LONG-TERM SALINITY PREDICTION WITH UNCERTAINTY ANALYSIS: APPPLICATION FOR THE COLORADO RIVER NEAR GLENWOOD SPRINGS, COLORADO

James Prairie, Civil Engineer, U.S. Bureau of Reclamation, Upper Colorado Regional Office, Salt Lake City, Utah; Balaji Rajagopalan, Associate Professor, Water Resource Division, Civil, Architectural, and Environmental Engineering Department, University of Colorado, Boulder, Colorado; Terry Fulp, Operations Research Analyst, U.S. Bureau of Reclamation, Lower Colorado Regional Office, Boulder City, NV

Abstract: Salinity on the Colorado River has been predicted by the U.S. Bureau of Reclamation using stochastic flows generated by the Index Sequential Method applied to historic flows, together with a natural salt model based on a weighted least square linear regression of the historic flow and salt data in the Colorado River Simulation System (CRSS). This approach limits flow predictions to historic flows. Recent attempts to re-calibrate the CRSS model using more recent data have found the model overpredicts historic salt mass over the period 1970-1995. As a result, a study has been done with the goal of improving the modeling (data, model, methodologies, etc.) used for projecting future salt mitigation needs for the Colorado River basin. Under this study, new techniques to generate stochastic flow data and a new salt model have been developed that improve the ability to predict salinity in the future. A modified K-nearest (KNN) neighbor stochastic natural flow model was developed to generate synthetic flows that maintain the statistical characteristics of the historic flows, but allow generation of possible future flows that are not reflected in the historic data. The new natural salt model utilizes a local linear nonparametric regression with KNN residual resampling. The improved salt model adds two new features: (1) it incorporates nonlinear features of the regression between flow and salt and (2) instead of simply generating salt mass directly from the regression, the improved salt model generates salt mass from the regression then perturbs the salt mass with a KNN technique that resamples a residual in the neighborhood of the salt mass from the regression. The residual resampling allows the model to incorporate the uncertainty of the regression when the Monte Carlo simulations are used to generate historic flow, salt mass, and concentration with CRSS.

A case study incorporates the results generated from the new flow and salt models into a submodel of CRSS consisting of the Upper Colorado mainstem down to the USGS stream gauge 09072500 (Colorado River near Glenwood Springs, CO). The submodel was used to compare multiple statistics including the probability density functions (PDF) of the synthetic natural flow and salt and the historic flow and salt. Additionally, we generated statistics describing the number of times the model exceeded a postulated salt mass and salt concentration standard. The submodel results for the time period 1941 to 1995 showed that the modified KNN flow and salt models were able to preserve the annual flow and salt mass PDFs and other statistical characteristics of the historical data for this period, and improved the accuracy of the salinity modeling.

INTRODUCTION

Salt concentration is important in the Colorado River Basin because it is a primary water quality parameter regulated by federal water quality standards. Federal water quality standards were set as a result of the Federal Water Pollution Control Act Amendments of 1972. The Amendments, interpreted by the Environmental Protection Agency, require numerical standards for salinity levels in the Colorado River basin. Modeling studies that predict long-term salinity levels under various operational scenarios and salinity control projects facilitate planning and operating the river to meet water quality standards. Operational and planning policies in the Colorado River basin are complicated because many laws, statutes, and court decrees govern the basin. To capture these complex policies in a modeling tool, the U.S. Bureau of Reclamation (USBR) developed the Colorado River Simulation System (CRSS) (USBR, 1987), a series of computer models for the entire basin. Developed in the early 1970's, one of the purposes of CRSS is to conduct long-term operational and planning studies that allow managers to understand the effects of future development on salinity throughout the Colorado River Basin. The original CRSS was replaced in 1996 using the RiverWare modeling framework (Zagona et al., 2001). The newer model uses the same methodologies and has been shown to reproduce the results of the original model (Fulp et al., 1999).

CRSS includes a simulation model of the entire Colorado River system. It also includes a stochastic natural flow model to generate future stochastic flows and a salt regression model that estimates natural salinity associated with natural flows. The stochastic natural flow model used for many years is the index sequential method (ISM). ISM is limited to generating flows and flow sequences that have occurred historically, which limits the ability of planning studies to consider flows that are statistically possible but have not occurred. The salt regression model consists of a series of 12 monthly regressions of natural salt mass as a function of natural flow that were developed by the USGS (Mueller and Osen, 1988). These two models provide the natural flow and natural salt input data for the CRSS simulation model.

To ensure the CRSS simulation model is calibrated, it is periodically used to simulate a historic period and the results are compared to the observed historic record. Recently, the historic runs have indicated that the simulation system overpredicts salt throughout the basin. The existing CRSS over-predicts the historical salt mass at USGS stream gauge 09072500 (Colorado River near Glenwood Springs, CO), form 1970 to 1990, by an average 140,000 tons/year. The over prediction could result from;

- salinity pickup from agriculture being too high,
- natural salt loading being too high.

From 1970 to 1990, the historic salt mass in the river passing gauge 09072500 averaged 570,300 tons/year. The regressions developed by the USGS estimates average annual natural salt of 583,000 tons/year. For CRSS to simulate the historic salt mass, the human-induced salinity pickup sources would need to *remove* salt from the river. Current estimates, as reflected in CRSS, are that human-induced sources contribute 137,000 tons/year from agriculture salinity pickup and exports remove an average 44,000 tons/year. The estimate for salinity pickup by agriculture is developed from an extensive study that quantified estimates of natural and human-induced salt (Iorns et al., 1965). The report estimates that, in 1957, natural sources contributed 516,200 tons/year, and human-induced sources contributed 138,881 tons/year from agriculture salinity pickup and removed 15,881 ton/year by exports above Glenwood Springs. These values were adjusted for current basin conditions then input in the CRSS simulation model.

Using these numbers, if human-induced sources contributed no salt above gauge 09072500, the existing CRSS would still overpredict salt mass. The Iorns report indicates that the human-induced sources of salinity are not removing salt, but are adding significant amounts. These findings point to an overestimation of natural salt by the USGS model. To correct or refine the USGS model would require a reanalysis of the detailed data on which the regressions were based. However, this data is not readily available. Therefore, we propose a new technique to relate natural flow to natural salt that is more accurate and for which the uncertainty can be quantified.

To improve CRSS performance; newly developed models replaced both the stochastic natural flow model and the natural salt regression model. We bring these new models together in a simulation model to demonstrate that they more accurately reflect variability by improving flow variability prediction and quantification of risks, fully reflecting future salt variability by improving the relationship between natural flow and salt, and reproducing historical results. These developments result in an improved analysis of future salinity and quantified risks.

We performed our study on the upper mainstem of the Colorado River, at USGS stream gauge 09072500 Colorado River near Glenwood Springs, CO. We chose the upper Colorado River mainstem because this part of the basin contributed more than 51 percent of the total annual historic salt load seen in the outflow from Lake Powell from 1941 to 1990 (Prairie and Fulp, 2000). Gauge 09072500 also exhibited an overprediction of historic salt mass in the calibration runs from 1970 to 1990 by 20 percent. One of our primary goals was to develop techniques to correct the overprediction. While investigating the cause and means to correct the overprediction, we ensured our solutions were portable and easily implemented at the 28 remaining gauges throughout the basin.

EXISTING CRSS

CRSS is used to simulate future periods and model proposed development and changing operational and planning policy. It is also used to simulate flow and salt over a historic period to verify that the model is calibrated. The model uses different input data for different simulated time periods.

Stochastic Planning Runs: Long-term operational and planning studies are conducted using CRSS. For these runs, the CRSS simulation model uses natural flows generated by the ISM. The ISM resamples the 90 years of historic calculated natural flows, generating 90 individual time series sequences of *synthetic natural flows*. The CRSS simulation model uses the 90 synthetic natural flow time series sequences, or traces, to produce 90 simulated results, which are then used to generate statistical probabilities of various events.

The USBR has used the index sequential technique to generate synthetic hydrology for the CRSS since the inception of the CRSS in the 1970’s. Various studies have found that the ISM generates “statistically faithful” synthetic streamflow sequences (Kendall and Dracup, 1991; Ouarda et al., 1997). The Colorado River System is well suited to using the index sequential modeling system because of the extensive historic time series in the basin (water year 1906 to 1995).

The simulation model also requires the *synthetic natural salt mass* associated with each synthetic natural flow sequence. A series of 12 regressions, developed by the USGS, are used to compute the associated synthetic natural salt mass for a given synthetic natural flow. Thus, 90 individual synthetic natural salt traces are found using the regressions, one for each synthetic natural flow trace.

Figure 1 depicts the CRSS simulation model's water and salt balance with a line diagram of the basin above USGS stream gauge 09072500 when the simulation model simulates future flow and salt mass. The inputs are at the top of the diagram: synthetic natural flow, from the stochastic natural flow model and associated synthetic natural salt mass from the USGS salt model. The model routes these inputs through the river reach above USGS gauge 09072500, where the projected future monthly depletions from agriculture, exports, municipal and industrial uses are removed from the river reach. Projected salt mass is added with agricultural returns and removed with exports.

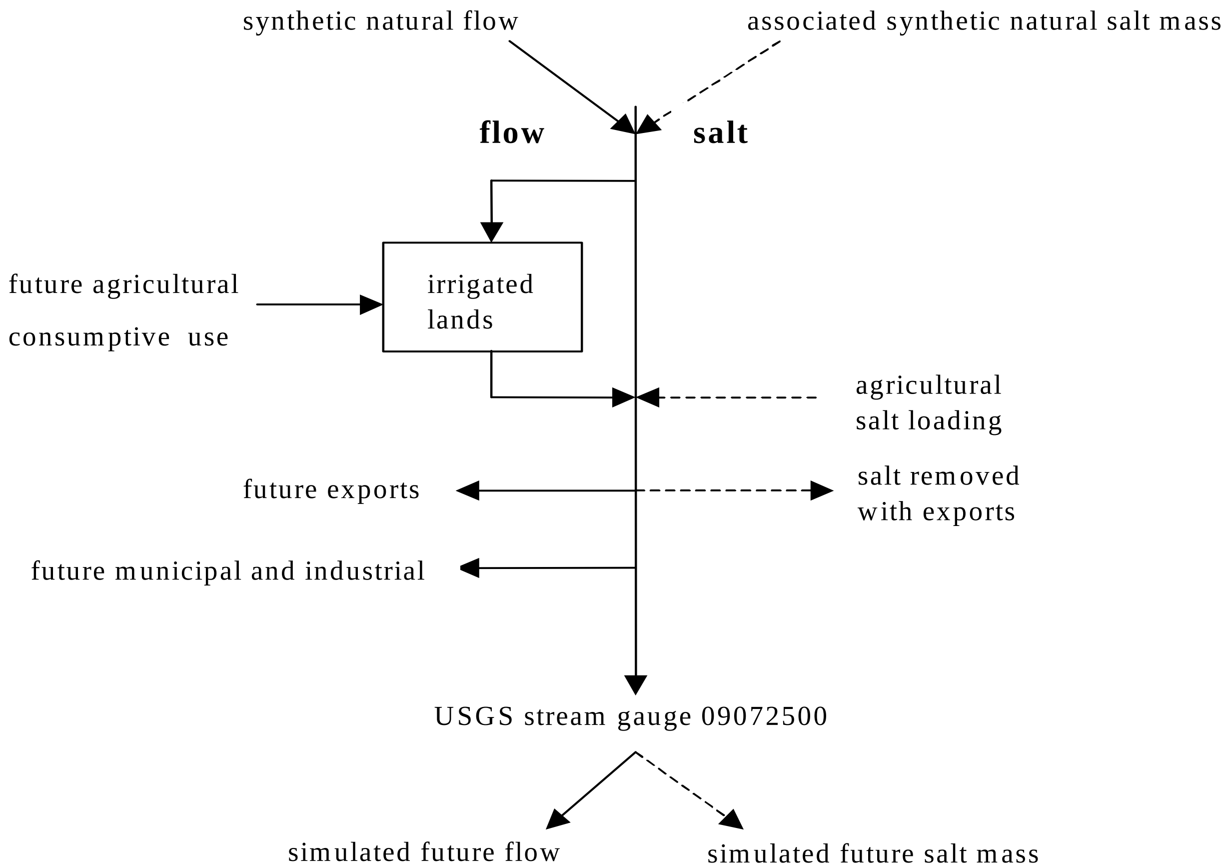


Figure 1: Line diagram of the CRSS simulation model and data for stochastic planning runs

The left side of the diagram shows how simulated future flow is related to synthetic natural flow. The right side shows how simulated future salt mass is related to synthetic natural salt mass. The simulation model performs Monte Carlo simulations by running each synthetic natural flow and associated salt time series. The existing CRSS has 90 synthetic natural flow and associated salt time series that are each run through the simulation model, one at a time, calculating 90 simulated future flow and associated salt time series. Together, the 90 simulations of simulated future flow, salt, and concentration can be used to approximate the PDF for the predicted future flow, salt, and concentration.

Historical Verification: To verify that the stochastic planning runs simulate the observed historic PDFs traces of simulated flows and associated salt mass, the CRSS simulation model was populated with historic monthly depletions, instead of projected future depletions. Then, PDFs of the observed historic flow and salt mass were compared to the simulated historic flow and salt mass from the CRSS simulation model. If the simulation statistics preserved the observed historic flow and salt mass statistics, the CRSS is deemed verified.

We performed stochastic planning runs with both historic depletions to verify the simulation runs and projected future depletions, to compare the existing CRSS, explained previously, and the modified CRSS, explained next.

MODIFIED CRSS

To address the limitations of the existing CRSS modeling system, we developed a modified stochastic natural flow model and statistical nonparametric natural salt model. A description of these models follows. We used both models to generate data for the CRSS simulation model. We compared the results from the simulation model to results from simulations using the ISM natural flow and USGS natural salt regression model described previously.

Modified Stochastic Natural Flow Model: A drawback of ISM is that it cannot generate synthetic sequences that did not occur in the past record. An alternate nonparametric method, K-NN, eliminates this drawback. Traditional K-NN methods resample values from the historic time series, one at a time (Lall, 1995). Because samples are not taken as blocks, as in the ISM, this technique is able to produce time series sequences that did not occur in the historic data. However, values are limited to those in the data set. We further developed the traditional K-NN technique to provide the ability to create values not seen in the historic record. The modified K-NN method developed a regression relationship between successive months' flows and saved the residuals from the regression. The succeeding month's flow was first calculated from the appropriate regression. Then, the k-nearest flows to the flow from the regression were found. A residual from the k-nearest flows was resampled and added to the flow from the regression to produce a new value. This scheme allowed the K-NN method to perturb the historic data within its representative neighborhood and allowed extrapolation beyond the sample, while maintaining the residuals of the data. Like the ISM, the modified K-NN has the advantage of nonparametric models; there is no need to transform the data to fit an assumed probability density function. It has the further advantage of being able to generate synthetic time series containing numbers and sequences that have not occurred in the past, but are "statistically faithful" to the original time series.

Statistical Nonparametric Natural Salt Model: To replace the USGS model, a statistical nonparametric natural salt model is developed that computes a natural salt mass given a natural flow, either historic or generated by the K-NN flow model. The computation is achieved using a nonparametric local regression fit to a scatter plot of calculated natural flow versus calculated natural salt, both from 1941 to 1995. The calculated natural flow and salt are calculated from historic gauged data and salt load data from the CRSS simulation model as follows:

calculated natural flow = observed historic flow
+ agricultural consumptive use
+ exports
+ municipal and industrial uses
± effects of reservoir regulation

calculated natural salt = observed historic salt
+ salt with water exported out of the basin
- salinity pickup from agriculture (values based on CRSS)

The value found from the look up is perturbed with a local residual chosen by a K-NN technique that resamples a residual from the nonparametric regression.

An analysis of the results shows that the annual natural salt PDF is preserved best by an annual regression model. Unfortunately, CRSS requires natural flow and salt data be entered at a monthly time step to accommodate the operational rule set in CRSS. Work to move the operational ruleset to an annual time step is being considered. When the work is complete, the natural salt mass PDF would be preserved better with an annual regression salt model, avoiding the summation of monthly salt mass values.

Simulation Model: To test the modified stochastic nonparametric natural flow model, we used a segment of the CRSS simulation model that includes USGS gauge 09072500 (Colorado River Basin near Glenwood Springs, CO). The modified stochastic natural flow model and statistical nonparametric natural salt model generated inputs for the CRSS simulation model. The inputs included calculated or synthetic natural flow and associated natural salt mass. Additionally, human-induced depletions and salt loading were entered in the CRSS simulation model, as explained in Section " EXISTING CRSS ." Total depletions are a sum of depletions from agriculture, municipal and industrial sources, exports, and reservoir regulation. Human-induced salt results from agricultural salinity pickup and the salt removed with exports. We concluded that reservoir regulation has minimal effect on salt at an annual scale; therefore, reservoir regulation did not model salt. The time period simulated by the model dictated the depletion and salt loading data used in the run. For example, if the run simulated the historic time period from 1941 to 1995, the model used the depletions and human-induced salt loadings from the 1941 to 1995 historic record. If the model simulated the future time period of 2002 to 2062, it used projected depletions and salt loadings.

CRSS performs multiple simulations by running each trace of synthetic natural flow and associated salt time series through the CRSS simulation model. Our modified CRSS used 100 synthetic natural flow traces and associated salt mass traces that were each run through the simulation model, one at a time, calculating 100 simulated historic or future flows and associated simulated salt mass and concentration time series. These simulations could be used to approximate the PDF for the observed historic or predicted future flow, salt, and concentration.

We first used the CRSS simulation model to validate the modified CRSS system by showing that it could preserve the observed PDF for historic flow, salt mass, and concentration from 1941 to 1995. The inputs were 100 traces of synthetic natural flow and associated salt mass traces generated with the modified stochastic natural flow model, using the 1941 to 1995 calculated natural flow, and the statistical nonparametric natural salt model, using the synthetic natural flow traces generated from the modified stochastic natural flow model. The total depletions were based on 1941 to 1995 historic records.

The results were compared to the observed historic PDF, as discussed previously. The results indicated the observed historic flow and salt mass PDFs were preserved best using an annual flow and salt regression relationship and incorporating a relationship for total depletion, which is dependent on natural flow when simulating historic flow.

Stochastic Planning Run Simulations: Once the modified CRSS was validated the existing CRSS and the modified CRSS were used to simulate stochastic planning runs. The runs approximate the projected future flow, salt mass, and concentration from 2002 to 2061 using the most extensive data available. Using the entire calculated historic flow time series allowed the existing CRSS and the modified CRSS to be compared.

Stochastic planning runs were developed from a monthly time step simulation model. A synthetic natural flow time series of 60 years length was generated based on 1906 to 1995 calculated natural flow. The associated synthetic natural salt mass was computed using the statistical nonparametric natural salt model. The total depletions were projected future depletions from 2002 to 2061. Running the simulation for 60 years ensured the simulation was run until the future depletions reached full development.

The upper graph in Figure 2 shows the PDF created from the median for boxplots of probability at a given projected future flow. Both systems generated projected future flow similarly. The lower graph shows the PDF created from the median for boxplots of probability at a given projected future salt mass. The modified CRSS produced lower salt mass than the existing CRSS.

Figure 3 reiterates these results when the two modeling systems are compared with simulated policy analysis. To gain an understanding of how using the modified CRSS model could influence policy decisions, we tested policies using fictional salt mass and concentration standards. Salinity concentration standards are mandated at three locations in the lower basin of the Colorado River. Our segment of the CRSS simulation model did not include these locations, but developing fictional standards in the modeled segment facilitated comparing the performance of the existing and modified CRSS. These standards occur in the tails of the PDF, where extreme events, such as high salt mass or concentration, occur. The upper graph shows the fictional salt mass standard that determined the number of times 750,000 tons or more of salt occurred at the simulated gauge. The lower graph shows the number of times a salinity standard of 600 mg/L was violated. Again, the estimates of the existing CRSS exceeded the estimates of the modified CRSS.

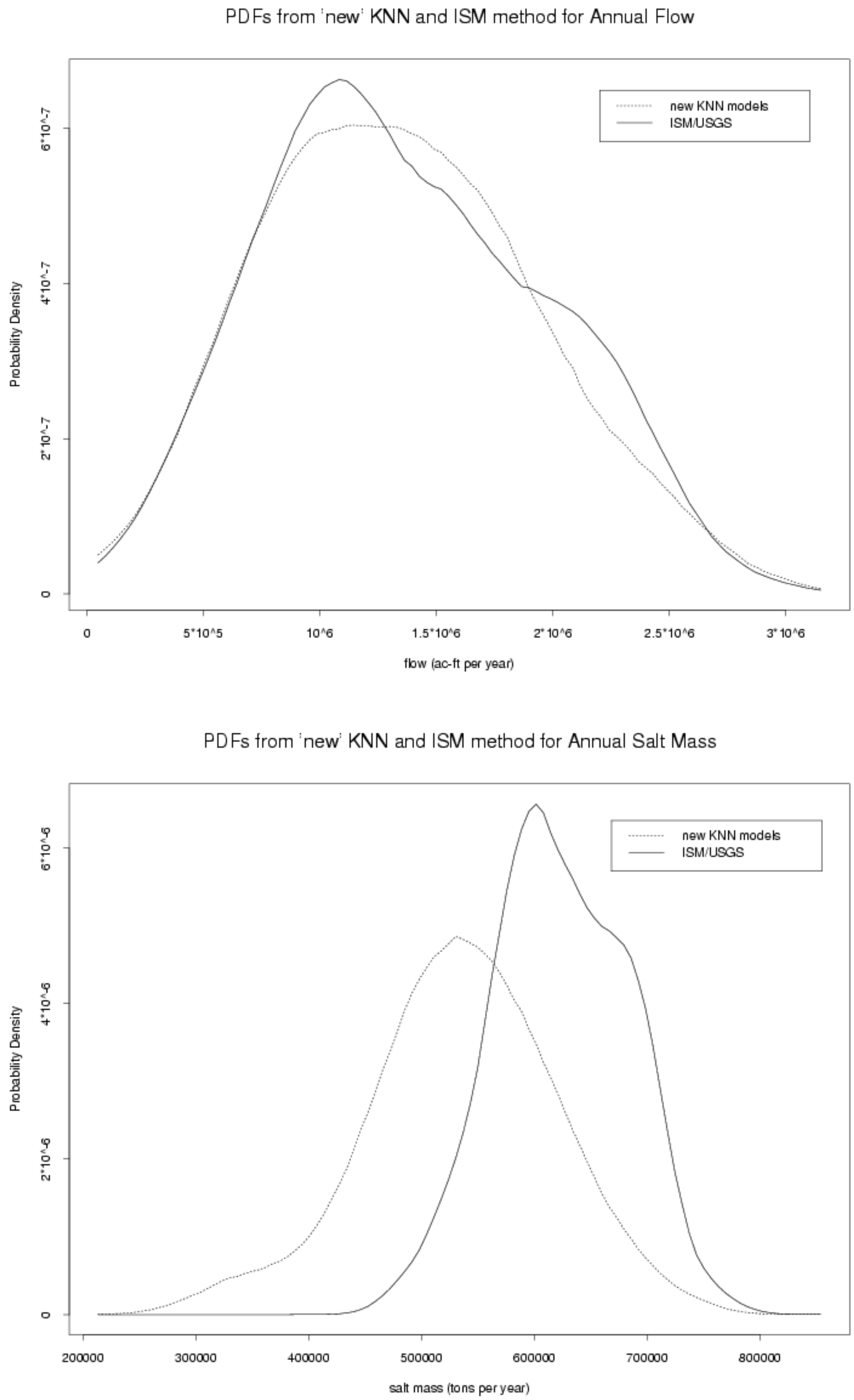
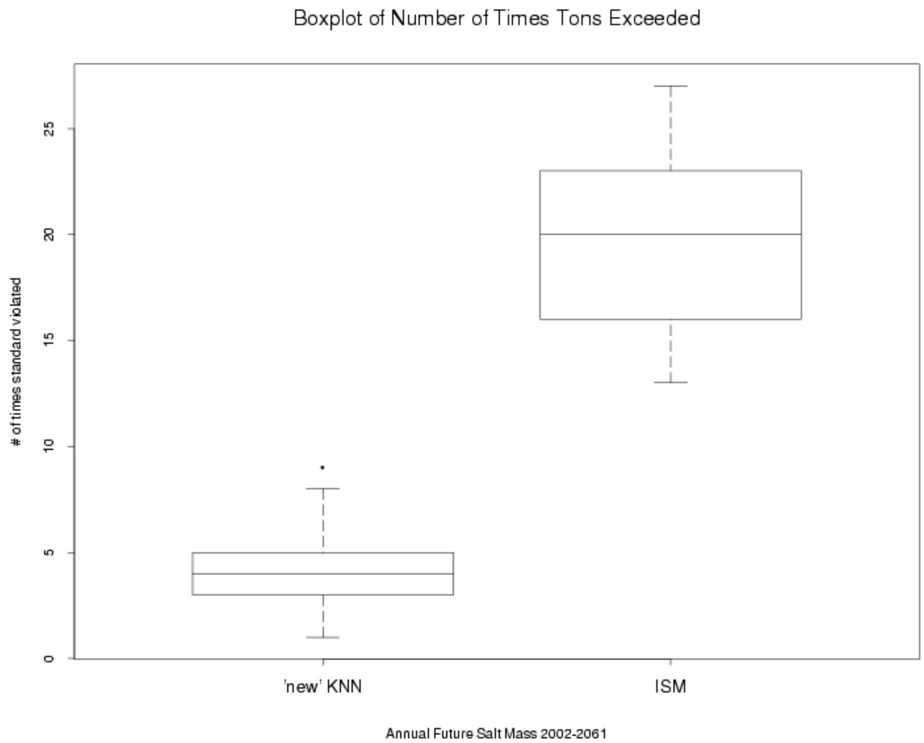


Figure 1: The upper graph shows the PDF created from the median of boxplots of probability at a given projected future flow generated from the modified and existing CRSS. The lower graph shows the projected future salt mass from the modified and existing CRSS. The shift in the projected future salt mass PDF shows that the existing CRSS generated greater salt mass.



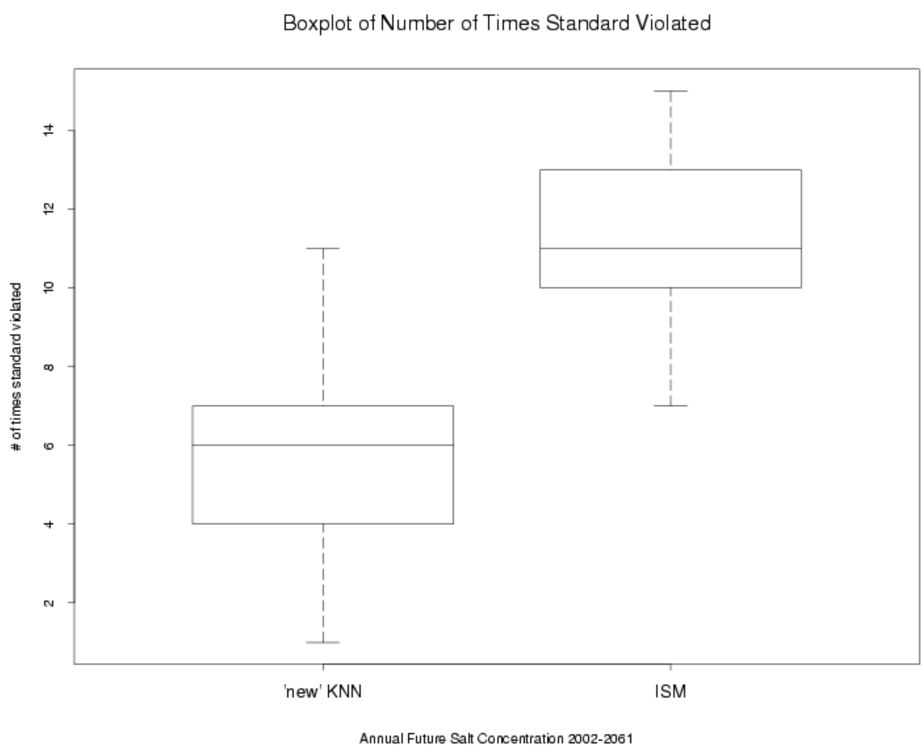


Figure 2: The upper graph shows the number of times the modified and existing CRSS results exceeded a tons standard of 750,000 tons, while the lower graph shows number of times a salt standard of 650mg/L was exceeded and, therefore, violated during water year 2002 to 2062. For both graphs the existing CRSS results exceeded the modified CRSS results.

DISCUSSION AND CONCLUSIONS

We outlined and developed a modified CRSS modeling system to simulate long-term flow, salt mass, and concentration at a single gauge in the upper Colorado River mainstem. The modified modeling system consisted of two modified models. The first model was a nonparametric K- NN model with resampling that generated synthetic natural flow. The second model used the generated natural flow and looked up an associated natural salt mass from a nonparametric local linear regression with K-NN resampling. A simulation model used the generated synthetic natural flow and salt mass to perform Monte Carlo simulations, which model flow, salt mass, and concentration with uncertainty.

We validated our modified modeling system by ensuring the modeling system reproduced all the distributional properties, i.e. the PDF of the observed historic record. Reproducing the PDF of the historic record ensured the mean, standard deviation, and skewness were all preserved. Validation was performed by using the model to generate synthetic streamflows for the period 1941 to 1995.

We used two time scales to develop a monthly and annual nonparametric local linear regression of natural salt mass as a function of natural flow with K-NN residual resampling model. At a monthly time scale, 12 regressions were developed - one for each month. To calculate the annual natural salt mass, the results from the 12 regressions were summed to an annual value. Using this method, we lost the ability to preserve the tails of the observed annual distribution. For policy analysis, with regard to long-term salinity mitigation, the annual time step is more important than the monthly time step. A monthly time step is only used to facilitate the use of the "rules" to simulate river basin policy. If the simulation model were run at an annual time step, an annual statistical nonparametric natural salt model would generate better results due to limitations of the monthly data. We found that using a single annual natural salt mass as a function of natural flow regression to generate annual natural salt mass best allowed the model to preserve the observed annual natural salt mass PDF.

After completing our validation, we used the modeling system at the monthly time step, accommodating the simulation model's requirement to model operational policy rules. Projected future flow, salt mass, and concentration were simulated from 2002 to 2061. We compared the modified and existing CRSS, using the entire natural flow dataset from 1906 to 1995. The comparison showed that the modified CRSS simulated lower future projected salt mass than the existing CRSS. The results were reiterated by the policy analysis demonstration. The existing CRSS generated more salt mass (88,000 tons lower) and concentration (57mg/L lower) violations than the modified CRSS.

Another strength of the modified CRSS is that the nonparametric K-NN stochastic flow model can be easily used to generate high- or low- flow scenarios to test policies under adverse conditions. To add this ability to the K-NN framework, the flow database that is used for resampling could be broken in thirds. The first third would be high flows, the second normal flows, and the third low flows. To generate a low-flow scenario, the K-NN model would generate a future streamflow time series by resampling only from the low flows in the database, thereby generating a synthetic time series with a sustained drought. This could also be performed by resampling from only the high flows to generate the time series.

Our modified modeling system generated historic flow and salt mass that matched the observed time period from which they were generated. We showed that our modified CRSS generated lower projected future salt mass and salt concentration than the existing CRSS. We developed the new modeling system from data used to run the simulation model. By using the data from the simulation model, we kept our methods to develop data to drive the simulation model consistent with the data used in the simulation model. Keeping the data consistent allowed the modified CRSS to reproduce historical results and to fully reflect variability by improving the relationship between natural flow and salt mass.

REFERENCES

Fulp. T., Vickers, B., Williams, B., King, D., 1999. Replacing an Institutional Model: The Colorado River Simulation System Example. paper presented at the WaterPower 99 conference, American Society of Civil Engineers, Las Vegas, Nevada.

Iorns, W.V., Hembree, C.H., Oakland, G.L., 1965. Water Resources of the Upper Colorado River Basin - Technical Report. U.S. Geological Survey, Professional Paper 441.

Kendall, D.B., Dracup, J.A., 1991. A Comparison of Index-Sequential and AR(1) Generated Hydrologic Sequences. Journal of Hydrology 122, 335-352.

Lall, U. 1995. Recent Advances in Nonparametric Function Estimation: Hydraulic Applications. Reviews of Geophysics 33 Part 2, Suppl. S, 1093-1102.

Mueller, D.K., Osen, L.L., 1988. Estimation of Natural Dissolved-Solids for the Upper Colorado River Basin. U.S. Geological Survey, Water Resources Investigation Report 87-4069.

Ouarda, T., Labadie, J.W., Fontane, D.G., 1990. Index Sequential Hydrologic Modeling For Hydropower Capacity Estimation. Journal of the American Water Resources Association 33:6, 1337-1349.

Prairie, J.R., Fulp, T.J., 2000. Verification and Calibration of the Monthly Colorado River Basin Salinity Data and Model. Report dated June 21. CADSWES.

U.S. Bureau of Reclamation. May 1987. Colorado River Simulation System, System Overview. U.S. Department of Interior, Bureau of Reclamation, Denver, Colorado.

U.S. Department of Interior, 2001. Quality of Water Colorado: Progress Report 20 [WWW Document]. URL http://www.uc.usbr.gov/progact/salinity/index.html (visited 2001, October 15).

Zagona, E.A., Fulp, T.J., Shane, R., Magee, T., Goranflo, H.M., 2001. RiverWare: Generalized Tool For Complex Reservoir System Modeling. Journal of the American Water Resources Association 37:4, 913-929.