Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Engineering
Department of Civil and Environmental Engineering

Britta Rachel Solveig Benz

Online Monitoring of Biofilm Growth in Premise Plumbing Systems

Master's thesis in Civil and Environmental Engineering
Supervisor: Cynthia Hallé
Co-supervisor: David Steffelbauer
June 2022

NTNU
Norwegian University of
Science and Technology

Britta Rachel Solveig Benz

Online Monitoring of Biofilm Growth in Premise Plumbing Systems

Master's thesis in Civil and Environmental Engineering
Supervisor: Cynthia Hallé
Co-supervisor: David Steffelbauer
June 2022

Norwegian University of Science and Technology
Faculty of Engineering
Department of Civil and Environmental Engineering

NTNU
Norwegian University of
Science and Technology

Abstract

The gram-negative bacteria *Legionella* can cause illness in humans, collectively termed legionellosis, which is potentially fatal. The global incidence of legionellosis is rising. *Legionella spp.* are native to both anthropogenic and natural water systems. Biofilm is considered a requisite for the survival and proliferation of *Legionella* in premise plumbing. Non-invasive, online biofilm monitoring thus carries a large potential in providing valuable, continuous information about biofilm development that more established methods, such as manual collection of biofilm biomass, cannot. However, currently, there exists no monitoring system that provides a complete set of information on biofilm growth. To reduce the shortage of documentation on the performance of these online monitoring systems, biofilm electrochemical signals recorded by four ALVIM Srl A001S3 sensors were analysed. The sensors were installed in four of in total twelve distal pipes, which each were connected to outlets of a premise plumbing pilot. The sensors proved highly sensitive to elevated temperatures around 49 and 60 °C, as well as rapid changes in flow rate. The ALVIM sensor data was used to give evidence on the efficacy of the following control measures applied to the pilot: increased hot water temperatures, flushing, and the disinfectant monochloramine. No growth was discerned during the phase in which a combination of flushing at 60 °C and addition of 1 $\frac{mg}{L}$ monochloramine was employed. Whilst the sensors provided an insight into biofilm growth, the sensor data could not have been interpreted as confirmation of biofilm growth with high enough certainty without supporting evidence provided by biofilm biomass sampling.

Sammendrag

Den gramnegative bakterieslekten *Legionella* kan forårsake sykdom med samlebenevningen legionellose. Legionellose kan gi alvorlig sykdom og er potensielt dødelig. Globalt er forekomsten økende. *Legionella spp.* forekommer i menneskeskapte så vel som naturlige vannsystemer. Biofilm er en tilnærmet nødvendighet for overlevelsen av *Legionella* i lokale drikkevannsystemer i bygg, som blant annet omfatter varme-, ventilasjons- og sanitærsystemer. Nettbaserte, ikke-inngripende biofilmovervåkningssystemer har et stort potensial til å bidra med å tette informasjonshull på biofilmvekst som ikke dekkes av mer anerkjente metoder som manuell innsamling av biomasse fra biofilm for DNA-analyse. I skrivende stund finnes det imidlertid ingen nettbaserte overvåkningssystemer som gir komplett og helhetlig informasjon om biofilmvekst. For å bidra til å redusere dokumentasjonsknappheten på slike overvåkningssystem, har denne studien samlet og analysert data fra fire ALVIM Srl A001S3 sensorer. Sensorene måler elektrokjemiske signal fra biofilm. Hver av sensorene ble installert i individuelle rør koblet til utløp i en modell av et lokalt drikkevannsanlegg. Sensorene viste seg å være høyst sensitive til forhøyede temperaturer (49 og 60 °C), samt hurtige endringer i strømningshastighet. Data fra ALVIM-sensorene ble brukt til å gi indikasjoner på virkningen av følgende tiltak for å forebygge eller redusere forekomsten av *Legionella* og biofilm: økte vanntemperaturer, spyling, og tilsetning av desinfeksjonsmiddelet monokloramin. Sannsynlig biofilmvekst ble observert i alle operasjonsfasene av piloten, med unntak av der en kombinasjon av spyling med en varmtvannstemperatur på 60 °C og tilsetning av 1 $\frac{mg}{L}$ monokloramin ble benyttet. Til tross for at sensorene ga en innsikt i biofilmvekst, så kunne ikke sensordataene alene brukes til å bekrefte veksten og tilstedeværelsen av biofilm. Til dette behøvdes komplementære overvåkningsmetoder som DNA-analyse av biomasse fra biofilm.

Preface

This project is a master's thesis in the course *Water Supply and Wastewater Systems - TVM4905*, at the Norwegian University of Science and Technology (NTNU), spring 2022. The work load accounts for 30 study credits.

First of all I would like to thank my main supervisor, Associate Professor Cynthia Hallé, for providing me with continuous advice, and bearing with me throughout the whole thesis process. Your knowledge on and experience with the subject has been invaluable.

Secondly, thanks is due to my co-supervisor, Associate Professor David Steffelbauer, who provided me with much appreciated guidance in data processing, as well as taking the time out of his very busy schedule to answer any questions I had.

Next, I would like to thank PhD candidate Charuka Meegoda for always taking the time to answer my countless questions, and provide me with all the laboratory information I needed and more. Your patience is highly valued.

Furthermore, thanks to Dr. Giovanni Pavanello, Director of ALVIM Srl, who really went out of his way to provide answers to any technical questions I had about the sensors used. Your knowledge is much appreciated.

Additionally, thanks to Professor Tone Merete Muthanna and Associate Professor Franz Tscheikner-Gratl for providing feedback and treasured advice on the writing process.

Thanks also to Researcher Michael B. Waak at Sintef, for sharing his schematics with me and allowing me to utilise them in this paper.

Thanks to the University of Minnesota, Twin Cities (UMN) for being a collaboration partner in this study.

Last but absolutely not least, thank you so much to family and friends who have supported me unequivocally this last half year, as well as throughout my studies, everyone in room 234 who has made the writing process an enjoyable one and lifted my spirits whenever I needed it, as well as everyone else who has helped me throughout my academic journey.

Britta Rachel Solveig Benz
June 8., 2022

Table of Contents

List of Figures

List of Tables

Acronyms

***Lp*1** *Legionella pneumophila* serogroup 1. 1, 2, 5, 22, 40

AIC Akaike information criterion. 11, 17, 19, 24

AOC assimilable organic carbon. 4, 13, 16

BES biofilm electrochemical signals. 1, 3, 14, 15, 17, 19, 22–26, 39

BIC Bayesian information criterion. 11, 17, 19, 24

EABs electrochemically active biofilms. 1, 4, 22, 23

FLA free-living amoeba. 1, 2, 5, 13, 16, 25

NIPH Norwegian Institute of Public Health. 4, 5

OPs opportunistic pathogens. 1, 5, 13, 19, 23

RMSE Root Mean Square Error. 11, 19, 21, 24

TCC total cell concentration. 3, 4

VBNC viable but non-culturable. 2, 5

1 Introduction

Legionella is a gram-negative genus of bacteria that exists ubiquitously in natural and man-made water systems (Falkinham et al., 2015). Man-made water systems suitable for *Legionella* occurrence and growth, such as premise plumbing systems, are suspected to be the main source of legionellosis (Falkinham et al., 2015; Kirschner, 2016). Legionellosis is the collective term for disease caused by *Legionella spp.*. Most commonly, it appears as Legionnaires' disease, a severe form of pneumonia, or milder Pontiac fever (Falkinham et al., 2015). The mortality rate for legionellosis remains high, in particular for vulnerable groups such as the immunocompromised and elderly (Herwaldt et al., 2018). Within the EU, *Legionella* is considered the waterborne pathogen with the highest health burden (European Parliament, Council of the European Union, 2020). Furthermore, *Legionella* has been reported to be a good indicator for other opportunistic pathogens (OPs) in drinking water (C. Zhang et al., 2021). *Legionella pneumophila* serogroup 1 (*Lp*1) is documented as the main cause of legionellosis, and thus frequently used to identify the presence of *Legionella spp.* (Lau et al., 2009; Miyashita et al., 2020). However, other legionellosis-causing species are considered under-represented in testing and detection (Chambers et al., 2021). Globally, the incidence of legionellosis is rising, which cannot solely be accredited to increased testing and detection (Falkinham et al., 2015; Graham et al., 2020; Herwaldt et al., 2018). An alteration in water consumption patterns and use, such as more frequent showers and the surge in use of water-based indoor cooling systems, as well as climate change, are important contributions to increased *Legionella* exposure (Voulvoulis, 2018; Walker, 2018).

Planktonic growth of *Legionella* is feasible, but considered rare (Vu et al., 2009). Amoeba, or free-living amoeba (FLA), and biofilm, often in conjunction, are deemed central for the proliferation of *Legionella spp.* in premise plumbing (Taylor et al., 2009). Both aid in providing shelter and nutrients for *Legionella* (Thomas et al., 2011; Wingender et al., 2011). Monitoring biofilm is thus an important factor in the control of *Legionella* (Lau et al., 2009; Pereira et al., 2021). Intracellular growth of *Legionella spp.* in amoeba has been reported to result in increased virulence and disinfectant resistance (Dupuy et al., 2011; Lau et al., 2009; Taylor et al., 2009). Additionally, amoeba are pathogenic on their own (Dupuy et al., 2011).

As a result of the strong association between *Legionella spp.* and biofilm occurrence in premise plumbing, improvements in biofilm monitoring methods are needed to advance the control of *Legionella* (Herwaldt et al., 2018; Lau et al., 2009; Pavanello et al., 2011; Thomas et al., 2011). Currently, there are no monitoring systems that provide information covering all aspects of biofilm development in premise plumbing systems. The same is true for the species *L. pneumophila*; invasive testing must be done to identify its presence. Required actions are almost solely based upon detection of *L. pneumophila* above a specified threshold level (European Centre for Disease Prevention and Control, 2018; European Parliament, Council of the European Union, 2020; Pettersen, 2015a). In contrast, biofilm monitoring is not a standard inclusion in monitoring protocols (Pereira et al., 2021). A multitude of biofilm monitoring systems make use of the electroactivity associated with biofilms. Bacteria, archaea and eukarya all contain electroactive microorganisms, aggregating in electrochemically active biofilms (EABs), with diverse compositions and functioning conditions. EABs are more frequently exploited in wastewater systems, for example for denitrification, or in microbial fuel cells, for harvesting electrical power, than for biosensors in drinking water (Erable et al., 2010; Logan et al., 2019; Patil et al., 2010; Patil et al., 2011; Rabaey et al., 2009). Online biofilm sensors can monitor biofilm development through the measurement of biofilm electrochemical signals (BES) emanated by EABs (Faimali et al., 2011; Pavanello et al., 2011).

There are currently no confirmed methods to eliminate *Legionella spp.* from water systems, making preventative actions, such as biofilm control, a requisite for *Legionella* control (Falkinham et al., 2015; Pereira et al., 2021). Despite the commercial availability of biofilm sensors, there still exists a significant knowledge gap in how to best utilise and analyse the BES-data generated, as well as the serviceability of biofilm sensors (Falkinham et al., 2015; LeChevallier, 2021; Pereira et al., 2021). The primary goal of this investigative study was to determine whether biofilm sensors can be used to monitor biofilm growth in premise plumbing. To achieve this, we used four ALVIM Srl A001S3 sensors in a pilot system, designed to evaluate the impact of temperature, pipe material, flow conditions and disinfection on biofilm and *Legionella* growth.

2 Background Theory

2.1 Legionella

Legionella spp. in premise plumbing display the following characteristics: (i) disinfectant resistance, (ii) pipe surface adherence, and (iii) the ability for growth within amoeba and biofilm (Falkinham et al., 2015). *Legionella* require specific nutrients such as nitrogen for growth. Colonisation of biofilm and (often simultaneous) parasitism of FLA is considered the main growth strategy (Taylor et al., 2009).

Legionella spp. interact with FLA in a species-specific manner (Thomas et al., 2011). *Acanthamoeba* is the most commonly identified host of *Legionella spp.*, but is also subject to identification skewness (Thomas et al., 2011). At least 20 amoeba species have been documented as *Legionella* hosts (Lau et al., 2009). Intracellular growth of *Legionella* is associated with a subsequent increase in virulence and infectivity, as well as resistance to biocides and higher temperatures (Lau et al., 2009; Taylor et al., 2009). FLA may enter a dormant cyst form, which acts as an additional buffer against unfavourable environmental conditions for *Legionella* cells harbouring within (Thomas et al., 2011). Similarly, *Legionella spp.* may form viable but non-culturable (VBNC) cells, reported to withstand higher temperatures and disinfectant doses than regular *Legionella* cells (Casini et al., 2018).

In Norway there are several central laws dictating the control of *Legionella*: the Norwegian Public Health Act (Lovdata, 2011), the Working Environment Act (Lovdata, 2006), the Communicable Diseases Act (Lovdata, 1995), the Bathing Facilities, Pool Baths and Saunas Health Act (Lovdata, 1996), the Planning and Building Act (Lovdata, 2008), the Drinking Water Act (Lovdata, 2016), and Regulations on technical requirements for construction works (TEK17) (Norwegian Building Authority, 2017). TEK17 includes a guiding document that details how to fulfill specific laws of TEK17. Regarding *Legionella* control, guidance on $\S\,15-5$ in TEK17 is particularly relevant. The Norwegian Public Health Act places the responsibility of ensuring the health of inhabitants upon municipalities, but the owners of water systems are responsible for ensuring that their systems do not infect users.

Human-to-human transmission of *Legionella* is considered rare, and zoonotic transmission has not been documented, establishing *Legionella* as an environmental pathogen (Graham et al., 2020). *L. pneumophila* has been documented to be the causative agent in 95 percent of cases of Legionnaires' disease (Chambers et al., 2021; Graham et al., 2020; Walker et al., 2021). Whilst *Lp*1 is the main legionellosis cause, it is subject to a positive bias; diagnosis is mainly conducted through the use of antigen detection tests explicitly made for *Lp*1 (Chambers et al., 2021; Graham et al., 2020; Miyashita et al., 2020; Vaccaro et al., 2016). Of the more than 60 *Legionella spp.* discovered so far, 19 have been reported to cause at least one infection (Walker et al., 2021). The total number of species as well as the number of pathogenic species of *Legionella* is increasing (De Giglio et al., 2019; Dennis et al., 1993; Fields, 1996; Gomes et al., 2018; Lesnik et al., 2015; Mercante et al., 2015; Muder et al., 2002; Pascale et al., 2020; Pearce et al., 2012; Wu et al., 2019).

The disease burden caused by non-*L. pneumophila* species such as *Legionella anisa*, is underrecognised and thus overlooked in testing, detection, monitoring and control of *Legionella* (Chambers et al., 2021; Logan-Jackson et al., 2021; Walker et al., 2021). Whilst drinking water is considered a major legionellosis source, others, such as soils, are gaining attention for their pathogenic role (Chambers et al., 2021). The health risk tied to drinking-water-related legionellosis is majorly impacted by the efficacy of disinfectants and protection of biofilm (Shen et al., 2017). *L. pneumophila*-associated biofilms have been credited for legionellosis outbreaks (Khweek et al., 2018). Similarly, release vesicles originating from FLA have the potential of transmitting *Legionella* (Lau et al., 2009).

2.2 Biofilm

Biofilm is considered the principal form of bacterial life (Flemming et al., 2019). Biofilm, defined as an aggregation of various bacteria, algae, fungi and protozoa, is estimated to house $40-80$

percent of all cells on earth (Flemming et al., 2019; Vu et al., 2009). One central component of biofilm, extracellular polysaccharides, (EPS), is responsible for adherence and support, as well as enabling multicellular functionality (Flemming, 2001; Mayer et al., 1999).

Biofilms have the following benefits compared to planktonic life forms: (i) increased mechanical strength, (ii) easy access to nutrients, (iii) horizontal gene transfer, and (iv) communication as well as extracellular digestion of nutrients, conducted by EPS (Wingender et al., 2011). Mayer et al., 1999 outlined the most important biofilm characteristics to be: (i) location and extent, (ii) quantity (mass and thickness), (iii) deposit nature (organic/inorganic, chemical composition etc.), and (iv) kinetics of formation and removal.

Biofilm composition varies with age, organism type, and the following environmental conditions: (i) oxygen and nitrogen levels, (ii) temperature, (iii) pH, (iv) nutrient availability, (v) flow conditions, (vi) desiccation event (Mayer et al., 1999; Vu et al., 2009). The outcome of studies done by Martiny et al., 2003 and Boe-Hansen et al., 2002 established that younger biofilms have a community composition more akin to bulk water than older biofilms. The results of G. Liu et al., 2014 support this, revealing that bacterial communities in the bulk water phase differ significantly in composition from pipe wall biofilms, suspended solids and loose deposits. Dias et al., 2019 concluded that the bacterial composition in premise plumbing was to a greater extent determined by pipe material, disinfectant residuals and hydraulic residence time, as opposed to the bacterial community composition of the treated water.

Increases in build-up rate of biofilm is an early warning sign of pathogenic risk (Pavanello et al., 2011). Likewise, the last stage of biofilm development, dispersal, signifies the ability to spread to other parts of water systems (Hall-Stoodley et al., 2005; Pereira et al., 2021). Dispersal of a large number of cells at once is believed to be among the mechanisms that give pathogens originating from biofilms a higher likelihood of causing disease than planktonic pathogens (Hall-Stoodley et al., 2005). To remove biofilms, their mechanical strength must be overcome (Mayer et al., 1999). In contrast to dispersal as a biofilm development stage, biofilm sloughing off caused by rapid increases in flow rate (flushing), aims to reduce biofilm growth, but can cause an increase in the amount of microorganisms released from the biofilms in the time period right after (Hozalski et al., 2020; Muhammad et al., 2020; Shen et al., 2015)

Biofilms can also contribute to biocorrosion, also referred to as microbially influenced corrosion (MIC), causing water quality alterations that result in unwanted taste, changes in colour, turbidity, odour, and reductions in insulating properties of heat exchangers (Mayer et al., 1999; Prest et al., 2016). Biofilm occurrence may mobilise or immobilise metals, in particular through the corrosion of copper, resulting in structural damages and elevated copper levels in drinking water (Beech et al., 2004; Vargas et al., 2017). MIC is associated with soft water with little disinfection (Vargas et al., 2017). There exists uncertainty in regards to what influences the biofilm in drinking water pipes more: the microbiome in the local drinking water, or the source water, but Chan et al., 2019 concluded that biofilm in pipes is the main reason for increases in total cell concentration (TCC) in water at greater distances from treatment plants in water distribution systems.

2.3 Theoretical Principle of ALVIM Srl A001S3 Sensors

BES measured by ALVIM sensors are a result of cathodic depolarisation, which describes the phenomenon of cathodic current increases as biofilm grows (Faimali et al., 2008; Faimali et al., 2010). ALVIM A001S3 sensors are composed of a three-electrode system: a working electrode of stainless steel, in contact with the water, and a zinc electrode that acts both as a counter and a reference electrode (Pavanello et al., 2011). The BES-data generated by the sensors is a measure of the difference in potential between the working and counter electrode, given the use of the potentiostatic mode of the ALVIM sensors (ALVIM, 2020; Pavanello et al., 2011). Sensitivity is directly proportional to the current that flows between the electrodes of the sensor, and is provided as a pre-calibrated, system-specific value (Faimali et al., 2011).

Biofilms have been shown to be capable of accelerating electron transfer cathodically (from metal to external acceptor) and anodically (from external donor to metal) (Faimali et al., 2010). However,

there exist no mechanism which describes the process of cathodic depolarisation in its entirety. According to some literature, it is caused by the reduction of new cathodic reactants produced by certain bacteria in biofilm (Faimali et al., 2011). Other literature states that the increase is due to oxygen reduction catalysis by organometallic substances (Brink et al., 2010; Johnsen et al., 1985), or enzymes generated by bacteria and immobilised on the metal surface (Huang et al., 2001; Lai et al., 2000; Lai et al., 1999; Scotto et al., 1985). Results from a study conducted by Faimali et al., 2011, are in agreement with the hypothesis that the current increase is a result of electron transfer from metal to oxygen, catalysed by enzymes generated by settled bacteria, and immobilised on the stainless steel surface. Regardless of magnitude or biofilm development stage, Faimali et al., 2011 suggested that the limiting factor of current increase is the enzyme turnover rate.

Cathodic current increases have been shown to be proportional to the percentage of the metal surface covered by biofilm (Faimali et al., 2008). The increase in cathodic current is independent of the activity of specific bacteria, highlighting the large phylogenetic diversity of EABs (Faimali et al., 2010; Vandecandelaere et al., 2009). By extension, individual biofilms cannot be distinguished based on their electrochemical activity, as they produce similar signals regardless of composition (Faimali et al., 2010).

2.4 Premise Plumbing

Premise plumbing, also referred to as domestic drinking water distribution systems or building water systems, are the sections of water distribution systems that are connected to the main system through service lines, encompassing drains, devices such as hot water tanks, and fixtures such as showers (United States Environmental Protection Agency (EPA), n.d.). They are characterised by a high surface-area-to-volume-ratio, low concentrations of assimilable organic carbon (AOC), diverse plumbing material, low or no disinfection residuals and frequent periods and areas of stagnation (Falkinham et al., 2015; Leslie et al., 2021). In Norway, concentrations of AOC vary, even though they are generally reported as low (Charnock et al., 2000; Krzeminski et al., 2019).

The Norwegian Institute of Public Health (NIPH) states that the following systems carry the largest risk of legionellosis, largely attributable to their capacity for aerosol-production: (i) cooling towers, (ii) air scrubbers, (iii) public bubble baths, (iv) heating, ventilation and air conditioning (HVAC) systems in buildings with residents at high risk, (v) indoor fountains, and (vi) certain humidification systems (Pettersen, 2015a). Shower outlets provide nearly-ideal environments for biofilm growth; Proctor et al., 2018 concluded that *Legionella* occurrence had a positive correlation with biofilm TCC, which in turn was found to be influenced by the frequency of use and disinfectant choice, but not the shower hose age. Thus, shower outlets present a particularly high risk of harbouring and spreading *Legionella* (Pettersen, 2015a).

Bédard et al., 2018 documented that points of use with higher surface-to-volume ratios were associated with higher bacterial loads, presumably due to a larger attachment area for biofilm. A study performed in Tuscon, Arizona determined that the bacteria ingested by consumers mainly stemmed from the residency distribution system (Pepper et al., 2004). Materials with high biofilm formation potential, elevated temperatures (above 20 °C) and large volumes of stagnant water also contribute to favourable biofilm growth conditions (Bédard et al., 2018; Pettersen, 2015a). In water distribution systems, storage units are suspected to provide favourable growth conditions for *Legionella* (Walker et al., 2021). The most distal branches of water distribution systems, i.e. those entering premise plumbing, reportedly contain more biofilm than other parts (Falkinham et al., 2015).

Maintaining sufficient water temperatures and avoiding stagnant areas is particularly challenging in large buildings with complex water systems, or poorly-designed systems, due to a higher occurrence of dead-ends and stagnant areas, as well as shifting water temperatures (Pereira et al., 2021; Pettersen, 2015a; Walker et al., 2021). Stagnation in premise plumbing is associated with significant decreases in water quality, and is a critical factor for enhanced growth potential of *Legionella spp.* (Bédard et al., 2018; Pereira et al., 2021). In particular, periodically stagnant water, combined with temperatures between $25 - 50$ °C is ideal for growth of *Legionella spp.* (Bédard et al., 2018; Gavaldà et al., 2019; Pettersen, 2015b). Lipphaus et al., 2014 found infrequently used taps to be

associated with the highest bacterial load, whilst stagnation tended to increase the bacterial load in the frequently used taps. Temperatures around $55 - 60$ °C are deemed sufficient for reducing *Legionella* numbers significantly (Bédard et al., 2015; Cervero-Aragó et al., 2019; Gavaldà et al., 2019; Rasheduzzaman et al., 2020).

VBNC cells of *Legionella* have been shown to tolerate temperatures up to 70 °C, but the ability to infect FLA decreases significantly above 50 °C (Cervero-Aragó et al., 2019). The ability of *Legionella* to proliferate in amoeba is also weakened at temperatures below 25 °C (Taylor et al., 2009). Buse et al., 2017 reported a higher relative abundance of *L. pneumophila* compared to other *Legionella spp.* at 37 °C, than at slightly lower temperatures, indicating that higher temperatures select for *L. pneumophila*. In situations with *Legionella*-outbreaks, NIPH recommends heat-shock-flushing at 80 °C for 30 minutes, whilst flushing at 70 °for at least 5 minutes is considered satisfactory as a preventative action (Pettersen, 2015b).

Disinfectant levels are impacted by system hydrodynamics, and interact with water flow, temperature and structural components (LeChevallier, 2021). Sydnor et al., 2012 credited this interaction as the cause for the higher and more frequent occurrence of OPs, such as *Legionella*, in automatic faucets compared to manual. Flushing, which causes increased shear stress along pipe walls, can ensure disinfection is distributed evenly throughout water systems, reduce the occurrence of stagnant areas, as well as enhance disinfection through the increased flux of disinfectant into biofilms (Bishop et al., 1997; Pereira et al., 2021).

Monochloramine is a common disinfectant in drinking water systems. Buse et al., 2019 report that free chlorine has a higher efficacy in inactivating planktonic *Lp*1 cells than monochloramine, whilst monochloramine has a higher inactivation efficacy on biofilm-associated *Lp*1 in copper pipes. Additionally, *Lp*1 cells shed from biofilm are associated with higher disinfectant resistance than planktonic *Lp*1 cells (Buse et al., 2019). However, Lee et al., 2018 concluded that whilst monochloramine provided significantly more rapid penetration of biofilm, this penetration did not correlate with an immediate loss in viability and aerobic activity, whereas it did in the case of biofilm penetration by free chlorine.

Monochloramine persists for a longer duration in water distribution systems than free chlorine, given that it does not experience accelerated decay due to nitrification (Y. Zhang et al., 2009). Systems using monochloramine contain free ammonia, which acts as a substrate for ammonia-oxidising bacteria, causing increased nitrification (Lee et al., 2018). Nitrification causes an accelerated decomposition of monochloramine and loss of disinfection residuals, reduction of dissolved oxygen, and thus, increased potential for biofilm growth (Hossain et al., 2022).

Changing water consumption patterns, in particular water reuse and increased water ages in e.g. so-called green buildings, lead to a decrease or an absence of chlorine residuals; Rhoads et al., 2016 concluded that these residuals decay up to 144 times faster in premise plumbing water of a higher water age, compared to water in water distribution systems. Premise plumbing systems are thus often associated with the absence of disinfection residuals. Non-targeted disinfection practices, combined with *Legionella*s' lower susceptibility to chlorination compared to other OPs in premise plumbing such as *Escheria Coli* (Y. E. Lin et al., 2011; Y. Lin et al., 1998) may result in dominance of *Legionella spp.* in premise plumbing. Additionally, mixed species biofilms may gain increased resistance to disinfectants by altering their composition (Shen et al., 2017).

3 Materials and Method

3.1 Case Description: Premise Plumbing Pilot

The premise plumbing pilot (Figure 1), consists of 2 rigs, each with 6 distal pipes, referred to as lines, connected to the outlets; 3 hot, marked H, and 3 cold supply lines, marked C. These are divided into group A, B or C. Each of the lines has a diameter of 1.27 cm, whilst the rest of the pilot has a diameter of 2.5 cm. Rig 1 is entirely made of copper, whilst Rig 2 has lines of PEX-A. Copper has biocidal properites (Lehtola et al., 2004). Therefore, the concentration of copper was monitored as part of an independent study administered on the same pilot. These values are presented in Krakk, 2021.

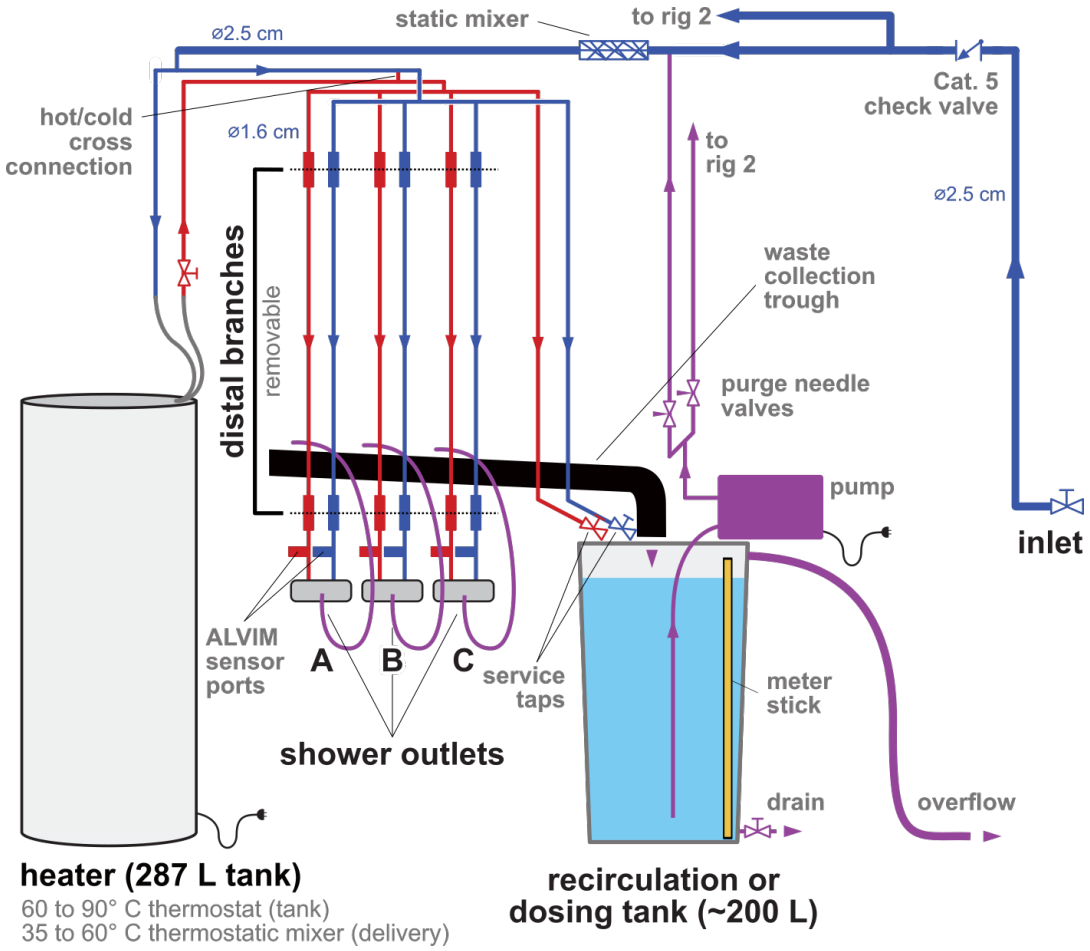


Figure 1: Schematic of the premise plumbing pilot. Credit: Michael B. Waak, used with permission.

The operational design of the premise plumbing pilot is presented in Figure 2, Figure 3 and Figure 4. Prior to the data collection encompassing the 4 phases, the premise plumbing system was subject to an acclimation period of 8 weeks, in which re-circulation was employed to stimulate the formation of biofilm. Between phases, a transition period of minimum 2 weeks was applied, see Figure 2.

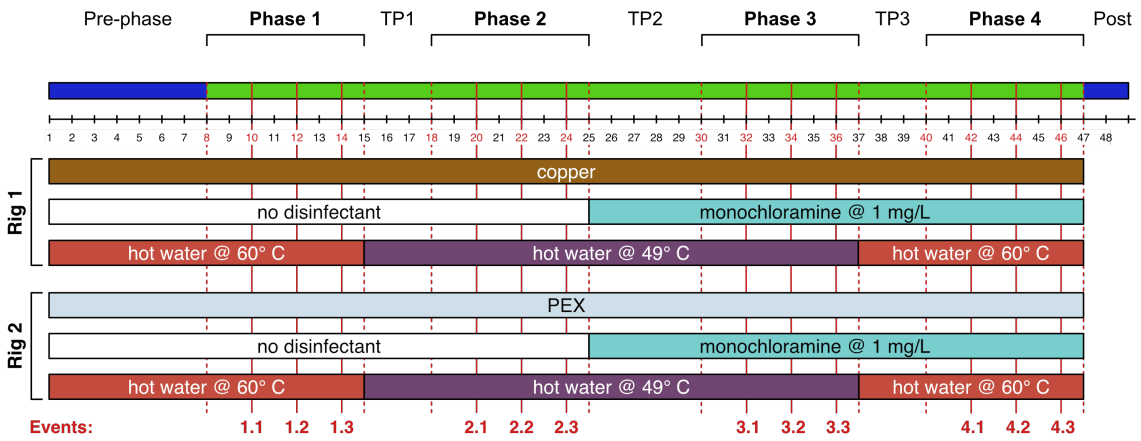


Figure 2: Operational design of the pilot. Credit: Michael B. Waak, used with permission.

During Phase 1 and Phase 2, no chlorine residuals were present in the system. During Phase 3 and Phase 4, 1 $\frac{mg}{L}$ monochloramine was added to the system, see Section 3.2. In Phase 1 and Phase 4, the hot water temperature was set to 60 °C, whilst it was at 49 °C during Phase 2 and 3. Each phase was subject to a different combination of events. Each event is characterised by the combination of daily treatments applied. Daily treatments are the procedures: flush, shower, or stagnation, conducted once on a given day, see Figure 3 and Figure 4.

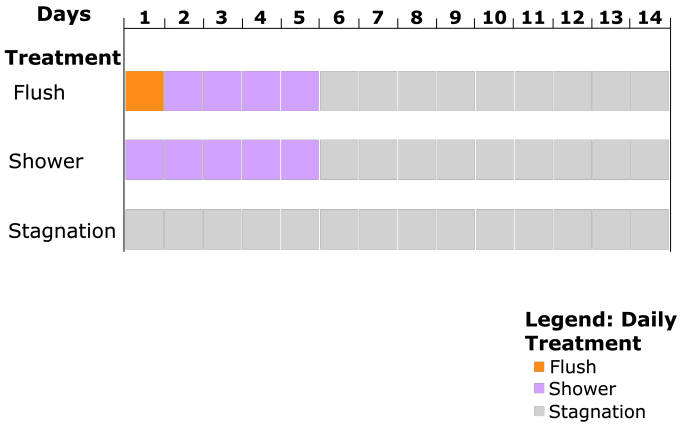
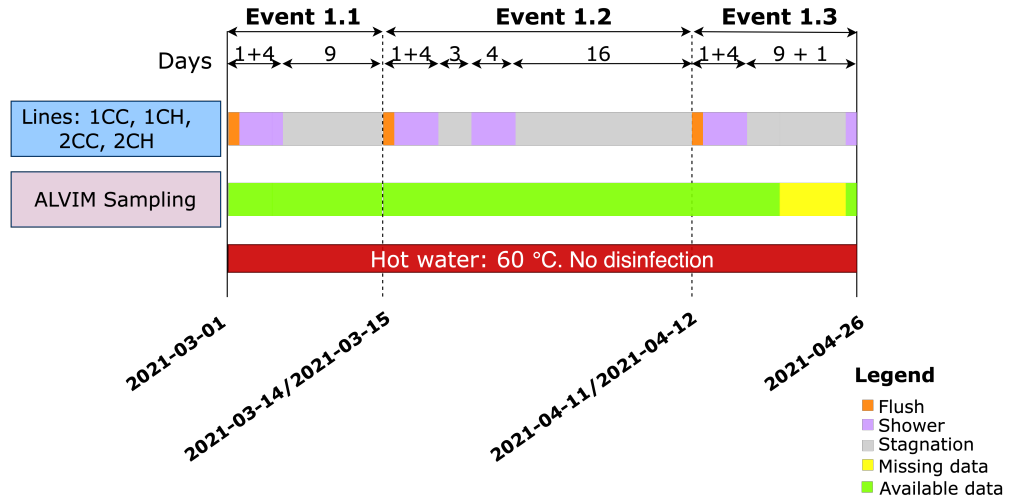


Figure 3: Operational design of events.

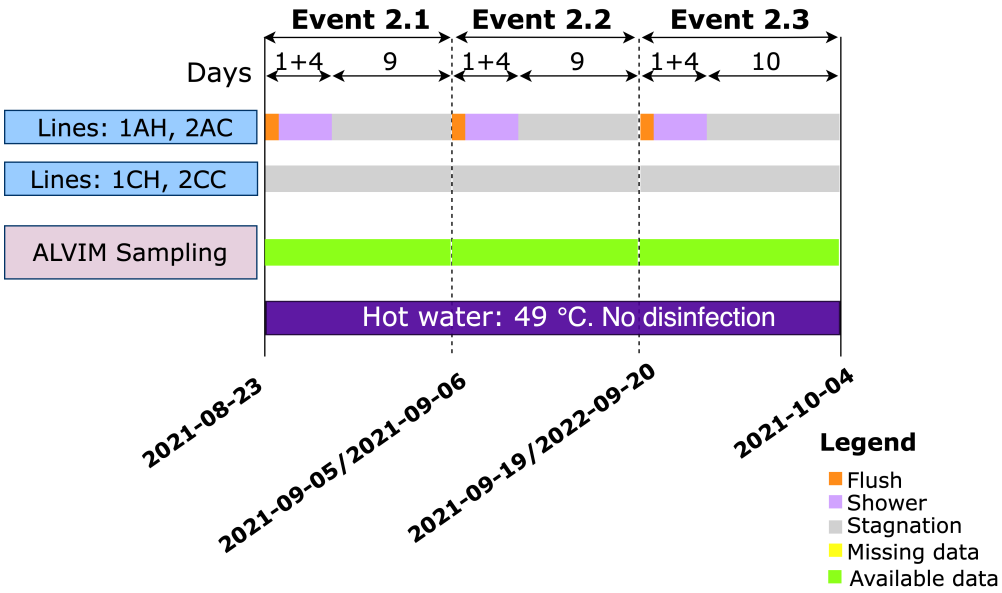
ALVIM Srl A001S3 sensors are online, non-destructive tools for biofilm monitoring (Pavanello et al., 2011; Pereira et al., 2021). In total four ALVIM sensors were installed, two in Rig 1, and two in Rig 2, in four individual lines. All four sensors were continuously in use, and their placement was kept consistent, with the exception of one configuration change, which occurred between Phase 1 and Phase 2, see Figure 4. The placement and events were used to name all the sampled data sets: for example 1AH 1.1 refers to Event 1.1 conducted in line 1A Hot (flushing).

Flushing was in this study conducted as unobstructed flow (without an aerator) for five minutes, at temperatures of 49 or 60 °C in hot lines, and at inlet temperatures for cold lines. The daily treatment flush was done in one line at a time, unlike the daily treatment shower. Thus, unlike the daily treatment shower, only one line at a time had elevated flow rates at once during flushing. During showers, which had a duration of 7.8 $\pm$ 0.5 min and a flow rate of 7.2 $\frac{L}{min}$, both the hot and cold line within the same group were in use, to achieve a blend of 40 °C at the outlet. Flow rate was calculated using a meter stick, see Figure 1, calibrated in 10 L increments.

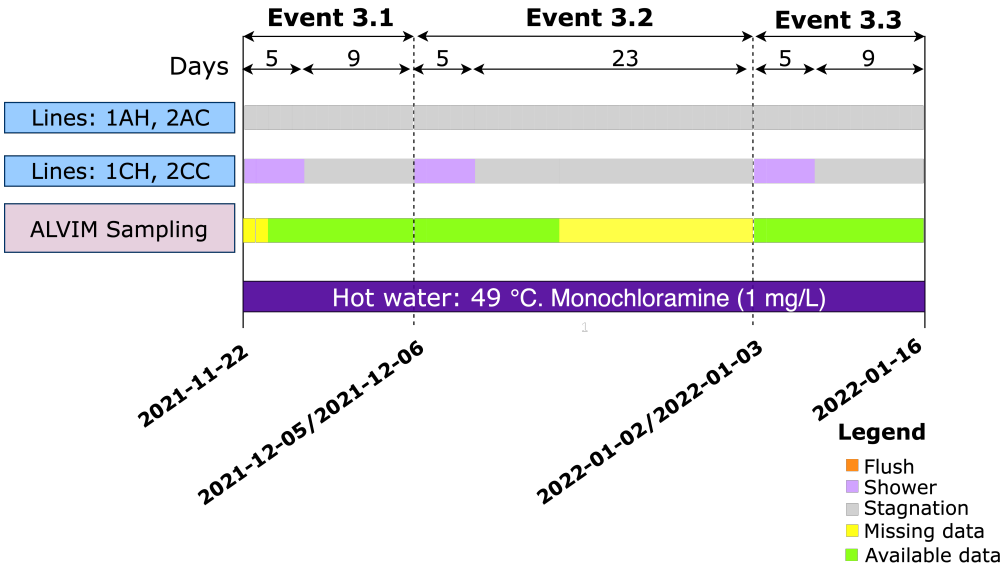
Various malfunctions of the ALVIM sensors occurred during the study, resulting in a loss of data, marked in Figure 4. Data recorded in transition phases was not used in this study.



(a) Phase 1

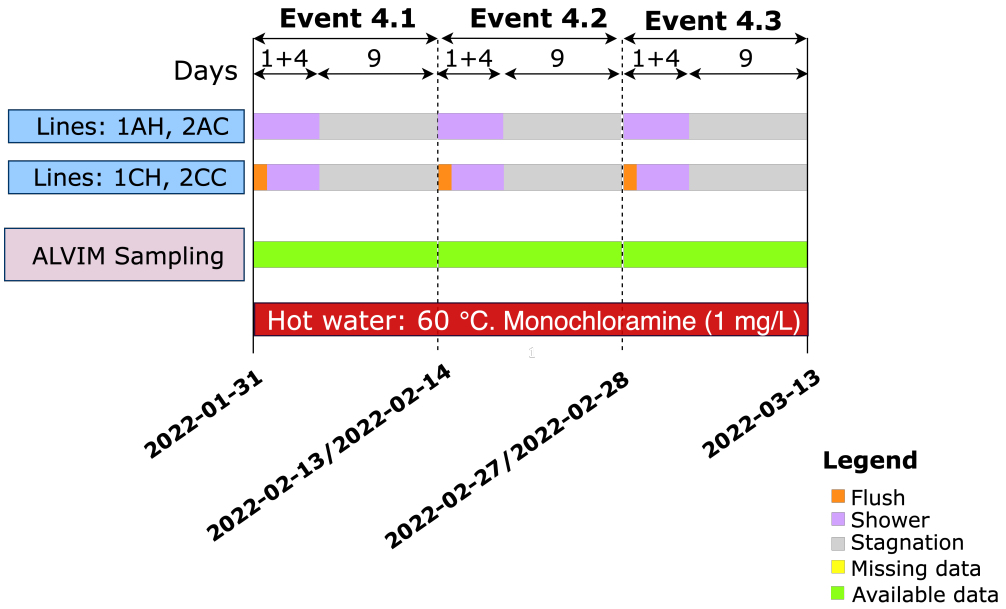


(b) Phase 2



(c) Phase 3

Figure 4: Operational timeline of phases: (a) Phase 1, (b) Phase 2, (c) Phase 3, and (d) Phase 4.



(d) Phase 4

Figure 4: Operational timeline of phases: (a) Phase 1, (b) Phase 2, (c) Phase 3, and (d) Phase 4.

In an independent study on the premise plumbing pilot, biofilm biomass was sampled in the distal branches of the rigs, see Appendix D.

3.2 Disinfection

The premise plumbing pilot was operated with effluent water from Vikelvdalen drinking water treatment plant in Trondheim, which on average uses $0.5 - 0.8\ \frac{mg}{L}$ of the disinfectant sodium hypochlorite, $NaOCL$. $NaOCL$ can appear as hypochlorous acid, $HOCL$ or hypochlorite ion, OCL^-. At $pH = 8$, $HOCL$ only constitutes approximately 20 percent of free chlorine (Trondheim kommune, 2017). At pH values below 7.5, there will be higher relative concentrations of $HOCL$, resulting in higher disinfection efficacies (Choi et al., 2022).

Monochloramine, NH_2Cl, was added to the pilot and hot water tanks in Phase 3 and Phase 4, as well as during the transition period between Phase 2 and Phase 3, initially at a concentration of $0.5\ \frac{mg}{L}$, which was increased to $1.0\ \frac{mg}{L}$. Anhydrous ammonia, NH_3, reacts with $HOCL$ or OCL^- in water to form chloramines: depending on pH, temperature, chlorine dose and other factors, this reaction will form monochloramine, dichloramine and/or trichloramine (Harp, 1995).

To evaluate the decay of monochloramine, bulk water samples i.e. water samples taken directly after a shower, and placed in a clean bottle kept in the dark, as well as water samples taken from the lines after a period of stagnation, were administered. The bulk water samples represent the decay solely due to water, whilst the water samples represent the decay due to both water and biofilm. Chlorine decay in these samples was calculated through the use of first order kinetics, see Equation 1.

$$\frac{C}{C_0} = e^{-K_b * t} \tag{1}$$

In Equation 1, K_b is the decay constant, t is time in hours, and C and C_0 are the final and initial concentrations respectively.

3.3 Biofilm Monitoring: ALVIM A001S3 Srl Sensors

The ALVIM Srl A001S3 sensors were used in the potentiostatic mode for this study. This mode uses Equation 2 to model the cathodic current, $i(E,t)$, at a fixed potential, E, on the stainless steel (working) electrode (Pavanello et al., 2011). $\Theta(t)$ is the fraction of surface covered by biofilm, $i_1(E)$ is the current density of the fraction of the electrode surface not covered by biofilm, whilst $i_2(E)$ is the current density of the fraction $\Theta(t)$, also referred to as active sites for oxygen reduction (Faimali et al., 2008). ALVIM sensors using the potensiostatic technique have been shown to provide detailed information on biofilm development down to $\Theta(t) = 0.01\%$ (Faimali et al., 2008; Faimali et al., 2010; Mollica et al., 1997). The sensors sampled at a frequency of 3 seconds during daily treatments, and 10 minutes at other times.

$$i(E,t) = i_1(E) + [i_2(E) - i_1(E)] * \Theta(t) \tag{2}$$

Equation 2 is based on a model that assumes biofilm develops according to the following steps: (a) planktonic bacteria in seawater settle on the stainless steel surface with a given probability, (b) settled bacteria can, later on, generate another bacteria which in turn settles, yielding the apparent duplication time (Faimali et al., 2008).

One drawback of the potentiostatic technique is the possibility of carbonate precipitation associated with mature biofilm, given that a high cathodic current (in the order of 50 $\frac{\mu A}{cm^2}$) is sustained for a long period of time (Pavanello et al., 2011). This can cause a decrease of the active sensor area, but is preventable through periodic acid cleaning, which was conducted in the pilot between Phase 1 and Phase 2 (Pavanello et al., 2011). Additionally, the drinking water quality parameters indicate that the water is too soft for this to be an issue, see Table 2 in Appendix A.

3.4 Data Processing and Analysis

First, the raw data was organised after the principles of tidy data (Wickham, 2014). The raw data was taken from dated excel-files, sorted into column after line sampled, merged with files containing metadata such as daily treatment, event type and disinfection. The Python (Van Rossum et al., 2009) scripts, for both data processing and analysis can be found in Appendix G. Note that in some cases, not the entire code is written out, instead, only an example in which the input name can be altered, is provided.

Probable biofilm growth curves were identified manually for each event in each line. These are characterised by a gradual increase in ALVIM sensor data of 150 mV, or more (ALVIM, 2020). However, gradual increases of 100 mV and less were also extracted as the ALVIM technology developers specified this to be an indication of early biofilm development.

In addition to the biofilm growth curves, the stagnation period after each of the different event types was extracted and curve fitted. For showers and stagnation events, all the stagnation period data sets were analysed, but for flushing, only data sets containing 8 or 9 days were used, as those shorter than this (2 days) omitted important trends. The flushing event data sets with a duration of 22 days (all from event 1.2) were also omitted due to missing data and significant differences in behaviour, but were analysed in the extraction of probable biofilm growth. Neither the shower, nor the stagnation data sets were analysed at full lengths, even though all the data sets were used. For the shower data sets analysis was done until and including day 6, meaning two days were omitted from a minority of the data sets. The same applied to stagnation events, in which 11 days were deemed sufficient, despite the occurrence of some data sets containing 13 or 14 days of data. This selection was used throughout the data analysis for consistency purposes. The stagnation event data sets consisted of more days than the other event data sets, as data from the whole event was usable. The first five days of flushing and shower events contained daily treatments, as detailed in Figure 3.

3.4.1 Root Mean Square Error

The objective function Root Mean Square Error (RMSE) was used to do an initial quantification and visualisation of the difference between data sets. Each data set consisted of a unique combination of line and event sampled. RMSE is typically used to evaluate model performance, and is defined by Equation 3 (Bennett et al., 2013; W. Wang et al., 2018):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{N}} \tag{3}$$

$\hat{y}_i$ are the predicted/modeled values, y_i are the actual/measured values, and N is the number of values. RMSE has the benefit over Mean Square Error of yielding results in input units (mV), making RMSE-values more easily interpretable (Bennett et al., 2013). The range of RMSE is $(0, \infty)$, and the lower the RMSE-value is, the better the function fits the data (Bennett et al., 2013; W. Wang et al., 2018). RMSE squares the residuals before calculating the mean, making all contributions positive, thus hindering cancellation of positive and negative errors, as well as penalising large errors more (Bennett et al., 2013).

3.4.2 Curve Fitting

Fitting curves was conducted using the free software python package lmfit 1.0.3 (Newville et al., 2021), which makes use of the non-linear least-squares method for minimisation. The method of least squares aims to minimise the error; for a dataset consisting of $(x_1, y_1), ...(x_n, y_n)$, represented by the linear equation $y = ax + b$, the error is given by Equation 4 (Miller, 2006).

$$E(a, b) = \sum_{i=1}^{N}(y_n - (ax_n + b))^2 \tag{4}$$

The squared term in Equation 4 results in heavier weighting of large errors, thus penalising them more (Miller, 2006).

The best fit was selected on the basis of the results of the following goodness of fit measures: Akaike information criterion (AIC), Bayesian information criterion (BIC), reduced chi-square, χ_r^2, as well as the relative error of the coefficients. χ_r^2, i.e. chi-squared divided by the number of degrees of freedom, is used to determine if a significant difference exists between two probability distribution functions; a small value indicates a high degree of correlation between the observed and modeled values (Pearson, 1900; Rossi et al., 2020).

AIC is defined in Equation 5 (Akaike, 1998; Bennett et al., 2013), and BIC in Equation 6 (Bennett et al., 2013; Schwarz, 1978). Both information criteria favour simple models, punishing complexity.

$$AIC = -2ln(L) + 2(k) \tag{5}$$

$$BIC = -2ln(L) + k * ln(n) \tag{6}$$

In Equations 5 and 6, L refers to the likelihood (maximum likelihood) of the model, given the input data, k to the number of parameters estimated, and n to the number of data points used in calibration.

The main difference between the equations lies in the penalty for model complexity: $2k$ for AIC, compared to $k * ln(n)$ for BIC. BIC thus penalises additional parameters in a model more than AIC. As BIC and AIC depend on maximum likelihood, the results depend on sample size, and therefore models using two different sample sizes cannot be compared (Akaike, 1998; Schwarz, 1978). AIC aims to find models that predict future data accurately, whilst BIC aims to identify

models with the highest probability of being the true model for the input data, however, they are commonly used in conjunction, in particular for finding models that perform favourable for both criteria (Kuha, 2004).

Bacterial growth is characterised by an initial exponential growth, which becomes logistic in nature as it slows due to limited nutrient availability (Cunningham et al., 2004; Tsai, 2005). Each data set was curve fitted using the following three functions: an exponential (Equation 7), a polynomial (Equation 8) and a linear (Equation 9).

$$y = a * e^{-b*t} + c \tag{7}$$

$$y = a * t^3 + b * t^2 + c * t + d \tag{8}$$

$$y = a * t + b \tag{9}$$

In Equation 7, Equation 8 and Equation 9, t refers to time (hours), and y to the ALVIM sensor output data (mV). A combination of Equation 7 and 9, as well as a power law and log-based equation were initially included, but discarded after initial tests showed a frequent lack of convergence and/or unsatisfactory results. The input values for parameters a, b, c, and d were altered in the event of non-convergence.

4 Results and Discussion

Biofilm microbiomes and their development are impacted by temperature, biofilm age, drinking water quality, flow conditions, material, transient intrusions of pathogens as well as the presence and abundance of FLA (Buse et al., 2017; Mayer et al., 1999; Pereira et al., 2021). Temperature, drinking water quality, material, as well as flow conditions varied in the pilot, with the ALVIM sensor data enabling the possible comparison of these conditions in regards to their impact on biofilm growth. A summary of the ALVIM sensor data collected in Phase 2 and processed, is presented in Figure 5. The other phase plots are available in Appendix B.

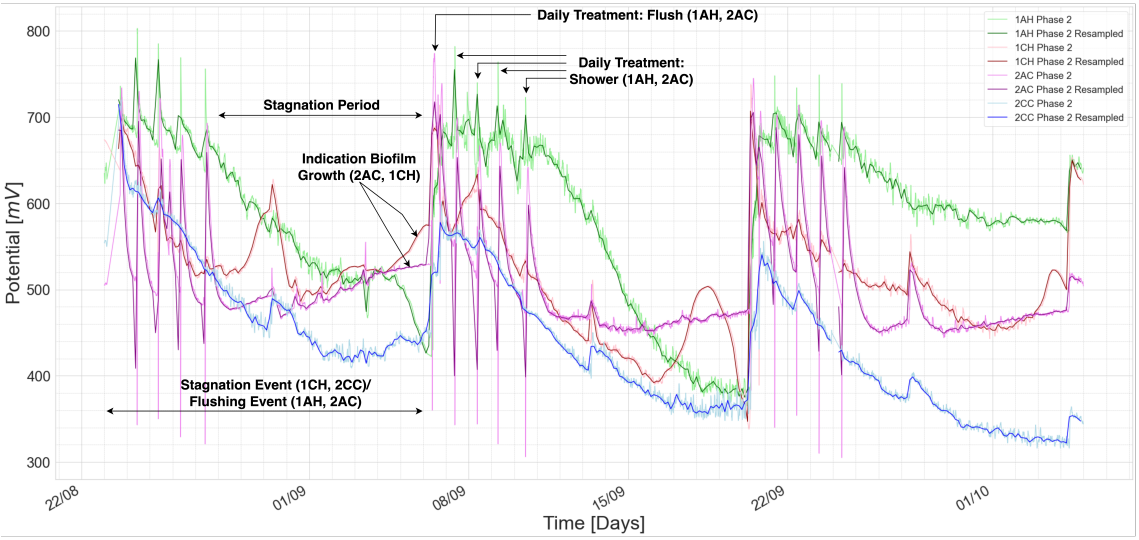


Figure 5: Phase 2 processed ALVIM sensor data.

The increase in potential from the stagnation period of Event 2.2 to Event 2.3 in Figure 5 is a sign of shifting baseline, which, according to the ALVIM technology developers, can be influenced by a number of local environmental conditions. The baseline signal of ALVIM sensors for freshwater is $450 - 650\,mV$, but this is indicative, not absolute (ALVIM, 2020). Gradual decreases and values below $400\,mV$ have in previous studies been linked to low levels of oxygen caused by either: (i) low oxygen concentrations in the water ($< 1\,ppm$), often related to frequent and/or extended no-flow conditions, or (ii) deposit formation at the sensor surface (ALVIM, 2020). Whilst this study contained frequent and extended no-flow conditions, Johansen, 2018 found the drinking water at Valgrinda (where the premise plumbing pilot is located) to have an AOC of $28 \pm 3.7\,\frac{\mu g}{L}$. Drinking water was by van der Kooij, 1992 determined to be biologically stable at AOC-concentrations below $10\,\frac{\mu g}{L}$. A reduction in AOC and dissolved oxygen (DO) indicates growth and activity of biofilm and bacteria (Walker et al., 1991). A biologically stable system is highly unfavourable for microbial growth, therefore, at the recorded AOC, the microbial growth will not reduce DO-levels enough to cause oxygen deficiency in the premise plumbing pilot (Hammes et al., 2010; X. Liu et al., 2015). However, as OPs in drinking water are adapted to nutrient-scarce conditions, reducing the organic carbon levels in drinking water has been deemed by Williams et al., 2015 as an insufficient strategy for control of OPs in drinking water.

4.1 Flushing Events

Flushing was the event most often conducted; during Phase 1, Phase 2, and Phase 4 (Figure 4), and is a central component of both reactive and preventative biofilm control measures. The results of the stagnation periods of flushing events in Figure 6 exhibit on average higher potentials for hot lines than cold, with the exception of Phase 1.

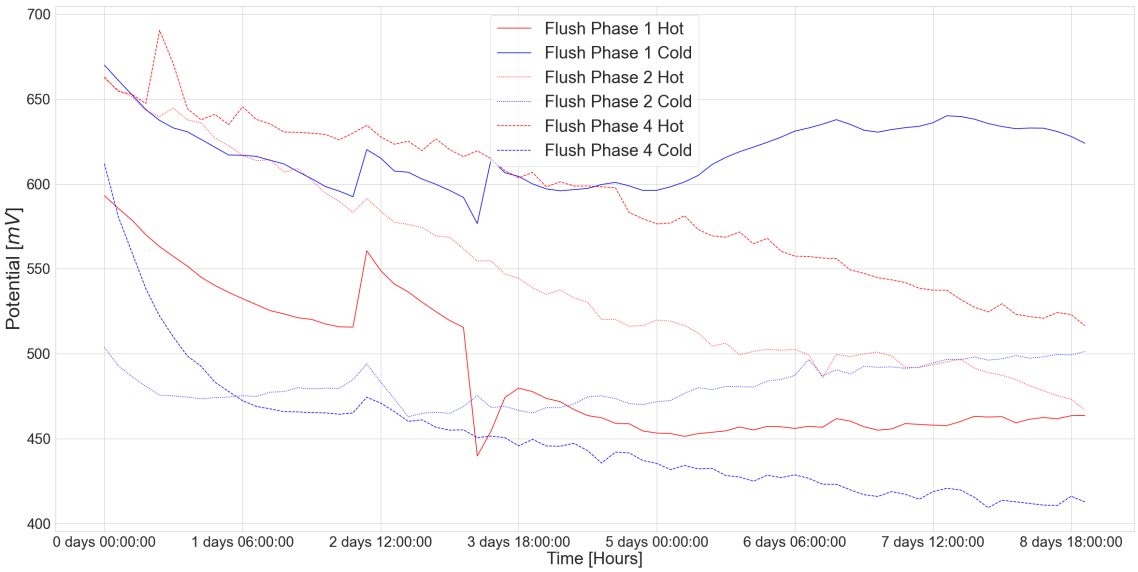


Figure 6: Changes in ALVIM sensor data during stagnation periods of flushing events in response to phase-specific conditions.

During Phase 1, all lines underwent the same event: flushing. From the ALVIM data shown in Figure 6, it was observed that the cold lines consistently had a higher signal than the hot for Phase 1. From Figure 12 in Appendix B it was determined that lines 1CC and 2CC both contributed to this impact. In contrast to this, the Phase 2 BES-data collected by ALVIM sensors (Figure 5), showed that the hot lines, in particular line 1AH, had higher potentials than the cold, except for towards the end of the stagnation periods, in which the cold lines exhibited an increase in potential (figure 6). The ALVIM data from the stagnant lines (1CH, 2CC), revealed a hydraulic connection between lines: the individual daily treatments of flushing were observed in the stagnant lines, despite the no-flow conditions that were ment to prevail there, see Figure 15 in Appendix C. This hydraulic impact was also observed in the stagnant lines (1AH, 2AC) of Phase 3, in relation to the shower events in lines 1CH, 2CC. The effect was less prominent for showers compared to flushing, but nonetheless, the impact was still significant. Consistent with Phase 2 results, the hot lines of Phase 3 had higher values for BES-data recorded by the ALVIM sensors than the cold (Figure 13 in Appendix 14).

The hydraulic connection between lines was not observed in Phase 4 (14 in Appendix B), presumably due to the lack of stagnant lines (Figure 4d). In contrast to the BES-data of the cold lines from Phase 1 and Phase 2 (Figure 6), the hot and cold lines in Phase 4 showed no increase in potential at all throughout the stagnation period.

Despite the impact of flow rate alterations on the ALVIM sensor data examined in Figure 5, Figure 12, Figure 13, and Figure 14, the higher potentials observed in all phases except Phase 1, cannot be attributed to changing flow rates. Whilst the daily treatment flush in Phase 2 had a temperature of 49 °C in hot lines, the daily treatment shower resulted in temperatures close to 60 °C in the hot lines due to thermal mass balance. See Table 3 in Appendix A for temperatures during Event 1.1. The BES-data from the ALVIM sensors did not show any sign of being affected by material (copper vs PEX-A) during any of the phases. The distribution of events across phases as well as their total frequency was uneven, which enabled the comparison of various conditions within each phase, but also prompted difficulties in comparing data sets across phases.

4.2 Biofilm Growth

Identification of biofilm growth was a primary objective of this study. Biofilm growth curves are indicated by clear, gradual increases in BES, in the range of $100 - 150\,mV$ or less. Values below $100\,mV$ values were interpreted as early signs of biofilm development, and dominate the probable

biofilm growth identified. An example of how the biofilm growth curves were extracted is provided in Figure 7. The numeric results of the biofilm growth curves selected appear in Table 1.

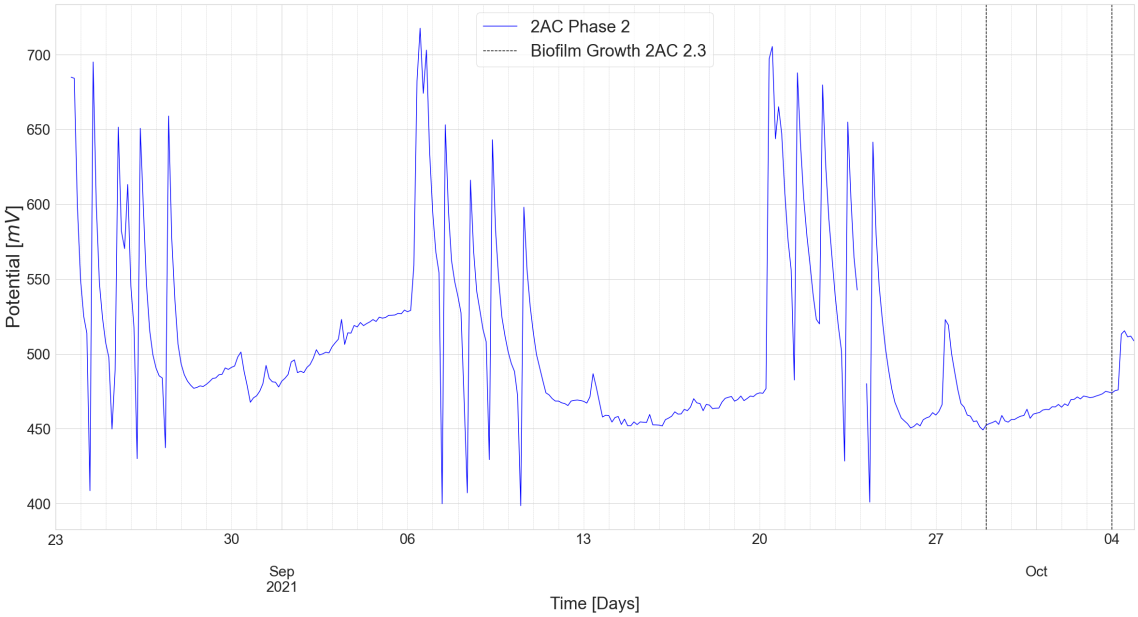


Figure 7: Biofilm growth identified in line 2AC during Event 2.3.

Due to the resampling performed on the curves prior to their extraction, significant outliers were not considered an issue. However, the occurrence of large fluctuations made interpretation of probable biofilm growth challenging at times.

Table 1: Results of ALVIM Sensor Data From Probable Biofilm Growth Curves

Event	**Initial [mV]**	**Maximum [mV]**	**Increase [mV]**	**Rate of Increase [$\frac{mV}{day}$]**
1CH 1.1	491.7	531.7	40.0	10.3
1CC 1.1	646.0	699.0	53.0	20.2
2CC 1.1	551.3	582.6	31.3	8.1
1CH 1.2	412.7	505.2	92.6	11.6
1CC 1.2	542.5	625.7	83.2	15.1
2CC 1.2	496.4	736.2	239.8	24.3
1CH 2.1	498.3	567.9	69.6	14.3
2AC 2.1	491.1	529.3	38.3	5.6
2CC 2.1	423.9	446.7	22.8	11.4
2AC 2.2	458.9	473.3	14.4	2.5
1CH 2.3	452.8	522.7	69.8	26.6
2AC 2.3	452.4	475.0	22.6	4.8
2AC 3.1	376.4	409.5	33.1	6.3
2AC 3.2	381.4	397.7	16.3	6.9
2CC 3.2	412.5	434.4	21.9	11.7
1AH 3.3	551.9	580.7	28.8	8.2
1CH 3.3	656.8	676.5	19.7	13.1
2AC 3.3	368.1	390.7	22.6	5.5

The three largest increases in BES recorded by ALVIM sensors in Figure 1: 239.8, 92.6 and 83.2mV, all occurred during event 1.2, and were ascribed to the prolonged stagnation period of 22 days, providing sufficient time for extended biofilm growth. However, lines 1CC and 1CH displayed a

gradual downwards curve after an initial increase. Nutrient exhaustion or oxygen deficiency after 10 (1CC) or 11 (1CH) days of stagnation was anticipated as the likely cause. During stagnation, few or no new nutrients entered the lines, consequently, nutrient exhaustion is most plausible. This is further supported by the low AOC-values recorded at Valgrinda, as discussed previously. Generally, the rate of increase in potential was larger for curves spanning longer time periods than curves of a short duration.

The curves selected from Phase 3 all had an increase in potential below 35 mV, remarkably lower than Phase 1 and Phase 2. This pattern concurred with the introduction of monochloramine to the system. Norwegian drinking water distribution systems contain no chlorine residuals upon reaching points of use (Table 2 in Appendix A), despite the importance of disinfection residuals in inhibiting biofilm proliferation (LeChevallier, 2021; Pettersen, 2015a; Taylor et al., 2009). Thus, no disinfection residuals were present prior to the addition of monochloramine. The results in Table 1 thus support the hypothesis that monochloramine is efficient and imperative in controlling biofilm. In a study conducted in copper premise plumbing pipes, Buse et al., 2017 found that temperature as well as inoculation of *L. pneumophila* and the amoeba species *Vermamoeba vermiformis* increased the diversity of biofilm composition, with major changes occurring in the period $1-7$ months, as opposed to $0-1$ months. In the same study, *L.pneumophila* persisted for up to 1 month at 20 °C, and up to 7 months at 37 °C at alternating flow and stagnation periods.

For the sake of reproducibility, the FLA *Acanthamoeba* and *Vermamoeba* (previously known as *Hartmanella*), commonly associated with *Legionella spp.*, were monitored in the pilot through biofilm biomass sampling in the lines as part of an independent study (Donlan et al., 2005; Dupuy et al., 2011; Lau et al., 2009; Thomas et al., 2011). Unlike the ALVIM sensors, the biofilm biomass sampling was manual and to some extent, destructive. See Appendix D for the sampling protocol and the impact sampling had on ALVIM sensor data. Biofilm biomass sampling was also conducted to monitor the presence of *Legionella spp.*: subsequent analysis of ssrA and 16S rRNA, conducted in an independent study, revealed the presence of *Legionella* in all samples, see Table 4 in Appendix D for further specifications on gene concentrations. The results of that study showed a significant reduction of *Legionella*-ssrA upon addition of monochloramine to the pilot. However, Hozalski et al., 2020 results led to the conclusion that even in chloraminated plumbing environments biofilms may contain *Legionella spp.*. DNA analysis of biofilm sampling also disclosed the presence of *Vermamoeba*. Donlan et al., 2005 demonstrated that *L. pneumophila* co-cultured with *Vermamoeba vermiformis* in biofilm had a lower susceptibility to monochloramine. In dormant cyst form, FLA have been reported to withstand temperatures up to 80 °C for up to 10 minutes, as well as withstanding 100 ppm free chlorine (Taylor et al., 2009; Thomas et al., 2011). Ohno et al., 2008 reported that at temperatures below 20 °C, the amoeba species *Acanthamoeba castellanii* ingested *L. pneumophila*, whilst at temperatures above 25 °C, *L. pneumophila* infected *A. castellanii*. Buse et al., 2017 confirm a similar behaviour for *V. vermiformis*, additionally concluding that at 37 °C, *L. pneumophila*-colonisation of biofilm was most likely not due to the presence of *V. vermiformis*.

In Phase 4, no biofilm growth was identified, not even in the cold lines, which dominated the probable biofilm growth in Phase 1, Phase 2 and Phase 3. This result conforms to the hypothesis that the combination of 60 °C flushing and monochloramine is sufficient in impeding the growth of biofilm. The lack of growth in cold lines highlights the important role of monochloramine in hindering biofilm growth. Simultaneously, the data confirms that flushing single-handedly is insufficient as a biofilm control measure. In partial opposition to these results, Z. Wang et al., 2021's study revealed that even concentrations of $2-4\ \frac{mg}{L}$ were insufficient in eliminating biofilm unassisted. However, in this study, monochloramine addition was, for a majority of events, done in conjunction with flushing or showers, and thus seldom administered alone.

The rate of bacterial growth has been shown to be lower in copper pipes than other materials, attributed to coppers' biocidal properties (Lehtola et al., 2004; Song et al., 2021; Vargas et al., 2017). The copper concentration in the pilot was therefore monitored, however, the difference in material of the lines of Rig 1 (copper) and Rig 2 (PEX-A) did not have any impact on biofilm growth, as indicated by the ALVIM sensor data in Table 1, which shows no correlation between biofilm growth or the size of it with which rig it occurred in. This is not evidence that copper lacks biocidal properties, but rather indicates an even distribution of biocidal effects as a consequence

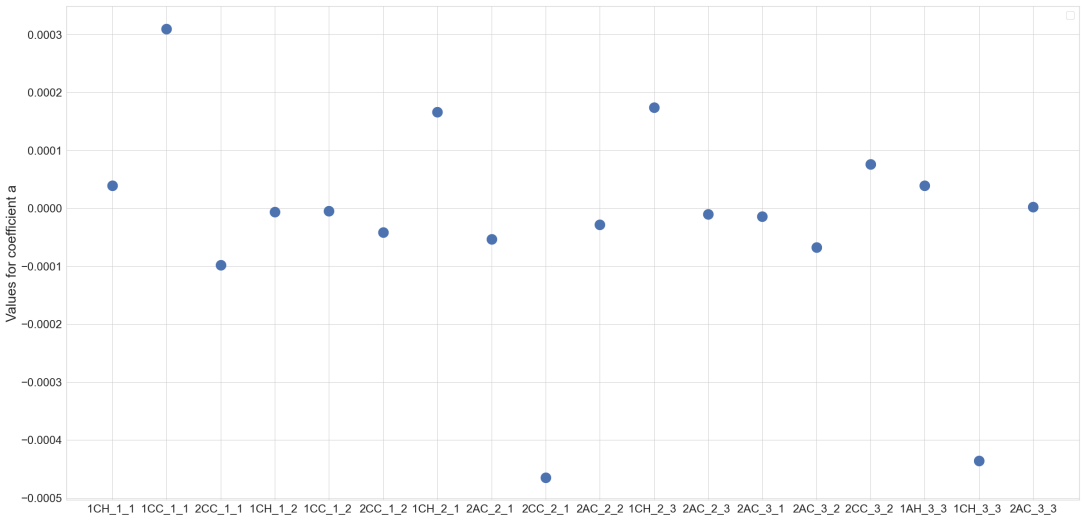
of the pilot material, which, aside from the lines of Rig 2, consisted entirely of copper. Song et al., 2021 concluded that the form of *Legionella* that exits outside of hosts, may be more resistant to copper than previously thought. This is presumably extendable to *Legionella*-associated biofilm. Baron et al., 2014 observed significant reductions in culturable total bacteria and *Legionella spp.* following monochloramine treatment. Marchesi et al., 2012 concluded that on-site monochloramine disinfection in a hospital hot water distribution system significantly reduced the number of sites contaminated with *L. pneumophila*, with no detected changes in levels of nitrate or nitrite. Revetta et al., 2013 demonstrated changes in composition of biofilms over time. Regardless of disinfection, biofilms display a large diversity in build-up, both across and within, as well as over time, which may alter the biofilms' properties (Mayer et al., 1999; Prest et al., 2016; Vu et al., 2009). A change in composition of biofilms can alter BES: Faimali et al., 2010 reported on the challenge of correlating specific bacteria and electrochemical activity. As a consequence of composition changes, biofilm structure also changes, which in turn impacts the likelihood of sloughing off: for example, filamentous biofilm structures have been associated with less sloughing off (Wagner et al., 2010). Filamentous biofilm structures, with mat-like morphologies, have been showcased by Piao et al., 2006 and Khweek et al., 2018 to be a consequence of formation at higher temperatures, specifically 42 °C. In contrast, at 25 °C, biofilms had a mushroom-like structure with water channels, at 38 °C, they were thicker in structure and lacked water channels (Khweek et al., 2018; Piao et al., 2006).

One explanation behind differences in the BES-data recorded by ALVIM sensors from event to event is thus that the biofilm composition and structure is dynamic. This is supported by the results in Table 1: for example, the maximum potential in 1CH 1.1 ($531.7\,mV$) is much closer to the maximum in 1CH 2.1 ($567.9\,mV$) than to 1CH 3.3 ($676.5\,mV$), or 1CC 1.1 ($699.0\,mV$). Buse et al., 2017 confirm temperature as the main driver of microbial diversity.

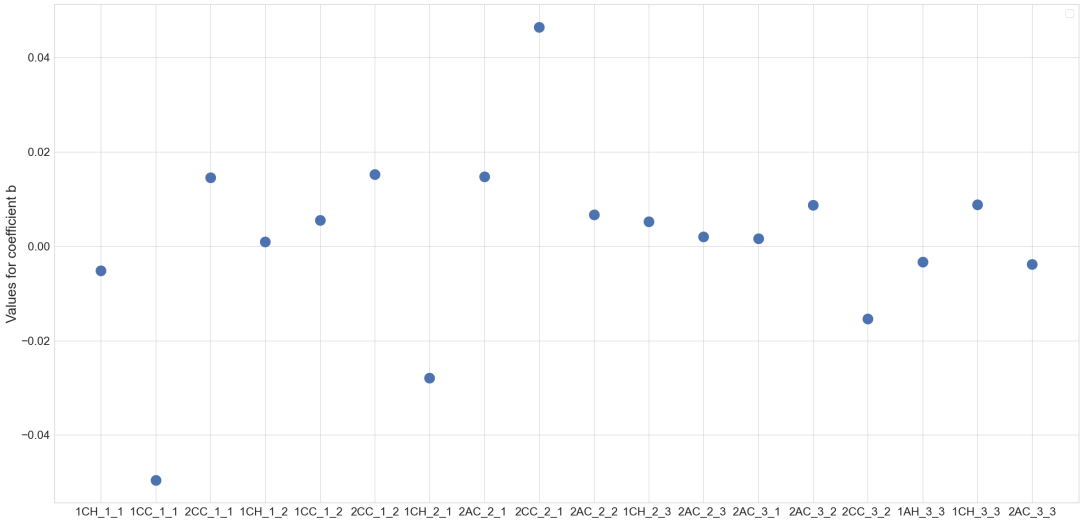
The ALVIM sensors do not provide information on biofilm thickness, instead solely detecting bacteria settling on the working electrode surface, following Equation 2. The thickness, as well as structure of biofilms is significantly influenced by flow conditions and velocity (Schmid et al., 2002). Sudarsan et al., 2016 suggested that even weak biofilms may form tall structures in the case of high spatial proximity to other colonies of biofilm, whose proximitity gives additional protection, reducing drag by $50-100\%$. Thus, the inability of the sensors to investigate thickness is a significant limitation.

4.2.1 Curve Fitting Biofilm Growth Data

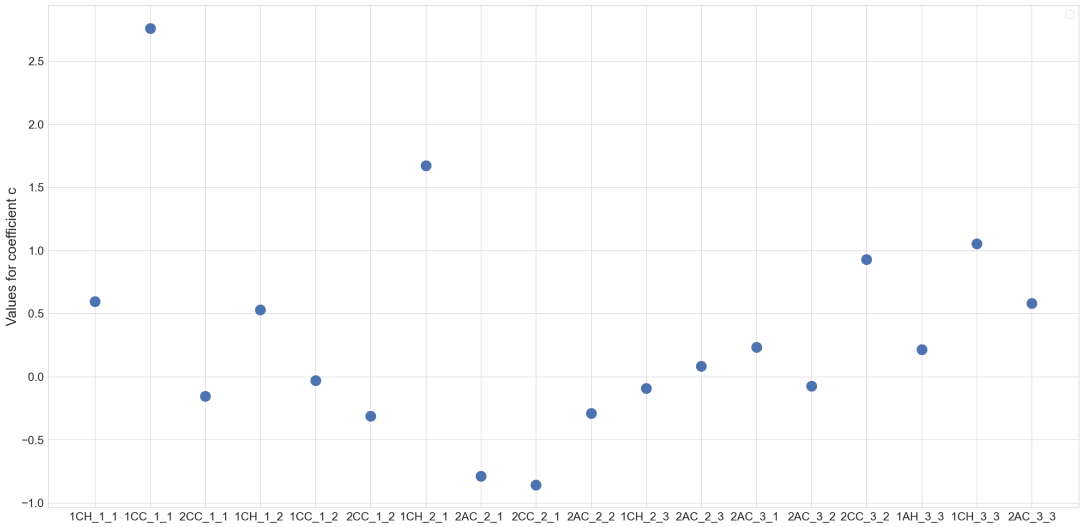
Upon fitting the data to Equation 7, Equation 8, and Equation 9, the polynomial curve (Eq. 8) performed best, on average producing the lowest values for AIC, BIC and χ^2_r, see Table 5 in Appendix E. Figure 8 shows the spread in values for coefficients a - d for the polynomial fit on the biofilm growth curves identified in Table 1.



(a) Coefficient a

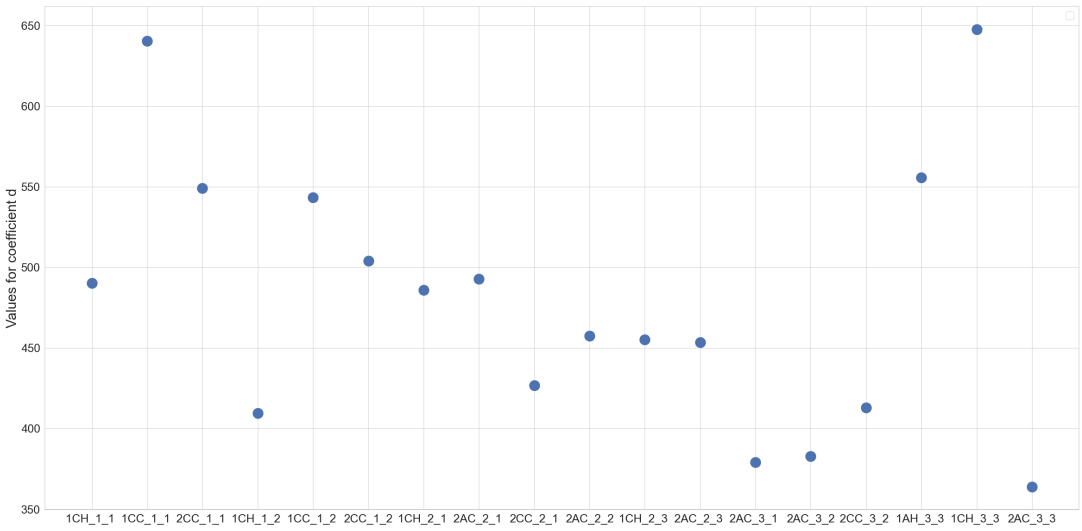


(b) Coefficient b



(c) Coefficient c

Figure 8: Spread of values of coefficients a (a), b (b), c (c) and d (d) for the polynomial function fit on biofilm growth data.



(d) Coefficient d

Figure 8: Spread of values of coefficients a (a), b (b), c (c) and d (d) for the polynomial fit on biofilm growth data.

The individual nature and large variability of the biofilm growth curves is demonstrated in the polynomial fit coefficients in Figure 8. There are no data sets that explicitly distinguish themselves from the rest. This variability exists across event type, line, material and temperature of hot lines.

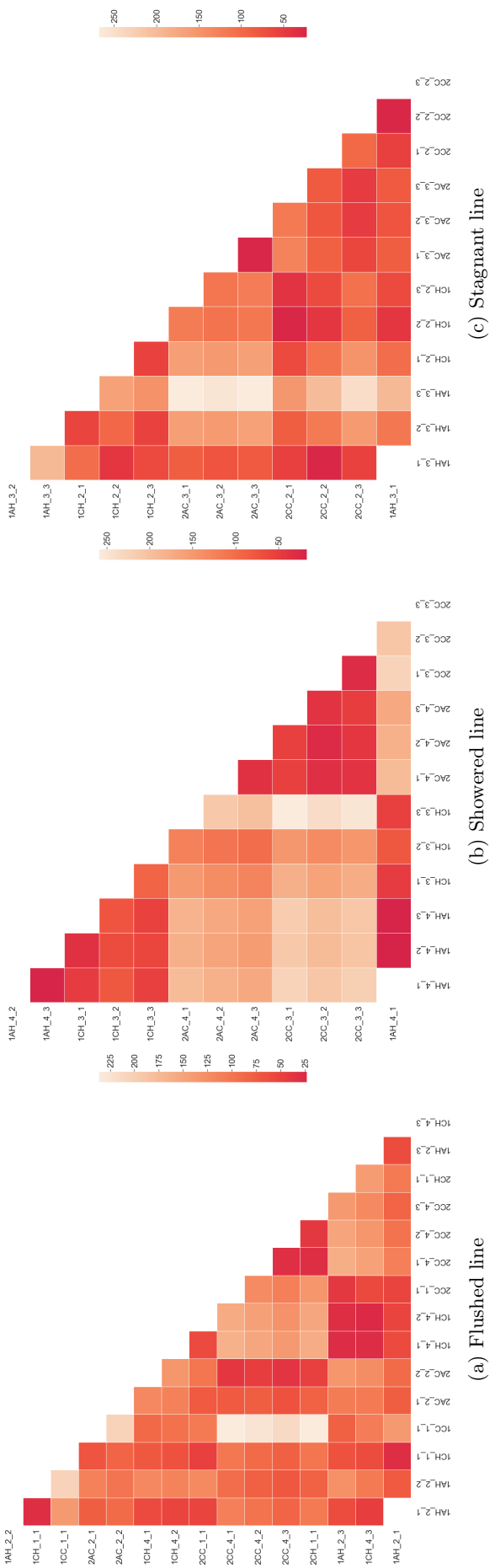
Whilst outliers impacted some of the averaged values of goodness of fit measures for the exponential and linear functions, shown in Table 5 in Appendix E, overall, the other functions performed abysmally in comparison to the polynomial function. Relative errors were an exception to this: as expected, the linear function had the lowest relative errors for all coefficients, whilst the polynomial and exponential had high relative errors. See Table 5 in Appendix E for more values of goodness of fit measures. BIC was consistently slightly higher, displaying values approximately 3 - 10 units more than AIC, which can be attributed to BIC penalising more for complexity than AIC.

4.3 Investigating the Impact of Temperature on ALVIM Sensor Data

Disinfection, temperature and flushing are all widely acknowledged control measures of OPs in premise plumbing, and were investigated in the pilot. Altering the following environmental conditions may contribute to biofilm detachment: (i) temperature, (ii) pH, (iii) nutrient availability, (iv) oxygen deficiency, (v) flow rate (Kostakioti et al., 2013). Removing biofilm or parts of biofilm through sloughing off is an important element in impeding biofilm growth (Mayer et al., 1999).

As discussed in Section 4.1, the BES-data recorded by the ALVIM sensors revealed that the hot lines almost continuously had higher potentials than the cold lines. This trend was temporal: the discrepancy in potential was largest directly subsequent of the last daily treatment of an event, and decreased as the stagnation period progressed.

Heatmaps of RMSE-values, displayed in Figure 9, were constructed between each of the stagnation period data sets of the same event type, specifically: flush (Figure 9a), shower (Figure 9b) and stagnation (Figure 9c). These were created to provide an initial indication and quantification of similarity. Data sets that are more similar to each other have a lower shared RMSE, shown in Figure 9 as a strong, red colour.



(a) Flushed line (b) Showered line (c) Stagnant line

Figure 9: Heatmaps of RMSE-values for the stagnation period data sets of (a) flushed line, (b) showered line, and (c) stagnant line.

The RMSE-heatmaps shown in Figure 9 showcase the lack of impact material had on ALVIM sensor data through the lack of observable trends connected to the rig number, which further supports the aforementioned statements on material in Section 4.2. This is exemplified by examining the data from specific lines: comparison of the stagnation period data sets of 1CH 1.2 and 1CC 1.2 results in $RMSE = 120.3$, whilst the comparison of 1CH 1.2 and 2CH 1.2 gives $RMSE = 97.7$. The higher RMSE between the data sets from rig 1, compared to those of the same temperature, but different rigs, indicates that temperature was more paramount than material in impacting the ALVIM sensor data.

Figure 9 supports the observations made between hot and cold lines previously: hot lines are more similar to each other than to cold lines, and vice versa. This is particularly distinct in the case of shower events (Figure 9b). The flow rates in cold lines were lower than in hot during showers, see Figure 3 in Appendix A, due to the necessary thermal mass balance to achieve the desired output shower temperature of 40 °C.

These data sets were grouped by temperature and event and averaged. In Figure 10, the discrepancy between ALVIM sensor data from hot and cold lines is demonstrated: the mean initial value of hot averages was $620.6\,mV$, the final was $509.9\,mV$, whilst the cold averages started at a mean value of $509.3\,mV$ and ended with $436.9\,mV$ respectively. These values, as well as visual inspection of Figure 10, exhibit how the discrepancy decreased as the stagnation period progressed. This is in support of the proposal that the trend is caused by an exposure to higher temperatures in the hot lines during showering and flushing, rather than an being an attribute of faster biofilm growth in hot lines.

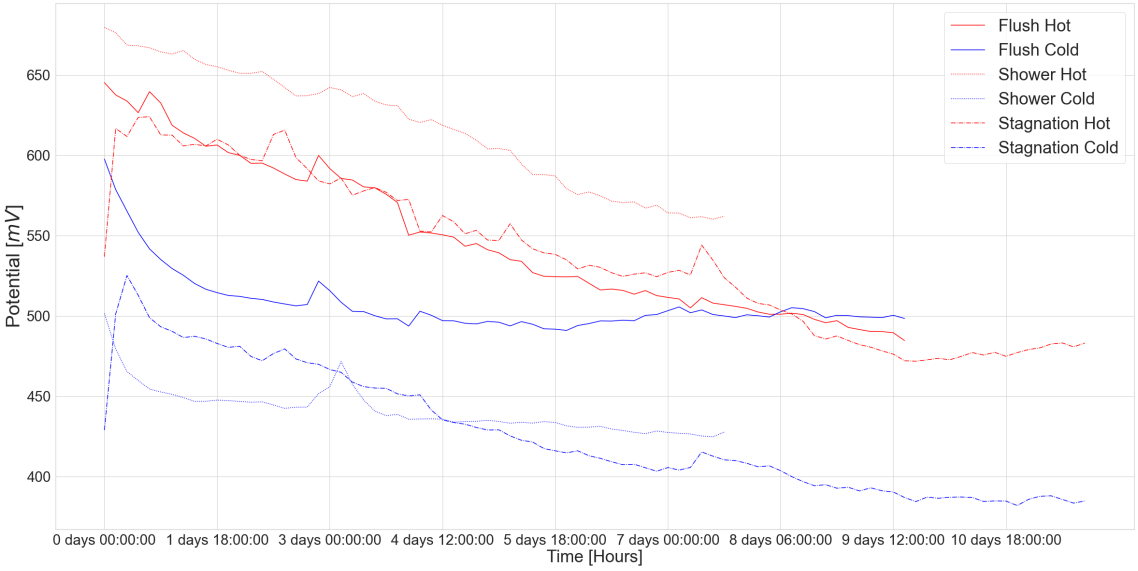


Figure 10: Effect of temperature on ALVIM sensor data during stagnation periods.

Biofilms formed under stagnant conditions are more vulnerable to increases in flow rate, thus showers and flushing have a higher likelihood of dispersing biofilm than stagnation, especially of dispersing biofilms formed under the antecedent stagnation period (Mayer et al., 1999). According to Wagner et al., 2010, high, local shear appears to be the main reason for the sloughing off of biofilm. Tsai, 2005 reported a significant reduction in maximum biofilm biomass upon an increase in flow velocity to $60\,\frac{cm}{s}$, but not between $20-40\,\frac{cm}{s}$, with a concurrent increase in the concentration of bulk water bacteria, an indication of dispersal. Lehtola et al., 2006 found that the formation of biofilms and total bacteria in biofilm increased with increased flow velocity, up to a flow rate of $0.8\,\frac{l}{min}$ in drinking water copper pipes due to increased supply of nutrients. Comparatively, given a line diameter of $1.27\,cm$, calculations making use of the flow data in Table 3 in Appendix A, revealed that the flow velocity in this pilot was much higher than the values specified by Tsai, 2005 for flushing, but not for showering. For example, the flow velocity during flushing of line 1CH, Event 1.1 was $211\,\frac{cm}{s}$, whilst the flow velocity during showering of line 1BC, Event 1.1 was

50 $\frac{cm}{s}$. The flow rates, 16.0 $\frac{L}{min}$ for 1CH 1.1 and 3.8 $\frac{L}{min}$ for 1BC 1.1, clearly surpass the upper boundary of increased biofilm formation specified by Lehtola et al., 2006. Thus, the flow rates and velocities of showers and in particular flushing presumably do not stimulate biofilm growth, but rather result in biofilm detachment.

This is supported by the results of Lipphaus et al., 2014 that report flushing as responsible for the largest decline of microbial load in water in cold water taps of infrequent use. Therefore it is surprising that flush, cold in Figure 10 displayed a slight increase from 6 days onwards, as well as accommodating the highest average potential out of the cold events. However, these results can in part be explained by Hozalski et al., 2020, whose study displayed a decrease of chlorine residuals and a contemporary increase in bacteria of water samples, reaching pre-flushing levels after 6 - 7 days after the last flush. This underlines the necessity for repeating flushes frequently enough, as well as highlighting that stagnation periods shorter than 6 days may be too short for observable bacterial regrowth. Bédard et al., 2018 concluded that rapid increases in culturable cell density after a stagnation period of 1 day originates from bacterial dispersion of biofilm rather than cell growth.

The patterns displayed in Figure 10 were generally repeated upon grouping by temperature, disinfection and event, see Figure 11. One exception was cold, flush, no disinfection, which had a higher final potential than initial. Shower, hot, monochloramine remained at the highest potential, both on average, and in terms of final and initial values, whilst flush, cold, no disinfection had the lowest final value. Lines without disinfection were expected to have a higher recorded BES-value than those with, but this was not always the case: for example stagnation, cold, no disinfection had the lowest final BES-value, which may be an indication of the aforementioned temperature impact. Monochloramine is reported to have a higher efficacy than free chlorine in inactivating biofilm-associated *Lp*1 in copper pipes (Buse et al., 2019).

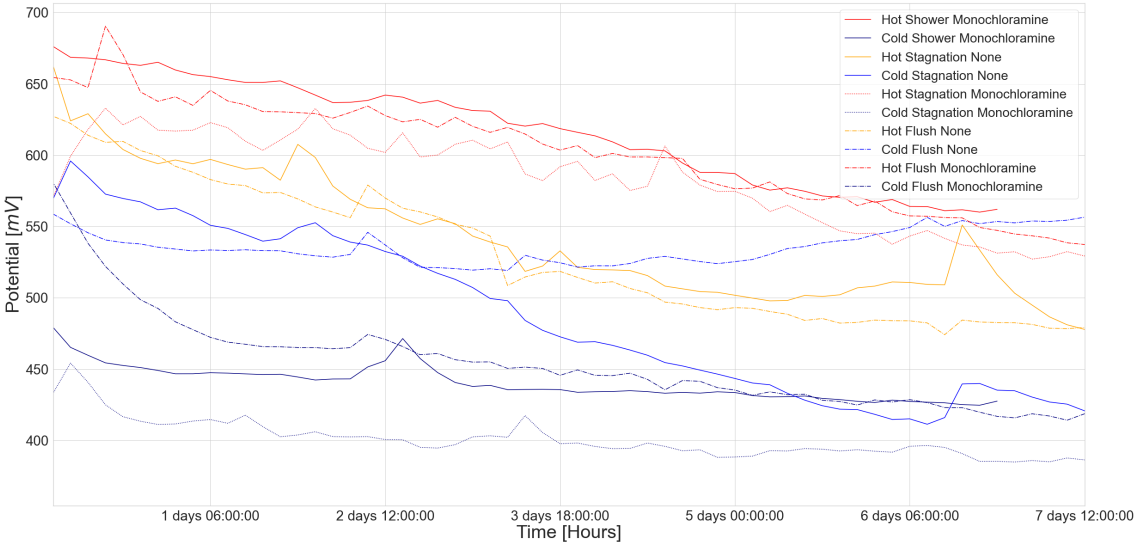


Figure 11: Effect of temperature and NH_2Cl-addition on ALVIM sensor data during stagnation periods.

Flush, cold, no disinfection, displayed an increase in potential after 5 days, see Figure 11. Stagnation, hot, no disinfection also displayed a slight increase after 10 day approximately. As both increases in potential were observed in lines without monochloramine, this is further support of the positive impact disinfection had on the control of biofilm growth.

The elevated BES-data recorded by ALVIM sensors in hot lines was conceivably caused by a combination of various mechanisms, stemming from both the sensors as well as inherent properties of EABs.

Firstly, the temperatures during flushing and showers in hot lines surpassed the ideal working temperatures of the ALVIM sensors. ALVIM sensors have been tested in multiple instances, often

in desalination plants, with temperatures ranging between $11-15$ °C, $11.6-16$ °C and $23.5-27$ °C for in-plant testing (Faimali et al., 2011; Faimali et al., 2008; Faimali et al., 2010; Mollica et al., 1997; Pavanello et al., 2011). By comparison, inlet water temperatures during Event 1.1 ranged between $7.6-12.5$ °C, whilst temperatures in stagnant lines, ranged between $21.3-22.1$ °C during Event 1.1 (Figure 2 in Appendix A). As examined and confirmed by Zlatanovic et al., 2017, inlet and ambient temperatures are the most critical factors for the water temperature inside premise plumbing. The results from their study, combined with the knowledge that water temperatures in the non-insulated pipes of this study reached ambient temperatures at the onset of the stagnation periods due to thermal mass balance, leads to the conclusion that ambient temperature was more critical than inlet temperature in this premise plumbing pilot. Whilst the temperature working range for ALVIM sensors is $[-10, 60]$ °C, the sensor performance has been reported to decrease at temperatures above $35-40$ °C (ALVIM, 2020; Faimali et al., 2011; Pavanello et al., 2011).

The decrease in sensor performance above $35-40$ °C has been partially attributed to accelerated thermal degradation of enzymes that contribute to BES (Amaya et al., 1997; Axelsen et al., 2001; Faimali et al., 2011; Féron et al., 1997; Holthe et al., 1987; Mollica et al., 1988). Enzymatic inhibitors also reportedly impede cathodic current increases attributable to biofilm growth (Faimali et al., 2011). Faimali et al., 2011 demonstrated decreases in cathodic current in response to temperature shocks using water temperatures of 53, 44, 40, and 35 °C. Higher temperatures and longer exposure caused a larger decay. This would however cause lower BES in hot lines than in cold, which is only observed in Phase 1. However, it cannot be ruled out that bacteria growing above 40 °C have enzymes resistant to high water temperatures, causing an ascent in BES. Additionally, the duration of thermal enzymatic degradation is unknown, and may be subject to individual biofilm composition. It is, in this case, indistinguishable from other mechanisms such as increased growth due to favourable growth conditions, biofilm detachment, thermal inactivation of bacteria, and/or changes in biofilm composition.

Thirdly, according to the ALVIM technology developers, higher temperatures are associated with an increase in recorded BES, estimated to be roughly $100\,mV$ for a temperature increase from 20 °C to 60 °C, considered normal for electrochemical sensors.

Furthermore, despite the use of elevated temperatures as a control measure against biofilm growth, to a certain extent elevated temperatures stimulate biofilm growth. From the results of a study by Ahmad et al., 2021, higher water temperatures (25 and 30 °C) resulted in a faster primary biofilm colonisation in the water distribution system studied, but did not yield higher biomass density or biofilm diversity. Temperatures above ambient are associated with higher numbers of *Legionella* and other OPs in premise plumbing: (i) Logan-Jackson et al., 2021 found hot-water samples to contain a higher diversity of *Legionella spp.* than cold-water samples (temperature range: $16.1-32.7$ °C), (ii) hot water taps were reported as a place of amplification of *Legionella* already in 1985 (Bollin et al., 1985), and (iii) higher numbers of both *Legionella spp.* and *L. pneumophila* have been reported during the summer (Proctor et al., 2017; Rhoads et al., 2016; Sharaby et al., 2019), presumably due to higher temperatures (Kao et al., 2015; Katz et al., 1987; L. Liu et al., 2019; Ohno et al., 2003; Whiley et al., 2015). This alone does not guarantee that a multispecies biofilm will experience an overall increase in growth. However, Patil et al., 2010 reported that EABs in wastewater at a temperature range of $0-50$ °C, displayed an increase in bio-electrolytic performance provided temperatures were elevated during initial formation, and that biofilms grown at higher temperatures had a larger decrease in performance when temperatures were lowered.

Next, elevated temperatures caused a more rapid decay of monochloramine, creating conditions beneficial for biofilm growth, thus increasing BES. Monochloramine decay was investigated using the decay constant from Equation 1, see Section 3.2, yielding the following results in bulk water and water samples, respectively: $0.03\,\frac{1}{h}$ and $3.76\,\frac{1}{h}$ for hot lines and $0.01\,\frac{1}{h}$ and $1.16\,\frac{1}{h}$ for cold lines. This data is supported by Hozalski et al., 2020's investigation of flushing in premise plumbing, which disclosed much higher chlorine decay rates in hot and cold water supply lines ($1.04\,\frac{1}{day}$ and $0.74\,\frac{1}{day}$ respectively) than clean glass bottles ($0.05\,\frac{1}{day}$). Xu et al., 2018 showcased how pipes containing biofilms that were allowed to develop had a higher level of free chlorine decay than pipes where this was not the case. The presence of biofilms gives rise to an increased decay of chloramine, especially biofilms harbouring nitrifying bacteria (Hozalski et al., 2020; Y. Zhang et al., 2009). Thus, the

differences in decay constant give evidence of biofilm presence in the pilot, as well as more rapid monochloramine decay in hot lines than cold.

According to Sacher et al., 2019, the decay rate of monochloramine in river and reservoir water increased significantly at elevated temperatures. The decay of monochloramine due to biofilm growth and elevated temperature is to some degree self-reinforcing: monochloramine decays faster at higher temperatures, enabling faster biofilm growth, which in turn results in accelerated monochloramine decay, (Lee et al., 2018; Sacher et al., 2019). However, the monochloramine in this pilot decayed rapidly in both hot and cold lines: after 5 hours of stagnation, almost all chlorine was gone from both sets of lines. Thus, despite a faster decay of monochloramine in hot lines, this alone does not explain the temperature impact on ALVIM sensor data.

Furthermore, the temperature impact could not come from a systematic error in the sensors installed in hot lines, due to the change in sensor configuration between Phase 1 and 2, in which one sensor was moved from 1CC to 1AH, and another from 2CH and 2AC, as the discrepancy in temperature did not follow the sensors.

Lastly, the following environmental conditions may have been more beneficial for biofilm growth in hot lines, yielding a higher BES: light, chlorine residual, material and prior control measures (Pavanello et al., 2011). This is considered unlikely due to the temporal proximity of discrepancy in BES-data to the daily treatments, as well as the data not showing discrepancy with other conditions stimulating or inhibiting biofilm growth such as material.

4.3.1 Curve Fitting Stagnation Period ALVIM Data

Each of stagnation period data sets, grouped by event types: stagnation, shower, flush were fitted using the three functions described in Section 3.4.2 (Equation 7, Equation 8, and Equation 9), resulting in a total of nine attempted curve fittings. Finding a function of best fit was frequently challenging due to the highly event-specific behaviour of the BES-data produced by the ALVIM sensors.

Whilst the polynomial function had the best fit for the stagnation periods subsequent of showers, the exponential function (Equation 7) was of greater interest due to its very distinct demonstration of the temperature discrepancy between lines. All hot lines had negative values for coefficient b, in the range $[-0.00000129, -0.0251]$, whilst all cold lines had positive values in the range $[0.00000317, 0.177]$. The same is true for values of coefficient a (Table 6 in Appendix F), highlighting the differences in curve shape, in addition to the final and initial values discussed previously. This behaviour is in support of observations made from the temperature-grouped and averaged data, see Figure 10, as well as the RMSE-heatmap (Figure 9b). The shower data sets that performed comparatively well on AIC, BIC and χ^2_r, did not always perform well in terms of relative error and vice versa, see Table 6, and Table 7 in Appendix E). For all fits, 2AC 4.1 performed poorly due to a sudden, large spike. As is true for both flushing and stagnation events, the linear function performed well in terms of relative error, but poorly in regards to AIC, BIC and χ^2_r.

It was not feasible to select an overall function of best fit for the stagnation events. This was in large part due to the presence of sudden spikes, that showcased the hydrodynamic connections between lines: the daily treatments in other lines frequently appeared in the ALVIM sensor data of the stagnant lines, see Figure 15 in Appendix C. The cell density of biofilms and the area covered by biofilm was likely highly irregular in time periods of close proximity to daily treatments, which contributed to irregularities in the ALVIM sensor data. Additionally, some data sets deviated significantly in behaviour from the rest, highlighting their individual and event-specific nature. The same is true for data sets of the stagnation periods subsequent of flushing events. Despite the polynomial function denoting itself with a distinctly better performance than the linear and exponential functions, there was a large spread in coefficient values, and very large relative errors for the exponential as well as the polynomial function. The linear function was not selected either, despite its better performance in terms of relative error, due to the poor performance in terms of goodness of fit measures.

5 Conclusion

This study demonstrated the variability in BES-data recorded by ALVIM Srl A001S3 sensors, in particular in relation to the control measures explored in the pilot: flushing, elevated temperatures and monochloramine addition. Rapid flow rate alterations, hydraulic connections within the pilot, and temperatures above $35 - 40$ °C created unfavourable working environments for the sensors. No conclusive biofilm growth was measured in the stagnation periods. All probable biofilm growth curves identified, except one (2CC 1.2), had a gradual increase in ALVIM sensor data well below the specified $100\,mV$. Others (1CH 1.2, 1CC 1.2, 1CH 2.1 and 1CH 2.3) were substantially closer to $100\,mV$ than the remaining data sets. Despite this, none could have been confirmed without the manual biofilm biomass sampling performed in an independent study on the same premise plumbing pilot.

One of the largest benefits of online monitoring systems, are the potential to complement other monitoring methods, by providing new insights into biofilm development in premise plumbing systems. In addition, such systems can act as early warning systems in a way that is not feasible with other monitoring methods, which often are manual. These benefits were not fully displayed in this study, due to the difficulties surrounding the interpretation of ALVIM sensor data. However, the sensors did show the ability to provide information that biomass sampling could not. Replacing invasive, time consuming methods such as biofilm biomass sampling is currently not desirable, nor is it feasible, as they provide different informational aspects about biofilm growth. Despite not demonstrating a high degree of applicability in this pilot, the sensors have useful application areas elsewhere; previous studies such as the one conducted by Pavanello et al., 2011 show that they perform well in steady flow environments in desalination plants. Thus, it can be established that they are more suitable for use under constant flow conditions as well as temperatures below 35 °C. In conclusion, the current ALVIM Srl A001S3 sensor is not suitable for measuring biofilm development with high enough certainty in premise plumbing systems similar to the one utilised in this study, but do provide useful indications of how biofilm growth responds to the various control measures explored.

The probable biofilm growth data support the hypothesis that the combination of 1 $\frac{mg}{L}$ monochloramine and flushing at 60 °C is successful in hindering further biofilm growth. Increasing the flushing temperature from 49 to 60 °, without the addition of monochloramine, proved to be insufficient as a biofilm control measure, highlighting the necessity of adding disinfection or maintaining disinfection residuals in premise plumbing. This is further highlighted by biofilm growth in the hot lines, which also received water at temperatures of 60 °C during showers. In contrast, the addition of 1 $\frac{mg}{L}$ monochloramine resulted in a significant reduction of biofilm growth, in particular in the cold lines, which contained the majority of the probable biofilm growth prior to monochloramine addition. Despite these results, and the support they have in literature, disinfection residuals are largely absent from most Norwegian water distribution and premise plumbing systems.

The material of the lines did not impact the ALVIM sensor data, however, as the majority of the pilot piping consists of copper, also in Rig 2 prior to the PEX-A lines, this does not exclude the presence of copper's biocidal properties. Materials can contribute to changes in biofilm composition and structure over time, together with inlet water quality, disinfection, presence of FLA and other microorganisms, water age and water temperature (Boe-Hansen et al., 2002; Buse et al., 2017; Faimali et al., 2010; Khweek et al., 2018; Piao et al., 2006; Revetta et al., 2013; Shen et al., 2017; Wagner et al., 2010). From this study, no conclusive statements could be made about the biofilm composition and its dynamic nature, apart from the confirmation of *Legionella spp.* and *Vermamoeba* presence in the pilot through biofilm biomass sampling.

It is readily apparent that temperature had a significant impact on the BES-data recorded by the ALVIM Srl A001S3 sensors. Specifically, higher temperatures were associated with elevated potentials, as well as different curve shapes, attributed to a slower descent throughout the stagnation periods, even after restoration of ambient temperatures in the stagnation periods following treatments with elevated temperatures. No conclusions about the exact temperature-related mechanisms, their interaction and importance can be made. However, presumably the temperature impact observed is combination of the following: (i) temperatures of daily treatments surpassed the ideal working conditions of the ALVIM sensors, (ii) accelerated thermal degradation of enzymes,

(iii) presence of enzymes resistant to thermal degradation, counteracting (ii), (iv) increased electrochemical signals with increased temperatures, as specified by the ALVIM technology developers, (v) stimulation of biofilm growth due to increased temperatures, (vi) higher electrolytic performance of biofilms grown at higher temperatures (close and up to 50 °C), (vii) accelerated monochloramine decay at higher temperatures, stimulating biofilm growth, (viii) systematic equipment errors, (ix) more favourable biofilm growth conditions in the hot lines. Mechanisms (viii) and (ix) are deemed least likely due to a lack of support in the results.

5.1 Recommendations for Future Research

The ALVIM sensors responded to rapid alterations in flow rate, and have predominantly been tested under constant flow conditions previously (Pavanello et al., 2011). In the future, studies conducted with a variety of flow conditions should be carried out, in particular in systems associated with significant periods or areas with extended periods of stagnation, and intermittent rapid flow rate increases. This should be done without surpassing temperatures of 35 °C, to distinguish between various interacting mechanisms.

Due to the significant temperature impact on the BES-data recorded by the ALVIM sensors, the sensors should be used with caution at water temperatures above 35 °C. Further research and testing is required to distinguish what mechanism(s) are responsible for the temperature impact observed in this study.

Furthermore, the duration as well as composition of control measures should be examined further. For example, a variety of disinfectants, at a range of concentrations and addition frequencies, could be tested to examine the response of the system in terms of biofilm growth. The length of stagnation periods could also be increased: the results of this study indicate that at least 7 days, preferably more than 10 days, of stagnation are necessary for biofilm growth to become pronounced enough to be identified with greater certainty. This would provide information on the necessary frequency of various control measures, ensuring that the frequency is sufficiently high, without wasting resources. Additionally, it would provide an indication of how system-dependent the length of stagnation periods are before regrowth of biofilm is observed.

Overall, studies investigating the conditions impacting biofilm growth, such as light, temperature, dissolved oxygen, nutrient availability as well as flow rate should be investigated further and in more diverse premise plumbing system models, particularly ones with more complex configurations containing dead-ends, loops and manifolds. Conducting these test with ALVIM Srl A001S3 sensors would provide evidence as to in which systems the sensors have a sufficient degree of serviceability, as well as creating more extensive guidelines regarding their application area.

Bibliography

Ahmad, J. I., Dignum, M., Liu, G., Medema, G. & van der Hoek, J. P. (2021). Changes in biofilm composition and microbial water quality in drinking water distribution systems by temperature increase induced through thermal energy recovery. *Environmental Research*, *194*, 110648. https://doi.org/10.1016/j.envres.2020.110648

Akaike, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle. In E. Parzen, K. Tanabe & G. Kitagawa (Eds.), *Selected Papers of Hirotugu Akaike* (pp. 199–213). Springer New York. https://doi.org/10.1007/978-1-4612-1694-0_15

ALVIM. (2020). *ALVIM - Model A001S3 USER MANUAL*. ALVIM Srl.

Amaya, H. & Miyuki, H. (1997). Mechanism of microbially influenced corrosion on stainless steels in natural seawater and the effect of surface treatment on corrosion resistance. *Zairyo-to-Kankyo*, *46*(8), 481–490. https://doi.org/10.3323/jcorr1991.46.481

Axelsen, S. B. & Rogne, T. (2001). Microbially influenced corrosion of stainless steels in sea water: Critical parameters. *MIC of industrial materials*. https://efcweb.org/efcweb2019_media/Documents/WP_TF/WP10/micinsw.pdf

Baron, J. L., Vikram, A., Duda, S., Stout, J. E. & Bibby, K. (2014). Shift in the Microbial Ecology of a Hospital Hot Water System following the Introduction of an On-Site Monochloramine Disinfection System. *PLOS ONE*, *9*(7), 1–9. https://doi.org/10.1371/journal.pone.0102679

Bédard, E., Fey, S., Charron, D., Lalancette, C., Cantin, P., Dolcé, P., Laferrière, C., Déziel, E. & Prévost, M. (2015). Temperature diagnostic to identify high risk areas and optimize *Legionella pneumophila* surveillance in hot water distribution systems. *Water Research*, *71*, 244–256. https://doi.org/10.1016/j.watres.2015.01.006

Bédard, E., Laferrière, C., Déziel, E. & Prévost, M. (2018). Impact of stagnation and sampling volume on water microbial quality monitoring in large buildings. *PLOS ONE*, *13*(6), 1–14. https://doi.org/10.1371/journal.pone.0199429

Beech, I. B. & Sunner, J. (2004). Biocorrosion: Towards understanding interactions between biofilms and metals. *Current Opinion in Biotechnology*, *15*(3), 181–186. https://doi.org/10.1016/j.copbio.2004.05.001

Bennett, N. D., Croke, B. F. W., Guariso, G., Guillaume, J. H. A., Hamilton, S. H., Jakeman, A. J., Marsili-Libelli, S., Newham, L. T. H., Norton, J. P., Perrin, C., Pierce, S. A., Robson, B., Seppelt, R., Voinov, A. A., Fath, B. D. & Andreassian, V. (2013). Characterising performance of environmental models. *Environmental Modelling & Software*, *40*, 1–20. https://doi.org/10.1016/j.envsoft.2012.09.011

Bishop, P. L., Gibbs, J. T. & Cunningham, B. E. (1997). Relationship Between Concentration and Hydrodynamic Boundary Layers over Biofilms. *Environmental Technology*, *18*(4), 375–385. https://doi.org/10.1080/09593331808616551

Boe-Hansen, R., Albrechtsen, H. J., Arvin, E. & Jørgensen, C. (2002). Dynamics of biofilm formation in a model drinking water distribution system. *Journal of Water Supply: Research and Technology. AQUA*, *51*, 399–406. https://doi.org/10.2166/aqua.2002.0036

Bollin, G. E., Plouffe, J. F., Para, M. F. & Hackman, B. (1985). Aerosols containing *Legionella pneumophila* generated by shower heads and hot-water faucets. *Applied and environmental microbiology*, *50*(5), 1128–1131. https://doi.org/10.1128/aem.50.5.1128-1131.1985

Brink, F., Barendrecht, E. & Visscher, W. (2010). The cathodic reduction of oxygen: A review with emphasis on macrocyclic organic metal complexes as electrocatalysts. *Recueil des Travaux Chimiques des Pays-Bas*, *99*, 253–262. https://doi.org/10.1002/recl.19800990902

Buse, H. Y., Ji, P., Gomez-Alvarez, V., Pruden, A., Edwards, M. A. & Ashbolt, N. J. (2017). Effect of temperature and colonization of *Legionella pneumophila* and *Vermamoeba vermiformis* on bacterial community composition of copper drinking water biofilms. *Microbial Biotechnology*, *10*(4), 773–788. https://doi.org/10.1111/1751-7915.12457

Buse, H. Y., Morris, B. J., Struewing, I. T., Szabo, J. G. & Dudley, E. G. (2019). Chlorine and Monochloramine Disinfection of *Legionella pneumophila* Colonizing Copper and Polyvinyl Chloride Drinking Water Biofilms. *Applied and Environmental Microbiology*, *85*(7), e02956–18. https://doi.org/10.1128/AEM.02956-18

Casini, B., Baggiani, A., Totaro, M., Mansi, A., Costa, A. L., Aquino, F., Miccoli, M., Valentini, P., Bruschi, F., Lopalco, P. L. & Privitera, G. (2018). Detection of viable but non-culturable

Legionella in hospital water network following monochloramine disinfection. *Journal of Hospital Infection*, *98*(1), 46–52. https://doi.org/10.1016/j.jhin.2017.09.006

Cervero-Aragó, S., Schrammel, B., Dietersdorfer, E., Sommer, R., Lück, C., Walochnik, J. & Kirschner, A. (2019). Viability and infectivity of viable but nonculturable *Legionella pneumophila* strains induced at high temperatures. *Water Research*, *158*, 268–279. https://doi.org/10.1016/j.watres.2019.04.009

Chambers, S. T., Slow, S., Scott-Thomas, A. & Murdoch, D. R. (2021). Legionellosis Caused by Non-*Legionella pneumophila* Species, with a Focus on *Legionella longbeachae*. *Microorganisms*, *9*(2). https://doi.org/10.3390/microorganisms9020291

Chan, S., Pullerits, K., Keucken, A., Persson, K. M., Paul, C. J. & Rådström, P. (2019). Bacterial release from pipe biofilm in a full-scale drinking water distribution system. *Biofilms and Microbiomes*, *5*(1). https://doi.org/10.1038/s41522-019-0082-9

Charnock, C. & Kjønnø, O. (2000). Assimilable organic carbon and biodegradable dissolved organic carbon in Norwegian raw and drinking waters. *Water Research*, *34*(10), 2629–2642. https://doi.org/10.1016/S0043-1354(00)00007-5

Choi, Y., Byun, S. H., Jang, H. J., Kim, S. E. & Choi, Y. J. (2022). Comparison of disinfectants for drinking water: Chlorine gas vs. on-site generated chlorine. *Environmental Engineering Research*, *27*(1), 200543–. https://doi.org/10.4491/eer.2020.543

Cunningham, W. P. & Cunningham, M. A. (2004). *Principles Of Environmental Science: Inquiry And Applications 2nd Edition*. McGraw-Hill Higher Education.

De Giglio, O., Fasano, F., Diella, G., Lopuzzo, M., Napoli, C., Apollonio, F., Brigida, S., Calia, C., Campanale, C., Marzella, A., Pousis, C., Rutigliano, S., Triggiano, F., Caggiano, G. & Montagna, M. T. (2019). *Legionella* and legionellosis in touristic-recreational facilities: Influence of climate factors and geostatistical analysis in southern italy (2001–2017). *Environmental Research*, *178*, 108721. https://doi.org/10.1016/j.envres.2019.108721

Dennis, P. J., Brenner, D. J., Thacker, W. L., Wait, R., Vesey, G., Walt, A. G. S. & Benson, R. F. (1993). Five New *Legionella* Species Isolated from Water. *International Journal of Systematic and Evolutionary Microbiology*, *43*(2), 329–337. https://doi.org/10.1099/00207713-43-2-329

Dias, V. C. F., Durand, A. A., Constant, P., Prévost, M. & Bédard, E. (2019). Identification of Factors Affecting Bacterial Abundance and Community Structures in a Full-Scale Chlorinated Drinking Water Distribution System. *Water*, *11*(3). https://doi.org/10.3390/w11030627

Donlan, R., Forster, T., Murga, R., Brown, E., Lucas, C., Carpenter, J. & Fields, B. (2005). *Legionella pneumophila* associated with the protozoan *hartmannella vermiformis* in a model multi-species biofilm has reduced susceptibility to disinfectants. *Biofouling*, *21*(1), 1–7. https://doi.org/10.1080/08927010500044286

Dupuy, M., Mazoua, S., Berne, F., Bodet, C., Garrec, N., Herbelin, P., Ménard-Szczebara, F., Oberti, S., Rodier, M. .-., Soreau, S., Wallet, F. & Héchard, Y. (2011). Efficiency of water disinfectants against *Legionella pneumophila* and *Acanthamoeba*. *Water Research*, *45*(3), 1087–1094. https://doi.org/10.1016/j.watres.2010.10.025

Erable, B., Duţeanu, N. M., Ghangrekar, M. M., Dumas, C. & Scott, K. (2010). Application of electro-active biofilms. *Biofouling*, *26*(1), 57–71. https://doi.org/10.1080/08927010903161281

European Centre for Disease Prevention and Control. (2018). *Legionnaires' disease* [In: ECDC. Annual epidemiological report for 2018. Accessed: 18.04.2022]. Stockholm.

European Parliament, Council of the European Union. (2020). Directive (EU) 2020/2184 of the European Parliament and of the Council of 16 December 2020 on the quality of water intended for human consumption. Retrieved 23rd March 2022, from https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32020L2184

Faimali, M., Benedetti, A., Giovanni, P., Chelossi, E., Wrubl, F. & Mollica, A. (2011). Evidence of enzymatic catalysis of oxygen reduction on stainless steels under marine biofilm. *Biofouling*, *27*(4), 375–384. https://doi.org/10.1080/08927014.2011.576756

Faimali, M., Chelossi, E., Garaventa, F., Corrà, C., Greco, G. & Mollica, A. (2008). Evolution of oxygen reduction current and biofilm on stainless steels cathodically polarised in natural aerated seawater. *Electrochimica Acta*, *54*(1), 148–153. https://doi.org/10.1016/j.electacta.2008.02.115

Faimali, M., Chelossi, E., Pavanello, G., Benedetti, A., Vandecandelaere, I., Vos, P. D., Vandamme, P. & Mollica, A. (2010). Electrochemical activity and bacterial diversity of natural marine

biofilm in laboratory closed-systems. *Bioelectrochemistry (Amsterdam, Netherlands)*, *78*, 30–80. https://doi.org/10.1016/j.bioelechem.2009.04.012

Falkinham, J. O., Pruden, A. & Edwards, M. (2015). Opportunistic premise plumbing pathogens: Increasingly important pathogens in drinking water. *Pathogens*, *4*(2), 373–386. https://doi.org/10.3390/pathogens4020373

Féron, D., Dupont, I. & Novel, G. (1997). Influence of micro-organisms on the free corrosion potentials of stainless steels in natural sea water. *European Federation of Corrosion*, (22), 103–122. https://stanford.idm.oclc.org/login?url=https://app.knovel.com/hotlink/toc/id:kpAMICEFCF/aspects-of-microbially?kpromoter=marc

Fields, B. S. (1996). The molecular ecology of legionellae. *Trends in Microbiology*, *4*(7), 286–290. https://doi.org/10.1016/0966-842X(96)10041-X

Flemming, H. C. (2001). Relevance of microbial extracellular polymeric substances (EPSs)–Part I: Structural and ecological aspects. *Water science and technology : a journal of the International Association on Water Pollution Research*, *43*(6), 1–8. https://pubmed.ncbi.nlm.nih.gov/11381954/

Flemming, H. C. & Wuertz, S. (2019). Bacteria and archaea on Earth and their abundance in biofilms. *Nature Reviews Microbiology*, *17*(4), 247–260. https://doi.org/10.1038/s41579-019-0158-9

Gavaldà, L., Garcia-Nuñez, M., Quero, S., Gutierrez-Milla, C. & Sabrià, M. (2019). Role of hot water temperature and water system use on *Legionella* control in a tertiary hospital: An 8-year longitudinal study. *Water research*, *149*, 460–466. https://doi.org/10.1016/j.watres.2018.11.032

Gomes, T. S., Gjiknuri, J., Magnet, A., Vaccaro, L., Ollero, D., Izquierdo, F., Fenoy, S., Hurtado, C. & del Águila, C. (2018). The influence of *Acanthamoeba–Legionella* interaction in the virulence of two different *Legionella* species. *Frontiers in Microbiology*, *9*. https://doi.org/10.3389/fmicb.2018.02962

Graham, F. F., Hales, S., White, P. S. & Baker, M. (2020). Review Global seroprevalence of legionellosis - a systematic review and meta-analysis. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-63740-y

Hall-Stoodley, L. & Stoodley, P. (2005). Biofilm formation and dispersal and the transmission of human pathogens. *Trends in Microbiology*, *13*(1), 7–10. https://doi.org/10.1016/j.tim.2004.11.004

Hammes, F., Berger, C., Köster, O. & Egli, T. (2010). Assessing biological stability of drinking water without disinfectant residuals in a full-scale water supply system. *Journal of Water Supply: Research and Technology-Aqua*, *59*(1), 31–40. https://doi.org/10.2166/aqua.2010.052

Harp, D. L. (1995). *Current Technology of Chlorine Analysis for Water and Wastewater*. HACH Company. https://books.google.no/books?id=H5oTPwAACAAJ

Herwaldt, L. A. & Marra, A. R. (2018). *Legionella*: a reemerging pathogen. *Current Opinion in Infectious Diseases*, *31*(4). https://journals.lww.com/co-infectiousdiseases/Fulltext/2018/08000/Legionella__a_reemerging_pathogen.10.aspx

Holthe, R., Gartland, P. O. & Bardal, E. (1987). Influence of the microbial slime layer on the electrochemical properties of stainless steel in sea water, 617–623.

Hossain, S., Chow, C. W. K., Cook, D., Sawade, E. & Hewa, G. A. (2022). Review of Nitrification Monitoring and Control Strategies in Drinking Water System. *International journal of environmental research and public health*, *19*(7), 4003. https://doi.org/10.3390/ijerph19074003

Hozalski, R. M., LaPara, T. M., Zhao, X., Kim, T., Waak, M. B., Burch, T. & McCarty, M. (2020). Flushing of Stagnant Premise Water Systems after the COVID-19 Shutdown Can Reduce Infection Risk by *Legionella* and *Mycobacterium spp. Environmental Science & Technology*, *54*(24), 15914–15924. https://doi.org/10.1021/acs.est.0c06357

Huang, R. & Hu, N. (2001). Direct electrochemistry and electrocatalysis with horseradish peroxidase in eastman aq films. *Bioelectrochemistry*, *54*(1), 75–81. https://doi.org/10.1016/S1567-5394(01)00113-X

Johansen, I. E. (2018). *Assimilerbart organisk karbon i drikkevann i Trondheim* (Master's thesis). Norwegian Unviversity of Science and Technology, NTNU. Norway. http://hdl.handle.net/11250/2565290

Johnsen, R. & Bardal, E. (1985). Cathodic properties of different stainless-steels in natural seawater. *CORROSION*, *41*(5), 296–302. https://doi.org/10.5006/1.3582007

Kao, P. M., Hsu, B. M., Chang, T. Y., Hsu, T. K., Tzeng, K. J. & Huang, Y. L. (2015). Seasonal variation of *Legionella* in taiwan's reservoir and its relationships with environmental factors. *Environmental science and pollution research international*, *22*(8). https://doi.org/10.1007/s11356-014-3819-2

Katz, S. M. & Hammel, J. M. (1987). The effect of drying, heat, and pH on the survival of *Legionella pneumophila*. *Annals of clinical and laboratory science*, *17*(3), 150–156.

Khweek, A. A. & Amer, A. O. (2018). Factors Mediating Environmental Biofilm Formation by *Legionella pneumophila*. *Frontiers in Cellular and Infection Microbiology*, *8*. https://doi.org/10.3389/fcimb.2018.00038

Kirschner, A. K. (2016). Determination of viable legionellae in engineered water systems: Do we find what we are looking for? *Water Research*, *93*, 276–288. https://doi.org/10.1016/j.watres.2016.02.016

Kostakioti, M., Hadjifrangiskou, M. & Hultgren, S. J. (2013). Bacterial biofilms: Development, dispersal, and therapeutic strategies in the dawn of the postantibiotic era. *Cold Spring Harb Perspect Med.*, *3*(4). https://doi.org/10.1101/cshperspect.a010306

Krakk, T. (2021). *Evaluating the Effect of Flushing to Reduce Bacterial Loads in Plumbing Systems* (Master's thesis). Norwegian Unviversity of Science and Technology, NTNU. Norway. https://hdl.handle.net/11250/2826367

Krzeminski, P., Vogelsang, C., Meyn, T., Köhler, S., Poutanen, H., de Wit, H. & Uhl, W. (2019). Natural organic matter fractions and their removal in full-scale drinking water treatment under cold climate conditions in Nordic capitals. *Journal of Environmental Management*, *241*, 427–438. https://doi.org/10.1016/j.jenvman.2019.02.024

Kuha, J. (2004). AIC and BIC: Comparisons of Assumptions and Performance. *Sociological Methods & Research*, *33*(2), 188–229. https://doi.org/10.1177/0049124103262065

Lai, M. E. & Bergel, A. (2000). Electrochemical reduction of oxygen on glassy carbon: Catalysis by catalase. *Journal of Electroanalytical Chemistry*, *494*(1), 30–40. https://doi.org/0.1016/S0022-0728(00)00307-7

Lai, M. E., Scotto, V. & Bergel, A. (1999). Analytical characterization of natural marine biofilm: A tool for understanding biocorrosion of s.s. in seawater. *Proceedings of the 10th International Congress on Marine Corrosion and fouling*, 38–46.

Lau, H. Y. & Ashbolt, N. J. (2009). The role of biofilms and protozoa in *Legionella* pathogenesis: implications for drinking water. *Journal of Applied Microbiology*, *107*(2), 368–378. https://doi.org/10.1111/j.1365-2672.2009.04208.x

LeChevallier, M. W. (2021). Guidance on Developing a *Legionella pneumophila* Monitoring Program for Utility Distribution Systems. *Health Education and Public Health*, *3*(6), 357–364. https://doi.org/10.31488/HEPH.158

Lee, W. H., Pressman, J. G. & Wahman, D. G. (2018). Three-Dimensional Free Chlorine and Monochloramine Biofilm Penetration: Correlating Penetration with Biofilm Activity and Viability. *Environmental Science & Technology*, *52*(4), 1889–1898. https://doi.org/10.1021/acs.est.7b05215

Lehtola, M. J., Laxander, M., Miettinen, I. T., Hirvonen, A., Vartiainen, T. & Martikainen, P. J. (2006). The effects of changing water flow velocity on the formation of biofilms and water quality in pilot distribution system consisting of copper or polyethylene pipes. *Water Research*, *40*(11), 2151–2160. https://doi.org/10.1016/j.watres.2006.04.010

Lehtola, M. J., Miettinen, I. T., Keinänen, M. M., Kekki, T. K., Laine, O., Hirvonen, A., Vartiainen, T. & Martikainen, P. J. (2004). Microbiology, chemistry and biofilm development in a pilot drinking water distribution system with copper and plastic pipes. *Water Research*, *38*(17), 3769–3779. https://doi.org/10.1016/j.watres.2004.06.024

Leslie, E., Hinds, J. & Hai, F. I. (2021). Causes, Factors, and Control Measures of Opportunistic Premise Plumbing Pathogens—A Critical Review. *Applied Sciences*, *11*(10). https://doi.org/10.3390/app11104474

Lesnik, R., Brettar, I. & Höfle, M. (2015). *Legionella* species diversity and dynamics from surface reservoir to tap water: From cold adaptation to thermophily. *The ISME journal*, *10*. https://doi.org/10.1038/ismej.2015.199

Lin, Y. E., Stout, J. E. & Yu, V. L. (2011). Controlling *Legionella* in hospital drinking water: An evidence-based review of disinfection methods. *Infection control and hospital epidemiology*, *32*(2), 166–73. https://doi.org/10.1086/657934

Lin, Y., Stout, J. E., Yu, V. & Vidic, R. D. (1998). Disinfection of water distribution systems for *Legionella*. *Seminars in respiratory infections*, *13*(2), 147–59.

Lipphaus, P., Hammes, F., Kötzsch, S., Green, J., Gillespie, S. & Nocker, A. (2014). Microbiological tap water profile of a medium-sized building and effect of water stagnation. *Environmental Technology*, *35*(5), 620–628. https://doi.org/10.1080/09593330.2013.839748

Liu, G., Bakker, G. L., Li, S., Vreeburg, J. H. G., Verberk, J. Q. J. C., Medema, G. J., Liu, W. T. & Dijk, J. C. V. (2014). Pyrosequencing Reveals Bacterial Communities in Unchlorinated Drinking Water Distribution System: An Integral Study of Bulk Water, Suspended Solids, Loose Deposits, and Pipe Wall Biofilm. *Environmental Science & Technology*, *48*(10), 5467–5476. https://doi.org/10.1021/es5009467

Liu, L., Xing, X., Hu, C. & Wang, H. (2019). One-year survey of opportunistic premise plumbing pathogens and free-living amoebae in the tap-water of one northern city of China. *Journal of environmental sciences (China)*, *77*, 20–31. https://doi.org/10.1016/j.jes.2018.04.020

Liu, X., Wang, J., Liu, T., Kong, W., He, X., Jin, Y. & Zhang, B. (2015). Effects of Assimilable Organic Carbon and Free Chlorine on Bacterial Growth in Drinking Water. *PLOS ONE*, *10*(6), 1–11. https://doi.org/10.1371/journal.pone.0128825

Logan, B. E., Rossi, R., Ragab, A. & Saikaly, P. E. (2019). Electroactive microorganisms in bioelectrochemical systems. *Nature Reviews Microbiology*, *17*(5), 307–319. https://doi.org/10.1038/s41579-019-0173-x

Logan-Jackson, A. R., Flood, M. & Rose, J. B. (2021). Enumeration and characterization of five pathogenic *Legionella* species from large research and educational buildings. *Environ. Sci.: Water Res. Technol.*, *7*(2), 321–334. https://doi.org/10.1039/D0EW00893A

Lovdata. (1995). Act relating to control of communicable diseases [Last updated: 04/04/2022. Accessed: 02/06/2022]. https://lovdata.no/dokument/NL/lov/1994-08-05-55?q=vern%20mot%20smittsomme%20sykdommer

Lovdata. (1996). Bathing facilities, pool baths and saunas health act [Last upated: 04/10/2021.Accessed: 02/06/2022]. https://lovdata.no/dokument/NL/lov/2008-06-27-71

Lovdata. (2006). Act relating to working environment, working hours and employment protection, etc. [working environment act] [Last updated: 08/04/2022. Accessed: 02/06/2022]. https://lovdata.no/dokument/NL/lov/2005-06-17-62?q=lov%20om%20arbeids

Lovdata. (2008). The planning and building act [Accessed: 16/05/2022]. https://lovdata.no/dokument/NL/lov/2008-06-27-71

Lovdata. (2011). The norwegian public health act [Accessed: 16/05/2022]. https://lovdata.no/dokument/NL/lov/2011-06-24-29

Lovdata. (2016). Regulations on water supply and drinking water (drinking water regulations) [Accessed: 16/05/2022]. https://lovdata.no/dokument/SF/forskrift/2016-12-22-1868/KAPITTEL_1#KAPITTEL_1

Marchesi, I., Cencetti, S., Marchegiano, P., Frezza, G., Borella, P. & Bargellini, A. (2012). Control of *Legionella* contamination in a hospital water distribution system by monochloramine. *American journal of infection control*, *40*(3), 279–281. https://doi.org/10.1016/j.ajic.2011.03.008

Martiny, A. C., Jørgensen, T. M., Albrechtsen, H. J., Arvin, E. & Molin, S. (2003). Long-term succession of structure and diversity of a biofilm formed in a model drinking water distribution system. *Applied and environmental microbiology*, *69*(11), 6899–6907. https://doi.org/10.1128/AEM.69.11.6899-6907.2003

Mayer, C., Moritz, R., Kirschner, C., Borchard, W., Maibaum, R. & Flemming, H. C. (1999). The role of intermolecular interactions: Studies on model systems for bacterial biofilms. *International journal of biological macromolecules*, *26*(1), 3–16. https://doi.org/10.1016/s0141-8130(99)00057-4

Mercante, J. W. & Winchell, J. M. (2015). Current and Emerging *Legionella* Diagnostics for Laboratory and Outbreak Investigations, journal = Clinical Microbiology Reviews. *28*(1), 95–133. https://doi.org/10.1128/CMR.00029-14

Miller, S. J. (2006). The Method Of Least Squares. https://web.williams.edu/Mathematics/sjmiller/public_html/BrownClasses/54/handouts/MethodLeastSquares.pdf

Miyashita, N., Higa, F., Aok, Y., Kikuchi, T., Seki, M., Tateda, K., Maki, N., Uchino, K., Ogasawara, K., Kiyota, H. & Watanabe, A. (2020). Distribution of *Legionella* species and serogroups in patients with culture-confirmed *Legionella* pneumonia. *Journal of Infection and Chemotherapy*, *26*(5), 411–417. https://doi.org/10.1016/j.jiac.2019.12.016

Mollica, A., Traverso, E. & Thierry, D. (1997). On oxygen reduction depolarisation induced by biofilm growth on stainless steel in sea water. *European Federation of Corrosion publication*, *22*, 51–63.

Mollica, A., Trevis, A., Traverso, E., Ventura, G., De Carolis, G. & Dellepiane, R. (1988). Crevice Corrosion Resistance of Stainless Steels in Natural Seawater in the Temperature Range of 25 to 40 °C. *Corrosion*, *44*(4), 194–198. https://doi.org/10.5006/1.3583924

Muder, R. R. & Victor, L. Y. (2002). Infection Due to *Legionella* Species Other Than *L. pneumophila*. *Clinical Infectious Diseases*, *35*(8), 990–998. https://doi.org/10.1086/342884

Muhammad, M. H., Idris, A. L., Fan, X., Guo, Y., Yu, Y., Jin, X., Qiu, J., Guan, X. & Huang, T. (2020). Beyond Risk: Bacterial Biofilms and Their Regulating Approaches. *Frontiers in Microbiology*, *11*. https://doi.org/10.3389/fmicb.2020.00928

MWE. (n.d.). Polywipe™range [Date accessed: 07/06/2022]. https://www.mwe.co.uk/product/polywipe-range/

Newville, M., Otten, R., Nelson, A., Ingargiola, A., Stensitzki, T., Allan, D., Fox, A., Carter, F., Michał, Osborn, R., Pustakhod, D., lneuhaus, Weigand, S., Glenn, Deil, C., M., Hansen, A. L. R., Pasquevich, G., Foks, L., ... Persaud, A. (2021). *Lmfit/lmfit-py: 1.0.3* (Version 1.0.3). Zenodo. https://doi.org/10.5281/zenodo.5570790

Norwegian Building Authority. (2017). Regulations on technical requirements for construction works [Accessed: 16/05/2022]. https://dibk.no/regelverk/byggteknisk-forskrift-tek17/

Nøst, T. (2020). Monitoring of water resources in trondheim 2021. results [Date accessed: 03/06/2022]. https://www.trondheim.kommune.no/tema/veg-vann-og-avlop/vann-og-avlop/om-vann-og-avlop/overvaking-av-vannkvalitet-og-miljotilstand/

Ohno, A., Kato, N., Sakamoto, R., Kimura, S. & Yamaguchi, K. (2008). Temperature-dependent parasitic relationship between *Legionella pneumophila* and a free-living amoeba (*Acanthamoeba castellanii*). *Applied and environmental microbiology*, *74*(14), 4585–8. https://doi.org/10.1128/AEM.00083-08

Ohno, A., Kato, N., Yamada, K. & Yamaguchi, K. (2003). Factors influencing survival of *Legionella pneumophila* serotype 1 in hot spring water and tap water. *Applied and environmental microbiology*, *69*(5), 2540–2547. https://doi.org/10.1128/AEM.69.5.2540-2547.2003

Pascale, M. R., Mazzotta, M., Salaris, S., Girolamini, L., Grottola, A., Simone, M. L., Cordovana, M., Bisognin, F., Dal Monte, P., Sabattini, S. M. A. B., Viggiani, M. & Cristino, S. (2020). Evaluation of MALDI–TOF Mass Spectrometry in Diagnostic and Environmental Surveillance of *Legionella* Species: A Comparison With Culture and Mip-Gene Sequencing Technique. *Frontiers in Microbiology*, *11*. https://doi.org/10.3389/fmicb.2020.589369

Patil, S. A., Harnisch, F., Kapadnis, B. & Schröder, U. (2010). Electroactive mixed culture biofilms in microbial bioelectrochemical systems: The role of temperature for biofilm formation and performance. *Biosensors and Bioelectronics*, *26*(2), 803–808. https://doi.org/10.1016/j.bios.2010.06.019

Patil, S. A., Harnisch, F., Koch, C., Hübschmann, T., Fetzer, I., Carmona-Martínez, A. A., Müller, S. & Schröder, U. (2011). Electroactive mixed culture derived biofilms in microbial bioelectrochemical systems: The role of pH on biofilm formation, performance and composition. *Bioresource Technology*, *102*(20), 9683–9690. https://doi.org/10.1016/j.biortech.2011.07.087

Pavanello, G., Faimali, M., Pittore, M., Mollica, A., Mollica, A. & Mollica, A. (2011). Exploiting a new electrochemical sensor for biofilm monitoring and water treatment optimization. *FEMS Microbiology Ecology*, *45*(4), 1651–1658. https://doi.org/10.1016/j.watres.2010.12.003

Pearce, M. M., Theodoropoulos, N., Mandel, M. J., Brown, E., Reed, K. D. & Cianciotto, N. P. (2012). *Legionella cardiaca* sp. nov., isolated from a case of native valve endocarditis in a human heart. *International Journal of Systematic and Evolutionary Microbiology*, *62*, 2946–2954. https://doi.org/10.1099/ijs.0.039248-0

Pearson, K. (1900). X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have

arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *50*(302), 157–175. https://doi.org/10.1080/14786440009463897

Pepper, I. L., Rusin, P., Quintanar, D. R., Haney, C., Josephson, K. L. & Gerba, C. P. (2004). Tracking the concentration of heterotrophic plate count bacteria from the source to the consumer's tap. *International Journal of Food Microbiology*, *92*(3), 289–295. https://doi.org/10.1016/j.ijfoodmicro.2003.08.021

Pereira, A., Silva, A. R. & Melo, L. F. (2021). *Legionella* and biofilms—integrated surveillance to bridge science and real-field demands. *Microorganisms*, *9*(6). https://doi.org/10.3390/microorganisms9061212

Pettersen, J. E. (2015a). *Forebygging av legionellasmitte – en veiledning 4. utgave* [Accessed: 10.03.2022]. Norwegian Institute of Public Health. Oslo, Norway.

Pettersen, J. E. (2015b). *Forebygging av legionellasmitte – en veiledning 4. utgave* (tech. rep.) [Accessed: 11.03.2022]. Norwegian Institute of Public Health. Oslo, Norway.

Piao, Z., Sze, C. C., Barysheva, O., Iida, K. & Yoshida, S. (2006). Temperature-regulated formation of mycelial mat-like biofilms by legionella pneumophila. *Applied and Environmental Microbiology*, *72*(2), 1613–1622. https://doi.org/10.1128/AEM.72.2.1613-1622.2006

Prest, E. I., Hammes, F., van Loosdrecht, M. C. M. & Vrouwenvelder, J. S. (2016). Biological Stability of Drinking Water: Controlling Factors, Methods, and Challenges. *Frontiers in Microbiology*, *7*. https://doi.org/10.3389/fmicb.2016.00045

Proctor, C. R., Dai, D., Edwards, M. A. & Pruden, A. (2017). Interactive effects of temperature, organic carbon, and pipe material on microbiota composition and *Legionella pneumophila* in hot water plumbing systems. *Microbiome*, *5*(130). https://doi.org/10.1186/s40168-017-0348-5

Proctor, C. R., Reimann, M., Vriens, B. & Hammes, F. (2018). Biofilms in shower hoses. *Water Research*, *131*, 274–286. https://doi.org/10.1016/j.watres.2017.12.027

Rabaey, K., Angenent, L., Schröder, U. & Keller, J. (2009). *Bioelectrochemical Systems: From Extracellular Electron Transfer to Biotechnological Application*. IWA Publishing. https://doi.org/10.2166/9781780401621

Rasheduzzaman, M., Singh, R., Haas, C. N. & Gurian, P. L. (2020). Required water temperature in hotel plumbing to control *Legionella* growth. *Water Research*, *182*, 115943. https://doi.org/10.1016/j.watres.2020.115943

Revetta, R. P., Gomez-Alvarez, V., Gerke, T. L., Curioso, C., Santo Domingo, J. W. & Ashbolt, N. J. (2013). Establishment and early succession of bacterial communities in monochloramine-treated drinking water biofilms. *FEMS Microbiology Ecology*, *86*(3), 404–414. https://doi.org/10.1111/1574-6941.12170

Rhoads, W. J., Pruden, A. & Edwards, M. A. (2016). Survey of green building water systems reveals elevated water age and water quality concerns. *Environmental Science: Water Research & Technology*, *2*(1), 164–173. https://doi.org/10.1039/C5EW00221D

Rossi, R., Murari, A., Gaudio, P. & Gelfusa, M. (2020). Upgrading Model Selection Criteria with Goodness of Fit Tests for Practical Applications. *Entropy (Basel, Switzerland)*, *22*(4). https://doi.org/10.3390/e22040447

Sacher, F., Gerstner, P., Merklinger, M., Thoma, A., Kinani, A., Roumiguières, A., Bouchonnet, S., Richard-Tanaka, B., Layousse, S., Ata, R., Marolleau, F. & Kinani, S. (2019). Determination of monochloramine dissipation kinetics in various surface water qualities under relevant environmental conditions - Consequences regarding environmental risk assessment. *Science of The Total Environment*, *685*, 542–554. https://doi.org/10.1016/j.scitotenv.2019.05.364

Schmid, T., Panne, U., Haisch, C., Hausner, M. & Niessner, R. (2002). A Photoacoustic Technique for Depth-Resolved In Situ Monitoring of Biofilms. *Environmental Science & Technology*, *36*(19), 4135–4141. https://doi.org/10.1021/es0158657

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Scotto, V., Dicintio, R. & Marcenaro, G. (1985). The influence of marine aerobic microbial film on stainless-steel. *CORROSION SCIENCE*, *25*(3), 185–194. https://doi.org/10.1016/0010-938X(85)90094-0

Sharaby, Y., Rodríguez-Martínez, S., Höfle, M. G., Brettar, I. & Halpern, M. (2019). Quantitative microbial risk assessment of *Legionella pneumophila* in a drinking water supply system in

Israel. *The Science of the total environment*, *671*, 404–410. https://doi.org/10.1016/j.scitotenv.2019.03.287

Shen, Y., Huang, C., Lin, J., Wu, W., Ashbolt, N. J., Liu, W. & Nguyen, T. H. (2017). Effect of Disinfectant Exposure on *Legionella pneumophila* Associated with Simulated Drinking Water Biofilms: Release, Inactivation, and Infectivity. *Environmental Science & Technology*, *51*(4), 2087–2095. https://doi.org/10.1021/acs.est.6b04754

Shen, Y., Monroy, G. L., Derlon, N., Janjaroen, D., Huang, C., Morgenroth, E., Boppart, S. A., Ashbolt, N. J., Liu, W. & Nguyen, T. H. (2015). Role of Biofilm Roughness and Hydrodynamic Conditions in *Legionella pneumophila* Adhesion to and Detachment from Simulated Drinking Water Biofilms. *Environmental Science & Technology*, *49*(7), 4274–4282. https://doi.org/10.1021/es505842v

Song, Y., Pruden, A., Edwards, M. A. & Rhoads, W. J. (2021). Natural organic matter, orthophosphate, ph, and growth phase can limit copper antimicrobial efficacy for *Legionella* in drinking water. *Environmental Science & Technology*, *55*(3), 1759–1768. https://doi.org/10.1021/acs.est.0c06804

Sudarsan, R., Ghosh, S., Stockie, J. M. & Eberl, H. J. (2016). Simulating Biofilm Deformation and Detachment with the Immersed Boundary Method. *Communications in Computational Physics*, *19*(3), 682–732. https://doi.org/10.4208/cicp.161214.021015a

Sydnor, E. R. M., Bova, G., Gimburg, A., Cosgrove, S. E., Perl, T. M. & Maragakis, L. L. (2012). Electronic-eye faucets: *Legionella* species contamination in healthcare settings. *Infection control and hospital epidemiology*, *33*(3), 235–240. https://doi.org/10.1086/664047

Taylor, M., Ross, K. & Bentham, R. (2009). *Legionella*, protozoa, and biofilms: interactions within complex microbial systems. *Microbial Ecology*, *58*(3), 538–547. https://doi.org/10.1007/s00248-009-9514-z

Thomas, J. M. & Ashbolt, N. J. (2011). Do Free-Living Amoebae in Treated Drinking Water Systems Present an Emerging Health Risk? *Environmental Science & Technology*, *43*(3), 860–869. https://doi.org/10.1021/es102876y

Trondheim kommune. (2017). Kommunedelplan for vannforsyning for 2017-2028. https://www.trondheim.kommune.no/globalassets/10-bilder-og-filer/10-byutvikling/kommunalteknikk/hovedplaner/pi15768-komplett.pdf

Tsai, Y. P. (2005). Impact of flow velocity on the dynamic behaviour of biofilm bacteria. *Biofouling*, *21*(5-6), 267–277. https://doi.org/10.1080/08927010500398633

United States Environmental Protection Agency (EPA). (n.d.). Premise plumbing decontamination [Last updated: 03/03/2022. Accessed: 31/05/2022]. https://www.epa.gov/emergency-response-research/premise-plumbing-decontamination

Vaccaro, L., Izquierdo, F., Magnet, A., Hurtado, C., Salinas, M., Gomes, T., Angulo, S., Salso, S., Pelaez, J., Tejeda, M. I., Alhambra, A., Gómez, C., Enríquez, A., Estirado, E., Fenoy, S. & del Aguila, C. (2016). First Case of Legionnaire's Disease Caused by *Legionella anisa* in Spain and the Limitations on the Diagnosis of *Legionella* non-*pneumophila* Infections. *PLOS ONE*, *11*(7), 1–12. https://doi.org/10.1371/journal.pone.0159726

van der Kooij, D. (1992). Assimilable Organic Carbon as an Indicator of Bacterial Regrowth. *Journal AWWA*, *84*(2), 57–65. https://doi.org/10.1002/j.1551-8833.1992.tb07305.x

Van Rossum, G. & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.

Vandecandelaere, I., Nercessian, O., Faimali, M., Segaert, E., Mollica, A., Achouak, W., Vos, P. D. & Vandamme, P. (2009). Bacterial diversity of the cultivable fraction of a marine electroactive biofilm. *Bioelectrochemistry*, *78*(1), 62–66. https://doi.org/10.1016/j.bioelechem.2009.07.004

Vargas, I. T., Fischer, D. A., Alsina, M. A., Pavissich, J. P., Pastén, P. A. & Pizarro, G. E. (2017). Copper Corrosion and Biocorrosion Events in Premise Plumbing. *Materials*, *10*(10). https://doi.org/10.3390/ma10091036

Voulvoulis, N. (2018). Water reuse from a circular economy perspective and potential risks from an unregulated approach. *Current Opinion in Environmental Science & Health*, *2*, 32–45. https://doi.org/10.1016/j.coesh.2018.01.005

Vu, B., Chen, M., Crawford, R. J. & Ivanova, E. P. (2009). Bacterial Extracellular Polysaccharides Involved in Biofilm Formation. *Molecules*, *14*(7), 2535–2554.

Waak, M. B., LaPara, T. M., Hallé, C. & Hozalski, R. M. (2018). Occurrence of *Legionella spp.* in Water-Main Biofilms from Two Drinking Water Distribution Systems [PMID: 29902377].

Environmental Science & Technology, *52*(14), 7630–7639. https://doi.org/10.1021/acs.est.8b01170

Wagner, M., Manz, B., Volke, F., Neu, T. R. & Horn, H. (2010). Online assessment of biofilm development, sloughing and forced detachment in tube reactor by means of magnetic resonance microscopy. *Biotechnology and Bioengineering*, *107*(1), 172–181. https://doi.org/10.1002/bit.22784

Walker, J. T. (2018). The influence of climate change on waterborne disease and *Legionella*: a review. *Perspective Public Health*, *138*(5). https://doi.org/10.1177/1757913918791198

Walker, J. T., Dowsett, A. B., Dennis, P. J. L. & Keevil, C. W. (1991). Continuous culture studies of biofilm associated with copper corrosion [Special Issue Extracellular Microbial Products]. *International Biodeterioration*, *27*(2), 121–134. https://doi.org/10.1016/0265-3036(91)90004-B

Walker, J. T. & McDermott, P. J. (2021). Confirming the Presence of *Legionella pneumophila* in Your Water System: A Review of Current *Legionella* Testing Methods. *Journal of AOAC INTERNATIONAL*, *104*(4), 1135–1147. https://doi.org/10.1093/jaoacint/qsab003

Wang, W. & Lu, Y. (2018). Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model. *IOP Conference Series: Materials Science and Engineering*, *324*, 012049. https://doi.org/10.1088/1757-899x/324/1/012049

Wang, Z., Li, L., Ariss, R. W., Coburn, K. M., Behbahani, M., Xue, Z. & Seo, Y. (2021). The role of biofilms on the formation and decay of disinfection by-products in chlor(am)inated water distribution systems. *Science of The Total Environment*, *753*, 141606. https://doi.org/10.1016/j.scitotenv.2020.141606

Whiley, H., Keegan, A., Fallowfield, H. & Bentham, R. (2015). The presence of opportunistic pathogens, *Legionella spp.*, *L. pneumophila* and mycobacterium avium complex, in south australian reuse water distribution pipelines. *Journal of water and health*, *13*(2). https://doi.org/10.2166/wh.2014.317

Wickham, H. (2014). Tidy Data. *Journal of Statistical Software*, *59*(10), 1–23. https://doi.org/10.18637/jss.v059.i10

Williams, K., Pruden, A., Falkinham, J. O. & Edwards, M. (2015). Relationship between Organic Carbon and Opportunistic Pathogens in Simulated Glass Water Heaters. *Pathogens*, *4*(2), 355–372. https://doi.org/10.3390/pathogens4020355

Wingender, J. & Flemming, H. .-. (2011). Biofilms in drinking water and their role as reservoir for pathogens. *INTERNATIONAL JOURNAL OF HYGIENE AND ENVIRONMENTAL HEALTH*, *214*(6), 417–423. https://doi.org/10.1016/j.ijheh.2011.05.009

Wu, H. Y., Yan, H., Zheng, M. L., Sun, M. M., Wang, Q., Hu, C. M., Zhan, X. Y., Yuan, M. G., P. H, Q. & Hu, C. H. (2019). *Legionella qingyii* sp. nov., isolated from water samples in china. *International Journal of Systematic and Evolutionary Microbiology*, *69*(7), 2017–2022. https://doi.org/10.1099/ijsem.0.003421

Xu, J., Huang, C., Shi, X., Dong, S., Yuan, B. & Nguyen, T. H. (2018). Role of drinking water biofilms on residual chlorine decay and trihalomethane formation: An experimental and modeling study. *Science of The Total Environment*, *642*, 516–525. https://doi.org/10.1016/j.scitotenv.2018.05.363

Zhang, C., Struewing, I., Mistry, J. H., Wahman, D. G., Pressman, J. & Lu, J. (2021). *Legionella* and other opportunistic pathogens in full-scale chloraminated municipal drinking water distribution systems. *Water Research*, *205*, 117571. https://doi.org/10.1016/j.watres.2021.117571

Zhang, Y. & Edwards, M. (2009). Accelerated chloramine decay and microbial growth by nitrification in premise plumbing. *Journal - American Water Works Association*, *101*, 51–+. https://doi.org/10.1002/j.1551-8833.2009.tb09990.x

Zlatanovic, L., Moerman, A., van der Hoek, J. P., Vreeburg, J. & Blokker, M. (2017). Development and validation of a drinking water temperature model in domestic drinking water supply systems. *Urban Water Journal*, *14*(10), 1031–1037. https://doi.org/10.1080/1573062X.2017.1325501

ZYMO RESEARCH. (n.d.). ZymoBIOMICS DNA Miniprep Kit [Date accessed: 07/06/2022]. https://www.zymoresearch.com/products/zymobiomics-dna-miniprep-kit

Appendix

A Drinking Water Quality for Event 1.1

The drinking water quality parameters shown in the "Water post-treatment" section of Table 2, characterise the outlet water from Vikelvdalen water treatment plant in Trondheim. The "Building water" section outlines the drinking water quality parameters measured at the building inlet at Valgrinda to the premise plumbing pilot in this study. BDOC stands for biodegradable dissolved organic carbon, and TOC for total organic carbon.

Table 2: Municipal Drinking Water Quality

Parameter	Unit	Median	Range	Ref.
Water post-treatment				
Temperature	°C	4.1	1.0 - 5.5	*
pH		8.1	7.8 - 8.6	*
Hardness	$\frac{mg}{L}\ CaCO_3$	54	53 - 55	*
Alkalinity	$\frac{mg}{L}\ HCO_3^-$	62		*
Conductivity	$\frac{\mu S}{cm^2}$	128	113 - 137	*
Turbidity	FNU	0.1	0.1 - 3.6	*
TOC	$\frac{mg}{L}\ C$	3.0	2.0 - 4.0	*
BDOC	$\frac{mg}{L}\ C$	0.5		Charnock et al., 2000
$AOC_P - 17/NOX$	$\frac{\mu g}{L}\ C$	118	72 - 268	Waak et al., 2018, Johansen, 2018
Ammonium	$\frac{mg}{L}\ N$	0.005		*
Nitrate	$\frac{mg}{L}\ N$	0.24	0.23 - 0.24	*
Nitrite	$\frac{mg}{L}\ N$	0.003		*
Phosphorus, total	$\frac{mg}{L}\ P$	< 0.001		Waak et al., 2018
Chlorine, total	$\frac{mg}{L}\ Cl$	0.07	0.05 - 0.12	*
Copper	$\frac{\mu g}{L}\ Cu$	45	3 - 87	*
Building water				
Temperature	°C	10.0	7.6 - 12.5	Simultaneous pilot study
pH		8.0	7.8 - 8.3	Simultaneous pilot study
Conductivity	$\frac{\mu S}{cm^2}$	129	124 - 139	Simultaneous pilot study
Turbidity	FNU	0.6	0.2 - 4.4	Simultaneous pilot study
TOC	$\frac{mg}{L}\ C$	2.9	2.7 - 3.7	Simultaneous pilot study
AOC	$\frac{\mu g}{L}\ C$	96	28 - 244	Waak et al., 2018, Johansen, 2018
Chlorine, total	$\frac{mg}{L}\ Cl$	< 0.02		Waak et al., 2018, Johansen, 2018
Copper	$\frac{\mu g}{L}\ Cu$	180	48 - 1270	Simultaneous pilot study

*parameters were taken from the drinking water quality parameters of Trondheim municipality (Nøst, 2020). Hardness was estimated from the Ca^{2+} and Mg^{2+} data available.

Due to the consistency and standardisation of pilot operations, it is reasonable to assume that the flow rates and temperatures shown in Table 3 are an accurate representation of the conditions in Phase 2, Phase 3, and Phase 4.

Table 3: Empirical Flow Rates and Temperatures From the Pilot During Event 1.1

Rig	Outlet	Supply	Operation	Pipe material	Temp., °C	Total vol., L	Flow rate, $\frac{L}{min}$
1	A	Cold	Stagnant	Copper	21.3	< 1	0.0
		Hot	Stagnant	Copper	21.3	< 1	0.0
	B	Blend	Shower	Copper	40.8	90	11.3
		Cold	Shower	Copper	10.6	30	3.8
		Hot	Shower	Copper	57.4	60	7.5
	C	Cold	Flush	Copper	12.5	83	16.6
		Hot	Flush	Copper	57.4	80	16.0
2	A	Cold	Stagnant	PEX-A	22.1	< 1	0.0
		Hot	Stagnant	PEX-A	22.1	< 1	0.0
	B	Blend	Shower	PEX-A	40.3	73	9.1
		Cold	Shower	PEX-A	10.6	26	3.3
		Hot	Shower	PEX-A	58.7	46	5.8
	C	Cold	Flush	PEX-A	8.7	88	17.5
		Hot	Flush	PEX-A	58.7	79	15.8

The total shower volumes of the hot and cold lines in Table 3 were calculated through thermal mass balance.

B Processed ALVIM Sensor Data from Phase 1, Phase 3, and Phase 4

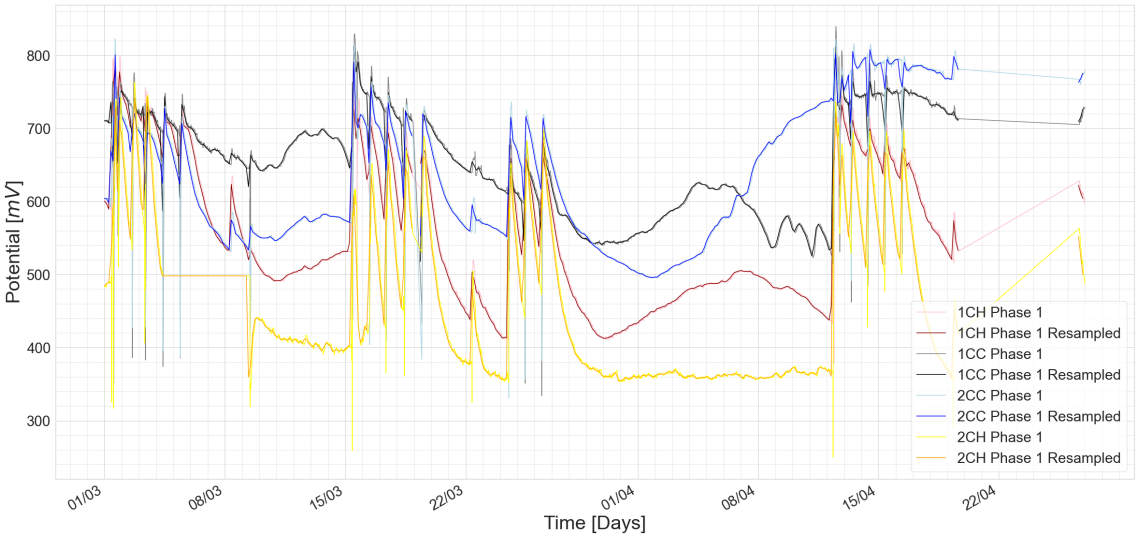


Figure 12: Phase 1 processed ALVIM sensor data.

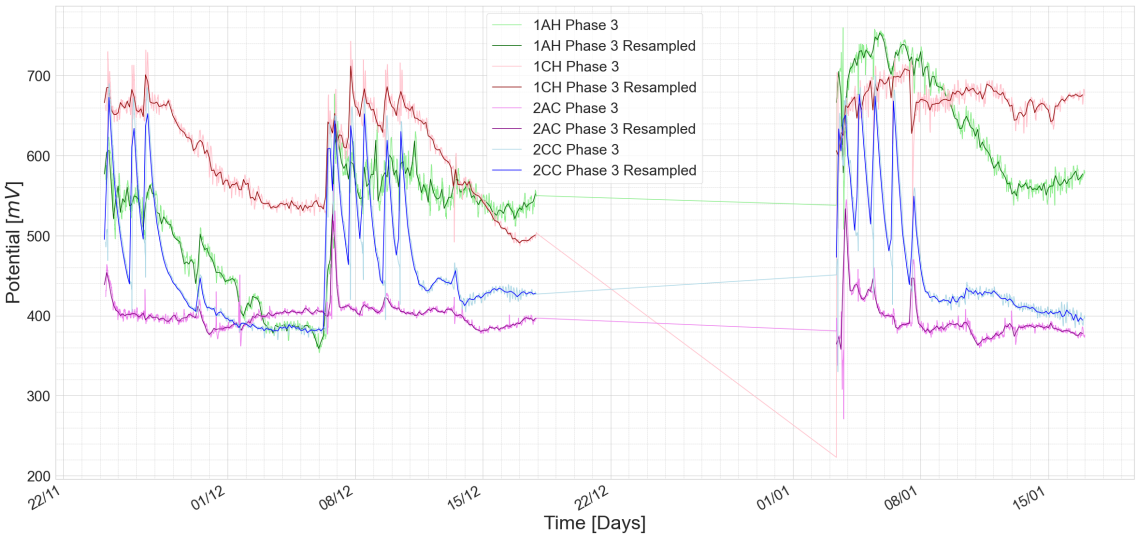


Figure 13: Phase 3 processed ALVIM sensor data.

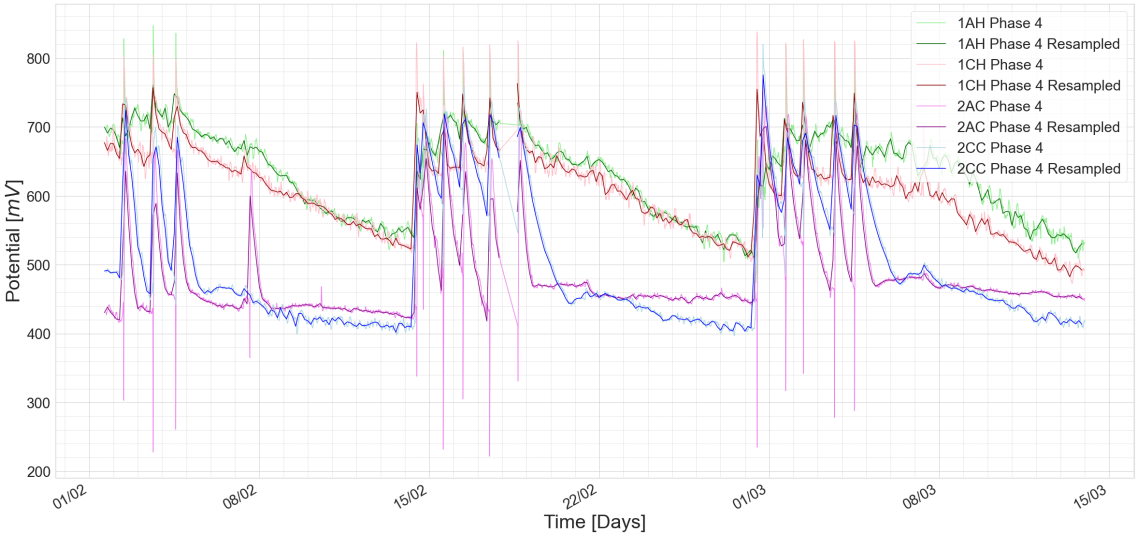


Figure 14: Phase 4 processed ALVIM sensor data.

C Hydraulic Impacts of Daily Treatments on ALVIM Sensors in Stagnant Lines

The ALVIM sensors deployed in stagnant lines proved sensitive to the hydraulic actions in other lines, demonstrated for Phase 2 in Figure 15. The implications this had for biofilm growth in the impacted lines is unclear. The same trend was observed during Phase 3, in which showers were conducted simultaneously as stagnation. This behaviour was only observed in the stagnant lines, and thus not exhibited in Phase 1 and Phase 4.

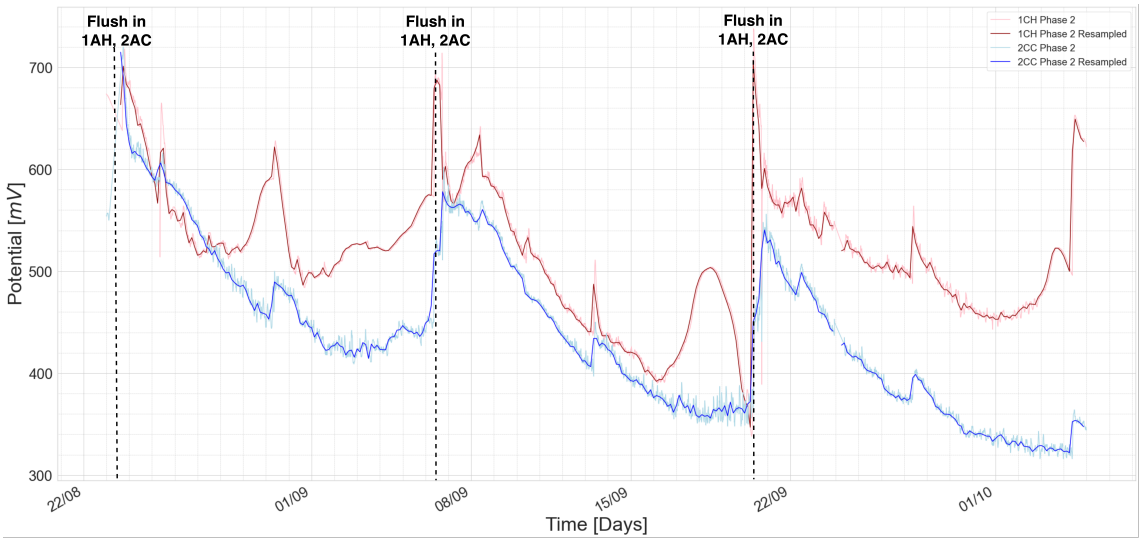


Figure 15: Hydrodynamic impacts of daily treatments (flush and shower) in lines 1AH, 2AC on the ALVIM sensor data of the stagnant lines 1CH, 2CC during Phase 2.

D Biofilm Biomass Sampling and Subsequent DNA Extraction and Analysis

Biomass samples of biofilm were collected for DNA analysis using scraping according to an adjusted CDC Protocol. Specifically, in Phase 1, sterilised micro-spatulae were used to scrape the line (distal pipe) interior gently in 30 circular motions clockwise, and 30 counter-clockwise, resulting in a total of 60 rotations. In Phase 2, a POLYWIPE™sponge (MWE, n.d.) was cut into approximately 4 cm x 2 cm pieces. A sponge piece was attached to a custom-made flexible steel loop, which was flame-sterilised. The loop was then inserted into a drained line (distal pipe) carefully, without touching the pipe exterior, before an industrial drill was used to spin it 30 seconds clockwise and 30 seconds counterclockwise.

The biofilm samples were collected upon stabilisation in BES, approximately 5 ± 1 hours after each daily treatment (i.e. not during stagnation periods), as indicated by the ALVIM sensors (Pavanello et al., 2011). From Figure 16, it was observed that the impact on BES-data was short-lived, with the ALVIM sensor signals recovering after a period of roughly 2 hours surrounding the collection procedure. Scraping did not have a noticeable impact on biofilms' ability for (re)growth.

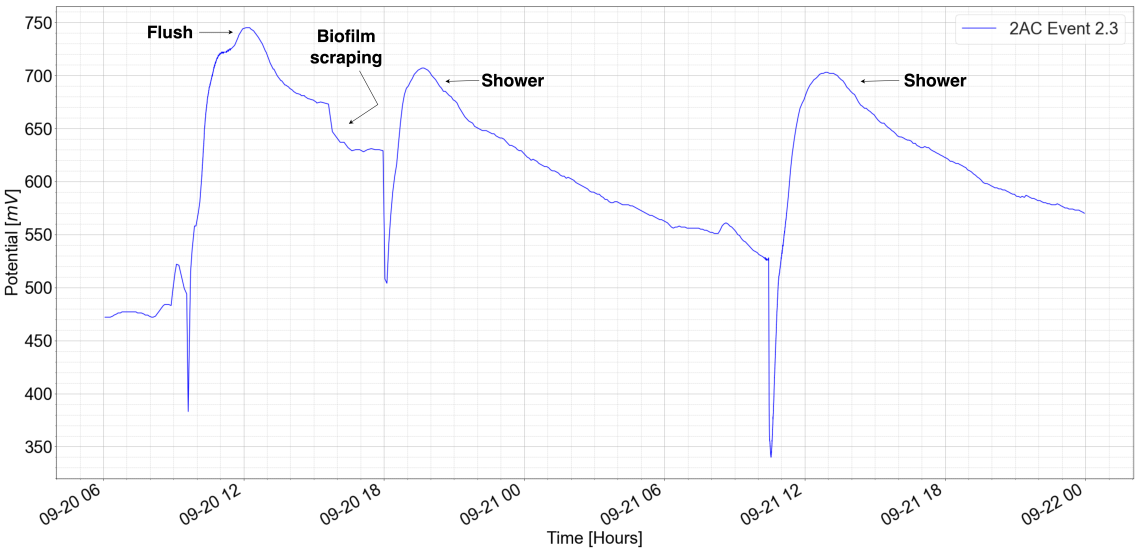


Figure 16: Impact of biofilm biomass collection on ALVIM sensor data for Event 2.3 in line 2AC.

The time period of 5 ± 1 hours between flushing and collection of biofilm biomass may have resulted in recovery of bacterial growth. However, this is not believed to be of significance due to the following reasons: (i) slow-growing *Legionella spp.* were not affected, (ii) various control measures required sufficient time to reveal possible effects/benefits, and (iii) as this was done for all, the results represented the same day-effects/-benefits of flushing, if any.

DNA-extraction was executed through the use of ZymoBIOMICS DNA Miniprep Kit (ZYMO RESEARCH, n.d.), with subsequent analysis using reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR) to test for *Legionella*, *L. pneumophila*, *Lp*1 as well as *Vermamoeba* and *Acanthamoeba*. The gene concentrations targeted and their specifications are detailed in Table 4.

Table 4: Gene Concentrations Measured and Their Corresponding Specifications

Gene concentration measured	**Specification**
16S rRNA	All bacteria (domain)
18S rRNA	*Acanthamaoeba* (genus)
	Vermamoeba (genus)
ssrA	*Legionella spp.* (genus)
mip	*L. pneumophila* (species)
wzm	*L. pneumophila sg1* (serogroup)

E Biofilm Growth Curves: Goodness of Fit Measures

Table 5: Polynomial Curve Fit on Biofilm Growth Curves: Goodness of Fit Results and Corresponding Standard Deviations

Goodness of fit	**Linear**	**Polynomial**	**Exponential**
$\boldsymbol{\chi_r^2}$			
Mean ± SD	38.9 ± 40.1	14.0 ± 10.2	26.2 ± 35.8
Max	153.9	30.2	152.8
Min	2.0	1.9	2.0
AIC			
Mean ± SD	127.0 ± 88.2	94.7 ± 60.8	111.2 ± 88.3
Max	404.9	276.4	405.3
Min	28.8	29.0	29.7
BIC			
Mean ± SD	130.1 ± 88.8	101.0 ± 61.9	116.0 ± 89.2
Max	409.6	286.0	412.4
Min	32.2	34.7	34.1
Mean relative error [%] ± SD			
a	9.2 ± 6.7	117.8 ± 270.9	8401.6 ± 35403.2
b	0.4 ± 0.2	89.2 ± 126.7	8389.6 ± 35392.1
c		89.6 ± 125.2	8484.6 ± 35909.0
d		0.4 ± 0.2	

F Curve Fitting ALVIM Sensor Data From the Stagnation Periods of Shower Events Using the Exponential Function

The data sets from the stagnation periods of shower events each consisted of 56 data points. Nearly all were curve fitted using initial coefficient values of 0.0001. However, adjustments were made for those who did not reach convergence with these, e.g. 2CC 3.3, whose initial values were: $a = 100$, $b = 0.0001$, and $c = 10000$.

Table 6: Goodness of Fit Measures and Coefficient Values Resulting From Curve Fitting the Exponential Function on the Stagnation Period Shower Event Data

Event	**AIC**	**BIC**	$\boldsymbol{\chi^2_r}$	**a**	**b**	**c**
1AH 4.1	239	245	67.8	-173	-0.00388	879
1AH 4.2	246	252	76.4	-86.3	-0.00578	767
1AH 4.3	273	279	124	-14.4	-0.0142	686
1CH 3.1	259	265	97.4	-545000	-0.00000160	545000
1CH 3.2	261	267	101	-833000	-0.00000129	833000
1CH 3.3	257	263	93.2	-0.311	-0.0251	675
2AC 4.1	374	380	750	47600	0.00000317	-47200
2AC 4.2	224	230	52.0	40.1	0.0196	449
2AC 4.3	215	221	43.9	77.4	0.00300	410
2CC 3.1	254	260	88.1	133	0.0304	387
2CC 3.2	242	248	71.3	59.2	0.118	430
2CC 3.3	238	244	66.7	38.4	0.177	419

Table 7: Relative Errors for Coefficients Presented in Table 6

Event	**Relative error a [%]**	**Relative error b [%]**	**Relative error c [%]**
1AH 4.1	39.2	29.0	7.99
1AH 4.2	36.6	23.7	4.43
1AH 4.3	36.1	14	1.18
1CH 3.1	17500	17600	17600
1CH 3.2	9760	9750	9750
1CH 3.3	262	64.3	0.450
2AC 4.1	903000	934000	912000
2AC 4.2	9.09	26.6	0.700
2AC 4.3	119	154	23.0
2CC 3.1	3.95	8.72	0.620
2CC 3.2	12.5	22.0	0.290
2CC 3.3	20.1	37.5	0.280

G Python Scripts

```
#!/usr/bin/env python
# coding: utf-8

# In[1]:


import numpy as np
import pandas as pd
from datetime import datetime, timedelta
import matplotlib.pyplot as plt
import seaborn as sns
import sys
import os
from pathlib import Path
import pickle
import matplotlib


# # Reading in Raw Data from the ALVIM Sensors
# **Note that code cannot be replicated without the input files and the pilot operations file.**

# In[2]:


first_config = ({'ModBus_1': '1C Hot', 'ModBus_2': '1C Cold',
'ModBus_3': '2C Hot', 'ModBus_4': '2C Cold'})
second_config = ({'ModBus_1.0': '1C Hot', 'ModBus_2.0': '1A Hot',
'ModBus_3.0': '2A Cold', 'ModBus_4.0': '2C Cold'})
second_config2 = ({'ModBus_1': '1C Hot', 'ModBus_2': '1A Hot',
'ModBus_3': '2A Cold', 'ModBus_4': '2C Cold'})
sheets = [0, 1, 2, 3]


# In[3]:


def read_excel_files(filename, configuration = first_config):
    data = []
    for s in sheets:
        d = pd.read_excel(filename, index_col=1, header=0, sheet_name = s, parse_dates = [0])
        ModBus = d['ModBus'].iloc[1]
        ModBus = f'ModBus_{ModBus}'
        d.drop("ModBus", inplace = True, axis = 1)
        d.drop("Sensivity", inplace = True, axis = 1)
        d = d.squeeze()
        d.name = ModBus
        data.append(d)
    data_all = pd.concat(data, axis=1)
    data_all.index = pd.to_datetime(data_all.index, format='%Y%m%d%H%M%S')
    data_all = data_all.sort_index()
    data_all = data_all.rename(columns = configuration)
    return data_all


# In[4]:


data_all_1 = read_excel_files('ALVIM_Data/R_ALVIM 2021.02.01-2021.04.07.xls')
data_all_2 = read_excel_files('ALVIM_Data/R_ALVIM 2021.02.03-2021.02.08.xls')
data_all_3 = read_excel_files('ALVIM_Data/R_ALVIM 2021.02.03-2021.04.19 (2).xls')
data_all_4 = read_excel_files('ALVIM_Data/R_ALVIM 2021.02.26-2021.03.03.xls')
```

```
data_all_5 = read_excel_files('ALVIM_Data/R_ALVIM 2021.02.28-2021.03.16.xls')
data_all_6 = read_excel_files('ALVIM_Data/R_ALVIM 2021.05.01-2021.06.01.xls')
data_all_7 = read_excel_files('ALVIM_Data/R_ALVIM 2021.07.27-2021.08.02.xls')
data_all_8 = read_excel_files('ALVIM_Data/R_ALVIM 2021.08.02-2021.08.16.xls')
data_all_9 = (read_excel_files('ALVIM_Data/R_ALVIM 2021.08.16-2021.09.03.xls',
configuration = second_config2))
data_all_10 = (read_excel_files('ALVIM_Data/R_ALVIM 2021.08.30-2021.10.01.xlsx',
configuration = second_config))
data_all_11 = (read_excel_files('ALVIM_Data/R_ALVIM 2021.10.01-2022.01.17.xlsx',
configuration = second_config2))
data_all_12 = read_excel_files('ALVIM_Data/R_ALVIM 2021.04.26-2021.05.28.xls')
data_all_13 = (read_excel_files('ALVIM_Data/R_ALVIM 2021.12.01-2022.03.25A.xls',
configuration = second_config2))


# ## Organising Data by Lines
# The input data was organised by time period of sampling and sensor.

# In[5]:


R_1AH = [data_all_9['1A Hot'], data_all_10['1A Hot'], data_all_11['1A Hot'], data_all_13['1A Hot'] ]
R_2AC = [data_all_9['2A Cold'], data_all_10['2A Cold'], data_all_11['2A Cold'], data_all_13['2A Cold'] ]
R_1CH = ([data_all_1['1C Hot'], data_all_2['1C Hot'], data_all_3['1C Hot'], data_all_4['1C Hot'],
data_all_5['1C Hot'], data_all_6['1C Hot'], data_all_7['1C Hot'], data_all_8['1C Hot'],
data_all_9['1C Hot'], data_all_10['1C Hot'], data_all_11['1C Hot'],
data_all_12['1C Hot'], data_all_13['1C Hot'] ])
R_1CC = ([data_all_1['1C Cold'], data_all_2['1C Cold'], data_all_3['1C Cold'], data_all_4['1C Cold'],
data_all_5['1C Cold'], data_all_6['1C Cold'],
data_all_7['1C Cold'], data_all_8['1C Cold'], data_all_12['1C Cold'] ])
R_2CC = ([data_all_1['2C Cold'], data_all_2['2C Cold'], data_all_3['2C Cold'], data_all_4['2C Cold'],
data_all_5['2C Cold'], data_all_6['2C Cold'], data_all_7['2C Cold'], data_all_8['2C Cold'],
data_all_9['2C Cold'], data_all_10['2C Cold'], data_all_11['2C Cold'],
data_all_12['2C Cold'], data_all_13['2C Cold'] ])
R_2CH = ([data_all_1['2C Hot'], data_all_2['2C Hot'], data_all_3['2C Hot'], data_all_4['2C Hot'],
data_all_5['2C Hot'], data_all_6['2C Hot'], data_all_7['2C Hot'],
data_all_8['2C Hot'], data_all_12['2C Hot']])


# In[24]:


d1AH = pd.concat(R_1AH, axis=0)
d2AC = pd.concat(R_2AC, axis=0)
d1CH = pd.concat(R_1CH, axis=0)
d1CC = pd.concat(R_1CC, axis=0)
d2CC = pd.concat(R_2CC, axis=0)
d2CH = pd.concat(R_2CH, axis=0)


# ## Dataframes for Each Line
# First, input data was sorted,
#then metadata from the pilot operations file was extracted and related to relevant sections.
# These were later used to identify stagnation periods etc.

# In[25]:


def simple_df(df):
    df_simple = pd.DataFrame(df)
    name = df_simple.columns[0]
    df_simple['Rig'] = name[0]
    df_simple['Line'] = name[1]
```

```
    df_simple['Supply'] = name.split(' ')[1]
    df_simple = df_simple.rename(columns = {name: 'Potential'})
    df_simpled = df_simple.dropna()
    df_simple = df_simple.sort_index()
    return df_simple


# In[28]:


df_1AH = simple_df(d1AH)
df_2AC = simple_df(d2AC)
df_1CH = simple_df(d1CH)
df_1CC = simple_df(d1CC)
df_2CC = simple_df(d2CC)
df_2CH = simple_df(d2CH)


# In[29]:


#Dataframes containing metadata
file_po = 'ALVIM_Data/Pilot Operation .xlsx'

line_1A_Hot = pd.read_excel(file_po, index_col=0, header=0, sheet_name = 1)
line_1C_Hot = pd.read_excel(file_po, index_col=0, header=0, sheet_name = 9)
line_1C_Cold = pd.read_excel(file_po, index_col=0, header=0, sheet_name = 10)
line_2A_Cold = pd.read_excel(file_po, index_col=0, header=0, sheet_name = 4)
line_2C_Cold = pd.read_excel(file_po, index_col=0, header=0, sheet_name = 11)
line_2C_Hot = pd.read_excel(file_po, index_col=0, header=0, sheet_name = 12)

copy_list = ['Weekday', 'Phase', 'Event', 'Treatment', 'Daily.Treatment', 'Disinfection']


# In[30]:


def metadata_df(df, pilot_op, copy_list = copy_list):
    for row_index, row in pilot_op.iterrows():
        date = row_index.date()
        for c in copy_list:
            for c in copy_list:
                df.loc[str(date), c] = row[c]
    df = df.sort_index()
    return df


# In[31]:


dfm_1AH = metadata_df(df_1AH,line_1A_Hot)
dfm_1CH = metadata_df(df_1CH,line_1C_Hot)
dfm_1CC = metadata_df(df_1CC,line_1C_Cold)
dfm_2AC = metadata_df(df_2AC,line_2A_Cold)
dfm_2CC = metadata_df(df_2CC,line_2C_Cold)
dfm_2CH = metadata_df(df_2CH,line_2C_Hot)


# In[34]:


#To avoid many rows with NaN data
dfm_1AH = dfm_1AH.dropna()
```

```
dfm_1CH = dfm_1CH.dropna()
dfm_1CC = dfm_1CC.dropna()
dfm_2AC = dfm_2AC.dropna()
dfm_2CC = dfm_2CC.dropna()
dfm_2CH = dfm_2CH.dropna()


# ### Saving the Metadata-Dataframes
#Uncomment lines to run
# In[ ]:


#dfm_1AH.to_csv('ALVIM_CSV_TS_Meta/dfm_1AH.csv')
#dfm_1CH.to_csv('ALVIM_CSV_TS_Meta/dfm_1CH.csv')
#dfm_1CC.to_csv('ALVIM_CSV_TS_Meta/dfm_1CC.csv')
#dfm_2AC.to_csv('ALVIM_CSV_TS_Meta/dfm_2AC.csv')
#dfm_2CC.to_csv('ALVIM_CSV_TS_Meta/dfm_2CC.csv')
#dfm_2CH.to_csv('ALVIM_CSV_TS_Meta/dfm_2CH.csv')


# In[2]:


#Note that 1CC, 2CH used 1.0 not 1 for Phase, Events etc.
#This is relevant when extracting specific columns.

dfm_1AH = pd.read_csv('ALVIM_CSV_TS_Meta/dfm_1AH.csv', index_col= 0, header = 0, parse_dates = [0])
dfm_1CH = pd.read_csv('ALVIM_CSV_TS_Meta/dfm_1CH.csv', index_col= 0, header = 0, parse_dates = [0])
dfm_1CC = pd.read_csv('ALVIM_CSV_TS_Meta/dfm_1CC.csv', index_col= 0, header = 0, parse_dates = [0])
dfm_2AC = pd.read_csv('ALVIM_CSV_TS_Meta/dfm_2AC.csv', index_col= 0, header = 0, parse_dates = [0])
dfm_2CC = pd.read_csv('ALVIM_CSV_TS_Meta/dfm_2CC.csv', index_col= 0, header = 0, parse_dates = [0])
dfm_2CH = pd.read_csv('ALVIM_CSV_TS_Meta/dfm_2CH.csv', index_col=0, header = 0, parse_dates = [0])


# ### Zero-Indexing and Organising Data into Specific Phases, Lines, Events

# #### Functions used:

# In[3]:


def start_exact(df):
    s = df.first_valid_index()
    st = str(s)
    return st


# In[4]:


def zero_index(df):
    start_time = start_exact(df)
    start_time = pd.to_datetime(start_time)
    dfc = df.copy()
    dfc.index = dfc.index-start_time
    dfc = dfc.sort_index()
    return dfc


# In[5]:


def concat_phases(e1, e2, e3):
```

```
    p = pd.concat([e1, e2, e3])
    p = p.sort_index()
    return p


# In[6]:


#Used for resampling events for each line to a specific frequency (3 hours).
def resample_phases(df, re_window):
    re = df['Potential'].resample(re_window).mean()
    re = re.sort_index()
    return re


# In[7]:


#Use resample_phases as well as extracting stagnation periods.
def stag_re(df, re_window):
    dfc = df.query('`Daily.Treatment` == "Stagnation"')
    dfc = resample_phases(dfc, re_window)
    return dfc


# In[8]:


#Same as stag_re, but in a few cases,
#stagnation was labelled "Stagnant" instead of "Stagnation".
def stag_re_2(df, re_window):
    dfc = df.query('`Daily.Treatment` == "Stagnant"')
    dfc = resample_phases(dfc, re_window)
    return dfc


# #### Register of start and end times of each Phase, Event.

# In[9]:


#Dates for phases, events:
start_1_1 = '2021-03-01 00:00:00'
end_1_1 = '2021-03-14 23:59:59'
start_1_2 = '2021-03-15 00:00:00'
end_1_2 = '2021-04-11 23:59:59'
start_1_3 = '2021-04-12 00:00:00'
end_1_3 = '2021-04-26 23:59:59'

start_2_1 = '2021-08-23 00:00:00'
end_2_1 = '2021-09-05 23:59:59'
start_2_2 = '2021-09-06 00:00:00'
end_2_2 = '2021-09-19 23:59:59'
start_2_3 = '2021-09-20 00:00:00'
end_2_3 = '2021-10-04 23:59:59'

start_3_1 = '2021-11-22 00:00:00'
end_3_1 = '2021-12-05 23:59:59'
start_3_2 = '2021-12-06 00:00:00'
end_3_2 = '2022-01-02 23:59:59'
start_3_3 = '2022-01-03 00:00:00'
end_3_3 = '2022-01-16 23:59:59'
```

```
start_4_1 = '2022-01-31 00:00:00'
end_4_1 = '2022-02-13 23:59:59'
start_4_2 = '2022-02-14 00:00:00'
end_4_2 = '2022-02-27 23:59:59'
start_4_3 = '2022-02-28 00:00:00'
end_4_3 = '2022-03-13 23:59:59'


# In[10]:


#1AH
dfm_1AH_2_1 = dfm_1AH[start_2_1: end_2_1]
z_1AH_2_1 = zero_index(dfm_1AH_2_1)
dfm_1AH_2_2 = dfm_1AH[start_2_2: end_2_2]
z_1AH_2_2 = zero_index(dfm_1AH_2_2)
dfm_1AH_2_3 = dfm_1AH[start_2_3: end_2_3]
z_1AH_2_3 = zero_index(dfm_1AH_2_3)
dfm_1AH_2 = concat_phases(dfm_1AH_2_1, dfm_1AH_2_2, dfm_1AH_2_3)

dfm_1AH_3_1 = dfm_1AH[start_3_1: end_3_1]
z_1AH_3_1 = zero_index(dfm_1AH_3_1)
dfm_1AH_3_2 = dfm_1AH[start_3_2: end_3_2]
z_1AH_3_2 = zero_index(dfm_1AH_3_2)
dfm_1AH_3_3 = dfm_1AH[start_3_3: end_3_3]
z_1AH_3_3 = zero_index(dfm_1AH_3_3)
dfm_1AH_3 = concat_phases(dfm_1AH_3_1, dfm_1AH_3_2, dfm_1AH_3_3)

dfm_1AH_4_1 = dfm_1AH[start_4_1: end_4_1]
z_1AH_4_1 = zero_index(dfm_1AH_4_1)
dfm_1AH_4_2 = dfm_1AH[start_4_2: end_4_2]
z_1AH_4_2 = zero_index(dfm_1AH_4_2)
dfm_1AH_4_3 = dfm_1AH[start_4_3: end_4_3]
z_1AH_4_3 = zero_index(dfm_1AH_4_3)
dfm_1AH_4 = concat_phases(dfm_1AH_4_1, dfm_1AH_4_2, dfm_1AH_4_3)


# In[11]:


#1CH
dfm_1CH_1_1 = dfm_1CH[start_1_1: end_1_1]
z_1CH_1_1 = zero_index(dfm_1CH_1_1)
dfm_1CH_1_2 = dfm_1CH[start_1_2: end_1_2]
z_1CH_1_2 = zero_index(dfm_1CH_1_2)
dfm_1CH_1_3 = dfm_1CH[start_1_3: end_1_3]
z_1CH_1_3 = zero_index(dfm_1CH_1_3)
dfm_1CH_1 = concat_phases(dfm_1CH_1_1, dfm_1CH_1_2, dfm_1CH_1_3)

dfm_1CH_2_1 = dfm_1CH[start_2_1: end_2_1]
z_1CH_2_1 = zero_index(dfm_1CH_2_1)
dfm_1CH_2_2 = dfm_1CH[start_2_2: end_2_2]
z_1CH_2_2 = zero_index(dfm_1CH_2_2)
dfm_1CH_2_3 = dfm_1CH[start_2_3: end_2_3]
z_1CH_2_3 = zero_index(dfm_1CH_2_3)
dfm_1CH_2 = concat_phases(dfm_1CH_2_1, dfm_1CH_2_2, dfm_1CH_2_3)

dfm_1CH_3_1 = dfm_1CH[start_3_1: end_3_1]
z_1CH_3_1 = zero_index(dfm_1CH_3_1)
dfm_1CH_3_2 = dfm_1CH[start_3_2: end_3_2]
z_1CH_3_2 = zero_index(dfm_1CH_3_2)
dfm_1CH_3_3 = dfm_1CH[start_3_3: end_3_3]
z_1CH_3_3 = zero_index(dfm_1CH_3_3)
```

```
dfm_1CH_3 = concat_phases(dfm_1CH_3_1, dfm_1CH_3_2, dfm_1CH_3_3)

dfm_1CH_4_1 = dfm_1CH[start_4_1: end_4_1]
z_1CH_4_1 = zero_index(dfm_1CH_4_1)
dfm_1CH_4_2 = dfm_1CH[start_4_2: end_4_2]
z_1CH_4_2 = zero_index(dfm_1CH_4_2)
dfm_1CH_4_3 = dfm_1CH[start_4_3: end_4_3]
z_1CH_4_3 = zero_index(dfm_1CH_4_3)
dfm_1CH_4 = concat_phases(dfm_1CH_4_1, dfm_1CH_4_2, dfm_1CH_4_3)


# In[12]:


#1CC
dfm_1CC_1_1 = dfm_1CC[start_1_1: end_1_1]
z_1CC_1_1 = zero_index(dfm_1CC_1_1)
dfm_1CC_1_2 = dfm_1CC[start_1_2: end_1_2]
z_1CC_1_2 = zero_index(dfm_1CC_1_2)
dfm_1CC_1_3 = dfm_1CC[start_1_3: end_1_3]
z_1CC_1_3 = zero_index(dfm_1CC_1_3)
dfm_1CC_1 = concat_phases(dfm_1CC_1_1, dfm_1CC_1_2, dfm_1CC_1_3)


# In[13]:


#2AC
dfm_2AC_2_1 = dfm_2AC[start_2_1: end_2_1]
z_2AC_2_1 = zero_index(dfm_2AC_2_1)
dfm_2AC_2_2 = dfm_2AC[start_2_2: end_2_2]
z_2AC_2_2 = zero_index(dfm_2AC_2_2)
dfm_2AC_2_3 = dfm_2AC[start_2_3: end_2_3]
z_2AC_2_3 = zero_index(dfm_2AC_2_3)
dfm_2AC_2 = concat_phases(dfm_2AC_2_1, dfm_2AC_2_2, dfm_2AC_2_3)

dfm_2AC_3_1 = dfm_2AC[start_3_1: end_3_1]
z_2AC_3_1 = zero_index(dfm_2AC_3_1)
dfm_2AC_3_2 = dfm_2AC[start_3_2: end_3_2]
z_2AC_3_2 = zero_index(dfm_2AC_3_2)
dfm_2AC_3_3 = dfm_2AC[start_3_3: end_3_3]
z_2AC_3_3 = zero_index(dfm_2AC_3_3)
dfm_2AC_3 = concat_phases(dfm_2AC_3_1, dfm_2AC_3_2, dfm_2AC_3_3)

dfm_2AC_4_1 = dfm_2AC[start_4_1: end_4_1]
z_2AC_4_1 = zero_index(dfm_2AC_4_1)
dfm_2AC_4_2 = dfm_2AC[start_4_2: end_4_2]
z_2AC_4_2 = zero_index(dfm_2AC_4_2)
dfm_2AC_4_3 = dfm_2AC[start_4_3: end_4_3]
z_2AC_4_3 = zero_index(dfm_2AC_4_3)
dfm_2AC_4 = concat_phases(dfm_2AC_4_1, dfm_2AC_4_2, dfm_2AC_4_3)


# In[14]:


#2CC
dfm_2CC_1_1 = dfm_2CC[start_1_1: end_1_1]
z_2CC_1_1 = zero_index(dfm_2CC_1_1)
dfm_2CC_1_2 = dfm_2CC[start_1_2: end_1_2]
z_2CC_1_2 = zero_index(dfm_2CC_1_2)
dfm_2CC_1_3 = dfm_2CC[start_1_3: end_1_3]
z_2CC_1_3 = zero_index(dfm_2CC_1_3)
```

```
dfm_2CC_1 = concat_phases(dfm_2CC_1_1, dfm_2CC_1_2, dfm_2CC_1_3)

dfm_2CC_2_1 = dfm_2CC[start_2_1: end_2_1]
z_2CC_2_1 = zero_index(dfm_2CC_2_1)
dfm_2CC_2_2 = dfm_2CC[start_2_2: end_2_2]
z_2CC_2_2 = zero_index(dfm_2CC_2_2)
dfm_2CC_2_3 = dfm_2CC[start_2_3: end_2_3]
z_2CC_2_3 = zero_index(dfm_2CC_2_3)
dfm_2CC_2 = concat_phases(dfm_2CC_2_1, dfm_2CC_2_2, dfm_2CC_2_3)

dfm_2CC_3_1 = dfm_2CC[start_3_1: end_3_1]
z_2CC_3_1 = zero_index(dfm_2CC_3_1)
dfm_2CC_3_2 = dfm_2CC[start_3_2: end_3_2]
z_2CC_3_2 = zero_index(dfm_2CC_3_2)
dfm_2CC_3_3 = dfm_2CC[start_3_3: end_3_3]
z_2CC_3_3 = zero_index(dfm_2CC_3_3)
dfm_2CC_3 = concat_phases(dfm_2CC_3_1, dfm_2CC_3_2, dfm_2CC_3_3)

dfm_2CC_4_1 = dfm_2CC[start_4_1: end_4_1]
z_2CC_4_1 = zero_index(dfm_2CC_4_1)
dfm_2CC_4_2 = dfm_2CC[start_4_2: end_4_2]
z_2CC_4_2 = zero_index(dfm_2CC_4_2)
dfm_2CC_4_3 = dfm_2CC[start_4_3: end_4_3]
z_2CC_4_3 = zero_index(dfm_2CC_4_3)
dfm_2CC_4 = concat_phases(dfm_2CC_4_1, dfm_2CC_4_2, dfm_2CC_4_3)


# In[15]:


#2CH
dfm_2CH_1_1 = dfm_2CH[start_1_1: end_1_1]
z_2CH_1_1 = zero_index(dfm_2CH_1_1)
dfm_2CH_1_2 = dfm_2CH[start_1_2: end_1_2]
z_2CH_1_2 = zero_index(dfm_2CH_1_2)
dfm_2CH_1_3 = dfm_2CH[start_1_3: end_1_3]
z_2CH_1_3 = zero_index(dfm_2CH_1_3)
dfm_2CH_1 = concat_phases(dfm_2CH_1_1, dfm_2CH_1_2, dfm_2CH_1_3)


# #### Example of saving and reading in: zero-indexed 1AH 2.1 file
# Same format used for other sets, just adjust the file name.

# In[16]:


#z_1AH_2_1.to_csv('ALVIM_CSV_Zero/z_1AH_2_1.csv')
#z_1AH_2_1 = pd.read_csv('ALVIM_CSV_Zero/z_1AH_2_a', index_col= 0, header = 0, parse_dates = [0])


# #### Example of reading in specific event
# Note that the folder path must be adjusted to your local working environment.

# In[17]:


#dfm_1AH_2_1 = pd.read_csv('ALVIM_CSV_Phase/dfm_1AH_2_1.csv',
#index_col= 0, header = 0, parse_dates = [0])


# ## Sample Curve Showcasing the Impact of Biofilm Scraping:  Line 2AC, Event 2.3

# In[18]:
```

```
fig, ax = plt.subplots(figsize = (40,20))
sns.set(font_scale=2, style="whitegrid")
ax.minorticks_on()

ax.grid(which='minor', linestyle=':', linewidth='0.5', color='grey')
ax.xaxis.set_minor_locator(matplotlib.dates.HourLocator())

s_2AC_2_1 = dfm_2AC_2['2021-09-20 06:00:00':'2021-09-21 23:59:59']
s_2AC_2_1['Potential'].plot(ax=ax, label = '2AC Event 2.3', color = 'blue', fontsize = 35)

plt.ylabel('Potential' ' ' r'[$mV$]', fontsize = 35)
plt.legend(fontsize = 35)
plt.xlabel('Time [Hours]', fontsize = 35)


# ## Sample Curve Phase 2 All Lines
#

# In[17]:


fig, ax = plt.subplots(figsize = (40,20))
sns.set(font_scale=2, style="whitegrid")
ax.minorticks_on()
date_form = matplotlib.dates.DateFormatter("%d/%m")
ax.xaxis.set_major_formatter(date_form)

ax.grid(which='minor', linestyle=':', linewidth='0.5', color='grey')
ax.xaxis.set_minor_locator(matplotlib.dates.DayLocator())

dfm_1AH_2['Potential'].plot(ax=ax, label = '1AH Phase 2', color = 'lightgreen', fontsize = 30)
re_1AH_2 = resample_phases(dfm_1AH_2, '3H')
re_1AH_2.plot(ax=ax, label = '1AH Phase 2 Resampled', color = 'darkgreen', fontsize = 30)


dfm_1CH_2['Potential'].plot(ax=ax, label = '1CH Phase 2', color = 'pink', fontsize = 30)
re_1CH_2 = resample_phases(dfm_1CH_2, '3H')
re_1CH_2.plot(ax=ax, label = '1CH Phase 2 Resampled', color = 'darkred', fontsize = 30)

dfm_2AC_2['Potential'].plot(ax=ax, label = '2AC Phase 2', color = 'violet', fontsize = 30)
re_2AC_2= resample_phases(dfm_2AC_2, '3H')
re_2AC_2.plot(ax=ax, label = '2AC Phase 2 Resampled', color = 'purple', fontsize = 30)


dfm_2CC_2['Potential'].plot(ax=ax, label = '2CC Phase 2', color = 'lightblue', fontsize = 30)
re_2CC_2= resample_phases(dfm_2CC_2, '3H')
re_2CC_2.plot(ax=ax, label = '2CC Phase 2 Resampled', color = 'blue', fontsize = 30)


plt.ylabel('Potential' ' ' r'[$mV$]', fontsize = 40)
plt.legend(fontsize = 20)
plt.xlabel('Time [Days]', fontsize = 40)


# ## Extracting Stagnation Periods from Each Event: Stagnation, Shower, Flush
# **Note: used stag_re for all phases except phase 1, where stag_re2 was used.**

# In[18]:


#1AH
```

```
sr_1AH_2_1 = stag_re(dfm_1AH_2_1, '3H')
zsr_1AH_2_1 = zero_index(sr_1AH_2_1)
zsr_1AH_2_1 = zsr_1AH_2_1.reset_index().rename(columns = {'Potential':'1AH_2_1'}).set_index('ReadTime')

sr_1AH_2_2 = stag_re(dfm_1AH_2_2, '3H')
zsr_1AH_2_2 = zero_index(sr_1AH_2_2)
zsr_1AH_2_2 = zsr_1AH_2_2.reset_index().rename(columns ={'Potential':'1AH_2_2'}).set_index('ReadTime')

sr_1AH_2_3 = stag_re(dfm_1AH_2_3, '3H')
zsr_1AH_2_3 = zero_index(sr_1AH_2_3)
zsr_1AH_2_3 = zsr_1AH_2_3.reset_index().rename(columns = {'Potential':'1AH_2_3'}).set_index('ReadTime')

sr_1AH_3_1 = stag_re(dfm_1AH_3_1, '3H')
zsr_1AH_3_1 = zero_index(sr_1AH_3_1)
zsr_1AH_3_1 = zsr_1AH_3_1.reset_index().rename(columns = {'Potential':'1AH_3_1'}).set_index('ReadTime')

sr_1AH_3_2 = stag_re(dfm_1AH_3_2, '3H')
zsr_1AH_3_2 = zero_index(sr_1AH_3_2)
zsr_1AH_3_2 = zsr_1AH_3_2.reset_index().rename(columns = {'Potential':'1AH_3_2'}).set_index('ReadTime')

sr_1AH_3_3 = stag_re(dfm_1AH_3_3, '3H')
zsr_1AH_3_3 = zero_index(sr_1AH_3_3)
zsr_1AH_3_3 = zsr_1AH_3_3.reset_index().rename(columns = {'Potential':'1AH_3_3'}).set_index('ReadTime')

sr_1AH_4_1 = stag_re(dfm_1AH_4_1, '3H')
zsr_1AH_4_1 = zero_index(sr_1AH_4_1)
zsr_1AH_4_1 = zsr_1AH_4_1.reset_index().rename(columns = {'Potential':'1AH_4_1'}).set_index('ReadTime')

sr_1AH_4_2 = stag_re(dfm_1AH_4_2, '3H')
zsr_1AH_4_2 = zero_index(sr_1AH_4_2)
zsr_1AH_4_2 = zsr_1AH_4_2.reset_index().rename(columns = {'Potential':'1AH_4_2'}).set_index('ReadTime')

sr_1AH_4_3 = stag_re(dfm_1AH_4_3, '3H')
zsr_1AH_4_3 = zero_index(sr_1AH_4_3)
zsr_1AH_4_3 = zsr_1AH_4_3.reset_index().rename(columns = {'Potential':'1AH_4_3'}).set_index('ReadTime')


# In[19]:


#1CH
sr_1CH_1_1 = stag_re_2(dfm_1CH_1_1, '3H')
zsr_1CH_1_1 = zero_index(sr_1CH_1_1)
zsr_1CH_1_1 = zsr_1CH_1_1.reset_index().rename(columns = {'Potential':'1CH_1_1'}).set_index('ReadTime')

sr_1CH_1_2 = stag_re_2(dfm_1CH_1_2, '3H')
zsr_1CH_1_2 = zero_index(sr_1CH_1_2)
zsr_1CH_1_2 = zsr_1CH_1_2.reset_index().rename(columns = {'Potential':'1CH_1_2'}).set_index('ReadTime')

sr_1CH_1_3 = stag_re_2(dfm_1CH_1_3, '3H')
zsr_1CH_1_3 = zero_index(sr_1CH_1_3)
zsr_1CH_1_3 = zsr_1CH_1_3.reset_index().rename(columns ={'Potential':'1CH_1_3'}).set_index('ReadTime')


sr_1CH_2_1 = stag_re(dfm_1CH_2_1, '3H')
zsr_1CH_2_1 = zero_index(sr_1CH_2_1)
zsr_1CH_2_1 = zsr_1CH_2_1.reset_index().rename(columns = {'Potential':'1CH_2_1'}).set_index('ReadTime')

sr_1CH_2_2 = stag_re(dfm_1CH_2_2, '3H')
zsr_1CH_2_2 = zero_index(sr_1CH_2_2)
zsr_1CH_2_2 = zsr_1CH_2_2.reset_index().rename(columns = {'Potential':'1CH_2_2'}).set_index('ReadTime')

sr_1CH_2_3 = stag_re(dfm_1CH_2_3, '3H')
```

```
zsr_1CH_2_3 = zero_index(sr_1CH_2_3)
zsr_1CH_2_3 = zsr_1CH_2_3.reset_index().rename(columns = {'Potential':'1CH_2_3'}).set_index('ReadTime')

sr_1CH_3_1 = stag_re(dfm_1CH_3_1, '3H')
zsr_1CH_3_1 = zero_index(sr_1CH_3_1)
zsr_1CH_3_1 = zsr_1CH_3_1.reset_index().rename(columns = {'Potential':'1CH_3_1'}).set_index('ReadTime')

sr_1CH_3_2 = stag_re(dfm_1CH_3_2, '3H')
zsr_1CH_3_2 = zero_index(sr_1CH_3_2)
zsr_1CH_3_2 = zsr_1CH_3_2.reset_index().rename(columns = {'Potential':'1CH_3_2'}).set_index('ReadTime')

sr_1CH_3_3 = stag_re(dfm_1CH_3_3, '3H')
zsr_1CH_3_3 = zero_index(sr_1CH_3_3)
zsr_1CH_3_3 = zsr_1CH_3_3.reset_index().rename(columns = {'Potential':'1CH_3_3'}).set_index('ReadTime')


sr_1CH_4_1 = stag_re(dfm_1CH_4_1, '3H')
zsr_1CH_4_1 = zero_index(sr_1CH_4_1)
zsr_1CH_4_1 = zsr_1CH_4_1.reset_index().rename(columns = {'Potential':'1CH_4_1'}).set_index('ReadTime')

sr_1CH_4_2 = stag_re(dfm_1CH_4_2, '3H')
zsr_1CH_4_2 = zero_index(sr_1CH_4_2)
zsr_1CH_4_2 = zsr_1CH_4_2.reset_index().rename(columns = {'Potential':'1CH_4_2'}).set_index('ReadTime')

sr_1CH_4_3 = stag_re(dfm_1CH_4_3, '3H')
zsr_1CH_4_3 = zero_index(sr_1CH_4_3)
zsr_1CH_4_3 = zsr_1CH_4_3.reset_index().rename(columns = {'Potential':'1CH_4_3'}).set_index('ReadTime')


# In[20]:


#1CC
sr_1CC_1_1 = stag_re_2(dfm_1CC_1_1, '3H')
zsr_1CC_1_1 = zero_index(sr_1CC_1_1)
zsr_1CC_1_1 = zsr_1CC_1_1.reset_index().rename(columns = {'Potential':'1CC_1_1'}).set_index('ReadTime')

sr_1CC_1_2 = stag_re_2(dfm_1CC_1_2, '3H')
zsr_1CC_1_2 = zero_index(sr_1CC_1_2)
zsr_1CC_1_2 = zsr_1CC_1_2.reset_index().rename(columns = {'Potential':'1CC_1_2'}).set_index('ReadTime')

sr_1CC_1_3 = stag_re_2(dfm_1CC_1_3, '3H')
zsr_1CC_1_3 = zero_index(sr_1CC_1_3)
zsr_1CC_1_3 = zsr_1CC_1_3.reset_index().rename(columns = {'Potential':'1CC_1_3'}).set_index('ReadTime')


# In[21]:


#2AC
sr_2AC_2_1 = stag_re(dfm_2AC_2_1, '3H')
zsr_2AC_2_1 = zero_index(sr_2AC_2_1)
zsr_2AC_2_1 = zsr_2AC_2_1.reset_index().rename(columns = {'Potential':'2AC_2_1'}).set_index('ReadTime')

sr_2AC_2_2 = stag_re(dfm_2AC_2_2, '3H')
zsr_2AC_2_2 = zero_index(sr_2AC_2_2)
zsr_2AC_2_2 = zsr_2AC_2_2.reset_index().rename(columns = {'Potential':'2AC_2_2'}).set_index('ReadTime')

sr_2AC_2_3 = stag_re(dfm_2AC_2_3, '3H')
zsr_2AC_2_3 = zero_index(sr_2AC_2_3)
zsr_2AC_2_3 = zsr_2AC_2_3.reset_index().rename(columns = {'Potential':'2AC_2_3'}).set_index('ReadTime')
```

```
sr_2AC_3_1 = stag_re(dfm_2AC_3_1, '3H')
zsr_2AC_3_1 = zero_index(sr_2AC_3_1)
zsr_2AC_3_1 = zsr_2AC_3_1.reset_index().rename(columns = {'Potential':'2AC_3_1'}).set_index('ReadTime')

sr_2AC_3_2 = stag_re(dfm_2AC_3_2, '3H')
zsr_2AC_3_2 = zero_index(sr_2AC_3_2)
zsr_2AC_3_2 = zsr_2AC_3_2.reset_index().rename(columns = {'Potential':'2AC_3_2'}).set_index('ReadTime')

sr_2AC_3_3 = stag_re(dfm_2AC_3_3, '3H')
zsr_2AC_3_3 = zero_index(sr_2AC_3_3)
zsr_2AC_3_3 = zsr_2AC_3_3.reset_index().rename(columns = {'Potential':'2AC_3_3'}).set_index('ReadTime')

sr_2AC_4_1 = stag_re(dfm_2AC_4_1, '3H')
zsr_2AC_4_1 = zero_index(sr_2AC_4_1)
zsr_2AC_4_1 = zsr_2AC_4_1.reset_index().rename(columns = {'Potential':'2AC_4_1'}).set_index('ReadTime')

sr_2AC_4_2 = stag_re(dfm_2AC_4_2, '3H')
zsr_2AC_4_2 = zero_index(sr_2AC_4_2)
zsr_2AC_4_2 = zsr_2AC_4_2.reset_index().rename(columns = {'Potential':'2AC_4_2'}).set_index('ReadTime')

sr_2AC_4_3 = stag_re(dfm_2AC_4_3, '3H')
zsr_2AC_4_3 = zero_index(sr_2AC_4_3)
zsr_2AC_4_3 = zsr_2AC_4_3.reset_index().rename(columns = {'Potential':'2AC_4_3'}).set_index('ReadTime')


# In[22]:


#2CC
sr_2CC_1_1 = stag_re_2(dfm_2CC_1_1, '3H')
zsr_2CC_1_1 = zero_index(sr_2CC_1_1)
zsr_2CC_1_1 = zsr_2CC_1_1.reset_index().rename(columns = {'Potential':'2CC_1_1'}).set_index('ReadTime')

sr_2CC_1_2 = stag_re_2(dfm_2CC_1_2, '3H')
zsr_2CC_1_2 = zero_index(sr_2CC_1_2)
zsr_2CC_1_2 = zsr_2CC_1_2.reset_index().rename(columns = {'Potential':'2CC_1_2'}).set_index('ReadTime')

sr_2CC_1_3 = stag_re_2(dfm_2CC_1_3, '3H')
zsr_2CC_1_3 = zero_index(sr_2CC_1_3)
zsr_2CC_1_3 = zsr_2CC_1_3.reset_index().rename(columns = {'Potential':'2CC_1_3'}).set_index('ReadTime')

sr_2CC_2_1 = stag_re(dfm_2CC_2_1, '3H')
zsr_2CC_2_1 = zero_index(sr_2CC_2_1)
zsr_2CC_2_1 = zsr_2CC_2_1.reset_index().rename(columns = {'Potential':'2CC_2_1'}).set_index('ReadTime')

sr_2CC_2_2 = stag_re(dfm_2CC_2_2, '3H')
zsr_2CC_2_2 = zero_index(sr_2CC_2_2)
zsr_2CC_2_2 = zsr_2CC_2_2.reset_index().rename(columns = {'Potential':'2CC_2_2'}).set_index('ReadTime')

sr_2CC_2_3 = stag_re(dfm_2CC_2_3, '3H')
zsr_2CC_2_3 = zero_index(sr_2CC_2_3)
zsr_2CC_2_3 = zsr_2CC_2_3.reset_index().rename(columns = {'Potential':'2CC_2_3'}).set_index('ReadTime')

sr_2CC_3_1 = stag_re(dfm_2CC_3_1, '3H')
zsr_2CC_3_1 = zero_index(sr_2CC_3_1)
zsr_2CC_3_1 = zsr_2CC_3_1.reset_index().rename(columns = {'Potential':'2CC_3_1'}).set_index('ReadTime')

sr_2CC_3_2 = stag_re(dfm_2CC_3_2, '3H')
zsr_2CC_3_2 = zero_index(sr_2CC_3_2)
zsr_2CC_3_2 = zsr_2CC_3_2.reset_index().rename(columns = {'Potential':'2CC_3_2'}).set_index('ReadTime')

sr_2CC_3_3 = stag_re(dfm_2CC_3_3, '3H')
zsr_2CC_3_3 = zero_index(sr_2CC_3_3)
```

```
zsr_2CC_3_3 = zsr_2CC_3_3.reset_index().rename(columns = {'Potential':'2CC_3_3'}).set_index('ReadTime')

sr_2CC_4_1 = stag_re(dfm_2CC_4_1, '3H')
zsr_2CC_4_1 = zero_index(sr_2CC_4_1)
zsr_2CC_4_1 = zsr_2CC_4_1.reset_index().rename(columns = {'Potential':'2CC_4_1'}).set_index('ReadTime')

sr_2CC_4_2 = stag_re(dfm_2CC_4_2, '3H')
zsr_2CC_4_2 = zero_index(sr_2CC_4_2)
zsr_2CC_4_2 = zsr_2CC_4_2.reset_index().rename(columns = {'Potential':'2CC_4_2'}).set_index('ReadTime')

sr_2CC_4_3 = stag_re(dfm_2CC_4_3, '3H')
zsr_2CC_4_3 = zero_index(sr_2CC_4_3)
zsr_2CC_4_3 = zsr_2CC_4_3.reset_index().rename(columns = {'Potential':'2CC_4_3'}).set_index('ReadTime')


# In[23]:


#2CH
sr_2CH_1_1 = stag_re_2(dfm_2CH_1_1, '3H')
zsr_2CH_1_1 = zero_index(sr_2CH_1_1)
zsr_2CH_1_1 = zsr_2CH_1_1.reset_index().rename(columns = {'Potential':'2CH_1_1'}).set_index('ReadTime')

sr_2CH_1_2 = stag_re_2(dfm_2CH_1_2, '3H')
zsr_2CH_1_2 = zero_index(sr_2CH_1_2)
zsr_2CH_1_2 = zsr_2CH_1_2.reset_index().rename(columns = {'Potential':'2CH_1_2'}).set_index('ReadTime')

sr_2CH_1_3 = stag_re_2(dfm_2CH_1_3, '3H')
zsr_2CH_1_3 = zero_index(sr_2CH_1_3)
zsr_2CH_1_3 = zsr_2CH_1_3.reset_index().rename(columns = {'Potential':'2CH_1_3'}).set_index('ReadTime')


#Remove the comments below to save the stagnation period data in own files.

# In[24]:


zsr_flush = (pd.concat([zsr_1AH_2_1, zsr_1AH_2_2, zsr_1AH_2_3, zsr_1CH_1_1, zsr_1CH_1_2, zsr_1CH_1_3,
zsr_1CC_1_1, zsr_1CC_1_2, zsr_1CC_1_3, zsr_2AC_2_1, zsr_2AC_2_2, zsr_2AC_2_3, zsr_1CH_4_1, zsr_1CH_4_2,
zsr_1CH_4_3, zsr_2CC_1_1, zsr_2CC_1_2, zsr_2CC_1_3, zsr_2CC_4_1, zsr_2CC_4_2, zsr_2CC_4_3, zsr_2CH_1_1,
zsr_2CH_1_2, zsr_2CH_1_3], axis = 1))
#zsr_flush.to_csv('ALVIM_Stag_Re3/zsr_flush.csv')


# In[25]:


zsr_show = (pd.concat([zsr_1AH_4_1, zsr_1AH_4_2, zsr_1AH_4_3, zsr_1CH_3_1, zsr_1CH_3_2, zsr_1CH_3_3,
zsr_2AC_4_1, zsr_2AC_4_2, zsr_2AC_4_3, zsr_2CC_3_1, zsr_2CC_3_2, zsr_2CC_3_3], axis = 1))
#zsr_show.to_csv('ALVIM_Stag_Re3/zsr_show.csv')


# In[26]:


zsr_stag = (pd.concat([zsr_1AH_3_1, zsr_1AH_3_2, zsr_1AH_3_3, zsr_1CH_2_1, zsr_1CH_2_2, zsr_1CH_2_3,
zsr_2AC_3_1, zsr_2AC_3_2, zsr_2AC_3_3, zsr_2CC_2_1, zsr_2CC_2_2, zsr_2CC_2_3], axis = 1))
#zsr_stag.to_csv('ALVIM_Stag_Re3/zsr_stag.csv')


# # Curve Fitting Stagnation Periods
#
```

```
# In[47]:


zsr_shower = pd.read_csv('ALVIM_Stag_Re3/zsr_show.csv', index_col= 0, header = 0, parse_dates = [0])
zsr_flushes = pd.read_csv('ALVIM_Stag_Re3/zsr_flush.csv', index_col= 0, header = 0, parse_dates = [0])
zsr_stagnation = pd.read_csv('ALVIM_Stag_Re3/zsr_stag.csv', index_col= 0, header = 0, parse_dates = [0])


# ### Shower Events

# In[48]:


d_zsr_shower = zsr_shower.dropna()
#Two of the data sets contain 6 rather than 8 days of stagnation. None contain the anticipated 9.
shower_6d = pd.concat([zsr_shower['1CH_3_2'], zsr_shower['2CC_3_2'] ], axis = 1)
shower_6d = shower_6d.dropna()
shower_8d = (pd.concat([zsr_shower['1AH_4_1'], zsr_shower['1AH_4_2'], zsr_shower['1AH_4_3'],
zsr_shower['1CH_3_1'], zsr_shower['1CH_3_3'], zsr_shower['2AC_4_1'], zsr_shower['2AC_4_2'],
zsr_shower['2AC_4_3'], zsr_shower['2CC_3_1'], zsr_shower['2CC_3_3']], axis = 1))
shower_8d = shower_8d.dropna()


# In[49]:


#Grouping and averaging by temperature and event (and disinfection).
#Note that all showers from Phase 3 and 4, i.e. all shower events contained monochloramine.

hot_shower = (pd.concat((d_zsr_shower['1AH_4_1'], d_zsr_shower['1AH_4_2'], d_zsr_shower['1AH_4_3'],
d_zsr_shower['1CH_3_1'], d_zsr_shower['1CH_3_2'], d_zsr_shower['1CH_3_3']), axis = 1).mean(axis = 1))
cold_shower = (pd.concat((d_zsr_shower['2AC_4_1'], d_zsr_shower['2AC_4_2'], d_zsr_shower['2AC_4_3'],
d_zsr_shower['2CC_3_1'], d_zsr_shower['2CC_3_2'], d_zsr_shower['2CC_3_3'] ), axis = 1).mean(axis = 1))


# ### Flushing Events

# In[50]:


d_zsr_flushes = zsr_flushes.dropna()


# In[51]:


flush_2d = (pd.concat([zsr_flushes['1CH_1_3'], zsr_flushes['1CC_1_3'], zsr_flushes['2CC_1_3'],
zsr_flushes['2CH_1_3'], zsr_flushes['2AC_2_3']], axis = 1))
flush_2d = flush_2d.dropna()
flush_8d = (pd.concat([zsr_flushes['1AH_2_1'], zsr_flushes['1AH_2_2'], zsr_flushes['1CH_1_1'],
zsr_flushes['1CC_1_1'], zsr_flushes['2AC_2_1'], zsr_flushes['2AC_2_2'], zsr_flushes['1CH_4_1'],
zsr_flushes['1CH_4_2'], zsr_flushes['2CC_1_1'], zsr_flushes['2CC_4_1'], zsr_flushes['2CC_4_2'],
zsr_flushes['2CC_4_3'], zsr_flushes['2CH_1_1'] ], axis = 1))
flush_8d = flush_8d.dropna()
flush_9d = pd.concat([zsr_flushes['1AH_2_3'], zsr_flushes['1CH_4_3']], axis = 1)
flush_9d = flush_9d.dropna()
flush_22d = (pd.concat([zsr_flushes['1CH_1_2'], zsr_flushes['1CC_1_2'], zsr_flushes['2CC_1_2'],
zsr_flushes['2CH_1_2']], axis = 1))
flush_22d = flush_22d.dropna()

flush_8_more = pd.concat([flush_8d, flush_9d], axis = 1)
flush_8_more = flush_8_more.dropna()
```

```
# In[52]:


#Grouping and averaging by temperature and event
hot_flush = (pd.concat((flush_8_more['1AH_2_1'], flush_8_more['1AH_2_2'], flush_8_more['1CH_1_1'],
flush_8_more['1CH_4_1'], flush_8_more['1CH_4_2'], flush_8_more['2CH_1_1'], flush_8_more['1AH_2_3'],
flush_8_more['1CH_4_3'] ), axis = 1).mean(axis = 1))
cold_flush = (pd.concat((flush_8_more['1CC_1_1'], flush_8_more['2AC_2_1'], flush_8_more['2AC_2_2'],
flush_8_more['2CC_1_1'], flush_8_more['2CC_4_1'], flush_8_more['2CC_4_2'], flush_8_more['2CC_4_3']),
axis = 1).mean(axis = 1))


# In[53]:


#Grouping and avergaing by temperature, disinfection and event
Hot_no_d_8mf = (pd.concat([flush_8_more['1AH_2_1'], flush_8_more['1AH_2_2'], flush_8_more['1CH_1_1'],
flush_8_more['2CH_1_1'], flush_8_more['1AH_2_3']], axis = 1).mean(axis = 1))
Cold_no_d_8mf = (pd.concat([flush_8_more['1CC_1_1'], flush_8_more['2AC_2_1'], flush_8_more['2AC_2_2'],
flush_8_more['2CC_1_1'], flush_8_more['2AC_2_1']], axis = 1).mean(axis = 1))
Hot_d_8mf = (pd.concat([flush_8_more['1CH_4_1'], flush_8_more['1CH_4_2'], flush_8_more['1CH_4_3']],
axis = 1).mean(axis = 1))
Cold_d_8mf = (pd.concat([flush_8_more['2CC_4_1'], flush_8_more['2CC_4_2'], flush_8_more['2CC_4_3']],
axis = 1).mean(axis = 1))


# ### Stagnation Events

# In[54]:


d_zsr_stag = zsr_stagnation.dropna()


# In[55]:


stag_11d = (pd.concat([zsr_stagnation['1AH_3_1'], zsr_stagnation['1AH_3_2'], zsr_stagnation['2AC_3_1'],
zsr_stagnation['2AC_3_2']], axis = 1))
stag_11d = stag_11d.dropna()
stag_13d = (pd.concat([zsr_stagnation['1CH_2_1'], zsr_stagnation['1AH_3_3'], zsr_stagnation['1CH_2_2'],
zsr_stagnation['2AC_3_3'], zsr_stagnation['2CC_2_1'], zsr_stagnation['2CC_2_2']], axis = 1))
stag_13d = stag_13d.dropna()
stag_14d = pd.concat([zsr_stagnation['1CH_2_3'], zsr_stagnation['2CC_2_3']], axis = 1)
stag_14d = stag_14d.dropna()
#Used d_zsr_stag regardless, but was of interest to see distribution of days.


# In[56]:


hot_stag = (pd.concat((d_zsr_stag['1AH_3_1'], d_zsr_stag['1AH_3_2'], d_zsr_stag['1AH_3_3'],
d_zsr_stag['1CH_2_1'], d_zsr_stag['1CH_2_2'], d_zsr_stag['1CH_2_3']),
axis = 1).mean(axis = 1))
cold_stag = (pd.concat((d_zsr_stag['2AC_3_1'], d_zsr_stag['2AC_3_2'], d_zsr_stag['2AC_3_3'],
d_zsr_stag['2CC_2_1'], d_zsr_stag['2CC_2_2'], d_zsr_stag['2CC_2_3']),
axis = 1).mean(axis = 1))


# In[57]:
```

```
Hot_no_d_st = (pd.concat((d_zsr_stag['1CH_2_1'], d_zsr_stag['1CH_2_2'], d_zsr_stag['1CH_2_3'] ),
axis = 1).mean(axis = 1))
Cold_no_d_st = (pd.concat((d_zsr_stag['2CC_2_1'], d_zsr_stag['2CC_2_2'], d_zsr_stag['2CC_2_3']),
axis = 1).mean(axis = 1))
Hot_d_st = (pd.concat((d_zsr_stag['1AH_3_1'], d_zsr_stag['1AH_3_2'], d_zsr_stag['1AH_3_3']),
axis = 1).mean(axis = 1))
Cold_d_st = (pd.concat((d_zsr_stag['2AC_3_1'], d_zsr_stag['2AC_3_2'], d_zsr_stag['2AC_3_3']),
axis = 1).mean(axis = 1))
```

 # ## Plot of Temperature-Grouped Events

 # In[74]:

```
fig, ax = plt.subplots(figsize = (40,20))
sns.set(font_scale=2, style="whitegrid")
ax.grid(which='minor', linestyle=':', linewidth='0.5', color='grey')

hot_flush.plot(ax=ax, label = 'Flush Hot', fontsize = 30, color = 'red')
cold_flush.plot(ax=ax,  label = 'Flush Cold', fontsize = 30, color = 'blue')

hot_shower.plot(ax=ax, label = 'Shower Hot', fontsize = 30, color = 'red', linestyle = ':')
cold_shower.plot(ax=ax,  label = 'Shower Cold', fontsize = 30, color = 'blue', linestyle = ':')

hot_stag.plot(ax=ax, label = 'Stagnation Hot', fontsize = 30, color = 'red', linestyle = '-.')
cold_stag.plot(ax=ax,  label = 'Stagnation Cold', fontsize = 30, color = 'blue', linestyle = '-.')

plt.ylabel('Potential' ' ' r'[$mV$]', fontsize = 40)
plt.legend(fontsize = 35)
plt.xlabel('Time [Hours]', fontsize = 35)
```

 # ## Plot of Temperature-Disinfection-Grouped Events

 # In[73]:

```
fig, ax = plt.subplots(figsize = (40,20))
sns.set(font_scale=2, style="whitegrid")
ax.grid(which='minor', linestyle=':', linewidth='0.5', color='grey')

hot_shower.plot(ax = ax, label = 'Hot Shower Monochloramine', fontsize = 30, color = 'red')
cold_shower.plot(ax = ax, label = 'Cold Shower Monochloramine', fontsize = 30, color = 'navy')

Hot_no_d_st.plot(ax=ax, label = 'Hot Stagnation None', fontsize = 30, color = 'orange')
Cold_no_d_st.plot(ax=ax, label = 'Cold Stagnation None', fontsize = 30, color = 'blue')
Hot_d_st.plot(ax=ax, label = 'Hot Stagnation Monochloramine', fontsize = 30, color = 'red',
linestyle = ':')
Cold_d_st.plot(ax=ax, label = 'Cold Stagnation Monochloramine', fontsize = 30, color = 'navy',
linestyle = ':')


Hot_no_d_8mf.plot(ax = ax, label = 'Hot Flush None', fontsize = 30, color = 'orange', linestyle = '-.')
Cold_no_d_8mf.plot(ax = ax, label = 'Cold Flush None', fontsize = 30, color = 'blue', linestyle = '-.')
Hot_d_8mf.plot(ax = ax, label = 'Hot Flush Monochloramine', fontsize = 30, color = 'red',
linestyle = '-.')
Cold_d_8mf.plot(ax = ax, label = 'Cold Flush Monochloramine', fontsize = 30, color = 'navy',
linestyle = '-.')

ax.set_xlim(1, 60)
plt.ylabel('Potential' ' ' r'[$mV$]', fontsize = 40)
plt.legend(fontsize = 28)
```

```
plt.xlabel('Time [Hours]', fontsize = 40)
#slightly adjusted version with only 6 days displayed in thesis, but is based on this set.


# ## Plot of Flushing Events by Phase

# In[38]:


flush_ph1_h = pd.concat((flush_8_more['1CH_1_1'], flush_8_more['2CH_1_1']), axis = 1).mean(axis = 1)
flush_ph1_c = pd.concat((flush_8_more['1CC_1_1'], flush_8_more['2CC_1_1']), axis = 1).mean(axis = 1)

flush_ph2_h = (pd.concat((flush_8_more['1AH_2_1'], flush_8_more['1AH_2_2'], flush_8_more['1AH_2_3']),
axis = 1).mean(axis = 1))
flush_ph2_c = pd.concat((flush_8_more['2AC_2_1'], flush_8_more['2AC_2_2']), axis = 1).mean(axis = 1)

flush_ph4_h = (pd.concat((flush_8_more['1CH_4_1'], flush_8_more['1CH_4_2'], flush_8_more['1CH_4_3']),
axis = 1).mean(axis = 1))
flush_ph4_c = (pd.concat((flush_8_more['2CC_4_1'], flush_8_more['2CC_4_2'], flush_8_more['2CC_4_3']),
axis = 1).mean(axis = 1))



# In[45]:


fig, ax = plt.subplots(figsize = (40,20))
sns.set(font_scale=2, style="whitegrid")
ax.grid(which='minor', linestyle=':', linewidth='0.5', color='grey')

flush_ph1_h.plot(ax=ax, label = 'Flush Phase 1 Hot', fontsize = 30, color = 'red', linestyle = '-')
flush_ph1_c.plot(ax=ax, label = 'Flush Phase 1 Cold', fontsize = 30, color = 'blue', linestyle = '-')

flush_ph2_h.plot(ax=ax, label = 'Flush Phase 2 Hot', fontsize = 30, color = 'red', linestyle = ':')
flush_ph2_c.plot(ax=ax, label = 'Flush Phase 2 Cold', fontsize = 30, color = 'blue', linestyle = ':')

flush_ph4_h.plot(ax=ax, label = 'Flush Phase 4 Hot', fontsize = 30, color = 'red', linestyle = 'dashed')
flush_ph4_c.plot(ax=ax, label = 'Flush Phase 4 Cold', fontsize = 30, color = 'blue', linestyle = 'dashed')

plt.ylabel('Potential' ' ' r'[$mV$]', fontsize = 40)
plt.legend(fontsize = 35)
plt.xlabel('Time [Hours]', fontsize = 35)


# # Creation of RMSE-heatmaps

# In[95]:


from sklearn.metrics import mean_squared_error
from itertools import permutations


# In[96]:


#Functions MSE, RMSE
def mse(y1, y2):
    return np.sum((y1 - y2) ** 2) / len(y1)
def rmse(y1, y2):
    return np.sqrt(mean_squared_error(y1, y2))


# In[97]:
```

```
#Make grid of RMSE-values
def rmse_grid(df):
    sensors = df.columns
    result = dict()
    for sensor in sensors:
        result[sensor] = dict()
    for sens1, sens2 in list(permutations(sensors, 2)):
        m = rmse(df[sens1], df[sens2])
        result[sens1][sens2] = m
    result = pd.DataFrame(result)
    return result


# In[98]:


rmse_shower = rmse_grid(d_zsr_shower)

mask = np.triu(np.ones_like(rmse_shower, dtype=bool))
f, ax = plt.subplots(figsize=(25, 20))
cmap = sns.color_palette("rocket", as_cmap=True)
sns.heatmap(rmse_shower, center=0, mask=mask, cmap=cmap,
            square=True, linewidths=.5, cbar_kws={"shrink": .5});


# In[99]:


rmse_flush = rmse_grid(flush_8_more)
mask = np.triu(np.ones_like(rmse_flush, dtype=bool))
f, ax = plt.subplots(figsize=(25, 20))
cmap = sns.color_palette("rocket", as_cmap=True)
sns.heatmap(rmse_flush, center=0, mask=mask, cmap=cmap,
            square=True, linewidths=.5, cbar_kws={"shrink": .5});


# In[102]:


rmse_stag = rmse_grid(d_zsr_stag)
mask = np.triu(np.ones_like(rmse_stag, dtype=bool))
f, ax = plt.subplots(figsize=(25, 20))
cmap = sns.color_palette("rocket", as_cmap=True)
sns.heatmap(rmse_stag, center=0, mask=mask, cmap=cmap,
            square=True, linewidths=.5, cbar_kws={"shrink": .5});


# In[105]:


#Example of comparison of rigs (materials) vs temperature using RMSE
Rig_1_1_2_rmse = rmse(flush_22d['1CH_1_2'], flush_22d['1CC_1_2'])
Rig_2_and_1_1_2_rmse = rmse(flush_22d['1CH_1_2'], flush_22d['2CH_1_2'])
print(f'Same material, but hot vs cold: {Rig_1_1_2_rmse}')
print(f'Different materials, but both hot: {Rig_2_and_1_1_2_rmse}')


# # Curve Fitting Using the package lmfit()

# In[106]:
```

```
from scipy.optimize import curve_fit
from lmfit import Model
import numpy as np


# In[107]:


def func_exp(x, a, b, c): #keep x as first argument, needed for lmfit
    a = float(a)
    b = float(b)
    c = float(c)
    x = x.astype('float')
    return a * np.exp(-b * x) + c


# In[108]:


def curve_fit_exp(df, y1, a1, b1, c1):
    f = pd.to_timedelta(df.index)
    t = f.total_seconds()
    t = t/ 3600
    f1_model = Model(func_exp)
    params = f1_model.make_params(a = a1, b= b1, c=c1)
    result = f1_model.fit(y1, params, x = t)
    plt.plot(t, y1, 'o')
    plt.plot(t, result.best_fit, '-', label='best fit')
    plt.legend()
    plt.show()
    return result


# In[111]:


def func_poly(x, a, b, c, d):
    a = float(a)
    b = float(b)
    c = float(c)
    d = float(d)
    x = x.astype('float')
    return a*x**3 + b*x**2 + c*x + d


# In[112]:


def curve_fit_poly(df, y1, a1, b1, c1, d1):
    f = pd.to_timedelta(df.index)
    t = f.total_seconds()
    t = t/ 3600
    f2_model = Model(func_poly)
    params = f2_model.make_params(a = a1, b= b1, c=c1, d = d1)
    result = f2_model.fit(y1, params, x = t)
    plt.plot(t, y1, 'o')
    plt.plot(t, result.best_fit, '-', label='best fit')
    plt.legend()
    plt.show()
    return result


# In[113]:
```

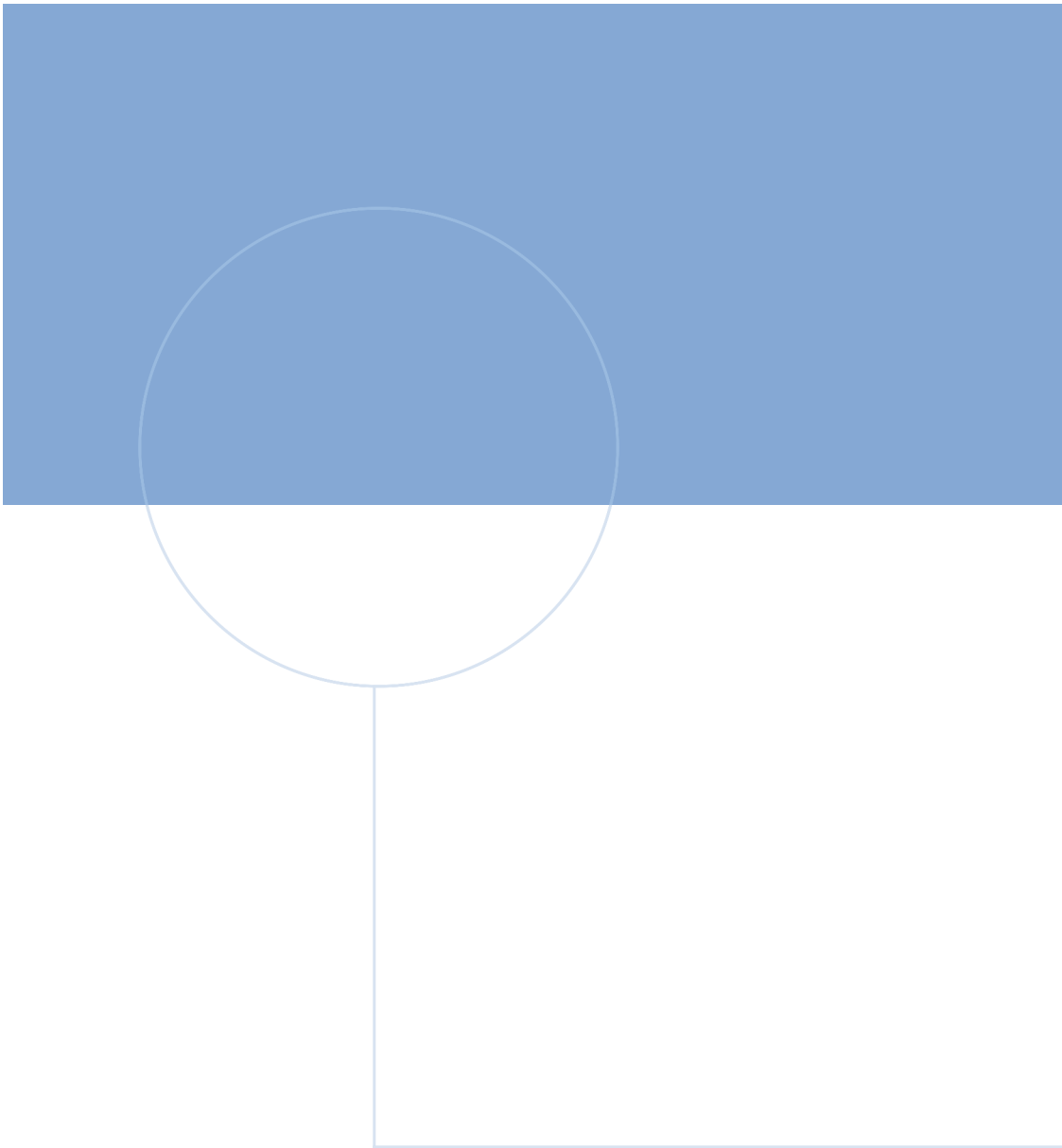
```
def func_line(x, a, b):
    a = float(a)
    b = float(b)
    x = x.astype('float')
    return x*a + b


# In[114]:


def curve_fit_line(df, y1, a1, b1):
    f = pd.to_timedelta(df.index)
    t = f.total_seconds()
    t = t/ 3600
    f3_model = Model(func_line)
    params = f3_model.make_params(a = a1, b= b1)
    result = f3_model.fit(y1, params, x = t)
    plt.plot(t, y1, 'o')
    plt.plot(t, result.best_fit, '-', label='best fit')
    plt.legend()
    plt.show()
    return result
    #print(result.fit_report())


# ### Other Functions Initially Tested, Unused
#

# In[ ]:


#def func_line_exp(x, a, b, c, d):
    #a = float(a)
    #b = float(b)
    #c = float(c)
    #d = float(d)
    #x = x.astype('float')
    #return a * np.exp(-b * x) + c * x + d


# In[ ]:


#def power_law(x, a, b):
    #a = float(a)
    #b = float(b)
    #x = x.astype('float')
    #return a* x^b


# In[ ]:


#def log_growth(x, a, b, c):
    #a = float(a)
    #b = float(b)
    #c = float(c)
    #x = x.astype('float')
    #return (a)/(1+ c*np.exp(-b*x))
```

```
# In[120]:


#sample fitting: same procedure for the rest, except the number of coefficients.
#AIC, BIC, reduced chi-square, coefficients as well as relative error of coefficients noted manually
#for each function on each event: 3 x 3 = 9 excel sheets total.
curve_fit_exp(flush_8_more, flush_8_more['2AC_2_2'], 0.01, 0.0001, 0.001)


# ## Curve Fitting the Biofilm Growth Data

# In[125]:


#Phase 1
re_1CH_1_1= resample_phases(dfm_1CH_1_1['2021-03-11':'2021-03-16'], '3H')
re_1CC_1_1= resample_phases(dfm_1CC_1_1['2021-03-11':'2021-03-16'], '3H')
re_2CC_1_1= resample_phases(dfm_2CC_1_1['2021-03-10':'2021-03-16'], '3H')
re_1CH_1_2 = resample_phases(dfm_1CH_1_2['2021-03-30':'2021-04-11'], '3H')
re_1CC_1_2 = resample_phases(dfm_1CC_1_2['2021-03-30':'2021-04-08'], '3H')
re_2CC_1_2= resample_phases(dfm_2CC_1_2['2021-04-02':'2021-04-11'], '3H')


# In[126]:


#Phase 2
re_1CH_2_1= resample_phases(dfm_1CH_2_1['2021-09-01':'2021-09-05'], '3H')
re_2AC_2_1= resample_phases(dfm_2AC_2_1['2021-08-31':'2021-09-05'], '3H')
re_2CC_2_1= resample_phases(dfm_2CC_2_1['2021-09-03':'2021-09-05'], '3H')
re_2AC_2_2 = resample_phases(dfm_2AC_2_2['2021-09-14':'2021-09-19'], '3H')
re_1CH_2_3= resample_phases(dfm_1CH_2_3['2021-10-01':'2021-10-03'], '3H')
re_2AC_2_3= resample_phases(dfm_2AC_2_3['2021-09-29':'2021-10-03'], '3H')


# In[127]:


#Phase 3
re_2AC_3_1= resample_phases(dfm_2AC_3_1['2021-11-30':'2021-12-06'], '3H')
re_1AH_3_2= resample_phases(dfm_1AH_3_2['2021-12-17':'2021-12-18'], '3H')
re_2AC_3_2= resample_phases(dfm_2AC_3_2['2021-12-15':'2021-12-18'], '3H')
re_2CC_3_2= resample_phases(dfm_2CC_3_2['2021-12-14':'2021-12-18'], '3H')
re_1AH_3_3= resample_phases(dfm_1AH_3_3['2022-01-13':'2022-01-17'], '3H')
re_1CH_3_3= resample_phases(dfm_1CH_3_3['2022-01-15':'2022-01-17'], '3H')
re_2AC_3_3= resample_phases(dfm_2AC_3_3['2022-01-11':'2022-01-16'], '3H')

#No observable biofilm growth curves found for Phase 4.


# In[128]:


#Zero-indexing the biofilm growth curves:
ze_1CH_1_1 = zero_index(re_1CH_1_1)
ze_1CC_1_1 = zero_index(re_1CC_1_1)
ze_2CC_1_1 = zero_index(re_2CC_1_1)
ze_1CH_1_2 = zero_index(re_1CH_1_2)
ze_1CC_1_2 = zero_index(re_1CC_1_2)
ze_2CC_1_2 = zero_index(re_2CC_1_2)

ze_1CH_2_1 = zero_index(re_1CH_2_1)
ze_2AC_2_1 = zero_index(re_2AC_2_1)
```

```
ze_2CC_2_1 = zero_index(re_2CC_2_1)
ze_2AC_2_2 = zero_index(re_2AC_2_2)
ze_1CH_2_3 = zero_index(re_1CH_2_3)
ze_2AC_2_3 = zero_index(re_2AC_2_3)

ze_2AC_3_1 = zero_index(re_2AC_3_1)
ze_2AC_3_2 = zero_index(re_2AC_3_2)
ze_2CC_3_2 = zero_index(re_2CC_3_2)
ze_1AH_3_3 = zero_index(re_1AH_3_3)
ze_1CH_3_3 = zero_index(re_1CH_3_3)
ze_2AC_3_3 = zero_index(re_2AC_3_3)


# In[131]:


#Function used to find the initial potential, maximum potential and change in potential.
def pot_change(df):
    change = df.max()-df[0]
    return "{:.1f}".format(df[0]), "{:.1f}".format(df.max()), "{:.1f}".format(change)


# In[ ]:


#Example of how to find index of maximum value, for calculating #rate of biofilm change:
ze_1CH_1_1[ze_1CH_1_1==ze_1CH_1_1.max()]


# In[152]:


#Plot of biofilm growth curves, zero-indexed, not used in paper:
fig, ax = plt.subplots(figsize = (40,20))
sns.set(font_scale=2, style="whitegrid")
ax.grid(which='minor', linestyle=':', linewidth='0.5', color='grey')

ze_1CH_1_1.plot(ax=ax, label = '1CH 1.1', color = 'orangered')
ze_1CC_1_1.plot(ax=ax, label = '1CC 1.1', color = 'lightseagreen')
ze_2CC_1_1.plot(ax=ax, label = '2CC 1.1', color = 'cadetblue')
ze_1CH_1_2.plot(ax=ax, label = '1CH 1.2', color = 'orange')
ze_1CC_1_2.plot(ax=ax, label = '1CC 1.2', color = 'darkcyan')
ze_2CC_1_2.plot(ax=ax, label = '2CC 1.2', color = 'cyan')

ze_1CH_2_1.plot(ax=ax, label = '1CH 2.1', color = 'maroon')
ze_2AC_2_1.plot(ax=ax, label = '2AC 2.1', color = 'deepskyblue')
ze_2CC_2_1.plot(ax=ax, label = '2CC 2.1', color = 'steelblue')
ze_2AC_2_2.plot(ax=ax, label = '2AC 2.2', color = 'dodgerblue')
ze_1CH_2_3.plot(ax=ax, label = '1CH 2.3', color = 'salmon')
ze_2AC_2_3.plot(ax=ax, label = '2AC 2.3', color = 'blue')

ze_2AC_3_1.plot(ax=ax, label = '2AC 3.1', color = 'slateblue')
ze_2AC_3_2.plot(ax=ax, label = '2AC 3.2', color = 'darkblue')
ze_2CC_3_2.plot(ax=ax, label = '2CC 3.2', color = 'lightsteelblue')
ze_1AH_3_3.plot(ax=ax, label = '1AH 3.3', color = 'red')
ze_1CH_3_3.plot(ax=ax, label = '1CH 3.3', color = 'firebrick')
ze_2AC_3_3.plot(ax=ax, label = '2AC 3.3', color = 'cornflowerblue')

plt.ylabel('Potential' ' ' r'[$mV$]')
plt.legend()
plt.xlabel('Time [Days]')
```

```
# ### Evaluation of Goodness of Fit Measures: Biofilm Growth Data

# In[148]:


biofilm_exp = (pd.read_excel('Curve_fitting/Biofilm_curve_fitting.xlsx',
index_col=0, header=0, sheet_name = 0, parse_dates = [0]))
biofilm_poly = (pd.read_excel('Curve_fitting/Biofilm_curve_fitting.xlsx',
index_col=0, header=0, sheet_name = 1, parse_dates = [0]))
biofilm_lin = (pd.read_excel('Curve_fitting/Biofilm_curve_fitting.xlsx',
index_col=0, header=0, sheet_name = 2, parse_dates = [0]))


# In[149]:


coeff_a_poly = biofilm_poly['a'].dropna()
coeff_b_poly = biofilm_poly['b'].dropna()
coeff_c_poly = biofilm_poly['c'].dropna()
coeff_d_poly = biofilm_poly['d'].dropna()


# In[151]:


#Plot of coefficients of biofilm growth data fitted with polynomial function
fig, ax = plt.subplots(figsize = (40,20))
sns.set(font_scale=2, style="whitegrid")
ax.grid(which='minor', linestyle=':', linewidth='0.5', color='grey')

x = biofilm_poly.index.dropna()


for label in (ax.get_xticklabels() + ax.get_yticklabels()):
    label.set_fontsize(25)

ax.scatter(x, coeff_a_poly, s = 500)
#ax.scatter(x, coeff_b_poly, s = 500)
#ax.scatter(x, coeff_c_poly, s = 500)
#ax.scatter(x, coeff_d_poly, s = 500)

plt.ylabel('Coefficient Values', fontsize = 30)
plt.xlabel('Events', fontsize = 30)
plt.legend()
plt.xlabel('', fontsize = 30)


# ### Evaluation of Goodness of Fit Measures: Stagnation Periods

# In[153]:


#Reading in excel-files with results from curve fitting, that were created manually
Flush_exp = (pd.read_excel('Curve_fitting/Flush_curve_fitting.xlsx',
index_col=0, header=0, sheet_name = 0, parse_dates = [0]))
Flush_poly = (pd.read_excel('Curve_fitting/Flush_curve_fitting.xlsx',
index_col=0, header=0, sheet_name = 1, parse_dates = [0]))
Flush_lin = (pd.read_excel('Curve_fitting/Flush_curve_fitting.xlsx',
index_col=0, header=0, sheet_name = 2, parse_dates = [0]))


# In[154]:
```

```
Shower_exp = (pd.read_excel('Curve_fitting/Shower_curve_fitting.xlsx',
index_col=0, header=0, sheet_name = 0, parse_dates = [0]))
Shower_poly = (pd.read_excel('Curve_fitting/Shower_curve_fitting.xlsx',
index_col=0, header=0, sheet_name = 1, parse_dates = [0]))
Shower_lin = (pd.read_excel('Curve_fitting/Shower_curve_fitting.xlsx',
index_col=0, header=0, sheet_name = 2, parse_dates = [0]))


# In[155]:


Stag_exp = (pd.read_excel('Curve_fitting/Stag_curve_fitting.xlsx',
index_col=0, header=0, sheet_name = 0, parse_dates = [0]))
Stag_poly = (pd.read_excel('Curve_fitting/Stag_curve_fitting.xlsx',
index_col=0, header=0, sheet_name = 1, parse_dates = [0]))
Stag_lin = (pd.read_excel('Curve_fitting/Stag_curve_fitting.xlsx',
index_col=0, header=0, sheet_name = 2, parse_dates = [0]))


# ## Example of Identification of Biofilm Growth: 2AC, Phase 2
#

# In[96]:


fig, ax = plt.subplots(figsize = (40,20))
sns.set(font_scale=2, style="whitegrid")
ax.grid(which='minor', linestyle=':', linewidth='0.5', color='grey')


re_2AC_2 = resample_phases(dfm_2AC_2, '3H')
re_2AC_2.plot(ax=ax, label = '2AC Phase 2', fontsize = 30, color = 'blue', linestyle = '-')

#Note that for the lines below: the end date deviates by a day compared to extraction,
#this is because of how lines are assigned, e.g. '2021-09-19' will show as a line at the very beginning
#of '2021-09-19', however for data extraction, it includes all data up until '2021-09-20'.

plt.axvline('2021-09-29', linestyle = '--', color = 'black', label = 'Biofilm Growth 2AC 2.3')
plt.axvline('2021-10-04', linestyle = '--', color = 'black')

plt.ylabel('Potential' ' ' r'[$mV$]', fontsize = 40)
plt.legend(fontsize = 35)
plt.xlabel('Time [Days]', fontsize = 35)
```



NTNU

Norwegian University of
Science and Technology