

Iceland 
Liechtenstein
Norway grants

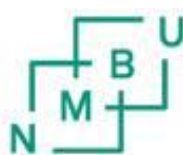
“Working together for a **green**,
competitive and **inclusive** Europe”

Project: *Digitalisation of water sector and water education - DIGIWATRO*,
Contract: 20-COP-0050

Intellectual Output 1: *Enhancing Digital Surveillance, Automation and Cybersecurity in Water Utilities Boosting*

Disclaimer

This document was realized with the EEA Financial Mechanism 2014-2021 financial support. Its content (text, photos, videos) does not reflect the official opinion of the Programme Operator, the National Contact Point and the Financial Mechanism Office. Responsibility for the information and views expressed therein lies entirely with the authors.



Norwegian University
of Life Sciences

Introducere

Primul rezultat intelectual al proiectului DigiwatRO și-a propus să dezvolte, să testeze și să valideze concepte ilustrative pentru supravegherea digitală, automatizarea și securitatea cibernetică în sectorul apei.

Impactul rezultatelor obținute a fost observat la trei niveluri: științific, academic și industrial.

- (1) Dintr-un punct de vedere științific, impactul poate fi observat prin intermediul articolelor rezultate din proiectul curent. Cercetarea întreprinsă în acest proiect formează deja baza pentru cercetările conexe. În plus, metodele de analiză dezvoltate în acest proiect vor fi utilizate în cercetările viitoare în domeniul digitalizării sectorului apei.
- (2) La nivel academic, impactul rezultatelor proiectului plasează echipele de cercetare în peisajul digitalizării sectorului apei. Fiecare membru al echipei a acumulat experiență în acest domeniu foarte promițător, un impact care se va vedea pe termen lung. Acesta este unul dintre rezultatele cele mai semnificative, deoarece echipele vor putea propune noi proiecte internaționale în domeniul digitalizării sectorului apei. Expertiza echipelor poate fi demonstrată prin rezultatele diseminate din acest proiect.
- (3) La nivel industrial, expertiza echipelor este folosită pentru transferul de tehnologie către sectorul privat. Workshopul organizat la ETFA 2023 în cadrul acestui proiect a implicat prezența companiilor din sectorul de automatizare.

Sectorul apei, și în special utilitățile de apă, trec printr-o digitalizare rapidă, conducând la noi investiții, echipamente noi și generând o cantitate uriașă de date online, cu o creștere a monitorizării și controlului la distanță. Informațiile generate de acest proiect sunt subutilizate în mod considerabil pentru controlul proceselor. Prin urmare, un aspect este să se exploreze și să se dezvolte concepte pentru îmbunătățirea calității datelor și utilizarea acestora pentru o monitorizare mai bună a proceselor. Un alt mare pericol cu care se confruntă utilitățile de apă este riscul amenințărilor cibernetice. Cele mai multe utilități de apă nu sunt conștiente de aceste pericole, iar obiectivul este de a crește conștientizarea, pregătirea și recuperarea rapidă în cazul unui eveniment.

În cadrul problemelor generate de sectorul apei, în acest proiect au fost identificate și investigate trei domenii științifice principale:

- Sistemele de Monitorizare și Control al Proceselor
- Senzorii Virtuali
- Analiza Incidentelor de Securitate Cibernetică în Sectorul Apei.

Pentru stabilirea acestor domenii, au fost luate în considerare structura și echipamentele stațiilor de tratare din România (Galați) și Norvegia (Oslo). Trei lucrări au rezultat din investigarea acestor domenii științifice și au fost prezentate la 28th International Conference on Emerging Technologies and Factory Automation – ETFA 2023, Sinaia, 12-15.09..2023 ([https://2023.ieee-etfa.org/main/static/files/program/ConferenceProgram Complete.pdf](https://2023.ieee-etfa.org/main/static/files/program/ConferenceProgram%20Complete.pdf)):

- Ratnaweera H, Nair A, Hykkerud A, Sivchenko N, Ratnaweera D, Condrachi L. Achieving legislative requirements in wastewater treatment using digital tools. ETFA 2023
- Ghinea LM, Miron M, Ratnaweera H. A Deep Learning Approach for Faults Recognition of Dissolved Oxygen Sensor in Wastewater Treatment Plants. ETFA 2023
- Țîru AE, Vasiliev I, Diaconu L, Vilanova R, Voipan D, Ratnaweera H. Integration of ANN for Accurate Estimation and Control in Wastewater Treatment. ETFA 2023.

1. Sisteme de Monitorizare și Control al Procesului

Diagnosticarea defectelor în stațiile de tartare a apelor uzate (WWTP) este importantă pentru a proteja comunitățile și ecosistemele de elemente toxice eliminate în apă. În acest sens, identificarea defectelor la senzori joacă un rol important, deoarece aceștia sunt componentele cheie ale sistemelor de control ale instalațiilor de tratare a apei, în special pentru că legislația de mediu este foarte strictă în ceea ce privește defecțiunile sau anomaliile în WWTP.

1.1. Abordare de Deep Learning pentru Recunoașterea Defectelor Sensorului de Oxigen Dizolvat în Stațiile de Tratare a Apelor Uzate

Acest studiu de caz evaluează performanțele a doi algoritmi de Deep Learning (FFNN - Feedforward Neural Network și 1DCNN - Convolutional Neural Network) în identificarea a patru defecte mecanice diferite care pot apărea în senzorul de oxigen dizolvat (DO) dintr-un proces de tratare a apelor uzate (SEAU). Toate defectele au fost analizate folosind Benchmark Simulation Model No. 2 (BSM2), dezvoltat de IWA Task Group (Alex et al., 2008). Structura instalației este identică cu cea stației de tratare a apelor uzate de la nivelul Municipiului Galați. În acest scop, au fost implementate blocuri de defecte în Matlab Simulink 2022a și au fost dezvoltate două clasificatoare neuronale în mediul Google Colaboratory (Colab) cu bibliotecile open-source Python: Scikit-Learn 1.2.2 și TensorFlow 2.12 cu Keras, o interfață Deep Learning API Integrated.

În acest studiu sunt analizate următoarele scenarii de defecte:

1. Defect de tip Bias – apare datorită unei decalări constante (v) în măsurătorile de ieșire ale senzorului. Injectarea de defect de tip bias poate fi implementată prin adăugarea unei valori constante la ieșirea senzorului. Ca rezultat, se produce o deviere față de valoarea normală la ieșirea de oxigen dizolvat (DO), definită matematic:

$$s(t) = h(t) + \eta + u, u = constant \quad (1.1)$$

unde $s(t) = h(t) + \eta$ reprezintă ieșirea așteptată a senzorului fără prezența defectelor, $h(t)$ este ieșirea senzorului la momentul de timp t și η este zgomotul.

2. Defect de tip Stuck - apare atunci când un senzor rămâne "blocat" într-o stare sau poziție specifică, nereușind să răspundă la modificările sistemului. Injectarea unui defect de tip stuck presupune ca ieșirea senzorului să fie blocată la o valoare fixă v pentru o perioadă temporară sau permanentă. Ca rezultat, se produce o defecțiune completă la ieșirea de oxigen dizolvat (DO), definită matematic:

$$s(t) = u, u = constant \quad (1.2)$$

3. Defect de tip Spike – când apar vârfuri de amplitudine mare la ieșirea senzorului de oxigen dizolvat (DO). Injectarea de defecte de tip spike se realizează, așa cum indică și numele, prin vârfuri de amplitudine mare la intervale de timp constante (r). Pentru a defini matematic un defect de tip vârf, se adaugă o deviere constantă b_t la elementele semnalului normal, după cum este prezentat mai jos:

$$s(t) = h(t) + \eta + b_t \quad (1.3)$$

$$\forall t \in u \times r$$

unde $v = \{1, 2, \dots\}$ este un set de numere naturale și r este intervalul în care defectele de tip spike au loc în ieșirea senzorului, cu $r \geq 2$.

- Defect de tip Precision degradation (PD) – apare ca o pierdere de precizie în senzorii sau sistemele de control folosite pentru monitorizarea și controlul procesului de tratare. Injectarea de defect de tip PD se realizează prin adăugarea unui zgomot cu medie zero și varianță mare la ieșirea senzorului de oxigen dizolvat (DO), definită matematic:

$$s(t) = h(t) + \eta + v \sim N(0, \delta_v^2), \delta_v^2 \gg \delta_n^2 \quad (1.4)$$

unde δ_v^2 este zgomotul cu medie zero și varianță mare.

Tabelul 1.1 arată durata și ziua de început a fiecărui scenariu de defect analizat.

Tabel 1.1 Defectele senzorului de oxigen dizolvat (DO)

Fault	Start [day]	Duration [hours]
Bias	280	480
Stuck	350	600
Spike	400, 420, 440, 460	48
PD	500	720

Conform (Alex et al., 2008), modelul instalației este simulat pe o perioadă de 609 zile. Datele sunt evaluate la intervale de 15 minute, începând cu ziua a 245-a. În studiul nostru au fost analizate următoarele scenarii de defecte (Fig. 2.1-2.4): bias, stuck, spike și precision degradation (PD).

Scenariul de defect de tip bias din Fig. 1.1 este generat în senzorul de oxigen dizolvat (DO) pe o perioadă de 20 de zile. Acest tip de defect este caracterizat de o diferență constantă între valoarea reală și ieșirea eronată a senzorului DO de +1.5 mg/L (Liu et al., 2022).

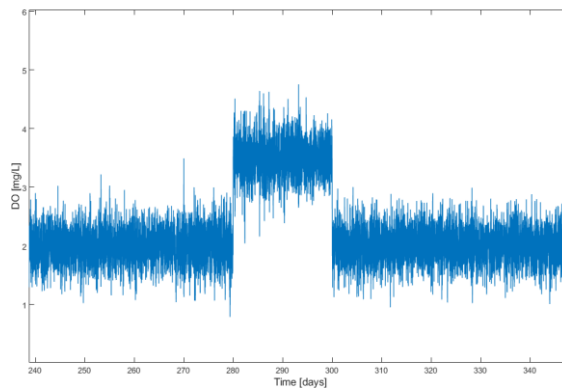


Fig. 1.1. Defect de tip bias al senzorului de DO

Scenariul de tip stuck din Fig. 1.2 este injectat în senzorul de DO pe o perioadă de 25 zile. Acest tip de defect presupune ca măsurătorile senzorului de oxigen dizolvat (DO) să rămână la o valoare fixă, în acest caz, la 2 mg/L, și astfel nu mai răspund la nicio variație a concentrației de DO. Acest defect este considerat o defecțiune completă, care poate fi o problemă temporară sau permanentă.

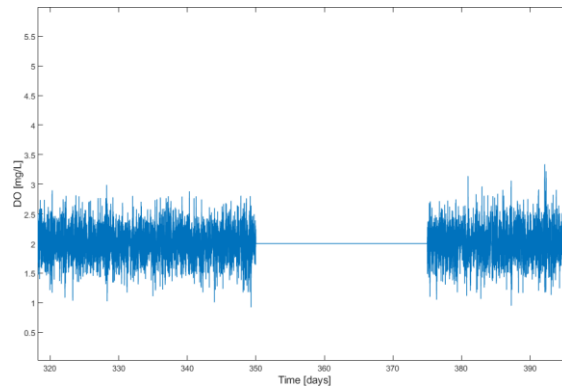


Fig. 1.2. Defect de tip stuck al senzorului de DO

Defectul de tip spike din Fig. 1.3 este generat în senzorul de DO în 4 intervale de timp, având amplitudini diferite care constau în 2 zile fiecare.

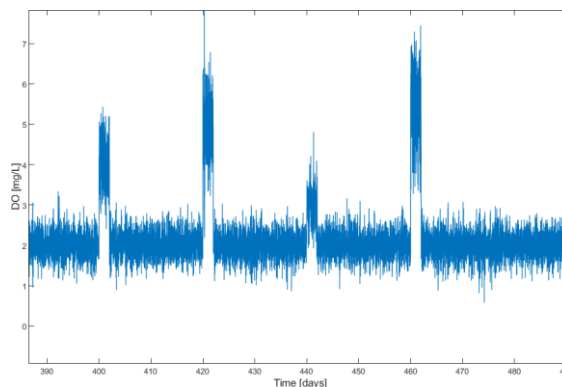


Fig. 1.3. Defect de tip spike al senzorului de DO

Scenariul de defect de tip PD din Fig. 1.4 este implementat în senzorul de DO pe o perioadă de 30 de zile. Acest tip de defect constă în adăugarea unui zgomot cu medie zero și varianță mare la ieșirea senzorului.

Tehnicile de Deep Learning sunt intens utilizate în domeniul diagnosticării defectelor în stațiile de epurare a apelor uzate (WWTPs). Mai mult, aceste procese sunt complexe, dinamice și neliniare, adesea predispuse la defecțiuni, incertitudini și perturbări. Din această perspectivă, este important să avem instrumente performante, cum ar fi Rețelele Neuronale (Neural Networks - NNs), pentru a identifica cu precizie și eficiență orice defecte mecanice care pot apărea în aceste procese, în special în cazul senzorilor utilizați pentru controlul consumului de energie și calitatea deversării (Chi and Guo, 2019; Mamandipoor et al., 2020).

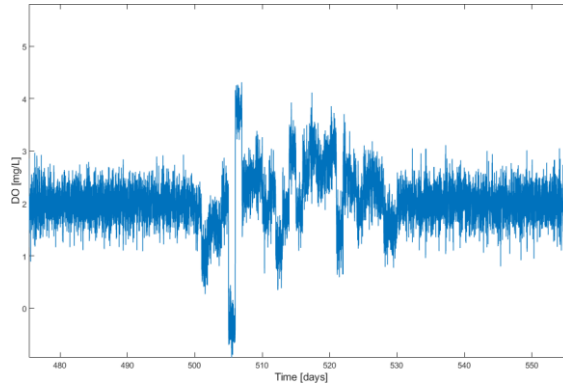


Fig. 1.4. Defect de tip precision degradation (PD) al senzorului de DO

De exemplu, în cazul senzorilor de oxigen dizolvat, sistemele de aerare depind de valorile măsurate de senzorul DO, astfel încât orice defecțiune în acest semnal poate afecta funcționarea normală a sistemului (Salles et al., 2023).

În acest sens, studiul de față compară două modele neuronale, FFNN și 1DCNN. Scopul este să se stabilească ce clasificator DL este capabil să identifice cu precizie cele 5 stări de funcționare ale defectelor senzorului DO în stațiile de epurare a apelor uzate: normal (clasă 0), bias (clasă 1), stuck (clasă 2), spike (clasă 3) și precision degradation (PD) (clasă 4).

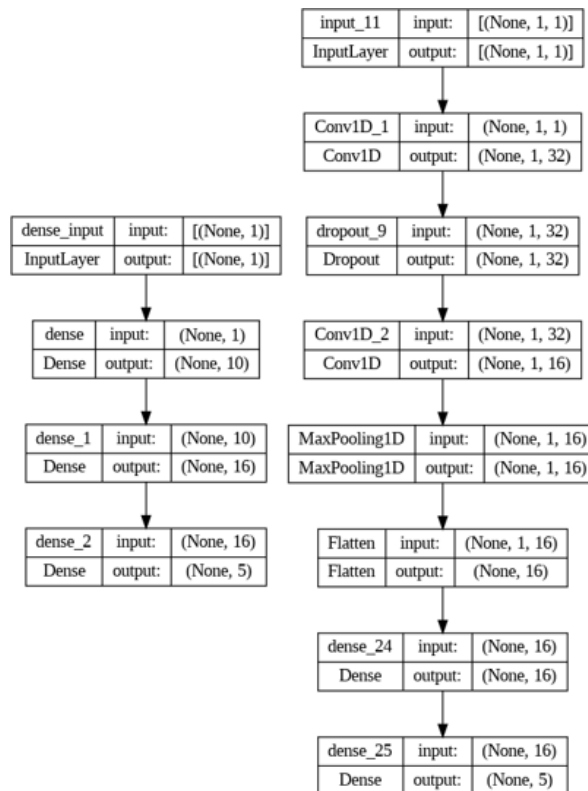


Fig. 1.5. Arhitecturi de Deep Learning: a) FFNN și b) 1DCNN

Arhitectura fiecărei metode de Deep Learning este prezentată în Fig. 1.5. În cazul FFNN, straturile sunt adăugate secvențial: input layer (10 neuroni, ReLU – funcție de activare), 1 hidden layer (16 neuroni și ReLU – funcție de activare) și output layer (5 neuroni și Softmax – funcție de activare). Pentru a nu apărea fenomenul de overfitting, regularizarea activității este

setată la 0.01 în input layer al modelului neuronal. Pe de altă parte, în cazul 1DCNN, straturile sunt: Conv1D layer (filters=32, kernel_size=1, activation='relu'), Dropout layer (dropout rate=0.2), Conv1D layer (filters=16, kernel_size=1, activation='relu'), MaxPool1D layer (pool_size=1), Flatten layer, Dense layer (16 neuroni, activation='relu') și Dense layer (5 neuroni, activation='softmax').

Ambele metode de Deep Learning sunt compilate cu funcția Categorical Cross Entropy Loss, Adam optimizer (rata de învățare=0.001) și metrică de acuratețe pentru evaluarea modelului în timpul antrenării și validării. De asemenea, antrenarea se face cu metode fit() din Keras API integrate în 100 epochs. Din setul de date sunt selectate 80% date pentru antrenare și 20% date pentru testare.

Hyperparametrii rețelelor neuronale au fost selectați în jurul valorilor sugerate de KerasTuner, o bibliotecă generală de tuning a hyperparametrilor.

Metricile de performanță utilizate pentru evaluarea modelului neuronal sunt:

- Acuratețe

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1.5)$$

unde TP – true positive, TN – true negative, FP – false positive și FN – false negative, toate obținute din matricea de confuzie.

- Precizie

$$Precision = \frac{TP}{TP+FP} \quad (1.6)$$

unde Precision – reprezintă datele pozitive prezise.

- Recall

$$Recall = \frac{TP}{TP+FN} \quad (1.7)$$

unde Recall – calculează procentul de date identificate într-o clasă relevantă.

- Scor F1

$$F1 - score = 2 \frac{Precision \cdot Recall}{Precision + Recall} = \frac{2TP}{2TP+FP+FN} \quad (1.8)$$

unde F1-score – reprezintă media armonică a preciziei și recall-ului modelului.

Metodele de Deep learning au obținut: FFNN - 98.32% acuratețe generală de antrenare și 98.30% acuratețe generală de testare, iar 1DCNN - 92.76% acuratețe generală de antrenare și 92.90% acuratețe generală de testare. Matricele de confuzie ale celor două metode de Deep Learning sunt prezentate în Fig. 1.6. Matricea de confuzie a FFNN arată că modelul clasifică corect într-o proporție de 98.3% (11494 date din totalul de 11693) și eronat într-o proporție de 1.7% (199 date din totalul de 11693). Pe de altă parte, matricea de confuzie a 1DCNN arată că modelul clasifică corect într-o proporție de 92.90% (10863 date din totalul de 11693) și eronat într-o proporție de 7.1% (830 date din totalul de 11693). Bazându-ne pe aceste rezultate, devine evident că performanța FFNN depășește performanța 1DCNN. Acest lucru a fost evidențiat în raportul de clasificare din Figura 1.7, cu o linie punctată roșie.

normal	10020.0	2.0	29.0	0.0	37.0	normal	9917.0	0.0	194.0	0.0	7.0
bias	4.0	394.0	0.0	0.0	0.0	bias	31.0	357.0	0.0	0.0	0.0
stuck	13.0	0.0	448.0	0.0	0.0	stuck	0.0	0.0	473.0	0.0	0.0
spike	19.0	0.0	0.0	139.0	0.0	spike	20.0	61.0	0.0	65.0	0.0
pd	95.0	0.0	0.0	0.0	493.0	pd	466.0	50.0	1.0	0.0	51.0
	normal	bias	stuck	spike	pd		normal	bias	stuck	spike	pd
	a)						b)				

Fig. 1.6. Matricele de confuzie ale metodelor de Deep Learning: a) FFNN și b) 1DCNN

Conform Fig. 1.7, cele mai bune rezultate de clasificare sunt obținute în cazul scenariului de defect Bias cu FFNN (99.49% Precision, 98.99% Recall și 99.24% F1-score) iar cele mai scăzute rezultate sunt obținute în cazul scenariului de defect PD cu 1DCNN (87.93% Precision, 0.08% Recall și 0.16% F1-score). Acest lucru indică faptul că datele de suport pentru PD din 1DCNN sunt dezechilibrate în comparație cu celelalte clase. Prin urmare, ar putea fi necesară colectarea de mai multe date pentru a îmbunătăți performanța modelului 1DCNN. Cu toate acestea, studiul nostru arată că FFNN are performanțe și adaptabilitate mai bune decât 1DCNN și reprezintă o unealtă puternică de învățare profundă pentru îmbunătățirea recunoașterii defectelor senzorilor DO în stațiile de epurare a apelor uzate.

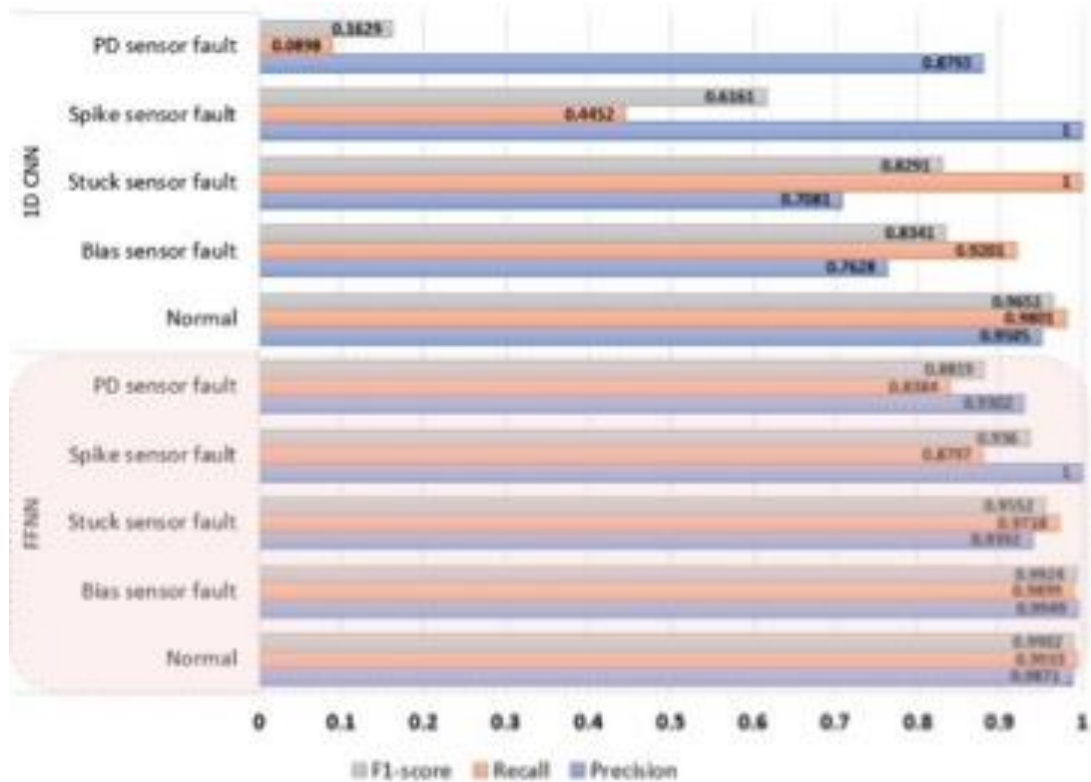


Fig. 1.7. Metricile (F1-score, Recall, Precision) pentru FFNN și 1DCNN

Modelul FFNN este foarte eficient în comparație cu 1DCNN pentru îmbunătățirea recunoașterii defectelor senzorului de oxigen dizolvat.

1.2. Controlul unui proces de tratare a apelor uzate folosind control predictiv liniar și neliniar.

Controlul procesului de tratare a apelor uzate nu este o sarcină trivială, deoarece sistemul este neliniar, prezintă constante de timp mari și întârzieri, iar interacțiunea dintre variabile este importantă (Luenberger, 1979). Prin urmare, tehnica Model Predictive Control (MPC) este o abordare bună pentru o astfel de sarcină exigentă, datorită unui avantaj important: MPC are capacitatea de a gestiona constrângeri complexe și neliniarități, care sunt specifice proceselor de tratare a apelor uzate; în plus, MPC poate gestiona modificări ale procesului și perturbări, cum ar fi modificări ale calității și debitului influent, prin ajustarea dinamică a acțiunilor de control (Necoara, 2008). Prin urmare, MPC este larg utilizat în industria proceselor actuale datorită capacității sale de a gestiona probleme de control cu constrângeri și multivariable (Socha, 2007). În cazul sistemelor liniare, problema MPC este de obicei formulată ca o problemă convexă pătratică (QP), în timp ce pentru sistemele neliniare este necesar să se rezolve o problemă de optimizare extrem de neconvexă (DeCarlo, 1989).

ACADO Toolkit este un mediu de software și o colecție de algoritmi scrise în C++ pentru controlul automat și optimizarea dinamică (Ariens et al., <http://www.acadotoolkit.org>). Oferă un cadru general pentru utilizarea unei mari varietăți de algoritmi pentru controlul optimal direct, inclusiv controlul predictiv al modelului, precum și estimarea stării și a parametrilor. De asemenea, furnizează (în mod stand-alone) integratoare eficiente implementate de tip Runge-Kutta și BDF pentru simularea ecuațiilor diferențiale ordinare (ODE) și a ecuațiilor diferențiale

algebrice (DAE) (Ariens et al., 2010). Pentru problema neliniară de MPC a procesului de tratare a apelor uzate, folosim ACADO pentru a rezolva problema corespunzătoare. Cu toate acestea, deoarece problema este extrem de nonconvexă, ACADO necesită perioade lungi de timp ale CPU-ului pentru a o rezolva. Prin urmare, propunem o alternativă în care liniarizăm sistemul la fiecare moment de eșantionare, derivăm problema corespunzătoare de MPC liniar și o rezolvăm cu un solver pătratic din Matlab. Din simulări, observăm că cea de-a doua abordare este mai eficientă, deoarece comportamentele în buclă închisă sunt similare pentru abordările de MPC liniară și neliniară, dar abordarea bazată pe liniarizare este mai rapidă în ceea ce privește timpul de CPU decât ACADO.

Algoritmul de Model Predictive Control neliniar

Un sistem neliniar este un sistem în care schimbarea ieșirii nu este proporțională cu schimbarea intrării (Dasarathy, 1970). Scriem astfel sistemul neliniar cu care lucrăm (Khalil 2002):

$$\dot{x} = f(x, u) \quad (1.11)$$

unde $x \in \mathbb{X}$ și $u \in \mathbb{U}$. Astfel, considerăm următoarea problema de control optimal:

$$\min_{u(\cdot)} \int_0^T L(x, u) dt + V(x(T)) \quad (1.12)$$

supusă la următoarele constrângeri de egalitate:

$$\begin{aligned} \dot{x} &= f(x, u) \\ x(0) &= x_0 \text{ given } x \in \mathbb{X} \text{ and } u \in \mathbb{U} \end{aligned}$$

unde termenul $L(x, u)$ reprezintă costul intermediar, și $V(x(T))$ costul final. În termeni abstracți, aceasta este o problemă de optimizare cu constrângeri. Pentru urmărirea referințelor, luăm în considerare semnalele de referință x^{ref} și u^{ref} . Atunci, $L(x, u) = \|x - x^{ref}\|_{Q_x}^2 + \|u - u^{ref}\|_{R_u}^2$ și $V(x(T)) = \|x_T - x_T^{ref}\|_{P_x}$. Pentru a simplifica algoritmi pe care îi utilizăm, vom aproxima sistemul neliniar cu care lucrăm folosind principiul seriilor Taylor. Începem cu un sistem de forma:

$$\dot{x} = f(x, u), x \in \mathbb{X}, u \in \mathbb{U} \quad (1.13)$$

Fie (\bar{x}, \bar{u}) un punct de echilibru, astfel $f(\bar{x}, \bar{u}) = 0$. Apoi, aplicând seriile Taylor pe funcția f în jurul punctului de echilibru, obținem (Socha, 2007):

$$f(x, u) \approx f(\bar{x}, \bar{u}) + \frac{\partial f}{\partial x}(\bar{x}, \bar{u})(x - \bar{x}) + \frac{\partial f}{\partial u}(\bar{x}, \bar{u})(u - \bar{u}) \quad (1.14)$$

Notăm $\delta x = x - \bar{x}$ și $\delta u = u - \bar{u}$ și obținem liniarizarea $\delta \dot{x} = \frac{\partial f}{\partial x}(\bar{x}, \bar{u})\delta x + \frac{\partial f}{\partial u}(\bar{x}, \bar{u})\delta u$, care poate fi scrisă și sub forma:

$$\delta \dot{x}(t) = A_x \delta x(t) + B_u \delta u(t) \quad (1.15)$$

unde $A_x = \frac{\partial f}{\partial x}(\bar{x}, \bar{u})$ și $B_u = \frac{\partial f}{\partial u}(\bar{x}, \bar{u})$. Dacă sistemul este discretizat cu perioada de eșantionare ΔT , el poate fi scris și sub forma (Jacod and Protter, 2011):

$$\delta x_{k+1} = A_x \delta x_k + B_u \delta u_k \quad (1.16)$$

Algoritmul de Model Predictive Control liniar

În general, un system cu constrângeri liniare pe stări și intrări are următoarea formă (Chen, 1984):

$$\begin{aligned} x_{k+1} &= A_x x_k + B_u u_k \\ lb_x &\leq x_k \leq ub_x, \forall k \in \{0, \dots, N-1\} \\ C_u u_k &\leq d_u \end{aligned} \quad (1.17)$$

unde $C_u \in \mathbb{R}^{n_i \times n_x}$ și $d_u \in \mathbb{R}^{n_i}$. Aunci, problema de control optimal cu referințe devine:

$$\begin{aligned} \min_{x_k, u_k} \frac{1}{2} \sum_{k=0}^{N-1} \|x_k - x_k^{ref}\|_{Q_x}^2 + \sum_{k=0}^{N-1} \|u_k - u_k^{ref}\|_{R_u}^2 + \|x_N - x_N^{ref}\|_{P_x} \\ x_0 = x, x_{k+1} = A_x x_k + B_u u_k \\ lb_x \leq x_k \leq ub_x, C_u u_k \leq d_u, \forall k \in \{0, \dots, N-1\} \end{aligned} \quad (1.18)$$

unde x_k^{ref} și u_k^{ref} sunt anumite valori de referință pentru vectorii de stare și variabile de control ale sistemului, pe întregul orizont de predicție, iar norma ponderată are următoarea formă:

$$\|v_k - v_k^{ref}\|_V^2 = (v_k - v_k^{ref})^T V (v_k - v_k^{ref})$$

De asemenea, matricele Q_x și R_u sunt ambele considerat semipozitiv definite, $\forall k$.

Problema considerată anterior poate fi tratată ca o problemă de optimizare convexă pătratică. Pentru a o rezolva, notăm întâi variabila de decizie cu $x \in \mathbb{R}^{n_x + n_u}$ și putem scrie (Necoara, 2013):

$$x = [u_0^T x_1^T u_1^T x_2^T \dots u_{N-1}^T x_N^T] \quad (1.19)$$

Constrângerile de egalitate ale problemei de optimizare sunt liniare și provin din definirea dinamicii sistemului controlat. Pentru a formula problema pătratică, acestea pot fi concatenate într-o singură constrângere de bloc, având în vedere modul în care a fost definită variabila de decizie. Din punct de vedere matematic, acest lucru înseamnă:

$$A_x = b, A \in \mathbb{R}^{Nn_x \times N(n_x+n_u)}, b \in \mathbb{R}^{Nn_x} \quad (1.20)$$

Matricele A și b au următoarea formă:

$$A = \begin{bmatrix} -B_u & I_{n_x} & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & -A_x & -B_u & I_{n_x} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -A_x & -B_u & I_{n_x} \end{bmatrix}$$

$$b = \begin{bmatrix} A_x x_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Această concatenare are sens deoarece ecuațiile care descriu dinamica sistemului pe întregul orizont de predicție pot fi scrise astfel:

$$\begin{aligned} x_1 &= A_x x_0 + B_u u_0 \Leftrightarrow -B_u u_0 + I_{n_x} x_1 = A_x x_0 \\ x_2 &= A_x x_1 + B_u u_1 \Leftrightarrow -A_x x_1 - B_u u_0 + I_{n_x} x_2 = 0 \\ &\vdots \\ x_N &= A_x x_{N-1} + B_u u_{N-1} \Leftrightarrow -A_x x_{N-1} - B_u u_{N-1} + I_{n_x} x_N = 0 \end{aligned}$$

Structura matricei A este un bloc tridiagonal, iar singurul termen nenul al vectorului b depinde de starea inițială, care va fi actualizată la fiecare iterație a algoritmului. Constrângerile de inegalitate pot, de asemenea, să fie rescrise sub forma unei singure constrângeri $Cx \leq d$. Inegalitățile corespunzătoare stărilor sistemului pot fi scrise astfel:

$$\begin{bmatrix} I_{n_x} \\ -I_{n_x} \end{bmatrix} x_k \leq \begin{bmatrix} ub_x \\ -lb_x \end{bmatrix}$$

Astfel, matricele $C \in \mathbb{R}^{N(2n_x+n_i) \times N(n_x+n_u)}$ și $d \in \mathbb{R}^{N(n_x+n_u)}$ au următoarea formă:

$$C = \begin{bmatrix} C_u & 0 & 0 & \cdots & 0 & 0 \\ 0 & C_x & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & C_u & 0 \\ 0 & 0 & 0 & \cdots & 0 & C_x \end{bmatrix}$$

$$d \in \begin{bmatrix} d_u \\ d_x \\ \vdots \\ d_u \\ d_x \end{bmatrix}$$

În ceea ce privește rescrierea funcției obiectiv în forma standard convexă a problemei QP, trebuie construit un vector cu o structură similară cu variabila de decizie x , dar care concatenează valorile de referință pe întregul orizont de predicție. Îl denumim x^{ref} iar forma sa este următoarea:

$$x^{ref} = [(u_0^{ref})^T (x_1^{ref})^T (u_1^{ref})^T (x_2^{ref})^T \dots (u_{N-1}^{ref})^T (u_N^{ref})^T] \quad (1.21)$$

Dacă luăm în considerare și matricea diagonală $Q \in \mathbb{R}^{N(n_x+n_u) \times N(n_x+n_u)}$, scrisă:

$$Q = \begin{bmatrix} R_u & 0 & 0 & \cdots & 0 & 0 \\ 0 & Q_x & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & R_u & 0 \\ 0 & 0 & 0 & \cdots & 0 & P_x \end{bmatrix}$$

atunci funcția cost în problema de optimizare poate fi rescrisă ca:

$$\begin{aligned} \frac{1}{2} \|x - x^{ref}\|_Q^2 &= \frac{1}{2} (x - x^{ref})^T Q (x - x^{ref}) \\ &= \frac{1}{2} (x^T Q x - x^T Q x^{ref} - (x^{ref})^T Q x + (x^{ref})^T Q x^{ref}) \\ &= \frac{1}{2} x^T Q x - x^T Q x^{ref} + \frac{1}{2} (x^{ref})^T Q x^{ref}. \end{aligned}$$

Apoi, notăm $q = -Qx^{ref}$ și neglijând termenul $\frac{1}{2} (x^{ref})^T Q x^{ref}$, obținem problema în formă convexă QP:

$$\begin{aligned} \min_{x \in \mathbb{R}^{N(n_x+n_u)}} \quad & \frac{1}{2} x^T Q x + q^T x \\ \text{s.t.} \quad & Ax = b, \quad Cx \leq d. \end{aligned} \quad (1.22)$$

Pentru a avansa în cercetarea noastră, am decis să reproducem rezultatele obținute în ACADO prin scrierea unui algoritm în MATLAB. Scopul este același: stabilizarea sistemului de tratare a apelor uzate în punctele de echilibru. Astfel, folosim liniarizarea asupra sistemului continuu, apoi aplicăm discretizarea, după cum este prezentat în schema de mai jos:

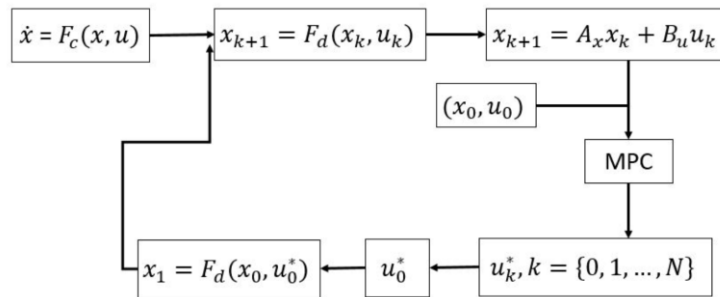


Fig. 1.8. Schema de liniarizare și discretizare a sistemului

Fig. 1.8. afișează discretizarea și liniarizarea sistemului pentru procesul de control predictiv.

În esență, luăm sistemul continuu, îl transformăm într-unul discret și apoi aplicăm liniarizarea. Din punct de vedere matematic, avem următoarele forme:

$$\dot{x} = F_c(x, u) \quad (1.23)$$

$$x_{k+1} = F_d(x_k, u_k) \quad (1.24)$$

Notăm cu F_c funcția continuă și cu F_d pe cea discretă.

Primul pas ar fi să discretizăm folosind metoda Euler. Astfel, obținem:

$$x_{k+1} = x_k + \Delta T F_c(x_k, u_k) \quad (1.25)$$

Considerând $x_{k+1} = F_d(x_k, u_k)$, obținem:

$$F_d(x, u) = x + \Delta T F_c(x, u) \quad (1.26)$$

Următorul pas este liniarizarea. Aproximarea liniară a unei funcții este dezvoltarea în serie Taylor de ordinul întâi în jurul punctului de interes. Prin urmare, putem scrie:

$$\begin{aligned}
F_d(x_k, u_k) &\approx F_d(x_0, u_0) + \frac{\partial F_d}{\partial(x_k, u_k)}(x_0, u_0) \begin{bmatrix} x_k - x_0 \\ u_k - u_0 \end{bmatrix} \\
&= F_d(x_0, u_0) + \begin{bmatrix} \frac{\partial F_d}{\partial x_k}(x_0, u_0) & \frac{\partial F_d}{\partial u_k}(x_0, u_0) \end{bmatrix} \begin{bmatrix} x_k - x_0 \\ u_k - u_0 \end{bmatrix} \\
&= F_d(x_0, u_0) + \frac{\partial F_d}{\partial x_k}(x_0, u_0)(x_k - x_0) + \frac{\partial F_d}{\partial u_k}(x_0, u_0)(u_k - u_0).
\end{aligned}$$

Notăm:

$$\begin{aligned}
A_x &= \frac{\partial F_d}{\partial x_k}(x_0, u_0) \\
B_u &= \frac{\partial F_d}{\partial u_k}(x_0, u_0).
\end{aligned}$$

Deoarece $F_d(x, u) = x + hF_c(x, u)$, matricele pot fi scrise și în forma:

$$\begin{aligned}
A_x &= \frac{\partial(x + hF_c)}{\partial x} = I_n + h \frac{\partial F_c}{\partial x} \\
B_u &= \frac{\partial(x + hF_c)}{\partial u} = h \frac{\partial F_c}{\partial u}.
\end{aligned} \tag{1.27}$$

A fost folosit un model matematic simplificat pentru a descrie comportamentul stației de epurare a apelor uzate (Ghinea et al., 2023). Modelul are 4 variabile de stare: X – concentrația biomasei, S – concentrația de substrat (încarcarea organică), DO – oxigenul dizolvat și X_r – concentrația biomasei recirculate.

Toolkit-ul ACADO

ACADO aplică practic controlul optimal pe un sistem neliniar. Pentru a face acest lucru, luăm în considerare următorul cost:

$$\begin{aligned}
\min_{x_k, u_k} & \frac{1}{2} \sum_{k=1}^N \|x_k - x_k^{ref}\|_{Q_k}^2 + \sum_{k=0}^{N-1} \|u_k - u_k^{ref}\|_{R_k}^2 \\
\text{s.t.} & : x_0 = x, \quad x_{k+1} = A_x x_k + B_u u_k \\
& lb_x \leq x_k \leq ub_x \\
& C_u u_k \leq d_u, \quad \forall k = \{0, 1, \dots, N-1\}.
\end{aligned} \tag{1.28}$$

Aici, x reprezintă stările, $x = [X \ S \ DO \ X_r]^T$ și u reprezintă variabila de intrare, $u = W$. Scopul nostru este să obținem valori de până la 40 mg/l pentru substratul S , și în același timp

să menținem valorile oxigenului dizolvat în intervalul 1 – 3 mg/l. Mai mult, variabila de intrare, W , trebuie menținută sub $100 \text{ m}^3/\text{h}$. Astfel, considerăm constrângerile pe intrare $0 \leq u \leq 100$, constrângerile pe ieșire $0 \leq y \leq 40$ și constrângerile pe stare $1 \leq x(3) \leq 3$ Astfel, se obține calitatea apei conform standardelor naționale în domeniu, menținând totodată un consum rezonabil de electricitate. Acest consum, care este consumul principal în cazul stațiilor de epurare a apelor uzate, este determinat de suflantele utilizate pentru aerare.

În ACADO, creăm o matrice Q care conține Q_k și R_k din costul de mai sus,

$$Q_k = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ și } R_k = 1. \text{ În principiu, dorim să stabilizăm sistemul nostru folosind}$$

starea S și variabila de control W . Acest lucru presupune să luăm în considerare costul care are aproximativ forma $\min_W \int S^2 + W^2$. Tot ce rămâne este să alegem punctele de echilibru pentru sistem. Luăm în considerare, de asemenea, valoarea pentru pasul de eșantionare $\Delta T = 0.5$.

Considerăm două seturi de puncte de echilibru:

Case 1. $rr = [223.8 \ 38.9 \ 2.6 \ 447.6 \ 33]$.

Case 2. $rr = [228.7 \ 37.1 \ 2.9 \ 457.4 \ 35]$.

Observăm că atât stările, cât și variabila de intrare ating valorile pe care dorim să le obținem, indiferent de punctele de echilibru pe care le stabilim.

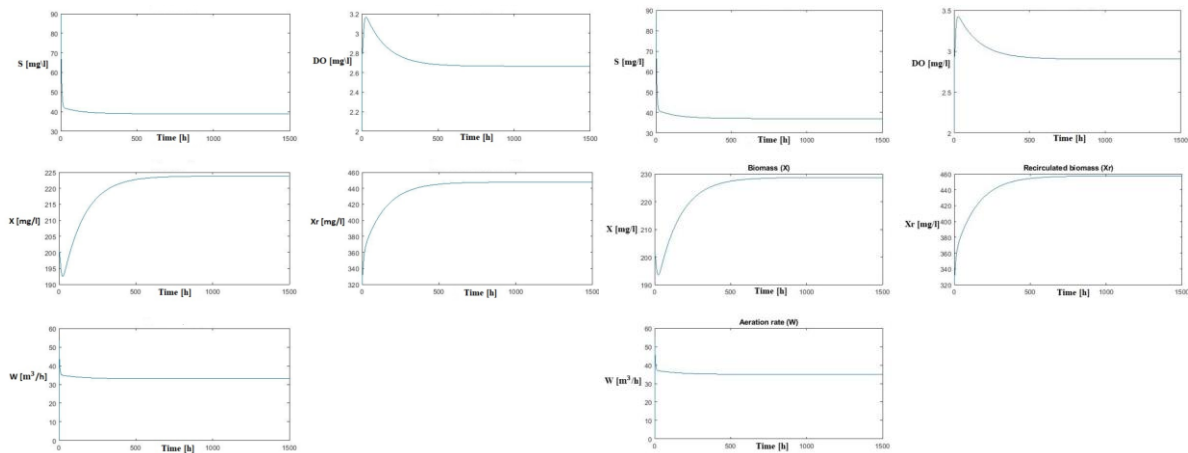


Fig. 1.9 Simularea ACADO pentru Cazul 1.

Fig. 1.10 Simularea ACADO pentru Cazul 2.

Control Optimal în Matlab

În Matlab, simulăm problema de optimizare pentru forma următoare:

$$x_{k+1} = A_x x_k + B_u u_k + a$$

unde $a = F_c(x_0, u_0) - A_x x_0 - B_u u_0$.

Pentru a face acest lucru, construim mai întâi o funcție care calculează matricele pe care le denumim A_x și B_u folosind (1.27), apoi utilizăm cele două matrice, împreună cu $Q_x = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ și $R_u = 1$ pe care le-am folosit în ACADO (deoarece costul este același) într-o altă funcție care creează elementele necesare pentru forma sparse a problemei QP. Această a doua funcție rezultă în elementele Q, q, A, b, C, d care compun următoarea formă QP:

$$\min_{\psi} \frac{1}{2} \psi^T Q \psi + q^T \psi \quad (1.29)$$

Având constrângerile $A\psi \leq b$ și $C\psi \leq d$.

În scriptul care apelează aceste funcții, inițializăm mai întâi stările și variabilele de control ca $x_0 = [200; 90; 2; 320]$ și $u_0 = 30$. Apoi, conform schemei din Figura 1.8, calculăm u_k^* , $k = \{0, 1, \dots, N\}$ folosind fie comanda *quadprog* (un solver pentru funcții obiectiv pătratice cu constrângeri liniare), sau Software-ul de Matlab CVX (un sistem de modelare bazat pe Matlab pentru optimizare convexă; CVX transformă Matlab într-un limbaj de modelare, permițând specificarea constrângerilor și obiectivelor folosind sintaxa standard a expresiilor Matlab). Indiferent pe care dintre aceste două decidem să o folosim, pașii următori sunt aceiași.

Utilizăm doar primul element din u_k^* , $k = \{0, 1, \dots, N\}$, adică u_0^* , apoi calculăm valorile următorului vector de stare, x_1 folosind forma discretă F_d calculată pentru valoarea curentă a lui x , care este x_0 și u_0^* pe care tocmai l-am determinat. Tot ce ne rămâne de făcut este să setăm x_0 și u_0 pe care urmează să le folosim la următoarea iterație:

$$\begin{aligned} x_0 &= x_1 \\ u_0 &= u_0^* \end{aligned}$$

Exact ca în ACADO, setăm $\Delta T = 0.5$ și rulăm programul pentru aceleași două cazuri de puncte de echilibru.

Observăm că atât stările, cât și variabila de intrare ating valorile pe care dorim să le obținem, independent de punctele de echilibru pe care le stabilim.

În concluzie, comparăm valorile obținute cu ACADO Toolkit și cu algoritmul de Control Optimal pe care l-am creat.

În ceea ce privește timpul de rulare (CPU time), algoritmul pe care l-am construit în Matlab este mult mai rapid decât cel creat cu ACADO Toolkit (200 de secunde față de 2 zile complete). În ceea ce privește stabilizarea, este clar din graficele din secțiunea anterioară că algoritmul din ACADO stabilizează mai rapid decât controlul optimal; ambii algoritmi sunt capabili să atingă punctele de echilibru pentru stările și variabila de intrare pe care le-am stabilit.

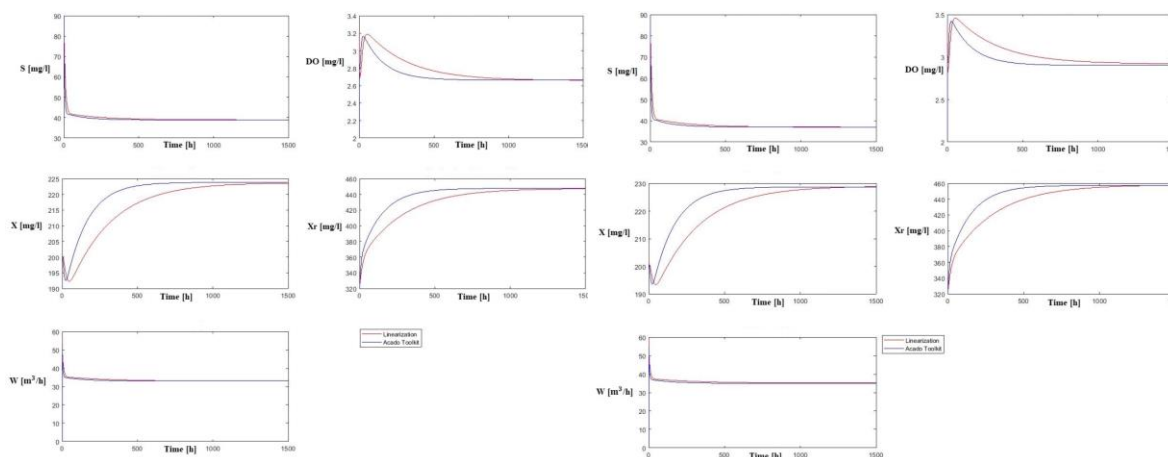


Fig. ACADO vs. Linearization pentru Cazul 1.11

Fig. ACADO vs. Linearization pentru Cazul 1.12

1.3. Supravegherea și Controlul Procesului în Procesele de Tratare cu Membrană

Datele mari disponibile și îmbunătățirea rapidă a puterii de calcul au atenuat progresul recent în modelele de optimizare numerică, permițând optimizarea în spații hiperdimensionale unde soluțiile analitice sunt prea costisitoare.

Viitorul digitalizării designului și optimizării materialelor membranelor implică o reconstrucție și generare a structurii cu ajutorul inteligenței artificiale, prezicerea proprietăților și a performanței în serviciu pentru materialele membranelor.

Procesul MBR are două dezavantaje principale: încărcarea membranei și consumul mare de energie. Cea mai mare parte a energiei necesare este pentru curățarea cu aer. Prevenirea încălcării membranei și consumul de energie în MBR pot fi echilibrate prin:

- monitorizarea permeabilității membranelor în timp real
- reglarea fluxului de curățare cu aer în consecință

Se poate obține o reducere medie a ratei de curățare cu aer de 13% - 20%, corespunzătoare unei economii de energie de 14% - 22% (SmartAirMBR).

Ultrafiltrarea dinamică permite atenuarea problemelor legate de încărcarea membranei și permite operarea cu fluxuri mai mari decât în cazul operațiilor convenționale cu flux transversal. Cu toate acestea, încărcarea nu este complet evitată, astfel încât este necesară o estimare a fluxului pentru determinarea celor mai bune scenarii de operare. Datorită înțelegerii limitate a sistemului dinamic, se propun modele de învățare automată. Metodele de învățare automată aplicabile includ: rețele neuronale, arbori de decizie, random forest, procese Gaussiene de regresie și metode clasice precum spațiu-stare și ARMAX (Autoregressive Moving Average with Exogenous Inputs - Medie Mobilă Autoregresivă cu Intrări Exogene). Noutatea abordării constă în combinarea propusă a mai multor modele pentru a obține o predicție cu o incertitudine redusă.

Selectarea materialelor optime poate fi realizată prin intermediul învățării automate. Învățarea automată poate prezice parametrii inerenti ai membranei (respingere sau flux) folosind doar informații moleculare structurale. Informațiile moleculare structurale ale soluțiilor, solventilor și membranelor pot fi folosite pentru interpretarea caracteristicilor soluției, solventului și membranei care afectează respingerea și fluxul. Prin utilizarea rețelelor neuronale grafice, se poate modifica efectul grupurilor funcționale, inelelor sau chiar a legăturilor și atomilor individuali dintr-o moleculă - ducând la structuri optime. O bază de date (www.osndatabase.com) prezintă caracteristicile a peste 500 de compuși chimici diferiți în 10 solvenți diferiți și patru membrane diferite, împreună cu unele modele.

Chimia polimerilor joacă un rol vital în performanța separării membranei, procesabilitatea în solvenți și capacitatea de a captura selectiv gaze precum CO₂. Parametrii de performanță pot fi variabili precum solubilitatea, difuzivitatea, permeabilitatea, selectivitatea ideală și multicomponent, umflarea și miscibilitatea în solvenți comuni. Sistemele bazate pe învățarea automată pentru screening rapid pot funcționa chiar și cu informații foarte limitate, cum ar fi chimia moleculară.

Materialele poroase pentru energie sunt componente esențiale ale multor dispozitive și sisteme energetice, dezvoltarea cărora a fost mult timp afectată de două provocări principale:

- "blestemul dimensionalității", adică relațiile complexe dintre structură și proprietăți ale materialelor energetice sunt în mare măsură determinate de un spațiu de parametri multidimensional.
- eficiența scăzută a tehnicilor de optimizare/descoperire pentru noi materiale energetice.

Soluția constă în digitalizarea materialelor energetice poroase, transformând toate informațiile despre materiale în spațiul digital prin utilizarea datelor de reconstrucție și imagini și integrându-le cu diverse metode computaționale. Caracterizarea rapidă, predicția proprietăților și optimizarea autonomă a noilor materiale de membrane pot fi realizate prin utilizarea algoritmilor matematici avansați combinați cu diverse instrumente de rețele neuronale artificiale (ANN).

Detectarea/predicția încărcării membranei poate fi realizată prin spectroscopie de amprentă. Monitorizarea online a potențialului de încărcare permite controlul inteligent pentru gestionarea încărcării. Metodele spectroscopice sunt susceptibile să fie un concept cheie, dar încă nu s-au definit indicatorii. Au fost scanate 11 indicatori UV-Vizibili, 19 Matrice de Emisie-Excitație în Fluorescență (FEEM) și 2 indicatori de dispersie în infraroșu pe o perioadă de 151 de zile. S-a constatat că blocarea porilor membranei și potențialul de încărcare cu gel/tortă sunt bine corelate cu UV-VIS și 15 FEEM. Scanarea UV-VIS și FEEM poate fi integrate într-un model pentru a prezice și detecta încărcarea. Scanarea în timp real a UV-VIS și FEEM permite acum detectarea mai rapidă și chiar în timp real.

Un protocol de avertizare timpurie pentru curățarea membranei prin predicția, diagnosticarea și generarea de avertismente (de exemplu, fenomenele de bioîncărcare în instalațiile MBR) poate fi implementat. Progresul bioîncărcării a fost prevăzut recursiv utilizând metoda filtrului Kalman - pentru identificarea mecanismului dominant de încărcare, prin incorporarea unui algoritm genetic. Regula de avertizare pentru curățarea membranei se bazează pe controlul sumei cumulate de încărcare - semnale de alarmă pentru eșecul operațional în instalația țintă.

2. Senzori virtuali

2.1. Integrarea Rețelelor Neuronale Artificiale pentru Estimare și Control în Tratarea Apelor Uzate

Studiul curent se concentrează pe implementarea unui Soft Sensor bazat pe Rețele Neuronale Artificiale (ANN) în scopul prognozării concentrațiilor de poluanți în Stațiile de Epurare a Apelor Uzate (WWTPs). Scopul principal al soft sensorului luat în considerare este facilitarea strategiilor de control, asigurându-se că nivelurile de poluanți rămân în limitele prescrise. Precizarea precisă a încălcărilor limitelor de deversare și gestionarea proactivă a consecințelor lor poate duce la reducerea cheltuielilor operaționale și la îmbunătățirea performanței generale a stațiilor de epurare a apelor uzate.

Cercetarea în desfășurare implică utilizarea unei Rețele Neuronale Artificiale (ANN) pentru o aplicație specifică. Intrările în ANN includ măsurătorile influentului și efluentului din Stațiile de Epurare a Apelor Uzate (WWTPs), colectate pe o perioadă de un an, cu un interval de eșantionare de 15 minute. Măsurătorile au fost obținute prin utilizarea BSM2M, folosind un model simulat pe computer al unei stații de epurare versatilă a apelor uzate (vezi Figura 2.1). Structura instalației este identică cu cea stației de tratare a apelor uzate de la nivelul Municipiului Galați. O înțelegere mai detaliată a problemei poate fi găsită în (Satin et al., 2016; Pisa et al., 2021; Pisa, 2022).

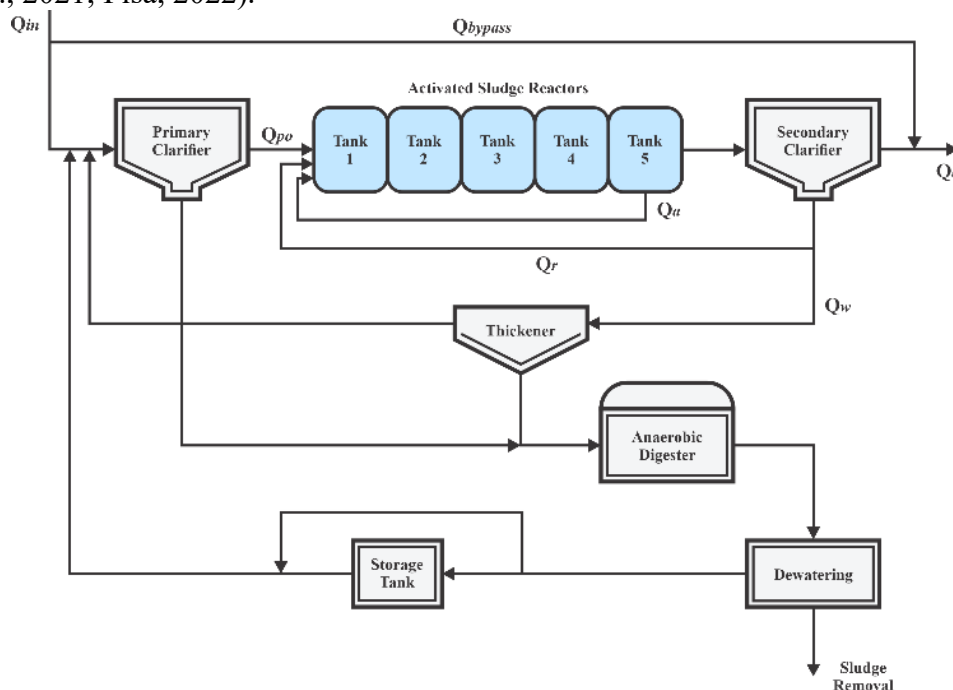


Fig. 2.1. Arhitectura BSM2 a unei Stații de Tratare a Apelor Uzate, afișând strategii de control standard (Default Control). Q_{po} reprezintă deversarea primară a clarificatorului, Q_a și Q_r fluxurile de recirculare interne și externe, iar Q_{in} și Q_e influentul și respectiv efluentul

În Figura 2.2, ilustrăm problema investigată. Intrările în sistem includ influentul și măsurătorile disponibile obținute de la Stațiile de Epurare a Apelor Uzate (WWTPs). Aceste măsurători includ fluxurile de apă și concentrațiile de nutrienți, cum ar fi concentrația de amoniu în al cincilea rezervor bioreactor ($S_{NH,5}$), fluxul de ieșire din primul clarificator (Q_{po}), temperatura ambientală (T_{as}), și totalul solidelor în suspensie (TSS) (Barbu et al., 2018). Scopul principal utilizează o Rețea Neuronală Artificială (ANN) pentru a prezice concentrațiile de efluent (\hat{y}_t).

Intrările alese pentru această aplicație sunt analizate pentru a determina capacitățile lor predictive, și se bazează pe o metodă de predicție care implică utilizarea a două celule two-stacked Long-Short Term Memory (LSTM).

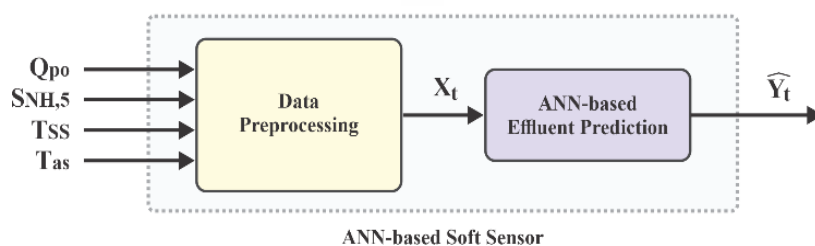


Fig. 2.2. Soft Sensor bazat pe Rețele Neuronale Artificiale (ANN). Datele utilizate pentru intrare și ieșire sunt reprezentative pentru măsurătorile influentului și respective ale efluentului.

Aceste celule sunt un tip de rețea neuronală artificială caracterizată de mecanismul său de gating. Conform rezultatelor cercetărilor prezentate în referințe (Goodfellow et al., 2016; Pisa, 2022), fiecare celulă LSTM este formată din 75 de neuroni ascunși per ANN care este inclusă în celulă.

Analiza este centrată pe examinarea concentrațiilor de amoniu ($S_{NH,e}$) care se numără printre concentrațiile de nutrienți poluanți cel mai des întâlnite în stațiile de epurare a apelor uzate (WWTPs). Au fost observate apariții rare de vârfuri în concentrațiile de amoniu ($S_{NH,e}$). Activitatea de predicție realizată de Soft Sensor bazat pe Rețele Neuronale Artificiale prezintă o provocare semnificativă din cauza incidenței rare a acestor evenimente și a distribuției neechilibrate a datelor.

Pregătirea datelor

Optimizarea performanței și reducerea complexității în Rețelele Neuronale Artificiale (ANNs) depind în mare măsură de implementarea tehnicilor de preprocesare a datelor (Naduvil-Vadukootu et al., 2017). Problema datelor neechilibrate în procesele industriale reprezintă o provocare pentru abordările convenționale în tratarea problemelor de regresie. Pentru a aborda aceste provocări, s-a dezvoltat o nouă abordare a preprocesării datelor. Acest set inovator de mecanisme va fi examinat în detaliu în secțiunile ulterioare, unde caracteristicile și avantajele lor unice vor fi explicate în detaliu.

A. Fereastră alunecătoare (Sliding Window)

Propunem implementarea unei metodologii noi de fereastră alunecătoare pentru a îmbunătăți aranjarea și menținerea coerenței temporale în cadrul setului de date. Metodologia actuală include două variabile fundamentale, și anume lungimea ferestrei (Window Length - WL) și orizontul de predicție (Prediction Horizon - PH). În studiul curent, configurația WL și PH a fost stabilită după cum urmează:

- O lungime a ferestrei (WL) de 10 ore a fost selectată ca un cadru temporal adecvat pentru a păstra valorile observate la fiecare moment de eșantionare și pentru a include măsurătorile

anterioare. Această abordare are ca scop capturarea unui context istoric amplu al datelor analizate.

- S-a luat în considerare un orizont de predicție (PH) de 4 ore. Acest parametru definește perioada în care predicțiile concentrațiilor de efluent pot fi furnizate în avans, facilitând astfel luarea deciziilor proactive.

Utilizarea mecanismului de fereastră alunecătoare într-un sistem ANN permite integrarea în mod eficient atât a măsurătorilor curente, cât și a celor anterioare, valorificând astfel corelația temporală semnificativă prezentă în date.

Tehnica fereastră glisantă este implementată într-un mod în care se generează o măsură nouă la fiecare deplasare a ferestrei, în timp ce cea mai veche măsură este eliminată conform principiului First-In-First-Out (FIFO). După cum este ilustrat în Figura 2.3, sistemul asimilează datele înregistrate în precedentele 10 ore pentru fiecare măsurare nouă. Specificațiile arhitecturale ale stației de epurare a apelor uzate (WWTP) luate în considerare au determinat cerințele de timp de retenție să fie de 14 ore. Configurația specifică a parametrilor ferestrei alunecătoare a fost analizată pentru a îndeplini aceste cerințe. Informații suplimentare despre arhitectura subiacentă pot fi găsite consultând referința (Satin et al., 2016).

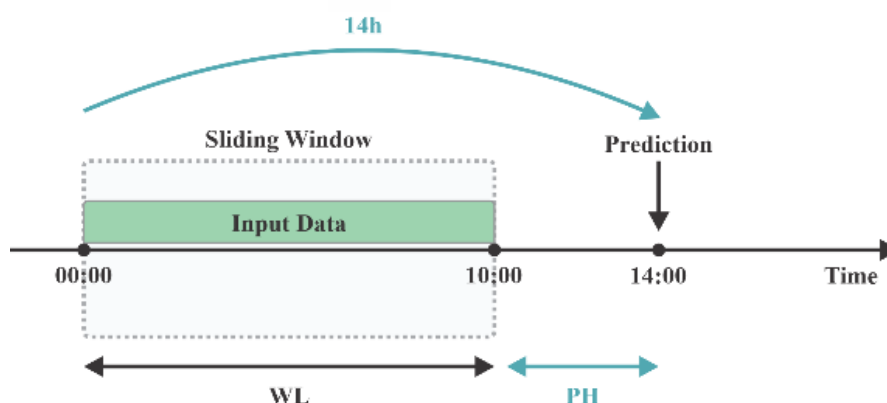


Fig. 2.3. Structura unei ferestre alunecătoare. Timpul de retenție al stației de epurare a apelor uzate (WWTP) este egal cu suma dintre Lungimea Ferestrei (WL) și Orizontul de Predicție (PH).

B. Normalizarea Datelor

Propunem utilizarea tehnicilor de normalizare a datelor ca o soluție potențială pentru abordarea issue-ului de heterogenitate a datelor. Metoda de ajustare a scorului Z este una dintre modalitățile prin care facem acest lucru. Atunci când apar evenimente rare, acestea pot fi susceptibile la bias (Garcia et al., 2015). Procesul de normalizare a scorului Z este calculat astfel:

$$x_t = \frac{x - E[x]}{\sqrt{E[(x - E[x])^2]}} \quad (2.1)$$

Astfel, se arată datele care trebuie normalizate și datele care au fost normalizate. Prin utilizarea acestei metode de normalizare, sperăm să standardizăm modul în care datele sunt distribuite și

să reducem efectul valorilor extreme, făcând intrarea în ANN mai echitabilă.

C. Antrenare bazată pe K-Fold

Când lucrăm cu seturi de date mari de serii temporale, nu este neobișnuit să întâlnim date dezechilibrate, iar semnalele de efluent sunt un exemplu elocvent al acestei disparități. Neprezentarea adecvată a problemei datelor neechilibrate poate duce la predicții distorsionate în favoarea valorilor reprezentate mai frecvent, în special a celor care cad sub anumite praguri.

Pentru a aborda această provocare, metoda K-Fold a apărut ca o tehnică valabilă de preprocesare a datelor în timpul etapei de învățare (Barbu et al., 2018). Această abordare funcționează pe baza a două principii fundamentale: împărțirea setului de date în subseturi de dimensiuni egale și executarea proceselor de antrenament. În cazul nostru specific, setul de date cuprinde măsurătorile influentului și ale efluentului din modelul BSM2 al stației de epurare a apelor uzate (WWTP).

Un număr de 5 K-Fold-uri a fost ales cu grijă pentru a alocă 70% din setul complet de date pentru antrenarea rețelelor neuronale artificiale (ANNs), în timp ce se rezervă 30% pentru teste și validare. În cadrul acestui 30%, 15% este destinat validării, în timp ce restul de 15% servește ca subset de testare. Scopul este de a obține modele distincte de predicție prin fiecare proces de antrenament, rezultând într-un total de modele. Setul de date este utilizat pentru fiecare model, cu submulțimi utilizate pentru antrenament și o submulțime dedicată testării și validării.

În cele din urmă, modelul care prezintă cea mai mare precizie de predicție în toate procesele de antrenament este selectat pentru aplicația finală. O reprezentare vizuală a metodologiei K-Fold este prezentată în Figura 2.4..

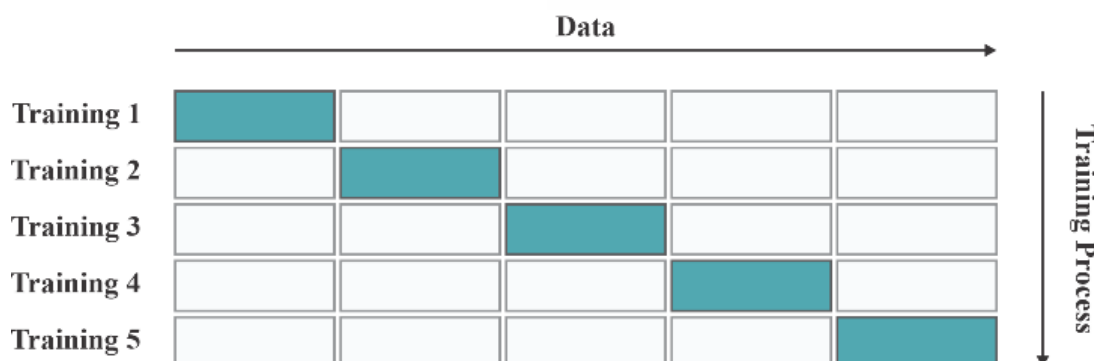


Fig. 2.4. Metoda K-Fold o tehnică frecvent utilizată în învățarea automată pentru validarea și selecția modelelor. Datele colorate.

Evaluarea soft sensor-ului

Evaluarea performanței (Figura 2.5) a tehnicilor propuse de preprocesare a datelor pentru date industriale de tip serie temporală se bazează pe predicțiile generate de Soft Sensor-ul bazat pe ANN. Prezentul studiu utilizează trei metrici distincte pentru a evalua performanța modelului. Aceste metrici include Eroarea Medie Patratică - Root Mean Squared Error (*RMSE*), Eroarea Medie Procentuală Absolută - Mean Absolute Percentage Error (*MAPE*), și Coeficientul de Determinare - Coefficient of Determination (R^2).

RMSE este o metrică frecvent utilizată în analiza statistică pentru a măsura diferența dintre

valorile prezise și cele reale. Este calculată cu formula următoare:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (2.2)$$

unde \hat{y}_i reprezintă predicția de la pasul i a concentrației efluentului, iar y_i denotă valoarea țintă corespunzătoare. O valoare mai mică a RMSE indică predicții mai precise.

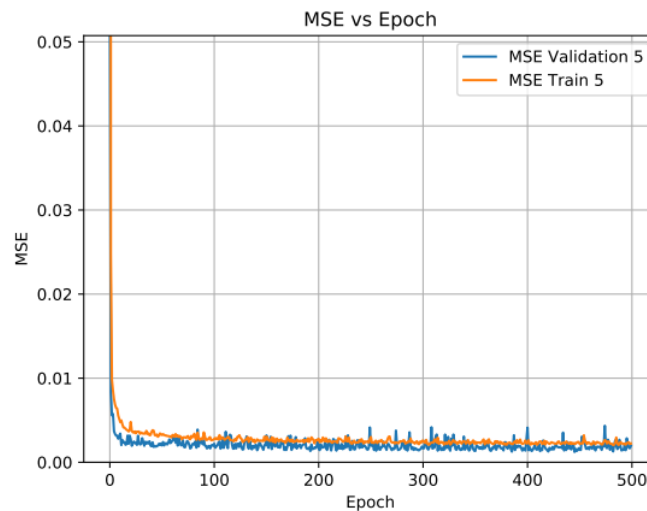


Fig. 2.5. Evaluarea performanței pentru al cincilea K-Fold, care prezintă rezultatul ideal pentru setul de date aplicat (ANN cu TSS), și valorile optime ale MSE pentru datele noastre..

MAPE este o altă metrică frecvent utilizată în domeniul modelării predictive. Este calculată luând media erorilor procentuale absolute în procesul de predicție:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100 \quad (2.3)$$

Această metrică este adesea preferată față de alte metrice de eroare, deoarece oferă o înțelegere mai intuitivă a exactității modelului.

Coeficientul de determinare, cunoscut în mod obișnuit sub denumirea de R^2 , este o măsură statistică care evaluează capacitatea unui model ANN de a explica variabilitatea observată în date. Această metrică cuantifică proporția variației totale a variabilei dependente care poate fi explicată de variabilele independente incluse în model. Calculul se realizează prin utilizarea unei formule specifice:

$$R^2 = \frac{(\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}}) \cdot (y_i - \bar{y}))^2}{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2 \cdot \sum_{i=1}^N (y_i - \bar{y})^2} \quad (2.4)$$

Simbolul $\bar{\hat{y}}$ reprezintă media valorilor prezise, iar \bar{y} reprezintă media valorilor țintă. Această

notație este frecvent utilizată în analiza statistică și modelele de învățare automată.

Combinarea acestor metrici oferă o evaluare detaliată a eficacității Soft Sensor-ului bazat pe Rețele Neuronale Artificiale (ANN).

Rezultatele evaluării indică faptul că modelul ANN prezintă o precizie remarcabilă în predicțiile sale. Studiul actual a dezvoltat un model capabil să explice aproximativ 97% din variația datelor, așa cum este indicat de valoarea R^2 . Acest nivel ridicat de variație explicat de model a condus la predicții precise, susținute și de valorile scăzute ale RMSE și MAPE. Mai precis, modelul ANN a generat valori RMSE și MAPE de 0.048 și, respectiv, 3.54%. Aceste constatări sugerează că modelul dezvoltat este foarte precis și de încredere în predicția rezultatelor de interes. Precizia predicțiilor este demonstrată în Figura 2.6 și Figura 2.7, în ciuda prezenței unor predicții nefavorabile, unde ANN afișează o precizie remarcabilă în prezicerea valorilor S_{NH} . Studiul actual evidențiază faptul că îmbunătățirile menționate anterior se bazează pe prognoze derivate din setul complet de date pe o perioadă de un an ca intrare.

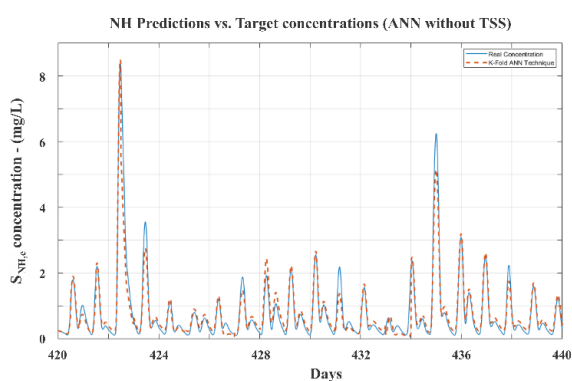


Fig. 2.6 Predicții ale $S_{NH,e}$ utilizând soft sensor-ul bazat pe ANN și implementare K-Fold fără să luăm în considerare TSS pentru rețeaua ANN antrenată.

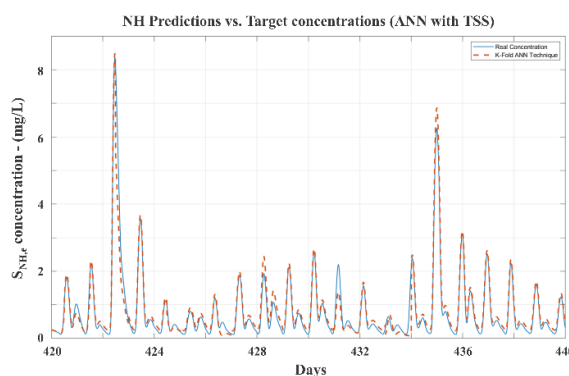


Fig. 2.7 Predicții ale $S_{NH,e}$ utilizând soft sensor-ul bazat pe ANN și implementare K-Fold, luând în considerare TSS ca intrare.(Fig.2.2).

2.2.Îndeplinirea cerințelor legislative în tratarea apelor uzate cu ajutorul instrumentelor digitale

Directiva privind tratarea apelor uzate urbane (Urban Wastewater Treatment Directive - UWWTD) solicită statelor membre să elimine fosfații și materia organică din apele uzate menajere în funcție de dimensiunea stației de epurare a apelor uzate (WWTP) și de sensibilitatea mediului înconjurător al locației (Directiva Consiliului 91/271/CEE din 21 mai 1991). Materia organică din apele uzate, măsurată în mod obișnuit ca Chemical Oxygen Demand - Consumul Chimic de Oxigen (COD) sau Biochemical Oxygen Demand - Consumul Biochimic de Oxigen (BOD), poate fi eliminată eficient cu ajutorul proceselor biologice de tratare a apelor uzate. În timp ce procesul chimic de tratare - coagularea - este extrem de eficient în eliminarea fosfaților, el nu elimină fracțiunile dizolvate de COD sau BOD din apele uzate (Ratnaweera and Fetting, 2015). Aceasta a generat o provocare semnificativă pentru stațiile de tratare care nu dispun de etape biologice.

Deși Norvegia nu este membră a Uniunii Europene, este obligată să îndeplinească cerințele Directivei privind Gestionarea Apelor Urbane. 23% din populația Norvegiei este conectată la stații de epurare chimică, 42% la stații de epurare chimico-biologică și doar 9% la stații de

epurare biologică pură (<https://www.ssb.no/en>). Acest lucru a generat ample discuții cu privire la îndeplinirea Directivei UE privind Gestionarea Apelor Urbane, iar ca măsură intermediară, autoritățile de mediu norvegiene au permis stațiilor lor de epurare să nu se concentreze pe cerința de eliminare a COD/BOD. Cu toate acestea, în ultimii ani, această practică a fost modificată, iar toate stațiile de epurare din zonele cu peste 10 000 de locuitori sunt acum obligate să respecte Directiva UE privind Gestionarea Apelor Urbane.

Directiva privind Gestionarea Apelor Urbane (UWWd) impune o eliminare de peste 75% sau o concentrație sub 125 mg/l COD în efluent și o eliminare de peste 70% sau o concentrație sub 25 mg/l BOD în efluent (Directiva Consiliului 91/271/CEE din 21 mai 1991). Deoarece coagularea poate înlătura fracția suspendată a COD/BOD, iar această fracție este în mod obișnuit între 40-60%, percepția generală este că este imposibil să se atingă cerințele UWWd. Consecința este extinderea stațiilor de epurare existente din Norvegia cu etape biologice, ceea ce va costa sute de milioane de euro.

Cu toate acestea, deoarece fracția pur suspendată, măsurată sub formă de solide suspendate (SS) mai mari de 1 μm , coagularea poate elimina eficient coloizii, care sunt mai mari de 10 μm . Acest lucru ar putea deschide noi posibilități pentru o eliminare extinsă a COD/BOD pentru a atinge nivelurile de eliminare necesare.

În timpul procesului de coagulare, se adaugă la apele uzate unele săruri anorganice, adesea aluminiu sau fier. Procesul rezultat de hidroliză cu fosfați elimină coloizii și particulele, în timp ce reacțiile chimice elimină fosfații.

Eficiența procesului de coagulare depinde critic de doza de coagulant, care trebuie să fie optimă pentru calitatea apei de intrare în orice moment, iar aceasta se schimbă rapid pe parcursul zilei. Doza optimă de coagulant depinde de debitul de apă uzată, pH-ul său, conținutul de particule și fosfați. Cu toate acestea, în practică, se ia în considerare numai debitul de intrare atunci când se calculează doza, uneori cu un pH de suprasciere pentru a asigura un proces de hidroliză optim. Această practică duce la rezultate suboptimale din cauza subdozării sau supradozării.

Deși necesitatea de a lua în considerare particulele și fosfații în plus față de debit și pH este bine recunoscută, lipsa senzorilor accesibili și preciși pentru a măsura acești parametri în timp real, precum și lipsa modelelor conceptuale eficiente pentru coagularea apelor uzate, au împiedicat implementarea lor.

Acest raport prezintă experiența cu utilizarea sistemelor surogat pentru monitorizarea calității apei și utilizarea acestora în controlul procesului pentru a îndeplini cerințele privind COD/BOD într-o instalație de coagulare, eliminând necesitatea extinderii cu o etapă biologică.

Descrierea procesului

Experimente la scară reală au fost efectuate la stația de epurare Søndre Follo (SFR) din Vestby, situată la 60 km sud de Oslo. SFR tratează ape uzate provenite de la 24 000 de persoane și procesează între 3500 și 9000 m³/zi, în funcție de vreme și precipitații. Permisul actual de deversare cere o eliminare de peste 90% a fosfatelor totale, în timp ce planul trebuie să îndeplinească cerința de tratare secundară. SFR este o instalație de tratare mecanico-chimică.

La stație au fost instalați mai mulți senzori online fizici care măsoară debitul, solidele suspendate, conductivitatea, pH-ul și temperatura la intrare, pH-ul după coagulare și solidele suspendate după sedimentare (efluentul). Datele de supraveghere sunt colectate la fiecare 10 minute și stocate în sistemul SCADA. Fosfații totali au fost măsurați în probe medii zilnice la stația de epurare și săptămânal de laboratoare acreditate. COD și BOD au fost măsurați în 24 de probe zilnice de laboratoare acreditate anual, așa cum cere legislația. Aceste date au fost

completate de mai multe campanii de prelevare care au măsurat probe orare în timpul vremii uscate și umede, și în zilele de săptămână și weekend, la laboratorul DOSCON. Probele au fost analizate folosind Metodele Standard Norvegiene. Rezultatele analitice au fost validate încrucișat folosind mai multe metode analitice.

Datele analitice au fost utilizate pentru a calibra senzori hibridi pentru fosfați totali și COD total și BOD total, folosind un concept prezentat de (Nair et al., 2022). Datele de supraveghere de la senzorii fizici și senzorii software au fost integrate și utilizate pentru analize suplimentare.

Doza de coagulant a fost determinată folosind un concept dezvoltat de DOSCON, utilizând statistici multivariate combinate (Manamperuma et al., 2017), care a fost îmbunătățit ulterior cu ajutorul Rețelelor Neuronale Artificiale. Conceptul ia în considerare toți parametrii critici pentru procesul de coagulare, spre deosebire de dozarea proporțională la debit cu anulare a pH-ului, astfel încât particulele și fosfații sunt incluși direct sau indirect.

Acuratețea software sensor-ului

Chiar dacă conceptul DOSCON a asigurat o eliminare de peste 90% a fosfaților, performanța SFR în ceea ce privește materia organică (COD și BOD) nu era conformă cu legislația. Pentru a optimiza în continuare eliminarea de COD/BOD, am luat în considerare integrarea supravegherii fosfaților totali și COD total în algoritmul de dozare a coagulantului.

Figurile 2.8 și 2.9 prezintă precizia modelului pentru senzorii hibridi pentru fosfați totali și COD total. În aceste figuri, folosim R^2 ca indicator al exactității predicției față de valorile măsurate (Date reale). Predicțiile au fost satisfăcătoare având în vedere variațiile semnificative în calitatea apei uzate de intrare.

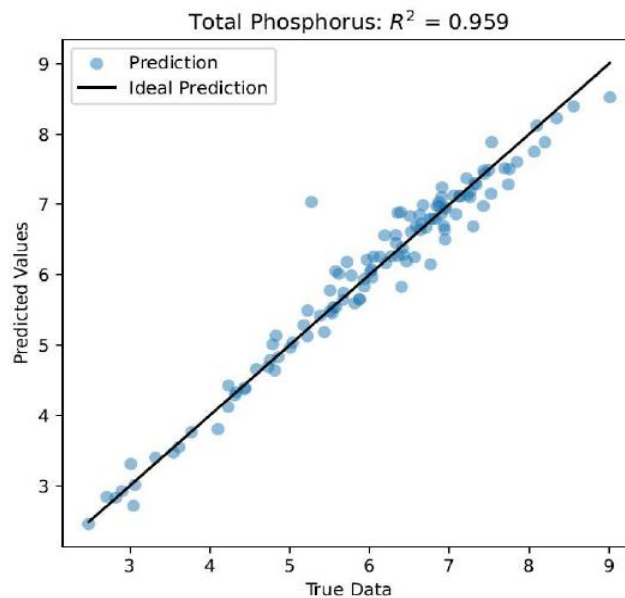


Fig 2.8. Acuratețea de predicție a modelului pentru Total-P

Prin ajustări suplimentare ale algoritmilor de dozare, instalația a înregistrat performanțe conforme cu cerințele de tratare secundară, așa cum sunt prezentate în Figurile 2.10, 2.11 și 2.12, pentru total-P, COD și BOD, respectiv.

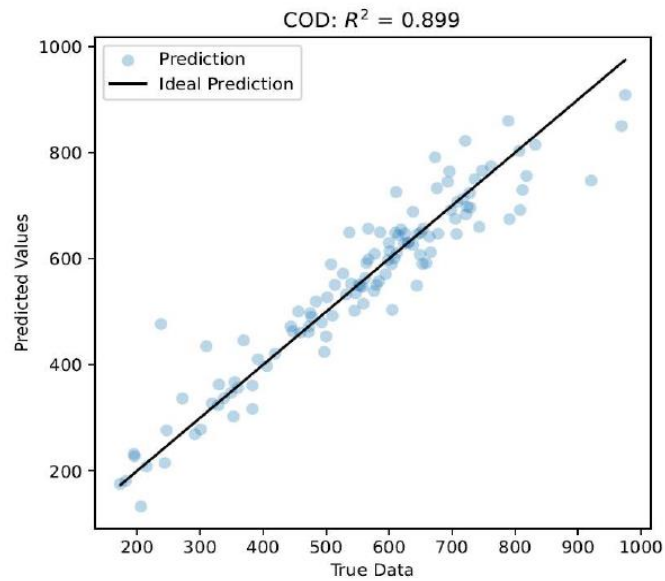


Fig 2.9. Acuratețea de predicție a modelului pentru total-COD

Precizia modelelor obținute pentru cei doi parametri a fost considerată satisfăcătoare și a fost integrată în optimizarea ulterioară a determinării dozei de coagulant.

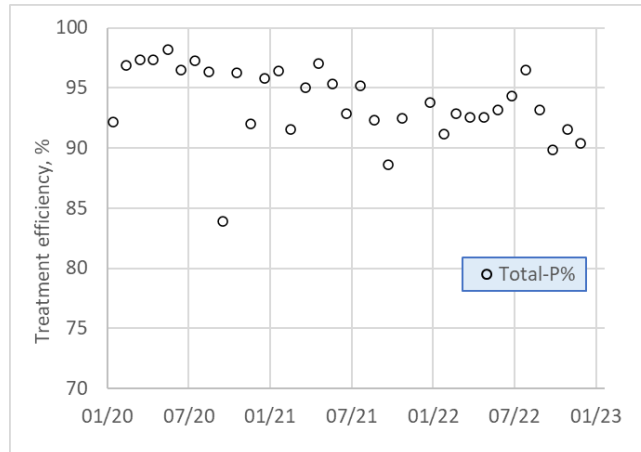


Fig 2.10. Eficiența de înlăturare a Total-P. Cerința de conformitate este >90%.

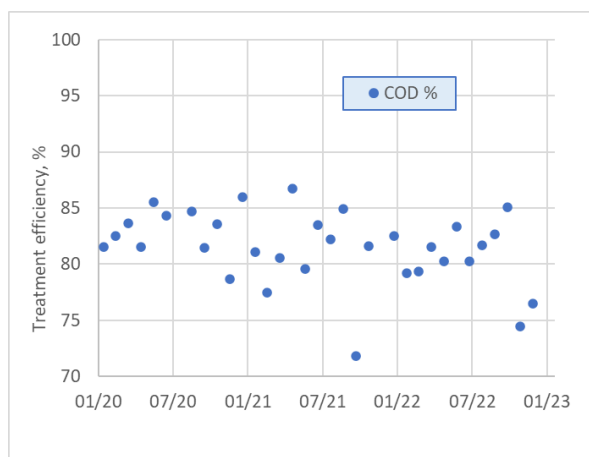


Fig 2.11. Eficiența de înlăturare a COD. Cerința de conformitate este >75%

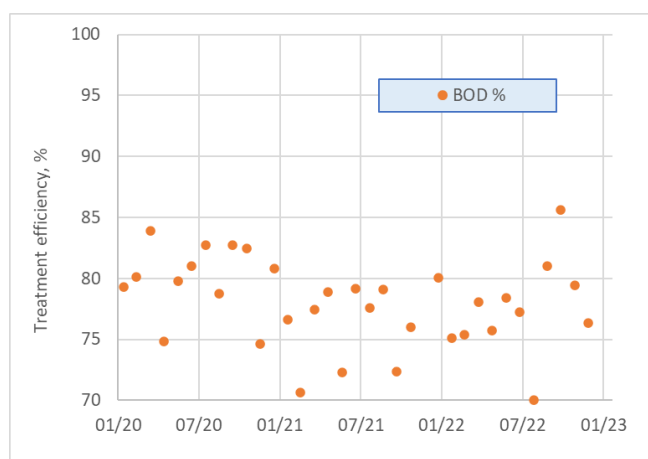


Fig 2.12. Eficiența de înlăturare a BOD. Cerința de conformitate este >70%

Calitatea apei influente și apei uzate variază rapid pe parcursul zilei și numai cu măsurători în timp real, se poate răspunde rapid pentru a aborda orice slăbiciuni ale performanței. Prin urmare, este valoros să avem supraveghere în timp real cu privire la parametrii de conformitate, cum ar fi total-P, COD și BOD. În timp ce există senzori fizici pentru COD pe piață, nu există senzori similari pentru fosfați. Singura opțiune aproape în timp real a fost să se utilizeze analizoare online. Astfel de echipamente sunt destul de scumpe (15 000-50 000€ fiecare), în timp ce senzorii hibridi oferă o alternativă bună pentru o fracțiune din cost. Figurile 2.8 și 2.9 confirmă validitatea modelelor folosite în senzorii hibridi și furnizează rezultate comparabile cu valorile de laborator acreditate.

Figura 2.8 prezintă precizia modelului sensorului hibrid pentru Total-P, care pare să fie bună. Figura 2.9 prezintă precizia modelului sensorului hibrid COD. Deși nu este la fel de bun ca modelul Total-P, îndeplinește scopul de a optimiza eficiențele de înlăturare a COD. Precizia mai scăzută a modelului este rezultatul datelor de supraveghere limitate disponibile pentru calibrarea modelului, și odată cu creșterea datelor de supraveghere, modelul devine mai bun și mai robust.

Senzorii hibridi sunt integrați în algoritmi operaționali de estimare a dozelor în timp real. În loc să avem solide suspendate ca parametru țintă, a fost introdus Total-P. Figura 2.10 prezintă valorile raportărilor de conformitate pentru Total-P, de două ori pe lună, timp de 2 ani. Acestea

sunt mult peste nivelurile de conformitate (>90%), cu excepția valorilor mai mici ocazionale, care sunt și ele permise. Investigații suplimentare asupra acestor condiții excepționale au relevat că acestea se datorează fie fluxurilor extraordinare care reduc drastic timpul de sedimentare, fie activităților de întreținere a instalației.

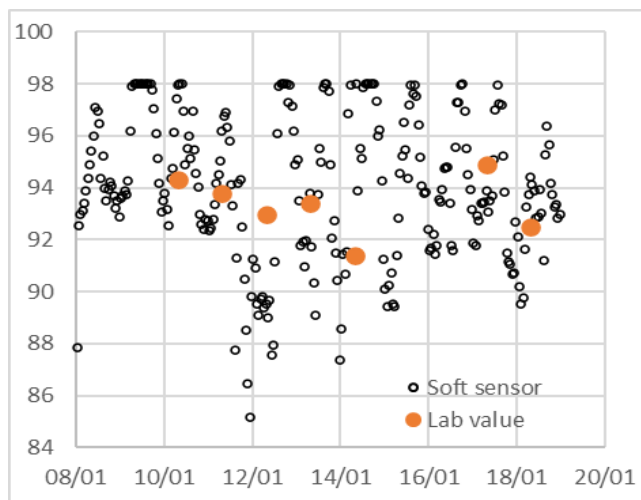


Fig 2.13. Eficiența eliminării Total-P cu date în timp real

Figura 2.13 prezintă datele în timp real pentru Total-P timp de 12 zile, comparate cu valorile de laborator. Valorile în timp real de la senzorii hibridi au arătat că eficiența tratamentului poate varia între 85-98%, conducând la rezultate medii de 91-95%. Aceste valori au fost afișate continuu la stația de epurare, făcând ca operatorii să fie mai conștienți de condiții, în timp ce echipa de modelare putea lucra mai concret pentru a reduce evenimentele cu rezultate mult sub nivelurile de conformitate. Lucrările sunt în curs de desfășurare pentru eliminarea eficiențelor de tratare prea scăzute și prea mari, iar supravegherea în timp real și senzorii hibridi virtuali facilitează acest proces.

Figura 2.11 prezintă eficiența de înlăturare a COD pe parcursul a 2 ani, bazată pe raportările de conformitate. Cerința de conformitate este fie >75% înlăturare, fie <125 mg/l în efluent. Figura 5 prezintă rezultate similare pentru BOD. Cu excepția a două ocazii în decurs de doi ani, rezultatele COD au fost în limitele de conformitate, în timp ce toate valorile BOD au fost în limitele permise.

Percepția generală este că realizarea cerințelor de tratare secundară pentru înlăturarea materiei organice este imposibilă doar cu coagulare. Cu toate acestea, rezultatele prezentate mai sus atestă capacitatea de a îndeplini cerințele de conformitate într-o instalație mecanico-chimică. Când fracțiunea suspendată a COD și BOD sunt în mod normal <60%, realizarea a >75% și 70% respectiv a necesitat o elaborare mai detaliată asupra mecanismelor subiacente.

Coagularea înlătură nu numai fracțiunea suspendată (>1 μ), ci și fracțiunea coloidală (0,1-1 μ m). Prin urmare, este rezonabil să presupunem că și fracțiunea coloidală a COD și BOD va fi înlăturată în timpul coagulării. Încă nu am analizat fracțiunea coloidală a COD în apele uzate de la stația de epurare SFR, dar analiza din literatură este prezentată în Tabelul 2.1.

Tabel 2.1. Frațiuni COD raportate în alte studii

Total COD	Suspended COD	Colloidal COD	Dissolved COD	Colloidal + suspended	Reference
699	409	71	219	69 %	(Tawfik et al., 2010)
733	485	64	185	75 %	
803	603	49	151	81 %	
334	137	39	80	76 %	(Drewnowski et al., 2020)
650	413	24	213	67 %	
630	442	36	152	76 %	
590	408	39	143	76 %	
548	384	33	131	76 %	
451	313	32	106	76 %	
821	630	37	154	81 %	
620	444	44	132	79 %	
1390	1143	73	174	87 %	
1001	765	64	172	83 %	
797	632	53	112	86 %	
750	558	44	148	80 %	

Conform datelor prezentate în Tabelul 2.1, fracțiunea totală suspendată și coloidală a apelor uzate poate atinge chiar și 87%. Prin urmare, pare logic să ne așteptăm la o înlăturare de COD >75%.

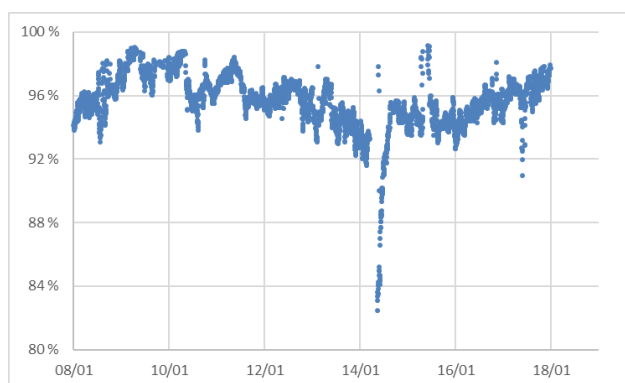


Fig 2.14. Eficiența de înlăturare a solidelor suspendate cu valori în timp real

Cu toate acestea, chiar și după un proces cuprinzător de destabilizare cu coagulare, înlăturarea finală a acelei fracțiuni va depinde de cuprinsul procesului de separare. Chiar și în situații bine funcționabile, ratele de înlăturare a solidelor suspendate variau între 92-98%. Figura 2.14 prezintă eficiențele de înlăturare a solidelor suspendate la stația de epurare SFR, care arată că eficiența de înlăturare a variat între 91-99%, cu excepția unei perioade mai scurte cu 82%. Prin urmare, pentru a asigura o înlăturare a COD >75%, este necesar să avem fracțiuni suspendate + coloidale de peste 80%.

Algoritmul de control al dozării la stația de epurare SFR se bazează pe solidul suspendat la intrare și ieșire. Având în vedere reflectările asupra înlăturării fracțiunilor coloidale în plus față de fracțiunea suspendată, am putea anticipa o prognoză chiar mai bună a dozei optime, dacă ambele fracțiuni ar putea fi luate în considerare. Cu toate acestea, măsurarea în timp real a

fracțiunilor coloidale ale particulelor sau materiei organice cu senzori fizici nu este realistă în prezent. Cu toate acestea, credem că ar fi posibil să o estimăm folosind un senzor hibrid soft. Particulele din apele uzate sunt măsurate atât ca turbiditate, cât și ca solide suspendate, așa că ar putea exista o posibilitate de a obține o estimare pentru suma fracțiunilor coloidale și suspendate folosind o combinație a acestor doi parametri, completată cu alți senzori fizici în dezvoltarea unui senzor hibrid soft. Dacă ar fi posibil să derivăm fracțiunea totală de coloid și suspendat doar cu un senzor de solide suspendate și un senzor de turbiditate este, de asemenea, o întrebare interesantă, care trebuie dezvoltată în viitor.

3. Analiza Incidentelor de Securitate Cibernetică în Sectorul Apei

Serviciile publice de apă sunt esențiale pentru sănătate, siguranță și bunăstare, fiind considerate infrastructuri critice (IC), ale căror perturbări pot duce la pierderi semnificative din punct de vedere economic, al siguranței publice și al mediului (Igre et al., 2006; Moraitis et al., 2020). Sistemele de control și supervizare a datelor - Supervisory Control and Data Acquisition (SCADA) furnizează capacități de monitorizare și control pentru operațiunile în timp real ale serviciilor publice de apă (Rasekh et al., 2016). În trecut, sistemele SCADA se bazau pe rețele izolate și protocoale non-standard pentru a le proteja împotriva atacurilor cibernetice (Igre et al., 2006). În prezent, aceste rețele au fost conectate la rețelele corporative și la internet. De asemenea, au existat progrese în utilizarea protocoalelor standard de rețea pentru comunicare (Cheung et al., 2006). Aceste schimbări au făcut ca sistemele SCADA să fie mai accesibile pentru atacatori, care le pot viza de la distanță prin Internet de oriunde în lume (Christiansson and Luijff, 2008).

Deși integrarea rețelelor SCADA și corporative a îmbunătățit semnificativ eficiența sistemelor, a crescut, de asemenea, suprafețele de atac și a expus infrastructurile fizice și cibernetice la atacuri (Housh and Ohar, 2018; Ramotsoela et al., 2019; Rasekh et al. 2016). De exemplu, un atac reușit în sectorul apei poate duce la contaminare chimică, deteriorare fizică sau perturbarea comunicațiilor între elementele de rețea și sistemul SCADA. Mai multe atacuri cibernetice, precum cele asupra instalației de apă din Harrisburg (2006), sistemului de apă din Northern Colorado (2019), sistemului de apă din Israel (2020), și stației de tartare a apelor reziduale din Florida (2021) (Shardeen et al., 2021) exemplifică problemele de securitate cibernetică cu care se confruntă sectorul apei. Numărul atacurilor cibernetice raportate în sectorul apei este în creștere, situându-l ca al treilea cel mai vizat sector (ICS-CERT 2016).

Înțelegerea diferitelor dimensiuni ale acestor incidente de securitate cibernetică și modul în care au evoluat în timp poate oferi perspectiva necesară pentru dezvoltarea de strategii eficiente de prevenire sau atenuare a atacurilor similare în viitor (Miller and Rowe, 2012). Hassanzadeh et al., 2020 au revizuit incidentele de securitate cibernetică în sectorul apei. În timp ce studiul lor contribuie semnificativ la înțelegerea naturii incidentelor de securitate cibernetică în sectorul apei, gama de opțiuni luate în considerare în schema lor de clasificare și numărul limitat de incidente de securitate cibernetică revizuite limitează capacitatea de a obține o înțelegere mai profundă a tehnicilor, tendințelor și modelelor legate de incidentele anterioare de securitate cibernetică (Hassanzadeh et al., 2020). În plus, analiza lor nu se corelează cu cadrele internaționale, cum ar fi MITRE's ATT & CK pentru ICS.

Incidente de Securitate Cibernetică

În această secțiune, vom examina incidentele de securitate cibernetică din sectorul apei și vom analiza mecanismele de atac (comportamentul adversarial) utilizând tactici și tehnici din MITRE ATT&CK (Alexander et al., 2020; Mitre, 2022). Tabelul 3.6 furnizează incidente cibernetice cuprinzătoare în sectorul apei, iar fiecare incident conține un nume, anul producerii, actorul amenințării confirmat/suspectat, tehnica și tipul atacului, componenta/stratul vizat și impactul.

1. Incidentul canalizării Maroochy Shire
 - a. Rezumatul incidentului

În martie 2000, Consiliul Maroochy Shire a întâmpinat probleme cu noul său sistem de apă uzată (Slay and Miller 2008). Comunicările trimise prin legături radio către stațiile de pompare ale apei uzate întâmpinau dificultăți. Pompele nu funcționau corect, iar alarmele care trebuiau să notifice inginerii sistemului despre defecțiuni nu se activau așa cum era de așteptat (Hemsley and Fisher 2018). Inițial, au crezut că problema se datora noului sistem, dar după un timp, au observat că sistemul fusese atacat. Un inginer care monitoriza fiecare semnal care trecea prin sistem a descoperit că cineva se infiltra în sistem și provoca în mod deliberat problemele. Serviciul public de apă a angajat o echipă de detectivi privați care l-au identificat pe atacator și au alertat poliția. Atacatorul, Vitek Boden, a fost arestat și ulterior închis; el era angajat al unei firme contractante care furniza tehnologie IT/sistem de control.

b. Mecanismul atacului

Domnul Boden a utilizat un laptop și un emițător radio pentru a prelua controlul a 150 de stații de pompare a apei uzate (Slay and Miller 2008). Pe o perioadă de trei luni, a eliberat un milion de litri de apă uzată netratată într-un canal de apă pluvială, de unde aceasta a ajuns în cursurile locale de apă. Atacatorul a fost localizat fizic de autorități și s-a constatat că avea asupra sa un laptop cu un software furat pentru reconfigurarea SCADA, împreună cu un radio bidirecțional Motorola M120 și dispozitive de control PDS. Dovezile recuperate de pe laptop au indicat și faptul că comenzile din programul sistemului au fost rulate cel puțin de 31 de ori, ceea ce se potrivea cu comportamentul observat în jurnalele companiei (Makrakis et al., 2021). Atacul a fost motivat de dorința de răzbunare din partea domnului Boden, după ce acesta nu a reușit să obțină un loc de muncă la Consiliul Maroochy Shire.

Tacticile și tehnicile adversarilor folosite în timpul acestui atac cibernetic sunt cartografiate folosind cadrul MITRE ATT&CK în Tabelul 3.1.

Tabel 1.1 Tactici și tehnici adversare folosite în incidentul de canalizare din Maroochy Shire

Tactic	Technique (ID)
Initial access	Wireless Compromise (T0860)
Execution	
Persistence	
<i>Privilege Escalation</i>	
<i>Defense evasion</i>	
<i>Credential Access</i>	
Discovery	Remote system discovery (T088) T0808
Lateral Movement	Default credentials (T0812)
Collection	Automated collection (T0802)
Command and Control	
Inhibit Response Function	
Impair Process Control	
Impact	

2. Incidentul barajului de apă Bowman Avenue

a. Rezumatul incidentului

Hackerii iranieni au obținut acces neautorizat la sistemele SCADA ale barajului de pe Bowman Avenue din New York în 2013, ceea ce le-a permis să adune informații despre

nivelurile apei, temperatura și starea ușii de deversare (Shimon et al., 2015). Barajul Bowman controlează valurile de furtună, iar sistemul său SCADA era conectat la internet printr-un modem celular (Hemsley and Fisher 2018). Acest acces i-ar fi permis atacatorului să opereze și să manipuleze ușa de deversare a barajului de la distanță. Cu toate acestea, în acest caz, ușa de deversare fusese deconectată manual în timpul lucrărilor de întreținere la momentul atacului.

b. Mecanismul atacului

Atacatorii au exploatat o vulnerabilitate pentru a identifica un computer neasigurat care controla ușile de deversare și alte funcții ale barajului. Atacatorul a detectat vulnerabilitatea prin "Google Dorking", un proces de efectuare a căutărilor avansate pe Google pentru a depista vulnerabilități, și apoi a implementat alte tehnologii pentru a exploata cu succes sistemul (Germano, 2019).

Tacticile și tehnicile adversarilor folosite în timpul acestui atac cibernetic sunt cartografiate folosind cadrul MITRE ATT&CK în Tabelul 3.2.

Tabel 3.2 Tactici și tehnici adversare utilizate pentru a desfășura incidentul de la barajul Bowman Avenue

Tactic	Technique (ID)
Initial access	Internet accessible device (T0883)
Execution	
Persistence	
<i>Privilege Escalation</i>	
<i>Defense evasion</i>	
<i>Credential Access</i>	
Discovery	
Lateral Movement	
Collection	
Command and Control	
Inhibit Response Function	Activate Firmware update(T0800)
Impair Process Control	Brute force(T0806)
Impact	Damage to property (T0879)

3. Incidentul companiei de apă Kemuri

a. Rezumatul incidentului

Soluțiile de securitate Verizon au raportat că o companie de apă nedeclarată a fost victima unui atac cibernetic asupra sistemelor sale de control industrial (ICS) de către un presupus grup de hacktiviști sirieni în 2016 (Andreeva et al. 2016; Hemsley and Fisher, 2018). Verizon a atribuit un pseudonim de "Kemuri" companiei de apă (KWC) pentru a-și proteja identitatea din motive de securitate. Evaluarea a avut loc după ce angajații au devenit suspicioși în legătură cu o intruziune datorită comportamentului neobișnuit al valorilor și conduitei (Makrakis et al. 2021). Atacatorul a reușit să manipuleze sistemul pentru a modifica cantitatea de substanțe chimice care intră în sistemul de alimentare cu apă și să afecteze capacitățile de tratare și producție a apei, determinând creșterea timpilor de recuperare a alimentării cu apă (Hemsley and Fisher 2018).

b. Mecanismul atacului

Stația KWC a utilizat un sistem SCADA bazat pe IBM AS/400 pentru gestionarea PLC-urilor în vederea reglării fluxului de apă și substanțelor chimice prin controlul supapelor și conductelor în instalație. Conform rapoartelor publicate de firma de securitate Vericlave (Vericlave, 2018) și alte surse (Adepu et al., 2020; Alladi et al., 2020) vectorii principali de atac utilizați în încălcarea securității sistemului AS/400 intern al KWC ar fi putut fi un atac de injectare Structured Query Language (SQL) și phishing prin email. Atacatorii au extras datele de conectare pentru sistem de pe serverul web frontal pentru a accesa software-ul de control al apei al instalației, care rula, de asemenea, pe același sistem AS/400. Deoarece acest sistem era central pentru majoritatea operațiunilor IT din această instalație, accesul la acest sistem de control le-a permis hackerilor să controleze cea mai mare parte a celorlalte echipamente din instalație (Desc4). Cu toate că atacatorii au reușit să manipuleze supapele care controlează fluxul chimic, nu s-a observat niciun impact asupra operațiunilor instalației. S-a raportat, de asemenea, că informațiile personale ale aproximativ 2,5 milioane de clienți au fost compromise și au fost scurse din baza lor de date (Alladi et al., 2020).

Tacticile și tehnicile adversare utilizate în timpul acestui atac cibernetic sunt afișate folosind framework-ul MITRE ATT&CK în Tabelul 3.3.

Tabel 3.3 Tactici și tehnici adversare utilizate pentru incidentul companiei de apă Kemuri

Tactic	Technique (ID)	Description
Initial access	-Exploit public facing application (T819) -Phishing (T1566)	-SQL injection on payment portal website, combined with spear-phishing for missing info
Execution		
Persistence		
<i>Privilege Escalation</i>		
<i>Defense evasion</i>	Exploitation for Evasion(T0820)	
<i>Credential Access</i>		
Discovery	<i>Remote system discovery (T0486)</i>	
Lateral Movement	-Valid Accounts (T0859)	- Webserver held files with plaintext credentials and IP of internal IBM AS/400 server
Collection		
Command and Control	Commonly used port (T0885)	
Inhibit Response Function	Activate firmware update (T0800)	
Impair Process Control	-Modify parameter (T0836)	-Attackers manipulated chemical flow twice, but alarms caused operators to intervene and prevent real impact (Modify parameter)
Impact	ATT&CK Enterprise (Exfiltration)	- Attackers exfiltrated 2.5 million customer data records

4. Incidentul sistemelor de tratare a apei din Israel

a. Rezumatul incidentului

În aprilie 2020, a avut loc un atac cibernetic asupra sistemelor de control industrial ale facilității de tratare a apei a Autorității de Apă din Israel - Israel Water Authority's water treatment facility (Kovacs, 2020b). Conform unei declarații din partea Direcției Naționale pentru Cibernetică a Israelului - Israel's National Cyber Directorate, atacul intenționat viza sistemele de comandă și control ale instalațiilor de tratare a apelor uzate, stațiilor de pompare și infrastructurii de canalizare a Autorității de Apă (Kovacs, 2020a). Atacatorul a obținut acces la sistemele de tratare a apei și a încercat să modifice nivelurile de clor din apă. Autoritățile israeliene au emis o alertă solicitând facilităților de tratare a apei afectate să-și reseteze imediat parolele și operatorilor de sisteme de control industrial să ia în considerare măsuri suplimentare de securitate pentru sistemele lor operaționale accesibile publicului.

b. Mecanismul atacului

Tacticile și tehnicile adversare utilizate în timpul acestui atac cibernetic sunt afișate folosind framework-ul MITRE ATT&CK în Tabelul 3.4.

Tabel 3.4 Tactici și tehnici adversare utilizate pentru incidentul cu sistemul de tratare al apelor în Israel

Tactic	Technique (ID)	Description
Initial access	-Internet Accessible Device (T883)	-Sites likely internet-connected using cellular gateways, rendering PLCs either directly or indirectly internet exposed
Execution		
Persistence		
<i>Privilege Escalation</i>		
<i>Defense evasion</i>		
<i>Credential Access</i>		
Discovery		
Lateral Movement	Default credentials (T0812)	- Reportedly no or default credentials on gateways and PLCs
Collection		
Command and Control	-Commonly Used Port (T0885) -Standard application layer protocols (T0869)	-Attacker communicated with PLCs using standard protocols on common ports (Siemens S7comm (102/TCP), Modbus TCP (502/TCP), GE SRTP (18245/TCP, 18246/TCP))
Inhibit Response Function		
Impair Process Control	-Modify Parameter (T836)	Attackers modified control logic and parameters
Impact	-Manipulation of control (T831)	-Attempted to increase chlorine levels

5. Incidentul sistemelor de tratare a apelor în Florida

a. Rezumatul incidentului

În februarie 2021, un actor de amenințare cibernetică neidentificat a încercat să manipuleze nivelurile de hidroxid de sodiu (lichid de curățare) în facilitatea de tratare a apei din orașul Oldsmar, Florida, SUA (FBI 2021; Mass 2021). A fost recunoscut public faptul că o mașină operator avea instalat un pachet de software de acces la distanță și era accesibilă prin Internet. Acest lucru a permis ca atacul să fie efectuat prin obținerea accesului la stația de lucru la distanță folosind TeamViewer și manipularea punctelor de control pentru rata de dozare a hidroxidului de sodiu (NaOH) în apa potabilă. Atacatorul a crescut punctul de setare a dozei de NaOH de la setarea normală de 100 de părți per milion (ppm) la 11.100 ppm - *The attacker raised the NaOH does setpoint from its normal setting of 100 part-per-million (ppm) to 11,100 ppm* (Serino and Miller, 2021). Operatorul a observat rapid cursorul mouse-ului deplasându-se pe ecran și a revenit nivelurile de hidroxid de sodiu la parametrii de operare normali, astfel încât alarmele de monitorizare a pH-ului să nu detecteze un nivel dincolo de parametrii acceptabili.

b. Mecanismul atacului

Cyber actorii au avut probabil acces la sistem prin exploatarea vulnerabilităților de securitate cibernetică, inclusiv o securitate deficitară a parolelor și un sistem de operare învechit. Conform avertizării comune de securitate cibernetică emise de FBI, Cybersecurity and Infrastructure Security Agency (CISA), Environmental Protection Agency (EPA) și Multi-State Information Sharing and Analysis Center (MS-ISAC) (FBI 2021), este posibil ca un software de partajare a desktopului, cum ar fi TeamViewer, să fi fost folosit pentru a obține acces neautorizat la sistem.

Tacticile și tehnicile adversare utilizate în timpul acestui atac cibernetic sunt afișate folosind framework-ul MITRE ATT&CK în Tabelul 3.5.

Tabel 3.5 Tactici și tehnici adversare utilizate pentru incidentul cu sistemelor de tratare a apei din Florida

Tactic	Technique (ID)	Description
Initial access	-Exploit public-facing application (T0819) -External remote services (T0866)	
Execution	Execution through Graphical User Interface (T0823)	
Persistence	Valid accounts (T0859)	
<i>Privilege Escalation</i>		
<i>Defense Evasion</i>		
<i>Credential Access</i>		
Discovery		
Lateral Movement		
Collection		
Command and Control		
Inhibit Response Function		
Impair Process Control	-Modify parameter (T0836) -Unauthorized command message (T0855)	

Impact	-Loss of safety (T0880)	-Safety of lives drinking that water would have been lost
	-Manipulation of control (T0831)	-the chemical level control put in place was manipulated.

6. Rezumatul incidentelor de securitate cibernetică în sectorul apelor

Tabelul 3.6 prezintă un rezumat al incidentelor de securitate cibernetică în sectorul apei. Tabelul evidențiază numele incidentului, anul și județul în care a fost raportat atacul. De asemenea, sunt listate componenta/stratul țintă (rețea corporativă - CN, rețea supervizorală - SN, rețea de control local - LN, rețea de teren sau de instrumentare - FN), impactul atacului, tehnica/metoda folosită pentru lansarea atacului și componenta vulnerabilă.

Tabel 3.6 Rezumatul incidentelor de securitate cibernetică în sectorul apelor

Cyber Incident Name	Year	Attack location	Target	Threat actor	ATT & CK technique (initial compromise)	Attack type
Salt River Project	1994	United States	CN	Individual	Internet accessible device (T0883)	Malware(backdoor)
Maroochy Shire Sewage Spill	2000	Australia	SN	Individual	Wireless compromise (T0860)	Disgruntled Employee
Baseline Audit Uncovers Virus in Water Control System	2003	Australia	LN	Unknown		Virus
Trojan Backdoor on Water SCADA System	2004	Canada	SN	Unknown	<i>Spear phishing attachment (T0865)</i>	Malware (trojan backdoor)
Routine Audit of SCADA Laptop Identifies Virus	2005	Australia	CN	Unknown		Virus
Pennsylvania Water Company Hack	2006	United States	CN	Unknown	Internet accessible device (T0883)	malware
California Canal System Hack	2007	United States	SN	Individual (Insider)		Disgruntled employee
Wastewater Treatment District Hacked	2012	United States	CN	Individual (insider)		Disgruntled employee

Iranian cyberattack on New York dam	2013	United States	SN	Organized group	Internet accessible device (T0883)	Targeted attack
TGB water station hacks	2014	United States	CN	Individual (insider)		Disgruntled employee
European Public utility services attacked	2014					
Drinking water utility cellular routers hacked	2016	United States	LN		Wireless compromise (T0860)	
Suspicious network traffic data undisclosed location	2016	United States	LN			malware
“Kemuri” Water Company	2016	United States	SN	Organized group	Exploit public-facing application (T0819)	SQL injection and phishing
Clark County water hack	2016	United States				Ransomware
City of Atlanta ransomware attack	2018	United States	CN	Organized group/cyber criminals	-Drive-by-compromise (T0817) / Spearphishing attachment (T0865)	Ransomware
Crypto miner European water utility company	2018	Europe	SN		Drive-by compromise (T0817)	Cryptocurrency malware (Cryptojacking attacks)
Onslow Water and Sewer Authority	2018	United States	CN	Organized groups/cyber criminals		Ryuk Ransomware attack
Riviera Beach Water Utility ransomware	2019	United States	CN, LN, FN	Organized groups	Spearphishing attachment (T0865)	Ransomware attack
Israel’s water system facilities	2020	Israel	LN	Nation state		Targeted attack
Israel’s agricultural water pumps	2020	Israel				
Recycled water reservoir	2020	Israel				
Volue ASA Ransomware attack	2021	Norway	CN	Unknown	Spearphishing (T0865)	Ransomware attack
San Francisco Bay Area	2021	United States	LN	Unknown	Exploit public-facing application (T0819)	

water treatment incident						
Florida Water Treatment Plant	2021	United States	SN	Unknown	-Exploit public-facing application (T0819) -Exploitation of remote services (T0866)	

Analiza atacurilor cibernetice

Această secțiune furnizează o analiză a incidentelor de securitate cibernetică din diferite perspective (de exemplu, actori de amenințare cibernetică, ținte, tehnici, impact etc.). Tabelul 3.6 arată numărul de incidente de securitate cibernetică și tehnicile lor de atac, componentele vizate și impactul acestora. Utilizând investigația noastră a acestor incidente de securitate cibernetică, prezentăm următoarele atribute cheie, precum și modelele și tendințele care pot fi folosite pentru apărare și informații despre amenințări în sectorul apei și în alte infrastructuri critice.

1. Factori de amenințare

Figura 3.1 reprezintă *factorii de amenințare* pentru incidentele de securitate cibernetică analizate în secțiunea anterioară. Când analizăm tendințele în ceea ce privește actorii de amenințare, este important să notăm și motivația din spatele acestor atacuri. Am identificat *i* indivizi (atât interni, cât și externi), *j* grupuri organizate, *k* state naționale, *l* actori de amenințare necunoscuți. Mulți indivizi au desfășurat atacuri din motive personale, fie pentru câștig financiar, fie ca metode de răzbunare.

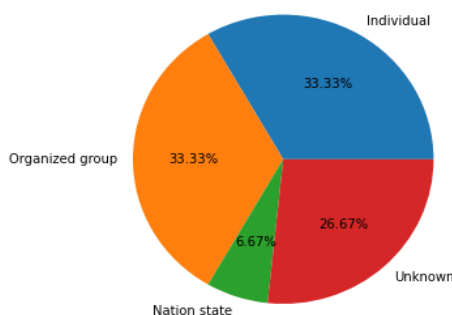


Figura 3.1 Factori de amenințare

2. Accesul inițial

Tehnicile de acces inițial din framework-urile MITRE ATT&CK și ATT&CK ICS Frameworks sunt folosite atunci când se clasifică tehnicile identificate în fiecare incident de securitate cibernetică din Secțiunea 5 (Mitre 2022a, 2022b). Figura 3.2 detaliază *tehnicile de atac* folosite în incidentele de securitate cibernetică din Tabelul 3.6. Dintre incidentele analizate, 5 au utilizat *un dispozitiv accesibil prin internet*, *y* au profitat de *un compromis wireless*, *z* dintre atacuri au implicat *exploatarea unei aplicații cu acces public*, *x* incidente au folosit *spear phishing*, iar *y* altele au utilizat *drive-by-compromise*.

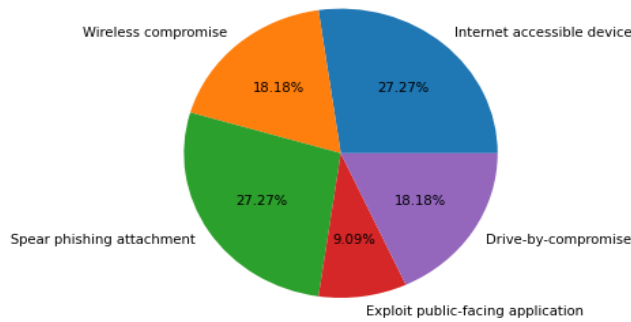


Figura 3.2 Tehnici de acces inițial

3. Niveluri țintă

Figura 3.3 prezintă distribuția incidentelor de securitate cibernetică în funcție de stratul țintă ICS. X incidente în sectorul apei au vizat stratul de rețea corporativă, y rețeaua supervizorală, z rețeaua de control local, dar k au vizat rețeaua de dispozitive de teren.

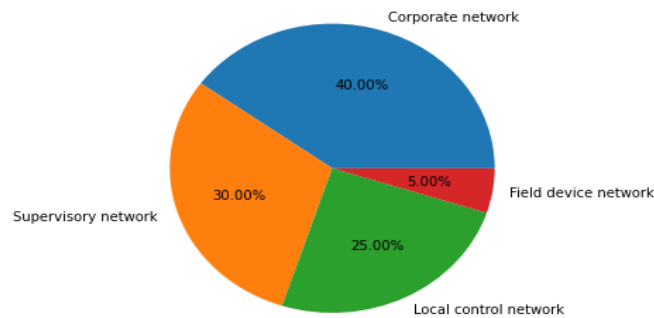


Figura 3.3 Niveluri țintă

Concluzii

Monitorizarea și controlul eficient al procesului sunt esențiale pentru îndeplinirea cerințelor în tratarea apelor uzate. Instrumentele digitale care pot fi utilizate în sectorul apei sunt numeroase, dar extrem de utili sunt senzorii software care pot fi utilizați în supraveghere și control. Necesitatea de a utiliza senzori software derivă din faptul că senzorii fizici (de exemplu, pentru încărcătura organică) sunt inexistenți sau foarte costisitori.

Supravegherea și controlul procesului sunt suficiente pentru obținerea unei calități bune a efluentului, dar nu protejează instalația împotriva atacurilor cibernetice și trebuie să fie cuplate cu proceduri de securitate cibernetică. Este, așadar, important ca industria să înțeleagă riscurile și să acționeze în consecință. În plus, materialele/cursurile despre supravegherea procesului, control și securitate cibernetică trebuie dezvoltate și utilizate pentru pregătirea inginerilor din stațiile de tratare a apelor uzate.

References

- Adepu, Sridhar, Venkata Reddy Palleti, Gyanendra Mishra, and Aditya Mathur. 2020. "Investigation of Cyber Attacks on a Water Distribution System." Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12418 LNCS(0): 274–91
- Alex J, Benedetti L, Copp J, Gernaey K, Jeppsson U, Nopens I, Pons M, Rosen C, Steyer J, Vanrolleghem P (2008) Benchmark simulation model no. 2 (bsm2). Report by the IWA Taskgroup on benchmarking of control strategies for WWTPs pp 1–99.
- Alexander, Otis, Misha Belisle, and Jacob Steele. 2020. "MITRE ATT&CK ® for Industrial Control Systems: Design and Philosophy."
- Alladi, Tejasvi, Vinay Chamola, and Sherali Zeadally. 2020. "Industrial Control Systems: Cyberattack Trends and Countermeasures." Computer Communications 155: 1–8.
- Andreeva, Oxana et al. 2016. "INDUSTRIAL CONTROL SYSTEMS VULNERABILITIES STATISTICS." <https://securelist.com/blog/research/73440/blackenergy-apt-attacks-in-ukraine-employ-spearphishing-with-word-> (August 11, 2022)
- Ariens, D., B. Houska, H. Ferreau, ACADO toolkit website. <http://www.acadotoolkit.org>.
- Ariens, D., B. Houska, H. Ferreau, F. Logist, "ACADO for Matlab User's Manual", Optimization in Engineering Center (OPTEC), 1.0 beta edition, May 2010.
- Barbu, M., I. Santin, and R. Vilanova, "Applying Control Actions for Water Line and Sludge Line to Increase Wastewater Treatment Plant Performance," Industrial and Engineering Chemistry Research, vol. 57, no. 16, pp. 5630-5638, 2018.
- Bergmeir, C., R. J. Hyndman, and B. Koo, "A note on the validity of cross-validation for evaluating autoregressive time series prediction," Computational Statistics & Data Analysis, vol. 120, 2018.
- Chen, C.T. "Linear System Theory and Design", Philadelphia, PA, USA: Saunders College Publishing, 1984.
- Cheung, Steven et al. 2006. "Using Model-Based Intrusion Detection for SCADA Networks" <http://digitalbond.com> (May 12, 2022)
- Chi B, and L. Guo, "Wastewater treatment sensor fault detection using RBF neural network with set membership estimation," Chinese Control and Decision Conference (CCDC), Nanchang, China, pp. 2685-2690, 2019.
- Christiansson, Henrik, and Eric Luijff. 2008. "Creating a European SCADA Security Testbed." IFIP International Federation for Information Processing 253: 237–47.
- Council Directive 91/271/EEC of 21 May 1991 concerning urban waste-water treatment The Urban Waste Water Treatment Directive (UWWTD)
- Dasarathy B., "Optimal control of nonlinear systems", IEEE Transactions on Automatic Control, 15(6), 690-692, 1970.
- DeCarlo, R. "Linear Systems: A State Variable Approach with Numerical Implementation", Prentice Hall, NJ, 1989.
- Drewnowski J, Szelağ B, Xie L, Lu X, Ganesapillai M, Deb CK, Szulzyk-Cieplak J, Łagód G. The Influence of COD Fraction Forms and Molecules Size on Hydrolysis Process Developed by Comparative OUR Studies in Activated Sludge Modelling. *Molecules*. 2020 Feb 19;25(4):929. doi: 10.3390/molecules25040929.
- FBI, CISA, EPA, and MI-ISAC. 2021. "TLP: WHITE Compromise of U.S. Water Treatment Facility SUMMARY." www.fbi.gov/contact-us/field-offices, (August 11, 2022)
- Germano, Judith H. 2019. "Cybersecurity Risk & Responsibility in the Water Sector Prepared By." : 6–19. <https://www.awwa.org/Portals/0/AWWA/Government/AWWACybersecurityRiskandResponsibility.pdf?ver=2018-12-05-123319-013>
- Ghinea LM, Miron M, Ratnaweera H. A Deep Learning Approach for Faults Recognition of Dissolved Oxygen Sensor in Wastewater Treatment Plants. ETFA 2023
- Goodfellow I., Y. Bengio, and A. Courville, "Deep Learning". MIT Press Cambridge, 2016, vol. 1.

- Hassanzadeh, Amin et al. 2020. "A Review of Cybersecurity Incidents in the Water Sector." *Journal of Environmental Engineering* 146(5): 03120003
- Hemsley, Kevin E, and Ronald E Fisher. 2018. "History of Industrial Control System Cyber Incidents." INL/CON-18-44411-Revision-2 (December): 1–37. <https://www.osti.gov/servlets/purl/1505628>
- Housh, Mashor, and Ziv Ohar. 2018. "Model-Based Approach for Cyber-Physical Attack Detection in Water Distribution Systems." *Water Research* 139: 132–43
- Igure, Vinay M., Sean A. Laughter, and Ronald D. Williams. 2006. "Security Issues in SCADA Networks." *Computers & Security* 25(7): 498–506
- Jacod J., P. Protter, "Discretization of Processes", Springer, 2011.
- Khalil H. K., "Nonlinear Systems", Third Edition, Prentice hall Upper Saddle River, NJ, 2002.
- Kovacs E. 2020a. "Hackers Knew How to Target PLCs in Israel Water Facility Attacks: Sources | SecurityWeek.Com." <https://www.securityweek.com/hackers-knew-how-target-plcs-israel-water-facility-attacks-sources> (August 12, 2022).
- . 2020b. "Israel Says Hackers Targeted SCADA Systems at Water Facilities | SecurityWeek.Com." <https://www.securityweek.com/israel-says-hackers-targeted-scada-systems-water-facilities> (August 12, 2022)
- Liu Y., Y. Jiang, and I. Bortone, "A Scheme for Anaerobic Digestion Modelling and ADM1 Model Calibration", *Chemical Engineering Transactions*, vol. 96, 2022.
- Luenberger D.G., "Optimal Control. Introduction to Dynamic Systems", New York: John Wiley & Sons, pp. 393–435, 1979.
- Makrakis, Georgios Michail et al. 2021. "Industrial and Critical Infrastructure Security: Technical Analysis of Real-Life Security Incidents." *IEEE Access* 9: 165295–325
- Mamandipoor B, M. Majd, S. Sheikhalishahi, C. Modena, and V. Osmani, "Monitoring and detecting faults in wastewater treatment plants using deep learning", *Environ. Monit. Assess.* vol. 192, issue 2, 2020.
- Manamperuma, L., Wei, L., Ratnaweera, H. Multi-parameter based coagulant dosing control. *Water Science and Technology* 75 (9), 2017, 2157-2162
- Mass, Massachusetts Department of Environmental Protection. 2021. "Cybersecurity Advisory for Public Water Suppliers | Mass.Gov." <https://www.mass.gov/service-details/cybersecurity-advisory-for-public-water-suppliers> (August 11, 2022)
- Miller, Bill, and Dale C. Rowe. 2012. "A Survey of SCADA and Critical Infrastructure Incidents." *RIIT'12 - Proceedings of the ACM Research in Information Technology*: 51–56
- Mitre. 2022. "ATT&CK® for Industrial Control Systems." https://collaborate.mitre.org/attackics/index.php/Main_Page (August 8, 2022).
- Moraitis, Georgios et al. 2020. "Quantifying Failure for Critical Water Infrastructures under Cyber-Physical Threats." *Journal of Environmental Engineering* 146(9): 04020108. <https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29EE.1943-7870.0001765> (May 11, 2022).
- Naduvil-Vadukootu, R. A. Angryk, and P. Riley, "Evaluating pre-processing strategies for time series prediction using deep learning architectures," *The Thirtieth International Flairs Conference*, 2017.
- Nair, A., Hykkerud, A., Ratnaweera, H. Estimating Phosphorus and COD Concentrations Using a Hybrid Soft Sensor: A Case Study in a Norwegian Municipal Wastewater Treatment Plant. *Water* 14 (3), 2022, 332
- Necoara, I. "Model predictive control for hybrid systems: piecewise affine and maxplus-linear systems", *VDM*, 2008.
- Necoara, I. "Metode de optimizare numerica", Editura Politehnica Press, Bucuresti, 2013.
- Pisa I., A. Morell, R. Vilanova, and J. L. Vicario, "Transfer Learning in Wastewater Treatment Plant Control Design: From Conventional to Long Short-Term Memory-Based Controllers," *Sensors*, vol. 21, no. 18, p. 6315, Sep. 2021, doi: 10.3390/s21186315.
- Pisa I., "Artificial Neural Networks in the Wastewater Industry From Conventional to Data-based Industrial Control", Ph.D. Thesis in Electronic and Telecommunication Engineering, Universitat Autònoma de Barcelona (UAB), Spain, 2022.

- Ramotsoela, Daniel T., Gerhard P. Hancke, and Adnan M. Abu-Mahfouz. 2019. "Attack Detection in Water Distribution Systems Using Machine Learning." *Human-centric Computing and Information Sciences* 9(1): 1–22
- Rasekh, Amin et al. 2016. "Smart Water Networks and Cyber Security." *Journal of Water Resources Planning and Management* 142(7): 01816004
- Ratnaweera, H., Fetting, J. State of the Art of Online Monitoring and Control of the Coagulation Process. *Water* 2015; Volume 7.(11) p. 6574-6597.
- Salles R., J. Mendes, R. P. Ribeiro, and J. Gama, "Fault Detection in Wastewater Treatment Plants: Application of Autoencoders Models with Streaming Data", section in book "Machine Learning and Principles and Practice of Knowledge Discovery in Databases", pp. 55- 70 2023.
- Santin I., C. Pedret, R. Vilanova, and M. Meneses, "Advanced decision control system for effluent violations removal in wastewater treatment plants," *Control Engineering Practice*, vol. 49, no. 2, 2016
- Serino G, Ben Miller. 2021. "Recommendations Following the Oldsmar Water Treatment Facility Cyber Attack | Dragos." <https://www.dragos.com/blog/industry-news/recommendations-following-the-oldsmar-water-treatment-facility-cyber-attack/> (August 11, 2022)
- Sharmeen, Shaila et al. 2021. "An Advanced Boundary Protection Control for the Smart Water Network Using Semi Supervised and Deep Learning Approaches." *IEEE Internet of Things Journal*
- Shimon Prokupez, Tal Kopan and Sonia Moghe. 2015. "Official: Iranians Hacked into New York Dam - CNNPolitics." <https://edition.cnn.com/2015/12/21/politics/iranian-hackers-new-york-dam/index.html> (August 12, 2022)
- Slay, Jill, and Michael Miller. 2008. "LESSONS LEARNED FROM THE MAROOCHY WATER BREACH."
- Socha L., "Linearization Methods for Stochastic Dynamic Systems", Springer, 2007.
- Statistics Norway. <https://www.ssb.no/en>. Accessed 10.06.2023
- Tawfik, A., El-Gohary, F., Temmink, H. Treatment of domestic wastewater in an up-flow anaerobic sludge blanket reactor followed by moving bed biofilm reactor. *Bioprocess Biosyst Eng* (2010) 33:267–276. DOI 10.1007/s00449-009-0321-1
- Vericlave. 2018. "Vericlave™-The Kemuri Water Company Hack." www.sentryo.com (August 11, 2022)