Task: Subtask 5.17.1.5 – 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.

### Budget Spent (YCA)

### Summary (Kawata)

# Desktop Analysis based on OCWD

* 1. **RO Optimization Overview**

Basic optimization story is to decrease the chemical dosing cost (e.g., sulfuric acid or inhibitor) while satisfying water quality standards or monitoring a RO membrane scaling.

RO optimization is to improve the operation in some period. This type of optimization problem is the scheduling problem, which is formulated and solved according to procedure as shown in Figure 1.1.1. The procedure consists of drawing flow chart, formulation optimization problem, and calculation operational schedule. Therefore, first and second are required to construct a simulation model.

* + 1. **Flow Chart**

Figure 1.1.2 [Figure1.1.2 TBU] shows the flow chart for RO system configuration. Figure 1.1.2 (a) is the flow chart limited to the part related to permeate TOC prediction and Figure 1.1.2 (b) is the flow chart limited to the part related to permeate EC prediction. In Figure 1.1.2, arrows connecting objects are assigned variables. The classification of each variable (e.g., optimization variable or fixed variable) is described below. The storage object means exogenous variable, the hexagon object means mathematical model, and the square object means predictive variable (e.g., water quality data). The arrow toward the object is input data and the arrow out from the object is output data. The color of arrows is different for each object to which it is connected. Especially, mathematical models have both input and output data. Output data from the hexagon object toward the square object is calculated by the prediction model in the later section.

The RO system configuration and tag name list of OCWD is provided as “Additional Materials A1” at the appendix of this report.



**Figure 1.1.1: Procedure for Solving Scheduling Problem**

* + 1. **Formulation Optimization Problem**

The optimization problem is formulated as minimizing objective function satisfying constraints. In the optimization problem, denotes the optimization variable and denotes the objective function (cost function). A detail of the problem is provided at the next section.

Table 1.1.1 shows the optimization variable list in the OCWD optimization model. “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. In OCWD optimization model, the optimization variable is sulfuric acid and threshold inhibitor usage, and the intermediate variables are the permeate electric conductivity (EC) at stage 1, stage 2, and stage 3, combined permeate EC, and combined permeate TOC calculated by water quality prediction model.



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



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

**Figure 1.1.2: Flow Chart for RO Optimization Simulation (OCWD)**

**Table 1.1.1: Optimization Variable List (OCWD)**

|  |  |  |  |
| --- | --- | --- | --- |
| ID No. | Description | Unit | Opt. Variable / Fixed Parameter |
| ID0000 | Sulfuric Acid Usage | ton/hour | Optimization Variable |
| ID0001 | Threshold Inhibitor Usage | ton/hour | 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 |
| ID0102 | Stage 1 Permeate Flow Rate | gpm | Fixed Parameter |
| ID0200 | Stage 2 Feed Pressure | psi | Fixed Parameter |
| ID0201 | Stage 2 Feed Flow Rate | gpm | Fixed Parameter |
| ID0202 | Stage 2 Permeate Flow Rate | gpm | Fixed Parameter |
| ID0300 | Stage 3 Feed Flow Rate | gpm | Fixed Parameter |
| ID0301 | Stage 3 Permeate 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 | Combined Permeate EC |  | Intermediate Variable |
| ID4001 | Combined Permeate TOC | ppm | Intermediate Variable |

* 1. **Water Quality Prediction** (Imoto)

水質予測の目的を説明

* + 1. **Outliers Removal**
* 外れ値の削除方法を説明
* 外れ値前後のデータを比較（代表的な変数2～3個ピックアップ）
  + 1. **Prediction Method**
* 重回帰分析とランダムフォレストで水質を予測
* 重回帰分析だけでなくランダムフォレストでも予測した理由を説明
* 逐次的に学習期間を変えて予測、評価期間などの計算条件を説明
* 重回帰分析のアルゴリズムの簡単な説明があるとベター
* ランダムフォレストのアルゴリズムの簡単な説明があるとベター
  + 1. **Prediction Results**
       1. **Prediction by Multiple Linear Regression Model**
* 重回帰分析で水質を予測した結果
  + - 1. **Prediction by Random Forest**
* ランダムフォレストで水質を予測した結果
  + - 1. **Comparison of the Model Accuracy**
* 重回帰分析とランダムフォレストのモデル精度を比較
* 考察
  1. **RO Membrane Scaling Estimation Model Using Xact Data (Soya)**
* スケーリングを推定する目的を説明
* Xactを簡単に説明
  + 1. **Method to Estimate Scaling**
* 濃度分極とSaturation Indexを用いたスケーリング推定計算法の概要説明
* 計算式を紹介（SiとCaを計算）
  + 1. **Results**
       1. **Conditions**
* 計算条件（Silica／Ca計算式に入れるパラメータ）
  + - 1. **Results for Concentration Polarization and Saturation Index**
* 計算結果
* SilicaとCalciumの濃度と温度の関係
* 濃度分極とSaturation Indexの計算からanti-scalantの注入量との関係を示せないか検討中
  1. **RO Optimization Simulation**

We developed a RO optimization simulator and calculated the improvement effect of saving cost. But there is the little improvement in the case of not using the RO membrane scaling model.

* + 1. **Optimization Problem for RO Operational Scheduling**

RO optimization story is to decrease the chemical dosing cost (e.g., sulfuric acid or inhibitor) while satisfying water quality standards or monitoring a RO membrane scaling. But this report uses only the water quality prediction model because the RO scaling model needs more validation in some periods and conditions.

RO optimization problem is formulated as the scheduling problem. The operational schedule is denoted by (). denotes the -th variable schedule at time , denotes the -th variable schedule in all periods, denotes the length (step size) of the optimization period, and denotes the number of variables, respectively.

The objective function is sum of the threshold inhibitor usage and sulfuric acid usage during the optimization period. The constraint conditions are given by lower and upper limit, and water quality standards limit based on Logarithmic Reduction Value (LRV). Table 1.4.1 shows only the major constraints. LRV of water quality data (e.g., electric conductivity and TOC) as follows:

Here, denotes RO feed quality data and denotes RO combined permeate quality data . The constraint’s parameters shown in Table 1.4.1 are basic values, modified somewhat in some cases.

The mathematical prediction model is based on the Multiple Linear Regression Model (MLR) because the amount of calculation time using the Random Forest is about 300 times longer than MLR in the optimization.

**Table 1.4.1: Major Constraints (basic upper/lower limit and LRV)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Description** | **Units** | **Lower** | **Upper** |
| Stage 1 Permeate EC |  | 10 | 95 |
| Stage 2 Permeate EC |  | 20 | 200 |
| Stage 3 Permeate EC |  | 60 | 350 |
| Permeate Combined EC |  | 60 | 350 |
| Permeate Combined TOC | ppm | 0 | 0.15 |
| Permeate Combined EC LRV | - | 1.5 | 2.3 |
| Permeate Combined TOC LRV | - | 1.5 | 2.5 |

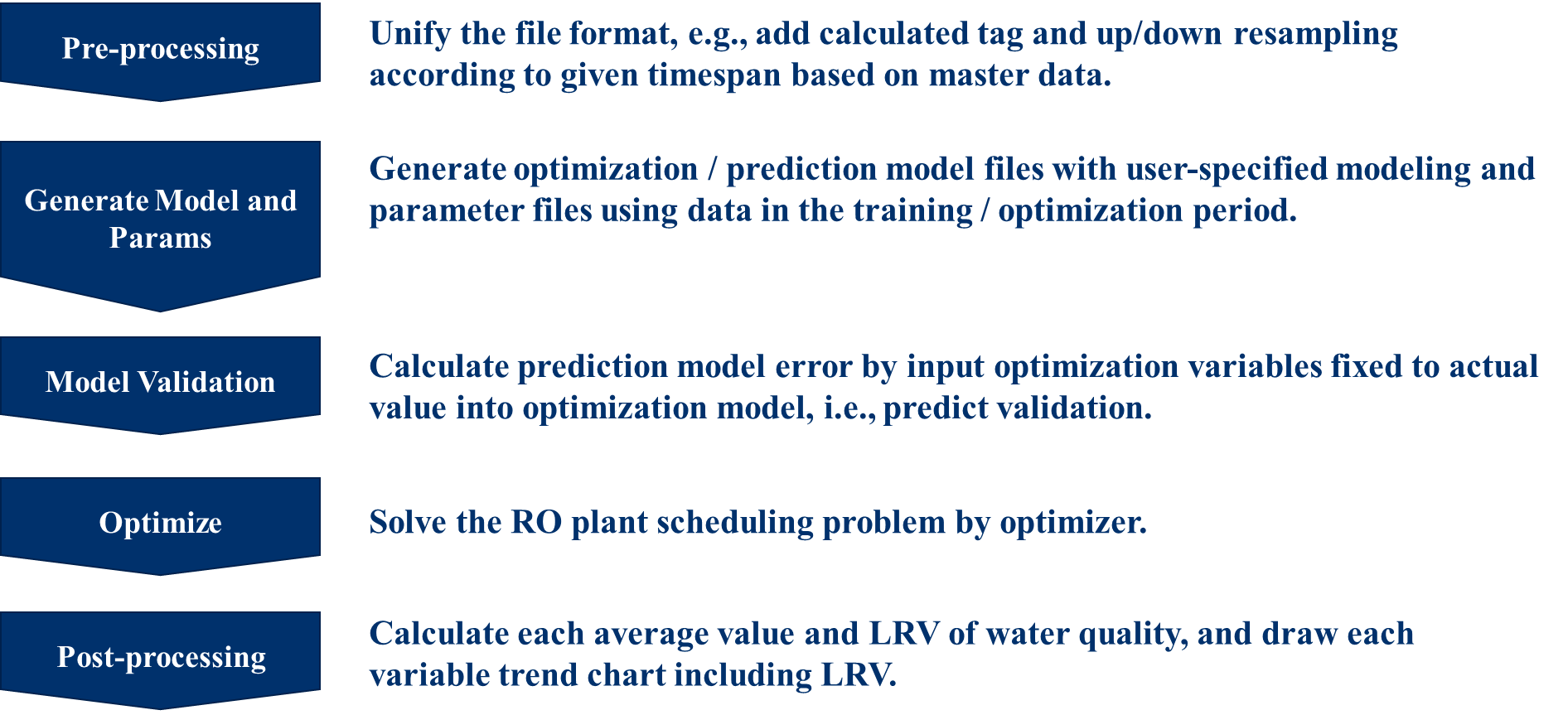
* + 1. **Optimization Algorithm**

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.

* + 1. **RO Optimization Simulator**

We implemented the RO optimization simulator by writing python source code from scratch. This simulator solves the above optimization problem by using the above algorithm and calculate the optimized RO operational schedule.

Figure 1.4.1 shows the functional configuration and flow chart of RO optimization simulator. The procedure consists of 5 steps: step 1 is data pre-processing, step 2 is model construction, step 3 is model validation, step 4 is optimization, and step 5 is data post-processing. We will implement the RO simulator in the pilot-scall based on this.



**Figure 1.4.1: Functional Configuration and Flow Chart of RO Optimization Simulator**

* + 1. **Simulation Results**

This part shows the optimization simulation results. The method trains one week's data and optimizes the following week. The training and optimization periods are shifted sequentially and the entire of the consolidated optimization period is 5 weeks in this report.

The prediction results by the RO simulator are shown in Figures 1.4.2 and 1.4.3 [Figures 1.4.2 and 1.4.3 TBU]. Figure 1.4.2 is prediction trend of permeate combined TOC and LRV of that, and Figure 1.4.3 is prediction trend of permeate combined conductivity (EC) and LRV of that. The results are …

The optimization results are shown in Figures 1.4.4, 1.4.5, and 1.4.6 [Figures 1.4.4, 1.4.5, and 1.4.6 TBU]. Figure 1.4.4 is optimized trend of chemical dosage, i.e., sulfuric acid and inhibitor, Figure 1.4.5 is optimized trend of permeate combined TOC and LRV of that, and Figure 1.4.6 is optimized trend of permeate EC and LRV of that.

**グラフ

自動的に生成された説明**

1. **Permeate Combined TOC**

**グラフィカル ユーザー インターフェイス, グラフ, テーブル

中程度の精度で自動的に生成された説明**

1. **LRV of Permeate Combined TOC**

**Figure 1.4.2: Prediction Results of TOC in RO simulator (OCWD)**

**グラフ

自動的に生成された説明**

1. **Permeate Combined EC**

**グラフ

自動的に生成された説明**

1. **LRV of Permeate Combined EC**

**Figure 1.4.3: Prediction Results of EC in RO simulator (OCWD)**

**グラフ

自動的に生成された説明**

1. **Sulfuric Acid Usage**

**グラフ

自動的に生成された説明**

1. **Inhibitor Usage**

**Figure 1.4.4: Optimization Results of Chemical Dosage in RO simulator (OCWD)**

**グラフ

自動的に生成された説明**

1. **Permeate Combined TOC**

**グラフ

自動的に生成された説明**

1. **LRV of Permeate Combined TOC**

**Figure 1.4.5: Optimization Results of TOC in RO simulator (OCWD)**

**グラフ

自動的に生成された説明**

1. **Permeate Combined EC**

**グラフ

自動的に生成された説明**

1. **LRV of Permeate Combined EC**

**Figure 1.4.6: Optimization Results of EC in RO simulator (OCWD)**

The optimization effect shows Table 1.4.2 [Table 1.4.2 TBU].

**Table 1.4.2: Optimization Effect**

|  |  |  |  |
| --- | --- | --- | --- |
| **Description** | **Units** | **Lower** | **Upper** |
| Stage 1 Permeate EC |  | 10 | 95 |
| Stage 2 Permeate EC |  | 20 | 200 |
| Stage 3 Permeate EC |  | 60 | 350 |
| Permeate Combined EC |  | 60 | 350 |
| Permeate Combined TOC | ppm | 0 | 0.15 |
| Permeate Combined EC LRV | - | 1.5 | 2.3 |
| Permeate Combined TOC LRV | - | 1.5 | 2.5 |

* 1. **Future Tasks (Kamada)**
* Optimization
  + We will estimate the improvement cost by introducing the scaling model to the optimization model and considering a potential of RO membrane scaling.
  + We need the way to reduce the optimization calculation time in the case of using the Random Forest.

1. **Desktop Analysis based on LVMWD** **(Kamada)**

* LVMWDのRO膜システムの概要説明
* モデル作成のポイントを説明
  1. **RO Optimization Overview** (Kumagai)
     1. **Flowchart**
     2. **Formulation Optimization Problem**
  2. **Water Quality Prediction** **(Imoto)**
     1. **Outliers Removal**
* 外れ値の削除方法を説明（OCWDと同じであれば、その旨を記述）
* 外れ値前後のデータを比較（代表的な変数2～3個ピックアップ）
  + 1. **Prediction Method**
* 重回帰分析とランダムフォレストで水質を予測（OCWDと同じであれば、その旨を記述）
* 重回帰分析だけでなくランダムフォレストでも予測した理由を説明（OCWDと同じであれば、その旨を記述）
* 逐次的に学習期間を変えて予測、評価期間、変数などの計算条件を説明
  + 1. **Prediction Results**
       1. **Prediction by Multiple Linear Regression Model (Imoto)**
* 重回帰分析で水質を予測した結果
  + - 1. **Prediction by Random Forest (Imoto)**
* ランダムフォレストで水質を予測した結果
  + - 1. **Comparison of the Model Accuracy (Imoto)**
* 重回帰分析とランダムフォレストのモデル精度を比較
* 考察
  1. **RO Membrane Fouling Estimation Model (Imoto)**
* Foulingを推定する目的を説明
  + 1. **Relationships between Feed Pressure and Total Chlorine**
* 1stと3rdのFeed Pressureの比率データトレンドから、CIPやRO膜詰まりと関連がある仮説を言及
  + 1. **Method to Optimize Total Chlorine based on Fouling Estimation**
* Pressure比からTotal Chlorineの量を最適化する考えを書く
* Total ChlorineとFeed Pressure比の散布図を描いて、その可能性を考察
* ただし、受領データには、データが悪いのか、方法が悪いのかを区別する分量や条件が不足しているため、様々な条件の計測データで追加検証する必要がある。そのため、今回の最適化シミュレーションでは導入していない。
  1. **RO Optimization Simulation (Kumagai)**
* 最適化のコンセプトを紹介（Water Prediction ModelとMembrane Fouling Modelを組み込む）
* Membrane Fouling Modelはまだ検討中なので、Water Prediction Modelを組み込んだ結果を紹介すると前振り
  + 1. **Optimization Problem for RO Operational Scheduling**
    2. **Optimization Algorithm**
    3. **Simulation Results**
* 最適化計算の結果
* コスト削減率（良さそうな結果があれば）
* 各変数間の関係（薬品量と水質の関係など）
  1. **Future Tasks (Kamada)**
* 各パートの課題を書く
* 解決策もあるとベター
* これを踏まえて、Action Planを立てるとベスト
* 各プラントで共通性が高いなら、4章に書くべきでは？（Scale-Up Selectionなので）

1. **Desktop Analysis based on WBMWD (Kawata)**

* WBMWDのプロセス概要説明
* モデル作成のポイントを説明
* Hourly Average ModelとDaily Average Modelを作成
  1. **Model Creation by DDMO**
* DDMOの説明
  1. **Hourly Average Model**
* Hourly Average Modelの概要説明
  + 1. **Calculation Status**
* 計算条件を記述
* 変数、学習期間、評価期間を説明
  + 1. **Model Accuracy**
* モデル精度確認の説明
* 今回はRMSEで評価
  + 1. **Optimization Calculation**
* 最適化計算結果を記述
* 各変数間の動き、関係性について記述
  1. **Daily Average Model**
* Daily Average Modelの概要説明
* 手分析値を採用した旨を記述
  + 1. **Calculation Status**
* 計算条件を記述
* 変数、学習期間、評価期間を説明
  + 1. **Model Accuracy**
* モデル精度確認の説明
* 今回はRMSEで評価
  + 1. **Optimization Calculation**
* 最適化計算結果を記述
* 各変数間の動き、関係性について記述
  1. **Suggestions**

流入濁度を測定することで水質の変化に応じて薬品注入量を調整できるようになるため、薬品注入量を削減できる可能性がある旨を記述

1. **Scale-Up Plan**

* Scale-Upする上での共通課題を書く
* それらの対策案を書く
* Task 2で組み込んだアクションプランを書く

# Appendix for Additional Materials

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

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)