

Research Proposal 2023

Table of Contents

RESEARCH PROPOSAL 2023	1
1. RESEARCH QUESTION.....	2
2. ABSTRACT	2
3. BACKGROUND	2
4. AIMS AND OBJECTIVES.....	2
5. LITERATURE ANALYSIS	3
5.1 THE APPLICATION OF ARTIFICIAL INTELLIGENCE TECHNIQUES TO BETTER MANAGE IRON IN DRINKING WATER DISTRIBUTION SYSTEMS.....	3
5.2 INFLUENCE OF HYDRAULIC MODEL COMPLEXITY ON RESULTS OF WATER AGE AND QUALITY SIMULATION IN MUNICIPAL WATER SUPPLY SYSTEMS	4
5.3 MODELLING AND FIELD WORK TO INVESTIGATE THE RELATIONSHIP BETWEEN THE AGE AND THE QUALITY OF DRINKING WATER AT CUSTOMER'S TAPS	4
6. METHODOLOGY	6
7. ANALYSIS AND RESULTS	7
8. CONCLUSION	8
9. SUBMISSION REFLECTION	9
10. REFERENCE.....	10
11. APPENDIX	11

1. Research Question

In clean water hydraulic modelling, is turbidity modelling a better indication of water quality than water age?

2. Abstract

This research will investigate the pros and cons of using turbidity modelling in a pressurised potable water network. Currently at Welsh Water, water age would be used to identify where a water quality issue might spread. Water age is currently preferred as it's an easier calculation to make however, it does not take into account dilution and several other considerations.

3. Background

Being able to accurately predict how water quality events will travel through a water network is crucial. This allows measures to be put in place to reduce the impact on the network and customers supplied off it. On most occasions, there is normally a limited time to react when a water quality event happens, so being able to accurately and quickly predict where that water will travel is key to ensuring that the smallest number of customers are affected.

Currently the favoured method of predicting where water will travel would be to use a water age calculation, which can be done very simply using a spreadsheet or utilising modelling software such as Synergi. The benefit to using water age to predict travel is that it is easy and quick to carry out. What will be looked at is if the added time spent building a turbidity model gives more accurate travel, therefore allowing for a more informed solution. This needs to be compared with the additional time taken.

Turbidity modelling has been developed mainly by Sheffield University and their Prediction of Discolouration in Distribution Systems (PODDS) theory. Their research has become the widely accepted method for modelling water quality in a potable pressurised water network.

4. Aims and Objectives

The aim of this research is to identify if it is worth the extra time and effort to calibrate. At the moment water age is the go-to method for calculating various things, as it gives an idea of where in the water network the water should be. However, this does not add any consideration to the dilution of the water or the chlorine levels.

5. Literature analysis

As this is a fairly niche topic with clean water pressurised hydraulic modelling is still a very small industry there is not a great deal of literature around the topic, however the quality of the literature which does exist is very good. The majority of the literature around this topic have come from Sheffield University. One of the most widely used modelling software now uses the PODD's system developed by Sheffield University.

5.1 The application of artificial intelligence techniques to better manage iron in drinking water distribution systems

The Authors of this paper have used machine learning to try and find patterns using messy historical data. Using machine learning to analyse the data is very appealing as most water companies have lots of historical water event data, however it is often in a very messy unformatted form. Therefore, not being much use for a person to manually process.

As stated by the authors, the sampling of water is often at a single location at various times. The sample is normally taken at a set interval, usually once a month. The issue is that the sample then gets analysed at a laboratory, the results could be days after the sample is taken. Once a sample has failed, more frequent samples will usually be taken. However, this is after the event which caused the failure. This makes it very difficult to retrospectively find patterns.

The data outputs from the machine learning appear to be promising, figure 6 shows an excellent trend where the last 10% of areas see a sharp increased probability of having iron failures. My concern is how to verify the outputs of the machine model, or to check the model has been correctly set up. This is evident in the data showing that the presence of cast iron pipe in the nearby network is not necessarily an indication that the probability of an iron failure will be high. This goes against the current consensus.

The machine learning did not use velocities or water age, which are key indicators of water quality. This data would have been very difficult to add to the machine learning, however it is still a crucial measure of water quality issues.

5.2 Influence of Hydraulic Model Complexity on Results of Water Age and Quality Simulation in Municipal Water Supply Systems

The methodology used by the authors of this paper, is very thorough which means the test could be repeated to verify their results. However, the authors seem to have complicated the modelling aspect, and their results are not clear. Having read their paper, I am still unsure if they have answered their question.

5.3 Modelling and field work to investigate the relationship between the age and the quality of drinking water at customer's taps

This paper provides a very impressive view into the benefits of using computer models to predict the next water quality event. Improvements in affordable high-performance computers has allowed for large scale mathematical calculations to be made using a standard laptop. Modelling software has also evolved to not just model water travelling through a pressurised network. The software can now be set up to model each packet of that water. The authors of this paper look to show the benefits over previously used methods. The widely used method of calculating where and when decreased water quality would occur was to use a pen and paper method. Taking the average diameter of the mains and taking a flow reading, you can then calculate the velocity of the water, once you have the velocity you can then work out how far the water will travel per metre usually in seconds. This method is quick and easy to do however it has some serious limitations.

The authors used a water quality modelling software to compare the benefits over using just water age to predict water quality events.

The importance of accurately modelling water quality is driven by Ofwat who are the regulators for drinking water for England and Wales. The regulators are putting far greater pressure on drinking water providers to increase the quality of the drinking water they supply to customers. This is enforced with large fines if there are failures. Most water mains in the United Kingdom are ferrous, mainly cast iron. They are also like to be over a hundred years old. As the writers of this paper state the main cause of water quality issues in potable water mains is high iron content.

The authors of this paper used two consistent water networks and historical water quality data to evaluate the benefits of water age over turbidity modelling. It has been previously proven that the age of water has an impact on water quality as it has a longer contact time with the iron of the water main. *Burlingame (1995)*.

As water quality is a very complicated topic, and many different factors can contribute to water quality at customers taps. Taking this into consideration the age of the water is an excellent substitute for water quality modelling and prediction.

Over the last two decades the modelling of water quality has progressed hugely. As CPU capacity has increases it has allowed the modelling software to carry out more mathematical equations. Therefore, letting the software predict to a far greater accuracy.

Although the software can start to accurately predict water quality events, it is still reliant on a well calibrated model. To calibrate a model to be used for water quality is still very difficult as there are many factors which are difficult to accurately calibrate. The authors allude to the difficulty in being able to acquire some of the data critical to the calibration of the water quality model. Since this report was written there has been considerable gains in water quality monitoring devices, this does allow for easier pipeline coefficient calibration. However, they still have limitations, as sensors become smaller and cheaper there should be significant gains in this area.

The authors of the paper compared water age to regulatory historical water quality samples. Looking back at previous events is a good way to test how well a model works. you can look back at previous events and calibrate the model to emulate the same results. Once you can reproduce those conditions you can have a confidence in your model to predict how the water network might perform in future. This is dependent on the network remaining the same and sediment not building up on the inside of the pipe. Having a calibrated model using the methods the authors used seems a sensible approach to answer the questions of their paper. They took sample data from various parts of the network, with differing water age to see if the model can reproduce the results.

The limitations of the way the writers carried out the tests mean that this way of comparing water age with the modelling of water quality means it will only work in a very specific system set up. They have used mean water age and have assumed the water age at inlets is 0. Ideally, they should have modelled the age back from the water treatment works. as the water could be days old by the time it gets to the inlet of their selected district metered area. The other thing which could affect their results is that they do not mention the material of the trunk main system. Ideally the trunk main which supplies their chosen district metered area could already have poor water quality by the time it arrives. The authors should have had a sample point at the part of the network they are calling zero.

It would have been nice to see a broader sample of networks used, the networks that have been used had very few water quality issues. If the authors had selected an area of network that had historical water quality failures it would have been easier to compare water age with water quality failures. with the failures being so infrequent it could point to other factors for failures that are unrelated to water ages or increases in velocity.

Throughout the paper there are a lot of uncontrollable parameters which then leave the data subject to discrepancy. Because the authors used a real in use potable water network there are many changes every day. Which are very difficult to record and control. The internal condition of the network is also unknown or has not been mentioned. Cutting out sections of the network would have been good however that is expensive and has operational problems but could have given a more accurate view into the results.

The by using a small sample area the authors have limited what the research can tell us.

6. Methodology

To gather the data for this research an analysis of a water quality event was carried out. The data was required for the clean water regulator (Ofwat). Water age was used for the information required for the regulator and to ensure the samples taken at customers taps was a representative sample. For example travel times of water in a distribution network can be over 100 hours therefore, when water companies know of a water quality event they need to take samples downstream. Using water age the samples can be verified as representative, if the travel time was 60 hours and the sample was taken 12 hours after the event the sample taken would not be the same water from the event. The first sample should be taken after 60 hours at that location to ensure the correct sample is taken.

This is very useful information as it means that the samples can show the regulator that the samples are safe to drink, or shows that unsafe water was supplied to customers. Utilising turbidity modelling.

To calibrate a turbidity model the model must be calibrated, this is achieved by first understanding the material. The model is then run at peak flow which defines the pipe wall shear strength. The model before this needs to be calibrated for pressure and flow. This is achieved by using data loggers to record pressure and flow around the network. The pressure is recording at multiple points around the network, this is then added to their elevation to show if there is any head loss in the system. If the pressure logger is lower than expected then that is an indication of either a restriction or the mains diameter is smaller than in the model. The flow is calibrated based on the input to that part of the network then it is usually divided by the larger export users and each property utilising daily average flows.

7. Analysis and Results

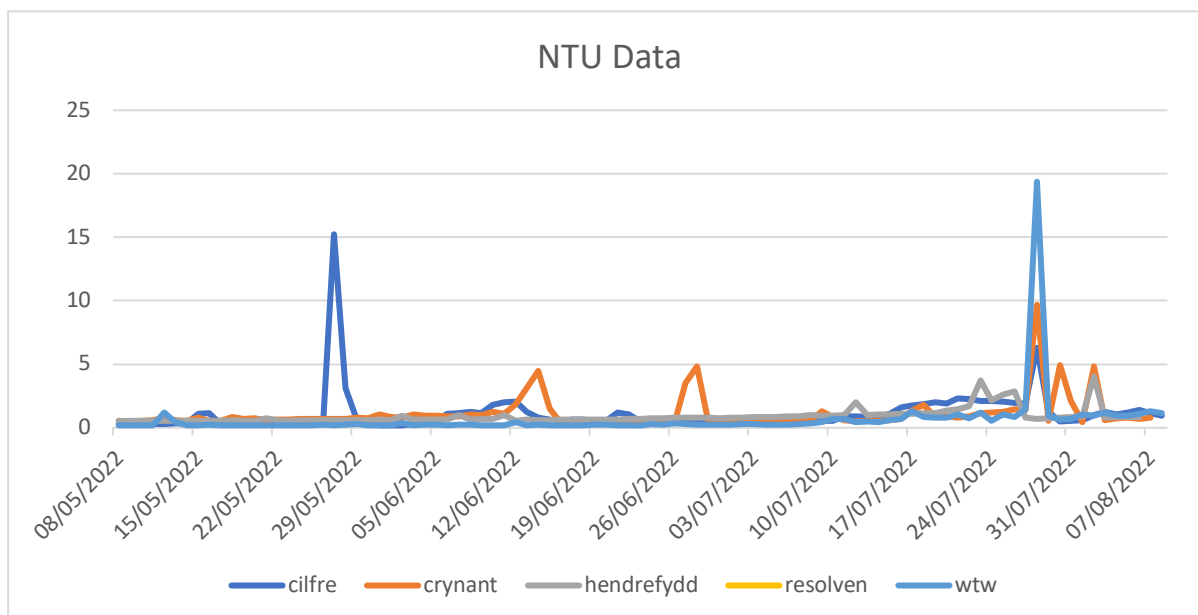


Figure 1 – NTU chart showing modelled NTU spikes compared with actual NTU spikes.

When comparing modelled turbidity data with water age it is two very different data sets that are telling you different things however, the benefits of being able to model turbidity is clearly evident.

Water age data is the current best practice for modelling and predicting water quality in a pressurised clean water network. What it shows is how long it will take for that packet of water to travel from one point to another, this is very useful information to have. For example, if there was a contamination at the start of the system, know when and where the contaminant will travel to is crucial to prevent serious issues, the contaminant could be isolated to stop it traveling downstream. It is a very simple calculation to make, and modelling software can easily and efficiently model water age and travel times.

Turbidity modelling data shows the huge benefits of being able to accurately model water quality and other contaminants is revolutionary. Where turbidity modelling excels is that it shows concentrations of what is being modelled, this could be particles, chlorine concentrations and many other things. The results give you a much greater indication of where will be affected most and how they will be affected. It allows operators to make data driven decisions on how to respond to incidents in the network. The two data sets are not really directly comparable as they are telling two different things waterage is just indicating how long it would likely take for that packet of water to get to a downstream location, the biggest negative for the water age results is that you need to know a spike in NTU's has happened and then can calculate how long it will be until it reaches downstream locations. The turbidity model and calibrate the sheer strength of the pipe and calibrate what the mains are conditioned for, then if additional flow is added and velocities increase, the model can predict if it will cause any water quality issues. It can also model the same way that water age would, by that if you know of a spike in NTU's at the start of the system you can track its travel

time and where it will travel to. It can go one step further than that, it can then tell you the dilution of that sediment in specific parts of the network. This again allows operators to make far more accurate decisions on where to focus resources and where not to. As clean drinking water is so relied upon and can easily affect millions of people know this kind of information is incredible. Where water scarcity is becoming a bigger issue with climate change the frequency of contamination issues will rise therefore, making this kind of modelling even more crucial.

8. Conclusion

There are clear benefits to being able to predict where and when water quality issues will arise. It is something which would not be required all the time and does have some specific criteria and information to model it effectively. Therefore, it is not a replacement for water age calculation instead, it is another tool to support the correct decisions.

The extra time taken to model water quality and turbidity needs to be factored in. With this there are also some assumptions of the shear stress of the sediment attached to the inside of the pipe. The same can be said for water age calculation however, as there are fewer variables it means there are fewer things to get wrong.

With improvements being made to NTU loggers all the time, that data is becoming more reliable which is critical to calibrating a turbidity model correctly.

The issue with using water age is that water age changes constantly depending on velocity. Therefore, a max and min is used to give an indication of the travel time. This kind of information is reliant on the person utilising this data to have a deep understanding of how the network works and behaves. They will then have to use judgment to predict when or if they are likely to be affected. Comparing this to turbidity modelling where it takes out the guess work. The calibrated turbidity model will give network engineers an excellent easy to use, well displayed batch of information. The model can calculate the dilution rate of whatever is being modelled. Below one NTU would not normally be detectable by a normal customer therefore, if the calibrated turbidity model suggests that the water will be less than one NTU the network operators could then choose to not waste resource on that particular part of the network and instead focus efforts to parts which will see high dilution of sediment. With this knowledge and with travel times sometimes reaching 24 hours, there is time to get people in to areas before the patch of dirty water arrives. This allows for bottled water stations to be set up, text messages to be sent out. Vulnerable customers can have water delivered to their door. Boil notices can be issued and this all happens before the potentially contaminated water arrives. This has the potential to save people's lives although this is a very rare event, they do happen and people do catch infections like cryptosporidium which can cause long lasting health issues or death. The model would not tell you exactly where would be exposed to something like cryptosporidium but it would give an indication as to where might see it. This is a huge improvement compared to current methods which would be everything from the outlet of a water treatment works which could be hundreds of thousands of customers.

This is only the start of turbidity modelling, water companies around the United Kingdom are starting to utilise communication devices to feed live data into control rooms. Currently

most will have data updates every 30mins. Currently most companies use a calibrated model which could have been built many years before hand. With advancements in machine learning and data being more available and cheaper the future of modelling is likely to be a live model. Which will be linked to customers meters which then sends data back to the model the model can then automate the network for you. There are water companies in the united kingdom currently trialling this technology on small sample areas. The limitations is having good up to date data. For water companies such as Thames Water data collection is easier due to good phone reception and the proximity of their customer base. In comparison to Dwr Cymru Welsh Water who's customers range from densely populated in Cardiff to rural Pembrokeshire with very poor phone reception it will take significant investment to achieve a live model or digital twin.

Once live models become a real possibility this could then be combined with turbidity modelling. The possibilities then are huge, as travel times change every second as flows are not constant in a domestic distribution network, with live data the model can make updates on the fly giving network operators accurate information to make the best informed decision.

9. Submission Reflection

This exercise has taught me several things about myself both personally and professionally. This is the first time I have carried out an exercise like this and picking a topic which is relevant to my work which made it very interesting. The problem with the topic I picked and if I had to do it again I would consider something that was not so new. As the concept of modelling turbidity is a new idea, pioneered by a small group it meant the literature was limited however, the literature that was out there was of excellent quality.

Since starting this assignment I have used turbidity modelling as part of my role as a clean water hydraulic modeller which resulted in accurate data being passed to senior members of Welsh Water and reporting the findings to the regulator for clean drinking water.

My knowledge of the topic has significantly increased thanks to carrying out this assignment.

Having researched this topic I now would like to further my understanding of the topic and due to the limited literature around the topic I have started my own research in work with the aim of presenting it at the next Clean Water Modelling Advisory Group meeting.

(CWMAG)

Having carried out this research it has really shown me the true potential of modelling, and as the topic has very little in the way of research and literature there appears to be lots of opportunities to study and develop things further. With water resources being strained and stressed around the world as the effects of global warming continue this problem will only become worse. The importance of modelling and finding efficiencies is imperative.

10. Reference

Influence of Hydraulic Model Complexity on Results of Water Age and Quality Simulation in Municipal Water Supply Systems

Artur Zajkowski 1, , Wojciech Kruszyński 1,* , Izabela Bartkowska 1 , Łukasz Wysocki 2 and Anna Krysztopik 3*

1 Department of Water Supply and Sewerage, Bialystok University of Technology, 15-351 Bialystok, Poland

2 Water Energy Systems (WES), 15-256 Bialystok, Poland

3 Podlaskie Voivodeship Marshal's Office, 15-888 Bialystok, Poland

** Correspondence: artur.zajkowski@doktoranci.pb.edu.pl (A.Z.); w.kruszynski@pb.edu.pl (W.K.);*

Tel.: +48-571-443-143 (W.K.)

Modelling and field work to investigate the relationship between the age and the quality of drinking water at customer's taps

John Machell 1 (corresponding author) and Joby Boxall 2

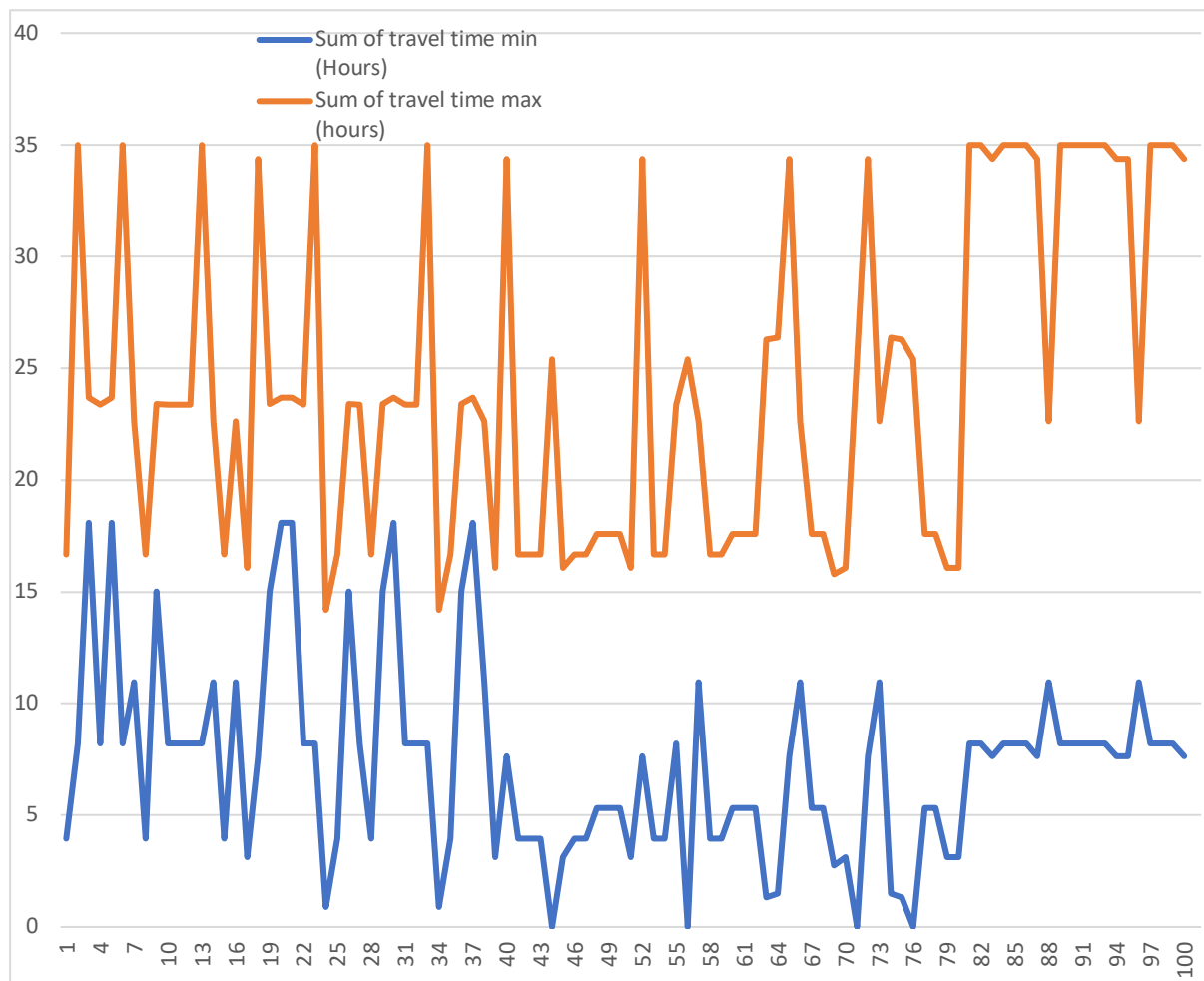
The application of artificial intelligence techniques to better manage iron in drinking water distribution systems

Joby Boxall¹, Katrina Flavell², Vanessa Speight¹, Grigorios Kyritsakas¹, Ehsan Kazemi¹, Stewart Husband¹, Samuel Bright², Sam Ledger², Luke Montgomery²

1 Department of Civil and Structural Engineering, University of Sheffield, Sheffield. UK

2 Yorkshire Water Services, Bradford. UK

11. Appendix



Appendix 1 – shows the age of water for each sample location from the data set

Row Labels	Sum of travel time min (Hours)	Sum of travel time max (hours)
1	3.93979	16.67637
2	8.2	35
3	18.08981	23.69214
4	8.2	23.36177
5	18.08981	23.69214
6	8.2	35
7	10.955	22.61619
8	3.93979	16.67637
9	15.02722	23.39623
10	8.2	23.36177
11	8.2	23.36177
12	8.2	23.36177
13	8.2	35

14	10.955	22.61619
15	3.93979	16.67637
16	10.955	22.61619
17	3.11409	16.08569
18	7.64317	34.38462
19	15.02722	23.39623
20	18.08981	23.69214
21	18.08981	23.69214
22	8.2	23.36177
23	8.2	35
24	0.88123	14.19473
25	3.93979	16.67637
26	15.02722	23.39623
27	8.2	23.36177
28	3.93979	16.67637
29	15.02722	23.39623
30	18.08981	23.69214
31	8.2	23.36177
32	8.2	23.36177
33	8.2	35
34	0.88123	14.19473
35	3.93979	16.67637
36	15.02722	23.39623
37	18.08981	23.69214
38	10.955	22.61619
39	3.11409	16.08569
40	7.64317	34.38462
41	3.94716	16.68058
42	3.94716	16.68058
43	3.94716	16.68058
44	0.00452	25.38917
45	3.11409	16.08569
46	3.94716	16.68058
47	3.94716	16.68058
48	5.33883	17.59425
49	5.33883	17.59425
50	5.33883	17.59425
51	3.11409	16.08569
52	7.64317	34.38462
53	3.93979	16.67637
54	3.93979	16.67637
55	8.2	23.36177
56	0.00452	25.38917
57	10.955	22.61619
58	3.94716	16.68058

59	3.94716	16.68058
60	5.33883	17.59425
61	5.33883	17.59425
62	5.33883	17.59425
63	1.33177	26.29342
64	1.49905	26.37443
65	7.64317	34.38462
66	10.955	22.61619
67	5.33883	17.59425
68	5.33883	17.59425
69	2.75466	15.79347
70	3.11409	16.08569
71	0.00452	25.38917
72	7.64317	34.38462
73	10.955	22.61619
74	1.49905	26.37443
75	1.33177	26.29342
76	0.00452	25.38917
77	5.33883	17.59425
78	5.33883	17.59425
79	3.11409	16.08569
80	3.11409	16.08569
81	8.2	35
82	8.2	35
83	7.64317	34.38462
84	8.2	35
85	8.2	35
86	8.2	35
87	7.64317	34.38462
88	10.955	22.61619
89	8.2	35
90	8.2	35
91	8.2	35
92	8.2	35
93	8.2	35
94	7.64317	34.38462
95	7.64317	34.38462
96	10.955	22.61619
97	8.2	35
98	8.2	35
99	8.2	35
100	7.64317	34.38462

Appendix 2 – the table above is the sample data showing wate age max and min from each sample point.

