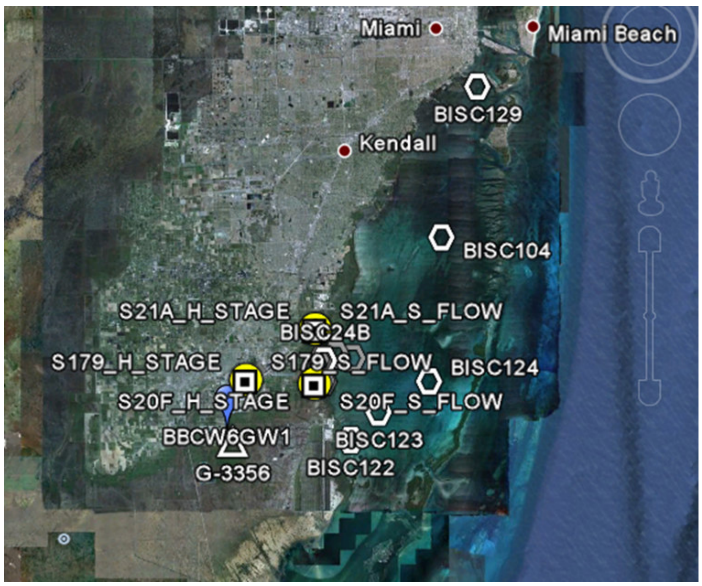
South Miami-Dade Statistical Data Analyses

PREPARED FOR

South Florida Water Management District
3301 Gun Club Road
West Palm Beach, Florida 33410



PREPARED BY

Eric P. Smith
906 Allendale Court
Blacksburg, VA 24060

DECEMBER 30, 2010

Purchase Order Number: P4500052127

Summary

This report describes the compilation of various clean data sets into a single file containing daily observations of flow, stage, well depth and water quality at numerous locations in south Florida. The data were used to compute correlations between the various types of measurements in order to assess the degree of association between the sets of measurements. Correlations were computed for the full set of observations as well as for the wet and for the dry seasons. In addition to calculations based on daily values, correlations using monthly mean values were also computed. Both parametric and nonparametric correlations were used to summarize the associations. An example using the data to explore relationships between flow at three stations and well, stage and salinity at neighboring stations is presented.

Table of Contents

List of Tables

List of Figures

1. Introduction and Background

The South Florida Water Management District (District) has provided clean data sets for South Miami-Dade County projects on hydrology, water quality, geographic and regulatory data. The objective of this project is to perform an exploratory statistical analysis of the cleaned data. The data has been collected over different time periods by different agencies and covers a period from 1/1/2000 through 12/31/2009. Most of the data are at a daily scale although some is a more frequent interval and others are at an irregular interval.

This document describes exploratory analyses that were carried out and summarizes the output. Due to the size of the files that were produced, it is not possible to provide printed copies of the output; these were saved to disk. Some summaries are provided in this document as well as descriptions of the output. These include correlations that measure the association between the time series of measurements. The correlations were calculated for measurements on well, flow, stage, rain and water quality. Correlations for daily data as well as monthly averages were calculated. Graphical displays and multivariate methods were used to understand some of the main attributes of the data. Correlation tables and summaries were saved in files as they are too large to produce as printed output. Some summaries are presented below to help understand the analyses that were undertaken. An example exploring the relationship between flow at S20F and well, stage and salinity at several sites is presented to illustrate usage of the data.

2. Scope of work

The following six analyses were carried out:

1. Calculation of the correlation between stations of the same data types such as groundwater to groundwater.
2. Calculation of the correlations between stations of different types.
3. Conduct evaluations of the data sets to investigate importance of lags for correlations between variables of the same type or different types
4. Repeat the correlations between the variables of the same type during wet and dry seasons
5. Repeat the correlations between the variables of different types during wet and dry seasons
6. Evaluate relationships at a given station as a function of neighboring stations

In addition, additional analyses were made to try to identify patterns in the sites based on correlations. These include

1. Factor analysis and principal components analysis of the correlation matrix (Spearman) to identify groups of sites that are related in pattern. This was done for different types.
2. Where possible, create a biplot displays of sites and times
3. Where possible, use regression analysis to identify possible relationships between variables.

The first step in the analysis was to compile the data into a single file. Then SAS was used to perform the analyses associated with the above tasks.

3. Compilation of data sets

The first step in the analysis was the compilation of data sets. The initial data files were converted to .csv files. These files were created using consistent formats for the dates, the names of the variables and the period of the collection. Note that the word "variable" will be used to refer to a measurement at a site of a particular type. In some cases it will refer to a derived value such as the mean of several measurements.

After generating the *.csv* files, these were imported into SAS. A problem in combining the data files into one large file was noted. In some cases, a measurement was made at a particular site for a certain period of time in one study and for a second period of time in a different study. These sites were identified and the data combined into one site/variable. Following this data reduction, a "wide" file was created. The wide file is an array of times by variables. The wide data file consists of over 1000 variables measured at 3652 times. Variable names were created by combining the station name with the type of measurement. For example, S_123_Rain is the rain measurement from station S_123. Stations that included the measurement type were not altered (e.g. ANGEL_WELL_km). A master list of variables was created to use in subsequent analyses.

Because the time series for the water quality data were irregular and sparse, some of the analyses were found to be non-informative. A second data file was created in which all of the data collected for a month were averaged. This resulted in a wide data file of 120 monthly observations. Some of the water quality measurements (salinity, temperature and conductivity) were made at different depths for the same time and station. The value at the greatest depth (referred to as the bottom) was used in these cases.

A summary of the data sets including the quality codes and their frequency as well as summary information on the measurements at each site such as the mean (averaged over time), median, number of times with observations, start and stop dates of sample in given in Appendix 1. Several of the files had some irregularities and these are described in the Appendix 2. For example, site S-122, a flow station, was zero for all times and was deleted.

4. Analyses

The information below gives details on the analyses that were undertaken. The analyses in general produced rather large summary tables that are too large to reproduce in this document. Therefore, only some of the analyses are reported here. Others are provided as EXCEL spreadsheets (csv), PDFs or Word files (doc or rtf). Rather than produce correlation arrays describing relationships between different types of variables a single large array was calculated for items 2 and 5. This array includes information about the correlations between variables of the same type as well as different types. A summary of these files is given Appendix 3.

For the correlation analyses, three correlations were calculated Pearson, Spearman and Kendall. The traditional correlation measure is the Pearson measure. This measure provides a good summary of similarity in pattern of time series when the data are from a normal distribution and the relationship is linear. However, the measure is sensitive to odd observations and nonlinear relationships. The Spearman correlation is a rank-based measure of association. Ranks of the observations are computed for each time series and the correlation between the ranks is calculated. The Spearman measure does not require the assumption of normality for inference and is not sensitive to odd observations. The Kendall measure also does not require the data to be normal for good results and is not sensitive to odd observations. The Kendall correlation measure is a good measure of association for monotonic relationships between series. All three measures range from -1 to 1 with values close to one indicating strong relationships. Because the data consist of time series of observations, statistical inference (i.e. tests) on the correlations were not considered.

The analyses that were undertaken are described below:

1. Calculation of the correlation between stations of the same data types such as groundwater to groundwater.

   Three types of correlations were computed: Kendall, Spearman and Pearson. The correlations were calculated for each of the five data types (flow, rain, stage, water quality and well) and stored in comma separated files. The resulting files are arrays with columns associated with measurements at each site. The first three rows of the file contain summary statistics including the mean, standard deviation and sample size. The remainder of the file gives the correlations with other variables of the same type. In the case of water quality the variables are salinity, temperature and conductivity.

   The correlation matrices are stored in separate files for each type and correlation measure. For example the file *Rain Kendall.csv* has the Kendall correlation coefficients for all pairs of rain variables. The files are stored in the folder *Report Items 1-6\Item1_correlations.*

2. Calculation of the correlations between stations of different types.

   This analysis was done using the full matrix of observations and by calculating the full correlation matrix. The elements of this matrix are the correlations between a measurement at a particular site and all other measurements at the same site or different sites. Again, three correlation matrices were produced using different types of correlations.

   The correlations are in three *csv* files for the overall data set: *dayAll_kendall.csv, dayAll_spearman.csv and dayAll_pearson.csv*. The name indicates that the data are from daily measurements using all the data with a specified correlation. The files are located in the folder *Report Items 1-6\Item2_correlations.*

3. Conduct evaluations of the data sets to investigate importance of lags for correlations between variables of the same type or different types

   To evaluate lags, a data set was generated with new variables that corresponded to the original measurements plus lags from1 to 5 days (the lag 1 measurement is the measurement from the previous day). For each lag, a prefix was added to the variable name that was L1 through L5 corresponding to the number of days. Correlations were then computed for this matrix of values. This was carried out by writing a SAS macro that looped through the variables and calculated correlations. For each variable that was considered, correlations were evaluated for all other variables, at lags of 1 through 5. The correlation with the original variable (i.e. no lag) was also computed. The 10 greatest (in absolute value) correlations were selected and output into a table. The file *LagCorrs.docx* is a Word file containing the ten greatest (absolute) correlations for each variable. Three tables were computed, for Kendall, Spearman and Pearson type correlations. Because some pairs of variables had small sample sizes, some correlations were equal to 1.0. These were dropped from the table to attempt to give useful results.

   A sample table is given below for the variable stage at station _3B_SE_B_Stage_D (Table 1). The columns are as follows: _TYPE_ is the type of correlation that was calculated, _NAME_ is the name of the variable that was most correlated with the variable of interest. The next column's name is the name of the variable of interest and the correlations are given in the column. The last column gives the absolute value of the correlation. The name of the best variable is the variable name with a prefix added for the lag. Thus, for example, _3B_SE_B_Stage_D refers to stage at station 3B_SE_B. In the table below, _3BS1W1_H_Stage_D (no lag) had the greatest absolute Pearson correlation with the stage data at 3B_SE_B, followed by the well variable at site 3BS1W2_G. The next two variables are also at lag 0. The fifth variable is the stage variable at 3B_SE_B. Note that sample size is not listed and some of the sample sizes could be small. Also, the periods where data are observed will overlap for the variables in the table with the variable of interest however the overlap is potentially small and could be different for different pairs. The correlations should only be interpreted as measures of association between the measurements and not as cause-effects measures.

Table 1. Example of output for lag analysis. Obs refers to the rank of the correlation pair, _TYPE_ is the type of correlation, _NAME_ is the name of the variable that was correlated with the variable of interest, _3BS1W1_H_Stage_D is the variable of interest with the correlation coefficient given below, absc is the absolute value of the correlation.

Obs	_TYPE_	_NAME_	_3B_SE_B_Stage_D	absc
1	Spearman	_3BS1W1_H_Stage_D	0.99948	0.99948
2	Spearman	_3BS1W2_G_WELL_km	0.99928	0.99928
3	Spearman	_3BS1W4_G_WELL_km	0.99889	0.99889
4	Spearman	_3BS1W3_G_WELL_km	0.99886	0.99886
5	Spearman	L1_3B_SE_B_Stage_D	0.99866	0.99866
6	Spearman	L1_3BS1W2_G_WELL_km	0.99813	0.99813
7	Spearman	L1_3BS1W1_H_Stage_D	0.99795	0.99795
8	Spearman	L1_3BS1W3_G_WELL_km	0.99773	0.99773
9	Spearman	L1_3BS1W4_G_WELL_km	0.99750	0.99750
10	Spearman	L2_3B_SE_B_Stage_D	0.99585	0.99585

The complete set of correlations is given in the file *lagcorrrs.doc* (located in the Item 3 folder).

4. Repeat the correlations between the variables of the same type during wet and dry seasons

The basic data set was divided into two data sets based on season. The wet season was defined as the period from May 1 through October 14 and the dry season was from October 15 through April 30. Once the data set was generated, correlations were computed for each period using SAS. Again, three types of correlations were computed. The results are in the folder: *Item 4 Correlations WetDry by Type*. The analyses were run both on the full data and on the monthly mean data. For each type there will be two folders, one for the full data and one for the monthly data (e.g. the folder *flow* has three files for the full data with a correlation matrix for the wet season and one for the dry season in each spreadsheet. *Monthly Flow* is the folder that has similar results using monthly means.)

5. Repeat the correlations between the variables of different types during wet and dry seasons

The correlations for the complete set of variables for item 5 are contained in the files *pearsonWD.csv, kendallWD.csv* and *spearmanWD.csv,* located in the folder *Item 5 Correlations Wet Dry all.* In addition, there is a folder *correlations with Monthly means* that contains the correlations based on monthly data.

6. Evaluate relationships at a given station as a function of neighboring stations

   To evaluate relationships with neighbors the metadata file that was provide with the latitudes and longitudes of the sites was used to calculate distances between sites. Some sites did not have location information and this was obtained through web search. A SAS macro was written to do the following: the wide data set was transposed to be of the form of variables by time. The latitude and longitude was added to the data from each site. For a given site of interest, the distance to other sites was calculated and the seven nearest neighbors were obtained. Correlations were then calculated for each of these variables. This was then repeated for all of the variables.

   The output from the analysis is a table of correlations for each site and is described in Table 2. The columns of the table correspond to the variables from the stations closest to the station of interest. Three rows are given for each table, corresponding to the correlations using Spearman, Kendall and Pearson correlations. The column labeled *ctype* gives the correlation type. The site name is listed in the column labeled as *site_name*. An example using ANGEL_WELL_km is given below.

Table 2. Table of correlations for sites that are closest to a given site. The columns are: *Obs* is the observation number in the file, the next seven columns give the sites that are closest, ctype is the type is correlation and site_name identifies the variable of interest. The type of variable was not included to allow the table to fit in the margins. The correlations are only measures of associated and do not represent measures of causation.

Obs	G_3628	G_3626	G_596	G_596_B	S331W	S331_H	S331_R	ctype	site_name
1	0.92835	0.79754	0.90155	0.89084	0.21266	-0.11627	0.21889	Spearman	ANGEL_WELL_km
2	0.77680	0.62284	0.74742	0.73290	0.15999	-0.08600	0.16631	Kendall	ANGEL_WELL_km
3	0.89712	0.81452	0.89512	0.87026	0.13403	0.06230	0.13006	Pearson	ANGEL_WELL_km

The results of this analysis are in the file *neighbors correlations.pdf* in the folder *Report Items 1-6\Item 6_Correlations Neighbors.* In addition, the monthly data was used to repeat the analysis and these correlations are provided in *neighborsYM.pdf*, located in the same folder. A pdf file was used to allow indexing.

Regression analysis

The data from different stations were also evaluated using regression methods. The exploratory approach that was taken used a stepwise regression procedure to find the top variables (five or fewer) that best explained another variable. When this approach was applied generally, the method often did not produce good results because of sample size issues. When multiple variables are considered in a regression analysis, if one site has a missing value then all the sites are deleted. Thus, for example, if one site was measured in the wet season and another in the dry season then the sample size is zero after missing values are deleted and regression cannot be used. Since some sites only have a few observations, the sample size for the regression was greatly reduced. Given the large number of sites that were considered for a regression, the missing data problem greatly reduces the sample size and can lead to uninformative models or the inability to fit a model. To deal with this issue, sites with smaller sample sizes were eliminated. Data sets were generated using different minimum sample sizes and stepwise regression analyses were run. For each variable, the model selected the top variables related to the variable of interest. The minimum sample size of 3500 was selected for reported analysis and output was stored in an EXCEL file (*RegsN_3500* in the *regression* folder). A sample of the table using three variables is given below in Table 3.

Table 3. Summary of regression analysis. The two tables below are partial tables from the full EXCEL spreadsheet. *_DEPVAR_*is the name of the variable that is analyzes, *_RMSE_* is the square root of the mean square error, *Intercept* gives the value of the intercept in the model independents is the number of variables selected. Next is the list of variables that could be or are in the model. In the second table,*_IN_* is the number of variables in the model, *_P_* is the number of parameters in the regression model, _E*Df_* is the degrees of freedom for the model, *_RSQ_* is the r-square for the model and _AIC_ is the Akaike's Information Criterion. In the table a blank indicates the variable was not selected, otherwise the coefficient is given. The regressions represent possible relationships and should not be interpreted as causal relationships.

DEPVAR	_RMSE_	Intercept	S179_H_Stage_J	S20F_H_Stage_J	S121_C_Flow_SK
S179_H_Stage_J	0.05425	0.07560	-1		-0.003703435
S20F_H_Stage_J	0.03370	-0.03821	0.424981417	-1	
S20F_S_flow	100.066	25.72			

DEPVAR	_IN_	_P_	_EDF_	_RSQ_	_AIC_
S179_H_Stage_J	5	6	3636	0.986853176	-21220.36263
S20F_H_Stage_J	5	6	3636	0.986758797	-24686.99846
S20F_S_flow	5	6	3636	0.853982852	33554.91974

The table lists the dependent variable, the root mean square error of the model *(_RMSE_* or standard deviation), the number of independent variables in the final model, *p* is the number of parameters, *df* the degrees of freedom of the model, *Rsq* is the R^2 of the model. The rest of the table is the list of the parameter values. The remaining columns contain all the variables that could potentially have been in the model. Most of the entries are blank indicating the variable is not in the model. If there is a -1 then this indicates the dependent variable. For example, the first row describes the model for stage at station S179_H. The model includes 5 variables and resulted in an R^2 of 0.98, the intercept is 0.075 and the variable S121_C_Flow_SK has a coefficient of -0.0037. S20F_H_Stage_J has no coefficient, indicating that it is not important, given the other variables in the model. The other variables that are not in the model are not listed in the above table because of space. The full set of models is given in the *csv* table *RegsN_3650A*. The full model for flow at S179_H is:

S179_H_Stage_J = 0.076 - 0.0037 S121_C_Flow_SK - 0.00015 S179_S_flow - 0.0003 S21A_S_flow - 0.081 S165_T_Stage_D + 1.051 S166_T_Stage_D.

An overall summary of the files that were generated is described in Table 4.

Table 4. Table of deliverables. This summarizes the files that are included in the report but not printed due to size.

Item	File	Summary
Report Item 1-6\Item 1 Correlations by type Correlations with variables of the same type	Five folders, each containing files kendall.csv, spearman.csv pearson.csv. with name of type (e.g. *flow_kendall.csv)*	Parametric and nonparametric correlations
Report Item 1-6\Item2 correlations by type Correlations with variables of the different type	DayAll_kendall.csv, DayAll_spearman.csv DayAll_pearson.csv.	Parametric and nonparametric correlations
Report Item 1-6\Item 3 Correlations Lags Correlations with lags of variables of the same and different type	lagcorrrsYM.pdf. lagcorrs	Parametric and nonparametric correlations with lag variables. Best correlations are listed for lags up to 5 days. YM is for monthly data.
Report Item 1-6\Item4 WetDry by type Correlations between variables of the same type for wet and dry seasons	Folders for daily and Monthly summaries. Each folder has a file for the different correlation type pearsonWDday.csv, kendallWDday.csv spearmanWDday.csv.	Parametric and nonparametric correlations calculated separately for wet and dry seasons
Report Item 1-6\Item5 WetDry All Correlations between all variables for wet and dry seasons	All_pearsonWD.csv, All_kendallWD.csv All_spearmanWD.csv.	Parametric and nonparametric correlations calculated separately for wet and dry seasons
Report Item 1-6\Item 6Correlations Neighbors Calculated correlations with neighboring stations	Neighbors correlations.pdf NeighborsYM.pdf	Parametric and nonparametric correlations for data from stations close to a given station. Correlations for seven closest stations are reported.
Regression models	RegN_3650.csv. Regression models.sas	Regression summaries based on variable selection models.
Heatmaps	Folders with programs and output Heatmap folder contains heatmaps in RTF file or as a GIF file. For each type, there are two	Heatmaps of Spearman correlation matrices SAS programs to obtain the figures

	files, on for daily data and one for monthly	
Factor analysis	Files for each type summarizing factor analyses The general form is type_factor for daily data or typeYM factor for monthly data	Files with scree plots, rotated factor pattern (used to group sites) and a biplot display (where possible).
Data and programs	A folder for data and one for programs	Important data sets and SAS programs used in the analyses

5. Summary of some results

The interpretation of the correlation matrices and other analyses is difficult due to the number of variables in the analysis. In some cases, there are some additional analyses that might help summarize the data and relationships between the variables. Two such analyses are based on principal components analysis/factor analysis and heat maps.

Principal components analysis (PCA) is a useful way to summarize a correlation matrix by reducing the matrix using eigenvector and eigenvalues. It is typically applied to a matrix calculated from a full data set (no missing values) but may also be applied to a matrix of calculated correlations. In the examples below, the Spearman rank correlation matrix was used. The basic idea behind PCA is to form a combination or weighted average (called a *component)* of the variables that best represents the information in the data. If all the variables are closely related then a simple average of the variables would provide a good representation of the data. In general, there are likely to be several of these averages that represent different aspects of the data.

The PCA solution can be rotated to try to group variables together. These groups of variables define what are referred to as “factors” and the analysis is referred to as factor analysis. The analysis requires a selection of the number of components prior to rotation. To select the number of factors the plot of the eigenvalues versus the eigenvalue number (scree plot) was produced (see, for example, Figure 1). The number of factors with eigenvalues above one was used as a criterion, selecting fewer factors when eigenvalues were close to one. Variables or sites important to a factor were selected based on the correlations between the site measurements and the factors. These eigenvectors from the PCA were rotated using a varimax rotation to group together variables and correlations with the rotated factors computed. To facilitate summarization, these are multiplied by 100 and rounded, then sorted. Values that were “small” (as determined by SAS) were replaced with blank values to aid in interpretation.

Another way to summarize results is through a biplot display (Figure 2). This graphical display summarizes sites and times in one display. The center of the plot corresponds to the mean of the data. To interpret the display one would look for sites that are close together to represent sites that are similar in measurement pattern and times that are not near the center to identify the times that have more extreme measurements. In addition, the times may be projected onto the sites to give an approximation to the value of measurement at the site (for that time). The program uses the correlation matrix (Pearson) and the SAS program that was used requires the number of usable observations (times) to be larger than the number of variables (sites). While a useful display, interpretations may be difficult when there are missing values.

A third way to summarize the information in the correlations between variables is through a “heat map” (Figure 3). The heat map is a graphical display that tries to summarize relationships by assigning different colors to correlations of different size. In an ideal situation one might have two sets of variables that would yield high within block correlations and low between block correlations. The map attempts to put variables that have similar, large correlations into blocks that have the same colors. In the case of two blocks, this would yield a heat map with two blocks of one color on the diagonal and off-diagonal blocks of a different color.

Note that the PCA and heatmaps are based on Spearman (rank) correlations while the biplot display is based on the Pearson (parametric) correlation. This is because the PCA analysis can be run on a correlation matrix while the biplot requires a data matrix. Because of this, the sample sizes between pairs of variables are critical and are likely to affect interpretations. Variables with missing values are omitted in the biplot display and this fact results in patterns that might be hard to interpret or connect to the PCA. For example, with the rain data, there are several sites that did not record rain during peak rain periods. In these cases, all sites during that period are not included. The result is that only one date appears to be extreme in the plot and the plot is not representative of the entire period.

The above analyses were applied to data that satisfied conditions to allow calculations. Specifically, variables that were constant were omitted, variables that were perfectly correlated with other variables were omitted, and if a value was missing from one site, the time was omitted from all sites. Heat maps are only feasible for smaller sets of variables (under 150 variables).

The plots and table below are for rain and were included because the number of variables are small relative to other types (stage, well and water quality), hence are easy to display. The types stage, well and flow are summarized in separate documents, located in the *heatmap* folder. The files are *well heatmap.doc, flow heatmap.doc and stage heatmap.doc*. These files are for correlation matrices computed using daily data; there are also files for the analysis on a monthly basis. The other graphical displays and summary analyses that are based on PCA/factor analyses are in the folder *factor analysis* and there will be two files for each type (daily and monthly).

The results from the analysis of the correlation matrices generally suggest that while there are a large number of stations associated with each type, there are some sets of sites with a moderate degree of similarity in measurement. This is indicated by the relatively high amount of variance explained by a few components and would be expected as the measurements are over time.

Figure 1. Scree plot and variance explained plot for the rain data. The plot gives the eigenvalues of the correlation matrix versus the number of factors. Look for an "elbow" or values to drop below 1. The variance explained plot displays the proportion of the total variance for each component and the cumulative proportion of the variance explained.

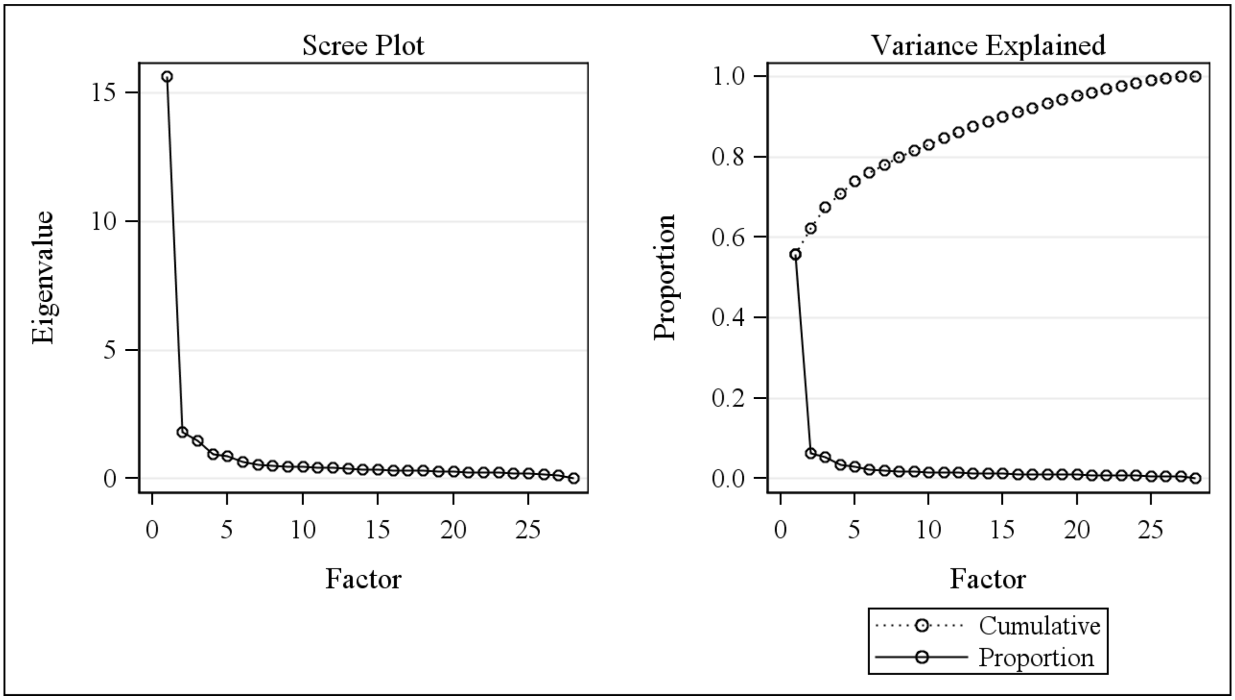


Table 5. Sets of variables identified by the rotated factor pattern for the rain data. Variables with similar values flagged with a star are put into a group. The numerical values are correlations between the factor and the variable and have been multiplied by 100, rounded and reported as two digits. The cutoff value is determined based on the root mean square error of the loadings.

Rotated Factor Pattern						
	Factor1		Factor2		Factor3	
S20G_Rain	77	*	27		23	
S20F_Rain	77	*	28		22	
S179_Rain	76	*	30		24	
S21A_Rain	75	*	26		20	
S21_Rain	72	*	35		24	
S20_Rain	71	*	23		22	
HOMESFS_rain	71	*	35		26	
S165_Rain	71	*	40		21	
S167_Rain	70	*	38		24	
S18C_Rain	69	*	29		18	
S177_Rain	67	*	38		17	
MDTS_rain	67	*	18		15	
tpts_rain	67	*	14		16	
S332_Rain	63	*	38		19	
S174_Rain	62	*	45		21	
JBTS_rain	62	*	21		17	
S123_rain	61	*	39		23	
S334_rain	29		88	*	20	
S335_rain	29		88	*	20	
S336_Rain	36		83	*	21	
S338_Rain	48	*	65	*	22	
S331W_rain	53	*	62	*	23	
S331_Rain	48	*	60	*	22	
S26_rain	47		48	*	28	
miamai_ap_rain	22		21		82	*
TAMI_AIR_rain	22		19		82	*
PERRINE_Rain	27		22		81	*
miami_2_Rain	25		15		77	*
Printed values are multiplied by 100 and rounded to the nearest integer. Values greater than 0.474122 are flagged by an '*'.						

Figure 2. Principal components biplot for the rain data (complete cases). The axes correspond to the first (x-axis) and second (y-axis) principal component with the proportion of the variance of standardized variables explained in parentheses. The red names correspond to the different sites; the blue numbers correspond to the day, starting from 1/1/2000.

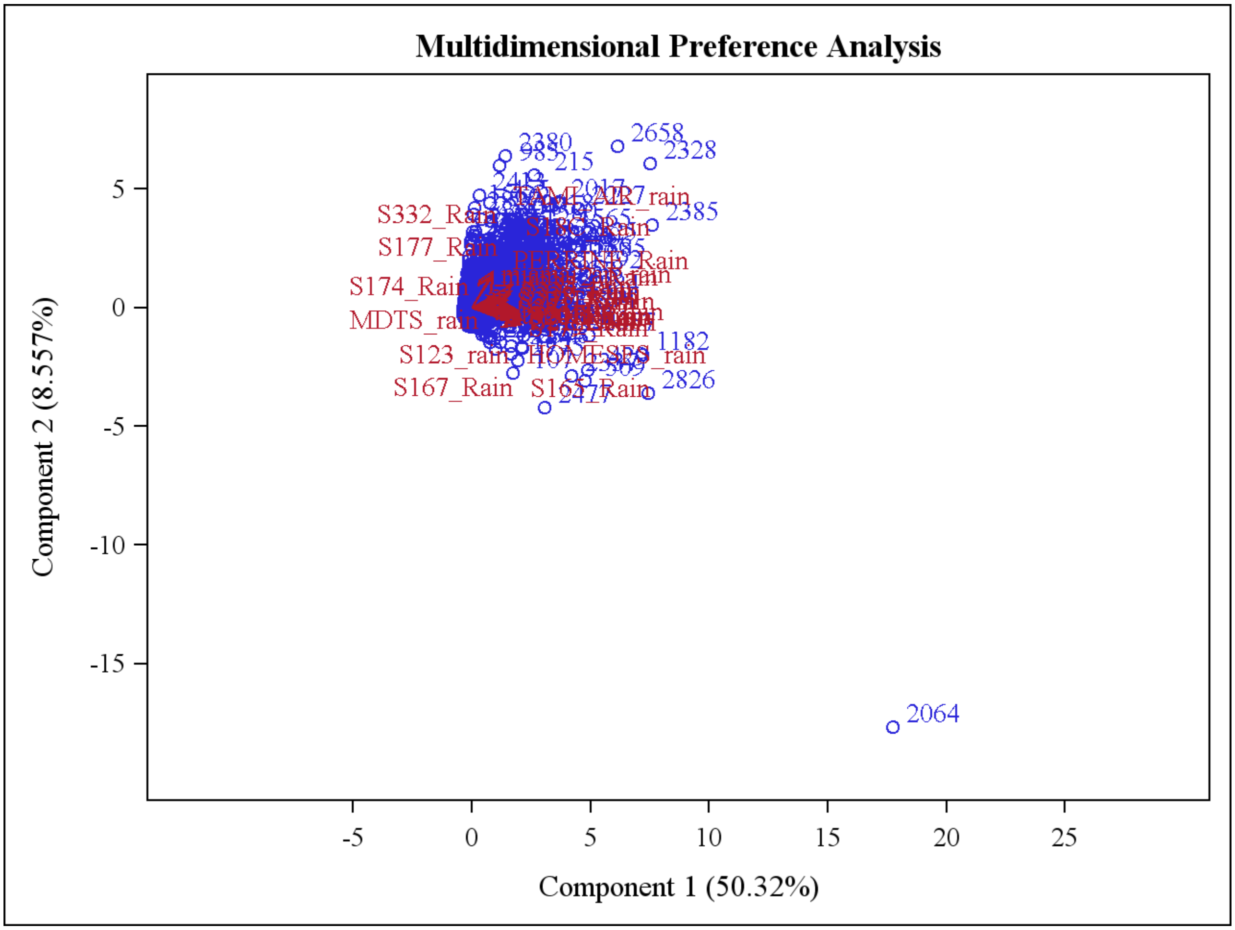
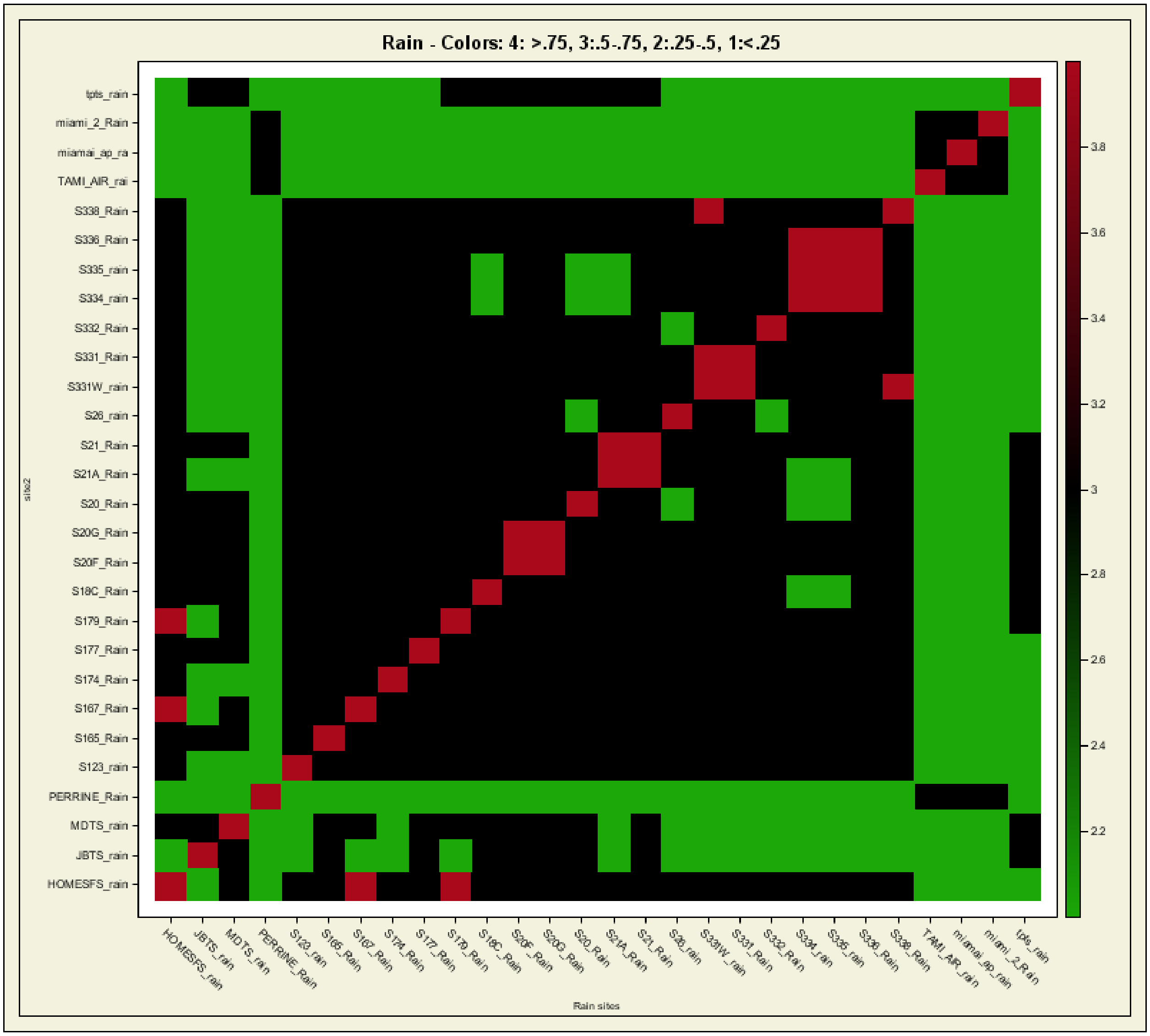


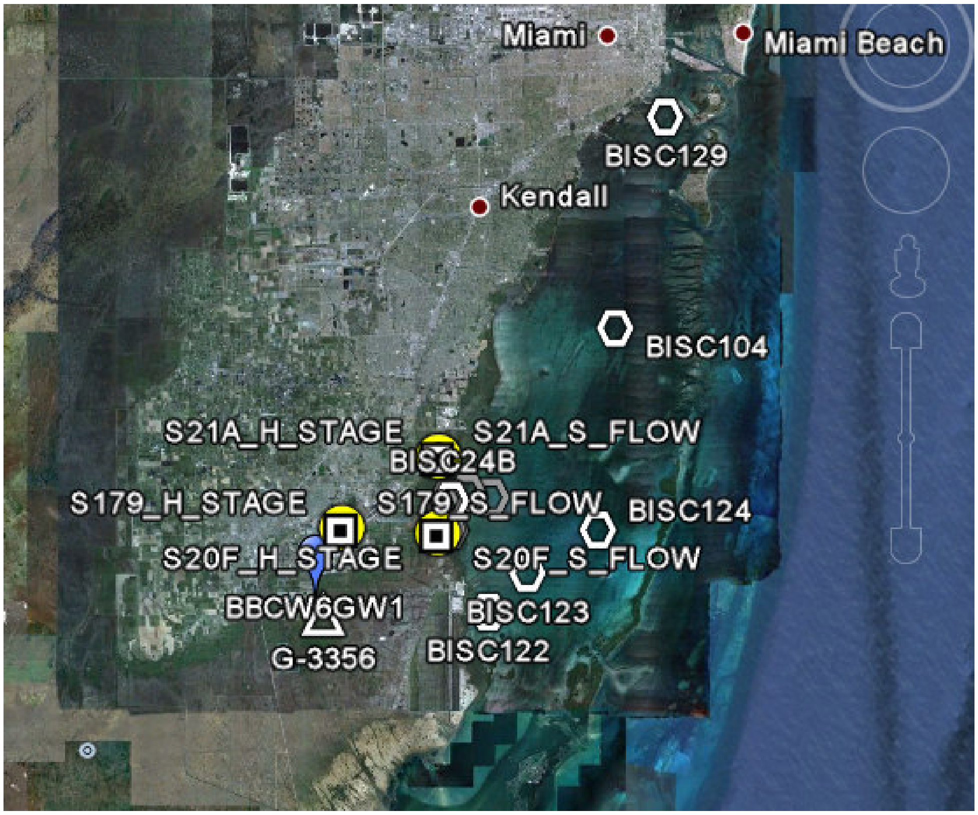
Figure 3. Heat map for rain variables using Spearman correlations. Colors are associated with different sizes of correlation using different colors. Note that in the case of rain, no correlations were below 0.25 so only three colors are needed.

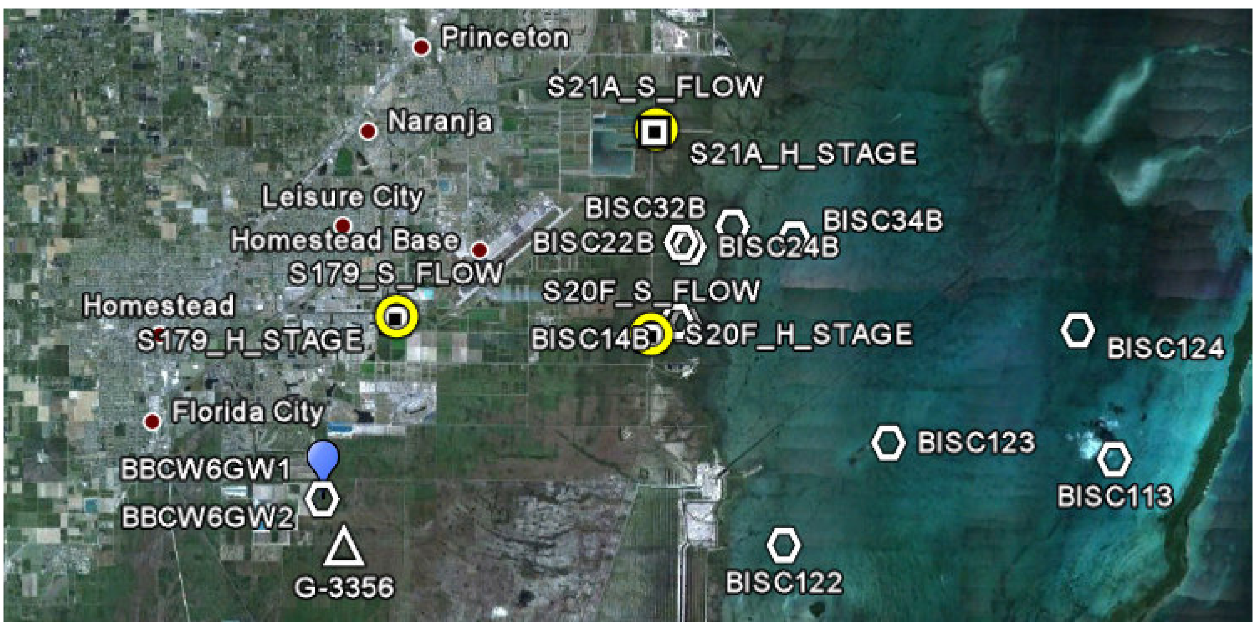


6. Example analysis: S20F flow and salinity

To illustrate how the compiled data set may be used to aid in understanding and exploring relationships among the different data types at specific sites, consider the relationship between flow at S20F and two neighboring stations and salinity, well and stage. The main question of interest is whether or not the flow out of S20F causes changes in salinity, stage and well levels. This is a complex question as factors other than those measured may be required for a full assessment. Here a simpler question is addressed: Is there evidence that the salinity, stage and well levels are related to flow. To address this, correlations between flow variables S20F_S_flow, S21A_S_flow, S179_S_flow and well, stage and salinity are calculated. The variables considered are the well variables BBCW6GW1_Well_D, BBCW6GW2_Well_D and G_3356_Well, and the stage variables S179_H_Stage_J, S20F_H_Stage_J, and S21A_H_Stage_D. Salinity at six locations close to the flow station (BISC16_Sal_BOT, BISC24_Sal_BOT, BISC34_Sal_BOT, BISC14_Sal_BOT, BISC32_Sal_BOT, and BISC22_Sal_BOT) are considered as well as four sites farther away from the S20F site, but still potentially influenced by S20F flows (BISC113_SALD_bot, BISC124_SALD_bot, BISC123_SALDbot, BISC122_SALD_bot). In addition, two sites located north of S20F that are unlikely to be influenced by S20F flows, BISC129 and BISC104, are used to estimate the degree of correlation that is likely to be associated with other factors. These sites are expected to have correlations with S20F flows that are close to zero. If the correlations are non-zero, it possibly indicates that flow and salinity are partly connected through another factor. Correlations are calculated using both daily and monthly values.

Figure 4. Maps of sites used in the analysis. The first map displays all the locations (some are hidden) while the second focuses on the locations closer to S20F.





To display the relationships scatterplots are used (Figures 5-9). In Figures 5 and 6, the three flow variables are plotted along with the log transformed flow for S20F versus the well and stage data. The data exhibit nonlinear relationships and there are a variety of extreme values, both for flow and well depth. The nonlinear pattern is expected as well and stage measurements have maximum values (so an increasing or decreasing pattern followed by a flattening pattern in expected). The log transformation of flow does improve the linearity of the relationship and there are quite a few zero values that make modeling the relationship difficult. This suggests that parametric correlations are not likely to adequately capture the associations and that nonparametric correlation is more appropriate. Kendall's correlation will be used to evaluate the strength of association between flows and the other variables.

Figure 5. Scatterplot of well data versus flow data. Ls20F is the log of S20F flow.

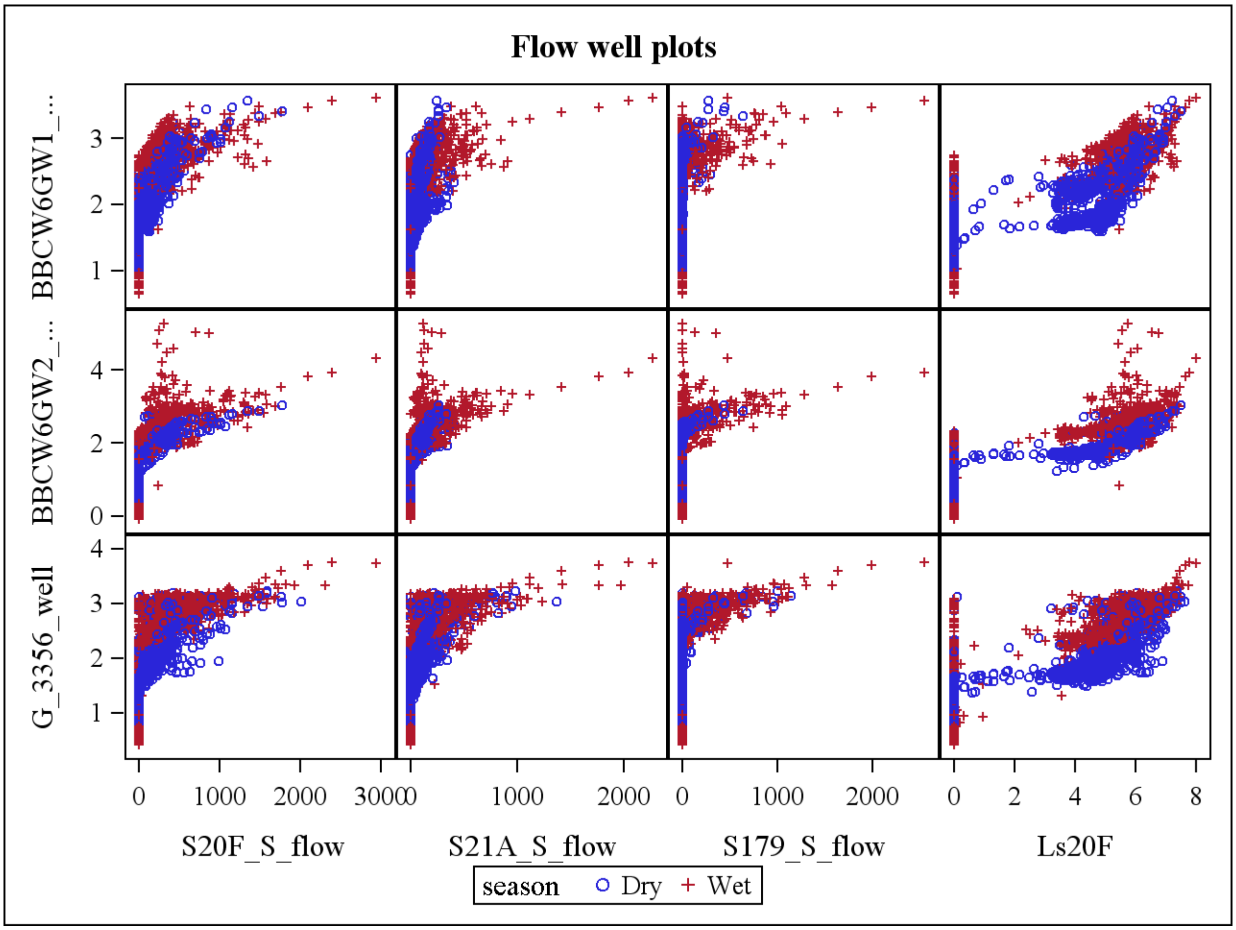
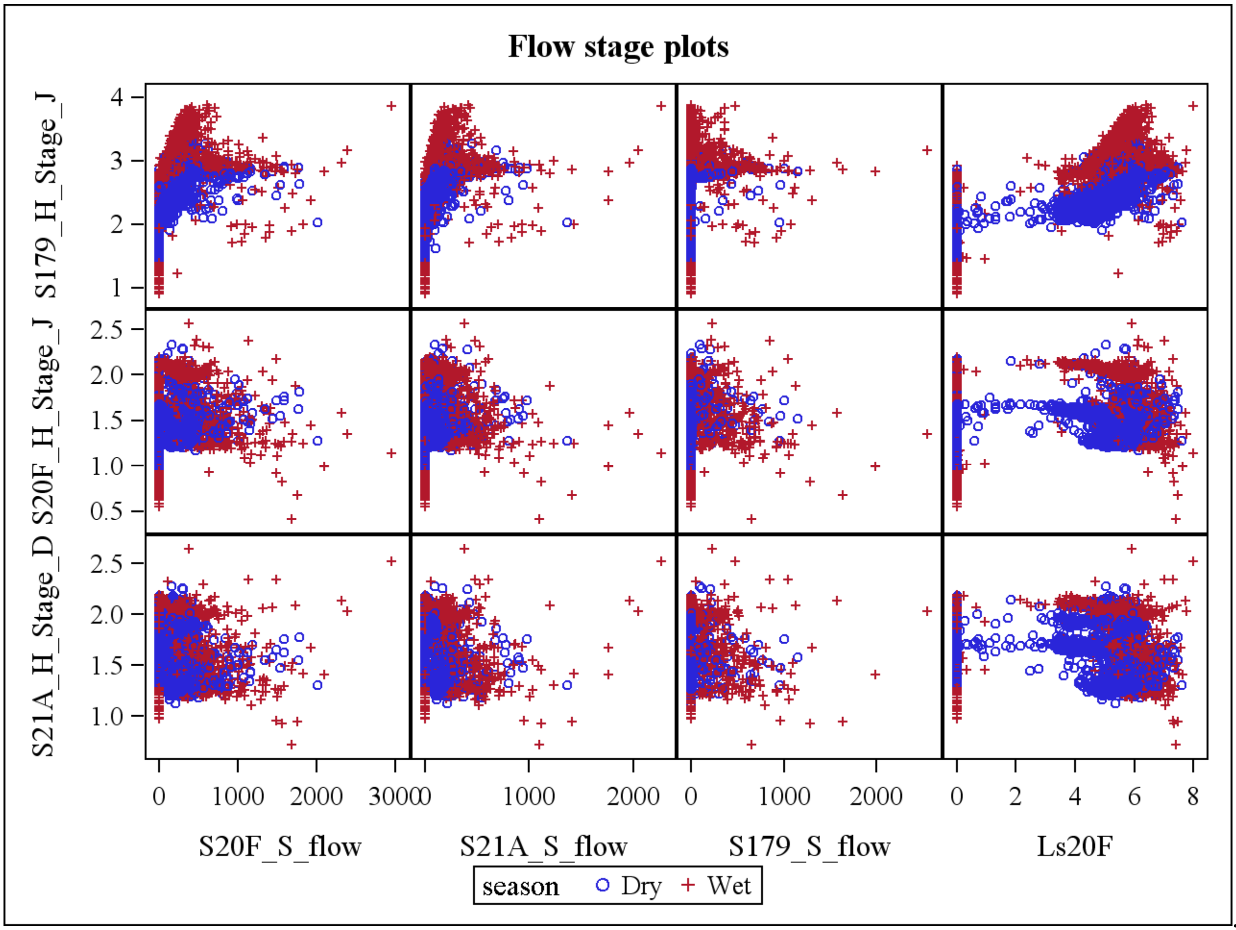


Figure 6. Scatterplot of stage data versus flow data. Ls20F corresponds to the log transformed flow at S20.



Tables 6 and 7 give the Kendall correlations for the flow measurements and measurements for well, stage and salinity for daily data with separate calculations for wet and dry seasons. The well values are moderately correlated for BBCW6GW1_Well_D BBCW6GW2_Well_D and G_3356_Well. Correlations are strongest with S20F and weak for S179. The correlation is also moderate between the flow variables and S179_H_Stage_J. The other correlations with stage are negative and low. The wet season tends to have slightly lower correlations with flow and stage.

Table 6. Summary statistics and Kendall correlation coefficients for well and stage variables with flow. The summary statistics are for the flow variables.

Obs	season	_TYPE_	_NAME_	S20F_S_flow	S21A_S_flow	S179_S_flow
1	Dry	MEAN		152.42	85.43	13.30
2	Dry	STD		194.84	102.38	78.46
3	Dry	N		1983.00	1983.00	1983.00
4	Dry	CORR	BBCW6GW1_Well_D	0.53	0.32	0.25
5	Dry	CORR	BBCW6GW2_Well_D	0.71	0.59	0.32
6	Dry	CORR	G_3356_well	0.70	0.58	0.28
7	Dry	CORR	S179_H_Stage_J	0.62	0.54	0.26
8	Dry	CORR	S20F_H_Stage_J	-0.28	-0.22	0.10
9	Dry	CORR	S21A_H_Stage_D	-0.34	-0.51	-0.03
10	Wet	MEAN		275.47	159.22	72.17
11	Wet	STD		312.04	200.17	196.30
12	Wet	N		1670.00	1670.00	1670.00
13	Wet	CORR	BBCW6GW1_Well_D	0.50	0.44	0.35
14	Wet	CORR	BBCW6GW2_Well_D	0.52	0.49	0.35
15	Wet	CORR	G_3356_well	0.59	0.54	0.40
16	Wet	CORR	S179_H_Stage_J	0.41	0.41	0.09
17	Wet	CORR	S20F_H_Stage_J	-0.21	-0.18	-0.22
18	Wet	CORR	S21A_H_Stage_D	-0.31	-0.34	-0.28

Scatterplots of salinity versus flow are given in Figures 7-9. Plots for the sites close to S20F (near sites) BISC16, 24, 34, 14, 22 and 32 are in Figure 7. A general decline in salinity with flow is observed in the plots with S20F and S21A, less so with S179. The pattern is not linear although there is a general decline in salinity with flow. There are several dates with extreme flows. Interestingly the salinity is not always low for these high flow periods (for example, BISC34 has moderate salinity for a period in August-September 2005, even though the flow is high). A similar pattern is present in the plots for BISC14, 22, 32. Figure 8 plots flow and salinity for BISC113, 122, 123 and 124 (far sites). The plots show little pattern. Even though the plots are of daily observations, there are not many observations for most of the sites. Figure 9 plots the data for the two northern sites and the scatterplots for these sites look similar to those of the three far sites. Note that the sites with fewer observations were not measured during the highest flows and the range of salinity is smaller. It is also evident from the plots that sites near were sampled more frequently than the other sites.

Figure 7. Scatterplot of salinity data versus flow for near sites.

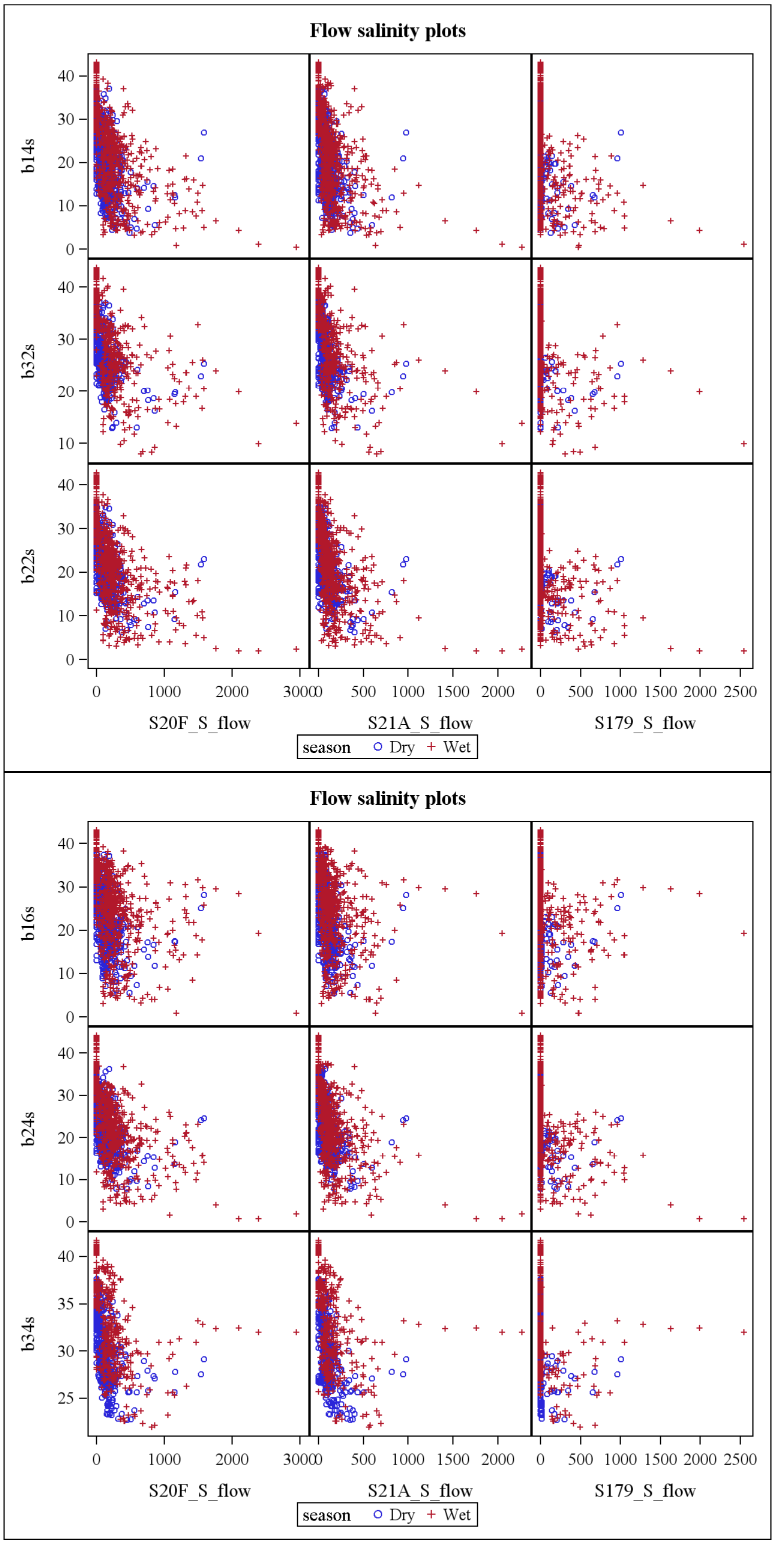


Figure 8. Scatterplot of salinity data versus flow for far sites.

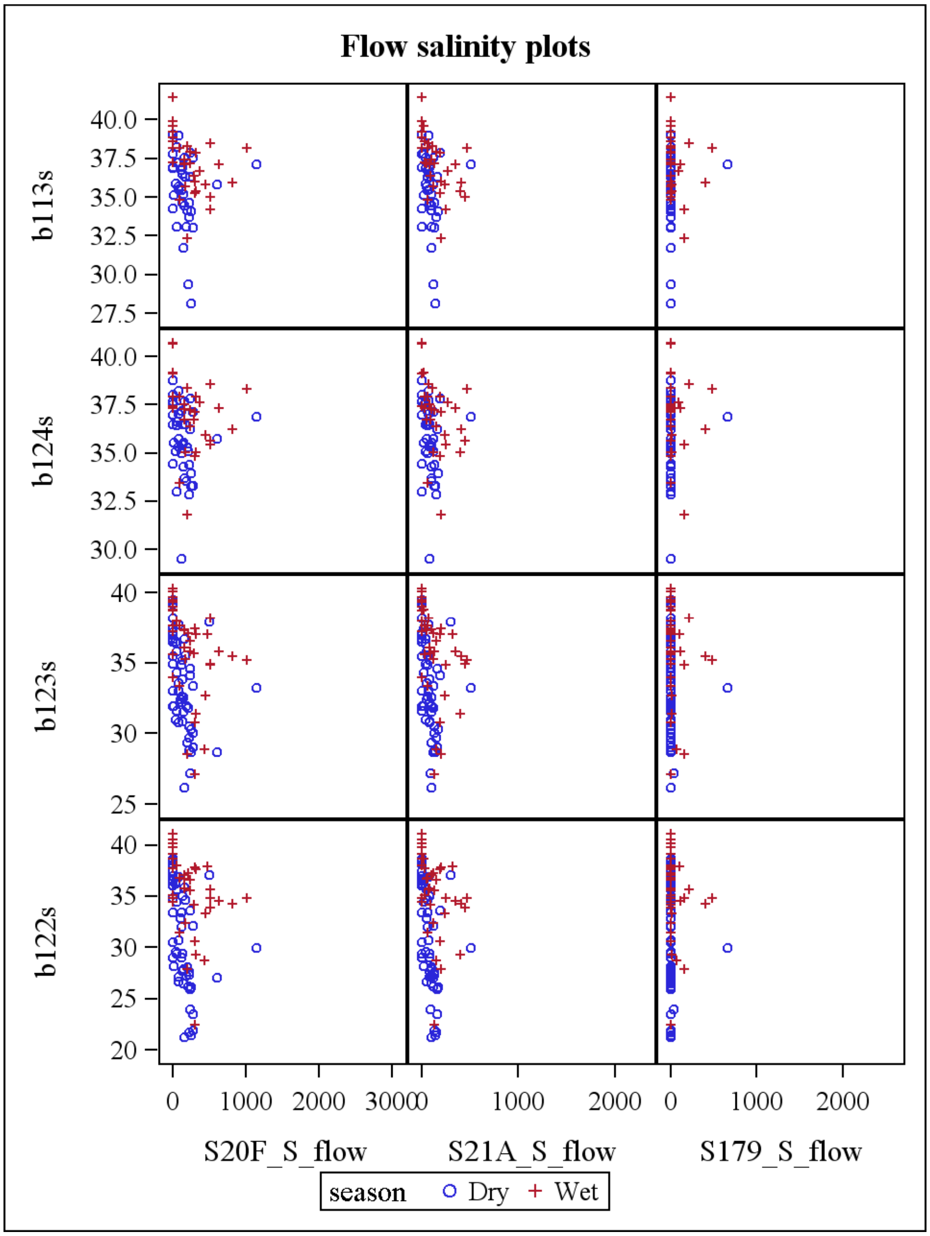
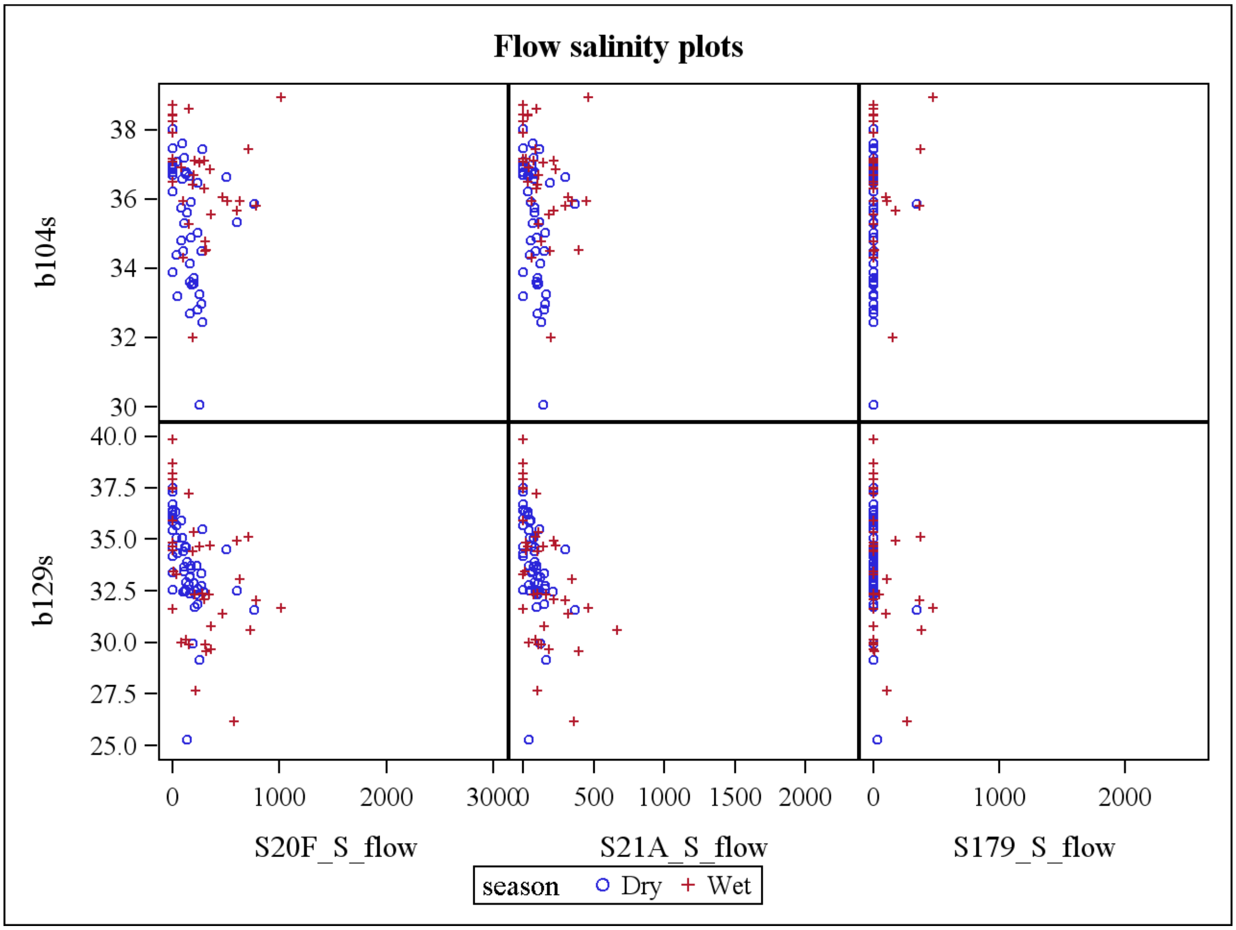


Figure 9. Scatterplot of salinity data versus flow for two northern sites.



Kendall correlation coefficients for the salinity data are given in Table 7. For the dry period there is little difference in the correlations except perhaps for BISC113 and BISC124. Correlations tend to be moderate in size for S20F and S21A. Correlations for S179 flow with salinity are quite low. For the wet period the pattern is similar although correlations are higher for S179 flow with salinity relative to the dry period.

Sample sizes for the variables are given in Table 8 and indicate considerable differences in sample size. The sample sizes for the well and stage variables tend to be close to or the same as the flow variables. Sample sizes for the salinity variables are moderate for the stations close to S20F but quite small for the other stations. Because of the sample size issue and autocorrelations, monthly averages were calculated and used in a second analysis.

Plots of the monthly means for well and stage data are given in Figures 10 and 11 and Table 9 gives the correlation coefficients. Again a nonlinear pattern is present, partly due to a flattening of the relationship for high flows with higher levels present in the wet period. Patterns for the three well sites are somewhat similar.

Table 7. Kendall correlation coefficients for salinity and flow.

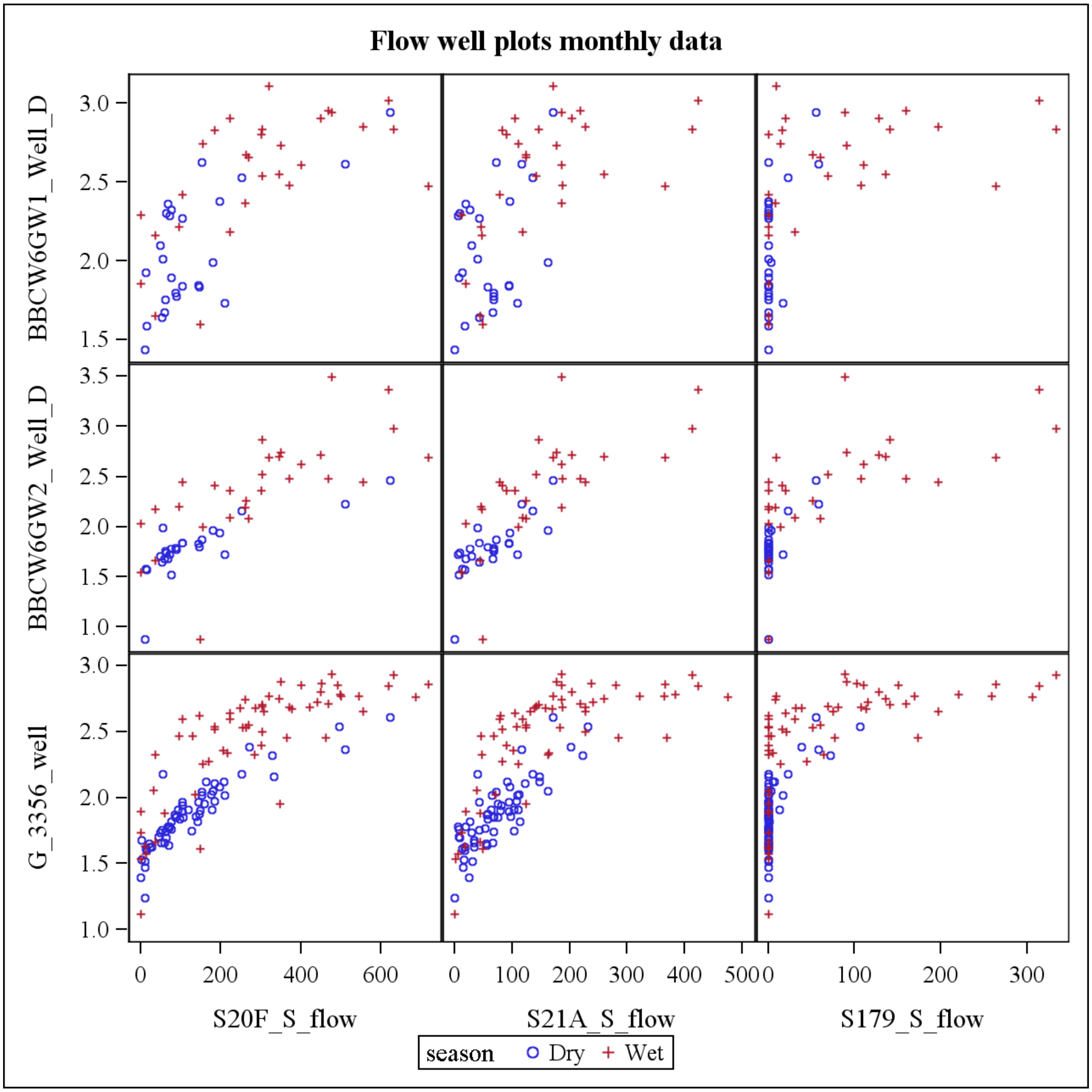
Obs	season	_TYPE_	_NAME_	S20F_S_flow	S21A_S_flow	S179_S_flow
1	Dry	MEAN		152.42	85.43	13.30
2	Dry	STD		194.84	102.38	78.46
3	Dry	N		1983.00	1983.00	1983.00
4	Dry	CORR	BISC113_SALD_bot	-0.29	-0.30	0.11
5	Dry	CORR	BISC124_SALD_bot	-0.24	-0.26	0.09
6	Dry	CORR	BISC123_SALD_bot	-0.45	-0.44	-0.08
7	Dry	CORR	BISC122_SALD_bot	-0.48	-0.50	-0.06
8	Dry	CORR	BISC16_Sal_BOT	-0.44	-0.44	-0.20
9	Dry	CORR	BISC24_Sal_BOT	-0.47	-0.44	-0.20
10	Dry	CORR	BISC34_Sal_BOT	-0.45	-0.38	-0.27
11	Dry	CORR	BISC14_Sal_BOT	-0.46	-0.44	-0.18
12	Dry	CORR	BISC32_Sal_BOT	-0.49	-0.43	-0.14
13	Dry	CORR	BISC22_Sal_BOT	-0.48	-0.45	-0.20
14	Dry	CORR	BISC104_SALD_bot	-0.36	-0.38	0.07
15	Dry	CORR	BISC129_SALD_bot	-0.53	-0.51	-0.25
16	Wet	MEAN		275.47	159.22	72.17
17	Wet	STD		312.04	200.17	196.30
18	Wet	N		1670.00	1670.00	1670.00
19	Wet	CORR	BISC113_SALD_bot	-0.41	-0.51	-0.25
20	Wet	CORR	BISC124_SALD_bot	-0.27	-0.41	-0.14
21	Wet	CORR	BISC123_SALD_bot	-0.44	-0.50	-0.31
22	Wet	CORR	BISC122_SALD_bot	-0.39	-0.42	-0.29
23	Wet	CORR	BISC16_Sal_BOT	-0.42	-0.38	-0.31
24	Wet	CORR	BISC24_Sal_BOT	-0.49	-0.47	-0.37
25	Wet	CORR	BISC34_Sal_BOT	-0.48	-0.46	-0.36
26	Wet	CORR	BISC14_Sal_BOT	-0.51	-0.47	-0.38
27	Wet	CORR	BISC32_Sal_BOT	-0.54	-0.51	-0.39
28	Wet	CORR	BISC22_Sal_BOT	-0.53	-0.50	-0.39

Obs	season	_TYPE_	_NAME_	S20F_S_flow	S21A_S_flow	S179_S_flow
29	Wet	CORR	BISC104_SALD_bot	-0.31	-0.35	-0.19
30	Wet	CORR	BISC129_SALD_bot	-0.39	-0.43	-0.31

Table 8. Sample sizes for variables.

Simple Sizes		
Variable	**N dry**	**N wet**
BISC113_SALD_bot	40	31
BISC124_SALD_bot	40	31
BISC123_SALD_bot	54	38
BISC122_SALD_bot	54	38
BISC16_Sal_BOT	659	642
BISC24_Sal_BOT	667	641
BISC34_Sal_BOT	443	410
BISC14_Sal_BOT	742	657
BISC32_Sal_BOT	538	491
BISC22_Sal_BOT	655	668
BISC104_SALD_bot	46	34
BISC129_SALD_bot	53	39
S20F_S_flow	1983	1670
S21A_S_flow	1983	1670
S179_S_flow	1983	1670
BBCW6GW1_Well_D	834	768
BBCW6GW2_Well_D	834	775
G_3356_well	1961	1656
S179_H_Stage_J	1983	1670
S20F_H_Stage_J	1983	1670
S21A_H_Stage_D	1983	1670

Figure 10. Scatterplot of well data versus flow data for monthly data.



The scatterplots for the stage variables is in Figure 11. For S179, there is an indication of relationships with flows at S20F and S21A. Again the pattern is nonlinear with an increase in stage with flow followed by a flattening for higher flows. There is less of a pattern for the other stage sites. Correlations for the well and stage variables are given in Table 9. Correlations between flow and well levels at BBCW6GW2 and G_3356 are moderately strong and positive as expected from the plots. Stage at S179 also exhibits a strong correlation. The correlation for the other stage sites is weaker and negative. Again correlations are weaker with S179 than with the other sites.

Figure 11. Scatterplot of stage data versus flow for monthly values.

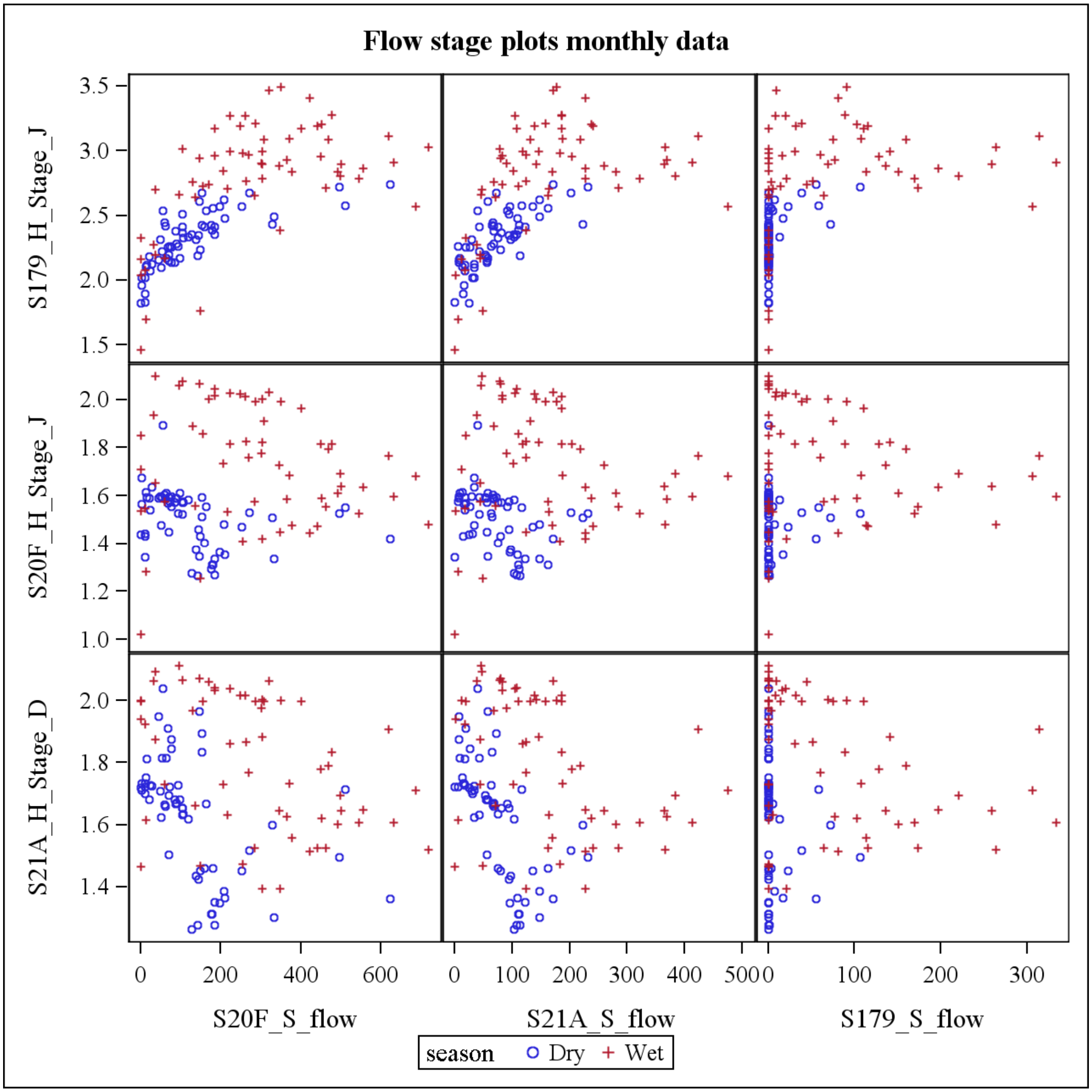


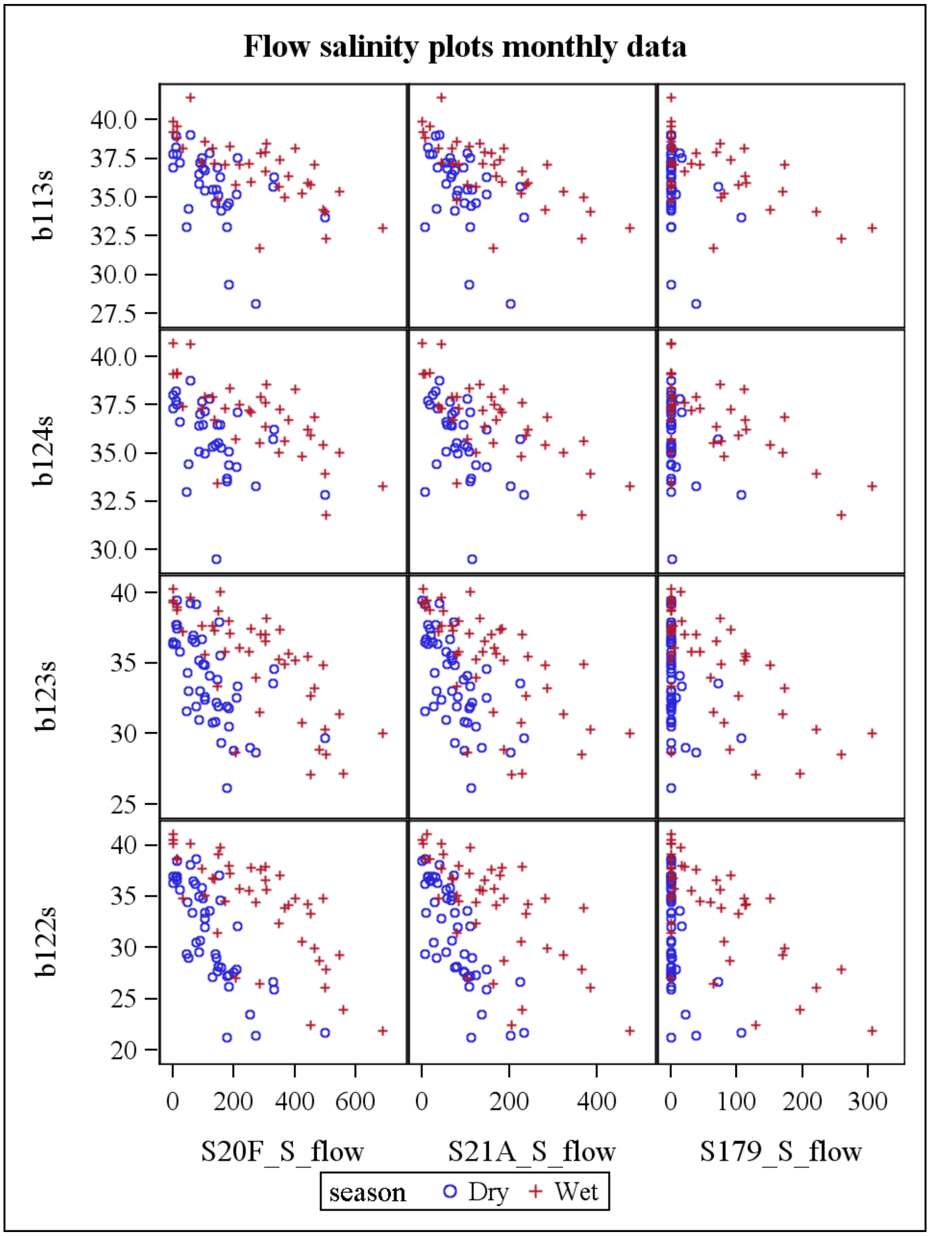
Table 9. Kendall correlations for stage and well variables using monthly data. Summary statistics are for the flow variables.

Obs	season	_TYPE_	_NAME_	S20F_S_flow	S21A_S_flow	S179_S_flow
1	Dry	MEAN		129.928	74.544	6.6707
2	Dry	STD		123.520	53.966	19.6785
3	Dry	N		60.000	60.000	60.0000
4	Dry	CORR	BBCW6GW1_Well_D	0.391	0.200	0.1024
5	Dry	CORR	BBCW6GW2_Well_D	0.655	0.612	0.3988
6	Dry	CORR	G_3356_well	0.785	0.621	0.3562
7	Dry	CORR	S179_H_Stage_J	0.662	0.568	0.3520
8	Dry	CORR	S20F_H_Stage_J	-0.355	-0.349	-0.1535
9	Dry	CORR	S21A_H_Stage_D	-0.443	-0.577	-0.1577
10	Wet	MEAN		286.331	163.280	73.3169
11	Wet	STD		186.227	115.915	90.5962
12	Wet	N		60.000	60.000	60.0000
13	Wet	CORR	BBCW6GW1_Well_D	0.478	0.394	0.4141
14	Wet	CORR	BBCW6GW2_Well_D	0.616	0.562	0.5396
15	Wet	CORR	G_3356_well	0.640	0.621	0.5901
16	Wet	CORR	S179_H_Stage_J	0.360	0.359	0.2808
17	Wet	CORR	S20F_H_Stage_J	-0.146	-0.149	-0.0987
18	Wet	CORR	S21A_H_Stage_D	-0.287	-0.297	-0.2259

Plots of monthly flow and salinity are given in Figures 12-14 and associated Kendall correlations are in Table 10. The plots and correlations indicate a general decline in salinity levels at all locations as flow increases. Note however that the axes do not have the same scales. Declines are over a greater range in the sites that are close to S20F than at sites that are farther away. In addition, the range is greater in the wet season than in the dry season.

Correlations between salinity and flow are moderately large and negative, ranging from -0.390 to 0.733. Correlations tend to greater at sites closer to S20F than those that are farther away, especially in the dry season. The greater correlations occur for sites BISC14, 24, 34, 16, 22, and 32. Correlations with sites BISC 113, 122, 123, and 124 are smaller in the dry period but are roughly the same in the wet period. Note also that sites that are not expected to be correlated with S20F have correlations around -0.45 (BISC104 and 129). These correlations suggest that a relationship with flow may be partly due to another factor, perhaps at a larger spatial scale (possibly due to weather patterns or seasonal factors).

Figure 12. Scatterplot of salinity data versus flow for close sites using monthly data



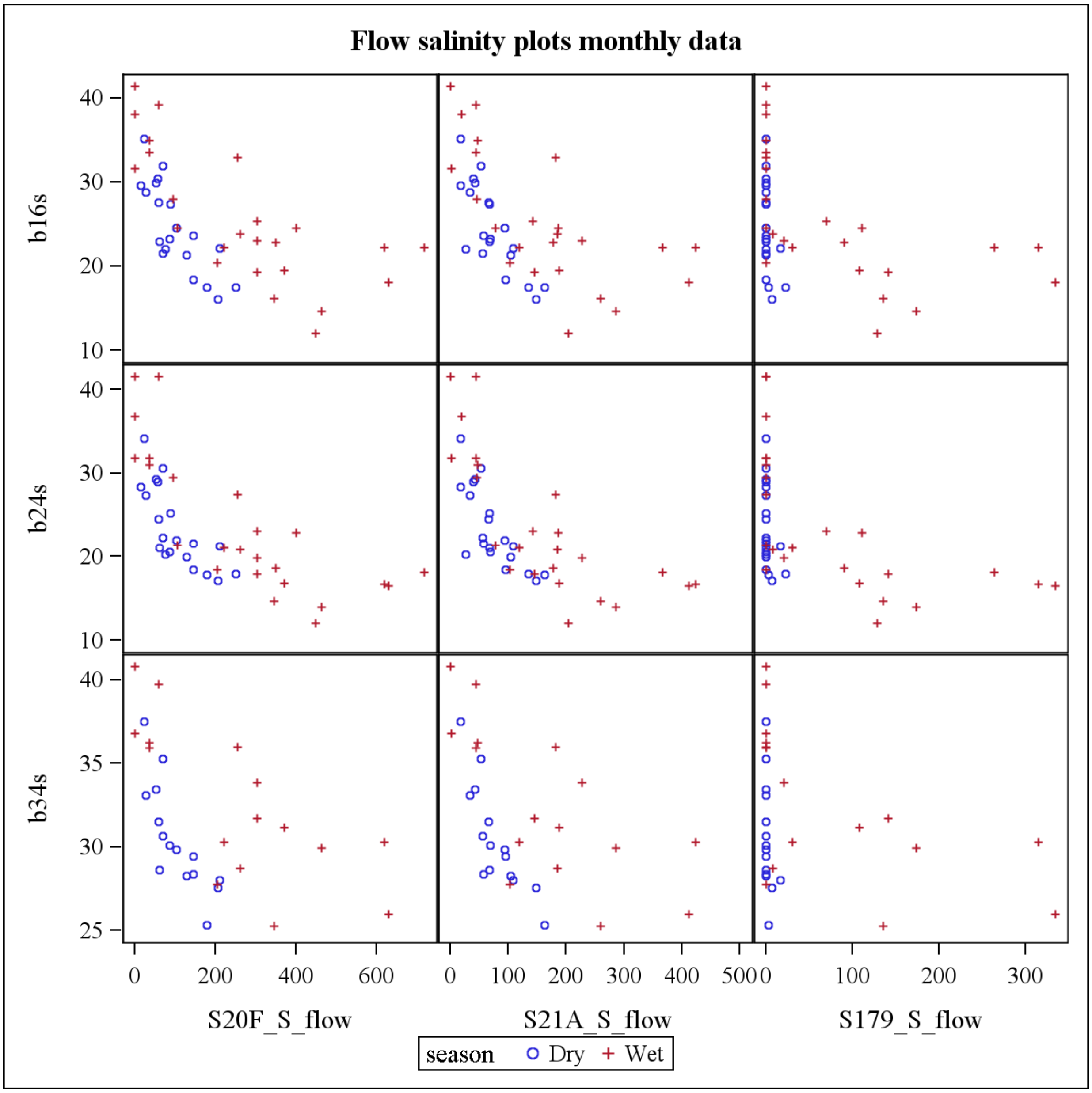

Flow salinity plots monthly data
b16s
b24s
b34s
S20F_S_flow
S21A_S_flow
S179_S_flow
season Dry Wet

Figure 13. Scatterplots of salinity data versus flow for far stations using monthly data.

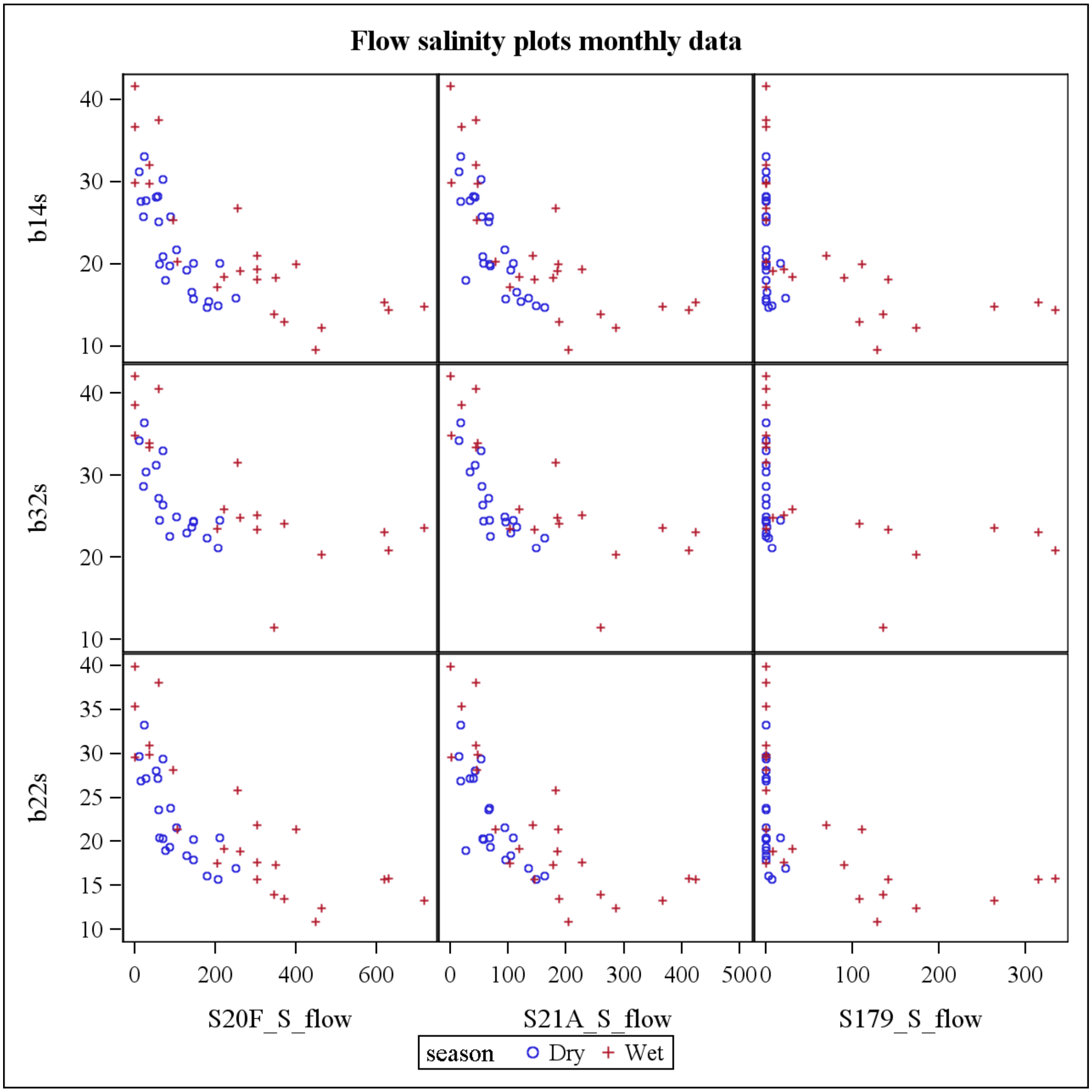


Figure 14. Scatterplots of salinity versus flow for northern stations using monthly data.

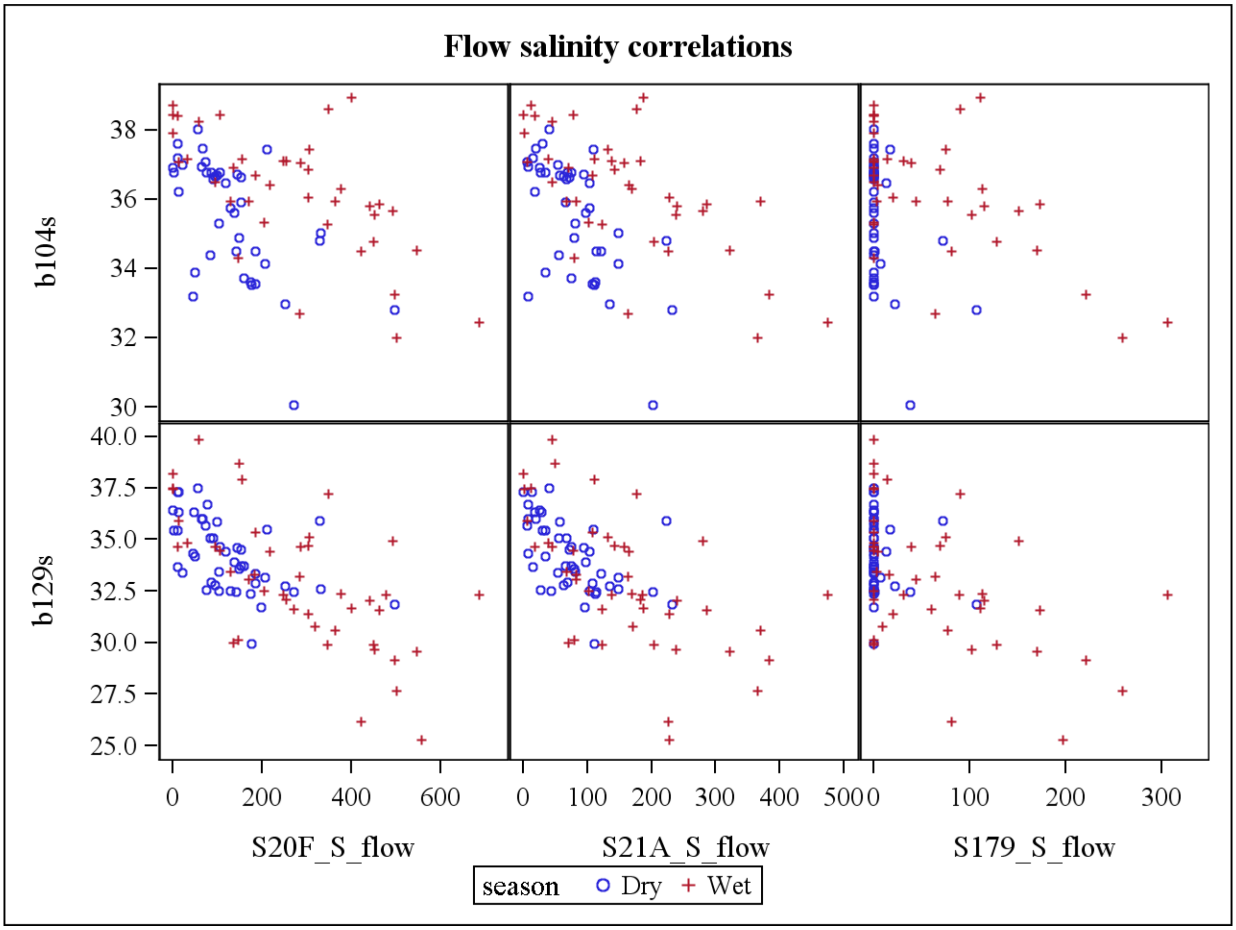


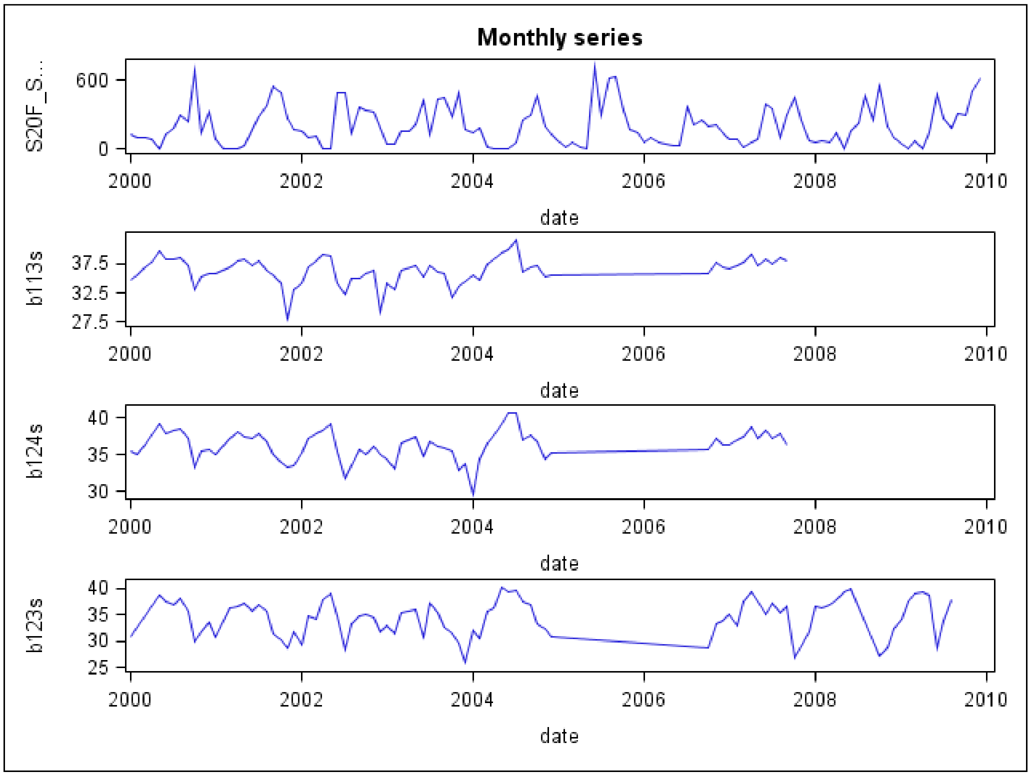
Table 10. Correlations between flow and salinity variables using monthly data. Summary statistics and sample sizes are for the flow variables.

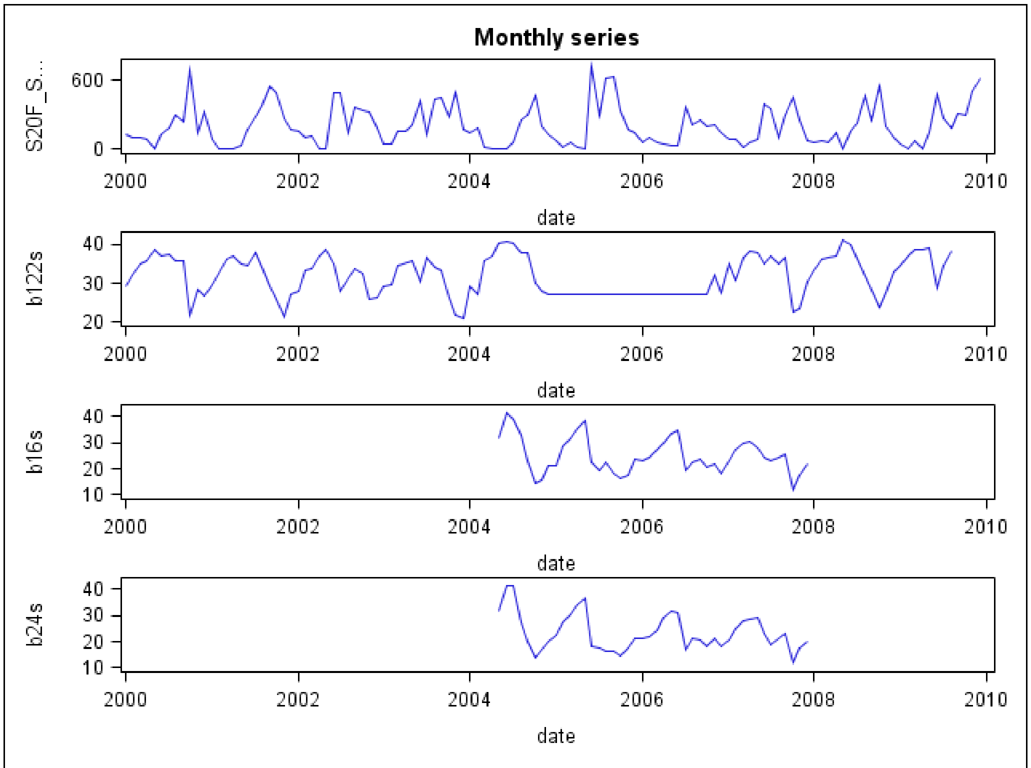
Obs	season	_TYPE_	_NAME_	S20F_S_flow	S21A_S_flow	S179_S_flow
1	Dry	MEAN		129.928	74.544	6.6707
2	Dry	STD		123.520	53.966	19.6785
3	Dry	N		60.000	60.000	60.0000
4	Dry	CORR	BISC113_SALD_bot	-0.413	-0.373	-0.0506
5	Dry	CORR	BISC124_SALD_bot	-0.390	-0.437	-0.1540
6	Dry	CORR	BISC123_SALD_bot	-0.465	-0.484	-0.1797
7	Dry	CORR	BISC122_SALD_bot	-0.597	-0.611	-0.2136
8	Dry	CORR	BISC16_Sal_BOT	-0.611	-0.600	-0.2806
9	Dry	CORR	BISC24_Sal_BOT	-0.632	-0.621	-0.2689
10	Dry	CORR	BISC34_Sal_BOT	-0.733	-0.771	-0.3703
11	Dry	CORR	BISC14_Sal_BOT	-0.623	-0.696	-0.2825
12	Dry	CORR	BISC32_Sal_BOT	-0.608	-0.752	-0.4122
13	Dry	CORR	BISC22_Sal_BOT	-0.657	-0.610	-0.3062
14	Dry	CORR	BISC104_SALD_bot	-0.479	-0.466	-0.2698
15	Dry	CORR	BISC129_SALD_bot	-0.449	-0.472	-0.2047
16	Wet	MEAN		286.331	163.280	73.3169
17	Wet	STD		186.227	115.915	90.5962
18	Wet	N		60.000	60.000	60.0000
19	Wet	CORR	BISC113_SALD_bot	-0.554	-0.586	-0.4390
20	Wet	CORR	BISC124_SALD_bot	-0.541	-0.509	-0.4349
21	Wet	CORR	BISC123_SALD_bot	-0.612	-0.550	-0.5129
22	Wet	CORR	BISC122_SALD_bot	-0.596	-0.514	-0.5146
23	Wet	CORR	BISC16_Sal_BOT	-0.580	-0.580	-0.5984
24	Wet	CORR	BISC24_Sal_BOT	-0.681	-0.681	-0.6372
25	Wet	CORR	BISC34_Sal_BOT	-0.533	-0.550	-0.5078
26	Wet	CORR	BISC14_Sal_BOT	-0.652	-0.623	-0.6217
27	Wet	CORR	BISC32_Sal_BOT	-0.647	-0.660	-0.6474

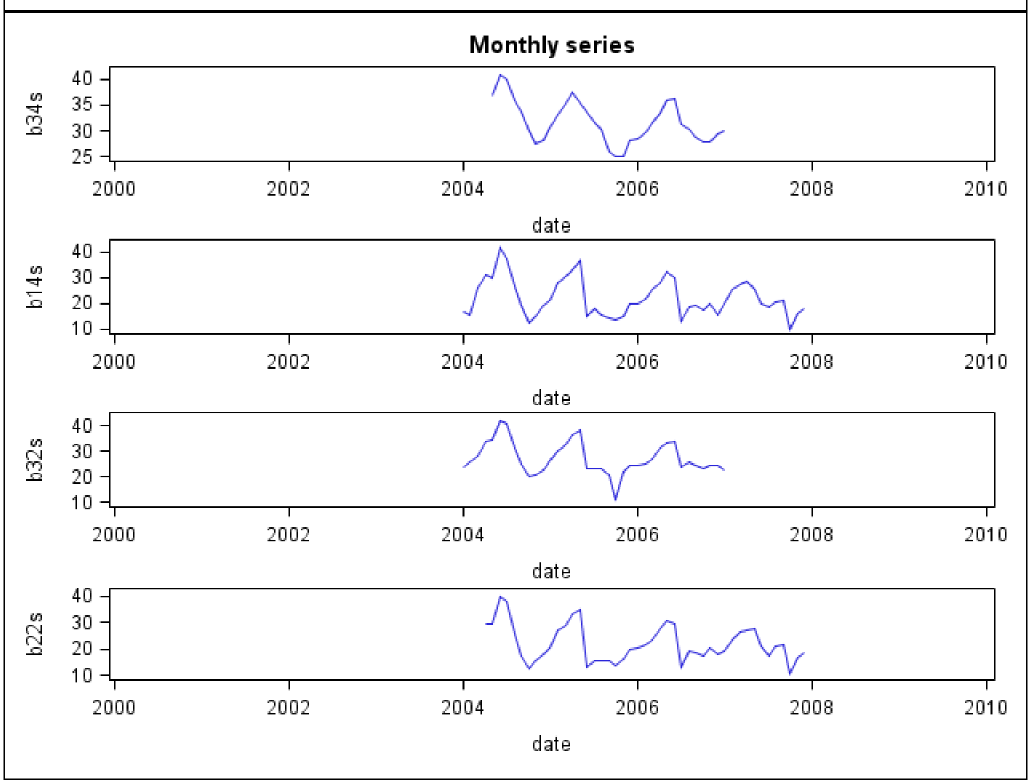
Obs	season	_TYPE_	_NAME_	S20F_S_flow	S21A_S_flow	S179_S_flow
28	Wet	CORR	BISC22_Sal_BOT	-0.696	-0.667	-0.6139
29	Wet	CORR	BISC104_SALD_bot	-0.473	-0.435	-0.3725
30	Wet	CORR	BISC129_SALD_bot	-0.499	-0.501	-0.3723

The correlations are based on data with differing sample size and this may be important if the sampling periods are quite different. Figure 15 below plots the data for salinity and flow at S20F over time for the different stations. Note that one group of stations is mostly observed in the 2004-2008 period while the other group is missing during some of that period. This suggests that care is needed in comparing the correlations across stations.

Figure 15. Time series plots of monthly data for flow and salinity.




Monthly series
S20F_S...
600
0
b122s
40
30
20
b16s
40
30
20
10
b24s
40
30
20
10
2000
2002
2004
2006
2008
2010
date


Monthly series
b34s
40
35
30
25
b14s
40
30
20
10
b32s
40
30
20
10
b22s
40
30
20
10
2000
2002
2004
2006
2008
2010
date

The above analyses show how the data and correlations can be used to investigate relations between measurements of different types. The analyses were run using both daily data and monthly averages. Note that in interpreting the results the actual times of sampling are important and might limit interpretation. As indicated in the tables, the sample sizes are different for different pairs of measurements indicating that the times of sampling are not matched. Correlations can only be compared when the times are matched or the times may be viewed as a random sample of times. For example, if station BISC16 is sampled only in the wet season and station BISC32 only in the dry season, the number of samples is likely to be the same but sampling dates completely mismatched. An analysis comparing the correlations would not be useful in that case.

The analysis suggests that there is a relationship between flow at S20F and salinity levels at nearby stations. The relationship is strongest at the stations closer to the S20F location. However, the strength of the relationship is difficult to assess as there may be other factors involved in the relationship. This is suggested by the fact that salinity at stations farther from S20F is also correlated with flow at S20F, although the correlation is slightly weaker. The reason for this correlation is unclear but suggests that care needs to be taken in interpretation of relationships. The above analysis is viewed as exploratory and more complex analyses may be required to try to provide a better estimate of the relationship, especially if daily data are to be used.

Appendix 1. Data and summaries

The following data was supplied on DVDs from the District:

1. July 19 2010 DVD
   SoMiami Dade final as document and PDF
   Stage - 47 EXCEL files
   Well - 2 files:
      G-1362
      USGS_daily_data_Scrub (a large file)
2. Data DVD
   Flow_station - 23 EXCEL files
   Rain – one EXCEL file
   Stage – one file with all stage data (date is before)
   Well data – 32 EXCEL files
   Final salinity graphs
   Original salinity graphs
   SoMiamiDadeScrubFinal – Access file dated 7/8/2010
3. Scrub documents
   SoMiamiDade_final (doc and pdf)
   SoMiamiDadeScrubFinalRpt_FinalsTable Revised (large file of data quality data)
4. Wells
   40 EXCEL files of well water levels
5. Additional Flow data on 32 sites was added in an EXCEL file

Summaries of the data files created and combined from the supplied files are described below. Frequency tables provide counts of codes for each data file. No code refers to the number of observations that have no code. Summaries include the basic summary statistics as well as the start and end dates for the site measurements. Freq is the number of days with measurements.

Well data

Code	Description	Frequency	Percent	Cumulative Frequency	Cumulative Percent
A	accumulated	34971	74.64	34971	74.64
E	estimated	8737	18.65	43708	93.29
M	missing	1951	4.16	45659	97.45
N	Data not processed	14	0.03	45673	97.48
P	Partial record	1179	2.52	46852	100.00

No code = 17391

Obs	site	Freq	mean_well	std_well	mindate	maxdate
1	G-3698_Drawdow	79	1.97636	0.20305	14OCT2009	31DEC2009
2	G-3699_Drawdow	79	1.69934	0.24792	14OCT2009	31DEC2009
3	G-3700_Drawdow	79	2.30647	0.26260	14OCT2009	31DEC2009
4	G-3701_Drawdow	79	1.97317	0.23984	14OCT2009	31DEC2009
5	G-3897_RT_Well	174	3.55103	0.75740	10APR2009	30SEP2009
6	G3900_RT_Well	190	2.31370	0.73274	25MAR2009	30SEP2009
7	BBCW10GW1_Well	1689	0.74479	0.45309	18MAY2005	31DEC2009
8	BBCW10GW2_Well	1689	0.75153	0.43767	18MAY2005	31DEC2009
9	BBCW1_Well	1689	2.86924	0.49720	18MAY2005	31DEC2009
10	BBCW7GW1_Well	1689	1.65092	0.37472	18MAY2005	31DEC2009
11	BBCW7GW2_Well	1689	1.58473	0.41694	18MAY2005	31DEC2009
12	BBCW8GW1_Well	1689	0.88887	0.45298	18MAY2005	31DEC2009
13	BBCW8GW2_Well	1689	0.76872	0.42795	18MAY2005	31DEC2009
14	BBCW2_Well	1690	1.92258	0.30324	17MAY2005	31DEC2009
15	BBCW3GW1_Well	1690	1.95971	0.37566	17MAY2005	31DEC2009
16	BBCW3GW2_Well	1690	1.93909	0.34545	17MAY2005	31DEC2009
17	BBCW4_Well	1690	1.67301	0.33666	17MAY2005	31DEC2009
18	BBCW5_Well	1690	1.78453	0.39265	17MAY2005	31DEC2009

Obs	site	Freq	mean_well	std_well	mindate	maxdate
19	BBCW6GW1_Well	1690	2.32755	0.51603	17MAY2005	31DEC2009
20	BBCW6GW2_Well	1690	2.10427	0.58340	17MAY2005	31DEC2009
21	BBCW9GW1_Well	1690	1.73517	0.27336	17MAY2005	31DEC2009
22	BBCW9GW2_Well	1690	1.74406	0.29621	17MAY2005	31DEC2009
23	G-1183_RT_Well	3653	2.01714	0.49372	01JAN2000	31DEC2009
24	G-3551_RT_Well	3653	5.43032	0.80250	01JAN2000	31DEC2009
25	G-3552_RT_Well	3653	4.83199	0.75811	01JAN2000	31DEC2009
26	G-3553_RT_Well	3653	4.66896	0.74742	01JAN2000	31DEC2009
27	G-3557_RT_Well	3653	5.37248	0.79474	01JAN2000	31DEC2009
28	G-3558_RT_Well	3653	4.74076	0.68863	01JAN2000	31DEC2009
29	G-3559_RT_Well	3653	5.57855	0.63940	01JAN2000	31DEC2009
30	G-3574_RT_Well	3653	5.83156	0.71887	01JAN2000	31DEC2009
31	G-3575_RT_Well	3653	5.82702	0.89634	01JAN2000	31DEC2009
32	G-580A_RT_Well	3653	2.57055	0.50382	01JAN2000	31DEC2009

Well19USGS data

Obs	site	Freq	mean_well	std_well	mindate	maxdate
1	G-3897_well	174	3.55103	0.75740	10APR2009	30SEP2009
2	G-3900_well	190	2.31370	0.73274	25MAR2009	30SEP2009
3	G-3328_well	893	2.15101	0.40216	01JAN2000	11JUN2002
4	MRS-HOPC1_well	1274	4.88216	1.23132	07JUL2006	31DEC2009
5	MRS-HOPD1_well	1281	4.63713	1.31592	30JUN2006	31DEC2009
6	MRS-HOPC2_well	1282	5.08557	1.22570	29JUN2006	31DEC2009
7	MRS-HOPC3_well	1283	5.08543	1.22293	28JUN2006	31DEC2009
8	MRS-HOPB1_well	1284	5.16728	1.20824	27JUN2006	31DEC2009
9	G-976_well	1769	4.37862	1.00246	01JAN2000	03NOV2004
10	G-3336_well	2374	2.20570	0.63170	03JUL2003	31DEC2009
11	G-3473_well	2515	4.07662	0.69058	01JAN2000	19NOV2006

Obs	site	Freq	mean_well	std_well	mindate	maxdate
12	G-3354_well	3219	1.88068	0.51406	01JAN2000	23OCT2008
13	G-3353_well	3224	1.26652	0.44953	01JAN2000	28OCT2008
14	G-3568_well	3614	3.17671	0.46981	01JAN2000	22NOV2009
15	G-1362_well	3630	4.05851	0.70850	24JAN2000	31DEC2009
16	G-3466_well	3633	1.03854	0.70064	21JAN2000	31DEC2009
17	G-789_well	3640	3.74308	0.72068	01JAN2000	18DEC2009
18	G-3577_well	3649	5.99174	1.03846	01JAN2000	27DEC2009
19	F-179_well	3653	2.19417	0.61046	01JAN2000	31DEC2009
20	F-319_well	3653	2.41072	0.43828	01JAN2000	31DEC2009
21	F-358_well	3653	2.85940	0.69315	01JAN2000	31DEC2009
22	G-1074B_well	3653	-2.37208	2.68089	01JAN2000	31DEC2009
23	G-1183_well	3653	2.01714	0.49372	01JAN2000	31DEC2009
24	G-1251_well	3653	2.02633	0.59413	01JAN2000	31DEC2009
25	G-1363_well	3653	3.88663	0.78288	01JAN2000	31DEC2009
26	G-1486_well	3653	2.67219	0.55318	01JAN2000	31DEC2009
27	G-1487_well	3653	5.46842	0.83847	01JAN2000	31DEC2009
28	G-1488_well	3653	6.19594	0.68582	01JAN2000	31DEC2009
29	G-1502_well	3653	6.20510	0.97659	01JAN2000	31DEC2009
30	G-3074_well	3653	0.93968	1.47522	01JAN2000	31DEC2009
31	G-3272_well	3653	6.20654	0.97752	01JAN2000	31DEC2009
32	G-3327_well	3653	1.86446	0.37868	01JAN2000	31DEC2009
33	G-3329_well	3653	2.61294	0.42323	01JAN2000	31DEC2009
34	G-3355_well	3653	2.71260	0.63998	01JAN2000	31DEC2009
35	G-3356_well	3653	2.16957	0.55022	01JAN2000	31DEC2009
36	G-3437_well	3653	5.50493	1.14454	01JAN2000	31DEC2009
37	G-3439_well	3653	4.00275	0.67039	01JAN2000	31DEC2009
38	G-3465_well	3653	1.80825	0.58054	01JAN2000	31DEC2009
39	G-3467_well	3653	1.67307	0.44816	01JAN2000	31DEC2009
40	G-3549_well	3653	1.84734	0.26661	01JAN2000	31DEC2009
41	G-3550_well	3653	1.67639	0.33663	01JAN2000	31DEC2009
42	G-3551_well	3653	5.43032	0.80250	01JAN2000	31DEC2009
43	G-3552_well	3653	4.83199	0.75811	01JAN2000	31DEC2009

Obs	site	Freq	mean_well	std_well	mindate	maxdate
44	G-3553_well	3653	4.66896	0.74742	01JAN2000	31DEC2009
45	G-3554_well	3653	4.86237	0.72514	01JAN2000	31DEC2009
46	G-3555_well	3653	4.29196	0.69847	01JAN2000	31DEC2009
47	G-3556_well	3653	4.63476	0.73510	01JAN2000	31DEC2009
48	G-3557_well	3653	5.37248	0.79474	01JAN2000	31DEC2009
49	G-3558_well	3653	4.74076	0.68863	01JAN2000	31DEC2009
50	G-3559_well	3653	5.57849	0.63948	01JAN2000	31DEC2009
51	G-3560_well	3653	4.61805	0.74337	01JAN2000	31DEC2009
52	G-3561_well	3653	4.08322	0.63312	01JAN2000	31DEC2009
53	G-3563_well	3653	2.99852	0.46342	01JAN2000	31DEC2009
54	G-3564_well	3653	1.92093	0.65777	01JAN2000	31DEC2009
55	G-3565_well	3653	3.18059	0.52682	01JAN2000	31DEC2009
56	G-3570_well	3653	2.95303	0.82782	01JAN2000	31DEC2009
57	G-3572_well	3653	3.42915	0.47173	01JAN2000	31DEC2009
58	G-3574_well	3653	5.83156	0.71887	01JAN2000	31DEC2009
59	G-3575_well	3653	5.82668	0.89647	01JAN2000	31DEC2009
60	G-3576_well	3653	6.51129	0.78626	01JAN2000	31DEC2009
61	G-3578_well	3653	6.28917	0.86650	01JAN2000	31DEC2009
62	G-3619_well	3653	2.48420	0.53740	01JAN2000	31DEC2009
63	G-3620_well	3653	2.34831	0.53097	01JAN2000	31DEC2009
64	G-3621_well	3653	2.09764	0.48887	01JAN2000	31DEC2009
65	G-3622_well	3653	4.42026	0.99292	01JAN2000	31DEC2009
66	G-3626_well	3653	4.57934	0.57067	01JAN2000	31DEC2009
67	G-3627_well	3653	4.22012	0.54125	01JAN2000	31DEC2009
68	G-3628_well	3653	4.72737	0.76967	01JAN2000	31DEC2009
69	G-3676_well	3653	4.25429	0.73822	01JAN2000	31DEC2009
70	G-551_well	3653	-0.24703	1.59445	01JAN2000	31DEC2009
71	G-553_well	3653	3.19792	0.65426	01JAN2000	31DEC2009
72	G-580A_well	3653	2.57055	0.50382	01JAN2000	31DEC2009
73	G-596_well	3653	5.12269	0.68244	01JAN2000	31DEC2009
74	G-613_well	3653	2.42439	0.54057	01JAN2000	31DEC2009
75	G-614_well	3653	3.39476	0.83659	01JAN2000	31DEC2009

Obs	site	Freq	mean_well	std_well	mindate	maxdate
76	G-618_well	3653	6.92284	0.52509	01JAN2000	31DEC2009
77	G-620_well	3653	6.41530	0.79073	01JAN2000	31DEC2009
78	G-757A_well	3653	4.14115	0.69497	01JAN2000	31DEC2009
79	G-855_well	3653	4.48677	0.72021	01JAN2000	31DEC2009
80	G-860_well	3653	2.58234	0.40652	01JAN2000	31DEC2009
81	G-864A_well	3653	2.41284	0.80089	01JAN2000	31DEC2009
82	G-864_well	3653	2.53483	0.79078	01JAN2000	31DEC2009
83	S-182A_well	3653	2.68231	0.45356	01JAN2000	31DEC2009
84	S-196A_well	3653	3.58572	0.73779	01JAN2000	31DEC2009
85	S-19_well	3653	1.52917	0.65198	01JAN2000	31DEC2009
86	S-68_well	3653	0.29401	0.99075	01JAN2000	31DEC2009

South Florida Well data

Code	Description	Frequency	Percent	Cumulative Frequency	Cumulative Percent
>	Value is greater than	2	0.01	2	0.01
E	Estimated	20847	89.94	20849	89.95
M	Missing	2288	9.87	23137	99.82
P	Partial record	42	0.18	23179	100.00

No code = 95874

Obs	site	Freq	mean_well	std_well	mindate	maxdate
1	3BS1W2_G_WELL_km	3532	6.56140	0.91290	01MAY2000	31DEC2009
2	3BS1W3_G_WELL_km	3524	6.56452	0.91217	09MAY2000	31DEC2009
3	3BS1W4_G_WELL_km	3524	6.60799	0.88946	09MAY2000	31DEC2009
4	ANGEL_WELL_km	3653	5.67014	0.97958	01JAN2000	31DEC2009
5	C2GSW1_GW1_WELL_	2603	1.41669	1.25354	16NOV2002	31DEC2009
6	C2GSW1_GW2_WELL_	2603	-0.19695	2.58224	16NOV2002	31DEC2009
7	C2GW1_GW1_WELL_k	2603	0.46064	2.33199	16NOV2002	31DEC2009
8	C2GW1_GW2_WELL_k	2603	0.65454	2.62047	16NOV2002	31DEC2009
9	C4GW1_WELL_km(2)	1711	4.76646	0.91268	26APR2005	31DEC2009
10	DUCLOS_G_WELL_km	3653	3.96051	0.72598	01JAN2000	31DEC2009
11	F-179_WELL_km	2738	2.16969	0.60883	01JAN2000	30JUN2007
12	FRGPD2_G_WELL_km	3653	3.66316	0.67460	01JAN2000	31DEC2009
13	FROGP_G_WELL_km(	3653	4.06487	1.00978	01JAN2000	31DEC2009
14	G-1183_WELL_km	3104	2.00538	0.49073	01JAN2000	30JUN2008
15	G-1251_B_WELL_km	3653	2.04760	0.56961	01JAN2000	31DEC2009
16	G-1362_G_WELL_km	3653	4.07860	0.70828	01JAN2000	31DEC2009
17	G-1363_WELL_km	3653	3.88658	0.78922	01JAN2000	31DEC2009
18	G-1486_G_WELL_km	3652	2.67795	0.57710	01JAN2000	31DEC2009
19	G-1487_G_WELL_km	3104	5.46949	0.82803	01JAN2000	30JUN2008
20	G-3272_WELL_km	3653	6.20590	0.97774	01JAN2000	31DEC2009

21	G-3338_G_WELL_km	2248	2.45198	0.48837	06NOV2003	31DEC2009
22	G-3339_G_WELL_km	2248	2.45034	0.51169	06NOV2003	31DEC2009
23	G-3349_G_WELL_km	2248	1.99026	0.53160	06NOV2003	31DEC2009
24	G-3350_G_WELL_km	2248	2.22408	0.50907	06NOV2003	31DEC2009
25	G-3778_WELL_km	2150	5.59605	0.71382	12FEB2004	31DEC2009
26	G-3779_WELL_km	2185	5.57575	0.70200	08JAN2004	31DEC2009
27	G-3780_WELL_km	2185	5.58774	0.68404	08JAN2004	31DEC2009
28	G-3781_WELL_km	2185	5.55256	0.68331	08JAN2004	31DEC2009
29	G-3784_WELL_km	2185	5.47504	0.88010	08JAN2004	31DEC2009
30	G-3785_WELL_km	2185	5.31484	0.85357	08JAN2004	31DEC2009
31	G-3786_WELL_km	2185	5.39274	0.88575	08JAN2004	31DEC2009
32	G-3787_WELL_km	2150	5.47171	0.93453	12FEB2004	31DEC2009
33	G-551_G_WELL_km(	3653	-0.26451	1.58881	01JAN2000	31DEC2009
34	G-613-B_WELL_km	3104	2.34240	0.52746	01JAN2000	30JUN2008
35	G-757A_WELL_km	3653	4.14208	0.69725	01JAN2000	31DEC2009
36	G-789_G_WELL_km	3653	3.74087	0.71967	01JAN2000	31DEC2009
37	G-855_G_WELL_km	3653	4.48675	0.72022	01JAN2000	31DEC2009
38	G-860_G_WELL_km	3104	2.58745	0.39939	01JAN2000	30JUN2008
39	G-864A_WELL_km	3653	2.41284	0.80089	01JAN2000	31DEC2009
40	G-864_G_WELL_km	3653	2.53306	0.78174	01JAN2000	31DEC2009

Stage data

CODE	Description	Frequency	Percent	Cumulative Frequency	Cumulative Percent
!	Normal limits exceeded	94	0.31	94	0.31
<	Less than	137	0.46	231	0.77
E	Estimated	23443	77.88	23674	78.65
M	Missing	6246	20.75	29920	99.40
N	Data not processed	144	0.48	30064	99.87
P	Partial record	38	0.13	30102	100.00

No code = 313512

Obs	site	Freq	mean_Stage	std_Stage	mindate	maxdate
1	G422_H_Stage	1404	3.40772	0.47204	27FEB2006	31DEC2009
2	G422_T_Stage	1404	3.54083	0.51487	27FEB2006	31DEC2009
3	S332CS_T_Stage	1444	6.11134	0.80181	18JAN2006	31DEC2009
4	C4SW1_Stage	1535	4.66827	0.73934	19OCT2005	31DEC2009
5	S21AV_Stage	1626	1.77517	0.28213	20JUL2005	31DEC2009
6	C4SW2_Stage	1711	4.28849	0.65359	26APR2005	31DEC2009
7	C4SW3_Stage	1711	4.78371	0.90409	26APR2005	31DEC2009
8	S22V_Stage	1711	2.81680	0.37471	26APR2005	31DEC2009
9	S123V_Stage	1751	2.58596	0.38919	17MAR2005	31DEC2009
10	S21V_Stage	1751	2.01609	0.23915	17MAR2005	31DEC2009
11	L31E_H_Stage	1766	1.57722	0.27263	01JAN2000	31OCT2004
12	L31E_T_Stage	1766	1.56180	0.18551	01JAN2000	31OCT2004
13	G420_H_Stage	1845	3.49632	0.53064	13DEC2004	31DEC2009
14	G420_T_Stage	1845	3.94095	0.77099	13DEC2004	31DEC2009
15	S25BM_H_Stage	1940	2.26363	0.47325	09SEP2004	31DEC2009
16	S25BM_T_Stage	1940	1.08719	0.45859	09SEP2004	31DEC2009
17	S380_H_Stage	2015	4.23418	0.69663	26JUN2004	31DEC2009
18	S380_T_Stage	2015	3.72513	0.62523	26JUN2004	31DEC2009
19	L31NN_Stage	2044	6.16287	1.03886	28MAY2004	31DEC2009
20	L31NS_Stage	2044	5.81305	1.09854	28MAY2004	31DEC2009
21	S332BN_T_Stage	2107	6.40486	1.35552	26MAR2004	31DEC2009
22	D53_Stage	2291	6.14054	1.16043	24SEP2003	31DEC2009
23	MRMS1_Stage	2300	0.89949	0.43451	15SEP2003	31DEC2009
24	MRMS4_Stage	2301	0.93012	0.42402	14SEP2003	31DEC2009
25	C2GW1_Stage	2603	0.46397	2.33118	16NOV2002	31DEC2009
26	C2SW1_Stage	2603	3.13090	0.31245	16NOV2002	31DEC2009
27	C2SW2_Stage	2603	2.88850	0.33101	16NOV2002	31DEC2009
28	G69_H_Stage	2630	7.27633	0.58560	01JAN2000	14MAR2007
29	G-596_B_Stage	2738	5.17602	0.65365	01JAN2000	30JUN2007
30	G-618_B_Stage	2738	6.96469	0.51152	01JAN2000	30JUN2007

Obs	site	Freq	mean_Stage	std_Stage	mindate	maxdate
31	G1502_Stage	2738	6.31533	0.87501	01JAN2000	30JUN2007
32	NP_RG1_Stage	2738	5.59251	0.98472	01JAN2000	30JUN2007
33	NP_RG2_Stage	2738	5.25297	1.02524	01JAN2000	30JUN2007
34	S123_Stage	2738	2.51336	0.37574	01JAN2000	30JUN2007
35	S332D_H_Stage	3305	4.35535	0.47540	14DEC2000	31DEC2009
36	S332D_T_Stage	3305	5.80718	2.43454	14DEC2000	31DEC2009
37	S332B_H_Stage	3318	4.43379	0.47020	01DEC2000	31DEC2009
38	S332B_T_Stage	3318	6.41354	1.30614	01DEC2000	31DEC2009
39	3BS1W1_H_Stage	3524	6.66524	0.90190	09MAY2000	31DEC2009
40	3B-SE_B_Stage	3653	6.65860	0.97252	01JAN2000	31DEC2009
41	C4_CORAL_Stage	3653	2.86500	0.30254	01JAN2000	31DEC2009
42	C6_NW36_Stage	3653	2.31662	0.40214	01JAN2000	31DEC2009
43	G119_H_Stage	3653	5.77530	0.67490	01JAN2000	31DEC2009
44	G119_T_Stage	3653	4.32080	0.68660	01JAN2000	31DEC2009
45	G211_H_Stage	3653	5.45664	0.56686	01JAN2000	31DEC2009
46	G211_T_Stage	3653	4.56850	0.46319	01JAN2000	31DEC2009
47	G334_H_Stage	3653	6.95818	0.48865	01JAN2000	31DEC2009
48	G334_T_Stage	3653	5.44073	0.59373	01JAN2000	31DEC2009
49	G93_H_Stage	3653	2.50142	0.37271	01JAN2000	31DEC2009
50	G93_T_Stage	3653	0.95930	0.46541	01JAN2000	31DEC2009
51	L31NT_Stage	3653	4.41126	0.47945	01JAN2000	31DEC2009
52	NESRS2_Stage	3653	6.69857	0.62746	01JAN2000	31DEC2009
53	NESRS3_Stage	3653	6.68373	0.92266	01JAN2000	31DEC2009
54	S118_H_Stage	3653	3.27754	0.55394	01JAN2000	31DEC2009
55	S118_T_Stage	3653	2.51924	0.37785	01JAN2000	31DEC2009
56	S119_H_Stage	3653	3.11402	0.56728	01JAN2000	31DEC2009
57	S119_T_Stage	3653	2.57138	0.37600	01JAN2000	31DEC2009
58	S120_H_Stage	3653	2.99188	0.59733	01JAN2000	31DEC2009
59	S120_T_Stage	3653	2.55111	0.38281	01JAN2000	31DEC2009
60	S121_H_Stage	3653	3.03858	0.35877	01JAN2000	31DEC2009

Obs	site	Freq	mean_Stage	std_Stage	mindate	maxdate
61	S121_T_Stage	3653	3.11495	0.57965	01JAN2000	31DEC2009
62	S122_H_Stage	3653	2.49152	0.35413	01JAN2000	31DEC2009
63	S122_T_Stage	3653	1.94481	0.28300	01JAN2000	31DEC2009
64	S123_H_Stage	3653	2.51799	0.38423	01JAN2000	31DEC2009
65	S123_T_Stage	3653	0.86711	0.40258	01JAN2000	31DEC2009
66	S148_H_Stage	3653	3.89905	0.69770	01JAN2000	31DEC2009
67	S148_T_Stage	3653	2.06758	0.23821	01JAN2000	31DEC2009
68	S149_H_Stage	3653	2.72430	0.50539	01JAN2000	31DEC2009
69	S165_H_Stage	3653	3.54899	0.66041	01JAN2000	31DEC2009
70	S165_T_Stage	3653	1.82795	0.32014	01JAN2000	31DEC2009
71	S166_H_Stage	3653	2.98006	0.61879	01JAN2000	31DEC2009
72	S166_T_Stage	3653	2.54462	0.48390	01JAN2000	31DEC2009
73	S167_H_Stage	3653	3.45237	0.76341	01JAN2000	31DEC2009
74	S167_T_Stage	3653	2.58099	0.50403	01JAN2000	31DEC2009
75	S174_H_Stage	3653	4.37508	0.47002	01JAN2000	31DEC2009
76	S174_T_Stage	3653	4.45333	1.14577	01JAN2000	31DEC2009
77	S176_H_Stage	3653	4.35422	0.47136	01JAN2000	31DEC2009
78	S176_T_Stage	3653	3.48343	0.55428	01JAN2000	31DEC2009
79	S194_H_Stage	3653	4.33890	0.52031	01JAN2000	31DEC2009
80	S194_T_Stage	3653	3.78768	0.70925	01JAN2000	31DEC2009
81	S195_H_Stage	3653	2.60230	0.44493	01JAN2000	31DEC2009
82	S195_T_Stage	3653	1.76889	0.36060	01JAN2000	31DEC2009
83	S196_H_Stage	3653	4.29725	0.47429	01JAN2000	31DEC2009
84	S196_T_Stage	3653	3.54290	0.77620	01JAN2000	31DEC2009
85	S20G_H_Stage	3653	1.70096	0.30201	01JAN2000	31DEC2009
86	S20G_T_Stage	3653	0.83031	0.41112	01JAN2000	31DEC2009
87	S21A_H_Stage	3653	1.71101	0.28256	01JAN2000	31DEC2009
88	S21A_T_Stage	3653	0.81601	0.41259	01JAN2000	31DEC2009
89	S21_H_Stage	3653	1.99319	0.24838	01JAN2000	31DEC2009
90	S21_T_Stage	3653	0.88775	0.40443	01JAN2000	31DEC2009

Obs	site	Freq	mean_Stage	std_Stage	mindate	maxdate
91	S22_H_Stage	3653	2.82225	0.40382	01JAN2000	31DEC2009
92	S22_T_Stage	3653	0.82227	0.41479	01JAN2000	31DEC2009
93	S25B_H_Stage	3653	2.27949	0.47817	01JAN2000	31DEC2009
94	S25B_T	3653	1.05577	0.45266	01JAN2000	31DEC2009
95	S25_H_Stage	3653	1.84553	0.23356	01JAN2000	31DEC2009
96	S25_T_Stage	3653	0.98896	0.43937	01JAN2000	31DEC2009
97	S26_H_Stage	3653	2.30360	0.40974	01JAN2000	31DEC2009
98	S26_T_Stage	3653	1.01302	0.43109	01JAN2000	31DEC2009
99	S331_H_Stage	3653	4.56650	0.47822	01JAN2000	31DEC2009
100	S331_T_Stage	3653	4.56203	0.55049	01JAN2000	31DEC2009
101	S335_H_Stage	3653	6.08528	0.74986	01JAN2000	31DEC2009
102	S335_T_Stage	3653	5.48009	0.58836	01JAN2000	31DEC2009
103	S336_H_Stage	3653	5.47832	0.59831	01JAN2000	31DEC2009
104	S336_T_Stage	3653	5.67857	0.68530	01JAN2000	31DEC2009
105	S338_H_Stage	3653	5.35897	0.53334	01JAN2000	31DEC2009
106	S338_T_Stage	3653	4.33177	0.76731	01JAN2000	31DEC2009
107	SHARK1_H_Stage	3653	7.44363	0.61150	01JAN2000	31DEC2009
108	T5_H_Stage	3653	3.20226	0.39903	01JAN2000	31DEC2009
109	s179_T_Stage	3653	1.66796	0.29745	01JAN2000	31DEC2009

Stage July19 data

Obs	site	Freq	mean_stage	std_stage	mindate	maxdate
1	NP-112_Stage	1403	4.19530	1.05035	01JAN2000	03NOV2003
2	NP-C27_Stage	1403	2.07674	0.48363	01JAN2000	03NOV2003
3	NP-C50_Stage	1403	1.86482	0.49893	01JAN2000	03NOV2003
4	NP-CV5N_Stage	1403	1.92947	0.50888	01JAN2000	03NOV2003
5	NP-EV8_Stage	1403	1.92522	0.47948	01JAN2000	03NOV2003
6	NP-ROB_Stage	1403	3.05035	0.70912	01JAN2000	03NOV2003
7	NP-TS2_Stage	1403	3.45714	1.08070	01JAN2000	03NOV2003
8	S20FNV_Stage	1612	1.71224	0.29177	03AUG2005	31DEC2009
9	S20FWV_Stage	1626	1.70930	0.30386	20JUL2005	31DEC2009
10	S20FSV_Stage	1631	1.75709	0.30205	15JUL2005	31DEC2009
11	BERM3_H_Stage	2234	3.23831	0.95041	20NOV2003	31DEC2009
12	BERM3_T_Stage	2234	3.28089	0.86009	20NOV2003	31DEC2009
13	NP-31W_Stage	2738	4.34472	1.08530	01JAN2000	30JUN2007
14	NP-EP1_Stage	2738	1.64888	0.44924	01JAN2000	30JUN2007
15	NP-EPS_Stage	2738	1.47352	0.33084	01JAN2000	30JUN2007
16	NP-EV6_Stage	2738	1.99826	0.44502	01JAN2000	30JUN2007
17	NP-EV7_Stage	2738	2.12856	0.42534	01JAN2000	30JUN2007
18	NP-N10_Stage	2738	4.57341	1.13463	01JAN2000	30JUN2007
19	NP-NTS1_Stage	2738	4.32008	1.09061	01JAN2000	30JUN2007
20	NP-TSB_Stage	2738	3.29692	1.10887	01JAN2000	30JUN2007
21	SWEVER5A_Stage	3266	1.09155	0.49504	01JAN2000	09DEC2008
22	SWEVER5B_Stage	3266	1.46161	0.22994	01JAN2000	09DEC2008
23	JBTS_Stage	3653	0.93764	0.42533	01JAN2000	31DEC2009
24	MBTS_Stage	3653	3.16507	0.41844	01JAN2000	31DEC2009
25	MDTS_Stage	3653	2.60322	0.42429	01JAN2000	31DEC2009
26	S175_H_Stage	3653	4.35178	1.14665	01JAN2000	31DEC2009
27	S175_T_Stage	3653	2.52372	0.58927	01JAN2000	31DEC2009
28	S177_H_Stage	3653	3.45267	0.51361	01JAN2000	31DEC2009
29	S177_T_Stage	3653	2.21755	0.45035	01JAN2000	31DEC2009

Obs	site	Freq	mean_stage	std_stage	mindate	maxdate
30	S178_H_Stage	3653	2.46319	0.55574	01JAN2000	31DEC2009
31	S178_T_Stage	3653	2.22356	0.43192	01JAN2000	31DEC2009
32	S179_H_Stage	3653	2.55952	0.47246	01JAN2000	31DEC2009
33	S18C_H_Stage	3653	2.20573	0.41880	01JAN2000	31DEC2009
34	S18C_T_Stage	3653	1.89847	0.54004	01JAN2000	31DEC2009
35	S197_H_Stage	3653	1.92920	0.53237	01JAN2000	31DEC2009
36	S197_T_Stage	3653	0.99615	0.41522	01JAN2000	31DEC2009
37	S20F_H_Stage	3653	1.61551	0.29270	01JAN2000	31DEC2009
38	S20F_T_Stage	3653	0.81127	0.40944	01JAN2000	31DEC2009
39	S20_H_Stage	3653	1.69618	0.42733	01JAN2000	31DEC2009
40	S20_T_Stage	3653	1.05151	0.52812	01JAN2000	31DEC2009
41	S332_H_Stage	3653	4.42287	1.14499	01JAN2000	31DEC2009
42	S332_T_Stage	3653	4.51336	0.91092	01JAN2000	31DEC2009
43	SWEVER1_Stage	3653	1.54193	0.35052	01JAN2000	31DEC2009
44	SWEVER2A_Stage	3653	1.91860	0.53993	01JAN2000	31DEC2009
45	SWEVER3_Stage	3653	2.11598	0.48632	01JAN2000	31DEC2009

Rain data

CODE	Description	Frequency	Percent	Cumulative Frequency	Cumulative Percent
A	Accumulated	13	0.16	13	0.16
E	Estimated	3186	39.36	3199	39.52
M	Missing	4319	53.36	7518	92.88
N	Data not yet available	429	5.30	7947	98.18
P	Partial record	74	0.91	8021	99.10
X	Included in next amount marked "A"	73	0.90	8094	100.00

No code = 97189

Obs	site	Freq	mean_rain	std_rain	mindate	maxdate
1	S20_Rain	2999	0.12718	0.45791	01JAN2000	17MAR2008
2	HOMESFS_rain	3653	0.14048	0.49321	01JAN2000	31DEC2009
3	JBTS_rain	3653	0.08430	0.35178	01JAN2000	31DEC2009
4	MBTS_rain	3653	0.09823	0.38951	01JAN2000	31DEC2009
5	MDTS_rain	3653	0.10770	0.42641	01JAN2000	31DEC2009
6	PERRINE_Rain	3653	0.16192	0.48362	01JAN2000	31DEC2009
7	S123_rain	3653	0.11863	0.42282	01JAN2000	31DEC2009
8	S165_Rain	3653	0.13521	0.45125	01JAN2000	31DEC2009
9	S167_Rain	3653	0.13541	0.43687	01JAN2000	31DEC2009
10	S174_Rain	3653	0.13992	0.42764	01JAN2000	31DEC2009
11	S177_Rain	3653	0.12385	0.41217	01JAN2000	31DEC2009
12	S179_Rain	3653	0.14406	0.51021	01JAN2000	31DEC2009
13	S18C_Rain	3653	0.16408	0.53703	01JAN2000	31DEC2009
14	S20F_Rain	3653	0.12063	0.39140	01JAN2000	31DEC2009
15	S20G_Rain	3653	0.12724	0.42663	01JAN2000	31DEC2009
16	S21A_Rain	3653	0.11843	0.41156	01JAN2000	31DEC2009
17	S21_Rain	3653	0.10968	0.38639	01JAN2000	31DEC2009
18	S26_rain	3653	0.15513	0.47096	01JAN2000	31DEC2009

CODE	Description	Frequency	Percent	Cumulative Frequency	Cumulative Percent	
19	S331W_rain	3653	0.15677	0.49667	01JAN2000	31DEC2009
20	S331_Rain	3653	0.12489	0.41322	01JAN2000	31DEC2009
21	S332_Rain	3653	0.13088	0.40591	01JAN2000	31DEC2009
22	S334_rain	3653	0.16673	0.50793	01JAN2000	31DEC2009
23	S335_rain	3653	0.16673	0.50793	01JAN2000	31DEC2009
24	S336_Rain	3653	0.13599	0.40999	01JAN2000	31DEC2009
25	S338_Rain	3653	0.13205	0.40593	01JAN2000	31DEC2009
26	TAMI_AIR_rain	3653	0.14289	0.44576	01JAN2000	31DEC2009
27	miamai.ap_rain	3653	0.17298	0.51191	01JAN2000	31DEC2009
28	miami_2_Rain	3653	0.14259	0.44435	01JAN2000	31DEC2009
29	tpts_rain	3653	0.11873	0.43534	01JAN2000	31DEC2009

Flow data

Code	Description	Frequency	Percent	Cumulative Frequency	Cumulative Percent
<	Less than	4	0.03	4	0.03
>	Greater than	1	0.01	5	0.04
E	Estimated	9345	68.04	9350	68.08
M	Missing	4223	30.75	13573	98.83
N	Data not yet processed	122	0.89	13695	99.72
P	Partial record	39	0.28	13734	100.00

No code = 58555

Obs	site	Freq	mean_flow	std_flow	mindate	maxdate
1	G422_P_Flow_SK	1404	0.663	5.152	27FEB2006	31DEC2009
2	S21V_I_Flow_SK_R	1654	250.537	363.423	22JUN2005	31DEC2009
3	S25B_P_FLOW_SK	1979	3.893	29.189	01AUG2004	31DEC2009
4	S332B1_P_Flow_SK	2107	60.498	89.762	26MAR2004	31DEC2009
5	S332B2_P_FLOW_SK	3044	61.251	79.779	01SEP2001	31DEC2009
6	C4.CORAL_Flow_SK	3653	195.863	101.113	01JAN2000	31DEC2009
7	C6.NW36flow	3653	183.402	211.876	01JAN2000	31DEC2009
8	S118_S_FLOW_SK	3653	12.592	72.405	01JAN2000	31DEC2009
9	S119_S_FLOW_SK	3653	9.499	37.202	01JAN2000	31DEC2009
10	S120_C_FLOW_SK	3653	7.099	36.701	01JAN2000	31DEC2009
11	S121_C_Flow_SK	3653	0.166	4.161	01JAN2000	31DEC2009
12	S123_S_FLOW_SK	3653	41.935	152.930	01JAN2000	31DEC2009
13	S148_S_Flow_SK	3653	173.180	248.468	01JAN2000	31DEC2009
14	S178__C_Flow_SK	3653	-4.783	32.112	01JAN2000	31DEC2009
15	S178__S_Flow_SK	3653	-3.808	20.305	01JAN2000	31DEC2009
16	S178__W_Flow_SK	3653	0.006	0.186	01JAN2000	31DEC2009
17	S195_C_Flow_sk	3653	9.485	45.016	01JAN2000	31DEC2009

18	S20G_S_Flow_SK	3653	20.184	51.658	01JAN2000	31DEC2009
19	S20_S_Flow_SK	3653	12.895	55.887	01JAN2000	31DEC2009
20	S20_S_Flow_SKa	3653	12.895	55.887	01JAN2000	31DEC2009
21	S25B_S_FLOW_SK_R	3653	257.283	236.114	01JAN2000	31DEC2009
22	S25_C_Flow_SKREV	3653	10.421	21.440	01JAN2000	31DEC2009

Flow data for additional sites

code	Description	Frequency	Percent	Cumulative Frequency	Cumulative Percent
E	Estimated	16971	99.99	16971	99.99
P	Partial record	1	0.01	16972	100.00

No code = 89169

Obs	site	FREQ	mean_flow	std_flow	mindate	maxdate
1	G420_P	383	3.9039765013	27.598700416	14DEC2004	31DEC2005
2	G421_S	384	0.1566692708	1.4394705229	13DEC2004	31DEC2005
3	S26_P	1813	2.0334225041	19.473536765	14JAN2005	31DEC2009
4	S332B_	2192	117.88774498	143.33013951	01JAN2000	31DEC2005
5	S197_C	3104	23.047021585	161.43496834	01JAN2000	30JUN2008
6	S331+S	3287	285.55588774	260.34589029	01JAN2000	31DEC2009
7	G211_C	3653	271.87904654	231.95895146	01JAN2000	31DEC2009
8	G93	3653	15.220852724	38.103520361	01JAN2000	31DEC2009
9	S165_S	3653	17.202182863	63.406315582	01JAN2000	31DEC2009
10	S166_S	3653	5.381689844	28.324498839	01JAN2000	31DEC2009
11	S167_S	3653	10.333117985	41.32192704	01JAN2000	31DEC2009
12	S173_C	3653	44.067447851	57.347656833	01JAN2000	31DEC2009
13	S174_S	3653	12.20290282	34.361627968	01JAN2000	31DEC2009
14	S175_C	3653	3.4065647413	30.798018194	01JAN2000	31DEC2009
15	S176_S	3653	41.597506159	90.228522464	01JAN2000	31DEC2009
16	S177_S	3653	164.13817082	211.89433861	01JAN2000	31DEC2009
17	S179_S	3653	40.210740214	147.6866738	01JAN2000	31DEC2009
18	S18C_S	3653	214.21861675	273.47928865	01JAN2000	31DEC2009
19	S194_C	3653	41.322895428	64.515800372	01JAN2000	31DEC2009
20	S196_C	3653	18.95560772	35.210025939	01JAN2000	31DEC2009
21	S20F_S	3653	208.67674186	262.41428998	01JAN2000	31DEC2009
22	S21A_S	3653	119.16445469	159.22441937	01JAN2000	31DEC2009

23	S21_S	3653	204.16835094	242.43017986	01JAN2000	31DEC2009
24	S22_S	3653	139.90754777	243.45496179	01JAN2000	31DEC2009
25	S26_S	3653	205.50025322	250.51146459	01JAN2000	31DEC2009
26	S331	3653	227.41288174	288.88536764	01JAN2000	31DEC2009
27	S332	3653	8.0329635916	61.619000512	01JAN2000	31DEC2009
28	S332D_	3653	152.54576622	182.9137812	01JAN2000	31DEC2009
29	S334_S	3653	98.280341637	177.08898935	01JAN2000	31DEC2009
30	S335_S	3653	144.94410758	117.38886421	01JAN2000	31DEC2009
31	S336_C	3653	1.5017829181	11.366719607	01JAN2000	31DEC2009
32	S338_C	3653	145.2435089	133.47055702	01JAN2000	31DEC2009

FlBay Salinity data

Obs	site	Freq	mean_salin	std_salin	minday	maxday
1	FLAB01	94	32.1418	4.2470	24JAN2000	24SEP2009
2	FLAB02	94	30.3026	5.0455	24JAN2000	24SEP2009
3	FLAB03	94	28.5886	5.4584	24JAN2000	24SEP2009
4	FLAB04	94	31.3252	4.2896	24JAN2000	24SEP2009
5	FLAB05	93	31.2472	4.6266	24JAN2000	01DEC2009
6	FLAB06	93	25.5006	7.6275	24JAN2000	01DEC2009
7	FLAB07	93	14.5275	12.6443	24JAN2000	01DEC2009
8	FLAB08	93	21.6208	9.2890	24JAN2000	01DEC2009
9	FLAB09	93	30.6663	5.2936	24JAN2000	01DEC2009
10	FLAB10	93	15.8341	11.4242	24JAN2000	01DEC2009
11	FLAB11	93	24.8731	7.3230	24JAN2000	01DEC2009
12	FLAB12	93	33.8430	6.9367	25JAN2000	07DEC2009
13	FLAB14	93	35.5035	7.3098	25JAN2000	07DEC2009
14	FLAB47	93	10.5559	8.2790	11JAN2000	03DEC2009

Diel Salinity data

Obs	site	Freq	mean_salin	std_salin	minday	maxday
1	BISC00	1218	30.2787	5.60573	15MAY2004	31DEC2007
2	BISC01	856	29.2461	6.99415	13JUL2004	31DEC2006
3	BISC06	861	32.0280	5.49648	15MAY2004	31DEC2007
4	BISC10	1192	36.1563	1.09948	01JAN2004	23JUL2007
5	BISC12	1159	32.7526	3.94115	07MAY2004	31DEC2007
6	BISC13	546	33.3169	4.70760	13JUL2004	31DEC2006
7	BISC14	1399	22.2515	8.80407	27JAN2004	31DEC2007
8	BISC16	1301	24.8708	8.11483	14MAY2004	31DEC2007
9	BISC18	999	32.2478	4.23633	07MAY2004	29MAY2007
10	BISC19	548	32.4603	4.73840	15JUL2004	31DEC2006
11	BISC20	1033	34.8710	3.28678	01JAN2004	13FEB2007
12	BISC22	1323	22.1198	7.99210	16APR2004	31DEC2007
13	BISC24	1308	23.1595	7.81000	14MAY2004	31DEC2007
14	BISC26	969	28.0905	6.67520	14MAY2004	08JAN2007
15	BISC28	1380	22.0778	8.16197	27JAN2004	31DEC2007
16	BISC30	1314	22.4188	7.88038	14MAY2004	31DEC2007
17	BISC32	1029	27.9192	6.75188	01JAN2004	08JAN2007
18	BISC34	853	31.3258	4.40500	07MAY2004	08JAN2007
19	BISC35	725	30.8894	4.65932	20JUL2004	31DEC2006
20	BISC36	846	34.8056	2.72536	07MAY2004	17APR2007
21	BISC37	670	34.1694	2.80516	20JUL2004	31DEC2006
22	BISC40	1393	21.2203	8.33554	27JAN2004	31DEC2007
23	BISC42	1306	21.2925	8.29950	14MAY2004	31DEC2007
24	BISC44	960	28.1401	5.78844	07MAY2004	19APR2007
25	BISC45	812	26.5882	6.55905	20JUL2004	31DEC2006
26	BISC46	1115	28.6969	6.00977	01JAN2004	19APR2007
27	BISC48	897	33.5342	3.13655	07MAY2004	19APR2007
28	BISC50	1231	24.6435	7.76595	28JAN2004	31DEC2007
29	BISC52	1303	24.6086	7.31710	14MAY2004	31DEC2007

Obs	site	Freq	mean_salin	std_salin	minday	maxday
30	BISC54	922	29.8699	4.78189	07MAY2004	19APR2007
31	BISC55	674	28.6077	5.74270	22JUL2004	31DEC2006
32	BISC56	1410	26.0004	7.61654	28JAN2004	31DEC2007
33	BISC58	1315	26.0896	7.41016	14MAY2004	31DEC2007
34	BISC60	867	30.0167	5.14031	16MAY2004	27FEB2007
35	BISC61	841	29.9257	4.74266	22JUL2004	31DEC2006
36	BISC62	1364	26.6249	6.58792	29JAN2004	31DEC2007
37	BISC64	1242	28.5507	5.83360	14MAY2004	31DEC2007
38	BISC66	908	31.8970	4.25831	01JAN2004	27FEB2007
39	BISC67	801	30.6065	4.33594	22JUL2004	31DEC2006
40	BISC68	544	32.9094	3.58489	01JAN2004	27FEB2007
41	BISC69	325	31.2946	4.76470	26AUG2004	25AUG2005
42	BISC70	907	34.0784	3.22156	01JUN2004	27FEB2007
43	BISC72	161	35.7147	2.31487	01JUN2004	04JAN2005

Biscayne Salinity data

Obs	site	Freq	mean_salin	std_salin	minday	maxday
1	AC01	109	32.5377	2.58072	03JAN2000	03AUG2009
2	AC02	109	30.9851	2.94510	03JAN2000	03AUG2009
3	AC03	77	16.1133	8.06691	06JAN2003	03AUG2009
4	AR01	112	25.6893	6.11434	05JAN2000	05AUG2009
5	AR03	109	0.5425	0.82881	05JAN2000	05AUG2009
6	BB01	86	28.4697	4.89737	03JAN2000	10SEP2007
7	BB02	112	27.0115	6.30444	03JAN2000	03AUG2009
8	BB03	110	27.7328	7.47985	03JAN2000	03AUG2009
9	BB04	112	31.7066	3.68276	03JAN2000	03AUG2009
10	BB05A	106	31.6847	4.74932	03APR2000	03AUG2009
11	BB06	109	34.1763	2.42548	03JAN2000	03AUG2009
12	BB07	108	33.9479	3.19707	03JAN2000	03AUG2009
13	BB09	112	31.4015	3.49686	03JAN2000	03AUG2009
14	BB10	109	32.0156	2.97715	03JAN2000	03AUG2009
15	BB11	109	29.1461	4.17666	03JAN2000	03AUG2009
16	BB14	112	30.6936	3.40529	03JAN2000	03AUG2009
17	BB15	110	30.7361	3.29395	03JAN2000	03AUG2009
18	BB16	109	31.9070	3.29413	04JAN2000	04AUG2009
19	BB17	112	31.6969	2.58933	04JAN2000	04AUG2009
20	BB18	109	32.3176	2.31901	04JAN2000	04AUG2009
21	BB19	109	31.1746	3.62173	04JAN2000	04AUG2009
22	BB22	112	33.0424	2.86331	04JAN2000	04AUG2009
23	BB23	110	32.8634	2.94270	04JAN2000	04AUG2009
24	BB24	112	34.5089	1.85500	04JAN2000	04AUG2009
25	BB25	109	34.3146	2.16141	04JAN2000	04AUG2009
26	BB26	109	32.8520	2.60775	04JAN2000	04AUG2009
27	BB27	112	32.9454	3.08984	06JAN2000	06AUG2009
28	BB28	109	34.2687	2.58240	06JAN2000	06AUG2009
29	BB29	112	32.7262	3.34819	06JAN2000	06AUG2009

Obs	site	Freq	mean_salin	std_salin	minday	maxday
30	BB31	109	33.4484	3.25794	06JAN2000	06AUG2009
31	BB32	109	33.9192	2.90386	06JAN2000	06AUG2009
32	BB34	112	31.4216	4.03053	06JAN2000	06AUG2009
33	BB35	109	34.6385	2.88779	06JAN2000	06AUG2009
34	BB36	112	35.0594	2.67998	06JAN2000	06AUG2009
35	BB37	109	35.3595	2.18182	06JAN2000	06AUG2009
36	BB38	112	35.3148	2.80506	05JAN2000	05AUG2009
37	BB39A	109	29.4322	5.67447	05APR2000	05AUG2009
38	BB41	111	34.1165	3.58340	05JAN2000	05AUG2009
39	BB44	111	34.7294	3.40558	05JAN2000	05AUG2009
40	BB45	108	34.0601	3.68526	05JAN2000	05AUG2009
41	BB47	111	33.8833	3.20879	05JAN2000	05AUG2009
42	BB48	108	32.5038	4.04484	05JAN2000	05AUG2009
43	BB50	111	31.3563	4.65138	05JAN2000	05AUG2009
44	BB51	112	28.5531	5.50657	05JAN2000	05AUG2009
45	BB52	105	25.7012	7.07728	06APR2000	06AUG2009
46	BB53	106	21.8500	8.69589	05APR2000	05AUG2009
47	BISC101	92	30.2431	5.99452	29JAN2000	12AUG2009
48	BISC102	92	28.7505	5.76494	29JAN2000	12AUG2009
49	BISC103	71	29.9779	5.61073	29JAN2000	20SEP2007
50	BISC104	80	35.8871	1.75836	28JAN2000	12JUN2008
51	BISC108	92	36.0545	1.36321	28JAN2000	23SEP2009
52	BISC109	71	36.0393	1.85339	28JAN2000	20SEP2007
53	BISC110	92	33.7047	3.88440	29JAN2000	12AUG2009
54	BISC111	91	36.2002	1.90235	28JAN2000	15JUL2009
55	BISC112	80	36.3217	1.23507	28JAN2000	12JUN2008
56	BISC113	71	36.2432	2.27249	28JAN2000	21SEP2007
57	BISC116	71	35.6647	2.16344	28JAN2000	21SEP2007
58	BISC121	92	34.8370	2.87876	28JAN2000	12AUG2009
59	BISC122	92	32.8430	5.05896	28JAN2000	12AUG2009

Obs	site	Freq	mean_salin	std_salin	minday	maxday
60	BISC123	92	34.5131	3.45489	28JAN2000	12AUG2009
61	BISC124	71	36.2840	2.00219	28JAN2000	21SEP2007
62	BISC126	71	34.6434	2.54118	28JAN2000	20SEP2007
63	BISC127	92	33.9803	3.44347	28JAN2000	23SEP2009
64	BISC128	92	33.0080	3.17294	28JAN2000	23SEP2009
65	BISC129	92	33.6286	2.63952	28JAN2000	23SEP2009
66	BISC130	92	33.8497	2.65453	28JAN2000	23SEP2009
67	BISC131	92	32.3922	2.66236	28JAN2000	23SEP2009
68	BISC132	71	31.3799	2.80977	28JAN2000	20SEP2007
69	BISC133	92	32.0576	3.30335	28JAN2000	23SEP2009
70	BISC134	91	33.7768	3.17280	28JAN2000	23SEP2009
71	BISC135	92	33.8992	3.52353	28JAN2000	12AUG2009
72	BL01	113	18.5978	9.29018	05JAN2000	05AUG2009
73	BL02	80	17.6839	9.59738	08JAN2003	05AUG2009
74	BL03	83	0.4295	0.92525	08JAN2003	05AUG2009
75	BL12	79	0.2856	0.03705	08JAN2003	08JUL2009
76	BS01	113	25.7307	6.80186	03JAN2000	03AUG2009
77	BS04	83	2.2729	2.90657	06JAN2003	03AUG2009
78	BS10	78	0.2940	0.01688	06JAN2003	03AUG2009
79	CD01A	83	27.0611	6.97352	09JAN2003	06AUG2009
80	CD02	80	0.4497	0.78437	09JAN2003	06AUG2009
81	CD05	83	0.2352	0.03755	09JAN2003	06AUG2009
82	CD09	80	0.2592	0.03214	09JAN2003	06AUG2009
83	CG01	111	27.9583	4.27845	06JAN2000	06AUG2009
84	CG07	80	0.2591	0.03203	09JAN2003	06AUG2009
85	FC03	80	0.3261	0.15111	08JAN2003	05AUG2009
86	FC15	83	0.2698	0.03363	08JAN2003	05AUG2009
87	GL02	70	15.3289	9.49551	08JAN2003	05AUG2009
88	GL03	68	0.5193	0.17812	08JAN2003	05AUG2009
89	LR01	113	23.2660	5.77625	03JAN2000	03AUG2009

Obs	site	Freq	mean_salin	std_salin	minday	maxday
90	LR03	80	18.4516	8.72045	06JAN2003	03AUG2009
91	LR06	82	0.3245	0.33329	06JAN2003	03AUG2009
92	LR10	77	0.4557	1.25172	06JAN2003	03AUG2009
93	MI01	78	19.9670	8.54499	08JAN2003	05AUG2009
94	MI02	106	1.6419	3.28880	05APR2000	05AUG2009
95	MI03	80	0.7239	1.82532	08JAN2003	05AUG2009
96	MR01	113	26.5682	4.98043	04JAN2000	04AUG2009
97	MR02	60	16.8970	7.02918	07JAN2003	11SEP2007
98	MR03	80	14.3427	7.11870	07JAN2003	04AUG2009
99	MR04	80	13.3910	7.46072	07JAN2003	04AUG2009
100	MR06	82	7.6836	7.75538	07JAN2003	04AUG2009
101	MR07	77	6.8805	7.28568	07JAN2003	04AUG2009
102	MR08	83	0.2978	0.04649	07JAN2003	04AUG2009
103	MR15	83	0.3208	0.04216	07JAN2003	04AUG2009
104	MW01	113	21.3165	9.48221	05JAN2000	05AUG2009
105	MW04	83	1.2372	1.49146	08JAN2003	05AUG2009
106	MW05	80	0.3153	0.11502	08JAN2003	05AUG2009
107	MW13	82	0.2894	0.04606	08JAN2003	05AUG2009
108	NO07	74	0.2816	0.02479	08JAN2003	04FEB2009
109	OL03	80	3.7526	4.27725	06JAN2003	03AUG2009
110	PR01	111	15.8032	9.66759	05JAN2000	05AUG2009
111	PR03	83	0.3760	0.33830	08JAN2003	05AUG2009
112	PR04A	80	0.3045	0.00732	08JAN2003	05AUG2009
113	PR08	83	0.2918	0.04166	08JAN2003	05AUG2009
114	SK01	109	20.6409	9.21693	03JAN2000	03AUG2009
115	SK02	114	1.8100	2.47244	03JAN2000	03AUG2009
116	SK09	83	0.2908	0.02813	06JAN2003	03AUG2009
117	SP01	113	23.1962	9.08153	06JAN2000	06AUG2009
118	SP04	83	0.2613	0.03570	09JAN2003	06AUG2009
119	SP08	80	0.2690	0.03200	09JAN2003	06AUG2009

Obs	site	Freq	mean_salin	std_salin	minday	maxday
120	TM02	78	4.7004	6.50719	07JAN2003	04AUG2009
121	TM03	79	0.2542	0.03559	07JAN2003	04AUG2009
122	TM08	77	0.2505	0.03834	07JAN2003	04AUG2009
123	WC02	82	6.5932	6.99966	07JAN2003	04AUG2009
124	WC03	80	4.5267	5.26689	07JAN2003	04AUG2009
125	WC04	80	1.1399	1.44934	07JAN2003	04AUG2009

Appendix 2. Data preparation, decisions and comments.

Well data

Problem: G-3786_well_km has a missing value as the first value. This will cause the column to be read as a nominal variable rather than a continuous variable. A value of zero replaced the missing value. A code of DEL was used to indicate that this value should be deleted. In the SAS program observations with code=DEL were deleted.

Problem: G-613-B_well_km – line 3106 was "query returned 3104 records". This line was deleted.

Problem: G_1486_well_km: the date 20011022 occurs twice, once in record 639 and once in 662. The values are the same, the first one is deleted.

Problem: G_551_well: the date code is of a different variety (EXCEL style). I added a column with the dates in the regular format and renamed. The last line has entry "query returned 3". This was deleted.

Well Data USGS

The drawdown files have date in the format "yyyymmdd hour". The other files have daily dates (no hour) so this created a problem with input. The column for date was copied and year, month, day, and hour were extracted and used to create day. This was not a date formatted column but will suffice for calculations.

USGS data from 7-19-2010

The sheet for 254917080143301 (next to last sheet) has data on several additional sites. This does not seem to have been processed. All other sheets consist of two sets of columns, one for the original data and the other for the cleaned data. However, this sheet seems to have six sites in one set and one in the other. The sites are 254943080121501, 254950080171202, 254950080180801, 254951080194901, 255008080161801 (a total of 21,960 records).

Flow Data

Two of the sites have data past 12-31-. The files were truncated to be consistent with other data.

S21V_I_Flow_SK_R	1779	250.537	363.423	22JUN2005	05MAY2010
S332B2_P_FLOW_SK	3170	62.115	81.013	01SEP2001	06MAY2010

One site S_122 had the same value (zero) for all times and was omitted.

Bay data

Some of the data sets have several observations for each data. These are values at different depths (for example, salinity at three depths). Means were computed to produce one observation per time.

Appendix 3. Summary of data files, program files and output files. YM indicates monthly data.

Task	Output file	SAS file or other file	Data file
Task1 : correlations by type	Various files	*varnames.sas* individual matrices are calculated from the import file	All programs use all.ssd (daily), ymdata.ssd (monthly)
Task2: general correlations	Kendall.xls Spearman.xls Pearson.xls	*Generate AllCorrelations3* *generate all correlations* *pairwise samplesize.sas* *YM/sample sizes.sas*	All.ssd, ymdata.ssd
Task 3 Lag	lagCorrs.doc	*Correlations with lags.sas* *YMcorrelations lags*	
Tasks 4,5 Wet-Dry season	KendallWD.xls SpearmanWD.xls Pearson WD.xls	*Generate all correlations3 wet dry* *YM/Generate all correlationswet dry.sas*	Allwetdry.ssd
Task 6 Distance	neighbors.doc	*macro var names11-2.xls* *latlong3.csv* *Neighbor corr 11-2.sas* *Neighbor CorrYM11*-2.sas	Generates dist.ssd, LL.ssd