2001 Ghausi Hall
University of California
1 Shields Avenue
Davis, CA 95616
(530) 752-0586

11 December 2019

Thomas Young, Senior Water Quality Engineer
Riverside Water Treatment Facility
32 Arsenic Drive
Riverside, CA 95674
(123) 456-7891

Dear Dr. Young:

Enclosed is the preliminary design of an adsorption-based arsenic removal system CPMD Engineering was contracted to produce on behalf of the City of Riverside on 25 September 2019.

As requested, we evaluated three candidate media, GFH, E33, and MetSorb, and produced a preliminary design based on our findings. The report includes costs associated with installing, maintaining, and operating the treatment system. In addition to documenting uncertainty, the report details two treatment scenarios, one of which is a conservative model that creates a safety margin to ensure the health of Riverside's customers.

My colleagues and I at CPMD Engineering hope that this report answers all questions you may have regarding the feasibility of removing arsenic from your drinking water via adsorption. Please contact us with any concerns you may have regarding the contents of this report.

Sincerely,

Patrick Beckett
Morgan Deangelis
Charles Hammond
Dalia Rakha
CPMD Engineering

Encl.: *Arsenic Removal from Riverside Drinking Water: Preliminary Design Report*

Arsenic Removal from Riverside Drinking Water

Preliminary Design Report

CPMD ENGINEERING

Water Quality by Quality People

Prepared for: Riverside (32) Water Treatment Facility

Prepared by: Patrick Beckett
Morgan Deangelis
Charles Hammond
Dalia Rakha

Submitted on: 11 December 2019

Executive Summary

Riverside's drinking water source (specifically #32) contains arsenic levels above the EPA maximum contaminant level of 10 μg/L, and therefore requires a treatment system to ensure the health and safety of Riverside's customers. Adsorption systems, where contaminated water is run through a "bed" of media to remove arsenic, are a standard approach. Normally, these systems are designed with the aid of rapid small-scale column tests (RSSCTs) that provide reliable information on how long the media will last. However, running RSSCTs is not an option in this case; thus, the design is based on a model of an extensive set of RSSCT data compiled at the University of California, Davis. Riverside's source water contains silica, vanadium and phosphate concentrations that compete with arsenic to reach adsorption sites. It was therefore also necessary to determine the effects of these three parameters on arsenic removal. This preliminary design report includes the evaluation of three candidate adsorbents (GFH, E33, and MetSorb) and provides the dimensions and estimated costs of a full-scale system designed to treat the entire design flow of 900 gallons per minute.

The inherent uncertainty in the modeling approach was addressed by assessing both an ideal scenario (containing extrapolated values) and a conservative scenario (limited to interpolated values). The conservative scenario provides estimates based on water quality parameters that are as close as possible to the actual numbers, without leaving the boundaries of the model. Additionally, 95% prediction intervals were calculated for both scenarios, providing upper and lower bounds for media lifetime and costs.

The recommended design consists of a five-reactor (four active/one standby) treatment system using GFH media and sulfuric acid pH adjustment (from 7.7 to 7.0) prior to treatment. The cylindrical reactors have an ideal inner diameter of 1.96 meters and a media bed height of 85 centimeters. The pH adjustment is expected to add $21,000 per year over 20 years in costs, but is estimated to save approximately $135,000 per year in media replacement costs. The estimated capital cost is $623,000 and the annual media replacement cost is estimated to be between $113,000 (extrapolated) and $160,000 (conservative) per year. pH levels had the greatest positive affect on annual media costs. Arsenic and silica concentrations also had significant influence on the media costs. As the safety of Riverside's customers is the top priority and because clean water is a human right, a rigorous arsenic monitoring program informed by the prediction intervals and the conservative estimates is recommended for the first batch of media.

Contents

Executive Summary **i**

Nomenclature **iii**

1 Introduction **1**

2 Methods **2**
- 2.1 Adsorbent Selection 2
- 2.2 Model Error and Uncertainty 4
- 2.3 Safety Factor 5
- 2.4 Economic/Alternatives Analysis 5
- 2.5 Scaling to Full-Size 5
- 2.6 Operational Configuration 8
- 2.7 Sensitivity Analysis 9
- 2.8 pH Adjustment 10

3 Results & Discussion **10**
- 3.1 Configuration and Operation 12

4 Design Recommendation **13**

References **16**

Appendix A - Python Code for Multiple Linear Regression **17**

Appendix B - Economic/Alternatives Analysis **26**
- B.1 Media Cost 26
- B.2 Capital Cost 26

Appendix C - Sensitivity Analysis Graphs **28**

Appendix D - Scale Up Example Calculations for GFH **31**

Nomenclature

Symbol/Initialism	Meaning
GFH	Granular Ferric Hydroxide
RSSCT	Rapid Small Scale Column Test
As	Arsenic
Si	Silica
PO_4^{3-}	Phosphate
P	Phosphorus
V	Vanadium
BV	Bed Volume
EBCT	Empty Bed Contact Time
MAE	Mean Absolute Error
DFPSDM	Dispersed-Flow Pore & Surface Diffusion Model
NIH	National Institute of Health
WHO	World Health Organization
EPA	United States Environmental Protection Agency
BV_{10}	Bed Volumes to 10 μg/L arsenic breakthrough

1 Introduction

The City of Riverside obtains its drinking water from underground sources, specifically from the San Bernardino, Bunker Hill, and Riverside basins [1]. Groundwater usually requires less treatment than surface water, as it is naturally filtered through the soil. However, groundwater can be rich in mineral deposits due to leaching from surrounding rocks. One of the most common groundwater pollutants of concern is arsenic, a naturally occurring element found in the Earth's crust in both its organic and inorganic form. Drinking water contaminated with arsenic usually has either arsenate (As(V)) or arsenite (As(III)), depending on the pH and redox potential of the water [2]. Arsenate typically occurs as $H_2AsO_4^-$ and $HAsO_4^{2-}$ in oxygenated waters, while arsenite occurs mostly as H_3AsO_3 in waters with lower redox potential [2].

Inorganic arsenic is a known carcinogen (according to the EPA, WHO, and NIH) and has been linked to skin, lung, bladder, kidney, and liver cancer [3] [4]. It affects a range of organs, including the nervous system, immune system, respiratory system, the skin, and developmental processes [4]. Acute affects from exposure include pigmentation and skin lesions [5]. Even at low levels, arsenic has been shown to be an endocrine disruptor. For young children, this exposure can be especially harmful. Researchers at the University of California, Berkeley found that adults who had been exposed to high levels of arsenic at an early age had increased incidents of lung and bladder cancer, even when their exposure had ceased 40 years prior. It has also been linked to reduced intellectual development in young children, as well as metabolic diseases like diabetes [4].

High levels of arsenic in drinking water sources is a global problem. Bangladesh, Chile, China, Vietnam, Taiwan, India, and the United States are among the affected countries [4]. While the US Environmental Protection Agency (EPA) limits arsenic levels to 10 μg/L, many water treatment plants struggle to meet this standard. [6]. Riverside's own arsenic concentration of 23 μg/L, as shown in Table 1.1 along with other constituents, far exceeds EPA's limit of 10 μg/L. Thus, the design of a treatment system to protect the health of Riverside's customers is needed.

Table 1.1: Riverside Drinking Water Data; "actual" is the raw water quality, "extrapolated" is the ideal model scenario, and "conservative" is the evaluation considering limitations of the model for safety [6].

Parameter	**Actual**	**pH Adjusted (Extrapolated)**	**Conservative**
Design flow (gpm)	900	-	-
pH	7.7	7.0	7.0
Alkalinity (mg/L as $CaCO_3$)	66	66	66
Arsenic (μg/L as As)	23	23	23
Silica (mg/L as SiO_2)	17	17	33
Phosphate (μg/L as P)	26	26	55
Vanadium (μg/L as V)	3.4	3.4	21

Three metallic oxide adsorbents known to be effective at removing arsenic were evaluated: Granular Ferric Hydroxide (GFH), Bayoxide E33, and MetSorb. GFH and E33 are iron-based, while MetSorb is titanium dioxide based [7]. Linear regression on breakthrough data from small-scale columns was used to estimate the lifetime of each media type under the conditions found in Table 1.1, and the optimal adsorbent was selected. The other constituents listed in Table 1.1 are known as interfering substances. They reduce the efficacy of the media by blocking or impeding arsenic adsorption in some way. Silica is one of the most important interfering substances, as silicate can pre-coat the media and reduce arsenic adsorption [8]. Additionally, pH has a dramatic effect on media lifetime; therefore, pH adjustment is considered in this analysis. The small-scale columns' parameters were then used to size a full-scale treatment system.

This report presents the preliminary design and estimated costs of an adsorption system developed to ensure that Riverside's customers receive drinking water with arsenic levels below the EPA's limit. The preliminary design has five treatment vessels (four active/one standby) operating in parallel, uses GFH media, and implements a sulfuric acid pH adjustment system. Capital costs are estimated to be $623,000, pH adjustment is estimated to both cost $21,000 per year (assuming a 20 year lifetime of the equipment) and save $135,000 per year (in reduced media costs), and media is estimated to cost between $113,000 and $160,000 per year. The following sections detail the methods, calculations, and assumptions used to reach these conclusions.

2 Methods

Various analytical methods and assumptions were used to evaluate the candidate adsorbents and produce a preliminary design. This section outlines those methods and documents the necessary assumptions.

2.1 Adsorbent Selection

The optimal adsorbent was chosen based on media cost alone, since the labor and maintenance costs are similar for all three choices. The media costs were estimated using multiple linear regression on the results of 81 rapid small-scale column tests (RSSCTs) carried out by Nguyen in his UC Davis PhD dissertation (27 per adsorbent) [9]. Nguyen evaluated the arsenic breakthrough characteristics of three adsorbents, E33, GFH, and MetSorb, as a function of pH, silica (as SiO_2), vanadium (V), phosphate (as P), and arsenic (As) concentrations. Bed volumes (BV) of water to breakthrough to 10 μg/L arsenic were recorded for 81 unique RSSCTs, each of which was configured using partial factorial experimental design. Unfortunately, the vanadium, silica, and phosphate concentrations found in Riverside's source water are not within the boundaries of Nguyen's data. Thus, extrapolation in combination with interpolation was used to estimate the breakthrough behavior of each candidate adsorbent under Riverside's water quality conditions. Additionally, a conservative estimate that stays within the bounds of the data is included.

Multiple Linear Regression

Multiple linear regression is used when multiple variables, or predictors, influence a result of interest [10]. In this case, the predictors are pH, and V, P, SiO_2, and As concentrations. The result of interest is the column bed volumes to 10 μg/L As breakthrough (BV_{10}). Thus, a function approximating the relationship between these predictors and BV_{10} takes the form

$$BV_{10} = \beta_o + \beta_1 pH + \beta_2[SiO_2] + \beta_3[V] + \beta_4[P] + \beta_5[As] + \epsilon \tag{2.1}$$

where

$[X]$ = Concentration of species X ($\mu g/L$, mg/L for SiO_2)
β_n = Model coefficients
ϵ = Error term

Taking a matrix approach, the model can be expressed as

$$\mathbf{BV_{10}} = \mathbf{X}\beta + \epsilon \tag{2.2}$$

where

$$\mathbf{BV}_{10} = \begin{bmatrix} BV_{10,1} \\ BV_{10,2} \\ \vdots \\ BV_{10,27} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_5 \end{bmatrix} \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_{27} \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} 1 & pH_1 & [SiO_2]_1 & [V]_1 & [P]_1 & [As]_1 \\ 1 & pH_2 & [SiO_2]_2 & [V]_2 & [P]_2 & [As]_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & pH_{27} & [SiO_2]_{27} & [V]_{27} & [P]_{27} & [As]_{27} \end{bmatrix}$$

The model coefficients may be estimated using the method of least squares [10]. The least squares function, given by Equation 2.3, must be minimized to find the best fit.

$$L = \sum_{i=1}^{n} \epsilon_i^2 = \left(\mathbf{y} - \mathbf{X}\beta\right)' \left(\mathbf{y} - \mathbf{X}\beta\right) \tag{2.3}$$

The solution for $\frac{\partial L}{\partial \beta} = 0$ is the least squares estimator $\hat{\beta}$, which is given by

$$\hat{\beta} = (X'X)^{-1} X' BV_{10} \tag{2.4}$$

The resulting coefficients from Equation 2.4 can be used to interpolate and extrapolate BV_{10} values for a given set of conditions. The Scikit-Learn library for the Python Programming Language was used to perform this regression in addition to other linear regression variants, including ridge, lasso, elastic net, and stochastic gradient descent regression. Based on a comparison of the R-squared values for each modeling approach, the ordinary least squares approach yielded the best results. Further improvement in the model's fit was achieved by performing the linear regression in linear-log space, meaning the predictor variables were left in linear space, while the bed volume results were log-transformed (base e) prior to regression. Goodness-of-fit calculations were performed after transforming log-space bed volume predictions back to linear space. The python script that performs this analysis is located in Appendix A.

2.2 Model Error and Uncertainty

The quality of the model for each adsorbent type was determined from R^2 and Mean Absolute Error (MAE) values which are reported in the Results section in Table 3.1. The GFH model had the highest R^2 value, but also had the highest MAE; however, this does not necessarily indicate the GFH model is inferior. Since the original data for GFH has BV_{10} values many times larger than any reported for either E33 or Metsorb, errors in the model's prediction of these large BV_{10} values contribute to this higher MAE. The MAEs for large GFH BV_{10} values were greater in absolute magnitude than equivalent percent errors for the smaller BV_{10} values of the other media types. The fact that the GFH model has the highest MAE is not incompatible with it also having the highest R^2 value.

To estimate the uncertainty in the regression's prediction of BV_{10} for GFH, prediction intervals were calculated to provide upper and lower bounds with 95% confidence. The lower bound of this prediction interval may be viewed as a "worst case" BV_{10} value, an important consideration for safety factor recommendations. The upper bound of the interval suggests that actual BV_{10} values may be higher than the predicted BV_{10} value, and thus less costly than the conservative projections our recommendations are based on.

The equation to calculate the upper and lower bounds of the prediction interval [11] can be expressed as

$$x_0'\hat{\beta} - t_{\alpha/2,n-p}\sqrt{\hat{\Omega}^2(1 + x_0'(X'X)^{-1}x_0} \leq BV_{10} \leq x_0'\hat{\beta} + t_{\alpha/2,n-p}\sqrt{\hat{\Omega}^2(1 + x_0'(X'X)^{-1}x_0} \tag{2.5}$$

where

$$x_0 = \begin{bmatrix} 1 \\ pH \\ Si \\ V \\ P \\ As \end{bmatrix}$$

is a column vector containing the specific water quality properties for the model under consideration (i.e. conservative or non-conservative). $t_{\alpha/2,n-p}$ is a parameter taken from the t-distribution for a model with n samples, p parameters, and a $100(1-\alpha)\%$ prediction interval.

and

$$\hat{\Omega}^2 = \text{variance estimator}$$

The prediction interval was calculated in python based on output data from the Multiple Linear Regression model, Nguyen's experimental data, and a $t_{\alpha/2,n-p}$ value taken from [11]. Source code for the prediction interval calculation is provided in Appendix A. It should be noted that due to the use of log (base e) transformations of the BV_{10} data, the prediction interval is not symmetric about the predicted BV_{10} value.

2.3 Safety Factor

In all public works the safety of the consumer is paramount. The highest risk for this treatment system is the risk of media exhaustion occurring earlier than anticipated, causing unsafe water to be distributed to customers. This can be avoided by monitoring the effluent arsenic levels on a weekly basis during the first run. The predicted time-to-exhaustion for the extrapolated scenario contains unavoidable uncertainty due to the limitations of linear regression, thus, the conservative scenario provides guidance on the earliest estimations for media exhaustion, since it has interfering substances at far higher concentrations than any found in Riverside's source water. Additionally, the 95% prediction intervals also provide guidance on potential breakthrough timelines. In the case that media does not last as long as anticipated, it must be replaced. If monitoring is done correctly, the worst-case scenario involves greater cost to the City of Riverside, through increased media replacement frequency, but will not result in violations or negative health consequences.

2.4 Economic/Alternatives Analysis

After estimating the bed volumes to breakthrough for each adsorbent, the total annual media cost was estimated using media unit costs obtained from the EPA and from a manufacturer. For this analysis, the unit cost of E33 and GFH are assumed to be \$5,509/m^3 and \$8,476/m^3, respectively [12]. The cost of MetSorb, \$19,500/m^3, was obtained from Pacific Water Technology [13]. The resulting annual media cost was then compared between the three candidate adsorbents and the most economical choice was selected for the final design. Approximately 80% of the operations and maintenance (O&M) costs of small adsorption-based arsenic removal systems are due to replacement of exhausted media, thus, total O&M costs were estimated by multiplying the media cost by 1.25 [14]. The capital cost associated with each media was assumed equal and the procedure to obtain the annual and capital costs associated with each media is located in Appendix B.

2.5 Scaling to Full-Size

The breakthrough characteristics used to determine the optimal adsorbent were obtained from rapid small-scale column tests (RSSCT) conducted by Nguyen [9]. The column dimensions of these RSSCTs are defined in Table 2.1.

Table 2.1: Design and operating parameters for the RSSCTs

Parameter	**E33 and GFH**		**MetSorb**	
	Small	Large	Small	Large
Mean particle size (mm)	0.127	1.16	0.127	0.67
I.D. (cm)	0.7	-	0.7	-
Bed volume (m^3)	3.9×10^{-6}	-	3.4×10^{-6}	-
Media Depth (cm)	10.2	-	10.2	-
EBCT (min)	0.33	3	0.57	3
Flow rate (m^3/min)	12×10^{-6}	-	6	-
Hydraulic Loading Rate (m/min)	0.312	0.283	0.156	0.283

The RSSCTs were designed using the scaling procedure developed in the Dispersed-Flow Pore and Surface Diffusion Model (DFPSDM) [15]. This model dictates that dispersive perfect similarity must be maintained between the performance of the RSSCT and the full-scale column. This perfect similarity is obtained by setting the dimensionless groups describing adsorbate transport in the RSSC equal to those of the full-scale column. According to the DFPSDM, if perfect similarity is maintained, the two columns will produce identical breakthrough profiles. Therefore this relationship can be used to evaluate media selection for a full-scale adsorption unit based on the results of a RSSCT.

To scale the RSSC dimensions, Nguyen assumed a linear relationship between the surface diffusion coefficient of the adsorbents and the particle diameter. This assumption results in Equation 2.6, which can be used to calculate the large column's empty bed contact time (EBCT) [16]. Using this equation, the large scale EBCT for all three adsorbents is found to be 3 min. This result is consistent with the design parameters originally assumed by Nguyen, as 3 minutes was the ideal EBCT recommended by the adsorbent manufacturers [9].

$$\frac{EBCT_{SC}}{EBCT_{LC}} = \left[\frac{d_{p,SC}}{d_{p,LC}}\right] \tag{2.6}$$

where

$EBCT$ = Empty bed contact time

d_p = Particle diameter

LC = Indicates a large column parameter

SC = Indicates a small column parameter

To scale the hydraulic loading rates between the large and small columns, Equation 2.7 is used. Once the EBCT and the scaled hydraulic loading rate is determined, the media depth can be found for that column using Equation 2.8.

$$\frac{V_{SC}}{V_{LC}} = \left[\frac{d_{p,SC}}{d_{p,LC}}\right] \tag{2.7}$$

$$L = EBCT \times V \tag{2.8}$$

where

V = Hydraulic loading rate

L = Media Depth

However, the design of the small column when using the assumed large column dimensions determined in Nguyen, using Equation 2.7 and Equation 2.8 return a small column with the same length as the large column. Because such a long small column design would result in significant head loss, this result is impractical. In order to solve this, Equation 2.9 is employed. The relationship is based off the dominance of internal mass transfer over external mass transfer in an adsorption column. The small column's Reynolds number (Re_{SC}) does not impact the breakthrough curve, as long as internal mass transfer dominates external mass transfer [15]. The $Re_{SC,min} \times Sc$ is assumed to be 2,000. This value has been found to be reasonable for most small-scale columns [16].

$$\frac{V_{SC}}{V_{LC}} = \left[\frac{d_{p,SC}}{d_{p,LC}}\right] \times \frac{Re_{LC} \times Sc}{Re_{SC,min} \times Sc} \tag{2.9}$$

$$\frac{V_{SC}}{V_{LC}} = \left[\frac{d_{p,SC}}{d_{p,LC}}\right] \times \frac{Re_{LC} \times Sc}{2,000} \tag{2.10}$$

$$Re = \frac{V \times \rho_L \times d_p}{\mu} \tag{2.11}$$

$$Sc = \frac{\mu}{D_L \times \rho_L} \tag{2.12}$$

where

Re = Reynolds number

$Re_{SC,min}$ = Reynolds number defining the minimum velocity in the small column

Sc = Schmidt number

ρ_L = Density of water at 20°C, 998 kg/m^3

μ = Dynamic viscosity of water at 20°C, $8.891 \times 10^{-4} \frac{kg}{m \cdot s}$

Because the Reynolds number is dependent on the loading rate, Equation 2.10 can only be used to scale down a full-size column. Therefore, the large column's loading rate must be the same as the originally assumed large column's hydraulic loading rate in Nguyen's RSSCT design (17 m/hr). Once the hydraulic loading rate was determined to be 17 m/hr, the large column's media depth is found to be 85 cm using Equation 2.8.

Using the EBCT and the large scale flow rate as shown in Equation 2.13, the full scale bed volume for each of the three adsorbents is calculated to be 10.22 m^3.

$$BV = EBCT \times Q \tag{2.13}$$

where

BV = Bed volume

Q = Flow rate

The large column internal diameter is calculated using the hydraulic loading rate and the flow rate as shown in Equation 2.14. The large column Diameter is found to be 3.91 m.

$$A = \left[\frac{\pi}{4}\right] D^2 = \frac{Q}{V} \tag{2.14}$$

where

A = Cross sectional area of the column

D = Internal diameter of the column

An example of the above equations used to scale a large column for E33 is presented in Appendix D.

2.6 Operational Configuration

The EPA recommends an adsorption system operating with beds in parallel. While operating the columns in series may increase specific throughput (the volume of water treated per unit weight of adsorbent) by 20-50%, single beds operated in parallel are the least expensive option when 80% removal or less is required [14]. Additionally, operating columns in series exposes media to interfering substances for longer than necessary, for example, silicate can pre-coat the media, which can result in reduced efficacy [8]. Operating beds in parallel can significantly increase the specific throughput and can reduce the required amount of adsorbent. Beds in parallel are especially cost effective as effluent from exhausted columns can be blended with flow from fresh columns.

The EPA also recommends the inclusion of optional bypass piping. For systems where the raw water arsenic concentration is slightly above the arsenic max contaminant level (MCL), bypassing and reblending a portion of the raw water with the remaining treated effluent can extend treatment media's use and can reduce the total operating cost. Additionally, a redundant column is included to allow for continuous operation during media replacement and for emergency or contingency situations.

The proposed system layout is depicted in Figure 3.2. To adjust the "scaled up" large column dimensions for this configuration, the media depth and hydraulic loading rates have to remain 85 cm and 17 m/hr respectively in order to maintain perfect dispersive similarity between the large columns and the RSSCT. Flow rate, however, (assuming no bypass use) will be split evenly between the four columns and thus according to Equation 2.14, the area of the individual columns will be a quarter of the original area. The internal diameters of the individual columns are then calculated to be 1.956 m using Equation 2.14. A summary of the dimensions for each of the five columns (four active, one standby) is presented in Table 2.2 and the calculations used to determine these values are presented in Appendix E.

Table 2.2: Design and operating parameters for individual columns in Parallel

Parameter	**Value**
Flow rate (m^3/min)	0.853
Hydraulic Loading Rate (m/min)	0.283
Media Depth (cm)	85
I.D. (m)	1.96
Area (m^2)	3.00
Bed volume (m^3)	2.56
EBCT (min)	3

2.7 Sensitivity Analysis

Estimates of how source water quality affects annual media cost are given by a sensitivity analysis considering each variable of interest (pH, V, SiO_2, P, As). This was done by first measuring their modeled effect on bed volumes to breakthrough, defined by when effluent arsenic levels reach 10 μg/L. To determine how each of these characteristics individually affected the BV values, the isolated variable was altered while the other four remained constant at the specified scenario's default values.

For each variable, two ranges were selected. The first range, termed "Model data", describes the values that fit within the boundaries of the regression model. The second range consists of values requiring extrapolation. For pH, the extrapolated data is based on Xue et al.'s [17] statement that the average pH range in most drinking water treatment plants is between 6.6 to 8.6. The extrapolated data points for phosphate, vanadium, and silica filled the concentration gap between zero and the lowest value in Nguyen's range. This is because Riverside's source water had very low concentrations of these three parameters, all of which fell below Nguyen's range. For arsenic concentrations, the extrapolated points filled the gap between 0 and 14 μg/L to cover the range of concentrations recorded by the Riverside's water system data from 2018 [1]. In addition to these two ranges, a data point was added to represent the empirical source water data obtained from the Riverside water system [6]. This point is displayed as an "X" on the figure. In the pH graph, an additional data point was added to represent the adjusted pH level.

From the corresponding BV values, the longevity of the media based on the incoming flow rate and original arsenic concentration was calculated. In this way, we were able to determine how many times within a year the media would need to be changed. The price per volume of the media was then used to calculate the annual media cost to treat water from this system to meet the specified arsenic requirements.

Two sets of BV data were analyzed for each variable. The first set, labeled "extrapolated," utilize the pH adjusted source water data (see Table 1.1) [6]. The second set, labeled "conservative," redefined these variables to values closest to the actual water quality as possible given the data from Nguyen's experiments. Both the extrapolated and conservative analysis utilize the adjusted pH value of 7 to calculate BV values.

2.8 pH Adjustment

The decision to adjust the pH of the source water depends on the potential savings compared to the cost of implementing and operating the adjustment system. For Riverside's source water, the pH adjustment would be downward to a pH of approximately 7. The capital cost of implementing a pH adjustment system in conjunction with a typical adsorption system is estimated to be \$135,000 (2003 dollars) for a 570 gpm system [18]. Assuming cost scales linearly with flow rate, and adjusting to 2019 dollars, the estimated capital cost of implementing the adjustment system is given by

$$Cost_{adj} = \$135,000 \times \frac{900\, gpm}{570\, gpm} = \$213,158 \tag{2.15}$$

And adjusting to 2019 dollars using the Bureau of Labor Statistics' CPI inflation calculator, the capital cost of an appropriate pH adjustment system is approximately \$302,000 [19]. Assuming an equipment lifetime of 20 years, an average annual maintenance cost of 1% of the capital cost, and an annual inflation rate of 2%, and an internal rate of return of 6% the equivalent annual cost (AC) of pH adjustment is found as shown below, using the equations found in Table 2.11 of Fundamental of Engineering Economics [20] [11]. First the net present cost (NPC) of the initial capital cost and the maintenance costs adjusted for inflation are calculated by

$$NPC_{pH} = \$302,000 + \$3,020 \times \left[\frac{1-(1+0.02)^{20}(1+0.06)^{-20}}{0.06-0.02}\right] = \$342,500 \tag{2.16}$$

The NPC is then annualized as follows, using the interest factors for discrete compounding found in [11].

$$AC_{pH} = \$342,500 \times (A/P, 2\%, 20) = \$342,500 \times 0.0612 = \$21,000/y \tag{2.17}$$

The savings associated with adjusting the pH were estimated via the regression model. The difference in annual operating costs (extrapolated version) between pH 7 and pH 7.7 is approximately \$247,499 – \$112,500, or \$135,000/y.

3 Results & Discussion

The multiple linear regression yielded well-fitting results for all three adsorbents, as shown by the the R^2 values and the mean average errors in Table 3.1. Two scenarios were considered for the analysis: "extrapolated" means the inputs to the model are the given (pH adjusted) water quality parameters shown in Table 1.1; "conservative" means the inputs to the model are as close to the actual values as possible without leaving the boundaries of the model's data. Because extrapolation is prone to error and uncertainty, the "conservative" scenario estimates costs and bed volumes to breakthrough with higher confidence; however, the real-life results will almost certainly be better than the conservative scenario predicts, since the actual values for SiO_2, P, and V are all significantly lower than the lower boundaries of the model.

Table 3.1: Results of multiple linear regression on RSSCT data from Nguyen's dissertation.

Parameter		GFH	E33	MetSorb
	R^2	0.97	0.89	0.91
	Mean Absolute Error (BVs)	2,799	1,967	1,985
Bedvolumes to 10 ppb breakthrough:	Extrapolated	61,298	38,965	28,473
	Conservative	42,770	28,288	19,012

Table 3.2 displays the costs associated with adsorption systems using each type of media. The costs of GFH and E33 were obtained from an EPA document on "US EPA Arsenic Removal Technology Demonstration Program" [12]. The cost of MetSorb was obtained from the website of the vendor, Pacific Water Technology [13]. As Table 3.2 shows, GFH is the most economical choice, and is therefore the recommended adsorbent media. To obtain the total operations and maintenance cost (O&M), the media cost was multiplied by 1.25, as media is on average 80% of the total O&M cost of smaller iron-bed adsorption systems [14].

Table 3.2: Predicted cost of media for each type of treatment system for both scenarios.

Parameter	GFH	E33	MetSorb
Unit Media Cost ($\$/m^3$)	8,476	5,509	19,500
Media Cost, Extrapolated ($/year)	113,000	157,000	486,000
Media Cost, Conservative ($/year)	160,000	217,000	712,000

Table 3.3 shows the estimated bed volumes to breakthrough and corresponding media lifetime for GFH for both the extrapolated and the conservative scenarios. The estimates include the lower and upper bounds of the 95% prediction interval. Notice the conservative scenario's predicted values are similar to the extrapolated scenario's lower boundary. In terms of days to exhaustion, the conservative scenario provides a safety factor of 1.4. But rather than blindly replacing the media every 197 days, a monitoring program should use the conservative scenario as guidance on sampling frequency and timing.

Table 3.3: Estimates of the media lifetime, including uncertainty; breakthrough is defined as the point when effluent from the treatment system has 10 μg/L arsenic or greater.

Scenario	Parameter	Lower	Predicted	Upper
Extrapolated	Bed Volumes to Breakthrough	95,650	134,380	188,800
	Days to Exhaustion	201	281	394
Conservative	Bed Volumes to Breakthrough	69,360	93,760	126,750
	Days to Exhaustion	146	197	263

As Figure 3.1 shows, pH dramatically affects the annual media cost. Because the annualized cost of pH adjustment is $21,000/y while the estimated savings are $135,000/y, the clear choice is to implement a pH adjustment system. Figure 3.2 demonstrates how the pH adjustment system would be integrated with the design.

The sensitivity analysis revealed that pH has the most significant influence on annual media cost. For the

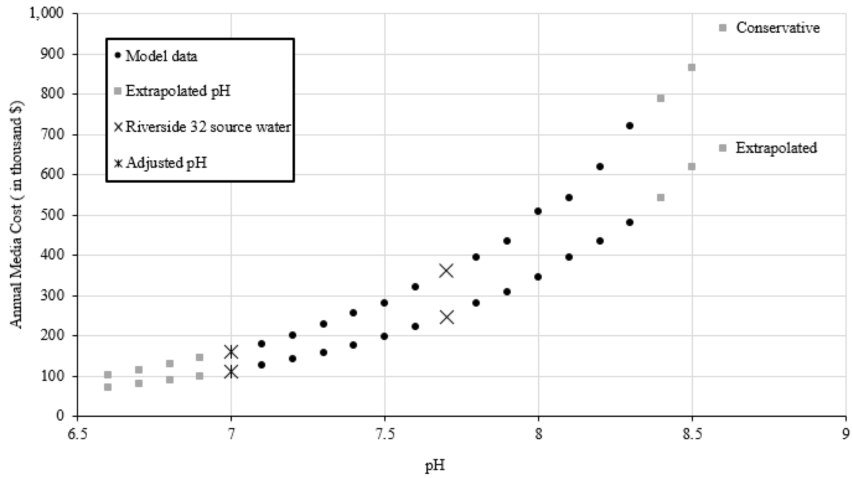

Figure 3.1: Sensitivity of annual media cost to pH.

conservative values (those values altered to fit within Nguyen's data), the annual media cost ranged from approximately $103,000 to $962,000. For the extrapolated values (the Riverside source water data), the annual media cost ranged from approximately $72,000 to $666,000. In adjusting the pH, annual media cost was reduced by $200,000 in the conservative scenario and by $135,000 in the extrapolated scenario.

Silica and arsenic concentrations significantly affected annual media cost. While not as drastic as pH, both parameters showed similar nonlinear relationships with media cost. The sensitivity plots for the remaining variables are included in Appendix C and an overview of the findings are summarized below in Table 3.4. With the exception of pH, the low value was at zero concentration and the high value was the largest concentration in Nguyen's data. From these two values, a general scope was calculated to describe the range of costs presented by changing each variable individually. As Table 3.4 indicates, it is evident that pH concentrations had the greatest affect on annual media cost. All five variables had a positive correlation between concentration and annual media cost.

Table 3.4: Annual media cost (in thousand $) for varying source water parameters

Parameter	**Conservative**			**Extrapolated**		
	Low	High	Range	Low	High	Range
pH	103	962	859	72	666	594
Silica (mg/L as SiO_2)	94	309	215	86	279	194
Vanadium (μg/L as V)	147	192	46	111	144	33
Phosphate (μg/L as P)	155	173	19	111	126	14
Arsenic (μg/L as As)	95	333	238	67	234	167

3.1 Configuration and Operation

The proposed system configuration is depicted in Figure 3.2. The design consists of four active treatment vessels running in parallel and includes one standby vessel to allow for media replacement and contingencies. The design incorporates an optional bypass line and includes a sulfuric acid pH adjustment system.

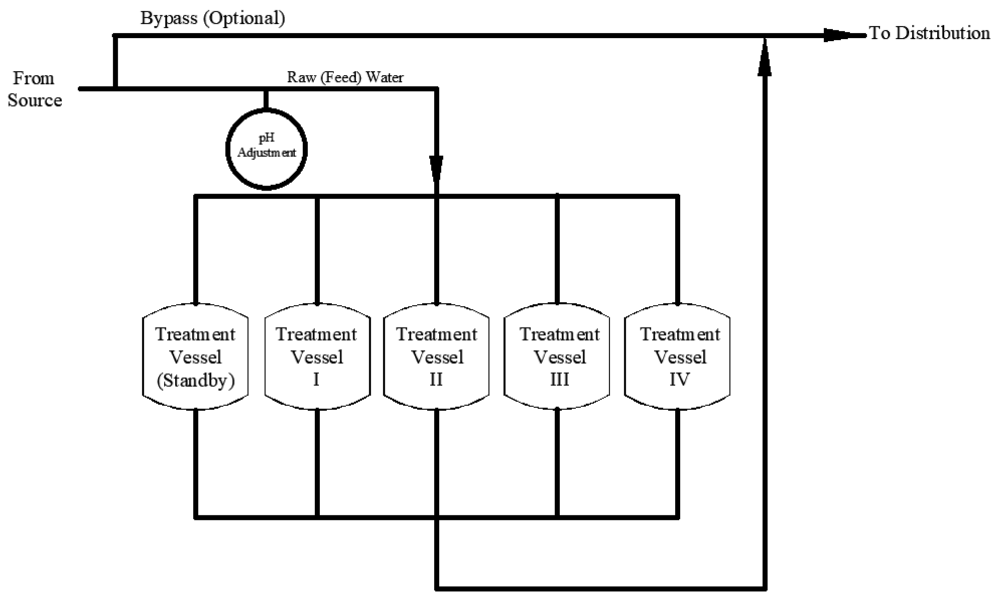

Figure 3.2: Proposed Adsorption System Layout; based on EPA recommended process and instrumentation layout for the proposed adsorption system [18].

The individual column dimensions presented in Figure 3.3 are based on our calculated large column dimensions split between four active columns as discussed in section 2.6 of this report. These values were then applied to the EPA recommended vessel configuration found in [18].

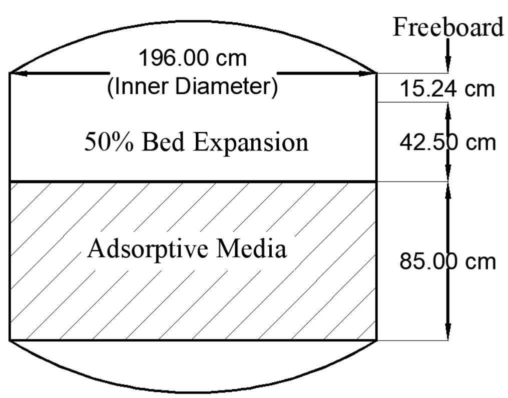

Figure 3.3: Dimensions for the proposed adsorption columns; based on EPA recommended treatment vessel specifications [18].

4 Design Recommendation

Based on the modeling results and the economic analysis, CPMD Engineering recommends GFH as the media of choice. It offers the best combination of unit media cost and bed volumes to arsenic breakthrough under the source water conditions, both of which are contained in the design summary shown in Table 4.1. The design will consist of four active columns each treating one quarter of the total flow, or 225 gpm each. One additional column will serve as standby in case of emergency malfunction or maintenance needs. For the proposed design layout see Figure 3.2.

Table 4.1: Design specifications for adsorptive media arsenic removal system.

Parameter	Design
Media	GFH
Unit Cost ($/m^3)	8,476
Total Flow Treated (gpm)	900
Flow per Column (gpm)	225
Active Columns	4
Standby Columns	1
Column Inner Diameter (m)	1.96
EBCT (min)	3.0
Bed Depth (cm)	85
Media Volume per Column (m^3)	2.56

The pH will be adjusted using an automated sulfuric acid dosing system. As shown in Table 4.2, the reduced cost of media replacement far outweighs the equivalent annual cost of installing and maintaining the pH adjustment equipment. The media cost for the first year is higher than other years because the redundant column must be supplied with media. To ensure the media does not age excessively, the redundant column should not sit unused for more than a year.

Table 4.2: Costs (in thousands of dollars) for both the extrapolated and conservative scenarios and including uncertainty intervals.

Scenario	Parameter	Lower	Predicted	Upper
	Capital Cost	-	623	-
	pH Adjustment, Equivalent Annual Cost	-	21	-
	Annual Savings Due to pH Adjustment	-	135	-
	Media, First Year	102	135	179
Extrapolated	Annual Media Replacement	80	113	157
	Additional Annual O&M	20	28	39
	Media, First Year	142	182	239
Conservative	Annual Media Replacement	120	160	217
	Additional Annual O&M Cost	30	40	54

To ensure the safety of Riverside's customers and mitigate the uncertainty inherent in this design, the conservative values for media exhaustion may be used to inform monitoring of arsenic levels. The conservative approach is effectively a safety factor of 1.4 on the extrapolated scenario, providing a safety margin in which Riverside's operators can determine the true operating characteristics of the treatment system. Because Riverside's water contains V, P, and SiO_2 at much lower levels than used in the conservative scenario, the real performance will likely be superior.

References

[1] “Water Quality Report 2018.” https://www.riversideca.gov/utilities/pdf/Water

[2] M. Badruzzaman, P. Westerhoff, and D. Knappe, “Intraparticle diffusion and adsorption of arsenate onto granular ferric hydroxide (GFH),” *Water Research*, July 2004.

[3] “Arsenic.” Environmental Working Group, https://www.ewg.org/tapwater/reviewed-arsenic.php, October 2019.

[4] National Institute of Health, “Arsenic Fact Sheet,” 2014.

[5] World Health Organization, “Arsenic,” 2018.

[6] Thomas Young, “Design project.” Provided on ECI 243A Canvas website, 2019.

[7] P. Westerhoff, M. D. Haan, A. Martindale, and M. Badruzzaman, “Arsenic Adsorptive Media Technology Selection Strategies,” *Water Quality Research*, 2006.

[8] M. Kanematsu, T. Young, K. Fukushi, P. Green, and J. Darby, “Individual and combined effects of water quality and empty bed contact time on as(v) removal by fixed-bed iron oxide adsorber: Implication for silicate precoating,” *Water Research*, July 2012.

[9] V. L. Nguyen, *Effects of Interferences on the Breakthrough of Arsenic: Rapid Small-Scale Column Tests*. PhD thesis, University of California Davis, 2011.

[10] D. Montgomery and G. Runger, *Applied Statistics and Probability for Engineers*. Wiley, 6th ed., 2014.

[11] C. S. Park, *Fundamentals of Engineering Economics*. Pearson, 3rd ed., 2013.

[12] Lili Wang and Abraham Chen, “Costs of Arsenic Removal Technologies for Small Water Systems: US EPA Arsenic Removal Technologies Demonstration Program,” 2011.

[13] Pacific Water Technology, “Heavy Metals and Arsenic Filter Media.” `https://pacificwater.com.au/product/metsorb-arsenic-heavy-metals-removal/`, 2019.

[14] T. Sorg, A. Chen, L. Wang, and R. Kolisz, “Regenerating an Arsenic Removal Iron-Bed Adsorptive Media System, Part 1: The Regeneration Process,” *Journal of American Water Works Association*, May 2017.

[15] J. Crittenden, J. Berrigan, D. Hand, and B. Hykins, “Design of rapid fixed-bed adsorption tests for nonconstant diffusivities,” *Journal of Environmental Engineering-ASCE*, 1987.

[16] P. Westerhoff, D. Highfield, M. Badruzzaman, and Y. Yoon, “Rapid small-scale column tests for arsenate rmoval in iron oxide packed bed columns,” *Journal of Environmental Engineering-ASCE*, 2005.

[17] R. Xue, H. Shi, Y. Ma, J. Y. B. Hua, E. C. Inniss, C. D. Adams, and T. Eichholz, “Evaluation of thirteen haloacetic acids and ten trihalomethanes fomration by peracetic acid and chlorine drinking water disinfection,” *Chemosphere*, 2017.

[18] F. Rubel, *Design Manual: Removal of Arsenic from Drinking Water by Adsorptive Media*. US Environmental Protection Agency, 2003.

[19] Bureau of Labor Statistics, "CPI Inflation Calculator." https://data.bls.gov/cgi-bin/cpicalc.pl?cost1=213157.89&year1=200301&year2=201910, 2019.

[20] US Inflation Calculator, "Current US Inflation Rates: 2009 to 2019." https://www.usinflationcalculator.com/inflation/current-inflation-rates/, 2019.

Appendix A - Python Code for Multiple Linear Regression

```
import numpy as np
from sklearn import linear_model
from sklearn.linear_model import SGDRegressor
import sklearn.metrics as metrics
import math as m
import statsmodels
from regression_results import error_stats
from prettytable import PrettyTable

#----------------------------------------------------------------
## Nguyen DATA
#----------------------------------------------------------------

pH = (np.array([8.3,8.3,8.3,8.3,8.3,8.3,8.3,8.3,8.3,7.65,7.65,
        7.65,7.65,7.65,7.65,7.65,7.65,7.65,7.00,7.00,7.00,7.00,7.00,
                7.00,7.00,7.00,7.00,])) # S.U.
Si = np.array([33,53,73,33,53,73,33,53,73,33,53,73,33,53,73,33,
        53,73,33,53,73,33,53,73,33,53,73]) # Silica as mg/L as SiO2
V = np.array([21,41,61,41,61,21,61,21,41,41,61,21,61,21,41,21,41,
        61,61,21,41,21,41,61,41,61,21]) # Vanadium, ug/L as V
P = np.array([55,55,55,155,155,155,105,105,105,105,105,105,55,55,55,155,
        155,155,155,155,155,105,105,105,55,55,55]) # ug/L as P
As = np.array([15,15,15,35,35,35,55,55,55,15,15,15,35,35,35,55,55,55,15,
        15,15,35,35,35,55,55,55]) # ug/L as As
BV10_GFH = np.log(np.array([30000,16900,12500,13300,9200,9000,10200,
                6600,5300,51000,25200,23000,26300,20700,16000,19700,11100,8300,
                        96000,88000,55000,65000,40600,29900,42800,29200,27100]))
BV10_E33 = np.log(np.array([30100,17100,13900,7200,5500,5100,8500,6000,
                4000,31500,20100,17500,19300,16000,15100,10100,7500,7200,33500,
                        31200,30900,21900,18300,16100,22600,16500,14100]))
BV10_MET = np.log(np.array([13500,8800,4600,5000,2300,2000,4300,1500,
                1400,22100,14900,7200,13900,8400,5800,7700,5000,3100,38900,36800,
                        29000,25900,18600,12000,16600,15000,11700]))

Xp = np.array([pH,Si,V,P,As])
# print('Xp:'+ str(Xp.shape))
X = Xp.T
#print('X:'+ str(X.shape))

#----------------------------------------------------------------
## Design Criteria!
#----------------------------------------------------------------

pH_m = 7
```

```
 # S.U.
Si_m = 17 # mg/L as SiO2
V_m = 3.4 # ug/L as V
P_m = 80*30.974/(4*15.999+30.974) # ug/L as P (coversion shown here)
As_m = 23 # ug/L as As

## Conservative Values, no extrapolation here.

pH_m_c = 7 # S.U.
Si_m_c = 33 # mg/L as SiO2
V_m_c = 21 # ug/L as V
P_m_c = 55 # ug/L as P (coversion shown here)
As_m_c = 23 # ug/L as As


#---------------------------------------------------------------
## GFH
#---------------------------------------------------------------
print("For ordinary least squares regression on log-transformed BV10s: ")
print()

lm = linear_model.LinearRegression()
model = lm.fit(X,BV10_GFH)
PredictionsORD_GFH = lm.predict(X)

MAE_GFH = np.around(metrics.mean_absolute_error(np.exp(BV10_GFH),
          np.exp(PredictionsORD_GFH)))
r2_GFH=np.around(metrics.r2_score(np.exp(BV10_GFH),
        np.exp(PredictionsORD_GFH)),2)

Coef = lm.coef_

BV10_m = np.around(np.exp(lm.intercept_ + Coef[0]*pH_m +
        Coef[1]*Si_m + Coef[2]*V_m + Coef[3]*P_m+Coef[4]*As_m))
BV10_m_c = np.around(np.exp(lm.intercept_ + Coef[0]*pH_m_c +
        Coef[1]*Si_m_c + Coef[2]*V_m_c + Coef[3]*P_m_c+Coef[4]*As_m_c))

# Prediction Interval calcuations

# modified Xp for PI calculations
ones = np.array([1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1])
Xp_ones = np.array([ones,pH,Si,V,P,As])
Xp_ones_T = Xp_ones.T


D = np.array([1, pH_m, Si_m, V_m, P_m, As_m])
D_c = np.array([1, pH_m_c, Si_m_c, V_m_c, P_m_c, As_m_c])
D_T = np.array([[1], [pH_m], [Si_m], [V_m], [P_m], [As_m]])
```

```
D_T_c = np.array([[1], [pH_m_c], [Si_m_c], [V_m_c], [P_m_c], [As_m_c]])
B_column = np.array([[lm.intercept_], [Coef[0]], [Coef[1]], [Coef[2]],
        [Coef[3]], [Coef[4]]])
B_row = B_column.reshape(1, -1)
GFH_column = BV10_GFH.reshape(-1, 1)

# this is the product of x0 and estimators
xB = np.matmul(D, B_column)
xB_c = np.matmul(D_c, B_column)

# non conservative
omega_GFH_pt1 = np.matmul(BV10_GFH, GFH_column)
omega_GFH_pt2 = np.matmul(Xp_ones, GFH_column)
omega_GFH_pt3 = np.matmul(B_row, omega_GFH_pt2)
omega_GFH_f1 = omega_GFH_pt1 - omega_GFH_pt3
omega_GFH_f2 = omega_GFH_f1/21

S_pt1 = np.matmul(Xp_ones, Xp_ones_T)
S_pt2 = np.matmul(np.linalg.inv(S_pt1), D_T)
S_pt3 = np.matmul(D, S_pt2)
S_final = 1 + S_pt3

B_low_GFH = np.exp(xB - 2.080*np.sqrt(omega_GFH_f2*S_final))
B_high_GFH = np.exp(xB + 2.080*np.sqrt(omega_GFH_f2*S_final))

print("The lower bound for non-conservative PI is equal to:",
        B_low_GFH)
print("The upper bound for non-conservative PI is equal to:",
        B_high_GFH)

# conservative

S_pt2_c = np.matmul(np.linalg.inv(S_pt1), D_T_c)
S_pt3_c = np.matmul(D_c, S_pt2_c)
S_final_c = 1 + S_pt3_c

B_low_GFH_c = np.exp(xB_c - 2.080*np.sqrt(omega_GFH_f2*S_final_c))
B_high_GFH_c = np.exp(xB_c + 2.080*np.sqrt(omega_GFH_f2*S_final_c))

print("The lower bound for non-conservative PI is equal to:",
        B_low_GFH_c)
print("The upper bound for non-conservative PI is equal to:",
        B_high_GFH_c)

#---------------------------------------------------------------
## E33
#---------------------------------------------------------------
```

```
lm = linear_model.LinearRegression()
model = lm.fit(X,BV10_E33)
PredictionsORD_E33 = lm.predict(X)

MAE_E33 = np.around(metrics.mean_absolute_error(np.exp(BV10_E33),
        np.exp(PredictionsORD_E33)))
r2_E33=np.around(metrics.r2_score(np.exp(BV10_E33),
        np.exp(PredictionsORD_E33)),2)

Coef = lm.coef_

BV10_m_E33 = np.around(np.exp(lm.intercept_ + Coef[0]*pH_m +
                Coef[1]*Si_m + Coef[2]*V_m + Coef[3]*P_m+Coef[4]*As_m))
BV10_m_E33_c = np.around(np.exp(lm.intercept_ + Coef[0]*pH_m_c +
                Coef[1]*Si_m_c + Coef[2]*V_m_c + Coef[3]*P_m_c+Coef[4]*As_m_c))

#-----------------------------------------------------------------
## METSORB
#-----------------------------------------------------------------

lm = linear_model.LinearRegression()
model = lm.fit(X,BV10_MET)
PredictionsORD_MET = lm.predict(X)

MAE_MET = np.around(metrics.mean_absolute_error(np.exp(BV10_MET),
                np.exp(PredictionsORD_MET)))
r2_MET=np.around(metrics.r2_score(np.exp(BV10_MET), np.exp(PredictionsORD_MET)),2)

Coef = lm.coef_

BV10_m_MET = np.around(np.exp(lm.intercept_ + Coef[0]*pH_m + Coef[1]*Si_m +
                Coef[2]*V_m + Coef[3]*P_m+Coef[4]*As_m))
BV10_m_MET_c = np.around(np.exp(lm.intercept_ + Coef[0]*pH_m_c + Coef[1]*Si_m_c +
                Coef[2]*V_m_c + Coef[3]*P_m_c+Coef[4]*As_m_c))

#-----------------------------------------------------------------
## COST CALCULATION - Data from Table 2-10 in EPA document (2011)
#-----------------------------------------------------------------

Cost_GFH = np.around(240*35.315) # $/m3
Cost_E33 = np.around(156*35.315) # $/m3
Cost_MET = np.around(3900/200*1000) # $/m3


BV_GFH = 10.22 # m3/BV, bed volume of the large scale reactor for GFH
BV_E33 = 10.22 # m3/BV
BV_MET = 10.22  # m3/BV
```

```
Q = 900*60*24*365/264.172 # m3/year
print(Q)

BV_Treated_GFH = Q/BV_GFH # BV/year
BV_Treated_E33 = Q/BV_E33 # BV/year
BV_Treated_MET = Q/BV_MET # BV/year

Time_to_repl_GFH = np.around(BV10_m/BV_Treated_GFH,2) # Time to
# replacement for media, years
Time_to_repl_GFH_c = np.around(BV10_m_c/BV_Treated_GFH,2)

Time_to_repl_GFH_low = np.around(B_low_GFH/BV_Treated_GFH,2) # Time to
# replacement for media, years
Time_to_repl_GFH_high = np.around(B_high_GFH/BV_Treated_GFH,2) # Time to
# replacement for media, years

Time_to_repl_GFH_c_low = np.around(B_low_GFH_c/BV_Treated_GFH,2)
Time_to_repl_GFH_c_high = np.around(B_high_GFH_c/BV_Treated_GFH,2)

Cost_per_year_GFH = np.around(BV_GFH*Cost_GFH/Time_to_repl_GFH)
# $/year
Cost_per_year_GFH_c = np.around(BV_GFH*Cost_GFH/Time_to_repl_GFH_c)
# $/year

Cost_per_year_GFH_low = np.around(BV_GFH*Cost_GFH/Time_to_repl_GFH_high)
# $/year
Cost_per_year_GFH_high = np.around(BV_GFH*Cost_GFH/Time_to_repl_GFH_low)
# $/year
Cost_per_year_GFH_c_low = np.around(BV_GFH*Cost_GFH/Time_to_repl_GFH_c_high)
# $/year
Cost_per_year_GFH_c_high = np.around(BV_GFH*Cost_GFH/Time_to_repl_GFH_c_low)
# $/year

Time_to_repl_E33 = np.around(BV10_m_E33/BV_Treated_E33,2) # Time
# to replacement for media, years
Time_to_repl_E33_c = np.around(BV10_m_E33_c/BV_Treated_E33,2)

Time_to_repl_MET = np.around(BV10_m_MET/BV_Treated_MET,2) # Time
# to replacement for media, years
Time_to_repl_MET_c = np.around(BV10_m_MET_c/BV_Treated_MET,2)

Cost_per_year_E33 = np.around(BV_E33*Cost_E33/Time_to_repl_E33)
# $/year
Cost_per_year_E33_c = np.around(BV_E33*Cost_E33/Time_to_repl_E33_c)
 # $/year

Cost_per_year_E33_total = np.around(BV_E33*Cost_E33
        /Time_to_repl_E33/.8) # $/year
```

```
Cost_per_year_E33_c_total = np.around(BV_E33*Cost_E33
        /Time_to_repl_E33_c/.8) # $/year

Cost_per_year_MET = np.around(BV_MET*Cost_MET
        /Time_to_repl_MET) # $/year
Cost_per_year_MET_c = np.around(BV_MET*Cost_MET
        /Time_to_repl_MET_c) # $/year

Cost_per_year_MET_total = np.around(BV_MET*Cost_MET
        /Time_to_repl_MET/.8) # $/year
Cost_per_year_MET_c_total = np.around(BV_MET*Cost_MET
        /Time_to_repl_MET_c/.8) # $/year

x = PrettyTable()

x.field_names = ["Parameter", "GFH", "E33", "MetSorb"]

x.add_row(["R2", r2_GFH, r2_E33, r2_MET])
x.add_row(["MAE", MAE_GFH, MAE_E33, MAE_MET])
x.add_row(["BV to breakthrough, extrapolated", BV10_m,
        BV10_m_E33,
        BV10_m_MET])
x.add_row(["BV to breakthrough, extrapolated, low", np.around(B_low_GFH),
        "-", "-"])
x.add_row(["BV to breakthrough, extrapolated, high", np.around(B_high_GFH),
        "-", "-"])
x.add_row(["BV to breakthrough, conservative", BV10_m_c,
        BV10_m_E33_c,
        BV10_m_MET_c])
x.add_row(["BV to breakthrough, conservative, low", np.around(B_low_GFH_c),
        "-", "-"])
x.add_row(["BV to breakthrough, conservative, high", np.around(B_high_GFH_c),
        "-", "-"])
x.add_row(["Unit Cost ($/m3)", Cost_GFH, Cost_E33, Cost_MET])
x.add_row(["Time to replacement, extrapolated (days)",
        np.around(365*Time_to_repl_GFH),
        np.around(365*Time_to_repl_E33), np.around(365*Time_to_repl_MET)])
x.add_row(["Time to replacement, extrapolated, low (days)",
        np.around(365*Time_to_repl_GFH_low),
        "-", "-"])
x.add_row(["Time to replacement, extrapolated, high (days)",
        np.around(365*Time_to_repl_GFH_high),
        "-", "-"])
x.add_row(["Time to replacement, conservative (days)",
        np.around(365*Time_to_repl_GFH_c),
        np.around(365*Time_to_repl_E33_c), np.around(365*Time_to_repl_MET_c)])
x.add_row(["Time to replacement, conservative, low (days)",
        np.around(365*Time_to_repl_GFH_c_low),
```

```
        "-", "-"])
x.add_row(["Time to replacement, conservative, high (days)",
        np.around(365*Time_to_repl_GFH_c_high),
        "-", "-"])
x.add_row(["Media Cost, extrapolated ($/year)",
        Cost_per_year_GFH,
        Cost_per_year_E33, Cost_per_year_MET])
x.add_row(["Media Cost, conservative ($/year)",
        Cost_per_year_GFH_c,
        Cost_per_year_E33_c, Cost_per_year_MET_c])
x.add_row(["Media Cost, extrapolated, lower bound ($/year)",
        Cost_per_year_GFH_low,
        "-", "-"])
x.add_row(["Media Cost, extrapolated, upper bound ($/year)",
        Cost_per_year_GFH_high,
        "-", "-"])
x.add_row(["Media Cost, conservative, lower bound ($/year)",
        Cost_per_year_GFH_c_low,
        "-", "-"])
x.add_row(["Media Cost, conservative, upper bound ($/year)",
        Cost_per_year_GFH_c_high,
        "-", "-"])


print(x)

#----------------------------------------------------------------
## Design of System - E33
#----------------------------------------------------------------

HLR = 17/60 # m/min
EBCT = 3 # min
Media_Depth = HLR*EBCT # m
Q = 900/264.172/4# m3/min, per column, of which there are 4
Area = Q/HLR # m
Diameter = m.sqrt(4/m.pi*Area)
Media_Vol = Area*Media_Depth

x1 = PrettyTable()

x1.field_names = ["Parameter", "Value"]

x1.add_row(["HLR (m/min)", np.around(HLR,3)])
x1.add_row(["EBCT (min)", np.around(EBCT,1)])
x1.add_row(["Media Depth (m)", np.around(Media_Depth,3)])
x1.add_row(["Q (m3/min)", np.around(Q,2)])
x1.add_row(["Area (m2)", np.around(Area,2)])
x1.add_row(["Diameter (m)", np.around(Diameter,3)])
```

```
x1.add_row(["Media Volume (m3)", np.around(Media_Vol,2)])

print(x1)

#################################################################
################ OTHER METHODS BELOW, NOT USED #################
#################################################################

## Which models were tested determined from: https://scikit-learn.
#   org/stable/tutorial/machine_learning_map/index.html


# #-----------------------------------------------------------------
# ## Ridge
# #-----------------------------------------------------------------

# ridge = linear_model.Ridge(alpha = 0.5)
# ridge_fit = ridge.fit(X,BV10)
# #print("For ridge regression: ")
# Predictions = ridge.predict(X)
# #print(Predictions)
# #error_stats(BV10, Predictions)

# #-----------------------------------------------------------------
# ## Lasso
# #-----------------------------------------------------------------

# lasso = linear_model.Lasso(alpha = 1)
# lasso_fit = lasso.fit(X,BV10)
# #print("For lasso regression: ")
# Predictions_lasso = lasso.predict(X)
# #print(Predictions_lasso)
# #error_stats(BV10, Predictions_lasso)

# #-----------------------------------------------------------------
# ## Elastic Net
# #-----------------------------------------------------------------

# EN = linear_model.ElasticNet(alpha = .1)
# EN_fit = EN.fit(X,BV10)
# #print("For Elastic Net regression: ")
# Predictions_EN = EN.predict(X)
# #print(Predictions_lasso)
# #error_stats(BV10, Predictions_EN)

# #-----------------------------------------------------------------
# ## Stochastic Gradient Descent
# #-----------------------------------------------------------------
```

```
# SGD = SGDRegressor(loss="huber", penalty="l2", max_iter=500000)
# SGD.fit(X,BV10)
# #print("For Stochastic Gradient Descent: ")
# Predictions_SGD = SGD.predict(X)
# #print(Predictions_lasso)
# #error_stats(BV10, Predictions_SGD)
```

Appendix B - Economic/Alternatives Analysis

B.1 Media Cost

First, find the total volume to be treated per year (y).

$$Q = 900\,gpm \times \frac{60min}{hr} \times \frac{24hr}{d} \times \frac{365d}{y} \times \frac{1m^3}{264.172gal} = 1,790,651.5m^3/y \tag{B.1}$$

The total bed volumes to be treated per year is then the total flow in a year divided by the volume of the reactor, which is given by

$$BV_{annual} = \frac{1,790,651.5\ m^3/y}{10.22\ m^3/BV} = 175,210.5\ \frac{BV}{y} \tag{B.2}$$

Thus, the time to exhaustion (t_E) for a given media is the bed volumes to breakthrough divided by BV_{annual}. For example, for GFH,

$$t_E = \frac{BV_{10,GFH}}{BV_{annual}} = \frac{61,298\ BV}{175,210.5\ BV/y} = 0.35\ y \tag{B.3}$$

And so the cost per year of a given media, in this case GFH (extrapolated), is given by

$$C_{annual} = \frac{\$8,476}{m^3} \times \frac{10.22\ m^3}{BV} \times \frac{1\ BV}{0.77\ y} = \frac{\$112,500}{y} \tag{B.4}$$

The total O&M cost can be found by multiplying C_{annual} by 1.25 [14].

B.2 Capital Cost

The capital cost was estimated using Table 3-1 in [18], which gives costs for a 570 gpm system in 2003 dollars, and adjusting the values in the same way as used to adjust the cost of pH adjustment to 2019 dollars. This analysis assumes cost scales linearly with gpm.

Treatment Vessels

Cost of treatment vessels was calculated as follows.

$$C_V = \$78,000 \times \frac{900\,gpm}{570\,gpm} = \$123,158 \tag{B.5}$$

Which is then $174,432 in 2019 dollars using [19].

Process Piping, Valves, and Accessories

No option is given in Table 3-1 in [18] that describes our system exactly, so a medium cost of $50,000 for a 570 gpm system is assumed; scaling in the same way as used earlier:

$$C_{PPV\&A} = \$50,000 \times \frac{900\,gpm}{570\,gpm} = \$78,947 \tag{B.6}$$

Which is then $111,815 in 2019 dollars using [19].

Instruments and Controls

The instruments and controls were assumed to cost slightly less than the costs given in Table 3-1 in [18] for the automatic system with regeneration and pH control, so $60,000 was chosen. Additionally, because the instrumentation is probably not that different for 560 gpm than it is for 900 gpm, the cost is merely adjusted for inflation, giving $84,979.

Process Equipment Installation and Miscellaneous Installed Items

These items were assumed to cost the same as for the automatic system with regeneration and pH control (including 20% contingency), minus the regeneration wastewater surge tank. So,

$$C_{PE\&MI} = \$112,000 \times \frac{900\,gpm}{570\,gpm} = \$178,105 \tag{B.7}$$

Which is then $252,254 in 2019 dollars using [19].

Final Capital Cost

The final capital cost was calculated assuming everything except the media costs will not be significantly different with one additional tank for redundancy.

$$C_{Capital} = \$624,000 \tag{B.8}$$

Appendix C - Sensitivity Analysis Graphs

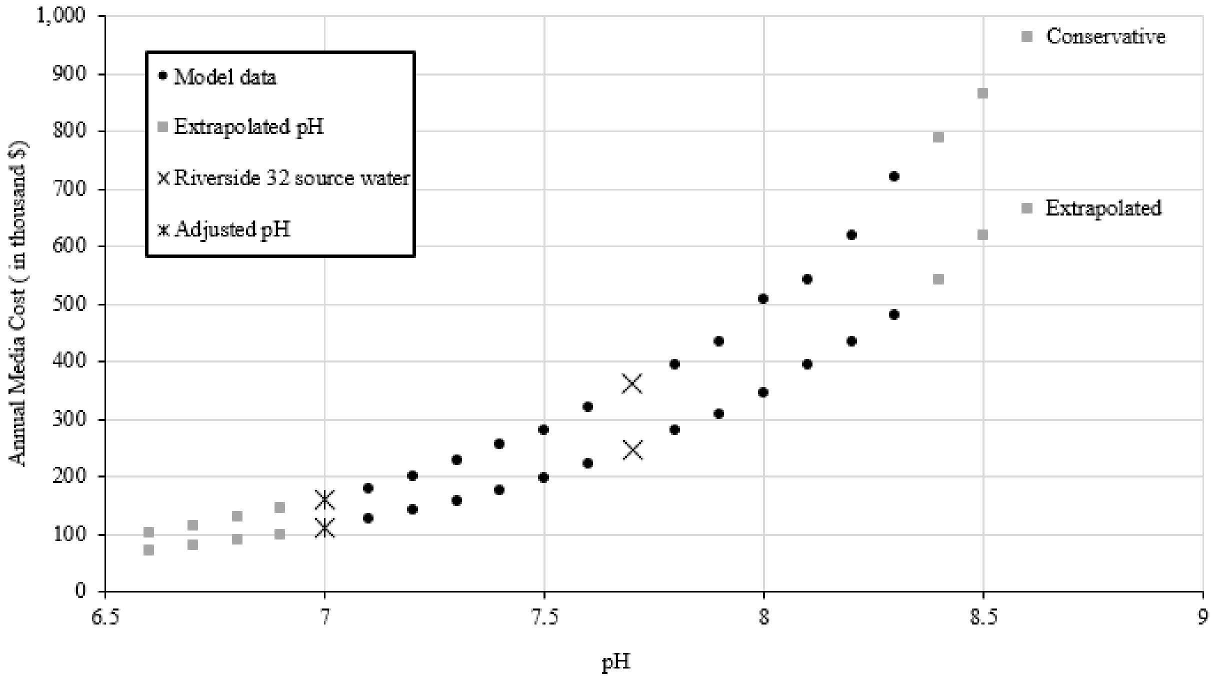

Figure C.1: pH and annual media cost

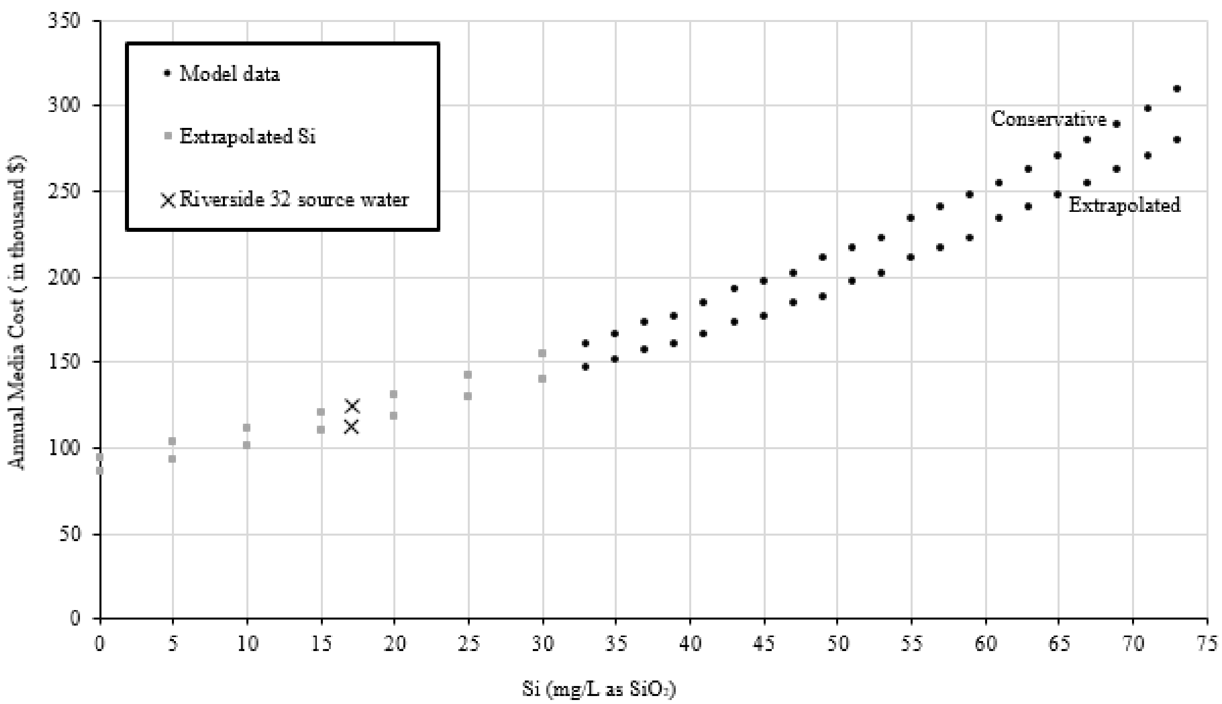

Figure C.2: Silica concentrations and annual media cost

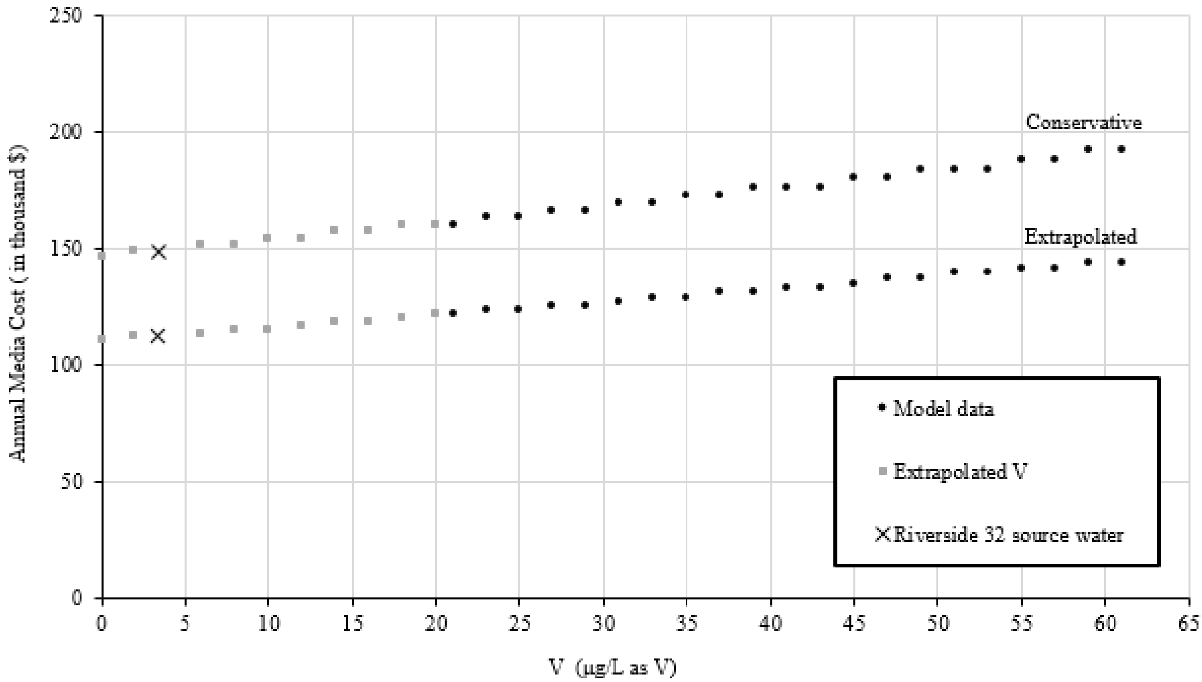

Figure C.3: Vanadium concentrations and annual media cost

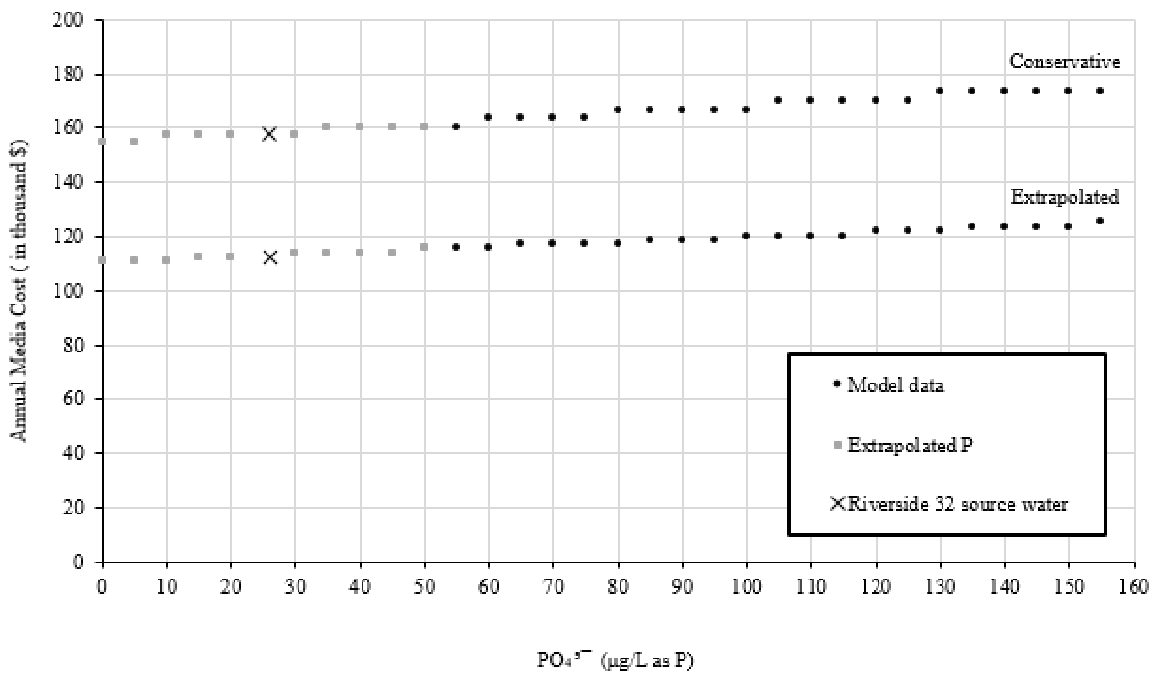

Figure C.4: Phosphorous concentrations and annual media cost

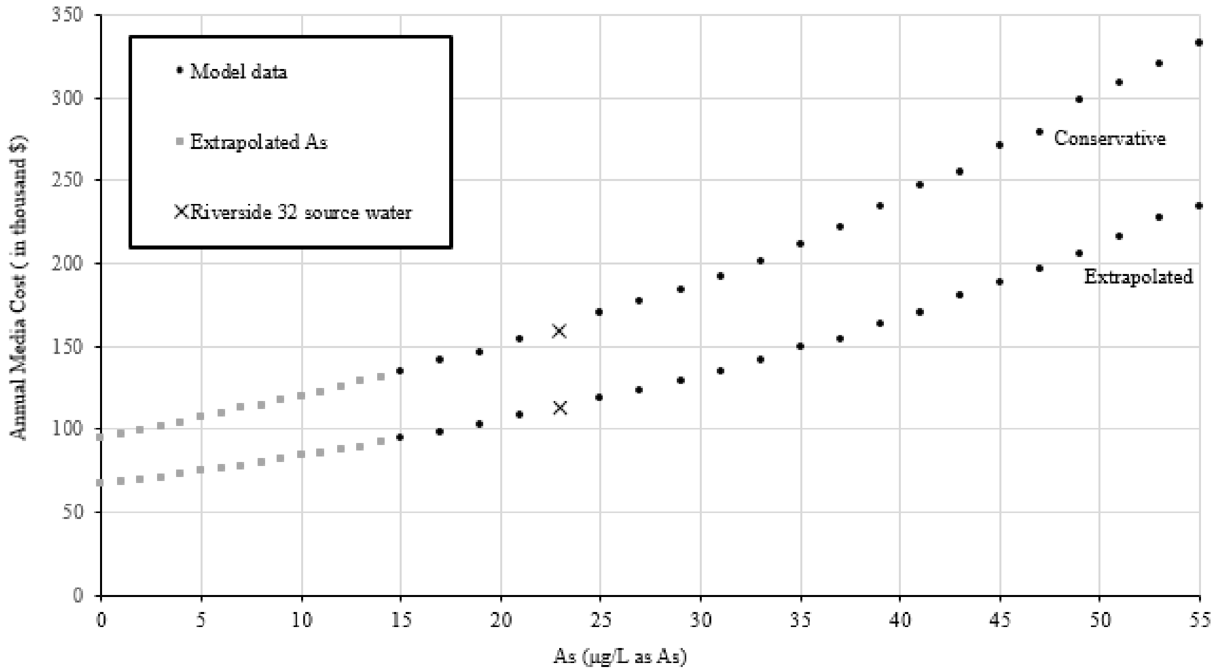

Figure C.5: Arsenic concentrations and annual media cost

Appendix D - Scale Up Example Calculations for E33

Find large column empty bed contact time from Equations 2.6 and **??**:

$$\frac{EBCT_{SC}}{EBCT_{LC}} = \left[\frac{d_{p,SC}}{d_{p,LC}}\right] \tag{D.1}$$

$$EBCT_{LC} = 0.33 min \times \left[\frac{1.16 mm}{0.127 mm}\right] = 3 min \tag{D.2}$$

Find large column media depth given EBCT and hydraulic loading rate using Equation 2.8:

$$L = EBCT \times V = 3 min \times 0.283 \frac{m}{min} = 0.85 m = 85 cm \tag{D.3}$$

Find large column bed volume using Equation 2.13:

$$BV = EBCT \times Q = 3 min \times 3.41 \frac{m^3}{min} = 10.22 m^3 \tag{D.4}$$

Find the large column internal diameter from flow rate and hydraulic loading rate using Equation D.5:

$$A = \left[\frac{\pi}{4}\right] D^2 = \frac{Q}{V} \tag{D.5}$$

$$D = \sqrt{\frac{4 \times Q}{\pi \times V}} = \sqrt{\frac{4 \times 3.41 (\frac{m^3}{min})}{\pi \times 0.283 (\frac{m}{min})}} = 3.91 m \tag{D.6}$$

Appendix E - Columns in Parallel: Characteristics of an Individual Column Example Equation

Empty Bed contact time, hydraulic loading rate, and media depth must remain the same to retain perfectly similar dispersion characteristics. As show by Equation D.8 the area must be divided by the same number as flow rate in order to keep hydraulic loading rate constant.

$$V = \frac{Q}{A} \tag{D.7}$$

To divide the flow rate into 4 parallel columns:

$$Q_p = \frac{Q}{4} = \frac{3.41\frac{m^3}{min}}{4} = 0.853\frac{m^3}{min} \tag{D.8}$$

To determine the individual internal diameters of the 4 parallel columns:

$$\frac{A}{4} = 3.005m^2 \tag{D.9}$$

$$D = \sqrt{\frac{3.005(m^2) \times 4}{\Pi}} = 1.96m \tag{D.10}$$