The NOAA Climate Prediction Center
African Rainfall Estimation Algorithm
Version 2.0

Beginning January 1, 2001, the African Rainfall Estimation Algorithm Version 2 (RFE 2.0) was moved to operational status, replacing the previous algorithm (RFE 1.0) (Herman et.al. 1997), used from June 1, 1995 to December 31, 2000. The merging technique, the backbone of RFE 2.0, has been shown to significantly reduce bias and random error compared to individual precipitation data sources, thus increasing the accuracy of the rainfall estimates (Xie and Arkin, 1996). Due to modernization of data sources and programming techniques, RFE 2.0 exhibits many improvements over its predecessor. Along with improved accuracy, increased speed and convenient portability make RFE 2.0 a much better method to estimate daily precipitation over Africa, although orographic rainfall effects are not incorporated. The core programs execute in approximately 8 minutes on a standard Linux PC. RFE 1.0 required a workstation and almost 6 hours to complete. Following are some essential facts about RFE 2.0.

Input data used for operational rainfall estimates are from 4 sources; 1) Daily GTS rain gauge data for up to 1000 stations 2) AMSU microwave satellite precipitation estimates up to 4 times per day 3) SSM/I satellite rainfall estimates up to 4 times per day 4) GPI cloud-top IR temperature precipitation estimates on a half-hour basis. The three satellite estimates are first combined linearly using predetermined weighting coefficients, then are merged with station data to determine the final African rainfall. Daily binary and graphical output files are produced at approximately 3pm EST with a resolution of 0.1° and spatial extent from 40°S-40°N and 20°W-55°E. Additional data sets of 10-day, monthly, and seasonal rainfall totals are created by accumulating daily data. Seven other daily binary output fields are produced using various combinations of input data, but these are not considered operational and will be discussed later. Resources needed for the algorithm to function include the four data source files, the Linux operating system, 64MB RAM, minimum 3GB hard disk space, and a Fortran 77 or 90 compiler. By default, each output field is created in GrADS format, so the GrADS graphics package will allow for easy image creation. Inputs to RFE 2.0 will be discussed in somewhat more detail next.

RFE 2.0 uses four types of input data, including three satellite sources, to create the final rainfall estimates. Special Sensor Microwave/Imager (SSM/I) rain rate estimation data is available up to four times per day depending on the geographical location, as governed by ascending and descending Defense Meteorological Satellite Program polar satellite tracks. The NOAA SSM/I rainfall algorithm uses the 85V GHz channel to detect the scattering of upwelling radiation by precipitation sized ice particles within the cloud layer (Ferraro and Marks, 1995); (Ferraro et.al., 1996). Returned scattering patterns are compared against previously derived rainfall amounts, and instantaneous rain rates are determined with a horizontal resolution of 0.25°. High emissivity of land surfaces necessitates the use of scattering techniques versus emissions, and although they are generally less accurate, this high frequency approach increases the spatial resolution. Since water surfaces have a relatively low emissivity, emission algorithms using lower frequency waves are used over the ocean as long as insignificant scattering is detected. The instrument measures the brightness temperature increases due to the presence of liquid precipitation. As of June 2001, AMSU-B data will be input to RFE 2.0, replacing older AMSU-A precipitation estimates. Like SSM/I, AMSU-A rain rates were based on a scattering algorithm over land and an emission algorithm over ocean (Zhao et. al., 2000). Over land, rain rates were determined from measuring brightness temperature differences due to ice concentration, while over the ocean, the measurement of cloud liquid water determined the rain rate. The rain rate retrieval process used for AMSU-B is very similar to that used with the SSM/I instrument, and produces double the spatial and temporal resolution of its AMSU-A predecessor. Half-hourly GOES Precipitation Index (GPI) rainfall amounts are derived from Meteosat IR cloud top temperatures and input into RFE 2.0. Empirical methods have determined that cloud top temperatures less than 235K in the tropics are generally expected to produce stratiform rainfall at the rate of 1.5mm/half hour (Arkin and Meisner, 1987). Thus, all IR segments are combined by explicit time integration and total daily GPI rainfall is input into RFE 2.0 with a resolution of 4km. Finally, Global Telecommunications Station (GTS) precipitation data is input to the algorithm in a final step in obtaining the daily RFE values. Up to approximately 1000 stations are available for the African continent on any given day, although the number used is usually less than 500 due to poor station maintenance or erroneous data. The need for satellite-estimated precipitation arises from this non-dependable, poorly spatially distributed rainfall data.

The data input and conversion process is the first step in the RFE 2.0 algorithm. SSM/I and GTS data require the least manipulation to prepare for input, as the data set is available in a binary format that can be directly reduced to an Africa grid and assimilated. HDF formatted global AMSU precipitation data is ftp'd to the working site and then converted to binary format via Fortran scripts. This process is quick and immensely reduces the file size so that data may be made available to outside users at their request. Half-hourly Meteosat IR data must be cut to an Africa grid and areas of cloud-top temperature below 235K determined before data is input to RFE 2.0. This process of determining 'temperature counts' is the most time consuming of the entire algorithm design, taking more than an hour to complete, although most time is spent copying files to the working machine. All input data sets are finally used to produce distinct individual data with resolutions of 5.0°, 2.5°, 1.0°, 0.5°, 0.25°, and 0.1°. It should be mentioned that all inputs in their original resolution are archived to CD for later use. After each data set is prepared for incorporation to RFE 2.0, the merging process begins.

The two-step merging process is the essence of RFE 2.0, as it is here that all inputs are combined and rainfall estimates are produced. Merging is needed due to the fact that separately, each input source is incomplete in spatial coverage and contains non-negligible random error and systematic bias. As will be shown later, the method used to combine data improves these aspects significantly. The first step involved in the merging process is to reduce random error of the satellite precipitation estimates. This is done by linearly combining GPI, SSM/I, and AMSU data through a maximum likelihood estimation method. Using the equation:

$$W_i = \frac{\sigma_i^{-2}}{\sum_{i-1}^{3} \sigma_i^{-2}}$$ where W_i = weighting coefficient, σ^2 = random error

weighting coefficients for each satellite data type are calculated from their random errors which are determined by comparing the estimated precipitation to actual rain gauge values on a daily basis. It can be seen that each weighting coefficient is inversely proportional to the random error of the corresponding satellite method, thus giving increasingly accurate estimations greater leverage. After weighting coefficients are determined, the rainfall estimates are combined to produce a precipitation estimate with reduced random error:

$$S = \sum_{i-1}^{3} W_i S_i$$ where S_i = individual satellite rainfall estimate

Precipitation estimates are calculated for grid resolutions of 5.0°, 2.5°, 1.0°, 0.5°, 0.25°, and 0.1° from input data sets created earlier.

The second step of the merging process compares the satellite-estimated precipitation in step one with GTS rain gauge data to remove as much bias as possible. For a complete explanation of this process, refer to Xie and Arkin, 1996. In short, the shape of the precipitation is given by the combined satellite estimates, while the magnitude is inferred from GTS station data. Immediately surrounding a GTS station, the final precipitation estimates retain the station's rain gauge value, while as distance from a station increases, rainfall estimates rely more heavily on satellite precipitation.

Cross validation of the RFE 2.0 rainfall estimates was performed for the period from December 1-30, 1999, prior to the algorithm becoming operational. The merging process was performed ten times, removing 10% of the GTS stations for each run, until all stations had been removed exactly once. Each processing of the data used the remaining 90% of the GTS stations as inputs, and precipitation estimates were created. The rainfall estimates were then compared each time to the station recorded precipitation values that were removed. Statistics were generated upon completion of all processes and the results are shown in table 1.

Data	Bias (mm/day)	Correlation
GPI only	2.26	0.345
SSM/I only	-0.24	0.321
AMSU-A only	-0.15	0.095
GTS+GPI+SSM/I+AMSU-A	-0.15	0.501
GTS+GPI	-0.04	0.467

Table 1: Output Bias and Correlation of Various Input Data Combinations

GTS, GPI, SSM/I, and AMSU-A data are considered the operational inputs of the RFE 2.0 algorithm. It can be seen that this set of inputs produces the highest correlation between estimated precipitation and actual station rainfall amounts, with a relatively small bias. Using GTS and GPI input data produced similar results, although a slightly lower correlation was obtained. Due to this adequate accuracy and the availability of data, a climatology is currently being created from 1982-current that uses these two inputs. For the operational product, the introduction of AMSU-B microwave precipitation estimates will likely improve the output accuracy further, but tests have not been completed to determine the overall improvement.

As mentioned earlier, there are a total of eight outputs produced by RFE 2.0, although only one is considered operational. Other outputs include combined GTS rainfall data, a GTS sampling distribution, as well as merged and unmerged precipitation estimates using various data sets. One output uses the same input data as the operational product, but adds six-hourly Global Data Assimilation System (GDAS) precipitation estimates to produce the final rainfall estimate. Only GTS and GPI rainfall data are used for the final set of outputs, and table 1 shows reasonable accuracy for these results. Due to excellent availability of input data along with low bias and relatively high correlation of estimated results, the merged output of these input data sets are used for an African rainfall climatology that is currently being created.

References

Arkin, P. A. and B. N. Meisner, 1987: The Relationship between Large-Scale Convective Rainfall and Cold Cloud over the Western Hemisphere during 1982-84. *Mon. Wea. Rev.* Vol. 115, No. 1, pp. 51–74.

Ferraro, R. R., and G. F. Marks, 1995: The Development of SSM/I Rain Rate Retrieval Algorithms Using Ground-Based Radar Measurements. *J. Atmos. Oceanic Technol.*, **12**, 775-780.

Ferraro, R. R., N. C. Grody, F. Weng, and A. Basist, 1996: An Eight-Year (1987–1994) Time Series of Rainfall, Clouds, Water Vapor, Snow Cover, and Sea Ice Derived from SSM/I Measurements. *Bulletin of the Amer. Metr. Soc.*: Vol. 77, No. 5, pp. 891–906.

Herman, A., V. B. Kumar, P.A. Arkin, and J.V. Kousky, 1997: Objectively Determined 10-Day African Rainfall Estimates Created for Famine Early Warning Systems. *Int. J. Remote Sensing*, **18**, 2147-2159.

Xie, P., and P. A. Arkin, 1996: Analysis of Global Monthly Precipitation Using Gauge Observations, Satellite Estimates, and Numerical Model Prediction, *J. Climate*, **9**, 840-858.

Zhao, L., R. Ferraro, and D. Moore, 2000: Validation of NOAA-15 AMSU-A Rain Rate Algorithms. *10th Conf. on Satellite Metr.* 192-195.