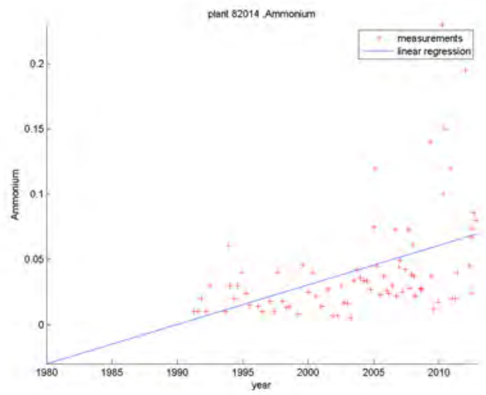


(a) Hellebæk Vandværk (83344), Hovedstaden

(b) Frederiksgade Vandværk (83381), Hovedstaden

Figure 6.8: Measurements of ammonium in mg/L for two water works in Hovedstaden

Ammonium's most frequent pattern shows an abrupt decrease during the second half of the 1980s, especially in Midtjylland. That is also the region with largest number of high averages, so in general the mean computed values suffer from the high values during the



(a) Odense Vandselskab A/S: Holmehavaværket (82014), Hovedstaden

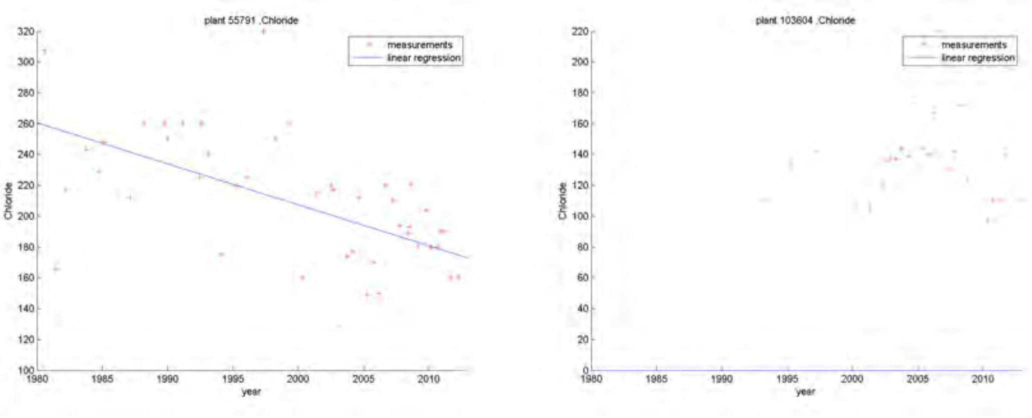
Figure 6.9: Measurements of ammonium in mg/L for one plant in Syddanmark with increasing values.

first half of the 1980s. Syddanmark seems to be the region with all averages meeting the compliance levels, but its most frequent pattern is a scattering in the values starting from the 1990s.

6.4.2 Chloride

Chloride assumes levels largely below the threshold of 250 mg/L for all the regions; the region with the highest values of chloride is Sjælland, which has an average level of 77.79 mg/L over the time interval. Although the plant Nakskov Vandværk (55791) reaches 208.07 mg/L as mean chloride level, it has a negative trend in chloride and since 2000 all levels have been below 250 mg/L (-26.59 mg/L over 10 years). Another water works with a high average of chloride (134.90 mg/L) is SK Vand A/S, Korsør Erdrup (103604), but its measurements scatter a lot, thus no trend can be noticed (see Figure 6.10). The highest absolute slope recorded is in plant Holbæk Vand A/S (Knabstrup Enge, 103982), 22.61 mg/L, but it is based only on data from about 1993 to 1997.

However all the regions have significant trends in several plants, at least in half of them, but with low slopes and in general with low chloride values. Often the measures scatter a lot, and only in few cases the trends are clear, but there are not many outliers. Chloride can be said to have a trend all over Denmark, apart from Sjælland and another few plants (106648, 104361, and 71401), but they have a slight slope. Plant Hirtshals Vanderværk Vest (71401) has a negative trend and chloride values below 190 mg/L (negative variation of 14.4 mg/L over 10 years, with mean value 106.3 mg/L); plant Tårnby vv. (106648) has a positive trend (+8.9 mg/L over 10 years), but the maximum recorder value of chloride is 160 mg/L, and plant Greve Vandværk (104361) had high values of chloride



(a) Nakskov Vandværk (55791), Sjælland (b) SK Vand A/S, Korsør Erdrup (103604), Sjælland

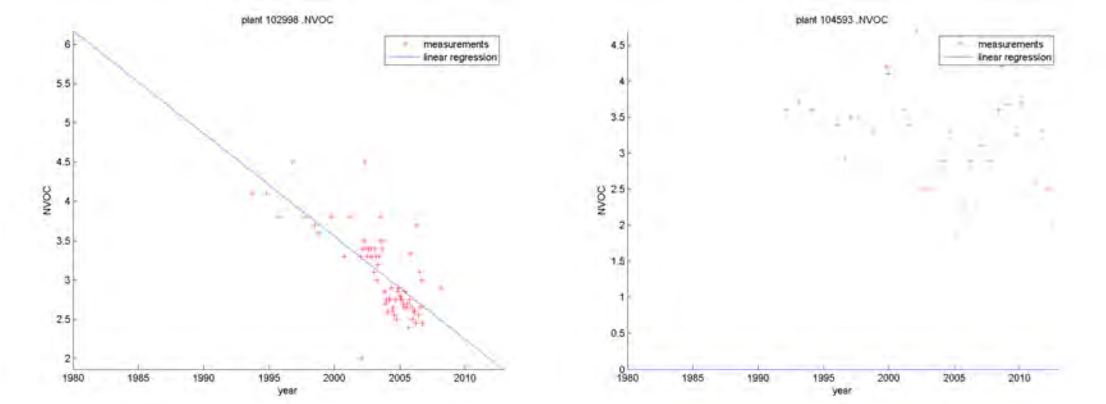
Figure 6.10: Measurements of chloride in mg/L for two water works in Sjælland

during the 1980s, but in the following years the law requirements were largely satisfied (negative variation of 43.3 mg/L over 10 years). In Hovedstaden, a general positive relationship of chloride over time can be noticed (from 2.3 mg/L to 14.7 mg/L over 10 years) but the average values are still largely within the prescription. In Midtjylland and Syddanmark the values of chloride are particularly low, most of plants have most of the values, respectively, below 30 mg/L and 70 mg/L. In Nordjylland, the plant Skagen vandværk (70624) has high chloride levels (chloride mean level is 194.9 mg/L) but no trend.

In conclusion, there are not noticeable variations or levels of concern for chloride. In fact, chloride is a conservative compound: it is not influenced by changing in the treatment and a stable value reflects the steady condition of the groundwater quality. The general picture of chloride levels in Denmark is small trends associated with low values. However, Sjælland experiences some high scattering level of chloride or relatively high slopes but associated with averages inside the requirements. No common pattern can be seen for chloride levels in Sjælland.

6.4.3 NVOC

NVOC presents a lot of significant trends only in Syddanmark, but the values are practically always under the maximum tolerable value of 4 mg/L. Nordjylland has one marked relative slope in plant Skagen Vandværk (70624), but only because of an outlier which moves the mean value from 6.17 to 19.44 and consequently the also absolute slope from -0.082 mg/L to -8.64 mg/L. Also plant Hirtshals vandværk Øst (71402) has relatively high values of NVOC; in fact the average is 2.69 mg/L but however largely below the limit. Plant Bagsværd (10311) has an high average (5.85 mg/L) but only because of an outlier, the other values are largely within the requirements. The only plant which has some values over the limit is Kalundborg Vandforsyning A/S, analyser vv, ledning (102998) in Sjælland; however the maximum reached value is about 4.5 and the trend is strongly negative (-1.31 mg/L over 10 years) (see Figure 6.11). The measures are taken starting from the 1990s. Plant Køge Vandværk (104593) has values of NVOC up to 4.7 and an average value of 3.23 mg/L but they do not follow any trend.



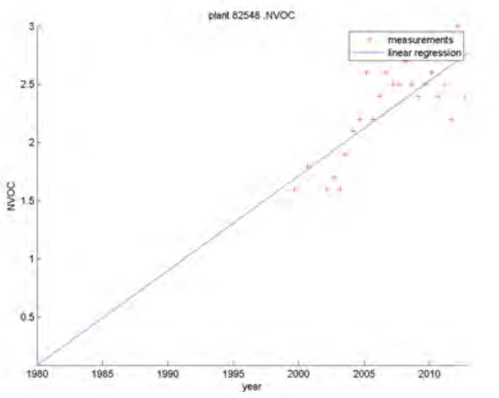
(a) Kalundborg Vandforsyning A/S, analyser vv, ledning (102998), Sjælland

(b) Køge Vandværk (104593), Sjælland

Figure 6.11: Measurements of NVOC in mg/L for two water works in Sjælland

A noticeable slope is present for plant Svendborg Vand: Skovmølleværket (82548) in Syddanmark, the absolute slope is 0.815 mg/L over 10 years and the average is 2.27 mg/L), but the measurements are only from 1999 (see Figure 6.12). The other trends with values above 3 mg/L are negative.

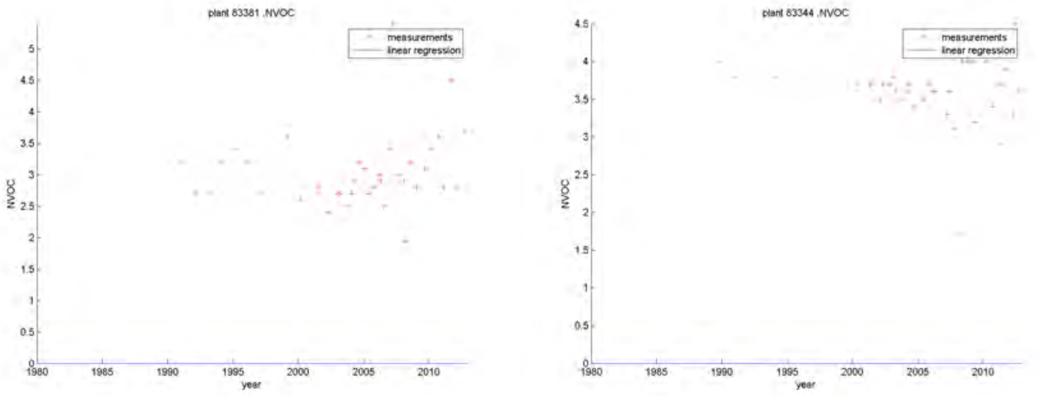
Despite the fact that NVOC levels in Hovedstaden generally satisfy the requirements, they are higher than in the other regions. Frederiksgade vandværk (83381) and Hellebæk



(a) Svendborg Vand: Skovmølleværket, (82548) Syddanmark

Figure 6.12: Measurements of NVOC in mg/L for one water works in Syddanmark

vandværk (83344) have chloride measurements around the threshold of 4 mg/L (the average values are respectively 3.05 mg/L and 3.61 mg/L)(see Figure 6.13).



(a) Frederiksgade vandværk (83381), Hovedstaden (b) Hellebæk vandværk (83344), Hovedstaden

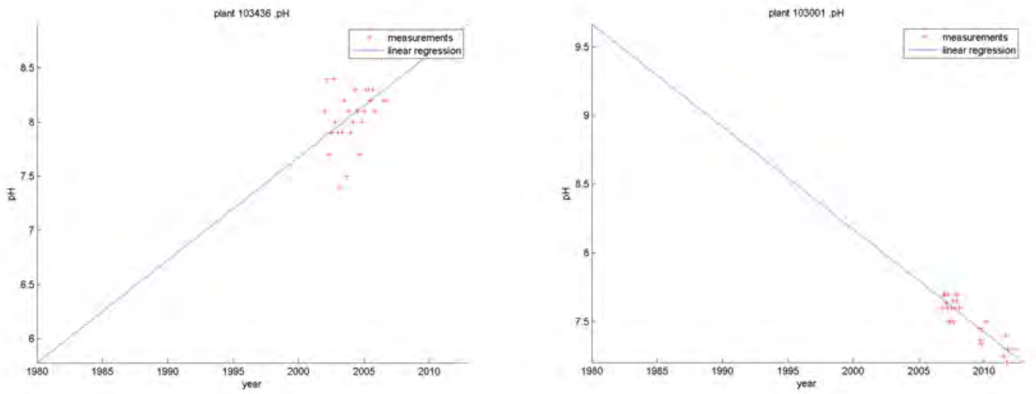
FIGURE 6.13: Measurements of NVOC in mg/L for two water works in Sjælland

In conclusion only the plant Svendborg Vand: Skovmllevrket in Syddanmark shows a noticeable trend and plants in Hovedstaden have values around the threshold.

6.4.4 pH

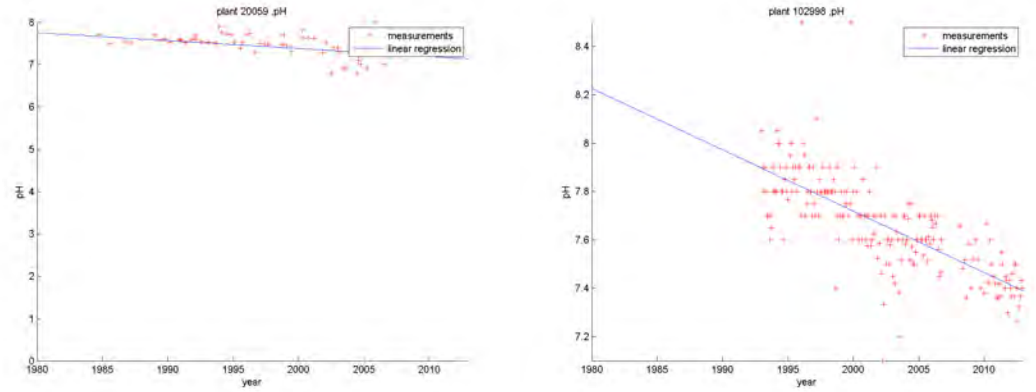
The average values of all the main plants in Denmark are inside the accepted range of 7-8.5. It has a significant trend in a large number of plants but its absolute slope is really low (among all 100 plants the range of the absolute slope for pH is [-0.74, 0.95]). Sjælland has two high slopes: one in plant Kalundborg Overfladevandforsyning A/S. Tissø (103436) and another in plant Kalundborg Vandforsyning A/S. Sultenkrog (103001) (see Figure 6.14). The first has a positive trend, but data are available only from 2002; it has two values between 2007 and 2009 above the upper limit 8.5. The second one has a strong negative trend, but similarly it has measures of pH from 2007; however, all its values are inside the range. The plant Marbjerg Værket (20059) has an outlier, a null value is present as a measure, probably inserted as a mistake, but does not have a big influence on the trend because it has the same order of magnitude of the measurements (the corrected absolute slope is -0.14 instead of -0.19 with an average of 7.35). The same happens for plant Køge Vandværk (104593), and probably for other plants, but for the previous reason the check can be ignored. Plant Kalundbord Vandforsyning A/W, analyser vv, ledning (102998) has measurements from 1993 characterized by a negative trend.

In Midtjylland, a relatively strong positive trend in pH is noticeable in plants 97059 and 80571. Plant Nordre Vandværk (97059) has data only from 1995, the absolute slope is



(a) Kalundborg Overfladevandforsyning A/S. Tissø(103436), Sjælland

(b) Kalundborg Vandforsyning A/S. Sultenkrog (103001), Sjælland



(c) Marbjerg Værket, Sjælland

(d) Kalundbord Vandforsyning A/W, analyser vv, ledning (102998), Sjælland

FIGURE 6.14: Measurements of pH in mg/L for four water works in Sjælland

0.19 over 10 years with an average of 8.01. Plant Lyngbyværket (80571) has an absolute slope of 0.14 over 10 years and a mean value of 7.96. Plant Hvinningdal Vandværk (79983) has some values out of the required range (the pH absolute slope is 0.15 and the average is 7.75). In Syddanmark, the only noticeable trend is in plant Skindermarkens Vandværk (53083), which has an absolute slope of 0.33 with an average of 7.69. In Nordjylland, the plant AKV-Kongshøj (69781) has a high negative slope (-0.38 with a mean value of 7.77), but data have been recorded since 2004 and during this period all values are inside the tolerable range.

pH trends are affected more by the leakage of data then from the outlier because usually they have values not so far away from the average. For many plants data are available only from the 1990s or even later. In general, drinking water supplied conform to the legislation.

6.4.5 Iron

Iron has 0.1 mg/L as maximum tolerable concentration according to the Danish law, it is a really strict requirement (in fact, the Australian guideline has 0.3 mg/L as iron limit and the WHO does not have any guideline). This is because iron is a good indicator of the sand filter operation, thus a low limit for iron may cause a benefit in other compounds removed from the sand filters (as manganese and ammonium). Lots of plants have average values above the limit, especially in Midtyjlland (15 on 20 plants). Nevertheless, most of plants that have a significant trend in iron show a decrease in its level. Plotting on a map only the average of iron which do not satisfy the requirements (in mg/L), it can be seen that high concentrations of iron are concentrated in the central area of Denmark apart from few plants (see Figure 6.15). Plants 14253, 103436 and 80582

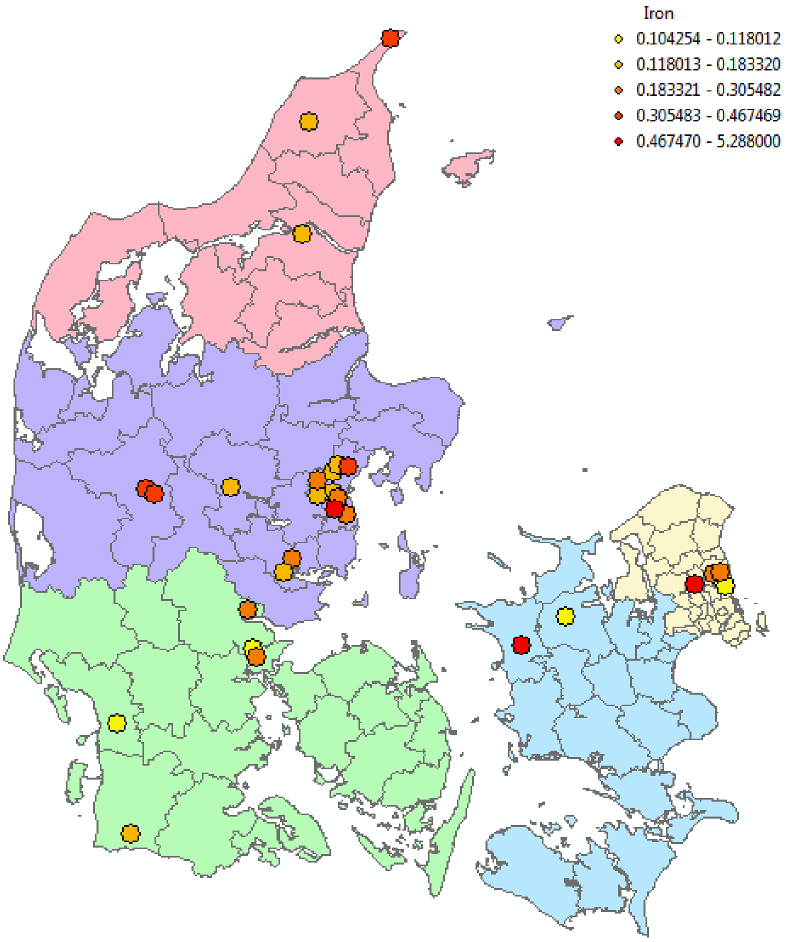
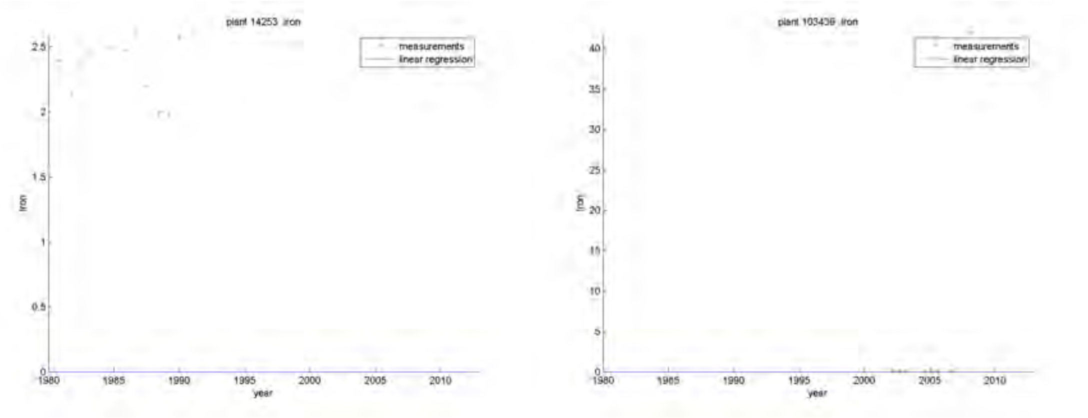


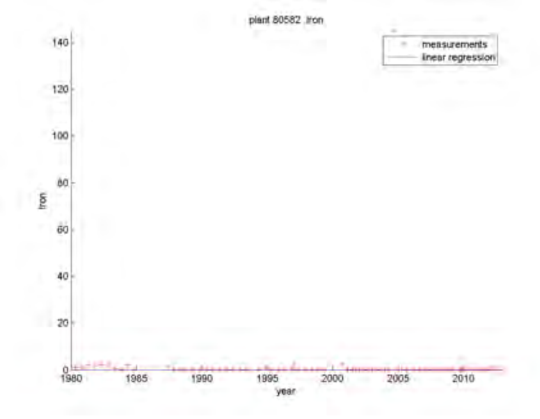
FIGURE 6.15: Measurements of iron in mg/L for water works which have noticeable average values (above 0.1 mg/L).

have particularly high values in iron averages: 2.35 mg/L, 5.28 mg/L and 2.16 mg/L. Plant Søndersø Øst (14253) has measurements only during the 1980s and all of them have high values. The averages of plants Kalundborg Overfladevandforsyning A/S Tissø (103436) and plant Østerbyværket (80582) are so high owing to outliers (see Figure 6.16). The only two noticeable positive slopes are in Sjælland in plants SK Vand A/S, Skælskør Nordre (103743) and SK Vand A/S, Valbygårdsværket (103821) but they are not reliable because of two outliers in the samples. These water works have iron level characterized by two peaks: one around 1995 and another around 2000; apart from those the level is stable and within the limit (see Figure 6.17). The other noticeable trends are all negative and they generally mirror a steep reduction which appears in Midtjylland's plants during 1987-1988 and after 2000 in the plants in Hovedstaden and Nordjylland. Although there is a general decrease in iron level, plants Nordre Vandværk (97057) and Østre Vandværk (97059) have high values during the last years but without a significant trend (see Figure 6.18).



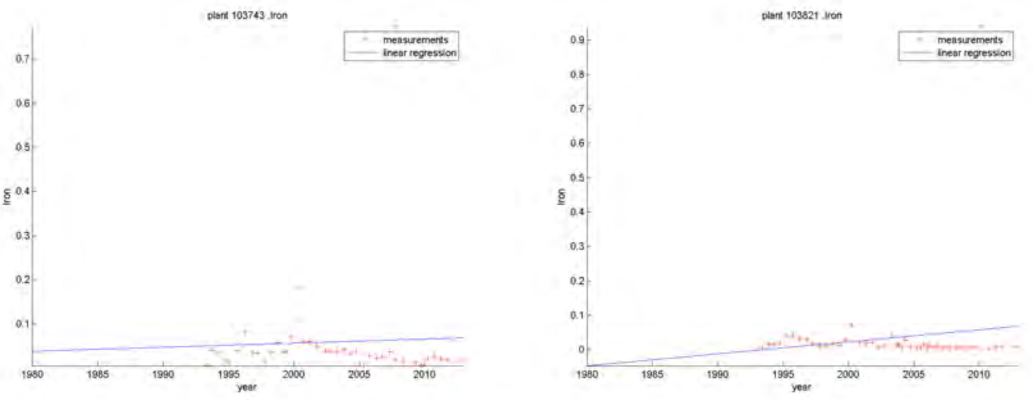
(a) SøndersøØst (14253), Hovedstaden

(b) Kalundborg Overfladevandforsyning A/S Tissø(103436), Sjælland



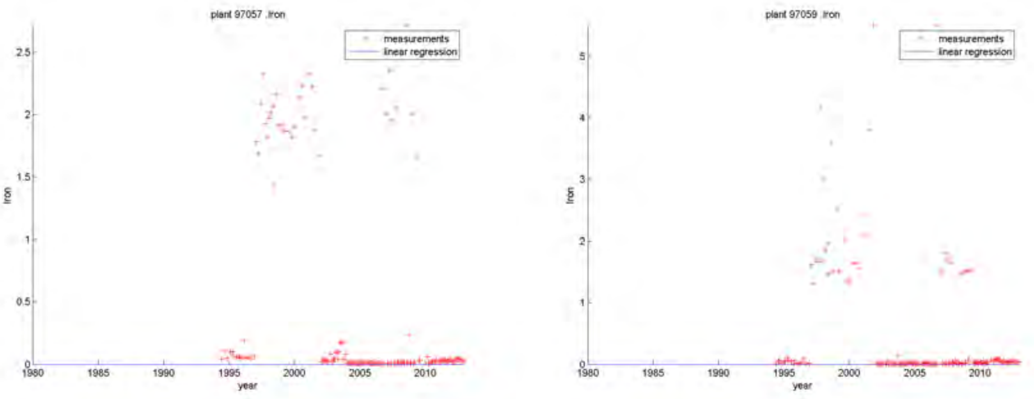
(c) Østerbyværket (80582), Midtjylland

FIGURE 6.16: Measurements of iron in mg/L for water works which have noticeable average values.



(a) SK Vand A/S, Skælskør Nordre (103743), Sjælland (b) SK Vand A/S, Valbygårdsværket (103821), Sjælland

Figure 6.17: Measurements of iron in mg/L for water works which have peaks in iron level in Sjælland.



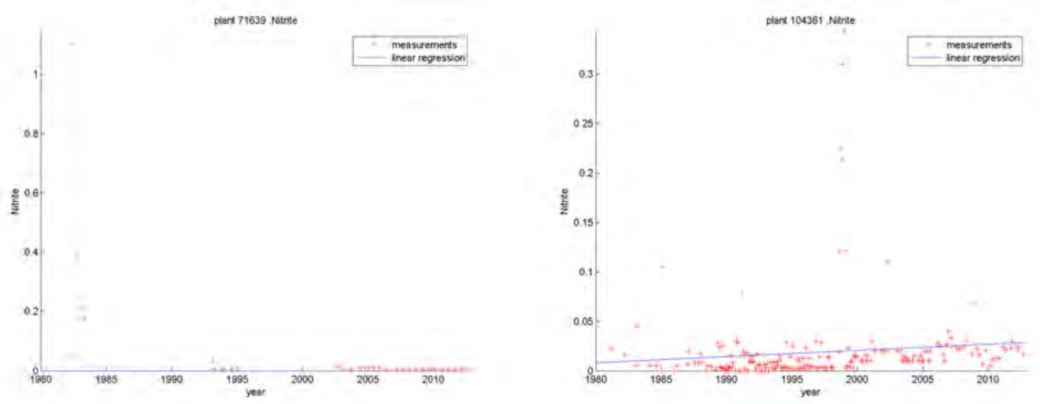
(a) Nordre Vandværk (97057), Midtjylland (b) Østre Vandværk (97059), Midtjylland

Figure 6.18: Measurements of iron in mg/L for water works which have high values during last years in Midtjylland.

6.4.6 Nitrite

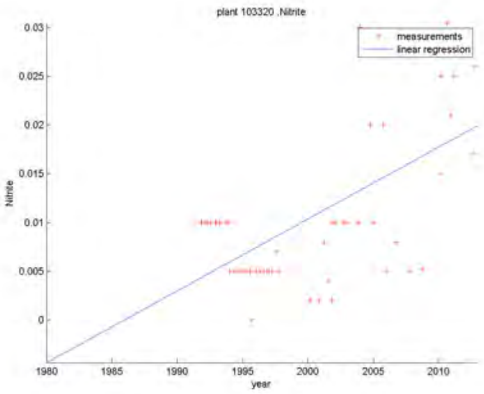
Nitrite's maximum content in drinking water is 0.01 mg/L according to Danish law. Nitrite results from the nitrification of ammonia in the sand filters, thus it is a good indicator of sand filter's operation (as iron). Nitrite does not have many significant slopes, but it is present quite often in high values. The highest values of nitrite are in plant Tolne Vandværk, Frederikshavn Forsyning (71639, Nordjylland), but they are concentrated in the period from 1980 to 1985. In Syddanmark there are many plants with some high values during 1995-2000.

Plants Greve Vandværk in Hovedstaden (104361) and Faxe Forsyning, Haslev (Bækvej) (103320) have samples characterized by high values and positive trends (plant 104361 has a nitrite mean value of 0.18 mg/L and an absolute slope of 0.006 mg/L; plant 103320 has a nitrite mean value of 0.011 mg/L and an absolute slope of 0.007 mg/L)(see Figure 6.19).



(a) Tolne Vandværk, Frederikshavn Forsyning (71639), Nordjylland

(b) Greve Vandværk (104361), Hovedstaden



(c) Faxe Forsyning, Haslev(Bækvej) (103320), Sjælland

FIGURE 6.19: Measurements of nitrite in mg/L for water works which have high values during last years in Midtjylland.

In some plants, ammonium and nitrite tend to follow the same pattern. In nitrite the effect of the water treatment is particularly evident , because the level of nitrite is often forced to be 0.01 mg/L.

6.4.7 Manganese

The manganese limit accoding to Danish law is 0.02 mg/L. Manganese results in having high averages in many plants, especially in Midtjylland and Nordjylland. Plants Lysholt Vandværk (73312), Kalundborg Overfladevandforsyning A/S Tissø (103436) and Højballegårdværket A/S (72322) respectively have averages 0.65 mg/L, 1.04 mg/L and 0.06 mg/L owing to outliers. Plant Søndersø Øst (14253, Hovedstaden) has disperse manganese values but still around the relatively high mean. In plant 103436, manganese was measured only in 10 months and one of these measurements is an outlier; calculating the average without the outgoing value it is still does not satisfy the request, but its percentage become a 248% instead of 3183%. Plant 72322 has a manganese average above the limit also neglecting that value (the percentage becomes 182%). The same happens for plant 73312, whose corrected mean percentage is 170%. Likely there are other mean values affected by outliers, but their presence is less noticeable than the previous cases. In general, manganese has no trend or if there is, it is really flat or negative. The cause is that the decrease is sharp most of the time, especially in Midtjylland. In fact, the high manganese averages in Midjylland reflect the high compound level concentrated during the first years.

The previous observations are summed up, noting if there are trends or measurements out of the tolerable range (Table 6.3). Only the plants with a trend leading to levels of concern or with many non-compliance measurements during the last years are listed. In the table both the significant trend and the average value are reported, highlighting in the column “notes” if the compound level or trend (or both) is particularly noticeable. Ammonium and pH have the largest number of cases causing concern. In particular, Sjælland has plant Kalundborg Overfladevandforsyning A/S Tissø with high average ammonium level and a significant positive trend. Another 5 plants with high recorded values are reported, though not all of them have a trend, on the contrary the values often scatter. As far as pH is concerned, both positive and negative trends are highlighted, because pH also has a lower limit. All its average values comply with the requirements, in fact trends in this parameter are the reason of concern. NVOC has only plant Køge Vangværk with some recorded values out of the requirements, but the average value is still within the limit of 4 mg/L. It is in the list because all plants experience low values of NVOC apart from the said water works. As a matter of fact, water in Denmark is at least 3 years old and the content of dissolved organic matter is generally low (Hilbert et al. [2010]). Iron has many measurements out of the limit 0.1 mg/L and a common pattern is hardly recognizable. Thus, in general there is not a significant trend but measures scatter a lot around high values. Similarly, nitrite experiences high values, but it does not have many significant trends.

Table 6.3: Noticeable plants, in the column "notes" it is reported "trend" if the slope is relevant or "levels" if some values of the compound are out of the requirements

Compound	**Region**	**Plant ID**	**Notes**	**Average**	**Change over 10 yy**
NH_4^+	Sjælland	103436	trend-level	0.095 mg/L	0.324 mg/g
NH_4^+	Sjælland	103982	trend-level	0.045 mg/L	0.244 mg/L
NH_4^+	Syddanmark	82014	trend-level	0.042 mg/L	0.030 mg/L
NH_4^+	Midtjylland	97059	level	0.021 mg/L	-
NH_4^+	Hovedstaden	83344	level	0.028 mg/L	-
NH_4^+	Hovedstaden	83381	level	0.028 mg/L	-
NVOC	Sjælland	104593	level	3.23 mg/L	-
pH	Sjælland	103436	trend-level	8.09	0.95
pH	Sjælland	103001	trend	7.52	-0.75
pH	Sjælland	102998	trend	7.67	-0.25
pH	Midtjylland	97059	trend	8.01	-0.18
pH	Midtjylland	80571	trend	7.96	0.14
pH	Syddanmark	53083	trend	7.69	0.33
pH	Nordjylland	69781	trend	7.77	-0.38
Fe^{2+}	Hovedstaden	14253	levels	2.36 mg/L	-
Fe^{2+}	Midtjylland	97057	levels	0.43 mg/L	-
Fe^{2+}	Midtjylland	97059	levels	0.47 mg/L	-
Fe^{2+}	Midtjylland	63364	trend-level	0.06 mg/L	0.01 mg/L
NO_2^-	Hovedstaden	104361	trend-level	0.02 mg/L	0.01 mg/L
NO_2^-	Sjælland	103320	trend-level	0.01 mg/L	0.01 mg/L

6.5 Discussion of trends

This paragraph presents the main identifiable patterns followed by the samples and infers possible causes of the changes in compound levels. It is important to notice that it is not the main aim of this work to find an explanation for all the changes in water compound contents for all 100 plants. Nevertheless, looking at common behaviors in the samples helps to understand how many factors influence inorganic compounds in drinking water. Often it is difficult to find a common pattern for water chemical compounds because several drivers occur simultaneously and manifest themselves with different speeds.

Ammonium, manganese and iron show a steep decrease in some plants in 1987. That change is likely due to the introduction of stricter requirements in drinking water quality. Low levels of ammonium, iron, nitrite and nitrate are good indicators of the sand filter's operation. The deposition of particles, air bubbles and precipitation of iron and manganese clog sand filters rapidly, causing uneven backwashing (Lopato et al. [2011]). In fact, in some cases, ammonium and nitrite follow the same pattern. A change in operation of the water treatment (probably in the sand filter) can be the cause. When there is a slow change in the said compounds, it can be caused by a deterioration in some components of the sand filter or by a change in the well use, because they influence the water quality for longer periods. Nitrite has higher values if it is measured just after backwashing in the sand filter.

Chloride has a general positive temporal trend, especially in the region Hovedstaden (up to 14.7 mg/L over 10 years); while in Syddanmark most of the plants have chloride level under 65-70 mg/L. Also sulphate has a generally positive trend, in Midtjylland in particular (it has all positive trends but the maximum slope is only 6.3 mg/L over 30 years). For the plants near the coast, a possible change reason is a salt intrusion from the sea. An increment in sulphate content could be caused by the nitrate front movement in the ground and the solphur oxides content because of air pollution. Other possible causes of chloride increase could be the salt used for road de-icing or leakages from landfills.

NVOC has only three plants with levels around the threshold of 4 mg/L: two plants are in Hovedstaden but the samples follow no trend and one plant in Sjælland, whose high values are concentrated in the 1980s. NVOC is a good indicator for groundwater flow variations.

6.6 Reliability of the calculated trends

The aim of this section is to have an overall look at the set of the trends, comparing them with the results of the Kendall's test. The test was applied to each dataset (for all the 100 plants and the 17 plants, that means 1700 datasets) to select the significant trend. A first selection of data was made before the test, filtering out the dataset with only one or two measurements and those one with many measurements with the same value. In fact, if data assume the same value (tie) too many times, the denominator in the equation of the Kendall's tau becomes impossible to calculate, being the square root of a negative number.

The outputs of the test are the tau-b and the relative p-value. The first measures the association of the values of the measurements with time, while the second is the probability associated to the hypothesis that there is not a trend over time. Hence, to approach the results it was firstly checked that there is an actual negative relationship between the value of tau-b and its p-value (Figure 6.20). A strong correlation between time and data (high module of tau-b) implies a low probability that the level of compound is constant in time (low p-value) and vice versa, as can intuitively be foreseen.

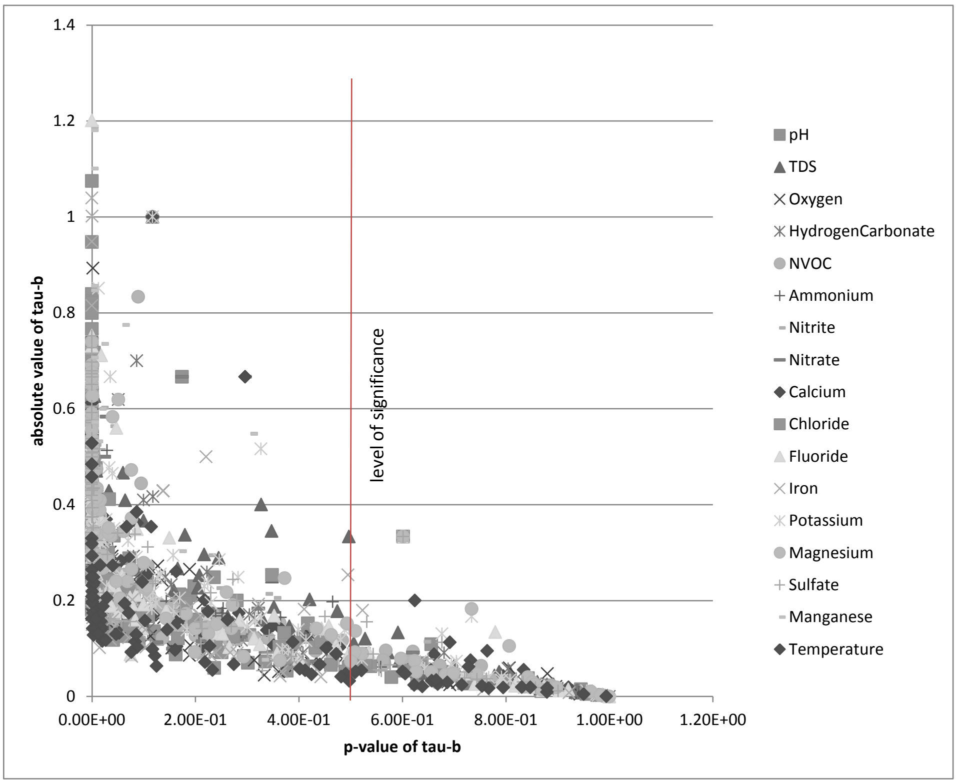


FIGURE 6.20: Negative relationship between tau-b and p-value. All the considered trends are below the level of significance, marked in the figure.

Then the values of the significant relative slopes are considered (that is those which led to a p-value of the tau-b below 0.05). They are compared with the p-values of the respective tau-b: if there is a negative proportionality, it would mean that datasets with low p-value have a big change over time. As can be foreseen, there is no correlation between the tau's p-value and the magnitude of the absolute value of the relative slope (calculated as the absolute 10-year slope divided for the average of the relative plant and compound) (Figure 6.21). (In the graph, only the significant slopes are plotted, i.e. those with p-value relative to the linear regression below 0.05). In fact, tau-b gives the measurement of how monotonic the trend is, independently from the sloping of the regression. The only common aspect is the sign: if the trend is increasing, both slope and tau-b are positive and vice versa (but in the figure the absolute value is considered).

Many datasets have a p-value below 0.005, so they are highly reliable. Generally, trends with p-value of the tau-b between 0.025 and 0.05 have a relative slope below 60%, while high relative slopes (>60%) are characterized by low p-values of tau-b (<0.025) (Figure 6.21). The graph has the x-axes limited at 160% as relative slope, because

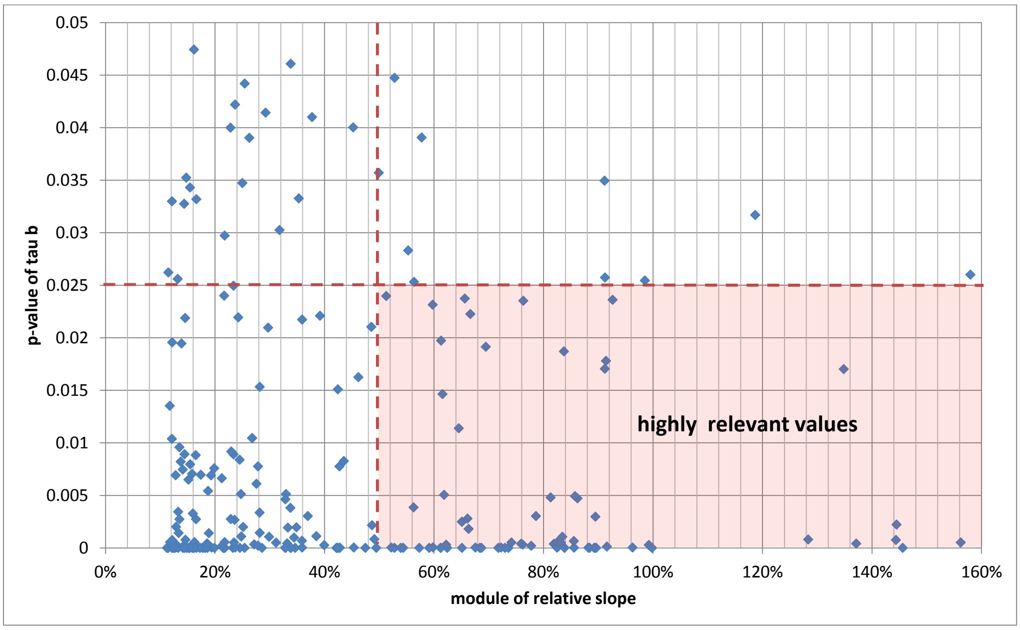


FIGURE 6.21: The datasets with relative slope above 60% and with p-value below 2.5% are highlighted among all significant trends as "highly relevant values".

the trends with slope above 160% have few measurements and some outliers, thus they are not reliable points to look to (see Section 6.1). The graph is divided in quadrants to emphasize the most interesting trends (marked as *highly relevant results*). Indeed those datasets have a big change over time associated with a high reliability.

Only the datasets marked as "highly relevant" are then considered (with relative slope $>$50% and p-value of tau-b $<$0.025). Their relative slopes are compared with the related plants effluent (Figure 6.22). Ammonium and iron have a large number of high trends. Chloride and potassium have trends varying from 50% to 75% for both the biggest water works (106294 and 25132). However, if smaller slopes are considered more trends are noticeable for those plants, but with smaller values.

In conclusion, many plants have significant trends in more then one compound. For that reason only the trends associated with lowest p-values were chosen. Comparing those with the relative effluent, ammonium and iron result in being the compounds with more high trends associated with a low p-value. Nevertheless, a relationship between slope of the regression and effluent discharge cannot be inferred from the data.

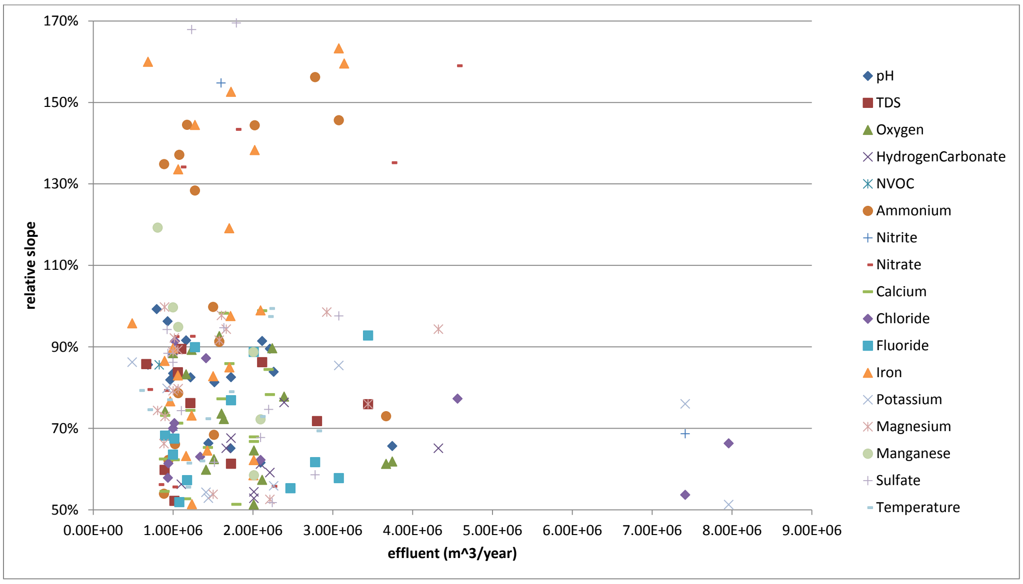


Figure 6.22: The datasets with relative slope above 60% and with p-value below 2.5% are compared with the respective effluents.

6.7 Basic parameters representing pipe corrosiveness, dental health and water taste

In this paragraph the results about the chosen parameters representing water corrosivity, dental health and water taste are presented. They are all based on the average measurements over the 32 years, thus the aim is to gain an overall idea of the effects of water quality in the three last decades on the said fields.

6.7.1 Hardness

The hardness was calculated as the sum of calcium and magnesium ion concentration expressed in equivalents. The graph represents the hardness values based on the average values of calcium and magnesium. Only plant 21582 has a mean hardness value out of the range, owing to the high average in magnesium (69 mg/L). But that value is not reliable because there is only one measurement of magnesium for that plant (see Figure 6.23). All the other plants have values classified as *good quality* (60-200 mg/L) or *increasing scaling problems* (200-500 mg/L) in the Australian guidelines (water quality management [2011]). Syddanmark has the highest frequency of plants having a good quality in water hardness (6 out of 20). In general, most of plants have a relatively high hardness (83%), while only 16% of plants have a good water hardness. Hovedstaden and

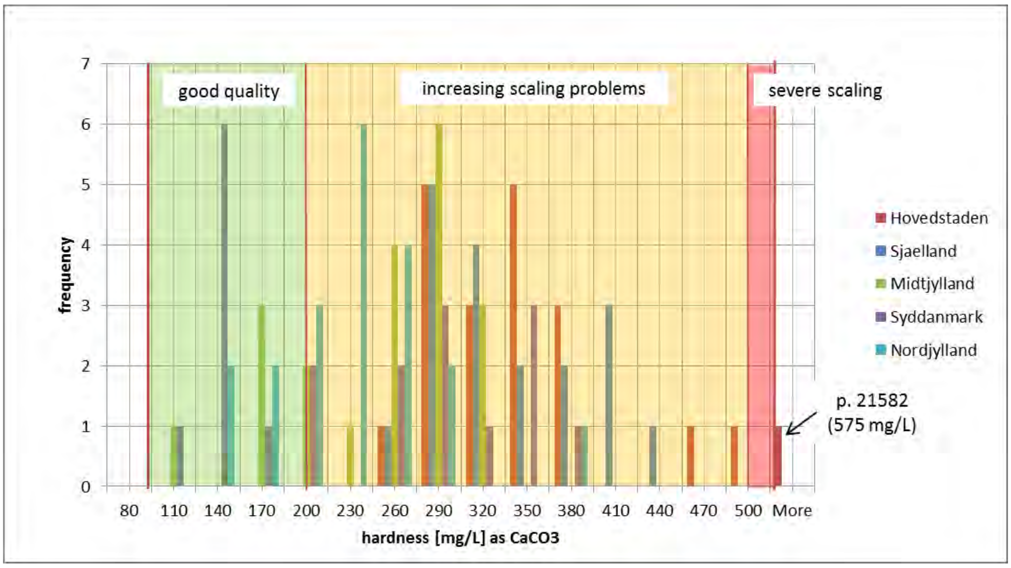


FIGURE 6.23: Histograms of hardness for all the selected plants divided by region. The categories given by the Australian guidelines are reported. (water quality management [2011])

Midtjylland are the regions with more plants in the range *increasing scaling problems*, and they are also the first regions for amount of water supplied.

About the distribution of hardness values in Denmark, a map is drawn overlapping the graduated points representing the averages with the figure having the water hardness averaged by region (see Figure 6.24,Hilbert et al. [2010]). Water with high levels of hardness characterizes all the area around København and it becomes softer moving to the west part of Denmark.

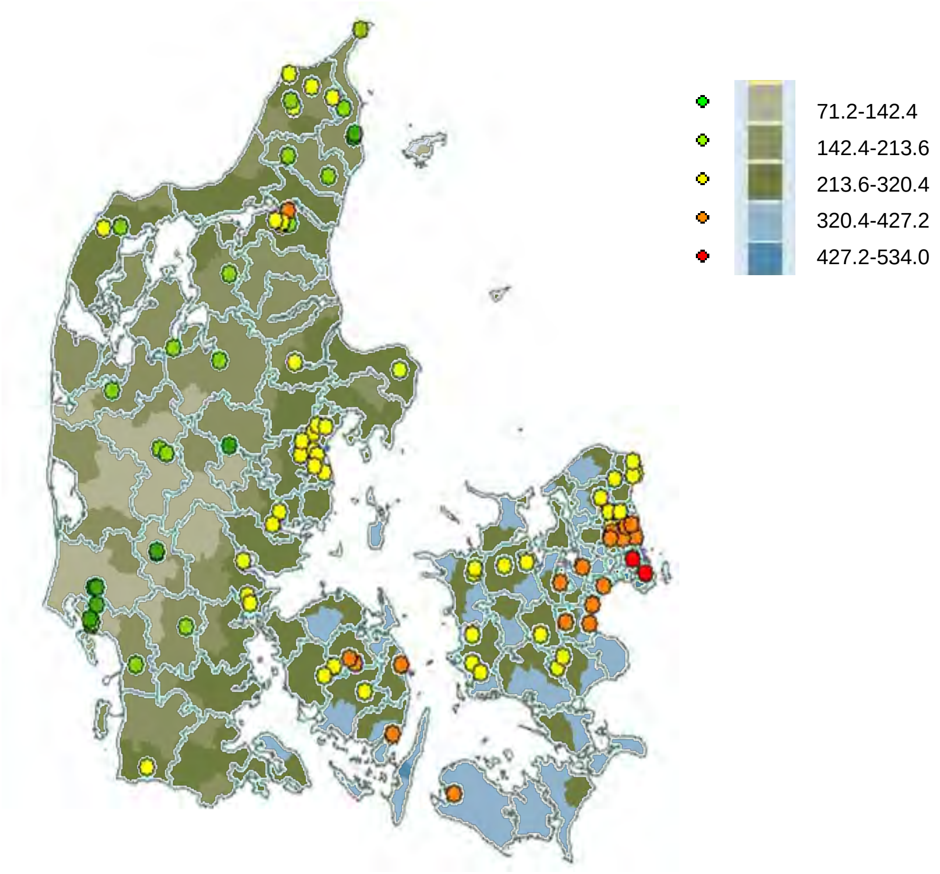


FIGURE 6.24: Map of water hardness in Denmark expressed as mg/L of $CaCO_3$

6.7.2 Alkalinity

The distribution of the average levels of alkalinity shows many plants with high values of alkalinity (see Figure 6.25). The alkalinity mean value all over Denmark is 220.34 mg/L.

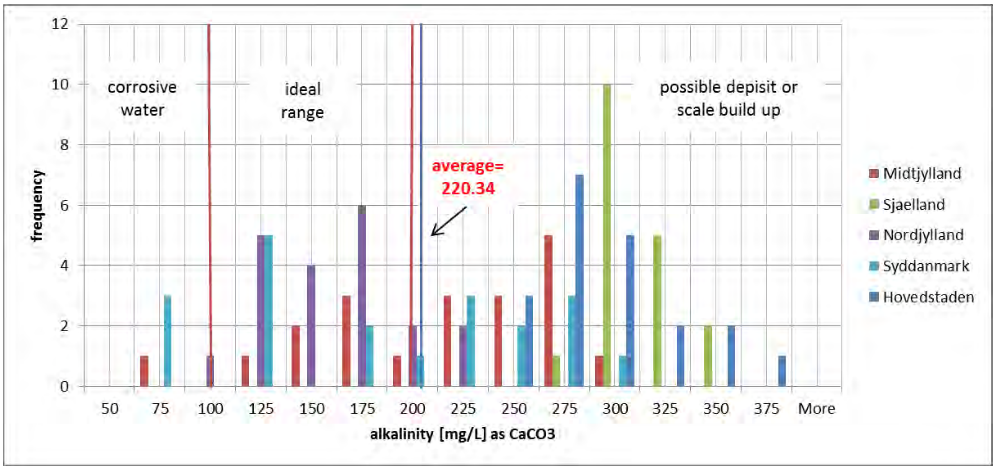


Figure 6.25: Histogram of average values of alkalinity for all selected plants divided by region. The limits of alkalinity ranges are reported according to the test on corrosiveness made by the University of Rhode Island (University of Rhode Island)

Hovedstaden and Sjælland have meanly high values of alkalinity, while Nordjylland is the region with the largest number of plants inside the ideal range. In the figure 6.26 the plants within the ideal range of alkalinity are represented.

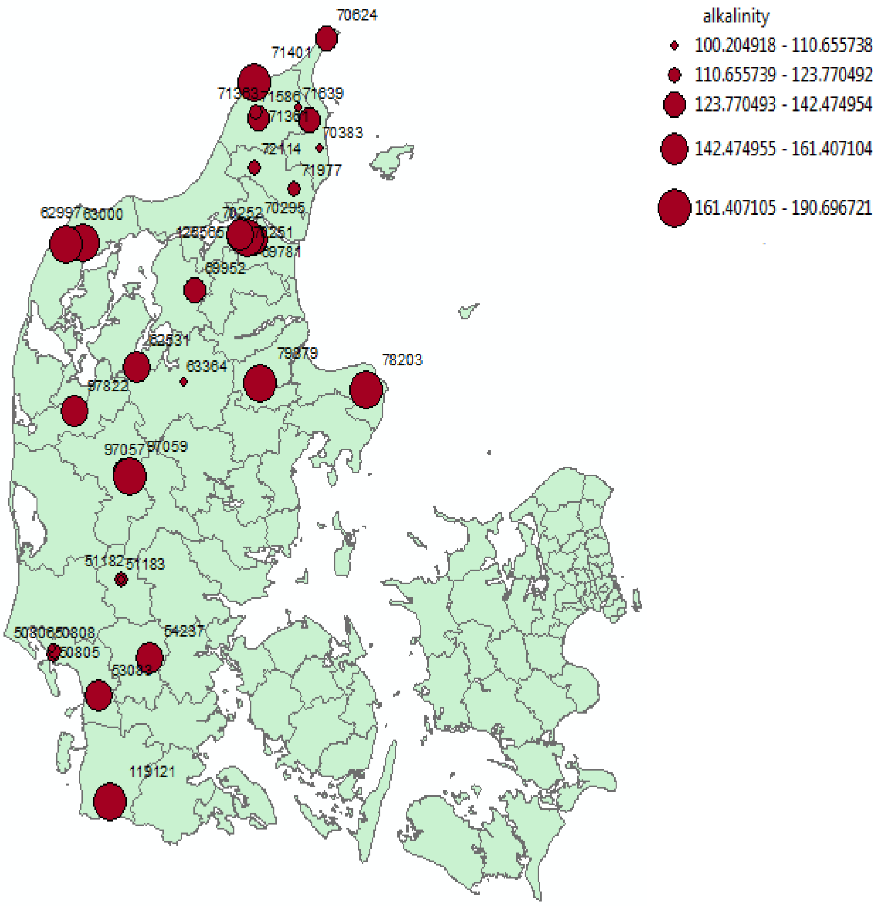


Figure 6.26: Map of the alkalinity averages within the ideal range (in mg/L). The optimum levels of alkalinity are located in Jutland.

The stability of water flowing through the pipes influences the extent of any corrosion rate (Mountain Empire). In order to evaluate the average water stability in the selected plants, the mean alkalinity and pH values are overlapped to the Baylis curve (see Figure 6.27, Mountain Empire). The Baylis curve represents the relationship between the alkalinity, pH and chemical stability of water. In the graph, the area below the lines represents corrosive water while above the line the water is scale-forming. The red dots represent the mean values of pH and alkalinity of the selected plants. A tendency to form scale can be noticed in most of the selected plants.

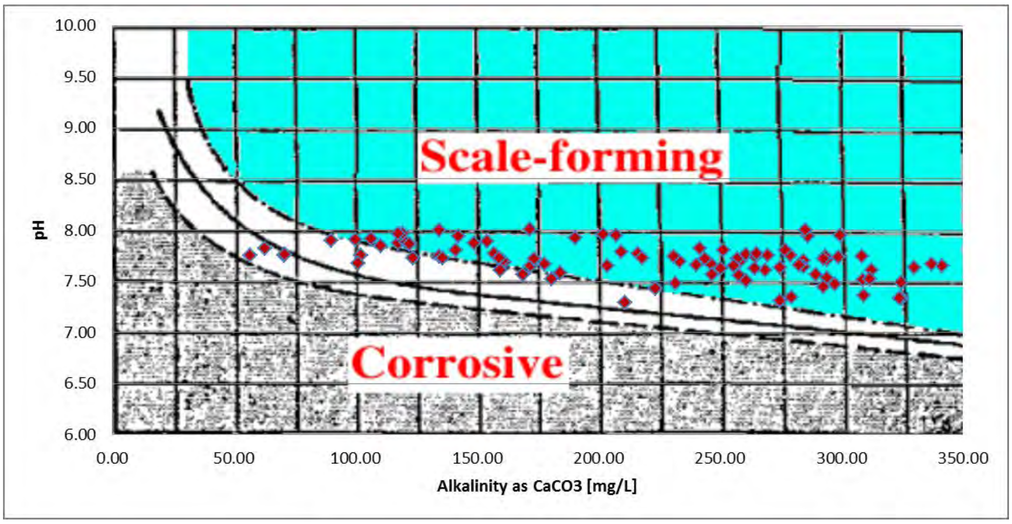


Figure 6.27: Water stability in relation to pH and Alkalinity (Baylis curve). A disposition in scale forming can be noticed for most of the selected plants.

6.7.3 Larson ratio

The average Larson ratios over time are drawn in the following figure. A tendency to high values of LR can be observed. The high value of LR in plant 21582 is caused by the chloride, whose average is not reliable because only one data is available and it is likely wrong (670 mg/L, while the other values are around 60 mg/L). The Larson index is above the maximum value of 1 in few plants in Hovedstaden, Sjaelland and Midtjylland. Its mean value in all Denmark is 0.61 (see Figure 6.29). Most of the plants with a low LR are in the area of Horsens and Kalundborg, while in the area close to København and Aarhus there are plants with high LR, even still inside the limit of one. The most plants denoted with high corrosive potential are in Nordjylland.

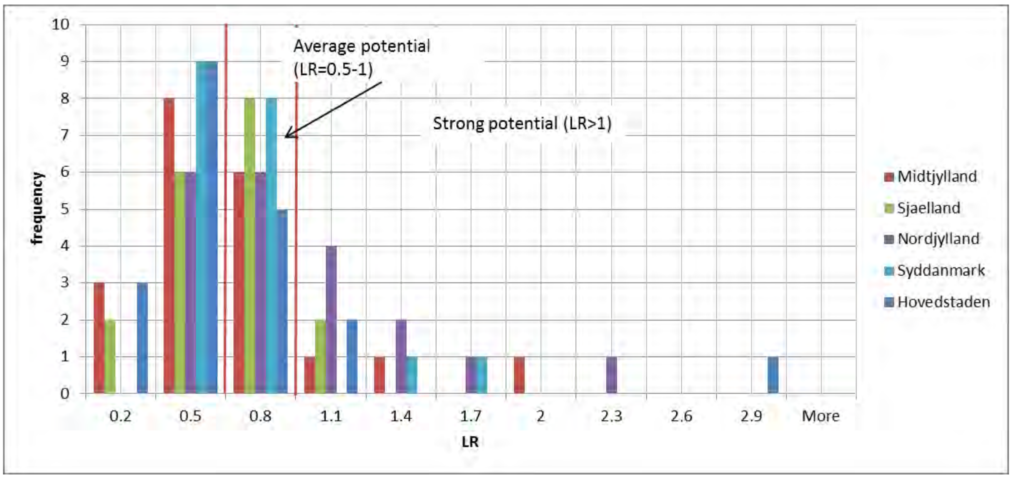


Figure 6.28: Larson ratios for the selected plants, classified regarding to the classification given by Delion et al. Meanly a tendence to high corrosiveness in drinking water is noticeable.

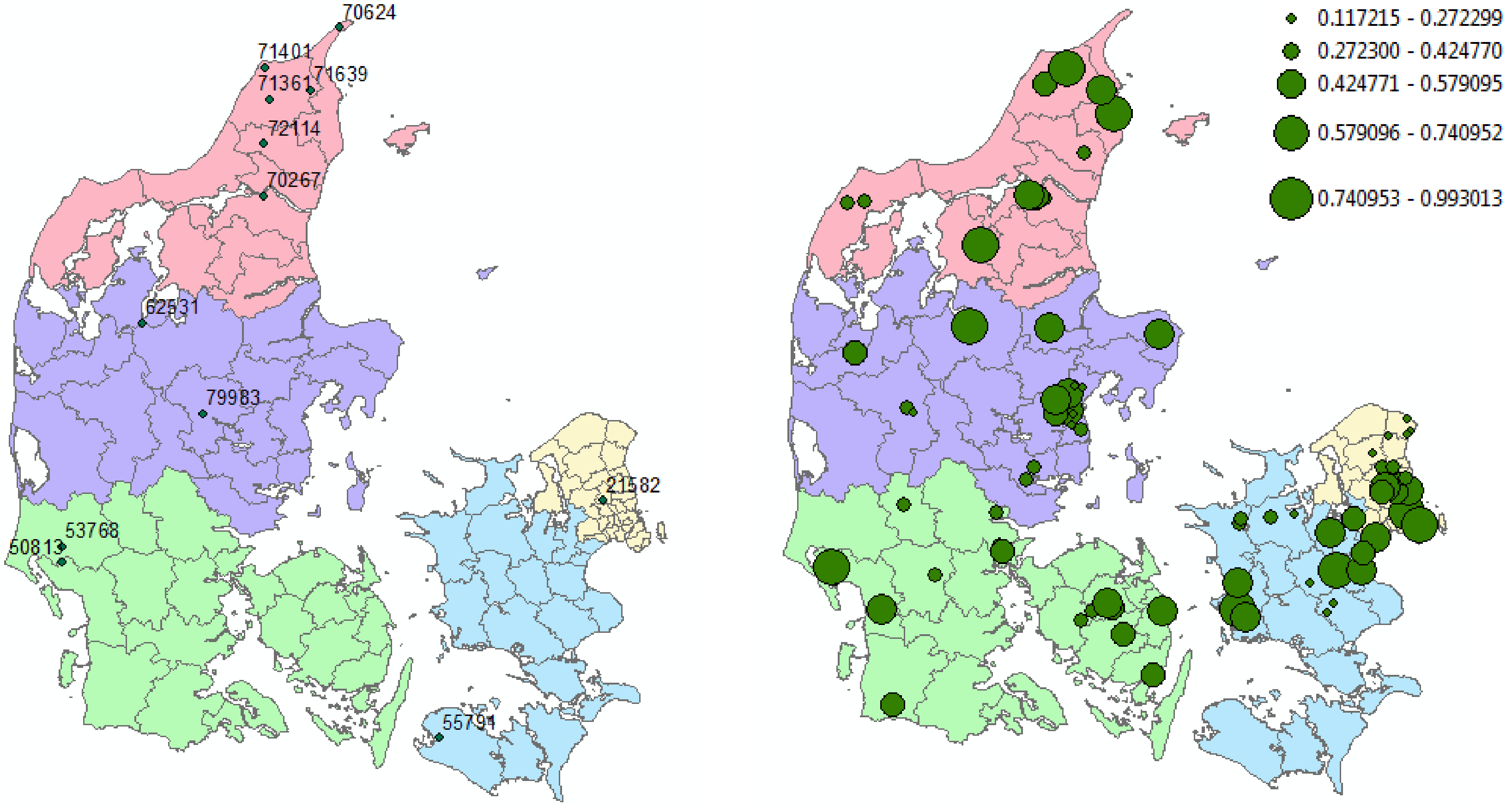


FIGURE 6.29: In the map on the left only the plants with LR above 1 are represented, while in the map on the right the values of LR are represented with proportional points. Smaller is the size of the point in the map, lower is the corrosive potential associated at the relative plant.

6.7.4 Dental health

Dental health is evaluated through the number of decayed, missing and filled surfaces calculated with the model presented in Section 2.2.2, in reference to the averages of compounds over the 30 years. The results are presented in the histogram (see Figure 6.30). The number of DMF-S tend to be higher than the ideal value of 2. In particular, plants in Syddanmark have 12 out of 19 plants with a calculated number of DMF-S above 3 (for one plant there are no measurements of fluoride), with values above 4. Also in Nordjylland there are several plants with more than 3 calculated DMF-S (18 on 20) but with a maximum value of almost 4 only in one plant. Midtjylland has 13 out of 20 plants with more than 3 DMF-S from the model, but only one plant has a value above 4. Hovedstaden and Sjælland are the regions with less DMF-S computed, (all the plants have meanly DMF-S below 3). The minimum number of DMF-S computed is in plant Greve Vandværk (104361, Sjælland), while the maximum is 4.19 in plant Varde Forsyning A/S, Lerpøtvej Vandværk (53768, Syddanmark).

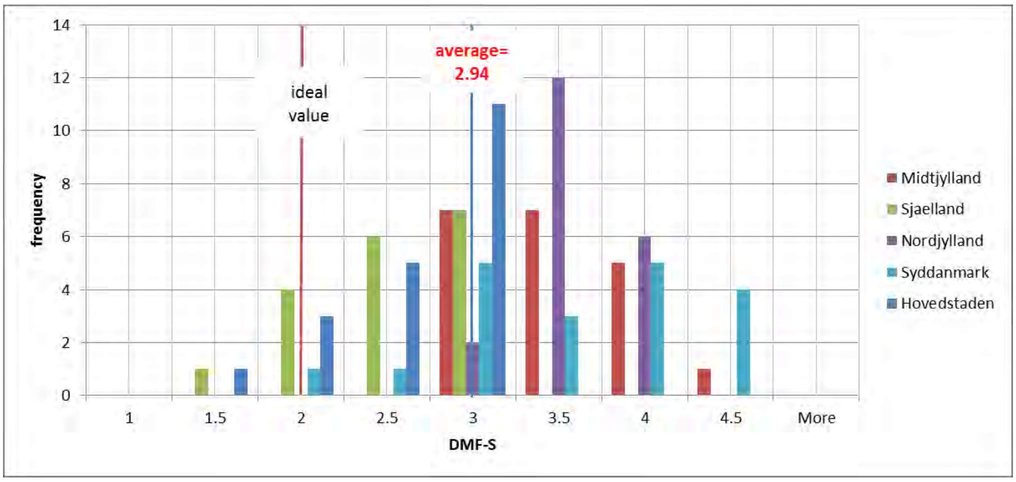


Figure 6.30: Histograms of calculated DMF-S. Most of plants has a number of DMF-S above 2 (suggested by Bruvo as ideal value (Bruvo et al. [2008])).

An high number of DMF-S reflects low levels in fluoride or calcium (or both of them), in relation to the suggested combination of 0.75 mg/L of F^- and 90 mg/L of Ca^{2+}. The blue and the orange lines in the graph represent respectively DMF-S=3 and DMF-S=2; the red point denoting the optimum combination belongs to the orange line. Looking at the averages of those two compounds in detail, the high number of DMF-S in Nordjylland is caused mainly by a low level of fluoride, as well as for Syddanmark and Midtjylland (see Figure 6.31). Sjælland and Hovedstaden have plants with calcium/fluoride combination closest to the optimum. On the other hand, fluoride levels above 1 may lead to fluorosis (McDonagh [2000]), but only the waterworks Greve Vandværk (104361), Nakskov Vandværk (55791) and Nærum Vandværk (106559) have average number of DMF-S in that range and however below 1.4 mg/L. Midtjylland, Nordjylland and Syddanmark have most plants below the blue line, further confirm of the previous considerations.

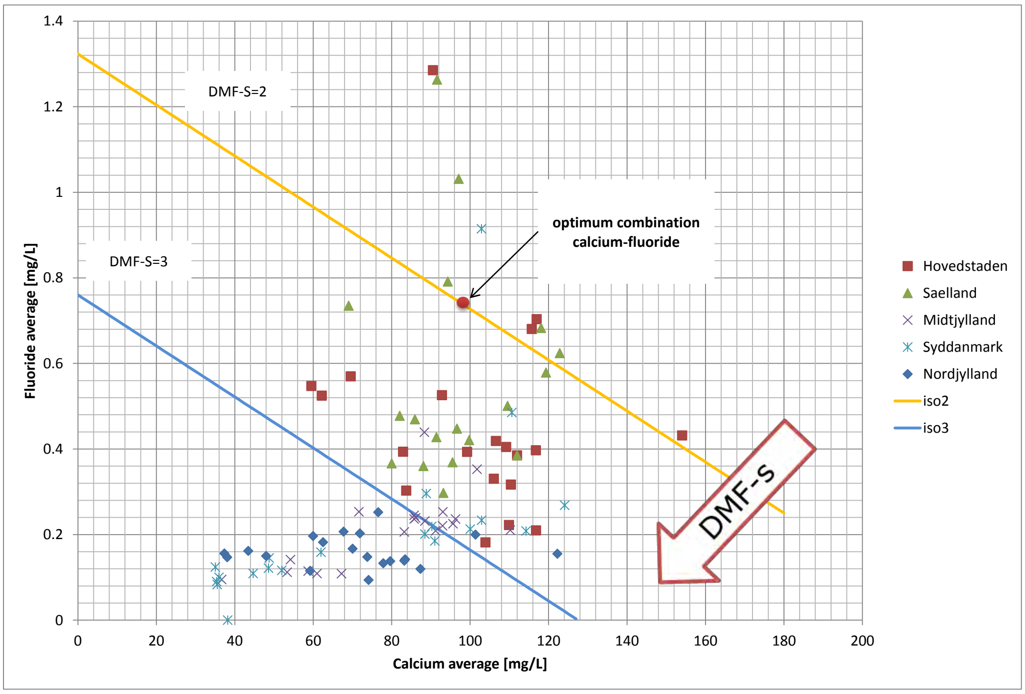


FIGURE 6.31: Average combinations of calcium/fluoride in the selected compound gathered by region. (suggested by Bruvo as ideal value Bruvo et al. [2008]).

On the whole, the drinking water quality supplied by the main water works is good from a dental health point of view if compared with the given range of 1-9 DMF-S (Ekstrand et al. [2003]). Thus, considering the range of variation in all Denmark, values up to 4-5 can be considered tolerable.

6.7.5 Water taste

The factor with the biggest influence on taste suggested in literature can be found in the inorganic compounds, TDS and pH. TDS is considered the most important factor. In this section, the method applied to evaluate the taste of the drinking water supplied from the selected water works is described. Firstly, the results of the taste index are presented; secondly, there are some considerations about the concentrations of the compounds effecting water taste; and finally, TDS and pH are considered to formulate an overall assessment.

In order to check the reliability of the taste index (TI), the TI was calculated for the water samples used by Platikanov in his survey (Platikanov et al. [2013]). It has to be noticed that all the three first samples are from bottled water (and they are the most liked) and the fourth one is from tap water (and it is the most disliked). The first value (0.99, marked in bold) is chosen as value of reference because it is the biggest; while the TI 0.74 is taken as a lower limit of reference, because it is related to the less liked sample by the panelists. The histogram representing the TIs calculated for each plant and gathered per region is reported. For three water works it has not been possible to evaluate the TI because of a lack of data. 33 out of 97 water works have a TI above 0.9 and 13 of them are located in Midtjylland. Only three plants have a taste index below the threshold of 0.75 but the maximum TI is 0.94. Nordjylland has all TIs below 0.9 but above 0.75 (apart from one plant with a TI of 0.58). Sjælland has TI's values distributed within the range [0.72-0.93]. Consequently, it seems that the water with the best taste is supplied in the region of Midtjylland. In general, the distribution of the TI has a negative skew indicating an overall good water taste.

TABLE 6.4: Labels and mineral contents of some samples used as references from a chemometric analysis on water taste (Platikanov et al. [2013]). For the same samples the taste index was calculated in order to find the upper and lower values to use as references

Label	**pH**	$\mathbf{HCO_3^-}$	$\mathbf{NO_3^-}$	$\mathbf{Ca^{2+}}$	$\mathbf{Cl^-}$	$\mathbf{K^+}$	$\mathbf{Mg^{2+}}$	$\mathbf{SO_4^{2-}}$	**TI**
CaSO4min2	7.8	63.98	0.32	68.9	2.9	0.6	11.6	116.7	**0.99**
. Ca(HCO3)min1	8.1	124.2	0.9	21.3	4.2	0.5	12.6	12.2	0.968
maximum 27	8.1	399	31.4	290.1	506.5	32.2	85.6	608.4	0.735
HMtap4	7.2	33402	31.4	290.1	506.5	10.1	85.6	608.4	0.737

The TI embeds a high number of compounds in one number; in order to see if the TI represents the actual situation in terms of taste, the compounds are analyzed singularly.

Due to the high number of plants, the average concentration over each region was calculated in order to get a handier dataset. This choice is criticizable because an administrative division not necessarily reflects geological or chemical differences in Danish

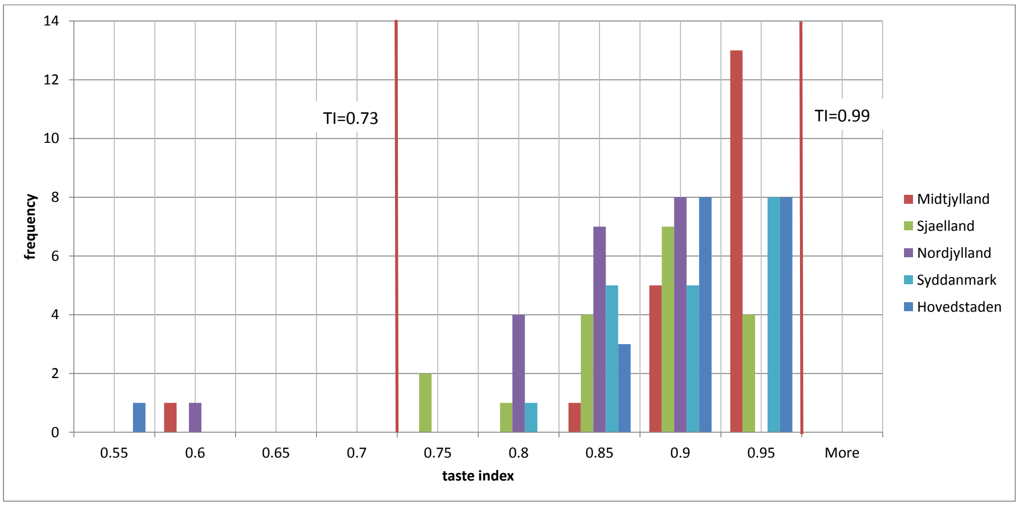


FIGURE 6.32: Histogram representing the mean TI for each plant grouped by region. The higher number of plants with high TI are in Midtjylland.

underground water sources. Although that allows a rough idea about the area in which water can be considered more tasty. The average concentration of the different compounds are not comparable as they are, thus they are normalized by the respective average all over Denmark. The normalized averages are expressed in percent and they can assume values higher than 100% (in all cases in which the average in the region is higher than the average all over Denmark). The gained percentages are plotted in a radar graph (see Figure 6.33). The radar graph gives further information in respect to the histogram with the TIs: that is the quantity of each compound and not only the proportion between positive and negative components. In fact, the resulting considerations are completely different:

- Sjælland and Hovedstaden seem to have better tastes because they are characterized by high percentages of Mg^{2+}, HCO_3^- and SO_4^{2-} and a low level of NO_3^-. Nevertheless they have high values of K^+ and Cl^-.
- Midtjylland, Syddanmark and Nordjylland have low percentages (<80% in relation to the national mean values) of all compounds with positive effects on taste; furthermore Midtjylland and Nordyjlland have a relatively high value in NO_3^- (>110%).

Despite the fact that the last figure is able to give a better picture of the water taste situation, it does not consider the supplied discharge by each plant. The goal now is to evaluate the water taste in relation to the real amount of water supplied grouped by

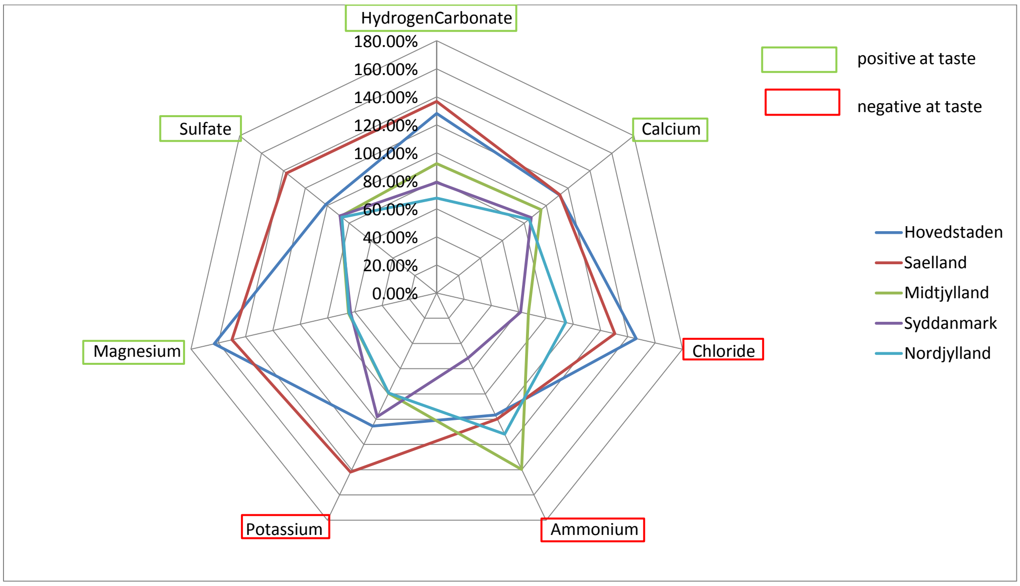


FIGURE 6.33: Diagram representing the average values of the compounds considered in the TI in each region. The compounds with a positive effect on taste are squared in green color and the compounds with a negative effect in red color.

region. For this purpose, the previous percentages are weighted with the normalized effluent as explained in Section 2.2.4. The radar diagram representing the described percentages can be compared with the latter one only for the shape, because the weights representing the discharge change the content expressed in per cent (see Figure 6.34). The graphs of the regions Midtylland, Syddanmark and Nordjylland keep approximately the same shape while the graphs of the regions Sjælland and Hovedstaden change theirs. In Hovedstaden, ammonium and especially sulphate are in bigger proportion than previously, compared to the graphs of the other regions. Despite the fact that all the 20 weights concerning the discharge sum is 1 for each region, Hovedstaden continues to be the region with the highest overall percentage of compounds and Syddanmark has the lowest. Looking at only the second radar graph (see Figure 6.34):

- The region Sjælland is the region supplying water with the highest concentration of all compounds which influence water palatability. In particular, the drinking water supplied in Sjælland is rich in sulfate, which has a positive impact on the taste.
- The region Hovedstaden, similarly, has approximately the same percentages in bicarbonate, calcium, chloride, ammonium and magnesium but a lower percentage in sulphate and potassium (and they have the opposite effect on taste).

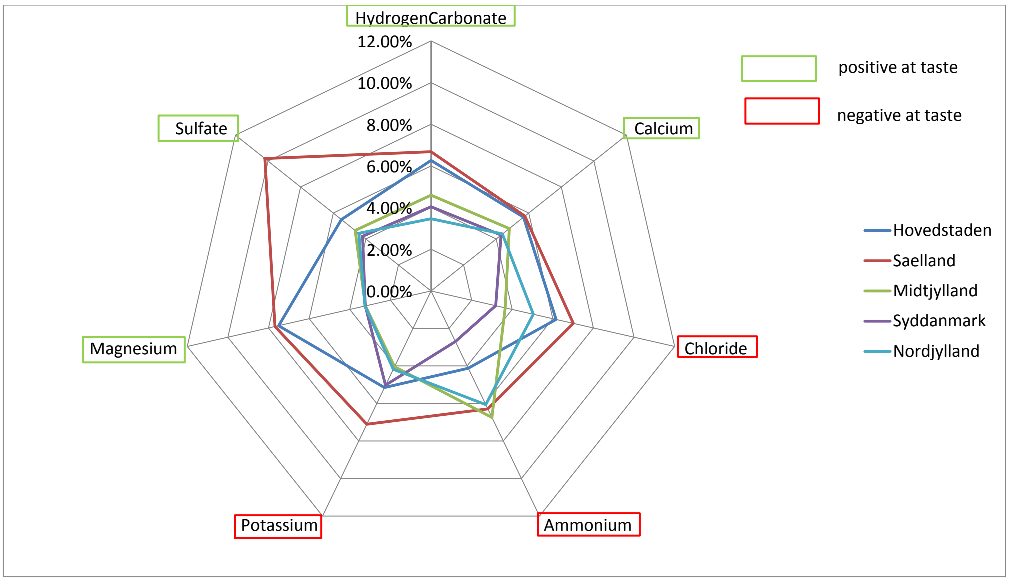


Figure 6.34: Diagram representing the average values of the compounds considered in the TI in each region weighted by the regional mean effluent. The compounds with a positive effect on taste are squared in green color and the compounds with a negative effect in red color.

- Regions Midtjylland, Syddanmark and Nordjylland have lower percentages. Despite that it has generally low percentages, Midtjylland has the highest percentage of ammonium, which has bad effect on taste.

Therefore, Sjælland and Hovedstaden have the most tasty drinking water.

Other important parameters for the water palatabity are TDS and pH. The histograms of both parameters are reported (see Figures 6.35 and 6.36). The survey led by Platikonov suggested levels of TDS up to 400, although among the selected Danish water works, 40 supply water with higher TDS content. However, the average all over Denmark is 406 mg/L, which is close to the taste limit but still far from the threshold of 200 mg/L suggested by Rygaard (Rygaard et al. [2011]). Hovedstaden and Sjælland are the regions with the lowest number of water works with TDS average below 400 mg/L (8 on 18 in Hovedstaden and 4 on 19 in Sjælland).

Considering the pH, the ideal range for taste is 7.5-8.1 and 88% of the water works are inside this range. The remaining plants have lower pH, the minimum average level is 7.3, and considering that the difference in taste is really appreciable only below 7, all plants supply water with a good level of pH. In conclusion, there is not an area with water taste that is particularly noticeable. In fact, if Sjælland and Hovedstaden show

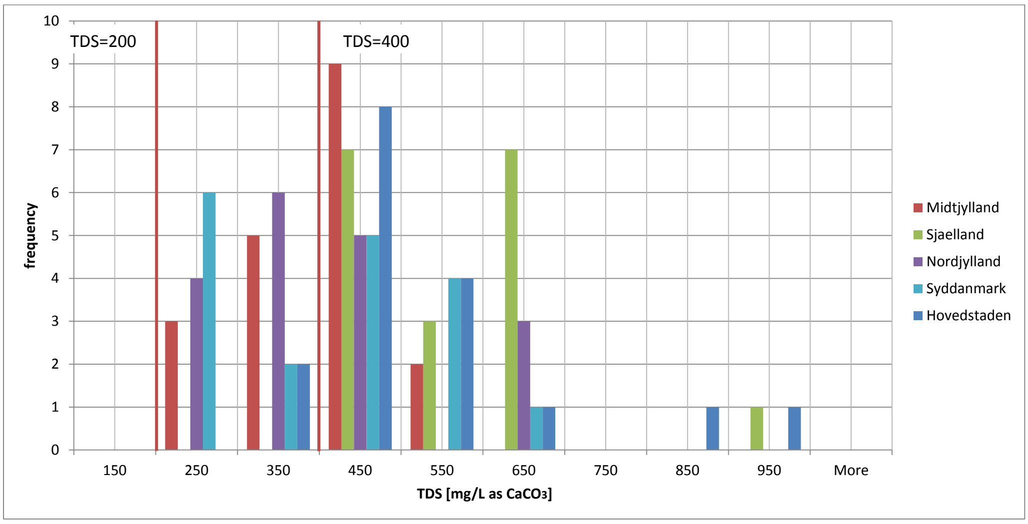


FIGURE 6.35: Histogram representing the mean TDS for each plant grouped by region. High levels of TDS can be noticed, the average is about 410 mg/L.

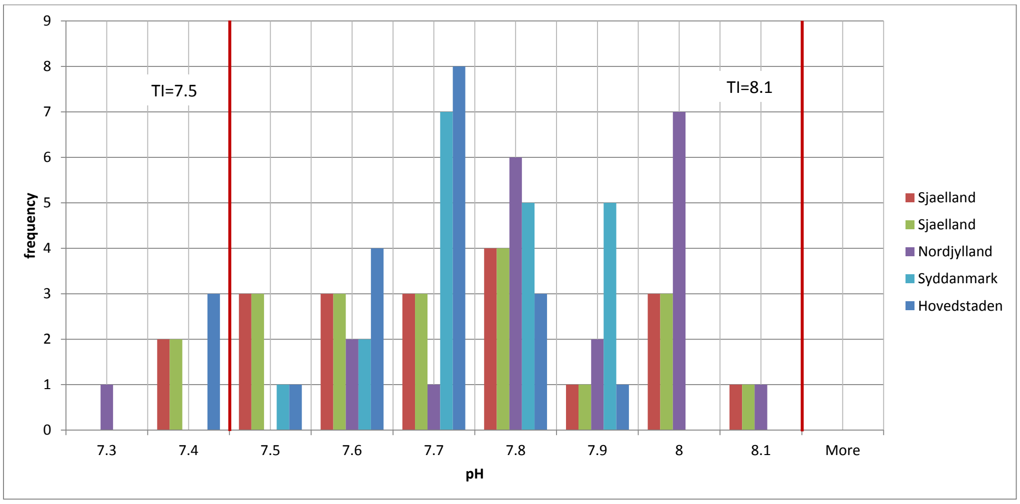


FIGURE 6.36: Histogram representing the mean pH for each plant grouped by region. pH assumes good values for the drinking water taste.

good levels for the inorganic compounds influencing the taste, they have also high mean TDS values, which influence taste in a negative way.

The taste index does not give a complete idea of the water taste because it has no information about the compound content but only the proportion between compounds which have a positive impact on taste and all the compounds which influence taste. Despite that the chemometric analysis was led by trained panelists, it has to be mentioned that a perceived difference in taste can have different effects on people (some people like drinking water that they are used to and some prefer trying different kinds).

6.7.6 Final considerations about the evaluated parameters

From the assessed parameters evaluated in average, it is difficult to point out an area in which the water quality is evidently better than the other. In fact, alkalinity experiences the highest averages in Hovedstaden and Sjælland, while Nordjylland is the region with most plants within the ideal range of alkalinity. On the other hand, Nordjylland is the region with the biggest number of plants with high Larson Ratio; while in Midtjylland, especially in the area of Horsens and Kalundborg there are good Larson Ratio values. Hardness is particularly high in Hovedstaden and Midtilland (respectively 80% and 70% of water works categorized as having "increasing scaling problems"). In general, Denmark is supplied with particularly hard water (most of plants has hardness between 15 and 20 dH), but with good total dissolved solid (the TDS average over time all over Denmark is about 410 mg/L). As far as dental health is concerned, Nordjylland is the region with the largest number of high DMF-S values, but Syddanmark reaches the biggest number of DMF-S in average (4.17). Sjælland seems the region having the best impact on water taste thanks to the high content in sulphate. However, Sjælland and Hovedstaden show many plants with high values in TDS, that has a negative consequence on water taste. Each considered field has different areas that can be considered good in average. If a particular area has to be chosen for the supplied water quality, the most influential parameter have to be identified. TDS influences both the formation of deposit inside the pipes and taste and high hardness is a characteristic of drinking water in Denmark. Contrariwise, Larson ratio represents corrosion in cast iron and steel but as the Baylis curve showed, the main problem in Danish distribution network is the scale build-up more than the corrosion. Thus, LR can be consider as a secondary parameter. From this point of view, Syddanmark can be considered the area with the best water supplied. As a matter of fact, in that region alkalinity has low value as well as hardness and TDS does not assume particularly high values. The discordant observation to that conclusion is that Syddanmark has low average levels in fluoride which lead to relatively high number of calculated DMF-S. Considering the singular plants, the water works which have more than one parameter assuming a high value is pointed out:

- Plant Frederiksberg Vandværk (44357): Both hardness and TDS assume high mean values: 490.3 mg/L as $CaCO_3$ (category "increasing scaling problems") and 923.3 mg/L (category "poor taste"), but TDS has only one measurement. Larson ratio is 0.99, close to the limit of strong corrosive potential. In the same plant also ammonium and nitrite assume high values (respectively 0.05 mg/L and 0.01

mg/L), but nitrite experiences lower values in the last years. Chloride assumes high scattering values (with average 118 mg/L).

- Plant Tårnby vv (106648): Hardness mean value is 464.7 ("increasing scaling problems") and alkalinity mean value is 341.7 mg/L (category "possible deposit or scale build up"); thus both those two parameters indicate water quality causing scale forming. The high hardness value is caused both by high calcium levels (with average 115 mg/L) and high magnesium levels (with average 42 mg/L). Moreover TDS has an average of 751.1 mg/L, that is still in the taste category "fair" but far from the threshold suggested by Rygaard of 200 mg/L. Also ammonium average is close to the threshold (0.05 mg/L), as well as nitrite (0.01 mg/L).
- Plant Søndersø Vest (21582): It has high hardness (575.4 mg/L) and high alkalinity (371.3 mg/L), thus the water supplied likely causes scaling problems. Nevertheless the magnesium average is not reliable because only one data is available. Also oxygen and chloride have apparently high levels but they have only one measurement. In general, the levels of this water works suffer from a leakage of data.
- Plant Nakskov Vandværk: It has high TDS mean value (850.9 mg/L) but associated to a negative trend. Larson ratio has also high average: 1.1, indicating high corrosive potential; likely it is caused by a relatively high level in chloride (208.1 mg/L). Ammonium has also high average value (0.04 mg/L) but its measurements scatter a lot.Nitrite assumes value close to the limit (the average is 0.01 mg/L) but with lower values during the last years.
- Plant Hvinningdal Vandværk (79983): It has high average value in Larson ratio (1.28), caused by a low average value in hydrogen carbonate (69.0 mg/L) and number of DMF-S (more than 4). The mean NVOC is particularly low (0.63 mg/L), while nitrite and iron assume meanly high values (respectively 0.012 mg/L and 0.14 mg/L).

Chapter 7

Future studies

This thesis is one of the first studies on drinking water quality in Denmark considering a long-term time correlation of all inorganic compounds, thus several are the possible improvements and developments:

- A criterion to filter out the outliers should be implemented in the code in order to have more reliable data sets. Also a selection of only the plants with a minimum number of measurements is suggested, in order to prevent high trends based on few values. In the program, an automatic comparison between average and significant trend slope could be introduced to find the plants with high slope associated to high average in short time.
- Apply the same analysis during only the last years (for instance 10 or 20 years), with the aim to see if the non-compliance levels are still present.
- Find an other method to build up the taste index, considering also TDS and pH. It has to include also the absolute mean values of the compounds and not only a ratio of them.
- The common problem all big databases have is the difficulty in selecting the wanted data. The implemented procedure allowed the researcher to gain a handy set of measurements, to develop some statistics from it and draw the graphs fairly automatically. Nevertheless, the method can be improved by importing the measurements in Matlab directly from Access to make the queries directly from there.
- Find the possible causes of the changes in compound levels.
- Compare the calculated number of DMF-S with the real data.

Chapter 8

Conclusion

The mean discharge analyzed is 168 million m^3 per year supplied from the main 20 water works in each region of Denmark (100 in total); it is about 50% of the total effluent of Danish water works. The compounds considered are most inorganic compounds in drinking water in addition to some parameters: TDS, NVOC, pH and temperature. From the investigation on drinking water effluent from the hundred main Danish water works, the following conclusions are drawn:

- For many plants, the number of measurements registered in Jupiter is lower than the expected one in compliance with the sample frequency given by Danish law.
- Practically all plants have significant trend in at least one compound. The time interval considered is 32 years, thus a change is foreseeable in such a long period. Hovedstaden and Midtjylland have respectively 39% and 43% of significant trends among the 340 data sets analyzed for each region.
- pH, oxygen, sulphate, chloride and nitrate are the compounds which have most significant trends, while iron, ammonium and nitrate are the compounds with highest slope values.
- Considering the absolute values of ammonium, iron, pH, nitrite and manganese, those compounds can be considered causes of some concern. Ammonium experiences high levels in the area of Horsens and Aarhus; it shows an abrupt decrease during the second half of 1980s especially in Midtjylland, while in Syddanmark there are many scattering value starting from the 1990s. pH has average values within the prescriptions but it has few measurements concentrated in a short time period. Nitrite has not many significant slopes but it experience a lot of values above the limit (7 waterworks have non-compliance mean level); for some periods

its measurements mirrors the treatment effect because it is forced to be equal to the limit.

- Manganese show a sharp decrease during the 1980s, following sometimes the same pattern as iron and ammonium.
- Ammonium and iron have the largest number of high trends associated to low p-values (that means their slopes are highly reliable).
- Despite ammonium, iron and manganese have averages in lots of plants which seem not to comply with the law, only ammonium, pH, iron, nitrite and NVOC in 20 data sets (1.2%) are pointed out as potentially harmful.

The considerations on taste, corrosion and dental health are based on the average values. It allows the reader to get a rough idea of the average situation during the last 30 years, but it cannot represent the changes over time. Looking at the mean values, the following inferences are derived:

- Hovedstaden and Sælland are the regions with the highest hardness in Denmark (in which water hardness is classified as *increasing scaling problems*). In fact, drinking water in the area around København is characterized by a high hardness (430-530 mg/L as $CaCO_3$) and it becomes softer moving to the west part of Denmark (reaching values up to about 70 mg/L as $CaCO_3$ in the area around Esbjerg).
- Carbon steel corrosiveness was assessed through the Larson Ratio. Most plants with low LR, indicating low corrosive potential, are in the area of Horsens and Kalundborg, while high values of LR are concentrated in the region of Nordjylland.
- Most part of the selected water works supply water with high mean values of alkalinity (the average value all over Denmark is 220 mg/L as $CaCO_3$ but two plants have alkalinity value above 350 mg/L). Despite that, in Nordjylland 18 out of 20 plants have alkalinity within the ideal range.
- Drinking water supplied by the main Danish plants seems to have good characteristics in terms of dental health: the calculated number of DMF-S (Decayed, Missing and Filled Surfaces) assumes values up to 4-5, which is relatively low in comparison with the range 1-9 of the mean caries experience in Denmark. Nordjylland is the region with highest number of DMF-S from the model (18 on the 20 plants have more than 3 assessed DMF-S and supply about 20 million m^3 of water) while Hovedstaden and Sjælland have a low number of DMF-S mainly because of the good fluoride content (0.24-1.24 mg/L). Only three plants have fluoride averages above 1, indicating a possible fluorosis risk.

- An attempt to build an index able to indicate water palatability in relation to some compounds' concentrations was led. However it is difficult to represent a subjective evaluation (as taste is) with only an index. Looking at all inorganic compounds having an effect on taste, in addition to TDS and pH, it remains difficult to point to an area with particularly good water taste. TDS highly influences water taste and it assumes mainly high values (40 water works supply water with TDS content above 400 mg/L). However TDS is partly made of compounds having a good influence on taste: calcium, magnesium, hydrogen carbonate and sulphate. As an example, Sjælland is the region with the water richest in sulphate but also with all TDS average above 400 mg/L.

- The Jupiter database is a good source of information but its potential is strongly reduced by the presence of some outliers. They reduce the likelihood of the results gained with an automatic procedure, and they force the researcher to go back to the original data series to express reliable conclusions.

The results of this thesis indicate that drinking water supplied by the main 100 danish water works in general supply good water. Most of the apparently non-compliance average values or particularly steep trend over time hide an adjustment of the levels during the years. A first assessment on corrosiveness indicates that in the area of Nordjylland water has higher corrosive potential than the other regions of Denmark. Hovedstaden and Sjælland especially are characterized by high values in hardness and alkalinity, indicating a disposition in scale build up. In general, the water supplied seems to have positive effects both on taste and dental health.

Appendix A

Plant selection

Table A.1: Tracking numbers, coordinates and names of the plant selection in Hovedstaden.

Plant id	effluent [million m^3/year]	Plant name
106294	7.41	SjælsøVandværk
14253	4.56	Søndersø Øst
2065	4.32	Søndersø Vandværk
106292	3.44	Ermelundværket
44357	2.39	Frederiksberg Vandværk
104361	2.01	Greve Vandværk
83344	2.01	Hellebæk Vandværk
83381	1.61	Frederiksgade Vandværk
83122	1.61	Birkerød Vandværk a.m.b.a.
83132	1.22	Farum Vandværk
21582	1.18	Søndersø Vest
83340	1.17	Espergærdeværket
83339	1.12	Snekkerstenværket
106558	1.00	Holte vv
106648	1.00	Tårnby vv
106559	0.97	Nærum vv
83080	0.90	Lillerød Andelsvandværk a.m.b.a.
106741	0.90	Værløse vandværk
106311	0.89	Bagsværd vv
83142	0.83	Endrup Vandværk

Table A.2: Tracking numbers, coordinates and names of the plant selection in Sjælland.

Plant id	effluent [million m^3/year]	Plant name
25132	7.96	Regnemark
103436	3.36	Kalundborg Overfladevandforsyning A/S. Tissø
20059	2.82	Marbjerg Værket
102998	2.20	Kalundborg Vandforsyning A/S, analyser vv, ledning
104361	2.01	Greve Vandværk
28315	2.01	HOFOR - Lejre Værket
104593	1.72	Køge Vandværk
103821	1.70	SK Vand A/S, Valbygårdsværket
2065	1.67	SøndersøVandværk
103653	1.22	Ringsted Vand A/S Tystevad Vandværk
55791	1.06	Nakskov Vandværk
103320	1.00	Faxe Forsyning Haslev (Bækvej)
103604	0.90	SK Vand A/S Korsør Erdrup
55874	0.89	Pindsobro Vandværk
105008	0.89	Solrød Vandværk a.m.b.a.
103369	0.88	Holbæk Vand A/S (Søndre Vandværk)
103743	0.81	SK Vand A/S Skælskør Nordre
103000	0.80	Kalundborg Vandforsyning A/S. Deigvad-nord
103001	0.80	Kalundborg Vandforsyning A/S. Sultenkrog
103982	0.78	Holbæk Vand A/S (Knabstrup Enge)

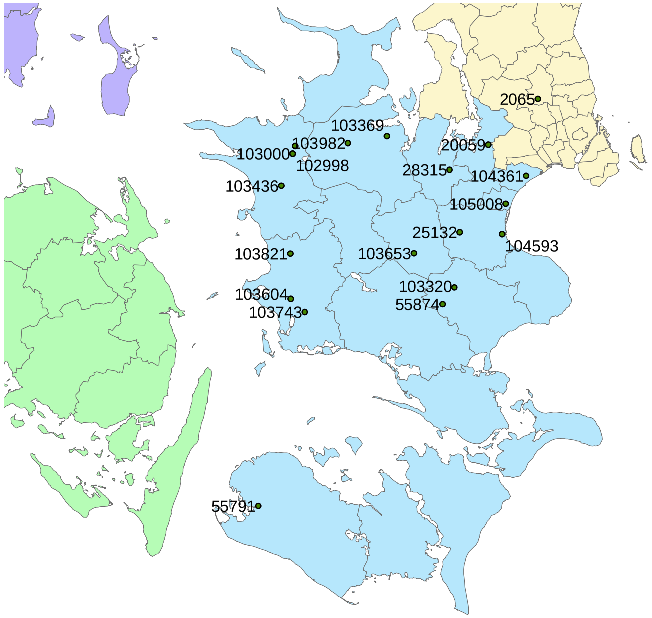


Table A.3: Tracking numbers, coordinates and names of the plant selection in Midtjylland.

Plant id	effluent [million m^3/year]	Plant name
80562	3.14	Stautrupværket
80567	3.07	Kastedværket
72322	2.93	Højballegårdværket A/S
79983	2.80	Hvinningdal Vandværk
80568	2.78	Bederværket
80572	2.47	Truelsbjergværket
62531	2.26	Skive Vandværk
79379	2.11	Oust Mølle Vandværk
80565	2.02	Åboværket
63364	1.82	Viborg Vand A/S - Nord
73312	1.73	Lysholt Vandværk
97057	1.69	Nordre Vandværk
72323	1.58	Rugballegårdværket
97822	1.44	Frøjk Vandværk
78203	1.34	Havdal Vandværk
80573	1.27	Elstedværket
80578	1.18	Vibyværket
97059	1.15	Østre Vandværk
80571	1.07	Lyngbyværket
80582	1.02	Østerbyværket

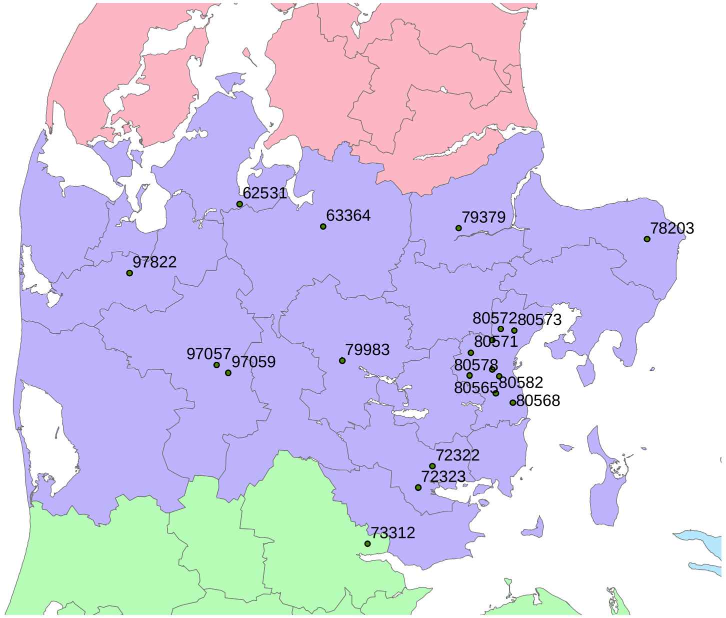


Table A.4: Tracking numbers, coordinates and names of the plant selection in Syddanmark.

Plant id	effluent [million m^3/year]	Plant name
50806	3.75	Spangsbjerg Vandværk
82014	3.67	Odense Vandselskab A/S: Holmehaveværket
82003	2.24	VandCenter Syd as: Lindvedværket
50805	1.79	Vognsbøl Vandværk
53768	1.64	Varde Forsyning A/S Lerpøtvej Vandværk
81725	1.52	NFS: Hjulby Bro Vandværk
74496	1.51	TRE-FOR Follerup Vandværk Fredericia
74497	1.50	TRE-FOR Kongsted Vandværk Fredericia
50808	1.43	V. Gjesing Vandværk
82001	1.41	VandCenter Syd as: Borrebyværket
53767	1.29	Varde Vandforsyning A/S (ledningsnet for Lerpøtve)
53083	1.23	Skindermarkens Vandværk
119121	1.16	Tønder Vandværk
51182	1.10	Grindsted Vandværk 2
82440	1.02	Midtfyns Vandforsyning-Vandgården/åværket
51183	0.97	Grindsted Vandværk 3
54237	0.97	Vejen Forsyning A/S østre
50813	0.94	Astrup Vandværk
82548	0.94	Svendborg Vand: Skovmølleværket
82010	0.93	VandCenter Syd as: Dalumværket

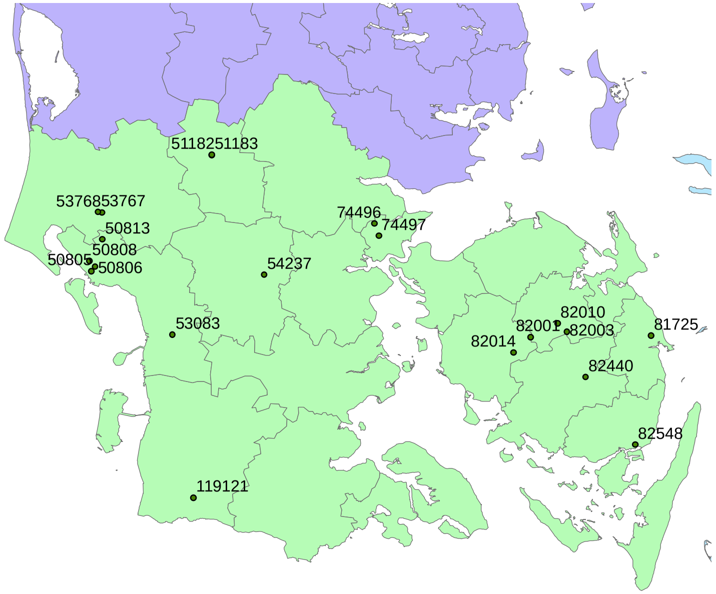
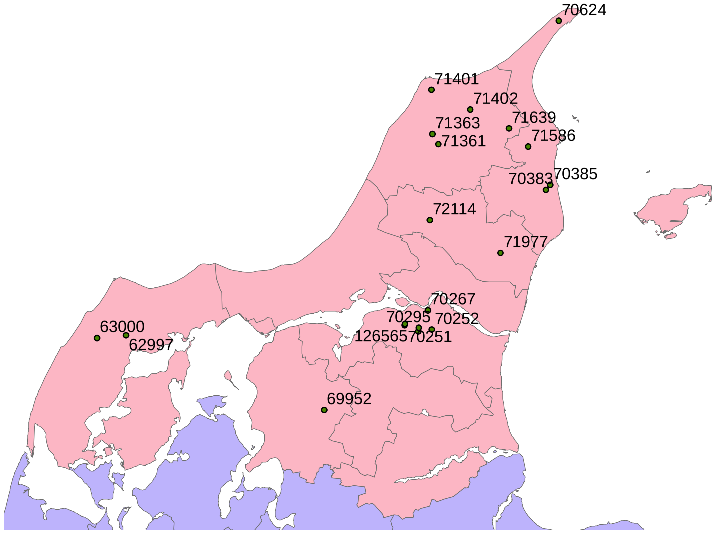


Table A.5: Tracking numbers, coordinates and names of the plant selection in Nordjylland.

Plant id	effluent [million m^3/year]	Plant name
62997	2.21	Thisted Drikkevand A/S Baun
71361	2.10	Bagterp Vandværk
70624	1.71	Skagen Vandværk
69781	1.60	AKV-Kongshøj
70251	1.45	AKV-Drastrup 1
71639	1.44	Tolne Vandværk Frederikshavn Forsyning
70295	1.37	AKV-Vissegård
71401	1.24	Hirtshals Vandværk Vest
71402	1.07	Hirtshals Vandværk øst
71586	1.03	åsted Vandværk
70252	0.94	AKV-Brunsted
63000	0.93	Thisted Drikkevand A/S Vang
126565	0.83	AKV-Drastrup 2
70267	0.69	Vejgaard Vandværk Administration
71363	0.69	Bredkær Vandværk
70383	0.66	ørnedalsværket
70385	0.62	Sæbygårdværket
69952	0.58	Aars Vand A.m.b.A
71977	0.55	Østvendsyssel Råvandsforsyning
72114	0.49	Brønderslev Søndre Vandværk



Appendix B

Kendall's test results

The table has to be read as follows:

- If the value in the p-th row and in q-th column is 1 it means that the p-th compound and the q-th plant has a significant trend at the α-level of 0.05, that is the null hypothesis had been rejected.
- If the value in the p-th row and in q-th column is 2 it means that the observed test statistic is outside the critical region, the null hypothesis cannot be rejected at the α-level of 0.05, thus the p-th compound and the q-th plant has probably no significant trend.
- If the value in the p-th row and in q-th column is NaN it means either there are only one ore no measurement at all or the number of ties in the statistic test is larger than the number of possible pairs in n elements, where n is the length of measurement array.

Table B.1: Results of the Kendall's test for Hovedstaden

plant id	pH	TDS	O_2	HCO_3^-	NVOC	NH_4	NO_2^-	NO_3^-	Ca^{2+}	Cl−	F−	Fe	K+	Mg^{2+}	Mn	SO_4^{2-}	Temp.
2065	1	2	1	1	2	NaN	NaN	1	2	1	1	1	1	1	NaN	2	1
14253	2	NaN	NaN	2	NaN	NaN	2	1	2	1	2	2	1	2	2	2	2
21582	2	NaN	NaN	NaN	NaN	NaN	2	2	NaN	NaN	2	2	NaN	NaN	2	NaN	NaN
44357	2	NaN	1	1	2	NaN	NaN	2	1	2	1	2	2	1	NaN	1	1
83080	NaN	2	2	1	2	NaN	NaN	NaN	1	1	1	2	2	2	NaN	2	2
83122	NaN	1	1	1	2	NaN	NaN	1	1	2	1	2	1	1	NaN	1	2
83132	2	1	2	1	2	NaN	NaN	1	1	1	1	1	2	2	NaN	2	2
83142	2	2	1	2	1	NaN	NaN	1	2	1	2	1	2	1	NaN	2	2
83339	2	2	1	1	2	NaN	NaN	2	2	2	2	1	2	2	NaN	2	2
83340	1	2	1	2	2	NaN	NaN	2	2	2	2	NaN	2	2	NaN	2	2
83344	2	2	2	2	2	NaN	NaN	2	2	2	2	1	2	2	1	2	2
83381	2	2	2	2	2	NaN	NaN	1	2	1	1	1	1	2	NaN	2	2
104361	1	2	2	1	1	1	1	1	1	1	1	2	1	1	1	1	2
106292	2	1	1	1	2	NaN	NaN	2	2	1	1	NaN	1	1	NaN	1	1
106294	1	1	2	1	1	1	1	1	2	1	1	2	1	1	NaN	1	1
106311	1	2	1	2	2	NaN	1	1	1	2	2	2	2	2	NaN	1	2
106558	1	2	1	2	2	NaN	NaN	1	1	2	1	2	1	1	NaN	2	2
106559	1	2	2	1	2	NaN	NaN	1	1	1	2	2	2	1	NaN	1	1
106648	2	1	1	2	2	NaN	NaN	1	2	1	1	1	1	1	NaN	1	1
106741	1	2	1	2	2	NaN	NaN	1	1	1	1	NaN	1	1	NaN	1	1

Table B.2: Results of the Kendall's test for Sjælland

plant id	pH	TDS	O_2	HCO_3^-	NVOC	NH_4	NO_2^-	NO_3^-	Ca^{2+}	Cl−	F−	Fe	K+	Mg^{2+}	Mn	SO_4^{2-}	Temp.
2065	1	2	1	1	2	NaN	NaN	1	2	1	1	1	1	1	NaN	2	1
20059	1	2	2	2	2	NaN	NaN	1	2	2	2	2	1	2	NaN	2	1
25132	1	1	2	1	2	1	NaN	2	1	1	2	NaN	1	2	NaN	1	2
28315	2	2	1	1	2	NaN	NaN	1	1	2	1	1	1	1	NaN	1	2
55791	2	1	2	1	2	NaN	NaN	1	2	1	2	1	2	1	NaN	2	1
55874	2	2	2	2	2	1	NaN	2	2	2	2	1	2	1	NaN	2	2
102998	1	1	1	1	1	NaN	NaN	1	1	1	1	2	2	1	NaN	1	1
103000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
103001	1	NaN	1	NaN	2	NaN	NaN	2	NaN	2	2	2	NaN	NaN	NaN	2	2
103320	1	2	1	2	2	2	1	1	2	1	2	2	2	2	1	2	NaN
103369	1	2	2	2	2	NaN	NaN	1	1	2	1	2	2	NaN	NaN	2	2
103436	1	2	NaN	NaN	NaN	1	NaN	2	NaN	1	NaN	2	NaN	NaN	1	2	2
103604	2	1	2	2	1	NaN	2	1	2	2	2	1	2	1	NaN	1	1
103653	1	2	2	2	2	2	NaN	1	2	2	2	2	2	2	NaN	1	NaN
103743	1	2	2	1	1	NaN	2	2	2	1	2	1	2	1	1	2	2
103821	1	2	2	2	2	1	2	2	1	1	2	1	2	2	NaN	2	1
103982	2	2	2	2	2	1	NaN	2	1	1	2	2	2	2	NaN	2	2
104361	1	2	1	1	1	1	1	1	1	1	1	2	1	1	1	1	2
104593	1	2	1	1	2	NaN	NaN	1	1	1	2	1	2	2	NaN	1	2
105008	2	2	1	1	2	1	NaN	1	1	2	2	NaN	1	1	NaN	1	2

Table B.3: Results of the Kendall's test for Midtjylland

plant id	pH	TDS	O_2	HCO_3^-	NVOC	NH_4	NO_2^-	NO_3^-	Ca^{2+}	Cl−	F−	Fe	K+	Mg^{2+}	Mn	SO_4^{2-}	Temp.
62531	1	2	1	2	2	2	NaN	1	2	2	2	NaN	1	2	NaN	1	2
63364	1	2	1	1	2	NaN	NaN	2	1	1	1	1	1	2	1	2	1
72322	1	2	2	2	2	2	NaN	2	1	2	1	2	2	1	NaN	1	2
72323	2	2	1	2	2	1	2	2	2	2	1	NaN	2	1	NaN	1	2
73312	1	1	1	1	2	2	NaN	2	2	2	1	1	1	1	NaN	1	1
78203	2	2	1	1	2	NaN	2	1	1	1	2	NaN	2	2	NaN	2	1
79379	1	1	1	2	2	2	NaN	1	1	1	2	NaN	2	2	NaN	1	2
79983	1	1	1	1	1	NaN	NaN	NaN	1	NaN	NaN	1	NaN	2	NaN	1	1
80562	1	2	1	2	2	2	NaN	1	2	2	2	1	2	2	NaN	1	2
80565	1	2	1	2	2	1	NaN	1	1	1	1	1	2	1	NaN	2	2
80567	1	2	1	1	1	1	NaN	NaN	2	1	1	1	1	2	NaN	1	1
80568	1	2	1	1	2	1	NaN	NaN	1	2	1	2	2	2	NaN	1	1
80571	1	1	NaN	2	1	1	NaN	2	1	1	1	NaN	1	1	NaN	1	1
80572	1	2	1	2	2	1	NaN	2	2	2	1	2	2	2	NaN	1	2
80573	1	1	1	2	2	1	NaN	2	1	1	1	1	2	2	NaN	1	2
80578	1	2	1	1	2	1	NaN	1	2	2	1	NaN	2	NaN	NaN	1	1
80582	1	1	1	1	2	2	NaN	1	1	1	1	NaN	1	1	NaN	1	1
97057	1	2	2	1	1	NaN	NaN	NaN	1	2	NaN	NaN	NaN	2	NaN	1	1
97059	1	2	2	2	2	NaN	NaN	NaN	2	NaN	NaN	NaN	NaN	2	NaN	2	1
97822	1	2	1	2	2	2	NaN	NaN	1	2	NaN	2	2	2	NaN	2	2

TABLE B.4: Results of the Kendall's test for Syddanmark

plant id	pH	TDS	O_2	HCO_3^-	NVOC	NH_4	NO_2^-	NO_3^-	Ca^{2+}	Cl−	F−	Fe	K+	Mg^{2+}	Mn	SO_4^{2-}	Temp.
50805	1	1	1	2	2	NaN	NaN	1	1	1	1	1	1	1	NaN	1	2
50806	1	1	1	2	2	NaN	NaN	1	2	1	1	2	2	1	NaN	2	2
50808	2	2	NaN	2	1	NaN	NaN	NaN	2	1	2	1	2	2	NaN	2	2
50813	2	1	1	2	1	1	NaN	NaN	2	1	2	2	2	1	NaN	1	1
51182	1	1	2	1	1	NaN	NaN	1	2	2	NaN	NaN	1	2	NaN	1	2
51183	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
53083	1	2	1	1	2	NaN	NaN	1	1	2	2	1	1	1	NaN	1	1
53767	2	NaN	2	2	NaN	2	NaN	2	1	NaN	NaN	2	NaN	2	NaN	NaN	2
53768	2	2	1	2	2	NaN	NaN	2	1	1	2	2	2	2	NaN	1	2
54237	1	2	1	1	1	2	NaN	NaN	2	1	1	1	2	2	NaN	2	2
74496	2	2	1	2	2	1	NaN	NaN	2	2	1	2	2	2	NaN	2	2
74497	2	2	2	2	1	1	NaN	1	2	1	1	1	2	1	NaN	1	2
81725	1	2	1	2	1	NaN	NaN	1	1	2	2	1	2	2	NaN	1	1
82001	2	2	1	2	2	2	NaN	1	2	1	2	2	1	1	NaN	2	1
82003	2	1	1	2	1	2	NaN	1	1	1	1	NaN	1	2	NaN	1	1
82010	1	1	1	1	2	2	1	2	2	1	2	NaN	2	2	NaN	1	1
82014	2	1	1	2	2	1	NaN	1	2	2	2	2	2	NaN	NaN	1	1
82440	NaN	2	2	1	1	2	NaN	1	1	1	2	NaN	2	2	1	1	2
82548	1	2	2	2	1	NaN	NaN	1	2	1	2	2	1	2	NaN	1	2
119121	1	1	2	1	1	NaN	NaN	NaN	1	1	2	1	1	2	1	2	1

Table B.5: Results of the Kendall's test for Nordjylland

plant id	pH	TDS	O_2	HCO_3^-	NVOC	NH_4	NO_2^-	NO_3^-	Ca^{2+}	Cl−	F−	Fe	K+	Mg^{2+}	Mn	SO_4^{2-}	Temp.
62997	1	NaN	1	1	1	NaN	NaN	1	1	1	1	NaN	2	1	NaN	1	1
63000	1	NaN	2	1	2	NaN	NaN	1	1	1	2	NaN	1	1	NaN	1	NaN
69781	1	2	2	2	2	2	1	1	2	2	2	NaN	2	2	NaN	2	2
69952	2	1	1	2	2	NaN	NaN	1	1	2	2	NaN	2	2	2	1	1
70251	1	1	1	2	1	NaN	NaN	2	2	2	1	NaN	1	2	NaN	2	2
70252	2	2	1	2	2	NaN	NaN	2	2	2	1	NaN	2	2	NaN	1	2
70267	1	2	1	2	2	NaN	NaN	1	2	2	NaN	1	1	2	NaN	2	1
70295	2	2	2	2	2	2	2	1	1	2	2	NaN	1	1	NaN	2	1
70383	2	1	2	2	2	1	NaN	2	2	1	2	1	1	2	NaN	1	2
70385	1	2	1	2	2	1	NaN	2	1	2	2	1	2	1	NaN	1	1
70624	1	1	1	2	1	1	NaN	NaN	1	2	2	1	1	2	2	1	1
71361	1	1	1	1	2	2	NaN	1	1	1	1	1	1	1	1	1	1
71363	2	2	1	2	2	2	NaN	2	2	2	2	2	1	2	NaN	2	2
71401	2	2	2	2	2	1	NaN	1	2	1	2	1	2	1	NaN	1	2
71402	2	2	2	2	2	1	NaN	1	1	2	2	1	2	1	1	1	2
71586	2	2	2	2	2	1	NaN	2	2	1	1	2	2	2	NaN	1	2
71639	1	1	1	1	2	2	NaN	1	1	1	1	2	1	1	NaN	2	2
71977	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
72114	2	1	2	1	2	2	NaN	2	1	1	2	1	1	1	NaN	1	2
126565	2	2	2	2	2	2	1	1	2	2	1	NaN	2	2	2	2	2

Appendix C

Averages

Table C.1: Averages for all 17 compounds and 20 plants selected in Hovedstaden

plant id	pH	TDS	O_2	HCO_3^-	NVOC	NH_4	NO_2^-	NO_3^-	Ca^{2+}	Cl−	F−	Fe	K+	Mg^{2+}	Mn	SO_4^{2-}	Temp.
2065	7.50	507.80	9.16	318.24	2.18	0.01	0.00	1.78	111.98	60.85	0.38	0.03	3.09	15.20	0.01	59.97	9.54
14253	7.31	NaN	0.00	335.50	3.13	NaN	0.01	0.02	110.42	47.08	0.32	2.36	3.01	13.33	0.10	47.33	9.11
21582	7.47	NaN	1.50	453.00	NaN	NaN	0.00	2.33	117.00	670.00	0.70	0.08	10.00	69.00	0.01	93.00	9.00
44357	7.36	923.33	8.83	377.67	1.54	0.05	0.01	2.48	154.05	117.97	0.43	0.05	4.21	25.67	0.01	135.36	11.53
83080	7.69	367.50	10.48	285.27	2.31	0.04	0.01	0.83	103.93	36.06	0.18	0.05	1.69	6.64	0.01	44.04	11.27
83122	7.57	403.92	7.82	353.54	2.66	0.05	0.01	2.02	99.26	35.08	0.39	0.06	3.44	16.93	0.01	28.25	10.91
83132	7.63	370.75	9.82	305.79	2.23	0.04	0.01	1.29	109.93	37.01	0.22	0.07	2.16	8.62	0.02	39.00	10.92
83142	7.70	350.29	8.68	347.16	2.73	0.02	0.01	5.45	82.90	21.23	0.39	0.04	2.68	14.96	0.01	3.28	10.12
83339	7.69	346.93	7.56	315.40	2.12	0.02	0.01	2.62	59.55	38.46	0.55	0.05	2.74	23.67	0.01	1.43	11.09
83340	7.75	358.07	9.08	340.90	1.75	0.01	0.01	1.93	62.20	30.73	0.52	0.04	2.04	27.79	0.01	1.60	10.78
83344	7.55	340.25	8.43	314.88	3.61	0.03	0.01	4.46	83.73	25.19	0.30	0.07	3.19	12.45	0.01	2.24	10.85
83381	7.81	372.62	8.74	338.25	3.05	0.03	0.01	2.59	69.58	43.30	0.57	0.07	4.03	24.07	0.01	5.47	10.66
104361	7.34	565.05	7.14	396.91	1.38	0.08	0.02	2.96	90.54	104.16	1.29	0.04	4.30	35.56	0.01	58.64	10.35
106292	7.66	509.56	9.00	300.64	1.50	0.02	0.01	1.44	106.59	55.94	0.42	0.12	2.89	18.12	0.02	72.85	10.43
106294	7.76	429.48	9.54	318.50	2.66	0.03	0.01	2.27	92.84	42.41	0.53	0.07	3.72	18.03	0.02	45.95	12.25
106311	7.66	473.72	8.91	293.56	5.85	0.01	0.01	2.09	116.81	51.38	0.21	0.06	2.61	9.92	0.01	50.99	10.04
106558	7.61	492.93	8.94	381.52	2.16	0.06	0.01	1.52	116.77	61.20	0.40	0.23	3.61	18.15	0.02	39.46	9.88
106559	7.63	461.36	9.80	335.42	1.82	0.06	0.02	0.90	106.07	54.29	0.33	0.22	3.15	14.49	0.02	37.14	11.08
106648	7.65	751.07	8.92	416.88	2.44	0.05	0.01	1.51	115.73	126.76	0.68	0.05	6.72	42.77	0.01	108.41	10.58
106741	7.53	392.14	8.81	358.92	2.28	0.05	0.02	1.32	109.22	83.52	0.40	0.03	3.58	25.21	0.01	57.81	10.08

TABLE C.2: Averages for all 17 compounds and 20 plants selected in Sjælland

plant id	pH	TDS	O_2	HCO_3^-	NVOC	NH_4	NO_2^-	NO_3^-	Ca^{2+}	Cl−	F−	Fe	K+	Mg^{2+}	Mn	SO_4^{2-}	Temp.
2065	7.50	507.80	9.16	318.24	2.18	0.01	0.00	1.78	111.98	60.85	0.38	0.03	3.09	15.20	0.01	59.97	9.54
20059	7.35	480.00	8.66	341.28	1.47	0.01	0.01	1.42	109.58	34.37	0.50	0.02	3.04	21.11	0.01	80.06	9.19
25132	7.45	625.11	8.75	357.16	2.84	0.07	0.00	2.79	119.36	106.64	0.58	0.01	5.77	21.61	0.01	87.15	9.16
28315	7.52	611.40	7.70	377.05	1.78	0.02	0.01	4.05	94.33	102.31	0.79	0.02	5.20	23.09	0.01	60.27	9.22
55791	7.63	850.91	9.64	402.95	1.73	0.04	0.01	3.14	97.12	208.07	1.03	0.06	8.82	36.36	0.01	50.60	11.73
55874	7.95	390.27	10.84	365.68	2.05	0.03	0.02	2.40	69.05	45.68	0.73	0.04	3.74	27.94	0.01	2.82	9.25
102998	7.67	505.94	8.68	411.82	3.10	0.02	0.01	4.10	88.11	64.28	0.36	0.06	5.71	17.12	0.01	22.06	11.23
103000	7.90	NaN	8.70	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.10	NaN	NaN	NaN	NaN	9.80
103001	7.52	396.50	9.00	380.50	2.94	0.01	0.01	3.55	80.00	58.15	0.37	0.06	3.90	12.00	0.01	13.84	11.09
103320	7.94	363.80	9.91	349.50	2.00	0.05	0.01	2.25	85.93	24.36	0.47	0.08	4.89	16.39	0.01	10.35	9.83
103369	7.71	412.48	8.87	359.14	2.41	0.09	0.01	2.52	95.53	33.30	0.37	0.09	3.62	12.33	0.01	20.92	9.39
103436	8.09	397.22	NaN	NaN	NaN	8.01	NaN	4.54	NaN	38.27	NaN	5.29	NaN	NaN	1.04	414.50	11.21
103604	7.74	593.05	10.87	358.13	2.40	0.01	0.01	2.26	99.75	134.90	0.42	0.02	5.23	15.44	0.01	38.32	9.45
103653	7.75	400.04	7.96	376.72	2.56	0.01	0.01	2.06	82.07	30.11	0.48	0.03	4.33	16.46	0.01	12.84	9.33
103743	7.64	554.75	9.61	347.45	1.72	0.04	0.01	1.65	96.68	86.90	0.45	0.06	4.75	15.72	0.01	79.07	9.27
103821	7.49	634.89	7.51	396.37	2.33	0.07	0.01	3.21	91.42	148.34	0.43	0.03	6.08	17.06	0.01	22.61	9.18
103982	8.00	409.25	9.54	348.25	2.40	0.04	0.01	2.74	93.20	49.87	0.30	0.11	3.53	11.25	0.01	17.40	9.20
104361	7.33	565.04	7.32	395.57	1.41	0.08	0.02	2.91	91.59	102.50	1.26	0.04	4.27	34.73	0.01	59.52	10.42
104593	7.73	617.15	10.42	364.96	3.23	0.01	0.00	6.65	122.88	88.28	0.62	0.03	4.52	25.91	0.01	81.07	11.32
105008	7.47	544.72	9.52	363.04	1.74	0.01	0.00	6.78	118.11	60.87	0.68	0.02	7.80	25.07	0.01	65.29	11.03

Table C.3: Averages for all 17 compounds and 20 plants selected in Midtjylland

plant id	pH	TDS	O_2	HCO_3^-	NVOC	NH_4	NO_2^-	NO_3^-	Ca^{2+}	Cl−	F−	Fe	K+	Mg^{2+}	Mn	SO_4^{2-}	Temp.
62531	7.87	290.75	9.50	181.89	0.89	0.01	0.01	1.04	54.20	179.28	0.14	0.04	3.26	5.66	0.01	18.80	11.44
63364	7.75	251.03	9.08	124.70	0.72	0.01	0.01	2.22	58.72	28.83	0.11	0.06	1.01	3.10	0.01	47.75	10.83
72322	7.80	388.33	9.68	307.07	1.26	0.07	0.01	1.87	85.73	30.08	0.24	0.29	3.95	11.00	0.06	30.50	10.71
72323	7.48	373.33	8.68	282.76	1.42	0.08	0.01	1.99	71.69	45.78	0.25	0.14	3.61	9.66	0.02	23.59	10.90
73312	7.79	331.46	9.51	255.51	1.33	0.04	0.01	0.84	83.24	22.40	0.21	0.26	2.56	7.67	0.65	43.56	9.48
78203	7.67	391.00	5.06	217.30	1.24	0.02	0.01	14.51	91.21	34.36	0.21	0.03	2.03	12.33	0.01	72.04	8.89
79379	7.65	396.40	4.81	210.23	0.75	0.03	0.01	10.28	93.00	41.54	0.22	0.02	2.64	7.47	0.02	66.33	10.76
79983	7.75	173.80	8.34	69.04	0.63	0.01	0.01	1.16	36.69	20.26	0.09	0.14	0.99	2.45	0.01	41.91	9.65
80562	7.74	527.87	8.02	323.18	1.46	0.06	0.01	1.25	101.80	98.45	0.35	0.14	3.85	13.47	0.03	43.69	9.37
80565	7.72	427.37	7.60	297.76	1.37	0.07	0.01	1.10	88.36	39.35	0.44	0.18	3.34	10.56	0.05	60.91	8.96
80567	7.77	477.50	8.61	263.84	1.48	0.06	0.01	0.82	110.25	34.97	0.21	0.16	2.99	10.21	0.04	103.89	8.93
80568	7.72	385.41	8.13	313.76	1.47	0.06	0.01	1.02	93.09	27.16	0.25	0.22	3.56	11.31	0.03	39.47	9.23
80571	7.96	405.43	9.34	246.67	1.38	0.07	0.01	0.82	95.67	36.92	0.23	0.24	2.51	9.89	0.04	73.76	8.79
80572	7.82	334.58	8.95	295.13	1.47	0.07	0.01	0.84	85.75	21.55	0.24	0.14	2.51	8.12	0.05	19.74	8.68
80573	7.77	349.32	8.56	323.60	1.65	0.10	0.01	0.87	88.47	23.07	0.23	0.35	2.82	9.39	0.05	9.87	9.26
80578	7.72	411.85	9.00	357.12	1.43	0.11	0.01	1.13	96.32	28.92	0.24	0.26	4.39	14.36	0.05	32.98	9.20
80582	7.76	351.59	8.39	329.38	1.64	0.11	0.01	1.33	86.15	22.57	0.24	2.16	4.14	12.72	0.05	15.85	8.84
97057	7.75	227.80	8.55	164.28	0.99	0.02	0.01	0.49	53.36	21.13	0.11	0.43	1.59	4.42	0.03	18.32	8.71
97059	8.01	231.25	9.08	209.61	0.96	0.02	0.00	0.51	61.00	16.11	0.11	0.47	1.64	4.77	0.04	4.87	8.98
97822	7.89	287.80	10.55	188.11	1.24	0.02	0.00	0.48	67.23	28.17	0.11	0.09	1.57	6.16	0.02	39.90	8.77

TABLE C.4: Averages for all 17 compounds and 20 plants selected in Syddanmark

plant id	pH	TDS	O_2	HCO_3^-	NVOC	NH_4	NO_2^-	NO_3^-	Ca^{2+}	Cl−	F−	Fe	K+	Mg^{2+}	Mn	SO_4^{2-}	Temp.
50805	7.87	228.48	9.70	143.67	0.72	0.01	0.01	5.63	44.71	32.26	0.11	0.09	2.23	5.98	0.01	15.58	9.54
50806	7.88	235.56	9.94	147.72	0.67	0.01	0.01	5.85	48.59	30.98	0.12	0.08	1.86	5.31	0.01	18.01	9.32
50808	7.68	222.32	11.08	123.48	0.47	0.01	0.01	0.63	48.77	31.44	0.15	0.07	1.59	3.70	0.01	33.25	8.89
50813	7.76	189.30	9.93	86.21	0.35	0.01	0.01	0.35	35.06	31.19	0.12	0.09	1.65	4.23	0.01	37.02	8.78
51182	7.86	176.86	10.12	149.45	0.76	0.01	0.01	2.55	35.39	15.82	0.09	0.03	2.26	8.03	0.01	4.72	9.41
51183	7.84	188.33	10.42	135.00	0.84	0.01	0.01	7.43	36.00	18.33	0.10	0.03	2.73	8.22	0.00	10.82	8.47
53083	7.69	310.00	8.93	196.92	1.42	0.06	0.01	2.66	51.98	50.87	0.12	0.10	4.87	10.70	0.02	23.90	10.59
53767	7.76	NaN	10.54	86.69	NaN	0.01	0.01	21.33	38.21	NaN	NaN	0.05	NaN	7.08	0.00	NaN	11.15
53768	7.82	226.21	10.85	76.50	0.46	0.01	0.01	15.09	35.54	37.63	0.08	0.06	2.25	7.11	0.00	43.88	9.35
54237	7.73	251.04	7.66	193.96	0.85	0.02	0.01	1.16	62.00	22.86	0.16	0.05	2.38	5.30	0.01	21.54	10.34
74496	7.73	361.82	9.93	266.25	1.13	0.03	0.01	1.84	90.24	25.26	0.22	0.11	11.78	8.73	0.01	47.68	8.84
74497	7.65	370.76	10.56	248.89	1.32	0.04	0.01	2.33	88.47	31.73	0.20	0.31	3.36	7.90	0.02	53.15	8.98
81725	7.66	575.74	9.12	344.95	2.33	0.01	0.01	3.11	102.92	108.30	0.91	0.05	5.82	19.43	0.01	44.91	11.43
82001	7.75	385.88	9.19	282.03	1.71	0.01	0.01	1.77	100.05	21.97	0.21	0.05	2.53	7.74	0.01	54.66	9.25
82003	7.60	487.29	8.95	328.10	1.86	0.02	0.02	1.97	110.66	49.72	0.49	0.02	5.01	13.80	0.01	67.69	9.24
82010	7.63	528.00	9.12	322.77	1.37	0.02	0.01	0.94	124.10	54.49	0.27	0.02	3.67	12.50	0.01	88.88	10.00
82014	7.65	399.31	8.57	312.39	1.69	0.04	0.01	1.50	88.86	40.72	0.30	0.04	3.74	12.04	0.01	29.95	9.21
82440	7.42	408.89	8.71	272.65	1.24	0.03	0.01	0.98	102.97	27.43	0.23	0.03	2.55	10.00	0.02	60.26	9.99
82548	7.57	450.63	9.04	301.17	2.27	0.01	0.01	3.24	114.29	28.96	0.21	0.04	2.63	8.87	0.01	65.71	9.34
119121	7.58	452.21	9.68	224.93	2.48	0.02	0.01	0.84	91.01	36.51	0.19	0.18	1.48	3.79	0.03	50.45	10.79

Table C.5: Averages for all 17 compounds and 20 plants selected in Nordjylland

plant id	pH	TDS	O_2	HCO_3^-	NVOC	NH_4	NO_2^-	NO_3^-	Ca^{2+}	Cl−	F−	Fe	K+	Mg^{2+}	Mn	SO_4^{2-}	Temp.
62997	7.56	NaN	9.82	206.29	0.70	0.01	0.00	22.34	77.87	33.46	0.13	0.04	0.88	3.00	0.01	10.87	9.63
63000	7.52	NaN	8.56	220.56	0.89	0.01	0.00	27.90	87.33	40.93	0.12	0.03	1.80	4.36	0.01	18.30	8.98
69781	7.77	331.44	8.23	191.89	0.73	0.01	0.01	36.22	79.67	33.50	0.14	0.01	1.27	4.34	0.00	22.06	8.50
69952	7.73	317.81	6.03	165.86	0.79	0.04	0.02	11.77	67.76	49.92	0.21	0.04	1.84	4.87	0.03	44.67	9.77
70251	7.70	336.54	5.29	196.02	0.70	0.01	0.01	20.41	101.38	44.64	0.20	0.02	2.04	6.06	0.01	25.92	10.06
70252	7.72	307.30	6.84	194.52	0.71	0.01	0.01	23.27	73.81	25.75	0.15	0.04	1.28	5.58	0.00	16.82	9.68
70267	7.29	649.00	5.11	257.52	1.15	0.08	0.01	32.69	122.26	122.05	0.16	0.15	3.70	16.31	0.01	76.15	10.60
70295	7.72	364.78	6.75	212.00	0.80	0.01	0.01	32.18	83.45	43.24	0.14	0.02	6.70	7.14	0.00	24.58	8.49
70383	7.91	210.17	10.63	122.25	0.92	0.01	0.01	1.81	37.38	41.68	0.16	0.04	2.26	6.06	0.01	14.28	10.09
70385	7.89	202.70	10.24	109.70	0.79	0.01	0.00	1.36	38.08	36.42	0.15	0.04	1.60	5.62	0.01	19.64	10.27
70624	7.94	586.23	10.24	173.82	19.44	0.11	0.01	12.35	60.00	194.88	0.20	0.36	6.63	13.56	0.03	41.61	9.41
71361	7.80	370.66	7.98	172.39	1.31	0.07	0.01	7.09	74.12	52.50	0.09	0.15	2.64	12.25	0.07	71.07	10.41
71363	7.97	221.67	10.00	145.86	1.09	0.05	0.01	3.21	43.46	26.38	0.16	0.02	2.74	9.67	0.00	30.75	8.79
71401	7.92	554.53	9.03	232.65	2.23	0.06	0.01	4.62	83.24	106.28	0.14	0.04	4.13	18.25	0.02	91.04	9.74
71402	7.95	442.21	8.95	252.95	2.69	0.22	0.01	5.47	70.08	94.27	0.17	0.05	4.42	15.42	0.04	39.64	9.43
71586	8.00	278.00	10.53	164.08	1.10	0.01	0.01	3.74	62.52	29.16	0.18	0.03	1.52	7.78	0.01	43.80	7.78
71639	7.91	393.24	8.59	130.21	1.02	0.08	0.12	3.97	71.90	43.75	0.20	0.02	3.59	11.22	0.02	92.68	9.20
71977	7.73	230.00	10.00	151.00	NaN	0.00	0.00	1.30	48.00	22.00	0.15	0.03	2.00	6.00	0.00	20.00	10.00
72114	7.97	294.54	10.57	143.41	1.01	0.01	0.00	0.75	59.23	35.24	0.12	0.04	1.70	10.12	0.02	65.91	8.84
126565	7.61	371.00	4.74	194.80	0.54	0.01	0.00	21.91	76.60	50.11	0.25	0.01	2.06	5.90	0.00	28.44	8.95

Appendix D

SQL statement

D.1 Plant selection

SELECT First(DRWPLANT.XUTM32EUREF89) AS FirstOfXUTM32EUREF89, First(DRWPLANT.YUTM32EUREF89) AS FirstOfYUTM32EUREF89, DRWPLANT.PLANTID, Avg(WRRCATCHMENT.AMOUNT) AS AvgOfAMOUNT, Last(DRWPLANT.PLANTNAME) AS LastOfPLANTNAME FROM (DRWPLANT INNER JOIN DRWPLANTCOMPANYTYPE ON DRWPLANT.PLANTID = DRWPLANTCOMPANYTYPE.PLANTID) INNER JOIN WRRCATCHMENT ON DRWPLANT.PLANTID = WRRCATCHMENT.PLANTID	Relationship among the selected tables (inner join)
GROUP BY DRWPLANT.PLANTID	Grouping criteria (per plant)
HAVING (((First(DRWPLANT.XUTM32EUREF89)) Not Like ””) AND ((Avg(WRRCATCHMENT.AMOUNT))>800000) AND ((Last(DRWPLANT.PLANTNAME)) Not Like ”*kilde*”) AND ((First(DRWPLANT.ACTIVE))=1) AND ((First(DRWPLANTCOMPANYTYPE.COMPANYTYPE))=”V01” Or (First(DRWPLANTCOMPANYTYPE.COMPANYTYPE))=”V02”))	Filter criteria
ORDER BY Avg(WRRCATCHMENT.AMOUNT) DESC;	Way to order the output (in order of discharge)

D.2 Pivot table construction

```
SELECT DRWCHEMSAMPLE.PLANTID, DRWCHEMSAMPLE.SAMPLEDATE, DRWCHEMANALYSIS.COMPOUNDNO, DRWCHEMANALYSIS.AMOUNTREPORTED FROM DRWCHEMSAMPLE INNER JOIN DRWCHEMANALYSIS ON DRWCHEMSAMPLE.SAMPLEID = DRWCHEMANALYSIS.SAMPLEID
```

Relationship among the selected tables (inner join)

```
WHERE (((DRWCHEMSAMPLE.PLANTID)=106294
Or (DRWCHEMSAMPLE.PLANTID)=14253 [...]
Or (DRWCHEMSAMPLE.PLANTID)=83142) AND
((DRWCHEMANALYSIS.COMPOUNDNO)=41 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=125 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=251 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=305 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=380 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=1011 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=1051 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=1176 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=1551 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=1591 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=2022 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=2041 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=2056 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=2081 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=2086 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=2142 Or
(DRWCHEMANALYSIS.COMPOUNDNO)=9902))
```

Criteria on the tracking numbers of the selected plants and compounds (replace the plants number to gain a different plant selection)

D.3 Procedure to gain the input table for the Matlab code

1. Download the 5 regional databases from `http://data.geus.dk/JupiterWWW/downloadpcjupiter.jsp?xl=1`, and extract them from the compressed folder.

2. Open the database in Access, create a new query. In the SQL view, copy and paste the SQL statement "plant selection" given in this paragraph. Run the query and adjust the threshold on the effluent discharge in order to select the wanted plants for trials and errors. Save the query as *plant_selection_[RegionName].xls*.

3. Create another new query. In the SQL view copy and paste the SQL statement "pivot table construction"' given in this paragraph. Change manually the number of the plants (those marked with the red colour) inserting the ones gained from the previous query.

4. Move to PivotTable view. Select the fields to include in the table: in the rows from the left PLANTID, years and months; in the columns COMPOUNDNR and in the body of the table AMOUNTREPORTED. Deselect the subtotals from all the fields included. Form PivotTable Tools select Export to Excel. Save the table in the same directory where the Matlab file is stored as $PivotTable\_[RegionName]$.

5. Open the table in Excel. Remove the first two empty rows. Drop again the field AMOUNTREPORTED in the table. In Field *list/values* select Value Field Settings and change Sum of AMOUNTREPORTED in Average of AMOUNTREPORTED. In $Fieldlist/Rowlabel/Months/Field\ Settings/Layoutandprint$ select *Show items* with no data. In the column Years, in the filter drop-down menu deselect all dates before 1980 and after 2012. Save the resultant table (it is necessary that all fields are extended (not reduced as a single row) before the save).

Appendix E

Kendall's statistic hypothesis testing

Conclusions on the probability of a linear trend presence is expressed in statistics as a two-tailed hypothesis test, where the independence of X and Y represents the null hypothesis ($H_0 : \tau = 0$), while the alternative hypothesis $H_1 : \tau \neq 0$ presents when X and Y are correlated. It is a two-tailed test because the test statistic could assume either positive or negative values. Theoretically, since Kendall's tau depends only on the order of pairs, the p-value associated with each possible value of τ can be always computed. One rank order is taken as a reference point and with two rank orders provided on n objects, there are $n!$ possible outcomes (one for each possible order) to consider for computing the sampling distribution of τ. That can be always calculated as the dimension of the rank is finished but it would require to compute $n!$ correlation coefficients, so a too heavy computational load. Nevertheless, it is still possible to evaluate the p-value because the sampling distribution of τ asymptotically converges to a normal distribution with mean 0 and variance $\sigma_\tau^2 = (2(2n+5))/(9n(n-1))$. Hence, for n larger than 10, τ can be reduced into a Z value

$$Z_\tau = \frac{\tau}{\sigma_\tau} = \frac{\tau}{\sqrt{\frac{2(2n+5)}{9n(n-1)}}}$$

that follows a standard normal distribution. The meaning of the p-value, which will be computed later on, is now presented. Let α=5% the given level of significance, taking into the cumulative density function of the normal distribution the calculated value of Z_τ along the x axis, the probability of that particular value of τ is given by the correspondent ordinate. That value have to be compared with $\alpha/2$, because this is a two-tailed test:

- If p-value= $cdf(-|Z_\tau|) < \alpha/2$ then you reject the null hypothesis $H_0 : \tau = 0$, and a trend of the considered compound in time is likely.
- If p-value= $cdf(-|Z_\tau|) > \alpha/2$ then you fail to reject the null hypothesis, the evidence is insufficient to support a conclusion and probably there is no trend of data in time.

In fact, in the cumulative probability function, $\alpha/2$ taken as an ordinate (x_{crit}) represent the area subtended under the standard normal pdf in the left tail of the graph. Thus if the value $-|Z_\tau|$ fall at the left of x_{crit} the statistic parameter is out of the acceptance area, and you have to reject H_0. The absolute value of Z is considered because the distribution is symmetric and the comparison is always done with the left tail.

Appendix F

Compounds' description

In this section all analysed compounds will be presented one by one with reference to their effects on health, corrosion and taste in addition to the WHO and Danish guidelines (Gorchev and Ozolins [1984], Ikrafttr [2014]).

F.1 Ammonia

With ammonia both non-ionized (NH_3) and ionized (NH_4^+) species are named. It originates from metabolic, agricultural and industrial processes and from disinfection with chloramine. In groundwater and surface water, natural level of ammonia is usually below 0.2 mg/l. Concentrations up to 10 mg/L appears in water contaminated with animal waste. Its content can increase due to contamination of source water or through microbial metabolism.

Health effect: Above daily levels of 32 mg ammonium per kilogram body weight it shifts acid-base equilibrium in human body, with impacts of glucose tolerance and tissue sensitivity to insulin. Ammonia occurs in drinking water at levels with no relevance for health because only an extremely small portion of ammonia originates directly from food or water.

Corrosion effect: ammonia influences the disinfection efficiency (not applied in Denmark) in distribution systems and compromises the removal of manganese. It causes copper stains on sanitary ware. If its concentration reaches 0.5 mg/l corrosion of copper pipes and fittings may occur.

Taste effect: Ammonia can cause taste and odour problems if the level is above 1.5 mg/l because it can compromise the disinfection efficiency and the removal of manganese.

Limits: basing on health considerations, WHO does not establish any guideline value, while Australian Water Guideline suggests 0.5 mg/l as the maximum concentration. Ammonia should not exceed 0.05 mg/l for the Danish law.

F.2 Calcium

See hardness.

F.3 Hardness

Water's hardness is the concentration of multivalent cations in water, predominantly calcium (Ca^{2+}) and magnesium (Mg^{2+}) ions expressed as a calcium carbonate equivalent. Hardness of water measures the capacity of water to react to soap, where hard water requires more soap. For people lacking in calcium and magnesium, a right hardness level contributes to intake the proper amount of those minerals (Gorchev and Ozolins [1984]).
Health effect: epidemiological studies show a protective effect of magnesium or hardness on cardiovascular mortality, but available data are insufficient to define a minimum or a maximum level. Hardness levels below 75 mg/l , may negatively affect mineral balance.
Corrosion effect: Australian guideline identifies the following categories of hardness (water quality management [2011]):

<60 mg/L $CaCO_3$	Soft but possibly corrosive
60-200 mg/L $CaCO_3$	Good quality
200-500 mg/L $CaCO_3$	Increasing scaling problems
>500 mg/L $CaCO_3$	Severe scaling

In addition to hardness level, also pH, alkalinity and dissolved oxygen concentration influence pipe corrosion. **Taste effect**: -
Limits: No limit is suggested from WHO directives. Australian guideline suggests levels above 200 mg/l to minimize objectionable build-up of scale. For the Danish law water hardness should be between 5 and 30 dH.

F.4 Iron

In water iron occurs in oxides forms as ferric (Fe(III)) or ferrous (Fe(II)) compounds. In natural fresh water, iron levels range between 0.5 to 50 mg/l.

Health effect: Iron requirement varies with age and sex, but not enough data are available to evaluate a health-based guideline. Nevertheless iron does not become a health concern unless the concentration is above 3 mg/l.

Corrosion effect:iron level can increase due to the use of iron coagulants and corrosion of steel and cast irons pipes in the water distribution system.

Taste effect:the test threshold for iron is about 0.3 mg/l in drinking water and become undesirable above 3 mg/l. Rust-brown colour appears in water when iron high level occurs.

Limits: WHO does not establish any guideline value, while the Australian limit is 0.3 mg/l. The maximum level for the Danish law is 0.1 mg/l. The Danish law establishes 0.02 mg/L as maximum level.

F.5 Magnesium

See hardness.

F.6 Manganese

Manganese has three different forms: divalent (Mn(II)), tetravalent (Mn(IV)) that is insoluble, and heptavalent (Mn(VII)) that is soluble. It occurs in many of surface and groundwater aquifers, especially in anaerobic or low oxidation conditions (the most frequent).

Health effect: manganese is one of the least toxic elements, but a health-based value of 0.4 mg/l can be derived on the upper range value of manganese intake of 11 mg/day (given by dietary surveys).

Corrosion effect: Concentrations over 0.1 mg/l strains plumbing fixtures and laundry. Even at concentrations of 0.02 mg/l manganese forms a coating on pipes that can turn into black ooze.

Taste effect: manganese has an undesirable taste effect for concentrations exceeding 0.1 mg/l.

Limits: Australian guideline is 0.1mg/l based on practical experience, while WHO suggests levels up to 0.4 mg/l but no guideline value is established.

F.7 Potassium

Potassium is one of the most important elements for human diet, the suggested daily intake is greater than 3000 mg. It commonly occurs in natural waters. In drinking water it is a result of water treatment when potassium permanganate is used (not in Danish water works).

Health effect: typical potassium concentrations in municipally treated drinking water are well below those of health concern. Some health effects are noticed only in susceptible individuals.

Corrosion effect: -

Taste effect: Potassium is an aesthetic parameter; its presence influences taste, color and smell.

Limits: Danish law gives a limit of 10 mg/l.

F.8 Chloride

Chloride is often originated from disinfectant and bleach both domestically and industrially used. Surface water has usually concentration lower than 100 mg/l, although groundwater has higher concentration, especially if there is salt water intrusion.

Health effect:Chloride affects the osmotic activities of body fluid. Large salt intake leads to blood pressure increase but this change is due to sodium rather than chloride.

Corrosion effect:-

Taste effect: users can identify fell the taste of chlorine if the concentration is above 5 mg/l. (WHO) The taste threshold of chloride varies with the associated cation between 200-300 mg/l.

Limits: the WHO guideline limit is 5 mg/l. This level derives is the whole TDI (total daily intake) for an adult with a weight of 60 kg, who take 2 litres/day. While basing on aesthetic assessment, Australian guidelines give a threshold of 250 mg/l. The limit given by the Danish law is 250 mg/L.

F.9 Fluoride

Fluoride occurs in high concentrations in groundwater. Because of the fluoride-containing minerals, well water can reach fluoride levels up to 10 mg/l.

Health effect: the content of fluoride has several consequences on health. Indeed, it rapidly absorbed from the body and it is quickly incorporated into teeth and bones, with practically no storage in soft tissues. It helps to prevent dental caries in a range between

0.5 mg/l and 2 mg/l. Nevertheless, it can lead to dental fluorosis if levels are in a range between 0.9 mg/l and 1.2 mg/l. Fluoride content between 3 and 6 mg/l causes skeletal fluorosis that becomes crippling above 10 mg/l. Sign of intoxication appear for doses of about 1 mg per kilogram of body weight (Gorchev and Ozolins [1984]). **Corrosion effect**: -
Taste effect: -
Limits: WHO guideline and Danish law has 1.5 mg/l as limit value.

F.10 Hydrogen carbonate

Not mentioned both in WHO guidelines and in Danish prescription.

F.11 Nitrite

Nitrite may be originated from the microbial reduction of nitrate. Otherwise it derives from the stagnation of nitrate-containing and oxygen-poor drinking water in galvanized steel pipes.
Health effect:methaemoglobinaemia in bottle-fed infants (short-term exposure) for levels above 3 mg/L.
Corrosion effect:-
Taste effect:-
Limits:WHO guideline value is 3 mg/l as nitrite ion (or 0.9 mg/l as nitrite-nitrogen).Danish law gives 0.01 as limit value.

F.12 Nitrate

Nitrate is a fundamental plant nutrient and it is present in the nitrogen cycle. It is mainly produced from agricultural activities (inorganic nitrogenous fertilizers and manures), but it comes also from wastewater disposal and oxidation from nitrogenous derived from several waste products. Leaching from natural vegetation influences groundwater contamination as well. Nevertheless, the most part of nitrate and nitrite incomes through vegetables and meat I the diet.
Health effect:methaemoglobinaemia in bottle-fed infants (short-term exposure)
Limits: WHO guideline value is 50 mg/l as nitrate ion and 11 mg/l as nitrate-nitrogen. Danish law gives 50 mg/l as a limit for nitrate.

F.13 Sulphate

Sulfate is often originated from natural sources. Despite food is the principal source, also air and drinking-water provide part of the daily intake of sulphate, that is about 500 mg.
Corrosion effect: sulphate contributes to pipe corrosion.
Taste effect: individuals notice the presence of sulphate in drinking-water taste.
Limits: WHO do not give any guideline but points out that concentration of 1000-1200 mg/l causes laxative effects. Danigh law gives 250 mg/L as a limit for sulphate.

F.14 pH

pH is operationally very important for water quality. Values up to 7.0 identify an acid reaction. One point decrease in pH-value means the solution became 10 times more acid, because pH is expressed in logarithmic scale.
Health effect:water with pH level below 7.0 has relevant effects on health.
Corrosion effect:pH should be between 6.5 and 8.5 to reduce corrosion in pipes and fittings (water quality management [2011]).
Taste effect:Bitter metallic taste if pH is less than 6.5. Above 8.5 it causes a slippery feeling and soda-like taste. (EPA)
Limits:No limits are suggested from WHO. Australian guidelines accept pH values varying from 6.5 to 8.5. Danish law gives the range 7-8.5 for pH.

F.15 TDS (evaporation residue)

Total Dissolved Solids (TDS) parameter identifies all inorganic salts (mainly calcium, magnesium, potassium, sodium, bicarbonates, chlorides and sulfates) and small quantities of inorganic matter dissolved in water. Sewage, industrial waste water, salts used for road de-icing, natural sources and urban runoff contribute to TDS content.
Corrosion effect: -
Health effect: -
Taste effect: Bed effect on taste.
Limits:No limits are suggested from WHO. Australian guidelines propose 600 mg/L as TDS limit for good water palatability. Danish law gives 1500 mg/L as TDS limit.

F.16 Temperature

Cool water is generally more palatable than warm water, and temperature will impact on the acceptability of a number of other inorganic constituents and chemical contaminants that may affect taste. High water temperature enhances the growth of microorganisms and may increase taste, odour, colour and corrosion problems. Temperature above 20C in drinking water makes the number of complaints to increase (water quality management [2011]).
Limits:No limits are suggested both from WHO and Australian guidelines.

F.17 Oxygen

The dissolved oxygen content of water is influenced by the source, raw water temperature, treatment and chemical or biological processes taking place in the distribution system. Depletion of dissolved oxygen in water supplies can encourage the microbial reduction of nitrate to nitrite and sulfate to sulfide. It can also cause an increase in the concentration of ferrous iron in solution, with subsequent discoloration at the tap when the water is aerated. No health-based guideline value is recommended.
Health effect: -
Corrosion effect: high levels of dissolved oxygen speed up corrosion in water pipes.
Taste effect:high levels of dissolved oxygen affect positively the taste.
Limits:For the Australian guidelines dissolved oxygen concentration in drinking water shoul be greater than 85% saturation. $>$ 8 mg/L for the Danish law.

Appendix G

Attached files

A compressed folder is given as attachment to my thesis, it is named *ThesisFile.zip*. The list of its content to follow:

- **Folder** *databases*: It contains the databases downloaded from the Jupiter web page with the implemented queries, one for each region.
- **One folder for each region**: it contains all the implemented Matlab codes, the inputs for the program, the outputs and the graphs.
- **Excel file** *IreneThesis.xlsx*: it includes the mean spreadsheets used in the analysis.

G.1 Folders with the names of the regions

- *data.txt*, *PivotTable_[RegionName]*, *pl_selection_[RegionName]*: inputs for the program.
- *time_trend.m* and *kendf.m*: Matlab code and the attached function as described in chapter 5.
- *wspace*08_01*.mat*: workspace containing the output variables.
- Folder *measurements*: All the data sets gathered by plant; each columns have the measurements related to onu compound in the same order of the program.
- Folder *GISinput*: Outputs of the program (the values of the absolute slopes in the file *absslope.txt*, the averages in the file *avercoord.txt*, the relative slopes in the file *relslopes.txt* and the average effluent of each pant in *Q_pumped.txt*.

- Folder *graphs*: graphs gathered by compound. The most important graphs are discussed in the report.

G.2 Excel spreadshits in *IreneThesis.xlsx*

- **measurements**: number of measurements recorded in Jupiter for each plant and each compound
- **kend**: results of the Kendall's test
- **relslopes**: relative changes of compound levels in a decade
- **averages**: average values of all compounds in all plants
- **absslopes**: changes of compound levels in a decade
- **plants**: names and average effluents of all plants
- **histogram**: histograms presented in the report
- **requirements**: requirements according to WHO guidelines, Danish law and Australian guidelines

Bibliography

Klage Straf Ikrafttr. Bekendtgørelse om vandkvalitet og tilsyn med vandforsyningsanlæg, 2014. URL `https://www.retsinformation.dk/Forms/R0710.aspx?id=138647`.

National water quality management. *Australian Drinking Water Guidelines 6*. Natural Resource Management Ministerial Council, 2011. ISBN 1864965118.

University of Rhode Island. Tests for Corrosiveness. Technical report, University of Rhode Island, Rhode Island. URL `http://www.uri.edu/ce/wq/has/PDFs/Testsforcorrosivenesssum.pdf`.

M Bruvo, K Ekstrand, E Arvin, H Spliid, D Moe, S Kirkeby, and A Bardow. Composition for Caries Control. *Journal of Dental Research (JDR)*, 87(Table 1):340–343, 2008. URL `http://jdr.sagepub.com/content/87/4/340.short`.

H G Gorchev and G Ozolins. WHO guidelines for drinking-water quality. *WHO chronicle*, 38(3):104–8, January 1984. ISSN 0042-9694. URL `http://www.ncbi.nlm.nih.gov/pubmed/24286868`.

N. Delion, G. Mauguin, and P. Corsin. Importance and impact of post treatments on design and operation of SWRO plants. *Desalination*, 165:323–334, August 2004. ISSN 00119164. doi: 10.1016/j.desal.2004.06.037. URL `http://linkinghub.elsevier.com/retrieve/pii/S0011916404002413`.

WHO. Total dissolved solids in Drinking-water Background document for development of WHO Guidelines for Drinking-water Quality. Technical report, WHO, Geneva, 1996. URL `http://www.who.int/water_sanitation_health/dwq/chemicals/tds.pdf`.

M Rygaard, E Arvin, a Bath, and P J Binning. Designing water supplies: Optimizing drinking water composition for maximum economic benefit. *Water research*, 45(12): 3712–22, June 2011. ISSN 1879-2448. doi: 10.1016/j.watres.2011.04.025. URL `http://www.ncbi.nlm.nih.gov/pubmed/21565384`.

Ori Lahav and Liat Birnhack. Quality criteria for desalinated water following post-treatment. *Desalination*, 207(1-3):286–303, March 2007. ISSN 00119164. doi: 10.

1016/j.desal.2006.05.022. URL `http://linkinghub.elsevier.com/retrieve/pii/S0011916407000306`.

Stefan Platikanov, Veronica Garcia, Ignacio Fonseca, Elena Rullán, Ricard Devesa, and Roma Tauler. Influence of minerals on the taste of bottled and tap water: a chemometric approach. *Water research*, 47(2):693–704, February 2013. ISSN 1879-2448. doi: 10.1016/j.watres.2012.10.040. URL `http://www.ncbi.nlm.nih.gov/pubmed/23200507`.

GEUS. Water supply in Denmark. Technical report, Danish ministry of environment, Copenhagen, 2008. URL `http://www.ecoinnovation.dk/NR/rdonlyres/E4D4BD37-82E9-413D-87D8-D6AECD6B7E79/0/Vandforsyning_artikel.pdf`.

Naturstyrelsen. Kvaliteten af det danske drikkevand for perioden 2008-2010. Technical report, Naturstyrelsen, 2012. URL `http://www.naturstyrelsen.dk/NR/rdonlyres/066551BF-0347-4B1C-851C-E48D5063A0B2/149212/Indberetningsrapportdrikkevand20082010.pdf`.

Dansk Vand-og Spildevandforening (DANVA). Water in figures. Technical report, DANVA, 2011. URL `http://www.e-pages.dk/danva/100/`.

Martin Rygaard, Erik Arvin, and Philip J Binning. The valuation of water quality: effects of mixing different drinking water qualities. *Water research*, 43(5):1207–18, March 2009. ISSN 0043-1354. doi: 10.1016/j.watres.2008.12.014. URL `http://www.ncbi.nlm.nih.gov/pubmed/19136136`.

K.R. Ekstrand, M.E.C. Christiansen, and V. Qvist. Influence of Different Variables on the Inter-Municipality Variation in Caries Experience in Danish Adolescents. *Caries Research*, 37(2):130–141, 2003. ISSN 1421-976X. doi: 10.1159/000069021. URL `http://www.karger.com/doi/10.1159/000069021`.

M. S McDonagh. Systematic review of water fluoridation. *Bmj*, 321(7265):855–859, October 2000. ISSN 09598138. doi: 10.1136/bmj.321.7265.855. URL `http://www.bmj.com/cgi/doi/10.1136/bmj.321.7265.855`.

L R Hilbert, H J Albrechtsen, and A Andersen. Effect of material and water quality on disinfection and risks of corrosion. In *The Annual Congress of the European Federation of Corrosion Eurocorr*, pages 1–17, 2010.

I.G. Grossman and L. B. Yarger. Water resources of the Rochester area, New York, 1963. URL `https://play.google.com/books/reader?id=0ZsvAAAAYAAJ&printsec=frontcover&output=reader&authuser=0&hl=it&pg=GBS.PR1`.

Thomas S Ferguson, Christian Genest, and Marc Hallin. Kendall ' s tau for autocorrelation. *Department of Statistics Papers*, 2011. URL `http://escholarship.org/uc/item/3p91609d`.

R. Huth. Testing for Trends in Data Unevenly Distributed in Time. *Theoretical and Applied Climatology*, 64(3-4):151–162, December 1999. ISSN 0177-798X. doi: 10.1007/s007040050119. URL `http://link.springer.com/10.1007/s007040050119`.

Hervé Abdi. The Kendall Rank Correlation Coefficient. In *Encyclopedia of Measurement and Statistics*, pages 508–510. Sage, Dallas, 2007. URL `https://www.utdallas.edu/~herve/Abdi-KendallCorrelation2007-pretty.pdf`.

Jens Christian Refsgaard, Anker Lajer Hø jberg, Ingelise Mø ller, Martin Hansen, and Verner Sø ndergaard. Groundwater modeling in integrated water resources management–visions for 2020. *Ground water*, 48(5):633–48, 2010. ISSN 1745-6584. doi: 10.1111/j.1745-6584.2009.00634.x. URL `http://www.ncbi.nlm.nih.gov/pubmed/19788560`.

Landskabsstyrelsen. Summary and conclusions - Filtre påmindre vandværker, 2009. URL `http://www2.blst.dk/udgiv/Publikationer/2009/978-87-7091-018-7/html/kap02.htm`.

DELTA Laboratories. Physical, Chemical and Biological Parameters, 2008. URL `http://www.adopt-a-stream.org/parameters.php`.

Laure Lopato, Zofia Galaj, Sébastien Delpont, Philip J Binning, and Erik Arvin. Heterogeneity of Rapid Sand Filters and Its Effect on Contaminant Transport and Nitrification Performance. *Journal of Environmental Engineering*, 137(April):248–257, 2011. doi: 10.1061/(ASCE)EE.1943-7870.0000321.

Mountain Empire. Causes of Corrosion. URL `http://water.me.vccs.edu/concepts/corrosioncauses.html`.