

# Data Processing

## Making the numbers play nice with VB

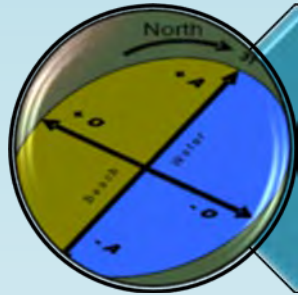
### Module 1



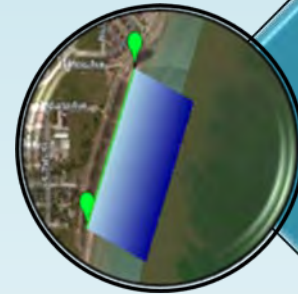
great lakes  
observing system



## Variable Types



## Variable Combinations



## Variable Transformations

# Independent Variables

- MUST be quantitative for Virtual Beach
- Should influence or explain bacteria somehow
- Can be categorical numbers (ranked 1-3)

# Variable Types

## Quantitative

- Ratio
- Interval

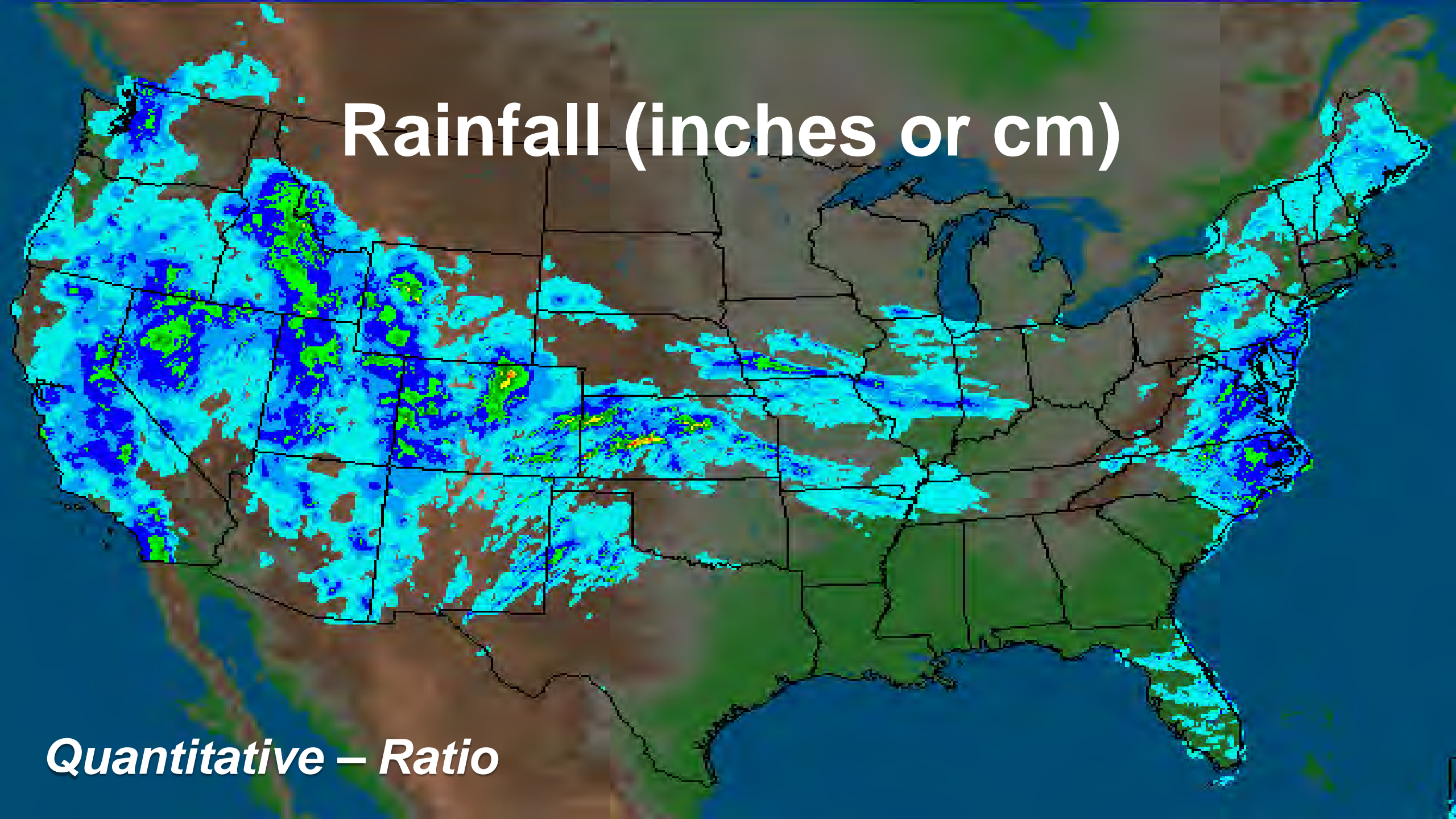
75 °F  
75 NTU

## Qualitative

- Ordinal
- Nominal

Low  
Cloudy

# Rainfall (inches or cm)



*Quantitative – Ratio*

# Turbidity



**<10, 200, 1500 NTU**

***Quantitative – Ratio***

# Temperature ( $^{\circ}\text{C}$ , $^{\circ}\text{F}$ )



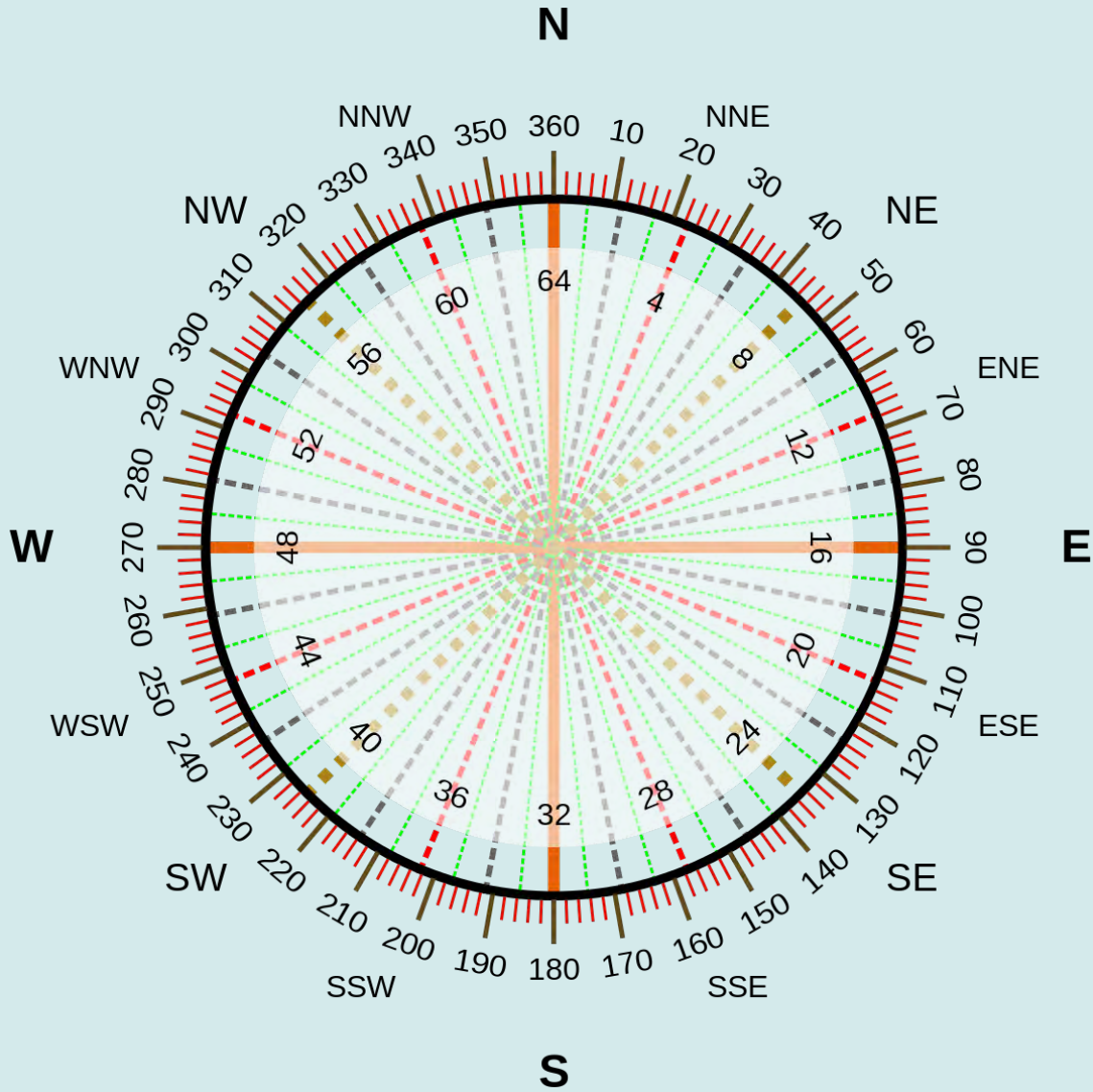
***Quantitative – Interval***

Variable Types

Variable Combinations

Variable Transformations

# Wind Direction (degrees)



**Quantitative – Interval**



Source: NOAA



# Sky Conditions - sunny, partly cloudy, cloudy

Ranked Categories	sunny	mostly sunny	partly sunny	mostly cloudy	cloudy
Numerical	0	1	2	3	4

***Qualitative – Ordinal***

# Variable Combinations

## Interaction Terms

Multiply terms that interact with each other:

Gulls & Wave Height

Tributary Input & Alongshore current

## Combine Categorical Value

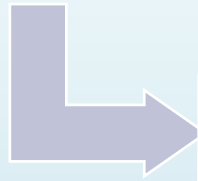
Sum categories converted to numbers

## Change-in-flow Variables

Subtract variables that occur over time

# Interaction Terms

Explanatory variables are assumed to be **independent** from each other



**Collinear** (non-independent) variables may over influence model



Variables may interact: e.g. river discharge & current



Solution: Combine collinear variables into **“Interaction term”**

# Combined Categories

Turbidity	Clear	Slightly Turbid	Turbid	Opaque
Day 1	0	0	1	0
Day 2	0	1	0	0
Day 3	0	0	0	1

**Turbid + Opaque = ...**

Turbidity	Clear	Slightly Turbid	Turbid
Day 1	0	0	<b>1</b>
Day 2	0	1	0
Day 3	0	0	<b>1</b>

# Change flow variables

Subtract one continuous data point from another to create proxy variables for changes in flow

EX.: Minimum and maximum stream flow over a day

**(Trib24 hr max) – (Trib24 hr min) = stream “flashiness”**

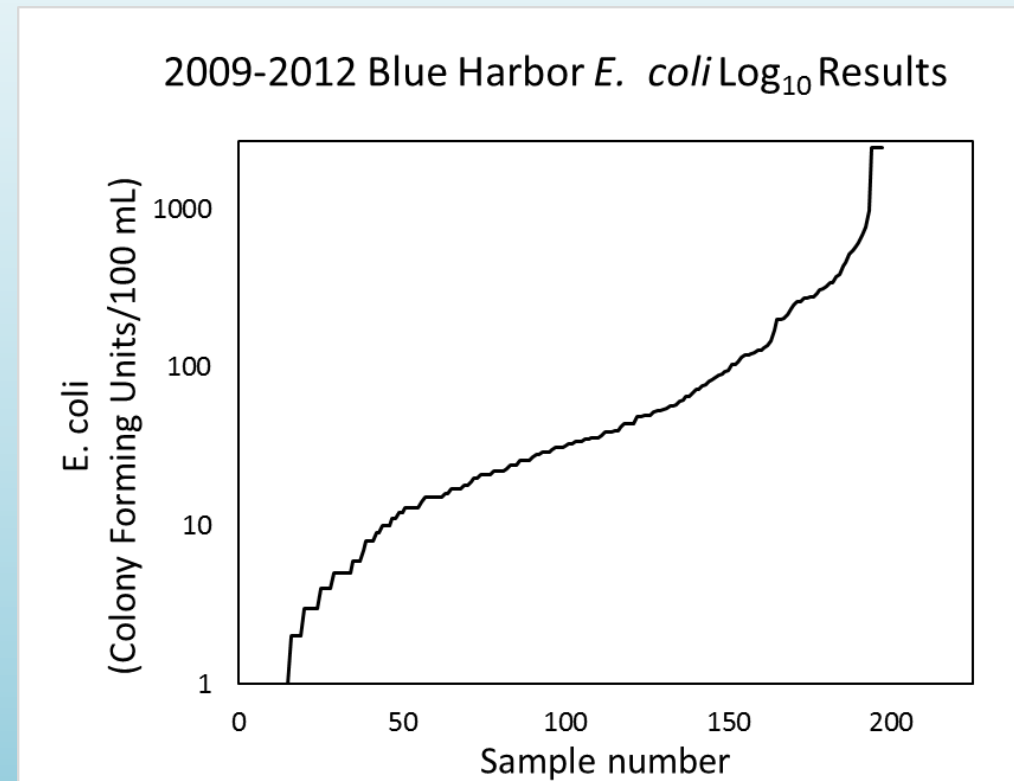
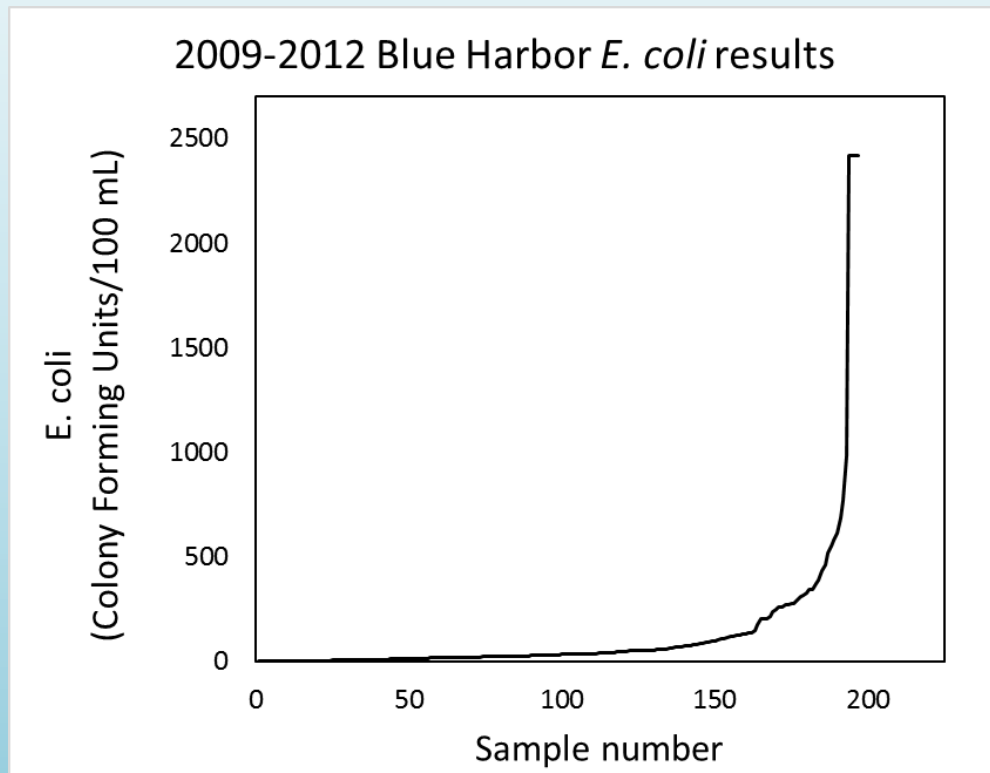
# Variable Transformations

Including directional data ~

Making non-linear variables “more linear”

# Nonlinearity

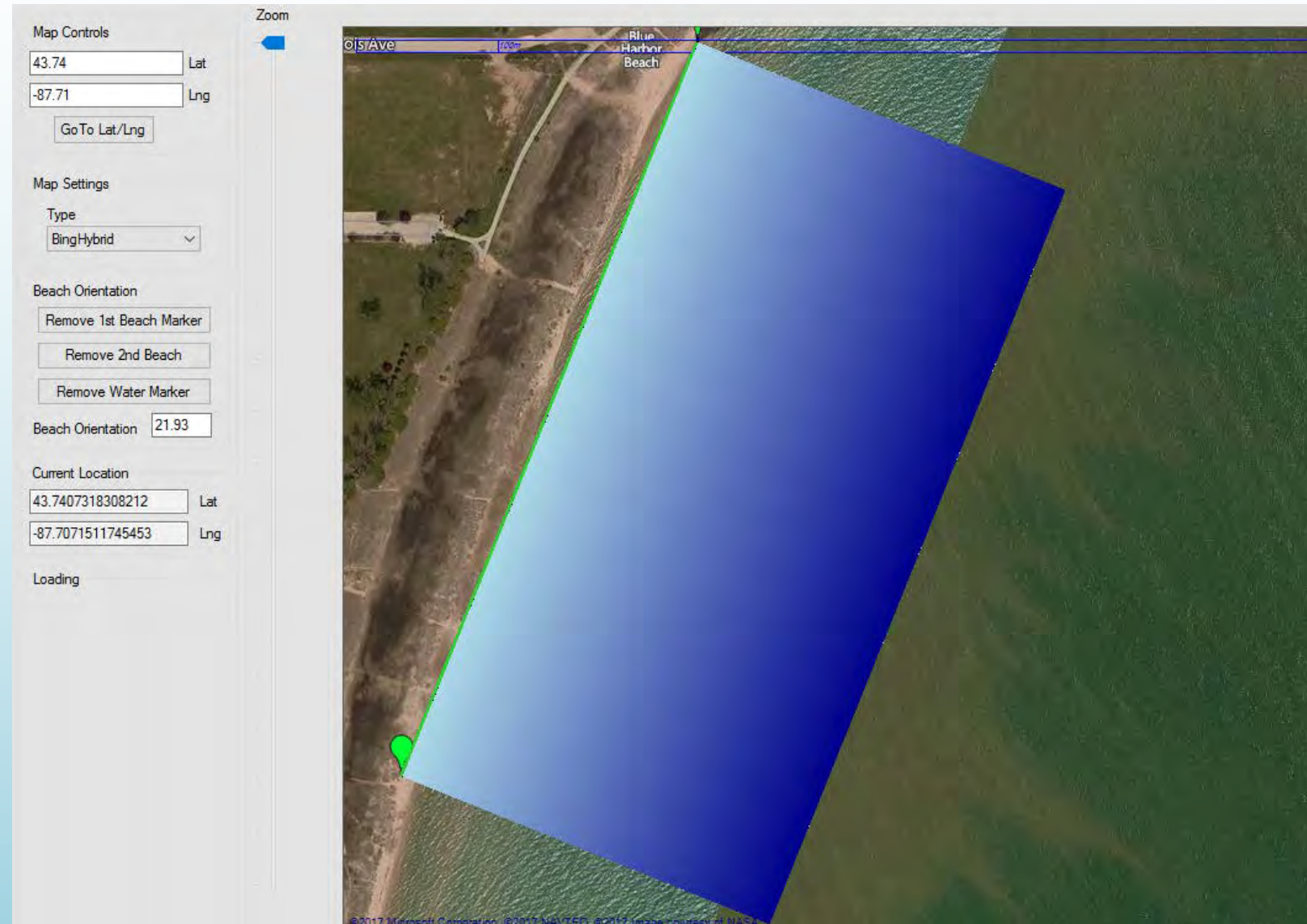
Since changes in *E. coli* concentrations can occur over orders of magnitude, it is a useful modeling practice to transform this variable



# Directional Data

Beach Orientation

Optional, but useful!

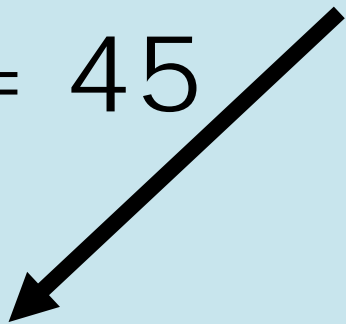




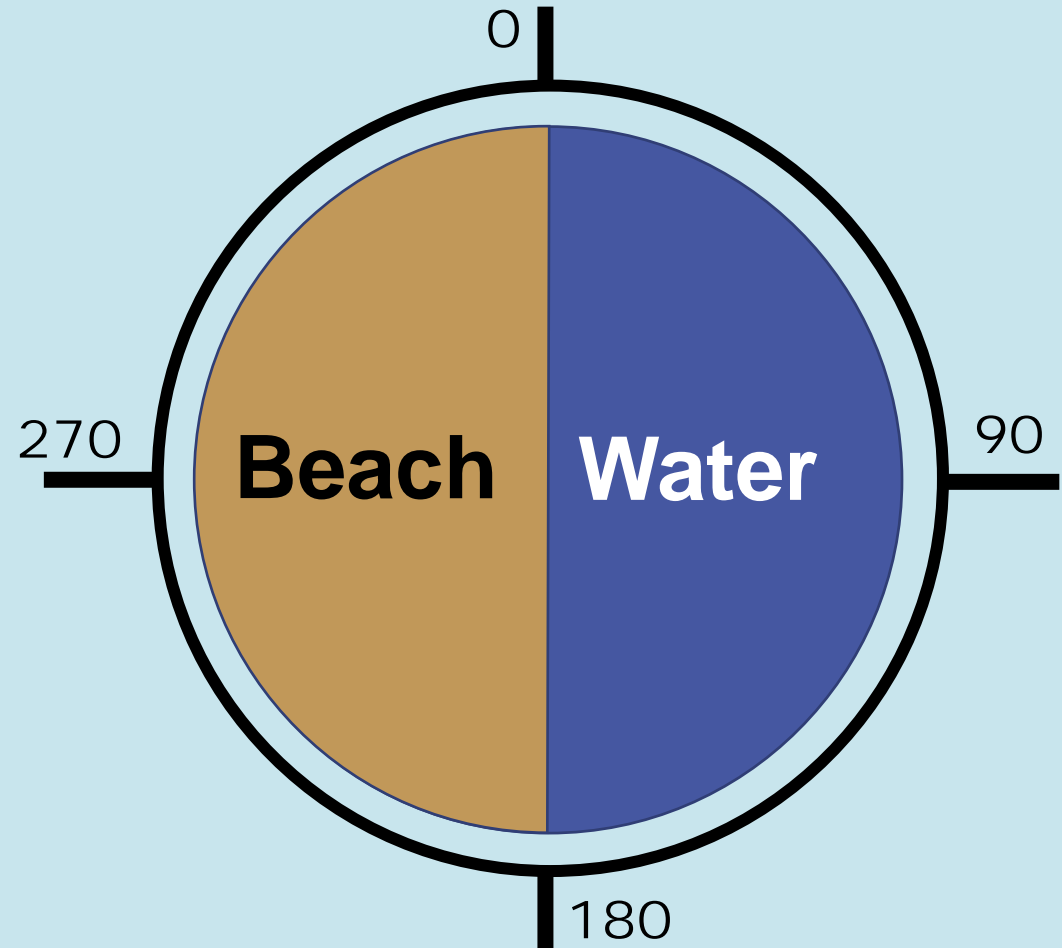
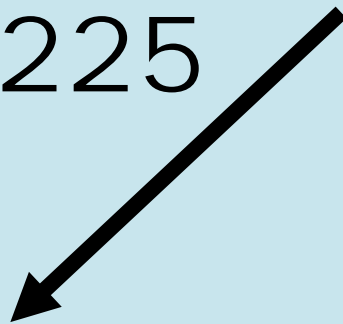
# Beach Orientation in Compass Degrees

Wind & Current  
Directions given in  
numbers

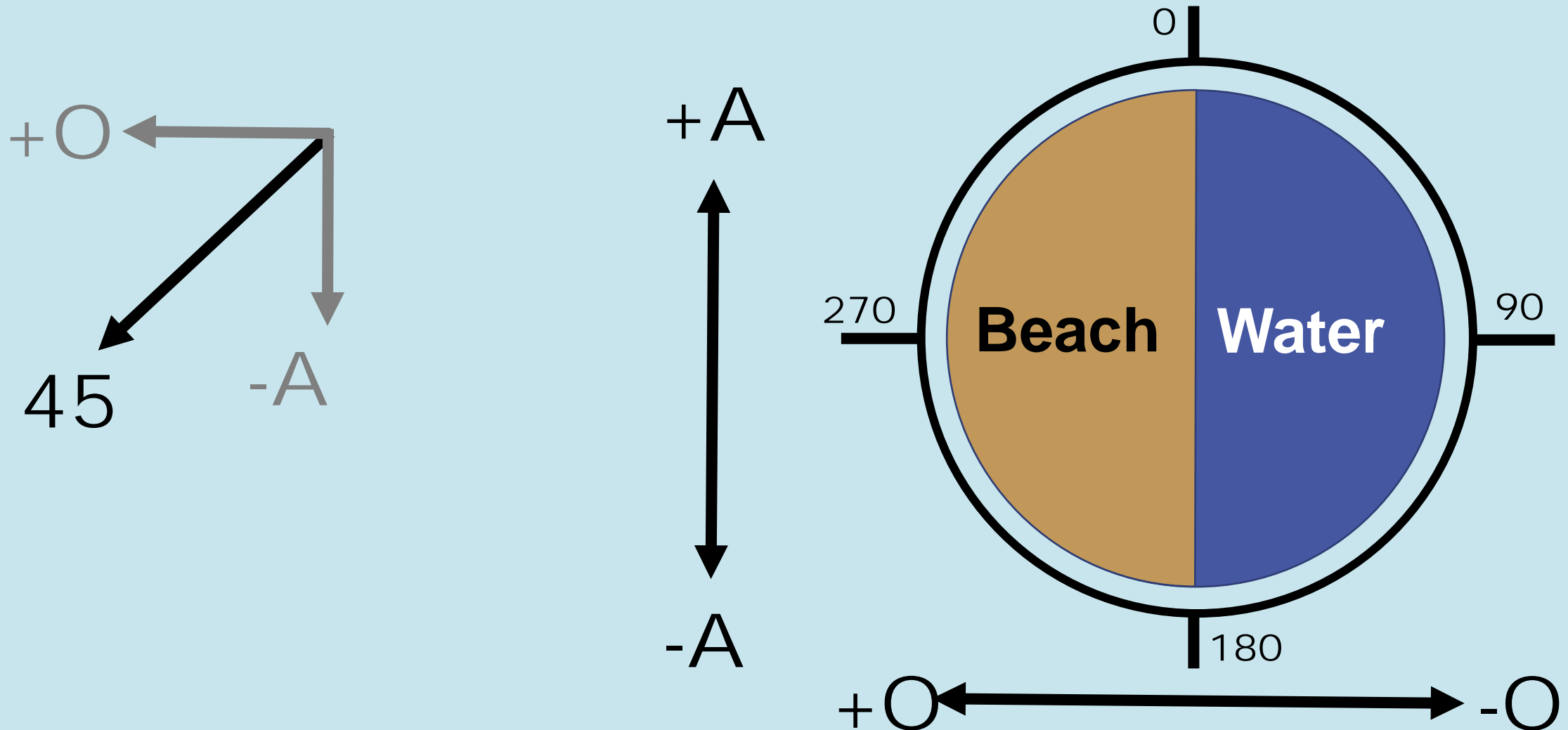
Wind  
= 45



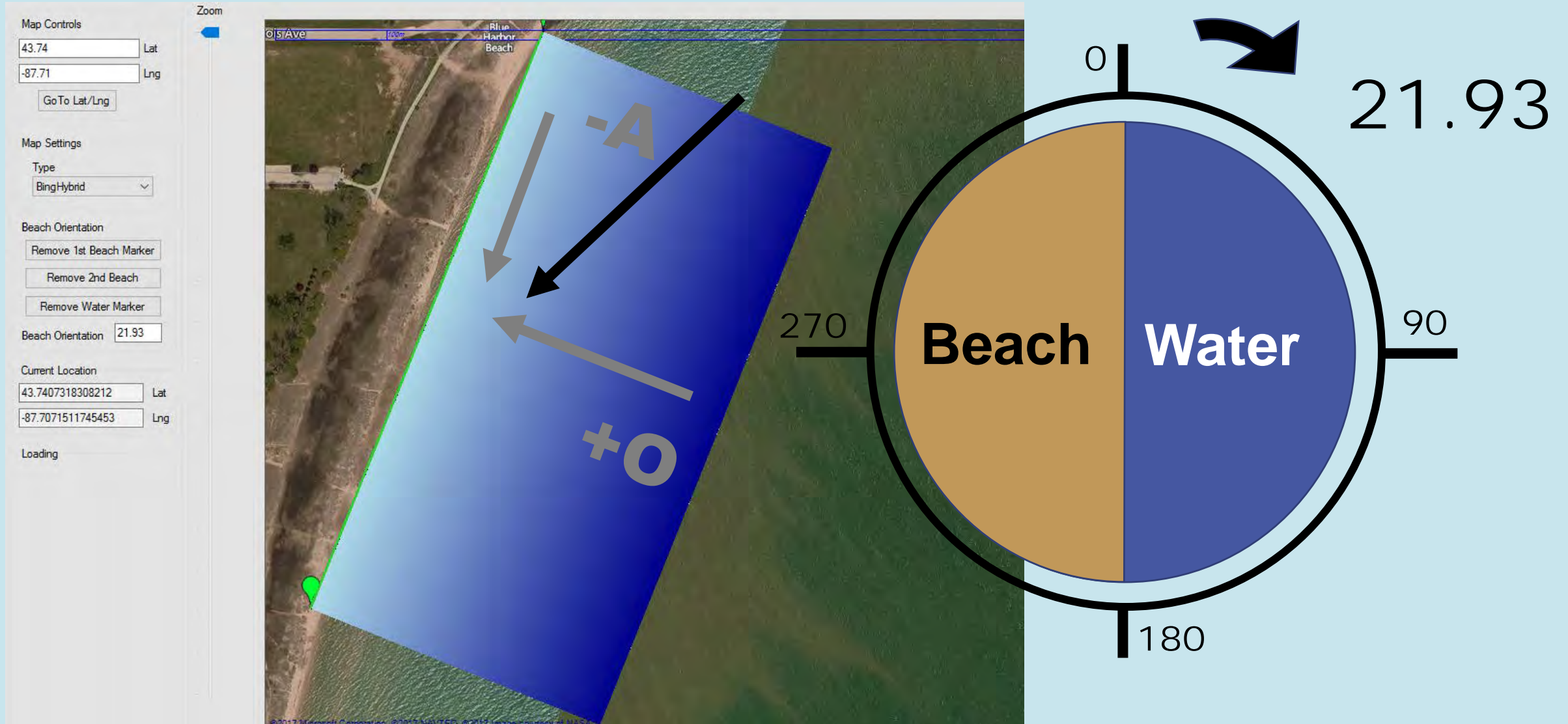
Current  
= 225



# Computing Alongshore & Onshore components



# Computing Alongshore & Onshore components



# Computing Alongshore & Onshore components

