Multiparametric Water Quality Sensor: Characterization, Calibration, and Calibration Transfer

Thesis submitted in partial fulfilment

of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

Submitted by

BALAKUMARA VIGNESH M

(Reg. No. 186107106)

DEPARTMENT OF CHEMICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY, GUWAHATI

GUWAHATI-781039, ASSAM, INDIA

April 2024

This thesis is dedicated to

"Amma & Appa"

"If a man will begin with certainties, he shall end in doubts; but if he will be content to begin with doubts, he shall end in certainties."

- Francis Bacon

Indian Institute of Technology Guwahati
Guwahati-781039, Assam, India
Department of Chemical Engineering

DECLARATION

I hereby certify that the work presented in this thesis entitled **"Multiparametric Water Quality Sensor: Characterization, Calibration, and Calibration Transfer"** is the outcome of my original research work carried out at the Department of Chemical Engineering, Indian Institute of Technology Guwahati, under the supervision of **Prof. Senthilmurugan Subbiah** and **Dr. Bérengère Lebental**. The results documented in this thesis are not submitted to any other university or institute for the award of any degree or diploma. Due acknowledgment has been made wherever the work described is based on the findings of investigations of others with supporting references.

Date: 01 Dec 2024

Place: Guwahati, India

Balakumara Vignesh M
Roll No. 186107106
Department of Chemical Engineering
Indian Institute of Technology Guwahati
India

Indian Institute of Technology Guwahati

Guwahati-781039, Assam, India

Department of Chemical Engineering

CERTIFICATE

This is to certify that the work contained in this thesis entitled **"Multiparametric Water Quality Sensor: Characterization, Calibration and Calibration Transfer"** is being submitted by **Balakumara Vignesh M (Roll No. 186107106)** for the award of Ph.D. degree, is a record of bonafide original research carried out by him at Department of Chemical Engineering, Indian Institute of Technology Guwahati, under our guidance and supervision. This work embodied in this thesis has not been submitted to any other University or Institute for the award of any other degree or diploma.

Date: 01-Dec-2024

Place: Guwahati

Prof. Senthilmurugan Subbiah
Professor,
Department of Chemical Engineering,
Indian Institute of Technology Guwahati,
India

Dr. Bérengère Lebental
Research Director, COSYS/IMSE
Université Gustave Eiffel,
Champs-sur-Marne,
France

ACKNOWLEDGMENT

I am delighted to express my heartfelt thanks to everyone who has supported me on my doctoral journey.

First, I would like to thank my thesis supervisors, Prof. Senthilmurugan Subbiah and Dr. Bérengère Lebental, for their incredible support and patience during my Ph.D. I'm thankful to them for providing me with an excellent lab, many resources, and a great working environment. I'm incredibly grateful for their constant encouragement and motivation, which helped me stay focused on my research goals during the crests and troughs of this project work.

I would like to thank my doctoral committee members, **Prof. Pallab Ghosh**, **Prof. Prabirkumar Saha**, **Dr. Rashmi Dutta Baruah, Dr. Prithwijit Guha,** and **Dr. Reshmi Suresh,** for their kind words, constant encouragement, and constructive inputs during my progress seminars and my synopsis seminar which helped shape my thesis to its present form.

I am grateful to Prof. Kaustubha Mohanty, HoD, Department of Chemical Engineering, and Prof. Anugrah Singh, HoD (Former) Department of Chemical Engineering, for providing me with all the facilities during the course of my work. I would also like to thank the Department of Chemical Engineering staff for their valuable assistance.

I'm extremely grateful to Dr. Roberto Linguerri, MCF, HDR, and Dr. Aghiad Khadour, Directeur de Recherche from Département Composants et Systèmes, Instrumentation Modélisation Simulation et Expérimentations – IMSE at Université Gustave Eiffel, France and members of the Laboratoire: Dr. Rachida Chakir, Dr. Stéphane Laporte, Dr. Guillaume Perrin, Dr. Anne Raus, Mr. Yan Ulanowski, Dr. Julien Waeytens for their immense support, for a conducive environment and the warm welcoming for an international student, when I was doing my research work there.

I want to take this opportunity to thank all my colleagues from the Water-and-Energy Nexus Lab research group and Laboratoire Instrumentation, Modélisation, Simulation et Expérimentation (IMSE). I would like to thank Dr. Vishal K. Verma, Dr. Vigneshwaran K., Dr. Habtom Teklu, Dr. Arunkumar Chandrasekaran, Dr. Aanisha Akhtar, Dr. Viswanth Ramba, Dr. Nivedhitha S., Dr. Muniraja Tippa, Dr. Surendhar G., Dr. Senthil Selva, Mr. Munubarthi Kranthi K., Dr. Ananya Bardhan, Dr. Dinesh K. Gautam, Dr. Priyamjeet, Ms. Neelam Dutta, Ms. Seema, Mr. Shanmugam, Mr. Bijoyendra, Mr. Venkatesh, Mr. Sai Mukesh Reddy, Mr. Bedanta, Mr. Chibibharath, and Mr. Dhivakar for their constant support, laughter, and

motivations. I would also like to thank Mr. Koushik Ghosh, Mr. Kombiah Pandian, and Mr. Ramanan for their consistent help during the PhD journey. I would also like to thank our lab assistants, Mr. Banajit Saloi, Mr. Rupam, and Mr. Bishnu Sharma, for their timely assistance.

I would also like to thank my colleagues from laboratoire Instrumentation, Modélisation, Simulation et Expérimentation (IMSE) Mr. Nicholas Aubert, Mr. Hadi Nasser, Ms. Marine Dumon, Mr. Leshan Usgodaarachchi, Ms. Ayad Tilleli, and Mr. Tsubasa Hamada.

I would also like to express my gratitude to the Indian Institute of Technology Guwahati, India, and Université Gustave Eiffel, France, for the opportunity to do research work. This work has been part of the Indo-European project LOTUS (Project No: DST/IMRCD/India EU/LOTUS/208/(G) and EU Grant No 820881). I would like to convey my sincere thanks to the LOTUS project team for the technical assistance provided.

I am grateful to my seniors and friends, Dr. Sumita, Ashwin, Sathish R, Sathishraja, S.R. Vignesh, Jeeva, Vishwa, Indrajeet, Laxmi, Sameran, Sarat, Aishwarya, Biswajit, Vinita, Sanhita, Remisha and Akanksha for their constant love, support, and motivations. I would also like to thank my strong supporters, Mohammed Azharuhdeen, and Vijayamanikandan, for cheering me. I would like to sincerely thank all my dear friends and well-wishers for their constant support and encouragement.

Finally, as special thanks, I dedicate this thesis to my parents, Vijayalakshmi M, & Muppidathi G, and my brother, Sakthi Rajan M, for their sacrifice, support, and belief in me.

(Balakumara Vignesh M)

Department of Chemical Engineering,

Indian Institute of Technology Guwahati,

India

ABSTRACT

Ensuring the supply of clean and safe water through the distribution networks necessitates effective water disinfection and precise monitoring of key water quality parameters like pH, Active Chlorine, conductivity, and Total Dissolved Solids (TDS) adhering to World Health Organization (WHO) and national standards. Sustainable Development Goal 6 underlines the global effort to supply citizens with safe drinking water, motivating the development of a multi-parametric water quality sensor in this thesis. The sensor aims to monitor water quality and can also aid in regulating disinfectant levels to maintain a safe water supply.

The aim of this thesis is to develop a multiparametric water quality sensor that uses Carbon Nanotubes (CNT) based ink to measure multiple parameters. Developed as a part of an Indo-EU Horizon 2020, project LOTUS (LOw-cost innovative Technology for water quality monitoring and water resources management for Urban and rural water Systems in India), the LOTUS sensor, integrates a CNT ink-based sensor head, an Analog Front End (AFE), and IoT communication solutions for real-time monitoring. In addition to developing the sensor, a bypass structure has been devised to protect the LOTUS sensor in the network from suspended solids and to ensure controlled water flow for accurate measurements of pH, conductivity, temperature, and active chlorine; facilitating easy maintenance and installation.

The developed LOTUS sensor goes through several stages of testing, from small-scale lab experiments to a controlled water loop and, finally, a more extensive water loop similar to a water distribution network. Based on the outcomes, the sensor was improved from a silicon version to a plastic version that is ready for field use to meet the Indian market's needs for real-time performance and cost-effectiveness. The silicon-based LOTUS sensor demonstrated precision in measuring active chlorine within ±0.13 mg/l, pH within ±0.12 units, conductivity within ±70 µS/cm, and temperature within ±0.2°C, covering ranges from 0 to 5 mg/l for chlorine, pH from 7.6 to 8.3, conductivity from 3000 to 5000 µS/cm, and temperature from 17.5°C to 20°C. The results obtained from the cost-effective plastic version – LOTUS PCB (LPCB) showed accuracy in predicting active chlorine within ±0.05 mg/l, pH within ±0.22 units, conductivity within ±68 µS/cm, and temperature within ±0.7°C, spanning ranges from 0 to 5 mg/l for chlorine, pH from 7.2 to 8.2, conductivity from 200 to 2100 µS/cm, and temperature from 10°C to 40°C. However, the LPCB's performance has not yet met the market requirement specifications due to the challenges in the ink deposition process, where the drop-casting printer with a 0.25 mm diameter nozzle could not deposit a consistent volume of the

CNT ink uniformly on the chemiresistor IDEs. To further this complication, the mapping software used in DMD100 could not maintain the gap between the nozzle and substrate when the substrate was not uniformly flat, leading to scratches in the IDEs while depositing the ink. These issues caused baseline resistance variations, negatively impacting the sensor performance. On the other hand, the silicon version was printed using a dot matrix printer, which was comparatively more precise.

The results discussed in this thesis are based on testing only 2 LPCB sensors over 130 days, indicating that further improvements in fabrication and additional testing with more sensors are needed to establish the generalized performance of LPCBs. Nonetheless, the preliminary results highlight the potential of an affordable multiparametric sensor for the real-time monitoring of active chlorine and pH in field conditions. Furthermore, lifecycle testing of the LOTUS sensor suggests that it could sustain functionality beyond five months without significant sensitivity loss despite minimal sensor drift. Developing robust models supported by extensive experimental data could mitigate these anticipated challenges.

This thesis also proposes a calibration transfer technique-based methodology for the mass production of multiparametric sensors. This methodology is assessed with 3 multisensor datasets, which cover the transfer of calibration model among different units in the same and different environments and the same unit after significant drift. From the assessment of 10 transfer techniques in these datasets, recalibration and direct standardization were found to be optimal transfer techniques, and the selection between them depends on the features of the dataset, the structure of the calibration model chosen and the availability of transfer standards. The same techniques can be applied to the LOTUS sensor.

This thesis successfully developed and designed an optimized framework and a lab-tested prototype for a low-cost multiparametric water quality sensor suitable for the Indian market.

***Keywords**: Carbon nanotubes; chemiresistors; multiparametric sensor; water quality; calibration transfer;*

ABBREVIATIONS

ADC	Analog Digital Converter
AFE	Analog Front-end
AMC	Arithmetic Mean Correction
API	Application Programming Interface
APP	Average Peak-to-Peak
BDD	Boron Doped Diamond
BIS	Bureau of Indian Standards
BOD	Biological Oxygen Demand
CAPEX	Capital Expenditures
CNT	Carbon Nanotube
CO	Carbon Monoxide
COD	Chemical Oxygen Demand
DAQ	Data Acquisition
DO	Dissolved Oxygen
DPD	N, N-Diethyl-P-Phenylenediamine
DS	Direct Standardization
EM	Electromagnetic
EU	European Union
FET	Field Effect Transistor
FFUR	Polyfluorene polymer
FIA	Flow Injection Analysis
FIR	Finite Impulse Response
FS	Full Scale

GEP	Gaussian Error Propagation
HPLC	High-Performance Liquid Chromatography
IEC	International Electrotechnical Commission
IITG	Indian Institute of Technology Guwahati
ISFET	Ion-Sensitive Field-Effect Transistor
KNN	K-nearest Neighbours
LOC	Lab-on-Chip
LOD	Limit of Detection
LOTUS	LOw-cost innovative Technology for water quality monitoring and water resources management for Urban and rural water Systems in India
LPCB	LOTUS PCB
LR	Linear Regression
MAE	Mean Absolute Error
MAFE	Monarch AFE
MAPE	Mean Absolute Percentage Error
MMC	Multiplicative mean correction
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
MOX	Metal Oxide
MSC	Multiplicative Signal Correction
MSE	Mean Squared Error
MWCNT	Multiwalled CNT
NRV	No Return Valve
ORP	Oxidation - Reduction Potential
OSC	Orthogonal Signal Correction
PCA	Principal Component Analysis

PCAT	Phenyl-Capped Aniline Tetramer
PCB	Printed Circuit Board
PCNT	Pristine CNT
PEDOT: PSS	Poly(3,4-ethylene dioxythiophene): Poly(styrene sulfonate)
PI	Polyimide
PLS	Partial Least Squares
PT	Procrustes Transform
PVC	Polyvinyl Chloride
PVT118	Polyvinyltriazole Polymer
RF	Radio Frequency
RFID	Radio Frequency Identification
RH	Relative Humidity
RMS	Root Mean Square
RMSE	Root Mean Squared Error
RO	Reverse Osmosis
SAR	Successive-Approximation Resistor
SBC	Slope Bias Correction
SCSS	Sense-City Supervision System
SDG	Sustainable Development Goal
SH	Sensor Housing
SSS	Single Sensor Standardization
SVR	Support Vector Regression
SWCNT	Single-Walled Carbon Nanotubes
TAFE	Terramatics AFE
TDS	Total Dissolved Solids

TRL	Technology Readiness Level
TS	Transfer Standards
TSS	Total Suspended Solids
UAFE	Uni AFE
UART	Universal Asynchronous Receiver / Transmitter
UN	United Nations
USB	Universal Serial Bus
UV	Ultraviolet Rays
WDN	Water Distribution Network
WHO	World Health Organization
WQM	Water Quality Monitoring

Table of Contents

List of figures

List of Tables

1

Introduction

Abstract: *This chapter provides a comprehensive review of the importance and obstacles of monitoring water quality, especially in the context of developing countries, and explores the current market of water quality sensors, highlighting how multiparametric sensors address existing shortcomings. It also introduces the LOTUS project, which forms an integral part of this thesis work.*

1 Introduction

1.1 Overview of Water Quality Monitoring

Monitoring water quality is essential for safeguarding the health and welfare of communities, particularly in countries like India, where water shortage and pollution pose significant challenges. Access to safe drinking water is a basic human necessity, and continuous water quality monitoring offers vital information for effective water resource management and timely pollution control.

The 2021 UN-Water report [1] stresses the urgent requirement to enhance water quality monitoring to attain sustainable development objectives, especially in developing nations. It emphasizes the obstacles posed by insufficient water quality, which can hinder endeavors from enhancing public health, preserving biodiversity, and promoting sustainable economic advancement. The report highlights the significance of real-time water quality monitoring systems in swiftly pinpointing sources of pollution and facilitating prompt actions to avert the degradation of water sources. This up-to-the-minute data is crucial for more efficient management of water resources, guaranteeing drinking water safety, and aiding in decision-making processes concerning water governance and policy formulation.

Furthermore, the report emphasizes the need for increased investment in innovative technologies and infrastructure to facilitate the implementation of real-time water quality monitoring in developing areas. It also highlights the disparity in monitoring capabilities between developed and developing nations, as depicted in Figure 1.1. The proportion marked in grey lack adequate real-time data for classifying water quality, with many falling into the developing category; with 51% in Asian-African countries. This emphasizes the importance of international collaboration in sharing knowledge, technology, and financial resources. The goal is to equip developing countries with the necessary tools and expertise to effectively monitor water quality, safeguarding human health and the environment. The UN-Water 2021 report ultimately calls for a unified global endeavor to prioritize water quality, proposing that real-time monitoring could revolutionize water protection as a valuable resource for future generations.

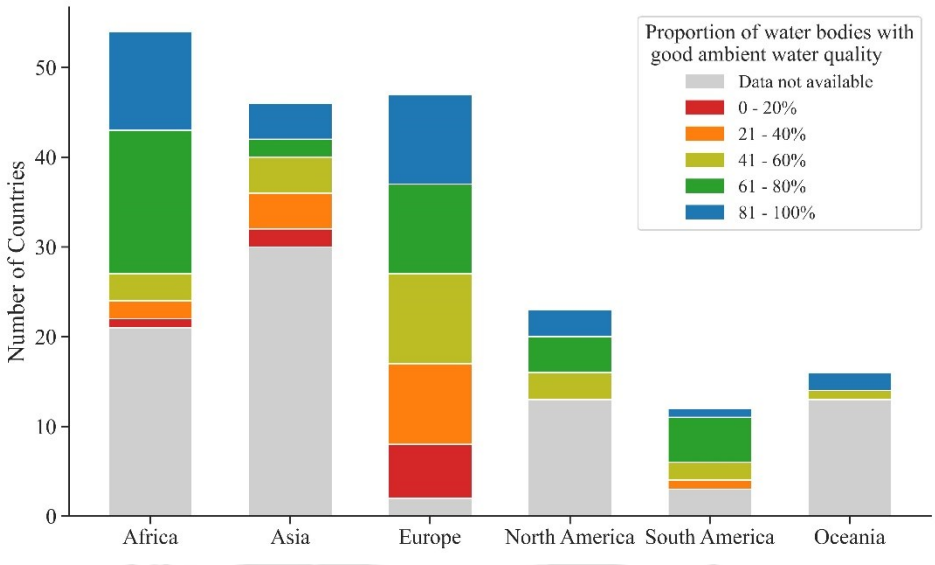

Figure 1.1: Proportion of water bodies with good ambient water quality classified. Based on data provided by UNEP, GEMStat, downloaded from the UN-Water SDG 6 Data Portal (https://sdg6data.org/).

The United Nations' Sustainable Development Goal 6 (SDG 6) [2] is a global initiative to ensure the availability and sustainable management of water and sanitation for all by 2030. This goal encompasses various targets, including improving water quality by reducing pollution, eliminating dumping and minimizing the release of hazardous chemicals, increasing water recycling and safe reuse globally, and significantly expanding international cooperation and support for water- and sanitation-related activities and programs in developing countries. SDG 6 focuses not only on the physical availability of water but also on its quality, the efficiency of its use, and the sustainability in water resource management to prevent water scarcity and contamination.

1.2 Challenges in Water Quality Monitoring in India as a Developing Country

Providing safe drinking water in developing nations continues to be a significant hurdle. Despite good intentions, efforts to offer government-subsidized drinking water can unintentionally lead to inadequate maintenance and operational efficiency of utilities [3]. Financial restrictions limit these utilities from operating, as they have minimal resources due to the absence of regular income. Though capital expenditures (CAPEX) were initially covered by government funding, the utilities later had to generate revenue. Also, the reduced

price of water due to government subsidies hinders their capacity to generate reliable income, which is essential for enhancing water quality and treatment plant facilities [4], [5].

Urban areas in developing nations are expanding rapidly, often outpacing the development of essential infrastructure. This unplanned expansion leads to reliance on untreated groundwater or tanker water supplies due to the absence of a centralized water distribution network. In contrast, rural areas frequently lack any water supply system, although initiatives like India's Jal Jeevan Mission Scheme [6] are beginning to address this gap. Utility operators have limited access to advanced technology due to affordability constraints. The primary sources for these emerging rural water schemes are local water bodies or, more commonly, groundwater. The operation of these systems often relies on manual control by unskilled [7], [8].

The Jal Jeevan Mission [6], launched by the Indian government in August 2019, aligns with SDG 6's objectives and seeks to provide safe and adequate drinking water through individual household tap connections to all rural households in India by 2024. The mission's strategy involves creating a community-driven approach to water management, focusing on the sustainability of water sources, and implementing robust water supply systems to ensure potable water for every rural home. The Jal Jeevan Mission aims to build a resilient water management ecosystem capable of withstanding climatic and demographic challenges by emphasizing the management of local water resources and promoting water conservation practices for rural India.

Under this circumstance, the challenge of ensuring safe drinking water in developing countries, particularly India, is exacerbated by a complex paradox. On one side, governments provide subsidized water, yet this water often fails to meet drinking quality standards. Conversely, households face the financial burden of securing safe drinking water, incurring exorbitant costs. The price of procuring or producing potable water at the household level can reach around one rupee per litre, and it is significantly higher than the cost of drinking water supplied in WDN by developed nations (for example, in the EU water price varies between 1 – 10 euros per 1000 L [9]). This dichotomy highlights a critical inefficiency in the current water supply model, where the potential for government-provided, quality water could significantly alleviate household expenses dedicated to water procurement [10], [11].

Adding to the complexity is the dependence on household water purifiers and storage tanks due to sporadic water supply. Families store water, which, despite being treated prior to

storage, faces the risk of contamination before being consumed, requiring additional treatment [12], [13]. This recurring process of treating and storing water compromises its quality and imposes an extra financial strain on households. The extra layer of filtration/treatment done at household levels does not always match the required standards [14].

The solution lies in a radical transformation of the water supply infrastructure. The government can achieve multiple benefits by establishing a continuous (24x7) supply of drinking-quality water through the pipe network [15]. First, it would eliminate the need for household-level water purification and storage practices, directly reducing the financial strain on families. Second, discontinuing localized water treatment operations could significantly reduce carbon emissions associated with filter manufacturing and operation. Third, a reliable and quality water supply system could lead to increased revenue for the government, ensuring the sustainability and maintenance of the water supply infrastructure [16], [17].

To achieve this, deploying an appropriate water quality monitoring solution equipped with advanced water quality sensors throughout the distribution network is paramount, as visualized in Figure 1.2. This approach guarantees safe drinking water delivery and positions the government to realize considerable economic benefits through increased revenue and system stability.

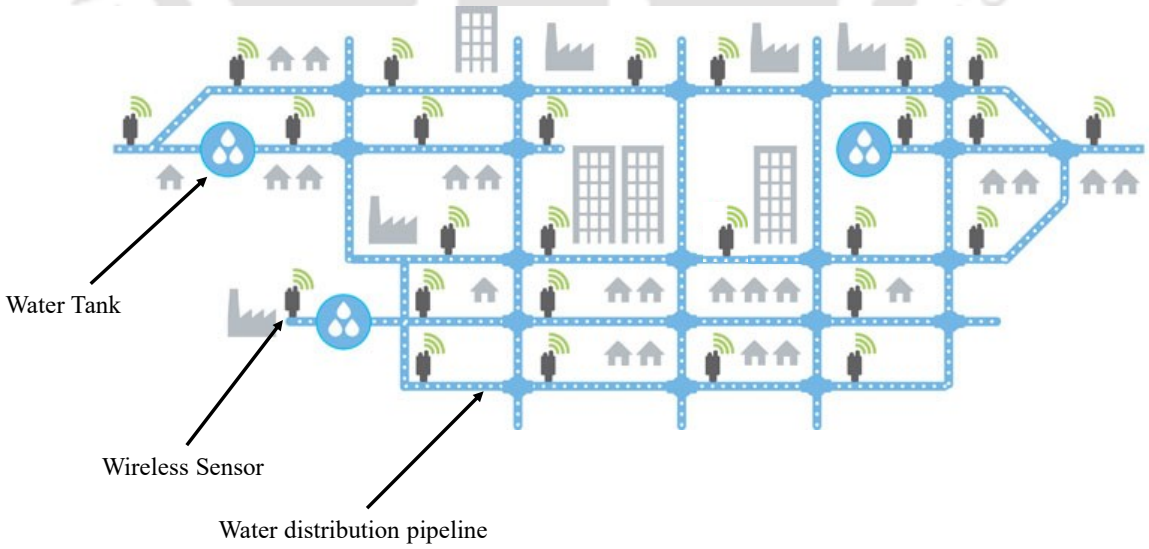

Figure 1.2: Schematic of a water distribution pipeline network populated with wireless sensors to monitor the water quality. Adapted from ref [18].

One significant and cost-determining factor in deploying optimal water quality monitoring solutions is the availability of reliable, affordable, and scalable water quality sensors [19], specifically in developing nations like India [20].

The table below (Table 1.1) outlines the different aspects in the scheme of sustainable development goals 6 of the UN, and that of the Jal Jeevan mission, which is on a regional scale.

Table 1.1 Outline of different aspects between UN SDG 6 and Jal Jeevan Mission

	Sustainable Development Goals	**Jal Jeevan Mission**
Objective	To monitor and improve the global water quality by assessing ambient water quality parameters	To provide safe and adequate drinking water to all rural households ensuring water quality and regularity
Scope	Global, focusing on broad range of water bodies, across multiple countries and regions	National, specifically rural households and public institutions of India, focusing on piped water supply
Parameters monitored	Ambient water quality parameters; chemical, physical and biological parameters	Water quantity, regularity of water supply, potability (including chlorine, pH, and TDS)
Data Collection	Data submissions from multiple countries	Structured and manual sampling approach, supported by surveys
Improvements	Enhanced global engagement with better data submission. Capacity building in low-GDP countries for better data quality.	Expansion of IoT-based monitoring and increased coverage of water quality testing infrastructure. Increased local manpower training.
Coverage	Increased participation from Central, Southern and Western Asia countries. Broader inclusion of diverse water body types.	Focus on ensuring functional tap connection and regular monitoring nationwide.

		Coverage of water supply infrastructure, ensuring quality and quantity across diverse conditions.
Benefits	Improved global understanding of water quality trends. Strengthened ability to identify and mitigate global water quality issues.	Significant reduction in waterborne diseases through improved access to safe drinking water. Empowerment of local communities, especially women, in managing and monitoring water resources.

1.3 Current State of Water Quality Sensors

Water quality sensors are essential for ensuring the safety and reliability of water for various purposes, such as drinking water and industrial processes like cement fuel production. These sensors are responsible for monitoring important water quality parameters, offering vital data to uphold specific standards for different applications. This thesis explores critical elements of water quality sensors, such as their creation and validation, specifically emphasizing essential parameters for potable water, the sensor technologies employed, their advantages and disadvantages, and the economic impacts of their implementation, especially in developing countries. Additionally, it investigates the potential of lab-on-chip technology in transforming water quality monitoring by enabling cost-effective and efficient multi-parameter sensing.

In the context of drinking water, the main parameters of focus include Total Dissolved Solids (TDS), pH levels, and active chlorine concentration [21]. These parameters are crucial for ensuring the safety and taste quality of drinking water.

TDS: Represents the total concentration of dissolved substances in water. High levels of TDS can affect the taste of water and may indicate the presence of harmful minerals or chemicals.

pH: A measure of the acidity or alkalinity of water. A balanced pH level is crucial for preventing corrosion in water pipes and maintaining water's taste. An overview of the evolution of pH measurement systems is depicted in Figure 1.3.

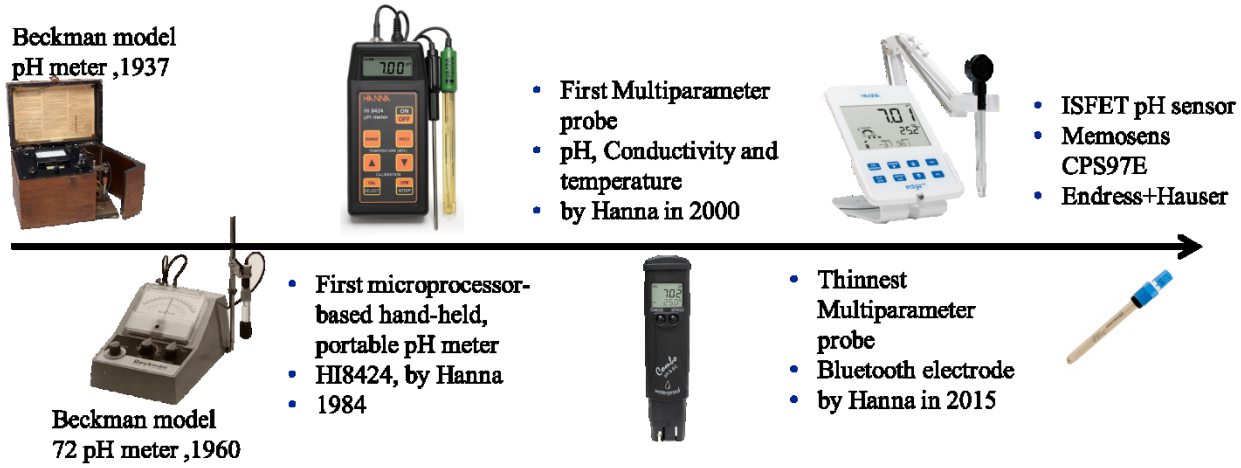

Figure 1.3: An Overview of the evolution of pH sensors

Active Chlorine: Used as a disinfectant, its concentration is vital for eliminating harmful pathogens while ensuring the water does not acquire an unpleasant taste or odor. The most common detection methodologies employed in commercial sensors are either DPD-based colorimetry or amperometry (Figure 1.4).

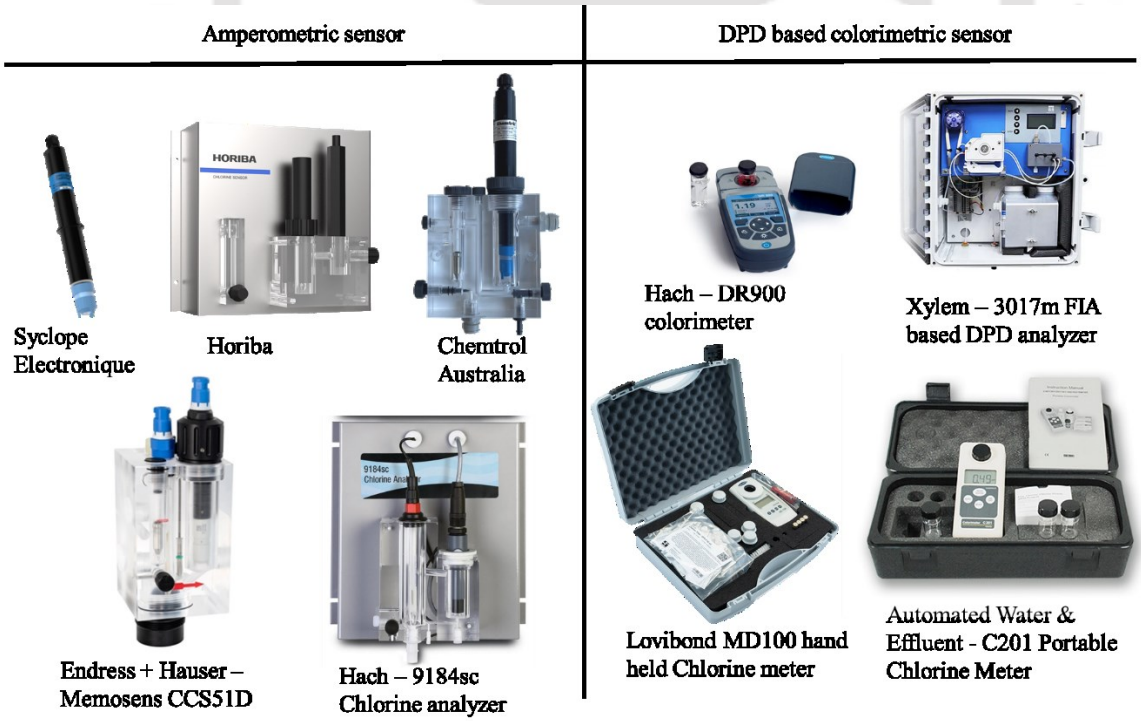

Figure 1.4: Most common commercial Chlorine Sensors for Water Quality

Several sensing technologies are available for measuring these parameters, especially pH, active chlorine, and TDS, each with advantages and limitations. They are discussed in detail in Chapter 2, but the numerous sensor types can be generally grouped as follows:

Electrochemical sensors: These are the most conventional type of sensors and are widely used for pH and chlorine measurements. They offer high sensitivity and specificity but may require regular calibration and maintenance.

Optical sensors: These are employed chiefly for TDS and some types of chlorine sensors. These sensors are known for their long life and minimal maintenance but can be more costly upfront.

Electrical sensors: These form a relatively newer group of sensors (compared to the traditional methods – electrochemical and optical). They are comparatively more affordable and require less maintenance than their matured competitors.

The water quality sensing technologies vary in terms of sensor lifespan, installation convenience, and real-time sensing capacity. Immediate monitoring is essential to quickly identify any deviations from safe levels, enabling timely corrective measures.

The debate between using individual sensors for each parameter versus a single multiparametric sensor encompasses considerations of deployment, cost, and benefits. Single-parameter sensors can offer high specificity and accuracy but may increase complexity and cost in systems requiring multiple parameters monitoring. Conversely, multiparametric sensors simplify installation and maintenance but might compromise on sensitivity and specificity for individual parameters [22], [23].

The cost of sensors is a significant factor, especially in the context of deploying water quality monitoring systems in developing countries. High initial costs and maintenance expenses can be prohibitive, limiting the widespread adoption of advanced water quality monitoring systems.

The advent of lab-on-chip technology presents a promising avenue for reducing the cost and complexity of water quality monitoring. By integrating multiple sensors on a single chip, this technology offers a compact, efficient, and cost-effective solution for multi-parameter analysis. Coupled with advancements in the Internet of Things (IoT), low-cost mobile processors, and electronic circuits, it is now feasible to develop low-cost, high-performance water quality monitoring systems [24]. These systems can transmit real-time data for immediate analysis, ensuring the continuous supply of safe drinking water [25], [26].

The need for robust, efficient, and cost-effective water quality monitoring systems is more critical than ever, especially in the face of growing water scarcity and pollution challenges. Advances in sensor technology, particularly through lab-on-chip and IoT innovations, are paving the way for affordable and comprehensive water quality monitoring solutions. These advancements are vital for ensuring the accessibility of safe drinking water worldwide, highlighting the importance of continued research and development in this field [27].

1.4 Classification of Water Quality Sensors

Water quality sensors can be categorized according to different criteria [28], [29], including the parameters they monitor, their intended uses, and their technological methods (Figure 1.5). Here is a primary classification:

Based on Measured Parameters

Physical sensors measure the physical characteristics of water, such as temperature, conductivity, and turbidity. Chemical sensors detect particular chemical attributes like pH, dissolved oxygen, and levels of pollutants or nutrients. Biological sensors keep track of biological markers such as chlorophyll, blue-green algae, and bacterial contamination.

Based on Application

Monitoring of Drinking Water: Sensors created to evaluate the quality of water in treatment facilities and distribution networks

Environmental Monitoring: Utilized in rivers, lakes, and seas to analyze ecological well-being and levels of pollution

Monitoring Industrial Processes: Employed to oversee water quality in industrial operations and discharges.

Based on the Transduction method

Electrochemical Sensors

Electrochemical sensors operate based on the principle of electrochemistry, which involves chemical reactions producing electrical signals. These sensors typically consist of electrodes (conductors) that interact with specific ions or gases dissolved in water. The interaction causes a measurable change in electrical properties, such as voltage or current, which is directly related to the concentration of the target analyte. There are several types of electrochemical sensors, including:

Ion-selective electrodes (ISEs): Designed to measure the concentration of a specific ion in water, such as nitrate, fluoride, or potassium. They work by having a selective membrane that only allows the target ion to pass through, causing a potential difference that can be measured.

Amperometric sensors: Often used to measure dissolved oxygen levels. They operate by applying a voltage between two electrodes and measuring the current that flows as a result of the oxygen reduction reaction on the cathode.

Potentiometric sensors: Measure the voltage between two electrodes without passing a current. This method is commonly used for pH measurements, where the sensor measures the hydrogen ion activity in water.

Optical Sensors

Optical sensors measure the interaction between light and the water sample to deduce the concentration of certain parameters. These interactions can include absorption, fluorescence, and scattering of light, each of which provides information about different water quality parameters:

Absorption-based sensors: Measure the amount of light absorbed by dissolved substances or suspended particles. This method is widely used for measuring concentrations of organic compounds, as many absorb light at specific wavelengths.

Fluorescence sensors: Detect the fluorescence emitted by certain substances when they are excited by light of a specific wavelength. This technique is beneficial for detecting chlorophyll, blue-green algae, and certain organic compounds.

Turbidity sensors: Utilize light scattering to measure the cloudiness or haziness of water, which indicates the presence of suspended particles. The sensor shines a light into the water and measures the intensity of light scattered by the particles.

Electronic Sensors

There sensors act as passive electronic elements like a resistor, capacitor, or transistor. On exposure to the measurand, these sensors change their nature (for instance, the conductivity). This change in nature is measured by exciting the sensor with a signal and acquiring the output (for instance, potential drop across the resistive device after exciting the device with a signal). Based on the passive element they are made to represent, they are called chemitransistors [30] or chemiresistors [31].

Other types

Mass Sensors: Typically based on piezoelectric materials, they detect the change in mass to provide information on the target analyte in the chemical reaction. Thermal Sensors use the heat generated by the chemical reaction to provide analytical information about the reaction.

These sensors are typically thermistor-based and detect changes in temperature as a change in the resistance of the sensing material. Acoustic Sensors, or sonar sensors, use sound waves to measure various properties of water. These sensors emit sound pulses and analyze the echo received back, which changes based on the characteristics of the water and its contents, such as sediment concentration, where the intensity of the sound reflected to the sensor can indicate the concentration of sediments in the water.

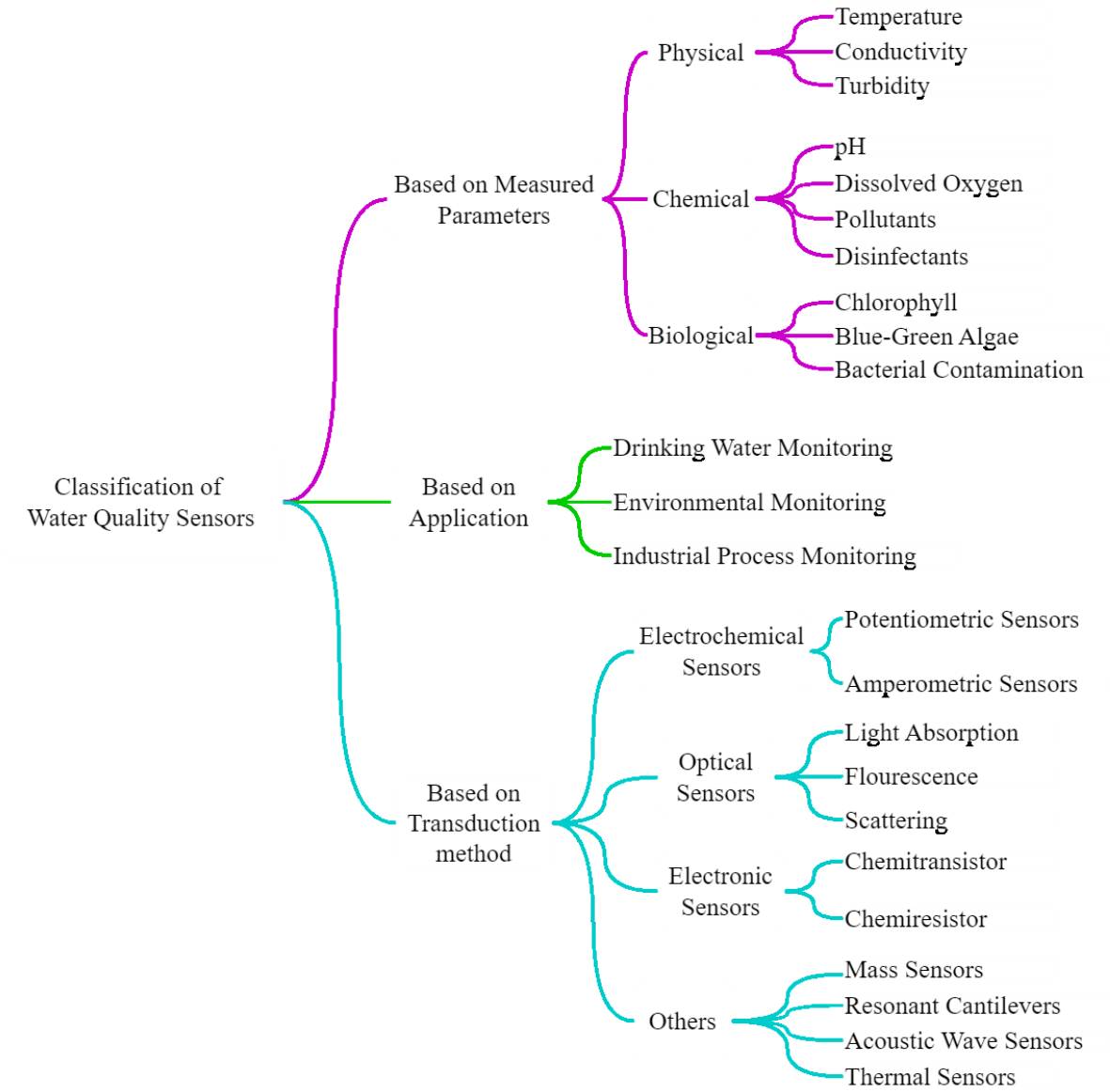

Figure 1.5: Classification of Multiparametric Water Quality Sensors

Each of these sensor technologies offers unique advantages and is suited to specific applications within the broad field of water quality monitoring.

Apart from these, there are several other classification schemes for water quality sensors considering sensor structure – macro vs micro, measurement parameters – monoparametric (targeting only one parameter) vs multiparametric (single sensor array measuring multiple parameters of interest), response data dimensionality – single (for monoparametric sensors) vs. multi-dimensional (for multiparametric sensors), technological sophistication – conventional vs. advanced, data integration and analysis – standalone vs. integrated networks, application focus – generic vs. targeted contaminant detection, and operational environment – in-situ vs. laboratory-based monitoring. This nuanced classification enables tailored sensor selection based on specific monitoring needs and environmental conditions.

1.5 Overview of Multiparametric Water Quality Sensors

Multiparametric water quality sensors represent a rapidly evolving technology that provides comprehensive and real-time monitoring of various water quality parameters. These technologies are crucial for identifying pollutants, evaluating the efficiency of water treatment processes, and ensuring adherence to environmental standards. This overview outlines multiparametric sensors' primary classifications, integration methods, technological foundations, and applications, drawing from extensive research and the literature discussed in Chapter 2.

Integrated Single Parameter Sensing Technologies

This category combines individual single-parameter sensors into a cohesive system. Each sensor measures water quality parameters, such as pH, turbidity, chemical oxygen demand (COD), or heavy metal concentrations. The sensors work in tandem, allowing for the simultaneous assessment of multiple parameters. A unified data acquisition system collects, processes, and displays the data, enhancing monitoring efficiency and accuracy by utilizing specialized sensors for each parameter [32]. The only downside is the bulky footprint due to the integration of multiple sensors, as visualized in Figure 1.6.

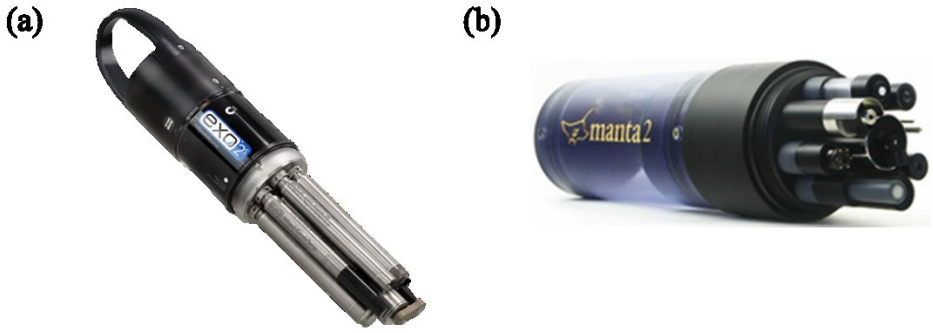

Figure 1.6: Industrial Integrated Multiparametric Water Quality Sensors. (a) YSI Exo Multiparameter Sonde, (b) Eureka Manta2 Water Quality Recorder

Lab-on-Chip (LOC) Technology

LOC technology signifies a move towards the miniaturization and integration of multiparametric sensing on a single chip [33], [34]. It combines multiple sensor functions on a compact platform, using microfluidics to channel small water volumes across different sensing areas. This approach is noted for its innovative chemical detection methods and the measurement of physical properties like temperature.

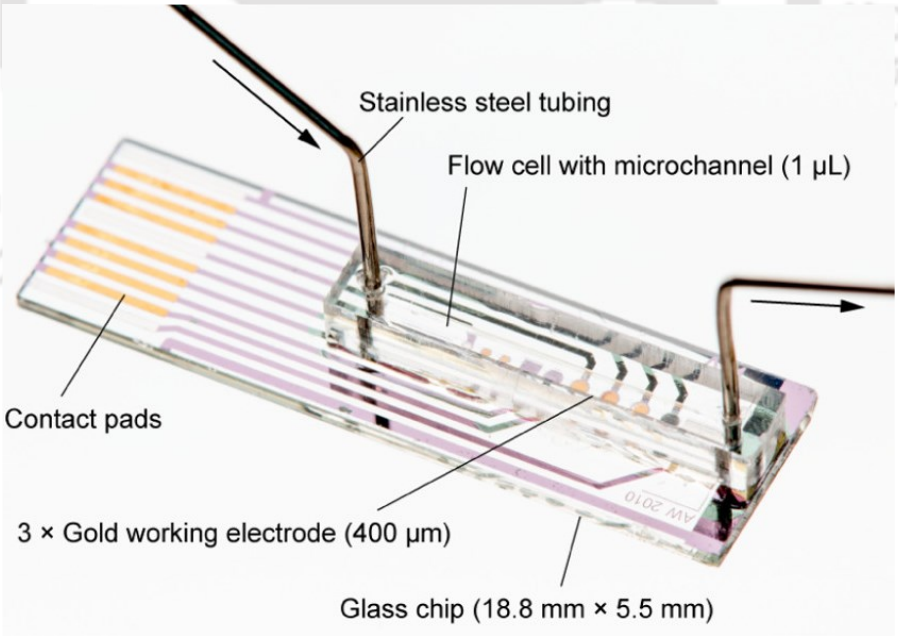

Figure 1.7: Microfluidic sensor with a flow cell and gold working electrodes. Reprinted (adapted) with permission from [35]. Copyright 2021 American Chemical Society.

Multiparametric sensors are designed to simultaneously monitor various water quality parameters, integrating different sensing elements into one device [36]. This integration

facilitates real-time, continuous monitoring of water bodies, crucial for pollution detection, ecosystem health assessment, and water resource management. Advances in sensor technology, data analytics, and wireless communication have significantly improved the precision, reliability, and accessibility of water quality data.

The integration of these sensors into water quality monitoring systems allows for real-time, continuous assessment of water bodies, which is essential for detecting pollution events, understanding ecosystem health, and managing water resources effectively. Moreover, advancements in sensor technology, data analytics, and wireless communication have enhanced the precision, reliability, and accessibility of water quality data.

1.6 The LOTUS Project: Aims and Progress

The LOTUS project [37], funded by the European Union's Horizon 2020 Research and Innovation Programme and the Indian Government, aims to address India's water and sanitation challenges through innovative, low-cost sensing solutions developed in partnership between EU and Indian organizations. Project LOTUS, elaborated as LOw-cost innovative Technology for water quality monitoring and water resources management for Urban and rural water Systems in India, focuses on creating a multiparametric sensor-based system to improve water quality monitoring in both rural and urban India, featuring tailor-made decision support and a co-creative approach for design and production.

The sensor technology developed under LOTUS [38] will support the early detection of water quality issues and optimize water management with continuous improvements through real-world testing. The LOTUS sensor incorporates a multifaceted approach to water quality monitoring, leveraging carbon nanotube (CNT) based sensing technology for measuring pH and Active chlorine concentration alongside temperature and conductivity measurements [39], [40], [41].

Given this comprehensive functionality, the LOTUS sensor aligns with several types in the classification of multiparametric water quality sensors [28], [29]:

1. Based on Measured Parameters:

Chemical Sensors: They can detect specific chemical parameters like pH, free chlorine, and Arsenic, leveraging the unique properties of CNTs and specific polymer functionalization for target chemicals.

Physical Sensors: The sensor's functionality to measure temperature and conductivity also places it within this category, offering insights into the physical characteristics of water that are essential for quality assessment.

2. Based on Application:

Environmental Monitoring: Given its target functionality to detect a range of parameters critical to water quality assessment, the sensor would be well-suited for monitoring natural water bodies, contributing to studies on ecological health and pollution levels.

Drinking Water Monitoring and Industrial Process Monitoring: Its capability to measure parameters critical to water safety and quality makes it relevant for both drinking water assessment and monitoring water in industrial processes.

3. Based on the Transduction method:

Electronic Sensors: The fundamental principle of the LOTUS sensor, particularly in its use of CNTs for detecting chemical parameters [39], classifies it as an electronic sensor. This technology relies on the interaction between the sensing material (CNT) and the analyte to produce a measurable change (conductivity of the sensing material) indicative of the analyte's concentration.

In conclusion, the LOTUS sensor embodies characteristics of both chemical and physical sensors, leveraging electronic principles for its operation. Its design and functionality indicate its applicability across environmental monitoring, drinking water quality assessment, and industrial process oversight, making it a versatile tool in multiparametric water quality sensors.

1.7 Objectives of the Thesis

The brief discussion above highlights that multiparametric sensing technology for monitoring water quality presents numerous advantages over traditional methods, which typically rely on multiple single-parameter sensors or offline laboratory analysis. Given that the primary goal of the LOTUS project is the creation of a multiparametric sensor, this focus aligns perfectly with the scope of the current thesis.

Within the LOTUS project's framework, the Université Gustave Eiffel team is responsible for developing and manufacturing sensor chips based on carbon nanotubes to measure pH, active chlorine, temperature, and conductivity.

The core objective of this thesis is to create a comprehensive, integrated sensor platform designed explicitly for monitoring water quality, utilizing these sensor chips. From this perspective, the following research objectives were formulated:

1. Development of the Sensor Platform: Achieving full integration and demonstrating consistent functionality of a multiparametric water quality sensor that employs Carbon Nanotubes (CNTs).
2. Development of the Calibration Model: Formulating experimental procedures and developing calibration models for the CNT-based multiparametric sensor to optimize its long-term performance.
3. Optimization and Streamlining of a Mass Calibration Process: Employing Calibration Transfer Techniques on Multiparametric sensors to facilitate efficient and rapid calibration processes on a large scale.

1.8 Outline of the Thesis

Chapter 1: Introduction: This chapter offers a brief examination of the significance and challenges associated with water quality monitoring, particularly within the framework of developing nations. Subsequently, it delves into a brief discussion on the existing landscape of water quality sensors in the market and describes how multiparametric sensors could tackle the prevailing deficiencies. Additionally, this section also provides a brief introduction to the LOTUS project.

Chapter 2: Literature Review and Research Objectives: This chapter aims to review the literature on several vital areas thoroughly: the historical evolution of water quality monitoring technologies, the progress and challenges of monoparameter sensors for chlorine and pH, the state of multiparametric sensing technologies for water quality, and detailed considerations on calibrating these sensors, including the use of calibration transfer techniques. Within the context of the LOTUS project's multiparametric water quality sensor, significant research gaps have been identified that have informed the objectives of the current study, with these insights discussed towards the end of the chapter.

Chapter 3: Sensor Platform Development: This chapter presents the end-to-end integration of the LOTUS sensor from the sensor chip to the final data visualization platform. It also explains the development, fabrication, and assembly of the LOTUS multiparametric water quality sensor. This chapter also describes the evolution of the data acquisition (DAQ) system developed for activating and acquiring data from the LOTUS sensor and the optimized methodology to use with the DAQ system to improve the reliability of sensor response.

Chapter 4: Sensor Characterization and Calibration: This chapter details the experiments carried out on the sensors and the sensor characterization. This chapter also elaborates on the calibration model development for target parameters: Temperature, Conductivity, Active chlorine, and pH. The comprehensive analysis and outcomes derived from the experiments conducted are also presented at the conclusion of this chapter. The results showing the performance of CNT-based sensors on par with those reported in the literature and much better than those for real-time deployment are also discussed in this chapter.

Chapter 5: Calibration Transfer: This chapter discusses the need for calibration transfer and highlights the calibration transfer techniques that can be applied to LOTUS sensors. This chapter assesses the 10 most common calibration transfer techniques discussed in the literature by applying them to 3 different gas sensor datasets, each with a different scenario invalidating the initial laboratory calibration. Based on the results, the calibration transfer technique applicable to the LOTUS sensor is discussed.

Chapter 6: Conclusion and Scope of Future Work: This chapter outlines the key conclusions drawn from the dissertation work. This chapter concludes by highlighting potential areas for future exploration in the context of current work.

2

Literature Review

Abstract: *This chapter aims to conduct an in-depth review of water quality monitoring technologies, examining the progress and limitations of monoparameter sensors for chlorine and pH detection, including calibration practices, and exploring the state-of-the-art multiparametric sensing technologies along with their calibration and calibration transfer techniques. Based on the findings of the literature survey and within the context of the LOTUS multiparametric water quality sensor, significant research gaps were identified, which served as the basis for establishing the objectives of the current research work.*

2 Literature Review

2.1 Evolution of Drinking Water Quality Monitoring

The evolution of drinking water quality monitoring has transitioned significantly from traditional methods to advanced sensor-based technologies, marking a profound shift in how water quality is assessed and managed. Initially, water quality monitoring relied heavily on offline sample collection followed by laboratory analysis, a time-consuming and labour-intensive process [42], [43]. This method, while accurate, posed limitations in terms of the frequency and speed of data collection, making it challenging to detect and respond to water quality issues in real time [44]. Maintenance and equipment requirements are also expensive, especially for developing nations, where large-scale monitoring is mandatory for significant impact. These highlight the inefficiency of traditional methods for water quality monitoring [19], [45].

The development of in situ sensors began to change this landscape, offering the ability to continuously monitor water quality parameters such as pH, turbidity, dissolved oxygen, and various chemical contaminants directly at the source. However, these advancements were initially limited by the need for manual data retrieval and the lack of integration into broader monitoring networks. The introduction of Internet of Things (IoT) technologies has significantly enhanced sensor-based monitoring, enabling real-time data collection, analysis, and remote access to water quality information [46], [47]. This integration has facilitated a more dynamic approach to water quality management, allowing immediate responses to potential contamination events and providing a comprehensive overview of water quality trends over time [48], [49].

The water quality parameter can be grouped under physical (color, taste, odour, temperature, conductivity, turbidity, etc.), chemical (pH, active chlorine, dissolved oxygen, COD, BOD, presence of heavy metals, etc.), and biological parameters (algae, bacteria, viruses etc.). World Health Organization (WHO) has developed guidelines, with the latest version being the fourth edition of 2022, that summarize the health implications and provide the standards and permissible limits of water quality parameters [50].

The primary parameters for water quality across regions and countries mostly align with WHO standards, with few adjustments based on water availability and specific usage scenarios [51], [52]. Temperature, conductivity, and turbidity are among the most significant physical parameters of potable water quality, each providing critical information on the

physical and chemical characteristics of water bodies [50], [53], [54], [55], [56]. Temperature is a fundamental parameter that influences the solubility and reaction rates of chemicals in water, affecting biological activity and the ecosystem's health. It is a primary indicator of thermal pollution and can signal water quality changes affecting aquatic life. Conductivity, measuring the water's ability to conduct electrical current, is directly related to the concentration of ionized substances or total dissolved solids (TDS) in the water. This parameter is essential for assessing the overall ionic content of water, indicating the presence of dissolved salts, which can be pollutants or vital nutrients. The Bureau of Indian Standards (BIS) has set an acceptable limit of TDS to be 500 mg/L and a permissible limit of 2000 mg/L (if an alternate drinking water source is available) [57]. Turbidity, which measures the clarity of water, can indicate the presence of suspended particles, including industrial discharge, sediment, or microbial growth, affecting the aesthetic and safety aspects of water. It is also directly related to the presence of microorganisms. For potable water, 5 nephelometric turbidity units (NTU) is the threshold set by the WHO as well as the BIS.

On the other hand, several chemical parameters are recommended by the WHO, starting with pH, hardness, biological oxygen demand (BOD), chemical oxygen demand (COD), total suspended solids (TSS), dissolved oxygen (DO), residual disinfectant, i.e., active chlorine, presence of other elements, contaminants, ions, heavy metals, drugs, and other pollutants, etc. WHO and BIS standards (in the Indian context) provide a maximum permissible limit for each of these (and several other) contaminants commonly found in drinking water. Within the spectrum of chemical parameters, both pH and chlorine stand out as critical indicators due to their profound implications for both environmental health and human safety. pH, a measure of the acidity or alkalinity of water, is pivotal in regulating the solubility and biological availability of chemical compounds, including metals and nutrients. Its value affects the toxicity of pollutants and microbes [58] and aquatic life health, making it a foundational though indirect parameter for water quality assessment [59], [60]. Active chlorine is a primary disinfectant that plays a crucial role in controlling microbial contamination in drinking water [61], [62], [63]. However, monitoring its levels is vital to ensure they remain within safe limits to prevent harmful by-products. The balance between adequate disinfection and minimizing chemical hazards emphasizes the significance of chlorine as a water quality parameter.

2.2 State of the Art on Monoparametric Sensors for pH and Active Chlorine

As discussed in Chapter 1, Water Quality sensors can be generally classified as Optical, Electrochemical, and Electrical (and others) based on their transduction method. The significant differences among the 3 types are outlined in Table 2.1.

Table 2.1: Difference between the 3 types of Water Quality Sensors

	Optical	**Electrochemical**	**Electrical**
Operation	Relies on interaction with photons and changes in the electronic structure of receptor molecules.	Involves a current flow between working, reference, and counter electrodes in an electrolytic solution.	Measures changes in conductivity or resistance directly due to analyte interaction.
Types	Absorbance, fluorescence, luminescence, light scattering, colorimetry.	Potentiometry, voltammetry, amperometry, electrochemical impedance spectroscopy.	Chemiresistors, Gated field-effect devices (ChemFETs), Gate-less field-effect devices.
Commercial availability	Widely available for both personal and professional use.	Commonly available, especially for parameters like pH, dissolved oxygen, and ORP.	Emerging, with some applications in gas phase sensing and portable devices.
Typical use case	Personal care, healthcare applications (e.g., pregnancy tests,	Water quality monitoring, Environmental monitoring, industrial	Water quality monitoring, gas sensing, and environmental

	diabetes monitoring), water quality testing.	processes, and healthcare diagnostics.	monitoring are potentially in smart textiles and wearable devices.
Cost	Varies; disposable indicator strips are low-cost, but precision sensors and optoelectronic devices can be expensive.	Moderate; requires reference and counter electrodes but can be miniaturized.	Potentially lower due to simpler geometry and no need for reference electrodes; cost-effective for large-scale deployment.
Advantage	Can measure multiple parameters simultaneously; easy to deploy and use by non-experts.	Highly specific and sensitive; it can measure a wide range of analytes directly involved in electrochemical reactions.	Simple and robust; does not require a liquid electrolyte or reference electrode, allowing for simpler sensor designs.
Disadvantage	Potential interference from deposits, turbidity, or other colored species may require visual interpretation.	Needs a stable reference electrode and is prone to maintenance and frequent calibration.	Requires direct contact with the analyte; some types may have increased complexity or require more expensive electronics.

Water quality monitoring has traditionally relied on colorimetric (optical) and electrochemical sensors or large laboratory instruments, each with limitations such as the need for reagents or reference electrodes [28]. Chemiresistors, in contrast, offer a reagent-free, low-maintenance alternative that does not require reference electrodes, overcoming the common drawbacks of existing technologies [64]. However, chemiresistors are not without their challenges, which include adapting them to aqueous environments, like preventing

electrical shorts through the water and maintaining low sensing voltages to avoid unwanted electrochemical reactions or causing water splitting, as most literature discusses their application in gas sensing. Nevertheless, their potential for simple and economical fabrication and the affordable cost of operation and maintenance position them as a promising solution for continuous, online water quality monitoring. Though much of the sensing literature focuses on ChemFETs (the most common type of electrical transduction category), where an applied gate voltage modulates the active layer's conductivity to elucidate sensing mechanisms, most of the sensing is performed at zero gate voltage [65] effectively simplifies a ChemFET to function as a chemiresistor.

2.2.1 Monoparametric Sensors for pH

The pH value of an aqueous solution is calculated as the negative decimal logarithm of the hydronium ion concentration (-log [H3O+]). The drinking water pH usually falls within the range of 6.5–9. The pH also influences the disinfection capability of disinfectants (most having better disinfection capability at lower pH), increases the corrosion rate of metal pipelines at acidic pH levels, and affects the ecosystem of aquatic organisms [51]. The predominant methods for measuring pH in water quality monitoring traditionally rely on pH meters equipped with electrodes initially developed as glass electrodes by F. Haber and Z. Klemensiewicz in 1909 [66]. These meters were conceptualized by Arnold Beckman in 1934 and brought to the market in 1936 [67]. Over the subsequent decades, various pH sensors have been developed, falling into categories such as chemo-mechanical, electrochemical, and optical sensors, with glass-membrane-based electrochemical electrodes remaining the most widely utilized even in contemporary applications [68]. Despite their affordability, these instruments require frequent maintenance and calibration, and are fragile, thereby rendering them expensive to maintain in continuous real-time monitoring.

In the literature, the pH sensors (those without glass electrodes) can be grouped into different categories based on the transduction mechanism like optical [69], electrochemical [70], [71], voltammetric [72], impedimetric [73], diode-based [74], potentiometric, ISFETs (Ion-Sensitive Field-Effect Transistor) [71], [75] and chemiresistive [28], [76] based sensors. A standard potentiometric pH sensor comprises a two-electrode structure, with one electrode designated as the sensing electrode and the other as the reference electrode, wherein Ag/AgCl is the most frequently used in micro-scale pH sensors [77], [78], [79]. Compared to slightly complex design structures like optical, diode-based, or ISFET, these sensors are

simple in configuration and provide a response in the mV range. They are widely used in conjunction with a commercial voltmeter on a laboratory scale.

The Ion-Sensitive Field-Effect Transistor (ISFET) was introduced by Bergveld [71] in 1970 as a metal-oxide-semiconductor field-effect transistor (MOSFET) with its gate physically separated from the chip by the solution under examination. In this configuration, a reference electrode replaces the gate metal, and the dielectric layer significantly influences sensing performance [80], [81], [82]. The pH variation can be determined by recording the threshold voltage or drain current changes for thin-film transistor-based ISFETs and Ion-Sensitive Potentiometric Transistors (ISPETs). These devices, commonly fabricated on doped silicon substrates, utilize insulation materials such as metal oxides and semiconductor-based ceramics for the pH-sensitive layer.

Based on a chemiresistor, a pH sensor is a two-terminal device featuring a pH-sensitive material deposited between two electrodes. The sensing material's intrinsic electrical resistance/conductance undergoes alterations upon exposure to solutions with varying pH values, as illustrated in Figure 2.1. Notably, the chemo-resistor has a straightforward physical structure and does not necessitate a reference electrode, making it conducive to sensor miniaturization. A suitable voltmeter/DAQ system measures this change in resistance and correlates it to the sample's pH.

Lee et al. (2009) [83] developed a chemo-resistor with Pd-resistors and gold electrodes, wherein, based on the pH level, the concentration of hydrogen gas produced varies, which in turn influences the formation of Pd-H hydride. This increases the resistance, which is measured to estimate the pH level. Similarly, Lee et al. (2014) [84] used TiO_2 nanowires at the surface of Ti/C nanofibers, wherein the presence of H3O+, the depletion layer reduces, giving rise to a change in conductivity.

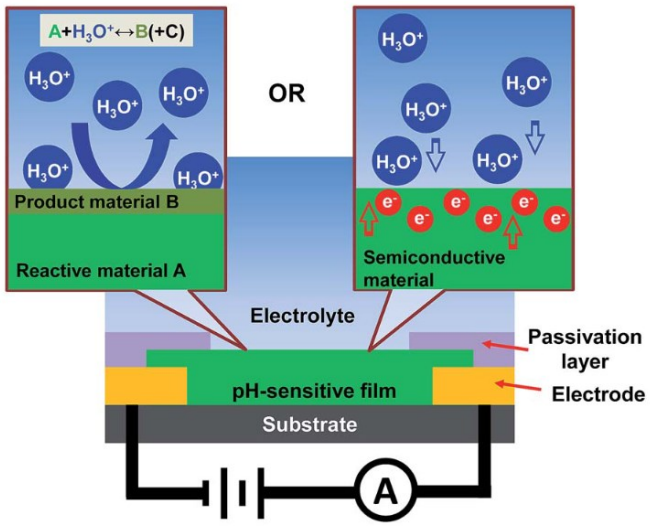

Figure 2.1: Outline of a Chemiresistive pH sensor and its sensing mechanisms.

Source: Qin et al. 2015 [85].

Kwon et al. (2006) [86] demonstrated the spray method for fabricating SWCNT-based pH sensors, displaying amperometric responses in buffer solutions across various hydroxyl ion concentrations. This sensor exhibited lower sensitivity for pH below 6 than above (3.4% /pH for pH < 6 and 9.3%/pH for pH > 6) and was active throughout the range of 1 to 11 pH units. However, the testing was done by dropping the pH buffer solutions on top of the sensor rather than inserting the sensor into a water beaker or stream.

Li et al. (2011) [87] developed a pH sensor with dielectrophoresis-aligned SWNTs, which showed linear sensitivity with a minimal change in the sensitivity over 10 days of testing. They report a sensitivity of 0.253 %/pH units over the range of 5 – 9pH, with the response time varying between 2 to 24 seconds depending on the pH level (higher pH, longer response time). The authors also discuss the problem of developing identical sensors with the same resistance between batches. The problems, like depositing an exact amount of ink onto the sensor each time to match the quantity of SWCNTs is not easily achievable, and the uncertainty of the film composition adds to the differences among the sensors. The authors included sensor response normalization into the data preprocessing to alleviate these issues. However, this adds to another uncertainty as in the literature, different authors use slightly different formulae to normalize the resistance, and choosing a consistent baseline in real-time is especially difficult when the baseline has to be varied each time. However, this negates the effect of batch-batch differences as well as sensor drift to a great extent.

Lei et al. (2011) [88] reported the fabrication and characterization of graphene-based chemiresistors as pH sensors (Figure 2.2). The sensor was tested with pH buffers between 4 and 9, and it showed linear sensitivity of -0.157 %/pH units and minimal hysteresis of ~2.3% for the average slope of 3.48 kΩ/pH, though the sensitivity is lower than that of a 3-terminal field-effect transistor type (FET type) sensor with the same material. As this article focused on fabrication and characterization, the sensors' long-term stability and continuous testing were not discussed. Liao et al. (2011) [89] studied the effect of pH on the one-dimensional nanofibers of SWCNT – polyaniline composites (which shows a linear response between 4 – 12 pH) for developing a gas chemosensor with tuneable conductivity.

Lei et al. (2012) [90] developed a carbon nanotube (CNT) based pH sensor on a filter paper. The authors discuss the potential applications of paper-based microfluidics with CNT-based sensors, which realize more bio-assays. As this development was focused on bioassay, the sensing framework is micro-exposure to the analyte rather than stream immersion as required for water quality monitoring. Jung et al. (2014) [91] synthesized a pH sensor with MWCNT doped with Nickel particles. Though the fabrication process was simple, the pH sensing properties relied heavily on the size of Nickel particles, a complex parameter to maintain during upscaling. On the other hand, Gou et al. (2014) [92] developed an in-vivo pH sensor based on oxidized SWCNT functionalized with a conductive polymer. This sensor was designed as an implantable sensor with a passive RFID tag that transmits the pH data through the skin.

Qin et al. (2016) [93] inkjet printed SWCNTs onto a glass substrate (Figure 2.3) and demonstrated fast response, linear sensitivity (of 48.1 mV/pH), small hysteresis (of 8% of 48.1 mV/pH), and stable pH sensing even after 14 days of inactive storage, highlighting the potential of using SWCNT for pH sensing. Goh et al. (2018) [94] reported fabricating a pH sensor CNT-based ink deposited on silver electrodes. The sensor exhibits outstanding biocompatibility essential for human skin-based applications.

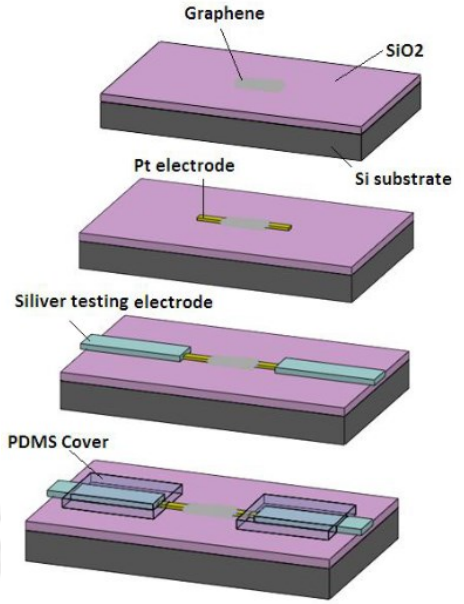

Figure 2.2: Fabrication process of the Graphene-based chemiresistive type pH sensor.
Source: Lei et al. (2011) [88].

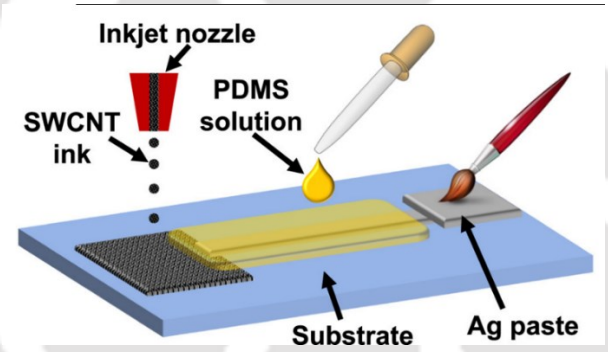

Figure 2.3: Schematic of the fabrication process of a SWCNT-based pH sensor.
Source: Qin et al. (2016) [93]

Jeon et al. (2020) [95] demonstrated the real-time application of a flexible pH sensor developed from nanocomposites of SWCNTs and Nafion, which was added to make the SWCNTs more robust to irreversible physical damage that could be caused by bending. The sensor's sensitivity varied from 0.018 units/pH to 0.035 units/pH for alkaline and 0.019 units/pH to 0.039 units/pH for acidic pH conditions, where the variation depended on the number of layers of ink deposition. The authors demonstrated the sensor's real-time applicability but integrated it with a drone-based water pH monitoring system to monitor pH at any location in the stream. Nevertheless, the authors did not establish the stability and robustness of sensing in continuous water streams. Emami et al. (2021) [96] developed a flexible pH sensor for wound monitoring, using the CNT – Polyaniline nanocomposites as a chemiresistive sensor. They reported linearity in the 2 – 10 pH range, with a sensitivity of

2.72 Ω/pH and a response time of 70 seconds. Though this article does not discuss other applications for the pH sensor, the repeated usage of CNT-based chemiresistors for pH sensing in this as well as in the previous articles, in situations demanding flexibility and affordability as well as reliability, shows the potential of CNT-based chemiresistive sensor for pH sensing application as well as highlights the significant gap in affordable pH sensing in water quality monitoring.

The currently commercially available technologies face notable hurdles related to stability, consistency, and degradation of materials, both physically and chemically. Factors such as the physical breakdown of sensing components due to increased water flow rates and chemical changes induced by substances like chlorine pose significant challenges. Microfabricated pH sensors have emerged as pivotal tools in water quality monitoring, albeit facing considerable challenges that hinder their broad utilization [85], [97], and chemiresistive type sensors offer more flexibility in adoption into water quality monitoring applications due to their simpler fabrication process; affordable fabrication, assembly, maintenance, and DAQ systems; lower operation cost; and easy scalability [64], [98], [99].

Table 2.2: Summary of details of Current pH Sensing Technologies

Types	Market Availability	Sensitivity and Selectivity	Stability and Durability	Cost and Ease of Production	Reference
Glass Electrode pH Meters	Widely Available	High sensitivity and selectivity. It can measure a wide pH range without interference.	High durability but requires careful handling due to fragility. Requires frequent calibration.	Moderate cost. Production is well-established but involves sophisticated glasswork.	[66], [67], [68], [100], [101]
ISFETs	Available	Good sensitivity and selectivity. Less affected by external ions compared to glass electrodes.	Good stability. Performance can degrade over time due to surface fouling.	Higher cost due to semiconductor fabrication processes. Moderate ease of production.	[71], [75], [80], [81], [82], [102]
Optical Sensors	Limited	High sensitivity and good selectivity, especially with specific dyes and indicators.	High. It is not prone to electrical interference and has no direct contact with the sample.	High cost. Production can be complex due to the need for specialized equipment.	[69], [103], [104]
Voltammetric and Impedimetric	Limited	High sensitivity. Selectivity depends on the electrode material and setup.	Good stability, but it can be affected by the buildup of substances on the electrode.	Moderate to high cost. Production complexity depends on electrode material and design.	[72], [73], [105]
Diode-based	Research Stage	Potentially high sensitivity. Selectivity may be enhanced with specific designs.	Early stages of development, so long-term stability is not well-established.	Currently, high cost and complex production due to being in the research and development phase.	[74]
Metal Oxide-based Sensors	Limited	Good sensitivity and selectivity can be achieved with specific metal oxides.	Good, but performance can be influenced by environmental conditions and chemicals.	Moderate cost. Production involves standard materials but requires precise control.	[106], [107]
Chemiresistor type Sensors	Emerging	Good sensitivity and selectivity can be improved by using specific sensing materials	Good stability, and reproducibility, but long-term studies are not well-established.	Moderate cost and simpler fabrication techniques.	[83] - [95]

2.2.2 Monoparametric Sensors for Active Chlorine

Active chlorine (or total chlorine) sensor measures the concentration of either hypochlorous acid (HOCl) or the hypochlorite ion (OCl) (or total chlorine). Colorimetric and electrochemical methods are widely utilized compared to many reported analytical methods for active chlorine detection. Other analytical techniques used for selective application are iodometric titration [108], chromatography [109], chemiluminescence [110], amperometric [111] and voltammetric [112], and potentiometric two-terminal (chemiresistors) and three-terminal (chemFETS).

The colorimetric method-based sensors rely on a chemical reagent's colour change after a reaction with chlorine. In contrast, electrochemical sensors measure the current generated by the redox reaction of chlorine across the electrode. The choice of sensor for a given application often depends on factors such as sensitivity, specificity, operational complexity, ability to measure in real-time installation at the pipeline, and the cost associated with the monitoring process.

The colorimetric method, particularly absorptiometry using DPD (N, N-diethyl-p-phenylenediamine) [113], is favoured for its high-sensitivity application. This method operates on the principle that active chlorine oxidizes DPD to produce a magenta compound whose intensity indicates chlorine concentration and is photometrically measurable. The DPD-based method is the standard analytical technique approved by the United States Environmental Protection Agency for chlorine monitoring [56]. This reagent-based methodology is typically confined to manual operations or laboratory settings due to its reliance on reagents, which also introduces specificity issues in the presence of multiple redox-active species. Methods like titration (iodometric or amperometric) and chemiluminescence are unsuitable for continuous monitoring, and unaffordable when scaling up, because of the use of specialized instruments.

Electrochemical methods offer a more straightforward approach, eliminating the need for reagents and generating direct electrical signals that support autonomous and continuous monitoring. However, electrochemical sensors require frequent calibration to compensate for flow rate variations and electrode aging. The significant size of the complete system of a conventional electrochemical sensor deters their use in online water quality monitoring

solutions despite their reliability and widespread use in laboratory conditions. This prompts the need for more compact and integrated solutions.

However, electrochemical methods that eliminate a few of these problems and improve the pre-existing solutions have been actively researched for decades. Starting with, Pletcher et al. (1991) [114] studied the response of platinum and gold electrodes to changes in chlorine concentration in an aqueous phosphate buffer. This study demonstrated an excellent linear relationship for lower concentrations of chlorine with respect to output current. Still, extensive testing for deployment in water quality analysis was not discussed, while Berg et al. (1993) [115] fabricated a 3-electrode silicon-based active chlorine sensor for water quality monitoring. This sensor utilized a chloridized silver electrode as the silver electrode, and the sensor did not suffer from the interference of dissolved oxygen. However, the authors did not provide information on the effect of pH or TDS on the sensor response. They also did not discuss continuous monitoring of chlorine in flowing water.

Kodera et al. (2005) [112] demonstrated a free chlorine sensor based on voltammetry using platinum, gold, and glass carbon electrodes having detection capability in agreement with iodometric titration with a range of 4 – 400 mg/l, while Campo et al. (2005) [111] improved the free chlorine sensing by replacing platinum electrodes with gold electrodes. In this study, the authors also demonstrated 7 days of continuous free chlorine measurement in tap water, wherein the gold electrode performed better than the platinum electrode. They demonstrated the functionality within the limited range of 0 to 0.5 mg/l but with a detection limit of 0.22 (±0.09) mg/l without electrochemical cleaning and 0.08 (± 0.08) mg/l with electrochemical cleaning of the sensor.

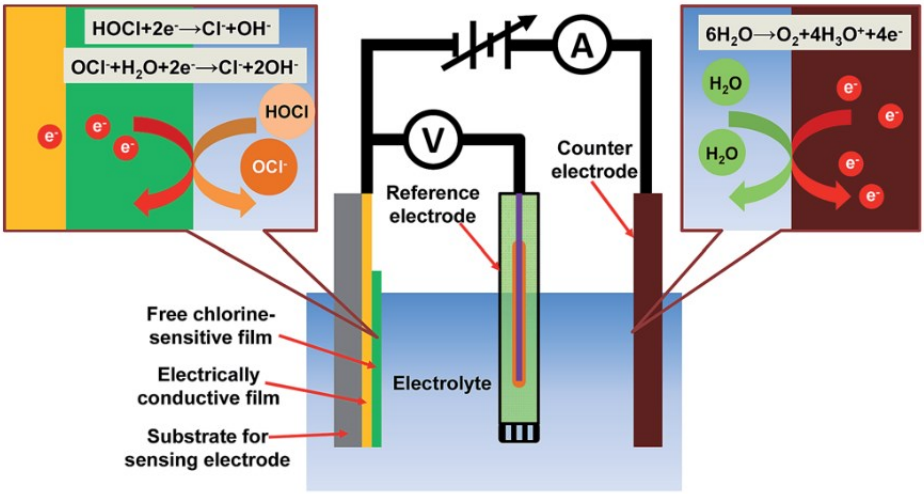

Figure 2.4: Schematic of Electrochemical (Amperometric type) Active Chlorine Sensor and its working principle. Source: Qin et al. 2015 [85].

Pan et al. (2015) [116] devised a low-cost free chlorine sensor using pencil lead as the electrode, using ammonium carbamate to modify its surface electrochemically. Though the faster response time and negligible hysteresis prove the feasibility of this sensor for real-time monitoring, the authors did not provide an extensive analysis of its long-term stability.

Qin et al. (2016) [117] introduced a paper-based hand-drawn electrochemical free chlorine sensor with stability, reusability, and wider sensing range. The chemiresistive type sensor had poly(3,4-ethylene dioxythiophene): poly(styrene sulfonate) (PEDOT: PSS) as the conductive ink that is deposited within silver electrodes. The authors identified a variation of less than 15% in sensitivity after 30 days of storage, leading to the question of long-term storage contributing to the degradation of the sensor. Also, the authors did not discuss continuous monitoring of flowing water. Davis et al. (2007) [118] developed a disposable sensor to determine free chlorine and total chlorine using microelectrodes rapidly. Though they provide a faster response time and resolution of 0.02 mg/l within a usable range of 0 – 20 mg/l, this sensor cannot be used for continuous chlorine concentration monitoring as they are developed for single-use applications.

Yen et al. (2020) [119] improved the PEDOT: PSS inked paper-based sensor developed by Qin et al. (2016) [117] by doping the PEDOT: PSS with graphene. This nanohybrid sensor demonstrated a LOD of 0.18 mg/l in the 0.1 to 500 mg/l range. The authors also established the real-time application of the sensor by interfacing it with an electrical readout that provides the chlorine concentration to a smartphone via Bluetooth. Although this provides a low-cost, compact solution, extended testing on its long-term stability was not discussed,

and the methodology is not suitable for continuous real-time monitoring of chlorine concentration.

Murata et al. (2008) [120] employed Boron-doped diamond (BDD) electrodes alongside voltammetry to quantitatively assess active chlorine, revealing a nearly linear correlation with chlorine concentration. However, the high costs of sensor production and the data acquisition system, the necessity for a pretreatment process, and the limited research supporting this method hinder its application for continuous real-time monitoring of chlorine levels. Yang et al. (2009) [121] demonstrated the potential for an electronic-grade CNT-based active chlorine sensor in an integrated microfluidic chamber. The sensor exhibits a detection limit lower than 5ppb, with a longer response time (in minutes) and even longer recovery time (in hours) at room temperature. This acts as a bottleneck when deploying this sensor for real-time chlorine monitoring applications. Also, the authors did not provide information on the sensor's sensitivity to the water sample's pH.

Olive-Monllau et al. (2010) [122] demonstrated the use of composite electrodes with MWCNTs within a Flow Injection Analysis (FIA) system for low-level free chlorine detection in various water samples, including swimming pool and tap water, maintaining stability over 30 days. However, the widespread application for real-time monitoring across water distribution networks is hindered by the cost-effectiveness of the measurement system. This limitation suggests the need for further development to reduce the sensor cost.

Senthilkumar et al. (2014) [123] developed a disposable active chlorine sensor based on a polymelamine-modified screen-printed carbon electrode, which provided a wider linear dynamic range of 0.5 to 300 mg/l and 0.3 mg/l LOD, which is better than using graphite or platinum or BDD or gold microelectrodes as reported in the previous studies. This technology can be used for routine analysis of active chlorine in tap water and swimming pools but cannot be used for continuous monitoring.

Hsu et al. (2015) [63] doped SWCNTs with phenyl-capped aniline tetramer (PCAT) to develop a simple active chlorine sensor (Figure 2.5) with a range of 0.06 to 60mg/l, where the sensor exhibited a linear response for the range of 0.06 to 6 mg/l. The response to chlorine is due to the oxidation of the PCAT, so the authors proposed an electrochemical reset before each subsequent measurement. The authors demonstrated the potential applicability of SWCNT-based sensor for continuous real-time chlorine monitoring by experimenting with

this sensor for 30 hours. The authors also reported the need for further investigation to assess potential interference from other oxidant species; nevertheless, the sensor remains suitable for measuring free chlorine in the water distribution network, where only free chlorine is allowed as the oxidant species.

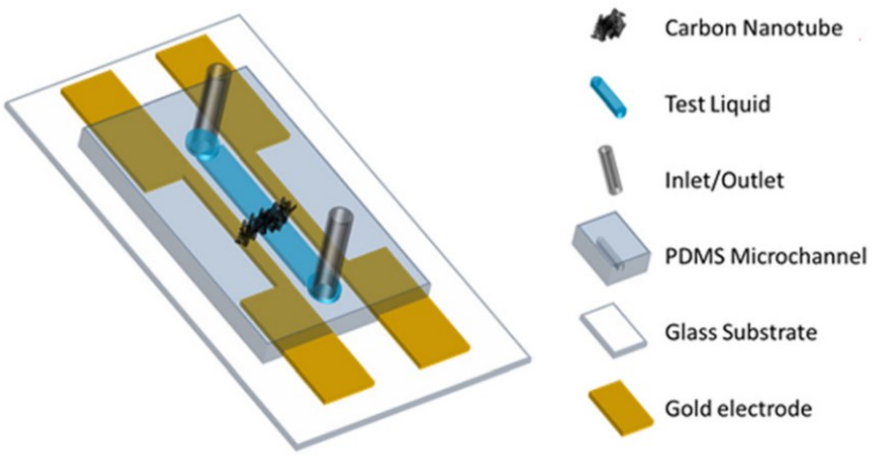

Figure 2.5: Outline of a Chemiresistive Chlorine sensor (PCAT-SWCNT type).
Source: Hsu et al. (2015) [63]

Building on the previous work with PCATs and pencil lines, Mohtasebi et al. (2017) [124] and Hoque et al. (2017) [76] used PCAT-doped pencil lines to measure active chlorine in water, improving on the methodology developed by Hsu et al. [63]. This sensor showcases a broad detection range for aqueous free-chlorine from 0.06 mg/L to 60 mg/L, with a stable and nearly linear response unaffected by common interferences. Its design eliminates the need for electrochemistry or a reference electrode, making it robust, cost-effective, and ideal for applications with limited resources. However, the need for a reset procedure after every measurement (to bring the sensor back to sensing condition) affects its feasibility in continuous chlorine concentration monitoring.

Wilson et al. (2019) [125] elaborately reviews different chlorine sensors, specifically in the context of continuous monitoring in water quality applications. The author argues that compared to reagent-based techniques (colorimetric, luminescent, and fluorescent methods), the electrochemical techniques, specifically the amperometric methods, are ideal chlorine sensors because of being reagent-free, accurate, easy to use, and low maintenance. The review also highlights the issues of current electrochemical sensors (commercial sensors and reported in the literature), like the degradation of precious metal electrodes, microfouling by

organisms, mineralization, and instability in long-term use. The authors conclude by defining an ideal chlorine sensor as one that should have the simplicity and reagent-less operation of electrochemical sensors while being equally robust as the offline analysis. These sensors should be tested in actual distribution network conditions to ensure their stability and robustness. They should also be easily deployable, with minimal maintenance.

Alam et al. (2021) [126] discussed a reagent-free, reusable active chlorine sensor using gold electrodes, which, being a robust and stable sensing material, provided comparatively superior performance in actual samples of tap water and swimming pool water than the graphite-based alternatives. Apart from superior performance, the authors reported immunity to common interferents, low drift, and low pH dependence between pH 7 and 8. However, the sensor measurement was performed with stable water in beaker tests; it was not subjected to continuous water monitoring.

On the other hand, Zubiarrain-Laserna et al. (2022) [127] identified the need for continuous real-time chlorine monitoring and demonstrated a graphene-like carbon film (functionalized with PCAT) based chlorine sensor that operates without any reset procedure for continuous monitoring. Though the sensor range is limited to 0.01 – 1.4 mg/l, the limit of detection of < 1 µg/l and the absence of a resetting process enables the usage of this sensor in continuous chlorine concentration monitoring, with the only caveat being the longer response time.

Siddiqui et al. (2023) [128] extensively reviewed several carbon-based electrochemical active sensors for active chlorine concentration measurement. The authors presented the limited methodologies with affordable fabrication as the major bottleneck in adopting carbon-based chlorine sensors for water quality monitoring.

The simple design of chemiresistive sensors and low-cost materials used in the fabrication highlight chemiresistors-based chlorine sensors as the solution for affordable chlorine monitoring systems. Other concerns, including reversibility, interference, sensor degradation, and fouling, will have to be addressed to fully realize the continuous chlorine monitoring with a chemiresistive sensor instead of offline sampling or electrochemical-based chlorine sensors currently in use.

Table 2.3 Summary of key aspects of each sensor transduction methodology with respect to affordable water quality monitoring

Types	**Affordability**	**Robustness**	**Ease of deployment**	**Real-time monitoring**	**Scalability**
Colorimetric (DPD-based)	●	✖	NA	✖	✖
Iodometric Titration	●	✖	NA	✖	✖
Chemiluminescence	✖	●	✖	●	✖
Amperometry	✖	●	✖	✖	✖
Electrochemical	✖	●	✖	✖	✖
CNT-based electronic sensors	●	●	●	●	●

Table 2.4: Summary of details of Current Active Chlorine Sensing Technologies

Types	Market Availability	Typical range and LOD	Device Cost & Ease of Production	Ease of Installation	Major concerns	Reference
Colorimetric (DPD-based)	Available	0 to 5 ppm, 4 ppb	Low-cost reagents and devices, Easier production	Manual measurement, Increases operational cost	Continuous monitoring is not possible, and higher operational costs	[113], [129], [130], [131], [132]
Iodometric Titration	High (Lab use)	1 to 15 ppm, 150 ppb	Low, primary cost is for reagents	Laboratory setting	Not suitable for real-time infield monitoring	[108]
Chemiluminescence	High	0 to 28 ppm, 18 ppb	Increased cost due to the use of expensive light sources and devices	Requires specialized equipment in most cases	Not suitable for continuous monitoring	[28], [133]
Amperometry	High	0 to 20 ppm, 5 ppb	Moderate to high depended on the materials used for the fabrication of electrodes	Subject to system design for integration, requires test cells, deters compact installation	Known interference with metal ions, oxidants, biofilms, and scale	[111], [122], [134], [135]
Electrochemical (General)	High	0.1 to 400 ppm, 20 ppb	Moderate to high depended on the materials used for the	Subject to system design for integration, requires	Frequent recalibration, high dependence on flow rate, aging of electrodes	[111], [112], [114],

Types	Market Availability	Typical range and LOD	Device Cost & Ease of Production	Ease of Installation	Major concerns	Reference
			fabrication of electrodes	test cells, deters compact installation		[115], [120], [129]
Electrochemical (Graphene-based)	Varied	0.01 to 1.4 ppm, 1 ppb	Low-cost materials, compact design	Often requires microfluidic channels	Long-term stability and continuous monitoring are questionable, limited lifetime of coated membrane	[127]
Electrochemical (paper-based)	Prototype	0.5 to 50 ppm, 500 ppb	Low production cost, compact design	Often requires microfluidic channels	Long-term stability and continuous monitoring are questionable, limited lifetime of coated membrane	[116], [117], [118], [119]
CNT-based electronic sensors	Research	0.05 – 60 ppm, 60 ppb	Low production cost, compact design	Easier than an electrochemical counterpart	Information on long-term stability and continuous monitoring is scarce in the literature	[63], [121], [122], [124]

2.2.3 Focusing on CNT-based chemiresistive sensors

Chemiresistive sensors stand out as the optimal technology for pH and active chlorine detection in water quality analysis due to their fundamental attributes of simplicity, durability, affordability, and cost-effective manufacturing and operational processes [136]. These sensors circumvent the necessity for reference electrodes, a requirement in electrochemical sensors, and present an efficient, economically viable solution as opposed to optical sensors that necessitate complex and expensive setups for accurate readings. Their operation involves direct contact with the analyte, thereby facilitating real-time monitoring. This direct interaction mechanism enables chemiresistive sensors to provide more consistent and reliable measurements for critical water quality parameters like pH and active chlorine.

Carbon Nanotube (CNT) based sensors are increasingly recognized in water quality monitoring, particularly for pH and active chlorine sensing, thanks to their exceptional electrical, chemical, and mechanical properties [137]. Compared to other carbon-based materials like graphene, CNTs exhibit unique advantages, including higher surface area-to-volume ratios and distinct electron transport characteristics, which enhance sensitivity and selectivity towards specific analytes. These properties make CNTs especially adept at detecting subtle changes in water quality through chemiresistive changes. The operating principle for Active chlorine and pH sensing is based on chemo-physisorption and is mostly reversible [97]. For broader water quality monitoring, CNTs' versatility in functionalization allows for detecting a wide range of pollutants beyond pH and active chlorine, including heavy metals and organic compounds, positioning them as comprehensive tools for ensuring water safety and compliance with environmental standards.

Integrating CNTs into chemiresistive sensors significantly enriches the pH and active chlorine monitoring methodology, providing a synergistic approach to water quality assessment [138]. The high sensitivity of CNTs to chemical changes enables the detection of minute concentrations of contaminants, while their rapid response times facilitate immediate water quality assessment. This combination ensures that CNT-based chemiresistive sensors can offer detailed and dynamic insights into water quality, surpassing the capabilities of conventional sensors. Moreover, the ability of CNT sensors to operate effectively in complex water matrices without the need for frequent recalibration or reagent replenishment underscores their practical advantage for continuous monitoring applications.

Despite the promising advancements and detailed literature surrounding CNT-based chemiresistive sensors for water quality monitoring, the transition from research to real-world application faces significant hurdles, primarily due to the scarcity of commercial solutions that utilize CNTs. This scarcity underlines a significant gap between technological innovations and their market realization, necessitating augmented research, development, and commercialization efforts to facilitate the transition from laboratory achievements to effective field deployment.

Compounding these challenges is the lack of long-term stability data and comprehensive monitoring studies under real-world conditions, which are critical for assessing the sensors' performance and reliability. The existing literature primarily focuses on short-term laboratory experiments. So it often falls short of providing essential insights into the sensors' behavior in varying water environments, like their maximum saturation limits for diverse analytes, and the factors affecting their stability and sensitivity like (but not limited to) type and concentration of analytes, the pH and temperature of the water, the functionalization and alignment of CNTs, the electrode material and configuration, and the presence of interfering substances or biofouling. Addressing these gaps through detailed research on the long-term durability, optimal operational parameters, and calibration techniques of CNT-based sensors is crucial for advancing water quality monitoring technologies and eventually unlocking the full potential of CNTs to safeguard public health and the environment.

2.3 State of the Art on Multiparametric Sensing Technologies for water quality monitoring

Water quality monitoring (WQM) involves assessing various parameters like chlorine, pH, temperature, and conductivity to ensure the safety and suitability of water for use. Traditionally, this has been achieved by utilizing individual sensors for each parameter to collate and analyze the data. This method works well for monitoring specific locations or for straightforward applications, offering a direct approach to gathering necessary water quality data.

Vijayakumar et al. (2015) [32] designed a simple water quality monitoring system by combining several individual sensors for measuring pH, temperature, conductivity, turbidity, and dissolved oxygen into a single unit. A Raspberry Pi was used as a core controller to collect data from all these sensors, process them, and upload them to a cloud platform. So, any user with a smartphone could connect to the cloud platform and monitor these parameters in real-time. However, this system was limited to a particular area, and upscaling can become exponentially cumbersome due to too many components being interfaced, which can also exponentially increase the associated costs. Ilie et al. (2017) [139] utilized a commercially available smart water kit (Libelium Smart water kit [140]) to configure a wireless sensor network by integrating 3 smart kits, each having electrodes to measure pH, Oxidation-Reduction Potential (ORP), conductivity, and temperature. This article discussed the advantages and problems faced with integrating multiple electrodes with a sensor node kit and the ease of upscaling facilitated by Libelium. However, using electrodes increased the cost of the unit and the price for upscaling, apart from requiring regular maintenance and calibration.

Several such systems are reported in the literature, focusing on combining individual sensors for each parameter into a single WQM unit. Cui et al. (2018) [141] demonstrated a WQM unit that collected data from sensors for free chlorine, pH, conductivity, and ORP, as is then transmitted wirelessly to a primary station where the real-time values are displayed; He et al. (2021) [142] developed an alarm system that alerts the farmers about the degradation of water quality, estimated as water quality index derived from the measurement of temperature, pH, dissolved oxygen, turbidity, ammonia, nitrogen, and lead. The system integrated electrode-type sensors for each of the above parameters; Okpara et al. (2022) [143] provided a detailed review of several such related works in the literature where the improvement is done to upscale the WQM units over a water distribution network via

wireless sensor networks. However, all the WQM units are based on a combination of several individual sensors. As an alternative to the electrode type, microfabricated sensors were used, as proposed by Gimenez-Gomez et al. (2015) [144]. Here, platinum thin film electrodes were used to develop sensors for conductivity, ORP, and temperature sensors.

This conventional method of multiple sensor systems has several limitations, including its bulkiness, high cost, and complexity, which pose challenges for scaling up and deploying across extensive water distribution networks. These drawbacks have prompted the exploration of more innovative approaches, such as developing electronic tongues. According to IUPAC, an electronic tongue is an analytical tool comprising an array of non-selective chemical sensors with partial specificity and a pattern recognition system designed to identify the qualitative and quantitative makeup of both simple and complex solutions [145].

Electronic tongues represent a promising alternative for water quality monitoring, offering a more compact and cost-effective solution than traditional methods. Their simpler fabrication process, affordability, and flexible integration into larger water distribution networks make electronic tongues especially appealing for modern water quality assessment needs. The oldest and most significant work in the development and application of electronic tongue in water quality can be traced back to the work done by Legin et al. (1999) [146], where they developed the electronic tongue with 29 potentiometric chemical sensors to classify between mineral waters, between contaminated and pure water, and between samples of wines from different vineyards. Though this was a qualitative analysis, several applications for quantitative analysis in different domains were already discussed in the literature, like the work done by Vlasov et al. (1996) [147], in which they developed a multisensor system to estimate the concentrations of heavy metal cations in multi-component aqueous solutions. Witkowska Nery et al. (2015) [148] presented a paper-based, affordable electronic tongue that can classify tap and lake water from 11 mineral water samples using PCA.

Electronic tongues, in general, can be classified based on the sensing mechanism of the sensors into voltammetric, potentiometric, amperometric, and chemiresistive.

Carbo et al. (2018) [149] developed a voltammetric electronic tongue comprising electrodes of iridium, rhodium, platinum, and gold to quantify the pH and concentrations of nitrate, sulphate, fluoride, chloride, and sodium. Whereas, Cuartero et al. (2022) [150] utilized 6

potentiometric electrodes to classify between different samples of natural mineral water, as well as quantify the concentration of significant ions like chloride, nitrate, sulphate, bicarbonate, magnesium, sodium, and calcium. Kirsanov et al. and team developed various potentiometric electronic tongues for several applications ranging from quantification of rare earth mixtures [151] to water toxicity assessment [152], [153], [154], water quality control [155], [156].

With the exponential usage of smartphones (a computing device), the technology stack for interfacing with multiparametric sensors has become more adaptable [157]. Alam et al. (2021) [158] discusses a multiparameter water quality monitoring system that measures pH, temperature, free chlorine concentration, and bisphenol A on-demand and transmits the data via Bluetooth to a mobile application. Alam et al. (2022) [36] improved the electrochemical sensor array for in-situ water quality monitoring by fabricating sensors for pH, free chlorine, and temperature onto glass slides.

With advancements in chemiresistive sensor technology, electronic tongues are increasingly becoming viable for widespread application, demonstrating the significant potential for enhancing water quality monitoring infrastructure. Several articles have reported the application of chemiresistive sensors for electronic nose systems in environmental monitoring, as discussed in the reviews by Chiu et al. (2013) [159], Moon et al. (2014) [160], and John et al. (2021) [161]. However, very few articles discuss the usage of chemiresistive-based electronic tongues for water quality applications. Darestani-Farahani et al. (2022) [162] proposed using a solid-state chemiresistive sensor array to quantify nitrate, nitrite, and ammonium ion concentration in water. They demonstrated that these sensor arrays are equally sensitive and more selective than equivalent potentiometric sensors. Since chemiresistive sensors do not require reference electrodes, they are more robust for continuous water quality monitoring.

Chemiresistive sensors, with their simplistic design, present a cost-effective and scalable solution for water quality monitoring. Their affordability makes them accessible for widespread use, potentially allowing for an extensive network of monitoring stations. This scalability is crucial for covering large distribution systems ensuring comprehensive water quality assessment across diverse geographical areas. Unlike more complex and expensive technologies, chemiresistive sensors can be easily deployed and maintained, making them ideal for resource-limited settings. Their simplicity does not compromise their effectiveness; these sensors can provide accurate and reliable data, essential for informed decision-making.

As a result, chemiresistive sensors stand out as a practical choice for expanding water quality monitoring efforts, making them an attractive option for environmental scientists and policymakers alike. Their high sensitivity, selectivity, and simple fabrication process positions CNT-based sensors as prime candidates for developing electronic tongues[138], [163].

2.4 Overview of Calibration and Calibration Transfer

The development of multiparametric sensors demands the formulation of highly accurate calibration models to correctly identify and measure each analyte based on the sensors' response. This calibration is crucial for the sensors to perform effectively, especially when dealing with various substances in an environment that could interfere with accurate detection. To build such calibration models, it is important to design calibration experiments specifically tailored to address these challenges, aiming to minimize the influence of environmental disturbances on the sensor's accuracy. We can improve the sensor array's performance, validate sensor reliability, and ensure precise data collection through optimal experiment designs [164], [165] with respect to the sensor application, range, and environment.

2.4.1 Sensor Calibration

Calibration is an essential process involving the development of a mathematical model known as a calibration model that effectively maps the sensor's output signal to a specific target parameter. The development of a calibration model in sensor technology is guided by two fundamental approaches [166], [167], [168]: the classical approach and the inverse approach. The classical approach constructs the model by defining the sensor response (y) as a function of the target parameter (x), typically through a simple linear regression model such as $y = a_1x + b_1$, where b_1 represents the measurement error associated with the sensor. Conversely, the inverse approach models the target parameter as a function of the sensor response, expressed as $x = a_2y + b_2$, where b_2 denotes the error in measuring the target parameter (by the reference sensor).

In either approach, two distinct calibration models, physics-based or first principles based and data-driven, can be developed. Physics-based models derive from the fundamental principles governing the sensor's operation, aiming to capture the underlying physical relationships through calibration [169], [170], [171]. This approach models the sensor's behavior based on its design and operating conditions, providing a theoretically grounded

understanding of its outputs. These models are specific in nature and can be utilized only with those sensors whose underlying behavior can be explained as per the principles assumed. Here, the model structure is predefined, and only the coefficients are estimated from the sensor response data.

Conversely, data-driven models focus on identifying the mathematical function that best fits the calibration data, without requiring prior knowledge of the system's internal workings. These models learn from data to establish relationships between inputs and outputs, prioritizing empirical accuracy over theoretical understanding. They leverage statistical techniques and machine learning algorithms, such as linear regression, support vector machines, and neural networks, to optimize performance [172], [173], [174]. The calibration process typically involves fitting the data to several models and selecting the best-performing one based on criteria like prediction accuracy, robustness, and interpretability. This approach offers flexibility and can be applied broadly across various sensor types, making it a robust choice for diverse calibration tasks. The selection between physics-based and data-driven methodologies depends on the specific requirements of the calibration task, including the nature of the sensor, the target parameter, and the desired level of accuracy [175].

Monoparametric sensors, characterized by their ability to measure a single parameter, typically utilize univariate calibration to map a single sensor response vector to a corresponding target vector. This approach predominantly employs simple linear regression ($y = mx + c$) as its foundational technique, although the model can be adjusted based on the specific physics underlying the sensor's operation. The literature on univariate calibration is extensive [176], [177], covering both specialized models tailored for particular types of sensors and more generalized approaches for model development. The developed calibration model is essential for effectively using monoparametric sensors across various applications.

In contrast, the calibration of multiparametric sensors, capable of measuring multiple parameters simultaneously, encompasses a wider range of techniques that vary in complexity. The simplest among these is multiple linear regression, which extends the concept of linear regression to accommodate multiple inputs and outputs. However, as the complexity of the sensor responses increases, more sophisticated linear/non-linear methods are employed. These advanced techniques include projecting the sensor response onto a hyperplane and developing a regression model based on those scores, unfolding the data to introduce extra dimensions, or utilizing computational approaches like neural networks

[178], [179], [180], [181], [182], [183]. These methods are designed to capture the relationships between multiple sensor responses and their corresponding target parameters. The diversity of calibration techniques for multiparametric sensors reflects the challenges and opportunities in harnessing their full potential for comprehensive and nuanced data analysis.

2.4.2 Calibration Transfer

Calibrating multiparametric sensor systems is an intricate, time-intensive process that demands considerable resources. The calibration models produced are closely associated with the specific sensors and the environmental conditions during the calibration process. When these sensors are introduced to new environments, or as their performance naturally declines over time—a phenomenon known as sensor drift—the original calibration models become less precise and eventually outdated [184], [185], [186], [187]. This issue is compounded by intrinsic factors, such as changes within the sensor units themselves, including alterations in the sensing layer [188], [189], the binding of inhibiting substances to metal oxide layers [190], and the transformation of parts of polymers into non-conductive states due to redox reactions [191]. These factors contribute to the difficulty of applying a calibration model developed for one multisensor system based on a calibration experiment to a new sensor.

The complexity of creating a universal calibration model, which would be applicable across sensors developed in the same or different production batch, is substantial. To address this, calibration transfer techniques have emerged in chemometrics, offering strategies to adapt calibration models for use across different sensor units and settings [192]. These methods were initially crafted for spectral data, facilitating the standardization between primary and secondary spectroscopic instruments to reuse calibration models [193], [194], [195], [196], [197], thus lowering the costs of recalibration and accelerating sensor deployment. However, the application of these techniques to non-spectral, multisensor data remains relatively unexplored in the scientific literature, with existing studies focusing on specific sensor types and their novel applications, such as electronic noses/tongues[186], [198], [199], [200], [201], gas sensor arrays [202], and potentiometric sensors [185], [203], without examining their broader applicability.

In the context of CNT-based chemiresistive sensors, ensuring the selectivity and sensitivity of chemiresistors across different production batches to meet the stringent requirements of

calibration models often necessitates advanced CNT ink deposition technologies. This requirement can significantly increase the fabrication costs. An alternative strategy involves adapting the calibration model to accommodate variations between batches through minor modifications. This process, known as calibration transfer, allows cost-effective calibration adjustment without repeating the extensive experiments performed for initial model development. By leveraging calibration transfer, manufacturers can maintain sensor performance across batches more economically, mitigating the need for high-end fabrication processes.

2.4.3 Need for Calibration Transfer for CNT-based Chemiresistors

Calibration model development for a CNT-based chemiresistor sensor array is not straightforward. Fabrication of CNT-based chemiresistors does not yield identical sensors in all the batches [87], [88] unless advanced ink deposition technologies are utilized. The primary characterizing feature of CNT-based chemiresistors is their baseline resistance and minimal capacitance effect. Due to various reasons like the variations in the CNT during ink preparation, non-homogenous mixture of CNT in the ink volume deposited, uneven ink spreading onto the substrate, microscopic variations in the ink during deposition and drying, environmental effects, etc accelerate the variations among the chemiresistors within the batch [90]. Between each batch, a few more factors like ink age, conditions during ink storage and transport, minute variations within the substrate, variations in the printer between batches, etc, also add to the differences in the chemiresistors. These differences ultimately lead to each chemiresistor having a signature baseline resistance, which may or may not be within the negligible range.

Therefore, the best-fit model developed for one chemiresistor sensor array may not necessarily be the best-fit model for another set or batch of sensor arrays. This highlights the need for calibration transfer applications in chemiresistor-based sensor arrays.

For a monoparametric sensor, a simple slope and offset correction would have been sufficient to eradicate all the above-discussed differences, but for a multiparametric sensor, that is not feasible. Especially in the case of a CNT-based chemiresistive type sensor, which responds to the target analyte with different gains and may provide a differential response in the presence of interferents, multiparametric calibration model development is significantly challenging. The complexity of this challenge grows exponentially unless curbed by calibration transfer.

2.4.4 Different Approaches to Calibration Transfer

The calibration transfer methodology typically involves three approaches: model updating, transfer based on signal processing, and standardization.

In the model updating approach, the calibration model undergoes retraining through the incorporation of new samples or data derived from a secondary sensor developed in a new batch [203], [204], [205]. The success of the updated model hinges on the secondary data's quality and ability to accurately reflect the variability observed across the entirety of the secondary sensor's data. This method of implementing updates to the model is notably more straightforward than other calibration transfer techniques, making it a preferred choice for addressing issues of sensor drift by leveraging the relationship between the datasets of the primary and secondary sensors. As time progresses, these newly added data samples begin supplanting older samples within the model's training dataset, as illustrated in Figure 2.6 (b), where the primary calibration model is retrained with the new samples. Nonetheless, this strategy falls short in more complex scenarios, such as those involving non-linear relationships, where the need for an increased volume of secondary sensor data necessitates a full recalibration of the sensor. Additionally, the requirement to compute a distinct model for each sensor unit—due to each model's specific association with its respective unit—poses a significant challenge, particularly when the model computation is computationally demanding.

Signal processing-based calibration transfer methodology involves various pre-processing techniques to eliminate variations specific to the sensors or their operating environments, thereby highlighting the variations pertinent to the samples or targets.

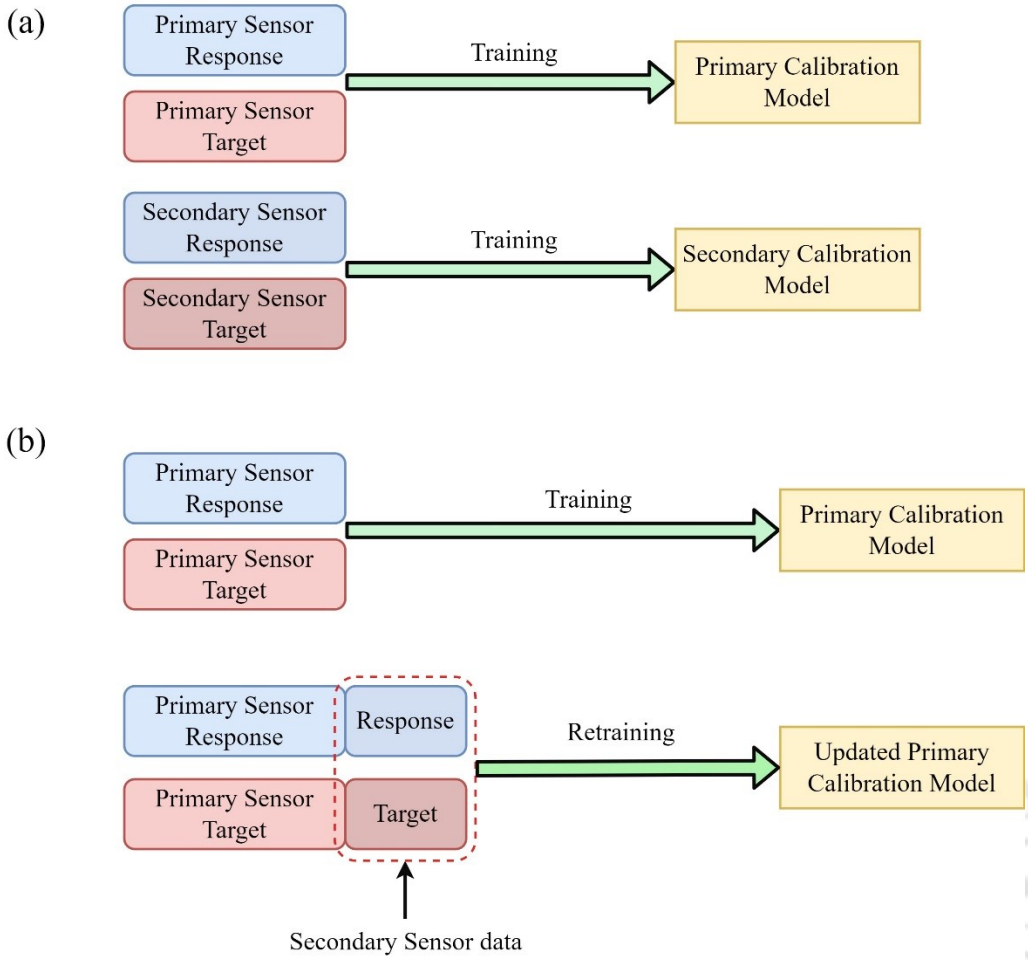

Figure 2.6: Calibration Transfer Methodology: Model Updating

Techniques range from fundamental operations such as baseline correction, scaling, smoothing, noise reduction, and sample normalization to more sophisticated methods like finite impulse response (FIR) filtering, generalized moving window multiplicative signal correction (MSC) [206], [207], orthogonal signal correction (OSC) [208], [209], [210], and wavelet-based pre-processing [211]. Following this preparatory stage, the calibration model is constructed using the processed signals (Figure 2.7 top), which helps insulate the model from discrepancies stemming from different sensors or environmental conditions. Once the calibration model is developed and the signal processing steps are identified, predictions with any other sensor can be made with the same calibration model after repeating the same signal processing steps.

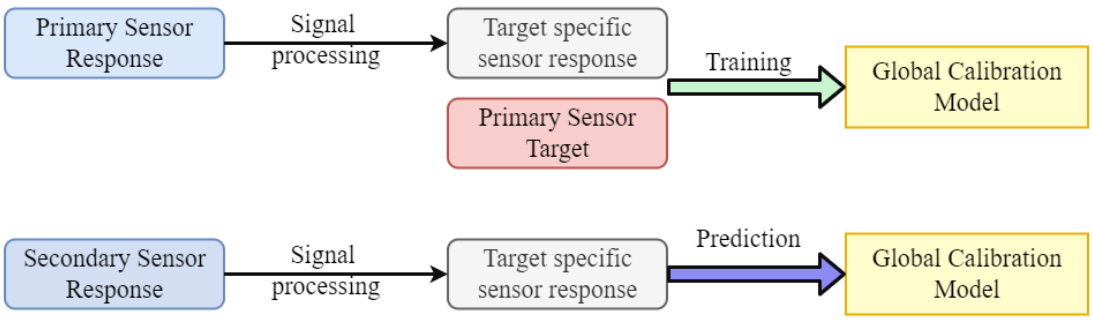

Figure 2.7: Calibration Transfer Methodology: Signal Processing

Nonetheless, the challenge of precisely pinpointing and segregating these sensor-specific variations limits the applicability of this strategy, confining its use to scenarios where such differences can be accurately identified and quantified.

Standardization techniques involve the adjustment of the secondary sensor's response to align with the characteristics of the primary sensor's response. This alignment is achieved through modifying sensor responses or model predictions via mathematical operations. The standardization process is executed in two phases: initially, a mapping is created between the responses of both the primary and secondary sensors through the application of transfer standards (Figure 2.8). Subsequently, this established relationship is employed to predict the measurand value based on new outputs from the secondary sensor.

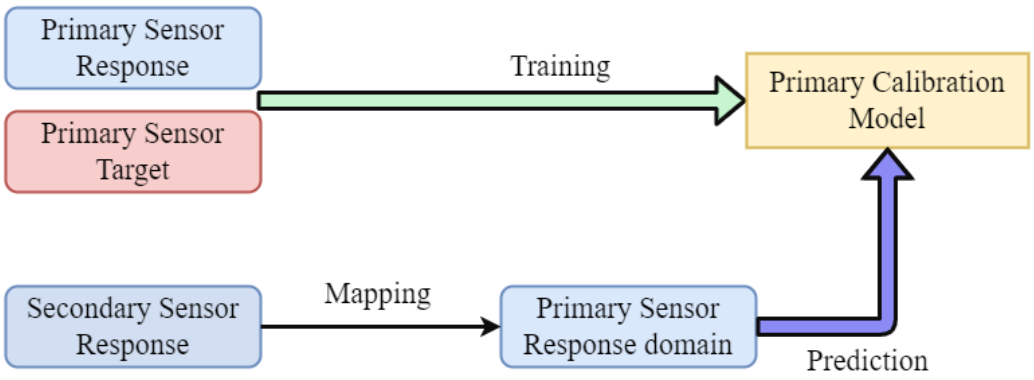

Figure 2.8: Calibration Transfer methodology: Standardization

Within the scope of multisensor systems, the direct standardization approach is frequently discussed for transferring calibrations [185], [186], [198], [202], [203]. Developed methods, including single wavelength(sensor) standardization [185], [186], model adjustment through sample weighting, Tikhonov regularization, joint Y-partial least squares regression [203], the application of transfer learning (transfer sample-based coupled task learning) [201], Spectral space transformation [212], and Variable Penalty dynamic time warping method [213] have been tailored to the unique aspects of their respective applications. These

techniques, however, often lack the universality required for broad application across different sensor systems, highlighting the ongoing challenge of achieving effective calibration transfer in diverse sensing scenarios.

2.5 Identification of Literature Gaps and Future Directions

The literature review identifies several vital gaps in the field of water quality monitoring:

- Due to high initial and operational costs, current pH and active chlorine sensing technologies on the market are expensive for widespread implementation within water distribution networks, particularly in developing nations like India.
- While academic research has proposed more affordable sensors, there is a noticeable lack of information on their long-term stability, deployment outcomes, and capability for sustained operation, making their immediate real-world application uncertain.
- There is significant research interest in developing sensors based on carbon nanotube (CNT) technology for monitoring active chlorine, pH, and other water quality indicators, showing an apparent inclination toward utilizing CNTs for water quality assessment.
- However, the literature on CNT-based chemiresistive sensors predominantly focuses on individual water quality parameters, with a minimal exploration into their potential as multiparametric sensors for a holistic approach to water quality monitoring.
- No commercial versions of these comprehensive multiparametric sensor systems (on CNT-based chemiresistors) have been developed or made available in the market.
- Developing a universal calibration model for multiparametric sensor systems poses difficulties due to differences between units produced in the same batch, highlighting an urgent need for methods that allow calibration models to be applied across various multiparametric sensors.
- There is limited discussion on using Calibration Transfer Techniques for systems with multiparametric sensors, and this discussion typically focuses on individual types of sensors, indicating a lack of comprehensive analysis and application in this area.

The core research objective can be established as the intersection of continuous and real-time monitoring of water quality parameters using a Multiparametric sensor based on chemiresistors made out of carbon nanotubes, visualized in Figure 2.9.

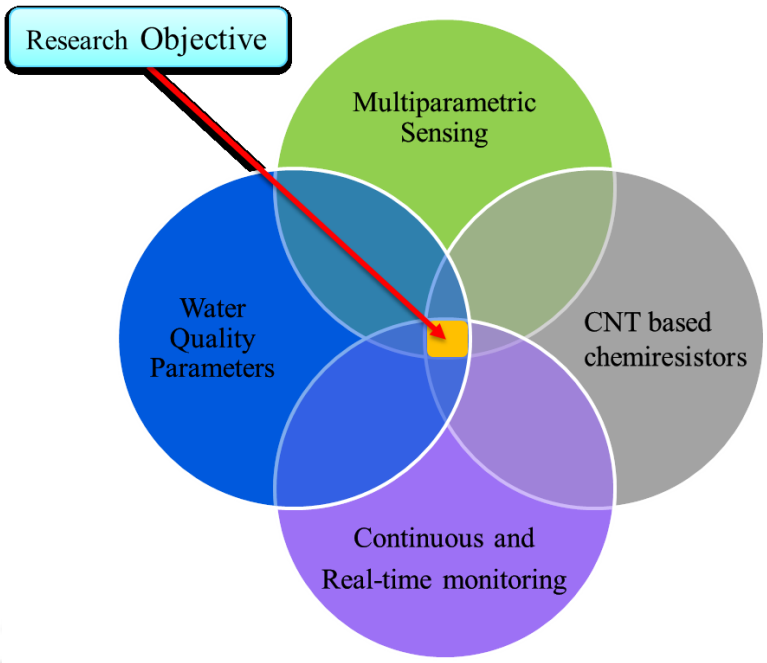

Figure 2.9: Visualization of the core research objective

In the context of the project LOTUS, Sensor chips are developed and fabricated by the LOTUS project team at Université Gustave Eiffel [39], [214]. Developing an End-to-End integrated sensor platform for the sensor chip for water quality monitoring constitutes the primary intent of this thesis. From this perspective, based on the literature survey conducted on the current water quality monitoring technologies, the following research objectives were formulated:

1. *Development of a Sensor Platform with the LOTUS sensor chip*:

    - End-to-end integration of the LOTUS sensor chip – a CNT-based chemiresistive sensor for pH and Active chlorine sensing
    - Improving the pre-existing data acquisition system (DAQ) to interface with the chemiresistive sensor, considering the typical issues pertaining to electrical sensing in water
    - Development of software tools to interface with the DAQ to collect the data, process the data, and provide a minimal data visualization platform for user interface

- Designing, developing, and validating the deployment strategy of using a bypass structure to protect the sensor from fouling and provide easy installation

2. *Characterization and Calibration of LOTUS sensor:*

- Formulating the experiment protocols for laboratory setup and in a water loop representing a typical water distribution network
- Developing calibration models for the LOTUS sensor based on the generated experimental data
- Characterizing the sensitivity, accuracy, and uncertainty associated with the sensors and calibration model

3. *Optimizing and Streamlining Mass Calibration Process*:

- Assessing the Calibration Transfer techniques reported in the literature for use with multiparametric sensors, using available gas sensor datasets to identify the optimal strategy for application to LOTUS sensor
- Developing a methodology for deploying the identified calibration transfer technique to LOTUS sensor aiding in Mass Calibration

3

Sensor Platform Development

Abstract: *This chapter presents the end-to-end integration of the LOTUS sensor – a multiparametric sensor with several CNT-based chemiresistors that can measure active chlorine and pH, thermistors that can measure temperature, and capacitive sensors that can measure conductivity. The chapter introduces the sensor chip, explains the fabrication and assembly of the chip into the sensor, and discusses the evolution of the DAQ system developed explicitly to optimally acquire response from this chip and transfer it to a connected host system. This chapter concludes with a brief discussion of the minimal dashboard developed for data visualization.*

3 Sensor Platform Development

3.1 Introduction

The sensor platform of the LOTUS project comprises three main components, each playing a crucial role in the functionality and data collection process of the system (Figure 3.1):

Sensor Head: This component houses the assembled sensor chip. The sensor chip is the core element responsible for detecting and measuring the various parameters related to water quality. It is designed to accurately sense the water conditions based on Active chlorine, pH, conductivity, and temperature.

Sensor DAQ: The Sensor (Data Acquisition) DAQ is the intermediary electronic card connecting the sensor head to the host system. Also known as Analog Front End (AFE), this DAQ is a microprocessor-driven data acquisition system that communicates with each sensor device in the sensor head. The AFE triggers the sensors, gathers their outputs, and pre-processes the data before sending it to the host system. The precision of the data collected significantly influences the sensor performance.

Host System: The Host System is the central hub for data transmission and user interaction. The Sensor DAQ communicates the collected sensor data to the host system through a USB/UART interface. Depending on the specific application or use case, the host system serves as the user interface for end-users, allowing them to monitor and analyze the water quality data in real-time. The host system can also function as a data transmitter, sending the collected data to the cloud for storage, analysis, or further processing.

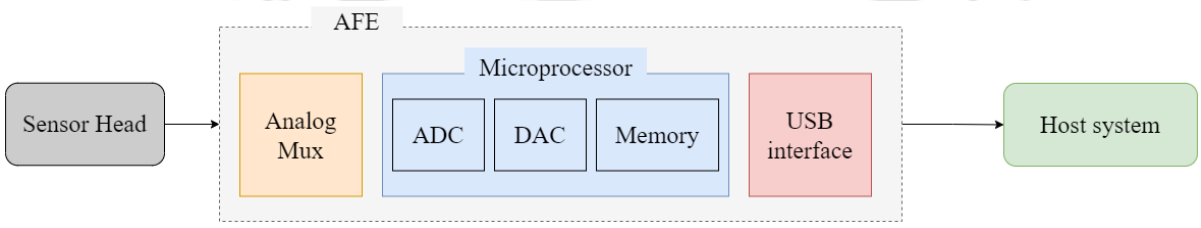

Figure 3.1: LOTUS sensor system - Block diagram

The sensor casing is a crucial component of the sensor unit, protecting the internal components from water (as the sensor will be immersed in a continuous water stream) and houses the sensor chip and the AFE system. The sensor casing is divided into parts made from PVC material: the sensor head casing and the sensor body casing. The sensor head casing, as shown in Figure 3.2 on the left, is where the substrate is attached. In contrast, the sensor body casing, depicted on the right in Figure 3.2, encloses the AFE responsible for

gathering data from the sensor. This sensor unit is designed to be exposed to water flow, enabling real-time monitoring of water quality parameters. The casings are designed for ease of installation and maintenance.

Figure 3.2: Sensor casing: left = sensor head casing and right = sensor body casing

In brief, the LOTUS project's sensor platform comprises the sensor head for assembling the sensor chip, the sensor DAQ (i.e., AFE) for acquiring and processing data, and the host system for transmitting data and facilitating user interaction. These elements collaborate harmoniously to guarantee precise data gathering, analysis, and monitoring of water quality parameters. The sensor resulting from this platform is referred to as the LOTUS sensor, which aligns with the project's title.

3.2 Development of Sensor Head

The fundamental unit of the LOTUS sensor is a chemical resistor, which is developed by depositing a sensing material on top of a pair of inter-digitated electrodes (IDEs). This sensing material interacts with the environment and changes the resistance or conductance. This change is measured by a data acquisition system which transmits the data for further processing.

In the case of the LOTUS sensor, the sensing material is made of Carbon Nanotube (CNT) inks – either pristine (non-functionalized) or functionalized with a polymer. This functionalization modifies the sensitivity of the ink to specific analytes, thereby leading to detecting different analytes.

3.2.1 Overview of Chemiresistor Fabrication

The fabrication of the chemiresistor can be summarised as follows:

1. Synthesis of functionalization polymer (e.g., FFUR or PVT118)
2. Preparation of carbon nanotube ink and FFUR functionalized CNT ink

3. Deposition of the ink onto the IDEs in the sensor chip.

Lebental et al. (2017) [214] discuss the procedure for synthesis of the CNT ink and its functionalization. Pristine CNT and FFUR (Polyfluorene) inks were initially developed as part of the Proteus project, a predecessor to the LOTUS project, and further improved throughout LOTUS. Based on the results from Proteus project (deliverable 5.3) [215], the selectivity of the chemiresistor (ink type) to the target is mentioned in Table 3.1.

Table 3.1: Summary of Ink selectivity with respect to target parameters

Ink type	pH	Active Chlorine
Pristine CNT	High	Low
CNT + FF-UR	Low	High

The PVT118 (Polyvinyltriazole) was introduced as part of the LOTUS project through a partnership with the LCPO laboratory (CNRS/Université de Bordeaux). They are demonstrated to have selective sensing capability primarily for arsenic and then for active chlorine and pH. The ink printing on the chemiresistor substrate/chip was done via manual drop casting [216] for initial testing. This technique, though unreliable in providing reproducible electrodes, did provide a quick-to-test option that was then evaluated in the laboratory environment. The ink deposition was then done using an industrial ink printer for the field-testing version.

3.2.2 Sensor chip v1 - Silicon version

The first version of the LOTUS sensor chip consists of a sensor array with 20 CNT chemiresistors, of which 10 are pristine CNT (labelled as 1A to 1E and 2A to 2E in Figure 3.3)) and 10 are functionalized with a polymer FFUR (labelled as 3A to 3E and 4A to 4E). This polymer was identified to be more selective to pH than the pristine CNT-based devices [40], [41], as will be discussed in detail in Chapter 4. This difference enabled quantifying the pH and active chlorine concentration using only the two inks. Ink deposition onto the IDEs to fabricate the chemiresistors was carried out with Dimatix Materials Printer DMP – 2850 [217]. The number of layers of the ink to be added was determined by having the IDEs exhibit a resistance level below 100 kΩ.

In addition to the 20 sensors mentioned, the chip incorporated two temperature sensors (T1 and T2 in the diagram) and three conductivity sensors (C1, C2, and C3). The temperature sensors were designed in a serpentine shape using metal-based thermistors, while the conductivity sensors were constructed as parallel-line capacitance sensors. The variations in

sample conductivity were indicated by changes in the dielectric constant, resulting in alterations in capacitance. This shift, detected as a modification in potential difference, was directly linked to the sample's conductivity level. The sensor response for chemiresistor and thermistor is in resistance (ohms) and the same for conductivity sensor is in voltage difference (mV), for practical reasons.

The target specifications for each type of sensor are summarized in Table 3.2.

Table 3.2: Requirement Specification for Sensor Chip v1 - Silicon version

Parameter	Chemiresistor	Temperature sensor	Conductivity sensor
Response range	1 kΩ to 100 kΩ	940 Ω to 1200 Ω	0.5 mV to 7mV
Target range	Chlorine: 0 – 5mg/l pH: 5 – 10 units	-20°C to 60°C	500 µS/cm to 10 mS/cm
Response precision	~0.1% full scale	~ 0.02 Ω	~ 70 µV
Target precision	0.01 mg/l, 0.1 pH	0.1°C	100 µS/cm
Measurement protocol	2-wire probe, DC 5 µA signal, 10 ms duration	4-wire probe, DC 125 µA signal, 10 ms duration	4-wire probe, AC 0.8V, 5kHz signal for 3 seconds

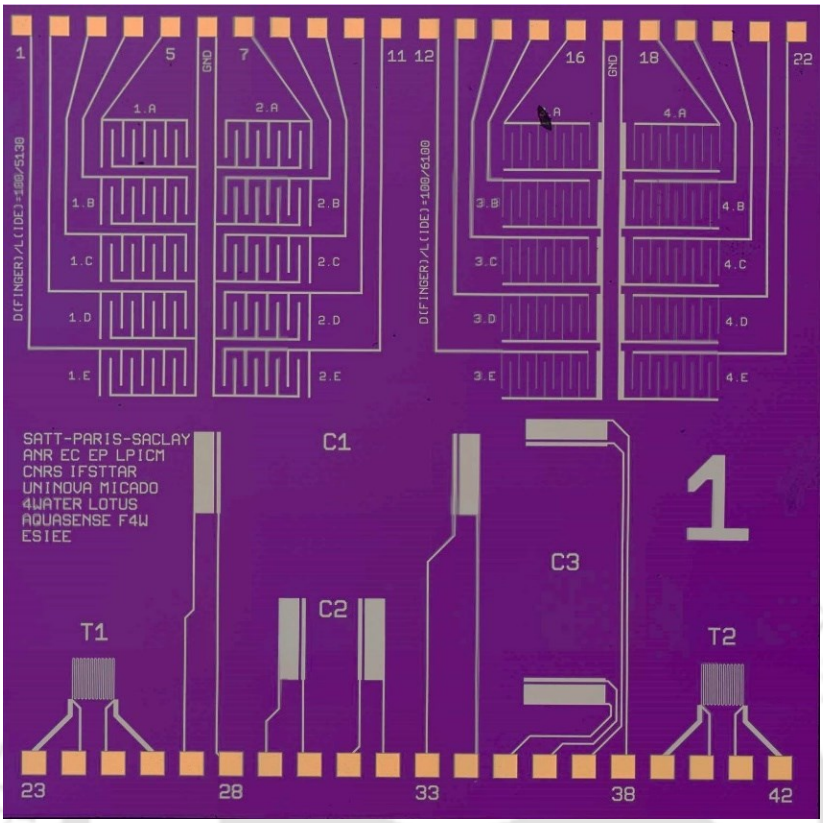

Figure 3.3: Sensor chip - Overview

In this initial version of the sensor chip, sensor array components, such as chemiresistors, temperature, and conductivity sensors, were developed on a silicon substrate and subsequently mounted onto a support Printed Circuit Board (PCB) for wire bonding. This wire bonding enables the sensor DAQ to interact with each of the 25 devices and collect the sensor response. A glob-top was necessary to protect the wire bonding from continuous water exposure and other minute particles (Figure 3.4). The support PCB can be directly connected to the sensor DAQ system via the head casing.

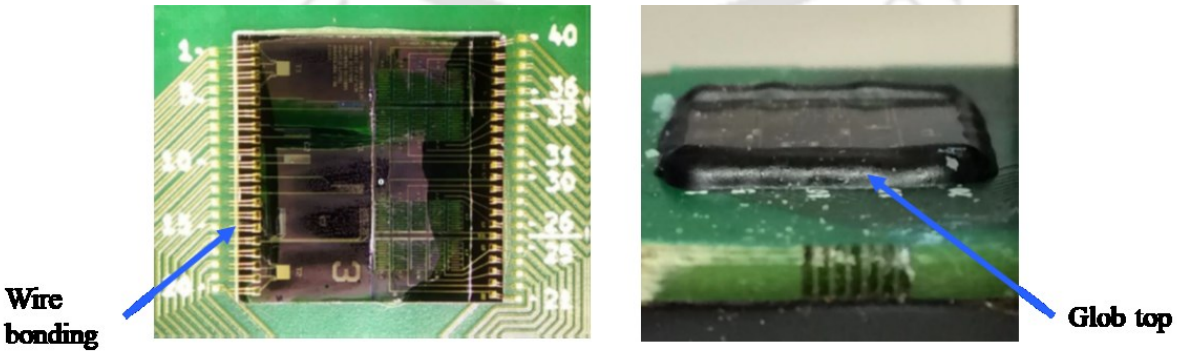

Figure 3.4: Silicon version sensor chip: Left - Top view showing wire bonding of the chip onto the support PCB, right

3.2.3 Sensor chip v2 – Plastic version

Originating from a silicon chip, the original LOTUS Sensor concept was devised to integrate essential electrodes for CNT-based chemiresistors and sensor circuits for conductivity and temperature measurements. While effective, this design faced various limitations mainly attributed to the expensive raw material cost of the silicon chip, the requirement of a support PCB and wire bonding, and glop topping. These increased the sensor's overall complexity and manufacturing expenses, necessitating multiple quality control protocols to ensure the product's quality.

The LOTUS team replaced the silicon chip with a plastic substrate to solve these challenges. This new version uses gold-nickel-copper lines for the sensor electrodes, developed onto a single polyimide (PI) sheet, significantly streamlining the assembly process [218]. This advancement permits integrating sensing elements and connectivity on a singular substrate, enabling straightforward attachment to the sensor head with epoxy resin. Consequently, this innovative approach not only simplifies the manufacturing process but also substantially reduces the need for additional components and processes, such as PCBs and wire bonding, thereby enhancing product reliability and lowering production costs by up to 10 times (as the bare chip cost for silicon version is estimated to be above 10€/cm^2 and below 1€/cm^2 for the plastic version).

The transition to a plastic chip was executed in two distinct phases, beginning with the production of substrates that feature single chemiresistors consisting of pairs of interdigitated electrodes coated with CNT inks, referred to as single-electrodes, with manual drop-casting of CNT ink (illustrated in Figure 3.5 (a)). In the initial stage of plastic chip manufacturing, the compatibility of the metal utilized in creating the chemiresistor was fine-tuned to enhance resistance to chemicals in water environments and ensure robust mechanical strength. This preliminary phase was supported by collaborations with external vendors and technology licensing partners, aiming to evaluate the new design's performance and reliability in the lab and real-world environments.

These single electrode devices were assembled into sensor heads of 2, 3, or 6 devices (Figure 3.5 (b)) and subjected to various experiments to assess their performance. Based on the results from lab scale experiments, the final field-ready version was developed (Figure 3.5 (c)), integrating temperature and conductivity measurements with multiple chemistors.

This field-ready version (referred to as LPCB) consists of 12 CNT-based chemiresistors, wherein pristine CNT ink was deposited on 4 of them and two different functionalized inks were deposited on four sensors each. As illustrated in Figure 3.6 (a), the first column from the left, which has four chemiresistors (labelled as R11, R12, R13, and R14 from top to bottom), has pristine CNT ink deposited. The two functionalized inks, FFUR – CNT and PVT118 – CNT, are deposited in the following two columns (R21 to R24 and R31 to R34, respectively). Additionally, this version has two temperature (T1 and T2) and two conductivity sensors (C1 and C2).

The ink deposition onto the chemiresistor IDEs were performed by the DMD 100 drop casting printer with a nozzle size of 0.25 mm diameter. This higher diameter nozzle dispersed the ink improperly over the chemiresistor IDEs. Secondly, the gap between the nozzle and the substrate was controlled by the mapping software used in DMD100 printer, but due to incompatibility the software tool was not able to maintain the gap accurately, which lead to needle tapping as depicted in Figure 3.6 (b). Therefore, some chemiresistors were damaged due to the needle tapping, yielding a higher unstable resistance for those.

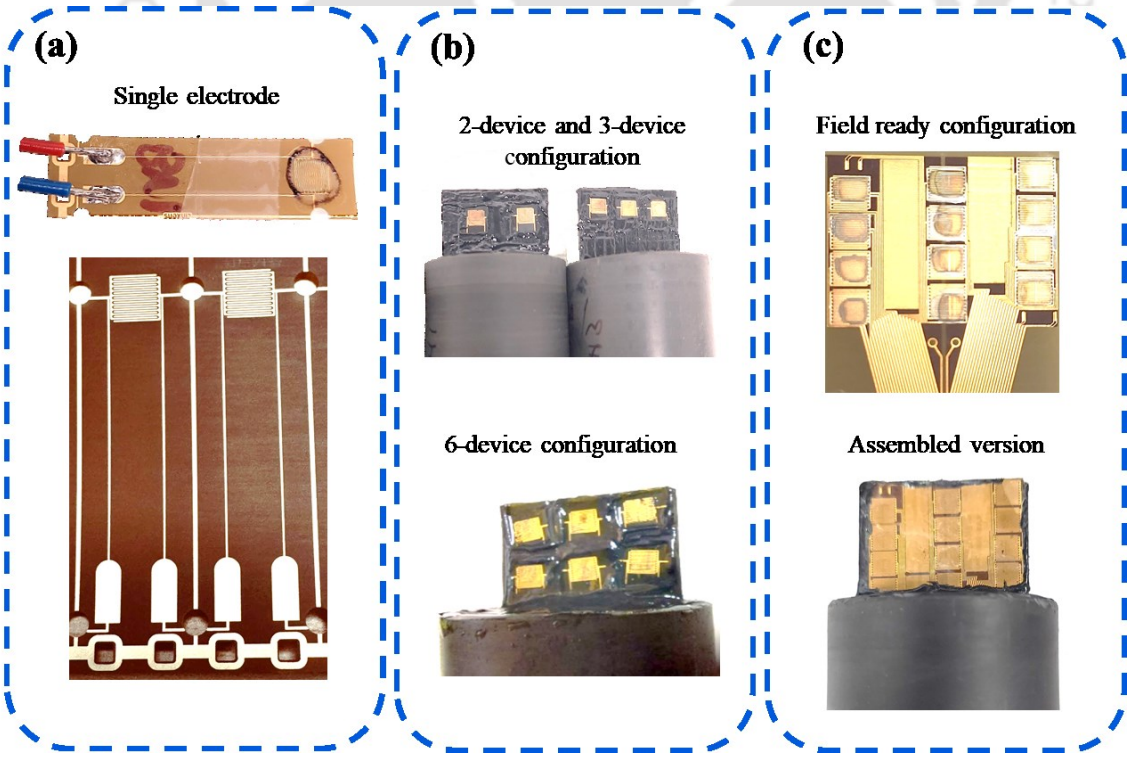

Figure 3.5: Sensor chip v2 - Plastic version iterations

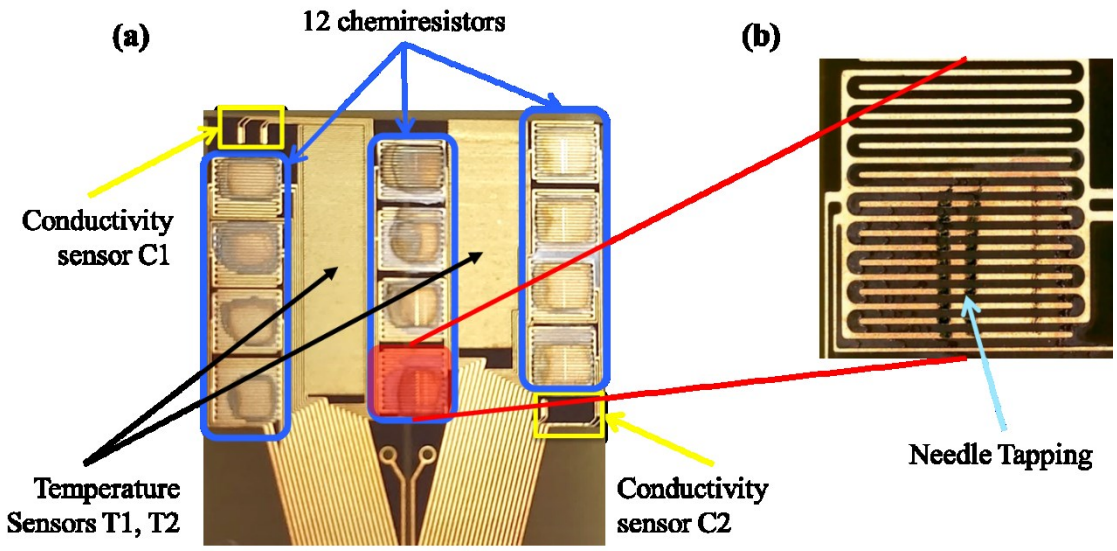

Figure 3.6: Sensor chip v2 - Field ready version (LPCB). (a) Showing the individual sensor types, (b) highlighting the needle tapping on IDEs.

3.2.4 Sensor head assembly

Each single electrode device (after ink deposition) was soldered with cables to connect to the Sensor DAQ for the sensor head assembly. After the wire connection, the electrodes were placed onto the sensor head casing based on the design requirement (2-device, 3-device or 6-device). Then, an epoxy adhesive is used to secure the devices onto the sensor head, as seen in Figure 3.7 (a). The sensor head and the plastic substrates were further strengthened by filling the interior of the sensor head with epoxy as well, as shown in Figure 3.7 (b). After the required curing time, the sensor heads were employed for experiments to measure water quality.

The assembly of the LPCB chip into the sensor head is much simpler compared to the individual electrodes case (2-device, 3-device, or 6-device), as this is a single plastic sheet that is firstly glued to the sensor head (Figure 3.8 (b)) and then as discussed before, epoxy is poured inside the sensor head (Figure 3.8 (c)) to complete the assembly.

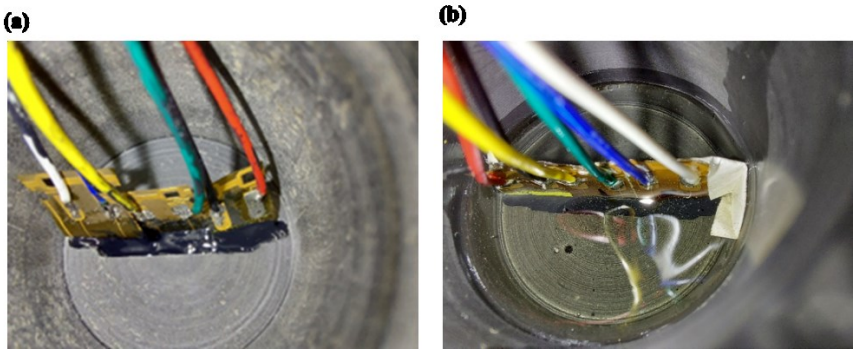

Figure 3.7: Sensor head assembly: (a) after glueing the single-electrode sensors to the sensor head, (b) after adding epoxy to the sensor head

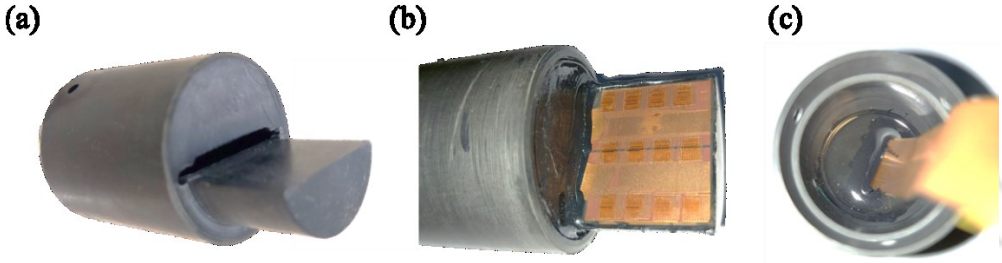

Figure 3.8: Sensor head assembly - LPCB. (a) Blank sensor head casing, (b) LPCB glued to the sensor head, and (c) Epoxy filled sensor head

3.3 Analog Front-End (AFE)

The sensor DAQ system, named Analog Front End (AFE), consists of a microprocessor at its heart, which activates each sensor using a specific signal and measures the response via an analog-digital converter (ADC). The microprocessor connects to each sensor via a multiplexer, thereby reducing the microprocessor's need for multiple input-output (IO) channels. This AFE then transmits the data to a host system, processing and visualizing the final data.

3.3.1 Sensor Excitation signal

As discussed earlier, the sensor chip consists of 3 types of sensors: chemiresistors, thermistors, and capacitance-based conductivity sensors. Accordingly, each type had limitations and specifications and required its excitation signal and data acquisition protocol.

An overview of the specifications for each sensor is summarised in Table 3.3. For the silicon version, the excitation signal for chemiresistors and thermistors was a constant DC pulse of 5 μA and 125 μA, respectively, for 1 second. The AFE measured the voltage drop across these resistive sensors and was then converted to resistance. This value is then utilized as the sensor response for further processing. On the other hand, a sine wave of 5kHz frequency and 0.8 V peak-to-peak amplitude is used as an excitation signal for the conductivity sensor. This signal is applied to the external electrode of the sensor, and the voltage across the internal electrodes is measured (referring to Figure 3.3, for sensor C2, pin numbers 29 and 32 correspond to external electrodes, and pin numbers 30 and 31 correspond to internal electrodes).

Table 3.3: Sensor Type Specifications

	Chemiresistors	Thermistors	Conductivity sensors
Sensing parameter	Active chlorine, pH	Temperature	Conductivity
Nominal range (Silicon version)	1 kΩ to 100 kΩ	1kΩ to 1.5 kΩ	5 μF to 50 μF
Activation signal (Silicon version)	DC, <5μA	DC, <1mA	AC Voltage, 0.8V
Nominal range (Plastic version)	1 kΩ to 100 kΩ	10 Ω to 50 Ω	20 nF to 5 μF
Activation signal (Plastic version)	Rectangular wave, <5μA	Rectangular wave, <1mA	AC Voltage, 0.8V

The optimized excitation conditions are determined by evaluating experiments in laboratory settings over short and long durations. Further information on establishing these optimal conditions is detailed in Chapter 4.

The response from the sensor is deemed of good quality when it has very low noise, which is characterized by a coefficient of variation (CoV) below 0.1%. The CoV is calculated as the ratio of standard deviation to the mean of the data, which in the present case is considered as the sensor response for a 10-minute window. Initially, the excitation signals developed for the silicon version were used in the plastic version, but these proved harmful to the

sensors as the excitation signal consists of applying a constant DC across IDEs, which may induce electrolysis.

Apart from the general constraint of using current < 5μA for chemiresistors, the voltage across any sensor should not exceed 1V as it causes water splitting [219]. The AFE utilizes a constant current source, so higher resistance (> 200 kΩ) at the sensors, as well as mechanical damages in the IDEs (~ open circuit), will also cause water splitting. The prolonged exposure to 5 μA current also accelerates this issue apart from damaging the CNTs.

The water splitting is harmful because the gases formed as a result oxidise the nickel layer, exposing the copper lines, which corrodes in water, leading to a green deposition (Figure 3.9 (b)), which increases the sensor resistance and leads to the death of the sensor.

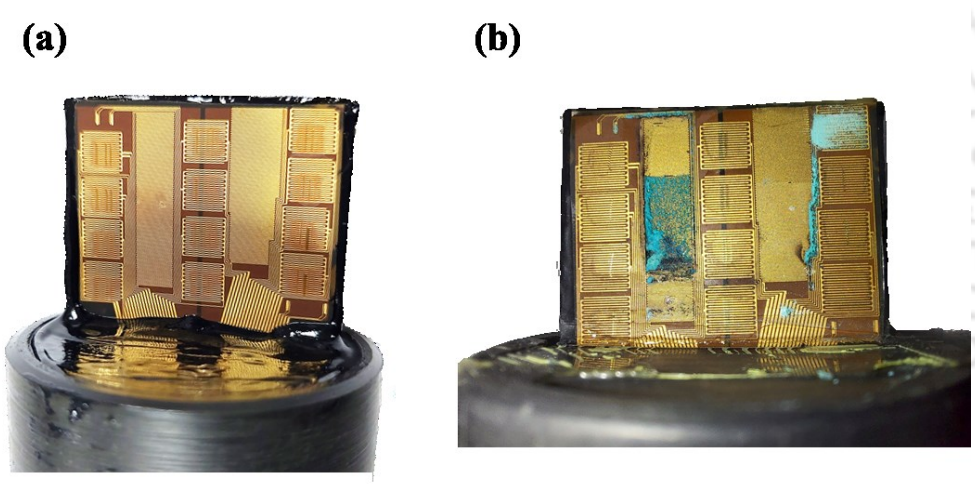

Figure 3.9: Sensor chip before (a) and after (b) corrosion due to overvoltage

After several iterations, the excitation signal that provided much better CoV was a sine wave and rectangular wave compared to a constant DC pulse for chemiresistors. The signal improvement may be attributed to the dynamic nature of the signal (which prevents the capacitance saturation – since the CNT-based chemiresistor is a resistive sensor with minimal capacitance). Further investigation is required to confirm this.

For preliminary testing, the rectangular wave was chosen as the excitation signal, and five parameters were chosen to define the rectangular wave, as visualized in Figure 3.10.

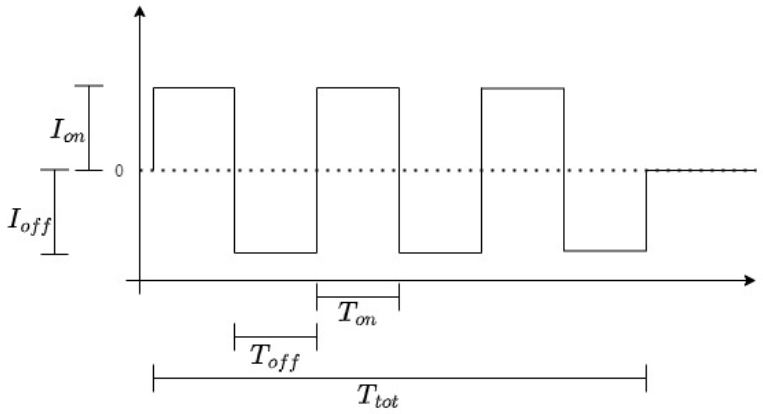

Figure 3.10: Sensor excitation signal: Rectangular wave

Table 3.4 shows the default values of the rectangular wave used for chemiresistors and thermistors in further testing and experimentations. The selected excitation signal for the conductivity sensor is a sine wave (0.8V, 0.8kHz).

Table 3.4: Sensor Excitation signal: Rectangular wave parameters

Parameter	Chemiresistor	Thermistor
On current, I_{ON}	5 µA	1 mA
Off current, I_{OFF}	0 A	0 A
On current duration, T_{ON}	2.5 ms	2.5 ms
Off current duration, T_{OFF}	2.5 ms	2.5 ms
Total acquisition duration, T_{TOT}	1 s	1 s

3.3.2 Sensor Response Signal

Chemiresistors exhibit characteristics beyond mere resistive behaviour, including a noticeable capacitance effect. This capacitance influence is evident in the sensor's response signal, as depicted in Figure 3.11. The signal's steady-state region is distinguished by a segment where the measured values remain within 2% of the calculated average. This identification of the steady state is critical for accurate analysis and interpretation of the sensor's response data.

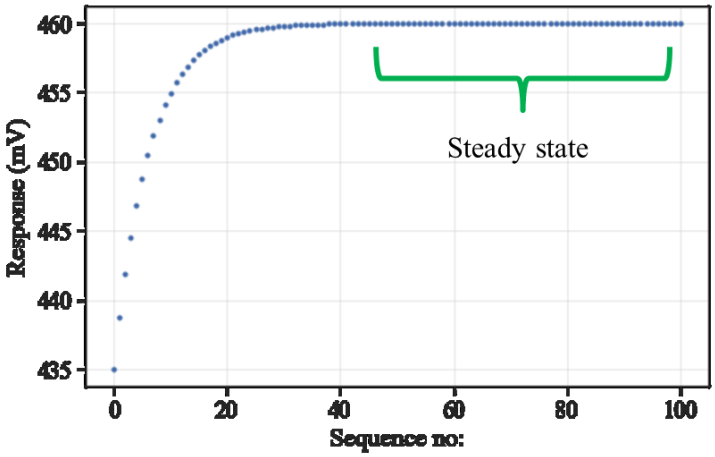

Figure 3.11: Sensor response signal

Signal acquisition for both chemiresistors and thermistors is conducted at predetermined intervals, with the silicon-based sensors collecting data throughout their activation period (since the signal is a constant DC) and the plastic-based sensors gathering data during their ON current phases (since the signal used is a rectangular wave with ON and OFF phases). In the conducted experiments, 100 data points were collected from the silicon version, while the plastic version yielded 200 for each excitation cycle. These numbers are fixed based on the different versions of AFE, as discussed in the next section. The approach to calculating the sensor response varies by sensor type. For thermistors, the response is determined by the potential drop across the device, calculated as the average of the last 10 data points measured at the thermistor's terminals during signal excitation.

For chemiresistors, the sensor's response is estimated as the average of the final 10% of the total data points collected during an excitation cycle (10 data points for the silicon version and 20 data points for the plastic version), even though the sensor's response may stabilize before reaching this threshold. Regarding conductivity sensors, the silicon version utilizes the root mean square (RMS) voltage of the last 10% of data points as the response metric. Conversely, in the plastic version, this RMS calculation is substituted with the average peak-to-peak (APP) voltage of the final 10% of data points. The peak-to-peak voltage was preferred primarily because the computation of RMS values for each cycle was computationally expensive and time-consuming for the AFE. Essentially, it is a tradeoff for acquiring more data in less time.

3.3.3 AFE first version – UAFE

The initial version of the AFE, developed under the Proteus project [215], utilized the Cypress PSoC (CY8C5888LTI), integrating a 32-bit ARM Cortex M3 core processor with a 20-bit Delta Sigma ADC (analog to digital converter). In this thesis, this AFE version is referred to as Uni-AFE (UAFE) because the Project team at Université Gustave Eiffel in Paris designed this. The precision of the analog-to-digital conversion, a critical aspect of sensor performance, hinges on the bit count of the ADC—the higher the number, the finer the resolution. By opting for a Delta Sigma type ADC, the AFE aimed to enhance signal quality by reducing noise and improving resolution, albeit at the cost of speed, increased expense, and higher power consumption compared to alternative ADC types, such as the Successive-Approximation Resistor (SAR).

3.3.4 AFE specification and test case development for upgraded AFE versions

The initial version of the AFE faced challenges that rendered it inadequate for the evolving needs of excitation and response signal processing within the project. These limitations were not only related to its core functionality but were exacerbated by its inability to effectively reduce noise and mitigate interference from environmental electromagnetic sources. The impact of such interference was notably documented in experiments, as detailed in Figure 3.12, which captured the sensor response of a chemiresistor immersed in tap water over an extended period.

Two AFEs were placed nearby, and data from 2 different chemiresistors in 2 beakers were continuously measured. Around 71 hours into the experiment, marked by a red dotted line on the plot, the signal dispersion significantly decreased upon deactivating the second AFE. This occurrence highlighted the vulnerability of the first AFE version to electromagnetic interference (produced by the second AFE in this particular experiment), leading to the distorted accuracy of the sensor's readings. This realization prompted a shift in design priorities, emphasizing adhering to stringent standards such as IEC 61326, which outlines requirements for electrical equipment's immunity and emissions concerning electromagnetic compatibility. By aligning with these standards, subsequent versions of the AFE aimed to mitigate interference effects, ensuring more stable and reliable sensor data collection.

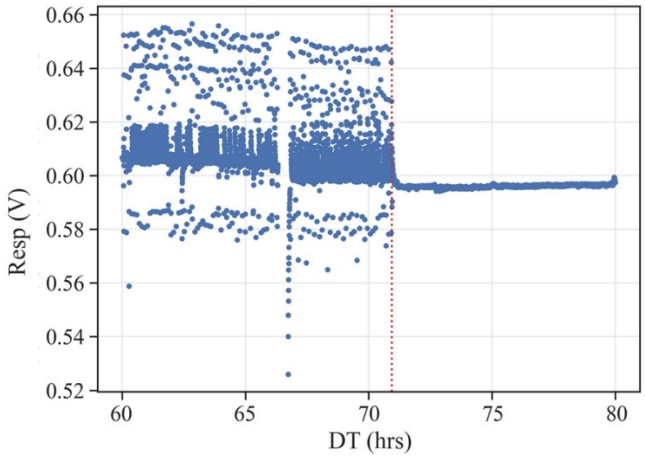

Figure 3.12: Sensor response signal highlighting the effect of electromagnetic noise

The AFE requirements were, thus, defined as follows:

- The AFE aims to communicate with a versatile sensing platform: multiple sensors of different types. Consequently, the AFE should be versatile and compliant with evolving sensor chip designs.
- For the chemiresistor channels, the AFE should be able to measure resistance between the range of 1kΩ to 100kΩ, with a CoV below 0.1%, with a maximum permissible source amplitude of 5μA
- For temperature channels, the AFE should be able to measure the resistance between the range of 10 to 100 Ω, with a CoV below 0.02%, with a maximum permissible source amplitude of 1mA
- Resolution for chemiresistors should be in the range of 100Ω (0.1% of full-scale reading), and for temperature sensors should be in the range of 0.1Ω (0.1% of the full-scale reading).
- Under no circumstances should the voltage drop across any sensor channel cross 0.9V.
- Electromagnetic (EM) shielding to comply with IEC 61326 EM standards.

3.3.5 AFE Evolution

The initial experimentation phase revealed that the UAFE fell short of meeting the specific Data Acquisition (DAQ) requirements, prompting the need for an advanced version. In response to this inadequacy, the UAFE was upgraded to the second version, developed by

Terramatic Solutions Pvt Ltd with the support of IITG and Université Gustave Eiffel, and is referred to as TAFE (Figure 3.13).

The TAFE was engineered from the ground up, incorporating a more advanced microprocessor, the Xtensa LX6 processor equipped with an ESP32-WROOM-32 module. This upgrade introduced several benefits, including a 24-bit ADC for enhanced data precision, integrated Bluetooth and Wi-Fi for improved connectivity, and a significant reduction in power consumption compared to its predecessor. Despite these advancements, the TAFE encountered its own set of challenges, such as the absence of a direct current source, which necessitated the use of a voltage source as an indirect means to activate the sensors. Furthermore, the TAFE's design adjustments limited the adaptability seen in the UAFE, particularly in sensor routing capabilities, due to multiple hardware modifications aimed at meeting the DAQ requirements. The TAFE version also lacked overvoltage protection and, therefore, damaged the sensors when used longer.

To address the limitations encountered with the TAFE, the development trajectory led to another upgrade in collaboration with Monarch Innovation Pvt Ltd and Hydroscope Technologies Pvt Ltd, culminating in the development of the MAFE (Figure 3.13). This iteration saw the introduction of an improved microprocessor from the Cypress family (the same microprocessor family as that of UAFE) coupled with an external 24-bit ADC to enhance the quality of sensor data collection further. The transition to MAFE also included a series of software optimizations aimed at refining sensor excitation and response signal processing, thereby overcoming the drawbacks of its predecessors. Moreover, the MAFE introduced shielding mechanisms to protect against EM interference, a critical feature for ensuring the integrity of sensor data in varied environmental conditions. As defined in the requirements, MAFE also provided a Zener-based overvoltage protection for the sensor.

This suite of improvements positioned the MAFE as a more functional, deployment-ready version of the AFE, embodying the project's commitment to achieving high data accuracy and reliability standards in sensor operations. A brief comparison of the significant differences between the 3 versions of AFE is summarised in Table 3.5.

Table 3.5: Brief comparison between different iterations of AFE in the application context

	UAFE	TAFE	MAFE
Microcontroller	Cypress CY8C5888AXI	Xtense LX6 Espressif	Cypress CY8C5688AXI

		WROOM 32D	
ADC	20-bit Delta-Sigma	24-bit external	24-bit external
Active power consumption	15.4mA at 48MHz	30mA at 80MHz	15.4mA at 48MHz
WiFi + Bluetooth	No	Yes	No
EMI shielding	No	No	Yes
Overvoltage protection	Not required (VCC < 1V)	Not provided	Zener-based
Data points collected per cycle and duration	100 points, 1 second	200 points, > 2 seconds	200 points, 1.1 seconds
Total duration per cycle (data from all sensor types)	~30 seconds	>150 seconds	~ 45 seconds

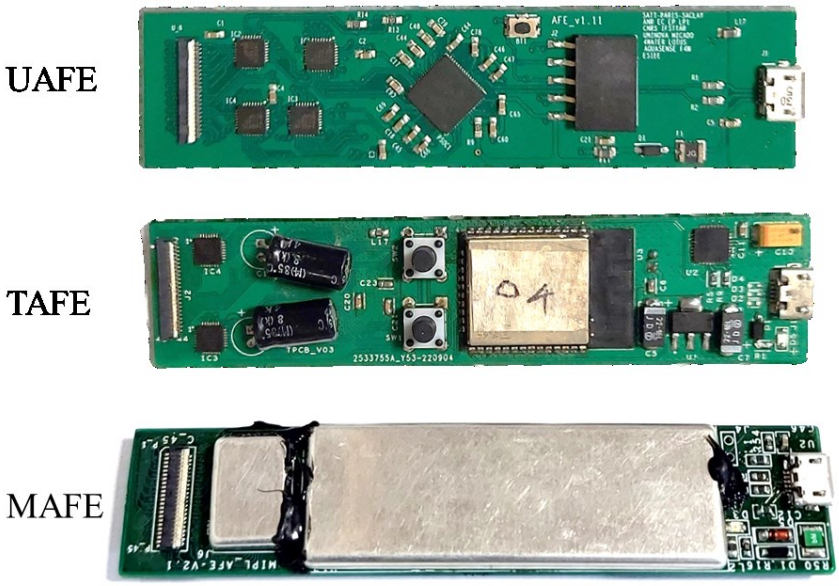

Figure 3.13: Different versions of AFE (From top: UAFE, TAFE, MAFE)

3.3.6 Deployment strategy: Bypass structure and sensor housing

Integrating sensor units into water distribution systems presents two primary methodologies: direct installation into the pipeline or incorporation via a bypass line. Direct installation

boasts several advantages, including minimal requirements for additional components—only a sensor flange is necessary—and straightforward installation procedures. However, this approach introduces significant challenges, notably the need to disrupt the water distribution network during the installation and maintenance operations. Such disruptions can lead to temporary service interruptions, which might not be feasible for critical or continuously operating water distribution systems.

Alternatively, deploying sensors through a bypass structure offers a solution that circumvents the drawbacks associated with direct pipeline integration. This method facilitates ease of installation and maintenance, allowing for sensor access without necessitating a halt in water distribution. The bypass approach involves a dedicated channel that diverts water through the sensor unit, thus maintaining uninterrupted service in the primary pipeline. This method enhances operational flexibility and significantly reduces the risk of service disruptions, making it an attractive option for water systems prioritizing continuous flow and service reliability. Moreover, the bypass installation can be designed to accommodate easy sensor retrieval and maintenance, thereby extending the operational lifespan of the sensor unit within the water distribution framework.

Fouling—a common issue wherein a layer accumulates on the sensor surface, diminishing its sensitivity and reducing its effective lifespan—poses a significant challenge for sensors installed directly in water pipelines. To address this, we have developed a bypass structure featuring four valves and a membrane-encased sensor housing designed to shield the sensor from fouling agents. Detailed in Figure 3.14, this structure incorporates a filter-sandwiched sensor housing strategically positioned within the bypass, surrounded by valves that facilitate the reversal of water flow. The entry and exit nodes are connected to the water distribution pipeline. The pump inside the bypass structure creates a pressure difference, ensuring continuous water flow through the structure. The flow direction is reversed by turning on pairs of valves (V1 & V4, V2 & V3).

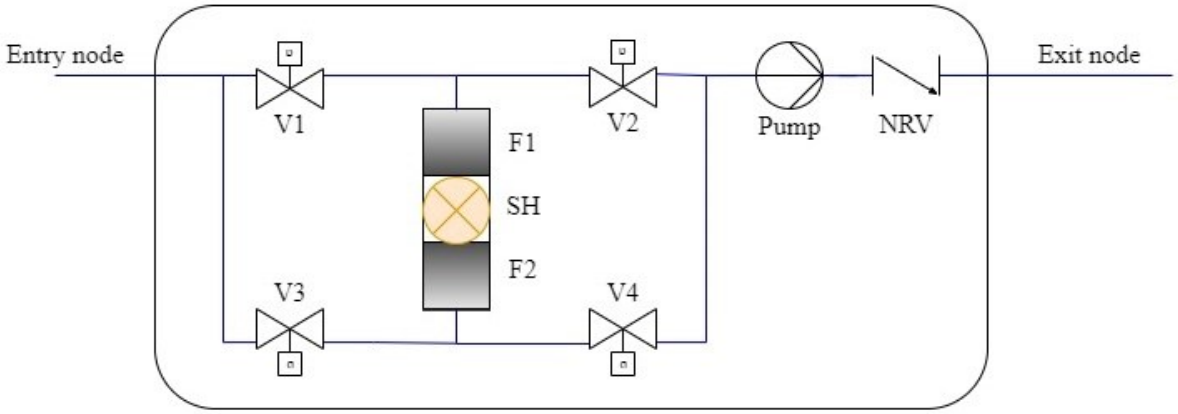

Figure 3.14: Bypass structure outline: V1-V4 = valves; F1, F2 = Filters; SH = sensor housing; NRV = no-return valve

This ingenious design allows one filter to capture fouling agents while the other undergoes backwashing, effectively cleaning the sensor environment. An exploded view of this assembly, presented in Figure 3.15, showcases the intricate design of the sensor housing fabricated in stainless steel (Figure 3.16) for increased EM shielding and easier food-compliance certification.

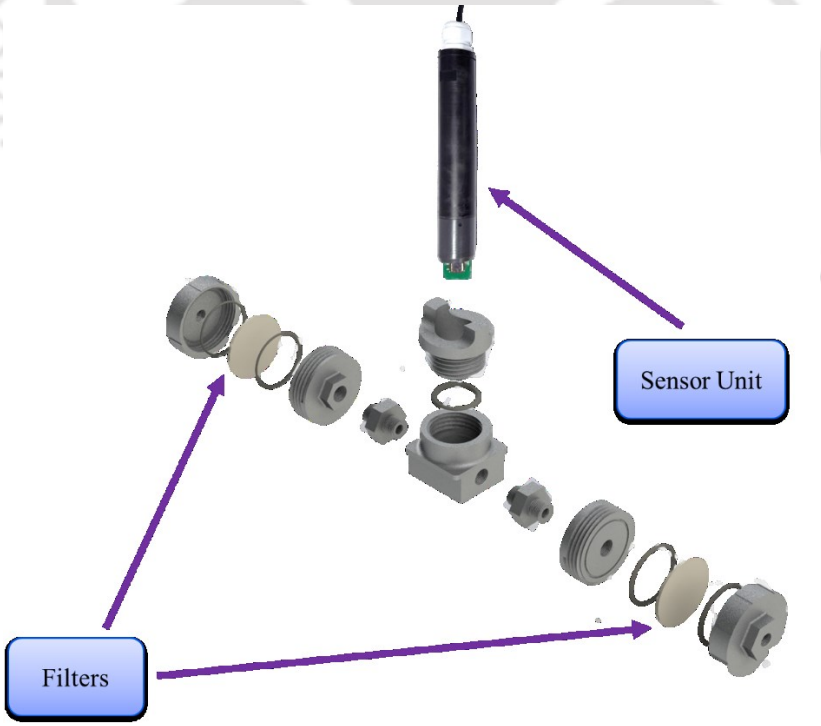

Figure 3.15: Exploded view of the filter-sandwiched sensor housing.

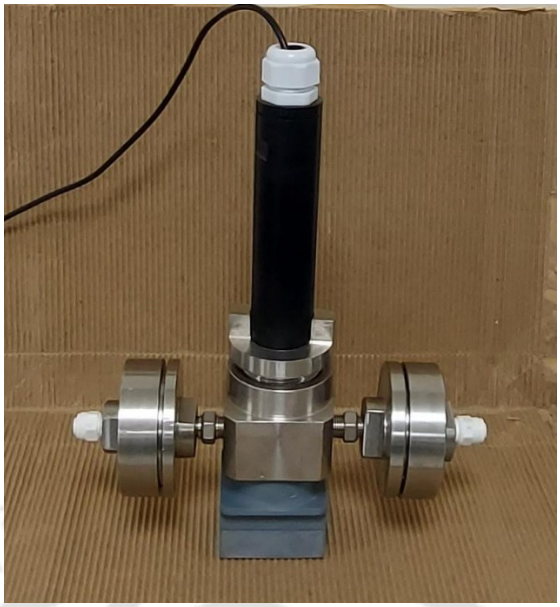

Figure 3.16: Sensor housing fabricated in stainless steel.

Figure 3.17 reveals the compactness of the overall bypass structure. This design enhances sensor longevity by mitigating fouling and ensures that maintenance and sensor access are straightforward, supporting continuous, uninterrupted water monitoring within distribution systems.

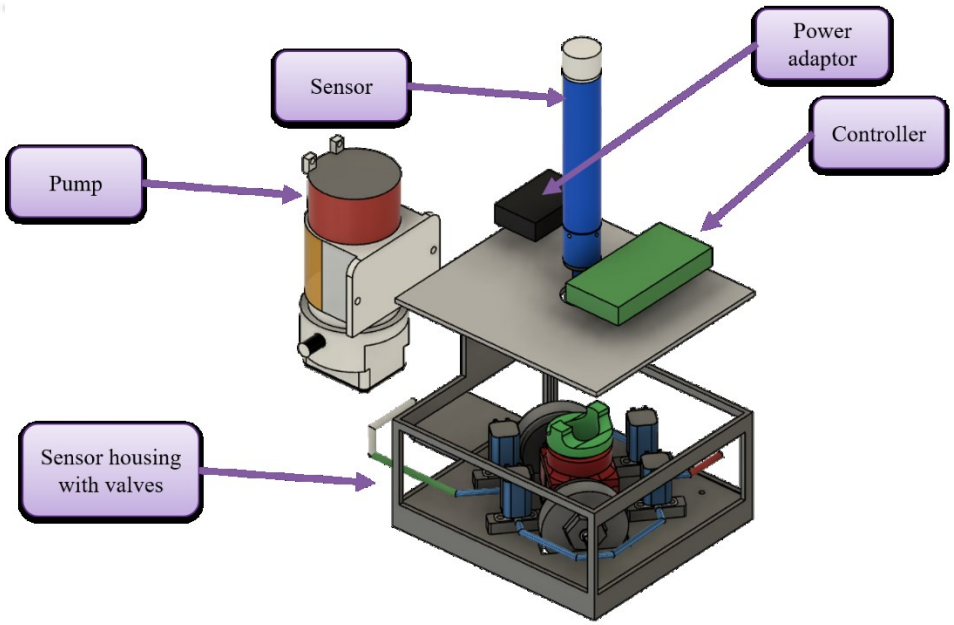

Figure 3.17: Compact design of the sensor bypass structure

3.4 Sensor Data Visualization

The sensor unit, equipped with a sensor chip interfaced with the AFE, establishes a connection to a host system through cables. This host system utilizes custom-developed Python code explicitly tailored to communicate with the AFE's API, thereby enabling data retrieval from the sensor. To accommodate the evolving designs of the AFE, distinct versions of the Python code were developed for each iteration. Once collected, the sensor data is stored in InfluxDB, an open-source time-series database designed for high-performance data storage and retrieval. InfluxDB features an integrated dashboard that facilitates the real-time visualization of the collected data, as shown in Figure 3.18, enhancing the user's ability to monitor sensor outputs efficiently. This dashboard displayed the raw data during field trials and presented the target values computed using the sensor's calibration model which will be discussed in sections 4.3 to 4.5. This approach allowed immediate access to processed data alongside raw measurements, streamlining the analysis and monitoring process.

3.5 Summary

In this chapter, we discussed the initial version of the LOTUS sensor. We also explored the transition from silicon-based to plastic sensor chips and the pivotal role of analog front ends (AFE) in enhancing sensor performance. The initial challenges posed by silicon chips led to the innovative switch to plastic to reduce costs and simplify manufacturing processes, thus making water quality sensors more accessible and cost-effective, particularly in countries like India. We delved into the various iterations of AFE— from UAFE to TAFE and finally to MAFE—each upgrade targeting improvements in sensor excitation and response signal processing, along with integrating features to shield against electromagnetic interference, thereby significantly enhancing data accuracy and sensor longevity.

The discussion also covered the practical aspects of sensor installation, emphasizing the advantages and limitations of direct insertion into pipelines versus employing a bypass structure, allowing seamless maintenance without disrupting water distribution. Additionally, creating a bypass structure to solve the problem of sensor fouling employing a reverse water flow mechanism ensures the longevity and reliability of sensor readings in real-world applications.

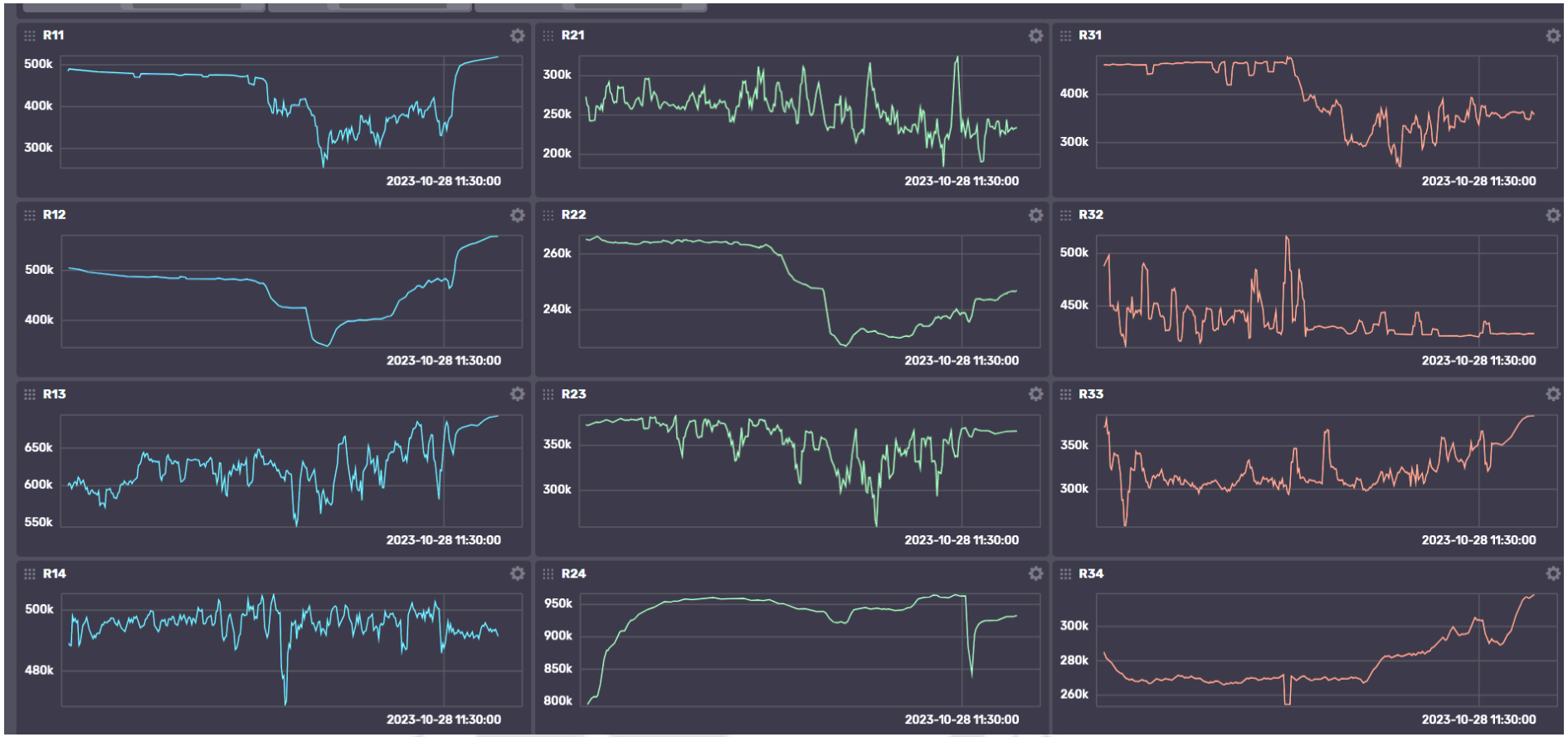

Figure 3.18: InfluxDB dashboard for sensor response

Finally, we discussed the development of a simple data pipeline and minimal visualization framework, utilizing Python for data acquisition from the AFE and InfluxDB for data storage and real-time display of raw sensor data, as well as the target parameters obtained as a result of calibration models.

4

Sensor Characterization and Calibration

Abstract: *This chapter details the experiments with the LOTUS sensors for sensor characterization and calibration. Initial experiments were conducted with the silicon version, and the results, although they meet the target numerically, lack reliability. The experiments with the plastic version, from the single electrode sensor to the field-ready LPCB version, show promising signs for monitoring active chlorine and pH along with conductivity and temperature. However, the inconsistency in the sensor performance in Sense-city and the higher uncertainty associated with the sensing necessitates more extensive experiments with multiple devices and further refinement of the models. Though the results lack consistency and are derived from only 2 multiparametric sensors in the field and only from an experiment duration of 5 months, of which only 60% are useful (~ 468 hours of total 3760 hours of data has errors), the information gained, and conclusions drawn from these experiments and analysis shows the potential of the sensor in real-time multiparametric sensing and consequently, will be used in the upscaling and productizing the LOTUS sensor.*

4 Sensor Characterization and Calibration

Sensor characterization is essential for developing and deploying single-parameter and multi-parameter sensors engineered to measure distinct environmental or physical factors. This process involves evaluating key performance indicators of a sensor, including its sensitivity, selectivity, range, and response time, which becomes particularly critical when sensors are tailored for precise tasks like monitoring water quality. In applications of this nature, factors such as the sensor's capability to detect pollutants at low levels, its long-term stability, and its resistance to interference from non-target substances are of utmost importance. Characterization arises from guaranteeing the sensors' dependability, precision, and suitability for real-world scenarios. By comprehending a sensor's operational boundaries and behaviour under diverse circumstances, optimizing its utilization in specific settings becomes feasible, ensuring the data's accuracy and reliability.

Identifying and quantifying these characteristics typically involves a series of standardized tests and procedures under controlled conditions to simulate real-world applications as closely as possible. For example, sensitivity can be assessed by comparing the sensor's output to known amounts of a target analyte. The sensor can measure lower limits with high precision because of its high sensitivity. Selectivity is measured by how the sensor reacts to the analyte to detect when other substances could interfere with it. Stability and response time are tested by exposing the sensor to the analyte under different conditions. These steps are essential to establish the working range of the sensor and make sure that it suits the particular needs of its application.

On the contrary, sensor calibration sets a defined correlation between the sensor's reaction and the actual value of the measured parameter during characterization. This typically includes subjecting the sensor to established reference values and aligning its output to correspond precisely. The developed mathematical correlation is called a calibration model, and it also rectifies the inherent deviations or inaccuracies in the actual sensor's output due to external noise, which will guarantee the accuracy and reliability of the collected data. Calibration of the sensor is carried out prior to deployment and at regular intervals during the sensor's operational lifespan to address any alterations in sensor performance as time progresses.

When dealing with multi-parametric sensors capable of measuring multiple parameters simultaneously, the characterization and calibration process becomes inherently more

complex. Unlike monoparametric sensors, multi-parametric sensors must be characterized and calibrated for each parameter they measure, often requiring a multidimensional approach to capture the interactions among the sensing elements and environmental conditions. This complexity introduces challenges not applicable in monoparametric sensor calibration, such as cross-sensitivity and the potential for interference between parameters. Applying monoparametric characterization and calibration techniques to multi-parametric sensors may lead to inaccurate measurements and compromised data quality. Therefore, it is essential to use multi-parametric calibration techniques that consider the interactions among the sensing elements and environmental conditions to ensure accurate measurements and high-quality data.

Currently, the characterization and calibration of multi-parametric sensors involve advanced methodologies emphasizing data fusion, multivariate techniques, and machine learning algorithms. Data fusion plays a critical role in integrating and interpreting the complex datasets generated by multi-parametric sensors, allowing for a more comprehensive understanding of the monitored environmental or physical parameters. Multivariate methods reveal the complex connections between multiple sensor's responses, allowing the precise identification and quantification of specific variables within a complicated data collection. Machine learning algorithms enhance calibration by learning from the sensor's output and making predictive adjustments to compensate for cross-sensitivity and interference between parameters. The combination of multivariate and machine learning algorithms can provide a better calibration model for multi-parametric sensors.

4.1 Sensor Characterization experiments

The characterization of the LOTUS sensor involves three distinct stages. The first stage called the beaker experiments, consists of conducting experiments in a static environment by utilizing a beaker containing a known volume and concentration of target analytes in tap water. This phase aims to establish the fundamental functionality and validate the sensor concept. Moving on to the second stage, known as the lab-loop experiments, tests are carried out under flowing conditions by installing the sensor in a bypass structure (Figure 4.2) connected to a tank and pump with appropriate valves. This setup includes the LOTUS sensor and reference sensors to monitor each target parameter. The third stage encompasses real-world pipeline experiments where a LOTUS sensor is integrated into an active water pipeline system using a hot tapping technique. This system replicates a controlled environment similar to an operational water distribution network. The experiments were

conducted within Sense-City, an experimental facility at Gustave Eiffel University that provides a realistic setting for water distribution network (WDN) simulations and sensor deployment scenarios. The subsequent section offers a detailed overview of each phase involved in characterizing the sensors.

4.1.1 Laboratory beaker experiments

The initial phase of the study involved submerging the sensors in a beaker filled with a specified amount of tap water, as shown in Figure 4.1, to monitor their responses over an extended period. This setup aimed to reach a point of sensor stabilization, characterized by minimal changes in sensor output, a phase influenced by the sensor's air resistance, which varies with the number of ink layers applied. Specific analytes were incrementally introduced into the beaker to examine the sensor's responsiveness to variations in critical parameters such as temperature, conductivity, active chlorine, and pH. Aqualabo Portable pH and Conductivity probe was used to measure the pH and Conductivity of the sample [220]. This reference sensor provided a pH range of 0-14 with an accuracy of ±0.1 pH, a conductivity range of 0 to 2000 µS/cm with a ±1% accuracy, and a temperature range of 0 to 50°C with an accuracy of ±0.5°C. The active chlorine concentration was measured with the photopod provided by Aqualabo, using the DPD-based colorimetric method, having a range of 0.5 to 60 mg/L and an accuracy of ±0.1mg/L. This phase also explored the effects of environmental factors like light and external noise and the challenges in achieving a homogeneous analyte distribution without manual stirring. Experiments were devised for each parameter and performed to gain insights into the sensor's performance.

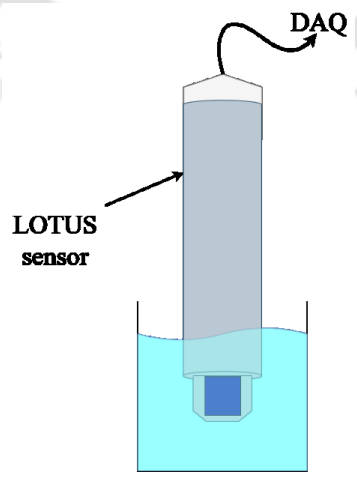

Figure 4.1: Preliminary sensor testing: In a beaker

4.1.2 Laboratory experiments

The laboratory flow experiments were conducted using two setups: (1) a basic configuration connected to a 10-litre water tank and bypass structure, including a reference sensor, pump, and data acquisition through a host system, and (2) a second setup integrated with a real WDN pipeline, along with a bypass structure, industrial DAQ, and commercial sensors.

4.1.2.1 Sensor Bypass unit without WDN pipeline

After establishing the sensor response data from beaker experiments, the characterization experiments are advanced to investigate the sensor's performance under flowing water conditions. This was facilitated by establishing a simplistic flow loop system, where water was circulated from a 10-litre recirculation tank through the filter (to remove suspended particles from damaging the sensor), LOTUS sensors and reference sensors arranged within the loop and then back to the tank. An outline of the simplistic laboratory water loop is visualized in Figure 4.2, and the components' specifications are summarized in Table 4.1. Throughout this continuous operation, analytes were manually added to the tank to study the sensor's behaviour in a dynamic fluid environment.

Table 4.1: Specifications of the Laboratory water loop setup

Sno	Components	Specification	Supplier
1	Circulation pump	24 VDC, 140 psi, RO booster pump, flow range 0 to 60L/h	Hi-Tech
2	Reference chlorine sensor	Model: BH5530D, 0 – 2mg/L, 0 – 10 mg/L, 0 – 20 mg/L Accuracy: ±1% FS	Fluidings
3	Reference pH sensor	Model: BH1701D, 0 – 14 pH, ±0.05 pH	Fluidings
4	Ref. Conductivity and Temperature sensor	Model: BH3701D, 0 – 30 mS/cm and 0 – 80°C, ±1%	Fluidings
5	Heating and Cooling unit	5L circulating water bath, -10°C to 100°C, ±0.5°C	ANM Industries WB2000V
6	Host System	Windows desktop, I5 processor, 32GB RAM	DELL

The circulation pump was operated at 15V to circulate water throughout the loop. The reference sensors were connected to the host system via ModBus protocol.

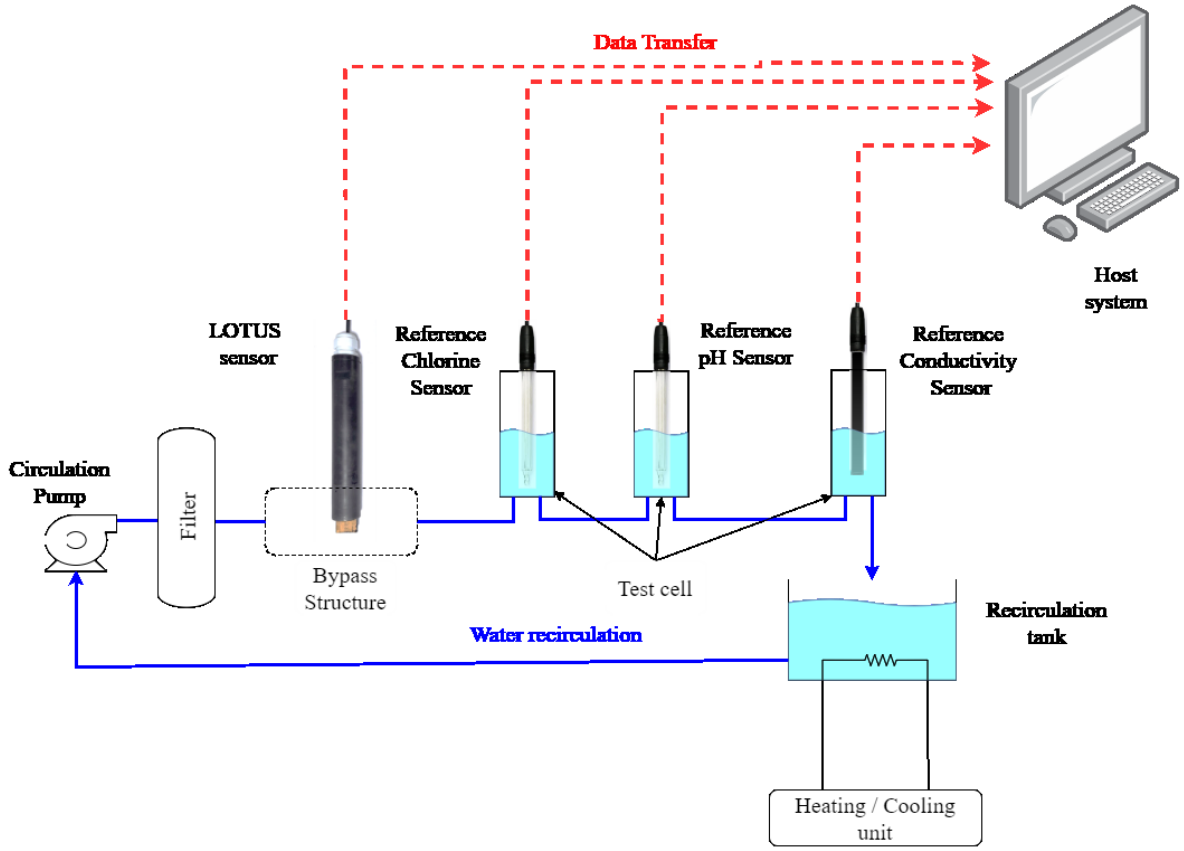

Figure 4.2: Laboratory Water Loop

4.1.2.2 IITG Water lab – Water loop

The IITG Water Lab – Water Loop, an 80-litre laboratory water loop, has been established at the Indian Institute of Technology, Guwahati. Its objective was to conduct thorough mono-parametric and multi-parametric studies on LOTUS sensors placed in bypass structure amidst varying water flow conditions and across different water model simulations. In addition to its experimental setup, the IITG Water Lab provides access to high-end analytical tools, including Ion Chromatography, Mass Spectrometry, Atomic Spectroscopy, and High-Performance Liquid Chromatography (HPLC). The lab also has advanced computing systems to handle the demands of extensive computational tasks and intricate data analysis processes.

As outlined in Figure 4.3, the lab's test bench system is designed to accommodate up to three LOTUS sensors within individual bypass structures alongside reference online sensors. A summary of the components of the IITG Waterloop is provided in Table 4.2.

The assembly is automated and managed via computer systems, facilitating seamless data collection. The operation of the test bench is divided into two primary phases: The solution Preparation process and the Sensing Cycle.

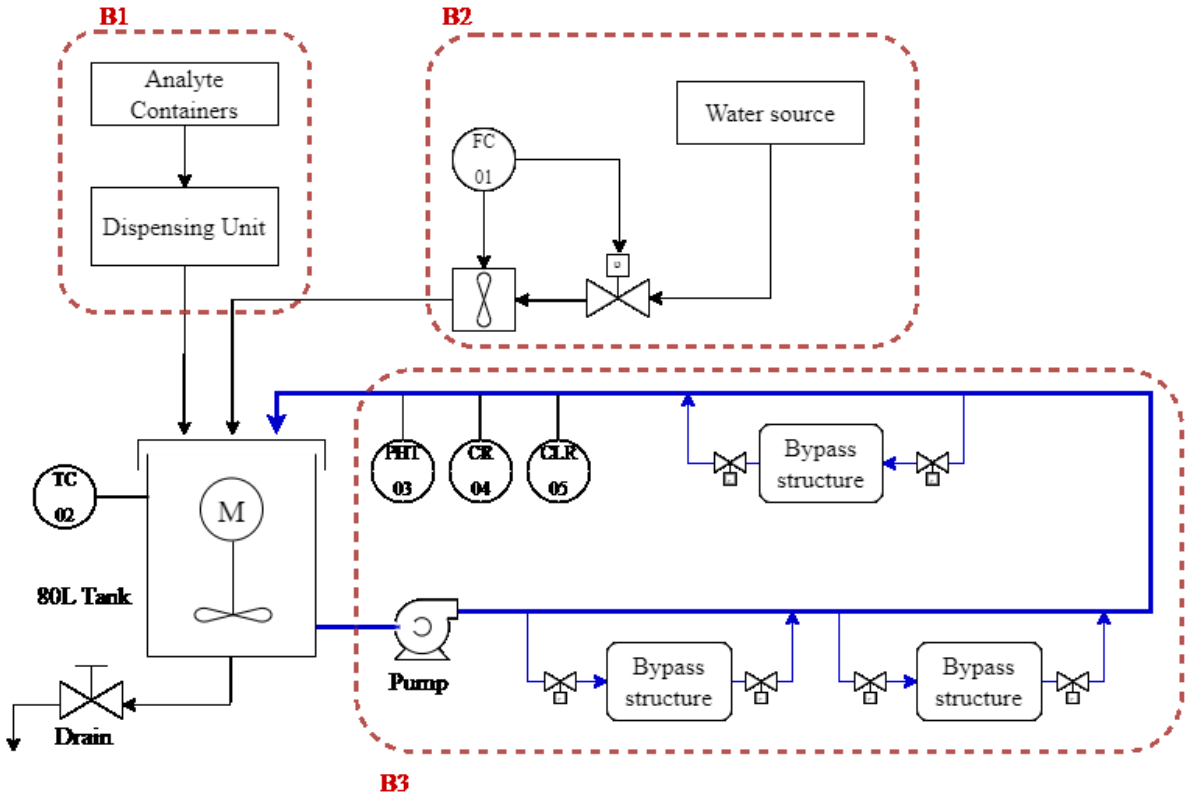

Figure 4.3: IITG Water lab - Water Loop

Solution Preparation Process:

- Analytes (six or more as needed) are prepared in designated quantities and stored in separate containers, referred to as Analyte Containers in Figure 4.3 – B1.
- A dispensing mechanism (Figure 4.3 – B1), equipped with Peristaltic pumps, accurately measures and dispenses these analytes into an 80-litre collection tank controlled by a computer through analog input (milliamps or mV).
- Water is introduced into the 80-litre tank from a chosen water source—distilled, tap, or a special lab-prepared matrix (Figure 4.3 – B2).
- The water volume in the tank is controlled by the flow controller FC01 via the PC
- After adding water and analytes, the solution is mixed to achieve a uniform distribution. This mixing is accomplished without contact, and temperature

adjustments are made as per the temperature controller, all under computer control to reach the desired homogeneity and temperature within a specific range.

- The conclusion of this stage is marked by the prepared solution, adjusted to the required analyte concentrations and temperature settings.

Table 4.2: Specifications of IITG Waterlab - Water loop setup

Sno	Components	Specification	Supplier
1	Peristaltic pumps	0.2 to 3600 ml/min, Resolution: ~0.01ml Accuracy: ±0.5% ModBus RTU protocol	Shenchen
2	Flow control system	To control the water in the 80L tank, up to 100mL precision	Avgarde Systems Pvt Ltd
3	Tank Temperature controller	Control range: 20°C to 60°C, accuracy: ±0.5 °C	Avgarde Systems Pvt Ltd
4	Pump	Rated flow: 3 m^3/h Rated head: 92.5 m Max pressure: 25 bar	GRUNDFOS
5	PHT 03 – pH transmitter	2 – 12, accuracy: ±0.2 units	Avgarde Systems Pvt Ltd
6	CR 04 – Conductivity Recorder	0 – 3000 µS/cm, accuracy: ±1%	Avgarde Systems Pvt Ltd
7	CLR 05 – Chlorine sensor	BH5530D, 0 – 2mg/L, 0 – 10 mg/L; accuracy: ±1%	Fluidings
8	80L Tank	Food grade, Stainless steel tank	Avgarde Systems Pvt Ltd
9	DAQ system	PCIe-6535B Digital I/O Device, PCIe-6353 Multifunction I/O Device	National Instruments

Sensing Cycle (Data Acquisition Phase):

- Completion of the solution preparation triggers the start of the sensing cycle. The automated system activates the pump, allowing the mixed solution to circulate through the sensing system (Figure 4.3 - B3) and back to the 80-litre tank, establishing a continuous flow loop.
- The water flow pathway encompasses the pump through the sensing system—covering reference and LOTUS sensors and returning to the 80-litre tank.
- Throughout this cycle, parameters such as temperature, pressure, flow rate, conductivity, and pH are monitored and recorded at set intervals, along with readings from the LOTUS sensors installed in the bypass structure, ensuring comprehensive data capture for analysis.

4.1.3 Controlled Small-scale Water Distribution Network: Sense-city

Located at the Université Gustave Eiffel in Marne-la-Vallée, within the eastern Paris area, the Sense-City [221] water loop is a remarkable facility for validating the reliability of water sensors under quasi-real conditions. The successful completion of trials within the Sense-City setup is pivotal for validating Technology Readiness Level (TRL) 5 for sensors designed for drinking water networks. Figure 4.4 provides a detailed schematic view of the water loop within the Sense-City chamber, highlighting its integral components.

The hydraulic setup is a crucial feature of the Sense-City water loop, consisting of a 1000L tank filled with tap water, which is in contact with a heat exchanger having the thermal regulator to maintain temperature control. The water loop extends to a total length of 44 meters from the tank outlet to the inlet. The internal diameter of the pipes is 93.8mm, accommodating a total water volume of 310L, with pipes made from PVC/PEDH material. A 1.20 m deep visit chamber is strategically placed in the middle of the Sense-City chamber, equipped with two saddle clamps for mounting the LOTUS sensor units. For draining, the water can be directed into Sense-City's wastewater network, which leads to the groundwater reservoir or the public wastewater system.

The thermal regulation within the loop is achieved through a system that circulates a calorific fluid via copper heat exchanger loops at the bottom of the tank. This setup allows for temperature adjustments ranging from 5°C to 40°C, ensuring a controlled environment for testing the sensors under various thermal conditions.

The loop can be operated at desired pressure and flowrate by manipulating the valve, and it is designed to operate with a flow range from 10 m^3/h (0.1 m/s) to 80m^3/h (0.8m/s), with a maximum pressure of 16 bar at 0.1 m/s.

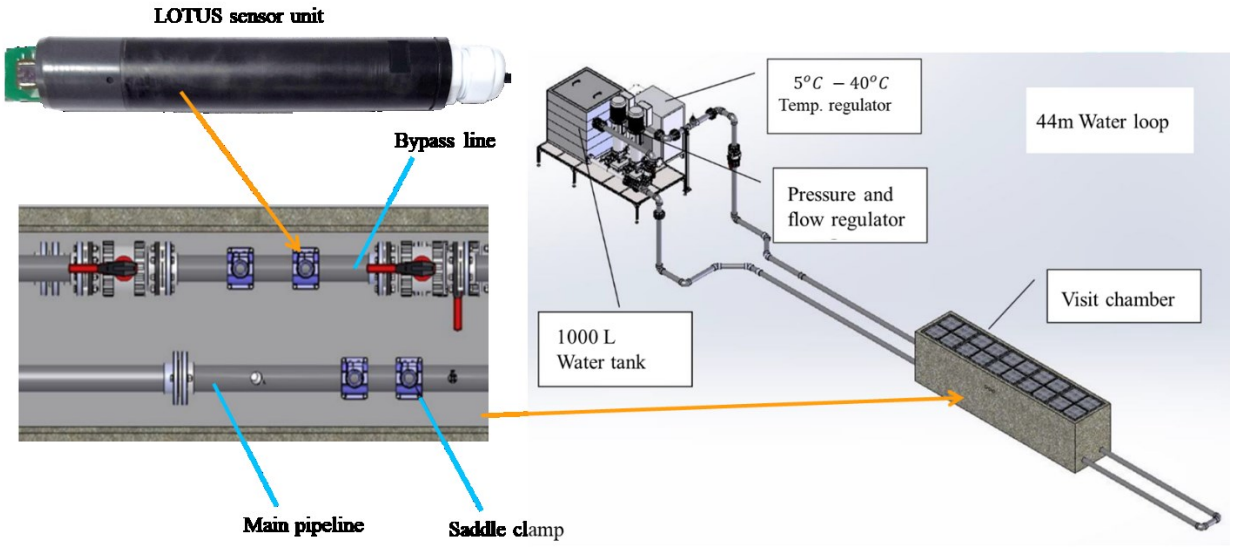

Figure 4.4: Schematics of Sense-city

With reference sensors, online monitoring is integral to the Sense-City water loop. Continuous data collection is facilitated through the Sense-City supervision system (SCSS), allowing for real-time monitoring and analysis. The collected data can be exported in CSV format or accessed directly from the SCSS, offering comprehensive insights into sensor performance under simulated operational conditions.

4.1.4 Significance of Sense-City Experiments

The experiments performed within the Sense-city were designed to accomplish a series of technical objectives vital for enhancing the sensor's water quality monitoring capability:

Real-Time Sensor Performance Assessment: Initial experiments were dedicated to examining the sensor's dynamic response to predetermined variations in active chlorine concentrations and pH levels, providing data for calibration model construction.

Sensor Performance Evaluation in Field Condition: Subsequent experimental phases were tailored to replicate field conditions, notably intermittent flow scenarios characteristic of Indian water distribution networks, to ascertain sensor efficacy in actual operational environments.

Sensor Characterization Studies: A detailed analysis of sensor properties critical for deployment scenarios was conducted using the collected experimental data. This encompassed examining the sensor's resilience to external variables, including temperature variations, conductivity shifts, and electromagnetic disturbances, and assessing key performance indicators such as sensor longevity, dependability in sustained flow conditions, durability in static conditions with a subsequent reintroduction to flow, and susceptibility to sensor drift.

Calibration Model Validation: The culmination of the experiments aimed at the real-time validation of the devised calibration model, ensuring its accuracy and applicability in reflecting precise water quality metrics under diverse environmental conditions.

4.2 Experiments with LOTUS sensor- Silicon version

In a previous study, the LOTUS sensor silicon version was initially characterized in a beaker setup by researchers at Gustave Eiffel University. The results of calibration experiments conducted in beakers are documented in the referenced article [38], [222], where a Bayesian calibration model based on previous experiments is presented. Due to the extended time required to reach a steady state, the researchers carried out only a few experiments, and the calibration model developed was in the context of small data.

The main objective of the Sense-city experiments was to evaluate the performance and reliability of the LOTUS sensor silicon version in a realistic urban water network. The experiments aimed to test the sensor's response to various water quality parameters, such as temperature, conductivity, chlorine, and pH, and the sensor's stability and durability over time. The experiments also sought to compare the sensor's readings with reference and laboratory measurements and assess the sensor's calibration model under different operational conditions. The sensor was exposed to simulated scenarios that mimic the challenges and opportunities of smart water management in urban environments by conducting experiments at the Sense-city facility. During the experiment, the sense-city water loop was not operated for short and long durations to replicate the intermitted water supply commonly observed in the Indian water distribution network. The sensor performance is analyzed only during the sense city operation.

The silicon version of the LOTUS sensor was tested for real-world performance at the Sense-City facility, where six distinct units were assessed over an extended period. The setup could

simultaneously host a maximum of two sensor units in an active pipeline using a hot-tapping mechanism. Despite some sensors being in operation for several weeks to over a year, they were only intermittently exposed to flowing water due to operational limitations. When water flow was initiated, it reached speeds of approximately 0.3 m/s with a pressure of around 1 bar. Throughout the various testing phases, the sensors were exposed to fluctuating environmental conditions, including temperatures ranging from 15°C to 25°C, conductivity levels between 870 µS/cm and 1270 µS/cm, chlorine concentrations from 0 to 5 mg/L, and pH values fluctuating between 7 and 8.5. Due to the nature of the ON-OFF controller operation, the temperature varied within a range of ± 2°C for the given setpoint. This variation introduced a periodic interference pattern in the sensor response, which had to be removed during preprocessing.

When the loop is not running, the temperature controller is also disabled, which brings the temperature of the water in the loop to the environment temperature, which was relatively low (10°C) due to the winter season. The Conductivity variations seen in the dataset can be attributed to the varying water quality, environmental conditions, and cross-contamination via minute leakage in the loop (which was fixed after the conclusion of the experiments). The pH levels varied by adding hydrochloric acid (HCl) and sodium hydroxide (NaOH). The chlorine concentration was varied by adding a predetermined amount of Sodium Hypochlorite (NaOCl). These chemical additions were made to the main tank directly.

4.2.1 Description of Dataset

Sense-city experiments with Version 1 of the Lotus Sensor have provided an extensive dataset, as outlined in Table 4.3. The experiments were performed as part of the LOTUS project by the research team at Gustave Eiffel University. As per the experimental protocol, a known quantity of NaOCl was added to the tank at a frequent interval, the loop was under circulation, and the temperature was maintained using the ON-OFF controller. During this period, the performance of six LOTUS sensors was evaluated.

The dataset obtained from a single LOTUS sensor, consisting of 20 chemiresistors, three conductivity sensors, and two temperature sensors, was collected through the initial version of AFE (UAFE). Data was collected consistently at intervals ranging from every 3 minutes, independent of whether the loop was on or off. These conditions represent the intermittent water supply commonly found in Indian WDNs.

The data collection period extended from March 2021 to July 2022, providing a comprehensive overview of the sensor performance timeline. The table specifies the LOTUS sensor number, the duration of data collection in days and the total number of data points recorded for each sensor. Additionally, it outlines the operational ranges observed for different parameters during the testing period, such as pH and chlorine concentrations and temperature and conductivity levels.

Table 4.3: Summary of the dataset collected from 6 units of LOTUS sensor (silicon) in Sense-city

Year	**LOTUS sensor No**	**Days**	**Dataset size**	**pH**	**Chlorine (mg/L)**	**Temperature (°C)**	**Conductivity (µS/cm)**
2021	141	35	15,000	7.6 – 8.7	0 – 5	14 – 20	980 – 1050
	80	5	5,000	8.1 – 8.4	0 – 1.1	14 – 20	990 – 1030
	35	9	4,000	8.1 – 8.5	0 – 1.4	17 – 21	970 – 1030
2022	AS001	55	37,000	7.0 – 8.4	0 – 5	15 – 24	750 – 1270
	141	48	35,000	7.0 – 8.1	0 – 5	14 – 24	1010 – 1090
	SH007	42	31,000	7.2 – 8.1	0 – 5	15 – 24	1030 – 1100
	SH005	10	6,000	7.0 – 8.4	0 – 4.2	15 – 23	1030 – 1270

Table 4.4: Sensor testing sequence in the Sense-City experiments

2021			**2022**		
Date	Slot A	Slot B	Date	Slot A	Slot B
25 Mar	35	-	19 May	-	141
31 Mar	35	80	20 May	AS001	-
02 Apr	-	80	23 May	AS001	SH005
09 Apr	141	80	25 May	AS001	-
11 Apr	141	-	07 June	AS001	SH007
13 May	-	-	06 July	-	-

The Sense-city can accommodate 2 sensors at a time, so the damaged ones in the slots were replaced with new ones. The sequence of sensors in the sense-city loop is outlined in Table 4.4.

Due to variations in water quality and different quantities of chlorine addition, the actual concentration of active chlorine varies from 1 – 5 mg/l since few of the sensors were operated when the maximum actual concentration reached around 1.5 mg/l (like 80, 35), so the range of these sensors is observed to be lower than the others.

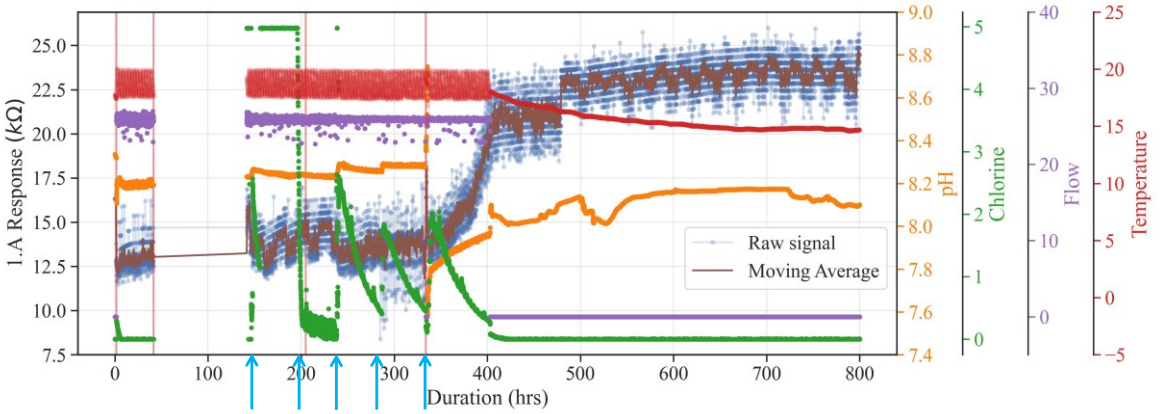

Figure 4.5: Response from 1 chemiresistor (1A) from the silicon version LOTUS sensor (chip 141) during experiments in Sense-city

For illustrative purposes, Figure 4.5 showcases data from one sensor - specifically, the output (measured in kΩ on the left vertical axis) from chemiresistor 1A on chip 141 plotted over time (indicated by blue dots) alongside its moving average (represented by a brown line). Concurrently, the graph illustrates environmental conditions within the water loop over time, using a quartet of color-coded axes on the right side for pH (orange), chlorine (green), flow rate (violet), and temperature (red).

4.2.2 Data processing

The data collected with the Lotus sensor in sense-city encountered challenges due to insufficient EM shielding around the AFE and certain design flaws. These shortcomings led to considerable signal noise, consistently exceeding 1%. The intense EM interference experienced in the SenseCity test environment and the design issues were pinpointed as primary contributors to this problem.

The impact of this noise manifested in the sensor's raw data in several disruptive ways, affecting the precision of the readings. Notably, it caused:

- Issues with multiline signals complicate the interpretation of the data.

- Baseline shifts, which abruptly altered the reference point for measurements.
- The combination of double lines and noise adds complexity and reduces the sensor output's clarity.
- Persistent oscillatory noise (one of the contributors to this periodic noise is the ON-OFF controller used for temperature control in the loop) threatens the long-term consistency of the measurements.
- Variable noise levels complicate the task of setting a stable benchmark for analysis.
- Elevated noise that could mask the accurate sensor data.

The same has been illustrated in Figure 4.6. Recognizing that such disturbances could also arise in field conditions, it became clear that addressing these issues was crucial for the next version of the device. Consequently, the Lotus Sensor version 2 incorporates strategies to reduce EM interference, significantly improving the sensor's performance, reliability, and accuracy in capturing real-world data.

Significant efforts were made to address these challenges through various data preprocessing methods. Initially, baseline jumps were tackled by re-aligning the data baseline to its original position. This was achieved by regressing the mean of the data, segmented into windows, against the time axis. The original and adjusted data are illustrated in Figure 4.6, with the bottom part displaying the moving standard deviation. This visualization confirms that while the signal's mean has been adjusted, its noise structure remains unchanged.

Following the correction of baseline jumps, the signal underwent further refinement. This involved smoothing through a moving average technique and filtering out noise [41]. Specifically, high-frequency disturbances were attenuated using a low-pass filter, and 'salt and pepper' type noise was eliminated with a median filter.

The resulting dataset was then transformed into a relative response R_{rel} using the formula:

$$R_{rel} = \frac{R_s - R_0}{R_0} \tag{4.1}$$

where R_s represents the raw signal response, and R_0 denotes the baseline response. This relative change is most widely used in the domain of chemical sensing [127], [137]. This processed data forms the basis for subsequent analyses.

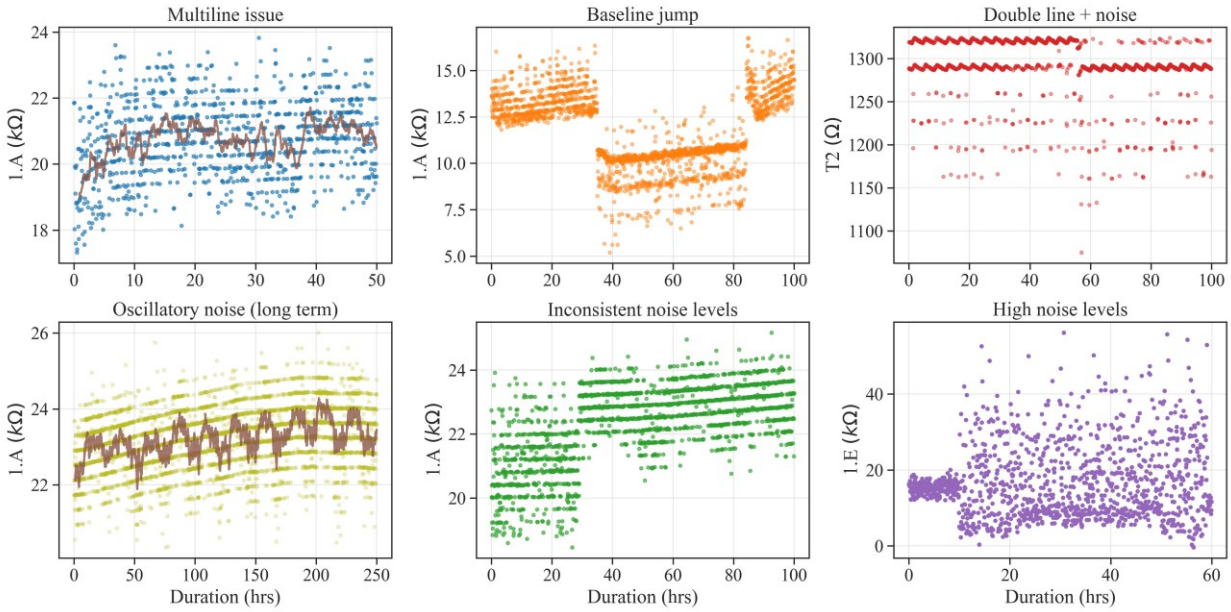

Figure 4.6: Signal artefacts in the dataset collected during Sense City experiments

Baseline Jump correction:

The baseline jump was corrected by using simple linear regression. Initially, a few anchor points are selected before the start of the baseline jumps. These anchor points were then regressed against the time axis (any linear index would suffice). This regression is then used to correct the raw data's anchor points. Once these anchor points are identified, each data point is pulled toward the closest anchor point (closeness is calculated according to the linear index, which, in this case, is the time axis). In the present case, the anchor points are chosen as the average of the signal in every 5-hour time block. This 5-hour mark was chosen as a balance between the computational overhead and loss of information (lower time blocks lead to more computation, while higher time blocks would lead to losing the information in the signal).

In Figure 4.7, the smaller blue dots represent the actual signal, and orange crosses represent the corrected signal. The bigger black dots represent the chosen anchor points, and the purple triangles show the corrected anchor points.

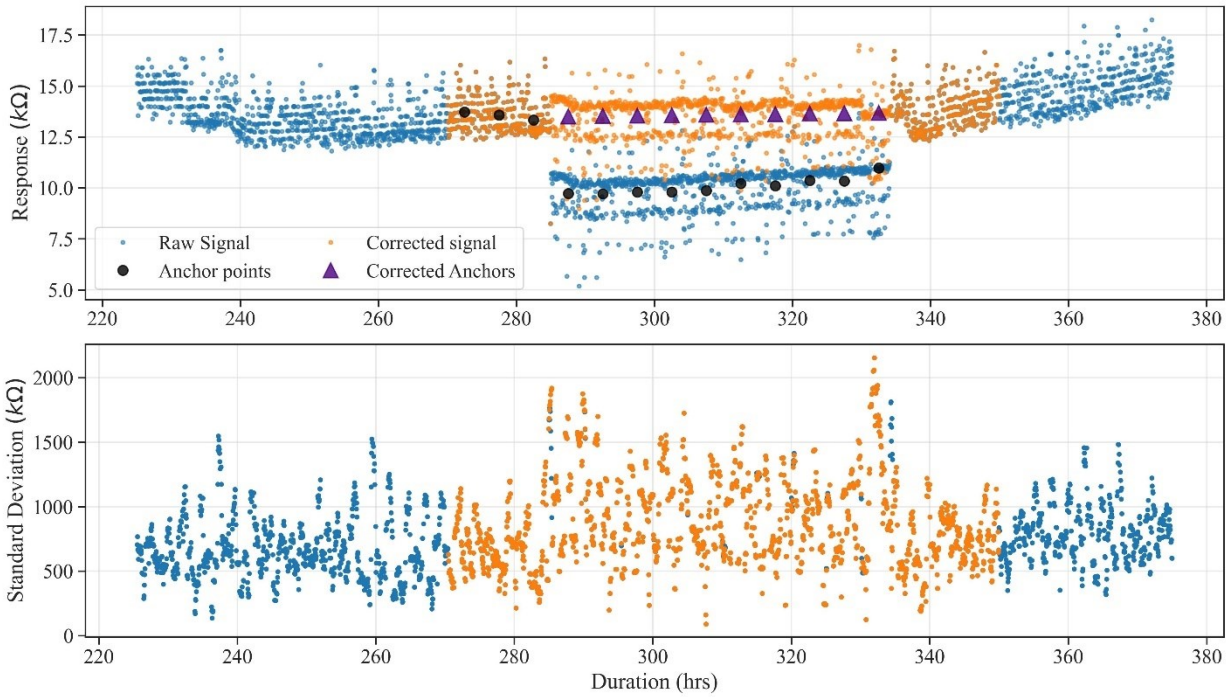

Figure 4.7: Baseline Jump correction in Sense-city data

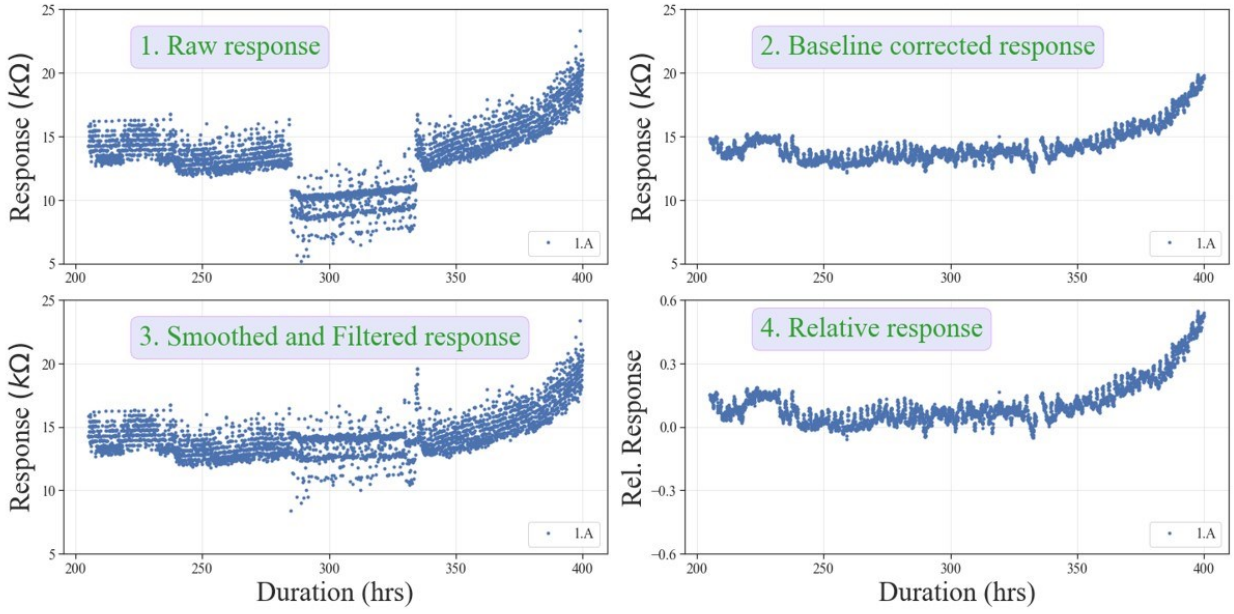

Figure 4.8: Data pre-processing steps adopted in the model development pipeline

4.2.3 Calibration Model Development

The processed data from the silicon version LOTUS sensor underwent further analysis to develop models to estimate temperature, conductivity, active chlorine, and pH levels. This involved dividing the dataset randomly into training and testing subsets twenty times, maintaining a 70-30 split. The training subset was used to estimate model parameters, whereas the test dataset was employed to assess the model performance.

The model structure was defined using first principles, starting with a simple linear relationship and increasing complexity based on sensor response. Optimal parameters for each target were identified by assessing explained variance, and predictions were made using the selected models. The accuracy of each model for the target was evaluated based on the Mean Absolute Error (MAE), which is calculated by averaging the absolute differences between the actual observed values y_i and the model's predictions $\hat{y}_i$.

- Mean Absolute Error:

$$MAE = \left(\frac{1}{n}\right) \sum |y_i - \hat{y}_i| \tag{4.2}$$

- Compared with Mean prediction (Naïve model with $R^2 = 0$):

$$Naive\ MAE = \left(\frac{1}{n}\right) \sum |y_i - \bar{y}| \tag{4.3}$$

In this context, y_i represents the observed or measured value of interest, $\hat{y}_i$ denotes the model's prediction, and $\bar{y}$ is the average of all observed values. The efficacy of our model was gauged against the naïve model, which predicates that every prediction mirrors the mean of all observations (yielding an R^2 value of 0) to evaluate our model's performance.

The model development discussed above is affected by the level of data pre-processing carried out. Though these processes removed artefacts and improved the signal (reducing the noise), was conducted manually rather than automated. Secondly, the performance was estimated using random splits and not with temporal splits (where the history data is used to develop the model and present data is used to estimate the performance, which can be generalized for future performance with some uncertainty). These can lead to the risk of overfitting of the models.

4.2.4 Note on Model development

In the silicon version, we utilized simple linear models for the calibration of temperature and conductivity sensors, and higher-order linear models for predicting chlorine and pH levels. The primary rationale for limiting the model set to linear models stems from the limited computational power of the Analog Front End (AFE). In this initial setup, all target computations were performed directly on the AFE, and only the results were transmitted to the host system for user display. This setup necessitated models that were computationally less demanding to ensure efficient real-time processing.

In the upgraded versions, even though the computational capabilities of the AFE improved compared to the earlier setup, we continued to use linear models to maintain a lower memory footprint and ensure real-time monitoring. This decision led us to exclude advanced filtering techniques and complex machine learning models, such as those based on neural networks and deep learning. Looking ahead, we plan to further enhance the DAQ system and explore the potential of more advanced models as computational resources allow.

4.3 Calibration Models for LOTUS Sensor - Silicon version

4.3.1 Temperature Sensor Model

Figure 4.9 compares the sensor's response to temperature variations in scenarios with and without water flow. The data illustrate that flow significantly increases the noise level in the sensor's output. Despite this, it is feasible for both scenarios to establish a linear relationship between the reference temperature T_{ref} and the sensor's output T_i (T_1 and T_2), as shown in eqn (4.4).

$$T_{ref} = T_i \times slope + bias \tag{4.4}$$

Such a model facilitates temperature prediction from the sensor data, yielding an MAE of 0.34°C in flowing conditions and 0.09°C when the water is static.

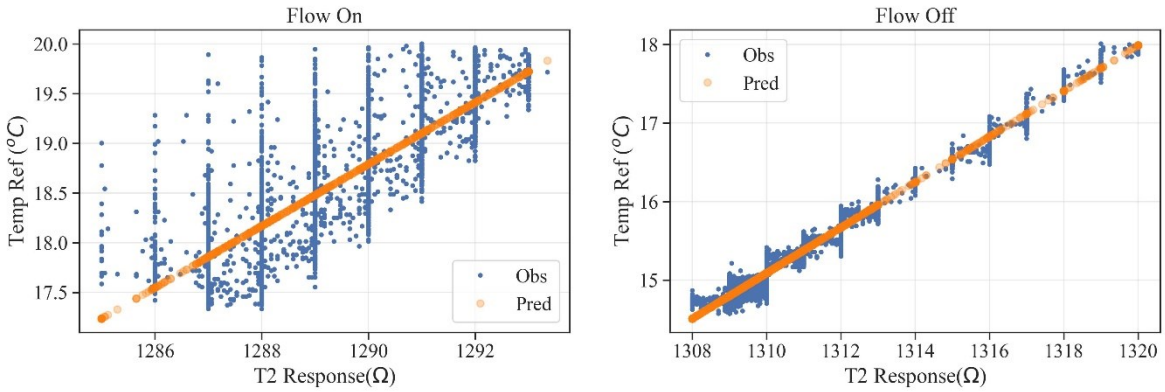

Figure 4.9: Temperature sensor (T2) response and impact of water flow condition on sensor response signal (blue dots: Obs – raw sensor signal, orange dots: Pred – model prediction)

Table 4.5 summarizes the outcomes for three distinct sensor chips, examining conditions with and without water flow. The model's parameters, slope and bias (as summarized in Table 4.5), are remarkably consistent across both conditions, deviating by less than 10% from one another. This consistency highlights the robustness of the modelling approach employed.

Across all setups, the sensors exhibit superior predictive performance compared to a naïve baseline, as evidenced by lower MAE. Predictions in the absence of flow consistently outperform those in flowing conditions, although the degree of difference is contingent upon the specific chip in question. The outperformance in no flow condition can be attributed to having a static measurand (in this case, the water, which is not in a chaotic state) and the

reduction of EM noise due to the pump and temperature regulator (and other related equipment) being turned off. Overall, MAEs span from 0.09°C to 0.4°C, a crucial parameter that decides the prediction accuracy of other target parameters (Conductivity, pH, and Chlorine concentration).

Table 4.5: Model parameters and performance metrics for temperature sensors (T1, T2) in different chips

		Flow on		**Flow off**	
	Parameters	T1	T2	T1	T2
C141 (2021) On: 17°C – 20°C Off: 14°C – 18°C	Slope (°C/Ω)	0.28	0.29	0.22	0.29
	Bias (°C)	-362	-362	-279	-366
	MAE (°C)	0.34	0.35	0.09	0.08
	Mean MAE (°C)	0.60	0.60	0.69	0.51
C80 On: 17°C – 20°C Off: 14°C – 22°C	Slope (°C/Ω)	0.40	0.43	0.49	0.49
	Bias (°C)	-501	-521	-615	-611
	MAE (°C)	0.23	0.20	0.14	0.13
	Mean MAE (°C)	0.60	0.59	1.38	1.38
C35 On: 17°C – 20°C Off: 17°C – 21°C	Slope (°C/Ω)	Sensor not available	0.44	Sensor not available	0.31
	Bias (°C)		-535		-366
	MAE (°C)		0.23		0.19
	Mean MAE (°C)		0.65		0.95

4.3.2 Conductivity Sensor Model

Two distinct models were analyzed for conductivity sensors, incorporating a linear temperature dependency as frequently noted in existing research. The models explore both direct linear relationships with conductivity and inverse relationships with conductivity. Given the sensor's physical principles, an inverse model is theoretically supported:

$$V_{Ci} = a_0 + a_1 T_{ref} + a_2 / C_{ref} \quad (4.5)$$

where V_{Ci} is the voltage drop across the conductivity sensor Ci, T_{ref} and C_{ref} are the reference temperature and conductivity values.

However, considering the limited variability of conductivity in the dataset, the study also assesses whether a linear approximation might effectively represent the nonlinear relationship between the sensor response and conductivity, as given by the formulae:

$$V_{Ci} = a_0 + a_1 T_{ref} + a_2 C_{ref} \tag{4.6}$$

After estimating the coefficients, the target conductivity is predicted C_P, by inverting the model equation as shown in Table 4.6, replacing the reference temperature with the temperature sensor model response.

Table 4.6: Two chosen models for Conductivity sensor and model inversion for prediction

	Model 1	Model 2
Direct model for coefficient estimation	$V_{Ci} = a_0 + a_1 T_{ref} + a_2 C_{ref}$	$V_{Ci} = a_0 + a_1 T_{ref} + a_2 / C_{ref}$
Inversion for prediction	$C_P = \frac{V_{Ci} - a_0 - a_i T_{ref}}{a_2}$	$C_P = \frac{a_2}{V_{Ci} - a_0 - a_i T_{ref}}$

The dataset for just one sensor chip (of the 6 chips experimented in Sense-city) has a broad enough range of conductivity values (from 750 to 1270 µS/cm) to test the proposed models adequately. As anticipated, due to the limited variability in conductivity, Models 1 and 2 exhibit closely matched performance and model coefficients, as detailed in Table 4.7. Surprisingly, the sensitivity to changes in conductivity is lower than to temperature variations by an order of magnitude of one to two, leading to relatively inconsistent performance outcomes. This discrepancy is likely attributable to the high thermal sensitivity of the silicon substrate and the wire bonding connections between the sensor chip and its PCB (previously highlighted as vulnerability points in the first version of the LOTUS sensor and addressed in the design of version 2).

Despite these limitations in version 1, the best MAEs recorded in static and dynamic water conditions stand at 31 µS/cm (for sensor C1) and 73 µS/cm (for sensor C3), respectively.

Table 4.7: Conductivity sensor model parameters and performance metrics

	Flowing water			Static water		
C_{ref} = 750 to 1270 µS/cm	C1	C2	C3	C1	C2	C3
Model 1, a1 (µV/°C)	-189	-100	-93	-20	6.6	-27
Model 1, a2 (µV/ µS/cm)	0.2	-1.5	-0.7	-6.4	-0.6	0.1
MAE (µS/cm)	204	**73**	128	**31**	153	656
Model 2, a1 (µV/°C)	-189	-96	-94	-19	7	-27
Model 2, a2 (V. µS/cm)	-0.2	1.4	0.7	4.6	0.4	-0.1
MAE (µS/cm)	323	**98**	188	**31**	124	5133
Naïve MAE (µS/cm)	182	141	103	222	92	76

4.3.3 Active Chlorine and pH Sensor Model

The models for chlorine and pH are exclusively developed for conditions with water flow since chlorine measurements are not taken in the absence of flow. The dataset from chip 141, encompassing approximately 200 hours of data collected in 2021, was utilized to create and validate these models. The presence of flow introduces a significant level of noise to the dataset. In water, chlorine splits into active chlorine and hypochlorite ions. The LOTUS sensor is designed to detect active chlorine concentration within the Sense-City pipeline. Consequently, the calibration models for the sensors are formulated based on the concentrations of the actual chemical entities found in the water.

In water, chlorine dissociates into HClO (active chlorine) and ClO^- (hypochlorite ion), where the equilibrium concentration depends on temperature and pH [222]. For a chlorine concentration ($C_{chlorine}$) mg/l, pH and Temperature (T) ^oC, the active chlorine concentration C_{HClO} mg/l, and concentration of hypochlorite ions C_{ClO^-} mg/l, is given as:

$$C_{chlorine} = C_{HClO} + C_{ClO^-} \tag{4.7}$$

$$C_{HClO} = C_{chlorine}\left(1 + \left(0.98 \times 10^{pH\ -pKa(T)}\right)\right)^{-1} \tag{4.8}$$

$$pKa(T) = 7.5 - 0.01(T - 30) \tag{4.9}$$

Eleven calibration models (as detailed below) were explored for their ability to predict active chlorine and pH levels, applying these models across twenty chemiresistors, where R_i corresponds to the relative response of ith chemiresistor selected, $\alpha_i, \beta_i, \gamma_i, \zeta_i, \lambda_i, \eta_i, \delta_i$ are the coefficient parameters, C_{HClO} denotes active chlorine concentration, C_{ClO} denotes the concentration of hypochlorite ions, T denotes the temperature, and ϕ_i denotes the error.

$$R_i = \alpha_i C_{HClO} + \phi_i \tag{4.10}$$

$$R_i = \alpha_i C_{HClO} + \beta_i\, pH + \phi_i \tag{4.11}$$

$$R_i = \gamma_i C_{ClO} + \phi_i \tag{4.12}$$

$$R_i = \gamma_i C_{ClO} + \beta_i\, pH + \phi_i \tag{4.13}$$

$$R_i = \alpha_i C_{HClO} + \gamma_i C_{ClO} + \beta_i\, pH + \phi_i \tag{4.14}$$

$$R_i = \alpha_i C_{HClO} + \beta_i\, pH + \zeta_i T + \phi_i \tag{4.15}$$

$$R_i = \alpha_i C_{HClO} + \lambda_i C_{HClO}^2 + \beta_i\, pH + \phi_i \tag{4.16}$$

$$R_i = \gamma_i C_{ClO} + \eta_i C_{ClO}^2 + \beta_i\, pH + \phi_i \tag{4.17}$$

$$R_i = \lambda_i C_{HClO}^2 + \eta_i C_{ClO}^2 + \phi_i \tag{4.18}$$

$$R_i = \alpha_i C_{HClO} + \delta_i \exp(pH) + \phi_i \tag{4.19}$$

$$R_i = \gamma_i C_{ClO} + \delta_i \exp(pH) + \phi_i \tag{4.20}$$

Considering both Coefficient of determination and computational efficiency, the selection process led to the choice of model M2, given by equation (4.11).

Figure 4.10 illustrates the Coefficient of determination for the explored linear models (4.10) - (4.15) and non-linear models (4.16) - (4.20), respectively. It is important to note that while model M2 (4.11) is not the top performer, its performance is nearly identical to that of the best-performing model, M5 (4.14), across almost all chemiresistors. Moreover, M2 has fewer degrees of freedom, minimizing the risk of overfitting. From this model, the chemiresistors having R^2 above 60% and the coefficients for chlorine and pH having CoV below 5% are chosen, namely: 1A, 1B, 1C, 1D, 2A, 2B, 2C, 1F, 1G, 1H, 1I, 2G; with the initial seven having non-functionalized CNT ink and the remaining five featuring functionalized ink.

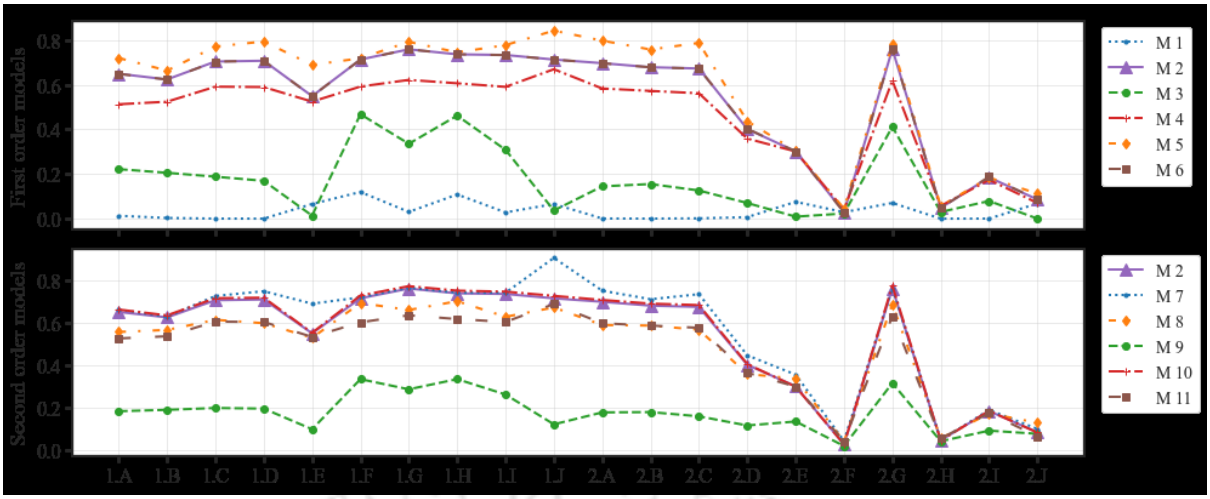

Figure 4.10: Coefficient of determination estimated from the dataset of the explored models. Top – Linear models and Bottom – Non-linear models

Table 4.8 offers a summarized overview of the sensors' sensitivities, highlighted by:

Sensitivity modulation: The introduction of FFUR functionalization altered the chemiresistors' sensitivity compared to those with pristine CNT. This effect is more pronounced in pH sensitivity (where PCNT has 80.9% sensitivity compared to FFUR's 48% per pH unit).

Device-to-device reproducibility: Sensitivity variation among devices is consistent between those with CNT and FFUR functionalization. Variability is marginally less for active chlorine than for pH, which could be attributed more to the testing conditions than the chemical composition of the devices. Notably, a 20% variability (CoV) in active chlorine sensitivity matches well with findings from previous research on printed CNT devices [223].

Table 4.8: Sensitivity analysis of the chemiresistors: PCNT – 7 pristine CNT devices, 1A, 1B, 1C, 1D, 2A, 2B, 2C; FFUR – 6 functionalized CNT devices, 1F, 1G, 1H, 1I, 2G

Name of the CNT Ink	Coefficient for Active chlorine, α (%/ mg/l)		Coefficient for pH, β (%/ pH)	
	Average	Standard deviation	Average	Standard deviation
PCNT	-48.1	9.8	-80.9	22.6
FFUR	-41.6	5.7	-48.0	12.3

Following the choice of M2 as the calibration model for the chemiresistors, developing the prediction model involves an inversion process. Since the model targets two parameters but utilizes more than two chemiresistors, a pseudo-inverse matrix technique is employed for model inversion. This approach not only facilitates the inversion but also enables data fusion.

An exhaustive search of the optimal subset from the selected chemiresistors yields 1A, 1B, 1C, 1D, and 1H (4 PCNT, 1 FFUR sensor). The prediction model is constructed as given below:

$$\begin{bmatrix} 1A & 1B & 1C & 1D & 1H \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} HClO & pH \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} \alpha_i & \cdots \\ \beta_i & \cdots \end{bmatrix} + [\phi_i \quad \cdots] \tag{4.21}$$

In shorthand notation, this can be written as:

$$\mathrm{R} = \mathrm{TC} + \phi \tag{4.22}$$

from which the target parameters are estimated by matrix inversion, given by:

$$\mathrm{T} = (\mathrm{R} - \phi)\mathrm{C}^{+} \tag{4.23}$$

where C^+ is the Moore-Penrose inverse.

Figure 4.11 illustrates the outcomes of the model inversion: the model predictions track the changes in chlorine and pH levels. However, the model fails to detect the smallest step change in active chlorine (occurring at 90 hours) and exhibits drifts in pH following each pH adjustment. The Mean Absolute Errors (MAEs) for active chlorine and pH stand at 0.13 mg/l and 0.12, respectively, compared to the naïve MAEs at 0.14 mg/l and 0.17. The error margin for active chlorine is above the ideal threshold for its intended application (where 0.05 mg/l would be preferable), whereas the error for pH is within acceptable limits.

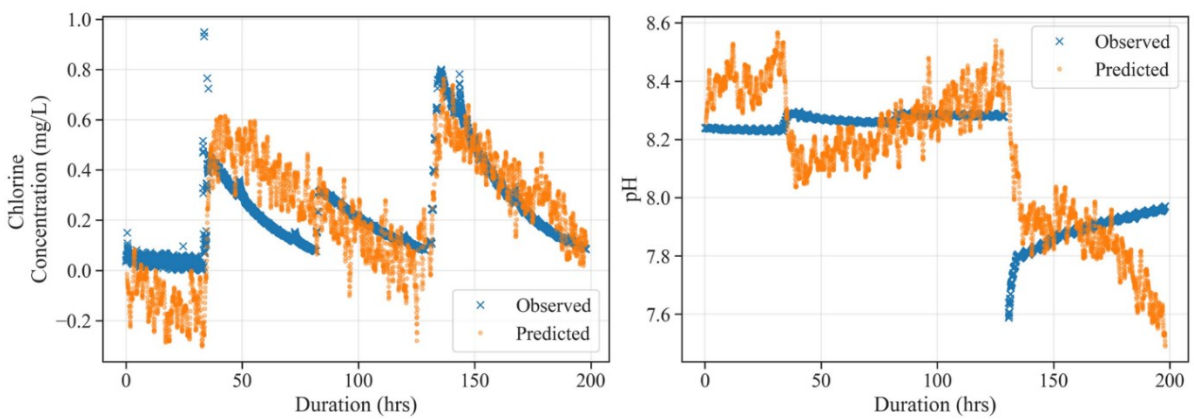

Figure 4.11: Time series of predicted and observed HClO and pH values

4.3.4 Uncertainty quantification of LOTUS sensors

Uncertainty quantification is a process that involves systematically calculating and expressing the degree of variability associated with measurements, models, and predictions. It aims to comprehensively assess the confidence or reliability of the outcomes derived from experiments and models. This process encompasses identifying sources of uncertainty, quantifying them through statistical measures (such as standard deviation), and then analyzing their impact on the final predictions. In essence, uncertainty quantification aids in understanding the potential range of outcomes and the likelihood of different scenarios.

In a sensor system, two primary types of general uncertainty are commonly encountered: measurement uncertainty and model uncertainty. Measurement uncertainty arises from the sensor due to factors like sensor drift, calibration errors, or environmental influences affecting the sensor's readings. Model uncertainty, on the other hand, stems from the process of using a calibration model to interpret sensor data, including inaccuracies in the model's structure, the estimation of model parameters, the appropriateness of the model for capturing complex real-world phenomena, and the uncertainties in target data. Typically, the uncertainty can be estimated via different techniques, and they are influenced to a great extent by the calibration model. If the calibration model has a closed analytical form, we can use the Gaussian Error Propagation (GEP) formulae [224], [225], [226] to derive an uncertainty equation and estimate it. However, if the model is complex, has no closed form, or is a black box model (like Neural networks), bootstrap methods [227] can be used to estimate uncertainty.

In the case of the LOTUS sensor, we are focusing on the following factors contributing to the uncertainty: noise or response variability in the LOTUS sensor, inaccuracies and random noise in the target sensors, and inadequacies of the model developed. For Temperature and Conductivity, we have explicit closed-form simple analytical equations, so we are using GEP formulae, while for chlorine and pH, we developed a bi-parameter matrix inversion model, for which the bootstrap method would be a simpler approach.

Uncertainty in Temperature Sensing

The dataset generated from experiments with LOTUS sensor 141 during 2021 is utilized for this analysis. The uncertainty associated with the LOTUS sensor response ($\sigma_{Response}$) and the model coefficients are outlined in Table 4.9 for the temperature testing over the range of 17.5°C to 20°C.

Table 4.9: Uncertainties associated with the sensor response and model coefficients

Flow condition	Sensor	$\sigma_{Response}$ (Ω)	Model coefficients			
			Slope (°C/Ω)	σ_{Slope} (°C/Ω)	Bias (°C)	σ_{Bias} (°C)
On	T1	1.50	0.284	0.002	-362.35	3.17
	T2	1.47	0.295	0.002	-362.05	2.52
Off	T1	0.50	0.222	0.0003	-277.09	0.33
	T2	0.37	0.291	0.0004	-365.59	0.57

Using the error propagation, the uncertainty of the sensor prediction ($\sigma_{Prediction}$) can be calculated as per the following equations. For the calibration model given in (4.24), T_P refers to the predicted temperature, R_{Ti} referes to the response from temperature sensor T1, T2 based on i = 1 or 2, and α_i, β_i denotes the slope and bias of the model for sensor Ti.

$$T_P = \alpha_i R_{Ti} + \beta_i \tag{4.24}$$

$$\sigma_{T_P} = \sqrt{\left(\frac{\partial T_P}{\partial \alpha_i}\,\sigma_{\alpha_i}\right)^2 + \left(\frac{\partial T_P}{\partial R_{Ti}}\,\sigma_{R_{Ti}}\right)^2 + \left(\frac{\partial T_P}{\partial \beta_i}\,\sigma_{\beta_i}\right)^2} \tag{4.25}$$

$$\Rightarrow \sigma_{T_P} = \sqrt{\left(R_{Ti}\,\sigma_{\alpha_i}\right)^2 + \left(\alpha_i \sigma_{R_{Ti}}\right)^2 + \sigma_{\beta_i}^2} \tag{4.26}$$

Based on equation (4.26), the uncertainty associated with predicting temperature σ_{T_P} with LOTUS sensor T1 and T2 is summarized in Table 4.10. The uncertainty provided is much higher than the respective MAE reported, as the MAE considered only the uncertainty in the sensor response and assumed the model coefficients to be fixed.

Table 4.10: Uncertainty of the Temperature prediction by LOTUS sensor no: 141 in the temperature range 17.5°C to 20°C

Sensor	Flow ON		Flow OFF	
	Range (Ω)	σ_{T_P} (^oC)	Range (Ω)	σ_{T_P} (^oC)
T1	1335.1 – 1345.0	4.18	1321 – 1337	0.53
T2	1285.1 – 1293.4	3.63	1308.1 – 1319.9	0.78

Uncertainty in Conductivity Sensing

In the case of the conductivity sensor, as discussed earlier, only the LOTUS sensor no: AS001 has enough data for an approximate model. In the model for temperature parameters, we utilized the actual temperature instead of the temperature predicted by sensors T1 and T2 because the baseline resistance of the sensor in AS001 was at 22 kΩ compared to the targeted 1 kΩ to 1.5 kΩ. So, to provide an estimate, the uncertainty associated with the temperature parameter is taken from Table 4.10. From Table 4.7, we can see that for both models, sensor C2 was performing better in flow ON condition, and C1 was performing better in no flow condition. Therefore, uncertainty quantification is only done for these sensors. The uncertainty for conductivity sensing σ_{CP}, for both models, with respect to the factors can be expressed as:

$$\sigma_{CP} = \sqrt{\left(\frac{\partial C_P}{\partial V_{Ci}}\,\sigma_{V_{Ci}}\right)^2 + \left(\frac{\partial C_P}{\partial a_0}\,\sigma_{a_0}\right)^2 + \left(\frac{\partial C_P}{\partial a_1}\,\sigma_{a_1}\right)^2 + \left(\frac{\partial C_P}{\partial a_2}\,\sigma_{a_2}\right)^2 + \left(\frac{\partial C_P}{\partial T_{ref}}\,\sigma_{T_{ref}}\right)^2} \quad (4.27)$$

For model 1, each term in the above equation will be,

$$\frac{\partial C_P}{\partial V_{Ci}} = \frac{1}{a_2}, \frac{\partial C_P}{\partial a_0} = -\frac{1}{a_2}, \frac{\partial C_P}{\partial a_1} = -\frac{T_{ref}}{a_2}, \frac{\partial C_P}{\partial T_{ref}} = -\frac{a_1}{a_2}, \frac{\partial C_P}{\partial a_2} = -\frac{\lambda}{a_2^2},$$

$$where\ \lambda = V_{Ci} - a_0 - a_1 T_{ref}$$

Table 4.11: Uncertainty in Conductivity sensing, using model 1 (4.6), in flow and no flow condition

		C2 (best in flow condition)	C1 (best in no flow condition)
Model Coefficients	a_0 (µV)	7182.81	9437.43
	a_1 (µV/°C)	-100.31	-19.67
	a_2 (µV/ µS/cm)	-1.53	-6.36
Uncertainty in model coefficients	σ_{a_0} (µV)	531.95	1162.13
	σ_{a_1} (µV/°C)	10.98	5.62
	σ_{a_2} (µV/ µS/cm)	0.30	1.37
Model inputs and their uncertainties	V_{Ci} (µV)	3541.08	3647.68
	$\sigma_{V_{Ci}}$ (µV)	137.60	338.84
	T_{ref} (°C)	20.90	20.49
	$\sigma_{T_{ref}}$ (°C)	3.62	0.52
Uncertainty with sensing	Range (µV)	3243 – 3960	3005 – 5000
	σ_{CP} (µS/cm)	497.1	257.4
	σ_{CP}/C_P	0.50	0.31

For model 2, each term will be:

$$\frac{\partial C_P}{\partial V_{Ci}} = \frac{-a_2}{\lambda^2}, \frac{\partial C_P}{\partial a_0} = \frac{a_2}{\lambda^2}, \frac{\partial C_P}{\partial a_1} = \frac{a_2 T_{ref}}{\lambda^2}, \frac{\partial C_P}{\partial T_{ref}} = \frac{a_1 a_2}{\lambda^2}, \frac{\partial C_P}{\partial a_2} = \frac{1}{\lambda}$$

Table 4.12: Uncertainty in Conductivity sensing, using model 2 (4.5), in flow and no flow condition

		C2 (best in flow condition)	C1 (best in no flow condition)
Model Coefficients	a_0 (µV)	4192.17	-1375.29
	a_1 (µV/°C)	-96.29	-19.15
	a_2 (V µS/cm)	1.37	4.58
Uncertainty in model coefficients	σ_{a_0} (µV)	160.67	1171.28
	σ_{a_1} (µV/°C)	11.68	5.63
	σ_{a_2} (V µS/cm)	0.38	0.98
Model inputs and their uncertainties	V_{Ci} (µV)	3541.08	3647.68
	$\sigma_{V_{Ci}}$ (µV)	137.60	338.84
	T_{ref} (°C)	20.90	20.49
	$\sigma_{T_{ref}}$ (°C)	3.62	0.52
Uncertainty with sensing	Range (µV)	3243 – 3960	3005 – 5000
	σ_{CP} (µS/cm)	392.5	255.3
	$\frac{\sigma_{CP}}{C_P}$	0.40	0.30

Uncertainty in Chlorine and pH sensing

In the case of chlorine and pH sensing, we use a bi-parameter mode with matrix inversion. For this model, the bootstrap-based method is preferred over developing an analytical uncertainty equation from the GEP, as the model is a complex, closed form though analytical.

For the case of bootstrapping, we have used the model with random data selected from the available dataset (sensor no: 141 from 2021 experiments), and the coefficients are obtained from a Gaussian distribution with the mean and standard deviation, as obtained from model development (repeated iterations with random 70% of data as described in 4.3.3).

Based on equation *(4.21)*, the uncertainties associated with the coefficients of the selected chemiresistors (1A, 1B, 1C, 1D, and 1H) are outlined in Table 4.13.

Table 4.13: Uncertainties associated with the coefficients of Active Chlorine and pH model

Coefficients →	HClO (%/ mg/l)		pH (%/ pH)		Bias
Chemiresistors ↓	α_i	σ_{α_i}	β_i	σ_{β_i}	σ_{bias}
1A	-0.425	0.010	-0.560	0.011	0.096
1B	-0.366	0.008	-0.530	0.010	0.086
1C	-0.450	0.010	-0.698	0.013	0.110
1D	-0.499	0.013	-0.811	0.016	0.133
1H	-0.369	0.007	-0.358	0.006	0.054

The uncertainty associated with chlorine and pH sensing using bootstrap methodology with the information from Table 4.13 for chlorine is estimated to be 0.235 mg/l and for pH 0.210 pH units for the range 0 – 5mg/l and 7.6 to 8.3, respectively.

4.3.5 Summary of LOTUS sensor - silicon version validation in Sense-city

Table 4.14 comprehensively summarizes the optimal Mean Absolute Errors (MAEs) obtained from the Sense-city dataset following data refinement and model formulation. These results are deemed satisfactory for the applications envisaged by LOTUS, except for chlorine predictions, where a reduction in MAE by half would be more desirable. It is anticipated that enhancements in system noise reduction will straightforwardly address this concern. It is important to highlight that while Table 4.14 sheds light on the system's potential in practical settings, these outcomes should not be interpreted as real-time or in-field performance metrics. This is because the models have not been trained on datasets with temporal structuring (for instance, training with initial days' data and predicting for subsequent days).

Table 4.14: Best MAE values achieved over the dataset for LOTUS sensor version 1

Parameter	**Average Error of Detection**		**Field requirement**
	Flowing water	Static water	
Temperature	0.2^oC	0.1^oC	$< 0.1^oC$
Conductivity	73.2 $\mu S/cm$	31.3 $\mu S/cm$	$< 30\ \mu S/cm$
Active Chlorine	0.13 mg/L		< 0.05 mg/L
pH	0.12 pH units		< 0.1 pH

4.4 Experimentation with LOTUS Sensor - Plastic Version

4.4.1 Laboratory experiments for Chlorine and pH

Before comprehensive characterization, the laboratory protocol for characterization was refined to address specific issues like interference from external light and delayed homogenization.

Light sensitivity was noted in Pristine CNT (strongly) and FFUR (moderately) chemiresistors, a phenomenon not encountered with previous MWCNT-based inks. The second version's inks, which utilize SWCNTs, are known for their light sensitivity, and the FFUR polymer is known for its UV sensitivity. This accounts for the observed light sensitivity, which does not affect field usability since the sensors are deployed within pipes shielded from light exposure. However, this finding significantly influences laboratory characterizations typically conducted in glass beakers (as seen in Figure 4.12), and the tube lights used in the laboratory (commercial 40W) influenced the dynamic response of chemiresistors. After this realization, characterizations were performed with a cover over the setup, or lights turned off.

Magnetic stirring, a common laboratory technique for rapid solution homogenization, was found to interfere substantially with AFE readings, rendering it unsuitable for consistent use. Experimentation revealed that short bursts of magnetic stirring for 100 seconds followed by immediate cessation allowed sensors to reach a steady state, indicating that their response time in laboratory conditions is under 100 seconds (refer to Figure 4.12). Due to the protocol's impracticality with magnetic stirring, it was largely excluded from the presented results. Consequently, the response times discussed indicate solution homogenization within the beaker rather than intrinsic sensor response characteristics.

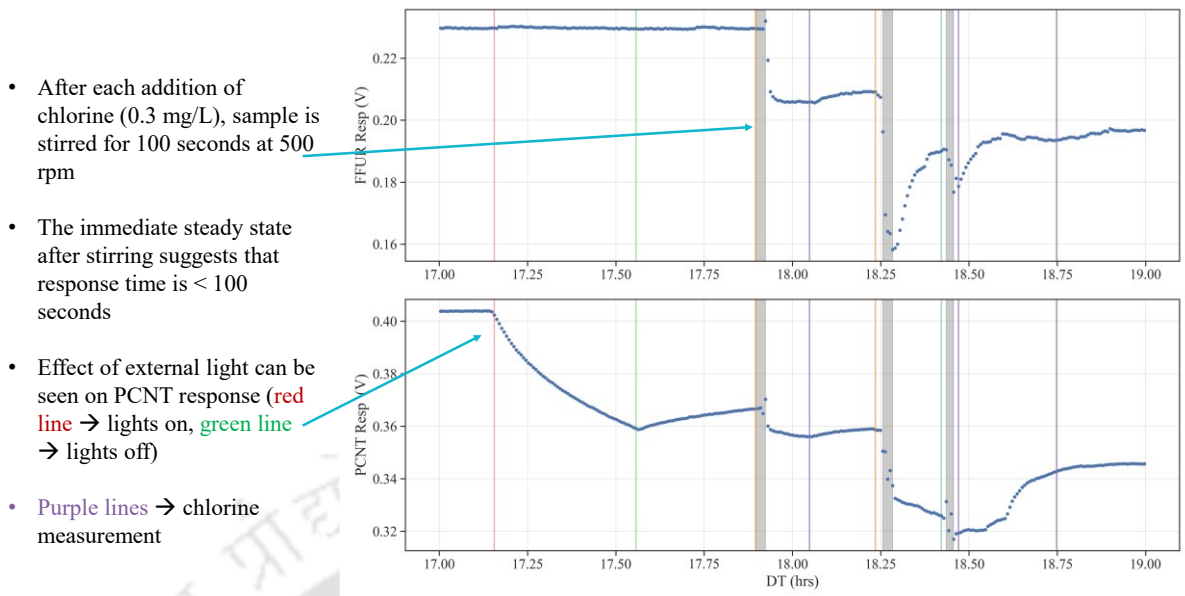

Figure 4.12: Effect of magnetic stirring (applied during the interval of times plotted in grey) on the response time. Effect of the PCNT light sensitivity (bottom, between 17:20 and 17:55). Measurements carried out with Alternating DC protocol

Figure 4.13 illustrates a one-day laboratory characterization of chlorine sensitivity across different pH levels, conducted using AC excitation with the second version of the LOTUS Analog Front End (UAFE), which at that time was not voltage-limited. Key observations from this figure include:

- Adjustments in chlorine (addition of NaOCl) and pH levels (addition of HCl) visibly impact sensor readings, albeit with a somewhat extended response time, possibly due to challenges in achieving uniform beaker homogenization.
- No real-time chlorine measurements were available, as during the initial phase, we used a commercial DPD-based colorimeter Aqualabo Photopod [228], and no real-time pH measurements were also made, as the pH electrode created interference in the LOTUS sensor response measurement.
- The immediate response to chemical additions is evident in the coefficient of variation (CoV) plots (shown in the second row of Figure 4.13). These plots reveal the sensors' rolling CoVs, with a distinct peak corresponding to each addition.
- Between these additions, the CoV stabilizes at 0.1% or lower, aligning with the predefined objectives for the AFE.

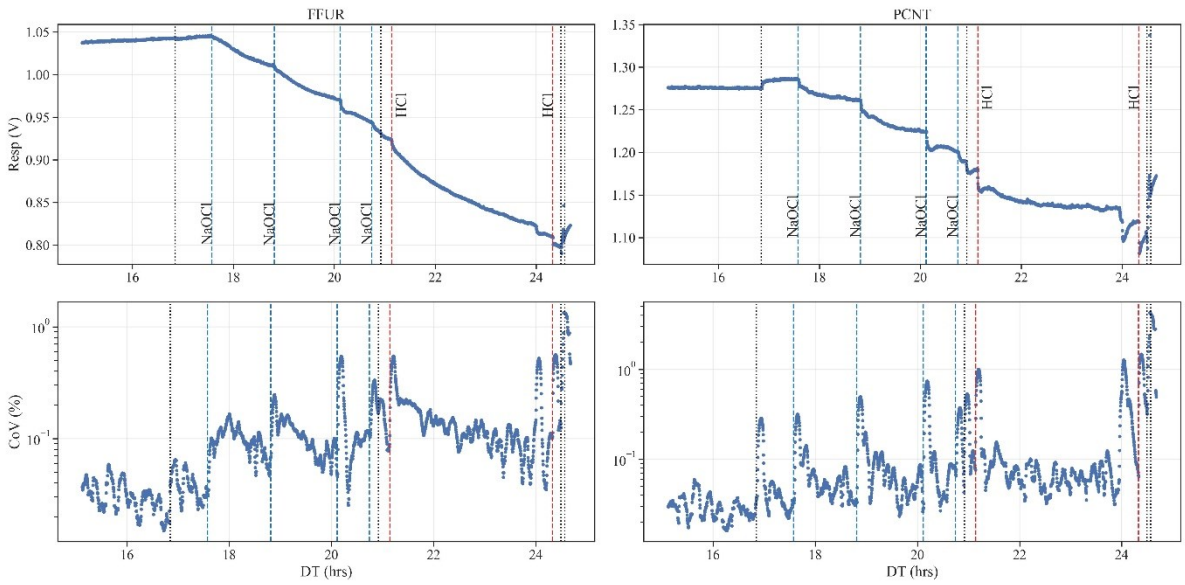

Figure 4.13: Example of 1-day chlorine/pH experiment under AC excitation. Top: raw sensor outputs. Bottom: coefficients of variation (e.g. noise)

Encouraged by this proof of concept, a comprehensive series of calibration experiments was subsequently executed using the UAFE with AC excitation protocol for the chemiresistors with the experimental details as depicted in Table 4.15. After processing the data to express outputs in terms of relative variation of resistance, the results were correlated with the system's main variables (actual chlorine, active chlorine, hypochlorite ions, and pH).

Table 4.15: Series of lab characterization experiments used to build calibration model.

S.no	**Chlorine range (mg/l)**	**pH range**	**FFUR ink**	**PCNT ink**	**Environmental effect**
1	0 - 0.75	8.5	Detrended response + Median response at steady state	Relative response + Median response at steady state	Lab Lights on
2	0 - 2	8.5			
3	0 - 2	7.2 - 8.5			
4	0 - 3	7.5 - 8.5	Relative response + Median response at steady state		Lab Lights off
5	0 - 1.9	7.3 – 8.5			

The construction of calibration models was approached as follows:

- Dataset Composition:
    - Total 30 samples, 20 iterations of data, split into 70% training and 30% testing
- Model Formulation:
    - Utilization of generalized linear regression
- Evaluation Metrics:
    - Coefficient of determination, with minor variations (~0.01), deemed insignificant
    - Standard deviation across splits, noting high variability in test outcomes attributed to the limited dataset size

Apart from the models analyzed for the Silicon version (models M1 to M11, corresponding to equations (4.10) - (4.20)), a few other model equations were also analyzed for this dataset. These models were chosen step by step, wherein at each step, one of the 4 parameters (actual chlorine CL, active chlorine HCLO, hypochlorite ions OCL, and pH) was introduced in first order and second order. We also examined the case with Hydrogen ion concentrations instead of pH (which is estimated as 10^{pH}) along with the other models. The Coefficient of determination obtained from all these models are visualized in Figure 4.14.

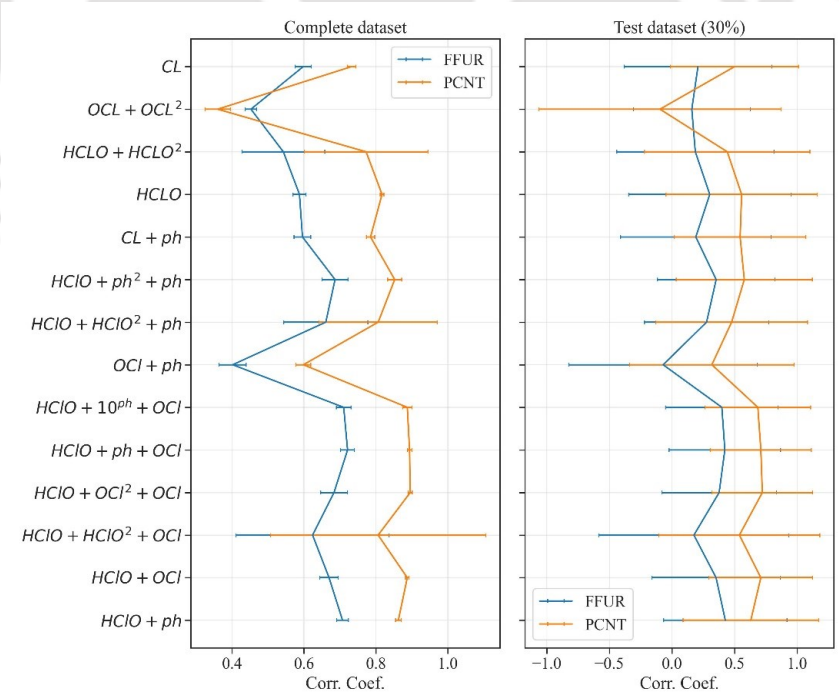

Figure 4.14: Coefficient of determination achieved for different models. The error bars indicate the variability over the 20 different splits. Left: training sets. Right: testing sets. Blue: FCNT. Orange: PCNT.

Figure 4.14 reveals that the most stable models identified based on their consistent performance as:

$$PCNT = c\,HClO + d\,OCl + \epsilon_1 \tag{4.28}$$

$$FFUR = a\,HClO + b\,pH + \epsilon_0 \tag{4.29}$$

The characteristics of these models are detailed in Table 4.16 and Figure 4.15. While some alternatives exhibited marginally superior Coefficient of determination, they were disqualified due to their more significant standard deviation across data splits or an excessive number of parameters, which could increase the risk of overfitting.

Table 4.16: Properties of the best model combination. The R^2 and MAE provided here relate to the direct model (predicting a given sensor output knowing active chlorine, pH, or OCl).

Coefficients	FFUR	PCNT
HClO (/ mg/l)	0.15 (0.01)	0.074 (0.004)
pH (/pH unit)	0.06 (0.01)	-
OCl (/ mg/l)	-	0.017 (0.004)
ϵ_0, ϵ_1	-0.46 (0.08)	-0.015 (0.003)
R2	0.71 (0.02)	0.89 (0.005)
MAE	0.024 (0.001)	0.001 (0.0003)
Naïve MAE	0.042	0.026

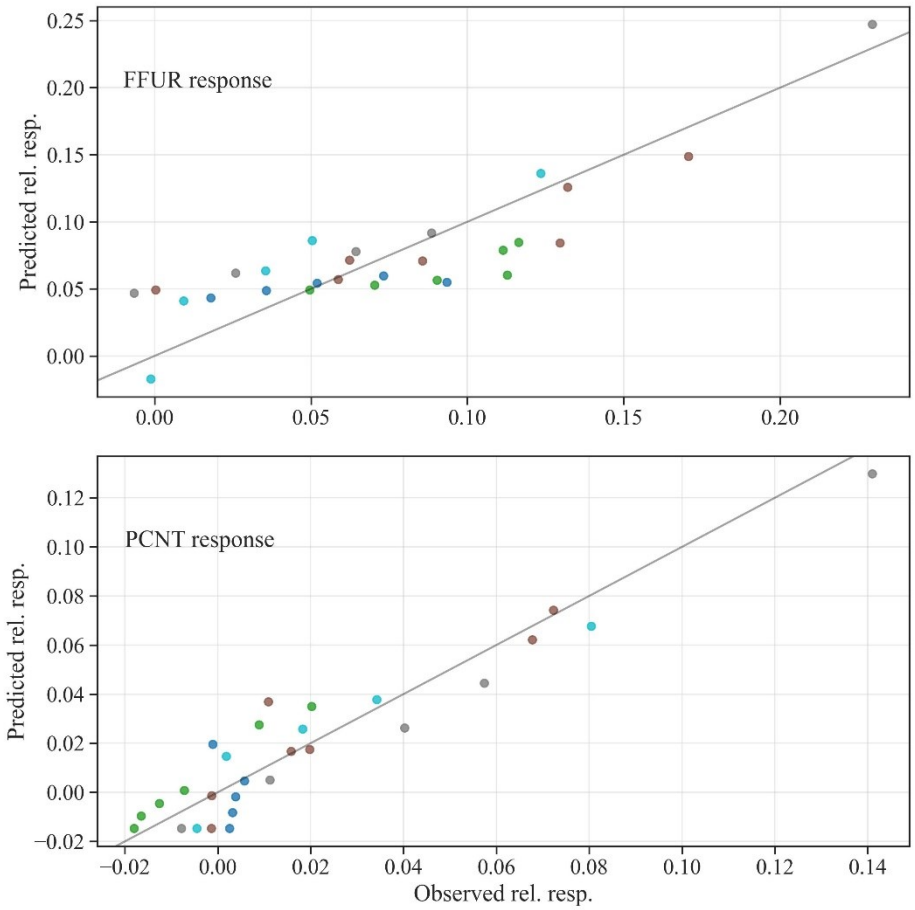

Figure 4.15: Scatterplot predicted (vertical axis) versus actual (horizontal axis) for the direct calibration model.

Two different strategies were assessed for model inversion:

1. Single Parameter Inversion: This approach involves inverting for one parameter at a time, assuming the value of the other parameter is known, applied individually to each of the two chemiresistors.
2. Bi-Parameter Inversion: This method attempts a simultaneous inversion for both parameters, using both chemiresistors together, with various algorithms evaluated for effectiveness.

Given the dataset's baseline, with naive MAEs at 0.29 mg/l for active chlorine and 0.41 for pH, improvements in these figures were anticipated. The outcomes are synthesized in Table 4.17:

- By leveraging both chemiresistors concurrently without prior knowledge, an MAE of 0.16 mg/l for active chlorine was attainable.
- With PCNT chemiresistor and known values of the complementary variable, MAEs can be reduced to 0.01 mg/l for active chlorine or 0.2 for pH.
- The FFUR chemiresistor model shows lesser efficacy, especially in predicting pH, which accounts for the success of bi-parameter inversion solely for HClO and not for pH.

Table 4.17: Inversion performances. Naive MAEs: HClO - 0.29mg/l; pH: 0.41. "Uncertainty" refers to the standard deviation over 20 splits.

Model Input	Prediction	MAE	Uncertainty
C_{HClO}, R_{PCNT}	pH	0.21	0.06
pH, R_{PCNT}	Chlorine (mg/L)	0.01	0.01
C_{HClO}, R_{FFUR}	pH	0.41	0.08
pH, R_{FFUR}	Chlorine (mg/L)	0.16	0.01
R_{PCNT}, R_{FFUR}	pH	0.43	0.004
R_{PCNT}, R_{FFUR}	Chlorine (mg/L)	0.16	0.02

These findings validate the practicality of the chemiresistors printed on plastic substrates for applications within the LOTUS framework. A subsequent study on chlorine sensitivity was performed using alternating DC excitation to verify its applicability in the LOTUS setting. The results, depicted in Figure 4.16, demonstrate a pronounced sensitivity to chlorine (FFUR sensitivity: 26%/mg/l, R^2=0.89 and PCNT sensitivity: 13%/mg/l, R^2=0.71) alongside adequate linearity, which led to adopting this alternating DC – rectangular wave signal as the excitation signal for further experiments.

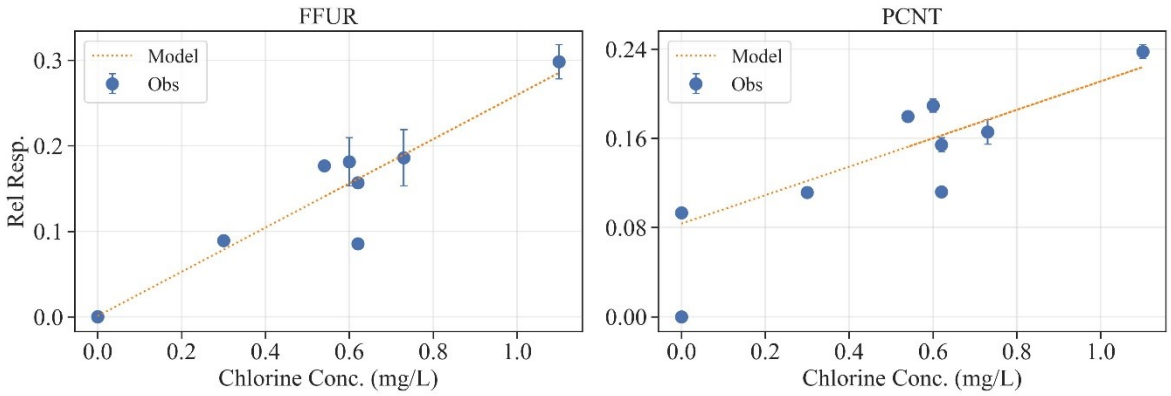

Figure 4.16: Relative response as a function of initial chlorine concentration, showing both observed data (Obs) and simple linear model (Model).

4.4.2 Sense-city experiments

4.4.2.1 Experimental parameters and Dataset description

The sensor was deployed in the SenseCity water loop, operating at a flow rate of 0.3 m/s and a pressure of 1.5 bar for approximately 432 hours. Temperature control was ensured at 25.5°C through an On-Off controller set within a dead band of 25 to 26°C. The experimental conditions varied target parameters within 7 to 8.5 for pH and 0 to 5 mg/L for chlorine concentration. The sensor comprised six units (6-device configuration - Figure 4.17), including pairs designated for PCNT, FFUR, and PVT inks, explicitly labelled 1E, 1I, 2A, 1F, 2D, and 2J. An unexpected complication in the sensor assembly led to damage in device 1F, restricting data analysis and model development to the remaining five functional devices.

Figure 4.17: 6-device configuration sensor head: installed in Sense-city

The sensor utilized the UAFE, employing an enhanced excitation protocol of DC pulses lasting 15 ms, with each sensor's total activation spanning 1.1 seconds. Sensor activation occurred in successive cycles, each lasting 1.1 seconds.

The collected dataset spans 432 hours and includes the following key events:

- Three experiments of chlorine addition at varying pH levels, alongside minor pH adjustments
- A single pH variation experiment conducted in the absence of chlorine
- A gradual increase in pH from 7 to 8.5
- Notable changes in water turbidity throughout the experiment, though this parameter was not formally tracked.

The correlation between the target parameters during the experiments is visualized in Figure 4.18.

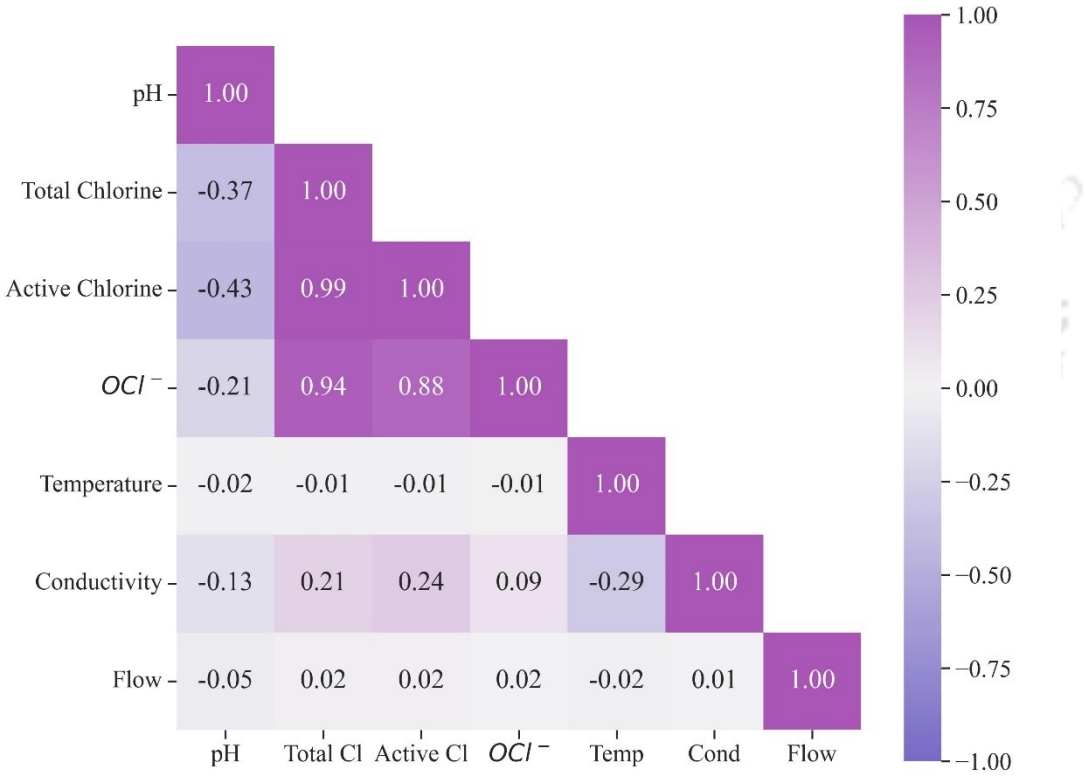

Figure 4.18: Correlation between the experimental parameters using commercial sensors during the complete experiment.

4.4.2.2 Data Pre-processing

Using an on-off temperature controller led to the emergence of a periodic triangular pattern in the response data of both the LOTUS sensor and the pH reference sensor, as illustrated in Figure 4.19. This pattern could stem from the temperature changes directly affecting the sensor readings or from pH fluctuations induced by temperature variations. While the direct impact of temperature on the sensor responses is the more plausible explanation, the current dataset lacks the variability to definitively attribute the cause, leading to the decision to eliminate this triangular waveform from the data of both the reference pH sensor and the LOTUS sensor. FFUR chemiresistor has a lagged response compared to PCNT chemiresistor; this lag can be validated by the triangular effect seen in Figure 4.20 even after removing the temperature effect, as seen in PCNT.

Furthermore, the responses from the Lotus sensors were refined using a 5-point rolling average for smoothing, with outliers removed through a median filter. The corrected responses, with the triangular pattern significantly reduced, are presented in Figure 4.20. The coefficients of variation (CoV) following data pre-processing are showcased in Figure 4.21.

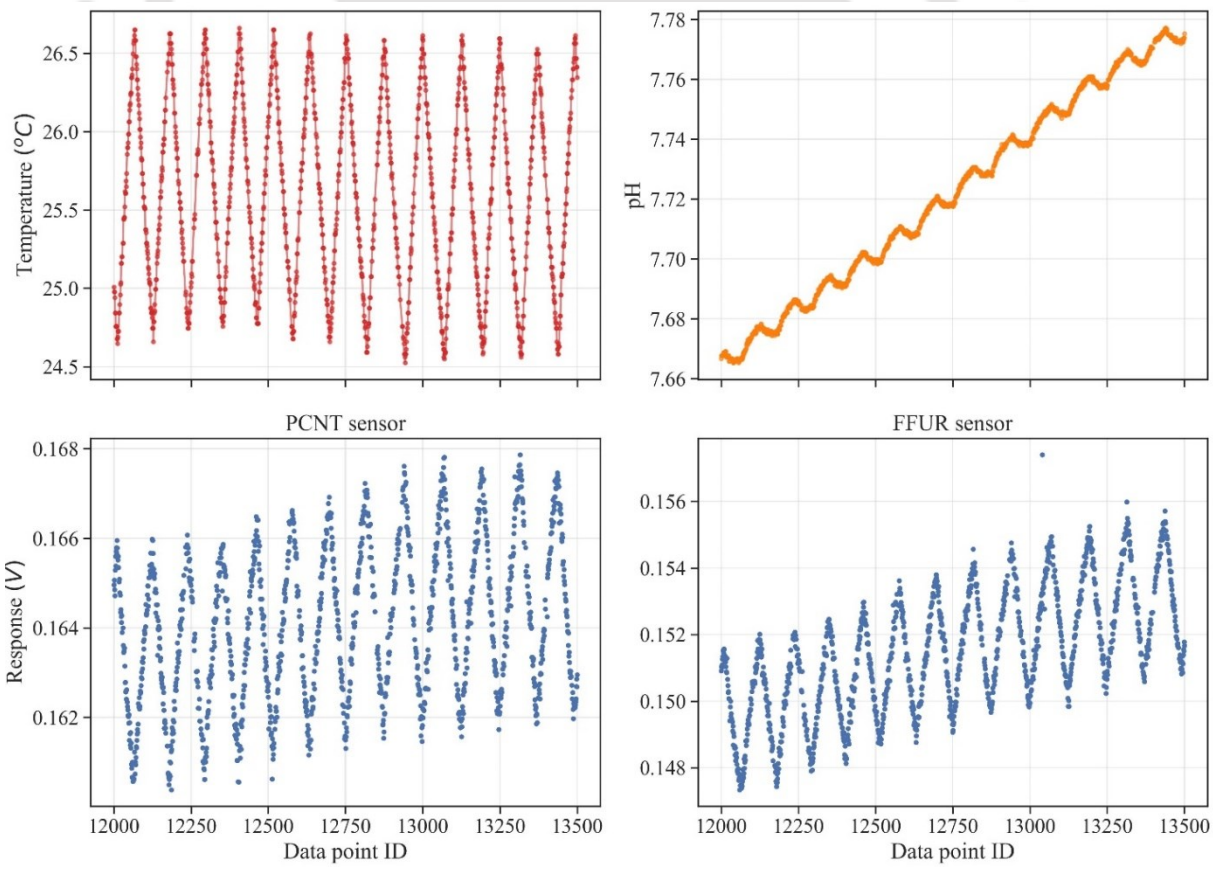

Figure 4.19: Effect of Temperature Controller: top left – reference temperature sensor; top right: reference pH sensor; bottom left: 1 PCNT sensor; bottom right: 1 FFUR sensor.

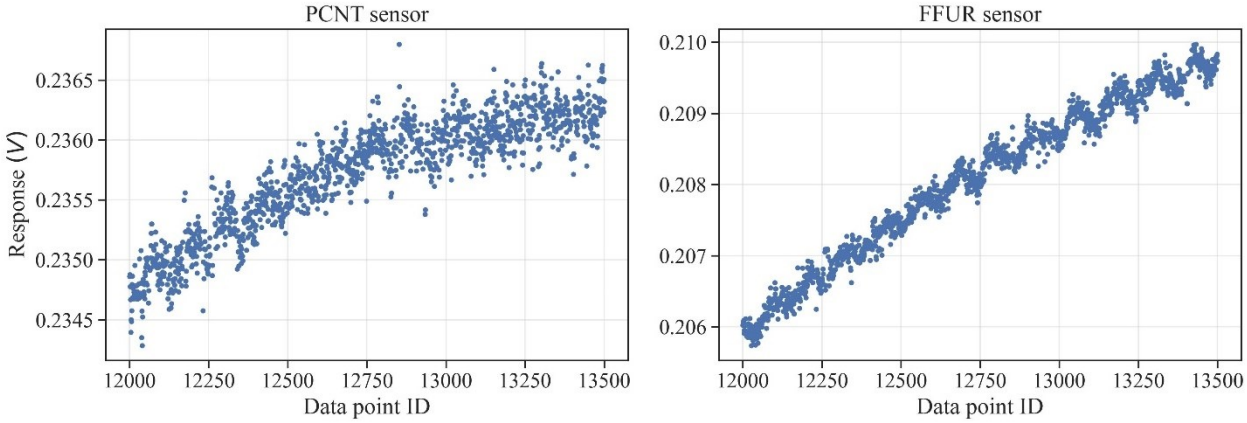

Figure 4.20: Corrected sensor response (after removing the temperature controller artefact)

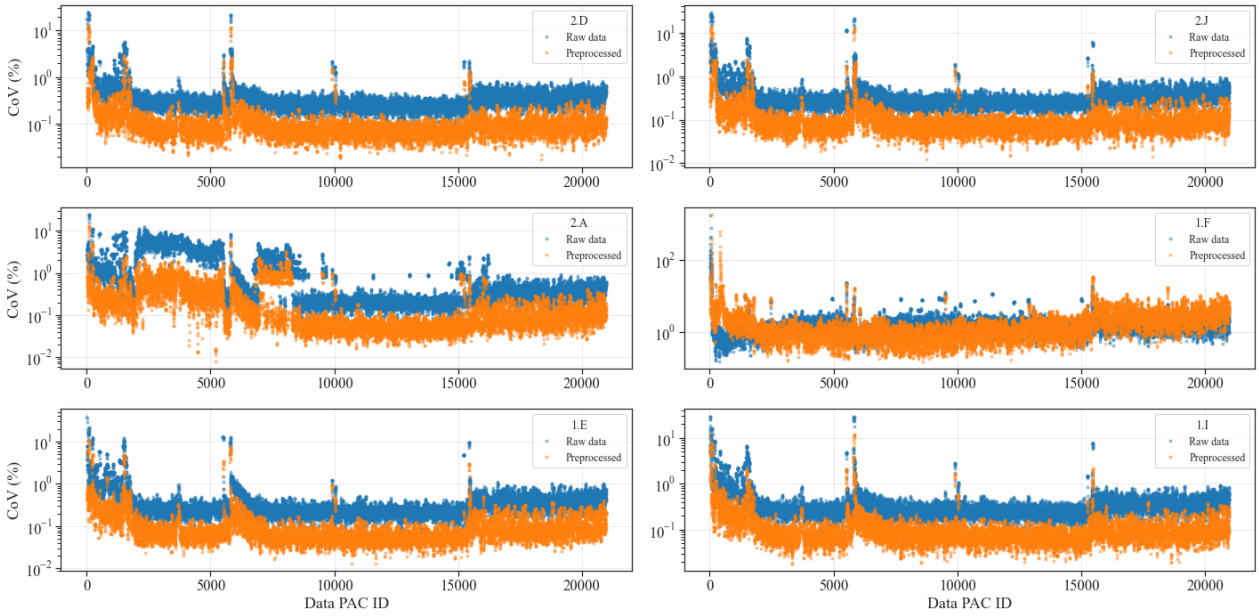

Figure 4.21: CoV of raw data and pre-processed data

4.4.2.3 Model Development and Metrics

In this study, we explore two methodologies for model development: random splits, where the dataset is divided into random subsets of 70%-30% to estimate model coefficients from the 70% portion, and time-structured splits, which leverage data from an initial phase (often the first few hours of the experiment) to maintain chronological integrity in model development. Random splits gauge the system's baseline accuracy, offering insights into the general calibration effectiveness of the sensors without reflecting their capacity for real-time operation. In contrast, time-structured splits aim to mimic the real-world application of the

sensors, predicting their performance in live conditions where models are established shortly after deployment and used subsequently without further calibration.

We continue to utilize the same evaluative metrics, namely the R^2 (Coefficient of determination) and MAE (Mean Absolute Error), with MAE benchmarked against the Naïve MAE of a baseline model that assumes all predictions equal the mean target value, rendering R2 as 0.

Additionally, the study investigates the mutual influences of chlorine on pH predictions and vice versa. This involves analyzing specific subsets of the dataset: one where chlorine is present (HClO concentration ≥ 0.01 mg/L) and another absent chlorine (HClO < 0.01 mg/L). Moreover, a subset containing data with only decreasing chlorine conditions is also evaluated to examine how the sensors' response times are affected under conditions of gradually decreasing chlorine levels. Figure 4.22 and Figure 4.23 provide the scatter plots of each chemiresistor against the target parameters – active chlorine and pH.

Analyzing the scatter plot data reveals several pertinent trends and characteristics of sensor performance. Firstly, PCNT chemiresistors and FFUR chemiresistors (the top row, middle column is the damaged chemiresistor) demonstrate an initial linear response to active chlorine concentration that eventually plateaus, indicating a saturation point. This saturation phenomenon necessitates identifying operational concentration ranges to ensure accurate readings. Moreover, the scatter plots indicate that sensors yield a more distinct second-degree response in active chlorine at lower pH levels, whereas at higher pH levels, the response appears saturated. This suggests that pH levels influence sensor responsiveness to chlorine.

In the context of pH responsiveness, PCNT chemiresistors present a linear response up to a pH of 7.4, after which a saturation effect is observed, resuming a linear trend post pH 7.8. This biphasic behaviour needs further exploration to develop calibration models accurately representing non-linearities across varying pH ranges. The working FFUR chemiresistor, in contrast, exhibits a second-degree response curve, with saturation occurring between pH levels of 7.4 and 7.8, followed by a re-emergence of linear response. However, the plots for PVT ink in the scatter plot data do not display any clear relationship, suggesting complexity in sensor behaviour that may require more experiments to unravel.

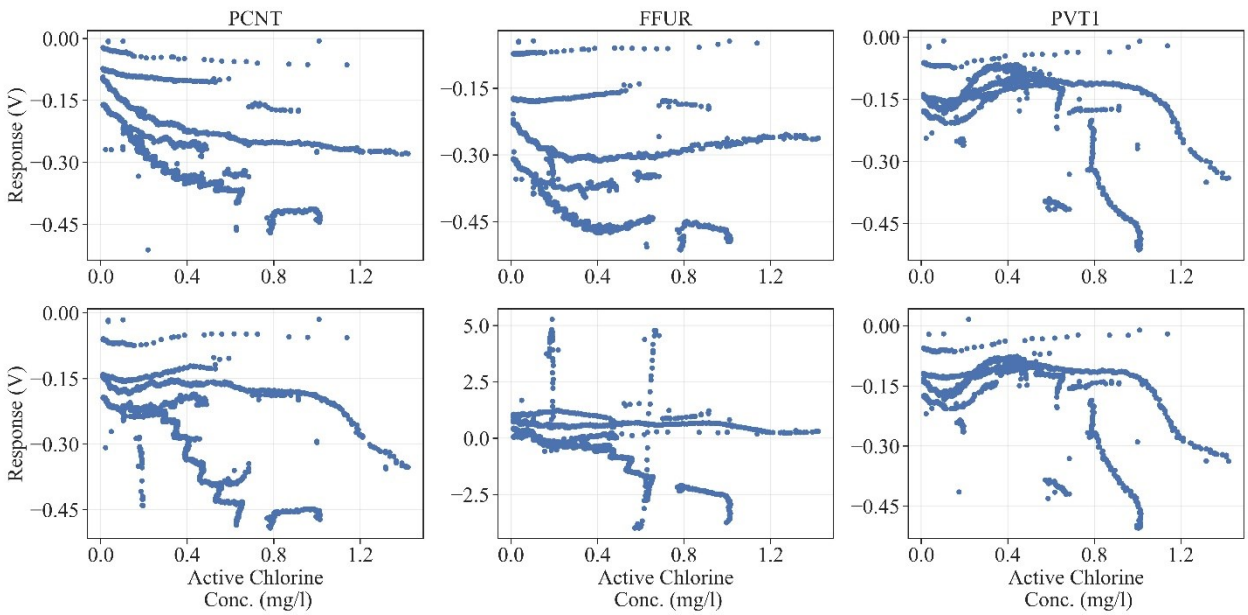

Figure 4.22: Active chlorine vs sensor response (with pH parameter as hue)

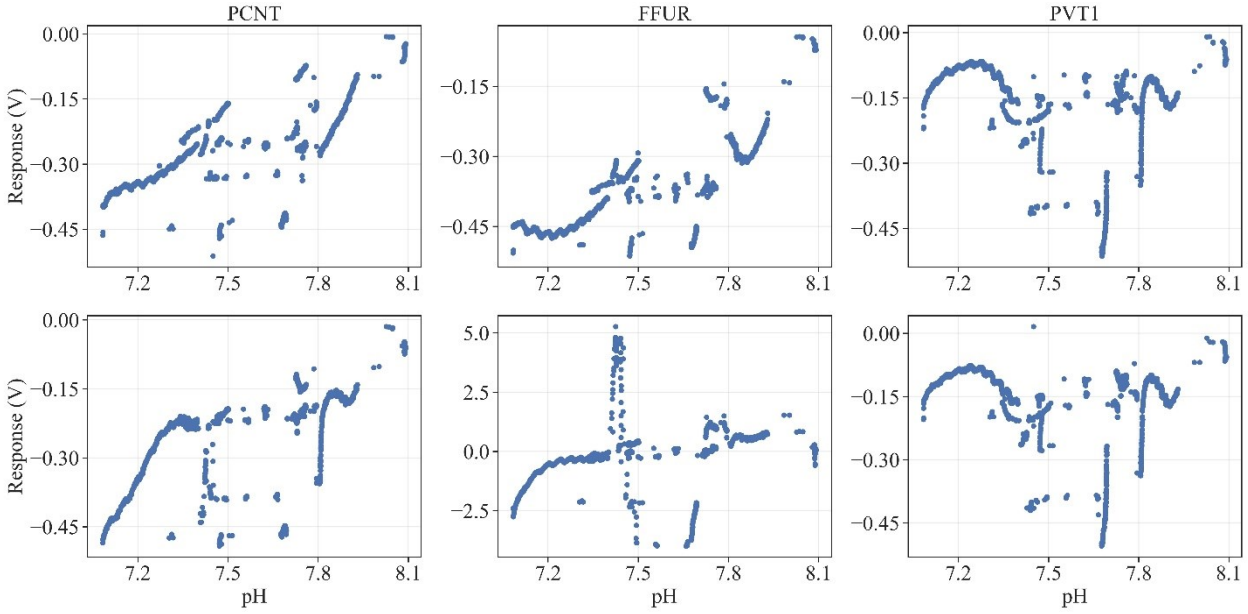

Figure 4.23: pH vs sensor response (with active chlorine parameter as hue)

4.4.2.4 Random splits

Here, the complete 432-hour dataset is randomly split into 70% – 30%, and the model development is repeated 50 times on the 70%, evaluated on the complete data to get the generalized performance. Based on the information from previous experiments, we utilized 3 models: a simple linear model for pH and active chlorine concentration estimation, the second-degree model for the same, and a simple linear model only for pH estimation.

Predicting pH and Active Chlorine from a simple linear model

To simultaneously predict pH and active chlorine levels, the initial approach involves a bi-parameter linear model chosen for its promising results during lab tests. The sensor responses R_1, R_2 are related to the target pH and active chlorine via coefficients a_i, b_i and c_i as shown below:

$$\begin{bmatrix} R_1 & R_2 \\ \vdots & \vdots \end{bmatrix} = \begin{bmatrix} pH & C_{HClO} \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} + \begin{bmatrix} c_1 & c_2 \end{bmatrix} \tag{4.30}$$

$$\Rightarrow R = TC + \phi \tag{4.31}$$

$$\Rightarrow T = (R - \phi)C^{+} \tag{4.32}$$

where C^{+} denotes the Moore-Penrose inverse.

Subsequent inversion across the entire dataset allows for the concurrent estimation of pH and active chlorine values. The performance of this bi-parameter model inversion is showcased in Table 4.18, highlighting the top three outcomes. The chemiresistor pair 1E and 2A, in particular, demonstrate effective overall performance for both pH and active chlorine, though the improvement in active chlorine prediction is minimal compared to the naïve MAE. According to the time series analysis presented in Figure 4.24, while the model capably tracks the overall trends, it occasionally fails to identify specific pH changes.

Table 4.18: pH Active Chlorine Bi-parameter linear model: Random splits. The values in parenthesis are the standard deviations over the 50 random splits

Selected Chemiresistors		pH		Active Chlorine	
		R2	MAE	R2	MAE (mg/l)
1E (PCNT)	2A (FFUR)	0.49 (0.003)	0.18 (0.001)	0.53 (0.01)	0.08 (0.001)
1I (PCNT)	2A	-1.1	0.30 (0.01)	-2.9	0.21 (0.01)
1E	1I	-0.6	0.26 (0.01)	-0.6	0.14 (0.003)
Naïve			0.24		0.09

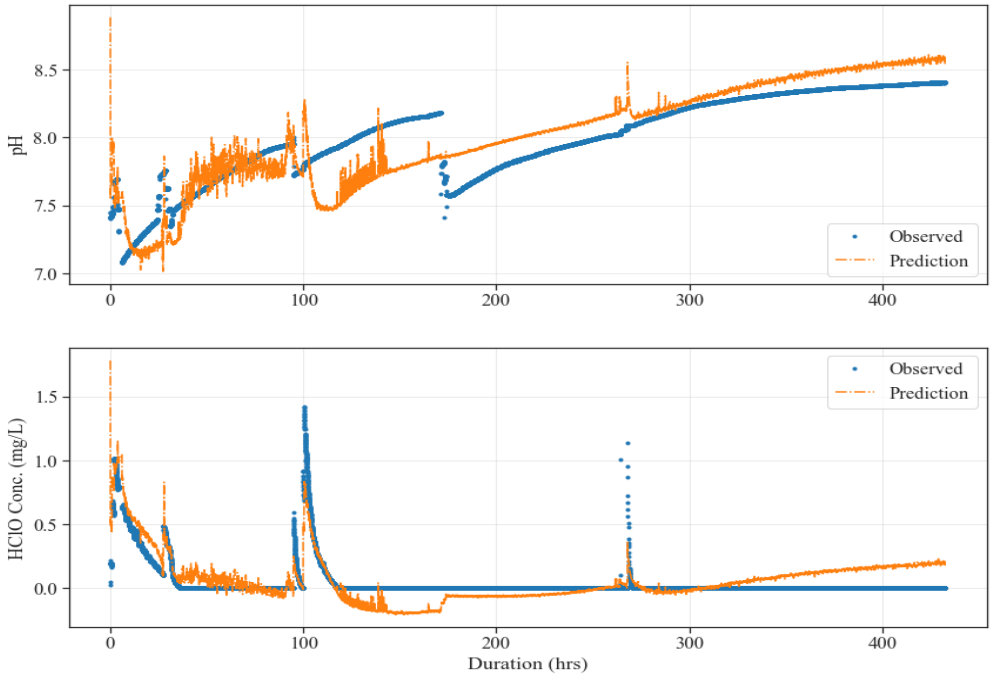

Figure 4.24: pH - Active Chlorine model: Prediction from using chemiresistors 1E and 2A

Predicting pH and Active Chlorine from a linear second-degree model

The scatterplot presented in Figure 4.22 suggests that the active chlorine response is non-linear. So, the following calibration model was identified to represent this scenario:

$$R_i = \alpha_i\, pH + \beta_i\, C_{HClO} + \gamma_i\, C_{HClO}^2 + \phi_i \tag{4.33}$$

Details of the model's performance are outlined in Table 4.19. This model's inversion process, which incorporates the computation of a square root due to the second-degree term of active chlorine, causes numerical instability, necessitating the exclusion of some data

points from the analysis. Consequently, the Naïve MAE for active chlorine varies with each case. While employing this model enhances the MAE and linearity for active chlorine predictions, it does so at the expense of pH estimation accuracy, exhibiting a comparable MAE but a decline in R^2. A time series plot from a randomly chosen split is depicted in Figure 4.25, showing results broadly similar to those from the linear model, with overall trends captured but certain steps missed.

Table 4.19: pH Active Chlorine Bi-parameter higher-order model

Selected Chemiresistors		pH			Active Chlorine		
		R2	MAE	Naïve MAE	R2	MAE (mg/l)	Naïve MAE
1E (PCNT)	2A (FFUR)	0.28 (0.008)	0.20 (0.001)	0.23	0.73 (0.002)	0.05 (0.0003)	0.07
1E	1I	0.13 (0.01)	0.18 (0.001)	0.23	0.50 (0.007)	0.04 (0.0002)	0.07
1E	2D (PVT)	-0.9	0.24 (0.001)	0.22	0.29 (0.003)	0.04 (0.0002)	0.06

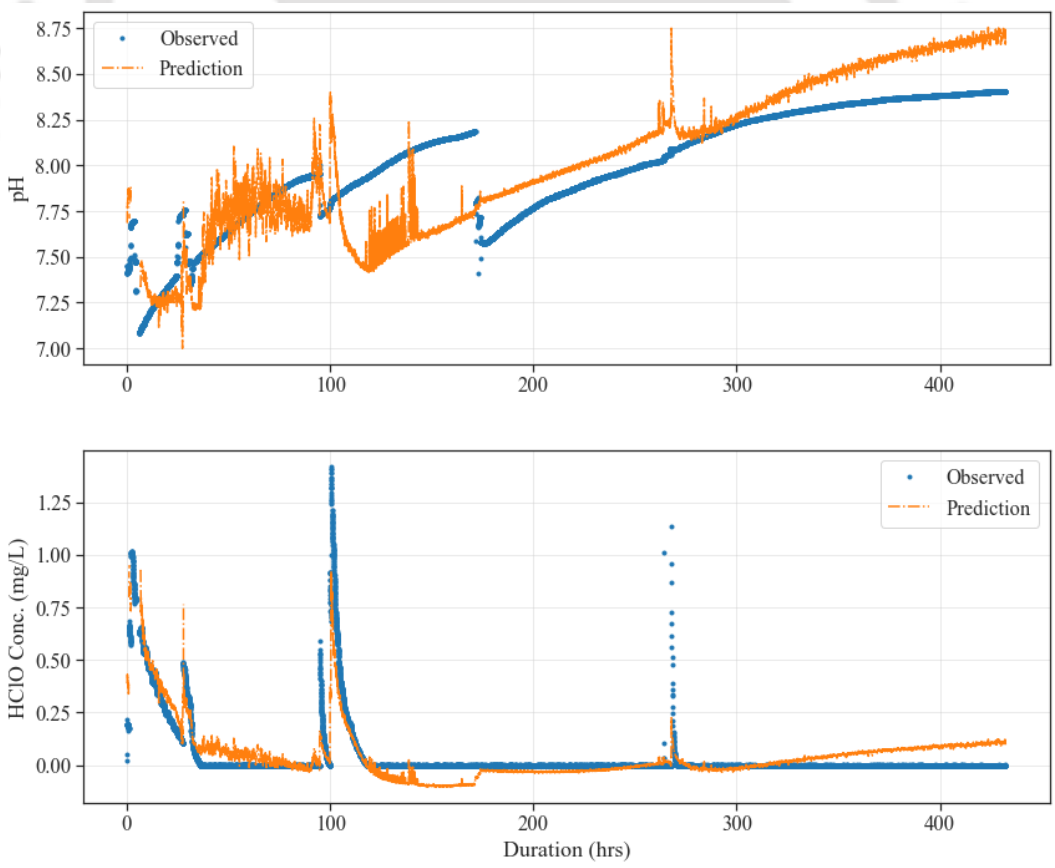

Figure 4.25: pH Chlorine Bi-parameter2nd order model: target prediction

Predicting only pH from Chemiresistor responses

To assess how chlorine impacts pH predictions, we applied a linear model for pH that excluded the active chlorine variable. Table 4.20 details the coefficients for each chemiresistor and the outcomes of the pH prediction via model inversion, with the naïve model MAE being 0.24. This approach improves pH prediction accuracy, as evidenced by lower MAEs, making it a promising strategy for pH estimation. Meanwhile, a bi-parameter, second-order model incorporating pH and active chlorine could be reserved for chlorine prediction.

Table 4.20: Metrics for the pH-only model developed with random splits of data

Chemiresistor	R2	MAE	Device Ink
1E	0.17 (0.008)	0.20 (0.0005)	PCNT
1I	0.44 (0.004)	0.15 (0.0004)	PCNT
2A	0.55 (0.003)	0.15 (0.0003)	FFUR
2D	-0.19 (0.013)	0.22 (0.001)	PVT118
2J	-0.09 (0.01)	0.20 (0.001)	PVT118

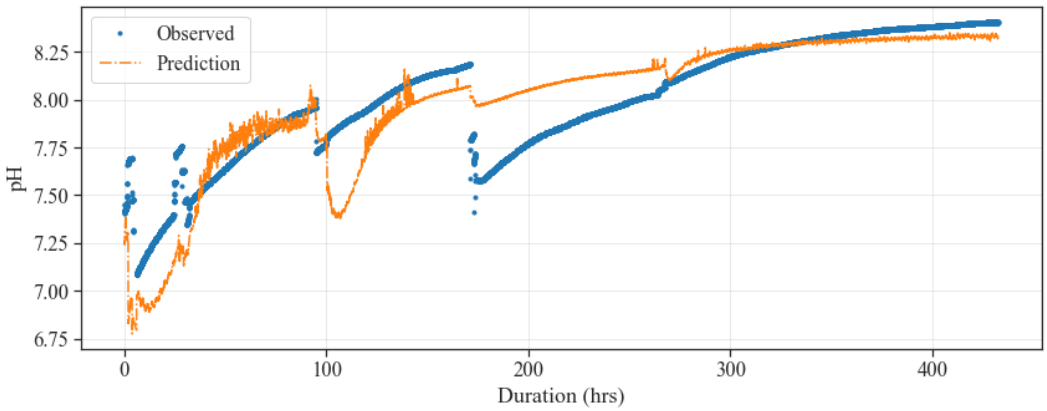

Figure 4.26: pH only model: pH prediction with chemiresistor 1E (PCNT)

Table 4.20 highlights device 2A (FFUR) for its superior performance, marked by the lowest MAE and highest R^2. The pH predictions derived from this device, as visualized in Figure 4.26, show the model's ability to track pH trends and detect shifts in pH levels. However, it does not precisely mirror every change.

Data fusion techniques were explored to enhance these predictions. Specifically, we considered the weighted average predictions from two selected devices as the final

estimation. The top three selections out of twenty possible combinations are presented in Table 4.21. Figure 4.27, which features model predictions using chemiresistors 1I and 2A over time, demonstrates the advantage of data fusion compared to the earlier model (Figure 4.26). While the performance metrics remain similar, the plot from Figure 4.27 shows a noticeable narrowing of the gap between the observed (blue) and predicted (orange) pH values, illustrating the method's effectiveness.

Table 4.21: pH only Model: Metrics of the model with data fusion

Selected Chemiresistors		R2	MAE
1I	2A	0.53 (0.003)	0.15 (0.0003)
2J	2A	0.52 (0.003)	0.15 (0.0003)
2D	2A	0.52 (0.003)	0.15 (0.0004)

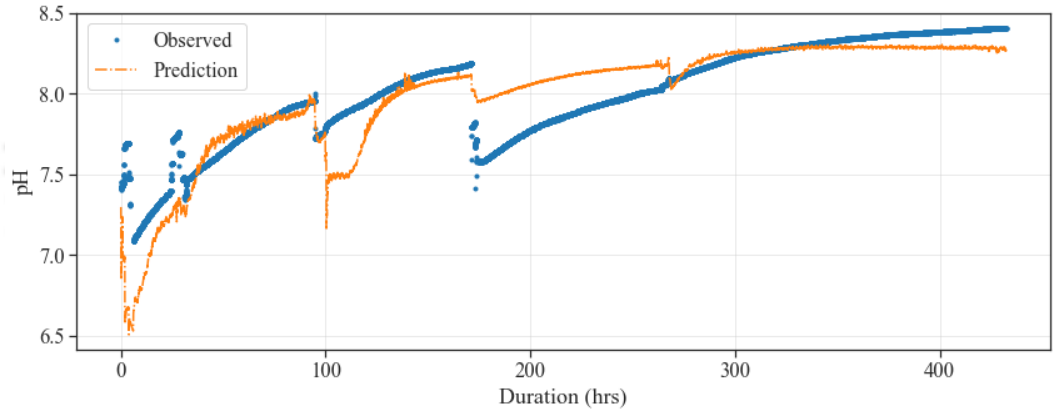

Figure 4.27: pH-only model: pH prediction with data fusion

4.4.2.5 Time Structured Splits

In a real-world setting, model development relies on historical data, with its effectiveness validated against current data to assess performance upon deployment. To approximate this real-time performance, model coefficients were derived using the initial 70 hours of the dataset and then tested across the entirety of the 432-hour dataset. Standard deviations are not included in the subsequent tables since this process was conducted in a single trial.

Predicting pH

Reflecting on the findings from the prior section, it is established that pH can be accurately predicted using a model dedicated to pH. This section explores the model's efficacy when applying a time-structured data split.

Table 4.22 outlines the model's performance across 432 hours of data, with coefficients determined from the first 70 hours. Additionally, the table contrasts predictions for dataset segments with and without chlorine to evaluate prediction accuracy under varying chlorine conditions. Figure 4.28 displays the pH predictions based on the response from chemiresistor 2A, showcasing similarities to the outcomes of random data splits. Generally, the performance aligns with that seen in random splits, with the time series plot revealing comparable issues. While overarching trends are captured, minor variations are overlooked or inaccurately timed. Notably, the pH prediction accuracy remains consistent regardless of chlorine presence, indicating that prediction errors are not directly attributable to chlorine interference.

Table 4.22: pH only Model: Time structured splits

Chemiresistor	Complete data		Learning on the first 70h, predicting on the full dataset with HClO >= 0.01		Learning on the first 70h, predicting on the full dataset with HClO < 0.01	
	R2	MAE	R2	MAE	R2	MAE
1E	0.46	0.17	0.19	0.18	0.26	0.17
1I	0.62	0.15	0.14	0.19	0.55	0.14
2A	0.66	0.15	0.39	0.16	0.56	0.14
2D	-68	2.17	-25	0.86	-112	2.4
2J	-1604	10.6	-813	4.79	-2557	11.6
Naïve MAE		0.24		0.25		0.21

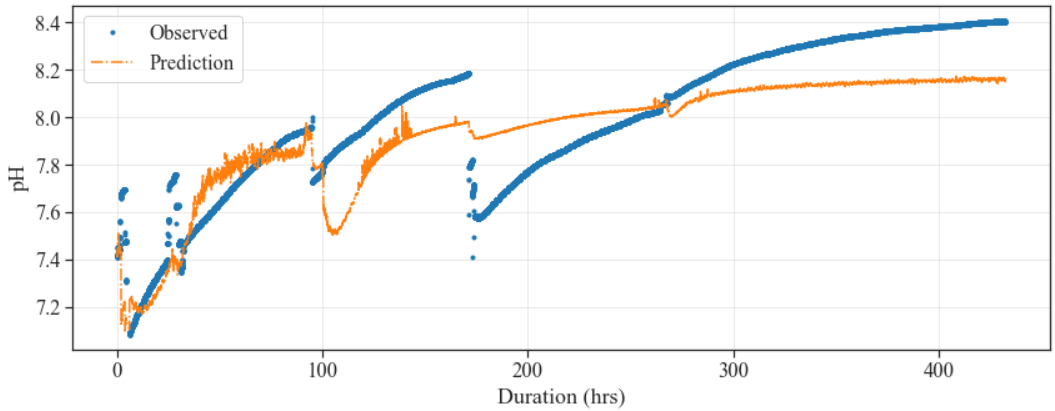

Figure 4.28: pH prediction from pH-only model with time-structured data splits

The enhancements in performance attributed to data fusion for time-structured data splits do not sufficiently justify its implementation despite the noticeably improved signal quality of pH predictions derived from data fusion compared to predictions from individual models. This observation is substantiated by Table 4.23 and illustrated in Figure 4.29, where the predictions are based on chemiresistors 1I (PVT) and 2A (FFUR) combined responses.

Table 4.23: Metric of pH-only Model with data fusion, using Time structure data splits

Chemiresistors		R2	MAE
1I	2A	0.66	0.14
2A	2J	0.65	0.15
1I	2J	0.62	0.15
1E	1I	0.57	0.16
1E	2A	0.58	0.16

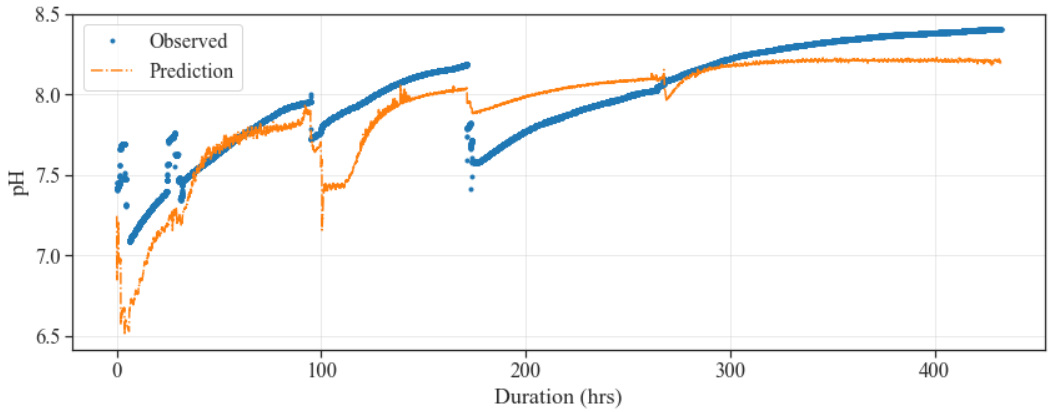

Figure 4.29: pH predicted from the pH-only model with data fusion using time-structured data splits.

Predicting pH and Active Chlorine from a simple linear model

Model coefficients derived from the initial 70 hours of data were applied to analyze performance across a 432-hour dataset. The pH and active chlorine predictions utilizing chemiresistors 1E and 2A, as depicted in Figure 4.30 and summarized in Table 4.24, indicate that the model is not ideally suited for time-structured splits typical of real-world conditions, particularly struggling under zero-chlorine scenarios.

Table 4.24: pH chlorine bi-parameter estimation using time-structured data splits

Chemiresistors		pH		Active Chlorine	
		R2	MAE	R2	MAE (mg/l)
1I	2A	0.42	0.18	-0.63	0.17
1E	2A	0.42	0.2	0.17	0.11
2A	2J	-0.7	0.33	-8.1	0.43
Naïve			0.236		0.088

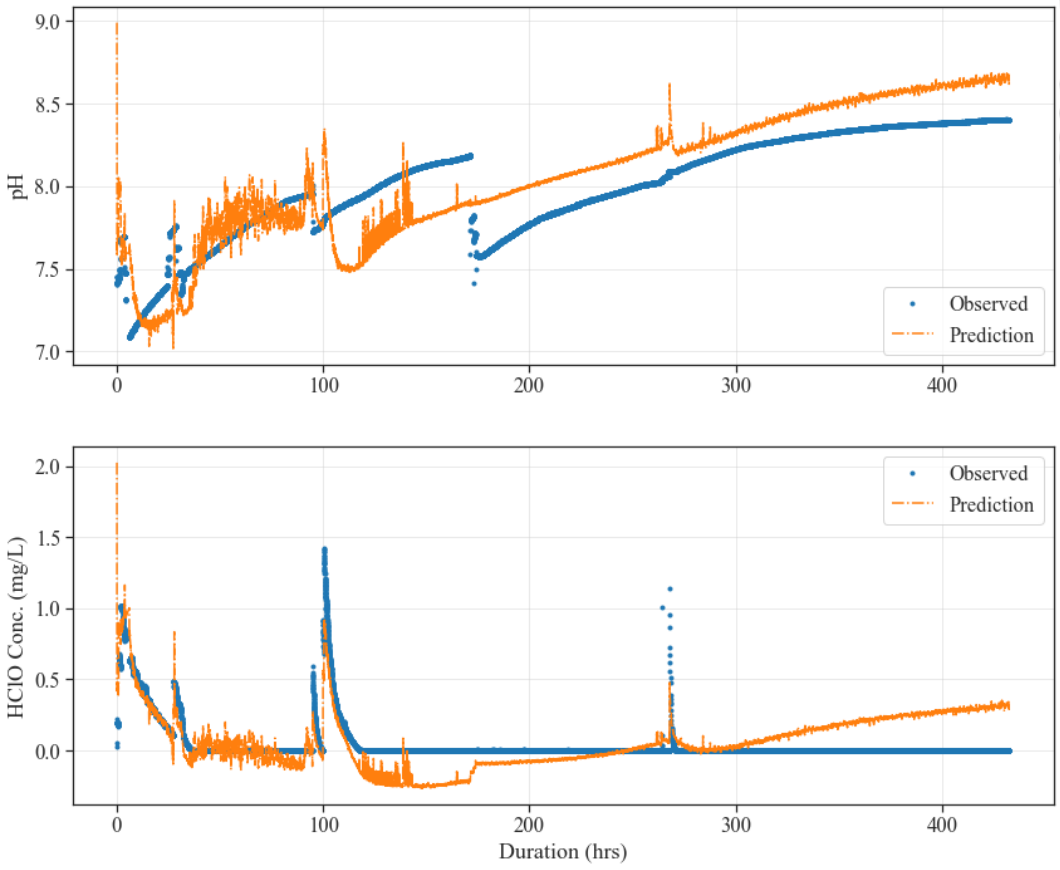

Figure 4.30: pH and active chlorine prediction from the bi-parameter model developed with time-structured data splits

Predicting pH and Active Chlorine from a linear second-degree model

Advancing to a model that includes pH, active chlorine, and a second-order term for active chlorine, coefficients again based on the first 70 hours of data showed improved tracking and error reduction in signal patterns. This enhancement is evident when comparing Figure 4.31 with Figure 4.30, notably mitigating an erroneous uptrend observed in predictions post-300 hours. Although this refined model does not perfectly capture sharp active chlorine peaks, it effectively mirrors broader trends and fluctuations in active chlorine and pH levels, as shown in chemiresistors 1E and 1I in Figure 4.31.

Table 4.25: pH active chlorine bi-parameter estimation (second-degree model) using time-structured data splits

Chemiresistors		pH			Active Chlorine (mg/l)		
		R2	MAE	Naïve	R2	MAE	Naïve
1E	1I	0.30	0.15	0.23	0.37	0.06	0.07
1E	2A	-6.9	0.69	0.24	-1.1	0.18	0.09
1I	2D	-4.2	0.37	0.18	-1.9	0.32	0.21

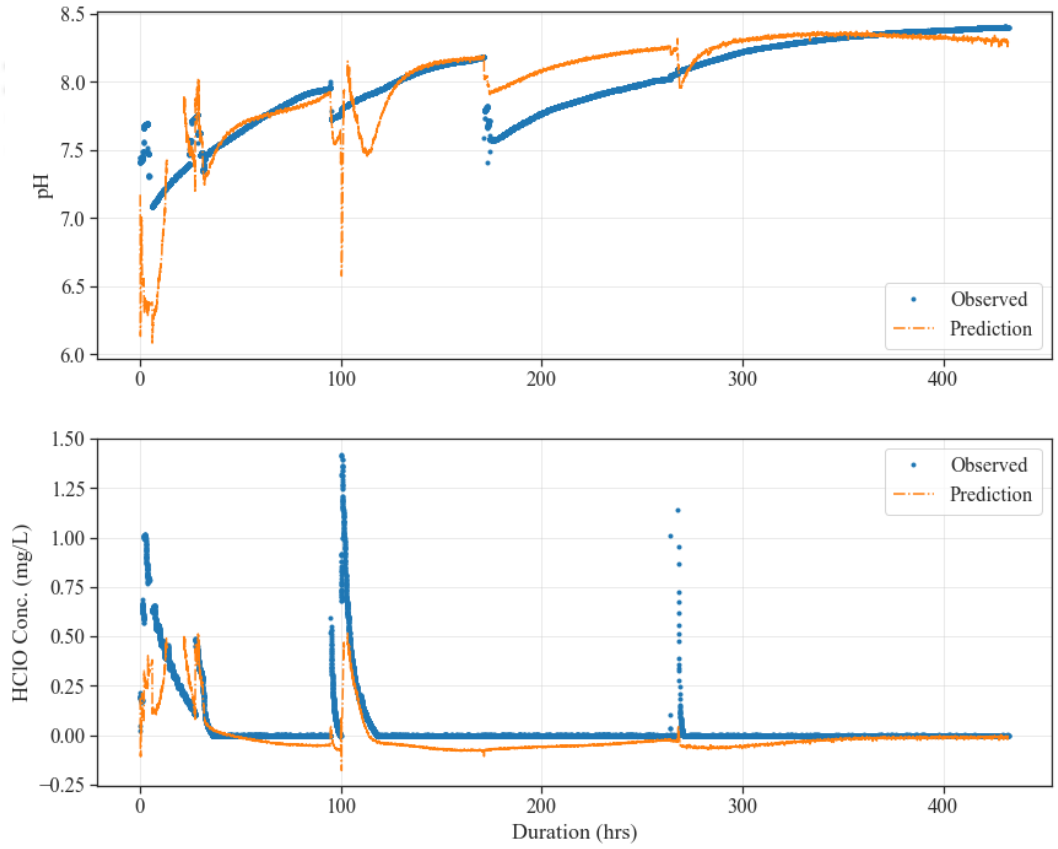

Figure 4.31: pH and chlorine prediction with the bi-parameter higher-order model.

4.4.2.6 Summary of 6D sensor head Model analysis

Table 4.26 compiles the sensor performance within the Sense-City environment over 18 days, excluding naïve MAE values, as all chosen configurations exhibit superior MAE. Figure 4.32 and Figure 4.33 present time series plots from the best model following a time-structured split. The real-time performance metrics post 70 hours of initial learning—0.14 in pH MAE and 0.06 mg/L in active chlorine MAE—satisfy the criteria for both measured parameters. A second-order model for active chlorine combined with a linear model for pH is necessary for accurate active chlorine prediction, whereas pH predictions are most reliable using a pH-only model. The similar outcomes between random and time-structured splits hint at minimal baseline drift in sensor accuracy, which is encouraging.

However, it is critical to note that, despite generally good performance, the models sometimes fail to closely follow rapidly changing conditions, especially with decreasing chlorine levels. This identifies a potential area for further model refinement, to account for sensor response timing. Moreover, the accuracy of the reference chlorine sensor during swift chlorine fluctuations is questionable due to its dependency on 5-minute-long colorimetric assessments, leaving room for speculation about its precision during such fast-changing events.

Table 4.26: Performances of the sensor in Sense-City - summary

		Second-degree model			pH-only model (with data fusion)		
		Chemiresistors	R^2	MAE	Chemiresistors	R^2	MAE
Time-structured split	pH	1E, 1I	0.30	0.15	1I, 2A	0.66	0.14
	Active Chlorine		0.37	0.06 mg/L			
Random splits	pH	1E, 2A	0.28	0.20	1I, 2A	0.53	0.15
	Active Chlorine		0.73	0.05 mg/L			

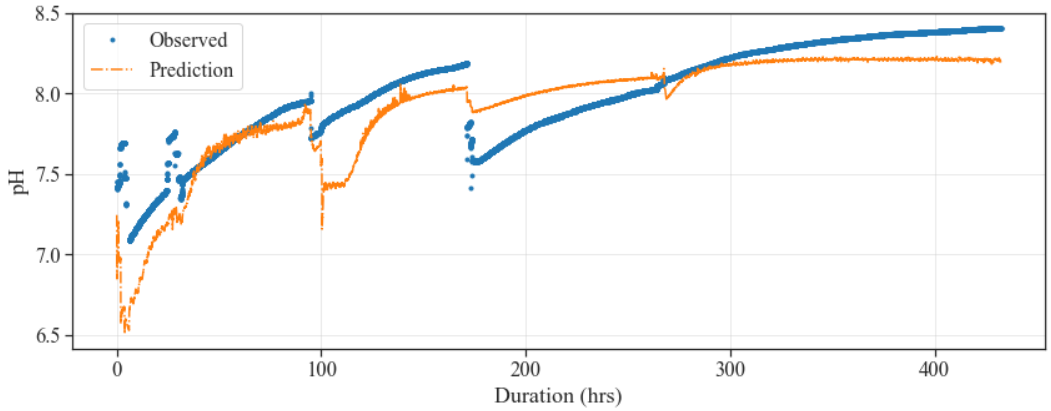

Figure 4.32: Best pH prediction from the pH-only model with data fusion, using time-structured data splits

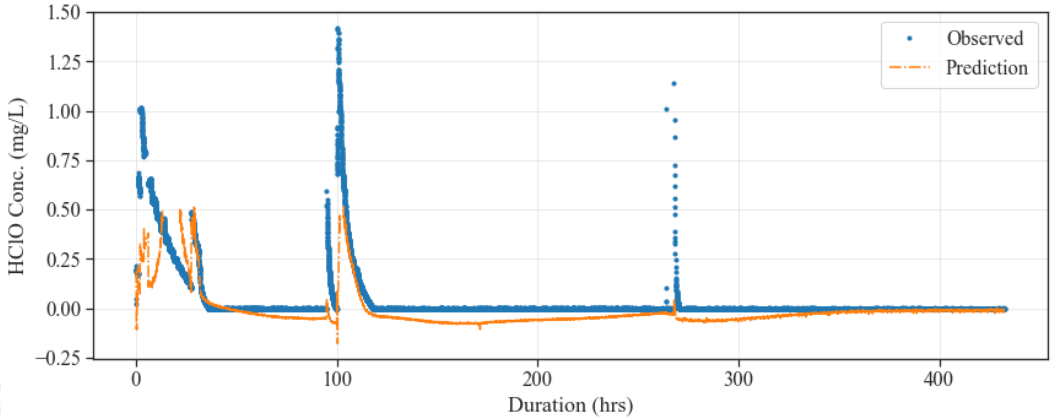

Figure 4.33: Best chlorine prediction with the bi-parameter higher-order model, time-structured split

4.5 Experimentation with field-ready version (LPCB)

4.5.1 Experiments in Sense-city for Chlorine-pH Model development

In the Sense-City water loop (Figure 4.4), two field-ready version LOTUS sensors (LPCB1 and LPCB 2), as illustrated in Figure 4.37, connected to the updated UAFE version, were deployed from September 2023 through February 2024, spanning about 3740 hours. An outline of the significant parts of the data is summarized in Table 4.27, with reference to Figure 4.35.

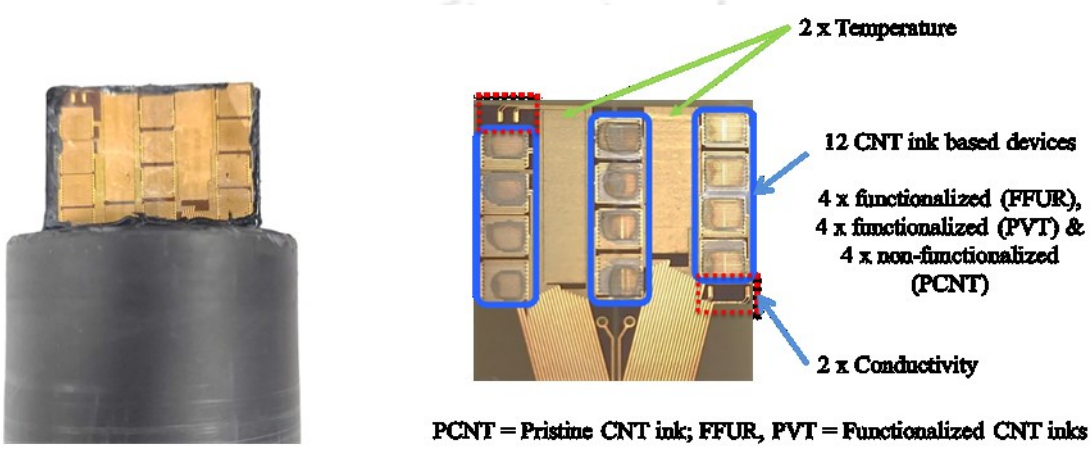

Figure 4.34: Left: assembled sensor head with 12-chemiresistor sensor chip. Right: detailed composition of a three-ink sensor chip for the pre-commercial LOTUS sensor version

The calibration model utilized here remains the previously described second-degree model, with the distinction in the inversion process. Previously, the model was algebraically inverted to estimate target values. However, an optimization strategy is employed in the current approach rather than inverting the equation. Here, the target values (active chlorine and pH) are treated as decision variables, and the error in fitting the direct model equation serves as the objective function. With the relative response from each channel represented as R_i, the calibration coefficients (ranging from $M_{0,i}$ to $M_{4,i}$) are determined directly from the training data through the calibration equation:

$$R_i = M_{0,i}pH + M_{1,i}C_{HClO} + M_{2,i}C_{HClO}^2 + M_{3,i} \tag{4.34}$$

Then, the objective function is defined as the error of fitting the above direct calibration equation for every channel:

$$\min \sum_{i=1}^{12} \left(R_i - M_{0,i}pH + M_{1,i}C_{HClO} + M_{2,i}C_{HClO}^2 + M_{3,i}\right)^2 \tag{4.35}$$

Where C_{HClO} and pH are the decision variables.

Table 4.27: Outline of significant subsets of the complete dataset from the Sense-city experiment

Time range (hrs)	Active Chlorine	pH	Phase
0 – 700	Manual chlorine addition, Chlorine decay ratio varies from 1.8 mg/l/hr to 0.4 mg/l/hr	5 manual pH additions	1a
1000 – 1350			1b
1350 – 1750	Chlorine addition via Chlorine pump, Decay ratio = 0.65 mg/l/hr to 0.75 mg/l/hr	No pH addition	2
1750 – 1950	Temperature regulator was not working		
1950 – 2050	Reference sensor error		
2050 – 3100	Chlorine addition via Chlorine pump, Decay ratio = 0.27 mg/l/hr to 0.55 mg/l/hr, but lower concentration	2 manual pH additions	3
3100 – 3290	Low chlorine concentration (< 0.05 mg/l)	No pH addition	X
3290 – 3740	Water loop stopped		

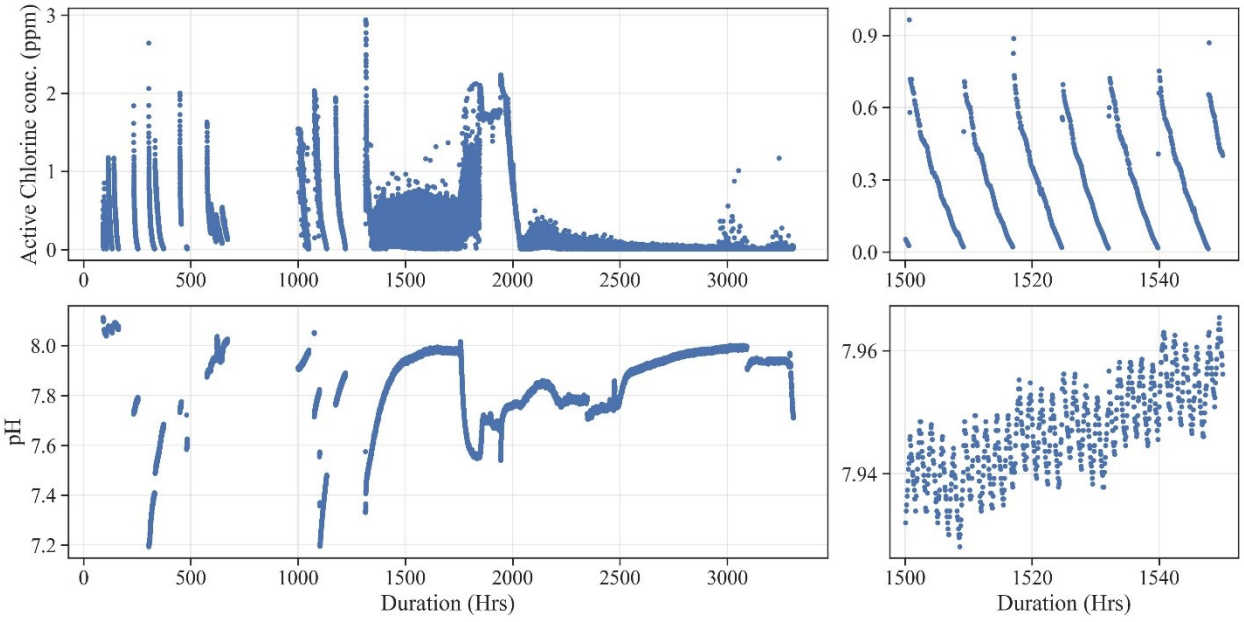

Figure 4.35: Complete dataset collected from Sense-city between September 2023 and February 2024: The top plots correspond to Active chlorine, and the bottom corresponds to pH; the left plots show the complete 3300 hours of data, and the right plots show the data during 1500-1550 hours (where the effect of chlorine pump can be seen).

4.5.1.1 Model development and Validation in Phase 1a

During the first phase of experiments, changes in chlorine concentrations and pH levels were manually varied by adding sodium hypochlorite (bleach) and hydrochloric acid to the tank. The sensors were left undisturbed in the loop for stabilization for the first 100 hours (~ 4 days), so no experiments were performed.

Phase 1a denotes the duration from 0 to 700 hours, first 30 days. The data from this duration was extensively studied to upgrade the active chlorine-pH model. The model training was done using the first 10 days (~230 hours), and the model performance was evaluated using the dataset till the end of Phase 1a, 700 hours. After smoothing the data, the model calibration was conducted up to 1mg/l of chlorine (0.4mg/l active chlorine), as at higher chlorine levels, the sensors exhibited saturation. The chemiresistors were chosen with the criteria of having $R^2 > 75\%$ and the parameter coefficient variability below 10% in the direct model. The coefficients of the chosen sensors are outlined in Table 4.28, and the model's performance built with the Phase 1a dataset is shown in Table 4.29.

Table 4.28: Coefficients of selected chemiresistors for active chlorine pH estimation, obtained using Phase 1 dataset (0 – 700 hrs) from LPCB 1

Chemiresistor	pH (%/pH unit)	HClO (%/mg/L)	$HClO^2$ $(\%/mg/L)^2$	R^2
R11	0.12	-2.08	1.92	0.98
R14	0.34	-1.85	1.59	0.98
R21	0.40	-1.87	1.94	0.89
R32	0.34	-0.93	0.75	0.96
R33	0.28	-2.29	2.24	0.92
R34	0.55	-1.76	1.73	0.97

Table 4.29: Performance of Active chlorine – pH model developed from Phase 1 dataset of LPCB 1

	pH	Active Chlorine (mg/l)
R^2	0.44	0.91
Model MAE	0.15	0.03
Naïve MAE	0.22	0.10

In this phase, the sensors displayed accurate tracking of chlorine levels and good adherence to pH steps, though pH predictions exhibited fluctuations across steps, as seen in Figure 4.36. These were attributed to persisting instabilities in the upgraded UniAFE used in the experiment. Over the assessed month, the performance criteria were met numerically without applying algorithms developed for compensating sensor drift and baseline shifts.

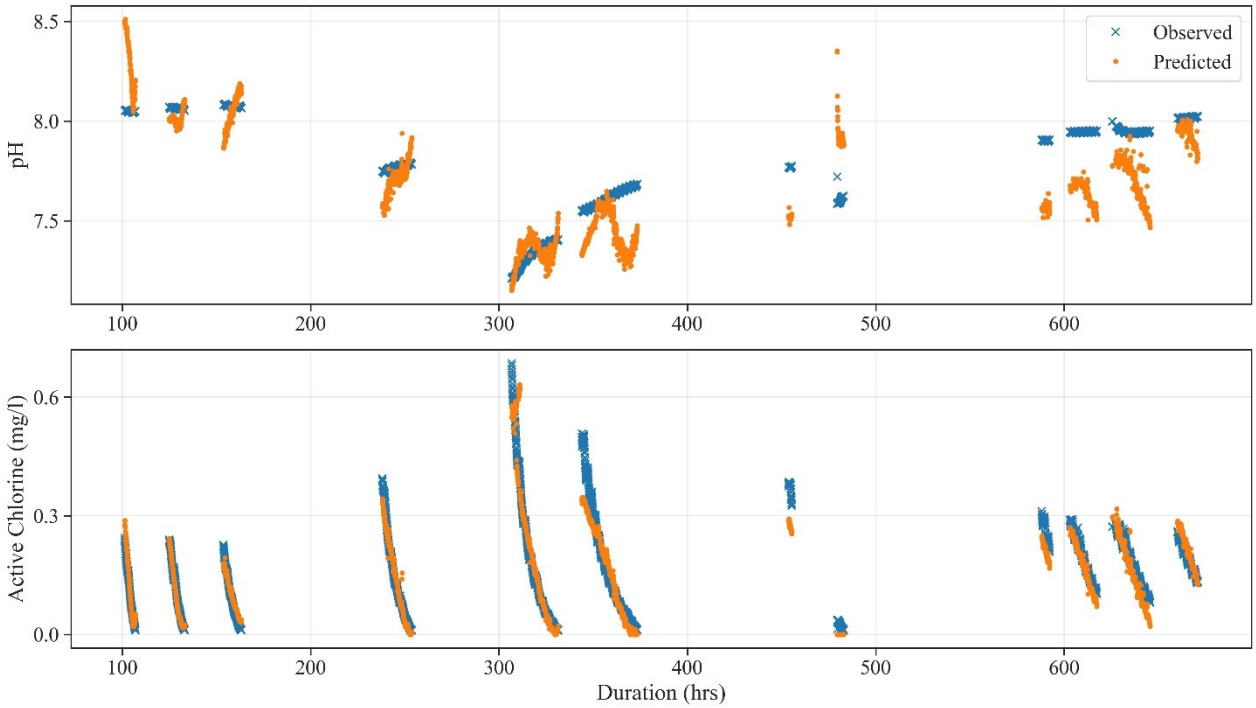

Figure 4.36: Sensor chlorine and pH prediction over Phase 1a experiments in Sense-City

This result (in Table 4.29) highlights the potential of the multiparametric sensor. The obtained results, in terms of average error in chlorine sensing and pH sensing, are comparable with those reported in the literature, and the longevity of 30 days in continuous water testing is more than reported in the state of the art for both chlorine and pH sensing using CNT based sensors. The differences in average calibration coefficients observed across the three inks and 2 LPCB sensors during the first 10 days (dataset used in model calibration) highlight the different ink's role in modulating sensitivity, as presented in Table 4.30. This variability aids in quantifying the different target parameters, which are the active chlorine concentration and pH levels.

Table 4.30: Variability in calibration coefficients for both LPCB sensors with 3 inks

	pH	HClO	$HClO^2$	Intercept
CNT	36%	3.8%	2.1%	28%
FFUR	19%	1.9%	7.7%	10%
PVT	28%	34%	26%	16%

Assuming a standard deviation in resistance below 6kΩ for resistances under 40kΩ (a variance of less than 15%), the expected standard deviation in sensitivity would be below 6% for pH and HClO (up to 35kΩ) in pristine CNT. For FFUR, this deviation remains under 30% across various pH, HClO, and $HClO^2$ conditions. At the same time, PVT stays below 7% for HClO and below 20% for $HClO^2$, indicating desirable resistance ranges for maintaining below 30% variability in sensitivity. Specifically, resistances under 15kΩ are recommended for PCNT and PVT inks to ensure less than 15% variability with these two inks alone.

In summary, it is feasible to amplify sensitivity distinctions between different inks and induce significant sensitivity variations within a single ink type by controlling resistance. Aiming for specific resistance ranges (below 15kΩ for PVT and pristine CNT and above 40kΩ for FFUR) minimizes sensitivity variability, enabling pronounced sensitivity differentiation. These findings are pivotal for device productization, though they stem from a limited sample size of PCBs and devices, introducing uncertainties in baseline trend predictions. Hence, these preliminary results call for cautious application and further validation in larger device batches in the future.

The dataset after 700 hrs till the 3740 was obtained from the Sense-city database at a very later stage, so we could not do extensive analysis. Also, between each phase, significant events change the chemiresistor baseline and sensitivity, rendering the model developed with Phase 1a ineffective for continuing ahead. However, the model can be retrained in each phase and evaluated until it fails, highlighting the sensor's usability. In future work, this retraining will be automated so the model developed initially can be utilized throughout the experiment duration. The model will also be improved to accommodate for the baseline changes.

Therefore, the results presented in the following sections are not representative of the technology developed in this thesis but are provided only for completeness.

4.5.1.2 Model Validation Extended to Phase 1b

In the gap between Phase 1a and Phase 1b (~ 12 days), the LOTUS sensor data collection system failed, and no data was collected. This meant that the sensor was effectively inactive in the continuous water loop, which may have led to a baseline change in the sensor. This particular batch of sensor also had complications from the ink fabrication and needle tapping (as discussed in Chapter 3 Section 3.2.3). For continuity, the same model developed in Phase 1a was extended to evaluate in the 15 days in Phase 1b. The performance of the same model in Phase 1b compared to Phase 1a is outlined in Table 4.31.

Table 4.31: Performance of Model developed using data from Phase 1a in Phase 1a and 1b

Parameter	Phase	R^2	Model MAE	Naïve MAE
pH	1a	0.44	0.15	0.22
	1b	-2	0.35	0.23
Active Chlorine (mg/l)	1a	0.91	0.03	0.10
	1b	-0.08	0.11	0.10

This poor performance may be attributed to the change in the chemiresistor baseline, to the noisy response in the reference sensor (only during Phase 1b), or to the complications in the sensor (fabrication) or a combination thereof. Due to the lack of a varied dataset (from more than 2 LPCBs) and lack of extensive analysis (the dataset was obtained after Feb 2024), it was impossible to identify the exact cause.

Owing to similar reasons, along with the change in the water matrix and introduction of the chlorine pump in Phases 2 and 3, the model developed in Phase 1a performed very poorly in those phases, with R^2 being negative in all the cases.

4.5.1.3 Model development and Validation in Phase 2

In the second phase of experiments, an automatic chlorine addition pump was introduced to continuously add a predetermined quantity of bleach solution to the tank at a regular interval (every 8 hours). Also, before the start of the experiment, the loop was cleaned, thereby leading to a different water matrix. Consequently, the sensor baseline and sensitivity

exhibited variations, which were inconsistent across all the chemiresistors. The first 1-week data was ignored in this phase due to sensor baseline stabilization. So, the data from the remaining 300 hours (1450 to 1750 hours) was utilized to develop and validate the model.

In the 12.5 days dataset, model coefficients were estimated using the initial 4 days and evaluation was performed over the complete duration. From Figure 4.37, We can see that there was very minimal change in the pH levels after 1550 hrs (since 1450 to 1550 is used in training), and this led to a poor prediction performance in pH (as shown in Table 4.33).

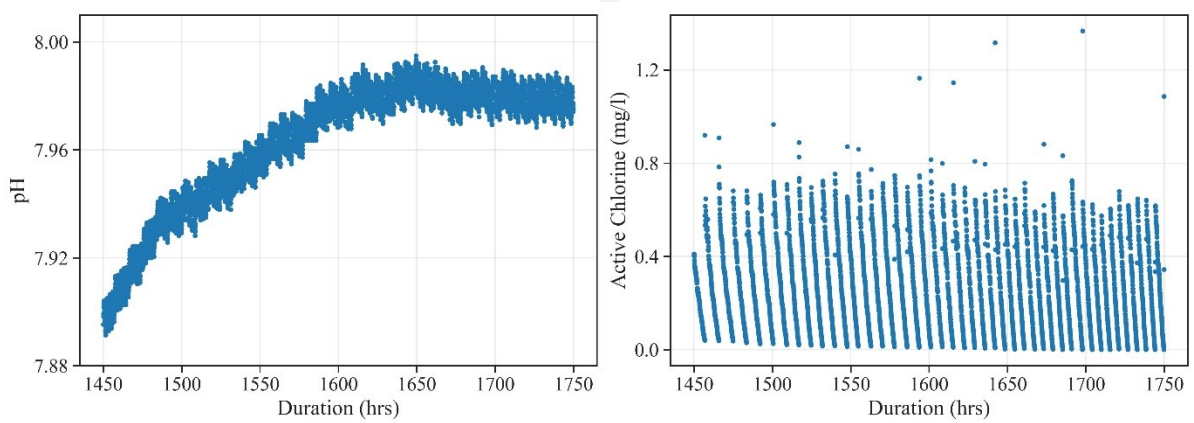

Figure 4.37: Measurements from pH (left) and Active chlorine(right) from the reference sensors in Sense-city during Phase 2

Six chemiresistors were chosen from the dataset (Table 4.32) based on the same criteria defined for model development for Phase 1a, and the performance metrics for evaluation in Phase 2 and Phase 3 are shown in Table 4.33. The data analysis showed that compensating for drift improved the model performance drastically.

Table 4.32: Coefficients of selected chemiresistors for active chlorine pH estimation, obtained using Phase 2 dataset (1450 - 1750 hrs) from LPCB 1

Chemiresistor	pH (%/pH unit)	HClO (%/mg/L)	$HClO^2$ $(\%/mg/L)^2$	R^2
R11	-0.08	-1.48	1.68	0.80
R13	0.44	-4.20	5.27	0.80
R14	0.71	-0.79	1.01	0.76
R21	-0.16	-1.61	1.90	0.80
R24	0.92	-0.48	0.65	0.79
R33	0.21	-1.51	1.75	0.79

Table 4.33: Performance of Active chlorine – pH Model developed in Phase 2 and evaluated in Phase 2 and Phase 3, LPCB 1 dataset.

Parameter	Phase	R^2	Model MAE	Naïve MAE
pH	2	-0.70	0.03	0.02
	3	-1	0.33	0.04
Active Chlorine (mg/l)	2	0.05	0.04	0.08
	3	-1	0.06	0.05

There were no pH experiments during Phase 2, so in the evaluation dataset (1550 to 1750 hours), the pH was almost linear, with an average value of 7.98 and a standard deviation of 0.01 units. This is the primary reason for the negative R^2 score in the performance. Also, before Phase 3, between 1750 and 2050 hours, the temperature regulator failed, heating the water in the loop to above 45°C for about 200 hours. This prolonged exposure to hot water may also have modified the sensor baseline, which can also be one of the contributing factors that led to the poor performance of previously developed models, evaluated in Phase 3 (both Phase 1a model and Phase 2 model gave negative R^2).

4.5.1.4 Model development and Validation in Phase 3

In the third phase of experiments, the chlorine concentration in the bleach solution was reduced (via evaporation) to a deficient concentration, the maximum level of which deteriorated rapidly from 1 mg/l per addition to 0.4 mg/l per addition within 2 days. Then, it slowly reduced to 0.2 mg/l per addition after about 17 days. Apart from this reduced chlorine levels, only 2 pH level changes were performed during the first 600 hours, and after that, the pH variations were limited between 7.9 and 8.0 units. The first 100 hours (~ 4 days) of the phase 3 dataset was utilized to estimate the model coefficients, and the prediction was performed till the end of phase 3 (3100 hours ~ 36 days). The selected chemiresistors are shown in Table 4.34, and the performance, depicted in Table 4.35, shows that the chemiresistors could estimate chlorine with an accuracy of ± 0.03 mg/l, in the range of 0 to 1 mg/l.

Table 4.34: Coefficients of selected chemiresistors for active chlorine pH estimation, obtained using Phase 3 dataset (2040 - 3100 hrs) from LPCB 1

Chemiresistor	pH (%/pH unit)	HClO (%/mg/L)	$HClO^2$ $(\%/mg/L)^2$	R^2
R11	3.26	-0.78	1.02	0.91
R21	3.30	-0.11	-0.67	0.92
R31	3.05	-0.24	0.20	0.90
R32	3.41	0.76	-3.17	0.92
R33	2.57	0.19	-1.42	0.89
R34	3.63	0.04	-1.16	0.90

Table 4.35: Performance of Active chlorine – pH Model developed in Phase 3 using LPCB 1 dataset.

Parameter	R^2	Model MAE	Naïve MAE
pH	-1	0.42	0.08
Active Chlorine (mg/l)	0.24	0.03	0.05

4.5.1.5 Synthesis of Model development in the Sense-city dataset

The chemiresistors showed promising results in terms of tracking chlorine with minimal error. The sensor performance was better in Phase 1a for pH monitoring than in the other phases. The performance degradation of the model throughout the phases can be attributed to several reasons, primary ones being: change in the water matrix, introduction of chlorine pump, extended exposure to hot water stream, longer inactive periods of the chemiresistors, very low chlorine concentrations, not enough pH variations for the model to learn from and complications in sensor fabrication line improper ink deposition, uneven ink dispersion onto the substrate, or needle tapping.

However, within each phase, we can see that the sensor's performance continues longer. In phase 1a, the model was developed with only 10 days of data but was tracking chlorine for the next 30 days. Similarly, in phase 2, the model was valid for 10 days after being retrained with 4 days data and in phase 3, the model was active for 36 days after being retrained with only 4 days data. This highlights the viability of the sensor even after ~130 days, which is much higher than reported in the literature (< 30 days). Similar performances were also obtained with the LPCB 2 sensor dataset.

4.5.2 Temperature Sensor characterization at IITG Waterlabs testbench

Regarding temperature measurement accuracy, the performance of two different LPCB sensor heads was assessed. One LPCB featured temperature sensors with a resistance of 14Ω (Type B), while the other LPCB had temperature sensors with resistances between 36Ω and 39Ω (Type A). The sensors were excited with a sine wave signal of 5kHz frequency and 1mA amplitude. This design adjustment in Type A sensors—achieved by thinning and narrowing the copper lines of the temperature sensor on the PCB—was intended to explore performance variations. Experiments involved placing the sensors directly into a chiller (ANM Industries, WB2100V, -20°C to 100°C [229]) filled with reverse osmosis water and systematically varying the water temperature between 10°C and 40°C. Calibration models were linear, with non-zero intercepts, and exhibited exceptional linearity, consistently above 0.95 (Table 4.36). The average error across the five devices ranged from 0.7°C to 0.9°C, closely matching the reference sensor's accuracy, indicating consistent performance.

Table 4.36: Temperature Sensors Prediction Performance

	Sensitivity (mV/°C)	Noise in response (mV)	Linear model, R2	Linear model, MAE (°C)
Type A, T1	0.079	0.071	0.97	0.75
Type A, T2	0.078	0.071	0.97	0.74
Type A	**0.079**	**0.071**	**0.97**	**0.75**
Type B, T1	0.031	0.034	0.95	0.89
Type B, T2	0.031	0.033	0.95	0.90
Type B	**0.031**	**0.034**	**0.95**	**0.90**

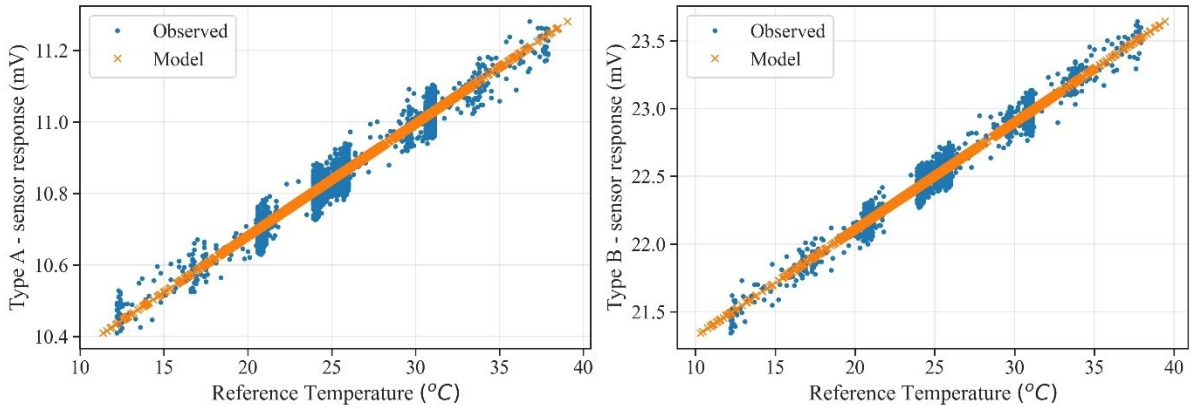

Figure 4.38: Temperature sensor model development

It was observed that the AFE noise for the temperature channel increased with decreasing sensor resistance. As a result, the Type A sensor variant was proposed to counteract this issue. However, due to the Type A sensors' increased sensitivity compared to Type B, this adaptation merely balanced out the increased AFE noise (as seen in Figure 4.38), offering no net benefit. Moreover, the thinner copper lines of Type A sensors potentially compromise their durability. A third sensor variant, designed to maintain the original resistance levels while preserving copper line dimensions similar to Type B (thus aiming to retain sensitivity), is under development to further enhance MAE performance.

4.5.3 Effect of Temperature on Chemiresistors

At Sense-City, temperatures are tightly controlled between 24 to 26 degrees Celsius, which has a negligible effect on chemiresistor performance and their integration within models. However, the IITG Waterlabs facility provided a broader range of temperature fluctuations during temperature sensor characterization experiments, allowing for a detailed assessment of temperature effects on chemiresistors. Owing to a scarcity of devices within the optimal operational resistance range (below 100 kΩ), the study at IITG Waterlabs was performed using a single device incorporating two distinct inks.

The analysis revealed that chemiresistors with FFUR ink exhibited a decrease in resistance of 1.27kΩ for every 1°C increase in temperature, while those with CNT ink showed a change of 1.15kΩ. On average, the temperature coefficient was determined to be 1.22 kΩ/°C, with a standard deviation of 0.18 kΩ/°C. These findings are derived from a single device, so they should be considered illustrative and primarily for demonstration purposes. In light of these

insights, future upgrades to the model will reincorporate temperature compensation measures to mitigate the identified temperature-related effects on sensor performance.

4.5.4 Conductivity Sensor characterization at IITG Waterlabs testbench

The performance in conductivity measurement was assessed using two different LPCBs, B2 and A1, as mentioned previously. Each LPCB is equipped with two operational conductivity sensors, C1 and C2, varying in size. Additionally, consistent with earlier observations, the copper thickness in the A1 conductivity sensors is halved compared to those in the B2 devices, affecting the baseline capacitance by reducing it for the C1 and C2 sensors in A1.

The experiments were conducted in the same chiller as before, with changes in conductivity achieved by altering the NaCl concentration in reverse osmosis water, within the range of 200 to 2100 µS/cm. Conductivity measurements were taken using an instrument with 1% accuracy, translating to an approximate accuracy of 4 to 6 ppm across the datasets under consideration, maintaining the chiller at 2 different temperature ranges, 15°C, and 30°C. It was observed that during separate testing phases for both PCBs, noise levels for both C1 and C2 sensors would unexpectedly spike before normalizing; steps exhibiting these anomalies were omitted from the evaluation.

Performance results are compiled in Table 4.37, showcasing calibration outcomes across the entire dataset. The sensors' outputs, either from C1 or C2, demonstrated a linear relationship (with a non-zero intercept) with the inverse square root of conductivity ($V_{Ci} \propto \frac{1}{\sqrt{Conductivity}}$), exhibiting high linearity (>0.95) with minimal temperature interference. The derived calibration equation for predicting conductivity C_P, from sensor response V_{Ci} is:

$$C_P = \left(\frac{a_0}{V_{Ci} - a_1}\right)^2 \tag{4.36}$$

Performance metrics ranged between 15 to 50 ppm (30 to 100 µS/cm), except for B2F C2, which can be attributed to the sensor noise. Overall, Type A devices outperformed Type B, with C1 sensors showing better accuracy than C2. Unfortunately, the response for sensor C2, for the experiments at temperatures 15°C and 25°C, was so noisy that no significant signal was identified. Consequently, the same is removed from the analysis.

Table 4.37: Conductivity Sensors Prediction Performance

Type	Temperature (°C)	Sensitivity (V µS/cm$^{0.5}$)	Noise (mV)	R2	MAE (µS/cm)
B2F C1	15	2.63	5.4	0.97	49.6
	25	2.30	8.5	0.70	70.3
	30	1.78	4.9	0.90	82.2
B2F C2	30	0.57	2.1	0.70	198.7
A1 C1	15	2.93	2.5	0.98	31.5
	25	2.45	2.5	0.99	11.8
	30	2.44	6.2	0.94	91.3
A1 C2	30	0.83	2.5	0.96	73.5

We also see from Table 4.37 the influence of temperature on sensitivity, so including temperature in the model equation (4.36) yielded the final calibration equation after an exhaustive search of other linear variants with sub-standard performance in lower R2. The final calibration model for Conductivity sensing with temperature compensations from temperature sensor T_j (T1 or T2):

$$C_P = \left(\frac{\alpha T_j + \beta}{V_{Ci} - a_1}\right)^2 \tag{4.37}$$

Where α and β are the coefficients and a_1 is the bias. The performance metrics of the conductivity sensor after using the calibration model defined in *(4.37)* are outlined in Table 4.38.

Table 4.38: Performance metrics of Conductivity sensor C1 with Temperature input

Type – Temperature Sensor	Model Coefficients			R2	MAE (µS/cm)	NMAE (µS/cm)
	α (µS/cm)$^{0.5}$	β (V (µS/cm)$^{0.5}$)	a_1 (V)			
Type B – T1	-1380.1	17.2	-0.027	0.84	117.2	368.7
Type B – T2	-1414.2	17.4	-0.027	0.85	116.6	368.7
Type A – T1	-513.6	14.1	-0.025	0.96	67.6	369.3
Type A – T2	-520.1	14.0	-0.025	0.96	67.9	369.3

4.5.5 Uncertainty Quantification of field ready (LPCB) version

For temperature sensors, the same model *(4.4)* described in the silicon version is utilized here; therefore, via the GEP formulae, the uncertainty is calculated using the equation *(4.26)*. The individual uncertainties are tabulated in Table 4.39, and based on the final equations, the uncertainty in Temperature sensing for each type is outlined in Table 4.40

Table 4.39: Uncertainties associated with the sensor response and model coefficients

Type	Chemiresistor	$\sigma_{Response}$ (Ω)	Model Uncertainty			
			Slope (°C/mV)	σ_{Slope} (°C/mV)	Bias (°C)	σ_{Bias} (°C)
A	T1	0.07	12.67	0.025	260.18	0.57
A	T2	0.07	12.86	0.026	259.31	0.57
B	T1	0.03	31.91	0.074	320.72	0.81
B	T2	0.03	32.37	0.071	322.25	0.76

Table 4.40: Uncertainty of the Temperature prediction by the LPCB version

Sensor	Temperature Range (°C)	Type A		Type B	
		Range (Ω)	σ_{T_P} (^oC)	Range (Ω)	σ_{T_P} (^oC)
T1	12.2 – 39.5	21.34 – 23.73	1.21	10.41 – 11.32	1.59
T2	12.2 – 39.5	20.97 – 23.37	1.22	10.29 – 11.21	1.54

For the conductivity sensor, the uncertainty equation using the GEP formulae for the model *(4.37)* is given by:

$$\sigma_{C_P} = \sqrt{\left(\frac{\partial C_P}{\partial \alpha}\sigma_\alpha\right)^2 + \left(\frac{\partial C_P}{\partial T_j}\sigma_{T_j}\right)^2 + \left(\frac{\partial C_P}{\partial \beta}\sigma_\beta\right)^2 + \left(\frac{\partial C_P}{\partial V_{Ci}}\sigma_{V_{Ci}}\right)^2 + \left(\frac{\partial C_P}{\partial a_1}\sigma_{a_1}\right)^2} \quad (4.38)$$

The corresponding uncertainties of each model parameter and sensor response estimated from the data are tabulated in Table 4.41.

Table 4.41: Uncertainties associated with the sensor response and model coefficients

Type – Temperature Sensor	Uncertainty in model coefficients			$\sigma_{V_{Ci}}$ (V)	σ_{Tj} (mV)
	σ_α (µS/cm)$^{0.5}$	σ_β (V(µS/cm)$^{0.5}$)	σ_{a_1} (V)		
Type B – T1	15.15	0.16	0.0005	0.008	0.03
Type B – T2	17.64	0.18	0.0005	0.008	0.03
Type A – T1	4.70	0.10	0.0003	0.006	0.07
Type A – T2	4.91	0.10	0.0003	0.006	0.07

From Table 4.41, using equation *(4.38)*, the uncertainty in conductivity at room temperature (T = 25 °C) is summarized in Table 4.42.

Table 4.42: Uncertainty in Conductivity sensing by sensor C1, for the range 230 – 2100 µS/cm at temperature T = 25°C

	σ_{C_P} (µS/cm)	$\frac{\sigma_{C_P}}{C_P}$
Type B – T1	349.4	0.24
Type B – T2	382.9	0.27
Type A – T1	189.6	0.14
Type A – T2	177.8	0.16

For Chlorine and pH estimation, the calibration model is based on optimizing an objective function *(4.35)* for a given set of sensor responses. Therefore, the bootstrap-based uncertainty estimation was performed instead of using the GEP formulae. The uncertainty associated with the coefficients of the chemiresistors is tabulated in Table 4.43. Using the tabulated information, the uncertainty for active chlorine and pH was estimated to be 0.24 mg/l and 0.25 pH units.

Table 4.43: Uncertainties associated with the coefficients of Active Chlorine and pH model using Phase 1 dataset

Coefficients →	pH (%/ pH)		HClO (%/ mg/l)		$HClO^2$ (%/ mg/l)^2		Bias
Chemiresistors ↓	$M_{0,i}$	$\sigma_{M_{0,i}}$	$M_{1,i}$	$\sigma_{M_{1,i}}$	$M_{2,i}$	$\sigma_{M_{2,i}}$	$\sigma_{M_{3,i}}$
R11	0.12	0.003	-2.08	0.02	1.92	0.04	0.02
R14	0.34	0.003	-1.85	0.02	1.59	0.03	0.02
R21	0.40	0.01	-1.87	0.05	1.94	0.08	0.06
R32	0.34	0.01	-0.93	0.02	0.75	0.03	0.02
R33	0.28	0.003	-2.29	0.04	2.24	0.06	0.05
R34	0.55	0.003	-1.76	0.03	1.73	0.06	0.03

The uncertainties shown in Table 4.43 were estimated for the model developed with Phase 1 dataset, and similar uncertainties in the model inputs were estimated for Phase 2 and Phase 3. The uncertainty in active chlorine and pH sensing throughout the 3 phases based on model input uncertainties is tabulated in Table 4.44.

Table 4.44: Uncertainties in Active chlorine and pH sensing throughout the Phase 1 (a, b), 2 and 3 of Sense-city experiments

	Phase 1a, 1b	Phase 2	Phase 3
pH	0.257	0.024	0.095
Active Chlorine (mg/l)	0.125	0.133	0.064

4.6 Summary of Performances

Initial experiments with the silicon version of the LOTUS sensor in the Sense-city's simulated WDN laid the groundwork for understanding its basic functionality and potential. This allowed for a comprehensive evaluation of the sensor's performance under varying conditions, revealing that temperature and conductivity sensitivities presented challenges due to thermal sensitivity at wire bonding connections. Notably, the predictive performance for pH and active chlorine, although meeting the requirements for LOTUS applications, highlighted areas for real-time and field applicability improvement due to the lack of temporal structuring in model training.

Further exploration involved the development of the plastic version of the LOTUS sensor to address the silicon version's limitations and improve its affordability. The laboratory conditions confirmed the practicality of sensors printed on plastic substrates through AC and alternating DC excitation studies. This validation moved forward into the real-world testing phase at Sense-City, emphasizing the potential of data fusion techniques to refine predictions for chlorine and pH levels. The exploration of these techniques underlined the importance of adapting model strategies to better align with rapidly changing environmental conditions, pointing to the need for further model refinement and questioning the accuracy of reference sensors during swift chlorine fluctuations.

The development of a field-ready version was made possible by the several experiments done with single-electrode to 6 electrodes sensor head. The deployment of the field-ready version, characterized by a comprehensive dataset gathered over five months at Sense-City, provided valuable insights into the sensor's long-term performance and reliability. The introduction of an automated chlorine pump partway through the study suggested the necessity of model adjustments. At the same time, observations on device durability and performance inconsistency emphasized the critical need for ongoing sensor evaluation and adaptation in response to environmental changes.

Through various studies and experiments, calibration models have been refined for predicting active chlorine and pH and estimating conductivity and temperature, leading to enhanced sensor performance despite noise challenges. While achieving target performance levels, the silicon version of the sensor faced limitations in real-time applications due to insufficient data.

In contrast, the field-ready version, although not yet fully deployed in the field, shows promising signs of becoming a viable solution shortly. The chlorine prediction by the chemiresistors were on par with those reported in literature, while the longevity of the sensor (with 130 days of active performance) is about 4 times higher than reports in the state of the art. However, these conclusions are drawn from a limited number of devices with fabrication complications, and therefore, they lack consistency, necessitating further validation through experiments with a broader array of sensors to confirm performance consistency in real-world settings. Furthermore, there is a recognized need for ongoing refinement of sensor models to improve their accuracy and reliability amid environmental and operational variabilities.

5

Calibration Transfer

Abstract: *This chapter discusses the need for calibration transfer, examines select calibration transfer techniques widely used in literature with 3 gas sensor datasets, and provides insights on the applicability of these techniques. This chapter also highlights the calibration transfer techniques that can be applied to LOTUS sensors.*

5 Calibration Transfer

One of the challenges of deploying environmental sensor solutions is ensuring the measurements' accuracy and reliability across different sensor units and environments. For instance, when producing a new multi-parametric water quality sensor in large quantities, each sensor unit may have slightly different characteristics due to manufacturing variations and sensitivity changes [87], [88], [90]. Moreover, the water samples to be measured may have different compositions, temperatures, pH levels, or salinities, depending on the location and time of sampling. These factors can affect the sensor response and introduce measurement errors or biases. To overcome these issues, calibration transfer techniques can be used to adjust the sensor responses and make them consistent and comparable with a reference sensor unit or a standard calibration model. Calibration transfer reduces the effort in calibrating each sensor unit as per the original calibration protocol, which may require expensive reference instruments, chemicals, or skilled operators. Calibration transfer also enables the use of the same calibration model across different sensor units and environments, thereby improving the efficiency and scalability of the sensor deployments. Calibration transfer can also enhance the long-term stability and performance of the sensors by compensating for drift or degradation over time.

Calibration transfer techniques were developed in the chemometrics literature [192] to ease the transfer of calibration models from unit to unit and from environment to environment. These techniques were primarily designed for spectral datasets wherein one spectroscopic instrument is standardized with another to enable reusing the calibration model [230], [231]. The general applicability of these techniques to environmental sensors and sensor arrays has been little discussed in the literature. Articles often show the successful application of a specific method on a specific sensor, for instance, electronic noses/tongues [232], [233], gas sensor arrays [234], and potentiometric sensors [235], [236], without tackling the portability of the technique to different sensor scenarios. Transfer methodologies follow one of three approaches: model updating, signal processing-based transfer, and standardization. In model updating-based techniques, the calibration model is re-trained by adding samples or data obtained from the secondary sensor [204], [205], [236]. In signal processing-based techniques, the signals from primary and secondary sensors are processed before building the calibration model to remove sensor-specific and environment-dependent variations. Pre-

processing techniques vary from simple mathematical operations like baseline correction or sample normalization to advanced operations such as orthogonal signal correction (OSC) [208] or wavelet-based pre-processing [211]. Standardization techniques consist of modifying the response of the secondary sensor to match the features of the primary sensor response [10]. This modification is realized by either updating the model coefficients, the sensor responses, or the model predictions by mathematical operations.

5.1 Background Information

5.1.1 Definitions

"Sensor" refers to a multisensor system or sensor array, while a "sensing unit" denotes individual devices reacting to specific analytes. The term calibration function refers to the mathematical relationship that links the outputs of each sensing unit in the sensor to the target measurand (e.g., gas concentration, temperature). The primary sensor refers to a sensor for which an accurate calibration model is available thanks to extensive experiments. Calibration transfers aim to adapt the primary calibration model to a secondary sensor, for which usually less data is available. Most calibration transfer techniques require transfer standards: they are the dataset constituted by the outputs of the primary and secondary sensors measuring the same target value of the measurand.

The data from the primary and secondary sensors is respectively denoted by X_P and X_S. The corresponding transfer standards are denoted by T_P, T_S. The target response for primary and secondary sensors is denoted by Y_P, Y_S. $\widehat{X_S}$ refers to the corrected secondary sensor response, i.e., the secondary sensor response transferred to the primary sensor response domain using standardization techniques. K denotes the number of individual sensing units, with i being the index variable and f_P denotes the primary calibration model.

5.1.2 Description of compared standardization methods

Table 5.1 outlines 10 calibration transfer techniques analyzed, focusing on methods that do not rely on sensor-specific features, unlike single wavelength standardization [235], [237], model expansion with sample weighting, Tikhonov regularization, joint Y-partial least squares regression [236], or transfer sample-based coupled task learning) [233]. Transfer techniques that exploit the spectral nature of the data, like Spectral space transformation

[212] or the Variable Penalty dynamic time warping method [213], are not considered either. Deep Learning-based calibration transfer are not addressed as they require high data volume.

Table 5.1: Selected Calibration Transfer Techniques

No.	Technique	Reference
1	Direct Standardization (DS)	[230], [234], [238]
2, 3	Single Sensor Standardization (SSS) + Regression (LR, SVR)	[235], [239], [240]
4	PLS-based standardization (PLS + DS)	[241], [242], [243]
5	PC-based standardization (PC + DS)	[242]
6	Procrustes transform (PT)	[244]
7	Procrustes transform (PT) with PC factors (PC + PT)	[245], [246]
8	Slope Bias Correction (SBC)	[236], [247]
9, 10	Mean correction, additive (AMC) and multiplicative (MMC)	[205], [248]

5.1.2.1 Direct standardization

This method builds a linear mapping between the primary and secondary sensor domain by using the transfer standards of primary and secondary sensor responses. The transformation matrix is defined as $F_{DS} = T_S^+ T_P$, where T_S^+ is the Moore-Penrose inverse of T_S [230], [234], [238]. The secondary sensor response is standardized to the primary sensor response by applying the formula $\widehat{X_S} = X_S F_{DS}$. This corrected secondary sensor response $\widehat{X_S}$ can be used with the primary calibration model to get the target:

$$Y_S = f_P(\widehat{X_S}) = f_P(X_S F_{DS}) \tag{5.1}$$

5.1.2.2 Single sensor standardization

Single sensor standardization (SSS) is a transfer technique proposed by Khaydukova et al. [235] based on single wavelength standardization as developed by Nørgaard et al. [239]. Each wavelength in the secondary instrument is regressed against the corresponding primary instrument, thereby creating the mapping function. Khaydukova replaced the wavelength

with the response from the sensing units in the sensor. Each sensing unit in the secondary sensor is mapped to the corresponding sensing unit in the primary sensor. This mapping can be done via linear regression (SSS + LR) or Support Vector Regression (SSS + SVR).

The parameters required for SSS + LR are estimated from the transfer standards by:

$$t_p^i = m_i t_s^i + c_i \quad \dots \quad \forall\, i \in K \tag{5.2}$$

Then, the corrected response is estimated as:

$$\widehat{x_s^i} = m_i x_s^i + c_i \quad \dots \quad \forall\, i \in K \tag{5.3}$$

For SSS + SVR, the SVR function from the scikit-learn (1.4.0) library in Python is utilized.

5.1.2.3 PLS-based and PC-based standardization

Partial least squares (PLS) [249] and principal component (PC) [250] project a dataset of input/output variables into a hyperplane maximizing correlation (for PLS) or variance (for PC) between variables. In PLS-based (PLS + DS) and PC-based standardization (PC + DS), the PLS or PC transformation matrix is calculated for the transfer standards dataset, and then the direct standardization method is applied to the PLS- or PC-transformed sensor's data [241], [242], [243]. Assuming an operator $h(.)$ computes the PLS or PC transformation, the mapping is obtained as:

$$h(T_S)\, F_{PLS} = h(T_P) \tag{5.4}$$

So, the corrected response is then estimated via:

$$\widehat{X_S} = h^{-1}(h(X_S) F_{PLS}) \tag{5.5}$$

5.1.2.4 Procrustes transforms

Procrustes analysis [244] is a statistical technique that performs mathematical operations like translation, rotation, scaling, and reflection to match different shapes/geometry. Andrade et al. [245] proposed mapping other spectral instruments by applying the Procrustes transform to the PC scores (PC + PT) of the dataset, and Sergeant et al. [42] used it to water quality data from different data sources to identify regions with similar water quality. To

apply Procrustes transform, the MATLAB function Procrustes was ported to Python as detailed in [251].

5.1.2.5 Slope-Bias Correction

The slope-bias correction technique estimates a univariate mapping between the primary calibration model prediction with transfer standards from primary and secondary sensors [236], [247]. For the case of univariate regression, given the primary and secondary transfer standard T_P, T_S and primary calibration model f_P, the mapping is defined as:

$$f_P(T_P) = \alpha\, f_P(T_S) + \beta \tag{5.6}$$

This mapping is then applied to the primary calibration model prediction with secondary sensor response to get the corrected response mathematically:

$$Y_S = \alpha f_P(X_S) + \beta \tag{5.7}$$

5.1.2.6 Mean Correction

The mean correction technique aims at aligning the probability distributions of the primary and secondary sensor datasets. Two simple realizations of the methodology for matching probability distributions are Additive Mean Correction (AMC) and Multiplicative Mean Correction (MMC) [205], [248]. In additive correction, the mean of the secondary response dataset, $\overline{X_S}$ is subtracted from the secondary dataset, X_S and the mean of primary dataset, $\overline{X_P}$ is added:

$$\widehat{X_S} = X_S - \overline{X_S} + \overline{X_P} \tag{5.8}$$

Similarly, in a multiplicative correction, the mean of the secondary response dataset, $\overline{X_S}$ is divided from the secondary dataset, X_S and the mean of the primary dataset, $\overline{X_P}$ is multiplied. The operations are performed elementwise for multidimensional data:

$$\widehat{x_s^l} = x_s^i * \left(\frac{\overline{x_p^l}}{\overline{x_s^l}}\right) \ldots \ \forall\, i \in K \tag{5.9}$$

where $\overline{x_p^l}$ is the mean of the ith sensing unit in the primary sensor response dataset, and $\overline{x_s^l}$ is the mean of the ith sensing unit in the secondary sensor response dataset. This corrected dataset can then be directly used with the primary calibration model.

5.1.3 Implementation of compared standardization methods

The following notations are used: let X_P denote the response from primary sensor and X_S denote the response from secondary sensor. Their corresponding target variables, are denoted by Y_P and Y_S. Transfer standards corresponds to the same target variables Y, for which the both the primary and secondary sensor have response. They are respectively denoted T_P and T_S. The calibration model developed by utilizing the primary sensor response and primary sensor target, termed as primary calibration model, is represented as f_p. The secondary sensor response (which is not compatible with the primary calibration model) will be corrected using the calibration transfer techniques, to give the corrected secondary sensor response denoted as $\widehat{X_S}$. The notation $A \times B$ represent matrix multiplication, and $A \,.\, B$ represent scalar/element wise multiplication.

All the code was implemented in Python version 3.9.0 with libraries numpy (1.22.1), pandas (1.4.0), scikit-learn (1.4.0), and scipy (1.7.3).

5.1.3.1 Direct Standardization (DS):

Here a simple mapping is developed between the primary and secondary sensor response using the transfer standards.

1. Given: (X_P, Y_P), (X_S, Y_S), (T_P, T_S, Y)
2. Calculate: $F_{DS} = Inverse\ (T_S) \times T_P$
3. For new response, Compute: $\widehat{X_S} = X_S \times F_{DS}$
4. Target output: $f_P(\widehat{X_S})$

The $Inverse$ () function implemented in numpy library of python was utilized here (which calculates the inverse of a matrix using its singular-value decomposition).

5.1.3.2 Single Sensor Standardization (SSS):

If the sensor (primary or secondary) response matrix consists of K columns, where each column represents the response from each individual sensing unit within the sensor, then the matrix X can be denoted as $X = [x_1, x_2, \dots x_i \dots x_K]$.

In this Single sensor standardization technique, each response of each sensing unit from the primary sensor is mapped to the response of the corresponding unit in the secondary sensor. The mapping can be any function, but here we considered only Linear Regression (LR), and Support Vector Regression (SVR).

If we denote this mapping as $g(.)$, then the pseudocode for this technique with $g(.)$ as linear regression will be as follows:

1. Given: (X_P, Y_P), (X_S, Y_S), (T_P, T_S, Y)
2. Where, $X_P = [x_{p1}, x_{p2}, \dots x_{pi} \dots x_{pk}]$, similarly $T_P = [t_{p1}, t_{p2} \dots t_{pi} \dots t_{pk}]$
3. For every sensing unit i,

   Calculate: $t_{pi} = m_i \,.\, t_{si} + c_i$
4. Generate matrix: $M = \begin{bmatrix} m_1 & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & m_k \end{bmatrix}, C = [c_1 \quad \dots \quad c_k]$
5. For new response, Compute: $\widehat{X_S} = X_S \times M + C$
6. Target output: $f_P(\widehat{X_S})$

For the case where SVR is used as the mapping function for $g(.)$, only steps 3-5 will change as follows:

3. For every sensing unit i,

   Calculate: $t_{pi} = g_i(t_{si})$
4. Generate matrix: $G = \begin{bmatrix} g_1(.) & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & g_k(.) \end{bmatrix}$
5. For new response, Compute: $\widehat{X_S} = G \times X_S$

For SSS + SVR, the SVR function from the scikit-learn library in python was utilized.

5.1.3.3 PLS-based (PLS+DS) and PC-based Standardization (PC+DS):

Here, let the function $h(.)$ denote the function that generates the latent variables using the Partial Least squares (PLS) algorithm. Similarly let the function $q(.)$ denote the function that generates the PC scores of the given data using the principal component analysis (PCA) algorithm. The inverse functions represented as $h^{-1}(.)$ and $q^{-1}(.)$ are used to compute the actual data from the latent variables and PC scores respectively. In this technique, for the case of PLS+DS, the standardization is performed between the latent variables obtained from the PLS algorithm for both primary and secondary sensor data corresponding to transfer standards. For PC+DS, the standardization is performed between the PC scores of primary and secondary sensor response data corresponding to transfer standards. This standardization is then utilized to correct the secondary sensor response. The pseudocode for PLS+DS:

1. Given: (X_P, Y_P), (X_S, Y_S), (T_P, T_S, Y)
2. Calculate latent variables: $L_{TP} = h_P(T_P, Y)$ and $L_{TS} = h_S(T_S, Y)$
3. Calculate: $F_{PLS} = Inverse\ (L_{TS}) \times L_{TP}$
4. For new response:
    - Transform to latent variables: $L_{XS} = h_S(X_S)$
    - Standardize the latent variables: $\widehat{L_{XS}} = L_{XS} \times F_{PLS}$
    - Transform to sensor response domain: $\widehat{X_S} = h_p^{-1}(\widehat{L_{XS}})$
5. Target output: $f_P(\widehat{X_S})$

In the case of PC + DS,

1. Given: (X_P, Y_P), (X_S, Y_S), (T_P, T_S, Y)
2. Calculate PC scores: $P_{TP} = q_P(T_P)$ and $P_{TS} = q_S(T_S)$
3. Calculate: $F_{PC} = Inverse\ (P_{TS}) \times P_{TP}$
4. For new response:
    - Transform to PC scores: $P_{XS} = q_S(X_S)$
    - Standardize the PC scores: $\widehat{P_{XS}} = P_{XS} \times F_{PC}$
    - Transform to sensor response domain: $\widehat{X_S} = q_p^{-1}(\widehat{P_{XS}})$
5. Target output: $f_P(\widehat{X_S})$

The functions $h(.)$ and $h^{-1}(.)$ were utilized from the PLSRegression method and the functions $q(.)$ and $q^{-1}(.)$ were utilized from the PCA method in scikit-learn library.

5.1.3.4 Procrustes Transform (PT):

This technique can be either applied directly or applied to the PC scores. The inbuild function from MATLAB is ported to python and then utilized for this technique application. Procrustes transform between 2 matrices, provides a set of parameters: translation parameter (τ), scaling parameter (λ) and rotational parameter (R) which minimizes the difference between the 2 matrices; mathematically:

$$PT(A, B) \approx \tau, \lambda, R$$

Such that, the difference between matrix B and resultant matrix $\hat{A} = (A - \tau).\lambda \times R$ is minimum. The pseudocode for the calibration transfer technique employing Procrustes transform follows:

1. Given: (X_P, Y_P), (X_S, Y_S), (T_P, T_S, Y)
2. Calculate PT parameters of the transfer standards: $\tau, \lambda,\ R = PT(T_S, T_P)$
3. For new response, $\widehat{X_S} = (X_S - \tau).\lambda \times R$
4. Target output: $f_P(\widehat{X_S})$

For the PC+PT technique, the PT parameters are estimated for the PC scores instead:

1. Given: (X_P, Y_P), (X_S, Y_S), (T_P, T_S, Y)
2. Calculate PC scores: $P_{TP} = q_P(T_P)$ and $P_{TS} = q_S(T_S)$
3. Calculate PT parameters of the transfer standards: $\tau, \lambda,\ R = PT(P_{TS}, P_{TP})$
4. For new response:
    - Transform to PC scores: $P_{XS} = q_S(X_S)$
    - Procrustes transform inversion: $\widehat{P_{XS}} = (P_{XS} - \tau).\lambda \times R$
    - Transform to sensor response domain: $\widehat{X_S} = q_p^{-1}(\widehat{P_{XS}})$
5. Target output: $f_P(\widehat{X_S})$

5.1.3.5 Slope-Bias correction (SBC):

In this technique the secondary sensor response is used as is with the primary calibration model and the output is the corrected to get the target output.

1. Given: (X_P, Y_P), (X_S, Y_S), (T_P, T_S, Y)
2. Estimate: α, β, such that $f_P(T_P) = \alpha . f_P(T_S) + \beta$
3. For new response, Target output: $\alpha . f_P(X_S) + \beta$

5.1.3.6 Mean correction – Additive (AMC) and Multiplicative (MMC):

This technique aims to match the mean of both primary and secondary sensor resposnse via addition or multiplication. Assume, $\overline{x_{pi}}$ is the mean of the response of the ith sensing unit in the primary sensor and $\overline{x_{si}}$ is the mean of the response of the ith sensing unit in the secondary sensor, the pseudo code for both techniques is given below.

For AMC:

1. Given: (X_P, Y_P), (X_S, Y_S), (T_P, T_S, Y)
2. Where, $X_P = [x_{p1}, x_{p2}, \dots x_{pi} \dots x_{pk}]$, similarly $T_P = [t_{p1}, t_{p2} \dots t_{pi} \dots t_{pk}]$
3. For every sensing unit i,

   Calculate: $\widehat{x_{si}} = x_{si} - \overline{x_{si}} + \overline{x_{pi}}$
4. Generate: $\widehat{X_S} = [\widehat{x_{s1}}, \widehat{x_{s2}}, \dots \widehat{x_{si}} \dots \widehat{x_{sk}}]$
5. Target output: $f_P(\widehat{X_S})$

Similarly, for MMC:

1. Given: (X_P, Y_P), (X_S, Y_S), (T_P, T_S, Y)
2. Where, $X_P = [x_{p1}, x_{p2}, \dots x_{pi} \dots x_{pk}]$, similarly $T_P = [t_{p1}, t_{p2} \dots t_{pi} \dots t_{pk}]$
3. For every sensing unit i,

   Calculate: $\widehat{x_{si}} = x_{si} . (\overline{x_{pi}}/\overline{x_{si}})$
4. Generate: $\widehat{X_S} = [\widehat{x_{s1}}, \widehat{x_{s2}}, \dots \widehat{x_{si}} \dots \widehat{x_{sk}}]$
5. Target output: $f_P(\widehat{X_S})$

5.2 Dataset description

5.2.1 Dataset 1 – Multiple sensors of the same make

The first case study uses the dataset generated by Fonollosa et al. [234], [252] concerning gas concentration measurement with metal-oxide (MOX) based sensor arrays. The dataset comprises 5 stand-alone sensors (B1 to B5), each equipped with 8 individual gas sensing units, for classifying 4 different gases (ethanol, ethylene, methane, and carbon monoxide) at various concentration levels over 22 days. Each step includes the sensors' steady-state response. This dataset being small, its size was increased by using the Clifford-Tuma model [253], [254]:

$$\log\left(\frac{R_0 - R_s}{R_0}\right) = \log(s) + \beta \log(c) \tag{5.10}$$

where R_s corresponds to the response of a MOX sensor at concentration c and R_0 corresponds to the response in clean air (at 0 concentration, baseline), s, β refers to sensor-dependent parameters. These parameters were estimated for each MOX sensing unit by bootstrap curve fitting, and then the responses were simulated for 9 additional concentration values (midpoints of the existing data) (More details in Appendix A). The primary and secondary sensors were chosen arbitrarily as B1 and B5, respectively, with carbon monoxide results reported due to similar outcomes across gases.

5.2.2 Dataset 2 – Time drift dataset

This dataset examines the long-term drift of a single MOX gas sensor array with 16 sensing units [255], [256], [257]. The sensor was exposed to six gases (ammonia, acetaldehyde, acetone, ethylene, ethanol, and toluene) at varying concentrations over 36 months, with continuous experiments conducted from the first to the 30th month and a 5-month inactive period before the final data collection in the 36th month. The drift of sensing units ranged from -6% to 17% per month, normalized by their baseline resistance. We focus on the baseline-normalized steady-state response to ammonia, evaluating sensor calibration transfer from month 1 to month 36, with primary sensor data comprising 90 points and secondary sensor data comprising 600 points.

5.2.3 Dataset 3 –sensors of the same make in different environments

This case study utilizes data from two prototype Radio Frequency (RF)-based humidity sensors (S1 and S2) [258]. While sensor S1 underwent controlled laboratory calibration experiments, sensor S2 was exposed to real-world field conditions. RF-based humidity sensors operate based on resonance frequency modulation induced by changes in the dielectric constant of the humidity-sensitive polymer layer. The sensor response, characterized by the differential magnitude and frequency, elucidates alterations in resonance characteristics influenced by moisture presence and concentration.

Calibration experiments for sensor S1 entailed controlled environmental settings to ascertain the relationship between sensor output and humidity levels across a broad temperature spectrum. In contrast, sensor S2 was deployed outdoors alongside a reference weather station, Vaisala WXT 520 [221], [259], within the Sense-City platform for a two-day experiment. Sensor S1 was subjected to varying humidity levels (25% to 70% RH) and temperatures (25°C – 45°C) over 44 hours. This setup facilitated the examination of sensor performance under fluctuating environmental conditions, encompassing temperature variations, ambient humidity levels, and exposure to diverse weather conditions.

5.3 Methods

5.3.1 Performance assessment

The assessment process is as follows: the primary calibration model was first developed with the primary sensor calibration data. This model was then applied to the secondary sensor data without correction to provide the baseline performance to improve upon using calibration transfer. Transfer standards (TS) were selected for each case study as detailed in section IV.C. The secondary sensor was also fully recalibrated with just the TS for comparison.

Only two representative error metrics were selected to quantify performance: Mean Absolute Error (MAE) and Coefficient of determination (R^2). MSE (Mean Squared Error), RMSE (Root Mean Squared Error), or MAPE (Mean Absolute Percentage Error) were calculated but are not provided here for ease of understanding, as they yield similar conclusions. Unless otherwise specified, the provided metrics are the mean and standard deviation (displayed within parenthesis after the mean value) of MAE and R^2 over Q splits of the dataset (random train/test splits, random selection of transfer standards). Q is determined based on the dataset size.

In each case, the MAE of the naïve model is provided: by definition, this model has $R^2=0$, which is achieved by returning a constant value equal to the mean of the dataset. The better the metrics are compared to the naïve model (lower MAE, higher R^2), the better the performance of the method. In cases where R^2 is negative (the model performance is worse than the naïve model), the standard deviation is not reported, and the R^2 values are capped at -1.

The ideal calibration transfer method provides metrics as good as or better than the primary calibration model applied to the primary sensor. To be relevant, a calibration transfer method should also have comparable or better metrics than recalibration with the transfer standards.

For the hyperparameters in each calibration transfer technique, we used the default parameters provided by the library unless otherwise specified. While these defaults may not be the optimal choice for the specific dataset, this approach establishes a baseline performance across all cases, enabling a preliminary comparison.

5.3.2 Primary calibration model

Linear regression (LR) and K-nearest neighbours regression (KNN) were selected as model-based and model-free calibration methods for all three case studies, as these two methods are frequently used in sensor array calibration and perform well. Naturally, linear regression does not perform optimally in the case of non-linear sensors. This is the case of the third dataset, where the physical model is exponential. However, it was elected to still use the linear model in the third case study for the sake of comparison between datasets. It also allows to evaluate calibration transfer methods in the case of an imperfect primary calibration model. The linear regression and KNN models were developed in the Python scikit-learn (1.4.0) library using the LinearRegression function and the KNeighborsRegressor function. The hyperparameter 'k' for the KNN regression was estimated by the elbow method with a 60%-20%-20% split of the total available data points in the primary sensor response.

5.3.3 Transfer standard selection

Case study 1: The primary and secondary sensors have the same 19 gas concentration values. As a result, any point in the dataset can be used as a transfer standard. Transfer standards from 3 to 11 (from 20% to 60% of the TS dataset) were thus randomly selected out of the dataset. For each calibration transfer technique and each TS ratio, 20 random data splits were generated.

Case study 2: Only 6 common concentration levels (10, 50, 100, 150, 200, 250 ppm) are found in the primary and secondary datasets for this case study. An increasing number (from 3 to 6) of transfer standards was randomly selected, using all the data available for each concentration level (6 points and 100 points for each concentration for the primary and secondary dataset).

Case study 3: Here, the primary sensor response has 3500 data points, and the secondary sensor response has 1400 data points, with 290 target values (Relative humidity values) common between the two. The number of transfer standards used increased from 14 to 174 points (5% to 60% ratio).

5.4 Application to the case studies

5.4.1 Case Study I

In this case study, the primary and secondary sensors are different units of the same build measuring the concentration of CO in the air. Table 5.2 shows the performance of the different calibration transfer techniques with 20% of transfer standards (Results with more TS are provided in Appendix B). Figure 5.1 shows the MAE as a function of the number of transfer standards for the methods with non-negative R^2. The naïve model has 59ppm MAE. With LR, only DS, PC+DS, and PLS+DS yield non-negative R^2 with performances comparable with recalibration with TS. With 60% standards, their MAE is similar to the primary sensors. By comparison, KNN performs worse overall than LR (even for the primary sensor) because of the limited size of the dataset. Acceptable transfer methods ($R^2>0$ and comparable or better than recalibration with TS) are DS and PLS+DS, which perform ~10 times better than recalibration with TS. From 40% of TS upward, the MAE becomes lower than the primary sensor MAE. For this case study, DS and PLS+DS appear to be the most robust methods, working well in all the different configurations.

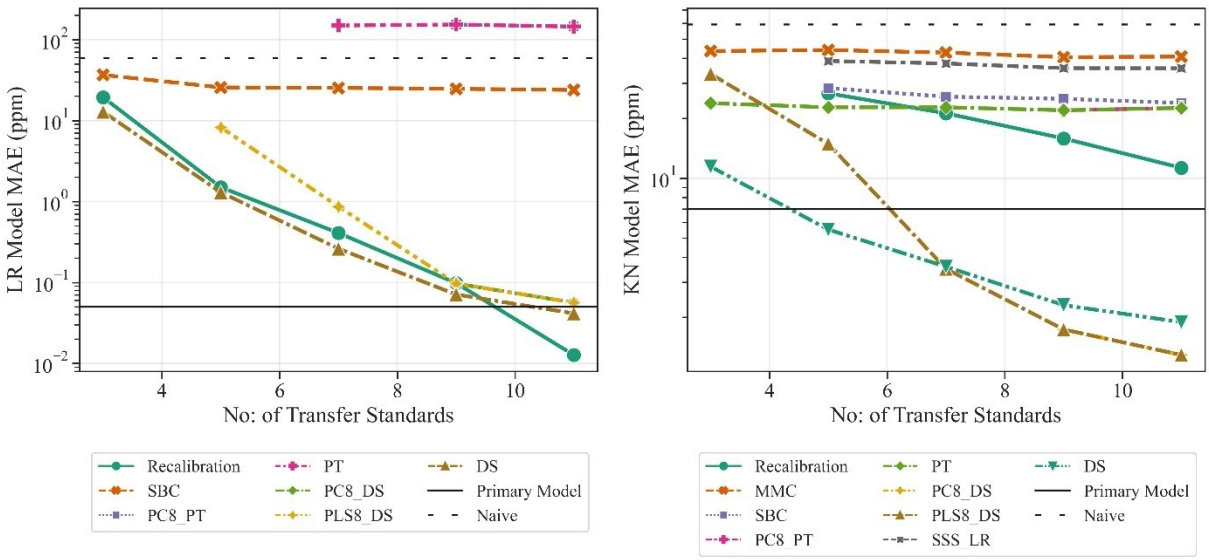

Figure 5.1: Case study I: Average MAE of the different calibration transfer methods achieved with either LR or KNN as primary calibration as a function of the number of transfer standards.

Table 5.2: Results of calibration transfer for the case study I with only 20% (3 points) of the transfer standards.

	LR		KNN	
	MAE (ppm)	R2	MAE (ppm)	R2
Primary sensor with primary calibration	0.05 (0.03)	1.00	7.0 (1.8)	0.96 (0.0)
Secondary sensor with primary calibration	3348	-1	103	-1
DS	13 (9.2)	0.9 (0.2)	11 (6.4)	0.9 (0.1)
SSS + LR	1193	-0.6	42	-0.4
SSS + SVR	93	-1	57	-1
PLS (4) + DS	43	-1	33 (4.2)	0.5 (0.4)
PLS (8) + DS	43	-1	33 (4.2)	0.5 (0.4)
PC (4) + DS	43.3	-1	33 (4.2)	0.5 (0.4)
PC (8) + DS	43.9	-1	35	-1
PT	3028	-1	24 (4.8)	0.7 (0.3)
PC (8) + PT	141	0	24 (4.8)	0.7 (0.3)
SBC	37 (18)	0.3 (1.2)	34	-1
AMC	995	-1	61	-1
MMC	1013	-1	44	0
Recalibration with TS	19 (15)	0.7 (0.7)	43	-1
Full Recalibration	0.01 (0.01)	1.00	9.9 (2.2)	0.90 (0.03)

5.4.2 Case Study II

Here, the calibration transfer is applied on a single sensor to compensate for its time drift. The data from month 1 is primary, and the data from month 36 is secondary. Table 5.3 shows the performance of the different calibration transfer methods for 3 transfer standards (other TS ratios are provided in Appendix B). Figure 5.2 shows the evolution of performances as a function of the number of transfer standards, with methods having $R^2>0$. The naïve model has 78ppm MAE.

Here, for the case of the LR model, the performance of PLS+DS and PC+DS closely follows the recalibration performance, reaching primary calibration performance with 6 concentration levels. DS works slightly less well but reaches comparable performance at 6 concentration levels. It appears that PLS and PC play their expected role of reducing dataset dimensionality, compared to DS. In the case of KNN regression, recalibration has negative R2 for 3 distinct concentration levels, but it remains the optimal technique for larger concentration levels, followed by PLS+DS and PC+DS, then DS. With KNN, no technique reaches the performance of primary calibration due to the limited number of transfer standards.

Table 5.3: Results of calibration transfer for the case study II with only 3 concentration levels as transfer standards.

	LR		KNN	
	MAE (ppm)	R2	MAE (ppm)	R2
Primary sensor with primary calibration	12 (4)	0.9 (0.1)	6.3 (1.4)	0.9 (0.04)
Secondary sensor with primary calibration	185	-6.3	65	-0.23
DS	56 (26)	0.4 (0.3)	43(18)	0.4 (0.4)
SSS + LR	65	-1	45	-1
SSS + SVR	112	-1	44	-1
PLS (8) + DS	33 (12)	0.6 (0.4)	32 (13)	0.4 (0.8)
PLS (16) + DS	60 (32)	0.4 (0.3)	44 (20)	0.4 (0.5)
PC (8) + DS	31 (13)	0.6 (0.4)	31 (15)	0.5 (0.9)
PC (16) + DS	60 (32)	0.4 (0.3)	44 (20)	0.4 (0.5)
PT	88	-1	58	-1
PC (8) + PT	78	-1	58	-1
PC (16) + PT	88	-1	58	-1
SBC	46	-1	63	-1
AMC	126 (7.8)	0.4 (0.1)	53.6 (2.5)	0.2 (0.1)

MMC	108 (32)	0.2 (0.2)	55.5 (3.9)	0.4 (0.1)
Recalibration with TS	36 (16)	0.6 (0.3)	37	-1
Full Recalibration	11.1 (0.4)	0.96 (0.0)	2.9 (0.3)	0.98 (0.0)

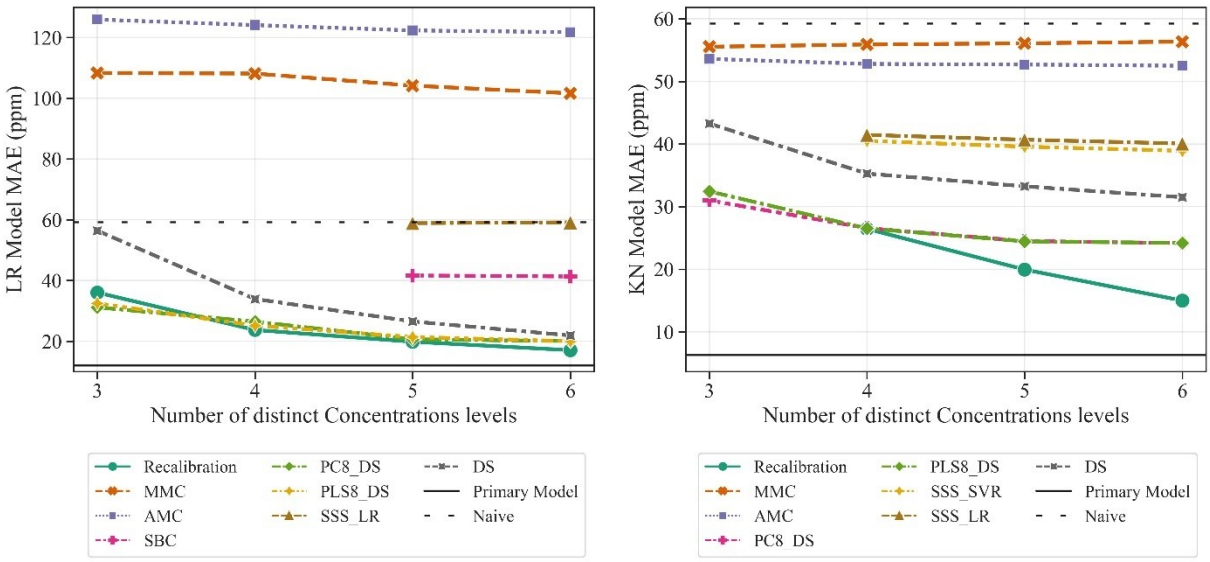

Figure 5.2: Case study II: Average MAE of the different calibration transfer methods achieved with either LR or KNN as primary calibration as a function of the number of transfer standards.

5.4.3 Case Study III

In this dataset, the secondary sensor differs from the primary sensor, and the calibration and test environments exhibit variations. Table 5.4 shows the performance of the different calibration transfer methods for the 5% TS ratio (other TS ratios provided in Appendix B), with the naïve model having 13%RH MAE, and Figure 5.3 shows the evolution of performances as a function of the number of transfer standards.

Overall, for both KNN and LR, a larger number of methods (than in the other case studies) perform well (e.g., with R^2>0.7 and with improved performance compared to calibration with the primary model). This is explained by the larger number of TS compared to the other studies. On the other hand, all methods (except KNN-based recalibration with TS) saturate very quickly in performance with the increasing number of TS (from 15 to 50), without reaching the performances of the primary model on the primary sensor, which is attributed to the intrinsic differences in sensors and environments. Given the exponential nature of the physical model, LR notably underperforms compared to KNN across all methods and quantities of transfer standards utilized.

Table 5.4: Results of calibration transfer for the case study III with only 5% (14 points) of the transfer standards.

	LR		KNN	
	MAE (%RH)	R2	MAE (%RH)	R2
Primary sensor with primary calibration	2.4 (0.01)	0.9 (0.0)	0.5 (0.01)	0.99
Secondary sensor with primary calibration	13	-0.04	13	-0.1
DS	7.7 (1.2)	0.7 (0.1)	6.3 (0.6)	0.2 (0.3)
SSS+ LR	8.2 (1.2)	0.7 (0.1)	5.7 (0.5)	0.4 (0.2)
SSS + SVR	8.1 (0.7)	0.7 (0.03)	5.4 (0.4)	0.6 (0.2)
PLS (2) + DS	8.9 (1.1)	0.6 (0.04)	6.1 (0.5)	0.3 (0.3)
PC (2) + DS	682	-1	29	-1
PT	8.5 (2.7)	0.7 (0.1)	5.8 (0.4)	0.5 (0.1)
PC (2) + PT	8.5 (2.7)	0.7 (0.1)	5.8 (0.4)	0.5 (0.1)
SBC	8.4 (1.2)	0.7 (0.1)	5.6 (0.7)	0.6 (0.5)
AMC	8.8	-1	10.2	-1
MMC	7.4 (1.1)	0.7 (0.1)	6.4 (0.2)	0.2 (0.03)
Recalibration with TS	7.0 (1.3)	0.7 (0.1)	6.4 (0.9)	0.3 (0.4)
Full Recalibration	5.2 (0.0)	0.8 (0.0)	1.6 (0.1)	0.9 (0.0)

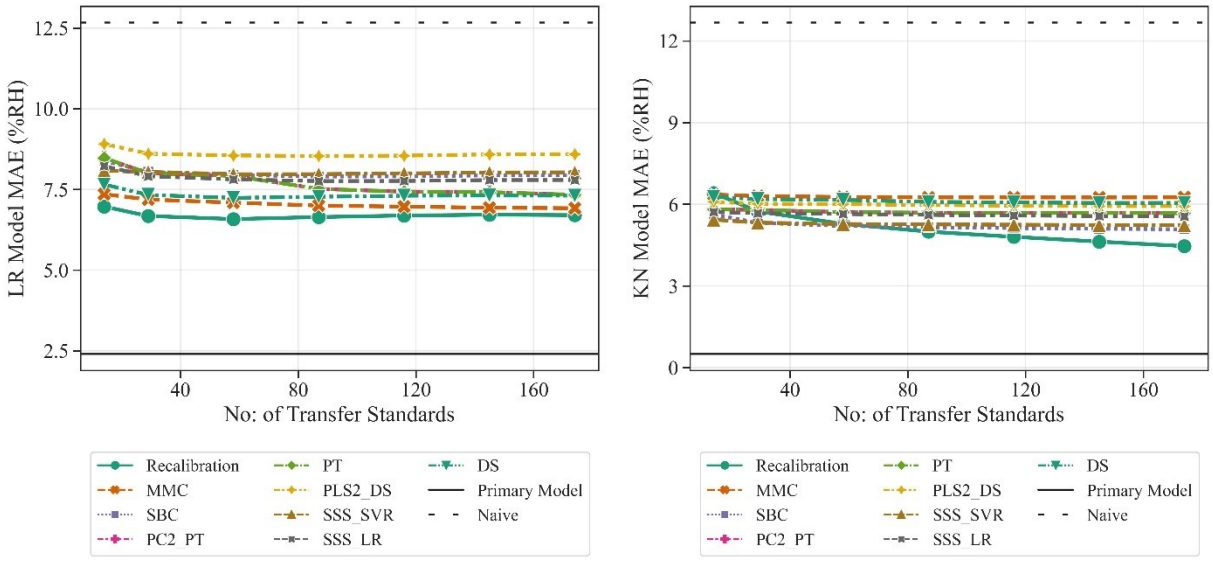

Figure 5.3: Case study III: Average MAE of the different calibration transfer methods achieved with either LR or KNN as primary calibration as a function of the number of transfer standards.

5.5 Results and Discussion

The results are synthesized in Table 5.5, classifying the methods by MAE achieved compared to naïve MAE, to recalibration with TS, and primary calibration. For generality, a 30% margin around the MAE values achieved with primary calibration or recalibration with TS is considered. Notation: '--': $R^2<0$ or MAE>Naïve MAE; '-': $R^2>0$ and MAE<Naïve MAE and MAE > 1.3*Recalibration MAE; '+' $R^2>0$ and MAE≤ 1.3*Recalibration MAE; '++' : $R^2>0$ and MAE≤ 1.3*Primary sensor MAE;'*': MAE≤ 0.7*Recalibration MAE.

Table 5.5: Performance classification of methods by MAE achieved compared to naïve MAE, to recalibration with TS, and primary calibration

	Case Study 1		Case Study 2		Case Study 3	
	LR	KNN	LR	KNN	LR	KNN
Secondary sensor with primary calibration	--	--	--	--	--	--
DS	++	++*	+	-	+	+
SSS+ LR	--	--	--	-	+	+
SSS + SVR	--	--	--	-	+	+
PLS + DS	++	++*	+	+	+	+
PCA + DS	++	-	+	+	--	--
PT	--	+	--	--	+	+
PCA + PT	--	+	--	--	+	+
SBC	-	+	-	--	+	+
AMC	--	--	--	-	--	--
MMC	--	-	-	-	+	+
Recalibration with TS	++	+	+	+	+	+

The results show that recalibration with TS works best in most of these cases, even for a minimal TS ratio. While it is attributed to the relatively simple calibration laws (logarithmic or exponential) involved in the present case studies, it means that, in general, recalibration with TS should be the primary approach tested in the calibration transfer study. Then, it appears that PLS+DS works at least as well as recalibration for all the case studies, followed closely by DS. The good performance of PLS+DS is explained by the PLS regression capability to reduce dimensionality while being resilient to missing data. At the same time, the DS approach works well with and without PLS due to its minimalistic approach, which limits the risk of overfitting.

On the other hand, as they rely on stronger hypotheses on dataset structuration, the other methods (SSS, AMC, MMC, PC, and PT) feature minimal performance. They should be tried on a case-by-case basis, particularly when a large number of transfer standards are available.

Finally, no drastic difference is observed between KNN and LR, though, as expected, the limited data availability leads to somewhat reduced performances in some cases.

5.6 Methodology for mass calibration of LOTUS sensor

Developing a calibration model for the LOTUS multiparametric sensor necessitates rigorous experimentation due to the complex nature of multi-analyte detection. However, a model perfectly tailored to one sensor unit may not retain the same level of accuracy when applied to another, even within the same production batch, due to slight variations inherent in the manufacturing process, as discussed earlier.

Our previous analysis suggests that calibration transfer techniques must be selected based on the specificity of the sensor and the model complexity. For the LOTUS multiparametric sensor, simple recalibration may suffice for temperature and conductivity sensors. For chlorine and pH sensors, direct standardization would prove to be more effective, with the anticipation that Partial Least Squares Standardization (PLS+DS) may be better suited as we expand the range of detectable analytes with additional inked chemiresistors.

The facilitation of calibration transfer in mass production is predicated on acquiring ample data for transfer standards. This requirement entails a substantial testing environment and an expansive data acquisition (DAQ) network capable of simultaneous data collection from multiple sensors within the loop, alongside reference sensors to monitor the target parameters, considering the dynamic nature of the water loop in the data modelling.

Post-data collection, temperature, and conductivity models for each multiparametric sensor can undergo recalibration, with new coefficients updated in the AFE (sensor's DAQ system). For chemiresistors, a sufficient baseline of transfer standards must be established before proceeding with calibration transfer. On a case-by-case basis, calibration transfer methods and recalibration can be implemented, the choice of which hinges on the variability among

sensor units. For nominal differences, recalibration may suffice if performance metrics meet the defined tolerances. Larger discrepancies may necessitate calibration transfer techniques to achieve acceptable standards.

The challenge lies in obtaining an optimal amount of transfer standard data for effective calibration transfer without excessive resource consumption—particularly water in this context. This task also demands considerable infrastructure to support a test bench capable of connecting multiple LOTUS sensor units concurrently, a individual host system for data collection from each AFE connected to the sensor, and a sophisticated server-side data architecture for aggregation and analysis of loop-wide sensor data.

5.7 Summary

This study compares 10 standardization-based calibration transfer techniques commonly employed in non-spectral sensor data contexts. These techniques were assessed across three distinct datasets that address issues such as sensor variability, time drift, and environmental fluctuations, all within the constraints of limited availability of transfer standards. Across most evaluated configurations, recalibration utilizing transfer standards demonstrated superior performance, suggesting its merit as a baseline consideration in future applications before resorting to calibration transfer methods. Notably, Partial Least Square standardization and Direct standardization emerged as the most robust approaches, exhibiting capabilities to surpass recalibration in performance. However, the efficacy of other methods remains variable and warrants a case-by-case evaluation. Furthermore, the augmentation of data availability broadens the spectrum of viable transfer methods.

6

Conclusion

Abstract: *This chapter outlines the key conclusions drawn from the dissertation work and presents the most substantial impact of the thesis. This chapter concludes by highlighting potential areas for future exploration in the context of current work.*

6 Conclusions and Scope of Future Work

The primary objective of this thesis was to develop a multiparametric water quality sensor for real-time measurement of pH, active chlorine, conductivity, and temperature in drinking water. This consisted of three main components: (1) Sensor platform, (2) Sensor calibration model, and (3) Methodology for mass calibration of the sensors. The three components were successfully developed, and the integrated sensor was verified in laboratory conditions.

As part of the sensor platform development, the plastic version replaced the silicon version to lower the sensor cost (mainly driven by the cost of raw materials) and to improve the sensor durability. The LOTUS project partners at Gustave Eiffel University completed the substrate fabrication and ink deposition of the sensor. The improvement of the AFE from the initial version developed in the previous project was crucial for increasing the signal-to-noise ratio, boosting sensor performance, and prolonging sensor's operational lifespan. The problems in the previous version were solved by changing the wiring of the circuitry, adding overvoltage protection with Zener bypass, and by modifying the excitation signal that activates the sensor. A data visualization platform was also developed and deployed to show the real-time data from the sensor as well as the estimated target parameters.
Both versions of the sensors were characterized by using methodology established in this thesis. The characterization results of LOTUS sensors by using the Sense-city data have created a solid basis for the sensor's potential use in real-time water quality monitoring. These experiments, especially with the plastic version, demonstrate the sensor's ability to detect critical parameters like active chlorine and pH despite the challenges in achieving consistent and reliable performance.

The silicon version sensor predicted the chlorine level within ±0.12 mg/l and the pH level within ±0.13 units for the testing range of 0 to 5mg/l of chlorine and 7.6 to 8.3 pH. The uncertainties for the chlorine and pH sensing for the silicon version sensor were 0.24 mg/l and 0.21 units respectively. The LPCB sensor in Sense-city had an average error for the chlorine and pH sensing varied in the range [±0.04 mg/l, ±0.03 mg/l] and [±0.42 units, ±0.03 units] respectively, over 3100 hours, with their uncertainties being 0.24 mg/l decreasing to 0.06 mg/l and 0.25 units decreasing to 0.10 units. For the temperature sensor, the silicon version had an average error of 0.2 °C, in the range of 17°C – 20°C with an uncertainty of 3.6 °C (~19% FS) while the LPCB sensor had an average error of 0.75°C, in the range of 10°C – 40°C with an uncertainty of 1.21°C (~3% FS).

Similarly, the conductivity sensor for the silicon version had an average error of ±73 µS/cm with an uncertainty of 50% in the range of 3000 to 5000 µS/cm, while the LPCB sensor had an average error of ±67.6 µS/cm with an uncertainty of 15%, which included the temperature data from the temperature sensor. By using the plastic version, the conductivity and temperature sensors were found to be more precise and overall sensor cost reduced by 40X compared to silicon version. The sensor operational life also has been considerable improved (~4X) than the state of the art reported for chlorine sensing [122], with a comparable performance. However, the gap between lab success and field deployment is still wide, with only a few multiparametric sensors tested and providing limited consistent data. This contrast between excellent results and the need for more model improvement shows the iterative nature of sensor development. The experiments point out a clear direction: increasing the attention on model calibration and the long-term dependability of the sensors in different and changing environmental conditions.

To account for the variability in the sensor ink deposition within the batch, which can pose a challenge to developing a universal calibration model, this work uses a calibration transfer technique, to apply the calibration model developed with one multiparametric sensor, to other multiparametric sensors. Additionally, 10 calibration transfer techniques were evaluated, using 3 multisensor datasets and the performance metrics were compared to find the best technique for the LOTUS sensor. Two different calibration models were used in the analysis, one being the linear regression, which captures the global behavior of the data, and the other model is KNN regression, which uses the local features of the dataset.

The LOTUS sensor does not have the data needed for mass calibration, so the public domain datasets for gas sensor array were used to test the calibration transfer techniques. The first dataset, estimating gas concentration from gas sensor arrays, aimed to transfer the calibration models from one multisensor unit to another multisensor unit, where both units had the same experiments and environment. In this case, Direct Standardization (DS) technique did much better than recalibration (with an MAE of 13 ppm compared to the recalibration MAE of 19 ppm), or any other techniques (which had an MAE higher than the naïve MAE of 59 ppm). Similarly, for the second dataset, estimating gas concentration from a group of 16 MoX sensors, the calibration transfer was done between the same unit but correcting for drift. This unit was tested for 36 months, and the dataset from month 1 was used to build the calibration model, which was transferred to use with the dataset from month 36. Here, both calibration

models did well with PC + DS, PLS + DS and recalibration than with the other techniques, followed by DS (the MAEs 31 ppm, 32 ppm, 36 ppm, 56 ppm respectively).

There was a significant improvement in the performance for KNN based calibration model than LR calibration model, showing the dependence on the model structure. For the third case, dataset from two radio frequency-based humidity sensors were used. Here the model built with the sensor in laboratory conditions was transferred to use with the dataset from the sensor tested in field conditions (similar units, but different environments). In this case, recalibration technique (MAE of 7% RH) did much better than all the other techniques, when used with LR calibration model. On the other hand, with KNN calibration model, 7 out of 10 techniques did better than recalibration technique (MAE 6.4% RH).

The study concluded that the efficiency of the calibration transfer depends significantly on the dataset features and the structure of the calibration model. Based on this, the methodology identified for mass calibration of LOTUS sensors was identified, to use recalibration technique for temperature and conductivity sensors, while exploring the DS and PLS+DS options for chlorine and pH estimation. This requires an extensive testing environment, coupled with a sophisticated data acquisition and analysis infrastructure.

The results showed that the multiparametric LOTUS sensor had great potential to measure various water quality parameters in real time at a low cost, and it could be commercialized by enhancing the ink deposition technology, testing with LOTUS sensors, refining the calibration model and evaluating in real field situations.

6.1 Critical Outlook

Though the plastic LPCB version of the LOTUS sensor features the potential to be used in long-term real-time monitoring of active chlorine, pH, conductivity, and temperature, the performance results from the experiments suffer from inaccuracy, inconsistency, and inadequacy. The prediction capability of the LOTUS sensor for active chlorine and temperature is comparatively good, and the metrics of active chlorine reach the target metric; however, the conductivity estimation did not meet the target set, albeit it was really good in tracking the changes. The performance of pH estimation was the lowest of all the four target parameters. Apart from the lower performance, the inconsistent performance of the sensors (in pH estimation) in sense-city, evidenced by the fluctuating sensitivity of chemiresistors throughout the experiment in Sense-city, and is a matter of concern for field deployment.

These complications may be attributed to the improper fabrication of sensor (uneven dispersion and sensor damage by needle tapping), the model's inability to capture the sensor's dynamic behavior fully and to the lack of varied data in the experiments conducted.

Finally, the inadequacy is a significant gap in this work. Only 2 LPCBs were experimented with in sense-city and only for 3100 hours (~130 days). The results and conclusions are drawn from a minimal subset of sensors, therefore, the minor variations in the sensor play a significant role in determining the performance. Consequently, these results are not representative of technology. Most of the chemiresistors in the evaluated sensors, exhibited their baseline resistance at the maximum threshold, and few were beyond that, thereby not providing a generalized outlook on the performance of the technology.

6.2 Next Steps and Future Work

The path forward for enhancing the LOTUS sensor technology emphasizes refining the ink printing process crucial to the sensor's manufacturing, aiming to elevate precision and operational efficacy. Rigorous real-world validation is a critical milestone for confirming the sensor's reliability and effectiveness.

Future research initiatives are set to extend the LOTUS sensor's capabilities significantly. This includes broadening the sensor's detection and quantification range through additional inks to detect a broader spectrum of analytes, such as arsenic, fluoride, and iron, enhancing its utility in diverse water quality monitoring scenarios. Parallel efforts will optimize the sensor activation protocol and investigate the potential of virtual or soft sensors to augment the sensor's functionality innovatively. Moreover, strategic modifications to the sensor's structural design are planned to accommodate more sensory devices, aiming to include measurements of parameters like turbidity. Improvements in the computing capability of the AFE are also planned for future versions, enabling the implementation of advanced filtering, smoothing, and denoising techniques, as well as allowing for more complex models. Additionally, we will explore the use of deep learning-based methods to further enhance the sensor's performance and expand its detection range.).

These advancements signal a comprehensive approach to refining the LOTUS sensor's performance and broadening its application, marking a significant leap in multiparametric sensing technology for environmental monitoring.

References

[1] United Nations, "UN Water Annual Report," United Nations, 2021. Accessed: Mar. 01, 2024. [Online]. Available: https://www.unwater.org/sites/default/files/app/uploads/2022/06/Annual-Report-2021_30June2022.pdf

[2] United Nations, "Sustainable Development Goals Report," United Nations, 2022. [Online]. Available: https://unstats.un.org/sdgs/report/2022/The-Sustainable-Development-Goals-Report-2022.pdf

[3] V. Sancin and M. Juhart, "The Right to Safe Drinking Water in International Law and in Slovenia's legal framework and implementation," *J. Agric. Environ. Law Agrár. És Környjog.*, vol. 18, no. 34, pp. 106–124, Jun. 2023, doi: 10.21029/JAEL.2023.34.106.

[4] L. A. Patterson and M. W. Doyle, "Measuring water affordability and the financial capability of utilities," *AWWA Water Sci.*, vol. 3, no. 6, p. e1260, Nov. 2021, doi: 10.1002/aws2.1260.

[5] C. Ortiz, C. Salcedo, and J. Saldarriaga, "Assessment of the Effects of COVID-19 Pandemic Stay-at-Home Measures on Potable Water Consumption Patterns, Location, and Financial Impacts for Water Utilities in Colombian Cities," *Water*, vol. 14, no. 19, p. 3004, Sep. 2022, doi: 10.3390/w14193004.

[6] Department of Drinking water and Sanitation, "Drinking Water Quality Monitoring and Surveillance Framework," Ministry of Jal Shakti, Oct. 2021. Accessed: Mar. 01, 2024. [Online]. Available: https://jaljeevanmission.gov.in/sites/default/files/guideline/WQMS-Framework.pdf

[7] M. C. Peter, S. Adeshina, O. Idowu-Bismark, O. Osanaiye, and O. Oyeleke, "Digital Control and Management of Water Supply Infrastructure Using Embedded Systems and Machine Learning," *Int. J. Intell. Syst. Appl.*, vol. 15, no. 5, p. 1.

[8] J. Shijie, S. Liyin, and Z. Li, "Empirical Study on the Contribution of Infrastructure to the Coordinated Development between Urban and Rural Areas: Case Study on Water Supply Projects," *Procedia Environ. Sci.*, vol. 11, pp. 1113–1118, Jan. 2011, doi: 10.1016/j.proenv.2011.12.168.

[9] "Europe's Water in Figures," The European Federation of National Associations of Water Services, Brussels, Belgium. Accessed: Mar. 29, 2024. [Online]. Available: https://www.eureau.org/resources/publications/eureau-publications/5824-europe-s-water-in-figures-2021/file

[10] S. Roy *et al.*, "Assessing and modelling drinking water quality at the railway stations of Tripura, India, with a possible strategic solution," *Arab. J. Geosci.*, vol. 16, no. 2, p. 98, Feb. 2023, doi: 10.1007/s12517-022-11130-1.

[11] S. V. Berg, "Conceptual and Practical Aspects of Water Regulation in Developing Countries," in *Oxford Research Encyclopedia of Global Public Health*, Oxford University Press, 2021. doi: 10.1093/acrefore/9780190632366.013.291.

[12] M. Manga *et al.*, "The effect of household storage tanks/vessels and user practices on the quality of water: a systematic review of literature," *Environ. Syst. Res.*, vol. 10, no. 1, p. 18, Feb. 2021, doi: 10.1186/s40068-021-00221-9.

[13] M. Manga *et al.*, “Impacts of storage tanks under the indirect cold water supply system on household water quality: a case of Wakiso District, Uganda,” *Water Supply*, p. ws2021411, Dec. 2021, doi: 10.2166/ws.2021.411.

[14] W. Lin, C. Ye, L. Guo, D. Hu, and X. Yu, “Analysis of microbial contamination of household water purifiers,” *Appl. Microbiol. Biotechnol.*, vol. 104, no. 10, pp. 4533–4545, May 2020, doi: 10.1007/s00253-020-10510-5.

[15] R. R. Auliya, N. Nurkholis, and M. T. Sabirin, “Public-private partnership water supply project in Indonesia: A public sector review,” *Int. J. Bus. Econ. Manag.*, vol. 6, no. 2, pp. 214–222, Jun. 2023, doi: 10.21744/ijbem.v6n2.2150.

[16] A. W. Utami, D. N. Martono, and H. Agustina, “Water Quality Assessment of West Tarum Canal for Drinking Water Supply,” *E3S Web Conf.*, vol. 277, p. 04003, 2021, doi: 10.1051/e3sconf/202127704003.

[17] A. EludoyiN, “Domestic water quality and the discriminatory influence of socio-economic stratification on accessibility to safe water in a part of Osun State, Nigeria,” *Ege Coğrafya Derg.*, vol. 32, no. 1, pp. 19–31, Jun. 2023, doi: 10.51800/ecd.1253076.

[18] Robert Wurm and Andreas Weingartner, “Innovative Online Water Quality Monitoring in Drinking Water Networks.” Apr. 2019. Accessed: Mar. 31, 2024. [Online]. Available: https://www.s-can.at/wp_contents/uploads/2021/09/iet_innovative_online_water_quality_monitoring_in_drinking_water_networks-april_2019.pdf

[19] T. P. Lambrou, C. G. Panayiotou, and C. C. Anastasiou, “A low-cost system for real time monitoring and assessment of potable water quality at consumer sites,” in *2012 IEEE Sensors*, Taipei, Taiwan: IEEE, Oct. 2012, pp. 1–4. doi: 10.1109/ICSENS.2012.6411190.

[20] J. Crocker and J. Bartram, “Comparison and Cost Analysis of Drinking Water Quality Monitoring Requirements versus Practice in Seven Developing Countries,” *Int. J. Environ. Res. Public. Health*, vol. 11, no. 7, Art. no. 7, Jul. 2014, doi: 10.3390/ijerph110707333.

[21] J. DeZuane, *Handbook of Drinking Water Quality*. John Wiley & Sons, 1997.

[22] L. Manjakkal, L. Lorenzelli, and M. Willander, *Sensing Technologies for Real Time Monitoring of Water Quality*, 1st ed. Wiley, 2023. doi: 10.1002/9781119775843.

[23] J. O’Grady, D. Zhang, N. O’Connor, and F. Regan, “A comprehensive review of catchment water quality monitoring using a tiered framework of integrated sensing technologies,” *Sci. Total Environ.*, vol. 765, p. 142766, Apr. 2021, doi: 10.1016/j.scitotenv.2020.142766.

[24] A. Roy and J. J Kizhakkethottam, “Real Time Water Quality Monitoring of River Pamba (India) using Internet of Things,” *E3S Web Conf.*, vol. 477, p. 00093, 2024, doi: 10.1051/e3sconf/202447700093.

[25] D. P.M, S. S. K, K. A.J, S. R.S, P. M.E, and M. A, “IOT Based Real Time River Water Quality Monitoring and Control System,” *E3S Web Conf.*, vol. 399, p. 04013, 2023, doi: 10.1051/e3sconf/202339904013.

[26] R. Wiryasaputra, C.-Y. Huang, Y.-J. Lin, and C.-T. Yang, “An IoT Real-Time Potable Water Quality Monitoring and Prediction Model Based on Cloud Computing Architecture,” *Sensors*, vol. 24, no. 4, p. 1180, Feb. 2024, doi: 10.3390/s24041180.

[27] M. K. Ch, M. S. K. Munaga, C. S. Kolli, and S. K. Maddila, "Real-Time Water Quality Tracking and Alert System with IoT Integration," in *2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN)*, Salem, India: IEEE, Jun. 2023, pp. 1013–1018. doi: 10.1109/ICPCSN58827.2023.00172.

[28] P. Kruse, "Review on water quality sensors," *J. Phys. Appl. Phys.*, vol. 51, no. 20, p. 203002, May 2018, doi: 10.1088/1361-6463/aabb93.

[29] J. Janata, *Principles of Chemical Sensors*. Springer Science & Business Media, 2010.

[30] G. M. Lazzerini, L. M. Strambini, and G. Barillaro, "Self-tuning porous silicon chemitransistor gas sensors," in *2013 IEEE SENSORS*, Baltimore, MD, USA: IEEE, Nov. 2013, pp. 1–4. doi: 10.1109/ICSENS.2013.6688369.

[31] M. Shaik, V. K. Rao, M. Gupta, K. S. R. C. Murthy, and R. Jain, "Chemiresistive gas sensor for the sensitive detection of nitrogen dioxide based on nitrogen doped graphene nanosheets," *RSC Adv.*, vol. 6, no. 2, pp. 1527–1534, 2016, doi: 10.1039/C5RA21184K.

[32] N. Vijayakumar and R. Ramya, "The real time monitoring of water quality in IoT environment," in *2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, Coimbatore, India: IEEE, Mar. 2015, pp. 1–5. doi: 10.1109/ICIIECS.2015.7193080.

[33] S. A. Jaywant and K. M. Arif, "A Comprehensive Review of Microfluidic Water Quality Monitoring Sensors," *Sensors*, vol. 19, no. 21, p. 4781, Nov. 2019, doi: 10.3390/s19214781.

[34] A. Kapoor, S. Balasubramanian, P. Muthamilselvi, V. Vaishampayan, and S. Prabhakar, "Lab-on-a-Chip Devices for Water Quality Monitoring," in *Nanosensor Technologies for Environmental Monitoring*, Inamuddin and A. M. Asiri, Eds., Cham: Springer International Publishing, 2020, pp. 455–469. doi: 10.1007/978-3-030-45116-5_15.

[35] B. Uka, J. Kieninger, G. A. Urban, and A. Weltin, "Electrochemical Microsensor for Microfluidic Glyphosate Monitoring in Water Using MIP-Based Concentrators," *ACS Sens.*, vol. 6, no. 7, pp. 2738–2746, Jul. 2021, doi: 10.1021/acssensors.1c00884.

[36] A. U. Alam, D. Clyne, H. Jin, N.-X. Hu, and M. J. Deen, "Fully Integrated, Simple, and Low-Cost Electrochemical Sensor Array for in Situ Water Quality Monitoring," *ACS Sens.*, vol. 5, no. 2, pp. 412–422, Feb. 2020, doi: 10.1021/acssensors.9b02095.

[37] "LOTUS – LOw-cost innovative Technology for water quality monitoring and water resources management for Urban and rural water Systems in India." Accessed: Dec. 04, 2023. [Online]. Available: https://www.lotus-india.eu/

[38] B. Lebental *et al.*, "Water and air quality monitoring with multiparameter chemical sensors Managing non-idealities from lab to field," in *2022 IEEE Sensors*, Oct. 2022, pp. 01–04. doi: 10.1109/SENSORS52175.2022.9967256.

[39] Lebental, Bérengère, "Project LOTUS (#820881) Deliverable D2.2 first version of lotus prototype: Public summary," Ecole Polytechnique, Palaiseau, France, Project Deliverable D2.2, 2022. [Online]. Available: https://www.lotus-india.eu/wp-content/uploads/2022/06/D2.2-First-version-lotus-prototype-public-summary.pdf

[40] B. V. Muppidathi, S. Laporte, Y. Ulanowski, S. Subbiah, and B. Lebental, "Performance of a Multiparametric Water Quality Sensor in a Small-Scale Water

Distribution Network," in *2023 IEEE SENSORS*, Vienna, Austria: IEEE, Oct. 2023, pp. 1–4. doi: 10.1109/SENSORS56945.2023.10324917.

[41] B. V. Muppidathi, S. Laporte, Y. Ulanowski, S. Subbiah, and B. Lebental, "Multiparametric water quality sensor based on carbon nanotubes: Performance assessment in realistic environment," presented at the EGU General Assembly 2023, in EGU23-16053. Vienna, Austria, Apr. 2023. doi: https://doi.org/10.5194/egusphere-egu23-16053.

[42] C. M. Amrita and D. Babiyola, "ANALYSING THE WATER QUALITY PARAMETERS FROM TRADITIONAL TO MODERN METHODS IN AQUACULTURE".

[43] M. N. Barabde and S. R. Danve, "Continuous Water Quality Monitoring System for Water Resources at Remote Places," vol. 3, no. 3, 2015.

[44] A. De Das, A. Pramanik, and A. Adak, "Evolution in Water Monitoring Technology," in *Proceedings of the 21st International Conference on Distributed Computing and Networking*, Kolkata India: ACM, Jan. 2020, pp. 1–5. doi: 10.1145/3369740.3372773.

[45] S. Zhuiykov, "Solid-state sensors monitoring parameters of water quality for the next generation of wireless sensor networks," *Sens. Actuators B Chem.*, vol. 161, no. 1, pp. 1–20, Jan. 2012, doi: 10.1016/j.snb.2011.10.078.

[46] P. M. Pujar, H. H. Kenchannavar, R. M. Kulkarni, and U. P. Kulkarni, "Real-time water quality monitoring through Internet of Things and ANOVA-based analysis: a case study on river Krishna," *Appl. Water Sci.*, vol. 10, no. 1, p. 22, Dec. 2019, doi: 10.1007/s13201-019-1111-9.

[47] J.-Y. Lin, H.-L. Tsai, and W.-H. Lyu, "An Integrated Wireless Multi-Sensor System for Monitoring the Water Quality of Aquaculture," *Sensors*, vol. 21, no. 24, Art. no. 24, Jan. 2021, doi: 10.3390/s21248179.

[48] P. Mahajan and P. Shahane, "An IoT Based System for Water Quality Monitoring," *SSRN Electron. J.*, 2021, doi: 10.2139/ssrn.3769765.

[49] S. Pasika and S. T. Gandla, "Smart water quality monitoring system with cost-effective using IoT," *Heliyon*, vol. 6, no. 7, p. e04096, Jul. 2020, doi: 10.1016/j.heliyon.2020.e04096.

[50] "Guidelines for drinking-water quality: fourth edition incorporating the first and second addenda," World Health Organization, Geneva, 2022. Accessed: Mar. 04, 2024. [Online]. Available: https://www.who.int/publications-detail-redirect/9789240045064

[51] C. E. Boyd, *Water Quality: An Introduction*. Cham: Springer International Publishing, 2015. doi: 10.1007/978-3-319-17446-4.

[52] E. R. Alley, *Water quality control handbook*, 2nd ed. in McGraw-Hill handbooks. New York : Alexandria, Va: McGraw-Hill ; WEF Press, 2007.

[53] J. Cotruvo, *Drinking Water Quality and Contaminants Guidebook*, 1st ed. Boca Raton : Taylor & Francis, a CRC title, part of the Taylor &: CRC Press, 2018. doi: 10.1201/9781351110471.

[54] N. Hassan Omer, "Water Quality Parameters," in *Water Quality - Science, Assessments and Policy*, K. Summers, Ed., IntechOpen, 2020. doi: 10.5772/intechopen.89657.

[55] F. R. Spellman, *The Drinking Water Handbook*, 3rd ed. Boca Raton : Taylor & Francis, CRC Press, 2018: CRC Press, 2017. doi: 10.1201/9781315159126.

[56] EPA Water Security Division, "Online Water Quality Monitoring in Distribution Systems for Water Quality Surveillance and Response Systems," United States Environmental Protection Agency, 2019.

[57] "Indian Standard: Drinking water - Specification," Bureau of Indian Standards, New Delhi, India, IS 10500 : 2012, 2015.

[58] S. Liu, C. Gunawan, N. Barraud, S. A. Rice, E. J. Harry, and R. Amal, "Understanding, Monitoring, and Controlling Biofilm Growth in Drinking Water Distribution Systems," *Environ. Sci. Technol.*, vol. 50, no. 17, pp. 8954–8976, Sep. 2016, doi: 10.1021/acs.est.6b00835.

[59] H. A. Mohammed and S. F. Ismail, "Design and implementation of remotely monitoring system for pH level in Baghdad drinking water networks," *TELKOMNIKA Telecommun. Comput. Electron. Control*, vol. 19, no. 3, p. 1030, Jun. 2021, doi: 10.12928/telkomnika.v19i3.12921.

[60] Y. Villasana, N. Moradi, C. Navas-Cárdenas, and G. S. Patience, "Experimental methods in chemical engineering: pH," *Can. J. Chem. Eng.*, vol. 100, no. 8, pp. 1703–1717, 2022, doi: 10.1002/cjce.24393.

[61] D. E. Helbling and J. M. VanBriesen, "Continuous monitoring of residual chlorine concentrations in response to controlled microbial intrusions in a laboratory-scale distribution system," *Water Res.*, vol. 42, no. 12, pp. 3162–3172, Jun. 2008, doi: 10.1016/j.watres.2008.03.009.

[62] N. Janudin *et al.*, "Sensing Techniques on Determination of Chlorine Gas and Free Chlorine in Water," *J. Sens.*, vol. 2022, pp. 1–27, Jun. 2022, doi: 10.1155/2022/1898417.

[63] L. H. H. Hsu, E. Hoque, P. Kruse, and P. Ravi Selvaganapathy, "A carbon nanotube based resettable sensor for measuring free chlorine in drinking water," *Appl. Phys. Lett.*, vol. 106, no. 6, p. 063102, Feb. 2015, doi: 10.1063/1.4907631.

[64] P. Kruse, "(Invited) Chemiresistive Water Quality Sensors: Challenges and Progress," *ECS Meet. Abstr.*, vol. MA2022-01, no. 52, p. 2135, Jul. 2022, doi: 10.1149/MA2022-01522135mtgabs.

[65] A. Zubiarrain-Laserna and P. Kruse, "Review—Graphene-Based Water Quality Sensors," *J. Electrochem. Soc.*, vol. 167, no. 3, p. 037539, Jan. 2020, doi: 10.1149/1945-7111/ab67a5.

[66] B. Marczewska and K. Marczewski, "First Glass Electrode and its Creators F. Haber and Z. Klemensiewicz – On 100th Anniversary," *Z. Für Phys. Chem.*, vol. 224, no. 05, pp. 795–799, May 2010, doi: 10.1524/zpch.2010.5505.

[67] D. S. Tarbell and A. T. Tarbell, "The development of the pH meter," *J. Chem. Educ.*, vol. 57, no. 2, p. 133, Feb. 1980, doi: 10.1021/ed057p133.

[68] M. H. Banna *et al.*, "Online Drinking Water Quality Monitoring: Review on Available and Emerging Technologies," *Crit. Rev. Environ. Sci. Technol.*, vol. 44, no. 12, pp. 1370–1421, Jun. 2014, doi: 10.1080/10643389.2013.781936.

[69] D. Wencel, T. Abel, and C. McDonagh, "Optical Chemical pH Sensors," *Anal. Chem.*, vol. 86, no. 1, pp. 15–29, Jan. 2014, doi: 10.1021/ac4035168.

[70] A. Fog and R. P. Buck, “Electronic semiconducting oxides as pH sensors,” *Sens. Actuators*, vol. 5, no. 2, pp. 137–146, Feb. 1984, doi: 10.1016/0250-6874(84)80004-9.

[71] P. Bergveld, “Thirty years of ISFETOLOGY: What happened in the past 30 years and what may happen in the next 30 years,” *Sens. Actuators B Chem.*, vol. 88, no. 1, pp. 1–20, Jan. 2003, doi: 10.1016/S0925-4005(02)00301-5.

[72] D. K. Kampouris, R. O. Kadara, N. Jenkinson, and C. E. Banks, “Screen printed electrochemical platforms for pH sensing,” *Anal. Methods*, vol. 1, no. 1, pp. 25–28, Sep. 2009, doi: 10.1039/B9AY00025A.

[73] N. Mzoughi *et al.*, “Characterization of novel impedimetric pH-sensors based on solution-processable biocompatible thin-film semiconducting organic coatings,” *Sens. Actuators B Chem.*, vol. 171–172, pp. 537–543, Aug. 2012, doi: 10.1016/j.snb.2012.05.029.

[74] L.-M. Kuo, K.-N. Chen, Y.-L. Chuang, and S. Chao, “A Flexible pH-Sensing Structure Using WO3/IrO2 Junction with Al2O3 Encapsulation Layer,” *ECS Solid State Lett.*, vol. 2, no. 3, pp. P28–P30, Dec. 2012, doi: 10.1149/2.004303ssl.

[75] A. Das *et al.*, “Highly sensitive palladium oxide thin film extended gate FETs as pH sensor,” *Sens. Actuators B Chem.*, vol. 205, pp. 199–205, Dec. 2014, doi: 10.1016/j.snb.2014.08.057.

[76] E. Hoque, L. H. H. Hsu, A. Aryasomayajula, P. R. Selvaganapathy, and P. Kruse, “Pencil-Drawn Chemiresistive Sensor for Free Chlorine in Water,” *IEEE Sens. Lett.*, vol. 1, no. 4, pp. 1–4, Aug. 2017, doi: 10.1109/LSENS.2017.2722958.

[77] R. C. Faria and L. O. S. Bulhões, “Hydrogen ion selective electrode based on poly(1-aminoanthracene) film,” *Anal. Chim. Acta*, vol. 377, no. 1, pp. 21–27, Dec. 1998, doi: 10.1016/S0003-2670(98)00522-4.

[78] H. Kaden, H. Jahn, and M. Berthold, “Study of the glass/polypyrrole interface in an all-solid-state pH sensor,” *Solid State Ion.*, vol. 169, no. 1, pp. 129–133, Apr. 2004, doi: 10.1016/S0167-2738(03)00216-9.

[79] M. A. Rahman, P. Kumar, D.-S. Park, and Y.-B. Shim, “Electrochemical Sensors Based on Organic Conjugated Polymers,” *Sensors*, vol. 8, no. 1, Art. no. 1, Jan. 2008, doi: 10.3390/s8010118.

[80] A. Bonfiglio, I. Manunza, P. Cosseddu, and E. Orgiu, “Detection of Chemical and Physical Parameters by Means of Organic Field-Effect Transistors,” in *Organic Semiconductors in Sensor Applications*, vol. 107, D. A. Bernards, G. G. Malliaras, and R. M. Owens, Eds., in Materials Science, vol. 107. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 185–212. doi: 10.1007/978-3-540-76314-7_6.

[81] C. Bartic and G. Borghs, “Organic thin-film transistors as transducers for (bio) analytical applications,” *Anal. Bioanal. Chem.*, vol. 384, no. 2, pp. 354–365, Jan. 2006, doi: 10.1007/s00216-005-0031-8.

[82] M. W. Shinwari and M. J. Deen, “Impedance Modelling of FET-Based Biosensors,” *J. Electrochem. Soc.*, vol. 158, no. 6, p. J189, Apr. 2011, doi: 10.1149/1.3579509.

[83] Y. T. Lee, E. Lee, J. M. Lee, and W. Lee, “Micro-sized pH sensors based on patterned Pd structures using an electrolysis method,” *Curr. Appl. Phys.*, vol. 9, no. 4, Supplement, pp. e218–e221, Jul. 2009, doi: 10.1016/j.cap.2009.06.022.

[84] "Hierarchically Structured Suspended TiO2 Nanofibers for Use in UV and pH Sensor Devices | ACS Applied Materials & Interfaces." Accessed: Mar. 05, 2024. [Online]. Available: https://pubs.acs.org/doi/10.1021/am501563v

[85] Y. Qin, H.-J. Kwon, M. M. R. Howlader, and M. J. Deen, "Microfabricated electrochemical pH and free chlorine sensors for water quality monitoring: recent advances and research challenges," *RSC Adv.*, vol. 5, no. 85, pp. 69086–69109, 2015, doi: 10.1039/C5RA11291E.

[86] J.-H. Kwon, K.-S. Lee, Y.-H. Lee, and B.-K. Ju, "Single-Wall Carbon Nanotube-Based pH Sensor Fabricated by the Spray Method," *Electrochem. Solid-State Lett.*, vol. 9, no. 9, p. H85, Jun. 2006, doi: 10.1149/1.2217131.

[87] P. Li, C. M. Martin, K. K. Yeung, and W. Xue, "Dielectrophoresis Aligned Single-Walled Carbon Nanotubes as pH Sensors," *Biosensors*, vol. 1, no. 1, Art. no. 1, Mar. 2011, doi: 10.3390/bios1010023.

[88] N. Lei, P. Li, W. Xue, and J. Xu, "Simple graphene chemiresistors as pH sensors: fabrication and characterization," *Meas. Sci. Technol.*, vol. 22, no. 10, p. 107002, Aug. 2011, doi: 10.1088/0957-0233/22/10/107002.

[89] Y. Liao *et al.*, "Carbon Nanotube/Polyaniline Composite Nanofibers: Facile Synthesis and Chemosensors," *Nano Lett.*, vol. 11, no. 3, pp. 954–959, Mar. 2011, doi: 10.1021/nl103322b.

[90] K. F. Lei, K.-F. Lee, and S.-I. Yang, "Fabrication of carbon nanotube-based pH sensor for paper-based microfluidics," *Microelectron. Eng.*, vol. 100, pp. 1–5, Dec. 2012, doi: 10.1016/j.mee.2012.07.113.

[91] D. Jung, M.-E. Han, and G. S. Lee, "pH-sensing characteristics of multi-walled carbon nanotube sheet," *Mater. Lett.*, vol. 116, pp. 57–60, Feb. 2014, doi: 10.1016/j.matlet.2013.10.095.

[92] P. Gou *et al.*, "Carbon Nanotube Chemiresistor for Wireless pH Sensing," *Sci. Rep.*, vol. 4, no. 1, p. 4468, Mar. 2014, doi: 10.1038/srep04468.

[93] Y. Qin *et al.*, "Inkjet-printed bifunctional carbon nanotubes for pH sensing," *Mater. Lett.*, vol. 176, pp. 68–70, Aug. 2016, doi: 10.1016/j.matlet.2016.04.048.

[94] G. L. Goh, S. Agarwala, Y. J. Tan, and W. Y. Yeong, "A low cost and flexible carbon nanotube pH sensor fabricated using aerosol jet technology for live cell applications," *Sens. Actuators B Chem.*, vol. 260, pp. 227–235, May 2018, doi: 10.1016/j.snb.2017.12.127.

[95] J.-Y. Jeon, B.-C. Kang, and T.-J. Ha, "Flexible pH sensors based on printed nanocomposites of single-wall carbon nanotubes and Nafion," *Appl. Surf. Sci.*, vol. 514, p. 145956, Jun. 2020, doi: 10.1016/j.apsusc.2020.145956.

[96] H. Emami, S. Mahinnezhad, A. Al Shboul, M. Ketabi, A. Shih, and R. Izquierdo, "Flexible Chemiresistive pH Sensor Based on Polyaniline / Carbon Nanotube Nanocomposite for IoT Applications," in *2021 IEEE Sensors*, Sydney, Australia: IEEE, Oct. 2021, pp. 1–4. doi: 10.1109/SENSORS47087.2021.9639621.

[97] G. Cho, S. Azzouzi, G. Zucchi, and B. Lebental, "Electrical and Electrochemical Sensors Based on Carbon Nanotubes for the Monitoring of Chemicals in Water—A Review," *Sensors*, vol. 22, no. 1, Art. no. 1, Jan. 2022, doi: 10.3390/s22010218.

[98] M. McGarrity and F. Zhao, “Graphene-Based Chemiresistor Sensors for Drinking Water Quality Monitoring,” *Sensors*, vol. 23, no. 24, p. 9828, Dec. 2023, doi: 10.3390/s23249828.

[99] A. Z. Laserna, “Nanocarbon Based Chemiresistive Water Quality Sensors,” 2019. Accessed: Apr. 01, 2024. [Online]. Available: https://www.semanticscholar.org/paper/Nanocarbon-Based-Chemiresistive-Water-Quality-Laserna/765ef7b4870927023bf0fedff9ecf0c1477a468c

[100] “edge pH and ORP Meter - HI2002-01 | Hanna Instruments.” Accessed: Apr. 01, 2024. [Online]. Available: https://www.hannainst.com/edge-dedicated-ph-orp-meter.html

[101] HORIBA Advanced Techno, Co., Ltd., “LAQUA Horiba - pH meters.” Accessed: Apr. 01, 2024. [Online]. Available: https://www.horiba.com/int/water-quality/parameters/ph/

[102] Endress+Hauser AG, “Digital non-glass pH sensor Memosens CPS77E.” Accessed: Apr. 01, 2024. [Online]. Available: https://www.endress.com/en/field-instruments-overview/liquid-analysis-product-overview/digital-pH-sensor-Memosens-CPS77E

[103] “PyroScience GmbH.” Accessed: Apr. 01, 2024. [Online]. Available: https://www.pyroscience.com

[104] “Optical pH Sensors.” Accessed: Apr. 01, 2024. [Online]. Available: https://www.presens.de/products/ph/sensors

[105] M. Lu and R. G. Compton, “Voltammetric pH sensing using carbon electrodes: glassy carbon behaves similarly to EPPG,” *Analyst*, vol. 139, no. 18, pp. 4599–4605, Aug. 2014, doi: 10.1039/C4AN00866A.

[106] G. M. da Silva, S. G. Lemos, L. A. Pocrifka, P. D. Marreto, A. V. Rosario, and E. C. Pereira, “Development of low-cost metal oxide pH electrodes based on the polymeric precursor method,” *Anal. Chim. Acta*, vol. 616, no. 1, pp. 36–41, May 2008, doi: 10.1016/j.aca.2008.03.019.

[107] L. Manjakkal, D. Szwagierczak, and R. Dahiya, “Metal oxides based electrochemical pH sensors: Current progress and future perspectives,” *Prog. Mater. Sci.*, vol. 109, p. 100635, Apr. 2020, doi: 10.1016/j.pmatsci.2019.100635.

[108] G. C. White, *The handbook of chlorination*, 2nd ed. New York: Van Nostrand Reinhold Co, 1986.

[109] J. Ellis and P. L. Brown, “Determination of residual chlorine by derivatisation with 2,6-dimethylphenol and gas chromatographic separation,” *Anal. Chim. Acta*, vol. 124, no. 2, pp. 431–436, Mar. 1981, doi: 10.1016/S0003-2670(01)93593-7.

[110] T. Nakagama, M. Yamada, and T. Hobo, “Chemiluminescence sensor with uranine immobilized on an anion-exchange resin for monitoring free chlorine in tap water,” *Anal. Chim. Acta*, vol. 231, pp. 7–12, Jan. 1990, doi: 10.1016/S0003-2670(00)86390-4.

[111] F. J. D. Campo, O. Ordeig, and F. J. Muñoz, “Improved free chlorine amperometric sensor chip for drinking water applications,” *Anal. Chim. Acta*, vol. 554, no. 1, pp. 98–104, Dec. 2005, doi: 10.1016/j.aca.2005.08.035.

[112] F. Kodera, M. Umeda, and A. Yamada, “Determination of free chlorine based on anodic voltammetry using platinum, gold, and glassy carbon electrodes,” *Anal. Chim. Acta*, vol. 537, no. 1, pp. 293–298, Apr. 2005, doi: 10.1016/j.aca.2005.01.053.

[113] G. Gordon, D. L. Sweetin, K. Smith, and G. E. Pacey, "Improvements in the *N*,*N*-diethyl-*p*-phenylenediamine method for the determination of free and combined residual chlorine through the use of FIA," *Talanta*, vol. 38, no. 2, pp. 145–149, Feb. 1991, doi: 10.1016/0039-9140(91)80122-G.

[114] D. Pletcher and E. M. Valdes, "Studies of a microelectrode sensor for monitoring chlorine in water supplies," *Anal. Chim. Acta*, vol. 246, no. 2, pp. 267–273, Jun. 1991, doi: 10.1016/S0003-2670(00)80960-5.

[115] A. van den Berg, A. Grisel, E. Verney-Norberg, B. H. van der Schoot, M. Koudelka-Hep, and N. F. de Rooij, "On-wafer fabricated free-chlorine sensor with ppb detection limit for drinking-water monitoring," *Sens. Actuators B Chem.*, vol. 13, no. 1, pp. 396–399, May 1993, doi: 10.1016/0925-4005(93)85410-C.

[116] S. Pan, M. J. Deen, and R. Ghosh, "Low-Cost Graphite-Based Free Chlorine Sensor," *Anal. Chem.*, vol. 87, no. 21, pp. 10734–10737, Nov. 2015, doi: 10.1021/acs.analchem.5b03164.

[117] Y. Qin, S. Pan, M. M. R. Howlader, R. Ghosh, N.-X. Hu, and M. J. Deen, "Paper-Based, Hand-Drawn Free Chlorine Sensor with Poly(3,4-ethylenedioxythiophene):Poly(styrenesulfonate)," *Anal. Chem.*, vol. 88, no. 21, pp. 10384–10389, Nov. 2016, doi: 10.1021/acs.analchem.6b03211.

[118] F. Davis, S. D. Collyer, D. D. Gornall, K. A. Law, D. W. Mills, and S. P. J. Higson, "New Techniques in Monitoring Water Pollution - Development of Sonochemically Fabricated Microarrays for the Determination of Pollutants," *Chim. Oggi-Chem. Today*, vol. 25, pp. 28–31, 2007.

[119] Y.-K. Yen, K.-Y. Lee, C.-Y. Lin, S.-T. Zhang, C.-W. Wang, and T.-Y. Liu, "Portable Nanohybrid Paper-Based Chemiresistive Sensor for Free Chlorine Detection," *ACS Omega*, vol. 5, no. 39, pp. 25209–25215, Oct. 2020, doi: 10.1021/acsomega.0c03366.

[120] M. Murata, T. A. Ivandini, M. Shibata, S. Nomura, A. Fujishima, and Y. Einaga, "Electrochemical detection of free chlorine at highly boron-doped diamond electrodes," *J. Electroanal. Chem.*, vol. 612, no. 1, pp. 29–36, Jan. 2008, doi: 10.1016/j.jelechem.2007.09.006.

[121] L. Yang, M. Li, Y. Qu, Z. Dong, and W. J. Li, "Carbon nanotube-sensor-integrated microfluidic platform for real-time chemical concentration detection," *ELECTROPHORESIS*, vol. 30, no. 18, pp. 3198–3205, 2009, doi: 10.1002/elps.200900126.

[122] R. Olivé-Monllau *et al.*, "Flow injection analysis system based on amperometric thin-film transducers for free chlorine detection in swimming pool waters," *Talanta*, vol. 77, no. 5, pp. 1739–1744, Mar. 2009, doi: 10.1016/j.talanta.2008.10.015.

[123] K. Senthilkumar and J.-M. Zen, "Free chlorine detection based on EC' mechanism at an electroactive polymelamine-modified electrode," *Electrochem. Commun.*, vol. 46, pp. 87–90, Sep. 2014, doi: 10.1016/j.elecom.2014.06.018.

[124] A. Mohtasebi, A. D. Broomfield, T. Chowdhury, P. R. Selvaganapathy, and P. Kruse, "Reagent-Free Quantification of Aqueous Free Chlorine via Electrical Readout of Colorimetrically Functionalized Pencil Lines," *ACS Appl. Mater. Interfaces*, vol. 9, no. 24, pp. 20748–20761, Jun. 2017, doi: 10.1021/acsami.7b03968.

[125] R. E. Wilson, I. Stoianov, and D. O'Hare, "Continuous Chlorine Detection in Drinking Water and a Review of New Detection Methods," *Johns. Matthey Technol.*

Rev., vol. 63, no. 2, pp. 103–118, Apr. 2019, doi: 10.1595/205651318X15367593796080.

[126] A. U. Alam, D. Clyne, W. Lush, and M. J. Deen, "A reusable, reagent-less free chlorine sensor using gold thin film electrode," *Analyst*, vol. 146, no. 8, pp. 2626–2631, Apr. 2021, doi: 10.1039/D1AN00038A.

[127] A. Zubiarrain-Laserna, S. Angizi, M. A. Akbar, R. Divigalpitiya, P. R. Selvaganapathy, and P. Kruse, "Detection of free chlorine in water using graphene-like carbon based chemiresistive sensors," *RSC Adv.*, vol. 12, no. 4, pp. 2485–2496, Jan. 2022, doi: 10.1039/D1RA08264G.

[128] J. Siddiqui, M. Taheri, M. Nami, I. A. Deen, M. Packirisamy, and M. J. Deen, "Carbon-Based Electrochemical-Free Chlorine Sensors," *Adv. Mater. Technol.*, vol. 8, no. 19, p. 2300717, 2023, doi: 10.1002/admt.202300717.

[129] E. H. Seymour, N. S. Lawrence, and R. G. Compton, "Reaction with N,N-Diethyl-p-phenylenediamine: A Procedure for the Sensitive Square-Wave Voltammetric Detection of Chlorine," *Electroanalysis*, vol. 15, no. 8, pp. 689–694, 2003, doi: 10.1002/elan.200390086.

[130] "DPD Chlor-Analyzer Chlorine 3017M - WTW." Accessed: Apr. 01, 2024. [Online]. Available: https://www.xylemanalytics.com/en/en/general-product/id-1787/dpd-chlor-analyzer-chlorine-3017m----wtw

[131] "DPD Chlorine Analyzer for Municipal Water | ysi.com." Accessed: Apr. 01, 2024. [Online]. Available: https://www.ysi.com/3017m

[132] "Chlorine Analyzers - HACH India." Accessed: Apr. 01, 2024. [Online]. Available: https://in.hach.com/chlorine-analyzers/family?productCategoryId=22216783833

[133] M. Szili, I. Kasik, V. Matejec, G. Nagy, and B. Kovacs, "Poly(luminol) based sensor array for determination of dissolved chlorine in water," *Sens. Actuators B Chem.*, vol. 192, pp. 92–98, Mar. 2014, doi: 10.1016/j.snb.2013.10.080.

[134] "Best Chlorine Sensor? - DPD vs. Amperometric Chlorine Analyzers," Halogen Systems Inc. Accessed: Apr. 01, 2024. [Online]. Available: https://halogensys.com/chlorometric-vs-amperometric-vs-halogen/

[135] "Data sheet - ChloroStar Cell | Free and total chlorine cell." Accessed: Apr. 01, 2024. [Online]. Available: https://search.abb.com/library/Download.aspx?DocumentID=DS%2fACL410&LanguageCode=en&DocumentPartId=&Action=Launch

[136] M. Hartwig, R. Zichner, and Y. Joseph, "Inkjet-Printed Wireless Chemiresistive Sensors—A Review," *Chemosensors*, vol. 6, no. 4, Art. no. 4, Dec. 2018, doi: 10.3390/chemosensors6040066.

[137] V. Schroeder, S. Savagatrup, M. He, S. Lin, and T. M. Swager, "Carbon Nanotube Chemical Sensors," *Chem. Rev.*, vol. 119, no. 1, pp. 599–663, Jan. 2019, doi: 10.1021/acs.chemrev.8b00340.

[138] P. Cousin *et al.*, "Improving Water Quality and Security with Advanced Sensors and Indirect Water Sensing Methods," in *Instrumentation and Measurement Technologies for Water Cycle Management*, A. Di Mauro, A. Scozzari, and F. Soldovieri, Eds., in Springer Water. , Cham: Springer International Publishing, 2022, pp. 251–277. doi: 10.1007/978-3-031-08262-7_11.

[139] A. M. C. Ilie, C. Vaccaro, J. Rogeiro, T. E. Leitão, and T. Martins, "Configuration, programming and implementation of 3 Smart Water network wireless sensor nodes for assessing the water quality," in *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, Aug. 2017, pp. 1–8. doi: 10.1109/UIC-ATC.2017.8397442.

[140] "Smart Water Management IoT Solution | Water Quality Sensors Certified," Libelium. Accessed: Mar. 10, 2024. [Online]. Available: https://www.libelium.com/iot-solutions/smart-water/

[141] L. Cui, M. Qiu, H. Fan, L. Wang, S. Zhao, and J. Du, "Design and Research of Water Quality Monitoring System Based on nRF24L01 Wireless Communication," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 208, no. 1, p. 012074, Dec. 2018, doi: 10.1088/1755-1315/208/1/012074.

[142] Y. He, "Design of water quality monitoring platform based on Embedded System," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 631, no. 1, p. 012020, Jan. 2021, doi: 10.1088/1755-1315/631/1/012020.

[143] E. C. Okpara, B. E. Sehularo, and O. B. Wojuola, "On-line water quality inspection system: the role of the wireless sensory network," *Environ. Res. Commun.*, vol. 4, no. 10, p. 102001, Oct. 2022, doi: 10.1088/2515-7620/ac9aa5.

[144] P. Giménez-Gómez, R. Escudé-Pujol, C. Jiménez-Jorquera, and M. Gutiérrez-Capitán, "Multisensor Portable Meter for Environmental Applications," *IEEE Sens. J.*, vol. 15, no. 11, pp. 6517–6523, Nov. 2015, doi: 10.1109/JSEN.2015.2460011.

[145] Y. Vlasov, A. Legin, A. Rudnitskaya, C. D. Natale, and A. D'Amico, "Nonspecific sensor arrays ('electronic tongue') for chemical analysis of liquids (IUPAC Technical Report)," *Pure Appl. Chem.*, vol. 77, no. 11, pp. 1965–1983, Jan. 2005, doi: 10.1351/pac200577111965.

[146] A. Legin, A. Rudnitskaya, Y. Vlasov, C. Di Natale, E. Mazzone, and A. D'Amico, "Application of Electronic Tongue for Quantitative Analysis of Mineral Water and Wine," *Electroanalysis*, vol. 11, no. 10–11, pp. 814–820, 1999, doi: 10.1002/(SICI)1521-4109(199907)11:10/11<814::AID-ELAN814>3.0.CO;2-7.

[147] Yu. G. Vlasov, A. V. Legin, A. M. Rudnitskaya, and K. D. Natale, "A multisensor system based on a group of chemical sensors and artificial neuron neworks electronic language for quantitative analysis of multi-component aqueous solutions," *Zhurnal Prikl. Khimii*, vol. 69, no. 6, pp. 958–964, 1996.

[148] E. Witkowska Nery, J. A. Guimarães, and L. T. Kubota, "Paper-Based Electronic Tongue," *Electroanalysis*, vol. 27, no. 10, pp. 2357–2362, Oct. 2015, doi: 10.1002/elan.201500054.

[149] N. Carbó *et al.*, "Quantitative Determination of Spring Water Quality Parameters via Electronic Tongue," *Sensors*, vol. 18, no. 1, Art. no. 1, Jan. 2018, doi: 10.3390/s18010040.

[150] M. Cuartero, A. Ruiz, M. Galián, and J. A. Ortuño, "Potentiometric Electronic Tongue for Quantitative Ion Analysis in Natural Mineral Waters," *Sensors*, vol. 22, no. 16, Art. no. 16, Jan. 2022, doi: 10.3390/s22166204.

[151] D. Kirsanov, M. Khaydukova, L. Tkachenko, A. Legin, and V. Babain, “Potentiometric Sensor Array for Analysis of Complex Rare Earth Mixtures,” *Electroanalysis*, vol. 24, no. 1, pp. 121–130, 2012, doi: https://doi.org/10.1002/elan.201100439.

[152] D. Kirsanov, O. Zadorozhnaya, A. Krasheninnikov, N. Komarova, A. Popov, and A. Legin, “Water toxicity evaluation in terms of bioassay with an Electronic Tongue,” *Sens. Actuators B Chem.*, vol. 179, pp. 282–286, 2013, doi: 10.1016/j.snb.2012.09.106.

[153] D. Kirsanov, O. Zadorozhnaia, A. Legin, A. Krasheninnikov, N. Komarova, and A. Popov, “Water toxicity assessment with potentiometric multisensor system,” in *Proceedings of the International Conference on Sensing Technology, ICST*, IEEE, 2012, pp. 611–613. doi: 10.1109/ICSensT.2012.6461752.

[154] E. Legin *et al.*, “Rapid Evaluation of Integral Quality and Safety of Surface and Waste Waters by a Multisensor System (Electronic Tongue),” *Sensors*, vol. 19, no. 9, Art. no. 9, Jan. 2019, doi: 10.3390/s19092019.

[155] D. Kirsanov, E. Legin, A. Zagrebin, N. Ignatieva, V. Rybakin, and A. Legin, “Mimicking Daphnia magna bioassay performance by an electronic tongue for urban water quality control,” *Anal. Chim. Acta*, vol. 824, pp. 64–70, 2014, doi: 10.1016/j.aca.2014.03.021.

[156] D. Kirsanov *et al.*, “A Pencil-Drawn Electronic Tongue for Environmental Applications,” *Sensors*, vol. 21, no. 13, Art. no. 13, Jan. 2021, doi: 10.3390/s21134471.

[157] X. Niu, N. Cheng, D. Du, and Y. Lin, “Smartphone-Based Sensors for On-Site Water Quality Monitoring,” 2022, pp. 331–348. doi: 10.1142/9789811245770_0011.

[158] A. U. Alam, D. Clyne, and M. J. Deen, “A Low-Cost Multi-Parameter Water Quality Monitoring System,” *Sensors*, vol. 21, no. 11, Art. no. 11, Jan. 2021, doi: 10.3390/s21113775.

[159] S.-W. Chiu and K.-T. Tang, “Towards a Chemiresistive Sensor-Integrated Electronic Nose: A Review,” *Sensors*, vol. 13, no. 10, Art. no. 10, Oct. 2013, doi: 10.3390/s131014214.

[160] H. G. Moon, Han, Soo Deok, Kang, Min-Gyu, Jang, Ho-Won, and Ryu, Kwang-Su, “Chemiresistive Sensor Array Based on Semiconducting Metal Oxides for Environmental Monitoring,” *J. Sens. Sci. Technol.*, vol. 23, no. 1, pp. 15–18, Jan. 2014, doi: 10.5369/JSST.2014.23.1.15.

[161] A. T. John, K. Murugappan, D. R. Nisbet, and A. Tricoli, “An Outlook of Recent Advances in Chemiresistive Sensor-Based Electronic Nose Systems for Food Quality and Environmental Monitoring,” *Sensors*, vol. 21, no. 7, Art. no. 7, Jan. 2021, doi: 10.3390/s21072271.

[162] M. Darestani-Farahani, P. R. Selvaganapathy, and P. Kruse, “Development of Solid-State Chemiresistive Devices for Simultaneous Detection of Nitrate, Nitrite and Ammonium Ions in Aqueous Solutions,” *ECS Meet. Abstr.*, vol. MA2022-01, no. 52, p. 2139, Jul. 2022, doi: 10.1149/MA2022-01522139mtgabs.

[163] A.-M. Nasture, E. I. Ionete, F. A. Lungu, S. I. Spiridon, and L. G. Patularu, “Water Quality Carbon Nanotube-Based Sensors Technological Barriers and Late Research Trends: A Bibliometric Analysis,” *Chemosensors*, vol. 10, no. 5, Art. no. 5, May 2022, doi: 10.3390/chemosensors10050161.

[164] S. B. Crary, “Optimal Design of Experiments for sensor calibration,” in *IEEE*, 1991, pp. 404–407.

[165] R. G. Brereton, “Multilevel Multifactor Designs for MultivariateCalibration,” *The Analyst*, vol. 122, no. 12, pp. 1521–1529, 1997, doi: 10.1039/a703654j.

[166] C. Osborne, “Statistical Calibration: A Review,” *Int. Stat. Rev. Rev. Int. Stat.*, vol. 59, no. 3, p. 309, 1991, doi: 10.2307/1403690.

[167] E. V. Thomas, “A Primer on Multivariate Calibration,” *Anal. Chem.*, vol. 66, no. 15, pp. 795–804, 1994.

[168] J. Tellinghuisen, “Inverse vs. classical calibration for small data sets,” *Fresenius J. Anal. Chem.*, vol. 368, no. 6, pp. 585–588, Nov. 2000, doi: 10.1007/s002160000556.

[169] B. Ouyang, “First-Principles Algorithm for Air Quality Electrochemical Gas Sensors,” *ACS Sens.*, vol. 5, no. 9, pp. 2742–2746, Sep. 2020, doi: 10.1021/acssensors.0c01129.

[170] A. Majid, B. Khadim, M. Alkhedher, S. Haider, and M. S. Akhtar, “Modeling of Inert Gas Sensors Using First Principles Methods,” *IEEE Sens. J.*, vol. 23, no. 16, pp. 18118–18124, Aug. 2023, doi: 10.1109/JSEN.2023.3283959.

[171] I. T. Cameron and K. Hangos, *Process Modelling and Model Analysis*. Elsevier, 2001.

[172] Y. Zhang, L. O. H. Wijeratne, S. Talebi, and D. J. Lary, “Machine Learning for Light Sensor Calibration,” *Sensors*, vol. 21, no. 18, Art. no. 18, Jan. 2021, doi: 10.3390/s21186259.

[173] S. Gupta, M. Mittal, and A. Padha, “Predictive Analytics of Sensor Data Based on Supervised Machine Learning Algorithms,” in *2017 International Conference on Next Generation Computing and Information Systems (ICNGCIS)*, Dec. 2017, pp. 171–176. doi: 10.1109/ICNGCIS.2017.12.

[174] W. Zhao, A. Bhushan, A. D. Santamaria, M. G. Simon, and C. E. Davis, “Machine Learning: A Crucial Tool for Sensor Design,” *Algorithms*, vol. 1, no. 2, Art. no. 2, Dec. 2008, doi: 10.3390/a1020130.

[175] L. Rajulapati, S. Chinta, B. Shyamala, and R. Rengaswamy, “Integration of machine learning and first principles models,” *AIChE J.*, vol. 68, no. 6, p. e17715, 2022, doi: 10.1002/aic.17715.

[176] M Thompson, “Is my calibration linear?,” *The Analyst*, vol. 119, no. 11, p. 2363, 1994, doi: 10.1039/an9941902363.

[177] Norman R. Draper and Harry Smith, *Applied Regression Analysis*. Wiley.

[178] R. G. Brereton, “Introduction to multivariate calibration in analytical chemistry,” *The Analyst*, vol. 125, no. 11, pp. 2125–2154, 2000, doi: 10.1039/b003805i.

[179] V. Centner *et al.*, “Comparison of Multivariate Calibration Techniques Applied to Experimental NIR Data Sets,” *Appl. Spectrosc.*, vol. 54, no. 4, pp. 608–623, Apr. 2000, doi: 10.1366/0003702001949816.

[180] R. E. Shaffer, S. L. Rose-Pehrsson, and R. A. McGill, “A comparison study of chemical sensor array pattern recognition algorithms,” *Anal. Chim. Acta*, vol. 384, no. 3, pp. 305–317, Apr. 1999, doi: 10.1016/S0003-2670(98)00780-6.

[181] E. V. Thomas and N. Ge, "Development of Robust Multivariate Calibration Models," *Technometrics*, vol. 42, no. 2, pp. 168–177, May 2000, doi: 10.1080/00401706.2000.10485996.

[182] J. H. Kalivas, "Multivariate calibration, an overview," *Anal. Lett.*, vol. 38, no. 14, pp. 2259–2279, 2005, doi: 10.1080/00032710500315904.

[183] M. Forina, S. Lanteri, and M. Casale, "Multivariate calibration," *J. Chromatogr. A*, vol. 1158, no. 1–2, 2007, doi: 10.1016/j.chroma.2007.03.082.

[184] A. Vergara, S. Vembu, T. Ayhan, M. A. Ryan, M. L. Homer, and R. Huerta, "Chemical gas sensor drift compensation using classifier ensembles," *Sens. Actuators B Chem.*, vol. 166–167, pp. 320–329, May 2012, doi: 10.1016/j.snb.2012.01.074.

[185] M. Khaydukova, V. Panchuk, D. Kirsanov, and A. Legin, "Multivariate Calibration Transfer between two Potentiometric Multisensor Systems," *Electroanalysis*, vol. 29, no. 9, pp. 2161–2166, 2017, doi: 10.1002/elan.201700190.

[186] V. Panchuk, V. Semenov, L. Lvova, and A. Legin, "Response Standardization for Drift Correction and Multivariate Calibration Transfer in 'Electronic Tongue' Studies," in *Biomimetic Sensing: Methods and Protocols*, vol. 2027, 2019, pp. 181–194.

[187] V. Panchuk *et al.*, "Extending electronic tongue calibration lifetime through mathematical drift correction: Case study of microcystin toxicity analysis in waters," *Sens. Actuators B Chem.*, vol. 237, pp. 962–968, 2016, doi: 10.1016/j.snb.2016.07.045.

[188] G. Korotcenkov and B. K. Cho, "Instability of metal oxide-based conductometric gas sensors and approaches to stability improvement (short survey)," *Sens. Actuators B Chem.*, vol. 156, no. 2, pp. 527–538, 2011, doi: 10.1016/j.snb.2011.02.024.

[189] N. A. Woody, R. N. Feudale, A. J. Myles, and S. D. Brown, "Transfer of Multivariate Calibrations between Four Near-Infrared Spectrometers Using Orthogonal Signal Correction," *Anal. Chem.*, vol. 76, no. 9, pp. 2595–2600, May 2004, doi: 10.1021/ac035382g.

[190] H. Meixner and U. Lampe, "Metal oxide sensors," *Sens. Actuators B Chem.*, vol. 33, no. 1–3, pp. 198–202, 1996, doi: 10.1016/0925-4005(96)80098-0.

[191] R. D. Marco *et al.*, "Evidence of a water layer in solid-contact polymeric ion sensors," *Phys. Chem. Chem. Phys.*, vol. 10, no. 1, pp. 73–76, 2008, doi: 10.1039/b714248j.

[192] O. E. de Noord, "Multivariate calibration standardization," *Chemom. Intell. Lab. Syst.*, vol. 25, no. 2, pp. 85–97, Nov. 1994, doi: 10.1016/0169-7439(94)85037-2.

[193] R. N. Feudale, N. A. Woody, H. Tan, A. J. Myles, S. D. Brown, and J. Ferré, "Transfer of multivariate calibration models: A review," *Chemom. Intell. Lab. Syst.*, vol. 64, no. 2, pp. 181–192, 2002, doi: 10.1016/S0169-7439(02)00085-0.

[194] F. A. Honorato, B. D. B. Neto, M. N. Martins, R. K. H. Galvão, and M. F. Pimentel, "Calibration transfer in multivariate methods," *Quimica Nova*, vol. 30, no. 5, pp. 1301–1312, 2007, doi: 10.1590/s0100-40422007000500044.

[195] Y. Mou, L. Zhou, S. Yu, W. Z. Chen, X. Zhao, and X. You, "Robust calibration model transfer," *Chemom. Intell. Lab. Syst.*, vol. 156, pp. 62–71, 2016, doi: 10.1016/j.chemolab.2016.05.014.

[196] J. J. Workman, “A Review of Calibration Transfer Practices and Instrument Differences in Spectroscopy,” *Appl. Spectrosc.*, vol. 72, no. 3, pp. 340–365, Mar. 2018, doi: 10.1177/0003702817736064.

[197] Y. Y. SHI, J. Y. LI, and X. L. CHU, “Progress and Applications of Multivariate Calibration Model Transfer Methods,” *Chin. J. Anal. Chem.*, vol. 47, no. 4, pp. 479–487, 2019, doi: 10.1016/S1872-2040(19)61152-7.

[198] M. Khaydukova, C. Medina-Plaza, M. L. Rodriguez-Mendez, V. Panchuk, D. Kirsanov, and A. Legin, “Multivariate calibration transfer between two different types of multisensor systems,” *Sens. Actuators B Chem.*, vol. 246, pp. 994–1000, 2017, doi: 10.1016/j.snb.2017.02.099.

[199] S. Deshmukh *et al.*, “Calibration transfer between electronic nose systems for rapid In situ measurement of pulp and paper industry emissions,” *Anal. Chim. Acta*, vol. 841, pp. 58–67, Sep. 2014, doi: 10.1016/j.aca.2014.05.054.

[200] K. Yan and D. Zhang, “Improving the transfer ability of prediction models for electronic noses,” *Sens. Actuators B Chem.*, vol. 220, pp. 115–124, 2015, doi: 10.1016/j.snb.2015.05.060.

[201] K. Yan and D. Zhang, “Calibration transfer and drift compensation of e-noses via coupled task learning,” *Sens. Actuators B Chem.*, vol. 225, pp. 288–297, Mar. 2016, doi: 10.1016/j.snb.2015.11.058.

[202] J. Fonollosa, L. Fernández, A. Gutiérrez-Gálvez, R. Huerta, and S. Marco, “Calibration transfer and drift counteraction in chemical sensor arrays using Direct Standardization,” *Sens. Actuators B Chem.*, vol. 236, pp. 1044–1053, Nov. 2016, doi: 10.1016/j.snb.2016.05.089.

[203] A. Rudnitskaya, A. M. S. Costa, and I. Delgadillo, “Calibration update strategies for an array of potentiometric chemical sensors,” *Sens. Actuators B Chem.*, vol. 238, pp. 1181–1189, Jan. 2017, doi: 10.1016/j.snb.2016.06.075.

[204] C. F. Pereira, M. F. Pimentel, R. K. H. Galvão, F. A. Honorato, L. Stragevitch, and M. N. Martins, “A comparative study of calibration transfer methods for determination of gasoline quality parameters in three different near infrared spectrometers,” *Anal. Chim. Acta*, vol. 611, no. 1, pp. 41–47, Mar. 2008, doi: 10.1016/j.aca.2008.01.071.

[205] B. Malli, A. Birlutiu, and T. Natschläger, “Standard-free calibration transfer - An evaluation of different techniques,” *Chemom. Intell. Lab. Syst.*, vol. 161, no. March 2016, pp. 49–60, 2017, doi: 10.1016/j.chemolab.2016.12.008.

[206] H. Martens and E. Stark, “Extended multiplicative signal correction and spectral interference subtraction: New preprocessing methods for near infrared spectroscopy,” *J. Pharm. Biomed. Anal.*, vol. 9, no. 8, pp. 625–635, Jan. 1991, doi: 10.1016/0731-7085(91)80188-F.

[207] A. Kohler, M. Zimonja, V. Segtnan, and H. Martens, “Standard Normal Variate, Multiplicative Signal Correction and Extended Multiplicative Signal Correction Preprocessing in Biospectroscopy,” in *Comprehensive Chemometrics*, Elsevier, 2009, pp. 139–162. doi: 10.1016/B978-044452701-1.00102-2.

[208] J. A. Westerhuis, S. de Jong, and A. K. Smilde, “Direct orthogonal signal correction,” *Chemom. Intell. Lab. Syst.*, vol. 56, no. 1, pp. 13–25, Apr. 2001, doi: 10.1016/S0169-7439(01)00102-2.

[209] M. Padilla, A. Perera, I. Montoliu, A. Chaudry, K. Persaud, and S. Marco, “Drift compensation of gas sensor array data by Orthogonal Signal Correction,” *Chemom. Intell. Lab. Syst.*, p. 8, 2010.

[210] R. Laref, D. Ahmadou, E. Losson, and M. Siadat, “Orthogonal Signal Correction to Improve Stability Regression Model in Gas Sensor Systems,” vol. 2017, 2017, doi: https://doi.org/10.1155/2017/9851406.

[211] B. Walczak, E. Bouveresse, and D. L. Massart, “Standardization of near-infrared spectra in the wavelet domain,” *Chemom. Intell. Lab. Syst.*, vol. 36, no. 1, pp. 41–51, 1997, doi: 10.1016/S0169-7439(96)00075-5.

[212] W. Du *et al.*, “Maintaining the predictive abilities of multivariate calibration models by spectral space transformation,” *Anal. Chim. Acta*, vol. 690, no. 1, pp. 64–70, 2011, doi: 10.1016/j.aca.2011.02.014.

[213] C. Zou *et al.*, “Analytical Methods dynamic time warping for near-infrared,” pp. 4481–4493, 2019, doi: 10.1039/c9ay01139k.

[214] B. Lebental *et al.*, “Carbon nanotube sensor array for water monitoring with conjugated polymers,” in *C’NANO 2017, The Nanoscience Meeting*, Lyon, France, Dec. 2017, p. 18p. Accessed: Mar. 14, 2024. [Online]. Available: https://hal.science/hal-01811007

[215] “AdaPtive micROfluidic- and nano-enabled smart systems for waTEr qUality Sensing | PROTEUS Project | Fact Sheet | H2020,” CORDIS | European Commission. Accessed: Apr. 01, 2024. [Online]. Available: https://cordis.europa.eu/project/id/644852

[216] A. Kaliyaraj Selva Kumar, Y. Zhang, D. Li, and R. G. Compton, “A mini-review: How reliable is the drop casting technique?,” *Electrochem. Commun.*, vol. 121, p. 106867, Dec. 2020, doi: 10.1016/j.elecom.2020.106867.

[217] “Dimatix Materials Printer DMP-2850 | Fujifilm [United States].” Accessed: Mar. 14, 2024. [Online]. Available: https://www.fujifilm.com/us/en/business/inkjet-solutions/inkjet-technology-integration/dmp-2850

[218] K. Bock, *Heterointegration Technologies for System in a foil*. 2008.

[219] H. Saada *et al.*, “Photoelectrochemical Sensing of Hydrogen Peroxide on Hematite,” in *ChemElectroChem*, Mar. 2020, pp. 1155–1159. doi: 10.1002/celc.202000028.

[220] “Portable meter pH-Conductivity 3m ODEON 2 inputs.” Accessed: Mar. 27, 2024. [Online]. Available: https://en.aqualabo.fr/set-odeon-open-x-pheht-3-c4e-3-b3624.html

[221] “Sense-City.” Accessed: Dec. 04, 2023. [Online]. Available: https://sense-city.ifsttar.fr/en/

[222] G. Perrin and B. Lebental, “Uncertainty-Based Calibration Method for Environmental Sensors—Application to Chlorine and pH Monitoring With Carbon Nanotube Sensor Array,” *IEEE Sens. J.*, vol. 23, no. 5, pp. 5146–5155, Mar. 2023, doi: 10.1109/JSEN.2023.3238900.

[223] F. Michelis, L. Bodelot, Y. Bonnassieux, and B. Lebental, “Highly reproducible, hysteresis-free, flexible strain sensors by inkjet printing of carbon nanotubes,” *Carbon*, vol. 95, pp. 1020–1026, Dec. 2015, doi: 10.1016/j.carbon.2015.08.103.

[224] J. R. Taylor, *Introduction To Error Analysis: The Study of Uncertainties in Physical Measurements*. University Science Books, 1997.

[225] N. Trajchevski, M. Tomov, and V. Sarac, “An Approach of Empirical Model Uncertainty Estimation During Physical Quantities Investigation in Turning,” *J. Adv. Manuf. Syst.*, vol. 21, no. 04, pp. 851–868, Dec. 2022, doi: 10.1142/S0219686722500329.

[226] National Aeronautics and Space Administration, *Measurement Uncertainty Analysis Principles and Methods*. in NASA Measurement Quality Assurance Handbook. Accessed: Apr. 10, 2024. [Online]. Available: https://standards.nasa.gov/sites/default/files/standards/NASA/Baseline/0/nasa-hdbk-873919-3.pdf

[227] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. New York: Chapman and Hall/CRC, 1994. doi: 10.1201/9780429246593.

[228] “Aqualabo Photopod.” Accessed: Apr. 03, 2024. [Online]. Available: https://en.aqualabo.fr/-b3314.html

[229] “Circulating Water bath, 5 Ltr. Vertical, -20C to 100C WB 2100 V in India | Biomall.” Accessed: Apr. 04, 2024. [Online]. Available: https://www.biomall.in/product/circulating-water-bath-5-ltr-vertical-20c-to-100c-wb-2100-v

[230] R. N. Feudale, N. A. Woody, H. Tan, A. J. Myles, S. D. Brown, and J. Ferré, “Transfer of multivariate calibration models: A review,” *Chemom. Intell. Lab. Syst.*, vol. 64, no. 2, pp. 181–192, 2002, doi: 10.1016/S0169-7439(02)00085-0.

[231] J. J. Workman, “A Review of Calibration Transfer Practices and Instrument Differences in Spectroscopy,” *Appl. Spectrosc.*, vol. 72, no. 3, pp. 340–365, 2018, doi: 10.1177/0003702817736064.

[232] S. Deshmukh *et al.*, “Calibration transfer between electronic nose systems for rapid In situ measurement of pulp and paper industry emissions,” *Anal. Chim. Acta*, vol. 841, pp. 58–67, 2014, doi: 10.1016/j.aca.2014.05.054.

[233] K. Yan and D. Zhang, “Calibration transfer and drift compensation of e-noses via coupled task learning,” *Sens. Actuators B Chem.*, vol. 225, pp. 288–297, 2016, doi: 10.1016/j.snb.2015.11.058.

[234] J. Fonollosa, L. Fernández, A. Gutiérrez-Gálvez, R. Huerta, and S. Marco, “Calibration transfer and drift counteraction in chemical sensor arrays using Direct Standardization,” *Sens. Actuators B Chem.*, vol. 236, pp. 1044–1053, 2016, doi: 10.1016/j.snb.2016.05.089.

[235] M. Khaydukova, V. Panchuk, D. Kirsanov, and A. Legin, “Multivariate Calibration Transfer between two Potentiometric Multisensor Systems,” *Electroanalysis*, vol. 29, no. 9, pp. 2161–2166, 2017, doi: 10.1002/elan.201700190.

[236] A. Rudnitskaya, A. M. S. Costa, and I. Delgadillo, “Calibration update strategies for an array of potentiometric chemical sensors,” *Sens. Actuators B Chem.*, vol. 238, pp. 1181–1189, 2017, doi: 10.1016/j.snb.2016.06.075.

[237] V. Panchuk, V. Semenov, L. Lvova, and A. Legin, “Response Standardization for Drift Correction and Multivariate Calibration Transfer in ‘Electronic Tongue’ Studies,” *Biomim. Sens. Methods Protoc.*, vol. 2027, pp. 181–194, 2019, doi: https://doi.org/10.1007/978-1-4939-9616-2_15.

[238] V. Panchuk, D. Kirsanov, E. Oleneva, V. Semenov, and A. Legin, “Calibration transfer between different analytical methods,” *Talanta*, vol. 170, pp. 457–463, Aug. 2017, doi: 10.1016/j.talanta.2017.04.039.

[239] L. Nørgaard, “Direct standardisation in multi wavelength fluorescence spectroscopy,” *Chemom. Intell. Lab. Syst.*, vol. 29, no. 2, pp. 283–293, 1995, doi: 10.1016/0169-7439(95)80103-G.

[240] F. Sales, M. P. Callao, and F. X. Rius, “Multivariate standardization techniques on ion-selective sensor arrays,” *Analyst*, vol. 124, no. 7, pp. 1045–1051, 1999, doi: 10.1039/a902585e.

[241] M. Forina *et al.*, “Transfer of calibration function in near-infrared spectroscopy,” *Chemom. Intell. Lab. Syst.*, vol. 27, no. 2, pp. 189–203, 1995, doi: 10.1016/0169-7439(95)80023-3.

[242] K. S. Park, Y. H. Ko, H. Lee, C. H. Jun, H. Chung, and M. S. Ku, “Near-infrared spectral data transfer using independent standardization samples: A case study on the trans-alkylation process,” *Chemom. Intell. Lab. Syst.*, vol. 55, no. 1–2, pp. 53–65, 2001, doi: 10.1016/S0169-7439(00)00115-5.

[243] F. A. Honorato, R. K. H. Galvão, M. F. Pimentel, B. D. B. Neto, M. C. U. Araújo, and F. R. D. Carvalho, “Robust modeling for multivariate calibration transfer by the successive projections algorithm,” *Chemom. Intell. Lab. Syst.*, vol. 76, no. 1, pp. 65–72, 2005, doi: 10.1016/j.chemolab.2004.09.006.

[244] C. Goodall, “Procrustes Methods in the Statistical Analysis of Shape,” *J. R. Stat. Soc. Ser. B Methodol.*, vol. 53, no. 2, pp. 285–321, 1991, doi: 10.1111/j.2517-6161.1991.tb01825.x.

[245] J. M. Andrade, M. P. Gómez-Carracedo, W. Krzanowski, and M. Kubista, “Procrustes rotation in analytical chemistry, a tutorial,” *Chemom. Intell. Lab. Syst.*, vol. 72, no. 2, pp. 123–132, 2004, doi: 10.1016/j.chemolab.2004.01.007.

[246] C. J. Sergeant, E. N. Starkey, K. K. Bartz, M. H. Wilson, and F. J. Mueter, “A practitioner’s guide for exploring water quality patterns using principal components analysis and Procrustes,” *Environ. Monit. Assess.*, vol. 188, no. 4, p. 249, Mar. 2016, doi: 10.1007/s10661-016-5253-z.

[247] A. Rudnitskaya, “Calibration Update and Drift Correction for Electronic Noses and Tongues,” *Front. Chem.*, vol. 6, p. 433, Sep. 2018, doi: 10.3389/fchem.2018.00433.

[248] B. Malli, T. Natschläger, M. Pawliczek, T. Reischer, and W. Kantner, “Application-oriented Standard-free Methods for Calibration Transfer,” *Lenzing. Berichte*, vol. 92, pp. 33–46, 2015.

[249] M. Forina *et al.*, “Transfer of calibration function in near-infrared spectroscopy,” *Chemom. Intell. Lab. Syst.*, vol. 27, no. 2, pp. 189–203, 1995, doi: 10.1016/0169-7439(95)80023-3.

[250] Wegelin, Jacob A, “A Survey of Partial Least Squares (PLS) Methods, with Emphasis on the Two-Block Case,” University of Washinton, Technical Report 371, Mar. 2000. Accessed: Feb. 09, 2024. [Online]. Available: https://stat.uw.edu/research/tech-reports/survey-partial-least-squares-pls-methods-emphasis-two-block-case

[251] "Python - Procrustes Analysis with NumPy? - Stack Overflow." Accessed: Mar. 25, 2024. [Online]. Available: https://stackoverflow.com/questions/18925181/procrustes-analysis-with-numpy/18927641#18927641

[252] J. Fonollosa, E. Neftci, R. Huerta, and S. Marco, "Evaluation of calibration transfer strategies between Metal Oxide gas sensor arrays," *Procedia Eng.*, vol. 120, pp. 261–264, 2015, doi: 10.1016/j.proeng.2015.08.601.

[253] P. K. Clifford and D. T. Tuma, "Characteristics of semiconductor gas sensors I. Steady state gas response," *Sens. Actuators*, vol. 3, pp. 233–254, 1983.

[254] P. K. Clifford and D. T. Tuma, "Characteristics of semiconductor gas sensors. II. Transient response to temperature change," *Sens. Actuators*, vol. 3, pp. 255–281, 1983.

[255] J. Fonollosa, I. Rodríguez-Luján, and R. Huerta, "Chemical gas sensor array dataset," *Data Brief*, vol. 3, pp. 85–89, Jun. 2015, doi: 10.1016/j.dib.2015.01.003.

[256] A. Vergara, S. Vembu, T. Ayhan, M. A. Ryan, and M. L. Homer, "Chemical Chemical gas sensor drift compensation using classifier ensembles," *Sens. Actuators B Chem.*, vol. 166–167, pp. 320–329, 2012, doi: 10.1016/j.snb.2012.01.074.

[257] I. Rodriguez-Lujan, J. Fonollosa, A. Vergara, M. Homer, and R. Huerta, "On the calibration of sensor arrays for pattern recognition using the minimal number of experiments," *Chemom. Intell. Lab. Syst.*, vol. 130, pp. 123–134, 2014, doi: 10.1016/j.chemolab.2013.10.012.

[258] B. B. Ngoune *et al.*, "Comparison of calibration strategies for a high sensitivity PEI-based RF humidity sensor," *IEEE Sens. J.*, pp. 1–1, 2024, doi: 10.1109/JSEN.2024.3367594.

[259] B. Lebental, "Sense-City, Smart technologies for sustainable cities," in *CoE Technology Domain Workshop on Urban Solutions & Ressource Resilience*, Singapore, France, Mar. 2016, p. 32p. Accessed: Nov. 14, 2023. [Online]. Available: https://hal.science/hal-01810251

Appendix A

The dataset for case study 1 has 8 columns corresponding to each sensing unit (MOX sensor) and 10 rows corresponding to the 10 concentrations of Carbon Monoxide gas.

For the data upsampling, the first step is to estimate the deterministic model parameters and information for the noise model. For this case (for MOX sensors), the Clifford-Tuma model is chosen as the deterministic model. From the raw dataset, the sensor baseline response is estimated (which corresponds to the average of the sensor response when the sensor array B1-B5 is exposed to clean air before introducing the analyte). This baseline is then used to normalize the actual sensor response by the formulae:

$$R_s' = \frac{R_0 - R_s}{R_0}$$

The Clifford-Tuma model for the normalized sensor response R_s', and gas concentration c , is given by:

$$\log(R_s') = \log(s) + \beta \log(c)$$

Where s, β are sensor dependent parameters.

This equation is a linear regression with 2 unknowns; a minimum of 2 data points is necessary to estimate the sensor parameters. In our case study, for each MOX sensor (for each column), we randomly selected 3 data points from the 10 data points available and used them to estimate the parameters s and β. This process is repeated 100 times, each time having different random data points (there are 120 ways of selecting 3 data points from a total of 10), and the parameters for each set of data points are recorded. From this collection, the average and standard deviation of each parameter (s, β) for each MOX sensor is estimated. This information is then used with a Gaussian model (where the mean and standard deviation of the model are the mean and standard deviation of the parameters estimated in the previous step) to generate sensor response for new gas concentrations. These gas concentrations are selected as midpoints of the already available gas concentrations.

Appendix B

The tables below correspond to the R^2 and MAE values for each transfer technique for each case study, as detailed in Chapter 5. In all the case studies, we observe that as transfer standards increase, the Mean Absolute Error (MAE) decreases while the R^2 value correspondingly increases. In some instances, the MAE plateaus and does not further decrease with additional transfer standards. This can be attributed to the fact that the extra data does not provide significant information to enhance the model.

Case Study I

Table B. 1 Results of application of calibration transfer techniques and recalibration for the same transfer standards for dataset 1

		LR		*KNN*	
		MAE	R^2	*MAE*	R^2
Direct Standardization	20% - 3	12 (9)	0.86 (0.25)	11 (6.4)	0.91 (0.14)
	30% - 5	1.3 (0.6)	0.99	5.5 (1.2)	0.98 (0.01)
	40% - 7	0.3 (0.4)	0.99	3.6 (1.6)	0.99
	50% - 9	0.07 (0.04)	0.99	2.3 (0.9)	0.99
	60% - 11	0.04 (0.02)	1	1.9 (0.4)	1
SSS + LR	20% - 3	1193	-0.6	42	-0.4
	30% - 5	948	-0.4	38 (3.7)	0.2 (0.2)
	40% - 7	872	-0.3	37 (3.1)	0.2 (0.1)
	50% - 9	847	-0.2	35 (2.2)	0.2 (0.2)
	60% - 11	790	-0.2	35 (2)	0.2 (0.1)
SSS + SVR	20% - 3	92	-1	57	-1
	30% - 5	163	-1	48	-1
	40% - 7	230	-1	43	-1
	50% - 9	237	-1	39	-1
	60% - 11	193	-1	37	-1
PLS (4 components) + DS	20% - 3	43	-1	33 (4)	0.4 (0.4)
	30% - 5	11 (9.0)	0.9 (0.3)	13 (8)	0.8 (0.5)
	40% - 7	8.8 (2.5)	0.96 (0.02)	13 (3)	0.9 (0.05)
	50% - 9	7.8 (1.2)	0.96 (0.01)	11 (2)	0.9 (0.03)

		LR		*KNN*	
		MAE	R^2	*MAE*	R^2
	60% - 11	7.9 (1.2)	0.97 (0.01)	12 (2)	0.9 (0.02)
PLS (8 components) + DS	20% - 3	43	-1	33 (4)	0.45 (0.4)
	30% - 5	8.3 (6.6)	0.9 (0.3)	15 (12)	0.71 (0.7)
	40% - 7	0.89 (0.68)	0.99	3.5 (1.6)	0.99
	50% - 9	0.1 (0.08)	0.99	1.7 (0.6)	0.99
	60% - 11	0.06 (0.04)	1.0	1.3 (0.3)	1.0
PCA (4) + DS	30% - 5	2.1 (0.7)	0.99	8.3 (2.9)	0.96 (0.03)
	40% - 7	1.8 (0.5)	0.99	8.4 (1.7)	0.97 (0.01)
	50% - 9	1.6 (0.5)	0.99	8.0 (0.9)	0.97 (0.01)
	60% - 11	1.5 (0.1)	1.0	8.1 (0.8)	0.97 (0.01)
PCA (8) + DS	50% - 9	0.1 (0.08)	0.99	1.7 (0.6)	0.99
	60% - 11	0.06 (0.04)	1.0	1.3 (0.3)	1.0
Procrustes transform	20% - 3	3028	-0.4	24 (5)	0.7 (0.3)
	30% - 5	172	-0.07	23 (2)	0.7 (0.1)
	40% - 7	149	0.09	23 (2)	0.7 (0.1)
	50% - 9	153	0.09	22 (2)	0.8 (0.1)
	60% - 11	145 (35)	0.15 (0.2)	23 (1)	0.7 (0.1)
PCA (8) + Procrustes	50% - 9	153 (48)	0.1 (0.2)	22 (2)	0.8 (0.1)
	60% - 11	145 (35)	0.2 (0.2)	23 (1)	0.8 (0.1)
SBC	20% - 3	37 (18)	0.3 (1.2)	33 (9)	-0.15
	30% - 5	26 (3)	0.7 (0.1)	28 (5)	0.7 (0.2)
	40% - 7	25 (2)	0.7 (0.1)	26 (3)	0.7 (0.1)
	50% - 9	25 (1)	0.7 (0.1)	25 (2)	0.7 (0.05)
	60% - 11	24 (1)	0.8 (0.1)	24 (2)	0.8 (0.06)
AMC	50% - 9	882	-0.19	55	-1
	60% - 11	875	-0.17	57	-1
MMC	50% - 9	879	-0.12	41 (3)	0.3 (0.1)
	60% - 11	863	-0.09	41 (3)	0.3 (0.1)
Recalibration	20% - 3	19 (14)	0.7 (0.7)	43	-1
	30% - 5	1.5 (0.6)	0.99	27 (5)	0.3 (0.5)

		LR		*KNN*	
		MAE	R^2	*MAE*	R^2
	40% - 7	0.4 (0.4)	0.99	21 (5)	0.6 (0.3)
	50% - 9	0.1 (0.2)	0.99	16 (4)	0.8 (0.2)
	60% - 11	0.01 (0.01)	1.0	11 (3)	0.9 (0.1)

Case Study II

Table B. 2 Results of application of calibration transfer techniques and recalibration for the same transfer standards for dataset 2

		Linear Regression		*KNN Regression*	
		MAE	R^2	*MAE*	R^2
Direct Standardization	3 targets	56 (26)	0.4 (0.3)	43 (18)	0.4 (0.4)
	4 targets	34 (10)	0.6 (0.2)	35 (14)	0.6 (0.3)
	5 targets	27 (6)	0.8 (0.1)	33 (10)	0.6 (0.2)
	6 targets	22 (4)	0.8 (0.1)	32 (6)	0.6 (0.1)
SSS + LR	3 targets	65	-1	45	-1
	4 targets	60	-0.4	41 (3)	0.1 (0.5)
	5 targets	59 (6)	0.3 (0.3)	41 (2)	0.2 (0.3)
	6 targets	59 (6)	0.4 (0.2)	40 (1)	0.3 (0.2)
SSS + SVR	3 targets	112	-1	44	-1
	4 targets	117	-1	40 (3)	0.2 (0.3)
	5 targets	97	-1	40 (2)	0.3 (0.2)
	6 targets	98	-1	39 (1)	0.3 (0.1)
PLS (8 components) + DS	3 targets	33 (12)	0.6 (0.4)	32 (13)	0.4 (0.8)
	4 targets	25 (7)	0.8 (0.2)	26 (9)	0.7 (0.3)
	5 targets	21 (5)	0.8 (0.1)	24 (6)	0.7 (0.2)
	6 targets	20 (3)	0.8 (0.1)	24 (5)	0.7 (0.1)
PLS (16 components) + DS	3 targets	60 (32)	0.4 (0.3)	44 (20)	0.4 (0.5)
	4 targets	36 (16)	0.6 (0.2)	36 (15)	0.5 (0.4)
	5 targets	27 (6)	0.8 (0.1)	33 (9)	0.6 (0.2)
	6 targets	22 (4)	0.8 (0.1)	30 (6)	0.6 (0.1)
PCA (8 components) + DS	3 targets	31 (13)	0.6 (0.3)	31 (15)	0.5 (0.9)

		Linear Regression		*KNN Regression*	
		MAE	R^2	*MAE*	R^2
	4 targets	26 (13)	0.7 (0.2)	27 (10)	0.7 (0.2)
	5 targets	21 (4)	0.8 (0.1)	24 (6)	0.7 (0.1)
	6 targets	20 (3)	0.8 (0.1)	24 (4)	0.7 (0.1)
PCA (16 components) + DS	3 targets	60 (32)	0.4 (0.3)	44 (20)	0.4 (0.5)
	4 targets	36 (16)	0.6 (0.2)	36 (15)	0.5 (0.4)
	5 targets	27 (6)	0.8 (0.1)	33 (9)	0.6 (0.2)
	6 targets	22 (4)	0.8 (0.1)	30 (6)	0.7 (0.1)
Procrustes transform	3 targets	88	-1	58	-1
	4 targets	95	-0.8	54	-1
	5 targets	91	-0.3	52	-1
	6 targets	89	-0.06	52	-0.9
PCA (8) + Procrustes	3 targets	78	-1	58	-1
	4 targets	77	-1	54	-1
	5 targets	70	-0.7	52	-1
	6 targets	66	-0.4	52	-0.9
PCA (16) + Procrustes	3 targets	88	-1	58	-1
	4 targets	95	-0.8	54	-1
	5 targets	91	-0.3	52	-1
	6 targets	89	-0.1	52	-0.9
SBC	3 targets	46	-1	64	-1
	4 targets	43	-0.4	60	-1
	5 targets	42 (4)	0.1	58	-0.9
	6 targets	41 (3)	0.1	56	-0.7
AMC	3 targets	126 (8)	0.4 (0.1)	54 (2)	0.2 (0.1)
	4 targets	124 (4)	0.4 (0.04)	53 (1)	0.3 (0.03)
	5 targets	122 (2)	0.4 (0.02)	53 (1)	0.3 (0.03)
	6 targets	122 (1)	0.4 (0.01)	52 (1)	0.3 (0.02)
MMC	3 targets	108 (32)	0.2 (0.2)	56 (4)	0.4 (0.1)
	4 targets	108 (36)	0.2 (0.2)	56 (3)	0.4 (0.1)
	5 targets	104 (22)	0.2 (0.1)	56 (3)	0.4 (0.1)

		Linear Regression		*KNN Regression*	
		MAE	R^2	*MAE*	R^2
	6 targets	101 (19)	0.2 (0.1)	56 (3)	0.4 (0.1)
Recalibration	3 targets	36 (16)	0.6 (0.3)	37	-0.4
	4 targets	24 (8)	0.8 (0.1)	26 (6)	0.6 (0.3)
	5 targets	20 (5)	0.9 (0.1)	20 (4)	0.8 (0.1)
	6 targets	17 (4)	0.9 (0.03)	15 (2)	0.9 (0.02)

Case Study III

Table B. 3 Results of application of calibration transfer techniques and recalibration for the same transfer standards for dataset 3

		Linear Regression		*KNN Regression*	
		MAE	R^2	*MAE*	R^2
Direct Standardization	5% - 14	7.6 (1.2)	0.7 (0.05)	6.3 (0.6)	0.2 (0.3)
	10 % - 29	7.3 (0.2)	0.7 (0.03)	6.2 (0.5)	0.2 (0.2)
	20% - 58	7.2 (0.4)	0.7 (0.02)	6.2 (0.3)	0.2 (0.1)
	30% - 87	7.3 (0.3)	0.7 (0.02)	6.1 (0.3)	0.2 (0.1)
	40% - 116	7.3 (0.3)	0.7 (0.02)	6.1 (0.3)	0.3 (0.1)
	50% - 145	7.3 (0.2)	0.7 (0.02)	6.1 (0.3)	0.3 (0.1)
	60% - 174	7.3 (0.2)	0.7 (0.01)	6.1 (0.2)	0.3 (0.1)
SSS + LR	5% - 14	8.2 (1.2)	0.7 (0.04)	5.7 (0.5)	0.5 (0.2)
	10 % - 29	7.9 (0.6)	0.7 (0.03)	5.7 (0.3)	0.5 (0.1)
	20% - 58	7.8 (0.4)	0.7 (0.01)	5.6 (0.2)	0.5 (0.1)
	30% - 87	7.8 (0.3)	0.7 (0.01)	5.6 (0.2)	0.5 (0.1)
	40% - 116	7.8 (0.2)	0.7 (0.01)	5.6 (0.2)	0.5 (0.05)
	50% - 145	7.8 (0.2)	0.7 (0.01)	5.6 (0.1)	0.5 (0.04)
	60% - 174	7.8 (0.2)	0.7 (0.01)	5.2 (0.0)	0.6 (0.01)
SSS + SVR	5% - 14	8.1 (0.7)	0.7 (0.03)	5.4 (0.4)	0.6 (0.2)
	10 % - 29	8.1 (0.5)	0.7 (0.02)	5.3 (0.2)	0.6 (0.1)
	20% - 58	8.0 (0.3)	0.7 (0.01)	5.3 (0.1)	0.6 (0.0)

		Linear Regression		*KNN Regression*	
		MAE	R^2	*MAE*	R^2
	30% - 87	8.0 (0.2)	0.7 (0.01)	5.3 (0.1)	0.6 (0.0)
	40% - 116	8.0 (0.2)	0.7 (0.01)	5.3 (0.1)	0.6 (0.0)
	50% - 145	8.0 (0.2)	0.7 (0.01)	5.2 (0.0)	0.6 (0.0)
	60% - 174	8.0 (0.2)	0.7 (0.01)	5.2 (0.0)	0.6 (0.0)
PLS (2 components) + DS	5% - 14	8.9 (1.1)	0.6 (0.03)	6.1 (0.5)	0.3 (0.3)
	10 % - 29	8.6 (0.6)	0.6 (0.02)	6.1 (0.3)	0.3 (0.2)
	20% - 58	8.6 (0.4)	0.6 (0.01)	6.0 (0.2)	0.3 (0.1)
	30% - 87	8.5 (0.3)	0.6 (0.01)	6.0 (0.2)	0.3 (0.1)
	40% - 116	8.5 (0.3)	0.6 (0.01)	6.0 (0.2)	0.3 (0.1)
	50% - 145	8.6 (0.3)	0.6 (0.01)	5.9 (0.1)	0.3 (0.1)
	60% - 174	8.6 (0.2)	0.6 (0.01)	5.9 (0.1)	0.3 (0.01)
PCA (2 components) + DS	5% - 14	683	-0.15	29	-1
	10 % - 29	672	-0.15	30	-1
	20% - 58	670	-0.15	31	-1
	30% - 87	672	-0.15	31	-1
	40% - 116	674	-0.15	31	-1
	50% - 145	678	-0.15	31	-1
	60% - 174	678	-0.15	31	-1
Procrustes transform	5% - 14	8.5 (2.7)	0.7 (0.1)	5.8 (0.4)	0.5 (0.1)
	10 % - 29	8.0 (2.2)	0.7 (0.1)	5.8 (0.3)	0.5 (0.1)
	20% - 58	7.9 (2.0)	0.7 (0.1)	5.7 (0.2)	0.5 (0.1)
	30% - 87	7.5 (1.8)	0.7 (0.1)	5.7 (0.2)	0.5 (0.1)
	40% - 116	7.4 (1.7)	0.7 (0.1)	5.7 (0.1)	0.5 (0.1)
	50% - 145	7.4 (1.7)	0.7 (0.1)	5.7 (0.1)	0.5 (0.1)
	60% - 174	7.3 (1.7)	0.7 (0.1)	5.7 (0.1)	0.5 (0.1)
PCA (2) + Procrustes	5% - 14	8.5 (2.7)	0.7 (0.1)	5.8 (0.4)	0.5 (0.1)
	10 % - 29	8.0 (2.2)	0.7 (0.1)	5.8 (0.3)	0.5 (0.1)
	20% - 58	7.9 (2.0)	0.7 (0.1)	5.7 (0.2)	0.5 (0.1)
	30% - 87	7.5 (1.8)	0.7 (0.1)	5.7 (0.2)	0.5 (0.1)
	40% - 116	7.4 (1.7)	0.7 (0.1)	5.7 (0.1)	0.5 (0.1)

		Linear Regression		*KNN Regression*	
		MAE	R^2	*MAE*	R^2
	50% - 145	7.4 (1.7)	0.7 (0.1)	5.7 (0.1)	0.5 (0.1)
	60% - 174	7.3 (1.7)	0.7 (0.1)	5.7 (0.1)	0.5 (0.1)
SBC	5% - 14	8.4 (1.2)	0.7 (0.04)	5.6 (0.7)	0.6 (0.5)
	10 % - 29	8.0 (0.6)	0.7 (0.02)	5.3 (0.3)	0.7 (0.1)
	20% - 58	7.9 (0.4)	0.7 (0.02)	5.2 (0.2)	0.7 (0.05)
	30% - 87	7.9 (0.3)	0.7 (0.01)	5.1 (0.1)	0.7 (0.04)
	40% - 116	7.9 (0.2)	0.7 (0.01)	5.1 (0.1)	0.7 (0.03)
	50% - 145	7.9 (0.2)	0.7 (0.01)	5.1 (0.1)	0.7 (0.03)
	60% - 174	7.9 (0.2)	0.7 (0.01)	5.1 (0.1)	0.7 (0.02)
AMC	5% - 14	8.8	-1	10.1	-1
	10 % - 29	8.8	-1	10.1	-1
	20% - 58	8.8	-1	10.1	-1
	30% - 87	8.8	-1	10.1	-1
	40% - 116	8.8	-1	10.1	-1
	50% - 145	8.8	-1	10.1	-1
	60% - 174	8.8	-1	10.1	-1
MMC	5% - 14	7.4 (1.1)	0.7 (0.05)	6.4 (0.2)	0.2 (0.03)
	10 % - 29	7.2 (0.8)	0.7 (0.04)	6.3 (0.1)	0.2 (0.02)
	20% - 58	7.1 (0.5)	0.7 (0.03)	6.3 (0.1)	0.2 (0.01)
	30% - 87	7.0 (0.4)	0.7 (0.02)	6.3 (0.06)	0.2 (0.01)
	40% - 116	7.0 (0.3)	0.7 (0.02)	6.3 (0.04)	0.2 (0.01)
	50% - 145	6.9 (0.2)	0.7 (0.01)	6.3 (0.03)	0.2 (0.01)
	60% - 174	6.9 (0.2)	0.7 (0.01)	6.3 (0.03)	0.2 (0.01)
Recalibration	5% - 14	7.0 (1.3)	0.7 (0.1)	6.4 (0.9)	0.2 (0.4)
	10 % - 29	6.7 (0.6)	0.8 (0.03)	5.7 (0.6)	0.4 (0.2)
	20% - 58	6.6 (0.4)	0.8 (0.02)	5.3 (0.3)	0.6 (0.1)
	30% - 87	6.6 (0.3)	0.8 (0.02)	5.0 (0.3)	0.6 (0.04)
	40% - 116	6.7 (0.2)	0.8 (0.02)	4.8 (0.2)	0.6 (0.03)
	50% - 145	6.7 (0.2)	0.8 (0.02)	4.6 (0.2)	0.6 (0.03)
	60% - 174	6.7 (0.2)	0.8 (0.01)	4.5 (0.1)	0.7 (0.02)

The tables below provide an overview of the results from the above tables. In the table, for each instance, the '--' denote poor performance, as evident by negative R^2 or MAE higher than naïve MAE; '-' for the cases where R^2 is positive and MAE is lower than naïve MAE, but is 30% more than recalibration MAE. Similarly for the cases where R^2 is positive and MAE is within the tolerance range (MAE ≤ 1.3*Recalibration MAE), are indicated by '+' and for MAE ≤ 1.3*Primary MAE by '++'. The '*' indicates the instances where the MAE ≤ 0.7*Recalibration MAE, as in calibration transfer is better alternative to recalibration.

Case Study I

Table B. 4 Synthesis of results for the dataset from Case Study I, at different transfer standards ratio (20% to 60%)

Transfer techniques	Linear Regression model					KNN model				
	20%	30%	40%	50%	60%	20%	30%	40%	50%	60%
Secondary sensor with primary calibration	--	--	--	--	--	--	--	--	--	--
DS	++*	*	*	*	++	++*	*	*	*	*
SSS+ LR	--	--	--	--	--	--	-	-	-	-
SSS + SVR	--	--	--	--	--	--	--	--	--	--
PLS(8) + DS	++	-	-	-	-	++*	*	*	*	*
PC(8) + DS	++	--	--	+	-	-	--	--	++*	++*
PT	--	--	-	-	-	+	+	+	-	-
PC(8) + PT	--	--	--	-	-	+	--	--	-	-
SBC	-	-	-	-	-	+	+	+	-	-
AMC	--	--	--	--	--	--	--	--	--	--
MMC	--	--	--	--	--	--	--	--	-	-
Recalibration with TS	+	+	+	+	+	+	+	+	+	+

Case Study II

Table B. 5 Synthesis of results for the dataset from Case Study II, at different counts of concentration levels (from 3 to 6 distinct concentration levels)

Transfer techniques	LR Model				KNN Model			
	3	4	5	6	3	4	5	6
Secondary sensor with primary calibration	--	--	--	--	--	--	--	--
DS	-	-	-	+	+	+	-	-
SSS+ LR	--	--	-	-	--	-	-	-
SSS + SVR	--	--	--	--	--	-	-	-
PLS(8) + DS	+	+	+	+	+	+	+	-
PC(8) + DS	+	+	+	+	+	+	+	-
PT	--	--	--	--	--	--	--	--
PC(8) + PT	--	--	--	--	--	--	--	--
SBC	--	--	-	-	--	--	--	--
AMC	-	-	-	-	-	-	-	-
MMC	-	-	-	-	-	-	-	-
Recalibration with TS	+	+	+	+	--	+	+	+

Case Study III

Table B. 6 Synthesis of results for the dataset from Case Study III, at different transfer standards ratio (5% to 60%)

Transfer techniques	LR Model							KNN Model						
	5 %	10 %	20 %	30 %	40 %	50 %	60 %	5 %	10 %	20 %	30 %	40 %	50 %	60 %
Secondary sensor with primary calibration	--	--	--	--	--	--	--	--	--	--	--	--	--	--
DS	+	+	+	+	+	+	+	+	+	+	+	+	-	-
SSS+ LR	+	+	+	+	+	+	+	+	+	+	+	+	+	+
SSS + SVR	+	+	+	+	+	+	+	+	+	+	+	+	+	+
PLS + DS	+	+	+	+	+	+	+	+	+	+	+	+	+	-
PC + DS	--	--	--	--	--	--	--	--	--	--	--	--	--	--
PT	+	+	+	+	+	+	+	+	+	+	+	+	+	+
PC + PT	+	+	+	+	+	+	+	+	+	+	+	+	+	+
SBC	+	+	+	+	+	+	+	+	+	+	+	+	+	+
AMC	--	--	--	--	--	--	--	--	--	--	--	--	--	--
MMC	+	+	+	+	+	+	+	+	+	+	+	-	-	-
Recalibration with TS	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Research Output

- Conference Publications

  - **B. Muppidathi**, S. Laporte, Y. Ulanowski, S. Subbiah, B. Lebental, Multiparametric water quality sensor based on carbon nanotubes: Performance assessment in realistic environment, EGU General Assembly 2023 held at Vienna, Austria, EGU23-16053 from 23-28 April 2023 (doi: https://doi.org/10.5194/egusphere-egu23-16053).
  - **B. Muppidathi**, Assessing the effectiveness of a low-cost multiparametric sensor for real-time water quality monitoring in a controlled lab-scale water distribution network, LOTUS scientific workshop held at Université Gustave Eiffel, Marne La Vallée, France on 04 July 2023.
  - **B. Muppidathi**, S. Laporte, Y. Ulanowski, S. Subbiah, B. Lebental, Performance of a Multiparametric Water Quality Sensor in a Small-Scale Water Distribution Network, IEEE Sensors 2023 held at Vienna, Austria from 29 October 2023 to 01 November 2023 (doi: 10.1109/SENSORS56945.2023.10324917).
  - L. Usgodaarachchi, **B. Muppidathi**, B. Piro, S. Subbiah, B. Lebental. Polyvinyltriazole-functionalized SWCNT ink for printed Arsenic (III) chemistors for drink water monitoring applications. Réunion de Printemps du Club Microcapteurs chimiques 2024, Club MicroCapteurs Chimiques, Mar 2024, Grenoble, France. ⟨hal-04605900⟩

- Journal Publications

  - **B. Muppidathi**, B. B. Ngoune, H. Hallil, S. Subbiah, B. Lebental, Assessment of 10 Standardization-Based Calibration Transfer Techniques for Gas Sensors in a Small Data Context, IEEE Sensors, DOI: 10.1109/JSEN.2024.3466235.
  - B. B. Ngoune, M. Dumon, **B. Muppidathi**, B. Bondu, S. Subbiah, G. Perrin, S. Bila, C. Dejous, B. Lebental, H. Hallil, "Comparison of calibration strategies for a high sensitivity PEI-based RF humidity sensor," in IEEE Sensors Journal, DOI: 10.1109/JSEN.2024.3367594).

- Patents filed

  - **B. Muppidathi**, B. Lebental, S. Mishra, S. Subbiah, L. Usgodaarachchi, L. Challier, G. Koushik, Dispositif pour l'analyse d'un fluide comprenant une tête de sonde. France, Patent no: FR2315497. 2023. ⟨hal-04438019⟩

  - B. Lebental, **B. Muppidathi**, S. Mishra, S. Subbiah, L. Usgodaarachchi, L. Challier, G. Koushik, Dispositif pour l'analyse d'un fluide comprenant une portion de conduite. France, Patent no: FR2315496. 2023. ⟨hal-04438000⟩

  - **B. Muppidathi**, S. Subbiah, Method and system for online and at-line sensor fouling protection utilizing microporous barriers and automated flow direction reversal. India, Patent reference no.- 202131012436, 2021