Final Report

DATA QUALITY CONTROL, TIME SERIES ANALYSIS AND STATISTICAL MODELING OF SALINITY AND CANAL DISCHARGES IN BISCAYNE NATIONAL PARK

(NPS-CESU Agreement: H5000000494/0002)

Submitted by:

Jiangang Luo[1] & Joseph E. Serafy[1,2]

[1]University of Miami, RSMAS/MBF
4600 Rickenbacker Causeway
Miami, FL33149

[2]NOAA/NMFS/SEFSC
75 Virginia Beach Drive
Miami, FL33149

Submitted to :

Biscayne National Park

October 18, 2002

DATA QUALITY CONTROL, TIME SERIES ANALYSIS AND STATISTICAL MODELING OF SALINITY AND CANAL DISCHARGES IN BISCAYNE NATIONAL PARK

Summary

This document constitutes the final report deliverable as specified under an NPS-CESU agreement (NPS H5000000494/0002). Funds were granted to pursue the following objectives: (1) perform data quality control on water quality data collected by Biscayne National Park (BNP) from 1990 to 1999; (2) acquire relevant data sets (e.g., canal discharges, wind vectors, rainfall); and (3) perform preliminary time series analysis on the data. As part of this study, we corrected many obvious errors in earlier data compilations. While BNP's intent was for data to be collected continuously over the entire time period, in fact, there was only one case where a continuous data series was obtained that spanned longer than one year (i.e., TG06: June 1997 to August 1998). We found that the overall data availability for the AR, BB, BS, CR, TP, and TR sites ranged from 58% to 69% of the entire time period. Of these data, high quality data represented 68% to 98%. It appears that salinity data quality was strongly impacted by probe fouling and mis-calibration. To solve these and other problems, we make the following recommendations:

(1) To reduce data compilation errors, we suggest annual data compilation be passed to a second party to test data readability and to generate data plots for visual inspection for ouliers and other problems.

(2) To reduce data (time series) gaps, we suggest use of two salinity measuring units at the same site with different, overlapping rotation schedules. In addition to serving a "back-up" in case of unit failure, this strategy also provides opportunities for cross-validation of salinity readings.

(3) Mis-calibration problems appear to have occurred mostly prior to 1997. Whatever training and/or personnel changes that occurred from 1997 forward should be recognized and retained, if possible.

(4) Our examination of data strongly implicates that BNP may still be deploying one or two bad units. An effort to identify such unit(s) is recommended

(5) Probe fouling appears to be the major cause of poor data quality. Unless instruments can be treated with anti-fouling agents, we suggest frequent probe cleaning within the instrument unit rotation period and/or the reduction of unit deployment periods. Data indicated that fouling can occurs in as little as two weeks. If such a short deployment period is impractical, we recommend a 4-week deployment period with in situ cleaning every 2 weeks.

Finally, we conducted a some preliminary time series analyses on data that we judged to be of adequate quality. Among the insights gained was the result that canal discharges can affect Biscayne Bay salinities for as long as 15 days. More detailed and complete analysis will be continued in our next project. Included in this report is a CD containing Microsoft Access database files and yearly plots of temperature and salinity for each sites.

INTRODUCTION

Freshwater flows into the bays and estuaries of southeastern Florida are strongly affected by human modifications to, and activities on, the watershed. The Comprehensive Everglades Restoration Plan (CERP) is expected to alter land runoff into Biscayne National Park, altering its salinity and nutrient regimes and other aspects of the ecosystem. Among the objectives of the CERP is the replacement of certain canals and related structures with sheet-flow-producing structures for the ultimate restoration of more natural freshwater delivery (and thus salinity) regimes. Currently, seven canals discharge fresh water directly or indirectly into Biscayne National Park (BNP) waters. One of the most profound changes anticipated to occur is the alteration of salinities within the Park, especially in the nearshore habitats along the mainland coast (Serafy et al. 2001). These alterations may have important consequences on plant and animal diversity, distribution and production. While the prediction of restoration impacts is difficult, this process begins with developing an understanding of how the current hydrographic system is affecting Bay water quality todat. Hydrodynamic simulation models are useful in this regard, but such models typically conceal the significance of each of their constituent variables. For management purposes, therefore, it is important to understand each process by examining empirical relationships among salinity variation and several other variables, both singly and in combination. This approach provides insight into the dynamic ans complex interelationships among canal discharges, salinity regimes and biological resources.

Biscayne National Park (BNP) has been deploying benthic-mounted YSI datasondes for over 10 years. These instruments are capable of collecting continous (hourly) water quality data. Until now, these data have not been examined thoroughly for quality nor have they been extensively comapared against other relevant data sets for management applications. In the present study, funds were granted to pursue the following objectives: (1) to perform quality control on 1990 to 2000 water

quality data (i.e., salinity, temperature, and dissolved oxygen) collected on an hourly basis at several locations within BNP and adjacent waters; (2) to acquire relevant data sets (i.e., canal discharge, wind vector, rainfall); and (3) to perform preliminary time series analyses and statistical modeling methods to data sets of adequate quality and temporal scope.

METHODS

Water Quality Data

Biscayne National Park has been collecting water quality data using benthic-mounted YSI datasondes at several locations for over 10 years (**Fig 1**) . The datasondes are capable of measuring, on an hourly basis, water temperature, conductivity, depth and dissolved oxygen concentration (DO). Some datasonde models also record pH, turbidity, and chlorophyll levels.

Canal Discharge Data

The South Florida Water Management District (SFWMD) has been monitoring and managing the canal discharges since the major construction the canal system began over 40 years ago. Currently, 13 canals discharge fresh water into the Biscayne Bay, seven discharge fresh water directly into BNP waters (**Fig 1**). The actual discharge rate values are estimated from head water and tail water levels with an experimentally-calibrated empirical model. According to Swain et al. (1997), the accuracy of discharge rate estimates ranges from 10 to 20%. Our resuts in the present study, therefore, should be tempered accordingly.

Hydrodynamic Model Output

Two hydrodynamic models have been generated for Biscayne Bay. Only one was made available for the present study, namely, a finite-element characteristic model first developed by Wang et al. (1988). This model has undergone substantial refinement in recent years including calibration using empirical data collected from 1995 through 1998 (Wang et al 1988, Wang et al in review). The other model of Biscayne bay hydrodynamics was developed by the U.S. Army Corps of Engineers. This model still at the testing/calibration stage and was not available for the present study.

Calculation of Salinity from CTD Measurements

Typically, only temperature, specific conductivity, and depth were recorded by the datasondes, thus calculation of salinity from these variables was required. This was achieved by first converting specific conductivity to conductivity using the equation below (and provided in the YSI datasonde manual):

$$spcond = cond / (1+TC*(T-25))$$

$$cond = spcond*(1+TC*(T-25))$$

where *spcond* is specific conductivity at 25 ^{o}C, *cond* is conductivity, *TC* is a temperature coefficient (*TC*=0.0191), and *T* is water temperature. Next, we calculated the salinity according the algorithm published in the manual "Practical Salinity Scale Equations" (see IEEE, Journal of Oceanic Engineering, Vol. OE-5, No.1, January 1980, page 14.):

$$R = \frac{C(S, T, P)}{C(35, 15, 0)}$$

T in °C (IPTS '68), P in decibars.

$$R_T = \frac{R}{R_P r_T}; R_P = 1 + \frac{P \times (A_1 + A_2 P + A_3 P^2)}{1 + B_1 T + B_2 T^2 + B_3 R + B_4 R T}$$

$$r_T = c_0 + c_1 T + c_2 T^2 + c_3 T^3 + c_4 T^4$$

$A_1 = 2.070 \times 10^{-5}$ $B_1 = 3.426 \times 10^{-2}$

$A_2 = -6.370 \times 10^{-10}$ $B_2 = 4.464 \times 10^{-4}$

$A_3 = 3.989 \times 10^{-15}$ $B_3 = 4.215 \times 10^{-1}$

$B_4 = -3.107 \times 10^{-3}$

$c_0 = 6.766097 \times 10^{-1}$

$c_1 = 2.00564 \times 10^{-2}$

$c_2 = 1.104259 \times 10^{-4}$

$c_3 = -6.9698 \times 10^{-7}$

$c_4 = 1.0031 \times 10^{-9}$

$$S = \sum_{j=0}^{5} a_j R_T^{j/2} + \frac{(T-15)}{1 + k(T-15)} \sum_{j=0}^{5} b_j R_T^{j/2}$$

$a_0 = 0.0080$ $b_0 = 0.0005$ $k = 0.0162.$

$a_1 = -0.1692$ $b_1 = -0.0056$

$a_2 = 25.3851$ $b_2 = -0.0066$

$a_3 = 14.0941$ $b_3 = -0.0375$

$a_4 = -7.0261$ $b_4 = 0.0636$

$a_5 = 2.7081$ $b_5 = -0.0144$

where C(35,15,0) = 42.914 mmho/cm. We also compared calculated salinity values with datasonde-"calculated" (i.e., recorded) the salinity values in the database as an additional check of datasonde function.

Data Analyses

When performing time series analyses, there are typically two overall approaches that may be taken. One uses autocorrelation and partial autocorrelation functions to study the evolution of a time series through parameter models: this is known as *time domain analysis*. The second approach uses spectral functions to examine the nonparametric decomposition of a time series into its different

frequency components; this is known as *frequency domain analysis*. However, when analyzing environmental data, it is typical to have simultaneous observations of numerous variables, necessitating a multivariate approach. In analyzing multivariate time series data, the goal is to transfer function models that relate an output series to one or more input series. The fundamental tool used in the time domain approach is the cross-correlation function, and the fundamental tool used in the frequency domain approach is the cross-spectral function.

RESULTS

1. Water quality data

We have compiled water quality data for 18 stations located in BNP or in nearby waters (Table 1., Figure 1). These data span from 1990 to 2002. While the intention was continuous data collection, data gaps were unavoidable. The specific dates on which data are available are listed in Table 2 and Table 3.

Table 1. Station information

Station	Years	Latitude	Longitude	Depth(m)	Description
AR	1990-2002	25 23.190	-80 09.775	5.5	Alina's Reef
BB	1990-2002	25 19.097	-80 11.064	5.3	Ball Buoy
BS	1995-2002	25 29.172	-80 08.878	4.7	Bache Shoal
CR	1990-2002	25 23.790	-80 14.900	2.8	Caesar's Creek
TP	1993-2002	25 28.261	-80 17.071	2.5	Turkey Point
TR	1990-2002	25 28.333	-80 06.704	5.6	Triumph Reef
TG01	1997-1998	25 14.204	-80 24.662	1	Manatee Bay
TG02	1997-1998	25 17.310	-80 22.113	2	Card Sound Bridge 4th dolphin
TG03	1997-1998	25 21.167	-80 16.816	2	AngelFish Crk 15 yrds NW "14"
TG04	1997-2000	25 29.683	-80 16.702	2.1	Military Canal out
TG05	1997-2002	25 29.780	-80 19.531	1	Military Canal close
TG06	1997-2002	25 31.637	-80 18.243	1.6	Black Point mkr 10'N of "8"
TG07	1997-2002	25 36.245	-80 17.346	2	Snapper Crk
TG08	1997-2002	25 39.076	-80 15.575	2.4	Gables-by-the-Sea
TG09	1997-2000	25 39.267	-80 09.580	5.2	Biscayne Channel marker "10"
TG10	1997-1998	25 46.230	-80 10.910	3.3	Miami River # 59
TG11	1997-1998	25 48.828	-80 10.492	3.6	Julia Tuttle Csway Marker "35"
TG12	1997-1998	25 53.410	-80 08.390	2.1	Broad Csway marker "13"

Figure 1. Map showing datasonde sampling sites and freshwater discharge canals in the Biscayne Bay. Sampling sites are indicated by the stars; and canal structures are indicated by the white lines.

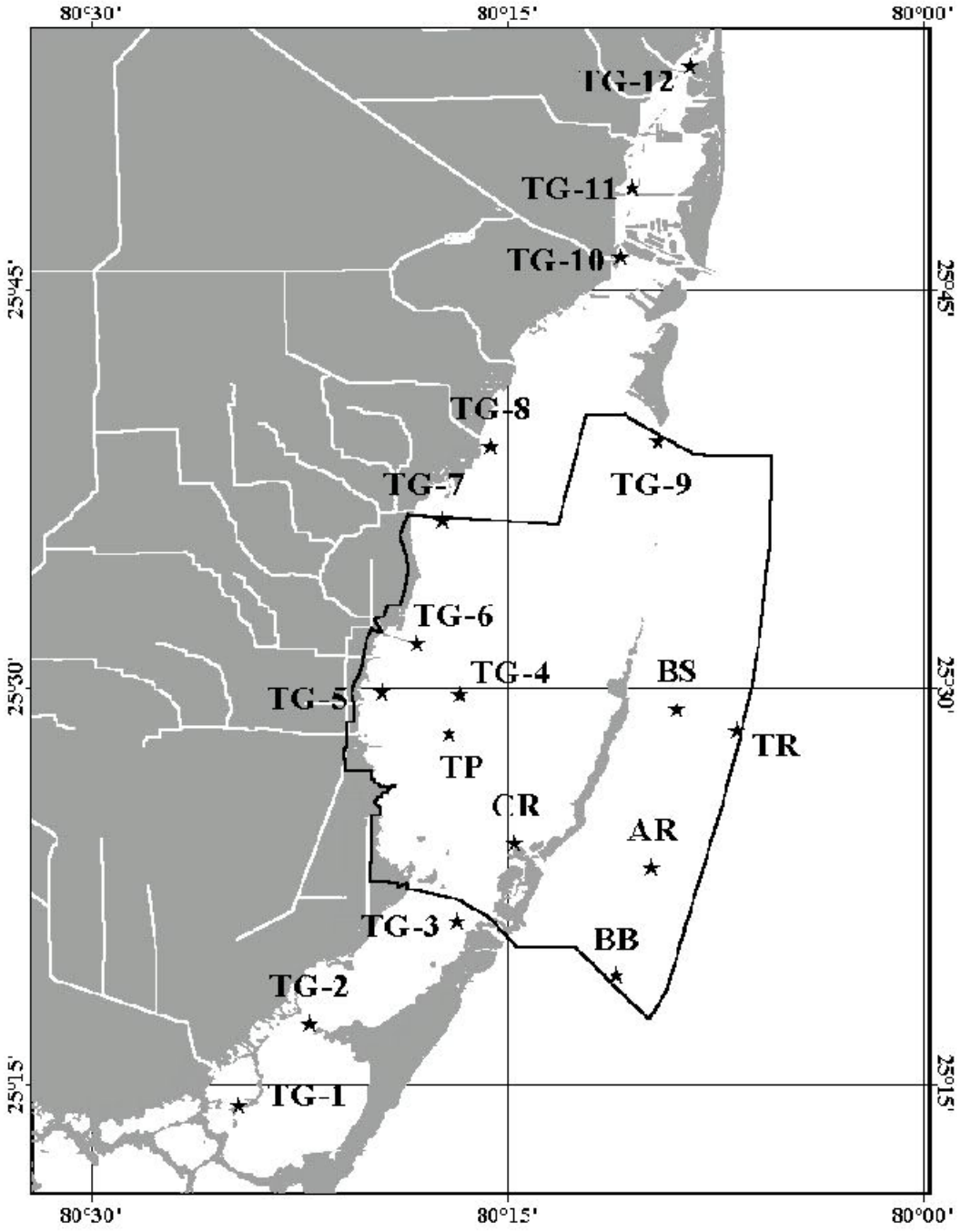

Table 2. Dates of continuous data for AR, BB, BS, CR, TP, TR stations

	Station Name											
Year	AR		BB		BS		CR		TP		TR	
1990	10-Aug	18-Oct	3-Apr	28-Apr			3-Apr	28-Apr			6-Apr	12-May
	4-Nov	10-Nov	10-Aug	6-Sep			19-Sep	18-Oct			23-Aug	26-Sep
							3-Nov	13-Dec				
			8-Dec				22-Dec					
1991	5-Jan	10-Feb		10-Feb				13-Apr			18-May	3-Jul
	10-Feb	11-May	10-Feb	1-Apr			29-Jun	3-Jul			10-Jul	27-Aug
	17-Jul	3-Oct	27-Apr	25-Jun			4-Jul	27-Jul			29-Aug	19-Nov
	3-Oct	4-Nov	29-Jun	13-Jul			6-Aug	4-Sep				
	4-Nov	11-Dec	13-Jul	19-Jul			12-Sep	5-Oct				
			29-Aug	31-Aug			6-Oct	11-Oct				
							15-Nov	14-Dec				
	11-Dec		19-Oct	27-Nov			14-Dec				6-Dec	
1992		4-Jan	6-Jan	31-Mar				22-Feb				20-May
	9-Jan	26-Feb	22-Apr	6-Jun			1-Mar	6-May			21-May	28-May
	27-Feb	10-Apr	7-Jun	10-Sep			20-May	24-Jul			17-Jun	9-Sep
	8-May	4-Jun					3-Aug	21-Aug				
	4-Jun	10-Sep					22-Aug	5-Sep			29-Nov	5-Dec
	14-Nov	14-Dec					6-Sep	9-Sep			6-Dec	18-Dec
			26-Nov				24-Nov				19-Dec	30-Dec
1993	7-Jan	1-Feb		7-Feb				7-Feb	22-Nov	29-Dec	1-Jan	9-Jan
	3-Feb	7-Feb	17-Feb	1-May			10-Feb	17-Mar			10-Jan	14-Feb
	10-Feb	17-Mar	12-May	21-Jul			18-Mar	28-Jun			17-Feb	24-Mar
	18-Mar	21-Jun	29-Aug	22-Sep			5-Sep	12-Oct			12-May	11-Jul
	10-Aug	19-Sep	23-Sep	30-Oct							11-Aug	10-Sep
	22-Sep	3-Dec	3-Nov	9-Nov							19-Sep	9-Nov
	9-Dec		9-Nov	10-Dec			14-Oct				22-Nov	29-Dec
1994		5-Jan	1-Jan	12-Jan				15-Jan	8-Feb	7-Mar	6-Jan	6-Feb
	20-Feb	27-Mar	12-Jan	21-Jan			19-Jan	24-Feb	13-Mar	24-Mar	7-Feb	14-Mar
	4-Apr	16-Apr	1-Feb	2-Feb			19-Mar	22-Mar	7-Apr	3-Aug	15-Mar	6-May
	3-May	3-Jun	3-Feb	31-Jul			20-Apr	27-May	20-Sep	31-Dec	19-May	18-Jul
	7-Jun	5-Dec					8-Jun	20-Sep			4-Aug	25-Sep
							21-Sep	1-Nov			22-Oct	19-Nov
											20-Nov	18-Dec
	11-Dec		18-Aug				27-Nov				18-Dec	
1995		29-Jan		3-Jan	11-Feb	24-Feb		3-Apr	1-Jan	13-Mar		22-Apr
	29-Jan	12-Apr	4-Jan	26-Mar	6-Mar	5-May	27-Nov	3-Apr	14-Mar	26-May	28-Apr	21-Jul
	26-Apr	23-May	26-Apr	15-May	9-Jun	10-Aug	26-Apr	27-Aug	20-Jun	16-Sep	22-Jul	26-Aug
	23-May	15-Jul	15-May	4-Oct	12-Aug	28-Aug	30-Aug	18-Sep	18-Sep	9-Oct	28-Aug	26-Sep
	11-Aug	29-Aug	9-Oct	27-Oct	30-Aug	14-Sep	19-Sep	10-Oct	11-Oct	1-Nov	4-Oct	5-Oct
	29-Aug	4-Oct	2-Nov	24-Nov	16-Sep	4-Oct	30-Oct	21-Nov	7-Nov	22-Nov	13-Oct	13-Oct
	7-Oct	27-Oct	24-Nov	15-Dec	7-Oct	28-Oct	22-Nov	24-Nov	22-Nov	13-Dec	16-Oct	2-Nov
	2-Nov	24-Nov			2-Nov	24-Nov	24-Nov	13-Dec			7-Nov	18-Dec
	24-Nov	16-Dec			27-Nov	18-Dec	15-Dec		15-Dec			
1996	11-Jan	1-Feb	11-Jan	28-Feb	12-Jan	14-Oct		5-Jan		3-Jan	14-Feb	28-Mar
	1-Feb	9-Mar	14-Mar	7-Jul			18-Jan	7-Feb	18-Jan	5-Dec	8-Apr	5-May
	15-Mar	2-Apr	6-Aug	29-Oct			7-Feb	25-Apr			22-May	26-Sep
	16-Apr	28-Apr	30-Oct	26-Nov			25-Apr	16-Oct			10-Oct	23-Nov
	29-Apr	9-May					16-Oct	5-Dec			3-Dec	21-Dec
	15-May	17-Jun					5-Dec	23-Dec			21-Dec	29-Dec
	17-Jun	20-Aug										
	9-Sep	30-Oct										
	1-Nov	1-Dec										
	3-Dec	14-Dec			16-Oct							

Table 2 continued.

1997	23-Jan	9-Apr	7-Jan	25-Jan		20-Jan	8-Jan	1-Feb	7-Jan	2-May	23-Jan	3-Mar
	14-May	16-Sep	28-Jan	21-Feb	23-Jan	7-Aug	27-Feb	27-Mar	2-May	7-Aug	6-Mar	31-Mar
			21-Feb	11-May	8-Sep	3-Nov	2-May	30-Dec			14-Apr	7-Aug
			5-Jun	10-Dec							8-Sep	29-Nov
	30-Sep		10-Dec		4-Nov				8-Sep		29-Nov	
1998		9-Dec		3-Jan		25-Dec	1-Jan	31-Dec		20-Jan		4-Jan
			8-Jan	4-Mar					22-Jan	22-Sep	20-Jan	2-Feb
			15-Apr	21-Sep					1-Oct	29-Dec	3-Feb	15-Feb
			8-Oct	25-Dec							16-Feb	10-Mar
											11-Mar	24-Sep
											25-Sep	25-Sep
											30-Sep	5-Nov
1999	10-Feb	5-Jul	10-Feb	1-Jun	19-Jan	31-Jan	1-Jan	5-Feb	19-Jan	13-Apr	20-Apr	17-Sep
	22-Jul	21-Nov	22-Jul	25-Sep	20-Apr	14-Sep	10-Feb	27-Sep	15-Apr	27-Sep	5-Oct	15-Oct
			20-Oct	21-Dec	13-Oct	17-Oct	13-Oct	18-Oct	12-Oct	20-Oct		
							21-Dec		21-Dec			
2000	11-Jan	30-Mar	11-Jan	14-Feb	20-Apr	23-May		24-Feb		19-Mar	20-Apr	11-Jul
	11-Jul	3-Oct	14-Feb	30-Mar	11-Jul	7-Aug	10-Jul	9-Sep				
	13-Nov	19-Dec	11-Jul	22-Sep								
			13-Nov									
2001	4-Aug	23-Oct		4-Feb	4-Aug	30-Aug	4-Aug	18-Aug	4-Aug	27-Aug	4-Aug	10-Aug
			4-Aug	21-Oct							30-Aug	4-Oct
			21-Oct		30-Aug		23-Oct		13-Sep		4-Oct	
2002	13-Feb	25-Jun		13-Jun		22-Apr		4-Jun		30-Jan		20-Jan
					4-Jun	25-Jun			5-Feb	9-Feb	20-Jan	4-Feb
									11-Mar	28-Jun	5-Feb	5-Feb
											27-Mar	27-Mar
											27-Mar	20-Jun

Table 3. Dates of continuous data for TG stations

	Station Name											
Year	**TG01**		**TG02**		**TG03**		**TG04**		**TG05**		**TG06**	
1997	6-Jun	19-Sep	13-Jun	19-Sep	6-Jun	19-Sep	6-Jun	8-Dec	6-Jun	10-Jun		
	22-Oct		22-Oct		22-Oct				10-Jul	2-Dec	6-Jun	
1998		14-Feb		14-Feb		24-Jan	7-Jan	5-Nov	2-Dec	28-Mar		15-Aug
	17-Feb	29-Mar	17-Feb	9-Nov	18-Feb	4-Aug	6-Nov	6-Nov	1-Apr	23-Nov	2-Sep	20-Dec
	1-Apr	9-Nov					6-Nov	6-Nov				
							9-Nov	31-Dec				
1999							15-Nov		15-Nov		15-Nov	
2000								24-Jan		10-Feb		5-Feb
							17-Feb	12-May	17-Feb	12-May	17-Feb	12-May
											14-Aug	23-Oct
2001									8-Nov		1-Nov	
2002										27-Jun		27-Jun

Continue of Table 3

	Station Name											
Year	**TG07**		**TG08**		**TG09**		**TG10**		**TG11**		**TG12**	
1997	10-Jul	15-Aug			12-Jun	21-Dec			15-Jul	26-Jul		
	10-Sep		12-Jun		21-Dec		15-Dec		1-Aug	3-Dec	18-Jun	
1998		2-Feb		18-Mar		24-Apr		21-Oct	27-Jan	4-Aug		4-Jun
	6-Feb	24-Mar	26-Mar	10-May	24-Apr	11-Oct						
	24-Mar	26-Mar	21-May	31-Dec	21-Oct	31-Dec						
	26-Mar	2-Sep										
	9-Nov	31-Dec										
1999	15-Nov		15-Nov		15-Nov							
2000		10-Feb		10-Feb		10-Feb						
	14-Aug	29-Oct	2-Mar	12-May	2-Mar	1-May						
			14-Aug	27-Oct								
2001	1-Nov	7-Nov	26-Nov									
2002	26-Nov	27-Jun		27-Jun								

Data Errors

Water quality data were collected and compiled in segments. From 1990 to 1998, output from instruments was compiled in a Microsoft Access database file by BNP personnel, however, from 1998 to 2002, only some data were compiled in this way. Upon examination of all data we identified the following data quality problems:

(1) The day February 29 was missing in leap years (1992, 1996, 2000). This problem was most likely caused by errant copying the date column of a non-leap year files. As a result, many data collectd after March 1 are designated to the wrong date. To correct this problem, we referred to the original datasonde outputs made the necessary changes.

(2) Data were often placed beneath the wrong column headings. This likely occurred because over the years different intruments were used that had different sensor configurations. For example, in November 1994 at station AR, pH values were stored under the conductivity column heading and conductivity values were stored under the salinity column heading. These problems were corrected.

(3) Records were often found to be incomplete. Where possible, through line-by-line checking, we filled data gaps by referring to original datasonde outputs.

(4) Breaks in the time series were common. While hourly or half-hourly measurements were the goal, thsi was not possible during all years or at all sites.

(5) Duplicate records were common. For many sites, the same records were repeated. We corrected this problem.

(6) Missing salinity values were often encountered. We were able to generate salinity values when both temperature and conductivity variables were measured using the algorithm described in the Methods section.

Instrument-Environment Effects on Data Quality

While many variables were measured over the years, only temperature and conductivity were measured consistently. In general, temperature data are reliable when recorded. Here, we focus on the quality of salinity data and identify possible instrument problems that resulted in erroneous or missing salinity records. First, we plotted both temperature and salinity time series for each year and each site to facilitate visual inspection of data quality. All such plots can be found in the CD included in this report. Upon identification of potential problem data points, we then searched for probable cause. Problem types and their likely causes are listed as below.

(1) Unusually low salinity was often observed at offshore sites. Such observations prompted us to refer back to the original instrument output files to reveal when was the instrument unit rotated. Figure 2 shows an example of a low-salinity series at site "AR" located on the reef tract. We conclude that the salinity drop was likely caused by probe fouling in July and August of 1998. We reached this conclusion because salinity values to immediately return to those typical of this area after the instrument rotation.

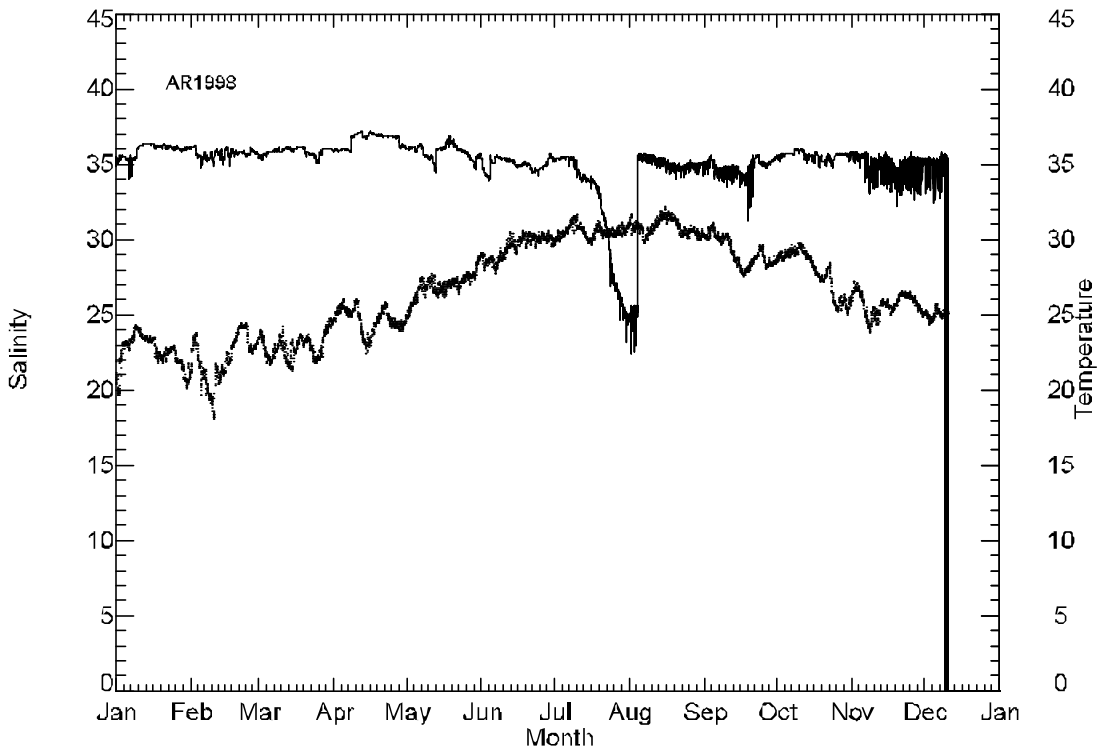

Figure 2. Example of unsual low salinity caused by bio-fouling at AR site 1998.

(2) On occasion, unusually high salinities were recorded at offshore sites. Hypersaline conditions can occur frequently Biscayne Bay waters during the dry season. But it is unlikely that hypersaline conditions would persist at offshore sites, especially when bay salinities are generally low. Figure 3 shows an example of unusually high salinity at the TR site in September 1995. Since all these unusually high salinities occured during the entire instrument rotation period and the rotations before and after all have what appear to be normal salinities, we conclude that this was a result of mis-calibration of the conductivity probe in the laboratory.

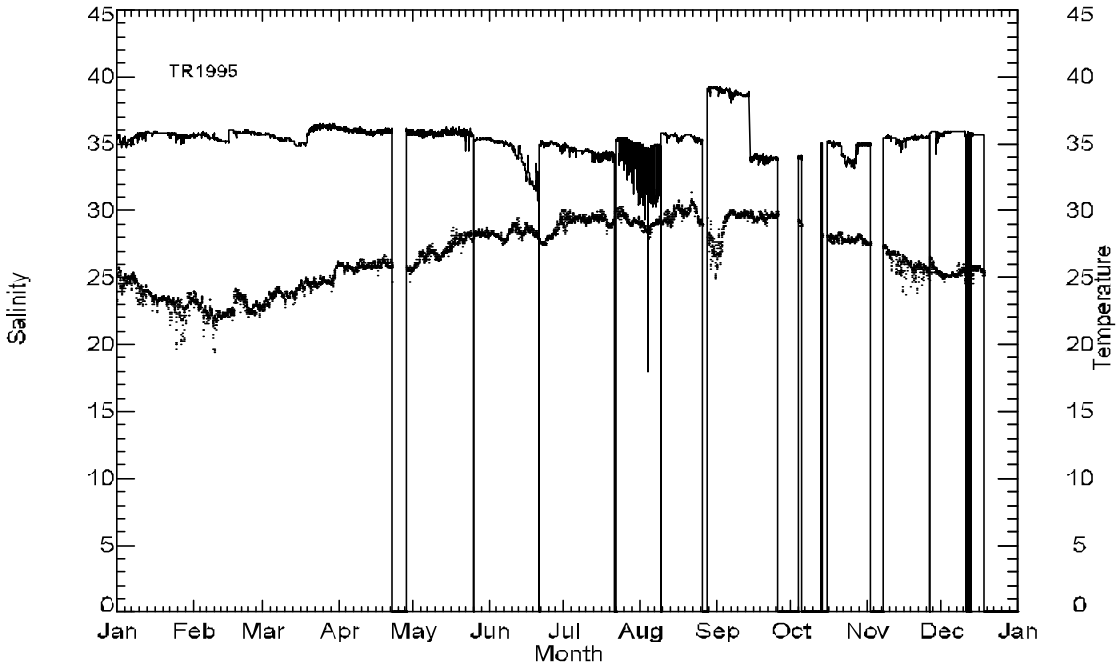

Figure 3. Example of unusually high salinity at TR site in 1995

(3) Unusual pattens of salinity oscilation. We found regular 24-hour salinity oscilations at almost all sites. The tidal cycle for the region is 12.4 hours (semi-diurnal), and this can be verified with the depth data recorded with the same instrument **(Figure 4**). We suspect that this 24-hour period of salinity oscilation was a result of fouling or a bad instrument unit because the oscilation usually disappeared upon new instrument rotation (**Figure 4**).

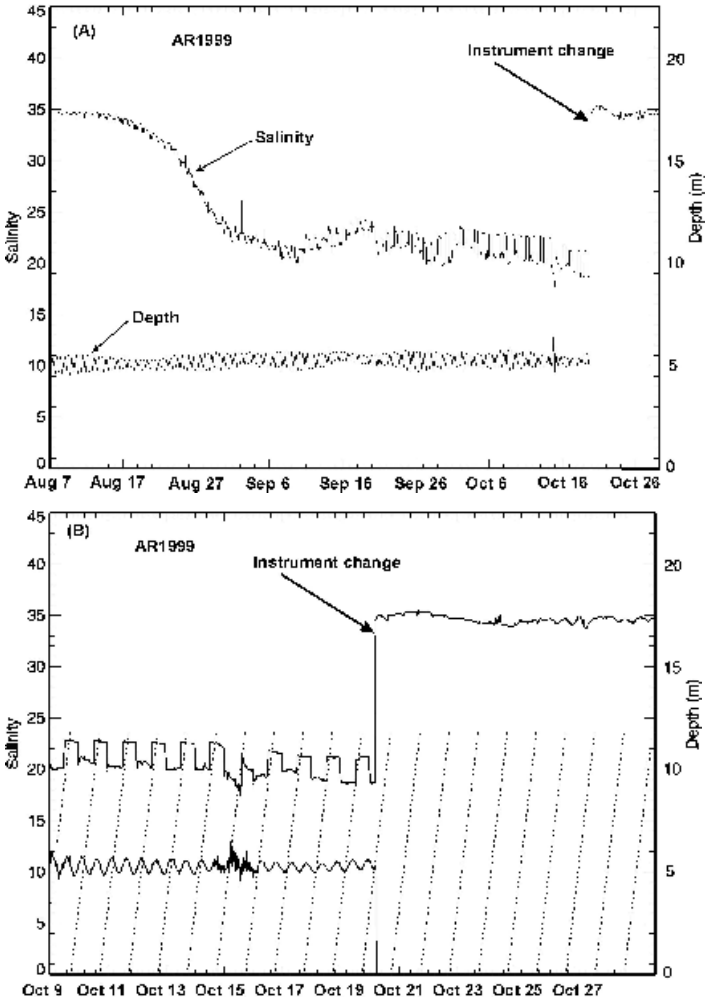

Figure 4. (A) An exmaple of unusaul salinity osilation at AR site from Aug to Oct of 1999. (B) Zoom in of (A) on Oct 20th 1999. The dotted lines indicate the hours of the day (0 to 24).

(4) Data skipping was another problem identified here. For example, during the compilation of the data from 1999-2002, we found that some data outputs from the instrument have skipped records. For example, the intent was to record data every 30 minutes interval, but often records were skipped (see Table 4). Often the message "Internal Error" was associated with record skipping and/or time shifts **(Table 4)** -- we uncertain of the meaning and cause of this message.

(5) Instrument time-shifting was also observed. For example, a half-hourly time series began with readings at 00:00, 00:30, 01:00, 01:30, 02:00 hrs, etc. However, in the middle of the rotation, the recording time inexplicably changed to 02:02:43, 02:32:43, 03:32:43. (**Table 4**, raw #16).

Table 4. Examples of instrument data output for BB site October 2001.

	Date	Time	Temp	Cond	Salin	DO	Depth
1	10/09/01	18:30:00	28.37	51.61	33.83	5.122	11.0
2	10/09/01	19:00:00	28.34	51.59	33.81	5.187	11.0
3	10/09/01	19:30:00	28.31	51.56	33.80	5.026	10.8
4	10/09/01	20:00:00	28.29	51.55	33.79	5.112	10.9
5	10/09/01	20:30:00	28.28	51.52	33.74	5.271	11.1
6	10/09/01	21:00:00	28.26	51.51	33.74	5.239	11.0
7	10/09/01	21:30:00	28.26	51.49	33.73	5.204	10.8
8	10/09/01	22:00:00 - Internal Error.					
9	10/09/01	22:30:00	28.20	51.66	33.86	5.332	11.0
10	10/09/01	23:00:00	28.16	51.73	33.93	5.387	11.2
11	10/09/01	23:30:00	28.06	51.88	34.02	5.445	10.9
12	10/10/01	00:00:00	28.04	52.02	34.14	5.510	11.1
13	10/10/01	00:30:00	28.02	52.02	34.14	5.481	11.1
14	10/10/01	01:00:00	28.00	52.09	34.18	5.580	11.1
15	10/10/01	01:30:00	27.99	52.12	34.20	5.620	10.9
16	10/10/01	02:02:43	27.96	51.94	34.10	5.524	10.8
17	10/10/01	02:32:43	27.96	51.96	34.10	5.471	10.7
18	10/10/01	03:02:43	27.98	51.93	34.09	5.426	10.8
19	10/10/01	03:32:43	27.96	51.80	33.98	5.509	10.9
20	10/10/01	04:03:17	27.94	51.69	33.89	5.322	10.7
21	10/10/01	04:33:17	27.92	51.62	33.84	5.342	11.0
22	10/10/01	05:03:17	27.90	51.53	33.79	5.288	11.1
23	10/10/01	05:33:17	27.88	51.45	33.72	5.201	11.0
24	10/10/01	06:03:17	27.86	51.34	33.65	5.120	10.7
25	10/10/01	06:33:17	27.82	51.33	33.65	5.105	10.7
26	10/10/01	07:03:17	27.80	51.31	33.63	5.049	10.7
27	10/10/01	07:33:17	27.76	51.25	33.57	4.935	10.7
28	10/10/01	08:03:17	27.82	51.21	33.54	5.037	10.9
29	10/10/01	08:51:14	27.89	51.21	33.54	5.079	11.0
30	10/10/01	09:21:14	27.89	51.17	33.52	5.082	11.1
31	10/10/01	09:51:14	27.89	51.17	33.52	5.165	11.0
32	10/10/01	10:21:14	27.87	51.28	33.59	5.261	11.1
33	10/10/01	10:51:14	27.89	51.31	33.63	5.308	10.9
34	10/10/01	11:21:14	27.90	51.36	33.66	5.424	11.0

Site-specific Data Quality

We have evaluated data for quality for all the sites and all years. They are summarized in **Table 5.** In general, data collected prior to 1997 appear to be inferior to data collected from 1997 forward. Below we provide some statistics as to data quality for selected sites.

At the AR site, some data were collected from 1990 to 2002, but the percentage of time that instruments were actually deployed/operating ranged from 21% in 1990 to 94% in 1998. The percentage of reliable (i.e., quality data) ranged from 21% to 100% with an average of 76%. For

At the BB site, data were collected from 1990 to 2002. The percentage of time when covered per year ranged from 26% in 1990 to 92% in 1994. Of the available data the percentage of reliable data ranged from 31% to 100% with an average of 68%.

A the BS site, data were collected from 1995 to 2002. The percentage of time covered ranged from 16% in 2000 to 96% in 1996. Of the available data the percentage of reliable data ranged from 44% to 100% with an average of 81%.

For CR site, data were collected from 1990 to 2002. The percentage of time covered ranged from 23% in 2001 to 100% in 1998. Of the available data, the percentage of relaible data ranged from 70% to 100% with an average of 91%.

For TP site, data were collected from 1993 to 2002. The percentage of time covered ranged from 10% in 1993 to 96% in 1998. Of the available data, the percentage of relaible data ranged from 88% to 100% with an average of 98%.

For TR site, data were collected from 1990 to 2002. The percentage of time covered ranged from 19% in 1990 to 87% in 1995. Of the available data, the percentage of reliable data ranged from 56% to 100% with an average of 85%.

For all the TG sites (TG01 to TG12), data collections started in 1997, some sites were terminated in 1998, and some sites dropped, but then restarted. However, data obtained at most of the TG sites are of good quality (**Table 5**). At first comparison of offshore and inshore site salinity records, it appears that the inshore sites had less fouling problems than the offshore sites. However, this may not be the case because at inshore sites, we cannot easily distinguish "real" salinity drops from those caused by fouling.

Table 5. Evaluation of Salinity data qaulity. Record is the # of data record, % total is the percent of time have data, Min is the minimum salinity, Max is the maximum salinity, Mean is the mean salinity,Std is the standard deviation, %good is percent of good from available data.

	Record	% total	Min	Max	Mean	Std	%good	Quality Status
AR			This is an off shore station				**76%**	Any salinity value over 37 are questionable for this site
1990	1818	21%	33.25	38.47	37.43	0.86	21%	most of the salinity values are over 37
1991	6736	77%	27.37	38.09	35.61	1.92	68%	bio-fouling on instrument, Jul, Aug, Sep
1992	5950	68%	28.95	38.69	35.6	1.12	63%	bio-fouling on instrument, May, Jul, Aug
1993	7026	80%	27.37	37.79	34.68	1.74	69%	bio-fouling on instrument, Feb, Mar, Apr, May
1994	6825	78%	27.23	36.74	34.65	1.38	79%	bio-fouling on instrument, Jun, Jul
1995	7169	82%	32.71	42.22	35.93	1.9	85%	high values (>38) in Sep, and Dec
1996	12845	87%	32.22	36.98	36.19	0.48	95%	bio-fouling on instrument, Jun
1997	7046	80%	33.88	37.21	35.85	0.46	100%	good data
1998	8224	94%	22.39	37.22	35.08	1.93	91%	bio-fouling on instrument, Jul
1999	12845	73%	15.52	36.72	31.07	6.12	54%	bio-fouling on instrument, Apr, May,Jun, Aug,Sep,Oct
2000	9594	55%	12.57	40.35	31.84	5.92	62%	bio-fouling on instrument, Aug, Sep
2001	3841	22%	31.12	35.65	34.83	0.63	100%	good data
2002	6336	75%	32.17	35.86	34.92	0.29	100%	good data
BB			This is an off shore station				**68%**	Any salinity value over 37 are questionable for this site
1990	2273	26%	33.91	39.07	37.29	0.92	36%	high values (>38) in Apr, bio-fouling in Dec
1991	6148	70%	27.44	39.52	36.05	1.87	64%	high value (>38) in Feb, Mar, bio-fouling in Jun
1992	6277	71%	27.59	42.88	35.46	0.9	65%	bio-fouling on instrument, Feb, Mar, Aug
1993	6694	76%	30.69	40.28	35.5	1.37	45%	bio-fouling on instrument, Jan, Apr, May, Jun
1994	8067	92%	23.62	36.59	34.77	1.12	55%	bio-fouling, Jan, Sep, Oct, Nov, Dec
1995	7316	84%	24.67	40.14	35.07	2.42	31%	bio-fouling, Jan, Feb, Mar, Jun, Jul, Aug
1996	6580	75%	23.69	38.4	36.2	1.23	78%	bio-fouling, May
1997	7904	90%	34.76	36.74	35.85	0.4	100%	good data
1998	7179	82%	18.79	36.57	34.1	3.58	80%	bio-fouling, Jul, Sep
1999	11462	65%	29.57	36.63	35.03	1.27	74%	bio-fouling, Apr
2000	9688	55%	22.18	37.11	34.98	2.33	85%	bio-fouling, Sep
2001	8837	50%	28.98	36.82	34.98	0.95	83%	bio-fouling, Jan
2002	7872	100%	32.01	36.75	35.39	0.61	83%	bio-fouling, Jan
BS		This is an off shore station					**81%**	Any salinity value over 37 are questionable for this site
1995	5942	68%	30.33	39.26	36.47	1.36	75%	bio-fouling, Jul
1996	8406	96%	30.07	37.05	35.82	0.88	83%	bio-fouling, Jun, Oct
1997	7915	90%	33.3	37.02	35.88	0.6	100%	good data
1998	16721	95%	29.85	36.87	35.13	0.97	91%	bio-fouling, Mar
1999	7840	45%	24.7	37.41	33.31	3.13	44%	bio-fouling, Jun,Sep
2000	2859	16%	35.43	37.43	36.37	0.48	100%	good data
2001	7172	41%	5.53	37.5	32.08	6.17	80%	bio-fouling, Aug
2002	6364	76%	31.07	35.75	34.59	0.69	78%	bio-fouling, Jan
CR		Caesar's Creek					**91%**	
1990	2491	28%	34.65	41.04	37.29	1.33	70%	high value (>40) in Dec
1991	5592	64%	27.95	41.2	36.67	1.92	74%	high value (>40) in Jan
1992	6160	70%	28.66	43.8	35.36	1.5	88%	a few high values
1993	6965	80%	28.66	39.75	35.08	2.38	100%	good data
1994	6507	74%	28.91	37.71	34.5	1.72	100%	good data
1995	7555	86%	22.78	36.66	33.19	1.88	100%	good data
1996	8261	94%	28	39.12	35.5	1.53	82%	bio-fouling Jun, Jul
1997	14176	81%	29.54	39.59	35.42	1.57	100%	good data
1998	17520	100%	24.63	38.29	35	1.66	83%	bio-fouling May, Sep
1999	13390	76%	28.36	39.77	36.07	1.33	89%	bio-fouling Apr
2000	5493	31%	30.73	39.89	36.01	1.49	100%	good data
2001	4028	23%	22.62	36.06	32.37	2.35	100%	good data
2002	7410	100%	30.44	39.39	35.43	1.47	100%	good data
TP		Turkey Point					**98%**	
1993	893	10%	29.46	33.55	32	0.63	100%	good data
1994	6195	71%	20.83	37.26	32.09	3.86	100%	good data
1995	7720	88%	19.8	35.74	29.16	2.88	100%	good data
1996	7784	89%	23.09	39.37	33.5	2.94	100%	good data
1997	7834	89%	25.42	38.68	32.56	3.06	91%	mostly good, problem in May
1998	16736	96%	23.04	38.14	32.24	3.86	100%	good data
1999	12815	73%	22.11	40.45	34.86	4.08	100%	good data
2000	3754	21%	28.96	35.34	32.42	1.29	100%	good data
2001	6363	36%	10.58	35.47	29.84	3.59	88%	mostly good, problem in Oct
2002	6841	80%	29.38	39.71	35.06	3.25	100%	good data

Table 5 continued.

TR			This is an off shore station				**85%**	Any salinity value over 37 are questionable for this site
1990	1629	19%	35.25	38.69	36.75	0.96	56%	high value Sep
1991	4822	55%	31.49	36.59	35.09	0.67	100%	good data
1992	6300	72%	28.81	38.09	35.62	0.67	100%	good data
1993	6179	71%	26.37	40.51	35.02	1.33	94%	mostly good
1994	7180	82%	28.23	37.83	35.44	1.03	90%	mostly good
1995	7638	87%	18.03	39.29	35.38	1.24	86%	bio-fouling Aug, high value Sep
1996	6389	73%	32.54	40.11	36.52	1	89%	bio-fouling Aug, high value Sep
1997	14094	80%	34.25	37.13	36.05	0.49	100%	good data
1998	13722	78%	30.16	36.93	34.96	1.25	79%	bio-fouling May, Jul
1999	7526	43%	31.33	36.66	35.86	0.87	81%	bio-fouling Sep
2000	3940	22%	32.11	36.74	35.56	0.97	62%	bio-fouling Jul
2001	6212	35%	25.99	36.3	34.93	1.2	76%	bio-fouling Oct
2002	5762	71%	32.37	35.34	34.7	0.44	88%	bio-fouling Apr
TG01	Manatee Bay						**95%**	
1997	8424	48%	18.46	34.21	23.06	2.22	100%	good data
1998	14730	84%	16.23	32.38	24.23	3.41	90%	bio-fouling Sep
TG02	Card Sound Bridge							
1997	8086	46%	20.24	30.71	25.24	1.66	100%	good data
1998	14876	85%	18.99	35.3	26.74	3.38	100%	good data
TG03	Angelfish Creek							
1997	8425	48%	26	39.65	33.84	2.08	100%	good data
1998	9141	52%	25.59	37.53	34.15	1.91	100%	good data
TG04	Military Canal Out							
1997	8873	51%	25.9	36.98	31.4	1.64	100%	good data
1998	17045	97%	22.52	38.33	32.79	3.74	100%	good data
1999	2218	13%	27.07	32.89	30.16	1.08	100%	good data
2000	5229	30%	29.11	35.1	32.39	1.21	100%	good data
TG05	Military Canal close							
1997	8546	49%	9.24	29.12	21.52	3.54	100%	good data
1998	15490	88%	7.52	37.58	26.16	6.96	100%	good data
1999	2216	13%	16.57	23.88	21.21	1.39	100%	good data
2000	6028	34%	12.99	34.71	25.23	4.99	100%	good data
2001	2561	15%	3.52	26.35	20.57	3.28	100%	good data
2002	8520	100%	12.24	41.15	27.89	6.49	100%	good data
TG06	Black Point							
1997	10002	57%	6.31	28.81	21.64	3.71	100%	good data
1998	16105	92%	7.79	37.13	24.93	6.45	100%	good data
1999	2216	13%	16.35	27.19	22.77	2.3	100%	good data
2000	9144	52%	12.26	35.9	27.02	3.35	100%	good data
2001	2904	17%	11.86	29.18	22.97	2.56	100%	good data
2002	8519	100%	8.95	41.21	27.88	6.11	100%	good data
TG07	Snapper Creek							
1997	7130	41%	18.35	33.39	27.34	2.77	100%	good data
1998	14054	80%	19.87	37.14	29.78	4.16	100%	good data
1999	2216	13%	24.02	29.52	26.6	1.21	100%	good data
2000	5603	32%	21.18	34.54	27.71	2.99	100%	good data
2001	2002	11%	18.61	30.93	27.5	2.92	100%	good data
2002	8518	100%	21.68	39.86	31.65	3.41	100%	good data
TG08	Gables-by-the-Sea						**98%**	
1997	9719	55%	20.95	32.71	28.64	2.45	100%	good data
1998	16620	95%	17.24	37.21	29.79	3.24	87%	bio-fouling Jul
1999	2216	13%	25.13	30.41	28.06	1.04	100%	good data
2000	8847	50%	25.1	35.84	31.76	2.22	100%	good data
2001	1703	10%	25.28	30.98	28.3	1.34	100%	good data
2002	8517	100%	25.46	39.86	32.38	2.82	100%	good data
TG09	Safety Valve							
1997	9715	55%	27.7	36.62	33.63	1.43	100%	good data
1998	17051	97%	25.95	36.9	33.97	1.95	100%	good data
1999	2215	13%	30.13	35.99	33.62	1.19	100%	good data
2000	4852	28%	31.29	37.11	34.97	1.46	100%	good data
TG10	Miami River							
1997	790	5%	17.75	35.22	29.55	3.16	100%	good data
1998	14089	80%	13.6	37.21	29.38	3.99	100%	good data
TG11	Julia Tuttle Csway						**94%**	
1997	6476	37%	22.18	33.79	29.05	2.37	100%	good data
1998	9067	52%	20.17	36.3	30.58	2.69	87%	bio-fouling Jun
TG12	Broad Csway							
1997	9427	54%	27.03	36.37	33.55	1.52	100%	good data
1998	7430	42%	26.3	35.99	33	1.73	100%	good data

Time Series Analysis

Before conducting a preliminary time series analysis, it was necessary to first determine the shortest common sampling interval. This was the hourly sampling interval. Next, we plotted each series for visual inspection. If a time series contains an obvious trend, then such variation were removed from the data to obtain a stationary time series before performing the spectrum analysis. Fast Fourier Transform (FFT) was used to transform the stationary time series data into its frequency domain spectrum. A periodogram was then used to interpret the results (**Figure 5**). From this periodogram, we can see a strong semi-diurnal signal indicated by the peak at about 2 cycles per day (1.93 cycles per day). The small peak around 4 is the second harmonic of the semi-diurnal tide. Further analysis of other data will be continued in future projects.

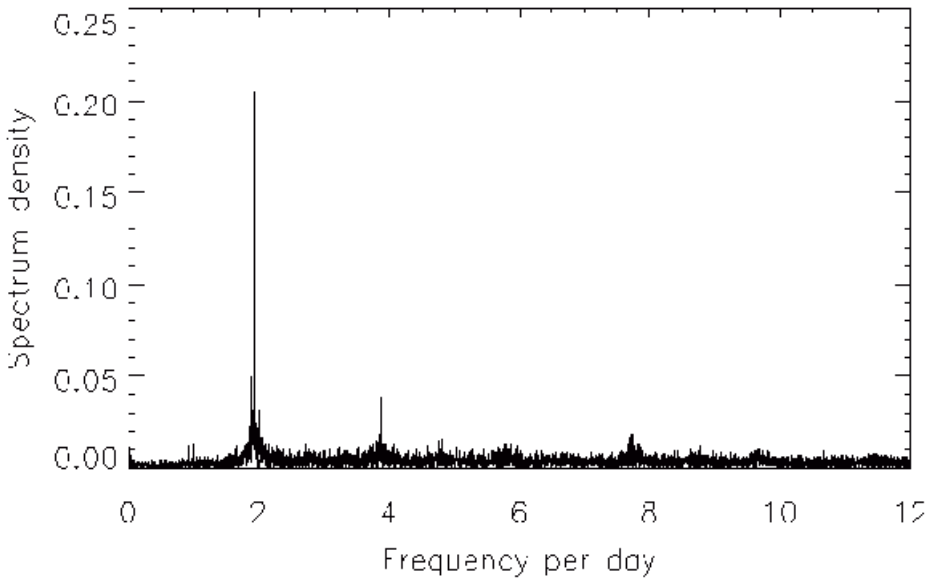

Figure 5. An example of a periodogram of salinity data from a site in Biscayne Bay after the removal of first order autoregression. This indicates a strong spectrum density at a frequency of about 2 cycles per day (1.93 per day), which is the signature of the semidiurnal tide. The second peak is the second harmonic of the tide.

In time domain analyses, we found strong cross-correlations between salinity and canal discharges from 1 to 100 hours lag time (**Figure 6**). The semi-diurnal tidal cycles were also clear. Perhaps the most important result was that effect of canal discharge on salinity at this site (TG06) was appreciable for up to 100 hrs (4 days) after which it gradually declined from 100 to 400 hrs (about 16 days). This indicates that canal discharges can affect the salinity of this site for as long as16 days, with strongest effects during the first 4 days. Further analyses of this type will be continued in our next project.

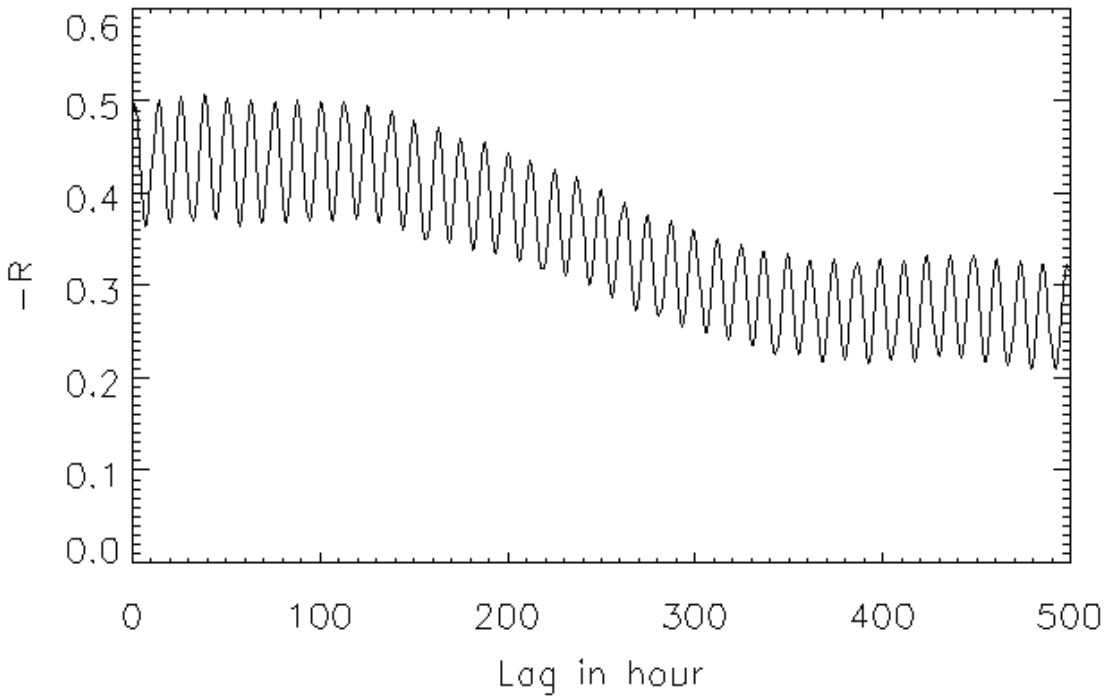

Figure 6. Cross-correlation between salinity and canal discharge. The y-axis is the negative of the correlation coefficients. The x-axis is the lagged hours before the observation.

Multiple regression modeling was also performed to estimate salinities from canal discharges and wind velocities (**Figure 7**). To predict the average daily salinity at site TG06, we used the average flow rates of previous 15 days from two nearby structures (S21 and S21A), wind speed from Fowey Rocks. As one can see, the simple regression model predicted the general trend of the average salinity, but not all of the peaks and troughs (see red line in Figure 7). Further statistical modeling of this type will be conducted in our next project.

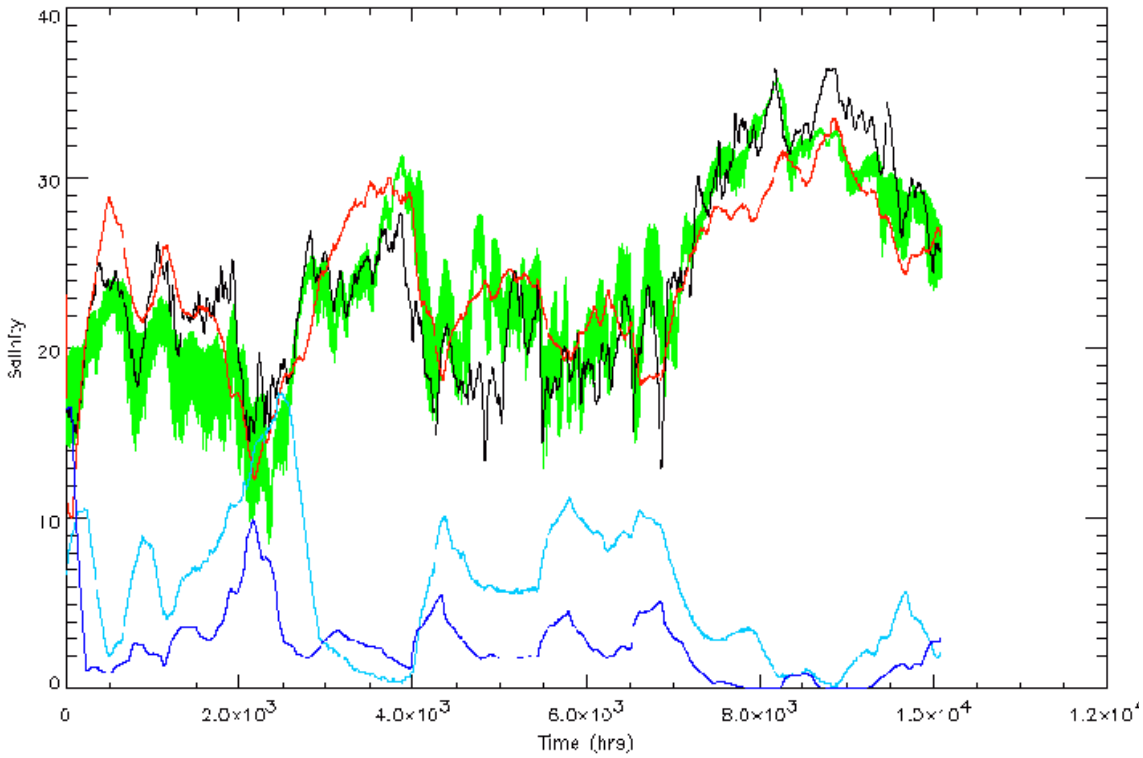

Figure 7. A preliminary statistical regression model predicting the average daily salinity from fresh water discharges of two canals, and the x, y components of wind vectors. The black line the average daily salinity, the red line is the regression model estimate, the green line is the estimate from the Wang et al. hydrodynamic model, the light blue line is the average flow rate of previous 15 days for canal A, and the dark blue line is the average flow rate of previous 15 days for canal B.

CONCLUSIONS AND RECOMMENDATIONS

It is important to note that 11 months passed before we finally obtained all original data; this hindered the timely completion of this report. In this study, we found many problems regarding the quantity and quality of the BNP water quality data. We exposed and in many cases corrected errors that were made during data compilation process and we also found considerable data gaps in the data time series which, therefore, limit the extent to which time series analyses could be conducted here and in the future. In fact, there was only on case where a data series spanned uninterrupted for longer than one year. Probe fouling and mis-calibration appear to be the major causes of poor data quality, these are especially apparent prior to the 1997. Of the available data, the percentages of reliable data ranged from 68% to 91%.

Most of the data, instrument and human problems identified here are to be expected given the size, complexity and nature of the task at hand. We offer the recommendations below to minimize problems in the future.

(1) To reduce data compilation errors, we suggest annual data compilation be passed to a second party to test data readability and to generate data plots for visual inspection for ouliers and other problems.

(2) To reduce data (time series) gaps, we suggest use of two salinity measuring units at the same site with different, overlapping rotation schedules. In addition to serving a "back-up" in case of unit failure, it also provides opportunities for cross-validation of salinity readings.

(3) Mis-calibration problems appear to have occurred mostly prior to 1997. Whatever training and/or personnel changes that occurred from 1997 forward should be recognized and retained, if possible.

(4) Our examination of data strongly implicates that BNP may still be deploying one or two bad units. An effort to identify such unit(s) is recommended .

(5) Probe fouling appears to be the major cause of poor data quality. Unless instruments can be treated with anti-fouling agents, we suggest frequent probe cleaning within the instrument unit rotation period and/or the reduction of unit deployment periods. Data indicated that fouling can occur in as little as two weeks. If such a short deployment period is impractical, we recommend a 4-week deployment period with in situ cleaning every 2 weeks.

REFERENCES

Serafy, J.E., J.S. Ault, P. Ortner and R. Curry (2001). Coupling Biscayne Bay's natural resources and fisheries to environmental quality and freshwater inflow management. In: Biscayne Bay Partnership Initiative Survey Team Final Reports. Pp. 163-174.

Swain, E.D., A. Kapadia, S. Kone, E. Damisse, D. Mtundu, and G.N. Tillis. 1997. Determing discharge-Coefficient ratings for Coastal Structures in Dade County, Florida, U.S. Geological Survey, Water-Resources Investigations Report 97-4079.

Wang, J.D., Cofer-Shabica, S.V., and J.Chin Fatt. 1988. Finite element characteristic adevction model, J. Hyd. Eng. 114(9):1098-1114.

Wang, J. D., Luo, J., Ault, J.S. (in review). Flows, salinity and some implications on larval transport in South Biscayne Bay, Florida. Bulletin of Marine Science.