Project Title: NAWI TASK 5.17 - (Salveson, Andy) Data-Driven Fault Detection and Process Control for Potable Reuse with Reverse Osmosis

**PROGRESS AND STATUS BY SUBTASK:**

# **Task: Task 5.17.1** – Desktop Evaluations

## **Task:** **Subtask 5.17.1.5 –** Data-Driven Model Optimization (DDMO) for Chloramine and Anti-Scalant Dosing

**Subtask Lead: Steve Hayden**

Research Questions:

* How much cost and energy could be saved across the reuse treatment train applying DDMO to adjust the pre-chloramine and antiscalant doses and predict a fault in real-time in response to water quality changes?

**PROGRESS AND STATUS:**

* Collected data from:
  + Las Virgenes Municipal Water District
    - All necessary data is available and shared among the team
  + Orange County Water District
    - All necessary data is available and shared among the team
  + West Basin Municipal Water District
    - Additional data from WBMWD is still needed.

【下記、Q3の構成】

### Budget Spent (YCA)

### Summary

In Task 5.17.1.5, data driven model optimization is used to optimize chloramine and anti-scalant dosing ahead of RO. In this report, we report data analysis underway for the model creation based on OCWD’s full scale plant and LVMWD’s demonstration plant. A key study objective is chemical optimization (e.g., anti-scalant) in the RO operation at both plants.

...

### Desktop Evaluation based on OCWD

#### Water Quality Prediction

We received data measured at 30-minute intervals in RO B01 unit from OCWD. The RO system configuration of OCWD can be found in Additional Materials A1. RO permeate electric conductivity (EC) and TOC, which are considered an indicator of RO membrane scaling, is focused to predict the behavior based on the statistical model. Each permeate EC is measured at each stage permeate in B01 unit and permeate TOC is measured at the whole RO permeate.

* Data Preprocessing

We used continuous data from May 13th 2022 to November 18th 2022 and excluded data from other periods because it contains data that is not properly measured or significantly different trends.

There are many noise components in the received raw data. The data preprocessing for them was needed before data analysis. Figure 1.1 shows a summary diagram of the data preprocessing. The flow chart in Figure 1.1 explains the data preprocessing approach including the outlier filtering, the smoothing, and the missing data imputation. The outlier filtering is to remove outlier data by imposing upper-lower bounds as a first step. The smoothing is to smooth the time series data by moving averages with XX window steps. The missing data imputation is to fill missing raw data by copying previous and next actual data. (i) 瞬時的に大きく外れる値を上下限を設けて除外, (ii) Hampleフィルターを用いて半窓サイズ48ステップ、3σの設定で補間, (iii) 前の値で欠損値を補完, (iv) 標準化 and, We converted the chemical dosing usage data with total daily interval into 30-minute interval data by dividing 48 steps.

Permeate EC [] at each stage and permeate TOC from [start day] to [end day] are plotted with raw data and the preprocessed in Figure 1.2 (a)-(f). By preprocessing the data, outliers are removed, and the data behavior can be extracted.

* Water Quality Prediction Model
* Mathematical Prediction Model

We constructed four models for permeate EC at 1st, 2nd, and 3rd stage and permeate TOC at the whole RO. The prediction model is based on the multiple regression model (MLR), which is linear statistical model. The mathematical model of MLR is shown as follows:

Here, () denotes explanatory variable, denotes predictive variable, () denotes coefficient parameter, and denotes bias parameter. The parameters () are estimated by actual data in the training period and are shown as follows:

Here, denotes the design matrix, denotes the -th explanatory vector, denotes the predictive vector, and denotes the -th training dataset. The standard coefficient means relative contribution of each variable for prediction and calculated by using the following equation:

Here, is the standard deviation of training data .

Table1.1 shows the variable list for each water quality prediction model. “X” means explanatory variable, “Y” means predictive variable, and “\*” means calculated variable. Tag name list is shown in Additional Materials A1.

The calculated 1st stage feed flow rate , 2nd stage feed flow rate , 3rd stage feed flow rate are shown in below:

Here, denotes concentrate flow rate at each stage, denotes permeate flow rate at each stage.

The law of conservation of mass is established around each RO stage. Based on the low, the 2nd stage feed EC and 3rd stage feed EC are calculated using following equations:

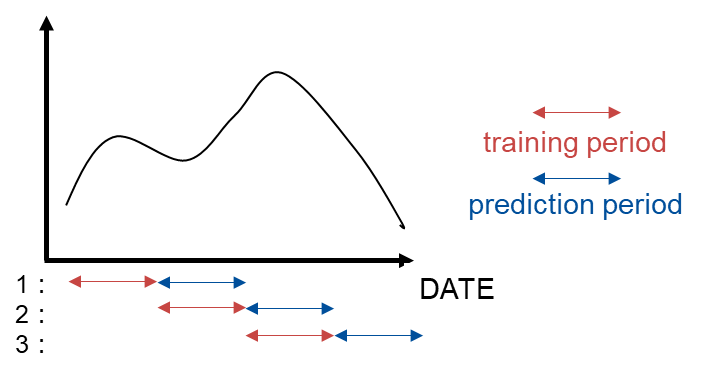
In the optimization phase, we need to predict permeate EC by the prediction model and calculate next stage feed EC in the prediction period by using the above equations and predicted permeate EC.

**Table 1.1: Variable List in Water Quality Prediction Model (OCWD)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Stage1 Permeate EC | Stage2 Permeate EC | Stage3 Permeate EC | Permeate TOC |
| RO B01 Feed Pressure | X |  |  | X |
| RO Feed EC | X |  |  | X |
| RO Feed Flow Rate \* | X |  |  | X |
| RO Feed Temperature | X | X | X | X |
| RO Feed pH | X |  |  | X |
| Sulfuric Acid Usage | X | X | X | X |
| Threshold Inhibitor Usage | X | X | X | X |
| RO B01 Stage 2 Feed Pressure |  | X |  | X |
| RO B01 Stage 2 Feed EC \* |  | X |  |  |
| RO B01 Stage 2 Feed Flow Rate \* |  | X |  |  |
| RO B01 Stage 3 Feed EC \* |  |  | X |  |
| RO B01 Stage 3 Feed Flow Rate \* |  |  | X |  |
| RO Feed TOC |  |  |  | X |
| RO B01 Stage 1 Permeate EC | Y |  |  |  |
| RO B01 Stage 2 Permeate EC |  | Y |  |  |
| RO B01 Stage 3 Permeate EC |  |  | Y |  |
| RO Permeate TOC |  |  |  | Y |

* Training and Period Term

We conducted sequential prediction by linear model and evaluated from May 2022 to November 2022. Figure 1.1 shows overview of sequential prediction and Table 1.2 shows training and prediction period in each term. For example, training period is from May 13th 2022 to May 19th 2022 (1week) and prediction period is from May 20th 2022 to May 27th 2022 (1week) in the first term. In the next term, we slide these periods to 1week later and conduct re-training (model re-tuning). Therefore, the size of training data set is and the size of prediction data set is in the single term.



**Figure 1.1: Overview of Sequential Prediction**

**Table 1.2: Training and Prediction Period in Each Term**

|  |  |  |
| --- | --- | --- |
| **Term** | **Trainig Period** | **Prediction Period** |
| 1 | from May 13th to May 19th | 5月20日～5月26日 |
| 2 | 5月20日～5月26日 | 5月27日～6月2日 |
|  |  |  |
| 27 | 10月28日～11月3日 | 11月4日～11月10日 |
| 28 | 11月4日～11月10日 | 11月11日～11月17日 |

* Prediction Evaluation Index

The evaluation indexes are Mean Absolute Percentage Error (MAPE) [%] and Root Mean Percentage Error (RMSE) . MAPE indicates a relative prediction error and RMSE is an absolute prediction error. They are formulated as follows:  
Here, is the actual data and is the predicted data at time . is the length of the prediction period. But we excluded outlier data in the data preprocessing from evaluation data.

* Predicton Result

Table1.4に各予測モデルにおける誤差評価結果を示す。Figure1.2に各予測結果のトレンドを示す。横軸が日付。青線が各目的変数の補正済み実績値、赤線が予測値。

Figure1.2の評価期間2022年6月10日～6月16日の結果を抜粋したものをFigure1.3に示す。Stageの導電率も周期的な変動を捉えることができた。逐次的にモデルを作成することで、評価する期間に近い傾向を学習できるため実績に近しい値が予測できた。

説明変数に大きな変動があったときや、結果変数である透過側導電率の傾向が変化した期間（例えば1st Stage 2022年7月中旬から8月上旬[グレーの領域]）では学習と評価期間の傾向が異なるため予測が難しい。

**Table 1.4: Prediction Accuracy Result (OCWD)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Index** | **Stage1 Permeate EC** | **Stage2 Permeate EC** | **Stage3 Permeate EC** | **Permeate TOC** |
| **RMSE** | 2.28 [uS/cm] | 6.26 [uS/cm] | 15.25 [uS/cm] | 0.014 [ppm] |
| **MAPE** | 10.33 [%] | 13.40 [%] | 15.59 [%] | 12.75 [%] |

グラフ

自動的に生成された説明

**(a): 1st Stage Permeate Conductivity**

グラフ, 棒グラフ

自動的に生成された説明

**(b): 2nd Stage Permeate Conductivity**

グラフ, 棒グラフ

自動的に生成された説明

**(c): 3rd Stage Permeate Conductivity**

グラフ, 折れ線グラフ

自動的に生成された説明

**(d): Permeate TOC**

**Figure 1.2: Prediction Trend in All Period (OCWD)**

グラフ, 折れ線グラフ

自動的に生成された説明

**(a): 1st Stage Permeate Conductivity**

グラフ, ヒストグラム

自動的に生成された説明

**(b): 2nd Stage Permeate Conductivity**

グラフ, 折れ線グラフ

自動的に生成された説明

**(c): 3rd Stage Permeate Conductivity**

グラフ, 折れ線グラフ, ヒストグラム

自動的に生成された説明

**(d): Permeate TOC**

**Figure 1.3: Prediction Trend from 10th to 16th June 2022 (OCWD)**

* + - 各モデルの時期ごとの回帰係数比較

Figure1.4に、1st Stage Permeate Conductivityと透過側のTOC予測モデルの回帰係数を示す。横軸に1週間で学習した時の最初の学習日を示している。Figure1.4(a)から、透過側の導電率の予測には、橙色線のFeedの導電率の値がどの期間でも正の値であり、透過側導電率の傾向が変化した期間を除き、大きな影響を持っている。一方、Figure1.4(b)から、透過側TOCの予測には、橙色線のFeed TOCの値の係数が常に正であり、大きな影響を持っている。  
　Figure1.5に、Figure1.4と同時期のFeed Flow RateとFeed pHの実績データトレンドを示す。Figure1.4(a)の水色線と灰色線から、1st　Stageの透過側導電率のモデルでは、Feedの流量とpHの回帰係数は小さいため、重回帰モデルでは変数としての影響度は低い。これは、Feedの流量とpH の値はおおよそ4080[gpm], 6.9で一定のためである。

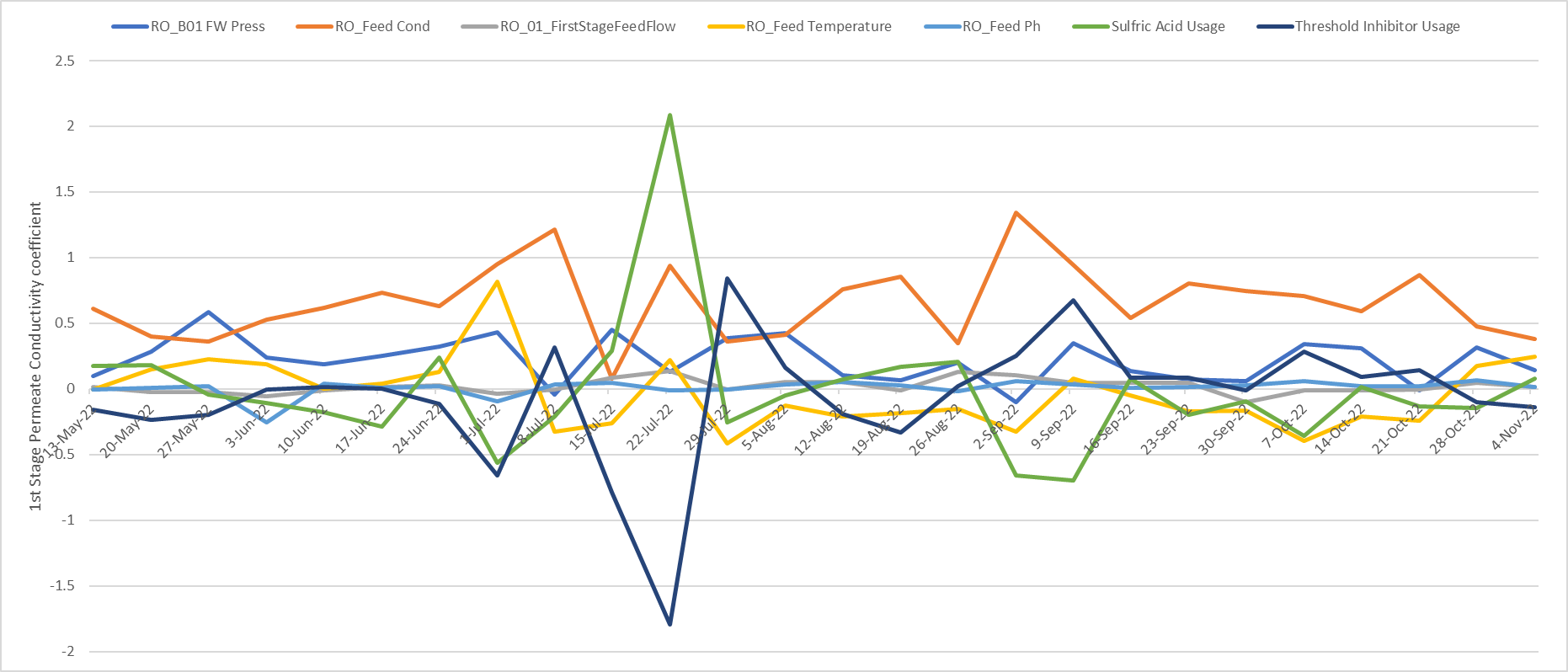
パソコンの画面

低い精度で自動的に生成された説明グラフ

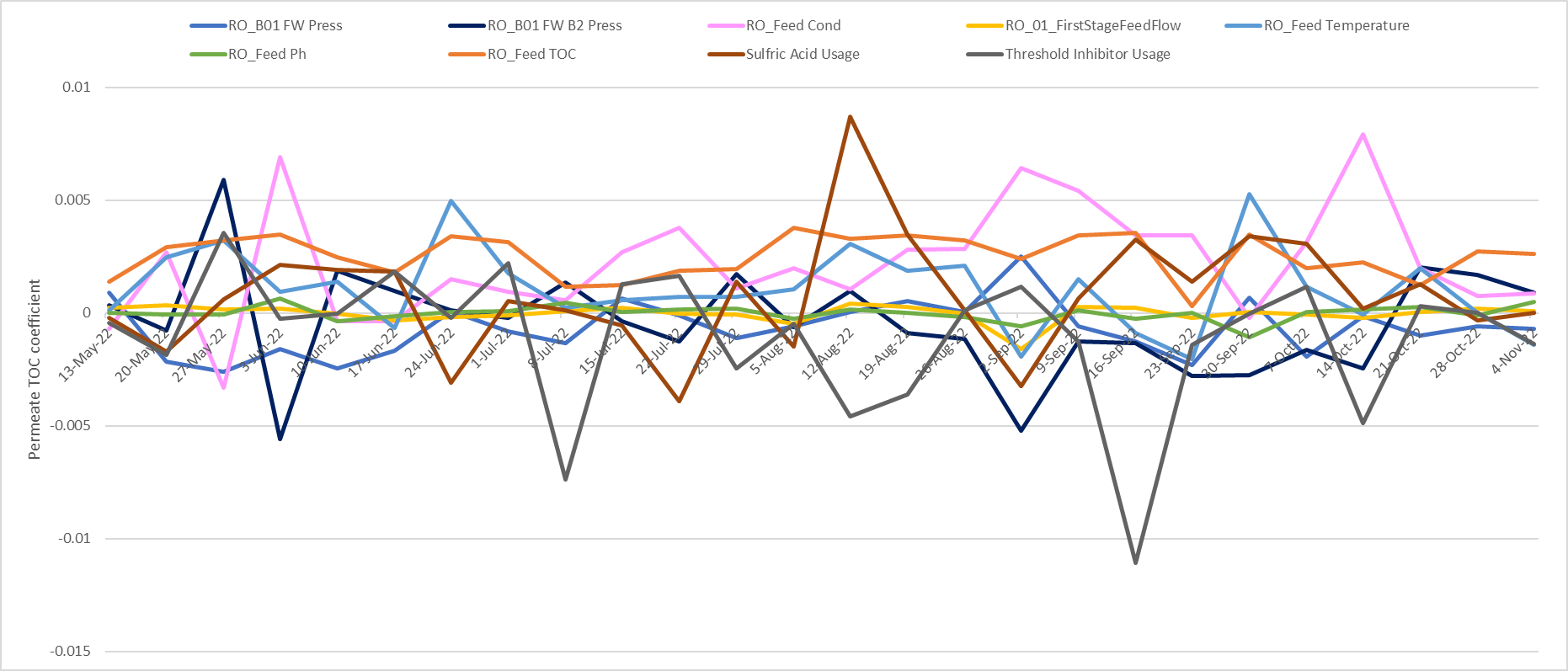
自動的に生成された説明

**(a): 1st Stage Feed Flow Rate (b): Feed pH**

**Figure 1.5: Actual Data Trend in RO Feed (OCWD)**



**(a): 1st Stage Permeate Conductivity**



**(b): Permeate TOC**

**Figure 1.4: Coefficients of Prediction Model Trend in All Period (OCWD)**

#### RO Membrane Scaling using Xact Data

* 基本方針

膜に供給水を通すと膜表面に近づくにあたって各成分の濃度が上昇する濃度分極という現象が発生するため、供給水の濃度と膜表面の濃度が一般的に異なることが知られている。そのため、RO膜表面の濃度を計算し、その結果に応じて各成分が析出するか理論的に計算できれば、膜表面における析出可能性に対応した薬液注入が可能になる。その中で本レポートにおいてはRO膜表面濃度算出について述べる。

* RO膜表面濃度計算アルゴリズム

膜表面濃度計算フローは以下の通りである。

* + 膜表面と透過水浸透圧を計算
  + 浸透圧をもとに、透過流束を計算
  + 浸透圧と透過流束から物質移動係数を計算
  + 物質移動係数と原液・透過水濃度から膜面濃度を計算

まず、浸透圧を計算することを考える。浸透圧と濃度間には式(1.1)のファントホッフの法則が成立する。

ただし、Rは気体定数、Ｔは温度、下付き文字に関しては、iはRO膜のStage数、jはどの位置(供給水、膜面、透過水)の濃度か、kはどの溶質かを表す。  
式(1.1)の通り、浸透圧を計算するためには各位置における濃度を計算する必要がある。ただし、膜面の濃度や2nd Stageの供給水・濃縮水濃度など計測していない項が数多く存在するため、以下を仮定する。

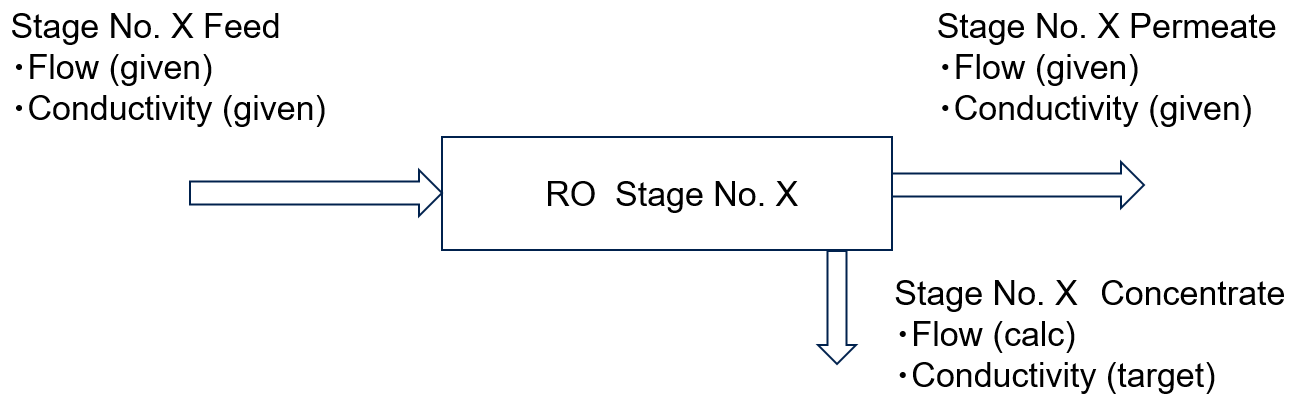
* 膜面の濃度は濃縮水の濃度とする
* X-1 Stageの濃縮水濃度とX Stage供給水濃度が等しい
* 1st Stage供給水側と各Stageの透過側濃度はX Actのデータを用いる
* 計測していない地点の濃度は式(1.2)に従う
* 計測していない地点の電気伝導率は次ページFigure 1.1と式(1.3)に従う

次に、透過流束を計算することを考える。溶媒と溶質それぞれの透過流束は次のように求めることができる。ただしは純水の透過係数である。

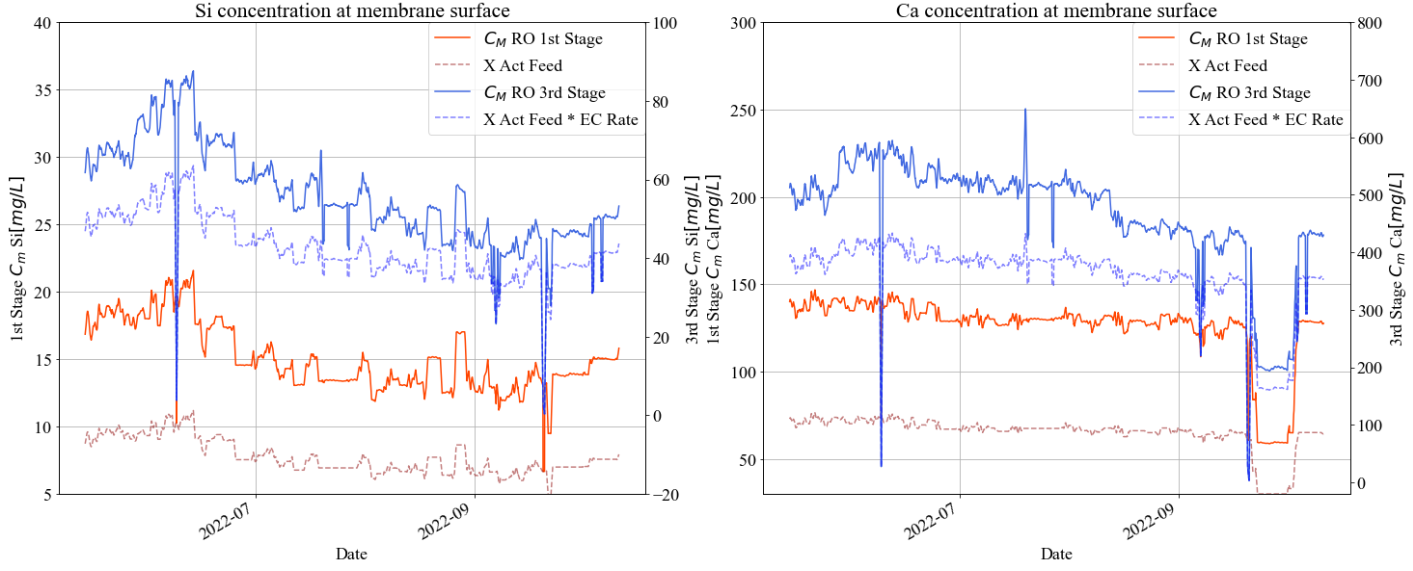
式(1.4,1.5)の計算結果を式(1.6)に代入すれば、Stage *i*,溶質*k*の物質移動係数を算出することができる。

最終的に、物質移動係数をもとに最終的に膜面の濃度は式(1.7)によって導出できる。

式(1.7)を用いて膜面の濃度を計算した結果が次ページのFigure1.2の通りであり、Figure1.2(a)がSiの膜面濃度、Figure1.2(b)がCaの膜面濃度を表現している。各グラフにて第一軸が1st Stageの濃度、第二軸が3rd Stageの濃度であり、Stage毎に見ると、点線で示した供給水の濃度と比較して実線の膜面濃度が150 ~ 200%程度濃縮されていることがわかる。またStage間でみると1st Stageより3rd Stageの濃度が高いことから、徐々に濃縮されてゆく過程をシミュレーションできていることが見て取れる。



**Figure 1.1:** 各Stage前後における流量と電気伝導率の関係



**(a):** Siの膜面濃度  **(b):** Caの膜面濃度

**Figure 1.2:** 1st&3rd Stageにおける供給水と膜面濃度の比較

#### RO Optimization Simulation Model

* Summary

We formulated RO optimization problem for OCWD and constructed a RO optimization simulation model implemented by the water quality prediction or scaling model. We will carry out the optimization simulation and calculate an improvement effect compared actual operation in the past period in the final report.

* RO Optimization Problem for OCWD

Basic optimization story for OCWD is to decrease the chemical dosing cost (e.g., antiscalant or pH adjustors) while satisfying water quality standards or monitoring a RO membrane scaling. Considering the above story, we formulated the RO optimization problem for OCWD.

* Scheduling Problem

RO optimization is to improve the operation in some period of time. This type of optimization problem is the scheduling problem, which is formulated and solved according to procedure as shown in Figure 1.1. The procedure consists of drawing flow chart, formulation mathematical optimization problem, and calculation operational schedule. Therefore, drawing a flow chart, variable, objective function, and constraint condition are required to construct a simulation model.

The operational schedule structure is shown as Table 1.1. denotes the -th variable schedule at time , denotes the -th variable schedule in all period, denotes all variables schedule in all period, denotes the length (step size) of the optimization period, and denotes the number of variables, respectively. For example, if the optimization period is 1 week of 30 minutes time interval data (e.g., 0:30, 1:00. 1:30, 2:00, …), we can estimate step.



**Figure 1.1: Procedure for Solving Scheduling Problem**

**Table 1.1: Operational Schedule Structure**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

* Flow Chart

Figure 1.2 shows the flow chart for OCWD based on the RO system configuration. The RO system configuration of OCWD can be found in Additional Materials A1. The storage shape means given data (fixed data), the hexagon shape means mathematical model, the square shape means demand data, especially mathematical models include the prediction model in the previous section.



1. **RO Total Flow Chart (Permeate TOC)**



1. **RO Each Stage Flow Chart (Permeate EC)**

**Figure 1.2: Flow Chart for RO Optimization Simulation (OCWD)**

* Mathematical Optimization Problem

Mathematical optimization problem is formulated as minimizing such that constraints (), where denotes the optimization variable, denotes the objective function (cost function), denote constraint functions, denotes the set of real numbers, denotes the number of total variables, respectively.

Optimization variable list is as shown in Table 1.2. “Optimization Variable” means cost or manipulated variable (chemical dose), “Fixed Parameter” is given by actual value, “Intermediate Variable” indirectly affects the objective function or constraints in the calculation (e.g., prediction value). In OCWD optimization model, “Optimization Variable” is “Sulfuric Acid Usage” and “Threshold Inhibitor Usage” and the number of the optimization variables .

**Table 1.2: Optimization Variable List (OCWD)**

|  |  |  |  |
| --- | --- | --- | --- |
| ID No. | Description | Unit | Opt. Variable / Fixed Parameter |
| ID0000 | Sulfuric Acid Usage | ton/day | Optimization Variable |
| ID0001 | Threshold Inhibitor Usage | ton/day | Optimization Variable |
| ID0002 | Feed EC |  | Fixed Parameter |
| ID0003 | Feed TOC | ppm | Fixed Parameter |
| ID0004 | Feed Temperature |  | Fixed Parameter |
| ID0005 | Feed pH | - | Fixed Parameter |
| ID0100 | Stage 1 Feed Pressure | psi | Fixed Parameter |
| ID0101 | Stage 1 Feed Flow Rate | gpm | Fixed Parameter |
| ID0200 | Stage 2 Feed Pressure | psi | Fixed Parameter |
| ID0201 | Stage 2 Feed Flow Rate | gpm | Fixed Parameter |
| ID0300 | Stage 3 Feed Flow Rate | gpm | Fixed Parameter |
| ID1000 | Stage 1 Permeate EC |  | Intermediate Variable |
| ID2000 | Stage 2 Feed EC |  | Intermediate Variable |
| ID2001 | Stage 2 Permeate EC |  | Intermediate Variable |
| ID3000 | Stage 3 Feed EC |  | Intermediate Variable |
| ID3001 | Stage 3 Permeate EC |  | Intermediate Variable |
| ID4000 | Permeate TOC | ppm | Intermediate Variable |

The objective function is formulated as follows:

Here, denotes threshold inhibitor usage at time , denotes the sulfuric acid usage at time , denotes cost coefficients, and denotes the length (step size) of optimization period, respectively.

Constraint conditions are as following:

Here, Equations (1.2), (1.3), (1.4), (1.5), (1.6), (1.7) denotes lower and upper limit, Equations (1.8), (1.9) denotes fluctuation range limit, Equations (1.10), (1.11), (1.12), (1.13) denotes water quality standards limit based on LRV, respectively.

Intermediate variables (Stage1 permeate conductivity), (Stage2 permeate conductivity), (Stage3 permeate conductivity), and (permeate TOC) are calculated by water quality prediction model as follows:

is Logarithmic Reduction Value (LRV) of water quality data (e.g., electric conductivity and TOC) as follows:

Here, denotes feed quality data and denotes permeate quality data at each stage.

The scale of the above problem can be roughly estimated such as number of total optimization variables and number of total constraints . For example, if the optimization period is 1week of 30 minutes time interval data, we can estimate steps, , and .

* Optimization Algorithm

The above problem is classified as constrained and black-box problem. We constructed and used an optimization algorithm by combining SHADE[[1]](#footnote-1) and Feasibility Rule[[2]](#footnote-2) for solving the problem. SHADE is known to the best unconstrained black-box optimization in the optimization field recently and often used in the BBOB competition[[3]](#footnote-3). Feasibility Rule is one of the constraint handling techniques, which enable unconstrained optimization algorithms to apply to the constrained optimization. We finished to write python source code from scratch and will apply it to the above problem.

### Desktop Evaluation based on LVMWD

#### Water Quality Prediction

We received data measured at 1minute or 1day intervals in RO unit from LVMWD. The RO system configuration of LVMWD can be found in Additional Materials A2. RO permeate electric conductivity (EC) and TOC, which are considered an indicator of RO membrane fouling, is focused to predict the behavior based on the statistical model. Each permeate EC is measured at each stage permeate and permeate TOC is measured at the whole RO permeate.

* Data Preprocessing

We used continuous data from July 1st 2021 to July 27th 2022 and excluded data from other periods because it contains data that is not properly measured or significantly different trends.

There are many noise components in the received raw data. The data preprocessing for them was needed before data analysis. The trend data is converted the chemical dosing usage data with 1minute interval into 30-minute interval data by using the method described in the 2nd quarterly report. After that, the data is preprocessed by using procedure as in the case of OCWD.

* Water Quality Prediction Model

We constucted water quality prediction model for LVMWD. They are MLR as in the case of OCWD model. Table2.1 shows the variable list for each water quality prediction model. “X” means explanatory variable, “Y” means predictive variable, and “\*” means calculated variable. Tag name list is shown in Additional Materials A2.

We conducted sequential prediction by linear model and evaluated from July 2021 to July 2022. Figure 1.1 shows overview of sequential prediction and Table 2.2 shows training and prediction period in each term as in the case of OCWD.

**Table 2.3: Variable List of each water quality prediction model (LVMWD)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Stage1 Permeate EC | Stage2 Permeate EC | Stage3 Permeate EC | Permeate TOC |
| RO Stage 1 Feed Pressure | X |  |  | X |
| RO Feed EC | X |  |  | X |
| RO Feed Flow | X |  |  | X |
| RO Feed Temperature | X | X | X | X |
| RO Feed pH | X |  |  | X |
| UF Filtrate Total Chlorine | X | X | X | X |
| RO Stage 2 Feed EC |  | X |  |  |
| RO Stage 2 Feed Flow |  | X |  |  |
| RO Stage 3 Feed Pressure |  |  | X | X |
| RO Stage 3 Feed EC |  |  | X |  |
| RO Stage 3 Feed Flow |  |  | X |  |
| RO Feed TOC |  |  |  | X |
| RO Stage 1 Permeate EC | Y |  |  |  |
| RO Stage 2 Permeate EC |  | Y |  |  |
| RO Stage 3 Permeate EC |  |  | Y |  |
| RO Combined Permeate TOC |  |  |  | Y |

* 短期的な予測の結果を確認すべく、1週間の学習期間と評価期間を連続で用意し、モデルの学習と評価を行った。Table2.1に学習期間と評価期間を示す。逐次予測における学習期間と予測期間の関係は、Figure1.1と同様にする。

**Table 2.1: 学習期間と評価期間**

|  |  |  |
| --- | --- | --- |
|  | **学習期間** | **評価期間** |
| 1 | 2021年7月1日～7月7日 | 7月8日～7月14日 |
| 2 | 7月8日～7月14日 | 7月15日～7月21日 |
|  |  |  |
| 54 | 2022年7月7日～7月13日 | 7月14日～7月20日 |
| 55 | 2022年7月14日～7月20日 | 7月21日～7月27日 |

* 透過側導電率と透過側TOCの予測モデルの使用変数

　Table2.2にTag Nameと変数名の対応関係を示す。また、Table2.3に各水質予測モデルの使用変数を示す。なお、〇X は説明変数として使用している変数で、〇Yは目的変数として使用している変数を表している。※1、※2は、計算Tagである。

* 評価指標

OCWDの時と同様に、RMSEとMAPEを使用。

* 結果

Table 2.4に各予測モデルにおける誤差評価結果を示す。Figure 2.1に各予測結果のトレンドを示す。横軸が日付。青線が各目的変数の補正済み実績値、赤線が予測値。

Figure 2.2の評価期間2022年3月10日～3月17日の結果を抜粋したものをFigure 2.2に示す。Stageの導電率も周期的な変動を捉えることができた。逐次的にモデルを作成することで、評価する期間に近い傾向を学習できるため実績に近しい値が予測できた。

**Table 2.4: Prediction Accuracy Result (LVMWD)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Index** | **Stage1 Permeate EC** | **Stage2 Permeate EC** | **Stage3 Permeate EC** | **Permeate TOC** |
| **RMSE** | 2.07 [uS/cm] | 2.36 [uS/cm] | 4.63 [uS/cm] | 0.030 [mg/L] |
| **MAPE** | 5.27 [%] | 5.04 [%] | 9.04 [%] | 19.46 [%] |

グラフ, ヒストグラム

自動的に生成された説明

**(a): 1st Stage Permeate Conductivity**

グラフ

自動的に生成された説明

**(b): 2nd Stage Permeate Conductivity**

グラフ

自動的に生成された説明

**(c): 3rd Stage Permeate Conductivity**

アンテナ, 立つ, グループ, 座る が含まれている画像

自動的に生成された説明

**(d): Permeate TOC**

**Figure 2.1: Prediction Trend in All Period (LVMWD)**

グラフ

自動的に生成された説明

**(a): 1st Stage Permeate Conductivity**

グラフ, 折れ線グラフ

自動的に生成された説明

**(b): 2nd Stage Permeate Conductivity**

グラフ, 折れ線グラフ

自動的に生成された説明

**(c): 3rd Stage Permeate Conductivity**

グラフ が含まれている画像

自動的に生成された説明

**(d): Permeate TOC**

**Figure 2.2: Prediction Trend from 10th to 17th March 2022 (LVMWD)**

Figure 2.3に、2021年7月～2022年7月の期間のFeed Conductivity、Feed TOC、Stage3透過側導電率の実績データトレンドを示す。Figure 2.3 (a)から、Feed Conductivity の灰色の領域（2021年7月～11月と2022年7月）は、～である。Figure 2.3 (b)から、Feed TOCの灰色の領域（2021年7月～11月と2022年6月～7月）は、～である。つまり、入口側のデータの振動が大きいため、予測値の振幅も大きい。

また、Figure 2.1 (c)から、Stage3透過側導電率の予測が大きく外れる期間(橙色)がある。これは、Figure 2.3 (c)から、この期間では導電率が大きく変化していることがわかる。つまり、学習期間と評価期間のデータ傾向が大きく変化すると適切な予測が難しい。

グラフ

自動的に生成された説明グラフ

自動的に生成された説明

**(a): Feed Conductivity (b): Feed TOC**

グラフ

自動的に生成された説明

**(c): 3rd Stage Permeate Conductivity**

**Figure 2.3: Actual Data Trend in RO Feed (LVMWD)**

* + 今後の課題

透過側TOCのデータについては、計測データが頻繁に0付近の微小な値をとる。そのため、今回の方法の前処理では、多くのデータが外れ値に該当し、生データのままでは、学習が難しいため外れ値の補間方法については再検討を行う。

#### RO Membrane Fouling (Soya-san)

* 基本方針

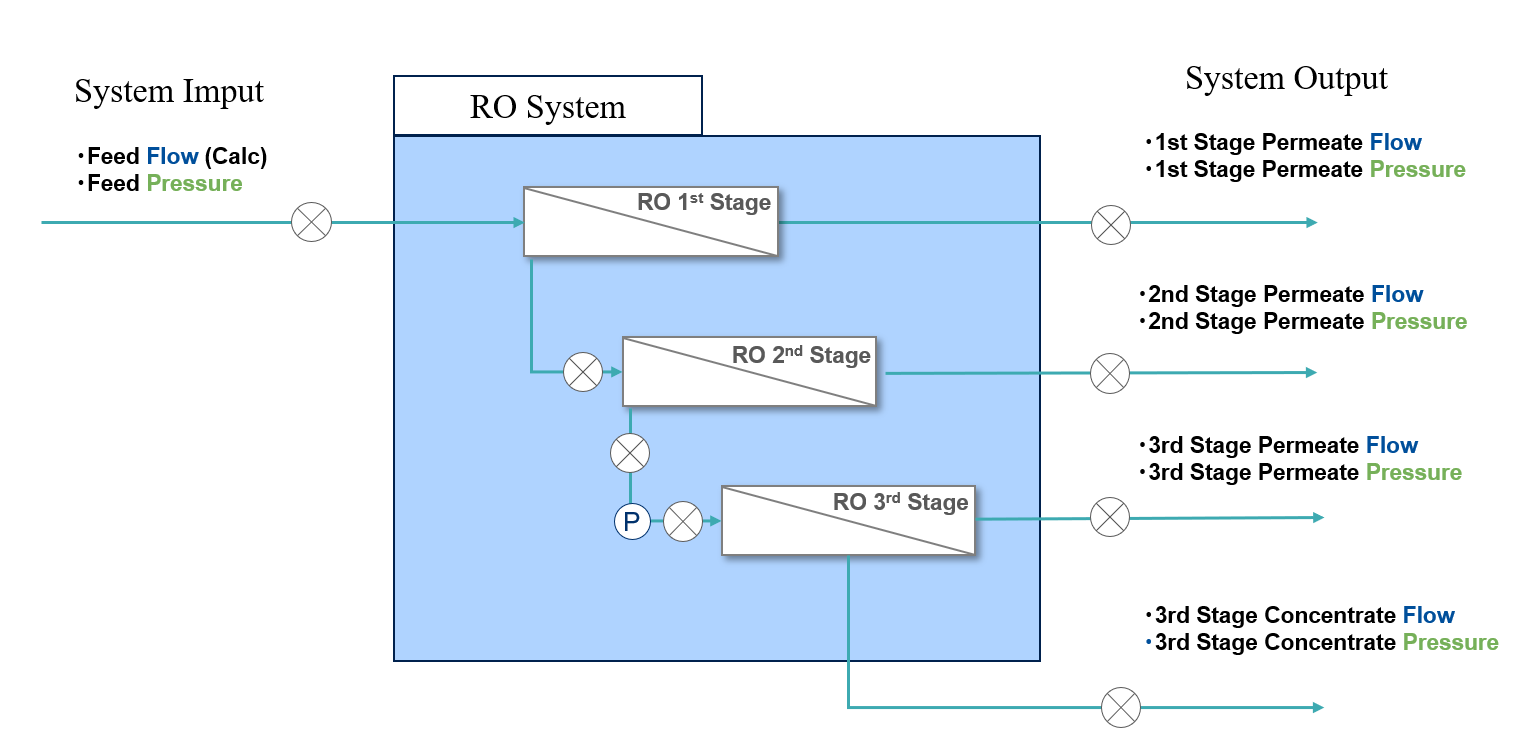
RO膜にFoulingを発生させる要因として大きく分類すると”物理的な負荷”と”化学的な負荷”2種類に分けられる。具体例としてRO膜への負荷を考える。前者に関しては、処理水の成分が一定とすれば、流量や圧力が大きいときに負荷が大きくなり、後者に関しては、流量等が一定とすれば、処理水がより高濃度のとき膜への負荷が大きいと考えられる。本レポートにおいては”物理的な負荷”にフォーカスする。

* ”物理的な負荷”のダイナミクス

RO膜への物理的な負荷を考えるために、流体のエネルギー保存則であるベルヌーイの定理を活用する。ベルヌーイの定理は管内層流条件のとき、各変数をとすれば式(2.1)の通りである。

図１のようにRO膜システムの入出力を定義する。RO膜システムの流体的抵抗がないと仮定すれば、式（1）は以下のように変形できる。ただし、配管の高さは変わらないものとする。

しかしながら、実際にはRO膜による抵抗があるためシステム出力側のエネルギーが減少し、式(2.2)の等号は成立しない。そのため、RO膜によるエネルギー減少分をとすれば、RO膜前後のエネルギー保存則は式(2.3)のように変形できる。



**Figure 2.1:** システムの入出力定義

ここで、とRO膜つまりの関係性を考える。RO膜につまりが発生すれば、RO膜による抵抗が大きくなり、流体が膜を通り抜けにくくなる。つまり、エネルギー減少分であるが大きくなる。そのため、はRO膜のつまりを推定するための有力な説明因子になりうると考えられる。次節からをどのように計算するか述べる。

* の推定方法

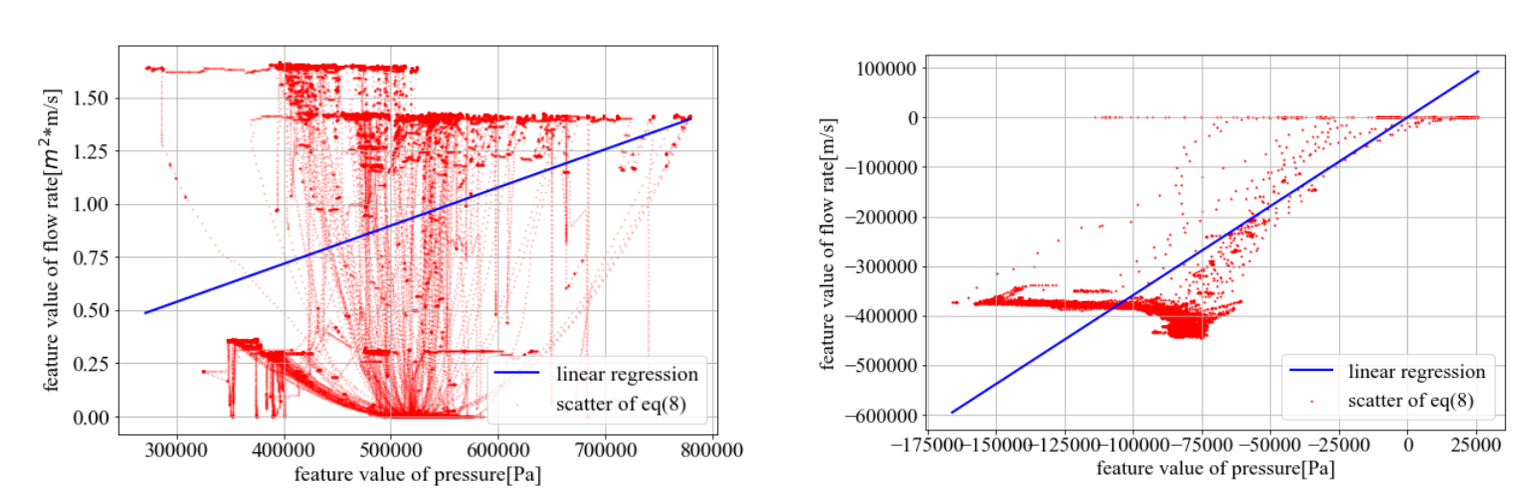
推定方法の概要は以下の３点である。

* + ベルヌーイの定理の各変数に計測値を変換して代入する
  + ベルヌーイの定理に含まれる未計測のパラメータをデータから推定
  + 推定したパラメータを用いてを推定

第一に、Figure2.1記載の入出力の測定値を式(2.3)の形へ変形することを考える。代表例としてFeed側を挙げて説明する。Feed側の圧力と流量の計測値（計算値）の値をそれぞれ,とすれば式(2.3)左辺の項はそれぞれ式(2.4),(2.5)のように求められる。ただし、は単位系の変換係数、はFeed側の配管の断面積である。

同様の変換を式(2.3)右辺に実施すれば、式(2.3)は計測値を用いて式(2.6)のように変形できる。ただし、は i th Stage Permeate Pressure , は i th Stage Concentrate Pressure,は i th Stage Permeate Flow,は i th Stage Concentrate Flowを表す。  
各配管径が未知であるため、配管の断面積がすべて等しいと仮定すれば下記の式になる。

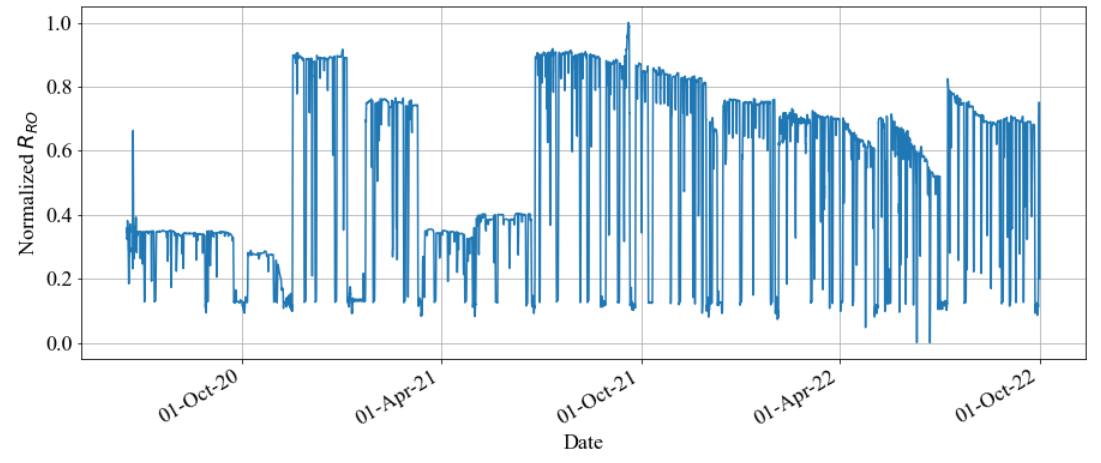
式(2.7)において未知のパラメータはである。を推定するために、一度と仮定して最小二乗法（切片=0の線形回帰）で解いた結果がFigure2.2の通りである。Figure 2.2 (a)がLVMWD, Figure 2.2 (b)がOCWDのデータを用いてパラメータを推定した結果である。ともに縦軸に式(2.8)右辺の流量側の特徴量を、横軸に式(2.8)左辺である圧力の特徴量をとっている。ともに密度、断面積ともにと正値かつ回帰係数が正となっていることから、物理的な関係を満たすパラメータ推定が達成されていることが分かる。また、LVMWDの赤点の分布よりOCWDの赤点の分布の方が相関が高いことが見て取れる。これはOCWDにおいては配管径が既知のためすべての可配管断面積と同一と仮定していないためである。そのため、LVMWDにおいても配管径が分かれば推定の精度が向上することが予想される。



**(a) LVMWD Data (b) OCWD Data**

**Figure 2.2:** 未知パラメータ推定結果

最後に、推定した未知パラメータを式(2.7)に代入し、を計算した上でつまりの説明変数となるように0-1正規化した結果がFigure2.3である。Figure2.3の縦軸は正規化した、横軸は時間である。各入力に従って流量的なエネルギー損失が時間的に変化していることが見て取れる。



**Figure 2.3:** 推定結果

#### RO Optimization Simulation Model (Kumagai)

* Summary

We formulated RO optimization problem for LVMWD and constructed a RO optimization simulation model implemented by the water quality prediction or fouling model. We will carry out the optimization simulation and calculate an improvement effect compared actual operation in the past period in the final report.

* RO Optimization Problem for LVMWD

We formulated the RO optimization problem for LVMWD. Basic optimization story for LVMWD is to decrease the chemical dosing cost (e.g., chloramines) while satisfying water quality standards or monitoring a RO membrane fouling. The problem is formulated as scheduling problem same as OCWD.

* Flow Chart

Figure 2.1 shows the flow chart for LVMWD based on the RO system configuration. The RO system configuration of LVMWD can be found in Additional Materials A2. This flow chart includes the prediction model in the previous section.



**Figure 2.1: Flow Chart for RO Optimization Simulation (LVMWD), TBD**

* Mathematical Optimization Problem

Optimization variable list is as shown in Table 2.1. “Optimization Variable”, “Fixed Parameter”, and “Intermediate Variable” have the same meaning as in the case of OCWD. In LVMWD optimization model, “Optimization Variable” is “UF Filtrate Total Chlorine” and the number of the optimization variables .

**Table 2.1: Optimization Variable List (LVMWD)**

|  |  |  |  |
| --- | --- | --- | --- |
| ID No. | Description | Units | Opt. Variable / Fixed Parameter |
| ID0000 | UF Filtrate Total Chlorine | tons | Optimization Variable |
| ID0001 | Feed EC |  | Fixed Parameter |
| ID0002 | Feed TOC | mg/L | Fixed Parameter |
| ID0003 | Feed Temperature |  | Fixed Parameter |
| ID0004 | Feed pH | - | Fixed Parameter |
| ID0100 | Stage 1 Feed Pressure | Psi | Fixed Parameter |
| ID0101 | Stage 1 Feed Flow Rate | gpm | Fixed Parameter |
| ID0200 | Stage 2 Feed Flow Rate | gpm | Fixed Parameter |
| ID0300 | Stage 3 Feed Pressure | psi | Fixed Parameter |
| ID0301 | Stage 3 Feed Flow Rate | gpm | Fixed Parameter |
| ID1000 | Stage 1 Permeate EC |  | Intermediate Variable |
| ID2000 | Stage 2 Feed EC |  | Intermediate Variable |
| ID2001 | Stage 2 Permeate EC |  | Intermediate Variable |
| ID3000 | Stage 3 Feed EC |  | Intermediate Variable |
| ID3001 | Stage 3 Permeate EC |  | Intermediate Variable |
| ID4000 | Permeate TOC | mg/L | Intermediate Variable |

The objective function is formulated as follows:

Here, denotes UF filtrate total chlorine at time , denote cost coefficients, and denotes the length (step size) of optimization period, respectively.

Constraint conditions are as following:

Here, Equations (2.2), (2.3), (2.4), (2.5), (2.6) denote lower and upper limit, Equation (2.7) denotes fluctuation range limit, Equations (2.8), (2.9), (2.10), (2.11) denote water quality standards limit based on LRV, respectively.

Intermediate variables (Stage1 permeate EC), (Stage2 permeate EC), (Stage3 permeate EC), and (permeate TOC) are calculated by water quality prediction model as follows:

Here, each prediction model is shown in details as previous section. is Logarithmic Reduction Value (LRV) of water quality data (e.g., electric conductivity and TOC) same as Equation (1.18).

The scale of the above problem can be roughly estimated such as number of total optimization variables and number of total constraints . For example, if the optimization period is 1week of 30 minutes time interval data, we can estimate steps, , and .

* Optimization Algorithm

We constructed and used the optimization algorithm by combination of SHADE and Feasibility Rule for solving the above constrained optimization problem same as the case of OCWD.

### Desktop Evaluation based on WBMWD (Kawata-san)

#### Target Process for Data Analysis and Measurements Points

#### Trend Charts

#### Title 22 Requirements

#### Predicting Turbidity from Title 22

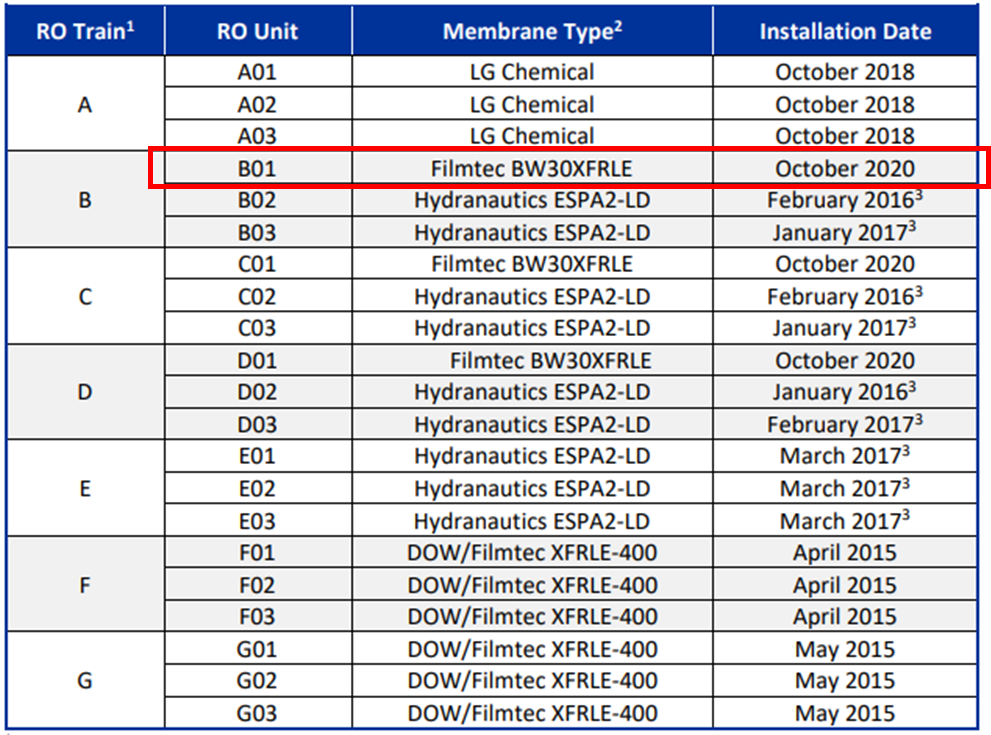
### Future Tasks (Kawata-san)

### Additional Materials

#### A1) RO System Configuration (OCWD)

The data set provided to this team is for the 100 mgd system. We note that the capacity has recently been upgraded to 130 mgd. The full-scale RO membrane system in OCWD, for the longer running 100 mgd of capacity, consists of 21 RO units (3 RO units x 7 RO trains) and each unit has 5 MGD capacity. As OCWD reported in Figure A.1**[[4]](#footnote-4)**, these 21 RO systems are being operated by different types of RO membranes. Current analysis focuses upon RO UNIT B01. As shown in the figure below, the membrane type of RO UNIT B01 is Filmtec BW30XFRLE installed in October 2020.

RO UNIT B01 is a 3-stage configuration, with flow rate, pressure, conductivity, and differential pressure measured as shown in the Figure A.2. 50 categories of data were provided including feed & permeate flow rate, feed & permeate water qualities (such as conductivity, TOC, turbidity), feed & permeate pressure, Xact, and chemical dosage. Tag name list in OCWD is shown in Table A.1.



**Figure A.1: RO System Membranes (OCWD)**



**(a): The Whole RO System**



**(b): RO Unit B01 System**

**Figure A.2: RO System Configuration (OCWD)**

**Table A.1: Tag Name List (OCWD)**

|  |  |
| --- | --- |
| Tag Name | Variable Name |
| RO\_B01 FW Press | RO B01 Feed Pressure |
| RO\_Feed Cond | RO Feed EC |
| RO\_01\_FirstStageFeedFlow | RO Feed Flow Rate |
| RO\_Feed Temperature | RO Feed Temperature |
| RO\_Feed Ph | RO Feed pH |
| Sulfric Acid Usage | Sulfuric Acid Usage |
| Threshold Inhibitor Usage | Threshold Inhibitor Usage |
| RO\_B01 FW B2 Press | RO B01 Feed Pressure |
| RO\_01\_SecondStageFeedConductivity | RO B01 Stage 2 Feed EC |
| RO\_01\_SecondStageFeedFlow | RO B01 Stage 2 Feed Flow Rate |
| RO\_01\_ThirdStageFeedConductivity | RO B01 Stage 3 Feed EC |
| RO\_01\_ThirdStageFeedFlow | RO B01 Stage 3 Feed Flow Rate |
| RO\_Feed TOC | RO Feed TOC |
| RO\_B01 Blank 1 Perm Cond | RO B01 Stage 1 Permeate EC |
| RO\_B01 Blank 2 Perm Cond | RO B01 Stage 2 Permeate EC |
| RO\_B01 Blank 3 Perm Cond | RO B01 Stage 3 Permeate EC |
| RO\_Permeat TOC | RO Permeate TOC |

#### A2) RO System Configuration (LVMWD)

The RO membrane system in LVMWD is a pilot scale system. The RO system consists of 3-stage configuration, utilized the Toray TMG10D membrane, and is shown in the Figure A.3. Tag name list in LVMWD is shown in Table A.2.



**Figure A.3: RO System Configuration and Measurement Points (LVMWD)**

**Table A.2: Tag Name List (LVMWD)**

|  |  |
| --- | --- |
| Tag Name | Variable Name |
| PT\_41095 | RO Stage 1 Feed Pressure |
| AIT\_40006 | RO Feed EC |
| RO\_01\_FirstStageFeedFlow | RO Feed Flow |
| TIT\_40005 | RO Feed Temperature |
| AIT\_40008 | RO Feed pH |
| AI\_31094 | UF Filtrate Total Chlorine |
| RO\_01\_SecondStageFeedConductivity | RO Stage 2 Feed EC |
| RO\_01\_SecondStageFeedFlow | RO Stage 2 Feed Flow |
| PT\_41347 | RO Stage 3 Feed Pressure |
| RO\_01\_ThirdStageFeedConductivity | RO Stage 3 Feed EC |
| RO\_01\_ThirdStageFeedFlow | RO Stage 3 Feed Flow |
| AIT\_40010 | RO Feed TOC |
| AIT\_41092 | RO Stage 1 Permeate EC |
| AIT\_41292 | RO Stage 2 Permeate EC |
| AIT\_41392 | RO Stage 3 Permeate EC |
| AIT\_41810 | RO Combined Permeate TOC |

#### A3) Temperature Correction Factor

Used mathematical equations for Temperature Correction Factor (TCF) or normalization in this report are as follows:

* Water Viscosity (Temperature dependency; Andrade equation):

Here, is the water temperature and is a specific coefficient. in LVMWD data, in OCWD data.

* Water Fluidity (Temperature dependency):

Here, is the water viscosity.

* Temperature Correction Factor (TCF):

Here, is the water temperature and is the standard temperature. Eq.(A3) is used in LVMWD data and Eq.(A4) is used in OCWD data.

* Temperature Correction Factor for Flow:

Here, is the water viscosity, is the water temperature and is the standard temperature.

* Temperature Correction Factor for Salt:

Here, is the water viscosity, is the water temperature and is the standard temperature.

* Net Driving Pressure (NDP)

Here, is the RO transmembrane pressure, is the differential RO module pressure, is the osmotic pressure, is the number of RO stages in the system, is the RO feed pressure, is the RO permeate pressure of entire train, and is the RO brine pressure. Unit of each pressure is bar in LVMWD data and psi in OCWD data. in OCWD data.

* Average Osmotic Pressure Differential:

Here, is the average of osmotic pressure between feed and brine flow, is the permeate osmotic pressure, is the log mean factor between feed and brine for TDS, is the permeate TDS, and are constant coefficients. Eq.(A12) is used in LVMWD.

* Log Mean Factor considering concentration polarization on RO surface: ,

Here, is the log mean factor between feed and brine for TDS, is the log mean factor for brine concentration, is the feed TDS, is the brine TDS, and is the recovery rate. Unit of each TDS is ppm in LVMWD data. Eq.(A15) is used in LVMWD data and Eq.(A16) is used in OCWD data.

* Recovery Rate:

Here, is the RO feed flow rate and is the RO permeate flow rate.

* Average Feed and Brine Concentration:

Here, is the RO feed concentration and is the log mean factor for brine concentration.

* Average Feed and Brine Flow Rate:

Here, is the RO feed flow rate, is the RO permeate flow rate, and is the RO brine flow rate. Unit of each flow rate is gpm (gallon per minute) in LVMWD data and OCWD data. Eq.(A19) is used in LVMWD data and Eq.(A20) is used in OCWD data.

* Calculated TDS:

Here, is the RO feed TDS, is the RO permeate TDS, is the RO brine TDS, is the RO feed flow rate, is the RO permeate flow rate, and is the RO brine flow rate. are conductivities. In Eqs.(A22) and (A23), are constant coefficients and given by as follows:

where is the feed conductivity.

1. R. Tanabe and A. Fukunaga: “Success-History Based Parameter Adaptation for Differential Evolution,” Proceedings of IEEE Congress on Evolutionary Computation 2013, pp.71–78 (2013) [↑](#footnote-ref-1)
2. K. Deb: “An Efficient Constraint Handling Method for Genetic Algorithms”, Computer Methods in Applied Mechanics and Engineering, Vol. 186 No. 2-4, pp.311-338 (2000) [↑](#footnote-ref-2)
3. The Black-box Optimization Benchmarking (BBOB) workshop (https://numbbo.github.io/workshops/) [↑](#footnote-ref-3)
4. GWRS 2020 ANNUAL REPORT (https://www.ocwd.com/wp-content/uploads/2020-gwrs-annual-report-appendices-1.pdf) [↑](#footnote-ref-4)