FINAL TASK REPORT

TASK 1 - INVESTIGATION OF EVAPORATION SURROGATES AND DIRECT RAINFALL FOR FLORIDA BAY SALINITY MODELING

Cooperative Agreement Number CA H5284-05-0006
Between
The United States Department of the Interior National Park Service
Everglades National Park
And
Cetacean Logic Foundation, Inc.

Frank E. Marshall, III
Principal Investigator

DeWitt Smith
Project Coordinator

August 6, 2008

CESI Project No. CA H5284-05-0006 for Everglades National Park

FINAL TASK REPORT

TASK 1 – INVESTIGATION OF EVAPORATION SURROGATES FOR FLORIDA BAY SALINITY MODELING

August 6, 2008

I. Introduction

A. General

This report describes a task for the current Critical Ecosystem Studies Initiative (CESI) project that was originally initiated by Cetacean Logic Foundation, Inc. (CLF) for Everglades National Park (ENP) in 2002. The work was continued in 2003 and 2004, and has been continued into 2005/2006. The subject of this on-going CLF work is the characterization and simulation of the salinity regime in the study area of Florida Bay and the southwest coast of the Gulf of Mexico (Gulf) within the Park.

For this phase of the CESI project, the scope of work has been broken into a number of tasks. This report describes the task of investigating surrogates for Florida Bay salinity modeling.

B. Task Objectives and Evaluation Methods

The objectives of this task are presented below as it appears in the contract for this CESI project approved by ENP:

"Investigate surrogates for evaporation that have the potential for inclusion in the models as candidate independent variables, and update models if needed. Also investigate the use of direct rainfall at the MMN salinity monitoring stations in the MLR salinity models so long as a long-term record (36-year) of local rainfall can be developed from regression models using Tavernier, Flamingo, and Royal Palm rainfall monitoring stations." (Marshall, 2004)

Although it has only been partly studied, evaporation is thought by most scientists and hydrologists to have a significant effect on salinity variation at times in south Florida, and this effect is thought to have a greater impact in the open water areas of Florida Bay, and during the dry season. Reliable, long-term, continuous data on Florida Bay evaporation do not exist for the period of MLR salinity model calibration and verification. Additionally, historical direct measurements of daily evaporation variability do not exist over the 36-year period for which SFWMD 2X2 model simulations are available for Comprehensive Everglades Restoration Plan alternative analysis. Therefore, this task will be focused on the evaluation of the correlative ability of evaporation

surrogates that are known to be available over the period of the assessments that are currently being made.

The same SARIMA correlation analysis that was used to choose the independent variables currently in the MLR salinity models will be used to evaluate surrogates for evaporation. The significance level for inclusion in the MLR salinity models is 0.999, which means that an evaporation surrogate must be highly correlated to salinity to be included in the models. If evaporation surrogates are found to definitively improve the MLR salinity models, the development of updated models will be considered.

In a similar manner, local rainfall measured at the MMN stations will be investigated as a potential candidate for an independent variable. However, the development of regression models using rainfall data at Tavernier, Royal Palm, and Flamingo stations must first be investigated in order to extend the period of available data to be compatible with the 36-year period of the other input data for salinity simulations.

II. Investigation of Evaporation as a Significant Independent Variable

Evaporation and evapotranspiration are difficult to measure in the field, even under the best conditions. The measurement of evaporation at a location that is in the open water of an expansive, remote estuarine/marine system such as Florida Bay or in the mangrove forest of the Everglades fringe is particularly difficult. For these and other reasons field-measured data on evaporation in Florida Bay and along the southwest Gulf coast are scarce. There are no long-term, field-measurements of evaporation that can be used for the 36-year evaluations that are needed for CERP evaluations. However, several models are available that use other more easily measured parameters to compute an estimate of evaporation.

Cosby, et al (2005) used three different approaches in estimating long-term evaporation for setting South Florida Water Management District (SFWMD) Minimum Flows and Levels salinity modeling purposes for Florida Bay. Air temperature, humidity and wind speed were used in a Dalton's Law calculation of evaporation as one approach. For the second approach, the temperature range was used to estimate atmospheric transmissivity in a radiation-based method. For the third approach, temperature alone was used to calculate potential evaporation based on the empirical Thornthwaite method. Each of these approaches was used to estimate monthly evaporation values.

Price, et al (2005) also used three methods to estimate evaporation, including the Dalton's Law method, the Priestly-Taylor method, and calibrated the "Simple" method described below for open water stations. Abtew, et al (2003) described an evaluation of a variety of methods for estimating potential evapotranspiration (as opposed to evaporation) throughout south Florida: Penman-Montheith,

Penman-Combination, Priestly-Taylor, Modified Turc, Radiation/Tmax, and "Simple" methods. It was noted that the Modified Turc, Radiation/Tmax, and Simple methods use only solar radiation and temperature and give comparable results to more complex methods. They compared the Modified Turc, Radiation/Tmax, and Simple methods using existing data and found that the Simple method provided reasonable estimate of potential evapotranspiration, and had the added advantage of being self-calibrated when used with the methodology presented in Allen (1997). The self-calibrating Simple method is being used to estimate evapotranspiration in a number of Everglades landscapes for SFWMM simulations (Irizarry-Ortiz, 2003) at the daily time step.

This evaluation of evaporation surrogates for MLR salinity modeling began with a preliminary analysis that was performed using the readily-available Joe Bay weather station data from the SFWMD DBHYDRO database and air temperature data that are available from the NOAA Key West station. The parameters that were included in this initial investigation dataset included all of the parameters previously used for MLR salinity modeling and the meteorological data described below. Salinity data were used for the preliminary evaluation from the Long Sound, Joe Bay, Little Madeira Bay, Terrapin Bay, North River, Whipray Basin, Duck Key and Butternut Key Marine Monitoring Network (ENP) stations. Meteorological data used in this preliminary evaluation included:

1. Joe Bay air temperature,
2. Joe Bay barometric pressure,
3. Joe Bay humidity,
4. Joe Bay water temperature,
5. Joe Bay air/sea temperature difference.
6. Key West air temperature (average),
7. Key West air temperature (minimum),
8. Key West air temperature (maximum), and
9. Joe Bay potential evapotranspiration (from the Simple method).

A model development exercise was conducted at each station in the same manner as the previous MLR salinity model development activities. The procedure begins with the evaluation of cross-correlation coefficient plots to identify the parameters in this meteorological dataset that were correlated with salinity (including lags). Correlated meteorological parameters were then tested for significance through step-wise regression to determine if the meteorological parameters improved the multivariate linear regression model. The significance level for inclusion of a parameter in a model was held at the 99.9% level, a very high level.

At all stations except Long Sound, one or more of the evaporation related parameters was found to be significant. In most cases, most of the original parameters (water level in the Everglades, sea level, and wind vectors) remained the most important in explaining the variation in salinity. The evaporation

parameters did not typically displace the salinity model parameters but instead became an added significant independent variable. However, in all cases the evaporation parameters only affected a minor improvement if any in the explanatory capability of the model as measured by the adjusted coefficient of determination (adjusted-R^2) which takes into account the number of variables in the created model. A "penalty" is included for the addition of independent variables. A point can be reached where additional parameters may be significant, but the number of variables may not improve the adjusted-R^2 value. The typical improvement seen in the model adjusted-R^2 value was on the order of 0.002 – 0.01, i.e. improving an adjusted-R^2 from 0.700 to 0.701 or 0.71. In some cases a minor improvement in the value of the root mean square was also seen. At some locations, five or more of the meteorological parameters were significant, but together they only improved the explanation of variation by the model in a minor way.

The conclusion of this preliminary investigation was that evaporation surrogate data are available, at least for a limited time period, and should be evaluated further for inclusion in the models. The evaporation-related meteorological data used in this preliminary investigation indicate that evaporation may be important as a minor player compared to the other independent variables that currently comprise the MLR salinity models: Everglades water level (stage), sea surface elevation (Key West water level), and wind vectors. Even so, including evaporation in the models may help the models to better simulate the higher salinity conditions, although the MLR salinity models are not particularly biased towards underestimation of higher salinity values.

Evaluation of the various methods presented above for estimating evaporation, evapotranspiration, and/or potential evapotranspiration determined that most of the methods required weather data that are not readily available on a long-term basis, over the entire Florida Bay and southwest Gulf coast region, and particularly not for 36 years (1965-2000) as needed for CERP evaluations. However, the self-calibrating Simple method for evapotranspiration (Abtew et al, 2003; Allen 1997, Price et al, 2005) uses only the difference between the maximum and minimum temperatures of the air and solar radiation as estimated by average values based on distance from the equator (latitude) (Allen, 1997). The equations that are used for the Simple method to estimate evapotranspiration as a surrogate for evaporation follow (Irizarry-Ortiz, 2003):

$$ET_p = \frac{K_1 * R_s}{\lambda}$$

ET_p: potential evapotranspiration [mm d^{-1}]
K_1 : coefficient
R_s : solar radiation received at the land surface [MJ m^{-2} d^{-1}]
λ : latent heat of evaporation [MJ kg^{-1}]

$$R_s = \tau R_a = K_r (T_{\max} - T_{\min})^{0.5} R_a$$

R_s : solar radiation received at the land surface [MJ m^{-2} d^{-1}]
τ : atmospheric transmissivity
K_r : empirical coefficient
T_{max} : mean daily maximum temperature over the period of interest [°C]
T_{min} : mean daily minimum temperature over the period of interest [°C]
R_a : extraterrestrial solar radiation [MJ m^{-2} d^{-1}]

$$R_a = \frac{24*60}{\pi} G_{sc} d_r (\omega_s \sin\varphi \sin\delta + \cos\varphi \cos\delta \sin\omega_s)$$

Ra : extraterrestrial solar radiation [MJ m^{-2} d^{-1}]
G_{sc} : solar constant = [MJ m^{-2} min^{-1}]
d_r : relative distance from the sun to the Earth
ω_s : sunset hour angle [rad]
φ : station latitude [rad]
δ : declination of the sun [rad]
J : Julian day of the year

$$d_r = 1 + 0.033\cos(\frac{2\pi J}{365})$$
$$\delta = 0.409\sin(\frac{2\pi J}{365} - 1.39)$$
$$\omega_s = ar\cos(\tan\varphi \tan\delta)$$

Therefore, it can be seen that the potential evapotranspiration (as modeled) is directly proportional to the solar radiation received at the land surface and inversely proportional to the latent heat of evaporation, as empirically calibrated. The solar radiation received at the land surface is estimated by the long-term extraterrestrial solar radiation and the square root of the difference in maximum and minimum air temperature values, as empirically calibrated. The extraterrestrial solar radiation is a function of the distance, angle and declination of the sun on a particular day of the year, and the latitude of the subject station. The empirically derived coefficients (K_1 and K_r) for the SFWMM activities were estimated as 0.53 for K_1 (mixed marsh, open water and shallow lakes) and 0.19 for K_r (coastal regions). It was noted that K_r was selected so that the long-term annual average potential evaporation matched an expected north to south gradient in south Florida, with higher values near the coast. The highest value of K_r is shown to be 0.203 at Flamingo City. However, the recommended coastal value of 0.19 was used for this exercise. The value of K_1 that was used was the recommended value of 0.53 (Irizarry-Ortiz, 2003), though Price et al (2005) have also calibrated this Simple model with evaporation measurements from the open water area of Florida Bay and estimated the value of K_1 at 0.63. The value of the solar constant that was used for this investigation was 0.08202 MJ m^{-2} min^{-1}

which is different from the value of this constant presented in Irizarry-Ortiz, 2003. The maximum and minimum air temperature values were taken from measurements at Key West because of the length of the record from that weather station.

Figures 1, 2, and 3 present 36-year reconstructions of daily, monthly, and annual potential evapotranspiration (etp) simulations at Joe Bay produced by the Simple method as described above. Monthly (Figure 2) and yearly (Figure 3) etp were computed by aggregating the daily simulations. As can be seen, the daily etp estimates ranged from 1 to 6 mm/day. Monthly etp ranges from about 60 to 140 mm/month, and the annual total ranges from 1050 to 1250 mm/year over the period 1965 to 2000. This compares favorably with the values for monthly and annual evaporation presented in Cosby et al, 2005 for Florida Bay MFL FATHOM model input data. There is little spatial variation in etp (Figure 4), though the etp at the northern near shore station (Joe Bay) has slightly lower yearly etp estimates than the southern open water station (Bob Allen Key).

Figure 1. Simulated daily values of potential evaporation (etp) at Joe Bay weather station produced by the Simple method (Abtew, et al, 2003)

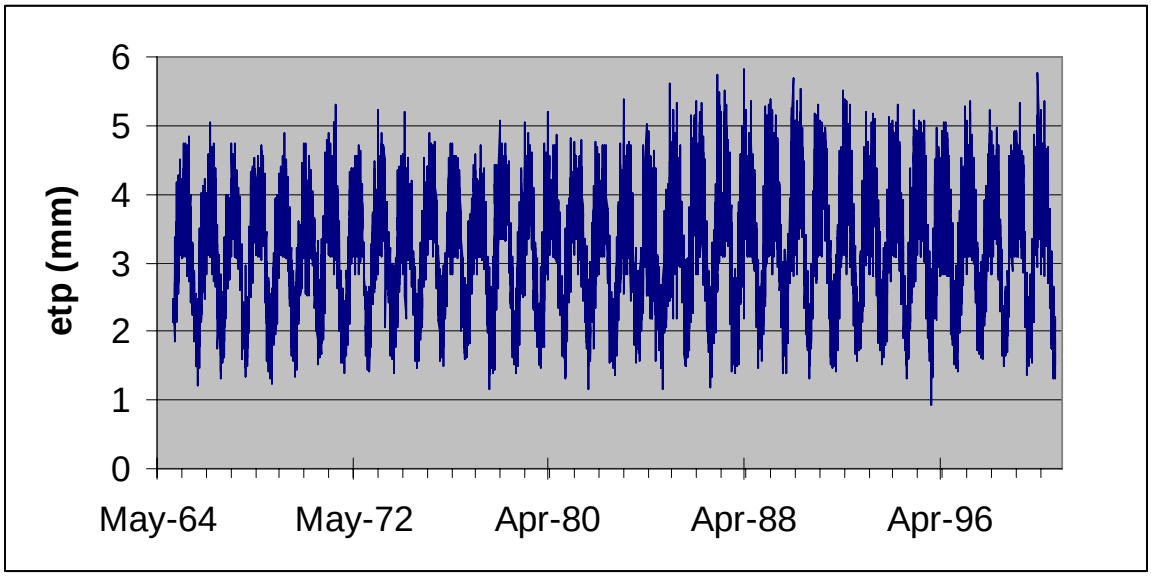


Figure 2. Simulated monthly values aggregated from daily values at Joe Bay weather station produced by the Simple method (Abtew et al, 2003).

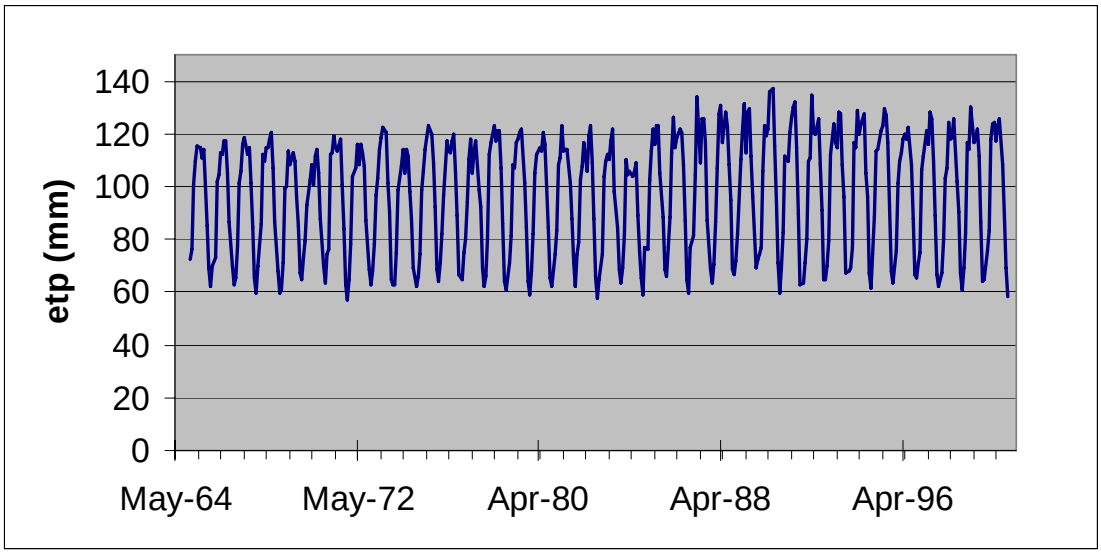


Figure 3. Simulated annual potential evapotranspiration aggregated from daily values at Joe Bay weather station produced by the Simple method (Abtew et al, 2003).

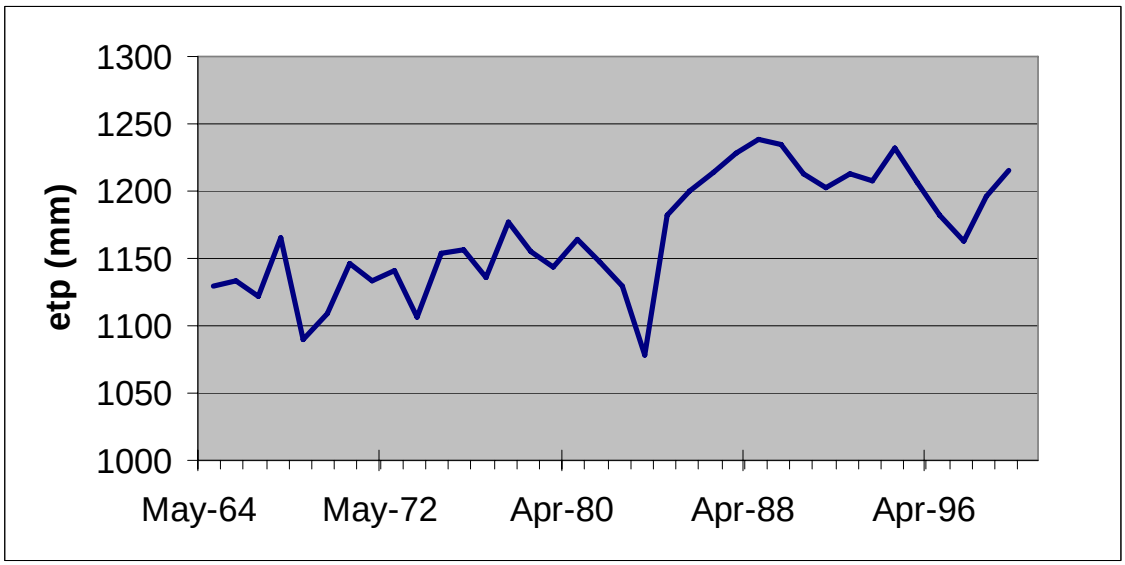
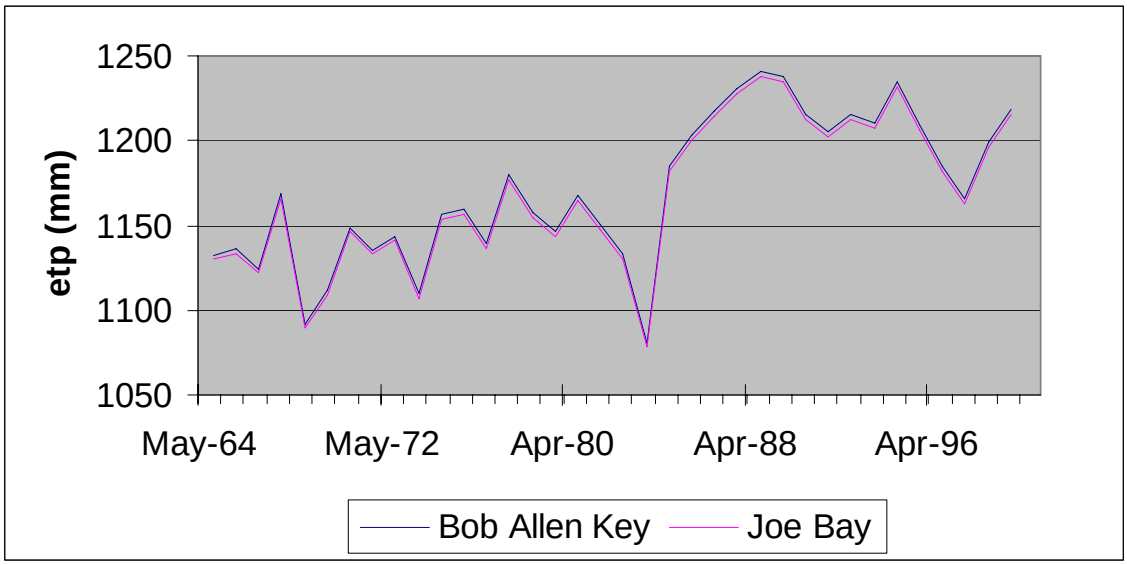


Figure 4. Comparison of Yearly ETP at Joe Bay and Bob Allen Key for 1965 – 2000.



The daily value of the potential evapotranspiration computed from the Simple method equations was used as a surrogate for the daily evaporation in evaluating the effect on the MLR salinity models of including salinity as a candidate independent variable. Unlagged and lagged values were submitted to the same step-wise regression procedure that was previously used for model development at the significance level of 0.999 needed for a candidate independent variable to be retained in the final MLR salinity model. For this evaluation there were three possible outcomes:

(1) evaporation was not a significant independent variable in the explanation of salinity variation when compared to the independent variables that are already in the model,
(2) evaporation as estimated in this manner may be an additional significant variable to those already determined to be significant, or
(3) evaporation may replace one or more of the already established significant parameters.

The primary measure of improvement, as it was for model goodness-of-fit and the preliminary evaporation surrogate evaluations, is the adjusted-R^2 value with the root mean square (RMS) error value as another measure of model improvement. To accomplish this evaluation, a complete model development procedure was undertaken for each station where MLR salinity models have been developed as of the date of this evaluation (19 stations).

Table 1 summarizes the results of the full-scale evaporation surrogate model development activities. As can be seen, at most locations evaporation was a significant variable at the 0.999 level of significance. At Long Sound, Duck Key,

and Highway Creek stations evaporation was not a significant variable. At 9 of 16 stations, other significant variables were eliminated from the previously derived models, while at the other 7 stations evaporation was added as a significant independent variable without the elimination of one of the previously derived significant independent variables.

When the adjusted-R^2 values are compared before and after including evaporation, 8 of 16 models showed a slight improvement in the adjusted-R^2 value, with 5 of 16 models showing a reduction in adjusted-R^2, and 3 of 16 models having an equivalent adjusted-R^2 value to the value before evaporation was considered. The most frequent independent variable that was replaced was one of the wind vector parameters, usually a lagged value of the parameter that explained very little (relatively) of the variation in salinity. Further evaluation of the SAS© step-wise output showed that evaporation was never the most significant variable, similar to the results of the preliminary evaluation using the estimated potential evapotranspiration at the Joe Bay weather station (also computed using the Simple method). The greatest increase in the value of adjusted-R^2 was seen at Butternut Key (0.65 to 0.70, a 0.05 increase), followed by a 0.03 increase at Bob Allen Key and Manatee Bay. This is an increase of about 8%, 3.8%, and 4.4%, respectively.

Table 1. Summary of statistics comparing models that included evaporation as a significant variable with previously developed models that did not include evaporation.

MODEL	NO EVAP R-SQ	EVAP R-SQ	NO EVAP RMS	EVAP RMS	VARIABLES ELIMINATED
Joe Bay	0.75	0.76	5.1	4.9	none
Long Sound	0.8	n/a	3.9	n/a	evap not significant at 0.99 level
Little Madeira Bay	0.65	0.58	6.4	4.5	none
Terrapin Bay	0.75	0.71	5.7	6.5	none
North River	0.77	0.76	3.8	4.1	NP206(lag2)
Duck Key	0.71	n/a	3.1	n/a	evap not significant at 0.99 level
Butternut Key	0.65	0.7	3.3	3.13	ltmad(lag1), terbay(lag3)
Whipray Basin	0.8	0.81	2.7	2.7	terbay,vwndkw
Bob Allen	0.79	0.82	2.7	2.4	uwndkw, vwndkw(lag2)
Highway Creek	0.81	n/a	4.3	n/a	evap not significant at 0.99 level
Little Blackwater Sound	0.75	0.76	3.7	3.6	none
Taylor River	0.78	0.78	4.6	4.6	none
Manatee Bay	0.69	0.72	3.1	2.9	uwndkw(lag2), vwndkw(lag1), uwndmia(lag2)
Middle Key (Barnes Sound)	0.74	0.74	2.6	2.6	cp, uwndkw(lag2, uwndmia(lag2), little blackwater (lag3)
Whitewater Bay	0.74	0.73	3.1	3.2	uwndkw(lag2), kwwatlev(lag1)
Clearwater Pass	0.85	0.85	3.4	3.3	uwndkw, uwndmia(lag1)
Gunboat Island	0.85	0.84	3.4	3.3	none
Shark River	0.82	0.8	2.5	2.5	none
Garfield Bight	0.68	0.69	6.1	5.9	uwndkw(lag4), uwndmia(lag4)

With respect to the value of the root mean square (RMS) error (which has the same units of salinity and is often considered as a “tolerance” value), there were 9 reductions in the RMS error value out of 16, with 3 increases in RMS error, and 4 models where the RMS error remained essentially the same. The greatest improvement (a lower RMS error value) was seen at Little Madeira Bay where the RMS error value decreased by 1.9 psu. However, where there was an improvement, the average improvement (not including Little Madeira Bay) was only 0.2 psu. The greatest increase in RMS error, indicating an increase in uncertainty, was at Terrapin Bay where the RMS error value increased by 0.8 psu.

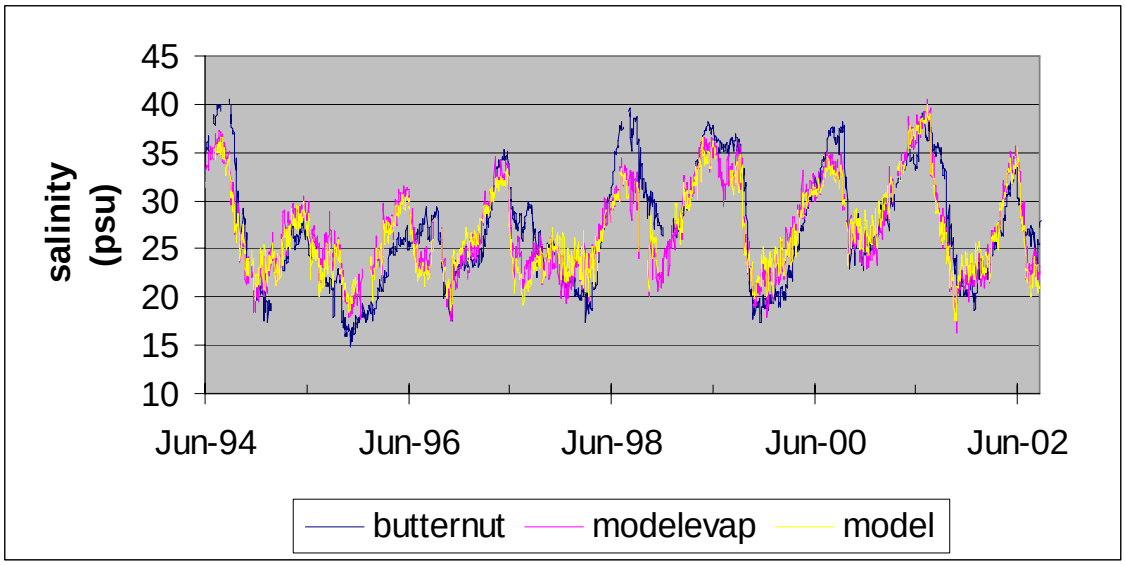
From a spatial perspective, there were no consistent results in improvement or degradation of the previously derived MLR model goodness-of-fit when evaporation was included in the model. It was expected that evaporation may provide an important improvement at the open water stations and improvement was seen at Butternut Key and Bob Allen Key, but at Whipray Basin there was neither improvement nor degradation even though evaporation replaced two previously significant variables. At Duck Key evaporation was not a significant variable at the 0.999 level, though it would have been at a significance level slightly below 0.999.

At Whitewater Bay and Clearwater Pass, also open water stations but located in the Shark Slough discharge region on the southwest Gulf coast as opposed to Florida Bay proper, there is virtually no difference in the goodness-of-fit of the models between the models that include evaporation and the models that do not include it. The results at Little Madeira Bay that indicate an important improvement in both adjusted-R^2 and RMS error values must be considered with caution, as the previously-derived Little Madeira Bay MLR salinity model was a relatively unstable model compared to the other 18 MLR salinity models that have been developed, and a difference in the model structure was seen during model development depending upon the length of the data record that was used for model development, a phenomena that was not seen at other stations.

It was thought at the onset of this investigation that evaporation may provide better simulations of high salinity periods. Figure 5 presents a comparison of the salinity simulations produced by the model that includes evaporation, the simulations produced by the original model, and the observed values for Butternut Key. Butternut Key was chosen for this comparison because the Butternut Key model with evaporation included showed the greatest increase in adjusted-R^2 and greatest decrease in RMS error value. It can be seen that the simulations produced by the model that includes evaporation are very similar to the simulations produced by the model without evaporation in that the evaporation model does not do much better than the non-evaporation model in meeting the highest salinity values when the previously derived model did not simulate well. For example, during the high salinity period of 1998, when the salinity values measured at the Butternut Key station reached close to 40 psu,

neither model does a very good job of simulating during this period. In contrast, during the high salinity period of 2001, both models do well in predicting the high salinity values that were measured, and the model with evaporation does not appear to do any better than the model without evaporation. Based on this limited evaluation, it does not appear that including evaporation in the models will help in simulating high salinity periods.

Figure 5. Comparison of model simulations with and without evaporation in the model and observed salinity values at Butternut Key.



III. Investigation of Direct Rainfall as a Significant Independent Variable

The second part of this investigation evaluated direct rainfall on the Bay as an independent variable in the MLR salinity models. For the original model development tasks (Marshall, 2003), rainfall as measured at the long-term weather data monitoring stations located on land was evaluated with other candidate independent variables. It was found that correlation of salinity with rainfall as measured at Flamingo, Royal Palm, and Tavernier monitoring stations was not very strong, and other parameters (water level in the Everglades, sea surface elevation, and wind) were capable of explaining a much greater portion of the variation in salinity such that rainfall was not a significant independent variable at the 0.999 level of significance.

Discussions of the MLR salinity models since the time of development has included the question of whether or not direct rainfall in Florida Bay, as opposed to rainfall measured at land-based stations near the south Florida mainland, may improve the MLR salinity models. Therefore this evaluation of direct rainfall on the Bay was undertaken. One of the most important requirements of the independent variables in the MLR salinity models was that data to conduct long-

term simulations must be available, either through direct observations (measurements) or through models that could be built using longer-term monitoring station. Therefore, before direct rainfall on Florida Bay can be considered as a candidate independent variable in the MLR salinity models, it must be directly or indirectly available.

Rainfall data from the Joe Bay, Flamingo, Tavernier, and Royal Palm weather stations were gathered into a database for evaluation. A plot of rainfall at all of the stations for the period of this evaluation (June 1991 – June 2001) is presented as Figure 6. Visually it is difficult to ascertain any correlations in the data. Even when the data from the first half of the evaluation period is plotted alone (Figure 7), and the range of values on the abscissa is reduced substantially, it does not appear that there are any correlations between Royal Palm, Flamingo, and Tavernier rainfall and the rainfall at Joe Bay at this daily resolution. In fact, when a multivariate linear regression model is developed at the daily time step using Royal Palm, Flamingo, and Tavernier rainfall to explain the variability in Joe Bay rainfall, the adjusted-R^2 value is poor, 0.39 which means that the "best" regression model for Joe Bay rainfall as developed from the data at these three stations is only capable of explaining about 39% of the variability in Joe Bay rainfall. To be able to be used for MLR salinity modeling at the daily level with independent variables that are significant at the 0.999 level, much better explanatory capability is needed for Joe Bay or any other weather station located on the Bay to be used on a long-term basis, even if the actual correlation between Joe Bay rainfall and salinity were quite high, which is not the case.

Figure 6. Comparison of daily rainfall at the Joe Bay, Royal Palm, Flamingo, and Tavernier weather stations for the period June 1991 – June 2001.

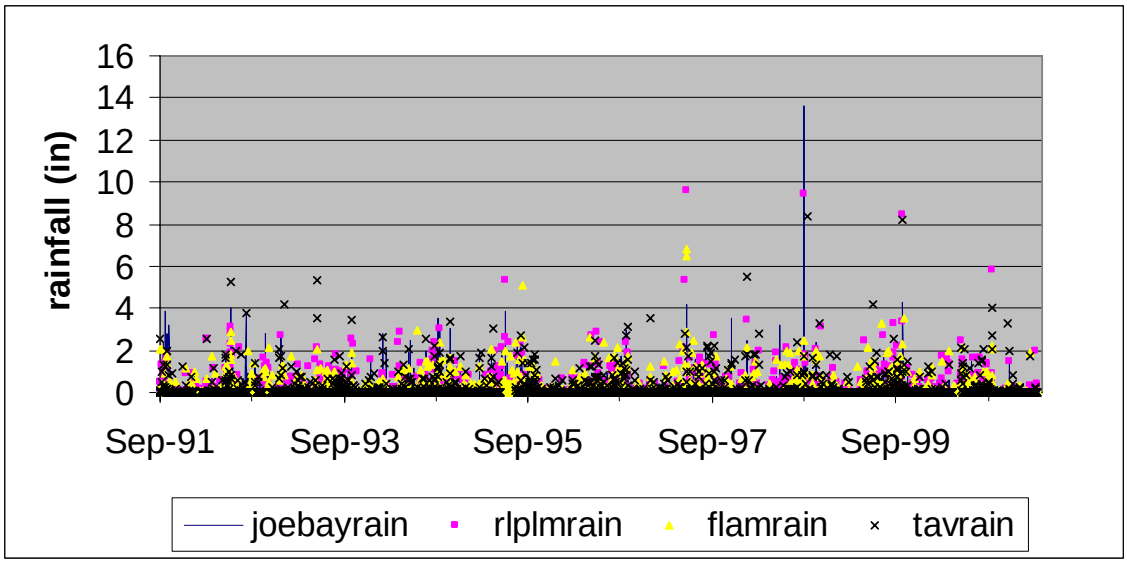
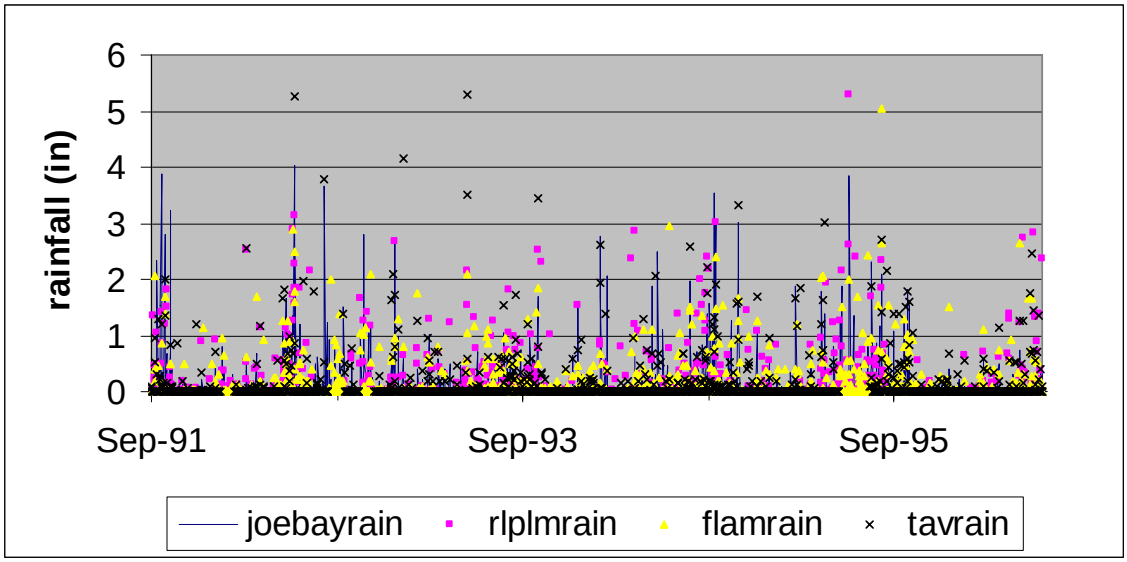


Figure 7. Comparison of daily rainfall at the Joe Bay, Royal Palm, Flamingo, and Tavernier weather stations for the period June 1991 – December 1995. Note that the scale on the rainfall axis is different from the previous figure.



However, when the daily values for rainfall at all four stations are aggregated into monthly values, relationships begin to appear (Figure 8). At the monthly time step, a general pattern can be discerned between all of the four stations throughout the year. When a multivariate linear regression model is developed for Joe Bay monthly rainfall as a function of the monthly rainfall at Royal Palm, Flamingo, and Tavernier, the explanatory level is significantly higher, with an adjusted-R^2 value of 0.68. Figure 9 presents a plot of the simulated rainfall from this regression equation compared to the observed values for Joe Bay rainfall. There are a number of missing values in the simulations due to missing data at one or more of the three stations. However, in general the fit at the monthly resolution of the model is good but not excellent.

Figure 8. Comparison of monthly rainfall values aggregated from daily values for Joe Bay, Royal Palm, Flamingo, and Tavernier weather stations.

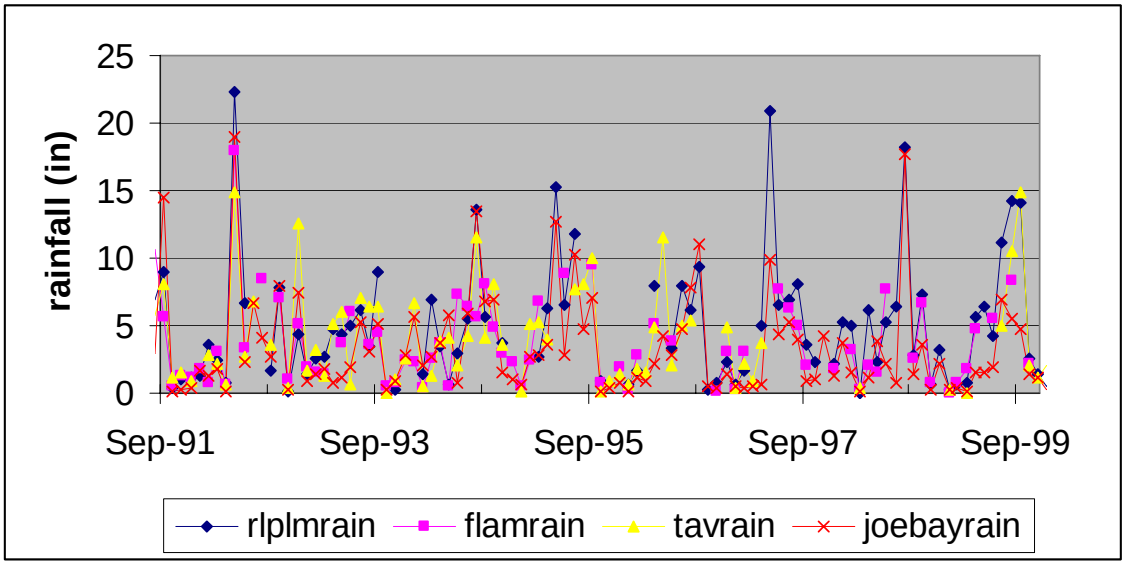
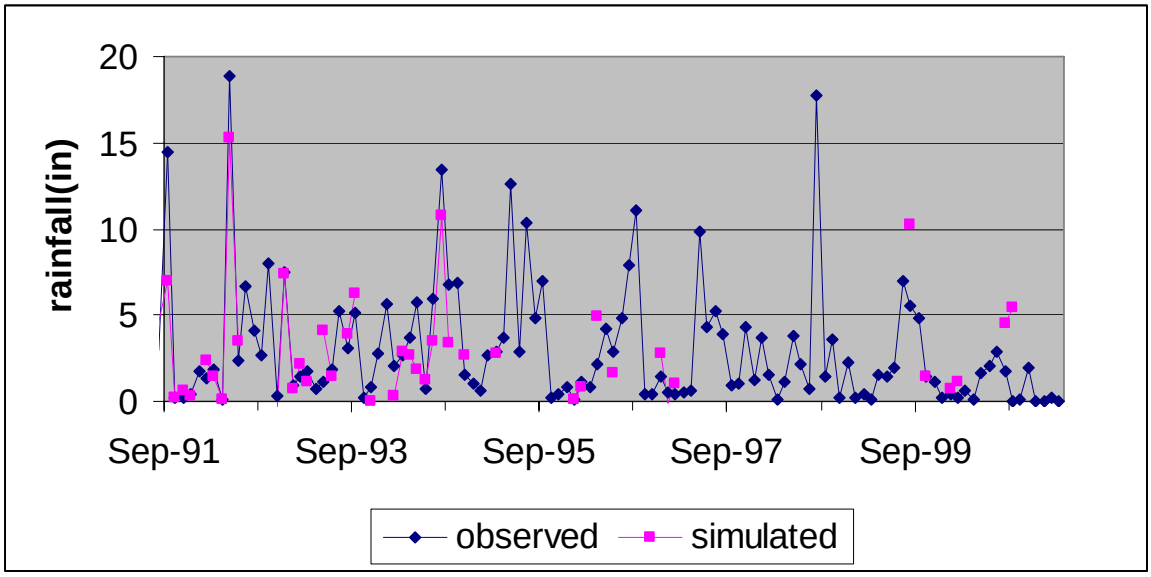


Figure 9. Comparison of monthly rainfall values aggregated from observed daily values and monthly values simulated by the multivariate linear regression model (adjusted-R^2 = 0.68).



V. Discussion and Summary

One of the primary objectives of this task is to re-visit consideration of evaporation as a candidate independent variable for MLR salinity models. In previous model development activities, the focus has been to find enough existing observed data to develop sound MLR salinity models at a high level of significance for long-term salinity simulation use. There are no long-term data available in the inventory of existing evaporation data for the south Florida region, though there are synoptic studies. Relative to the amount of observed data and the length of record for Everglades water levels, Key West sea surface water level, and wind (at Miami and Key West), the observed evaporation data are practically non-existent, particularly in Florida Bay and the mangrove fringe of he southwest Gulf coast. Therefore, evaporation was not included amongst the candidate independent variables during the MLR salinity model development tasks. However, it is well-established through water budget estimates and observations of hypersaline conditions that evaporation plays an important role in determining the salinity regime in parts of Florida Bay and in areas along the southwest Gulf coast. Because of this, it was decided that a further investigation of evaporation surrogates was warranted in the event that an indirect relationship with long-term climatic data could be established that would allow evaporation to be considered for inclusion in the MLR salinity models.

Various methods for estimating evaporation were examined. The SFWMD (Abtew et al, 2003; Irizarry-Ortiz, 2003) and others (Price et al, 2005) have found success using a self-calibrating method of estimating solar radiation, the key component of evaporation, and using it with maximum and minimum air temperature values to estimate the potential evapotranspiration, which can, in turn, be used as an evaporation surrogate for MLR salinity model tasks. This “Simple” method of estimating potential evapotranspiration is used at several weather stations by SFWMD, and values for the Joe Bay weather station are available on DBHYDRO, computed from the Joe Bay weather station air temperature values. A preliminary investigation using salinity data from a limited number of stations showed that there was some correlation between salinity and Joe Bay potential evapotranspiration computed using the “Simple” method.

Therefore, SAS© code was written to compute an estimate of potential evapotranspiration using the Simple method at each of the 19 stations that currently have MLR salinity models. Solar radiation was estimated using the latitude of the station and the time of the year. Temperature difference was based on the maximum and minimum daily air temperature value measured at the Key West weather station, which has a record of daily air temperature much longer than the 1965-2000 period needed for CERP evaluations. Historical reconstructions for the 1965-2000 period using the Simple method simulations indicate that during the period 1985 – 1992, a drought period, etp was higher than the previous 19 years. During this period salinity values, in general, over the Bay were also high. However, the high salinity period of the early 1970’s is

not expressed distinctly in the etp values. This was a period of low flow and the average evaporation at that time contributed to the hypersaline conditions during that period that were observed.

A salinity model development activity was then undertaken in which the estimated potential evapotranspiration (etp) was submitted to the step-wise model development process along with the parameters that had previously been determined to be the best independent variables for salinity simulations at each station. In all but 3 of the 19 cases, etp was shown to be a significant independent variable, sometimes replacing other independent variables and sometimes becoming an added independent variable. However, in all case where etp replaced an existing independent variable, the independent variable that was replaced only explained a small portion of salinity variability (usually less than 1%), usually being a lagged value, and the etp replacement also only explained a small portion of salinity variability. For some models the goodness-of-fit improved with the inclusion of the etp, and for others the replacement by etp degraded the adjusted-R^2 and root mean square error values. The greatest model improvement occurred at Butternut Key, where the new model with etp explained an additional 5% of the variation in salinity, with about 3% explanatory improvement at Bob Allen Key and Manatee Bay, based on adjusted-R^2 values. When etp was an added independent variable, such as at Terrapin Bay, model adjusted-R^2 and RMS error values were reduced because the model now included additional terms that did not explain enough additional salinity variability to overcome the penalty of additional terms added to the regression model.

Additionally, the model development results using this evaporation surrogate were somewhat inconclusive at the open water stations where it was expected that evaporation would improve the models even if no improvement was seen at the near shore stations. Although model improvement was seen at Butternut Key and Bob Allen Key with very slight improvement at Whipray Basin, etp was not a significant independent variable at Duck Key. Additional investigation with other open water salinity models (to be developed in the future) may prove that evaporation should be one of the independent variables in open water MLR salinity models. However, a comparison of the simulated values for salinity at Butternut Key (where the greatest improvement in model fit was seen using evaporation surrogates) estimated from the model that included etp with the values estimated from the previously derived MLR salinity model without etp indicates that the model with etp does no better at estimating high salinity values than the MLR salinity model that does not include etp.

Based on these results, it does not appear that including evaporation computed from the “Simple” method in the MLR salinity models is useful and that the models that have already been derived and are currently in use produce salinity simulations that are very similar to the salinity simulations produced by models that include evaporation surrogates.

With respect to rainfall directly on the Bay as compared to rainfall measured at a land-based meteorological station that may be remote to the Bay, previous attempts to correlate daily rainfall with salinity have not produced satisfactory models. For one reason, at the daily level, rainfall is an almost all-or-none discontinuous quantity, not a continuous independent variable like Everglades stage, sea surface elevation, and wind. In some manner, Everglades stage and sea surface elevation indirectly expresses the effect of rainfall accumulation. Because it is thought that rainfall on Florida Bay is spatially variable, to utilize direct rainfall as a candidate independent variable requires at least regional estimates of rainfall over a long-term period, or indirect data and a transfer function to relate rainfall at other stations to rainfall at a location on the Bay. The long term rainfall data that are available at Royal Palm, Flamingo, and Tavernier were used to produce a multivariate linear regression for rainfall measured at Joe Bay weather station, although the fit of the model at the daily level was poor, explaining only about 39% of the variation in Joe Bay rainfall, not acceptable for use with MLR salinity models as a rainfall surrogate. However, when the daily values are aggregated into monthly totals, a much better and acceptable multivariate linear regression model can be produced (adjusted-R^2 = 0.68). Therefore, rainfall is spatially discrete at the daily time step but becomes somewhat spatially continuous at the monthly step.

Therefore, even if daily rainfall at the Joe Bay weather station turned out to be a significant independent variable, it is not possible to extend the period of data by using the daily rainfall at other long-term weather stations because a reasonable transfer model can not be developed. Previous attempts to use daily rainfall to model salinity have also not been successful, though there is potential at the monthly resolution.

VI. Recommendations

Based on the evaluations that were completed as discussed above, the following recommendations are made:

1. It appears that using the “Simple” method to estimate potential evaporation using Key West air temperature provides reasonable estimates of evaporation at the daily level.
2. For near shore and Gulf coast models, evaporation is not needed in the MLR salinity models and should not be incorporated.
3. For the open water areas, additional evaluations of evaporation are needed when more open water models have been produced (2006).
4. It does not appear to be feasible at this time to use direct rainfall on the Bay for daily MLR salinity modeling.

VII. References

Abtew, W., J. Obeysekara, M. Irizarry-Ortiz, D. Lyons, A. Reardon. 2003. Evaporation Estimation for South Florida. Technical Paper EMA # 47, Proceedings of the World Water and Environmental Resources Congress 2003.

Allen, R. 1997. Self-calibrating Method for Estimating Solar Radiation from Air Temperature. Journal of Hydraulic Engineering. American Society of Civil Engineers, Vol 2., No. 2, April 1997.

Cosby, B., W. Nuttle, and F. Marshall. 2005. FATHOM Enhancements and Implementation to Support Development of MFL for Florida Bay. Final Report on Contract C-C-15975-WO05-05 for the South Florida Water Management District. Environmental Consulting & Technology, Inc. New Smyrna Beach, Florida.

Irizarry-Ortiz, M. 2003. Selected Methodology for Long-Term (1965-2000) Solar Radiation and Potential Evapotranspiration Estimation for the SFWMM2000 Update. Memorandum to Ken Tarboten, December 12, 2003. South Florida Water Management District, West Palm Beach, Florida.

Marshall, F. 2004 Statistical Salinity Models for Florida Bay and the Southwest Gulf Coast,Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. CESI Program Proposal.

Price, R., W. Nuttle, B. Cosby, and P Swart. 2005. Estimating Evaporation Rates in Time and Space Across Florida Bay. Abstract for the Florida Bay and Adjacent Marine Systems Science Program Conference, Duck Key, Florida, December, 11-14, 2005. Page 126.

FINAL TASK REPORT

TASK 2 - USE OF THE FATHOM FLORIDA BAY MFL MODEL WITH SICS/TIME INPUT

Cooperative Agreement Number CA H5284-05-0006
Between
The United States Department of the Interior National Park Service
Everglades National Park
And
Cetacean Logic Foundation, Inc.

Frank E. Marshall, III
Task / Project Manager

B. J. Cosby
Project Consultant (FATHOM)

DeWitt Smith
Project Coordinator

April 12, 2008

CESI Project No. CA H5284-05-0006 for Everglades National Park

FINAL TASK REPORT

TASK 2 - USE OF THE FATHOM FLORIDA BAY MFL MODEL WITH SICS/TIME INPUT

I. Introduction

A. General

This report describes the activities that were completed for a task in the current Critical Ecosystem Studies Initiative (CESI) project that was originally initiated by Cetacean Logic Foundation, Inc. (CLF) for Everglades National Park (ENP) in 2002. The goal of this on-going CLF research is the characterization and simulation of the salinity regime in Florida Bay and the estuaries along southwest coast of the Gulf of Mexico (Gulf) within the Park. It is intended that this research yield information regarding the link between the downstream salinity in the estuaries of ENP and the upstream freshwater hydrology of the Everglades.

The subject of this report is the use of the mass-balance model FATHOM as developed for the Florida Bay Minimum Flows and Levels (MFL) work with flows from SICS/TIME, the USGS model of the Everglades, as input. Included in this task is the upgrading of the model basin configuration. This task report describes the FATHOM update that was done.

B. Task Objectives and Evaluation Methods

The objective of this task is presented below as it appears in the modified contract for this CESI project approved by ENP:

> "The Principal Investigator will utilize the calibration run output from SICS / TIME as input to the South Florida Water Management District (SFWMD) Minimum Flows and Levels (MFL) FATHOM model to produce a time series of salinity that can be compared to observed data from the ENP MMN at Joe Bay, Little Madeira Bay, Terrapin Bay, Garfield Bight, and Whipray Basin. The Principal Investigator will coordinate with ENP on the development of revised Florida Bay basin GIS coverages. The revised basin layout will be utilized with the SICS / TIME input and output salinity time series will be compared with the MFL output.
>
> The deliverable is a task report documenting the use of the MFL FATHOM with SICS / TIME output for Joe Bay, Little Madeira Bay, Terrapin Bay, Garfield Bight, and Whipray Basin, development of a revised basin layout, use of SICS / TIME output with the revised configuration, and comparison with the MFL output."

II. Background

Salinity is a fundamental and key characteristic of the physical conditions of estuarine and coastal ecosystems. Salinity affects water quality, the make-up and spatial distribution of vegetative communities, and the life history of most animal species in these ecosystems. Simulations and forecasts of salinity are an important tool in the assessment of ecological resources in the Everglades, Florida Bay, and the estuaries on the Gulf of Mexico (CROGEE, 2002). Water managers use forecasts to evaluate the expected benefits and impacts of ecosystem restoration activities. Ecosystem restoration involves aspects of adaptive management (NRC 2004), uncertainty analysis (CERP 2002), and risk assessment (Thom et al. 2004), and these all rely on the application of predictive models.

The Everglades / Florida Bay hydrologic system is unique because of the vast area of freshwater marshes underlain by porous substrate that stores runoff before it enters the estuarine zone, as well as the spatial extent of estuarine conditions in Florida Bay. Standard riverine hydraulic models can not account accurately for the spatial and temporal variation in stored water and dispersed flows in the Everglades. Therefore, freshwater hydrology and wetland basin models have been developed to simulate the south Florida conditions required for use with salinity models. For statistical models, hydrology of the Everglades is described by the stage levels that are used as input for salinity simulating and forecasting.

The use of modeled input data for salinity simulations by mechanistic models is necessary because the standard period for evaluations of water management alternatives spans a 36-year period and observed data for some model input are not fully available. This increases the level of uncertainty in the salinity estimates produced by hydrodynamic models. The use of a 36-year period for south Florida simulations is warranted by the significant difference in wet and dry periods over years to decades, and the ecological implications of anthropogenic alterations that may only be expressed over longer periods of change in the salinity regime.

Considerable progress has been made in the development and refinement of salinity models since the report in 2002 by the Cadmus Group (Nuttle, 2002). To-date, the most widely used models for developing historical recreations and simulating salinity regimes for the evaluations of water management alternatives are the FATHOM mass balance model and the MLR salinity models. Because of their relative simplicity, development has occurred before full hydrodynamic model development has been completed. Mass balance and MLR salinity models have already been applied in a number of ways and are still being refined. However, the development of hydrodynamic models, particularly SICS/TIME and EFDC, is continuing, and use for historical salinity reconstructions and simulations is likely within the next several years.

The subject of this report is the use of the mass-balance model FATHOM as developed for the Florida Bay Minimum Flows and Levels (MFL) work with flows from SICS/TIME, the USGS model of the Everglades, as input. Included in this task is the upgrading of the model basin configuration.

Background information on FATHOM and the FATHOM MFL model can be found in the Task 4 Final Report completed for this CESI cooperative agreement, and in the project report which described the implementation of the FATHOM Florida Bay MFL model (Cosby et al, 2005). This Florida Bay MFL model was modified in this task by changing the basin configuration. The modifications made to the model are described in the Methods section.

III. Methods

Task 2 was initiated when the ENP project coordinator ran the MFL FATHOM model (47-basin configuration) with runoff input from TIME output. This TIME run was produced by Park staff using FTLOADDS version 3.1 code and input files for TIME run 157 provided by USGS TIME developers. The FATHOM output was reviewed by Dr. Cosby to ensure that the input parameters were being specified properly when TIME output is coupled to FATHOM.

Next, runoff estimates for the TIMEcv01 calibration run were provided to the project team by ENP. These flows were used as input to the FATHOM mass balance model to produce salinity simulations throughout Florida Bay. The FATHOM salinity output using the TIMEcv01 runoff was compared to the salinity output from FATHOM generated by the MFL Base Case run. The FATHOM configuration used for this comparison was the original 47-basin model domain.

As an additional task, an updated bathymetry for a new 58-basin configuration of FATHOM was delivered by ENP to Dr. Cosby in early June 2007. The new bathymetry was processed to prepare the input files for FATHOM. The new bathymetry files were compared to the original bathymetry files in a detailed quality assurance check to assure that bathymetry of the unaltered basins was not affected. The new bathymetry files were then used in a simulation of the MFL base case for comparison with simulation of the same inputs using the 47 basin version of FATHOM.

At the end of the project period (September 30, 2007), the TIME model was still under development. Because of this output from a revised calibration / verification run was not available so the revised calibration / verification run output could not be used with the 58-basin configuration of FATHOM.

IV. Results

Figure 1 through 4 present the comparison of the FATHOM 47-basin configuration output using the TIME runoff input, compared to both SERC and MMN observed data. The observed data collected by SERC are monthly grab samples, and ENP MMN data are monthly average values. Selected statistics and plots are presented as Figures 2 and 4 for SERC comparisons, and as Figures 3 and 5 for MMN comparisons to FATHOM output produced using the TIME run for flow inputs.

Figure 1. Selected statistics for comparing simulated and observed (SERC data) salinity values using the 47 basin configuration of FATHOM. The simulation is the TIME simulation covering the period 1991-2002.

Figure 3. Comparison of simulated and observed (SERC data) time series of salinity values using the 47 basin configuration of FATHOM. The simulation is the TIME simulation covering the period 1991-2002. Selected basins are displayed.

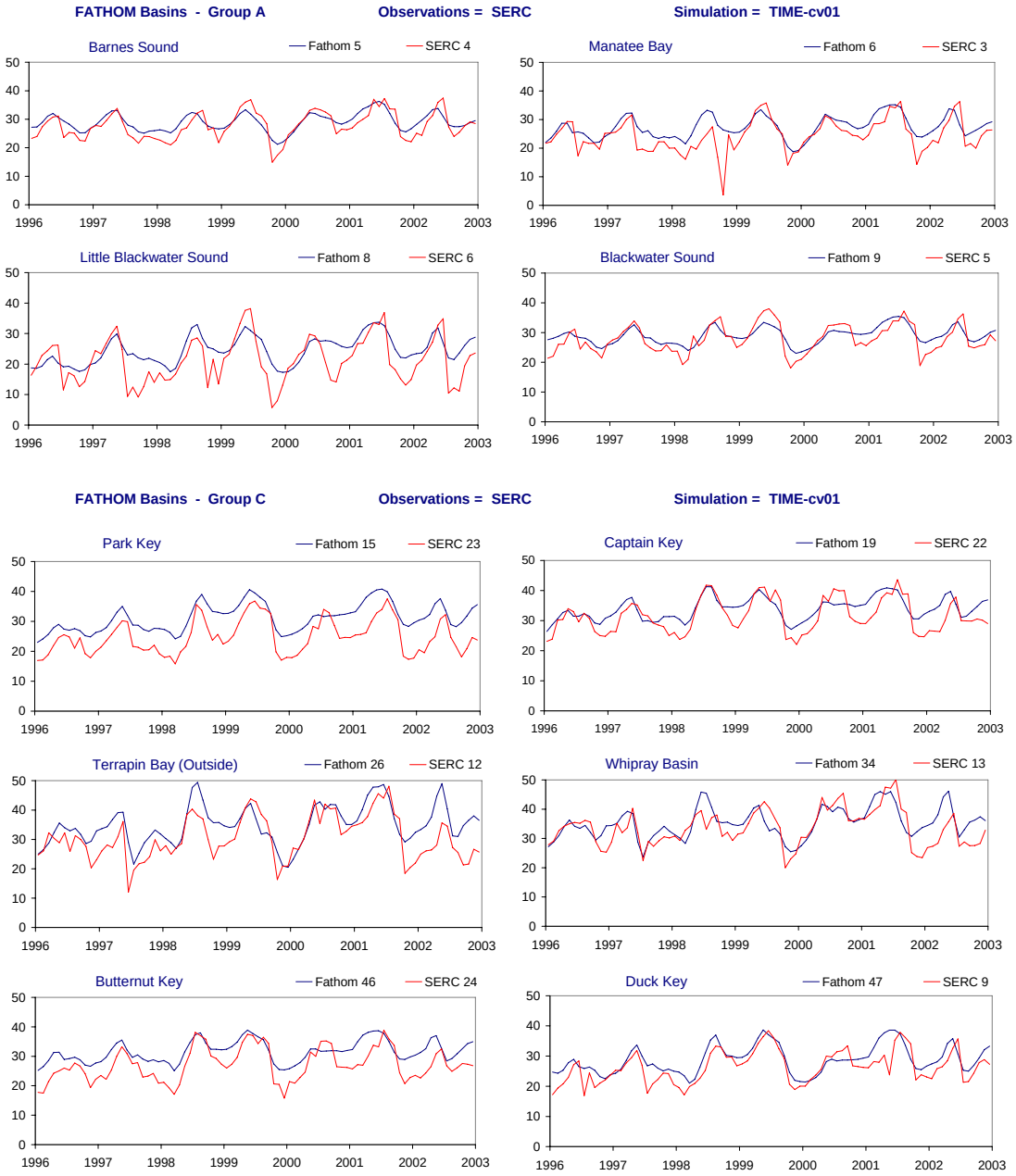


Figure 3. Selected statistics for comparing simulated and observed (MMN data) salinity values using the 47 basin configuration of FATHOM. The simulation is the TIME simulation covering the period 1991-2002.

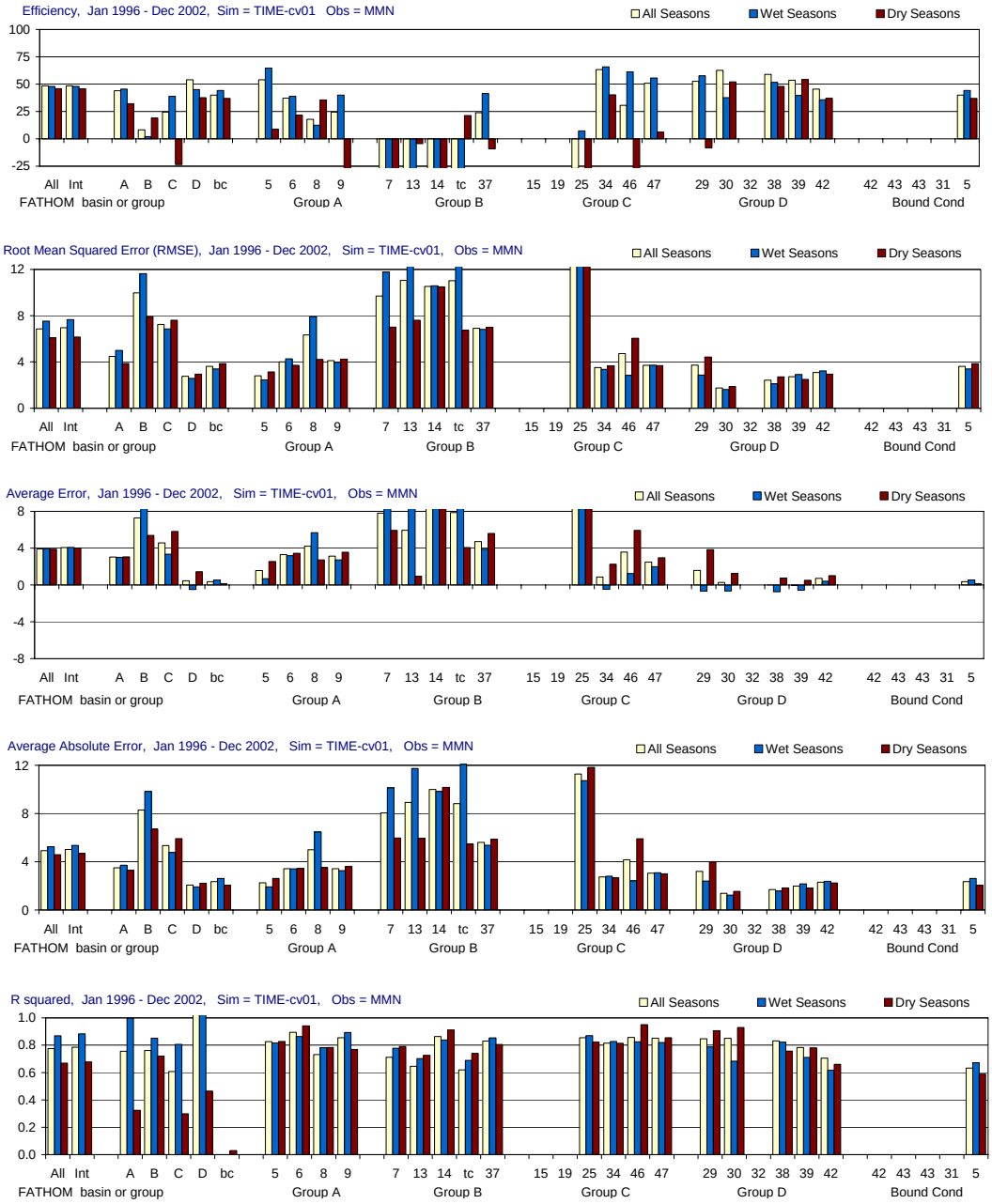
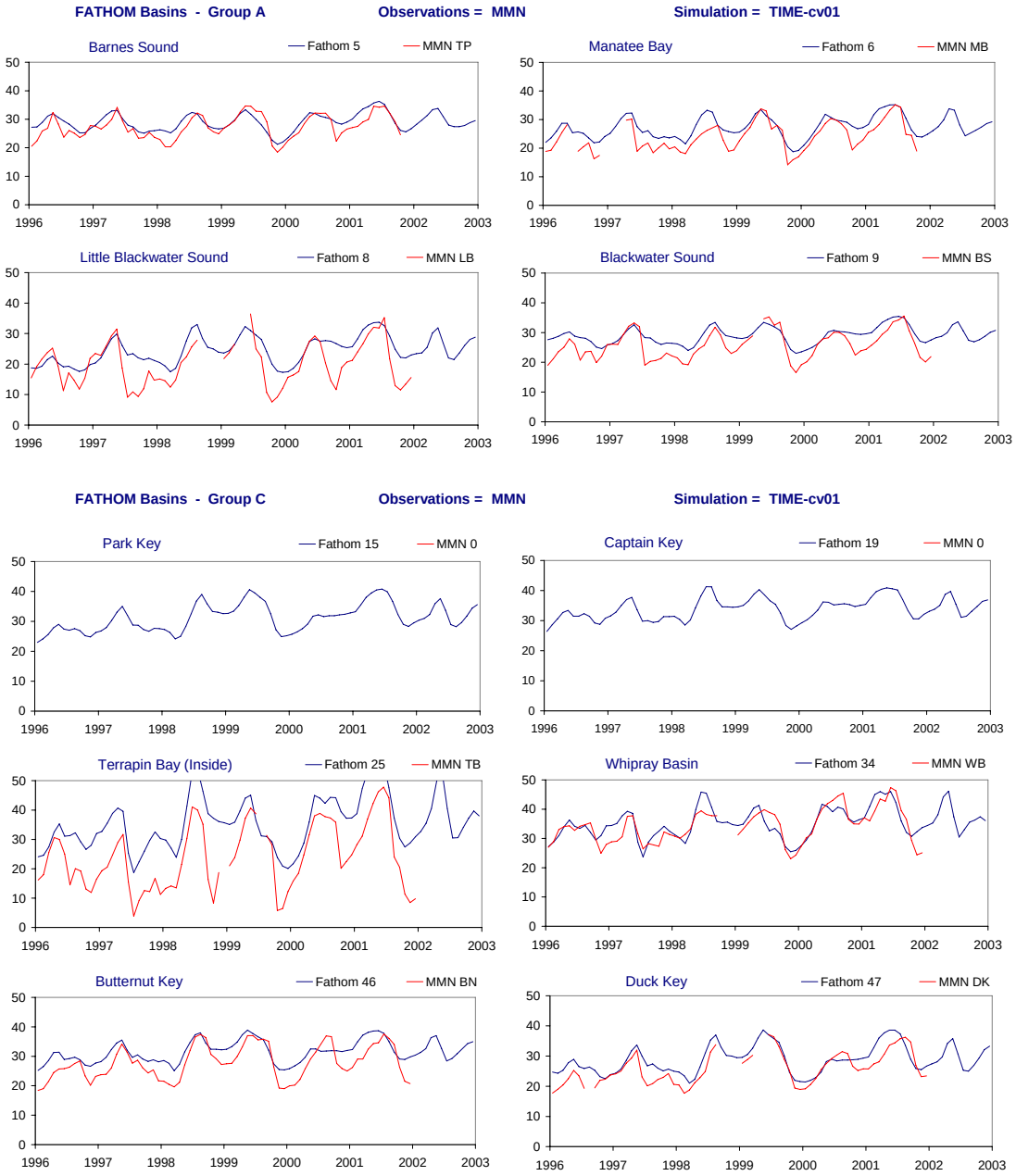


Figure 4. Comparison of simulated and observed (MMN data) time series of salinity values using the 47 basin configuration of FATHOM. The simulation is the TIME simulation covering the period 1991-2002. Selected basins are displayed



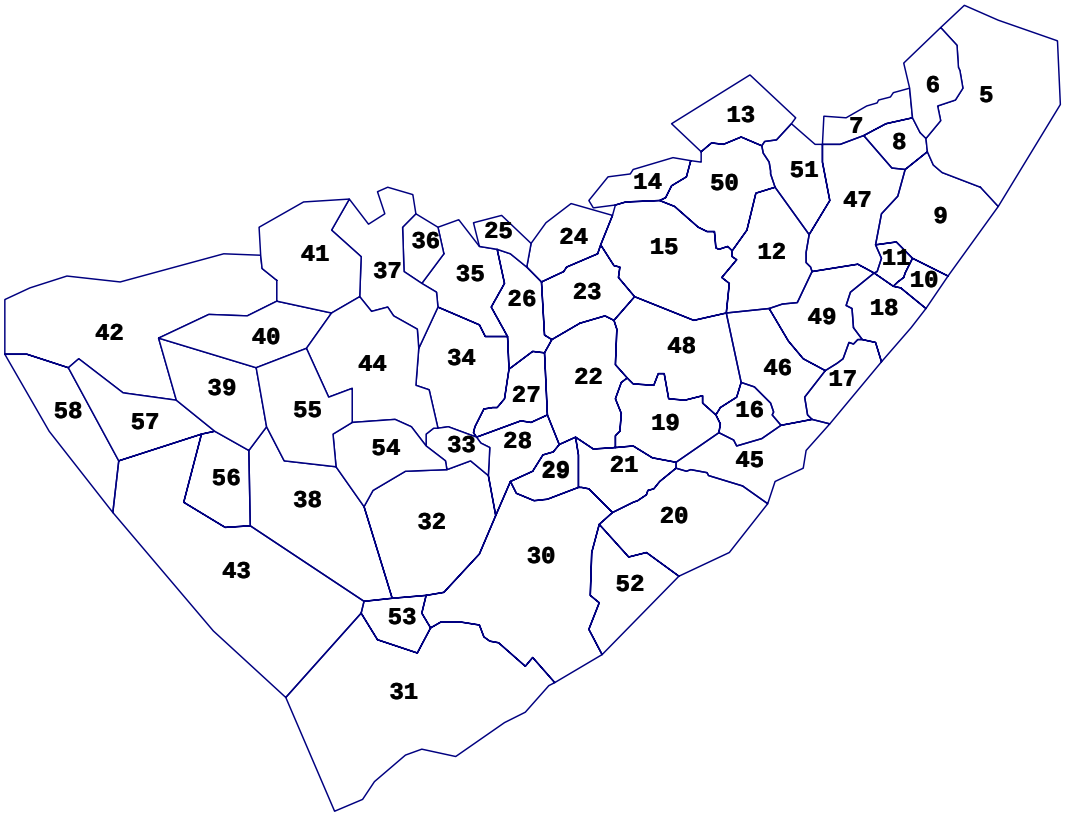
New Basin Configuration

As described the 47-basin configuration of the FATHOM model was reconfigured to 58 basins to:

- Better distribute runoff inputs, and
- Provide for additional mixing in near shore embayments and basins adjacent to boundary condition assignments.

The new 58 basin FATHOM configuration is shown in Figure 5.

Figure 5. The new 58 basin configuration for the FATHOM model.



The new 58 basin bathymetry files were inserted into the FATHOM code. FATHOM benchmark simulations were run to assure QA/QC with the new code. Then the final MFL base simulation using FATHOM with the new basin configuration was run. The overall model performance in simulating salinity with 58 basins was degraded compared to the 47 basin simulation (Figure 6).

Figure 6. Selected statistics for comparing simulated and observed (SERC data) salinity values using the 58 basin configuration of FATHOM. The simulation is the SFWMD MFL base simulation covering the period 1991-2002

V. Discussion

The FATHOM MFL model was run in the 47-basin configuration using FTLOADDS v 3.1 code executed by the Park using USGS 157 flow input to produce salinity simulations. The TIME-based salinity simulations were compared to SERC and MMN observed data at a monthly resolution. When the TIME-based salinity simulations were compared to the observed data (both datasets) a number of discrepancies were noted. A comparison of the inflows showed that the TIME runoff estimates are lower than the observed values. Therefore, though this task was successful in that the output from the TIME model was coupled with FATHOM to produce reasonable salinity estimates, it was not successful in producing reasonable simulations from TIME output, likely caused by problems in the TIME output and not the FATHOM modeling..

The bathymetry for the new 58 basin configuration of FATHOM was installed using a hybrid approach (as in the SFWMD MFL study). The bathymetry for those basins and shoals that were covered by the USGS bathymetry study was based on the trackline data. Bathymetry for other areas of the bay was based on the original ENP/FIU bathymetry (the areas east of Route 1 and the areas on the shallow banks in the northwest corner of the bay).

The benchmark tests for the installation were completed. The benchmarks consist of a series of simulations for which the model has been run in the old 47 basin configuration and for which model outputs and diagnostics are known. When applied to the new basin configuration, the benchmarks assure that the installation of the additional basins did not inadvertently alter other aspects of the model structure.

The model with the modified bathymetry was run with the SFWMD MFL inputs in two ways: 1) applying the Trout Creek flows to Davis Cove (western inputs); and 2) applying the Trout Creek flows to Trout Cove (eastern inputs). All other model inputs were applied as in the MFL study.

In the old 47 basin configuration (as applied in the MFL study), the western inputs of the Trout Creek flows (TCFs) produced a better agreement to observed SERC salinities in the eastern portion of the bay than did the eastern inputs of the TCFs. This was counter to the physical expectation – Trout Creek physically discharges in the eastern basin. The problem partially arose from the large size of the basins being used which allowed mixing to occur more rapidly than mixing to the west. The new 58 basin configuration, using smaller basins in the area to allow more discrimination of N-S and E-W gradients, was installed to alleviate this problem.

The new basin configuration was partially successful in this attempt. The differences in salinity between simulations using eastern and western inputs of

the TCFs are much smaller in the 58 basin configuration compared to the 47 basin configuration. In fact, the differences for the MFL inputs might almost be considered negligible (or at least unimportant). That is a positive result.

However, the overall model performance in simulating salinity with 58 basins is degraded compared to the 47 basin simulation. In general, the model efficiencies are smaller for the SERC sites in the northeast bay (but are essentially unchanged in the rest of the bay).

This negative result is not unexpected. In changing one set of inputs, it is reasonable to expect that other inputs might need to be adjusted in turn. The SFWMD MFL run (to which these new simulations were compared) was the "best" of a number of runs performed as part of a sensitivity analysis in that project. Changing the basin configuration and rearranging the TCFs has shifted the relative sensitivity of other inputs to the model. To be rigorous, it will be necessary to repeat the sensitivity analysis to examine the new pattern of effects of all inputs. A recent paper published by Nuttle et al (2007) suggests that evaporation should be increased by about 20% over that used in the SFWMD Florida Bay MFL FATHOM modeling activity.

VI. Findings and Recommendations

The task of coupling TIME model-produced flows with the FATHOM MFL model was partially successful. The flow outputs from a TIME calibration / verification run were used to produce salinity simulations but the simulated values did not compare well to the observed data. The problem was likely with the TIME data, and the problem could not be resolved during the period of this study. However, it was shown that FATHOM can successfully process TIME input.

The additional task implementing the new basin configuration in FATHOM has been completed. The new basin configuration runs successfully and corrects (at least partially) the conceptual problem of the location of the flow inputs from Trout Creek. In general, the model efficiencies are smaller for the SERC sites in the northeast bay, but are essentially unchanged in the rest of the bay. To analyze this result, it will be necessary to repeat the sensitivity analysis to examine the new pattern of effects of all inputs. However, modification of the model in this manner is beyond the scope of this task. Another CESI project is currently being initiated to begin this process.

In the meantime, to complete this task and use the TIME outputs with FATHOM the SFWMD Florida Bay MFL version of the model and inputs as already provided to ENP were utilized.

VII. References

Cosby, B.J., W.K. Nuttle, and J.W. Fourqurean. 1999. FATHOM – Flux Accounting and Tidal Hydrology at the Ocean Margin: Model Description and Initial Application to Florida Bay. Report to the Florida Bay Project Management Committee (PMC) and the Everglades National Park (ENP), National Park Service. Dept. of Environmental Sciences, University of Virginia, Charlottesville, VA.

Cosby, B.J., J.W. Fourqurean, and W.K. Nuttle, 2004. FATHOM Florida Bay Simulation 1991- 2002 Updated Model and Preliminary Results. Interim project report for Everglades National Park by Florida International University. January 2004.

Cosby, B., W. Nuttle, and F. Marshall. 2005 FATHOM Enhancements and Implementation to Support Development of MFL for Florida Bay. Final Report on Contract C-C-15975-WO05-05 for the South Florida Water Management District. Environmental Consulting & Technology, Inc. New Smyrna Beach, Florida.

Nuttle, W.K., J.W. Fourqurean, B.J. Cosby, J.C. Zieman, and M.B. Robblee, 2000. Influence of net freshwater supply on salinity in Florida Bay. Water Resources Research 36:1805-1822.

Price, R.M., W.K. Nuttle, B.J. Cosby, and P.K. Swart, 2007. Variation and Uncertainty in Evaporation from a Subtropical Estuary: Florida Bay. Estuaries and Coasts. Volume 30 3:497-506.

FINAL TASK REPORT

TASK 3 - DEVELOPMENT OF ADDITIONAL MULTIVARIATE LINEAR REGRESSION SALINITY MODELS FOR FLORIDA BAY AND THE SOUTHWEST GULF COAST, EVERGLADES NATIONAL PARK

Cooperative Agreement Number CA H5284-05-0006
Between
The United States Department of the Interior National Park Service
Everglades National Park
And
Cetacean Logic Foundation, Inc.

Frank E. Marshall, III
Principal Investigator

DeWitt Smith
Project Coordinator

January 5, 2008
Revised August 5, 2008

CESI Project No. CA H5284-05-0006 for Everglades National Park

FINAL TASK REPORT

TASK 3 - DEVELOPMENT OF ADDITIONAL MULTIVARIATE LINEAR REGRESSION SALINITY MODELS FOR FLORIDA BAY AND THE SOUTHWEST GULF COAST, EVERGLADES NATIONAL PARK

January 5, 2008
Revised August 5, 2008

I. Introduction

A. General

This report describes the activities that were completed for a task in the -going Critical Ecosystem Studies Initiative (CESI) project that was originally initiated by Cetacean Logic Foundation, Inc. (CLF) for Everglades National Park (ENP) in 2002. The goal of this CLF research is the characterization and simulation of the salinity regime in Florida Bay and the estuaries along southwest coast of the Gulf of Mexico (Gulf) within the Park, which receive freshwater drainage from the Everglades. It is intended that this research yield information regarding the link between the downstream salinity in the estuaries of ENP and the upstream freshwater hydrology of the Everglades.

The subject of the task and this report is the development of multivariate linear regression (MLR) salinity models for some the ENP Marine Monitoring Network (MMN) stations. MLR salinity models were previously developed for about half of the 33 stations in the MMN (Marshall, 2005a; 2005b). The development of models for the remaining MMN stations is described herein.

B. Task Objectives and Evaluation Methods

The objective of this task is presented below as it appears in the modified contract for this CESI project approved by ENP:

> "The year one CESI project focused primarily on the development of models for Joe Bay, Little Madeira Bay, Terrapin Bay, Garfield Bight, North River, and Whipray Basin. The IOP exercise updated some of those models and added a new model for Long Sound, Duck Key, and Butternut Key. The second year CESI project added new MLR salinity models for Taylor River, Highway Creek, Little Blackwater Sound, and Bob Allen Key. Work for the Southern Estuaries Sub-team developed new models for Whitewater Bay East, Clearwater Pass, Shark River, and Gunboat Island. New MLR salinity models will be developed using the longest period of data available at the remaining physical monitoring stations in the ENP MMN. Most of these stations are within the influence of Shark River Slough or relatively more distant from the

Everglades in the open water areas of Florida Bay. The development of new models assumes that the data are adequate and that there is a statistical relationship between salinity at a particular locations and the suite of independent variables that have been assembled for model development. The MMN stations for new models are Broad River, Buoy Key, Broad River Lower, Cane Patch, Harney River, Johnson Key, Lane River, Lostmans River, Little Rabbit Key, Murray Key, Peterson Key, Tarpon Bay East, and Willy Willy.

The deliverable is a task report describing model development activities and new MLR salinity models."

In addition to the stations above, MLR salinity models were also developed for Cannon Bay and Watson Place, in the extreme northwest area of ENP on the coast of the Gulf of Mexico.

II. Data for Model Development

Previous work has shown that daily stage values are useful for salinity modeling because they correlate well with daily salinity, though sometimes there is a lag (Marshall et al, 2003; Marshall, 2003; Marshall et al, 2004). The same is true for daily average sea surface elevation and wind. Coastal aquifer conceptual models and the well-known Ghyben-Herzberg principle provide examples of the connectivity of upstream water levels (both surface and ground water) and downstream salinity in the transition zone of an estuary, and the opposing hydraulic influence of the elevation of the sea surface.

In choosing the data that are to be included in the initial correlation analysis and ultimately the MLR salinity models, the end use of the models has to be considered. To-date, the MLR salinity models have been used primarily for the evaluation of Comprehensive Everglades Restoration Plan (CERP) water delivery alternatives using output from the South Florida Water Management Model (SFWMM, or 2X2 model). The 2X2 model produces daily estimates of stage (water level) and flow of freshwater throughout the Everglades for a number of CERP scenarios over a 36-year period (1965 – 2000). Therefore the MLR salinity models need to be easily utilized with 2X2 Model output as at least one of the applications of these models.

The CERP alternatives re-distribute and augment freshwater deliveries to the Everglades as expressed in the stage data of each of the 2X2 model CERP runs. The MLR salinity models that have already been developed use the 2X2 model stage output in conjunction with available long-term data for wind and sea surface water level to produce estimates of daily salinity for the 36-year period in Florida Bay, the southwest Gulf coast, and Barnes Sound and Manatee Bay.

The independent variable data used with the 2X2 model stage data must be available for most, if not all of the 36-year period in order to populate the models

and obtain estimates of salinity to be of use for the CERP evaluations. Long term sea surface elevation data (i.e. covering the entire 36-year period of record) are only available for at a small number of tide gauging stations along the coast of Florida, and the same is true for wind speed and direction in south Florida. Unfortunately, there are no long-term evaporation data available (particularly at the daily time step) for the 36-year period, because evaporation can play an important role in determining salinity during dry periods when freshwater inflows are at a minimum. Flow data from control structures and tidal creeks do not correlate as well to salinity compared to stage, and observed flow is not as useful for salinity model development and simulation purposes compared to stages (water levels) in the Everglades.

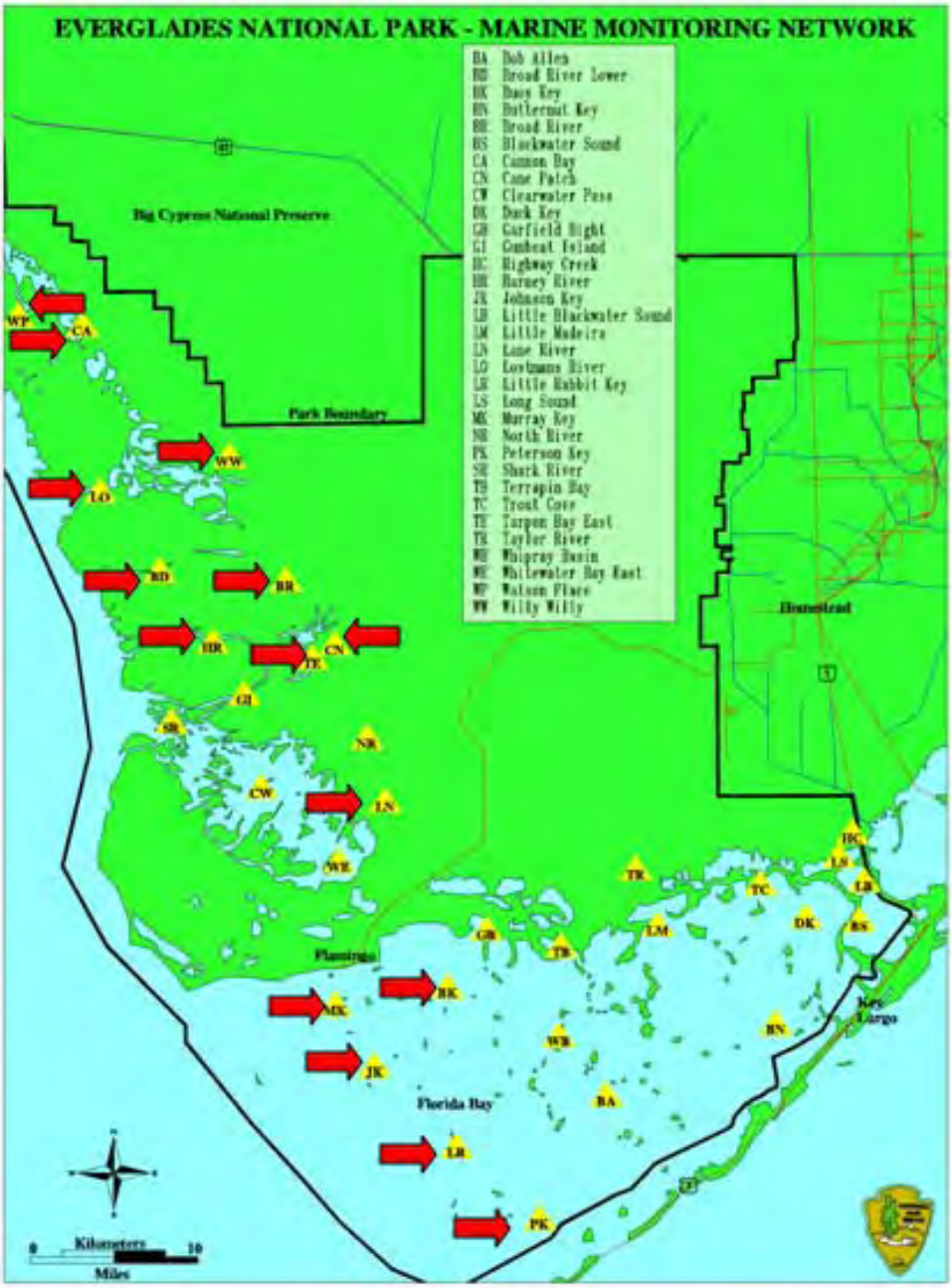
Although rainfall (like evaporation) is an important hydrologic parameter for seasonal salinity variation, rainfall at monitoring stations in the Everglades is not highly correlated with salinity at the daily level. Instead, the stochastic effect of rainfall falling on the Everglades and the upstream watershed is integrated by the coastal aquifer system and expressed adequately in stage data.

For model development, observed stage data are used. Model output data from the 2X2 model and SICS/TIME have previously been used for input to MLR salinity models for simulations. Model development and simulations use the same observed data for wind and sea level although the period of the simulation is longer.

The period of record for model development varied widely because the beginning date for data collection varied widely. Most series contained some missing values. No attempts were made to fill in data gaps or to eliminate outliers in either independent or dependent variable data sets. For all model development activities except the Willy Willy model, data were held from the calibration data and used for a verification exercise.

The models were developed from MMN observed data that have been collected at 15 to 60 minute increments and averaged to daily values. Salinity data were obtained from the ENP Marine Monitoring Network (MMN) data base, Table 1 (http://www.sfnrc.ever.nps.gov/portal/page?_pageid=53,1&_dad=portal&_schema=PORTAL – this website will not be available in the future). Details about these data can be found in Everglades National Park (1997a and 1997b), and Smith (1997, 1998, 1999, and 2001). A map showing the ENP MMN stations and the locations of the water level monitoring stations used for this study is presented as Figure 1. The stage data are ENP Physical Monitoring Network Everglades water levels, as available on the South Florida Water Management District DBHYDRO website (http://my.sfwmd.gov/portal/page?_pageid=2235,4688582&_dad=portal&_schema=PORTAL) . A limited number of continuous water level (stage) monitoring stations in the Everglades began recording data in the 1950's (see Table 2), but most stage records date from the 1990's.

Figure 1. Map of all ENP Marine Monitoring Stations showing stations where MLR salinity models were prepared during this study (red arrows). At all other locations, models were previously prepared as part of other studies



Tide data were obtained from the NOAA Tides Online website (http://tidesonline.nos.noaa.gov/). Because a number of the stations in this study were located north of Cape Sable along the southwest Gulf of Mexico coast, the data from the Naples and Fort Myers tide stations were evaluated for use. However, tide data at both of these stations have only been collected continuously since 1996. Prior to 1996, predicted daily average tide levels are available at these stations. An evaluation of the comparison plots for predicted and actual tide elevations showed considerable local variability likely due to wind or physical condition factors. Therefore, similar to the previous model development activities, the sea surface elevation measured at Key West was used instead as a potential independent variable, and the data span the entire 36-year period.

Wind data are available at a number of locations in the region. National Weather Service wind data used for MLR modeling were obtained from the Southeast Regional Climate Center. Wind data from Key West and Miami were used as these locations had the longest continuous records for wind and were considered to be representative of the regional wind patterns. Hourly wind speed and direction data were processed into vector quantities then daily averages were computed. The independent variables UWNDMIA and VWNDMIA are the *U* and *V* vectors of wind measured at the Miami weather station; UWNDKW and VWNDKW are the *U* and *V* vectors of wind measured at Key West. These components are computed as follows:

U = (Resultant wind speed) * Cosine (Resultant direction)

V = (Resultant wind speed) * Sine (Resultant direction).

For the MLR salinity model development the units of wind speed are ft/sec and the units of direction are bearing degrees.

Table 1. Summary of information about the monitoring stations and salinity data used in model development and verification for additional Florida Bay and southwest Gulf coast MLR salinity models. All data were collected by ENP.

Station Name	MMN ID	Location	Beginning 0f Record
Broad River	BR	Shark River Slough Estuary	01/18/1990
Broad River Lower	BD	Shark River Slough Estuary	04/12/1996
Buoy Key	BK	Central Florida Bay	04/27/1988
Cannon Bay	CA	Upper West Coast	09/20/2000
Cane Patch	CN	Shark River Slough Estuary	01/19/1990
Harney River	HR	Shark River Slough Estuary	03/13/1996
Johnson Key	JK	Western Florida Bay	01/01/1988
Lane River	LN	Whitewater Bay	04/18/1996
Lostmans River	LO	Upper West Coast	10/16/1997
Little Rabbit Key	LR	Western Florida Bay	04/27/1988
Murray Key	MK	Western Florida Bay	04/27/1988
Peterson Key	PK	Western Florida Bay	04/27/1988
Tarpon Bay East	TE	Shark River Slough Estuary	04/04/1996
Watson Place	WP	Upper West Coast	9/20/2000
Willy Willy	WW	Upper West Coast	8/25/97

Table 2. Data summary for the independent variable database used in model development and verification for Florida Bay MLR salinity models.

Variable Name	Variable Type	Units	Data Source	Location	Beginning Date of Data Record
Little Madeira Bay	Salinity	psu	ENP	North Central Florida Bay, near-shore embayment	04/28/1988
Terrapin Bay	Salinity	psu	ENP	North Central Florida Bay, near-shore embayment	09/12/1991
CP	Stage	Ft, NGVD 29	ENP	Craighead Pond	10/01/78
E146	Stage	Ft, NGVD 29	ENP	Taylor Slough	03/24/94
EVER4	Stage	Ft, NGVD 29	ENP	So. Of FL City	09/20/85
EVER6	Stage	Ft, NGVD 29	ENP	So. Of FL City	12/24/91
EVER7	Stage	Ft, NGVD 29	ENP	So. Of FL City	12/24/91
G3273	Stage	Ft, NGVD 29	ENP	East of S.R. Slough	03/14/84
NP206	Stage	Ft, NGVD 29	ENP	East of S.R. Slough	10/01/74
NP46	Stage	Ft, NGVD 29	ENP	Rocky Glades	01/15/66
NP62	Stage	Ft, NGVD 29	ENP	East of S.R. Slough	01/04/64
P33	Stage	Ft, NGVD 29	ENP	Shark River Slough	02/15/53
P35	Stage	Ft, NGVD 29	ENP	Shark River Slough	02/15/63
P37	Stage	Ft, NGVD 29	ENP	Taylor Slough	01/15/53
P38	Stage	Ft, NGVD 29	ENP	Shark River Slough	01/10/52
R127	Stage	Ft, NGVD 29	ENP	Taylor Slough	04/11/84
PA8	Stage	Ft, NGVD 29	SFWMD	Big Cypress Preserve	10/12/95
PA9	Stage	Ft, NGVD 29	SFWMD	Big Cypress Preserve	10/06/95
PA10	Stage	Ft, NGVD 29	SFWMD	Big Cypress Preserve	10/05/95
PA11	Stage	Ft, NGVD 29	SFWMD	Big Cypress Preserve	09/11/95
UWNDKW	E-W Wind	N/A	NWS	Key West	01/07/57
VWNDKW	N-S Wind	N/A	NWS	Key West	01/07/57
UWNDMIA	E-W Wind	N/A	NWS	Miami	01/07/57
VWNDMIA	N-S Wind	N/A	NWS	Miami	01/07/57
KWWATLEV	Sea Surface Elevation	Ft, MSL	NOS	Key West	01/19/13

III. Model Development

A step-wise multivariate linear regression process was used to determine the most appropriate linear combination of independent variables for each salinity model. To begin the model development procedure, all independent variables were subjected to a cross-correlation analysis with daily salinity using SARIMA techniques to determine which of the variables were correlated with salinity, to check for lagged relationships, and to evaluate the level of correlation. Lags up to 50 days were initially reviewed, though it was found that significant lagged correlations never exceeded six days. Then the observed data of the significant correlated variables (current and lagged values) were input to a SAS© PROC REG routine that uses a step-wise regression process to identify the most statistically significant parameters for a multivariate linear regression equation. To ensure that only the most highly significant parameters were selected by this process and to limit the number of variables in a model, the significance level for parameter inclusion in the model was set at 99.9%, a very high level. Parameter inclusion in a model was also manually controlled by eliminating any seemingly correlated variables that acted contrary to known physical relationships (such as an increasing stage in the Everglades indicating an increase in salinity) which can occur when there are cross-correlation effects. These parameters were eliminated, and the step-wise process re-run iteratively.

For the development of each model, some of the available data were held out of the calibration process and used for verification purposes. The periods for calibration and verification are presented in Table 3.

For the open-water stations in central Florida Bay and MMN stations along the open west boundary of the Bay with the Gulf, it was found that the best-fit salinity models included salinity at the near shore stations of Little Madeira Bay and Terrapin Bay in the model as well as stage in the Everglades, wind vectors and sea surface elevation. In fact, without the inclusion of Little Madeira Bay and Terrapin Bay salinity the open-water models were insufficient to explain variability beyond about 30%. This means that, for these stations, simulation will be a two-step process with the simulation of salinity at Little Madeira Bay and Terrapin Bay required before salinity at the open-water stations can be simulated.

For the stations that are located in Whitewater Bay, the Shark River Slough estuary, and along the northernmost Gulf coast in ENP, the use of the daily average sea surface elevation from the Naples, Florida station (ID No. 8725110) instead of the data from the Key West station (ID No. 8724580) improved model fit by about 3-5%. For the Florida Bay stations along the western boundary with the Gulf, both Key West and Naples data were significant. However, the Naples data are predicted values prior to 1995, so Key West water level was used for model development purposes as in the past.

Table 3. Periods used for calibration and verification for MLR salinity model development.

Station Name	Calibration Period	Verification Period
Buoy Key	09/07/1997 - 12/31/2000	01/01/2001 - 10/31/2002
Broad River	01/12/1996 - 12/31/1999	01/01/2000 - 12/31/2000
Broad River Lower	05/09/1996 - 12/31/1999	01/01/2000 - 12/31/2000
Cane Patch	05/09/1996 - 12/31/1999	01/01/2000 - 12/31/2000
Cannon Bay	10/16/1997 - 02/30/2002	01/01/2003 - 03/13/2003
Harney River	06/14/1996 - 12/31/1999	01/01/2000 - 12/31/2000
Johnson Key	08/19/1994 - 12/31/2000	01/01/2001 - 10/31/2002
Lane River	05/09/1996 - 12/31/1999	01/01/2000 - 12/31/2000
Little Rabbit Key	09/10/1997 - 12/31/2000	01/01/2001 - 10/31/2002
Lostmans River	10/16/1997 - 12/31/2000	01/01/2001 - 10/31/2002
Murray Key	10/21/1997 - 12/31/2000	01/01/2001 - 10/31/2002
Peterson Key	07/16/1994 - 12/31/2000	01/01/2001 - 10/31/2002
Tarpon Bay East	05/09/1996 - 12/31/1999	01/01/2000 - 12/31/2000
Watson Place	09/20/2000 - 12/31/2002	01/01/2003 - 03/13/2003
Willy Willy	08/25/1997 - 12/31/2001	01/01/2002 - 03/01/2003

Additional details on model development can be found in Marshall, et al (2003a; 2004) and Marshall (2003b; 2005). The daily salinity models that were developed for this task and the associated adjusted-R^2 value are shown below by sub-region. All independent variables in these models are significant at the 99.9% level, except for Cannon Bay (95% significance level).

Figures 2 to 17 present the salinity simulated by the models below compared to the observed data for the calibration and verification periods. These figures are presented at the end of the report following the references.

Central Florida Bay

Buoy Key = 24.82836 – (1.13942 cplag4) + (0.23472 terbay) + (0.14891 terbaylag4), adj-R^2 = 0.79

Western Florida Bay

Murray Key = 50.44369 – (3.15719 p33) + (0.14449 terbaylag4), adj-R^2 = 0.51

Johnson Key = 53.13962 – (3.53830 p33) + (0.09932 ltmadlag4) + (0.07948terbaylag4) + (0.04843 vwndmia) + (0.55111 kwwatlev), adj-R^2 = 0.55

Little Rabbit Key = 51.31785 – (3.18926 p33lag1) + (0.19900 ltmadlag4) + (0.59864 kwwatlev), adj-R^2 = 0.46

Peterson Key = 39.76767 – (1.57570 p33lag4) + (0.25200 ltmadlag1) - (0.04663 uwndkwlag1) - (0.06325 vwndkw) + (0.08712 vwndmia), adj-R^2 = 0.56

Whitewater Bay

Lane River = 56.42862 – (1.46788 p35lag2) – (2.82488 p33) - (4.71551 np206lag4) – (0.07508 uwndkwlag1) – (0.13990 vwndkwlag1) - 0.05170 vwndmialag3), adj-R^2 = 0.77

Shark River Slough Estuary

Cane Patch = 7.86280 – (1.20972 np206lag4) + (0.23731 vwndmia), adj-R^2 = 0.42

Tarpon Bay East = 22.18537 – (1.02605 p33lag4) – (2.23641 np206lag4), adj-R^2 = 0.57

Harney River = 51.09667 – (2.84042 p35lag4) – (6.12109 np206lag2) – (0.22464 vwndkw) - (0.11811 vwndkwlag1) – (0.17356 uwndmia) + (2.45632 kwwatlevlag1), adj-R^2 = 0.72

Broad River = 20.82753 – (3.11717 np206lag4) - (0.05780 vwndkwlag1), adj-R^2 = 0.53

Broad River Lower = 73.21431 – (4.82197 p35lag3) – (4.03692 p33) - (4.15505 np206lag2) - (0.31452 vwndkw) – (0.22519 uwndmia) + (0.28467 uwndmialag2) + (0.10788 vwndmialag3) + (4.41540 kwwatlevlag2), adj-R^2 = 0.74

Upper West Coast

Willy Willy = 19.10483 – (2.23072 pa8lag2) – (2.23460 pa11lag2), adj-R^2 = 0.71

Lostmans = 43.18695 – (6.08870 pa8lag2) – (3.11218 pa11lag2) – (0.14135 vwndkw) + (3.25077 kwwatlevlag1), adj-R^2 = 0.76

Cannon Bay = 38.12839 – (5.57603 pa8lag2) – (1.88696 pa11lag2), adj-R^2 = 0.68

Watson Place = 41.54103 – (9.59506 pa8lag1) – (0.10461 uwndkw) – (0.23667 vwndmia) adj-R^2 = 0.86 (Provisional model, limited data),

IV. Error Statistics for Models – Goodness-of-Fit Measures

A. Residuals Analysis

The ability of the MLR salinity models to simulate the observed conditions can be evaluated using a number of error statistics. Many error statistics rely on residuals (defined as the observed value minus the predicted value) as the basis for computation of the statistic. Residual plot characteristics can also be used to evaluate how well a model conforms to the assumptions of normally-distributed errors with a mean of 0. Two plots that are typically examined following model development are the predicted value/residual value plots and the normal probability plots. These two plots are presented in Appendix A for each model that was developed.

Most residual plots in Appendix A show no consistent deviations from typical behavior that would indicate a problem with the normal distribution assumptions. However, residuals for models with observed and predicted values that were in the range of 0 - 10 psu show the effects of a small range of observed values. Stations affected by the small range of salinity with low observed values include Broad River, Cane Patch, Tarpon Bay East, and Willy Willy. These stations with a large number of observed values below about 5 psu show systematic behavior

at predicted values less than 0 psu that can be seen in the predicted/residual value plots. This occurs because it is not possible for observed values to less than 0 psu, but it is possible for a predicted value from the model to be less than 0 psu. In the last step of the model development procedure any predicted value less than 0 psu is automatically set to equal 0 psu. However the SAS © procedure that produced the residual plots in Appendix A does not take this into account so the problem of systematic behavior of residuals seen in the plots when the predicted value is less than zero is not relevant to this analysis. The normal probability residual plots for these models are also affected by this issue.

The error statistics that were computed to measure model performance are described below.

B. Mean Error

The Mean Error is another measure of model uncertainty. It is defined as:

$$ME = \frac{1}{N}\sum_{n=1}^{N}\left(O^{(n)} - P^{(n)}\right)$$

where O=observed values, P=predicted values, and N= number of observations used to develop the model. Positive values of the mean error indicate that the model tends to over-predict, and negative values indicated that the model tends to under-predict.

C. Mean Square Error

The Mean Square Error, or MSE, is defined as the mean of the squares of all the errors, as follows:

$$MSE = \frac{1}{N}\sum_{n=1}^{N}(O^{(n)} - P^{(n)})^{2}$$

D. Root Mean Square Error

The Root Mean Square Error is defined as:

$$RMS = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(O^{(n)} - P^{(n)}\right)^{2}}$$

The Root Mean Square Error is a weighted measure of the error where the largest deviations between observed and predicted values contribute most to this uncertainty statistic. This statistic has units that are the same as the observed and predicted values. It is thought to be the most rigorous tests of absolute error.

E. Mean Absolute Error

The Mean Absolute Error is defined as:

$$MAE = \frac{1}{N}\sum_{n=1}^{N}\left|O^{(n)} - P^{(n)}\right|$$

Although the Mean Absolute Error tells nothing about over- or under-prediction, it is considered as another measure of the agreement between observed values and predicted values. It is preferred by some because it tends to cancel the effects of negative and positive errors, and is therefore less forgiving compared to the Mean Error.

F. Maximum Absolute Error

The Maximum Absolute Error is defined as:

$$MAX = \max\left|O^{(n)} - P^{(n)}\right| \quad : \quad n = 1, N$$

The Maximum Absolute Error is the largest deviation between observed and predicted values.

G. Relative Mean Error

Relative measures of error are not as extreme as the absolute measures presented above. Relative error statistics provide a measure of the error relative to the observed value. The Relative Mean Error is defined as:

$$RME = \frac{\sum_{n=1}^{N}\left(O^{(n)} - P^{(n)}\right)}{\sum_{n=1}^{N} O^{(n)}}$$

H. Relative Mean Absolute Error

The Relative Mean Absolute Error is defined as:

$$RMA = \frac{\sum_{n=1}^{N}\left|O^{(n)} - P^{(n)}\right|}{\sum_{n=1}^{N} O^{(n)}}$$

Caution must be applied in the use of these two statistics when there can be small values of the observed and predicted variable, and when they can have both positive and negative signs.

I. Relative Mean Square Error

The Relative Mean Square Error is not as prone to error by small values and/or the presence of both positive and negative values and is defined as:

$$RSE = \frac{\sum_{n=1}^{N}\left(O^{(n)} - P^{(n)}\right)^2}{\sum_{n=1}^{N}\left(\left(O^{(n)} - \overline{O}\right)^2 + \left(P^{(n)} - \overline{O}\right)^2\right)}$$

The Relative Mean Square Error has values between zero and one, with a model that predicts well having a Relative Mean Square Error close to zero.

J. Nash-Sutcliffe Efficiency

The Nash-Sutcliffe Efficiency (NSE) is a measure of model performance that is similar to R^2. It was first proposed for use with models in 1970 (Nash and Sutcliffe, 1970). It is defined as:

$$NSE = 1 - \frac{\sum_{n=1}^{N}(P^{(n)} - O^{(n)})}{\sum_{n=1}^{N}(O^{(n)} - \overline{O})}$$

The value of the NSE roughly corresponds to the percentage of variation that is explained by a model.

K. Coefficient of Determination - R^2

The Coefficient of Multiple Determination (R^2) is the most common measure of the explanatory capability of a multivariate regression model. It is defined as:

R^2 = Sum of Squares Regression/Sum of Squares Total, or
= 1- (Sum of Squares Error/Sum of Squares Total)

$$R^2 = 1 - \frac{\sum_{n=1}^{N}(O^{(n)} - P)^2}{\sum_{n=1}^{N}(O^{(n)} - \overline{O})^2} \qquad R^2 = \frac{\sum_{n=1}^{N}(P^{(n)} - \overline{O})^2}{\sum_{n=1}^{N}(O^{(n)} - \overline{O})^2}$$

R^2 measures the percentage reduction in the total variation of the dependent variable associated with the use of the set of independent variables that comprise the model. When there are many variables in the model, it is common to use the Adjusted Coefficient of Multiple Determination (adj-R^2), which is R^2 divided by the associated degrees of freedom.

L. Summary of Error Statistics

Table 4 presents a summary of the values of the error statistics for the models that were developed as part of this project. The R^2 values presented with the models above indicate that that the adjusted R^2 values for this suite of models range from 0.40 - 0.86. The error statistics in Table 4 indicate that Nash-Sutcliffe efficiencies of 0.41 – 0.90. The root MSE and the mean absolute error are between 1.23 – 6.12 psu, with most between 2 - 4 psu, meaning that estimation of daily point values produced by the models may have a potential error margin of 2 – 4 psu, on the average. As with all of the daily MLR salinity models produced to-date, the maximum absolute error is large, between about 5 – 20 psu, which means that there is the potential for a point estimate to have an error this large, though that is not considered to be typical.

Comparison of the error statistics in Table 4 with the statistics for other salinity models as presented in Marshall et al (2006) shows that, in general, these models perform as well and better than other salinity model types that have been developed such as hydrodynamic and hydraulic models. However, improvement may have been made for some hydrodynamic models such as USGS TIME and EFDC since the Marshall et al (2006) report was completed.

Table 4. Comparison of Model Error Statistics for MLR Salinity Models.

Station Name	N	Mean error	Mse	Root mse	Mean abs error	Max abs error	Rel mean error	Rel mse	Rel mean abs error	NSE
Broad River	1506	-0.04	5.11	2.26	1.26	13.91	-0.02	0.30	0.57	0.54
Broad Lower	1192	-0.05	15.48	3.94	3.04	16.65	0.00	0.14	0.26	0.75
Cane Patch	1602	-0.02	1.26	1.12	0.53	8.98	-0.02	0.41	0.67	0.42
Cannon Bay	794	0.01	37.50	6.12	4.75	22.87	0.01	0.05	0.26	0.90
Harney River	1600	-0.02	14.60	3.82	2.92	14.09	0.00	0.16	0.35	0.72
Johnson Key	1624	0.05	7.20	2.68	2.21	9.72	0.01	0.29	0.07	0.55
Lane River	1604	-0.14	8.58	2.93	2.13	12.55	-0.02	0.12	0.32	0.79
Little Rabbit	1121	0.09	5.84	2.41	1.92	8.58	0.01	0.37	0.05	0.45
Lostmans River	746	-0.03	21.33	4.62	3.72	12.80	-0.01	0.13	0.30	0.76
Murray Key	867	0.02	8.40	2.90	2.34	11.97	0.01	0.32	0.07	0.51
Peterson Key	1698	-0.01	3.90	1.98	1.58	5.79	-0.01	0.27	0.05	0.57
Tarpon Bay East	1601	-0.05	2.73	1.65	1.00	9.04	-0.02	0.27	0.48	0.58
Watson Place	716	-0.01	23.58	4.85	3.67	20.65	-0.01	0.07	0.18	0.86
Willy Willy	620	-0.31	6.62	2.57	1.57	10.86	-0.12	0.17	0.61	0.72

V. Discussion

The models presented herein for the ENP MMN stations, when added to the models previously prepared (Marshall, et al 2003a; 2004 and Marshall, 2003b; 2005), complete salinity model development for the MMN and provide a method of estimating salinity throughout all of the estuarine areas of Everglades National Park. While there are other salinity models available for salinity estimates in Florida Bay (FATHOM, EFDC, Nuttle's Four-box Model, SICS/TIME) the MLR salinity models developed by this study and the previous studies are the only salinity models available for Whitewater Bay, Shark River estuary, and upper west Gulf coast areas that are capable of producing a 36-year time series simulation. All of the MLR salinity models include the link to the upstream, freshwater hydrology of the Everglades through the stage independent variable, and include the important factors of sea surface level and wind variation. While evapotranspiration is not explicitly included in the model, another CESI study (Marshall, 2006) showed that the effects of evapotranspiration are included in the stage data and are so highly correlated with stage that little error is introduced at the daily time step for these models when evapotranspiration is not included. This is fortunate since long-term measurements of evapotranspiration in south Florida do not exist.

The initial MLR salinity models that were previously developed were in water body areas of specific interest. When the first models were developed they were shown to be capable of reasonably estimating daily salinity such that alternatives for CERP projects could be evaluated through the use of the SFWMD 2X2 model. However, it was noticed in the early modeling that the models are sensitive to both local and regional conditions.

The models presented herein provide additional evidence that the parameter selection process is capable of selecting independent variables that represent the relative effects of both local and regional conditions. For example, salinity at the MMN stations in Whitewater Bay and the Shark River estuary are related to the stage at monitoring stations in the Shark River Slough instead of stage stations in Taylor Slough. In particularly none of the west coast stations show any significant correlation to the stage at Craighead Pond, which was shown to be highly important for explaining the variation in salinity at Florida Bay MMN stations, and particularly the near-shore embayments. Similarly, the stations along the uppermost western Gulf coast within the Park are related to the water levels in Big Cypress Preserve as opposed to the stage in Shark River Slough.

All of the MLR salinity models for the MMN stations along the western open-water boundary of Florida Bay included P33 as a primary independent variable, either un-lagged or with a 4-day lag, as well as salinity in the near-shore embayments of Little Madeira Bay and Terrapin Bay. The importance of P33 provides evidence of the link between the Shark River discharge and the western

boundary stations as has been theorized by researchers who have examined circulation patterns in this area. All of the Shark River estuary salinity models included NP206 as a primary station, lagged and un-lagged. It is important to note that the selection criteria for inclusion of a parameter in the model was very high, 0.999, meaning that the consistent selection of independent variables as the most significant in explaining salinity variability provides additional evidence that certain stations are primary.

Wind vectors were important in explaining salinity variability at almost all stations. Each of these patterns shows that the MLR salinity models are reflecting the important driving forces for salinity in a way that seems to make physical hydrologic and hydraulic sense.

VI. Conclusions and Recommendations

The development of these additional MLR salinity models for the remaining stations in the MMN means that salinity models are now available for all of the estuarine areas within ENP. CERP alternatives can now be examined for their effect in all of the southern estuaries within ENP, providing a complete picture of the effects of meteorological events and water management as well as the potential for restoration. In a general manner the development of all MMN models have shown that there is a definitive link between the upstream hydrologic conditions in the freshwater marshes of the Greater Everglades and the estuaries into which the freshwater is being discharged, be it north or east of Cape Sable. This also means that the restoration of the estuarine areas within Everglades National Park can only happen with the restoration of freshwater levels and volumes in the Greater Everglades.

It is recommended that the models developed herein be used along with the previously developed models for estimating the salinity at the various MMN stations. The models can be used by any application that requires a time series of salinity as input data such as ecological models (e.g. seagrass and shrimp models), paleoecological evaluations, freshwater allocation activities such as minimum flows and levels determinations, and evaluating CERP alternatives.

VII. References

Marshall III, F.E.; D. Smith; and D. Nickerson. 2003. Salinity simulation models for north Florida Bay Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. 41 pp.

Marshall III, F.E. 2003. IOP salinity analysis using statistical models. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. 35 pp.

Marshall III, F. E.; D. Smith; and D. Nickerson. 2004. Using statistical models to simulate salinity variation and other physical parameters in north Florida Bay. Cetacean Logic Foundation, Inc., New Smyrna Beach, Florida.

Marshall III, F. E. 2005a. RECOVER Southern Estuaries performance measures: identification of hydrology-salinity relationships for coastal estuaries and analysis of interim CERP update scenarios. Environmental Consulting & Technology, Inc., New Smyrna Beach, Florida.

Marshall III, F.E.; and D. Smith. 2005. Investigation of evaporation surrogates for Florida Bay salinity modeling. Task Report for a Critical Ecosystems Initiative (CESI) Project. Cetacean Logic Foundation, Inc., New Smyrna Beach, Florida.

Marshall III, F.E.; D. Smith; and W. Nuttle. 2006. Simulating and forecasting salinity in Florida Bay: a review of models. Task Report for a Critical Ecosystems Initiative (CESI) Project. Cetacean Logic Foundation, Inc., New Smyrna Beach, Florida.

Figure 2. Comparison of predicted salinity compared to observed salinity for Buoy Key. See Table 3 for calibration and verification periods.

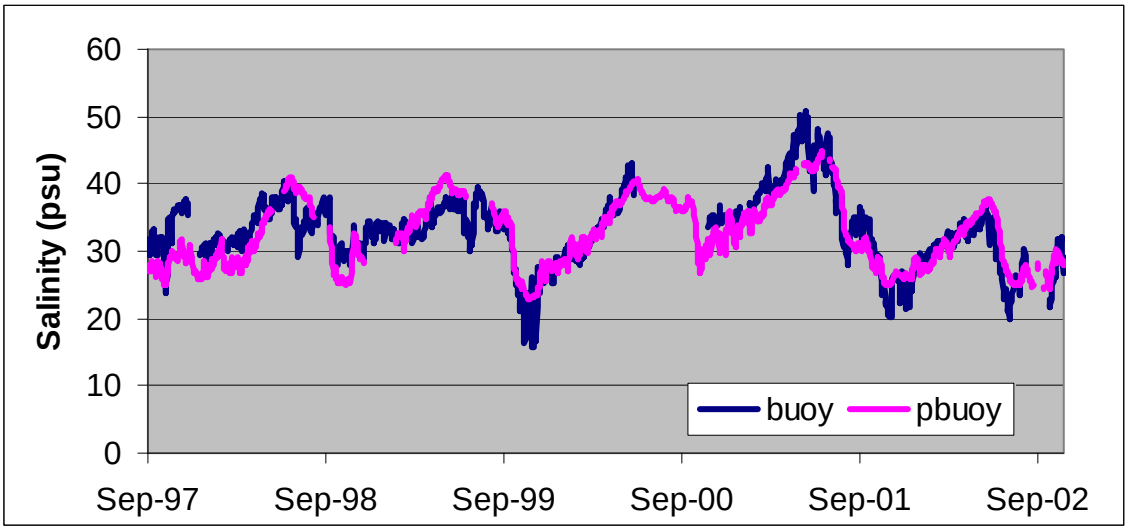


Figure 3. Comparison of predicted salinity compared to observed salinity for Murray Key. See Table 3 for calibration and verification periods.

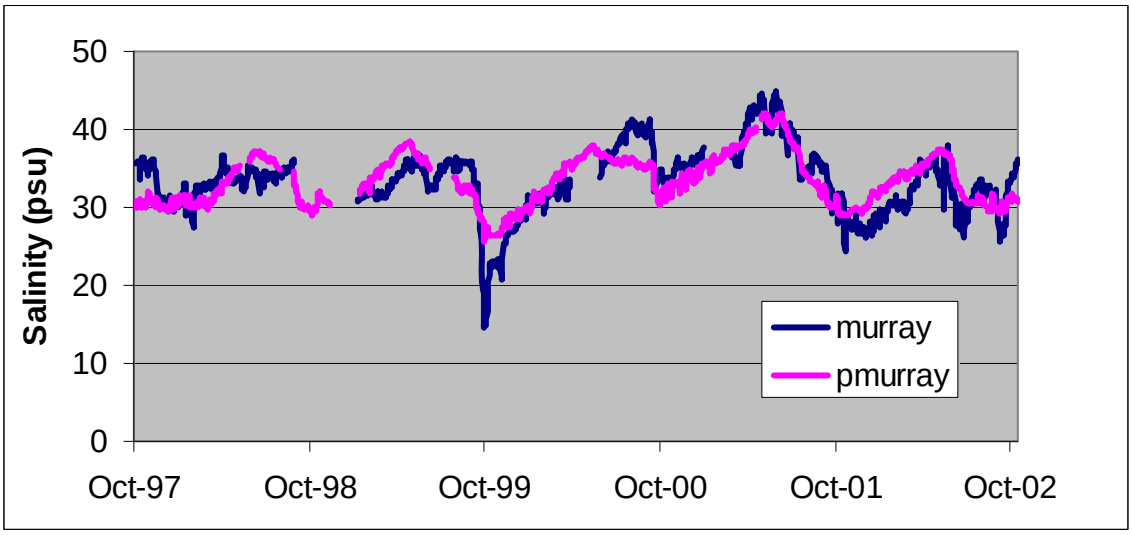


Figure 4. Comparison of predicted salinity compared to observed salinity for Johnson Key. See Table 3 for calibration and verification periods.

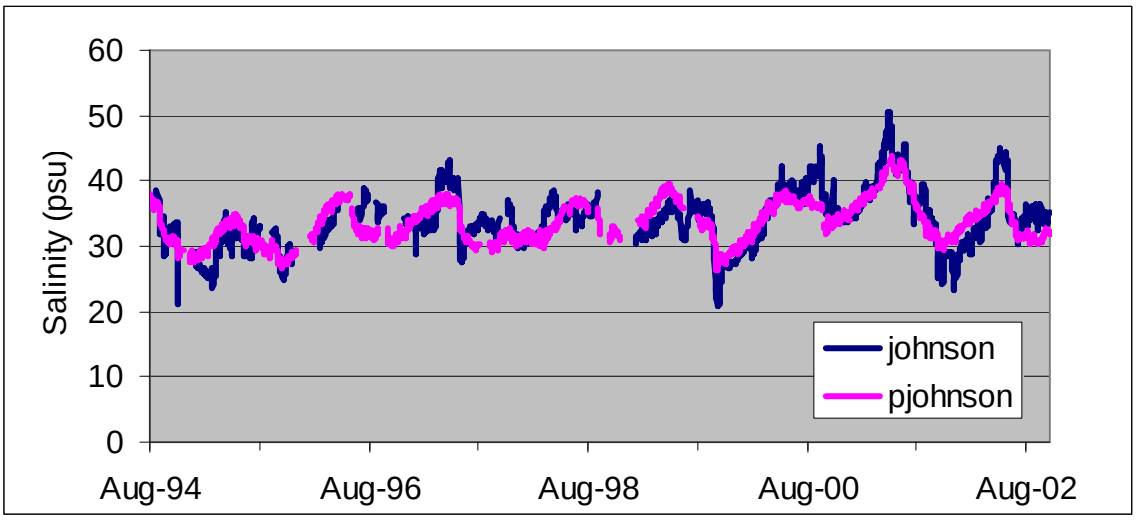


Figure 5. Comparison of predicted salinity compared to observed salinity for Little Rabbit Key. See Table 3 for calibration and verification periods.

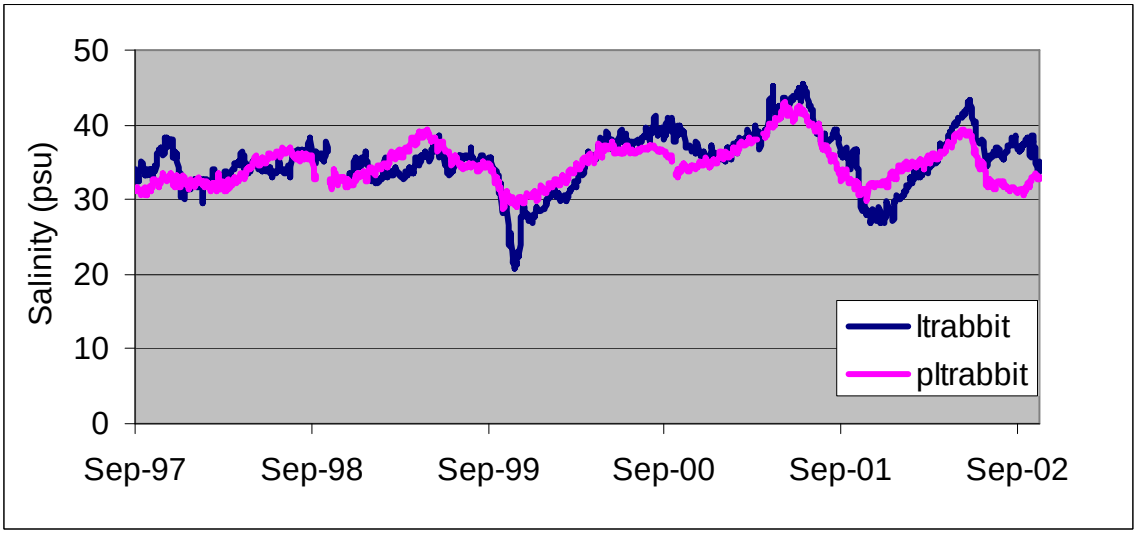


Figure 6. Comparison of predicted salinity compared to observed salinity for Peterson Key. See Table 3 for calibration and verification periods.

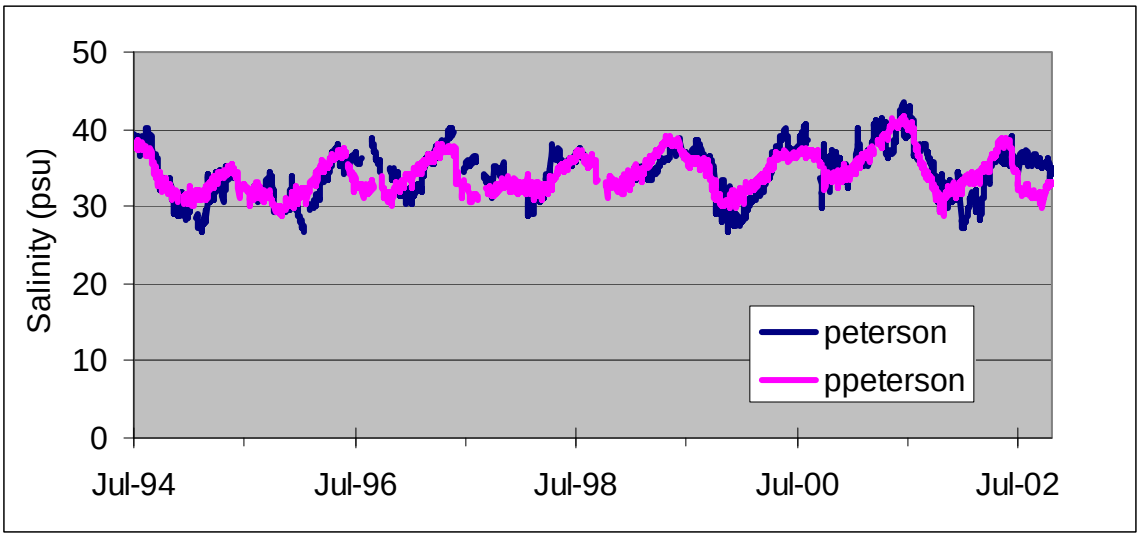


Figure 7. Comparison of predicted salinity compared to observed salinity for Lane River. See Table 3 for calibration and verification periods.

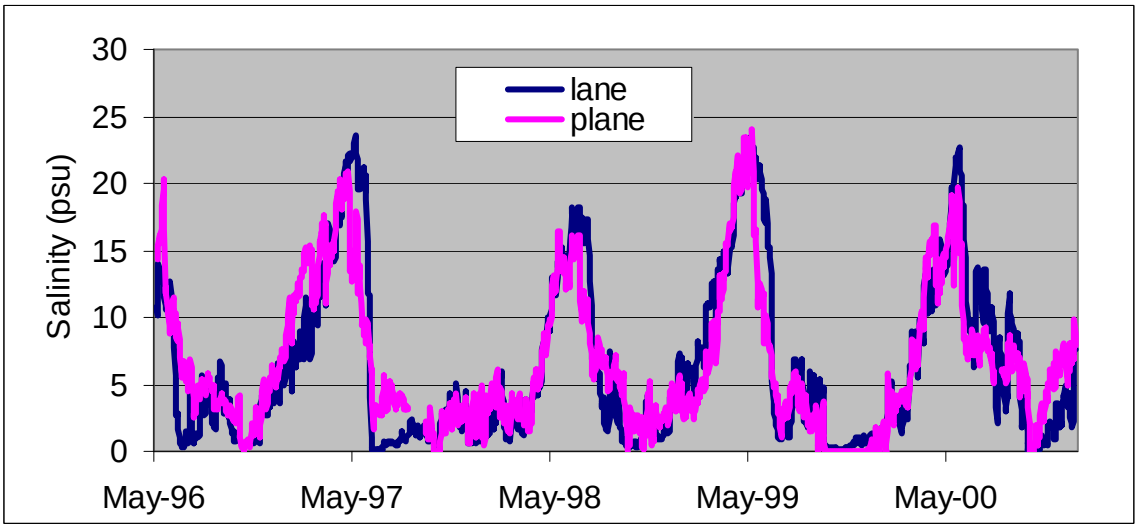


Figure 8. Comparison of predicted salinity compared to observed salinity for Cane Patch. See Table 3 for calibration and verification periods.

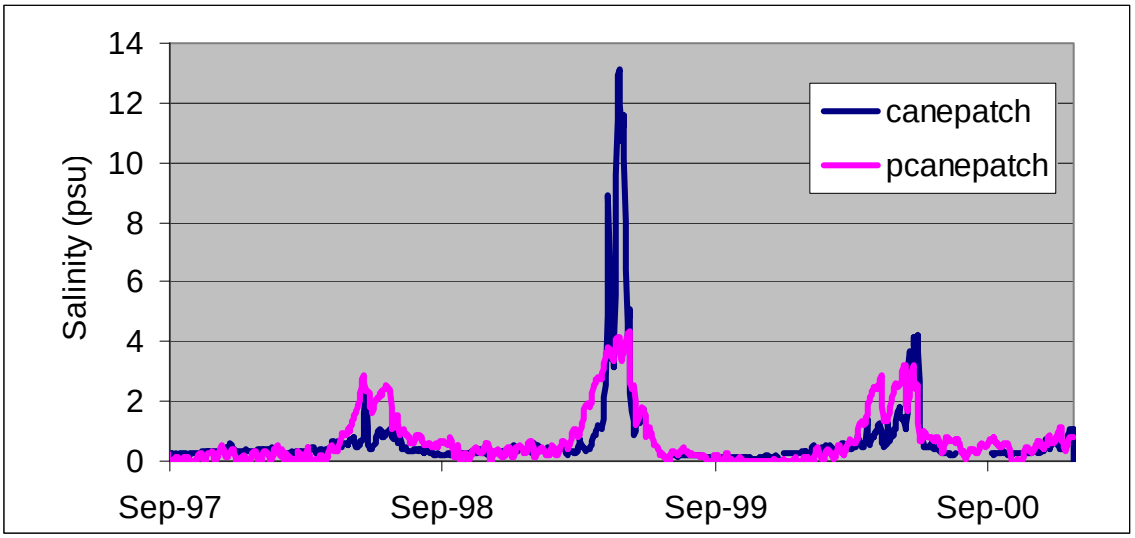


Figure 9. Comparison of predicted salinity compared to observed salinity for Tarpon Bay East. See Table 3 for calibration and verification periods.

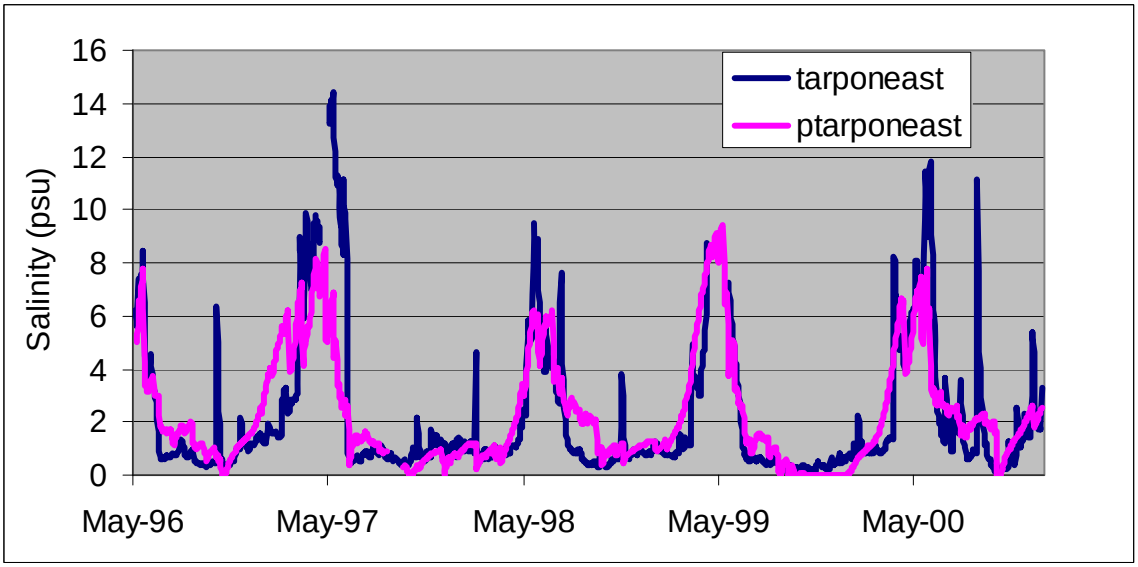


Figure 10. Comparison of predicted salinity compared to observed salinity for Harney River. See Table 3 for calibration and verification periods.

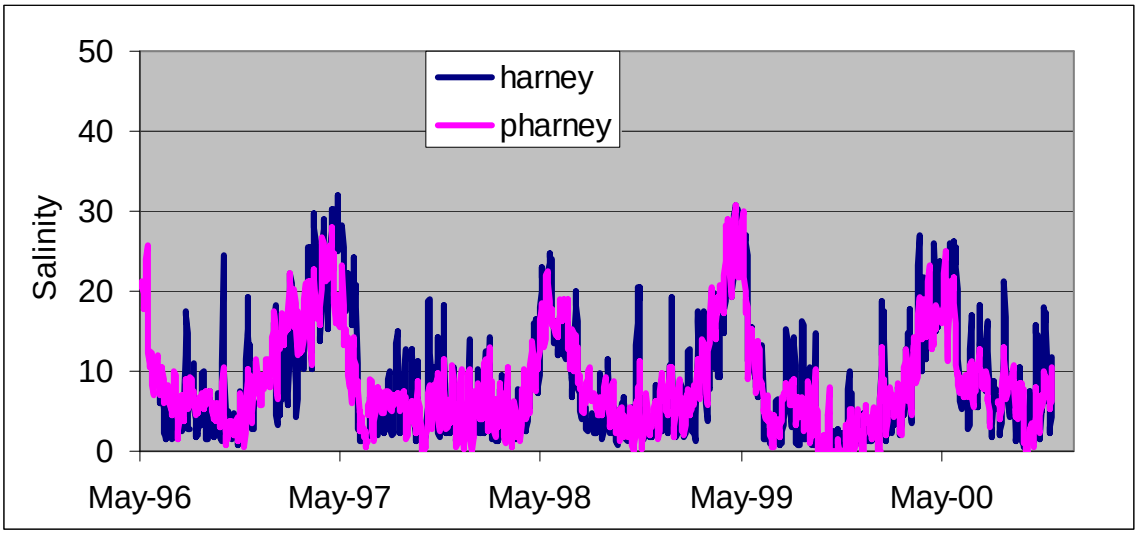


Figure 11. Comparison of predicted salinity compared to observed salinity for Broad River. See Table 3 for calibration and verification periods.

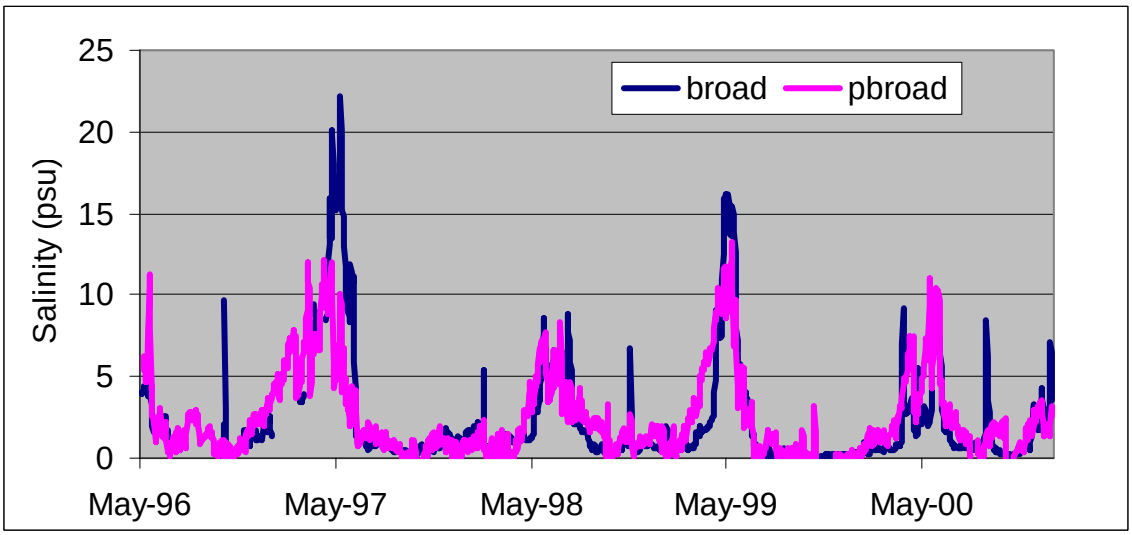


Figure 12. Comparison of predicted salinity compared to observed salinity for Broad River Lower. See Table 3 for calibration and verification periods.

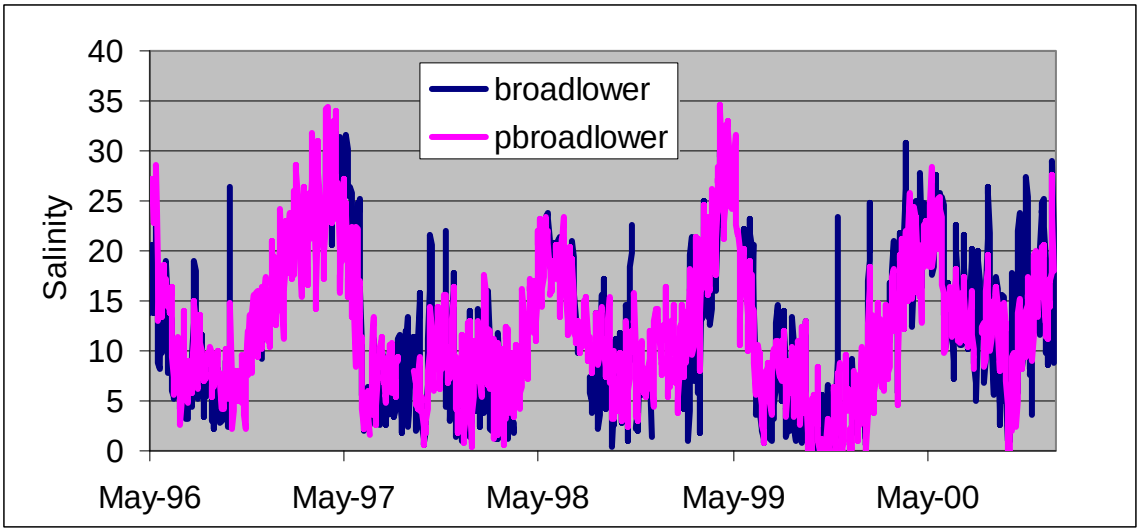


Figure 13. Comparison of predicted salinity compared to observed salinity for Willy Willy. See Table 3 for calibration and verification periods.

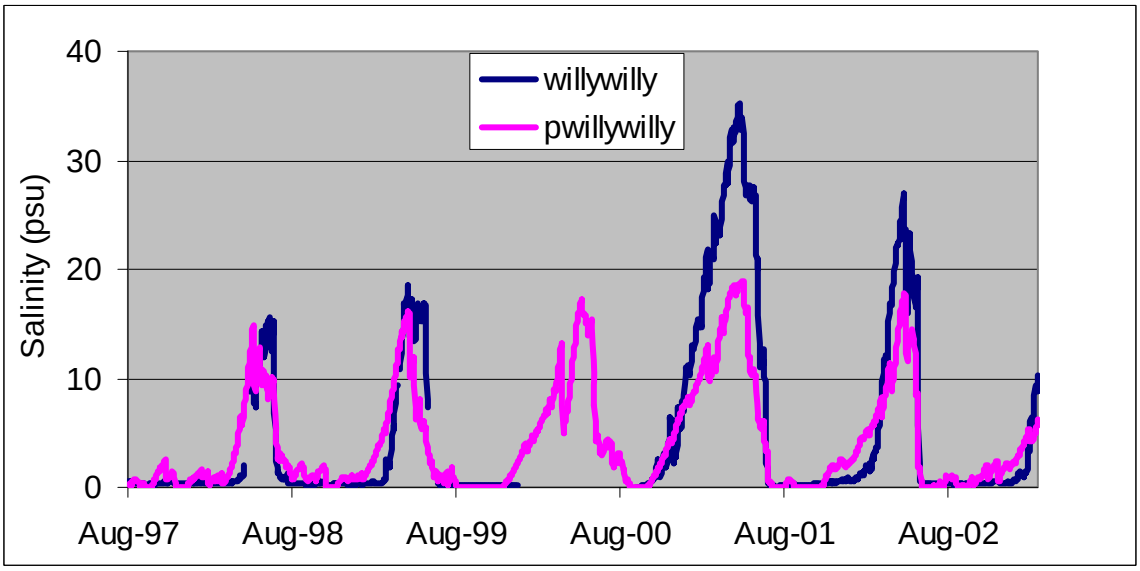


Figure 14. Comparison of predicted salinity compared to observed salinity for Lostmans River. See Table 3 for calibration and verification periods.

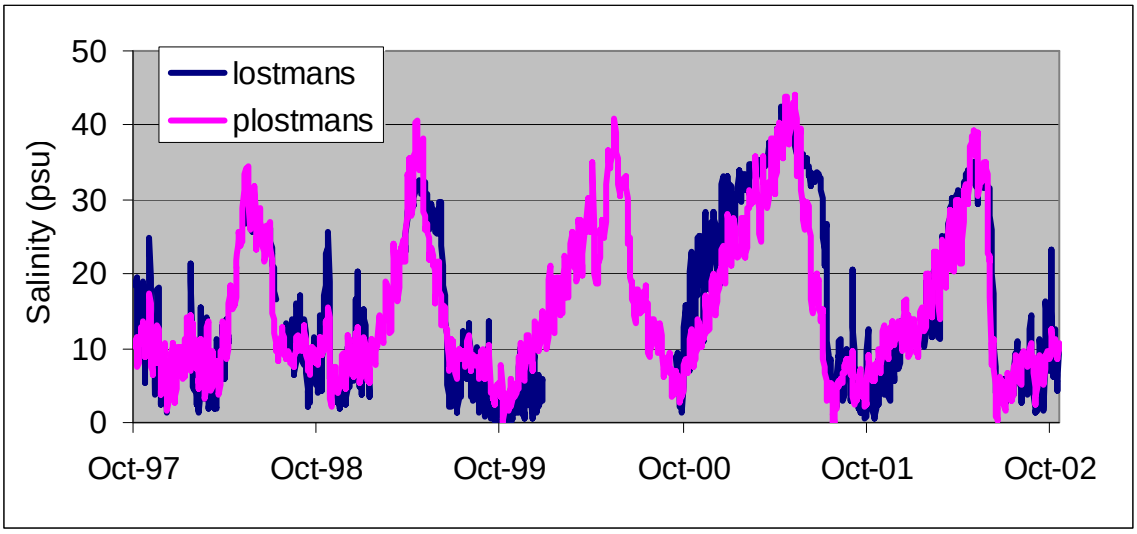


Figure 15. Comparison of predicted salinity compared to observed salinity for Cannon Bay. See Table 3 for calibration and verification periods.

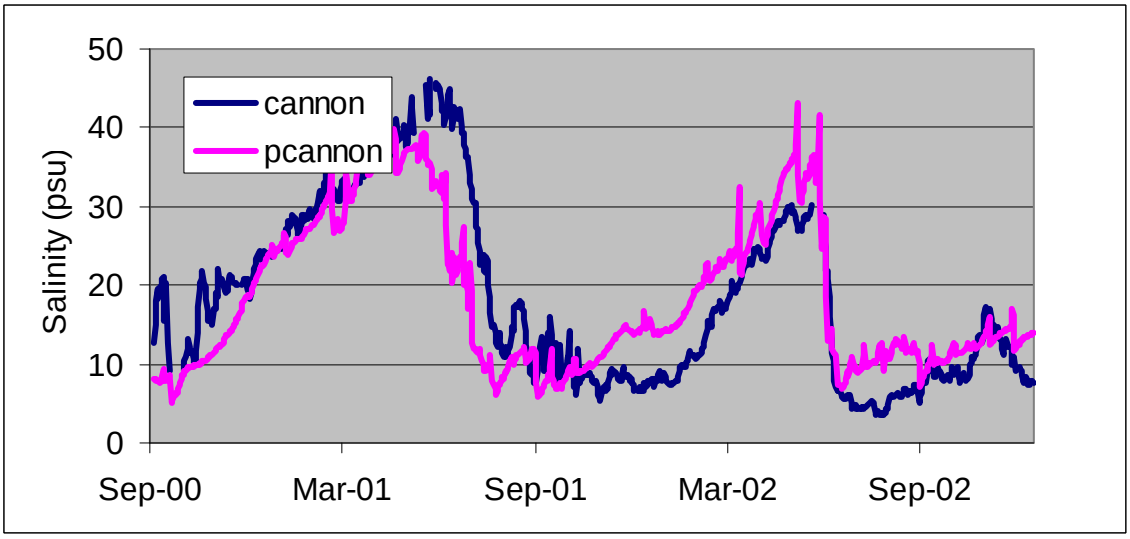
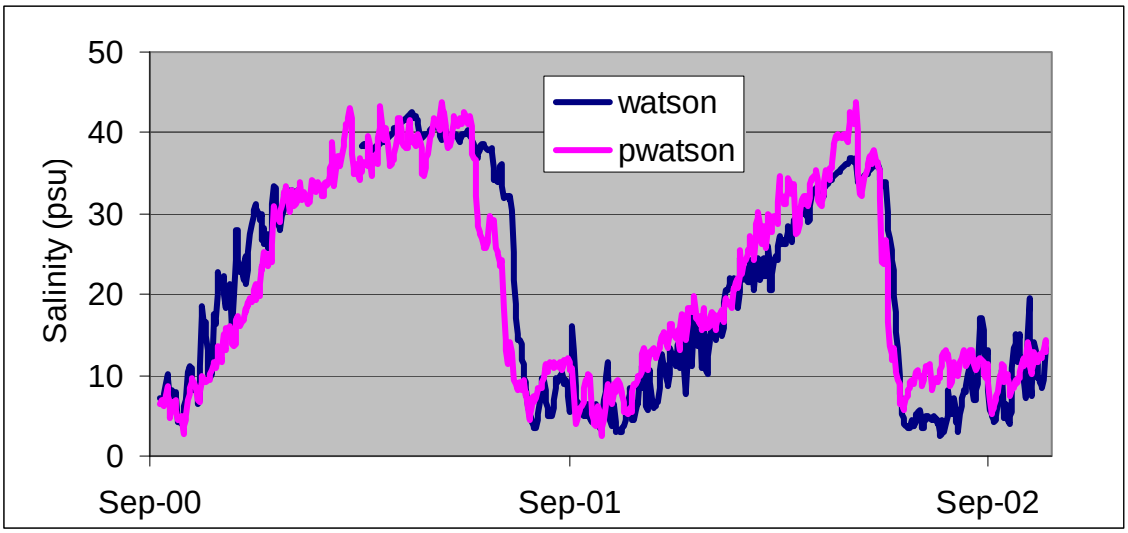
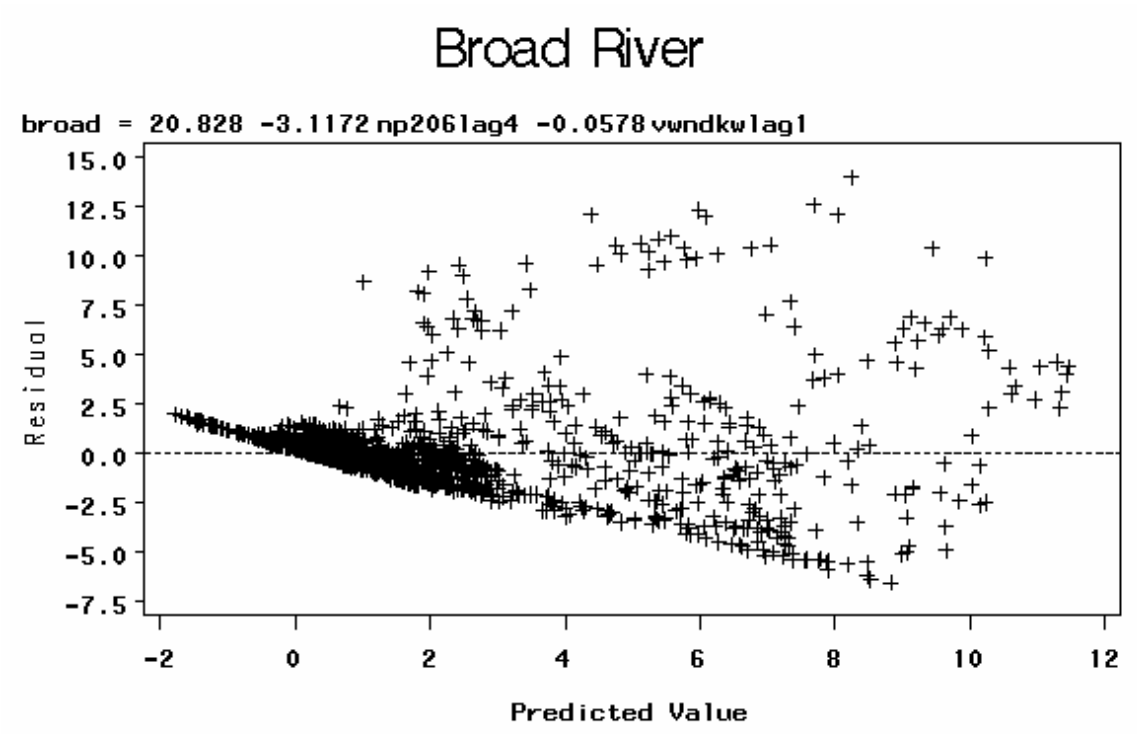


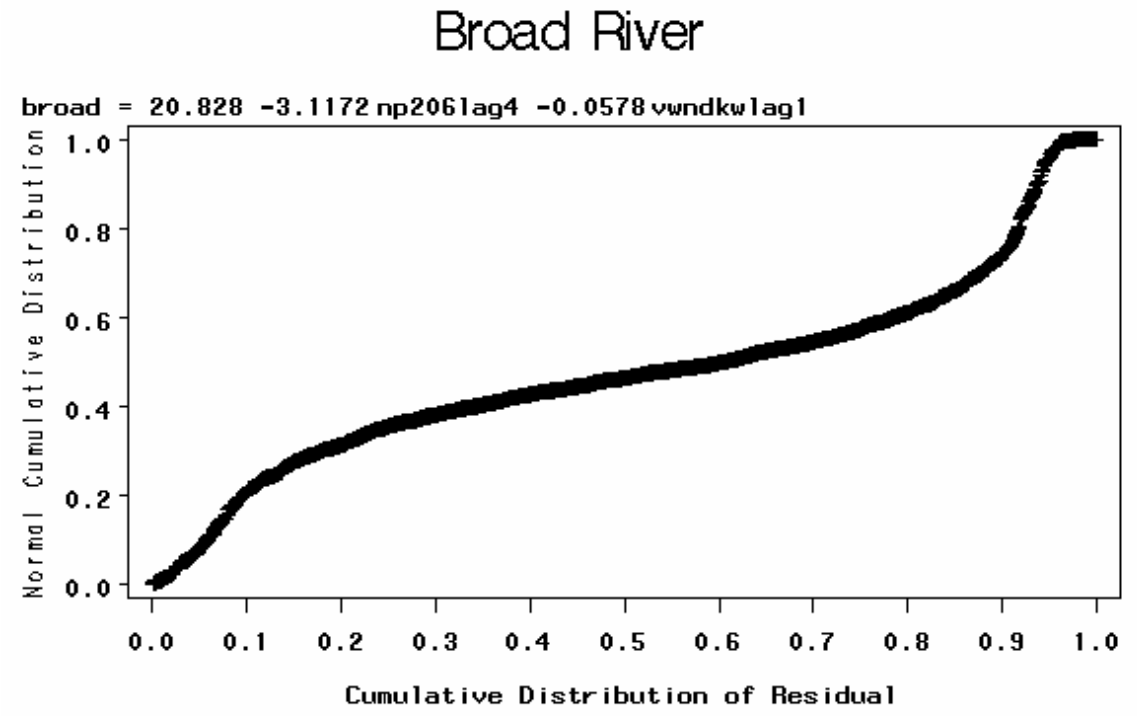
Figure 16. Comparison of predicted salinity compared to observed salinity for Watson Place. See Table 3 for calibration and verification periods.



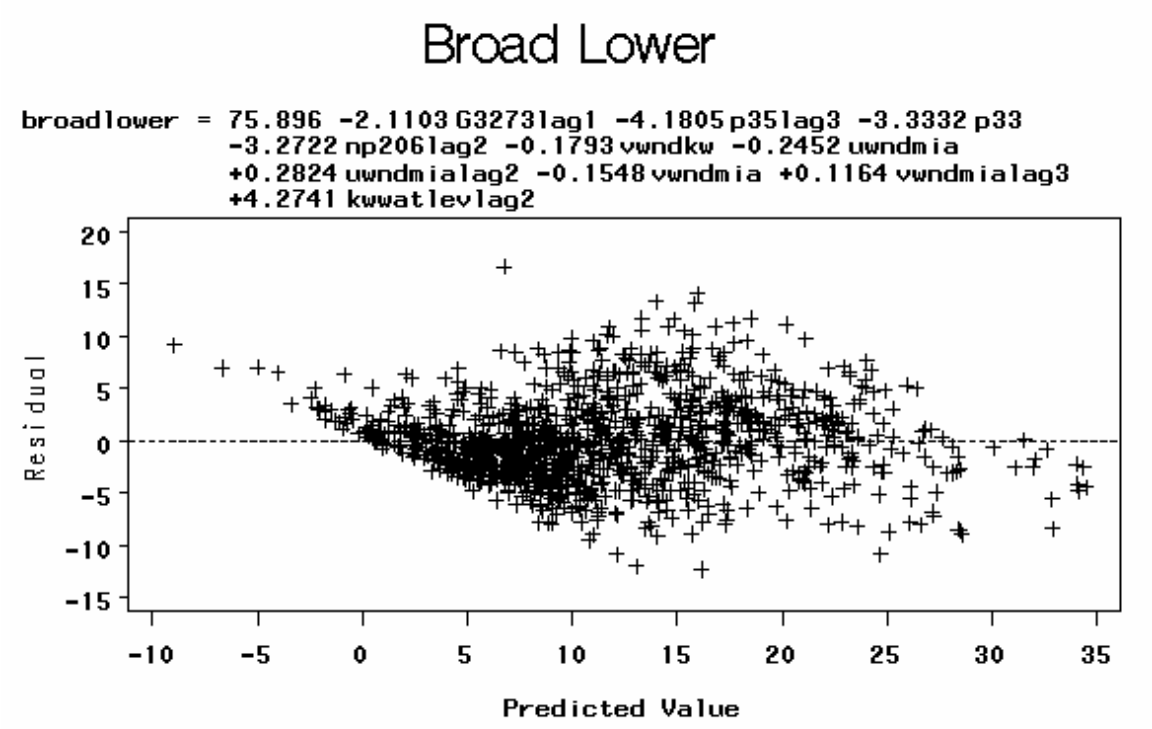
Appendix A. Residual Plots

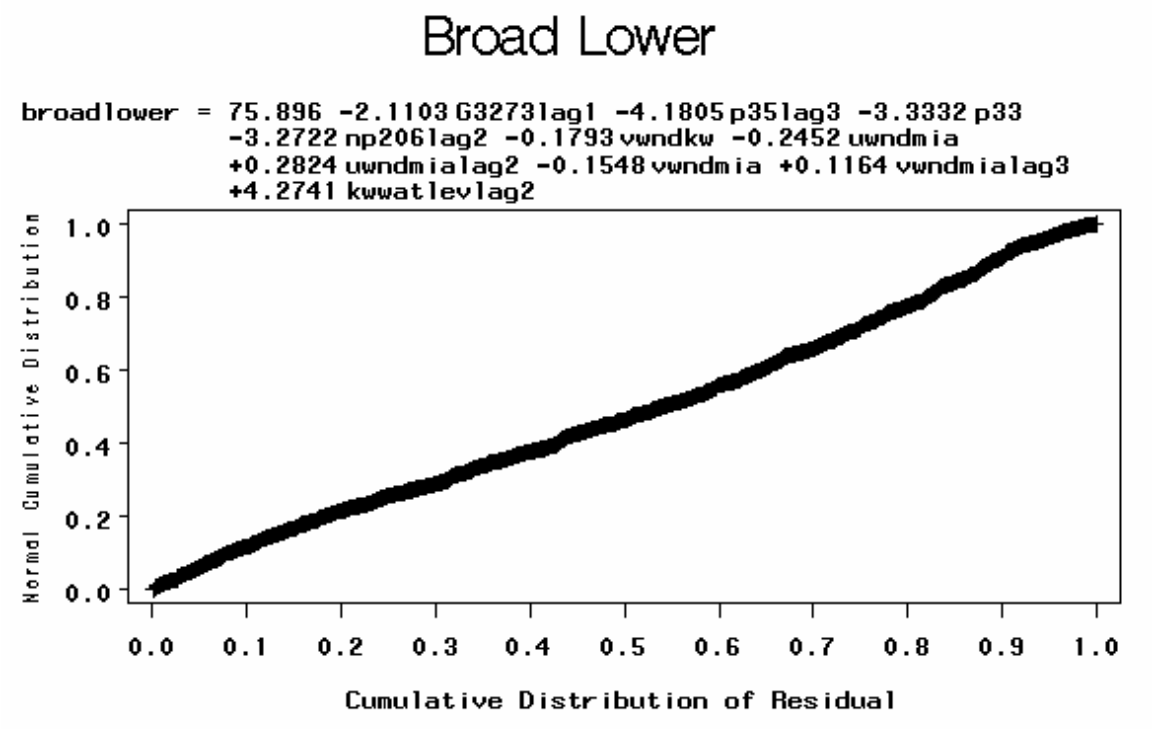
Broad River



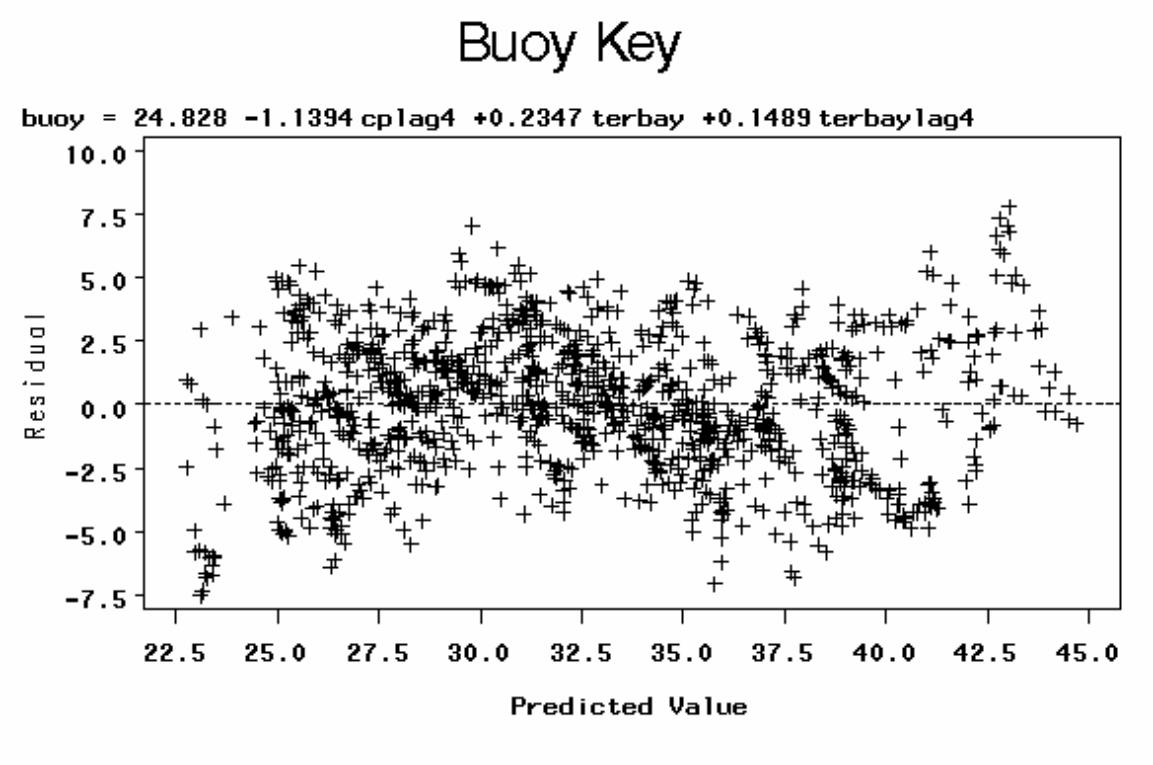


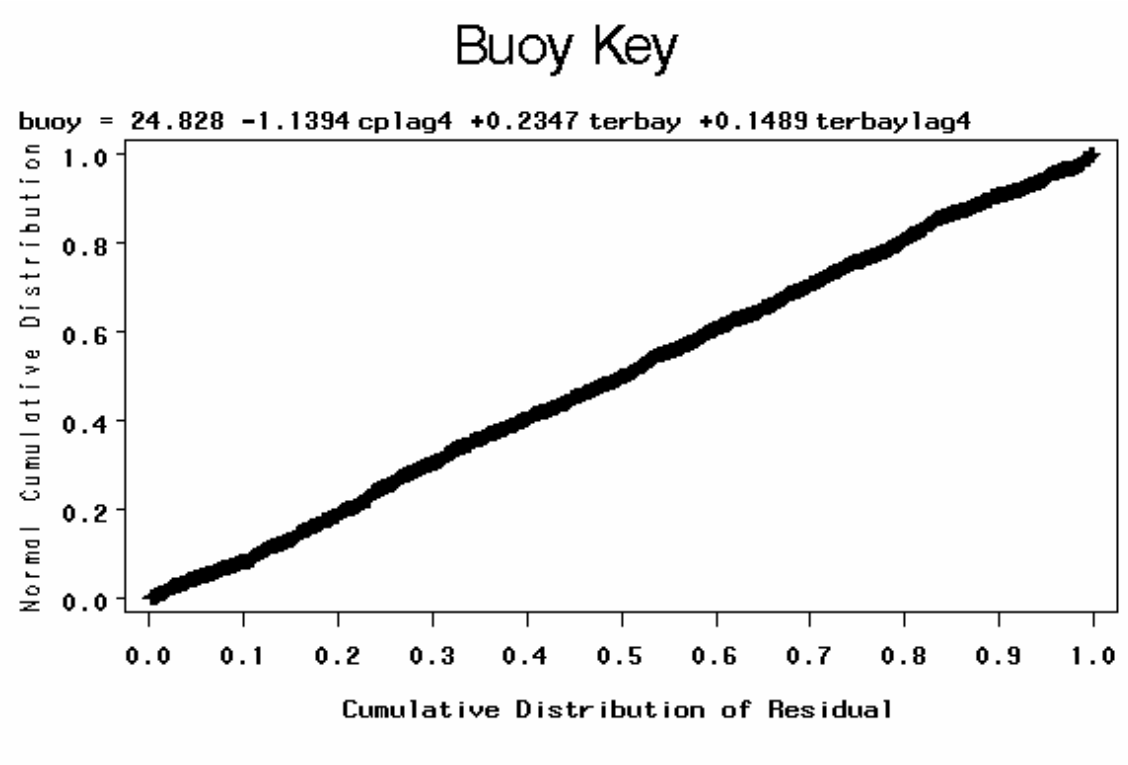
Broad River Lower



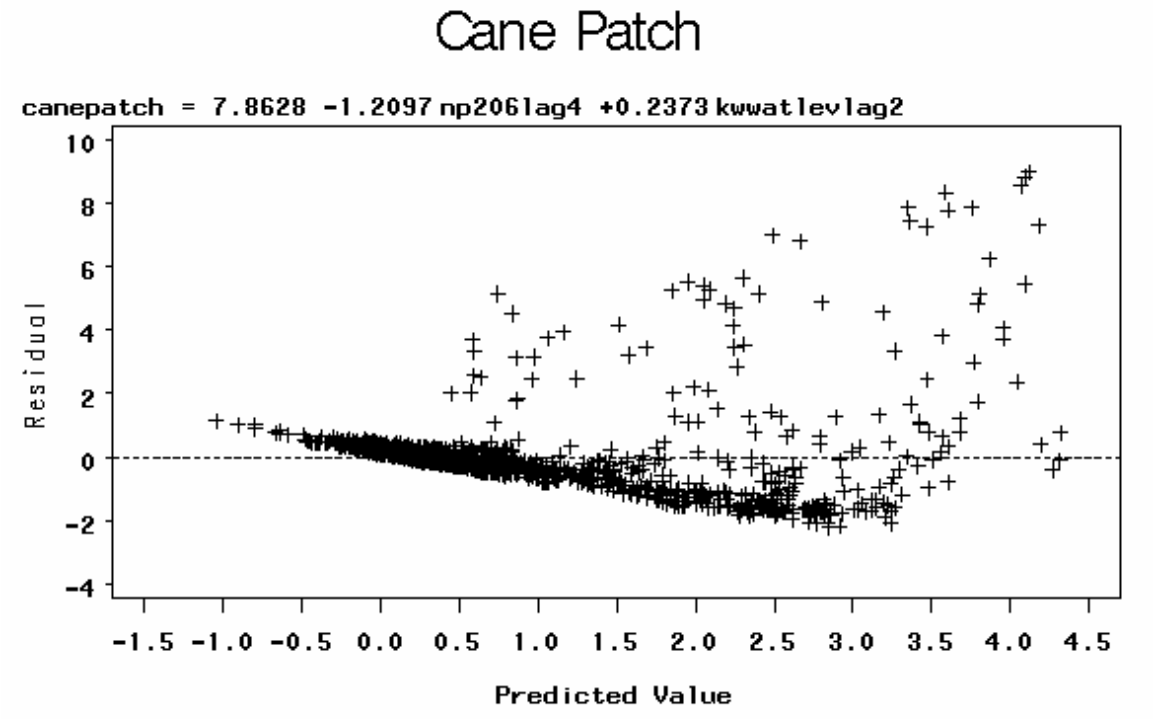


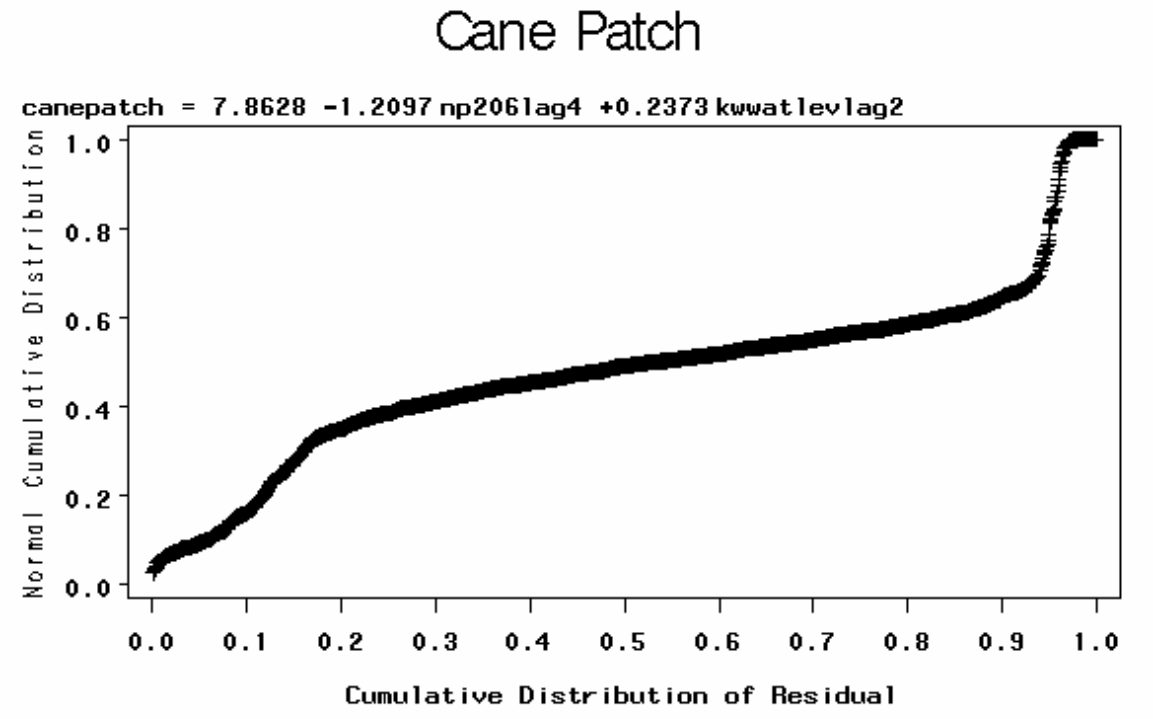
Buoy Key


Buoy Key
buoy = 24.828 -1.1394 cplag4 +0.2347 terbay +0.1489 terbaylag4
Residual
10.0
7.5
5.0
2.5
0.0
-2.5
-5.0
-7.5
22.5
25.0
27.5
30.0
32.5
35.0
37.5
40.0
42.5
45.0
Predicted Value

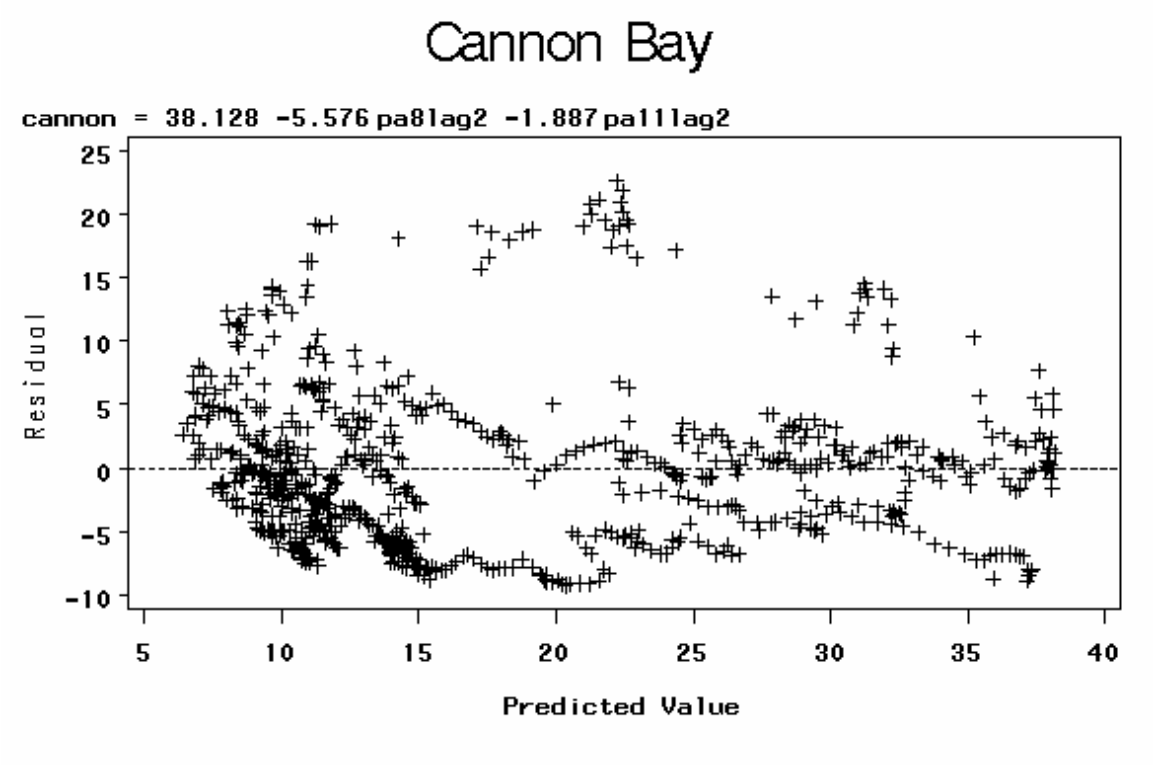

Buoy Key
buoy = 24.828 -1.1394 cplag4 +0.2347 terbay +0.1489 terbaylag4
Normal Cumulative Distribution
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Cumulative Distribution of Residual

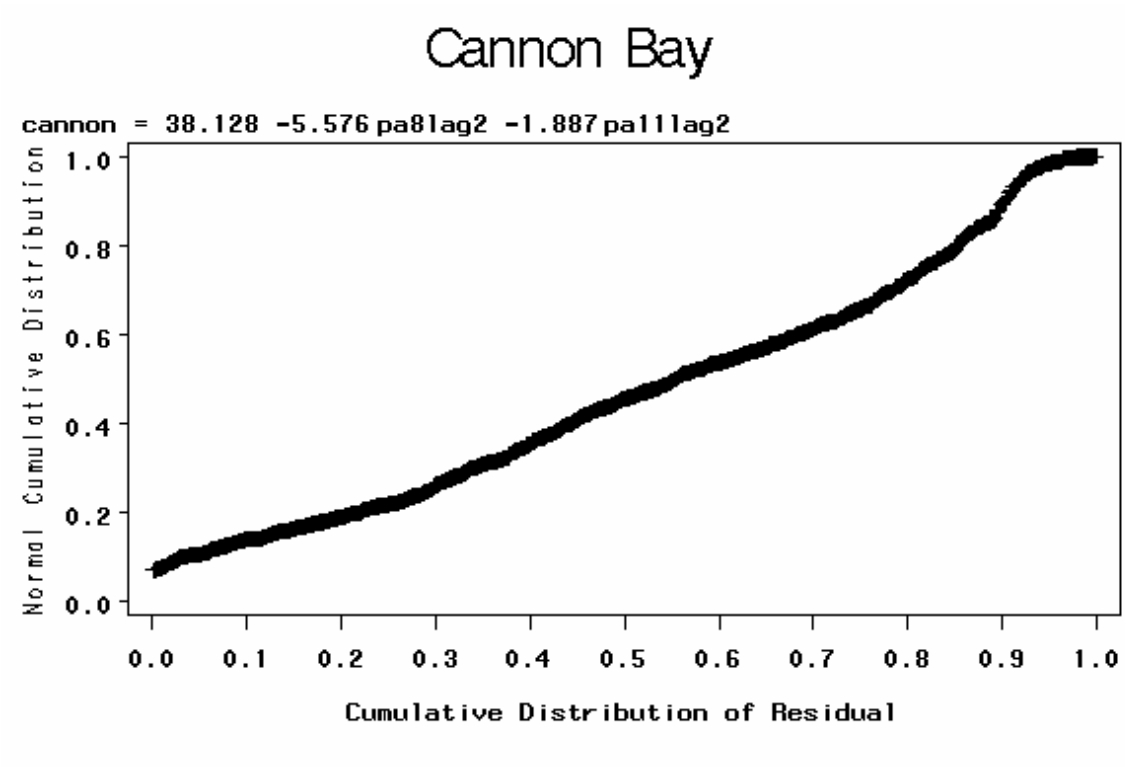
Cane Patch


Cane Patch
canepatch = 7.8628 -1.2097 np206lag4 +0.2373 kwwatlevlag2
Residual
10
8
6
4
2
0
-2
-4
-1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5
Predicted Value

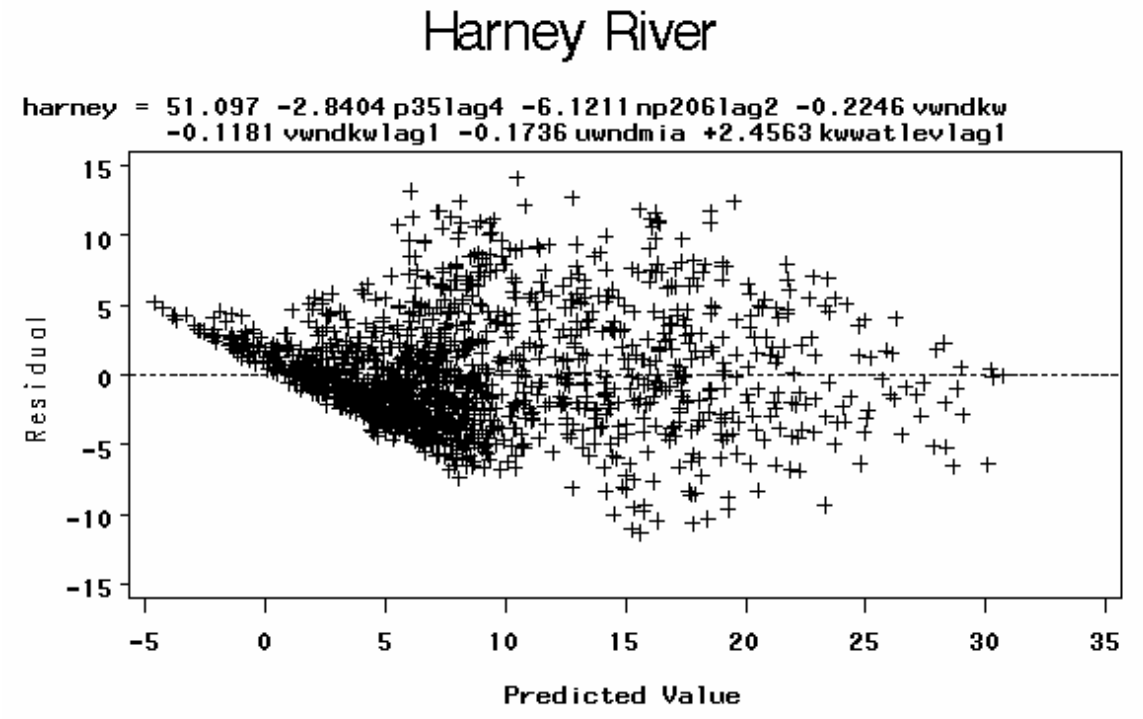

Cane Patch
canepatch = 7.8628 -1.2097 np206lag4 +0.2373 kwwatlevlag2
Normal Cumulative Distribution
1.0
0.8
0.6
0.4
0.2
0.0
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Cumulative Distribution of Residual

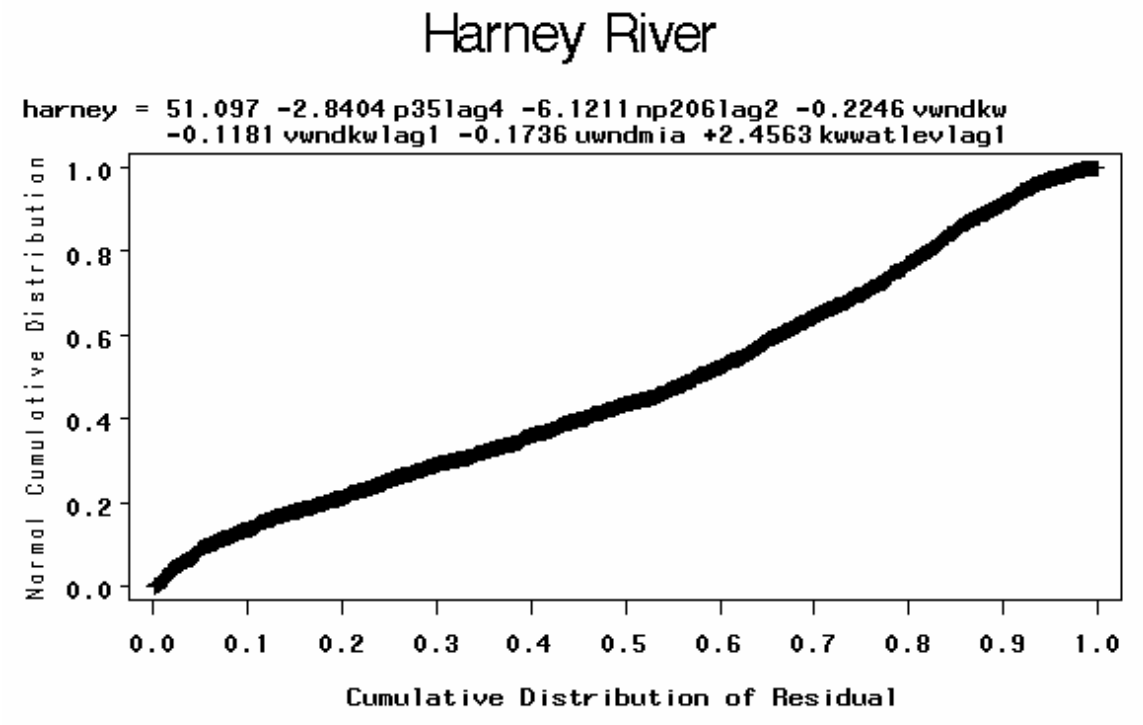
Cannon Bay


Cannon Bay
cannon = 38.128 -5.576 pa8lag2 -1.887 pa11lag2
Residual
Predicted Value
25
20
15
10
5
0
-5
-10
5
10
15
20
25
30
35
40

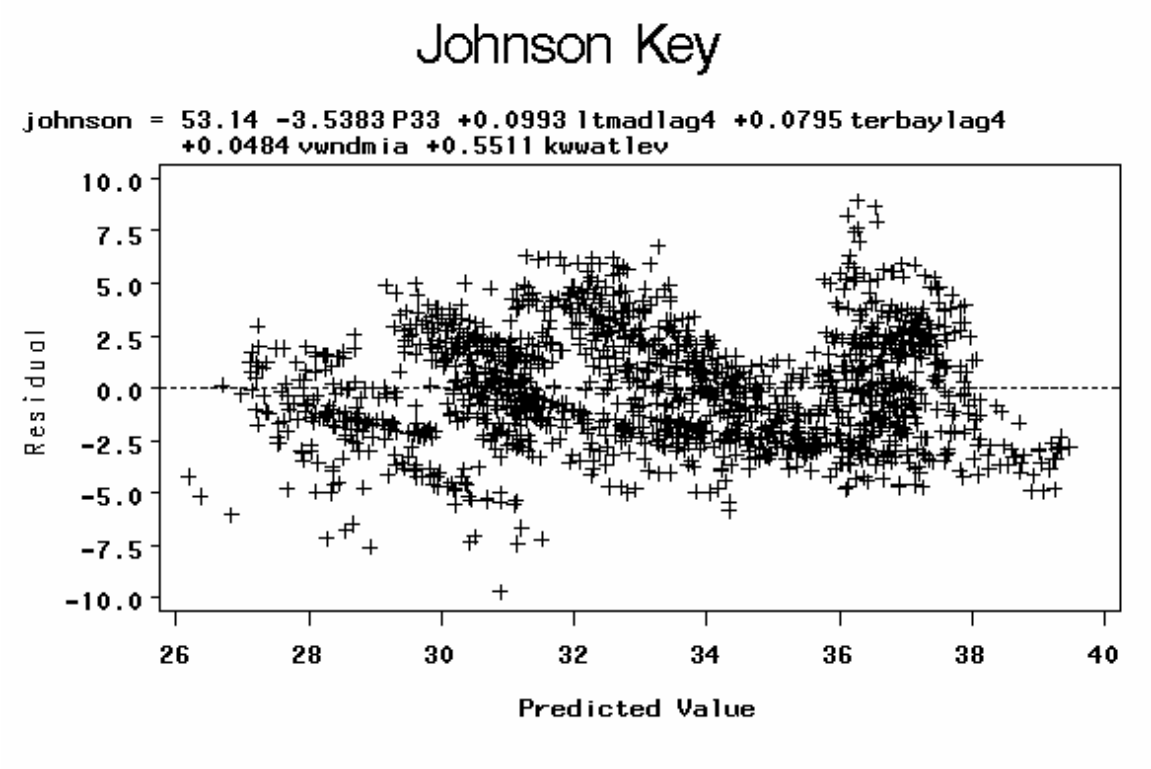

Cannon Bay
cannon = 38.128 -5.576 pa8lag2 -1.887 pa11lag2
Normal Cumulative Distribution
Cumulative Distribution of Residual
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

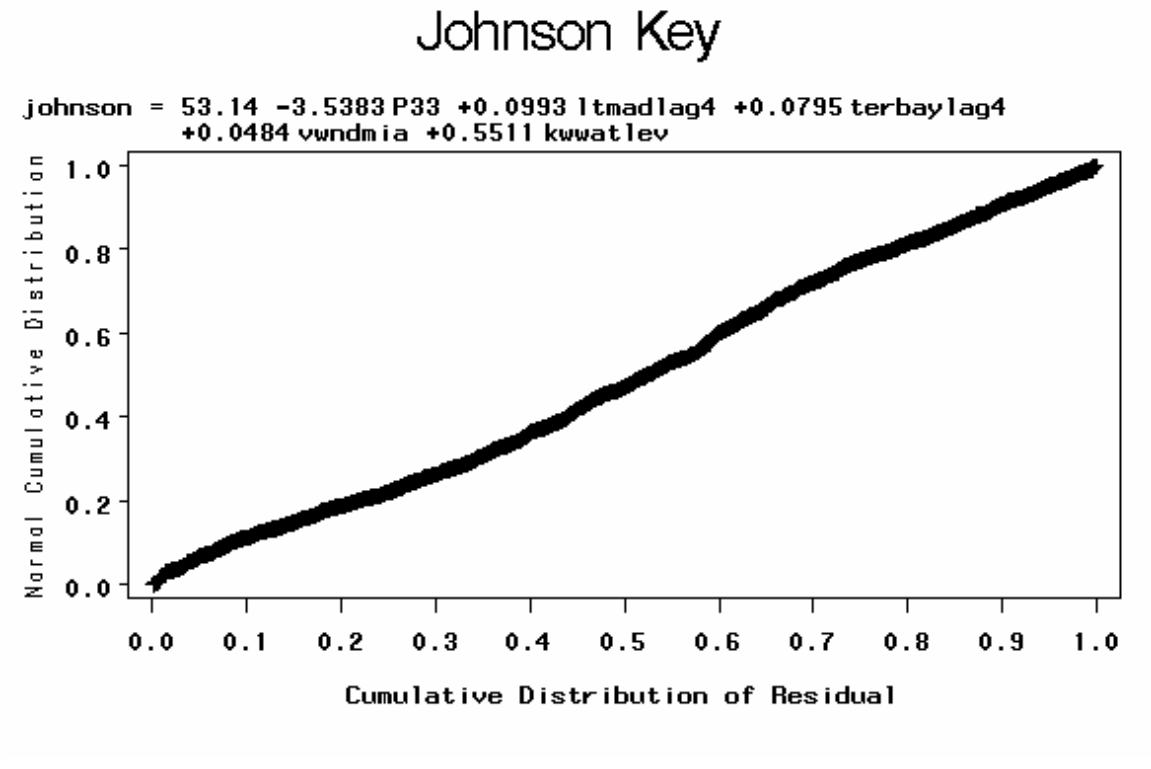
Harney River



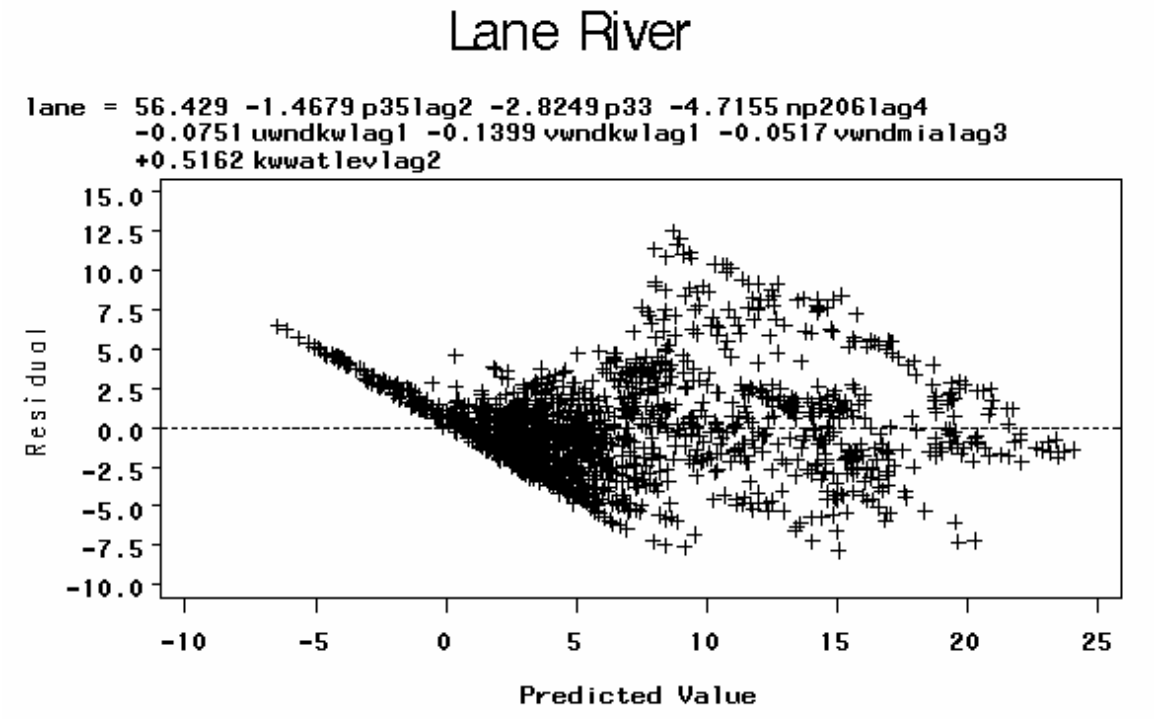


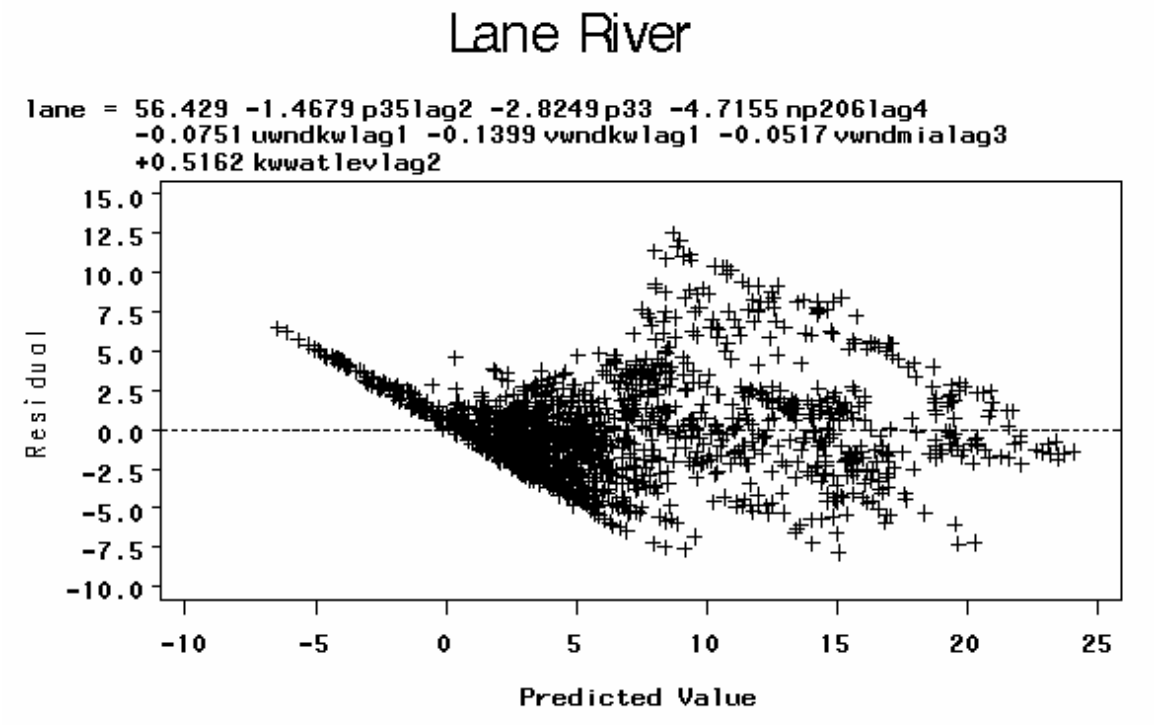
Johnson Key


Johnson Key
johnson = 53.14 -3.5383 P33 +0.0993 ltmadlag4 +0.0795 terbaylag4
+0.0484 vwndmia +0.5511 kwwatlev
Residual
10.0
7.5
5.0
2.5
0.0
-2.5
-5.0
-7.5
-10.0
26
28
30
32
34
36
38
40
Predicted Value

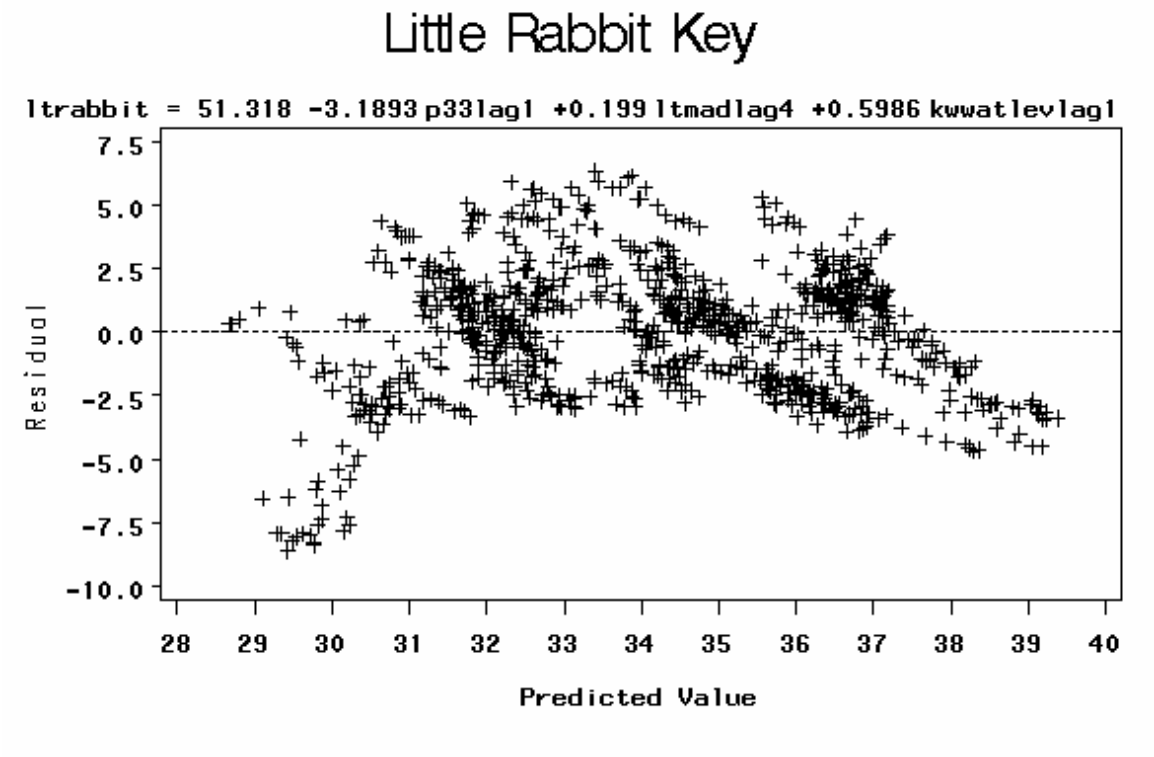

Johnson Key
johnson = 53.14 -3.5383 P33 +0.0993 ltmadlag4 +0.0795 terbaylag4
+0.0484 vwndmia +0.5511 kwwatlev
Normal Cumulative Distribution
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Cumulative Distribution of Residual

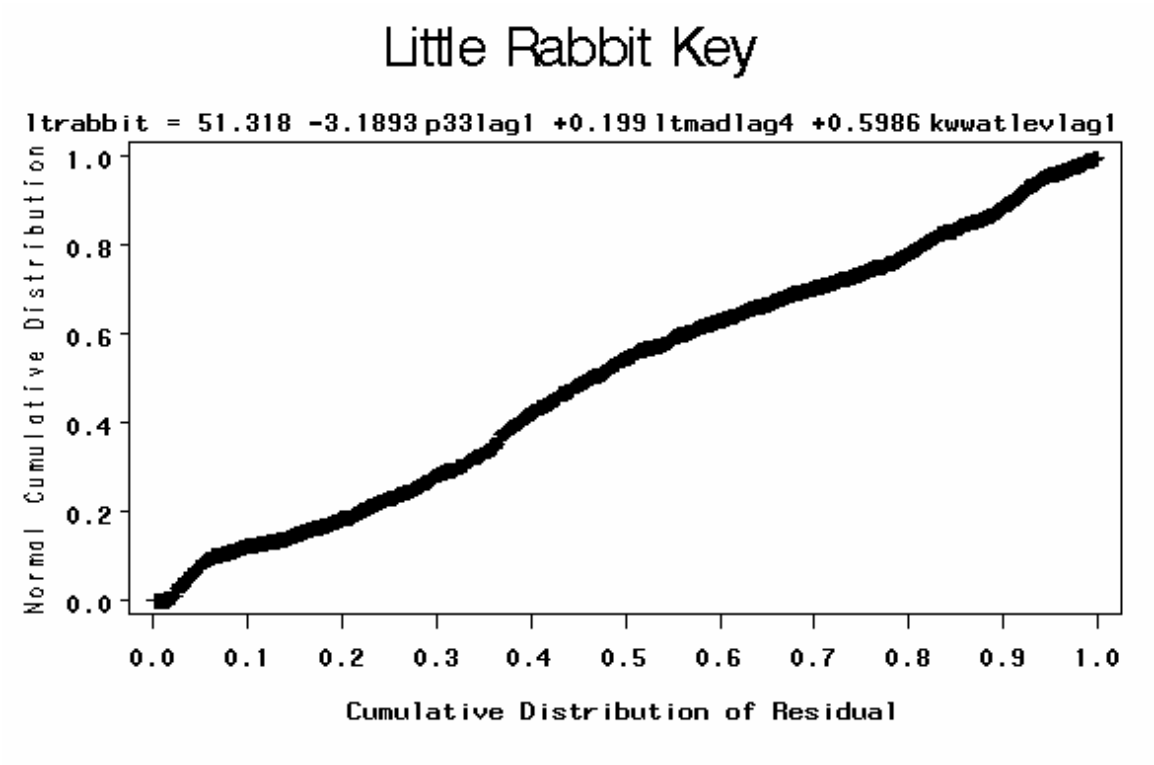
Lane River


Lane River
lane = 56.429 -1.4679 p35lag2 -2.8249 p33 -4.7155 np206lag4
-0.0751 uwndkwlag1 -0.1399 vwndkwlag1 -0.0517 vwndmialag3
+0.5162 kwwatlevlag2
Residual
Predicted Value

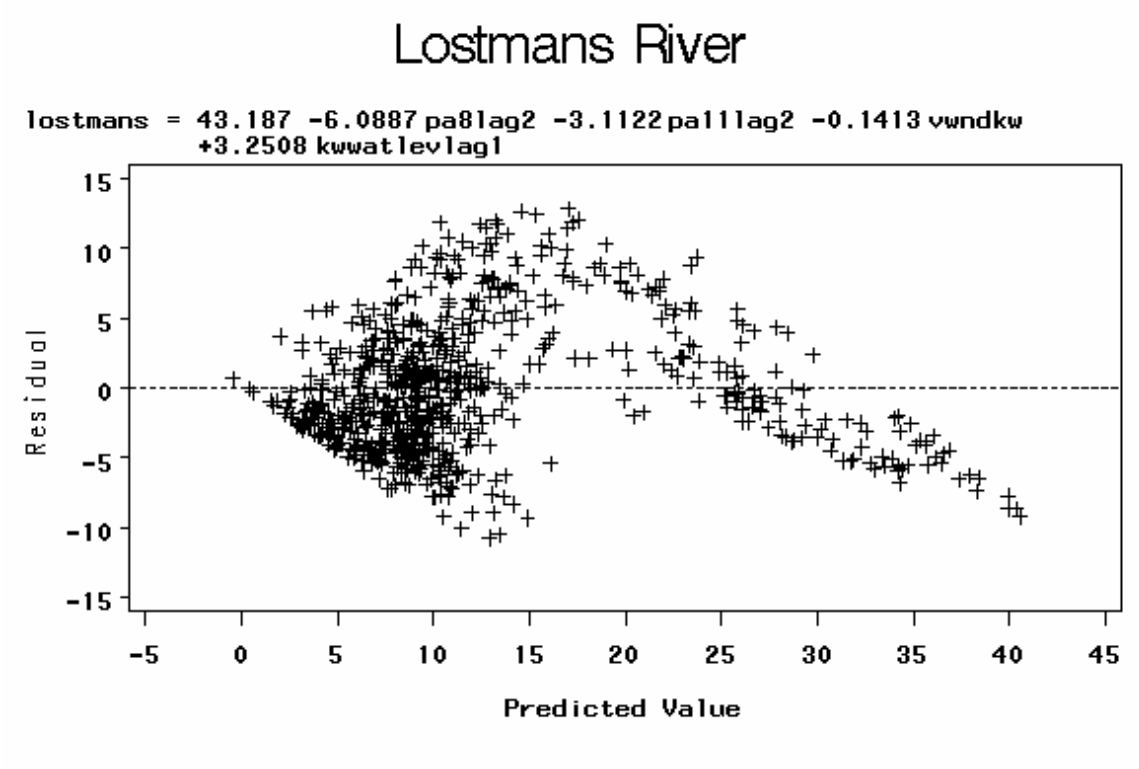

Lane River
lane = 56.429 -1.4679 p35lag2 -2.8249 p33 -4.7155 np206lag4
-0.0751 uwndkwlag1 -0.1399 vwndkwlag1 -0.0517 vwndmialag3
+0.5162 kwwatlevlag2
Residual
Predicted Value

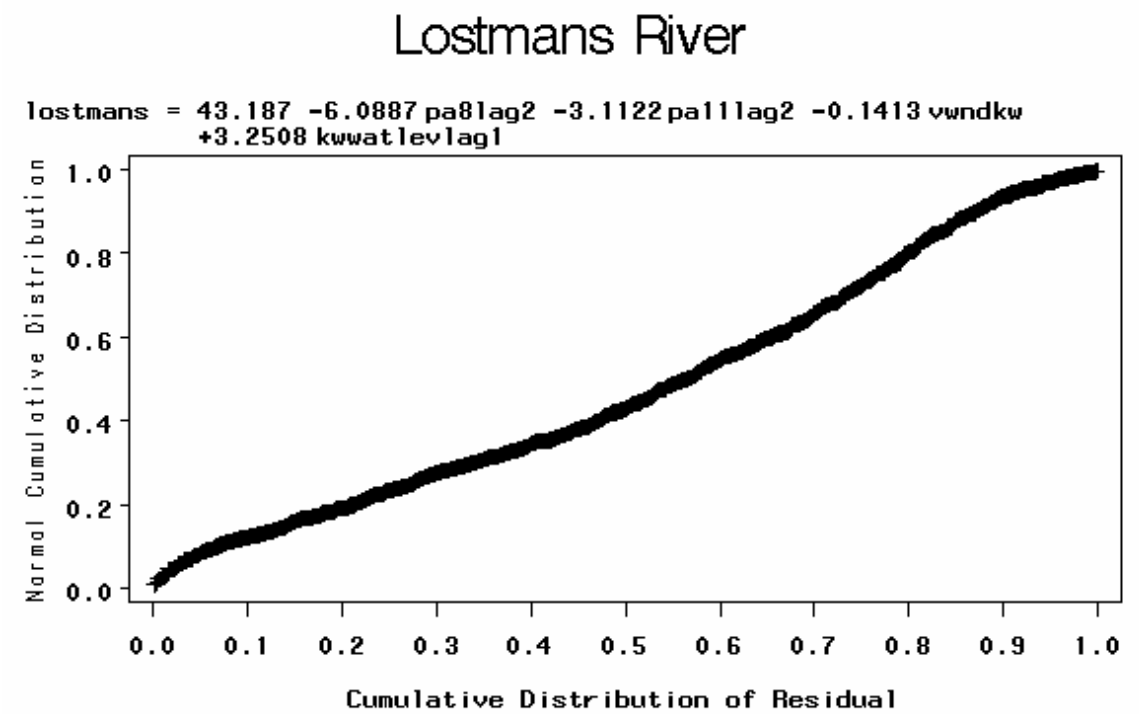
Little Rabbit Key



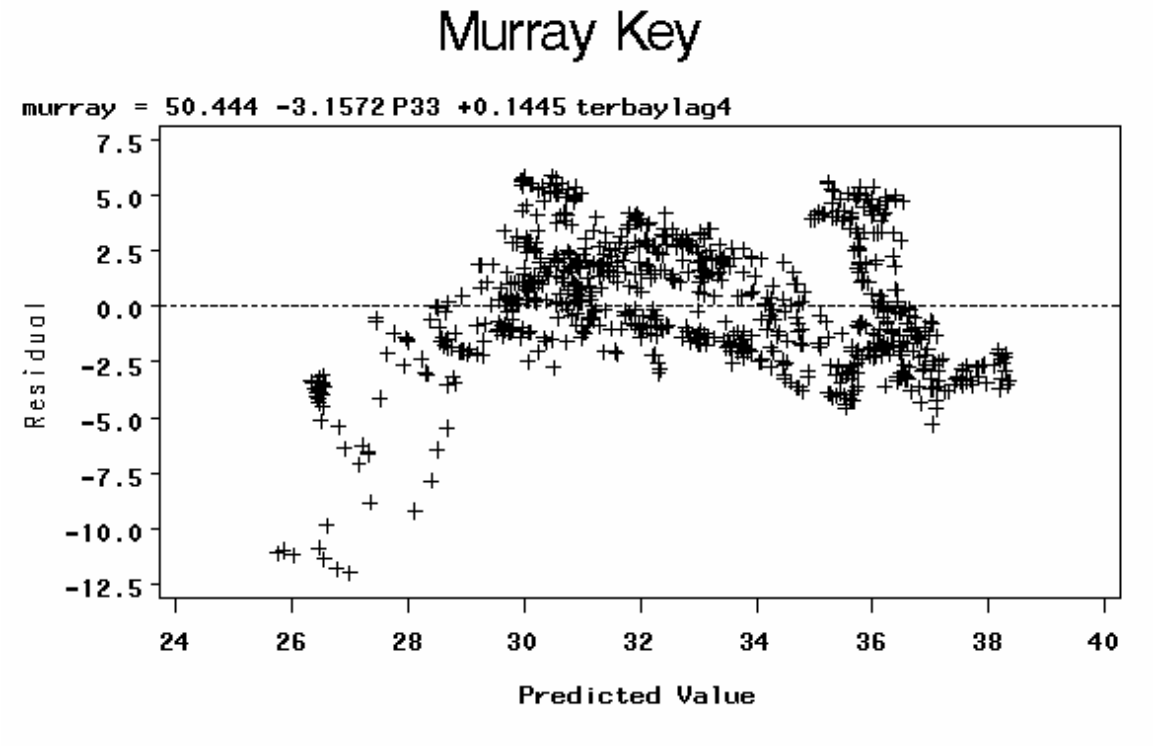


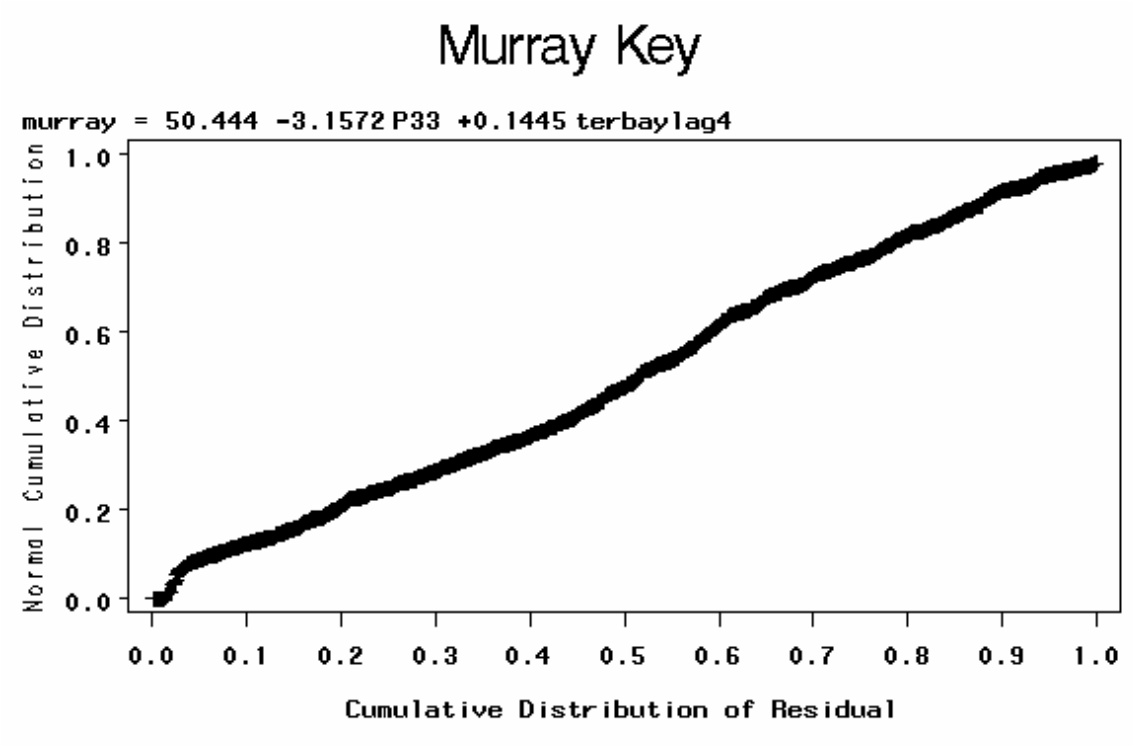
Lostmans River



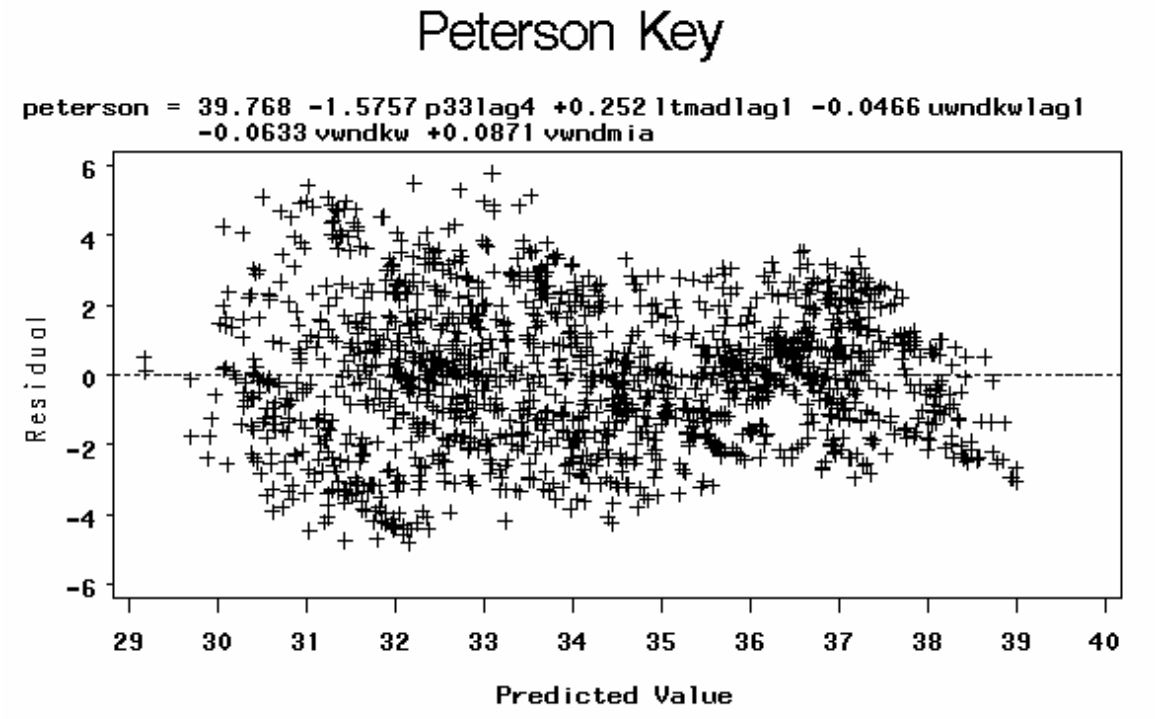


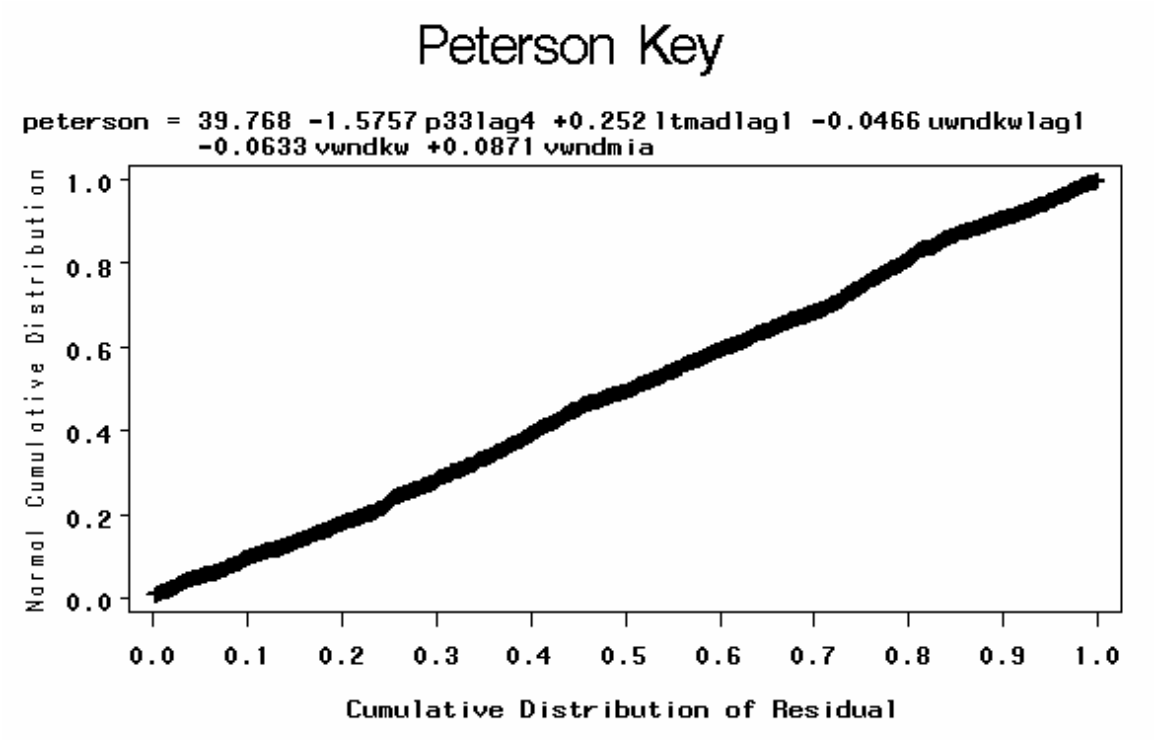
Murray Key


Murray Key
murray = 50.444 -3.1572 P33 +0.1445 terbaylag4
Residual
7.5
5.0
2.5
0.0
-2.5
-5.0
-7.5
-10.0
-12.5
24
26
28
30
32
34
36
38
40
Predicted Value

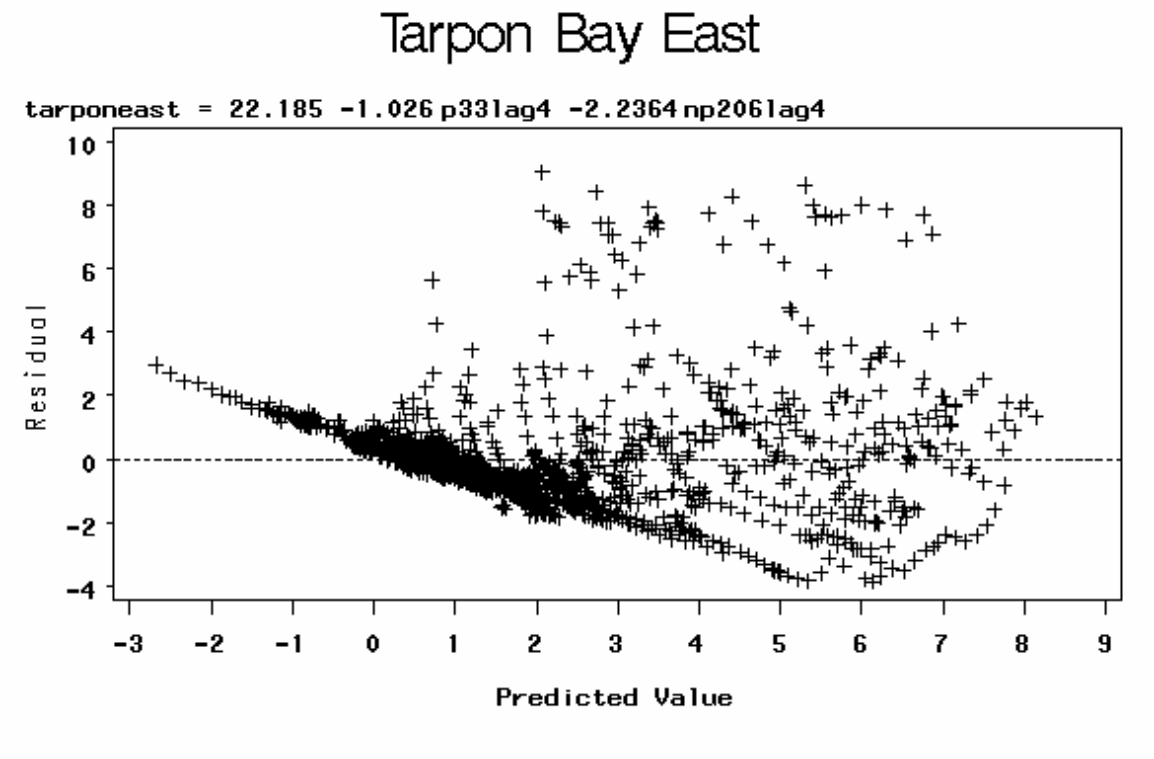

Murray Key
murray = 50.444 -3.1572 P33 +0.1445 terbaylag4
Normal Cumulative Distribution
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Cumulative Distribution of Residual

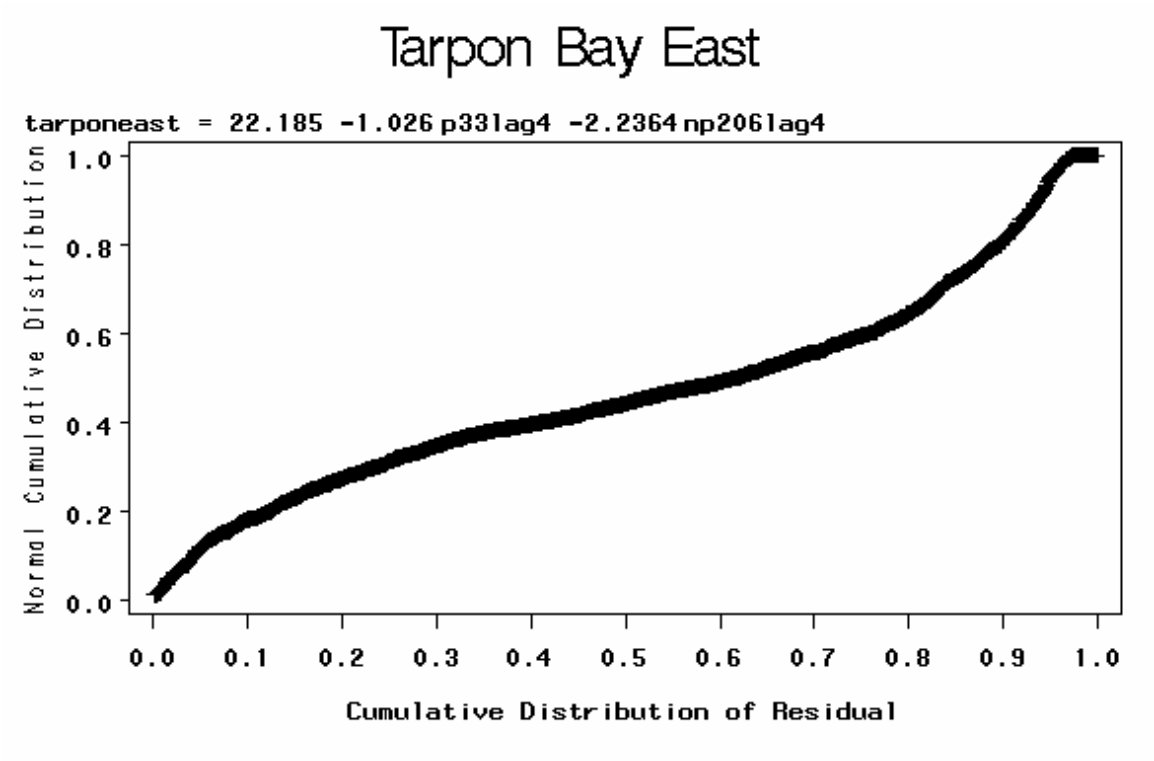
Peterson Key



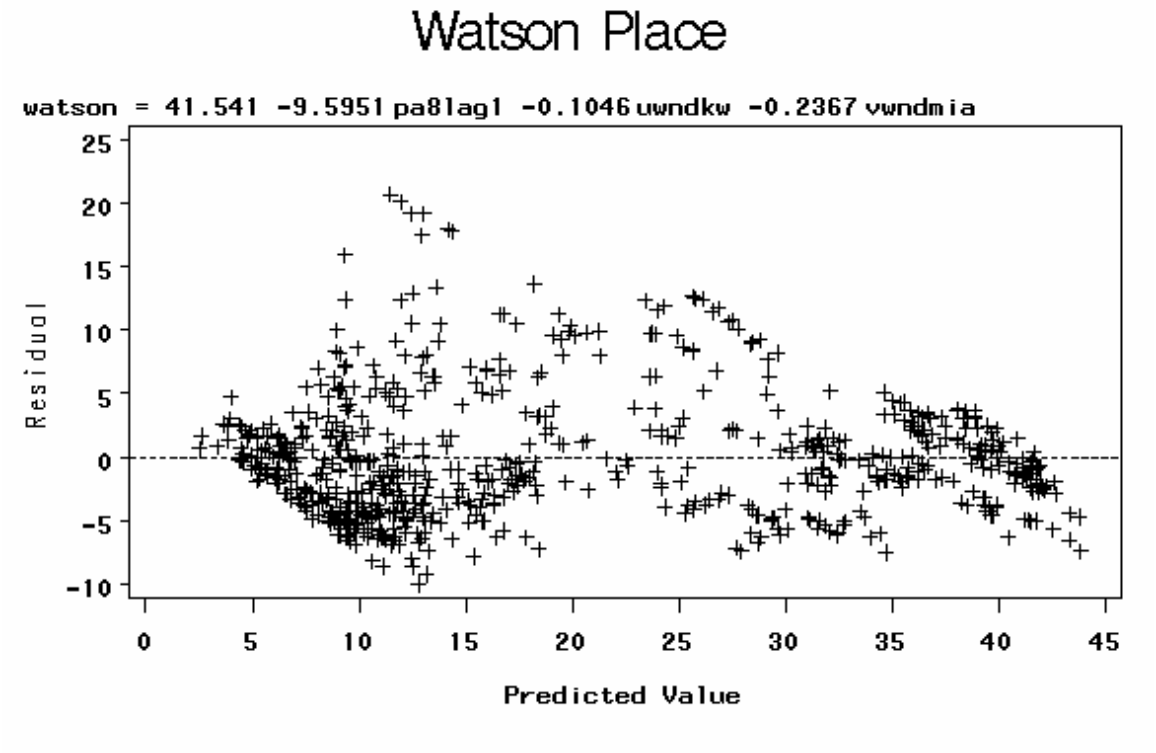


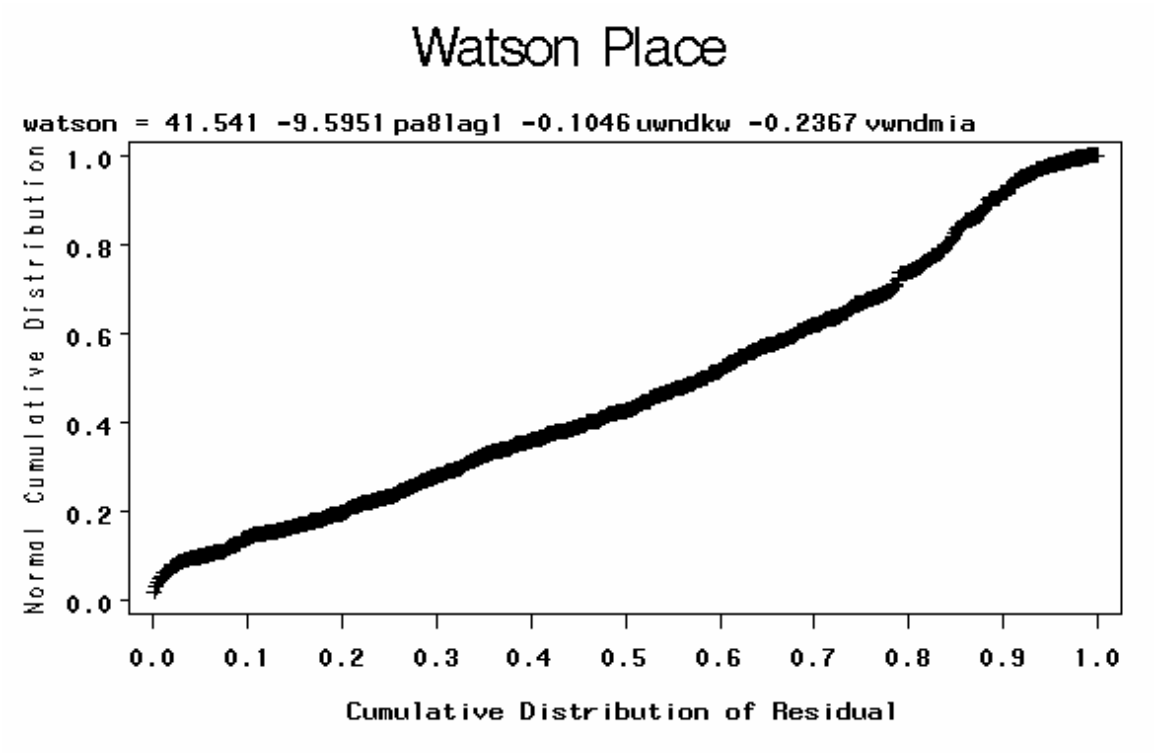
Tarpon Bay East



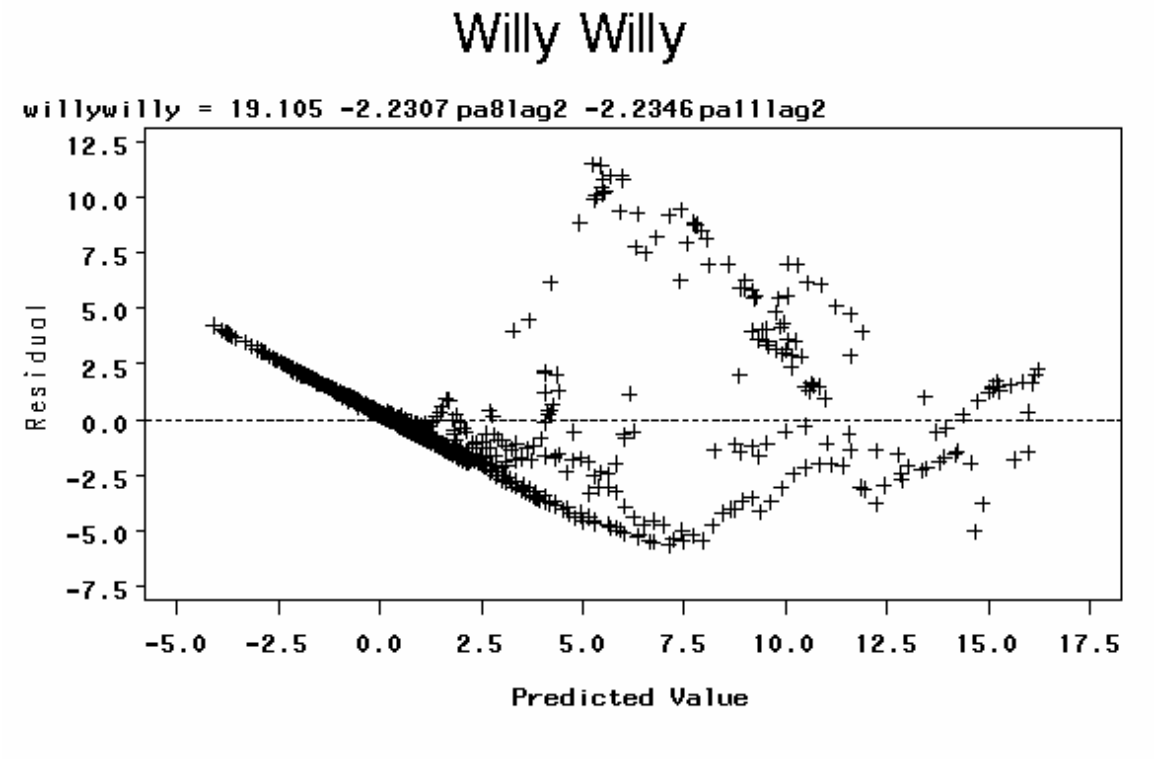


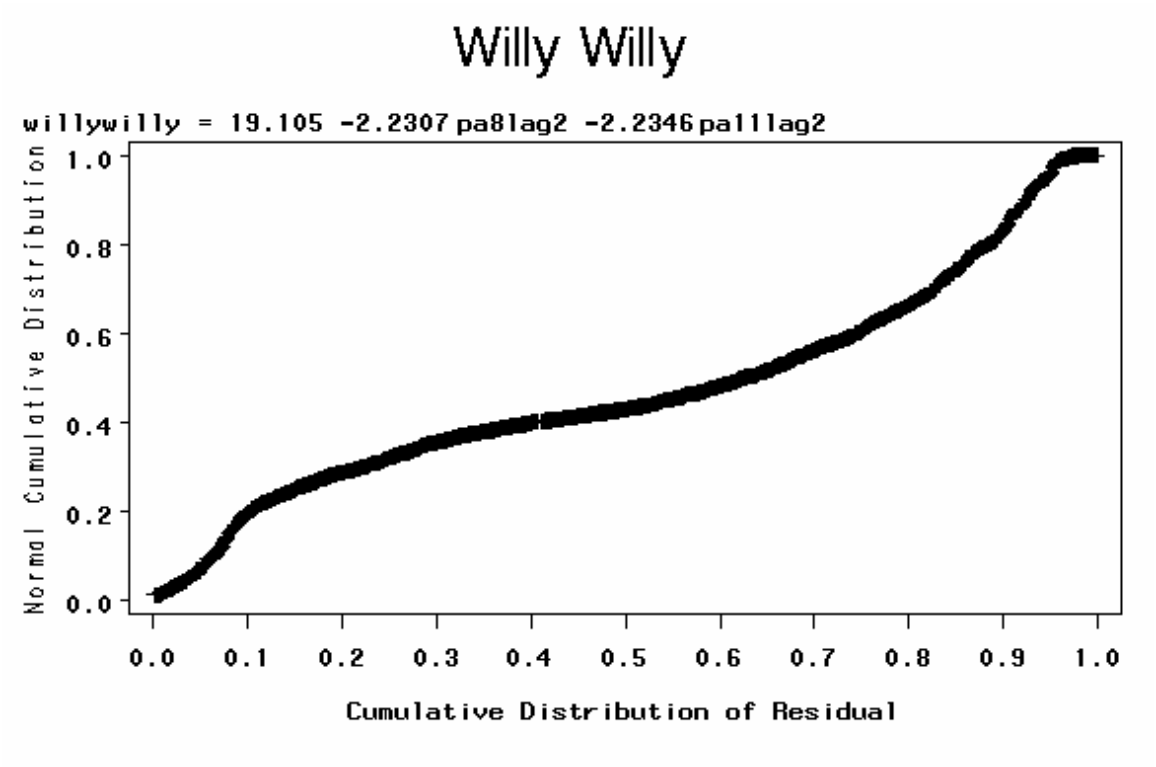
Watson Place





Willy Willy


Willy Willy
willywilly = 19.105 -2.2307 pa8lag2 -2.2346 pa11lag2
Residual
12.5
10.0
7.5
5.0
2.5
0.0
-2.5
-5.0
-7.5
-5.0
-2.5
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Predicted Value


Willy Willy
willywilly = 19.105 -2.2307 pa8lag2 -2.2346 pa11lag2
Normal Cumulative Distribution
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Cumulative Distribution of Residual

FINAL TASK REPORT

TASK 4 - ASSIST ENP WITH IMPLEMENTATION OF THE FATHOM FLORIDA BAY MFL MODEL

Cooperative Agreement Number CA H5284-05-0006
Between
The United States Department of the Interior National Park Service
Everglades National Park
And
Cetacean Logic Foundation, Inc.

Frank E. Marshall, III
Task / Project Manager

B. J. Cosby
Project Consultant (FATHOM)

DeWitt Smith
Project Coordinator

February 17, 2008

CESI Project No. CA H5284-05-0006 for Everglades National Park

FINAL TASK REPORT

TASK 4 - ASSIST ENP WITH IMPLEMENTATION OF THE FATHOM FLORIDA BAY MFL MODEL

I. Introduction

A. General

This report describes the activities that were completed for a task in the current Critical Ecosystem Studies Initiative (CESI) project that was originally initiated by Cetacean Logic Foundation, Inc. (CLF) for Everglades National Park (ENP) in 2002. The goal of this on-going CLF research is the characterization and simulation of the salinity regime in Florida Bay and the estuaries along southwest coast of the Gulf of Mexico (Gulf) within the Park. It is intended that this research yield information regarding the link between the downstream salinity in the estuaries of ENP and the upstream freshwater hydrology of the Everglades.

The subject of this report is the assistance that the project team provided to ENP in the implementation of the FATHOM model that had been updated by the South Florida Water Management District (SFWMD) for the Florida Bay Minimum Flows and Levels (MFL) project. This task report describes the assistance provided, the sample model runs that were made, and a comparison to observed data.

B. Task Objectives and Evaluation Methods

The objective of this task is presented below as it appears in the modified cooperative agreement for this CESI project approved by ENP:

> "The Principal Investigator will meet with ENP to review and discuss the current use of FATHOM and to review up to five (5) FATHOM input files provided by ENP. The Principal Investigator will interpret the output and provide the findings of the analysis in a report.
>
> The deliverable is a task report that discusses the current use of FATHOM, reviews FATHOM input files, and interprets the output provided by ENP after running FATHOM (ENP runs FATHOM)."

II. Background

FATHOM is a dynamic, spatially explicit, mass-balance model designed to investigate the response of salinity in Florida Bay to runoff, climate and variation in salinity on the Florida Shelf (Cosby et al. 1999, Nuttle et al. 2000, Cosby et al 2004). The model maintains a running account of the water and salt budgets in each well-mixed basin within the Bay. Circulation within Florida Bay and

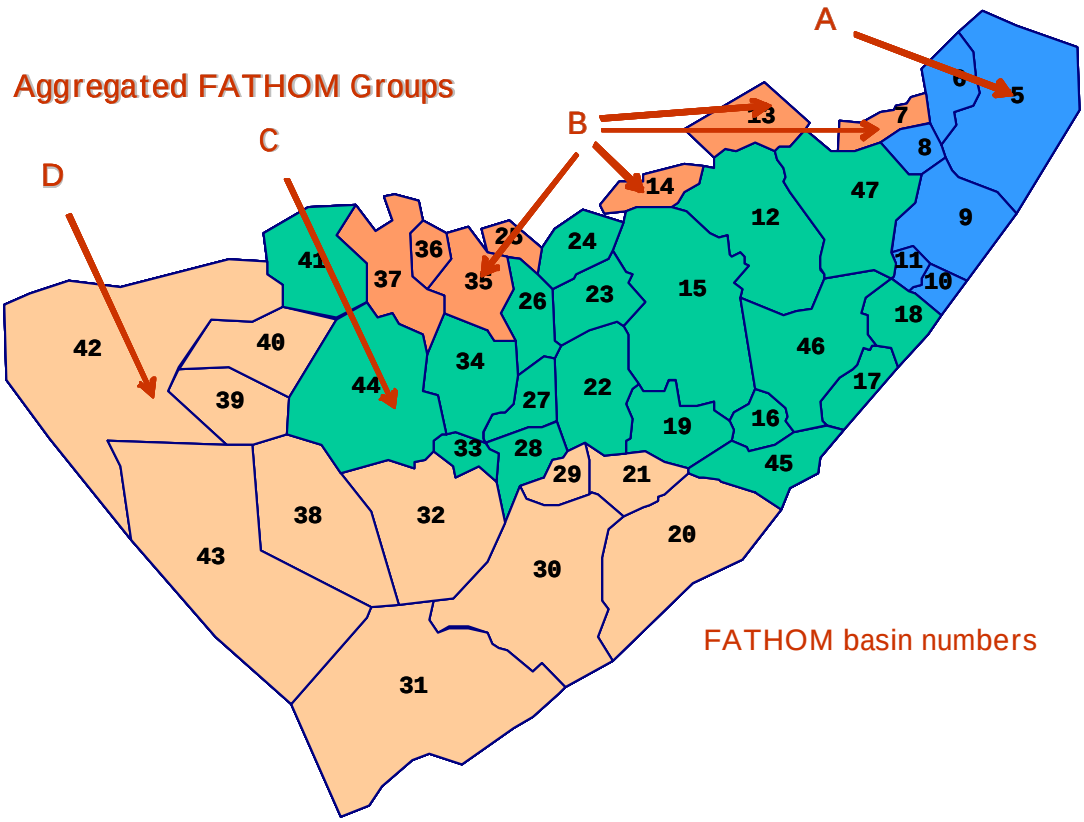
exchange with the Florida Shelf are controlled by the Bay's network of shallow banks. The basins defined by these banks offer a natural framework for mass-balance accounting.

FATHOM represents Florida Bay as a collection of well-mixed basins. Circulation and exchange are driven primarily by tides imposed along the western boundary. At each time step, the model solves for uniform hydraulic flow across each bank based on the depth, width, and frictional roughness of the bank, and water levels in the upstream and downstream basins. By this mechanism, tidal forcing at the boundary propagates into the Bay and drives the exchange of water and solutes among the basins. Solute fluxes are then calculated from water fluxes and the salinity of water on each bank. Details of the representation of flow over the banks and the hydraulic equations are given in Cosby et al. (1999).

Despite the model's computational simplicity, FATHOM requires highly detailed information about the bathymetry in Florida Bay. Bathymetric data are entered into a GIS database that classifies the depth for every 20 by 20 meter rectangle in the Bay (the pixel resolution of the data) into one of 11 classes: 1 land surface class and 10 depth classes. The depth classes covered the range from 0 to 10 feet in one-foot increments. Areas deeper than 10 feet occur locally in the Bay, but (based on these data) such areas are not extensive and occur infrequently. These depth-class data provide the basis for describing the hypsometry for each basin and the depth and widths of the shoals for the calculation of exchange fluxes between basins.

The total length of the line segments that describe the shallow banks is 626 km. Along the boundaries, the line segments follow the mainland coastline and the keys, and in the interior of the Bay the lines traverse or connect islands. About 276 km of the "shoals" defined for FATHOM are dry land (no-flow boundaries). Greater than 70% (252 km) of the wetted length of the shoals is shallower than 1 foot. A small proportion of the 21% of shoal length with depth greater than 3 feet represents cuts and channels through the shoals. Most of the deeper shoals are located in the western and southwestern part of the Bay. These were defined as shoals to establish a basin boundary in the model.

Figure 1. The numbered polygons identify the basins used in FATHOM calculations. The color-coded regions and the alphabetic destinations identify the FATHOM groups used to summarize the results of salinity calculations.



The solution scheme depends on the calculation of water velocity independently for each depth increment of each shoal. Conceptually, water flow from one basin to another over a bank in the Bay can be treated as flow over a broad crested weir with friction. Flow velocity across the weir depends on the difference in water levels on each side of the weir and (perhaps) the occurrence of critical flow. If critical flow occurs, then velocity does not depend on the water level on the outflow side. The influence of bottom friction may or may not have an important effect, depending on the velocity and the depth of flow. The difference in water levels across the weir provides the specific energy for the flow.

There is no direct simulation of wind shear on the water surface in FATHOM. A key aspect of the conceptual basis of FATHOM is the assumption that the major effects of wind shear on circulation in Florida Bay can be accounted for by the structural assumptions and the inputs to the model, without having to be modeled explicitly. Effects of wind have been incorporated in FATHOM in two ways: 1) by the assumption that each basin is well mixed (i.e., the "local" effect of the wind is to stir the basins, but not to move much water among them); and 2) by the incorporation of water level changes in addition to tides and sea-level patterns at the Gulf and Atlantic ocean boundaries (i.e., the "remote" effect of the wind is to cause a sea level setup of varying magnitude and duration along the different parts of the Bay boundaries).

As a consequence of the assumptions described above, salinity calculated by FATHOM represents a time-averaged value with a period of about one month. Even though circulation and exchange in the model are driven by tides, data on other processes such as rainfall and freshwater inflow are provided as monthly values. As well, the assumption of basins as well mixed imposes constraints on the time scale on which salinity calculations can be taken as comparable to observations at any particular location.

Bay-wide inputs required by FATHOM include time series of rainfall and evaporation for each basin in the Bay. The model structure allows these inputs to be specified individually for each basin to reproduce spatial gradients in these forcing functions. In practice, however, observed data are not sufficient to support more than a regional approach to the spatial distribution of climate inputs. That is, the Bay must be divided into a few regions for which climate inputs are applied uniformly to the model. Groundwater inputs to the basins can also be specified, but these have not been employed in the simulations performed for this project.

Time series of freshwater inflow volumes are required at the terrestrial boundaries of the Bay. Inflow (where it occurs) is specified as an input separately into each of the boundary basins along the Everglades coastline, though not all of the boundary embayments receive inflow. To compare the effects on the Bay of different runoff regimes, the distribution of inflow among the small embayments at the northern edge of the Bay was varied and analyzed. Along the

Keys, inflows of fresh water are small, and these are not included in the FATHOM inputs.

In addition to the runoff data at the terrestrial boundaries, FATHOM requires tide, sea level and salinity time series to set the open water boundary conditions for the Bay. The model allows these boundary conditions to vary spatially along the boundaries.

For the SFWMD Florida Bay MFL model, constructing the FATHOM input data sets beginning in 1970 entails making choices among alternative methods for assembling each component of the input data. For example, three methods for estimating the time series of rainfall over the Bay were considered. Model performance with different choices of inputs is evaluated by comparing calculated salinity with salinity measurements that are available for the period 1991 through 2002. A systematic sensitivity analysis verified the choice of the "best" combination of inputs that ultimately comprised the MFL base case, which was then used by SFWMD for MFLs. The sensitivity analysis (Marshall et al, 2005) quantified the uncertainty in calculated salinity related to the uncertainty in the input data sets.

The "best" input parameters from the iterative analyses were chosen to be the MFL base case. When the uncertainty statistics were computed, over the entire Bay the FATHOM MFL base case model was capable of explaining about 81% of the monthly salinity variability. However, for some of the basins, particularly in the Northeast and Central Regions, model fidelity showed improvement through the use of rainfall and inflow alternatives other than the base case parameters. This suggests that FATHOM may be able to be improved further through the use of area specific model parameters in some basins.

III. Methods

The objective of this task is for the FATHOM consultant (Dr. B. J. Cosby) to meet with ENP to review and discuss the current use of FATHOM and to review up to five (5) FATHOM input files provided by ENP. The FATHOM consultant interpreted the output and provide the findings of the analysis in a report. Listed below are the notes from those meetings and the results of the discussions.

<u>Sept 2006 Meetings.</u>

- Met with ENP staff at SFNRC office in Homestead.
- Delivered FATHOM model and all MFL base simulation inputs.
- Conducted a 2-day workshop to instruct ENP staff in the implementation of the FATHOM model.

- Instructed ENP staff in the preparation of inputs for the FATHOM model so that the output of the SICS/TIME model could be used as an input to FATHOM.
- Instructed ENP staff in the post-processing of FATHOM output, particularly the comparison of simulated salinity to observed salinity.
- Collected MMN data from ENP staff to prepare spreadsheet for comparison of FATHOM salinity to MMN observations. (meeting also related to Task 2)

September 2006 through July 2007 Activities.

- Participated in phone calls and exchanges of model inputs and outputs through ftp sites to assist ENP staff in the implementation FATHOM with SICS/TIME outputs.
- Examined and debugged FATHOM outputs produced by ENP staff. Advised ENP staff on procedures for adapting the SICS/TIME output to FATHOM.
- Posted a spreadsheet to the ENP ftp site for comparison of FATHOM output to MMN observed data.

July 2007 Meetings.

- Met with ENP staff in Homestead to review the data and programs developed and exchanged over the preceding 10 months.
- Discuss the new 58 basin configuration of FATHOM. (meeting also related to Task 2)

IV. Results

FATHOM MFL model output was compared to MMN observed data as part of the exercise of assisting ENP with full implementation of the FATHOM MFL model. Simulated salinities from those runs were compared to observed salinity for the period 1991-2002 in the following plots. Two observed salinity datasets are used, the SERC dataset and the MMN dataset. Coordination and assistance in use of the post-processing routines from the software to produce the comparisons of simulated salinity to MMN data is one of the objectives of this task.

Comparisons are in the form of selected statistics of goodness-of-fit for all FATHOM basins for which observed data are available, and in the form of time series plots of simulated and observed salinities for selected basins. Full details of these statistics and plots for all pertinent basins are included in the spreadsheets delivered to ENP. The figures below are extracted from those spreadsheets as examples. Figures 2 – 5 present example comparisons for the MFL base simulations. The SERC data are monthly grab samples and the MMN data used for the plots are monthly average data from continuous observations.

In general, there are few substantive differences when FATHOM MFL base case output is compared to either SERC or MMN data. The spatial coverage of the SERC data is broader than the spatial coverage of the MMN data set. The temporal coverage of the MMN data is broader then the temporal coverage of the SERC data because the MMN data are gathered more frequently than the SERC data, and the period of record is longer at most MMN stations.

From the statistics that were prepared it is noted that the values for efficiency for SERC and MMN data are similar but values of R^2 for SERC comparisons during the dry season are lower than MMN data. When the basins are grouped, the R^2 values for the MFL base case compared to the MMN data are lower than the comparisons to the SERC data. Exceptions to this are grouped basins C and D during the wet season.

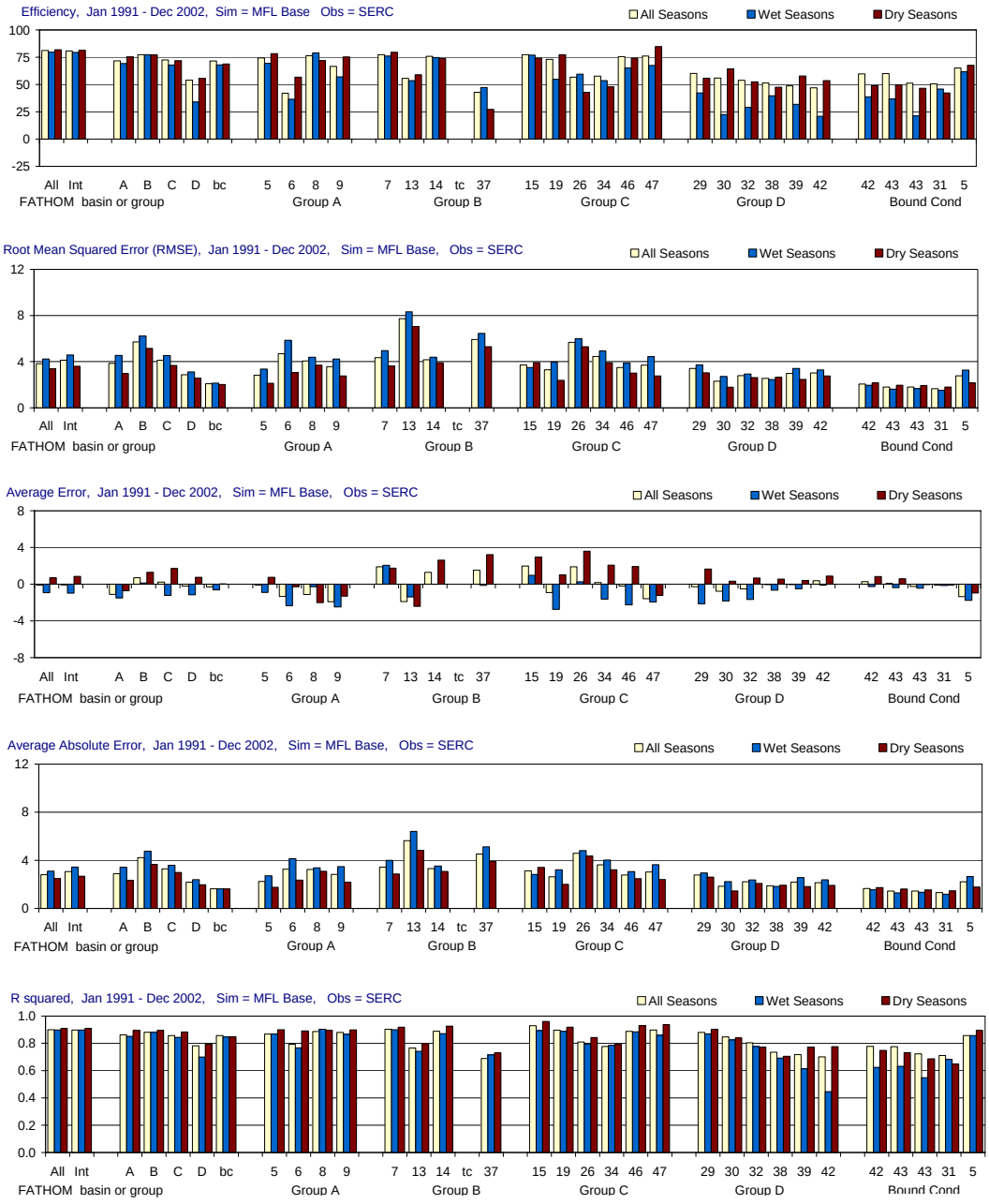


Figure 2. Selected statistics for comparing simulated and observed (SERC data) salinity values using the 47 basin configuration of FATHOM. The simulation is the SFWMD MFL base simulation covering the period 1991-2002.

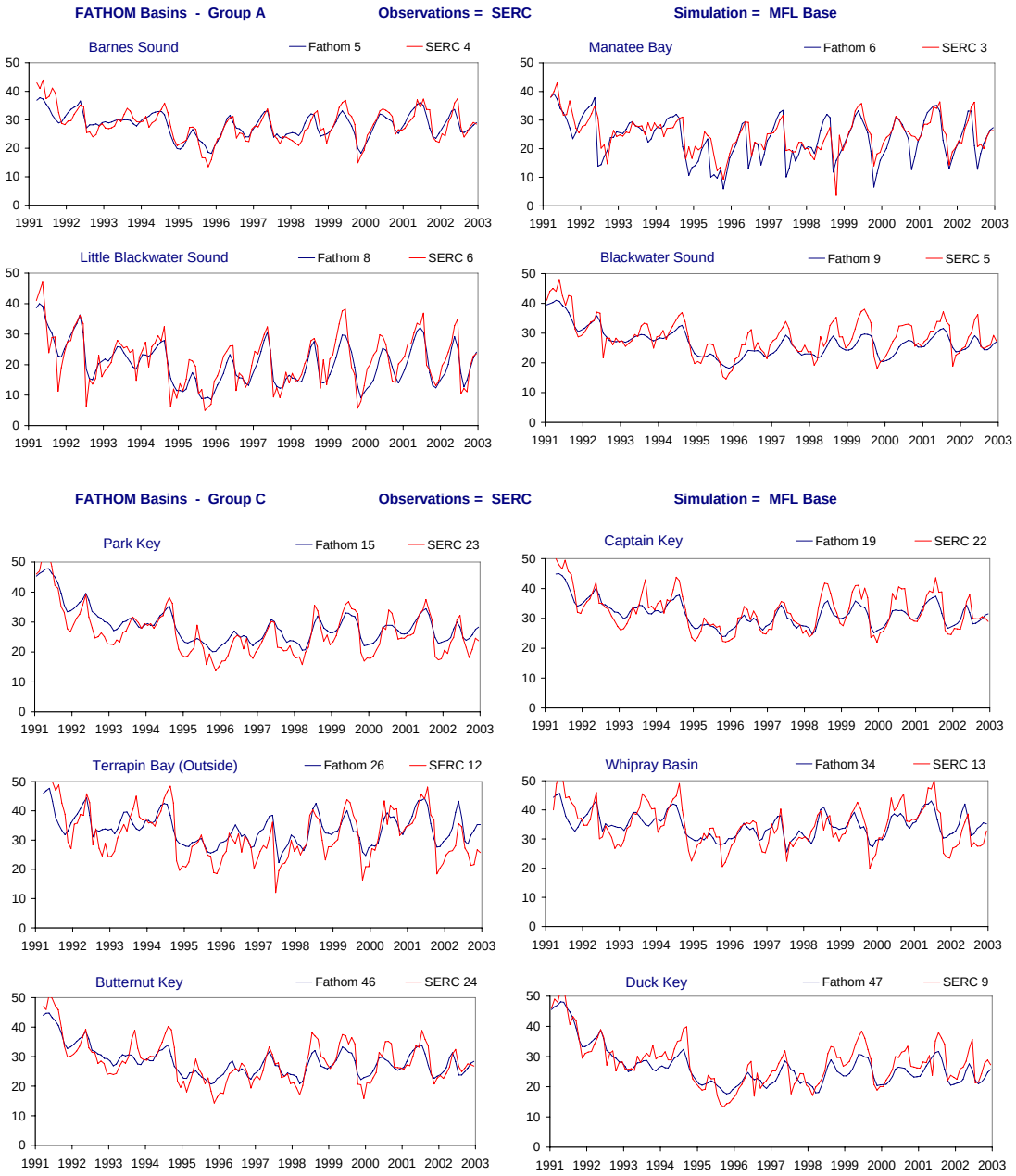


Figure 3. Comparison of simulated and observed (SERC data) time series of salinity values using the 47 basin configuration of FATHOM. The simulation is the SFWMD MFL base simulation covering the period 1991-2002. Selected basins are displayed

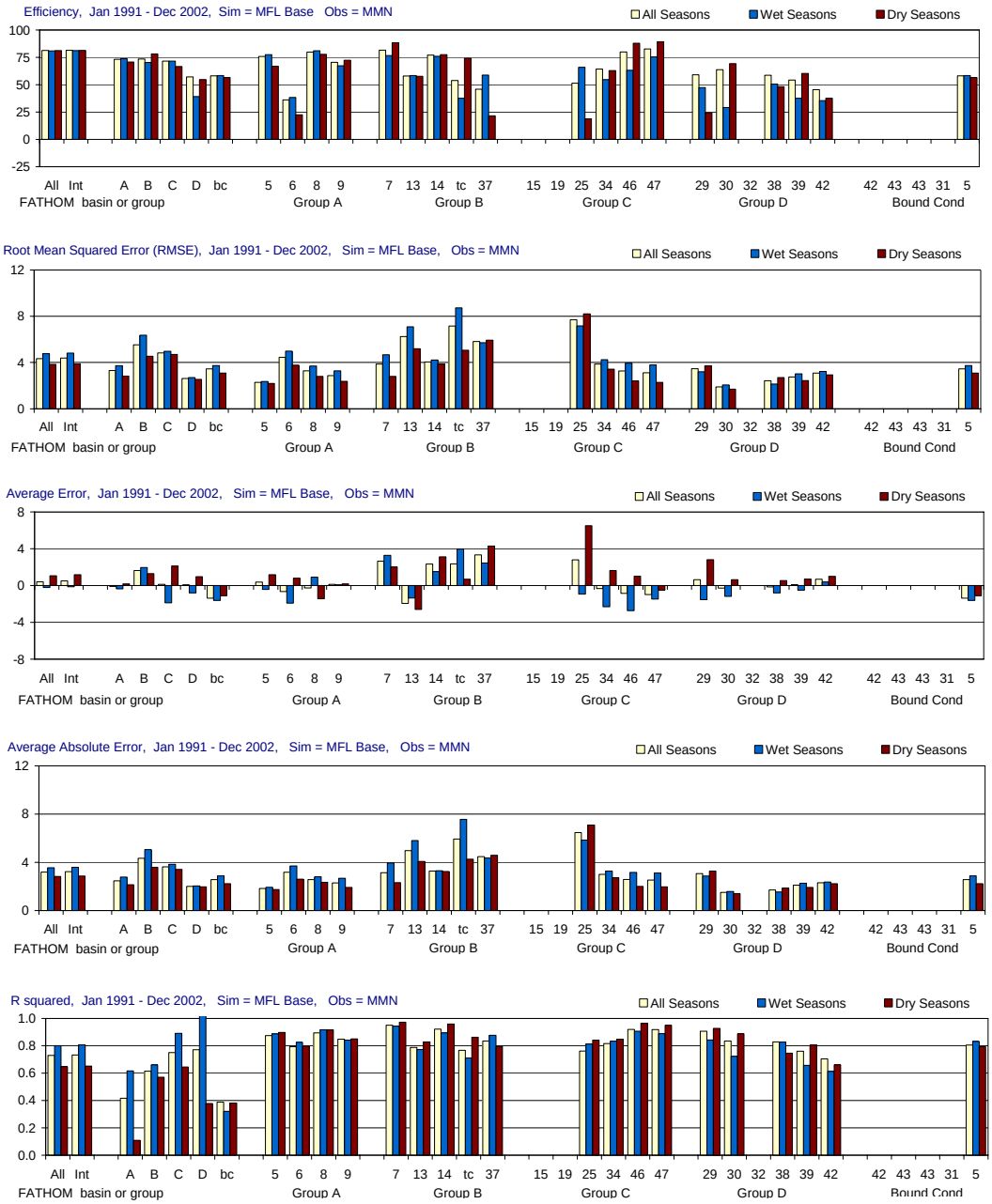


Figure 4. Selected statistics for comparing simulated and observed (MMN data) salinity values using the 47 basin configuration of FATHOM. The simulation is the SFWMD MFL base simulation covering the period 1991-2002.

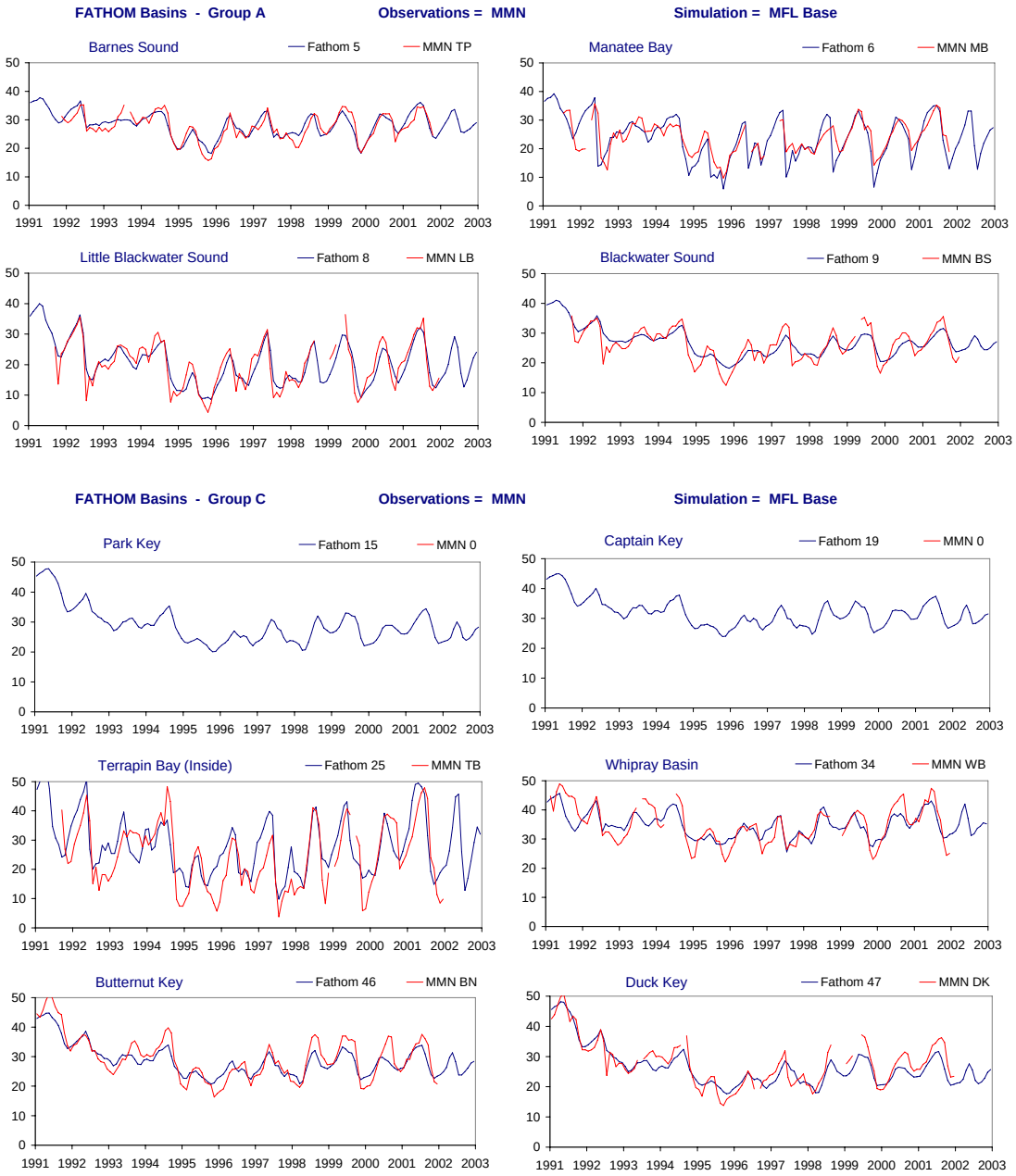


Figure 5. Comparison of simulated and observed (MMN data) time series of salinity values using the 47 basin configuration of FATHOM. The simulation is the SFWMD MFL base simulation covering the period 1991-2002. Selected basins are displayed

V. Findings

The FATHOM consultant met with ENP to review and discuss the current use of FATHOM and reviewed FATHOM input files provided by ENP. The FATHOM consultant assisted ENP in interpretation of the output, and provided a report that included statistics and plots from comparisons of FATHOM output from the MFL base case model to both SERC and MMN observed data.

The most significant accomplishment of this task was to assist ENP with implementing the FATHOM MFL software at South Florida Natural Resources Center. New spreadsheets were developed to analyze FATHOM results relative to ENP MMN data as well as SERC data.

VI. References

CERP, 2002. Model Uncertainty Workshop Report. Comprehensive Everglades Restoration Plan. Available from: <http://www.evergladesplan.org/pm/recover/recover_docs/mrt/052402_mrt_uncertainty_report.pdf> [cited 20 May 2005].

Cosby, B., W. Nuttle, and F. Marshall. 2005 FATHOM Enhancements and Implementation to Support Development of MFL for Florida Bay. Final Report on Contract C-C-15975-WO05-05 for the South Florida Water Management District. Environmental Consulting & Technology, Inc. New Smyrna Beach, Florida.

CROGEE. 2002. Florida Bay Research Programs and Their Relation to the Comprehensive Everglades Restoration Plan. The National Academies Press, Washington, D.C.

NRC, 2004. Adaptive Management for Water Resources Project Planning. National Academies Press, Washington, D.C. Available from: <http://www.nap.edu/catalog/10972.html> [cited 20 May 2005].

Thom, R.M., H.L. Diefenderfer, and K.D. Hofseth, 2004. A Framework for Risk Analysis in Ecological Restoration Projects. IWR Report 04-R-02.
Twilley, R. [n.d.] Evaluating the Response of Mangrove Wetlands to Alterations of Freshwater Flow in Everglades National Park. [online] Available from: <http://www.ucs.louisiana.edu/~rrt4630/mangrove-restudy.htm> [cited 20 May 2005].

FINAL TASK REPORT

TASK 5 - INVESTIGATE THE COUPLING OF PINK SHRIMP MODELS WITH MLR SALINITY MODELS

Cooperative Agreement Number CA H5284-05-0006
Between
The United States Department of the Interior National Park Service
Everglades National Park
And
Cetacean Logic Foundation, Inc.

Frank E. Marshall, III
Principal Investigator

Joan Browder, NOAA
Pink Shrimp Modeling

DeWitt Smith
Project Coordinator

February 17, 2008

CESI Project No. CA H5284-05-0006 for Everglades National Park

FINAL TASK REPORT

TASK 5 - INVESTIGATE THE COUPLING OF PINK SHRIMP MODELS WITH MLR SALINITY MODELS

I. Introduction

A. General

This report describes the activities that were completed for a task in the current Critical Ecosystem Studies Initiative (CESI) project that was originally initiated by Cetacean Logic Foundation, Inc. (CLF) for Everglades National Park (ENP) in 2002. The goal of this on-going CLF research is the characterization and simulation of the salinity regime in Florida Bay and the estuaries along southwest coast of the Gulf of Mexico (Gulf) within the Park. It is intended that this research yield information regarding the link between the downstream salinity in the estuaries of ENP and the upstream freshwater hydrology of the Everglades.

The subject of this report is the investigation into the use of salinity from multivariate linear regression (MLR) salinity models as input to the pink shrimp growth dynamics model of Browder et al (2002). This task report describes the coordination between the modelers, the MLR salinity model output that was provided to the pink shrimp modelers, and the output obtained from the pink shrimp model using the MLR salinity simulations as input.

B. Task Objectives and Evaluation Methods

The objective of this task is presented below as it appears in the modified contract for this CESI project approved by ENP:

> "MLR salinity models developed by years one and two and IOP CESI projects were capable of reasonably simulating salinity in parts of Florida Bay and have been identified as means to provide daily salinity estimates for ecological models (RECOVER, draft January 2004). This task will evaluate the coupling of MLR salinity models and other statistical models with a pink shrimp model that was developed by Browder, et al (2002). The data requirements for the model will be identified, and the mechanism for coupling the models will be determined. A period of time for the evaluation of MLR salinity model performance in this mode will be identified. If possible (determined by the ability of others to run the pink shrimp model with salinity simulations produced by this continuation project), the MLR salinity models will be coupled with the pink shrimp model at three locations as a proof-of-concept, leading to the next step of using the ICU simulations to estimate the broader response in pink shrimp ecosystem dynamics. Other ecological

models that utilize salinity as an input parameter will be reviewed and a list of models that could use the simulations will be developed.

The deliverable is a task report that describes the activities involved with coupling MLR salinity models with the pink shrimp model, and the presentation of three output files that can be directly input to the pink shrimp model to produce simulations by others. If the output from the pink shrimp model using the simulated salinity as input is made available before the final report is prepared, the results of the coupling will be presented in the final report. A list of ecological models that could potentially utilize MLR salinity simulations as input will also be provided."

II. Background

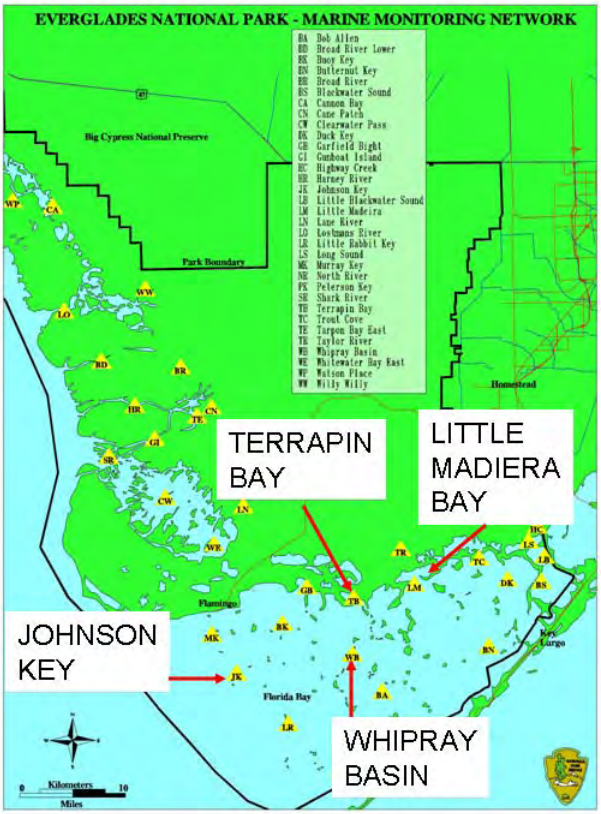
A simulation model of growth and survival was developed by Browder et al. (2002) to predict the response of pink shrimp (*Farfantepenaeus duorarum*) to changing salinity and temperature. This model has been applied in Florida Bay and other south Florida estuaries (Browder et al. 2005). Response variables include: average growth rate, average survival rate, and potential harvest (which integrates both growth and survival). In addition, the model provides estimates of the number of individuals at 30-day intervals, relating directly to density, once scaled in the calibration.

The pink shrimp stimulation models for Whipray Basin and Johnson Key (Figure 1) will be calibrated using part of the data from Robblee (unpublished) 18-station data set for juvenile pink shrimp in Florida Bay and observed and simulated salinity data for western Florida Bay and north-central Florida Bay. This will allow comparison of observed pink shrimp density and density-related output from the pink shrimp model. The model will then be used with the predicted paleosalinity regimes for Whipray Basin and Johnson Key to predict pre-drainage shrimp densities for Johnson Key and Whipray basins.

III. Methods

The data requirements for the shrimp model and the mechanism for using the various salinity runs were determined through discussions with pink shrimp modelers. It was arranged that the PI would provide salinity time series data (daily) for four scenarios. The Whipray Basin simulations had previously been generated. However, the Johnson Key models were not developed until recently and processing was necessary to prepare the Johnson Key models for use. Figure 1 shows the location of ENP Marine Monitoring Network (MMN) stations where observed data have been used for salinity modeling for this task.

Figure 1. Map showing the location of the ENP MMN salinity monitoring stations used for statistical salinity models for coupling with pink shrimp models.



The following data were sent to Joan Browder:

1. Whipray Basin observed salinity– MMN data collected by ENP beginning on April 6, 1989 and ending on Dec 31, 2000 (data still being collected).
2. Whipray Basin reconstructed salinity – Whipray Basin salinity reconstructed using long-term observed stage data, univariate linear regression models, and multivariate linear regression models, January 1, 1965 – December 31, 2000 (Marshall and Smith, 2007).
3. Whipray Basin paleo-based salinity – Paleoecology-based simulation of Whipray Basin salinity for January 1, 1965 through December 31, 2000 (Marshall, 2006; Marshall, 2007 draft).
4. Whipray Basin NSM 4.6.2-based salinity – Whipray Basin pre-development salinity regime produced from NSM 4.6.2 stage data using MLR salinity model, January 1, 1965 – December 31, 2000 (Marshall et al, 2004).
5. Johnson Key observed salinity– MMN data collected by ENP beginning on March 13, 1989 and ending on Dec 31, 2000 (data still being collected).
6. Johnson Key reconstructed salinity – Johnson Key salinity reconstructed using long-term observed stage data, univariate linear regression models,

and multivariate linear regression models, using methods presented in Marshall and Smith, 2007, January 1, 1965 – December 31, 2000.

7. Johnson Key paleosalinity salinity- Paleoecology-based simulation of Whipray Basin salinity for January 1, 1965 through December 31, 2000 (Marshall, 2006; Marshall, 2007 draft).
8. Johnson Key NSM 4.6.2 salinity – Johnson Key pre-development salinity regime produced for this task from NSM 4.6.2 stage data using MLR salinity model, similar to methods presented in Marshall et al, 2004, January 1, 1965 – December 31, 2000.

It is noted that the reconstructed, NSM 4.6.2-based, and paleo-based Whipray Basin salinity simulations were previously developed for other studies (Marshall and Smith, 2007; Marshall, 2005a; and Marshall, 2007 draft; respectively). The Johnson Key reconstructed and NSM 4.6.2-based salinity simulations were produced for this task. The Johnson Key paleo-based salinity was also previously produced (Marshall, 2007 draft). Details on the development of the salinity models used to make these runs can be found in the references.

Both Whipray Basin and Johnson Key multivariate linear regression models are functions of the stage at key stations in the freshwater marshes of the Everglades (CP and P33), wind vectors, Key West water level, and the salinity at Little Madeira Bay and Terrapin Bay. When the Whipray Basin and Johnson Key models are used for simulation, the salinity at Little Madeira Bay and Terrapin Bay must be simulated first using the MLR salinity models presented in Marshall et al, 2004.

Error statistics for the MLR salinity models used are presented below in Table 1. Plots showing a comparison of observed and simulated salinity data for the four models are presented as Figures 2 - 5.

Table 1. Error statistics for the MLR salinity models used for pink shrimp models.

station	mean sq error (mse), psu	root mse (rmse), psu	adj R-sq	mean error, psu	mean abs error, psu	max abs error, psu	Nash-Sutcliffe Efficiency
Little Madeira Bay	40.1	6.4	0.65	-0.66	5.1	22.6	-0.96
Terrapin Bay	32.6	5.7	0.75	-0.99	5.4	5.4	0.67
Whipray Basin	7.2	2.7	0.8	0.11	2.2	10.1	0.77
Johnson Key	0.05	2.68	0.55	0.05	2.2	9.72	0.55

Figure 2. Comparison of Observed and Simulated Data for the Little Madeira Bay Extended Period MLR Salinity Model

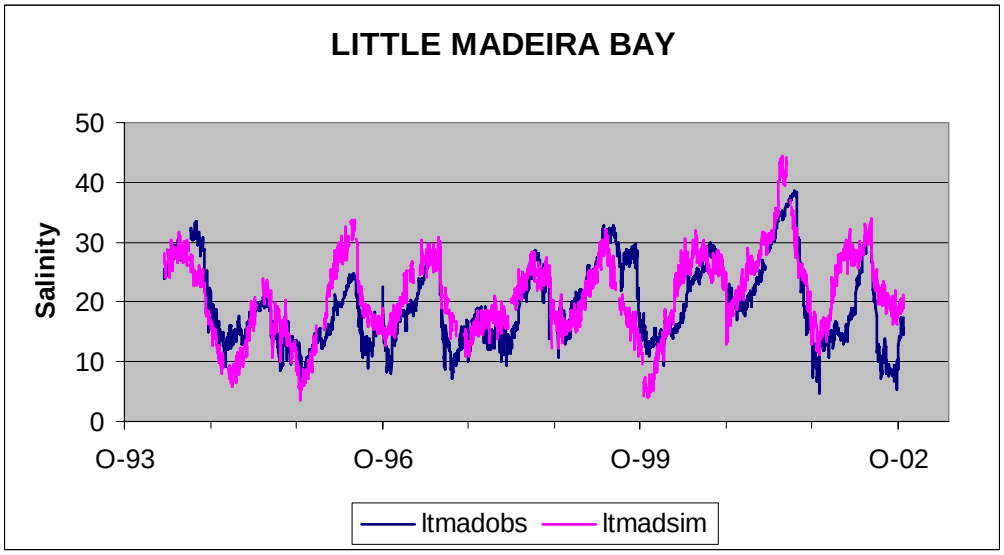


Figure 3. Comparison of Observed and Simulated Daily Data for the Terrapin Bay MLR Salinity Model.

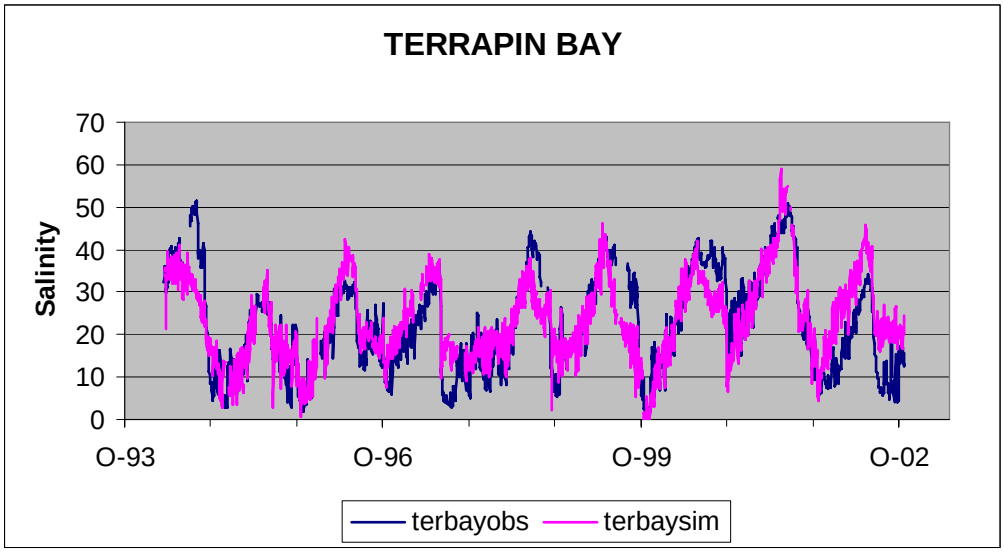


Figure 4. Comparison of Observed and Simulated Daily Data for the Whipray Basin MLR Salinity Model.

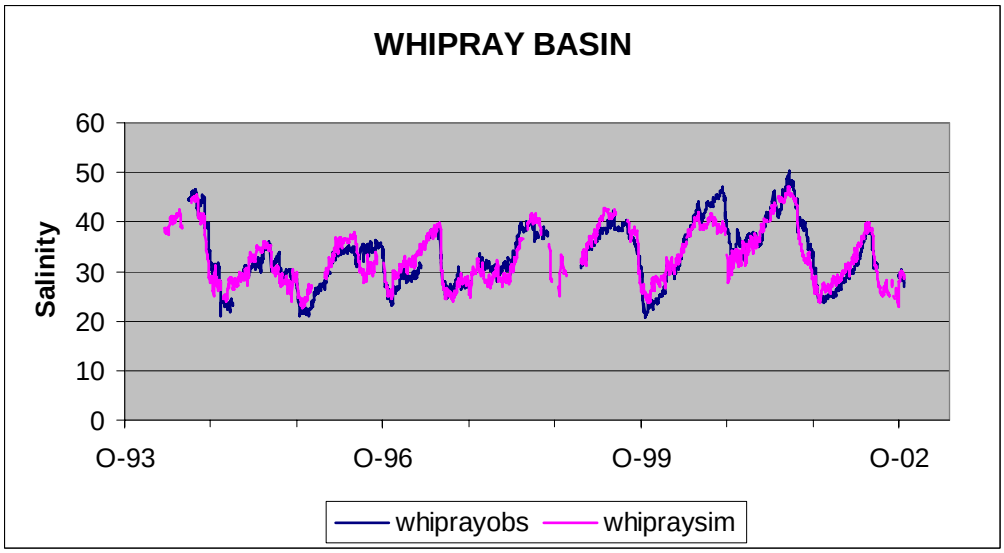
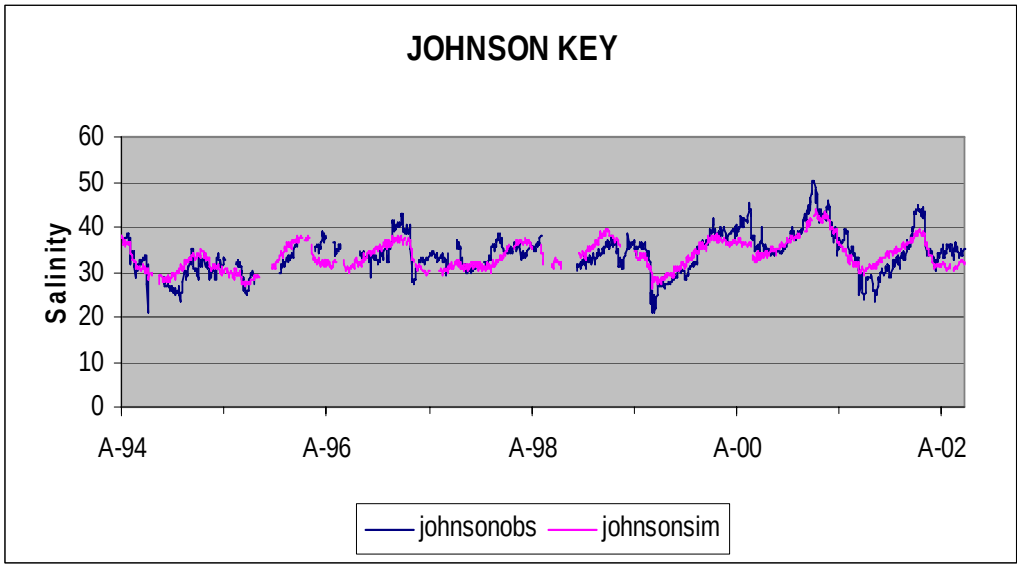


Figure 5. Comparison of Observed and Simulated Daily Data for Johnson Key MLR Salinity Model.



IV. Results

This task was a proof-of-concept exercise that focused on the coiupling of statistical salinity model output with pink shrimp models. Therefore, the results consist of output from the shrimp model with limited interpretation.

The data were sent by the Principal Investigator to Joan Browder (NOAA) and the pink shrimp model (Browder at al, 2002) was utilized for a number of preliminary runs, including the following:

1. Historical reconstruction salinity data (Marshall, 2007a),
2. Natural System Model-based (NSM 4.6.2) salinity from MLR model,
3. Paleo-based salinity from MLR model (Marshall, 2007b),
4. the average of reconstruction and NSM 4.6.2-based salinity (r+n/2), and
5. the average of NSM 4.6.2-based and paleo-based salinity (n+p/2).

These runs were intended to span the range of salinity values that were being examined at Whipray Basin and at Johnson Key.

Output produced by the pink shrimp model for Whipray Basin and Johnson Key include proportion of daily survival, average growth in mm/day, and potential harvest in tons kg/year. The products of the preliminary coupling of salinity produced by statistical models and pink shrimp models are represented by Figures 6 and 7 for Whipray Basin and Johnson Key.

Although present and reconstructed salinity regimes in the two basins differ, certain generalities apply to both basins regarding the simulated responses of pink shrimp to the various scenarios. As can be seen by the information in Figures 6 and 7, pink shrimp responses are lower (i.e., less favorable) under the paleo-based salinity scenario in the case of growth, survival, and potential harvests in Whipray Basin and in the case of survival and potential harvests in Johnson Key Basin. In contrast, growth rate was substantially higher for the paleo-based salinity regime compared to reconstitution salinity in Johnson Key Basin in these simulations. With the one exception, responses varied little among the reconstruction, NSM, and (r+n/2) scenarios. The (n+p/2) regime response was intermediate in responses between the paleo-based salinity regime and the other three scenarios.

Looking only at potential harvests which is an integrating parameter, in both Whipray Basin and Johnson Key the response to the paleo-based salinity scenario was higher than the response to the other scenarios in only two of the 36 years, 1971 and 1989, generally recognized as exceptionally “dry” years in the long-term south Florida rainfall records. The poorest responses were in the years 1966, 1968-1970, 1995-1997, and 1999, recognized as “wet” years in south Florida.

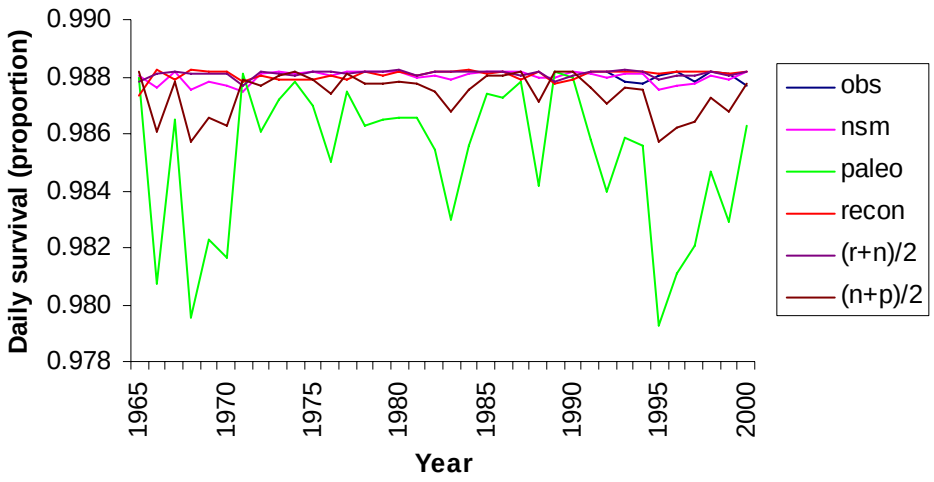
Differences in responses to the reconstruction, NSM, and (r+n)/2, were examined in more detail by calculating means and quartiles of the time series of responses.

NSM had the highest mean growth rates and (r+n)/2 had the highest mean survival rates and potential harvests. The quartiles showed their most striking differences between paleo-based salinity regime and the other scenarios. Differences among scenario responses increased inversely with quartile, causing the minimum to show the largest differences among scenarios.

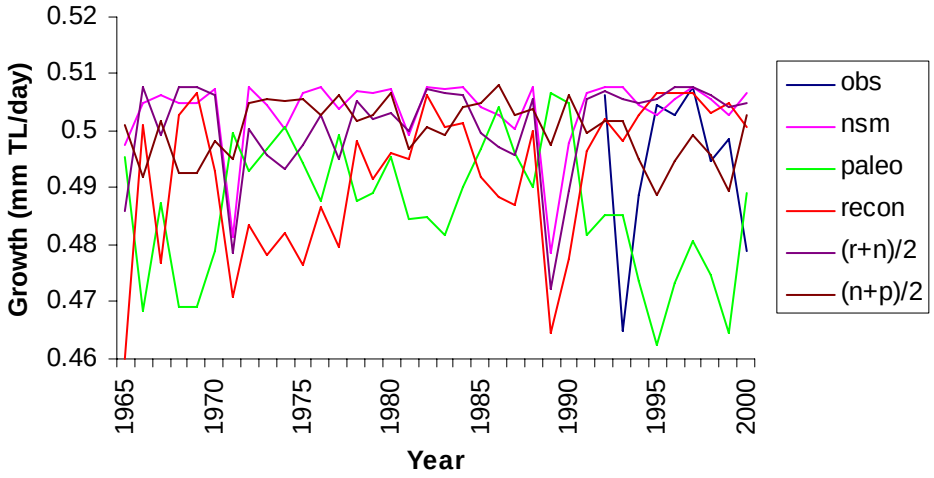
As another way of making comparisons, yearly differences between July-of-the-year responses were made (Table 2). An average value of the differences in each case was also computed (Table 2). In the table, cases where the counts do not sum to 36 indicate the presence of ties. According to annual averages, both NSM and (r+n)/2 outperform the reconstruction by a small margin in simulated growth rate, survival rate, and potential harvests. Based on the number of years, both NSM and (r+n)/2 outperform the reconstruction in growth rate and potential harvests in both basins. In survival rate, (r+n)/2 outperforms the reconstruction in both basins, and the reconstruction outperforms NSM in both basins.

Figure 6. Example output from pink shrimp model (Browder et al, 2002) for Whipray Basin produced from statistical salinity model simulations.

a. Daily Survival



b. Average Growth



c. Potential Harvest

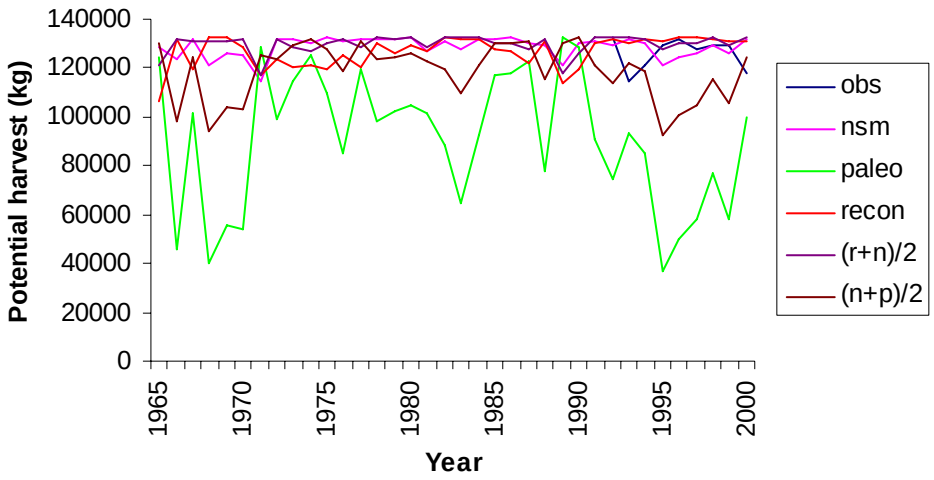
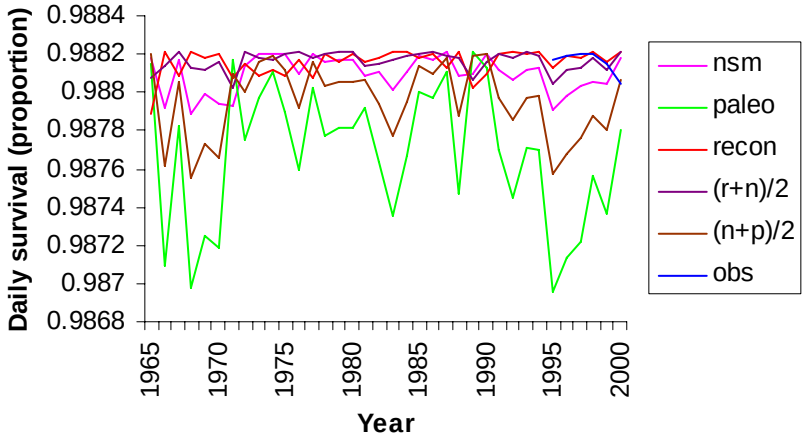
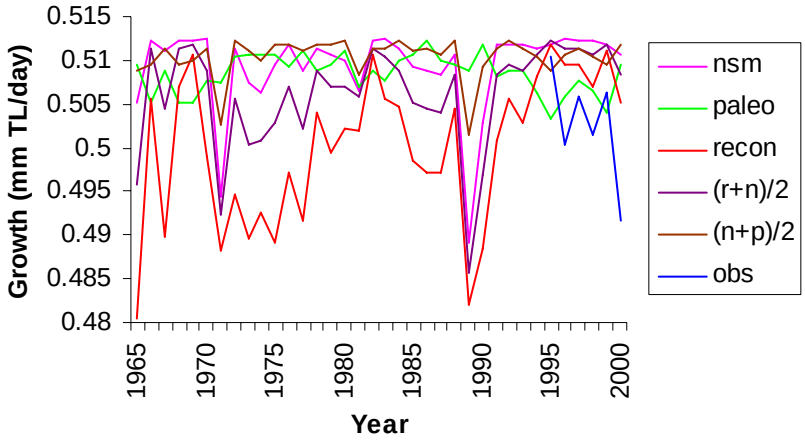


Figure 7. Example output from pink shrimp model (Browder et al, 2002) for Johnson Key produced from statistical salinity model simulations.

a. Daily Survival



b. Average Growth



c. Potential Harvest

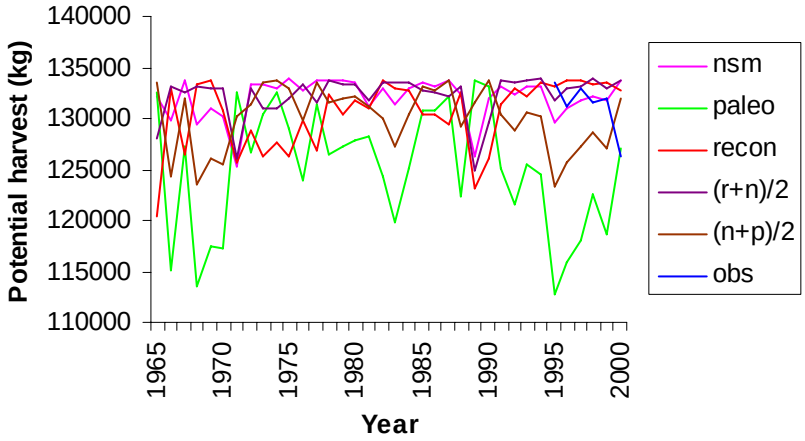


Table 2. Comparisons of simulated pink shrimp response levels of three parameters (potential harvest, daily growth rate, and daily survival rate) to the Reconstituted scenario in relation to NSM (top) and Intermediate-Reconstituted-and-NSM (bottom) scenarios in Johnson Key Basin and Whipray Basin across years for July cohorts Number of years is a count of years when the comparison was true. Averages are of positive values of the differences. Direction of the differences is indicated by the Comparison in column 1.

			Johnson Key Basin **NSM vs. Recon**		Whipray Basin **NSM vs. Recon**	
Comparison	Parameter	Units	Number of years	Ave. of positive differences	Years	Ave. of positive differences
NSM > Recon	Potential harvest	kg	21	3938	19	7526
		mm				
	Daily growth rate	TL/d	35	0.009544	32	0.012873
	Daily survival rate	percent	11	0.009990	14	0.022147
Recon > NSM	Potential harvest	kg	15	1726	17	4514
		mm				
	Daily growth rate	TL/d	0		4	0.002105
	Daily survival rate	percent	25	0.012889	22	0.026841
			(R+N)/2 vs. Recon		**(R+N)/2 vs. Recon**	
	Parameter	Units	Number of years	Ave. of positive differences	Years	Ave. of positive differences
(R+N)/2 > Recon	Potential harvest	kg	29	2389	29	4153
		mm				
	Daily growth rate	TL/d	36	0.006084	34	0.008935
	Daily survival rate	percent	19	0.005576	20	0.014139
Recon > (R+N)/2	Potential harvest	kg	7	646	7	1781
		mm				
	Daily growth rate	TL/d	0		2	0.000917
	Daily survival rate	percent	17	0.004740	16	0.008497

V. Discussion and Recommendations

The time series outputs from the statistical salinity models for three salinity scenarios at Whipray Basin and Johnson Key were used by the pink shrimp modelers to make five separate runs of the pink shrimp model. Pink shrimp model output presented in plots in the results section shows that the MLR salinity models can be coupled with the pink shrimp growth and production model to produce data on shrimp growth dynamics. Though the output is still being interpreted, the initial output produced these preliminary findings:

- The salinity data used for input to the pink shrimp model must not include missing values. The techniques identified in Task 8 for filling data gaps were useful for this activity.
- Historical reconstructions are useful for pink shrimp model input because they can be compared to the full period (1965 – 2000) for NSM 4.6.2 and paleo-based salinity runs.
- Historical reconstructions using modeled data only instead of observed data plus modeled data to fill gaps may be more useful for pink shrimp model coupling because the modeled data allow for control of one additional variable (salinity), allowing for model vs model comparisons. This does not downplay the importance of evaluations made with observed data.

Application of the pink shrimp model to Johnson Key Basin and Whipray Basin under different scenarios of salinity provided a view of the sensitivity of the pink shrimp model to variation in salinity regime. Output parameters relating to growth, survival, and potential harvests of the cohort of one month of the year (July of each year, 1965-2000) were compared among simulations (Figures 6. and 7.). Responses to the relatively short observed time series (1992-2000 in Whipray and 1995-2000 in Johnson) do not provide sufficient within-scenario salinity variation for robust comparison of the scenarios. are shown on the graphs with the other scenarios. Comparison of the other scenarios to the reconstruction scenario is more appropriate than comparison to the observed because, in the latter case, both salinity model effects and simulated pink shrimp responses are incorporated in the differences. The 36-year time series used in the simulations provides broader scope for responses.

The model results suggest that pink shrimp are favored by salinities on their nursery grounds that are neither extremely high nor extremely low. The Paleo-based salinity regime favors shrimp only in extremely dry years and provides salinities that are too low for good shrimp production in most other years. NSM performs better than the paleo-based salinity regime in wet years and offers some improvement over the reconstruction time series in dry years (although not in 1971). The model suggests that the scenario with salinities intermediate between NSM and the reconstruction consistently provides the most favorable salinity regime of those reviewed.

According to the results of this preliminary analysis (pink shrimp model only), if an objective were to optimize pink shrimp production in these two areas of Florida Bay, shifting the present salinity regime to create lower salinities across the board may not be as effective as reducing the frequency of the high salinities (>35 psu) without increasing the frequency of low salinities (<25 psu). While creating paleo-based salinity regimes in these two areas did not appear, in the simulations, to improve their nursery function for pink shrimp, more favorable salinities for pink shrimp than exist at present might become established in other parts of Florida Bay as a result of creating paleo-salinities in these two areas.

Because of this encouraging results, it is recommended that further coordination be continued between the PI and the pink shrimp modelers to finalize the interpretation of these runs and to increase the spatial extent of the pink shrimp simulations. However, this was a proof-of-concept exercise, so the success of the coupling of the models satisfies the contract requirements for this task.

This coupling exercise has shown that the salinity output from statistical models can be readily used as input to ecological models. Examples of ecological models other than the pink shrimp model that could potentially use statistical model-produced salinity time series include, but are not limited to, the following:

- Madden et al, 2006 (draft) seagrass model
- Trexler et al, 2003 fish models
- Lorenz and Serafy, 2006 fish assemblage models
- Johnson et al, 2005 statistical fish and crustacean models.

There are other efforts underway to develop ecological models that can use the time series salinity data produced by the statistical models. A comprehensive literature search was not performed, though it is known that other ecological modeling projects are under development.

It is also recommended that further cooperative efforts using statistical salinity model output with ecological models be commissioned so that the full benefit of the statistical models developed by this CESI project can be realized.

VI. VI. References

Browder, J. A., Z. Zein-Eldin, M.M. Criales, M. B. Robblee, S. Wong, T. L. Jackson and D. Johnson. 2002. Dynamics of pink shrimp (*Farfantepenaeus duorarum*) recruitment potential in relation to salinity and temperature in Florida Bay. Estuaries 25 (6B):1355-1371.

Browder, J. A., D. R. Johnson, R. Bennett, F. Marshall, and J. Wang. 2005. Application of a simulation model of pink shrimp growth and survival. (Abstract). Florida Bay Science Conference.

Johnson, D.R., J. A. Browder, and M. B. Robblee. 2005. Statistical models of FL Bay fishes and crustaceans to evaluate minimum flow levels in FL Bay. Contribution No. PRD 04/05-06. Protected Resources and Biodiversity Division, SEFSC. 474 pp.

Lorenz, J. and J, Serafy. 2006. Subtroprical wetland fish assemblages and changing salinity regimes: Implications for everglades restoration. In: J.C. Trexler, E.E. Gaiser & D.L. Childers (eds), *Interaction of Hydrology and Nutrients in Controlling Ecosystem Function in Oligotrophic Coastal Environments of South Florida*. Hydrobiologia (2006) 569:401–422.

Madden, C.J. and A.A. McDonald. 2006, draft. An Ecological Model of the Florida Bay Seagrass Community – Model Documentation. South Florida Water Management District, West Palm Beach, Florida.

Marshall, F.E. and D. Smith. 2007 (draft). Historical reconstructions of salinity in Northeast Florida Bay using multivariate linear regression models. Task report for Critical Ecosystems Studies Initiative project CA H5284-05-0006 for Everglades National Park. Cetacean Logic Foundation, Inc., New Smyrna Beach, Florida.

Marshall, F. E. 2007 (draft). Draft Historical Hydrology Analysis Report, Task 1: Estimating the Pre-drainage Everglades Hydrology and Florida Bay Salinity Using Paleoecology and Regression Models. Environmental Consulting & Technology, Inc., New Smyrna Beach, Florida.

Marshall, F. E. 2006. Freshwater Flow Estimates for the Florida Bay and Florida Keys Feasibility Study. Task report for Critical Ecosystems Studies Initiative project CA H5284-05-0006 for Everglades National Park. Cetacean Logic Foundation, Inc., New Smyrna Beach, Florida.

Marshall III, F. E. 2005a. RECOVER Southern Estuaries performance measures: identification of hydrology-salinity relationships for coastal estuaries and analysis of interim CERP update scenarios. Environmental Consulting & Technology, Inc., New Smyrna Beach, Florida.

Marshall III, F.E. 2005b. Using Statistical Models to Simulate Salinity Variability in Estuaries. *In* Estuarine Indicators, S. A. Bortone, Editor. CRC Press. Boca Raton, Florida. Pp 33-52.

Marshall III, F. E.; D. Smith; and D. Nickerson. 2004. Using Statistical Models to Simulate Salinity Variation and Other Physical Parameters in North Florida Bay. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. 36 pp.

Marshall III, F. E.; D. Smith; and D. Nickerson. 2003. Salinity Simulation Models for North Florida Bay Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. 41 pp.

Marshall III, F. E. 2003. IOP Salinity Analysis Using Statistical Models. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. 35 pp.

RECOVER. 2004, draft. Recommendation for Interim Goals and Interim Targets for the Comprehensive Everglades Restoration Plan Indicators and Prediction Methods. US Army Corps of Engineers, Jacksonville District. Jacksonville, Florida.

Trexler, J. C., W.F. Loftus, and J. H. Chick. 2003. Setting and Monitoring Restoration Goals in the Absence of Historical Data: The Case of Fishes in the Florida Everglades. In: D. Bush and J. Trexler (ed), *Monitoring Ecosystems – Interdisciplinary Approaches for Evaluating Ecological Initiatives*. Island Press, Washington, D.C.

FINAL TASK REPORT

TASK 6 - COORDINATE WITH THE IMC TO INCORPORATE THE NEW MLR SALINITY MODELS INTO THE POST-PROCESSING TOOLS

Cooperative Agreement Number CA H5284-05-0006
Between
The United States Department of the Interior National Park Service
Everglades National Park
And
Cetacean Logic Foundation, Inc.

Frank E. Marshall, III
Principal Investigator

DeWitt Smith
Project Coordinator

April 30, 2008

CESI Project No. CA H5284-05-0006 for Everglades National Park

FINAL TASK REPORT

TASK 6 - COORDINATE WITH THE IMC TO INCORPORATE THE NEW MLR SALINITY MODELS INTO THE POST-PROCESSING TOOLS

I. Introduction

A. General

This report describes the activities that were completed for a task in the current Critical Ecosystem Studies Initiative (CESI) project that was originally initiated by Cetacean Logic Foundation, Inc. (CLF) for Everglades National Park (ENP) in 2002. The goal of this on-going CLF research is the characterization and simulation of the salinity regime in Florida Bay and the estuaries along southwest coast of the Gulf of Mexico (Gulf) within the Park. It is intended that this research yield information regarding the link between the downstream salinity in the estuaries of ENP and the upstream freshwater hydrology of the Everglades.

The subject of this report are the steps that were taken by the Principal Investigator to assist the Interagency Modeling Committee to incorporate the most recently developed multivariate linear regression (MLR) salinity models (Task 3 of this CESI project) into the post-processing tools that IMC uses to evaluate Comprehensive Everglades Restoration Plan (CERP) water management alternatives. This task report describes the coordination between the groups and the effort taken to provide the new MLR salinity models to the IMC.

B. Task Objectives and Evaluation Methods

The objective of this task is presented below as it appears in the modified contract for this CESI project approved by ENP:

> "The IMC recently developed a post-processing tool for the Southern Estuaries Sub-team. The PI will coordinate with the IMC to include the new models in the post-processing tool.
>
> The deliverable for is a task report that describes the activities involved with coordinating with the IMC."

II. Background

Multivariate linear regression (MLR) salinity models are currently being used by the Southern Estuaries Sub-team of RECOVER to evaluate Comprehensive Everglades Restoration Plan (CERP) alternatives. Nineteen models developed as part of previous CESI tasks (Marshall, 2003; Marshal et al, 2003; Marshall et

al, 2004; Marshall, 2005b) or as part of a task for RECOVER (Marshall, 2005a) were entered into a computer code routine by the Interagency Modeling Center (IMC). When it is desired that a water management alternative be evaluated for performance, the IMC uses the stage output from the South Florida Water Management Model (SFWMM, or 2X2 Model) with observed wind and Key West water level data for the period 1965-2000 as input to the coded MLR salinity models to produce a daily time series simulation for the stations in Florida Bay, Whitewater Bay, the Shark River discharge estuaries, and Barnes Sound / Manatee Bay.

It is the objective of this project to make the new models that were developed as a part of this CESI project (Task 3) available to the IMC for incorporation into the salinity simulation routine if it is so desired by the IMC.

III. Methods

The new models that were developed in Task 3 were completed and reviewed on March 3, 2007 and a draft task report was submitted to the CESI Project Coordinator. An e-mail communication was sent to Jose Otero on June 16, 2007 notifying him that the models were ready to be incorporated and that the PI would assist as needed to insure that the models were coded appropriately. On July 5, 2007, a return communication indicated that the use of the new models should be coordinated through RECOVER.

Once RECOVER coordination through Andrew Gottlieb of South Florida Water Management District was completed, the PI met with Jose Otero and Michael Martin to begin placing the models in the IMC computer code in late March, 2008. The PI transmitted all new models from Task No. 3 of this CESI project to Michael Martin and assisted him with review of the final product. The QA/QC review was completed on April 30, 2008.

IV. Discussion, Findings, and Recommendations

This task was a perfunctory task that was intended to help the agencies that rely on the MLR salinity models to analyze CERP alternatives as part of the performance measure evaluations. All models developed for Task No. 3 of this CESI project have been imported by the IMC.

V. References

Marshall III, F. E. 2005a. RECOVER Southern Estuaries performance measures: identification of hydrology-salinity relationships for coastal estuaries and analysis of interim CERP update scenarios. Environmental Consulting & Technology, Inc., New Smyrna Beach, Florida.

Marshall III, F.E. 2005. Using Statistical Models to Simulate Salinity Variability in Estuaries. *In* Estuarine Indicators, S. A. Bortone, Editor. CRC Press. Boca Raton, Florida. Pp 33-52.

Marshall III, F. E.; D. Smith; and D. Nickerson. 2004. Using Statistical Models to Simulate Salinity Variation and Other Physical Parameters in North Florida Bay. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. 36 pp.

Marshall III, F. E.; D. Smith; and D. Nickerson. 2003. Salinity Simulation Models for North Florida Bay Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. 41 pp.

Marshall III, F. E. 2003. IOP Salinity Analysis Using Statistical Models. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. 35 pp.

VI. Appendix

FINAL TASK REPORT

TASK 7 - SIMULATING AND FORECASTING SALINITY IN FLORIDA BAY: A REVIEW OF MODELS

Cooperative Agreement Number CA H5284-05-0006
Between
The United States Department of the Interior National Park Service
Everglades National Park
And
Cetacean Logic Foundation, Inc.

Frank E. Marshall, III
Principal Investigator

DeWitt Smith
Project Coordinator

William Nuttle
Contributing Investigator

April 15, 2006
Updated August 6, 2008

Simulating and Forecasting Salinity in Florida Bay: A Review of Models

EXECUTIVE SUMMARY

Salinity is a fundamental and key characteristic of the physical conditions of estuarine and coastal ecosystems. Salinity affects water quality, the make-up and spatial distribution of vegetative communities, and the life history of most animal species in these ecosystems. Simulations and forecasts of salinity are an important tool in the assessment of ecological resources in the Everglades, Florida Bay, and the estuaries on the Gulf of Mexico (CROGEE, 2002). Water managers use forecasts to evaluate the expected benefits and impacts of ecosystem restoration activities. Ecosystem restoration involves aspects of adaptive management (NRC 2004), uncertainty analysis (CERP 2002), and risk assessment (Thom et al. 2004), and these all rely on the application of predictive models.

This report reviews models for which information is currently available on a broad basis (June 2006) for simulating and forecasting salinity in Florida Bay, Whitewater Bay, and the Gulf coast estuaries. For the salinity evaluations that have taken place thus far, there have been two general approaches to constructing such models. The first is empirical and relies on accurately describing observed salinity variations and correlative relationships. The second is mechanistic-based and relies on accurately accounting for the physical processes that drive changes in salinity. In both approaches the accuracy of the forecasts is limited by the data available to describe patterns of salinity variation and the driving processes.

Various statistical techniques can be employed in the empirical approach, the simplest being descriptive analysis. Both regression and time series modeling techniques have been applied to derive models for Florida Bay and Gulf coast salinity. Regression models exploit linear relationships in records of driving processes and systems response. Time series models utilize the serial correlation that is present in many hydrologic parameters. The statistical models that have been developed thus far for Florida Bay and the Gulf coast estuaries are based on a coastal aquifer conceptual model and have been used successfully for evaluating water management alternatives and for performance measure development.

Mechanistic salinity models for south Florida estuaries include both mass-balance models and more complex hydrodynamic models. Mass balance models of salinity, in their discretized numerical form, are similar in form to autoregressive time series models. Mass balance models account for the inputs and outputs of water from basins delineated by geomorphologic features. Mass

balance models have been used for ecological evaluations and for minimum flows and levels modeling.

Hydrodynamic models have been developed for both Everglades hydrology and the salinity in the downstream estuary. Hydrodynamic models are based on the solution of simultaneous differential equations of continuity and hydrodynamics (momentum) in one, two, or three dimensions, and can be used for both surface and groundwater applications. Hydrodynamic models have been used for modeling the freshwater portion of the Everglades / Florida Bay hydrologic system for about the past decade, and are in the process of being updated with better data and techniques. Only recently have hydrodynamic models been available that are capable of adequately simulating the salinity regime in south Florida Bay and the mangrove / salinity transition zone. Work is currently underway on the Florida Bay hydrodynamic models, while work on hydrodynamic models for the transition zone of the Gulf coast estuaries is still in preliminary stages.

A summary table presenting general model information and a summary evaluation table is included with this Executive Summary. In the evaluation table the Florida Bay Science Program model evaluation factors have been used and a score of 1 to 5 has been assigned to each model for each factor, with 1 being poor and 5 being excellent in application.

Summary of Salinity Models and Supporting Hydrologic Models Currently in Use For Simulating Florida Bay and Southwest Gulf Coast Salinity (Table 6 in report).

Model Name	Model Type	Simulated Parameters	Spatial Domain	Grid Size	Simulation temporal domain
SFWMM[1]	Freshwater Hydrology	Stage, Flow	Everglades	3.2km X 3.2km	1965-2000, daily
PHAST[2]	Wetland Basin	Flow	Lower Everglades and Mangrove Zone	regional	1965-2000, monthly
MLR[3]	Statistical	Salinity	Florida Bay, Whitewater Bay, southwest Gulf coast, Manatee Bay, Barnes Sound	N/A	1965-2000, daily
Four Box[4]	Mass Balance	Salinity	Florida Bay	regional	1993-1998, monthly
FATHOM[5]	Mass Balance	Salinity	Florida Bay, Manatee Bay, Barnes Sound	open-water basins	1965-2000, monthly
EFDC[6]	3-D Hydrodynamic	Salinity	Florida Bay, Whitewater Bay, southwest Gulf coast, Manatee Bay, Barnes Sound	variable	1965-2000, daily
SICS/TIME[7]	2D/3D Coupled surface and groundwater	Stage, Flow, Salinity	Florida Bay (SICS), southwest Gulf coast (TIME)	0.3km X 0.3km (SICS, 0.5km X 0.5km (TIME)	1996-2000, daily
SoFLA-HYCOM[8]	3-D Hydrodynamic ocean circulation model	Flow magnitude and direction	Gulf of Mexico, Florida Straits	6-7km X 6-7km	?

[1] http://www.sfwmd.gov/org/pld/hsm/models/sfwmm/index.html
[2] Nuttle and Teed 2002, Nuttle 2004
[3] Marshall, 2005
[4] Nuttle et al. (2000)
[5] Cosby et al. 1999, Nuttle et al. 2000, Cosby et al 2004
[6] Hamrick and Moustafa, 2003
[7] Swain, et al 2004 (SICS), Langevin, et al 2002 (TIME)
[8] Kourafalou, 2005

Summary evaluation of Florida Bay salinity and hydrology models using the Florida Bay Science Program evaluation factors (PMC 2000). Models with asterisk (*) are freshwater hydrology only models. Score is from 1=lowest to 5=highest (Table 7 in report).

Model	Portability	Validity	Fidelity	Focus	Ease of Use
SFWMM*	2	5	3	4	3
PHAST*	3	4	3	3	5
MLR	5	5	5	5	5
Four Box*	3	4	4	3	5
FATHOM	3	5	4	5	4
EFDC	2	5	3	5	3
SICS/TIME	2	5	4	5	3
SoFLA-HYCOM	2	3	?	3	3

CONTENTS

EXECUTIVE SUMMARY I

CONTENTS 1

1 INTRODUCTION 1
1.1 General 1
1.2 Background 2
1.2.1 Approaches to Simulation and Forecasting 3
1.2.2 Forecast Uncertainty 4
1.3 Previous Report 5
1.4 Objectives and Scope of this Study 6

2 SALINITY PATTERNS AND PROCESSES 8
2.1 Salinity Variation 8
2.2 Driving Processes 9
2.2.1 Net Freshwater Supply 9
2.2.2 Exchange Processes 10
2.2.3 Boundary Salinity 11
2.3 Available Salinity Data 11
2.4 Estimating Inflow from the Everglades to Florida Bay 13

3 STATISTICAL MODELS 16
3.1 Regression Models 16
3.2 Time Series and Frequency Domain Models 18

4 MECHANISTIC MODELS 22
4.1 Mass Balance (Box) Models 22
4.1.1 Four Box Model of Florida Bay 22
4.1.2 PHAST 24
4.2 Hydrodynamic Models 28
4.2.1 Initial Hydrodynamic Modeling Efforts 28
4.2.2 EFDC 28
4.2.3 SICS / TIME 33

5 SUMMARY AND DISCUSSION 38

6 FINDINGS 41

7 REFERENCES 45

FINAL TASK REPORT

TASK 7 - SIMULATING AND FORECASTING SALINITY IN FLORIDA BAY: A REVIEW OF MODELS

1 INTRODUCTION

1.1 General

Ecological forecast models play an essential role in efforts to restore and preserve natural resources. This role is analogous to the role that hydrologic forecast models have played in the management of water resources for human benefit (Lettenmaier and Wood 1993). Ecosystem restoration involves aspects of adaptive management (NRC 2004), uncertainty analysis (CERP 2002), and risk assessment (Thom et al. 2004), and these all rely on the application of predictive models. In south Florida, the Everglades and estuarine ecosystems in Everglades National Park have been altered by water supply and flood protection for agricultural and urban activities (CROGEE 2002). The restoration effort for these ecosystems is currently centered on the activities of the Comprehensive Everglades Restoration Plan (CERP).

For CERP, ecologists and water managers use salinity forecasts as one tool to evaluate the expected benefits and impact of restoration activities. These benefits and impacts to coastal ecosystems are reflected in potential future changes in wetland communities, estuarine water quality, and coastal fisheries that are expected for water management activities and for alternative management scenarios. Forecasts provide managers with quantitative information needed for evaluation of alternative actions under consideration and to choose the course of action that best meets objectives.

The study area for this review of models includes the freshwater marshes and mangrove eco-tone areas of the Everglades, the estuarine and near-marine basins of Florida Bay, the estuarine areas of Whitewater Bay, and the estuaries the discharge into the southeastern-most portion of the Gulf off Mexico (Gulf coast). The hydrologic features in the upstream Everglades that are important to salinity modeling include Shark River Slough, Taylor Slough, and the C-111 Canal system.

The ability to forecast how Everglades restoration will affect the ecology of Florida Bay, Whitewater Bay, and the Gulf coast of south Florida depends on first being able to forecast how changes in regional water management alter the bay's

salinity. Changes in salinity reflect changes in the amount or timing of the net supply of freshwater to an estuary, i.e. the sum of rainfall plus inflow minus evaporation, hydrodynamics and mixing, and exchange with the ocean. Salinity is a key characteristic of physical conditions (including water quality) in estuarine and coastal ecosystems; it affects the composition and spatial distribution of vegetative communities and life history of most animal species. Because salt is a conservative tracer, changes in salinity signal possible changes in the concentrations of other substances, such as nutrients and contaminants that enter estuarine and coastal waters through the inflow of freshwater or mixing with the coastal ocean.

This report reviews models for which information was currently available (June 2006) for forecasting salinity in Florida Bay. This constitutes part of the work being performed by the Cetacean Logic Foundation, Inc. for Everglades National Park (ENP) with support from the Critical Ecosystem Science Initiative (CESI) program. The purpose of this study is to update information in a similar report compiled for Everglades National Park by The Cadmus Group (Nuttle 2002). The present work expands the coverage of the earlier report by incorporating the recent improvements in hydrology and salinity modeling including statistical, mass balance, and hydrodynamic models.

1.2 Background

In general, the formulation and application of forecasting models serve three roles. First, the formulation and development of predictive models helps to confirm a common understanding of the system and its behavior in response to changes in driving processes. Second, the predictive model functions as one of the primary mechanisms for investigating possible future structure and behavior of the system that can may result from proposed restoration activities. Finally, predictive models are used to understand uncertainties about the present and future state of the system and the variation in driving processes, and translate these into corresponding uncertainties of meeting restoration goals.

Forecasting ability increases with improved scientific understanding through the synthesis of research results. Therefore, formulation and refinement of predictive models serves an essential function in the development of knowledge through research and in the application of that knowledge toward the practical goals of ecosystem restoration. This is the motivation for building predictive salinity models for the southern Everglades and Florida Bay region.

Recurrent patterns in the data, such as the annual cycle of wet and dry seasons, are predictive in their own right in the mode of a null model. The underlying assertion of a null model is that the mechanisms driving the phenomenon will continue unchanged. Models used in restoration planning must go beyond the

description of a null model, if only to test the proposition that the null model is or is not the best model for describing the observed data.

1.2.1 Approaches to Simulation and Forecasting

Making accurate forecasts of salinity in Florida Bay, the mangrove transition zone of the Everglades, Whitewater Bay, and the Gulf coast estuaries depends on knowledge of patterns of salinity variation in the past and of the underlying driving processes that produced them. Forecasts derive from the driving processes and a representation of the relationship between these processes and salinity in the bay. For Everglades restoration, driving processes include water management alternatives that affect freshwater inflow to the estuaries from the Everglades, tied to different proposed management strategies.

There are two general approaches currently employed to construct salinity models. The first is empirical and relies on accurately describing observed salinity variations. Analysis of the available data identifies basic patterns that characterize the phenomenon of interest. The second is mechanistically-based and relies on accurately portraying the processes that drive changes in salinity. Typically, a numerical model describes the physical relationship between driving processes and salinity. In both approaches the accuracy of the forecasts is limited by the data available to describe patterns of salinity variation and the driving processes.

Various statistical techniques, including descriptive analysis and correlation, are employed in the empirical approach. These techniques help in understanding the relationship between driving processes and resulting salinity variation and can be used in deriving a mathematical description embodied in a linear combination model. Correlation does not necessarily establish a causal link between characteristics of the ecosystem and the driving processes that incorporate the effects of human activities. However, descriptive analysis and correlation are the foundation for models capable of reproducing patterns of variation. Descriptive analysis also serves to diagnose bias and other problems related to the methods of observation and measurement. Patterns identified through descriptive analysis and correlations provide clues to the underlying mechanisms by their proximity in time and space.

Both regression and time series modeling techniques have been applied to models for Florida Bay salinity. Regression models exploit linear relationships in records of driving processes and systems response. A number of statistical modeling tools are available, ranging from simple linear regression to more complicated analytical techniques such as multivariate regression, linear transfer function models, and frequency domain models. Time series models utilize the distribution of variation with time and serial correlation to model system behavior. By nature, useful time series models require enough data such that the variation over time can be adequately analyzed statistically. Classical time series

modeling begins by allocating the variation in a set of data ordered by time into different components, such as mean, trend, seasonal, etc. Times series models also include an explicit representation for irreducible error represented ideally as uncorrelated white noise error term.

The mechanistic approach relies on knowledge of the physical processes that influence estuarine salinity. The structure of mechanistic models reflects this understanding. Explicit mathematical representation of cause and effect based on general physical principles means that a mechanistic model can predict the behavior of the system beyond the range that has been observed. For this approach, there are various models that exhibit different levels of complexity depending on the detail employed in the numerical description of the processes at work.

Mechanistic models have only been developed for Florida Bay and the southern Everglades mangrove zone. Mechanistic models include both relatively simple mass-balance models and more complex hydrodynamic models. Mass balance models of salinity, in their discretized numerical form, are similar in form to autoregressive time series models. Mass balance models ignore momentum effects which are negligible at time steps greater than daily. Complex mechanistic models are based on the solution of simultaneous differential equations of continuity and hydrodynamics (momentum). A hydrodynamic forecast model is used where additional temporal and spatial detail or coverage are required for forecasts.

1.2.2 Forecast Uncertainty

Uncertainty in salinity forecasts falls into the category of “Knowledge Uncertainty” (NRC, 2000a). Knowledge uncertainty encompasses sources of uncertainty from imperfect knowledge of processes, model structure, model parameter values and data used as input in generating the forecasts. These sources of uncertainty are often not independent of each other. Uncertainties in the data can be derived in part from the mismatch between the temporal and spatial scales represented by the model and the scales on which data are collected. And finally, the uncertainty in the data contributes to the uncertainty in the optimally selected model parameters. All sources of uncertainty must be considered when evaluating alternative approaches or making improvements to forecasting.

Uncertainty in forecasts can be characterized by various statistics calculated from the differences between measured and forecast values of salinity, i.e. the set of residuals (*R*). For this study, five error statistics are reported:

- average error (avg e),
- root mean squared error (rmse),
- average absolute error (abs e),
- coefficient of determination (r^2), and
- model efficiency (eff).

The average error, the average of *R*, measures bias between simulated and observed values; a mean error of zero means no bias. Even if the average error is zero there can still be significant differences between simulated and observed values; these differences may simply cancel out in the calculation of the average error.

The root mean squared error and the average absolute error are measures of the deviation between simulated and observed values, reported in the units of the simulated variable. The root mean squared error (RMSE) is calculated as the square root of the mean of the squared residuals (MSE). The average absolute error is calculated as the mean of the absolute values of the *R* values. These measures better reflect the expected magnitude of the difference between calculated and measured salinity at a particular location and time.

Model efficiency and the coefficient of determination, R-squared, or R^2 are similar. The coefficient of determination measures the fraction of the variance in the observations that can be explained by a linear transformation of the simulated salinity values; therefore it is a measure of the correlation between the simulated and observed values. In contrast, model efficiency is calculated from the mean square error normalized by the variance of the observed salinity:

$$\mathbf{\mathit{eff} = 100^* \ (1 - MSE / Var(obs))} \qquad 1$$

where MSE is the mean of the squared residual errors and Var(obs) is the variance of the observed salinity data.

Model efficiency, also known as the Nash-Sutcliffe efficiency (c.f. Nash and Sutcliffe 1970, Weglarczyk 1998), can be interpreted broadly as the percentage of the variance in the data that is accounted for directly by the model. A model efficiency of zero indicates that the model accounts for no more of the variation than does the mean of the data. An efficiency of 100 indicates that the model accounts for all of the variation in the data. However, model efficiency can take on negative values if, for example, the model produces a biased estimate of the data. In this case, the mean of the data offers a better forecast than the model.

1.3 Previous Report

The salinity modeling status report by Nuttle (2002) reviewed and evaluated work prior to 2002 that could be applied to forecast salinity in the coastal mangroves of Everglades National Park and in Florida Bay. The report focused on approaches for formulating models needed to support the development and application of ecological performance measures. The goal was to identify an approach for linking coastal salinity prediction to changes in Everglades hydrology that could be implemented quickly and so satisfy the immediate need for predictive tools in

planning. Different approaches were evaluated based on both predictive ability and practical considerations related to needs of the multi-agency planning process for ecosystem restoration in south Florida.

Accordingly, Nuttle (2002) evaluated the alternative approaches to forecasting salinity in Florida Bay based on the following set of practical requirements drawn from experience in the Florida Bay Science Program (PMC 2000):

- ***Portability*** – The model chosen should be widely available for evaluation and application.
- ***Validity*** – The predictive capability of the model must be generally known and accepted.
- ***Fidelity*** – The model must be consistent with understood mechanisms of cause and effect within the limitations of the underlying approach to prediction.
- ***Focus*** – Model predictions must relate directly to the ecosystem attributes defined as performance measures.
- ***Ease of use*** – Model results must be able to be obtained quickly within the typically short time period allotted for analysis of alternatives within the planning process.

The Nuttle (2002) report recommended adopting the mass balance modeling approach, and this recommendation led to the development of the aggregated wetland basin hydrology and estuarine basin salinity model (PHAST) for ENP, and more recently used as a planning tool for the Biscayne Bay Coastal Wetlands Project (Nuttle 2005).

1.4 Objectives and Scope of this Study

Resource managers need reliable salinity forecast models to use in evaluating the benefits of alternative project designs for water management through the CERP program. As planning progresses and the understanding of the system matures, modeling activities will focus more on assuring the predictive capability of the model. For example, the future activities of the SFWMD Interagency Modeling Center will extend to reviewing modeling needs and advising project management teams on the application of models used for planning activities of individual CERP projects (CERP 2004).

This CESI task report provides general information about the options currently available from models that can be applied at sub-regional levels. The models that will be described in this task report will have the following characteristics:

- The primary models that will be reviewed are salinity models; Everglades freshwater hydrology models are included to the extent they have been utilized with the salinity models being described;
- Status of modeling efforts will be reported as of June 30, 2006;

- Salinity models for the southern Everglades mangrove zone, Florida Bay, Whitewater Bay, and the southwest Gulf coast areas only will be reviewed (the subject area for this CESI project);
- Information will be gathered primarily from abstracts and papers that are available from the sofia.usgs.gov website, the evergladesplan.org website, the latest Greater Everglades Ecosystem Restoration (GEERS) conference proceedings, the latest Florida Bay Science Program conference proceedings, and from personal communications; and
- Models reviewed will include models currently in use by ENP as well as models not currently being used by ENP that have the potential for use.

2 SALINITY PATTERNS AND PROCESSES

2.1 Salinity Variation

Salinity in south Florida estuaries varies with time, and this variation can be expressed on a wide range of scales. In general, salinity varies with the annual wet and dry seasons driven by the regional precipitation and temperature patterns of sub-tropical south Florida. Salinity also responds to episodic meteorological events such as tropical storms and cold fronts. In addition, salinities in south Florida estuaries are influenced by interannual El Nino/La Nina cycles and by decadal variability in precipitation driven by Atlantic Ocean multi-decadal forcing.

A preliminary analysis of historical data from the Salinity Synthesis database (Robblee et al. 2001, Nuttle et al. 2000) shows that for the past few decades Florida Bay has behaved as a marine lagoon with, on a few occasions, salinities as high as 70 psu reported in central Florida Bay. During drought years of this period, salinity typically exceeded 40 psu over most of the Bay. Although estuarine, i.e. mesohaline, conditions in the open-water areas of Florida Bay are rare in recent history, precipitation has increased since the mid-1990's as has happened in the past when there is a change in the Atlantic Multidecadal Oscillation (AMO) a regional climate indicator (Enfield, et al 2001). Lowered salinity conditions are also associated with episodic events such as hurricanes and tropical storms or with other periods of above average rainfall, for example in 1994-1995. Increased water releases from the C-111 canal can also lower salinity across the Bay during relative dry years, as occurred 1983-1985.

Spatially within Florida Bay, salinity variability is greatest in the northeast and decreases to the west. Boyer et al. (1997) and Boyer and Jones (2001) have described a decadal trend in monthly salinity values collected by Florida International University (FIU)'s Southeast Environmental Research Center (SERC) in Florida Bay from 1989-1999. Over this interval salinity in the eastern, central and western region declined 13.6, 11.6, and 5.6 psu respectively, but this "trend" is due largely to very high values during the 1989-1990 drought and is not descriptive of substantial interannual variability. Following the wet period of 1994-1995 salinity increased, tropical storms in 1999 induced a salinity decrease, and data from the drought years of 2000 and 2001 show an increase. Therefore, the effect of precipitation patterns and episodic events can be traced in the salinity record.

In the estuaries of Whitewater Bay and the Gulf coast salinity variation has been affected by the same climatic factors as Florida Bay. However, the impact on salinity variation in these estuaries has been caused more by the decrease in freshwater supply to Shark River Slough than by water management features such as the C-111 Canal. The trend, if any, in salinity in these estuaries over the past decade has not been studied as it has in Florida Bay.

2.2 Driving Processes

Salinity in the study area reflects the shifting balance between the inflow of fresh water, the continual exchange of water with the coastal ocean and the Gulf, variation in evaporation, and physical circulation effects. The interplay of these hydrologic drivers creates a transition zone of increased salinity with distance downstream, from an upstream freshwater body to a marine downstream water body. The inflow of lower density fresh water from the Everglades dilutes the salinity in the estuary and moves the transition zone toward the bay, while exchange with the coastal ocean or Gulf replaces diluted estuarine water with water of higher salinity and greater density. As a consequence, changes in estuarine salinity, both in time and in space, are driven by the variation in three basic processes: the net supply of fresh water, the processes that drive mixing within the estuary (wind, geomorphological features, hydraulic effects), and the exchange of salinity with the coastal ocean.

2.2.1 Net Freshwater Supply

Net freshwater supply in Florida Bay and Whitewater Bay is the sum of rainfall over the area plus the inflow of fresh water through the coastal mangroves from the Everglades minus evaporation from the bay. Rainfall and evaporation dominate the freshwater budget in Florida Bay, but inflow from the Everglades is comparable in magnitude to the difference between rainfall and evaporation. For the Gulf coast estuaries, freshwater supply is dominated by the upstream contribution from Shark River Slough. An annual water budget for Florida Bay has been constructed using 31 years of salinity, hydrology and climate data (Cosby et al. 2005, Nuttle et al., 2000; 2001). From 1965-1995, annual runoff from the Everglades was one fifth of the annual direct rainfall into the Bay, and annual evaporation slightly exceeded annual rainfall. The freshwater budgets for Whitewater Bay and the Gulf coast estuaries has not been studied in detail.

On a seasonal basis rainfall, evaporation, and wetland inflow are not in phase. However, the overlap of rainfall and inflow and the opposition of evaporation lead to a strong seasonal pattern of salinity in Florida Bay (lowest in the fall, highest in the spring). Inter-annual variations in salinity appear to be affected primarily by fluctuations in rainfall both over the Everglades and over Florida Bay. These fluctuations influence salinity in the bay directly and also indirectly through variations in inflow to Florida Bay from the southern Everglades. Relative to the available data on rainfall and inflow of surface water, little is known directly about the rate of evaporation and its variation seasonally, year-to-year and spatially within the bay. Nuttle et al. (2001) describe an investigation that is designed to provide mean rates of evaporation and its variation both spatially and temporally in Florida Bay.

Freshwater inflows from Trout Creek and Taylor River (the two largest tributaries to Florida Bay) have significant influence on salinity patterns and variability in northeast Florida Bay. Low salinity values can be found near the creek mouths during the wet season, and strong gradients can occur in northeast Florida Bay.

In the fall, when sea level is relatively high, inter-basin exchange within the Bay is enhanced resulting in more of the inflowing fresh water to northeast Florida Bay reaching the central basin. When freshwater inflow to northeast Florida Bay is reduced by drought or water management practices, hypersaline conditions often develop in the central region during the dry season. Runoff from large tropical storm events can raise water levels in the mangroves and inject fresh water into the central region; this occurred notably following Hurricane Georges in 1998 (see Hurricane Georges Workshop report available at http://www.aoml.noaa.gov/flbay/hurgeocoverpage.html) and tropical storm Harvey and Hurricane Irene, both in 1999.

No large-scale effects of groundwater on salinity have been observed, but observing this input is made more difficult by the fact that the primary groundwater input is thought to be saline (Price et al. in press, Price 2001, Sutula et al. 2001). Estimates of groundwater discharge directly into the bay vary over orders of magnitude (from 1 to 16 cm/d) with higher values obtained during the dry season (Top et al., 2001; Corbett et al., 1999). Anecdotal evidence reports the existence of local springs and these have been observed by scientists in nearby Biscayne Bay within a few hundred meters of the shoreline (J. Proni, pers. comm.).

2.2.2 Exchange Processes

In northeast Florida Bay and the open water areas west (behind) the Florida Keys the lunar tides have only limited influence on driving the exchange of water with the adjacent coastal waters of the Atlantic Ocean and the Gulf of Mexico. The shallow bathymetry of the bay attenuates the influence of tides in the central and northeast regions of the bay. Here, wind-driven flows and longer term fluctuations in sea level are the dominant drivers of mixing and exchange processes. However, in the western open water areas and along the Gulf coast (including Whitewater Bay) the tidal influence also plays a role.

Wind-driven flows can affect salinity in Florida Bay both directly and indirectly. Direct effects include advection of the freshwater plumes discharging from the coastal creeks. Indirect effects include redirection of low salinity plumes from the Shark River Slough discharge into Gulf of Mexico waters into western Florida Bay. D. Smith (2001), Johns et al. (2001), and R. Smith et al. (2001) have recently shown that storm events cause significant and long-lasting (~months) changes in salinity patterns and turbidity in Florida Bay. In combination with upstream water management releases (due to flood control restrictions), storms can affect salinity by discharging pulses of fresh water into Florida Bay and the Gulf of Mexico even in the dry season. Tropical storms can cause Everglades water levels to rise rapidly, and then recede slowly as fresh water is discharged from upstream areas into northeast Florida Bay. Nor'easters drive large quantities of fresh water out of the Everglades into Florida Bay, usually during the dry season.

Fluctuations in sea level also affect exchange and mixing processes. First, annual sea level variations propagate throughout the bay essentially without attenuation. Annual changes in sea level within Florida Bay are on the order of 20 cm, representing the exchange of about 20 percent of the bay's total volume with surrounding coastal waters. Changes in water levels of this magnitude over the shallow topography have the potential to modulate the magnitude and patterns of exchange driven by wind and tides. During periods of low sea level, connections between adjacent basins are restricted thereby minimizing overbank exchange. Higher water levels facilitate the mixing of fresh water entering the northeast region from the Everglades across the shallow banks into the central and south regions of the bay.

Additionally, the increase in the elevation of higher density salt water relative to the lower density freshwater moves the broad transition zone between fresh and salt water bodies towards the upstream (inland) areas, until the increasing freshwater levels in the Everglades in the wet season can overcome the pressure created by salt water elevation and density in Florida Bay. In drought periods when freshwater inflow to the Everglades and direct rainfall are limited, saline water brought into the shallow near-shore embayments is not washed out, and high evaporation rates at the end of the dry season (late spring to early summer) can create hypersaline conditions in the near-shore areas.

2.2.3 Boundary Salinity

Flow from the rivers and tidal creeks along the southwest Florida coastline (primarily from Shark River Slough) can reduce salinities in the estuaries along the Gulf coast, the western perimeter of Florida Bay, and in its westernmost basins. Remote river discharges, i.e. the Mississippi River can be transported by coastal and boundary currents along the Florida Shelf to different parts of the south Florida coast (Lee et al., 2001a; 2001b; Ortner et al., 1995). Since most mechanistic models utilize the Gulf of Mexico as a boundary condition changes in salinity due to influence from distant sources can be a source of uncertainty in some cases and at certain times.

2.3 Available Salinity Data

In all instances, uncertainty in salinity modeling and forecasts depends on the available data. The amount and quality of the data available can affect the description of the variation in the driving processes and the resulting variation in salinity which may limit the predictive capacity of any model (CERP 2002). The models reviewed in this report rely on several important datasets that have been assembled by a number of agencies. Most models utilize data for some but not all of these processes to forecast salinity. The type and complexity of model governs the data that are used for model development and simulation.

The ENP Marine Monitoring Network (MMN) stations collect salinity and other data in Florida Bay and the southwest Gulf coast waters at 15 to 60 minute increments, which are averaged to reported hourly and daily values

(http://www.sfnrc.ever.nps.gov/). The entire hydrologic monitoring network within Everglades National Park includes 62 freshwater sites (Physical Monitoring Network) throughout the marsh and 37 marine/estuarine sites (MMN). At the freshwater sites, the oldest stations have been operating since 1949. For the MMN, salinity measurements began to be collected in 1988. Parameters measured include water level, rainfall, water temperature, and salinity. Collection of data is automated and the data are transmitted to a base station using telemetry. According to the website information, these data are used to determine the effect of the Central and Southern Florida Flood Control Project on the ENP natural resources, and to characterize the park's water resources. Details about these data can be found in Everglades National Park (1997a,1997b, 2001, 2003), and Smith (1997, and 1998).

As a separate effort, the Southeast Environmental Research Center (SERC) at Florida International University (FIU) collects monthly salinity data as part of a grab sample monitoring program that has been on-going since the early 1990s. This long-term program monitors water quality in the coastal waters of south Florida. This program visits 24 stations in Florida Bay and 21 stations along the southwest coast on a monthly basis. Water samples are analyzed for salinity along with a suite of nutrient and other water quality parameters. The report by Jones and Boyer (2001) summarizes the data and discusses long-term trends in water quality on a regional basis (http://serc.fiu.edu/sercindex/index.htm). There is also synoptic salinity data available, some of it recorded as data collected with a study other than salinity. Historical salinity data (prior to the mid-1980's) exist (Robblee et al, 2001), but continuous records are spotty (Nuttle et al. 2000). A description of some of the historical salinity data can be found in Orlando, et al (1998).

The South Florida Water Management District (SFWMD) collects stage data in the Everglades, and stage and flow data at each of the structures that are a part of the water management system of the Central and South Florida (C&SF) Project, including structures that affect freshwater delivery to ENP (http://www.sfwmd.gov/org/ema/dbhydro/index.html). The United States Geological Survey (USGS) collects flow data in the tidal creeks that flow into Florida Bay through ENP, and salinity data in some of the near-shore embayments (http://sofia.usgs.gov/).

To these data other time series data can be added, including wind and rain data from the National Weather Service (Southeast Regional Climate Center – (http://water.dnr.state.sc.us/climate/sercc/climateinfo/historical/historical_fl.html), and water level and tide data collected long-term at Key Westand Naples,and for shorter periods at other south Florida locations from the National Ocean Service (http://tidesonline.nos.noaa.gov/). Off-shore sea surface elevation data which are used for boundary conditions for some mechanistic models are available through the CMAN / SEAKEYS programs (http://www.ndbc.noaa.gov/cman.php; http://www.coral.noaa.gov/cman/cman_menu.html). Evaporation estimates are

often based upon the work of German (2000), or Nuttle (2001). Evaporation estimates have also been made by SFWMD for use with the South Florida Water Management Model (SFWMM) (Abtew, et al 2003).

2.4 Estimating Inflow from the Everglades to Florida Bay

Methods used to estimate fresh water inflow to Florida Bay from the Everglades deserve special attention because inflow is an important parameter for mechanistic salinity models, and can be incorporated into statistical models. Fresh water entering Florida Bay from the Everglades flows as diffuse overland flow, canal flow, and ground water. These various flow paths link water management activities of the Central and Southern Florida Project to the coastal ecosystems of South Florida. Understanding how changes to water management and restoration of the Everglades hydrology affect the ecology of the coastal ecosystems depends on information about this hydrological link between flow and salinity. Only in the last ten years have these flows been studied in detail.

There are eight estuarine creeks along the northern boundary of northeast Florida Bay. The U.S. Geological Survey has monitored flow and salinity continuously in five of these creeks since 1996. In addition to monitoring in the five creeks, Hittle et al. (2000) published empirical relationships for estimating instantaneous flow in three ungauged creeks from flow measurements in Taylor River (TR) and West Highway (WH) creek. Except for these empirical relationships, there appears to be no information from direct measurements on the magnitude of the ungauged discharge of fresh water from the Everglades directly into Florida Bay.

Everglades freshwater hydrology models such as the SFWMM do not directly estimate flow into Florida Bay. For that purpose, wetland basin models have been developed to estimate fresh water inflow into most areas of the estuarine and near shore areas in south Florida. For example, Walker (1998) implemented a set of watershed hydrology models in Everglades National Park to estimate nutrient loads to estuarine and nearshore waters. The PHAST models (Nuttle and Teed 2002, Nuttle 2004) cover the three wetland sub-basins in the Taylor Slough C111 wetland basin where discharge measured by the USGS in estuarine creeks can be used to verify the calculated wetland discharge. PHAST. Nuttle (2005) implements a series of wetland basin models as part of a screening tool for initial planning of the Biscayne Bay Coastal Wetland CERP project.

The most widely used freshwater hydrology model for south Florida was developed by the SFWMD and is called the South Florida Water Management Model (SFWMM), also known as the 2X2 model for the two mile by two mile (3.2 km X 3.2 km) grid size of the model. The SFWMM simulates the hydrology and the management of the water resources system from Lake Okeechobee to Florida Bay. It covers an area of 7600 square miles using a mesh of square cells.

The domain of this model stretches from Lake Okeechobee to the mangrove zone in Everglades National Park. The model utilizes inflows from Kissimmee River, and accounts for runoff and demands in the Caloosahatchee River and St. Lucie canal basins.

For each cell, the model simulates rainfall, evapotranspiration, infiltration, overland water flows, groundwater flow, canal flow, canal-groundwater seepage, levee seepage and groundwater pumping as well as water management control structures. SFWMM runs incorporate both current and proposed operational programs for structures in the water management system for urban, agricultural, and environmental water demands in south Florida. To evaluate water management alternatives for CERP, the SFWMM simulates Everglades hydrology on a daily basis using climatic data for 1965-2000 as model input. The model has been calibrated and verified using observed water level and flow data in the Everglades (http://www.sfwmd.gov/org/pld/hsm/models/sfwmm/index.html). Applications of the SFWMM have included the Initial CERP Update (ICU), the Restudy, the Lower East Coast Water Supply Plan, and the development of operational protocols for Lake Okeechobee. The SFWMM has been used by ENP for the Interim Structure and Operation Procedures (ISOP), the Interim Operation Procedures (IOP), and the Combined Structure and Operation Procedures (CSOP) evaluations (Santee, et al; 2003).

For FATHOM salinity modeling for minimum flows and levels for Florida Bay, an alternative approach was taken to constructing the inflow data (Cosby, et al 2005). Components of the wetland water budget are estimated from the available regional hydrological and climatic data. Freshwater inflow was estimated from the estimated water budget as monthly average flows. This alternative approach maintains the framework of the wetland water budget for combining information from long-term data sets. Uncertainty over the amount of ungauged fresh water inflow, inflow that occurs in addition to that measured by the USGS creek monitoring network, is dealt with explicitly in the manner of an unknown parameter in the salinity calculations (Cosby et al. 2005).

Hydrodynamic models under construction by the United States Geological Survey (USGS) and the SFWMD in conjunction with their ongoing measurement programs are expected to reduce the uncertainty concerning surface water discharges into Florida Bay. Hydrodynamic transport models have been developed that are capable of being linked to upland management models to address the impact of fresh water inflows on salinities and the conveyance of nutrients and contaminants to Florida Bay and southwest coastal estuaries. Langevin et al. (2002) of the USGS have developed a coupled surface / groundwater model of the southern Everglades (SICS / TIME) that is capable of simulating flow into the coastal embayments of northern Florida Bay in response to naturally occurring hydrologic events in the wetlands and the effects of upland management practices on fresh water releases.

The total groundwater discharge to the Everglades surface water regime on an annual basis for the entire SICS model domain is on the order of only one cubic meter per second (E. Swain, pers. com). This is a small contribution that can be ignored for some calculations. This estimate is based on analysis of the average net groundwater discharge calculated for each node of the SICS model within the Taylor Slough and C111 wetland basins. Net figures for groundwater exchange in the period 1995 through 1999 show patterns of upwelling to the surface in the upper reaches of Taylor Slough, north of Old Ingram Highway.

3 STATISTICAL MODELS

3.1 Regression Models

The first documented attempt to use statistical methods to establish the relationship between upstream hydrology and downstream salinity of Florida Bay and the Gulf coast estuaries is attributed to Durbin Tabb (1967). Tabb (1967) observed a close relationship between salinity and the elevation of the water level at two wells in ENP (P-35 and P-38) and one well in Homestead. Simple linear regression methods were used (including lagged values) to develop salinity prediction tables that were compared to observed data, and the correspondence was considered to be acceptable.

After the work by Tabb (1967), attempts were made to develop statistical relationships between water levels in monitored wells in the Everglades and Florida Bay salinity, primarily by the SFWMD using data that were collected monthly (Scully, 1986). Other statistical evaluations of salinity variation using correlation evaluations, simple linear regression models, and analysis of variance followed, including Cosby (1993). Though the performance of simple linear models was reasonable for salinity at some stations in Florida Bay nearest the mangrove fringe area, the effects of wind, tide, and local meteorological conditions, particularly at open Bay locations, limit the ability of the simple linear regression relationships to satisfactorily simulate salinity (Marshall, 2000). The use of other time-series modeling techniques, such as seasonal autoregressive integrated moving average models (SARIMA) that are robust to outliers and seasonality, was suggested as an improvement for simulation.

Nuttle (1997) was the first to implement more sophisticated statistical techniques using monthly data that included monitored C-111 Canal structure flows and rainfall. The resulting updated transfer functions solved some of the statistical problems that were associated with the initial simple linear models (such as a non-constant variance), but use of the models for predictive purposes was hampered for by the transformations needed to stabilize the non-constant variance of the monthly input data.

The conceptual model for the south Florida coastal aquifer system explains the relationship between estuarine and coastal shelf salinity, hydrology in the upstream watershed, and meteorology in the region. For this coastal aquifer system there is a dynamic balance between fresh and salt water bodies with a salinity transition zone from freshwater to sea water (Pandit, et al 1991). In south Florida, the salinity transition zone is wide because of the relatively small difference between upstream freshwater stage and downstream estuary water level. In most of the coastal aquifer examples in the literature, the focus of analysis is the water table aquifer, with the primary concern being the location of the transition zone as a water supply issue of saltwater intrusion. For salinity modeling in an estuary the focus is the salinity in the interface transition zone.

The Ghyben-Herzberg principle describes the location of this interface as function of the height of the freshwater surface in the watershed relative to the height of the sea surface above a common datum and the relative density of the water masses. In shallow estuaries like Florida Bay and Whitewater Bay wind can also cause the interface to translocate and mix. Therefore, Marshall (2002) hypothesized a correlation between salinity levels and these three factors (watershed water level or elevation, sea surface elevation, and wind), which is confirmed by a correlation analysis including lagged values on the order of days. However, each of these forcing factors (fresh water elevation, wind, and sea surface elevation) has a different pattern of variability over time

The MLR salinity models of Marshall (2002) used SFWMM model output for water levels in the Everglades and available long-term data for wind and sea surface water level to produce estimates of daily salinity for the 1965-2000 period. Although rainfall in the upstream watershed is an important hydrologic parameter for seasonal salinity variation, rainfall at monitoring stations in the Everglades are not highly correlated with salinity at the daily level. Instead, the stochastic effect of rainfall falling on the Everglades is integrated by the coastal aquifer system and expressed adequately in stage data.

The original MLR salinity models were developed for Joe Bay, Little Madeira Bay, Terrapin Bay, Garfield Bight, and North River (Marshall 2002). For the Interim Operating Plan (IOP) Congressional report, an additional station (Long Sound) was added (Marshall, 2002). Then, the second phase of the CESI project extended the spatial extent of MLR models to Highway Creek, Taylor River, Whipray Basin, Duck Key, Butternut Key, and Bob Allen Key (Marshall 2004).

In 2005, the Southern Estuaries Sub-team of RECOVER used the 12 existing MLR salinity models and developed new models for Gunboat Island, Shark River, Clearwater Pass, and Whitewater Bay, Barnes Sound and Manatee Bay (Middle Key station). By the end of 2006, statistical salinity models will be developed for the remaining stations in Florida Bay and on the Gulf coast. Figure 1 presents the MMN station locations with MLR salinity models.

A number of error statistics were computed (Marshall 2005) in order to quantify the uncertainty in the simulations produced by the MLR salinity models (See Tables 1 and 2). In general, the salinity in the near-shore embayments (Joe Bay, Little Madeira Bay, Terrapin Bay, and Garfield Bight) is observed to be more variable on a day-to-day basis than the salinity at the open water stations. Because of this, the development of MLR salinity models was more difficult and the R^2 and Nash-Sutcliffe Efficiency values are slightly lower compared to the open water MLR salinity models.

Error statistics indicate that the daily resolution MLR salinity models are capable of explaining on-the-order of 70 – 80% of the variation in salinity. However,

individual daily residuals can sometimes be large. At the weekly average level, large residuals are uncommon. This grouping of similar models by goodness-of-fit statistics follows closely the groups presented in Orlando et al (1998) from an archival salinity data set.

3.2 Time Series and Frequency Domain Models

Most of the statistical modeling of salinity completed prior to 2002 was done with monthly data. Marshall (2002) investigated the development of time series statistical models using daily time series data. Daily variability in salinity is valuable to a variety of biologists, because daily variability captures the "flashiness" of the system relative to changes in salinity.

When Marshall (2002) tried to apply SARIMA time series models to Florida Bay salinity, it was found that there were fewer limitations with multivariate linear regression (MLR) models for coupling with SFWMD model output to produce 36-year simulations. Therefore, Marshall (2002) adapted a SARIMA technique using cross-correlation coefficient analysis to efficiently identify significant variables and lagged values with MLR salinity models.

Nuttle and Marshall (unpub., 2005) applied spectral analysis to examine long-term salinity records from Florida Bay. Spectral analysis can be used as a diagnostic tool of system behavior by examining the characteristic spectral density function for Florida Bay salinity. Peaks in the spectral density function of salinity at periods of 12.5 hours and 25 hours signal the importance of diurnal and semi-diurnal tidal forcing on a number of estuarine processes such as water level, velocity of flow, and salinity. In addition to the information provided by peak values, trends in the spectral density function are also important. For example, a 5/3 slope found in the spectral density function for certain characteristics of marine ecosystems reveals the influence of turbulent mixing processes (Levine 1996). Changes in spectral density function slope can signal a change in the mode of behavior, such as between Eulerian and Lagrangian turbulence (Seuront el al. 1996), or in the underlying processes that control ecosystem structure and function (Holling 1992, 1996). The results obtained for Florida Bay indicate that although tides and wind-driven water movement are important over the short term, the variation in the supply of fresh water to the bay contributes significantly to changes in salinity over time scales greater than about a month. Spectral analysis allowed this change in the pattern of salinity variation with increasing time scale to be seen, an important finding for salinity modelers.

Figure 1: Map showing the location of Marine Monitoring Network stations with MLR salinity models (from Everglades National Park).

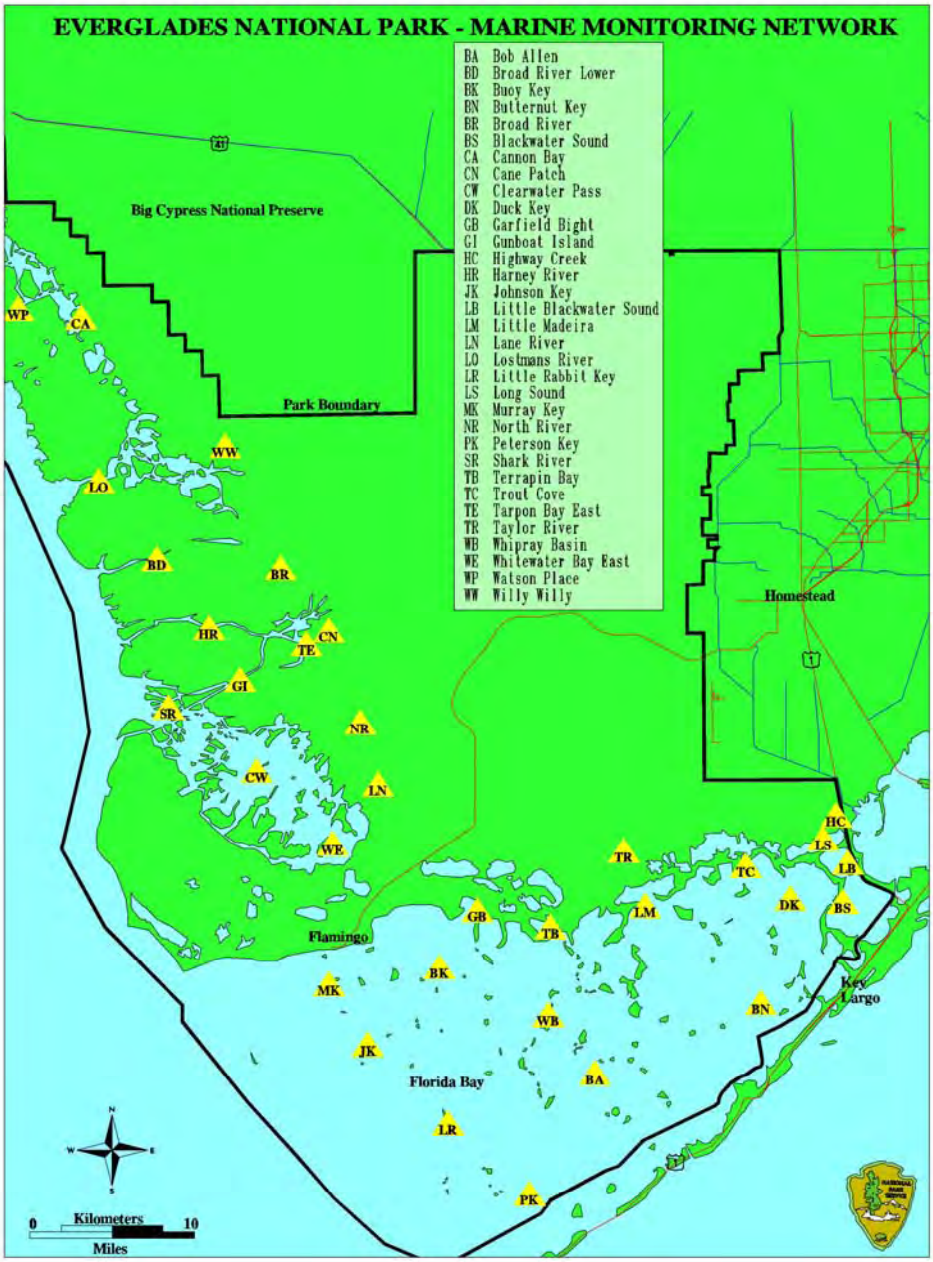


Table 1. Comparison of Model Uncertainty Statistics for IOP / CESI MLR Salinity Models (Marshall, 2004).

station	mse psu^2	root mse (rmse), psu	adj R-sq	mean error, psu	mean abs error, psu	max abs error, psu	Nash-Sutcliffe Effcy
Joe Bay	25.8	5.1	0.75	-0.14	3.7	20.6	0.76
Little Madeira Bay	40.1	6.4	0.65	-0.66	5.1	22.6	-0.96
Terrapin Bay	32.6	5.7	0.75	-0.99	5.4	5.4	0.67
Whipray Basin	7.2	2.7	0.8	0.11	2.2	10.1	0.77
Duck Key	9.7	3.1	0.71	-0.18	2.27	14.4	0.71
Butternut Key	10.7	3.3	0.65	0.1	2.7	11.3	0.66
Long Sound	15	3.9	0.8	0.31	2.7	18.9	0.81
Taylor River	21.4	4.6	0.78	-0.49	3.6	22.9	0.78
Highway Creek	18.2	4.3	0.81	-0.95	3.7	17.7	0.76
Little Blackwater Sound	14	3.7	0.75	-0.14	2.9	15.7	0.76
Bob Allen Key	7.2	2.7	0.79	0.3	2.1	9.2	0.81

Table 2. Comparison of Model Uncertainty Statistics for Southern Estuaries Sub-team MLR Salinity Models (Marshall, 2005).

station	mean square error	root mse	adj R-sq	mean error	mean abs error	max abs error	relative mean error	relative mean abs error	relative mean square error	Nash-Sutcliffe Effcy.
Garfield Bight	37.9	6.15	0.68	-0.36	4.75	21.1	-0.012	0.16	0.06	0.89
Clearwater Pass	11.60	3.40	0.85	-0.12	2.72	10.82	-0.01	0.16	0.08	0.85
Whitewater Bay	9.60	3.10	0.74	0.46	2.90	10.60	0.04	0.26	0.06	0.88
North River	14.30	3.80	0.77	0.56	3.23	17.92	0.08	0.45	0.04	0.92
Gunboat Island	11.50	3.40	0.85	1.03	3.02	13.28	0.09	0.27	0.05	0.89
Shark River	6.30	2.50	0.82	-0.11	2.02	9.11	0.00	0.08	0.06	0.89
Middle Key	6.88	2.60	0.74	-0.22	2.20	11.33	-0.01	0.09	0.16	0.71
Manatee Bay Stage	9.50	3.10	0.69	0.02	2.07	12.86	0.00	0.09	0.17	0.70

4 MECHANISTIC MODELS

4.1 Mass Balance (Box) Models

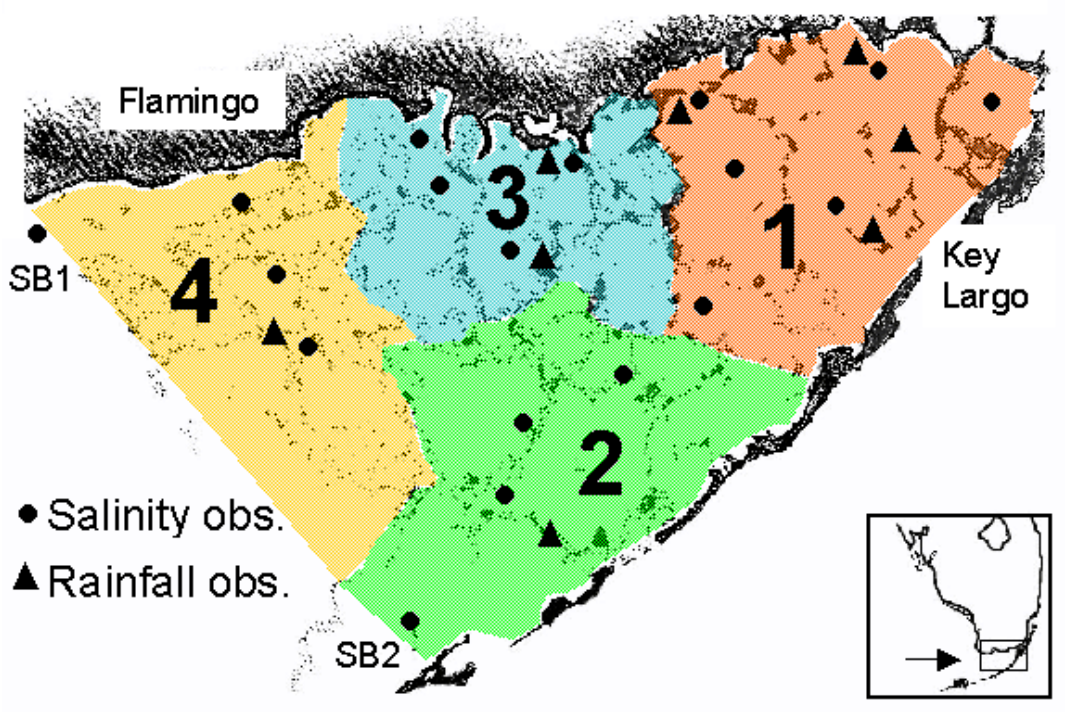
4.1.1 Four Box Model of Florida Bay

Nuttle et al. (2001) implemented a mass balance model in Florida Bay for the purpose of estimating evaporation. This model calculates salinity using monthly time steps from variation in the net supply of freshwater to and water exchange between each of four regions in the bay, Figure 2, and exchange with the Gulf of Mexico. The regions used in this model correspond to the regions defined from similarities in water quality (Boyer et al. 1997) and other attributes of the Florida Bay ecosystem. Rainfall and salinity used to drive the model are measured in the bay. Freshwater runoff is estimated from measured flows in Taylor Slough and the C111 canal that discharge into the mangrove wetlands north of the Florida Bay.

The four-box model by Nuttle et al. (2001) has been calibrated against salinity data for the period 1993 through 1995 and validated by comparison with salinity data in the period 1996 through 1998, Figure 2. Monthly salinity and rainfall data were aggregated within each region. Salinity at SB1 and SB2 provide boundary conditions for exchange with regions 2 and 4. Freshwater runoff also enters region 1. Evaporation and exchanges between regions were estimated by optimization. The standard error of prediction is about 2 ppt across all four regions. Calibration of the model produces estimates for the unknown seasonal evaporation rates and the exchange rates between basins and with the Gulf of Mexico. These exchange rates can be used to investigate residence times in the Bay, information that is needed to understand the processes that control nutrient concentrations and plankton blooms in the bay.

Nuttle et al. (2000) employ two different box models. One is the annual averaged version of the four-box model described above, which was used to estimate mean annual evaporation from Florida Bay. The other, FATHOM (described below), divides the Bay into about 40 basins based on morphology, and estimates exchanges between basins using tide-driven hydraulic calculations. FATHOM has been applied to analyze the influence of changing runoff into Florida Bay (Nuttle et al. 2000), but the calculated exchange rates and resulting residence times have yet to be validated by comparison with observation.

Figure 2: The box model divides Florida Bay into four regions based on observed patterns in water quality (Boyer et al 1997, Nuttle et al. 2000).



4.1.2 PHAST

The Nuttle (2002) report recommended adopting the mass balance modeling approach, and this recommendation led to the development of the aggregated wetland basin hydrology and estuarine basin salinity model (PHAST) for ENP, and more recently used as a planning tool for the Biscayne Bay Coastal Wetlands Project (Nuttle 2005). The domain of the PHAST models encompasses three wetland sub-basins in the Taylor Slough C111 wetland basin and the adjacent estuarine basins Long Sound, Joe Bay, Little Madeira Bay and Terrapin Bay. The PHAST models have been applied by Everglades National Park to simulate changes in salinity as a performance measure of restoration in the development of a spoonbill habitat suitability index model (Lorenz 2005) and in modeling studies to support development of minimum flows and levels criteria for Florida Bay (Cosby et al. 2005).

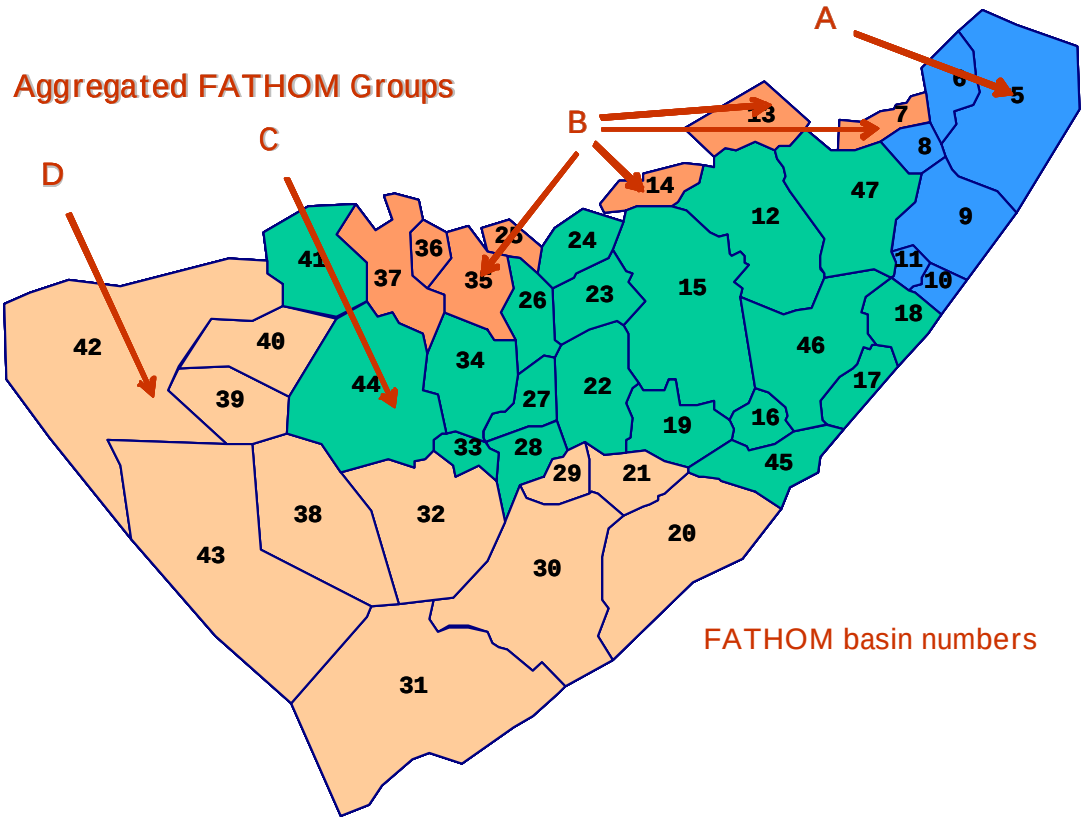
4.1.3 FATHOM

FATHOM is a dynamic, spatially explicit, mass-balance model designed to investigate the response of salinity in Florida Bay to runoff, climate, and variation in salinity on the Florida Shelf (Cosby et al. 1999, Nuttle et al. 2000, Cosby et al 2005). The model maintains a running account of the water and salt budgets in each of 41 well-mixed basins within the bay, Figure 3. Circulation within Florida Bay and exchange with the Florida Shelf are controlled by the network of shallow banks. The basins defined by these banks offer a natural framework for mass-balance accounting.

FATHOM represents Florida Bay as a collection of well-mixed basins. Circulation and exchange are driven primarily by tides imposed along the western boundary. At each time step, the model solves for uniform, hydraulic flow across each bank based on the depth, width, and frictional roughness of the bank and water levels in the upstream and downstream basins. By this mechanism, tidal forcing at the boundary propagates into the bay and drives the exchange of water and solutes among the basins. Solute fluxes are then calculated from water fluxes and the salinity of water on each bank. Details of the representation of flow over the banks and the hydraulic equations are given in Cosby et al. (1999).

Despite the model's computational simplicity, FATHOM requires highly detailed information about the bathymetry in Florida Bay. Bathymetric data are entered into a GIS database that classifies the depth for every 20 by 20 meter rectangle in the bay.

Figure 3: Map identifying the basins and the aggregated regions used in FATHOM for salinity calculations.



There is no direct simulation of wind shear on the water surface in FATHOM. Salinity calculated by FATHOM represents a time-averaged value with a period of about one month. Even though circulation and exchange in the model are driven by tides, data on other processes such as rainfall and freshwater inflow are provided as monthly values. The assumption of basins as well mixed imposes constraints on the time scale on which salinity calculations can be taken as comparable to observations at any particular location.

Bay-wide inputs required by FATHOM include time series of rainfall and evaporation for each basin in the bay. The model structure allows these inputs to be specified individually for each basin to reproduce spatial gradients in these forcing functions. In practice, however, observed data are not sufficient to support more than a regional approach to the spatial distribution of climate inputs. Instead, the bay must be divided into a few regions for each of which climate inputs are applied uniformly to the model to make long-term forecasts of salinity. Groundwater inputs to the basins can also be specified, but these have not been employed in the simulations performed for this project.

For FATHOM time series of freshwater inflow volumes are required at the terrestrial boundaries of the bay. Inflow is specified as an input separately into each of the boundary basins along the Florida Bay coastline. Along the keys, inflows of fresh water are small, and these are not included in the FATHOM inputs. In addition to the runoff data at the terrestrial boundaries, FATHOM requires tide, sea level and salinity time series to set the open water boundary conditions for the bay. The model allows these boundary conditions to vary spatially along the boundaries.

For the Florida Bay minimum flows and levels (MFLs) modeling for SFWMD, the bathymetry of Florida Bay was updates and freshwater inflows were improved using USGS observations and a sensitivity analysis approach. This effort produced 31-year (1970-2002) historical reconstructions of salinity in each of the 41 FATHOM basins. The salinity reconstruction for Little Madeira Bay was used as input to ecological models.

Error statistics for the FATHOM MFL base case calibration / verification run (1991-2000) are presented in Table 4 for monthly simulations. While not directly comparable, root mean squared error, absolute error, and r^2 for FATHOM are similar to the same statistics for the MLR daily salinity models (Tables 1 and 2). Average error is higher and efficiency values are less for FATHOM MFL base case model compared to daily MLR salinity models.

Table 3. : Error statistics for salinity simulations by FATHOM for the MFL base case model with monthly measurements over the period 1991 through 2002. (adapted with permission from Cosby, et al 2005).

station	root mse (rmse), psu	adj R-sq	mean error, psu	mean abs error, psu	Nash-Sutcliffe Effcy
Long Sound	4.3	0.9	1.9	3.4	0.77
Joe Bay	7.7	0.8	-1.9	5.6	0.56
Little Madeira Bay (mouth)	4.2	0.9	1.3	3.3	0.76
Park Key	3.7	0.9	2.0	3.1	0.77
Duck Key	3.7	0.9	-1.6	3.0	0.76
Butternut Key	3.5	0.9	-0.2	2.8	0.90
Garfield Bight / Rankin Bight	5.9	0.7	1.5	4.5	0.43
Whipray Basin	4.5	0.8	0.2	3.6	0.58
Rabbit Key Basin	2.6	0.7	0.0	1.9	0.51

4.2 Hydrodynamic Models

4.2.1 Initial Hydrodynamic Modeling Efforts

There were various preliminary detailed modeling efforts in the 1990's that were associated with Florida Bay restoration projects which were not carried forward. Examples include models developed or described by Wang et al (1994) and Wang (1998), Sheng et al (1995), and Cerco et al (2002). Because these modeling efforts are no longer active they were not included in the model summary, but they are worthy of note. Because of the spatial extent of the RMA-10-WES model developed by the US Army Corps of Engineers Waterways Experiments Station, it is discussed below.

The RMA-10-WES model is a two-dimensional version of a finite element hydrodynamic model that was used by the US Army Corps of Engineers in 1996 to simulate circulation in Florida Bay, and to be coupled with a water quality model, CE-QUAL-ICM (Cerco, et al; 2002). The RMA-10-WES grid mesh that was developed for Florida Bay consists of 19253 triangular elements and 40609 nodes, stretching from Barnes Sound to the Gulf of Mexico north of Johnston Key. When RMA-10-WES was coupled to CE-QUAL-ICM a number of issues surfaced. In addition, the grid for RMA-10-WES did not match the CE-QUAL-ICM grid. Several attempts were made to adapt the RMA-10-WES output for water quality use. Overall, the end result was not usable, and no documentation of any further effort could be found. The RMA-10-WES Florida Bay modeling activity by the Corps was notable because it was the first attempt in south Florida to link a hydrodynamic model to a water quality model, and it exposed the problems that have to be faced in that regard.

4.2.2 EFDC

The Environmental Fluid Dynamics Code (EFDC) is a hydrodynamic model that is used to simulate surface water systems in one, two, and three dimensions. EFDC is composed of stretched or sigma vertical coordinates and Cartesian or curvilinear, orthogonal horizontal coordinates to represent the physical characteristics of a water body. The code solves three-dimensional, vertically hydrostatic, free surface, turbulent averaged equations of motion for a variable-density fluid. Dynamically-coupled transport equations for turbulent kinetic energy, turbulent length scale, salinity and temperature are also solved. The EFDC model allows for drying and wetting in shallow areas by a mass conservation scheme.

The SFWMD EFDC model grid domain includes Florida Bay and extends westward into the Gulf of Mexico to the 81.9-degree latitude. The model was configured using NOAA and USGS bathymetry. Open boundary conditions

include tides and sea level, salinity, and temperature. Surface heat exchange is accomplished using spatially-varying wind and atmospheric data. Estimates of inflows, salinity, and temperature for canal, creek and river discharges are used for model input along northern Florida Bay and the eastern Gulf of Mexico.

The most current information on the EFDC model development as presented in the first draft of a calibration report to SFWMD in September, 2005 describes the a new multi-level grid for Florida Bay. This study describes the calibration of the hydrodynamic component of the EFDC model, building on the previous studies. In addition this new model configuration was used for a historical simulation from 1996 through 2002. The model calibration in this draft report describes the ability of the model to simulate sea level and currents frequencies at tidal and sub-tidal resolution, as well as temperature and salinity.

The new EFDC model has several grid resolutions. The coarse grid is shown by Figure 4. The EFDC grids consist of two configurations. One grid stops at the coast in northeast Florida Bay (nominal coast model), while the other configuration incorporates the mangrove area north of the Bay which is hydraulically connected to the open water areas of northeast Florida Bay (wetland model). According to the authors of the draft report, both configurations of the model performed well at reproducing observed sea level and currents at tidal frequency resolution. However, the grid that includes the mangrove zone is described as better in predicting the low frequency variation in water surface level in the northeast part of Florida Bay, including episodic events such as tropical storms.

As shown by Table 3, both grids (nominal coast model = NM, wetland model = WM) perform well in predicting salinity. According to Tetra Tech, Inc. (2005), the model is capable of reproducing seasonal variation as well as the extreme inflows that are caused by tropical storms. The smaller grid configuration (NM) apparently predicts better because groundwater is excluded, and there are problems depicting some mangrove zone features such as the Buttonwood embankment. Compared to daily MLR salinity models (Tables 1 and 2), monthly FATHOM model (Table 3), and the USGS SICS model (see below), daily salinity simulations by EFDC to-date contain significantly greater uncertainty (error) than the other 3 model systems. It is noted that the EFDC model development activity is on-going.

The EFDC modeling effort confirmed the physical processes at work in Florida Bay. For example the model shows that there is a shift in the tidal regime from macro-tidal in the western areas to micro-tidal in the central and eastern / northeast parts of the Bay. This transition is attributed to the mud banks that are said to attenuate tidal frequencies. In the east and in the central region sub-tidal frequency variations from the Gulf of Mexico and the Florida Straits and local winds are the primary water level drivers.

Because of the potential for problems specifying the open boundary conditions future EFDC efforts will utilize the ocean circulation model SoFLA-HYCOM (see below). This current draft report describes the initial work incorporating these boundary conditions. Additionally the surface and groundwater model TIME (see below) will be used in the future to simulate freshwater inflows to the mangrove zone.

Figure 4. Domain of the EFDC model, coarse grid (from TetraTech, 2003)

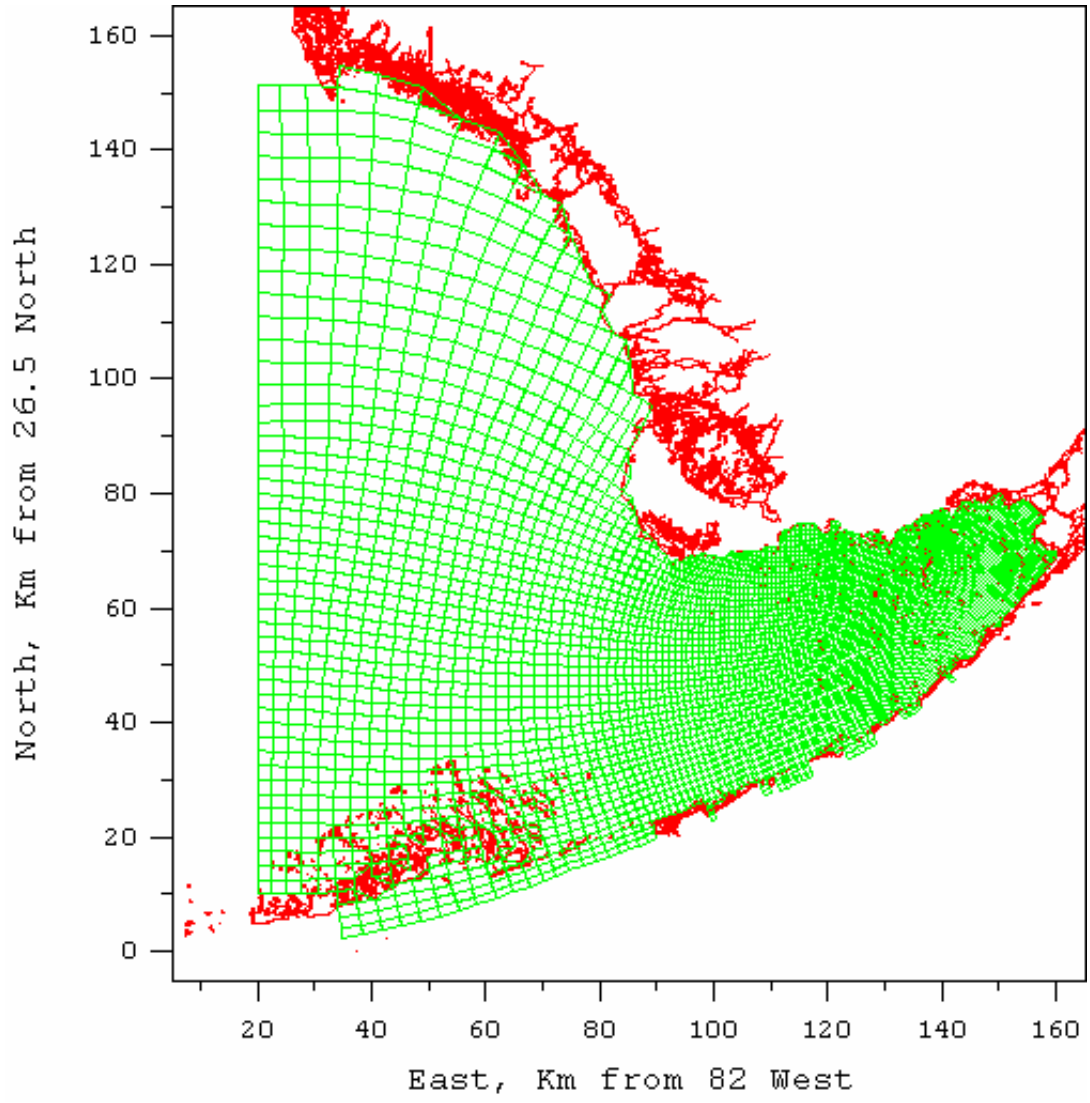


Table 4. Error statistics for EFDC salinity simulations at ENP MMN stations (Tetra Tech, Inc. 2005), nominal coast model = NM, wetland model = WM.

Station	Relative Error, PSU, NM	Relative Error, PSU, WM	Absolute Relative Error, PSU,NM	Absolute Relative Error, PSU,WM	Root Mean Squared Error, NM	Root Mean Squared Error, WM
Trout Cove	11.32	-3.57	11.34	6.27	14.18	8.08
Duck Key	1.64	-0.05	2.55	3.37	3.28	4.11
Little Madeira Bay	-2.35	-1.00	3.61	6.01	4.31	7.20
Butternut Key	0.13	-0.07	3.26	4.41	3.93	5.27
Terrapin Bay	-2.82	-1.10	5.27	5.67	6.46	7.15
Whipray Basin	2.04	3.30	3.40	4.64	4.04	5.64
Bob Allen Key	0.24	0.88	2.82	3.46	3.53	4.39
Garfield Bight	1.64	12.14	4.89	13.38	6.32	15.11
Buoy Key	2.29	6.26	4.63	7.36	5.70	9.13
Peterson Key	-1.15	-1.02	2.27	2.31	2.89	2.94
Murray Key	0.89	2.80	3.63	4.25	4.39	5.61
Johnson Key	1.10	2.71	3.80	4.59	4.69	5.97
Little Rabbit Key	-0.22	0.69	3.16	3.57	4.03	4.77
Shark River	4.46	4.78	5.14	5.38	5.88	6.10

4.2.3 SICS / TIME

The USGS has developed the Southern Inland and Coastal System (SICS) model (Swain et al 2004) and the Tides and Inflows in the Mangrove Ecotone (TIME) model (Langevin et al 2002). SICS has a smaller-domain and different grid-cell size than TIME, and there are other code differences. SICS and TIME adapt the USGS SWIFT2D two-dimensional hydrodynamic surface-water model coupled with SEAWAT, a three-dimensional ground-water model, to estimate freshwater flow and solute transport (including salinity) in the southern Everglades. The USGS developed a coupling model (FTLOADDS) to connect the two models.

The SICS model domain encompasses the Taylor Slough area and northeastern-most part of Florida Bay with a 305-m grid resolution (Figure 5). The TIME model has a coarser resolution (500 m) than SICS, but covers a larger area, including Shark and Taylor Sloughs, the Gulf of Mexico, and northern-most part of Florida Bay. Both models use the Flow and Transport in a Linked Overland/Aquifer Density Dependent System (FTLOADDS) computer code to couple surface water flow, groundwater flow, and solute transport. Both models produce flows, stages, and salinities in the wetlands and underlying aquifer system. The SICS and TIME simulations have been produced primarily for the 1996 through 2002 period.

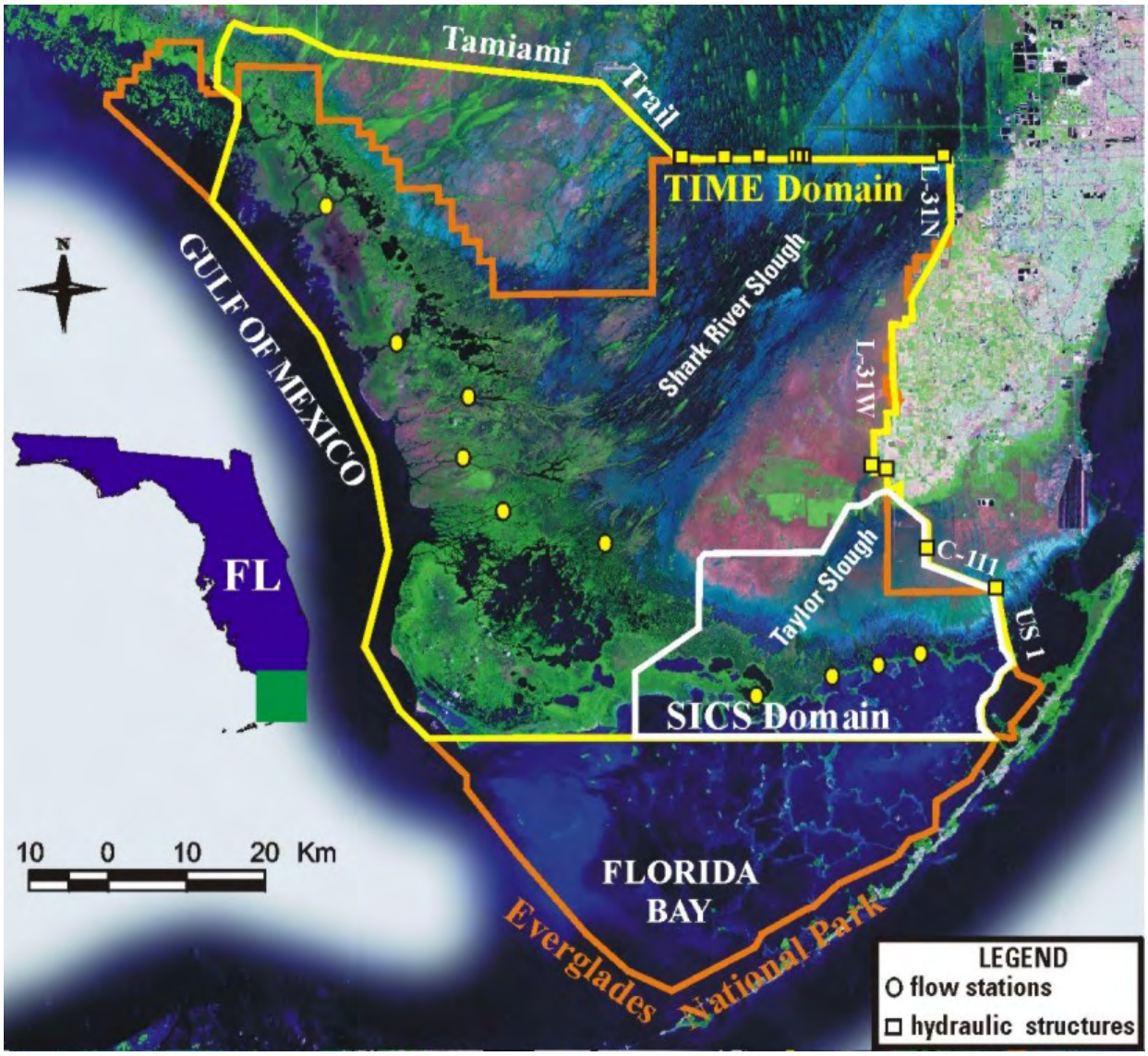
The SICS model can be driven by the SFWMM through the use of SFWMM stage values for SICS boundary conditions which allows for accurate prediction of freshwater flows to Florida Bay under restoration conditions. The SICS model has also produced salinities for use with the ATLSS models (Across Tropic Level System Simulation) to assess restoration effects on fish populations (Langevin et al, 2004a; Cline et al, 2006).

The SICS model has also been used for making daily salinity simulations near the coastal creeks that are being monitored by the USGS. Calibration statistics are presented in Table 3 from Langevin, et al 2004b. The SICS model was found to be better at simulating monthly salinity values ($r^2 = 0.76$) than daily salinity values ($r^2 = 0.67$) at Trout Creek.

Table 5. Error statistics for salinity for the calibration run of the USGS SICS model of Florida Bay and the adjacent coastal wetland (Langevin, et al 2004b)

Station	Mean Error	Mean Absolute Error	Root Mean Squared Error	N
McCormick Creek	2.76	7.14	9.43	2508
Mud Creek	2.10	3.95	5.08	2421
Trout Creek.	2.33	4.86	6.45	2529
Taylor River	4.95	6.35	7.70	2515
West Highway Creek	-1.43	4.60	5.57	2512

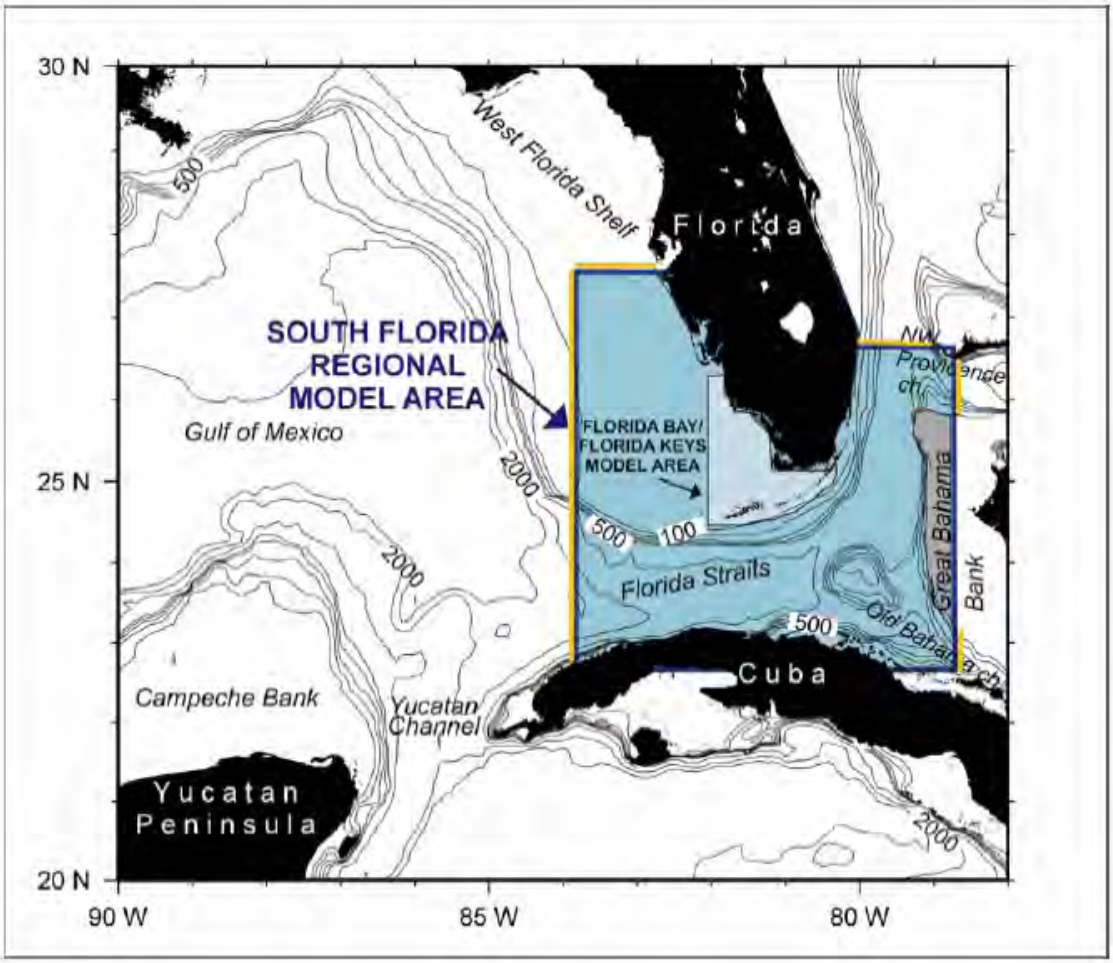
Figure 5: South Florida satellite image showing SICS and TIME model domains (USGS).



4.2.4 SoFLA-HYCOM

SoFLA-HYCOM is a three-dimensional hydrodynamic ocean circulation model for the south Florida coastal system with a domain that includes Florida Bay, the Florida Keys reef tract, and the southwest portion of the Florida shelf as shown by Figure 6. The model was developed to connect the south Florida estuaries and near-shore marine waters to the open-sea areas of the Florida Straits and the Gulf of Mexico. The South Florida (SoFLA) adaptation of the Hybrid Coordinate Ocean Model (HYCOM) simulates the complex circulation patterns of the seas in this region, including the interaction of coastal and offshore effects. SoFLA-HYCOM is coupled with larger scale models of the North Atlantic through nesting. The model is capable of resolving low salinity waters from remote sources, the prevailing Florida Current, the wind-driven southwestward flow along the Florida Keys, eddies that have been observed between the Florida Current and the Keys reef tract, and freshwater flows from rivers into Florida Bay and the Gulf coast estuaries (Kourafalou, 2005).

Figure 6. Map showing the domain of the SoFLA-HYCOM model (Kourafalou, 2005)



5 SUMMARY AND DISCUSSION

Considerable progress has been made in the development and refinement of salinity models since the report in 2002 by the Cadmus Group (Nuttle, 2002). The information presented herein reports on the current status of models that have been used to simulate the salinity in Florida Bay and in the backwaters of the southwest Gulf coast. The two primary salinity data sets used for modeling are:

(1) the SERC/FIU long-term monthly grab sample data set, and
(2) the ENP marine monitoring data set with observations at 10 to 60 minute intervals.

Other hydrologic and climate data sets are used as they are needed for model development and for model input for simulations.

The Everglades / Florida Bay hydrologic system is unique because of the vast area of freshwater marshes underlain by porous substrate that stores runoff before it enters the estuarine zone, as well as the spatial extent of estuarine conditions in Florida Bay. Standard riverine hydraulic models can not account accurately for the spatial and temporal variation in stored water and dispersed flows in the Everglades. Therefore, freshwater hydrology and wetland basin models have been developed to simulate the south Florida conditions required for use with salinity models. For statistical models, hydrology of the Everglades is described by the stage levels that are used as input for salinity simulating and forecasting.

The use of modeled input data for salinity simulations by mechanistic models is necessary because the standard period for evaluations of water management alternatives spans a 36-year period and observed data for some model input are not fully available. This increases the level of uncertainty in the salinity estimates produced by hydrodynamic models. The use of a 36-year period for south Florida simulations is warranted by the significant difference in wet and dry periods over years to decades, and the ecological implications of anthropogenic alterations that may only be expressed over longer periods of change in the salinity regime.

For salinity, the following models were presented and discussed:

1. Multivariate linear regression (MLR) models,
2. A four-box Florida Bay mass balance model;
3. FATHOM, a 41-basin dynamic mass balance models of Florida Bay;
4. RMA-10, a full three-dimensional hydrodynamic model of Florida Bay;
5. EFDC, a full three-dimensional hydrodynamic model of Florida Bay;
6. SICS/TIME, an integrated ground and surface water models that simulates hydrology in the Everglades and salinities in the near shore embayments of Florida Bay and the estuaries on the Gulf coast, and

7. SoFLA-HYCOM, a three-dimensional ocean circulation model that simulates circulation and salinity on the Gulf shelf, in the Florida Straits, and on the Keys reef tract.

A summary of general information about each of these models, except RMA-10, is presented in a Table 6. The RMA-10 model is no longer in use.

The following salinity model evaluation factors from the Florida Bay Science Program (PMC, 2004) were presented previously:

1. portability,
2. validity,
3. fidelity,
4. focus, and
5. ease of use.

Each of the models that were assigned a score for acheiving the desired result of each modeling factor (Table 7). The scale of scoring is from 1 = poor to 5 = excellent. For some models it was not possible to provide a score for a particular factor. From this summary it can be seen that the most complex models are the least portable and are rated lowest for ease of use. With respect to validity, all models rated high because the models are well-documented. Models that simulated salinity were rated highest for focus because salinity performance measures are the use for most of the salinity models. Finally, for model fidelity, daily MLR salinity models have the best performing error measures, followed by SICS / TIME, and monthly FATHOM MFL base case models.

To-date, the most widely used models for developing historical recreations and simulating salinity regimes for the evaluations of water management alternatives are the FATHOM mass balance model and the MLR salinity models. Because of their relative simplicity, development has occurred before full hydrodynamic model development has been completed. Mass balance and MLR salinity models have already been applied in a number of ways and are still being refined. However, the development of hydrodynamic models, particularly SICS/TIME and EFDC, is continuing, and use for historical salinity reconstructions and simulations is likely within the next several years.

By design hydrodynamic models are intended for detailed and spatially discrete applications because of the effort and cost to calibrate, validate, and run large-scale hydrodynamic models for regional scenarios. Statistical and mass balance models will likely remain in use for planning-level decisions on a regional basis. Where possible, it appears that it will be less-expensive and time-consuming to utilize both statistical and mass balance models together as multiple lines of evidence and corroboration compared to utilizing only one hydrodynamic model for regional evaluations.

A comparison was made of observed salinity data and forecasts made by MLR and FATHOM salinity models by plotting the following data for Long Sound and Whipray Basin for the period April, 1994 through October 2002 (Figures 7 and 8):

1. MMN observations averaged to monthly,
2. SERC monthly grab sample observations,
3. FATHOM monthly average estimates from SFWMD MFL work, and
4. MLR daily estimates averaged to monthly values.

It can be seen that the MMN monthly average and SERC grab sample observations correspond well, with fewer deviations at Long Sound than at Whipray Basin. It is important to note that the sampling locations for these two programs in these water bodies are not the same. It can also be seen that both FATHOM and MLR salinity models simulate monthly average salinity in both basins well. The MLR models appear to perform slightly better for Long Sound compared to FATHOM, and noticeably better at Whipray Basin, though the difference in the simulations by the two modeling procedures is small. These plots indicate that both the MMN and SERC data sets can be used interchangeably at the monthly level. It also shows that both FATHOM and MLR salinity models are capable of providing reasonable estimates of salinity at these stations.

6 FINDINGS

Based on this review of the current status (June, 2006) of the models available for simulating and forecasting salinity in Florida Bay, Whitewater Bay, and the Gulf coast estuaries, it is found that MLR salinity models, FATHOM, and the SICS / TIME models appear to be providing the most reasonable estimates of salinity at the time of this report, with corroborating results for salinity variation at the limited locations that were evaluated. In addition, these three models, and the EFDC model if model fidelity can be improved, meet most of the salinity modeling goals of the PMC (2004).

Table 6. Summary of Salinity Models and Supporting Hydrologic Models Currently in Use For Simulating Florida Bay and Southwest Gulf Coast Salinity

Model Name	Model Type	Simulated Parameters	Spatial Domain	Grid Size	Simulation temporal domain
SFWMM[1]	Freshwater Hydrology	Stage, Flow	Everglades	3.2km X 3.2km	1965-2000, daily
PHAST[2]	Wetland Basin	Flow	Lower Everglades and Mangrove Zone	regional	1965-2000, monthly
MLR[3]	Statistical	Salinity	Florida Bay, Whitewater Bay, southwest Gulf coast, Manatee Bay, Barnes Sound	N/A	1965-2000, daily
Four Box[4]	Mass Balance	Salinity	Florida Bay	regional	1993-1998, monthly
FATHOM[5]	Mass Balance	Salinity	Florida Bay, Manatee Bay, Barnes Sound	open-water basins	1965-2000, monthly
EFDC[6]	3-D Hydrodynamic	Salinity	Florida Bay, Whitewater Bay, southwest Gulf coast, Manatee Bay, Barnes Sound	variable	1965-2000, daily
SICS/TIME[7]	2D/3D Coupled surface and groundwater	Stage, Flow, Salinity	Florida Bay (SICS), southwest Gulf coast (TIME)	0.3km X 0.3km (SICS, 0.5km X 0.5km (TIME)	1996-2000, daily
SoFLA-HYCOM[8]	3-D Hydrodynamic ocean circulation model	Flow magnitude and direction	Gulf of Mexico, Florida Straits	6-7km X 6-7km	?

[1] http://www.sfwmd.gov/org/pld/hsm/models/sfwmm/index.html
[2] Nuttle and Teed 2002, Nuttle 2004
[3] Marshall, 2005
[4] Nuttle et al. (2000)
[5] Cosby et al. 1999, Nuttle et al. 2000, Cosby et al 2004
[6] Hamrick and Moustafa, 2003
[7] Swain, et al 2004 (SICS), Langevin, et al 2002 (TIME)
[8] Kourafalou, 2005

Table 7. Summary evaluation of Florida Bay salinity and hydrology models using the Florida Bay Science Program evaluation factors (PMC 2000). Models with asterisk (*) are freshwater hydrology only models. Score is from 1=lowest to 5=highest.

Model	Portability	Validity	Fidelity	Focus	Ease of Use
SFWMM*	2	5	3	4	3
PHAST*	3	4	3	3	5
MLR	5	5	5	5	5
Four Box*	3	4	4	3	5
FATHOM	3	5	4	5	4
EFDC	2	5	3	5	3
SICS/TIME	2	5	4	5	3
SoFLA-HYCOM	2	3	?	3	3

Figure 7. A comparison of observed salinity (SERC), monthly average salinity (MMN), and simulations by FATHOM and MLR salinity models (monthly average) at Long Sound.

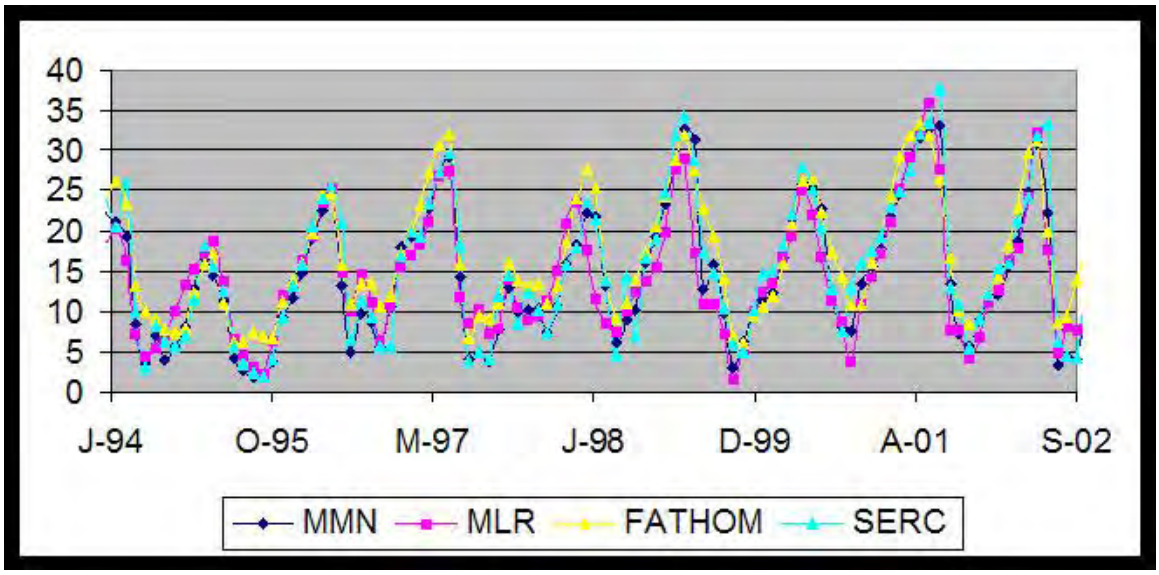
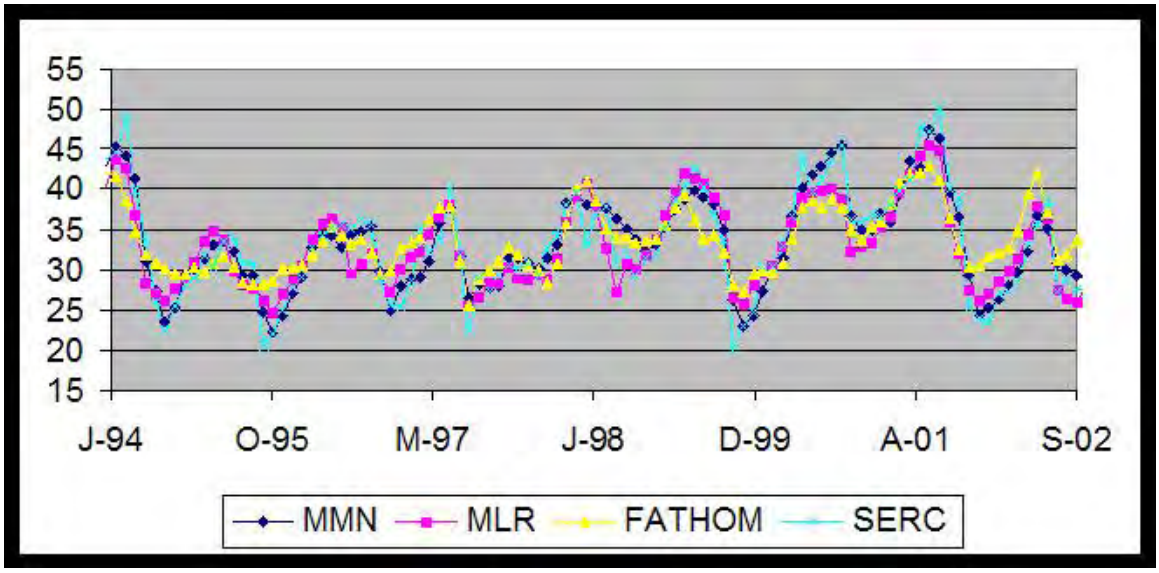


Figure 8. A comparison of observed salinity (SERC), monthly average salinity (MMN), and simulations by FATHOM and MLR salinity models (monthly average) at Whipray Basin.



7 REFERENCES

Abtew, W., J. Obeysekara, M. Irizarry-Ortiz, D. Lyons, A. Reardon. 2003. Evaporation Estimation for South Florida. Technical Paper EMA # 47, Proceedings of the World Water and Environmental Resources Congress 2003.

Boyer, J. N., J. W. Fourqurean, and R. D. Jones. 1997. Spatial characterization of water quality in Florida Bay and Whitewater Bay by multivariate analyses: zones of similar influence. Estuaries, 20:743-758.

Boyer, J. N. and R. D. Jones. 2001. Long-term trends in water quality of Florida Bay (June 1989 - June 1999). Abstract, Florida Bay Science Conference, Key Largo, FL, April 24-26, 2001.

Cline, J.C. and E. D. Swain. 2002. Coupling hydrologic and ecological modeling: SICS and ATLSS. Second Federal Interagency Hydrologic Modeling Conference, July 28-August 1, 2002, 12 pp.

Cerco, C.F., B.W. Bunch, A.M. Teeter, and M.S. Dortch. 2000. Water Quality Model of Florida Bay. Environmental Laboratory, ERDC/EL TR-00-10. U.S. Army Corps of Engineers. Vicksburg, Mississippi.

CERP, 2002. Model Uncertainty Workshop Report. Comprehensive Everglades Restoration Plan. Available from: <http://www.evergladesplan.org/pm/recover/recover_docs/mrt/052402_mrt_uncertainty_report.pdf> [cited 20 May 2005].

CERP, 2004. Interangency Modeling Center Program Management Plan. Comprehensive Everglades Restoration Plan. Available from: <http://www.evergladesplan.org/pm/progr_imc_plan.cfm> [cited 20 May 2005].

Cline J. C., J. J. Lorenz and E. D. Swain. 2006. Linking Hydrologic Modeling and Ecologic Modeling: Application of a Spatially-Explicit Species Index (SESI) Model for Adaptive Ecosystem Management in the Everglades Mangrove Zone of Florida Bay. Abstract for 2006 Greater Everglades Ecosystem Restoration Conference, Lark Buena Vista, Florida, June 5-9, 2006.

Corbett, D. R., J. Chanton, W. Burnett, K. Dillon, C. Rutkowski, and J. W. Fourqurean. 1999. Patterns of groundwater discharge into Florida Bay. Limnol. Oceanogr. 44:1045-1055.

Cosby, B.J. 1993. An Examination of the Relationships of Stage, Discharges and Meteorology in the Panhandle and Taylor Slough Areas of Everglades National Park to Salinity in Upper Florida Bay, Volumes 1-5. University of Virginia, Charlottesville, Virginia.

Cosby, B.J., W.K. Nuttle, and J.W. Fourqurean, 1999. FATHOM: Model Description and Initial Application to Florida Bay. (progress report submitted to Everglades National Park)

Cosby, B., W. Nuttle, and F. Marshall, 2005. FATHOM Enhancements and Implementation to Support Development of MFL for Florida Bay. Final Report on Contract C-C-15975-WO05-05 for the South Florida Water Management District. Environmental Consulting & Technology, Inc. New Smyrna Beach, Florida.

CROGEE. 2002. Florida Bay Research Programs and Their Relation to the Comprehensive Everglades Restoration Plan. The National Academies Press, Washington, D.C.

DeAngelis, D. L., S. Bellmund, W. M. Mooij, M. P. Nott, E. J. Comiskey, L. J. Gross, and W. M. Wolff. 2002. Modeling ecosystem and population dynamics on the South Florida hydroscape, p. 239-258. In J. W. Porter and K. G. Porter (eds.), The Everglades, Florida Bay and Coral Reefs of the Florida Keys: An Ecosystem Sourcebook, CRC Press, Delray Beach, FL.

Enfield, D.B., A.M. Mestas-Nunez, P.J. Trimble. 2001. The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S. Geophysical Research Letters, Vol. 28, No.10. pp. 2077-2080

Everglades National Park. 1997a. Everglades National Park Marine Monitoring Network 1994 Data Summary. Everglades National Park, Homestead, Florida. 67 pp.

Everglades National Park. 1997b. Everglades National Park Marine Monitoring Network 1995 Data Summary. Everglades National Park, Homestead, Florida. 67 pp.

German, E.R., 2000. Regional Evaluation of Evapotranspiration in the Everglades. Water-Resources Investigations Report 00-4217. U.S. Geological Survey, Tallahassee, FL.

Hamrick, J.H. and M.Z. Moustafa. 2003. Florida Bay Hydrodynamic and Salinity Model Analysis. Conference abstract from Joint Conference on the Science and Restoration of the Greater Everglades and Florida Bay Ecosystem.

Hittle, C., E. Patino, and M. Zucker, 2001. Freshwater Flow from Estuarine Creeks into Northeastern Florida Bay. Water-Resources Investigations Reports 01-4164, US Geological Survey, Tallahassee, Florida.

Hunt, J. and W.K. Nuttle (eds). Florida Bay Science Program. A Synthesis of Research on Florida Bay. Florida Marine Research Institute (in prep.).

Johns, E., P. Ortner, R. Smith, D. Wilson, T. N. Lee, and E. Williams. 2001. Salinity Variability in Florida Bay from Monthly Rapid High Resolution Surveys. Abstract, Florida Bay Science Conference, Key Largo, FL, April 24-26, 2001.

Jones, R.D., and J.N. Boyer, 2001. An Integrated Surface Water Quality Monitoring Program For The South Florida Coastal Waters: Fy2000 Cumulative Report to the South Florida Water Management District (C-10244) and Everglades National Park. Southeast Environmental Research Center, Florida International University, Miami, FL 33199. (http://serc.fiu.edu/wqmnetwork/)

Kourafalou, Villy. 2005. SoFLA-HYCOM (South Florida HYCOM) Regional Model around Florida Straits, Florida Bay and the Florida Keys: An overview. Rosenstiel School of Marine and Atmospheric Science, University of Miami Division of Meteorology and Physical Oceanography, Miami (http://hycom.rsmas.miami.edu/overview/SoFLA_HYCOM.pdf

Langevin, C.D., E. D. Swain, and M. A. Wolfert. 2002. Numerical simulation of integrated surface-water/groundwater flow and solute transport in the southern Everglades, Florida, Second Federal Interagency Hydrologic Modeling Conference, July 28-August 1, 2002, 12 p.

Langevin, C.D. E.D. Swain, J.D. Wang, M.A. Wolfert, R.W. Schaffranek, and A.L. Riscassi. 2004a. Development of coastal flow and transport models in support of everglades restoration. Fact Sheet 2004-3130. U.S. Department of the Interior, U.S. Geological Survey. October 2004.

Langevin, C.D. E.D. Swain, and M.A. Wolfert. 2004a. Simulation of integrated surface-water / groundwater flow and salinity for a coastal wetland and adjacent estuary. Open-File Report 2004-1097. U.S. Department of the Interior, U.S. Geological Survey. October 2004.

Lee, T. N., and E. Williams. 1999. Mean distribution and seasonal variability of coastal currents and temperature in the Florida Keys with implications for larval recruitment. Bulletin of Marine Science 64:35-56.

Lee, T. N., E. Johns, D. Wilson, and E. Williams. 1999. Site characterization for the Tortugas region: Physical Oceanography and Recruitment. Tortugas 2000: Report by the Florida Keys National Marine Sanctuary.

Lettenmaier, D.P. and E.F. Wood, 1993. Chapter 26: Hydrologic Forecasting. in D.R. Maidment (ed): Handbook of Hydrology. McGraw Hill, New York.

Lorenz, J. 2005. Habitat Suitability Index for Roseate Spoonbills Nesting in Northeastern Florida Bay. page 43 in Program and Abstract Book, 2005 Florida Bay and Adjacent Marine Systems Science Conference, December 2005, Duck Key, Florida.

Marshall, F.E. 2000. Florida Bay Salinity Transfer Function Analysis Volume 1 of 2: Final Report. Cetacean Logic Foundation, New Smyrna Beach, Florida. 24 pp.

Marshall III, F. E.; D. Smith; and D. Nickerson. 2004a. Using Statistical Models to Simulate Salinity Variation and Other Physical Parameters in North Florida Bay. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F.E. 2005. ICU Runs Summary Report, for Southern Estuaries Sub-team of RECOVER. Environmental Consulting & Technology, Inc. New Smyrna Beach, Florida.

Nash, J.E., and J.V. Sutcliffe, 1970. River flow forecasting through conceptual models: Part 1 – a discussion of principles. Journal of Hydrology 10:282-290.

NRC, 2002a. Risk Analysis and Uncertainty in Flood Damage Reduction Studies. Committee on Risk-Based Analysis for Flood Damage Reduction, Water Science and Technology Board, Commission on Geosciences, Environment, and Resources, National Research Council, National Academy Press, Washington, D.C.

NRC, 2002b. A Review of the Keys Carrying Capacity Study. Report of the Committee to Review the Florida Keys Carrying Capacity Study, Ocean Studies Board, Water Science and Technology Board, National Research Council, Washington, DC.

NRC, 2004. Adaptive Management for Water Resources Project Planning. National Academies Press, Washington, D.C. Available from: <http://www.nap.edu/catalog/10972.html> [cited 20 May 2005].

Nuttle, W.K., J.W. Fourqurean, B.J. Cosby, J.C. Zieman, and M.B. Robblee, 2000. The influence of net freshwater supply on salinity in Florida Bay. Water Resources Research 36:1805-1822.

Nuttle, W.K., R. Price, and P.K. Swart, 2001. Estimating evaporation from Florida Bay. pg 102 in Conference Abstracts, 16th Biennial Conference of the Estuarine Research Federation, November 2001, St. Pete Beach, Florida.

Nuttle, W.K. and R.S. Teed. 2002. Version 1: wetland hydrology and estuarine salinity models for the Taylor Slough/C111 area. Report prepared for the Everglades National Park. The Cadmus Group, Inc. Watertown, MA 02472.

Nuttle, W. K., December 1997. Central and Southern Florida Project Restudy: Salinity Transfer Functions for Florida Bay and West Coast Estuaries, Volumes 1 and 2. Southeast Environmental Research Program, Florida International University, Miami, Florida.

Nuttle, W.K. 2002. Review and Evaluation of Hydrologic Modeling Tools for the Coastal Mangroves and Florida Bay. Project report for Everglades National Park, April 2002. Available from: <http://www.eco-hydrology.com/wkn%20enp%201.pdf> [cited 20 May 2005].

Nuttle, W.K. 2005. Biscayne Bay Coastal Wetland Project Planning Tool Phase I: Hydrology and Salinity Calculations. Project report for Everglades National Park, January 2005. Available from: <http://www.eco-hydrology.com/wkn%20bisc%20tool%20fnl%20mod.pdf> [cited 20 May 2005].

Nuttle, W.K. and F.E. Marshall, 2005. Effect of time scale on patterns and processes of salinity variation in Florida Bay. Estuaries (in review).

Orlando, S.P., Jr.; M.B. Robblee; C. J. Klein. 1998. Salinity Characteristics of Florida Bay: A Review of Archived Data Set (1955-1995). National Oceanographic and Atmospheric Administration, Office of Ocean Resources Conservation and Assessment, Silver Spring Maryland. 32 pp.

Ortner, P. B., T. N. Lee, P. J. Milne, R. G. Zika, E. Clarke, G. Podesta, P. K. Swart, P. A. Tester, L. P. Atkinson, and W. R. Johnson. 1995. Mississippi River flood waters that reached the Gulf Stream. Journal of Geophysical Research 100:13595-13601.

Pandit, A.; C. El-Khazen; S.P. Sivaramapilli. 1991. Estimation of Hydraulic Conductivity Values in a Coastal Aquifer. Groundwater. Vol. 29, No.2. pp 175-180.

PMC 2000. Florida Bay Models Coordination Meeting. Report of a workshop held in May 2000.

Price, R.M., 2001. Geochemical Determinations of Groundwater Flow in Everglades National Park, University of Miami Dissertation, 307p.

Price, R.M., P.K. Swart, and J.W. Fourqurean. Coastal groundwater discharge – an additional source of phosphorus for the oligotrophic wetlands of the Everglades. Hydrobiologia, in press.

Robblee, M. B., D. Smith, R. Halley, and G. Clement. 2001. Temporal and Spatial Characterization of Salinity Pattern in Florida Bay. Abstract, Florida Bay Science Conference, Key Largo, FL, April 24-26, 2001.

Santee, E.R., L. Cadavid, W. Wilcox, R. Novoa, M. Irzarry-Ortiz, A. Ali, L Brehon, J. Barnes, A Montoya, J. Obetsekera, and K. Tarboton. 2003 Recent Enhancements to the South Florida Water Management Model (SFWMM). South Florida Water Management District, West Palm Beach, Florida.

Schaffranek, R.W., H. L. Jenter, and A. L. Riscassi. 2002. Overview of the “Tides and Inflows in the Mangroves of the Everglades” (TIME) project of the U.S. Geological Survey’s South Florida Ecosystem Program, Second Federal Interagency Hydrologic Modeling Conference, July 28-August 1, 2002.

Scully, S.P. 1986. Florida Bay Salinity Concentration and Groundwater Stage Correlation and Regression. South Florida Water Management District, West Palm Beach, Florida.

Sheng, Y. P., J. R. Davis, and Y. Liu. 1995. A Preliminary Model of Florida Bay Circulation. Final Report to the National Park Service. University of Florida, Gainesville, Florida.

Smith, D. 1997. Everglades National Park Marine Monitoring Network 1996 Data Summary. Everglades National Park, Homestead, Florida. 94 pp.

Smith, D. 1998. Everglades National Park Marine Monitoring Network 1997 Data Summary. Everglades National Park, Homestead, Florida. 100 pp.

Smith, D. 1999. Everglades National Park Marine Monitoring Network 1998 Data Summary. Everglades National Park, Homestead, Florida. 100 pp.

Smith, D. 1997. Everglades National Park Marine Monitoring Network 1996 Data Summary. Everglades National Park, Homestead, Florida. 94 pp.

Smith, D. 2001. Everglades National Park Marine Monitoring Network 1999 Data Summary. Everglades National Park, Homestead, Florida. 65 pp.

Sutula, M., J.W. Day, J. Cable, and D. Rudnick, 2001, Hydrological and nutrient budgets of freshwater and estuarine wetlands of Taylor Slough in Southern Everglades, Florida (U.S.A), Biogeochemistry, 56, pp. 287-310.

Smith, R. H., E. Johns, D. Wilson, T. N. Lee, and E. Williams. 2001. Moored Observations of Salinity Variability in Florida Bay and South Florida Coastal Waters on Daily to Interannual Time Scales. Abstract, Florida Bay Science Conference, Key Largo, FL, April 24-26, 2001.

Swain, E.D., Wolfert, M.A., Bales, J.D., and Goodwin, C.R., 2004, Two-dimensional hydrodynamic simulation of surface-water flow and transport to Florida Bay through the Southern Inland and Coastal Systems (SICS): U.S. Geological Survey Water-Resources Investigations Report 03-4287.

Tabb, D. 1967. Prediction of estuarine Salinities in Everglades National Park, Florida by the use of ground water records. PhD Dissertation, University of Miami, Miami, Florida. 107 pp.

Tetra Tech, Inc. 2005. Development of a Florida Bay and Florida Keys hydrodynamic and water quality model: hydrodynamic model calibration, interim status report. Tetra Tech, Inc. Fairfax, VA.0

Thom, R.M., H.L. Diefenderfer, and K.D. Hofseth, 2004. A Framework for Risk Analysis in Ecological Restoration Projects. IWR Report 04-R-02.
Twilley, R. [n.d.] Evaluating the Response of Mangrove Wetlands to Alterations of Freshwater Flow in Everglades National Park. [online] Available from: <http://www.ucs.louisiana.edu/~rrt4630/mangrove-restudy.htm> [cited 20 May 2005].

Top, Z., Brand, L. W., R.D. Corbett, W. Burnett, and J. Chanton, 2001, Helium and radon as tracers of groundwater input into Florida Bay, J. Coastal Research, 17, pp. 859-868.

Walker, W.W., 1998. Estimation of Inputs to Florida Bay. Report prepared for U.S. Army Corps of Engineers & U.S. Department of the Interior. Available from: http://wwwalker.net/flabay/ [cited 20 May 2005].

Wang, J.D., J. van de Kreeke, N. Krishnan, and D. Smith, 1994. Wind and Tide Response in Florida Bay. Bull. Marine Science, 54(3), 579-601.

Wang, J.D., 1998. Subtidal Flow Patterns in Western Florida Bay. Estuarine, Coastal and Shelf Science, 46(6), 901-915.

Weglarczyk, S. 1998. The interdependence and applicability of some statistical quality measures for hydrological models. Journal of Hydrology 206:98-103.

Wong, T.S.W., 2002. Discussion: Use of artificial flood events to demonstrate the invalidity of simple mixing models. Hydrological Sciences Journal 47:833-837.

FINAL TASK REPORT

EVALUATION OF METHODS FOR FILLING GAPS IN FLORIDA BAY SALINITY, STAGE, WATER TEMPERATURE, AND FLOW DATA

Task No. 8

Cooperative Agreement Number CA H5284-05-0006
Between
The United States Department of the Interior National Park Service
Everglades National Park
And
Cetacean Logic Foundation, Inc.

Frank E. Marshall, III
Principal Investigator

DeWitt Smith
Project Coordinator

February 21, 2008

CESI Project No. CA H5284-05-0006 for Everglades National Park

FINAL REPORT
TASK NO. 8

Evaluation of Methodologies for Filling Gaps in Florida Bay Salinity, Stage, Water Temperature, and Flow Data

I. Introduction

A. General

This report describes the activities that were completed by Cetacean Logic Foundation, Inc. (CLF) for Task No. 8 in the Scope of Work for CA H5284-05-0006, a Critical Ecosystem Studies Initiative (CESI) project for Everglades National Park (ENP). The current CESI project was originally initiated by CLF in 2002. The goal of this on-going CESI research is the characterization and simulation of the salinity regime in Florida Bay and the estuaries along southwest coast of the Gulf of Mexico (Gulf) within the Park.

This Task No. 8 report presents the methods and findings of an evaluation of the use of statistical models and other techniques to fill gaps in the time series data record for salinity, stage, flow and temperature. The methods evaluated include seasonal autoregressive integrated moving average (SARIMA) models, simple (univariate) linear regression models, and multivariate linear regression models.

B. Task Objectives and Evaluation Methods

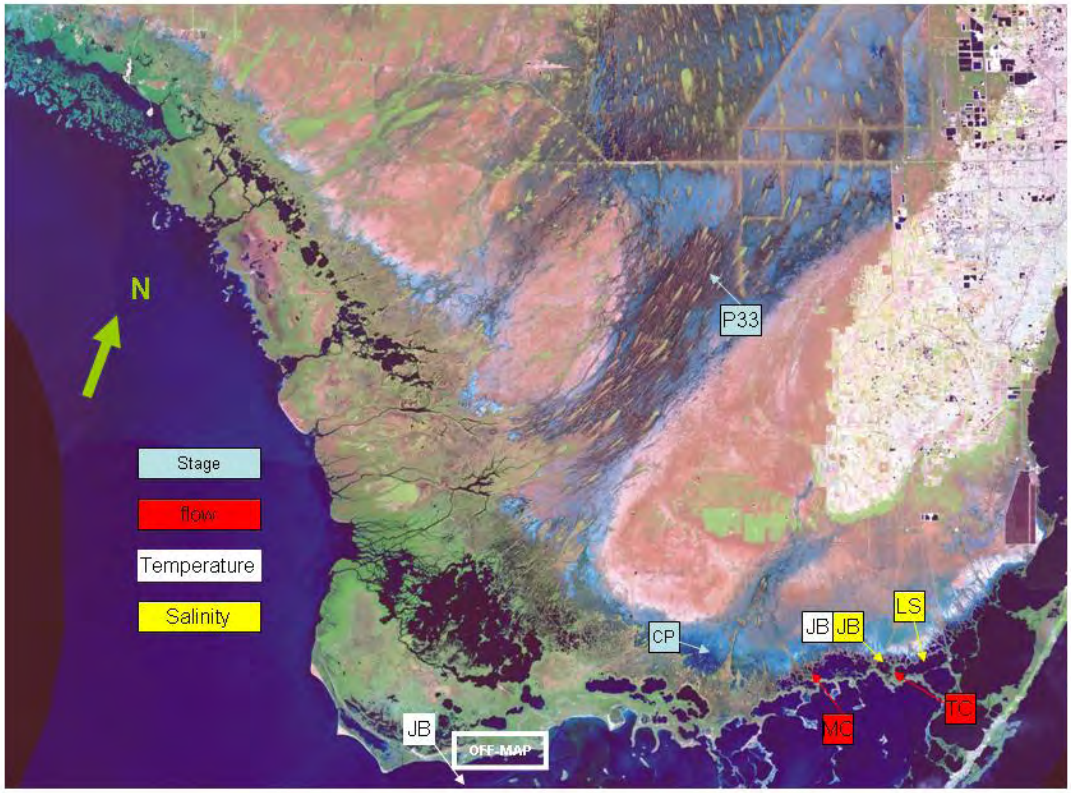
The objective of this task is presented below as it appears in the modified contract for this CESI project approved by ENP:

> "In the Year One CESI project it was determined that MLR models would be more useful than SARIMA models for coupling with 2X2 model output for salinity simulations. Even so, SARIMA models were identified as useful for filling in gaps in physical data. Therefore this task will develop a SARIMA modeling tool that can be used to fill data gaps from "nearest neighbor" MMN data.
>
> The deliverable is a task report on the development of a SARIMA modeling tool to fill data gaps in physical data being collected by the MMN program."

II. Study Area

The study area for this project includes the Greater Everglades freshwater marsh and the estuarine and marine areas of the Florida Bay region in south Florida within Everglades National Park. Various hydrologic monitoring stations within the study area were selected for the evaluation of models to fill gaps in the data record. The location of the monitoring stations depends somewhat on the parameter that was examined. For example, stage and flow data were collected in the freshwater marsh sloughs of the Greater Everglades area, and salinity and water temperature data were collected in Florida Bay. Figure 1 presents the study area and the locations of each monitoring station used in this evaluation.

Figure 1. Map of study area showing location of monitoring stations.



III. Data

To evaluate the use of time series models for filling data gaps, two monitoring stations with long periods of record were chosen as dependent variables for each parameter category: salinity, stage, water temperature, and freshwater flow. The stations chosen for the evaluation are:

1. Stage – Craighead Pond (CP) andP33
2. Salinity – Joe Bay (JB) and Long Sound (LS)
3. Water Temperature – Joe Bay and Johnson Key (JK)
4. Freshwater Flow – USGS stations Mud Creek (MUD) and Trout Creek (TRO).

Basic information on the stations are presented in Table 1 including station location, period of record, and responsible party for the data, station locations are in Figure 1. These stations were selected in part because of their proven utility for performance measure evaluations as part of the Comprehensive Everglades Restoration Plan (CERP) project alternative analysis.

Because of the length of record, the harsh environment of the study area, and problems inherent in remote monitoring of environmental data, all of the data used contain missing values at various levels. The percentage of missing values varies from .05% to over 17%.

One important characteristic of hydrologic data is that the data are usually serially correlated (i.e. salinity today is related to salinity yesterday) and also cross-correlated (i.e. Whipray Basin salinity is related to Joe Bay salinity). Because of this it can be said that hydrologic data are spatially and temporally related. This characteristic can be used to create statistical models using regression techniques to simulate time series data, and for use filling data gaps.

Table 1. Basic information on monitoring data used in the analysis of methods for filling gaps in hydrologic data.

Variable Name	Variable Type	Raw Data Units	Data Source	Location*	Beginning of POR
CP	Water Level	ft NGVD 29	ENP	Craighead Pond	10/01/78
P33	Water Level	ft NGVD 29	ENP	Shark River Slough	2/15/53
Mud Creek (MUD)	flow	cfs	USGS	SE Mangrove Zone	1/01/96
Trout Creek (TRO)	flow	cfs	USGS	SE Mangrove Zone	2/01/96
Joe Bay (JB)	Temperature	^{o}C	ENP	NE Florida Bay	3/24/94
Johnson Key (JK), top	Temperature	^{o}C	ENP	Western Florida Bay	3/24/94
Johnson Key (JK), bottom	Temperature	^{o}C	ENP	Western Florida Bay	3/24/94
Joe Bay (JB)	Salinity	psu	ENP	NE Florida Bay	7/14/88
Long Sound (LS)	Salinity	psu	ENP	NE Florida Bay	7/14/88

* SE= southeast, NE=northeast

IV. Models Utilized

Seasonal Autoregressive Integrated Moving Average (SARIMA) Models

Previous studies for ENP indicated that time series techniques may be useful for simulating salinity variability (Marshall, 2003a; 2003b). The types of models that take advantage of the serial correlation typical in physical data such as stage, flow, salinity, temperature, etc. include autoregressive, moving average (ARMA) models; autoregressive, integrated, moving average (ARIMA) models; and seasonal autoregressive, integrated, moving average (SARIMA) models. In the most sophisticated form of SARIMA models, the dynamic capability of one-step forward prediction is based on both serial and cross correlation, seasonal correlation, as well as correlation with errors (residuals). The SAS© model development procedure for SARIMA models provides the information necessary to choose the appropriate independent variables (lagged values of the subject independent variable) and allows the model developer to specify the method for applying the correlation effect, i.e. as a decay function or as a single pulse application.

The SARIMA analysis began by “pre-whitening” each series to explain autocorrelations and distill the series into one that is as close as possible to a white noise (random) series. Pre-whitening accentuates the cross-correlation relationship (or lack thereof) between the dependent variables and the independent variable (Brockwell and Davis, 1996).

This cross-correlation relationship was evaluated by examining a plot of the cross-correlation coefficient between a subject dependent variable and lagged values of each of the independent variables. Values of independent variables with cross-correlation coefficients that were greater than 2 standard deviations from the mean were then used for a conditional least squares evaluation to determine which independent variables should remain in the final model. A 95% significance level was used to keep or discard independent variables from the tentative model using the Student –T statistic.

The independent variable make-up of the “best” model was determined through the use of the Akiake Information Criterion (AIC). This criterion provides a credit when an independent variable explains some of the error in the data and a penalty when each variable is added to the model. The inclusion of variables with limited significance in a model tends to increase the value of the AIC for a particular model. The lowest AIC value indicates the most parsimonious model (i.e. a lower value of AIC means a better model) (Brockwell and Davis, 1996). The standard error of the estimate is also an important indicator of the ability of a SARIMA model to accurately predict or estimate salinity. The standard error of the estimate is equal to the standard deviation of the observed values about the

predicted values for the period of the model development. A smaller standard error is indicative of a more accurately estimating model.

The use of the SARIMA models for one-step forward predictive purposes requires that the previous observed values of the dependent variable for lagged periods be available for use in the model. When they are not, the prediction error becomes large after only several time periods, particularly when using an external dataset.

Multivariate Linear Regression (MLR) Models

Multivariate linear regression models (MLR) are linear combination models that utilize several independent variables to explain the variability of a dependent variable. MLR models are developed from unlagged and lagged values of other independent variables representing stations in the immediate surrounding area of the stations represented by the dependent variable.

MLR models may include cross-correlation effects that can sometimes express phenomena that are not physically-based. For the MLR models that were developed for this exercise, model development was controlled to insure that the model made physical sense. For example, for stage models, an increase in an independent stage variable must indicate an increase in a dependent stage variable, otherwise the independent variables were removed from the model. Inappropriate cross-correlation effects were only observed in independent variables that provided less significant explanatory capability, and elimination of them did not materially reduce the ability of the model to explain the variability of the subject dependent variable. Increasing the level of significance also helps to eliminate inappropriate cross-correlation effects.

MLR models were previously used for simulating salinity as a function of water levels (stage) in the Everglades marsh, the water surface elevation Key West, and wind speed and direction (Marshall, 2004; 2005). MLR models for this evaluation utilized independent variables that were of like kind to the dependent variable (i.e. stage was used to model stage) and also wind parameters.

Simple (Univariate) Linear Regression Models

In the case of temperature, univariate linear regression models were capable of explaining a very high level of variation. Therefore, simple regression models were developed for evaluation of filling data gaps. A correlation matrix analysis was used to select independent variables. The independent variable station used for a simple regression model was the station with the highest Pierson correlation coefficient value. A 95% Significance Level was used for development of the simple linear regression models.

Linear Interpolation Techniques

The SAS© time series software has several routines for filling gaps automatically (PROC EXPAND). A method for filling the gaps must be specified, and includes fitting a:

- Straight line (linear interpolation),
- Spline curve using a 3rd order polynomial,
- Straight line spline (successive straight line segments), and
- Discontinuous piece-wise constant curve (step function).

For this exercise, linear interpolation was used to fill data gaps.

Comparison Techniques

In order to evaluate the use of each technique to fill data gaps, a data set was developed for each subject variable by removing data values from an observed time series. Then the models were used to fill the gaps. The ability of the model to accurately simulate the observed data was evaluated by comparing the model data to the observed data. In order to evaluate the effect of various sizes of gaps (periods of time with missing data), several sets of missing data were assembled. Sets of data with gaps of 1 day, 2 days, 3 days, 5 days, 25 days, and six months were developed and model-simulated values were compared to observed values for the evaluation.

V. Results

Initial Results

This task was primarily intended to evaluate the utility of SAS © PROC ARIMA modeling techniques for filling gaps. It was thought that the use of the serial correlation and cross-correlation in data from neighboring stations may yield better results compared to multivariate techniques. However, when SARIMA model development was initiated, it was found that when the pre-whitening of independent variables was done, most of the explanatory capability for simulation of the dependent variable was lost. Pre-whitening of the independent variables leaves a time series that is supposed to be close to a white noise series, i.e. a random series. The SAS © PROC ARIMA procedure utilizes this remaining data to develop a linear model for the dependent variable. It was found that without serial correlation present in the independent variable series the remaining cross-correlation was small. This is primarily caused by the highly autoregressive nature of physical hydrologic data.

If the independent variables are not pre-whitened, the resultant model is a model that is almost identical to a multivariate linear regression model. However, the model has no integrated, autoregressive, or moving average parameters and thus uses none of the information that this category of models is designed to use. Therefore, there is no advantage to using time series (ARMA, ARIMA, or SARIMA) models over multivariate linear regression models.

Further examination of information on SARIMA models shows that SARIMA models are best used in a one-step forward prediction, as in control circuitry where the uncertainty in a prediction is replaced by an observed value for the next prediction in line to be made.

One-step forward modeling techniques, including SARIMA, may have value in operating the water management system in south Florida. For example, if system operational procedures specify a stage value at a monitoring station then a SARIMA model with the previous values from observed data can be used to determine if the current operational mode will produce the desired result, or if a change in operations is needed to increase or decrease the stage.

The emphasis of this task was shifted to the demonstration of multivariate linear regression models for filling data gaps after evaluating the modeling technique for stage and salinity because of the limitations of using time series techniques.

Stage

In order to evaluate models for filling gaps in stage data, two stage stations were selected as dependent variables for model development. The stage stations that were used as dependent variables were Craighead Pond (CP) for Taylor Slough

and P33 for Shark Slough. Both of these monitoring stations have long periods of record, beginning in 1952 for P33 and 1988 for CP. Stage stations in proximity to P33 and CP were selected for use in the models based on length of record and level of correlation, which varied widely.

The best model MLR for CP was developed from E146, NP206, and several wind variables. The greatest proportion of variability was explained by the two stage variables. The CP model R^2 is 0.96. For the ARMA model, the pre-whitened series used for the model included P33, EVER7, E146, EVER6, NP206, and R127.

Thee gaps that were created artificially in the 1999-2000 CP data were filled using ARMA models, linear interpolation (LI), and multivariate linear regression (MLR). Simple statistics on the observed data and the models data are compared in Table 2. There are additional missing values in the ARMA and MLR output due to the presence of missing values in the input data. The mean value for 322-370 values was relatively similar, differing by only -0.06 ft to +0.07 ft. However, there were larger relative differences in the other distribution descriptors.

For short gaps of less than five (5) days, the simulated values were compared to the observed values for the artificially-created gaps (Table 3). The linear interpolation estimates are very close to the observed values. The MLR estimates are also quite close, except for the period in June, 1999. The ARMA estimates are close when the observed data are close to the mean value, but when the observed value deviates from the mean, the ARMA estimates show considerable error.

Figures 2a and 2b present plots of the CP observed and simulated values for longer gaps. The MLR simulations closely track the observed values. If the data are continuously declining, the linear interpolation works well, but if the data change from declining to increasing or vice versa (monotonic) the simulation is not very close over time. None of the ARMA simulations worked very well.

Table 2. Simple statistics for observed and simulated values of CP used for the gap analysis.

	N	Mean	Std Dev	Minimum	Maximum
CP	370	1.37	0.51	0.33	2.74
ARMA	322	1.31	0.07	0.80	1.49
LI	370	1.44	0.35	0.53	2.35
MLR	365	1.33	0.51	0.20	2.68

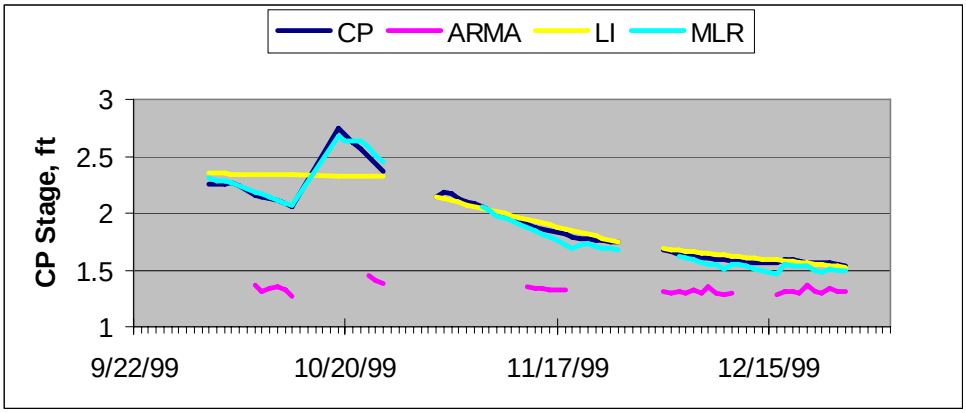


Figure 2a. Comparison of observed and simulated values of CP for 3 artificially-created gaps of about 25 Days each.

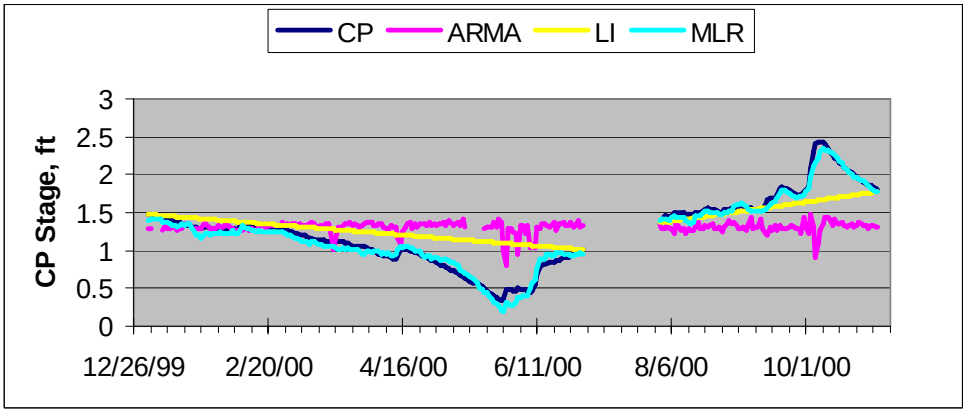


Figure 2b. Comparison of observed and simulated values of CP for 2 artificially-created gaps of 6 and 3 months each, continued.

Table 3. A comparison of simulated values of stage for monitoring station CP to observed values for artificially-created data gaps that were short in duration (<5 days). OBS-SIM is difference between observed and simulated values.

.

Date	Observed CP	ARMA	OBS-SIM	LI	OBS-SIM	MLR	OBS-SIM
2/1/1999	1.29	1.31	-0.02	1.3	-0.01	1.27	0.02
3/1/1999	1.11	1.33	-0.22	1.11	0	1.06	0.05
4/5/1999	0.72	1.28	-0.56	0.72	0	0.75	-0.03
4/6/1999	0.7	1.32	-0.62	0.7	0	0.74	-0.04
4/7/1999	0.68	1.35	-0.67	0.68	0	0.71	-0.03
4/11/1999	0.61	1.36	-0.75	0.61	0	0.62	-0.01
4/12/1999	0.59	1.34	-0.75	0.59	0	0.58	0.01
4/13/1999	0.57	1.36	-0.79	0.57	0	0.55	0.02
4/14/1999	0.55	1.36	-0.81	0.55	0	0.54	0.01
4/15/1999	0.53	1.37	-0.84	0.53	0	0.47	0.06
6/1/1999	0.64	1.28	-0.64	0.64	0	0.81	-0.17
6/2/1999	0.64	1.29	-0.65	0.64	0	0.84	-0.2
6/5/1999	0.68	1.28	-0.6	0.73	-0.05	0.83	-0.15
6/6/1999	0.69	1.26	-0.57	0.8	-0.11	0.8	-0.11
6/7/1999	0.76	1.39	-0.63	0.86	-0.1	0.8	-0.04
6/11/1999	1.14	1.28	-0.14	1.08	0.06	0.86	0.28
6/12/1999	1.23	1.24	-0.01	1.12	0.11	0.92	0.31
6/13/1999	1.25	1.24	0.01	1.17	0.08	0.97	0.28
6/14/1999	1.25	1.26	-0.01	1.21	0.04	1.01	0.24
6/15/1999	1.26	1.22	0.04	1.25	0.01	1.06	0.2
7/11/1999	1.58	1.38	0.2	1.59	-0.01	1.52	0.06
7/12/1999	1.57	1.29	0.28	1.58	-0.01	1.53	0.04
7/13/1999	1.56	1.37	0.19	1.58	-0.02	1.52	0.04
7/14/1999	1.54		1.54	1.58	-0.04	1.5	0.04
7/15/1999	1.55		1.55	1.57	-0.02	1.56	-0.01
8/1/1999	1.48	1.28	0.2	1.51	-0.03	1.41	0.07

For P33, the stations that are included in the MLR model are R127 and G3273 as well as wind parameters. Similar to CP, the bulk of the explanation of the variability in P33 came from the stage parameters. The R^2 value for the P33 MLR model is 0.78, not as high as the CP MLR model. The relationship between observed stage at P33 and G3273confirm that water management is affecting water levels in Shark River Slough. G3273 is a station that is used to make decisions on the flow of water under Tamiami Trail into northeast Shark River Slough. This control of the stage may also affect the R^2 value for the MLR model.

Simple statistics for observed and simulated values of P33 are presented in Table 4. Deviations from the observed mean value for the observed data as seen in the simulations were similar in magnitude to CP. Linear interpolation and MLR work well for the shorter duration gaps (Table 5). As for CP, when the simulated and observed values were close to the mean value, the ARMA simulation was also close to the observed value. Figure 3a and b show that the MLR simulations are not as close to the observed data for P33 for the longer duration gaps (>5 days) compared to CP, but the fit was still good. The ARMA simulations for longer gaps in stage were not very good, and the linear interpolation simulations were only reasonable when the data were not monotonic or vice versa.

Table 4. Simple statistics for observed and simulated values of P33.

	N	Mean	Std Dev	Minimum	Maximum
P33	379	6.66	0.57	5.83	8.08
ARMA	321	6.56	0.08	6.07	6.81
LI	379	6.81	0.46	5.91	7.73
MLR	379	6.72	0.47	5.68	7.74

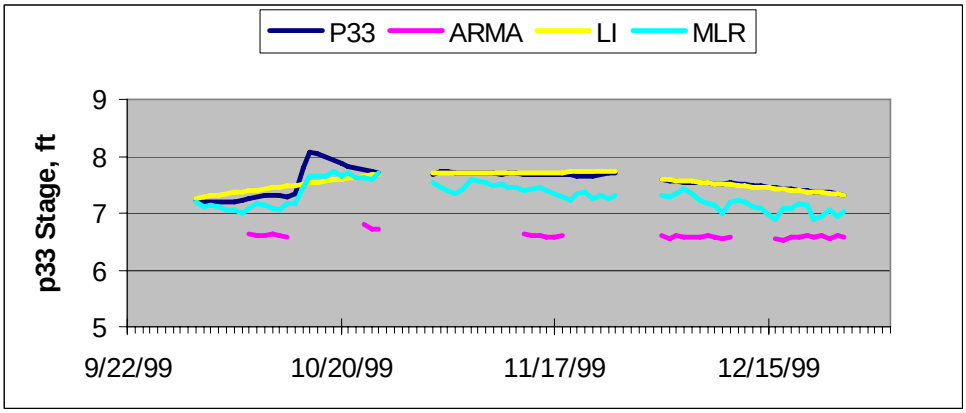


Figure 3a. Comparison of observed and simulated values of P33 for artificially-created gaps of about 25 Days each..

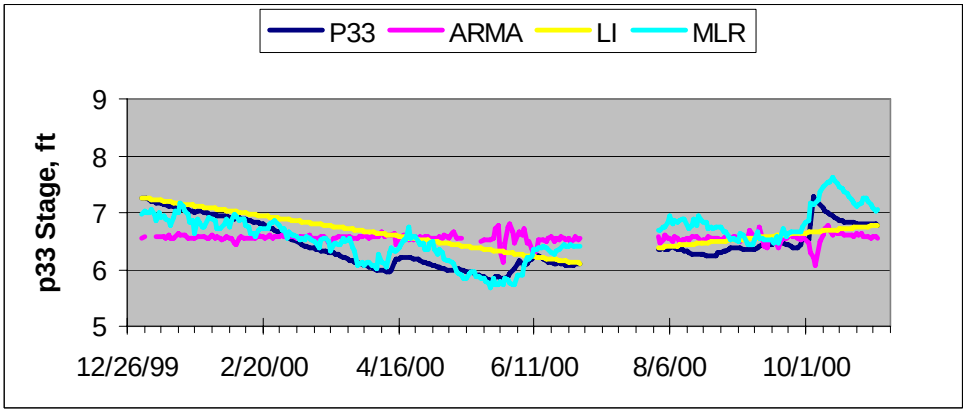


Figure 3b. Comparison of observed and simulated values of CP for artificially-created gaps of 6 and 3 months each, continued.

Table 5. A comparison of simulated values of stage for monitoring station P33 to observed values for artificially-created data gaps that were short in duration (<5 days). OBS-SIM is difference between observed and simulated values. OBS-SIM is difference between observed and simulated values.

Date	P33	ARMA	OBS-SIM	LI	OBS-SIM	MLR	OBS-SIM
2/1/1999	6.74	6.57	0.17	6.74	0	6.73	0.01
3/1/1999	6.45	6.57	-0.12	6.45	0	6.42	0.03
4/1/1999	6.08	6.56	-0.48	6.08	0	5.94	0.14
4/2/1999	6.07	6.56	-0.49	6.07	0	5.95	0.12
4/11/1999	5.97	6.59	-0.62	5.97	0	5.91	0.06
4/12/1999	5.95	6.58	-0.63	5.95	0	5.8	0.15
4/13/1999	5.94	6.58	-0.64	5.94	0	5.88	0.06
4/14/1999	5.93	6.6	-0.67	5.93	0	5.93	0
4/15/1999	5.91	6.61	-0.7	5.91	0	5.83	0.08
6/1/1999	6.14	6.52	-0.38	6.17	-0.03	6.28	-0.14
6/2/1999	6.32	6.8	-0.48	6.25	0.07	6.34	-0.02
6/5/1999	6.28	6.53	-0.25	6.32	-0.04	6.44	-0.16
6/6/1999	6.26	6.51	-0.25	6.34	-0.08	6.43	-0.17
6/7/1999	6.31	6.45	-0.14	6.36	-0.05	6.5	-0.19
6/11/1999	6.31	6.38	-0.07	6.33	-0.02	6.39	-0.08
6/12/1999	6.29	6.38	-0.09	6.33	-0.04	6.36	-0.07
6/13/1999	6.27	6.53	-0.26	6.33	-0.06	6.38	-0.11
6/14/1999	6.27	6.55	-0.28	6.32	-0.05	6.41	-0.14
6/15/1999	6.32	6.48	-0.16	6.32	0	6.4	-0.08
7/1/1999	6.79		6.79	6.79	0	6.86	-0.07
7/2/1999	6.84		6.84	6.84	0	6.85	-0.01
7/5/1999	6.83		6.83	6.83	0	6.94	-0.11
7/6/1999	6.81		6.81	6.81	0	7	-0.19
7/7/1999	6.79		6.79	6.79	0	6.94	-0.15
7/11/1999	6.72	6.66	0.06	6.72	0	6.8	-0.08
7/12/1999	6.7	6.52	0.18	6.71	-0.01	6.83	-0.13
7/13/1999	6.69	6.58	0.11	6.71	-0.02	6.85	-0.16
7/14/1999	6.68		6.68	6.7	-0.02	6.8	-0.12
7/15/1999	6.68		6.68	6.69	-0.01	6.79	-0.11
8/1/1999	6.79	6.57	0.22	6.84	-0.05	6.73	0.06

Salinity

Salinity models for filling data gaps in salinity data were evaluated in the same manner as for stage. The MMN stations Joe Bay (JB) and Long Sound (LS) were selected as dependent variables for the evaluation. A time series with artificially-created data gaps was developed and the gaps were filled by an ARMA model, a SAS© linear interpolation technique, and an MLR model, then simulated values were compared to the observed values.

The Joe Bay MLR model is comprised of Long Sound and Taylor River as well as wind parameters. The R^2 value for the Joe Bay MLR model was 0.95. The Joe Bay ARMA model included Long Sound and Taylor River.

Table 6 presents the simple statistics for the values that were simulated for the Joe Bay artificial gaps as well as for the observed data. As can be seen the MLR average value was close to the observed average value, the linear interpolation average was a little low, and the ARMA average was quite a bit higher.

When the short and long period data gaps for Joe Bay are examined (Table 6 and Figures 4a and b) it can be seen that there are greater differences in observed and modeled values compared to stage. Although the MLR models provided the best fit to the observed data, there are times when the simulated value differs greatly from the MLR model simulation. For the short period data (Table 7), the MLR simulations are generally lower than the observed values, but for the longer duration gaps (Figures 4a and b) the MLR values are both larger and smaller than the observed data. However, the simulations by linear interpolation and ARMA models are poor to extremely poor.

Table 6. Simple statistics for observed and simulated values of salinity at the Joe Bay MMN station.

	N	Mean	Std Dev	Minimum	Maximum
Observed	346	11.80	9.37	0.33	36.98
MLR	382	11.89	8.44	0.00	33.35
ARMA	382	15.80	0.54	12.68	19.22
LI	408	11.25	7.26	0.57	36.53

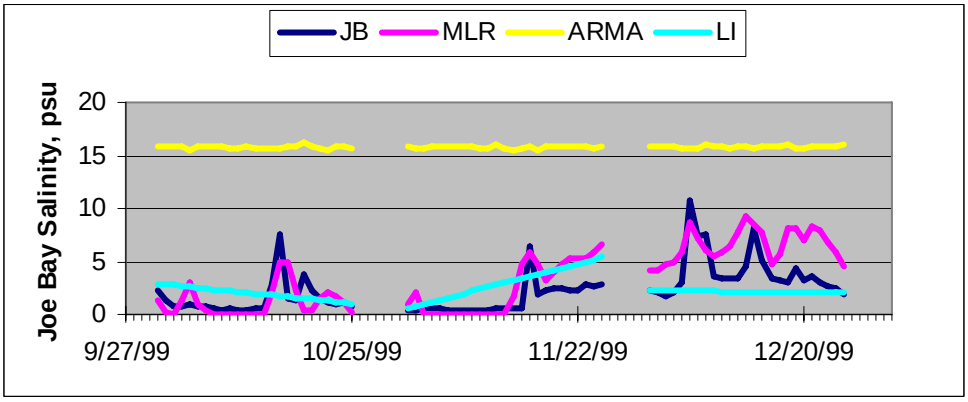


Figure 4a. Comparison of observed and simulated values of Joe Bay for artificially-created gaps of about 25 Days each.

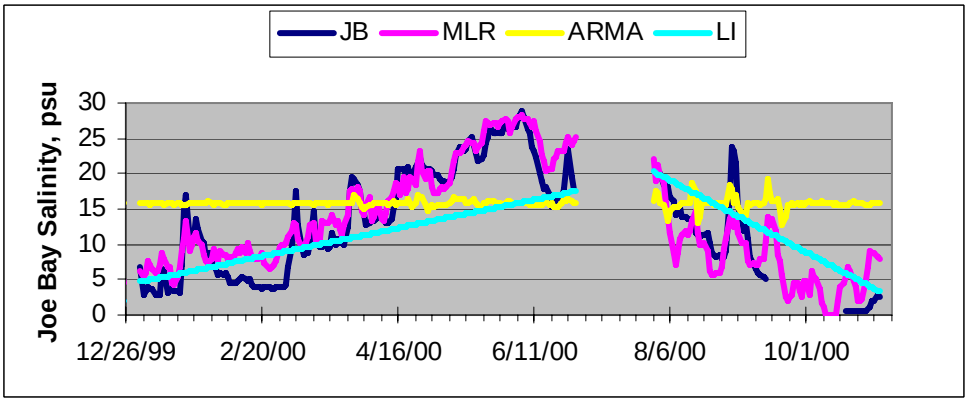


Figure 4b. Comparison of observed and simulated values of Joe Bay for artificially-created gaps of 6 and 3 months each.

Table 7. A comparison of simulated values of salinity for Joe Bay to observed values for artificially-created data gaps that were short in duration (<5 days). OBS-SIM is difference between observed and simulated values.

date	JB	ARMA	OBS-SIM	LI	OBS-SIM	MLR	OBS-SIM
2/1/1999	14.14	15.74	-1.60	15.46	-1.32	12.99	1.15
3/1/1999	23.36	16.63	6.73	23.94	-0.58	20.24	3.12
4/1/1999	24.25	15.95	8.30	24.79	-0.54	22.93	1.32
4/2/1999	24.48	15.46	9.02	24.80	-0.32	22.64	1.84
4/5/1999	25.15	15.53	9.62	25.01	0.14	22.91	2.24
4/6/1999	25.50	16.09	9.41	24.86	0.65	22.80	2.70
4/7/1999	25.38	16.47	8.91	24.70	0.68	22.83	2.55
4/11/1999	28.85	16.68	12.17	29.67	-0.82	26.23	2.62
4/12/1999	28.68	16.22	12.46	30.27	-1.59	26.36	2.32
4/13/1999	29.30	16.01	13.29	30.88	-1.58	25.93	3.37
4/14/1999	30.51	16.65	13.86	31.49	-0.98	27.04	3.47
4/15/1999	32.00	16.06	15.94	32.09	-0.09	30.54	1.46
6/1/1999	35.50	15.22	20.28	35.59	-0.09	30.81	4.69
6/2/1999	34.76	15.89	18.87	35.33	-0.57	30.60	4.16
6/5/1999	34.30	15.64	18.66	34.48	-0.18	28.88	5.42
6/6/1999	33.77	15.64	18.13	34.34	-0.57	29.10	4.67
6/7/1999	33.57	15.57	18.00	34.20	-0.63	29.22	4.35
6/11/1999	36.53	16.66	19.87	35.79	0.75	32.38	4.15
6/12/1999	36.80	15.91	20.89	35.97	0.83	32.50	4.30
6/13/1999	36.98	15.97	21.01	36.16	0.82	31.75	5.23
6/14/1999	36.50	15.19	21.31	36.34	0.16	32.31	4.19
6/15/1999	36.71	15.78	20.93	36.53	0.19	33.35	3.36
7/1/1999	19.94	16.27	3.67	19.38	0.56	14.20	5.74
7/2/1999	18.05	15.86	2.19	18.01	0.04	14.08	3.97
7/5/1999	14.21	15.75	-1.54	15.68	-1.47	11.75	2.46
7/6/1999	12.72	16.16	-3.44	14.79	-2.07	10.38	2.34
7/7/1999	13.31	15.86	-2.55	13.89	-0.58	11.63	1.68
7/11/1999	9.32	15.87	-6.55	9.07	0.25	8.00	1.32
7/12/1999	10.22	15.84	-5.62	8.68	1.54	7.56	2.66
7/13/1999	9.18	15.72	-6.54	8.29	0.90	7.90	1.28
7/14/1999	8.73	15.89	-7.16	7.89	0.84	7.85	0.88
7/15/1999	7.63	15.80	-8.17	7.50	0.13	7.18	0.45
8/1/1999	26.94	15.79	11.15	22.52	4.42	13.96	12.98

The Long Sound MLR model is comprised of Joe Bay, Highway Creek, and Blackwater Sound as well as wind parameters. The R^2 value for the Long Sound MLR model was 0.99. The Long Sound ARMA model included Joe Bay and Taylor River.

The simple statistics for the Long Sound simulations by the models compared to the observed data are presented in Table 8. The average value for the MLR simulations is similar to the observed average but the linear interpolation and the ARMA average values are not close.

Table 9 presents the comparison of observed data for Long Sound to simulated values for the short duration gaps (<5 days). Many of the linear interpolation values were similar to the observed values. Although most of the MLR simulated values are missing because of missing input data, the MLR simulations also compare favorably to the observed data. The ARMA simulations are poor.

Figures 5a and b present the comparison for longer gaps. The MLR simulations for Long Sound are better than the MLR simulations for Joe Bay and are considered good. The linear interpolation simulations for Long Sound are also good until the salinity begins to fluctuate, in which case the linear interpolation simulation is poor, as expected. The ARMA simulations are poor.

Table 8. Simple statistics for observed and simulated values of salinity at the Long Sound MMN station.

	N	Mean	Std Dev	Minimum	Maximum
LS	382	14.51	8.04	1.62	33.90
ARMA	344	18.47	0.36	17.03	20.80
LI	382	15.64	6.63	2.69	33.90
MLR	293	14.70	7.13	3.19	33.91

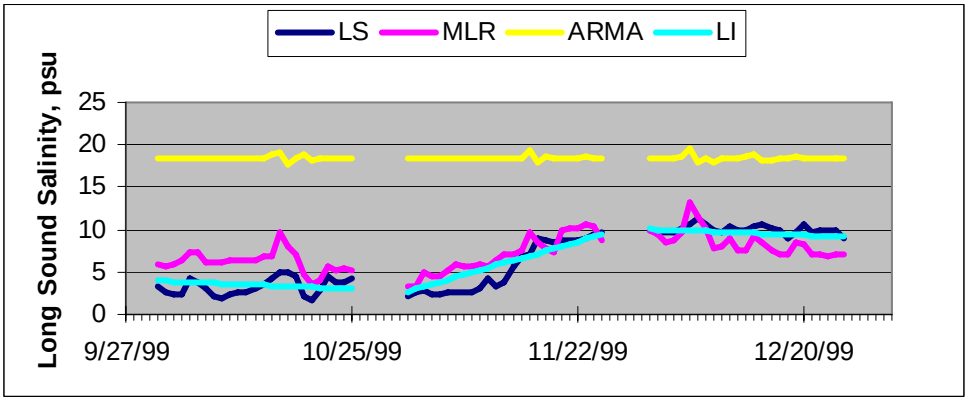


Figure 5a. Comparison of observed and simulated values of Long Sound for artificially-created gaps of about 25 Days each

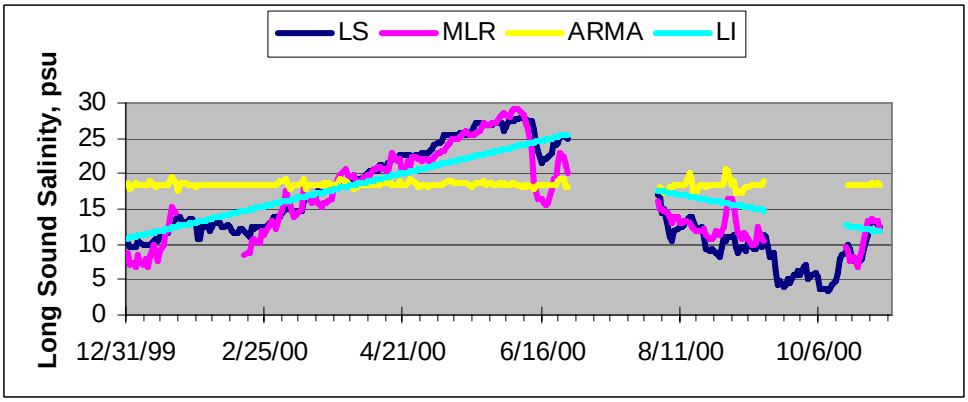


Figure 5b. Comparison of observed and simulated values of Long Sound for artificially-created gaps of 6 and 3 months each .

Table 9. A comparison of simulated values of salinity for Joe Bay to observed values for artificially-created data gaps that were short in duration (<5 days). OBS-SIM is difference between observed and simulated values.

date	LS	ARMA	OBS-SIM	LI	OBS-SIM	MLR	OBS-SIM
2/1/1999	17.19	18.48	-1.29	17.11	0.08	16.17	1.02
3/1/1999	22.28	18.64	3.64	22.35	-0.06	23.43	-1.15
4/1/1999	25.14	18.49	6.65	25.25	-0.11		
4/2/1999	25.50	18.37	7.13	25.51	-0.01		
4/5/1999	25.97	18.37	7.60	25.85	0.12		
4/6/1999	26.13	18.68	7.45	26.11	0.02		
4/7/1999	26.34	18.81	7.53	26.36	-0.02		
4/11/1999	27.07	18.90	8.17	27.11	-0.04		
4/12/1999	27.20	18.67	8.53	27.33	-0.13		
4/13/1999	27.44	18.68	8.76	27.55	-0.11		
4/14/1999	27.54	19.08	8.46	27.77	-0.23		
4/15/1999	27.98	18.84	9.14	27.99	-0.01		
6/1/1999	33.49	18.14	15.35	33.38	0.11	33.91	-0.42
6/2/1999	33.15	18.39	14.76	33.28	-0.13	33.44	-0.29
6/5/1999	33.10	18.34	14.76	33.17	-0.07	31.10	2.00
6/6/1999	33.08	18.32	14.76	33.13	-0.05	31.54	1.54
6/7/1999	33.15	18.34	14.81	33.08	0.07		
6/11/1999	33.57	19.04	14.53	33.57	0.00		
6/12/1999	33.72	18.58	15.14	33.65	0.07		
6/13/1999	33.70	18.58	15.12	33.74	-0.03		
6/14/1999	33.90	18.12	15.78	33.82	0.08		
6/15/1999	33.79	18.48	15.31	33.90	-0.11		
7/1/1999	18.46	18.33	0.13	20.95	-2.49		
7/2/1999	17.93	18.20	-0.27	20.26	-2.33		
7/5/1999	17.18	18.14	-0.96	16.52	0.66		
7/6/1999	14.97	18.26	-3.29	16.03	-1.06		
7/7/1999	14.57	18.55	-3.98	15.55	-0.98		
7/11/1999	10.65	18.45	-7.80	10.96	-0.31		
7/12/1999	10.37	18.59	-8.22	10.75	-0.38		
7/13/1999	10.83	18.33	-7.50	10.55	0.28		
7/14/1999	10.25	18.41	-8.16	10.35	-0.10		
7/15/1999	10.13	18.32	-8.19	10.14	-0.01		
8/1/1999	16.47	18.45	-1.98	16.55	-0.08	22.70	-6.23

Temperature

Based on the results of the use of ARMA modeling and linear interpolation method for simulating stage and salinity data to fill data gaps, the evaluation of these methods was not undertaken for either temperature or USGS creek flow data. Linear interpolation is an accepted procedure for filling small gaps in data when the last value before the gap and the first value after the gap are relatively close. Instead the effort was focused on evaluating the utility of using MLR models which showed the most promise for stage and salinity data.

In addition, for temperature, a separate comparison using MLR models was made with independent variables being (1) bottom temperature for top temperature data, and (2) temperature data from another station. The R^2 value for the bottom temperature model developed from the top temperature at Johnson Key was 0.9997, indicating a very high level of correlation between top and bottom temperature stations at Johnson Key. This is expected in a shallow sub-tropical estuarine/marine environment. The R^2 value for the Johnson Key bottom temperature linear regression model developed from the temperature at the Joe Bay station is 0.9961, also a very high level of correlation. These high R^2 values support the use of linear regression models to fill in temperature data gaps excusive of other methods.

Table 10 presents the simple statistics for the comparison between observed temperature data from the Johnson Key bottom station and the values simulated by the simple linear regression model using Johnson Key top station data. As expected, all statistics are very similar. Table 11 and Figures 6a and b indicate that the simulated values are very similar to observed values for both short and long duration data gaps. The R^2 value between the observed temperature values and the values simulated by the linear regression model is 0.99.

For temperature at Joe Bay, the simple (univariate) linear regression model utilizes the Long Sound temperature data as the independent variable for simulations. As shown by the simple statistics (Table 12) and the table and plots presenting the comparison of observed and simulated values for short and long period gaps (Table 13 and Figures 7a and b), the simulated values are not as close to the observed values as was the case for Johnson Key top and bottom station regression model. Individual values may be off by as much as 1.5^{o} C (6/19/1999, Table 13). In fact, the Pearson correlation coefficient between the observed and simulated values for the data that is filling the gaps is 0.9275, which is an R^2 value of 0.86, much less than the R^2 value of 0.99 for simulations using the top station to simulate the bottom station temperature at Johnson Key. While Joe Bay and Long Sound are located in relatively close proximity, they are not as tightly coupled (or correlated) as the Johnson Key top and bottom temperature data.

Table 10. Simple statistics for observed and simulated values of water temperature at the Johnson Key bottom station. For this comparison the MLR values are based on temperature from the Johnson Key top station as the independent variable.

	N	Mean	Std Dev	Minimum	Maximum
Jkbottom	380	26.17	3.64	14.87	32.56
Lin Reg	380	25.92	3.54	14.88	32.62

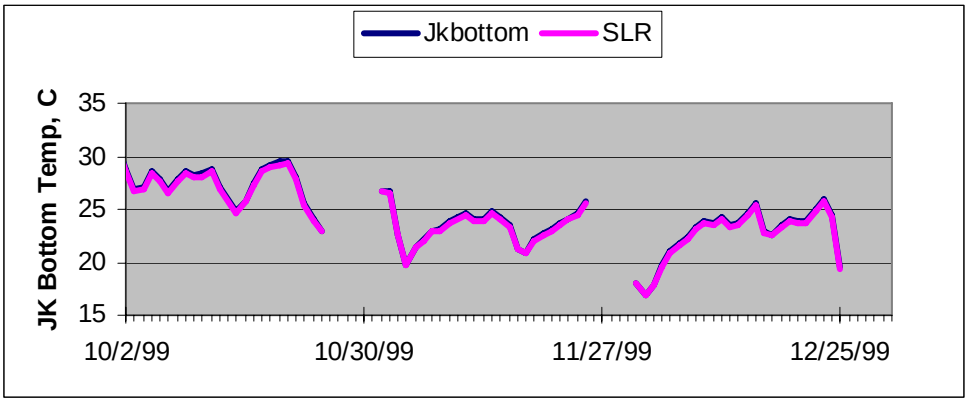


Figure 6a. Comparison of observed and simulated values of Johnson Key bottom temperature for artificially-created gaps of about 25 Days each

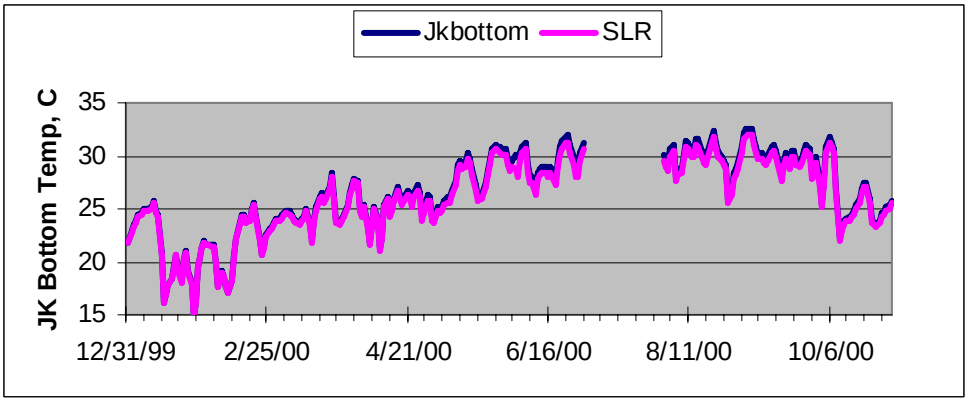


Figure 6b. Comparison of observed and simulated values of Johnson Key bottom temperature for artificially-created gaps of 6 and 3 months each .

Table 11. A comparison of simulated values of temperature (simple linear regression model from Johnson Key top station) for Johnson Key bottom station to observed values for artificially-created data gaps that were short in duration (<5 days). OBS-SIM is difference between observed and simulated values.

date	JK bottom	Lin Reg	OBS-SIM
2/1/1999	23.29	23.76	-0.47
3/1/1999	21.99	22.48	-0.49
4/5/1999	26.68	27.20	-0.52
4/6/1999	27.04	27.61	-0.57
4/7/1999	26.69	27.22	-0.53
4/11/1999	28.24	28.76	-0.52
4/12/1999	27.87	28.39	-0.52
4/13/1999	26.56	27.06	-0.50
4/14/1999	26.84	26.70	0.14
4/15/1999	26.59	27.13	-0.54
6/1/1999	27.66	29.27	-1.61
6/2/1999	28.83	30.39	-1.56
6/5/1999	28.91	30.48	-1.57
6/6/1999	27.69	29.24	-1.55
6/7/1999	28.07	29.64	-1.57
6/11/1999	31.09	32.62	-1.53
6/12/1999	28.48	30.08	-1.60
6/13/1999	28.51	30.04	-1.53
6/14/1999	29.52	31.00	-1.48
6/15/1999	30.47	31.93	-1.46
7/1/1999	29.30	29.96	-0.66
7/2/1999	28.52	28.32	0.20
7/5/1999	29.61	28.68	0.93
7/6/1999	29.30	28.36	0.94
7/7/1999	30.05	29.65	0.40
7/11/1999	29.62	29.62	0.00
7/12/1999	29.71	29.67	0.04
7/13/1999	29.85	29.82	0.03
7/14/1999	29.81	29.77	0.04
7/15/1999	29.98	29.95	0.03
8/1/1999	32.13	32.10	0.03

Table 12. Simple statistics for observed and simulated values of water temperature at the Joe Bay bottom station. For this comparison the univariate linear regression values are based on temperature from the Johnson Key bottom station as the independent variable.

	N	Mean	Std Dev	Minimum	Maximum
jbtemp	275	25.59	4.01	14.52	34.47
Lin Reg	387	26.05	3.61	14.80	32.41

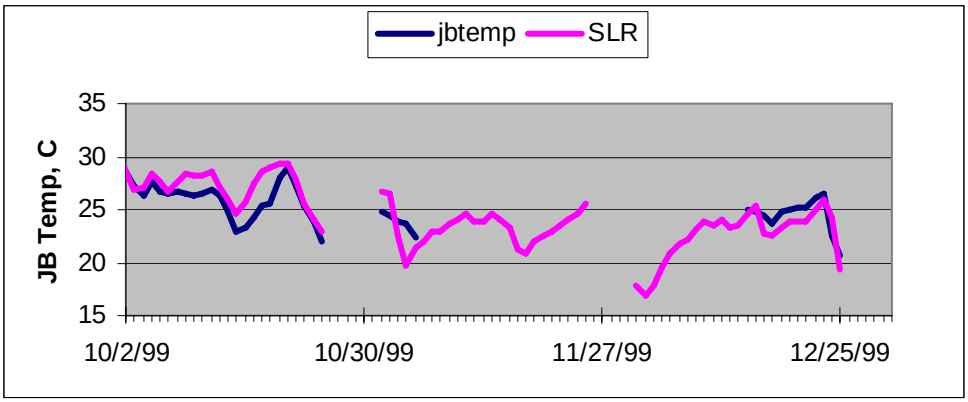


Figure 7a. Comparison of observed and simulated values of Joe Bay bottom temperature for artificially-created gaps of about 25 Days each

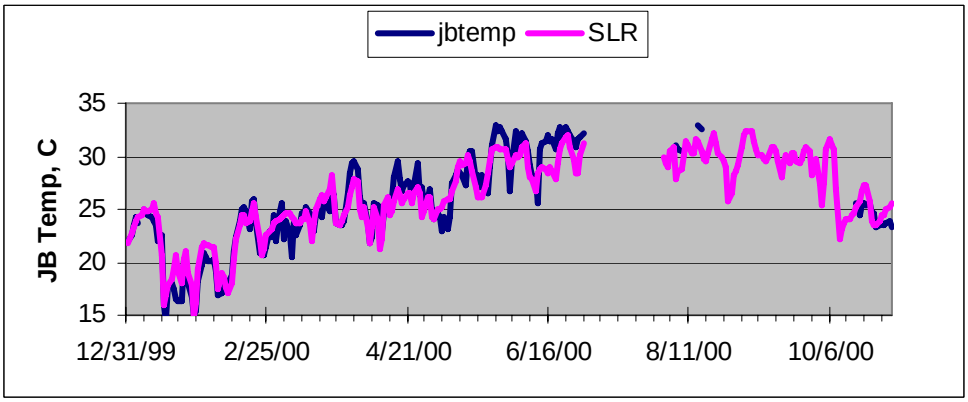


Figure 7b. Comparison of observed and simulated values of Joe Bay bottom temperature for artificially-created gaps of 6 and 3 months each .

Table 13. A comparison of simulated values of temperature (linear regression model from Johnson Key bottom station) for Joe Bay bottom station to observed values for artificially-created data gaps that were short in duration (<5 days). OBS-SIM is difference between observed and simulated values.

date	jbtemp	Lin Reg	OBS-SIM
2/1/1999	21.90	23.19	-1.29
3/1/1999	20.35	21.89	-1.54
4/1/1999	21.67	24.82	-3.15
4/2/1999	23.17	26.31	-3.14
4/11/1999	23.69	28.11	-4.42
4/12/1999	24.35	27.74	-3.39
4/13/1999	24.40	26.44	-2.04
4/14/1999	23.43	26.72	-3.29
4/15/1999	22.97	26.47	-3.50
6/1/1999	26.99	27.54	-0.55
6/2/1999	27.56	28.70	-1.14
6/5/1999	28.50	28.78	-0.28
6/6/1999	28.20	27.57	0.63
6/7/1999	28.47	27.94	0.53
6/11/1999	31.25	30.95	0.30
6/12/1999	28.99	28.35	0.64
6/13/1999	28.71	28.38	0.33
6/14/1999	29.70	29.39	0.31
6/15/1999	30.16	30.33	-0.17
7/1/1999	29.06	29.17	-0.11
7/2/1999	27.52	28.39	-0.87
7/5/1999	28.68	29.48	-0.80
7/6/1999	28.96	29.17	-0.21
7/7/1999	29.33	29.92	-0.59
7/11/1999	30.06	29.49	0.57
7/12/1999	30.49	29.58	0.91
7/13/1999	30.41	29.72	0.69
7/14/1999	30.39	29.68	0.71
7/15/1999	29.77	29.85	-0.08
8/1/1999	30.27	31.99	-1.72

Flow

Flow data are not as serially correlated as stage, salinity, and temperature, and are also not as highly spatially correlated, though the correlation on a spatial basis is still quite strong. Additionally, there is a large differential in the volume rate of flow between groups of stations. For example, the average flow for Mud Creek is about 34 cfs while the average flow for Trout Creek is 270 cfs. Flow is also suspected of being influenced by the wind, so wind parameters were included as candidate independent variables in MLR modeling. Because of the difference in magnitude, it was decided to utilize a low flow station (Mud Creek) and a high flow station (Trout Creek) separately for the evaluation. Attempts to model the low flow station with high flow data resulted in models with a high bias, although the correlation coefficient was relatively large.

For Mud Creek, Table 14 presents the simple statistics for the observed flow data and for the MLR model using Taylor River Entrance flow and wind as independent variables. The MLR model for Mud Creek has an R^2 value of 0.85, not very high compared to stage, salinity, and temperature regression models. The difference in the simple statistics reflects this, though the statistics for the simulated values of flow are somewhat similar to the statistics for the observed data. Table 15 presents the comparison of observed flow data from Mud Creek and the MLR simulations for short term gaps (< 5 days). Though the numerical differences seem large compared to stage, salinity, and temperature, when they are compared to the average value and the range the agreement is good but not exceptional. There were also several simulations with a flow direction opposite to the direction of the observed data but this was always for values close to 0. Figures 8a and b for the longer term gaps shows that, even though the simulations seem to be biased slightly high, the fit to the observed data is good and better than expected considering the stochastic nature of flow data.

For Trout Creek, the situation is similar to the Mud Creek simulation. The simple statistics in Table 16 show that the values simulated by the Trout Creek MLR model that includes Mud Creek and wind parameters compare reasonably well with the statistics for the observed data for the artificially created gaps. Table 17 indicates that, for short duration gaps, the difference between observed and simulated values can be several hundred cfs, though this is not unreasonable given the range of values that is about 3000 cfs. Similar to Mud Creek, some of the Trout Creek simulated values are opposite in sign from the observed when the flow is near 0. Figures 9a and b for the long period gaps shows that the fit to the observed data is quite good, better than for the Mud Creek case.

Table 14. Simple statistics for observed and simulated values of flow at the Mud Creek station. For this comparison the MLR model values are based on flow at the Taylor River East station and wind parameters as the independent variables.

	N	Mean	Std Dev	Minimum	Maximum
mud	366	34.38	66.52	-99.69	265.27
MLR	334	42.04	64.52	-82.29	299.92

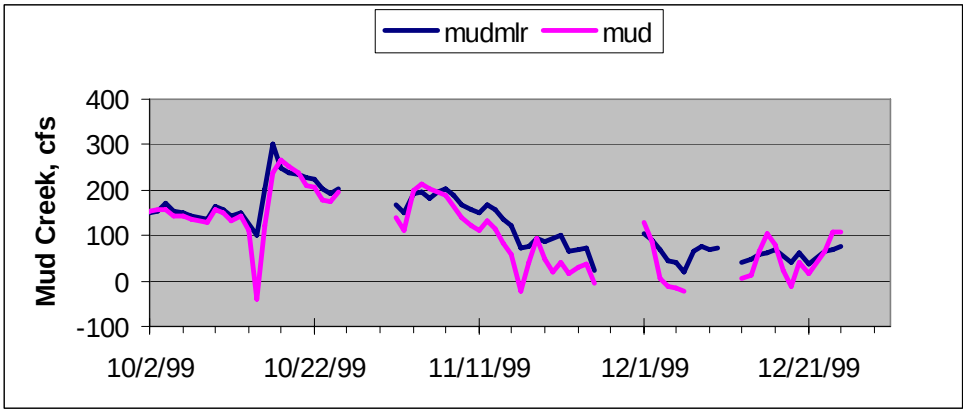


Figure 8a. Comparison of observed and simulated values of flow at Mud Creek for artificially-created gaps of about 25 Days each

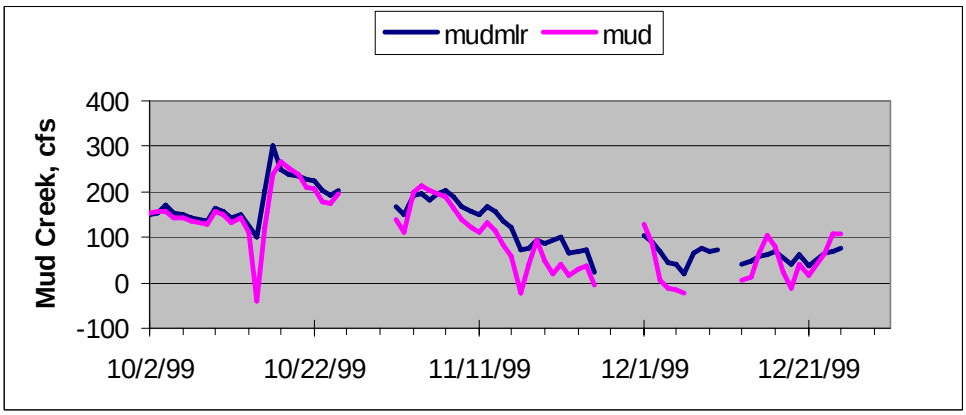


Figure 8b. Comparison of observed and simulated values of flow at Mud Creek for artificially-created gaps of 6 and 3 months each.

Table 15. A comparison of simulated to observed values of flow at the Mud Creek station for short duration, artificially-created data gaps (<5 days). OBS-SIM is difference between observed and simulated values.

date	mud	MLR	OBS-SIM
2/1/1999	-28.56	7.55	-36.11
3/1/1999	-7.90	-18.44	10.54
4/1/1999	16.21	17.53	-1.32
4/2/1999	19.36	3.77	15.59
4/5/1999	4.76	-4.44	9.20
4/6/1999	-4.92	-4.73	-0.19
4/7/1999	11.75	2.21	9.54
4/11/1999	-19.32	-29.23	9.91
4/12/1999	-7.68	-26.73	19.05
4/13/1999	-17.56	-40.31	22.75
4/14/1999	-54.24	-56.76	2.52
4/15/1999	-68.07	-56.88	-11.19
6/1/1999	41.41	40.72	0.69
6/2/1999	32.65	35.83	-3.18
6/5/1999	26.46	22.54	3.92
6/6/1999	35.06	23.62	11.44
6/7/1999	18.51	22.29	-3.78
6/11/1999	-6.80	-21.48	14.68
6/12/1999	-7.67	-23.08	15.41
6/13/1999	7.21	-11.33	18.54
6/14/1999	2.03	-0.95	2.98
6/15/1999	-20.91	-13.42	-7.49
7/1/1999	108.79	98.31	10.48
7/2/1999	92.81	94.56	-1.75
7/5/1999	115.84	103.03	12.81
7/6/1999	109.46	104.07	5.39
7/7/1999	87.41	98.69	-11.28
7/11/1999	90.86	78.86	12.00
7/12/1999	79.92	82.41	-2.49
7/13/1999	51.29	71.06	-19.77
7/14/1999	27.96	59.00	-31.04
7/15/1999	59.42	76.21	-16.79
8/1/1999	-11.89	-2.53	-9.36

Table 16. Simple statistics for observed and simulated values of flow at the Trout Creek station. For this comparison the MLR model values are based on flow at the Mud Creek station and wind parameters as the independent variables.

	N	Mean	Std Dev	Minimum	Maximum
trout	372	270.75	509.2	-984.59	2070
MLR	380	268.67	479.64	-710.86	1935

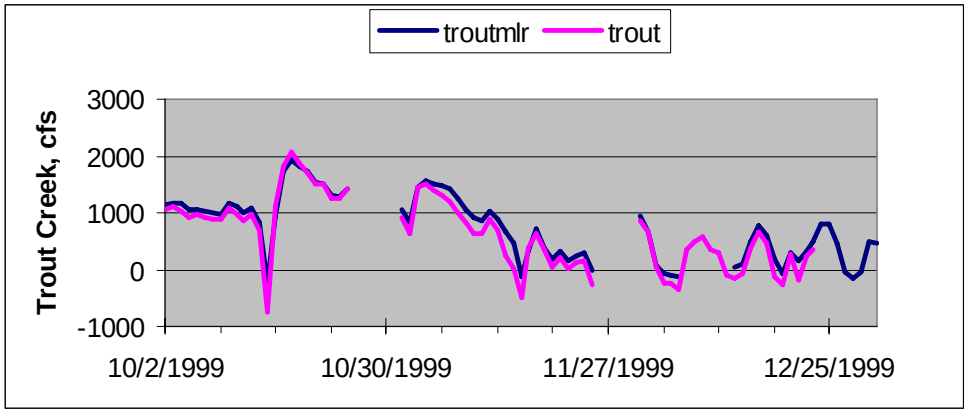


Figure 9a. Comparison of observed and simulated values of flow at Trout Creek for artificially-created gaps of about 25 Days each

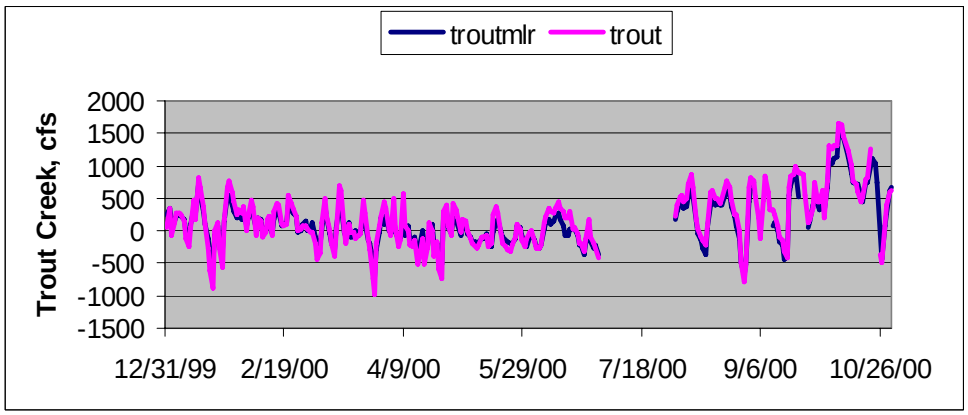


Figure 9b. Comparison of observed and simulated values of flow at Trout Creek for artificially-created gaps of 6 and 3 months each.

Table 17. A comparison of simulated values of flow at the Trout Creek station to observed values for artificially-created data gaps that were short in duration (<5 days). OBS-SIM is difference between observed and simulated values.

date	trout	MLR	OBS-SIM
2/1/1999	-216.34	-178.47	-37.87
3/1/1999	-130.10	-53.35	-76.75
			0.00
4/1/1999	-14.90	152.19	-167.09
4/2/1999	42.92	164.37	-121.45
4/5/1999	-102.64	62.50	-165.14
4/6/1999	-68.53	-12.14	-56.39
4/7/1999	154.65	117.71	36.94
4/11/1999	-319.65	-127.19	-192.46
4/12/1999	-244.73	-43.63	-201.10
4/13/1999	-222.80	-112.08	-110.72
4/14/1999	-576.25	-368.87	-207.38
4/15/1999	-803.65	-465.36	-338.29
6/1/1999	91.43	332.58	-241.15
6/2/1999	30.02	273.18	-243.16
6/5/1999	212.40	228.64	-16.24
6/6/1999	245.53	288.78	-43.25
6/7/1999	15.65	173.74	-158.09
7/1/1999	531.90	804.52	-272.62
7/2/1999	432.82	690.09	-257.27
7/5/1999	789.88	864.79	-74.91
7/6/1999	711.40	826.37	-114.97
7/7/1999	465.19	671.16	-205.97
7/11/1999	584.38	685.54	-101.16
7/12/1999	517.67	604.02	-86.35
7/13/1999	385.39	400.77	-15.38
7/14/1999	201.46	229.65	-28.19
7/15/1999	391.56	455.51	-63.95
8/1/1999	-308.15	-78.40	-229.75

VI. Discussion

One of the primary purposes of this analysis was to determine whether or not time series models would be useful for filling data gaps. It was hoped that the serial correlation in independent variables could be used as to prepare time series models for filling gaps in observed data.

Correlation matrices showed that stage and temperature are highly serially correlated. Salinity is serially correlated to a high degree though not at the higher significance level of stage and temperature. Flow is also correlated serially but at a lower level.

ARMA models for stage, salinity, temperature, and flow were constructed without pre-whitening the variables which meant that simulations made to fill gaps usually contained considerable error relative to the “actual” observed data unless the range of simulation was very close to the mean value. ARMA models proved unreliable for data gaps of more than 1-2 days.

ARIMA models were developed to take advantage of serial correlation, and particular the power of autocorrelation (correlation to the previous time-period value). The model development process is initiated by pre-whitening of the independent and dependent variables which filters out the serial correlation components. The remaining variation in the independent variables is then evaluated for cross-correlation with the dependent variable. If cross-correlation is significant, the independent variables are used for model development, along with the serial correlation in the dependent variable that was identified in pre-whitening. Typically, the first-order autocorrelation (correlation with the lag1 value) is the strongest, i.e. knowing the value of the dependent variable yesterday provides the best information about what the value will be today.

When the serial correlation in the hydrologic and temperature data is not included in the model but the pre-whitened signal of the independent variable remains for model development, the remaining unfiltered variation is generally not sufficient to develop a model that will explain more than about 20% of the variability as a maximum. If neither the independent nor the dependent variables are pre-whitened (no serial correlation used in modeling), the model is equivalent to a multivariate linear regression model and the advantages of ARIMA modeling are no longer being utilized. In addition the SAS© PROC ARIMA procedure is relatively inflexible meaning that utilization of multivariate methods provides a better overall process for filling data gaps compared to ARIMA models. After testing ARIMA models on stage and salinity and determining that the simulations were unreliable, further use of ARIMA models for temperature and flow was eliminated from the work.

Linear interpolation is one method of filling gaps that is common. Linear interpolation is generally considered to be acceptable if the gap is small (1-5

days), the day-to-day change is small, and there are no inflection point in the missing data. Larger gaps (in time) create a chance that an event of some kind may occur, which would change the rate of increase or decrease in a parameter (stage, salinity, temperature, flow) such that linear interpolation was no longer accurate for filling gaps. Similar to ARMA modeling, no additional time was spent evaluating linear interpolation for filling gaps in hydrologic or physical data after evaluating stage and salinity.

Multivariate linear regression (MLR) models proved to be the most successful technique for simulating hydrologic and physical data when there are gaps in the observed data record. MLR models were constructed from same-type data that was available from neighboring stations. Determining the candidate stations that would work best as an independent variable in an MLR model was straight-forward, accomplished through the use of a correlation matrix and step-wise development of the regression model. For stage and flow, use of the intercept term produced the best model, while a model with no intercept was best for salinity and temperature. However, use of a no-intercept model is only appropriate under certain circumstances, when one is confident that the distribution of values for the observed data and the distribution of values for the gap-filling data are equivalent, and the means are equal, which is rarely if ever known.

In addition, stage, salinity, and flow benefited from the inclusion of wind in the model. However, the benefit of including wind terms in a temperature model was very small, perhaps because the level of correlation with neighbor stations was so high with temperature. In all cases where wind was a significant parameter in the MLR model, the wind parameters generally only explained on the order of 1-5% of the variability in the dependent variable.

The use of multiple independent variables generally increased the explanatory capability of the model compared to a univariate model by about 1-5% for stage, 3-5% for salinity, and 3-5% for flow. The R^2 value for univariate temperature models was over 0.86 so multivariate modeling was not needed. The R^2 value for MLR models ranged form a high of 0.99 for salinity (Long Sound) to 0.78 for stage (P33).

VII. Conclusions

The conclusions of this effort to determine if time series models could be useful for filling gaps in time series data for stage, salinity, temperature, and flow are clear –ARIMA models are not very useful for this purpose. This in-depth analysis of ARIMA models brought to the surface the level to which this procedure relies on the availability of past values of the dependent variable, which are not available when the data gap exceeds one day for daily data. ARIMA models work reasonably well for gaps of 2-3 days particularly when the range of the data in the 2-3 day period is near the mean value for time series. However, the ability of ARIMA models to provide highly accurate one-step forward predictions using the serial correlation in the data means that ARIMA models should be considered for use as a predictive tool for water management system operation when the data are not missing.

Similarly, the obvious limitations of linear interpolation limit the use of this technique to small gaps. From the limited analyses done, it appears that gap-size limit for linear interpolation is about 5 days in duration for the 4 parameters of interest. Beyond a 5-day window, the chance that an event will occur that modifies the rate of change increases such that the use of linear interpolation may not be reliable.

Multivariate linear regression (MLR) models proved to be useful for filling data gaps in stage, salinity, and flow. Based on the limited sample of this evaluation MLR models appear to be able to explain about 85-95% of the variability in the stage, salinity, and flow data, and wind measured at Key West and Miami appears to be useful in modeling by explaining up to 5% additional variation over models that do not include wind. However, models that include wind tend to have a large number of independent variables even at the 0.999 level of significance, which may be the result of cross-correlation effects between independent variables.

In the case of temperature, the correlation of neighbor stations as independent variables to the dependent variables was so high that MLR models did not provide sufficient benefit to justify MLR use. Instead simple univariate linear regression models were more than adequate. The use of bottom measurements to model top measurements at the same station (and vice versa) provides an advantage over the use of neighbor data as an independent variable.

Based on this analysis the flexibility of MLR models was shown to be useful for filling data gaps. MLR models have the added advantage of being transparent and easy to understand by a wide audience. The use of neighbor independent variables and wind data in MLR models for filling data gaps is recommended when simple regression models are considered to be inadequate.

VIII. References

Brockwell, P.J. and R.A Davis. 1996. Introduction to Time Series and Forecasting. Springer-Verlag, New York, New York. 420 pp.

Marshall III, F.E.; D. Smith; and D. Nickerson. 2003a. Salinity Simulation Models for North Florida Bay Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. 41 pp.

Marshall III, F.E. 2003b. IOP Salinity Analysis Using Statistical Models. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. 35 pp.

Marshall III, F. E.; D. Smith; and D. Nickerson. 2004. Using statistical models to simulate salinity variation and other physical parameters in north Florida Bay. Cetacean Logic Foundation, Inc., New Smyrna Beach, Florida.

Marshall III, F. E. 2005a. RECOVER Southern Estuaries performance measures: identification of hydrology-salinity relationships for coastal estuaries and analysis of interim CERP update scenarios. Environmental Consulting & Technology, Inc., New Smyrna Beach, Florida.

FINAL TASK REPORT

TASK 5 - FRESHWATER FLOW ESTIMATES FOR THE FLORIDA BAY AND FLORIDA KEYS FEASIBILITY STUDY
SALINITY PERFORMANCE MEASURES

Cooperative Agreement Number CA H5284-03-0034
Between
The United States Department of the Interior National Park Service
Everglades National Park
And
Cetacean Logic Foundation, Inc.

Frank E. Marshall, III
Principal Investigator

DeWitt Smith
Project Coordinator

August 6, 2008

CESI FY2005 Project No. CA H5284-03-0034 for Everglades National Park

FINAL TASK REPORT

TASK 9 - ESTIMATION OF THE FLOW REQUIRED FOR THE FLORIDA BAY AND FLORIDA KEYS FEASIBILITY STUDY SALINITY PERFORMANCE MEASURES

November 30, 2006

I. Introduction

A. General

This report describes a task for the current Critical Ecosystem Studies Initiative (CESI) project that was originally initiated by Cetacean Logic Foundation, Inc. (CLF) for Everglades National Park (ENP) in 2002. The work was continued in 2003 and 2004, and has been continued into 2005/2006. This task is an FY2005 project. The subject of this on-going CESI work is the characterization and simulation of the salinity regime and the upstream freshwater hydrology in the study area of Florida Bay and the southwest coast of the Gulf of Mexico (Gulf) within the Park.

This task involves the development of models that can utilize a salinity target to produce an estimate of the flow required to accomplish the salinity target. To-date there has been no quantitative methods proposed to estimate the freshwater flows at various locations in the Everglades given a specified salinity regime. Therefore, this is a proof-of-concept exercise and the procedure described herein may be updated in the future as it is used. The purpose of this task is to develop the procedure.

This task has broad-reaching implications, including use of the developed procedure with paleosalinity information to develop estimates of the pre-drainage hydrologic conditions. This report describes the task of preparing models to investigate the flow required at specific monitoring stations to meet the salinity performance measures for the Florida Bay and Florida Keys Feasibility Study (REFERENCE).

B. Task Objectives and Evaluation Methods

The objectives of this task are presented below as they appear in the contract for this CESI project approved by ENP:

“Estimate the Flow Regime Required to Meet the Florida Bay and Florida Keys Feasibility Study Salinity Performance Measures

“For this task, the same combination of MLR and seasonal autoregressive integrated moving average (SARIMA) techniques that were used to produce the

current MLR salinity models will be used to develop regression models for flow at the S-12, S-18C, S-197, and S-175 structures as a function of observed stage at the statistically significant stations in the Everglades. These models are the same as the hydraulic rating curves used for weirs, gates and other flow control structures. These techniques will also be used to establish MLR relationships between the primary and secondary stage (water level) variables in the various MLR models. Model goodness-of-fit statistics will be computed and verification of all models will be performed using observed data for comparison.

"At least one MLR salinity model will be assigned to each FBFKFS zone in northeast Florida Bay, Zones 1, 2, 3, 5,and 14. For each of the zones, plots of the FBFKFS requirements will be used with plots of the NSM salinity simulation made for the Southern Estuaries sub-team to establish a salinity regime at a particular monitoring station that meets the performance measure requirements to the greatest extent possible. The NSM time series will be adjusted by adding or subtracting an equal amount to each NSM value over the 36-year period 1965-2000. In this manner, the NSM simulation will only be used to simulate the day-to-day salinity variation, not the magnitude of the salinity estimate. Then, a well-documented statistical procedure called inverse prediction (Neter et al. 1990) will be used with the MLR salinity models to establish a stage regime at the primary MLR model stations - Craighead Pond (CP), P33, and EVER7, using observed wind and water level and the previously developed regression models for the secondary stations.

"Then the flow regression models (hydraulic rating curves) developed previously will be used to estimate the required flow regime at the S12, S18C, S197, and S-175 structures for the 36-year simulation period 1965 through 2000. This will be done for each of the zones, Zones 1, 2, 3, 5, and 14." (Marshall 2005)

C. Background

In south Florida, the hydrologic cycle of freshwater is inextricably linked to the salinity regime in the estuarine regions of the coastal water bodies. Determining the quantity of freshwater that is needed in the Everglades hydrologic system to restore and maintain healthy biota (freshwater, estuarine, and marine) in ENP is a primary question for Everglades and Florida Bay researchers. Freshwater quantity in the Everglades is monitored through level (stage) measurements in wells and flow measurements at structures and key locations in the freshwater wetlands. Because stage is a response variable to flow and flow is highly stochastic at the daily level, salinity variability is better explained by stage in wells than by flow measurements. Because of this, multivariate linear regression (MLR) models using stage as one of the independent variables have been developed for 17 of the ENP Marine Monitoring Network (MMN) salinity stations in Florida Bay and along the southwest Gulf coast (Marshall, et al 2004; Marshall, 2005).

Consumptive use permitting of water (for example domestic water supply and agricultural irrigation) utilizes a volume of water over a specific time to allocate freshwater resources. Policy decisions regarding freshwater deliveries for the restoration of the Everglades/Florida Bay system require flow volumes as well as stage to insure that the right amount of water is being delivered at the right time and location. Water budgets are useful tools for analyzing hydrologic situations, and a primary element of a water budget is flow across a boundary.

Before the construction of the first drainage alterations in south Florida around the turn of the 20th century, flows into northeast Florida Bay were controlled by the rainfall and evaporation patterns that occurred within the large Everglades watershed, including Lake Okeechobee. The timing of the rainfall in the watershed and the volume were affected by both global and local climatic conditions, and the local conditions within the watershed can be quite variable spatially and temporally. For example, the upper part of the drainage basin (the Kissimmee River basin) may be experiencing a period of low rainfall at the same time that the Everglades are experiencing wet conditions. The spatial distribution of freshwater flowing into Florida Bay depends to some degree on the location of rainfall events, as well as the established natural flow ways and the effect of water management structures. Historically, most of the surface flows to northeast Florida Bay were directed through five creeks and Taylor River. For the west coast of ENP north of Cape Sable there are multiple flow ways, but most of the water originates in Shark River Slough.

Anthropogenic water management has changed the temporal and spatial hydrology and water budget of the Park. Since the beginning of the 1900's, water has been diverted from the natural system for a number of reasons. The effect has been a degradation of the ecosystems in both the Everglades and Florida Bay (Ogden et al. 2005). For ecosystem restoration, freshwater of sufficient quantity must be delivered at the right time and in the right locations. To understand the current situation, assess proposed alternatives, and specify flow volumes for restoration, the flow characteristics at the monitored locations need to be evaluated. For the Shark River Slough, water is now controlled primarily by the S-12 structures that flow beneath Tamiami Trail. For Florida Bay, freshwater deliveries are uncontrolled in Taylor Slough but the volume of water provided is monitored at the Taylor Slough bridge in the Park. Freshwater flows are also controlled and measured east of Taylor Slough through the C-111 Canal at the S-18C and S-197 structures. In general freshwater flow is uncontrolled at the coastal creek that discharge directly into Florida Bay. Since these creeks are downstream from S-18C and Taylor Slough, creek flows are also related somewhat to flows measured at S-18C and Taylor Slough bridge and any other flows from groundwater and additional direct rainfall.

Therefore, determining the volume of flow required at locations where flow has been measured for a particular salinity regime is an important first step in

estimating the long-term allocation of freshwater to ENP for restoration and protection of the unique natural resources of the Everglades and Florida Bay.

II. Data and Methods

In this procedure, salinity was used to specify stage, and stage was then used to estimate flow. The models used in this process were developed from observed data. The data used for this modeling activity includes daily values of stage at two locations in the southern part of the Everglades, Craighead Pond (CP) and EVER7. Both of these stations, along with P33 located in Shark River Slough, were identified as primary stage stations in the development of the MLR models (Marshall, et al 2004; Marshall, 2005) because they were consistently selected by the stepwise regression modeling process as the independent variable explaining the greatest portion of salinity variation, as compared to the data from other stage monitoring stations. In general, this means that the data from these primary stage stations is more highly correlated with salinity than the data from other stations.

The data also include salinity at the following MMN stations for which MLR salinity models have been developed and the FBFKFS Zone in which they are located:

1. FBFKFS Zone 1
   Long Sound
   Joe Bay
   Little Madeira Bay
2. FBFKFS Zone 2
   Duck Key
   Butternut Key
3. FBFKFS Zone 3
   Whipray Basin
4. FBFKFS Zone 5
   Terrapin Bay
   Garfield Bight

For stage and flow, data were obtained from the South Florida Water Management District's DBHYDRO database (http://www.sfwmd.gov/org/ema/dbhydro/index.html). Creek flows were obtained from the USGS website (http://sofia.usgs.gov/). Salinity data were obtained from the South Florida Natural Resources Center website (http://www.sfnrc.ever.nps.gov).

The methodology for estimating flow from target salinity values is described in the flow chart that is presented as Figure 1. As can be seen, the first step in the process is to develop an estimate of the salinity regime for a FBFKFS target. The FBFKFS performance measures describe a range of annual salinities that are desirable in a specific FBFKFS zone or basin. This desirable range was based on the best professional judgment of the scientists, ecologists,

hydrologists, and engineers that were involved in the FBFKFS. The annual salinity ranges used for this analysis are as follows:

1. Zone 1
    a. Long Sound: 5-15 psu
    b. Joe Bay: 5-15 psu
    c. Little Madeira Bay: 15-25 psu

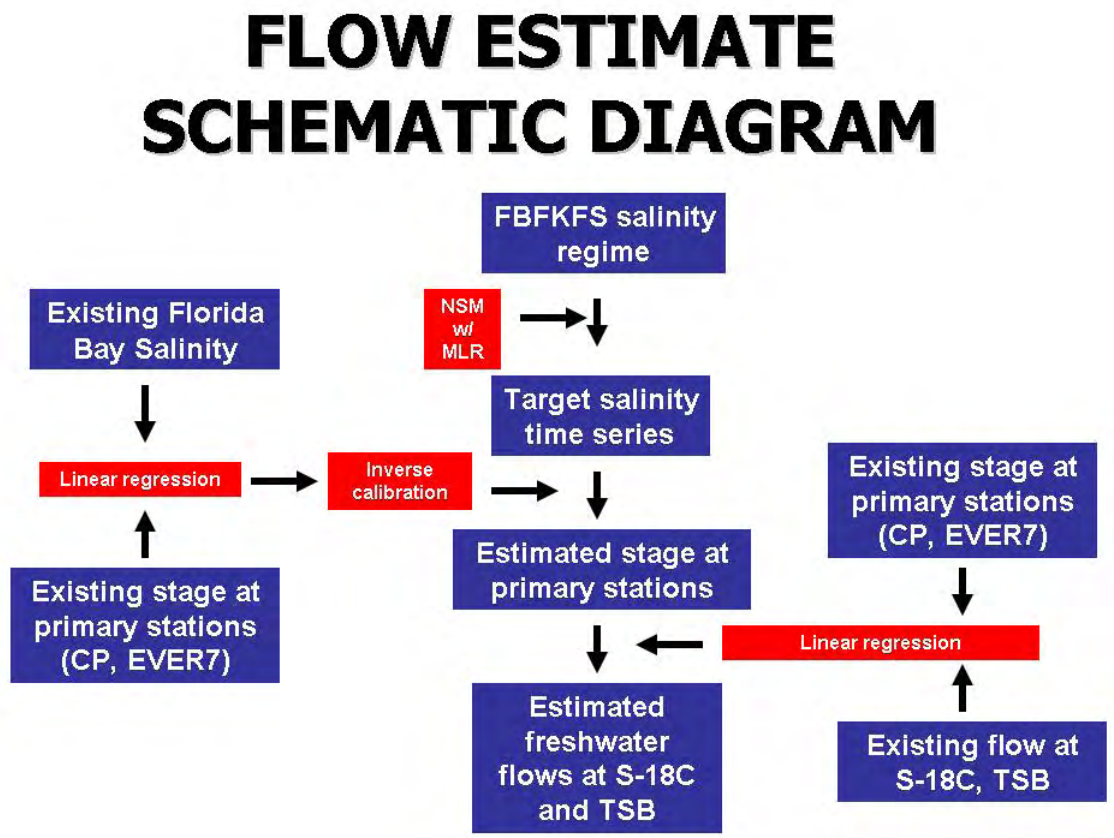
2. Zone 2
    a. Duck Key: 15-30 psu
    b. Butternut Key: 15-30 psu

3. Zone 3
    a. Whipray Basin: 25-35 psu

4. Zone 5
    a. Terrapin Bay: 15-35 psu
    b. Garfield Bight: 15-35 psu

5. Zone 6
    a. Bob Allen: 25-35 psu

6. Zone 14
    a. Little Blackwater Sound: 10-20 psu
    b. Blackwater Sound: 15-30 psu



For some of the zones the desired annual average salinity regime is the same for all basins in that zone (Zones 2, 3, 5, and 6). However, for Zones 1 and 14 the target range is different for different basins in the zone.

The NSM 4.6.2 daily salinity time series data that were produced for RECOVER using the MLR salinity models were averaged to annual values for the 36-year simulation period. The 36 values were plotted and compared to the FBFKFS desired salinity range. Where necessary, the NSM / MLR–based salinity regime was modified by adding/subtracting an equal value to/from each of the 36 annual mean values so that as many as possible of the 36 values fell within the FBFKFS desired salinity range. In some cases this meant having as many annual average values above the range as there were below the range. But for all cases it was relatively easy to assemble a 36-year mean annual value time series that satisfied the FBFKFS requirements.

Time series constructed in this manner were considered the targets for which flows are to be estimated. Six of the stations required no adjustment of the NSM / MLR annual series, while four of the stations required subtracting values to bring the NSM / MLR time series into the range, meaning that the NSM-based salinity estimates were higher for those locations than the FBFKFS salinity targets. The greatest amount that needed to be subtracted from the NSM / MLR annual average salinity series was 5 psu at Long Sound. For the four stations

that required adjustment of the NSM / MLR salinity series to meet the targets, the daily time series were modified in like manner to the annual series by subtracting the appropriate value from each daily value. For those basins that did not require adjustment, the daily salinity data produced by the MLR salinity models from NSM 4.6.2 were used. This created ten “target” time series of daily values for modeling purposes (Figures 2 – 15).

Figure 2. Long Sound Annual Average Salinity Based on NSM 4.6.2 compared to the FBFKFS Salinity Targets (15-25 psu).

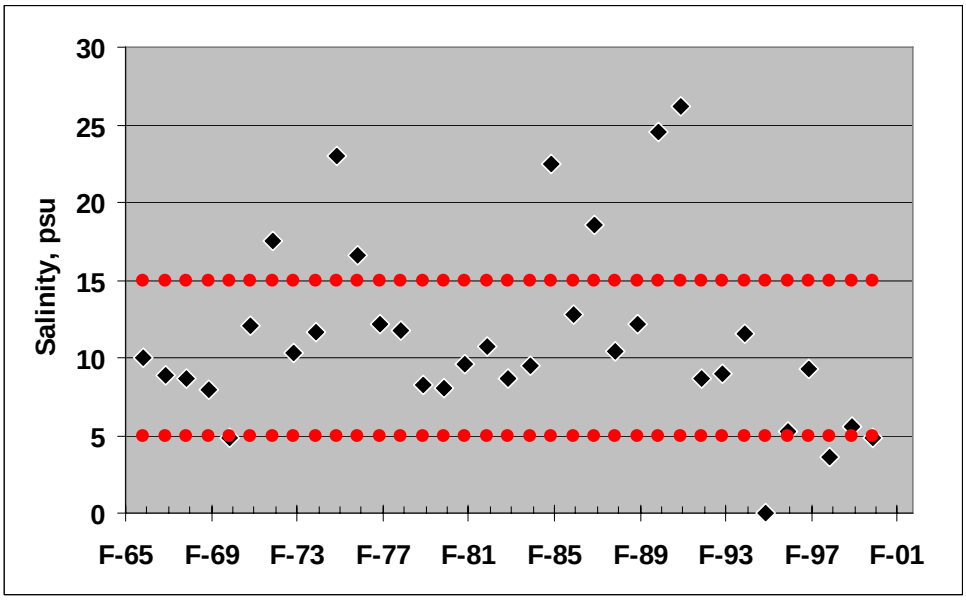


Figure 3. Long Sound Annual Average Salinity Based on NSM 4.6.2 *Minus 2.5 psu* compared to the FBFKFS Salinity Targets (15-25 psu).

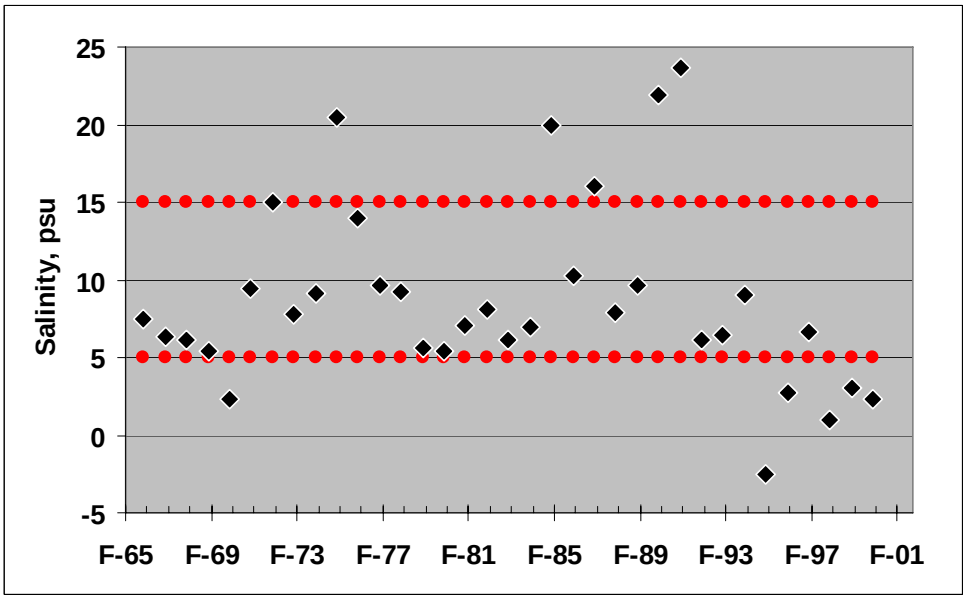


Figure 4. Joe Bay Annual Average Salinity Based on NSM 4.6.2 compared to the FBFKFS Salinity Targets (5-15 psu).

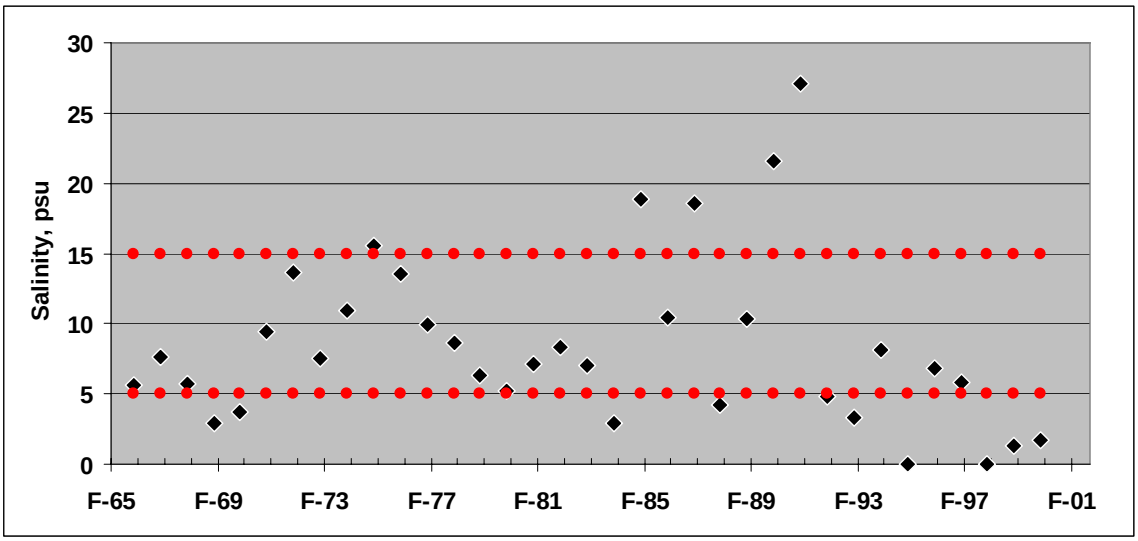


Figure 5. Little Madeira Bay Annual Average Salinity Based on NSM 4.6.2 compared to the FBFKFS Salinity Targets (15-25 psu).

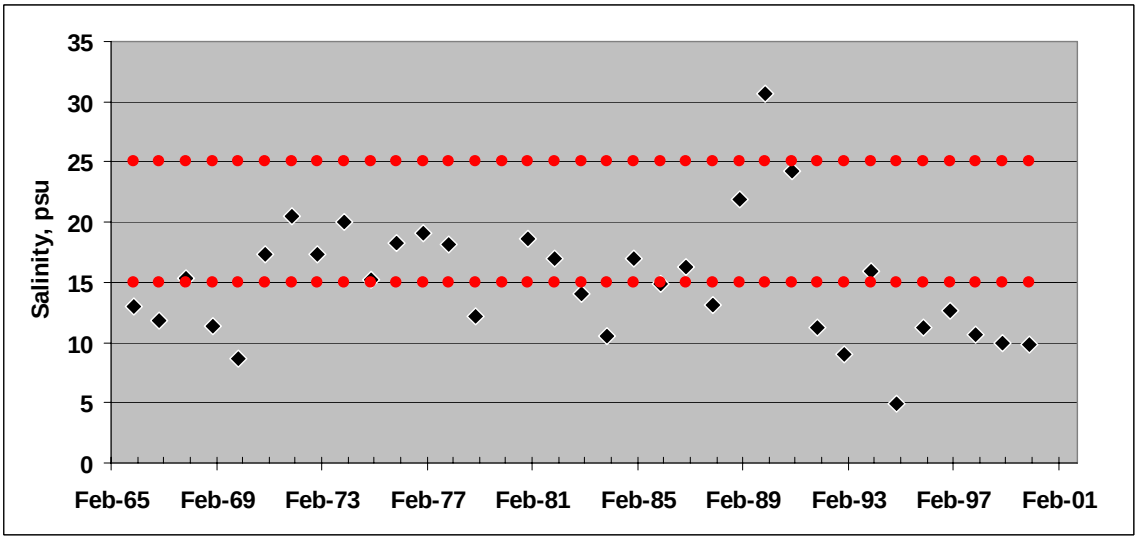


Figure 6. Duck Key Annual Average Salinity Based on NSM 4.6.2 compared to the FBFKFS Salinity Targets (15-30 psu).

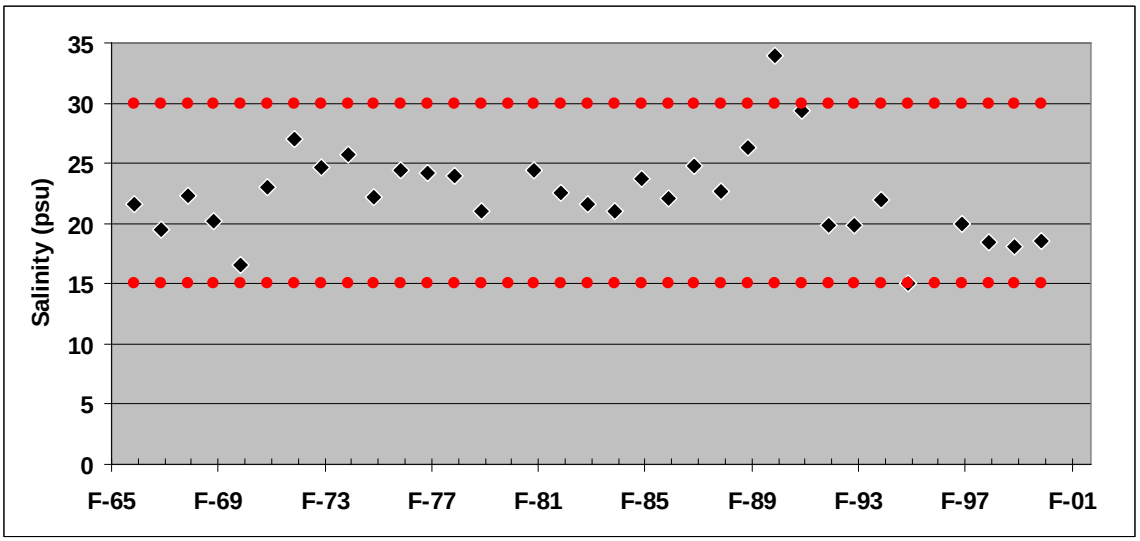


Figure 7. Butternut Key Annual Average Salinity Based on NSM 4.6.2 compared to the FBFKFS Salinity Targets (15-30 psu).

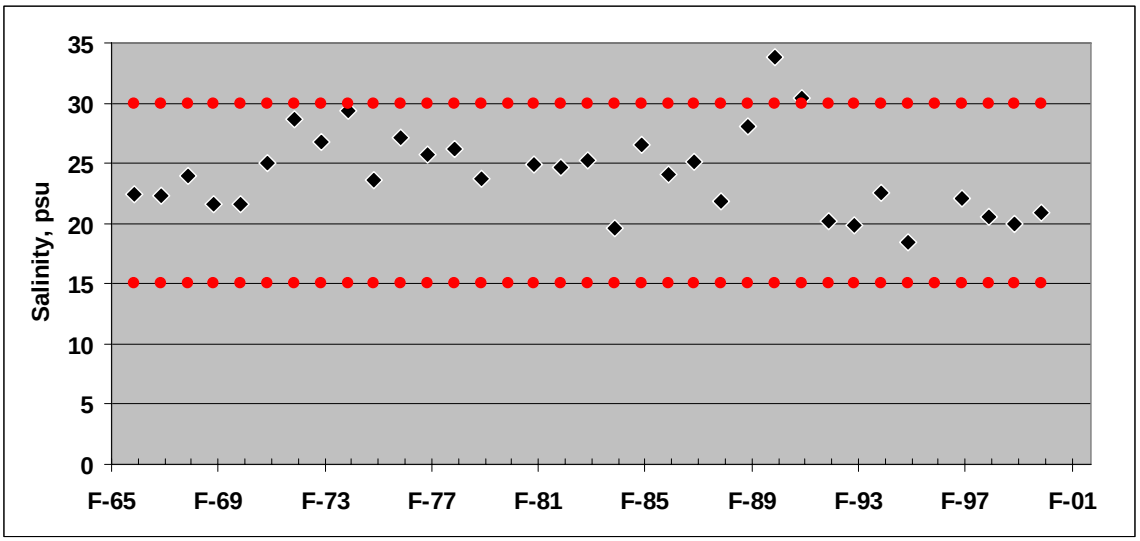


Figure 8. Butternut Key Annual Average Salinity Based on NSM 4.6.2 *Minus 3 psu* compared to the FBFKFS Salinity Targets (15-30 psu).

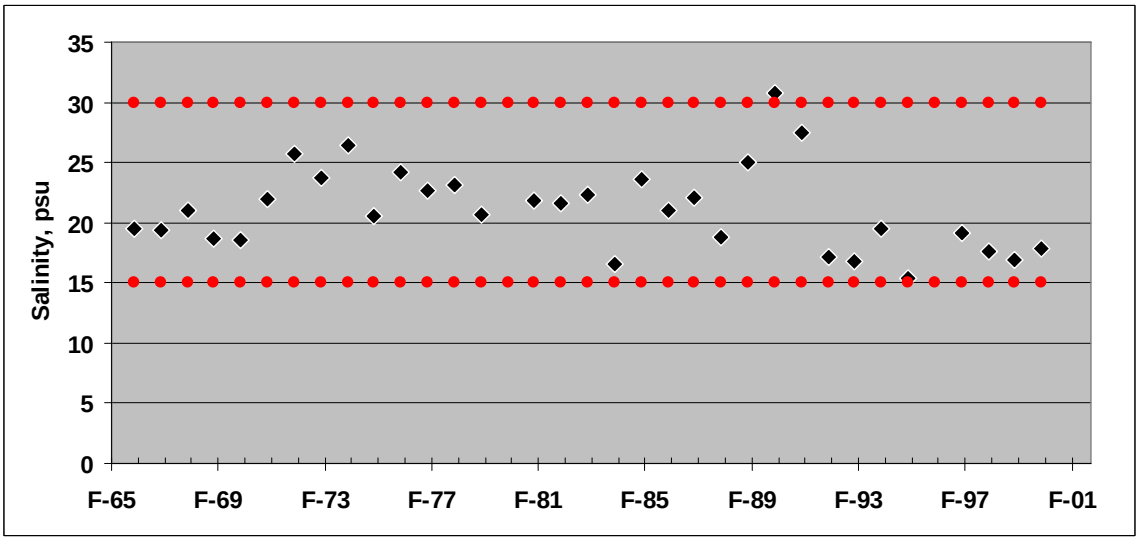


Figure 9. Whipray Basin Annual Average Salinity Based on NSM 4.6.2 compared to the FBFKFS Salinity Targets (25-35 psu).

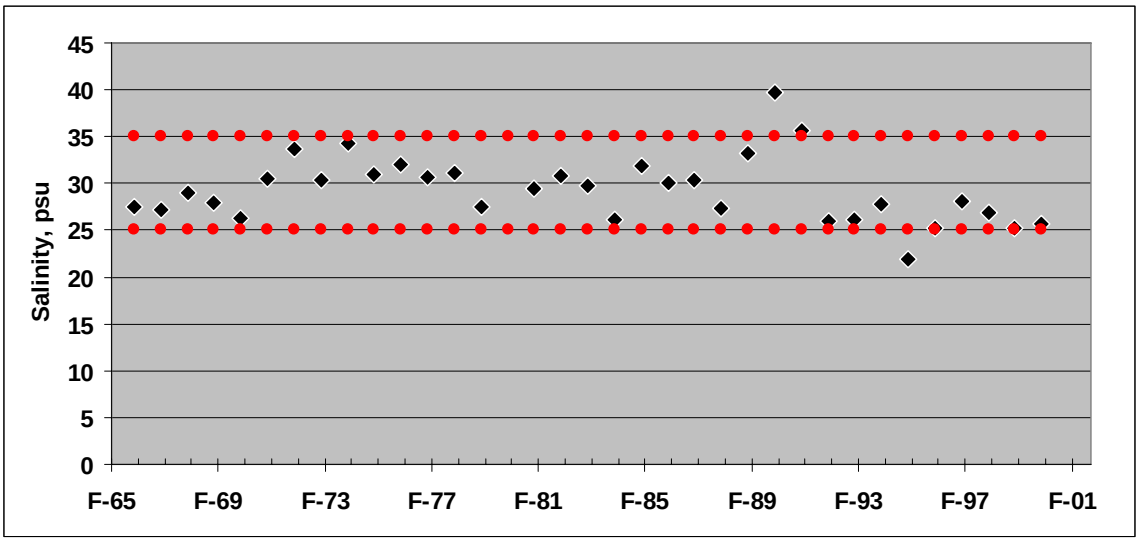


Figure 10. Terrapin Bay Annual Average Salinity Based on NSM 4.6.2 compared to the FBFKFS Salinity Targets (15-35 psu).

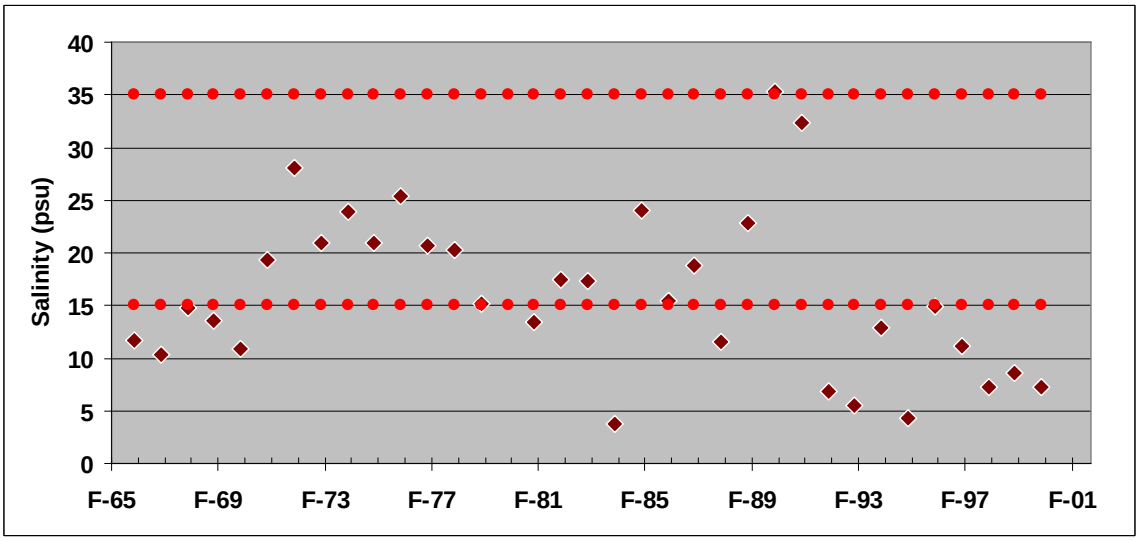


Figure 11. Garfield Bight Annual Average Salinity Based on NSM 4.6.2 compared to the FBFKFS Salinity Targets (15-35 psu).

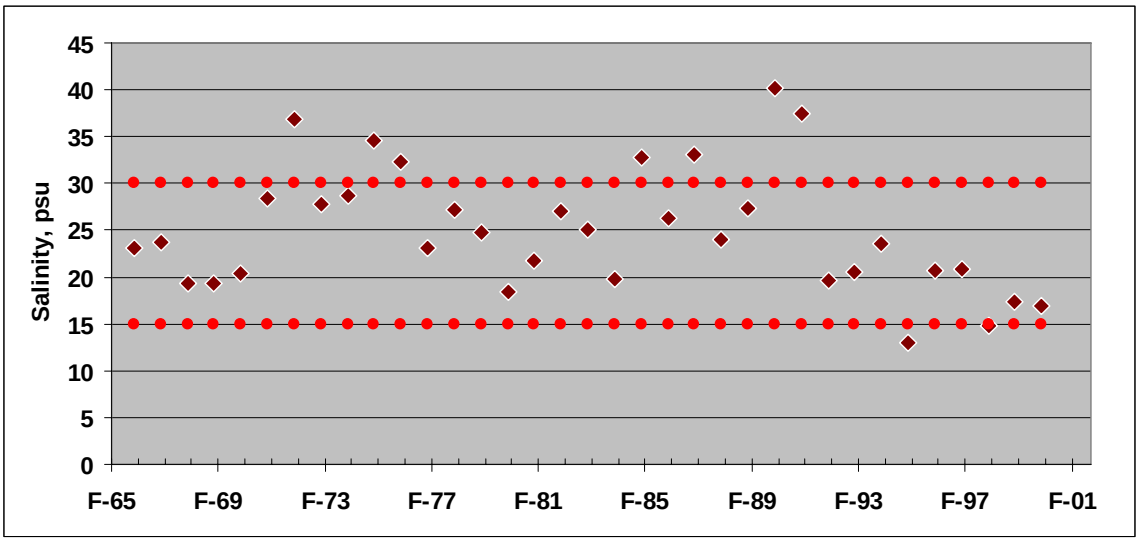


Figure 12. Garfield Bight Annual Average Salinity Based on NSM 4.6.2 *Minus 5 psu* compared to the FBFKFS Salinity Targets (15-35 psu).

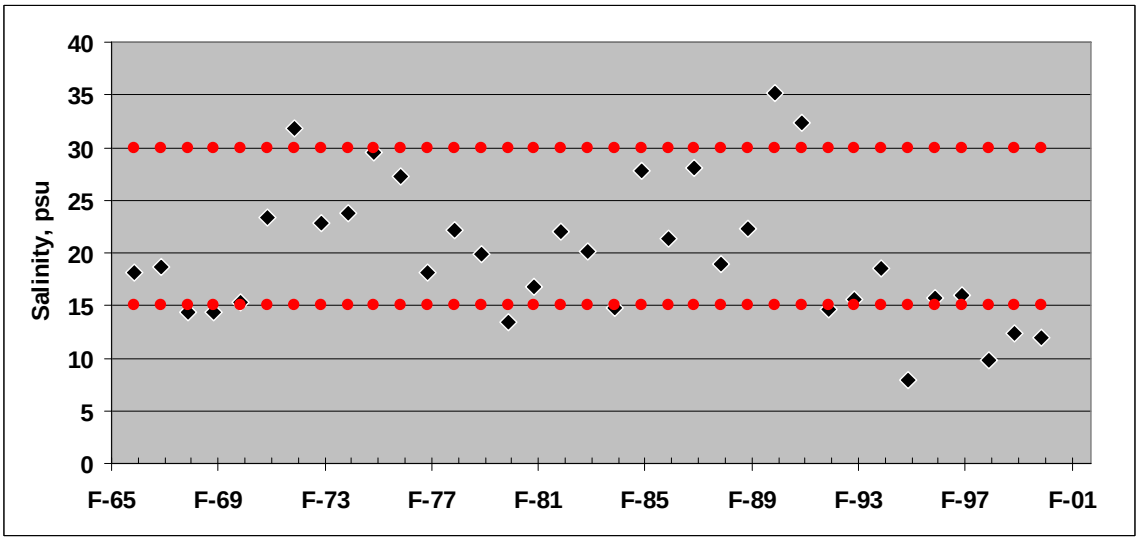


Figure 13. Bob Allen Key Annual Average Salinity Based on NSM 4.6.2 compared to the FBFKFS Salinity Targets (25-35 psu).

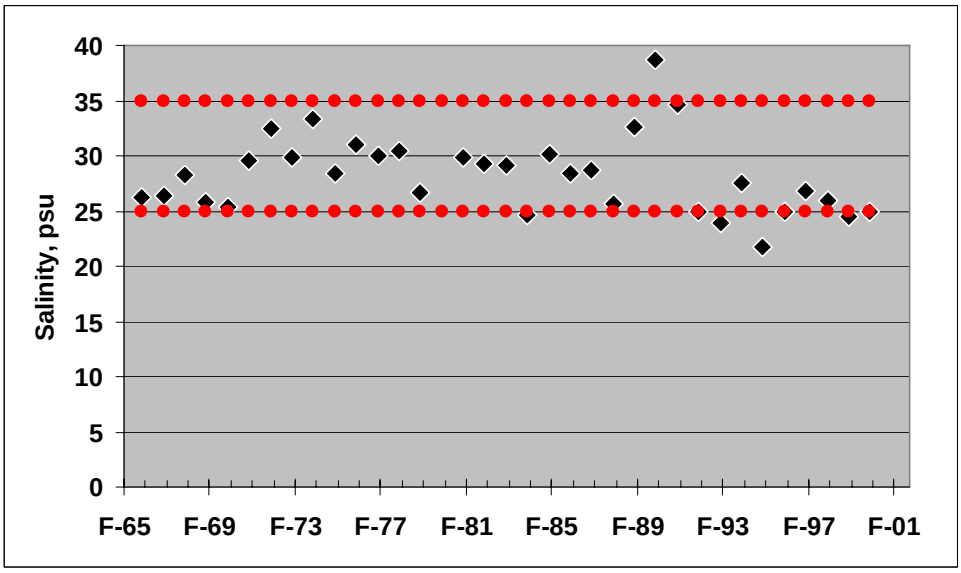


Figure 14. Little Blackwater Sound Annual Average Salinity Based on NSM 4.6.2 compared to the FBFKFS Salinity Targets (10-20 psu).

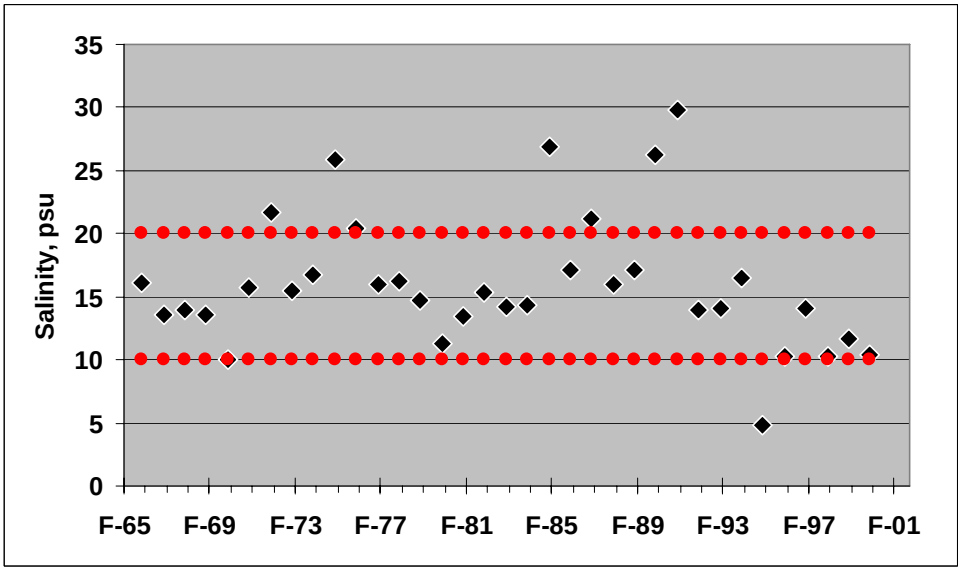
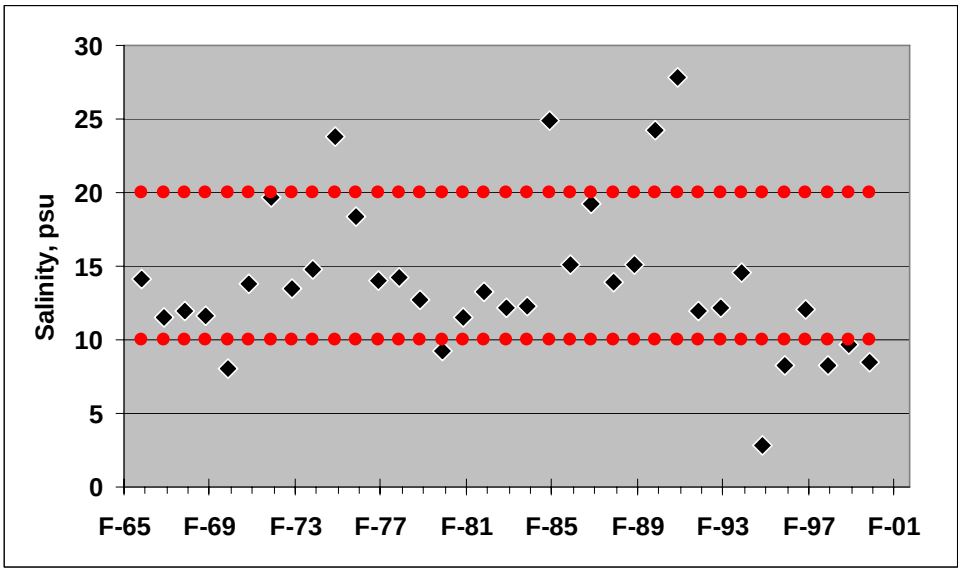


Figure 15. Little Blackwater Sound Annual Average Salinity Based on NSM 4.6.2 *Minus 2 PSU* compared to the FBFKFS Salinity Targets (10-20 psu).



Next, new simple (univariate) linear regression models were developed using observed stage to simulate salinity (dependent variable). For Little Madeira Bay, Terrapin Bay, Garfield Bight and Whipray Basin, Craighead Pond stage (CP) was the independent variable. For Joe Bay, Long Sound, and Little Blackwater Sound, stage at EVER7 was the independent variable. In the case of Butternut Key and Duck Key, it was not possible to fit regression models directly with stage values and produce an R^2 value that was acceptable. However, because the simple regression model for Long Sound produced a relatively large R^2 value, a two-step process was used to model these two stations. First, a simple regression model using Long Sound as the independent variable was developed at both stations. Then, the Long Sound/EVER7 model was used to estimate stage values for the purpose of estimating flow.

For Bob Allen Key, both CP and EVER7 were used separately as independent variables. For all models, the choice of which stage data to use for a particular linear regression model was based on physical location and the ability to develop a model with a reasonable adjusted R^2 value. For example, Joe Bay is downstream from EVER7 so EVER7 was used for the Joe Bay salinity model. In the case of Bob Allen Key it was possible to get reasonable relationships with both CP and EVER7, so this was done. The model parameters and adjusted R^2 value for each of the linear regression models is presented in Table 1.

Regression models were also developed between CP and P33 for the purposes of estimating the flow through the S12 structures (S12T) because the regressions

models using CP or EVER7 to simulate S12T flow were not capable of simulating the S12T flow as accurately as a two-step process.

The next step in the process (see Figure 1) was to develop regression models for flow from the observed daily stage values. Table 2 presents the model parameters and values of the Coefficient of Determination (R^2) for the flow regression models.

Table 1. Model Parameters and Coefficient of Determination (Adjusted) for Simple Linear Regression Models Relating Stage to Salinity and P33 Stage (units of salinity are PSU and units of stage are ft NGVD29)

Dependent Variable	Independent Variable	Lag	Coefficient	Intercept	R^2
Little Madeira Bay	**CP**	**5**	**-8.1**	**30.3**	**0.44**
Terrapin Bay	**CP**	**10**	**-17.5**	**46**	**0.57**
Garfield Bight	**CP**	**2**	**-13.9**	**47.7**	**0.51**
Whipray Basin	**CP**	**18**	**-6.6**	**42.8**	**0.31**
Bob Allen Key	**CP**	**35**	**-7.7**	**43.9**	**0.48**
Bob Allen Key	**EVER7**	**35**	**-9.8**	**56.1**	**0.42**
Joe Bay	**EVER7**	**7**	**-19.6**	**55.8**	**0.63**
Long Sound	**EVER7**	**9**	**-18**	**54.8**	**0.74**
Long Sound	**CP**	**6**	**-14.3**	**33.6**	**0.74**
Little Blackwater Sound	**EVER7**	**12**	**-16.8**	**57**	**0.68**
Butternut Key	**Long Sound***	**35**	**0.48**	**20.7**	**0.43**
Duck Key	**Long Sound***	**35**	**0.44**	**18.1**	**0.4**

* Simple regression with CP and EVER7 did not yield an acceptable relationship. A 2-step process using EVER and Long Sound models was used.

Table 2. Model Parameters and Coefficient of Determination (Adjusted) for Simple Linear Regression Models of Flow and Stage (units of stage are ft NGVD29, units of flow are cfs)

Dependent Variable	Independent Variable	Lag	Coefficient	Intercept	R^2
S18C	**EVER7**	**none**	**618.4**	**-1135**	**0.44**
TSB	**CP**	**none**	**181**	**-152**	**0.53**

The stage-based models in Table 1 were then used in an inverse calibration mode (Neter et al. 1991; Kashigan 1991) whereby the daily salinity values of the NSM 4.6.2 / MLR salinity models were utilized to estimate daily stage values at either CP or EVER7. For Butternut and Duck Keys, the daily NSM 4.6.2 / MLR salinity regime at each station was first used to estimate the daily salinity at Long

Sound, which was then used to estimate the daily stage at EVER7. As an example, the Joe Bay regression equation in Table 2 (Joe Bay as a function of EVER7) was solved for EVER7 (inverse calibration). Then the Joe Bay NSM 4.6.2 / MLR salinity regime was input to the inverse calibration model to estimate the stage at EVER7. This time series of EVER7 stage was then used with the S18C regression model in Table 3 to estimate the flow at S18C that meets the FBFKFS target for Zone 1. Table 4 compares the estimated average daily flows at TSB and S18C with the observed average daily flow values for the period December 24, 1991 – December 31, 2000, and presents the basis (model) of the estimate.

Table 4. Comparison of Simulated Flows (cfs) with Observed Flows (cfs) for the Period December 24, 1991 – December 31, 2000

station/model	FBFKFS Zone	target salinty - low (psu)	target salinity - high (psu)	NSM/MLR input	TSB	S18C	basis
Long Sound	1	5	15	minus 5		344	EVER7
Joe Bay	1	5	15	nsm		256	EVER7
Little Madeira Bay	1	15	25	nsm	228		CP
Duck Key	2	15	30	nsm		350	EVER7/Long Sound
Butternut Key	2	15	30	minus 3		254	EVER7/Long Sound
Whipray Basin	3	25	35	nsm	161		CP
Terrapin Bay	5	15	35	nsm	167		CP
Garfield Bight	5	15	35	minus 3	215		CP
Bob Allen	6	25	35	nsm	167	438	CP/EVER7
Little Blackwater Sound	14	10	20	minus 2		383	EVER7
Avg Daily Simulated Flow (cfs)					187.6	337.5	
Avg Observed Flow (cfs)					93.8	268.7	

The estimated average flow at TSB ranged from 161 to 228 cfs for the period of comparison, with a mean average of 187.6 cfs. By contrast, the observed average daily value over the period was 93.8 cfs. According to this modeling procedure the TSB flows over this period were about 50% of the flow needed to meet the FBFKFS targets, overall. At S18C, the estimated flows needed to meet the FBFKFS targets ranged from 254 – 438 cfs. The mean value of the average estimated daily flows was 337.5 cfs. The observed S18C flows over this period were therefore about 80% of the average daily flow needed to meet the targets.

In order to estimate the flow at S12T across Tamiami Trail it was necessary to utilize an additional regression model to estimate the daily stage at P33 from the daily stage at CP (R^2 = 0.63). Then, the daily flow at S12T can be estimated from the daily stage at P33 (R^2 = 0.65). Model parameters and the adjusted Coefficient of Determination for these models are presented in Table 5. Using

the link between (1) salinity and stage, and (2) stage and flow, the flow at S12T was estimated from the NSM / MLR-based daily salinity at Little Madeira Bay, Terrapin Bay, Garfield Bight, and Whipray Basin (Table 2). The range of estimated S12T flows to meet FBFKFS targets is 2095 to 2684 cfs, Table 6. The mean of these average daily flow values for S12T is 2426 cfs compared to the observed average daily flow of 1376 cfs over the comparison period. Based on this analysis the observed flows were about 56% of the FBFKFS target flows.

Table 5. Model Parameters and Coefficient of Determination (Adjusted) for Simple Linear Regression Models Relating CP Stage to P33 Stage, and P33 Stage to S12C Flow (units of P33 are ft NGVD29, units of flow are cfs).

Dependent Variable	Independent Variable	Lag	Coefficient	Intercept	R^2
S12T	**P33**	**none**	**2187.7**	**-13050**	**0.63**
P33	**CP**	**10**	**0.76**	**5.6**	**0.65**

Table 6. Comparison of simulated FBFKFS target S12T Flows (cfs) with Observed Flows (cfs) for the Period December 24, 1991 – December 31, 2000

station/model	FBFKFS Zone	target salinty low (psu)	target salinity high (psu)	nsm input	S12T (cfs)
Little Madeira Bay	**1**	**15**	**25**	**nsm**	**2684**
Whipray Basin	**3**	**25**	**35**	**nsm**	**2705**
Terrapin Bay	**5**	**15**	**35**	**nsm**	**2095**
Garfield Bight	**5**	**15**	**35**	**minus 3**	**2220**
Avg Daily Simulated Flow (cfs)					**2426**
Observed Flow (cfs)					**1376**

To evaluate the uncertainty in the estimated flows, observed values of salinity were input to the models to simulate flow, and compared to the observed flow values. Table 7 summarizes the results. In the worst case, the simulated TSB average daily flow over the period of this evaluation was within 16% of the observed TSB average daily flow, and the worst simulated S18C average daily flow was 56% of the observed flow. This worst case was for the Duck Key input, which required an additional step of using an intermediate (and extra) model for Long Sound. For S12T the worst case was a difference of just over 6%. If the simulated average daily TSB values from the Terrapin Bay, Little Madeira Bay, Long Sound, and Whipray Basin models are averaged, this mean simulated flow value is less than 2% different than the observed daily average flow value at TSB. For S18C, the mean calculated from Bob Allen Key, Joe Bay, Little Blackwater Sound, and Long Sound is within 8% of the observed S18C daily average flow value. The S12T mean value computed from Terrapin Bay,

Whipray Basin, and Garfield Bight models (with an equal number of values) is within 1% of the observed S12T flow.

Table 7. Summary comparison of observed flows to simulated flows produced by the univariate regression models from observed salinity values (p=simulated).

station/model	S18C	pS18C	TSB	pTSB	S12T	pS12T
Bob Allen Key	269	289				
Terrapin Bay			91.8	91.4	1376	1288
Butternut Key	269	156				
Duck Key	269	117				
Joe Bay	269	256				
Little Blackwater Sound	269	240				
Garfield Bight			99	84	1376	1410
Little Madeira Bay			91.8	77	1401	1340
Long Sound	269	198	91.8	94.9		
Whipray Basin			91.8	98	1376	1400

Figure 16 illustrates the differences at the daily time step between the simulated and observed flow values at TSB. For this comparison the observed salinity at Terrapin Bay was used to estimate the CP stage using the inverse calibration model. Then the simulated CP stage was used to simulate the TSB flows, which were compared to observed values. Figure 17 presents the same daily values averaged to monthly average values of daily flows. It can be seen that the modeled daily flows follow the overall trend of the observed daily flows for TSB. However, at the daily level the extreme high flow events in the observed flow record are muted considerably in the simulated time series. At the monthly time step, the average daily flow for a month tracks the observed average daily flow on a monthly basis much better.

As another example, the Long Sound univariate regression models were used with observed Long Sound salinity to estimate both S18C and TSB, because it was possible to develop good regression models relating Long Sound salinity to both EVER7 and CP. Figure 18 presents the estimated TSB flows and the observed TSB flows, and Figure 19 presents the same values for S18C. From the plots it appears that the simulated daily values of TSB more closely trace the observed daily flow values, compared to S18C. In both cases the models are not capable of predicting the highest flow values at either location.

To further evaluate the uncertainty in the models, the simulated overall average daily flow values at TSB and S18C from the above Long Sound input were compared to the observed values for an equal number of days, i.e. if simulated or observed values of either were missing that day was not included in the average. The results are presented in Table 8. As can be seen the average of the

simulated TSB values compares very well with the observed TSB values, within 3% of the observed. While the simulated S18C average values did not compare quite as well, the simulated average S18C is still within 6% of the observed average value. However, the Pearson's Correlation Coefficient for the daily values in Table 8 shows that even though the overall average values compare well, there are still considerable daily values differences between simulated and observed values.

Figure 16. Observed versus simulated daily values for TSB from Terrapin Bay salinity. Simulated values were based on observed values of Terrapin Bay salinity as original input to regression model to simulate CP, then CP is used as input to TSB flow model.

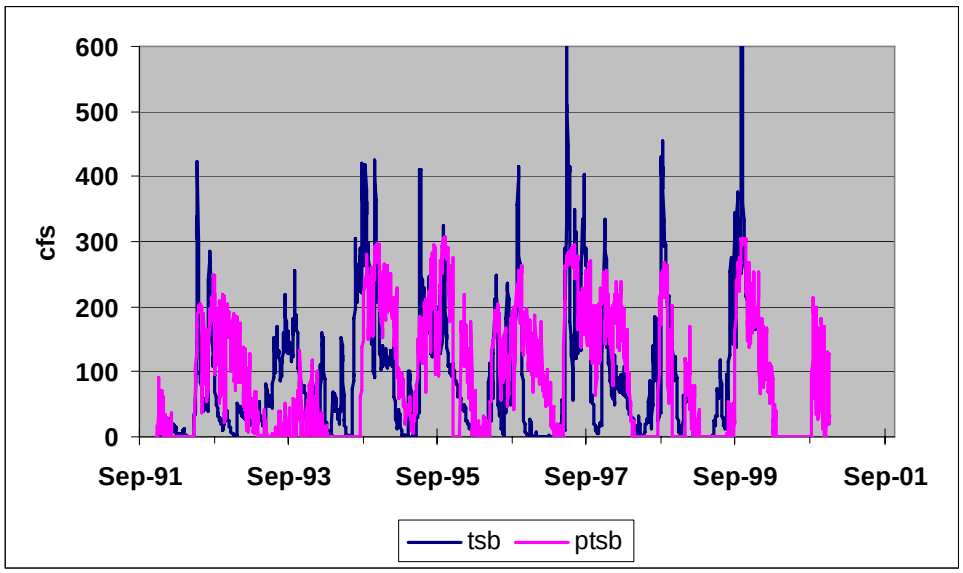


Figure 17. Same as above, except the values in Figure 16 have been averaged to monthly values.

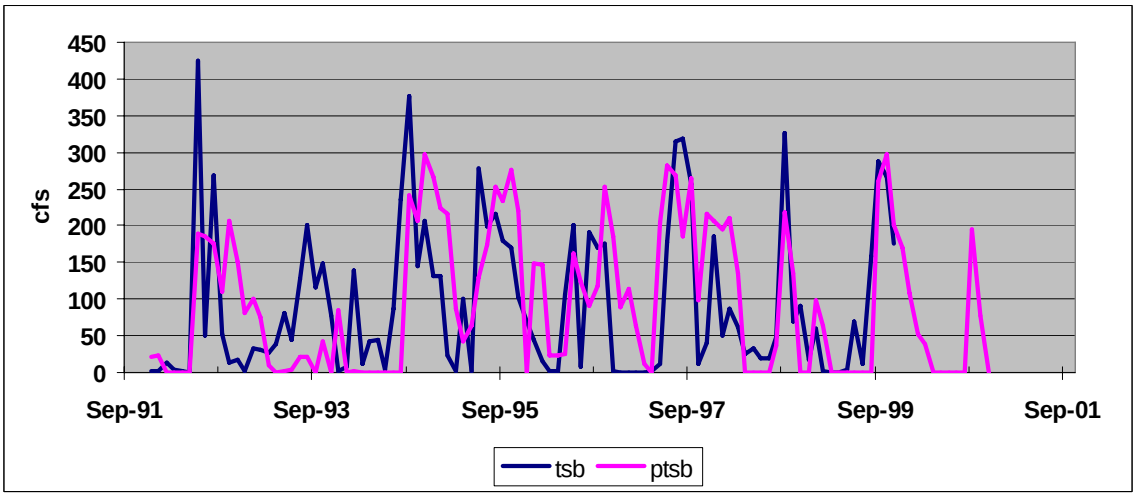


Figure 18. Observed versus simulated daily values for TSB from Long Sound salinity. Simulated values based on observed values of Long Sound salinity were used as original input to simulate CP, then CP is used as input to TSB flow model.

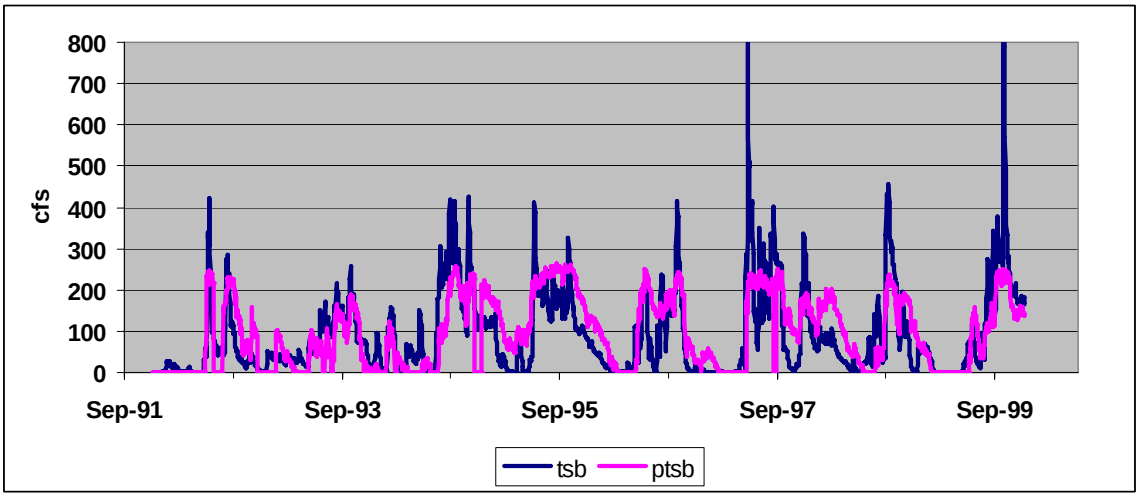


Figure 19. Observed versus simulated daily values for S18C based on Long Sound salinity. Simulated values based on observed values of Long Sound salinity as original input to simulate EVER7, then EVER7 is used as input to S18C flow model.

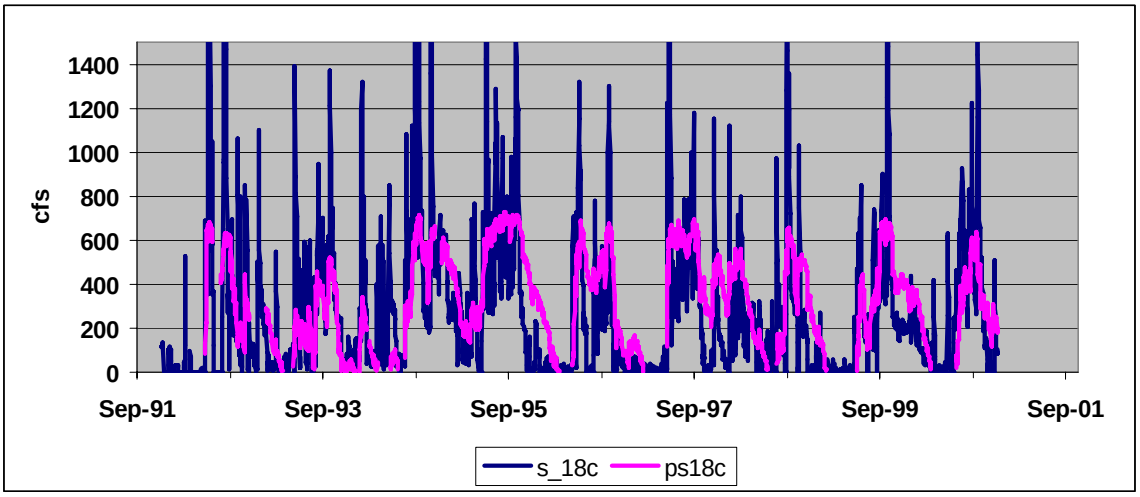


Table 8. Comparison of observed and simulated average daily flow for TSB and S18C. Simulated values from models using observed Long Sound salinity to estimate either CP (TSB) or EVER7 (18C).

	Observed	Simulated	Pearson's Corr. Coeff.
TSB	**91.8**	**94.9**	**0.68**
S18C	**332.2**	**352.3**	**0.54**

III. Discussion

This proof-of-concept exercise has shown that a suite of linear regression models and statistical techniques can be used to estimate the flow requirements at three locations in the Everglades hydrologic system for a particular salinity regime. In this case, the Florida Bay / Florida Keys Feasibility Study salinity targets were used with NSM 4.6.2 and MLR salinity models (Marshall 2004, 2005) to develop a FBFKFS daily salinity regime for eight locations in Florida Bay that were spread across several FBFKFS zones. Observed values of salinity and stage were regressed to produce salinity (independent variable) as a function of stage. Then the models were turned around to use stage as the independent variable. In this manner simulated time series were generate at CP and EVER7 from the FBFKFS salinity regimes. From CP and EVER7, flow at TSB and S18C are estimated using other regression models. An additional set of regression models produces stage at P33 from stage at CP, then P33 is used to estimate flow at S12T.

The daily simulated flow values can show large differences at times when compared to daily observed values, particularly at high flow values. However, the statistical power that comes from large data sets, such as the ones used in this study (about 2500 values used for model development and about 12,500 for FBFKFS runs) allows these simulations to be used best in a monthly average or daily average value mode. In this mode, with simulations made daily and then averaged to monthly mean or daily average value over the simulation period, the modeling procedure appears to work well. This means that the simulated daily values should be interpreted as long-term averages, over periods that have experienced a wide range of climatic conditions.

When the modeling procedure is used to simulate the flows required to meet various FBFKFS targets, the difference between the simulated FBFKFS target flows and observed flows is substantial. The TSB, S18C, and S12T observed average values were found to be about 50%, 80%, and 56%, respectively, of the FBFKFS target flows. This reflects the findings of Smith, et al (1988) using fluorescent techniques with coral banding on a piece of coral from near Peterson Key. These authors found that flow to northeast Florida Bay was reduced about 56% from historical conditions.

Future use of this technique appears to be planning-level decision making, such as for use with paleosalinity information, CERP evaluations using the 36-year dataset, and for water reservation purposes.

IV. Conclusions

From the above analysis, it is concluded that the procedure presented above using a suite of univariate linear regression models can be successfully used to simulate long-term average flow conditions (monthly mean and annual average), even though the residuals at the daily level of computation are often large. The procedure takes advantage of the statistical power in a large number of observations with which to develop models and a large number of input values for simulations.

It is also concluded that the current flows through the S-18C and S-12T structures and TSB are well below the flows needed to meet the FBFKFS salinity targets. However, the flows currently delivered by the C-111 Canal to the S-18C structure are closer to the target than the flows needed at S-12T and TSB, reflecting the flood management function.

V. References Cited

Kashigan, S.K. 1991. Multivariate Statistical Analysis. Radius Press. New York. pp117-192.

Marshall III, F. E.; D. Smith; and D. Nickerson. 2004. Using Statistical Models to Simulate Salinity Variation and Other Physical Parameters in North Florida Bay. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F.E. 2005a. RECOVER Southern Estuaries Performance Measues: Identification of Hydrology-Salinity Relationships for Coastal Estuaries and Analysis of Interim CERP Update Scenarios. Environmental Consulting & Technology, Inc. New Smyrna Beach, Florida.

Marshall, F.E. 2005b. Statistical Salinity Models for Florida Bay and the Southwest Gulf Coast, Everglades National Park – Continuation Project. Critical Ecosystems Studies Initiative Proposal to Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Neter, J.; W. Wasserman; M. H. Kutner. 1990. Applied Linear Statistical Models. Richard D. Irwin, Inc., Boston, Massachusetts.

Ogden, J. C., and S. M. Davis, K. J. Jacobs, T. Barnes, and H. E. Fling. 2005. The use of conceptual ecological models to guide ecosystem restoration in south Florida. *Wetlands* 25:795-809.

FINAL TASK REPORT

TASK 10 - THE USE OF SICS/TIME OUTPUT WITH MLR SALINITY MODELS

Cooperative Agreement Number CA H5284-05-0006
Between
The United States Department of the Interior National Park Service
Everglades National Park
And
Cetacean Logic Foundation, Inc.

Frank E. Marshall, III
Principal Investigator

DeWitt Smith
Project Coordinator

April 13, 2008

CESI Project No. CA H5284-05-0006 for Everglades National Park

FINAL TASK REPORT

TASK 10 - THE USE OF SICS/TIME OUTPUT WITH MLR SALINITY MODELS

I. Introduction

A. General

This report describes the activities that were completed for a task in the current Critical Ecosystem Studies Initiative (CESI) project that was originally initiated by Cetacean Logic Foundation, Inc. (CLF) for Everglades National Park (ENP) in 2002. The goal of this on-going CLF research is the characterization and simulation of the salinity regime in Florida Bay and the estuaries along southwest coast of the Gulf of Mexico (Gulf) within the Park. It is intended that this research yield information regarding the link between the downstream salinity in the estuaries of ENP and the upstream freshwater hydrology of the Everglades.

This report describes the steps that were taken by the Principal Investigator to implement the SICS/TIME model with multivariate linear regression (MLR) salinity models. The TIME output was also used with the FATHOM mass balance model as part of Task 2 of this CESI project. This task report describes the coordination between the Principal Investigator, ENP, the US Geological Survey, and others, and the use of the SICS/TIME output with MLR salinity models.

B. Task Objectives and Evaluation Methods

The objective of this task is presented below as it appears in the modified contract for this CESI project approved by ENP:

> "To generate simulations and predictions, the MLR salinity models require input data that includes the water level (stage) at a number of locations in the Everglades. In a similar manner, FATHOM, the mass balance model requires flow into specific basins. The data can be observed values as for a verification run, or synthetic data estimated by models or other means. For most CERP evaluations performed to-date, the synthetic water level data has been supplied by the South Florida Water Management Model (SFWMM) also known as the 2X2 Model for the grid size (2 miles), or a similar model, the Natural System Model (NSM). There have been a number of revisions and upgrades to the SFWMM and the NSM, and replacements for both models appear forthcoming.
>
> One modeling approach that has been in development for the last few years is the USGS SICS/TIME model. The USGS has constructed surface-water/ground-water models of the Everglades and parts of Florida Bay, utilizing SICS, a hydrodynamic surface-water model that is a coupling of SWIFT2D and SEAWAT (Langevin et al, 2004). The outputs for the model

are water level (stage), flow, and salinity. The model domain for SICS encompasses the Taylor Slough wetland area and the northern and central portions of eastern Florida Bay. When coupled with the USGS TIME model, the spatial domain will eventually extend into Shark Slough. It is understood that ENP has an executable copy of the model that can be used with specified input data to produce simulations of flow, stage, and salinity. At the current time, the model and, particularly the input data, are only set-up for running over the period 1996-2002.

For this task, the ENP version of the SICS/TIME model will be obtained and reviewed to understand the input needs, the details of running the model, and the form and content of the output. A meeting will be scheduled with the USGS to discuss the ENP version, particularly the input data, and the requirements for extracting stage and flow data from the model output. The input data file provided to ENP will be used to obtain a test simulation that can then be reviewed to determine if the model is being properly applied. If possible, the calibration period stage data produced by SICS/TIME will be used with MLR salinity models to simulate salinity for Joe Bay, Little Madeira Bay, Terrapin Bay, and Garfield Bight.

Flow data will also be extracted from the SICS/TIME output, and an input file will be prepared that can be used for a FATHOM model run for the period of the calibration/verification. If possible, the FATHOM calibration/verification period model run would be performed by ENP staff using the most up-to-date FATHOM model available to ENP.

It is anticipated that it will require more effort to process the flow data into a suitable input file for use with FATHOM than the effort required for the MLR salinity models, though at the time of preparation of this scope of work the level of effort required can not be estimated until the model information is reviewed. Therefore, it may not be possible to complete the preparation of a FATHOM input data file given the number of man-hours dedicated to this task. In that event, a future task in a future scope of work may need to include finalizing the input file and one or more FATHOM runs.

This task is intended to get the ENP version of the SICS/TIME model up and running, and to show that the SICS/TIME output can be used with the MLR salinity models and with FATHOM. Coordination requirements with the USGS will be determined so that future efforts can be devoted to using the SICS/TIME output for CERP water delivery alternatives evaluations.

The deliverable for is a task report that describes the ENP version of the SICS/TIME model, and a test run of the model using a test input file provided by ENP. MLR model runs will be made using SICS output to produce simulations for Joe Bay, Little Madeira Bay, Terrapin Bay, and Garfield Bight that can be compared to simulations previously made with 2X2 Model output. If possible given budget constraints, an input file from the SICS output will be prepared that can be input to the ENP FATHOM model by ENP staff to obtain salinity simulations for the calibration period."

II. Background

This investigation makes use of two modeling systems (SICS and TIME) developed by the USGS for simulating the hydrology and hydraulics of the Everglades (Langevin et al 2004a). The output of these modeling systems includes water depth values that can be converted to stage and used with the MLR salinity models to simulate salinity in Florida Bay. The SICS model represents the southeast coastal region that is connected through the mangrove zone with Florida Bay. The same modeling system has been expanded to the west and north using the TIME domain to represent the Shark Slough flows.

The USGS began the development of the Southern Inland and Coastal Systems (SICS) model in the mid-1990s. The SICS model domain includes the Taylor Slough area with a 305-meter grid resolution (Figure 1). The USGS expanded the SICS model to include both Taylor and Shark River Sloughs through development of the Tides and Inflows in the Mangroves of the Everglades (TIME) model. The TIME model domain encompasses a larger area than SICS but uses a 500-meter grid. Therefore, TIME is coarser in resolution than SICS but covers a larger area of south Florida (fig. 1).

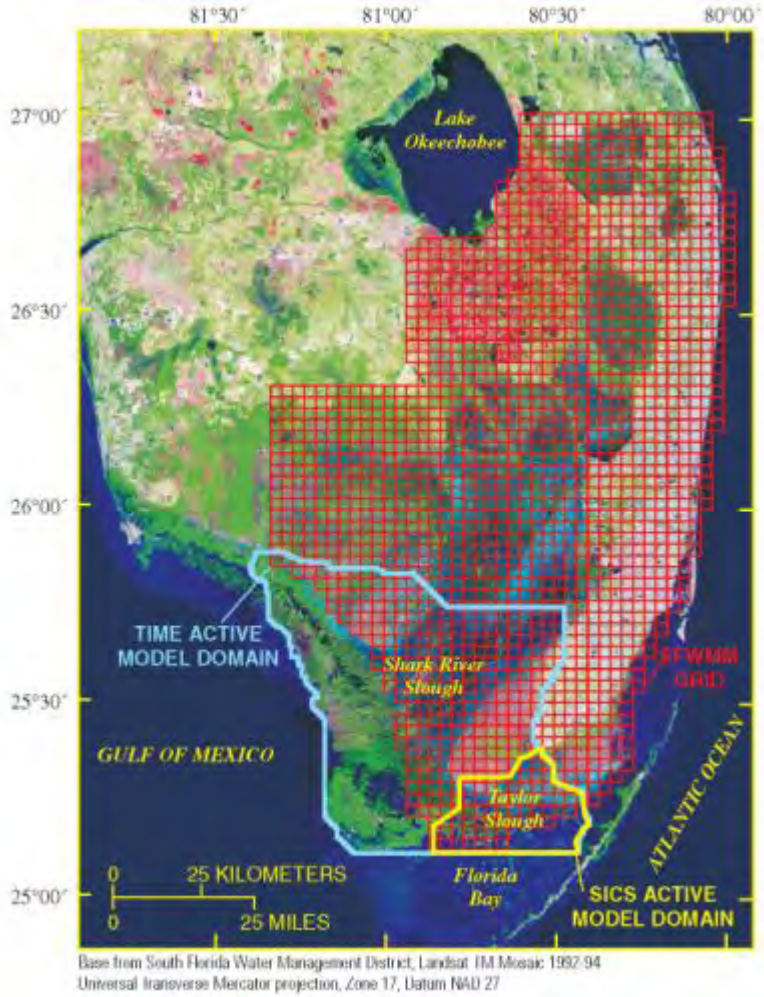


Figure 1. Map of southern Florida showing the active model domains of SICS and TIME and the SFWMM grid.

Note: Figure 1 was copied from Langevain et al, 2004a.

III. Methods

ENP obtained a copy of the SICS code in 2006 which was provided to the PI. The PI analyzed some examples of SICS output before the attention in the modeling effort turned to the further development and use of the TIME model. The PI was able to produce stage output from the SICS model using post-processing programs provided to ENP and the PI by the USGS. When the TIME model began to be used, the SICS post-processing tools were no longer applicable and the PI relied on ENP to provide post-processing files for input to the MLR salinity models.

A calibration / verification run from TIME (157 run) was made available to compare with observed data for stage. ENP provided stage output from the TIME 157 model run that has been corrected for model bias, relative to the NGVD29 datum. This adjusted TIME 157 run output (stage) was compared to the observed stage data, then used with MLR salinity models to produce salinity simulations, as present below.

The MLR salinity models used for simulating salinity using the adjusted TIME 157 run stage values for Joe Bay, Little Madeira Bay, and Terrapin Bay (Marshall et al, 2004) and for Garfield Bight (Marshall, 2005) are as follows:

JOE BAY = 37.1 - 3.1CP - 3.5 EVER6[lag6] - 10.5 E146[lag6] - 0.19 uwndkw - 0.09 uwndkw[lag2] - 0.1 vwndkw - 0.16 vwndmia[lag1], Adj-R^2 = 0.74

LITTLE MADEIRA BAY = 66.4 – 3.6 CP[lag2] - 6.3 P33[lag2] - 0.83(P33-NP206) – 0.21 uwndkw + 0.15 uwndmia - 0.14 vwndmia[lag1] + 0.8 kwwatlev[lag2], Adj-R^2 = 0.56

TERRAPIN BAY = 106.9 - 6.3 CP[lag1] - 11.1 P33[lag2] - 0.45 uwndkw - 0.23 uwndkw[lag1] - 0.2 uwndkw[lag2] - 0.14 vwndkw[lag2] + 0.46 uwndmia + 1.9kwwatlev[lag2], Adj-R^2 = 0.76

GARFIELD BIGHT= 56.1 - 9.2 CP[lag1] - 4.6 NP62[lag1] -0.46 UWNDKW[lag1] - 0.48 UWNDKW[lag4] + 0.35 UWNDMIA[lag1] + 0.64 UWNDMIA[lag4], Adj-R^2 = 0.68

IV. Results

Figure 2 presents the bias-adjusted TIME 157 stage data that were used for the MLR salinity model input. It can be seen that there are periods with missing data because the TIME data are output for the surface water module, only (see discussion below). When the conditions become dry enough for the simulated water level to drop below the elevation of the monitoring station, there is no TIME 157 output. Figures 3 through 8 present the comparison of the TIME 157 model output (stage) to the existing stage data for the period January 1, 1996 through December 31, 1999. Figures 9 through 12 present comparison plots between for salinity produced by MLR salinity models using the bias-adjusted TIME 157 output and observed data for Joe Bay, Little Madeira Bay, Terrapin Bay, and Garfield Bight, respectively.

Figure 2. Comparison of all stage values for the TIME R150 model run (as provided by ENP) used for salinity modeling.

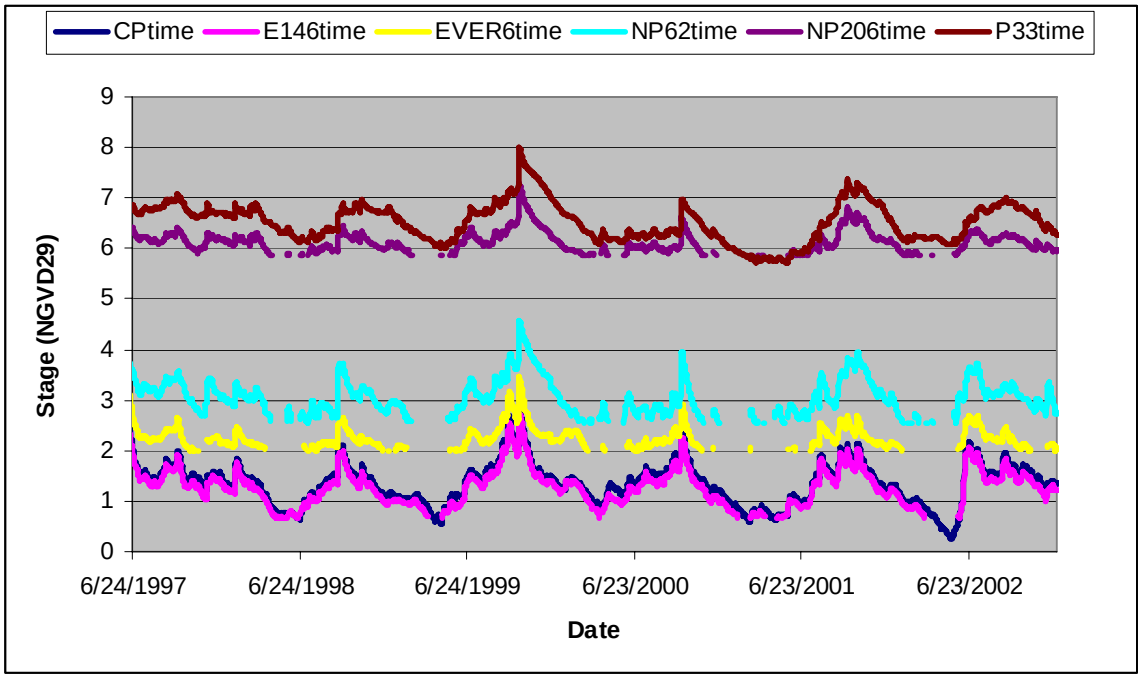


Figure 3. Comparison of observed data and stage values produced by the TIME 157 model run with CERP2000 values for CP.

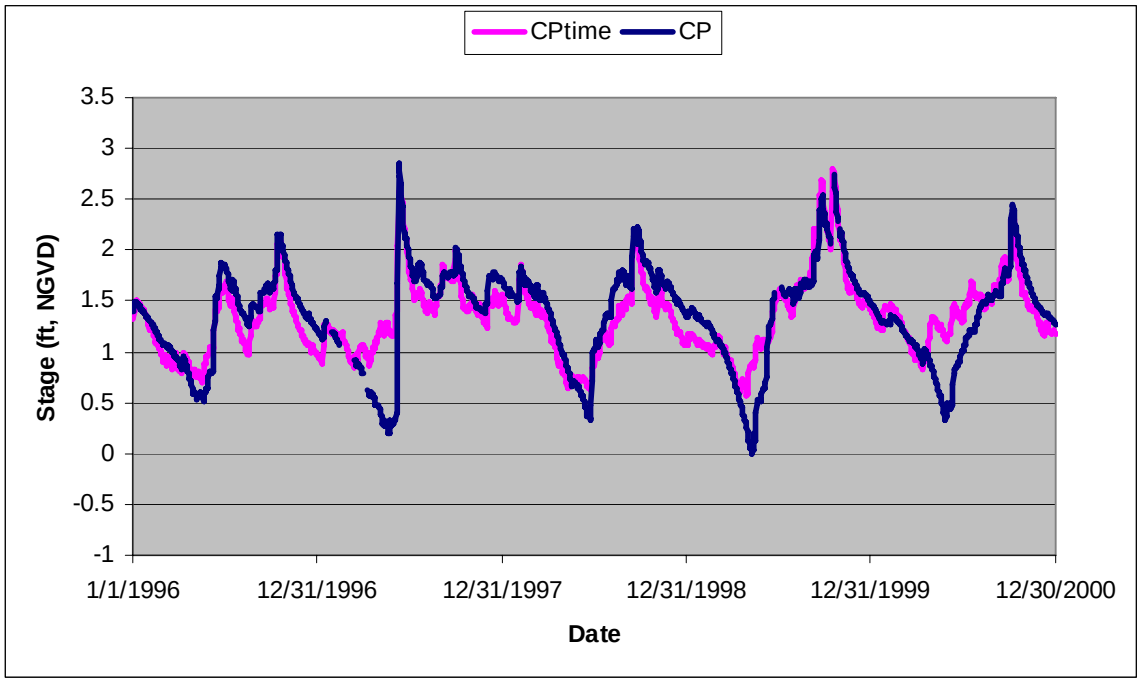


Figure 4. Comparison of observed data and stage values produced by the TIME 157 model run.

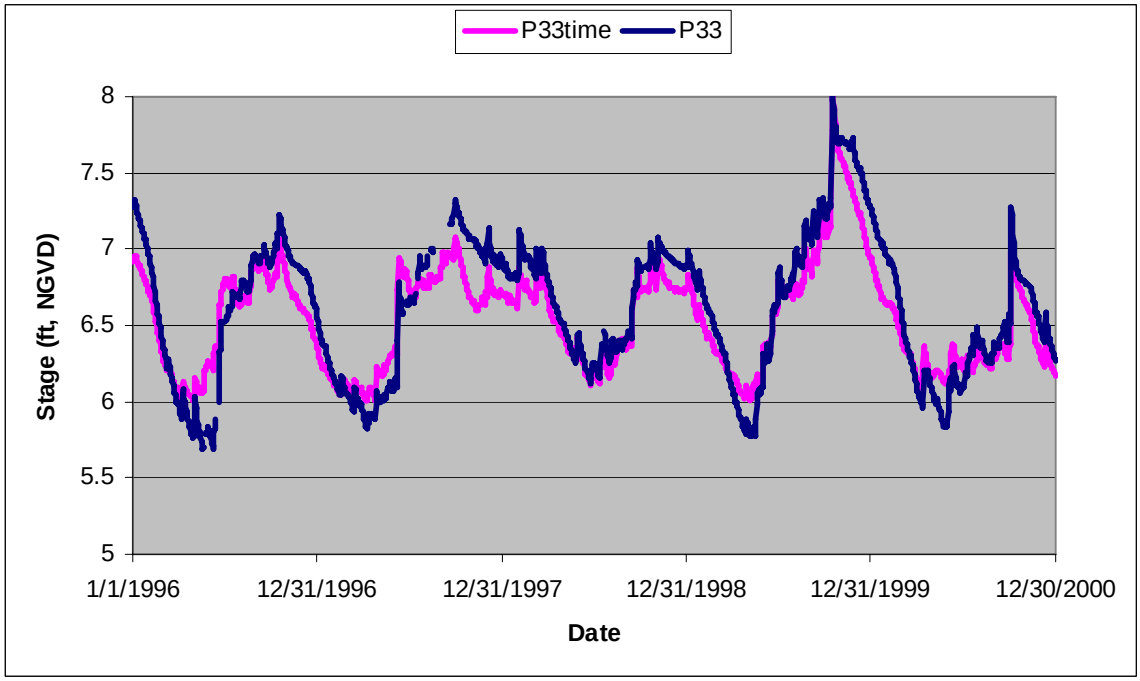


Figure 5. Comparison of observed data and stage values produced by the TIME 157 model run.

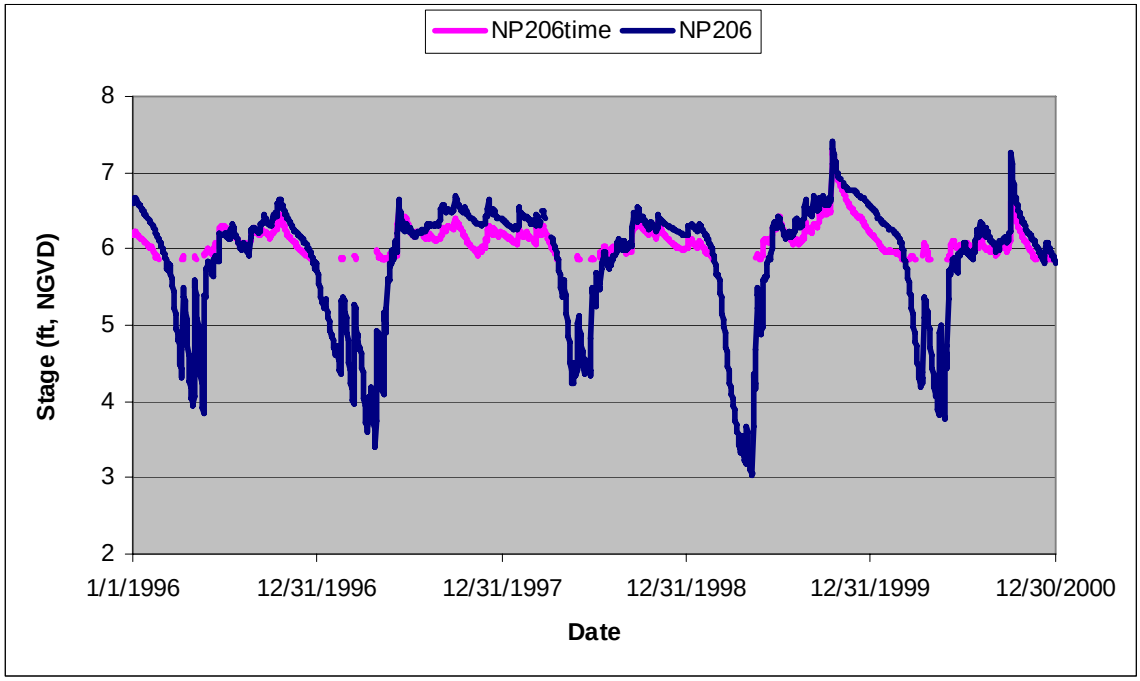


Figure 6. Comparison of observed data and stage values produced by the TIME 157 model run.

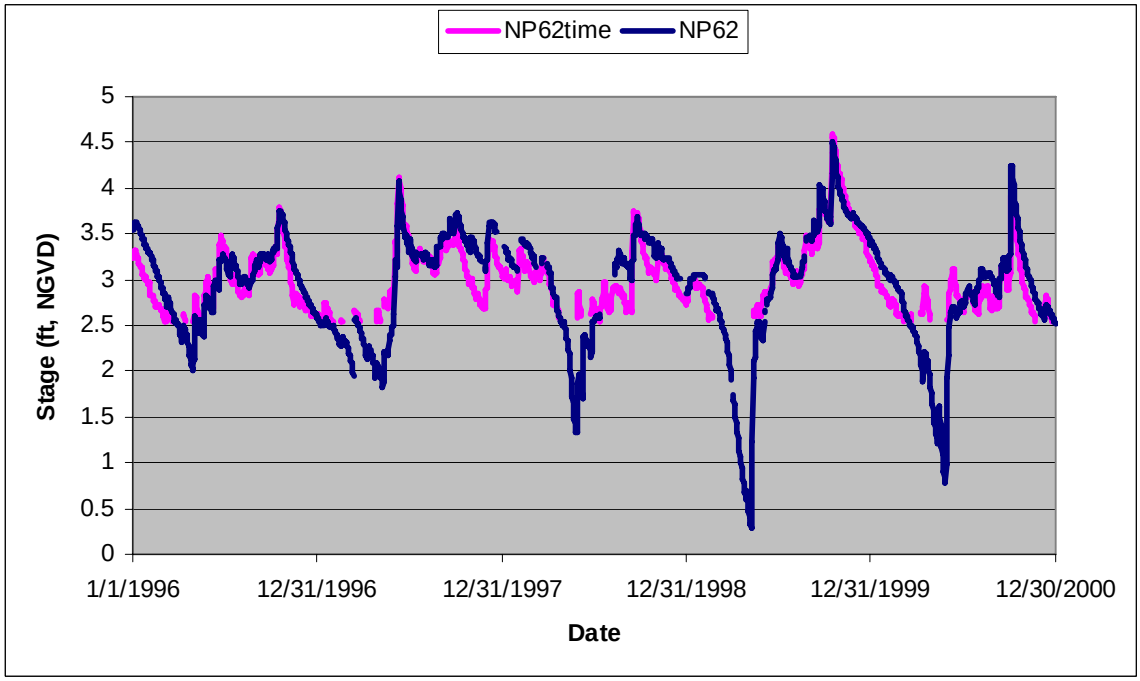


Figure 7. Comparison of observed data and stage values produced by the TIME 157 model run.

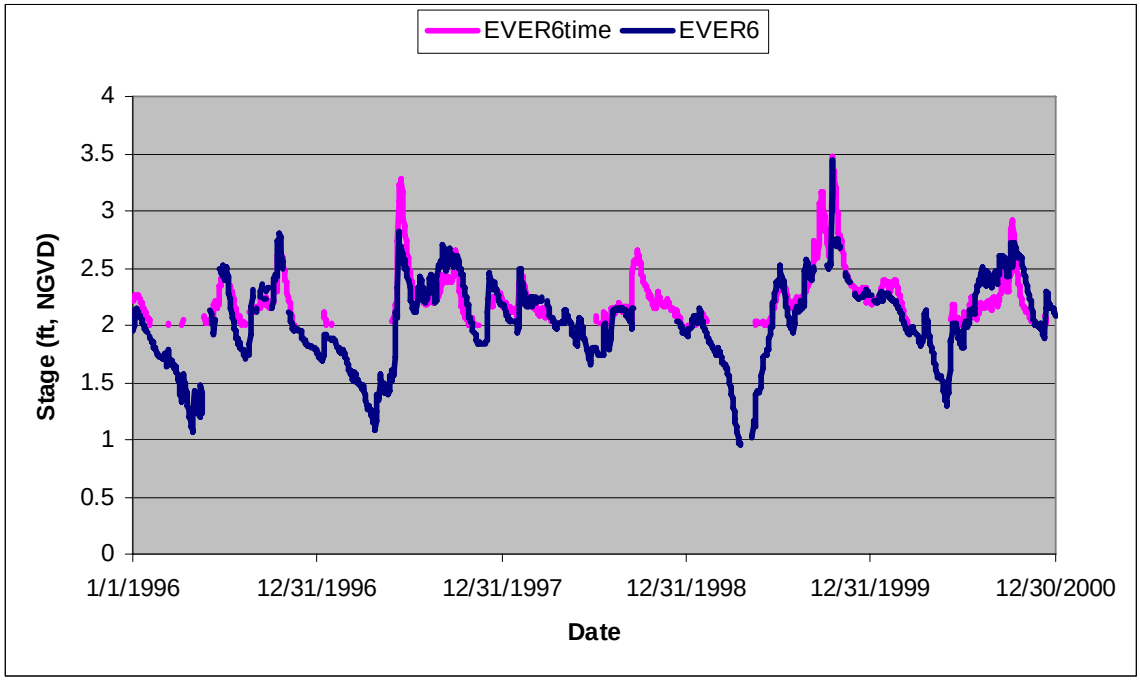


Figure 8. Comparison of observed data and stage values produced by the TIME 157 model run for E146.

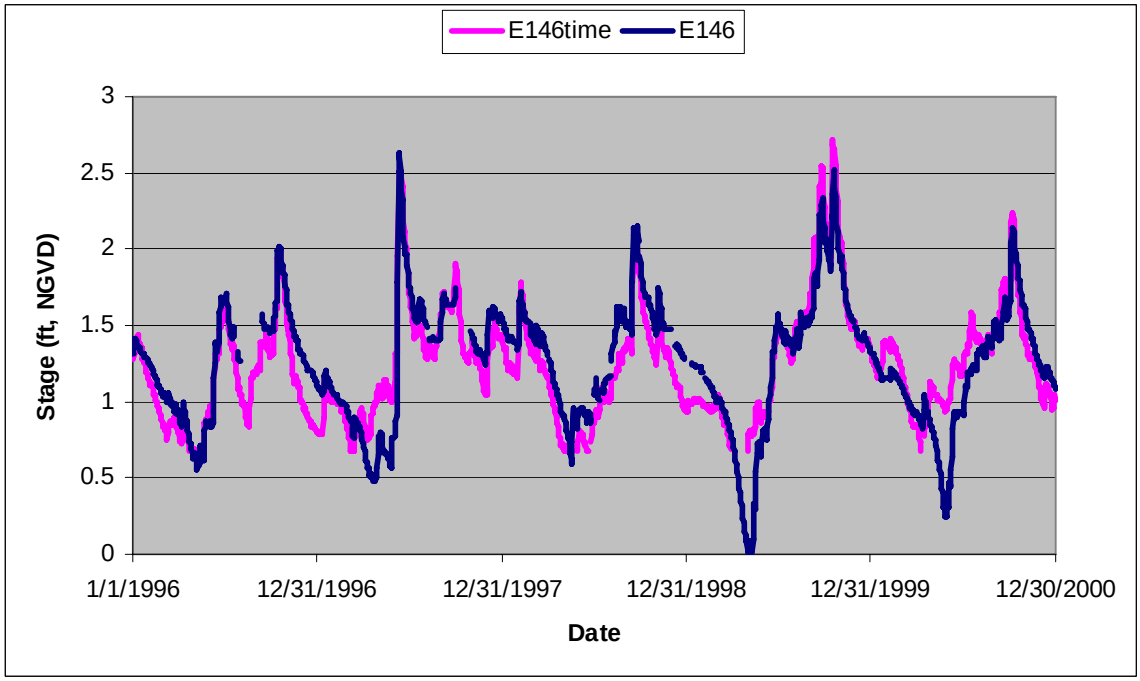


Figure 9. Comparison of observed data and salinity simulations for Joe Bay produced by MLR salinity models (Marshall et al 2004) using stage values from the TIME 157 model run.

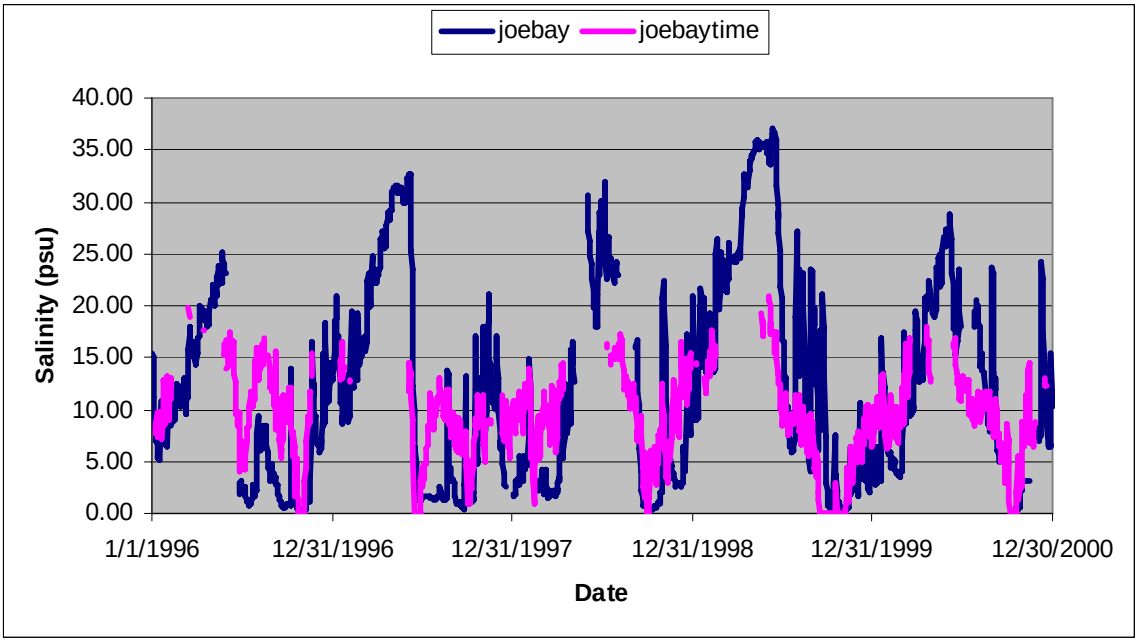


Figure 10. Comparison of observed data and salinity simulations for Little Madeira Bay produced by MLR salinity models (Marshall et al 2004) using stage values from the TIME 157 model run.

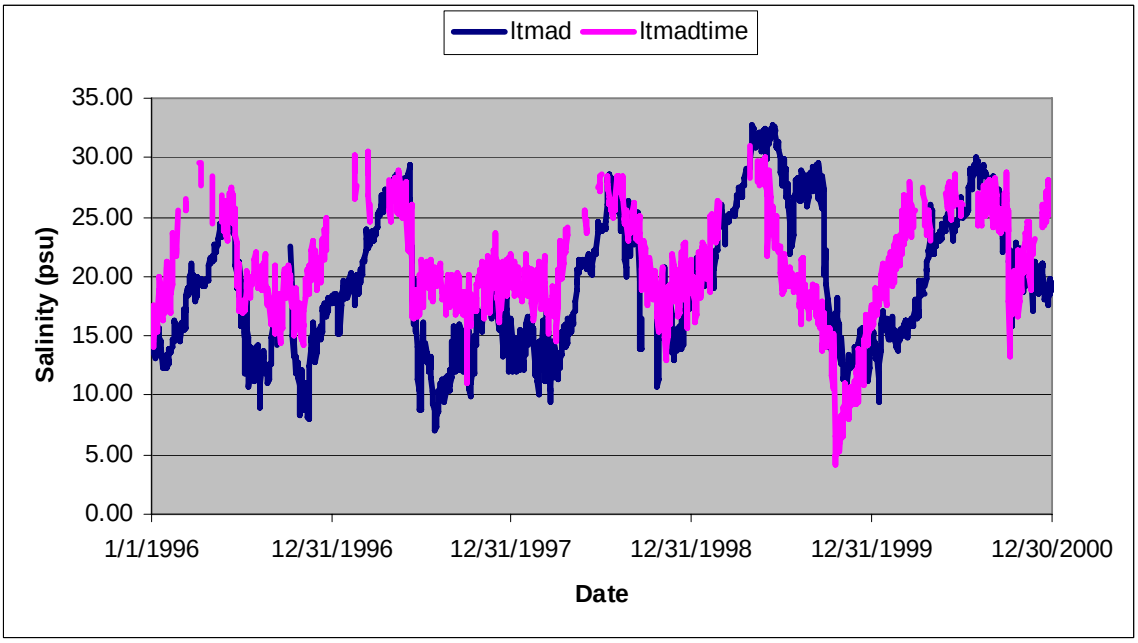


Figure 11. Comparison of observed data and salinity simulations for Terrapin Bay produced by MLR salinity models (Marshall et al 2004) using stage values from the TIME 157 model run.

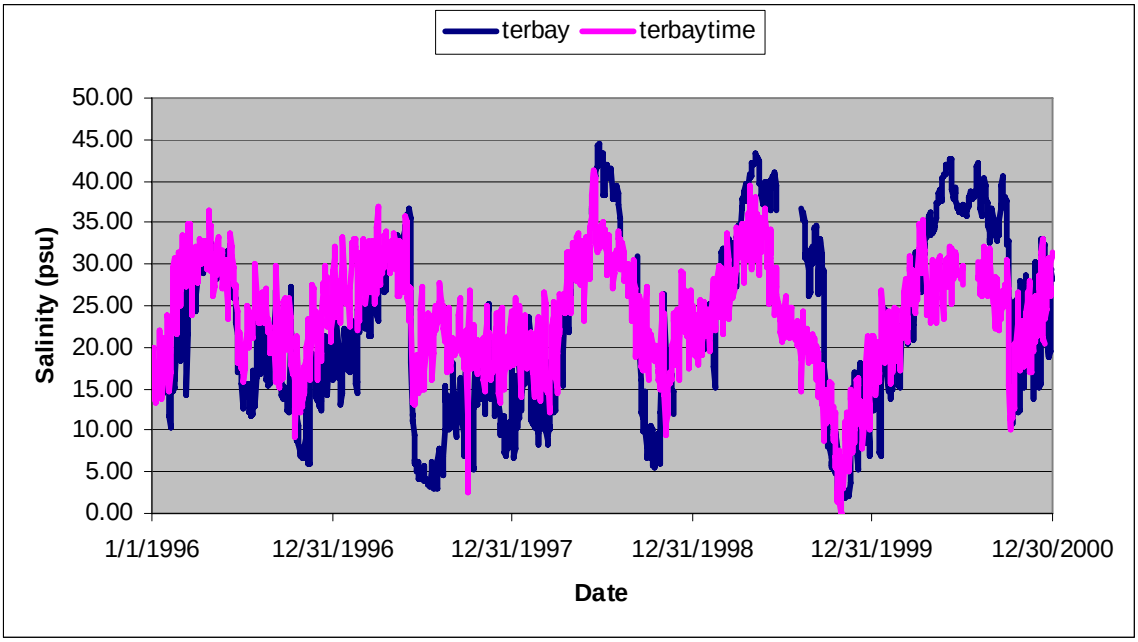
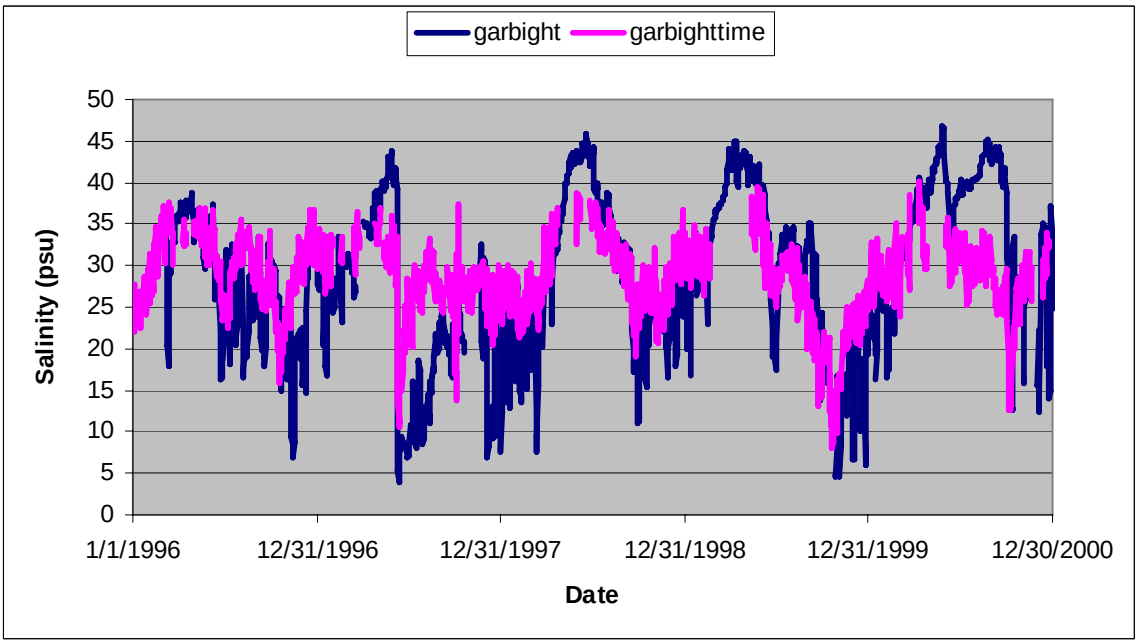


Figure 12. Comparison of observed data and salinity simulations for Garfield Bight produced by MLR salinity models (Marshall 2005) using stage values from the TIME 157 model run.



V. Discussion, Findings, and Recommendations

This task was the first task initiated for this part of the on-going CESI project, and the last to be finished. Mostly this was caused by the evolution of the SICS domain into the TIME domain, and the difficulty of the USGS in making the code revisions that were needed. However, it appears that the TIME model is now running dependably. Even so, stage output data are missing when the simulated water levels dropped below the ground surface elevation for the grid cell corresponding to a particular monitoring station needed for MLR salinity models. The ground water stage output is produced by the groundwater model, SEAWAT while surface water stage is produced by the SWIFT2D model. Output is only available from the SWIFT2D model at the surface cells.. An integrated time series of stage is needed to be able to fully couple TIME and MLR salinity models.

When the TIME 157 stage data are compared to the observed stage(Figures 3 through 8) it can be seen that the data are similar for both models during the wet season, but the dry season values show various levels of divergence. The greatest divergence of TIME 157 and observed dry season values is seen at NP62; the least divergence is seen at P33.

When both of these input data sets are used with wind and Key West water level data in the MLR salinity models, the comparison plots (Figures 9 through 12) show that salinity produced using the TIME 157 data and salinity produced using the observed data are also similar,. The exception may be Joe Bay, but there are many missing values for the dry season. Because of this drawing a conclusion is not feasible. EVER6 in the Joe Bay model input had many dry season missing values. All salinity simulations were affected by missing dry season data from TIME 157. Given the differences seen in the stage comparison plots, the similarity of the salinity simulations is somewhat surprising.

It is concluded that the TIME output may serve as input to MLR salinity models and the resulting salinity simulations are similar to salinity produced using the SFWMM output for similar modeling scenarios.

VI. References

Cline, J.C. and E. D. Swain. 2002. Coupling hydrologic and ecological modeling: SICS and ATLSS. Second Federal Interagency Hydrologic Modeling Conference, July 28-August 1, 2002, 12 pp.

Langevin, C.D., E. D. Swain, and M. A. Wolfert. 2002. Numerical simulation of integrated surface-water/groundwater flow and solute transport in the southern Everglades, Florida, Second Federal Interagency Hydrologic Modeling Conference, July 28-August 1, 2002, 12 p.

Langevin, C.D. E.D. Swain, J.D. Wang, M.A. Wolfert, R.W. Schaffranek, and A.L. Riscassi. 2004a. Development of coastal flow and transport models in support of everglades restoration. Fact Sheet 2004-3130. U.S. Department of the Interior, U.S. Geological Survey. October 2004.

Langevin, C.D. E.D. Swain, and M.A. Wolfert. 2004a. Simulation of integrated surface-water / groundwater flow and salinity for a coastal wetland and adjacent estuary. Open-File Report 2004-1097. U.S. Department of the Interior, U.S. Geological Survey. October 2004.

Marshall III, F. E.; D. Smith; and D. Nickerson. 2004. Using statistical models to simulate salinity variation and other physical parameters in north Florida Bay. Cetacean Logic Foundation, Inc., New Smyrna Beach, Florida.

Marshall III, F. E. 2005. RECOVER Southern Estuaries performance measures: identification of hydrology-salinity relationships for coastal estuaries and analysis of interim CERP update scenarios. Environmental Consulting & Technology, Inc., New Smyrna Beach, Florida.

Swain, E.D., Wolfert, M.A., Bales, J.D., and Goodwin, C.R., 2004, Two-dimensional hydrodynamic simulation of surface-water flow and transport to Florida Bay through the Southern Inland and Coastal Systems (SICS): U.S. Geological Survey Water-Resources Investigations Report 03-4287.

FINAL TASK REPORT

HISTORICAL RECONSTRUCTIONS OF SALINITY IN NORTHEAST FLORIDA BAY USING MULTIVARIATE LINEAR REGRESSION MODELS

Task 11

Cooperative Agreement Number CA H5284-05-0006
Between
The United States Department of the Interior National Park Service
Everglades National Park
And
Cetacean Logic Foundation, Inc.

Frank E. Marshall, III
Principal Investigator

DeWitt Smith
Project Coordinator

September 7, 2007

CESI Project No. CA H5284-05-0006 for Everglades National Park

FINAL REPORT
TASK NO. 11

HISTORICAL RECONSTRUCTIONS OF SALINITY IN NORTHEAST FLORIDA BAY USING MULTIVARIATE LINEAR REGRESSION MODELS

September 7, 2007

I. Introduction

A. General

This report describes the activities that were completed for a task in the current Critical Ecosystem Studies Initiative (CESI) project that was originally initiated by Cetacean Logic Foundation, Inc. (CLF) for Everglades National Park (ENP) in 2002. The goal of this on-going CLF research is the characterization and simulation of the salinity regime in the estuaries of Florida Bay and the southwest coast of the Gulf of Mexico (Gulf) within the Park. It is intended that this research yield information about salinity variation over time and the link between the downstream salinity in the estuaries of ENP and the upstream freshwater hydrology of the Everglades, which has been altered in recent times by water management activities.

The subject of this report is the development of historical time series reconstructions of the salinity regime for ENP Marine Monitoring Network (MMN) stations in northeast Florida Bay using multivariate linear regression (MLR) salinity models. MLR salinity models were previously developed for about half of the 33 stations in the MMN (Marshall, 2005a; 2005b), and the development of models for the remaining MMN stations was recently completed as part of another task in this CESI project. These MLR models will be used with other statistical methods to extend the time series of salinity back in time until around 1970 prior to the collection of continuous salinity measurements in ENP.

B. Task Objectives and Evaluation Methods

The objective of this task is presented below as it appears in the modified contract for this CESI project approved by ENP:

> "The SFWMD is using the Taylor River MLR salinity model for their Florida Bay MFL work. As part of that work the PI developed a historical salinity reconstruction for the period 1970 through 2003. Salinity measurements have been collected digitally at the Taylor River site by ENP since 1988. The Taylor River MLR salinity model was used to extend the salinity time series back in time to 1970, which is when the C-111 Canal system came

on line. To do so required using other MLR models to also extend certain input parameters.

Using the same methodology used by the PI at Taylor River for the SFWMD MFL activities, a historical salinity reconstruction will be developed at MMN stations in northeast Florida Bay (Joe Bay, Little Madeira Bay, Terrapin Bay, Garfield Bight, Long Sound, Highway Creek, Little Blackwater Sound, Duck Key, Butternut Key, Whipray Basin, and Bob Allen Key). A spectral analysis will then be performed using the historical reconstruction to determine if the dominate frequencies in the observed data appear also in the modeled data.

The deliverables are daily historical salinity reconstructions at Joe Bay, Little Madeira Bay, Terrapin Bay, Garfield Bight, Long Sound, Highway Creek, Little Blackwater Sound, Duck Key, Butternut Key, Whipray Basin, and Bob Allen Key, and a task report that describes the activities involved."

II. Data Used for Historical Reconstructions of Salinity

The foundation data used for 1970-2000 salinity reconstructions are data collected at Joe Bay (JB), Little Madeira Bay (LM), Terrapin Bay (TB), Garfield Bight (GB), Long Sound (LS), Highway Creek (HC), Little Blackwater Sound (LB), Duck Key (DK), Butternut Key (BN), Whipray Basin (WB), and Bob Allen Key (BA) by ENP (see Figure 1). Continuous salinity data are available at these MMN stations starting in 1988 at several locations in northeast Florida Bay, while the collection of salinity data at other stations may not have begun until 1997 (for example Bob Allen Key). The observed salinity data contain missing values, in addition to the missing values from January 1, 1970 through the beginning of salinity data collection, 1988 at the earliest. Gaps in the observed salinity data were filled using the MLR salinity models. The MLR salinity models are a linear combination of parameters that are correlated with salinity, including freshwater stage elevations in the Everglades, sea surface elevation measured at Key West, and wind vectors. Figure 1 shows the ENP MMN salinity monitoring stations and Figure 2 shows the stage monitoring stations used for this study.

For the MLR model input, the stage data are Everglades water levels from selected stations in the ENP Physical Monitoring Network. These observed stage data are available on the South Florida Water Management District DBHYDRO website (http://my.sfwmd.gov/portal/page?_pageid=2235,4688582&_dad=portal&_schema=PORTAL). The principal stage station used for MLR modeling is Craighead Pond (CP), with a data record that commences in 1978. The longest period of record for the continuous water level (stage) monitoring stations used in this study (P33) began in the 1950's (see Table 1).

Sea surface elevation data measured at Key West were obtained from the NOAA Tides Online website (http://tidesonline.nos.noaa.gov/). Wind data were obtained directly from the National Weather Service (Southeast Regional Climate Center) and provided to CLF by ENP for Key West and Miami stations. Wind data from Key West and Miami were used as these locations had the longest continuous records (until at least 1970) for wind and were considered to be representative of the regional wind patterns. The wind speed and direction data were processed by CLF into vector quantities. Additional information on the data used for this task can be found in Marshall 2003a and 2003b).

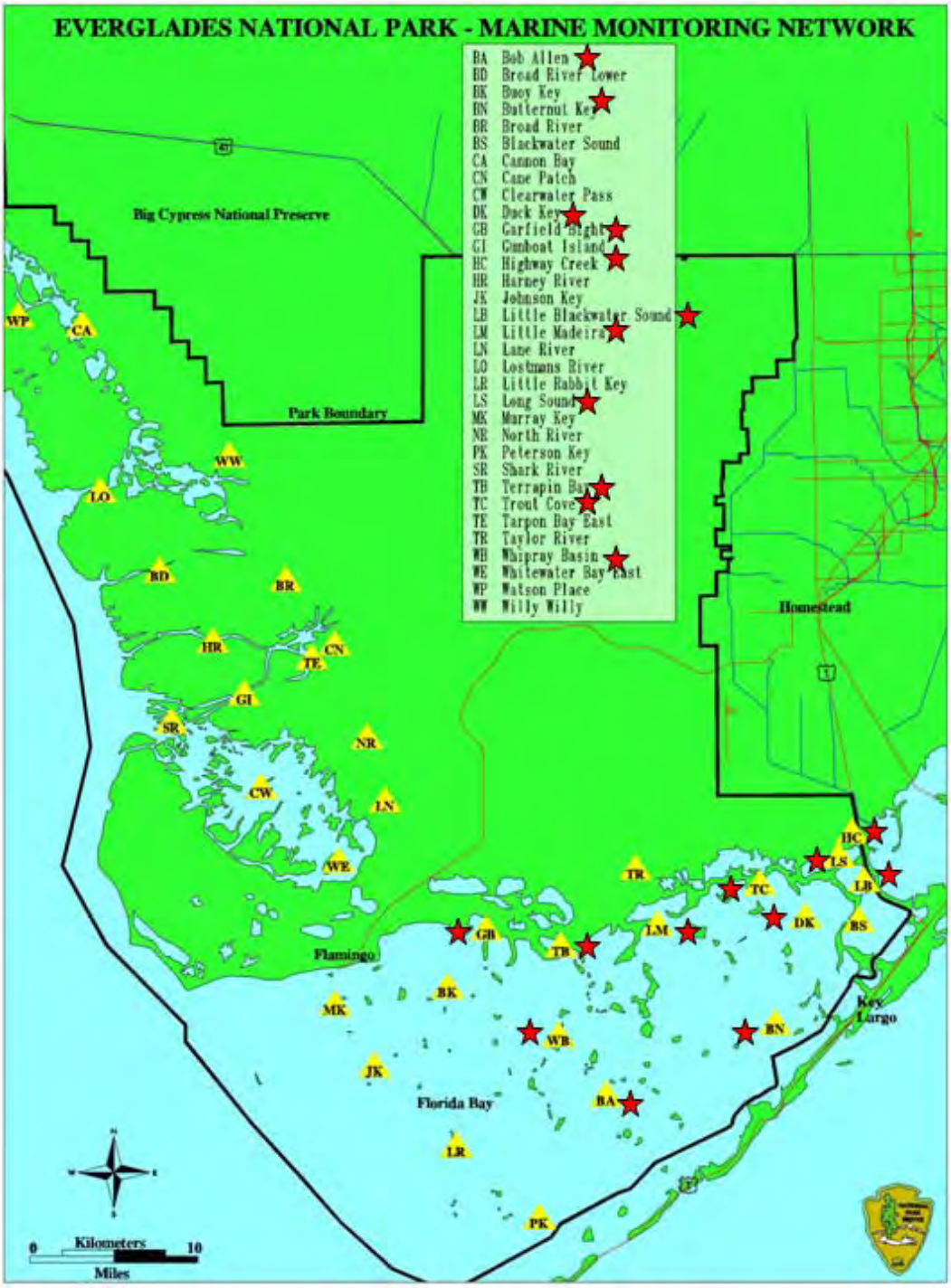


Figure 1. Location map showing the salinity monitoring stations in the ENP Marine Monitoring Network used for the historical restorations of salinity in Florida Bay, 1970-2000. Red stars indicate salinity stations for historical reconstructions.

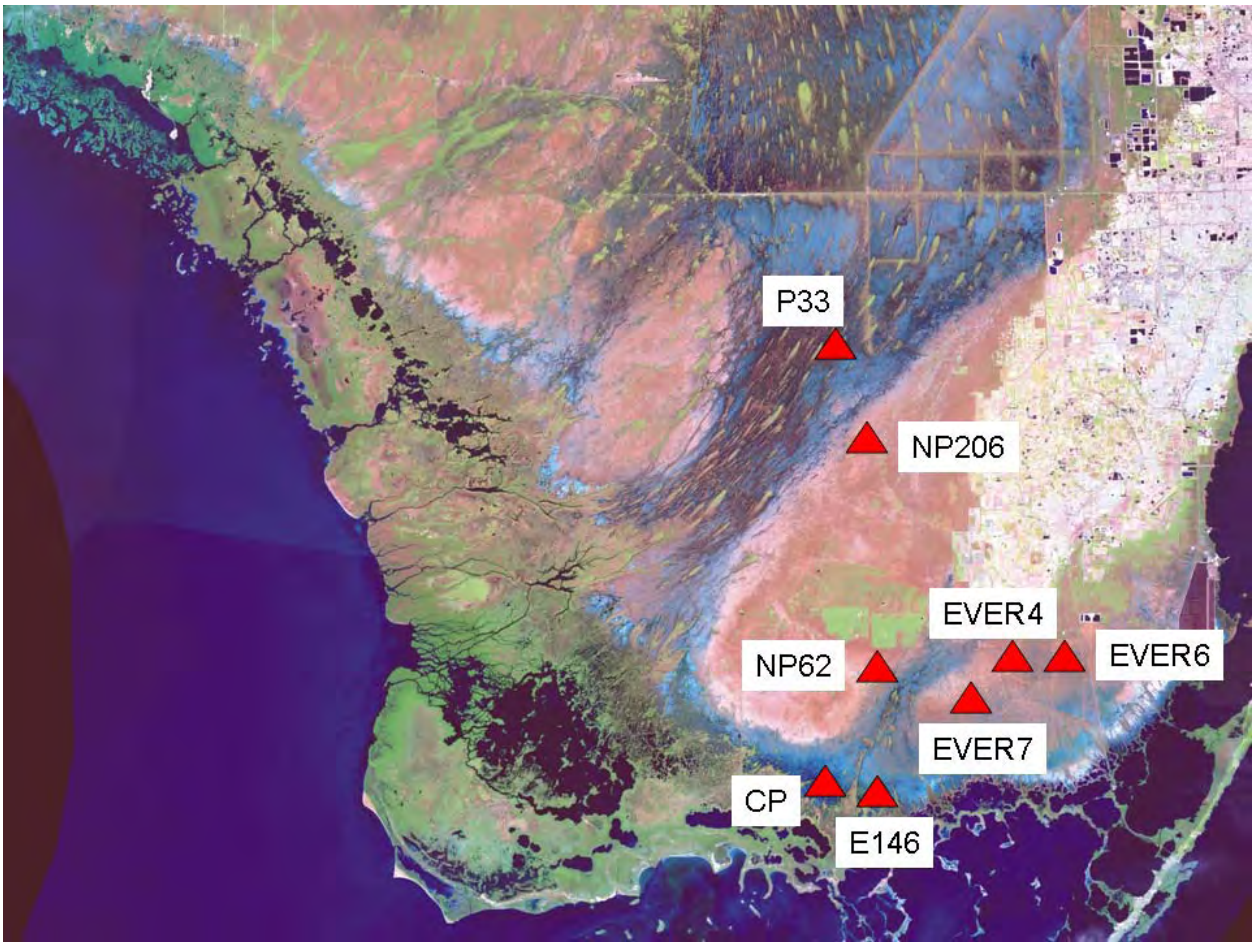


Figure 2. Location map showing the stage monitoring stations used for the historical restorations of salinity in Florida Bay, 1970-2000.

Table 1. Information on salinity and stage data used in the MLR models for historical reconstructions.

ID	Parameter	Period of Record Start	Station Name
JB	Salinity	07/14/88	Joe Bay
LM	Salinity	08/25/88	Little Madeira Bay
TB	Salinity	09/12/91	Terrapin Bay
GB	Salinity	07/03/91	Garfield Bight
LS	Salinity	07/14/88	Long Sound
HC	Salinity	07/14/88	Highway Creek
LB	Salinity	09/11/91	Little Blackwater Sound
WB	Salinity	04/06/89	Whipray Basin
DK	Salinity	07/14/88	Duck Key
BN	Salinity	02/08/90	Butternut Key
BA	Salinity	09/09/97	Bob Allen Key
CP	Stage	10/01/78	Craighead Pond
P33	Stage	10/01/52	
NP62	Stage	01/04/64	
E146	Stage	03/24/94	
EVER4	Stage	09/20/85	
EVER6	Stage	12/24/91	
EVER7	Stage	12/24/91	
NP206	Stage	10/01/74	

III. Historical Reconstruction and Spectral Density Methods

The objective of this study is to build a continuous daily time series of salinity for the period 1970 - 2000 (historical reconstruction) at the subject MMN stations, and to interpret the resulting data through spectral analysis. This requires filling gaps in the observed data, including the period before data were collected, if possible. Experience that was acquired through the implementation of Task 8 of this CESI project the subject of which was filling data gaps was most useful for developing salinity reconstructions. A set of guidelines was used to develop the historical reconstructions of salinity, as follows:

- 1970 was chosen as the point of beginning for the reconstructions because of data limitations and because this is about the time that the current configuration of Tamiami Trail and other water management facilities became operational which directly affected the delivery of water to the Everglades and Florida Bay as seen in the historic hydrologic data.
- For this study, gaps are defined as periods (days) during which monitoring was occurring and data are not available **and** the period from January 1, 1970 until monitoring was begun.
- The reconstructed salinity time series is constructed beginning with the observed data.
- For filling salinity data gaps, observed stage, Key West water level, and wind were used when available and as selected by the stepwise regression process.
- For the period of this study, there are no missing data in the Key West water level and wind data so it was not necessary to fill gaps.
- When there were gaps in the observed stage data they were filled using univariate linear regression models developed from observed CP data.
- When there were gaps in CP observed data they were filled using univariate linear regression models developed from NP62, including internal gaps after October 1, 1978 and the beginning-of-period gap from January 1, 1970 through October 1, 1978.
- When there were gaps in the observed NP62 data, they were filled using univariate linear regression models developed from observed P33 data.

A spectral analysis using each historical reconstruction was performed using the SAS© PROC SPECTRA algorithm. Program output including periodograms and plots of spectral density functions were examined.

For all salinity reconstruction, when observed data are not available, the following MLR salinity models were utilized (Marshall et al, 2003; Marshall 2003):

JOE BAY = 37.1 - 3.1CP - 3.5 EVER6 [lag6] - 10.5 E146 [lag6] - 0.2 UWNDKW - 0.09 UWNDKW [lag2] - 0.10 VWNDKW - 0.16 VWNDMIA [lag1]

LITTLE MADEIRA BAY = 106.1 - 0.3 CP [lag2] - 12.5 P33 [lag2] - 1.7 (P33-NP206) - 0.25 UWNDKW + 0.13 UWNDMIA - 0.19 VWNDMIA [lag1]

+ 0.95 KWWATLEV [lag2]

TERRAPIN BAY = 106.9 - 6.3 CP [lag1] - 11.1 P33 [lag2] - 0.45 UWNDKW
- 0.23 UWNDKW [lag1] - 0.2 UWNDKW [lag2] - 0.14 VWNDKW [lag2]
+ 0.46 UWNDMIA + 1.9 KWWATLEV [lag2]

GARFIELD BIGHT = 56.1 - 9.2 CP [lag1] - 4.6 NP62 [lag1]
- 0.46 UWNDKW [lag1] - 0.48 UWNDKW [lag4] + 0.35 UWNDMIA [lag1]
+ 0.64 UWNDMIA [lag4]

LONG SOUND = 42.2 - 9.5 CP [lag4] - 5.2 EVER7 [lag2] - 1.7 EVER6 [lag2]
- 0.04 VWNDMIA [lag1]

HIGHWAY CREEK = 49.9 - 5.3 CP - 16.3 EVER6 [lag4] + 0.2 UWNDMIA [lag3]
+ 0.73 KWWATLEV - 6.3 (EVER7 - EVER4) [lag2]

LITTLE BLACKWATER SOUND = 42.5 -7.65 CP [lag6] - 6.3 EVER7 [lag5]
+ 0.12 VWNDKW

WHIPRAY BASIN = 21.1 + 0.24 LM [lag3] + 0.2 TB
+ 0.15 TB [lag3] - 0.04 VWNDKW [lag2] - 0.5 KWWATLEV [lag2]

DUCK KEY = 10.2 + 0.3 LM [lag1] + 0.4 LM [lag3] + 0.10 UWNDKW [lag1]
+ 0.13 VWNDKW [lag2] + 0.5 KWWATLEV

BUTTERNUT KEY = 15.4 + 0.14 LM [lag1] + 0.44 LM [lag3]
+ 0.03 TB [lag3] - 0.08 UWNDKW - 0.10 UWNDKW [lag2] + 0.4 KWWATLEV

BOB ALLEN KEY = 19.4 + 0.3 LM + 0.25 LM [lag3] + 0.08 TB [lag3]
- 0.04 UWNDKW - 0.07 UWINDKW [lag2] - 0.06 VWNDKW [lag2]

For filling gaps in the stage data used for the above models, the following univariate linear regression models were used:

CP = - 0.20 + 0.5 NP62

NP62 = - 4.9 + 1.2 P33

P33 = 5.3 + 0.80 CP

E146 = 0.12 + 0.84 CP

EVER4 = 1.19 + 0.70 CP

EVER6 = 1.05 + 0.74 CP

$$EVER7 = 1.35 + 0.66\ CP$$

$$NP206 = 3.52 + 1.57\ CP$$

Error statistics for the MLR salinity models used for the historical reconstructions and for the univariate linear regression stage models used to fill the data gaps are presented in Tables 2 and 3, respectively. Verification plots for the univariate linear regression stage models are presented in Figures 3 through 9.

Table 2. Error statistics for MLR salinity models used for historical reconstructions.

station	mean error, psu	mean sq error (mse), psu^2	root mse (rmse), psu	mean abs error, psu	max abs error, psu	adj R-sq	Nash-Sutcliffe Efficiency
Joe Bay	-0.14	25.8	5.1	3.7	20.6	0.75	0.76
Little Madeira Bay	-0.66	40.1	6.4	5.1	22.6	0.65	-0.96
Terrapin Bay	-0.99	32.6	5.7	5.4	5.4	0.75	0.67
Garfield Bight	-0.36	37.9	6.15	4.75	21.1	0.68	0.89
Long Sound	0.31	15	3.9	2.7	18.9	0.8	0.81
Highway Creek	-0.95	18.2	4.3	3.7	17.7	0.81	0.76
Little Blackwater Sound	-0.14	14	3.7	2.9	15.7	0.75	0.76
Whipray Basin	0.11	7.2	2.7	2.2	10.1	0.8	0.77
Duck Key	-0.18	9.7	3.1	2.27	14.4	0.71	0.71
Butternut Key	0.1	10.7	3.3	2.7	11.3	0.65	0.66
Bob Allen Key	0.3	7.2	2.7	2.1	9.2	0.79	0.81

Table 3. Error statistics for univariate linear regression stage models used for historical reconstructions.

	N	mean error, psu	mean sq error (mse), psu^2	root mse (rmse), psu	mean abs error, psu	max abs error, psu	adj R-sq	Nash-Sutcliffe Efficiency
np62	10495	0.00	0.33	0.58	0.43	3.08	0.67	0.67
p33	7883	0.00	0.15	0.39	0.31	1.26	0.56	0.56
np206	7611	0.00	0.59	0.77	0.56	3.22	0.55	0.55
ever4	4877	0.00	0.08	0.28	0.20	1.55	0.66	0.66
ever6	3446	0.00	0.07	0.27	0.17	2.47	0.67	0.67
ever7	3380	0.00	0.03	0.18	0.16	0.97	0.79	0.79
e146	2879	0.00	0.01	0.11	0.08	0.64	0.95	0.95

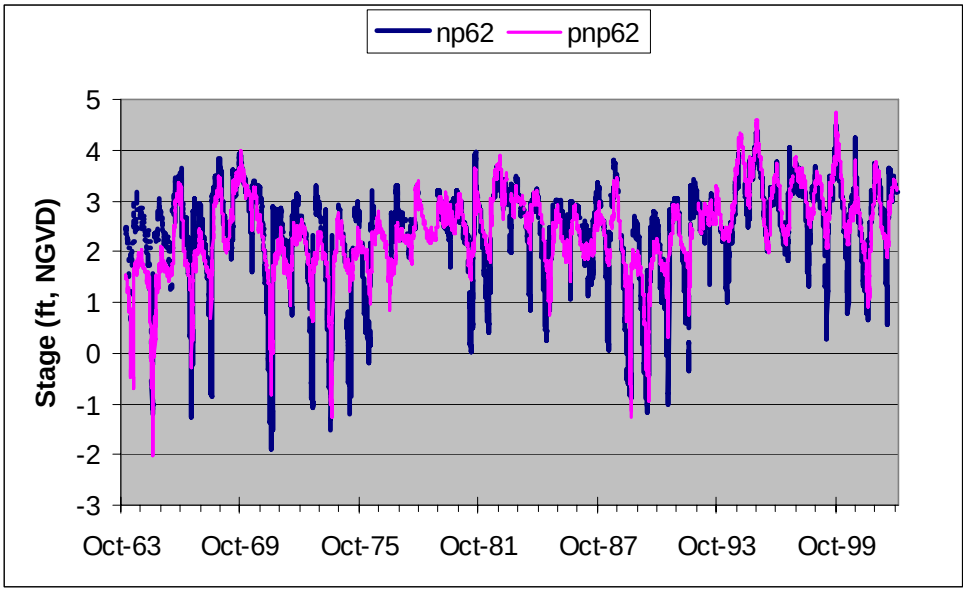


Figure 3. Verification plot for NP62 univariate linear regression model.

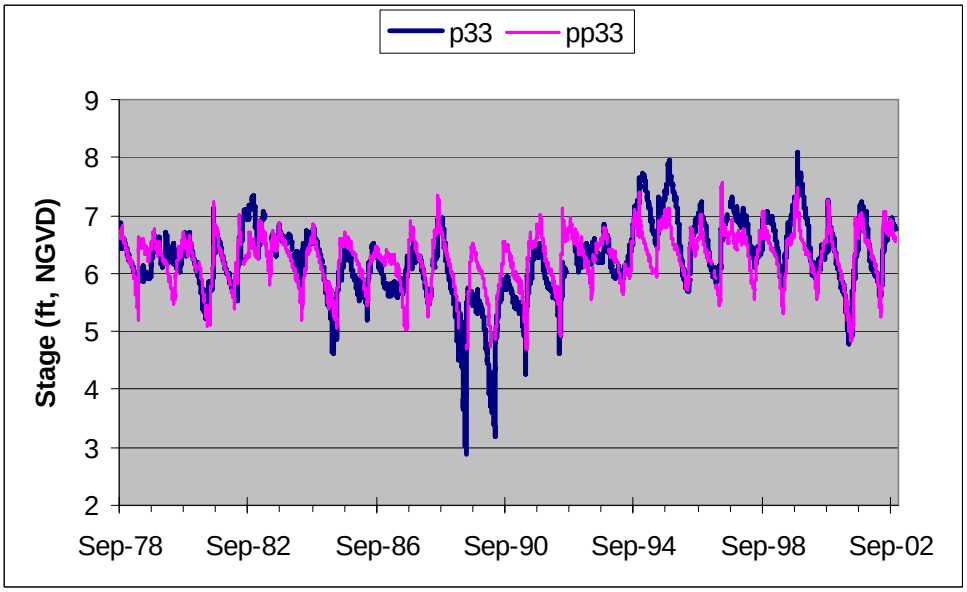


Figure 4. Verification plot for P33 univariate linear regression model.

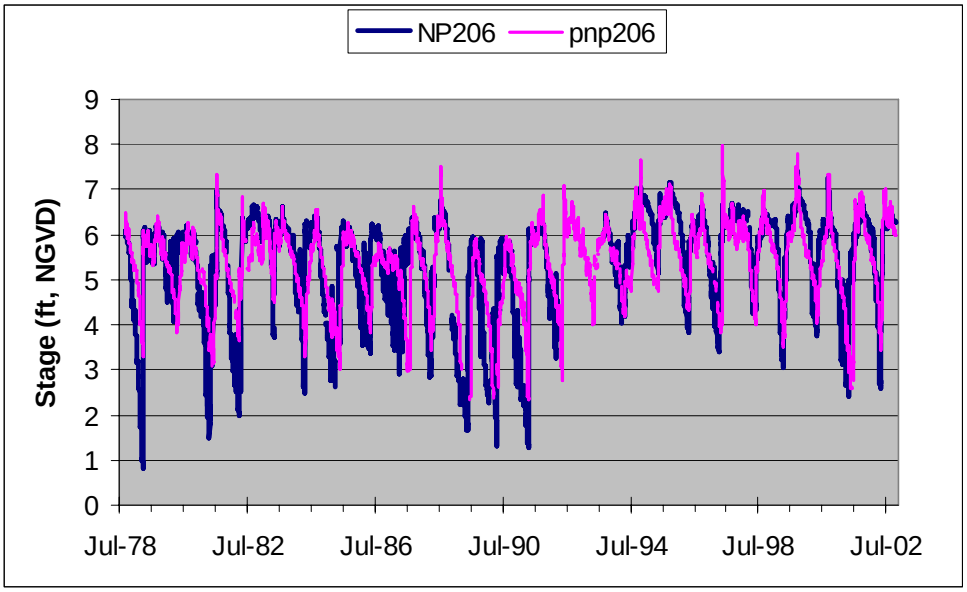


Figure 5. Verification plot for NP206 univariate linear regression model.

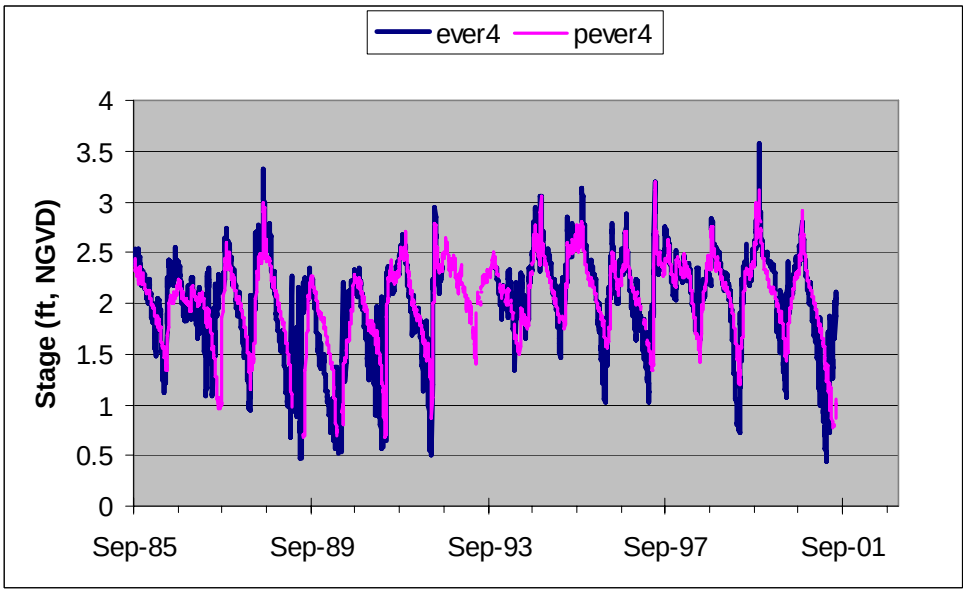


Figure 6. Verification plot for EVER4 univariate linear regression model.

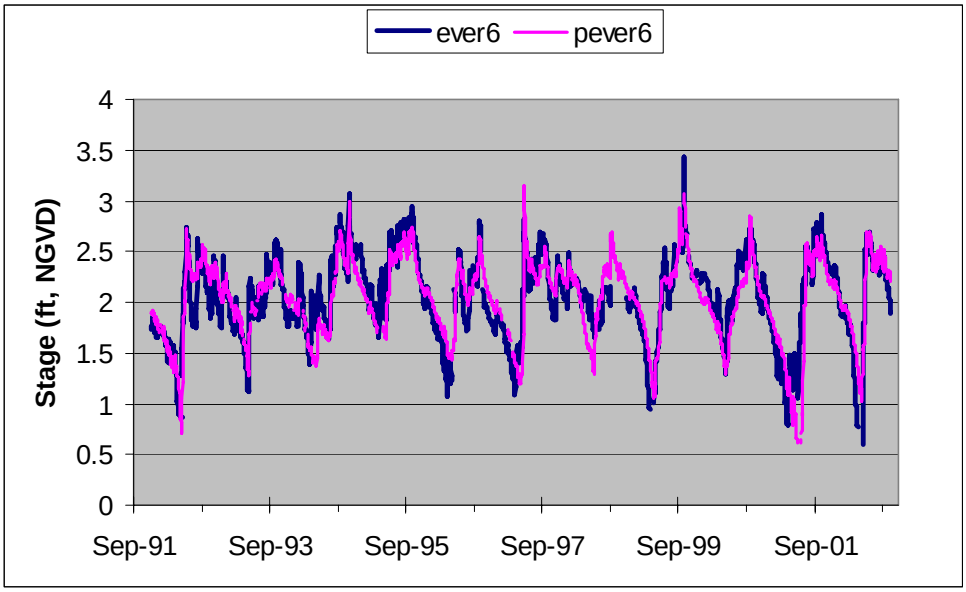


Figure 7. Verification plot for EVER6 univariate linear regression model.

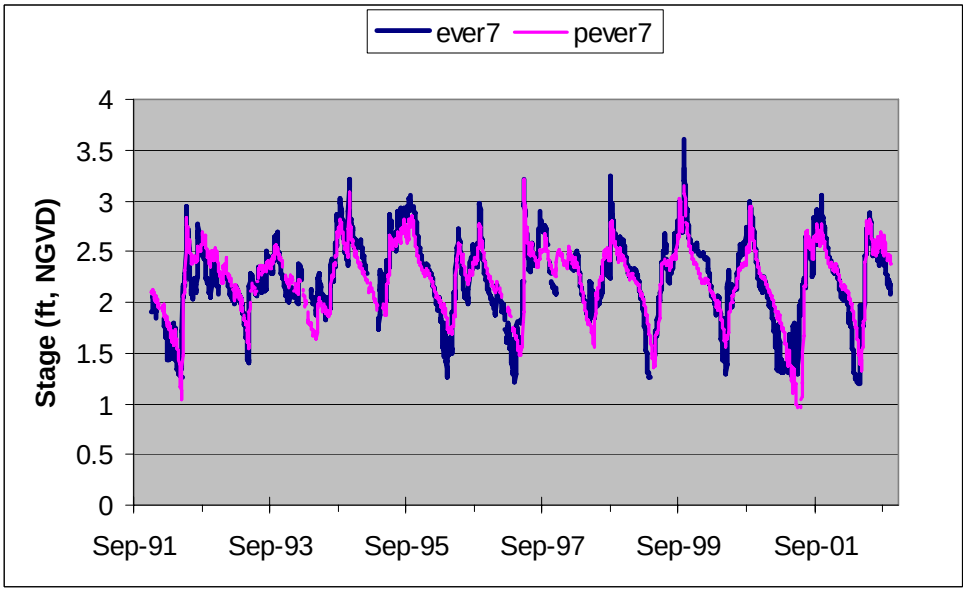


Figure 8. Verification plot for EVER7 univariate linear regression model.

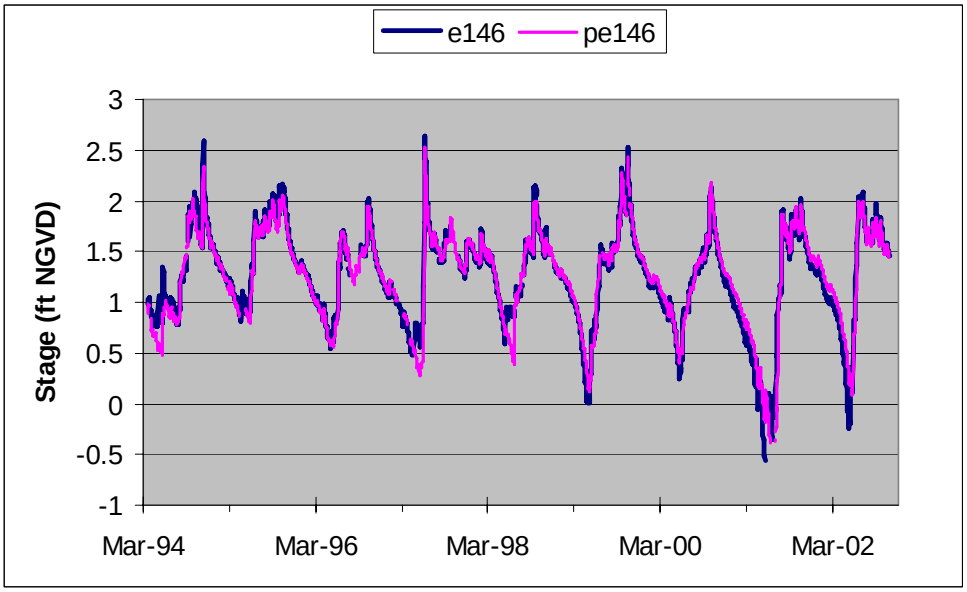


Figure 9. Verification plot for E146 univariate linear regression model.

IV. Historical Reconstructions

For the following MMN stations in northeast Florida Bay:

- Joe Bay,
- Little Madeira Bay,
- Terrapin Bay,
- Garfield Bight,
- Long Sound,
- Highway Creek,
- Little Blackwater Sound,
- Duck Key,
- Butternut Key,
- Whipray Basin, and
- Bob Allen Key

Historical reconstructions were constructed using the guidelines and models presented above. Plots of the daily time series are presented as Figures 10 through 20.

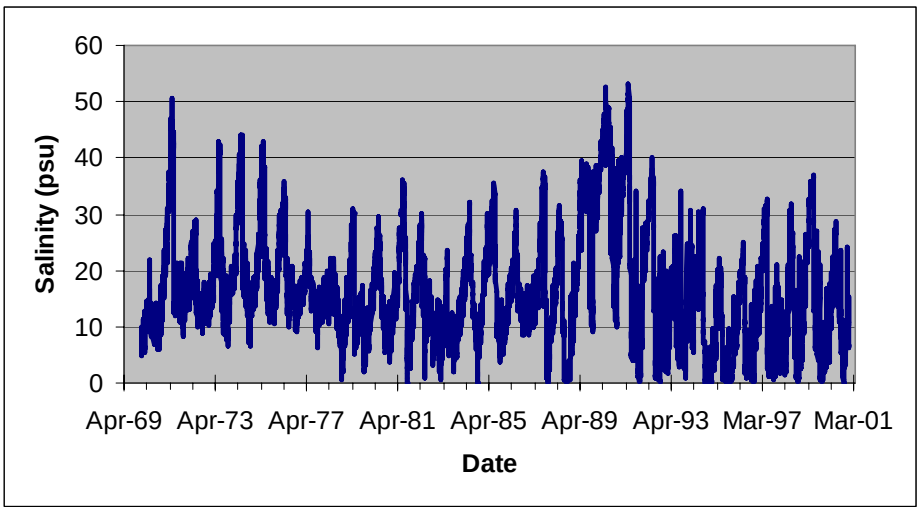


Figure 10. Historical reconstruction for Joe Bay. Observed data begin on July 14, 1988.

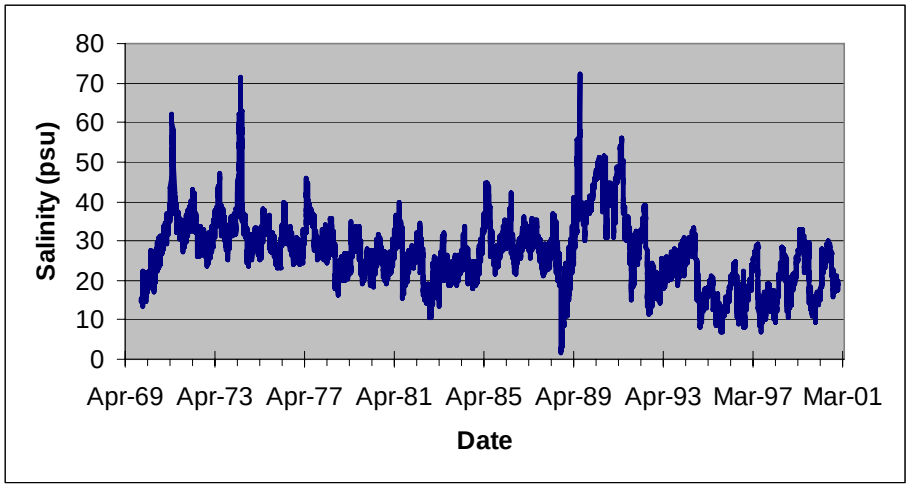


Figure 11. Historical reconstruction for Little Madeira Bay. Observed data begin on August 25, 1988.

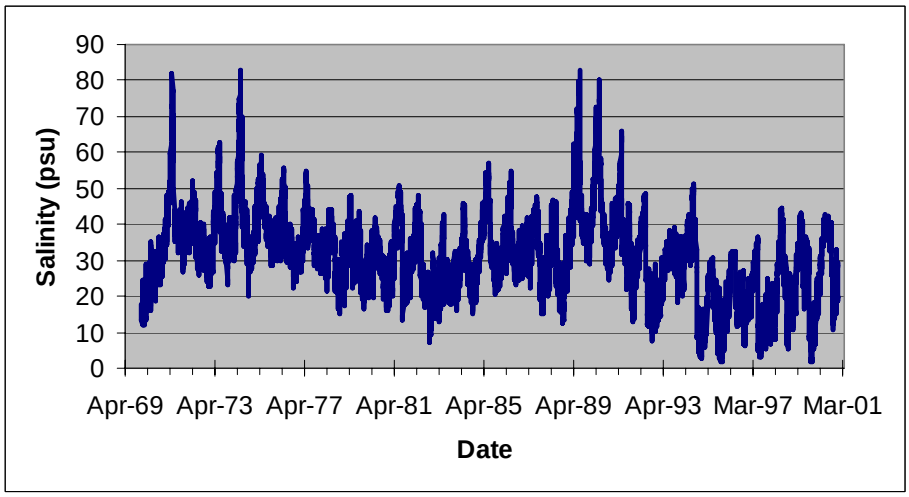


Figure 12. Historical reconstruction for Terrapin Bay. Observed data begin on September 12, 1991.

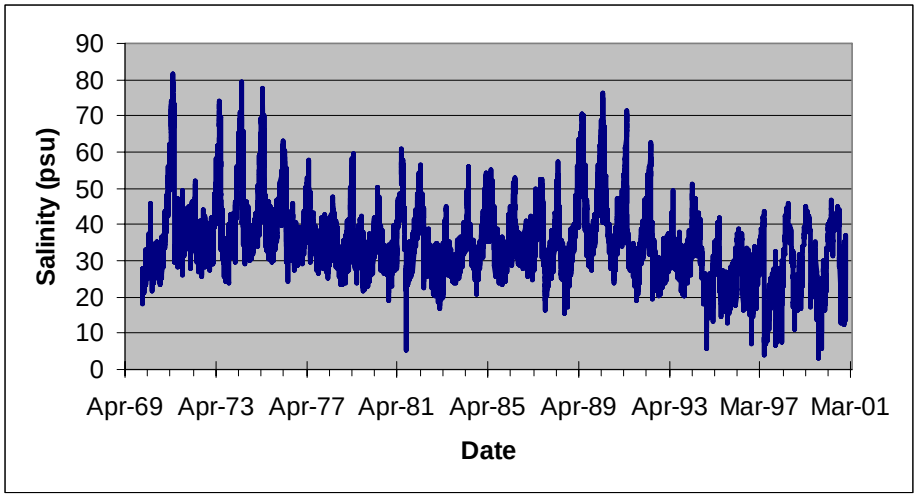


Figure 13. Historical reconstruction for Garfield Bight. Observed data begin on July 3, 1991.

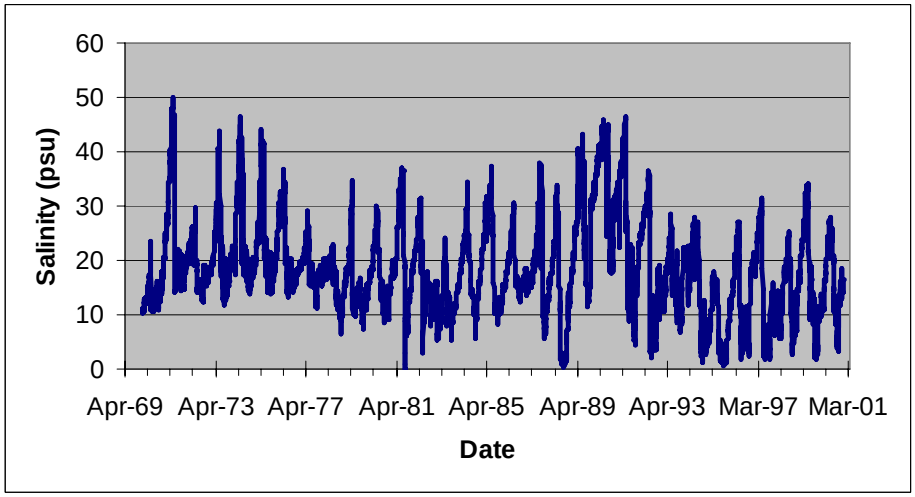


Figure 14. Historical reconstruction for Long Sound. Observed data begin on July 14, 1988.

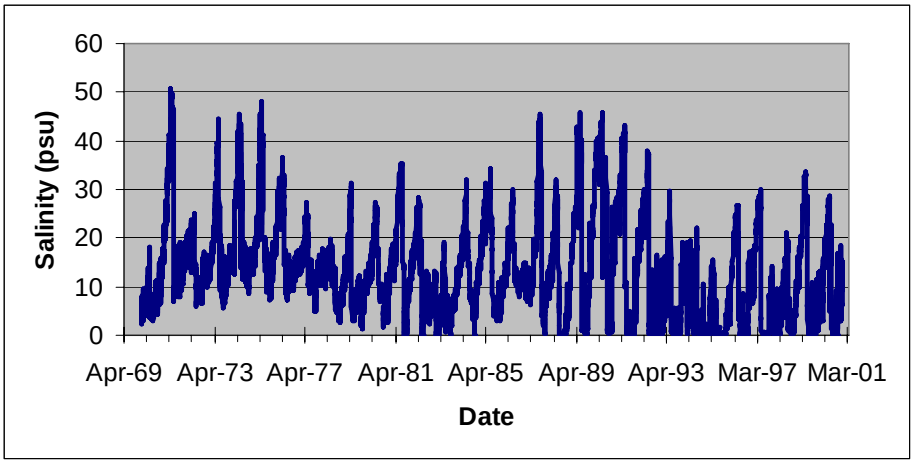


Figure 15. Historical reconstruction for Highway Creek. Observed data begin on July 14, 1988.

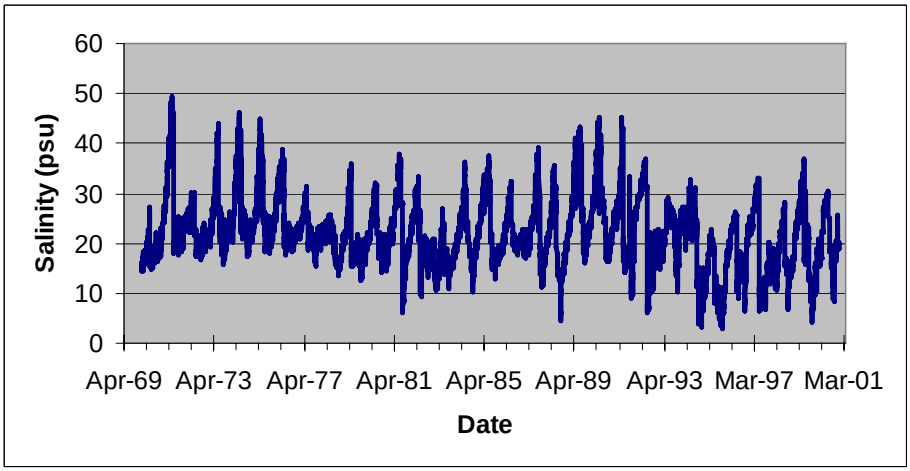


Figure 16. Historical reconstruction for Little Blackwater Sound. Observed data begin on September 11, 1991.

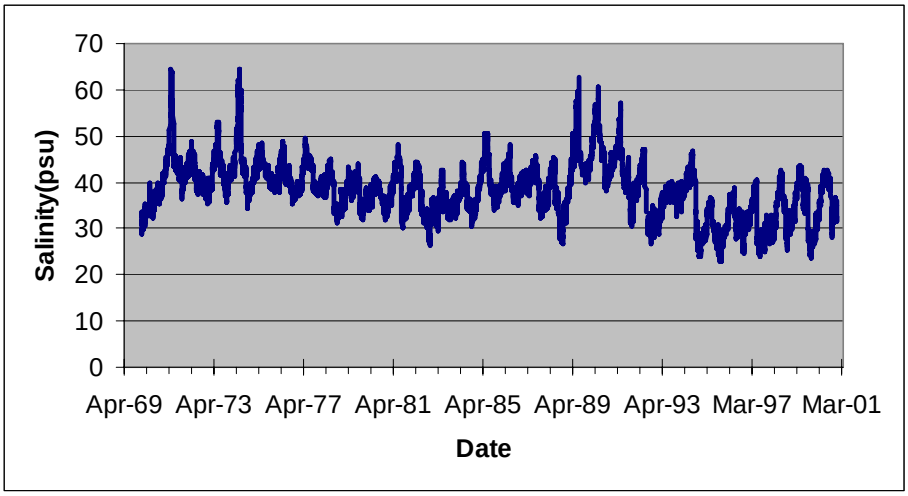


Figure 17. Historical reconstruction for Whipray Basin. Observed data begin on April 6, 1989.

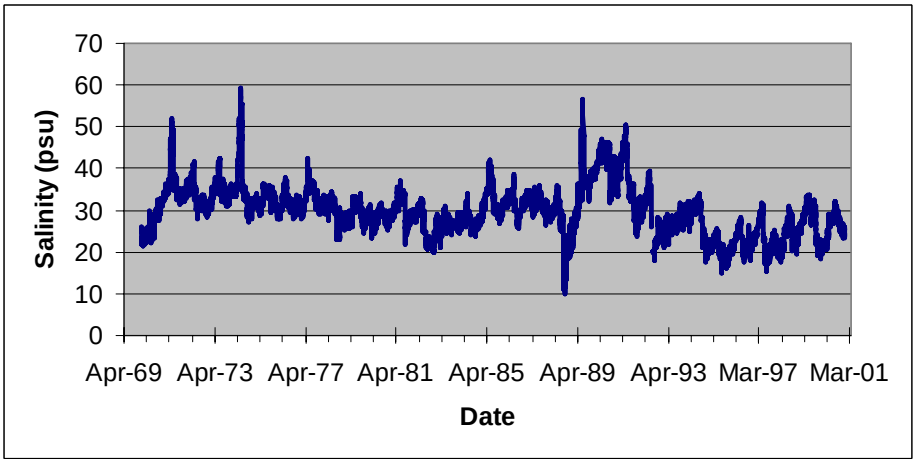


Figure 18. Historical reconstruction for Duck Key. Observed data begin on July 14, 1988.

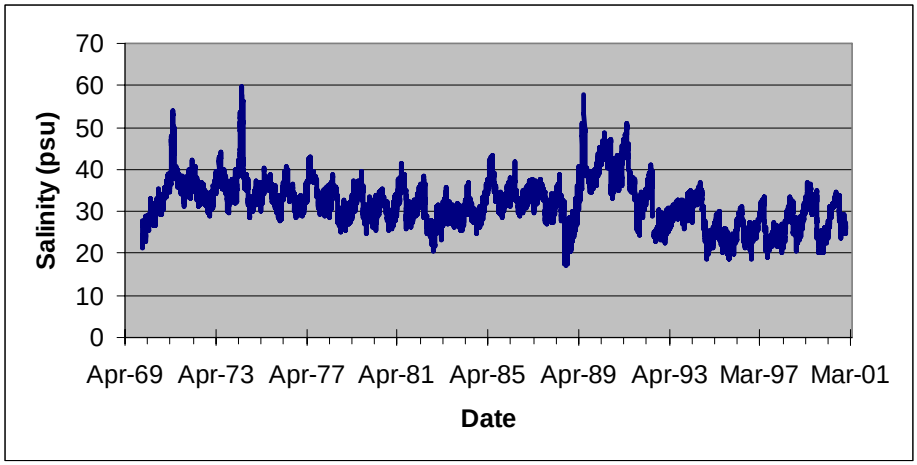


Figure 19. Historical reconstruction for Butternut Key. Observed data begin on February 8, 1990.

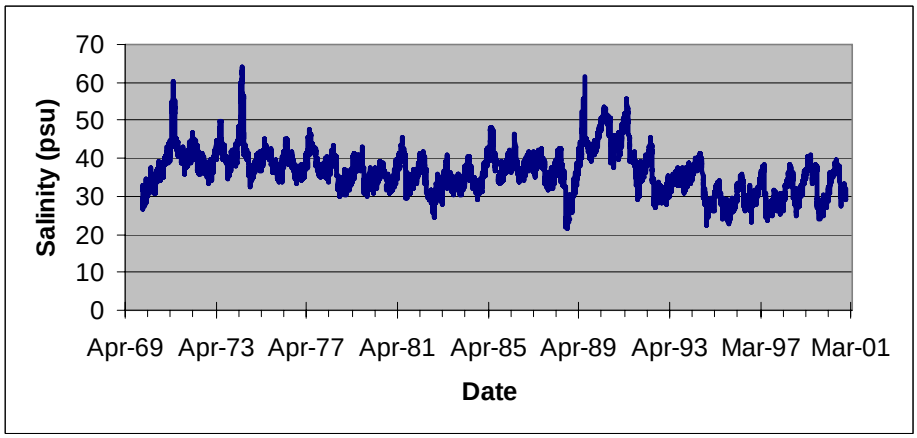


Figure 20. Historical reconstruction for Bob Allen Key. Observed data begin on September 9, 1997.

V. Spectral Analysis

A spectral analysis of each historical reconstruction was performed and program output including periodograms and plots of spectral density functions were examined. Plots of the spectral density function values for the period in days is presented as Figures 21. For ease of plotting, the mean has been subtracted from the salinity value before performing the Fourier decomposition. The spectral density functions of Joe Bay, Terrapin Bay, Garfield Bight, Long Sound, Highway Creek, and Little Blackwater Sound (near shore stations) are similar in the occurrence of maximum values. Likewise, the timing of maximum values is similar for Little Madeira Bay, Whipray Basin, Duck Key, Butternut Key, and Bob Allen Key (observed POR based on limited values at Bob Allen Key), and different in some respects than the other set of stations. The Little Madeira Bay spectral function behavior may be reflecting the location of the monitoring station, which is intermediate in location between near shore stations and open-bay stations.

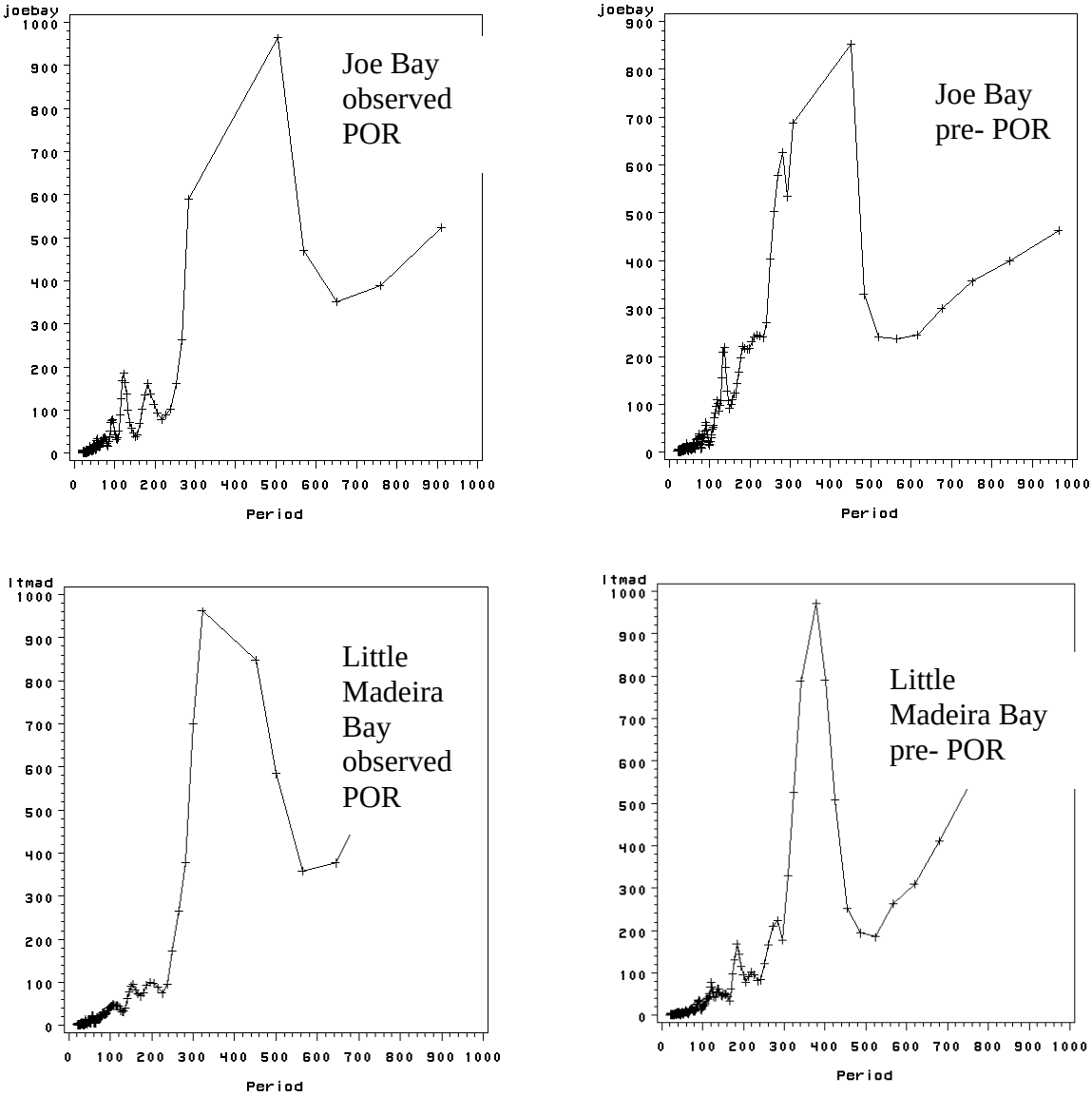


Figure 21. Comparison of spectral density function values for observed data (left column) and for hindcast reconstruction simulation data (right column).

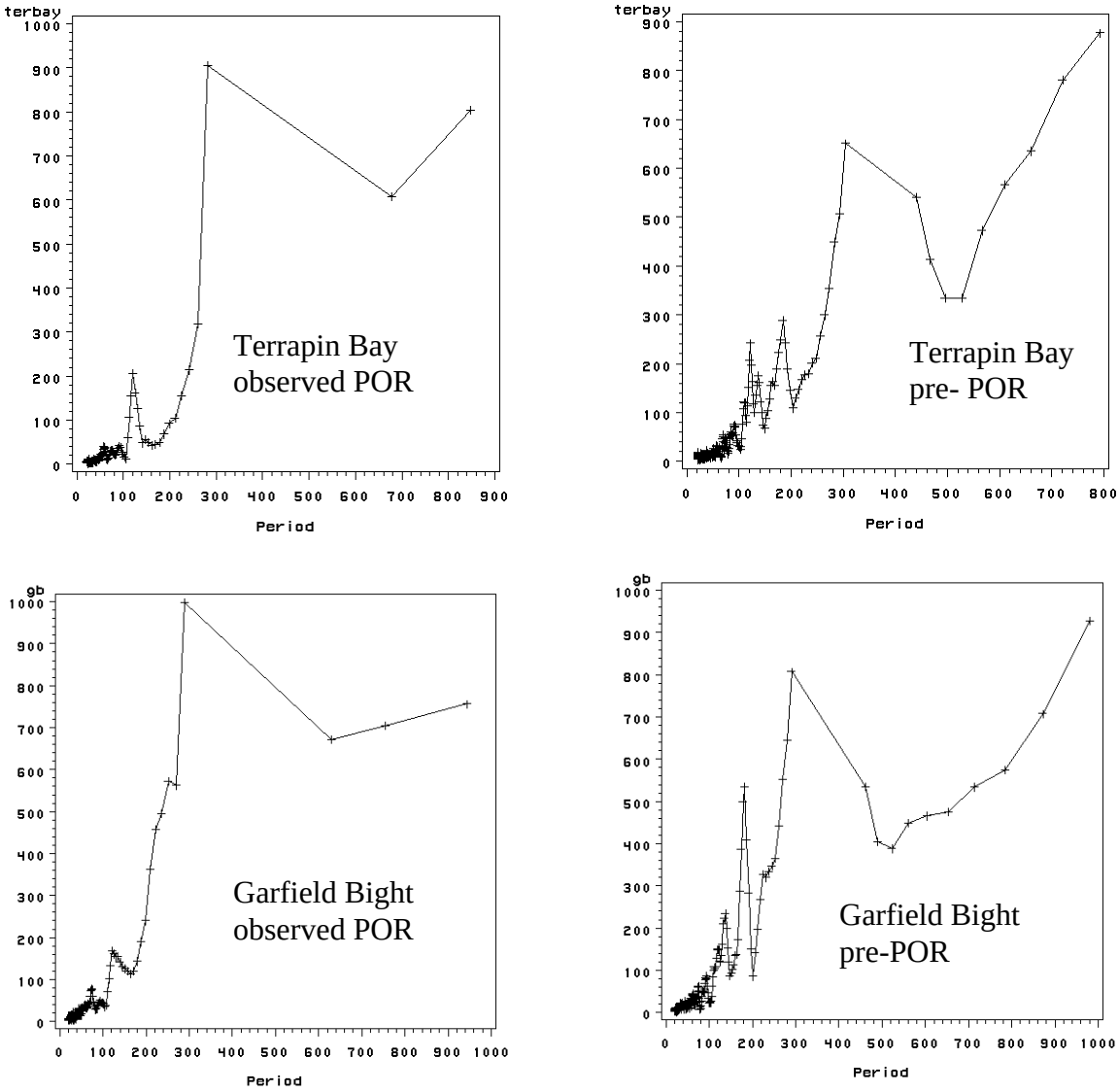


Figure 21, cont. Comparison of spectral density function values for observed data (left column) and for hindcast reconstruction simulation data (right column).

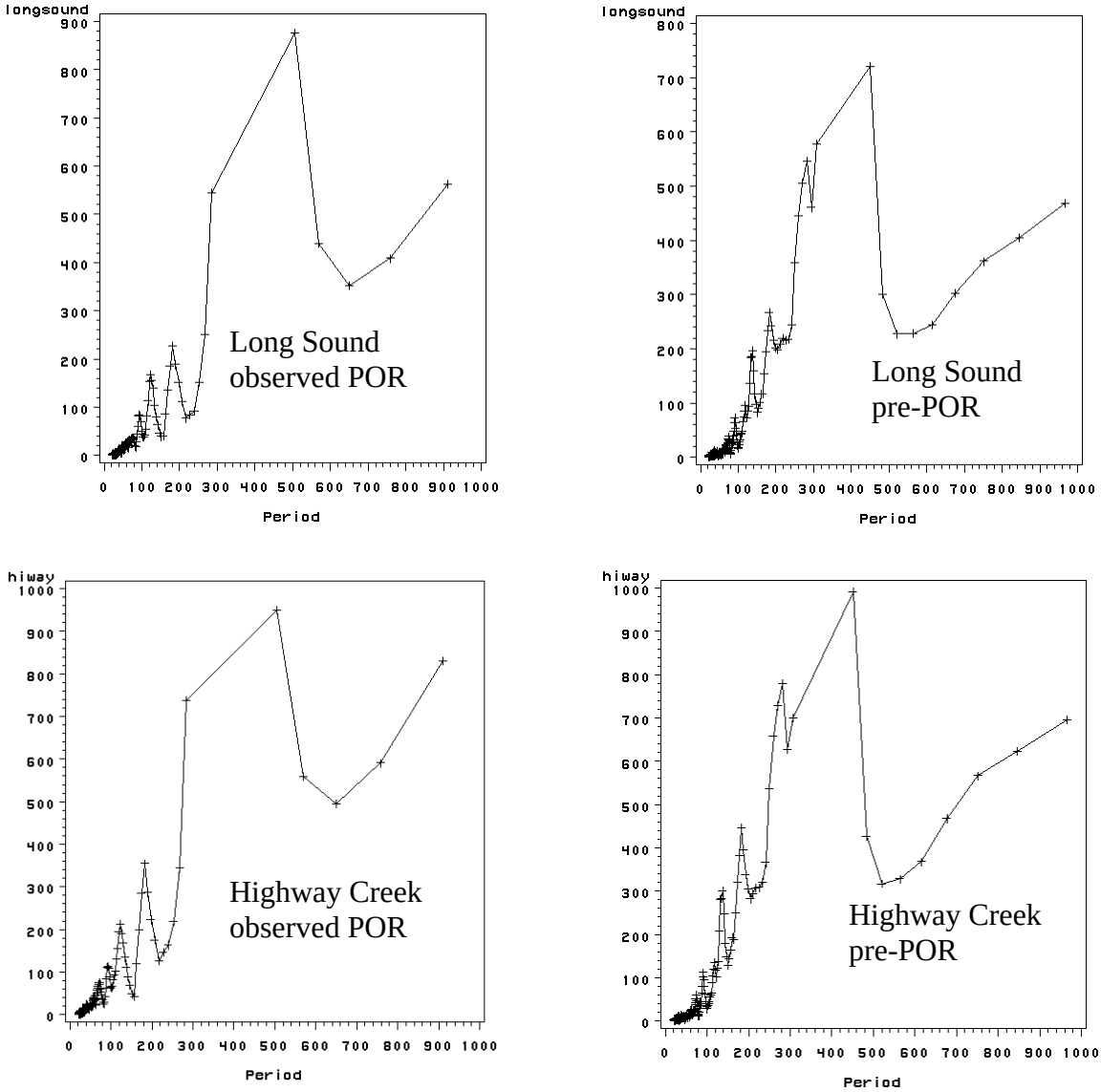


Figure 21, cont. Comparison of spectral density function values for observed data (left column) and for hindcast reconstruction simulation data (right column).

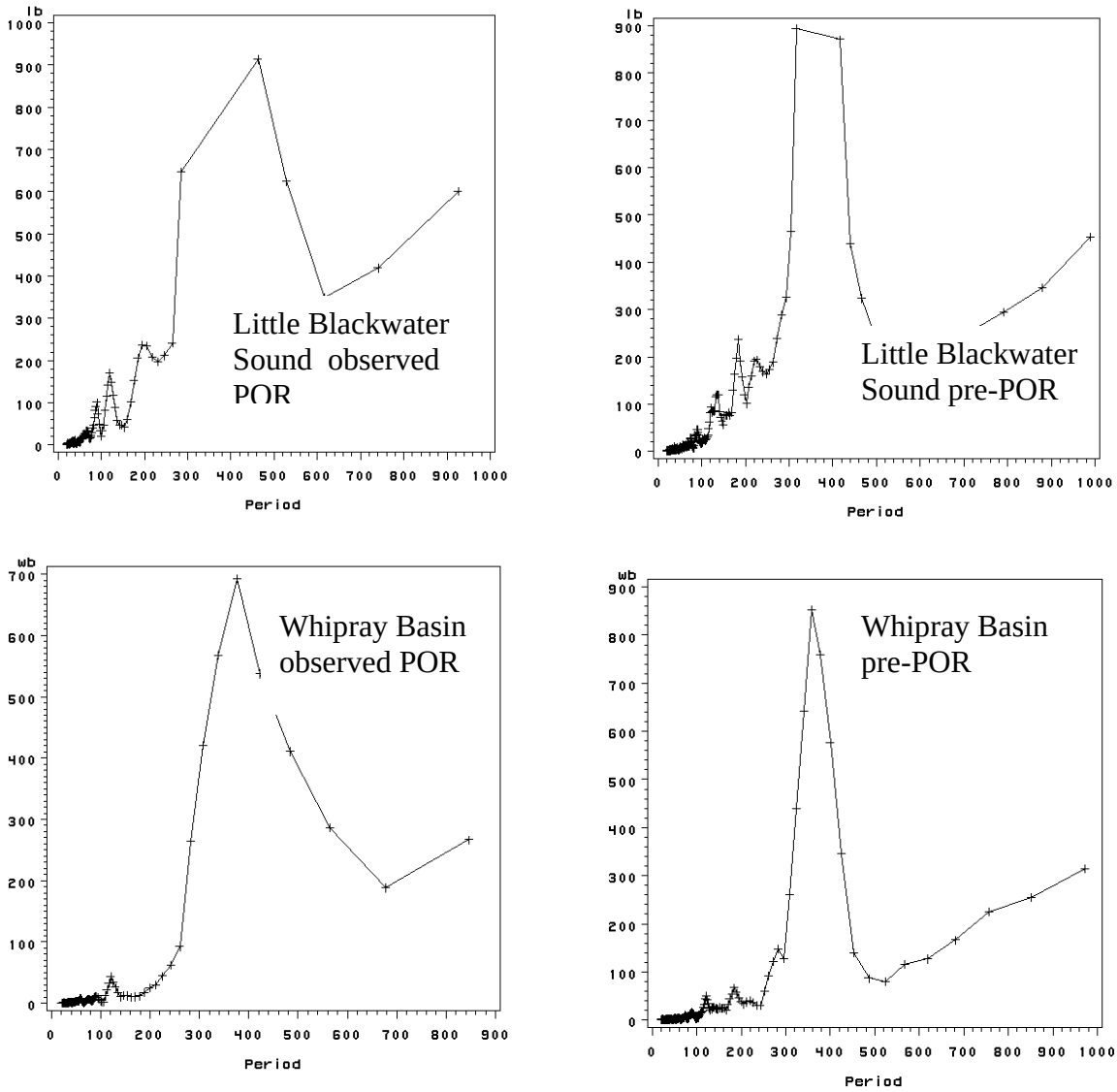


Figure 21, cont. Comparison of spectral density function values for observed data (left column) and for hindcast reconstruction simulation data (right column).

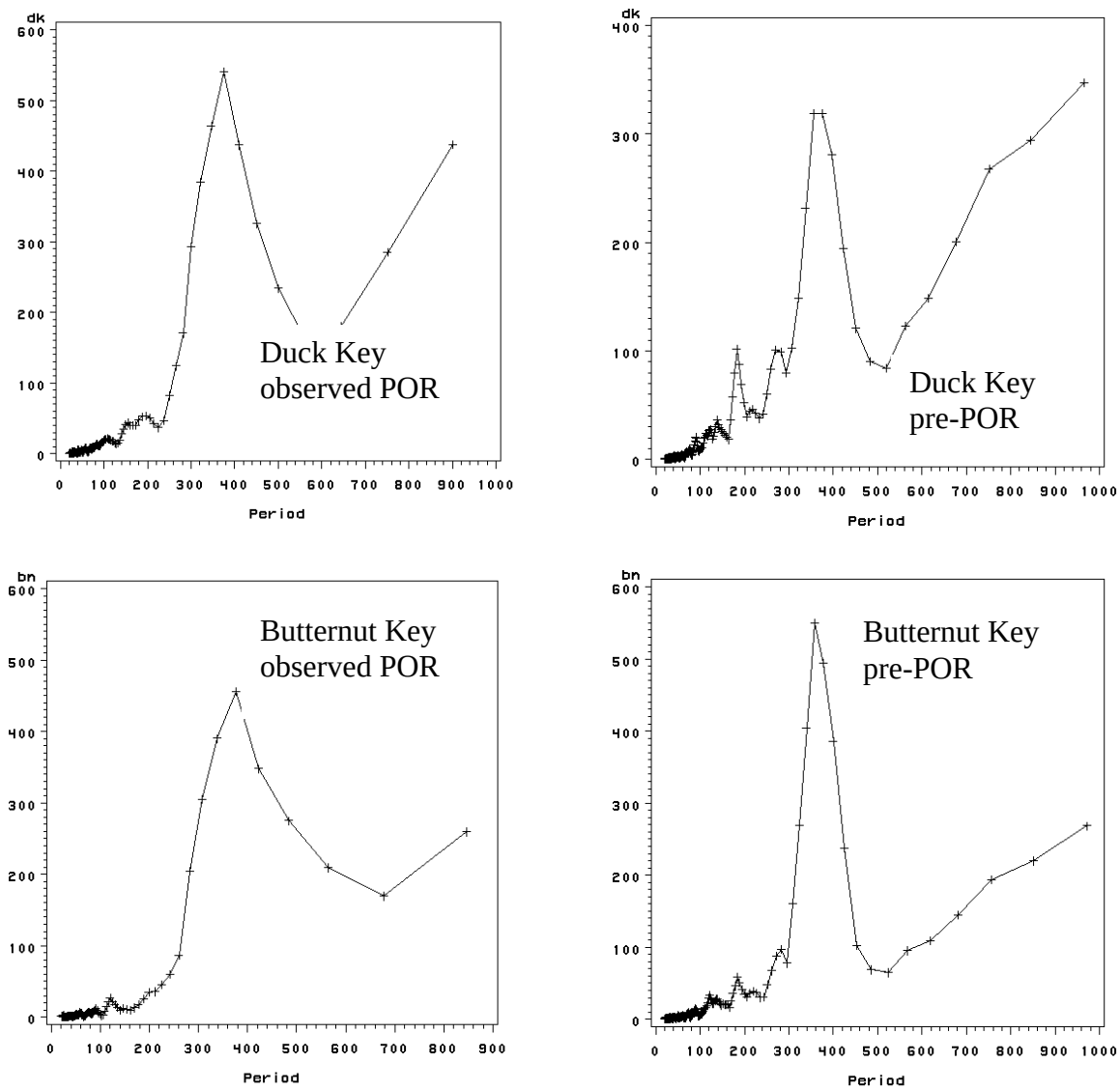
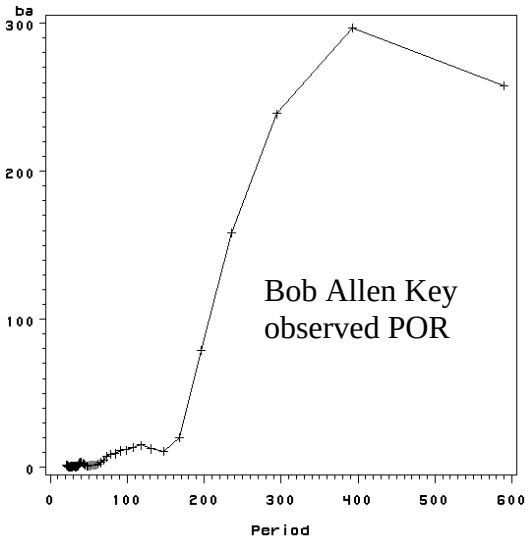


Figure 21, cont. Comparison of spectral density function values for observed data (left column) and for hindcast reconstruction simulation data (right column).



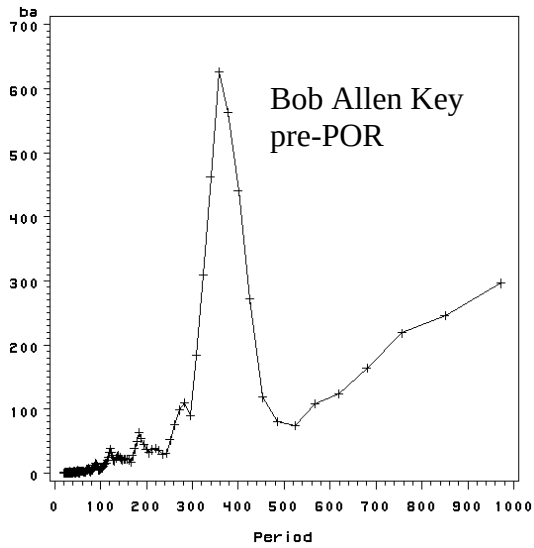


Figure 21, cont. Comparison of spectral density function values for observed data (left column) and for hindcast reconstruction simulation data (right column).

VI. Discussion

The purpose of this task was to prepare time series of salinity that extended back in time until January 1, 1970 which is beyond the beginning of the period of record for the MMN stations. In some cases, for example Joe Bay, Little Madeira Bay, Long Sound, and Highway Creek, the period of record begins in 1988. On the other end of the spectrum, the Bob Allen Key period of record doesn't begin until 1997. The primary tool for filling in the gaps - internally in the continuously-monitored observed data but also in the period prior to the date of data collection - are multivariate linear regression models for salinity that have been used successfully to-date for CERP performance measure evaluations. For CERP use, the South Florida Water Management Model (2X2 Model) providing the stage inputs and observed data were utilized for wind and Key West water level input parameters.

The 2X2 Model data begin on January 1, 1965. The observed wind and Key West water level data used for the MLR salinity simulations begin on the same date. For this CESI task it was beyond the scope of work to obtain additional wind or Key West water level data prior to 1965. Therefore, the earliest point in time that the reconstructions can begin with this data set is January 1, 1965. For the Florida Bay Minimum Flows and Levels modeling (ECT, 2005) the historical reconstruction from 1965 to 1970 was not particularly useful since the construction of Tamiami Trail in the current configuration including the S-12 structures was completed in the 1960's and water levels for next decade reflect an area in transition. Initial operations of these facilities interrupted almost completely the flow of water into Shark River Slough. However, due to intervention by ENP, additional flows to the Park were augmented through operational activities beginning in about 1970. This limitation was also realized for the historical reconstructions using the FATHOM model in Florida Bay and for the MLR model for Taylor River for the Florida Bay Minimum Flows and Levels (ECT, 2005). For these reasons, the historical reconstructions in this task begin on January 1, 1970 and end on December 31, 2000.

For the historical reconstructions to be useful, there must be confidence that the simulations prior to the observed data represent a reasonable estimate of the salinity regime that occurred during the period of the reconstruction. Occurrence of extreme high and low salinity values can be used to test the responsiveness of the models. For four of the MMN stations (Joe Bay, Little Madeira Bay, Long Sound, and Highway Creek) there are good examples of both high and low extremes in the observed record. For example, during the drought year of 1989, salinity values were recoded over 50 psu (Joe Bay), 70 psu (Little Madeira Bay), and 40 psu (Long Sound and Highway Creek). As a comparison to a simulated situation, at Bob Allen Key monitoring began in September 1997, meaning that the salinity values of the Bob Allen Key reconstruction in 1989 were from the MLR model. The plots show that simulated salinity at Bob Allen Key exceeded 60 psu in 1989, and values above 60 psu in the Bob Allen Key reconstruction are

only seen in the reconstruction 2 others years, 1971 and 1974. In fact, all of the reconstructions show an increase in salinity during this drought period of the 1970's, which means that the reconstructions are sensitive to dry period conditions.

At the other end of the salinity scale, lowest salinity values in the observed record of these 4 stations are seen in 1995 and the following years, though low salinity values at these stations during the observed record are much more common than high salinity values. At Bob Allen Key the simulated 1995 and 1996 salinity values are also low, as are the observed values in 1997 (when monitoring began) and in the ensuing years.

Comparisons of salinity reconstructions at the open-water stations with salinity at the near shoe stations indicates that there are times when high salinity conditions develop at the near shore stations in the spring at the end of the dry season, but the salinity conditions at the open-water stations remain at average levels. However, for some years (1971, 1974, and 1989 through 1991) the highest salinity values in the reconstruction at seen at both near shore and open-water locations. It appears that the dry conditions must reach a certain level of severity (or perseverance) before the open-water stations begin to experience hypersaline conditions. In average years, the highest salinity value of the year is about 20 psu higher at Bob Allen Key compared to Long Sound. However for the years of the highest salinity values when hypersaline conditions were recorded or simulated at both locations, the difference in the highest salinity values of the year at these two stations is only about 10 psu (Figure 22).

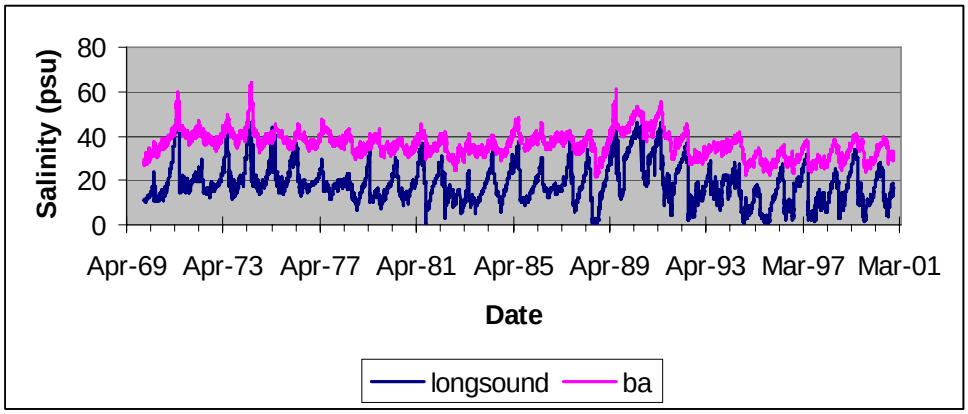


Figure 22. Comparison of historical salinity reconstructions from a near shore station (Long Sound) and an open-water station (Bob Allen Key, or BA).

Comparison of the spectral density plots for observed data and reconstructed data for the period prior to the observed period of record (POR) shows that the reconstructed time series have similar periodicities to the observed data, though some differences at some of the stations is noted (Figure 13) as follows:

- Joe Bay – The 120+/- day maximum seen in the observed POR is not seen in the reconstruction POR,
- Little Madeira Bay – The 183+/- day maximum is lower for the observed POR,
- Terrapin Bay – The 120+/- day maximum seen in the reconstruction POR is not seen in the observed POR,
- Garfield Bight – The reconstruction POR 183+/- day maximum value is much higher for the than the observed value,
- Long Sound – The observed POR and the reconstruction POR maximum values are similarly expressed but with different values,
- Highway Creek – Similar to Long Sound, the observed POR and the reconstruction POR maximum values are similarly expressed but with different values,
- Little Blackwater Sound - The observed POR and the reconstruction POR maximum values are similarly expressed but with different values,
- Whipray Basin – The 183+/- day maximum is expressed in the reconstruction POR but not in the observed POR,
- Duck Key - The 183+/- day maximum value is much higher in the reconstruction POR than in the observed POR,
- Butternut Key - The 183+/- day maximum value is much higher in the reconstruction POR than in the observed POR,
- Bob Allen – The Bob Allen observed POR does not begin until 1997, so the number of values is small and spectral density function values should only be considered preliminary values. However, the observed POR plot is similar (in general) to the other observed POR plots. The pre-POR plot expresses the maximum values at the typical periods (120+/-, 183+/-, 365+/- days).

Because the number of values used to compute the spectral density function for the longer periods is limited to about 20 or less, the differences between the observed POR and the reconstruction (pre-) POR are not considered to be significant. The important information is the recurring seasonality in the observed data is being expressed in the simulated reconstructions.

In addition, it appears that the near shore stations are subject to relatively significant cyclic forcing that is not on the order of yearly, semi-annually, or quarterly intervals. The open-water stations do not show the same behavior in the spectral density function plots.

The near shore locations appear to be subject to hydrologic forcing that is more complicated than Little Madeira Bay, Whipray Basin, Duck Key, Butternut Key, and Bob Allen Key, as expressed in the number of secondary maxima seen in

the spectral density function plots of the near shore stations at the more frequent periods (low values of period). In addition, at Whipray Basin, Duck Key, Butternut Key, and Bob Allen Key the value of the spectral density function does not increase to a secondary maximum at higher periods, as is seen in the plots of the near shore stations, including Little Madeira Bay. In the case of Joe Bay, Terrapin Bay, Long Sound, Highway Creek, and Little Blackwater Sound, the value is still increasing significantly at 1000 days. For Little Madeira Bay and Garfield Bight the spectral density function value has leveled off at a relatively high value at 1000 days. At Whipray Basin, Duck Key, Butternut Key, and Bob Allen Key the spectral density function levels off at a lower value of period.

This analysis shows that the reconstructions comprised of observed data and simulations using MLR salinity models and other statistical relationships provide a reasonable estimate of the historical conditions since 1970. These historical reconstructions of salinity are intended for use with ecologic models such as available fish models (ENP and others), pink shrimp models, and seagrass models by other researchers.

VII. References

Environmental Consulting & Technology, Inc. (ECT). 2005. Final Report FATHOM Enhancements and Implementation to Support Development of Minimum Flows and Levels for Florida Bay. Ed. By F. Marshall, W. Nuttle, and B. J. Cosby. ECT, New Smyrna Beach, Florida. Pages 126 – 127.

Marshall III, F.E.; D. Smith; and D. Nickerson. 2003. Salinity Simulation Models for North Florida Bay Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. 41 pp.

Marshall III, F.E. 2003. IOP Salinity Analysis Using Statistical Models. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. 35 pp.

FINAL TASK REPORT

TASK 12 - EVALUATE THE FEASIBILITY OF DEVELOPING MLR SALINITY MODELS FOR THE MANGROVE LAKES REGION NORTH OF GARFIELD BIGHT AND TERRAPIN BAY

Cooperative Agreement Number CA H5284-05-0006
Between
The United States Department of the Interior National Park Service
Everglades National Park
And
Cetacean Logic Foundation, Inc.

Frank E. Marshall, III
Principal Investigator

DeWitt Smith
Project Coordinator

April 12, 2008

CESI Project No. CA H5284-05-0006 for Everglades National Park

FINAL TASK REPORT

TASK 12 - EVALUATE THE FEASIBILITY OF DEVELOPING MLR SALINITY MODELS FOR THE MANGROVE REGION NORTH OF GARFIELD BIGHT AND TERRAPIN BAY

I. Introduction

A. General

This report describes the activities that were completed for a task in the current Critical Ecosystem Studies Initiative (CESI) project that was originally initiated by Cetacean Logic Foundation, Inc. (CLF) for Everglades National Park (ENP) in 2002. The goal of this on-going CLF research is the characterization and simulation of the salinity regime in Florida Bay and the estuaries along southwest coast of the Gulf of Mexico (Gulf) within the Park. It is intended that this research yield information regarding the link between the downstream salinity in the estuaries of ENP and the upstream freshwater hydrology of the Everglades.

The subject of this report is the initial statistical evaluation of salinity and the use of multivariate linear regression (MLR) models to simulate salinity in the mangrove lakes region north of Garfield Bight and Terrapin Bay. MLR salinity models were previously developed for the 33 stations in the Marine Monitoring Network (Marshall, 2005a; 2005b). Monitoring is currently on-going by ENP staff and others in the mangrove lakes area. The purpose of this study is to do a statistical analysis of the limited data and make a preliminary evaluation of the use of MLR salinity models developed from the data being collected.

B. Task Objectives and Evaluation Methods

The objective of this task is presented below as it appears in the modified contract for this CESI project approved by ENP:

> "The Florida Bay coastal embayments from Terrapin Bay to Garfield Bight and the mangrove zone lakes in this area are hydrologically connected to each other and to Florida Bay. MLR salinity models have already been prepared for both Terrapin Bay and Garfield Bight. Restoration of the ecology of these lakes (e.g., Seven Palms Lake, West Lake, and The Lungs) and the embayments are a focus of CERP activities. A project to collect submerged aquatic vegetation data, salinity, water level (stage), and other water quality data in the Florida Bay mangrove zone and coastal embayments of this central bay region has recently been initiated by ENP (Morrison, 2005). Data have been previously collected by ENP and the USGS in this area, and salinity data are available for a several year period that may be able to be used for statistical modeling in a similar manner to the MLR salinity models

already prepared. The feasibility of developing statistical salinity models for this area will be investigated.

The deliverable is a task report that describes the feasibility of developing statistical salinity models for the mangrove zone lake areas near Terrapin Bay and Garfield Bight."

II. Background

Work is currently underway in the Florida Bay mangrove lakes region to study submerged aquatic vegetation as part of a CESI project. As part of this project physiochemical data are being acquired. The research project specifics are as follows:

> Project name: Submerged aquatic vegetation and physicochemical monitoring in the Florida Bay mangrove zone for CERP assessments and targets refinement
>
> Cooperative Agreement CA H5297-05-0041
> ENP permit # EVER-2006-SCI-0022
>
> Investigators: Thomas A. Frankovich , James W. Fourqurean, Douglas Morrison

At the time of this evaluation, two reports had been filed to record project progress: Semi-annual Progress Report - October 18, 2006; and Annual Progress Report – March 23, 2007. Both of these reports were reviewed. Temperature and salinity data are presented for the period April/May 2006 – January/February 2007 for 8 stations in the region. This data are to be analyzed, if possible, for use with the development of MLR salinity models to determine if there is the potential to link the salinity of this region to the hydrology that affects it, the freshwater marsh of the Everglades.

III. Data

Figure 1 presents a site map of the Florida Bay mangrove lakes region where this research project is underway. The primary objective of this research is to document the submerged aquatic vegetation (SAV) in the area and to collect physiochemical data in this environment.

Figure 1. Site map copied from Frankovich et al, 2007. Continuous salinity and temperature data were collected at the stations indicated with a blue D. Stations in red are monthly water quality stations that were not used in this analysis.

Salinity and temperature were monitored continuously at Seven Palms Lake, West Lake, and The Lungs from July 2006 to February 2007, and monthly at . In general, at Seven Palms Lake salinity ranged from 4.1 ppt to 28.4 ppt; and at West Lake the salinity ranged from 12.2 ppt to 22.8 ppt; and at The Lungs the salinity ranged from 11.1 ppt to 44.4 ppt. Peak salinity was noted at 7 Palms Lake in July and in June at the other three stations. Minimum levels occurred in October at 7 Palms Lake and Middle Lake,and a month later at West Lake. Hypersaline conditions were recorded at The Lungs at the beginning of the period of recorded data, July 2006.

Salinity and temperature were monitored monthly at multiple stations in the following locations: 7 Palms Lake, Middle Lake, Monroe Lake, Terrapin Bay, West Lake, Long Lake, The Lungs, and Garfield Bight (Figure 1). The salinity ranges during the April 2006 to January 2007 monitoring period are as shown in Table 1.

Table 1. Maximum and minimum salinity from monthly monitoring runs for the April 2006 to January 2007 period. Values are approximate and were taken from Figures 8 and 9 of Frankovich et al, 2007. Garfield Bight data includes only 4 samples.

Station	Maximum Salinity, ppt	Month	Minimum Salinity, ppt	Month
7 Palms Lake	30	July	5	September
Middle Lake	37	May	5	September
Monroe Lake	42	July	5	September
Terrapin Bay	45	June	10	September
West Lake	29	May	15	October
Long Lake	52	June	12	September
The Lungs	52	June	11	September
Garfield Bight	50	April	20	September

The largest salinity range was observed in the water bodies closest to Florida Bay (Frankovich et al, 2007). Monthly salinity in those area ranged from hypersaline (40 to 50 ppt) in June and July, to hyposaline (0 to 20 ppt) in September/October.

IV. Methods and Results

Equipment problems affected the collection of continuous salinity data at the beginning of the project. Continuous salinity data were made available from June 27, 2006 through August 28, 2007, approximately 13 months. To begin the analysis, simple statistics were computed for all three salinity modeling stations. Then the data from Seven Palms Lake, The Lungs, and West Lake were used to prepare correlation matrices using the salinity data, stage data from upstream locations in the Everglades freshwater marsh, wind vectors from the Long Key C-Man station, and water level elevation from the Long Key C-Man station. Using guidance in parameter selection from the information in the correlation matrix, preliminary MLR salinity models were developed using stepwise regression procedures to explore the potential for MLR salinity model development once additional data have been collected.

The simple statistics on the data used for this analysis are presented in Table 2. There were between 215 and 301 salinity values available for use in the evaluation. The additional values in the dependent variable data result from differing beginning and ending dates. Mean salinity value over the period in the Lungs is noticeably higher than the mean salinity values at Seven Palms Lakes and West lake which are similar. The minimum salinity value at Seven Palms Lake was lower than at the other two stations, which had similar minimum salinity

values. The highest maximum value of salinity was at The Lungs which reached about 47 psu.

Table 2. Simple statistics for data used for the analysis of regression modeling for the mangrove lakes area of ENP.

Variable	N	Mean	Std Dev	Minimum	Maximum
Sal7Palms	301	13.09	6.20	4.15	27.78
SalLungs	215	27.26	8.53	13.95	47.01
SalWest	300	15.47	2.48	12.62	23.51
StgCP	428	1.35	0.42	0.56	2.16
StgP33	428	6.15	0.33	5.45	6.67
StgNP46	428	1.53	0.48	-0.03	2.45
StgNP62	415	2.76	0.43	1.46	3.56
StgNP206	428	5.59	0.67	3.34	6.35
StgP35	428	1.69	0.43	0.68	2.73
StgP37	428	1.56	0.39	0.65	2.33
StgP38	428	1.70	0.42	0.56	2.63
WlvLong1	400	0.70	0.26	0.19	1.35
WuLong1	400	-2.35	3.36	-9.98	8.59
WvLong1	400	-0.71	3.15	-9.78	8.70

The correlation matrix for unlagged values is presented in Table 3. The highest Pearson correlation coefficient for correlation with salinity was P33 for all stations. However the correlation coefficient was positive for P33 and West Lake, seemingly indicating that higher values of P33 stage were related to higher values of salinity in West Lake. Positive correlation coefficients are also seen for CP at both Seven Palms Lake and at West Lake. The correlation coefficients for CP at Seven Lakes Palms and West Lake show a medium level of correlation but they are opposite in sign. The correlation matrix for one-day-lag values is presented in Table 4, and the two-day-lag correlation matrix is presented in Table 5. Comparison of this correlation matrix with the correlation matrix for unlagged values shows that the correlation coefficient values were similar to the unlagged values.

Table 3. Correlation matrix for unlagged values of stage, wind, water level and salinity.

	Sal7Palms	SalLungs	SalWest
StgCP	0.3187	-0.44098	0.07688
StgP33	-0.50754	-0.7777	-0.41263
StgNP46	0.36571	-0.16816	0.165
StgNP62	0.23174	-0.46333	0.04888
StgNP206	0.09482	-0.52332	-0.06336
StgP35	0.2572	-0.22055	0.08097
StgP37	0.43856	-0.23116	0.15849
StgP38	0.2297	-0.4156	0.09965
WlvLong1	0.05757	-0.13833	-0.20612
WuLong1	-0.01508	0.01113	-0.12202
WvLong1	0.24254	0.1095	0.17167

Table 4. Correlation matrix for one-day-lag values of stage, wind, water level and salinity.

	Sal7Palms	SalLungs	SalWest
StgCPlag1	0.29177	-0.46833	0.04958
StgP33lag1	-0.52868	-0.78086	-0.43663
StgNP46lag1	0.34405	-0.1767	0.14476
StgNP62lag1	0.20749	-0.48326	0.02422
StgNP206lag1	0.07263	-0.53061	-0.08281
StgP35lag1	0.23592	-0.2394	0.05332
StgP37lag1	0.41373	-0.26111	0.13263
StgP38lag1	0.20427	-0.43469	0.07621
WlvLong1lag1	0.05504	-0.15206	-0.214
WuLong1lag1	-0.00064	0.02488	-0.10461
WvLong1lag1	0.23276	0.15125	0.14021

Table 5. Correlation matrix for two-day-lag values of stage, wind, water level and salinity.

	Sal7Palms	SalLungs	SalWest
StgCPlag2	0.26566	-0.49281	0.02291
StgP33lag2	-0.54874	-0.78329	-0.45917
StgNP46lag2	0.32317	-0.18093	0.12714
StgNP62lag2	0.18484	-0.5008	0.00222
StgNP206lag2	0.05195	-0.54095	-0.09971
StgP35lag2	0.21391	-0.27023	0.02566
StgP37lag2	0.39055	-0.28894	0.10804
StgP38lag2	0.18013	-0.45157	0.05458
WlvLong1lag2	0.05267	-0.18065	-0.21583
WuLong1lag2	0.00845	0.00405	-0.09019
WvLong1lag2	0.2384	0.169	0.13544

Preliminary models were developed for all three stations using the stepwise linear regression procedure. The significance level for including an independent variable in the model had to be set at 0.999 similar to other MLR salinity models to reduce the number of parameters in the model to a reasonable level. At the 0.950 significance level there were between 15 and 20 independent variables in the model. There is an additional benefit in setting the significance level at 0.999 – all independent variables in the final model are highly significant.

Tables 6 and 7 present the analysis of variance table for the Seven Palms Lake preliminary salinity model and the parameter estimates, respectively. Table 8 presents the partial R^2 values for the parameters, assuming the previous parameters were already included in the model. Table 9 presents the analysis of variance table for the Lungs and Table 10 presents the preliminary salinity model parameter estimates. Table 11 presents the partial R^2 values for the parameters of The Lungs MLR salinity model, assuming the previous parameters were already included in the model. For West Lake, Table 12 is the analysis of variance table, Table 13 is the parameter estimate table, and Table 14 presents the partial R^2 values.

Table 6. Analysis of variance table for Seven Palms Lake preliminary MLR salinity model.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	9012.62	1802.52	255.29	<.0001
Error	277	1955.83	7.06		
Corrected Total	282	10968.00			

Table 7. Parameter estimate table for Seven Palms Lake preliminary MLR salinity model.

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	113.05	3.21	8745.88	1238.66	<.0001
StgCP	4.91	1.05	153.59	21.75	<.0001
StgP33lag2	-19.64	0.59	7722.71	1093.75	<.0001
StgNP46lag2	3.60	0.75	162.12	22.96	<.0001
StgP35lag2	4.99	0.85	243.59	34.5	<.0001
WuLong1lag2	-0.18	0.05	85.40	12.09	0.0006

Table 8. Partial R^2 table for Seven Palms Lake preliminary MLR salinity model.

Variable	Partial R-Squar	Model R-Square	C(p)	F Value	Pr > F
StgP33lag2	0.31	0.31	1070.97	124.72	<.0001
StgCP	0.47	0.78	155.70	593.57	<.0001
StgNP46lag2	0.02	0.80	116.99	28.98	<.0001
StgP35lag2	0.02	0.81	89.74	22.42	<.0001
WuLong1lag2	0.01	0.82	76.56	12.09	0.0006

Table 9. Analysis of variance table for The Lungs preliminary MLR salinity model.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	11816	2954.066	157.55	<.0001
Error	199	3731.163	18.74956		
Corrected Total	203	15547			

Table 10. Parameter estimate table for The Lungs preliminary MLR salinity model.

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	115.53	7.03	5061.61	269.96	<.0001
StgCPlag2	-26.83	2.81	1705.20	90.95	<.0001
StgP33lag2	-14.86	1.20	2853.74	152.2	<.0001
StgNP46lag2	12.67	2.29	574.07	30.62	<.0001
StgP35lag1	11.00	1.45	1078.32	57.51	<.0001

Table 11. Partial R^2 table for The Lungs preliminary MLR salinity model.

Variable	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
StgP33lag2	0.62	0.62	208.61	323.41	<.0001
StgCPlag2	0.02	0.64	184.24	13.87	0.0003
StgP35lag1	0.08	0.72	98.30	59.76	<.0001
StgNP46lag2	0.04	0.76	61.06	30.62	<.0001

Table 12. Analysis of variance table for West Lake preliminary MLR salinity model.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	982.17	163.70	60.51	<.0001
Error	258	697.94	2.71		
Corrected Total	264	1680.11			

Table 13. Parameter estimate table for West Lake preliminary MLR salinity model.

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	44.58	2.01	1329.14	491.33	<.0001
StgP33lag2	-6.04	0.43	539.36	199.38	<.0001
StgNP62lag2	5.11	1.05	63.79	23.58	<.0001
StgNP206lag2	-2.00	0.48	46.18	17.07	<.0001
StgP35lag2	4.35	0.65	122.76	45.38	<.0001
WuLong1lag2	-0.14	0.04	45.77	16.92	<.0001
WlvLong1lag2	-2.71	0.51	76.83	28.4	<.0001

Table 14. Partial R^2 table for West Lake preliminary MLR salinity model.

Variable	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
StgP33lag2	0.21	0.21	413.99	69.5	<.0001
StgNP62lag2	0.23	0.44	215.94	110.36	<.0001
StgNP206lag2	0.06	0.50	166.49	31.71	<.0001
StgP35lag2	0.02	0.52	151.11	11.13	0.001
WlvLong1lag2	0.03	0.56	124.74	19.45	<.0001
WuLong1lag2	0.03	0.58	103.50	16.92	<.0001

The Seven Palms Lake preliminary MLR salinity model has an R^2 of 0.82. The Lungs preliminary MLR salinity model has an R^2 of 0.77, and the West Lake preliminary MLR salinity model has an R^2 of 0.58. The mean square error (MSE) for all three models is relatively low.

This was an initial exercise at preparing MLR models with this data. All parameters chosen by the SAS© stepwise procedure are reported. If this task included the preparation of a final model the parameters that were not consistent with physical reality would be removed, particularly positive values of coefficients for stage independent variables. Usually they can be removed with little or no penalty in explanation of variability, because the partial R^2 value is usually low. . However, for the Seven Palms Lake and West Lake preliminary models, one stage independent variable has a positive coefficient yet explains a relatively large portion of the variability. For Seven Palms Lake, CP has the positive coefficient; for West Lake the stage independent variable with a positive coefficient was NP62. In both models the day 2 lag term for P33 explains the greatest proportion of variability, perhaps suggesting that a hydraulic gradient relationship should be investigated.

V. Findings and Recommendations

Though there is only limited data available at this time from this project, it appears that the salinity data being collected ultimately will be useful for salinity modeling. The mangrove and lakes area of the north central part of Florida Bay where the salinity measurements are being taken represents an area with few monitoring stations. If extended continuous monitoring shows the salinity range seen for the limited monitoring, then this is an important area for monitoring because (1) salinity is being measured across a large salinity range, (2) hypersalinity is being measured, and (3), salinity values are not dropping to 0 ppt. At these locations, restoration in the form of additional freshwater flows towards western and central Florida Bay will be most noticeable, as will lower flows (if that

occurs) and increased salinity due to sea level rise. Since these stations do not go to zero they are useful for simulating changed conditions due to restoration or other alterations to freshwater flow and stage. If paleoecological investigations could be extended into this area, it may be useful for estimating the conditions before freshwater flow alterations around the turn of the 20th century.

The model development effort produced models that could be interpreted in a limited manner because of the short period of the data. The stage parameters that explained the greatest proportion of variability were P33 with a 2 day lag (as suggested by the correlation matrix), with lagged terms of CP, P35, NP46, filling out the Seven Palms Lake and The Lungs models.. The West lake model appears to be different than the other two – different stage parameters support the lagged P33 and P35 parameters (Np62, Np206). Wind parameters were significant at Seven Palms Lake and West Lake but not at the Lungs. Water level measured at Long Key was only significant in the West Lake model.

The preliminary modeling activities suggest that the data from these three stations should be capable of producing reasonable MLR salinity models using stage, wind, and water level parameters. The models produced herein are preliminary models prepared for analysis only. These models illustrate important relationships and should NOT be used for predictive purposes without further analysis. However, this initial evaluation suggests that this further analysis will yield models with R^2 values in the 0.55 to 0.80 range.

VI. References

Frankovich, Thomas A.; James W. Fourqurean, and Douglas Morrison. March 23, 2007. Submerged aquatic vegetation and physicochemical monitoring in the Florida Bay mangrove zone for CERP assessments and targets refinement. Annual Progress Report. Critical Ecosystem Studies Initiative Cooperative Agreement CA H5297-05-0041.

Marshall III, F. E. 2005a. RECOVER Southern Estuaries performance measures: identification of hydrology-salinity relationships for coastal estuaries and analysis of interim CERP update scenarios. Environmental Consulting & Technology, Inc., New Smyrna Beach, Florida.

Marshall III, F.E.; and D. Smith. 2005. Investigation of evaporation surrogates for Florida Bay salinity modeling. Task Report for a Critical Ecosystems Initiative (CESI) Project. Cetacean Logic Foundation, Inc., New Smyrna Beach, Florida.

FINAL TASK REPORT

TASK 13 - PREPARE AND SUBMIT A MANUSCRIPT FOR PUBLICATION IN A REFEREED JOURNAL

Cooperative Agreement Number CA H5284-05-0006
Between
The United States Department of the Interior National Park Service
Everglades National Park
And
Cetacean Logic Foundation, Inc.

Frank E. Marshall, III
Principal Investigator

DeWitt Smith
Project Coordinator

April 30, 2008

CESI Project No. CA H5284-05-0006 for Everglades National Park

FINAL TASK REPORT

TASK 13 - PREPARE AND SUBMIT A MANUSCRIPT FOR PUBLICATION IN A REFEREED JOURNAL

I. Introduction

A. General

This report describes the activities that were completed for a task in the current Critical Ecosystem Studies Initiative (CESI) project that was originally initiated by Cetacean Logic Foundation, Inc. (CLF) for Everglades National Park (ENP) in 2002. The goal of this on-going CLF research is the characterization and simulation of the salinity regime in Florida Bay and the estuaries along southwest coast of the Gulf of Mexico (Gulf) within the Park. It is intended that this research yield information regarding the link between the downstream salinity in the estuaries of ENP and the upstream freshwater hydrology of the Everglades.

The subject of this report is the preparation of a manuscript for publication in a refereed journal. Another task in this CESI project developed a methodology for estimating the flow needed at key locations in the Everglades to meet a salinity criterion in Florida Bay. This methodology was then used with paleoecological information about the salinity regime that existed prior to drainage alterations to estimate a “pre-drainage” flow for use as one benchmark for restoration of the Everglades and Florida Bay. The purpose of this task was to prepare a manuscript describing the methods used to estimate pre-drainage flow and make a preliminary flow estimate. This task report includes the manuscript as an appendix, so the task report will not present information that is included in the manuscript.

B. Task Objectives and Evaluation Methods

The objective of this task is presented below as it appears in the modified contract for this CESI project approved by ENP:

> “Although the work performed over the past several years on developing the salinity models for the MLR network has produced several project reports and a number of presentations at conferences and seminars, a formal paper presenting the development work has not been submitted to a scientific journal and subjected to peer review. The purpose of this task is to produce a manuscript that is suitable for submittal and review by a refereed journal, such as *Estuaries*, *Water Resources Research*, or *ASCE Proceedings*. The paper will discuss the development of MLR salinity models, present some if not all of the models, and present implementation examples as are appropriate and sufficiently finalized. Included in this task are site visits to

each of the monitoring stations in the Everglades where water levels are being measured, and to the salinity monitoring stations that have not yet been visited."

II. Background

A parallel task in this project (Task 9) developed a methodology for estimating the flow needed in the Everglades to meet the Florida Bay and Florida Keys Feasibility Study (FBFKFS) salinity criteria in Florida Bay. The flow estimating methodology was presented to the Southern Estuaries Sub-team in the fall of 2006. The methodology was of interest to this group since they are charged with evaluating how well the CERP alternatives perform compared to "restoration" standards. The Southern Estuaries performance measures are similar to the FBFKFS criteria. Therefore, the estimates that had been made for the hydrologic conditions needed to achieve the FBFKFS salinity criteria were relevant to their work.

At that time, the Southern Estuaries Sub-team was also evaluating the use of paleoecological data collected by the USGS to estimate the conditions in Florida Bay prior to the construction of water management structures and the alteration of the salinity regime. At the Florida Bay Science Program meeting on Duck Key in December 2005, the Principal Investigator and the ENP Project Coordinator met with the Southern Estuaries Sub-team and it was decided that it would be beneficial to all to use the flow estimating methodology of Task 9 with USGS Whipray Basin paleoecological information to make an estimate of the flow needed to achieve the freshwater flow conditions that may have resulted from the climatology of 1965-2000 assuming that no water management structures had been constructed. In order to have the estimate peer reviewed, the decision was made by the Principal Investigator and the ENP Project Coordinator to prepare the manuscript describing the flow estimation made using the CESI-developed flow estimation procedure and the USGS paleoecological information, instead of developing a manuscript for the MLR models.

At about the same time that the development of the flow estimating methodology was completed, the South Florida Water Management District began activities to implement Initial Water Reservations. For ENP this flow estimate was deemed useful for Initial Water Reservations discussions as a starting estimate of restored conditions.

III. Methods

The manuscript is included with this task report as Appendix A. The details on the statistical methods used to estimate the paleo-based hydrologic conditions (flow and stage in the Everglades, and salinity in Florida Bay) are presented in

the manuscript and will not be repeated in this task report. However, the steps that were taken to complete the process are summarized as follows:

(1) Whipray Basin paleoecological information and observed and modeled salinity data were interpreted by the US Fish and Wildlife Service to produce an average salinity for each month of the year necessary to sustain the mollusks and other animals that were recovered in the Whipray Basin sediment core and identified in the laboratory.
(2) A daily time series of Whipray Basin salinity was developed by the PI from the USFWS average monthly paleosalinity regime for the 36-year period 1965-2000 using the Whipray Basin MLR model (developed as a previous CESI task) with the Natural System Model output for the appropriate stage stations (adjusted to remove the bias). This produced the Whipray Basin paleo-based salinity regime.
(3) New linear regression models were developed that simulated Whipray Basin salinity from the primary stage stations (CP, P33, and EVER7).
(4) These models were then used in an inverse calibration mode to estimate the daily stage that would have produced the Whipray Basin daily paleo-based salinity regime.
(5) New linear regression models were developed that simulated the flow needed across the Tamiami Trail (S12T+s333-s334), and in Taylor Slough at the Taylor Slough Bridge to produce the P33, CP and EVER7 stage conditions estimated from the Whipray Basin paleo-based salinity regime.
(6) New linear regression models were developed to relate the stage at CP, P33, and EVER7 to the stage at other locations in the Everglades. These stage levels were used to estimate the annual hydroperiod at a variety of locations in the Everglades, which were then compared to the performance measures of the Greater Everglades Sub-team for evaluation of CERP alternatives.
(7) New linear regression equations were developed to relate the salinity at Whipray Basin with the salinity at other monitoring stations in Florida Bay in order to simulate the paleo-based salinity regime to throughout Florida Bay.
(8) The results are a paleo-based flow regime to the Shark River Slough and Taylor Slough at critical locations, paleo-based stage and hydroperiod conditions that compared favorably with the pre-drainage hydroperiod conditions from freshwater paleoecological work in the Everglades freshwater marshes, and an estimate of paleo-based salinity conditions throughout Florida Bay.

The steps above were described in a manuscript that was submitted to the journal *Estuaries and Coasts* in the beginning of March 2007.

IV. Discussion

The manuscript was reviewed by two reviewers. Their comments are presented below. Both reviewer used a structured format and answered questions. The second reviewer also marked-up a copy of the manuscript manually. A copy of the comments is presented as Appendix B.

First Reviewer

The first reviewer considered the contribution new and original and considered it of broad interest. The conclusions were considered to be valid. It was thought to be a model for other studies because there are few studies that use a combination of paleoecological and modeling as an approach. It is noted that the title of the papers utilizes paleoecology but the emphasis of the paper was modeling. The abstract was considered to be informative. A reference to a lake paleoecological study was presented for consideration. The illustrations and plots were considered adequate and is the reviewer stated that the agreement between simulated values and observed data was “obviously good”.

The overall assessment was that the paper needs major revision, not in the work that was done, but in describing the paleoecological basis upon which the analysis was based. The paper was recommended for publishing only with major revisions.

Second Reviewer

The second reviewer’s comments carry a theme that emphasizes the lack of paleoecological background information in the manuscript. Therefore, each question is answered in this context and there is not as much constructive criticism available compared to the first reviewer. For example, the reviewer commented that the title, which includes ‘paleoecology’, is not valid since little paleoecological information is included. The second reviewer considers the paper poorly written with too many figures and abbreviations. The abstract is said to be too long. The reviewer believes that the work is worthwhile, but considerable work is needed to include the needed background.

The most constructive criticism is contained in the additional comments, as follows:

> “It needs much further work on the writing and format although I would encourage the author to pursue this avenue as I think there does exist in it a novel approach of the use of paleoecological data and the ability to link inferred values with modeled events.”

Additional comments from the second reviewer are presented in a marked-up copy of the paper.

The manuscript was not recommended by the Editor for publishing in *Estuaries and Coasts*. The primary reason for this recommendation was the lack of information on the paleoecological activities that lead to the development of the paleosalinity regime by others. The reviewer suggested that the scientists who completed the paleoecological work become involved with the manuscript and that a significantly revised manuscript will need to be submitted.

Since it was beyond the scope and budget of this task to involve others as co-authors, no further work was done. This task was originally seen as a task that would involve one author or two at most. The budget for this task was 290 man-hours. By the time that the manuscript was submitted the time committed to this task was well over double the budget. One reason for the overrun was the unforeseen effort to add the components of hydroperiod and salinity at other Florida Bay stations to the analysis, which were both worthwhile additions that broadened the analysis considerably, albeit at a cost.

This analysis has been presented to the following groups since it was completed:

- The Greater Everglades Ecosystem Restoration (GEER) conference in November 2006;
- Southern Estuaries Sub-team Evaluation Team (ET) and Assessment Team (AT), meeting jointly in January 2007; and
- The National Conference on Ecosystem Restoration (NCER) in Kansas City, Missouri in April 2007.

At each presentation the comments were favorable on the methodology used and the basic need for this analysis in the Everglades and Florida Bay restoration effort.

V. Findings and Recommendations

The manuscript was not accepted for publication. Both reviewers considered the lack of information on the paleoecological data as unacceptable, even though both reviewers commented favorably on the use of paleo-data and modeling as a novel approach worth pursuing further. This task was originally formulated as a task to be performed by one principal investigator, and there was no budget for work by or with others. However, the Principal Investigator has been coordinating with Dr. Lynn Wingard (USGS) and Patrick Pitts (USFWS) to revise the manuscript and ultimately re-submit it for publication.

At the time of publication of this final report, the revised manuscript has been through several re-writes, and has been formally reviewed by Tom Cronin and Debra Willard of the USGS Reston office. It is anticipated that it will be re-submitted for publication by the end of June, 2008. A decision has not been made as to re-submittal to *Coasts and Estuaries*, or perhaps to another journal.

VI. Appendices – Included as electronic files.

A. Manuscript
Filename: Marshall Wingard Pitts Manuscript.doc

B. Editors Comments
Filename: ms4369bhr.pdf

FINAL PROJECT REPORT

SALINITY SIMULATION MODELS FOR FLORIDA BAY AND THE SOUTHWEST GULF COAST, EVERGLADES NATIONAL PARK

Cooperative Agreement Number CA H5280-05-0006
Between
The United States Department of the Interior National Park Service
Everglades National Park
And
Cetacean Logic Foundation, Inc.

Frank E. Marshall, III
Principal Investigator

DeWitt Smith
Project Coordinator

Collaborators on Specific Tasks
William K. Nuttle (Task 7)
B. J. Cosby (Tasks 2 and 4)
Joan Browder (Task 5)

August 6, 2008

Cetacean Logic Foundation, Inc.
109 Esther Street
New Smyrna Beach, Florida 32169
(386) 423-4278

TABLE OF CONTENTS

Section	Page
Abstract	1
I. Introduction	3
II. Methods and Findings by Task	
Task 1 – Investigate the use of surrogates for evaporation to improve the simulative capabilities of MLR salinity models	5
Task 2 – Run the FATHOM MFL model with SICS/TIME input and compare to observed data.	7
Task 3 – Complete the development of MLR salinity models at MMN stations not previously modeled.	9
Task 4 – Assist ENP with implementation of the FATHOM model that was developed for the SFWMD MFL project.	12
Task 5 – Investigate the coupling of pink shrimp models with MLR salinity models, and investigate other ecological models that may benefit from the use of MLR salinity simulations.	13
Task 6 - Coordinate with the IMC to incorporate the new MLR models in the recently developed post-processing tool.	15
Task 7 - Update the 2002 report by The Cadmus Group by incorporating the improvements in statistical modeling accomplished by CESI activities.	16
Task 8 - Develop a methodology for using SARIMA models to fill data gaps in salinity, stage, water temperature, and freshwater flow to Florida Bay	18
Task 9 - Estimate the Flow Regime Required to Meet the Florida Bay and Florida Keys Feasibility Study Salinity Performance Measures.	20
Task 10 – Implement the ENP Version of the SICS/TIME Model for Use with the MLR and FATHOM Salinity Models.	22
Task 11 - Develop historical reconstructions for the longest period possible for the MLR salinity models in northeast Florida Bay.	23
Task 12 – Evaluate the feasibility of developing MLR salinity models for the mangrove lakes region north of Garfield Bight and Terrapin Bay.	26

Task 13 – Prepare and submit a manuscript for publication in a refereed journal. 28

Task 14 – Prepare draft and final reports. 29

Task 15 - Attend and Conduct Meetings 29

III. Discussion of Overall Project Findings 30

IV. Recommendations 38

V. References 40

CESI Project No. CA H5284-05-0006 for Everglades National Park

FINAL PROJECT REPORT

SALINITY SIMULATION MODELS FOR FLORIDA BAY AND THE SOUTHWEST GULF COAST, EVERGLADES NATIONAL PARK

Abstract

A Critical Ecosystem Studies Initiative (CESI) project, Cooperative Agreement Number CA H5280-05-0006 was completed as a continuation project by Cetacean Logic Foundation, Inc. (CLF) for Everglades National Park (ENP). The goal of this project was the characterization and simulation of the salinity regime in Florida Bay and the estuaries along southwest coast of the Gulf of Mexico (Gulf) within the Park. This research provided important information regarding the link between the downstream salinity in the estuaries of ENP and the upstream freshwater hydrology of the Everglades. In all fifteen (15) tasks were completed.

Evaporation surrogates were investigated to determine if inclusion of evaporation would improve the simulative capabilities of multivariate linear regression (MLR) salinity models. Correlation analysis showed that evaporation surrogates were correlated with other independent variables, which likely means that the parameters that were selected by the stepwise process were already expressing the effects of evaporation. It was concluded that there were no substantial reasons to include evaporation in the MLR salinity models. Based on this, the existing MLR models remained unchanged and the new models will not include evaporation as an independent variable.

The FATHOM mass-balance model was upgraded to include additional basins so that the total number of basins or “cells” is now 58. The new basins are intended to improve the initial mixing of freshwater inflows, and improve the fit of the model compared to the 47-basin configuration. While the 58 basin configuration is an improvement, the geometry of the basin domain must be fully evaluated in the context of the other model input parameters, in particular the spatial distribution of freshwater flows to the central part of the bay.

An additional fifteen MLR salinity models were developed, which means that there are daily MLR salinity models throughout Florida Bay and the upstream mangrove areas, in the Whitewater Bay area and the upstream mangrove areas, in the Shark River discharge estuary, along the upper Gulf coast in the Ten Thousand Islands area, and in Barnes Sound / Manatee Bay. Inferences can now be made about the models based on the model development process and the resulting MLR salinity models.

Historical reconstructions were developed for stage data using regression models to extend the observed data encompass the full period of the evaluations being made for CERP, 1965 – 2000. In this manner, historical reconstructions of salinity are possible through the MLR salinity models. Additionally, the Interagency Modeling Center at the South Florida Water Management District coded the models into programs that can be readily accessed for alternative analysis for Comprehensive Everglades Restoration Plan activities.

MLR salinity models were used with historical reconstructions and a paleo-based salinity time series data as input to a pink shrimp growth dynamics model. Whipray Basin and Johnson Key historical salinity reconstructions were utilized to help calibrate the pink shrimp simulation model using existing data. The MLR salinity models were coupled with the pink shrimp model to simulate growth and survival as the response of pink shrimp (*Farfantepenaeus duorarum*) to changing salinity and temperature. This proof-of-concept task showed that the two models could be coupled.

When all of the current salinity models in use for south Florida estuaries were evaluated (not just MLR salinity models), it was found that considerable recent progress had been made in the development and refinement of salinity models. From the analysis performed it was seen that the most complex models are the least portable and are rated lowest for ease of use. With respect to validity, all models rated high because most models are well-documented. For model fidelity, daily MLR salinity models have the best performing error measures, followed by SICS / TIME, and monthly FATHOM MFL base case models. Statistical and mass balance models will likely remain in use for planning-level decisions on a regional basis. Where possible, it appears that it will be less-expensive and time-consuming to utilize both statistical and mass balance models together as multiple lines of evidence and corroboration compared to utilizing only one hydrodynamic model for regional evaluations.

A flow estimate procedure was developed for the purpose of estimating the flow needed to meet the performance measures of the Florida Bay and Florida Keys Feasibility Study. The procedures established a quantifiable link between salinity in the downstream waters and the upstream hydrology (stage and flow), a link that had previously not been quantified. This flow procedure was used outside of this project to estimate the pre-drainage hydrology in the Everglades and salinity in Florida Bay using paleoecological information from sediment cores retrieved by the USGS from Florida Bay.

A preliminary study of the mangrove lakes area within the ENP found that this region has highly variable salinity that may be important in providing information on restoration alternatives. Preliminary models suggest that this region of the Florida Bay transition zone behaves differently with respect to salinity than other areas of Florida Bay and should be investigated further.

I. Introduction

This Project Report describes all of the activities that were completed for the current Critical Ecosystem Studies Initiative (CESI) project, Cooperative Agreement Number CA H5280-05-0006. This is a CESI continuation project that was originally initiated by Cetacean Logic Foundation, Inc. (CLF) for Everglades National Park (ENP) in 2002. The goal of this on-going CLF research is the characterization and simulation of the salinity regime in Florida Bay and the estuaries along southwest coast of the Gulf of Mexico (Gulf) within the Park. It is intended that this research yield information regarding the link between the downstream salinity in the estuaries of ENP and the upstream freshwater hydrology of the Everglades.

There were 15 tasks in the latest contract modification, as follows:

Task 1 – Investigate the use of surrogates for evaporation to improve the simulative capabilities of MLR salinity models.

Task 2 – Run the FATHOM MFL model with SICS/TIME input and compare to observed data.

Task 3 – Complete the development of MLR salinity models at MMN stations not previously modeled.

Task 4 – Assist ENP with implementation of the FATHOM model that was developed for the SFWMD MFL project.

Task 5 – Investigate the coupling of pink shrimp models with MLR salinity models, and investigate other ecological models that may benefit from the use of MLR salinity simulations.

Task 6 - Coordinate with the IMC to incorporate the new MLR models in the recently developed post-processing tool.

Task 7 - Update the 2002 report by The Cadmus Group by incorporating the improvements in statistical modeling accomplished by CESI activities.

Task 8 - Develop a methodology for using SARIMA models to fill data gaps in salinity, stage, water temperature, and freshwater flow to Florida Bay

Task 9 - Estimate the Flow Regime Required to Meet the Florida Bay and Florida Keys Feasibility Study Salinity Performance Measures.

Task 10 – Implement the ENP Version of the SICS/TIME Model for Use with the MLR and FATHOM Salinity Models.

Task 11 - Develop historical reconstructions for the longest period possible for the MLR salinity models in northeast Florida Bay.

Task 12 – Evaluate the feasibility of developing MLR salinity models for the mangrove lakes region north of Garfield Bight and Terrapin Bay.

Task 13 – Prepare and submit a manuscript for publication in a refereed journal.

Task 14 – Prepare draft and final reports.

Task 15 - Attend and Conduct Meetings

All of these tasks were the subject of individual Task Reports that describes in detail the activities that were involved with a particular task. Those task reports are incorporated herein by reference (see Reference section). This Final Project Report summarizes the tasks that were performed and evaluates the project as a whole, as required by the contract.

II. Methods and Findings by Task

Task 1 – Investigate the use of surrogates for evaporation to improve the simulative capabilities of MLR salinity models (Marshall, 2008a)

Evaporation is thought by most scientists and hydrologists to have a significant effect on salinity variation in Florida Bay, and this effect is thought to have a greater impact in the open water areas of Florida Bay and at certain times, particularly during the dry season and at the beginning of the wet season. Reliable, continuous data on Florida Bay evaporation do not exist for the period of MLR salinity model calibration and verification. Additionally, historical direct measurements of daily evaporation variability over the period for which SFWMD 2X2 model simulations are available do not exist. Therefore, this task focused on the evaluation of the correlative ability of evaporation surrogates that are known to be available over the 31- and/or 36-year period of the assessments that are being made.

Quantities that were evaluated included air temperature at Key West and Miami, relative humidity, cloud cover, sea water temperature, and the difference between air and sea water temperature. A time series correlation analysis was used to evaluate surrogates for evaporation. The significance level for inclusion in the MLR salinity models is 0.999, which means that an evaporation surrogate must be highly correlated to salinity to be included in the models. In a similar manner, local rainfall measured at the MMN stations was investigated as a potential candidate independent variable.

One of the primary objectives of this task was to re-visit consideration of evaporation as a candidate independent variable for MLR salinity models. In previous model development activities, the focus has been to find enough existing observed data to develop sound MLR salinity models at a high level of significance for long-term salinity simulation use. There are no long-term data available in the inventory of existing evaporation data for the south Florida region, though there are synoptic studies. Relative to the amount of observed data and the length of record for Everglades water levels, Key West sea surface water level, and wind (at Miami and Key West), the observed evaporation data are practically non-existent, particularly in Florida Bay and the mangrove fringe of he southwest Gulf coast. Therefore, evaporation was not included amongst the candidate independent variables during the first MLR salinity model development tasks. However, it is well-established through water budget estimates and observations of hypersaline conditions that evaporation plays an important role in determining the salinity regime in parts of Florida Bay and in areas along the southwest Gulf coast.

Various methods for estimating evaporation were examined. The SFWMD (Abtew et al, 2003; Irizarry-Ortiz, 2003) and others (Price et al, 2005) have found success using a self-calibrating method of estimating solar radiation, the key

component of evaporation, and using it with maximum and minimum air temperature values to estimate the potential evapotranspiration, which can, in turn, be used as an evaporation surrogate for MLR salinity model tasks. A preliminary investigation using salinity data from a limited number of stations showed that there was some correlation between salinity and Joe Bay potential evapotranspiration computed using the “Simple” method.

Therefore, SAS© code was written to compute an estimate of potential evapotranspiration using the Simple method at each of the 19 stations that currently have MLR salinity models. Solar radiation was estimated using the latitude of the station and the time of the year. Temperature difference was based on the maximum and minimum daily air temperature value measured at the Key West weather station, which has a record of daily air temperature much longer than the 1965-2000 period needed for CERP evaluations. Historical reconstructions for the 1965-2000 period using the Simple method simulations indicate that during the period 1985 – 1992, a drought period, etp was higher than the previous 19 years. During this period salinity values, in general, over the Bay were also high. However, the high salinity period of the early 1970’s is not expressed distinctly in the etp values. This was a period of low flow and the average evaporation contributed to the hypersaline conditions during that period.

A salinity model development activity was then undertaken in which the estimated potential evapotranspiration (etp) was submitted to the step-wise model development process along with the parameters that had previously been determined to be the best independent variables for salinity simulations at each station. In all but 3 of the 19 cases, etp was shown to be a significant independent variable, sometimes replacing other independent variables and sometimes becoming an added independent variable. However, in all case where etp replaced an existing independent variable, the independent variable that was replaced only explained a small portion of salinity variability (usually less than 1%), usually being a lagged value, and the etp replacement also only explained a small portion of salinity variability. For some models the goodness-of-fit improved with the inclusion of the etp, and for others the replacement by etp degraded the adjusted-R^2 and root mean square error values. The greatest model improvement occurred at Butternut Key, where the new model with etp explained an additional 5% of the variation in salinity, with about 3% explanatory improvement at Bob Allen Key and Manatee Bay, based on adjusted-R^2 values. When etp was an added independent variable, such as at Terrapin Bay, model adjusted-R^2 and RMS error values were reduced because the model now included additional terms that did not explain enough additional salinity variability to overcome the penalty of additional terms added to the regression model.

Additionally, the model development results using this evaporation surrogate were somewhat inconclusive at the open water stations where it was expected that evaporation would improve the models even if no improvement was seen at the near shore stations. Although model improvement was seen at Butternut Key

and Bob Allen Key with very slight improvement at Whipray Basin, etp was not a significant independent variable at Duck Key. Additional investigation with other open water salinity models (to be developed in the future) may prove that evaporation should be one of the independent variables in open water MLR salinity models. However, a comparison of the simulated values for salinity at Butternut Key indicates that the model with etp does no better at estimating high salinity values than the MLR salinity model that does not include etp.

Based on these results, it does not appear that including evaporation computed from the "Simple" method in the MLR salinity models is useful and that the models that have already been derived and are currently in use produce salinity simulations that are very similar to the salinity simulations produced by models that include evaporation surrogates.

With respect to rainfall directly on the Bay as compared to rainfall measured at a land-based meteorological station that may be remote to the Bay, previous attempts to correlate daily rainfall with salinity have not produced satisfactory models. For one reason, at the daily level, rainfall is an almost all-or-none discontinuous quantity, not a continuous independent variable like Everglades stage, sea surface elevation, and wind. In some manner, Everglades stage and sea surface elevation indirectly expresses the effect of rainfall accumulation. Because it is thought that rainfall on Florida Bay is spatially variable, to utilize direct rainfall as a candidate independent variable requires at least regional estimates of rainfall over a long-term period, or indirect data and a transfer function to relate rainfall at other stations to rainfall at a location on the Bay. The long term rainfall data that are available at Royal Palm, Flamingo, and Tavernier were used to produce a multivariate linear regression for rainfall measured at Joe Bay weather station. Since the fit of the model at the daily level was poor, it was not considered acceptable for use with MLR salinity models as a rainfall surrogate. However, when the daily values are aggregated into monthly totals, a much better and acceptable multivariate linear regression model can be produced (adjusted-R^2 = 0.68). Rainfall is spatially discrete at the daily time step but becomes somewhat continuous at the monthly step.

Therefore, even if daily rainfall at the Joe Bay weather station turned out to be a significant independent variable, it is not possible to extend the period of data by using the daily rainfall at other long-term weather stations because a reasonable transfer model can not be developed. Previous attempts to use daily rainfall to model salinity have also not been successful, though there is potential at the monthly resolution.

<u>Task 2 – Run the FATHOM MFL model with SICS/TIME input and compare to observed data</u> (Marshall, 2008b)

Task 2 was initiated when the MFL FATHOM model (47-basin configuration) was used with runoff input from TIME output. The TIME output that was used was the

FTLOADDS version 3.1 code and input files for run 157 were provided to ENP by USGS TIME developers. The input data to FATHOM was reviewed by Dr. Jack Cosby to ensure that the input parameters were being specified properly when TIME output is coupled to FATHOM. Runoff estimates were provided to the project team by ENP to produce salinity simulations throughout Florida Bay. The FATHOM salinity output using the TIME 157 runoff was compared to the salinity output from FATHOM generated by the MFL Base Case run. The FATHOM configuration used for this comparison was the original 47-basin model domain.

The TIME-based salinity simulations were compared to SERC and MMN observed data at a monthly resolution. When the TIME-based salinity simulations were compared to the observed data (both datasets) a number of discrepancies were noted. A comparison of the inflows showed that the TIME runoff estimates are lower than the observed values. Therefore, though this task was successful in that the output from the TIME model was coupled with the 47-basin configuration of FATHOM to produce reasonable salinity estimates, it was not successful in producing reasonable simulations from TIME output.

As an additional task, an updated bathymetry for a new 58-basin configuration of FATHOM was delivered by ENP to Dr. Cosby in early June 2007. The new bathymetry was processed to prepare the input files for FATHOM. The new bathymetry files were compared to the original bathymetry files in a detailed quality assurance check to assure that bathymetry of the unaltered basins was not affected. The new bathymetry files were then used in a simulation of the MFL base case for comparison with simulation of the same inputs using the 47 basin version of FATHOM.

The bathymetry for the new 58 basin configuration of FATHOM was installed using a hybrid approach (as in the SFWMD MFL study) and benchmark tests for the installation were completed. The benchmarks consist of a series of simulations for which the model has been run in the old 47 basin configuration and for which model outputs and diagnostics are known. When applied to the new basin configuration, the benchmarks assure that the installation of the additional basins did not inadvertently alter other aspects of the model structure.

The overall model performance in simulating salinity with 58 basins is degraded compared to the 47 basin simulation. In general, the model efficiencies are smaller for the SERC sites in the northeast bay (but are essentially unchanged in the rest of the bay).

This negative result is not unexpected. In changing one set of inputs, it is reasonable to expect that other inputs might need to be adjusted in turn. The SFWMD MFL run (to which these new simulations were compared) was the “best” of a number of runs performed as part of a sensitivity analysis in that project. Changing the basin configuration has shifted the relative sensitivity of other inputs to the model. To be rigorous, it will be necessary to repeat the sensitivity analysis to examine the new pattern of effects of all inputs. A recent

paper published by Nuttle et al (2007) suggests that evaporation should be increased by about 20% over that used in the SFWMD Florida Bay MFL FATHOM modeling activity.

The new basin configuration runs successfully and corrects (at least partially) the conceptual problem of the location of the flow inputs from Trout Creek. In general, the model efficiencies are smaller for the SERC sites in the northeast bay, but are essentially unchanged in the rest of the bay. To analyze this result, it will be necessary to repeat the sensitivity analysis to examine the new pattern of effects of all inputs. However, modification of the model in this manner is beyond the scope of this task. Another CESI project is currently being initiated to begin this process.

Task 3 – Complete the development of MLR salinity models at MMN stations not previously modeled (Marshall, 2008c)

The year one CESI project focused primarily on the development of models for Joe Bay, Little Madeira Bay, Terrapin Bay, Garfield Bight, North River, and Whipray Basin. The IOP exercise updated some of those models and added a new model for Long Sound, Duck Key, and Butternut Key. The second year CESI project added new MLR salinity models for Taylor River, Highway Creek, Little Blackwater Sound, and Bob Allen Key. Work for the Southern Estuaries Sub-team developed new models for Whitewater Bay East, Clearwater Pass, Shark River, and Gunboat Island.

For this task, new MLR salinity models were developed using the longest period of data available at the remaining physical monitoring stations in the ENP MMN. Most of these stations are within the influence of Shark River Slough or relatively more distant from the Everglades in the open water areas of Florida Bay. The MMN stations for new models are Broad River, Buoy Key, Broad River Lower, Cane Patch, Harney River, Johnson Key, Lane River, Lostmans River, Little Rabbit Key, Murray Key, Peterson Key, Tarpon Bay East, and Willy Willy.

The models presented herein for the ENP MMN stations, when added to the models previously prepared (Marshall, et al 2003a; 2004 and Marshall, 2003b; 2005), complete salinity model development for the MMN and provide a method of estimating salinity throughout all of the estuarine areas of Everglades National Park. While there are other salinity models available for salinity estimates in Florida Bay (FATHOM, EFDC, Nuttle's Four-box Model, SICS/TIME) the MLR salinity models developed by this study and the previous studies are the only salinity models available for Whitewater Bay, Shark River estuary, and upper west Gulf coast areas that are capable of producing a 36-year time series simulation. All of the MLR salinity models include the link to the upstream, freshwater hydrology of the Everglades through the stage independent variable, and include the important factors of sea surface level and wind variation.

The initial MLR salinity models that were previously developed were in water body areas of specific interest. When the first models were developed they were shown to be capable of reasonably estimating daily salinity such that alternatives for CERP projects could be evaluated through the use of the SFWMD 2X2 model. However, it was noticed in the early modeling that the models are sensitive to both local and regional conditions. The models presented herein provide additional evidence that the parameter selection process is capable of selecting independent variables that represent the relative effects of both local and regional conditions.

For the stations that are located in Whitewater Bay, the Shark River Slough estuary, and along the northernmost Gulf coast in ENP, the use of the daily average sea surface elevation from the Naples, Florida station (ID No. 8725110) instead of the data from the Key West station (ID No. 8724580) improved model fit by about 3-5%. For the Florida Bay stations along the western boundary with the Gulf, both Key West and Naples data were significant. However, the Naples data are predicted values prior to 1995, so Key West water level was used for model development purposes as in the past.

Additional details on model development can be found in Marshall, et al (2003a; 2004) and Marshall (2003b; 2005; 2008c). The daily salinity models that were developed for this task and the associated adjusted-R^2 value are shown below by sub-region. All independent variables in these models are significant at the 99.9% level, except for Cannon Bay (95% significance level).

Central Florida Bay

Buoy Key = 24.82836 – (1.13942 cplag4) + (0.23472 terbay) + (0.14891 terbaylag4), adj-R^2 = 0.79

Western Florida Bay

Murray Key = 50.44369 – (3.15719 p33) + (0.14449 terbaylag4), adj-R^2 = 0.51

Johnson Key = 53.13962 – (3.53830 p33) + (0.09932 ltmadlag4) + (0.07948terbaylag4) + (0.04843 vwndmia) + (0.55111 kwwatlev), adj-R^2 = 0.55

Little Rabbit Key = 51.31785 – (3.18926 p33lag1) + (0.19900 ltmadlag4) + (0.59864 kwwatlev), adj-R^2 = 0.46

Peterson Key = 39.76767 – (1.57570 p33lag4) + (0.25200 ltmadlag1) - (0.04663 uwndkwlag1) - (0.06325 vwndkw) + (0.08712 vwndmia), adj-R^2 = 0.56

Whitewater Bay

Lane River = 56.42862 – (1.46788 p35lag2) – (2.82488 p33)
- (4.71551 np206lag4) – (0.07508 uwndkwlag1) – (0.13990 vwndkwlag1)
- 0.05170 vwndmialag3), adj-R^2 = 0.77

Shark River Slough Estuary

Cane Patch = 7.86280 – (1.20972 np206lag4) + (0.23731 vwndmia),
adj-R^2 = 0.42

Tarpon East = 22.18537 – (1.02605 p33lag4) – (2.23641 np206lag4),
adj-R^2 = 0.57

Harney River = 51.09667 – (2.84042 p35lag4) – (6.12109 np206lag2)
– (0.22464 vwndkw) - (0.11811 vwndkwlag1) – (0.17356 uwndmia)
+ (2.45632 kwwatlevlag1), adj-R^2 = 0.72

Broad River = 20.82753 – (3.11717 np206lag4) - (0.05780 vwndkwlag1),
adj-R^2 = 0.53

Broad River Lower = 73.21431 – (4.82197 p35lag3) – (4.03692 p33)
- (4.15505 np206lag2) - (0.31452 vwndkw) – (0.22519 uwndmia)
+ (0.28467 uwndmialag2) + (0.10788 vwndmialag3)
+ (4.41540 kwwatlevlag2), adj-R^2 = 0.74

Upper West Coast

Willy Willy = 19.10483 – (2.23072 pa8lag2) – (2.23460 pa11lag2), adj-R^2 = 0.71

Lostmans = 43.18695 – (6.08870 pa8lag2) – (3.11218 pa11lag2)
– (0.14135 vwndkw) + (3.25077 kwwatlevlag1), adj-R^2 = 0.76

Cannon Bay = 38.12839 – (5.57603 pa8lag2) – (1.88696 pa11lag2),
adj-R^2 = 0.68

Watson Place = 41.54103 – (9.59506 pa8lag1) – (0.10461 uwndkw)
– (0.23667 vwndmia) adj-R^2 = 0.86 (Provisional model, limited data),

Task 4 – Assist ENP with implementation of the FATHOM model that was developed for the SFWMD MFL project (Marshall, 2008d)

The objective of this task was for the FATHOM consultant (Dr. B. J. Cosby) to meet with ENP to review and discuss the current use of FATHOM and to review up to five (5) FATHOM input files provided by ENP. The FATHOM consultant interpreted the output and provided the findings of the analysis in a report. Listed below are the notes from those meetings and the results of the discussions.

Sept 2006 Meetings.

- Met with ENP staff at SFNRC office in Homestead.
- Delivered FATHOM model and all MFL base simulation inputs.
- Conducted a 2-day workshop to instruct ENP staff in the implementation of the FATHOM model.
- Instructed ENP staff in the preparation of inputs for the FATHOM model so that the output of the SICS/TIME model could be used as an input to FATHOM.
- Instructed ENP staff in the post-processing of FATHOM output, particularly the comparison of simulated salinity to observed salinity.
- Collected MMN data from ENP staff to prepare spreadsheet for comparison of FATHOM salinity to MMN observations. (meeting also related to Task 2)

September 2006 through July 2007 Activities.

- Participated in phone calls and exchanges of model inputs and outputs through ftp sites to assist ENP staff in the implementation FATHOM with SICS/TIME outputs.
- Examined and debugged FATHOM outputs produced by ENP staff. Advised ENP staff on procedures for adapting the SICS/TIME output to FATHOM.
- Posted a spreadsheet to the ENP ftp site for comparison of FATHOM output to MMN observed data.

July 2007 Meetings.

- Met with ENP staff in Homestead to review the data and programs developed and exchanged over the preceding 10 months.
- Discuss the new 58 basin configuration of FATHOM. (meeting also related to Task 2)

FATHOM MFL model output was compared to MMN observed data as part of the exercise of assisting ENP with full implementation of the FATHOM MFL model. Simulated salinities from those runs were compared to observed salinity for the period 1991-2002 in the following plots. Two observed salinity datasets are used, the SERC dataset and the MMN dataset. Coordination and assistance in use of the post-processing routines from the software to produce the comparisons of simulated salinity to MMN data is one of the objectives of this task.

Comparisons are in the form of selected statistics of goodness-of-fit for all FATHOM basins for which observed data are available, and in the form of time series plots of simulated and observed salinities for selected basins. Full details of these statistics and plots for all pertinent basins are included in the spreadsheets delivered to ENP.

In general, there are few substantive differences when FATHOM MFL base case output is compared to either SERC or MMN data. The spatial coverage of the SERC data is broader than the spatial coverage of the MMN data set. The temporal coverage of the MMN data is broader then the temporal coverage of the SERC data because the MMN data are gathered more frequently than the SERC data, and the period of record is longer at most MMN stations.

From the statistics that were prepared it is noted that the values for efficiency for SERC and MMN data are similar but values of R^2 for SERC comparisons during the dry season are lower than MMN data. When the basins are grouped, the R^2 values for the MFL base case compared to the MMN data are lower than the comparisons to the SERC data. Exceptions to this are grouped basins C and D during the wet season.

The most significant accomplishment of this task was to assist ENP with implementing the FATHOM MFL software at South Florida Natural Resources Center. New spreadsheets were developed to analyze FATHOM results relative to ENP MMN data as well as SERC data.

<u>Task 5 – Investigate the coupling of pink shrimp models with MLR salinity models, and investigate other ecological models that may benefit from the use of MLR salinity simulations</u> (Marshall, 2008e)

A simulation model of growth and survival was developed by Browder et al. (2002) to predict the response of pink shrimp (*Farfantepenaeus duorarum*) to changing salinity and temperature. This model has been applied in Florida Bay and other south Florida estuaries (Browder et al. 2005). Response variables include: average growth rate, average survival rate, and potential harvest (which integrates both growth and survival). In addition, the model provides estimates of the number of individuals at 30-day intervals, relating directly to density, once scaled in the calibration.

The pink shrimp stimulation models for Whipray Basin and Johnson Key was calibrated using part of the data from Robblee (unpublished) 18-station data set for juvenile pink shrimp in Florida Bay and observed and simulated salinity data for western Florida Bay and north-central Florida Bay. This allowed comparison of observed pink shrimp density and density-related output from the pink shrimp model. The model was used with the predicted paleosalinity regimes for Whipray Basin and Johnson Key to predict pre-drainage shrimp densities for Johnson Key and Whipray basins.

The data requirements for the shrimp model and the mechanism for using the various salinity runs were determined through discussions with pink shrimp modelers. It was arranged that the PI would provide salinity time series data (daily) for four scenarios. The Whipray Basin simulations had previously been generated. However, the Johnson Key models were not developed until recently and processing was necessary to prepare the Johnson Key models for use.

The following data were sent to Joan Browder:

1. Whipray Basin observed salinity, April 6, 1989 - Dec 31, 2000.
2. Whipray Basin reconstructed salinity, January 1, 1965 – December 31, 2000
3. Whipray Basin paleo-based salinity, January 1, 1965 through December 31, 2000
4. Whipray Basin NSM 4.6.2-based salinity, January 1, 1965 – December 31, 2000
5. Johnson Key observed salinity, March 13, 1989 - Dec 31, 2000
6. Johnson Key reconstructed salinity, January 1, 1965 – December 31, 2000
7. Johnson Key paleosalinity salinity, January 1, 1965 through December 31, 2000
8. Johnson Key NSM 4.6.2 salinity, January 1, 1965 – December 31, 2000.

The data were sent by the Principal Investigator to Joan Browder (NOAA) and the pink shrimp model (Browder at al, 2002) was utilized for a number of preliminary runs, including the following:

1. Historical reconstruction salinity data (Marshall, 2007a),
2. Natural System Model-based (NSM 4.6.2) salinity from MLR model,
3. Paleo-based salinity from MLR model (Marshall, 2007b),
4. the average of reconstruction and NSM 4.6.2-based salinity (r+n/2), and
5. the average of NSM 4.6.2-based and paleo-based salinity (n+p/2).

These runs were intended to span the range of salinity values that were being examined at Whipray Basin and at Johnson Key.

Output produced by the pink shrimp model for Whipray Basin and Johnson Key include proportion of daily survival, average growth in mm/day, and potential harvest in tons kg/year. Although present and reconstructed salinity regimes in

the two basins differ, certain generalities apply to both basins regarding the simulated responses of pink shrimp to the various scenarios. Pink shrimp responses are lower (i.e., less favorable) under the paleo-based salinity scenario in the case of growth, survival, and potential harvests in Whipray Basin and in the case of survival and potential harvests in Johnson Key Basin. In contrast, growth rate was substantially higher for the paleo-based salinity regime compared to reconstitution salinity in Johnson Key Basin in these simulations. With the one exception, responses varied little among the reconstruction, NSM, and (r+n/2) scenarios. The (n+p/2) regime response was intermediate in responses between the paleo-based salinity regime and the other three scenarios.

Looking only at potential harvests which is an integrating parameter, in both Whipray Basin and Johnson Key the response to the paleo-based salinity scenario was higher than the response to the other scenarios in only two of the 36 years, 1971 and 1989, generally recognized as exceptionally “dry” years in the long-term south Florida rainfall records. The poorest responses were in the years 1966, 1968-1970, 1995-1997, and 1999, recognized as “wet” years in south Florida.

Differences in responses to the reconstruction, NSM, and (r+n)/2, were examined in more detail by calculating means and quartiles of the time series of responses. NSM had the highest mean growth rates and (r+n)/2 had the highest mean survival rates and potential harvests. The quartiles showed their most striking differences between paleo-based salinity regime and the other scenarios. Differences among scenario responses increased inversely with quartile, causing the minimum to show the largest differences among scenarios.

As another way of making comparisons, yearly differences between July-of-the-year responses were made. According to annual averages, both NSM and (r+n)/2 outperform the reconstruction by a small margin in simulated growth rate, survival rate, and potential harvests. Based on the number of years, both NSM and (r+n)/2 outperform the reconstruction in growth rate and potential harvests in both basins. In survival rate, (r+n)/2 outperforms the reconstruction in both basins, and the reconstruction outperforms NSM in both basins.

This task was a proof-of-concept exercise that focused on the coupling of statistical salinity model output with pink shrimp models. Therefore, the results consist of output from the shrimp model with limited interpretation.

<u>Task 6 - Coordinate with the IMC to incorporate the new MLR models in the recently developed post-processing tool</u> (Marshall, 2008f)

The objective of this task was to make the new models that were developed as a part of this CESI project (Task 3) available to the IMC for incorporation into the salinity simulation routine used by the Southern Estuaries Sub-team of RECOVER for evaluation of CERP alternatives. The new models that were

developed in Task 3 were transmitted to the IMC staff. The PI assisted the IMC with review of the final product. This task was a perfunctory task that was intended to help the agencies that rely on the MLR salinity models to analyze CERP alternatives as part of the performance measure evaluations. All models developed for Task No. 3 of this CESI project have been imported by the IMC.

Task 7 - Update the 2002 report by The Cadmus Group by incorporating the improvements in statistical modeling accomplished by CESI activities (Marshall, 2008g)

A report prepared for ENP by the Cadmus Group (Nuttle, 2002) described the current (at that time) status of model development for correlative and mechanistic models. Since then, improvements and updates have been incorporated into the modeling procedures, hence that status report on modeling tools was outdated. The PI coordinated with the author of the 2002 report (William K. Nuttle) to update this information, and compare MLR salinity models and mass balance models (ex. FATHOM). The 2002 report was updated at the beginning of this project to include recent advancements in salinity models. The report was updated again as one of the final tasks for inclusion in the final report to include additional advancements made during this project, and by others.

The two primary salinity data sets used for salinity modeling in Florida Bay are:

(1) the SERC/FIU long-term monthly grab sample data set, and
(2) the ENP marine monitoring data set with observations at 10 to 60 minute intervals.

Other hydrologic and climate data sets are used as they are needed for model development and for model input for simulations.

For this analysis, models were reviewed for which information was available on a broad basis in June 2006 for simulating and forecasting salinity in Florida Bay, Whitewater Bay, and the Gulf coast estuaries. For the salinity evaluations that have taken place thus far, there have been two general approaches to constructing such models. The first is empirical and relies on accurately describing observed salinity variations and correlative relationships. The second is mechanistic-based and relies on accurately accounting for the physical processes that drive changes in salinity. In both approaches the accuracy of the forecasts is limited by the data available to describe patterns of salinity variation and the driving processes.

Various statistical techniques can be employed in the empirical approach, the simplest being descriptive analysis. Both regression and time series modeling techniques have been applied to derive models for Florida Bay and Gulf coast salinity. Regression models exploit linear relationships in records of driving processes and systems response. Time series models utilize the serial

correlation that is present in many hydrologic parameters. The statistical models that have been developed thus far for Florida Bay and the Gulf coast estuaries are based on a coastal aquifer conceptual model and have been used successfully for evaluating water management alternatives and for performance measure development.

Mechanistic salinity models for south Florida estuaries include both mass-balance models and more complex hydrodynamic models. Mass balance models of salinity, in their discretized numerical form, are similar in form to autoregressive time series models. Mass balance models account for the inputs and outputs of water from basins delineated by geomorphologic features. Mass balance models have been used for ecological evaluations and for minimum flows and levels modeling.

Hydrodynamic models have been developed for both Everglades hydrology and the salinity in the downstream estuary. Hydrodynamic models are based on the solution of simultaneous differential equations of continuity and hydrodynamics (momentum) in one, two, or three dimensions, and can be used for both surface and groundwater applications. Hydrodynamic models have been used for modeling the freshwater portion of the Everglades / Florida Bay hydrologic system for about the past decade, and are in the process of being updated with better data and techniques. Only recently have hydrodynamic models been available that are capable of adequately simulating the salinity regime in south Florida Bay and the mangrove / salinity transition zone. Work is currently underway on the Florida Bay hydrodynamic models, while work on hydrodynamic models for the transition zone of the Gulf coast estuaries is still in preliminary stages.

The following salinity model evaluation factors from the Florida Bay Science Program (PMC, 2004) were utilized:

1. portability,
2. validity,
3. fidelity,
4. focus, and
5. ease of use.

Each of the models that were assigned a score according to how well they achieved the desired result of each modeling factor. The scale of scoring is from 1 = poor to 5 = excellent. For some models it was not possible to provide a score for a particular factor.

From this summary (Table 7-2) it can be seen that the most complex models are the least portable and are rated lowest for ease of use. With respect to validity, all models rated high because the models are well-documented. Models that simulated salinity were rated highest for focus because salinity performance

measures are the use for most of the salinity models. Finally, for model fidelity, daily MLR salinity models have the best performing error measures, followed by SICS / TIME, and monthly FATHOM MFL base case models.

Additionally, a comparison was made of observed salinity data and forecasts made by MLR and FATHOM salinity models by plotting the following data for Long Sound and Whipray Basin for the period April, 1994 through October 2002 (Figures 7-1 and 7-2):

1. MMN observations averaged to monthly,
2. SERC monthly grab sample observations,
3. FATHOM monthly average estimates from SFWMD MFL work, and
4. MLR daily estimates averaged to monthly values.

The MMN monthly average and SERC grab sample observations correspond well, with fewer deviations at Long Sound than at Whipray Basin. It is important to note that the sampling locations for these two programs in these water bodies are not the same. Both FATHOM and MLR salinity models simulate monthly average salinity in both basins well. The MLR models appear to perform slightly better for Long Sound compared to FATHOM, and noticeably better at Whipray Basin, though the difference in the simulations by the two modeling procedures is small. These plots indicate that both the MMN and SERC data sets can be used interchangeably at the monthly level. It also shows that both FATHOM and MLR salinity models are capable of providing reasonable estimates of salinity at these stations.

Task 8 - Develop a methodology for using SARIMA models to fill data gaps in salinity, stage, water temperature, and freshwater flow to Florida Bay (Marshall, 2008h)

Autoregressive, moving average (ARMA) models for stage, salinity, temperature, and flow were first evaluated for filling data gaps. ARIMA models were developed to take advantage of serial correlation, and particular the power of autocorrelation (correlation to the previous time-period value). The evaluation determined that the SAS© PROC ARIMA procedure is relatively inflexible meaning that utilization of multivariate methods provides a better overall process for filling data gaps compared to ARIMA models. After testing ARIMA models on stage and salinity and determining that the simulations were unreliable, further use of ARIMA models for temperature and flow was eliminated from the task.

Linear interpolation is another method of filling gaps that is common. Linear interpolation is generally considered to be acceptable if the gap is small (1-5 days), the day-to-day change is small, and there are no inflection point in the missing data. Larger gaps (in time) create a chance that an event of some kind may occur, which would change the rate of increase or decrease in a parameter (stage, salinity, temperature, flow) such that linear interpolation was no longer

accurate for filling gaps. Similar to ARMA modeling, no additional time was spent evaluating linear interpolation for filling gaps in hydrologic or physical data after evaluating stage and salinity.

Multivariate linear regression (MLR) models proved to be the most successful technique for simulating hydrologic and physical data when there are gaps in the observed data record. MLR models were constructed from same-type data that was available from neighboring stations. Determining the candidate stations that would work best as an independent variable in an MLR model was straight-forward, accomplished through the use of a correlation matrix and step-wise development of the regression model. For stage and flow, use of the intercept term produced the best model, while a model with no intercept was best for salinity and temperature. However, use of a no-intercept model is only appropriate under certain circumstances, when one is confident that the distribution of values for the observed data and the distribution of values for the gap-filling data are equivalent, and the means are equal, which is rarely if ever known.

In addition, stage, salinity, and flow benefited from the inclusion of wind in the model. However, the benefit of including wind terms in a temperature model was very small, perhaps because the level of correlation with neighbor stations was so high with temperature. In all cases where wind was a significant parameter in the MLR model, the wind parameters generally only explained on the order of 1-5% of the variability in the dependent variable.

The use of multiple independent variables generally increased the explanatory capability of the model compared to a univariate model by about 1-5% for stage, 3-5% for salinity, and 3-5% for flow. The R^2 value for univariate temperature models was over 0.86 so multivariate modeling was not needed. The R^2 value for MLR models ranged form a high of 0.99 for salinity (Long Sound) to 0.78 for stage (P33).

The conclusions of this effort to determine if time series models could be useful for filling gaps in time series data for stage, salinity, temperature, and flow are clear –ARIMA models are not very useful for this purpose. This in-depth analysis of ARIMA models brought to the surface the level to which this procedure relies on the availability of past values of the dependent variable, which are not available when the data gap exceeds one day for daily data. ARIMA models work reasonably well for gaps of 2-3 days particularly when the range of the data in the 2-3 day period is near the mean value for time series. However, the ability of ARIMA models to provide highly accurate one-step forward predictions using the serial correlation in the data means that ARIMA models should be considered for use as a predictive tool for water management system operation when the data are not missing.

Similarly, the obvious limitations of linear interpolation limit the use of this technique to small gaps. From the limited analyses done, it appears that gap-size limit for linear interpolation is about 5 days in duration for the 4 parameters of interest. Beyond a 5-day window, the chance that an event will occur that modifies the rate of change increases such that the use of linear interpolation may not be reliable.

Multivariate linear regression (MLR) models proved to be useful for filling data gaps in stage, salinity, and flow. Based on the limited sample of this evaluation MLR models appear to be able to explain about 85-95% of the variability in the stage, salinity, and flow data, and wind measured at Key West and Miami appears to be useful in modeling by explaining up to 5% additional variation over models that do not include wind. However, models that include wind tend to have a large number of independent variables even at the 0.999 level of significance, which may be the result of cross-correlation effects between independent variables.

In the case of temperature, the correlation of neighbor stations as independent variables to the dependent variables was so high that MLR models did not provide sufficient benefit to justify MLR use. Instead simple univariate linear regression models were more than adequate. The use of bottom measurements to model top measurements at the same station (and vice versa) provides an advantage over the use of neighbor data as an independent variable.

Based on this analysis the flexibility of MLR models was shown to be useful for filling data gaps. MLR models have the added advantage of being transparent and easy to understand by a wide audience. The use of neighbor independent variables and wind data in MLR models for filling data gaps is recommended when simple regression models are considered to be inadequate.

Task 9 - Estimate the Flow Regime Required to Meet the Florida Bay and Florida Keys Feasibility Study Salinity Performance Measures (Marshall, 2008i)

Linear regression techniques were used to develop models for flow at the S-12, S-18C, S-197, and S-175 structures as a function of observed stage at the statistically significant stations in the Everglades. These models are the same as the hydraulic rating curves used for weirs, gates and other flow control structures. These techniques were used to establish relationships between the primary and secondary stage (water level) variables in the various MLR models. Model goodness-of-fit statistics was computed and verification of all models was performed using observed data for comparison.

At least one salinity model was assigned to each FBFKFS zone in northeast Florida Bay, Zones 1, 2, 3, 5,and 14. For each of the zones, plots of the FBFKFS requirements were used with plots of the NSM salinity simulation made for the Southern Estuaries sub-team to establish a salinity regime at a particular

monitoring station that meets the performance measure requirements to the greatest extent possible. The NSM time series was adjusted by adding or subtracting an equal amount to each NSM value over the 36-year period 1965-2000. In this manner, the NSM simulation was only be used to simulate the day-to-day salinity variation, not the magnitude of the salinity estimate. A well-documented statistical procedure called inverse prediction (Neter et al 1990) was used with the salinity models to establish a stage regime at the primary regression model stations - Craighead Pond (CP), P33, and EVER7, using observed water level and the previously developed regression models for the secondary stations.

Then the flow regression models (hydraulic rating curves) developed previously were used to estimate the required flow regime at the S12, S18C, S197, and S-175 structures for the 36-year simulation period 1965 through 2000. This was done for each of the zones, Zones 1, 2, 3, 5, and 14.

This proof-of-concept exercise has shown that a suite of linear regression models and statistical techniques can be used to estimate the flow requirements at three locations in the Everglades hydrologic system for a particular salinity regime. In this case, the Florida Bay / Florida Keys Feasibility Study salinity targets were used with NSM 4.6.2 and MLR salinity models (Marshall 2004, 2005) to develop a FBFKFS daily salinity regime for eight locations in Florida Bay that were spread across several FBFKFS zones. Observed values of salinity and stage were regressed to produce salinity (independent variable) as a function of stage. Then the models were turned around to use stage as the independent variable. In this manner simulated time series were generate at CP and EVER7 from the FBFKFS salinity regimes. From CP and EVER7, flow at TSB and S18C are estimated using other regression models. An additional set of regression models produces stage at P33 from stage at CP, then P33 is used to estimate flow at S12T.

The daily simulated flow values can show large differences at times when compared to daily observed values, particularly at high flow values. However, the statistical power that comes from large data sets, such as the ones used in this study (about 2500 values used for model development and about 12,500 for FBFKFS runs) allows these simulations to be used best in a monthly average or daily average value mode. In this mode, with simulations made daily and then averaged to monthly mean or daily average value over the simulation period, the modeling procedure appears to work well. This means that the simulated daily values should be interpreted as long-term averages, over periods that have experienced a wide range of climatic conditions.

When the modeling procedure is used to simulate the flows required to meet various FBFKFS targets, the difference between the simulated FBFKFS target flows and observed flows is substantial. The TSB, S18C, and S12T observed average values were found to be about 50%, 80%, and 56%, respectively, of the

FBFKFS target flows. This reflects the findings of Smith, et al (1988) using fluorescent techniques with coral banding on a piece of coral from near Peterson Key. These authors found that flow to northeast Florida Bay was reduced about 56% from historical conditions.

Task 10 – Implement the ENP Version of the SICS/TIME Model for Use with the MLR and FATHOM Salinity Models (Marshall, 2008j)

This investigation makes use of two modeling systems (SICS and TIME) developed by the USGS for simulating the hydrology and hydraulics of the Everglades (Langevin et al 2004a). The output of these modeling systems includes water depth values that can be converted to stage and used with the MLR salinity models to simulate salinity in Florida Bay.

ENP obtained a copy of the SICS code in 2006 which was provided to the PI. The PI analyzed some examples of SICS output before the attention in the modeling effort turned to the further development and use of the TIME model. The PI was able to produce stage output from the SICS model using post-processing programs provided to ENP and the PI by the USGS. When the TIME model began to be used, the SICS post-processing tools were no longer applicable and the PI relied on ENP to provide post-processing files for input to the MLR salinity models.

A calibration / verification run from TIME (157 run) was made available to compare with observed data for stage. ENP provided stage output from the TIME 157 model run that has been corrected for model bias, relative to the NGVD29 datum. This adjusted TIME 157 run output (stage) was compared to the observed stage data, then used with MLR salinity models to produce salinity simulations, as present below.

The MLR salinity models used for simulating salinity using the adjusted TIME 157 run stage values for Joe Bay, Little Madeira Bay, and Terrapin Bay come from Marshall et al, 2004 and for Garfield Bight the models are from Marshall, 2005.

When the bias-adjusted TIME 157 stage data were used for the MLR salinity model input there are periods with missing data because the TIME data are post-processed for the surface water module, only (see discussion below). When the conditions become dry enough for the simulated water level to drop below the elevation of the monitoring station, the stage elevation is a groundwater output, which is not currently produced by the post-processing tools.

This task was the first task initiated for this part of the on-going CESI project, and the last to be finished. Mostly this was caused by the evolution of the SICS domain into the TIME domain, and the difficulty of the USGS in making the code revisions that were needed. However, it appears that the TIME model is now running dependably. Even so, stage output data are not provided by the post-processing tools when the simulated water levels dropped below the ground

surface elevation. The ground water stage output is produced by the groundwater model, SEAWAT while surface water stage is produced by the SWIFT2D model. Output is only available from the SWIFT2D model at the surface cells using the post-processing tools available at the time of the analysis. An integrated time series of stage is needed to be able to fully couple TIME and MLR salinity models.

When the TIME 157 stage data are compared to the observed stage the data are similar for both models during the wet season, but the dry season values show various levels of divergence. The greatest divergence of TIME 157 and observed dry season values is seen at NP62; the least divergence is seen at P33.

When both of these input data sets are used with wind and Key West water level data in the MLR salinity models, the comparison plots the salinity produced using the TIME 157 data and salinity produced using the observed data are similar. The exception may be Joe Bay, but there are many missing values for the dry season. EVER6 in the Joe Bay model input had many dry season missing values from the TIME output. All salinity simulations were affected by missing dry season data from TIME 157. Given the differences seen in the stage comparison plots, the similarity of the salinity simulations is somewhat surprising.

It is concluded that the TIME output may serve as input to MLR salinity models and the resulting salinity simulations are similar to salinity produced using the SFWMM output for similar modeling scenarios.

Task 11 - Develop historical reconstructions for the longest period possible for the MLR salinity models in northeast Florida Bay (Marshall, 2008k)

The SFWMD used the Taylor River MLR salinity model from a previous CESI project for their Florida Bay MFL work. As part of that work a historical salinity reconstruction was developed for the period 1970 through 2003. Salinity measurements have only been collected at the Taylor River site by ENP since 1988. The Taylor River MLR salinity model was used to extend the salinity time series back in time to 1970, which is when the C-111 Canal system came on line. To do so required using other MLR models to also extend certain input parameters.

Using the same methodology used by the PI at Taylor River for the SFWMD MFL activities, a historical salinity reconstruction was developed at MMN stations in northeast Florida Bay (Joe Bay, Little Madeira Bay, Terrapin Bay, Garfield Bight, Long Sound, Highway Creek, Little Blackwater Sound, Duck Key, Butternut Key, Whipray Basin, and Bob Allen Key). A spectral analysis was then performed using the historical reconstruction to determine if the dominate frequencies in the observed data also appeared in the modeled data.

The purpose of this task was to prepare time series of salinity that extended back in time until January 1, 1970 which is beyond the beginning of the period of record for the MMN stations. In some cases, for example Joe Bay, Little Madeira Bay, Long Sound, and Highway Creek, the period of record begins in 1988. By comparison, the Bob Allen Key period of record doesn't begin until 1997. The primary tool for filling in the gaps - internally in the continuously-monitored observed data but also in the period prior to the date of data collection - are multivariate linear regression models for salinity that have been used successfully to-date for CERP performance measure evaluations. For CERP use, the South Florida Water Management Model (2X2 Model) providing the stage inputs and observed data were utilized for wind and Key West water level input parameters.

The 2X2 Model data begin on January 1, 1965. The observed wind and Key West water level data used for the MLR salinity simulations begin on the same date. For this CESI task it was beyond the scope of work to obtain additional wind or Key West water level data prior to 1965. Therefore, the earliest point in time that the reconstructions can begin with this data set is January 1, 1965. For the Florida Bay Minimum Flows and Levels modeling (ECT, 2005) the historical reconstruction from 1965 to 1970 was not particularly useful since the construction of Tamiami Trail in the current configuration including the S-12 structures was completed in the 1960's and water levels for next decade reflect an area in transition. Initial operations of these facilities interrupted almost completely the flow of water into Shark River Slough. However, due to intervention by ENP, additional flows to the Park were augmented through operational activities beginning in about 1970. This limitation was also realized for the historical reconstructions using the FATHOM model in Florida Bay and for the MLR model for Taylor River for the Florida Bay Minimum Flows and Levels (ECT, 2005). For these reasons, the historical reconstructions in this task begin on January 1, 1970 and end on December 31, 2000.

For the historical reconstructions to be useful, there must be confidence that the simulations prior to the observed data represent a reasonable estimate of the salinity regime that occurred during the period of the reconstruction. Occurrence of extreme high and low salinity values can be used to test the responsiveness of the models. For four of the MMN stations (Joe Bay, Little Madeira Bay, Long Sound, and Highway Creek) there are good examples of both high and low extremes in the observed record. For example, during the drought year of 1989, salinity values were recoded over 50 psu (Joe Bay), 70 psu (Little Madeira Bay), and 40 psu (Long Sound and Highway Creek). As a comparison to a simulated situation, at Bob Allen Key monitoring began in September 1997, meaning that the salinity values of the Bob Allen Key reconstruction in 1989 were from the MLR model. The plots show that simulated salinity at Bob Allen Key exceeded 60 psu in 1989, and values above 60 psu in the Bob Allen Key reconstruction are only seen in the reconstruction 2 others years, 1971 and 1974. In fact, all of the reconstructions show an increase in salinity during this drought period of the

1970's, which means that the reconstructions are sensitive to dry period conditions.

At the other end of the salinity scale, lowest salinity values in the observed record of these 4 stations are seen in 1995 and the following years, though low salinity values at these stations during the observed record are much more common than high salinity values. At Bob Allen Key the simulated 1995 and 1996 salinity values are also low, as are the observed values in 1997 (when monitoring began) and in the ensuing years.

Comparisons of salinity reconstructions at the open-water stations with salinity at the near shore stations indicates that there are times when high salinity conditions develop at the near shore stations in the spring at the end of the dry season, but the salinity conditions at the open-water stations remain at average levels. However, for some years (1971, 1974, and 1989 through 1991) the highest salinity values in the reconstruction are seen at both near shore and open-water locations. It appears that the dry conditions must reach a certain level of severity (or perseverance) before the open-water stations begin to experience hypersaline conditions. In average years, the highest salinity value of the year is about 20 psu higher at Bob Allen Key compared to Long Sound. However for the years of the highest salinity values when hypersaline conditions were recorded or simulated at both locations, the difference in the highest salinity values of the year at these two stations is only about 10 psu.

Comparison of the spectral density plots for observed data and reconstructed data for the period prior to the observed period of record (POR) shows that the reconstructed time series have similar periodicities to the observed data, though some differences at some of the stations is noted as follows:

Because the number of values used to compute the spectral density function for the longer periods is limited to about 20 or less, the differences between the observed POR and the reconstruction (pre-) POR are not considered to be significant. The important information is the recurring seasonality in the observed data is being expressed in the simulated reconstructions.

In addition, it appears that the near shore stations are subject to relatively significant cyclic forcing that is not on the order of yearly, semi-annually, or quarterly intervals. The open-water stations do not show the same behavior in the spectral density function plots.

The near shore locations appear to be subject to hydrologic forcing that is more complicated than Little Madeira Bay, Whipray Basin, Duck Key, Butternut Key, and Bob Allen Key, as expressed in the number of secondary maxima seen in the spectral density function plots of the near shore stations at the more frequent periods (low values of period). In addition, at Whipray Basin, Duck Key, Butternut Key, and Bob Allen Key the value of the spectral density function does

not increase to a secondary maximum at higher periods, as is seen in the plots of the near shore stations, including Little Madeira Bay. In the case of Joe Bay, Terrapin Bay, Long Sound, Highway Creek, and Little Blackwater Sound, the value is still increasing significantly at 1000 days. For Little Madeira Bay and Garfield Bight the spectral density function value has leveled off at a relatively high value at 1000 days. At Whipray Basin, Duck Key, Butternut Key, and Bob Allen Key the spectral density function levels off at a lower value of period.

This analysis shows that the reconstructions comprised of observed data and simulations using MLR salinity models and other statistical relationships provide a reasonable estimate of the historical conditions since 1970. These historical reconstructions of salinity are intended for use with ecologic models such as available fish models (ENP and others), pink shrimp models, and seagrass models by other researchers.

Task 12 – Evaluate the feasibility of developing MLR salinity models for the mangrove lakes region north of Garfield Bight and Terrapin Bay (Marshall, 2008l)

Work is currently underway in the Florida Bay mangrove lakes region to study submerged aquatic vegetation as part of a CESI project. As part of this project physiochemical data are being acquired.

At the time of this evaluation, two reports had been filed to record project progress: Semi-annual Progress Report - October 18, 2006; and Annual Progress Report – March 23, 2007. Both of these reports were reviewed. Temperature and salinity data are presented for the period April/May 2006 – January/February 2007 for 8 stations in the region. These data were analyzed for use with the development of MLR salinity models to determine if there is the potential to link the salinity of this region to the hydrology that affects it, the freshwater marsh of the Everglades.

Equipment problems affected the collection of continuous salinity data at the beginning of the project. Continuous salinity data were made available from June 27, 2006 through August 28, 2007, approximately 13 months. To begin the analysis, simple statistics were computed for all three salinity modeling stations. Then the data from Seven Palms Lake, The Lungs, and West Lake were used to prepare correlation matrices using the salinity data, stage data from upstream locations in the Everglades freshwater marsh, wind vectors from the Long Key C-Man station, and water level elevation from the Long Key C-Man station. Using guidance in parameter selection from the information in the correlation matrix, preliminary MLR salinity models were developed using stepwise regression procedures to explore the potential for MLR salinity model development once additional data have been collected.

The highest Pearson correlation coefficient for correlation with salinity was P33 for all stations. However the correlation coefficient was positive for P33 and

West Lake, seemingly indicating that higher values of P33 stage were related to higher values of salinity in West Lake. Positive correlation coefficients are also seen for CP at both Seven Palms Lake and at West Lake. The correlation coefficients for CP at Seven Lakes Palms and West Lake show a medium level of correlation but they are opposite in sign. Comparison of the correlation matrices shows that the correlation coefficient values were similar to the unlagged values.

Preliminary models were developed for all three stations using the stepwise linear regression procedure. The significance level for including an independent variable in the model had to be set at 0.999 similar to other MLR salinity models to reduce the number of parameters in the model to a reasonable level. At the 0.950 significance level there were between 15 and 20 independent variables in the model. There is an additional benefit in setting the significance level at 0.999 – all independent variables in the final model are highly significant.

This was an initial exercise at preparing MLR models with this data. All parameters chosen by the SAS© stepwise procedure are reported. Though there is only limited data available at this time from this project, it appears that the salinity data being collected ultimately will be useful for salinity modeling. The mangrove and lakes area of the north central part of Florida Bay where the salinity measurements are being taken represents an area with few monitoring stations. If extended continuous monitoring shows the salinity range seen for the limited monitoring, then this is an important area for monitoring because (1) salinity is being measured across a large salinity range, (2) hypersalinity is being measured, and (3), salinity values are not dropping to 0 ppt. At these locations, restoration in the form of additional freshwater flows towards western and central Florida Bay will be most noticeable, as will lower flows (if that occurs) and increased salinity due to sea level rise. Since these stations do not go to zero they are useful for simulating changed conditions due to restoration or other alterations to freshwater flow and stage. If paleoecological investigations could be extended into this area, it may be useful for estimating the conditions before freshwater flow alterations around the turn of the 20th century.

The preliminary modeling activities suggest that the data from these three stations should be capable of producing reasonable MLR salinity models using stage, wind, and water level parameters. The models produced herein are preliminary models prepared for analysis only. These models illustrate important relationships and should NOT be used for predictive purposes without further analysis. However, this initial evaluation suggests that this further analysis will yield models with R^2 values in the 0.55 to 0.80 range.

Task 13 – Prepare and submit a manuscript for publication in a refereed journal (Marshall, 2008m)

Although the work performed over the past several years on developing the salinity models for the MLR network has produce several project reports and a number of presentations at conferences and seminars, a formal paper presenting the development work has not been submitted to a scientific journal and subjected to peer review. The purpose of this task is to produce a manuscript that is suitable for submittal and review by a refereed journal, such as Estuaries and Coasts, Water Resources Research, or ASCE Proceedings. The paper discusses the development of statistical models for the estimate of salinity, stage, and flow; and presents the models and findings for an estimate of flow needed for restoration purposes. Included in this task are site visits to each of the monitoring stations in the Everglades where water levels are being measured, and to the salinity monitoring stations that have not yet been visited.

The manuscript was reviewed by two reviewers. The most constructive criticism is contained in the additional comments of one of the reviewers, as follows:

"It needs much further work on the writing and format although I would encourage the author to pursue this avenue as I think there does exist in it a novel approach of the use of paleoecological data and the ability to link inferred values with modeled events."

On the first submittal, the manuscript was not recommended by the Editor for publishing in *Estuaries and Coasts*. The primary reason for this recommendation was the lack of information on the paleoecological activities that lead to the development of the paleosalinity regime by others. As a result of this Lynn Wingard of the USGS and Patrick Pitts of the USFWS became co-authors and the paper has been re-written.

At the time of publication of this final report, the revised manuscript has been through several re-writes, and has been formally reviewed by Tom Cronin and Debra Willard of the USGS Reston office. It was re-submitted for publication by *Estuaries and Coasts at* the end of June, 2008.

Task 14 – Prepare draft and final reports.

Draft task reports were prepared by CLF for review by ENP. Comments on the draft reports were received and discussed with the COTR. Revisions were then made to finalize the final task report. Each task report has been provided as an appendix.

The information in the task reports was summarized in this draft project report, and interpreted. Once comments have been received, a final project report will be submitted.

Task 15 - Attend and conduct meetings

The Principal Investigator scheduled, attended, and conducted on-site meetings on a roughly bi-monthly basis Off-month telephone discussions were held to discuss progress, needs, schedule and other project items. Other telephone and electronic communications took place as-needed between the scheduled meetings.

Nine project meetings were held at the ENP office in Homestead. Summary meeting notes for each meeting were prepared and are provided in an appendix.

Bi-annual and annual progress reports were prepared and presented that were comprised of excerpts from the task reports completed to-date, as well as information on tasks that were still underway and had not yet reached the point of preparing a task report. The final project report represents the final bi-annual and annual progress report.

III. Discussion of Overall Project Findings

Salinity is a fundamental and key characteristic of the physical conditions of estuarine and coastal ecosystems. Salinity affects water quality, the make-up and spatial distribution of vegetative communities, and the life history of most animal species in these ecosystems. Simulations and forecasts of salinity are an important tool in the assessment of ecological resources in the Everglades, Florida Bay, and the estuaries on the Gulf of Mexico (CROGEE 2002). Water managers use forecasts to evaluate the expected benefits and impacts of ecosystem restoration activities. Ecosystem restoration involves aspects of adaptive management (NRC 2004), uncertainty analysis (CERP 2002), and risk assessment (Thom et al. 2004), and these all rely on the application of predictive tools such as statistical models.

The Everglades / Florida Bay hydrologic system is unique because of the vast area of freshwater marshes underlain by porous substrate that stores runoff before it enters the estuarine zone, as well as the spatial extent of estuarine conditions in Florida Bay. Standard riverine hydraulic models can not account accurately for the spatial and temporal variation in stored water and dispersed flows in the Everglades. Instead freshwater hydrology and wetland basin models have been developed that simulate the south Florida hydrology required for use with statistical salinity models. For the MLR salinity models, the hydrology of the Everglades is described by the stage levels that are used as model input with wind and sea surface elevation conditions for salinity simulating and forecasting.

This project has been a continuation of statistical salinity model development and application. The findings of the tasks that were completed, summarized above, are discussed further below.

There were several project tasks related to the further refinement and upgrading of the MLR salinity model development methodology, and to the development of additional models at MMN stations that were not previously included in the analysis. These similar tasks included Task 1 (evaluation of evaporation surrogates), Task 3 (development of new models), Task 6 (IMC coordination for new models), and Task 12 (evaluation of data from the mangrove lakes).

The evaluation of evaporation surrogates was developed as a task to determine if there were indirect measures of evaporation that could be used to improve the MLR salinity models. Evaporation had not been included in the models because there are few measurements of observed evaporation, much less a time series of daily values for 31- or 36-years. The only method for computing evaporation that accommodated the need for this long-term simulation through the use of data that were available was the SFWMD Simple Method, which is used for the SFWMM. While the method produced what seemed to be reasonable estimates of daily evaporation, demonstration of coherence with observed data could not be shown. Even so, when evaporation estimates from the Simple Method were

used for model development, little improvement of the fit of the MLR salinity models to observed data was seen. Correlation analysis showed that the evaporation estimates were correlated with other independent variables, which likely means that the parameters that were selected by the stepwise process were already expressing the effects of evaporation. For example, both stage and sea surface elevation record the reduction of water level due to evaporation. It was concluded that there were no substantial reasons to force the inclusion of evaporation in the MLR salinity models if simulative ability of the model is not improve. Based on this, the existing MLR models remained unchanged and the new models described below will not include evaporation as an independent variable.

Building on the success of the MLR models in the previous CESI projects, an additional fifteen MLR salinity models were developed, thereby completing the development of MLR salinity models for the MMN stations. Because of the work completed on this project there are currently daily MLR salinity models for the following regions of south Florida:

- Florida Bay and the upstream mangrove areas,
- Whitewater Bay area,
- Shark River estuary,
- the upper Gulf coast in the Ten Thousand Islands area, and
- Barnes Sound / Manatee Bay.

Inferences can now be made about the models based on the model development process and the resulting MLR salinity models.

All of the MLR salinity models for the MMN stations along the western open-water boundary of Florida Bay included P33, either un-lagged or with a 4-day lag, as well as salinity in the near-shore embayments of Little Madeira Bay and Terrapin Bay. The importance of P33 provides evidence of the link between the Shark River discharge and the western boundary stations as has been theorized by researchers who have examined circulation patterns in this area. All of the Shark River estuary salinity models included NP206, lagged.

Previous salinity model development efforts in Florida Bay showed that P33 and Craighead Pond (CP) were the primary stage stations in explaining the salinity variation. The additional Shark River estuary models indicate that NP206 is the primary station for explaining salinity variation at the MMN stations in the Shark River estuary. It is important to note that the selection criteria for inclusion of a parameter in the model was very high, 0.999, meaning that the consistent selection of independent variables as the most significant in explaining salinity variability provides additional evidence that certain stations are primary.

Wind vectors and sea surface elevation were important at a secondary level in explaining salinity variability at almost all stations. Each of these patterns shows

that the MLR salinity models are reflecting the important driving forces for salinity in a way that seems to make physical hydrologic and hydraulic sense.

The error statistics generated for the new salinity models indicate that, in general, the R^2 value for the new models is between 0.50 and 0.75 and the root mean square (RMS) value for all models is between 2-3 psu, with higher values for some of the models that are driven by the stage in Big Cypress National Preserve. These models perform as well and better than other salinity models that have been previously developed from ENP Marine Monitoring Network salinity data.

The development of these additional MLR salinity models for the remaining stations in the MMN means that salinity models are now available for all of the estuarine areas within ENP. This means that CERP alternatives can now be examined for their effect in all of the estuaries, providing a complete picture of the potential for restoration. In a general manner the development of all MMN models have shown that there is a definitive link between the upstream hydrologic conditions in the freshwater marshes of the Greater Everglades and the estuaries into which the freshwater is being discharged, be it north or east of Cape Sable. This also means that the restoration of the estuarine areas within Everglades National Park can only happen with the restoration of freshwater levels and volumes in the Greater Everglades.

The use of MLR models was also investigated by evaluating the data being gathered for a CESI project in the mangrove lakes area near Garfield Bight. Though the data are limited at this time due to operational difficulties, these data show promise for use in an area that may be important in recording restoration due to the variability of the salinity that extends over the full range of 0 – 35+ psu.

Applications involving the MLR salinity models comprised the majority of the remaining tasks. Historical reconstructions (daily estimates) of salinity were developed for eleven MMN stations using the MLR salinity models for the period 1965-2000. A small number of stage monitoring stations have a period of record that extends back in time beyond about 1985, including P33. Univariate (simple) linear regression models for stage at a location in the Everglades with very good fit to observed data are possible due to the cross-correlation that exists between stage measurements. Therefore, the stage input data were extended through the use of simple regression models. Wind and sea surface elevation (observed) are available for the entire period and were also used for the historical reconstructions.

All of the reconstructions were consistent in expressing the higher salinity conditions that existed during the extended drought periods of the early 1970s and 1980s. Spectral analysis of the reconstructions showed that there may be some minor errors in the historical reconstructions, but there are generally too few observations to determine any aberrant patterns. The historical

reconstructions will have the most value as input to ecological models as will be demonstrated further in this discussion.

Another application of MLR salinity models and other statistical models that was accomplished for this project was the use of historical reconstruction and paleo-based salinity time series data as input to the pink shrimp growth dynamics model developed by Browder et al (2002). This task was a proof-of-concept exercise that focused on the coupling of statistical salinity model output with pink shrimp models. The results consisted of output from the shrimp model with limited interpretation.

Another application of MLR models was the use of univariate and MLR models to fill gaps in stage, salinity, tidal creek flows, and water temperature data when it was determined that SARIMA models were only useful for gaps of one to about three values. The conclusions of this effort to determine if Autoregressive Moving Average (ARIMA) models could be useful for filling gaps in time series data for stage, salinity, temperature, and flow are clear: ARIMA models are not very useful for this purpose. This in-depth analysis of ARIMA models brought to the surface the level to which this procedure relies on the availability of past values of the dependent variable, which are not available when the data gap exceeds one day for daily data. ARIMA models work reasonably well for gaps of 2-3 days particularly when the range of the data in the 2-3 day period is near the mean value for time series. However, the ability of ARIMA models to provide highly accurate one-step forward predictions using the serial correlation in the data means that ARIMA models should be considered for use as a predictive tool for water management system operation when the data are not missing.

Similarly, the obvious limitations of linear interpolation limit the use of this technique to gaps that are less than 5 days in duration for the 4 parameters of interest. Beyond a 5-day window, the chance that an event will occur that modifies the rate of change increases such that the use of linear interpolation may not be reliable.

Multivariate linear regression (MLR) models proved to be useful for filling data gaps in stage, salinity, and flow. Based on the limited sample of this evaluation MLR models appear to be able to explain 85-95% (sometimes more) of the variability in the stage, salinity, and flow data, and wind appears to be useful in modeling by explaining up to 5% additional variation over models that do not include wind. However, models that include wind tend to have a large number of independent variables even at the 0.999 level of significance.

In the case of temperature, the correlation of neighbor stations as independent variables to the dependent variables was so high that MLR models did not provide any benefit. Instead simple univariate linear regression models were highly adequate. The use of bottom measurements to model top measurements

at the same station (and vice versa) provides an advantage over the use of neighbor data as an independent variable.

Based on this analysis the flexibility of MLR models was shown to be useful for filling data gaps. MLR models have the added advantage of being transparent and easy to understand by a wide audience. The use of neighbor independent variables and wind data in MLR models for filling data gaps is recommended when simple regression models are considered to be inadequate.

Another model application was the use of statistical models (both univariate and MLR) to estimate the flow needed to meet the performance measures for the Florida Bay and Florida Keys Feasibility Study. It was concluded that the procedure using a suite of univariate linear regression models can be successfully used to simulate long-term average flow conditions (monthly mean and annual average). The procedure takes advantage of the statistical power in a large number of observations with which to develop models and a large number of input values for simulations.

According to this analysis, the current flows through the S-18C and S-12T structures and TSB are well below the flows needed to meet the FBFKFS salinity targets. However, the flows currently delivered to the S-18C structure are closer to the target than the flows needed at S-12T and TSB, because of the water management operations at these structures.

Both MLR salinity models and the FATHOM mass-balance model were coupled with the output from the SICS / TIME model being developed by the USGS showing that these models can be coupled, though not directly at this time. The use of modeled input data for salinity simulations by mechanistic models is necessary because the standard period for evaluations of water management alternatives spans a 36-year period, and observed data for some model inputs are not fully available. The use of a 36-year period for south Florida simulations is warranted by the significant difference in wet and dry periods over years to decades, and the ecological implications of anthropogenic alterations that may only be expressed over longer periods of change in the salinity regime.

When the status of the current salinity models for Florida Bay were evaluated, it was found that considerable progress had been made in the development and refinement of salinity models since the report in 2002 by the Cadmus Group (Nuttle, 2002). To-date, the most widely used models for developing historical recreations and simulating salinity regimes for the evaluations of water management alternatives and ecological modeling are the MLR salinity models and the FATHOM mass balance model. Mass balance and MLR salinity models have already been applied in a number of ways and are still being refined. However, the development of hydrodynamic models, particularly the USGS TIME and the EFDC Florida Bay model by SFWMD, is continuing.

By design hydrodynamic models are intended for detailed and spatially discrete applications because of the effort and cost to calibrate, validate, and run large-scale hydrodynamic models for regional scenarios. Statistical and mass balance models will likely remain in use for planning-level decisions on a regional basis. Where possible, it appears that it will be less-expensive and time-consuming to utilize both statistical and mass balance models together as multiple lines of evidence and corroboration compared to utilizing only one hydrodynamic model for regional evaluations.

A comparison of observed salinity data and forecasts made by MLR and FATHOM salinity models by plotting the data for Long Sound and Whipray Basin. This comparison showed that both FATHOM and MLR salinity models are capable of providing reasonable estimates of salinity in Florida Bay. Additionally it was shown that ENP MMN and SERC (FIU) data sets can be used interchangeably at the monthly level.

Based on this review of the current status (June, 2006) of the models available for simulating and forecasting salinity in Florida Bay, Whitewater Bay, and the Gulf coast estuaries, it is found that MLR salinity models, FATHOM, and the SICS / TIME models appear to be providing the most reasonable estimates of salinity at the time of this report, with corroborating results for salinity variation at the limited locations that were evaluated. In addition, these three models, and the EFDC model if model fidelity can be improved, meet most of the salinity modeling goals of the PMC (2004).

Task 13 produced a manuscript that was submitted to *Estuaries and Coasts* and peer reviewed. The paper discussed the development of statistical models for the estimate of paleoecology-based salinity, stage, and flow. The paper presented the models and findings. The manuscript was reviewed by two reviewers, but revision were requested. Both reviewers considered the lack of information on the paleoecological data as a flaw, even though both reviewers commented favorably on the use of paleo-data and modeling as a novel approach worth pursuing further. Because of this the PI collaborated further with a paleoecologist and an estuarine ecologist to provide the background that the reviewers consider necessary. As of the date of this report, a revised manuscript has been submitted to *Estuaries and Coasts*.

When all of the tasks of this project are considered as a whole it is clear that most tasks worked well. The accomplishments of the tasks that worked well are summarized briefly below:

- The evaluation of the use of evaporation surrogates showed that evaporation was an important factor is explaining the variation in salinity within Florida Bay when the correlation between evaporation surrogates and salinity was considered. However, of all of the components of the water budget of Florida Bay, evaporation is the most difficult to directly

measure and is most frequently utilized in some sort of derived manner from other more easily quantified parameters. It was recommended that evaporation surrogates not be included in MLR salinity model development as the new models are being developed.

- The FATHOM mass-balance model was upgraded to include additional basins so that the total number of basins or “cells” is now 58.
- With respect to the development of new models, the work that was done completed the job of salinity model development at all MMN stations.
- All of the MMN MLR salinity models have been coded by the IMC and are available for CERP and any other applications.
- The FATHOM model developer met with ENP staff to insure that the model inputs were being applied appropriately. This assistance validated the model use and allowed ENP staff to have confidence in the predictions.
- The MLR salinity models were successfully coupled with pink shrimp growth and survival models in a proof-of-concept exercise. The interpretation of the output from the ecological model is on-going.
- The status of salinity modeling evaluation provided valuable information to ENP staff on the level of development of salinity models. The best models according to the model error statistics were the MLR salinity models, the Florida Bay Four Box model, FATHOM, SICS/TIME, and EFDC.
- The flow estimate procedure that was developed for the purpose of estimating the flow needed to meet the performance measures of the Florida Bay and Florida Keys Feasibility Study may be the task that has, to date, been the most successful. An application of the developed procedures followed as the Southern Estuaries Sub-team realized the commissioned the PI to develop a paleo-based salinity regime estimate.
- The historical reconstruction procedure was developed to extend the observed data to include the full period of the evaluations being made for CERP, 1965 – 2000.
- The mangrove lake salinity study showed that this area of highly variable salinity may be important in providing information on restoration alternatives because of the large range in salinity at the three stations that are being monitored.
- Though it required an extended effort, a revised journal manuscript has been submitted for the paleoecological adaptation of the flow estimation scheme. This was accomplished through additional collaboration between others involved in the analysis and has been submitted to *Estuaries and Coasts*.

While most tasks worked very well there was one task that was only partially successful, subject to being improved with some additional work that was beyond the scope of this project. Task 8 evaluated the use of seasonal autoregressive moving average models (SARIMA) to fill in gaps in observed data. It was found that autocorrelation (serial correlation) is prevalent in south Florida hydrologic and climatologic data. However, it is difficult to incorporate it into models that are

capable of hindcasting as is needed for south Florida restoration evaluations. Even so, the utility of multivariate linear regression models to fill in gaps (as was seen for historical reconstructions, task 12) was shown as MLR models were capable of providing reasonable estimates of salinity, stage, temperature, and (to a lesser extent) flow.

IV. Recommendations

Because this was a research project, follow-on projects and tasks beyond the current scope of work were identified as some of the tasks evolved. This follow-on work is discussed briefly below.

- Additional historical reconstructions

The historical reconstructions provide temporal evidence of past hydrologic conditions, both natural and anthropogenic. The use of historical reconstructions was shown through the Florida Bay MFL modeling work for the SFWMD (Hunt et al, 2005). The spatial extent of this evidence can be extended through the development of additional MLR salinity model reconstructions at the other MMN stations that were not a part of this project. The historical reconstructions are needed by the ecological modelers because some of the evidence of past ecological conditions comes from synoptic studies and anecdotal data that are infrequent. The historical reconstructions provide a way to bring those data into the analysis. Additionally, a body of data collected by Roblee (unpublished) on salinity from synoptic studies and anecdotal data would become useful with additional spatial coverage for historical salinity reconstructions.

- Work with other paleoecologists to verify the paleo-based salinity regime

The next CESI project (currently underway) includes work with Evelyn Gaiser of FIU. That task should be expanded to include the 3 additional stations USGS sediment cores that are being interpreted by Lynn Wingard of the USGS, because that paleoecological data can be used with ENP Marine Monitoring Network salinity data stations nearby –. Bob Allen Key holds the most promise.

The concept for the additional work is to develop the new regression relationships between Bob Allen Key, Little Madeira Bay, and Little Rabbit Key salinity and the primary upstream stations (primarily CP and P33), then use the transfer functions that have already been developed to estimate flow into Shark River Slough and Taylor Slough, stage at other stations in the Everglades (for hydroperiod calculations), and salinity in Florida Bay, all as was done previously.

The tasks involved are generally: (1) development of the salinity regime using the paleo data including coordination with Wingard and Pitts; (2) development of regression models for application of the paleo-salinity data to estimate upstream stage for a given salinity; (3) application of the existing models to estimate flow, hydroperiod, and salinity; and (4) interpretation of the results.

Then the results from all of the Florida Bay paleoecological analyses - Whipray Basin, Bob Allen Key, Little Madeira Bay, Little Rabbit Key, and Gaiser's work – should be combined into a manuscript with multiple authors. It will have a broadened spatial basis in paleoecology and a good description of the

interpretation of the paleoecological work that went into the establishment of the basis for the paleosalinity regime(s) that will be developed.

- Prepare manuscript for the MLR salinity models

There are now MLR salinity models for all stations in Florida Bay, Whitewater Bay, Shark River estuary, upper western Gulf coast (Big Cypress Preserve area), and Barnes Sound / Manatee Bay. The limited number of MLR salinity models that were available proved valuable in CERP alternative evaluations and for other analyses. However, the MLR salinity models have not been peer reviewed nor are they all readily available in one source for use by other researchers. At the present, given the lack of success in developing a comprehensive hydrodynamic model of Florida Bay, the MLR salinity models and FATHOM are the best tools available for salinity in the study area. It would be a relatively straight-forward task to document the MLR salinity model development procedure in a manuscript that would be well suited for process-oriented journals.

- Continue work with pink shrimp model

The output of the pink shrimp model coupled with observed salinity data, NSM 4.6.2, and paleo-based salinity was thought-provoking and needs further interpretation, which requires coordination with the shrimp modelers. It is recommended that the PI continue working with the shrimp modelers, perhaps to expand the spatial extent of the simulations using other MLR salinity models. If the paleoecological studies can be broadened as recommended above, the confidence in the paleo-based salinity regime being used with any ecological modeler will improve.

- Utilize the MLR salinity models with other ecological models

The historical reconstructions of salinity and the paleo-based salinity simulations can be used for freshwater fishes and other aquatic fauna and wading bird ecological models and Habitat Suitability Indices (HSI). The daily time step and the long period of the simulations (36 year) will allow the uncertainty in the historical simulations to be accurately characterized. The historical simulations can be analyzed by comparing the simulated conditions for wet, dry and normal rainfall years.

The models that have been produced by this and previous CESI studies provide: (1) a historical simulation of salinity and hydrology over the past 30-40 years, and (2) pre-drainage salinity and hydrology at key locations in Florida Bay and the Everglades. The developed salinity and hydrologic regimes can be coupled with ecological models and HSIs to improve and upgrade the indicators being used for evaluation and assessment performance measures. Assistance to and coordination with the researchers who have developed the ecological models and the developers of the indicators and performance measures should be part

of this effort to ensure that the simulation data are appropriately applied, and to assist with the interpretation of the results. The output of the coupling of the historical reconstruction of salinity and hydrology and the ecological models can be compared to observed conditions to provide a more complete baseline and to verify ecological models.

- Utilize TIME stage output with MLR salinity models and FATHOM

At this point, it appears that TIME is ready for use. However the post-processing tools to utilize the output are not fully developed. When this data becomes available it can be coupled with the MLR salinity models to provide a better tool for analysis of CERP alternatives and for other restoration studies.

- Continue working with mangrove lakes researchers

The limited work that was done with the mangrove lake data (because the project has just begun) shows much promise for these stations. This area is an important transition zone between fresh and salt water and may provide the most information for assessment of restoration projects based on the wide range of salinities that were observed.

V. References

Includes all references used for this project. See task reports in the appendices for references related to a specific task.

Abtew, W., J. Obeysekara, M. Irizarry-Ortiz, D. Lyons, A. Reardon. 2003. Evaporation Estimation for South Florida. Technical Paper EMA # 47, Proceedings of the World Water and Environmental Resources Congress 2003.

Allen, R. 1997. Self-calibrating Method for Estimating Solar Radiation from Air Temperature. Journal of Hydraulic Engineering. American Society of Civil Engineers, Vol 2., No. 2, April 1997.

Boyer, J. N., J. W. Fourqurean, and R. D. Jones. 1997. Spatial characterization of water quality in Florida Bay and Whitewater Bay by multivariate analyses: zones of similar influence. Estuaries, 20:743-758.

Boyer, J. N. and R. D. Jones. 2001. Long-term trends in water quality of Florida Bay (June 1989 - June 1999). Abstract, Florida Bay Science Conference, Key Largo, FL, April 24-26, 2001.

Brockwell, P.J. and R.A Davis. 1996. Introduction to Time Series and Forecasting. Springer-Verlag, New York, New York. 420 pp.

Browder, J. A., Z. Zein-Eldin, M.M. Criales, M. B. Robblee, S. Wong, T. L. Jackson and D. Johnson. 2002. Dynamics of pink shrimp (*Farfantepenaeus duorarum*) recruitment potential in relation to salinity and temperature in Florida Bay. Estuaries 25 (6B):1355-1371.

Browder, J. A., D. R. Johnson, R. Bennett, F. Marshall, and J. Wang. 2005. Application of a simulation model of pink shrimp growth and survival. (Abstract). Florida Bay Science Conference.

Cerco, C.F., B.W. Bunch, A.M. Teeter, and M.S. Dortch. 2000. Water Quality Model of Florida Bay. Environmental Laboratory, ERDC/EL TR-00-10. U.S. Army Corps of Engineers. Vicksburg, Mississippi.

CERP, 2002. Model Uncertainty Workshop Report. Comprehensive Everglades Restoration Plan. Available from: <http://www.evergladesplan.org/pm/recover/recover_docs/mrt/052402_mrt_uncertainty_report.pdf> [cited 20 May 2005].

CERP, 2004. Interangency Modeling Center Program Management Plan. Comprehensive Everglades Restoration Plan. Available from: <http://www.evergladesplan.org/pm/progr_imc_plan.cfm> [cited 20 May 2005].

Cline, J.C. and E. D. Swain. 2002. Coupling hydrologic and ecological modeling: SICS and ATLSS. Second Federal Interagency Hydrologic Modeling Conference, July 28-August 1, 2002, 12 pp.

Cline J. C., J. J. Lorenz and E. D. Swain. 2006. Linking Hydrologic Modeling and Ecologic Modeling: Application of a Spatially-Explicit Species Index (SESI) Model for Adaptive Ecosystem Management in the Everglades Mangrove Zone of Florida Bay. Abstract for 2006 Greater Everglades Ecosystem Restoration Conference, Lark Buena Vista, Florida, June 5-9, 2006.

Corbett, D. R., J. Chanton, W. Burnett, K. Dillon, C. Rutkowski, and J. W. Fourqurean. 1999. Patterns of groundwater discharge into Florida Bay. Limnol. Oceanogr. 44:1045-1055.

Cosby, B.J. 1993. An Examination of the Relationships of Stage, Discharges and Meteorology in the Panhandle and Taylor Slough Areas of Everglades National Park to Salinity in Upper Florida Bay, Volumes 1-5. University of Virginia, Charlottesville, Virginia.

Cosby, B.J., W.K. Nuttle, and J.W. Fourqurean, 1999. FATHOM: Model Description and Initial Application to Florida Bay. (progress report submitted to Everglades National Park)

Cosby, B.J., W.K. Nuttle, and J.W. Fourqurean. 1999. FATHOM – Flux Accounting and Tidal Hydrology at the Ocean Margin: Model Description and Initial Application to Florida Bay. Report to the Florida Bay Project Management Committee (PMC) and the Everglades National Park (ENP), National Park Service. Dept. of Environmental Sciences, University of Virginia, Charlottesville, VA.

Cosby, B.J., J.W. Fourqurean, and W.K. Nuttle, 2004. FATHOM Florida Bay Simulation 1991- 2002 Updated Model and Preliminary Results. Interim project report for Everglades National Park by Florida International University. January 2004.

Cosby, B.J., W.K. Nuttle, and F. Marshall, 2005. FATHOM Enhancements and Implementation to Support Development of MFL for Florida Bay. Final Report on Contract C-C-15975-WO05-05 for the South Florida Water Management District. Environmental Consulting & Technology, Inc. New Smyrna Beach, Florida.

CROGEE. 2002. Florida Bay Research Programs and Their Relation to the Comprehensive Everglades Restoration Plan. The National Academies Press, Washington, D.C.

DeAngelis, D. L., S. Bellmund, W. M. Mooij, M. P. Nott, E. J. Comiskey, L. J. Gross, and W. M. Wolff. 2002. Modeling ecosystem and population dynamics on

the South Florida hydroscape, p. 239-258. In J. W. Porter and K. G. Porter (eds.), The Everglades, Florida Bay and Coral Reefs of the Florida Keys: An Ecosystem Sourcebook, CRC Press, Delray Beach, FL.

Enfield, D.B., A.M. Mestas-Nunez, P.J. Trimble. 2001. The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S. Geophysical Research Letters, Vol. 28, No.10. pp. 2077-2080

Everglades National Park. 1997a. Everglades National Park Marine Monitoring Network 1994 Data Summary. Everglades National Park, Homestead, Florida. 67 pp.

Everglades National Park. 1997b. Everglades National Park Marine Monitoring Network 1995 Data Summary. Everglades National Park, Homestead, Florida. 67 pp.

Frankovich, T. A., J. W. Fourqurean, and D. Morrison. March 23, 2007. Submerged aquatic vegetation and physicochemical monitoring in the Florida Bay mangrove zone for CERP assessments and targets refinement. Annual Progress Report. Critical Ecosystem Studies Initiative Cooperative Agreement CA H5297-05-0041.

German, E.R., 2000. Regional Evaluation of Evapotranspiration in the Everglades. Water-Resources Investigations Report 00-4217. U.S. Geological Survey, Tallahassee, FL.

Hamrick, J.H. and M.Z. Moustafa. 2003. Florida Bay Hydrodynamic and Salinity Model Analysis. Conference abstract from Joint Conference on the Science and Restoration of the Greater Everglades and Florida Bay Ecosystem.

Hittle, C., E. Patino, and M. Zucker, 2001. Freshwater Flow from Estuarine Creeks into Northeastern Florida Bay. Water-Resources Investigations Reports 01-4164, US Geological Survey, Tallahassee, Florida.

Hunt, J. and W.K. Nuttle (eds). Florida Bay Science Program. A Synthesis of Research on Florida Bay. Florida Marine Research Institute (in prep.).

Irizarry-Ortiz, M. 2003. Selected Methodology for Long-Term (1965-2000) Solar Radiation and Potential Evapotranspiration Estimation for the SFWMM2000 Update. Memorandum to Ken Tarboten, December 12, 2003. South Florida Water Management District, West Palm Beach, Florida.

Johns, E., P. Ortner, R. Smith, D. Wilson, T. N. Lee, and E. Williams. 2001. Salinity Variability in Florida Bay from Monthly Rapid High Resolution Surveys. Abstract, Florida Bay Science Conference, Key Largo, FL, April 24-26, 2001.

Johnson, D.R., J. A. Browder, and M. B. Robblee. 2005. Statistical models of FL Bay fishes and crustaceans to evaluate minimum flow levels in FL Bay. Contribution No. PRD 04/05-06. Protected Resources and Biodiversity Division, SEFSC. 474 pp.

Jones, R.D., and J.N. Boyer, 2001. An Integrated Surface Water Quality Monitoring Program For The South Florida Coastal Waters: Fy2000 Cumulative Report to the South Florida Water Management District (C-10244) and Everglades National Park. Southeast Environmental Research Center, Florida International University, Miami, FL 33199. (http://serc.fiu.edu/wqmnetwork/)

Kashigan, S.K. 1991. Multivariate Statistical Analysis. Radius Press. New York. pp117-192.

Marshall III, F. E.; D. Smith; and D. Nickerson. 2004. Using Statistical Models to Simulate Salinity Variation and Other Physical Parameters in North Florida Bay. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Kourafalou, Villy. 2005. SoFLA-HYCOM (South Florida HYCOM) Regional Model around Florida Straits, Florida Bay and the Florida Keys: An overview. Rosenstiel School of Marine and Atmospheric Science, University of Miami Division of Meteorology and Physical Oceanography, Miami (http://hycom.rsmas.miami.edu/overview/SoFLA_HYCOM.pdf

Langevin, C.D., E. D. Swain, and M. A. Wolfert. 2002. Numerical simulation of integrated surface-water/groundwater flow and solute transport in the southern Everglades, Florida, Second Federal Interagency Hydrologic Modeling Conference, July 28-August 1, 2002, 12 p.

Langevin, C.D. E.D. Swain, J.D. Wang, M.A. Wolfert, R.W. Schaffranek, and A.L. Riscassi. 2004a. Development of coastal flow and transport models in support of everglades restoration. Fact Sheet 2004-3130. U.S. Department of the Interior, U.S. Geological Survey. October 2004.

Langevin, C.D. E.D. Swain, and M.A. Wolfert. 2004a. Simulation of integrated surface-water / groundwater flow and salinity for a coastal wetland and adjacent estuary. Open-File Report 2004-1097. U.S. Department of the Interior, U.S. Geological Survey. October 2004.

Lee, T. N., and E. Williams. 1999. Mean distribution and seasonal variability of coastal currents and temperature in the Florida Keys with implications for larval recruitment. Bulletin of Marine Science 64:35-56.

Lee, T. N., E. Johns, D. Wilson, and E. Williams. 1999. Site characterization for the Tortugas region: Physical Oceanography and Recruitment. Tortugas 2000: Report by the Florida Keys National Marine Sanctuary.

Lettenmaier, D.P. and E.F. Wood, 1993. Chapter 26: Hydrologic Forecasting. in D.R. Maidment (ed): Handbook of Hydrology. McGraw Hill, New York.

Lorenz, J. 2005. Habitat Suitability Index for Roseate Spoonbills Nesting in Northeastern Florida Bay. page 43 in Program and Abstract Book, 2005 Florida Bay and Adjacent Marine Systems Science Conference, December 2005, Duck Key, Florida.

Madden, C.J. and A.A. McDonald. 2006, draft. An Ecological Model of the Florida Bay Seagrass Community – Model Documentation. South Florida Water Management District, West Palm Beach, Florida.

Marshall, F.E. 2000. Florida Bay Salinity Transfer Function Analysis Volume 1 of 2: Final Report. Cetacean Logic Foundation, New Smyrna Beach, Florida. 24 pp.

Marshall III, F. E.; D. Smith; and D. Nickerson. 2003. Salinity Simulation Models for North Florida Bay Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. 41 pp.

Marshall III, F. E. 2003. IOP Salinity Analysis Using Statistical Models. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida. 35 pp.

Marshall III, F. E.; D. Smith; and D. Nickerson. 2004. Using Statistical Models to Simulate Salinity Variation and Other Physical Parameters in North Florida Bay. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F. E. 2005a. ICU Runs Summary Report, for Southern Estuaries Sub-team of RECOVER. Environmental Consulting & Technology, Inc. New Smyrna Beach, Florida.

Marshall III, F. E. 2005b. RECOVER Southern Estuaries performance measures: identification of hydrology-salinity relationships for coastal estuaries and analysis of interim CERP update scenarios. Environmental Consulting & Technology, Inc., New Smyrna Beach, Florida.

Marshall III, F. E. 2005c. Using Statistical Models to Simulate Salinity Variability in Estuaries. *In* Estuarine Indicators, S. A. Bortone, Editor. CRC Press. Boca Raton, Florida. Pp 33-52.

Marshall III, F .E. 2008a. Task 1 – Investigation of Evaporation Surrogates for Florida Bay Salinity Modeling. CESI Cooperative Agreement Number CA H5284-05-0006, Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F .E. 2008b. Task 2 – Use of the FATHOM Florida Bay MFL Model with SICS/TIME Input. CESI Cooperative Agreement Number CA H5284-05-

0006, Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F .E. 2008c. Task 3 – Development of Additional Multivariate Linear Regression Salinity Models for Florida Bay and the Southwest Gulf Coast. Everglades National Park. CESI Cooperative Agreement Number CA H5284-05-0006, Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F .E. 2008d. Task 4 – Assist ENP with Implementation of the FATHOM Florida Bay MFL Model. Everglades National Park. CESI Cooperative Agreement Number CA H5284-05-0006, Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F .E. 2008e. Task 5 – Investigate the Coupling of Pink Shrimp Models with MLR Salinity Models. Everglades National Park. CESI Cooperative Agreement Number CA H5284-05-0006, Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F .E. 2008f. Task 6 – Coordinate with the IMC to Incorporate the New MLR Salinity Models into the Post-processing Tools. Everglades National Park. CESI Cooperative Agreement Number CA H5284-05-0006, Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F .E. 2008g. Task 7 – Simulating and Forecasting Salinity in Florida Bay: A Review of Models. Everglades National Park. CESI Cooperative Agreement Number CA H5284-05-0006, Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F .E. 2008h. Task 8 – Evaluation of Methodologies for Filling Gaps in Florida Bay Salinity, Stage, Water Temperature, and Flow Data. CESI Cooperative Agreement Number CA H5284-05-0006, Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F .E. 2008i. Task 9 - Estimation of the Flow Required for the Florida Bay and Florida Keys Feasibility Study Salinity Performance Measures. CESI Cooperative Agreement Number CA H5284-05-0006, Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F .E. 2008j. Task 10 – The Use of SICS/TIME Output with MLR Salinity Models. CESI Cooperative Agreement Number CA H5284-05-0006, Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F .E. 2008k. Task 11 – Historical Reconstructions of Salinity in Northeast Florida Bay using Multivariate Linear Regression Models. CESI

Cooperative Agreement Number CA H5284-05-0006, Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F .E. 2008l. Task 12 – Evaluate the Feasibility of Developing MLR Salinity Models for the Mangrove Region North of Garfield Bight and Terrapin Bay. CESI Cooperative Agreement Number CA H5284-05-0006, Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Marshall III, F .E. 2008m. Task 13 – Prepare and Submit a Manuscript for Publication in a Referred Journal. CESI Cooperative Agreement Number CA H5284-05-0006, Everglades National Park. Cetacean Logic Foundation, Inc. New Smyrna Beach, Florida.

Nash, J.E., and J.V. Sutcliffe, 1970. River flow forecasting through conceptual models: Part 1 – a discussion of principles. Journal of Hydrology 10:282-290.

Neter, J.; W. Wasserman; M. H. Kutner. 1990. Applied Linear Statistical Models. Richard D. Irwin, Inc., Boston, Massachusetts.

NRC, 2002a. Risk Analysis and Uncertainty in Flood Damage Reduction Studies. Committee on Risk-Based Analysis for Flood Damage Reduction, Water Science and Technology Board, Commission on Geosciences, Environment, and Resources, National Research Council, National Academy Press, Washington, D.C.

NRC, 2002b. A Review of the Keys Carrying Capacity Study. Report of the Committee to Review the Florida Keys Carrying Capacity Study, Ocean Studies Board, Water Science and Technology Board, National Research Council, Washington, DC.

NRC, 2004. Adaptive Management for Water Resources Project Planning. National Academies Press, Washington, D.C. Available from: <http://www.nap.edu/catalog/10972.html> [cited 20 May 2005].

Nuttle, W.K., J.W. Fourqurean, B.J. Cosby, J.C. Zieman, and M.B. Robblee, 2000. The influence of net freshwater supply on salinity in Florida Bay. Water Resources Research 36:1805-1822.

Nuttle, W.K., R. Price, and P.K. Swart, 2001. Estimating evaporation from Florida Bay. pg 102 in Conference Abstracts, 16th Biennial Conference of the Estuarine Research Federation, November 2001, St. Pete Beach, Florida.

Nuttle, W.K. and R.S. Teed. 2002. Version 1: wetland hydrology and estuarine salinity models for the Taylor Slough/C111 area. Report prepared for the Everglades National Park. The Cadmus Group, Inc. Watertown, MA 02472.

Nuttle, W. K., December 1997. Central and Southern Florida Project Restudy: Salinity Transfer Functions for Florida Bay and West Coast Estuaries, Volumes 1 and 2. Southeast Environmental Research Program, Florida International University, Miami, Florida.

Nuttle, W.K. 2002. Review and Evaluation of Hydrologic Modeling Tools for the Coastal Mangroves and Florida Bay. Project report for Everglades National Park, April 2002. Available from: <http://www.eco-hydrology.com/wkn%20enp%201.pdf> [cited 20 May 2005].

Nuttle, W.K. 2005. Biscayne Bay Coastal Wetland Project Planning Tool Phase I: Hydrology and Salinity Calculations. Project report for Everglades National Park, January 2005. Available from: <http://www.eco-hydrology.com/wkn%20bisc%20tool%20fnl%20mod.pdf> [cited 20 May 2005].

Nuttle, W.K. and F.E. Marshall, 2005. Effect of time scale on patterns and processes of salinity variation in Florida Bay. Estuaries (in review).

Ogden, J. C., and S. M. Davis, K. J. Jacobs, T. Barnes, and H. E. Fling. 2005. The use of conceptual ecological models to guide ecosystem restoration in south Florida. *Wetlands* 25:795-809.

Orlando, S.P., Jr.; M.B. Robblee; C. J. Klein. 1998. Salinity Characteristics of Florida Bay: A Review of Archived Data Set (1955-1995). National Oceanographic and Atmospheric Administration, Office of Ocean Resources Conservation and Assessment, Silver Spring Maryland. 32 pp.

Ortner, P. B., T. N. Lee, P. J. Milne, R. G. Zika, E. Clarke, G. Podesta, P. K. Swart, P. A. Tester, L. P. Atkinson, and W. R. Johnson. 1995. Mississippi River flood waters that reached the Gulf Stream. Journal of Geophysical Research 100:13595-13601.

Pandit, A.; C. El-Khazen; S.P. Sivaramapilli. 1991. Estimation of Hydraulic Conductivity Values in a Coastal Aquifer. Groundwater. Vol. 29, No.2. pp 175-180.

PMC 2000. Florida Bay Models Coordination Meeting. Report of a workshop held in May 2000.

Price, R.M., 2001. Geochemical Determinations of Groundwater Flow in Everglades National Park, University of Miami Dissertation, 307p.

Price, R., W. Nuttle, B. Cosby, and P Swart. 2005. Estimating Evaporation Rates in Time and Space Across Florida Bay. Abstract for the Florida Bay and Adjacent Marine Systems Science Program Conference, Duck Key, Florida, December, 11-14, 2005. Page 126.

Price, R.M., P.K. Swart, and J.W. Fourqurean. Coastal groundwater discharge – an additional source of phosphorus for the oligotrophic wetlands of the Everglades. Hydrobiologia, in press.

RECOVER. 2004, draft. Recommendation for Interim Goals and Interim Targets for the Comprehensive Everglades Restoration Plan Indicators and Prediction Methods. US Army Corps of Engineers, Jacksonville District. Jacksonville, Florida.

Robblee, M. B., D. Smith, R. Halley, and G. Clement. 2001. Temporal and Spatial Characterization of Salinity Pattern in Florida Bay. Abstract, Florida Bay Science Conference, Key Largo, FL, April 24-26, 2001.

Santee, E.R., L. Cadavid, W. Wilcox, R. Novoa, M. Irzarry-Ortiz, A. Ali, L Brehon, J. Barnes, A Montoya, J. Obetsekera, and K. Tarboton. 2003 Recent Enhancements to the South Florida Water Management Model (SFWMM). South Florida Water Management District, West Palm Beach, Florida.

Schaffranek, R.W., H. L. Jenter, and A. L. Riscassi. 2002. Overview of the "Tides and Inflows in the Mangroves of the Everglades" (TIME) project of the U.S. Geological Survey's South Florida Ecosystem Program, Second Federal Interagency Hydrologic Modeling Conference, July 28-August 1, 2002.

Scully, S.P. 1986. Florida Bay Salinity Concentration and Groundwater Stage Correlation and Regression. South Florida Water Management District, West Palm Beach, Florida.

Sheng, Y. P., J. R. Davis, and Y. Liu. 1995. A Preliminary Model of Florida Bay Circulation. Final Report to the National Park Service. University of Florida, Gainesville, Florida.

Smith, D. 1997. Everglades National Park Marine Monitoring Network 1996 Data Summary. Everglades National Park, Homestead, Florida. 94 pp.

Smith, D. 1998. Everglades National Park Marine Monitoring Network 1997 Data Summary. Everglades National Park, Homestead, Florida. 100 pp.

Smith, D. 1999. Everglades National Park Marine Monitoring Network 1998 Data Summary. Everglades National Park, Homestead, Florida. 100 pp.

Smith, D. 1997. Everglades National Park Marine Monitoring Network 1996 Data Summary. Everglades National Park, Homestead, Florida. 94 pp.

Smith, D. 2001. Everglades National Park Marine Monitoring Network 1999 Data Summary. Everglades National Park, Homestead, Florida. 65 pp.

Sutula, M., J.W. Day, J. Cable, and D. Rudnick, 2001, Hydrological and nutrient budgets of freshwater and estuarine wetlands of Taylor Slough in Southern Everglades, Florida (U.S.A), Biogeochemistry, 56, pp. 287-310.

Smith, R. H., E. Johns, D. Wilson, T. N. Lee, and E. Williams. 2001. Moored Observations of Salinity Variability in Florida Bay and South Florida Coastal Waters on Daily to Interannual Time Scales. Abstract, Florida Bay Science Conference, Key Largo, FL, April 24-26, 2001.

Swain, E.D., Wolfert, M.A., Bales, J.D., and Goodwin, C.R., 2004, Two-dimensional hydrodynamic simulation of surface-water flow and transport to Florida Bay through the Southern Inland and Coastal Systems (SICS): U.S. Geological Survey Water-Resources Investigations Report 03-4287.

Tabb, D. 1967. Prediction of estuarine Salinities in Everglades National Park, Florida by the use of ground water records. PhD Dissertation, University of Miami, Miami, Florida. 107 pp.

Tetra Tech, Inc. 2005. Development of a Florida Bay and Florida Keys hydrodynamic and water quality model: hydrodynamic model calibration, interim status report. Tetra Tech, Inc. Fairfax, VA.0

Thom, R.M., H.L. Diefenderfer, and K.D. Hofseth, 2004. A Framework for Risk Analysis in Ecological Restoration Projects. IWR Report 04-R-02.
Twilley, R. [n.d.] Evaluating the Response of Mangrove Wetlands to Alterations of Freshwater Flow in Everglades National Park. [online] Available from: <http://www.ucs.louisiana.edu/~rrt4630/mangrove-restudy.htm> [cited 20 May 2005].

Top, Z., Brand, L. W., R.D. Corbett, W. Burnett, and J. Chanton, 2001, Helium and radon as tracers of groundwater input into Florida Bay, J. Coastal Research, 17, pp. 859-868.

Trexler, J. C., W.F. Loftus, and J. H. Chick. 2003. Setting and Monitoring Restoration Goals in the Absence of Historical Data: The Case of Fishes in the Florida Everglades. In: D. Bush and J. Trexler (ed), *Monitoring Ecosystems – Interdisciplinary Approaches for Evaluating Ecological Initiatives*. Island Press, Washington, D.C.

Walker, W.W., 1998. Estimation of Inputs to Florida Bay. Report prepared for U.S. Army Corps of Engineers & U.S. Department of the Interior. Available from: http://wwwalker.net/flabay/ [cited 20 May 2005].

Wang, J.D., J. van de Kreeke, N. Krishnan, and D. Smith, 1994. Wind and Tide Response in Florida Bay. Bull. Marine Science, 54(3), 579-601.

Wang, J.D., 1998. Subtidal Flow Patterns in Western Florida Bay. Estuarine, Coastal and Shelf Science, 46(6), 901-915.

Weglarczyk, S. 1998. The interdependence and applicability of some statistical quality measures for hydrological models. Journal of Hydrology 206:98-103.

Wong, T.S.W., 2002. Discussion: Use of artificial flood events to demonstrate the invalidity of simple mixing models. Hydrological Sciences Journal 47:833-837.