Techniques Used to Create CMAP and Their Potential Improvements

Pingping Xie[1)], Yelena Yarosh[1)], Mingyue Chen[1)],
John E. Janowiak[1),] and Phillip A. Arkin[2)]

[1)] NOAA/NWS/NCEP/Climate Prediction Center
5200 Auth Road
Camp Springs, MD 20746, USA
[2)] The Earth System Interdisciplinary Center (ESSIC)
University of Maryland,
College Park, MD 20742, USA

1. Introduction

The objective of this article is to describe the objective techniques used to create the CPC Merged Analysis of Precipitation (CMAP, Xie and Arkin 1996)and to discuss the problems in the current version of the technique. The authors expect that comments and advice from the workshop attendees will help the future improvements of the CMAP.

2. The CMAP Algorithm

The CPC Merged Analysis of Precipitation (CMAP, Xie and Arkin 1996, 1997a, 1997b) is defined by merging seven kinds of individual input data sources. These input data sets include the gauge data (the GPCC gauge-based analyses of Rudolf et al. 1994 over land and the atoll gauge data of Morrissey et al. 1995 over ocean),5 sets of satellite estimates derived from 1) the IR-based GPI (Arkin and Meisner 1987), 2) OLR-based Precipitation Index (OPI, Xie and Arkin 1998), 3) SSM/I scattering-based estimates of Ferraro and mark (1995), 4) SSM/I emission-based estimates of Wilheit et al. (1993), and 5)MSU-based estimates of Spencer (1993). Precipitation fields generated by the NCEP/NCAR reanalysis (Kalney et al. 1996) are also utilized as an additional source.

The merging of the individual input data sources is conducted in two steps. First, to reduce the random error, satellite estimates and reanalysis precipitation fields are combined linearly through the Maximum Likelihood Estimation method, in which the linear combination coefficients are inversely proportional to the squares of local random error of the individual data sources. Over global land areas, the individual random error is defined for each grid and for each time step (month/pentad) by comparing the data sources with the concurrent gauge-based analysis over the surrounding areas. Over global oceanic areas, it is defined by comparison with atoll gauge data (Morrissey et al. 1995) over tropics and by subjective assumptions regarding the error structures over the extra-tropics (Xie and Arkin 1997a).

Since the output of the first step contains bias passed through from the individual input data sources, a second step is included to remove it. For that purpose, the gauge-based analyses are combined with the output of the first step. Over land areas, the gauge data and the output of the first step are blended through the method of Reynolds (1988), in which the first-step-output and the gauge data are used to define the relative distribution (or “shape”) and the magnitude of the precipitation fields, respectively. Over the oceans, the bias in the first-step-output is ?removed? by comparison with the atoll gauge data over the tropics and by subjective assumptions regarding the bias structures over the extra-tropics.

The techniques described above have been applied to construct analyses of monthly and pentad precipitation on a 2.5o lat/lon grid over the globe for a 24-year period from 1979 to the present. Called the monthly and pentad CMAP (Xie and Arkin, 1997a and b), these analyses have been used widely for climate analysis, model verifications and other applications.

The CMAP techniques were designed and developed almost 10 years ago to ensure that analyses of precipitation are created with complete spatial coverage and reasonable accuracy by merging individual inputs with spatial resolution, availability and quality of1993/1994:

- Multiple satellite estimates are combined linearly to achieve maximum spatial/temporal coverage with reduced random error;

- Gauge observations were used over both land and ocean to assure stability in the magnitude of the final merged analyses throughout the data period;

- Precipitation fields generated by the NCEP/NCAR reanalysis were included to ensure complete global coverage and to complement the reduced quality of satellite estimates over mid- and hi-latitudes;

3. Problems and Potential Improvements

In recent years, several examinations and comparisons have been conducted for the CMAP datasets. While the data sets have proved to be useful for many applications in climate analysis and model verifications, several problems have been reported, most of which are due to the shortcomings of the objective techniques used to create the data sets.

3.1 Analyses over land

Over land areas, the precipitation distribution over regions subject to orographic effects is poorly represented in the current CMAP data sets, due to the sparse gauge networks there and lack of appropriate techniques to combine information from gauge observations, satellite observations and other sources.

Work is underway in NOAA/CPC to improve the quality of the merged analyses over the global land areas. As a first step, the optimal interpolation technique (OI, Gandin 1963) has been applied to create analyses of global monthly precipitation (Chen et al. 2002). In the OI-based precipitation analyses, the climatology of monthly precipitation is first defined for over 17,000 stations using the gauge observations of Global Historical Climatology Network (GHCN, NOAA/NCDC) for a 40-year period from1951-1990. Gridded analyses of monthly climatology are then defined for the 12 calendar months by interpolating the station climatology and used as the first guess fields in the OI. Monthly precipitation at a grid point is finally defined by adding increments determined from nearby gauge observations to the first guess.

Several tests have been conducted to examine the quality and robustness of the OI technique. Figure 1shows the sensitivity test performed for the OI and the interpolation algorithm of Shepard (1968) over the African continent for a 21-year period from 1950 to 1970 during which station observations are available from relatively dense gauge networks over the target domain. The Shepard (1968) technique is used to define the gauge-based analyses of GPCC, which determine the magnitude of the CMAP analyses over the land areas.

4 sets of monthly precipitation analyses were constructed by interpolating all and a subset of the GHCN gauge data using the OI and the Shepard (1968)techniques. Presented in figure 1 are comparison results for the 21-year mean precipitation distribution among the4 sets of analyses.
Fig.1 1950-1970 mean precipitation differences between analyses defined by interpolating all and a subset of the available gauge data using the OI and the Shepard techniques.

It is clear from the above figure that small differences are observed between the analyses created by the OI and the Shepard (1968) when gauge observations from all stations are used (left-top panel), indicating that the analysis is NOT sensitive to the interpolation algorithm used when the gauge network is dense. The 21-year mean precipitation in the OI-based analyses using fewer gauges is almost the same as that in the OI analyses using all gauges (left-bottom panel). Substantial differences, however, are observed when the analyses are generated from fewer gauges using the interpolation technique of Shepard (1968) (right panels). These preliminary results show that by making use of the climatology defined from dense gauge network, the OI is capable of representing the overall magnitude of precipitation fields better than a simple interpolation algorithm like Shepard (1968) in which only station observations available for the target month are used.

Further work is underway to develop an OI-based technique to combine the gauge observations with various satellite estimates.

3.2 Analyses over ocean

Over oceanic areas, uncertainty exists in the magnitude of the CMAP precipitation fields. The current CMAP techniques rely too much on the atoll gauge data. In particular, the CMAP procedures that use the atoll gauge data to adjust the oceanic bias may add noise to the combined satellite estimates and alias the global oceanic mean precipitation by assuming that the relative bias is identical over the entire tropical oceans and decreases gradually toward high latitudes. A comprehensive examination may be needed to check the bias structure for satellite estimates. Comparison with water budget calculation from reliable models may give us hints on the overall magnitude of precipitation, especially over mid- and high latitudes. New techniques need to be developed to ensure stable magnitude for the oceanic precipitation analyses.

3.3 Analyses over the high latitudes

Over high latitudes (polar caps), the CMAP analyses are basically defined as the same as the precipitation fields generated by the NCEP/NCAR reanalysis. While the circulation fields are reasonably well produced in the current version of the reanalysis, the precipitation fields are not represented as well. While collection of precipitation observations from more gauge stations will help improve the quality of the merged analysis, the final solution to the problem may lie in the development of completely new objective analysis techniques that combines the circulation fields from models with observations from gauges and satellites. Comments and advices from experts on this topics are highly appreciated.

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., 115, 51 - 74.

Chen, M., P. Xie, J.E. Janowiak, P.A. Arkin, 2002: Global Land Precipitation: A 50-yr Monthly Analysis Based on Gauge Observations. Journal of Hydrometeorology: 3, pp. 249??266.

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, 755 - 770.

Gandin, L. S., 1963: Objective analysis of meteorological fields. Gidrometeor. Isdaty., Leningrad. [Israel Program for Scientific Translations, Jerusalem,1965, 242 pp.] Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, Roy Jenne, and Dennis Joseph,1996: The NMC/NCAR 40-Year Reanalysis Project. Bull. Amer. Meteor. Soc., 77, 437-471.

Morrissey, M.L., M.A. Shafer, S.E. Postawko, and B. Gibson, 1995: Pacific rain gauge data. Water Resources Research, 31, 2111-2113.

Reynolds, R.W., 1988: A real-time global sea surface temperature analysis. J. Clim., 1, 75 - 86.

Rudolf, B., H. Hauschild, W. Rueth, and U. Sdchneider,1994: Terrestrial precipitation analysis: Operational method and required density of point measurements. NATO ASI Ser., 126, 173-186.

Shepard, D., 1968: A two-dimensional interpolation function for irregularly spaced data. 23rdNatl. Conf. Of Amer. Computing Machinery, Princeton, N.J.

Spencer, R.W., 1993: Global oceanic precipitation from MSU during 1979 - 91 and comparisons to other climatologies. J. Clim., 6, 1301 - 1326.

Wilheit, T., A.T.C. Chang, and L.S. Chiu, 1991: Retrieval of monthly rainfall indices from microwave radiometric measurements using probability distribution functions. J. Atmos. Oceanic Technol., 8, 118-136.

Xie, P., and P.A. Arkin, 1996: Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model precipitations. J. Clim., 9,840 - 858.

Xie, P., and P.A. Arkin, 1997a: Global Precipitation: A17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs. Bull. Amer. Meteor. Soc.,78,2539 - 2558.

Xie, P., and P.A. Arkin, 1997b: Global pentad precipitation analysis based on gauge observations, satellite estimates and model outputs. Amer. Geophys. Union 1997 Fall Meeting, December 8-12, 1997, San Francisco, CA, USA.

Xie, P., and P.A. Arkin, 1998: Global monthly precipitation estimates from satellite-observed outgoing longwave radiation. J. Clim., 11, 137- 164.