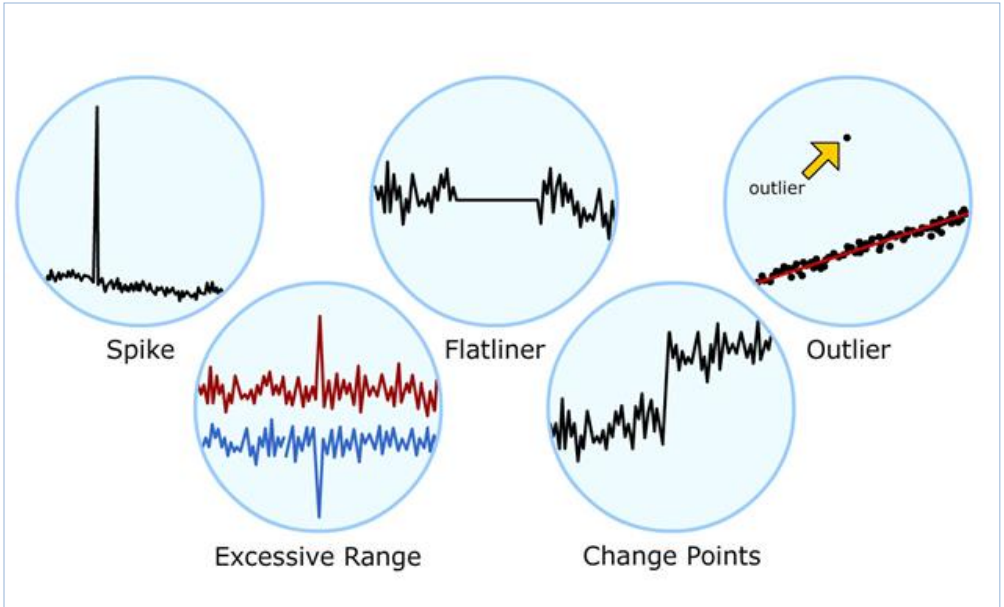
Quality Assessment of Point and Gridded Precipitation Data



Source: NOAA

RameshTeegavarapu Ph.D., P.E.
Professor,
Department of Civil, Environmental and Geomatics Engineering,
Florida Atlantic University, Boca Raton, Florida, 33431, USA
Director

HYDROSYSTEMS RESEARCH LABORATORY

ESEWG Meeting
August 20, 2019. Presentation

Precipitation Data & Issues

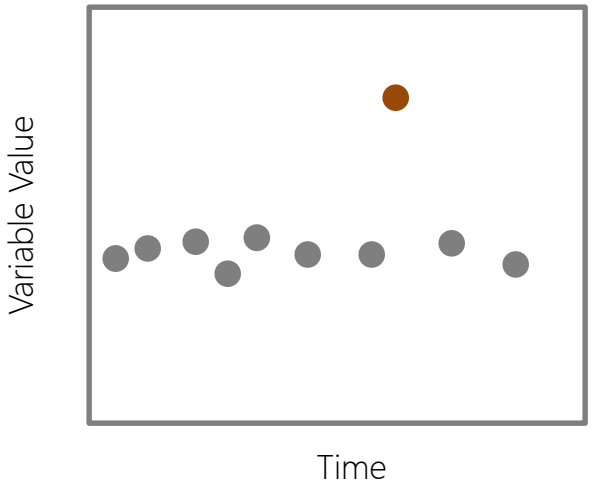
- Point Data: Precipitation data measured at rain gauges
- Gridded Data:
  - Radar/satellite data products
  - Spatially interpolated precipitation data products from either rain gauge data or radar/satellite sources
  - Multi-sensor data products

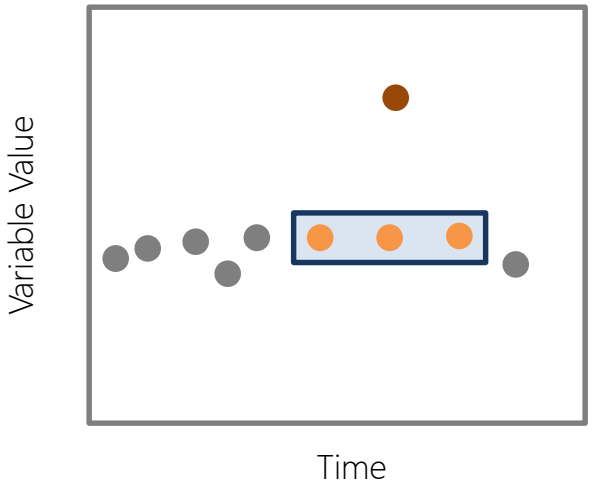
Issues with observations or estimations:

- Outliers and anomalies
  - Instrument –related
  - Gauge placement
  - Transcription or telemetry related
- Missing data (gaps in the data)
  - Systematic and random errors
  - Instrument malfunctions

Outlier and Anomalies

- The words "outlier" and "anomaly" are used interchangeably in many studies.
- An outlier is an unusual observation, numerically different from a set of observations.
- Anomaly refers to a pattern in a given data set that does not conform to an established normal behavior.
- Anomalies are more difficult to detect with statistical methods. Rule-based methods derived from expertise will help immensely.





Outlier

Anomalous observations

Data Issues and Use for Analysis

Form	Example	Meaning	Issues with the Data
Value	11.5	Reliable	No problems were found in the automatic quality control. The value was computed from a complete dataset.
Value)	11.5)	Quasi-Reliable	Only slight problems were found in the automatic quality control, or the value was computed from the dataset with a few missing data.
Value]	11.5]	Incomplete	The value was computed from a dataset with excessive missing data.
-	-	No phenomenon	No phenomenon was observed within the period.
X	X	Missing	No value is available due to problems with observation instruments, etc.
Blank		Out of observation	No observation was conducted.
*	31*	Most recent extreme values	The value is the most recently observed of those two or more identical daily extreme values in the period.
#	#	Suspicious	A serious quality problem was found in the value, treated as omitted from the statistics.

Source: JMA, as reported by JMA

Potential causes of anomalies

Type of Rain Gage	Potential Causes of Anomalies during Measurements	Anomalies in Measured Data
Tipping Bucket		
(recording)	Loss of rain during tipping in case of heavy rain	Underestimation of rain
	Evaporation loss	Underestimation of rain
	Water adherence (retention) to buckets	Underestimation of rain
	Water leaking during heavy rain	Overestimation of rain
	Delay in tipping	Time shift in reported rainfall value
	Data transmission issues	Under or overestimation or missing observations
	In and Out- splashing	Under or overestimation
	Wind field deformation around the gage	Underestimation
	Observational and instrumental errors	Under or overestimation
	Equipment failure	No measurement
Weighing Type		
(recording)	Wind field deformation around the gage	Underestimation
	Evaporation loss	Underestimation
	In and Out- splashing	Under or overestimation
	Observational and instrumental errors	Under or overestimation
	Equipment failure	No measurement
	Data transmission issues	Under or overestimation or missing observations
	Equipment failure	No measurement
Float-Type		
(recording)	Observational and instrumental errors	Under or overestimation
	Equipment failure	No measurement
	Data transmission issues	Errors in reporting process
	Equipment failure	No measurement
Standard Rain Gage		
(non-recording)	Evaporation loss (during light rain)	Underestimation
	Data transmission issues	Under or overestimation or missing observations
	In and Out- splashing	Under or overestimation
	Wind field deformation around the gage	Under or overestimation
	Observational and instrumental errors	Under or overestimation

Point Data Issues

- Gaps in the data
- Physically impossible values
- Constant values
- Values above pre-specified thresholds
- Improbable zero values
- Unusually low values (probably real values)
- Unusually high values (probably real values)

Outliers and Anomalies

- Gaps in the data
  - Missing precipitation values due to a number of reasons.
- Physically impossible values
  - Negative rainfall values, unusual intensities
- Constant values
  - Similar precipitation values over multiple time intervals
- Values above pre-specified thresholds
  - Values exceeding historic extremes (rare extremes ?)
- Improbable zero values
  - Rainfall in the area, however not registered at the gauge of interest
- Unusually low values (probably real values)
  - Low values (low totals) compared to nearby gauges
- Unusually high values (probably real values)
  - High values compared to nearby gauges

Issues with Precipitation Measurements

- An automated approach needs to developed to:
  - Detect suspect data
  - Identify any data errors due to outliers and anomalies
- Time Shifts
  - Recorded data shifted in either direction
- Errors
  - Values greater than a pre-specified value (defined by information from historical data)
  - Missing data

Evaluation of Timeseries

- Examination of rainfall time series
  - Graphical plotting
    - Evaluation of daily, historical and months values
  - Gaps, overlaps and relationships are identified.
  - Observations at the adjacent rain gauges in vicinity for the period-of-interest are analyzed
  - Check rainfall data at a selected site
    - Minimum
    - Maximum
    - Missing data
    - Overlapping period-of-record
    - Checking with threshold amounts from a reference site to ensure no operational problems exist.
  - Outlier identification – Use of boxplots recommended.

Methodologies

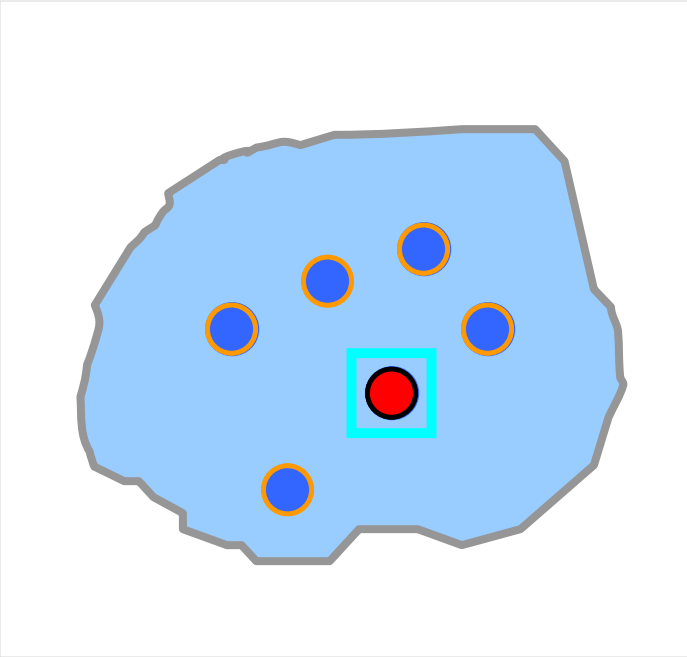
- Univariate time Series approach
  - Identification of any outliers and anomalies
- Neighborhood (multi-site approach)
  - Multiple rain gauges in the vicinity
    - Comparison of time consistent observations
    - Mass curves
    - Rainfall totals
  - Radar data
    - Evaluation using single collocated pixel (i.e. grid)
    - Evaluation using surrounding pixels, including the collocated pixel (i.e. grid) in which rain gauge is located.

Methodologies

- Univariate Approaches
  - Evaluation of time series of precipitation data at a single site
  - Identify outliers and anomalies
    - Using historical data from the same time period (month or season) for boundary consistency checks (range checks).
      - Threshold method
        - Limits are obtained from the historical data
        - Limits can be dynamic and temporally and spatially varying
        - Different threshold values can be used for different time scales
          - Physically impossible values
    - Simple rule-based methods – Precipitation data-specific rules
    - Statistical methods
      - Boxplot
      - Adjusted boxplot
      - Others that might be applicable to precipitation data

Methodologies

- Neighborhood Approaches
  - Using Rain Gauges



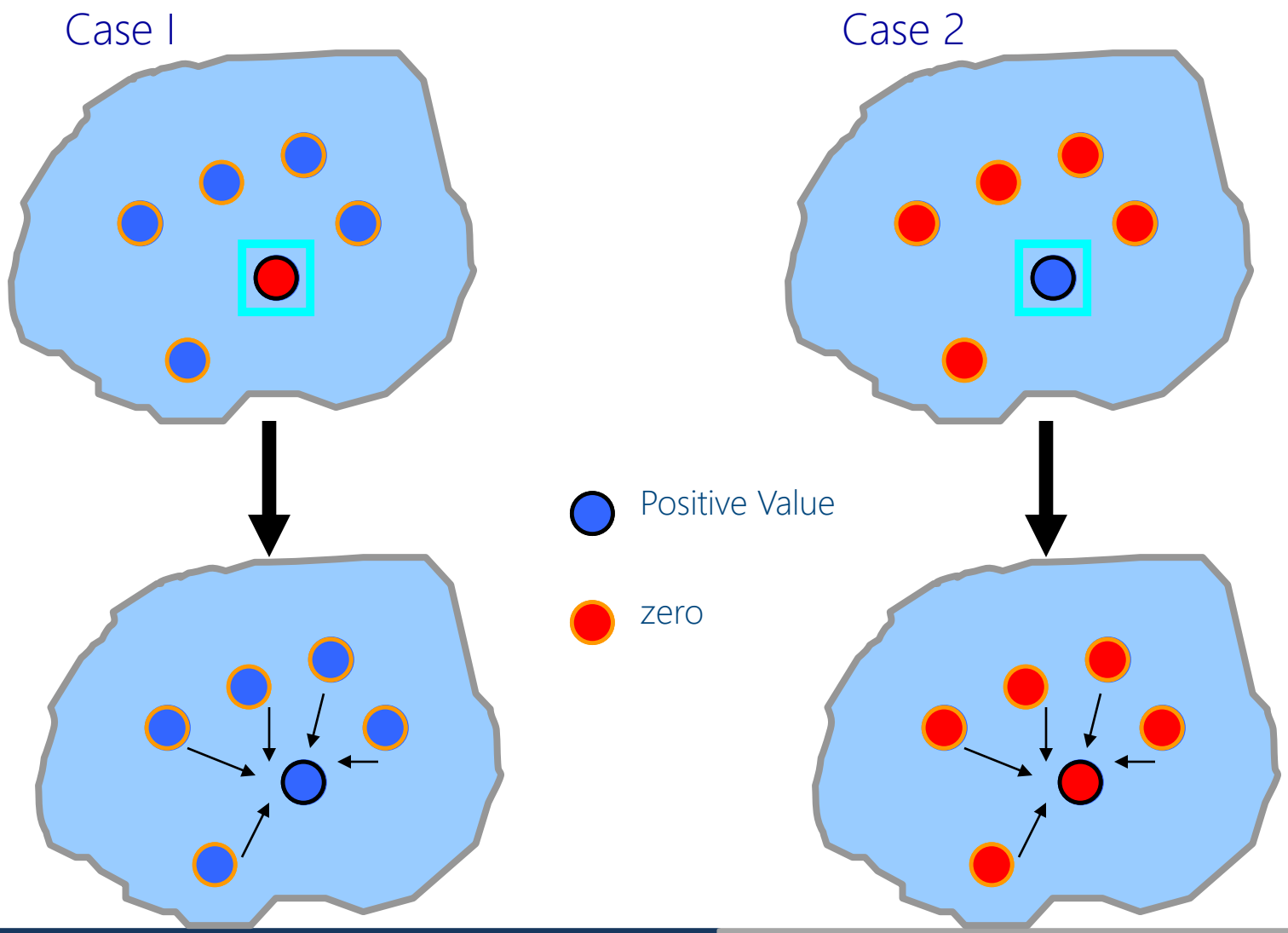
Selection of single or multiple rain gauges

Criteria for selection of rain gauges

- Proximity (distance)
- Proximity (correlation)
- Proximity (distribution similarity)

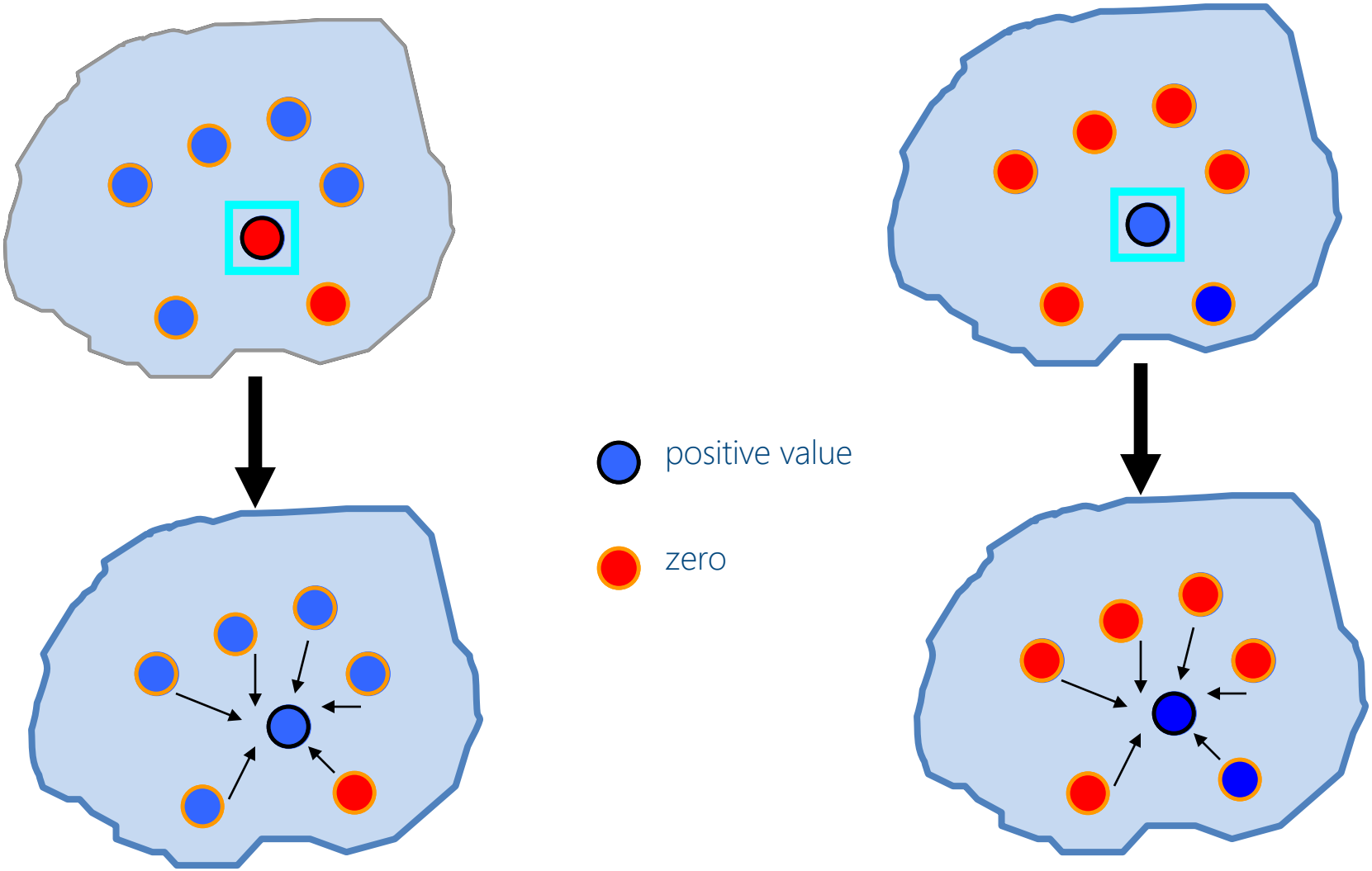
Using Information from Surrounding Rain gauges

Potential Problems



Potential Problems

Solution – using nearest neighbor

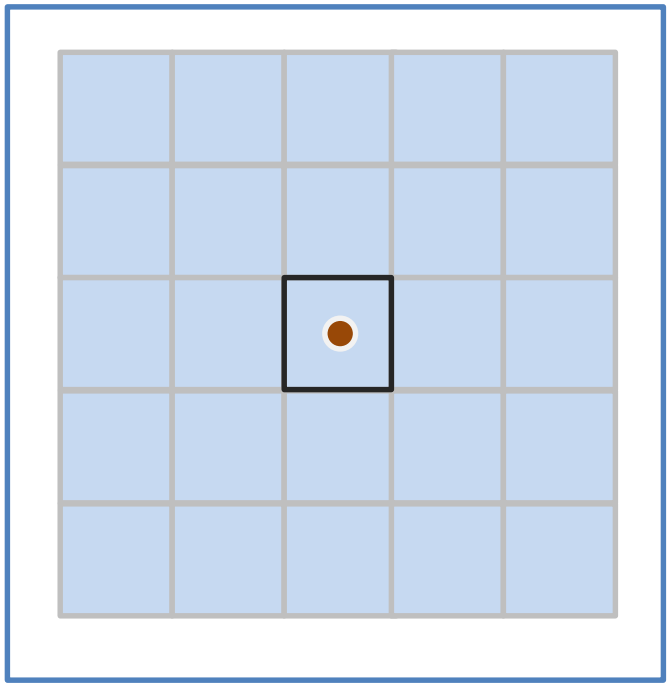


Neighborhood Approach

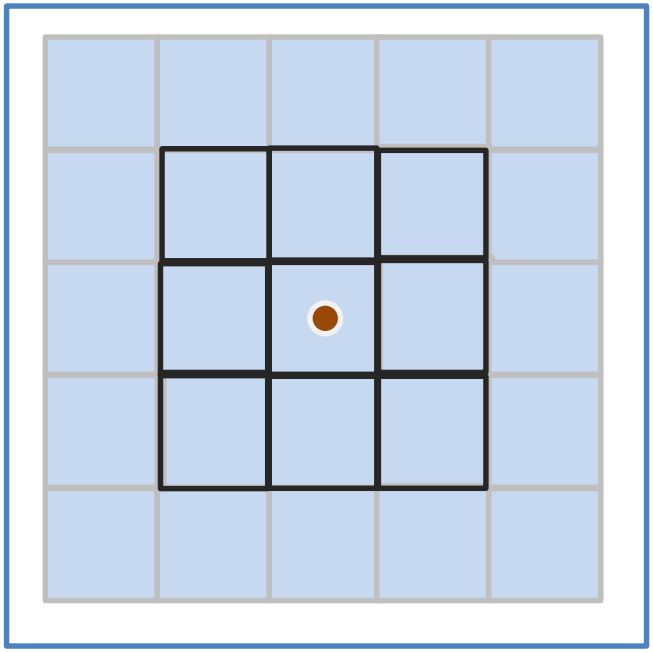
- Neighborhood Approach
  - Using Radar Data

Selection of nearby grids for assessment of outliers

- Collocated pixel (grid)

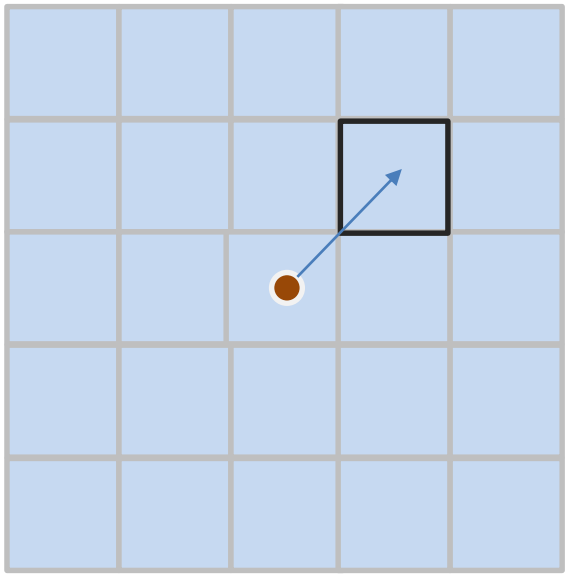


Nearby nine pixels (grids)



Rain gauge

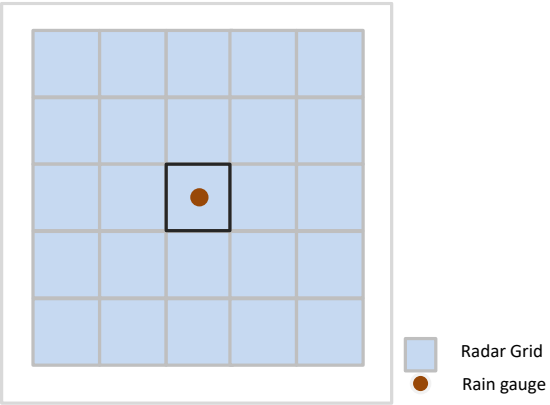
- Highest correlated pixel among the surrounding 9 pixels (grids) including the collocated grid.



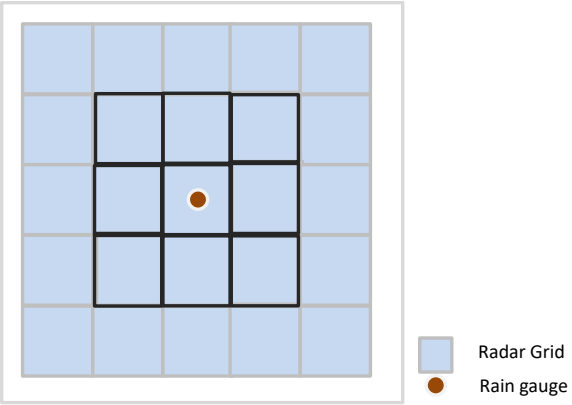
- Historical precipitation data from the rain gauge and radar data from 9 pixels (grids) need to be evaluated.
    - Highest correlation grid may change depending on:
        - Month
        - Season

Single Cell Approach

- Radar data available at the same temporal resolution as rain gage data can be used for quality control of rain gage data. Radar data can be used in three different ways:
- Single Cell Approach (SCA): Data available from radar data grid (or pixel) that is co-located with rain gage.
- In this approach a single radar grid (pixel) in which the rain gage is co-located is selected for quality control of rain gage data. Co-location of rain gage with radar grid can be achieved using any spatial analysis software (e.g. ArcGIS).
- Large difference between observed rain gage precipitation and radar-based estimate in a specific temporal interval can be considered for further investigation.



Fixed Multiple Cell Approach (FMCA)

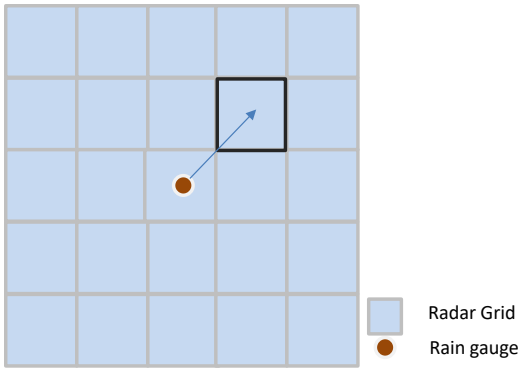


Fixed Multiple Cell Approach (FMCA): Data available from radar data grids (or pixel) surrounding rain gage (9 pixels) including the grid that is co-located with rain gage.

In this approach data from a fixed number of radar grids (pixels) surrounding the grid in which rain gage is co-located and the grid in which rain gage is located is selected for quality control of rain gage data. Large difference between observed rain gage precipitation and radar-based estimates in a specific temporal interval can be considered for further investigation.

Best Cell Approach (BCA)

- BCA method :
- Selection of one grid (i.e. best cell) in proximity to rain gage based on a criterion of highest correlation coefficient. The correlation coefficient is determined using historical rainfall observations at rain gage and grids.
- Selection of one grid using any proximity metric (i.e. a numeric metric that measures the similarity of two sets of observations). A total 20 proximity metrics were used in a recent study (Teegavarapu, 2012).
- Selection of one grid using an optimization formulation that minimizes the difference between rain gage measurements and radar-based estimates from any grid with a constraint that helps in selecting only one cell. This is enforced by using a binary constraint in the formulation. The optimization formulation is solved using historical rain gage and radar data for a specific period (e.g. month, season).
- The best cell selectin of BCA approach can be obtained using gage and radar-based estimates using a temporal window one hour or higher. The best cell may be identified for each site for different months or seasons



In this approach optimal selection of one of the immediately neighboring grids (including the grid in which rain gage is located) is carried. The selection can be carried by using historical rain gage and radar data and developing an optimization formulation or using highest correlated grid.

Anomaly and Outlier identification methods

Method	Suitability for Quality control of Precipitation Data	Conditions	Limitations
Univariate or Single Site Observation-based Methods			
• Statistical			
Control charts	Yes	No distributional assumptions.	Limited ability to deal with large data sets
Tukey' s Boxplot or derivatives	Yes	No assumptions about the distribution (non-parametric).	None Not a robust method
Adjusted Boxplot	Yes	None	None
3-Sigma method	Yes	None	Not a robust method
Median Rule and Mean Absolute Deviation method	Yes	No distribution specific assumptions are required.	None
Distribution-based	Yes	Need confirmation of specific distribution	None
• Rule-Based Method(s)			
Rule-based (based on rain gage measurement limits and physically possible rainfall totals for a specific duration • Boundary and spatial consistency checks, constant value checks	Yes	Requires data about rain gage type and historical region-specific precipitation extremes data for different durations.	Rules need to be exhaustive to consider all possible combinations.

Anomaly and Outlier identification methods

Method	Suitability for Quality control of Precipitation Data	Conditions	Limitations
Clustering (K-nearest neighbor method)	Yes	Requires determination of clusters.	Non-supervisory mode of cluster development may lead to non-robust method of anomaly detection. Method not suitable for constant data anomalies.
Neighborhood Value-based Methods			
Rain gage and radar data-based methods	Yes	Requires representative precipitation data from nearby gages or radar grids. SCA requires data from one cell and FMCA requires data multiple cells and BCA requires data from one cell.	The number of rain gages to be used as reference sites need to be established first. Radar grids need to be selected for three radar-based approached outlined.
Regression Models/Spatial Interpolation models/Thiessen Polygon Approach	Yes	Requires availability of reference sites that are representative of base site.	Accuracy depends on the results of the spatial interpolation.

Missing Data Estimation

- Identify anomalies and outliers associated with specific types of rain gages due to systematic errors and possible equipment malfunctions.

- Develop methods to estimate missing precipitation data after an initial evaluation of several estimation methods using spatial or temporal interpolation with the help of: (1) rain gage(s) and (2) available radar-based precipitation data. Spatial interpolation methods based on deterministic (e.g. Thiessen polygon) and stochastic interpolation methodologies (e.g. kriging) may be used.

- Evaluate the possibility of using raw reflectivity values available from different radar sites for preliminary estimation of: (1) magnitude of rain; (2) rain or no rain conditions and (3) spatial distribution of rainfall. Reflectivity-rainfall relationships can be improved and radar-based estimates can be used for improvement of spatial interpolation methods for estimation of missing precipitation data.

Observations/Conclusions

- To develop a data quality evaluation system we need to consider
  - Internal consistency (at a site of interest)
    - Boundary consistency
  - Spatial consistency
    - Based on Neighborhood stations
- All the outlier and anomaly detection methods for precipitation can be used at different temporal scales.
- Comprehensive assessment of historical data is critical to develop rule-based methods.

Methods available

Method(s)	Reference
Simple Rule-based Methods Boundary and spatial consistency checks, Basic value checks, constant values, ambiguity in date, time shifts in measurements, gaps in data; check for transmission errors	Nie Su-Ping et al. (2012); Hamada et al. (2011); Einfalt et al. (2008); Pelczer and Cisneros-Iturbe (2007); Einfalt et al. (2000) ; Einfalt et al. (2006); Shulski et al. (2014); Durre et al. (2010); Salvati and Brambilla (2008); Pajari (2014); Flemming et al. (2002); Zahumensky (2004), Teegavarapu (2014)
Statistical methods, outlier detection, spatial consistency checks	Krajewski (1986)
Radar data-based checks	Golz et al. (2006); Golz et al. (2004)
Polynomial interpolation, spatial regression, interpolation	Steinacker et al. (2011)
Objective Analysis, temporal consistency checks, spatial interpolation	Eischied et al. (1995)
Data mining, outlier detection, univariate methods	Wu et al. (2008)
Summation checks (daily, monthly), spatial evaluation, underestimation reported.	Hillaker and Andsager (2008)
Step change, spatial regression, persistence (constant values)	Hubbard et al. (2005)
Mass curves, evaluation of extremes	Jorgensen et al. (1998), Teegavarapu (2014)
Spatial evaluation, neighborhood methods	Kondragunta (2001); Kondragunta and Shresta (2006)
Manual assessments, extreme value checks	Kunkel et al. (2005)
Neighborhood approach for missing value estimation	Maul-Kotter and Einfalt (1998)
Simple rainfall evaluation methods, consistency checks	McKee et al. (2015)
Neural network approach – spatial interpolation	Sciuto et al. (2009)
Kriging and co-kriging – spatial interpolation	Yeung et al. (2014)
Statistical distributions test, neighborhood distribution comparison	You et al. (2007)

Publications

Ramesh S. V. Teegavarapu, Chandramouli, V. 2005. Improved weighting methods, deterministic and stochastic data-driven* models for estimation of missing precipitation records, Journal of Hydrology, 312, 191-206.

Ramesh S. V. Teegavarapu, 2007. Use of *universal function approximation* in variance-dependent interpolation technique: An application in Hydrology, Journal of Hydrology, 332, 16-29.

Ramesh S. V. Teegavarapu, 2009. Estimation of missing precipitation records integrating *surface interpolation techniques and spatio-temporal association rules*, Journal of Hydroinformatics, 11(2), 133 – 146.

Ramesh S. V. Teegavarapu, et al., 2009. *Optimal functional forms* for estimation of missing precipitation records, Journal of Hydrology, 374, 106-115.

Ramesh S. V. Teegavarapu, 2012. Spatial interpolation using *nonlinear mathematical programming models* for estimation of missing precipitation records, Hydrological Sciences Journal, 57(3), 383-406.

Precipitation Estimation and Corrections

- **Ramesh S. V. Teegavarapu**, Missing Precipitation Data Estimation using Optimal Proximity Metric-based Imputation, Nearest Neighbor Classification and Cluster-based Interpolation Methods, Hydrological Sciences Journal, 2013.

- **Ramesh S. V. Teegavarapu,** Statistical Corrections of Spatially Interpolated Precipitation Estimates, Hydrological Processes, 28(11), 3789–3808, 2014.

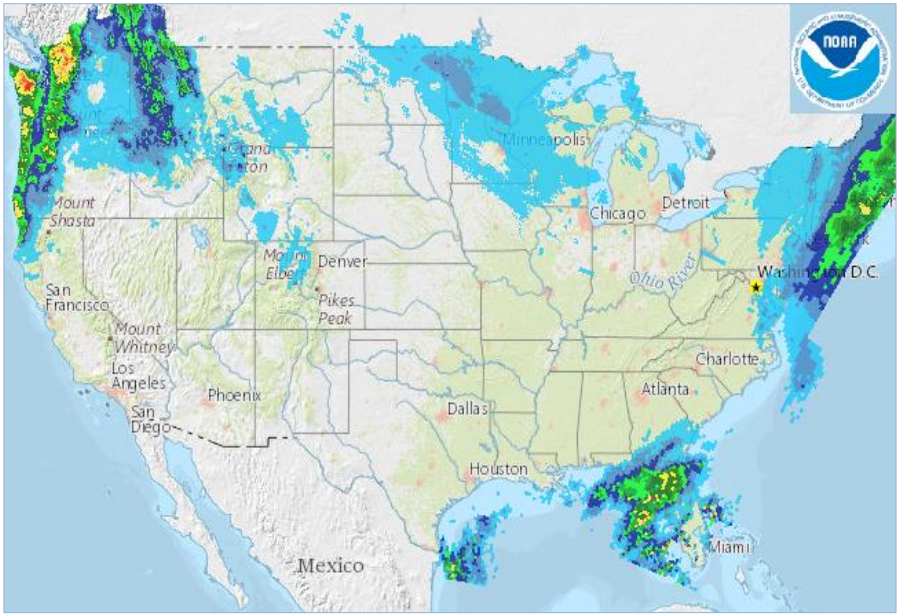
- **Ramesh S. V. Teegavarapu**, et al., Infilling Missing Precipitation Records using Variants of Spatial Interpolation and Data-Driven Methods: Use of Optimal Weighting Parameters and Nearest Neighbor-based Corrections, International Journal of Climatology, 2017. Published

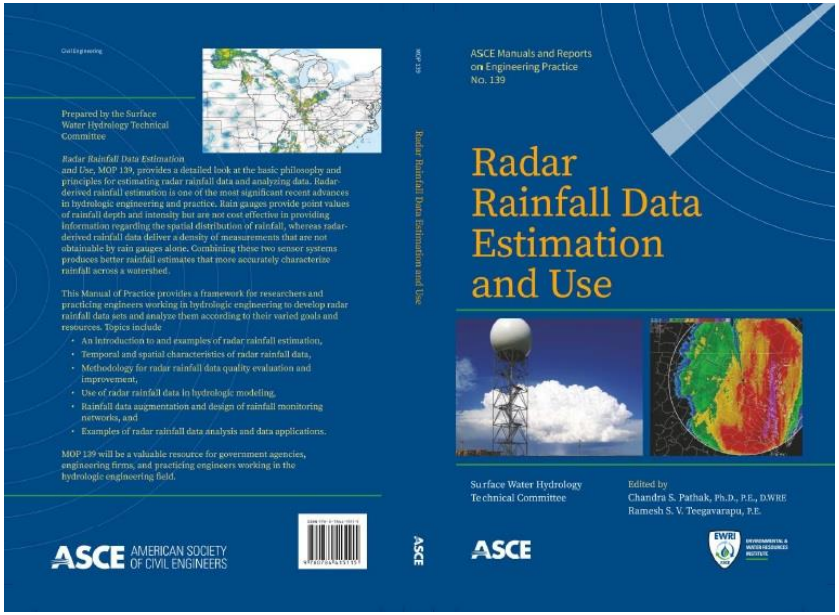
- **Ramesh S. V. Teegavarapu** and Anurag Nayak. Evaluation of Long-term Trends in Extreme Precipitation: Implications of Infilled Historical Data and Temporal Window-based Analysis, Journal of Hydrology, 2017. Published.

- **Ramesh S. V. Teegavarapu**, Spatial and Temporal Estimation and Analysis of Precipitation, Handbook of Applied Hydrology, McGraw Hill, 2016, Published.
- **Ramesh S. V. Teegavarapu**, Climate variability and changes in precipitation extremes and characteristics, Springer, 2016. Published.
- **Ramesh S. V. Teegavarapu**, Evaluation and Improvement of Radar-based Rainfall and Design of Monitoring Networks, Manual of Standard Practice for Radar Rainfall Data Estimation, ASCE, 2018..
- **Ramesh S. V. Teegavarapu**, Precipitation Data Augmentation and Analysis of Radar-based Rainfall Data, Manual of Standard Practice for Radar Rainfall Data Estimation, 2018, ASCE.
- **Ramesh S. V. Teegavarapu**, Statistical Analysis of Precipitation Extremes, Chapter in ASCE Book on Statistical Distributions in Hydrology: ASCE. 2019.

- **Ramesh S. V. Teegavarapu,** Tadesse Meskele and Chandra Pathak, Geo-Spatial Grid-based Transformation of Multi-Sensor Precipitation using Spatial Interpolation Methods, Computers and Geosciences, 2012.
- **Ramesh S. V. Teegavarapu**, Aneesh Goly, Qinglong Wu, Comprehensive Framework for Assessment of Radar-based Precipitation Data Estimates, Journal of Hydrologic Engineering, ASCE, 2016.
- Chandra S. Pathak, **Ramesh. S. V. Teegavarapu**, D. Curtis and C. Collier, Radar Rainfall and Operation Hydrology, Journal of Hydrologic Engineering, ASCE., 2016.

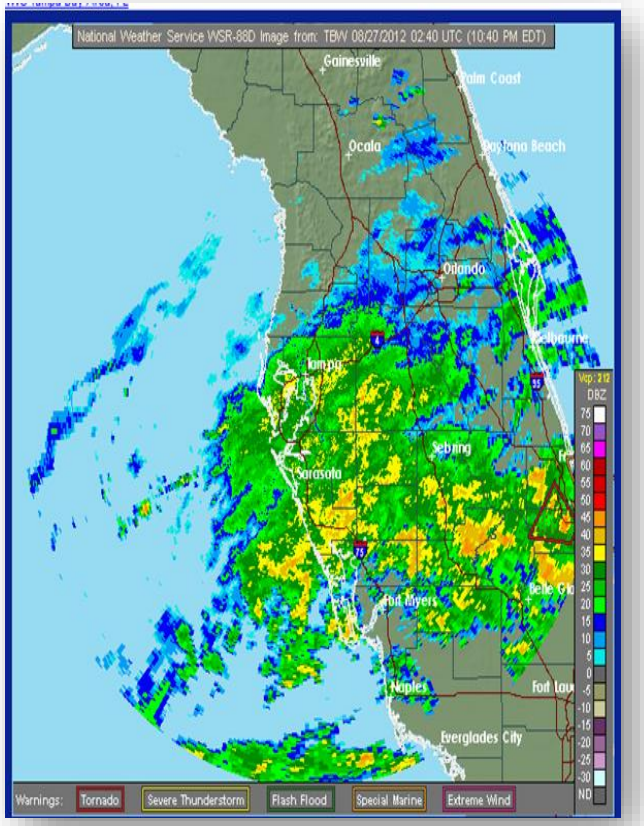
Assessment of Radar Data





Radar-based Precipitation Analysis and Estimation

- Radar-based rainfall estimation - Improvement using optimal Z-R relationships in space and time. Site-specific and seasonal relationships. Work related to S-band radar.
- Development of approaches for infilling missing rain-gauge data using radar-based precipitation estimates, gridded data generation°.
- Development of methods for use of Radar-based precipitation data for design of optimal rain gauge monitoring networks*
- Development of methodologies and indices for spatial-temporal analysis of biases in radar-based precipitation estimates - Development of tools+.
- Correction of biases in radar-based estimates
- Development of frameworks/tools for integrating radar-based precipitation data products to hydrologic simulation models
- Use of radar-data for identifying outliers and anomalies in rain gauge data sets.



*Teegavarapu et al., Journal of Spatial Science, 2015,
+Teegavarapu, et al., Journal of Hydrologic Engineering, 2015; ° Teegavarapu et al., 2012, Computers and Geosciences

Assessment of Radar Data

- The main question is : How good are these estimates?*
- Other questions :
  - How do we assess the data ?
  - What type of assessment framework is need ?
  - How comprehensive the assessment should be for specific application ?
  - How the assessment can help in improvement of radar-based data ?
  - How assessment results can be communicated to the users ?

Requirements :

- To assess radar data, error free, chronological, quality assured and quality controlled (QAQC) rain gauge data are required.
- Reasonable rain gauge density to facilitate assessment.
- Rain gauge data that is spatially uniform over the radar cover area
- Rain gauge data are available at the same temporal resolution as radar-based data.

* Krajewski and Smith (2002)

Assessments

- Visual
  - Visual measures are expected to provide a quick check of agreement between radar and rain gage datasets.
- Numeric
  - Contingency measures, skill scores, and a few new metrics are proposed. Quantitative measures are expected to provide information about errors, performance measures and skill scores assess the quality of radar data for dichotomous (rain and no-rain events) based on two datasets.
- Statistical
  - Summary statistics and hypothesis tests are expected to provide insights into distributional aspects of the data.

Questions to be Answered

- What is the overall quality of the radar data compared to rain gage data (i.e., ground truth)?
- Are the radar and rainfall data characteristics statistically similar?
- What is the overall bias?
- What is the error structure and how it is varying in space and time?
- What is variability of the bias in space and time?
- What is the skill of the method used for radar-based rainfall estimation in the definition of rain or no rain events?
- Do the radar-based rainfall estimates preserve the site (at a rain gage location) and regional (a set of rain gages) statistics?
- Are the rainfall fields generated by radar-based estimates and rain gage-based observations similar?
- Can a specific index or a set of indices help provide directions to improve the radar-based rainfall data?
- How do radar and rain gage data quantitatively compare at different temporal resolutions?
- Is radar data quality appropriate for distributed hydrologic modeling?

Objective
Development

Framework

- Assessment of bias in radar-based precipitation estimates compared to rain gage observations is a critical exercise to evaluate pre- and post-corrected (gage adjusted) radar-based precipitation data.
- A comprehensive bias assessment framework combining several visual, quantitative and statistical measures, indices and skill scores is proposed and developed for evaluation of radar-based precipitation estimates in space and time.
- Visual measures are expected to provide a quick check of agreement between radar and rain gage datasets. Contingency measures, skill scores and a few new metrics are proposed.
- Quantitative measures are expected to provide information about errors, performance measures and skill scores assess the quality of radar data for dichotomous (rain and no-rain events) based on two datasets.
- Summary statistics and hypothesis tests are expected to provide insights into distributional aspects of the data.

Bias Assessment Framework

- A comprehensive bias assessment framework is based on several measures adopted from:
  - Quantitative error and performance measures from hydrologic modeling
  - Skill scores from forecast verification
  - Proximity metrics from numerical taxonomy
  - Resampling methods
  - Exploratory data analysis and spatial analysis (GIS)
  - Quantitative and visual statistics
  - Distributional and hypothesis testing methods
  - Time series analysis
  - Analysis of residuals
  - Multi-model comparative analysis methods

Bias Evaluation Categories

Temporal

Temporal Resolution
- 15 minute
- Hour, Month

Seasonal
- Dry, Wet and Transitional

Annual
- Yearly, multi-year

Spatial

Extent

- Whole region
- Homogeneous Rainfall areas
- Specific Select Areas

Event

- Rainfall Threshold
- Selected Events
- Hurricane Rainfall Events
- Selected Events with Specific Storm Type

Point

Specific Site

- Select Rain Gages
- Rain Gage Clusters

Assessment of Data

Indices, Skill scores, Measures,

Measures, Indices and Evaluation Tests and Methods	Utility
Contingency Measures	Assessment of estimates using binary evaluations (states of the system)
Skill Scores	Assess the skill of method/estimate
Performance Measures (Error Measures)	Measures for evaluation of overall model accuracy and capabilities
Transitional Probabilities	Evaluation of persistence for different temporal state transitions
Association Measures and Plots	Bi-Variate assessment of data
Taylor Diagram	Assessment of multiple models and performances of models over time
Summary Statistical Measures	Simple measures for testing similarity in central tendency, range and variability between two data sets

Assessment of Data

Measures, Indices and Evaluation Tests and Methods	Utility
Serial Autocorrelation	Persistence at different temporal intervals
Quantile Plots	Similarity of two data distributions
Variograms	Assess spatial variability of data
Bootstrap Sampling based Confidence Intervals	Assessment of errors based on re-sampling of error measures
Infilling or Replacement Tests	Validity of data for replacing missing data
Evaluation of Residuals	Check for method bias, structure (temporal) and assessment of normality
Cumulative Plots	Check for consistency/deviations in data sets over time
Kernel Density Estimates	Non-parametric assessment of data for comparisons
Statistical Hypothesis tests	Two sample tests for confirmation of distribution characteristics of data
Data Distribution Evaluation	Similarity of Distributions

Critical Assumptions

- Radar-based precipitation estimates are adjusted based on rain gage observations.
- Radar-based precipitation estimates are not adjusted for each pixel (or spatial grid) in which rain gage is located or the adjustments are not based on one single rain gage.
- Bias correction procedures are generally applied to post-correction of radar-estimated data based on a specific Z-R (reflectivity-rainfall rate) relationship and by using rain gage data from a spatial network of gages.
- Observed precipitation data available from rain gages is not completely error free. However, the data available from the rain gages is assumed to be ground truth and error free without any systematic errors.
- The rain gage data should be critically evaluated before it can be used for bias analysis.

Assumptions…

- It is expected that radar-based observations improve over time starting from near real-time to end of the day and to finally end of the month. This is mainly due to improved rain gage data quality available from the time of collection to the point in time when these data are available are adjustment of radar data.

- Non-parametric statistical tests are valid if only if the samples from the two data sets are independent. As radar data is adjusted based on rain gage data the assumption of independence is not valid to carry out these tests.

- Radar-based rainfall estimates are considered as "model estimates" or "forecasts" as these estimates are surrogate measurements of observed rainfall amounts. This assumption also helps in application of forecast verification indices for evaluation of radar-based data.

Comparisons

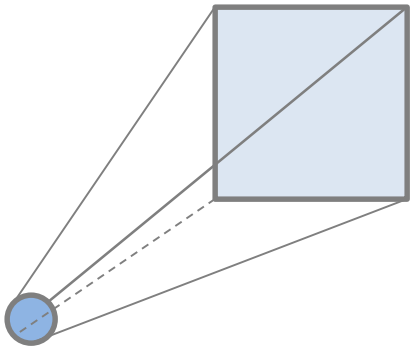
- Assessment of gridded radar data can be carried out using point (rain gage) and gridded (interpolated precipitation data from rain gage network).
- Many studies compare rain gage data (point data) with radar data available at a specific spatial resolution.
- Gridded precipitation data if available, can be used for assessment of radar data.
- Comparison of point (i.e., rain gage) and grid-based (i.e., radar) measurements is the only way to assess radar data when gridded precipitation data based on rain gage observations are not available at a specific temporal or spatial resolution.

(Livneh et al. 2013; Maurer et al. 2002)

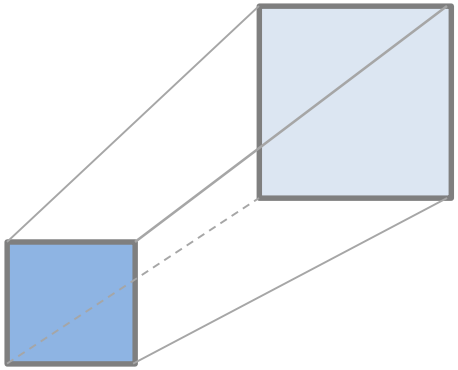
Comparison of Data sets

Spatial Scales

Point-to-Grid



Grid-to-Grid



Indices, Measures, Metrics, Scores and Tests

Visual

Scatter plots
Box Plots
Residual (Error) Plots
Accumulated Value Plots
Kernel Density Estimates
Combined Time Series Plots
Taylor Diagrams

Quantitative

Mean Error (ME)
Mean Absolute Error (MAE)
Multiplicative Bias (MB)
Mean Squared Error (MSE)
Root Mean Squared Error (RMSE)
Root Mean Squared Factor (RMSF)
Linear Error in Probability Space (LEPS)
Correlation Coefficient (CC)
Nash-Sutcliffe Efficiency Coefficient (NSEC)
Heidke Skill Score (HSS)
Contingency Measures:
Concordance, Error Rate, Specificity, Sensitivity.
Bootstrap Sample-based Error Measure

Statistical

Exploratory Data Analysis
Summary Statistics
Quantitle-Quantile Plots
Box Plots
Check for Distributions
Two Sample Kolmogorov-Smirnov Test
Wilcox Sum Test
Ansari-Bradley Test

Contingency Table

2 x 2 Matrix

		Rain Gage Precipitation	
Radar	$\theta_n > 0$	$\theta_r > 0$ [Hits]	$\theta_r = 0$ [False alarms]
Precipitation	$\theta_n = 0$	$\theta_r > 0$ [Misses]	$\theta_r = 0$ [Correct negatives]

$$if(\theta_{n,i} > 0, \theta_{r,i} > 0), then\ C_{11}^i = 1, else\ C_{11}^i = 0 \quad \forall i$$
$$if(\theta_{n,i} > 0, \theta_{r,i} = 0), then\ C_{10}^i = 1, else\ C_{10}^i = 0 \quad \forall i$$
$$if(\theta_{n,i} = 0, \theta_{r,i} > 0), then\ C_{01}^i = 1, else\ C_{01}^i = 0 \quad \forall i$$
$$if(\theta_{n,i} = 0, \theta_{r,i} = 0), then\ C_{00}^i = 1, else\ C_{00}^i = 0 \quad \forall i$$

$$C_{11} = \sum_{i=1}^{N} C_{11}^i$$
$$C_{10} = \sum_{i=1}^{N} C_{10}^i$$
$$C_{01} = \sum_{i=1}^{N} C_{01}^i$$
$$C_{00} = \sum_{i=1}^{N} C_{00}^i$$

$C_{11}, C_{10}, C_{01}, C_{00}$: Counts, N: number of time intervals

$\theta_{n,i}$: radar-based precipitation; $\theta_{r,i}$: rain gage observation

$$Concordance = \frac{C_{11} + C_{00}}{C_{11} + C_{10} + C_{01} + C_{00}}$$

$$Error\ rate = \frac{C_{01} + C_{10}}{C_{11} + C_{10} + C_{01} + C_{00}}$$

$$Sensitivity = \frac{C_{11}}{C_{11} + C_{10}}$$

$$Specificity = \frac{C_{00}}{C_{10} + C_{00}}$$

$$Bias\ Score = \frac{C_{11} + C_{10}}{C_{11} + C_{01}}$$

$$Probability\ of\ detection = \frac{C_{11}}{C_{11} + C_{01}}$$

$$Probability\ of\ false\ detection = \frac{C_{10}}{C_{10} + C_{00}}$$

Dr. Ramesh Teegavarapu, Professor, CEGE, Florida Atlantic University, Director, HRL

Skill Scores

- Heidke Skill Score (HSS)
- Critical Success Index (CSI)
- Peirce Skill Score (PSS)
- Gilbert Skill Score (GSS)
- Odd's Ratio

$$HSS = \frac{2\,(C_{11}C_{00} - C_{10}C_{01})}{(C_{11}+C_{01})(C_{01}+C_{00}) + (C_{11}+C_{10})(C_{01}+C_{00})}$$

$$CSI(threat\ score) = \frac{C_{11}}{C_{11}+C_{01}+C_{10}}$$

$$PSS = \frac{(C_{11}C_{00}) - (C_{10}C_{01})}{(C_{11}+C_{01})(C_{01}+C_{00})}$$

$$GSS = \frac{\boldsymbol{C_{11}-C_{re}}}{(\boldsymbol{C_{11}+C_{10}+C_{10}-C_{re}})} \qquad C_{re} = \frac{(C_{11}+C_{10})(C_{11}+C_{01})}{N}$$

$$Odd's\ Ratio = \frac{(C_{11}\ C_{00})}{(C_{10}\ C_{10})}$$

From forecast verification Literature

Transition Probabilities

Rain or no-rain states for determination of transition probabilities

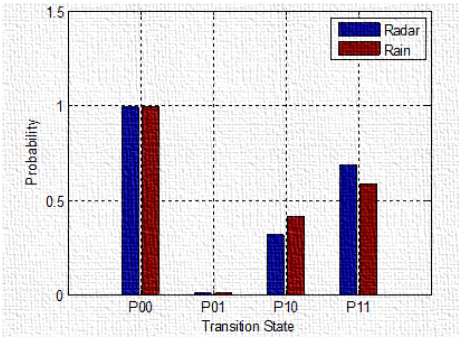
		Time interval (i+1)	
Time	$\theta_i > 0$	$\theta_{i+1} > 0$ [Wet-Wet]	$\theta_{i+1} = 0$ [Wet-Dry]
Interval (i)	$\theta_i = 0$	$\theta_{i+1} > 0$ [Dry-Wet]	$\theta_{i+1} = 0$ [Dry-Dry]

$$P_{11} = P_r(\theta_{i+1} > 0 \mid \theta_i > 0)$$

$$P_{10} = P_r(\theta_{i+1} = 0 \mid \theta_i > 0)$$

$$P_{01} = P_r(\theta_{i+1} > 0 \mid \theta_i = 0)$$

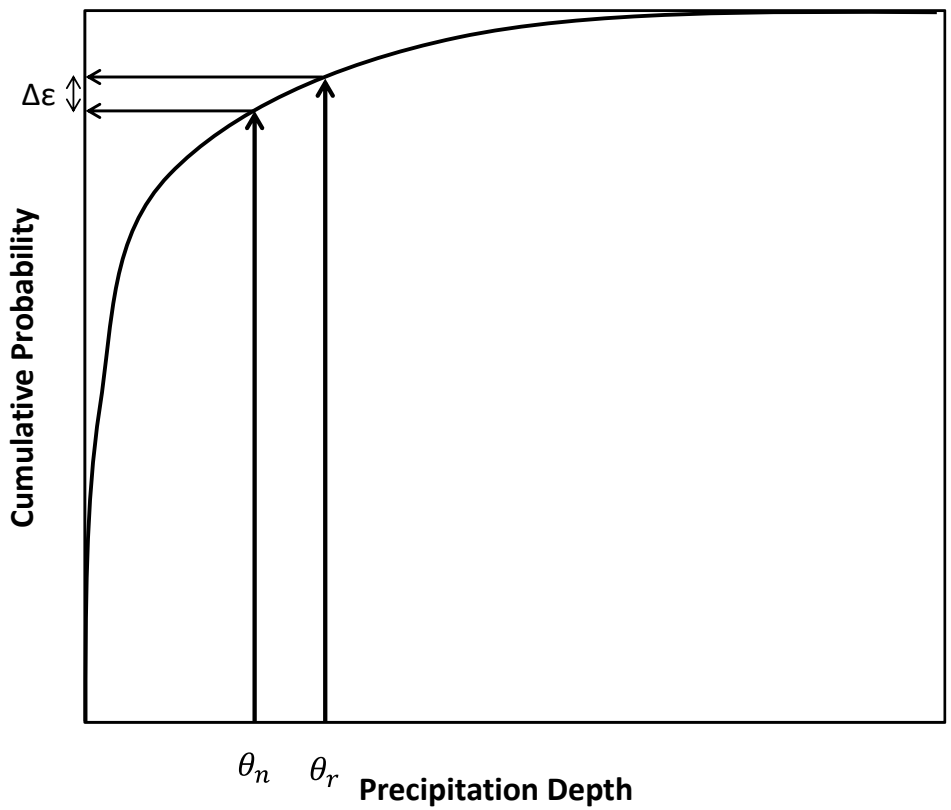
$$P_{00} = P_r(\theta_{i+1} = 0 \mid \theta_i = 0)$$



P_r: Probability, θ_i : Precipitation in time interval i;
θ_{i+1} : Precipitation in time interval $i+1$.

Linear Error in Probability Space

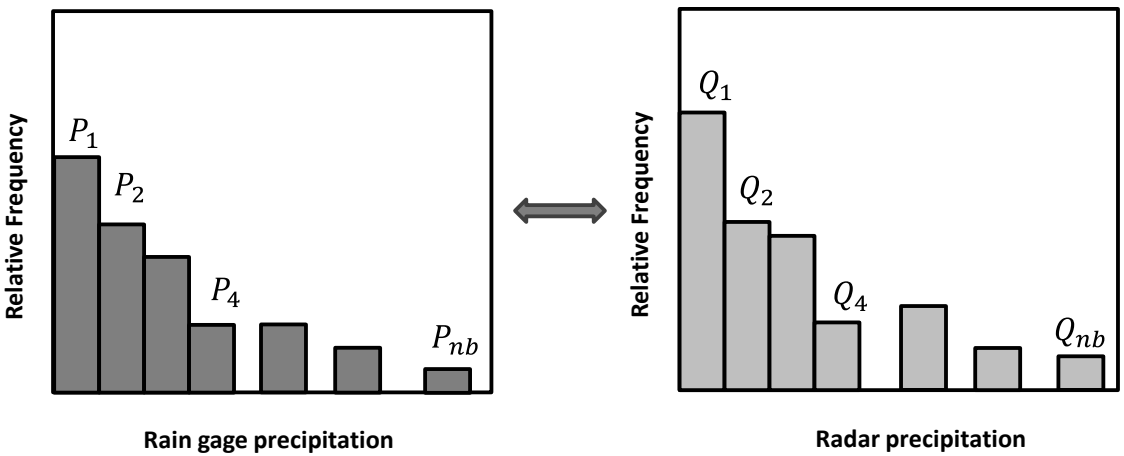
LEPS



$$LEPS = \sum_{i=1}^{N} \Delta\epsilon_i = \frac{1}{N}\sum_{i=1}^{N}\left|F_o\left(\theta_{r,i}\right) - F_o\left(\theta_{n,i}\right)\right|$$

$\theta_{r,i}$: *rain gage precipitation;* $\theta_{n,i}$: $radar - based\ precipitation$

Similarity of Probability Density Functions



P_j: *Relative frequency value for bin j for data set # 1(rain gage)*
Q_j : *Relative frequency value for bin j for data set 2 (radar)*
nb : number of bins
d : distance

$$d_{euc} = \sqrt{\sum_{j=1}^{nb} |P_j - Q_j|^2}$$

$$d_{cbl} = \sum_{j=1}^{nb} |P_j - Q_j|$$

$$d_{mnk} = \sqrt[pn]{\sum_{j=1}^{nb} |P_j - Q_j|^{pn}}$$

$$d_{che} = \max_j |P_j - Q_j|$$

$$d_{sor} = \frac{\sum_{j=1}^{nb} |P_j - Q_j|}{\sum_{j=1}^{nb} (P_j + Q_j)}$$

$$d_{gow} = (1/nb) \sum_{j=1}^{nb} |P_j - Q_j|$$

$$d_{soe} = \frac{\sum_{j=1}^{nb} |P_j - Q_j|}{\sum_{j=1}^{nb} \max_j (P_j, Q_j)}$$

$$d_{lor} = \sum_{j=1}^{nb} \ln(1 + |P_j - Q_j|)$$

$$d_{kul} = \frac{\sum_{j=1}^{nb} |P_j - Q_j|}{\sum_{j=1}^{nb} \min_j (P_j, Q_j)}$$

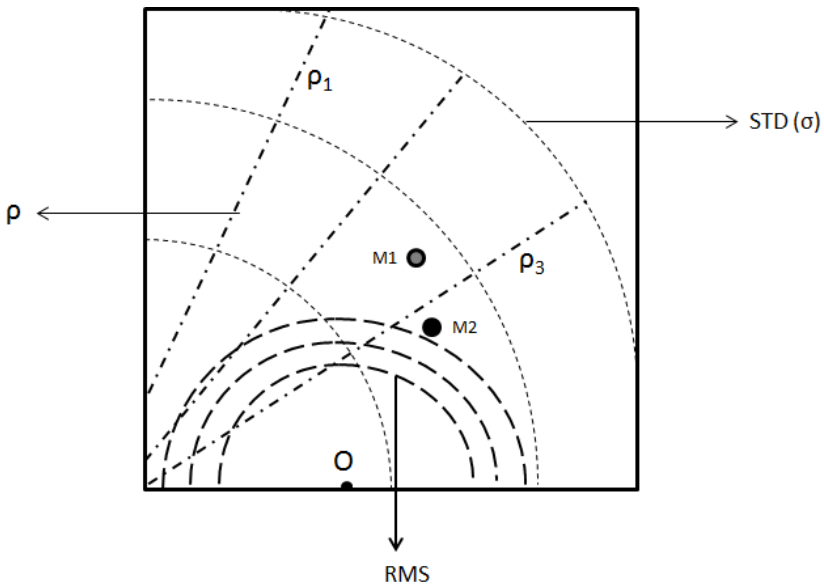
$$d_{can} = \sum_{j=1}^{nb} \frac{|P_j - Q_j|}{(P_j + Q_j)}$$

$$d_{int} = (1/2) \sum_{j=1}^{nb} |P_j - Q_j|$$

$$d_{whe} = \sum_{j=1}^{nb} \frac{|P_j - Q_j|}{\max(P_j, Q_j)}$$

- Euclidean (d_{euc})
- City Block (d_{cbl})
- Minkowski (d_{mnk})
- Chebyshev (d_{che})
- Sorensen (d_{sor})
- Gower (d_{gow})
- Soergel (d_{soe})
- Lorentzian (d_{lor})
- Kulczynski (d_{kul})
- Canberra (d_{can})
- Intersection (d_{int})
- Wave Hedges (d_{whe})

Taylor Diagram: A Multi Data Set Assessment Tool



$$\varepsilon = \frac{1}{N}\sum_{i=1}^{N}[(\theta_{n,i} - \bar{\theta}_n) - (\theta_{r,i} - \bar{\theta}_r)]^2$$

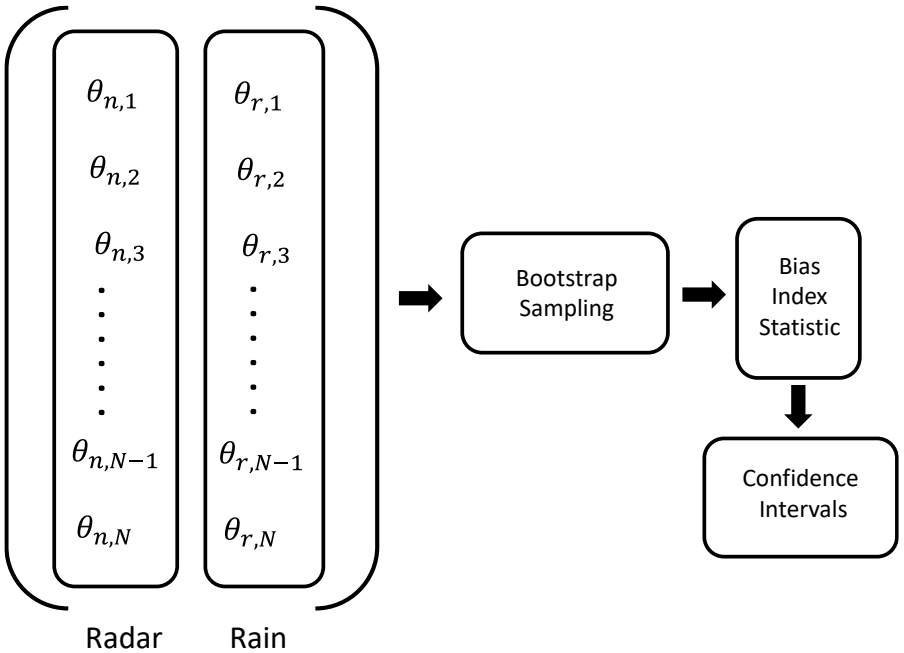
$$\varepsilon^2 = \sigma_n^2 + \sigma_r^2 - 2\sigma_n\sigma_r\rho$$

O: Observed
M1, M2 : Models

- Taylor diagram is used to summarize three different performance measures in one diagram based on the radar and rain gage data.
- Taylor diagram combines statistical measures such as standard deviations, centered root mean squared deviations or errors (RMS) and correlations based on multi-model estimations in one graph

Taylor, 2001

Confidence Intervals of Error or Performance Measures



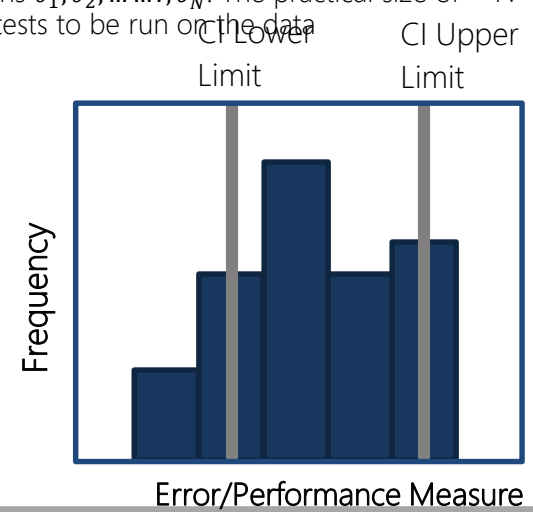
Samples values : $y_1, y_2, \dots\dots, y_n$ are the outcomes of independent and identically distributed ($\widetilde{iid}$) random variables. If the CDF of $Y_1, Y_2, \dots\dots, Y_n$ is *F*. Then estimate of *F* denoted by $\hat{F}$ is obtained using data $y_1, y_2, \dots\dots, y_n$.

1. Bootstrap (re) sample $y_1^*, y_2^*, \dots\dots, y_n^* \; \widetilde{iid} \; \hat{F}$ are obtained.

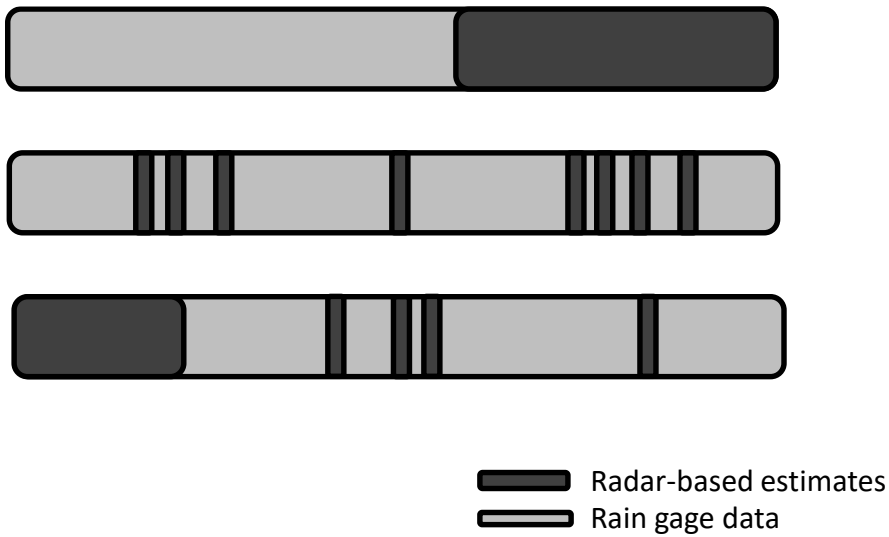
2. $\hat{F}$, an estimator of *F* is obtained non-parametrically using empirical distribution function (EDF) of the original data, i.e., by placing a probability of '*1/n*' at each data value from sample $y_1, y_2, \dots\dots, y_n$.

3. Sample mean statistic $\widehat{\theta^*}$ is computed from bootstrap sample $y_1^*, y_2^*, \dots\dots, y_n^*$.

4. The above steps are repeated '*N* times, to obtain *N* sample means $\widehat{\theta_1^*}, \widehat{\theta_2^*}, \dots\dots, \widehat{\theta_N^*}$. The practical size of '*N* depends on the tests to be run on the data



Replacement Tests

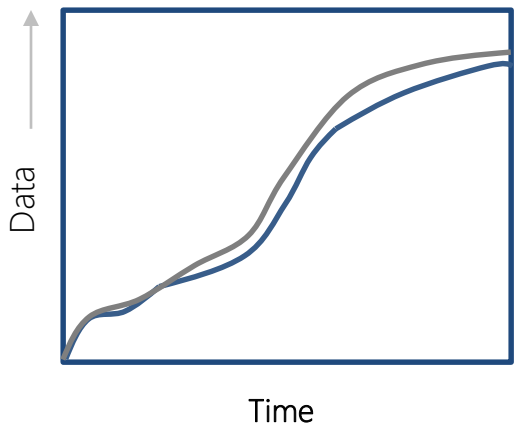


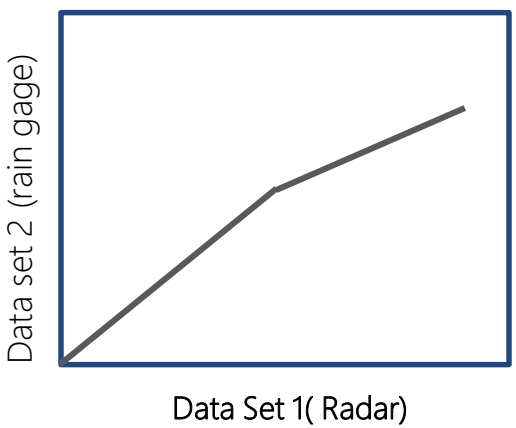
- Systematic
- Random
- Systematic and random

Replacement tests are proposed to check the validity of radar data when missing rain gage values are replaced in three possible configurations.
These configurations are : (1) systematic, (2) random and (3) systematic and random replacement. Once the replacement of data is completed, the data sets can be tested for homogeneity

Cumulative Plots

Accumulation over time





Comparison of cumulative plots of radar-based precipitation and rain gage observations can provide assessment of time consistent accumulations of these two data sets.

A cumulative plot is a variant of a double mass curve.

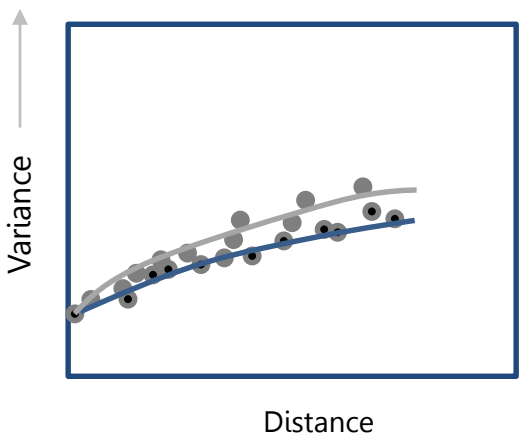
The rain gage and radar data sets when plotted together show how cumulative totals vary over time.

$$\theta_{r,1}^{ac} = \theta_{r,1}$$

$$\theta_{r,j+1}^{ac} = \theta_{r,j}^{ac} + \theta_{r,i} \quad \forall i, \qquad i = 2, \dots N \;\; \& \;\; j = 1 \dots N-1$$

Variance with Distance

Variogram Cloud



A variogram cloud is a plot of distances and variance or semi-variance values. Distance in this context refers to Euclidean distance between any two observation sites.

$$\gamma(d_{a,b}) = \frac{1}{N}\sum_{i=1}^{N}(\theta_{a,i} - \theta_{b,i})^2$$

The variable $\gamma(d_{a,b})$ is the average variance based on observations at site a and site b.

$\theta_{a,i}$: *radar estimates or rain gage observations*

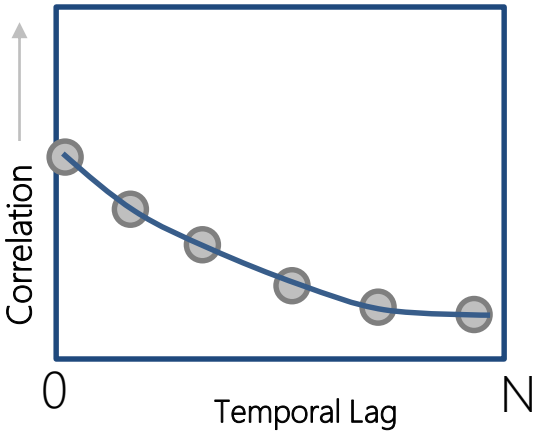
Summary Statistics / Exploratory Data Analysis (EDA)

- Summary Statistics
  - Measures of Central Tendency
  - Measures of dispersion
    - Interquartile Range (IQR)
    - Range (R)
    - Mean Absolute Deviation (MAD)
    - Measures of Shape
    - Box Plots

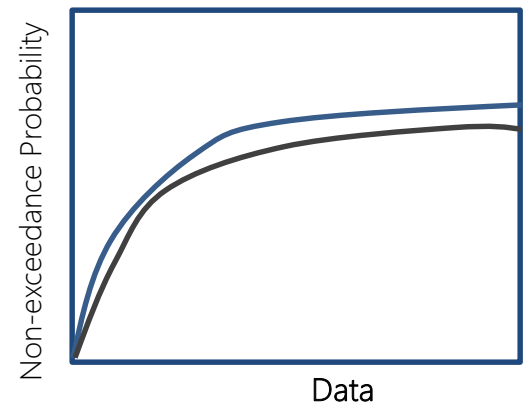
Mean . Median
Skewness . kurtosis
Quantile-Quantile Plots
Cumulative Distribution Function Plots

- Serial Autocorrelation

$$\rho_t = \frac{\sum_{i=1}^{N-t}(\theta_i - \bar{\theta}_{(1)})(\theta_{i+t} - \bar{\theta}_{(2)})}{\sqrt{\sum_{i=1}^{N-t}(\theta_i - \bar{\theta}_{(1)})^2}\sqrt{\sum_{i=2}^{N}(\theta_i - \bar{\theta}_{(2)})^2}}$$



Statistical Hypothesis Tests



Tests for Distributions of Data

- Kolmogorov-Smirnov (KS)Two Sample Test
- H_o : Two data sets are from the same continuous distribution.
- H_a : Two data sets are from different continuous distributions.
- Statistical significance level: 5%

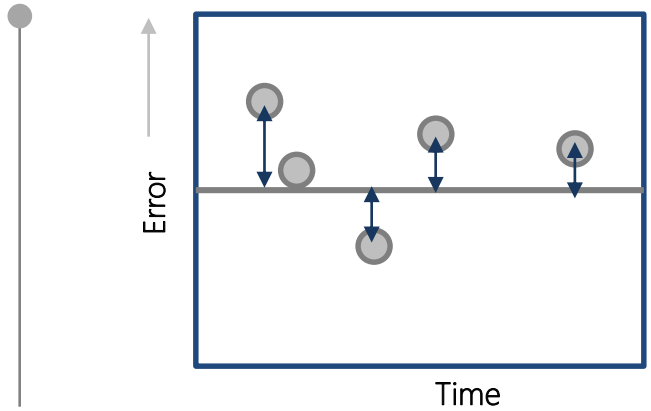
- Ansari-Bradley Test
- Data come from the same distribution (null hypothesis: H_o), against the alternative hypothesis: H_a that they come from distributions that have the same median and shape but different dispersions (e.g. variances). Significance level : 5%

- Wilcoxon Rank Sum Test
- The rank sum test can be used to evaluate the null hypothesis that data sets are from identical continuous distributions with equal medians, against the alternative that they do not have equal medians. Significance level : 5%.

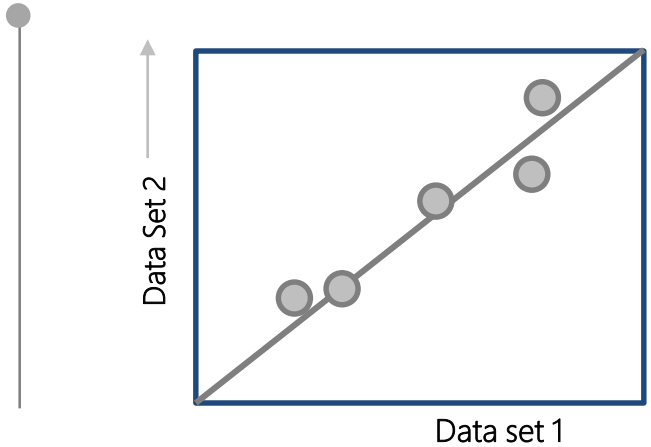
Quantitative and Association Measures

- Quantitative Error Measures
  - Mean Error (ME)
  - Mean Absolute Error (MAE)
  - Mean Squared Error (MSE)
  - Root Mean Squared Error (RMSE)
  - Root Mean Squared Factor (RMSF)
  - Nash-Sutcliffe Efficiency Coefficient (NSEC)

$$NSEC = 1 - \frac{\sum_{i=1}^{N}(\theta_{r,i} - \theta_{n,i})^2}{\sum_{i=1}^{N}(\theta_{r,i} - \bar{\theta}_r)^2}$$



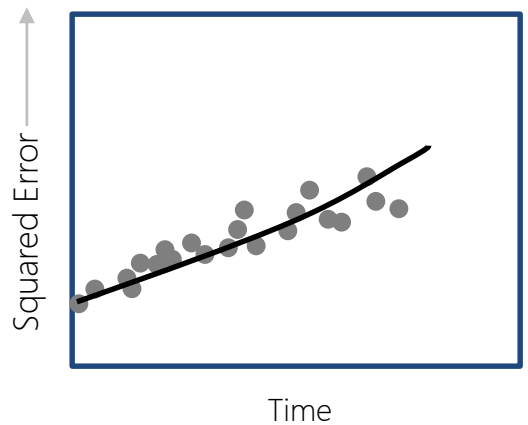
- Association Measures
  - Pearson Correlation Coefficient (ρ)
    - Linear association measure
  - Spearman Rank Correlation (ρ_{rank})
    - Based on ranked observations
  - Kendall's Tau (τ)
    - Concordant and discordant pairs

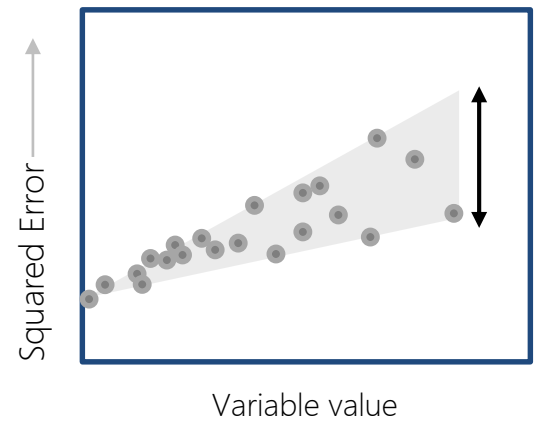


Residuals

- Evaluation of Residuals
- $e_i = \theta_{n,i} - \theta_{r,i} \qquad \forall i$

$$d^o = \frac{\sum_{i=2}^{N}(e_i - e_{i-1})^2}{\sum_{i=1}^{N} e_i^2}$$





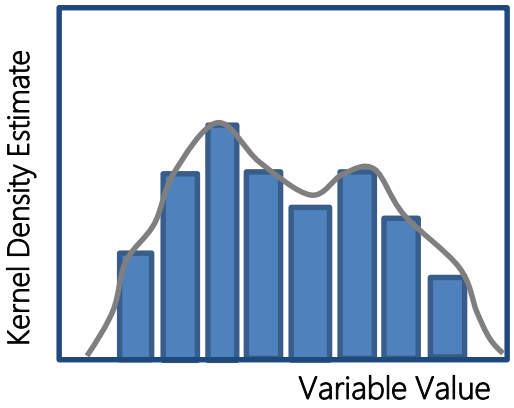
Test or Plot	Evaluation Type	Diagnostics/Inference
Time series plot of residuals	Time dependency	Random variation and no structured pattern in variation of residuals with time.
Autocorrelation Plot (autocorrelogram)	Check for serial correlation at different temporal lags	Low or negligible correlation at all lags
Probability Plot (Normal)	Normality of residuals	Linear plot of residuals
Histogram	Visual check for Normality	Gaussian distribution
Durbin-Watson Test (d^o)	Test for serial correlation	Durbin-Watson Statistic (d)

Dr. Ramesh Teegavarapu, Professor, CEGE, Florida Atlantic University, Director, HRL

Kernel Density Estimates

KDE

Scaled frequencies and density estimates are superimposed for comparison and optimal bandwidth is selected for the Gaussian kernel



$$\hat{f}_h(x) = \frac{1}{nh}\sum_{ip=1}^{n} K\left(\frac{x-x_{ip}}{h}\right)$$

Kernel Density Estimate

$$K\left(\frac{x-x_{ip}}{h}\right) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x-x_{ip}}{h}\right)^2\right) \quad \forall ip$$

Gaussian Kernel

Utility of Assessment Measures and Indices, Scores

Measures, Indices and Evaluation Tests and Methods	Utility for Operations [O], Modeling [M] and Short-term Planning [S]
Contingency Measures	Operations/short-term Decision Making and Evaluation [O, S]
Skill Scores	Short term Planning and Improvement [S]
Performance Measures (Error Measures)	Modeling, long-term assessment [M]
Transitional Probabilities	Short-term Evaluation & Modeling [S, M]
Association Measures and Plots	Operations, Short-term evaluation, modeling [O, S , M]
Taylor Diagram	Modeling and Long-term Assessment [M]
Summary Statistical Measures	Short term planning and Improvement and Modeling [S, M]
Data Distribution Analysis	Modeling [M]
Serial Autocorrelation	Modeling and long-term evaluation [M]
Quantile Plots	Modeling [M]
Variograms	Modeling [M]

Utility of Assessment Measures and Indices, Scores

Measures, Indices and Evaluation Tests and Methods	Utility for Operations [O], Modeling [M] and Short-term Planning [S]
Infilling or Replacement Tests	Operations, Short-term Planning [O, S]
Evaluation of Residuals	Long-term Assessment, modeling [M]
Cumulative Plots	Modeling [M]
Kernel Density Estimates	Modeling [M]
Statistical Hypothesis tests	Short-term evaluation, modeling [S, M]

Application of the Framework

Vieux and Associates Radar Product

- The Vieux and Associates Inc. (VAI) provide its rainfall database to the five WMDs and to FSU.
- The spatial resolution of VAI product is a 2×2 km Cartesian grid and the data are provided at 15 min. intervals.
- The VAI Corporation provides a near real-time product as well as an end-of-the-month product which has undergone further QC. VAI has provided the SFWMD with NEXRAD based precipitation estimates from years 2007 to present data.
- Radar data provide by VAI for the period of the record October 2007 – October 2011 is used.

Rain gage and Radar Datasets

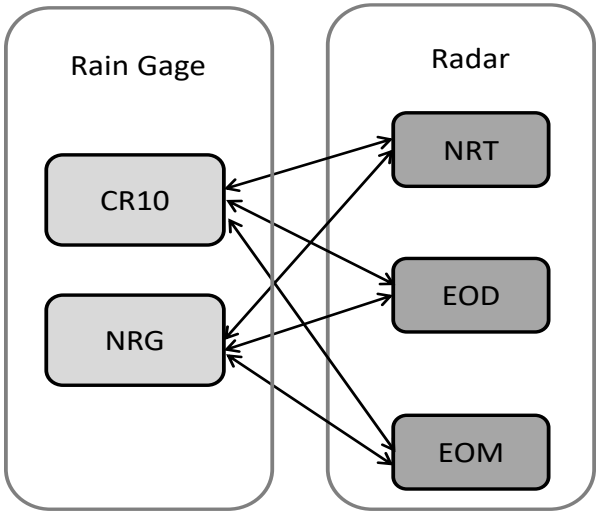
Data Source	Temporal Resolution	Missing data	Period of Record	Comments
Rain Gage Data				
CR10	15 minute data	Yes	Oct 2007 – Sept 2011	Missing Data is flagged (Tag: M)
NRG	15 minute data	Yes (May 2008)	Oct 2007- Sept, 2011	Missing Data is augmented and also filled with radar data (Tag: R)
Radar Data (NEXRAD)				
NRT	15 minute data	Yes	Oct 2007- Sept 2011	Near Real-Time Data
EOD	15 minute data	No	Oct 2007 – Sept 2011	Adjusted End of the day (EOD)
EOM	15 minute data	No	Oct 2007 – Sept 2011	Adjusted End of the Month (EOM)

Radar Data from Vieux and Associates Inc.

Radar and Rain Gage Data

- Three data sets derived from radar-based rainfall estimates and two data sets from rain gage network are used for the analysis of bias in the current study.
- The rain gage data sets have missing data and few anomalous values. The values exceeding a specific threshold are either limited to a maximum threshold value or not used for analysis.
- Rain gage augmented with radar data sets are also used for analysis in the current study. The time intervals in which the rain gage data is augmented are tagged. The percentage of augmented data ranged from 1% to 100%.
- A total of 189 rain gage sites were selected from an initial set consisting of 202 rain gages were selected for analysis. Rain gages with data augmented are also included in the analysis
- Data is split into four water years and analysis is carried out.

Data Sets : Combinations for Analysis

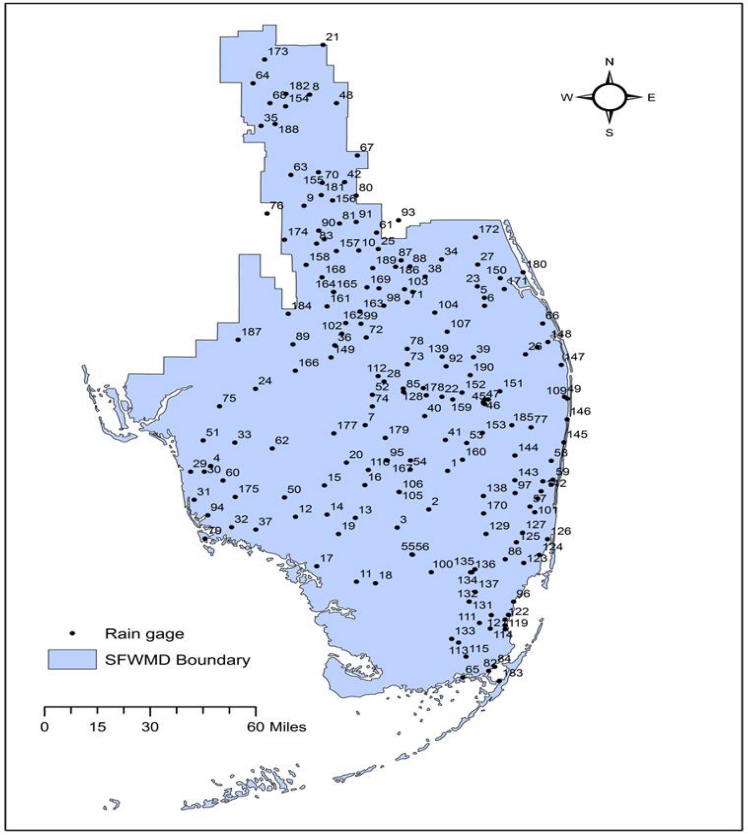


Rain gage data sets (CR10, NRG) and radar data sets (NRT, EOD and EOM) are analyzed for this study.

CR10: Rain gage (Breakpoint data)
NRG : Rain gage data augmented by radar data

NRT: Near Real-Time
EOD: End of the Day
EOM: End of the Month

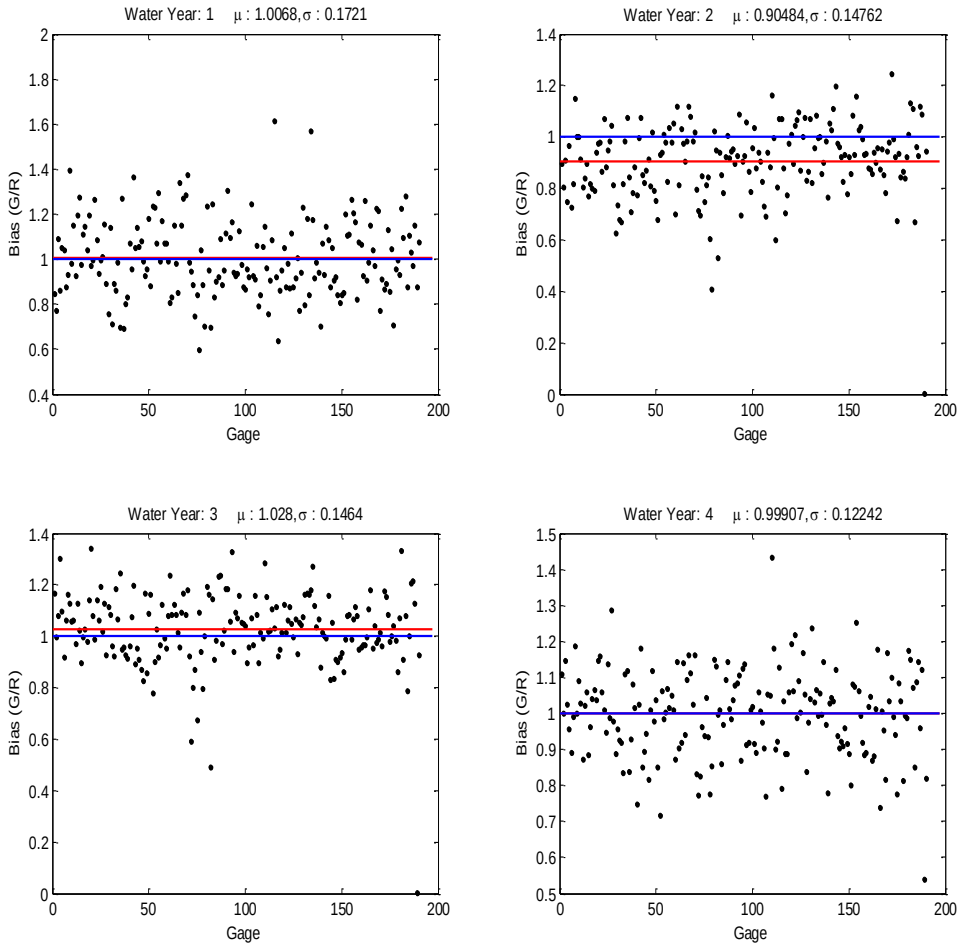
Rain Gages used for Evaluation



189 gages

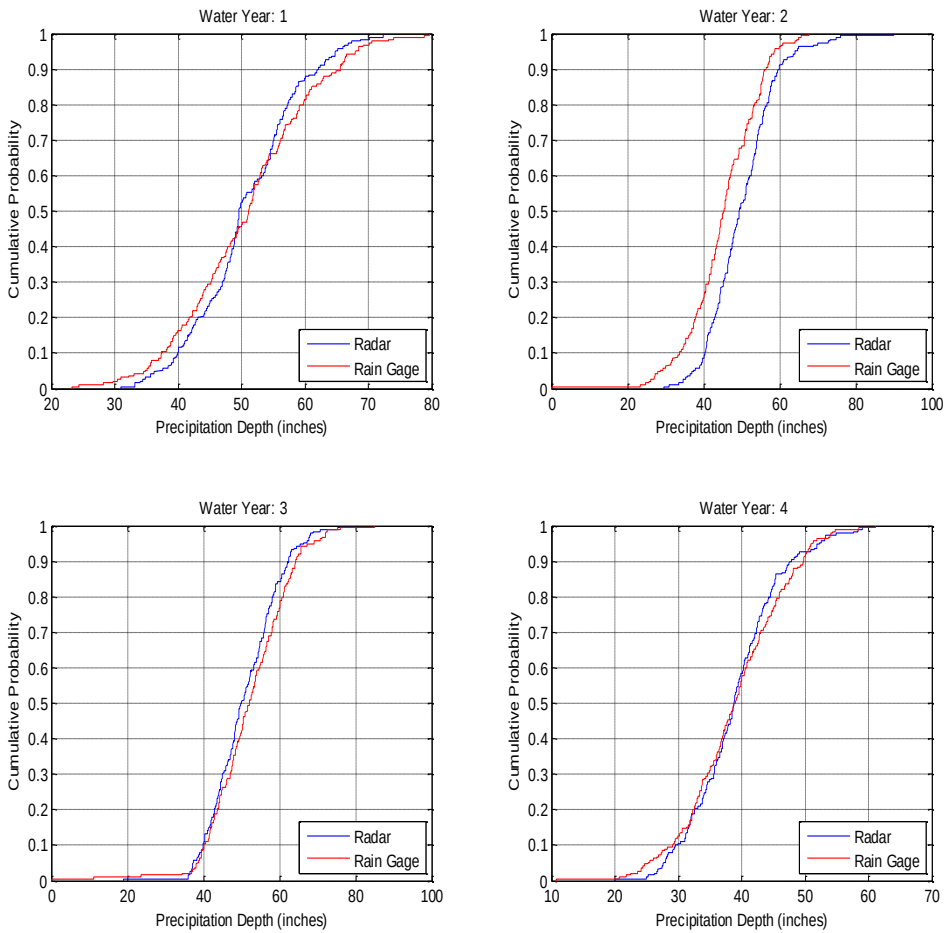
Bias

Gage/Radar Values

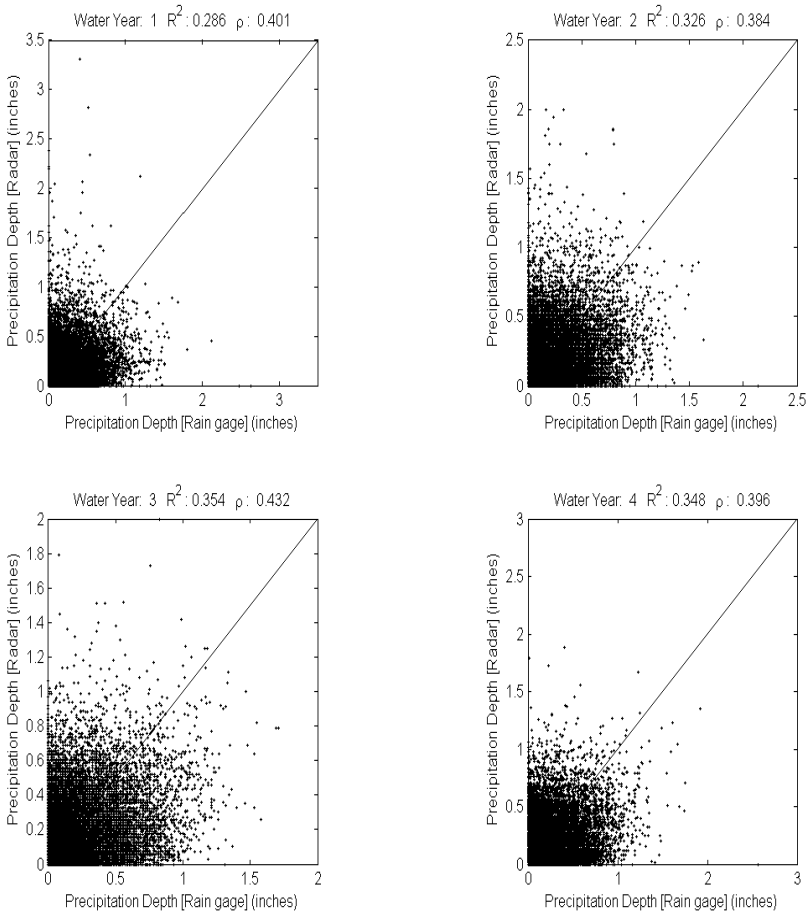


The average bias is improving when annual totals are considered.

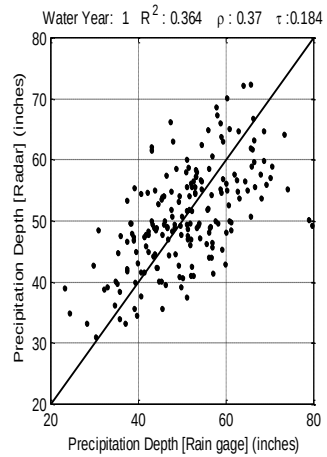
Annual Values

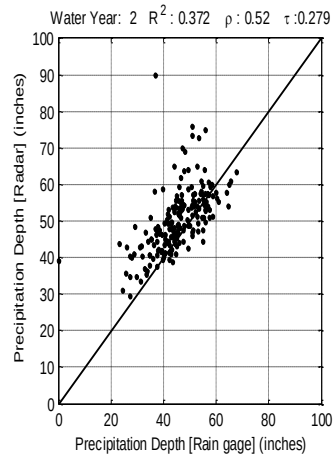


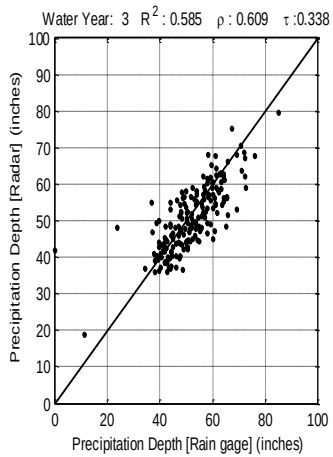
Cumulative densities of annual totals shows closeness of two data sets

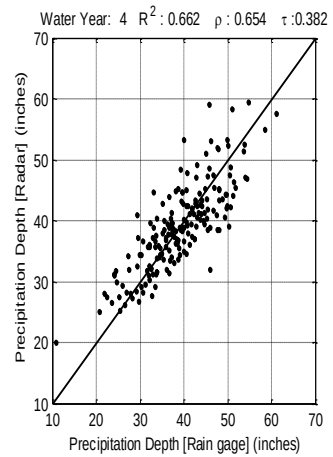


Scatter plots show good agreement between radar and rain gage data

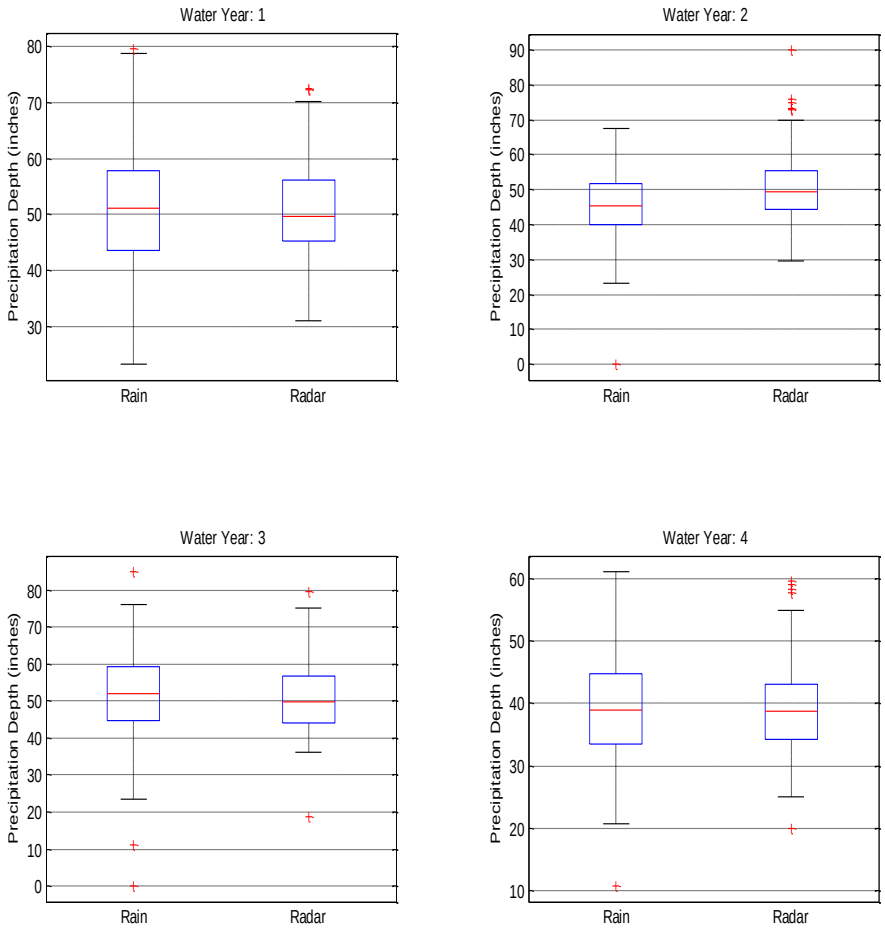






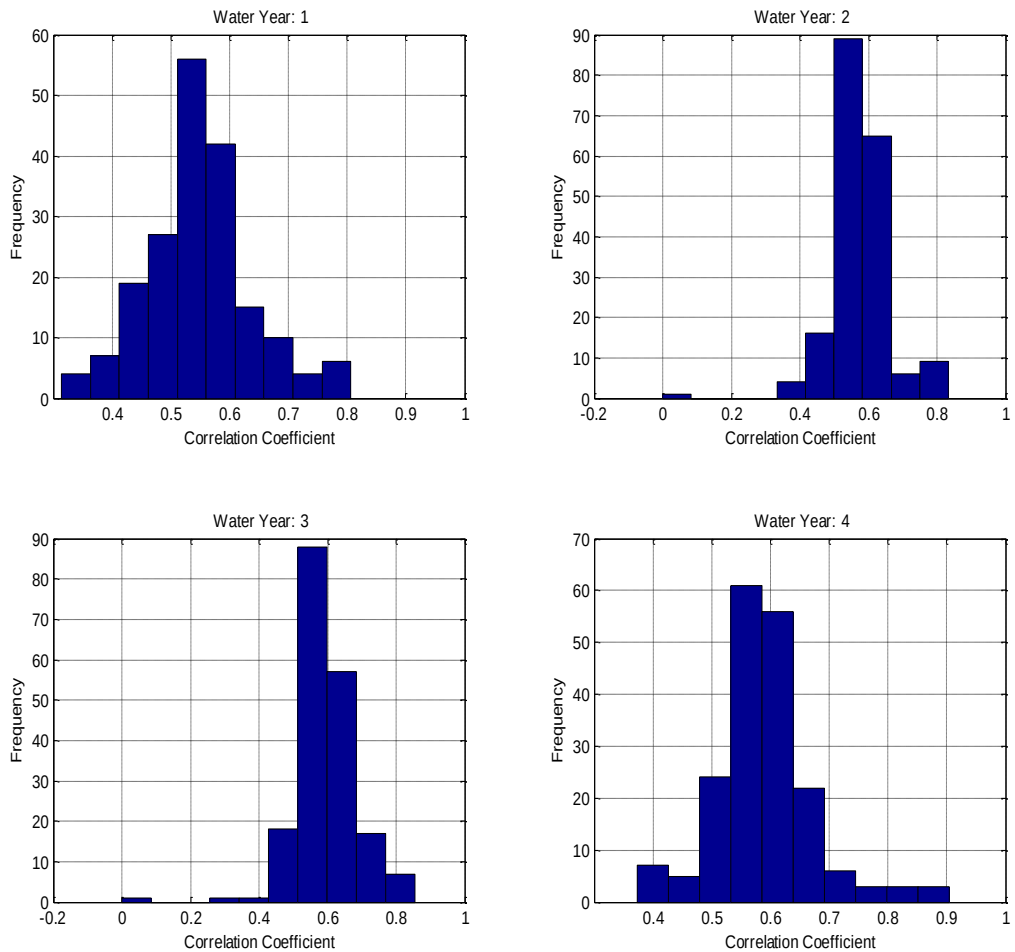


Scatter plots show good agreement between radar and rain gage data and improvement over time



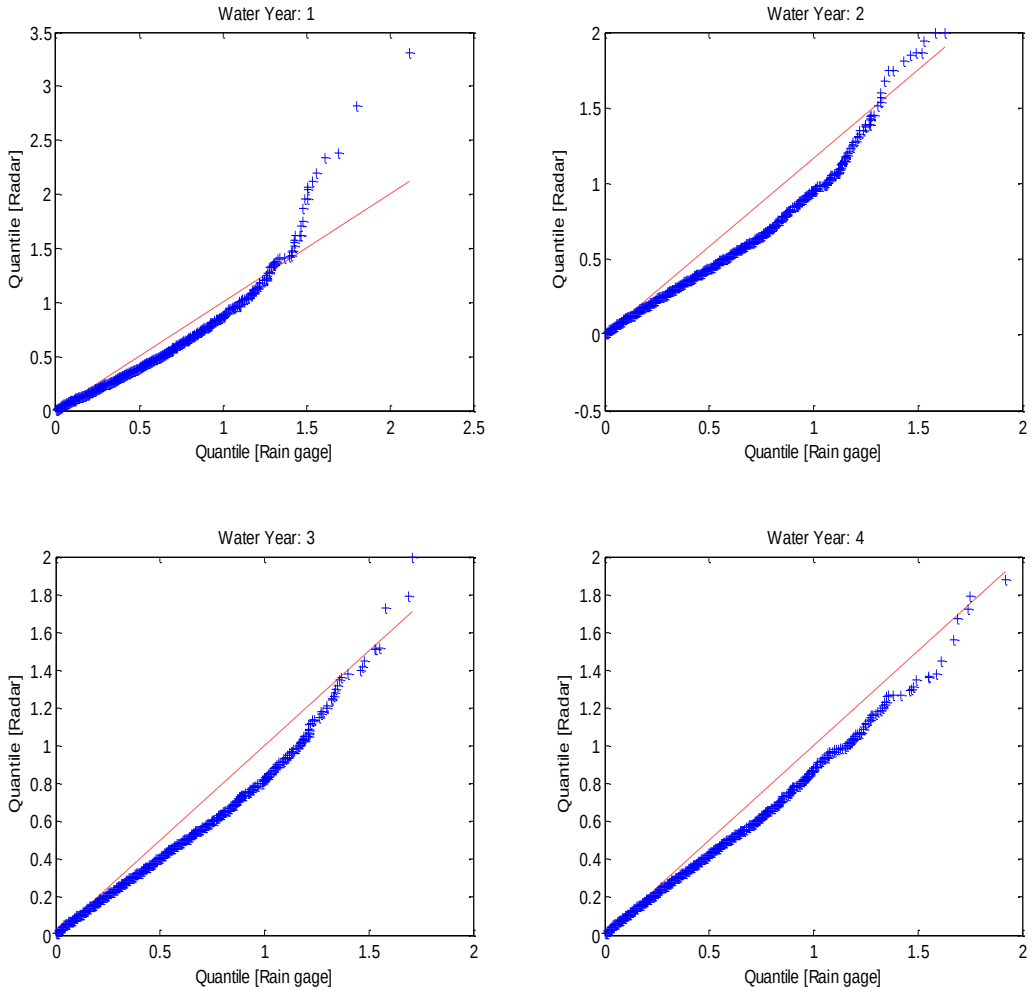
Good agreement between radar and rain gage data

Distributions of Correlations



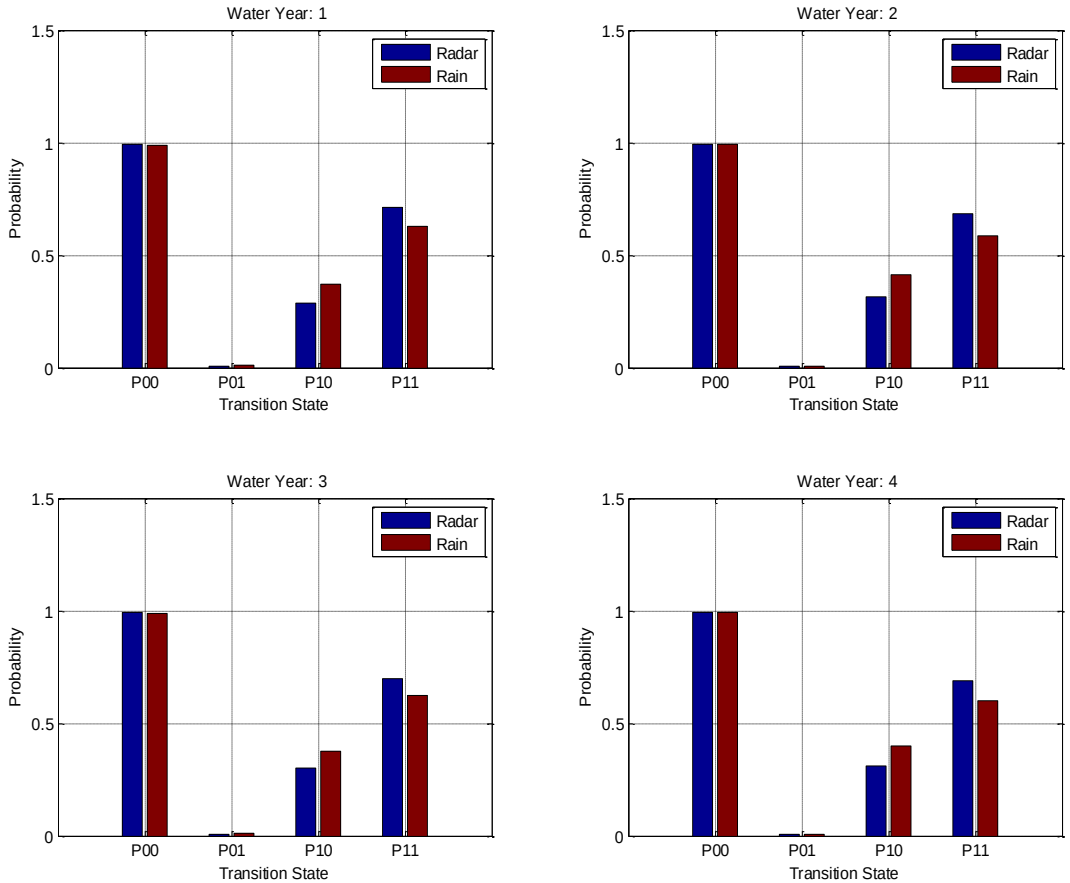
Distributions of the correlation coefficients based on radar-gage data sets are negatively skewed and show higher mean values over time

Quantile-Quantile Plots



Q-Q plots show good agreement between radar and rain gage data and improvement over time

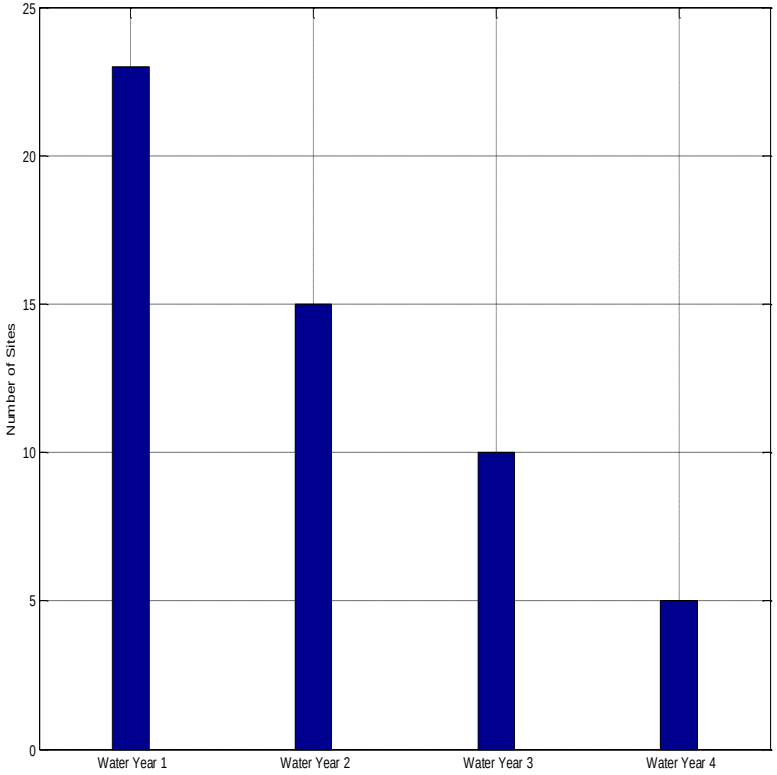
Transitional Probabilities



Underestimation of wet to dry transition.

Overestimation of wet to wet transition.

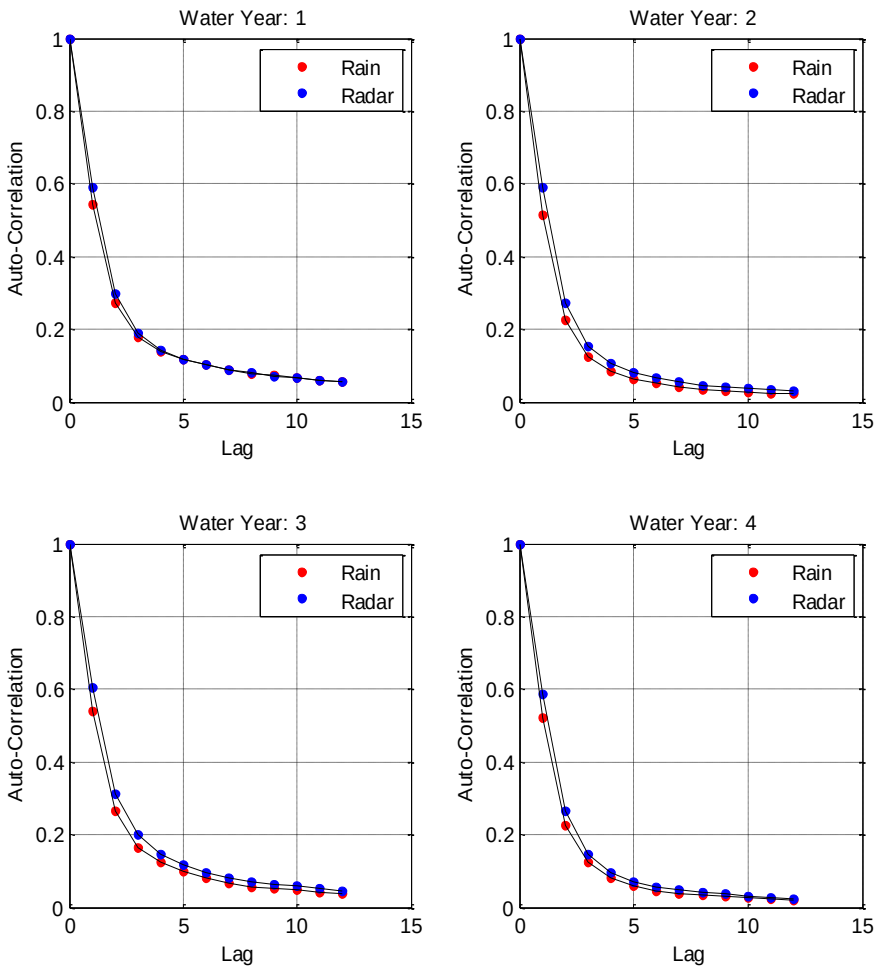
Statistical hypothesis tests



Number of sites failing "Distribution tests" reduces over time

Autocorrelation

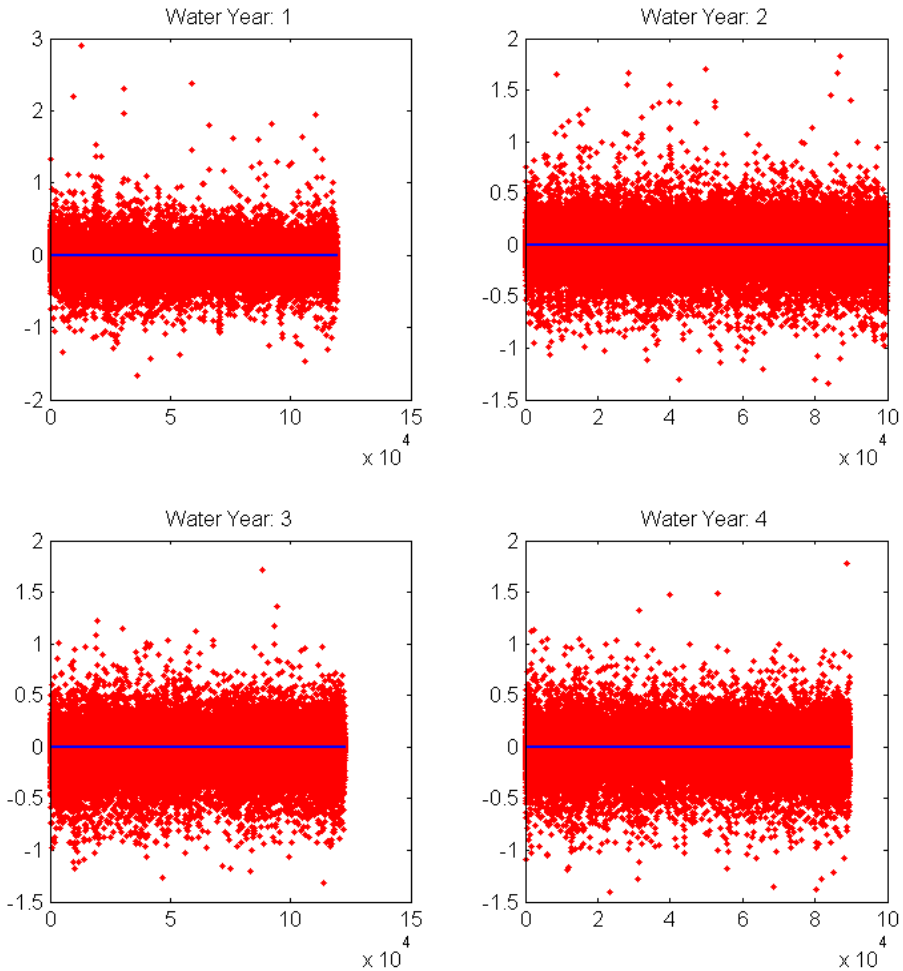
Persistence



Consistent over estimation of autocorrelation at different lags

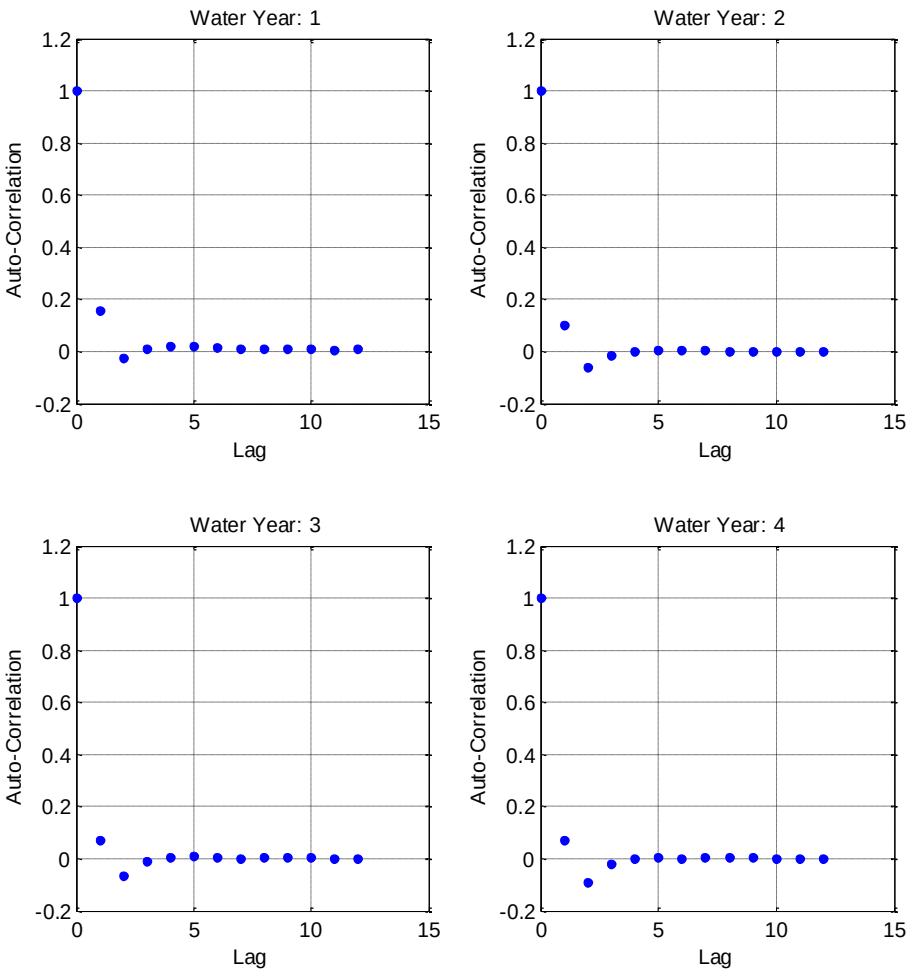
Radar data seems to be preserve persistence over time (for longer lagged periods.)

Evaluation of residuals



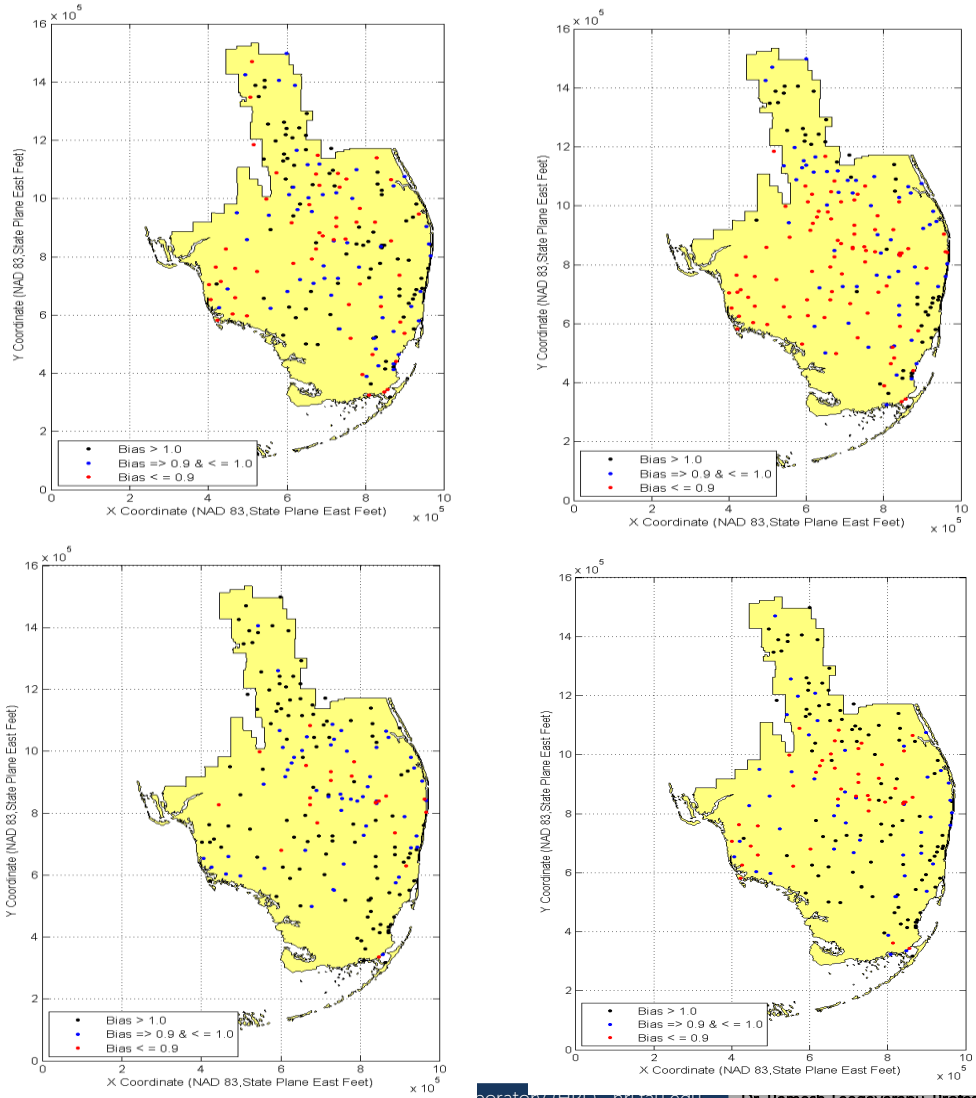
The residuals are random and no sign of time dependency or heteroscadasticity for different years under consideration.

Autocorrelation of residuals



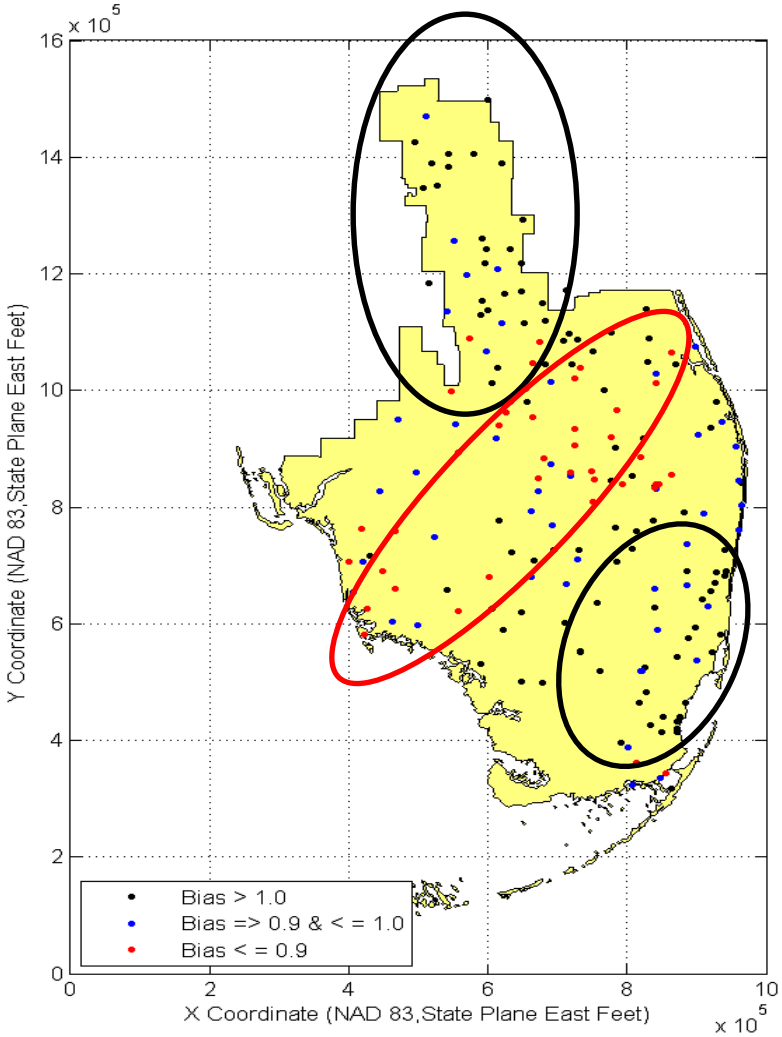
Very low autocorrelation at different lags for errors suggesting no time dependency

Spatial Variability of Bias



Few problem areas are identified where spatial patterns of biases are consistent

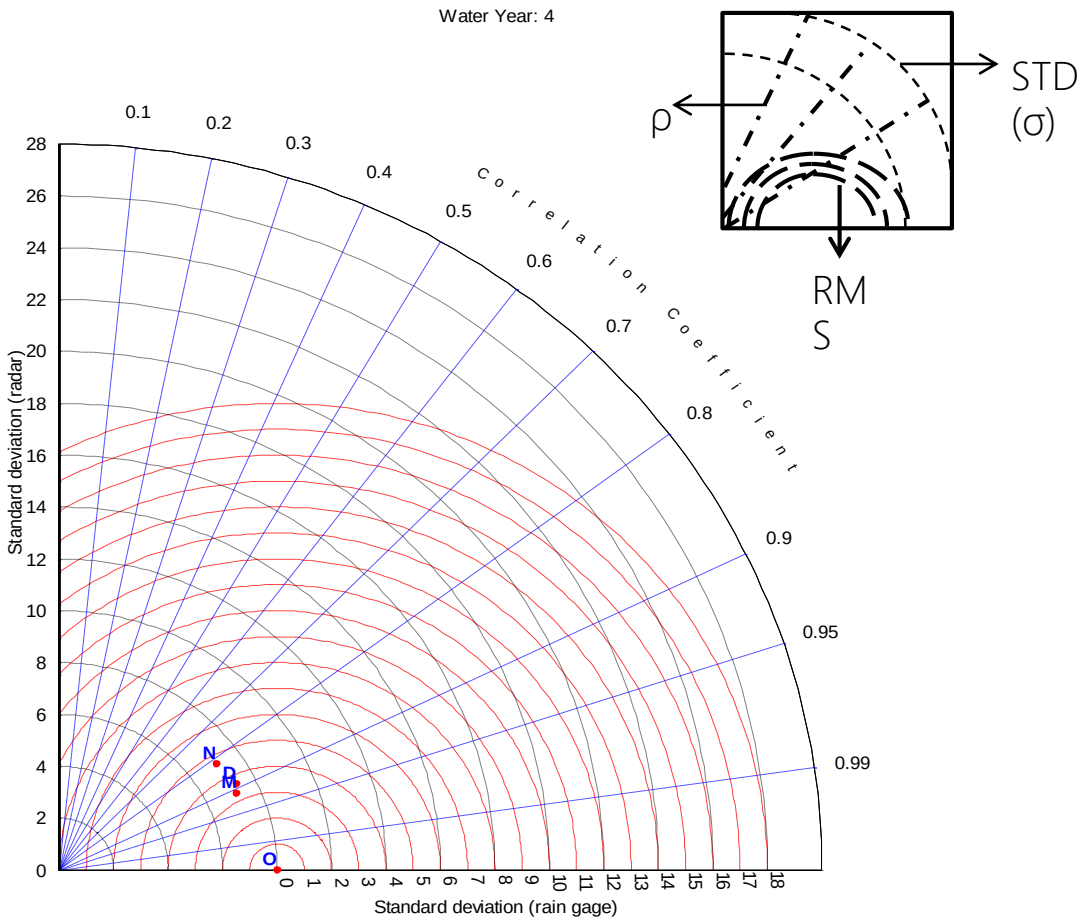
Issues



Underestimation of rainfall by radar is evident in the Northern parts and Southeastern parts of the District

Overestimation of rainfall by radar is evident in few regions of the District.

Taylor Diagram

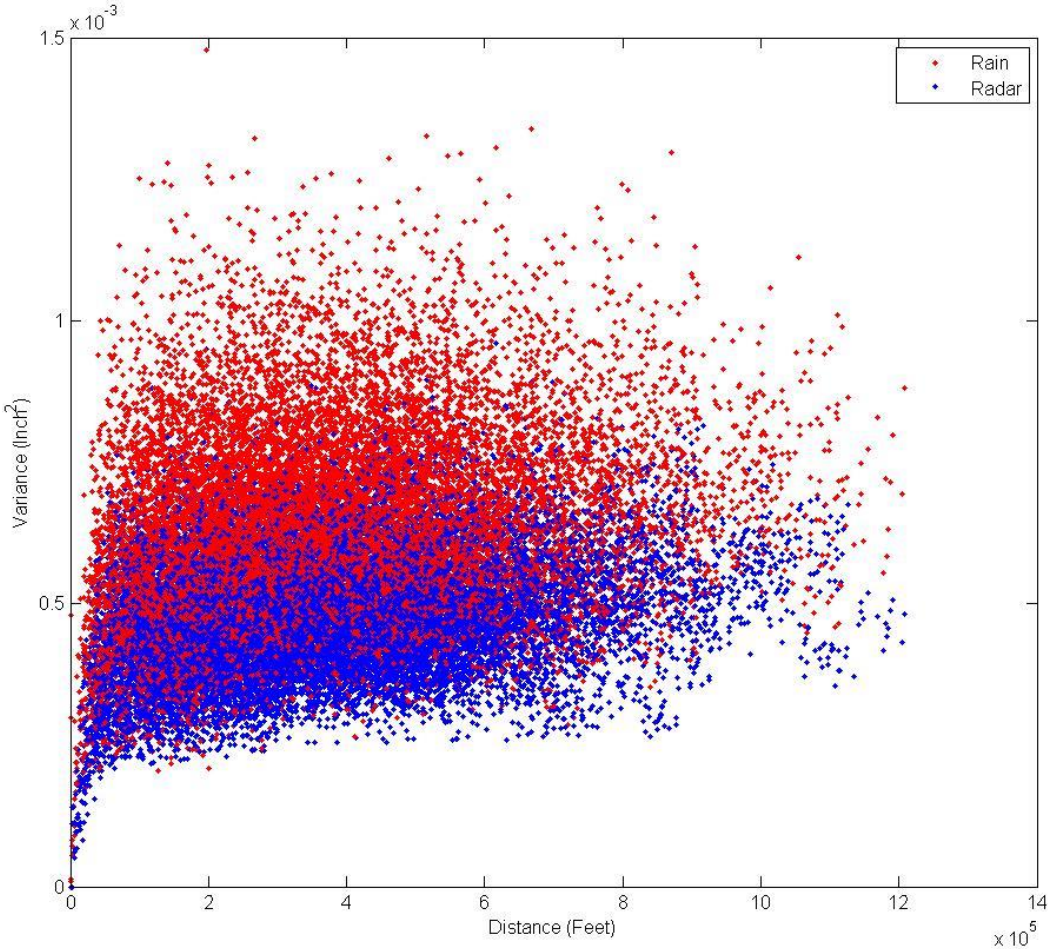


Taylor Diagram suggests improvements in the data from NRT to EOD to EOM

N: Near Real Time D: End of the day M: End of the Month O: Observed

Spatial variability of variance

Variogram cloud



Underestimation of spatial dependence and variability of radar data compared to rain gage data set.

CR10 - NRT

Performance Measure	Year 1	Year 2	Year 3	Year 4
Mean Error (ME)	-1.4E-06	0.000143	-3.7E-05	2.18E-06
Mean Absolute Error (MABE)	0.001572	0.001439	0.001459	0.001153
Multiplicative Bias (MB)	0.999026	1.110222	0.97553	1.001932
Mean Squared Error (MSQE)	0.000306	0.000293	0.00028	0.00024
Root Mean Squared Error (RMSE)	0.017488	0.017105	0.01672	0.015508
Root Mean Squared Factor (RMSF)	3.430207	3.564883	3.240621	3.403268
Linear Error in Probability Space (LEPS)	0.001573	0.001292	0.001478	0.001087
Correlation Coefficient (CORR)	0.535448	0.571098	0.595425	0.590233
Nash-Sutcliffe Efficiency Coefficient (NSEC)	0.169788	0.176609	0.282813	0.260504
Heidke Skill Score (HSS)	0.627252	0.613199	0.651726	0.625337
Contingency Measure: Concordance (CMC)	0.979108	0.981838	0.981181	0.985235
Contingency Measure: Error Rate (CMER)	0.020892	0.018162	0.018819	0.014765
Contingency Measure: Specificity (CMSP)	0.987822	0.988922	0.989443	0.991787
Contingency Measure: Sensitivity (CMSE)	0.670451	0.671277	0.682613	0.654693
Pierce Skill Score (PSS)	0.599183	0.572724	0.632659	0.605587
Gilbert's Skill Score (GSS)	0.456932	0.442168	0.483378	0.454902
Odd's Ratio (OR)	165.0287	182.299	201.5682	228.9617
Bias Score (BS)	1.101785	1.156875	1.064161	1.068989
Probability of Detection (POD)	0.670451	0.671277	0.682613	0.654693
False Alarm Ratio (FAR)	0.391486	0.41975	0.358544	0.387559
Critical Success Index (CSI)	0.46841	0.451857	0.494093	0.462911
Boot Strap CI (L)	0.054714	0.062625	0.053676	0.060738
Boot Strap CI (U)	0.056091	0.063829	0.054876	0.061933

CR10 - EOD

Performance Measure	Year 1	Year 2	Year 3	Year 4
Mean Error (ME)	1.99E-05	9.31E-05	5.7E-05	5.91E-05
Mean Absolute Error (MABE)	0.00157	0.001383	0.001449	0.001132
Multiplicative Bias (MB)	1.013418	1.071559	1.03801	1.052694
Mean Squared Error (MSQE)	0.000302	0.000261	0.000264	0.000222
Root Mean Squared Error (RMSE)	0.017375	0.01615	0.01625	0.014884
Root Mean Squared Factor (RMSF)	3.44396	3.451457	3.234682	3.385519
Linear Error in Probability Space (LEPS)	0.001547	0.001276	0.001416	0.00104
Correlation Coefficient (CORR)	0.541896	0.598819	0.625884	0.623374
Nash-Sutcliffe Efficiency Coefficient (NSEC)	0.174866	0.265958	0.322517	0.313891
Heidke Skill Score (HSS)	0.622719	0.612089	0.656464	0.62975
Contingency Measure: Concordance (CMC)	0.978946	0.981584	0.980847	0.985011
Contingency Measure: Error Rate (CMER)	0.021054	0.018416	0.019153	0.014989
Contingency Measure: Specificity (CMSP)	0.987883	0.988516	0.988337	0.990965
Contingency Measure: Sensitivity (CMSE)	0.662536	0.677737	0.710136	0.682465
Pierce Skill Score (PSS)	0.597412	0.566429	0.619485	0.591597
Gilbert's Skill Score (GSS)	0.452137	0.441014	0.48861	0.459588
Odd's Ratio (OR)	160.0594	181.0259	207.6152	235.7411
Bias Score (BS)	1.09155	1.181115	1.131629	1.141505
Probability of Detection (POD)	0.662536	0.677737	0.710136	0.682465
False Alarm Ratio (FAR)	0.393032	0.426188	0.372465	0.402136
Critical Success Index (CSI)	0.463632	0.45081	0.499571	0.467749
Boot Strap CI (L)	0.054963	0.059731	0.051483	0.057405
Boot Strap CI (U)	0.055969	0.060751	0.05244	0.059022

CR10 - EOM

Performance Measure	Year 1	Year 2	Year 3	Year 4
Mean Error (ME)	8.15E-05	4.91E-05	1.9E-05	1.54E-05
Mean Absolute Error (MABE)	0.001528	0.00133	0.001423	0.0011
Multiplicative Bias (MB)	1.055023	1.037705	1.012695	1.013753
Mean Squared Error (MSQE)	0.000265	0.000248	0.000262	0.000216
Root Mean Squared Error (RMSE)	0.01629	0.015746	0.016201	0.014713
Root Mean Squared Factor (RMSF)	3.394391	3.381147	3.231348	3.346904
Linear Error in Probability Space (LEPS)	0.001506	0.001261	0.001414	0.001034
Correlation Coefficient (CORR)	0.598413	0.613759	0.627474	0.627232
Nash-Sutcliffe Efficiency Coefficient (NSEC)	0.274664	0.30223	0.32665	0.32954
Heidke Skill Score (HSS)	0.619398	0.630963	0.671163	0.646592
Contingency Measure: Concordance (CMC)	0.978058	0.983166	0.982404	0.986353
Contingency Measure: Error Rate (CMER)	0.021942	0.016834	0.017596	0.013647
Contingency Measure: Specificity (CMSP)	0.986423	0.99032	0.990356	0.992641
Contingency Measure: Sensitivity (CMSE)	0.681917	0.669589	0.695017	0.666884
Pierce Skill Score (PSS)	0.577508	0.604561	0.657564	0.634177
Gilbert's Skill Score (GSS)	0.448643	0.460881	0.505076	0.477751
Odd's Ratio (OR)	155.7538	207.3242	234.0261	270.0263
Bias Score (BS)	1.162623	1.093894	1.043548	1.040806
Probability of Detection (POD)	0.681917	0.669589	0.695017	0.666884
False Alarm Ratio (FAR)	0.413467	0.387885	0.333987	0.359262
Critical Success Index (CSI)	0.460535	0.470116	0.515388	0.485387
Boot Strap CI (L)	0.052746	0.058852	0.051879	0.057584
Boot Strap CI (U)	0.053692	0.059912	0.053019	0.058958

Year 1

Performance Measure	NRT	EOD	EOM
Mean Error (ME)	-1.40E-06	1.99E-05	8.15E-05
Mean Absolute Error (MABE)	0.001572	0.00157	0.001528
Multiplicative Bias (MB)	0.999026	1.013418	1.055023
Mean Squared Error (MSQE)	0.000306	0.000302	0.000265
Root Mean Squared Error (RMSE)	0.017488	0.017375	0.01629
Root Mean Squared Factor (RMSF)	3.430207	3.44396	3.394391
Linear Error in Probability Space (LEPS)	0.001573	0.001547	0.001506
Correlation Coefficient (CORR)	0.535448	0.541896	0.598413
Nash-Sutcliffe Efficiency Coefficient (NSEC)	0.169788	0.174866	0.274664
Heidke Skill Score (HSS)	0.627252	0.622719	0.619398
Contingency Measure: Concordance (CMC)	0.979108	0.978946	0.978058
Contingency Measure: Error Rate (CMER)	0.020892	0.021054	0.021942
Contingency Measure: Specificity (CMSP)	0.987822	0.987883	0.986423
Contingency Measure: Sensitivity (CMSE)	0.670451	0.662536	0.681917
Pierce Skill Score (PSS)	0.599183	0.597412	0.577508
Gilbert's Skill Score (GSS)	0.456932	0.452137	0.448643
Odd's Ratio (OR)	165.0287	160.0594	155.7538
Bias Score (BS)	1.101785	1.09155	1.162623
Probability of Detection (POD)	0.670451	0.662536	0.681917
False Alarm Ratio (FAR)	0.391486	0.393032	0.413467
Critical Success Index (CSI)	0.46841	0.463632	0.460535
Boot Strap CI (L)	0.054714	0.054963	0.052746
Boot Strap CI (U)	0.056091	0.055969	0.053692

Year 2

Performance Measure	NRT	EOD	EOM
Mean Error (ME)	0.000143	9.31E-05	4.91E-05
Mean Absolute Error (MABE)	0.001439	0.001383	0.00133
Multiplicative Bias (MB)	1.110222	1.071559	1.037705
Mean Squared Error (MSQE)	0.000293	0.000261	0.000248
Root Mean Squared Error (RMSE)	0.017105	0.01615	0.015746
Root Mean Squared Factor (RMSF)	3.564883	3.451457	3.381147
Linear Error in Probability Space (LEPS)	0.001292	0.001276	0.001261
Correlation Coefficient (CORR)	0.571098	0.598819	0.613759
Nash-Sutcliffe Efficiency Coefficient (NSEC)	0.176609	0.265958	0.30223
Heidke Skill Score (HSS)	0.613199	0.612089	0.630963
Contingency Measure: Concordance (CMC)	0.981838	0.981584	0.983166
Contingency Measure: Error Rate (CMER)	0.018162	0.018416	0.016834
Contingency Measure: Specificity (CMSP)	0.988922	0.988516	0.99032
Contingency Measure: Sensitivity (CMSE)	0.671277	0.677737	0.669589
Pierce Skill Score (PSS)	0.572724	0.566429	0.604561
Gilbert's Skill Score (GSS)	0.442168	0.441014	0.460881
Odd's Ratio (OR)	182.299	181.0259	207.3242
Bias Score (BS)	1.156875	1.181115	1.093894
Probability of Detection (POD)	0.671277	0.677737	0.669589
False Alarm Ratio (FAR)	0.41975	0.426188	0.387885
Critical Success Index (CSI)	0.451857	0.45081	0.470116
Boot Strap CI (L)	0.062625	0.059731	0.058852
Boot Strap CI (U)	0.063829	0.060751	0.059912

Year 3

Performance Measure	NRT	EOD	EOM
Mean Error (ME)	-3.70E-05	5.70E-05	1.90E-05
Mean Absolute Error (MABE)	0.001459	0.001449	0.001423
Multiplicative Bias (MB)	0.97553	1.03801	1.012695
Mean Squared Error (MSQE)	0.00028	0.000264	0.000262
Root Mean Squared Error (RMSE)	0.01672	0.01625	0.016201
Root Mean Squared Factor (RMSF)	3.240621	3.234682	3.231348
Linear Error in Probability Space (LEPS)	0.001478	0.001416	0.001414
Correlation Coefficient (CORR)	0.595425	0.625884	0.627474
Nash-Sutcliffe Efficiency Coefficient (NSEC)	0.282813	0.322517	0.32665
Heidke Skill Score (HSS)	0.651726	0.656464	0.671163
Contingency Measure: Concordance (CMC)	0.981181	0.980847	0.982404
Contingency Measure: Error Rate (CMER)	0.018819	0.019153	0.017596
Contingency Measure: Specificity (CMSP)	0.989443	0.988337	0.990356
Contingency Measure: Sensitivity (CMSE)	0.682613	0.710136	0.695017
Pierce Skill Score (PSS)	0.632659	0.619485	0.657564
Gilbert's Skill Score (GSS)	0.483378	0.48861	0.505076
Odd's Ratio (OR)	201.5682	207.6152	234.0261
Bias Score (BS)	1.064161	1.131629	1.043548
Probability of Detection (POD)	0.682613	0.710136	0.695017
False Alarm Ratio (FAR)	0.358544	0.372465	0.333987
Critical Success Index (CSI)	0.494093	0.499571	0.515388
Boot Strap CI (L)	0.053676	0.051483	0.051879
Boot Strap CI (U)	0.054876	0.05244	0.053019

Year 4

Performance Measure	NRT	EOD	EOM
Mean Error (ME)	2.18E-06	5.91E-05	1.54E-05
Mean Absolute Error (MABE)	0.001153	0.001132	0.0011
Multiplicative Bias (MB)	1.001932	1.052694	1.013753
Mean Squared Error (MSQE)	0.00024	0.000222	0.000216
Root Mean Squared Error (RMSE)	0.015508	0.014884	0.014713
Root Mean Squared Factor (RMSF)	3.403268	3.385519	3.346904
Linear Error in Probability Space (LEPS)	0.001087	0.00104	0.001034
Correlation Coefficient (CORR)	0.590233	0.623374	0.627232
Nash-Sutcliffe Efficiency Coefficient (NSEC)	0.260504	0.313891	0.32954
Heidke Skill Score (HSS)	0.625337	0.62975	0.646592
Contingency Measure: Concordance (CMC)	0.985235	0.985011	0.986353
Contingency Measure: Error Rate (CMER)	0.014765	0.014989	0.013647
Contingency Measure: Specificity (CMSP)	0.991787	0.990965	0.992641
Contingency Measure: Sensitivity (CMSE)	0.654693	0.682465	0.666884
Pierce Skill Score (PSS)	0.605587	0.591597	0.634177
Gilbert's Skill Score (GSS)	0.454902	0.459588	0.477751
Odd's Ratio (OR)	228.9617	235.7411	270.0263
Bias Score (BS)	1.068989	1.141505	1.040806
Probability of Detection (POD)	0.654693	0.682465	0.666884
False Alarm Ratio (FAR)	0.387559	0.402136	0.359262
Critical Success Index (CSI)	0.462911	0.467749	0.485387
Boot Strap CI (L)	0.060738	0.057405	0.057584
Boot Strap CI (U)	0.061933	0.059022	0.058958

Proximity Metrics

Distance Measure	Water Year 1	Water Year 2	Water Year 3	Water Year 4
Euclidean	0.741	0.744	0.543	0.465
City Block	1.838	1.795	1.480	1.263
Minkowski	51.644	53.010	47.348	40.024
Chebyshev	0.640	0.664	0.457	0.395
Sorensen	0.919	0.898	0.740	0.632
Gower	0.006	0.006	0.005	0.004
Soergel	1.828	1.785	1.473	1.259
Lorentzian	1.837	1.793	1.479	1.263
Kulczynski	1.849	1.805	1.486	1.268
Canberra	5758.219	6225.606	5973.827	6047.979
Intersection	0.919	0.898	0.740	0.632
Wave Hedges	6446.285	6959.070	6666.633	6698.211

Distance measures for histogram comparison for different water years using rain gage (CR10) and radar (NRT) data

Observations

- Three data sets derived from radar-based rainfall estimates and two data sets from rain gage network are used for the analysis of bias in the current study.
- The rain gage data sets have missing data and few anomalous values. The values exceeding a specific threshold are either limited to a maximum threshold value or not used for analysis.
- Rain gage augmented with radar data sets are also used for analysis in the current study. The time intervals in which the rain gage data is augmented are tagged. The percentage of augmented data ranged from 1% to 100%.
- Data is split into four water years and analysis is carried out.

Main Conclusions and Summary of Bias Data Analysis

- The radar data is in good agreement with rain gage data at different temporal scales (i.e., 15 minute, hour, day, month and year).

- Scalar performance measures and skill scores calculated based on near real-time (NRT), end of the day (EOD) and end of the month (EOM) radar data sets and CR10 rain gage data indicate good agreement of radar data with rain gage data.

- As expected the performance of radar data sets based on all indices and skill scores progressively improved as temporal scale of radar data adjustment for analysis is changed from near real-time to the end of the day to the end of the month. The improvement is marginal when moving from near real-time to end of the day.

- The skill scores based on radar data compared with rain gage data also improved over different years and best performance often seen in the most recent water year.

- Visual and statistical tests also suggest that radar data at all levels (near real-time, end of the day and end of the month) are in good agreement with ground truth (i.e. rain gage data).

Conclusions and Summary of Bias Data Analysis

- In general the errors computed based on radar and rain gage data show no evidences of temporal dependence, major heteroskedasticity issues, non-normality of errors (residuals) and statistically significant autocorrelation at different lags.

- Replacements tests indicate that radar data can used for augmenting rain gage data or infilling missing data in different modes: 1) systematic; 2) random and 3) systematic and random replacement. Homogeneity tests need to be used to confirm that rain gage data is homogeneous after infilling.

- Spatial evaluation of bias indicates significant and consistent under estimation of precipitation by radar data in two areas of the SFWMD region. These areas are generally located in the North and Southeast corners of the SFWMD region.

Conclusions and Summary of Bias Data Analysis

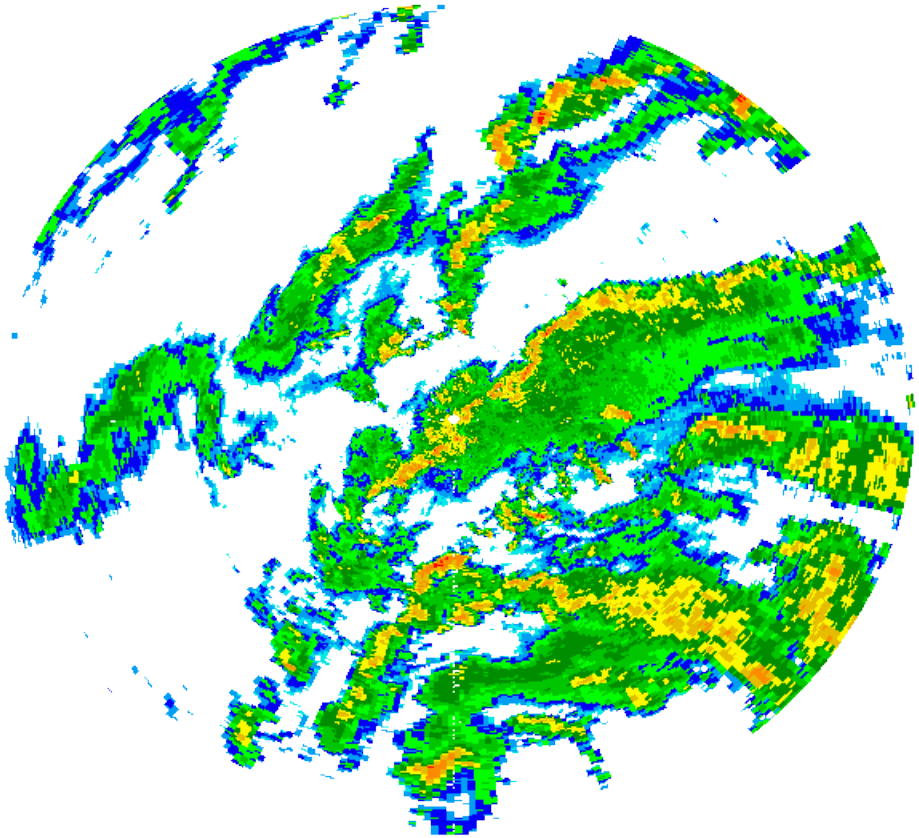
- Parametric and non-parametric tests used for evaluating similarity of distributions characterizing radar and rain gage data sets provided mixed results. The two sample Kolmogorov-Smirnov (KS), Ansari-Bradley and Wilcoxon Sum tests indicate that alternate hypothesis being true for observations at several sites.

- Transition probabilities P_{10} (probability of transition from wet to dry) and P_{11} (probability of transition from wet to wet) are underestimated and overestimated by radar data when compared to rain gage data for temporal scale of comparison of 15 minutes. However, a different picture of probabilities evolves when the temporal scale is increased from 15 minutes to an hour or a day.

- Serial autocorrelation of rain gage and radar data are evaluated at different lags. Radar data seems to overestimate the autocorrelation at first two to three lags.

- Accumulated radar and rain gage data plots show no significant deviation of radar from rain gage data for all the gages in all the water years.

Thank you



Range and perfect score values for different indices or measures.

Index/Measure or Skill Score	Possible Range of values	Perfect (ideal) score
Mean Error (ME)	$-\infty$ to $+\infty$	0
Mean Absolute Error (MAE)	0 to $+\infty$	0
Mean Squared Error (MSE)	0 to $+\infty$	0
Root Mean Squared Error (RMSE)	0 to $+\infty$	0
Nash-Sutcliffe Efficiency Criterion (NSCE)	0 to 1	1
Correlation Coefficient (ρ) [Pearson]	-1 to 1	1
Spearman Correlation Coefficient (ρ_{rank})	-1 to 1	1
Concordance	0 to 1	1
Error Rate	0 to 1	0
Specificity	0 to 1	1
Sensitivity	0 to 1	1
Probability of Detection (POD)	0 to 1	1
Probability of False Detection (POFD)	0 to 1	0
Bias Score (BS)	0 to $+\infty$	1
Heidke Skill Score (HSS)	$-\infty$ to 1	1
Critical Success Index (CSI)	0 to 1	1
Peirce' s skill score (PSS)	-1 to 1	1
Gilbert' s Skill Score (GSS)	-1/3 to 1	1
Odd' s ratio	0 to $+\infty$	$+\infty$

Skill Scores

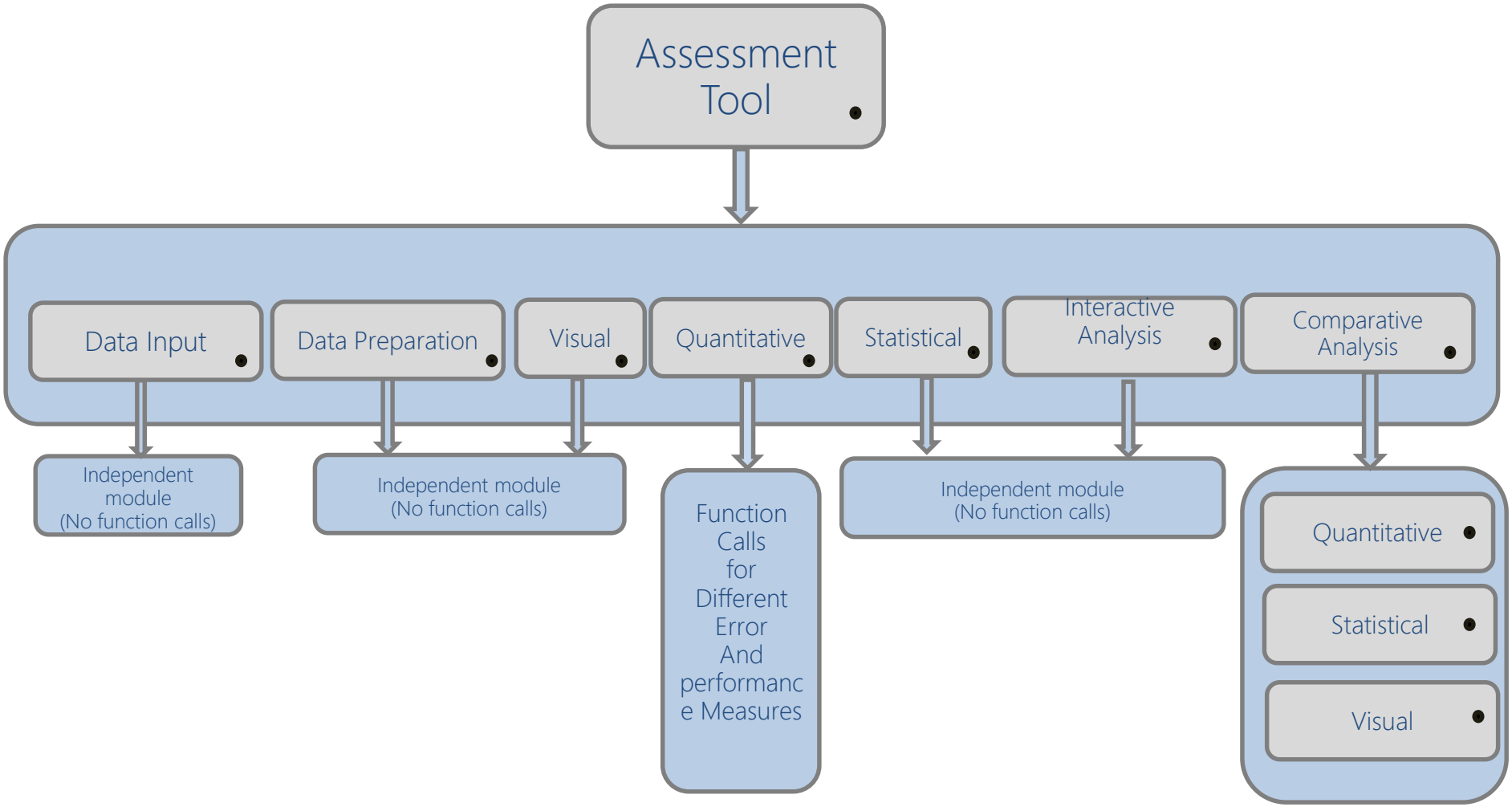
- Critical success index (CSI) provides information about how well the wet events obtained by radar-based precipitation data correspond to rain gage observed wet events.
- The Peirce skill score answers the question: What was the accuracy of the radar-based rainfall estimate in predicting the correct category, relative to that of random chance?
- The GSS referred to as equitable threat score answers the questions: How well did the wet events estimated by radar-based rainfall estimates correspond to the observed wet events (accounting for hits due to chance).
- Odds Ratio: What is the ratio of the odds of a wet event indicated by rain gage observation being correct based on radar-based estimation, to the odds of a wet event estimated by radar data being wrong? The score measures the ratio of the odds of making a hit to the odds of making a false alarm

The *RMSF* is the exponent of the root mean square error (multiplicative) of the logarithm of the data. The logarithmic transformation is mainly performed to smooth the data, reduce the discontinuities, and make the data more robust (CAWCR, 2012).

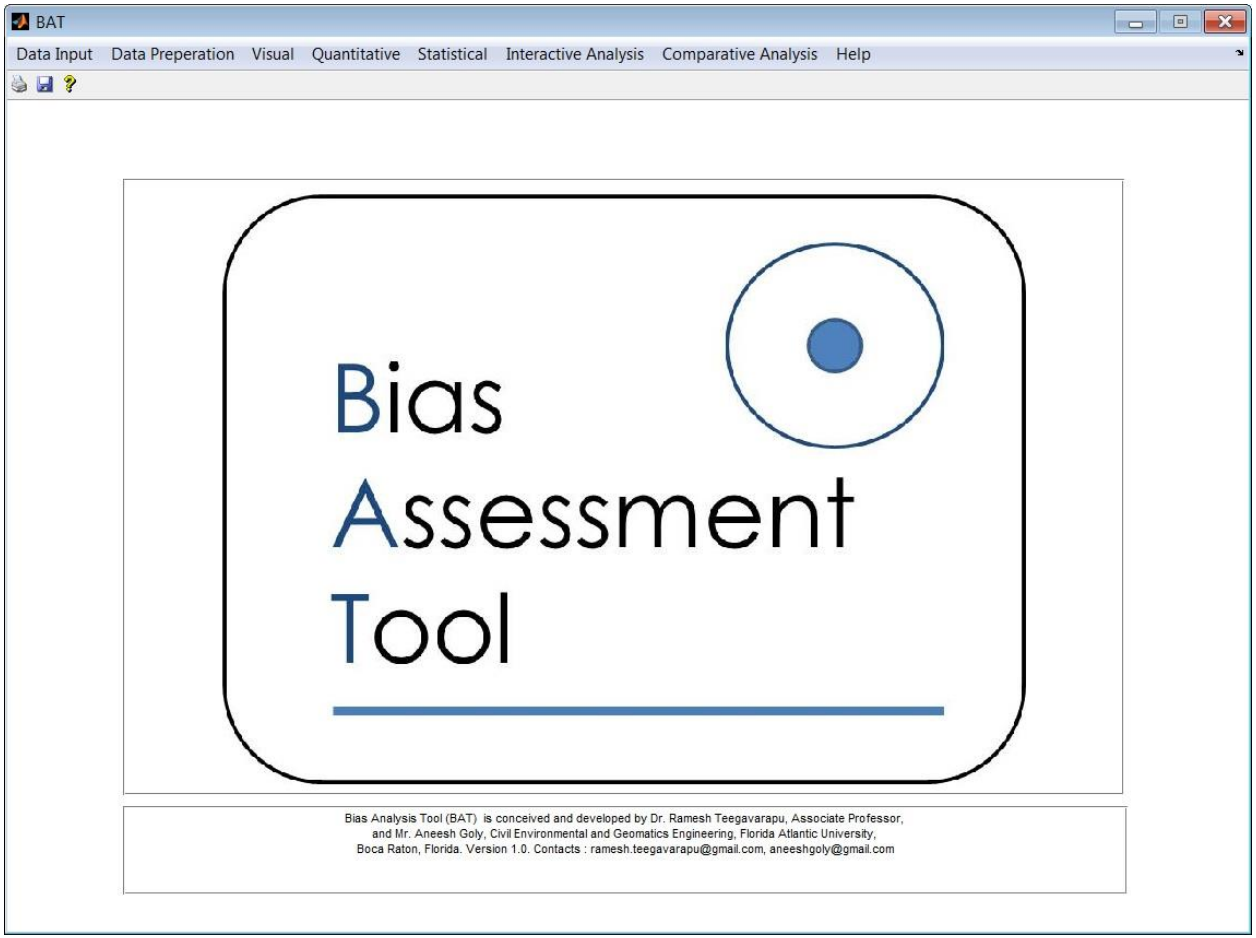
$$RMSF = exp\left\{\sqrt{\frac{1}{N}\sum_{i=1}^{N} ln\,(\theta_{n,i}/\theta_{r,i})^{\,2}}\right\}$$

An RMSE can be interpreted as giving a scale to the additive error; the RMSF can be interpreted as giving a scale to the multiplicative error (CAWCR, 2012).

Assessment Tool Architecture



Bias Assessment Tool (BAT)
Main Interface



Main Modules

Data Input

A module to accept data in a specified format

Data Preparation

This module prepares the data set for analysis and check the data format and completeness

Visual

This is a non-interactive module for evaluation of visual indices

Quantitative

This is a non-interactive module for evaluation of quantitative indices

Statistical

This is a non-interactive module for evaluation of statistical indices

Interactive Analysis

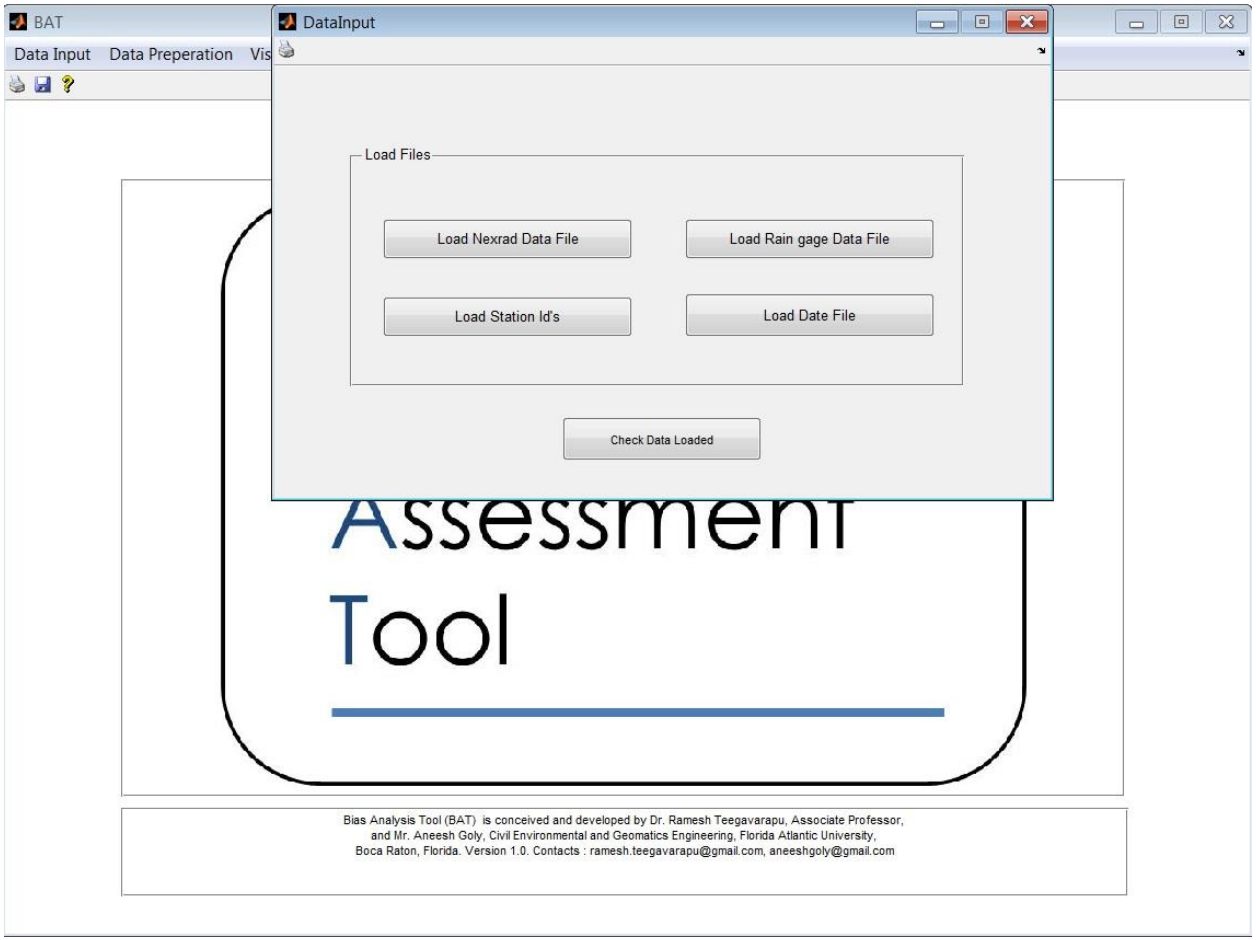
This is an interactive module as the name suggests for all the three (visual, quantitative and statistical indices)

Comparative Analysis

This is again an interactive module with option of comparative analysis option (side-by-side) for all indices for two different temporal time frames.

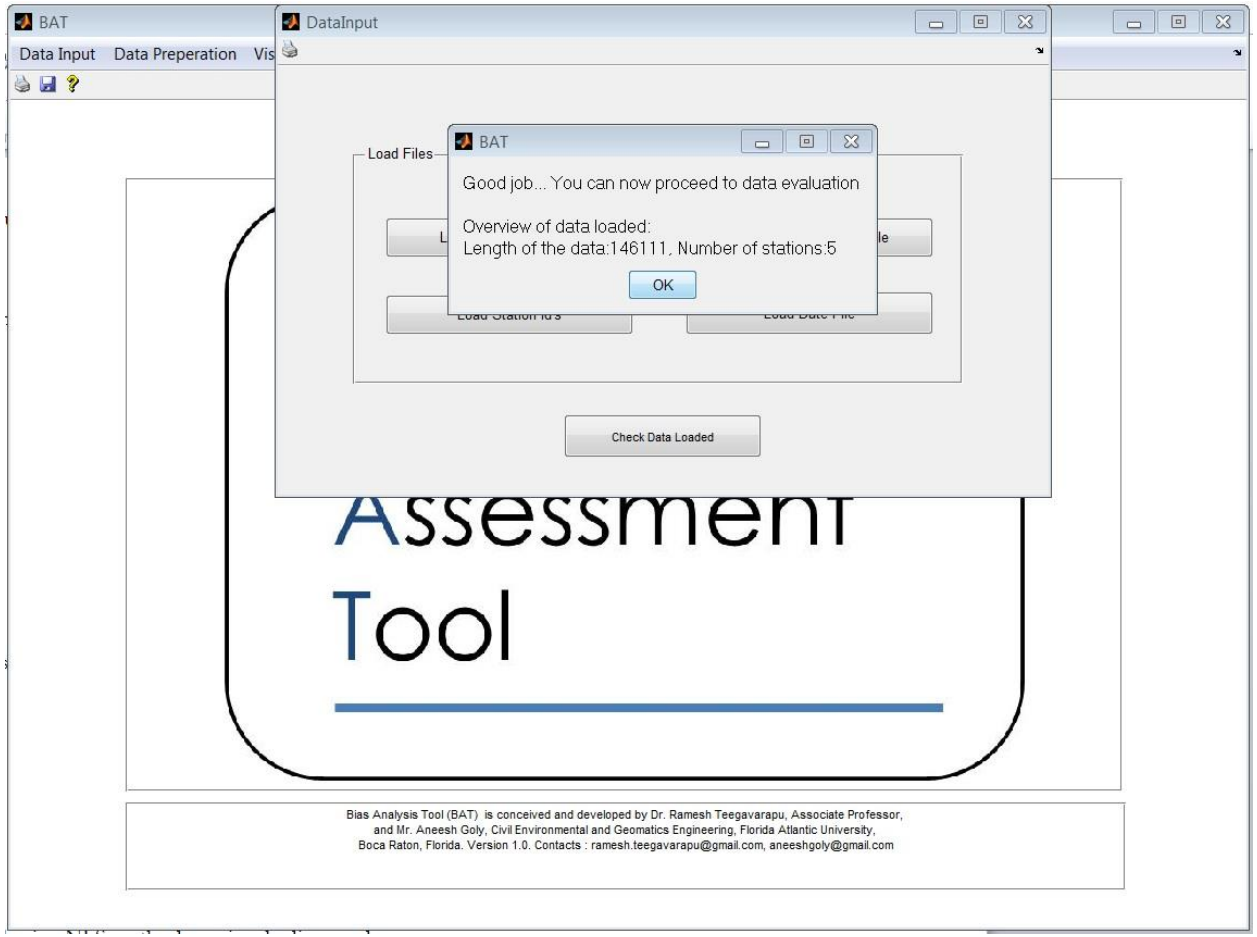
Bias Assessment Tool (BAT)

Data Input Module



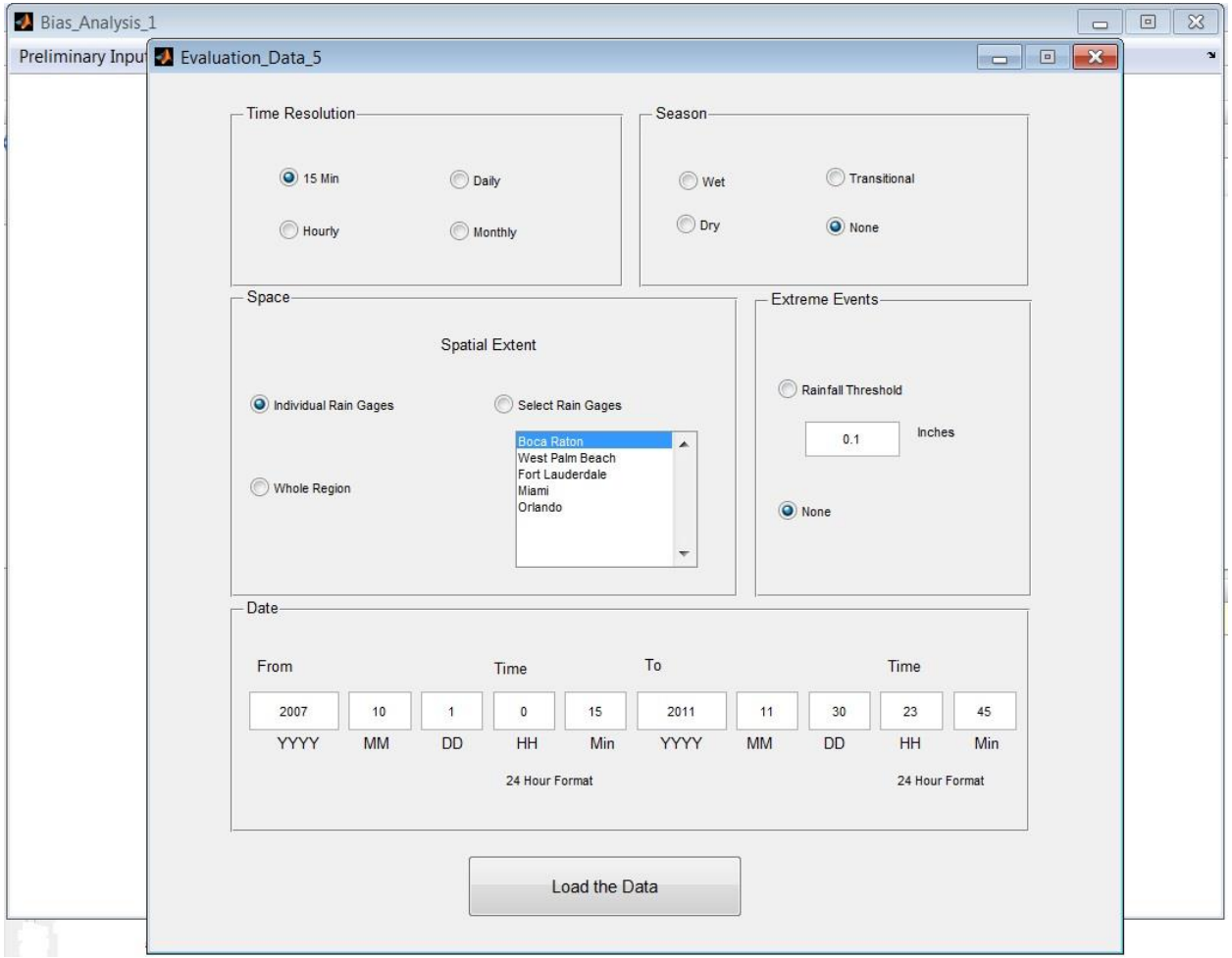
Bias Assessment Tool (BAT)

Data Input Module- Check



Bias Assessment Tool (BAT)

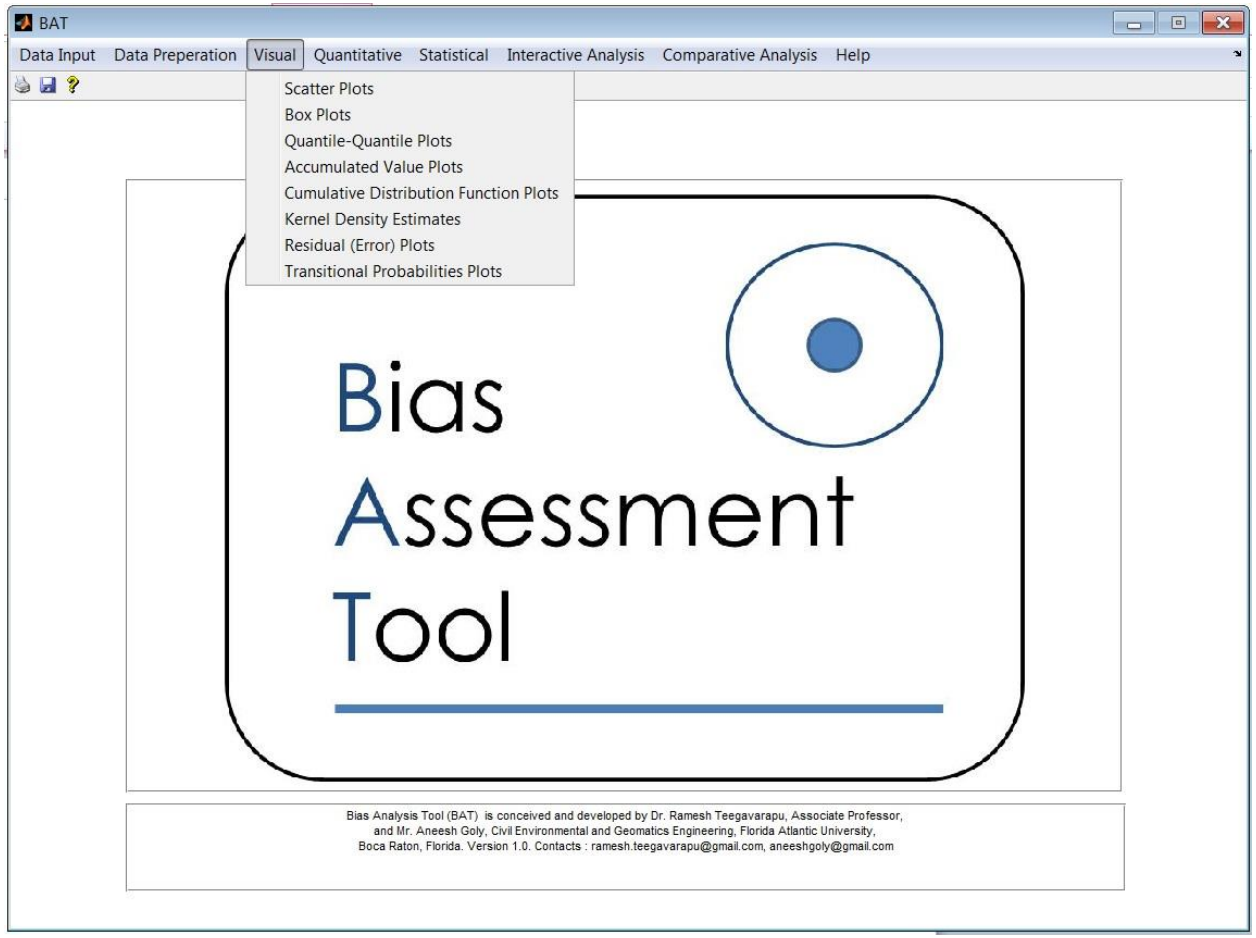
Data Preparation Module: Spatial and Temporal Resolution



Dr. Ramesh Teegavarapu, Professor, CEGE, Florida Atlantic University, Director, HRL

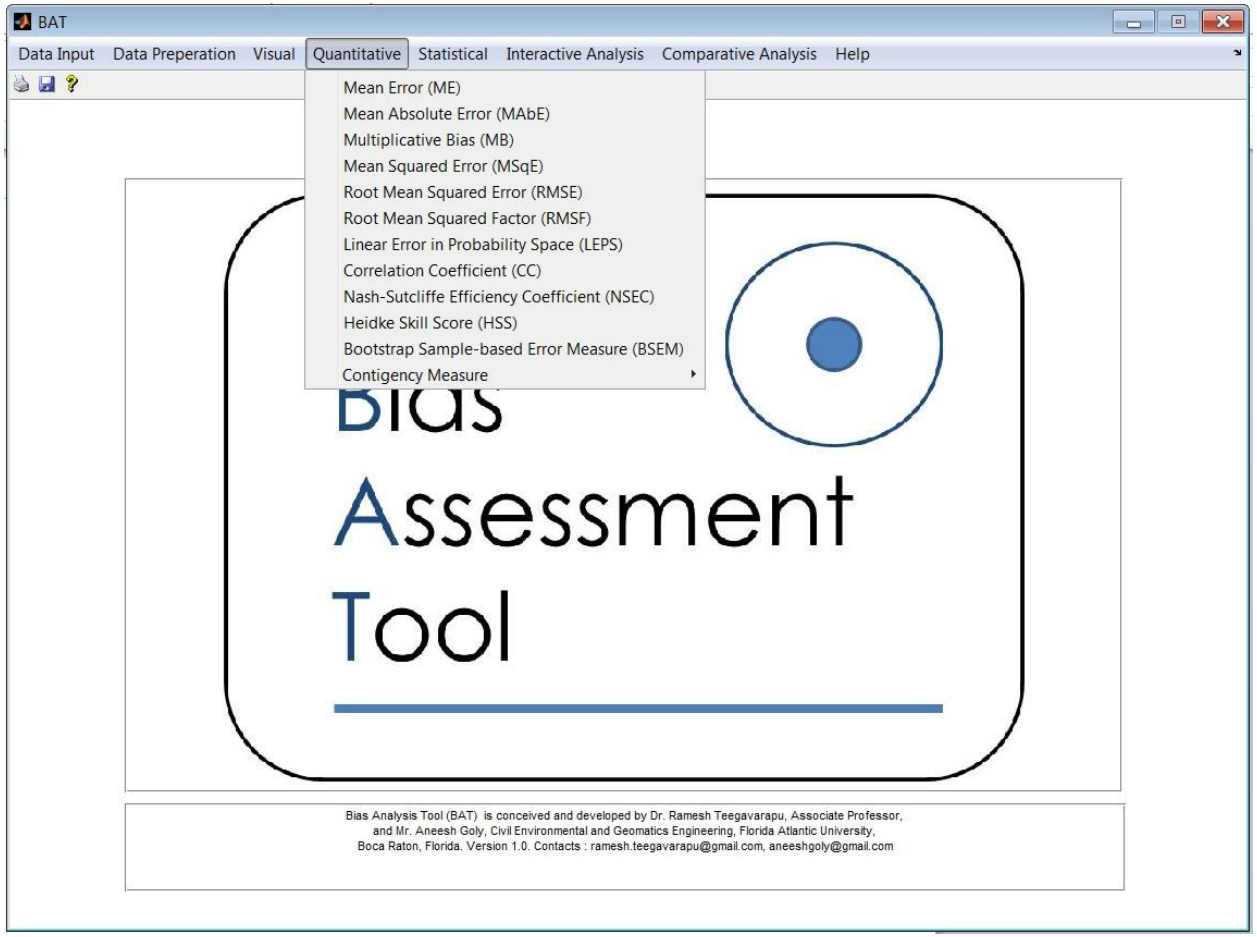
Bias Assessment Tool (BAT)

Visual Module: Different Assessment Indices



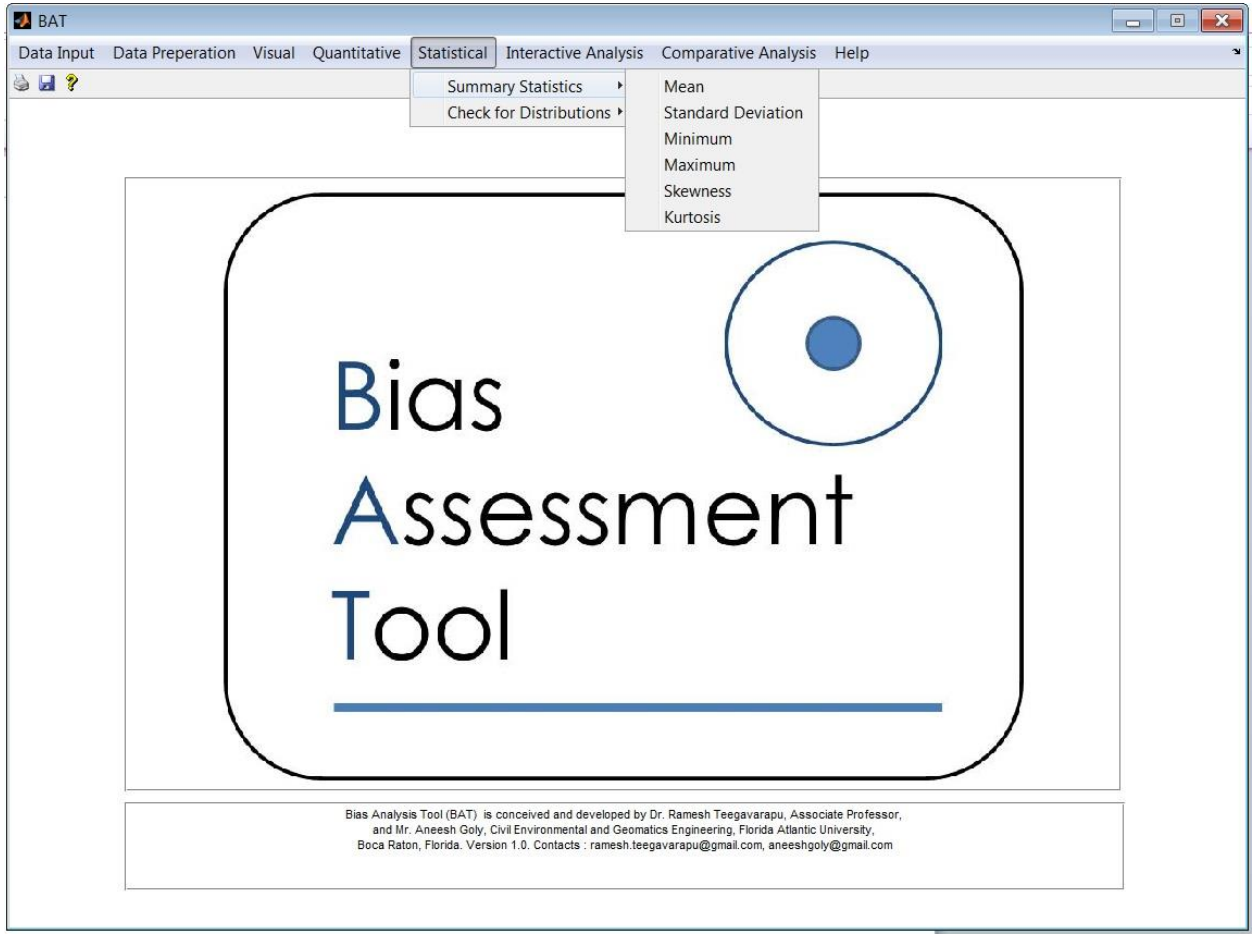
Bias Assessment Tool (BAT)

Quantitative Module: Different Assessment Indices



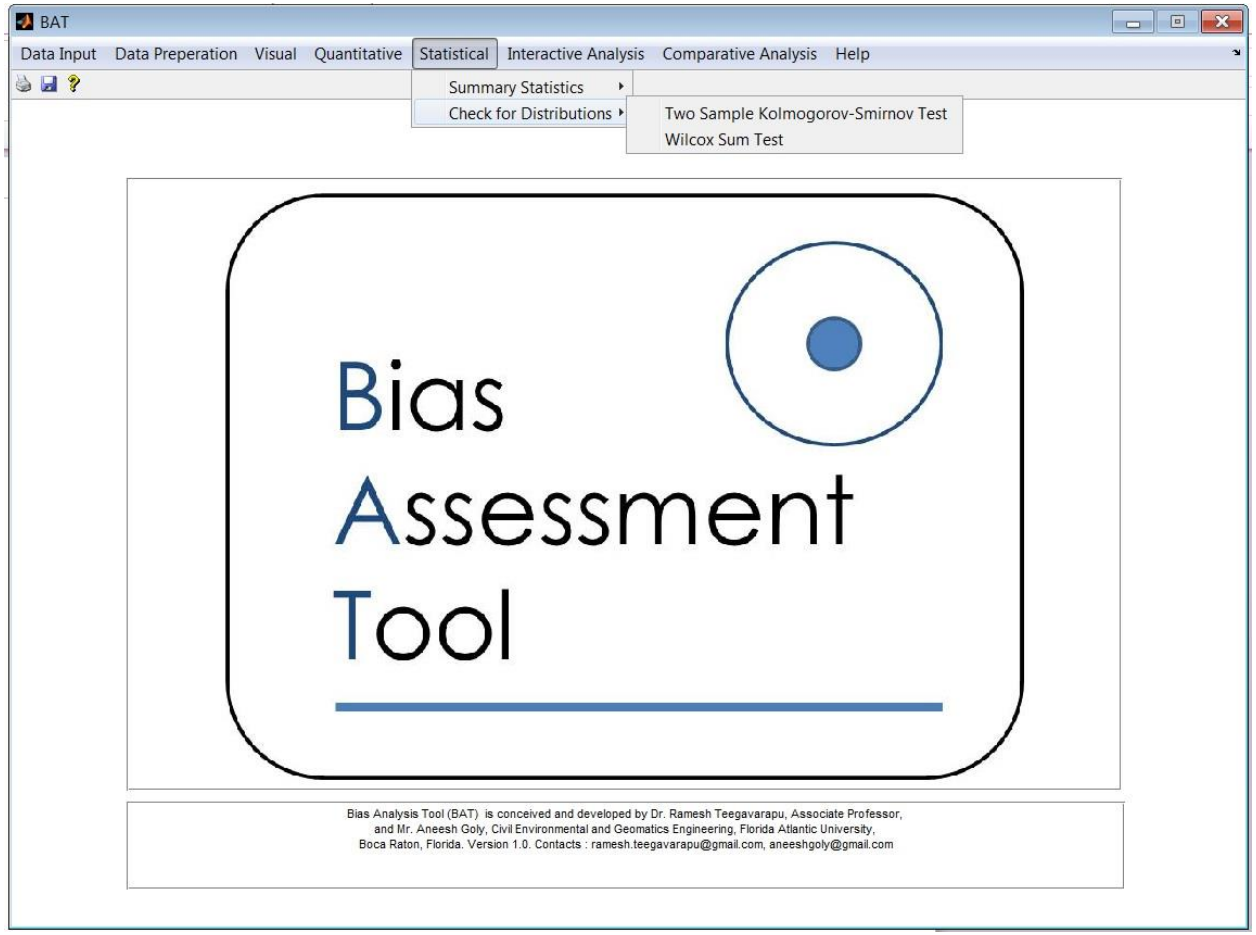
Bias Assessment Tool (BAT)

Statistical Module: Different Assessment Indices

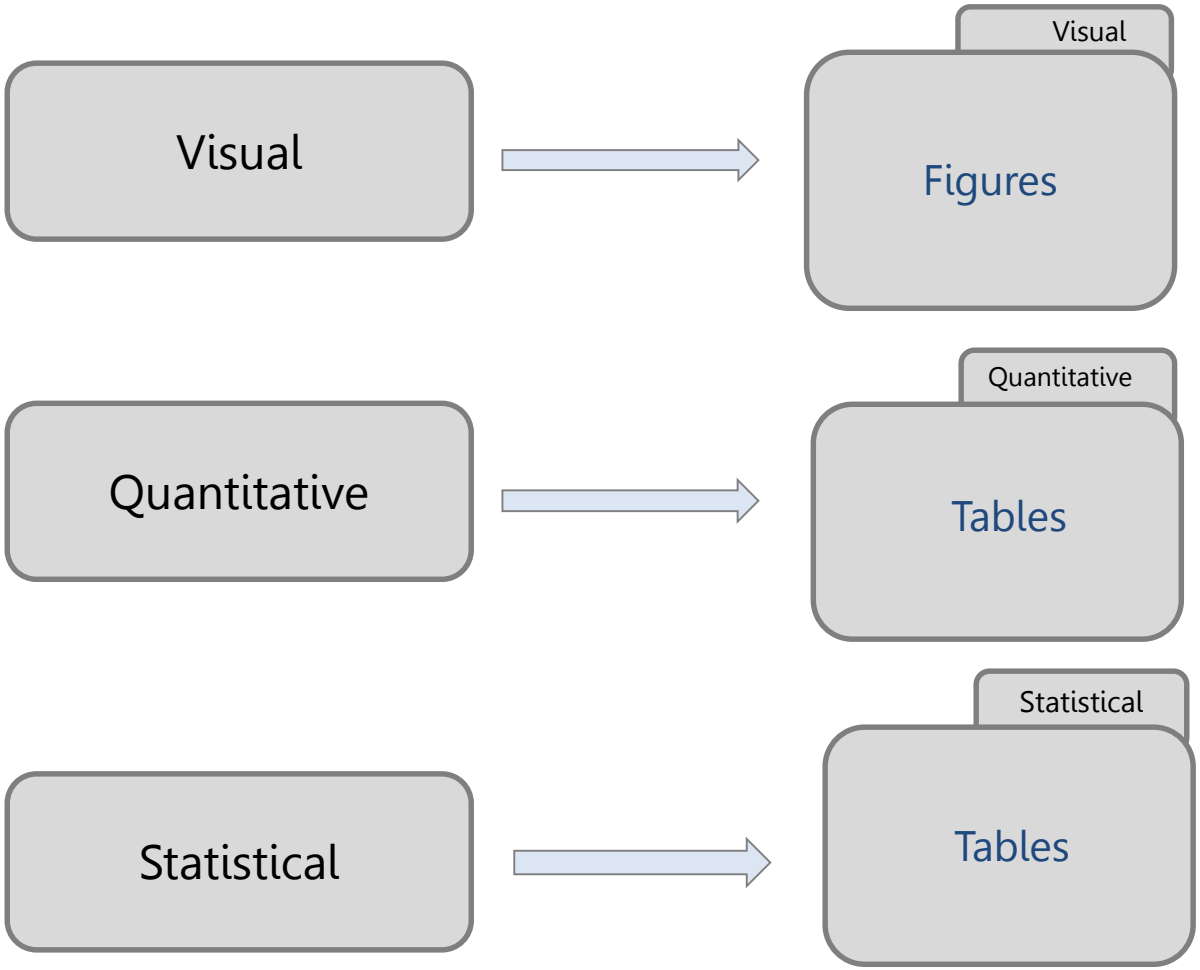


Bias Assessment Tool (BAT)

Statistical Module: Assessment of Distributions

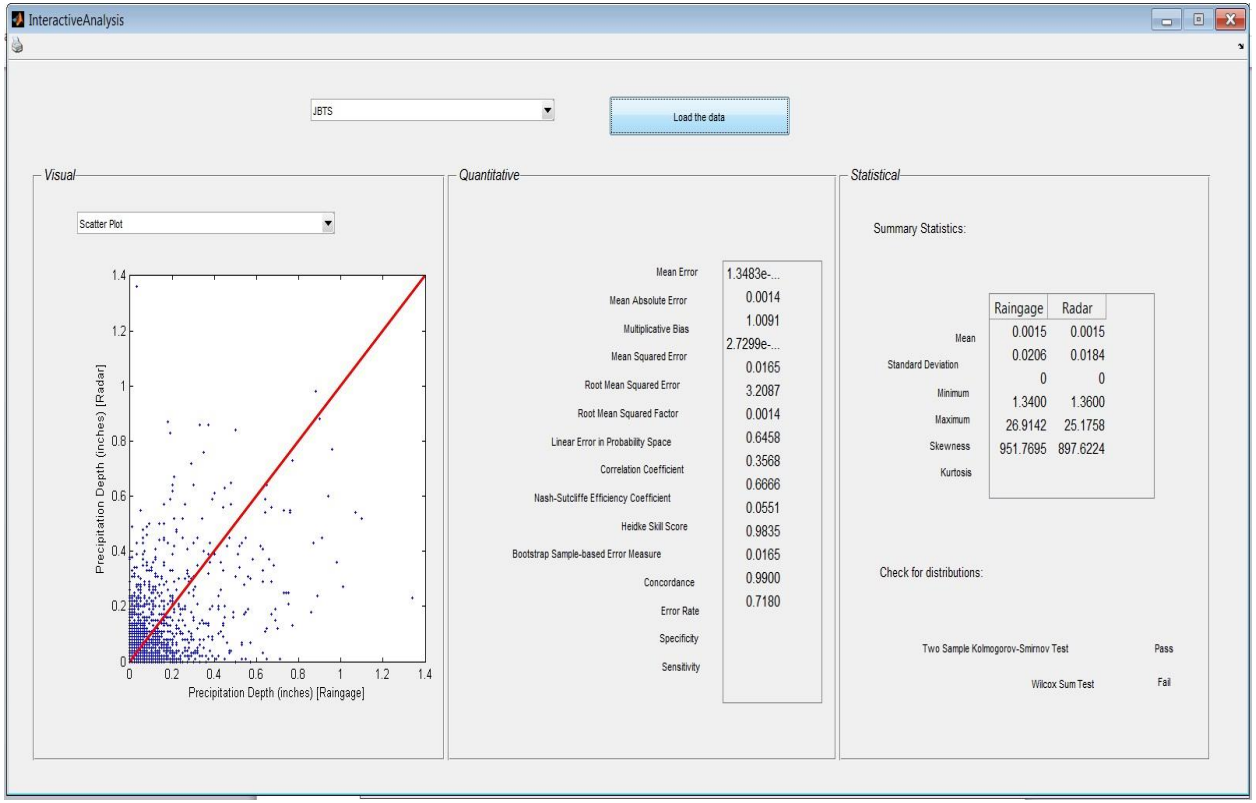


Background saving of results



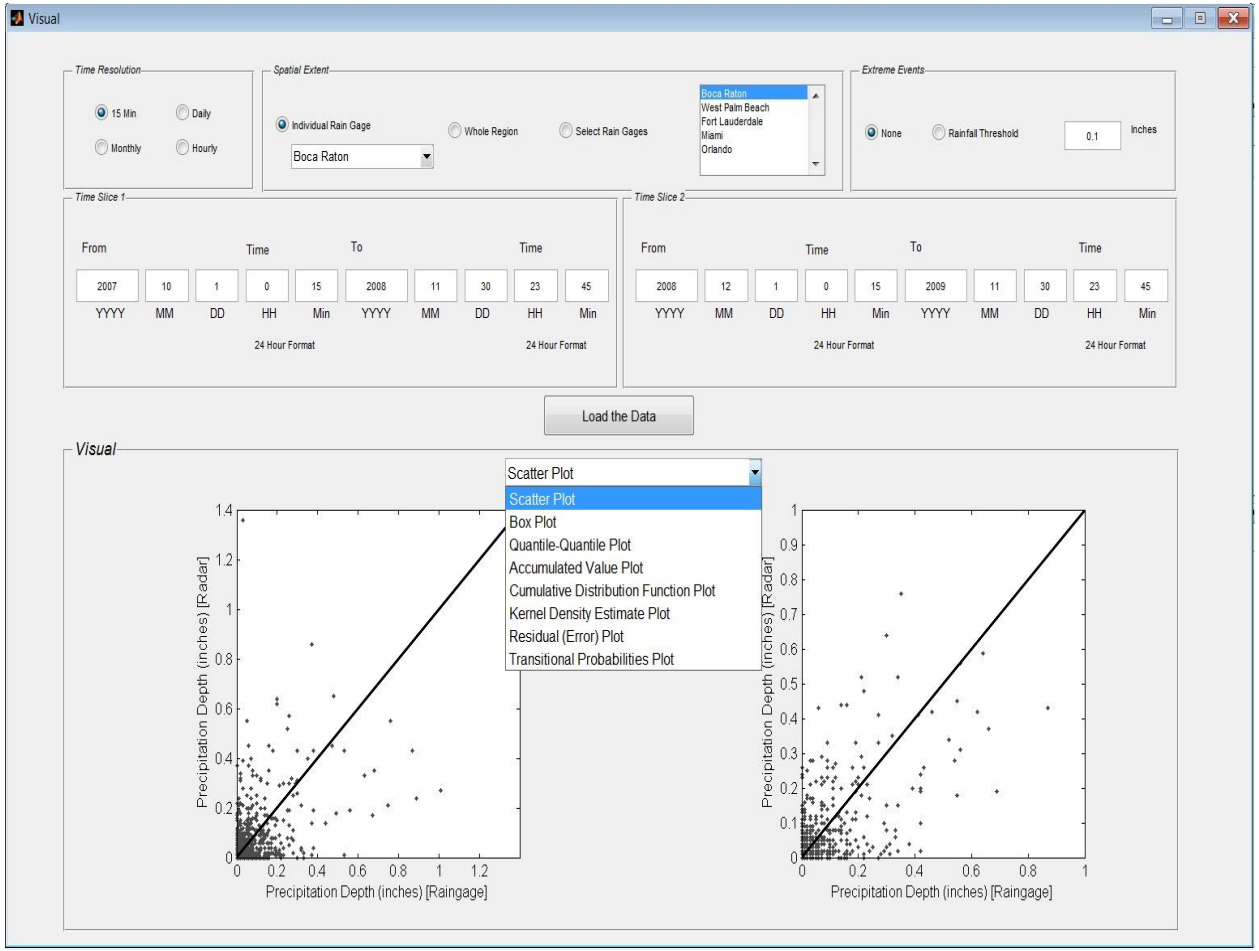
Bias Assessment Tool (BAT)

Interactive Assessment Module: Assessment of all indices



Bias Assessment Tool (BAT)

Comparative Assessment Module: Assessment of all indices



Dr. T. Ramesh Teegavarapu Ph.D., P.E