

Assessment of model adequacy and parameter identifiability for predicting contaminant transport in the Beaverlodge Lake area, Canada

C. E. Hamer¹, B. E. Halbert¹, M. Webster² & J. M. Scharer³

¹*SENES Consultants Limited, Richmond Hill, Ontario, Canada*

²*Cameco Corporation, Saskatoon, Saskatchewan, Canada*

³*University of Waterloo, Waterloo, Ontario, Canada*

Abstract

Contaminant dispersion modeling has been performed on a watershed within the Beaverlodge Lake region, an area which has been affected by past mining activities. The post-decommissioning period was modeled using a proprietary computer code called LAKEVIEW (SENES Consultants Limited) in order to predict future water column and sediment concentrations. The objective of the current study was to ascertain model adequacy and parameter identifiability. Using the Metropolis-Hastings algorithm of the Markov Chain Monte Carlo computational procedure, parameter calibration was performed at several lake locations in the watershed. The calibration process was informed by applying the modified Akaike Information Criterion (AICc) to assess model adequacy. A normalized parameter sensitivity matrix was then used to establish parameter identifiability and to rank the significance of the parameter estimates.

The calibrated model captured the time dependent trends for all key constituents throughout the calibration period. Results of this study indicate that flow through the system is the single most significant parameter for predicting uranium concentrations in the Fulton Creek Watershed while radium-226 concentrations are highly dependent on many parameters influencing transport from the sediments in addition to flow.

Keywords: constituent source modeling, parameter distribution, parameter identifiability, akaike information criterion, model adequacy.



1 Introduction

The Fulton Creek watershed (15.75 km² area), a former uranium tailings management site, is situated East-Northeast of Beaverlodge Lake in Northern Saskatchewan, Canada and contains six water bodies (Fookes Reservoir, Marie Reservoir, the Meadow, Unnamed Reservoir, Minewater Reservoir and Greer Lake) affected by uranium mining (see Figure 1). The main flow of Fulton Creek is from Fookes Reservoir to Marie Reservoir, the Meadow and Greer Lake from where Fulton Creek drains into Fulton Bay on Beaverlodge Lake. A second, smaller branch flows from Minewater Reservoir east to Unnamed Reservoir, which drains into the Meadow. Beaverlodge Lake is a large sized water body of approximately 57 km² surface area and 23 m mean depth. It supports fresh water sport fisheries of lake white fish and lake trout.



Figure 1: Study area watershed and downstream water bodies.

Previous studies identified dissolved radium-226, selenium, uranium, and total dissolved solids (TDS) as constituents of concern. Constituent sources in the watershed include mill tailings deposited underwater in Fookes and Marie reservoirs, and residual mine water slimes, mill tailings and barium-radium-sulphate co-precipitate sludge in sediments of Minewater Reservoir, Unnamed Reservoir, the Meadow and Greer Lake. Tailings deposited in Fookes and Marie reservoirs formed aerally exposed beaches at their points of discharge, which have been covered with waste rock.

2 The LAKEVIEW model

To address both current and future conditions, watershed dispersion modeling was carried out to predict levels of radium-226, selenium, uranium, and TDS in water and sediments. These constituents were modeled using a proprietary

computer code called LAKEVIEW that was developed by SENES Consultants Limited specifically for assessing mine waste effects on the environment. The code may be run in either a deterministic or a probabilistic manner and has been applied on numerous occasions in Canada, including in Northern Saskatchewan, to assess the impact of uranium and other mining operations.

LAKEVIEW is essentially a three-dimensional water quality transport and constituent speciation model for lakes and rivers in a watershed. Important processes incorporated into the LAKEVIEW model include horizontal (lateral) and vertical transport of dissolved species, chemical and biochemical reactions, settling of particulate matter, and sediment exchange. Transport between segments (water column–sediment, for example) is computed by solving ordinary differential equations while constituent speciation within a segment is computed by thermodynamic stability criteria for equilibrium processes and kinetic considerations [1, 2]. Model output for lakes includes surface water, hypolimnion and sediment porewater concentrations of dissolved constituents, constituents adsorbed to suspended matter and solid phase concentrations in the sediment. In accordance with the modified Tessier's extraction test [3] constituents in the sediment are partitioned as being adsorbed onto calcite, quartz and organics (ion-exchangeable), complexed with iron hydroxide (reducible), embedded in sulphide matrix (oxidizable) as well as being present as solid solution, pure mineral and recalcitrant (un-extractable) fractions.

3 Parameter discrimination

The development of a robust model that adequately describes the system often requires an examination of the parameter values. In general, the greater the number of parameters, the better the overall model fit. However, this “better fit” may be a result of spurious data variability rather than underlying scientific reasons; an example of this scenario is illustrated in Figure 2. Using observed annual averaged concentrations of dissolved uranium, radium-226, selenium and TDS in Fookes Reservoir (the largest water body in the watershed), the probability density function for the mass transfer coefficient between the water column and the sediment porewater (K_1) were generated by the Gibbs sampler version of the Metropolis-Hastings algorithm [4].

Although the distributions shown in Figure 2 differ, they overlap enough to suggest that the transport process characteristics for all dissolved constituents in Fookes Reservoir could be the same.

The possibility of replacement of the four K_1 distributions by a single one was tested by the Akaike information criterion adapted for small sample size (AICc) by Hurvich and Tsai [5]:

$$AICc = n \ln\left(\frac{SSE}{n}\right) + \frac{2np}{n - p - 1} \quad (1)$$

where: n = number of observations; SSE = sum of squared differences between observations and predictions; p = number of parameters.



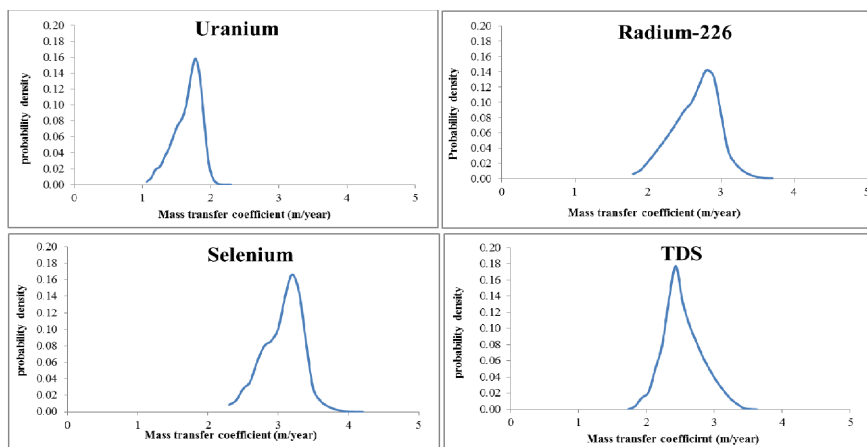


Figure 2: Probability density functions of the sediment water column mass transport parameter (K_1) in Fookes Reservoir as determined from uranium, radium-226, selenium, and total dissolved solids (TDS) monitoring data.

Altogether three parameters were investigated: the mass transfer coefficient between surface water and sediment porewater (K_1), the adsorption coefficient in the sediment for a given constituent (K_{D_SED}) and the adsorption coefficient of settling solids in the water column for a given constituent (K_{D_WAT}). The objective was to establish whether the individual “optimum” calibrated values can be replaced by a single average value that may be applied for all constituents in a lake (K_1 , for example) or for each constituent in all lakes (K_{D_SED} , for example). The procedure was as follows. The average parameter value of the distributions ($\theta_{avg,i}$) was used as the baseline value. The calibrated “optimum” value was treated as some increment (either positive or negative) from the model average [6]:

$$\theta_{i,opt} = \theta_{avg,i} + \sigma_i \quad (2)$$

Thus, the “optimum” configuration contained an extra term (σ_i) which results from including the additional model parameters. After re-running the model, the AICc scores were calculated according to eqn (1). Due to the logarithmic transformation of the SSE-s, the AICc scores can be either positive or negative. In either case, the best model corresponds to the lowest AICc score. The results for the Fookes Reservoir are summarized in Figure 3.

For each parameter, the minimum score corresponded to taking the average parameter value rather than the optimum value for either the water body (in case of K_1) or the constituent (in case of K_{D_SED} and K_{D_WAT}) in question. The penalty paid for introducing an extra parameter was more significant than the resulting reduction in the residual error between measurements and observations. Consequently, in all further calibration attempts average parameter values were applied.

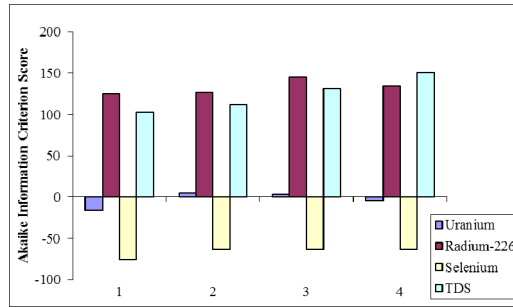


Figure 3: The Akaike Information Criterion (AICc) scores of model averaged and optimized parameters in Fookes Reservoir. 1) all parameters averaged (baseline), 2) sediment-water mass transfer coefficient (K_l) optimized, 3) contaminant adsorption coefficient to sediment solids (K_{D_SED}) optimized, and 4) contaminant adsorption coefficient to total suspended solids (TSS) (K_{D_WAT}) optimized.

4 Model calibration and long term predictions

As noted in the previous section, parameters estimated through calibration included: K_l , K_{D_SED} , and K_{D_WAT} . The other parameters were estimated based on a combination of applicable site specific data and values cited in literature.

In the model parameter calibration, parameters in each lake were sampled one at a time following the Metropolis-Hastings procedure using predicted and observed surface water concentrations as criterion. As a result of the Akaike Information Criterion analysis, discussed in Section 3, individual estimates of the calibrated parameters (K_{D_WAT} and K_{D_SED}) were then averaged on a “per constituent” basis across all water bodies while the sediment-to-water column mass transfer coefficient (K_l) was averaged on a “per lake” basis.

Detailed LAKEVIEW model calibration for radium-226, selenium, uranium and TDS was carried out using the entire set of available water quality data collected over an 11-year period (1998 to 2009).

The Metropolis-Hastings algorithm was applied as follows. First the parameter being investigated (K_l , K_{D_WAT} and K_{D_SED}) was assigned a symmetric, triangular prior (proposal) distribution with a minimum, mean and maximum parameter value. The triangular parameter spaces were sampled by random draw. The distribution of accepted parameter values was used to estimate the mean parameter value for the study lake.

The resulting calibrated K_l values are shown in Table 1. In general, there is fairly low variability seen between calibrated K_l values for similar water bodies: the small to mid-sized lakes found in the Fulton Creek Watershed range from 0.95 to 2.4 m/a. The higher K_l value seen in Fookes Reservoir may be due to the much larger volume of the water body which typically results in higher currents.

The K_{D_WAT} and K_{D_SED} values obtained through the calibration procedure are presented in Table 2. The high K_{D_WAT} values for uranium and selenium indicate

Table 1: Calibrated K_1 values for the Fulton Creek Watershed.

Segment	K_1 (m/a)
Fookes Reservoir	2.4
Marie Reservoir	1.15
The Meadow	1.15
Greer Lake	0.95

Table 2: Calibrated K_{D_WAT} and K_{D_SED} values for the Fulton Creek Watershed.

Constituent	K_{D_WAT} (m^3/kg)	K_{D_SED} (m^3/kg)
Radium-226	2.50	3.77
Selenium	6.37	4.82
TDS	0.1	0.1
Uranium	6.34	2.29

significant adsorption to iron mineral surfaces and organic matter in all lakes. The low K_{D_WAT} and K_{D_SED} values for TDS are due to the limited nature of TDS-sediment interactions.

The K_{D_SED} parameter combined with data on the leachable mass inventory present in the sediments determines the resulting porewater concentration of each constituent. The difference between the total leachable mass and the fraction of each constituent distributed in the porewater is treated as being adsorbed to organic and inorganic solids, if adsorption control is the limiting process, or as a chemical precipitate, if chemical solubility is the limiting process. The limiting process is checked within the model each time step based on mass balance calculations on the leachable constituent inventories.

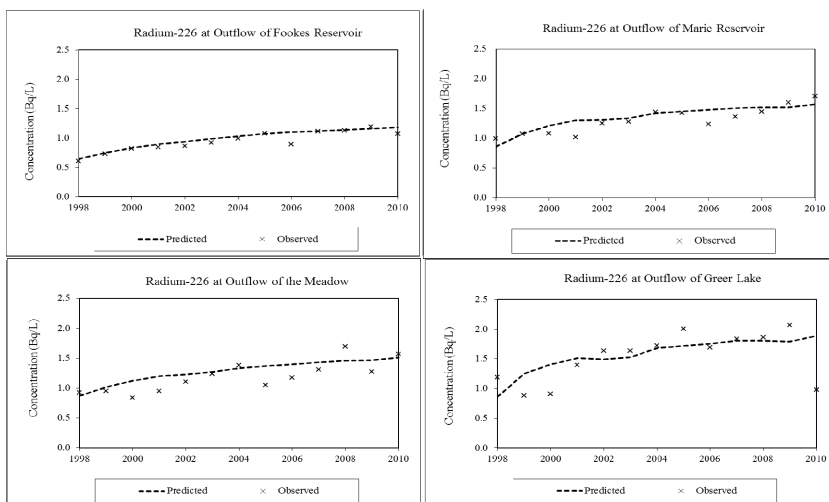


Figure 4: Measured and modeled radium-226 levels in the Fulton Creek Watershed.

Applying the calibrated values, the resulting comparison between measured and predicted water levels for radium-226, selenium, uranium, and TDS are shown in Figures 4 through 7 for the Fulton Creek Watershed.

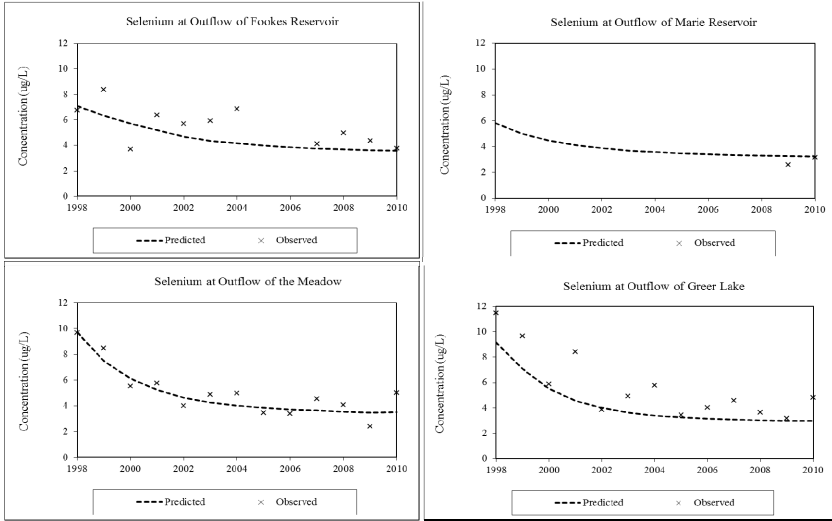


Figure 5: Measured and modeled selenium levels in the Fulton Creek Watershed.

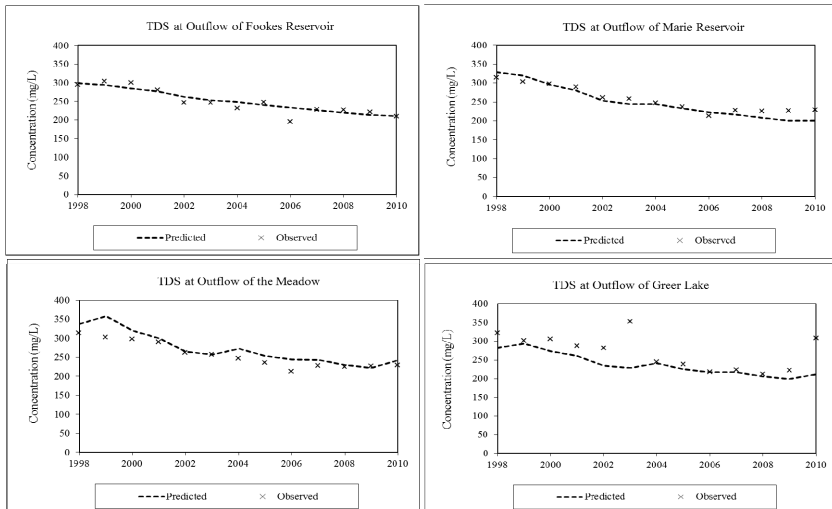


Figure 6: Measured and modeled total dissolved solids (TDS) levels in the Fulton Creek Watershed.



The calibrated LAKEVIEW model was run deterministically for 300 years, starting in 1998, to assess future trends in radium-226, selenium, uranium, and TDS levels in environmental media in the Fulton Creek Watershed. Only the first 150 years of modeled concentrations are presented in Figure 8 to allow for greater resolution of predictions within the initial time period.

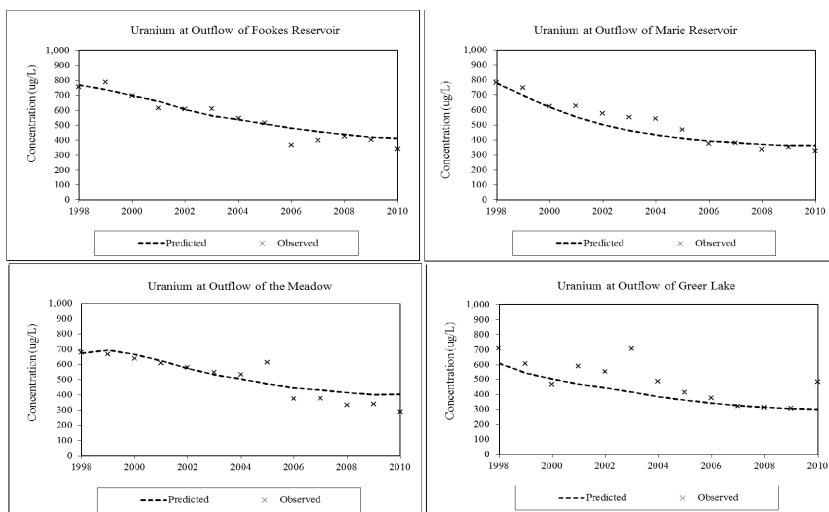


Figure 7: Measured and modeled uranium levels in the Fulton Creek Watershed.

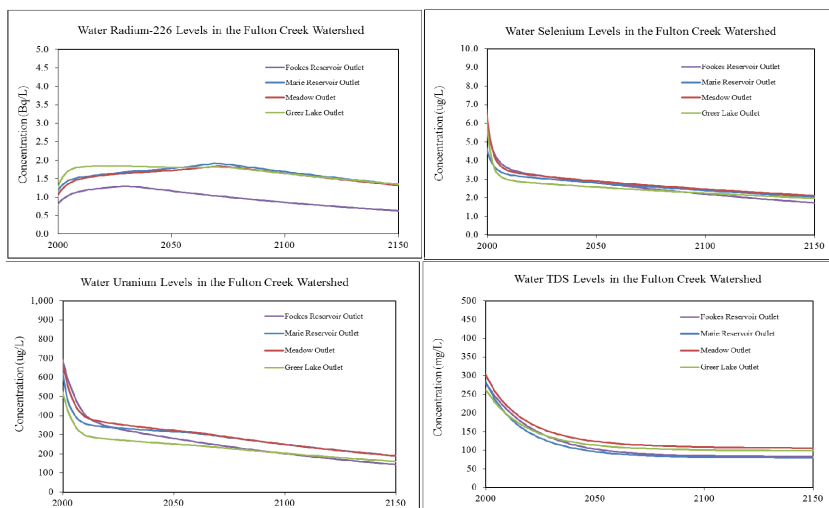


Figure 8: Long term water quality predictions in the Fulton Creek Watershed.

5 Parameter identifiability

After model calibration, an analysis of the model parameters' significance and reliability was carried out. Since the parameter values were derived from observations, their reliability depends on the quality of the database. In the present case, a parametric analysis known as structural identifiability was performed. Though structural identifiability analysis is based on the assumption of accurate model structure, it is useful to establish the relative physical significance of the parameters. The computational procedure proposed by Yao *et al.* [7] was employed. This method uses matrix orthogonalization to reduce the correlation in order to identify a set of estimable parameters. It also ranks the parameters from the most identifiable to the least.

The model parameters selected for analysis included the external contaminant load (load), the lake outflow (flow), total suspended solids concentration (TSS), TSS settling velocity (vel), contaminant adsorption coefficient to TSS in the water column (K_{D_WAT}), chemical mass transfer coefficient between the water column and sediment porewater (K_i), thickness of the sediment exchange zone (Xzone), and contaminant adsorption coefficient to sediment solids (K_{D_SED}). The external load and the outflow were parameterized by introducing a factor "f" to give "f x load" and "f x flow", respectively, as the appropriate functions. The expected value of "f" is unity. The total suspended solids concentration (TSS) and the outflow were based on measurements while the load, K_i , K_{D_SED} and K_{D_WAT} were calibrated values. The nominal thickness of the exchange zone was 5 cm as recommended by Thibodeaux and Mackay [8] for Brownian diffusion, bio-diffusion and bio-availability. Parameter identifiability studies were carried out for uranium and radium-226 in the entire watershed (Fookes Reservoir, Marie Reservoir, the Meadow, and Greer Lake).

First, a sensitivity score matrix was derived by evaluating the partial differentials of the predicted dependent variables (concentrations of a given constituent) with respect to the parameters (θ -s) as shown in eqn (3):

$$Z = \begin{bmatrix} \left. \frac{\delta Y_1}{\delta \theta_1} \right|_{t=t_1} & \dots & \left. \frac{\delta Y_1}{\delta \theta_k} \right|_{t=t_1} \\ \vdots & \ddots & \vdots \\ \left. \frac{\delta Y_n}{\delta \theta_1} \right|_{t=t_1} & \dots & \left. \frac{\delta Y_n}{\delta \theta_k} \right|_{t=t_1} \\ \left. \frac{\delta Y_1}{\delta \theta_1} \right|_{t=t_2} & \dots & \left. \frac{\delta Y_1}{\delta \theta_k} \right|_{t=t_2} \\ \vdots & \ddots & \vdots \\ \left. \frac{\delta Y_n}{\delta \theta_1} \right|_{t=t_m} & \dots & \left. \frac{\delta Y_n}{\delta \theta_k} \right|_{t=t_m} \end{bmatrix} \quad (3)$$

where:

Y_i = the dependent variable (concentration of a contaminant in the surface waters); θ_i = parameter being investigated; n = the total number of water bodies; k = total number of parameters.

The partial differentiation was performed numerically using the finite difference method. The partials in this analysis comprised the 5th year, 25th year and 100th year values obtained from 100-year simulations. Thus, the matrix Z had 12 rows (4 lakes x 3 time periods) and 9 columns (9 parameters).

The matrix operations, proposed by Yao *et al.* [7] can be summarized as follows:

- 1) Select column X_L ($L=1$) of Z with the highest total sum of squares (SS);
- 2) Calculate $Z_L = X_L (X_L^T X_L)^{-1} X_L^T Z$;
- 3) Calculate the residual $R_L = Z_L - Z$;
- 4) Select column X_{L+1} in Z with the next highest SS in corresponding R_L ;
- 5) Augment X_L with new column X_{L+1} ;
- 6) Repeat steps 2 to 5 until SS in columns of R are less than the cut-off.

The logarithm of the normalized (maximum = 100) scores are presented in Figure 9 for uranium and radium-226 in the Fulton Creek Watershed. The scores of all 9 parameters were higher than the designated cut-off value of 1%.

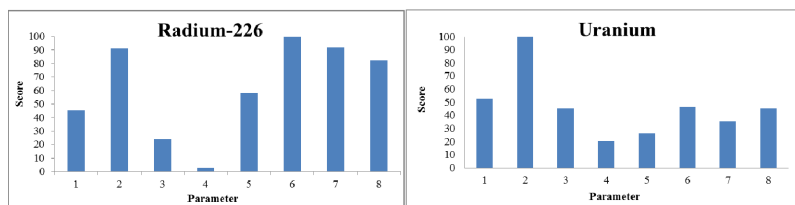


Figure 9: Parameter identifiability scores for radium-226 and uranium in the Fulton Creek watershed. 1) external contaminant load (load), 2) outflow (flow), 3) total suspended solids (TSS), 4) TSS settling velocity (vel), 5) contaminant adsorption coefficient to TSS (K_{D_WAT}), 6) sediment–water mass transfer coefficient (K_I), 7) sediment exchange zone thickness (Xzone), and 8) contaminant adsorption coefficient to sediment solids (K_{D_SED}).

It is evident that the outflow, or hydraulic residence time, is a very significant parameter for predicting water quality for both uranium and radium-226. The importance of flow ranked 1st for uranium and 2nd for radium-226. Sediment mass transfer (K_I), sediment chemistry (K_{D_SED}), and sediment morphology (Xzone) are also significant parameters for predicting radium-226 implying that the sediments are the primary contributors of dissolved radium-226 in the water column. This conclusion is supported by observed Ra-226/Ca carbonate co-precipitates in Fookes Reservoir and Marie Reservoir and the presence of Ra-

226/Ba sulphate (barite) co-precipitates in the Meadow and Greer Lake. Significant radium-226 deposits in the sediments are an inventory remaining from the operating period when Fookes Reservoir, Marie Reservoir, the Meadow and Minewater Reservoir acted as the Tailings Management Area. Unlike in the case of radium-226, the importance of sediment transport of uranium is less significant than other factors in the Fulton Creek Watershed. External loads are seen to be fairly important contributions for both constituents. The principal sources of external contaminant loads to the modeled water bodies consist of runoff from the tailings delta along the western shore of the Fookes Reservoir as well as flow into the Meadow from contaminated areas to the west. The contaminant removal flux resulting from adsorption to settling solids (TSS) had minor effect on both uranium and radium-226 concentrations in the surface waters.

6 Conclusions

Applying Akaike's information criterion, model complexity was simplified by applying model averaged values. The comparison between predicted and observed radium-226, selenium, uranium, and TDS levels in the water column of each modeled water body in the Fulton Creek Watershed showed that the model is able to accurately predict the behavior of the key parameters over the calibration period (1998 to 2009).

The resulting model was then used to predict radium-226, selenium, uranium, and TDS levels in the waters of the Fulton Creek Watershed into the future. The model predictions indicate that selenium and uranium levels are expected to decline steadily over the entire 150 year period while TDS levels are expected to exhibit a decreasing trend for the first 50 years before leveling off. Radium-226 concentrations, however, are driven by a solubility controlled solid phase radium-226 release from the sediments for the first 25 to 75 years followed by a gradual decline over the remainder of the modeled period controlled by surface sorption dependent processes. Since model identifiability analysis indicated that transport from sediment is a significant source of contaminants to the Fulton Creek Watershed, remedial measures which either reduce or eliminate the upward contaminant flux are expected to have the greatest impact.

References

- [1] Stumm, W. and Morgan, J.J. (1996). *Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters*, 3rd ed. Wiley and Sons Inc., USA.
- [2] Scharer, J.M., Nicholson, R.V., Halbert, B.E. and Snodgrass, W.J. (1994). A computer program to assess acid generation in pyritic tailings. *ACS Symposium Ser.* **550**, pp. 132-152.
- [3] Mossop, K.F. and Davidson, C.M. (2003). Comparison of original and modified BCR sequential extraction procedures for the fractionation of copper, iron, lead, manganese and zinc in soils and sediments. *Annal. Chim. Acta*, **478**, pp. 111-118.



- [4] Jitjareonchai J.J., Reilly P.M., Duever T.A. and Chambers, D.B. (2006). Parameter Estimation in the Error-in-Variables Models Using the Gibbs Sampler. *Canadian Journal of Chemical Engineering*. 84: 125-138.
- [5] Hurvich C.M. and Tsai C.L. (1989). Regression and time series model selection in small samples. *Biometrika*, *Biometrika* 76:297-307.
- [6] Cleaskens, G. and Hjort N.L. (2008). *Model selection and model averaging*. Cambridge University Press. London, UK.
- [7] Yao K.Z., Shaw B.M., Kou B., McAuley K.B. and Bacon D.W. (2003). Modeling Ethylene/Butene Copolymerization with Multi-site Catalysts: Parameter Estimability and Experimental Design, *Polymer Reaction Engineering*, 11: 563-588.
- [8] Thibodeaux, L. and Mackay, D. (2007). The importance of chemical mass transport coefficients in environmental and geochemical models. *Sentac Globe: Fate and Exposure Modeling*. p 29-31.

