LOTUS

LOW-COST INNOVATIVE TECHNOLOGY FOR WATER QUALITY MONITORING
AND WATER RESOURCES MANAGEMENT FOR URBAN AND RURAL WATER SYSTEMS IN INDIA

Deliverable D4.2

Tools Prototypes and documentation

Lead: IIT Bombay

Date: March 30th 2022

Public

LOTUS is co-funded by the European Commission under the Horizon 2020 research and innovation programme under Grant Agreement N° 820881 and by the Indian Government, Ministry of Science and Technology.

Project Deliverable

Project Number	Project Acronym	Project Title
820881	LOTUS	LOw-cost innovative Technology for water quality monitoring and water resources management for Urban and rural water Systems in India

Instrument:	Thematic Priority
Research and Innovation action	EU-India water co-operation

Title
D 4.2: Tools Prototypes and documentation

Contractual Delivery Date	Actual Delivery Date
July 31st, 2021 (M30)	March 30th, 2022 (M38) Update: June 17th, 2022 (M41)

Start Date of the project	Duration
February 1st, 2019 June 1st, 2019 (India)	48 months

Organisation name of lead contractor for this deliverable	Document version
IIT Bombay	V 2.0

Dissemination level		Deliverable Type	
Public	**X**	**Document/Report**	**X**
Confidential		Demonstrator	

LOTUS is co-funded by the European Commission under the Horizon 2020 research and innovation programme under Grant Agreement N° 820881 and by the Indian Government, Ministry of Science and Technology.

Authors (organisations)

IIT BOMBAY

Reviewers (organisations)

TUDO, IITG, UNEXE

Abstract

This document captures the first versions of the software tools and functionalities for real-time operational management of the water systems, where LOTUS (sensors and tools) will be deployed in each of the use case studies.

Keywords

Software tools; prototype, documentation

Disclaimer

LOTUS has been financed with support from the European Commission and the Indian Government, Ministry of Science and Technology.

The LOTUS Project

LOTUS is a project funded by DG Environment under the European Union Horizon 2020 Research and Innovation Programme and by the Indian Government. It brings together EU and Indian prominent organisations with the aim to co-create, co-design and co-develop innovative robust affordable low-cost sensing solutions for enhancing India's water and sanitation challenges in both rural and urban area.

The LOTUS solution is based on an innovative sensor and includes tailor-made decision support to exploit the capabilities of the sensor as well as a specific approach to co-creation. LOTUS aims to be co-designed and co-produced in India, and have a wide, diverse and lasting impact for the water sector in India due to intense collaborations with commercial and academic partners in India.

Based on the low-cost sensor platform, solutions for the early detection of water quality problems, decision support for countermeasures and optimal management of drinking and irrigation water systems, tailored on the functionalities of the new sensor, will be developed and integrated with the existing monitoring and control systems.

This sensor will be deployed in five different use cases: in a water-network, on ground-water, in irrigation, in an algae-based waste water treatment plant and in water tankers. The packaging of the sensor, as well as the online and offline software tools will be tailored for each of the use cases. These last will enable to test the sensors and improve them iteratively.

The project is based on co-creation, co-design and co-production between the different partners. Therefore, an important stakeholder engagement process will be implemented during the project lifetime and involve relevant stakeholders, including local authorities, water users and social communities, and will consider possible gender differences in the use and need of water. Broad outreach activities will take place both in India and in Europe, therefore contributing to LOTUS impact maximisation.

The further development and exploitation (beyond the project) of the novel sensor platform will be done in cooperation with the Indian partners. This will create a level playing field for European and Indian industries and SMEs working in the water quality area.

Table of Contents

LOTUS is co-funded by the European Commission under the Horizon 2020 research and innovation programme under Grant Agreement N° 820881 and by the Indian Government, Ministry of Science and Technology.

List of Figures

1 Executive Summary

This report outlines the various modelling tools and functionalities that have been developed for real-time operational management of water systems, where LOTUS sensors will be deployed. The software described here has been developed for assistance in the optimal operation of water systems of each use case in real time, or near real time (i.e. response times between minutes to few hours). To this end, a brief summary of developed modelling and simulation tools is provided whereever applicable in each use case. WP 4 focuses on the development of models/algorithms for simulating and real-time management of water systems, i.e. dynamically reacting to changing situations, demands, quality problems, failures, leakages, etc. while also ensuring the suitability of the tools with the use case instalments. The collection of technical requirements for the execution and integration of these real-time modelling and optimization tools with the LOTUS platform was already presented in last deliverable (D4.1), with specifics to each use case study.

The described framework of real-time modelling and simulation tools will aid the teams in decisions on subsequent steps for demonsration of these real time tools with the state-of-art platform where data will be collected, processed in real-time with these simulation/optimization tools and receives feedback for deciding optimal water systems management strategy.

2 Overview

2.1 List of real-time tools for all use cases

This section provides a brief description of various tools and functionalities for real-time operational management of the use case water systems, where LOTUS sensors will be deployed.

2.1.1 Water Distribution Network Guwahati

1. ***Water quantity (pressure, flow) monitoring and anomaly detection in urban water distribution systems***

The main objective of this tool is the real-time detection of anomalies and early warning for quantitative (flow and pressure) disturbances in the Guwahati piped water distribution network (WDN). These disturbances may occur after sudden physical damage to network components, due to e.g. pipe bursts (natural, accidental or intentional), pipeline leakages, component malfunctions (e.g. pumps, valves) etc. The detection will be based on pressure and flow sensors deployed in the network (existing sensors or LOTUS sensors) that are linked to the SCADA system. This functionality for real-time system operational management will involve models for event recognition (burst and leakage), leakage location identification and a holistic approach to leakage management by combining several types of alerts and approaches in a single Decision Support Platform to be developed in WP5.

Following software tools have been developed for this task and a brief description of algorithms and methodology involved in each tool is described in subsequent use case chapter :

a. **EPANET based Leak Detection Tool (ELD)** : This tool helps in detecting pipe burst events using pressure and/or flow sensors placed strategically in the water distribution system and a calibrated EPANET hydraulic model.
b. **EPANET based Leak Localization Tool (ELL)** : Subsequent to the detection of a burst event of possibilty from above tool, this tool will pinpoint the exact location of a pipe burst using a machine learning (ML) model.

2. ***Water quality monitoring and anomaly detection***

While the former tool focused on events related to water quantity, this tool will involve water quality monitoring in the Guwahati WDN and associated contamination anomaly detection functionalities. For the Guwahati WDN use case, this water quality monitoring tool will be based on and combined with the sensor placement module developed in WP3. The water quality parameters which will be considered in the tools will be described in a report from WP3.

3. *Real-time mitigation measures for water quantity and quality alerts*

Once a valid contamination alert (based on flow/ pressure/ water quality disturbance) has been generated from the anomaly detection tools, the real-time mitigation support tools act to assist water utility operators in deciding and optimizing mitigation plans for clean water supply. Further, the real-time intervention measures will also cover a control strategy for optimal disinfectant dosage to ensure that disinfectant levels are high enough to minimize microbial contaminant re-growth, while minimizing disinfectant by products formation risk.

Following software tools have been developed for this task and a brief description of algorithms and methodology involved in each tool is described in the text in corresponding use case chapter :

a. **EPANET Leak Detection (ELD) and Leak Location (ELL):** models to identify the occurrence of a leak (ELD) and the locations of valves (ELL) that need to be closed to isolate the burst pipe, described in section 3.1.1. and 3.1.2 respectively. This tool receives as input the ID of the burst pipe and produces the IDs of the valves to be closed.
b. **Coordinated Decentralization-based Optimization of Disinfectant Dosing in Large-scale Water Distribution Networks (CODD**): This tool is for responding to a microbial water quality contamination event alert, where an optimal response is predicted for disinfectant dosage at several booster stations in the network to control the further microbial population growth and restoration of water quality to acceptable limits in the contamination zone.

4. *Real-time alerts to the public*

These tools will involve modules for alerting the public, using mobile phone applications (apps). It comprises the design and development of the apps, as well as the system architecture of delivery to the public. The generic module will be adapted to the needs of each specific use case, with specifically designed displays and screens, in Indian languages.

2.1.2 Tanker-based water distribution network

1. *Short-term planning and operational management in real-time for tanker-based water distribution systems*

The objective of this tool is the real-time monitoring and management of tanker-based distribution systems. This includes, on the lower level, the monitoring of water quality based on data from LOTUS sensor deployed at the water sources, tankers and consumption points, and on the systems level the planning of treatment plant operations, logistic optimization, and scheduling of tankers in adaption to the dynamic demand of costumers.

Following software tools have been developed for this task and a brief description of algorithms and methodology involved in each tool is described in the text in corresponding use case chapter:

a. **Tanker water Scheduling and Planning Framework (IITB-TSP)**: This tool find the optimal tanker movement schedule (from sources to treatment plants to consumers) while minimizing the total operating cost for timely delivery of water to consumers and sustainable use of available water resources. The optimization tool rigorously incorporates major constraints related to operational nuances of different components (viz. water sources, treatment facilities, and consumers) in tanker-based water supply systems, various limitations of water treatment, transit time, and distribution aspects to provide the optimal target tanker schedule that is achievable in representative applications.

b. **Stochastic Optimization Model for Uncertain Demands in Tanker water supply systems in Urban Areas (SUD):** This tool involves a stochastic optimization framework for efficient planning and operations of tanker water supply systems exhibiting consumer water demand uncertainty.

2.1.3 Irrigation systems

1. Real-time optimal operation of irrigation systems based on quantity and quality monitoring

The objective of the decision support tools is to assist in the optimal use of water and in the optimal addition of fertilizer to the water for irrigation. The goals are to provide a continuous flow of water of sufficient quantity and to provide the right amount of fertilizer for optimum growth. This includes taking account of uncertainties (about the future inflow of water, in particular due to rainfall) for the informed decision making to manage the storage and retrieval of water.

Tool under development: Advisory system for optimum irrigation and fertigation

2.1.4 Groundwater and river water monitoring

1. Weather prediction tool for ground water and river water quality modelling

This tool will be concerned with time series based modelling of weather conditions to predict precipitation in the catchment areas of the river as well as other and ground water sources. Based on these prediction of the precipitation, in conjunction with other CFD/ hydrology-based models, the quality of the water stream will be predicted.

2.1.5 Wastewater treatment

1. Monitoring and control of algae-based wastewater treatment

The information about the quality of the outflow of the experimental algae-based waste-water treatment plant provided by the LOTUS sensor will be used for model-based monitoring and control of the operation of the plant with the goal of long-term stability, meeting of water quality standards and assuring the usability of the algae. This will be based upon the investigation of the observability of the process from different sensor information. A state estimator will be developed. Based on this information, strategies for the control of the algae based treatment will be designed and tested in simulations.

Tool under development: Model-based advisory for optimal operation of algae-based waste-water treatment

2. Control of activated sludge wastewater treatment processes

Algorithms are developed that employ the information provided by the LOTUS sensor about the concentrations of ammonia and nitrate in different basins of a conventional wastewater treatment plant to improve the quality of the effluent water and reduce the energy consumption of the plant.

Tool under development: Model-based control algorithm for activated sludge wastewater treatment

3 Guwahati Water Distribution Network

3.1 Subtask 4.2.1

Water quantity (pressure, flow) monitoring and anomaly detection in urban water distribution systems

This sub task addresses the real-time detection of flow and pressure disturbances in urban water distribution systems. The functionalities involved in the tool comprise anomaly detection and early warning for quantitative (flow and pressure) disturbances in water supply and distribution systems (WDS) in real time, occurring after sudden physical damage to network components: pipe bursts (natural, accidental or intentional), leakage alerts, component malfunctions (e.g. pumps). The detection will be based on the new LOTUS sensors and will also include information from existing pressure and flow sensors in the use case network. The leakage location models involved in this subtask exist for event (burst and leakage) recognition system management with UNEXE and its integration in LOTUS decision support platform is in progress, but a bit delayed due to pending deployment of LOTUS sensors in the network. The following section provides an overview of the leak detection and localization models developed.

3.1.1 EPANET Leak Detection (ELD)

The EPANET Leak Detection model focuses on detecting pipe burst events using pressure and/or flow sensors within a water distribution system and a calibrated EPANET hydraulic model. The OWA EPANET modelling software was selected since it is an open-source software, that is common within the industry, is able to perform pressure driven analysis (assisting with pipe break analysis) and is integrated into python package to assist with coding (Open Water Analytics, 2021).

The EPANET Leak Detection model must be calibrated for each specific water distribution system analyzed. Calibration of this model is performed by simulating leak and non-leak scenarios, optimization leak detection thresholds using an calibrated EPANET hydraulic model, calculating the exponential weighted moving average (EWMA) for sensor readings, and updating anomaly service and visualization for each sensor based on the calculated EWMA value. The entire EPANET Anomaly Detection methodological approach is highlighted below in **Figure 1**.

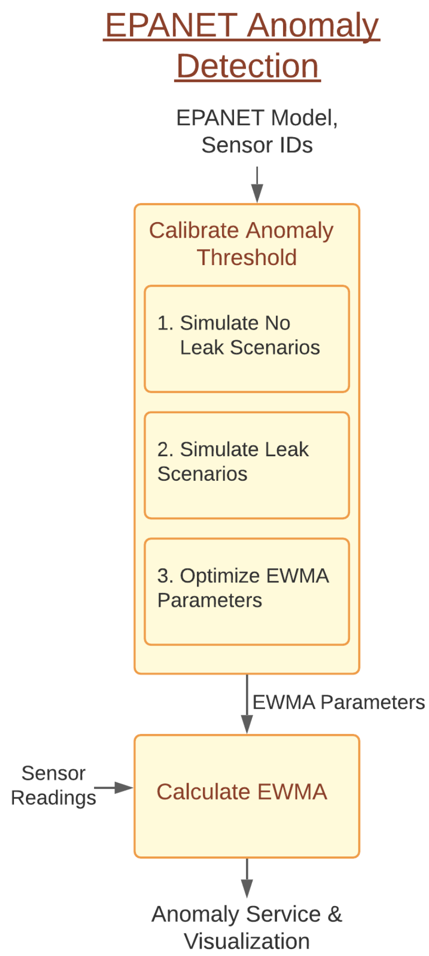

Figure 1: EPANET Anomaly Detection Schematic

The EWMA statistical method was chosen for the leak detection model as it has been shown to outperform many of the other methods in detecting anomalies within water distribution systems (Jung et al., 2015).

The EWMA equation is described below, with *a* representing the time weighted factor:

$$ewma_{i,t} = a * z_{i,t} + (1 - a) * \ ewma_{i,t-1}$$

where, $ewma_{i,t}$ is the exponentially weighted moving average at time t for the ith sensor. Note at t=0, $ewma_{i,0} = 0$.

The leak detection threshold for the EWMA is determined through the following equation:

$$EWMA\ threshold = +/- \ L * (\frac{a}{2-a})^{0.5}$$

Where L is a set failure threshold. If the absolute value of $ewma_{i,t}$ is greater than the EWMA threshold, a leak is detected.

A grid search optimization approach is employed to select optimal a and L parameters that maximize detection rate, while minimizing the false positive rate. Specifically, the optimized performance indicator is:

$$Balanced\ Accuracy = \frac{Detection\ Rate + (1-FPR)}{2}$$

Where FPR (False Positive Rate) represents the percentage of days with one or more false positive during the no leak scenario and the Detection Rate represents the percent of leak scenarios that successfully detected a leak, i.e. EWMA exceeded the EWMA threshold for one or more of the sensors within 48 hours of leak start.

The grid-search optimization approach runs through a number of possible values for a and L calculating the Balanced Accuracy based on the 100 simulated leak scenarios and the non-leak scenario. The a and L values that maximize the Balance Accuracy performance indicator are then selected from this approach and the EPANET Leak Detection Calibration step is completed.

Once the EPANET Leak Detection model is calibrated and the optimal EWMA threshold parameters are determined, the leak detection model can be used with live sensor data. As sensor data is uploaded, the EWMA value is calculated and compared with the EWMA threshold. If the EWMA value exceeds the threshold, the anomaly service associated with that sensor is updated. The specific steps are outlined below

1. New Sensor data associated with the water distribution system (pressure or flow sensors currently) is updated. This triggers the leak detection service
2. The leak detection service then calculates the residual of this data by comparing it to the no leak scenario hydraulic model simulation average and standard deviation for this time period and sensor.
3. The residual is then used to calculate the EWMA, being weighted by previous values within the time series for that specific sensor based on the optimized parameters.
4. If the EWMA exceeds the EWMA threshold a leak anomaly is triggered for that sensor.

3.1.1.1 Results

The EPANET Leak Detection Model was tested on an example EPANET network (no calibrated EPANET network was available from any of the case studies at the time of this report). The example EPANET network included 1004 nodes, 783 pipes, and 1 reservoir. Moreover, 2 flow and 5 pressure sensors were assumed to be installed throughout the network, with each sensor being associated with specific pipes and junctions within the EPANET file. Sensors were modelled to record pressure/flow at 15 min increments throughout the day. Noise was added to the hydraulic model's demands and simulated sensor readings.

The following simulations were performed to test the EPANET Anomaly Detection sub-model:

1. No-leak Scenario was simulated for six months. The average reading ($xavg_{i,t}$) and standard deviation ($\sigma_{i,t}$) were calculated for each time step for each day of the week for each sensor.
2. 100 leak scenarios (training dataset) were simulated to be used to calibrate the Leak Detection Model. Each leak scenario was simulated for 11 days, with the leak start time occurring randomly between day 2 and day 9.
3. To test the effectiveness of the calibrated mode, a separate testing dataset was generated containing 1 six-month non-leak scenario and 100 test leak scenarios.

To calibrate the EPANET Leak Detection model, an optimization grid search was performed that maximized the balanced accuracy based on the training scenarios. The optimization grid search attempted to optimize the EWMA parameters a, L using the following values:

a = [0.2, 0.3, 0.4 … 0.9]

L = [2, 2.2, 2.4 … 6]

The grid search optimization determined that a = 0.8 and L = 4.5 maximized the balance accuracy. The Overall training accuracy had an average False Positive Rate of 0%, Leak Detection Rate (LDR) of 100% and an average detection time of 42 minutes (or three time steps).

To test the accuracy the model the EWMA was then calculated for the hold-out testing datasets using the optimized EWMA parameters from the training set (a = 0.8, and L=4.5). A leak event was detected for all 100 leak scenarios, within 48 hours of the leak occurrence (LD = 100%) and only one day during the no-leak scenario was a leak triggered (i.e. false positive rate of 0.5%). The overall deterection time was 18min (or a single time step).

Overall, these results indicate that with a well calibrated hydraulic model the EPANET Leak Detection algorithm is effective at identifying when a pipe burst is to occur. However, since the sensors within a water distribution network are often spread out, identifying where the pipe burst may be occurring can be difficult. The Leak Localization Model is developed to assist with this task, identifying the area where a pipe burst is likely to have occurred.

3.1.2 EPANET Leak Location (ELL)

Many localization models developed attempt to pinpoint the exact location of a pipe burst. However, utility staff are still required to search in-person to confirm the leak location. If the leak is not at the predicted location, staff are left wondering where to search next. Several clustering approaches have been developed to improve leak localization. However, these methods are based on dividing the WDS network in general search areas, and then training a model to predict which area the leak is likely to occur in. This approach can help utilities identify areas to search but having set areas also limits the flexibility of the model to identify the accurate location. The model developed for this project advances leak localization research by developing search area clusters using the leak node probability predicted from the machine learning (ML) model.

3.1.2.1 Leak Localization Methodology

Similar to the EPANET Anomaly Detection model, the EPANET Localization model uses a EPANET hydraulic model to calibrate the prediction model by developing a labelled training dataset. A ML model is then trained using this dataset, outputting the probability of the event occurring at each node within the network. A Maximum Coverage Location Problem (MCLP) is then solved to identify the optimal search area(s) that would contain the location of the event. The following section of this report details the methodology for each of these steps.

Create a Labelled Dataset

The first step in developing the EPANET Localization model is to create a labelled training dataset that contains the sensor reads during the event, as well as the node id that the event was simulated at. This training dataset will be used to train the machine learning model to predict the location of the event.

To create an encompassing training dataset, six anomaly/pipe burst were simulated for every node within the EPANET network to study variations in pipe burst size and time of pipe burst. Pipe burst sizes were simulated by randomly selecting a low leak emitter (between 0.25-1) and randomly selecting a high leak emitter (between 1.5-3). Leak start times were randomly selected and pressure or flow readings at sensor locations were transformed into residuals using non-leak scenario averages and standard deviations (similar to EPANET Leak Detection model). A six-hour window of residuals are stored for each simulation, starting six hours before detection, and ending when the burst was identified. Note for testing purposes the leak detection time was randomly chosen to occur between 1-4 hours after leak occurrence.

Create a Machine Learning Model to Classify Anomaly Localization:

A Machine learning classifier is created using the labelled training dataset. Each node represents a class, and the input variables are the residuals for each simulated sensor reading during the six-hour detection window. A Random Forest (RF) machine learning model was selected because of its high accuracy, ease of hyper parameter tuning, and robustness against data outliers (Breiman, 2001).

With any machine learning classifier, the output is a probability score for each class (or in this case each node). The predicted probability for each node is then used by the Maximum Coverage Solution Problem to prioritize the search area(s).

Maximum Coverage Location Problem

To identify the optimal search areas for leak detection crews, the Maximum Coverage Location Problem (MCLP) is applied to the output from the machine learning classification model. Since the machine classification model outputs the predicted probability of the leak location occurring at each node within the network, the MCLP can be used to group these results geographically (using the node's GIS coordinates) and locate the search areas that maximizes the likelihood of containing the leak event.

The MCLP algorithm adapted to leak localization is as follows:

1. $Maximize\ z = \sum_{i \in I} a_i y_i,$
2. $Subject\ to:\ y_i \leq \sum_{j \in N\mathrm{i}} x_j,\ \ i \in I$
3. $\sum_{j \in J} x_j = p,$
4. $0 \leq y_i \leq 1\ \ i \in I,$
5. $x_j \in \{0,1\}\ j \in J.$

Where:

i, I The index and set of EPANET nodes (i.e. possible anomaly locations)

j, J The index and set of search area centroids

a_i The predicted probability at node i

$d_{i,j}$ The shortest distance from node i to search area centroid j

S The search area radius

N_i $\{j \mid d_{i,j} \leq S\}$ = the nodes *j* that are within a distance of *S* to node *i*

p The number of search areas to be determined

x_j A binary variable that equals one when the search centroid is located at the *j*th node.

y_i A binary variable which equals one if node *i is* within one or more search areas

(Church & Revelle, 1974).

In order to solve the MCLP algorithm a list of possible search area centroid locations are required. To provide this a grid of possible search area centroids is created to cover the entire network by selecting the EPANET nodes with maximum and minimum latitudinal and longitudinal coordinates to create a grid boundary. A grid of possible search area centroids is then created within this boundary using a set distance between centroids.

In addition to the grid spacing of each search area centroids, the search area radius and number of search areas must be provided to solve the MCLP. Various grid spacing, search area radiuses, and number of search areas are compared and discussed in the following evaluation section.

3.1.2.2 Leak Localization Results

The evaluation of the leak localization model is performed in two parts. First the Random Forest classification model's hyper parameters are tuned, and the accuracy of the model is calculated using an independent testing dataset. Second, after determining the effectiveness of the Random Forest classification model, the MCLP approach and overall leak localization model is compared with other popular leak localization models.

Random Forest Machine Learning Classification Model:

Before assessing the effectiveness of the Random Forest machine learning model, the tree depth and number of trees hyperparameters are adjusted to optimize accuracy (i.e. the percentage of accurately predicted node locations containing the simulated leak). A grid search methodology was employed, and results were tested using a validation dataset which included leak scenarios for 100 randomly selected leak nodes. The tuning parameters assessed include:

Tree depth: [100, 200, 300, 400, 500, 600]

Number of trees: [25, 50, 100, 150m 200, 250, 300].

The results from this analysis suggest the optimal tuning parameters were a tree depth of 250 with 300 number of decision trees developed.

Using these tuning parameters, the leak location random forest model was trained. The accuracy of the model was assessed using an independent test dataset which included leak scenarios for 100 randomly selected leak nodes. The overall accuracy of the Random Forest model on the independent test dataset was 7.5%, which is 75 times greater than random chance. However, this is still a relatively low accuracy which highlights the importance of using the MCLP approach to identify leak search areas that are more likely to contain the leak.

Evaluation of Complete Leak Localization Model (and MCLP aglorithm):

The complete leak localization model developed is compared with other leading localization approaches. Specifically, the Random Forest ML model with the MCLP approach detailed in this report was assessed using one search area (MCLP-1 SA) and two search area (MCLP – 2 SA) and compared with three other models:

1. A classification random forest model (Classification - 1 SA) - where the node with the highest predicted probability being chosen as the centroid for the search area. A single search area is expanded out from this centroid until 50% of the leaks from the test dataset are within the search areas.
2. A regression based random forest model (Regression -1 SA) - where the model predicts the latitude and longitude of the search area. A single circular search area is expanded outwards from the predicted latitude and longitude until 50% of the leaks from the test dataset are within the search areas.
3. An agglomerative hierarchical clustering model (Clustering - 2 SA) – where a random forest model classification model is trained to predict probabilities of the leak occurring at each node for each simulation. Using this output, an agglomerative hierarchical clustering algorithm is performed using a max linkage parameter set to the desired search diameter. The search diameter is expanded until two clusters include 50% of the leaks from the test dataset.

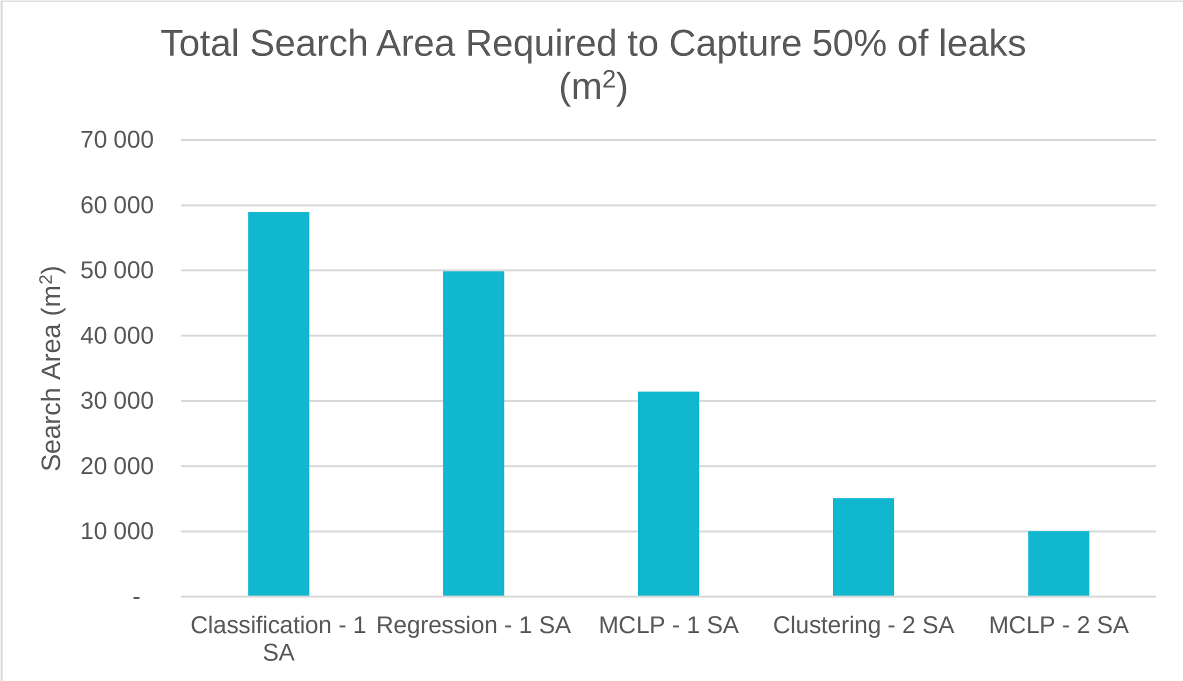

Figure 2: Comparing Search Areas for Various Localization Methods

Figure 2 identifies the geographical search area required by each model to accurately locate 50% of the simulated leaks within the test dataset. The results indicate that the Random Forest MCLP approach required the smallest single search area (35% less than leading regression-based model) and the smallest two search area (34% less than the clustering model). Overall, these results indicate the Leak localization model developed is very effective at reducing the search space required to locate the leak and outperforms the other leak localization models assessed.

3.2 Subtask 4.2.2

Water quality monitoring and anomaly detection

This sub-task is directly linked to the functionalities and capabilities of the sensor in WP2, demonstrating the advanced and flexible capabilities of the sensor, combined with the types of systems, where LOTUS will be implemented. For the Water Distribution System (WDS) it will be combined with the sensor placement module developed in WP3 and currently in progress following the deployment of the LOTUS sensors in the network.

LOTUS is co-funded by the European Commission under the Horizon 2020 research and innovation programme under Grant Agreement N° 820881 and by the Indian Government, Ministry of Science and Technology.

3.3 Subtask 4.2.3

Real time mitigation measures for water quality alerts

This sub-task includes the development of control strategy for online optimal disinfectant chemical dosage to ensure that disinfectant levels are high enough to minimize contaminant growth/ build-up, while not forming disinfectant by-products. For large-scale water distribution networks, the simulation and modelling tool available with IITB is tailored for Guwahati water distribution system and described below. Furthermore, the approach has been extended to optimize disinfectant dosage for intermittent water supply systems also.

3.3.1 Coordinated Decentralization-based Optimization of Disinfectant Dosing in Large-scale Water Distribution Networks (CODD)

Disinfectants are typically used in water distribution networks (WDNs) to maintain the microbiological quality of potable water throughout the network. Disinfectant residual levels in such large distribution networks need to be maintained within the prescribed bounds to address two major problems: (1) preventing microbial regrowth, and (2) minimizing harmful disinfection by-product (DBP) formation resulting from high levels of the disinfectant itself. These requirements pose a two-point constrained control problem of disinfectant residual levels for water quality risk mitigation measures. Another important aspect of consideration is that such WDNs exhibit spatial and temporal variations in water quality at different points in the network. Therefore, conventional systems engineering tools such as modeling, optimization, and control need to be adapted to accommodate these variations for water quality management in distribution networks. To this end, a novel zone control based approach to address the two-point quality control problem is developed. Further, to combat the spatial complexity, we combine the zone control approach with a decentralized optimization strategy, namely, the model coordination method, which has been widely proposed to solve large-scale optimization problems (Vadera and Gudi, 2013). The algorithm and optimization problem along with results is briefly described next, details can be found in (Maheshwari et al. 2018).

3.3.1.1 Model Coordination Method in the WDNs: Optimal Dosing Methodology

Due to large scale of problem, several challenges have been encountered in the tasks of modelling, analyzing, identifying cause-effect relationship and implementing appropriate control algorithms and decisions for disinfectant dosage in centralized control of disinfection process in WDN. These challenges include (i) significant efforts required in the design of centralized control system and communication overheads, as it consists of grouping information from all the parts of the system in order to create a centralize decision making structure, (ii) high dimensionality of variables in one single optimization problem is a difficulty for appropriate formulation and solution. This then requires choice of the most significant design variables, active design constraints and choice of the objective functions that reflects the needed design goal, and (iii) lack of robustness in case of failures in either the control or communication system. These challenges necessitate the following steps in the development of the overall approach for optimization and real-time control of disinfectant dosage in WDN:

a) Decomposition of the overall WDN into sub-systems using an appropriate partitioning approach

b) Identifying the nature and extent of interaction between the sub-systems

c) Accommodating the interactions into a formal optimization formulation at the local level with a view to converge on the global optimality in the decision making.

Therefore, a three step strategy, where the techniques (principal components analysis (PCA), partial correlation analysis and effective relative gai array analysis (ERGA) respectively) were employed sequentially on the data (400 simulated values of chlorine residual concentration at all nodes) in the strategic manner to develop a partitioning strategy for decentralized control of water quality (chlorine residuals and DBPs) problem in large-scale distribution networks. The above three steps help in partitioning a large sized WDN into local sub-systems which could be better controlled /optimized (with minimal interaction) to optimality. For the problem at hand, the merit of disinfection through booster stations, at several points in WDN can be used for designing decentralized controllers to make these actuators (disinfectant dosage booster stations) work in coordinated fashion. And, this design of the control system structure (i.e., which of the available system inputs are to be used to control each of the system outputs) has to be taken care at the planning level while defining the optimization problem.

Furthermore, in general, optimizing a single sub-system from a large system without taking effect of interactions from other sub-systems may lead to a solution which is not optimal for the entire system. However, if sub-systems are coupled through co-ordinating variables or parameters in a hierarchical manner as in model co-ordination, an overall optimal solution can be achieved. This type of decomposition is advantageous when first level sub-systems have certain independent variables and they are interacting with the second level only through common coordinating variables. Model co-ordination divides the optimization problem into two levels, namely lower level, and higher level. The partition is achieved such that coordinating variables of first level sub problems are free to be manipulated by the controller at the second level. The second level controller chooses coordinating variables such that all independent sub-systems in the first level take the system to overall optima corresponding to the chosen coordinating variable. Model co-ordination is directly applicable when the objective function of the integrated optimization problem is in additively separable form. Apart from the lower computational effort, it also improves the numerical performance of the optimizer, making it achieve the optimal solution in lesser iterations than an integrated optimization approach.

LOTUS is co-funded by the European Commission under the Horizon 2020 research and innovation programme under Grant Agreement N° 820881 and by the Indian Government, Ministry of Science and Technology.

The block diagram and optimization formulation of the two level decomposition, model co-ordination method, using co-ordination links resulted from partial correlation analysis for example network, is shown below in the Figure 3-1, and results also summarized in Figure 3-2.

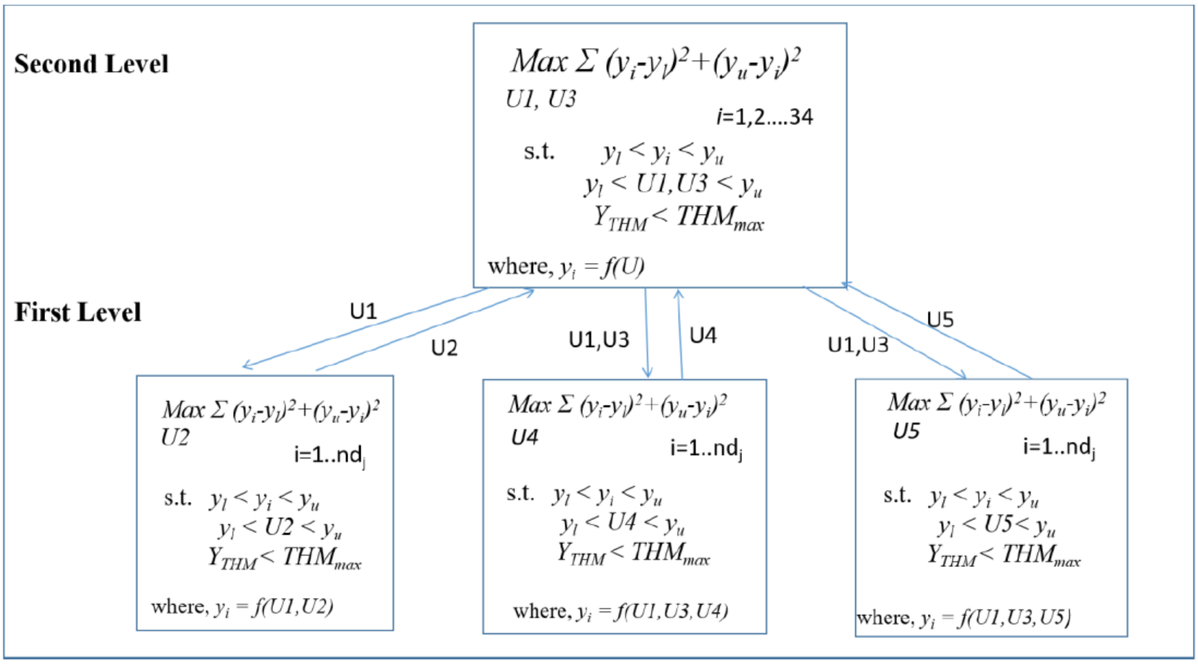

Figure 3-3: Block Diagram of Co-ordinated Decentralized Disinfectant Dosage Methodology

To start the algorithm, choose an initial guess value of coordinating variables as U1= U1° and U3= U3°. The algorithm proceeds in an iterative manner until the optimal solution of all sub problems at first level also corresponds to the optimized solution at the second level. It can be seen that in every sub-problem j, y_{ji} is a function of coordinating variables and local variable (U1, U3, Uj), reflecting the cause and effect relationship in WDN. This method is known as Model co-ordination because it modifies the mathematical model by adding one more equation of fixing coordinating variable at second level as C= C° in every iteration.

where, y_i and $Y_{i,THM}$ $= f(U)$.

y_i = chlorine concentration at i^{th} nodal point (mg/L)

nd = total number of nodes in the network

y_l = lower bound on chlorine concentration (mg/L)

y_u = upper bound on chlorine concentration (mg/L)

LOTUS is co-funded by the European Commission under the Horizon 2020 research and innovation programme under Grant Agreement N° 820881 and by the Indian Government, Ministry of Science and Technology.

$Y_{i,THM}$ = THM concentration at i^{th} node in each cluster (µg/L)

$THM_{\max}$ = max. concentration limit on THM concentration (80 µg/L)

U_j = disinfectant dosing concentration at j^{th} booster station (mg/L)

U = Total disinfectant dosage vector $[U_1, U_2, U_3, U_4, U_5]^T$

In terms of actual functionality, the above algorithm achieves regulation of chlorine levels within bounds as prescribed by drinking water standards; these bounds themselves reflect the need to ensure residual chlorine levels above a threshold to ensure dsinfection for preventing microbial (re)growth and below a maximum threshold that guards aginst formatin of harmful DBPs.

As an example, in the following network chlorine dosage levels were arrived at in an optimal manner to regulate chlorine levels within thresholds (0.2-2 mg/L) by coordination of disinfectant dosage at 5 booster stations (U_1-U_5) during sudden microbial contamination event as indicated in the Figure 3-2.

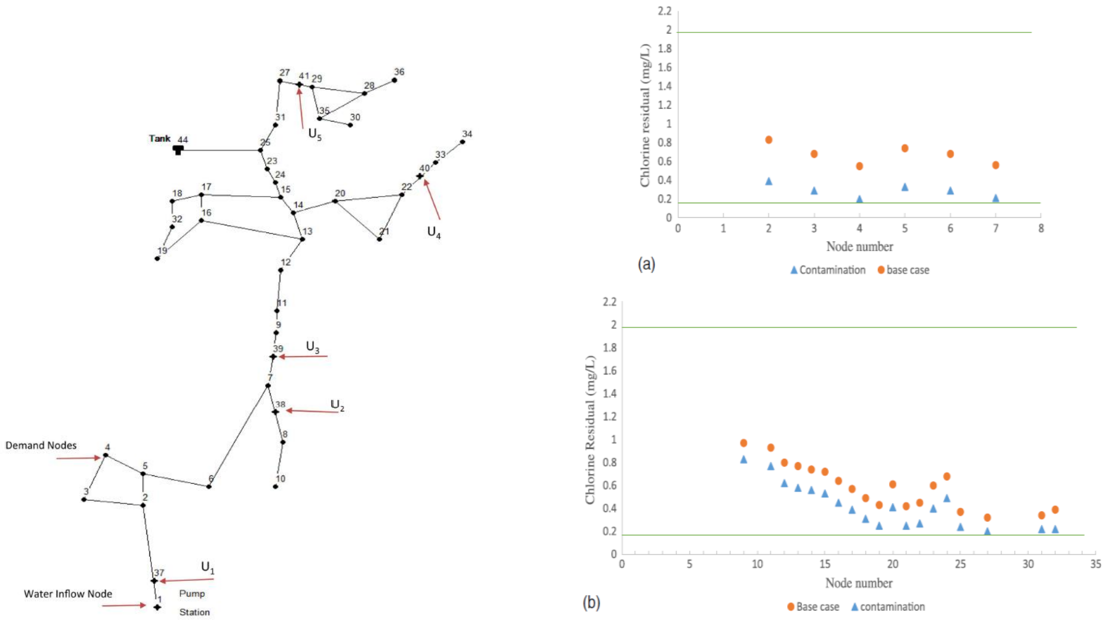

Figure 3-4: Results of Implementation of Algorithm on an Example Water Distribution Network of EPANET, (a) for nodes in cluster 1, (b) for nodes in cluster 3

References:

1. A. Maheshwari, A. A. Abokifa, R. D. Gudi, P. Biswas (2018). Coordinated Decentralization-Based Optimization of Disinfectant Dosing in Large-Scale Water Distribution Networks. Journal of Water Resources Planning and Management, ASCE, 144(10), p.04018066.

2. Vadera, D., & Gudi, R. D. (2013). "Scheduling Incorporating Waste Management using Decomposition Approaches". IFAC Proceedings Volumes, 46(32), 535–540.

4 Tanker-based Water Distribution Systems

4.1 Subtask 4.2.4

Management of tanker-based distribution systems

In this sub-task, solutions have been developed for the real-time monitoring and management of tanker-based distribution systems based on the LOTUS sensor. This includes, on the hardware level, the operation of the on-tanker disinfecting system, and on the system level the logistic optimization tool in adaption to the varying consumer demands. A brief description of the developed systems planning and optimization frameworks is provided in the next sections.

4.1.1 A Short-term Planning Framework for the Operation of Tanker -based Water Distribution System in Urban Areas

To efficiently operate tanker service systems, a large fleet of tanker trucks are required to transport water among several water sources, water treatment plants and consumers spanning across the regions. As shown in Figure 4-1, this requires tight coordination between water suppliers, treatment plant operations, and user groups to use available water resources in a sustainable manner, along with the assurance of water quality and timely delivery. This work proposes a novel formulation to assist decision-making for optimizing tanker-based water distribution systems and treatment operations, with an overall objective of minimizing the total operating cost such that all of the constraints related to the water demand, supply operations, and environmental and social aspects are honored while supplying water to a maximum number of users.

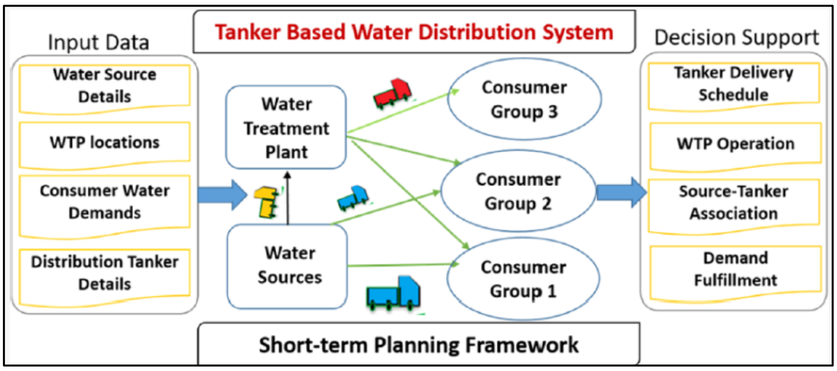

Figure 4-1: Problem Description of Tanker-based Water Distribution System

The problem is formulated and solved as a mixed integer linear programming (MILP) optimization framework and captures all of the nuances related to (i) water availability limitations and quality constraints from different sources, (ii) maintaining water quality as it transports via tankers, (iii) water demands for various end-use purposes, and (iv) transportation across a water supply chain. The proposed novel framework is applied to a realistic urban model to find the optimal tanker delivery schedule, ensuring appropriate treatment and timely delivery of water. The results of the case study conducted on a representative-scale problem also elucidate several aspects of treatment plant operation and consumer demand fulfillment for the efficient planning and management of tanker-based water distribution systems. A pictorial summary of the optimization model is shown in Figure 4-2 and mathematical details can be refeered to (Maheshwari et al., 2020).

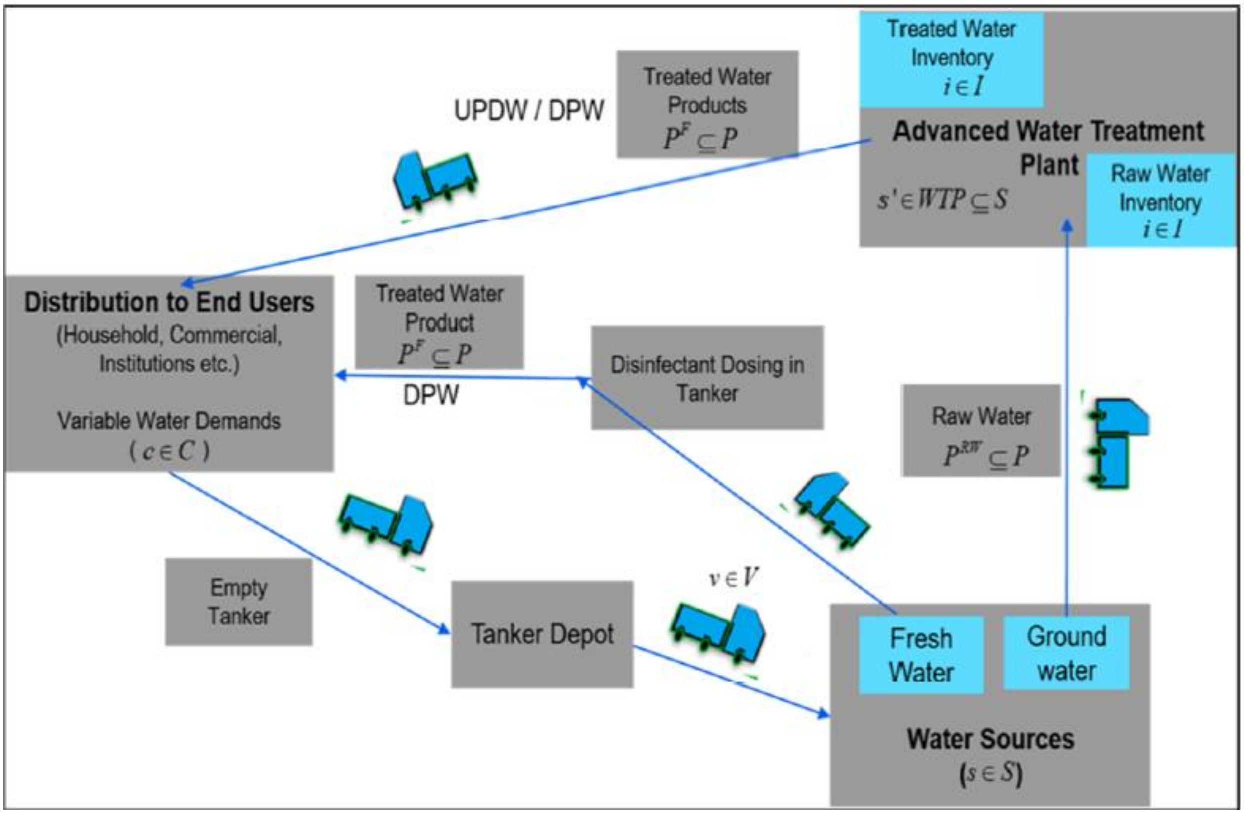

Figure 4-2: Optimization Formulation Design of a Tanker-based Water Distribution System

Furthermore, the efficacy of the proposed MILP optimization framework towards handling the peculiarities of planning and management of tanker based water supply system, is demonstrated here through the application on a representative case study. This case study is based on the practical insights about typical tanker water supply system operations from Just Paani Water Supply Solutions. Such enterprises that operate tanker water supply typically manage a network of tanker water vendors in different areas of the city. These vendors supply water either from ground water or fresh water sources in different capacity tankers. The consumers and vendors are mapped with each other by the enterprise company based on (i) the customer demand (quantity and time slot) data, (ii) delivery locations and (iii) availability of water tankers with several vendors, in the network across the city. However, in the present situation, due to the lack of any planning and scheduling framework for real-time operation of the tanker water supply system, the suppliers face challenges of untimely water delivery and water quality monitoring issues.

The results of framework are demonstrated to assist decision makers on various aspects of tanker water supply system operation to ensure quality water and fulfillment of all costumers demands, in the required turnaround times. Also, a typical scenario of a maintenance period for the treatment facility is simulated to present the ability of proposed formulation in assisting towards optimal decisions. The case study considers two fresh water sources (FW2, FW3), three groundwater sources (GW1, GW3, GW4) and two water treatment plants (TF1, TF2) to serve the water demands of the city as shown in Figure 4-3. Based on the location of these sources and consumers, the entire city is divided into three regions (R1, R2, and R3) having a different combination of water sources and treatment plants as depicted by Figure 4-3. And, consumers in each region are clustered in three types, namely, household consumers (HHC), commercial consumers (CC), and hospital and health care institutions (HC). The raw water from all the groundwater sources is first transported to either of the treatment facilities in region R1 and R2 and then treated water is delivered to consumers in different locations. To transport water in different sections, this case study considers tankers in two capacities 6000 L (6T) and 10,000 L (10T) for raw water whereas, treated water is supplied using tankers of three capacities (3000 L, 6000 L and 10,000 L).

The results in Figure 4-4 demonstrate the framework catering the demands of multiple consumers in an optimal manner in different regions from each water source. Further to understand the water supply service to each consumer type, Figure 4-5 shows the demand fulfilment pattern of DPW product for consumer CC1, in both morning time periods. This figure shows that in situations of increased water demands which cannot be fulfilled by the sources present in that region, the supply is planned so as to be met by sources of another region in an optimal manner. Taken together, these results highlights the framework's feature where a consumer is serviced by multiple sources depending on the quantity of water demand, demand timings, and tanker availability in the region. Furthermore, Figure 4-6 depicts that according to the water demand quantity in different time periods, tankers of various capacities are selected to supply water at minimum transportation cost. Thus, the framework provides a detailed understanding of questions such as how much quantity of water to be delivered from different suitable sources to a consumer by optimally utilizing multiple tanker capacities. Lastly, Figure 4-7 shows the percentage distribution of various cost components in the objective function.

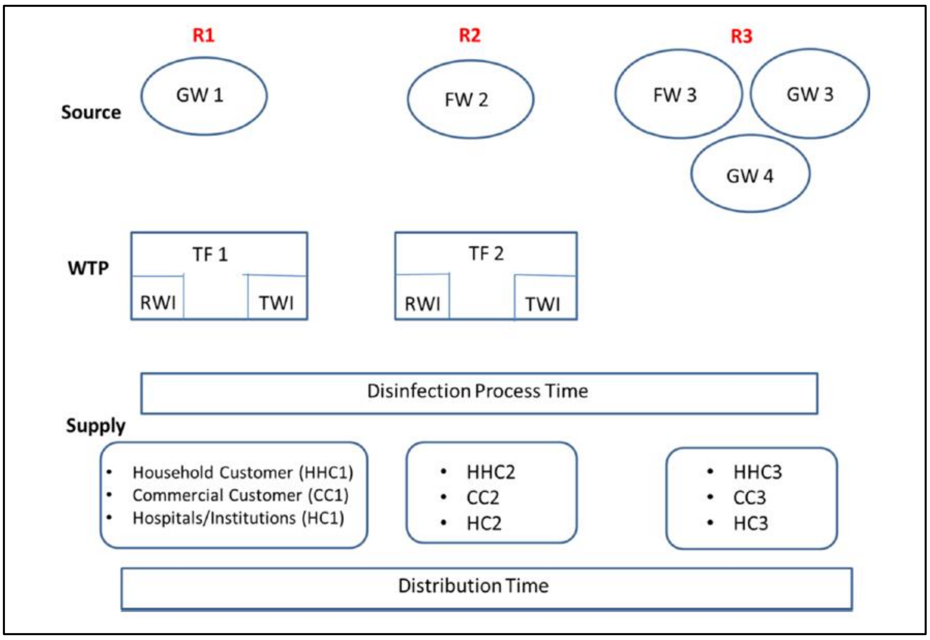

Figure 4-3: Description of Components in the Application Case Study

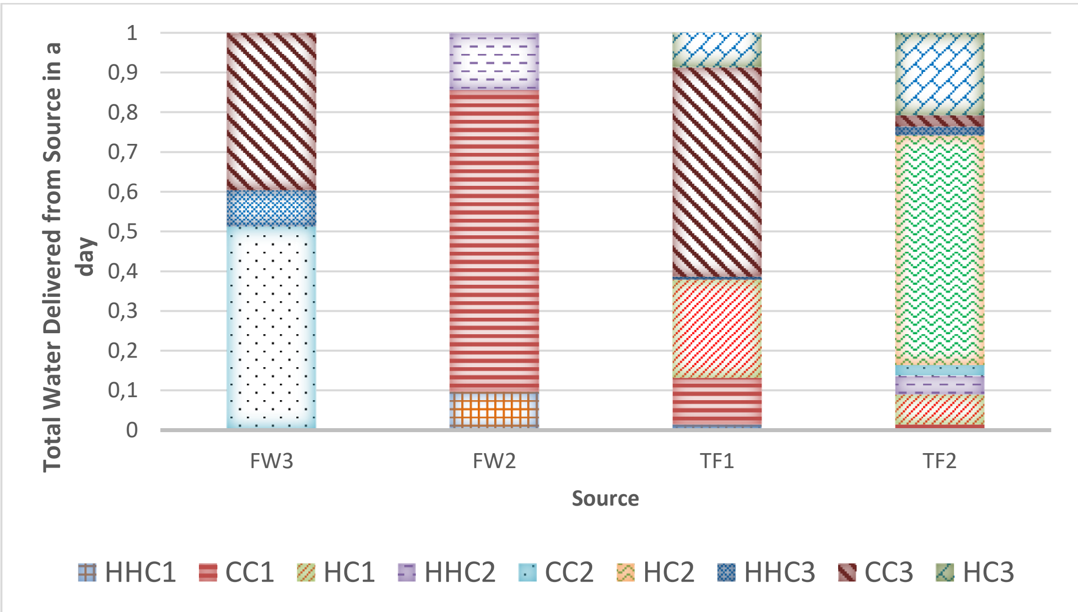

Figure 4-4: : Each Water Source Catering to Multiple Consumers

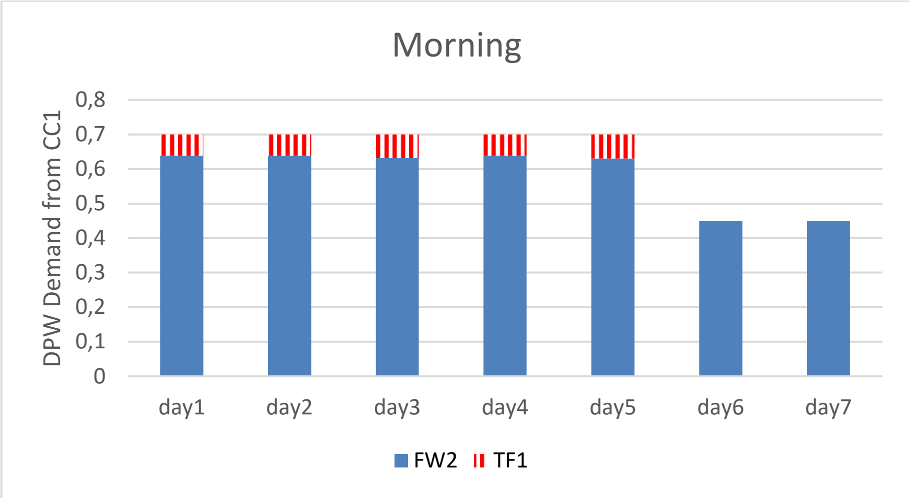

Figure 4-5: Daily Demand fulfilment pattern of CC1 Consumer in Morning period

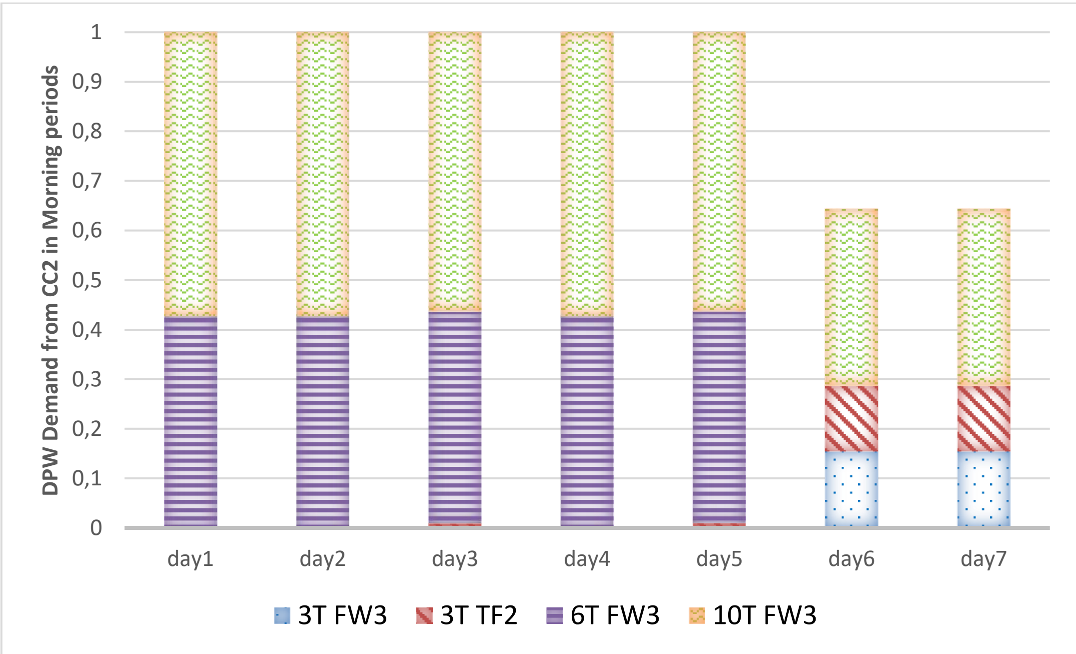

Figure 4-6: Source-Consumer-Truck association for demand fulfilment of CC2

LOTUS is co-funded by the European Commission under the Horizon 2020 research and innovation programme under Grant Agreement N° 820881 and by the Indian Government, Ministry of Science and Technology.

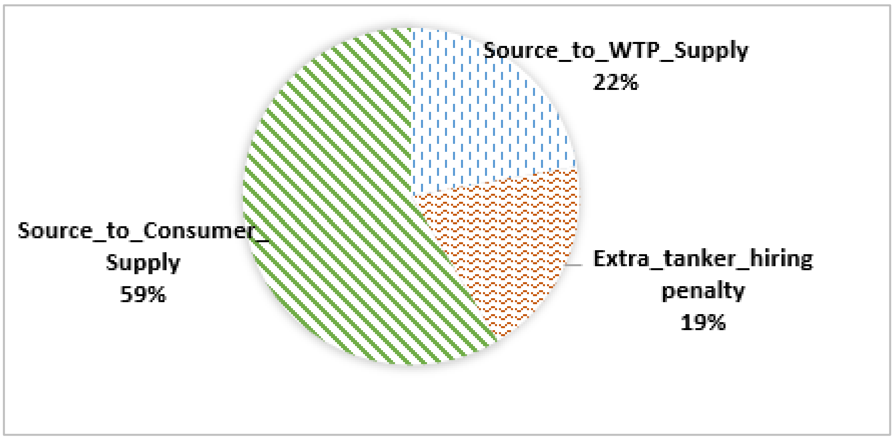

Figure 4-7:Percentage distribution of cost components in objective function corresponding to optimal solution

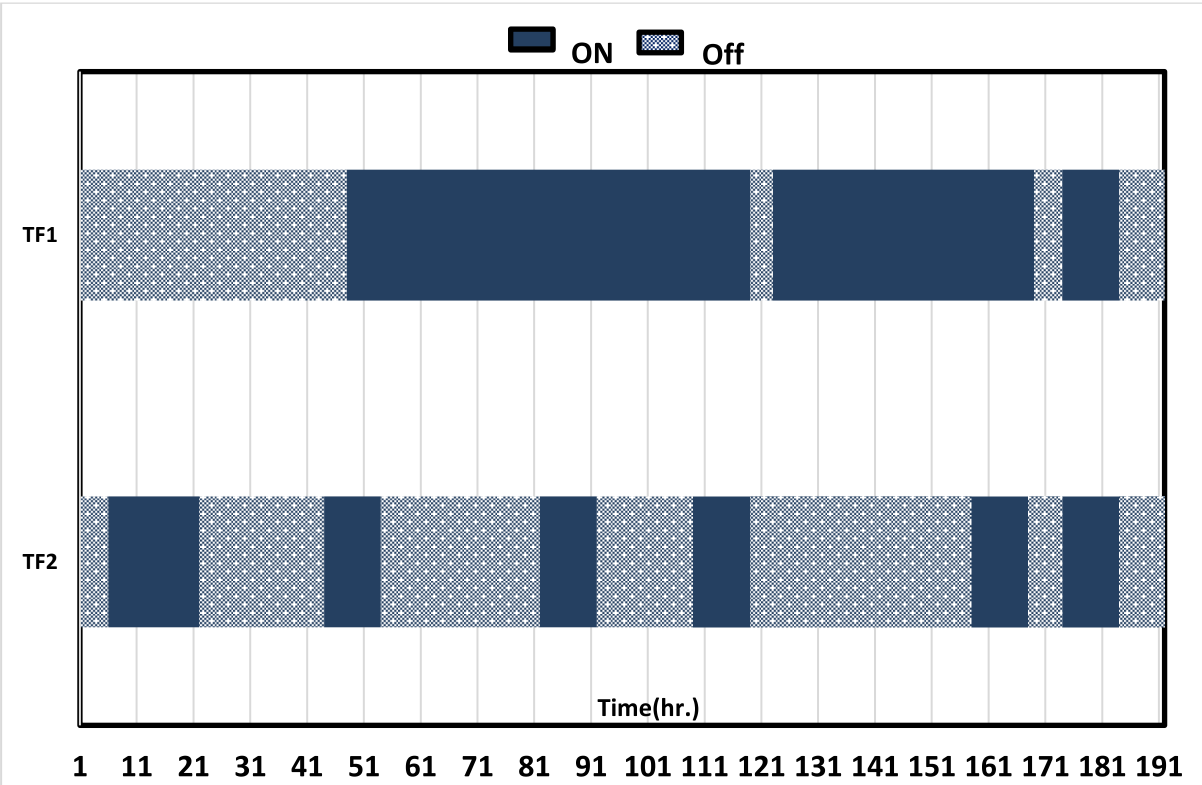

Figure 4-8:Optimal Schedule of Water Treatment Facilities operation

4.1.2 Stochastic Optimization Model for Short-term Planning of Tanker Water Supply Systems in Urban Areas

Accounting for the uncertain nature of customer demands while making the planning decisions not only aids in achieving minimum operational cost but is also important in reducing the water wastage. Herein, we developed a two-stage stochastic recourse programming model for an optimal planning and scheduling of tanker water supply system under daily demand uncertainty. The main objective is to supply water to maximum number of consumers with minimum total operating costs. A solution strategy combining Sample Average Approximation (SAA) and Monte-Carlo Simulation (MCS) methods, to generate an equivalent deterministic MILP (mixed integer programming problem) model with multiple scenarios of demand uncertainty realization, is adopted for problem solving. The proposed model is applied to an example tanker water supply system and the benefits of two-stage stochastic modelling in making agile decisions incorporating the effect of uncertainties are illustrated. The figure 4-10, below describes the overall optimization model in the tool and details can be refereed to (Maheshwari et al. 2022).

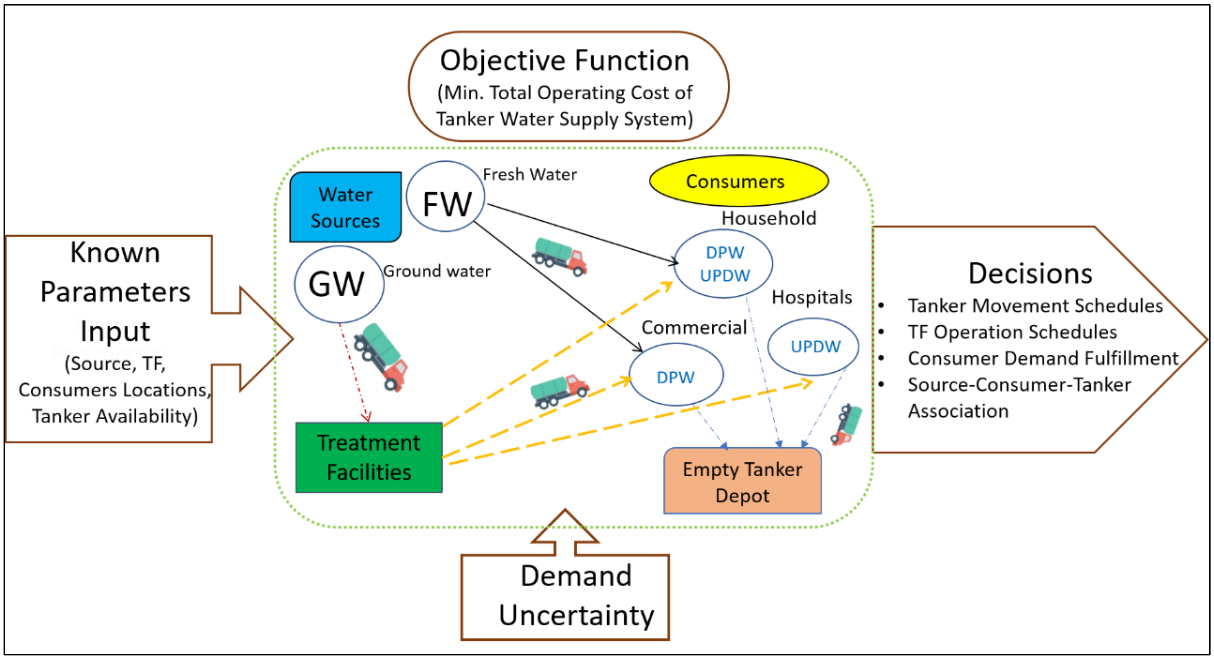

Figure 4-9:Schematic of Stochastic Demand Planning in Tanker Water Supply System

The developed stochastic MILP optimization model is demonstrated herein through an example system where water supply in urban area is divided into two regions (R1, R2). As shown in Figure 4-9, this system consists of one water source (GW/Surface water (FW)), one treatment facility and three type of consumer clusters in each region, viz. Households (HHC), Commercial (CC) and Hospitals (HC) having different water quality demands (DPW/ UPDW). Furthermore, two type of tanker capacities are considered, 6 kL and 10 kL to supply water from sources to treatment facilities and consumers and the total number of available tankers of each type is provided region-wise. The results of total operating cost, demand shortfall penalty costs and extra tankers hired in the objective function for stochastic problem with recourse (SPR) solution are shown in Figure 4-10. Furthermore, the value of stochastic solution (VSS), i.e. advantage of performing stochastic optimization over deterministic model with mean value is shown using a quantitative metric (Birge and Louveaux, 2011) calculated as shown in Equation 1. To calculate VSS, we fix the first stage solutions to the deterministic model optimal solutions and compared the model performance by solving each stage two problem separately. An expected total operating cost of 1.11 e+07 is obtained. This is also known as expected value solution (EVS) in the literature. As expected, this EVS is 13% higher than SPR solution for the example case study with the developed stochastic model in this paper. This is because the model in this first stage fixed case is oblivious of the high demand scenario and therefore hires the extra tanker capacity to be just enough to satisfy mean demand values. This led to high penalty costs on demand shortfall scenarios and consequently higher objective function value in EVS solution as also shown in results comparison in Fig. 4-10.

VSS= (SPR-EVS)/EVS (1)

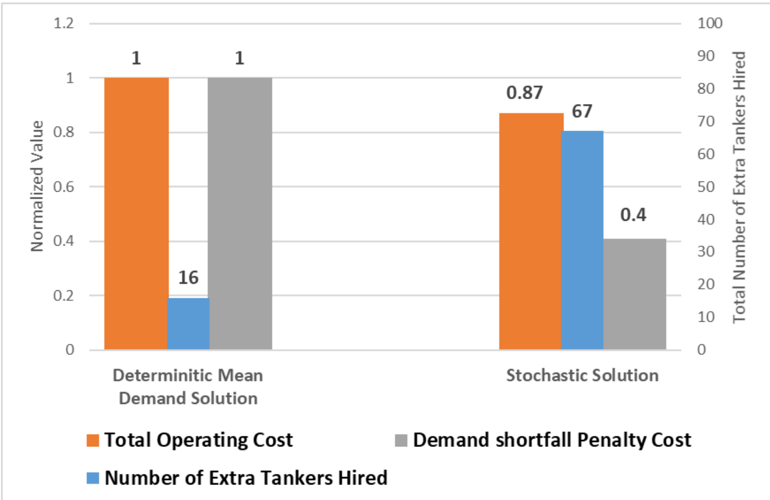

Figure 4-10:Comparison of stochastic solution (SPR) with expected value solution (EVS)

References:

1. A. Maheshwari, S. Misra, R. D. Gudi, S. Subbiah (2020). A Short-term Planning Framework for the Operation of Tanker based Water Distribution System in Urban Areas. Industrial Engineering & Chemistry Research 59 (20), 9575-9592.

LOTUS is co-funded by the European Commission under the Horizon 2020 research and innovation programme under Grant Agreement N° 820881 and by the Indian Government, Ministry of Science and Technology.

2. A. Maheshwari, S. Misra, R. D. Gudi, C.Laspidou (2022). Stochastic Optimization Model for Short-term Planning of Tanker Water Supply Systems in Urban Areas. Accepted for Publication in IFAC-PapersOnline of DYCOPS 2022.

5 Irrigation Systems

5.1 Subtask 4.2.5

Real-time optimal operation of irrigation systems regarding quantity and quality of water

In this sub-task, a solution will be developed for the optimal use of the available sources of water for irrigation in order to provide a continuous flow of water with the right amount of fertilizer. As a first step, the optimal usage of water and pumping energy for the growth of the crops was studied. Optimal irrigation can not only help plants to stay healthy but can also help to minimize the usage of water which is a valuable resource. An illustration of the approach can be found in Figure 5.1. Model-predictive control was chosen as the control strategy for the control of the amount of irrigation because it can integrate the weather forecast to obtain the optimal amount of irrigation water to achieve the optimum plant growth.

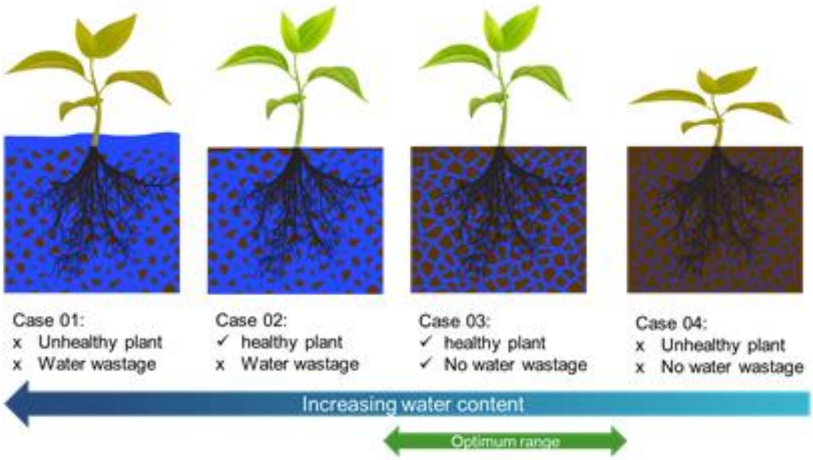

Figure 5-1: Illustration of plant health and water usage.

5.1.1 Model building

For the implementation of the model-predictive control scheme, a model that represents the evolution of the moisture content of the soil is required. The biosphere model 2 [1] is used for this purpose. The model equations are given below:

$$\frac{d\theta_1}{dt} = \frac{1}{D_1}\left[Pt + Irr - Q_{1,2} - E_g\right],$$

$$\frac{d\theta_2}{dt} = \frac{1}{D_2}\left[Q_{1,2} - Q_{2,3} - E_{tr}\right],$$

$$\frac{d\theta_3}{dt} = \frac{1}{D_3}\left[Q_{2,3} \; - \; Q_3\right],$$

where $\theta_1, \theta_2, \theta_3$ denote the moisture content of the top layer of the soil, of the root zone layer and of the deep layer of the soil with the associated depths D_1, D_2, D_3. Pt denotes precipitation, Irr denotes irrigation, $Q_{a,b}$ denotes the flow of water from layer a to layer b, E_g denotes the evaporation of water from the surface of the soil, E_{tr} denotes the transpiration of water from the plant.

Evapotranspiration is a combined term used for evaporation and transpiration. It refers to the amount of water transferred into the atmosphere, collectively from the soil surface and the plant or crop, during a certain duration of time. To represent the reality with reasonable accuracy, the estimation of evapotranspiration is vital. There are a number of factors that affect the amount of evapotranspiration: weather conditions, crop type and the stage of growth, soil factors, and other factors. We employ the methodology given in [2] to estimate the evaporation from the weather prediction data and employ MPC control.

5.1.1.1 Controller design for irrigation

The design of the control scheme for the irrigation of a field can be seen in Figure 5.2, where the constraints and objective function are derived from the desired plant growth targets. The model of the controller uses the weather forecast data to predict the required irrigation for the next two weeks. The value for the first predicted day is applied to the field, which is also influenced by the actual weather. The soil moisture is measured in the field to know its current value in the root zone layer. To achieve the control goal, a nonlinear model-predictive control scheme was designed using do-mpc, an open-source python library for the development of model predictive control applications created by TU Dortmund (www.do-mpc.com). It uses the open-source tools CasADi (https://web.casadi.org/) and Ipopt (https://coin-or.github.io/Ipopt/) for solving the optimization problems.

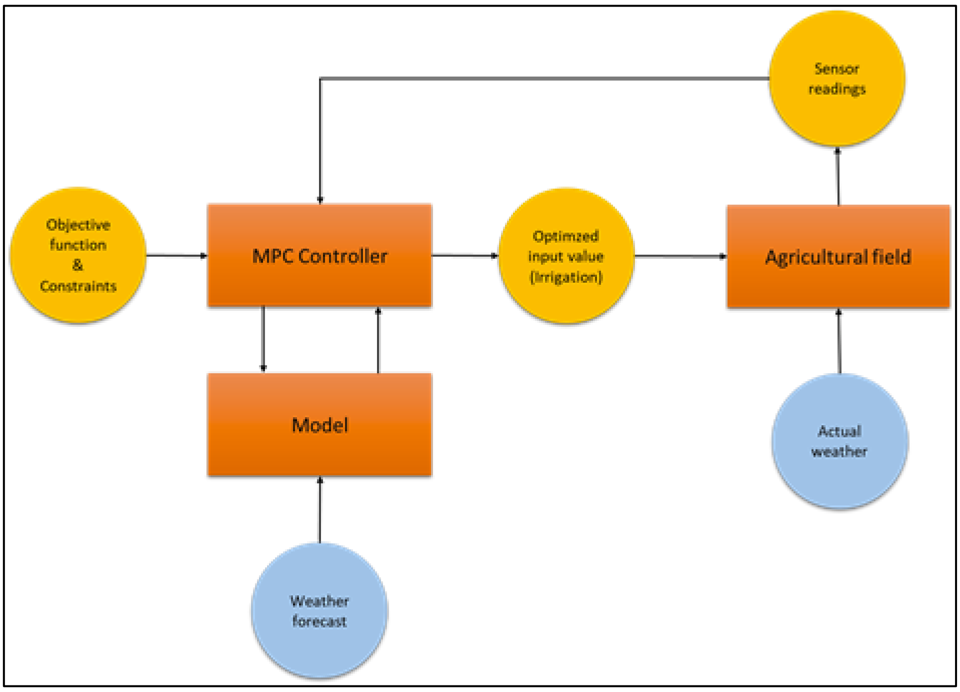

Figure 5-2: Control loop for irrigation.

5.1.1.2 Simulation study

The control goal is to keep the moisture content of the root zone layer at 0.2 units under varying weather conditions. Only a measurement of the moisture in the root zone layer was assumed to be available. An extended Kalman filter is used to estimate the states. The sampling time of the MPC controller is 1 day with a prediction horizon of 14 days. For conducting a realistic simulations study, weather data of the Jalgaon region of India is used.

The results of the simulation study are presented in Figure 5.4. It can be seen that the reference is tracked accurately until 200 days beyond which point in time the moisture is more than necessary. This is because of the monsoon season where the precipitation is much higher compared to the other time of the year. It can be seen that the irrigation is switched off to 0 during this period which is the best the controller can do.

In addition to the simulation study with the estimator, a simulation study in the presence of different soil types in the same agricultural field was performed assuming full-state information. Three types of soil are considered: clay loam, clay and sandy loam. The distribution of the soil in the field can be seen in Figure 5.3. The results of the simulation study can be seen in Figure 5.5. It can be seen that in the first 200 days, the soil moisture content of the different soil types can be kept close to the reference with a single irrigation strategy.

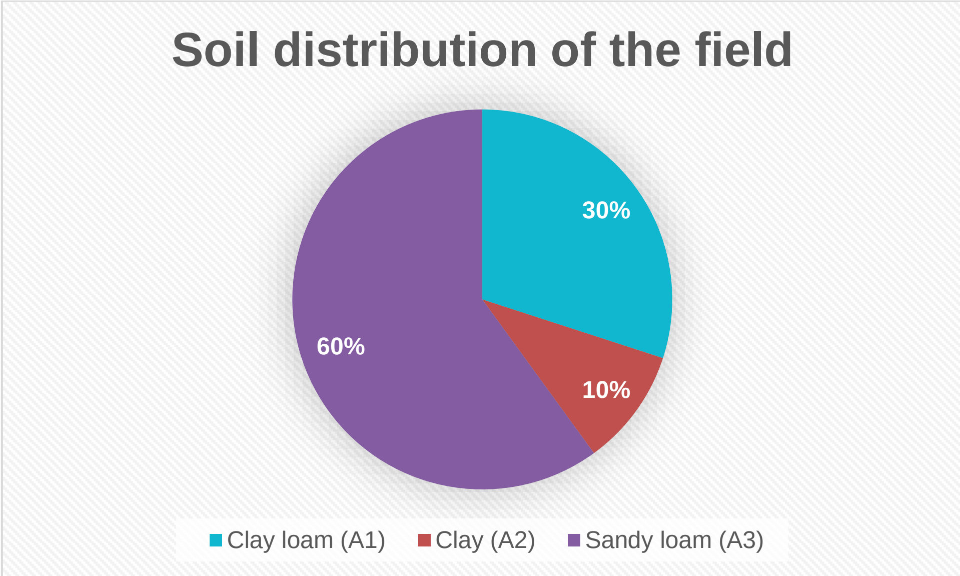

Figure 5-3: Distribution of the soil type considered in the simulation study.

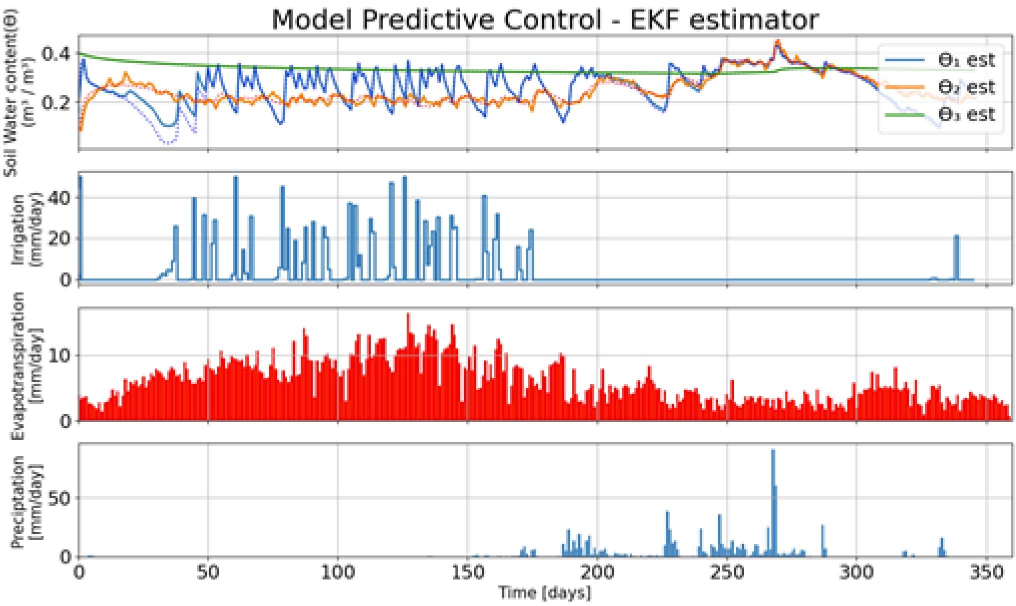

Figure 5-4: Simulation results obtained by applying the model-predictive control scheme for the irrigation use case estimating the states using the extended Kalman filter.

LOTUS is co-funded by the European Commission under the Horizon 2020 research and innovation programme under Grant Agreement N° 820881 and by the Indian Government, Ministry of Science and Technology.

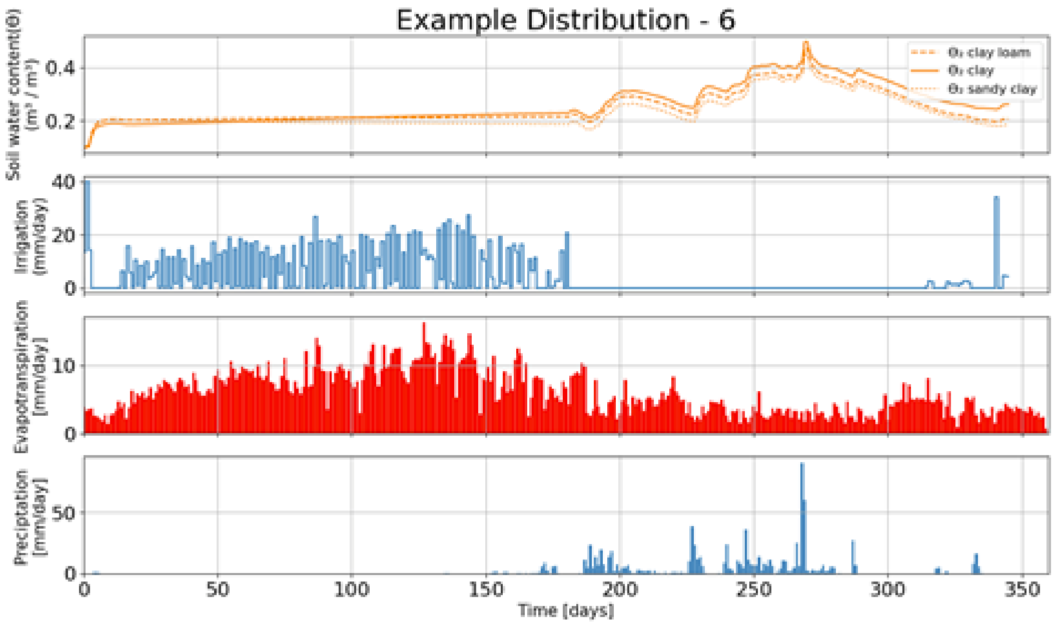

Figure 5-5: Simulation results obtained by applying the model-predictive control scheme for the irrigation use case for the field with three different soil types controlled using a single control input.

5.1.1.3 Current focus

Currently, the objective of the controller is chosen to minimize the irrigation water achieving the bound on the moisture level. This will be followed by the optimization of the growth of the crops and the optimization of fertigation strategies. The conrollers will be implemented using the LOTUS sensor together with commercially available sensors.

References:

[1] Sellers, P. J., Randall, D. A., Collatz, G. J., Berry, J. A., Field, C. B., Dazlich, D. A., ... & Bounoua, L. (1996). A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation. Journal of Climate, 9(4), 676-705.

[2] Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). FAO Irrigation and drainage paper No. 56. Rome: Food and Agriculture Organization of the United Nations, 56(97), e156.

6 Wastewater Treatment

6.1 Subtask 4.2.6

Two types of wastewater systems were studied: an algae-based wastewater treatment process and an activated sludge treatment process.

6.1.1 Activated Sludge Treatment Process

The activated sludge treatment process is a conventional water treatment that is used widely across the planet and also, mainly in the context of municipal wastewater treatment in various cities across India. One of the critical challenges in such treatment plants is to maintain the effluent water quality under various uncertainties, especially strongly varying inflows, and to minimize the consumption of electric power. Advanced control methodologies such as model predictive control (MPC) can reduce the power consumption and greatly improve the effluent water quality that is released into the freshwater streams because of its ability to predict the future behavior of the systems using the models while respecting the constraints. MPC based on state-space models (as opposed to conventional input-output-based MPC schemes) can be especially beneficial because it can help in controlling the inner states of the plant that are not measured explicitly. For the successful application of a state-space-model-based MPC scheme, three aspects are necessary: a model that reasonably represents the reality, a software tool that reliably solves optimization problems online, and reliable measurements of critical states of the system.

A reliable sensor is necessary to accurately estimate the various parameters of the model and to provide online measurement information of critical states. The LOTUS sensor is especially beneficial because it can measure various states at the same time while being cost-effective. Without the presence of the measurement information, MPC cannot operate the plant optimally to achieve the desired goals. In the following, the critical steps involved in the development of the online computation tool for efficient control are presented.

The activated sludge treatment process has have been studied for more than a century and detailed benchmark models are available. Benchmark simulation model No.1 is used here along with its parameters for the simulation study using the model predictive control framework. The benchmark model can be adapted to the local plant conditions by conducting experiments to adapt the parameters to the specific situation. LOTUS sensors together with commercially available sensors can be used to provide the data which is needed to estimate the parameters to a reasonable accuracy. The simulation studies have been performed for the original parameters provided in the benchmark model to test the applicability of the nonlinear MPC framework for this application.

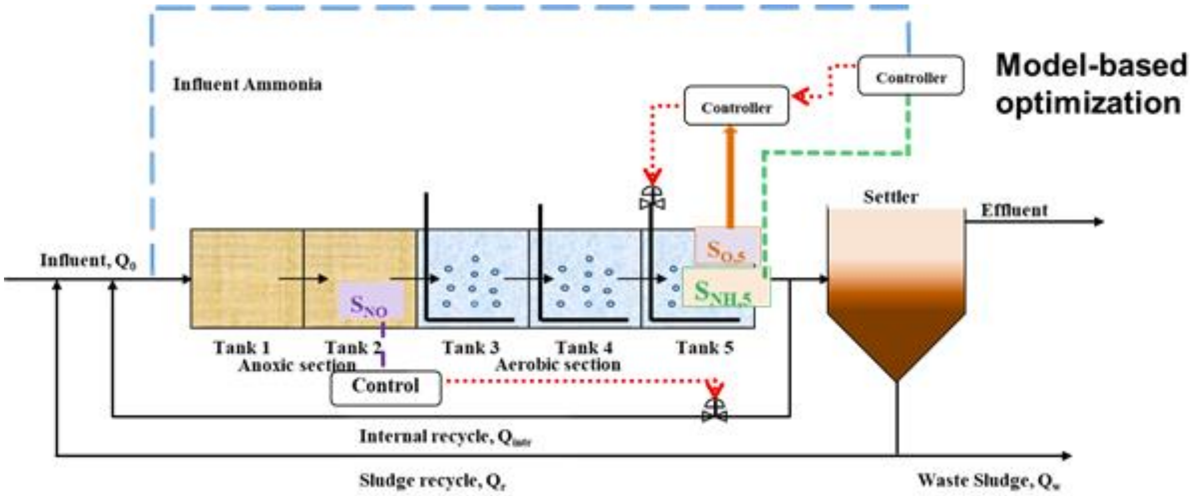

Figure 6-1: Illusration of the activated sludge process

The illustration of the process can be seen in Figure 6.1. It consists of five tanks with two sections: an anoxic section and an aerobic section. Each tank is represented by eight basic biochemical processes using 13 state variables. Note that the resulting model is nonlinear because the Monod equations are used to model the growth (as well as the decay) of both autotrophic and heterotrophic bacteria. The settler is modeled as a 10 layer non-reactive unit. This results in a total of 195 states to model the whole plant (more details can be found in [1]).

Control goals

As shown in Figure 6.1, there are two control loops the goals of which are to keep the nitrate and nitrite nitogen concentration at 1 g N/m3 in Tank 2 and dissolved oxygen concentration at 2 mg -COD/l in tank 5. To achieve these goals, a nonlinear model predictive control scheme was designed using do-mpc, an open-source python library for the development of model predictive control applications created by TU Dortmund (www.do-mpc.com). It uses the open-source tools CasADi (https://web.casadi.org/) and Ipopt (https://coin-or.github.io/Ipopt/) for solving optimization problems.

Simulation study under perfect information

Initially, full-state information was assumed to validate the performance of the controller. A sampling time of 15 minutes was chosen with a prediction horizon of 1 hour. The simulation is carried out for at first for one day of constant influent conditions followed by a realistic but challenging influent condition representing rainy weather. The dynamic influent conditions for 14 days are given in the plot below. The dynamic variations of the influent conditions can be clearly seen.

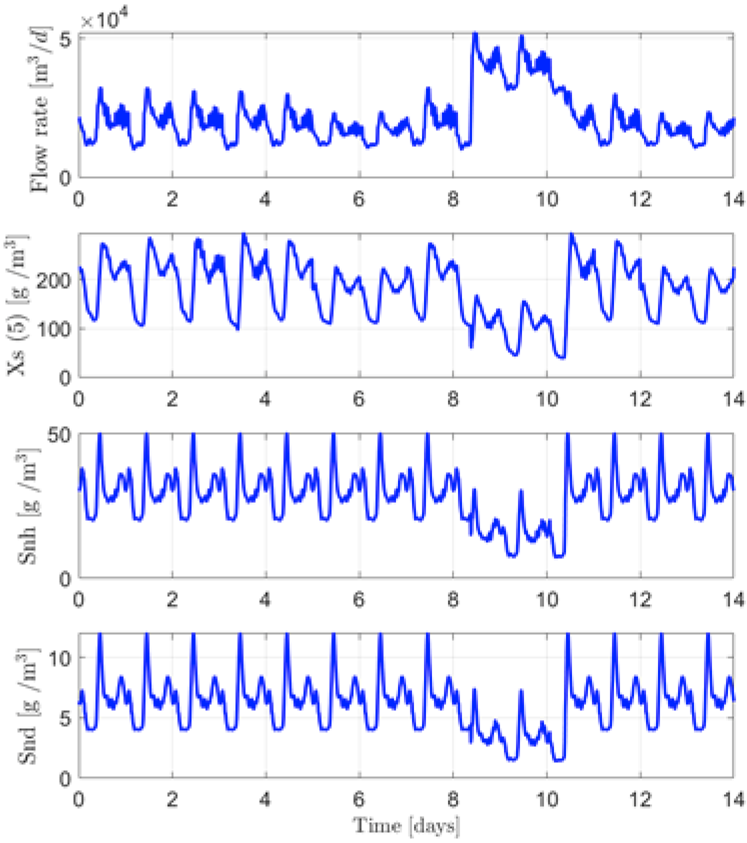

Figure 6-2: : Influent trajectories of flow rate, slowly biodegradable substrate (Xs), ammonia and ammonium nitrogen (Snh), soluble biodegradable organic nitrogen (Snd) for 14 days used for the simulation studies.

The optimization problem is formulated with a quadratic stage cost to track the concentration of S_{NO} in the reactor 2 and DO in reactor 5 and is given by $(S_{NO}(2) - 1)^2 + (DO(5) - 2)^2$. The simulation results are given below.

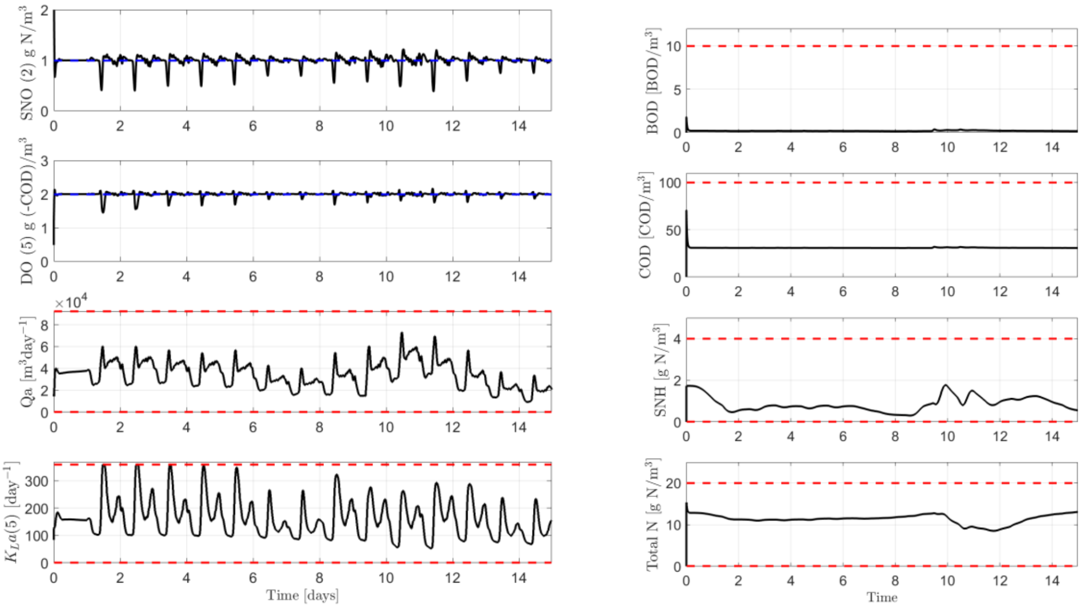

Figure 6-3: The closed-loop trajectories of nitrate and nitrite nitrogen concentration are reactor 2 S_{NO}, dissolved oxygen in reactor 5 DO, the internal recirculation flow rate Q_a and oxygen transfer rate in reactor 5 K_La obtained by applying the model-predictive controller (on the left). The resulting trajectories of effluents BOD, COD, ammonium and ammonia nitrogen (S_{NH}), total nitrogen (on the right).

From Figure 6.3 (left), it can be seen that the controller tracks the reference values of S_{NO} and DO perfectly under constant influent conditions and very well in the presence of dynamic variations of the influents. The manipulated variables (internal recirculation flow) and oxygen transfer rate are also shown. On the right of Figure 6.3, the effluent trajectories are shown with the required bounds shown in red-dashed lines. It can be seen that the effluents satisfy the required bounds at all time steps. Based on the variations in the influent conditions, the controller efficiently manipulates the control inputs to achieve the desired reference values and this results in the satisfaction of the constraints on the effluent trajectories at all time. The controller was also validated for dry weather and stormy weather conditions, where a similar performance of the controller was observed.

6.1.1.1 Benefits of the LOTUS sensor

In reality, not all state measurements are available. However, measurement information is crucial to estimate the current state of the system that can be then used in closed-loop control. Another set of simulation studies was performed by assuming that certain states are measured can either by using the LOTUS sensor or by using other commercially available sensors. The concentrations of dissolved oxygen, nitrate and nitrite nitrogen, ammonium and ammonia nitrogen, and alkalinity are assumed to be measured in all tanks. The measurement frequency is assumed to be 1 minute with a normally distributed measurement noise with a standard deviation of 0.1. The initial conditions are assumed to be uncertain by 20% in the state estimates. The states are estimated using an extended Kalman filter that combines both the model information and the measurement information at every time step to estimate the inner states of the system.

The simulation results can be seen in Figure 6.5. It can be seen that the plant trajectories track the references well and the results are very close to the full-state information case. Also, the effluent trajectories satisfy the required constraints at all times and the trajectories are similar to the full-state information case shown in Figure 6.4. This clearly shows the advantages that the LOTUS sensor can bring to the closed-loop control to achieve the desired results and not all states are required to be measured.

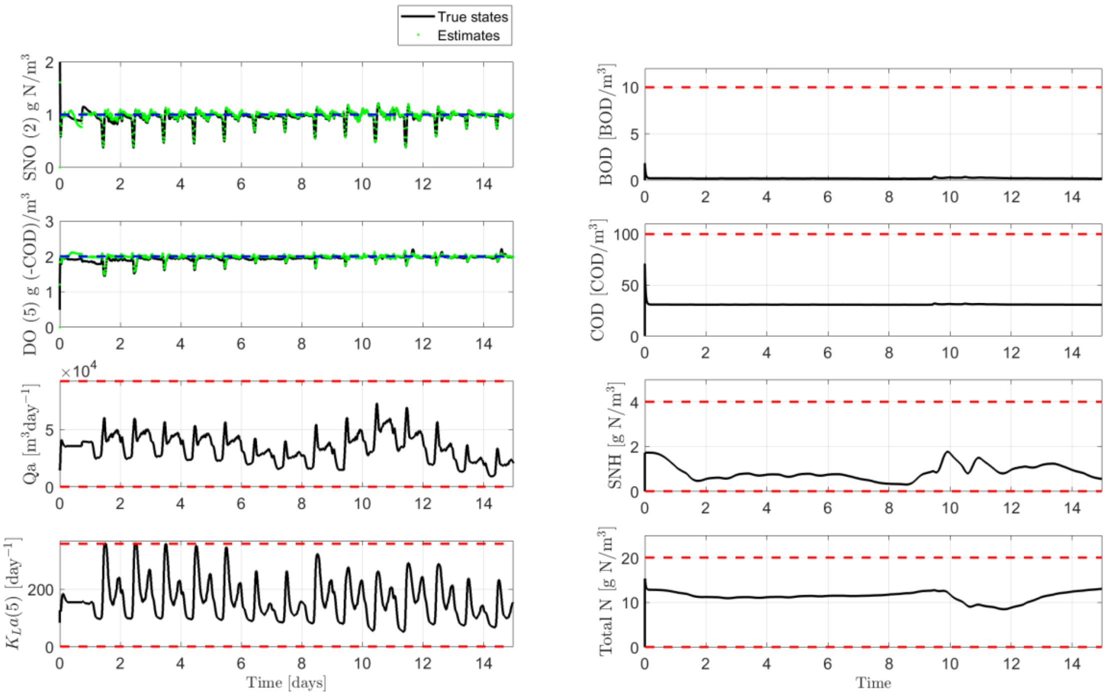

Figure 6-4: The closed-loop trajectories of nitrate and nitrite nitrogen concentration are reactor 2 S_{NO}, dissolved oxygen in reactor 5 DO, the internal recirculation flow rate Q_a and oxygen transfer rate in reactor 5 K_La obtained by applying the model predictive controller by measuring only certain states (on the left). The resulting trajectories of effluents BOD, COD, ammonium and ammonia Nitrogen (S_{NH}), total nitrogen (on the right).

Current focus

Currently, we are focussing on taking into account the uncertainty of the model that will be handled using the multi-stage NMPC scheme together with an economic objective to optimize energy usage.

6.1.2 Algae-based wastewater treatment

The algae-based wastewater treatment plant under construction at NEERI in Chennai is planned to be operated as a batch process, where the nutrients in the wastewater are used by the algae to increase their biomass. The control goal here is to maintain the desired pH level which is optimal for algae growth under various lighting and temperature conditions. As in conventional wastewater treatment, the measurement information from the LOTUS sensor together with commercially available sensors can help significantly in achieving this goal.

Model building

The algae growth depends on the local environmental conditions and varies across the globe. Hence a general benchmark formulation that can work in all conditions is difficult to obtain. However, various studies on such treatment processes are available. We investigated three such cases: the modified river water quality model as presented in [2], kinetic modeling for algae cultivation [3], and the algae-bacteria growth (ALBA) model in raceway the ponds [4]. In addition, a simple model based on first principles and nutrient availability are also developed. After thorough investigation, the model given in [2] will not be useful for the use case that we are investigating. Both [3] and [4] have the potential to model the plant that will be built at NEERI, Chennai. Model [4] is more detailed and consists of various interactions between bacteria and algae in open raceway ponds. The effects of light, temperature and pH are studied to aid algae growth.

Since the race way pond is planned to be implemented as a batch process, different manipulations to control the algae growth are investigated. For maintaining a suitable pH levels at the tank, weak acids as well as salts are considered to be the two control inputs that will be added to the tank at a reasonable interval. The frequency of manipulation is still under investigation. Rather than tracking a particular pH level, the upper and lower bounds for pH can be formulated and the bounds can be met using the model predictive control framework.

Secondly, during peak summer, the temperature of Chennai can reach as high as 42°C consistently. Though locally grown algae can withstand such temperature for a certain amount of time, prolonged exposure to heat can result in faster decay rate of algae leading to poor nutrient removal. Since, no manipulations are available for temperature control, a semi-batch type process is studied where in nutrient feed-rate can be a control input. Since the influents can be stored away from the sun, feeding additional nutrients can aid algae growth in addition to reducing the temperature of the raceway pond. In addition, it is also expected to provide a better penetration of light that can aid algae growth compared to a pure batch process (under the assumption that the turbidity of the wastewater is lower than that of raceway pond which also includes algal biomass).

6.1.2.1 Benefits of the LOTUS sensor

As in the case of a conventional wastewater treatment system, simulation studies were carried out to study the benefits of the sensor information. It was observed that by measuring the concentrations of the critical nutrients, all the inner states of the model can be estimated with reasonable accuracy. This can then be used to improve the growth of algae species quickly based on the available weather forecast. For example, in the kinetic model in [3], all the states can be estimated by measuring concentrations of phosphate, carbon and nitrates. Similarly, all 17 states of the ALBA model can be estimated just by measuring inorgonic carbon, phosphorus, ammonium and ammonia nitrogen, nitrates, and dissolved oxygen which can either be measured using the LOTUS sensor or commercially available sensors.

6.1.2.2 Current focus

Currently, the results of the estimation and control of the system are investigated with the goal to be adapted to user needs. The uncertainties in the plant model are studied to achieve required robustness against the uncertainties in the real plant. Model predictive control will be implemented with the goal of maximizing algae growth rate. In addition to achieving required effluent purity, the strategies to increase the bio-fuel yield will be investigated and the results will be documented.

References:

[1] Alex, J., Benedetti, L., Copp, J., Gernaey, K. V., Jeppsson, U., Nopens, I., ... & Winkler, S. (2008). Benchmark simulation model no. 1 (BSM1). Report by the IWA Taskgroup on benchmarking of control strategies for WWTPs, 19-20.

[2] Pierong, R., Nehrenheim, E., Carlsson, B., & Zambrano, J. (2016). Algae based wastewater treatment model using the RWQM1. In IWA World Water Congress & Exhibition.

[3] Eze, V. C., Velasquez-Orta, S. B., Hernández-García, A., Monje-Ramírez, I., & Orta-Ledesma, M. T. (2018). Kinetic modelling of microalgae cultivation for wastewater treatment and carbon dioxide sequestration. Algal Research, 32, 131-141.

[4] Casagli, F., Zuccaro, G., Bernard, O., Steyer, J. P., & Ficara, E. (2021). ALBA: A comprehensive growth model to optimize algae-bacteria wastewater treatment in raceway ponds. Water Research, 190, 116734.

LOTUS is co-funded by the European Commission under the Horizon 2020 research and innovation programme under Grant Agreement N° 820881 and by the Indian Government, Ministry of Science and Technology.