Estimation of Rain height from Rain Rate using Regression-based Statistical Model:

Application to SeaWinds on ADEOS-II

Bhaskar Natarajakumar

28th January 2004

Committee

Dr. Glenn Prescott (Chair)

Dr. Richard K. Moore

Dr. David Braaten

Outline

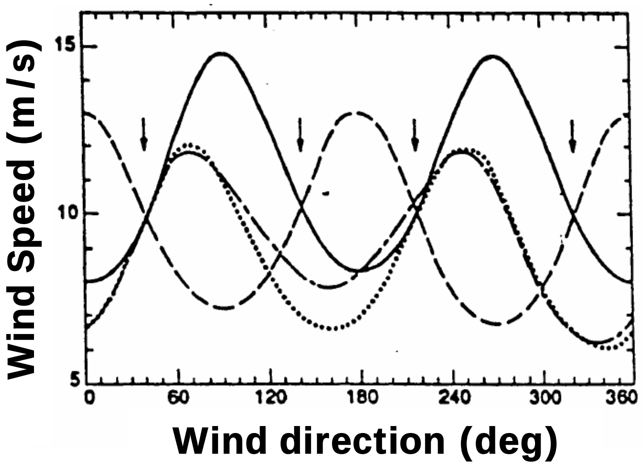
- **Introduction**
  - Thesis Objective
- **Methodologies for Rain Height (RH) Estimation**
  - Estimation using Rain Rate (RR)
- **Regression Analysis**
  - Log-Linear Combined Regression scheme
- **Regression-based Statistical Model**
- **Validation of Statistical Model**
- **Conclusion**

Introduction

- Winds are the single largest source of momentum for the ocean surface.
- They regulate crucial coupling between air and sea, which establishes and maintains global and regional climate.
- Wind vectors are crucial inputs for:
  - Global numerical weather prediction systems
  - Monitoring unusual climatological phenomena (hurricane, El-Niño)
  - Other meteorological and oceanographic applications
- Instruments on buoys and ships can provide measurement of surface wind vectors, but they are inaccurate and their coverage is insufficient.
- Only satellite-borne instruments can acquire wind data with global coverage, high spatial resolution and frequent sampling.

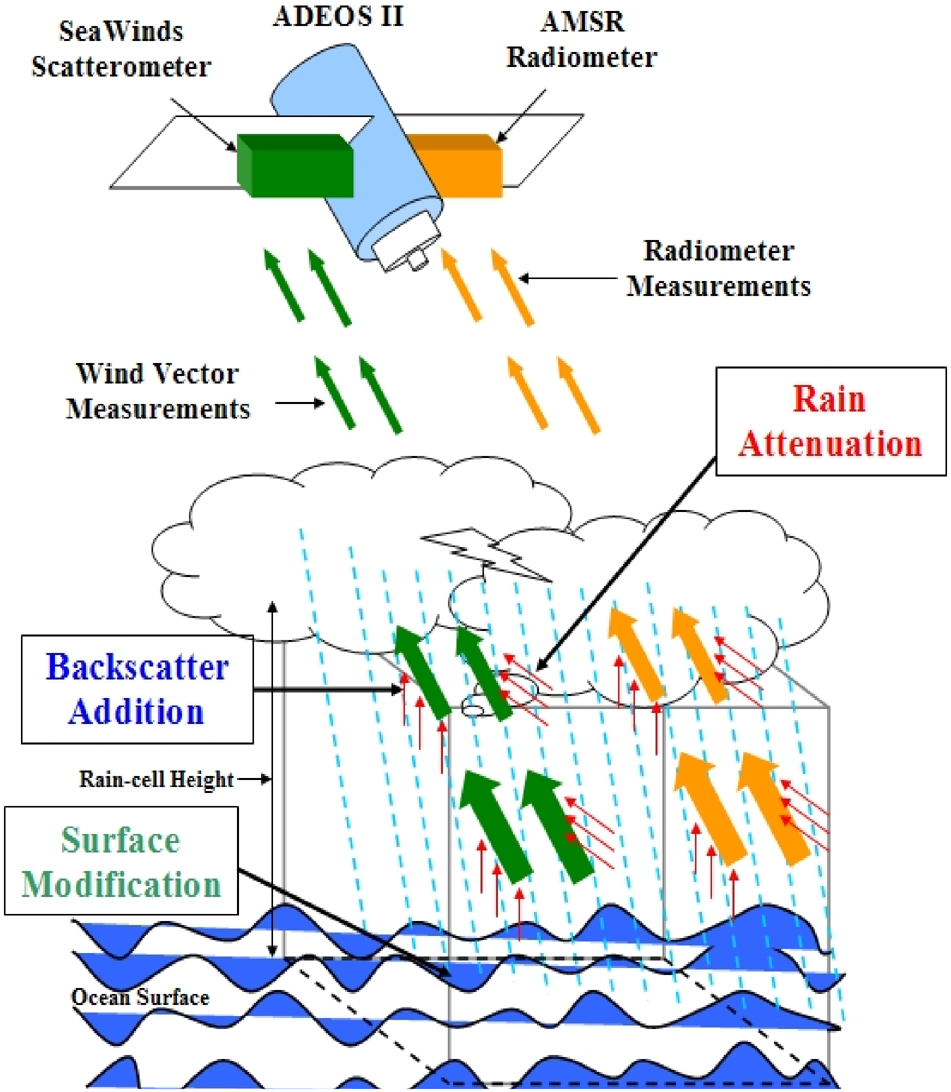
Scatterometry

- Scatterometry for Wind Vector Measurements
  - Scatterometer is a microwave radar sensor, which can be used to measure ocean near-surface wind vectors for all weather conditions.
  - Radar response (surface scattering coefficient) can be related to wind vectors using Geophysical Model Functions.
  - Need for Multiple Measurements.
- SeaWinds on ADEOS-II
  - Earlier scatterometer missions: SEASAT, ERS-1/2, NSCAT, QuikSCAT.
  - NASA SeaWinds scatterometer designed at Ku-band (13.4 GHz)
  - SeaWinds uses a rotating dish antenna, a pencil-beam type.



Rain Effects on Wind Data

- Rain types: convective and stratiform
- Rain effects on SeaWinds data
  - Rain Attenuation
  - Backscatter Addition
  - Surface Modification
- The energy backscattered by the rain can be a significant portion of the total backscatter power measured by radar.
- Most of the efforts of the research community were directed toward flagging rain-contaminated cells and only a few attempts were made to correct wind vector estimates for rain effects.



Rain Effects Correction Algorithm

- Our Research
  - Develop an algorithm to correct for the rain effects, when possible
  - Develop criterion (threshold) for rain flag when correction impossible

- Total Received power by the scatterometer is given as the sum of attenuated surface signal and the precipitation echo.

$$P_r = P_s \tau^2 + P_v \xi$$

where, P_r is the received signal

P_s is the surface signal

τ is the Total one-way transmissivity of atmosphere (<1)

P_v is the power from rain volume

ξ accounts for attenuation of rain echo

$$\sigma_r^0 A_s = \sigma_s^0 A_s e^{-2kR} + \eta A_s H \xi$$

Where σ_r^0 is scattering coefficient with no correction

A_s is footprint area

σ_s^0 is surface scattering coefficient

k is rain attenuation constant

H is Rain height - ($A_s H$) is scattering volume

R is slant range through the rain also related to rain height

η is volume scattering coefficient

Rain Effects Correction Algorithm (contd.)

- Correction Algorithm
  - To correct the signal, we invert the previous equation

$$\sigma_s^0 = (\sigma_r^0 - \eta H \xi) e^{2kR}$$

  - The correction of σ_s^0 from rain effects requires the following inputs:
    - Reflectivity factor (*Z*) related to volume scattering coefficient
    - Attenuation constant (*k*)
    - Rain height (*H*)
  - The reflectivity factor and attenuation constant can be obtained using empirical relationships with rain rate.

$$k = \alpha (RR)^{\beta} \qquad Z = a(RR)^{b}$$

  - α and β, and *a* and *b* (different for stratiform and convective), are found in the literature.
- The availability of rain rate from AMSR radiometer in ADEOS II made this approach possible for ADEOS-II, and presumably will do so for future missions.

Thesis Objective

- Rain height (RH) or storm height as used in our analysis, is the height of the top of the rain column above the mean sea level – first echo measured by radar

- Rain height is critical to get path length for attenuation and volume for backscatter.

- Unfortunately, Rain height is not available in any of the instruments carried by ADEOS-II.

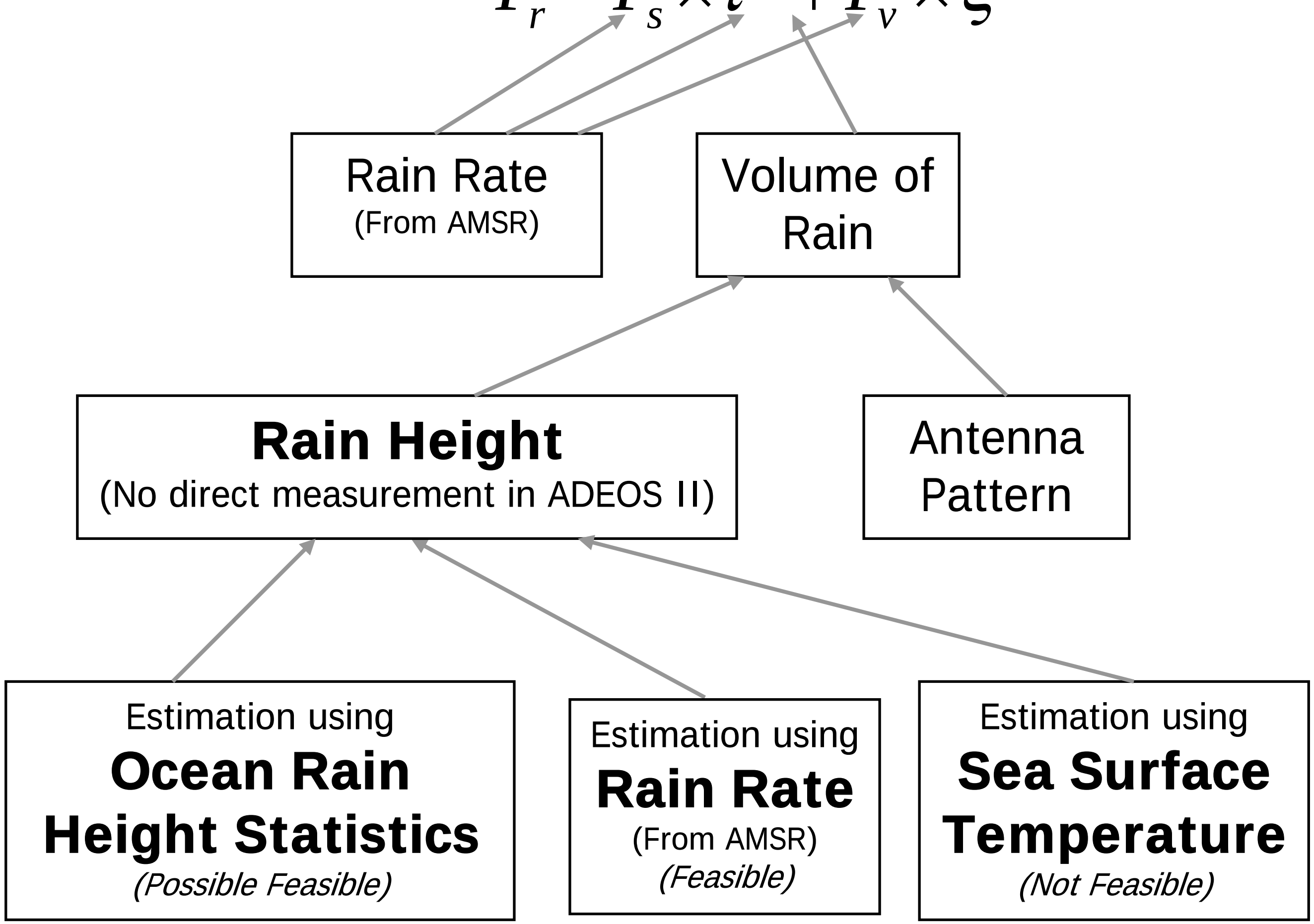
- ## Thesis Objective
  - To identify a methodology to estimate rain cell height over oceans using statistics or from any available ADEOS-II instrument's output
  - This estimated rain cell height should be valid over different seasons, rain types and regions of the globe.

Outline

- **Introduction**
  - Thesis Objective
- **Methodologies for Rain Height (RH) Estimation**
  - Estimation using Rain Rate (RR)
- **Regression Analysis**
  - Log-Linear Combined Regression scheme
- **Regression-based Statistical Model**
- **Validation of Statistical Model**
- **Conclusion**

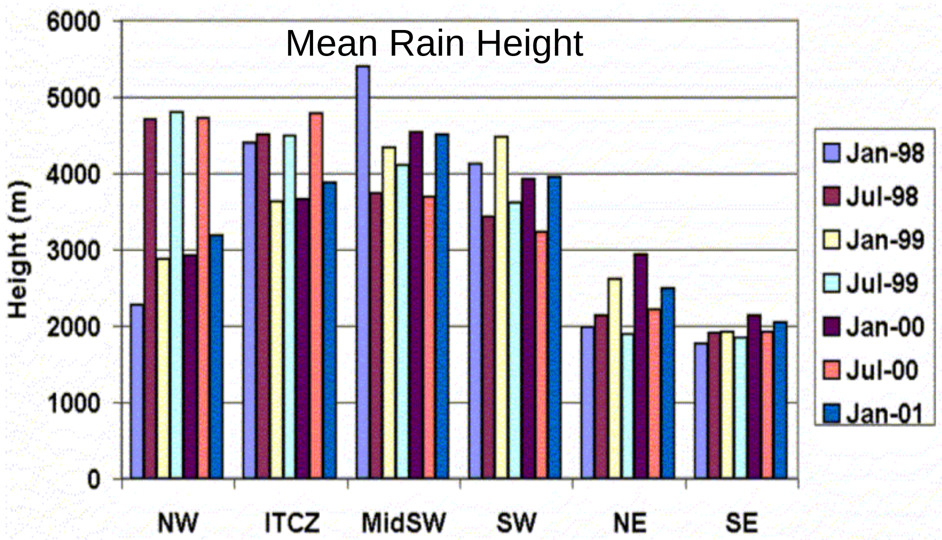
Methodologies for RH estimation

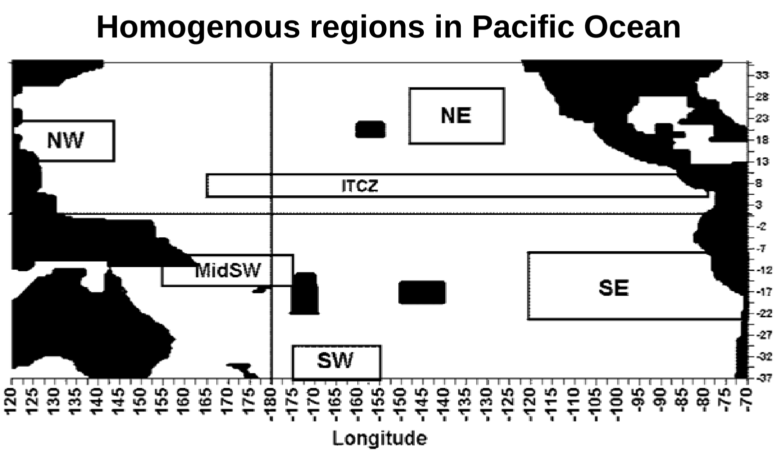
$$P_r = P_s \times \tau^2 + P_v \times \xi$$



Climatological Rain Height

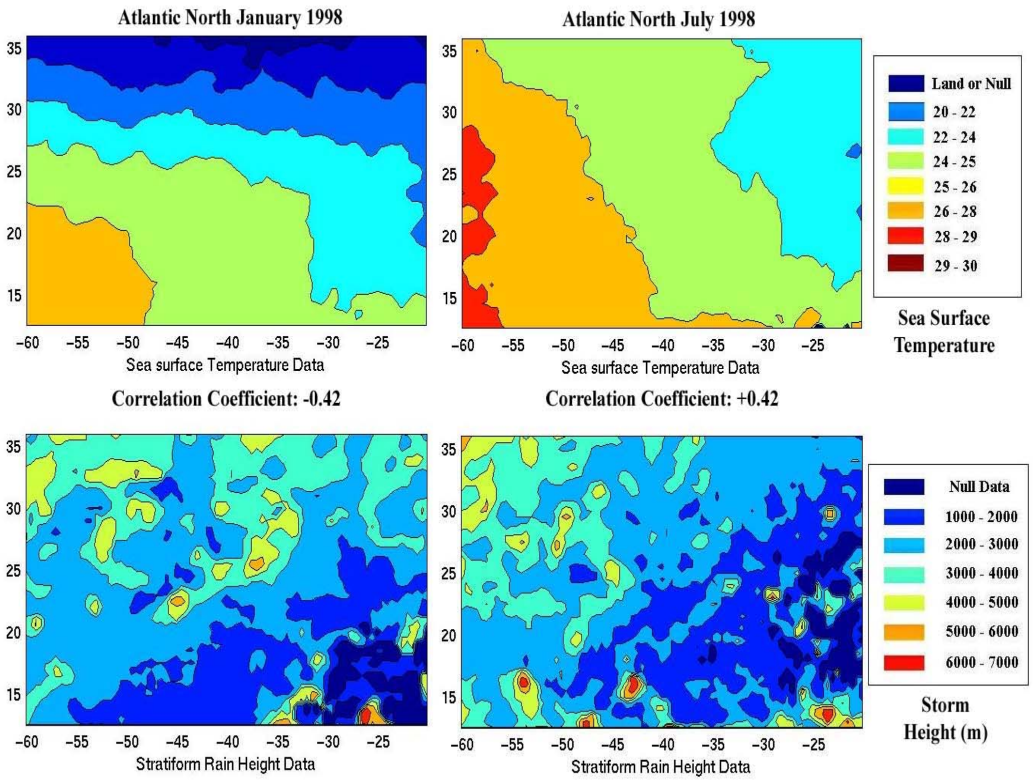
- Rain Height Statistics
  - Conventional statistics from weather services do not provide this information.
  - We attempted to analyze Level-3 TRMM precipitation radar monthly averaged rain-height data and generate mean rain-height tables.
  - Oceanic regions are subdivided into homogenous rain height regions to account for variation in regional ocean climate.
- This approach provides only a crude statistical measure of rain height.
- Another drawback of this approach, is the limited latitude coverage of TRMM (±35° Latitude). Hence, mean rain-height tables cannot be developed for higher latitudes.





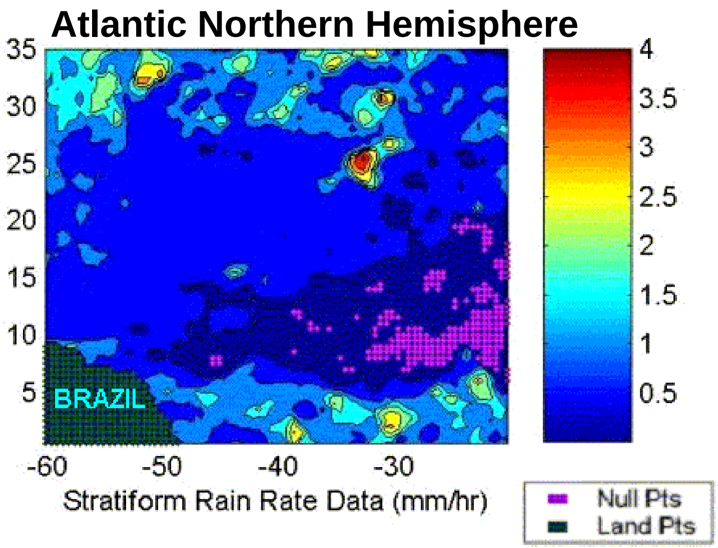
From Sea Surface Temperature

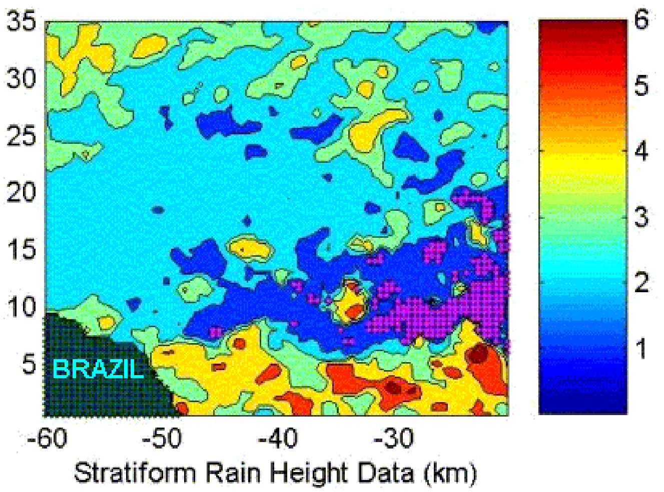
- Estimation of rain height from Sea-surface Temperature
  - TRMM PR rain height products and AVHRR sea-surface temperature (SST) data are correlated.
  - Although some regions show high positive correlations with SST, others sometimes correlate negatively, so the SST based rain-height estimation approach cannot be used throughout the world's ocean



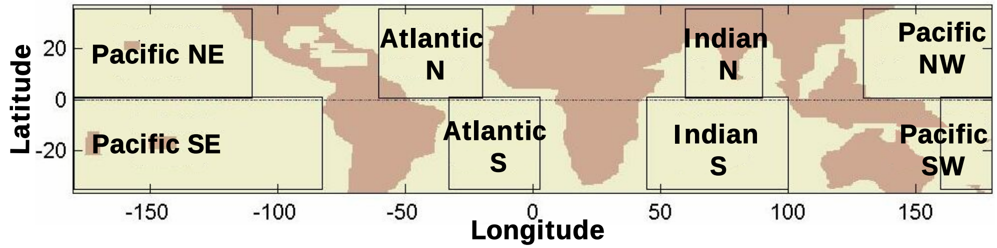
Estimation using Rain Rate

- Correlation between Rain-rate and Rain-height data
  - TRMM precipitation radar level-3 monthly averaged rain-rate and rain-height products are used in this study.
- Basic Data processing includes:
  - Same rainfall type is chosen from both the datasets (RR & RH).
  - Spurious point removal based on rain-count data.
  - Land Mask – All islands in major groups combined to single areas and the widely dispersed small islands are ignored.
  - Null-data and land masked points are omitted.



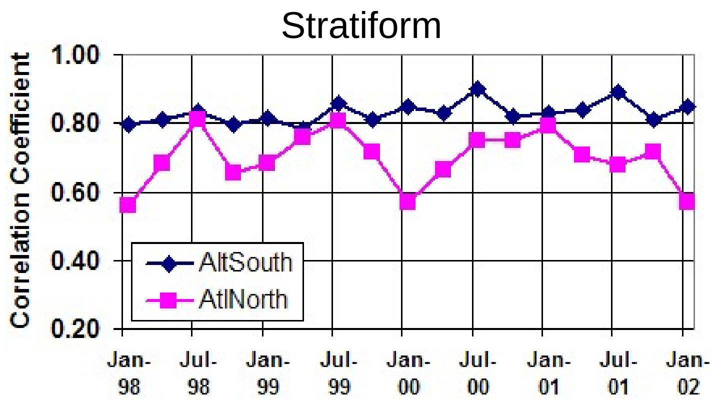


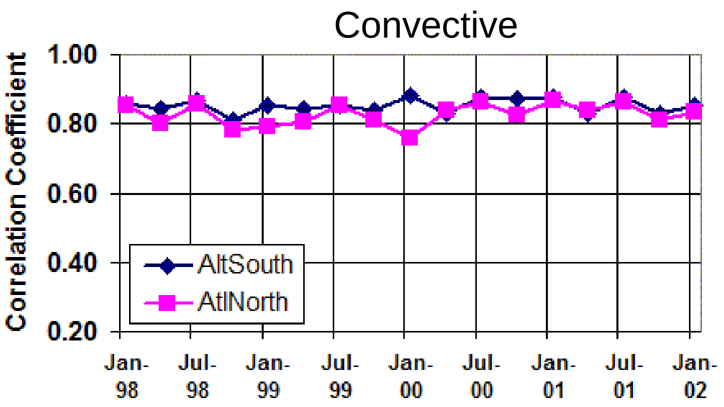
Estimation using Rain Rate



- Correlation is done on selected regions on quarterly basis
- Correlation coefficients are tabulated for further analysis

- Correlation coefficients are very high for most of the regions, seasons and rain types.
- The strong correlation between rate and height over the ocean demonstrates the feasibility of estimating rain height using rain rate.



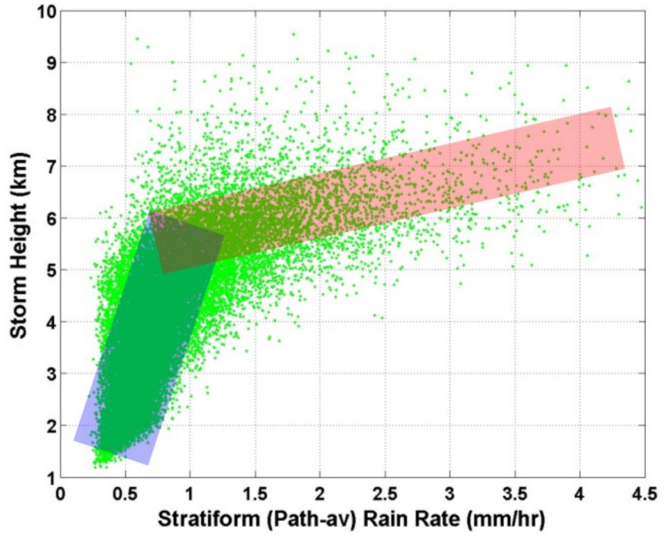


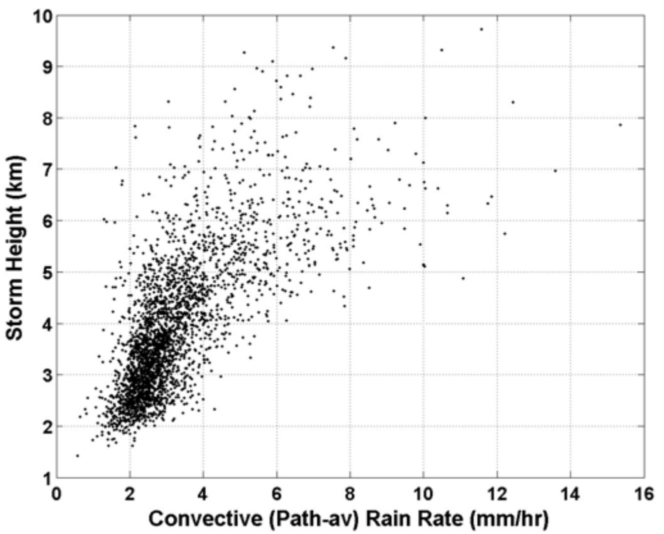
Outline

- **Introduction**
  - Thesis Objective
- **Methodologies for Rain Height (RH) Estimation**
  - Estimation using Rain Rate (RR)
- **Regression Analysis**
  - Log-Linear Combined Regression scheme
- **Regression-based Statistical Model**
- **Validation of Statistical Model**
- **Conclusion**

RR vs. RH scatter

- Mathematical dependency of rain height to rain rate is analyzed using regression analysis.
- TRMM PR's rain height and rain rate products are used in this analysis.
- Data processing of TRMM rain products are similar to the previous section.
- RR vs. RH scatter exhibits two different trends for different rain-rate regimes.
  - Lighter rain: RH vs. RR relation is steep
  - Heavier rain: scatter almost saturates
- RR vs. RH scatter exhibits similar trend for all rain types, and regional and hemispherical based data.



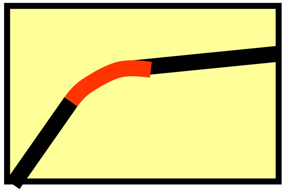


Objective of Regression methods

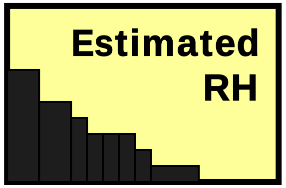
- Good fit over all RR
  - Importance in Lower RR section because of large number of scatter points.

Good Fit
∀RR

- Smooth Transition from one trend to another to maintain consistency.



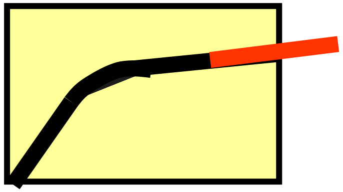
- Estimated RH distribution
  - The estimated rain-height distribution should be consistent with the original rain-height distribution.



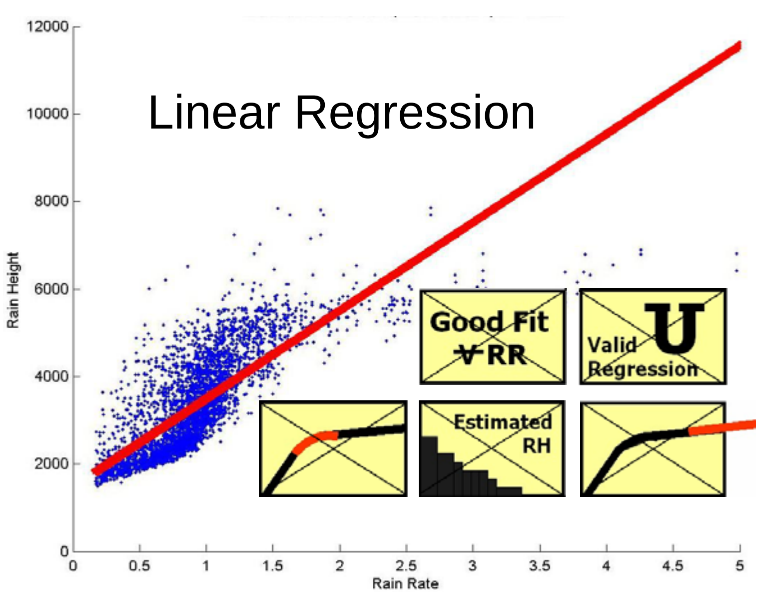
- Universal Regression scheme
  - It is important to have a regression scheme that applies to various regions, seasons, and rain types because, for simplicity, one should have a single regression model for final prediction of storm height from rain rate.

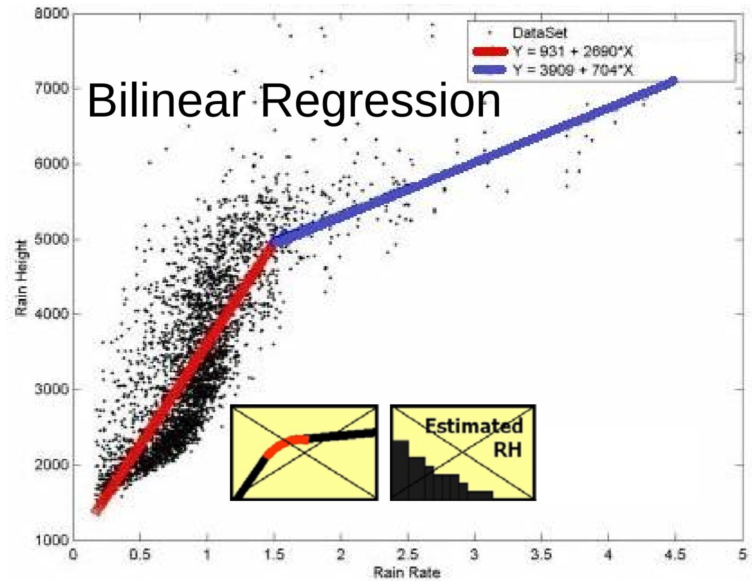
U
Valid
Regression

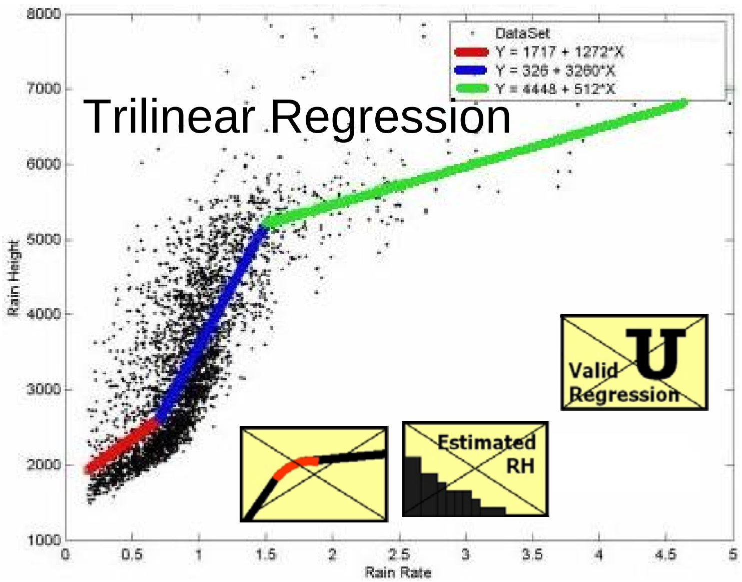
- Linear extrapolation
  - Mathematical relationship should be extendable.

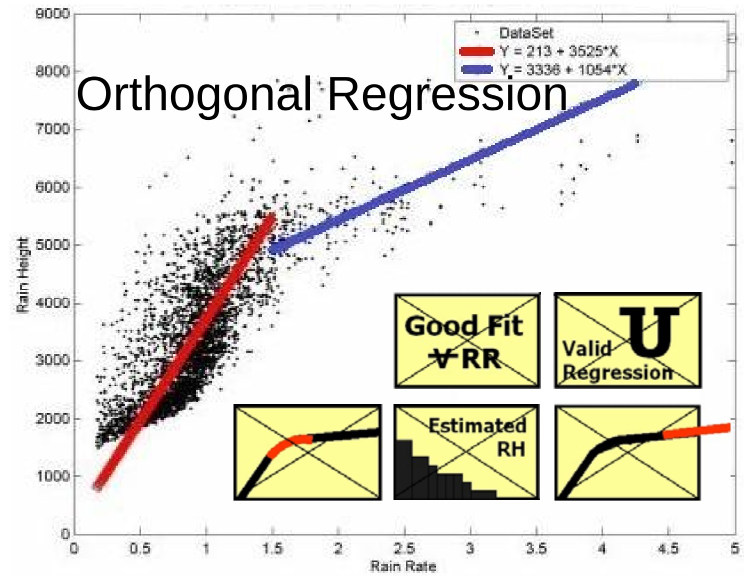


Linear Regression Schemes









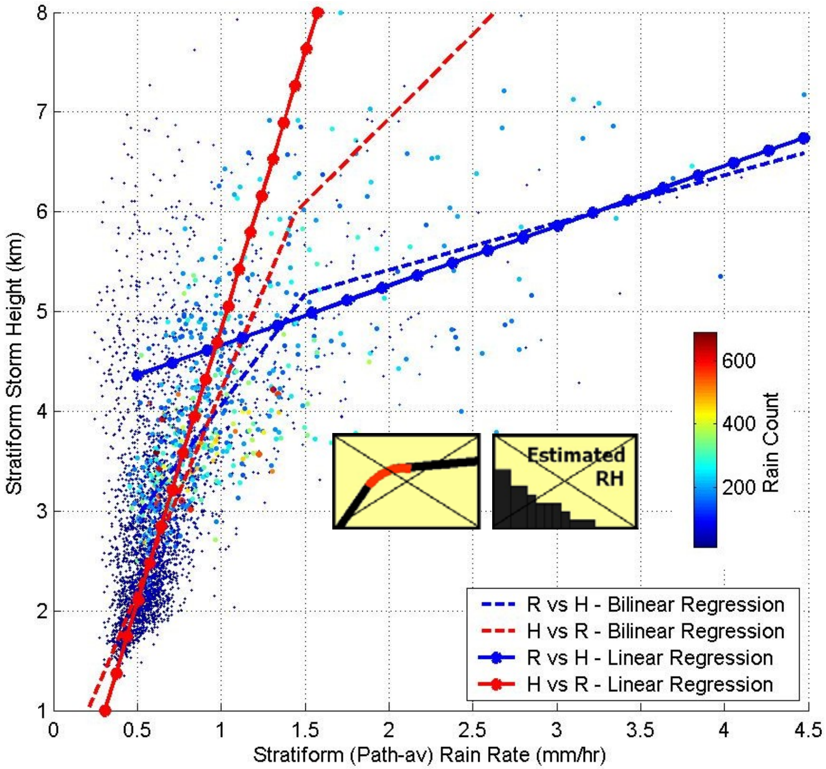
Combined Linear

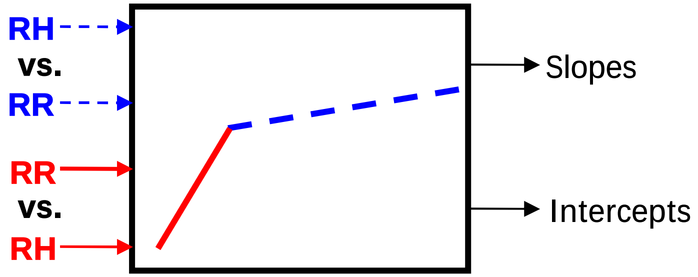
- Bilinear regression lines tend to influence each other, because of the additional constraints applied to make the lines meet at a point.

- Two linear regression lines:
  - RR vs. RH regression for lower rain rates (red solid)
  - RH vs. RR regression for higher rain rates (blue solid)
  - Individual linear regression lines seem to fit the scatter better

- *Combined Linear* regression
  - Two linear regression lines extended to all RR forms Combined Linear regression

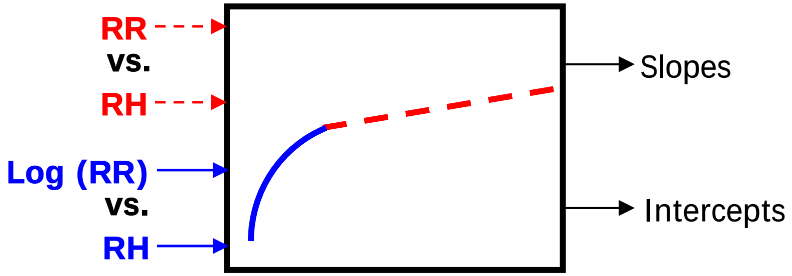
- Drawback: sharp transitions between lower and higher regression lines

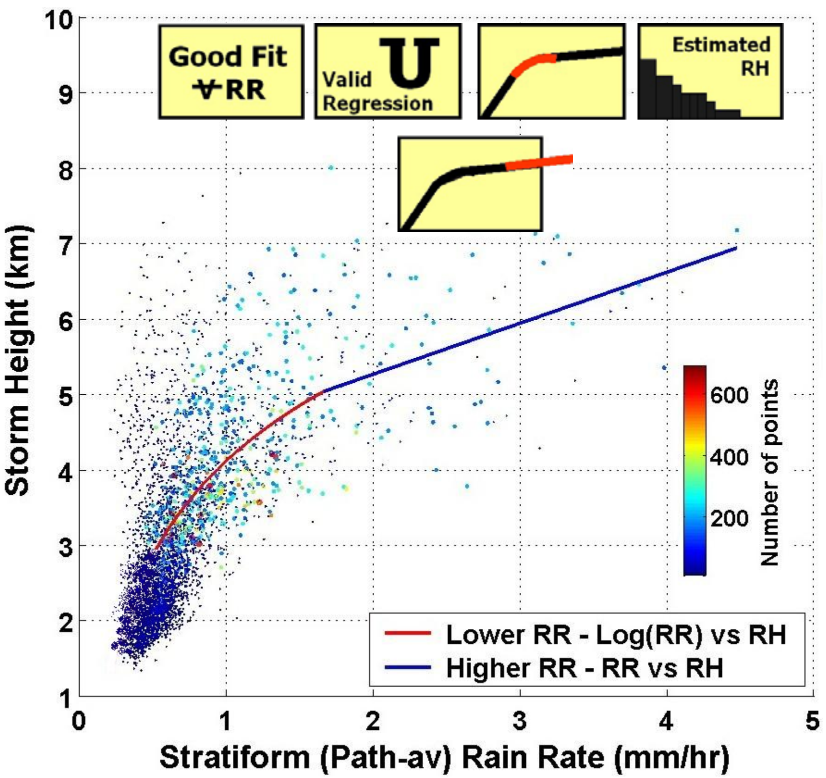




Log-Linear Combined

- *Log-linear Combined* regression is a modification to the *Combined Linear* regression to over come the drawbacks.
- In *Log-Linear Combined* regression, two linear regression lines:
  - Log(RR) vs. RH regression for lower rain rates.
  - RR vs. RH regression for higher rain rates.
- *Log-Linear combined* provides smooth transition between different trends and good fit for all regions and rain types.
- *Log-Linear Combined* is selected as the best regression scheme to relate rain rate and rain height.





Statistical Significance

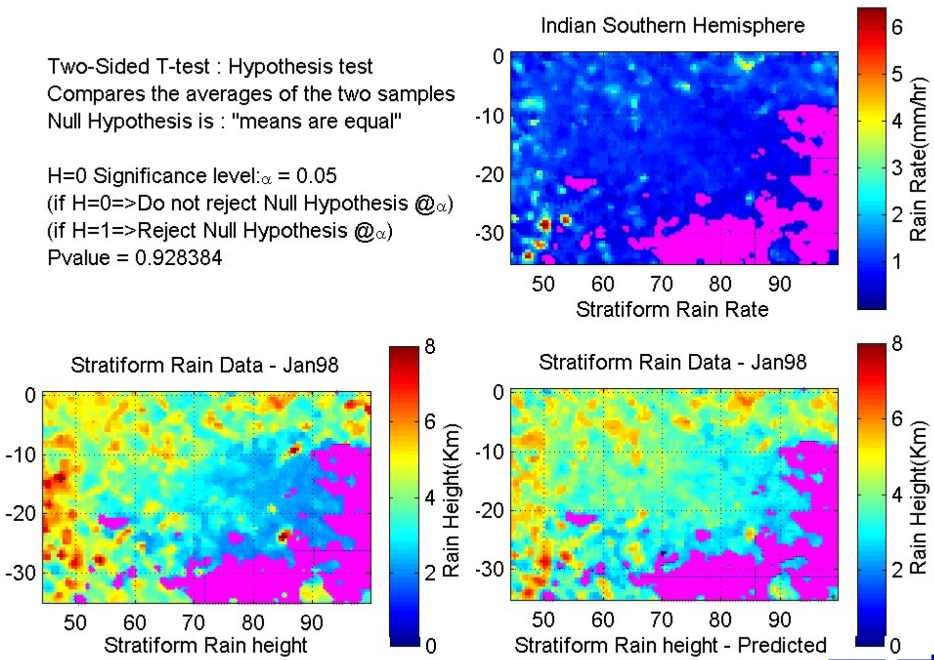
- Statistical measures
  - Standard Error of the estimate – square root of the average squared error in estimation
  - Goodness of fit is a measure of the extent to which the total variation of the dependent variable is explained by the regression scheme
- *Log-Linear Combined* regression shows best result compared to all other regression schemes
- To visualize the accuracy of our regression fit over all rain-rate values, we segmented the RR vs. RH scatter and estimated standard error in height for each segment. *Log-Linear Combined* regression showed consistently low standard error in height for all rain-rates segments

Regression schemes	Standard Error of Estimate (km)	Goodness of fit R^2
Linear Regression	2.5	0.4
Bilinear Regression	1.3	0.52
Weighted Bilinear Regression	1.2	0.55
Log(RR) vs. RH Regression	1.4	0.60
‘Combined Linear’ Regression	0.93	0.89
‘Log-Linear Combined’ Regression	**0.87**	**0.93**

Hypothesis T-test

- Hypothesis Paired-t test was used to find the closeness of the original TRMM rain height to that of the estimated rain height obtained from the regression model.
- The null hypothesis in this case is that the means of the two rain height data are equal.
- After analyzing the data based on T-statistics, the t-test produces a result to either “reject the null hypothesis” or “do not reject the null hypothesis”
- The p-value result from the t-test is the probability of the observing the given sample result under the assumption that the null hypothesis is true.

- The p-value for this test over Indian Southern Hemisphere is given as 0.9284.
- Higher p-value implies higher probability for the null hypothesis to be true.
- Similar results are observed for rain data from other stratiform regions.
- Convective rain data typically results in lower p-value, compared to the stratiform case.



Regional vs. Hemispherical based RR estimates

- Predicted rain-height from the regional-based RR vs. RH scatter is compared to hemispherical-based RR vs. RH scatter (t-test).
- The table shows higher confidence in northern and southern hemisphere scatter based rain-height prediction compared to region data based rain-height prediction.
- This result is very significant as one can omit the regional data based analysis for the rest of the statistical model development.
- Similar results were observed for convective case.

Stratiform

Region	p-value
Indian north	0.80
Atlantic north	0.91
Pacific northeast	0.78
Pacific northwest	0.75
Northern Hemisphere	**0.90**
Indian south	0.93
Atlantic south	0.92
Pacific southeast	0.89
Pacific southwest	0.85
Southern Hemisphere	**0.94**

Convective

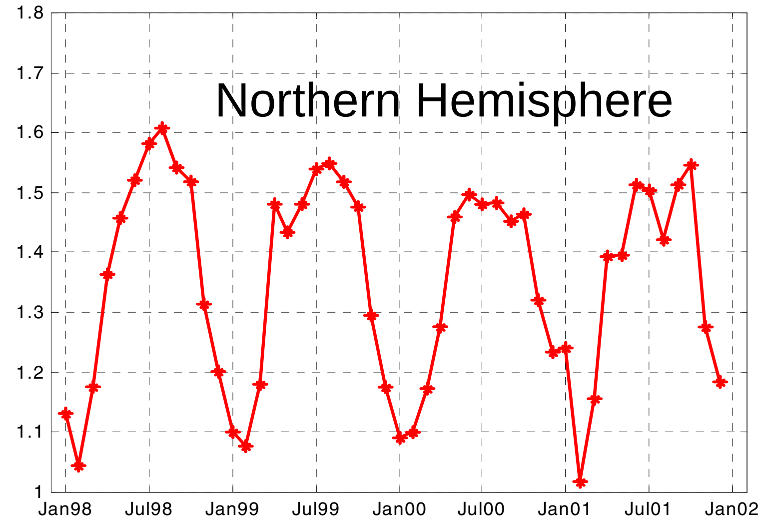
Region	p-value
Indian north	0.73
Atlantic north	0.75
Pacific northeast	0.77
Pacific northwest	0.69
Northern Hemisphere	**0.75**
Indian south	0.74
Atlantic south	0.72
Pacific southeast	0.69
Pacific southwest	0.70
Southern Hemisphere	**0.79**

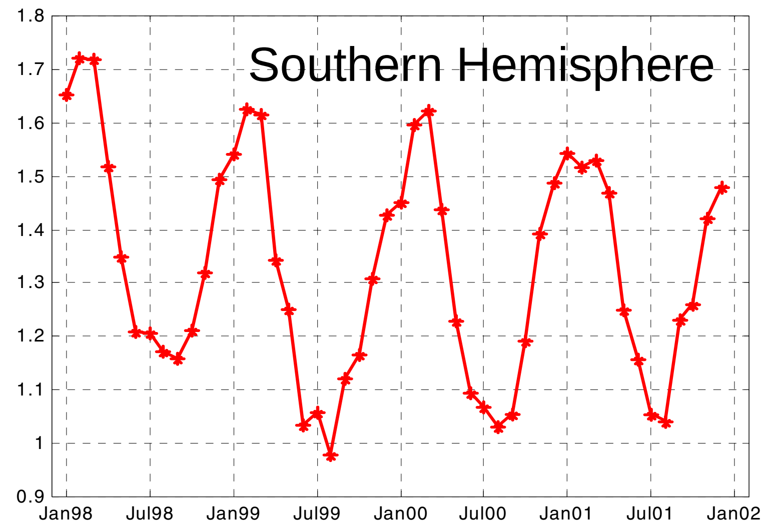
Outline

- **Introduction**
  - Thesis Objective
- **Methodologies for Rain Height (RH) Estimation**
  - Estimation using Rain Rate (RR)
- **Regression Analysis**
  - Log-Linear Combined Regression scheme
- **Regression-based Statistical Model**
- **Validation of Statistical Model**
- **Conclusion**

Slope and Intercept Statistics

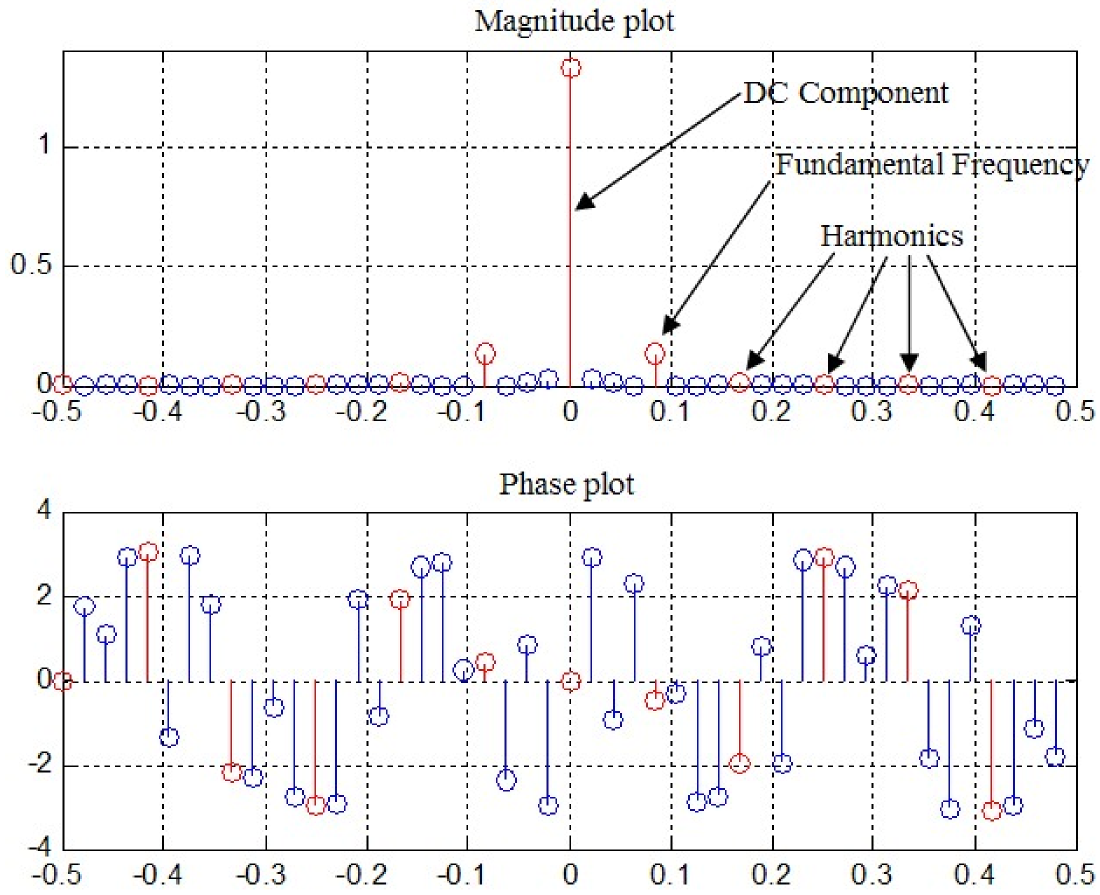
- Slopes and intercepts of *Log-Linear Combined* regression schemes are tabulated.
- Values of slopes from the regression fits clearly show sinusoidal trend over seasons with approximately one year as the period of each cycle.
- The northern and southern hemisphere slope curves are approximately out of phase.
- The seasonal trend in slopes and intercepts statistics have to be generalized to predict them for any future month.
- Mean-value substitution not possible since they affect the estimated rain-height.





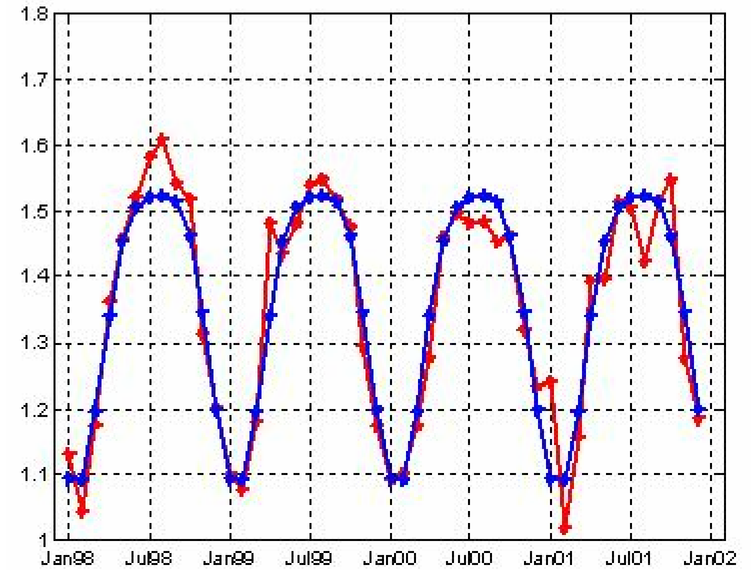
Fourier Analysis

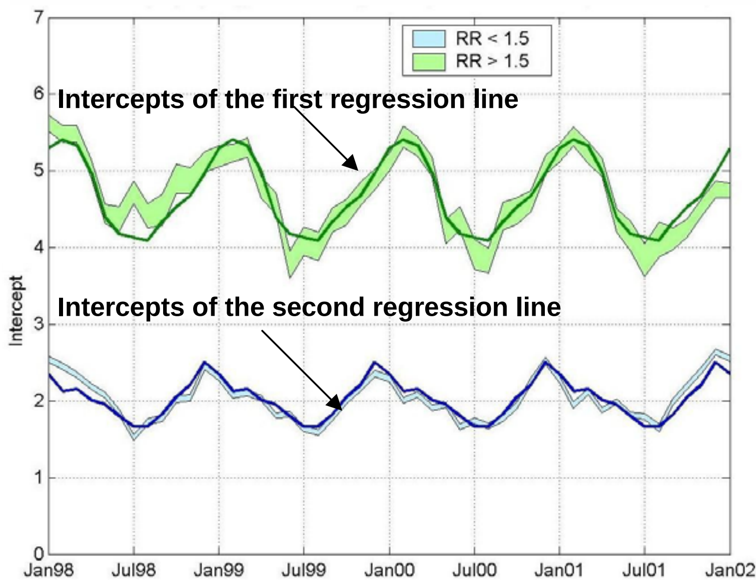
- Fourier Analysis can capture the seasonal trend exhibited by the slope and intercept statistics.
- Fourier coefficients can be obtained by analyzing the spectrum of the slope and intercept curves.



Fourier Synthesis

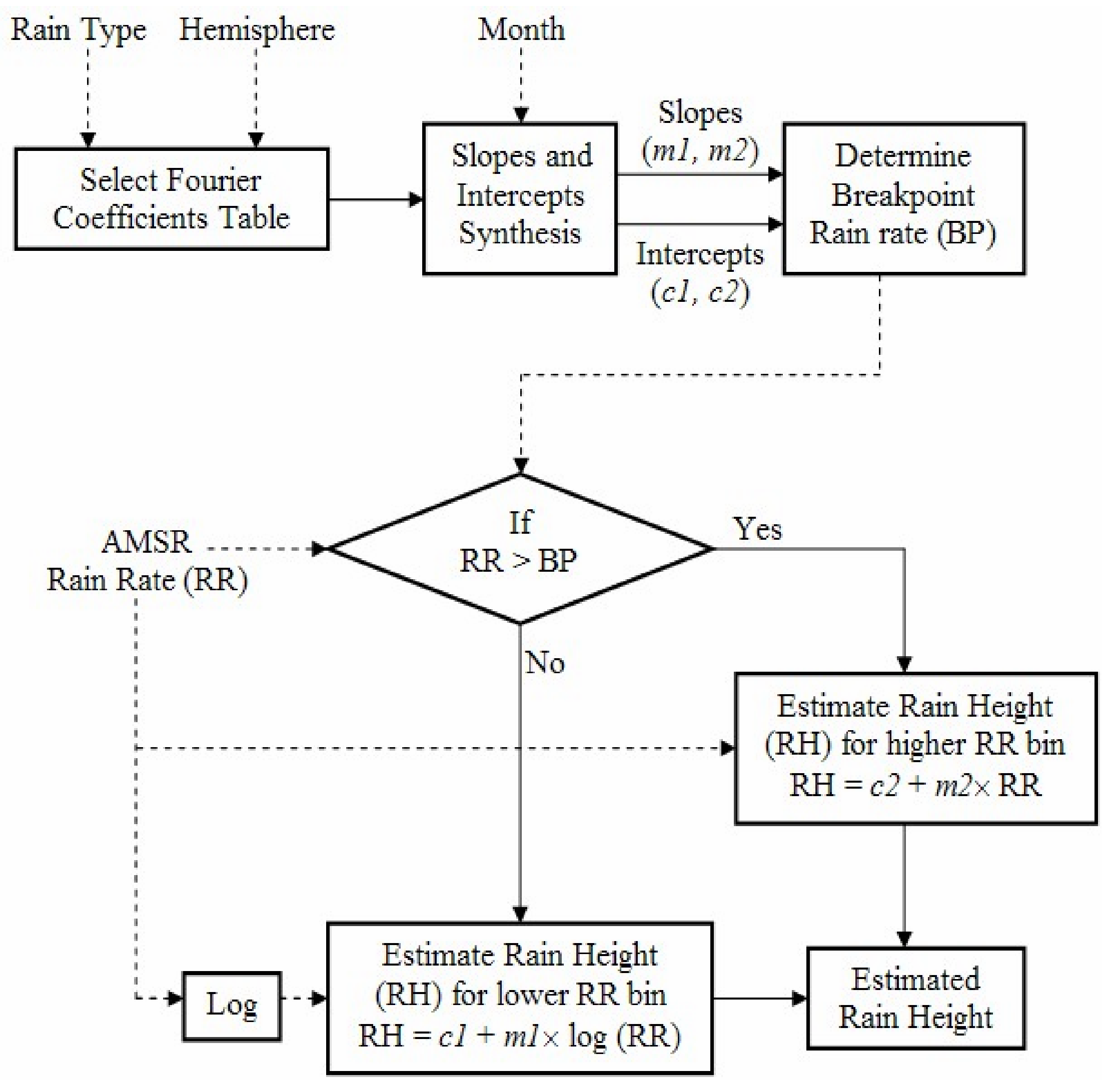
- After analyzing various slopes and intercept statistics, we concluded that the use of DC, fundamental and two harmonics are sufficient to regenerate the basic sinusoidal trend.
- Fourier synthesized curves along with the original slope statistics are shown.
- Using Fourier coefficients, the slopes and intercepts of the *Log-Linear Combined* regression scheme can be predicted for given month in the future.
- Table of Fourier coefficients are the final output of this analysis. Only four Fourier coefficient tables are required in the *Statistical Model*. They are:
  - Stratiform Northern Hemisphere
  - Stratiform Southern Hemisphere
  - Convective Northern Hemisphere
  - Convective Southern Hemisphere





Regression-based Statistical Model

- Slopes and intercepts are generated from the selected Fourier coefficient table based on rain type, hemisphere and month.
- The slopes and intercept values of the two regression lines are used to find the breakpoint rain-rate. This would split the scatter into two rain-rate sections.
- *Regression-based Statistical Model* uses the predicted slopes and intercepts of the *Log-Linear Combined* regression scheme to estimate rain height from AMSR rain rate.
- Final output of this thesis



Outline

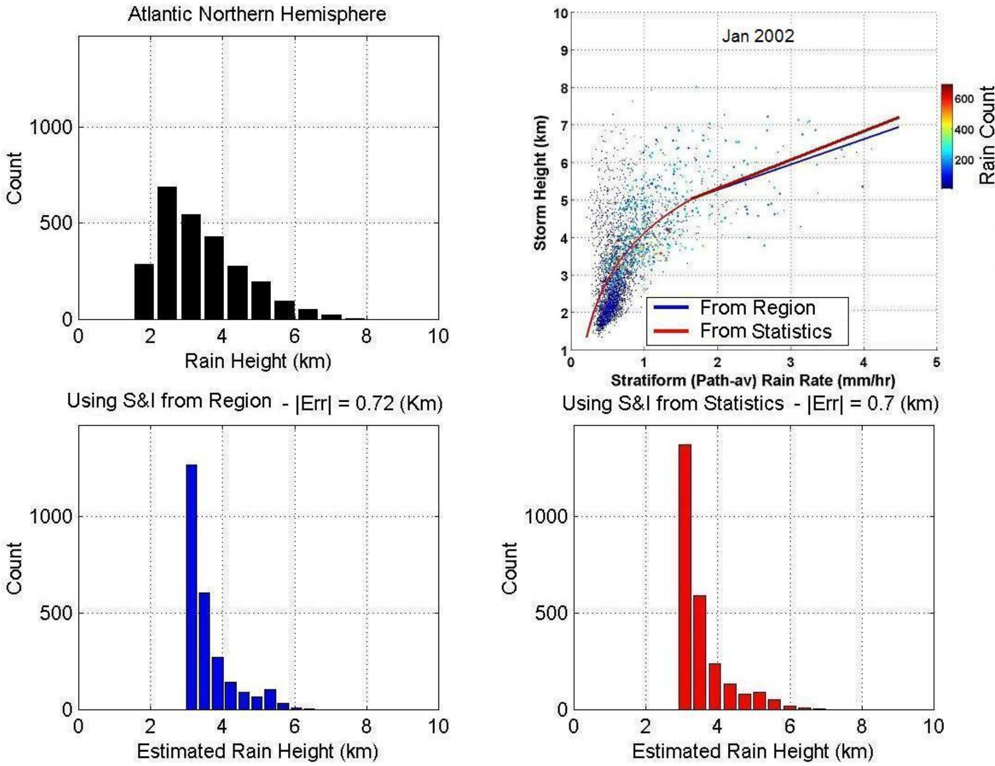
- **Introduction**
  - Thesis Objective
- **Methodologies for Rain Height (RH) Estimation**
  - Estimation using Rain Rate (RR)
- **Regression Analysis**
  - Log-Linear Combined Regression scheme
- **Regression-based Statistical Model**
- **Validation of Statistical Model**
- **Conclusion**

RR estimates from statistical model

- The rain height estimates obtained from *Log-Linear Combined* regression fits are compared with the rain height estimates derived from the statistical model.

- The Kolmogorov-Smirnov (ks-test) goodness-of-fit hypothesis test is used to compare the estimated rain-height distributions. The test results shows "not to reject the null hypothesis" and with a p-value of 0.92

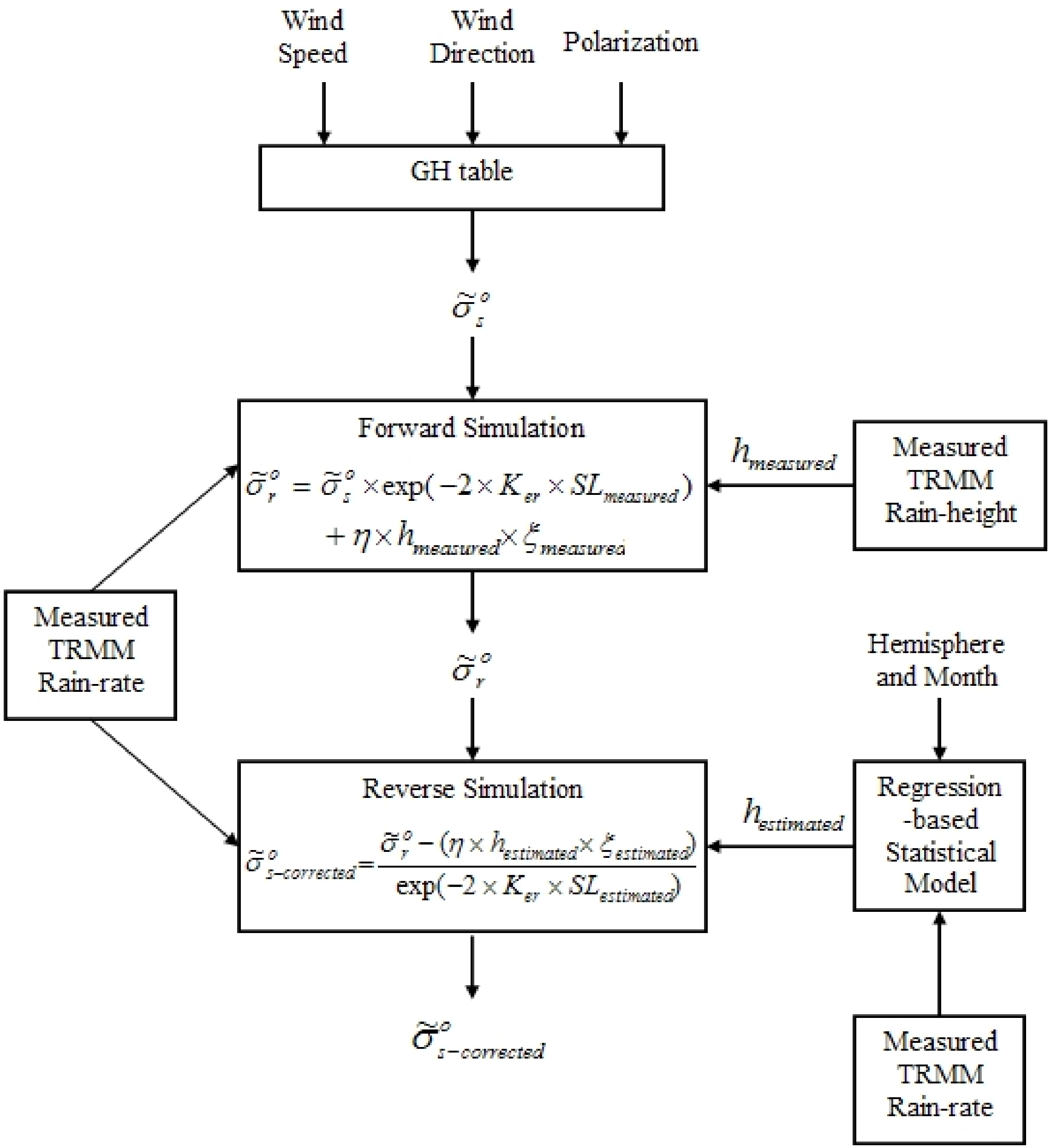
- This comparison highlights the effectiveness of the statistical model in predicting the slopes and intercepts of the regression lines and the subsequent rain-height estimation.



Simulation Procedure

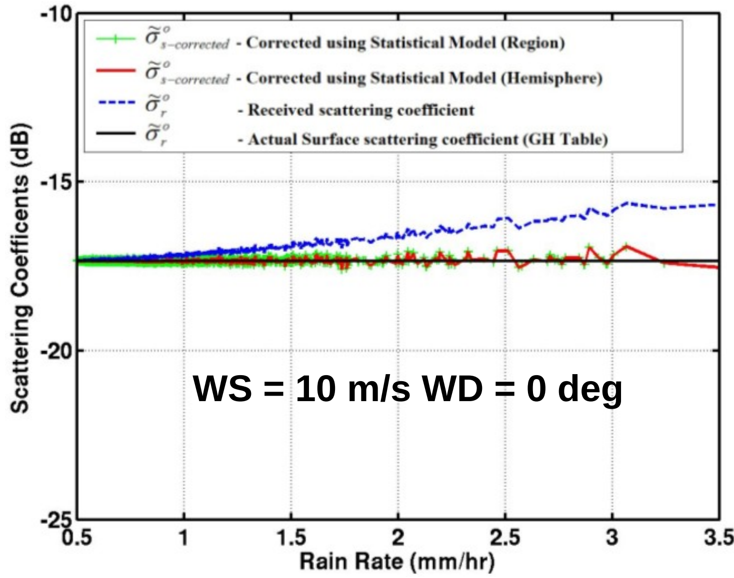
- Rain-height estimates from the *statistical model* will be in error compared to the actual rain-height values.
- Simulations were performed to determine the nature of the errors in σ_s^0 with the use of the estimated rain-height from the *regression-based statistical model.*
- Forward Simulation
  - Estimates the σ_r^0 as measured by the SeaWinds instrument.
  - The forward simulation models the rain effects in the SeaWinds received signal with the use of measured TRMM rain-height values $h_{measured.}$
- Reverse Simulation
  - The reverse simulation tries to correct rain effects in the σ_r^0 to estimate the corrected σ_s^0
  - The estimated rain-height $h_{estimated}$ from the *regression-based statistical model* is used in the reverse simulation to model the effects of rain.

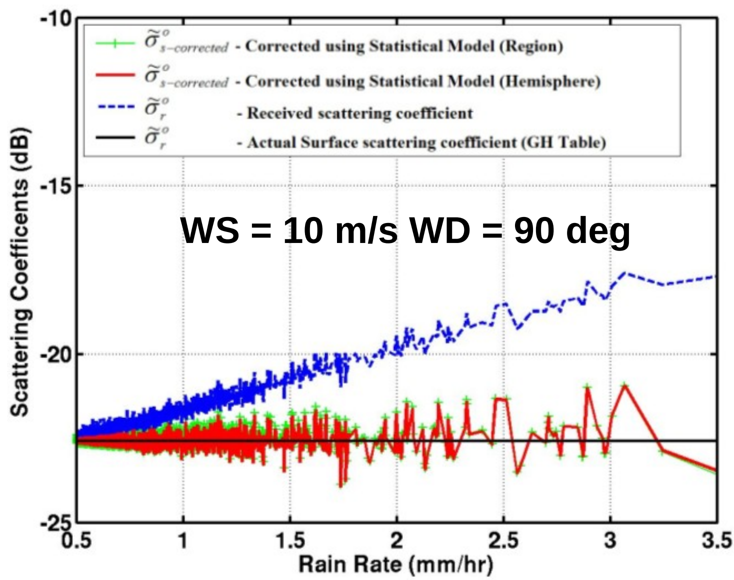
Simulation Procedure (contd.)



Simulation Results

- Simulations were performed for different combinations of wind speed, wind direction and polarization.
  - The errors in the corrected surface scattering coefficient are greater at higher rain rates – due to the large spread of data points for those rain rates.
  - Errors between the corrected and true scattering coefficients were greater at lower wind speeds than at higher wind speeds. Also we observe that the error increases as the look direction approaches crosswind.
- Corrected scattering signal obtained from statistical model developed based on regional (green) and hemispherical (red) data are very close to each other.
- The *regression-based statistical model* was found to be effective for rain-height estimation for most cases.
  - However, for some high rain-rates and other wind conditions, correction is not possible – rain flags





Conclusion

- SeaWinds scatterometer measurements are corrupted by rain attenuation and backscatter.
- SeaWinds on ADEOS-II had AMSR available to aid correction for rain effects.
- To estimate volume backscatter from rain, must have estimate of rain height - Unavailable in ADEOS-II.
- Our study of TRMM RR and RH showed good correlation.
- Due to the nature of the RR vs. RH scatter, *Log-Linear Combined* regression (log(RR) vs. RH for lower RR and RR vs. RH for higher RR) was selected.
- To apply this regression-based relationship for all seasons, the slope and intercept statistics were Fourier analyzed to capture the seasonal trends.
- Finally, we proposed the *regression-based statistical model*, which estimates the rain height from rain rate for a given hemisphere, season, and rain type.
- Simulation results showed good rain-effect correction for surface scattering coefficient for lower rain-rates, validating the use of statistical model in the *rain effect correction algorithm*.
- Although ADEOS-II failed, the methods for correction will be useful in future satellites carrying both wind-vector scatterometers and microwave radiometers

Thank you