Methodology for Flow and Salinity Estimates in the Sacramento-San Joaquin Delta and Suisun Marsh

36th Annual Progress Report
June 2015

Chapter 7
Calibrating the Martinez Boundary Salinity Generator Using PEST

Authors: Prabjot Sandhu and Yu Zhou
Delta Modeling Section
Bay-Delta Office
California Department of Water Resources

Contents

7 Calibrating the Martinez Boundary Salinity Generator Using PEST 7-1
7.1 INTRODUCTION 7-1
7.2 BACKGROUND 7-1
7.2.1 G-Model 7-2
7.2.2 Tidal Signal Incorporation 7-3
7.2.3 Model Calibration and Problems 7-3
7.3 METHODOLOGY AND STUDY CONFIGURATION 7-5
7.3.1 Theory 7-6
7.3.2 Configuration Setup 7-6
7.4 RE-CALIBRATION RESULTS AND DISCUSSION 7-8
7.4.1 Calibration for 1991-1993 7-8
7.4.2 Model Improvement and Re-calibration 7-8
7.4.3 Re-Calibration with Other Periods 7-11
7.5 SUMMARY AND FUTURE WORK 7-12
7.6 REFERENCES 7-13

Figures

Figure 7-1 Salinity at Martinez Boundary of the Bay-Delta 7-1
Figure 7-2 14-day Average Salinity as a Function of 14-day Average Net Delta Outflow (Q) and Antecedent Outflow (G). Data are for Water Year 1968-1986 7-2
Figure 7-3 Tidal Displacement Model as Embedded in the Full Salinity Model 7-3
Figure 7-4 Daily Averaged Salinity Comparison between Historical (blue), Old Salinity Generator (red) (enlarged for 1990-1992, 1993-1995) 7-5
Figure 7-5 Flow Chart for Inverse Problem (parameter estimation) with X as System Configuration 7-6
Figure 7-6 PEST Calibration Process of Martinez EC Generator 7-7
Figure 7-7 Daily Averaged Salinity Comparison Between Historical (blue), Old Generator (red), New Generator (green) Enlarged for 1990-1992, 1993-1995, 2002-2004, 2012-2014 7-9
Figure 7-8 Daily Averaged Salinity Comparison Between Old generator (green, with most of parts overlapped) and Modified Generator (red) 7-10

Tables

Table 7-1 Previous and New Calibration of Parameters of Martinez Salinity Planning Generator 7-4
Table 7-2 Part of Parameter Correlation Coefficient Matrix of Auto-calibration 7-10
Table 7-3 Parameters from Auto-calibration of the Modified Martinez Salinity Planning Generator 7-10
Table 7-4 Parameters of Auto-calibration from Observations of Different Periods 7-12

7 Calibrating the Martinez Boundary Salinity Generator Using PEST

7.1 Introduction

Martinez represents the stage-and-salinity boundary and the location for applying the Delta Simulation Model 2 (DSM2). The salinity at this location is estimated using the Net Delta Outflow (NDO) and stage. This chapter presents a re-calibration effort for the Martinez boundary salinity generator, with a mathematically based calibration software named PEST. This new calibration improves the performance of the model by better matching the historical salinity data, particularly at the higher value range. The performance of the current calibration has been a concern of water resources management, especially in the current drought crisis.

The chapter is organized with section 7.2 describing the background of Martinez boundary salinity generator, section 7.3 explaining methodology and configuration of PEST for model calibration, and section 7.4 presenting some preliminary findings.

7.2 Background

The Sacramento-San Joaquin Delta (Bay-Delta), Figure 7-1, is the nexus of California water system, which provides more than 30 million acre-feet of fresh water pumped to support drinking, agricultural, and industrial-use water. Martinez, with its critical estuarine location and continuous field measurement, serves as the downstream boundary in many modeling tools for this important Bay-Delta domain. The flow-salinity relationship at Martinez is the concern of research, because it could be used in real time or planning practice for various Delta operations and control strategies.

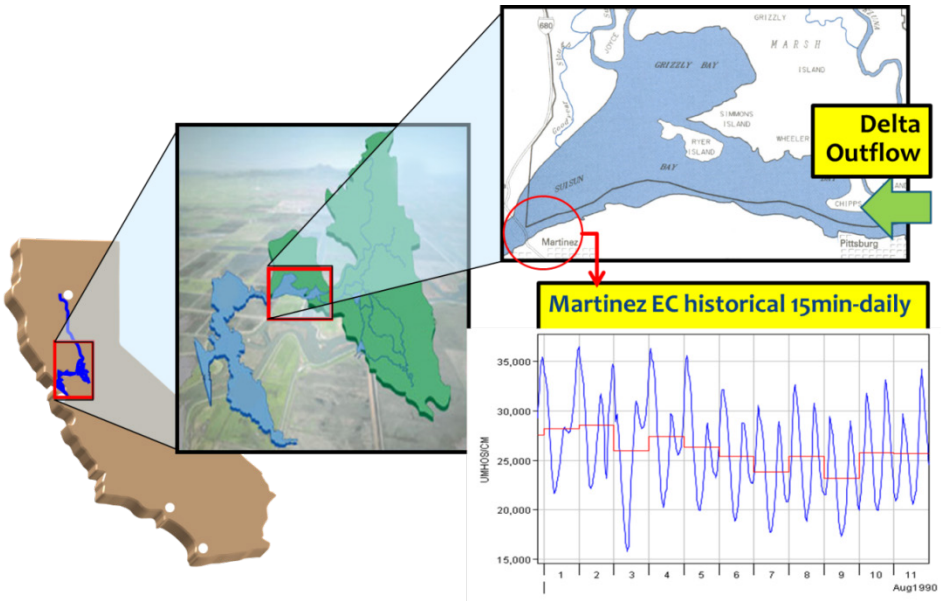

Figure 7-1 Salinity at Martinez Boundary of the Bay-Delta

7.2.1 G-Model

G-model is a conceptual-empirical model of salinity transport along the main stem of the Sacramento River (Ateljevich 2001b). It provides a 1-dimensional advective-diffusion solution for an infinitely long channel with a downstream ocean and an upstream river in its steady state (Denton and Sullivan 1993).

Its mathematical form could be written as an exponential longitudinal salinity profile

$$S(t) = (S_o - S_b) * \exp(-\alpha G^n(t)) + S_b.$$

Where

S_o and S_b are constant ocean and upstream river salinity respectively,

α is a dispersion parameter (consolidating upstream distance),

n is an additional empirical shape parameter,

G is a function representing the flow time-history, aka antecedent flow, with its math form as below

$$\frac{dG}{dt} = \frac{(Q-G)G}{\beta},$$

β is an empirically determined constant depicting system's reaction to Delta outflow,

Q is volumetric flowrate, which is NDO for Martinez,

t is time,

$\frac{dG}{dt}$ is the derivative of G with respect to time.

Figure 7-2 represents the advantages of using *G* over flow *Q* in flow-salinity relationship estimation. Details of the original model may be found in Dr. Denton's report (Denton and Sullivan 1993).

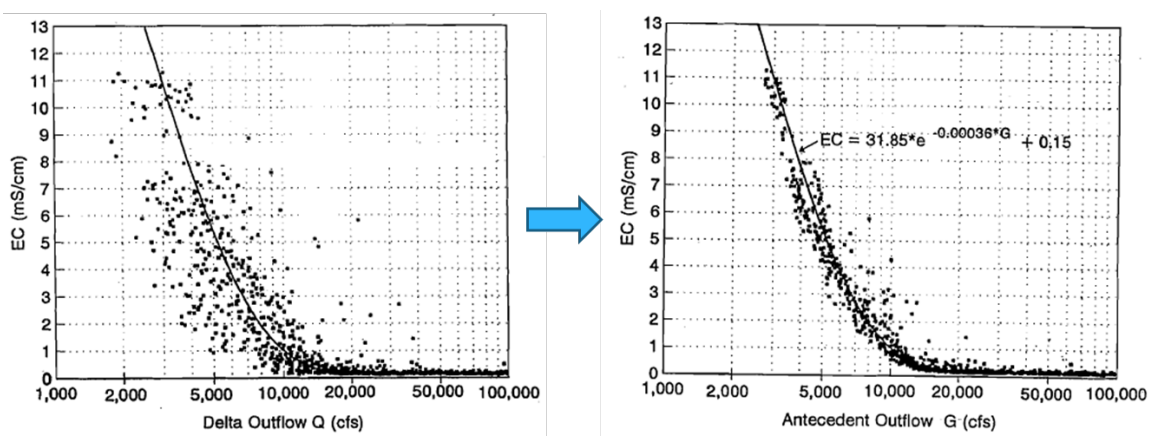

Figure 7-2 14-day Average Salinity as a Function of 14-day Average Net Delta Outflow (Q) and Antecedent Outflow (G). Data are for Water Year 1968-1986

Notice the corrections to NDO to account for both Delta filling and draining and island consumptive use estimates,

$$q_{corrected}(t) = q_{ndo}(t) - A\Delta z(t).$$

A represents the storage area filled by an incremental change in water surface height $\Delta z(t)$.

7.2.2 Tidal Signal Incorporation

To make predictions more accurately in a closer time interval, this salinity model was further developed to incorporate the tidally varying effect from the ocean. The newly applied methodology utilizes a linear filtration to model displacement on lagged values of stage, since the stage signal is directly observable, relatively noise free, and more perfectly harmonic in character (Ateljevich 2001b). The flow chart of the entire nonlinear model is depicted in Figure 7-3, with its full mathematical form modified as below.

$$S(t) = (S_o - S_b) * \exp(-\alpha G^n(t)x(t)) + S_b.$$

Where

x is a harmonic position, reflecting tidal effect, convolution filter,

S_o and S_b are constant ocean and upstream river salinity respectively,

α is a dispersion parameter (consolidating upstream distance),

n is an additional empirical shape parameter,

G is a function representing the flow time-history, aka antecedent flow,

t is time,

exp() is an exponential function.

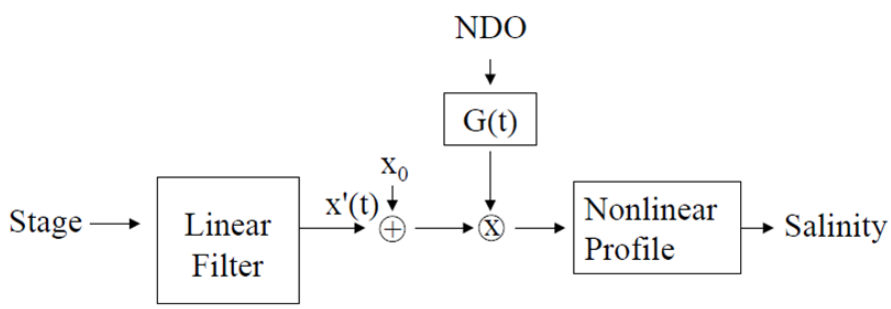

Figure 7-3 Tidal Displacement Model as Embedded in the Full Salinity Model

The final summarized equation is

$$ln\left[E\frac{s(t)-s_b}{s_0-s_b}\right] = \beta_0 + \beta_1 g^n(t) + g(t)\sum_{k=0}^{n_k-1} a_k z(t + k_0\Delta t - k\Delta t).$$

7.2.3 Model Calibration and Problems

The existing model was calibrated for 1991/8/20-1992/9/5 (yellow shaded area in Figure 7-4a), with parameters in the fourth column of Table 7-1 and validated for 1993-1994 (green shaded area in Figure 7-4a). The model output matches the main trend and tidal envelope, with a root mean squared error of 2,828 µs/cm during the validation period (Figure 7-4c).

	Description		Previous calibration	New calibration
Φ	The weighted least squares residuals sum		8.85E+11	4.43E+11
S_o	G-model related	ocean river salinity	32797	35414
S_b		upstream river salinity	200	2333.1
β		system's reaction to Delta outflow	600	419.7845
n		empirical shape parameter	0.75	0.77
A		Correction coefficient of NDO from Delta filling and draining	40000	53662.45
β_0	Tidal coefficient	Convolution filter coefficients for stage	2.76E-03	1.53E-01
β_1			-6.07E-05	-6.00E-05
a_0			1.52E-04	7.40E-05
a_1			-1.05E-05	-1.00E-05
a_2			-2.83E-06	-4.00E-05
a_3			4.96E-05	2.90E-06
a_4			-8.76E-05	-1.00E-04
a_5			7.21E-05	4.60E-05
a_6			-5.18E-05	-1.00E-04

Table 7-1 Previous and New Calibration of Parameters of Martinez Salinity Planning Generator

Newly collected field measurements of last 20 years show that the model underestimates electrical conductivity, (EC), up to 5,000 µs/cm (or the equivalent 5,000 µmhos/cm), at the higher salinity range. Furthermore, this seems to have an upper bound of around 25,000 µs/cm, which is possibly limited by its calibrated parameters.

This salinity underestimation is a concern for water resources operation and planning, especially in the recent continuous drought years. Accordingly, a new thorough and automated calibration is designed in this study.

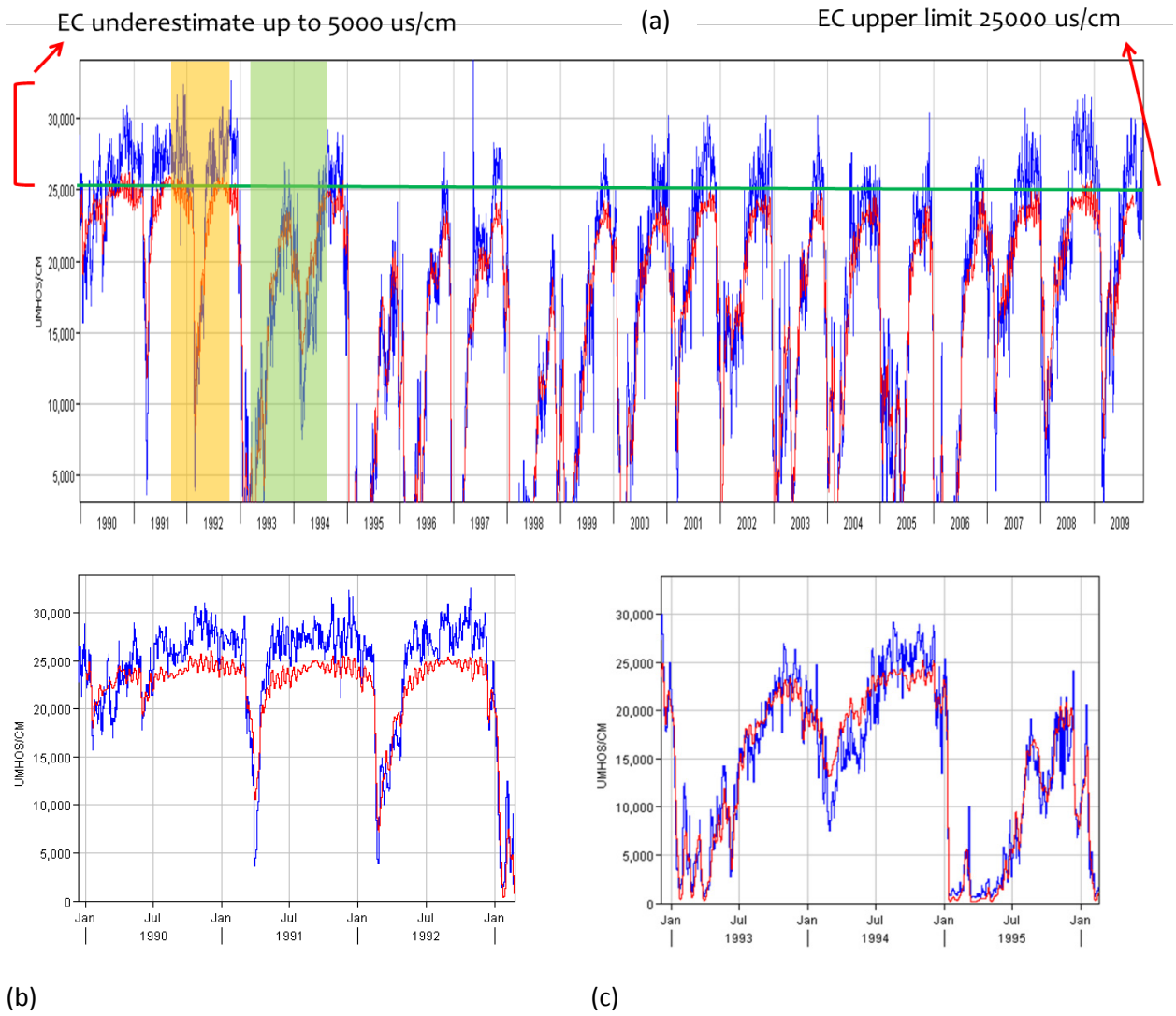

Figure 7-4 Daily Averaged Salinity Comparison between Historical (blue), Old Salinity Generator (red) (enlarged for 1990-1992, 1993-1995)

7.3 Methodology and Study Configuration

PEST, short for Parameter ESTimation, is the industry standard software package for parameter estimation and uncertainty analysis of complex environmental and other computer models, http://www.pesthomepage.org/. It is designed to relieve modelers from manual calibration, which is labor-intensive and subjective (Doherty 2010).

A major feature of PEST is that it is model-independent, so users just need to provide input and output in the required format (Figure 7-5). It has already been used in DWR's Modeling Support Branch for groundwater and surface water model calibrations (Kadir 2006; DiGiorgio 2009; Finch 2014). Details of this methodology and related utility can be found at http://www.pesthomepage.org/.

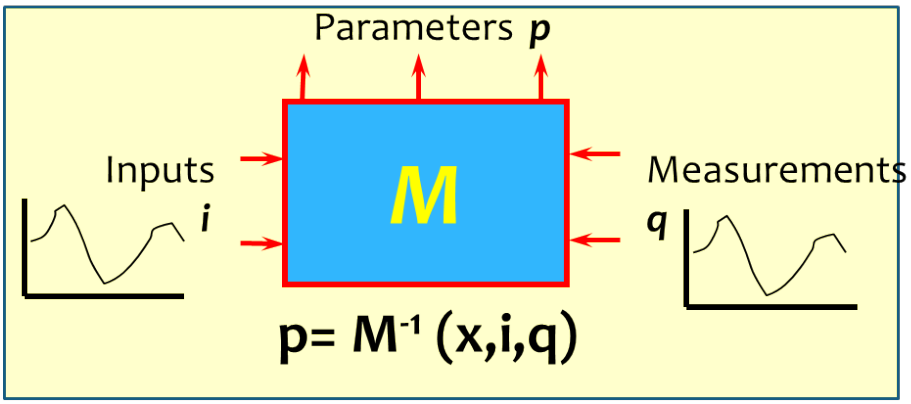

Figure 7-5 Flow Chart for Inverse Problem (parameter estimation) with X as System Configuration

7.3.1 Theory

PEST works by adjusting model parameters to minimize the discrepancies between model outputs and field/laboratory measurements. The discrepancy is formulated below as a single term, weighted least-squares-residuals sum.

$$\text{Objective function: } \Phi = \Sigma\, w_i^2 (q_i - o_i)^2 = \Sigma w_i r_i^2.$$

Where

q as field/lab measurement,

o is model output,

r is residual between field measurement and model output,

w is observation weight,

i is measurement index.

The nonlinear parameter estimation technique PEST uses is called Gauss-Marquardt-Levenberg method (Wikipedia, 2015). It basically constructs the Jacobian matrix of derivatives by using finite differences. It does this by perturbing each parameter independently and running the model to simulate the output at the new value. It then uses a gradient descent method damped by a trust region approach to calculate an updated set of parameters. This process is iterated until an optimal solution is approached.

$$o - o_0 = J\,(p - p_0).$$

$$J_{i,j} = \partial o_i / \partial p_j.$$

Where p is a parameter.

7.3.2 Configuration Setup

There are a couple of versions of the G-Model (described in section 7.2.1) in Bay-Delta Office. One version is used for long-term planning studies. The second version is used for near-term operational simulations. The recalibration described in this chapter using PEST was designed to first study the improvements for long-term planning studies. These improvements and results could be applied in the near-future forecasting G-Model version with Kalman function, which assimilates recent data from neighboring stations to increase the accuracy.

The planning Martinez salinity generator is written in Python scripts, and is included in the DSM2 package. Two of its inputs, Net Delta Outflow (NDO) and stage, are required to be in the format used in the U.S. Army Corps of Engineers' Hydrologic Engineering Center Data Storage System, HEC-DSS. Since PEST uses text files as input and output, wrapping functions are written to enable the data file conversion (Figure 7-6).

- NDO (daily time step) is calculated as the sum of all the Delta boundary flows and exports, with consumptive usage in the Delta Island Consumptive Use (DICU) model version of year 2014, based on California Department of Water Resources (DWR) Bay-Delta Office (BDO) historical data collection.
- Stage (15-minute time step) is the astronomical planning stage at Martinez, also generated by DWR BDO (Ateljevich 2001a).

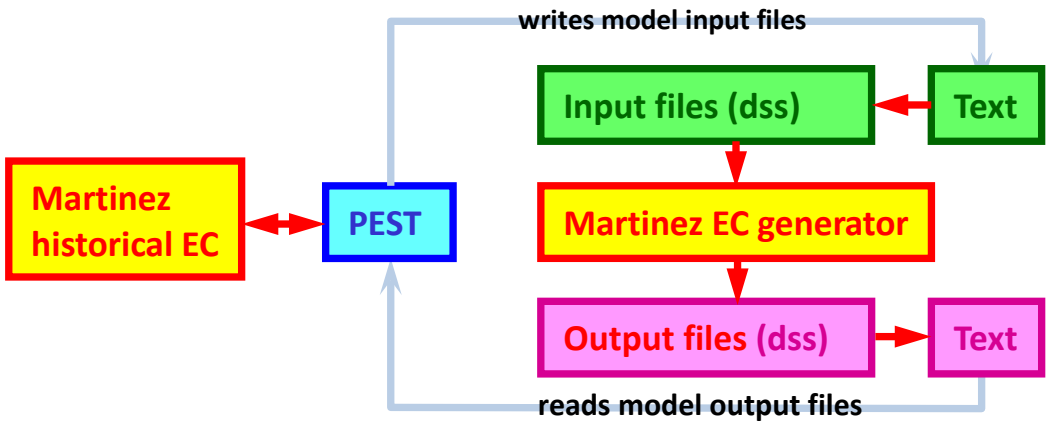

Figure 7-6 PEST Calibration Process of Martinez EC Generator

The model output EC at Martinez (15-minute time step) is compared with the estimation target, its relevant historical data, collected by DWR.

In this study, 14 parameters are determined as adjustable for calibration as shown in Table 7-1. Among which, five parameters are used for G-model daily salinity, and nine parameters are for stage tidal signal. Calibration is conducted first for the former group, with the latter group held constant; then all parameters are calibrated together to match the stage profile. This process is conducted iteratively until a stable result is achieved. Some details of calibration process are:

- Python and Windows-batch scripts have been utilized to develop pre- and post-processors to streamline and automate.
- As a result of the output data limit of the current PEST version, usually the calibration period is selected at less than 2 years and 10 months for 15-minute data.
- Parallel computing utility, BEOPEST, has been utilized on a multicore desktop machine to increase the processing efficiency. Calibration completes in about an hour with typically approximately 20 iterations.
- Different initial values have been tested and do not affect the final convergence much. Details could be referred to the discussion of sensitivity in section 7.4.2.

7.4 Re-Calibration Results and Discussion

The focus of this study is to improve the model suitability at high salinity range, especially the continuous high salinity drought years. Accordingly, 1990-1993 are first selected as the primary calibration period, and the recent 20 years are selected as the validation period.

7.4.1 Calibration for 1991-1993

As shown in Figure 7-7, this study selected 1991/1-1993/9 as the calibration period (yellow shaded area), and 1990-2014 as the validation period. Calibrated parameters are included in Table 7-1. Compared with the old validation outputs (red), the new validation outputs (green) better match the historical data (blue).

- New validation results still match the main trend and tidal envelope well at the moderate-EC range.
- For the low-EC range, new validation results show higher predictions, but this is deliberately ignored as prediction error at the lower range and is not a concern.
- New validation results improve the high-EC range prediction significantly, especially over 25,000 μs/cm (or the equivalent 25,000 μmhos/cm) for most years. Note that in some summers, especially during the recent high-EC period 2012-2014, the predictions from the new calibration still underestimate the observed EC by 2,000-3,000 μs/cm. The reason of this underestimate is still not known, although it is speculated that a more recent calibration data set may result in an improvement.

On the other hand, compared with historical data, model outputs tend to be smoother, while the real data is likely to have more abrupt changes (spikes), especially in the high-EC summer time. This could be caused by local environment variation, like terrain change and weather (e.g., wind) and the current model algorithm cannot represent these details.

7.4.2 Model Improvement and Re-calibration

PEST provides a convenient utility to explore correlation between parameters. Table 7-2 lists part (G-model) of the correlation matrix from calibration of section 7.4.1, with every cell indicating correlation between every two parameters (name of row and column). The closer a value is to 1 or -1, the higher the correlation between those two parameters.

Consequently, this information can be used to achieve parameter parsimony. In this case, S_o has -1 correlation with β_0, totally negatively correlated. It was found that β_0 was introduced with S_o and S_b fixed in the original calibration (Ateljevich 2001b), so it is worthwhile to remove (combine) β_0 to make the model simpler, as shown in the following equation. Calibration was conducted on the modified model. With both Φ and output plot (Figure 7-8) showing that the modified model produces a very similar result, it is safe to claim the validity of β_0 removal.

$$ln\left[\frac{s(t)-s_b}{s_0-s_b}\right] = \beta_1 g^n(t) + g(t)\sum_{k=0}^{n_k-1} a_k z(t + k_0\Delta t - k\Delta t).$$

Where
t is time,
s_0 and s_b are constant ocean and upstream salinities respectively,
β is an empirically determined constant depicting system's reaction to Delta outflow,
g is a function representing the flow-time history,
a, n, and k are convolution filter coefficients.

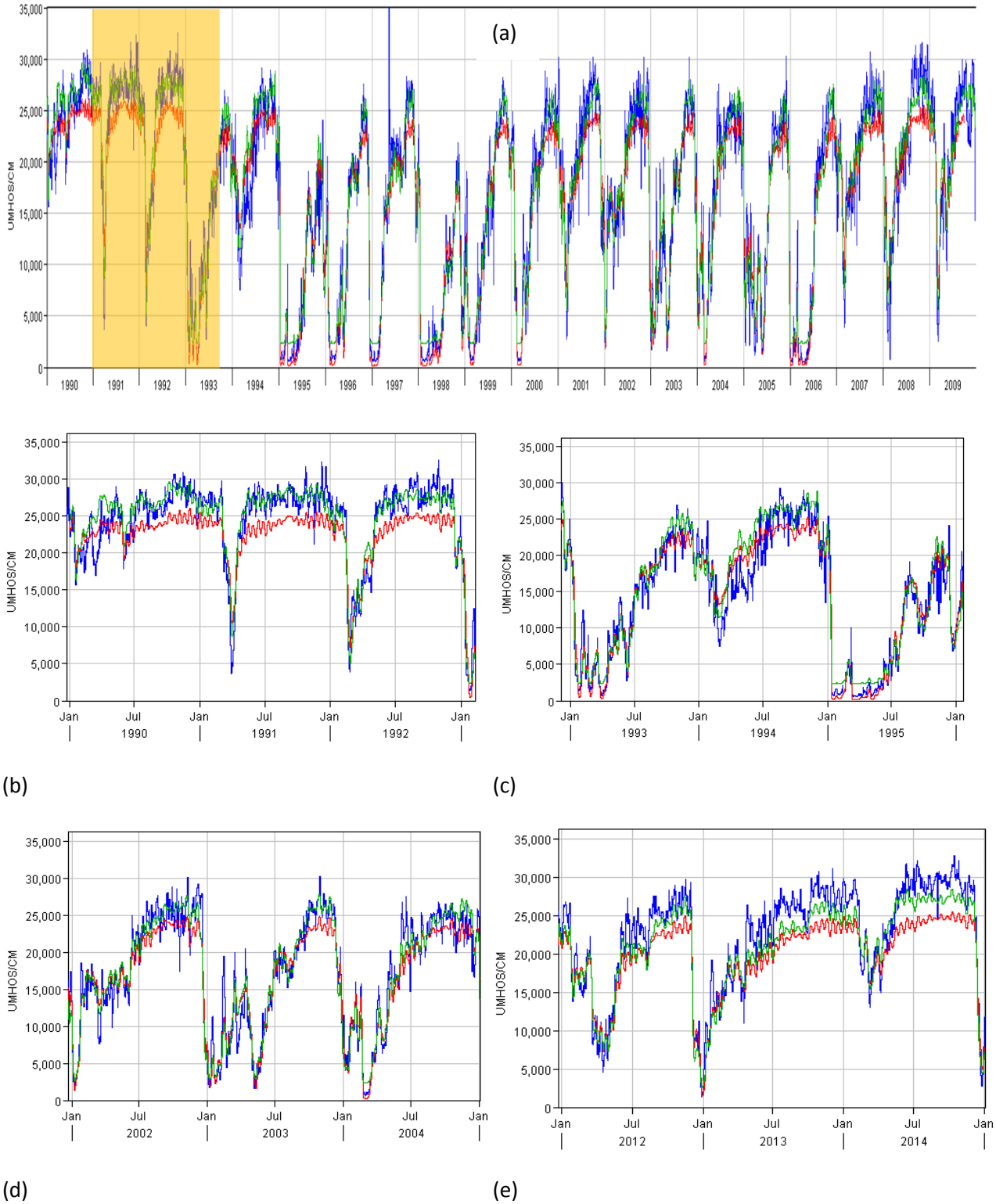

Figure 7-7 Daily Averaged Salinity Comparison Between Historical (blue), Old Generator (red), New Generator (green) Enlarged for 1990-1992, 1993-1995, 2002-2004, 2012-2014

	S_o	S_b	β	n	A	β_0
S_o	1	1.50E-03	-2.11E-04	-3.09E-03	1.90E-04	-1
S_b	1.50E-03	1	-5.88E-02	0.3235	-0.1598	-1.49E-03
β	-2.11E-04	-5.88E-02	1	-4.62E-02	0.2964	2.19E-04
n	-3.09E-03	0.3235	-4.62E-02	1	-0.2933	3.10E-03
A	1.90E-04	-0.1598	0.2964	-0.2933	1	-1.79E-04
β_0	-1	-1.49E-03	2.19E-04	3.10E-03	-1.79E-04	1

Table 7-2 Part of Parameter Correlation Coefficient Matrix of Auto-calibration

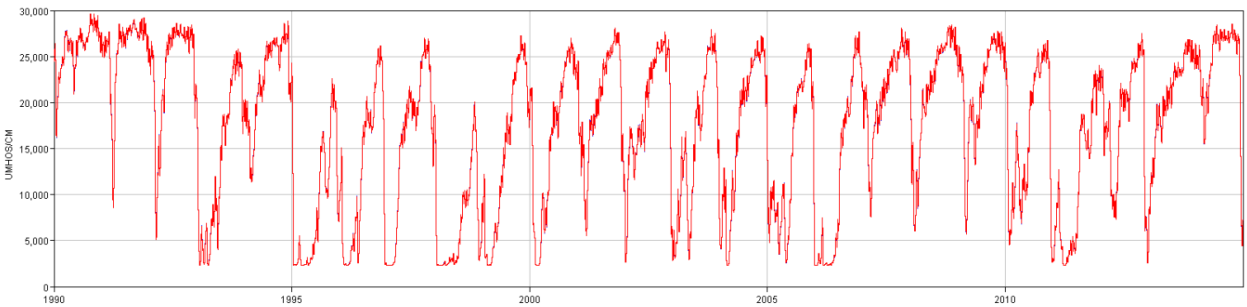

Figure 7-8 Daily Averaged Salinity Comparison Between Old generator (green, with most parts overlapped) and Modified Generator (red)

To quantify the parameters' uncertainty, their 95 percent confidence limits are also calculated, shown in the third and fourth columns of Table 7-3.

	Calibration values	95% percent confidence		Sensitivity	Rel. Sensitivity
Φ	4.46E+11	Lower limit	Upper limit		
S_o	37196	37122.6	3.73E+04	1.91E-03	71.168
S_b	2328.1	2273.29	2382.91	1.72E-03	4.00954
β	420	418.025	4.23E+02	1.90E-02	8.0004
n	0.775	7.69E-01	0.7813	104.992	8.14E+01
A	53411.1	52071.8	5.48E+04	3.91E-05	2.08748
β_1	-6.00E-05	-6.06E-05	-5.94E-05	462591	27.7555
a_0	7.30E-05	6.90E-05	7.70E-05	142668	10.4148
a_1	-1.00E-05	-1.33E-05	-6.74E-06	124834	1.24834
a_2	-3.00E-05	-3.54E-05	-2.46E-05	103390	3.10169
a_3	1.70E-06	-3.37E-06	6.77E-06	106941	0.1818
a_4	-1.00E-04	-1.07E-04	-9.25E-05	126659	12.6659
a_5	4.50E-05	4.09E-05	4.91E-05	122040	5.49179
a_6	-1.00E-04	-1.06E-04	-9.40E-05	110693	11.0693

Table 7-3 Parameters from Auto-calibration of the Modified Martinez Salinity Planning Generator

Correspondingly, the composite parameter sensitivities and relative composite sensitivities are shown in the fifth and sixth columns of Table 7-3. The mathematical definition of composite sensitivity

is shown as below, while the relative composite sensitivity is calculated by multiplying the magnitude of the value of the parameter. The latter represents a measure of the composite changes in model outputs that are incurred by a fractional change in the value of the parameter.

$$s_i = (J^t QJ)_{ii}{}^{1/2}/,$$

Where

J as the Jacobian matrix, and *Q* as the cofactor matrix,

s is sensitivity,

Q is the cofactor matrix.

Those parameters with the much lower sensitivities are regarded as insensitive, e.g., a_3, which corresponds to the larger confidence interval. They reflect that the results of a wider range of calibrated parameter values do not make difference on Φ. On the other side, higher sensitivity represents a narrower range calibrated parameter values, e.g., S_o, n, β_1, are the key parameters and dominate the model performance.

Different initial parameter values were also tested, and their results converge to the same magnitude, although some insensitive parameters, like a_3, may end in different values.

7.4.3 Re-Calibration with Other Periods

Several different calibration periods were selected as comparison groups and the calibrated parameter results are shown in Table 7-4. Some observations are:

- Different target data could result in different calibrated parameters, due to the regression nature of the salinity estimator. Usually high EC is represented by high S_o (ocean salinity boundary).
- These other control groups (2 years 10 months) do not show obvious improvement on objective target ϕ, i.e., match with observed data.
- Since different periods have different target data, they should not be compared directly.

For the purpose of getting a better estimate for high-EC, which is the concern of the study, results of 91/1-93/9 continuous drought period was still used as study target for a general planning estimator.

Calibration period	1991/1-1993/9	1994/1-1996/9	2000/1-2002/9	2006/1-2008/9	2009/1-2011/9	2012/4-2014/12
Φ	4.46E+11	6.06E+11	4.68E+11	7.49E+11	6.58E+11	4.79E+11
S_o	37196	29192	34692	34966	34215	38948
S_b	2328.1	616.58	1082.4	-501.2	988.71	2292.4
β	420	502	386	331	434	407
n	0.775	0.74	0.76	0.73	0.80	0.84
A	53411	51908	60968	39099	73051	48694
β_1	-6.00E-05	-6.0E-05	-6.0E-05	-6.0E-05	-6.0E-05	-8.0E-05
a_0	7.30E-05	1.5E-04	1.0E-04	1.5E-04	8.4E-05	8.0E-05
a_1	-1.00E-05	-1.0E-05	-3.0E-05	-1.0E-05	-4.0E-05	-5.0E-05
a_2	-3.00E-05	-3.0E-06	1.6E-05	-3.0E-06	3.7E-05	6.9E-05
a_3	1.70E-06	5.0E-05	-1.0E-05	5.0E-05	-3.0E-05	-6.0E-05
a_4	-1.00E-04	-9.0E-05	-5.0E-05	-9.0E-05	-2.0E-05	3.6E-05
a_5	4.50E-05	7.2E-05	3.3E-05	7.2E-05	8.7E-06	-2.0E-05
a_6	-1.00E-04	-5.0E-05	-6.0E-05	-5.0E-05	-3.0E-05	-1.0E-06

Table 7-4 Parameters of Auto-calibration from Observations of Different Periods

7.5 Summary and Future Work

In this study, PEST was applied to improve the Martinez salinity generator (G-Model).

- PEST formulates the calibration from manual to systematic automation, relieving people from time-consuming and subjective adjustment.
- PEST newly calibrated parameters produce outputs matching the historical field measurements better than previous calibrated ones.
- This model improvement is important because of its advance in high-EC range, which is the concern in continuous drought years.

Many benefits of PEST have been verified during this study.

- PEST is easy to set up and use for calibration.
- PEST, due to its systematic automation, makes it easy to try different nonlinear additional changes, large number of parameters, and ill-posed problems.
- PEST parallelism performance on modern multicore machines could largely increase efficiency.
- PEST outputs information that helps identify highly correlated parameters.

Future efforts will be spent in the following areas:

- Investigate recent historical information.
- Estimate NDO with another salinity-flow model, DSM2, and with other inputs such as consumptive use, Sacramento River flow, San Joaquin River flow, exports, and gate operations.
- Real-time operation use for forecasting runs.
- Water cost estimations for planning scenarios.

7.6 References

Ateljevich E. 2001a. "Planning Tide at the Martinez Boundary." In: *Methodology for Flow and Salinity Estimates in the Sacramento-San Joaquin Delta and Suisun Marsh*. Chapter 10. Sacramento (CA): California Department of Water Resources.

———. 2001b. "Improving Salinity Estimates at the Martinez Boundary."In: *Methodology for Flow and Salinity Estimates in the Sacramento-San Joaquin Delta and Suisun Marsh. 22nd Annual Progress Report*. Chapter 11. Sacramento (CA): California Department of Water Resources.

Denton R. and Sullivan G. 1993. "Antecedent Flow-salinity Relations: Application to Delta Planning Model. In: *Methodology for Flow and Salinity Estimates in the Sacramento-San Joaquin Delta and Suisun Mars. 34th Annual Progress Report*. Sacramento (CA): California Department of Water Resources.

DiGiorgio CL. 2009. "Data Quality Associated with EC Values in the DICU Model." In: *Methodology for Flow and Salinity Estimates in the Sacramento-San Joaquin Delta and Suisun Marsh. 30th Annual Progress Report*. Chapter 2. Sacramento (CA): California Department of Water Resources.

Doherty J. 2010. *PEST, Model-Independent Parameter Estimation:* 5th Edition. Watermark Numerical Computing.

Finch R. 2014. "Quantitative Calibration of DSM2." In: *Methodology for Flow and Salinity Estimates in the Sacramento-San Joaquin Delta and Suisun Marsh. 35th Annual Progress Report*. Chapter 2. Sacramento (CA): California Department of Water Resources.

Kadir T. 2006. "Estimates for Consumptive Water Demands in the Delta using DETAW.: In: *Methodology for Flow and Salinity Estimates in the Sacramento-San Joaquin Delta and Suisun Marsh. 27th Annual Progress Report*. Sacramento (CA): California Department of Water Resources.

Wikipedia. 2015. Gauss-Marquardt-Levenberg algorithm. Viewed online at:https://en.wikipedia.org/wiki/Levenberg%E2%80%93Marquardt_algorithm. Accessed: 2015.