Rain Heights Over the Oceans: Relation to Rain Rates

Bhaskar Natarajakumar, Vivek Kurisunkal, Richard K. Moore, David Braaten
Radar Systems & Remote Sensing Laboratory
Information and Telecommunication Technology Center
The University of Kansas
2335 Irving Hill Road
Lawrence, KS, 66045-7612, U.S.A.
rmoore@ku.edu

Introduction

Rain over the oceans causes attenuation of communication and radar signals for satellites and backscatter to satellite radars. To determine the attenuation and backscatter, one must know the height of the rain in a storm, and must be able to find it from a satellite measurement likely to be available. Microwave radiometers on satellites can provide estimates of rain rates, and therefore relating rain heights to these rain rates provides a way to estimate the heights.

We have been developing a method to correct the SeaWinds scatterometer data for rain effects, based on concurrent radiometer estimates of rain rates. As a part of this study, we examined the relation between rain rate and rain height, using measurements from the TRMM precipitation radar. Here we report on a phase of this study.

We find that use of a proper regression method allows estimation of rain heights with high confidence. The method selected relates low rain rates to rain height by regression of rain height (RH) *vs*. logarithm of rain rate [(log(RR)]. This is combined with linear regression of RH *vs*. RR for high rain rates to produce a combined result that best fits the observations.

Rain Height Data

The precipitation radar (PR) on the Tropical Rain Measurement Mission (TRMM) satellite provides many different outputs, including backscatter (expressed as Z values) at multiple heights within a storm. One of the outputs is the “storm height”, which is, in essence, the height of the rain, since the radar is not sensitive enough to measure cloud echoes. Another is the rain rate at each altitude and that averaged over the storm height. Although these values are available on a continuous basis, we used monthly averages on 0.5° x 0.5° grid squares for our analysis. Another useful output is the number of measurements used to produce each average value. The algorithm segregates data into stratiform, convective and “other” rains. When the signal exceeds about 40 dBZ, the rain is classified as convective. When a bright band is present, it is classified as stratiform. More complicated procedures apply when neither of these conditions is met. We used data from January, 1998, to December, 2001; and for comparison also used January, 2002, data.

Because convective rain cells seldom fill the footprint for the scatterometer, and often fail to do so even for the smaller TRMM PR footprint, the work reported here deals only with the stratiform rains. The results for convective rains are similar, but the errors are larger than those reported here.

Initially we divided the ocean into regions: North and South Atlantic; Northwest, Southwest, Northeast and Southeast Pacific; North and South Indian oceans. Comparison of results for the different regions showed that little error was introduced by combining all Northern Hemisphere and Southern Hemisphere regions, so much of what follows is based on the two hemisphere regions.

Regression Methods

A major problem was ascertaining the most suitable form for the regression relation between RH and RR, given the nature of the measurements. We tried numerous approaches before settling on a combination of RH *vs*. log(RR) for low rain rates and RR vs. RH for high rain rates. In the regression schemes tried, we used initially just the mean points from the TRMM PR. However, some of the averages used were for small numbers of measurements and others for large numbers of measurements. Therefore, we later used a weighted regression scheme that gives more weight to the average values representing many measurements than it gives to those representing only a few measurements.

Linear. The first attempt used linear regression of RH vs. RR. Fig. 1 is an example that shows that this does not work well because the RH increases rapidly with RR for low values of RR, but tends to saturate with higher values of RR.

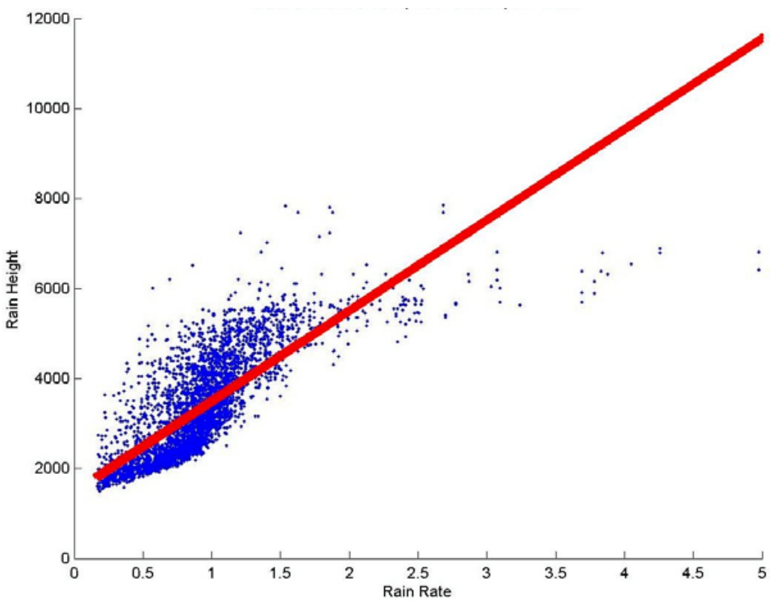


Fig. 1. Linear least square fit for stratiform North Atlantic data in January 1998

Multi-linear. The nature of the variation suggested that a bilinear regression would be appropriate, and it did improve the fit significantly, but the result was still poorer than we felt it should be. Fig. 2 shows an example. It seemed that a trilinear approach might work better, so we also tried this, as shown in Fig. 3. The trilinear did not seem to improve on the bilinear, so it was discarded

.

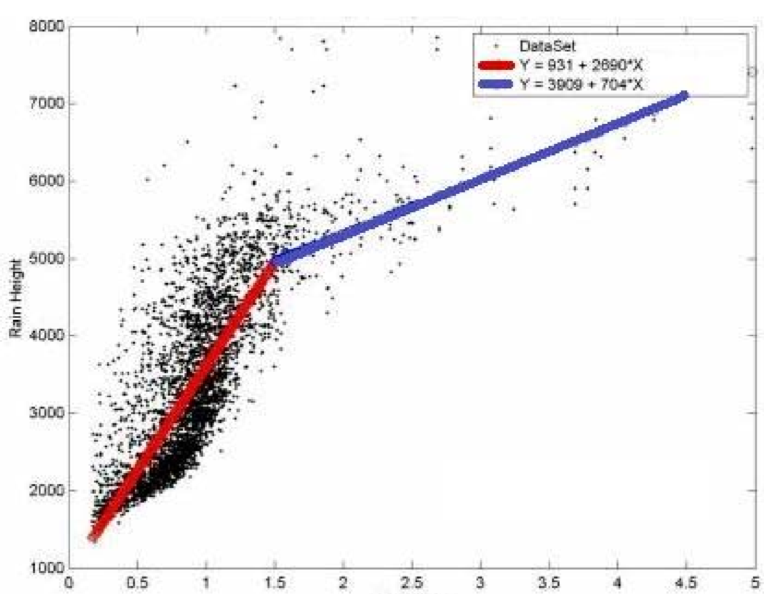
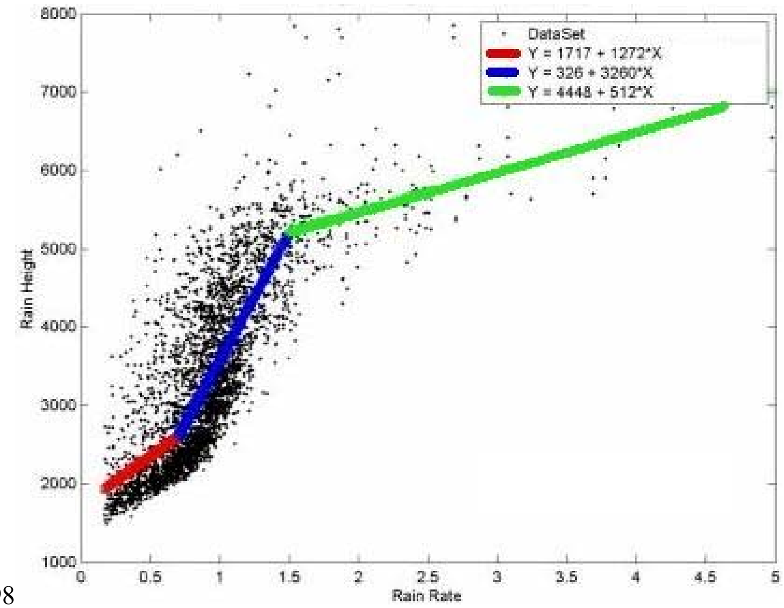


Fig. 2. Bilinear fit for RR vs. RH scatter of stratiform North Atlantic data in January



1998
Fig. 3. Trilinear fit for RR vs. RH scatter of stratiform North Atlantic data in January 1998

Initially we thought that bilinear regression would be the best solution. However, forcing the two lines to join at a fixed point causes the regression in the low-rain-rate region to influence that in the high-rain-rate region and vice versa. Consequently, before moving to log-linear regression, we showed that smaller errors occur when we used separate regressions in the two regions with the regression lines intersecting at different points depending on the results.

Log-linear. Observation of the scatter of points suggested that a logarithmic scale for RR might be more appropriate than a linear scale, so we tried this approach. The result was a fairly good fit at low rain rates, but it still could not handle the saturation at high rain rates.

Combined log-linear/linear. Since the log(RR) vs. RH regression worked well at low rain rates, it was retained for these values, but we decided that a linear regression at high rain rates would better fit the near-saturation condition there. Accordingly, we combined the two approaches for the different regions. The result was better than with any of the other approaches, so it was selected. An example is in Fig. 4.

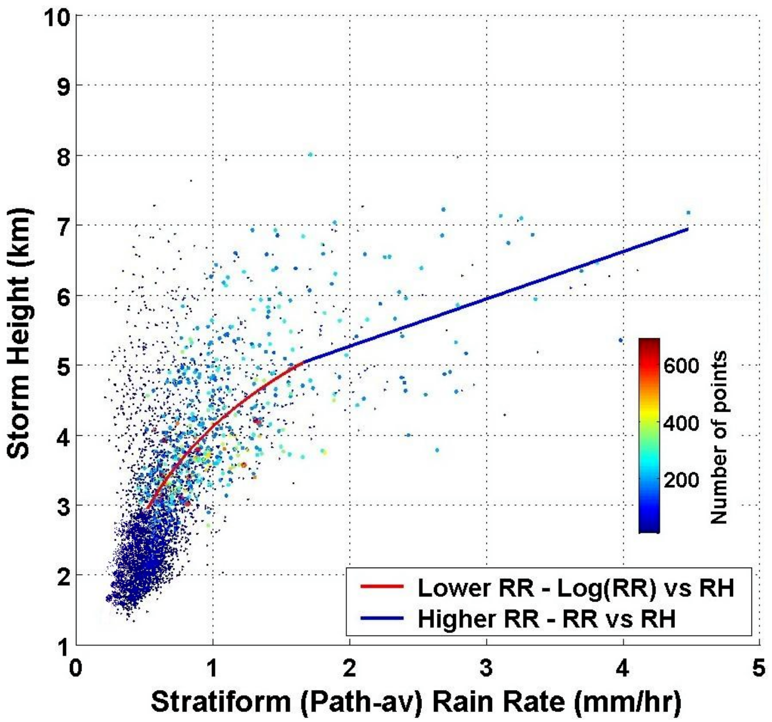


Fig. 4. Log-Linear Combined regression fit for RR vs. RH scatter of North Atlantic in January 1998.

We compared the standard errors of estimate and a goodness of fit parameter for the different regression schemes as shown in Table 1 for an example month (January 1998) and region (North Atlantic). Clearly the best method is the combined log-linear/linear.

Table 1. Standard Error of estimate and goodness of fit parameters of various regression schemes analyzed for RR vs. RH scatter for North Atlantic region in January 1998

Regression schemes	**Standard Error of Estimate (km)**	**Goodness of fit R^2**
Linear Regression	2.5	0.40
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**

Seasonal trends

We tabulated and plotted the slopes and intercepts for the regression lines as a function of time for the 4-year period studied and found that the rain-height *vs.* rain-rate relations depend on the seasons. As one would expect, these seasonal trends are opposite in the northern and southern hemispheres. The variations are quasi-sinusoidal, with 180° phase difference between the hemispheres. Since they were periodic, we could Fourier analyze them; Fig. 5 shows the

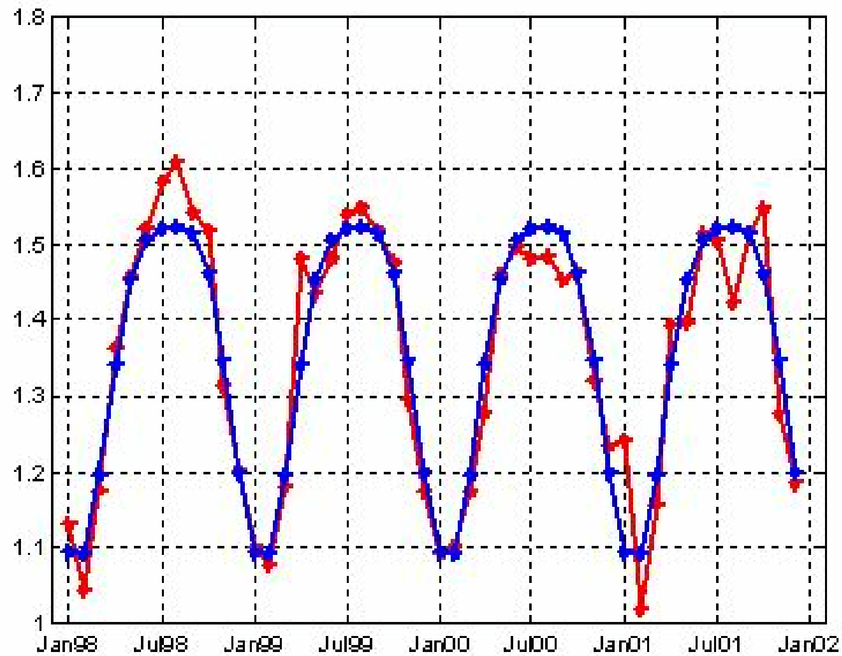


Fig. 5. Seasonal trends of slopes of regression lines for northern-hemisphere convective rain data. Values of slope based on Log(RR) vs. RH regression. The red line shows the measured data, and the blue line is the curve from Fourier analysis using the fundamental and two harmonic components.

measured slopes and the results of Fourier analysis, using only the fundamental and first two harmonics, superimposed for convective rain. Clearly the fit is good.

These Fourier series were used to create a "statistical model" for the rain-height variations when expressed in terms of rain rate. With this model we can apply future measurements of rain rate to determine the rain heights. This was tested for one case, as discussed below.

Confidence in Model

If the statistical model is to be used for future data, we must have confidence in its ability to predict the rain height from rain-rate data. Consequently we tested the model both against the data from which it was generated and against data from January, 2002 (the month after the end of the time series used to produce the model).

Student's t test with original data. The two-sided Student's t test was performed on the actual and predicted rain-height values for stratiform rain data over the southern Indian Ocean, using a significance level of $\alpha = 0.05$. Fig. 6 shows the stratiform rain-rate and rain-height maps for this case. The RR data from this region were used to predict RH values using the model, and the resulting predicted stratiform RH map is also shown. The two-sided *t* test, which compares the actual and predicted rain-height, does not have sufficient evidence to reject the null hypothesis that the means of actual and predicted rain-height data are equal. The p-value for this t-test over the southern Indian Ocean is 0.93, implying that the probability is high that the means are the same. Similar results were obtained in other regions for stratiform rains, although the results for convective storms were poorer.

Kolmogorov-Smirnov and Student's t test for prediction. A test of the predictive ability of the model compared model results with measurements from January, 2002. Fig. 7 shows the distribution of the rain-height estimates obtained using two different methods. The black histogram is the distribution of the measured TRMM rain-height values in the North-Atlantic region for January 2002. The regression lines on the scatter plot in Fig. 7 indicate the close correspondence of the statistical model and the regression fits. The blue histogram represents the distribution of the rain-height estimates obtained using the regression lines for TRMM measurements over the entire northern hemisphere.

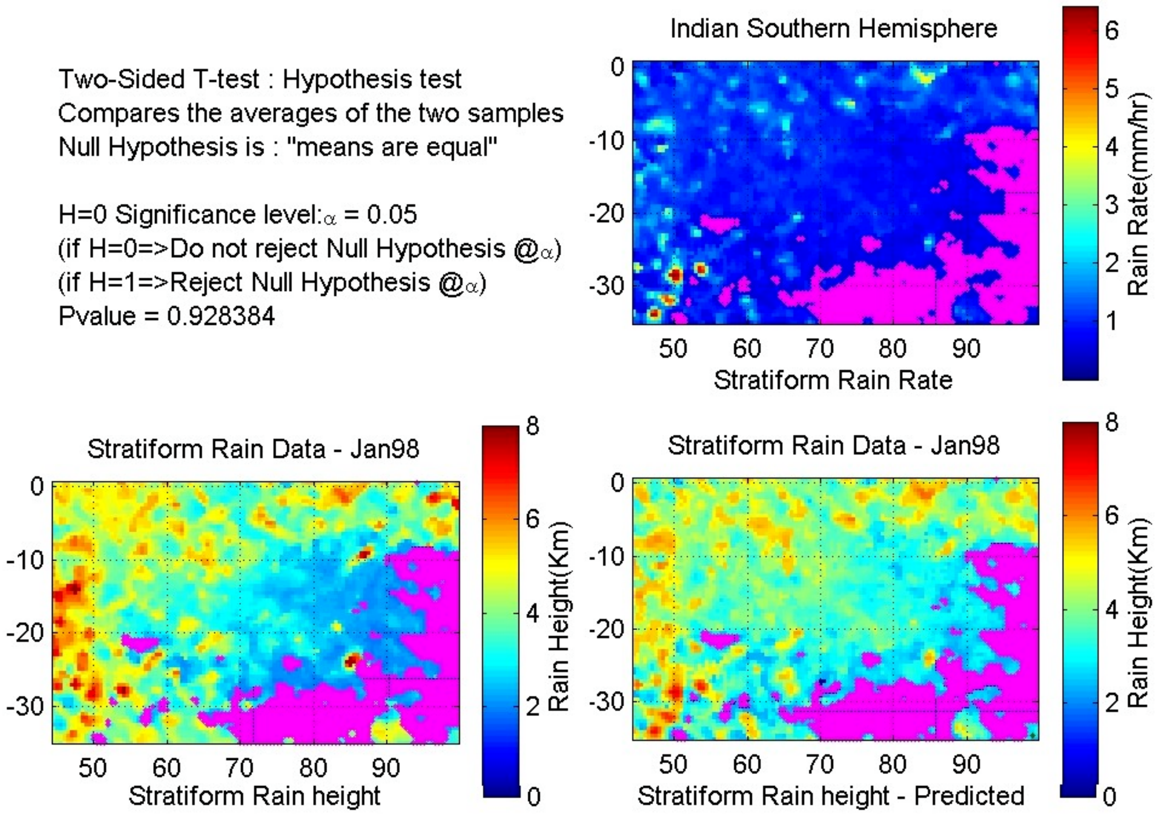


Fig. 6. Two-sided Student's t test result based on stratiform data from the southern Indian Ocean. The null points (no data) are shown in pink.

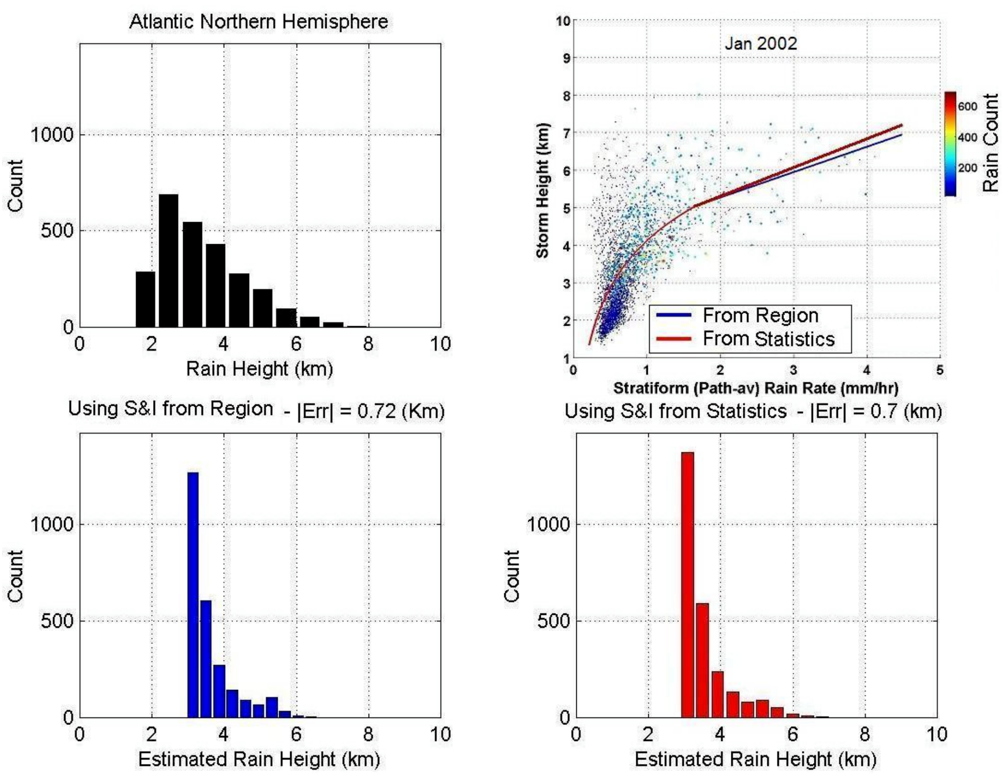


Fig. 7. Comparison of distribution of rain-height estimates obtained from the *Log-Linear* Combined regression fits and from the statistical model.

The red histogram represents their distribution obtained using the statistical model. This comparison highlights the effectiveness of the statistical model.

To quantify the confidence in the model, both the Student's *t* test and the Kolmogorov-Smirnov tests were applied to the data and the model prediction. The first test estimates whether the two processes (in this case the model and the data) have the same mean; it is best when the populations are normally distributed, which is not the case here. The second (K-S) test estimates whether two distributions are similar enough that the processes are the same. In this case, both tests suggest that the model reproduces the data well.

The measured rain height was compared with the estimated rain-height values obtained using both methods. The Student's *t* test results, with significance level (α) of 0.05, on measured and estimated rain-height suggests not to reject the null hypotheses with very high p-values as shown in Table 2. Similar results were obtained in the hypothesis testing using rain data from other regions and rain types. Thus, this test suggests that the model does, indeed, represent the data well.

Table 2. Student's t test results comparing the measured rain height and the estimated rain height

Student's t Test Comparison	**p-value**
Measured rain-height and estimated rain-height obtained from regression fits in the regional RR vs. RH scatter.	0.952
Measured rain-height and the rain-height estimates obtained from the model.	0.945

With a distribution that deviates significantly from Gaussian, the K-S test can be more reliable that the Student's *t* test. Accordingly, we applied this test to the data. In the stratiform North-Atlantic case (January 2002), the K-S test suggests that the null-hypothesis not be rejected and hence that the distributions are the same. It provides high confidence in the result with a p-value of 0.92. That is, it suggests that the distributions are from the same population.

Conclusions

This paper shows that rain rate can be used to estimate rain height over the oceans. Since microwave radiometers on satellites can provide estimates of rain rates, this means that satellites carrying radiometers can be used to study rain-height distributions, and satellites carrying both radiometers and radars can use the radiometer data to correct the radar data for rain effects.

Monthly average measurements on a 0.5° x 0.5° grid from the TRMM Precipitation Radar were used in the study. The plots of RH *vs.* RR showed considerable scatter, and several different regression schemes were tried to find the one that best described the observations. The approach finally adopted used regression of log(RR) *vs.* RH for small rain rates and of RH *vs.* RR for larger rain rates. Tests showed that this method led to estimated rain-height errors that were considerably under 1 km in all cases, at least for stratiform rain.

Although initial tests were done on oceanic regions, we found that we could use the entire northern hemisphere as a single region and the southern hemisphere as another, without significant loss of accuracy. Seasonal trends in the two regions were found to be 180º out of phase, as one would expect. Fourier analysis of the seasonal trends of the regression parameters showed that a series with a fundamental and two harmonics described the variation well, and this was used to produce a model that can be applied to future data.

Hypothesis testing using both the Student's t and Kolmogorov-Smirnov tests showed that the model described the processes well.

This work was intended to develop a correction scheme for the SeaWinds scatterometer on the ADEOS-II spacecraft. However, failure of the spacecraft makes this application moot. Nevertheless, the method can be applied on future radar-carrying spacecraft that also have radiometers.