D4.3 Water Quality Monitoring and Pollution Response

Authors: Martin Wagner (TZW), Theresia Meltzer (TZW)

Co-Authors: Stéphane DEVEUGHÈLE (3S), Michel JOUSSET (SUEZ)

Delivery Date May 2022

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant agreement No. 821036.

Disclaimer

This document reflects only the author's view. The European Commission is not responsible for any use that may be made of the information it contains.

Intellectual Property Rights

Project Consortium

Executive Summary

Monitoring and ensuring the quality of drinking water is one of the most important tasks of a water supplier, which requires sensors that can record important indicator parameters such as chlorine. However, due to the increasing use of sensors, the requirements for the evaluation of the recorded water quality data are also increasing. For the water supplier or operator, it is important to receive timely information about the changes in water quality or, if necessary, about possible sensor malfunctions. For this task, the use of event detection systems lends itself to automated evaluation of the data and thus supports the user in daily operation.

The accuracy and correctness of both the sensor data as well as the output of automated data evaluation models are prerequisites for efficient and safe operation of water works and distribution system.

In this deliverable two major aspects are shown:

1. The performance of a novel solid-state multiparametric electrochemical sensor (printed sensor) for free chlorine and temperature developed by Eurecat is tested under near real-life conditions in a model network (TZW).
2. The application of an event detection model (developed by TZW) to real-world water quality data obtained from the Cannes Demo Case (3S).

One important parameter in drinking water is free chlorine as an indicator for microbial safety as well as possible contaminations. Therefore, monitoring the chlorine concentration in the distribution network is an important task of a water supplier. For this purpose, a novel solid-state multiparametric electrochemical sensor (printed sensor) for free chlorine and temperature was developed by Eurecat and tested by TZW under near real-life conditions in a model network. In addition to the functional test of the sensor, tests were carried out to determine if software-based optimization of raw data processing can improve the performance of the sensors.

The experiments in the model network were carried out with drinking water of constant quality and constant hydraulic conditions, whilst the free chlorine concentration was changed by the addition of hypochlorous acid. First results show an indirect proportionality between the electric current that the sensor is measuring and the free chlorine concentration obtained by reference sensors.

Three different ways of improving the sensors performance due to software adaption where investigated. Namely the usage of a power function, of neural networks and of principal component regression to process the measurement data before the conversion in the actual concentration takes place. Of the three methods studied the Principal Components Regression showed the best improvement of accuracy through a pre-processing the raw data, not taking only one but a series of values into account.

To further evaluate the performance of the sensor, additional experiments covering the long-term stability, the influence of other water constituents and the influence of different hydraulic conditions on the printed sensor are necessary. If the further testing is successful, users in the water sector will have a reliable multiparameter sensor at their disposal, which, due to the printing technology, will only incur low costs.

For an easier evaluation of measured water parameter in real-time event detection tools are used. These warn the operator when significant changes in the water quality occur. The conventional early event detection systems issue an alarm when the monitored value exceeds or falls below a fixed

thresholds, whier the new approaches use artificial intelligence and machine learning to improve their performance. For example, by considering not only one but a combination of parameters or the time and range that single parameters are changing.

In this Deliverable the performance of an event detection model (developed by TZW) based on univariate signal evaluation as well as multivariate forecasting-based algorithms is evaluated. The data set used consists of measurements of seven different water quality parameters at four different measuring stations in the French Demo Case and covers a period of 12 month.

The model was successfully applied to the Cannes Demo Case. The performance of the model was measured both qualitatively through plausibility checks and quantitatively by means of F1 score (overall score of up to 0.85). It can be applied to time series of different characteristics (time series with stable signals that are subject to minor fluctuations as well as time series with strong oscillating signals) and can be flexibly adapted to the user's needs through numerous hyperparameters. In addition, different types of events can be detected (limit violations, violations of value ranges, abnormal signal changes, changes in the noise of the signal). This results in broad application possibilities for the model, and it can be assumed that the model can also be transferred to other application cases, such as water quality parameters of raw water monitoring and treatment in waterworks.

However, the outcomes are not only interesting for operators and water supply companies. Sensor manufactures can use the results to improve their own products and services, for example by extending their services with automated data evaluation and event detection.

Related Deliverables

Deliverable 3.2: FIWARE-enabled applications for Water Distribution

Deliverable 3.5: FIWARE-enabled Water Quality Sensors

Deliverable 4.2: FIWARE4_Leakage Management

Document Information

Programme	H2020 – SC0511-2018
Project Acronym	**Fiware4Water**
Project full name	FIWARE for the Next Generation Internet Services for the WATER sector
Deliverable	**D4.3: Water Quality Monitoring and Pollution Response**
Work Package	**WP4: Demonstrating Fiware4Water in the Real (Water) World**
Task	Task 4.2: FIWARE4_Water Distribution System Management in France
Lead Beneficiary	P8: TZW
Author(s)	Martin WAGNER (TZW), Theresia MELTZER (TZW)
Contributor(s)	Stéphane DEVEUGHÈLE (3S), Michel JOUSSET (SUEZ)
Quality check	Albert CHEN (UNEXE)
Planned Delivery Date	M35 (30/04/2022)
Actual Delivery Date	M36 (25/05/2022)
Dissemination Level	CONFIDENTIAL REPORT

Revision history

Version	Date	Author(s)/Contributor(s)	Notes
Draft1	06/05/22	Martin WAGNER (TZW), Theresia MELTZER (TZW)	
Draft2	09/05/22	Stéphane DEVEUGHÈLE (3S), Michel JOUSSET (SUEZ)	3S contribution: installation of 4 nano::stations on site
Final Draft	24/05/22	Albert CHEN (UNEXE)	Internal review
Final	24/05/22	Martin WAGNER (TZW), Theresia MELTZER (TZW)	

Table of content

Executive Summary 3

List of figures 7

List of tables 11

List of Acronyms/Glossary 12

Introduction 13

I. Monitoring of water quality with new sensors 14

I.1. Methodology 14

I.2. Results 19

II. Monitoring of water quality using Event Detection Algorithms at Cannes Demo Case 30

II.1. General context 30

II.2. Site description 32

II.3. Installation of multiparameter probes 33

II.4. Methodology 41

II.5. Analysis and interpretation of the measurement data 42

II.6. Results for Zone 801 50

II.7. Results for Zone 802 64

II.8. Results for Zone 803 76

II.9. Results for Zone 804 86

II.10. Discussion 94

Conclusion and Perspectives 99

European added value (EAV) and upscaling 100

References 101

Annex 102

List of figures

Figure 1 : Adapter for integration of printed sensor in TZWs model network 15
Figure 2: Raw data from three individual measurements.. 17
Figure 3: Raw data (left), selection of the last recorded current intensity per measurement (centre) and correlation of these current intensities with the chlorine concentration (right). 17
Figure 4: Example for a power function fitted to a current time series. The estimated function captures the overall slope and shape of the function while filtering out the high frequency noise components. 18
Figure 5: Comparison of the measurement curves of the 1-minute measurement for tests in beaker and in the model network 20
Figure 6: Time series of free chlorine concentration of sensor and reference measurement 21
Figure 7: Deviation of free chlorine, measured with printed sensor, from reference measurements as a function of free chlorine concentration 21
Figure 8: Deviation of free chlorine, measured with printed sensor, from reference measurements as a function of sensor run time 22
Figure 9. Residuals for a network trained for data from sensor 3 (R^2 of 0.869). Blue dots indicate predictions for time series the network was trained on, red dots indicate the predictions for the three time series that were not presented during training. The predictions are unremarkable, as they fit well with the predictions for training time series with similar chlorine concentrations............................. 24
Figure 10: Raw data of single measurements of a calibration. Each measurement is represented by one time series. The colour is equivalent to the concentration (blue to red, lowest to highest concentration). 25
Figure 11: Determined principal components of the calibration dataset.. 25
Figure 12: Comparison between reference calibration which uses the last recorded current value and the PCR model which takes the whole time series of each measurement into consideration. 26
Figure 13: Time series of reference sensor for chlorine determination and current of printed sensor (current in mA) under flow conditions.. 27
Figure 14: Raw data of printed sensor under flow conditions. The time series index on the y-axis corresponds with the vector that is recorded for each independent measurement. The time index on the x-axis corresponds to the measurement time (timestamp). The current values are shown on the z-axis (colormap). 27
Figure 15: Comparison of performance under flow conditions between benchmark which uses the last recorded current value and the PCR model which takes the whole time series of each measurement into consideration. 28
Figure 16: Comparison of the residuals with the chlorine concentration (left: benchmark, right: PCR model). 28
Figure 17: Comparison of the residuals with the duration of use of the printed sensor (left: benchmark, right: PCR model).. 29
Figure 18: Geographical location of the French Demo Case (DC2) – Cannes basin located in the south of France, on the shores of the Mediterranean Sea 32
Figure 19: Geographical location of the 4 nano::stations installed on the SICASIL drinking water distribution network.. 34
Figure 20: Water flow synoptic for the 4 nano::stations installed... 35
Figure 21: Site "Aubarède" – View of the room entrance (at the back on the right) 36

Figure 22: Site "Aubarède" – View of the nano::station and the on-line Hach CL17 chlorine sensor .. 36
Figure 23: Site "Suveret" – View of the room entrance 37
Figure 24: Site "Suveret" – View of the nano::station and the two on-line single chlori::lyser chlorine sensors (s::can) 37
Figure 25: Site "Espero Pax" – View of the room entrance 38
Figure 26: Site "Espero Pax" – View of the nano::station and the on-line Hach CL17 chlorine sensor 38
Figure 27: Site "Aigle" – View of the room entrance 39
Figure 28: Site "Aigle" – View of the nano::station 39
Figure 29: Basic structure of the model, which can be used to recognize four different types of events. 41
Figure 30: For the determination of recall, precision and F1 score, it is checked whether the detected events overlap with the manually classified events. A minute-by-minute comparison of individual time steps is not performed. FN : false negative, TP : true positive, FP : false positive............................. 42
Figure 31: Course of free chlorine, pH value and TOC. There is a negative correlation between chlorine and the two parameters... 44
Figure 32: Time series of free chlorine in the four zones 801 to 804. 44
Figure 33: Time series of free chlorine in August 2021 in all four zones. 45
Figure 34: Time series of temperature in zones 801 to 804. 46
Figure 35: Time series of electrical conductivity in zones 801 to 804.. 46
Figure 36: Time series of electrical conductivity in October 2021. 47
Figure 37: Time series of turbidity in all zones.. 48
Figure 38: Time series of opacity in zones 801 to 804 in June 2021. 48
Figure 39: Time series of UV absorption in zones 801 to 804.. 49
Figure 40: Manually labelled events in time series of free chlorine at zone 801. 50
Figure 41: Events (manual labelling) in time series of free chlorine in January 2021 (Zone 801). 51
Figure 42: Selected events (manual labelling) in time series of free chlorine. 51
Figure 43: Further examples of events in time series of free chlorine (Zone 801). 51
Figure 44: Detected chlorine events. Upper row: time series of free chlorine in zone 801. All other rows are different types of events where 0 indicates "no event" and 1 "is event". The last row shows the manual labelled events (yellow)... 52
Figure 45: Detected events of the type "is anomaly" for free chlorine in zone 801 for the period January to April 2021. 53
Figure 46: Detected events of the type "is anomaly" for free chlorine in zone 801 for the period May to December 2021. 54
Figure 47: Comparison of manually annotated events (top) and events detected by the model (bottom) in May in zone 801. 55
Figure 48: Comparison of manually annotated events (top) and events detected by the model (bottom) in September in zone 801.. 55
Figure 49: Detected chlorine events on 07.09. and 14.09. in zone 801... 56
Figure 50: Event of the type "limits violation" in which the upper limit value of 0.5 mg/L free chlorine was exceeded for a period of approx. two days (zone 801). 56
Figure 51: Detected event of the type "noise change" in the time series of free chlorine (zone 801). 57
Figure 52: Detected chlorine events when using a higher tolerance. Upper row: time series of free chlorine in zone 801. All other rows are different types of events where 0 indicates "no event" and 1 "is event". The last row shows the manual labelled events (yellow).. 58
Figure 53: Detected chlorine event where the concentration drops to 0 mg/L (zone 801). 58
Figure 54: Detected turbidity events. Upper row: time series of turbidity in zone 801. All other rows are different types of events where 0 indicates "no event" and 1 "is event". 59

Figure 55: All detected turbidity events of type « is anomaly » in zone 801. 60
Figure 56: Detected ultraviolet events. Upper row: time series of ultraviolet in zone 801. All other rows are different types of events where 0 indicates "no event" and 1 "is event". 61
Figure 57: Detected ultraviolet events in the first weeks of 2021. Upper row: time series of ultraviolet in zone 801. All other rows are different types of events where 0 indicates "no event" and 1 "is event". 62
Figure 58: Extract from the time series of ultraviolet (11.01. - 14.01.2021) to show fluctuations. 62
Figure 59: Detected ultraviolet events between May and December 2021. Upper row: time series of ultraviolet in zone 801. All other rows are different types of events where 0 indicates "no event" and 1 "is event". 63
Figure 60: Detected event in time series of ultraviolet at 15.05.2021. 63
Figure 61: Short peaks in time series of ultraviolet are not classified as events as their duration is too short. 64
Figure 62: Detected events in time series of ultraviolet in November 2021. 64
Figure 63: Manual annoted events in time series of chlorine in zone 802. 64
Figure 64: Outliers in time series of free chlorine during June 2021. 65
Figure 65: Examples of events identified in time series of free chlorine in zone 802. 65
Figure 66: Detected chlorine events. Upper row: time series of chlorine in zone 802. All other rows are different types of events where 0 indicates "no event" and 1 "is event". 66
Figure 67: Detected events of the type "is anomaly" for free chlorine in zone 802 for the period January to April 2021. 67
Figure 68: Detected "is anomaly" type events for free chlorine in Zone 802 for the period May to December 2021. 68
Figure 69: Time series of chlorine as well as manual labelled events (yellow, first row) and detected events (second row) in zone 802 for month May. 69
Figure 70: Time series of chlorine as well as manual labelled events (yellow, first row) and detected events (second row) in zone 802 for month June. 69
Figure 71: Time series of chlorine as well as manual labelled events (yellow, first row) and detected events (second row) in zone 802 for month July. 70
Figure 72: Time series of chlorine as well as manual labelled events (yellow, first row) and detected events (second row) in zone 802 for month August. 70
Figure 73:Time series of chlorine as well as manual labelled events (yellow, first row) and detected events (second row) in zone 802 for month September. 70
Figure 74: Time series of chlorine as well as manual labelled events (yellow, first row) and detected events (second row) in zone 802 for month October. 71
Figure 75: Time series of chlorine as well as manual labelled events (yellow, first row) and detected events (second row) in zone 802 for month November. 71
Figure 76: Time series of chlorine as well as manual labelled events (yellow, first row) and detected events (second row) in zone 802 for month December. 71
Figure 77: Events caused by a violation of the prescribed chlorine limit values (falling below 0.2 mg/L or exceeding 0.5 mg/L for a period of at least 24 h). 72
Figure 78: Detected events in time series of turbidity at zone 802 for the first six month of 2021. 73
Figure 79: Detected events in time series of turbidity at zone 802 for the last six month of 2021. 74
Figure 80: Detected events in the time series of ultraviolet in zone 802. 75
Figure 81: Manual annotated events in time series of free chlorine (upper figure: training phase, lower figure: validation phase). 76
Figure 82: Detected chlorine events. Upper row: time series of chlorine in zone 803. All other rows are different types of events where 0 indicates "no event" and 1 "is event". 77

Figure 83: Detected events of free chlorine in zone 803 from January to June 2021. 78
Figure 84: Detected events during July - December of free chlorine. 79
Figure 85: Detected events of type "noise changes" and "limits violation" in October and November in time series of free chlorine in zone 803. 80
Figure 86: Detected events in time series of turbidity during January - June in zone 803. 81
Figure 87: Detected events in time series of turbidity during July - December in zone 803 82
Figure 88: Time series of UV-Absorption in zone 803. 83
Figure 89: Detected events in time series of UV Absorption in zone 803 between January and June. 84
Figure 90: Detected events in time series of UV Absorption in zone 803 between July and December. 85
Figure 91: Manual annotated events in time series of free chlorine in zone 804 (upper figure: training phase, lower figure: validation phase). 86
Figure 92: Detected chlorine events. Upper row: time series of chlorine in zone 804. All other rows are different types of events where 0 indicates "no event" and 1 "is event". 87
Figure 93: Detected events of free chlorine in zone 804 from January to June 2021 88
Figure 94: Detected events of free chlorine in zone 803 from July to December 2021. 89
Figure 95: Time series of the free chlorine (top), the prediction error and the threshold (middle) as well as the classification of an event (bottom; 0: no, 1: yes). 90
Figure 96: Free chlorine (top), prediction error and threshold (middle) and classification result (bottom) for an event on 06.04.2021. 90
Figure 97: Results of event detection model for turbidity in zone 804. 91
Figure 98: Results of the event detection model for the time series of turbidity for the months January to June. Note that the y-axis has a different range of values for each month to better highlight details. 92
Figure 99: Results of the event detection model for the time series of turbidity for the months July to December. Note that the y-axis has a different range of values for each month to better highlight details. 93
Figure 100: Time series of UV-Absorption and detected events in zone 804. 94
Figure 101: Limits violations in time series of free chlorine in zone 804 during November 2021........ 96
Figure 102: Detection of a step event within a time series with oscillating values (free chlorine in zone 804). 96
Figure 103: Events that can be detected by the event detection model. The violation of limits assumes an upper limit of 0.5 mg/L (upper right graphic). 97
Figure 104: Examples of time series with different characteristics (strongly oscillating values at the top and relatively constant values at the bottom). 98

List of tables

Table 1: Example for calibration values and calibration curve 19
Table 2: Example for the check of the calibration in beaker before inserting the sensor in the model network 19
Table 3: Comparison of R^2 values of the original linear regression and the multiple regression based on the power function parameters. The results are mixed: For sensors 1 and 2, were the linear fit was not good to begin with, the multiple regression could improve the chlorine estimation, however, in case of a good linear fit (sensor 3), the multiple regression performed worse. 22
Table 4: Comparison of R^2 values of the original linear regression and the neural network predictions. In general, the fit is improved, however, the neural network needs much more data to train and the overall stability of the solution decreases the less data is available. Presented R^2 for the neural network are the best ones out of ten runs for the respective sensor. 23
Table 5: Comparison of the R2 score of the chlorine calibration before (original) and after (PCR) signal optimization. 26
Table 6: Four business issues for the French Demo Case 30
Table 7: NANOsensor: macro-planning and associated actions 30
Table 8: French Demo Case (DC2) – Correspondence between achievements and the concerned deliverable(s) 31
Table 9: Performance of event detection model for chlorine in zone 801. 52
Table 10: Performance of event detection model for chlorine in zone 802. 65
Table 11: Performance of event detection model for free chlorine in zone 803. 76
Table 12: Performance of event detection model for free chlorine in zone 804. 86
Table 13: Observed F1 scores for training, validation and overall period. 95
Table 14: Hyperparameters of event detection model for chlorine time series at Cannes Demo Case. 102
Table 15: Hyperparameters of event detection model for turbidity time series at Cannes Demo Case. 102
Table 16: Hyperparameters of event detection model for time series of ultraviolet at Cannes Demo Case. 103

List of Acronyms/Glossary

Acronym	Description
BIXX	Business Issue, one of the four French Business issues (BI01 to BI04)
CIRSEE	Centre International de Recherche Sur l'Eau et l'Environnement (in English, International Centre for Research on Water and Environment)
ECU	Electronic control unit
EGM	Easy Global Market, partner in the F4W project
EUT	Eurecat, partner in the F4W project
F4W	Fiware4Water project
HClO	Hypochlorous acid
NGI	Next Generation Internet *The Next Generation Internet (NGI) initiative, launched by the European Commission in the autumn of 2016, aims to shape the future internet as an interoperable platform ecosystem that embodies the values that Europe holds dear: openness, inclusivity, transparency, privacy, cooperation, and protection of data.*
SUEZ	Parent company of 3S and third party of 3S in the F4W framework
3S	SUEZ Smart Solutions, subsidiary of the worldwide group SUEZ, and partner in the F4W project
WPL	Work Packages Leaders
TZW	DVGW-Technologiezentrum Wasser, partner in the F4W project