

Data Processing and Analysis in Support of Development of Water Quality Indicators

Doug Miller
SMSP, CEOE at UD
For 04 December 2009
CIB STAC Meeting

Three Data Sets:

Citizens Monitoring Program, CISNET and DNREC's 305(b) Water Quality Inventory Report

Filename and Worksheet	Dates	Rows	Stations	Sampling Events	Columns	WQ Columns	Comments
Total_All_SDCorr_Final.xls	1/1/98 to 12/31/08	9029	122	9028	19	10	
CISNET Indicators.xls	3/17/99 to 10/19/01	932	932	931	35	21	Lat-Lon to be decoded
Inland Bays DNREC Data from 98 to present.xlsx	3/30/98 to 12/02/08	80100	38	≈2201	21	23	STORET data format ...
"Cast" Inland Bays DNREC Data from 98 to present.xlsx	3/30/98 to 12/02/08	≤2220	38	2220 ≈2220 w/ non-identical dups removed, 1912 w/ dups removed	34+	Up to 34	<i>Cast into wide format to match other data sets ✓</i>

Variable names

CMP

Site	site
Test Site Description	sitedesc
Site Type	sitetype
Lat	sitelat
Long	sitelon
Date	sampdate
Time	samptime
Air Temp °C	atemp
Water Temp °C	wtemp
DO mg/L	do
Secchi Depth m	secchi
Water Depth m	wdepth
Sal ppt	sal
pH	ph
TSS mg/L	tss
CHLa µg/L	chla
DIP µmol	dip
DIN µmol	din
Comments	comments
	datasource
	xlsflag
	recnum

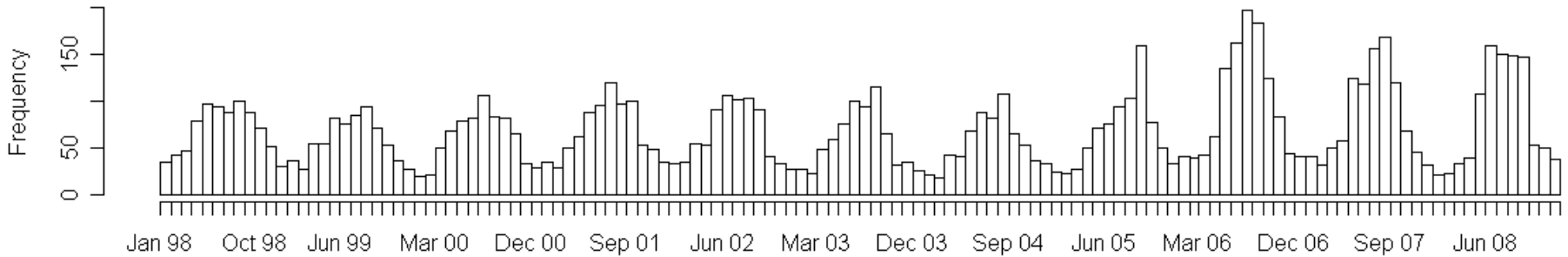
DNREC

Station ID	staid	staid
Station Name	staname	
Station Latitude	stalat	
Station Longitude	stalon	
Station Horizontal Datum	stadatum	
Visit Num	visitnum	
Activity ID	actid	
Activity Start	actstart	
Activity Start Zone	actstartzone	
Activity Type	acttype	
Activity Category-Rep Num	actcat	
Activity Depth	actdepth	
Activity Depth Unit	actdepthunit	
Characteristic Name	charname	
Sample Fraction	sampfrac	
Value Type	valtype	
Statistic Type	stattype	
Result Value as Text	resvalue	
Units	units	
Result Comment	rescomment	
Analytical Proc ID	analprocid	
	datasource	
	xlsflag	
	recnum	
Alkalinity, carbonate as CaCO3		wqp alka
Aluminum		aluminum
Arsenic		arsenic
Cadmium		cadmium
Carbonaceous biochemical oxygen demand, standard conditions		cbod
Chromium(VI)		chiv
Chlorophyll a		chla
Chloride		chloride
Copper		copper
Dissolved oxygen (DO)		do
Enterococcus		ecoccus
Flow		flow
Hardness, Ca, Mg		hardness
Iron		iron
Kjeldahl nitrogen		kjn
Lead		lead
Nitrogen, ammonia as N		nh4
Nitrogen, Nitrite (NO2) as N		no2
Nitrogen, Nitrite (NO2) + Nitrate (NO3) as N		no2no3
Nitrogen, Nitrate (NO3) as N		no3
Nutrient-nitrogen		nutrientn
Organic Nitrogen		organiczn
Organic carbon		orgc
Phosphorus, orthophosphate as P		orthop
Phosphorus as P		pasp
pH		ph
Pheophytin a		phaeophytin
Salinity		sal
Depth, Secchi disk depth		secchi
Specific conductance		spcond
Total suspended solids		tss
Turbidity		turbidity
Temperature, water		wtemp
Zinc		zinc

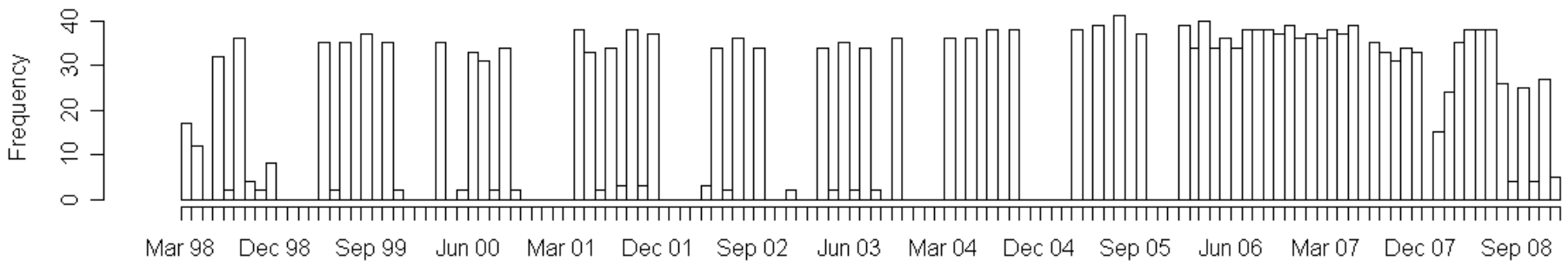
CISNET

BAY Sample	baysamp
Station Name	staname
Bay	bay
Tributary	trib
First_Sampled	firstsamp
SAMPLE TIME	samptime
Lat	samplatcoded
Lon	samploncoded
Comment	comment
Northing	sampnorth
Easting	sampeast
Station Name	staname2
Nitrate + Nitrite (uM)	no3no2
Phosphate (uM)	po4
Silicate (uM)	sio2
Ammonium (uM)	nh4
Diss Organic Carbon	doc
Persulfate N (uM)	persulfn
Persulfate P (uM)	persulfp
Particulate N	partn
Particulate C	partc
Particulate P	partp
Diss Organic N	don
Diss Organic P	dop
Salinity (ppt)	salppt
Station Name	staname3
Salinity	sal
Temp	temp
pH	ph
Diss O2 (mg/L)	disso2
Seston (mg/L)	stston
Chl a (ug/L)	chla
Secchi (m)	secchi
Total Depth (ft)	depthft
Field Notes	fieldnotes
	datasource
	xlsflag
	recnum

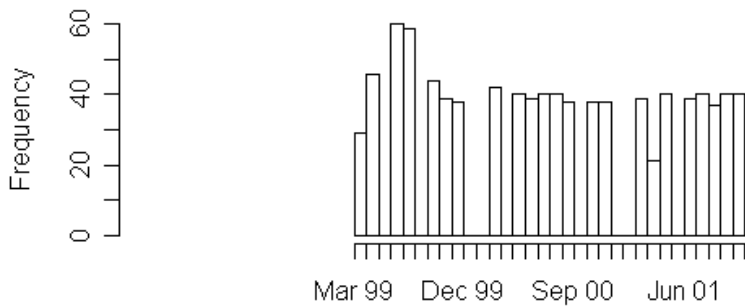
CMP



DNREC

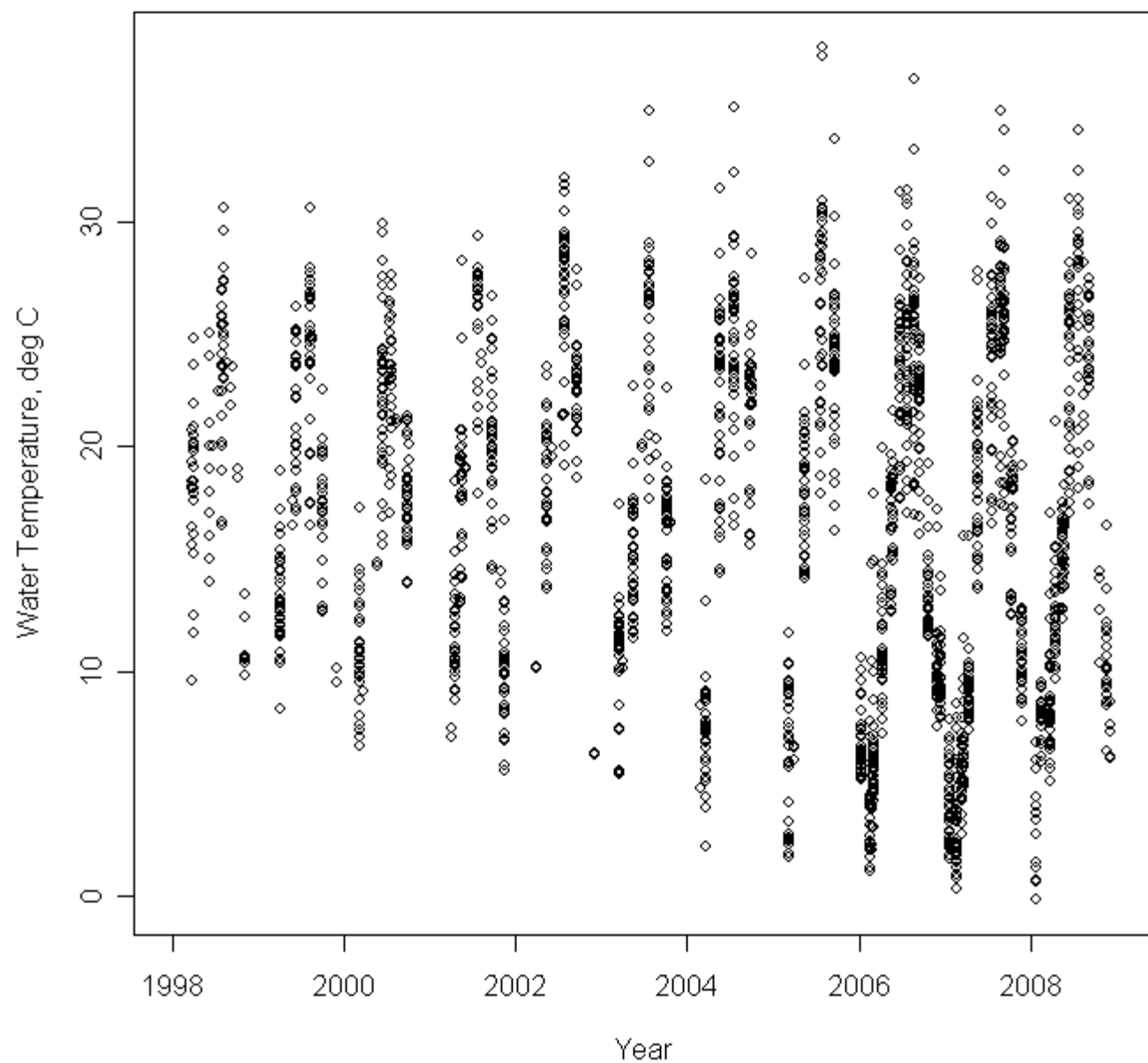


CISNET

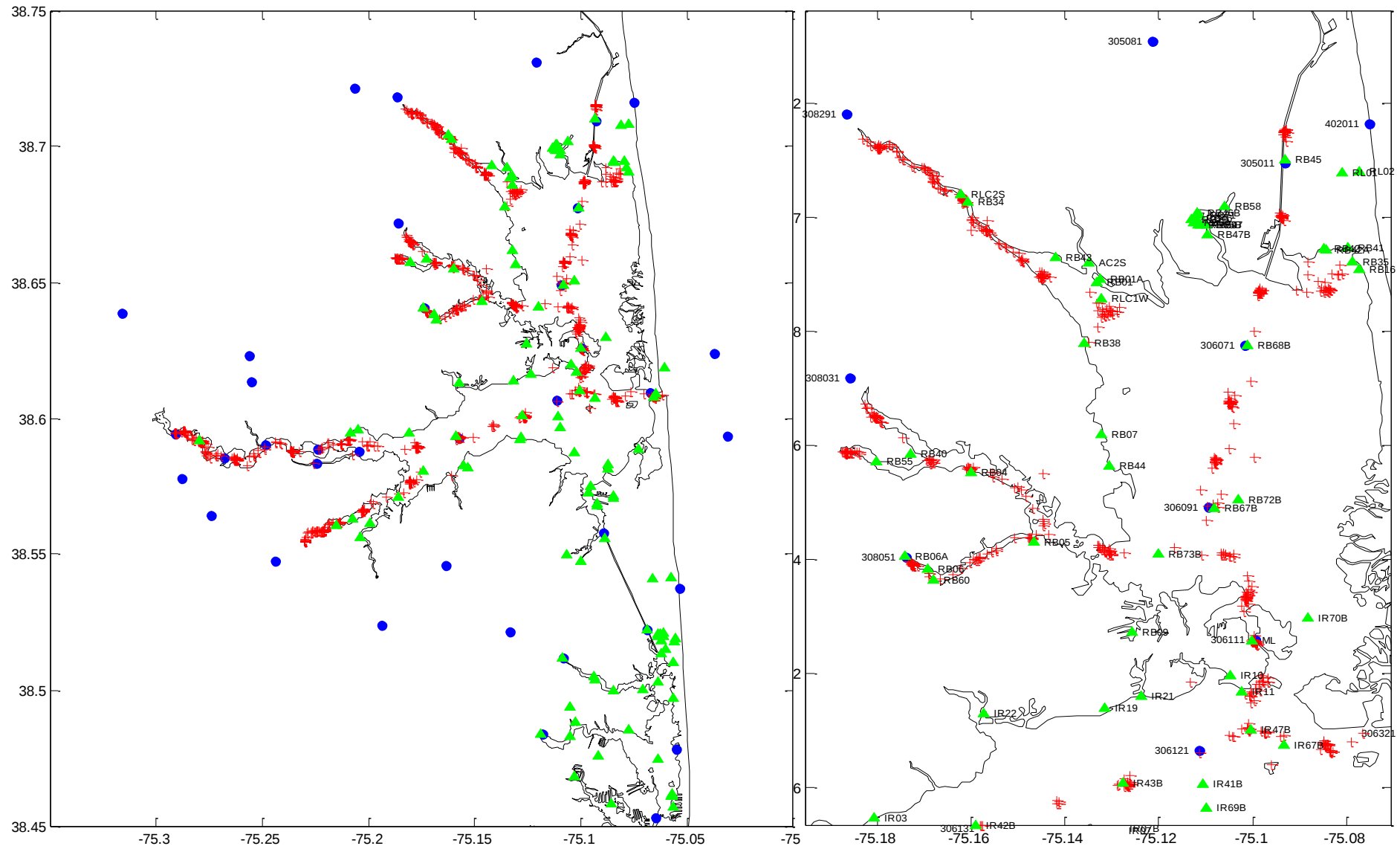


Sampling in Time ...

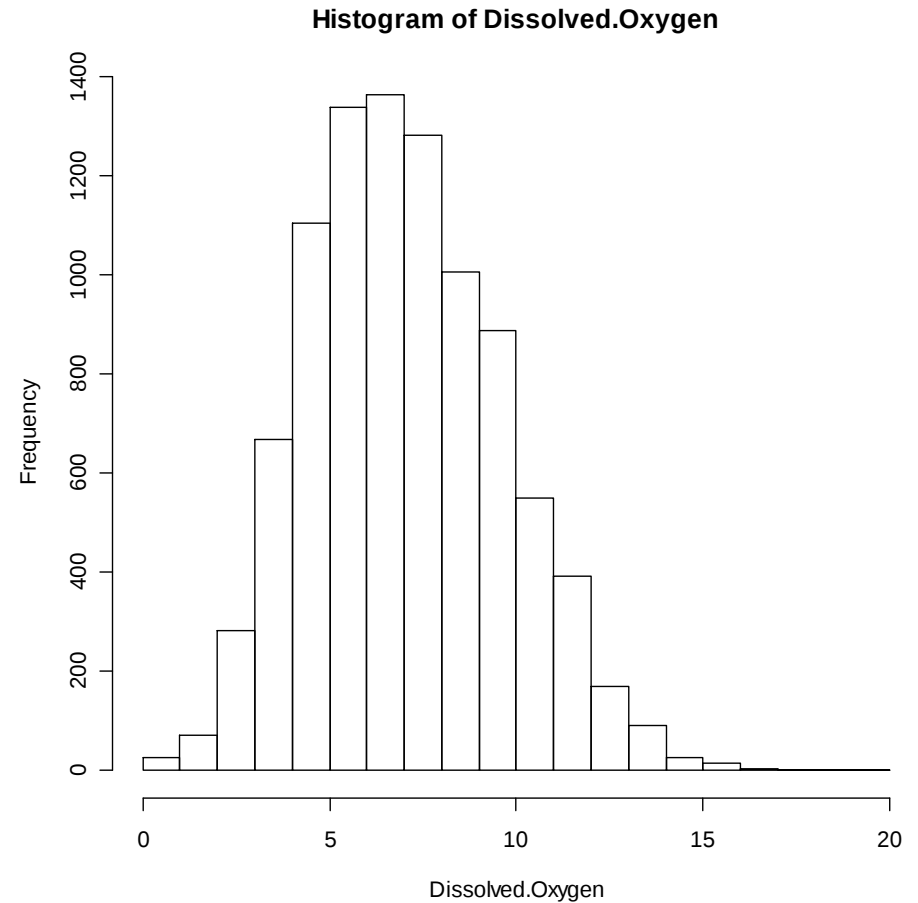
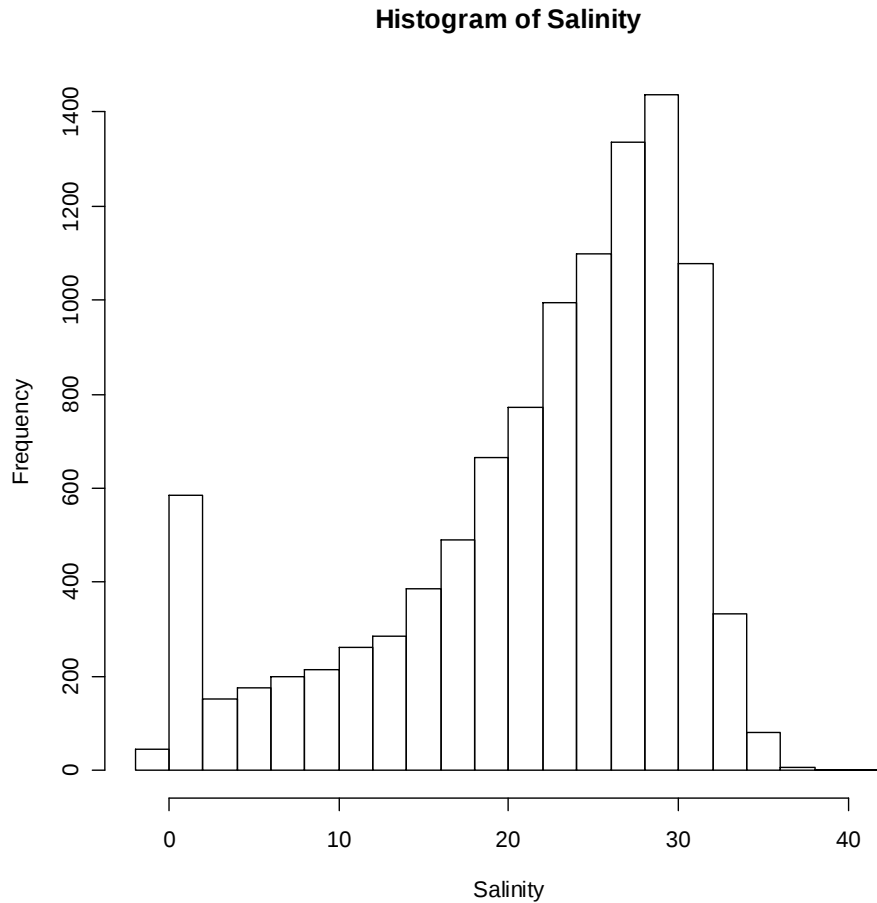
DNREC



Sampling Stations

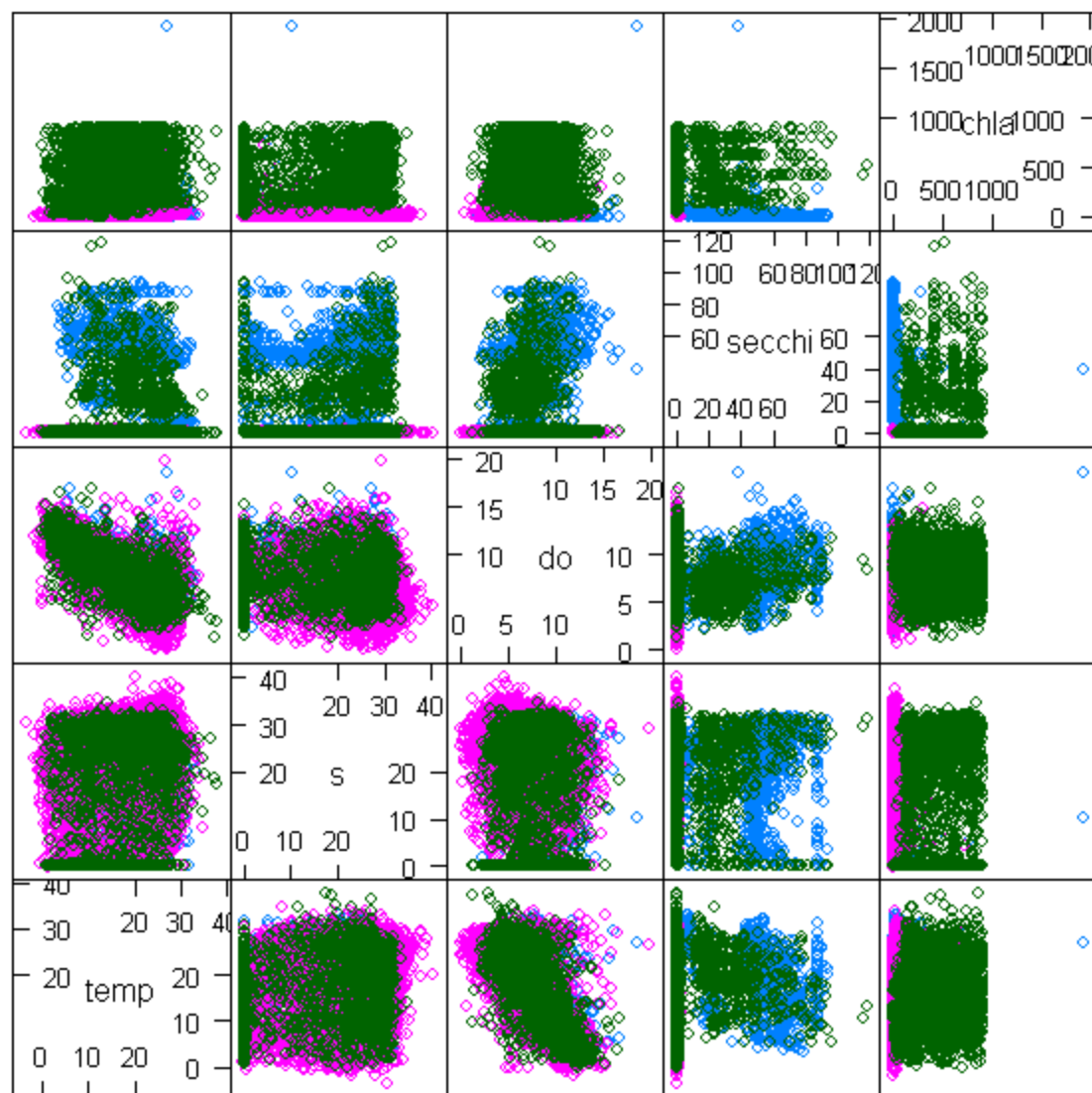


All (12,179), Salinity (10,598) and DO (9275)



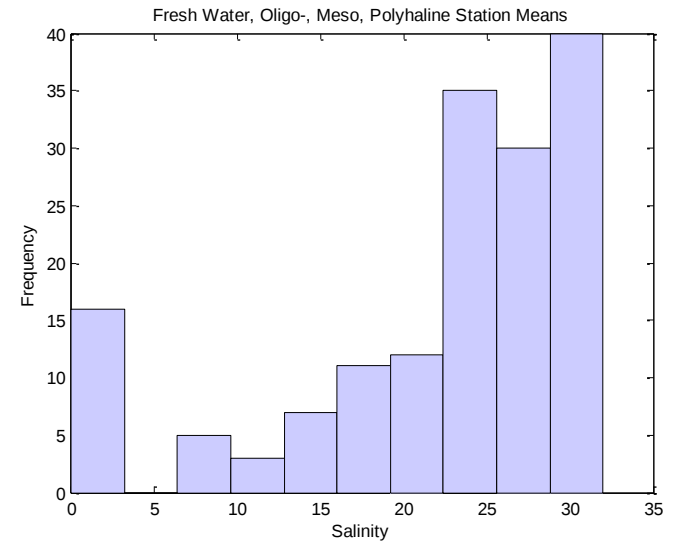
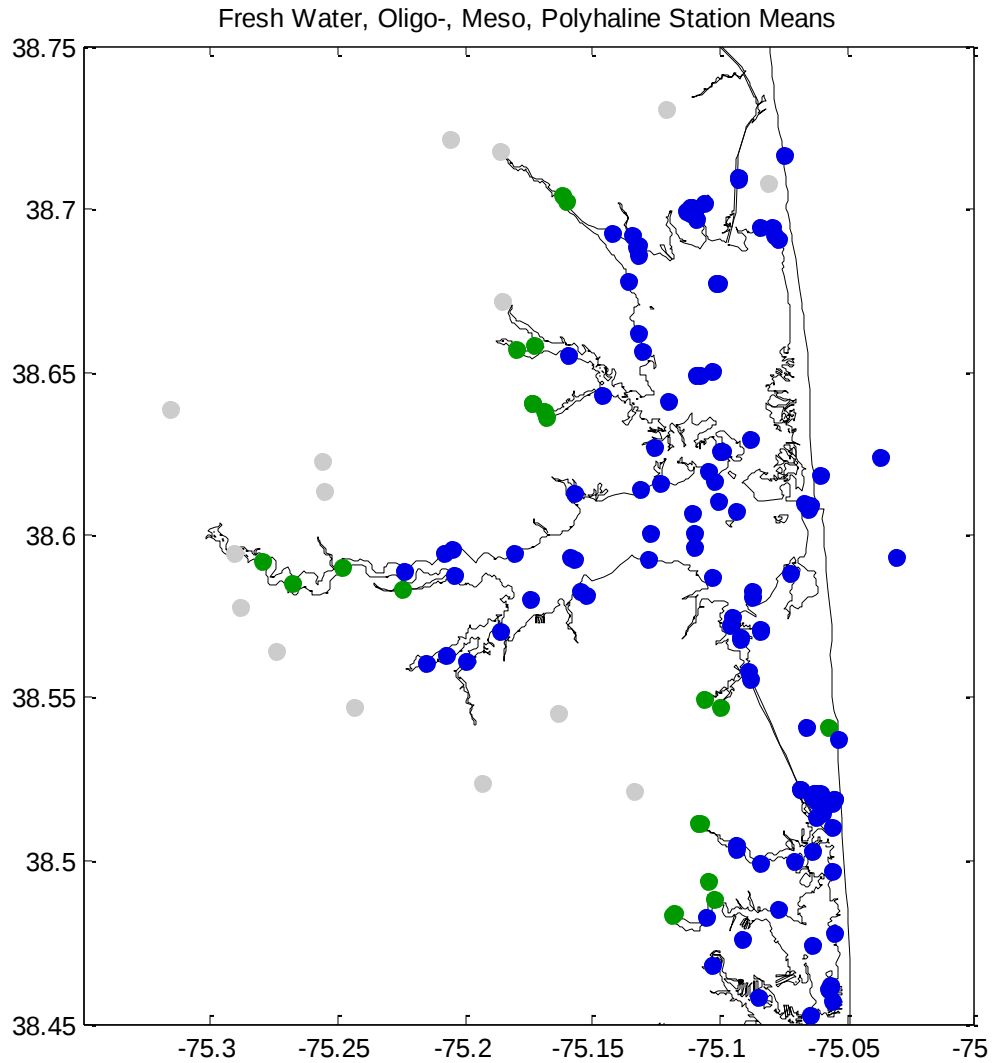
30,000' View

○ CMP ○ CISNET ○ DNREC

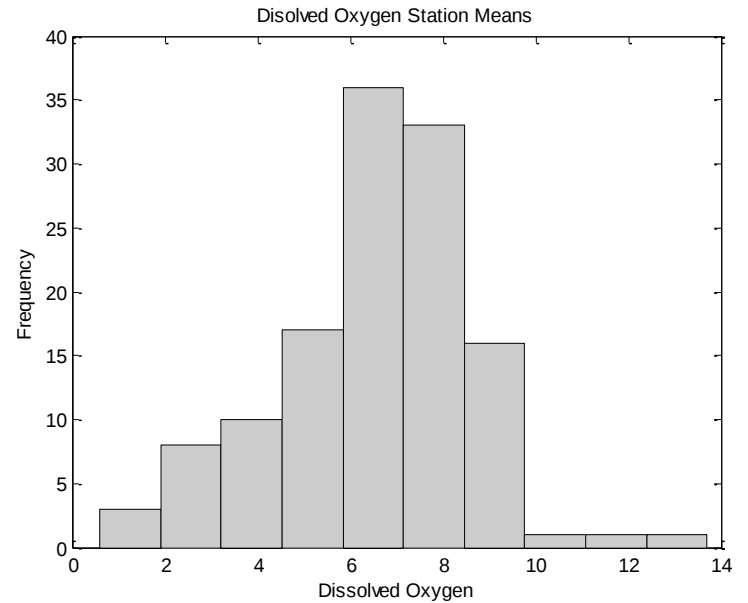
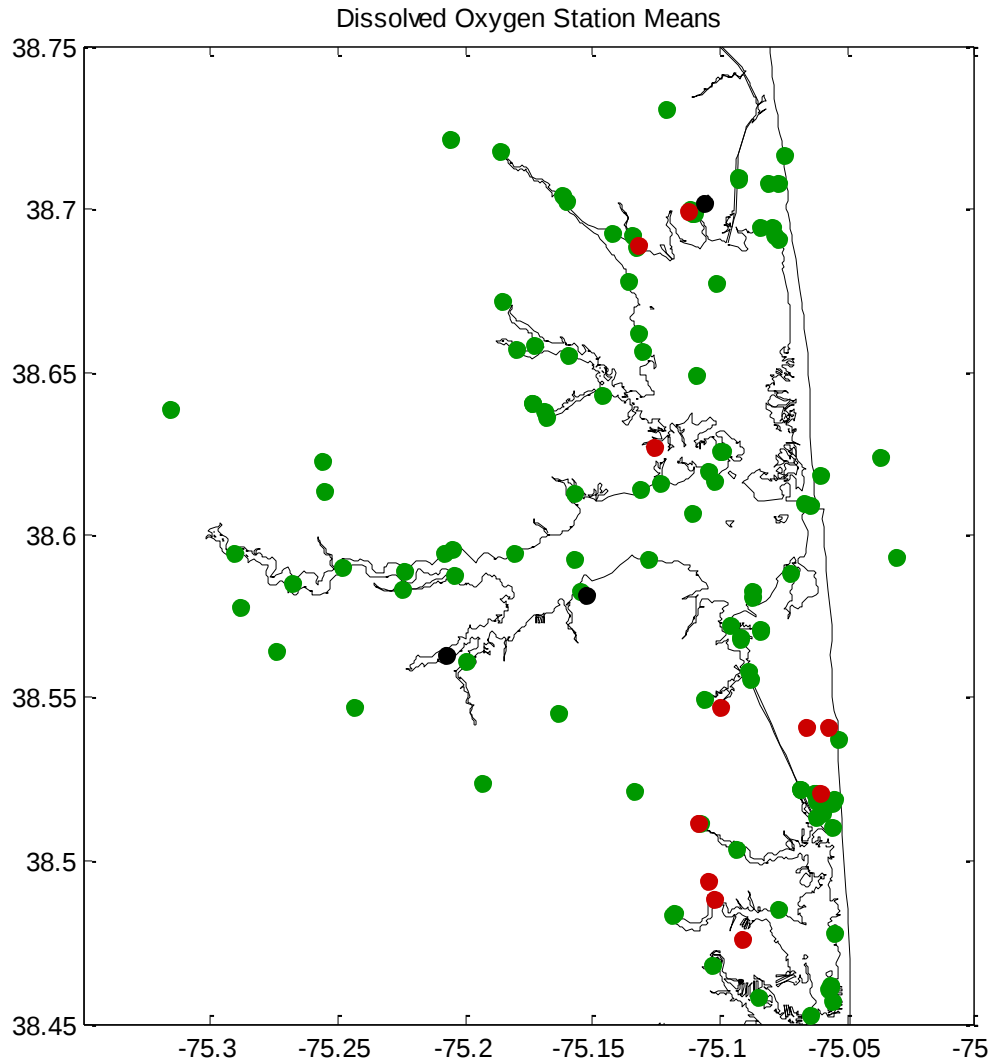


Scatter Plot Matrix

Fresh, Oligo-, Meso- and Polyhaline Station Means

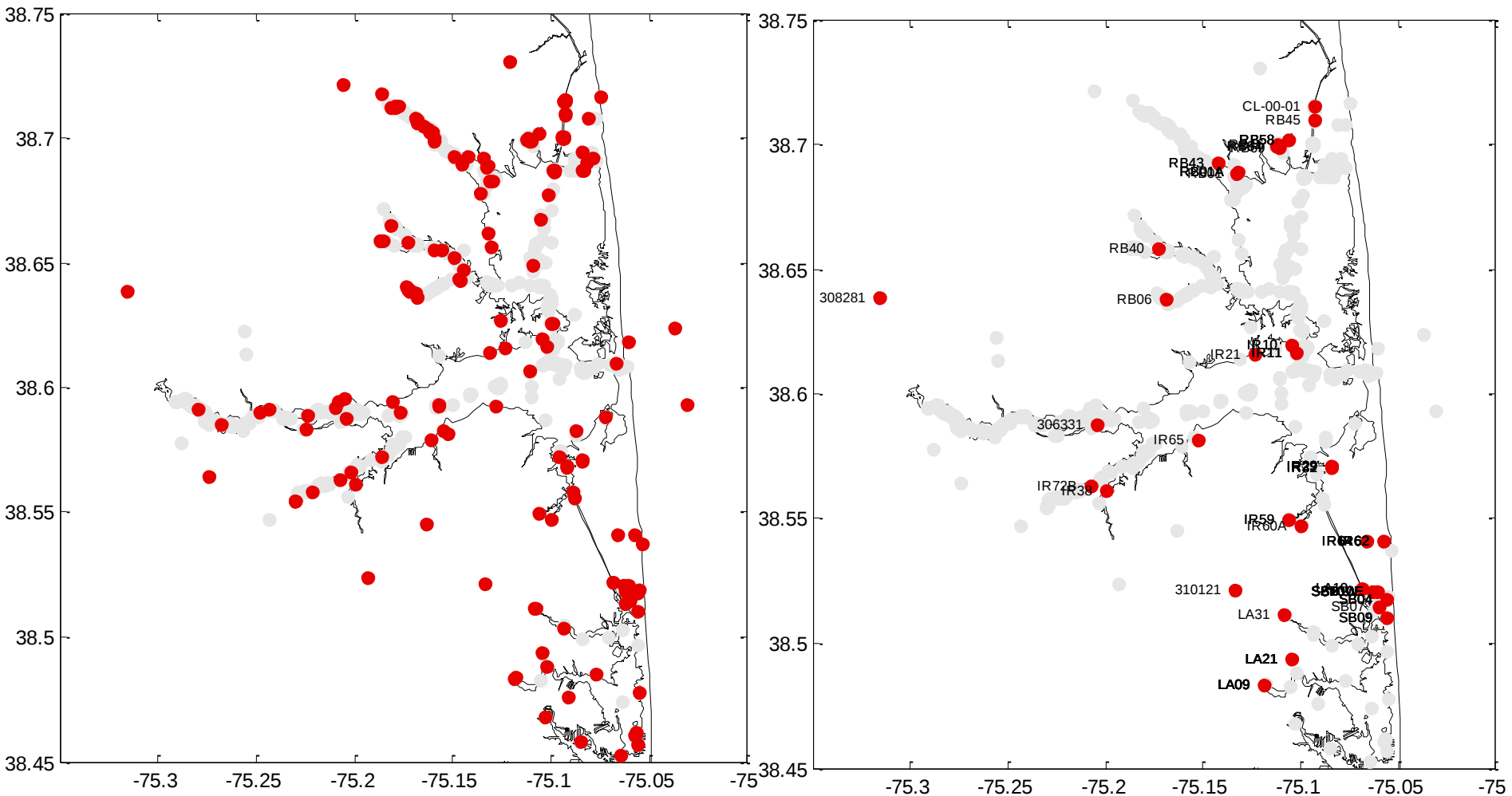


DO Station Means



DO < 5: 113 stations, 2009 observations

DO < 2: 31 stations, 86 observations

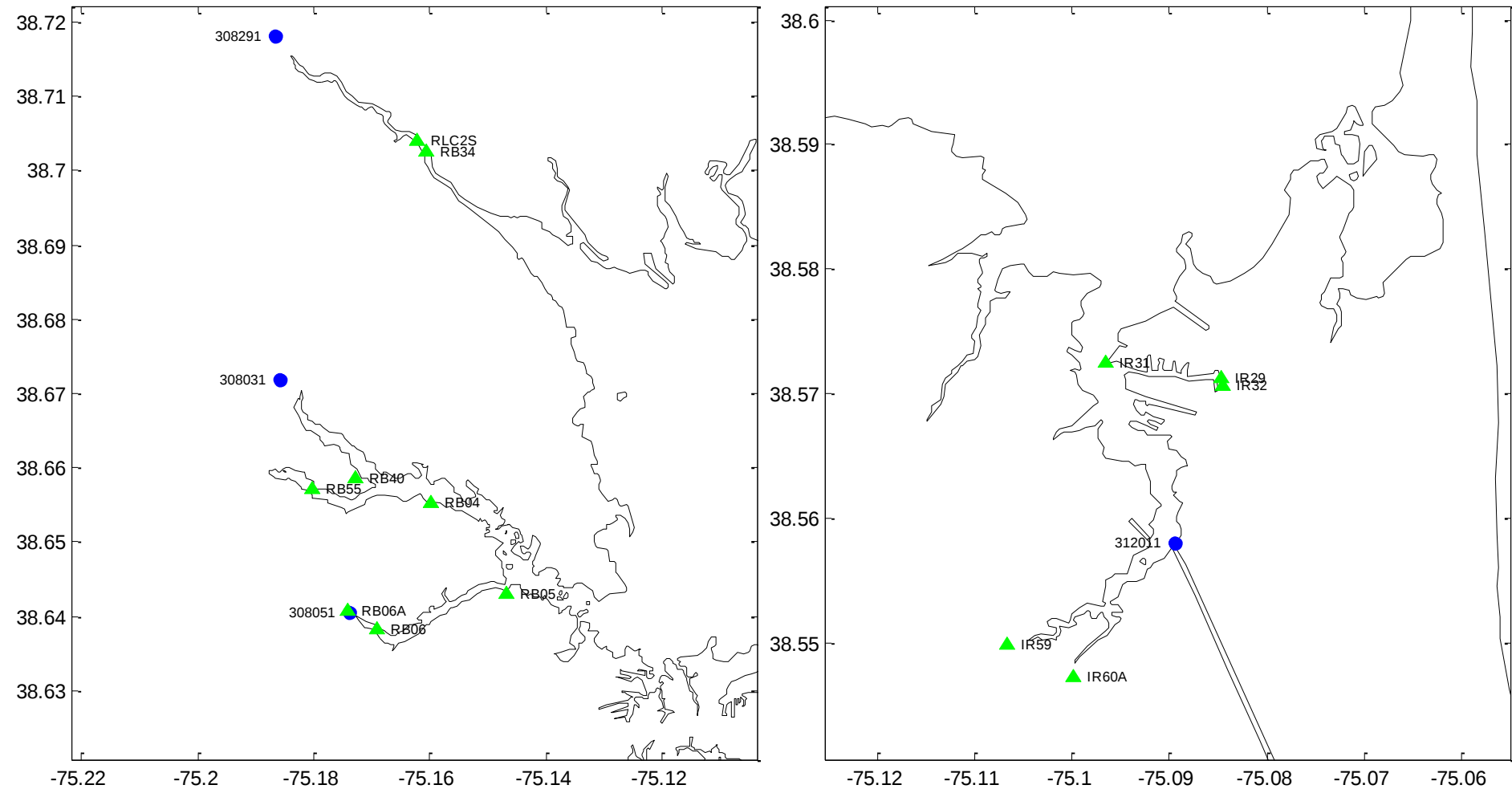


CMP+DNREC Metadata (n=11248)

Water Body	Stations	S Observations	DO Observations
Rehoboth Bay	40	3021	2522
Indian River Bay	50	2886	2445
Little Assawoman Bay	29	1613	1493
RB shore/open water	15	734	704
Love Creek	3	467	317
Herring	1	308	268
Guinea	4	877	685
IRB shore/open water	30	1979	1845
White	6	543	337
Pepper	4	52	64
LA shore/open water	9	469	408

*Including 931 CISNET “stations,” n = 12179

Love, Herring & White Creeks



Take-Home Messages

- Three data sets now in one, consistent, flat-file text format, usable in several packages (Excel, R, Matlab) ✓
- Parameters: salinity, DO, temperature, Chl a and Secchi depth ✓
- In this talk: metadata, sampling design, summary plots ✓
- Ahead this spring:
 - Salinity for segmented analysis (*in process*)
 - Geospatial: Mapping toolbox (Matlab) and Google Earth (KML), compatible w/ ESRI apps (ArcGIS)
 - Add nutrients parameters
 - Indicator development and performance, efficiency of sampling designs (*bootstrapping and resampling, training and validation sets, indicator performance and robustness ...*)