

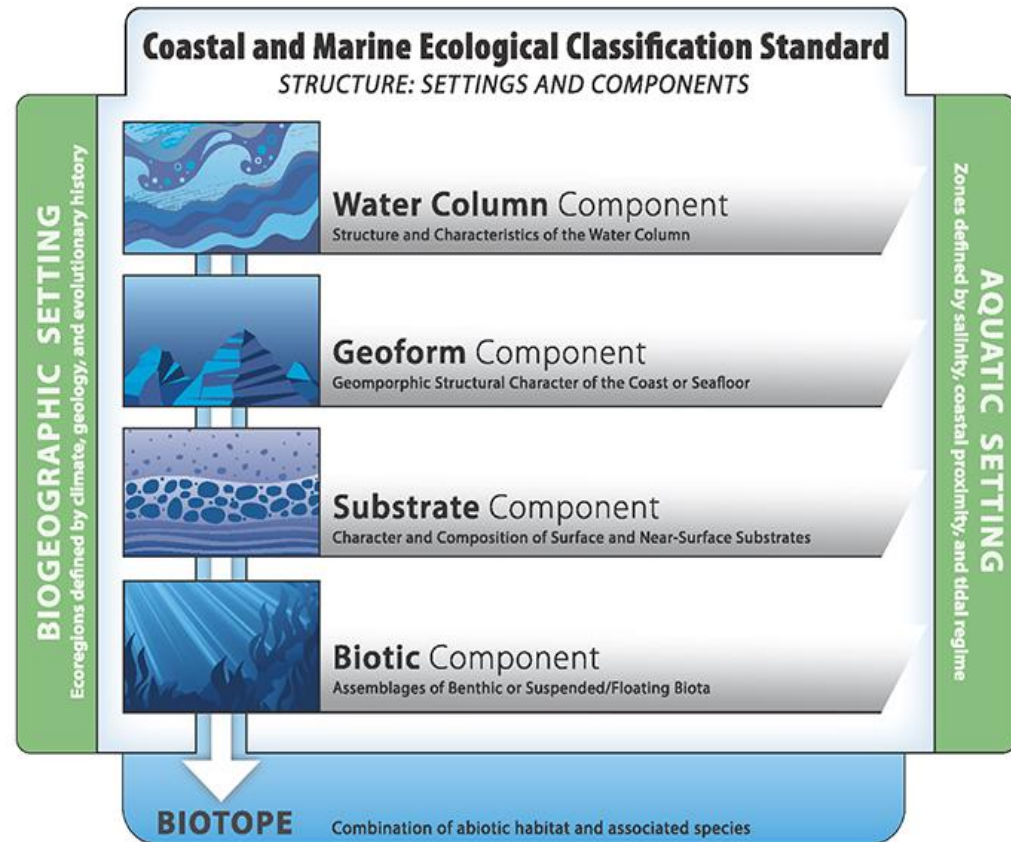
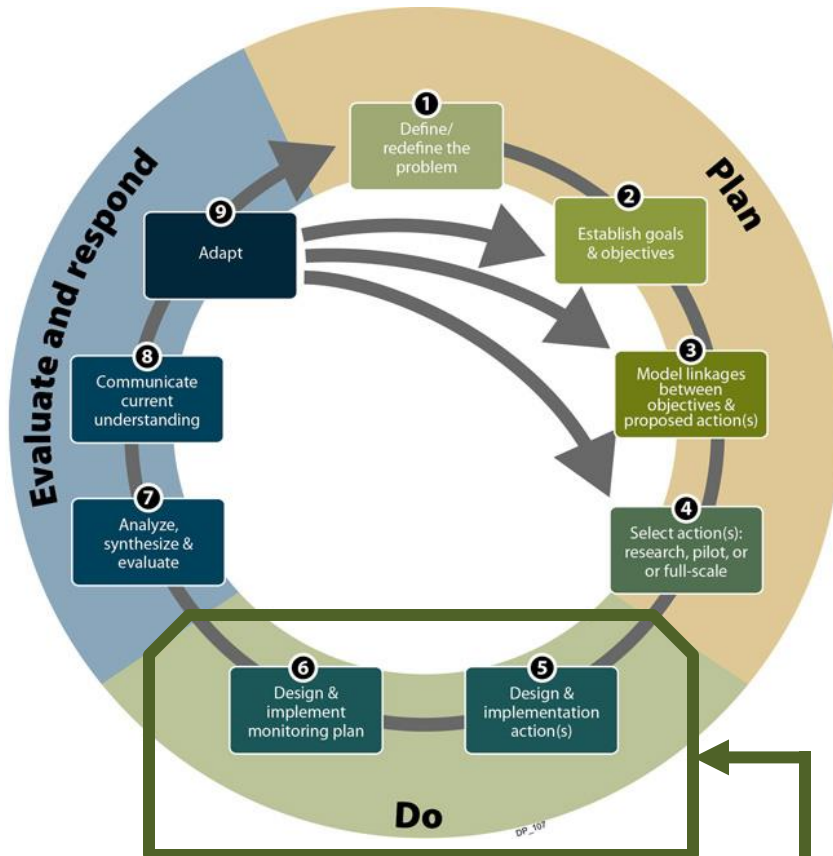
Effective and Simple Statistical Approaches for the Interpretation and Identification of Ecosystem Management Targets from Large Volume Data Sets

GYUNG SOO PARK

Department of Marine Biotechnology
Anyang University
Incheon, ROK



Backgrounds (adaptive management and ecological data)



California Natural Resources Agency

IOCM NOAA

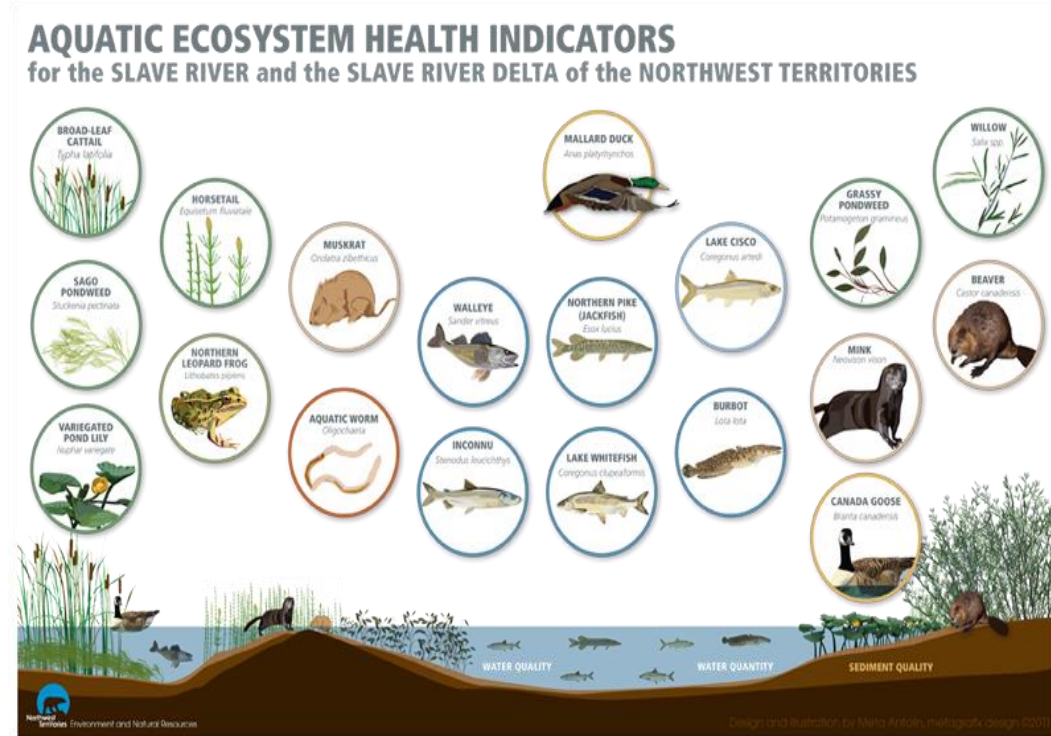
Identification of **MONITORING INDICATORS**

Human vs Ecosystem Health Check

Human Health Check



Ecosystem Health Check



Human vs Ecosystem Health Check

Men's Health Checklist



Have you had the following screenings?

- Blood Pressure
- Cholesterol
- Prostate
- Colon
- Diabetes
- Lung
- Skin Cancer



Have you had these immunizations?

- TDAP (Tetanus, Diphtheria, Pertussis)
- Hepatitis B
- Flu
- Pneumonia
- HPV



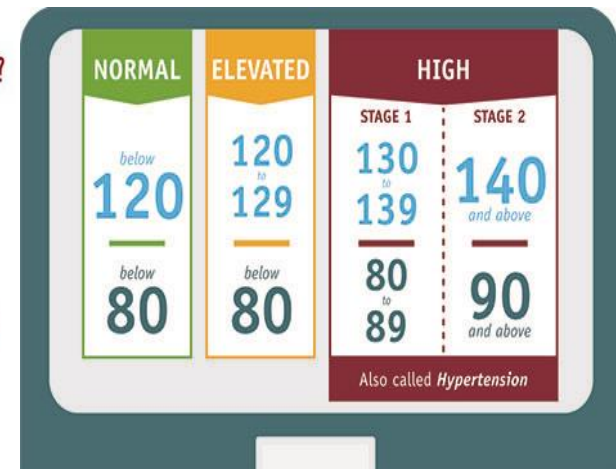
Have you been checked for...

- HIV
- TB
- STD's
- Hepatitis C

Things to think about...

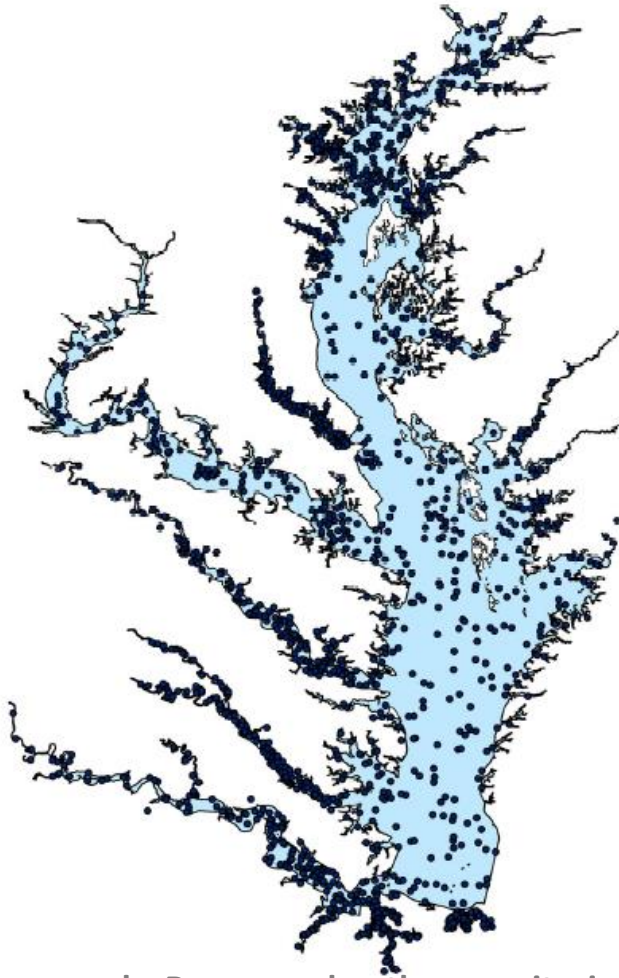
- Have you had a physical exam this year?
- Exercise 3x per week?
- Eat at least five servings of fruits and veggies daily?
- Do you know how to perform CPR?
- Consume any form of tobacco?

If you answered no to any of these questions, we can refer you to resources that will meet your needs. Please visit our website for more information... www.sullivanhealth.org



How to define ecological ranges of each parameters?

Marine ecosystem monitoring PG (CBP and HELCOM)

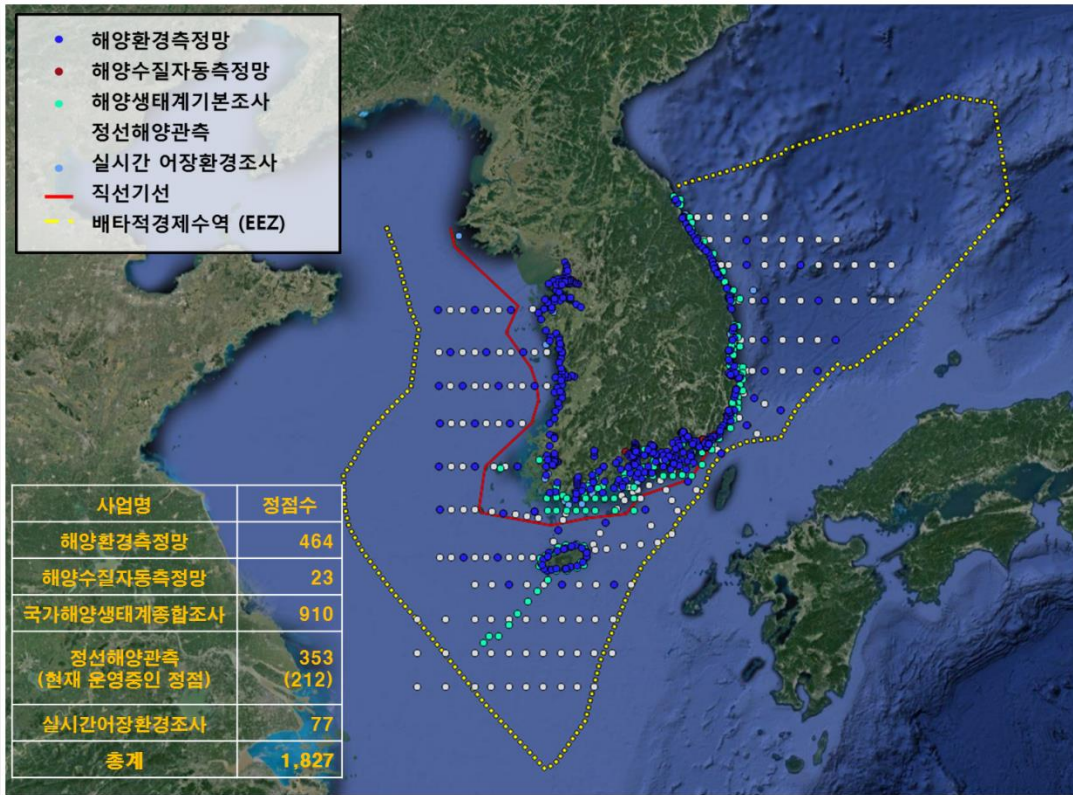


Chesapeake Bay macrobenthos monitoring stations



Baltic Sea environmental monitoring stations

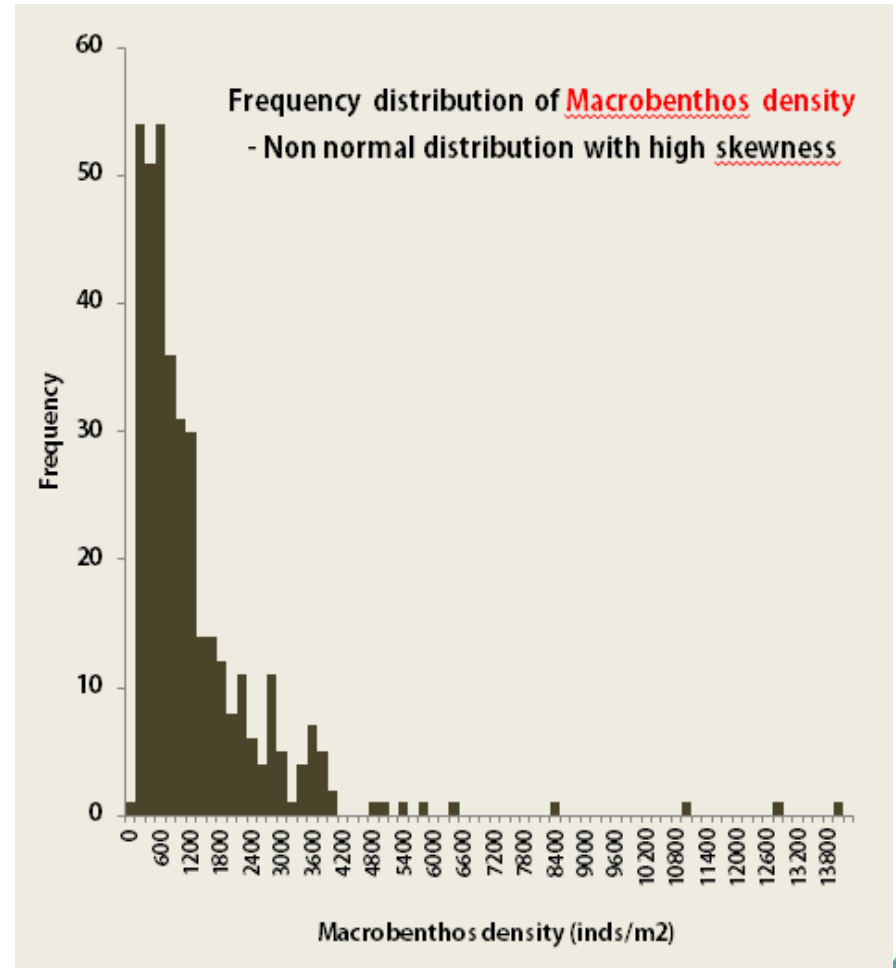
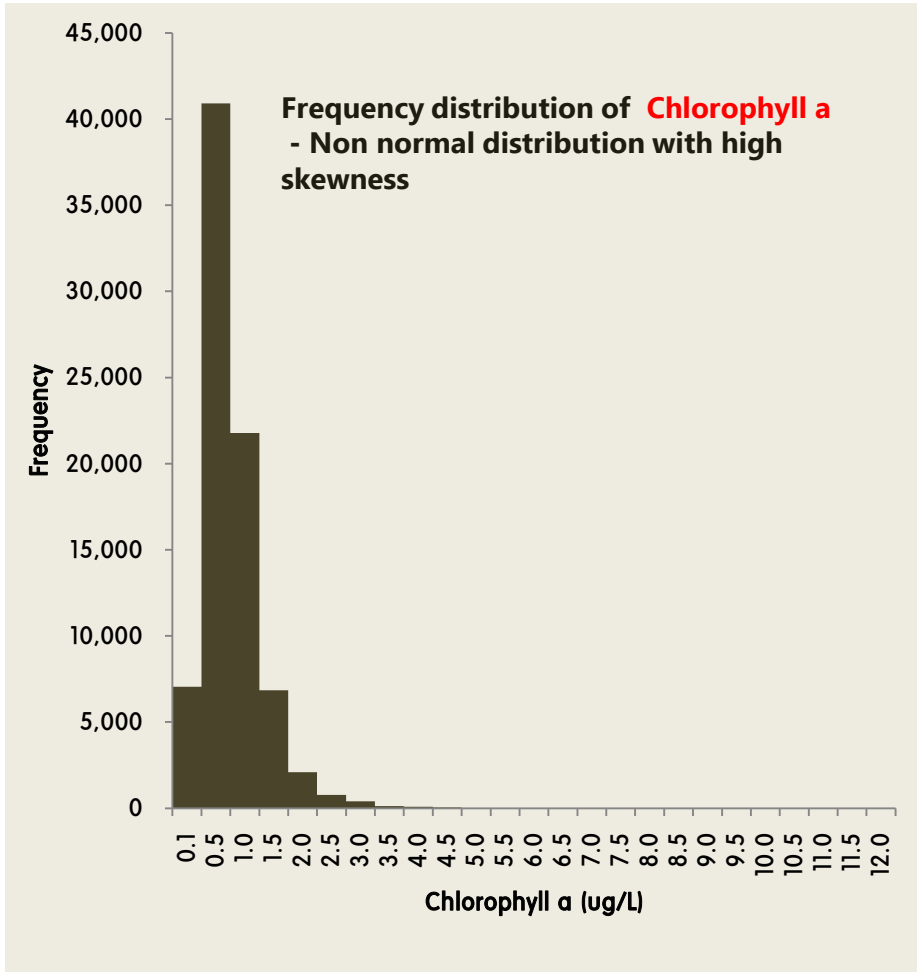
Marine ecosystem monitoring program (ROK)



Fields	Parameters
Water quality	WT, salinity, pH, transparency, nutrients, DO, SPM, POC/PON, metals
Sediment quality	PT, TOC, IL, TN, metals (Ni, Al, Cr, Fe, Cu, Zn, As, Cd, Hg, Pb)
Bacteria	number of cells
Phytoplankton	Chlorophylla (total/nano), species composition, density
Zooplankton	Chlorophylla (total/nano), species composition, density, biomass
Meiobenthos	species composition, density, biomass
Macrobenthos	species composition, density, biomass, health indices
Nekton	species composition, density, biomass, gut contents
Ichthyoplankton	species composition, density

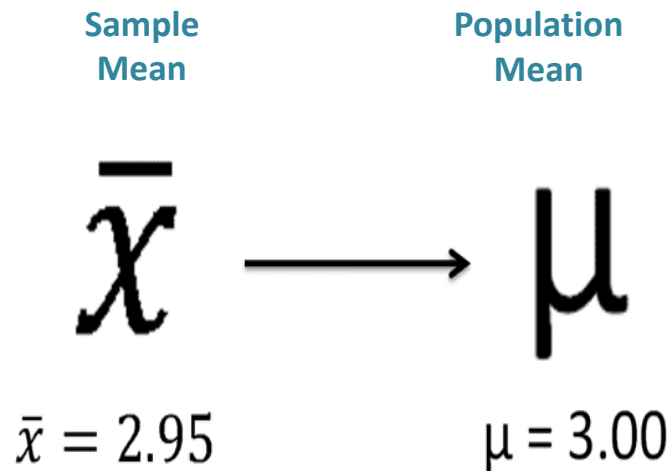
More than half million raw data per year in ROK marine ecosystem monitoring PGs

Characteristics of Coastal Ecosystem Data

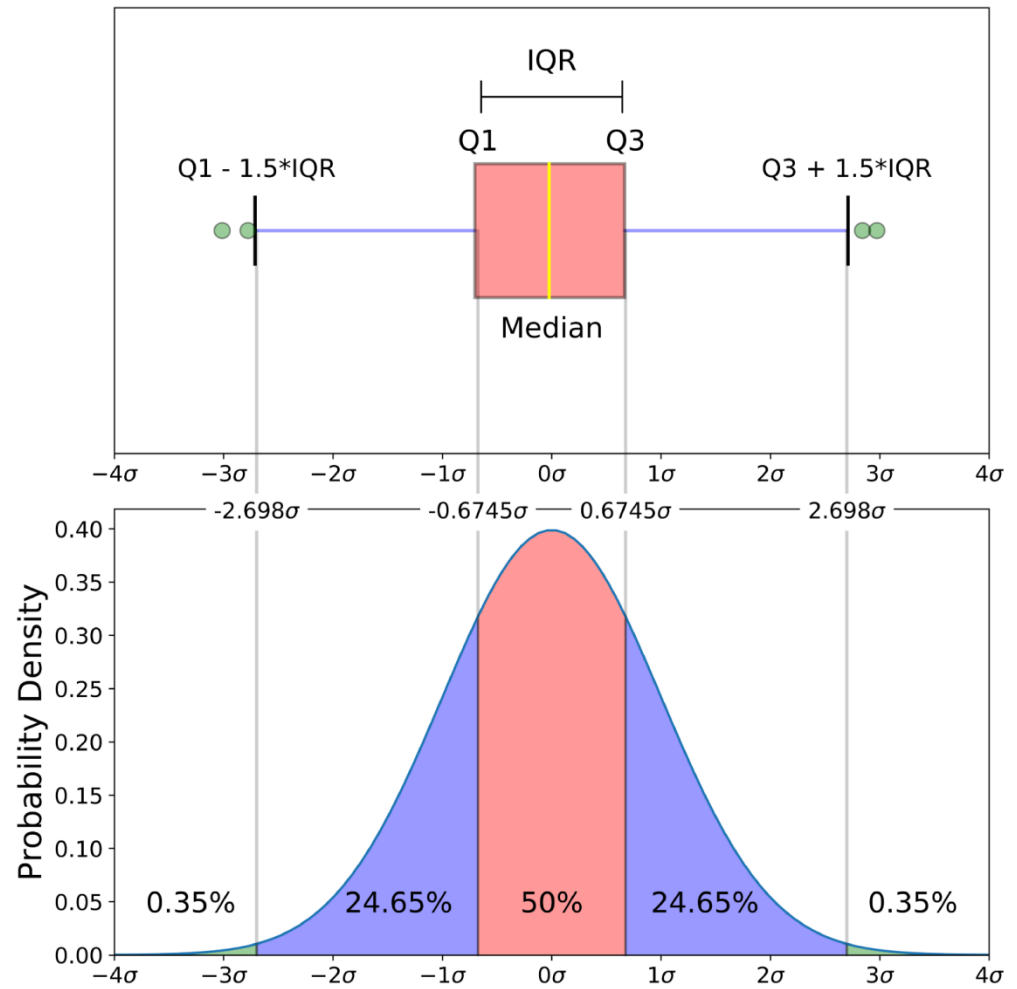


Coastal ecosystem data with high skewness and kurtosis (nonnormal distribution)

Estimation of ecological representations



An outlier is a value that is very different from the other data and can skew your results



Effect of outliers on descriptive statistics (zooplankton density)

* Outliers are only 12.5% of total number of data

Parameters	East Sea		Jeju Sea		South Sea		Yellow Sea		Mean
	Raw	Trimmed	Raw	Trimmed	Raw	Trimmed	Raw	Trimmed	
N	23	19	12	10	48	42	63	58	-
Mean	1,008	737	3,420	1,959	14,288	6,101	59,546	37,255	-
SD	703	382	3,598	1,246	24,199	4,484	103,766	51,729	-
Median	903	722	2,026	1,509	6,243	5,568	17,138	15,802	-
Max	2,541	1,654	11,143	4,296	101,196	19,776	665,895	229,439	-
Min	85	85	529	529	519	519	209	209	-
N (reduced %)	17.4		16.7		12.5*		7.9		13.6
Mean (reduced %)	26.9		42.7		57.3*		37.4		41.1
SD (reduced %)	45.7		65.4		81.5*		50.1		60.7

12.5% outliers overestimated samples mean of 57.3% in zooplankton density(27-57%) , 0-67% for phytoplankton and 0-40% for macrobenthos density

Estimation of background levels of heterotrophic bacterial density

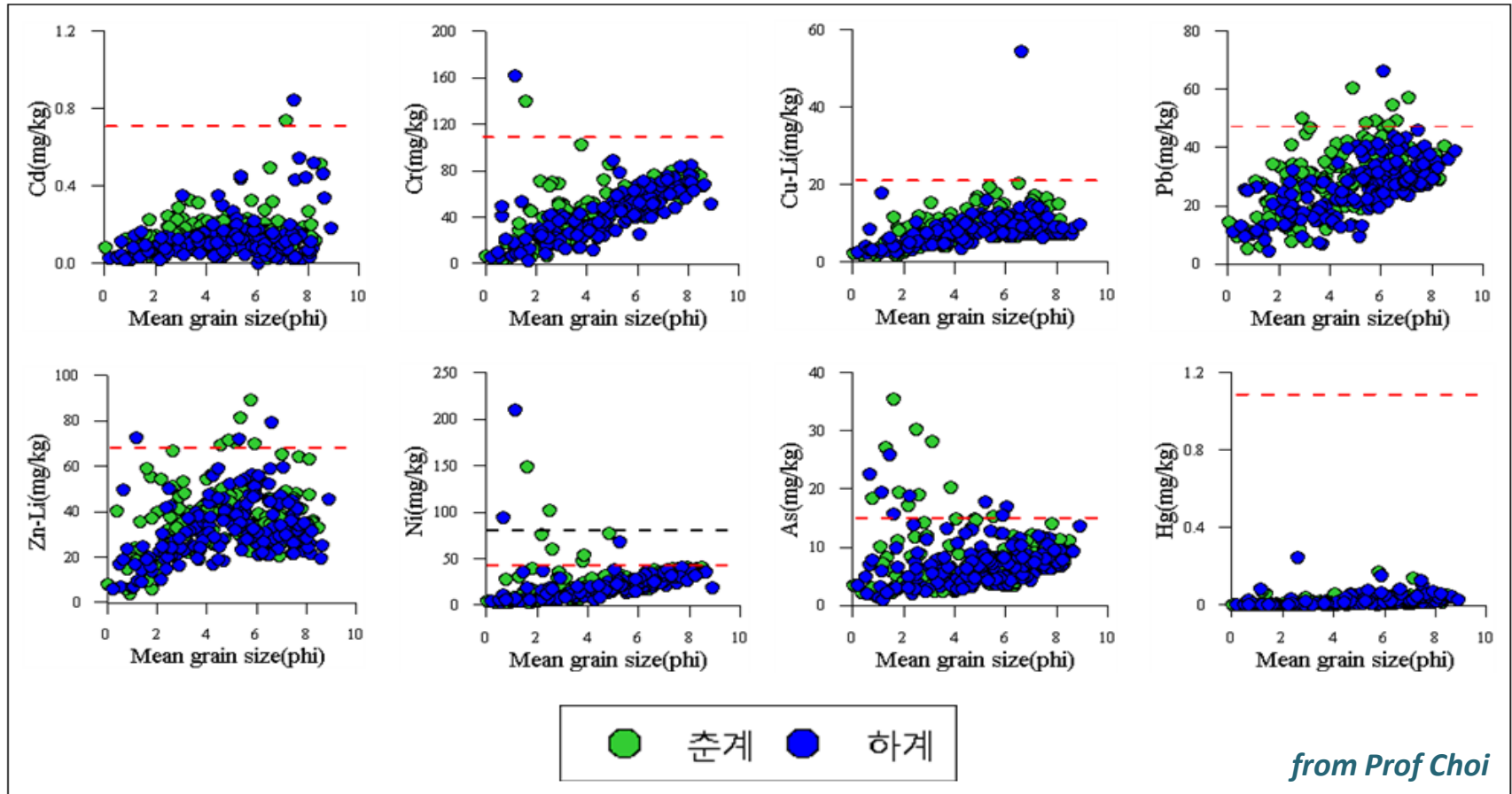
Months	Seas	Mean	Median	SD	CV(%)	Min	Max	Q1	Q3	Outlier	N
May	YS	680,275	611,000	346,969	51.0	238,470	1,678,459	401,870	869,568	7	112
	SS	637,646	576,519	321,490	50.4	206,956	1,470,192	388,725	817,168	9	91
	ES	792,341	757,942	297,571	37.6	346,718	1,485,781	555,696	1,015,264	0	47
	Jeju	516,051	487,124	240,162	46.5	158,543	1,277,516	354,109	688,396	0	28
Aug	YS	510,260	473,008	227,526	44.6	123,823	1,139,937	337,383	581,106	8	111
	SS	947,219	841,597	488,294	51.6	201,734	2,429,759	616,078	1,168,410	9	91
	ES	780,660	681,498	311,527	39.9	305,781	1,578,061	522,254	981,031	0	47
	Jeju	1,363,188	1,224,286	611,904	44.9	339,999	2,580,713	890,581	1,698,561	0	28



Use of outliers for the ID of management targets

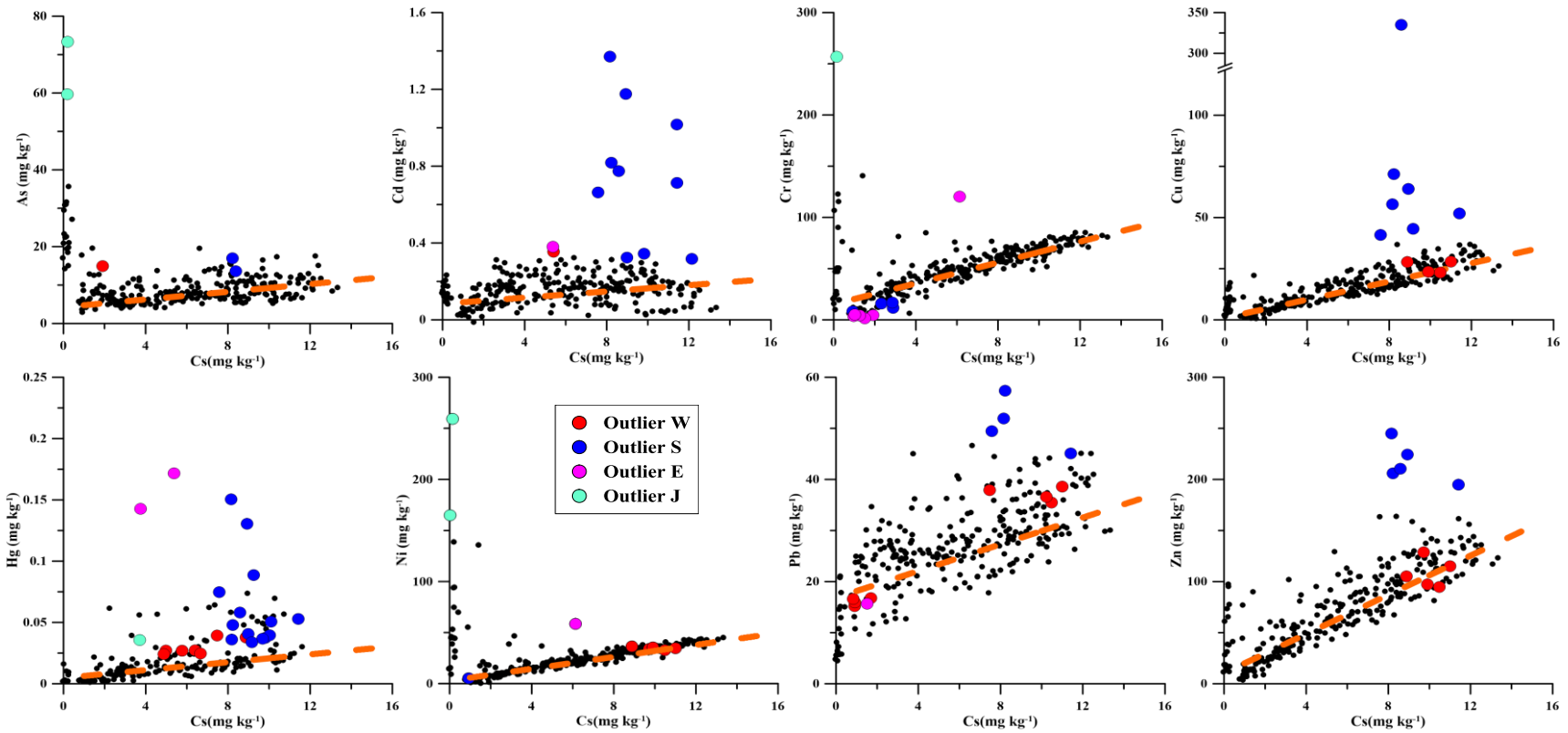
Indicators	2014	2015	2016	2017	2018	short/long
1. High species richness in Gangjin Bay	-	0	0	0	-	Short-term indicator
2. High chlorophyll in Han river estuary	0	0	0	0	0	Long-term indicators
3. High Lead conc in the mid area of East sea	-	-	0	-	0	Short-term indicator
4. Low oxygen in Masan bay	0	0	0	0	0	Long-term indicators
5. High zooplankton density in Kum river estuary	-	0	0	0	0	Long-term indicators
6. High Noctiluca density in Kum river estuary	-	0	-	X	-	Short-term indicator
7. High bacterial density in Youngil Bay	-	-	0	X	X	Terminated
8. High bacterial density in Jeju during summer	-	-	X	X	0	Short-term indicator

Feasibility of outliers for ecological monitoring indicators and management targets (regulatory levels)



Outliers can be “hidden” or not always be “odd” if they are under regulatory level

Feasibility of outliers for ecological monitoring indicators and management targets (cesium vs metals)



statistical outliers are not always “ecological” outliers

from Prof Choi

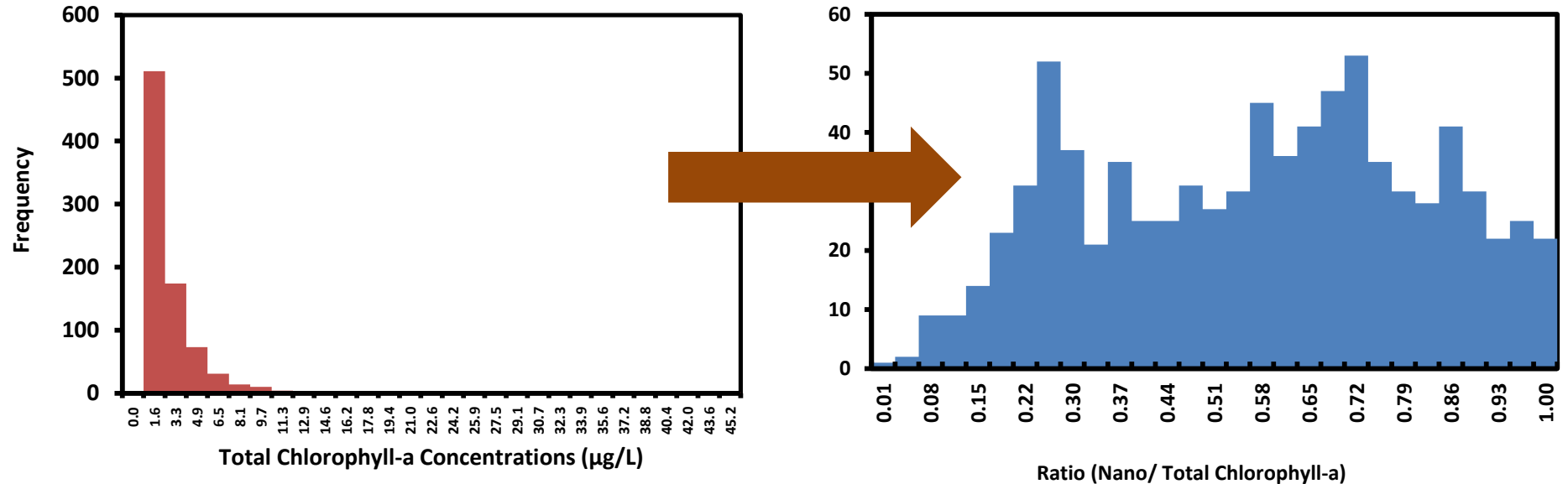
Feasibility of outliers for ecological monitoring indicators and management targets



Ecoregions divided into small relatively homogeneous area. In case of ROK based on current, tide, turbidity, and water depth

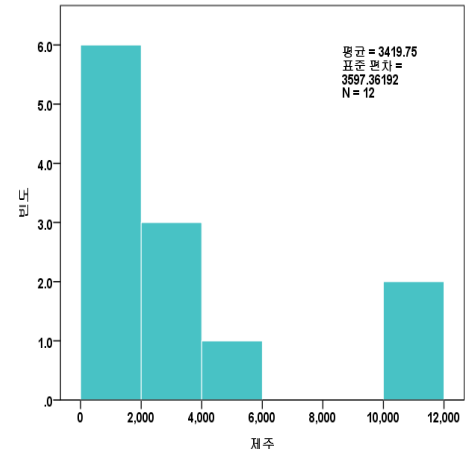
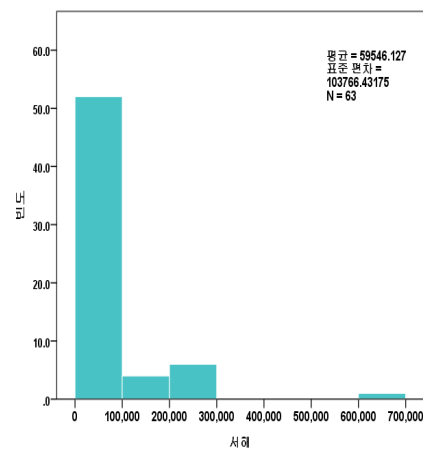
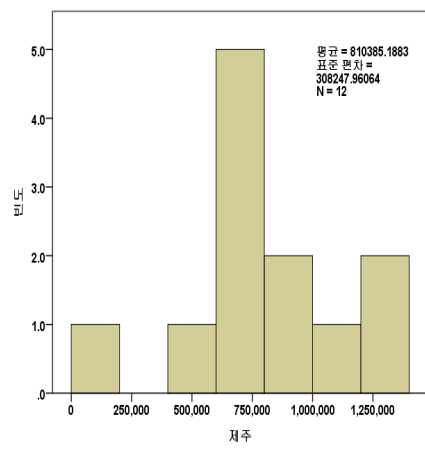
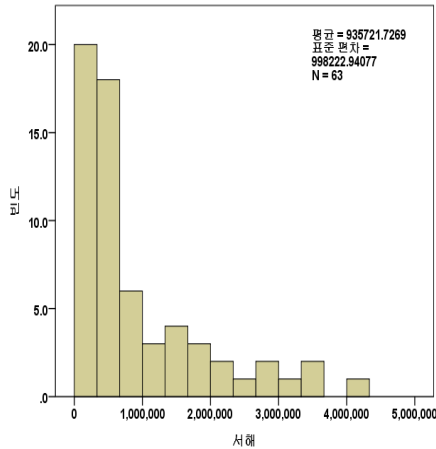
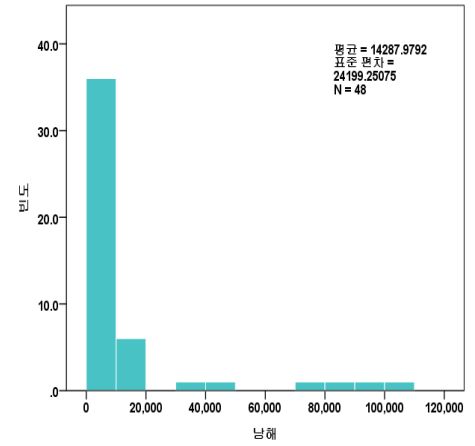
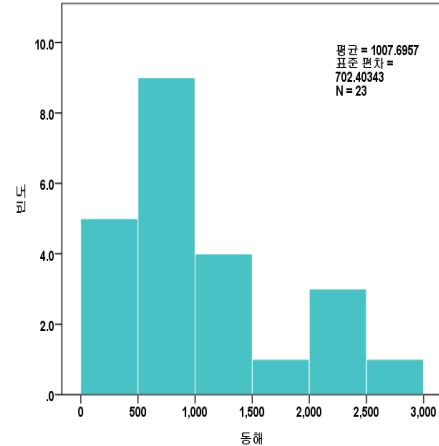
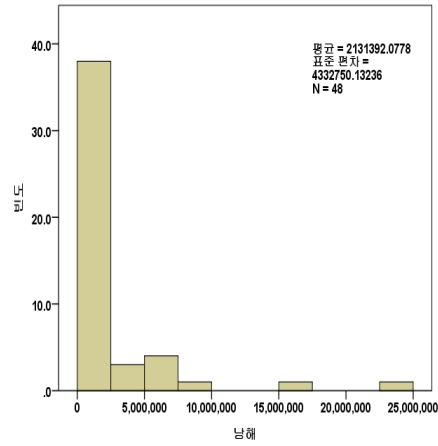
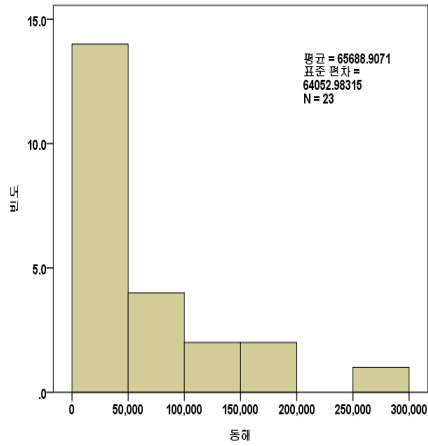
- Mid-west ecoregion : strong tide/high turbidity
- South-west ecoregion : strong tide/low turbidity
- Jeju ecoregion : subtropical climate and current
- Korea strait ecoregion ; Kuroshio current
- East Sea ecoregion : deep sea
- Five river estuaries and coastal upwelling area as hot spots

Use of derived data for normality (nano/total chl-a)



Chl-a	Mean	Median	Mode	SD	CV(%)	Kurt	Skew	Min	Max	N
Total	2.03	1.24	0.23	2.99	147.27	106.06	8.16	0.03	45.24	827
Nano	0.93	0.60	0.52	1.11	119.55	23.69	3.92	0.01	11.54	827
Ratio(%)	0.55	0.58	1.00	0.25	44.59	-1.05	-0.11	0.01	1.00	827

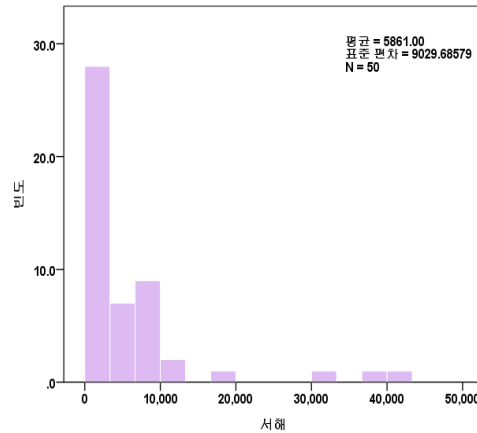
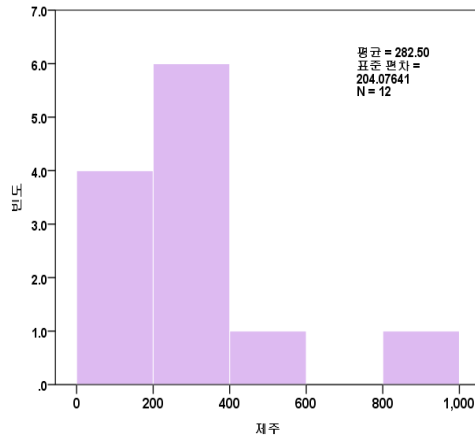
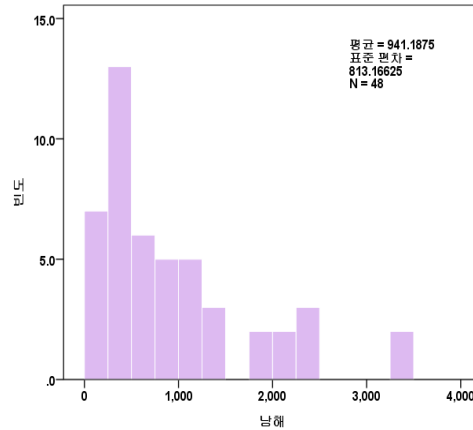
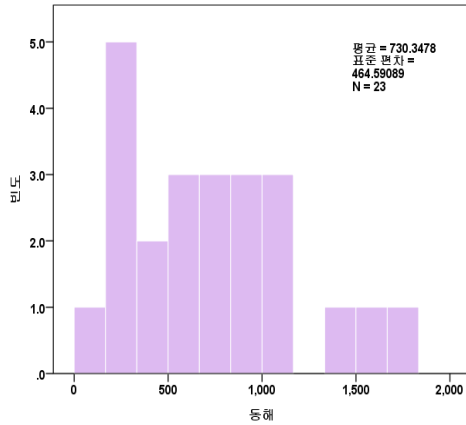
Frequency distributions of primitive data (plankton density)



Phytoplankton cell density

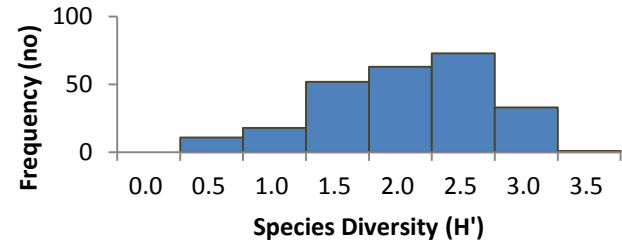
Zooplankton density

Use of derived data for normality (diversity indices)

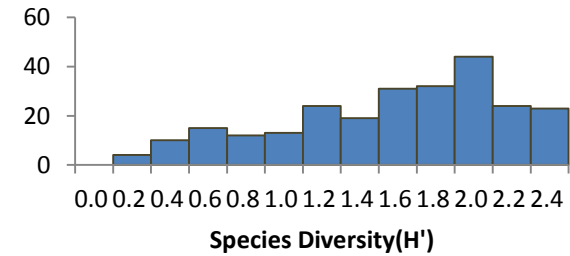


Macrobenthos density

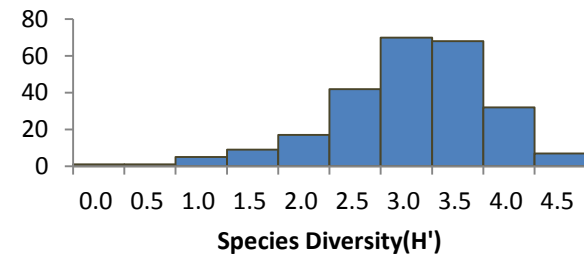
Phytoplankton Diversity



Zooplankton Diversity

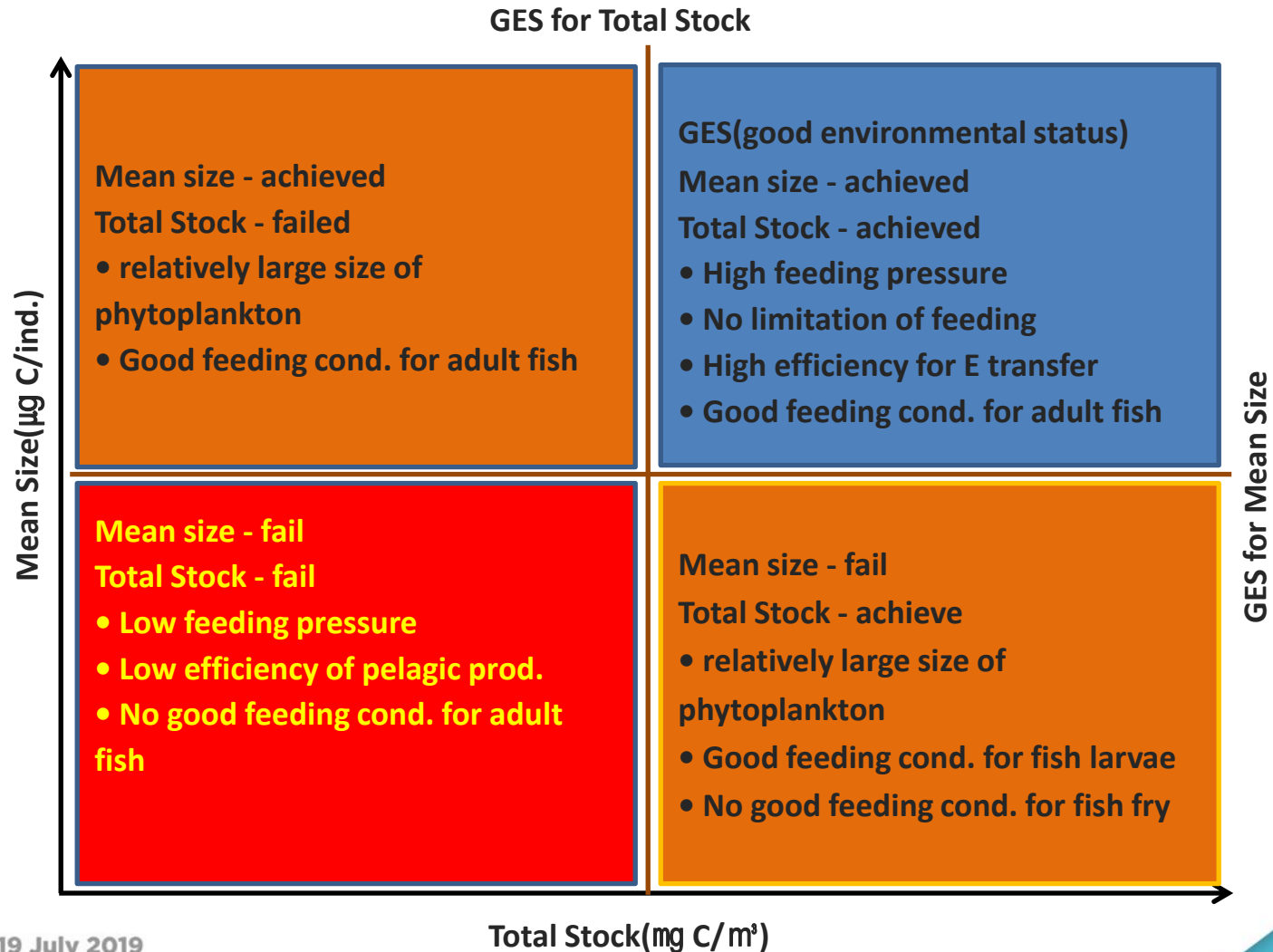


Macrozoobenthos Diversity



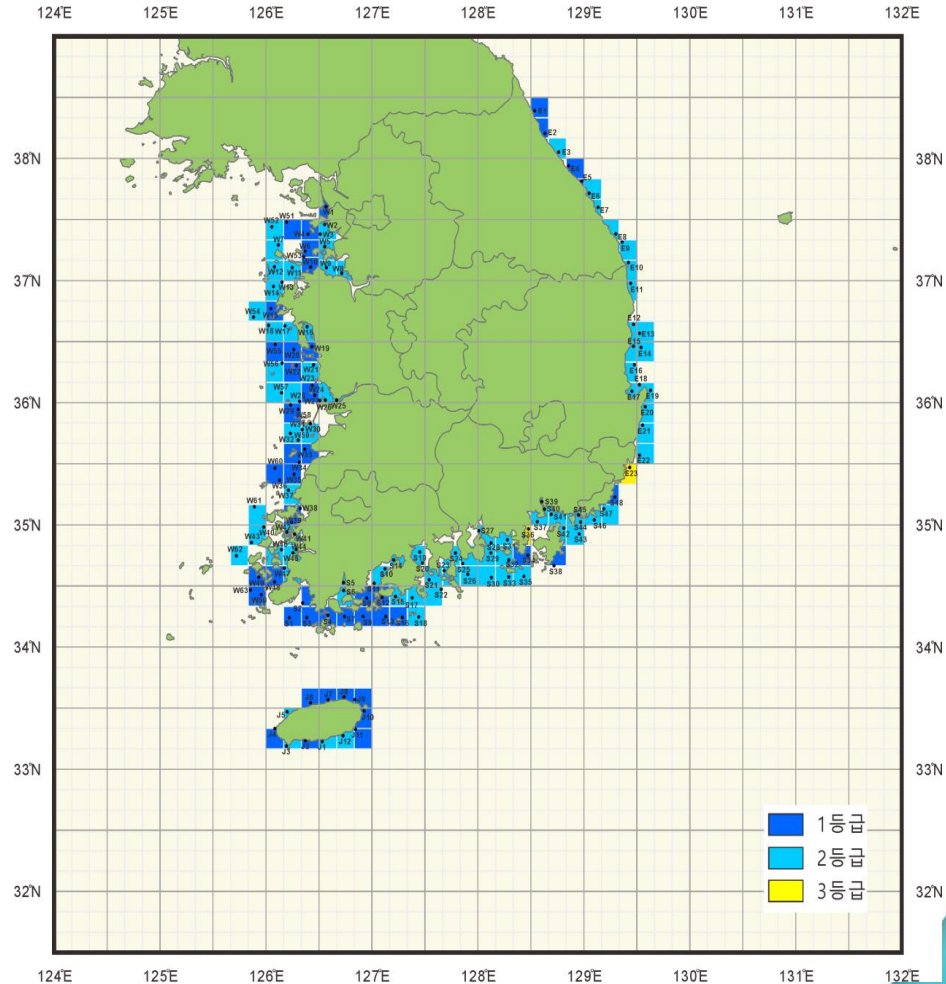
Use of derived data for normality (MSTS)

❖ Zooplankton MSTS (Mean Size and Total Stock, HELCOM)



Summary of data analysis (ecological index)

1. Protected species	2. Ecosystem health
<p>24 invertebrate species 7 sea grass and algae 2 fish species</p>	<p>Bottom DO% Metals in sediment Biomass of macrobenthos ISEP (macrobenthos) No of <i>Noctiluca</i> Chlorophyll a</p>
3. Biodiversity	4. Protected area
<p>Diversity index of phyto/zooplankton and macrobenthos No of phyto/zooplankton and macrobenthos species</p>	<p>Coastal wetlands Marine protected area (MPA) Marine ecosystem protected area</p>



Summary of presentation

1. Ecosystem data are highly skewed with extremely high deviations
 - highly biased mean values and no reflection of ecosystem status
2. Screening process such as outlier test can minimize the deviation, provide the easy way to downsize the large volume of data, and identify the ecological interpretation and management targets
 - statistical outliers, not always ecological outliers
 - first consideration of regulatory level if established
 - outlier testing within the relatively homogeneous ecoregions divided from other systems
 - outliers through in-depth verification selected as major ecosystem monitoring indicators in terms of space and time
3. Derived data from primitive values are more suitable to be analyzed for spatial interpretation and time series analysis because of the small deviation and relatively normal distribution – ex ratio and indices
4. Trimmed data like outlier testing can be applied to establish background levels of ecological parameters and further analysis like an ecological index estimation, cluster analysis and multivariate analysis etc



Thank you for your listening