

Comments Regarding the Ground Water Level (GWL) Percentile Ranking Index

Lucia Baldwin, PhD

Senior Scientist

Compliance & Reporting Section

FWRMC Salinity Network Workshop

November 9, 2016

Objectives

1. Discuss Percentile Rank functions in Excel
PERCENTRANK (Array, x)
2. Discuss the arguments of Percentile Rank function
 - Array / Data Set
 - Value (x) we want to rank against the Array
3. To ensure consistency across reporting agencies

1. Percentile Rank Functions in Excel

Array
21
22
22
23
23
24
25
28
37
45
48
67

What percentage of the data has values under **23**?

PERCENTRANK (array, **23**) = 27%

PERCENTRANK.**INC** (array, **23**) = **27%**

PERCENTRANK.**EXC** (array, **23**) = **31%**

Excel 2007 Excel

2010, 2013,
2016,

PercentRank.INC vs. PercentRank.EXC

Array
21
22
22
23
23
24
25
28
37
45
48
67

Array
21
22
22
23
23
24
25
28
37
45
48
67

PERCENTRANK.INC (array, 67) = 100% PERCENTRANK.INC (array, 21) = 0%

PERCENTRANK.EXC (array, 67) = 92% PERCENTRANK.EXC (array, 21) = 8%

2. Value (x) we want to rank

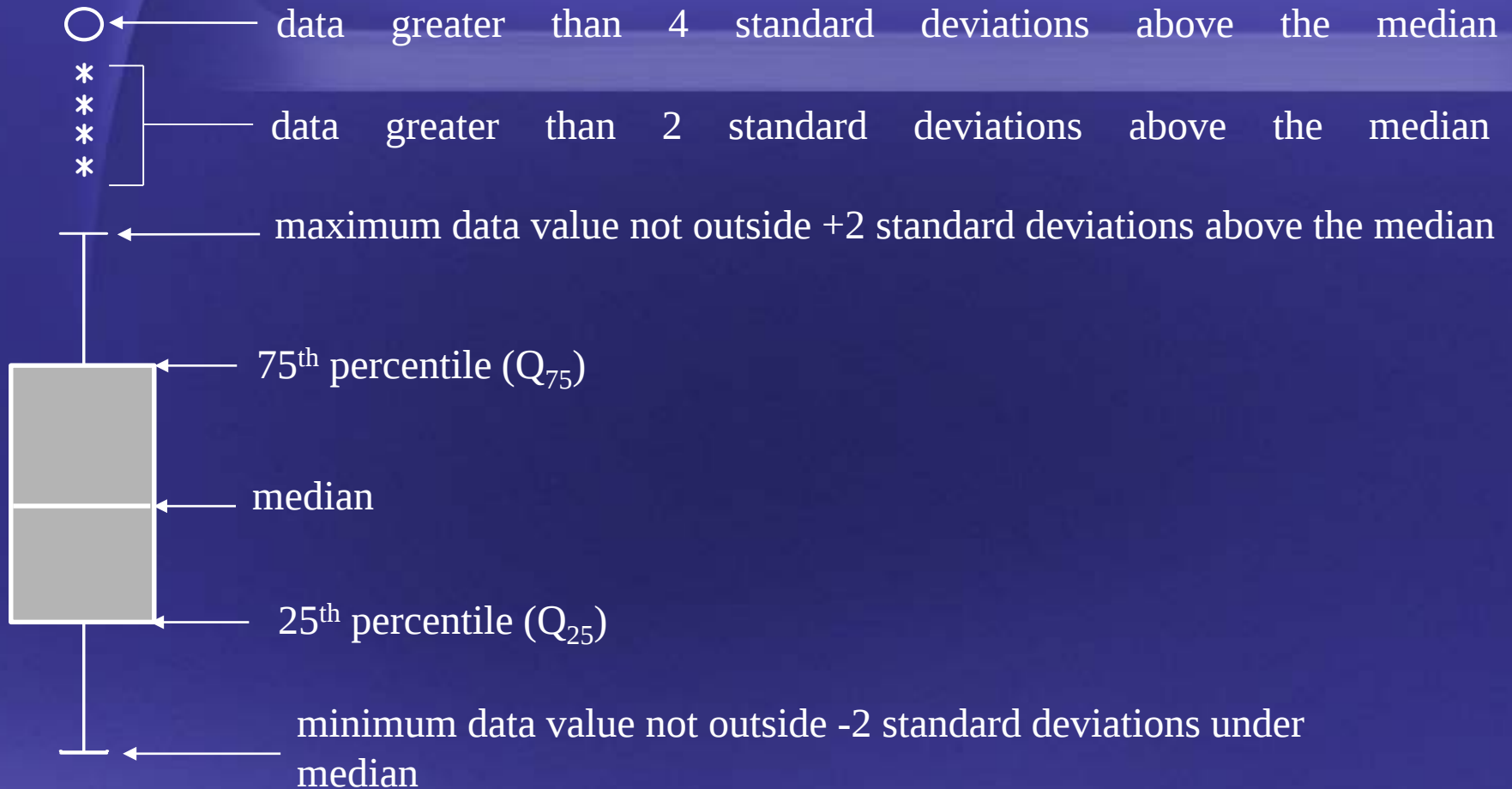
x is a GWL value representative for May / September of each year:

- monthly median of max daily values – **SFWMD**
(wells with max daily data and wells with 15-min. data)
- value registered nearest 15th of the month – **NFWMD**
(wells with daily, monthly or quarterly frequency data)
- daily mean on the 27th of the month – **Suwannee River WMD**
(wells with 15-min/ hourly data and manually-read well with a single values during first 2 weeks of the months)

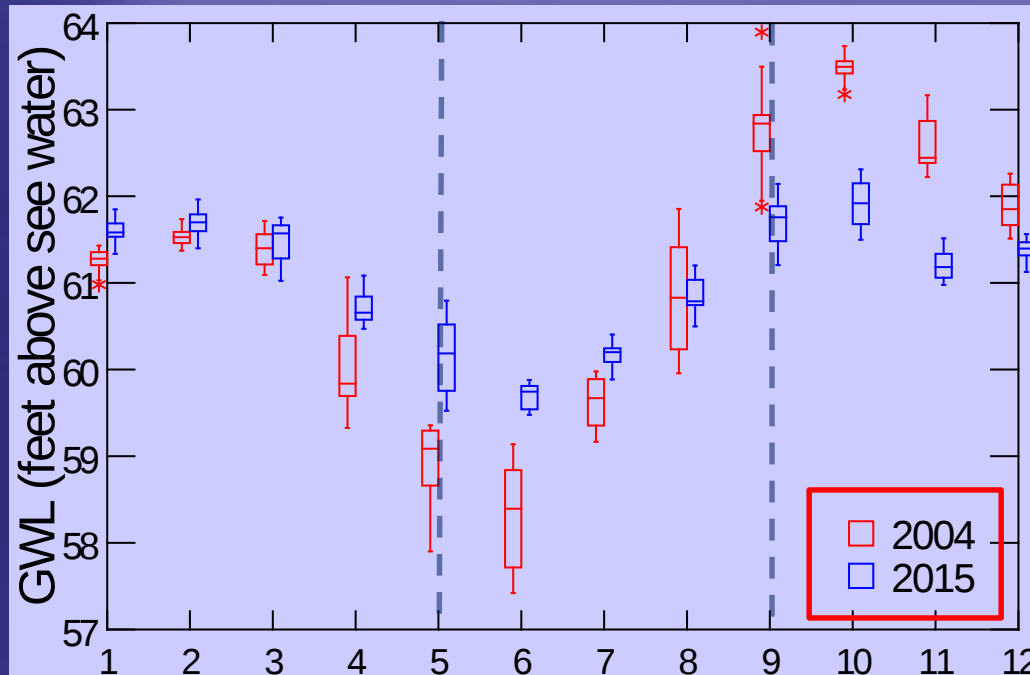
Option for Choosing the Array / Data Set Against Which Value x Is Ranked

- All similar values registered for entire period of records
- Similar values registered during the same month (week or season)

Box Plot Description



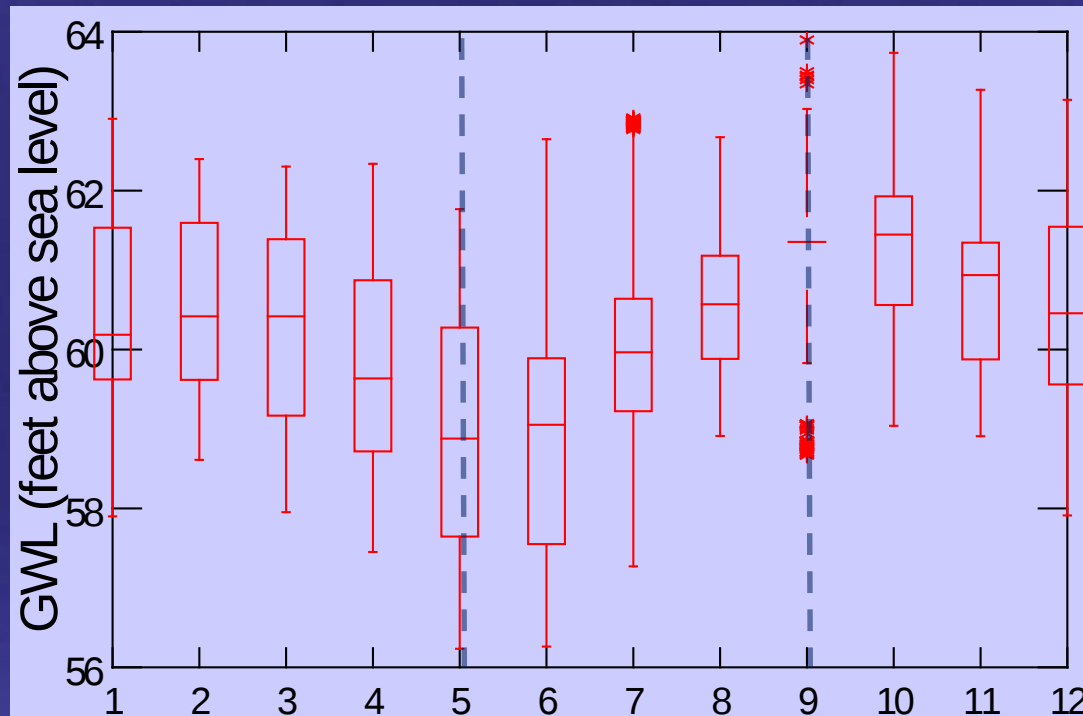
Daily Maximum GWL at Station POF-22 by Month



Note the seasonality of GWLs

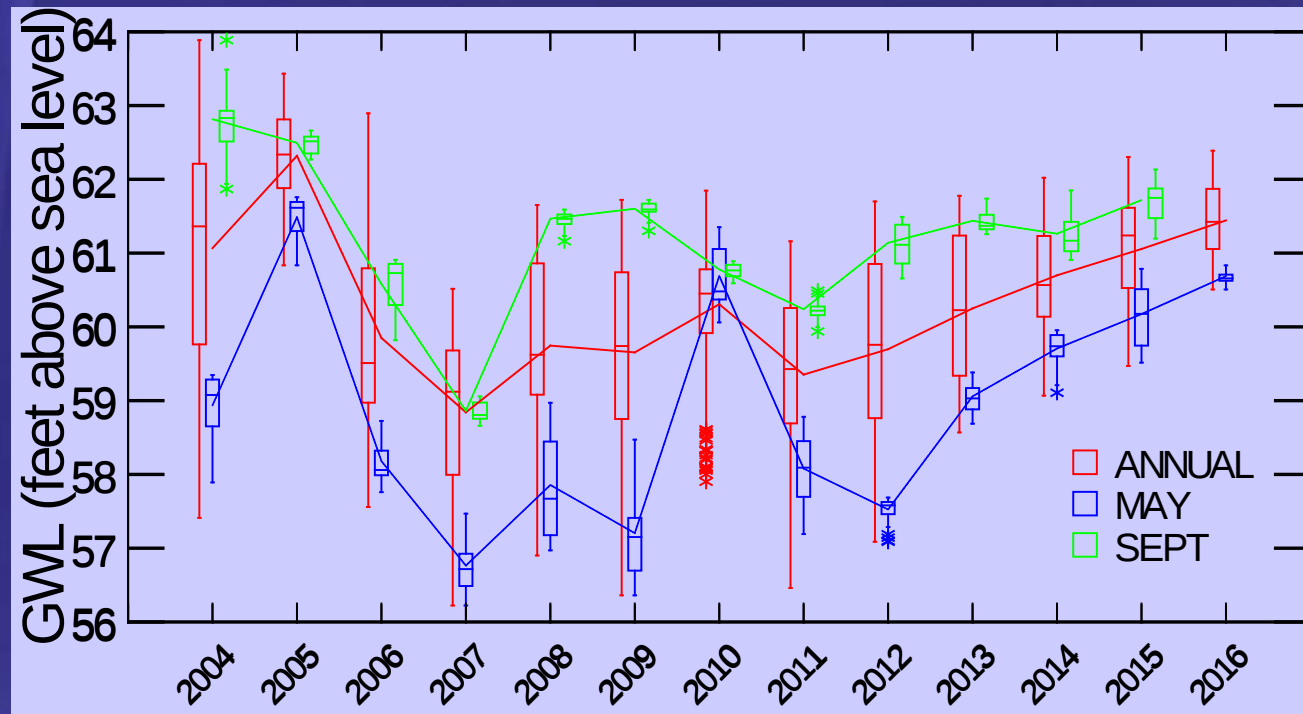
Low GWLs at the end /after dry season (May and June) High GWLs at the end /after wet season (September and October)

GWL at Station POF-22 by Month for Interval 1/1/2004 to 5/31/2016



Seasonality is the same for the entire period of records

Seasonality of GWL in S. Florida at Station POF-22 (2004 – 2016)



May GWLs are, in most cases, lower than annual GWLs
September GWLs are, in most cases, higher than annual

GWL

Percentile Rank of Median GWL for May 2016

Station	vs . all GWL	vs. May GWL
BF-6	39.7	68.1
G-3061	79.2	92.4
L2-PW2	38.8	67.1
MF-52	36.7	66.9
ORF-19	66.1	82.3
ORF-32	35.7	72.5
OSF-102	60.9	86.5
OSF-22	45.5	77.0
OSF-62	48.0	78.8

Station	vs . all GWL	vs. May GWL
OSF-64	48.1	78.1
OSF-66	47.3	79.7
PB-747	2.9	14.9
PBF-2	76.5	93.5
POF-22	58.8	84.8
SKYLAKE	25.1	62.9
SLF-21	43.9	84.0
SLF-75	38.7	81.5
SLF-76	46.0	82.5

Comparing the x value with lower values data set
yields higher percentile ranks.

Percentile Rank of Median GWL for September 2015

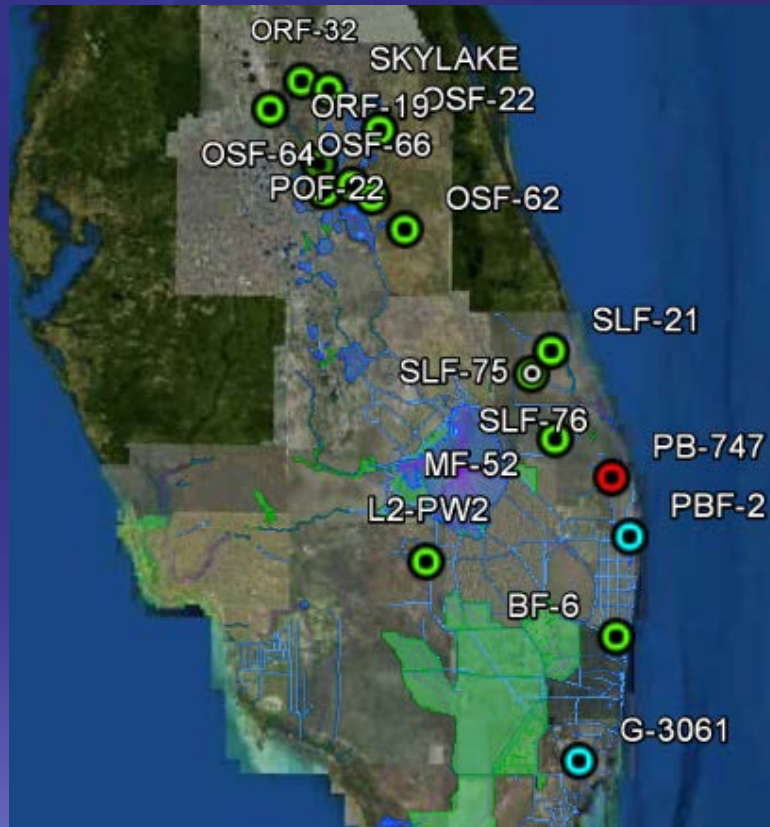
Station	vs. all GWL	vs. Sep. GWL
BF-6	69.6	38.6
G-3061	93.1	69.4
L2-PW2	75.6	34.4
MF-52	74.9	36.0
ORF-19	98.1	88.4
ORF-32	92.6	83.4
OSF-102	88.8	79.2
OSF-22	84.8	61.8
OSF-62	82.3	59.7

Station	vs. all GWL	vs. Sep. GWL
OSF-64	84.1	62.5
OSF-66	85.4	61.5
PB-747	37.9	6.0
PBF-2	89.3	80.5
POF-22	87.8	78.6
SKYLAKE	94.0	80.6
SLF-21	77.0	49.9
SLF-75	75.2	49.3
SLF-76	81.0	57.5

Comparing the x value with higher values data set yields lower percentile ranks.

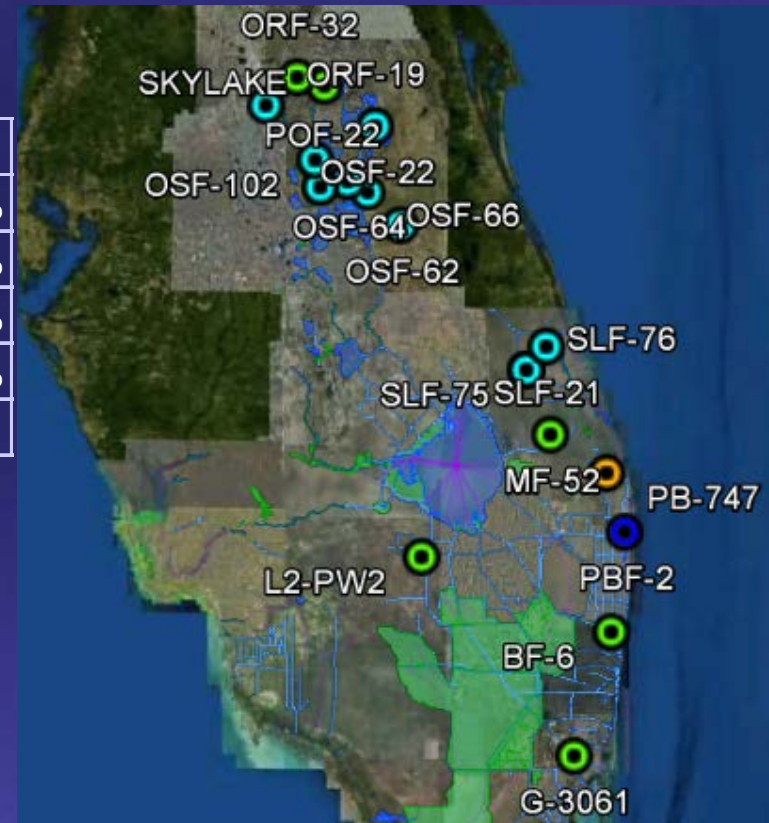
Mapped GWL Percentile Ranks of SFWMD Stations

PRs of May 2016 Median vs. **All**
Daily Values



Legend	
	> = 90 %
	76 - 89 %
	25 - 75 %
	11 - 24 %
	< = 10%

PRs of May 2016 Median vs. **May**
Daily Values



Reason to Use All GWL Data Set

- Missing values should be 5% or less of entire data set, but the percentage of missing value might be bigger for seasonal / restricted data sets.
- Seasonality might be different for different latitudes along Florida. S. Florida, **May** GWLs are among the lowest of the year and September GWLs are among the highest of the year.

Conclusion: Use all GWL data set not a seasonal subset

Questions and Comments

