Developing a Statistical Model to Improve Drinking Water Quality for Water Distribution System by Minimizing Heavy Metal Releases

Wei Peng and Rene V. Mayorga *

Faculty of Engineering, University of Regina, Regina, SK S4S 0A2, Canada; Wei.Peng@uregina.ca
* Correspondence: Rene.Mayorga@uregina.ca; Tel.: + 1-(306)-585-4397; Fax: + 1-(306)-585-4855

Abstract: This paper proposes a novel statistical approach for blending source waters in a public water distribution system to improve water quality (WQ) by minimizing the release of heavy metals (HMR). Normally, introducing a new source changes the original balanced environment and causes adverse effects on the WQ in a water distribution system. One harmful consequence of blending source water is the release of heavy metals, including lead, copper and iron. Most HMR studies focus on the forecasting of unfavorable effects using precise and complicated nonlinear equations. This paper uses a statistical multiple objectives optimization, namely Multiple Source Waters Blending Optimization (MSWBO), to find optimal blending ratios of source waters for minimizing three HMRs in a water supply system. In this paper, three response surface equations are applied to describe the reaction kinetics of HMR, and three dual response surface equations are used to track the standard deviations of the three response surface equations. A weighted sum method is performed for the multi-objective optimization problem to minimize three HMRs simultaneously. Finally, the experimental data of a pilot distribution system is used in the proposed statistical approach to demonstrate the model's applicability, computational efficiency, and robustness.

Keywords: water quality (WQ); blending; release of heavy metals (HMR); dual response surface optimization (DRSO); multiple source waters blending optimization (MSWBO)

1. Introduction

Mixing water from a new source may alter the original balanced environment in water networks. As a consequence, pipe corrosion, decreasing disinfectant residual and microbiological growth will deteriorate WQ. Optimizing the blending ratio of water coming from different sources may minimize this water deterioration [1–4]. However, an optimal blending ratio usually cannot be realized because an operational blending ratio is dependent on the water demand, capacities of water plates, and locations of water plates and the mixing point of source waters. Furthermore, the concentrations of heavy metals are unassignable by water plates. They are impacted by the chemical parameter concentrations (such as chlorine concentration) and physical parameter levels (such as turbidity), also influenced by the original balanced environment in water networks. It is a challenge to minimize the release of heavy metals (HMR) in a water distribution system.

Imran et al. [1,5] described a method that optimize the blending of multiple source waters for heavy metal corrosion abatement and monochloramine residual control in distribution systems. They built up a pilot water distribution system to simulate a full-scale operational water distribution system. They fitted the nonlinear empirical WQ models for copper, lead, apparent color, and monochloramine dissipation. In addition, then based on the regulations, the optimal water blending ratio ranges were found. However, this optimization may need to be further developed,

to provide some quantitative indicators to water utilities for improving the WQ in their water distribution systems. This paper proposes a new statistical optimization method for offering such quantitative indicators.

The optimization of this study is nonlinear because of the nonlinearity of the WQ. Conventional nonlinear optimization methodologies, such as gradient-based algorithms [6–8] and genetic algorithms [9], have some disadvantages that may not be suitable for use in the field of WQ. E.g., a gradient-based algorithm may have a large development cost and the linearization can be time-consuming; the genetic algorithms are time-consuming when be used in a large-scale nonlinear problem. Here, in this paper a second-order statistical model, namely the dual response surface optimization (DRSO), is introduced to conduct this nonlinear optimization.

The DRSO is based on the Response Surface Method (RSM). The RSM is a statistical technique used in empirical study. It approximates the true response surface and estimates the parameters; it works well for solving real response surface problems [10,11]. As a consequence, it searches for an optimal set of input variables to optimize the response by using a set of designed experiments. In the past few years, several robust DRSO techniques have been developed [12,13]. Yanikoglu et al. (2016) introduced Taguchi's Robust Parameter Design approach into DRSO to develop a method that uses only experimental data, and it can yield a solution that is robust against ambiguity in the probability of inputs [12].

The optimization is also required to run rapidly because it faces the dynamics of WQ, water demand, and source water capacity of water plates [14,15]. RSM programming can be quickly solved by any commercial solver because it contains a quadratic model rather than a high-order nonlinear model. All observations in RSM are typically assumed to have equal variation that may not be practically valid [16]. Vining and Myers [17] then proposed a DRSO model that includes two quadratic empirical models, one for the mean and another for the standard deviation, and then optimized one of the responses subjected to an appropriate constraint given by the other. Because the second empirical model considers the standard deviation, a DRSO approach can avoid misleading optimum results and produce robust results.

Harmful chemical reactions (especially for HMR) and microbiological growth are required to be controlled in water distribution systems. According to the previous studies, the biomass accumulation is influenced to a greater extent by the nature of the supporting material (such as unlined ductile iron) than by the secondary parameters of WQ [18]. Meanwhile, Imran et al. [1,5] confirmed that loss of disinfectant residual is affected to a greater extent by the delivery distance and retention time rather than the WQ. This paper only focuses on the relationships between the HMR and the chemical and physical parameters of WQ in the distribution system. Three heavy metals are iron, lead, and copper.

This paper proposes a Multiple Source Waters blending Optimization (MSWBO) methodology to find optimal blending ratios for multiple source waters according to the WQ parameters of these source waters (shown in Table 1) and a variety of scenario designs. The objectives of the study are to reduce three HMR of lead, copper, and iron in water distribution systems, and minimize the operational cost. A weighed sum method is used for the multi-objective function in the MSWBO to minimize HMR simultaneously. The dual response surface models are used to describe the reaction kinetics of three HMR, with tracking the standard deviations of the responses of these three release equations. An application is provided to demonstrate the applicability and advantages of this MSWBO.

Table 1. Water quality parameters of the source waters.

Parameter	GW	SW	DW
Alkalinity (mg/L) as $CaCO_3$	225	50	50
Calcium (mg/L) as $CaCO_3$	200	50	50
Dissolved oxygen (mg/L)	8	8	8
pH	7.9	8.2	8.3
Silica (mg/L)	14	7	1

Table 1. *Cont.*

Parameter	GW	SW	DW
HRT (day)	5	5	5
Sodium (mg/L)	10	15	30
Chloride (mg/L)	15	10	50
Sulfates (mg/L)	10	180	30
Temperature (°C)	25	25	25

Notes: Sourced from Imran et al. [1]; GW is groundwater, SW is surface water; DW is desalination water. HRT is hydraulic retention time.

2. Methodology

The development of the proposed MSWBO approach consists of three main steps: data acquisition, model fitting and optimization. Data acquisition includes data type identification, data range determination, experiment design, and experimental data recording.

2.1. Data Acquisition

2.1.1. Data Type Identification and Range Determination

Three HMR are fitted into the dual response surface models. For the dual response surface models, three mean responses are the release levels (RLs) of iron, lead, and copper in the mixed water; three standard deviation responses are the standard deviations of the responses of these three release equations; the input variables are the concentrations of the parameters of WQ in the source waters, respectively. These WQ parameters include alkalinity, calcium, dissolved oxygen, pH, silica, sodium chloride, sulfates, hydraulic retention time (HRT), and temperature.

For the MSWBO, the responses of the dual response surface models are treated as "the control variables"; the operational cost of the WQ adjustment of source waters is also treated as "the control variable"; the input variables of the dual response surface models are treated as "the state variables" or "the decision variables", which will be discussed in the next section; because the ratios of groundwater (GW), surface water (SW), and desalination water (DW) in the blending are given by the decision maker, they are treated as the state variables.

The value ranges of the variables related to WQ are limited by government standards such as National Primary Drinking Water Regulations (Primary Standards) and National Secondary Drinking Water Regulations (Secondary Standards) [19,20]. The sum of three ratios is equal to 1.

2.1.2. Experiment Design and Experimental Data Recording

Tampa Bay Water (TBW) operates a water distribution system and was required by the Southwest Florida Water Management District to reduce groundwater withdrawal for minimizing hydrological and ecological impacts to the riverine systems and ensuring that flows remain within the range of natural variability [21]. To complete this requirement, TBW completed the constructions of a surface water plant and a desalination plant [1,5]. Meanwhile, a study [1] was sponsored by TBW and the American Water Works Association (AWWA) Research Foundation to evaluate the WQ impacts from the new source water supplies. As a part of the above-mentioned study, a pilot distribution system (PDS) was constructed in Florida, U.S., to simulate the full-scale operational water distribution system of TBW [1,5]. Three source waters were selected in the PDS: a conventional groundwater (GW), surface water (SW) and desalinated water (DW). Table 1 describes their WQ information. The aged pipes of the PDS were removed from the water distribution system of TBW. The pipes were assembled on the pilot site and allowed to equilibrate with groundwater over a period of five months. After equilibrium was established, various blends were introduced into the PDS. Two years of continuous operation was performed on the PDS which was used to study the relationship between WQ and the distribution system.

2.2. Model Fitting

Box and Wilson [22] first used the Response Surface Method to study the relationship between a response and a set of input variables. Vining and Myers [11] first fitted second-order polynomial models for a mean and standard deviation separately and optimized one of the responses subjected to an appropriate constraint given by the other. Since the second empirical model considers the standard deviation, a MSWBO approach can avoid misleading optimum and produce robust results. A general dual response surface model is formulated as follows:

$$C_\mu = h_0 + S'h + S'HS \tag{1}$$

$$C_\sigma = g_0 + S'g + S'GS \tag{2}$$

where $g_0 = \alpha_0$, $h_0 = \beta_0$, $\boldsymbol{g} = (\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \ldots \boldsymbol{\alpha_k})'$, $\boldsymbol{h} = (\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \ldots \boldsymbol{\beta_k})'$, and

$$\boldsymbol{H} = \frac{1}{2} \begin{bmatrix} 2\beta_{11} & \beta_{12} & \cdots & \beta_{1k} \\ \beta_{21} & 2\beta_{22} & \cdots & \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{k1} & \beta_{k2} & \cdots & 2\beta_{kk} \end{bmatrix} \tag{3}$$

$$\boldsymbol{G} = \frac{1}{2} \begin{bmatrix} 2\alpha_{11} & \alpha_{12} & \cdots & \alpha_{1k} \\ \alpha_{21} & 2\alpha_{22} & \cdots & \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{k1} & \alpha_{k2} & \cdots & 2\alpha_{kk} \end{bmatrix} \tag{4}$$

where h_0, h, g_0, g, H and G are the appropriate scalars, $(k \times 1)$ vectors, and $(k \times k)$ matrices for the estimated coefficients; C_μ and C_σ are the mean and standard deviation; S and S' are $(k \times 1)$ vectors of the input variables and their transpose, respectively.

The proposed study needs to fit three couples of second-order polynomial models: (1) the lead release model and its standard deviation model; (2) the copper release model and its standard deviation model; and (3) the iron release model and its standard deviation model. It should be noticed that these three couples of models use same secondary WQ parameters (alkalinity, sodium, pH, conductivity, etc.) as input variables.

2.3. Dual Response Surface Optimization

The next step is to optimize the above response surface models simultaneously. The objectives represent the desire for minimizing (1) metal RLs; (2) standard deviations of metal RLs; (3) adjustment costs of WQ for the source waters. With the constraints related to the above control variables, a MSWBO model is formulated as follows:

$$\text{Min} \quad f = \sum_{j=1}^{n} \left\{ w_1 (C_\mu)_j^2 + w_2 (C_\sigma)_j^2 + w_3 [x_{ki}(U_{ki} - \hat{U}_{ki})]_j^2 \right\} \tag{5a}$$

s.t.

$$(C_\mu)_j = (h_o + S'h + S'HS)_j \leq (L_\mu)_j \tag{5b}$$

$$(C_\sigma)_j = (g_o + S'g + S'GS)_j \tag{5c}$$

$$S_k = \sum_{i=1}^{m} x_i U_{ki} \tag{5d}$$

$$\underline{U_{ki}} \leq U_{ki} \leq \overline{U_{ki}} \tag{5e}$$

$$0 \leq x_i \leq 1 \tag{5f}$$

$$\sum_{i=1}^{m} x_i = 1 \tag{5g}$$

where j is the index of the response of metal release, $j = 1, 2, \ldots, n$; i is the index of the source water, $i = 1, 2, \ldots, m$; L_μ is the maximum permissible metal RL; S_k is the concentration of the kth input variable in the blend, k is the index of input variables; x_i is the percentage of the source water i in the blend, i is the index of source waters; U_{ki} is the concentration of the kth input variable in the source water i; $\hat{U}_{ki}$ is the original concentration of the kth input variable in the source water i; U_{ki} and $\overline{U_{ki}}$ are the lower bound and upper bound of U_{ki}; w_1, w_2 and w_3 are the weights to reflect different priorities to the corrosion and the cost of WQ adjustment.

U_{ki} is the decision variable. It directly affects the WQ of the blend. Meanwhile, the lower the difference between U_{ki} and $\hat{U}_{ki}$, the low the operation cost of water utilities. In this study, C_μ, C_σ and $(U_{ki} - \hat{U}_{ki})$ are included in the objective function to reflect the trade-off between operation cost and WQ. This can be realized by adjusting values of w_1, w_2 and w_3 as shown in Equation (5a).

3. Application

3.1. Nonlinear Empirical Model

The previous studies formulated the statistical nonlinear corrosion models for copper, lead and color (iron) [1,5,23,24]. The nonlinear release models are listed as follows:

$$\mathrm{Cu} = (\mathrm{T})^{0.72}(\mathrm{Alk})^{0.73}(\mathrm{pH})^{-2.86}(\mathrm{SO_4})^{0.1}(\mathrm{SiO_2})^{-0.22} \tag{6}$$

$$\mathrm{Pb} = 1.027^{(\mathrm{T}-25)}(\mathrm{Alk})^{0.677}(\mathrm{pH})^{-2.726}(\mathrm{Cl})^{1.462}(\mathrm{SO_4})^{-0.228} \tag{7}$$

$$\Delta\mathrm{C} = 10^{-1.321}(\mathrm{T})^{0.813}(\mathrm{Alk})^{-0.912}(\mathrm{Cl})^{0.485}(\mathrm{Na})^{0.561}(\mathrm{SO_4})^{0.118}(\mathrm{DO})^{0.967}(\mathrm{HRT})^{0.836} \tag{8}$$

where Cu and Pb are the copper and lead concentrations in mg/L, ΔC is the increase in apparent color (measured in cpu: Co-Pt unit), T is temperature in °C, Alk is the concentration of alkalinity in mg/L as calcium carbonated ($CaCO_3$); SO_4 and SiO_2 are the concentrations of sulfates and silica in mg/L, respectively, Cl is the concentration of chlorides in mg/L, pH is the dimensionless measure of the acidity or alkalinity of a solution, and (HRT) is the hydraulic retention time in days.

Imran et al. (2006) concluded a strong relationship existed between the total iron (Fe) concentration (mg/L) and the apparent color (cpu) [1]:

$$Fe = 0.0132 \times Apparent\ Color \tag{9}$$

$$R^2 = 0.82$$

where R^2 is the correlation coefficient.

The major advantage of the above nonlinear models is explicitly describing the relationships between the metal RLs and the concentrations of secondary WQ parameters, but these nonlinear models will cause the optimization to be time consuming.

3.2. Dual Response Surface Model

The dual response surface model is composed of two second-order polynomial models, one for a mean and another for the standard deviation. The mean model of a metal release describes the relationship between the metal RLs and the concentrations of WQ parameters. The standard deviation model is to establish the relationship between the standard deviation and the WQ parameters. If it only uses the mean model for optimization, the result may have a large variation. The second empirical model of metal release considers the standard deviation; thus, the optimization can optimize the metal

release subjected to its standard deviation constraints. The dual response surface optimization can avoid generating such a large variation result and then can produce a robust result.

The WQ parameter of pH is non-conservative. It is defined as a negative decimal logarithm of the molar concentration of the hydrogen ion activity in a solution. Because the pH value of blends cannot be obtained from a mass balance equation, we chose the molar concentration of the activity of hydrogen ions A_{H+} as a substitution ($A_{H+} = 10^{-pH}$).

The next step is to convert the above metal nonlinear models into the dual response surface models. The corresponding dual response surface models are formulated as follows:

Iron-release response surface model

$$
\begin{aligned}
C_\mu^{Fe} = {}& \\
10^{-3}(& -9763 + 355T - 2Alk + 17Cl - 43Na - 6.5SO_4 + 1259DO + 699HRT \\
& - 0.1T{\bullet}Alk - 0.4T{\bullet}Cl + 1T{\bullet}Na + 0.1T{\bullet}SO_4 - 21T{\bullet}DO + 3.9T{\bullet}HRT \\
& + 0.1Alk{\bullet}Na - 0.3Alk{\bullet}DO - 0.2Alk{\bullet}HRT + 0.1Cl{\bullet}Na - Cl{\bullet}DO \\
& + 0.9Cl{\bullet}HRT + 1.9Na{\bullet}DO + 0.2Na{\bullet}HRT + 0.5SO_4{\bullet}DO \\
& + 0.2SO_4{\bullet}HRT + 8.2DO{\bullet}HRT \\
& - 4.2T^2 + 0.03Alk^2 - 0.04Cl^2 + 0.1Na^2 - 47.7DO^2 - 118.4HRT^2) \\
R^2 = {}& 0.817
\end{aligned}
\tag{10}
$$

$$
\begin{aligned}
C_\sigma^{Fe} = {}& \\
10^{-4}(& -65091 - 123T + 23Alk - 73Cl + 196Na - 39SO_4 + 2601DO + 37656HRT \\
& - 0.3T{\bullet}Alk + 1.7T{\bullet}Cl - 4T{\bullet}Na + 0.7T{\bullet}SO_4 - 20T{\bullet}DO + 11\ T{\bullet}HRT \\
& - 0.1Alk{\bullet}Na - 1.4Alk{\bullet}DO + 0.1Alk{\bullet}HRT \\
& - 0.1Cl{\bullet}Na + 5.2Cl{\bullet}DO + 0.9Cl{\bullet}HRT \\
& - 13.8Na{\bullet}DO - 0.1Na{\bullet}HRT \\
& + 2.7SO4{\bullet}DO + 0.2SO_4{\bullet}HRT + 0.3DO{\bullet}HRT \\
& + 6.9T^2 - 0.01Alk^2 - 0.04Cl^2 + 0.28Na^2 - 0.01SO_4{}^2 - 112.7DO^2 - 5431HRT^2) \\
R^2 = {}& 0.990
\end{aligned}
\tag{11}
$$

where C_μ^{Fe} is the mean response of iron release in mg/L, C_σ^{Fe} is the standard deviation response of iron release; T is the temperature in °C; Alk is the concentration of alkalinity in mg/L as calcium carbonate ($CaCO_3$); Na, SO_4 and Cl are the concentrations of sodium, sulfates and chlorides in mg/L, respectively; DO is the dissolved oxygen content in mg/L; and HRT is the hydraulic retention time in days. R^2 is the correlation coefficient; it should be noticed that the value of R^2 in Equation (10) is calculated by the quadratic equation responses against the results of Equation (8).

Copper-release response surface model

$$
\begin{aligned}
C_\mu^{Cu} = {}& \\
10^{-3}(& -252.3 + 35T + 2.4Alk - 12.3A_{H+} + 1.4SO_4 - 15SiO_2 \\
& + 0.1T{\bullet}Alk + 1.2T{\bullet}A_{H+} - 0.5T{\bullet}SiO_2 \\
& + 0.04Alk{\bullet}A_{H+} + 0.01Alk{\bullet}SO_4 - 0.14Alk{\bullet}SiO_2 \\
& + 0.04A_{H+}{\bullet}SO_4 + 0.34A_{H+}{\bullet}SiO2 - 0.08SO_4{\bullet}\ SiO_2 \\
& - 0.77T^2 - 0.68A_{H+}{}^2 + 1.34SiO_2{}^2) \\
R^2 = {}& 0.997
\end{aligned}
\tag{12}
$$

$$
\begin{aligned}
C_\sigma^{Cu} = \\
10^{-4}(-1132 + 124T + 4.8Alk - 77A_{H+} + 3SO_4 - 6SiO_2 \\
- 0.2T\bullet Alk + 2.6T\bullet A_{H+} - 0.1T\bullet SO_4 + 0.3T\bullet SiO_2 \\
+ 0.09Alk\bullet A_{H+} - 0.03Alk\bullet SiO_2 \\
+ 0.14A_{H+}\bullet SO_4 + 1.27A_{H+}\bullet SiO_2 - 0.17SO_4\bullet SiO_2 \\
- 2.92T^2 - 0.01Alk^2 - 0.81A_{H+}{}^2 - 0.07SiO_2{}^2) \\
R^2 = 0.999
\end{aligned}
\tag{13}
$$

where C_μ^{Cu} is the mean response of copper release in mg/L, C_σ^{Cu} is the standard deviation response of copper release; T is the temperature in °C; Alk is the concentration of alkalinity in mg/L as calcium carbonate ($CaCO_3$); SO_4 and SiO_2 are the concentrations of sulfates and silica in mg/L, respectively; A_{H+} is the molar concentration of the activity of hydrogen ions in 10^9 mol/L (in order to keep the same order of magnitude as the pH value). R^2 is the correlation coefficient.

Lead-release response surface model

$$
\begin{aligned}
C_\mu^{Pb} = \\
10^{-3}(6239 - 159T + 49Alk - 1527A_{H+} - 257Cl + 66SO_4 \\
-0.2T\bullet Alk + 35.3T\bullet A_{H+} + 8.3T\bullet Cl - 2T\bullet SO_4 \\
+ 3.76Alk\bullet A_{H+} + 1.25Alk\bullet Cl - 0.29Alk\bullet SO_4 \\
+ 11.6A_{H+}\bullet Cl + 1.8A_{H+}\bullet SO_4 - 1.4Cl\bullet SO_4 \\
-1.2T^2 - 0.26Alk^2 - 18.6A_{H+}{}^2 + 2.18Cl^2 + 0.12SO_4{}^2) \\
R^2 = 0.993
\end{aligned}
\tag{14}
$$

$$
\begin{aligned}
C_\sigma^{Pb} = \\
10^{-4}(-49578 + 4041T + 188Alk - 2854A_{H+} + 471Cl + 50SO_4 \\
-8.7T\bullet Alk + 62.3T\bullet A_{H+} - 5.8T\bullet Cl - 8T\bullet SO_4 \\
+ 11Alk\bullet A_{H+} - 1.5Alk\bullet Cl + 0.004Alk\bullet SO_4 \\
-5.4A_{H+}\bullet Cl + 10.3A_{H+}\bullet SO_4 + 0.4Cl\bullet SO_4 \\
-65.5T^2 - 0.22Alk^2 - 13.8A_{H+}{}^2 - 1.1Cl^2 + 0.05SO_4{}^2) \\
R^2 = 0.998
\end{aligned}
\tag{15}
$$

where C_μ^{Pb} is the mean response of lead release in mg/L, C_σ^{Pb} is the standard deviation response of lead release; T is the temperature in °C; Alk is the concentration of alkalinity in mg/L as calcium carbonate ($CaCO_3$); SO_4 and SiO_2 are the concentrations of sulfates and silica in mg/L, respectively; A_{H+} is the molar concentration of the activity of hydrogen ions in 10^9 mol/L (in order to keep the same order of magnitude as the pH value). R^2 is the correlation coefficient.

3.3. MSWBO Approach

3.3.1. Maximum Permissible Metal Release

The Lead and Copper Rule of US Environmental Protection Agency (EPA) for copper stipulates 90% of samples have a copper concentration less than 1.3 mg/L, and the concentration for lead is less than 15 μg/L [19]. The goal of this study is to find optimal blending ratios of source waters that can minimize the HMR in the distribution system. The minimized heavy metal concentrations should be less than the values of the regulations. Assuming these HMR followed a normal distribution; the maximum permissible metal RLs are their mean values. We know the standard deviations of the copper and lead RLs are 0.22 mg/L and 7 μg/L, respectively, and a one tailed Z statistic for 90% compliance is 1.28. The mean values of the copper and lead RLs are estimated on 1.02 mg/L and 6.04 μg/L, respectively. The actual government lead RL data indicated that the model (7) over-predicted the lead RL [1,5]. The maximum permissible lead release is then modified on a value of 10 μg/L.

The modification has been experimented with in the governments' distribution systems and shows a better reflection on the distribution function of lead release [1,5].

The EPA guides the secondary maximum contaminate level of iron is 0.3 mg/L [20]. We assume an influent iron RL of 0.1 mg/L, the maximum permissible iron RL within the distribution system is set at 0.2 mg/L. The metal release constraints are listed as:

$$L_{Cu} = 1 \; mg/L; \; L_{Pb} = 10 \; \mu g/L; \; L_{Fe} = 0.2 \; mg/L \tag{16}$$

where L_{Cu}, L_{Pb}, and L_{Fe} represent the maximum permissible concentrations of copper, lead, and iron in a water distribution system, respectively.

3.3.2. Variable Design

The task of this study is to find the optimal blending ratio of source waters, which consequently minimize all harmful reactions in the distribution system, as well as minimize the quality adjustment costs of source waters. The WQ chemical parameters of the source waters are considered as the decision variables. It is not necessary to use all of these chemical parameters as decision variables in the optimization model, such as some non-sensitive chemical parameters. Because these non-sensitive chemical parameters do not make much contribution to the metal release, e.g., the concentration of dissolved oxygen almost does not make contribution to the HMR, they are not necessary to be adjusted in the source waters. Hence, chemical parameters are considered as state inputs (certain value inputs). In practice, it is easy to increase the concentration of a chemical parameter in water rather than remove it from water, except for pH value. Therefore, the decision variables are designed to be varied from their initial value to a designed upper bound.

We found that only temperature, alkalinity and sulfates are common impacting parameters in the three metal release models (Equations (6)–(8)). These three common impacting parameters are selected as decision variables. Due to the parameter of temperature as an uncontrollable parameter which is dependent on ambient temperature; it is set on 25 °C in this study. Because the initial concentrations of alkalinity in the GW and sulfates in SW are high, they are considered as the state inputs.

The pH is a major importance in determining the corrosivity of water. In general, the higher pH value, the lower HMR. Equations (6) and (7) show the effects of increasing pH for copper and lead control. In water treatment process, addition of calcium hydroxide $Ca(OH)_2$ is a technology for water pH increasing. So that, in real-time operation increasing pH value of water is easier than reducing it. In this study, pH is replaced by A_{H+}, namely the molar concentration of the activity of hydrogen ions, due to the A_{H+} value in a blending source waters can be calculated by mass balance. The SMCL's for pH is 6.5 to 8.5, and the initial pH values for the GW, SW and DW are 7.9, 8.2 and 8.3, respectively. So that, we design the pH of GW is a decision variable, and the value range is from 7.9 to 8.5. The pH of the SW and DW are state inputs because they are high initial values.

Equation (6) also demonstrates that increasing concentration of silica contributes to lower copper release. Silica is then designed as a decision variable in the SW and DW rather than in the GW because of a high initial concentration of silica in the GW.

A higher chloride and sodium concentrations may be expected because of membrane deterioration in the desalination plant. Equations (7) and (8) display that a higher concentration of the chloride can cause a higher release of lead and iron, and a higher concentration of the sodium leads to a higher iron release. Because the primary objective of the MSWBO is to minimize releases of the metals, the optimization model will inhibit any increase in chloride and sodium concentrations. Chloride and sodium have to be treated as the state inputs. For the same reason, the parameters of HRT and dissolved oxygen are also considered as the state inputs (independent variables) to the model. Table 2 lists the initial values of the state inputs and the ranges of the decision variables.

Table 2. Value range of water quality parameter.

Parameter	GW	SW	DW
Alkalinity (mg/L) as $CaCO_3$	225	[50, 100]	[50, 100]
Dissolved oxygen (mg/L)	8	8	8
Silica (mg/L)	14	[7, 14]	[1, 14]
HRT (day)	5	5	5
Sodium (mg/L)	10	15	30
Chloride (mg/L)	15	10	50
Sulfates (mg/L)	[10, 80]	180	[30, 80]
pH	[7.9, 8.5]	8.2	8.3
Temperature (°C)	25	25	25

3.4. Results and Discussion

The dual response surface models are established for the releases of iron (Equations (10) and (11)), copper (Equations (12) and (13)) and lead (Equations (14) and (15)) in a particular water distribution system. On the basis of the ALs, the maximum copper and lead RLs are set at 1.0 mg/L and 10 µg/L, respectively. The maximum iron RL is established at 0.2 mg/L in accordance with the SMCLs. According to Model 5, the MSWBO approach for this particular water distribution system is developed when subjected to the dual response surface models and the relative constraints. The objectives are to minimize the metal corrosions of the pipes and minimize the adjustment costs of WQ for the multiple source waters. The decision variables of the MSWBO are selected from the WQ parameter, and the value ranges of these variables are listed in Table 2.

Three source waters are used to study the blending source waters problem. The WQ information of the three source waters are listed in Table 1. Theoretically, an optimal blending ratio may exist for a minimum release. A distribution system needs different blends of source waters, which depend on where the water demand is located on. It is not appropriate to set a constraint on source water percentages as long as there was no centralized point of blending. On the other hand, to study infinite blends of source waters is unrealistic. For brevity, only a few extreme situations and threshold blending combinations are discussed here.

The blending ratio design in the extreme situations considers the situation of single source water supply and one source off line; the threshold blending combinations are obtained from the conclusions of previous studies [1,5], such as GW contributions of >60% result in unacceptable copper release and <20% result in an unacceptable release in color. For the purpose of comparison, the release simulations for the base scenario (WQ information is shown in Table 1) are conducted though the three metal release models directly. The results are listed in Table 3.

In the objective function of the MSWBO, the term of the metal concentrations represents the WQ of blending source waters and the term of the changes of the variables represents the operation cost. The minimization of the objective reflects the trade-off between WQ and operation cost. This can be realized thought adjusting values of w_1, w_2 and w_3 as shown in Equation (5a). Here, two weighting sets are designed as w_1: w_2: w_3 = 1: 1: 1 and w_1: w_2: w_3 = 100: 100: 1.

Examination of the MSWBO indicates that the unacceptable copper release, because of the effect of high GW percentage, can be addressed by increasing the pH value and/or silica concentration in the GW. Even for 100% of the GW, the copper RL can meet its maximum permissible RL though adjusting the pH value from 7.9 to 8.3 (Tables 4 and 5). The results of the MSWBO show that high pH value could help in controlling copper and lead release.

Table 3. Results of release models.

x	y	z	Pb	Cu	Fe
1	0	0	3.53	**1.15**	0.08
0.8	0.2	0	2.34	**1.12**	0.01
0.6	0	0.4	8.13	0.96	0.03
0.43	0.41	0.16	2.91	0.97	0.01
0.2	0.8	0	2.74	0.88	0.04
0	0.1	0.9	6	0.66	**0.39**
0	1	0	4.2	0.76	**0.21**
0	0.9	0.1	3.26	0.75	**0.24**
0.2	0.6	0.2	2.54	0.86	0.1
0.1	0.4	0.5	3.28	0.76	**0.25**
0.7	0.2	0.1	3.48	**1.07**	0.01
0.1	0	0.9	8.34	0.7	**0.3**
0	0	1	7.64	0.65	**0.4**
0.19	0	0.81	8.75	0.75	**0.22**
0.1	0.3	0.6	4.11	0.75	**0.26**
0.5	0.1	0.4	6.71	0.94	0.03
0.7	0.1	0.2	5.27	**1.04**	0.01
0.1	0.9	0	3.39	0.83	0.11

Notes: x is the blending ratio of GW, y is the blending ratio of SW, z is the blending ratio of DW; Pb is in μg/L, others are in mg/L; Boldface number indicates the excess of maximum RL.

Table 4. Results of MSWBO with $w_1 = 1$, $w_2 = 1$ and $w_3 = 1$.

x	y	z	Pb	Cu	Fe	Alk SW	Alk DW	SiO$_2$ SW	SiO$_2$ DW	SO$_4$ GW	SO$_4$ DW	pH GW
1	0	0	3.05	1	0.08	-	-	-	-	10	-	8.3
0.8	0.2	0	1.94	1	0.01	100	-	14	-	10	-	8.35
0.6	0	0.4	7.72	0.89	0.03	-	50	-	14	20.3	36.9	7.95
0.43	0.41	0.16	2.99	0.89	0.01	50	50	7	1	10	30	7.93
0.2	0.8	0	2.75	0.86	0.04	50	-	8.56	-	10	-	7.9
0	0.1	0.9	9.4	0.93	**0.21**	100	100	7	1	-	30	-
0	1	0	4.2	0.72	0.2	51.1	-	8.7	-	-	-	-
0	0.9	0.1	3.33	0.75	0.2	58.5	50.9	8.7	1.19	-	30	-
0.2	0.6	0.2	2.54	0.84	0.1	50	50	8.24	1.41	10	30	7.9
0.1	0.4	0.5	3.8	0.8	0.2	62.1	65.2	7.91	2.14	10	30	7.9
0.7	0.2	0.1	3.01	1	0.01	100	100	7	1	10	30	8.45
0.1	0	0.9	10	0.89	0.2	-	89.9	-	3.4	11.8	46	7.9
0	0	1	10	0.7	**0.24**	-	100	-	10.4	-	53.8	-
0.19	0	0.81	9.26	0.75	0.2	-	58.9	-	2.65	10	30	7.9
0.1	0.3	0.6	5	0.81	0.2	61.9	73.7	7.71	2.42	10	30	7.9
0.5	0.1	0.4	6.71	0.93	0.03	50	50	7.22	1.89	10	30	7.9
0.7	0.1	0.2	5.37	0.99	0.01	64.2	78.4	14	14	10	30	7.95
0.1	0.9	0	3.39	0.79	0.11	50	-	8.64	-	10	-	7.9

Notes: x is the blending ratio of groundwater (GW), y is the blending ratio of surface water (SW), and z is the blending ratio of desalination water (DW); Pb is in μg/L, others are in mg/L; Boldface number indicates the excess of maximum RL.

Table 5. Results of MSWBO with $w_1 = 100$, $w_2 = 100$ and $w_3 = 1$.

x	y	z	Pb	Cu	Fe	Alk SW	Alk DW	SiO$_2$ SW	SiO$_2$ DW	SO$_4$ GW	SO$_4$ DW	pH GW
1	0	0	3.05	1	0.08	-	-	-	-	10	-	8.3
0.8	0.2	0	1.94	1	0.01	100	-	14	-	10	-	8.35
0.6	0	0.4	7.72	0.89	0.03	-	50	-	14	20.3	36.9	7.95
0.43	0.41	0.16	2.99	0.89	0.01	50	50	7	1	10	30	7.93
0.2	0.8	0	2.75	0.78	0.04	50	-	14	-	10	-	7.91
0	0.1	0.9	9.4	0.93	**0.21**	100	100	7	1	-	30	-
0	1	0	4.2	0.64	0.2	51.1	-	14	-	-	-	-
0	0.9	0.1	3.33	0.64	0.2	58.5	50.9	14	8.04	-	30	-
0.2	0.6	0.2	2.54	0.74	0.1	50	50	14	13.9	10	30	7.91
0.1	0.4	0.5	3.8	0.66	0.2	62.1	65.2	14	14	10	30	8.19
0.7	0.2	0.1	3.23	1	0.01	100	100	14	1	10	30	8.27
0.1	0	0.9	10	0.67	0.2	-	89.9	-	14	11.8	46	7.91
0	0	1	10	0.7	**0.24**	-	100	-	10.4	-	53.8	-
0.19	0	0.81	9.13	0.62	0.2	-	60.9	-	14	11.2	35.1	7.91
0.1	0.3	0.6	5	0.65	0.2	61.9	73.7	14	14	10	30	7.9
0.5	0.1	0.4	6.58	0.84	0.03	50	50	14	14	14.6	33.7	7.92
0.7	0.1	0.2	5.37	0.99	0.01	64.2	78.4	14	14	10	30	7.95
0.1	0.9	0	3.39	0.71	0.11	50	-	14	-	10	-	7.9

Notes: x is the blending ratio of groundwater (GW), y is the blending ratio of surface water (SW), and z is the blending ratio of desalination water (DW); Pb is in µg/L, others are in mg/L; Boldface number indicates the excess of maximum RL.

The results of the base scenario simulations (Table 3) demonstrate that the distribution system does not have a heavy metal release problem when it receives the SW that has a percentage over the 90% in the blend waters, and it is only mixed with GW. If the SW were mixed with DW and the percentage of DW excessed 10% in the blend waters; then, high iron release will occur. The test of the MSWBO shows that the SW can satisfy the original balance environment of the pipe system in any percentage, by slightly increasing the concentration of silica and/or alkalinity. For instance, 100% of SW can be distributed to the pipe system with adjusting the concentration of silica from 7 mg/L to 8.7 mg/L and the concentration of alkalinity from 50 mg/L to 51.1 mg/L (Table 4).

High DW percentages (>90%) will result in unacceptable iron release, no matter how the WQ parameters are adjusted. However, this unacceptable iron release is much lower than that of the base scenario. For example, 100% of DW would cause 0.24 mg/L of iron release after optimally adjusting its alkalinity concentration (Table 4); for the base scenario, 100% of DW would result in 0.4 mg/L of iron RL (Table 3).

Examination of the MSWBO also verifies a previous observation: "corrosion of copper and lead pipes is increased by increasing alkalinity, whereas increasing alkalinity is beneficial in reducing the release of iron products from pipes; increasing sulfates reduces lead release but increases iron release" [1–3]. These conflicting WQ requirements for release abatement can be addressed by the mathematical optimization technology of MSWBO. The trade-off between WQ and operation cost can be realized by adjusting the weighting values of w_1, w_2 and w_3. Tables 4 and 5 list the results of the MSWBO with different weighting sets of $w_1: w_2: w_3 = 1: 1: 1$ and $w_1: w_2: w_3 = 100: 100: 1$, respectively. The first weighting set shows that decision makers pay more attention to operation cost; the latter represents that decision makers are focused on WQ.

It is worth mentioning that the development of the MSWBO is based on two assumptions: (1) the pipes are of fixed geometry and the flow is in low conditions; (2) the changes of the WQ parameters (such as alkalinity and sulfates) have no effect (or have a very slight effect) on the pH value. The future studies could be: (a) develop a full-scale calibration of the model for the application of higher flow rates; (b) develop a rigorous approach to evaluating the pH of bends.

The MSWBO to the metal release problem assumes the means and variances of these metals are known. It is found that the optimal solutions are quite sensitive to these assumptions. We can use historical data to estimate the means and variances. However, the noises are rarely known and that would result in a non-optimal solution [25]. These unknown noises need to be considered in a future study.

4. Conclusions

The Multiple Source Waters Blending Optimization (MSWBO) has been developed for blending source waters to improve WQ in a distribution system by minimizing the HMR. The developed MSWBO approach has favorable advantages over other approaches in the current literature; in particular it has the following properties:

- It provides a quantitative optimal blending ratio of source waters to water utilities for minimizing the HMR in their water distribution systems.
- It has a wide range of applications, because the experiment of this model was designed in a variety of scenarios, even for some extreme situations.
- It shows a high computational efficiency, due to fact that the model only includes some second-order equations and inequalities, and also that the experiment showed that the run time is only several milliseconds.
- It exhibits a robust operation, since the MSWBO model considers the standard deviation and avoids ambiguity in the probability of inputs.

Author Contributions: Conceptualization, W.P. and R.V.M.; Methodology, W.P.; Formal Analysis, W.P. and R.V.M.; Investigation, W.P.; Writing-Original Draft Preparation, W.P.; Writing-Review & Editing, W.P. and R.V.M.; Supervision, R.V.M.; Funding Acquisition, R.V.M.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Imran, S.; Dietz, J.D.; Mutoti, G.; Xiao, W.; Taylor, J.S.; Desai, V. Optimizing Source Water Blends for Corrosion and Residual Control in Distribution Systems. *J. AWWA* **2006**, *98*, 107–115. [CrossRef]
2. Kanakoudis, V.; Tsitifli, S. Potable Water Security Assessment: A Review on Monitoring, Modelling and Optimization Techniques, Applied to Water Distribution Networks. *Desalin. Water Treat.* **2017**, *99*, 18–26. [CrossRef]
3. Peng, W.; Mayorga, R.V.; Imran, S. A rapid fuzzy optimization approach to source water blending problems in water distribution systems. *Urban Water J.* **2012**, *9*, 177–187. [CrossRef]
4. Peng, W.; Mayorga, R.V.; Imran, S. A robust optimization approach for real-time multiple source water blending problem. *J. Water Supply Res. Technol.* **2012**, *61*, 111–122. [CrossRef]
5. Imran, S.; Dietz, J.D.; Mutoti, G.; Taylor, J.S.; Randall, A.A.; Cooper, C.D. Red Water Release in Drinking Water Distribution Systems. *J. AWWA* **2005**, *97*, 93–100. [CrossRef]
6. Mehrez, C.; Percia, C.; Oron, G. Optimal Operation of a Multisource and Multi-quality Regional Water System. *Water Resour. Res.* **1992**, *128*, 1199–1206. [CrossRef]
7. Ostfeld, A.; Shamir, U. Design of Optimal Reliable Multi-Quality Water Supply Systems. *J. Water Resour. Plan. Manag. ASCE* **1996**, *122*, 322–333. [CrossRef]
8. Yang, S.L.; Sun, Y.-H.; Yeh, W.W.-G. Optimization of Regional Water Distribution System with Blend Requirements. *J. Water Resour. Plan. Manag. ASCE* **2000**, *26*, 229–235. [CrossRef]
9. Tu, M.Y.; Tsai, F.T.-C.; Yeh, W.W.-G. Optimization of Water Distribution and Water Quality by Hybrid Genetic Algorithm. *J. Water Resour. Plan. Manag. ASCE* **2005**, *131*, 431–440. [CrossRef]

10. Anderson-Cook, C.M.; Borror, C.M.; Montgomery, D.C. Response surface design evaluation and comparison. *J. Stat. Plan. Inference* **2009**, *139*, 629–641. [CrossRef]

11. Myers, R.H.; Montgomery, D.C.; Vining, G.G.; Borror, C.M.; Kowalski, S.M. Response surface methodology: A retrospective and literature survey. *J. Qual. Technol.* **2004**, *36*, 53–77. [CrossRef]

12. Yanikoglu, I.; den Denhertog, D.; Kleijnen, J.P.C. Robust Dual-Response Optimization. *IIE Trans.* **2016**, *48*, 298–312. [CrossRef]

13. Huang, T.L.; Song, X.W.; Liu, X.Y. The Multi-objective Robust Optimization of The Loading Path in the T-shape Tube Hydroforming Based on Dual Response Surface Model. *Int. J. Adv. Manuf. Technol.* **2016**, *82*, 9–12. [CrossRef]

14. Benayoun, R.; De Montgolfier, J.; Tergny, J.; Laritchev, O. Linear programming with multiple objective functions: STEP method (STEM). *Math. Program.* **1971**, *1*, 366–375. [CrossRef]

15. Kuhn, H.W.; Tucker, A.W. Nonlinear programming. In *Mathematical Statistics and Probabilities*; Neyman, J., Ed.; Berkeley Symphony: Berkeley, CA, USA, 1951.

16. Lin, D.; Tu, W. Dual response surface optimization. *J. Qual. Technol.* **1995**, *27*, 34–39. [CrossRef]

17. Vining, G.G.; Myers, R.H. Combining taguchi and response surface philosophies: A dual response approach. *J. Qual. Technol.* **1990**, *22*, 38–45. [CrossRef]

18. Chang, Y.C.; Jung, K. Effect of Distribution System Materials and Water Quality on Heterotrophic Plate Counts and Biofilm Proliferation. *J. Microbial. Biotechnol.* **2004**, *14*, 1114–1119.

19. The Lead and Copper Rule. Available online: https://www.epa.gov/dwreginfo/lead-and-copper-rule (accessed on 11 July 2018).

20. The Secondary Maximum Contaminant Levels. Available online: https://www.epa.gov/dwstandardsregulations/secondary-drinking-water-standards-guidance-nuisance-chemicals (accessed on 11 July 2018).

21. Tampa Bay Water (TBW). Available online: http://www.tampabaywater.org (accessed on 11 July 2018).

22. Box, G.E.P.; Wilson, K.B. On the Experimental Attainment of Optimum Conditions. *J. R. Stat. Soc.* **1951**, *B*, 1–45.

23. Taylor, J.; Dietz, J.D.; Randall, A.A.; Hong, S.K.; Norris, C.D.; Mulford, L.A.; Arevalo, J.M.; Imran, S.; Le Puil, M.; Liu, S.; et al. *Effects of Blend on Distribution System Water Quality*; AwwaRF, Repor No. 91065F; Denver; Water Environment Research Foundation: Alexandria, VA, USA, 2005.

24. Xiao, W. Effect of Source Water Blend on Copper Release in a Pipe Distribution System: Thermodynamic and Empirical Models. Ph.D. Thesis, University of Central Florida, Orlando, FL, USA, 2004.

25. Bingham, D.; Nair, V.N. Noise variable settings in robust design experiments. *Technometrics* **2012**, *54*, 388–397. [CrossRef]